



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

DEPARTAMENTO DE TOXICOLOGÍA

“Contribución de la carga ancestral regional a las variantes de la
arsénico-3 metiltransferasa (AS3MT) en poblaciones Nativas y
mestizas mexicanas contemporáneas”

Tesis que presenta:

LBD. REBECA ITZEL MONTERO DELGADO

Para obtener el grado de Maestría en Ciencias

En la especialidad de Toxicología

Director(a) de tesis:

Dra. María del Rocío Gómez Ortega

Este trabajo se realizó en el Laboratorio 14 del Departamento de Toxicología del Centro de Investigación y de Estudios Avanzados del IPN bajo la tutoría de la Dra. María del Rocío Gómez Ortega y fue financiado por el Consejo Nacional de Ciencia y Tecnología a través del Fondo Sectorial de Investigación en Salud y Seguridad Social (CONAYT-FOSISS)-No. 261268 y por el fondo de la *National Geographic Society*, *grant*: GR-000042849.

Resumen

La elevada exposición al arsénico inorgánico a través del agua de bebida es un problema de salud pública mundial. En algunas regiones de México, la población está expuesta hasta 100 veces al límite permitido por la Organización Mundial de la Salud (10 ppb). El arsénico inorgánico se metaboliza principalmente por la arsénico-3-metil transferasa (As3MT), enzima que adiciona grupos metilo de forma secuencial, para formar los metabolitos ácidos monometilarsénico y dimetilarsénico que se eliminan en la orina. Las proporciones altas de dimetilarsénico urinario se relacionan con una mayor capacidad de metilación y con un menor riesgo de toxicidad mediada por arsénico inorgánico. Ambas características se relacionan con la presencia de alelos considerados como de eficiente actividad enzimática en el gen de la *AS3MT*. En el presente estudio de tipo transversal, se relacionó la diversidad genética de cinco variantes polimórficas en el gen *AS3MT* y la carga ancestral a través de marcadores autosómicos bialélicos en 309 individuos muestreados a conveniencia: 225 mestizos y 84 Nativos Americanos pertenecientes a las etnias Mazahua, Me'Phaa y Tsotsil. Para ello se empleó el método de discriminación alélica mediante la reacción en cadena de la polimerasa cuantitativa en muestras de sangre y saliva (cada individuo firmó el consentimiento informado aprobados por comité de ética de Biología Molecular Diagnóstica S.A. de C.V. para mestizos y National Geogrpahic para los Nativos). Nuestros hallazgos mostraron que la población Me'Phaa, con prominente ancestría Nativo Americana (96.7%), presentó una mayor frecuencia de variantes relacionadas con la eficiencia (0.43). Por su parte, el resto de las poblaciones Nativas, así como los mestizos, no fueron consistentes con este comportamiento, posiblemente por el componente europeo. En general, nuestros resultados contribuyen a describir la variabilidad genética de la población Mexicana en la *AS3MT*, enfatizando la importancia del estudio de marcadores genéticos en nuestra población los cuales pueden tener utilidad predictiva como marcadores de susceptibilidad; en este caso en particular, con la metilación del iAs.

Abstract

High exposure to inorganic Arsenic through drinking water is a global public health problem. Several regions in Mexico, the population is exposed up to 100 times more than the limit recognised by the World Health Organisation (10 ppb). The inorganic arsenic is metabolised by the Arsenic 3+ methyltransferase (As3MT) enzyme, mainly, which adds methyl groups sequentially, generating two acid metabolites monomethylarsonic (MMA) and dimethylarsinic (DMA); being all eliminated in urine. DMA high proportions in urine have been related to a higher methylation capacity and, in turn, with lower arsenic toxicity risk. These two features have been associated with the presence of AS3MT specific alleles designated as efficient. In the present cross-sectional study, the gene diversity of five SNPs located on AS3MT and the autosomal ancestry background was related in 309 individuals (225 mestizos and 84 Native Americans belonged to Mazahua, Me'Phaa, and Tsotsil Native American groups). The quantitative polymerase chain reaction was used for allelic discriminations in blood and saliva samples (Each individual signed an informed consent form validated by the Ethics Committees of the Laboratorio Biologia Molecular Diagnostica for mestizos and National Geographic for Natives). Our findings pointed out, that the Me'Phaa population with remarkable Native American ancestry (96.7%), showed the highest frequencies of AS3MT variants related to efficiency (0.43). Nonetheless, this finding was not consistent in the rest of the populations studied, possibly due to European ancestry background. Overall, our results provided to broad the landscape about the arsenic-3-methyltransferase genetic diversity in Mexican individuals, emphasising the relevance of these studies in our population, which can have predictive value as susceptibility markers, in this particular case, with inorganic arsenic methylation.

Agradecimientos

A mi familia por tener su cariño y apoyo incondicional en todos los sentidos.

A la Dra María del Rocío Gómez Ortega por brindarme su guía, paciencia, tiempo, conocimiento y compromiso durante todo el proceso.

A los Dres, Carla Santana Torres y Gino Noris García por su brindarme su apoyo y confianza siempre.

A la Q.F.B. María de Lourdes López González por su apoyo y gran amistad, dentro y fuera del laboratorio.

Al los Dres, Libia Vega Loyo y Gabriel Obregón Molina por su apoyo y comentarios que enriquecieron el presente trabajo.

A los profesores del Departamento de Toxicología por su gran labor de enseñanza y apoyo en el programa de Maestría.

A mis amigos Ivan, Andrea y Ángel por el apoyo y todos los momentos compartidos.

Al CINVESTAV por brindar todos los medios para la realización de este posgrado.

A amigos que conocí durante el trayecto y permanecerán.

ÍNDICE

1. INTRODUCCIÓN	11
1.1 Arsénico inorgánico: generalidades	11
1.2 Situación del Arsénico en México	12
1.3 Toxicocinética del As	17
1.3.1 Vías de exposición	17
1.3.2 Absorción, distribución, metabolización y excreción del arsénico	18
1.4 Efectos del arsénico en la salud humana	20
1.5 Diversidad genética de las poblaciones humanas	21
1.5.1 Estructura poblacional	22
1.5.2 Estructura poblacional de México	22
1.6 Posible presión selectiva por la exposición a As y sus implicaciones en el metabolismo.	24
1.7 Justificación	29
1.8 Hipótesis	29
1.9 Objetivos	30
1.9.1 Objetivo General	30
1.9.2 Objetivos particulares	30
2. MÉTODOS	31
2.1 Diseño de estudio	31
2.1.1 Selección de sujetos	31
2.1.2 Criterios de inclusión	32
2.1.3 Criterios de exclusión	32
2.1.4 Criterios de eliminación	32
2.2 Descripción de las variables	32
2.3 Descripción operacional de las variables	33
2.4 Estrategia experimental	33
2.4.1 Diseño y evaluación del panel para ancestría	33
2.4.2 Obtención y procesamiento de muestras	35
2.4.3 Análisis moleculares	37
2.4.3.1 Reacción en cadena de la polimerasa cuantitativa.	37

2.4.3.2 Genotipificación de las variantes de la As3MT.	37
2.4.3.3 Determinación de AIMs en las poblaciones de estudio.	38
2.4.4 Análisis estadístico	38
2.4.4.1 Diversidad genética en AS3MT.	38
2.4.4.2 Inferencia de carga ancestral de las poblaciones de estudio.	38
2.4.4.3 Exploración de la relación entre la carga ancestral y las variantes del AS3MT.	39
3. RESULTADOS	40
4. DISCUSIÓN	53
5. CONCLUSIONES	64
6. PERSPECTIVAS	65
7. LITERATURA CITADA	66
8. ANEXO	78

ÍNDICE DE TABLAS

Tabla 1.	Estados de la República Mexicana con determinaciones de arsénico mayores a 25 ppb en aguas subterráneas.	14
Tabla 2.	Variantes genéticas del gen de la AS3MT (<i>Chr</i> : 10q24.32) que han sido relacionadas a una mayor capacidad de metilación.	25
Tabla 3.	Poblaciones de estudio pertenecientes a las etnias mestiza y Nativa Americana contemporáneas.	31
Tabla 4.	Distancias genéticas (<i>F_{ST}</i>) y su significancia estadística (<i>p</i>) de las poblaciones de estudio y del proyecto 1000 Genomas.	42
Tabla 5.	Frecuencias alélicas de las cinco variantes del AS3MT en las poblaciones de estudio y en las consideradas como parentales.	46
Tabla 6.	Frecuencias genotípicas de las cinco variantes del AS3MT en las poblaciones de estudio.	47
Tabla 7.	Frecuencias genotípicas de las variantes de AS3MT en Me'Phaas, mestizos y otras poblaciones de México.	48
Tabla 8.	Determinación del Equilibrio con respecto a la ecuación de Hardy-Weinberg en los <i>loci</i> de AS3MT en las poblaciones de estudio.	49
Tabla 9.	Frecuencias haplotípicas obtenidas mediante los cinco <i>loci</i> de AS3MT en las poblaciones de estudio y en las poblaciones del Este de Asia y Europa del proyecto 1000 Genomas.	50
Tabla A1.	Fuentes de revisión para el diseño del panel de AIMs.	81
Tabla A2.	Poblaciones empleadas para la evaluación <i>in silico</i> del panel de ancestría.	82
Tabla A3.	Panel de Marcadores Informativos de Ancestría para la inferencia de carga ancestral regional en las poblaciones de estudio	84
Tabla A4.	Variantes del gen de la AS3MT determinadas en las poblaciones de estudio.	85

ÍNDICE DE FIGURAS

Figura 1.	Mapa de la distribución mundial del arsénico en agua subterráneas.	13
Figura 2.	Esquema del metabolismo del arsénico por la arsénico-3-metiltransferasa.	19
Figura 3.	Mapa de las principales rutas migratorias en la colonización de América y la distribución del iAs en el mismo continente.	28
Figura 4.	Análisis Multidimensional a Escala mediante 28 Marcadores Informativos de Ancestría en las poblaciones de estudio y en los datos obtenidos del proyecto 1000 Genomas.	43
Figura 5.	Proporciones ancestrales de las poblaciones de estudio y de las obtenidas del proyecto 1000 Genomas con el programa STRUCTURE v.2.3.4.	45
Figura 6.	Análisis Multidimensional a Escala de las variantes eficientes de AS3MT en las poblaciones de estudio y en los datos obtenidos del proyecto 1000 Genomas	51
Figura 7.	Mapa de frecuencias haplotípicas de AS3MT (rs3740393, rs3740400, rs7085104, rs10748835, rs11191439) en las poblaciones de estudio y en algunas poblaciones obtenidas del proyecto de 1000 Genomas.	52
Figura A1.	Análisis de componentes factoriales mostrando la distribución de los individuos de las poblaciones de 1000 Genomas. El círculo en rojo señala a los individuos Mexicanos en Los Ángeles (MXL).	85
Figura A2.	Análisis Multidimensional a Escala (MDS) de la evaluación <i>in silico</i> con 28 Marcadores Informativos de Ancestría obtenido de los datos de las poblaciones reportadas en el proyecto de 1000 Genomas.	86
Figura 3A.	Análisis de componentes factoriales mostrando la distribución de los individuos de las poblaciones de 1000 Genomas con 34, 36 y Marcadores Informativos de Ancestría reportados.	87

ABREVIATURAS

AIMs – Marcadores Informativos de Ancestría
AMOVA – Análisis Molecular de Varianza
AS3MT – Gen de la Arsénico-3-metiltransferasa
BCL-2 – Linfoma de células B 2
Chr – Cromosoma
DMA – Ácido dimetilarsénico
DNA – Ácido desoxirribonucleico
DOF – Diario Oficial de la Federación
 F_{ST} – Índice de fijación
GI – Gastrointestinal
GSTMI – Gen de la Glutación S-transferasa Mu 1
GSTO – Gen de la Glutación S-transferasa Omega
GSTTI – Gen de la Glutación S-transferasa Theta 1
IARC – Agencia Internacional para la Investigación en Cáncer
iAs – Arsénico Inorgánico
IBD – Idénticos por descendencia
INEGI – Instituto Nacional de Estadística y Geografía
LD – Desequilibrio de Ligamiento
MDS – Análisis Multidimensional a Escala
MMA – Ácido monometilarsénico
mtDNA – DNA mitocondrial
NA – Nativa Americana
NOM – Norma Oficial Mexicana
NRY – Región no recombinante del Cromosoma Y
OMS – Organización Mundial de la Salud
ONU – Organización de las Naciones Unidas
PCA – Análisis de Componentes Principales
ppb – Partes por billón
QC – Quebrada Camarones
qPCR – Reacción en Cadena de la Polimerasa cuantitativa
SAC – San Antonio de los Cobres
SNP – Polimorfismo de cambio de un solo nucleótido
SSA – Secretaría de Salud

1. Introducción

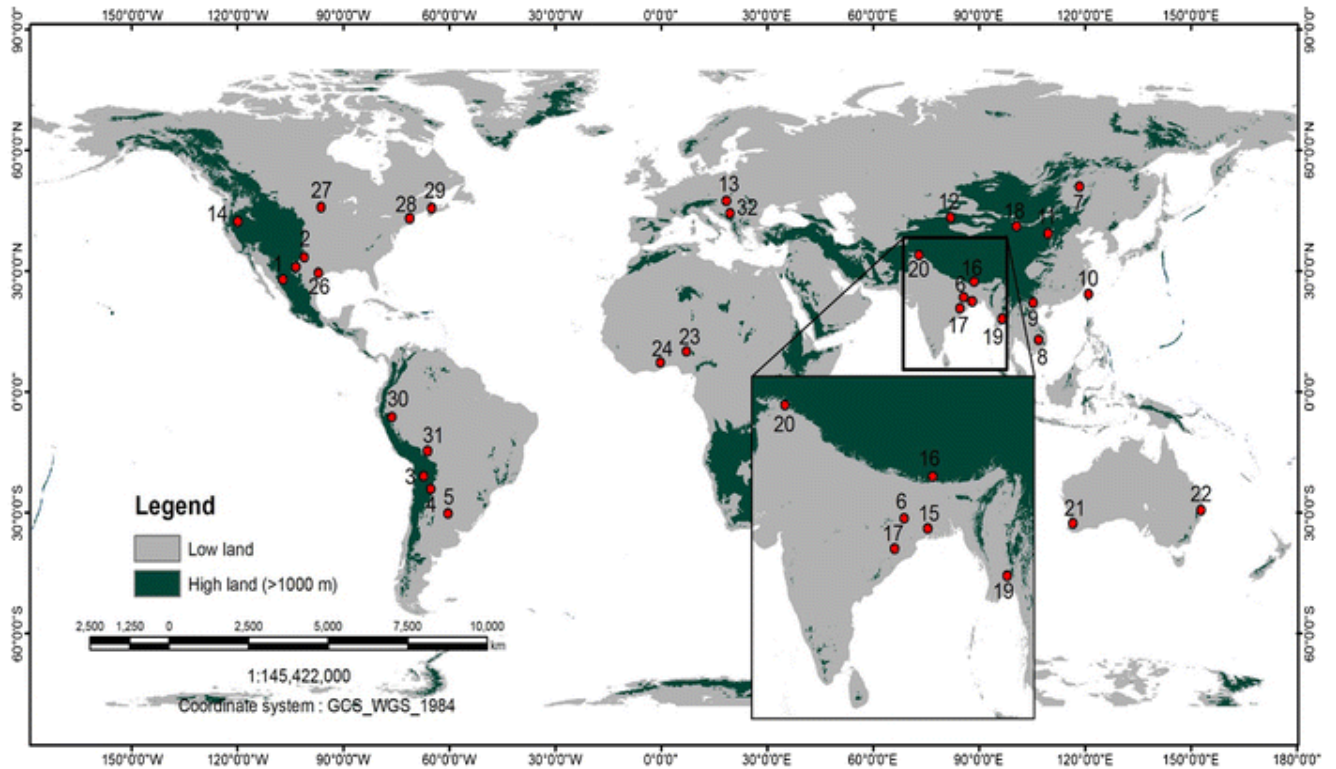
1.1 Arsénico inorgánico: generalidades

El arsénico (símbolo químico **As**, número atómico 33; 74.922 unidades de masa atómica) es considerado un metaloide por sus propiedades fisicoquímicas (www.rsc.org). De acuerdo con el potencial de hidrógeno y las condiciones de oxidación-reducción del medio, el As se puede encontrar en forma trivalente (As^{3+}) y pentavalente (As^{5+}), siendo la trivalente la especie predominante (<https://www.atsdr.cdc.gov/es/csem/arsenic/>). Este elemento químico, se utiliza ampliamente en la industria de transistores, láseres, semiconductores, procesamiento de vidrio, pigmentos, textiles, papel, adhesivos y preservantes de madera (www.who.int). De manera más limitada se utiliza como plaguicida, en aditivos de alimentos y fármacos (Alarcón, *et al*, 2013; www.who.int).

En su forma inorgánica (**iAs**), el As se encuentra ampliamente distribuido, ocupando el lugar veinte en abundancia a nivel mundial (Herath, *et al.*, 2016). Con respecto a su origen, éste puede ser geogénico (común en atmósfera, hidrósfera y biosfera) y antropogénico (principalmente por actividad minera). En ambos casos, el iAs es movilizadado al ambiente, siendo el agua subterránea y superficial los recursos más afectados, los cuales representan el 97% del agua como recurso de bebida alrededor del mundo (Montoya, *et al.*, 2015). Consecuentemente, se estima que entre 94 y 220 millones de personas en más de 50 países se encuentran potencialmente expuestas a iAs, entre los que destacan Argentina, Bangladesh, Chile, China, Estados Unidos, India y México (Figura 1); países en los que se concentra aproximadamente la mitad de la población en riesgo (Herath, *et al.*, 2016; Podgorski & Berg, 2020; www.int.org). De acuerdo con la Organización Mundial de la Salud (OMS), el nivel permisible de iAs en agua de bebida es de 10 $\mu\text{g/L}$ o ppb (partes por billón) (<https://www.who.int/en/news-room/fact-sheets/detail/arsenic>); en México la Norma Oficial Mexicana señala un límite permisible de hasta 25 $\mu\text{g/L}$ (NOM-127-SSA1-1994). De acuerdo con la Agencia Internacional para la Investigación en Cáncer (IARC, por sus siglas en inglés), el As y sus metabolitos han mostrado suficiente evidencia de capacidad carcinogénica en humanos, por lo que se encuentran clasificados en el grupo uno. (IARC, 2018).

1.2 Situación del Arsénico en México

El territorio mexicano cuenta con las dos regiones de mayor concentración de iAs en aguas subterráneas a nivel mundial: 1) la Comarca Lagunera y 2) la Región Centro-Norte (que abarca los estados de Guanajuato, Hidalgo y San Luis Potosí; [Alarcón, et al., 2013](#)). Asimismo, se han identificado tres principales ambientes hidrogeológicos con presencia de iAs: 1) los acuíferos aluviales (norte y centro del país), 2) áreas de actividad minera (zona norte y centro) y 3) las aguas geotérmicas asociadas a rocas volcánicas (zona centro en Jalisco, Michoacán y Puebla; [Alarcón, et al, 2013](#); [Montoya, et al, 2015](#); [Organización Mundial de la Salud, 2010](#)). Sin embargo, la mayor exposición se concentra en los estados áridos del centro-norte de México ([Alarcón-Herrera, et al., 2020](#)). Son muchos los estados en México con aguas contaminadas con iAs en donde los niveles de contaminación exceden el límite permitido por la NOM-127-SSA1-1994 y, por ende, las 10 ppb establecidas por la OMS ([Tabla 1](#)). Recientemente, se ha estimado que el número de personas sobreexpuestas a iAs a través de diversas fuentes de abastecimiento de agua en territorio nacional es cercano a 9 millones ([Alarcón-Herrera, et al., 2020](#); [Armienta, 2015](#)). En consecuencia, la exposición a iAs en México representa un problema de salud pública ([Alarcón-Herrera, et al., 2020](#)).



Location No	Location	Location No	Location
1	Lancunaria District, Mexico	17	Central India
2	North Central Mexico	18	Mongolia
3	Antofagasta, Chile	19	Myanmar
4	North West Argentina	20	Pakistan
5	Chaco-Pampean Plain, Argentina	21	Perth
6	Red River Delta	22	New South Wales
7	China	23	Nigeria
8	Mekong River Delta, Lao PDR, Cambodia	24	Ghana
9	Vietnam	25	Southern Texas
10	Taiwan	26	Northern Texas
11	Shanxi, China	27	Minnesota
12	Xinjiang, Tianshan plain	28	New England
13	Great Hungarian Plain, Hungary, Romania	29	Canada
14	Western USA	30	Peru
15	Bengal Basin	31	Bolivia
16	Nepal	32	Serbia

Figura 1. Mapa de la distribución mundial del iAs en agua subterráneas (Herath, *et al*; 2016).

Tabla 1. Estados de la República Mexicana con determinaciones de iAs mayores a 25 ppb en aguas subterráneas (Alfaro de la Torre, *et al.*, 2018).

Estado de la República	Municipio	Concentración de As (ppb)
Aguascalientes	Cosío	0.045
Baja California Sur	Comondú	0.030
	La Paz	0.025-0.500
	Loreto	0.030
	Los Cabos	0.005-0.100
	Mulegé	0.030-0.080
Chihuahua	La Paz	0.090-0.459
	Aldama	0.183
	Allende	0.029-0.048
	Ascensión	0.091
	Camargo	0.026-0.086
	Casas Grandes	0.032
	Chihuahua	0.114
	Coronado	0.032-0.045
	Delicias	0.027-0.053
	Galeana	0.029
	Julimes	0.042-0.105
	Jiménez	0.041-0.149
	La Cruz	0.111
	Meoqui	0.033-0.428
	Namiquipa	0.338
Nuevo Casas Grandes	0.025	
Ocampo	0.206	
Rosales	0.026-0.169	

	San Francisco de Conchos	0.030
	Satevó	0.144
	Saucillo	0.042-0.070
Coahuila de Zaragoza	Matamoros	0.031
	Sierra Mojada	0.079
	Torreón	0.209
	Viesca	0.046
Durango	Canatlán	0.039-0.103
	Cuencamé	0.100
	Durango	0.025-0.242
	El Oro	0.039-0.051
	General Simón Bolívar	0.048
	Gómez Palacios	0.121
	Indé	0.129
	Lerdo	0.042-0.136
	Mapimí	0.238
	Mezquital	0.068
	Nombre de Dios	0.026
	Nuevo Ideal	0.120
	Peñón Blanco	0.067-0.239
	Poanas	0.027
	Rodeo	0.039
	San Bernardo	0.051
	San Juan del Río	0.031
	Santiago Papasquiaro	0.040-0.370
Guanajuato	Tepehuanes	0.028
	Abasolo	0.025-0.029
	Cuerámaro	0.160-0.170
	Dolores Hidalgo	0.050-0.149
	Irapuato	0.050-0.077

	San Diego de la Unión	0.025-0.063
	San Luis de la Paz	0.028-0.120
	San Miguel de Allende	0.025-0.080
Hidalgo	Zimapán	0.046-0.358
Jalisco	Silao de la Victoria	0.033
	Atotonilco El Alto	0.045
	Ayotlán	0.051-0.127
	Ixtlahuacán de los Membrillos	0.029
México	Tenancingo	0.086
Michoacán de	Cuitzeo	0.036
Ocampo	Tlajomulco de Zúñiga	0.032
Nayarit	Acaponeta	0.045
Oaxaca	Heroica Ciudad de Huajuapán	0.025
	San Pablo Villa de Mitla	0.237
San Luis Potosí	Ahualulco	0.033-0.039
	Charcas	0.096
	Guadalcázar	0.025-0.051
	La Piedad	0.035
	Mexquitic de Carmona	0.062
	Salinas de Hidalgo	0.042
	Santo Domingo	0.025-0.062
	Villa de Guadalupe	0.061
	Villa de Ramos	0.042-0.195
Sinaloa	Ahome	0.029-0.089
	Mazatlán	0.027
	Navolato	0.029
Sonora	Caborca	0.034
	General Plutarco Elías Calles	0.028-0.077
	San Luis Río Colorado	0.068
	Santa Ana	0.069

	Trincheras	0.028
Tabasco	Cárdenas	0.031
Veracruz de I de la Llave	Veracruz	0.026
Zacatecas	Concepción del Oro	0.038
	El Salvador	0.061
	Hermosillo	0.048-0.050
	Fresnillo	0.032
	General Francisco R.Munguía	0.091
	General Pánfilo Natera	0.038
	Guadalupe	0.068
	Huanusco	0.028-0.051
	Juchipila	0.037
	Luis Moya	0.055-0.099
	Mazapil	0.179
	Miguel Azua	0.098
	Norial de Ángeles	0.055
	Ojocaliente	0.106
	Río Grande	0.094
	Trancoso	0.034
	Villa de Cos	0.400
	Villa González Ortega	0.033
	Zacatecas	0.030-0.038

1.3 Toxicocinética del As

1.3.1 Vías de exposición

Las principales rutas de exposición al As en el humano son por ingestión de agua o alimentos contaminados y por la vía inhalatoria ([Organización Mundial de la Salud, 2010](#)). En la exposición por agua de bebida, se estima un rango de consumo de 20-300 µg/día de As (www.monographs.iarc.fr). Mientras que en aquellas regiones en donde el As no se encuentra

de forma natural, el consumo es a través de alimentos contaminados, los cuales tienen 0.25 mg/kg de As en promedio, pudiendo contener hasta 100 mg/kg de As en algunos crustáceos (www.monographs.iarc.fr). La vía inhalatoria, , corresponde a la principal vía de exposición ocupacional aunque en menor medida que la oral, con ingesta de 20-200 ng en áreas rurales, 400-600 ng en ciudades (sin emisiones industriales) y de hasta 1µg/día en fumadores (www.monographs.iarc.fr; Organización Mundial de la Salud, 2010).

1.3.2 Absorción, distribución, metabolismo y excreción del arsénico

El iAs que ingresa al organismo (80-90%) es absorbido directamente en el tracto gastrointestinal (GI) como arsenito (AsIII) o arsenato (AsV) (<https://www.atsdr.cdc.gov/>). No obstante, la ingestión de As a través de los alimentos es de absorción variable, el cual se ha encontrado biodisponible hasta un 89.4% en el tracto GI (<https://www.atsdr.cdc.gov/>). Otras rutas de absorción son la pulmonar (60-90%), vía parenteral y absorción dérmica, si bien esta última no está bien determinada, se considera que se encuentra en muy bajo grado ([Kuivenhove & Mason, 2020](#)).

Una vez que el As ingresa al torrente sanguíneo se distribuye y acumula en diversos órganos a través de su unión con la globina de la hemoglobina eritrocitaria (www.monographs.iarc.fr). Entre los principales órganos de acumulación del As se encuentran el riñón, el pulmón, la vejiga, la piel, la sangre y el hígado; el As puede también depositarse en cabello, uñas, piel y, en menor grado, en huesos y dientes (<https://www.atsdr.cdc.gov/>).

Con respecto al metabolismo del As, éste se lleva a cabo mayoritariamente en el hígado, en donde en su forma tri o pentavalente es metilado por la arsénico-3-metiltransferasa (AS3MT), generando como productos ácidos de esta reacción a las formas monometilada (MMA) y dimetilada (DMA; [Drobná, et al., 2010](#); [Khairul, et al., 2017](#)). En esta reacción también participan la S-adenosilmetionina (SAM) y el glutatión (GSH), donadora del grupo metilo y cofactor de la AS3MT, respectivamente ([Figura 2](#)). Otras enzimas reductoras que pueden metabolizar el iAs son la Glutathion S-transferasa omega, mu-1 o theta-1 (GSTO, GSTM1 o GSTT1, respectivamente), con las que se forman conjugados arsénico-glutatión ([Khairul, et al., 2017](#); [Hirano & Watanabe, 2013](#)).

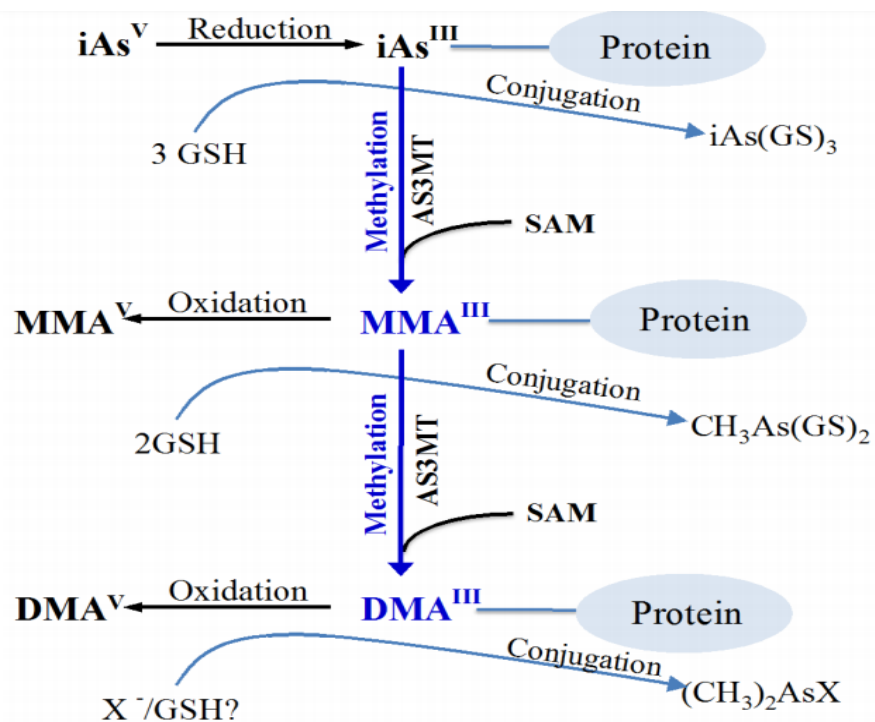


Figura 2. Esquema del metabolismo del iAs por la arsenico-3-metiltransferasa (Obtenida de Khairul, *et al.*, 2017).

Nota: GSH, Glutati3n; iAs, ars3nico inorg3nico; $iAs(GS)_3$, Triglutati3n de iAs; $CH_3As(GS)_2$, Diglutati3n metilars3nico, MMA, 3cido monometilars3nico; DMA, 3cido dimetilars3nico; $(CH_3)_2AsX$, Conjugado de dimetilars3nico; SAM, S-adenosilmetionina.

La capacidad citot3xica del iAs y sus metabolitos es completamente dependiente de su estado de oxidaci3n y estructura qu3mica. Las formas metiladas pentavalentes del iAs son menos citot3xicas, mientras que las metiladas trivalentes son altamente cito- y genot3xicas (Khairul, *et al.*, 2017). El da1o mediado por el iAs se ha relacionado con la generaci3n de especies reactivas de ox3geno, acumulaci3n de calcio, sobreexpresi3n de la caspasa-3, disminuci3n de genes anti-apopt3ticos como *BCL-2* (por sus siglas en ingl3s, *B cell lymphoma 2*) y con la deficiencia de la prote3na p53 (Ghosh & Sil, 2015). Por otro lado, tambi3n se han observado respuestas diferentes en cultivos de granulocitos y macr3fagos por su exposici3n a bajas o altas concentraciones de iAs evidencia del efecto conocido como hormesis (Ferrairo, *et al.*, 2008).

1.4 Efectos del arsénico en la salud humana

Debido al tamaño de la población que ingiere iAs a través del agua de bebida y su asociación con distintas enfermedades, la exposición a iAs es de interés mundial en salud pública (Fisher, *et al.*, 2015). Los efectos tóxicos por la ingesta de As pueden presentarse de manera aguda o crónica (www.who.int). La mayoría de los casos de intoxicación aguda ocurren mediante la ingestión accidental de plaguicidas o de manera menos común por intento de suicidio (Ratnaïke, 2003). La intoxicación por cantidades menores a 5 mg de As produce vómitos y diarreas que terminan en menos de doce horas sin necesidad de tratamiento (Ratnaïke, 2003). Otros síntomas clínicos relacionados son: náuseas, deshidratación, estados de *shock*, dolor abdominal, salivación excesiva, sarpullido, convulsiones y delirios; incluso se han reportado daños agudos en órganos como corazón, riñón y pulmón (Kuivenhoven & Mason, 2019; Ratnaïke, 2003). Por lo que se refiere a los efectos de la exposición crónica (por agua de bebida o alimentos), éstos pueden ser cambios de pigmentación o lesiones en la piel, así como hiperqueratosis en palmas de las manos y plantas de los pies; efectos que se observan tras una exposición de por lo menos cinco años (<https://www.who.int/ipcs/assessment/publichealth/arsenic/en/>). Los efectos en la piel pueden ser precursores de cáncer en el mismo órgano, aunque también se ha relacionado con cáncer de vejiga o pulmón, neuropatías, diabetes, enfermedades cardiovasculares, así como deficiencias en la capacidad cognitiva (Fisher, *et al.*, 2017). Al 2015, las enfermedades cardiovasculares y diabetes fueron las dos primeras causas de muerte en México, las cuales representan el 20 y 15% del total de fallecimientos en ese año, respectivamente, seguidas por los casos de cáncer (12%) (INEGI, 2015). Se ha estimado que cerca de 13,070 casos de cáncer en la población Mexicana se deben únicamente a la exposición a As (Alarcón-Herrera, *et al.*, 2020). Las manifestaciones clínicas de la intoxicación por As varían entre individuos, grupos poblacionales y áreas geográficas (Agusa, 2011). Sin embargo, se desconoce cuáles son los factores determinantes en la aparición de signos clínicos o qué sistema en específico del cuerpo humano será el afectado cuando hay exposición a este metaloide (Ratnaïke, 2003).

1.5 Diversidad genética de las poblaciones humanas

La diversidad genética dentro y entre las poblaciones es el resultado de la interacción de las fuerzas evolutivas (deriva génica, mutación, selección natural y migración)(Griffiths, *et al.*, 2000). Las variaciones genéticas son las diferencias en la secuencia del DNA (por sus siglas en inglés, *Deoxyribonucleic acid*) entre los individuos de una población (Griffiths, *et al.*, 2000). Consecuentemente, la determinación de estas diferencias permiten el estudio de la variación genética y del efecto de las fuerzas evolutivas sobre una población (Fontdevila, 2000).

A partir de la secuenciación completa del genoma humano finalizada en 2003 (<https://www.genome.gov/human-genome-project>), se han identificado variaciones genéticas no mayores al 15% entre los individuos de distintas poblaciones en los cinco continentes (The International HapMap Consortium, 2005). Las variantes genéticas más comunes son los polimorfismos de un solo nucleótido (SNP, por sus siglas en inglés) de los cuales se han descrito al menos 15 millones dentro del genoma humano (The 1000 Genomes Project Consortium, 2010). Los SNPs representan la sustitución de una base nucleotídica por otra dentro de una secuencia de DNA, por lo tanto, un mismo *locus* puede tener diferentes alelos (bialélico). Estas variantes se clasifican en (Rahim, *et al.*, 2008):

- Variantes comunes en donde el alelo de menor frecuencia se encuentra en al menos el 5% de la población) y
- Variantes raras, presentes en menos del 1% de la población.

En este sentido, la determinación de las frecuencias de las variantes alélicas constituye una medida de la variación genética poblacional y es determinada por técnicas de biología molecular (Griffiths, *et al.*, 2000). El estudio de las variantes permiten diferenciar entre los individuos de una población y, por ende, entre las diferentes poblaciones, proporcionando así información de la historia poblacional y de los patrones de migración (Witherspoon, *et al.*, 2007). En este sentido, los datos recabados de una población permiten describir su diversidad, antropología genética, filogeografía, mapas génicos, identificar marcadores de susceptibilidad e inferir la carga ancestral y la estructura poblacional, estas últimas dos utilizadas como covariables en estudios de asociación (Elhaik, *et al.*, 2014; Marchini, 2004).

1.5.1 Estructura poblacional

La estructura poblacional es uno de los aspectos más estudiados en genética de poblaciones que ha alcanzado una alta sofisticación matemática utilizando el análisis estadístico con profusión (Fontdevila, 2000). En los últimos 40 años, a través de la biología molecular, se han estudiado diferentes sistemas de marcadores genéticos para abordar este aspecto. Entre ellos se encuentran los patrones de variantes genéticas en marcadores uniparentales (región no recombinante del cromosoma Y -NRY- y en el DNA mitocondrial, mtDNA) así como en los autosómicos (Barrandeguy & García, 2014). Estos sistemas son los más utilizados para evaluar la diversidad genética e inferir la estructura poblacional. Sin embargo, el análisis poblacional a través de la NRY y mtDNA pueden mostrar diferencias debido a que solo representan un linaje parental por su heredabilidad; mtDNA en el caso materno y el cromosoma Y en el caso paterno (Salzano & Sans, 2014). Adicionalmente, el estudio con diferentes tamaños poblacionales y la variación en la tasa de migración entre hombres y mujeres, contribuyen a diferencias en la inferencia de la estructura genética (Ramachandran, *et al.*, 2004).

Los marcadores autosómicos como los marcadores informativos de ancestría (AIMs) son polimorfismos tipo SNP altamente abundantes en el genoma y cuyas frecuencias alélicas difieren con respecto a la región geográfica (Kosoy, *et al.*, 2009). Con base en estas características, estos marcadores han sido ampliamente utilizados para conocer los orígenes ancestrales de un individuo y describir la estructura poblacional. Sin embargo, carecen de un fundamento sistemático en su selección, validación y número mínimo de marcadores para realizar la inferencia ancestral (Galanter, 2012). Una de las ventajas del uso de este tipo de marcadores es que reflejan ancestría por regiones geográficas y no un linaje uniparental, como sucede en el análisis de la NRY y mtDNA (Galanter, 2012; Kidd, 2013).

1.5.2 Estructura poblacional de México

La población mexicana contemporánea posee una estructura genética compleja como consecuencia de la heterogeneidad de sus poblaciones parentales (Nativos Americanos, Europeos y esclavos Africanos) y de su reciente surgimiento (500 años; Johnson, *et al.*, 2011). Los estudios con marcadores uniparentales han evidenciado que desde el ángulo

patrilíneal los dos linajes más prominentes son el Europeo (proveniente principalmente de la península Ibérica) y el Nativo Americano, mientras que del lado materno predomina el componente Nativo Americano (Salazar-Flores, *et al.*, 2015). La mezcla poblacional entre estas dos poblaciones parentales dio origen a la población mestiza, la que constituye aproximadamente el 93% de la población mexicana actual (Martínez-Cortés, *et al.*, 2012). No obstante, el proceso del mestizaje no se llevó a cabo de forma homogénea, pues estuvo influenciado por el sesgo de género, por los orígenes de las diferentes poblaciones parentales de las que surgió y por los eventos demográficos experimentados por éstas (Silva-Solezzi, *et al.*, 2009). Entre los más destacables se encuentran los prominentes cuellos de botella -el más lejano hace al menos 10 mil años y el más próximo hace 500 años-, los procesos de diversificación en Beringia y Mesoamérica, el constante flujo genético propiciado por la riqueza de recursos del territorio mexicano y los movimientos migratorios contemporáneos entre las zonas rurales a las urbanas (Gómez, *et al.*, en prensa; González-Sobrino, *et al.*, 2016; Sandoval, *et al.*, 2012).

La otra cara de la diversidad mexicana está conformada por las más de 60 poblaciones Nativas provenientes de diversas familias lingüística, que al 2015 se representó por más de 7 millones de personas (el 6.6% del total de la población actual; INEGI, 2015). Los grupos Nativos de diferentes regiones geográficas han mostrado estratificación genética variable, por lo tanto, su contribución en la población mestiza tiene diferentes patrones genéticos regionales (Moreno-Estrada, *et al.*, 2014). Al menos tres migraciones provenientes de Asia se dispersaron y contribuyeron en la diversidad de las primeras poblaciones Nativas en América (Reich, *et al.*, 2012; Shurr & Sherry, 2004), posteriormente los eventos sociodemográficos pre-colombinos entre grupos Nativos ampliaron la diversificación genética (Martínez-Cortés, *et al.*, 2013; Moreno-Estrada, *et al.*, 2014). Sin embargo, su aislamiento durante y después de la colonización limitó el flujo génico, en comparación con el resto de la población mexicana (Belbin, *et al.*, 2018).

1.6 Posible presión selectiva por la exposición a As y sus implicaciones en el metabolismo.

La metilación del iAs es considerada la bioactivación y vía de eliminación de este metaloide (Watanabe & Hirano, 2013). El análisis *in silico* de la secuencia del gen *AS3MT* señala que la transferencia horizontal de genes (intercambio de información genética sin emparejamiento) de bacterias a eucariontes sucedió al menos cinco veces para este gen (Palmgren, et al., 2017). El estudio filogenético del gen también indica la necesidad de organismos de distintos reinos y, de complejidad variable, de convertir el iAs en compuestos más solubles para eliminarlos (Figura 2). Adicionalmente, los estudios *in vitro* han determinado variaciones en la capacidad de la generación de los productos metilados entre especies (Drobná, et al., 2010) Lo anterior sugiere que ambientes con altas concentraciones de iAs influyeron en la selección y mejoramiento de la capacidad de metilación por parte de la *AS3MT*, resultado que no ha sido observado en algunos organismos que se desarrollaron en ambientes poco contaminados con el metaloide (Palmgren, et al., 2017).

La variación fenotípica humana como el color de ojos y piel o el grupo sanguíneo es consecuencia directa de la diversidad genética, la que a su vez responde a la presión selectiva del ambiente y a factores conductuales (estilo de vida) (Rahim, et al., 2008). En la especie humana, la principal fuente de eliminación del As es la orina, en una proporción del 10-30% de iAs, 10-20% de MMA y 60-70% de DMA (Vhater, 1999). Estas proporciones han sugieren una variación interindividual, lo que se ha atribuido a variantes genéticas relacionadas con la capacidad de metilación del As (Janasik, 2015). En este sentido, se han descrito más de 30 SNPs en el gen que codifica a la *AS3MT* (Agusa, 2011; Tabla 2) el cual se encuentra localizado en el cromosoma diez (10q24.32; www.genecards.org). Algunos genotipos de estos SNPs se han relacionado con la variabilidad en la capacidad de metilación de esta transferasa y, por lo tanto, posiblemente con las diferencias en la susceptibilidad a los efectos tóxicos del As, principalmente en poblaciones Asiáticas y del Continente Americano (Agusa, 2011).

Tabla 2. Variantes genéticas del gen de la AS3MT (Chr: 10q24.32) que han sido relacionadas a una mayor capacidad de metilación.

Polimorfismo	Localización en el gen	Alelo de eficiencia	N	Población	Referencia
rs3740393 (C/G)	Intrón 1	G	135	Valle Yaqui, Sonora, México	Meza, <i>et al.</i> , 2007.
			148	San Antonio de los Cobres, Argentina	Engström, <i>et al.</i> , 2007
			1,078	Chiayi County, Taiwán.	Chung, <i>et al.</i> , 2009
			102	Comarca lagunera, México	García-Alvarado, <i>et al.</i> , 2018
rs3740400 (A/C)	Intrón 6	C	104	San Antonio de los Cobres, Argentina.	Engström, <i>et al.</i> , 2009
			176 y 500	Argentina y Bangladesh	Engström, <i>et al.</i> , 2011
			200	Gómez Palacio, Durango México	Drobná, <i>et al.</i> , 2016
			450	Wuyuan County, China	Chen, <i>et al.</i> , 2017
10748835 (A/G)	Intrón 10	A	148	San Antonio de los Cobres, Argentina.	Engström, <i>et al.</i> , 2007.
			100	Ha Nam, Vietnam	Agusa, <i>et al.</i> , 2009.
			172	San Antonio de los Cobres, Argentina	Engström, <i>et al.</i> , 2011
			361	Dhaka, Bangladesh	Engström, <i>et al.</i> , 2011
rs11191439 (T/C)	Exón 8	T	415	Ciudades de Hungría, Rumania y Eslovaquia.	Lindberg, <i>et al.</i> , 2007
			50	Chuquicamata, Chile	Hernández, <i>et al.</i> , 2008
			100	Ha Nam, Vietnam	Agusa, <i>et al.</i> , 2009
			1,062	Ciudades de Hungría, Rumania y Eslovaquia	Engström, <i>et al.</i> , 2015
rs7085104 (A/G)	Región 5' terminal	G	104	San Antonio de los Cobres, Argentina.	Engström, <i>et al.</i> , 2009
			122	Zimapán, Hidalgo México	Valenzuela, <i>et al.</i> , 2009

Son muchas las regiones del mundo en donde el humano ha tenido contacto con fuentes de agua contaminada con As durante miles de años. Tal es el caso de la región del altiplano de los Andes, el que ha estado habitado por al menos 12,000 años antes del presente (AAP; Jocas, *et al.*, 2018). Incluso, se ha encontrado As en la fauna de las regiones que conformaron Beringia, por lo que el contacto pudo darse antes del poblamiento de América. A pesar de que la información sobre la adaptación humana es limitada, varios grupos de estudio han explorado la diversidad del *AS3MT* en poblaciones Nativas ancestrales y contemporáneas de esta región mediante estudios paleogenéticos y genéticos, respectivamente (Apata, *et al.*, 2017, De Loma, *et al.*, 2019; Schlebusch, *et al.*, 2013, Schlebusch, *et al.*, 2015). Los resultados sugieren que las poblaciones expuestas a concentraciones altas de As podrían poseer un metabolismo más eficiente, en cuanto a la metilación del As, en comparación con aquellas que han estado poco expuestas (Apata, *et al.*, 2017; Eichstaedt, *et al.*, 2015; Schlebusch, *et al.*, 2015). Estas diferencias se ven reflejadas en la arquitectura genética en cuanto a los polimorfismos del *AS3MT*; las poblaciones de la región de los Andes que se encuentran sobreexpuestas a iAs (~200 ppb) exhiben una arquitectura distinta a aquellas poblaciones geográficamente cercanas pero cuya exposición es cercana al límite permisible (10 ppb; Eichstaedt, *et al.*, 2015; Schlebusch, *et al.*, 2015). Estos datos sugieren que la presión ambiental (sobreexposición a As) durante miles de años, pudo haber permitido que los pobladores de los Andes fueran seleccionados en el sentido de presentar variantes de la *AS3MT* con alta capacidad de metilación (Eichstaedt, *et al.*, 2015; Schlebusch, *et al.*, 2015). Esta selección positiva es conocida como la “Hipótesis arsenical”, mientras que al conjunto de alelos de los SNPs rs3740393-C, rs3740390-T y rs10748835-A, con una mayor capacidad de metilación, se les ha considerado como un posible haplotipo de protección (Apata, *et al.*, 2017; Schlebusch, *et al.*, 2013). Este haplotipo junto al alelo A del rs11191439 (CTTA) se ha encontrado en mayor frecuencia (0.68) en una población Nativa del Norte de Chile (Quebrada Camarones -QC-, perteneciente a las regiones de Arica y Paracota) en donde las concentraciones de iAs son cercanas a 1,312 ppb (Apata, *et al.*, 2017). Por su parte, la población Nativa de San Juan de la Costa (Provincia de Osorno, Región Los Lagos, Sur de Chile), expuestos a concentraciones de iAs <10 ppb, presentan una frecuencia del haplotipo de «protección» casi nueve veces menor (0.08) al de la población con concentraciones cercanas a 1,312 ppb. Con respecto a la población del Valle de Azapa, también al norte de Chile pero de menor exposición a iAs (10-20 ppb), la frecuencia de CTTA

(0.48) fue menor que en QC (Apata, *et al.*, 2017). Situaciones similares se han reportado en poblaciones de mujeres pertenecientes a etnias Nativas y geográficamente aisladas como sucede en Argentina y Bolivia, todas ellas expuestas a niveles mayores a 10 ppb de iAs. Estos grupos de mujeres mostraron un metabolismo eficiente con bajo porcentaje de MMA y altos niveles de DMA en orina. (De la Loma, *et al.*, 2019; Schlebusch, *et al.* 2015).

En el contexto del poblamiento de América, los orígenes de los pueblos Nativos han sido referidos a las regiones Centro-Sur y Noreste de Siberia, lo que refuerza la entrada al continente vía Beringia (Gómez R, *et al en revisión*; Shurr & Sherry, 2004). Los estudios paleogenéticos han sugerido que las primeras migraciones llegaron a América entre 14,600 a 17,500 AAP (Shurr & Sherry, 2004). Para que estas poblaciones hayan subsistido, debieron estar en contacto con mantos acuíferos. De modo interesante, los mapas que ilustran la distribución mundial de iAs en aguas subterráneas son muy semejantes a aquellos que sugieren las rutas de colonización que siguieron los primeros Americanos (Figuras 1 y 3). Lo anterior también sugiere una exposición prolongada al iAs a través del agua de bebida y, por consiguiente, el inicio del proceso de selección al iAs miles de años antes, como se sugiere en la región Andina. Hidrográficamente, las poblaciones Europeas no parecen tener exposición relevante a iAs (Podgorski & Berg, *et al.*, 2020), por el contrario, se estima que las poblaciones Asiáticas concentran el 94% del total de la población en contacto con iAs (Podgorski & Berg, *et al.*, 2020). Sin embargo, las poblaciones de América, debido al nivel de exposición de algunos individuos (hasta 1,000 ppb), también han sido ampliamente estudiadas (Herath, *et al.*, 2016), como es el caso de Argentina y Chile. En México, las zonas áridas del centro-norte de México es en donde se concentran la mayoría de los estudios poblacionales de exposición a As (Alarcón-Herrera, *et al.*, 2020; Herath, *et al.*, 2016).

La estructura genética y ancestral de una población son covariables importantes para el diseño e interpretación de estudios de asociación genética. (Marchini, 2004). El desconocimiento de estas características es una fuente importante de sesgo en el análisis estadístico de los datos pudiendo generar falsas asociaciones genotipo-fenotipo (Marchini, 2004). Las poblaciones de estructura genética compleja, como las Latinas, son el reflejo de diversos procesos demográficos con no más de 15 generaciones, favoreciendo una estratificación poblacional (Belbin, *et al.*, 2018; Johnson, *et al.*, 2011) y una compleja arquitectura genética, además de ser testimonio de los procesos ambientales a los que han

sido expuestos nuestros ancestros (Tishkoff, *et al.*, 2019). La población mexicana posee un importante y ampliamente diverso componente Nativo Americano que se mezcló con, el no menos diverso, componente europeo en conjunto con un minoritario componente africano. Estas poblaciones parentales han permeado, incluso, entre las poblaciones Nativas, algunas de las cuales han permanecido geográficamente aisladas y profundamente arraigadas a su pasado prehispánico (Martínez-Cortés, *et al.*, 2012). El mosaico de diversidad de la población mexicana de características multiétnicas, es una población ideal para estudiar la variación genética del gen *AS3MT* y relacionar las variantes alélicas que favorezcan el metabolismo de agentes potencialmente dañinos, para el organismo humano, con la carga ancestral de un grupo de individuos mestizos y Nativos Americanos contemporáneos, dando pie al presente proyecto.



Figura 3. Mapa de la distribución del iAs en el continente Americano (Modificado de Apata & Pfeifer, *et al.*, 2020). Las regiones con depósitos de iAs mayores a 10 ppb se indican en puntos rojos. Las flechas en línea oculta sugieren las rutas migratorias en la colonización de América.

1.7 Justificación

El As, un metaloide distribuido de forma heterogénea en México se localiza principalmente en ambientes hidrogeológicos, sobrepasando hasta en 60 veces, el límite permitido por la Organización Mundial de la Salud (10µg/L). La movilización del iAs al ambiente por procesos naturales y antropogénicos favorece su exposición por las vías inhalatoria y oral (principalmente por agua de bebida). Se estima que en México cerca de 9 millones de habitantes se encuentran expuestos a este elemento químico. Como consecuencia, el iAs ha sido relacionado con problemas de salud pública como el cáncer de piel, las enfermedades cardiometabólicas, pulmonares y los trastornos neurológicos, mismos que han incrementado sus tasas de incidencia en los últimos años.

Estudios previos han sugerido que la exposición crónica al iAs pudo haber favorecido la fijación de variantes genéticas relacionadas con un metabolismo del iAs más eficiente. . La arsénico-3-metiltransferasa, enzima encargada de metilar al iAs y favorecer su eliminación del organismo, presenta una elevada variabilidad genética, la que ha sido relacionada con diferencias en su actividad enzimática. Esta variabilidad genética ha sido escasamente estudiada en nuestra población, cuya compleja arquitectura genética, caracterizada por una heterogeneidad en los grados de mestizaje, aunado a la exuberante diversidad de los grupos Nativos, la convierte en el escenario indicado para relacionar la carga ancestral con las variantes genéticas de la *AS3MT* en poblaciones Mexicanas.

1.8 Hipótesis

Las variantes genéticas relacionadas con una mayor tasa metabólica del iAs serán más frecuentes en las poblaciones con mayor carga Nativa en comparación con los individuos de la población mestiza Mexicana.

1.9 Objetivos

1.9.1 Objetivo General

Relacionar la ancestría regional con las variantes polimórficas de la arsénico-3-metiltransferasa en poblaciones mestizas y Nativas contemporáneas mediante marcadores informativos de ancestría.

1.9.2 Objetivos particulares

- Determinar las variantes alélicas de cinco polimorfismos de la arsénico-3-metiltransferasa en una muestra de individuos Nativos y mestizos mexicanos contemporáneos y establecer sus frecuencias alélicas, genotípicas y haplotípicas.
- Diseñar y evaluar *in silico* un panel de marcadores tipo SNP para determinar la ancestría regional de las poblaciones de estudio, a través de la selección de marcadores previamente reportados.
- Establecer la relación entre las frecuencias haplotípicas del gen *AS3MT* y la carga regional ancestral de las poblaciones estudiadas.

2. Métodos

2.1 Diseño de estudio

El presente estudio es de tipo transversal, descriptivo y comparativo. La información utilizada fue recolectada en los estados de México, Guerrero, Guanajuato, Oaxaca, Puebla, Querétaro y San Luis Potosí durante el periodo 2001 a 2019 como parte de dos estudios originales: “Determinación de la diversidad genética de la población mexicana” y “Determinación de las rutas migratorias para dilucidar los orígenes del hombre Americano mediante el estudio de la diversidad genética de las poblaciones étnicas mexicanas”. Ambos estudios estuvieron conformados por hombres y mujeres, no relacionados entre sí, mayores de 18 años pertenecientes a las etnias mestiza y Nativa Americana contemporáneas (Tabla 3) que acudieron al laboratorio de Biología Molecular Diagnóstica para determinar parentesco mediante huella digital de material genético (mestizos), o que fueron recolectados a través de los muestreos llevados a cabo en diversas localidades Nativas de diferentes regiones de México.

Tabla 3. Poblaciones de estudio pertenecientes a las etnias mestiza y Nativa Americana contemporáneas.

Población	Localización	n	Familia lingüística ¹
Mazahuas	Estado de México	29	Oto-Mangue
Me'Phaas	Sierra Alta de Guerrero	35	Oto-Mangue
Mestizos	Querétaro, Guanajuato, y Puebla	225	Indo-Europeo
Tsotsiles	Chiapas	20	Oto-Mangue

¹De acuerdo al CATÁLOGO de las Lenguas Indígenas Nacionales, DOF 14/01/18 disponible en:

<https://www.inali.gob.mx/clin-inali/>

2.1.1 Selección de sujetos

El grupo mestizo se integró por 225 individuos (106 hombres y 119 mujeres) nacidos en México con al menos tres generaciones de ancestros nacidos en el mismo país, que se autoidentificaron como pertenecientes a la etnia mestiza mexicana a través de un cuestionario. Estos individuos fueron obtenidos aleatoriamente a partir de un banco de 1,640 individuos mestizo mexicanos provenientes de la Ciudad de México y los estados de Guanajuato, Querétaro, Jalisco y Puebla (Noris, *et al* 2012; Santana, *et al.*, 2014).

El grupo Nativo Americano estuvo integrado por 84 individuos (hombres y mujeres) nacidos en México con al menos tres generaciones de ancestros nacidos en el sitio de la toma de muestra y que hablaban un idioma perteneciente a las diferentes familias lingüísticas Nativas Americanas en al menos tres generaciones. Estos individuos se auto reconocieron como pertenecientes a una étnica Nativa a través de un cuestionario validado por la Universidad de Pensilvania (*Institutional Review Board # 8*). El tamaño de muestra de las etnias Nativas fue por conveniencia.

2.1.2 Criterios de inclusión

Hombres y mujeres no relacionados entre sí a través de tres generaciones de ancestros que:

- Se autoidentificaron como pertenecientes a las etnias descritas en la tabla 3.
- Hablaban la lengua correspondiente al grupo de autoidentificación al menos en tres generaciones.
- Haber firmado las cartas de consentimiento informado y de información genealógica.

2.1.3 Criterios de exclusión

- Hombres y mujeres que tuvieran alguna relación de parentesco igual o menor de tercer grado de acuerdo con el estudio de marcadores autosómicos previamente realizado por el laboratorio Biología Molecular Diagnóstica S.A. de C.V.

2.1.4 Criterios de eliminación

Muestras en las que no se pudo llevar a cabo el 100% de los análisis moleculares requeridos para esta investigación.

2.2 Descripción de las variables

- Independientes.
 - Carga ancestral regional
- Dependientes.
 - Polimorfismos del gen que codifica a la arsénico-3-metiltransferasa (*AS3MT*).

2.3 Descripción operacional de las variables

- Polimorfismo genético. Presencia de múltiples variantes alélicas en un mismo *locus*; los alelos deben presentar una frecuencia > al 1% de la población (Forrester, *et al.*, 2016).
- Marcador Informativo de Ancestría. Polimorfismos cuya frecuencia relativa de los alelos que los conforman difiere por continente o región geográfica, permitiendo trazar la ancestría de un individuo con respecto al sitio geográfico de donde proceden sus orígenes (<https://www.genome.gov/genetics-glossary>).
- Polimorfismo de cambio de un solo nucleótido. Tipo de polimorfismo que se refiere al cambio de una base nitrogenada por otra, siendo el tipo de variación genética más común entre las poblaciones humanas (<https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms>).
- Mestizo. Individuo de origen multiétnico (Nativo Americano y Español, principalmente) nacido en México con antecedentes familiares de al menos tres generaciones de ancestros nacidos en el mismo país y que poseen apellidos derivados del Español (Sánchez-Serrano, C., 1996).
- Nativo Americano. Individuos con continuidad histórica desde tiempos pre- y postcoloniales relacionados con el poblamiento de América, que difieren en su sistema social, económico y político, además de tener un lenguaje, cultura y creencias propias y un fuerte vínculo con su territorio y sus recursos (UN, 2006).
- Gen *AS3MT*. Gen que codifica a la enzima que cataliza la transferencia de un grupo metilo de la molécula S-adenosil-L-metionina a compuestos arsenicales trivalentes, produciendo compuestos arsenicales monometilados o dimetilados (metilarsenato y dimetilarsenato) (<https://www.uniprot.org/uniprot/Q9HBK9>).

2.4 Estrategia experimental

2.4.1 Diseño y evaluación del panel para evaluar ancestría

Para inferir la carga ancestral en las muestras poblacionales se diseñó un panel de AIMs. La estrategia para diseñar y seleccionar el panel fue la siguiente:

- Se revisaron y agruparon AIMs tipo SNPs reportados previamente tanto en la literatura como en diversas bases de datos (Anexo [Tabla A1](#)).
- La primera elección de los marcadores se llevó a cabo tomando como criterio las distancias génicas a través del índice de fijación (F_{ST} ; 997).

2.4.1.1 Cálculo de F_{ST} : El cálculo se llevó a cabo a través del programa Arlequin v.3.5 (Excoffier, *et al.*, 2005), el cuál determina la diferenciación poblacional por pares (diferencia entre dos poblaciones) y la presenta en una matriz a través de los cálculos descritos por Reynolds (1983) y Slatkin (1995) derivados del método más usado el estadístico F , desarrollado por Sewall Wright (1951). El método divide la variabilidad genética medida por los niveles de heterocigosidad de los individuos de una población entre la variación de la población (Sewall Wright, 1951). La fórmula más empleada para describir la heterocigosidad de una población es la siguiente (McVean, 2001):

donde:

$$F_{ST} = \frac{H_T - \bar{H}_S}{H_T}$$

F_{ST} : índice de fijación

H_T : heterocigosidad total

H_S : promedio de heterocigosidad entra la población

Generalmente, una población de un área geográfica se encuentra dividida en subpoblaciones debido al flujo génico. El estudio poblacional a través del índice F_{ST} puede describir la variabilidad dentro y entre subpoblaciones y dentro de la población total. De esta manera es posible agrupar a individuos por regiones geográficas (Fontdevila, 2000). Los valores del índice obtenido van de cero (no hay diferencia entre las poblaciones) a uno (las poblaciones son genéticamente diferentes; Fontdevila, 2000).

El valor de F_{ST} fue de al menos 0.15 de cada marcador, que indica la capacidad para diferenciar totalmente a una población humana de otra. Se seleccionaron aquellos con los mayores valores de diferenciación entre su mismo grupo reportado, mostrando el mayor contraste en las frecuencias de distribución de sus alelos entre las poblaciones históricas parentales: Asiática, Europea y Africana.

Con base en estos criterios se seleccionaron 200 marcadores, a partir de los cuales se formaron paneles con un número mínimo de AIMs que permitieron obtener una adecuada separación de las poblaciones parentales. La densidad de los marcadores por etnia parental en cada panel fue con proporción 5:4:1 (Nativa Americana:Europea:Africana) conforme a lo reportado por [Johnson *et al.*, 2011](#).

La eficiencia de los paneles se determinó mediante análisis estadístico *in silico* evaluado con los métodos de Análisis Multidimensional a Escala (**MDS**, por sus siglas en inglés), Análisis de Componentes Factoriales y el Análisis Molecular de Varianza (**AMOVA**). Para ello se emplearon 2,504 genotipos descritos en la [Tabla A2](#), pertenecientes a la base de datos del proyecto 1,000 genomas (www.internationalgenome.org).

La verosimilitud de la separación de las poblaciones se visualizó a través del MDS mediante el programa IBM SPSS Statistic v.22 ([IBM Corp. 2013](#)) y mediante un análisis de componentes factoriales con el programa Genètix v.4.05.2 ([Belkhir, *et al.*, 2004](#)). Estos análisis permitieron observar las distancias genéticas entre las poblaciones del proyecto 1000 Genomas de acuerdo con las frecuencias genotípicas con respecto a los 28 AIMs.

2.4.2 Obtención y procesamiento de muestras

2.4.2.1 El material genético de las poblaciones Nativas fue obtenido a partir de muestras de saliva contenidas en tubos de 15 mL con solución Scope®. Los tubos se sellaron con Parafilm y se transportaron a temperatura ambiente al laboratorio 14 del departamento de Toxicología (CINVESTAV-Zacatenco) donde se procesaron y realizaron alícuotas (10 ng/µl) de las muestras. En el caso de los individuos mestizos, se obtuvieron muestras de sangre periférica en tubos con EDTA como anticoagulante del sistema BD Vacutainer (Becton Dickinson, NJ, USA), los cuales se mantuvieron a 4 °C hasta su procesamiento en el laboratorio de Biología Molecular Diagnóstica S.A. de C.V. Los comités de ética de *National Geographic* y de Biología Molecular Diagnóstica S.A. de C.V. validaron los consentimientos informados firmados por los individuos de las poblaciones Nativas y la población mestiza, respectivamente.

2.4.2.2 Aislamiento de material genético. El DNA fue aislado a partir de muestras de sangre periférica y saliva mediante la técnica de *Salting-out* con el *kit* comercial *DNA Blood Preparation* ([Jena Bioscience; Lobstedter, Alemania](#)) siguiendo las instrucciones del fabricante ([la metodología en extenso se describe en el Anexo](#)).

2.4.2.3 Cuantificación y determinación de la pureza del DNA. Las muestras de DNA se cuantificaron y se determinó su pureza mediante el espectrofotómetro NanoDrop 2000™ ([ThermoScientific, Massachusetts, Estados Unidos](#)) siguiendo el manual de uso del instrumento ([la metodología en extenso se describe en el Anexo](#)).

Con base en la propiedad de absorción de luz de las bases aromáticas nitrogenadas del DNA a una longitud de onda de 260nm de absorbancia máxima, es posible determinar la concentración en una muestra líquida por espectrofotometría empleando la ecuación de Beer-Lambert (www.ibt.unam.mx). El equipo NanoDrop 2000™ determina la concentración de DNA mediante una modificación de la ecuación de Beer-Lambert que se describe a continuación:

$$c=(A*\epsilon)/b$$

donde:

c: concentración de ácido nucleico en ng/uL

A= Absorbancia

ϵ = coeficiente de extinción

b= longitud de trayectoria en cm (0.05-0.1 mm)

La pureza de la muestra se determinó con el índice de pureza calculado con una segunda lectura del espectrofotómetro a 280nm, que corresponde a la longitud de absorción máxima de los aminoácidos aromáticos. La relación 260/280 con valor de ~1.8 se consideró como muestra de concentración de DNA adecuada para continuar con los análisis posteriores ([Manual de usuario NanoDrop 2000™, 2009](#)).

2.4.2.4 Integridad de la muestra. La integridad del DNA se evaluó mediante electroforesis en geles de agarosa al 0.8%, empleando un intercalante fluorescente ([MIDORI^{Green} Advance DNA Stain, Nippon Genetics, Japón](#)), el que permitió la visualización

del material genético mediante la excitación del intercalante por medio de un transiluminador ultravioleta a 312 nm ([Spectroline modelo TM-312, Spectro Estados Unidos](#)). Se consideró una muestra íntegra a aquella que emitió una señal fluorescente sin signos de degradación (fluorescencia en forma de barrido en el gel de agarosa).

Las muestras de las poblaciones de estudio que obtuvieron al menos 10 ng/μL de DNA, pureza entre 1.6 y 2.0 e integridad de DNA fueron empleadas en los análisis moleculares.

2.4.3 Análisis moleculares

2.4.3.1 Reacción en cadena de la polimerasa cuantitativa. Los análisis moleculares se llevaron a cabo mediante la técnica de la Reacción en cadena de la polimerasa (PCR) de tipo cuantitativo o qPCR. La qPCR es una técnica de amplificación, detección y cuantificación de DNA reportada por una molécula fluorescente, en un termociclador para PCR cuantitativa. Esta señal es directamente proporcional al número de moléculas de DNA amplificadas. Por lo tanto, es una herramienta que permite determinar en una muestra, el número de copias de genes, cuantificación de carga viral o su discriminación alélica. La determinación se realiza por medio de intercalantes fluorescentes de DNA o sondas específicas como las de tipo Taqman[®], las que fueron empleadas en el presente estudio ([Anexo Tablas A3 y A4](#)). Este tipo de sondas están formadas por una secuencia de nucleótidos específica que detecta la región de interés. Tienen una molécula reportera en el extremo 5' y un *quencher* (molécula que disminuye la intensidad de la fluorescencia) en el extremo 3' para ser detectadas. Una vez alineada con su secuencia complementaria en el DNA e iniciada la amplificación, estas dos moléculas son hidrolizadas por la acción de la *Taq* DNA polimerasa. La separación de las moléculas permite la detección de la fluorescencia por el termociclador y su visualización a través de un *software* durante todo el proceso ([Jia, 2012](#)).

2.4.3.2 Genotipificación de las variantes del *AS3MT*. Los alelos de cinco SNPs del gen *AS3MT* (rs3740393, rs11191439, rs10748835, rs11191438, rs3740400, rs7085104) relacionadas -de acuerdo con la literatura- con una mayor actividad metabólica del iAs fueron discriminados por medio de la qPCR empleando sondas

Taqman[®] (ThermoFisher Scientific, Massachusetts, Estados Unidos) (Tabla A4 Anexo). Se realizaron pruebas de corrida a diferentes concentraciones de la muestra (2, 5 y 10 ng/ μ l) y diferentes temperaturas de ciclado para obtener la mayor eficiencia de la reacción.

2.4.3.3 Determinación de AIMs en las poblaciones de estudio. La discriminación alélica para inferir la carga ancestral también se realizó por medio de qPCR y sondas Taqman[®] (ThermoFisher Scientific, Massachusetts, Estados Unidos) del panel previamente diseñado (Tabla A3 Anexo). En ambas discriminaciones alélicas se empleó el termociclador CFX96 Real Time System[™] (BioRad[™], California, Estados Unidos). En todas las reacciones realizadas se empleó un control negativo (solución de dilución de muestras) por cada marcador analizado.

2.4.4 Análisis estadístico

2.4.4.1 Diversidad genética en *AS3MT*. Las frecuencias alélicas y genotípicas, así como las diversidades genética y haplotípica de cada uno de los *loci* en los individuos analizados se determinaron con el programa Arlequin v.3.5 (Excoffier, *et al.*, 2005). El equilibrio con respecto a la ecuación de Hardy-Weinberg (HW) se determinó con Genètx v.4.05.2 (Belkhir, *et al.*, 2004); el nivel de significancia fue $p \leq 0.01$ después de la corrección por Bonferroni. El desequilibrio de ligamiento (LD) fue determinado con el programa FSTAT v.2.9.4 (Goudet, *et al.*, 2009) a tres niveles nominales (0.05, 0.01, 0.001). Con base en los resultados obtenidos se construyeron los haplotipos, cuyas frecuencias fueron determinadas por conteo directo. El análisis molecular de varianza (AMOVA) y las distancias genéticas se determinaron con el programa Arlequin v.3.5 (Excoffier, *et al.*, 2005). Los resultados de las distancias genéticas se mostraron gráficamente a través de un análisis de escala multidimensional (MDS) mediante el programa SPSS v11.0 (IBM Corp. 2013).

2.4.4.2 Inferencia de carga ancestral. Este análisis se llevó a cabo mediante un método Bayesiano con el programa Structure v.2.3.4 (Pritchard, Stephens & Donnelly; 2000), con el que se determinaron las proporciones genéticas. El programa STRUCTURE v.2.3.4 identifica las proporciones ancestrales en las poblaciones

de estudio a partir de los genotipos obtenidos con los 28 AIMs. Este programa realiza correlaciones genotípicas entre marcadores no relacionados para determinar la estructura poblacional de acuerdo con su *background* genético de los individuos de estudio y asignándolos en subpoblaciones (k) para disminuir la estructura genética. El programa asume:

- Que los individuos heredaron sus genes de un *pool* de k poblaciones “no estructuradas” (valor de k desconocido).
- Las subpoblaciones son “no estructuradas” y en cada subpoblación todos los *loci* se encuentran en equilibrio de acuerdo con Hardy-Weinberg.
- Las frecuencias alélicas de cada *locus* dentro de cada población son desconocidas.

El *background* genético de cada individuo es representado por el vector, $q = q_1, \dots, q_k$, donde q_k es la proporción del genoma de cada individuo que se originó en la subpoblación k . El programa utiliza el método de Monte Carlo basado en cadenas de Markov (MCMC) para estimar el número de subpoblaciones, las frecuencias alélicas en cada subpoblación y el valor de q para cada individuo. Los parámetros de corrida fueron: 150,000 *Length of burning period* y 700,000 repeticiones MCMC después del *burn-in*.

El algoritmo de este programa identifica patrones de variación en el *set* de datos ingresados y posiciona a los individuos en el grupo con el patrón más similar; se utiliza principalmente en poblaciones en *admixture*. Los parámetros de F_{ST} obtenidos de los resultados del panel de ancestría en las poblaciones de estudio y del proyecto 1000 Genomas se representaron en proporciones de carga ancestral (Africana, Europea y Nativo Americana) identificados de acuerdo al patrón generado en las poblaciones parentales del proyecto 1000 Genomas.

2.4.4.3 Exploración de la relación entre la carga ancestral y las variantes del *AS3MT*. El análisis de la relación de la carga ancestral con las variantes alélicas de la enzima *AS3MT* se realizó de forma manual. Las frecuencias genotípicas también fueron comparadas con la base de datos de 1,000 genomas y con otras poblaciones que habían reportado las mismas variantes del *AS3MT* en poblaciones Mexicanas (Drobná, *et al.*, 2013; 2016; Gómez-Rubio, *et al.*, 2012; Xu, *et al.*, 2016; www.internationalgenome.org).

3. Resultados

3.1 Diseño y evaluación *in silico* del panel de ancestría

Con el objetivo de relacionar la ancestría regional con las variantes polimórficas del *AS3MT*, se diseñó un panel de AIMs. La selección se realizó a partir de los marcadores reportados en diferentes bases de datos y literatura (Tabla A1). La capacidad para separar a las poblaciones parentales mediante los diferentes marcadores seleccionados se analizó mediante los genotipos de acceso libre del proyecto 1000 Genomas (www.internationalgenome.org), los que se compararon con paneles de 34, 36 y 48 AIMs reportado previamente en la literatura (Kosoy, *et al.*, 2009; <https://alfred.med.yale.edu/>; <http://spsmart.cesga.es/snpforid.php>).

La evaluación estadística incluyó:

- La determinación de la distancia genética (F_{ST} ; matrices mostradas en el Anexo).
- El análisis molecular de varianza.
- Análisis de escala multidimensional y de componentes factoriales.

Los resultados permitieron seleccionar 28 marcadores, cuya eficiencia fue evaluada *in silico*. Las F_{ST} entre las poblaciones parentales y Latinas (Tabla 4) con respecto a la población Mexicana residente en Los Ángeles (MXL) permitieron observar la contribución de las tres poblaciones determinadas como parentales (Africana, AFR; Este de Asia, EAS; y Europea, EUR) en MXL. Asimismo, los 28 AIMs permitieron comparar a la población MXL con otras poblaciones con historias demográficas similares (Tabla A2, Anexo). De manera general, todas las poblaciones, tanto parentales como Latinas (Colombianos de Medellín - CLM, MXL, Peruanos de Lima-PEL y Puertorriqueños-PUR), mostraron distancias genéticas con diferencias estadísticas significativas ($p \leq 0.0001$). La mayor diferenciación genética se encontró entre las poblaciones AFR, EAS y EUR cuyos valores de F_{ST} fueron mayores a 0.5. La distancia entre estas poblaciones y MXL fue variable; el valor mayor de F_{ST} fue con respecto a AFR (0.4069), mientras que, con respecto a EAS (relacionado con la carga ancestral Nativo Americana) y EUR los valores fueron muy similares (0.3244 y 0.3314, respectivamente). En cuanto a las F_{ST} entre poblaciones Latinas y MXL, los CLM fueron los más cercanos (0.0585) a MXL, seguidos de PEL (0.0843) y PUR (0.1086).

Empleando los valores de la distancia genética se realizaron los análisis de escala dimensional que explican la variación genética con respecto de la media en proporciones

(%). El análisis de componentes factoriales formó tres grupos cohesivos pertenecientes a los individuos de las poblaciones parentales (Figura A1). Sin embargo, mostró una distribución heterogénea entre los individuos de las poblaciones Latinas, los que se extendieron entre las poblaciones AFR, EAS, EUR (Figura A1). Al comparar con los paneles de 34, 36 y 48 AIMs, los 28 marcadores seleccionados mostraron agrupaciones más consistentes entre las poblaciones parentales (Figura A3). La primera dimensión del MDS separó a las poblaciones EAS, MXL y PEL del resto de las poblaciones, mientras que la segunda dimensión separó a AFR y SAS del resto de las poblaciones. La verosimilitud del MDS ($S\text{-stress} = 0.01726$) y los resultados del AMOVA sustentaron la capacidad de los 28 AIMs seleccionados para determinar la ancestría regional (resultados no mostrados).

3.2 Determinación de la estructura genética en las poblaciones de estudio

Empleando los datos de las frecuencias obtenidas de los 28 AIMs, se calcularon los valores de F_{ST} entre las poblaciones de estudio (Mazahua, Me'Phaa, Mestiza y Tsotsil). De manera general, las poblaciones mestiza y Nativo Americanas mostraron valores de F_{ST} menores a 0.15 con valores estadísticamente significativos entre ellas (Tabla 4); la mayor distancia encontrada fue entre los mestizos y la población Me'Phaa ($F_{ST} = 0.2109$, $p \leq 0.0001$). Con base en las distancias genéticas, las poblaciones de estudio se distribuyeron en el MDS dentro de un mismo cuadrante (Figura 4). Sin embargo, se observó mayor cercanía de la población Tsotsil con la mestiza con respecto a la dimensión uno, seguida de la población Mazahua y con mayor distancia la población Me'Phaa. Entre las poblaciones Nativas, los Mazahuas y Tsotsiles fueron más cercanos entre ellos mientras que la población Me'Phaa tuvo mayor distancia respecto a las otras dos. En función de la dimensión dos y de manera marginal, los individuos Mazahuas y mestizos fueron más cercanos, así como los Me'Phaas de los Tsotsiles.

Tabla 4. Cálculo de las distancias genéticas (F_{ST}) y significancia estadística (p) de las poblaciones de estudio y de las obtenidas del proyecto 1,000 Genomas.

Pob.	MAZ	MEE	MES	TST	AFR	CLM	EAS	EUR	MXL	PEL	PUR	SAS
MAZ	-	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
MEE	0.1397	-	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
MES	0.1494	0.2109	-	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
TST	0.098	0.1069	0.0812	-	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
AFR	0.5694	0.6623	0.372	0.5624	-	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
CLM	0.3173	0.4095	0.0766	0.2464	0.3737	-	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
EAS	0.3123	0.4218	0.2824	0.3542	0.5312	0.4425	-	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
EUR	0.5992	0.6477	0.3208	0.5507	0.5156	0.1337	0.6248	-	<0.0001	<0.0001	<0.0001	<0.0001
MXL	0.1554	0.2709	0.016	0.1126	0.4069	0.0585	0.3244	0.3314	-	<0.0001	<0.0001	<0.0001
PEL	0.1134	0.2324	0.0891	0.0763	0.514	0.2342	0.2802	0.5297	0.0843	-	<0.0001	<0.0001
PUR	0.3776	0.4744	0.1261	0.319	0.3555	0.0119	0.4939	0.1045	0.1086	0.3021	-	<0.0001
SAS	0.4023	0.4964	0.1816	0.3749	0.2647	0.1049	0.4187	0.2672	0.1669	0.331	0.1058	-

Nota: AFR, África; CLM, Colombianos de Medellín; EAS, Este de Asia; EUR, Europeos; MAZ, Mazahuas; MEE, Me'Phaa; MES, Mestizos; MXL, Mexicanos en L.A; PEL, Peruanos de Lima; PUR, Puertorriqueños; SAS, Sur de Asia; TST, Tsotsiles.

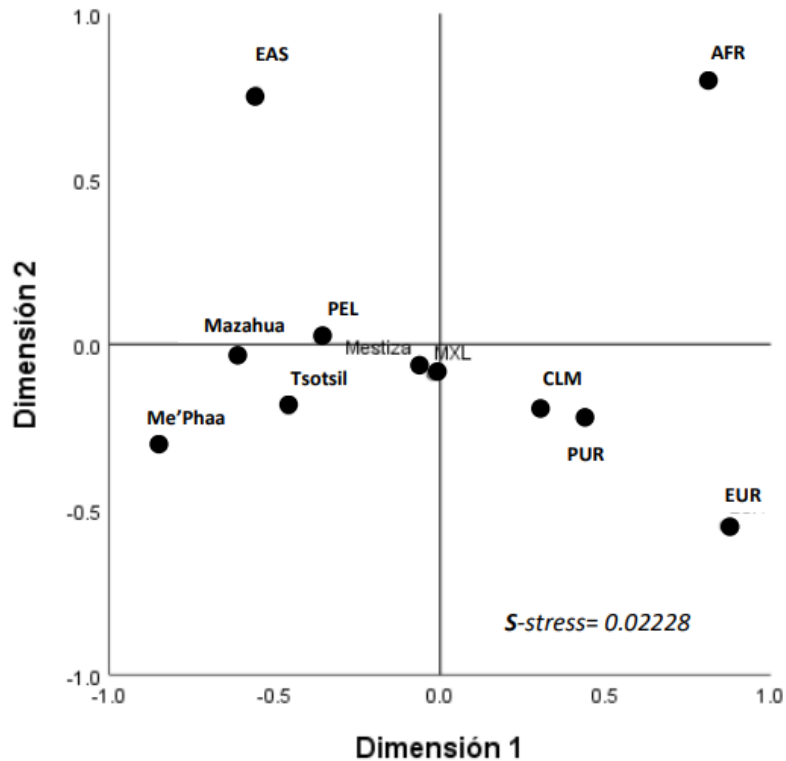


Figura 4. Análisis multidimensional a escala mediante 28 marcadores informativos de ancestría en las poblaciones de estudio y en los datos obtenidos del proyecto 1000 Genomas. AFR, África; CLM, Colombianos de Medellín; EAS, Este de Asia; EUR, Europeos; MXL, Mexicanos en L.A; PEL, Peruanos de Lima; PUR, Puertorriqueños.

Las proporciones de carga ancestral (individual y poblacional) se calcularon mediante el programa STRUCTURE v.2.3.4 (Pritchard, Stephens & Donnelly; 2000) empleando como referencia los patrones genéticos de las poblaciones consideradas como parentales obtenidas de la base de datos de 1000 Genomas (www.internationalgenome.org). Los resultados obtenidos se analizaron con el programa Structure Harvester v.0.6.9.4 (Earl & vonHoldt; 2012) el cual proporcionó los valores de probabilidad en múltiples valores de k (subpoblaciones) con el fin de detectar el número de grupos genéticos que mejor se ajustaron a los datos genotípicos introducidos. Las poblaciones de estudio se agruparon en tres subpoblaciones ($k=3$, $p < 0.0001$), con base en su afinidad con las poblaciones parentales: AFR, EAS y EUR se identificaron en colores verde, rojo y azul, respectivamente (Figura 5). Las poblaciones de estudio mostraron diferentes proporciones con respecto a estos tres

colores. Las proporciones ancestrales se determinaron en una escala de cero a uno en el eje de las ordenadas en los 309 individuos (cada individuo está representado por una barra vertical). Los individuos mestizos mostraron un promedio de 53.3% de ancestría NA, 41.6% de la Europea y 5.2% de la Africana. Las poblaciones Nativas mostraron, de forma general, una prominente ancestría NA (87.5%), mientras que las ancestrías Europea (8.1%) y Africana (4.3%) se presentaron en menor proporción; las proporciones ancestrales fueron variables dependiendo de la población. Con respecto a la carga ancestral NA, la población Me'Phaa fue la que mostró una mayor proporción (96.7%), seguida de la población Mazahua (86.9%) y la población Tsotsil (78.9%). Con respecto al componente Europeo, la población Tsotsil fue la que presentó una mayor proporción (16.4%), seguida de la población Mazahua (6.0%); la población Me'Phaa mostró la menor proporción (2.0%). El componente Africano se presentó en proporción menor: 7.1% en la población Mazahua, 4.6% en los Tsotsiles y 1.3% en la población Me'Phaa.

3.2.1 Comparación con otras poblaciones

Las poblaciones de estudio se compararon con poblaciones del proyecto 1000 Genomas (AFR, CLM, EAS, EUR, MXL, PEL, PUR). Las distancias genéticas obtenidas y el MDS se muestran en la [Tabla 4](#) y [Figura 4](#). En general, las cuatro poblaciones de estudio fueron próximas a PEL y MXL; mientras que mostraron mayor distancia con respecto a las poblaciones AFR, CLM, EAS, EUR y PUR. La primera dimensión del MDS separó a las poblaciones de estudio junto con MXL, PEL y EAS de las poblaciones AFR, CLM, EUR y PUR. La segunda dimensión separó a la población Mazahua, Me'Phaa, mestiza, Tsotsil, junto con MXL, CLM y PUR de las poblaciones AFR, EAS y PEL. En el caso de las poblaciones Latinas, los individuos MXL se mostraron en el mismo cuadrante que las poblaciones de estudio. Sin embargo, PEL mostró una menor distancia con las poblaciones Nativas Mexicanas de este estudio y se posicionó en el mismo cuadrante que EAS. Por otra parte, CLM y PUR compartieron cuadrante con EUR y mostraron la mayor distancia con respecto a las poblaciones de estudio, así como con PEL.

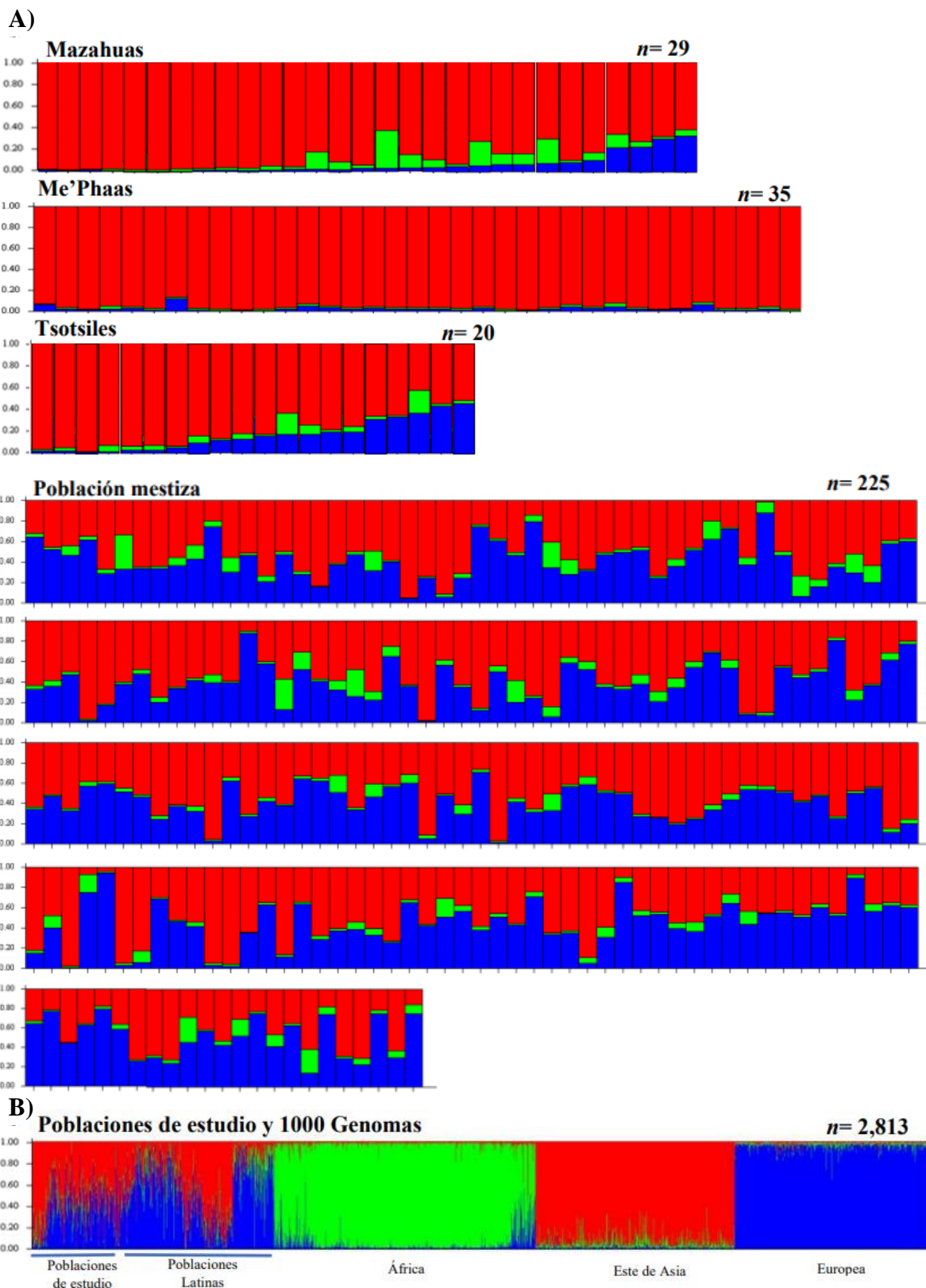


Figura 5. Proporciones ancestrales (A) de las poblaciones de estudio y (B) de las obtenidas del proyecto 1000 Genomas mediante el programa STRUCTURE v.2.3.4. n = muestra poblacional.

3.3 Variabilidad del gen de la *AS3MT* en las poblaciones de estudio y su comparación con las poblaciones consideradas como parentales

Una vez determinada la carga ancestral regional de las poblaciones de estudio, se analizó la distribución y frecuencia con respecto a las cinco variantes del *AS3MT* (Tabla 5). Las poblaciones Mazahua, Mestiza y Tsotsil mostraron distribuciones y frecuencia similares. Por su parte, la población Me'Phaa mostró diferencias importantes en cuanto a la distribución de estos alelos, principalmente al ser comparados con la población Mestiza. Las frecuencias genotípicas siguieron el mismo comportamiento que las alélicas, mostrando un exceso de individuos homocigotos en los alelos más frecuentes (Tabla 6). El alelo T del rs11191439 mostró frecuencias altas en todas las poblaciones por lo que fue considerado como poco informativo.

Tabla 5. Frecuencias alélicas de las cinco variantes del *AS3MT* en las poblaciones de estudio y en las consideradas como parentales.

Población		Mazahua	Me'Phaa	Mestizos	Tsotsil	África	Este de Asia	Europa
Locus	Alelo	29	35	225	20	661	504	503
3740393	C	0.24 (7)	0.47 (16) ^a	0.22 (50)	0.15 (3)	0.17 (112)	0.29 (146)	0.15 (75)
	G	0.76 (22)	0.53 (19)	0.78 (175)	0.85 (17)	0.83 (549)	0.71 (358)	0.85 (428)
3740400	T	0.57 (17)	0.44 (15)	0.59 (133)	0.7 (14)	0.67 (443)	0.41(207)	0.63 (317)
	G	0.43 (12)	0.56 (20) ^a	0.41 (92)	0.3 (6)	0.33 (218)	0.59 (297)	0.37 (186)
7085104	G	0.43 (12)	0.59 (21) ^a	0.39 (88)	0.2 (4) ^a	0.31 (205)	0.55 (277)	0.35 (176)
	A	0.57 (17)	0.41(14)	0.61 (137)	0.8 (16)	0.69 (456)	0.45 (227)	0.65 (327)
10748835	G	0.55 (16)	0.4 (14)	0.57 (128)	0.7 (14)	0.67 (443)	0.42 (212)	0.58 (292)
	A	0.45 (13)	0.6 (21) ^a	0.43 (97)	0.3 (6)	0.33 (218)	0.58 (292)	0.42 (211)
11191439	T	0.83 (24)	1.0 (35) ^a	0.91 (205)	0.95 (19)	0.91 (602)	0.98 (494)	0.89 (448)
	C	0.17(5)	0	0.09 (20)	0.05 (1)	0.09 (59)	0.02 (10)	0.11 (5)

^a $p \leq 0.03$ comparado con la población Mestiza Mexicana.

En negritas los alelos relacionados con un aumento en el metabolismo del *AS3MT*.

Tabla 6. Frecuencias genotípicas de las cinco variantes del *AS3MT* en las poblaciones de estudio.

SNP – <i>AS3MT</i>	Genotipo	Mazahua (29)	Me'Phaa (35)	Mestizos (225)	Tsotsil (20)
rs3740393	CC	0.14 (4)	0.29 (10) ^a	0.05 (1)	0.07 (15)
	CG	0.21 (6)	0.37 (13)	0.20 (4)	0.31 (70)
	GG	0.66 (19)	0.34 (12) ^a	0.75 (15)	0.62 (140)
rs3740400	TT	0.34 (10)	0.20 (7)	0.50 (10)	0.35 (79)
	TG	0.45 (13)	0.49 (17)	0.40 (8)	0.49 (110)
	GG	0.21 (6)	0.31 (11) ^a	0.10 (2)	0.16 (36)
rs7085104	GG	0.21 (6)	0.34 (12) ^a	0.05 (1)	0.16 (36)
	GA	0.45 (13)	0.49 (17)	0.30 (6)	0.47 (106)
	AA	0.34 (10)	0.17 (6) ^a	0.65 (13)	0.37 (83)
rs10748835	GG	0.34 (10)	0.14 (5)	0.50 (10)	0.31 (70)
	GA	0.41 (12)	0.51(18)	0.40 (8)	0.51 (115)
	AA	0.24 (7)	0.34 (12) ^a	0.10 (2)	0.18 (40)
rs11191439	TT	0.66 (19) ^a	1.00 (35) ^a	0.90 (18)	0.83 (187)
	CT	0.34 (10) ^a	-	0.10 (2)	0.16 (36)
	CC	-	-	-	0.01 (2)

^a $p \leq 0.039$ comparado con la población Mestiza Mexicana.

En negritas los alelos relacionados con un aumento en el metabolismo del *AS3MT*.

En la [Tabla 7](#), se describen las frecuencias genotípicas de dos o tres variantes (rs3740393, rs10748835 y rs11191439) determinadas en otras poblaciones Mexicanas obtenidas de estudios previos ([Drobná, et al., 2013; 2016; Gómez-Rubio, et al., 2012; Xu, et al., 2016](#)). No se encontró diferencia significativa en las frecuencias reportadas de poblaciones del centro-norte de México con respecto a las frecuencias de las poblaciones de estudio mestiza, Mazahua y Tsotsil. Sin embargo, se conservó la diferencia estadística con respecto a la población Me'Phaa en los genotipos homocigotos para el alelo.

El análisis con respecto a la ecuación de HW se describe en la [Tabla 8](#). Las cinco poblaciones de estudio mostraron estar en equilibrio, aunque se identificó un ligero exceso de individuos homocigotos, siendo la población Mazahua la que mostró el mayor índice de fijación ($F_{IS} = 0.1418$).

Tabla 7. Frecuencias genotípicas de las variantes de AS3MT en Me'Phaas, mestizos y otras poblaciones de México.

SNP – AS3MT	Genotipo	Me'Phaa (35)	Mestizos (225)	Zimapán, Hgo. (195) Drobná, <i>et al.</i> (2016)	Zimapán y Región Lagunera (255) Drobná, <i>et al.</i> (2013)	Chihuahua (500) Xu, <i>et al.</i> (2016)	Sonora (746) Gómez-Rubio <i>et al.</i> (2012)
rs3740393	CC	0.29 (10)*	0.07 (16)	0.09 (18)	0.05 (11)	0.05 (26)	0.05 (39)
	CG	0.37 (13)	0.31 (70)	0.32 (63)	0.41 (95)	0.40 (200)	0.32 (242)
	GG	0.34 (12)	0.62 (140)	0.59 (116)	0.55 (128)	0.55 (280)	0.62 (465)
rs3740400	TT	0.20 (7)	0.35 (79)	0.37 (73)			
	TG	0.49 (17)	0.49 (110)	0.47 (92)	ND	ND	ND
	GG	0.31 (11)*	0.16 (36)	0.16 (31)			
rs7085104	GG	0.34 (12)*	0.16 (36)	0.15 (30)			
	GA	0.49 (17)	0.47 (106)	0.46 (90)	ND	ND	ND
	AA	0.17 (6)	0.37 (83)	0.39 (76)			
rs10748835	GG	0.14 (5)	0.31 (70)	0.32 (62)	0.30 (76)	0.30 (151)	
	GA	0.51(18)	0.51 (115)	0.53 (104)	0.53 (134)	0.50 (251)	ND
	AA	0.34 (12)*	0.18 (40)	0.15 (30)	0.18 (45)	0.20 (98)	
rs11191439	TT	1.00 (35)*	0.83 (187)	0.83 (165)	0.87 (223)	0.82 (412)	0.89 (667)
	CT	-	0.16 (36)	0.14 (29)	0.11 (29)	0.17 (86)	0.10 (78)
	CC	-	0.01 (2)	0.03 (1)	0.02 (3)	0.01 (3)	0.01 (1)

ND: No se determinó.

* $p \leq 0.05$ comparado con la población mestiza de este estudio y las poblaciones de otros estudios.

Tabla 8. Determinación del Equilibrio con respecto a la ecuación de Hardy-Weinberg en los *loci* de *AS3MT* en las poblaciones de estudio.

Población (n) /Alelo	Mazahuas (29)	Me'Phaa (35)	Mestizos (225)	Tsotsiles (20)
rs3740393 $F_{IS}(p)$	0.449 (-0.03)	0.268 (-0.1)	0.102(-0.09)	0.24 (-0.35)
rs3740400 $F_{IS}(p)$	0.103 (-0.43)	0.03 (-0.56)	0.145 (-0.004)	0.073 (-0.57)
rs7085104 $F_{IS}(p)$	0.103 (-0.43)	0.014 (-0.61)	0.013 (-0.48)	0.088 (-0.58)
rs10748835 $F_{IS}(p)$	0.18 (-0.27)	-0.057 (0.75)	-0.038 (-0.76)	0.073 (-0.57)
rs11191439 $F_{IS}(p)$	-0.192 (-1)	Alelo monomórfico	0.014 (-0.537)	-0.027 (-1)
He	0.4329	0.4643	0.41	0.3097
Ho	0.3724	0.4964	0.3884	0.28
F_{IS}	0.1418	0.0655	0.0526	0.0983
EHW (p)	0.0694	0.2741	0.0406	0.2681

Nota: He, Heterocigosidad esperada; Ho, Heterocigosidad observada; F_{IS} , Índice de fijación; p , significancia estadística.

Con respecto a las frecuencias haplotípicas (Tabla 9), las más prominentes se distribuyeron entre los haplotipos TTAGG (cuyas mayores proporciones se presentaron entre los Tsotsiles, Mazahuas y Mestizos, respectivamente) y TGGAC, cuya mayor frecuencia se presentó en la población Me'Phaa (0.43). Vale la pena destacar que el haplotipo TGGAC, conformado por los cinco alelos descritos en la literatura como eficientes (Agusa, *et al.*, 2011), también mostró una frecuencia considerable en las poblaciones del Este de Asia (0.30), duplicando las proporciones de las poblaciones Tsotsil y Mestiza. Al analizar las frecuencias del haplotipo TGGAC con respecto a la carga ancestral NA, éste mostró un incremento de 1.56 veces entre los individuos Mestizos. La población Mestiza fue la que mostró una diversidad haplotípica mayor, mostrando 25 haplotipos y una diversidad haplotípica de 0.41.

Tabla 9. Frecuencias haplotípicas obtenidas mediante los cinco loci de AS3MT en las poblaciones de estudio y en las poblaciones del Este de Asia y Europa del proyecto 1000 Genomas.

Población (N_h) /Haplotipo	Mazahua (58)	Me'Phaa (70)	Mestizos (450)	Tsotsil (40)	Este de Asia (1,008)	Europa (1,006)
<u>TGGAC</u>	0.21 (12)	0.43 (30) ^{b,+,2}	0.16 (72) ^C	0.13 (5) ^C	0.29 (292)	0.10 (101)
<u>TGGGC</u>	-	0.02 (2)	-	-	-	-
<u>TGGAG</u>	0.05 (3)	0.06 (4)	0.11 (50)	0.03 (1)	0.22 (222)	0.13 (131)
<u>TTGAC</u>	-	0.03 (2)	-	-	-	-
<u>TGAAC</u>	-	-	-	0.03 (1)	-	-
<u>TGAAG</u>	-	-	-	0.03 (1)	0.05 (50)	0.02 (20)
<u>CGGAG</u>	0.17 (10)	-	0.07 (32)	0.05 (2)	0.02 (20)	0.11 (111)
<u>TTAAG</u>	0.02 (1)	0.08 (6)	-	0.05 (2)	-	-
<u>TTAGG</u>	0.52 (30)	0.33 (22) ^{+,1,2}	0.49 (220)	0.65 (26)	0.40 (404)	0.58 (583)
Otros	0.04 (2)	0.05 (4)	0.17 (76)	0.05 (2)	0.02 (20)	0.06 (60)
Diversidad Haplotípica (h)	6 (-0.433)	7 (-0.4)	25 (-0.41)	8 (-0.31)	10 (-0.38)	8 (-0.37)

^b $p < 0.01$ comparado con la población mestiza Mexicana, ¹Mazahua y ²Tsotsil.

^c $p < 0.01$ comparado con las poblaciones del Este de Asia (Proyecto 1000 Genomas).

⁺ $p < 0.01$ comparado con las poblaciones Europeas (Proyecto 1000 Genomas).

N_h : número de haplotipos.

h : Índice de diversidad Haplotípica.

Nota: Los alelos reportados en la literatura como representativos del haplotipo «eficiente» se encuentran subrayados y en negritas.

3.3.1 Comparación entre la carga ancestral regional y las variantes alélicas de AS3MT.

A partir de los datos de frecuencia obtenidos del análisis de los cinco loci del gen AS3MT se realizó un MDS. Las poblaciones MXL y PEL (del proyecto 1000 genomas) fueron incluidas dado que en los análisis con los 28 AIMs mostraron una relación con nuestras poblaciones de estudio (Figura 6). Las poblaciones con mayor carga NA (Me'Phaa y PEL) se agruparon con la población EAS ($p \geq 0.081$). Mientras que las poblaciones Mazahua, Mestiza, Tsotsil y MXL estuvieron más cercanas a la población EUR. No obstante, la población Mestiza mostró diferencia significativa con las poblaciones EAS, EUR, Me'Phaas y PEL ($p \leq 0.009$). En el caso de las poblaciones Nativas, la población Me'Phaa

fue significativamente diferente de los Mazahuas, Mestizos, Tsotsiles, EUR y MXL ($p \leq 0.001$).

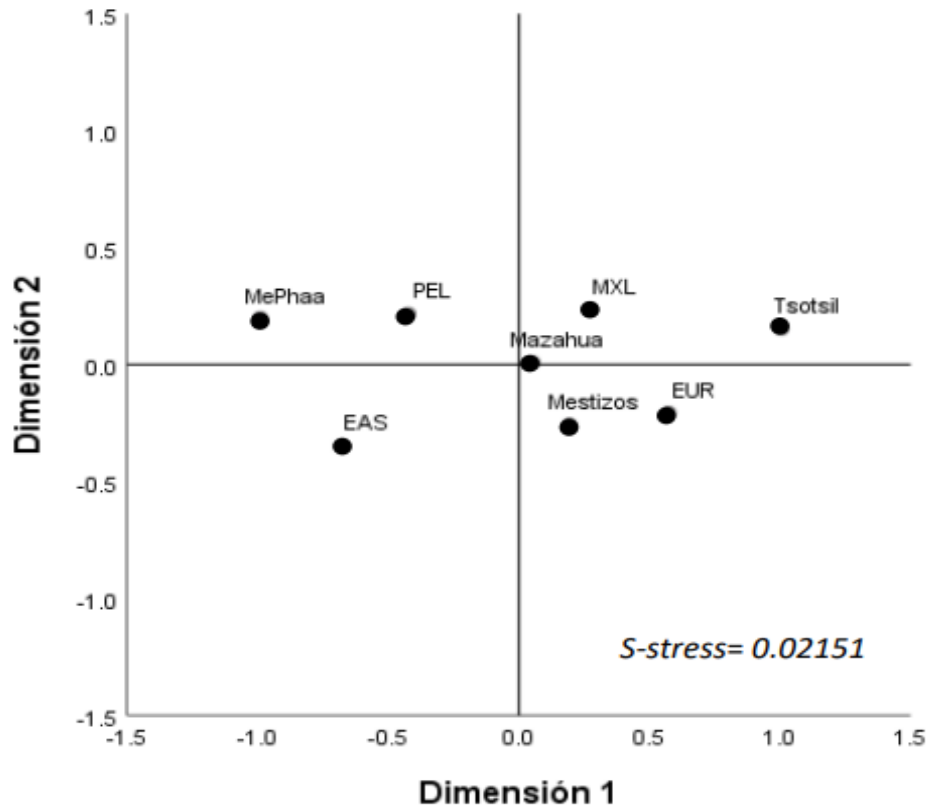


Figura 6. Análisis Multidimensional a Escala de las variantes eficientes de *AS3MT* en las poblaciones de estudio y en los datos obtenidos del proyecto 1000 Genomas. Los círculos indican dos grupos de poblaciones que mostraron diferencia significativa entre ellos.

Nota: EAS, Asiáticos del Este; EUR, Europeos; MXL, Mexicanos de L.A; PEL, Peruanos de Lima.

Finalmente, la distribución de haplotipos mostró similitud entre las poblaciones Me'Phaa, Ho Chi Minh (Vietnam), China, Bengali (Bangladesh) y PEL. Mientras que las frecuencias haplotípicas de las poblaciones Mazahua, Mestiza y Tsotsil se asemejan a las frecuencias observadas en la ancestría Europea (Figura 7).

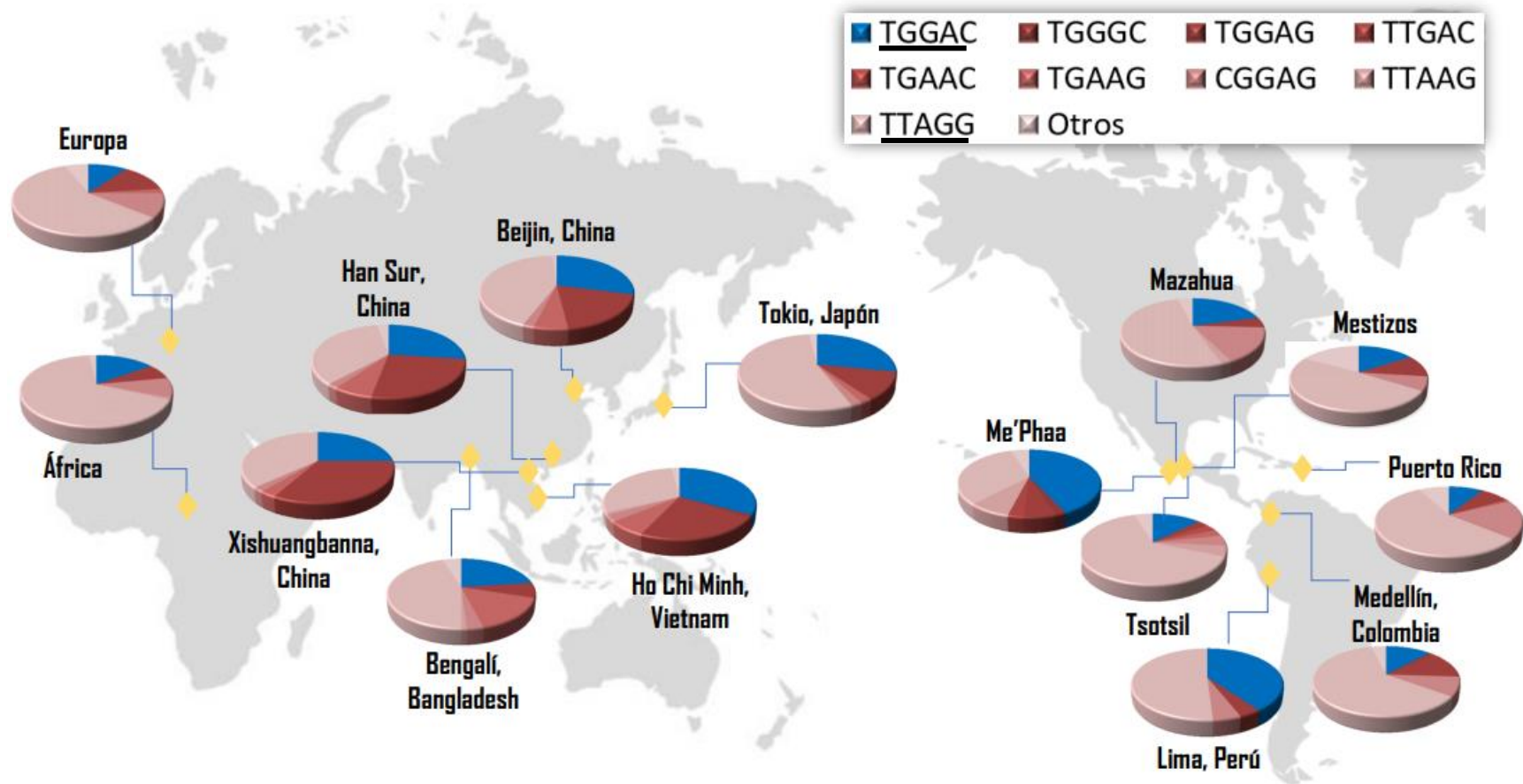


Figura 7. Mapa de frecuencias haplotípicas de AS3MT (rs3740393, rs40400, rs7085104, rs10748835, rs11191439) y en algunas poblaciones obtenidas del proyecto de 1000 Genomas.
Nota: Los haplotipos más frecuentes se encuentran subrayados.

4. Discusión

A pesar de que se conoce el metabolismo del iAs, se han observado variaciones en la susceptibilidad a efectos tóxicos relacionados con su capacidad de metilación en diferentes regiones en el mundo (Fujihara, *et al.*, 2009; 2010), lo que sugiere una posible variabilidad étnica en cuanto al metabolismo del iAs y las variantes alélicas en el gen *AS3MT* las cuales podrían tener un papel crítico (Agusa, *et al.*, 2011) en la susceptibilidad a los efectos de la exposición a iAs. Si bien, esto se ha observado en las poblaciones de Argentina y Chile principalmente, son escasos los estudios en poblaciones Latinas (incluido México) que relacionen la carga ancestral y las variantes eficientes de la *AS3MT*. (Apata, *et al.*, 2017; De la Loma, *et al.*, 2019; Eichstaedt, *et al.*, 2015; Schlebusch, *et al.*, 2013; 2015). Sin embargo, los hallazgos genéticos en la región Andina incluso han sugerido un proceso de selección positiva junto con la exposición crónica a iAs, lo cual podría constituir uno de los pocos ejemplos de adaptación biológica a metaloides (Apata, *et al.*, 2017).

México posee más de 60 grupos étnicos, pertenecientes a diversas familias lingüísticas que conectan el norte y el sur de América (Belbin, *et al.*, 2018, Bryc, *et al.*, 2010; Johnson, *et al.*,). Estudios previos han mostrado la contribución de los polimorfismos de la *AS3MT* a la predisposición de diferentes enfermedades y a la susceptibilidad a la toxicidad mediada por el metabolismo de iAs (D'Ambrosio, *et al.*, 2019; de la Rosa, *et al.*, 2017; Engström, *et al.*, 2014; Pierce, *et al.*, 2012; Song, *et al.*, 2020). No obstante, la caracterización de los patrones genéticos asociados al metabolismo del iAs, en poblaciones con diversa carga ancestral regional, ha sido desatendida. El presente trabajo buscó relacionar la ancestría regional obtenida a partir de 28 AIMS con cinco *loci* de la arsénico-3-metiltransferasa en poblaciones mestiza y Nativas Americanas contemporáneas.

4.1 Carga ancestral regional de las poblaciones mestizas y Nativas Americanas.

4.1.1 Población Mestiza

Los resultados obtenidos mediante los AIMs en la población Mestiza muestran 53.3% de componente ancestral primordialmente Nativo Americano, 41.6% Europeo y 5.2% de ancestría Africana. La proporción de la carga ancestral Africana de nuestro estudio corresponde a la proporción promedio descrita por otros grupos de investigación (Huerta-Chagoya, *et al.*, 2019; Kosoy, *et al.*, 2009; Martínez-Marignac, *et al.*, 2007; Silva Solezzi, *et al.*, 2009). Sin embargo, la ancestría Africana pudo haberse subestimado en este estudio debido a que esta proporción ancestral se concentra en la región sur costera del país (INEGI, 2015).

Las ancestrías NA y Europea reflejan los eventos demográficos que han enriquecido la estructura poblacional del Mestizo mexicano. Entre ellos, las migraciones Europeas (durante y después de la conquista) y más recientemente durante la Segunda Guerra Mundial y la Guerra Civil Española (Mitad del siglo XX y durante 1936-1960) (Gleizer, 2016; Hoyos-Puente, 2011); estas migraciones se concentraron, principalmente, en los estados con mayor interés económico. Por lo que se refiere a las poblaciones Nativas, éstas han atravesado por diversos cuellos de botella asociados a las epidemias, esclavitud, invasiones y hambrunas; los cuales actúan en función del número de generaciones partícipes de estos procesos (O'Fallon & Fehren-Schmitz, 2011; Salzano & Sans, 2014). Consecuentemente, la población Mestiza mexicana contemporánea presenta una intrincada estructura poblacional, la que también ha sido establecida por nuestro grupo de estudio a nivel autosómico y uniparental, principalmente en el Valle Central de México (Noris, *et al.*, 2012; Santana, *et al.*, 2014), así como por otros grupos de investigación que han retratado la carga ancestral regional de la Ciudad de México (56% NA; 40% Europea, 4% Africana) Juárez-Cedillo, *et al.*, 2008) y de

población en general empleando diferente número de marcadores (desde 32 hasta 1,814) (51-64% NA; 30-45% Europea; 2-5% Africana) (Chowdhry, *et al.*, 2006; Galanter, *et al.*, 2012, Huerta-Chagoya, *et al.*, 2019; Kosoy, *et al.*, 2009; Martinez-Marignac, *et al.*, 2007; Silva Solezzi, *et al.*, 2009). Dichos estudios, han descrito el claro *admixture* de la población Mexicana contemporánea y son evidencia de la importancia de una selección efectiva de AIMs para inferir carga ancestral regional como los aquí empleados. Nuestros resultados con respecto a la población Mestiza, no solo se asemejan a los reportados con AIMs, sino que incluso fueron consistentes a lo reportado con marcadores de la NRY, que es una de las herramientas más utilizadas en arquitectura genética (Cerde-Flores, *et al.*, 2002; Luna-Vázquez, *et al.*, 2005; Martínez-Cortés, *et al.*, 2012; Martínez-Fierro, *et al.*, 2009; Rangel-Villalobos, *et al.*, 2008; Rubí-Castellanos, *et al.*, 2009b; Salazar-Flores, *et al.*, 2010; Santana, *et al.*, 2014; Wang, *et al.*, 2008).

4.1.2 Poblaciones Nativas Americanas

Con respecto a las poblaciones NA, nuestros hallazgos también fueron consistentes, confirmando un proceso de mestizaje heterogéneo durante y después de la conquista (Gómez, *et al.*, 2020, en revisión; Gonzáles-Sobrino, *et al.*, 2016). Llama la atención las proporciones tan altas de ancestría NA en la población Me'Phaa (96.7%), lo que ha sido consistente con un estudio previo empleando 104 AIMs, en donde se reportó 79.9% de este componente ancestral (Cahua-Pablo, *et al.*, 2016). Cabe mencionar que nuestros resultados indican que esta población se encuentra en equilibrio con respecto a la ecuación de Hardy-Weinberg (con respecto a las variantes del gen *AS3MT*). Sin embargo, se observó una frecuencia importante de individuos homocigotos en la población y también observada por nuestro grupo de estudio

con marcadores multialélicos ($F_{IS} = 0.182$, $p \leq 0.0001$) (Camacho-Mejorado, *et al.*, 2020). La disminución de heterocigosis observada en los Me'Phaas puede deberse a su aislamiento geográfico y limitado acceso por cuestiones de narcotráfico; además de ser considerada una población libre de mestizaje (Bonilla, *et al.*, 2005; Camacho-Mejorado, *et al.*, 2015, 2020; Hernández-Tobías, 2016; López-Parra, *et al.*, 2016; <http://atlas.cdi.gob.mx/>). Consecuentemente, el flujo génico ha sido limitado lo cual contribuye a un comportamiento endogámico en la población por lo que los resultados en ella deben ser interpretados con cautela (Bonilla, *et al.*, 2005; Camacho-Mejorado, *et al.*, 2015, 2020; Hernández-Tobías, 2016; Lemes, *et al.*, 2018; López-Parra, *et al.*, 2016; <http://atlas.cdi.gob.mx/>). Los marcadores de *AS3MT* determinados en los individuos Me'Phaa no permiten determinar otras características demográficas poblacionales, por lo que se sugiere profundizar su estructura poblacional con un mayor número de AIMs u otro tipo de marcadores como los segmentos de DNA compartidos (idénticos por descendencia). (Kardos, Luikart & Allendorf, 2015).

Por otro lado, en una población aislada como la Me'Phaa la deriva génica también puede explicar la pérdida de heterocigosis, en donde la homocigosis aumenta con el tiempo (Bittles, 2008; Fontdevila, 2000; Willoughby, *et al.*, 2014). Esta fuerza evolutiva, afecta directamente la diversidad alélica en una población pequeña debido a que poseen un menor número de alelos efectivos (alelos que se presentan en futuras generaciones) y que, por azar puedan fijarse o perderse en la población (frecuencia 1 o 0) (Willoughby, *et al.*, 2014). Sin embargo, la deriva génica no deja de ser un emparejamiento al azar, consecuentemente, el ajuste de las frecuencias esperadas con respecto a Hardy-Weinberg es de una población en equilibrio (Fontdevila, 2000); cómo fue determinado en los individuos de la Sierra Alta de Guerrero.

En el caso de los individuos Mazahuas y Tsotsiles quienes pertenecen a las familias

lingüística Oto-Mangue y Maya, respectivamente, mostraron tintes de mestizaje a pesar de tener una muestra poblacional pequeña. La población Mazahua localizada para este caso en particular de Valle de Bravo, pertenece a la familia lingüística más diversa (Oto-Mangue) (<http://alin.inali.gob.mx/xmlui/>). Esta población mantiene una conexión comercial con las ciudades de Toluca y de México, por lo que es muy posible que mantenga flujo genético y consecuentemente el grado de carga ancestral NA se vea reducido (Cardoso-Saldaña, *et al.*, 2006). Previamente en nuestro grupo de estudio, también se identificó una estructura genética similar entre los Mazahuas y los Mestizos mediante marcadores multialélicos aunque de carácter no neutro (Camacho-Mejorado, *et al.*, 2015; 2020; Cardoso-Saldaña, *et al.*, 2006). No obstante, con respecto a nuestros 28 AIMs, la distancia entre Mazahuas y Mestizos ($F_{ST} = 0.1494$; $p < 0.0001$) fue significativa, pero no entre los Me'Phaas, Tsotsiles y Mazahuas dejando ver la conexión entre grupos Nativos.

La población Tsotsil perteneciente a la familia lingüística Maya, mostró la mayor variabilidad genética con el menor tamaño poblacional, además de la menor distancia genética con respecto a la población mestiza ($F_{ST} = 0.0812$; $p < 0.045$). Lo anterior sugiere que esta población (cabecera municipal en San Andrés Larrainzar, Chiapas) parece tener un intercambio genético con poblaciones no Nativas. Estudios previos reportan una estructura genética compleja al igual que en poblaciones Mayas (González, 2010; Martínez-Cortes, *et al.*, 2012; Moreno-Estrada, *et al.*, 2014; Santana, *et al.*, 2014; <http://alin.inali.gob.mx/xmlui/>; https://mexico.sil.org/es/lengua_cultura/maya); donde se han encontrado haplogrupos de cromosoma Y de orígenes diversos (Europeos, Euroasiáticos, Asiáticos y Africanos y Nativo Americanos), pudiendo explicar la mayor proporción de impronta Europea de las poblaciones Nativas de estudio (16.4%) (Martínez-Cortes, *et al.*, 2012; Santana, *et al.*, 2014). Asimismo, durante los siglos XIX-XX (González, 2010) se registraron oleadas migratorias de diversas

poblaciones Europeas (Alemania, Bélgica, Croacia, Francia, Grecia, Italia, Reino Unido, Suiza entre otras), además de poblaciones de Eurasia (China, Japón y Rusia), la polinesia (Kanakas), Medio Oriente (Líbano) y Norte América (Canadá y Estados Unidos de Norte América) (González, 2010). Con respecto a la carga Africana en los Tsotsiles, se puede explicar por la llegada de esclavos Africanos durante la colonización Española (www.inafed.gob.mx).

4.2 Diversidad genética del gen *AS3MT* y su relación con la carga ancestral regional

De manera general, las distribuciones de frecuencias de los cinco SNPs en *AS3MT* entre las poblaciones de estudio fueron muy similares. Estos hallazgos son consistentes a los reportados previamente en población general de otras ciudades en México (Drobná, *et al.*, 2013; 2016; Gómez-Rubio, *et al.*, 2012; Xu, *et al.*, 2016). Sin embargo, por lo que a los Nativos Americanos se refiere, éste es el primer reporte en poblaciones del centro y sur de México, siendo la población Me'Phaa la que presentó las mayores diferencias.

Con respecto a la exposición a iAs, las poblaciones del centro-norte, donde se ha reportado mayor exposición a iAs (Alarcón-Herrera, *et al.*, 2013; 2020; Cebrián, *et al.*, 1983; Del Razo, *et al.*, 2011; González-Horta, *et al.*, 2015; Meza, *et al.*, 2005; Sampayo-Reyes, *et al.*, 2010; Xu, *et al.*, 2016) han sido las más estudiadas, aunque con menor número de variantes a este estudio (Drobná, *et al.*, 2013; 2016; Gómez-Rubio, *et al.*, 2012; Xu, *et al.*, 2016; Tabla 7). En particular, el genotipo TT del SNP rs11191439 se consideró poco informativo debido a sus elevadas frecuencias en las poblaciones de estudio, y las reportadas en las poblaciones parentales. Por consiguiente, los resultados del alelo T en relación con el metabolismo del iAs deben ser interpretados con cautela puesto que posiblemente sea reflejo el fondo genético poblacional. Se destaca además la ausencia del alelo C de este SNP en la

población Me'Phaa, lo que podría estar relacionado con su comportamiento endogámico. En cuanto a las poblaciones Mazahua, Mestiza y Tsotsil, los heterocigotos (genotipo CT) tuvieron frecuencias variables, posiblemente a causa del *background* Europeo, dado que el alelo C presenta mayor frecuencia en poblaciones con esta ascendencia (www.internationalgenome.org). En los estados de Chihuahua, Coahuila, Hidalgo y Sonora también se han determinado frecuencias similares de CT (entre 10% y 17%). (Drobná *et al.*, 2013; 2016, Gómez-Rubio, *et al.*, 2012; Xu, *et al.*, 2016). En este sentido, los genotipos encontrados y los reportados previamente pueden reflejar la contribución de la ascendencia Europea identificada mediante los 28 AIMs en las poblaciones de estudio y en poblaciones de la región centro-norte a través de marcadores uniparentales (Martínez-Cortes, *et al.*, 2012; Santana, *et al.*, 2014; Gómez, *et al.* 2020, en revisión); o bien, podrían ser parte de la ventaja del heterocigoto. En lo que concierne a esto último, los individuos heterocigotos tienen una mayor ventaja de resistencia, haciéndolos posiblemente más eficientes (Sellis, *et al.*, 2011). En el caso de las poblaciones de Chihuahua, Ciudad Obregón (Sonora) y Zimapán (Hidalgo) están significativamente expuestas a iAs (entre 75 y 530 ppb de iAs; Alfaro de la Torre, *et al.*, 2018), esta presión del ambiente podría tener una cierta influencia genética. Lo anterior, se podría soportar por las frecuencias haplotípicas y por el número de alelos, teóricamente considerados como eficientes (rs3740393-C, rs10748835-A y rs11191439-T) en las poblaciones de estudio y aquellas de comparación (Tabla 5). No obstante, estos alelos también se presentan en altas frecuencias entre las poblaciones Europeas y Asiáticas (www.internationalgenome.org). La escasez en el número de estudios de variabilidad genética de *AS3MT* y la falta de acceso a sus bases de datos originales limitó la posibilidad de comparar los resultados a nivel haplotípico.

Por lo que respecta a la exposición crónica a iAs (≥ 200 ppb) en combinación con un

metabolismo eficiente, la poca evidencia de efectos tóxicos y las altas frecuencias de ciertos alelos en los *loci* de *AS3MT* han permitido sugerir un proceso de adaptación entre algunas poblaciones del sur del Continente, las que parecen tener un elevado componente Nativo Americano (Ameta, *et al.*, 2017; De la Loma, *et al.*, 2019; Eichstaedt, *et al.*, 2015; Engström, *et al.*, 2011; Schlebusch, *et al.*, 2013; 2015). La población Me'Phaa fue la que más se acercó al comportamiento de este tipo de poblaciones, tanto por su prominente componente NA como en las frecuencias de tres variantes reportadas como eficientes (Tabla 6), sin olvidar que puede ser producto de endogamia o deriva génica para estas poblaciones de ancestría Nativa. Adicionalmente, no se tiene un registro acerca de una alta exposición a iAs en el estado de Guerrero (Alarcón, *et al.*, 2020).

Finalmente, los análisis entre las frecuencias haplotípicas de las poblaciones estudiadas en el presente trabajo y aquellas reportadas en el proyecto de los 1000 genomas con respecto a las poblaciones parentales y otras poblaciones Asiáticas y Latinas podrían sugerir una cierta relación entre los alelos «eficientes» y la carga NA o Asiática. En este sentido el «haplotipo eficiente (rs1119143-T, rs3740400-G, rs7085104-G, rs10748835-A y rs3740393-C)» muestra una distribución similar en las poblaciones Me'Phaa y de Lima Perú por lo que al continente Americano se refiere. Pero también con las poblaciones de Bangladesh, China (Beijin, Han-Sur y Xishuangbanna), Japón y Vietnam (www.internationalgenome.org). De éstas, las poblaciones de Bangladesh, China y Vietnam, son las que están expuestas a altas concentraciones de iAs (200–1,180 ppb, 10-1,000 ppb; 1-632 ppb) (Figura 7) (McArthur, *et al.*, 2004; Rodríguez-Lado, *et al.*, 2013; Agusa, *et al.*, 2014). Estos resultados fueron similares a los de un estudio previo, en el que se relacionaron al ácido monometilarsénico con la carga ancestral. Este análisis encontró patrones genéticos compartidos entre la población de San Antonio de los Cobres y las poblaciones de fuerte

carga NA (PEL y MXL), además de las poblaciones de China, India (Gujarati), Japón y Vietnam (Schlebush, *et al.* 2015). Estos resultados sugirieron que el proceso de selección positiva en la subsistencia a iAs pudo haber sido incluso antes del asentamiento de las poblaciones en Sudamérica (al menos hace 12,000 años; Gómez *et al.*, 2020, en revisión, Jota, *et al.*, 2016; Schurr & Sherry, 2004). El estudio de la variabilidad genética en la población Me'Phaa debe ser ampliado para precisar su contribución a esta hipótesis.

Desde el punto de vista filogenético con respecto al gen *AS3MT*, se destaca la importancia de su actividad como un mecanismo de detoxificación del iAs debido a la presencia de este gen en los múltiples reinos biológicos (Palmgren, *et al.*, 2017; Sumi & Himeno, 2012). Incluso, el *AS3MT* se ha mantenido en el genoma a lo largo de las migraciones y expansiones poblacionales de diferentes especies que se encontraron en ambientes contaminados con iAs, a pesar de que implica la formación de metabolitos más tóxicos que el iAs *per se* (Palmgren, *et al.*, 2017). Es posible entonces que, el contacto con el iAs se incrementara conforme la especie humana migró a regiones de Asia, cruzó por Beringia y llegó a América como lo indican las teorías más soportadas de migraciones humanas y del poblamiento de América, favoreciendo posiblemente la fijación de haplotipos eficientes en el gen *AS3MT* (Gómez *et al.*, 2020, en revisión; Jota, *et al.*, 2016; de Menocal & Chris Stringer, 2016; Schurr & Sherry, 2004; Skoglund & Reich, 2016). Si bien la exposición a iAs reportada actualmente, no representa la exposición de hace miles de años, su origen geogénico (volcánico) nos asegura su presencia en la corteza terrestre e incluso ha sido encontrado en glaciares pleistocénicos en la región de los grandes lagos (Herath, 2016; Szrameka, Walter & McCall, 2004). Por otro lado, se ha sugerido la exposición crónica por hallazgos de iAs en el cabello y huesos de momias de la cultura Chinchorro en Sudamérica que datan de hace ~7,000 años (Arriaza, 2005; Pringle., 2009). Se debe considerar además

que los primeros Americanos debieron seguir rutas migratorias en donde el recurso hídrico no fuera un limitado (Goebel, Waters & O'Rourke, 2008).

Por el contrario, el haplotipo rs11191439-T, rs3740400-T, rs7085104-A, rs10748835-G y rs3740393-G fue más frecuente entre las poblaciones con mayor carga Europea, y por ende, estuvo más presente entre las poblaciones con más mestizaje de nuestro estudio (Mazahuas, Mestizos y Tsotsiles; Figura 7). Diversos estudios de susceptibilidad con respecto a la presencia de ciertos alelos en el *AS3MT* y el incremento en el riesgo de cáncer (e.g., vejiga y pulmón), enfermedades cardiovasculares, lesiones en la piel y esquizofrenia han sido llevados a cabo en poblaciones Asiáticas, Europeas y derivadas de esta última ancestría (D'Ambrosio, *et al.*, 2019; de la Rosa, *et al.*, 2017; Engström, *et al.*, 2014; Pierce, *et al.*, 2012; Song, *et al.*, 2020). Pero nuevamente, las minorías étnicas, como los grupos Nativos y las poblaciones en *admixture* (en las que se llevan a cabo el 3 y el 1% de los estudios, respectivamente) no han sido muy estudiadas, pese a que por su gran antigüedad y multiétnicidad podrían aportar datos que sirvan como predictores de susceptibilidad en la toxicidad mediada por iAs (Popejoy y Fullerton, 2016). Estudios como el presente, podrían contribuir al entendimiento e identificación de grupos susceptibles, así como a la estimación del riesgo debido a la exposición del iAs. Lo anterior toma aún más relevancia debido a que es la sobreexplotación del recurso hídrico, la que compromete su calidad, exponiendo a una mayor cantidad de personas a una mayor concentración de iAs (Alarcón-Herrera, *et al.*, 2020).

En el presente estudio observamos algunas limitaciones y fortalezas. Entre las limitaciones más importantes están: 1) la falta de determinación en los niveles de exposición a iAs en las poblaciones de estudio, 2) determinar el metabolismo de As en los individuos que las integran (no se tiene registro de una exposición mayor a 25 ppb en el estado de

Guerrero), 3) la falta de acceso a las bases de datos en poblaciones Mexicanas en las que sí se cuenta con los datos de los puntos uno y dos, lo que también limitó su comparación con nuestros resultados y 4) un tamaño reducido de la muestra y una baja representación de las poblaciones Nativas en regiones con altas concentraciones de As.

Por lo que se refiere a las fortalezas, se destacó la conjunción entre el *background* genético y la variabilidad del gen *AS3MT* en poblaciones Nativas y Mestiza, siendo de los primeros estudios en describir la distribución de variantes polimórficas relacionadas con cambios en la metilación del iAs en poblaciones Nativas Americanas. Además del diseño de un panel de ancestría con un número accesible de AIMs que permita una determinación precisa de la carga ancestral regional de poblaciones multiétnicas. Nuestros resultados permitieron evidenciar la complejidad de la estructura genética de las poblaciones mexicanas, la que se debería tomar en cuenta en los estudios de asociación con la finalidad de evitar los errores estadísticos tipo I y II.

5. Conclusiones

En conclusión, la identificación de marcadores genéticos como predictores de susceptibilidad en la toxicidad mediada por iAs pueden contribuir al entendimiento e identificación de grupos susceptibles, así como a la estimación del riesgo en las poblaciones vulnerables. Nuestros hallazgos no fueron suficientes para soportar la «hipótesis arsenical» pero sí puso en evidencia los contrastes en la distribución de haplotipos y su relación con la carga ancestral regional. Los 28 AIMS representaron un sistema técnicamente sencillo y confiable que permitió, por medio de técnicas genómicas, una buena opción cuando los recursos son limitados.

6. Perspectivas

A partir de los resultados obtenidos, la principal perspectiva del estudio es continuar con el retrato de la variabilidad genética de *AS3MT* en otras poblaciones Nativas, principalmente de la región norte de México donde la exposición a iAs es conocida, en un mayor tamaño muestral de la población Mestiza y en otras poblaciones del continente que puedan apoyar o no la hipótesis arsenical. Asimismo, conocer la capacidad metabólica al iAs y los niveles de exposición en las poblaciones Nativas y explorar su relación con los datos aquí reportados, ya que el conocimiento del nivel de exposición no representa, necesariamente, la exposición crónica que sugiera un proceso de selección.

Considerar la presencia de AIMs que permitan evaluar la ancestría de Medio Oriente.

Por último, sería interesante explorar a fondo un mayor tamaño de muestra de la población Me'Phaa y estudiarlos mediante marcadores que nos permitan averiguar su identidad por descendencia.

7. Literatura citada

- Agencia para Sustancias Tóxicas y el Registro de Enfermedades. 2007. «ARSENIC.» <https://www.atsdr.cdc.gov/toxprofiles/tp2-a.pdf>.
- Agencia para Sustancias Tóxicas y el Registro de Enfermedades. 2009. *Arsenic Toxicity*. <https://www.atsdr.cdc.gov/csem/csem.asp?csem=1&po=9>.
- Agencia Internacional para la Investigación del Cáncer. 2018. «Arsenic and Arsenic Compounds.» <https://monographs.iarc.fr/wp-content/uploads/2018/06/mono100C-6.pdf>.
- Agusa, T, P Thi, V Lan, D Anh, S Tanabe, y P & Berg, M Viet. 2014. «Human exposure to arsenic from drinking water in Vietnam.» *Science of the Total Environment* 562-569.
- Agusa, T, Fujihara, J, Takeshita, H y Iwata, H 2011. «Individual Variations in Inorganic Arsenic Metabolism Associated with AS3MT Genetic Polymorphisms.» *International Journal of Molecular Sciences* 2351-2382.
- Agencia para Protección Ambiental. 2010. «Toxicological Review if Inorganic Arsenic.» www.epa.gov/iris/IRIS_TOX_ARSENIC_EPR.PDF.
- Alfaro-de la Torre, MC, Medellín-Milán, P. 2018. «Requerimientos para la implementación del derecho humano al agua y al saneamiento en México: caso de estudio en el altiplano potosino» *Revista Legislativa de Estudios Sociales y de Opinión Pública*.
- Alarcón, MT, Leal, L, Miranda, S, Benavides, A y Domínguez, M. 2013. *Arsénico en Agua*. Ciudad de México: Centro de Investigaciones en Materiales Avanzados.
- Alarcón-Herrera, M, Martín-Alarcón, D, Gutierrez, M, Reynoso-Cuevas, L, Martín-Domínguez, A, Olmos-Márquez, M y Bundschuh, J. 2020. «Co-occurrence, possible origin, and health-risk assessment of arsenic and fluoride in drinking water sources in Mexico: Geographical data visualization.» *Science of the Total Enviroment*.
- Apata, M, y Pfeifer, S. 2020. «Recent population genomic insights into the genetic basis of arsenic tolerance in humans: the difficulties of identifying positively selected loci in strongly bottlenecked populations.» *Heredity* 253–262.
- Apata, M, Arriaza, B, Llop, E y Moraga, M. 2017. «Human adaptation to arsenic in Andean populations of the Atacama Desert.» *American Journal of Physical Anthropology* 192-199.
- Armienta. 2015. «Procesos Geológicos y Contaminación de Aguas Subterráneas en México.» https://www.gob.mx/cms/uploads/attachment/file/197304/26_8_Procesos_Geologicos_y_Contaminacion_de_aguas_Subterranneas_en_Mexico.pdf .
- Arriaza, B, Amarasiriwardenad, D, Cornejo, L, Standen, L, Byrned, S, Bartkusd, L, Bandak, B. 2010. «Exploring chronic arsenic poisoning in pre-Columbian Chilean mummies.» *Journal of Archaeological Science*. 1274-1278.

- Barrandeguy, M y García, M. 2014. «Quantifying genetic diversity: the starting point for population genetic studies using molecular markers.» *Journal of Genetics*. 587-589.
- Belbin, G, Nieves-Colón, M, Kenny, E, Moreno-Estrada, A y Gignoux, C. 2018. «Genetic diversity in populations across Latin America: implications for population and medical genetic studies.» *Current Opinion in Genetics & Development* 98-104.
- Belkhir, K, Borsa, P, Chikhi, L, Raufaste, N y Bonhomme, F. 2004. «GENETIX 4.05, logiciel sous Windows TM pour la genetique des populations. .»
- Bittles, A. 2008. «Consanguinity, genetic drift and genetic diseases in populations with reduced numbers of founders.» *Human Genetics - Principles and Approaches*. 4th Ed. Springer.
- Bonilla, C, Gutierrez, G, Parra, E, Kline, C y Shriver, M. 2005. «Admixture analysis of a rural population of the State of Guerrero, Mexico.» *American Journal of Physical Anthropology* 1-10.
- Bryc, K, Velez, C, Karafet, T, Moreno-Estrada, A, Reynolds, A, Auton, A, Hammer, M, Bustamante, C y Ostrer, H. 2010. «Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations.» *Proc Natl Acad Sci U S A* 8954-61.
- Cahua-Pablo, JA, Vidal, M, Cruz, P, Alarcón-Romero, L, Parra, E, Villerías-Salinas, S, Valladares-Salgado, A, Tello-Flores, V, Méndez-Palacios, A, Pérez-Macedonio, C y Flores-Alfaro, E. 2017. «Analysis of admixture proportions in seven geographical regions of the state of Guerrero, Mexico.» *American Journal of Human Biology*.
- Camacho-Mejorado, R, Gómez, R, Torres-Sánchez, Hernández-Tobías, E, Noris, G, Santana, C, Magaña, J, Orozco, L, de la Peña-Días, A, Arenas-Sordo, ML, Meraz-Ríos, MA y Majluf-Cruz, A. 2019. «ALOX5, LPA, MMP9 and TPO gene polymorphisms increase atherothrombosis susceptibility in middle-aged Mexicans.» *Royal Society Open Science* 1-14.
- Cardoso-Saldaña, G, De La Peña-Díaz A, Zamora-González, J, Gomez-Ortega, R, Posadas-Romero, C, Izaguirre-Avila, R, Malvido-Miranda, E, Morales-Anduaga, M y Anglés-Cano, E. 2006. «Ethnicity and lipoprotein(a) polymorphism in Native Mexican populations.» *Ann Hum Biol* 202-2012.
- Cebrián, M, Albores, A, Aguilar, M, Blakely, E. 1983. «Chronic arsenic poisoning in the north of Mexico. *Human Toxicology*.» 121-133.
- Cerda-Flores, RM, Villalobos-Torres, MC, Barrera-Saldaña, HA, Cortés-Prieto, LM, Barajas, LO, Rivas, F, Carracedo, A, Zhong, Y, Barton, SA y Chakraborty, R. 2002. «Genetic admixture in three mexican mestizo populations based on D1S80 and HLA-DQA1 Loci. » *Am. J. Hum. Biol.* 257-263.

- Chen, X, X Guo, P He, J Nie, X Yan, J Zhu, L Zhang, y otros. 2017. «Interactive Influence of N6AMT1 and As3MT Genetic Variations on Arsenic Metabolism in the Population of Inner Mongolia, China.» *Toxicological Sciences* 124-134.
- Choudhry, S, Coyle, N, Tang, H, Salari, K, Lind, D, Clark, S, Tsai, H, Naqvi, M, Phong, A, Ung, N, Matallana, H, Avila, P, Casal, J, Torres, A, Nazario, S, Castro, R, Battle, N, Perez-Stable, E, Kwok, P, Sheppard, D, Shriver, M, Rodriguez-Cintron, W, Risch, N, Ziv, E, González, E. 2006. «Population stratification confounds genetic association studies among Latinos.» *Human Genetics*. 652-664.
- Collins, F. *National Human Genome Research Institute*. <https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms>.
- Chung, CJ, YM Hsueh, CH Bai, YK Huang, YL Huang, MH Yang, y CJ Chen. 2009. «Polymorphisms in arsenic metabolism genes, urinary arsenic methylation profile and cancer.» *Cancer causes & control* 1653-1661.
- D'Ambrosio, E, Dahoun, T, Pardiñas A, Veronese, M, Bloomfield, M, Jauhar, S, Bonoldi, I 2019. «The effect of a genetic variant at the schizophrenia associated AS3MT/BORCS7 locus on striatal dopamine function: A PET imaging study.» *Psychiatry Research: Neuroimaging* 34-41.
- de la Loma, J, Tirado, N, Ascui, F, Levi, M, Vahter, M, Broberg, K, Gardon, J. 2019. «Elevated arsenic exposure and efficient arsenic metabolism in indigenous women around Lake Poopó, Bolivia.» *Sci Total Environ*. 179-186.
- de la Rosa, R, Steinmaus, C, Aker, N, Conde, L, Ferreccio, C, Kalman, D, Zhang, A, Chang, L, y Smith, M. 2017. «Associations between arsenic (+3 oxidation state) methyltransferase (AS3MT) and N-6 adenine-specific DNA methyltransferase 1 (N6AMT1) polymorphisms, arsenic metabolism, and cancer risk in a Chilean population.» *Environ Mol Mutagen* 411-422.
- de Menocal, P & Stringer, C. 2016. Climate and the peopling of the world. *Nature*. 49-50.
- del Razo, LM, García-Vargas, G, Valenzuela, O, Hernandez-Castellanos, E, Sánchez-Peña, L, Currier, J, Drobná, Z, Loomis, D, Stylbo, M. 2011. Exposure to arsenic in drinking water is associated with increased prevalence of diabetes: a cross-sectional study in the Zimapán and Lagunera regions in Mexico. *Environmental Health*.
- Drobná, Z, Waters, S, Walton, F, LeCluyse, E, Thomas, D y Stýblo, M. 2004. «Interindividual variation in the metabolism of arsenic in cultured primary human hepatocytes.» *Toxicology and applied pharmacology* 166-177.
- Drobná, Z, Martin, E, Kim, K, Smeester, L, Bommarito, P, Rubio-Andrade, M, García-Vargas, G, Stýblo, M, Zou, F, y Fryb, R. 2016. «Analysis of maternal polymorphisms in arsenic (+3 oxidation state)-methyltransferase AS3MT and fetal sex in relation to arsenic metabolism and infant birth outcomes: implications for risk analysis.» *Reproductive Toxicology* 28-38.

- Earl & vonHoldt. 2012. «STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.» *Conservation Genetics Resources*. 359-361.
- Eichstaedt, C, Antao, T, Cardona, A, Pagani, L, Kivisild, T y Mormina, M. 2015. «Positive selection of AS3MT to arsenic water in Andean populations.» *Mutation Research* 97-102.
- Elkhai, E, Tatarinova, T, Chebotarev, D, Piras, I, Calo, C, De Montis, A, Atzori, , Marini, M, Tofanelli, S, Francalacci, P, Pagani, L, Tyler-Smith, C, Xue, Y, Cucca, F, Schurr, T, Gaieski, J, Meedez, C, Vilar, M, Owings, A, Gómez, R, Fujita, R, Santos, F, Comas, D, Balanovsky, O, Balanovska, E, Zalloua, P, Soodyall, H, Pitchappan, R, GaneshPrasad, A, Hammer, M, Well, S y The Genographic Consortium. 2014. «Geographic population structure analysis of worldwide human populations infers their biogeographical origins. » *Nature Communications*.
- Engström, K, Vahter, M, Jurkovic, S, Concha, G, Nermell, N, Raqib, R, Cardozo, A y Broberg, K. 2011. «Polymorphisms in Arsenic(+III Oxidation State) Methyltransferase (AS3MT) Predict Gene Expression of AS3MT as Well as Arsenic Metabolism.» *Environmental Health Perspectives* 182-188.
- Engström, K, Vahter, M, Fletcher, T, Leonardi, G, Goessler, W, Gurzau, E, Koppova, K, Rudnai, P, Kumar, R y Broberg, K. 2014. «Genetic variation in arsenic (+3 oxidation state) methyltransferase (AS3MT), arsenic metabolism and risk of basal cell carcinoma in a European population.» *Environmental and Molecular Mutagenesis* 60-69.
- Engström, S, B Nermell, G Concha, U Strömberg, M Vahter, y K Broberg. 2009. «Arsenic metabolism is influenced by polymorphisms in genes involved in one-carbon metabolism and reduction reactions.» *Mutation Research* 4-14.
- Engström, S, K Broberg, G Concha, B Nermell, M Warholm, y M Vahter. 2007. «Genetic polymorphisms influencing arsenic metabolism: evidence from Argentina.» *Environmental Health Perspectives* 599-605.
- Ferrairo, D, Croera, C, Brustio, R, Collotta, A, Bowe, G, Vhater, M, Gribaldo, L. 2008. «Toxicity of inorganic arsenic and its metabolites on haematopoietic progenitors “in vitro”: Comparison between species and sexes. » *Toxicology*. 102-108.
- Fisher, A, López-Carrilo, L, Gamboa-Loira, B, y Cebrián, M. 2017. «Standards for arsenic in drinking water: Implications for policy in Mexico.» *Journal of Public Health Policy* 395-406.
- Fontdevila, A. Introducción a la Genética de Poblaciones. 2000. *Editorial Síntesis*. Madrid.
- Forrester, J, Dick, A, McMenamin, P, Roberts, F y Pearlman, E. 2016. «Genetics.» En *The Eye*, 130-156. W.B. Saunders.

- Fujihara J, Fujii Y, Agusa T, Kunito T, Yasuda T, Moritani T, Takeshita H. 2009. «Ethnic differences in five intronic polymorphisms associated with arsenic metabolism within human arsenic (+3 oxidation state) methyltransferase (AS3MT) gene.» *Toxicol. Appl. Pharmacol.* 41–46.
- Fujihara J, Soejima M, Yasuda T, Koda Y, Agusa T, Kunito T, Tongu M, Yamada T, Takeshita H. 2010. «Global analysis of genetic variation in human arsenic (+3 oxidation state) methyltransferase (AS3MT).» *Toxicol. Appl. Pharmacol.* 292–299
- Galanter, J, Fernandez-Lopez, J, Gignoux, C, Barnholtz, J, Fernández, C, Vía, M, Hidalgo, A, 2012. «<https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1002554>.» *PLOS Genetics*.
- García-Alvarado, F, Neri-Melendez, H, Pérez-Armendáriz, L y Rivera-Guillén, M. 2018. «Polimorfismos del gen arsénico 3 metiltransferasa (As3MT) y la eficiencia urinaria del metabolismo del arsénico en una población del norte de México.» *Revista Peruana de Medicina Experimental y Salud Publica* 72-76.
- Ghosh, J, y Sil, P. 2015. «Handbook of Arsenic Toxicology.» 203-231. India: Academic Press.
- Gleizer, D. 2016. «Las relaciones entre México y el Tercer Reich, 1933-1941.» *Tzintzun. Revista de estudios históricos*.
- González-Sobrino, BZ, Pintado-Cortina, AP, L Sebastián-Medina, F Morales-Mandujano, AV Contreras, YE Aguilar, y otros. 2016. «Genetic Diversity and Differentiation in Urban and Indigenous Populations of Mexico: Patterns of Mitochondrial DNA and Y-Chromosome Lineages.» *Biodemography and Social Biology* 53-72.
- Gómez, R, Magaña, J, Cisneros, B, Perez-Salazar, E, Faugeron, S, Martinez, D, Nava, GA Veliz, D, Santana, C y Noris, G. 2007. «Association of the estrogen receptor alpha gene polymorphisms with osteoporosis in the Mexican population.» *Clin. Genet* 574-581.
- Gómez, R, Schurr, TG, y Meraz-Ríos, MA. 2020. «Diversity of Mexican Paternal Lineages Reflects Evidence of Migration and 500 Years of Admixture.» En *Human Migration: Biocultural Perspectives*, de Muñoz Moreno ML y Crwaford MH. Oxford: En impresión.
- Gómez, R, Vilar, M, Meraz-Ríos, MA, Véliz, D, Zúñiga, G, Hernández-Tobías, EA, Figueroa-Corona, M, Owings A, Gaieski, J, Schurr, TG & The Genographic Consortium. 2020 «Y-chromosome Diversity in *Aztlán* Descendants and Its implications for the History of Central Mexico.» *En revisión*.
- Gómez-Rubio, P, Klimentidis Y, Cantu-Soto, E, Meza-Montenegro, M, Billheier, D, Lu, Z, Chen, X, y Klimecki, W. 2012. «Indigenous American Ancestry is associated with methylation efficiency in a admixed population of Northwest Mexico.» *Journal of Toxicology and Environmental Health* 36-49.

- Griffits, A. 2000. «Genética moderna.» *McGraw-Hill Ineramericana*.
- Hartl, D y Clark, A. 1997. «Principles of Population Genetics.» Massachusetts: Sinauer Associates, Inc.
- Herath, I, Vithanage, M, Bundschuh, J, Prakash, J y Bhattacharia, P. 2016. «Natural Arsenic in Global Groundwaters: Distribution and Geochemical Triggers for Mobilization.» *Current Populations Reports* 68-89.
- Hernández, A, Xamena, N, Surrallés, J, Sekaran, C, Tokunaga, H, Quinteros, D, Creus, A y Marcos, R. 2008. «Role of the Met(287)Thr polymorphism in the *AS3MT* gene on the metabolic arsenic profile.» *Mutation Research* 80-92.
- Hernández-Gutiérrez, S, Hernández-Franco, P, Martínez-Tripp, S, Ramos-Kuri, M y Rangel-Villalobos, H. 2004. «STR data for 15 loci in a population sample from the central region of Mexico.» *Forensic Science International* 97-100.
- Hoyos-Puente, J. 2011. «La Guerra Civil en los imaginarios del exilio republicano en México, 1939-1960.» *Amnis*.
- Huerta-Chagoya, Al, Moreno-Macias, H, Fernández-López, J, Ordóñez-Sánchez, ML, Rodríguez-Guillén, R, Contreras, A. «A panel of 32 AIMs suitable for population stratification correction and global ancestry estimation in Mexican mestizos.» *BMC Genetics* 1-11.
- IBM Statistical Package for the Social Sciences (SPSS). 2013. SPSS for Windows, v.22.0. Chicago, SPSS, Inc
- Instituto de Biotecnología de la Universidad Nacional Autónoma de México. *Unidad de Síntesis y Secuenciación de ADN*. <http://www.ibt.unam.mx/sintesis/cuantificacion.html>.
- Instituto Nacional de Estadística y Geografía. 2015. «Principales Resultados de la Encuesta Nacional 2015.» http://www.senado.gob.mx/comisiones/asuntos_indigenas/eventos/docs/etnicidad_240216.pdf.
- Instituto Nacional de Lenguas Indígenas. 2008. *Catálogo de las Lenguas Indígenas Nacionales*:. <http://www.inali.gob.mx/clin-inali/>.
- Jacovas, V, Couto-Silva, C, Nunes, K, Lemes, R, Oliveira, M, Salzano, F, Cátira, M y Hunemeier, T. 2018. «Selection scan reveals three new loci related to high altitude adaptation in Native Andeans.» *Scientific Reports Nature* 1-8.
- Janasik, B, Reszka, E, Stanislawska, M, Wiczorek, E, Fendler, W y Wasowicz, W. 2015. «Biological monitoring and the influence of genetic polymorphism of *As3MT* and *GSTs* on distribution of urinary arsenic species in occupational exposure workers.» *International Archives of Occupational and Environmental Health* 807-818.
- Jia, Y. 2012. «Real-Time PCR.» En *Methods in Cell Biology*, 55-68. Oregon: ElSevier.

- Johnson, N, Coram, M, Shriver, M, Romieu, I, Barsh, G, London, S, y Tang, H. 2011. «Ancestral Components of Admixed Genomes in a Mexican Cohort.» *PLoS Genetics*.
- Jota, M, D Lacerda, J Sandoval, P Vieira, S Santos-Lopes, R Bisso-Machado, V Paixao-Cortés, y otros. 2016. «New Native South American Y-chromosome lineages.» *J Hum Genet*.
- Juárez-Cedillo, T, Zuñiga, J, Acuña-Alonso, V, Pérez-Hernández, N, Rodríguez-Pérez, J, Banquera, R, Gallardo, G, Sánchez-Arenas, R, García-Peña, M, Granados, J y Vargas-Alarcón, G. 2008. «Genetic admixture and diversity estimations in the Mexican Mestizo population from Mexico City using 15 STR polymorphic markers.» *Forensic Science International: Genetics* 37-39.
- Kardos, M, G Luikart, y A Allendorf. 2015. «Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees.» *Heredity* 63-72.
- Khairul, I, Wang, Q, Jiang, Y, Wang, C y Naranmanrua, H. 2017. «Metabolism, toxicity and anticancer activities of arsenic compounds.» *Oncotarget* 23905-23926.
- Kidd, K, W Speed, A Pakstis, M Furtado, R Fang, A Madbouly, M Maires, M Middhaa, F Friedlaender, y J Kidd. 2014. «Progress toward an efficient panel of SNPs for ancestry inference.» *Forensic, Science International: Genetics* 23-32.
- Kidd, K, Speed, W, Pakstis, A, Furtado, M, Fang, R, Madbouly, A, Maiers, M, Middha, M y Kidd, J. 2014. «Progress toward an efficient panel of SNPs for ancestry inference.» *Forensic Science International: Genetics* 23-32.
- Kosoy, R, Nassir, R, Tian, C, White, PA, Buttler, LM, Silva, G, Kittles, R y otros. 2009. «Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America.» *Human Mutation* 69-78.
- Lawrence, y Brody. *National Human Genome Research Institute*. <https://www.genome.gov/genetics-glossary/Ancestry-informative-Markers>.
- Lemes, R. 2018. «Inbreeding estimates in human populations: Applying new approaches to an admixed Brazilian isolate.» *PLoS One*.
- Lindberg, A, Kumar, R, Goessler, W, Thirumaran, R, Gurzau, E, Koppova, K, Rudnai, P, Leonardi, G, Fletcher, T y Vahter, M. 2007. «Metabolism of low-dose inorganic arsenic in a central European population: influence of sex and genetic polymorphisms.» *Environmental Health Perspectives* 1081-1086.
- López-Parra, A, P Iturbe, Q Huang, y M Mesa. 2016. «Genetic polymorphism of 15 STR loci in 3 ethnics groups of Guerrero State, Mexico.» *Foresic Science International: Genetics* 8-9.
- Luna-Vázquez, A, Vilchis-Dorantes, G, Aguilar-Ruiz, M, Bautista-Rivas, A, Rojo-Nava, A, Rios-Barrios, E, y Rangel-Villalobos, H. 2005. «Population data for 15 loci (identifiler kit) in a samle from the Valley of Mexico.» *Legal Medicine* 331-333.

- Marchini, J, Cardon, L, Phillips, M y Donnelly, P. 2004. «The effects of human population structure on large genetic association studies.» *Nature Genetics* 512-517.
- Martínez-Cortes, G, Salazar-Flores, J, Fernandez-Rodríguez, L, Rubi-Castellanos, R, Rodríguez-Loya, C, Velarde-Felix, J, Muñoz-Valle, J, Parra-Rojas, I, Rangel-Villalobos, H. 2012. «Admixture and population structure in Mexican-Mestizos based on paternal lineages.» *Journal of Human Genetics* 568–574.
- Martínez-Cortés, G, Salazar-Flores, J, Haro-Guerrero, J, Rubí-Castellanos, R, Velarde-Félix, J, Muñoz-Valle, J, López-Casamichana, M, Carrillo-Tapia, E, Canseco-Avila, L, Bravi, C, López-Armenta, M, Rangel-Villalobos, H. 2013. «Maternal admixture and population structure in Mexican–Mestizos based on mtDNA haplogroups ». *American Journal of Physical Anthropology*. 526-537.
- Martínez-Fierro, M, Joke Beuten, Robin J Leach, Esteban J Parra, Miguel Cruz-Lopez, Hector Rangel-Villalobos, Lina R Riego-Ruiz, Rocio Ortiz-Lopez, Herminia G Martinez-Rodriguez, Augusto Rojas-Martinez. 2009. «Ancestry informative markers and admixture proportions in northeastern Mexico. » *Journal of human genetics*. 504-509.
- Martínez-Marignac, V, Valladares, A, Cameron, E, Chan, A, Perera, A, Globus-Goldberg, R, Wachter, N, Kumate, J, McKeigue, P, O'Donnell, D, Shriver, M, Cruz, M, Parra, E. 2007. «Admixture in Mexico City: implications for admixture mappin of Type 2 diabetes genetic risk factors.» *Human Genetics*. 807-819.
- McArthur, J, D Banerjee, K Hudson-Edwards, R Mishra, R Purohit, P Ravenscroft, A Cronin, R Howarth, A Chatterjee, y T & Chadha, D Talukder. 2004. «atural organic matter in sedimentary basins and its relation to arsenic in anoxic ground water: the example of West Bengal and its worldwide implications.» *Applied Geochemistry* 1255-1293.
- McVean, G. 2001. «Population structure.» *Department of Statistics, Oxford University*. <http://www.stats.ox.ac.uk/~mcvean/notes7.pdf>.
- Meza, M, Gandolfi, A y Klimecki, W. 2007. «Developmental and Genetic Modulation of Arsenic Biotransformation: A Gene by Environment Interaction?» *Toxicology and applied Pharmacology* 381-387.
- Montoya, E, Hernández, L, Escareño, M y Balagurusamy, N. 2015 «Impacto del arsénico en el ambiente y su transformación por microorganismos.» *Terra latinoamericana*. 103-118.
- Moreno-Estrada, A, Gignoux, C, Fernández, J, Zakharia, F, Sikora, M, Contreras, A, Acuña, V y otros. 2014. «The genetics of Mexico recapitulates Native American substructure and affects biomedical traits.» *Science* 1280-1285.
- Noris, G, Santana, C, Meraz-Ríos, MA,, Muñoz, ML, Majluf-Cruz, A, Magaña, JJ, Granados, J, Quezada, R, Revilla, MC, Martínez-Salas, S, Xihuitl, S, Martínez de la Escalera, G, Díaz-Badillo, A, Calderón-Aranda, ES, Gómez, R. 2012. «Mexican mestizo

- population sub-structure: effects on genetic and forensic statistical parameters.» *Mol Biol Rep.* 10139-56.
- O'Fallon BD, Fehren-Schmitz L. «Native Americans experienced a strong population bottleneck coincident with European contact.» 2011. *Proc Natl Acad Sci U S A.* 20444-8.
- Organización Mundial de la Salud. 2018. *Arsenic.* 15 de Febrero. <https://www.who.int/news-room/fact-sheets/detail/arsenic>.
- Organización Mundial de la Salud. 2011. «Arsenic in Drinking-water.» https://www.who.int/water_sanitation_health/dwq/chemicals/arsenic.pdf.
- Organización Mundial de la Salud. 2010. «Exposure to Arsenic: A Major Public Health Concern.» <https://www.who.int/ipcs/features/arsenic.pdf>.
- Palmgren, M, Engstrom, K, Hallstrom, B, Wahlberg, K, Sondergaard, D, Sall, T, Vahter, M, y Broberg, K. 2017. «AS3MT-mediated tolerance to arsenic evolved by multiple independent horizontal gene transfers from bacteria to eukaryotes.» *PLOS One.*
- Pfaffelhuber, P, Grundner-Culemann, F, Lippard, V y Baumdicker, F. 2019. «How to choose sets of ancestry informative markers: A supervised feature selection approach.» *Forensic Sciences International: Genetics.*
- Pierce BL, Kibriya MG, Tong L, Jasmine F, Argos M, Roy S. 2012. «Genome-Wide Association Study Identifies Chromosome 10q24.32 Variants Associated with Arsenic Metabolism and Toxicity Phenotypes in Bangladesh.» *PLoS Genet.*
- Podgorski, J y Berg, M. 2020. «Global threat of arsenic in groundwater.» *Science* 845-850.
- Popejoy, A, & Fullerton, S. 2016. «Genomics is failing on diversity.» *Nature.* 161–164.
- Pringle H. 2009. «Archaeology. Arsenic and old mummies: poison may have spurred first mummies.» *Science.*
- Pritchard, Stephens & Donnelly. 2000. «STRUCTURE 2.3.4.»
- Ramachandran, S, Rosenberg, NA, Zhivotovsky, L, Feldman, MW. 2004. «Robustness of the inference of human population structure: a comparison of X-chromosomal and autosomal microsatellites.» *Human Genomics.* 87-97.
- Rahim, N, Harismendy, O, Topol, E, Frazer, K. 2008. «Genetic determinants of phenotypic diversity in humans. » *Open Access.*
- Rangel, E, Montañez, L, Luévanos, M y Balagurusamy, N. 2015. «Impacto del arsénico en el ambiente y su transformación por microorganismos.» *Terra Latinoamericana* 103-118.

- Rangel-Villalobos, H, Muñoz-Valle, JF, González-Martín, A, Gorostiza, A, Magaña, A, Páez-Riberos, LA. 2008. «Genetic admixture, relatedness, and structure patterns among Mexican populations revealed by the Y-chromosome.» *American Journal of Physical Anthropology* 448-461.
- Ratnaik, R. 2003. «Acute and Chronic Arsenic Toxicity.» *Postgraduate Medical Journal*. 391-396.
- Reich, D, Patterson, N, Campbell, D, Tandon, A, Mazieres, S, Ray, N, Parra, M, Rojas, W, Duque, C, Mesa, N, García, L, Triana, O, Bair, S, Maestre, A, Dib, J, Bravi, C, Bailliet, G, Corach, D, Hunemeier, T, Bortolini, M, Salzano, F, Ruiz-Linares, A y otros. 2012. «Reconstructing Native American population history.» *Nature*. 370-374.
- Rodríguez-Lado, L, Sun, G, Berg, M, Zhang, Q, Xue, H, Zheng, Q, & Johnson, A. 2013. «Groundwater Arsenic Contamination Throughout China.» *Science* 866-868.
- Rosenberg, N, Li, R, Ward, R, y Pritchard, J. 2003. «Informativeness of Genetic Markers for Inference of Ancestry.» *American Journal of Human Genetics* 1402-1422.
- Rubí-Castellanos, R, Anaya-Palafox, M, Mena-Rojas, E, Bautista-España, D, Muñoz-Valle, Rangel-Villalobos, H. 2009. «Genetic data of 15 autosomal STRs (Identifiler kit) of three Mexican Mestizo population samples from the States of Jalisco (West), Puebla (Center), and Yucatan (Southeast).» *Elsevier*. 71-76.
- Salazar-Flores, J, Dondiego-Aldape, R, Rubí-Castellanos, R, Anaya-Palafox, M, Nuño-Arana, I, Canse, L, Flores-Flores, G, Morales, M, Barojas-Pérez, N, Muñoz-Valle, J, 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*. 401-409.
- Salzano, F y Sans, M. 2014. «Interethnic admixture and the evolution of Latin American populations.» *Genetics and Molecular Biology* 151-170.
- Sampayo-Reyes, A, Hernández, A, López-Campos, C, Mayet-Machado, E, Rincón-Castañeda, C, Limones-Aguilar, M, López-Campos, J, Bermudez, M, Marcos, R. 2010. «Arsenic Induces DNA Damage in Environmentally Exposed Mexican Children and Adults. Influence of GSTO1 and AS3MT Polymorphisms.» *Toxicological Sciences*. 63-71.
- Sánchez-Serrano, C. 2004. «Mestizaje y características físicas de la población mexicana.» *Arqueología*.
- Sandoval, K, Moreno-Estrada, I, Mendizabal, P, Underhill, M, Lopez-Valenzuela, R, Peñaloza-Espinosa, Marisol Lopez-Lopez, Leonor Buentello-Malo, y D Comas. 2012. «Sandoval, K., Moreno-Estrada, A., Mendizabal, I., Underhill, P. A., Lopez-Valenzuela, M., Peñaloza-Espinosa, R., ... Comas, D. (2012). Y-chromosome diversity in Native Mexicans reveals continental transition of genetic structure in the Americas. *American Journal of Physical Anthropology* 395-405.» *American Journal of Physical Anthropology* 395-405.

- Santana, C, G Noris, MA Meraz-Rios, JJ Magaña, ES Calderon-Aranda, L Muñoz, y R Gómez. 2014. «Genetic analysis of 17 Y-STRs in a Mestizo population from the Central Valley of Mexico.» *Human Biology* 289-312.
- Schlebusch, C, L Gattepaille, K Engström, M Vahter, M Jakobsson, y K Broberg. 2015. «Human adaptation to arsenic-rich environments.» *Molecular biology and evolution* 1544-1555.
- Scientific, Thermo Fischer. 2009. «Thermo Fischer.» <https://assets.thermofisher.com/TFS-Assets/CAD/manuals/NanoDrop-2000-User-Manual-EN.pdf>.
- Schurr, TG., & Sherry. 2004. «Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence.» *Am J Hum Biol* 420-439.
- Sellis, D., Callahan, B., Petrov, D., & Messer, P. 2011. «Heterozygote advantage as a natural consequence of adaptation in diploids.» *Proc Natl Acad Sci USA* 20666-20671.
- Silva-Zolezzi, I, Hidalgo-Miranda, A, Estrada-Gil, J, Fernandez, JC, Uribe-Figueroa, L, Contreras, A, Balam-Ortiz, E y otros. 2009. «Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico.» *PNAS* 8611-8616.
- Skoglund, P, y Reich, D. 2017. «A genomic view of the peopling of the Americas.» *Curr Opin Genet Dev* 27-35.
- Song, Y, D Jin, J Chen, W liang, y X Liu. 2020. «Effects of Arsenic (+3 oxidation state) Methyltransferase Gene Polymorphisms and Expression on Bladder Cancer: Evidence From A Systematic Review, Meta-analysis and TCGA dataset.» *Toxicological Sciences*.
- Soundararajana, U, Libing, Y, Meisen, Kidd, K y Shi. 2016. «Minimal SNP overlap among multiple panels of ancestry informative markers argues for more international collaboration.» *Forensic Science International: Genetics* 25-32.
- Sumi, D, & Himeno, S. 2012. «Role of Arsenic (+3 Oxidation State) Methyltransferase in Arsenic Metabolism and Toxicity.» *Biol. Pharm. Bull* 1870-1875.
- Szramek, K, Walter, L y McCall, P. 2004. «Arsenic mobility in groundwater/surface water systems in carbonate-rich Pleistocene glacial drift aquifers (Michigan).» *Applied Geochemistry* 1137-1155.
- Tian, C, DA Hinds, R Shigeta, SG Adler, A Lee, MV Pahl, G Silva, y otros. 2007. «A genomewide single-nucleotide-polymorphism panel for Mexican American admixture mapping.» *American Journal of Human Genetics* 1014-1023.
- Torrioni, A. and Wallace, D.C. 1995. «MtDNA haplogroups in Native Americans. *American Journal of Human Genetics*.» 1234-1238.
- Uniprot. s.f. *UniProtKB*. <https://www.uniprot.org/uniprot/Q9HBK9>.

- United Nation 5th Session. 2006. *Who are indigenous peoples?* New York: UN.
- Valenzuela, O, Drobná, Z, Hernández-Castellanos, E, Sánchez-Peña, L, García-Vargas, G, Borja-Aburto, V, Stýblo, M y Del Razo, LM. 2009. «Association of AS3MT polymorphisms and the risk of premalignant arsenic skin lesions.» *Toxicology and Applied Pharmacology* 200-207.
- Vahter, M. 1999. «Methylation of inorganic arsenic in different mammalian species and population groups.» *Science Progress* 69-88.
- Watanabe, T y Hirano, S. 2013. «Metabolism of arsenic and its toxicological relevance.» *Archives of Toxicology* 969-979.
- Willoughby, J, Fernández,N, Lamb, M, Ivy, J, Lacy, R, DeWoody, A. 2014. «The impacts of inbreeding, drift, and selection on genetic diversity in captive breeding populations.» *Molecular ecology*. 98-110.
- Witherspoon, D, Wooding, A, Rogers, A, Marchini, E, Watkins, W, Batzer, M, y Jorde, L. 2007. «Genetic Similarities Within and Between Human Populations.» *Genetics*. 351-359.
- Wright, S. 1951. «The genetical structure od populations.» *Annals of eugenics* 323-354.
- Xu X, Drobná Z, Voruganti VS, Barron K, González-Horta C, Sánchez-Ramírez B, Ballinas-Casarrubias L, Cerón RH, Morales DV, Terrazas FA, Ishida MC, Gutiérrez-Torres DS, Saunders RJ, Crandell J, Fry RC, Loomis D, García-Vargas GG, Del Razo LM, Stýblo M, Mendez MA. 2016. «Association Between Variants in Arsenic (+3 Oxidation State) Methyltransferase (AS3MT) and Urinary Metabolites of Inorganic Arsenic: Role of Exposure Level. *Toxicol Sci*. 112-123. »

8. ANEXO

Descripción de métodos del estudio

- *Extracción de ácidos nucleicos por el kit comercial: DNA Blood Preparation* (Jena Bioscience; Lobstedter, Alemania)

Procedimiento:

1. *Lisis celular:*

- 1.1 Pipetee 900 μL de solución de lisis de glóbulos rojos en un microtubo de 1.5 ml, agregue 300 μL de sangre total o médula ósea e invertir 10 veces.
- 1.2 Incube durante 3 min a temperatura ambiente con inversión ocasional. Nota: Para sangre fresca recolectada dentro de 1 hora antes de la preparación, aumente el tiempo de incubación a 10 minutos para asegurar la lisis completa de glóbulos rojos.
- 1.3 Centrifugar durante 30 segundos a 15.000 g.
- 1.4 Retire el sobrenadante con una pipeta dejando atrás el visible sedimento celular. Asegúrese de no exceder los 20 μL de líquido residual. Esta es un punto crítico para la precipitación efectiva de proteínas y ADN en los siguientes pasos.
- 1.5 Agite el tubo con fuerza durante 10 segundos para resuspender el blanco de células en el líquido residual. El sedimento de glóbulos blancos debe resuspenderse por completo.
- 1.6 Añada 300 μL de solución de lisis celular a las células resuspendidas y pipetear hacia arriba y hacia abajo para lisar las células hasta que no se vean grumos.
- 1.7 Para sangre tratada con heparina, caliente el sedimento de glóbulos blancos durante 10 minutos a 65 ° C para facilitar la lisis.

2. *Precipitación de proteínas:*

- 2.1 Añada 100 μL de solución de precipitación de proteínas al lisado celular.
- 2.2 Agite vigorosamente durante 20 segundos para mezclar bien. Pequeñas partículas de la proteína precipitada (sin grumos) debe ser visible.
- 2.3 Centrifugar a 15,000 g durante 1 min. Las proteínas precipitadas deben formar un sedimento oscuro y apretado. Si el el sedimento de proteína no está apretado, repita el vórtex, seguido de incubación en hielo durante 5 min y centrifugue nuevamente.

3. *Precipitación de ADN:*

- 3.1 Pipetee 300 μL de isopropanol > 99% en un microtubo limpio de 1.5 mL y agregue el sobrenadante.
- 3.2 Mezclar la muestra invirtiendo suavemente durante 1 min.
- 3.3 Centrifugar a 15,000 g durante 1 min. El ADN debe ser visible como un pequeño bolita blanca.
- 3.4 Desechar el sobrenadante y escurrir brevemente el tubo sobre papel absorbente limpio.
- 3.5 Añada 500 μL de tampón de lavado e invierta el tubo varias veces para lavar el sedimento de ADN.
- 3.6 Centrifugar a 15,000 g durante 1 min.
- 3.7 Deseche con cuidado el etanol y séquelo a temperatura ambiente durante aproximadamente de 10 a 15 min.

4. *Hidratación del ADN:*

- 4.1 Añada 50-100 μL de solución de hidratación de ADN.
- 4.2 Agite en el vórtex durante 5 segundos a velocidad media para mezclar.
- 4.3 Incube la muestra a 65 ° C durante 30 min para acelerar la rehidratación.
- 4.4 Almacene el ADN a 4 ° C. Para un almacenamiento prolongado, coloque la muestra a -20 ° C o -80 ° C.

- ***Cuantificación y determinación de pureza de DNA por mediante el espectrofotómetro NanoDrop 2000™ (ThermoScientific, Massachusetts, Estados Unidos)***

1. Pipetee 1 - 2 μL de muestra directamente en el pedestal de medición.
Consejo: se pueden utilizar alícuotas de 0.5 μL para muestras que tengan valores de absorbancia equivalentes altos, es decir en muestras concentradas (> 150 ng / μL dsDNA). Solo modelo NanoDrop 2000 / 2000c.
2. Baje el brazo de muestreo e inicie una medición espectral usando el software en la PC.
La tensión superficial se utiliza para mantener las muestras en su lugar entre dos fibras ópticas. La luz de una lámpara de destellos de xenón pasa a través de la fibra óptica superior, hacia abajo a través del líquido columna y es detectado por el espectrómetro interno.

3. Registre su lectura de la muestra.
4. Cuando se complete la medición, levante el brazo de muestreo y limpie la muestra de ambos los pedestales superior e inferior con un paño de laboratorio seco y sin pelusa.
5. Usando esta tecnología, los espectrofotómetros NanoDrop tienen la capacidad de medir muestras entre 50 y 200 veces más concentradas que las muestras medidas con un estándar en cubeta de 1 cm.

Tablas y figuras

Tabla A1. Fuentes de revisión para el diseño del panel de AIMS.

	SNPforID Browser http://spsmart.cesga.es/snpforid.php
	The ALlele FREquency Database https://alfred.med.yale.edu/alfred/index.asp
Base de	EUROFORGEN (Phillips, <i>et al.</i> , 2014)
Datos/Paneles	Bulbul's 32plex(22AIMS+10PI markers) (Bulbul, <i>et al.</i> , 2011)
de referencia	Daniele Podini's list of 32 AISNPs (Butler, <i>et al.</i> , 2014) ID34plex (Santos, <i>et al.</i> , 2015)
	Development of a Panel of Genome-Wide Ancestry Informative Markers to Study Admixture Throughout the Americas (Galanter, <i>et al.</i> , 2012).
	Ancestry Informative Marker Sets for Determining Continental Origin and Admixture Proportions in Common Populations in America (Kosoy, <i>et al.</i> , 2009).
Artículos	Analysis of genomic diversity in Mexican Mestizo populations to develop genomic medicine in Mexico (Silva-Solezzi, <i>et al.</i> , 2009).
	Analyses of a set of 128 ancestry informative single-nucleotide polymorphisms in a global set of 119 population samples (Kidd, <i>et al.</i> , 2011)
	An ancestry informative marker set for determining continental origin: validation and extension using human genome diversity panels (Nassir, <i>et al.</i> , 2009).
	Using ancestry-informative markers to identify fine structure across 15 populations of European origin (Huckins, <i>et al.</i> , 2014).
	Proportioning Whole-Genome Single-Nucleotide-Polymorphism Diversity for the Identification of Geographic Population Structure and Genetic Ancestry (Lao, <i>et al.</i> , 2016).
	A Genomewide Single-Nucleotide-Polymorphism Panel for Mexican American Admixture Mapping (Tian, <i>et al.</i> , 2007).
	A Genomewide Admixture Map for Latino Population (Price, <i>et al.</i> , 2007).
	A Genomewide Admixture Mapping Panel for Hispanic/Latino Populations (Mao, <i>et al.</i> , 2007).

A Genomewide Single-Nucleotide–Polymorphism Panel with High Ancestry Information for African American Admixture Mapping (Tian, *et al.*, 2006).

Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals (Ruiz-Linares, *et al.*, 2014)

Design and Analysis of Admixture Mapping Studies (Hoggart, *et al.*, 2004).

Evaluating the accuracy of AIM panels at quantifying genome ancestry (Pardo-seco, *et al.*, 2014).

Informativeness of Genetic Markers for Inference of Ancestry (Rosenberg, *et al.*, 2003)

Minimal SNP overlap among multiple panels of ancestry informative markers argues for more international collaboration (Soundararajan, *et al.*, 2016)

Mexican American ancestry-informative markers: examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians (Collins-Schramm, *et al.*, 2004)

AncestrySNPminer: A bioinformatics tool to retrieve and develop ancestry informative SNP panels. (Amirisetty, *et al.*, 2013; <https://research.cchmc.org/mershalab/AncestrySNPminer/login.php>.)

Comparing genetic ancestry and self-reported race/ethnicity in a multiethnic population in New York City (Lee, *et al.*, 2012).

Validation of a small set of ancestral informative markers for control of population admixture in African Americans (Ruiz-Narváez, *et al.*, 2011).

A minimum set of ancestry informative markers for determining admixture proportions in a mixed American population: the Brazilian set (Santos, *et al.*, 2016).

Comparing genetic ancestry and self-described race in african americans born in the United States and in Africa (Yaeger, *et al.*, 2008).

Tabla A2. Poblaciones empleadas para la evaluación *in silico* del panel de ancestría.

Población	Abreviación	N
Africana	AFR	661
Yoruba in Ibadan, Nigeria	YRI	108
Luhya in Webuye, Kenya	LWK	99
Gambian in Western Divisions in the Gambia	GWD	113
Mende in Sierra Leone	MSL	85
Esan in Nigeria	ESN	99
Americans of African Ancestry in SW USA	ASW	61
African Caribbeans in Barbados	ACB	96
Americana	AMR	347
Colombianos de Medellín	CLM	94
Mexicanos de Los Ángeles	MXL	64
Puertorriqueños	PUR	104
Peruanos de Lima	PEL	85
Este Asia	EAS	504
Han Chinese in Beijing, China	CHB	103
Japanese in Tokyo, Japan	JPT	104
Southern Han Chinese	CHS	105
Chinese Dai in Xishuangbanna, China	CDX	93
Kinh in Ho Chi Minh City, Vietnam	KHV	99
Europea	EUR	503
Utah Residents (CEPH) with Northern and Western European Ancestry	CEU	99
Toscani in Italia	TSI	107
Finnish in Finland	FIN	99
British in England and Scotland	GBR	91
Iberian Population in Spain	IBS	107
Sur de Asia	SAS	489
Gujarati Indian from Houston, Texas	GIH	103
Punjabi from Lahore, Pakistan	PJL	96
Bengali from Bangladesh	BEB	86
Sri Lankan Tamil from the U	STU	102
<i>Indian Telugu from the UK</i>	ITU	102

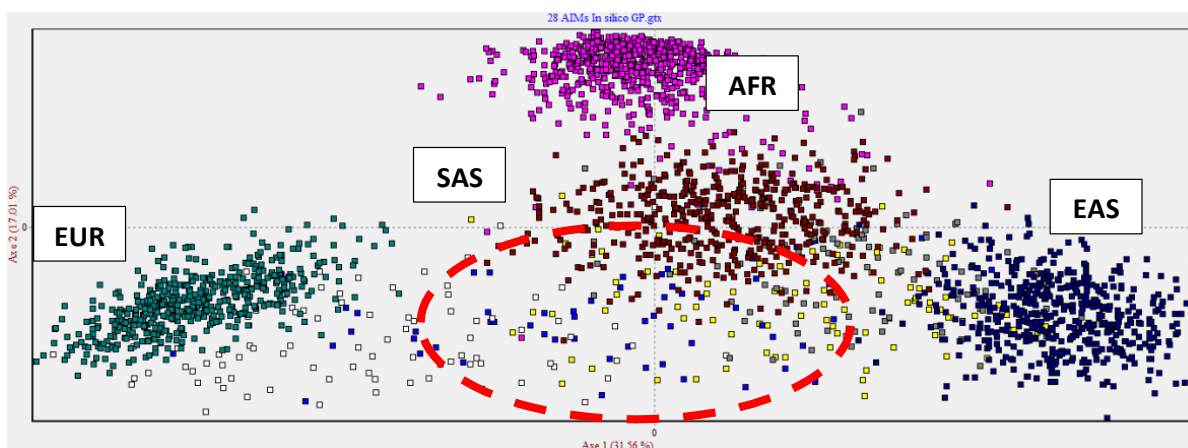
Tabla A3. Panel de Marcadores Informativos de Ancestría para la inferencia de carga ancestral regional en las poblaciones de estudio.

No.	SNP	AIM	Chr	C__
1	rs3737576	Nativo Americano	1	C__27471358_20
2	rs2065160	Asiático	1	C__1648531_10
3	rs1519654	Europeo	2	C__3231730_10
4	rs6754311	Europeo	2	C__11821925_10
5	rs1863086	Europeo	2	C__2862452_10
6	rs1876482	Asiático	2	C__11640969_10
7	rs260690	Europeo	2	C__790944_10
8	rs3827760	Asiático	2	C__27497202_10
9	rs12498138	Nativo Americano	3	C__211052_10
10	rs1344870	Nativo Americano	3	C__8767848_10
11	rs10079352	Africano	5	C__30545512_10
12	rs16891982	Europeo	5	C__2842665_10
13	rs6875659	Africano	5	C__31491320_20
14	rs192655	Asiático	6	C__996147_10
15	rs917115	Europeo	7	C__2621410_10
16	rs2196051	Europeo	8	C__1503192_10
17	rs1408801	Asiático	9	C__1911862_10
18	rs4918664	Asiático	10	C__27950606_20
19	rs2238151	Europeo	12	C__339070_20
20	rs9319336	Asiático	13	C__27328815_20
21	rs7997709	Asiático	13	C__30127919_20
22	rs9522149	Europeo	13	C__30502208_30
23	rs2065982	Asiático	13	C__11713156_10
24	rs730570	Europeo	14	C__2109090_10
25	rs12913832	Europeo	15	C__30724404_10
26	rs1426654	Europeo	15	C__2908190_10
27	rs881929	Asiático	16	C__7473941_10
28	rs3907047	Asiático	20	C__8948366_10

Tabla A4. Variantes del gen de la AS3MT determinadas en las poblaciones de estudio.

No.	SNP	Alelo ancestral	Alelo derivado	C_
1	rs11191439	T	C	C__31979150_10
2	rs3740400	G	T	C__27510174_10
3	rs7085104	A	G	C__3284563_10
4	rs10748835	A	G	C__2852804_10
5	rs3740393	C	G	C__25804287_10

Figura A1. Análisis de componentes factoriales mostrando la distribución de los individuos de las poblaciones de 1000 Genomas. El círculo en rojo señala a los individuos Mexicanos en Los Ángeles (MXL).



Nota: AFR, Africanos, CLM, Colombianos de Medellín, MXL, Mexicanos en Los Ángeles, PEL, Peruanos de Lima, PUR, Puertorriqueños, EAS, Asiáticos del Este, EUR, Europeos, SAS, Asiáticos

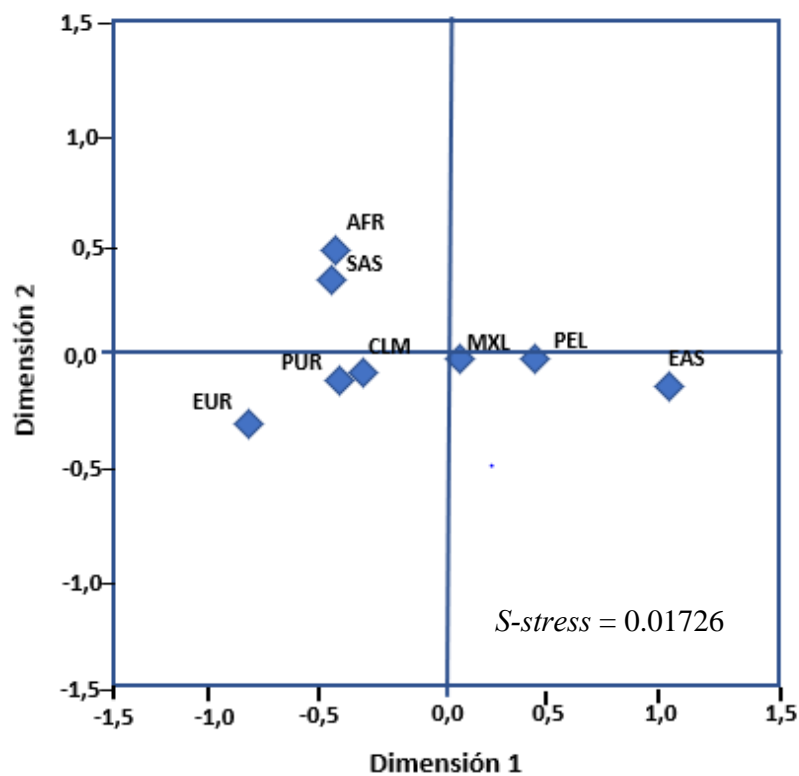
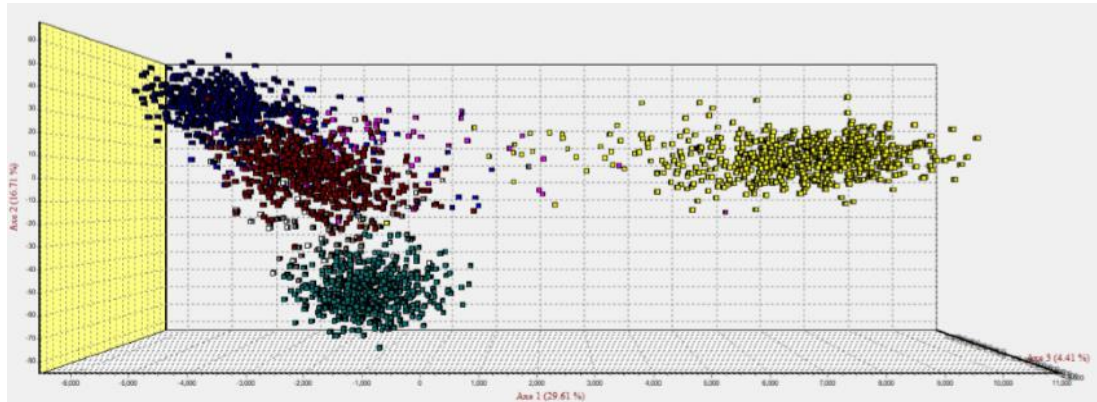


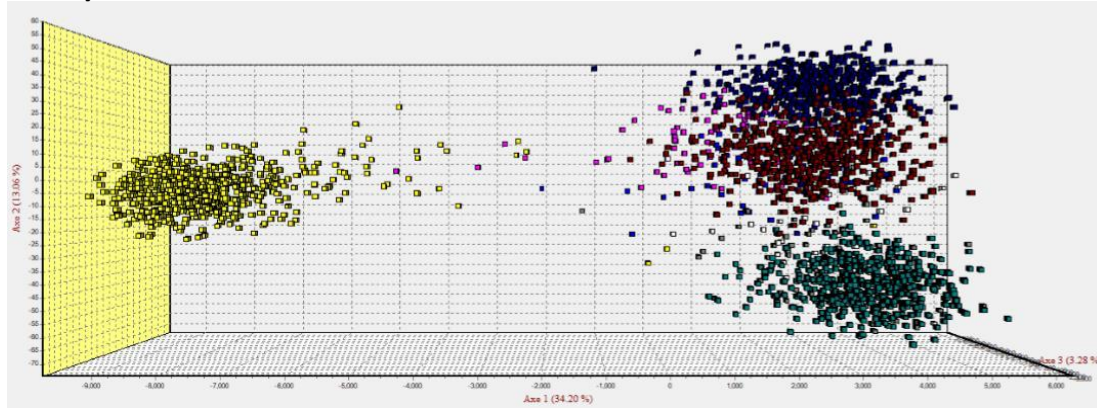
Figura A2. Análisis Multidimensional a Escala de la evaluación *in silico* con 28 Marcadores Informativos de Ancestría y las poblaciones de 1000Genomas.

Nota: AFR, Africanos, CLM, Colombianos de Medellín, MXL, Mexicanos en Los Ángeles, PEL, Peruanos de Lima, PUR, Puertorriqueños, EAS, Asiáticos del Este, EUR, Europeos, SAS, Asiáticos del Sur.

A) 34 AIMs



B) 36 AIMs



C) 48 AIMs

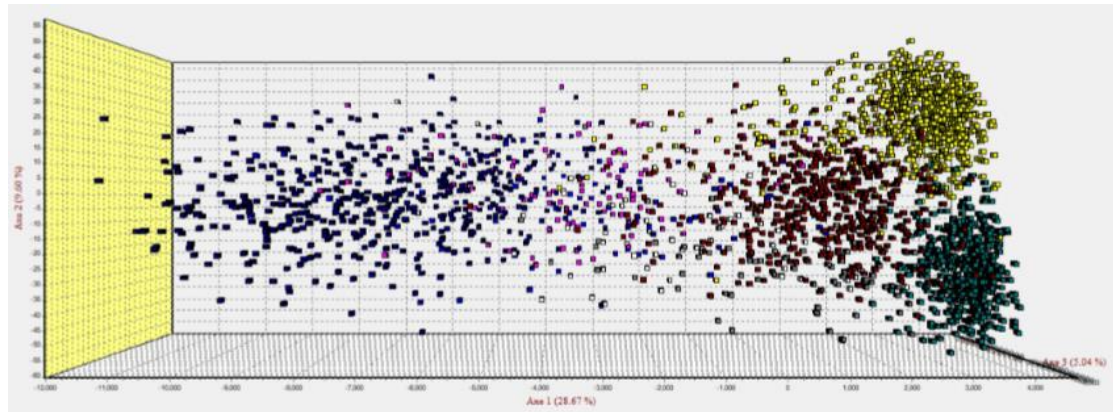


Figura A3. Análisis de componentes factoriales mostrando la distribución de los individuos de las poblaciones de 1000 Genomas con 34, 36 y 48 AIMs reportados (A) [Kosoy, et al., 2009](#); B) <http://spsmart.cesga.es/snpforid.php>; C) <https://alfred.med.yale.edu/>)