Int J Mol Epidemiol Genet 2022;13(3):32-41
www.ijmeg.org /ISSN:1948-1756/IJMEG0145131

Original Article

Pilot study: genetic distribution of AR, FGF5, SULT1A1 and CYP3A5 polymorphisms in male Mexican population with androgenetic alopecia

Daniela Martinez-Chapoy[1*], Francisco J Cruz-Arroyo[1*], Francisco D Ancer-Leal[1*], Regina A Rodriguez-Leal[1*], Bianka D Camacho-Zamora[1*], Daniela A Guzman-Sanchez[2*], Nelly A Espinoza-Gonzalez[3*], Lizeth Martinez-Jacobo[4*], Ivan A Marino-Martinez[4*]

[1]OmicronLab S.A. de C.V., Monterrey, Nuevo Leon, Mexico; [2]Dermatology Clinic, Guadalajara, Jalisco, Mexico; [3]Cumbres Monterrey Dermatology Center, Monterrey, Nuevo Leon, Mexico; [4]Universidad Autonoma de Nuevo Leon, Monterrey, Nuevo Leon, Mexico. [*]Equal contributors.

Received July 4, 2022; Accepted November 17, 2022; Epub December 15, 2022; Published December 30, 2022

Abstract: Genetics is responsible for 80% of androgenetic alopecia (AGA) predisposition. Several single nucleotide polymorphisms (SNPs) have been linked to AGA risk and the metabolism of its first-line therapies. Genotypic and allelic frequencies have not been described in Mexican individuals; therefore, the aim of this study was to describe the genetic distribution of SNPs associated with AGA predisposition and drug metabolism. Using Real Time-PCR, we genotyped SNPs rs4827528 (*AR*), rs7680591 (*FGF*5), rs1042028, rs1042157, rs788068 and rs6839 (*SULT*1*A*1) and rs776746 (*CYP*3*A*5) in 125 (controls = 60, cases = 65) male volunteers from Northern and Western Mexico. The *SULT*1*A*1 SNPs rs1042028 (C/T) and rs788068 (T/A/C) resulted in a 100% distribution of the ancestral allele C and mutated allele A, respectively; rs1042028 diverges from the previously reported frequency, while the rs788068 ancestral allele was found to be more predominant than the reported frequency. Rs1042028, rs788068 and rs4827528, were not in Hardy-Weinberg (HW) equilibrium; conversely, rs1042157 and rs6839, rs776746, and rs7680591 followed HW principles. A statistically significant difference ($P<0.05$) was obtained for the rs1042157 allelic frequency between cases and controls in Western Mexico. We reported the genotypic and allelic frequencies of seven polymorphisms in Mexican individuals from Northern and Western Mexico.

Keywords: Androgenetic alopecia, Mexican population, genotyping, genetic distribution, predisposition, drug metabolism

Introduction

Androgenetic alopecia (AGA) is an inheritable trait; it is the most frequent cause of hair loss in males and females [1]. In males, there is a typical pattern of progressive hair loss located in the frontal area and up to the vertex; female pattern hair loss presents with diffuse thinning, but the frontal hairline is preserved [2]. AGA prevalence depends on age and ethnicity, *e.g.* Caucasian males have a 50% chance of developing AGA by the age of 50, while by the age of 70 there is an 80% chance, however in Asian and African American populations, the probability of developing AGA is lower [3]. AGA severity is determined using the Hamilton-Norwood (HN) Scale for Male Pattern Baldness [2], or the Sinclair Hair Shedding Scale for Female Pattern Hair Loss [4]. Other diagnostic methods include dermatoscopy [5], scalp biopsy and trichogram, some of which are invasive techniques [6].

AGA is multifactorial and it is likely caused by hormonal (5-alpha-dihydrotestorone, DHT, binds to androgen receptors, which causes hair follicle miniaturization and shortening of the anagen phase of the hair cycle), lifestyle (daily habits, hair product usage, diet, etc.) and genetic factors [7]. Genetics is the most important factor, determining 80% of AGA predisposition [8].

AGA predisposition has been associated with numerous genes, including androgen receptor

gene (*AR*), the first one to be associated with AGA, and fibroblast growth factor 5 (*FGF*5) [7, 9-11]. *FGF*5 has been established as a predisposition gene because of its involvement in anagen phase regulation [9, 10]. The sulfotransferase family 1A member 1 (*SULT*1*A*1), and cytochrome P450 family 3 subfamily A member 5 (*CYP*3*A*5) genes code for proteins that regulate the metabolism of several first-line therapies used in AGA [12-14].

Several *AR* single nucleotide polymorphisms (SNPs) have been identified, including rs4827528, rs5919393 and rs6152. Rs4827528 (A/G/T/C) has the closest association with AGA (OR = 3.4) [10]. Androgen receptors have a fundamental role in the development of AGA as excessive androgen binding (*e.g.* testosterone and DHT) causes hair follicle miniaturization. Fibroblast growth factor 5 (FGF5) is an anagen phase inhibitor that promotes the transition from anagen to catagen phase. Furthermore, it regulates cellular proliferation and differentiation. Currently, only one *FGF*5 polymorphism, rs7680591 (T/A/G), shows a direct correlation with AGA. This marker increases gene expression and results in a shortened anagen phase [9, 10].

S*ULT*1*A*1 codes for a catalytic enzyme responsible for the bioactivation of xenobiotics, endogenous compounds, drugs, hormones and neurotransmitters. Bioactivation is carried out by the addition of a sulfonate group to a hydroxyl or amino group. Minoxidil, one of the first line treatment for AGA, is a *SULT*1*A*1 substrate [14]. SNPs like rs1042157 (G/A/C), rs6839 (T/C), rs788068 (T/A/C) and rs10420-28 or *SULT*1*A*1*2 (C/T) have been associated with decreased enzymatic activity [13, 15, 16].

*CYP*3*A*5 codes for a cytochrome P450 member, which is a monooxygenase that catalyzes a range of reactions involved in drug metabolism (*e.g.*finasteride), cholesterol, steroid, and other lipid syntheses [17]. SNPs such as rs15524 and rs776746 alter the enzyme function, also, both polymorphisms are related to a higher finasteride concentration, still the rs776746 or *CYP*3*A*5*3 (T/C) produces a cryptic splice site and the premature termination of the enzyme, thus the *CYP*3*A*5 expression is lost [12]. In this case, genotype C/C increases finasteride concentration, and therefore, it is a 5-alpha-reductase (responsible for the testosterone to DHT conversion) inhibitor.

Currently, there are no genotyping studies in native Mexican population related with AGA (predisposition and drug metabolism). The latter is relevant to prescribe a precise treatment to the AGA patient. So, the aim of this investigation was to evaluate the genetic distribution of AGA related polymorphisms rs4827528 (*AR*), rs7680591 (*FGF*5), rs1042028, rs10-42157, rs788068 and rs6839 (*SULT*1*A*1) and rs776746 (*CYP*3*A*5) in male volunteers from Northern and Western Mexico.

Materials and methods

Ethical statement

This protocol has been approved (register number CIDICIS/I-0127/20) by the Ethics, Research and Biosafety Committees of the Research and Development Center in Health Sciences (Centro de Investigación y Desarrollo en Ciencias de la Salud, CIDICS) of the Universidad Autonoma de Nuevo Leon (UANL) in Monterrey, Nuevo Leon, Mexico. In agreement with the pertinent research and health legislation (Reglamento de la Ley General de Salud), second title, chapter 1, article 17, section II, this kind of study is classified with the minimum risk. Therefore, regulatory approval by the Mexican Health Ministry is not required. This protocol has been conducted in agreement with the Good Clinical Practices Guidelines of the International Conference on Harmonization (ICH- E6-R1), the Mexican norms NOM-012-SSA3-2012 and NOM-220-SSA1-2016, and the current research and health legislation and other legal dispositions that applied. The volunteers' name and data are confidential, and they are kept as professional secret. All volunteers understood, accepted, and provided informed consent. The results, identity and medical records of the volunteers are protected according to the NOM-168-SSA1-1998.

Patients and samples

From June 2021 to September 2021, we recruited 125 male volunteers in Monterrey, Nuevo Leon and Guadalajara, Jalisco, México.

Inclusion criteria: All participants were from 18 to 65 years old. The volunteers were divided in two groups in each region (North and West): 1. Control group (volunteers without AGA) and 2. AGA or cases group (volunteers with AGA between category II and IV of the Hamilton-

Table 1. rhAmp SNP Assay design IDs (Integrated DNA Technologies, IDT)

Gene	SNP	Design ID
AR	Rs4827528	Hs.GT.rs4827528.G.1
FGF5	rs7680591	Hs.GT.rs7680591.A.1
CYP3A5	rs776746	Hs.ADME.rs776746.C.1
SULT1A1	rs1042028	CD.GT.CGFP1420.1
	rs1042157	CD.GT.NQFT9529.1
	rs6839	CD.GT.NTVL6110.1
	rs788068	CD.GT.VXLQ2918.1

SNP: Single Nucleotide Polymorphism.

Norwood Scale) (Hamilton, 1951). Both groups included participants whose parents, at least one paternal grandparent, and at least one maternal grandparent were born in Northern (Chihuahua, Coahuila, Nuevo Leon, Sonora, Tamaulipas, Durango, Baja California, Baja California Sur) or Western (Sinaloa, Nayarit, Aguascalientes, Jalisco, Colima, Michoacán, Guanajuato) Mexico.

Exclusion criteria: There were no female volunteers in this study. All volunteers, controls and cases, were exempt of a clinical diagnosis of alopecia areata, trichotillomania, traction alopecia, scarring alopecia, telogen effluvium, discoid or systemic erythematosus lupus, and thyroid disease present during the study or without control (at least 6 months of treatment with an endocrinologist), malignant neoplasm in the last 5 years (except non melanoma skin cancer). We did not recruit volunteers under hormonal replacement therapy or with any hormonal condition with a direct impact in the hair growth cycle, with a BMI <22 or with ancestors born in any other state than the ones mentioned.

Saliva samples were collected from the volunteers using a non-invasive saliva collection kit, Oragene DNA-OG 500 (DNA Genotek, Canada). DNA was extracted to determine the presence or absence of the mentioned SNPs. The genotyping study length was 11 months, including the recruitment, processing and analysis phases.

Hamilton-norwood evaluation

All patients were diagnosed and assigned to a control or case group by a certified dermatologist using the Hamilton-Norwood scale [2] as follows: *Type II:* The anterior border of the hairline in the frontoparietal regions has triangular areas of recession, which tend to be symmetrical and extend no farther posteriorly than a point of 3 cm, anterior to a line drawn in a coronal plane between the external auditory meatuses. Hair is also lost, or sparse, along the midfrontal border of the scalp, but the depth of the affected area is much less than in the frontoparietal regions.

Type III: Borderline cases were listed separately as Type III, which also included scalps in which classification is rendered inaccurately due to scars, lateral asymmetry in denudation, unusual types of sparseness and thinning of the hair, and other factors.

Type IV: It represents the minimal hair loss considered sufficient to represent baldness. There are deep frontotemporal recessions, usually symmetrical, and they are either bare or covered very sparsely by hair. These recessions extend farther posteriorly than a point, which lies 3 cm anterior to a coronal line drawn between the external auditory meatuses. If hair is sparse or lacking as a broad band along the entire anterior border of the hairline, it is classified as Type IVA.

rhApm SNP assay genotyping

For the DNA extraction and purification from the saliva samples the PureLink™ Genomic DNA Mini Kit (Thermo Fisher Scientific, USA) was used, following the supplier instructions (total volume after elution was 50 microliters, µL). Quantification was performed using the Qubit dsADN HS (Thermo Fisher Scientific, USA) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific, USA). For each reaction 2 µL of the sample and 198 µL of the solution stock (1 µL Qubit dsDNA HS Reagent + 199 µL Qubit dsDNA HS Buffer) were added.

Genotyping was performed with the rhAmp SNP Assay (Integrated DNA Technologies, IDT, USA) (**Table 1**) and the Real-time Polymerase Chain Reaction (qPCR) technique in a CFX96 Touch-Real Time PCR Detection System (Bio-Rad Laboratories, USA). As positive controls, gBlocks® Gene Fragments (Integrated DNA Technologies, IDT, USA) were used. Reaction components were stock solution 5.3 µL, rhAmp SNP Assay (20X) 0.05 µL, nuclease-free water 4.2 µL with a final volume of 10 µL. Thermocycling conditions were followed according to manufacturer.

Table 2. Control and cases group age distribution, mean BMI and volunteers per HW scale groups

Region	Control Group			Cases (AGA Group)		
	North	West	Total	North	West	Total
n (%)	30 (50.0%)	30 (50.0%)	60	32 (49.23%)	33 (50.77%)	65
Mean age (years)	31.5	28.23	29.87	37.81	31.72	34.77
n age groups (years)						
18-30	18 (30.0%)	22 (36.7%)	40 (66.7%)	7 (10.8%)	18 (27.7%)	25 (38.5%)
31-43	6 (10.0%)	6 (10.0%)	12 (20.0%)	15 (23.1%)	12 (18.5%)	27 (41.6%)
44-56	5 (8.3%)	2 (3.3%)	7 (11.7%)	9 (13.9%)	3 (4.6%)	12 (18.5%)
57-65	1 (1.7%)	0 (0.0%)	1 (1.7%)	1 (1.6%)	0 (0.0%)	1 (1.6%)
Mean BMI	26.2	28.61	27.41	28.23	26.03	27.13
n HN scale (%)						
II	n/a	n/a	n/a	5 (27.78%)	13 (72.22%)	18
III	n/a	n/a	n/a	14 (46.67%)	16 (53.33%)	30
IV	n/a	n/a	n/a	13 (76.47%)	4 (23.53%)	17

AGA: Androgenetic Alopecia; BMI: Body Mass Index; HN: Hamilton-Norwood; n/a: Not Applicable.

Statistical analysis

To statistically analyze the study results, Hardy-Weinberg Equilibrium (HW, $1 = p^2 + 2pq + q^2$) and square chi ($\chi^2 = \Sigma [(O-E)\ 2/E]$) were applied. The HW was used to determine if the populations were in equilibrium (*p-value* <0.05). The non-parametric square chi analysis was used to analyze the allelic and genotypic frequencies. Results were considered statistically significant with *p-value* <0.05. IBM SPSS package (v 2.0, IBM USA) was used to perform the analysis.

Results

Patients and samples

This case-control study included 125 volunteers (65 cases and 60 controls). The mean age was 29.87 years for the controls and 34.77 years for the cases, being 18-33 the group age with the highest frequency in the controls (66.7%) and the 31-43 in the cases group (41.6%). According to the HN scale, the 53.5% of volunteers within the cases presented a low moderate AGA (III). Control and cases group age distribution, mean BMI and volunteers per HW scale groups are shown in **Table 2**.

Genotypic and allelic frequencies

Considering the whole population (North/West) without separating by region, genotypic frequencies (Gf) were obtained (**Table 3**) for each SNP. In both groups, the polymorphisms rs1042028 and rs788068 (*SULT1A1*) got a 100% distribution of genotypes C/C and A/A, respectively. The rs4827528 (*AR*) presented a Gf of 0.90 for genotype G/G in controls and cases. Furthermore, genotypes G/A of rs1042157 and T/C of rs6839 (*SULT1A1*) and T/A of rs7680591 (*FGF5*) were predominant with 0.47, 0.45 and 0.53, respectively for controls and 0.61, 0.46 and 0.44, respectively for cases. Among both groups, SNP rs776746 (*CYP3A5*) there was a predominance for recessive genotype C/C.

Allelic frequencies (Af) were calculated for each SNP (**Table 4**). The ancestral allele for *SULT1A1* polymorphisms rs1042028 (C), rs1042157 (G), rs6839 (T) and rs788068 (T) was found in 100% (120/120, Af = 1.00), 45% (54/120, Af = 0.45), 51% (61/120, Af = 0.51) and 0% (0/120, Af = 0.00), respectively in the controls and 100% (130/130, Af = 1.00), 52% (67/130, Af = 0.52), 44% (58/130, Af = 0.44) and 0% (0/130, Af = 0.00) in the cases. In contrast, the mutated allele (T, A, C and A respectively) had a distribution of 0% (0/120, Af = 0.00), 55% (66/120, Af = 0.55), 49% (59/120, Af = 0.49) and 100% (120/120, Af = 1.00) in the controls and 0% (0/130, Af = 0.00), 48% (63/130, Af = 0.48), 55% (72/130, Af = 0.55) and 100% (130/130, Af = 1.00) in the AGA group. Ancestral alleles T, T and A were identified in 25% (30/120, Af = 0.25), 53% (64/120, Af = 0.53) and 10% (12/120, Af = 0.10) of the controls, and 18% (23/130, Af = 0.18), 43% (56/130, Af = 0.43) and 10% (14/130, Af = 0.10) of the cases for rs776746, 7680591 and rs4827528 polymorphisms; while mutated alleles C, A and G were

Table 3. Volunteers' genotypic frequencies showed by polymorphisms, group and region with their HW equilibrium *p-value*

SNP	Group	Region	n (Genotypic frequencies)			HW equilibrium
			AA	Aa	aa	*p*-value
rs1042028	Control	North	30 (0.50)	0 (0.0)	0 (0.0)	n/a
		West	30 (0.50)	0 (0.0)	0 (0.0)	n/a
		North/West	60 (1.00)	0 (0.0)	0 (0.0)	n/a
	Cases	North	32 (0.50)	0 (0.0)	0 (0.0)	n/a
		West	33 (0.50)	0 (0.0)	0 (0.0)	n/a
		North/West	65 (1.00)	0 (0.0)	0 (0.0)	n/a
rs1042157	Control	North	8 (0.13)	13 (0.22)	9 (0.15)	0.468
		West	5 (0.08)	15 (0.25)	10 (0.17)	0.875
		North/West	13 (0.21)	28 (0.47)	19 (0.32)	0.657
	Cases	North	4 (0.06)	19 (0.30)	9 (0.14)	0.219
		West	10 (0.15)	20 (0.31)	3 (0.04)	0.121
		North/West	14 (0.21)	39 (0.61)	12 (0.18)	0.104
rs6839	Control	North	6 (0.10)	18 (0.30)	6 (0.10)	0.273
		West	11 (0.18)	9 (0.15)	10 (0.17)	0.028
		North/West	17 (0.28)	27 (0.45)	16 (0.27)	0.439
	Cases	North	3 (0.04)	15 (0.23)	14 (0.22)	0.721
		West	11 (0.17)	15 (0.23)	7 (0.11)	0.656
		North/West	14 (0.21)	30 (0.46)	21 (0.33)	0.596
rs788068	Control	North	0 (0.0)	0 (0.0)	30 (0.5)	n/a
		West	0 (0.0)	0 (0.0)	30 (0.5)	n/a
		North/West	0 (0.0)	0 (0.0)	60 (1.00)	n/a
	Cases	North	0 (0.0)	0 (0.0)	32 (0.5)	n/a
		West	0 (0.0)	0 (0.0)	33 (0.5)	n/a
		North/West	0 (0.0)	0 (0.0)	65 (1.00)	n/a
rs776746	Control	North	3 (0.05)	11 (0.18)	16 (0.27)	0.594
		West	2 (0.03)	9 (0.15)	19 (0.32)	0.524
		North/West	5 (0.08)	20 (0.33)	35 (0.59)	0.389
	Cases	North	0 (0.0)	15 (0.23)	17 (0.26)	0.083
		West	0 (0.0)	8 (0.12)	25 (0.39)	0.428
		North/West	0 (0.0)	23 (0.35)	42 (0.65)	0.083
rs7680591	Control	North	6 (0.10)	18 (0.30)	6 (0.10)	0.273
		West	10 (0.17)	14 (0.23)	6 (0.10)	0.785
		North/West	16 (0.27)	32 (0.53)	12 (0.20)	0.580
	Cases	North	3 (0.4)	16 (0.25)	13 (0.20)	0.540
		West	11 (0.17)	12 (0.19)	10 (0.15)	0.118
		North/West	14 (0.21)	28 (0.44)	23 (0.35)	0.326
rs4827528	Control	North	3 (0.05)	0 (0.0)	27 (0.45)	<0.001*
		West	3 (0.05)	0 (0.0)	27 (0.45)	<0.001*
		North/West	6 (0.10)	0 (0.0)	54 (0.90)	<0.001*
	Cases	North	3 (0.04)	0 (0.0)	29 (0.45)	<0.001*
		West	4 (0.06)	0 (0.0)	29 (0.45)	<0.001*
		North/West	7 (0.10)	0 (0.0)	58 (0.90)	<0.001*

SNP: Single Nucleotide Polymorphism; HN: Hamilton-Norwood; A: Ancestral allele; a: Mutated allele; n/a: Not Applicable. *Significant result (*p*-value <0.05).

Table 4. Volunteers' allelic frequencies showed by polymorphisms, group and region with their *p-value*

SNP	Group	Allelic frequencies per region					
		North		West		North/West	
		A	a	A	a	A	a
rs1042028	Control	60 (0.50)	0 (0.0)	60 (0.50)	0 (0.0)	120 (1.00)	0 (0.0)
	Cases	64 (0.49)	0 (0.0)	66 (0.51)	0 (0.0)	130 (1.0)	0 (0.0)
	p-value	n/a		n/a		n/a	
rs1042157	Control	29 (0.24)	31 (0.26)	25 (0.21)	35 (0.29)	54 (0.45)	66 (0.55)
	Cases	27 (0.21)	37 (0.28)	40 (0.31)	26 (0.20)	67 (0.52)	63 (0.48)
	p-value	0.4919		0.0336*		0.3014	
rs6839	Control	30 (0.25)	30 (0.25)	31 (0.26)	29 (0.24)	61 (0.51)	59 (0.49)
	Cases	21 (0.16)	43 (0.33)	37 (0.28)	29 (0.22)	58 (0.44)	72 (0.55)
	p-value	0.0519		0.6212		0.3254	
rs788068	Control	0 (0.0)	60 (0.50)	0 (0.0)	60 (0.50)	0 (0.0)	120 (1.00)
	Cases	0 (0.0)	64 (0.49)	0 (0.0)	66 (0.51)	0 (0.0)	130 (1.00)
	p-value	n/a		n/a		n/a	
rs776746	Control	17 (0.14)	43 (0.36)	13 (0.11)	47 (0.39)	30 (0.25)	90 (0.75)
	Cases	15 (0.12)	49 (0.38)	8 (0.06)	58 (0.44)	23 (0.18)	107 (0.82)
	p-value	0.5335		0.1510		0.1578	
rs7680591	Control	30 (0.25)	30 (0.25)	34 (0.28)	26 (0.22)	64 (0.53)	56 (0.47)
	Cases	22 (0.17)	42 (0.32)	34 (0.26)	32 (0.25)	56 (0.43)	74 (0.57)
	p-value	0.0780		0.5623		0.1049	
rs487528	Control	6 (0.05)	54 (0.45)	6 (0.05)	54 (0.45)	12 (0.10)	108 (0.90)
	Cases	6 (0.04)	58 (0.45)	8 (0.06)	58 (0.45)	14 (0.10)	116 (0.90)
	p-value	0.9064		0.6150		0.8422	

SNP: Single Nucleotide Polymorphism; A: Ancestral allele; a: Mutated allele; n/a: Not Applicable. *Significant result (*p*-value <0.05).

reported respectively in 75.0% (90/120, Af = 0.75), 47% (50/120, Af = 0.47) and 90% (108/120, Af = 0.90) in the controls, and 82% (107/130, Af = 0.82), 57% (74/130, Af = 0.57) and 90% (116/120, Af = 0.90) in the AGA group.

A statistical analysis was performed to establish if the behavior of the alleles and genotypes complied with the HW equilibrium in the controls. It was observed that rs1042157 (χ^2 = 0.196, P = 0.657) and rs6839 (χ^2 = 0.597, *p-value* = 0.439) of *SULT1A1*, rs776746 (χ^2 = 0.740, *p-value* = 0.389) of *CYP3A5* and rs7680591 (χ^2 = 0.306, *p-value* = 0.580) of *FGF5*, were in HW equilibrium, considering the population as a whole. The rs1042028 and rs788068 of *SULT1A1*, and rs4827528 (χ^2 = 61, *p-value* <0.001) of *AR*, did not show any behavior according to what is postulated by the HW law.

Group comparative analysis

Using the χ^2 test, a comparative analysis of the genotypic and allelic frequencies of the seven SNPs under evaluation, was performed between the cases and the control groups. The comparison between groups was made in both ways: 1) without considering geographic areas and 2) separating the total population by geographic regions. The comparative analysis when considering the population as one, did not show significant differences between the allelic and genotypic distribution of the cases and the control groups (*p-value*s >0.05). However, when making the comparison considering the regions separately, there was a statistic difference in the West volunteers between the cases and the controls when the allelic frequency was analyzed, corresponding to rs1042157 (*SULT1A1*) (*P*<0.05).

Discussion and conclusion

Case-control studies have been used in genetic investigations to evaluate the genetic predisposition to human conditions and identify SNPs that can help to develop better prognosis and treatments [22]. To elucidate the influence of the genetic factor in AGA development, many

researchers have done the analysis of genes associated to processes involved in the AGA pathogeny, e.g. *AR* gene that has been identified to take part in the sexual steroids route [18], besides, evidence of high *AR* expression levels in AGA patients have been found [19-21]. The current study aimed to report the allelic and genotypic frequencies of seven SNPs possibly associated with AGA development and drug metabolism in Mexican population.

From the seven included SNPs, rs4827528 of *AR* gene and rs7680591 of *FGF*5 gene are recognized as AGA predisposition polymorphisms; while rs1042028, rs1042157, rs788068, and rs6839 of *SULT1A*1 gene, and rs775746 of *CYP3A*5 gene, may be associated to the drug metabolism used in AGA treatment.

Even though the Af are consistent with the ones reported in LA Mexican residents (California, U.S.A), Latin-American and American population, depending on the case, the *SULT1A*1 SNPs rs1042028 (C/T) and rs788068 (T/A/C) showed a 100% distribution of the ancestral allele C and mutated allele A, respectively; rs1042028 diverges from the previous reported frequency (100% for the mutated allele) [23], while the rs788068 ancestral allele was found to be more predominant than the reported frequency (45% for allele T) [24]. Neither rs1042028 nor rs788068 SNPs of *SULT1A*1 gene, plus the *AR* polymorphism rs4827528, were in HW equilibrium; conversely, rs1042157 and rs6839 of *SULT1A*1, rs776746 of *CYP3A*5 and rs7680591 of *FGF*5 followed the HW principles.

HW equilibrium is used as a quality control to detect genotyping errors in genetic studies, however disequilibrium (HWD) could occur for other reasons like chance, genome location, HW principles failure or the genetics of the condition per se [25]. Besides, the HWD can be expected in case-control studies in specific locus and small samples [26]. We believe that, it is the case of *AR* SNP (rs4827528), which is located in chromosome X and its distribution is different from autosomal SNPs [27]. While in other studies, the *AR* SNP has been found to be in HW equilibrium [7, 9, 10], in this protocol the HWD can be the result of the limited sample size and the SNP location in chromosome X.

Despite the vast SNP diversity analyzed and the limited sample size, we could identify a significant difference (p-value = 0.024) in Af (mutated allele A) of rs1042157 (*SULT1A*1) between cases and controls in the Western region. Nevertheless, interestingly, when we analyzed the full population (North and West together) and Northern region alone, we could not identify that difference. It may be explained by the population's racial composition in different geographic regions of Mexico.

Racial ancestry studies that have been conducted in Mexico tend to cluster the Northern and Western regions as one due to the similar ancestry proportions (~60-80% European, ~21-25% Amerindian, ~13-20% Eurasian and ~1-9% African) [28-30]. All studies concluded that, European ancestry is predominant in Northern and Western population, but there are significant differences in Amerindian and African ancestry percentages, e.g. 5% and 12% Amerindian ancestry for Northern and Western region, respectively [28].

Different results could be explained by the markers used in the study, their location and sample size. Regarding the markers and their genomic location, mitochondrial, autosomal and Y-linked markers elucidate the admixture bias resulting from sex. In Nuevo Leon state (Northern region), the main Y-chromosome contribution comes predominantly from European males (78%) [31]; in contrast, the maternal counterpart (mitochondrial DNA) had Amerindian predominance in Northern (89.1%) and Western (82%) Mexico [32]. It was previously stated that the paternal and maternal proportions are due to the fact that it was common for male Europeans and female Native Americans to breed, and not otherwise [30]. In addition, the use of autosomal markers produces different ancestry proportions; A study reported 56% Amerindian and 38% European ancestry in Nuevo Leon state [31], similar to the data published in another study for the northern state of Coahuila, where the estimated Amerindian ancestry was 55.6% [33]. Opposite results could be due to the genetic intrapopulation structure, which includes socioeconomic and educational factors as demonstrated in Mexico City population [32] and ethnic groups located in each region, like Chichimecas in the North [34] and Huicholes in the West

[29]. Our results, considering the analysis with one chromosome X, and six autosomal markers, and the regions chosen for the study, could be explained by the aforementioned data: the lack of significant differences when analyzing the North population alone, and the significant difference between cases and controls in the Af of Western Mexico.

Even though *SULT1A1* is not a gene that relates to biosynthetic routes or response to sexual steroids, the fact that we reported the significantly different polymorphism between cases and controls is relevant because of its participation in the minoxidil conversion to its bioactive form [14]. Certain SNPs could negatively affect the sulfotransferase activity of SULT1A1 enzyme, which could directly affect the AGA treatment response to minoxidil, being its active form conjugated to a sulfate group.

As we have recognized along the manuscript, the main shortcoming of the present study was the sample size. Although this was a pilot study, the number of cases and controls recruited allowed us to perform a full statistical analysis. However, we do not discard to increase sample size avoiding HWD and facilitating the observation of statistical differences which remain as tendences in the present study. For example, *AR* gene SNP was expected to show differences between study groups and only showed tendences ($P = 0.05$), although we must take into consideration that most cases were younger than 50 years old, so it's possible that they could be genetically predisposed to AGA but have not clinically present it yet.

The present study reported the genotypic and allelic frequencies of seven polymorphisms before mentioned in Northern and Western Mexican population, and the possible association for rs1042157 (*SULT1A1*). Now, it is necessary to increase sample size to investigate the possible association of the SNPs with the genetic AGA predisposition. And to perform a new protocol to study the possible association with the response to treatment using first line therapeutic drugs for AGA (Minoxidil and Finasteride) with the intention of personalizing AGA's treatment in Mexican population.

Acknowledgements

Special thanks to OmicronLab S.A. de C.V. for their unwavering support and numerous contributions.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Ivan A Marino-Martinez, Universidad Autonoma de Nuevo Leon, Monterrey, Nuevo Leon, Mexico. Tel: +52-81-1387-4103; E-mail: amarinomtz@gmail.com

References

[1] Heilmann-Heimbach S, Hochfeld LM, Paus R and Nöthen MM. Hunting the genes in male-pattern alopecia: how important are they, how close are we and what will they tell us? Exp Dermatol 2016; 25: 251-257.

[2] Hamilton JB. Patterned loss of hair in man: types and incidence. Ann N Y Acad Sci 1951; 53: 708-28.

[3] Lolli F, Pallotti F, Rossi A, Fortuna MC, Caro G, Lenzi A, Sansone A and Lombardo F. Androgenetic alopecia: a review. Endocrine 2017; 57: 9-17.

[4] Sinclair R, Jolley D, Mallari R and Magee J. The reliability of horizontally sectioned scalp biopsies in the diagnosis of chronic diffuse telogen hair loss in women. J Am Acad Dermatol 2004; 51: 189-199.

[5] Blume-Peytavi U and Vogt A. Current standards in the diagnostics and therapy of hair diseases-hair consultation. J Dtsch Dermatol Ges 2011; 9: 394-410; quiz 411-2.

[6] Vujovic A and Del Marmol V. The female pattern hair loss: review of etiopathogenesis and diagnosis. BioMed Res Int 2014; 2014: 767628.

[7] Heilmann-Heimbach S, Herold C, Hochfeld LM, Hillmer AM, Nyholt DR, Hecker J, Javed A, Chew EG, Pechlivanis S, Drichel D, Heng XT, Del Rosario RC, Fier HL, Paus R, Rueedi R, Galesloot TE, Moebus S, Anhalt T, Prabhakar S, Li R, Kanoni S, Papanikolaou G, Kutalik Z, Deloukas P, Philpott MP, Waeber G, Spector TD, Vollenweider P, Kiemeney LA, Dedoussis G, Richards JB, Nothnagel M, Martin NG, Becker T, Hinds DA and Nöthen MM. Meta-analysis identifies novel risk loci and yields systematic insights into the biology of male-pattern baldness. Nat Commun 2017; 8: 14694.

[8] Nyholt DR, Gillespie NA, Heath AC and Martin NG. Genetic basis of male pattern baldness. In J Invest Dermatol 2003; 121: 1561-1564.

[9] Hagenaars SP, Hill WD, Harris SE, Ritchie SJ, Davies G, Liewald DC, Gale CR, Porteous DJ, Deary IJ and Marioni RE. Genetic prediction of male pattern baldness. PLoS Genet 2017; 13: e1006594

[10] Pirastu N, Joshi PK, de Vries PS, Cornelis MC, McKeigue PM, Keum N, Franceschini N, Colombo M, Giovannucci EL, Spiliopoulou A, Fran-

ke L, North KE, Kraft P, Morrison AC, Esko T and Wilson JF. GWAS for male-pattern baldness identifies 71 susceptibility loci explaining 38% of the risk. Nat Commun 2017; 8: 1584

[11] Yap CX, Sidorenko J, Wu Y, Kemper KE, Yang J, Wray NR, Robinson MR and Visscher PM. Dissection of genetic variation and evidence for pleiotropy in male pattern baldness. Nat Commun 2018; 9: 5407

[12] Chau CH, Price DK, Till C, Goodman PJ, Chen X, Leach RJ, Johnson-Pais TL, Hsing AW, Hoque A, Tangen CM, Chu L, Parnes HL, Schenk JM, Reichardt JK, Thompson IM and Figg WD. Finasteride concentrations and prostate cancer risk: results from the prostate cancer prevention trial. PLoS One 2015; 10: e0126672.

[13] Daniels J and Kadlubar S. Pharmacogenetics of SULT1A1. Pharmacogenomics 2014; 1514: 1823-1838.

[14] Gamage N, Barnett A, Hempel N, Duggleby RG, Windmill KF, Martin JL and McManus ME. Human sulfotransferases and their role in chemical metabolism. Toxicol Sci 2006; 90: 5-22.

[15] Nagar S, Walther S and Blanchard RL. Sulfotransferase (SULT) 1A1 polymorphic variants *1, *2, and *3 are associated with altered enzymatic activity, cellular phenotype, and protein degradation. Mol Pharmacol 2006; 69: 2084-2092.

[16] Sanchez-Spitman AB, Dezentjé VO, Swen JJ, Moes DJAR, Gelderblom H and Guchelaar HJ. Genetic polymorphisms of 3'-untranslated region of SULT1A1 and their impact on tamoxifen metabolism and efficacy. Breast Cancer Res Treat 2018; 172: 401-411.

[17] Kuehl P, Zhang J, Lin Y, Lamba J, Assem M, Schuetz J, Watkins PB, Daly A, Wrighton SA, Hal SD, Maurel P, Relling M, Brimer C, Yasuda K, Venkataramanan R, Strom S, Thummel K, Boguski MS and Schuetz E. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Gen 2001; 27: 383-391.

[18] Jänne OA, Palvimo JJ, Kallio P and Mehto M. Androgen receptor and mechanism of androgen action. Ann Med 1993; 25: 83-89.

[19] Hibberts NA, Howell AE and Randall VA. Balding hair follicle dermal papilla cells contain higher levels of androgen receptors than those from non-balding scalp. J Endocrinol 1998; 156: 59-65.

[20] Randall VA. Role of 5 alpha-reductase in health and disease. Baillieres Clin Endocrinol Metab 1994; 8: 405-431.

[21] Sawaya ME and Price VH. Different levels of 5α-reductase type I and II, aromatase, and androgen receptor in hair follicles of women and men with androgenetic alopecia. J Invest Dermatol 1997; 109: 296-300.

[22] Hirschhorn JN, Lohmueller K, Byrne E and Hirschhorn K. A comprehensive review of genetic association studies. Genet Med 2002; 4: 45-61.

[23] Phan L, Jin H, Zhang W, Qiang E, Shekhtman D, Shao D, Revoe R, Villamarin E, Ivanchenko M, Kimura ZY, Wang L, Hao N, Sharopova M, Bihan A, Sturcke M, Lee N, Popova W, Wu C, Bastiani M, Ward JB, Holmes V, Lyoshin K, Kaur EM and B LK. ALFA: allele frequency aggregator. National center for biotechnology information, U.S. National Library of Medicine 2020.

[24] 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA and Abecasis GR. A global reference for human genetic variation. Nature 2015; 526: 68-74.

[25] Xu J, Turner A, Little J, Bleecker ER and Meyers DA. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? Hum Genet 2002; 111: 573-574.

[26] Wittke-Thompson JK, Pluzhnikov A and Cox NJ. Rational inferences about departures from HWE. Am J Hum Genet 2005; 76: 967-986.

[27] Graffelman J and Weir BS. Testing for Hardy-Weinberg equilibrium at biallelic genetic markers on the X chromosome. Heredity (Edinb) 2016; 116: 558-568.

[28] Martínez-Cortés G, Salazar-Flores J, Fernández-Rodríguez LG, Rubi-Castellanos R, Rodríguez-Loya C, Velarde-Félix JS, Muñoz-Valle JF, Parra-Rojas I and Rangel-Villalobos H. Admixture and population structure in Mexican-Mestizos based on paternal lineages. J Hum Genet 2012; 57: 568-574.

[29] Rangel-Villalobos H, Muñoz-Valle JF, González-Martín A, Gorostiza A, Magaña MT and Páez-Riberos LA. Genetic admixture, relatedness, and structure patterns among Mexican populations revealed by the Y-chromosome. Am J Phys Anthropol 2008; 135: 448-461.

[30] Rubi-Castellanos R, Martínez-Cortés G, Muñoz-Valle JF, González-Martín A, Cerda-Flores RM, Anaya-Palafox M and Rangel-Villalobos H. Prehispanic Mesoamerican demography approximates the present-day ancestry of Mestizos throughout the territory of Mexico. Am J Physical Anthropol 2009; 139: 284-294.

[31] Martinez-Fierro ML, Beuten J, Leach RJ, Parra EJ, Cruz-Lopez M, Rangel-Villalobos H, Riego-Ruiz LR, Ortiz-Lopez R, Martinez-Rodriguez HG and Rojas-Martinez A. Ancestry informative markers and admixture proportions in Mexico. J Hum Genet 2009; 54: 504-509.

[32] Martínez-Mariagnac VL, Valladares A, Cameron E, Chan A, Perera A, Globus-Goldberg R, Wacher N, Kumate J, McKeigue P, O'Donell D, Shriver MD, Cruz M and Parra EJ. Admixture in Mexico city: implications for admixture mapping of type 2 diabetes genetic risk factors. Hum Genet 2007; 120: 807-819.
[33] Lisker R, Ramirez E and Babinsky V. Genetic structure of autochthonous populations of Meso-America: Mexico. Hum Biol 1996; 68: 395-404.
[34] Barquera R, Bravo-Acevedo A, Clayton S, Munguía TJR, Hernández-Zaragoza DI, Adalid-Sáinz C, Arrieta-Bolaños E, Aquino-Rubio G, González-Martínez MDR, Lona-Sánchez A, Martínez-Álvarez JC, Arrazola-García MA, Delgado-Aguirre H, González-Medina L, Pacheco-Ubaldo H, Juárez-Barreto V, Benítez-Arvizu G, Escareño-Montiel N, Juárez-de la Cruz F, Jaramillo-Rodríguez Y, Salgado-Adame, A, Zúñiga J, Yunis EJ, Bekker-Méndez C and Granados J. Genetic diversity of HLA system in two populations from Nuevo León, Mexico: Monterrey and rural Nuevo León. Hum Immunol 2020; 81: 516-518.