

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITE SAAD DAHLEB – BLIDA 1



Faculté des Sciences de la Nature et de la Vie

Département de Biologie

Mémoire de fin d'études

En vue de l'obtention du diplôme de Master dans le domaine SNV

Filière Sciences Biologiques

Option : Génétique

Thème

**ESSAI D'IDENTIFICATION DE PERSONNE A PARTIR DE QUELQUES GENES
ESSENTIELS « DE SEXE, DE TAILLE ET DE MORPHOLOGIE DES CHEVEUX »
ETUDE IN SILICO**

Présenté par :

Date de soutenance : 12/ 09 /2022

ZERROUK MERIEM

AKEB OUARDIA

Devant le jury :

Nom	Grade	Lieu	Qualité
ROUAKI F	MCA	Blida	Présidente
ABDUL HUSSEIN	MCA	Blida	Examinatrice
MOHAMED SAID R	MCA	Blida	Promoteur
ZEGGAR H	MA	Blida	Co-promotrice

Promotion 2021-2022

Table des matières

Remerciement.....	9
Dédicaces	10
Résumé.....	12
Introduction :.....	15
Généralité :	18
1 La détermination du sexe :.....	22
1.1 Le chromosome Y :	22
1.1.1 L'évolution du chromosome Y :	22
1.1.2 La structure du chromosome Y :.....	24
1.1.2.1 PAR : La région pseudo-autosomale :.....	24
1.1.2.2 La région spécifique au mâle du chromosome Y :	24
1.2 La différenciation gonadique et l'identité du genre :	26
1.2.1 L'influence du chromosome Y sur la différenciation gonadique :.....	27
1.2.2 SRY : la région Y déterminant le sexe :.....	27
1.2.3 La protéine SRY :.....	27
2 La détermination de la taille :	30
2.1 L'influence de la plaque de croissance sur la taille :.....	30
2.1.1 La plaque de croissance :	30
2.2 La Génétique de la diversité de la taille :	32
2.2.1 Les études GWAS et GIANT :.....	32
3 La détermination de la morphologie des cheveux :	35
3.1 La Structure des cheveux :.....	35
3.1.1 Le follicule pileux :	35
3.1.2 La tige pileuse :	35
3.1.3 Les annexes de cheveu :	35
3.2 La Classification des cheveux humains :.....	37
3.3 La dynamique du cycle pileux :	38
3.3.1 La phase anagène : la pousse des cheveux :.....	38
3.3.2 La phase catagène : la fin de la croissance des cheveux :.....	39
3.3.3 La phase télogène : la maturité des cheveux :.....	39
3.4 La Génétique de la diversité capillaire :	41
Matériels et Méthodes Première étude : Détermination du sexe	

4	La détermination du sexe :.....	44
4.1	Extraction d'ADN à partir du sang total :.....	44
4.1.1	Le prélèvement du sang :	44
4.1.2	La lyse des globules rouges :.....	44
4.1.3	Traitement à la protéinase K :	44
	Après l'élimination du surnageant il faut ajouter :	44
4.1.4	La précipitation d'ADN :	45
4.1.5	Dosage et contrôle de la qualité de l'ADN :	45
4.2	Amplification par PCR : Polymerase Chain Reaction:	46
4.2.1	Les outils de la PCR :.....	46
4.2.2	Le principe de la PCR :	47
4.2.3	Contrôle du produit de la PCR :.....	47
4.2.4	Purification des produits de PCR :	47
4.3	Le séquençage de l'ADN :.....	48
4.3.1	La technique Sanger :.....	48
4.3.2	Le principe de Sanger :.....	48
4.4	La purification des produits de réaction de séquence :.....	49
4.4.1	L'électrophorèse capillaire :.....	49
4.4.2	L'analyse des données :.....	49
4.5	Les résultats attendus :.....	50
4.5.1	Premier cas :.....	50
4.5.2	Deuxième cas :	50
Matériels et Méthode Deuxième étude : Détermination de la taille		
5	La détermination de la taille :	52
5.1	Matériels biologique :.....	52
5.1.1	La collecte des échantillons :	52
5.2	Les Méthodes :.....	53
5.2.1	Génotypage :	53
5.3	Analyse statistique :.....	53
5.3.1	Test d'association :.....	53
5.4	Résultats :	54
Matériels et Méthode Troisième étude : Détermination de la morphologie des cheveux		
6	Matériels :	58
6.1	Matériel biologique :	58

6.1.1	La collecte des échantillons et phénotype :.....	58
6.2	Méthodes :	58
6.2.1	Le génotypage et contrôle de qualité :	58
6.2.2	Les analyses statistiques :.....	59
6.2.2.1	Le test d'association chez la population Européenne (EU) :.....	59
6.2.2.2	Le test d'association chez la population CANDELA :.....	59
6.2.2.3	Le test d'association chez la population chinoise :.....	60
6.3	Résultat :.....	61
6.3.1	Résultats de l'étude d'association chez les EU :.....	61
6.3.2	Les résultats de l'étude d'association chez la population CANDELA :	65
6.3.3	Résultat de l'étude d'association chez la population CHB :	71
	Discussion :	73
	Conclusion :.....	78
	Annexe :	80

Liste des figures

Figure 1 : L'évolution du chromosome Y	23
Figure 2: La différenciation des chromosomes X et Y initialement morphologiquement homogènes.....	23
Figure 3: La structure globale du chromosome Y	25
Figure 4: La structure moléculaire du chromosome Y.....	26
Figure 5: La plaque de croissance dans les os et la variation de la hauteur.	31
Figure 6: Image schématique du cartilage de la plaque de croissance à différents stades de son développement	32
Figure 7: Schéma représentant la structure du cheveu et de ses annexes	36
Figure 8: La glande sébacée (https://fr.dreamstime.com/).....	37
Figure 9: Une photo modifiée qui représente les huit types de cheveu	38
Figure 10: Les étapes clés du cycle pileux.....	40
Figure 11: Les différentes étapes de la technique Sanger	49
Figure 12: La comparaison des valeurs de taille chez les sujets avec différents génotypes pour les trois SNP significatifs chez les Caucasiens et chinois.....	56
Figure 13 : La combinaison de génotypes de SNP rs11803731, rs7349332, rs1268789.....	64
Figure 14: L'aperçu des résultats de l'étude d'association pangénomique (GWAS) des sept caractéristiques capillaires.....	66
Figure 15 : Les graphiques d'association pour quatre régions avec des SNP	69
Figure 16 : Les tailles d'effet pour l'allèle dérivé au niveau des SNP index.....	70
Figure 17: Le résultat du Diagramme de Manhattan et diagramme quantile-quantile montrant une analyse à l'échelle du génome de la raideur des cheveux chez les Chinois Han.	71
Figure 18: Le résultat du Diagramme de Manhattan et diagramme quantile-quantile montrant une analyse à l'échelle du génome de la raideur des cheveux chez les Ouïghours.	71

Liste des tableaux

Tableau 1: Un tableau récapitulatif de quelques gènes influençant la taille déterminés par des études GWAS.....	33
Tableau 2: Un tableau récapitulatif des SNPs associées à la diversité de morphologie des cheveux définie par une méta analyse.....	42
Tableau 3: Un tableau récapitulatif des données des deux populations différentes.....	52
Tableau 4: Tableau résume les résultats d'association pour 19 SNP dans HMGA2 avec la variation de taille adulte.	55
Tableau 5: Un tableau récapitulatif des données des trois études choisies.	58
Tableau 6 : L'association de huit haplotypes avec la morphologie des cheveux obtenue à partir de l'échantillon de la population polonaise.....	62
Tableau 7: Les résultats de l'analyse d'association univariée pour six SNPs prédictifs de la morphologie des cheveux.....	63
Tableau 8: Les résultats de l'analyse d'association multivariée pour les SNP prédictifs de la morphologie des cheveux.....	63
Tableau 9 : Les caractéristiques des SNP index montrant la plus forte association significative à l'échelle du génome (valeur $P < 5 \times 10^{-8}$) avec le cuir chevelu et caractéristiques des poils du visage examinées dans l'échantillon CANDELA.....	67

Liste des abréviations :

ADN: Acide désoxyribonucléique.

AR: Androgen receptor.

BET: Bromure d'éthidium...

ddNTPs: didésoxyribonucléotides triphosphate.

dNTPs: désoxyribonucléotides triphosphate.

ECM: extra cellular matrice.

EDAR : récepteur de l'ectodysplasine A1.

EDTA: Acide Ethylène Diamine Tétracétique.

EU : La population européenne.

EVC: External visible features.

FBN1: Fibrillin 1.

FRASI : Sous-unité 1 du complexe de matrice extracellulaire Fraser.

GB: Globule blanc.

GIANT: le consortium Genetic Investigation of Anthropometric Traits.

GR: Globule rouge.

GRS: Le score de risque génotypique

GWAS: Genome Wide Association Studies.

HMG: High Mobility Group.

HMG2: high mobility group AT-hook 2.

IHH: Indian hedgehog.

LINE1: Nuclear Elements 1.

MC: mélanocytes.

MgCl₂: Chlorure de Magnésium..

NCBI: National Center for Biotechnology Information.

Nm : nanomètre

NR**Y**: Non Recombining Y chromosome.

PAR : Pseudo Autosomal Region.

Pb: pair de base. **PCR** : Polymerase Chain Reaction.

SDS : dodécylsulfate.

SLR : Solution de Lyse des globules Rouges.

SNP: Single nucleotide polymorphism

SRY: sex-determining region on Y

STC₂: stanniocalcin 2.

STRs: Short tandem repeat

TBE: Tris Borate EDTA

TCHH: Trichohyaline.

TDF: determining factor of the testis

TM: Temperature de fusion

WNT10A: Wnt family member 10A

ZBTB38 : doigt de zinc et domaine BTB contenant 38.



Remerciement

Tout d'abord, nous ne pouvons pas faire ce travail sans la bénédiction et la sans force D'Allah subhanaho wa taala.

Nous commençons par exprimer nos vifs remerciements à Monsieur MOHAMED SAID RAMDANE, qui ce fut un honneur de diriger ce travail et nous tenons à exprimer notre plus profond respect.

Nous tenons également à remercier Madame ZEGGAR qui a fait l'honneur de diriger aussi ce travail, nous tenons à lui exprimer le profond respect.

Nous tenons également une reconnaissance aux membres du jury, Madame ROUAKI de nous avoir honoré en tant que présidente du jury de soutenance et Madame ABDUL HUSSEIN d'avoir accepté d'examiner le contenu de ce travail.

Nous remercions aussi SIDI-YAKHLEF MAHJOUBA NASSIMA pour son aide.





Dédicaces

J'ai l'honneur de dédier ce modeste travail réalisé grâce à l'aide d'Allah le tout-puissant. À la lumière de mes jours, la source de mes efforts, ma vie et mon bonheur que Dieu le garde pour moi, Aucune dédicace ne saura exprimer l'amour, l'estime et le respect que j'ai pour toi, mon cher père « Khaled ».

À la femme de ma vie, mon exemple éternel, mon soutien moral et la source de joie et de bonheur, la flamme de mon cœur, celle qui s'est toujours sacrifiée pour me voir réussir, que Dieu la garde pour moi, ma chère mère « Houria Choucha ».

À mes cher frères Alaa, Abd elmalek et Elmoatasim billah.

À mes chères sœurs Bouthaina, Imane, Salsabil, Ilhem et Ghofrane.

À toute ma grande famille, et ma tante Narimane qui m'a soutenu toujours

Moralement...

À la source de l'énergie positive ma très chère Tata Chafika

À mes chères amies G.Ikram, M.Ikram, Meriem, IBtisse, Ahlem, Maroua, NIhel, Nihed, Wided, Mina, Sabrina, Hiba, Hadil, Asma

Nadjet et hanane .

À mon meilleure camarade et binôme pendant cette 3ans Akeb Wardia.





Dédicaces

En premier j'adresse ce travail à mon cher papa que j'ai perdu toute petite j'aurai aimé que tu sois avec nous aujourd'hui, paix à ton âme.

J'ai l'honneur de dédier ce travail à la lumière et la femme de ma vie, mon exemple éternel, ma mère je tiens à te remercier pour ton amour, ton soutien et pour tous tes sacrifices, sans toi je serai jamais à cet endroit, que dieu te garde pour nous, je t'aime très fort maman.

À mon cher frère MOHAMED, merci beaucoup pour le rôle que tu joues dans ma vie, pour ton soutien et surtout pour ton amour que dieu te garde pour moi.

Je tiens à remercier mon cher fiancé FATHI bientôt mari pour ses encouragements, son amour et surtout pour la patience qu'il a envers moi pendant ces 5 ans.

À ma chère sœur SOUMIA, merci d'avoir fait partie de ma vie merci pour tous, je t'aime beaucoup.

À ma petite princesse SELMA à qui je tiens trop sans oublier tonton RABAH et ma tata FARIDA merci pour tous.

Je dédier ce travail à ma chère binôme et à mon amie MEREIM sans toi j'aurai jamais fait ce travail toute seule merci pour tous les moments quand a vécu ensemble je t'aime fort.

Je remercie aussi mes copines WARDIA, AMANI, MEREIM et RANIA pour leurs soutiens et leurs douaa.

À mon papi BELKACEM, et à mon meilleur oncle HAMID et un grand hommage à ma grand-mère TASSAHDIR qui nous a quitté cette année paix à son âme.

À toute la famille YERFA.

Un grand dédicace aussi pour mes petits amours, mes chats.



Résumé

Le phénotypage médico-légal de l'ADN a permis une meilleure compréhension des traits phénotypiques après il a été constaté que la méthode de typage ADN basée sur le SNP peut être informative ou avoir un grand potentiel pour résoudre un crime dans l'absence de profil d'ADN de suspect dans la base de données. Où il a pu concevoir un système pour prédire la couleur des cheveux, des yeux et de la peau (HIrisPlex). Actuellement il est en question de concevoir d'autres systèmes pour prédire les caractéristiques phénotypiques externes d'un individu de manière plus précise.

Notre présente étude porte sur l'étude des gènes responsables de la détermination du sexe et à détecter les variantes génétiques responsables de la détermination de la taille et de la morphologie des cheveux humains et leur contribution potentielle aux systèmes de prédiction du phénotype.

Des chercheurs dans le domaine des études d'association à l'échelle du génome ont conclu que le trait de la taille chez l'humain est un trait anthropologique polygénique qui reflète l'influence combinée de nombreux facteurs génétiques. Ces derniers ne sont pas encore connus, tandis que la variation de la forme des cheveux humains montre une variation frappante au sein et entre les populations humaines bien que la base génétique ne soit pas encore entièrement comprise. Par conséquent, les FDP n'ont pas révélé pour le moment une modèle de prédiction finale pour ces deux traits.

Contrairement à ces deux caractères, il est facile de prédire le sexe des individus à cause de la présence d'un seul gène responsable de la détermination des testicules de l'homme qui nous a permis de faire la différence entre les deux sexes humain.

Des études futures supplémentaires sont inévitables pour l'identification de nouveaux marqueurs génétiques pour une prédiction précise du phénotype ou des EVC via une association à l'échelle du génome (GWAS) et dans diverses populations mondiales.

Abstract

Forensic DNA phenotyping has led to a better understanding of phenotypic traits after it found that the SNP-based DNA typing method could be informative or have great potential to solve a crime in the absence of a profile of suspect DNA in the database. Where he was able to design a system to predict the color of hair, eyes and skin (HIrisPlex). He is currently looking to design other systems to predict individual external phenotypic characteristics more accurately.

The current study aimed to investigate genes responsible for sex determination and to detect genetic variants responsible for determining human hair size and morphology and their potential contribution to phenotype prediction systems.

Researcher's in future genome-wide association studies have concluded that height trait in humans is a polygenic anthropological trait that reflects the combined influence of many as yet unknown genetic factors, while Human hair shape variation shows striking variation within and between human populations although its genetic basis is not yet fully understood. Therefore, the FDPs are not yet revealing a final prediction model for these two traits.

Contrary to these two characters, it is easy to predict the sex of individuals because of the presence of a single gene responsible for the determination of the testicles of man, which allows us to make the difference between the two human sexes.

Further future studies are inevitable for the identification of new genetic markers for accurate prediction of phenotype or EVCs via genome-wide association (GWAS) and in various global populations.

ملخص

أدى التتميط الظاهري للحمض النووي في الطب الشرعي إلى فهم أفضل للسّمات المظهرية بعد أن وجد أن طريقة تصنيف الحمض النووي القائمة على تعدد أشكال نوكلّيوتيدة واحد يمكن أن تكون مفيدة أو لديها إمكانيات كبيرة لحل جريمة ما في حالة عدم وجود ملف تعريف للحمض النووي المشتبه به في قاعدة البيانات. حيث استطاع تصميم نظاماً للتنبؤ بلون الشعر والعينين والجلد ويتطلع حاليًا إلى تصميم أنظمة أخرى للتنبؤ بالخصائص المظهرية الخارجية للفرد بشكل أكثر دقة.

تهدف الدراسة الحالية إلى التحقيق في الحينات المسؤولة عن تحديد الجنس والكشف عن المتغيرات الجينية المسؤولة عن تحديد حجم و مورفولوجي اشعر الإنسان ومساهماتها المحتملة في أنظمة التنبؤ بالنمط الظاهري.

استنتج الباحثون في دراسات الارتباط على مستوى الجينوم في المستقبل إلى أن سمة الطول لدى البشر هي سمة أنثروبولوجية متعددة الجينات تعكس التأثير المشترك للعديد من العوامل الوراثية غير المعروفة حتى الآن، بينما يُظهر تباين شكل شعر الإنسان تباينًا صارخًا داخل السكان وبينهم على الرغم من أساسه الجيني لم يتم فهمه بشكل كامل بعد. لذلك لم يكشف التتميط الظاهري للحمض النووي الشرعي حتى الآن عن نموذج تنبؤ نهائي لهاتين السمتين.

على عكس هاتين الخاصيتين، من السهل التنبؤ بجنس الأفراد بسبب وجود جين واحد مسؤل عن تحديد خصيتي الرجل مما يسمح لنا بإحداث فرق بين الجنسين البشريين.

مزيد من الدراسات المستقبلية أمر لا مفر منه لتحديد الواسمات الجينية الجديدة للتنبؤ الدقيق بالنمط الظاهري أو بالخصائص المظهرية الخارجية عبر دراسة الارتباط على نطاق الجينوم وفي مجموعات سكانية عالمية مختلفة.

Introduction :

Le patrimoine génétique d'un organisme est préservé dans l'ADN présent dans toutes les cellules sachant qu'il est identique à 99,9 % chez tous les individus, ce qui signifie que les 0,1% restants rendent chaque personne vivante ou morte différente de toute autre par un ensemble de caractères physiques et morphologiques qui lui donne sa propre identité, tel que la couleur des yeux, la pigmentation de la peau, la taille, la voix, la morphologie et la couleur des cheveux...etc.

L'identification des identités humaines est devenue très importante pour diverses raisons, premièrement être capable d'identifier les personnes individuellement et en second lieu d'être en mesure de les identifier dans un grand groupe de personnes. Et aussi peut être d'une grande aide dans l'identification médico-légale puisque son importance longtemps limitée au crime s'est considérablement étendue avec la fréquence des accidents entraînant la carbonisation ou la destruction quasi totale des corps (comme les accidents d'avion), et peut également être utilisé dans les cas de perte, d'immigration illégale, de viol et de catastrophe de masse.

L'information génétique utilisée dans l'identification humaine peut être obtenus à partir du sang, de la salive, le sperme, les dents, cheveux (avec le bulbe), écouvillons vaginaux et les objets touchés, ainsi que les échantillons du sang séché, les traces et les tissus.

Dans certains cas, la quantité d'ADN présente dans l'échantillon peut être insuffisante ou bien trop dégradé pour pouvoir établir un profil génétique conventionnel à partir des STR exploitable. Et donc les enquêteurs se trouvent dans l'impossibilité de tirer profit des échantillons d'ADN recueillis sur une scène de crime ou encore le profil génétique déterminé par l'analyse de l'ADN recueilli peut être inconnu de la base de données. C'est dans cette optique que depuis quelques années, les laboratoires scientifiques et la police scientifique se sont intéressés à la possibilité d'établir des portraits-robots à partir des traces biologiques.

La compréhension génétique de l'apparence humaine de nombreux caractères visibles externes s'est améliorés ces dernières années ce qui a aider à développer le premier moyen validé de manière médico-légale (HIrisPlex-S) en 2018, afin de prédire la couleur des cheveux, des yeux et la pigmentation de la peau des personnes connus ou inconnus, basé sur l'identification des SNPs des gènes portés par l'échantillon d'ADN trouvé sur des scènes de crime ou scènes d'incidents.

Suite à une étude précédent dans le cadre de PFE sous le thème de la prédiction faciale à partir de l'expression de certains gènes dans le cas de la recherche d'enquête de crime ou disparition et dans le même contexte nous avons essayé de sélectionné quelques gènes déterminant le sexe, la taille et la morphologie des cheveux dans l'objectif de la contribution à la mise en place d'une méthode d'identification à partir de quelques gènes avec l'utilisation des outils bio-informatiques.

Généralité

Généralité :

L'ADN : Acide désoxyribonucléique :

L'acide désoxyribonucléique ou l'ADN est le support de l'information génétique. Ceci résume le fait que l'ADN contient l'information nécessaire à l'élaboration des protéines, constituants essentiels de tout être vivant. (Sicot, 2006)

D'après Michel Pugeat, chez l'homme, l'ADN est réparti sur 46 chromosomes (22 paires autosomes et une paire sexuelle) ; il porte des séquences codantes et non codantes dont certains jouent un rôle dans la régulation de l'expression des gènes.

Au dire de J. Lamoril, un gène est présent en plusieurs formes dénommés allèles et chaque caractère génétique possède deux allèles l'un transmis par le père biologique et l'autre par la mère biologique ; si les deux allèles sont identiques l'individu est dit homozygote pour le caractère par contre si les deux allèles se diffèrent l'individu est hétérozygote pour le caractère.

Le génotype et le phénotype :

Conformément aux Martin Mahner et Michael Kary, le génotype représente l'ensemble de traits génétiques propres à un individu et qui sont hérités par les parents. Et le phénotype est défini par l'assemblage de caractéristiques morphologiques et physiologiques observées chez un individu, qui est déterminé par les relations de dominances des allèles et la régulation du génotype, il peut également être influencé par des facteurs génétiques et environnementaux.

Le polymorphisme :

Selon Barkur S. Shastry, tous les individus d'une même espèce possèdent un taux de 99,9 % de la séquence d'ADN identique ce qui les rend génétiquement très proches toutefois, il existe des différences interindividuelles et personne ne ressemble à l'autre sauf les vrais jumeaux cela est due au 0.1% restants de l'ADN qui contient des variations qui sont à l'origine de l'individualité de chaque personne et qui représentent le polymorphisme.

STR : Séquences répétées en tandem courtes :

D'après Pham Hoai et ses collaborateurs, les séquences répétées en tandem courtes (STRs) nommées aussi microsatellites sont des motifs nucléotidiques répétés, ils sont généralement localisés dans des régions non codantes d'ADN et constitués de 2 à 6 paires de bases de nucléotides qui sont répétés 5 à 100 fois les uns après les autres, le nombre de répétition varie d'un individu à l'autre en raison de brassage génétique.

Leur fort pouvoir discriminant fait de ces marqueurs une option pour l'analyse médico-légale de l'ADN.

D'après les dires de Keyser et Petkovski en 2006, et bien que l'analyse des STRs soit facile à mettre en œuvre, elle peut se révéler inadaptée pour l'analyse des fragments d'ADN dégradé ainsi que la longueur de séquence à analyser (environ 400 pb) limite fortement la taille des fragments d'ADN amplifiables et leur détection ; c'est pour cela qu'il est préférable d'analyser un autre type de polymorphisme nommé les SNPs.

SNP : Polymorphisme d'un seul nucléotide :

Selon Keyser et Petkovski, le polymorphisme de séquence ou Snips sont des variations de la séquence d'ADN associées à un seul nucléotide, c'est-à-dire au niveau d'une seule paire de bases. Généralement ce sont des substitutions, insertions ou délétions de nucléotides qui surviennent à des positions uniques dans l'ensemble du génome des individus d'une même espèce ; ce sont les polymorphismes les plus répandus dans le génome humain, ils sont très abondants, stables et réparties à une fréquence d'environ 1 paire de bases sur 1000pb.

Il est à noter que généralement les SNPs ne modifient pas le phénotype du porteur et ils sont associés à la diversité de la population, elles présentent un taux de mutation bien inférieur à celui des STRs et permet d'étudier les acides nucléiques fragmentées parce que les régions ciblées pour l'analyse des SNP sont de petite taille.

Ainsi que d'après Shastry, de nombreuses études ont été établie suite à l'achèvement du séquençage de tout le génome humain en 2003 et à la découverte des SNPs pour définir les SNPs associés à la susceptibilité de certaines maladies et aux réponses individuelles aux traitements ainsi qu'à l'identification des gènes liés aux maladies en générale et multifactorielles.

Le polymorphisme de séquence et le phénotype :

Conformément à Keyser et Petkovski, de nombreuses études ont été faites dans le but de prédire les caractères phénotypiques individuelles en analysant les SNPs, notamment en médico-légales. En effet, le pouvoir prédictif à quoi ressemble une personne d'après l'ADN trouvé sur une scène de crime est d'une grande aide pour les enquêteurs, en limitant le nombre de suspects potentiels, et donne des orientations sur l'investigation lorsque les profils génétiques ne sont pas connus dans la base de données même lorsque l'ADN est fragmenté ou affecté par des facteurs environnementaux tels que la lumière, la chaleur ou l'humidité pour construire un profil génétique conventionnel (à partir des STR) exploitable. Cela pourrait être plus intéressant dans le contexte de découverte de cadavre ou dans le cadre de victimes de désastre, lorsque les corps sont décomposés, carbonisés et qu'il n'est plus possible d'avoir une description précise des traits morphologiques.

GWAS : l'étude d'association à l'échelle du génome :

D'après les dires de Gumpinger et ses collègues, l'étude d'association à l'échelle du génome est une approche de recherche qui permet d'identifier les variants génomique statistiquement associés à un risque de maladie à un trait. Elle consiste à l'étude des génomes de nombreuses personnes pour chercher les variantes génomiques qui se produisent fréquemment chez les personnes atteintes d'une maladie ou d'un trait spécifique.

Chapitre I

La détermination du sexe

1 La détermination du sexe :

D'après Pugeat, les observations des gènes humains indiquent qu'il existe deux genres, les individus portant les chromosomes 46XX représentent les femmes et les individus dont les chromosomes 46XY sont des hommes ; sur ses 46 chromosomes, deux sont considérés comme sexués autrement dit gonosomes, le chromosome X et le Y dont leur expression est limitée aux cellules germinales et sont responsables de la différenciation sexuelle.

Comme le soulignera la nature dicte que le chromosome X est transmis soit de la mère ou du père aux deux sexes, tandis que le chromosome Y n'est transmis qu'aux individus qui se verront attribuer comme étant un homme et se caractérise par une transmission uni-parentale qui veut dire du père au fils.

1.1 Le chromosome Y :

Au dire de Ravel et Siffroi, le chromosome Y autrement dit gonosome est chromosome qui s'exprime qu'au niveau des cellules germinales, il est responsable de la différenciation sexuelle, il est spécifique à l'homme.

1.1.1 L'évolution du chromosome Y :

Selon Ravel et ses collègues, les similitudes entre les séquences des chromosomes X et Y sont compatibles avec l'hypothèse d'origine commune. Le système avancé des chromosomes sexuels des mammifères est né il y a 300 millions d'années, les chromosomes X et Y étaient initialement des systèmes génétiquement largement homologues (Figure1).

Comme l'indique Gusmao et ses collaborateurs dans son article publié en 2008, l'évolution des chromosomes sexuels implique des mécanismes de restriction de la recombinaison, de la transposition et de la translocation des gènes. La séquence d'événements qui induisent la différenciation morphologique et génétique des chromosomes X et Y ainsi que l'inactivation génétique des gènes du chromosome Y n'a pas été entièrement élucidée. L'explication actuellement acceptée pour la différenciation des chromosomes X et Y initialement morphologiquement homogène implique un continuum de processus impliquant des étapes alternées de mutation, de restriction et de recombinaison (Figure2).

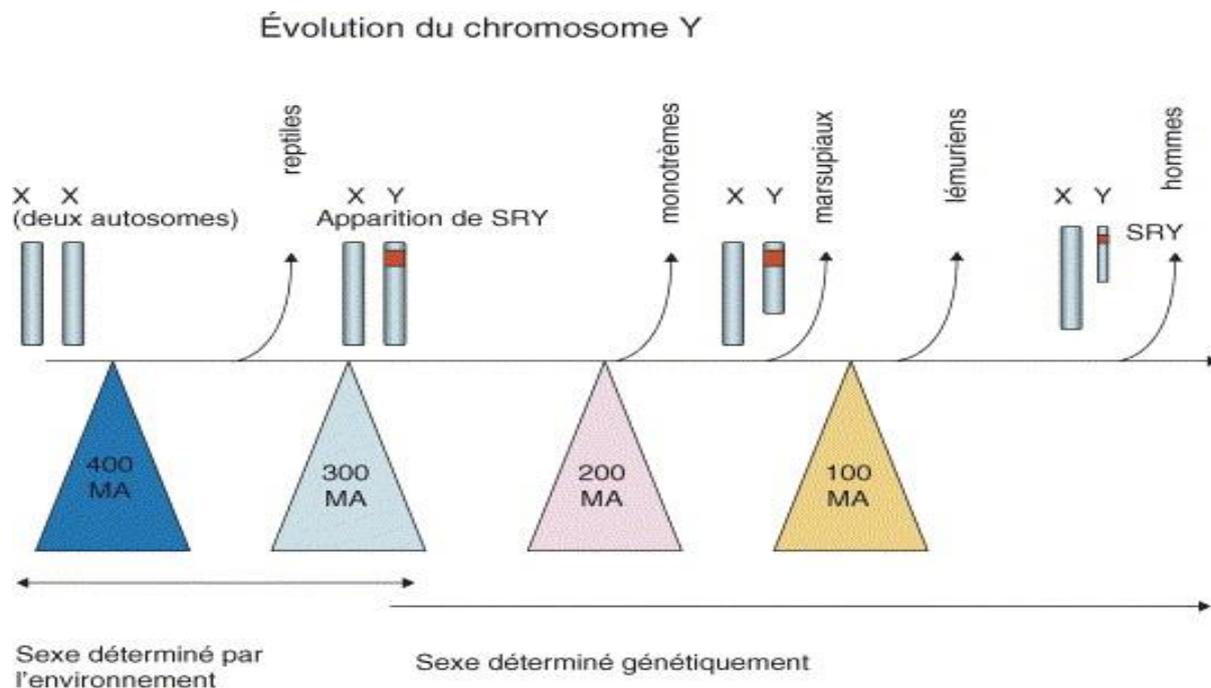


Figure 1 : L'évolution du chromosome Y

(Ravel et al., 2006)

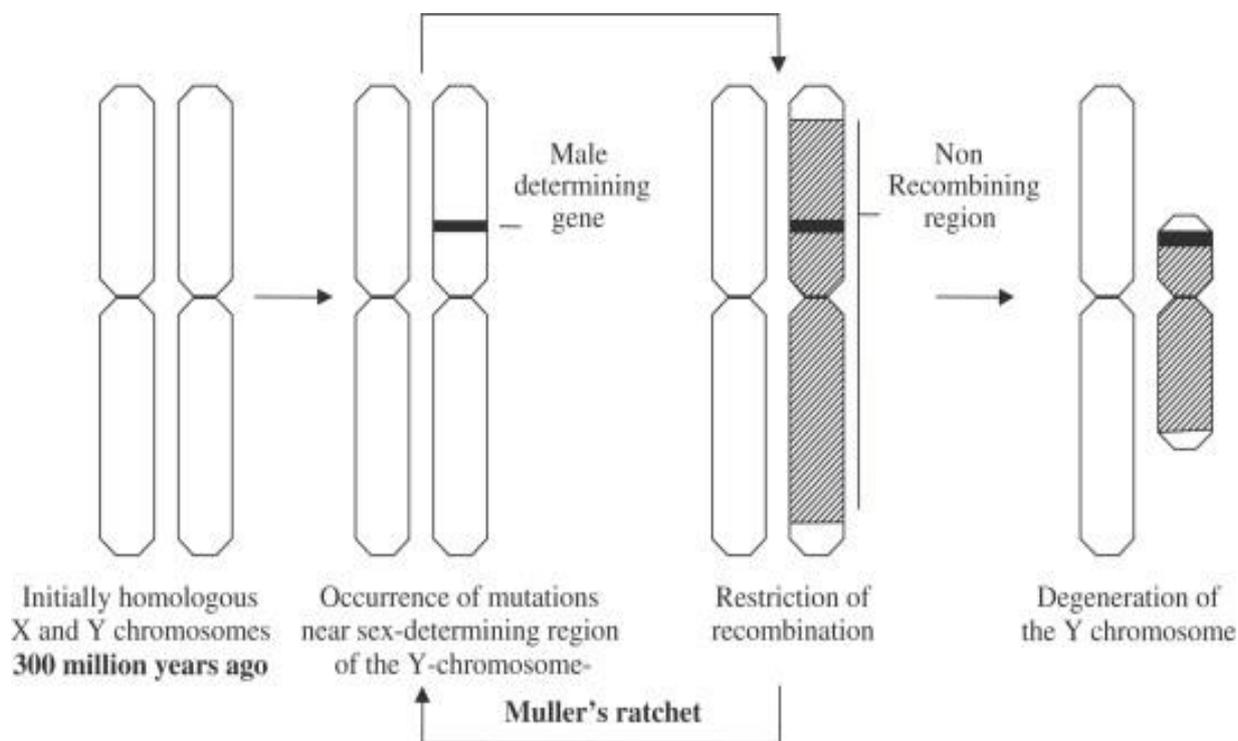


Figure 2: La différenciation des chromosomes X et Y initialement morphologiquement homogènes.

(Gusmão et al., 2008).

1.1.2 La structure du chromosome Y :

Comme ils le décrivent Ravel et ses collaborateurs dans les deux articles publiés en 2006 et 2009, le chromosome Y humain est considéré comme un petit chromosome, il fait le tiers de la taille du chromosome X et contient environ dix fois moins de gènes. (Figure 3) ; il est divisé en un bras long et un bras court qui sont reliés par un centromère, au niveau de ces extrémités possède deux régions pseudo-autosomales (PAR) entre lesquelles se trouve la région spécifique au mâle (MSY) :

1.1.2.1 PAR : La région pseudo-autosomale :

Les deux régions pseudo-autosomales nommées PAR (Pseudo Autosomal Region) flanquent le chromosome Y, et représentent 5 % de sa longueur ; elles se trouvent également aux extrémités du chromosome X et permettent la recombinaison inter-chromosomique.

Au niveau de ces deux régions, les chromosomes X et Y sont indiscernables et représentent les mêmes gènes.

1.1.2.2 La région spécifique au mâle du chromosome Y :

La région spécifique au mâle nommée MSY (Male Specific Y chromosome) précédemment NRY pour Non Recombining Y chromosome c'est une région qui normalement ne se recombine jamais avec le chromosome X ; elle est située entre les deux régions pseudo-autosomales et représente 95 % de la longueur du chromosome Y.

La région MSY est divisée en deux sous-régions :

- Une grande sous-région d'hétéro-chromatine compacte et non fonctionnelle située au niveau de la partie distale du bras long, qui est très riche en séquences répétées et présente des variations de taille importantes selon les hommes.
- Une deuxième sous-région dite euchromatique ou partie fonctionnelle s'étend sur environ 23 Mb, à la fois sur le bras court et sur la partie proximale du bras long, ce qui représente un peu moins de 1 % du génome humain. Elle contient 78 gènes qui codent pour des chaînes polypeptidiques, parmi ces gènes que l'on trouve exprimés chez l'homme et la femme et d'autres qui sont spécifiques aux hommes.

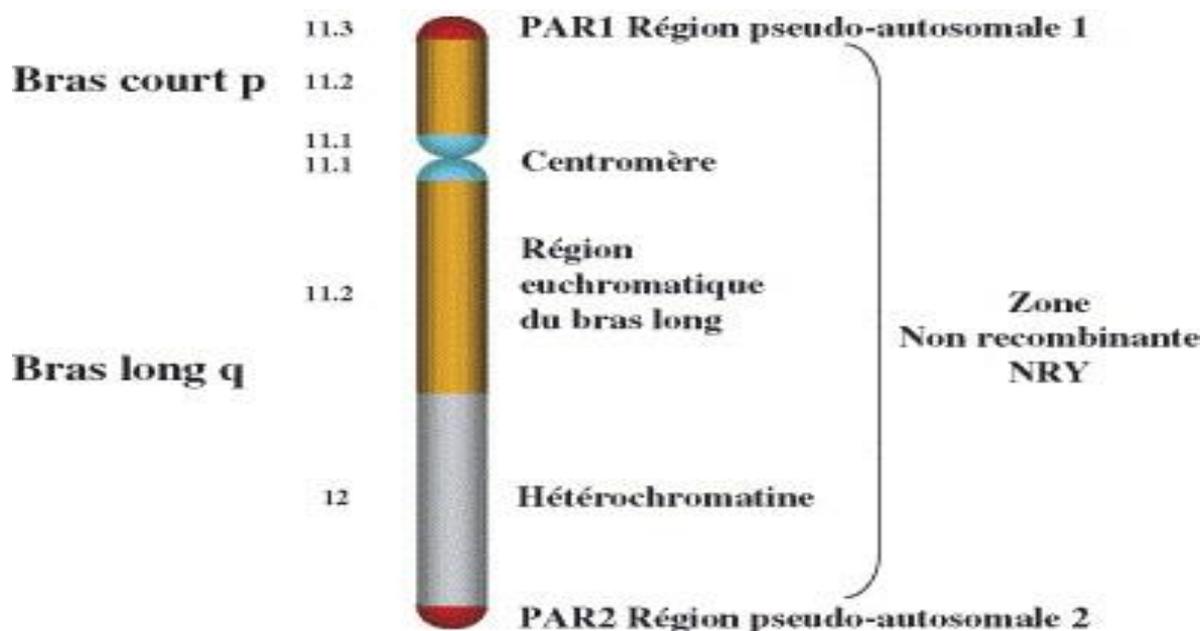


Figure 3: La structure globale du chromosome Y

(Ravel et al., 2006).

Conformément à l'article publié en 2006 par Ravel et ses collègues et suite au séquençage du chromosome Y, la structure moléculaire de la région MSY a été établie et trois types de séquences ont été identifiés :

- La première région, appelée « **X transposé** », a une taille de 3,4 Mb et partage 99 % d'homologie avec la région Xq21.2 sur le bras long du chromosome X. Cette région est pauvre en gènes mais riche en répétitions telles que de longs noyaux intercalés Élément séquence 1 (LINE 1).
- La deuxième région nommée « **X-dégénérées** », s'étend sur 8.5 Mb à la fois sur le bras court et le bras long, elle témoigne l'origine commune des deux gonosomes à fourche évolutive ; elle comporte en plus 16 gènes plus un certain nombre de pseudo-gènes sont en fait transcrits en protéines ou familles de protéines dont leur expression est plutôt ubiquitaire.
- La dernière région, appelée **amplicon**, est un bloc constitué de dizaines à des centaines de paires de bases répétitives avec une homologie presque parfaite. Ces amplicons constituent près d'un tiers de la région MSY euchromatique et ils sont situés sur les deux bras (long et court) du chromosome. Parmi ces amplicons, la plupart des gènes sont associés à la spermatogenèse, et leur expression est exclusive ou prédominante dans les testicules (Figure 4).

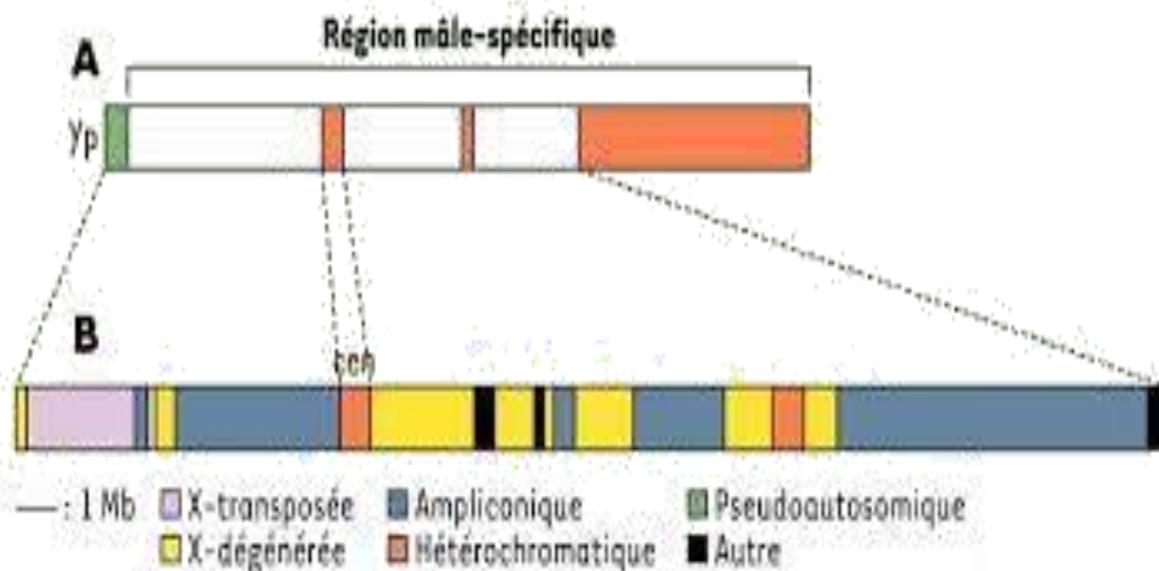


Figure 4: La structure moléculaire du chromosome Y
(Gilgenkrantz, 2004).

1.2 La différenciation gonadique et l'identité du genre :

Selon Pugeat, des études de différenciation sexuelle ont montré que l'intervention initiale des chromosomes sexuels est marquée par l'apparition du chromosome Y, qui contient une région déterminante du sexe, ou "sex determining region of the chromosome" (*SRY*). Une cascade de gènes localisés aléatoirement dans le génome sur différents chromosomes, dont l'expression est sous forme de protéines est déclenchée chronologiquement et conduit à la différenciation anatomique des gonades primitives en ovaires ou en testicules (Annexe1).

La différenciation des gonadiques est séparée en deux étapes :

La première étape est la détermination des gonades, elle est sous le contrôle des gènes inclus dans le sexe génétique.

Au cours de la cinquième semaine du développement embryonnaire, les premiers contours des gonades apparaîtront sous forme de gonade indifférenciée caractérisée par une orientation bi-potentielle qui a la capacité de suivre une voie de différenciation mâle ou femelle selon les facteurs exprimés.

À partir de la septième semaine de développement, les gonades indifférenciées commencent à se transformer en testicules chez les embryons humains de sexe masculin (46, XY), mais cette

transformation ne se produit pas chez les embryons de sexe féminin (46, XX) qu'à partir de la huitième semaine.

La deuxième étape est la différenciation sexuelle primaire, qui dépend de la sécrétion hormonale des gonades différenciées et de la sensibilité des tissus cibles à ces sécrétions hormonales et qui aboutit au sexe phénotypique.

1.2.1 L'influence du chromosome Y sur la différenciation gonadique :

La présence du facteur de différenciation sexuel *SRY* permet l'activation des facteurs impliqués directement dans le développement des testicules ou responsables de l'inhibition du développement ovariens.(Pugeat, 2021).

L'absence du facteur *SRY* permet le développement par un mécanisme d'inhibition du gène *SOX9* qui code pour un facteur indispensable à formation des testicules (Pugeat, 2021).

1.2.2 *SRY* : la région Y déterminant le sexe :

Selon les informations adoptés de la plateforme NCBI, le *SRY* région Y déterminant le sexe aussi connu sous le nom de TDF, TDJ, SRXX1, SRXY1, est situé sur le bras court du chromosome Y en position Yp11.2. C'est un gène mono-exonique qui veut dire un seul exon sans aucun intron, il s'étend sur 3.8 Kb dont le cadre ouvert de lecture est composé de 615 pb. Il code pour un facteur de transcription qui fait partie de la famille des boîtes HMG (High Mobility Group) des protéines de liaison à l'ADN. Cette protéine est le testis-determining factor (TDF), qui initie la détermination du sexe masculin. Les mutations de ce gène donnent naissance à des femelles XY atteintes de dysgénésie gonadique (syndrome de Swyer) ; la translocation d'une partie du chromosome Y contenant ce gène vers le chromosome X provoque le syndrome masculin XX.

1.2.3 La protéine *SRY* :

Selon l'article publié en 1997 par François Poulat et ses collègues, le gène *SRY* code pour une protéine de 204 acides aminés, ayant une masse moléculaire calculée de 24 kDa chez l'Homme. La protéine *SRY* se compose de trois parties principales :

- **Un domaine central** qui constitue la boîte HMG (High Mobility Group) qui comporte 79 acides aminés. La boîte HMG est la seule région de la protéine *SRY* qui est hautement conservée entre les différentes espèces. Cette boîte correspond à un domaine capable de se

lier, avec une haute affinité, à une séquence consensus A/TAACAAT/A au niveau de l'ADN.

- **Une partie N-terminale** qui est très peu conservée entre les différentes espèces.
- **Une partie C-terminale** cette partie est également très peu conservée entre les différentes espèces, elle est impliquée dans l'établissement d'une connexion entre la protéine SRY, déjà fixée à l'ADN, et une protéine du complexe transcriptionnel, appelée SIP-1.

Chapitre II

Détermination

de la taille

2 La détermination de la taille :

Selon He, la taille humaine est un trait anthropométrique complexe, classique, polygénique et hautement héréditaire, avec environ 80 à 90% de variation expliquée par des composants génétiques.

Dans les populations européennes, la taille moyenne est d'environ 178 cm pour les hommes et d'environ 165 cm pour les femmes,(Visscher, 2008).D'autres part, les tailles moyennes des hommes et des femmes américains sont respectivement de 176 cm et 163 cm (Gudbjartsson et al., 2008).

2.1 L'influence de la plaque de croissance sur la taille :

L'allongement des os interfère avec la variation de taille humaine, les études GWAS ont montré un fort enrichissement pour les voies de signalisation Hedgehog tel que de gène *IHH*, qui a été déjà impliquées dans le développement de la plaque de croissance.(Guo et al., 2018).

2.1.1 La plaque de croissance :

La plaque de croissance est une structure cartilagineuse située aux extrémités des os longs. Il se compose de cellules spéciales appelées chondrocytes qui sont disposées en des zones distinctes : zones de repos, prolifératives, pré-hypertrophiques et hypertrophiques (Figure 5). La régulation complexe au niveau de la plaque de croissance détermine la longueur des os longs et la taille humaine (Muthuirulan & Capellini, 2019).

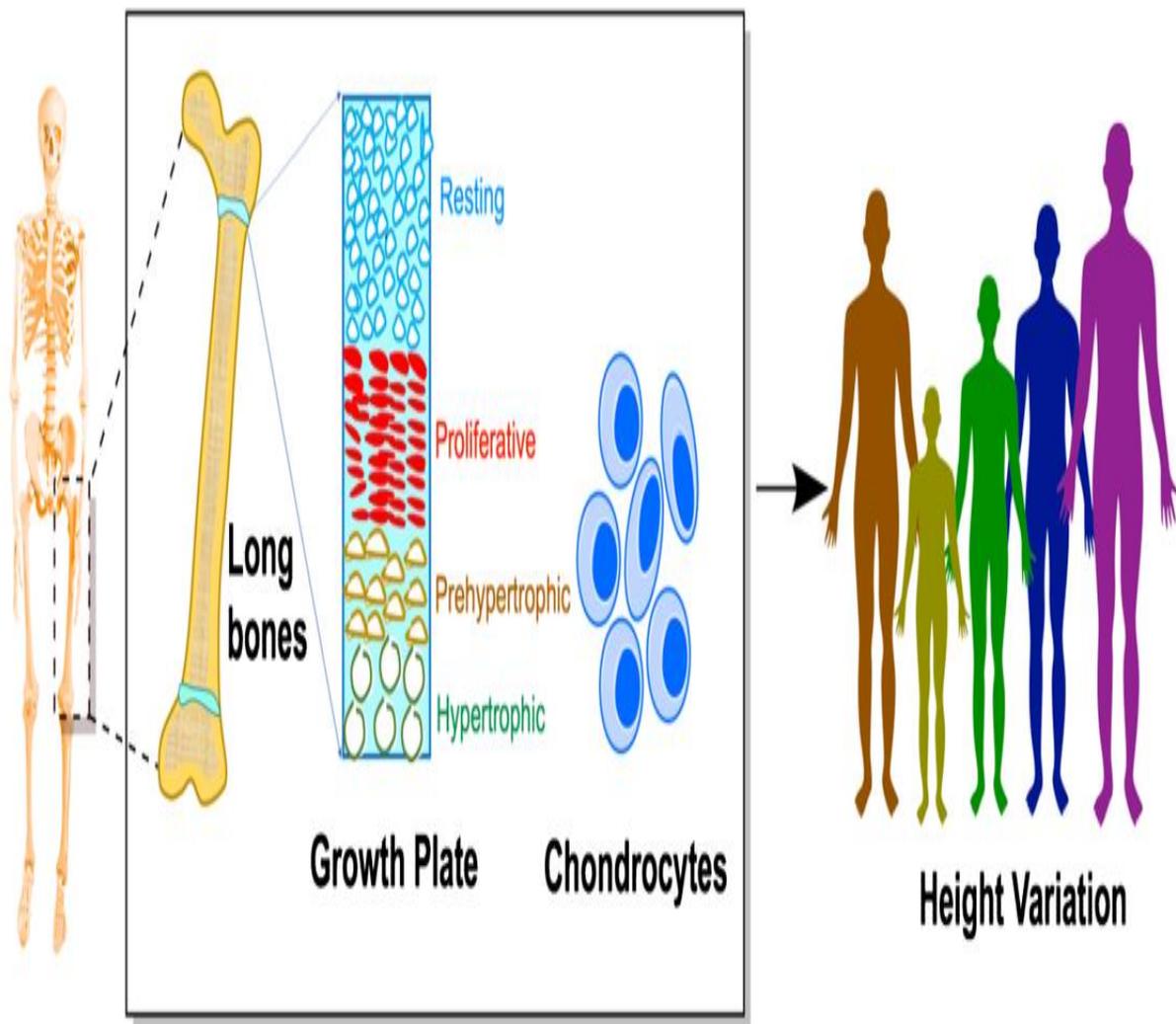


Figure 5: La plaque de croissance dans les os et la variation de la hauteur.

(Muthuirulan & Capellini, 2019)

Initialement, les chondrocytes de la zone de repos se répliquent à un rythme lent et donnent naissance à des chondrocytes prolifératifs, qui se divisent relativement rapidement en cellules filles apparentées par clonage. Ces cellules filles s'alignent généralement en colonnes le long de l'axe longitudinal et grâce à leur expansion et à leur prolifération, les os s'allongent.

Les chondrocytes prolifératifs finissent par cesser de se diviser et se différencient finalement en cellules hypertrophiques.

Les chondrocytes augmentent jusqu'à vingt fois, contribuant de manière significative à la croissance osseuse longitudinale. (Figure 6) (Muthuirulan & Capellini, 2019).

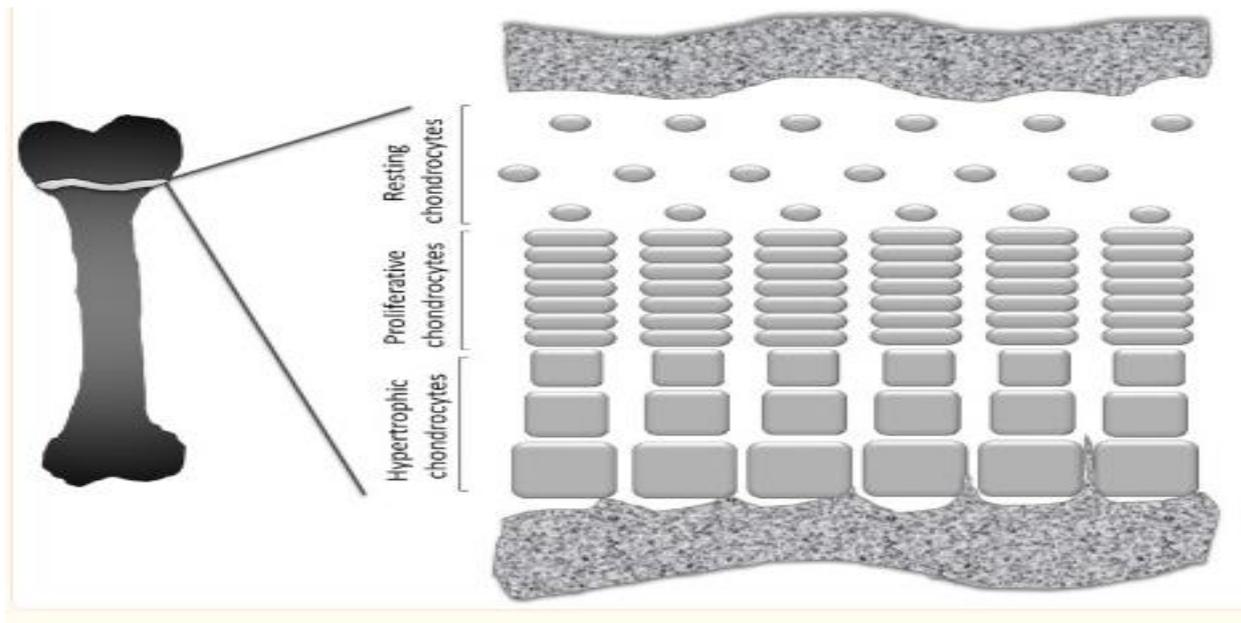


Figure 6: Image schématique du cartilage de la plaque de croissance à différents stades de son développement

(Karimian et al., 2012)

2.2 La Génétique de la diversité de la taille :

2.2.1 Les études GWAS et GIANT :

La contribution de la variation de la séquence d'ADN à la variation de la taille chez les adultes a été systématiquement évaluée par des études d'association à l'échelle du génome (GWAS).

Conformément à Guo, le premier GWAS analysant la taille adulte debout a été réalisé chez une population européenne. L'étude a identifié un variant génétique fortement associée au *HMGA2*. Cette découverte a donné la première indication que les GWAS pour la taille pouvaient identifier les gènes pertinents pour la croissance.

En 2010, le consortium Recherche Génétique des Traits Anthropométriques (GIANT) a montré que la variation la plus fortement associée à la taille était proche du gène *ZBTB38*. Toutefois chaque allèle de gène *ZBTB38* augmente la hauteur de ~0,4 cm.

En 2014 le consortium GIANT a effectué une série de GWAS afin de déterminer la taille d'une population européenne, l'étude a permis d'identifier 697 variants (SNPs) repartit dans 423 loci. Ces variants agissent indépendamment pour influencer la taille.

Gène	Chromosome	SNP	Position	Allèle	Référence
<i>HMGA2</i>	12	rs1042725	65964567	C > T	(Yang et al., 2010)
<i>ZBTB38</i>	3	rs2871960	141402972	A > C	(Muthuirulan & Capellini, 2019) + adopté de l'NCBI
<i>FBN1</i>	15	rs200342067	48481729	T > C	
<i>KCNQ1</i>	11	rs150199504	2793730	C > G	
<i>FADS3</i>	11	rs7115739	61874245	T > A	
<i>IHH</i>	2	rs142036701	219924961	G > T	(Marouli et al., 2017)
<i>STC₂</i>	5	rs148833559	172755066	C > A	
<i>AR</i>	X	rs137852591	66941751	C > G	

Tableau 1: Un tableau récapitulatif de quelques gènes influençant la taille déterminés par des études GWAS.

Chapitre III
Détermination de la
morphologie
des cheveux

3 La détermination de la morphologie des cheveux :

Selon Bernard en 2006, Le cheveu est l'un des traits caractéristiques de l'espèce humaine, il peut être considéré comme une sécrétion holocrine. Le nombre normal de chevelure d'une personne est de 100 000 à 150 000 cheveux renouvelés périodiquement.

3.1 La Structure des cheveux :

D'après Daoudi en 2005, Le cheveu est constitué d'un follicule pileux, d'une tige pileuse et d'annexes. (Figure 7).

3.1.1 Le follicule pileux :

Le follicule pileux du cheveu est une structure complexe composé d'une papille dermique, bulbe pileuse, une gaine épithéliale interne et externe.

3.1.2 La tige pileuse :

La tige pileuse est la partie apparente du cheveu qui se situe au niveau de la surface de l'épiderme. Elle est dépourvue des vaisseaux sanguins et des nerfs, et se présente sous la forme d'un cylindre constitué de trois couches : la cuticule, la corticale et la médullaire. Son diamètre est de 20 à 40 μm chez le nouveau-né et de 70 à 100 μm . chez l'adulte.

3.1.3 Les annexes de cheveu :

Les annexes de cheveu est composé d'une glande sébacée, un muscle arrecteur du poil, une vascularisation et une innervation :

- **La glande sébacée :** la glande sébacée est formée de plusieurs alvéoles dilatées en sec sécrète des gouttelettes lipidiques pour lubrifier le poil et la peau, elle possède une activité sécrétoire, (Figure 8).
- **Le muscle arrecteur du poil :** les fibres musculaires lisses se développent parallèlement à la formation du poil et de la glande sébacée. Elles sont insérées sur le follicule pileux et sur la glande sébacée, et sont sollicitées au cours de l'érection du poil ou « horripilation », de plus elle a un rôle dans la thermorégulation.
- **La vascularisation du follicule pileux :** est assurée par plusieurs capillaires localisés dans la papille, elle a un rôle dans la détermination du cycle pileux. Au cours du vieillissement physiologique, l'apport du sang artériel est réalisé par des vaisseaux issus du plexus artériel

dermique profond, tandis que le retour se fait grâce aux veinules qui rejoignent le plexus veineux dermique profond.

- **L'innervation** : le rôle des terminaisons nerveuses papillaires est l'intervenir dans le maintien de la trophicité et dans le renouvellement du poil et du cheveu au cours du cycle pileaire. Elle est importante chez certains animaux plus que l'homme.

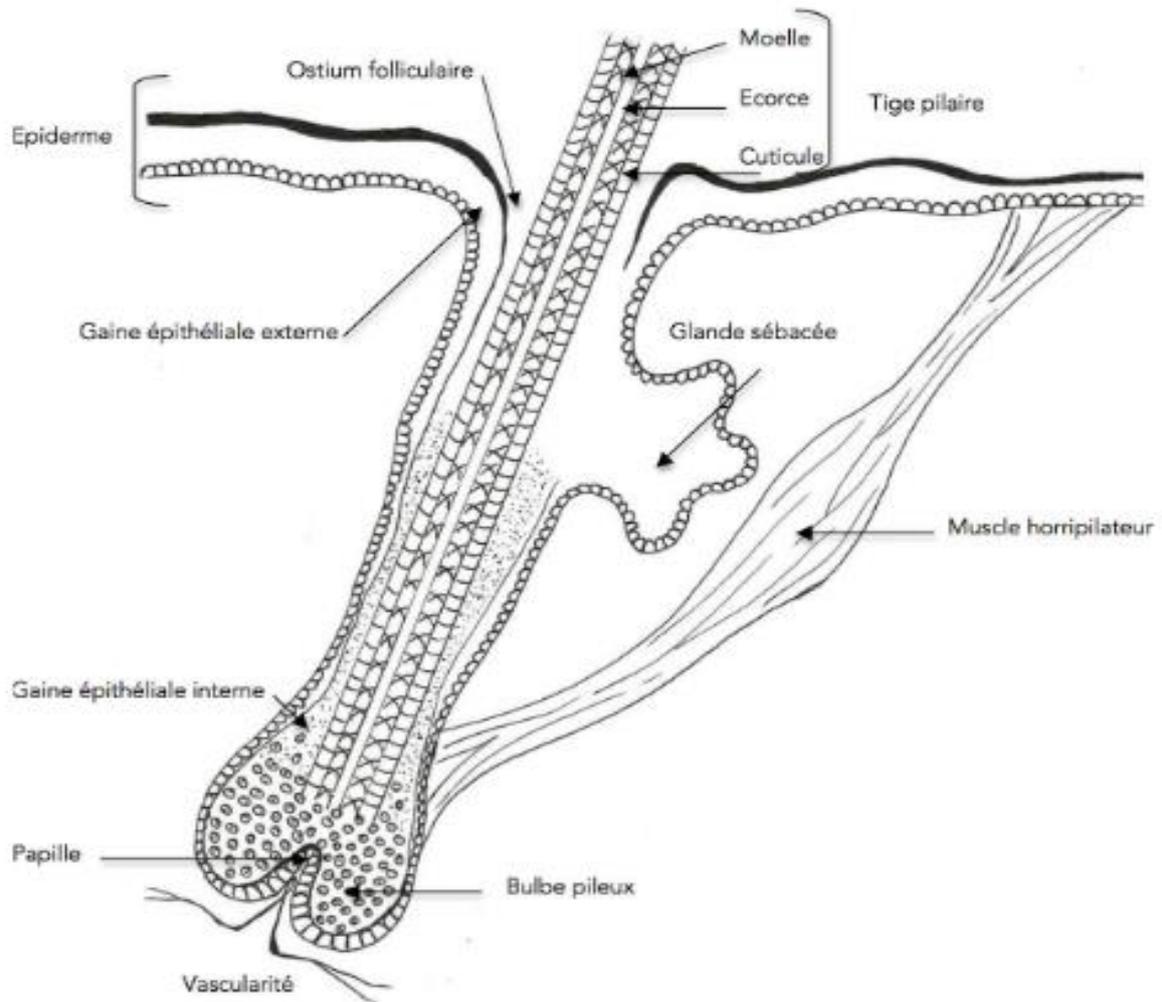


Figure 7: Schéma représentant la structure du cheveu et de ses annexes

(DAOUDI, K. (2005).

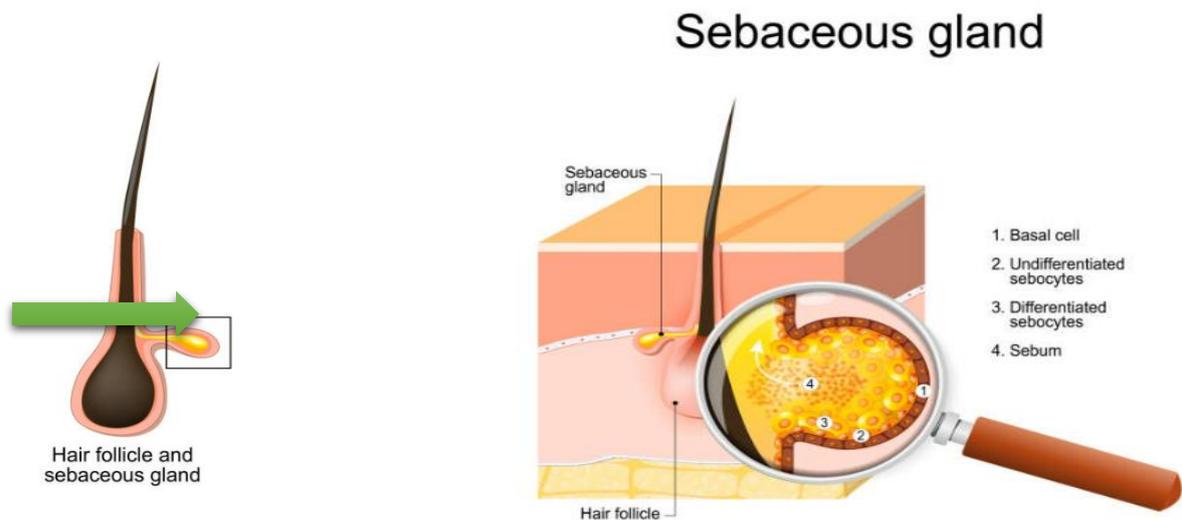


Figure 8: La glande sébacée (<https://fr.dreamstime.com/>)

3.2 La Classification des cheveux humains :

Afin de développer une vision plus factuelle et une approche scientifique basée sur les caractéristiques physiques des cheveux humains, un classement des cheveux humains a été effectué par Mettrie et ses collègues en 2007 selon trois sous-groupes : africain, asiatique et européens, qui aboutit à l'apparition de huit types de cheveux :

Type 1 : Cheveux raides et lisses.

Type 2 : Cheveux ondulés en forme de « S » : Le cheveu peut être raide à bouclé.

Type 3 : Cheveux des boucles lâches et détendues.

Type 4 : Cheveux avec des boucles ouvertes.

Type 5 : Chevelure avec des boucles fermées.

Type 6 : Cheveux avec des boucles fermées en forme de tire-bouchon.

Type 7 : Cheveux frisés et fragiles.

Type 8 : Cheveux crépus extrêmement fragiles.

Le degré de la raideur ou de la frisure des cheveux est déterminé par la forme du follicule pileux intégré. (Meng et al., 2019).(Figure 9).



Figure 9: Une photo modifiée qui représente les huit types de cheveu

(Mettrie et al., 2007)

3.3 La dynamique du cycle pileux :

Chaque follicule pileux passe perpétuellement par trois stades : croissance (anagène), involution (catagène) et repos (télogène). (Figure 10)

3.3.1 La phase anagène : la pousse des cheveux :

Suivant Daoudi en 2005, la phase anagène est une période de croissance continue des cheveux, elle est la conséquence du rythme de multiplication effréné des cellules matricielles du bulbe. La durée de cette phase varie de quelques mois à plusieurs années.

Au début de l'anagène, un sous-ensemble de cellules souches du follicule pileux s'active et commence à se diviser en cellules filles, ces dernières migrent vers la base du follicule, où elles commencent à se proliférer et deviennent des cellules matricielles activées de manière transitoire.

Des études de marquage et de transplantation in vivo ont confirmé la capacité des cellules souches renflées à donner naissance à toutes les lignées cellulaires du follicule pileux mature (Schneider et al., 2009).

3.3.2 La phase catagène : la fin de la croissance des cheveux :

La phase catagène est la transition dynamique entre l'anagène et la télogène, dans laquelle la partie inférieure de chaque follicule pileux dégénère complètement dans un processus qui comprend l'apoptose des cellules épithéliales dans le bulbe et la gaine externe de la racine (ORS). La base de la tige pileuse (HS) est scellée dans une structure circulaire appelée club, qui se déplace vers le haut jusqu'à ce qu'elle atteigne le follicule pileux supérieur acyclique permanent, qui reste ancré pendant le télogène (Alonso & Fuchs, 2006).

3.3.3 La phase télogène : la maturité des cheveux :

Conformément à Daoudi en 2005, Durant ce stade de cycle pileux, la profondeur du follicule est réduite à la moitié de ce qu'elle est lors de la phase anagène. ainsi que le poil télogène apparaît plus clair avec un diamètre inférieur à celui de poil anagène, il reste ancré dans le sac folliculaire et continu à monter lentement vers le haut jusqu'à ce qu'il atteigne l'infundibulum du follicule où il chute, au même temps au niveau de la partie profonde du follicule se dessine un nouveau bourgeon pileux et un nouveau cycle pileux est induit, qui fait entrer un nouveau poil en phase anagène.

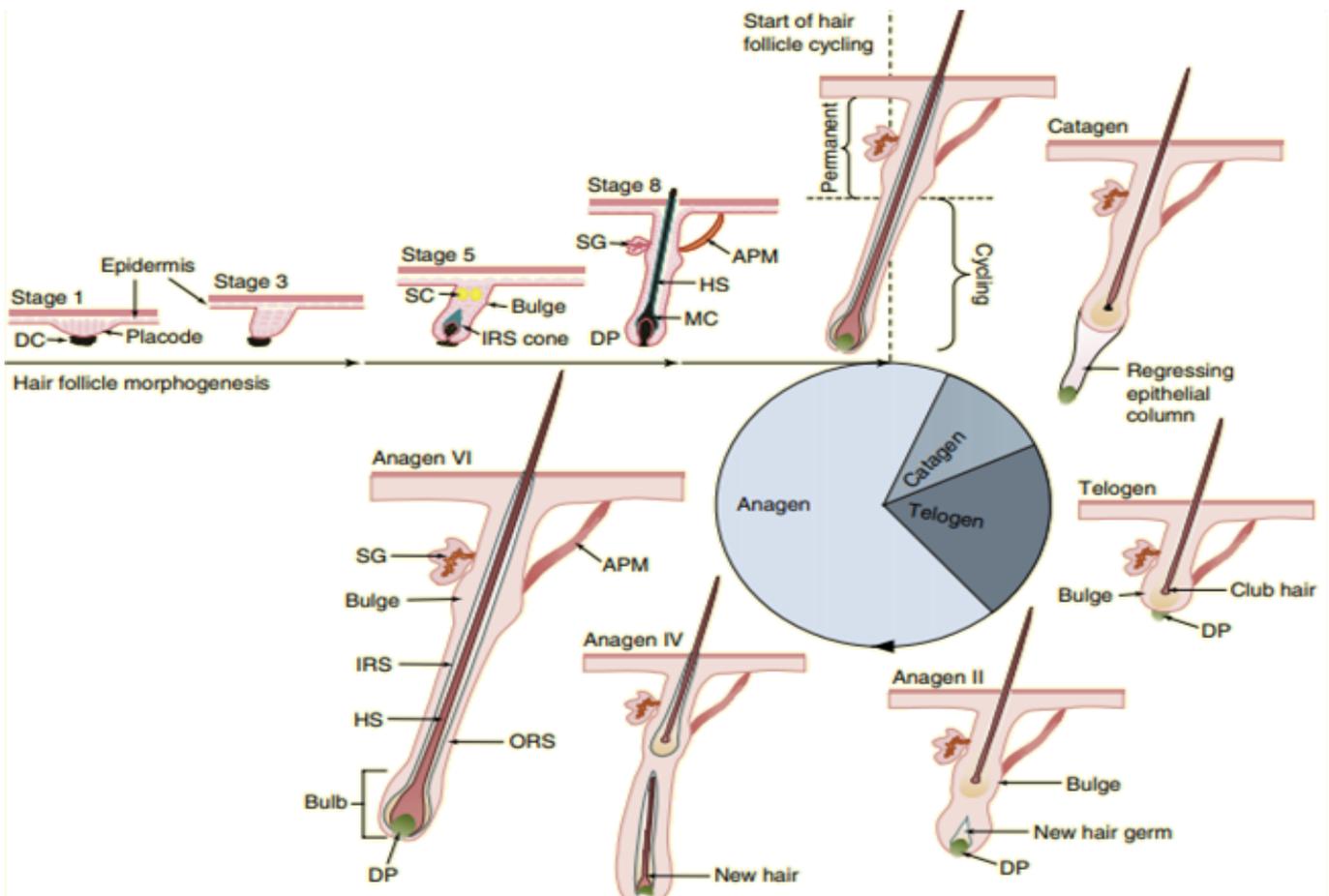


Figure 10: Les étapes clés du cycle pilaire.

(Schneider et al., 2009).

Les étapes du 1 à 8 du développement embryonnaire des cheveux.

La morphogénèse des cheveux postnatals mène au développement du follicule et à la production de la fibre capillaire qui émerge de la peau, une fois le follicule pileux mûré, il entre dans la phase de régression, au cours de laquelle la partie cyclable inférieure du follicule pileux est dégradée. Ce processus dirige la papille dermique à proximité du renflement, où résident les cellules souches ciliées (CSH) ces dernières interagissent avec la papille dermique pour former un nouveau follicule pileux (Paus & Cotsarelis, 1999; Schneider et al., 2009).

APM : muscle pili arrecteur **DC** : condensat dermique (vert) **DP** : papille dermique

HS : tige pilaire (brun) **IRS** : gaine radiculaire interne (bleue) **MC** : mélanocytes

ORS : gaine radiculaire externe **SC** : sébocytes (jaune) **SG** : glande sébacée.

3.4 La Génétique de la diversité capillaire :

D'après Pośpiech et ses collaborateurs, la morphologie des cheveux est l'une des caractéristiques visibles externes (EVC) les plus remarquables de la variation humaine qui a été réalisées dans les analyses de la génétique médico-légales.

La prédiction de la morphologie des cheveux quel que soit son type lisse, ondulé ou bouclé est basée sur une trace biologique (ADN) qui permette une reconstruction médico-légale plus précise de l'apparence extérieure et mène à une meilleure description d'une personne inconnu.

En 2021, Tozzo a réalisé un méta analyse sur les études d'association à l'échelle du génome (GWAS) qui a définie des SNPs au niveau des gènes qui sont avérés liés à la forme des cheveux, ces derniers ont décrit en-dessus (Tableau2).

Gènes	SNPs	Allèle	Chr	Position
<i>TCHH</i>	rs11803731	A > T	1	152110849
<i>WNT10A</i>	rs7349332	A > G	2	218891661
<i>FRAS1</i>	rs1268789	A > G	4	78359539
<i>EDAR</i>	rs3827760	A > G	2	108897145
<i>ERRF1/SLC45A1</i>	rs80293268	G > C	1	8147519
<i>PTK6</i>	rs310642	A > C	20	63530645
<i>LGR4</i>	rs2219783	G > A	11	27389751
<i>PRSS53</i>	rs11150606	T > C	16	31087690
<i>GATA3</i>	rs17143387	T > G	10	8272081
<i>OFCC1</i>	rs1556547	G > A	6	10270144
<i>LCE3E</i>	rs499697	G > A	1	152520678
<i>PEX14</i>	rs6658216	C > T	1	10501547
<i>PADI3</i>	rs11203346	A > G	1	17274327

<i>TGFA</i>	rs12997742	T > C	2	70559466
<i>LGR4</i>	rs2219783	G > A	11	27389751
<i>HOXC13</i>	rs11170678	T > G	12	53760390
<i>KRTAP</i>	rs11078976	T > C	17	41033108

Tableau 2: Un tableau récapitulatif des SNPs associées à la diversité de morphologie des cheveux définie par une méta analyse

(Tozzo et al., 2021) (NCBI).

Matériels

et

Méthodes

Première étude :

Détermination

du sexe

4 La détermination du sexe :

L'ADN peut être extrais à partir de la salive, sang, sperme, les dents, cheveux y compris le bulbe, écouvillons vaginaux et objets touchés, l'ADN tactile (traces) et d'échantillons de sang séché, ainsi que les tissus.

Dans cette étude, on a proposé un protocole de Service Des Maladies Congénitales Et Hérititaires Hôpital Charles Nicolle 2016/2017.

4.1 Extraction d'ADN à partir du sang total :

L'extraction d'ADN est une technique qui permet d'isoler des quantités suffisantes d'ADN de haute qualité. La technique est basée sur la précipitation à l'alcool de l'ADN, la post-lyse des globules rouges et blancs et la digestion enzymatique d'autres protéines à l'aide de la protéinase K.

4.1.1 Le prélèvement du sang :

Prélever une quantité de 10 à 30 ml de sang sur EDTA dans des tubes Falcon contenant les codes des échantillons, suivis d'une centrifugation à 3500 rpm/minute pour une période de 10 à 15 min ,récupérer le sérum à l'aide d'une pipette Pasteur en plastique en effectuant des mouvements circulaires sur les parois du tube afin d'éviter la récupération de la couche des globules blancs ou le Buffy coat ensuite le jeter dans un bocal rempli d'eau de javel.

4.1.2 La lyse des globules rouges :

Prélever le maximum de la couche des globules blancs (GB) et la déposer dans un tube Falcon puis ajouter 40 à 50 ml de la solution de Lyse des globules rouges (GR) le SLR, une solution hypotonique qui permet l'éclatement de la membrane cellulaire, laisser agir le mélange (GB, SLR) 20 mn à température ambiante et centrifuger les tubes à 3500 rpm/mn pendant 15 minutes puis jeter le surnageant, décoller le culot des GB en tapant légèrement le fond du tube Falcon.

Ajouter 15 ml du tampon de lyse des GB(SLB 10%), mélanger en vortexant, vérifier que tous les culots ont été dissouts puis centrifuger à 3500 rpm/mn pendant 15mn.SLB (10%)

4.1.3 Traitement à la protéinase K :

Après l'élimination du surnageant il faut ajouter :

- 5ml de tampon de lyse des GB.
- 250 µl de SDS 10%.

- 100 µl de protéinase K.
- Le **SDS** est un détergent qui lyse la membrane cellulaire des leucocytes en agissant sur les lipides membranaires.
- La **protéine K**, une serine protéinase non spécifique qui permet après un temps d'incubation d'une nuit à 37°C, de détruire les protéines liées à l'ADN (Annexe 2).

4.1.4 La précipitation d'ADN :

Mélanger en vortexant légèrement puis mettre les tubes dans un agitateur à 37°C et laisser agir tout une nuit, ensuite ajouter 2 à 3 fois le volume de la solution en éthanol absolu glacial, laisser agir pendant 15 mn.

Entre temps il faut préparer des tubes de 1.5ml avec les codes des patients et dans lesquels il faut ajouter 200 µl d'éthanol 80%, et à l'aide d'une pipette pasteur imbibée d'éthanol récupérer la méduse d'ADN et la mettre dans le tube 1.5ml mélanger en inversant 2 à 3 fois.

Après centrifuger le tube à 1000 rpm/mn pendant 3 mn, jeter l'éthanol dans la poubelle, égoutter le tube sur du papier essuie-tout propre afin d'éliminer toute trace d'éthanol laisser le tube sécher dans une étuve à 37° ou bien à l'air libre toute une nuit après avoir couvert le tube avec du parafilm.

Au final, ajouter 300 µl de T10E1 et laisser l'ADN se dissoudre, puis conservé à +4°C pour un stockage à court terme, et à -20°C pour une conservation à long terme.

4.1.5 Dosage et contrôle de la qualité de l'ADN :

Le contrôle qualité de l'ADN est une étape importante pour réussir la PCR elle se fait à l'aide d'un dispositif « NanoDrop » qui donne automatiquement la concentration ainsi que le rapport D.O 260nm/280nm. Le but de ce rapport est d'évaluer la qualité de nos échantillons ADN.

Si le ratio : $R = \frac{DO_{260nm}}{DO_{280nm}}$ est :

- Entre 1.7 et 1.9 : l'ADN est de bonne qualité ;
- à 2 : indique que l'ADN est contaminé par des traces de sel, dans ce cas un lavage à l'éthanol est nécessaire.
- < à 1.7 : indique que l'ADN est contaminé par des protéines ou des phénols.

Les contaminations nuisent à la qualité des séquences.

4.2 Amplification par PCR : Polymerase Chain Reaction:

La réaction de polymérisation en chaîne est une technique d'amplification enzymatique in vitro, mise au point par Kary Mullis (Prix Nobel 1993) au milieu des années 1980, qui permet à partir d'un fragment d'ADN minoritaire d'obtenir rapidement une quantité importante et exploitable de copies identiques à cette séquence spécifique.

Elle implique l'utilisation d'une paire d'amorces oligonucléotidiques synthétiques flanquant la séquence à amplifier.

4.2.1 Les outils de la PCR :

La PCR exige :

- **Molécule d'ADN d'intérêt** : généralement sous forme double brin, elle contient le fragment à amplifier.
- **Les désoxyribonucléotides triphosphates** (dNTPs: dATP, dGTP, dCTP, dTTP) sont des éléments constitutifs utilisés par la taq polymérase afin de synthétiser les brins d'ADN complémentaires.
- **MgCl₂ Chlorure de magnésium** : L'ion Mg²⁺⁺ est un cofacteur important de la Taq polymérase, il sert à stabiliser l'hybridation. Plus la concentration est élevée, plus l'hybridation est facilitée. La concentration du MgCl₂ doit être déterminée à chaque fois qu'un nouveau fragment est à amplifier.
- **Tampon** : utilisé pour tamponner le milieu réactionnel afin de maintenir des niveaux optimaux pour la Taq polymérase.
- **Les amorces sens et anti-sens** : encadrez la région à amplifier de sorte que chaque brin nouvellement synthétisé s'étende au-delà des amorces sur le brin opposé, créant de nouveaux sites de liaison sur ces brins d'amorce. Chaque amorce contient 18 à 30 nucléotides, pour assurer la spécificité de l'hybridation.
- **Taq polymérase** : ADN polymérase thermorésistante extraite de la bactérie *Thermus aquaticus*. Sa température optimale d'action est de 72°C, elle peut supporter un canal continu de 95°C, ce qui permet d'automatiser le procédé. Elle présente une activité exonucléasique 5' à 3', c'est-à-dire une fonction d'édition.
- **Adjuvant de type DMSO** : il est utilisé pour optimiser les réactions de PCR, notamment lorsqu'il s'agit de séquences riches en bases GC difficile à dénaturer. Ce réactif améliore la

dénaturation de deux brins d'ADN en favorisant le clivage des liaisons hydrogènes qui relient les deux brins de la molécule d'ADN.

- **H₂O.**

4.2.2 Le principe de la PCR :

Une PCR se décompose en trois étapes :

- **Dénaturation** : les deux brins d'ADN sont séparés par chauffage à 95 °C pendant 30 secondes à une minute.
- **Hybridation** : en abaissant la température (50-70 °C), des amorces constituées de courts fragments d'ADN reconnaissent leurs séquences complémentaires sur les brins d'ADN cible et s'hybrident chacune à son brin respectif, elle dure de 30 secondes à une minute.
- **Extension** : durant cette étape la température est augmentée à 72°C pendant 1 à 2 minutes, ce qui permet à la Taq polymérase de compléter la synthèse du brin d'ADN par l'ajout des nucléotides complémentaires aux amorces hybridées dans le sens 5' vers 3'.

En dernier, on porte le mélange à une température de 95°C pendant 20 secondes, afin que les régions d'ADN se séparent pour qu'ils servent à leur tour de matrice pour la prochaine synthèse.

Ces étapes de PCR sont répétées plusieurs fois au bout de n cycle pour avoir une quantité de 2n d'ADN. Et le produit sera conservé à T°= + 4°C.

4.2.3 Contrôle du produit de la PCR :

Une fois la réaction de la PCR est terminée, les produits sont testés sur un gel d'agarose à 2% ce qui signifie 2g d'agarose dans 100 ml de TBE 1% et BET. Le pourcentage d'agarose varie en fonction de la taille du fragment d'ADN à isolé.

Ajouter 8 µl de chaque produit PCR à 2 µl de bleu de dépôt (0,25% de xylène de Cyanol ; 0,25% de bleu de bromophénol et 30% de glycérol) afin de suivre la migration. Cette dernière est réalisée par électrophorèse horizontale (250A-160V) pendant 15 minutes dans une cuve contenant du TBE 1X et la lecture des résultats se fait en Gel DOC par le logiciel « Quantity One ».

4.2.4 Purification des produits de PCR :

La purification des produits PCR est réalisé afin d'éliminer l'excès d'amorces et des dNTPs.

4.3 Le séquençage de l'ADN :

Le séquençage de l'ADN est une technique qui a révolutionné la biologie moléculaire, elle permet l'identification de l'ordre d'enchaînement des nucléotides sur un fragment d'ADN cible ainsi la recherche des mutations.

Dans cette étude la technique de séquençage proposée est celle de Sanger.

4.3.1 La technique Sanger :

En 1977, Frederick Sanger et ses collègues ont développés une méthode de séquençage qui repose sur le procédé de la réplication de l'ADN, et qui consiste sur l'utilisation de nucléotides spécifiques nommées didésoxyribonucléotides (ddNTPs) qui bloquent la synthèse de l'ADN après leur incorporation par l'ADN polymérase. Ce blocage est dû à l'incapacité de ces nucléotides à former une liaison phosphodiester avec un autre nucléotide en raison de l'absence d'un groupe hydroxyle sur le carbone 3'.

Depuis 1977, les méthodes de séquençage ont considérablement évolué grâce à la mise au point de séquenceurs automatiques et de marquage des nucléotides à l'aide de fluorochromes. Mais malgré cela reste la technique Sanger reste une technique rapide et plus fiable et moins couteuse.

4.3.2 Le principe de Sanger :

Après la fragmentation de l'ADN en morceaux de quelques milliers de paires de bases ces derniers ont mis dans des tubes qui contient : les amorces, les 4 désoxyribonucléotides (dATP, dTTP, dCTP, dGTP), les 4 didésoxyribonucléotides (ddATP, ddTTP, ddCTP, ddGTP) marqués chacun par un fluorochrome différent dont le spectre d'émission est spécifique, la Taq polymérase, et un Tampon. (Figure 11).

L'ensemble subit une série de cycles de polymérisation au cours desquels l'ADN polymérase peut incorporer des désoxyribonucléotides ou des didésoxynucléotides à chaque nucléotide de la matrice d'ADN, un cas où l'ADN polymérase utilise aléatoirement les nucléotides présents dans le milieu pour copier le brin matrice en synthétisant un ADN de séquence complémentaire, du moment que la base (A, C, G ou T) est respectée.

L'autre cas l'ADN polymérase incorpore un didésoxynucléotide, la synthèse s'arrête et par conséquent on obtient des fragments de différentes tailles.

4.4 La purification des produits de réaction de séquence :

La purification des produits de réaction de séquence permet l'élimination des dNTPs et des didNTPs qui n'ont pas été incorporés.

4.4.1 L'électrophorèse capillaire :

L'électrophorèse capillaire est utilisée afin de séparer les produits de la réaction de séquence purifiés, ainsi que l'enregistrement et l'analyse spectrale des propriétés (du plus petit au plus grand) de la fluorescence spécifique du didNTP permettent d'assigner la base correspondante et la reconstruction de la séquence nucléotidique du brin matrice.

4.4.2 L'analyse des données :

Les signaux reçus par l'électrophorèse capillaire seront affichés sous forme d'électrophorégramme de pics de couleurs différentes : rouge pour le nucléotide T, bleu pour C, vert pour A et noir pour G. La lecture des séquences est faite par le logiciel informatique Sequencing Analysis v 6.0.

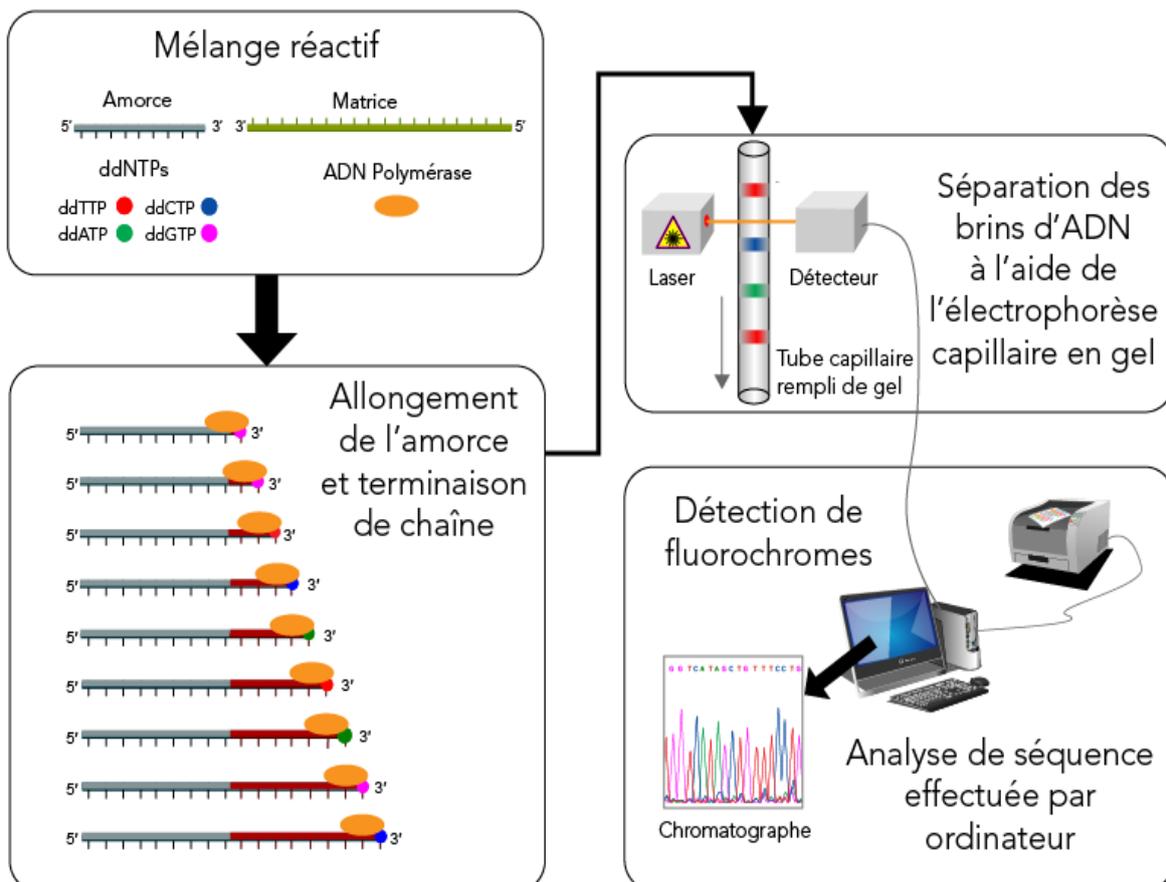


Figure 11: Les différentes étapes de la technique Sanger

4.5 Les résultats attendus :

Deux cas de figure sont possibles :

4.5.1 Premier cas :

Lors de la détection de la succession nucléotidique qui compose la séquence du gène facteur déterminant du testicule (*TDF*) qui se localise au niveau du chromosome Y, par la technique du séquençage Sanger prouve la présence de l'appareil génitale masculin donc l'échantillon d'ADN analysé appartient à un homme.

La succession nucléotidique du *TDF* : (adopté de la base de données NCBI)

> **NG_011751.1 :**

```
5'...AGGCATGTTTACACATTTTCATTAATTTAAGTACCAGAGAAGTATACTGAGAA
CCTACTACATGTCACACACTATCCTAGGAGCTGTAGACAGAAAATAATAAAATC
TCGTGTCTTTATTTAAATTTTGTGACTGCCACCCCAAGTGAGTCTCCTCACACA
CCCCGTCCTGTAATATATGCATCTGGGAGGTCTTTTTTGCCTTCTTAAAAACATAT
AGATGGTTGGACATATGTATATAAGAATATAAAATTCACCACTTTATCTTTTGTGA
ATGTGTGCTGTGAAGAACTCCTTTACTGGGGTGATGGAACCAGTGGCTACAAAGT
AAGGAGCTGGTTTACTGCTGTAAAGGGTTCGCGGCTTTGAATTTCAAGCTCTGGT
TCTGTGTCCTTGGGCACCTGCGCGTGAATCGTTGCCGCGAGGCTGGGCCAAGTTA
AGGCCCCACGCAGTTTGGCTTCCGGGCCAAGGAAGCCCCACAGGGTGCCCCACA
GGGTGAAGCCCCATGCCCTACAGGGTGAAGCGGCTGAAGCTGGTAGTGGTTCCG
AGGAAGCGGTCAAAGTCCCGCTCCAGAGGTTCTTCTTGGTTGTCACTCCCGG
AACCCCGCCAGGGGTCTGGC...3'
```

4.5.2 Deuxième cas :

L'analyse de l'échantillon d'ADN par la technique du séquençage Sanger ne détecte pas la présence de la succession nucléotidique du gène du facteur déterminant du testicule (*TDF*) qui signifie l'absence chromosome Y donc l'absence de l'appareil génitale masculin ce qui prouve que cet ADN appartient à une femme.

Matériels

et

Méthodes

Deuxième étude :

Détermination de

la taille

5 La détermination de la taille :

La taille humaine est un trait anthropométrique polygénique qui varie considérablement au sein des populations humaines ainsi qu'à travers le monde.

Pour identifier les variants génétiques influençant la taille humaine adulte, nous avons met en évidence l'étude de Yang en 2010 qui est considéré comme une réplication des études GWAS faite déjà par Weedon et Lettre et leurs camarades en 2007 et en 2008, et qui base sur l'utilisation des données de deux populations différentes caucasiennes américaines (apparenté et non apparenté) et chinoise.

5.1 Matériels biologique :

5.1.1 La collecte des échantillons :

La taille de chaque sujet a été mesurée sans chaussures à l'aide d'un stadiomètre mural standard dans la clinique par des infirmières (Tableau 3) (Annexe 3).

La population	Effectif	Référence
Caucasien non apparenté	998	(Yang et al., 2010)
Caucasien familial	8385	
chinoise	1638	

Tableau 3: Un tableau récapitulatif des données des deux populations différentes

5.2 Les Méthodes :

5.2.1 Génotypage :

L'ADN génomique a été extrait des leucocytes du sang périphérique en utilisant des protocoles standards. Pour les personnes non apparentées de race blanche et Échantillons chinois, le génotypage SNP a été réalisé en utilisant l'AffymetrixHumanMapping 500Karray set (Affymetrix, Santa Clara, CA, USA).

Dix-neuf SNP, dont rs1042725 dans le gène HMGA2, ont été génotypés avec succès dans tous les échantillons. SNP qui ont dévié de Hardy-Weinberg équilibre (HWE, $P < 0,01$) et avait une fréquence allélique mineure (MAF) $< 0,01$ ont été rejetés dans chaque ensemble d'échantillons.

5.3 Analyse statistique :

5.3.1 Test d'association :

Selon Yang et ses collègues en 2010 :

Pour les échantillons non liés, une régression linéaire mise en œuvre dans PLINK a été adaptée pour tester l'association en supposant un modèle d'héritage additif. Ils ont utilisé le génotype comme une covariable additive et la taille comme réponse, incluant simultanément l'âge et le sexe comme covariables dans le modèle de régression.

Pour l'échantillon basé sur la famille, des covariables significatives telles que l'âge et le sexe ont été utilisées pour ajuster les données brutes sur la taille. Résidus de hauteur étaient normalement distribués et ont été utilisés pour effectuer les tests d'association en utilisant la méthode QFAM implémentée dans PLINK. Les tests multiples ont été ajustés en adoptant la correction conservatrice de Bonferroni. Le seuil de signification a été fixé à une valeur P inférieure à $3,13 \times 10^{-3}$ (0,05/16 sous forme de 16 SNP ont été utilisés dans les analyses d'association). Les statistiques de la méta-analyse ont été générées à l'aide de la Z-scores pour quantifier les preuves globales d'association avec l'adulte variation de hauteur. Le Z-score individuel a été pondéré par la racine carrée de la taille de l'échantillon de chaque étude.

Ils ont ajouté ensemble les scores Z pondérés individuels dérivés de chaque échantillon et divisé par la racine carrée de la somme des tailles d'échantillon pour obtenir un score Z global et une valeur P combinée associée.

5.4 Résultats :

Les résultats de l'étude de Yang et ses collègues en 2017, sur des échantillons caucasiens non apparentés, caucasiens familial et échantillons chinois présentent une association significative de rs1042725 avec la taille, qui se réplique avec succès à la fois dans les échantillons caucasiens non apparentés ($P = 1,58 \times 10^{-3}$) et familiaux ($P = 3,0 \times 10^{-7}$), avec Chaque copie de l'allèle C à rs1042725 était associée à une augmentation de hauteur de $\sim 0,9$ et $0,7$ cm, respectivement.

En outre, le SNP rs1042725 dans HMGA2 était détecté comme nominalement significatif pour la taille adulte chez les chinois avec une valeur $P = 0,008$. Chaque copie de l'allèle C a été estimée d'être associée à une augmentation de la taille d'environ $0,6$ cm.

De plus de rs1042725, deux autres variants courantes, rs7968902 et rs7968682, a atteint le niveau de signification à la fois dans les échantillons caucasiens non apparentés et familiaux (rs7968902 : $P = 1,13 \times 10^{-3}$ et $3,4 \times 10^{-5}$, respectivement ; rs7968682 : $P = 1,55 \times 10^{-3}$ et $2,0 \times 10^{-7}$, respectivement. (Tableau 4)

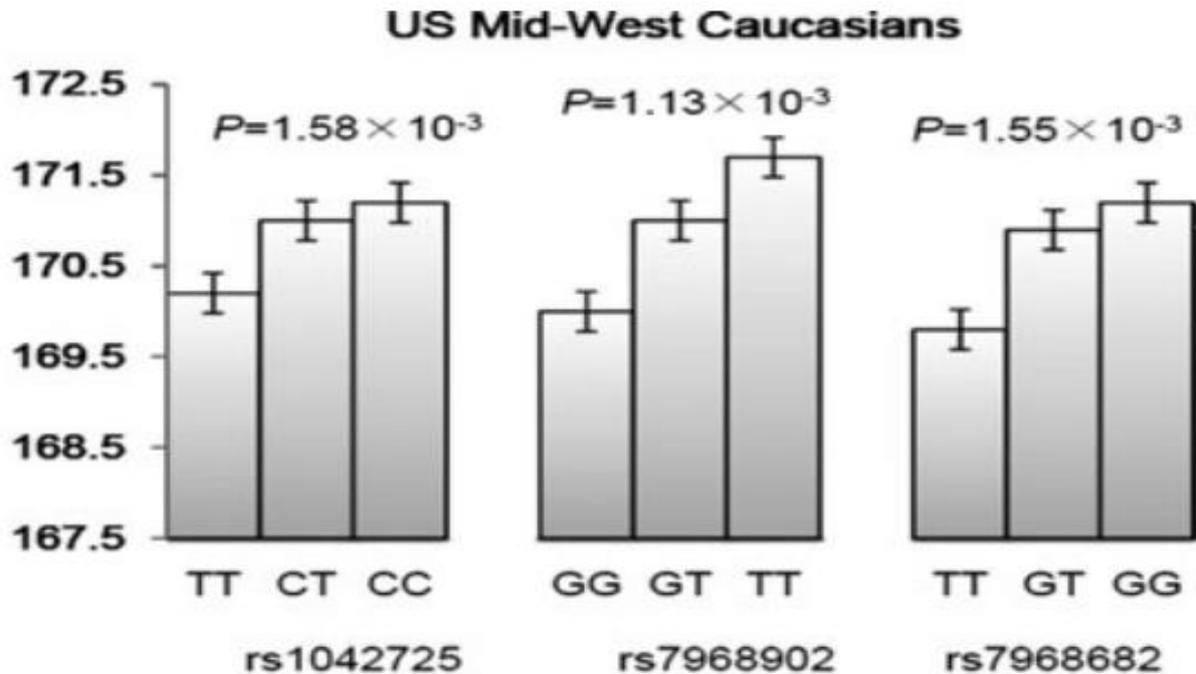
Par contre chez les chinois, le variant rs7968902 a montré une association légèrement significatif $P = 0.033$. Cependant, aucune association significative n'a été trouvée pour rs7968682. (Tableau 1)

Les différences phénotypiques chez les échantillons ayant des génotypes différents pour ces trois SNP sont illustrées en dessous (Figure 12)

Les résultats statistiques de la combinaison de tous les sujets caucasiens ensemble prouve l'association fort pour ces trois SNP (rs1042725 : $P = 4,25 \times 10^{-9}$; rs7968902 : $P = 6,34 \times 10^{-7}$; et rs7968682 : $P = 2,72 \times 10^{-9}$). La proportion de variance de taille expliquée par ces trois SNP était d'environ 1% , $1,07\%$ et $1,06\%$, respectivement, estimés en les Caucasiens non apparentés après ajustement pour l'âge et le sexe.

SNP	Physical position	Genic position	Alleles	Unrelated Caucasians			Family-based Caucasians			Chinese		
				MAF	P value	Effect size (cm)	MAF	P value	Effect size (cm)	MAF	P value	Effect size (cm)
rs7977687	64502340	5'UTR	G/A	0.041	0.171	0.975	0.045	0.133	0.485	0.196	0.120	0.647
rs2272046	64510728	intron2	G/T	0.028	0.425	0.699	0.025	0.041	0.840	0.088	0.571	0.192
rs12310312	64521025	intron3	A/G	0.045	0.084	1.172	0.062	0.030	0.619	0.212	0.029	0.524
rs2583937	64521731	intron3	C/T	0.069	0.931	0.051	0.097	0.804	0.056	0.114	0.618	0.154
rs2272047	64523002	intron3	C/T	0.049	0.084	1.147	0.058	0.027	0.641	0.211	0.030	0.241
rs2612060	64529769	intron3	G/A	0.068	0.823	0.132	0.099	0.923	0.022	0.115	0.788	0.082
rs17101839	64531835	intron3	T/C	0.024	0.115	1.500	0.049	0.023	0.754	-	-	-
rs11175944	64542662	intron3	C/G	0.064	0.045	1.149	0.089	3.65×10^{-3}	0.693	0.211	0.018	0.566
rs7959396	64543214	intron3	C/A	0.040	0.245	0.838	0.044	0.057	0.627	0.211	0.018	0.569
rs7973574	64558999	intron3	G/A	0.028	0.520	0.555	0.046	2.49×10^{-3}	1.141	0.036	0.412	0.434
rs10878346	64607140	intron3	A/G	0.226	0.223	0.420	0.246	1.0×10^{-6}	0.974	0.148	0.267	0.308
rs1156095	64613521	intron3	G/A	0.167	0.638	0.185	-	-	-	0.032	0.232	0.661
rs11175973	64620042	intron3	G/T	-	-	-	-	-	-	-	-	-
rs1480468	64628384	intron3	A/G	-	-	-	-	-	-	-	-	-
rs11175978	64629135	intron3	T/C	-	-	-	0.052	5.0×10^{-6}	1.993	0.092	0.025	0.763
rs1042725	64644614	3' UTR	T/C	0.499	1.58×10^{-3}	0.906	0.486	3.0×10^{-7}	0.683	0.803	0.008	0.614
rs7968902	64649337	3' UTR	T/G	0.429	1.13×10^{-3}	0.938	0.414	3.4×10^{-5}	0.533	0.153	0.033	0.586
rs1480464	64656838	3'downstream	T/C	0.196	0.153	0.518	0.190	0.379	0.142	0.252	0.064	0.418
rs7968682	64658147	3'downstream	G/T	0.499	1.55×10^{-3}	0.939	0.468	2.0×10^{-7}	0.761	0.107	0.167	0.438

Tableau 4: Tableau résume les résultats d'association pour 19 SNP dans HMGA2 avec la variation de taille adulte.



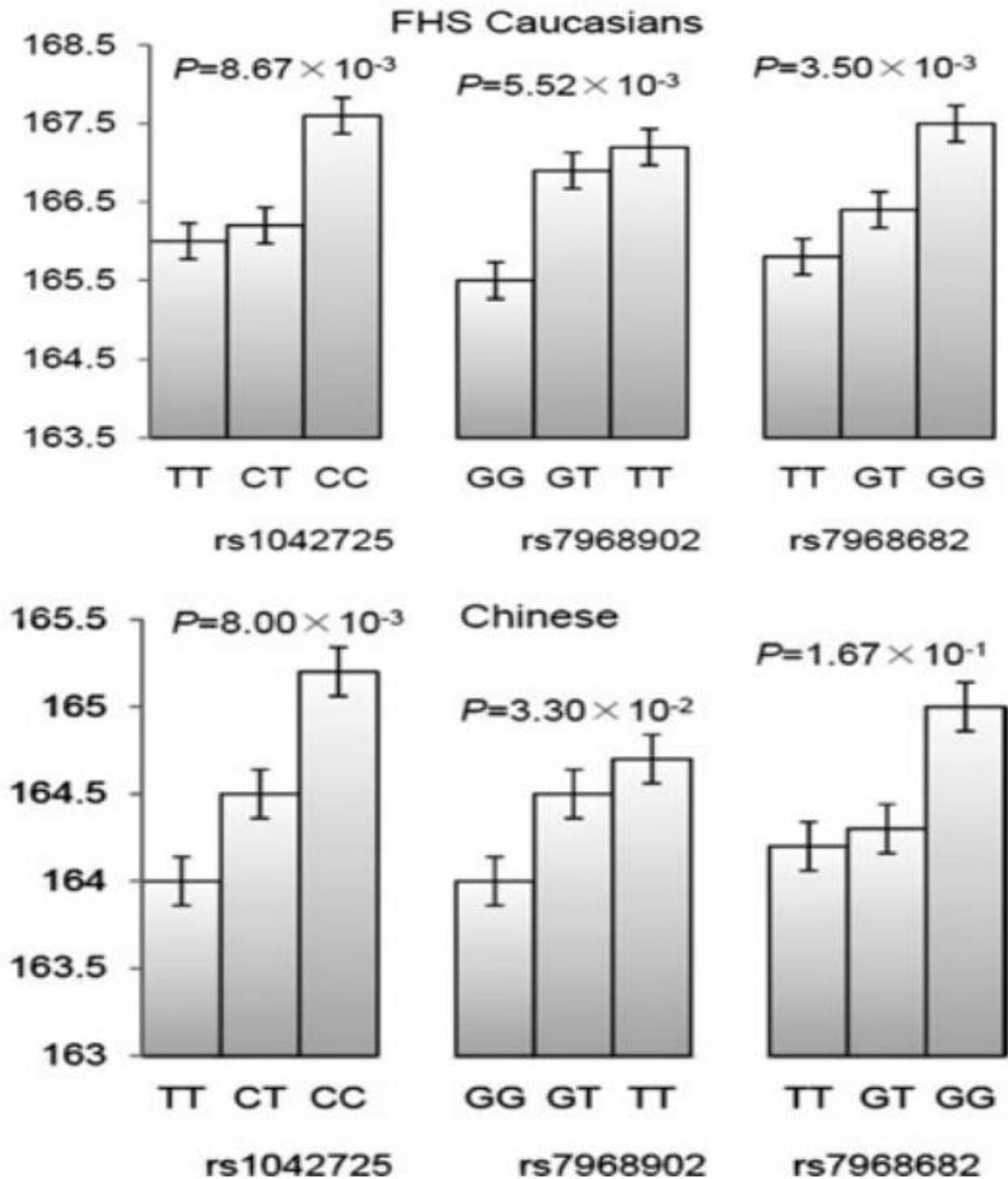


Figure 12: La comparaison des valeurs de taille chez les sujets avec différents génotypes pour les trois SNP significatifs chez les Caucasiens et chinois.

Matériels

et

Méthodes

Troisième étude :

Détermination de

la morphologie

des cheveux

6 Matériels :

6.1 Matériel biologique :

6.1.1 La collecte des échantillons et phénotype :

Selon trois études, des échantillons de 10493 ont été collectés à partir de trois populations différentes (Tableau 5) :

Population	Pays	Effectif	Effectif globale	Phénotype étudié	Référence de l'étude
Européen	La Pologne	ni = 528	N= 528	Raides Ondulés Bouclés	(Pośpiech et al., 2015)
Américaine latine	Chili	ni = 1745	N = 6357	Raides Ondulés Bouclés crépu	(Adhikari et al., 2016)
	Colombie	ni = 1507			
	Pérou	ni = 1247			
	Mexique	ni = 1207			
Les asiatique de l'est	Han chinois	ni = 2899	N = 3608	Raides Ondulés Bouclés	(Wu et al., 2016)
	Ouïghours	ni = 709			

Tableau 5: Un tableau récapitulatif des données des trois études choisies.

6.2 Méthodes :

6.2.1 Le génotypage et contrôle de qualité :

Des échantillons d'ADN des participants ont été génotypés à l'aide de système de Snapshot qui a été utilisé par Pośpiech et ses collègues en 2015, Puce Illumina HumanOmniExpress par Adhikari et ses adjoints en 2016 et Pucés Illumina HumanOmniZhongHua-8 utilisé par Wu et ses camarades en 2016.

6.2.2 Les analyses statistiques :

6.2.2.1 Le test d'association chez la population Européenne (EU) :

L'étude de Pospiech et ses camarades en 2015, centrée sur six SNP situés dans trois gènes *TCHH*, *WNT10A* et *FRAS1* initialement identifiés par Medland et ses collègues) avec un échantillon de 528 individus de Pologne. (Annexes 4).

Les SNP candidats ont été testés pour leur association avec la morphologie des cheveux, en utilisant la méthode de régression logistique binaire avec IBM SPSS Statistique v. 22 (SPSS Inc., Chicago, Illinois, États-Unis). (Pośpiech et al., 2015).

D'abord, une analyse d'association univariée pour les six SNP clés pour la morphologie des cheveux (rs17646946, rs11803731, rs4845418, rs12130862, rs1268789, rs7349332) a été faite, suivie d'une analyse de régression multivariée impliquant des tests simultanés des six SNP ainsi que les haplotypes *TCHH* inférés.

Au final, un test de l'effet de la combinaison de l'application des trois SNP les plus fortement associés sélectionnés à partir d'une analyse de régression multivariée a été évalué en estimant le score de risque génotypique (GRS). Le GRS résultant a ensuite été testé pour l'association avec la morphologie des cheveux en estimant les rapports de cotes, l'IC à 95 % et les valeurs P.

6.2.2.2 Le test d'association chez la population CANDELA :

Pour approfondir la compréhension de la base génétique de la variation des cheveux humains, (Adhikari et al., 2016), ont réalisé une étude d'association à l'échelle du génome (GWAS) chez plus de 6 357 Latino-Américains : des individus d'ascendance mixte européenne, amérindienne et africaine (Annexe 5) qui présentent une grande diversité génétique et une variation importante de l'apparence des cheveux .

Cette étude a porté sur le cuir chevelu : forme, couleur, grisonnement, calvitie et des poils du visage : épaisseur de la barbe, monosourcil, épaisseur des sourcils. Nous avons sélectionné que la partie qui s'intéresse aux types de cheveux .

PLINK 1.9 a été utilisé pour effectuer des tests d'association à l'échelle du génome pour chaque phénotype en utilisant une régression linéaire multiple avec un modèle génétique additif incorporant l'âge, le sexe et cinq PC génétiques comme covariables.

Une méta-analyse a été réalisée pour les SNP index identifiés dans le GWAS. En testant l'association séparément dans chaque échantillon de pays et combinant les résultats à l'aide du logiciel de méta-analyse METAL22.

Des parcelles forestières ont été produites avec MATLAB combinant toutes les régressions coefficients et erreurs types. La statistique Q de Cochran a été calculée pour chaque trait.

Afin de tester l'hétérogénéité de la taille de l'effet entre les échantillons de pays. Pour les SNP avec hétérogénéité significative, un modèle à effets aléatoires a été utilisé pour la méta-analyse.

6.2.2.3 Le test d'association chez la population chinoise :

Précédentes études d'association à grande échelle sur la raideur des cheveux se sont concentrées sur les populations européennes et latino-américaines et aucune analyse à l'échelle du génome de la raideur des cheveux n'avait été réalisée chez les Asiatiques de l'Est pour cela, Wu et ses collègues ont fait la première étude en 2016, sur la population de l'Est d'Asie qui comprenait 2899 Chinois Han et 709 Ouïghours.

Ils ont utilisé PLINK 1.7 pour effectuer les tests primaires d'association à l'échelle du génome utilisant la régression d'oreille linéaire multiple avec un modèle génétique additif incorporant le sexe, l'effet de lot et quatre CP génétiques comme covariables.

Les parcelles quantiles-quantiles et le degré de génomique l'inflation (λ) a été utilisée pour tous les tests d'association afin de tester pour les signes d'inflation.

Un test d'épistasie a également été réalisée à l'aide de PLINK 1.7 Les diagrammes quantile-quantile et Manhattan ont été créés dans R.

Ils ont utilisé un package R rehh pour effectuer une analyse de sélection positive. (Annexe 6)

Le cheveu est l'un des caractéristiques les plus remarquables de la variation humaine et l'un des ressources les plus importantes dans l'analyse médico-légale.

Les gènes de la morphologie des cheveux peuvent aider de manière médico-légale à caractériser des personnes connues ou inconnues et pour cela, nous avons mis en évidence trois études de Pospiech, Adhikari et Wu et leurs collègues basées sur l'identification des gènes responsables à la différence de la morphologie des cheveux chez trois populations différentes (EU/CANDELA/CHB).

6.3 Résultat :

6.3.1 Résultats de l'étude d'association chez les EU :

Les résultats montrent que les haplotypes des quatre SNPs rs12130862, rs17646946, rs11803731 et rs4845418 dans ou près de *TCHH* ont révélé une association de deux haplotypes TGAG le plus fréquent (72,3 %) avec des cheveux bouclés/ondulés, et de l'haplotype AATC (22,0 %) avec des cheveux lisses (Tableau 6). L'effet le plus fort obtenu pour l'haplotype AATC de OR = 2,1 (1,5–2,9), $P = 1,87 \times 10^{-5}$ et r^2 Nagelkerke = 5,1 %, était similaire à l'effet de rs11803731 OR = 2,0 (1,5–2,7), $P = 9,77 \times 10^{-6}$ et r^2 Nagelkerke = 5,4% (Tableau 7).

Les valeurs d'association enregistrées pour rs7349332 (*WNT10A*) et rs1268789 (*FRAS1*) se sont avérées inférieures à ceux des SNP dans *TCHH* avec OR = 1,6 (1,1–2,4), $P = 0,018$, r^2 Nagelkerke = 1,4% et OR = 1,4 (1,1–1,8), $P = 0,013$, r^2 Nagelkerke = 1,6%, respectivement (Tableau 7).

Analyse d'association multivariée, y compris simultanée l'analyse des six SNP et haplotypes *TCHH* a confirmé l'association avec la morphologie des cheveux pour *TCHH*, *WNT10A* et *FRAS1*, avec l'effet le plus fort noté pour rs11803731 (OR = 2,0, IC 95 % = 1,5–2,8, valeur $P = 1,12 \times 10^{-5}$) et des effets plus faibles révélés pour rs7349332 (OR = 1,7, IC à 95 % = 1,1–2,5, P value = 0,015) et rs1268789 (OR = 1,4, IC à 95 % = 1,0–1,8, P -value = 0,022), respectivement.

L'effet combiné de ces trois SNPs donne une signification de $P = 3,85 \times 10^{-7}$ et explique 8,2 % de la variation totale des cheveux chez les polonais (Tableau 8). Une analyse détaillée de combinaisons de génotypes pour trois loci et génotype sélectionnés combinaisons donnant la

plus grande confiance de prédiction ont été choisis. Quatre combinaisons de génotypes avec des fréquences entre 0,76 % et 13,26 % ont fourni une probabilité > 70 % de cheveux raides, la confiance la plus élevée de la prédiction était trouvée avec la combinaison TTGGGG donnant 82 à 88% de probabilités de cheveux raides, selon le type de modèle de prédiction (Figure 13).

Haplotype analysis for rs12130862, rs17646946, rs11803731, rs4845418 (TCHH)

No.	Haplotype	Frequency (%)	Association testing for straight vs. wavy/curly hair		
			OR (95% CI)	P-value	Nagelkerke r^2 statistic
1	TCAG	764 (72.3)	0.6 (0.4-0.8)	3.21×10^{-4}	3.5%
2	TCAC	1 (0.1)	NT	NT	NT
3	TAAG	9 (0.9)	0.3 (0.1-1.2)	0.079	NT
4	TATC	2 (0.2)	NT	NT	NT
5	ACAG	4 (0.4)	NT	NT	NT
6	AAAG	2 (0.2)	NT	NT	NT
7	AATC	41 (3.9)	1.3 (0.6-2.5)	0.471	NT
8	AATC	233 (22.0)	2.1 (1.5-2.9)	1.87×10^{-5}	5.1%
Total		1056 (100.0)	-	-	-

OR: Odds ratios estimated for straight hair; NT: not tested.

Tableau 6 : L'association de huit haplotypes avec la morphologie des cheveux obtenue à partir de l'échantillon de la population polonaise.

Les valeurs en gras indiquent les deux haplotypes les plus fortement associés aux cheveux lisses.

SNP	Gene	Chr	Chr position GRCh37.p13	Alleles	MAF	Effect allele ^a	Univariate association analysis		
							OR (95% CI) for the effect allele	P-value	Nagelkerke r ²
rs12130862	Near TCHH	1	152027015	A/T	A 0.27	A	1.8 (1.4-2.5)	7.73 × 10 ⁻⁵	4.3%
rs17646946	TCHHL1	1	152062767	G/A	A 0.27	A	1.8 (1.3-2.4)	1.02 × 10 ⁻⁴	4.1%
rs11803731	TCHH	1	152083325	A/T	T 0.26	T	2.0 (1.5-2.7)	9.77 × 10 ⁻⁶	5.4%
rs4845418	Near TCHH	1	152136230	C/G	C 0.22	C	2.1 (1.5-2.9)	1.76 × 10 ⁻⁵	5.1%
rs7349332	WNT10A	2	219756383	G/A	A 0.11	G	1.6 (1.1-2.4)	0.018	1.4%
rs1268789	FRAS1	4	79280693	G/A	A 0.31	G	1.4 (1.1-1.8)	0.013	1.6%

^a The allele associated with straight hair.

Tableau 7: Les résultats de l'analyse d'association univariée pour six SNPs prédictifs de la morphologie des cheveux.

SNP	Gene	Chr	Allele variants	Effect allele ^a	Multivariate association analysis	
					OR (95% CI) for the effect allele	P-value
rs11803731	TCHH	1	A/T	T	2.0 (1.5-2.8)	1.12 × 10 ⁻⁵
rs7349332	WNT10A	2	G/A	G	1.7 (1.1-2.5)	0.015
rs1268789	FRAS1	4	G/A	G	1.4 (1.0-1.8)	0.022
P-value of a model	3.85 × 10 ⁻⁷					
Nagelkerke r ² statistic	8.2%					
GRS ^b association	OR = 2.7, 95% CI = 1.9-3.9, P = 6.64 × 10 ⁻⁸					

Tableau 8: Les résultats de l'analyse d'association multivariée pour les SNP prédictifs de la morphologie des cheveux.

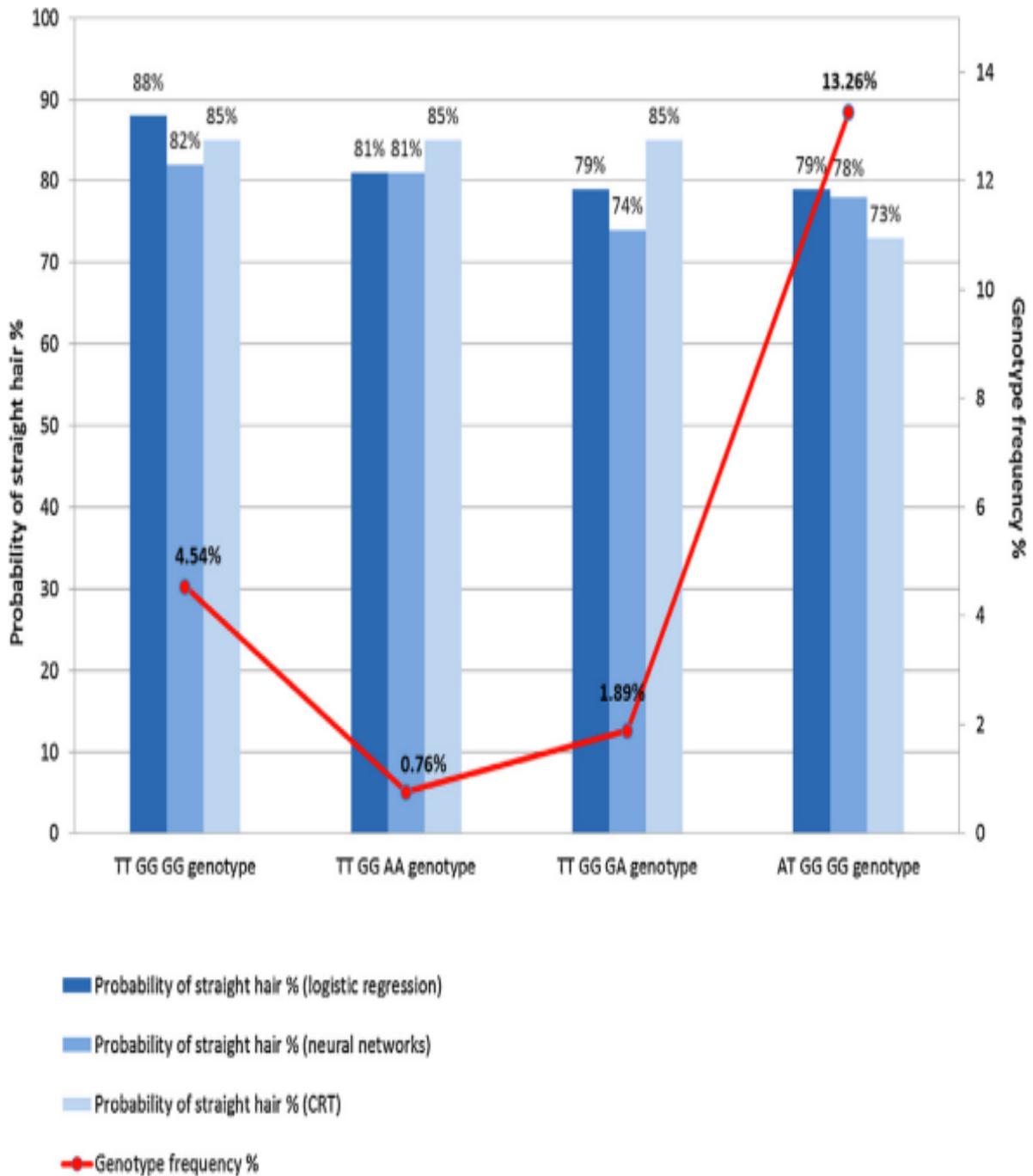


Figure 13 : La combinaison de génotypes de SNP rs11803731, rs7349332, rs1268789 qui donnent 70 % de probabilité pour les cheveux raides et leurs fréquences.

6.3.2 Les résultats de l'étude d'association chez la population CANDELA :

Les résultats d'examinations des sept caractéristiques capillaires dans l'échantillon de l'étude de la population CANDELA (Figure 14)

Les résultats d'association de GWAS identifier quatre SNPs significatifs associés à la morphologie des cheveux, sont : rs11803731 dans *TCHH* , rs3827760 dans *EDAR* , rs17143387 dans *GATA3* et rs11150606 dans *PRSS53* avec P value de : 3×10^{-12} ; 3×10^{-119} ; 4×10^{-8} ; 7×10^{-9} , successivement (Tableau 9) et (Figure 15).

Le résultat de la réexaminations des signaux d'association pour chaque index SNP dans chaque échantillon de pays séparément et ils ont combiné les résultats sous forme de méta-analyse, dénoter que pour chaque SNP les effets significatifs étaient dans la même direction dans tous les pays, la variabilité de la taille de l'effet entre les pays reflétant la taille de l'échantillon, (Figure 16).

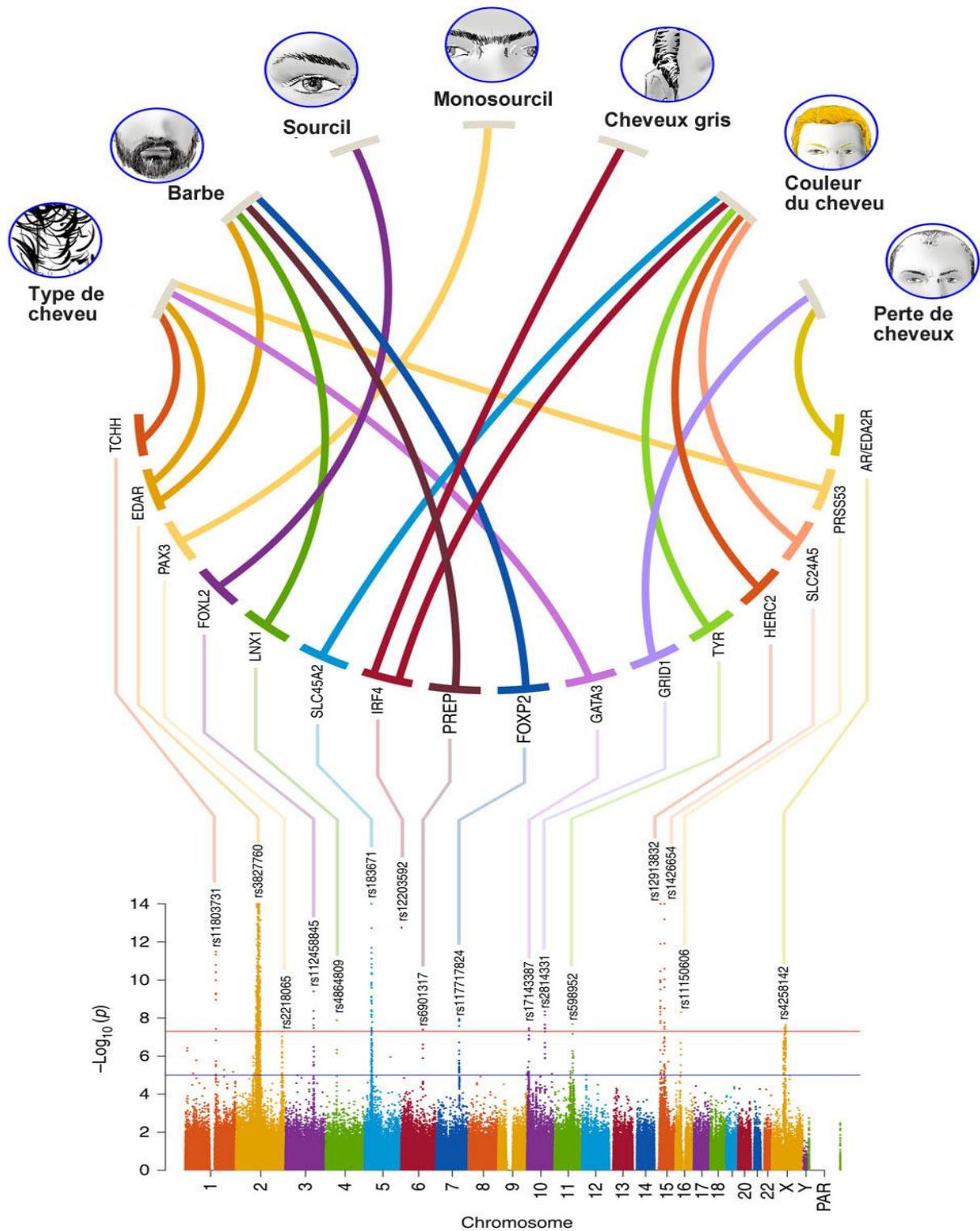


Figure 14: L’aperçu des résultats de l’étude d’association pangénomique (GWAS) des sept caractéristiques capillaires examinées dans l’échantillon de l’étude CANDELA.

En haut figurent des dessins illustrant les sept caractéristiques capillaires examinées dans l'échantillon de l'étude CANDELA.

Les lignes épaisses relient ces caractéristiques aux gènes candidats identifiés dans les régions où les SNP atteignent une association significative à l'échelle du génome.

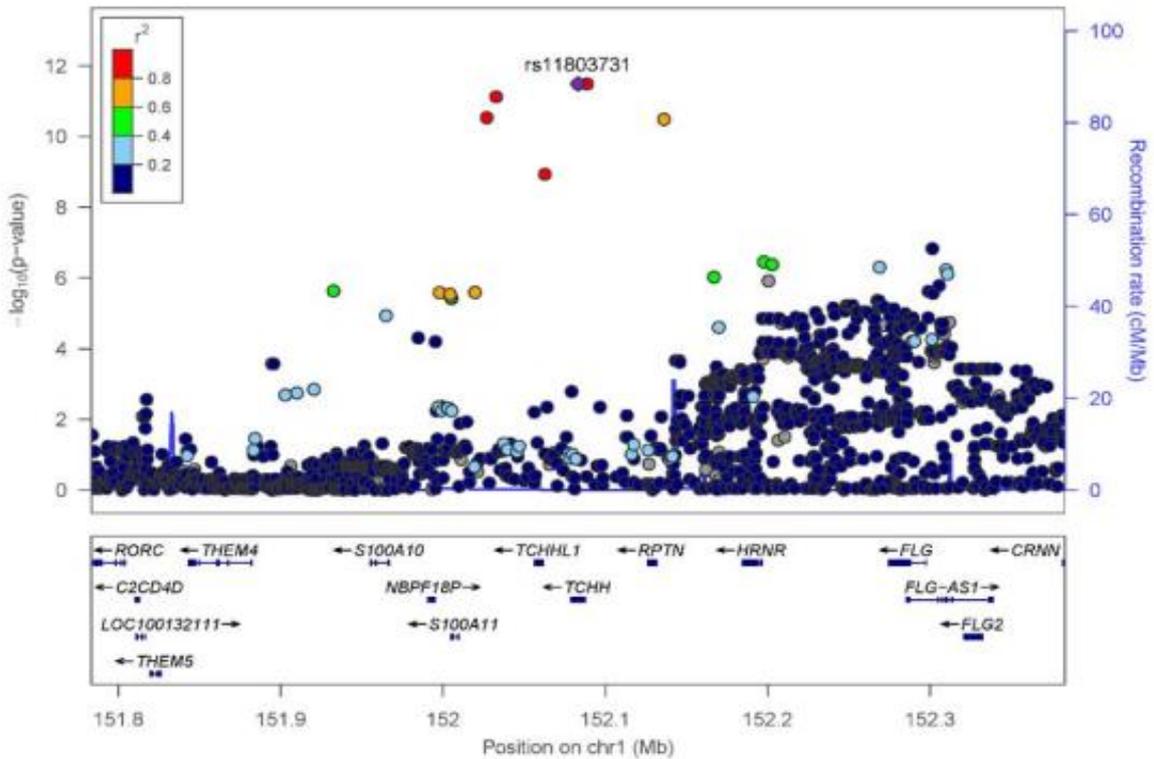
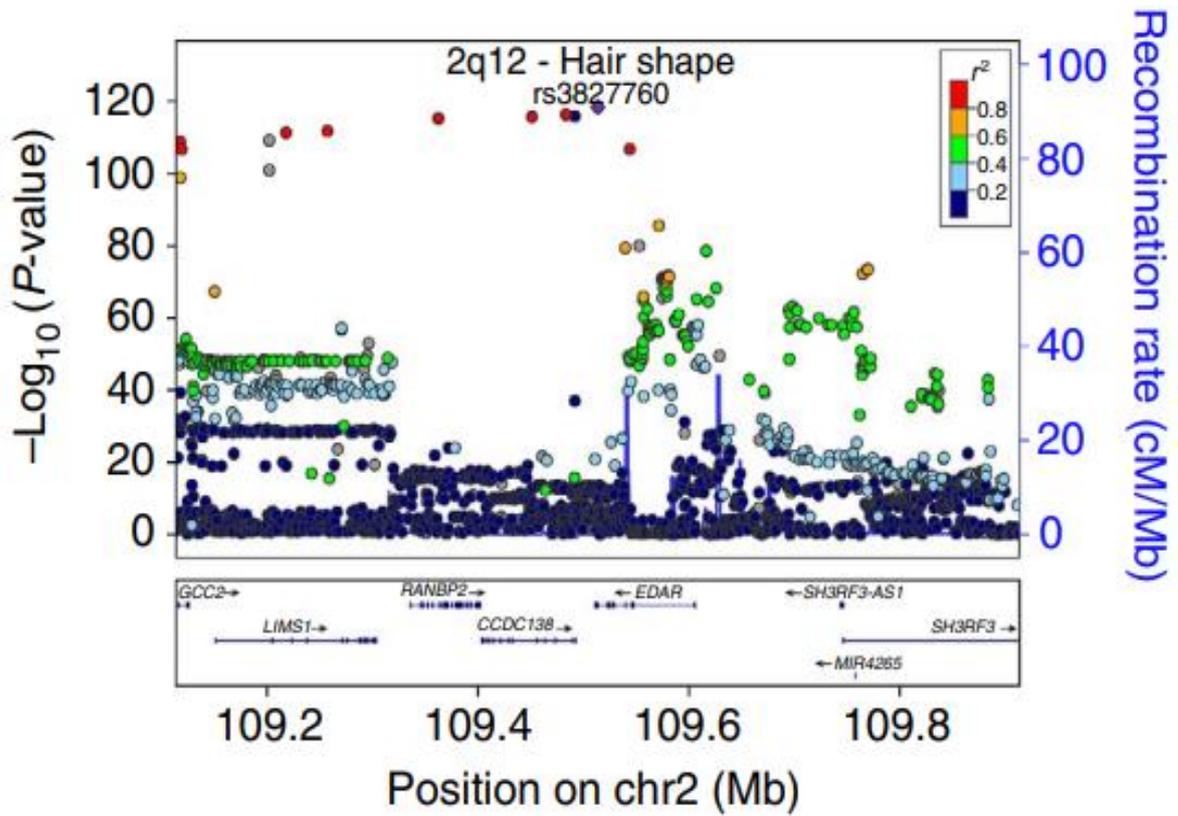
Au fond est ont montré un diagramme composite de Manhattan affichant tous les SNP significativement associés pour les caractéristiques capillaires examinées. Le nombre du SNP avec la plus petite valeur P est indiqué en haut de chaque pic d'association.

Les panneaux composites de cette figure et des suivantes ont été réalisés à l'aide de Photoshop.

Trait	Region	SNP	Closest gene	P value	Alleles ancestral > derived
Hair shape	1q21	rs11803731	<u>TCHH</u>	3×10^{-12}	A>T
Hair shape	2q12	rs3827760	<u>EDAR</u>	3×10^{-119}	A>G
Hair shape	10p14	rs17143387	GATA3	4×10^{-8}	T>G
Hair shape	16p11	rs11150606	<u>PRSS53</u>	7×10^{-9}	T>C
Beard thickness	2q12	rs365060	<u>EDAR</u>	1×10^{-15}	G>C
Beard thickness	4q12	rs4864809	LNX1	1×10^{-8}	A>G
Beard thickness	6q21	rs6901317	PREP	4×10^{-8}	G>T
Beard thickness	7q31	rs117717824	FOXP2	1×10^{-8}	G>T
Eyebrow thickness	3q22	rs112458845	FOXL2	4×10^{-10}	G>A
Monobrow	2q36	rs2218065	PAX3	5×10^{-8}	A>G
Hair greying	6p25	rs12203592	<u>IRF4</u>	3×10^{-13}	C>T
Hair colour	5p13	rs183671	SLC45A2	2×10^{-59}	T>G
Hair colour	6p25	rs12203592	IRF4	1×10^{-13}	C>T
Hair colour	11q14	rs598952	TYR	2×10^{-8}	T>A
Hair colour	15q13	rs12913832	<u>HERC2/OCA2</u>	5×10^{-104}	A>G
Hair colour	15q21	rs1426654	SLC24A5	1×10^{-18}	G>A
Balding	10q22	rs2814331	<u>GRID1</u>	4×10^{-9}	C>T
Balding	Xq12	rs4258142	AR/EDA2R	2×10^{-8}	C>T

SNP, single-nucleotide polymorphism.
Regions with intragenic index SNPs are shown in bold (the gene name is underlined if this index SNP is coding). SNPs in regions reaching only genome-wide suggestive association, P value $< 10^{-5}$, are listed in Supplementary Table 6.

Tableau 9 : Les caractéristiques des SNP index montrant la plus forte association significative à l'échelle du génome (valeur $P < 5 \times 10^{-8}$) avec le cuir chevelu et caractéristiques des poils du visage examinées dans l'échantillon CANDELA.



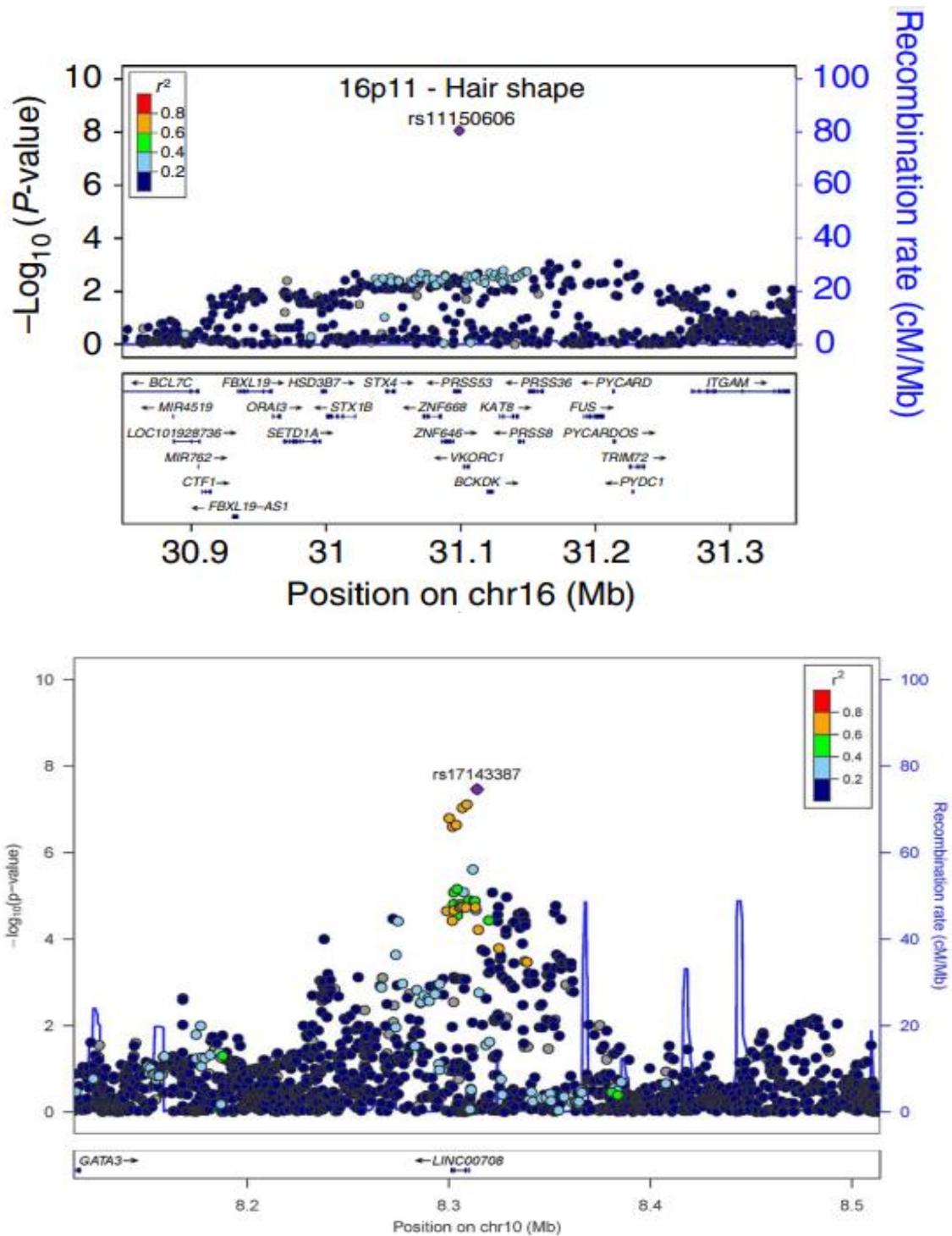


Figure 15 : Les graphiques d'association pour quatre régions avec des SNP montrant une association significative à l'échelle du génome avec la morphologie des cheveux.

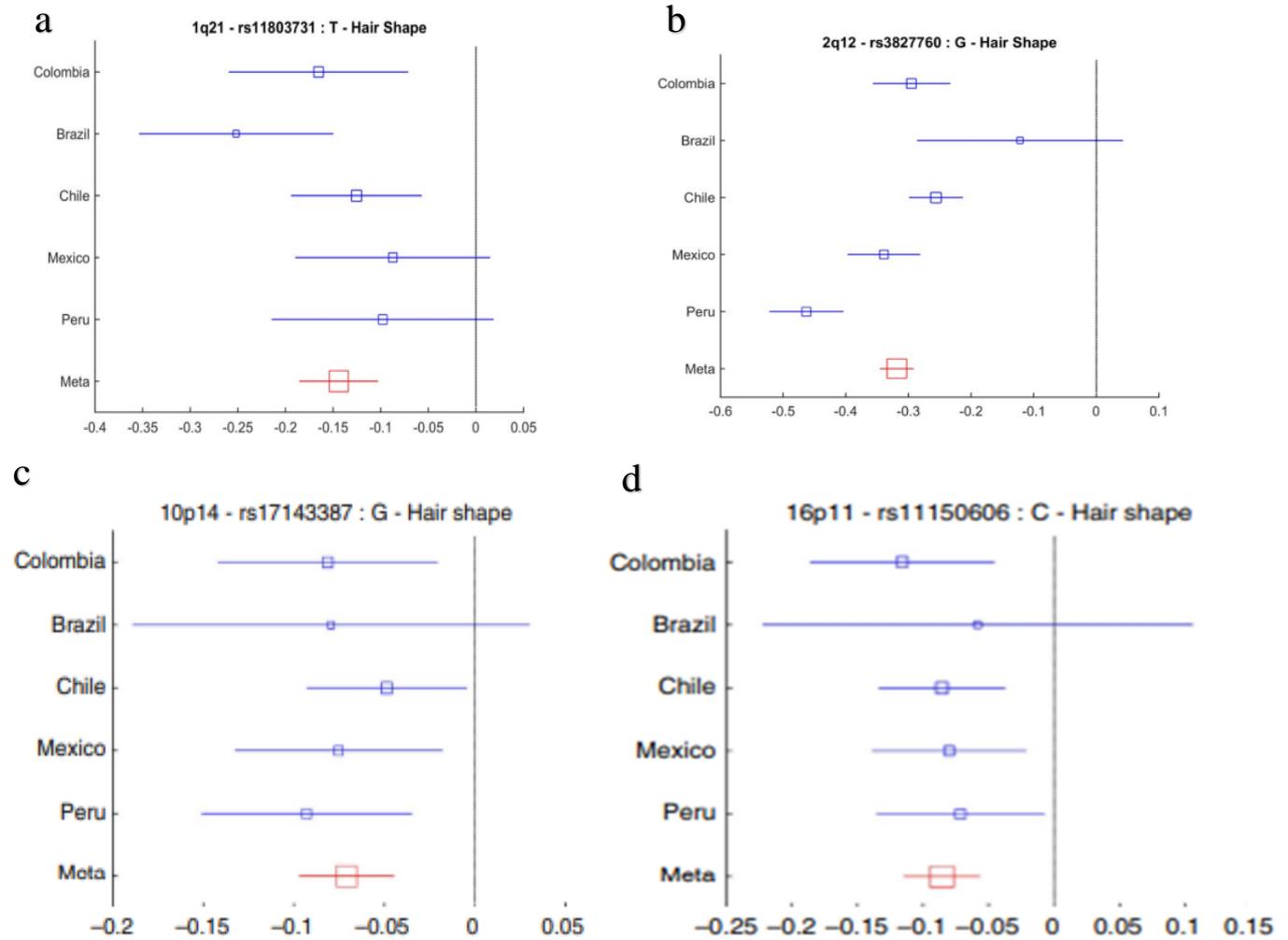


Figure 16 : Les tailles d'effet pour l'allèle dérivé au niveau des SNP index (Tableau 6) dans deux régions génomiques associées (a, b) et non associées auparavant à la forme des cheveux (c, d).

(a) forme de cheveux 10p14.

(c) forme de cheveux 1q21.

(b) forme de cheveux 16p11.

(d) forme de cheveux 2q12.

Les cases bleues représentent les coefficients de régression (axe x) estimés dans chaque pays. Les cases rouges représentent les tailles d'effet estimées dans la méta-analyse combinée. La taille des cases bleues est proportionnelle à la taille de l'échantillon. Les barres horizontales indiquent un intervalle de confiance à 95 % de largeur égale à $2 \times$ erreurs standard.

6.3.3 Résultat de l'étude d'association chez la population CHB :

Les résultats de l'analyse à l'échelle du génome ont révélé que l'*EDAR* était le principal gène associé à la raideur des cheveux chez les Chinois Han et les Ouïghours, avec rs3827760 : $P = 4,67 \times 10^{-16}$ (Figure 17) ; $P = 1,75 \times 10^{-12}$ (Figure 18), respectivement.

Wu et ses camarades ont trouvé aucune preuve d'épistasie entre le SNP *EDAR* et tout autre SNP à travers le génome.

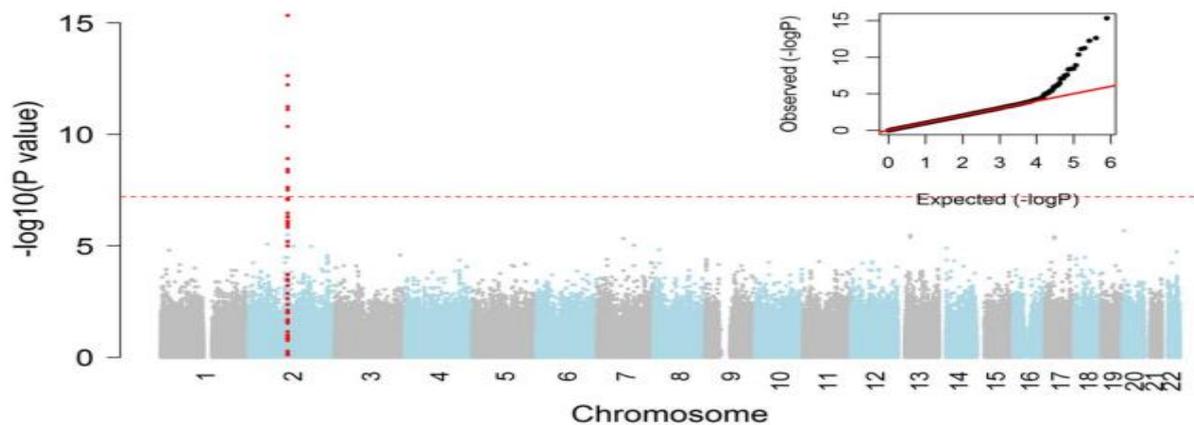


Figure 17: Le résultat du Diagramme de Manhattan et diagramme quantile-quantile montrant une analyse à l'échelle du génome de la raideur des cheveux chez les Chinois Han.

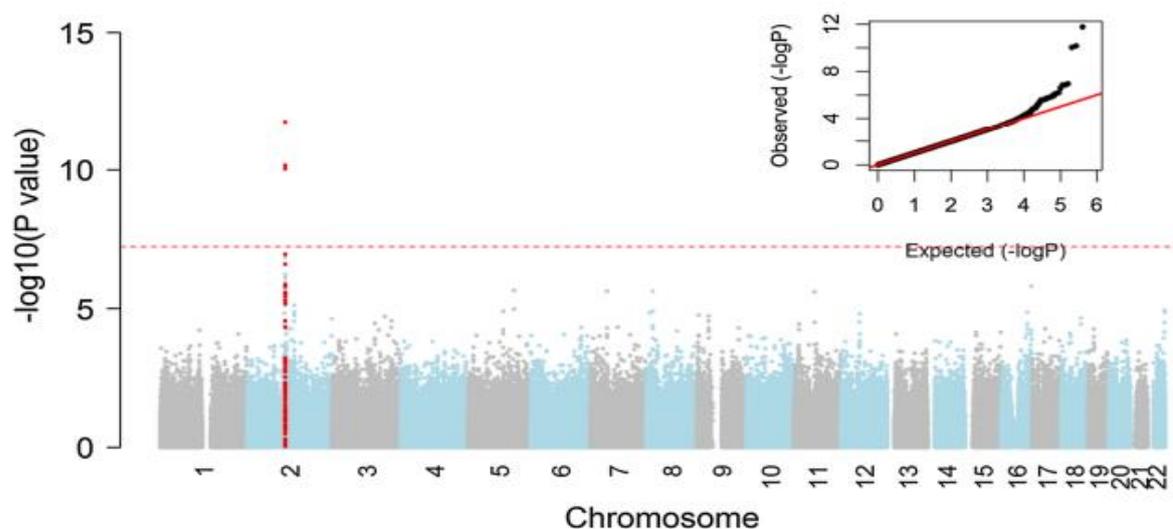


Figure 18: Le résultat du Diagramme de Manhattan et diagramme quantile-quantile montrant une analyse à l'échelle du génome de la raideur des cheveux chez les Ouïghours.

Discussion

Discussion :

La technique de séquençage de Sanger proposée, en plus d'être moins coûteuse par rapport aux autres techniques, elle permet l'identification de séquences nucléotidiques avec une longueur de lecture d'environ 1 Kb par séquence ce qui est compatible avec la longueur du gène *SRY*.

Chez les êtres vivants, la différenciation sexuelle est conditionnée par la présence ou l'absence du principal déterminant du sexe qui est le gène *SRY*, porté par le chromosome Y.

Lors du premier stade du développement embryonnaire, il n'y a aucune différence visible entre l'appareil génital féminin et masculin : on parle alors de gonades primordiales. Les gonades sont indifférenciées jusqu'à l'activation d'une protéine TDF (Testis Determining Factor). Cette protéine est produite par le gène *SRY* qui est considéré comme un initiateur des stades précoces du déterminisme testiculaire.

Chez l'homme, la protéine codée par *SRY* agit sur une population de la crête génitale et la transforme en cellules de Sertoli. Celles-ci envoient un signal aux autres lignages cellulaires dont la différenciation et l'organisation histologique aboutissent à la formation d'un testicule tandis que chez la femme, la détermination gonadique est plus tardive, Elle a lieu en l'absence du TDF ce qui met en place le développement ovarien et l'obtention du phénotype femelle.

La réplication des associations génétiques dans des populations indépendantes est essentielle pour réduire les résultats faussement positifs et explorer davantage le rôle de ces variantes dans les traits complexes. Cependant, ces études ont été réalisées par les mêmes groupes, et les échantillons ont été principalement du nord de l'Europe. Des répliques dans plusieurs populations supplémentaires par des groupes indépendants sont nécessaires pour évaluer la crédibilité et la généralité de cette constatation.

Puisque la génétique et l'environnement sont à l'origine des variations entre les différents groupes ethniques, par conséquent, le but de cette étude était de tenter de répliquer les associations entre les polymorphismes dans le gène *HMGA2* et la taille adulte à la fois chez la population caucasienne américaine et chinoise.

Le rapport de l'étude GWAS de Yang et ses collaborateurs ont réussi à découvrir que le rs1042725 du gène *HMGA2* est associé à la taille adulte dans les populations caucasiennes

américaines non apparentée et familiales et dans la population chinoise. Ils ont également identifié deux autres SNP significatifs rs7968902 et rs7968682 au niveau du même gène (*HMGA2*) dans la population caucasiennes (non apparentée et familiale) mais pas dans la population chinoise, donc il serait intéressant d'évaluer des études d'association GWAS dans des populations d'ascendance différente, puisque la variation génomique est plus grande lorsqu'elle est comparée entre les populations, ce qui peut accroître la crédibilité des résultats.

Il a été suggéré que le SNP rs1042725 du gène *HMGA2* a un effet sur la croissance longitudinale générale du squelette, il est considéré comme un candidat biologique fort pour la taille, où il a été observé que les humains présentant des mutations à fort impact dans le gène *HMGA2* avaient une petite ou une grande taille selon l'effet de la mutation. Tous ces trois SNPs sont localisés dans le 3-UTR de gène *HMGA2*, où la plupart des variants résident dans les régions non codantes du génome n'affectent pas directement la structure ou la fonction des protéines, mais modulent probablement l'expression d'un gène voisin, par contre les variants du gène *HMGA2* ils peuvent affecter l'ARNm de *HMGA2* de manière directe ou indirecte.

Les tests de génotypage GWAS traditionnels sont très précis pour les variants courants dans des régions non codantes, tandis qu'il est attirant que dernièrement GIANT ont mis en évidence des études GWAS qui concentrent sur l'étude des variantes rares dans des régions codantes qui peuvent affecter directement l'expression ou la fonction des protéines aux chondrocytes de la plaque de croissance ou d'autres types cellulaires liés à la taille humaine.

L'étude GWAS faite en 2015, par Posèpiech et ses collaborateurs sur des échantillons d'origine européenne a révélé que les quatre SNPs (rs12130862, rs17646946, rs11803731, rs4845418) du gène *TCHH* sont associées à la forme des cheveux. Parmi ses SNPs, le SNP exonique rs11803731 a montré une signification plus forte que les autres SNPs, sachant qu'il provoque le remplacement de l'acide aminé lysine par la méthionine dans la protéine *TCHH* (p.L790M) qui forme des complexes réticulés avec elle-même et avec les filaments intermédiaires de la kératine afin de fournir une résistance mécanique à la gaine interne de la racine du follicule pileux. Tandis que les trois SNPs restants n'ont pas fournis plus d'informations ce qui rend le SNP rs11803731 le variant fonctionnel le plus probable dans la détermination de la morphologie des cheveux.

L'étude a montré aussi que rs7349339 du gène *WNT10A* et le rs1268789 du gène *FRAS1* sont notés comme des signaux associés à la forme des cheveux chez les européens, mais avec une signification plus faible par rapport aux SNPs du gène *TCHH*.

L'effet combiné des SNPs (rs11803731, rs7349339, rs1268789) donne une explication à 8% de la variation de la morphologie des cheveux chez la population polonaise.

La combinaison génotypique des six SNPs précédents a sélectionné un meilleur prédicteur qui est le TTGGGG avec une probabilité supérieure à 80% d'avoir des cheveux raides.

D'autre part, l'étude sur la raideur des cheveux faite par Adhikari et ses adjoints sur un mélange d'européens, africains et d'américains (CANDELA) indique que le SNP rs11803731 du gène *TCHH*, le rs3827760 du gène *EDAR*, le rs17143387 du gène *GATA3* et le rs11150606 du gène *PRSS53* sont les SNPs les plus significatifs et les plus associés aux cheveux raides.

Par ailleurs, et suivant l'étude de Wu et ses collègues qui est considérée comme la première étude GWAS sur les cheveux raide chez les asiatiques de l'est et qui a révélé que le SNP rs3827760 du gène *EDAR* joue un rôle prédominant dans la raideur des cheveux. Ce variant mononucléotidique, code pour EDAR (pVal370Ala) qui est associée à des fibres capillaires plus épaisses.

Avec l'approche de ces trois études on révèle que les variants rs11803731 de gène *TCHH*, le rs7349339 du gène *WNT10A*, le rs1268789 du gène *FRAS1*, le rs3827760 du gène *EDAR*, le rs17143387 du gène *GATA3* et le rs11150606 du gène *PRSS53* sont associées à la raideur des cheveux chez les trois populations différentes (européenne, américaine et celle de l'Asie de l'est) où le variant de *TCHH* affecte les cheveux lisses dans les deux populations européennes et Latino-Américains mixtes. Le variant d'*EDAR* affecte la raideur des cheveux dans la population Latino-Américains mixtes et l'Asie de l'Est, alors que les variants des gènes *GATA3* et *PRSS53* sont spécifiques à la population Latino-Américaines seulement.

Ils n'ont également trouvé aucune interaction entre les gènes *TCHH* et *EDAR* dans leur effet sur la raideur des cheveux. Étant donné que la trichohyaline confère une résistance mécanique à la gaine de la racine interne du follicule pileux, tandis que le gène *EDAR* du récepteur de l'ectodysplasine A fait partie d'une voie importante affectant le développement des dérivés

ectodermiques, il est probable que l'*EDAR* et le *TCHH* affectent la raideur des cheveux par différents mécanismes.

L'apparition de différents gènes dans la population latino-américaine est probablement référer à la présence des échantillons dans leurs ancêtres sont mixtes (africain, européens et américains). La présence d'une hétérogénéité génétique substantielle entre diverses populations, et il est possible que la raideur des cheveux soit affectée par différents mécanismes dans d'autres populations.

Les études GWAS sur l'identification des gènes responsable à la variation de la morphologie des cheveux sont récentes et leurs résultats ne sont pas suffisants afin de déterminer ou de prédire la forme des cheveux en raison du nombre limités des échantillons ainsi que le pourcentage des échantillons ayant des cheveux lisse est élevé par rapport aux échantillons ayant d'autres types des cheveux.

Conclslion

Conclusion :

On conclut que le phénotypage et l'anthropologie avance jour en jour dans le domaine de la prédiction des caractères visibles extérieurs (EVC) des individus, où les chercheurs ont réussi à identifier quelques gènes responsables de la variation de la forme des cheveux, de la taille et du sexe des individus entre eux.

Les progrès de la biologie moléculaire et du séquençage de l'ADN ont permis de mieux développer et comprendre comment l'apparence d'un individu (son phénotype) est affectée de manière directe ou indirecte. Par conséquent les chercheurs ont pu arriver à déterminer le gène responsable de la différenciation du sexe qui est le facteur déterminant des testicules (TDF).

L'illustration la plus prometteuse de phénotypage médico-légale est largement limitée aux traits de pigmentation humaine. Les connaissances concernant les déterminants génétiques des caractères visibles extérieurs humains sont encore à leurs balbutiements et plusieurs études d'associations (GWAS) supplémentaires sont nécessaires pour identifier de nouveaux marqueurs dans diverse populations mondiales pas seulement dans les populations européennes et chinoises afin de prédire avec précision le phénotype d'un individu, c'est pour cela qu'aucun portrait perceptible a été fait pour le moment.

Annexes

Annexe :

Annexe 1 :

La différenciation sexuelle sous le contrôle des chromosomes sexués :

1. Le stade de la gonade indifférenciée :

Sur les 46 chromosomes humains deux sont dits sexués : le chromosome X et le chromosome Y. Ces chromosomes sont responsables d'une différenciation de genre, femelle pour l'individu porteur de la formule 46XX chromosomes et male pour l'individu porteur de la formule 46XY chromosomes. On remarquera que la nature a prévu la transmission d'un chromosome X dans les deux genres en provenance soit de la mère soit du père et que le chromosome Y n'est transmis qu'à l'individu à qui sera attribué le sexe masculin.

L'apparition du chromosome Y, qui se différencie des autres chromosomes en perdant du matériel génétique pour devenir plus petit, se concentre sur sa fonction de différenciation sexuelle par la présence d'un gène de différenciation nommé *SRY* «sex determining region of the Y chromosome » *SRY* permet la mise en place du testicule pendant la vie embryonnaire.

La sécrétion de testostérone par le testicule permet le développement des organes génitaux Internes à partir des canaux de Wolff (canaux excréteurs, épидидyme et canal déférent) et Externes avec la croissance du pénis et la descente des testicules dans les bourses. De plus, Par la sécrétion d'une hormone nommée «Anti-Müllerian Hormone » ou AMH, le testicule induit la régression des canaux de Müller et ainsi l'absence de développement des trompes Et de l'utérus (2). Tout est en place dès la naissance. Ainsi, à la naissance, la différence de Phénotype des organes génitaux est évidente et permet l'attribution du sexe masculin ou Féminin.

Le sexe de l'embryon est déterminé dès la fécondation, c'est à dire après fusion d'un Spermatozoïde porteur d'un unique chromosome X ou Y et d'un ovocyte comportant un seul Chromosome X. La première ébauche de gonade n'apparaît, dans l'espèce humaine, qu'à la Cinquième semaine de développement sous la forme d'une gonade indifférenciée qui garde Un bi-potentiel d'orientation.

Les cellules germinales primordiales sont à l'origine commune des spermatozoïdes et des Ovules et donc de la lignée germinale. Elles sont diploïdes avec 2n chromosomes, et se Multiplient par mitoses non seulement lors de leur migration, mais aussi dans la crête Génitale.

Plusieurs gènes jouent un rôle essentiel dans la formation des crêtes génitales, en particulier, *WT1* (Willems tumoral gène 1) et *SFI* (Steroidogenic Factor 1) qui codent pour des protéines impliquées dans la prolifération, la différenciation et la survie des cellules somatiques, en soutien de la gonade-bi potentielle. Elles vont permettre la synthèse des hormones stéroïdes des gonades (ovaire ou testicule) mais aussi de la corticosurrénale qui fabrique le cortisol connu sous le terme d'hormone du stress.

Dès la septième semaine de développement, chez les embryons humains de sexe génétique masculin (46, XY), la gonade indifférenciée commence à se transformer en testicule, alors que chez les embryons de sexe génétique féminin (46, XX), la transformation ne commence qu'à partir de la huitième semaine. Les ovaires ne sont reconnaissables qu'à la dixième semaine au moment de la prophase méiotique. A ce stade, la gaméto-genèse s'accompagne de la réduction du nombre de chromosome de 2n des cellules germinales à 1 n pour chacune des gamètes : spermatozoïde ou ovocyte.

Annexe 2 :

Service Des Maladies Congénitales Et Héritaires Hôpital Charles Nicolle 2016/2017. Protocole D'extraction D'ADN A Partir Du Sang Total :

JOUR 1 :

- Préparer des tubes Falcon avec les codes des patients
- Centrifuger les prélèvements sanguins à 3500 rpm/mn pendant 10 à 15 minutes
- A l'aide d'une pipette Pasteur (en plastique) récupérer le sérum en effectuant des mouvements circulaires sur les parois du tube (afin d'éviter la récupération de la couche des Globules blancs ou le Buffy coat)
- Jeter le sérum dans un bocal rempli d'eau de javel
- Prélever le maximum de la couche des globules blancs (GB) il est possible de prélever des Globules Rouges GR avec le Buffy coat (il est conseillé d'agiter légèrement le tube pour décoller les GR qui risquent de se coller sur les parois)
- Déposer la couche de GB dans le tube Falcon
- Ajouter qsp 45-50 ml de la Solution de Lyse des Globules Rouges (SLR)

- Rincer la pipette Pasteur avec les bulles d'air formées par la SLR afin de récupérer le max de GR
- Mélanger la solution 2 à 3 fois puis jeter la pipette Pasteur
- NB : Quand le volume de sang prélevé n'est pas important (<5ml) ou que le prélèvement est hémolysé mettre la totalité du prélèvement dans le tube Falcon et commencer à partir de l'étape 5
- Laisser agir le mélange (GB, SLR) 20 minutes à température ambiante
- Centrifuger les tubes à 3500 rpm/mn pendant 15 minutes puis jeter le surnageant
- Ajouter 20ml de la SLR et laisser agir pendant 20mn à température ambiante
- Centrifuger 3500 rpm/mn pendant 15 mn et jeter le surnageant
- Décoller le culot des GB en tapant légèrement le fond du tube Falcon
- NB : Procéder à l'étape suivante sinon garder le culot à +4°C
- Ajouter 15 ml du tampon de lyse des Globules Blancs
- Mélanger en vortexant
- Vérifier que tous les culots ont été dissouts
- Centrifuger à 3500 rpm/mn pendant 15mn puis jeter le surnageant
- Ajouter : 5ml de tampon de lyse des GB
- 250 µl de SDS 10%
- 100 µl de protéinase K
- Mélanger en vortexant légèrement
- Mettre les tubes dans un agitateur à 37°C et laisser agir tout une nuit
- NB : Les tubes peuvent être gardés dans une étuve à 56°C pendant 2h, n'oublier pas de mélanger de temps en temps

JOUR 2 :

- Retirer les tubes de l'agitateur
- Ajouter 2 à 3 fois le volume de la solution en éthanol absolu glacial
- Pour les culots 'pas très propres' commencer par ajouter un peu d'éthanol absolu glacial au culot mélanger puis compléter par l'éthanol glacial jusqu'à obtenir 2 à 3 fois le volume initial
- Laisser agir pendant 15 mn

- Préparer des tubes 1.5ml avec les codes des patients
- Mettre 200 µl d'éthanol 80% dans les tubes 1.5ml
- A l'aide d'une pipette pasteur (imbibée d'éthanol) récupérer la méduse d'ADN et la mettre dans le tube 1.5ml
- Mélanger en inversant 2 à 3 fois
- Centrifuger le tube à 1000 rpm/mn pendant 3 mn, jeter l'éthanol dans la poubelle
- Egoutter le tube sur du papier essuie-tout propre afin d'éliminer toute trace d'éthanol
- Laisser le tube sécher dans une étuve à 37°C sinon laisser sécher à l'air libre toute une nuit après avoir couvert le tube avec du parafilm

JOUR 3 :

- Ajouter 300 µl de T10E1 et laisser l'ADN se dissoudre
- Doser la solution finale d'ADN

Préparation de la solution de lyse des globules rouge (SLR) :

Pour la préparation de 2 Litres de SLR à 20x nous avons besoin de :

- NH₄CL (Ammonium Chlorure)
- KHCO₄ (Potassium Hydrogène Carbonate)
- EDTA (Ethylène Diamine Tétra-acétique Acide) (0.5 M ; PH= 7.5)
- H₂O Stérile qsp 2L
- Dans une éprouvette graduée mettre :
- NH₄CL 331,6 g
- KHCO₄ 40g
- EDTA 80 ml
- H₂O qsp 2L
- Autoclave la solution de SLR
- Diluer la solution mère avant chaque extraction

Préparation du tampon de lyse des globules blancs :

Pour la préparation de 1 Litre de Tampon de lyse des GB à 10x nous avons besoin de :

- 200 ml Tris Hcl (1M ; PH=7.5)

- 100 EDTA (0.5 M ; PH=7.5)
- H₂O qsp 1L
- Préparation du Tris HCl (1M ; PH= 7.5)
- 121g dans 1L de H₂O stérile -----> Ajuster le PH à l'aide du Hcl jusqu'à atteindre 7.5
- Préparation de l'EDTA (0.5M ; PH=7.5)
- 186.1g dans 1L de H₂O stérile -----> Ajuster le PH à l'aide de la NaOH jusqu'à atteindre 7.5

Annexe 3 :

Échantillon caucasien non apparenté :

Un échantillon aléatoire de 998 sujets sains non apparentés (500 femmes et 498 hommes) a été identifié à partir de notre base de données établie et en expansion contenant plus de 10 000 sujets. Tous les sujets identifiés étaient des Caucasiens américains d'Europe du Nord d'origine, vivant à Omaha, Nebraska et ses régions environnantes dans le Midwest des États-Unis.

Échantillon caucasien basé sur la famille :

L'échantillon familial provenait du Framingham HeartStudy (FHS), qui est une étude longitudinale de 14 277 sujets phénotypés pour identifier les facteurs de risque de maladies cardiovasculaires. Les sujets éligibles pour cette investigation sont issus du projet FHS SNP Health Association Resource (SHARe).

Nous avons les mesures de taille adulte sur 8385 sujets caucasiens phénotypés, 1307 de la cohorte originale (521 hommes et 786 femmes), 3189 de la cohorte de deuxième génération (1491 hommes et 1698 femmes) et 3889 de la cohorte de troisième génération (1821 hommes et 2068 femmes).

Échantillon chinois :

L'échantillon chinois était composé de 1638 sujets non apparentés, dont 810 hommes et 828 femmes. Les sujets ont été recrutés des adultes Han du nord de la Chine vivant dans la ville de Xi'an, Shanxi Province.

	Unrelated Caucasians	Family-based Caucasians	Chinese
Number	998	8385	1638
Gender (M/F)	498/500	3833/4552	810/828
Age (years)	50.3 (18.3)	45.5 (11.4)	34.5 (13.2)
Weight (kg)	80.2 (17.8)	76.5 (17.5)	60.2 (10.5)
Height (cm)	170.8 (9.8)	168.7 (9.5)	164.3 (8.2)

Note: Data are shown as mean (standard deviation, SD).

Tableau : Les caractéristiques des sujets d'étude

Annexe 4 :

les analyses d'association ont utilisé l'échantillon polonais de 333 individus aux cheveux raides et 195 aux cheveux bouclés/onduleux (75/195 et 120/195, respectivement). la figure suivante montre la répartition de la morphologie capillaire des cheveux bouclés, ondulés et raides dans l'échantillon d'étude polonais.

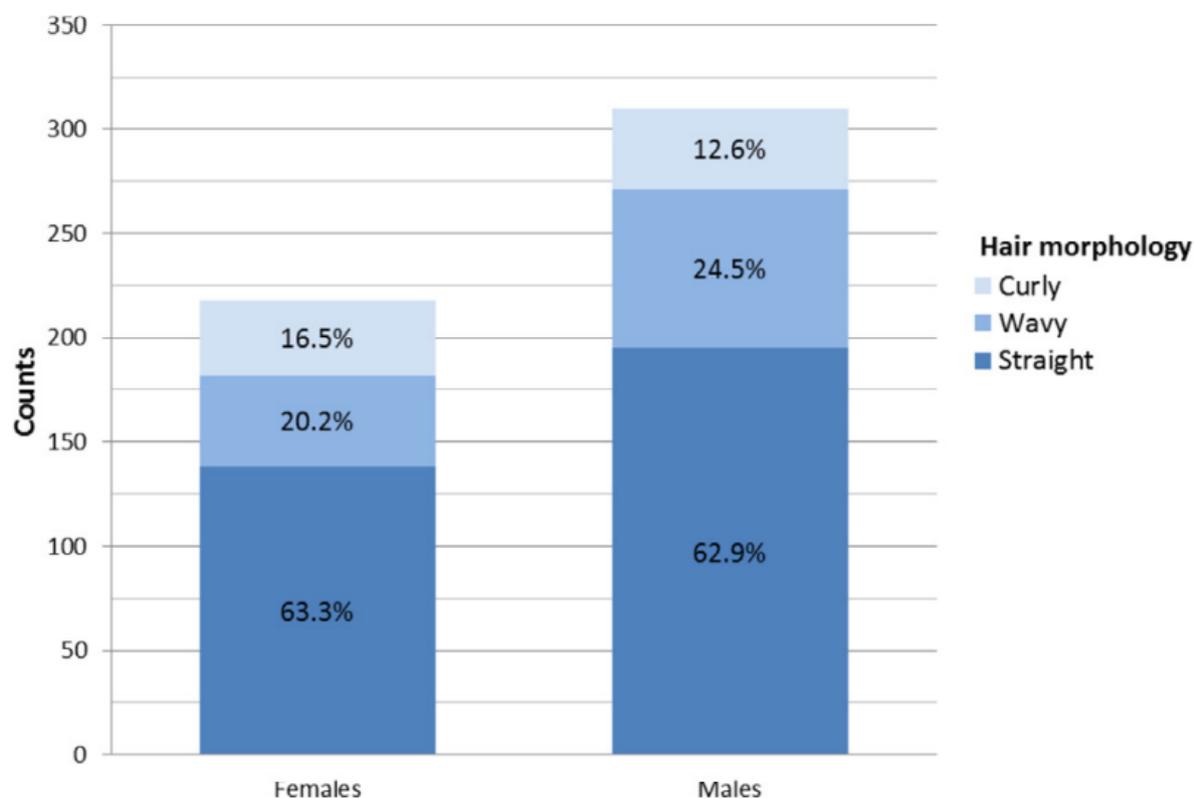
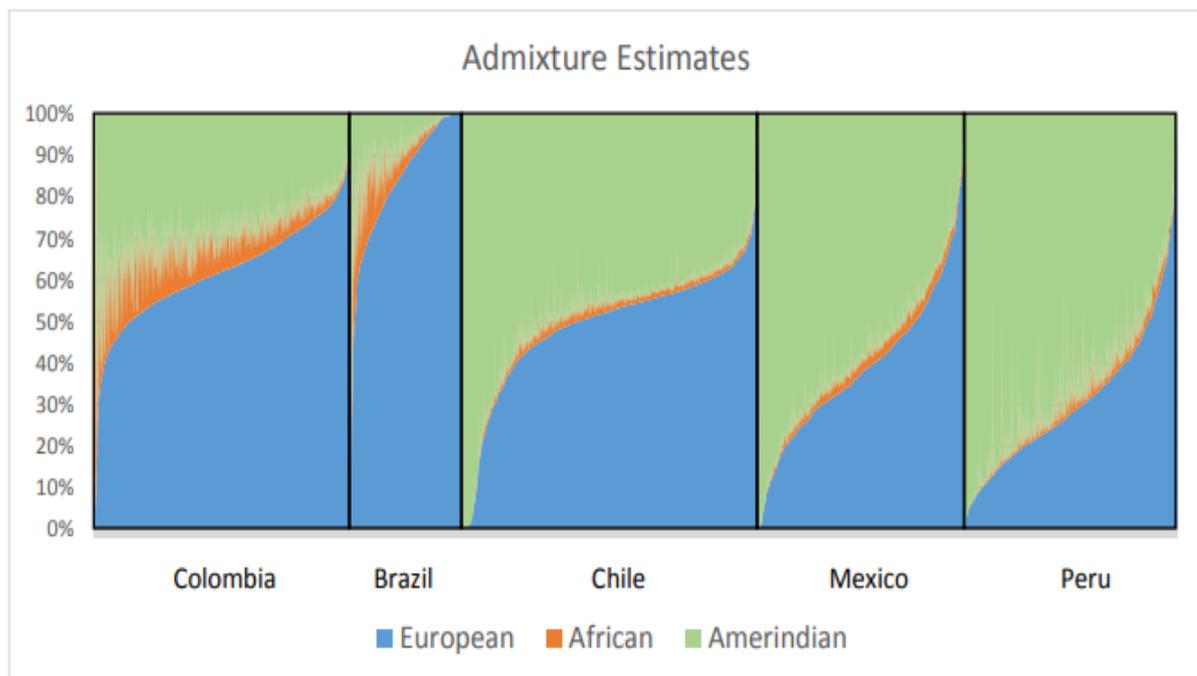


Figure :Distribution de la morphologie des cheveux dans l'échantillon d'étude polonais.

Annexe 5 :

Les barplots d'ascendance individuels pour chaque pays sont présentés ci-dessous. Les individus de chaque pays sont triés par augmentation de l'ascendance européenne.

Les proportions moyennes de mélange autosomique pour les échantillons de CANDELA ont été estimées comme suit : 48 % d'Européens, 46 % d'Amérindiens et 6 % d'Africains, mais avec une variation interindividuelle substantielle.



Mean ancestry estimates for each country and overall:

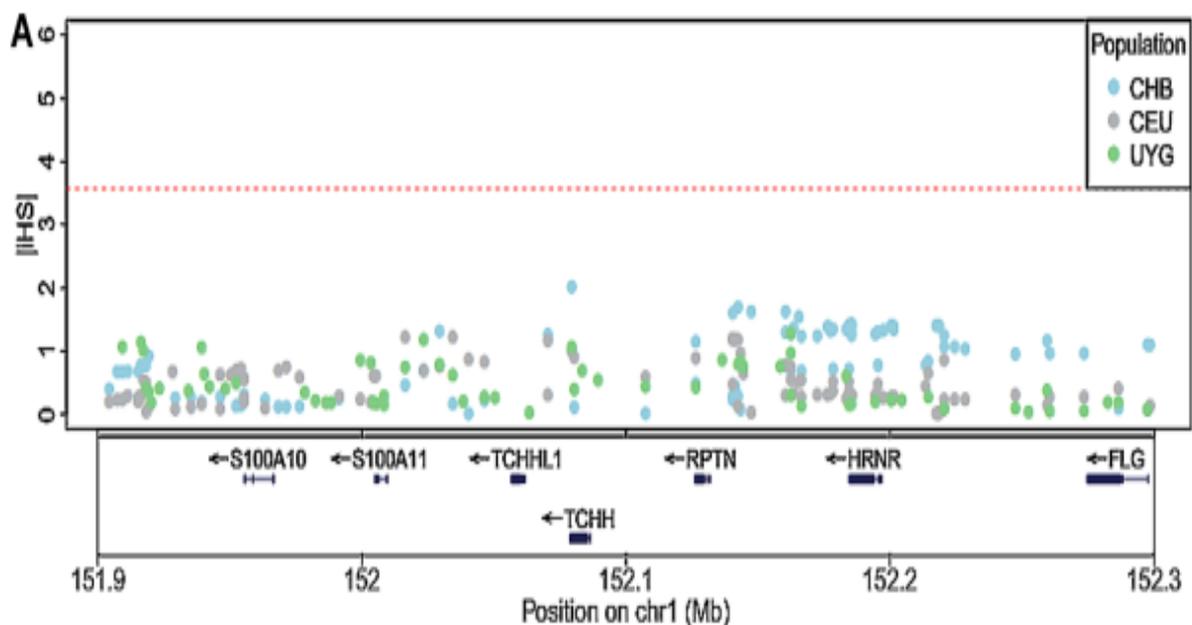
Country	African	European	Native American
Colombia	10%	61%	29%
Brazil	9%	79%	12%
Chile	5%	46%	49%
Mexico	5%	38%	58%
Peru	5%	31%	65%
Overall	6%	48%	46%

Annexe 6 :

Ils ont utilisé le package R *rehh* (Gautier et Vitalis 2012) pour effectuer une analyse de sélection positive (Sabeti et al. 2002).

Tout d'abord, ils ont calculé l'homozygotie étendue des haplotypes (EHH) pour tous les SNP jusqu'à $EHH < 0,05$ dans CEU, CHB et notre Échantillons ouïghours. Ensuite, le score d'haplotype intégré (iHS) a été calculé sur tous les SNP, avec un bin de fréquence allélique de 0,05 pour normaliser les scores iHS par rapport aux autres SNP de sa classe de fréquence dans la région. Enfin, nous avons calculé les valeurs P en supposant une distribution gaussienne des scores iHS sous le modèle neutre, qui a été vérifié en traçant les valeurs contre une distribution gaussienne (Voight et al. 2006).

Ils ont aussi effectué une analyse CMS à l'échelle du génome en Afrique (YRI), Populations européennes (CEU) et d'Asie de l'Est (JPT + CHB) du projet 1000 Genomes (Grossman et al. 2010) pour vérifier le résultat. utilisé HaploPS (Liu et al. 2013) pour détecter l'haplotype fondateur de *EDAR* et *TCHH* dans différentes populations en utilisant nos données et d'autres populations de les données de 1000 génomes, et d'identifier l'ensemble commun de SNP présents dans tous les haplotypes.



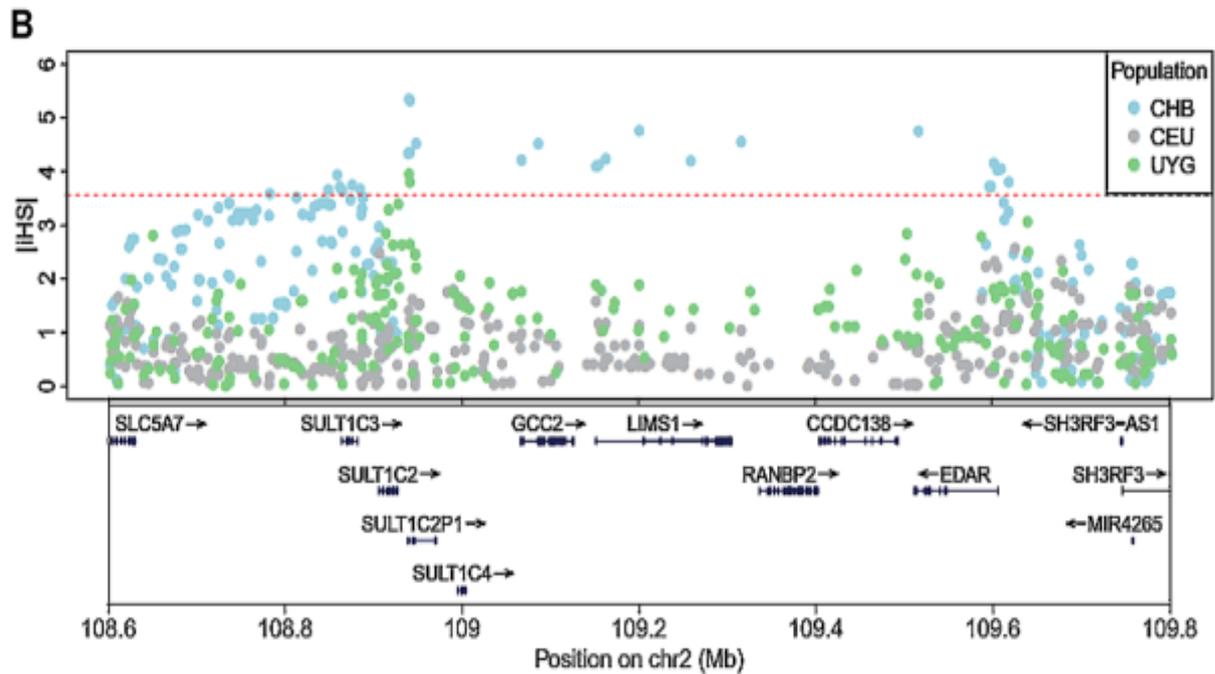


Figure : Signatures de sélection naturelle sur *TCHH* et *EDAR*

A : The $|iHS|$ results for the *TCHH* region.

B : The $|iHS|$ results for the *EDAR* region.

Nous avons constaté que *TCHH* n'est pas soumis à la sélection dans les populations CHB, UYG et CEU (la plupart des SNP dans la région *TCHH* $|iHS| < 2$), tandis que *EDAR* est sous forte sélection positive dans CHB mais pas dans CEU et UYG (il y a une apparente enrichissement pour les valeurs élevées $|iHS| > 3,56$ (top 0,1 %) dans CHB mais aucun pour CEU et UYG).

Références Bibliographiques

Référence :

- Adhikari, K., Fontanil, T., Cal, S., Mendoza-Revilla, J., Fuentes-Guajardo, M., Chacón-Duque, J.-C., Al-Saadi, F., Johansson, J. A., Quinto-Sanchez, M., Acuña-Alonzo, V., Jaramillo, C., Arias, W., Barquera Lozano, R., Macín Pérez, G., Gómez-Valdés, J., Villamil-Ramírez, H., Hunemeier, T., Ramallo, V., Silva de Cerqueira, C. C., ... Ruiz-Linares, A. (2016). A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. *Nature Communications*, 7(1), 10815. <https://doi.org/10.1038/ncomms10815>
- Alonso, L., & Fuchs, E. (2006). The hair cycle. *Journal of Cell Science*, 119(3), 391-393. <https://doi.org/10.1242/jcs.02793>
- Bernard, B. A. (2006). La vie révélée du follicule de cheveu humain. *médecine/sciences*, 22(2), 138-143. <https://doi.org/10.1051/medsci/2006222138>
- DAOUDI, K. (2005). Les alopecies et leur traitements [Thèse de doctorat, Université de Nantes].
- Gilgenkrantz, S. (2004). Au cœur de l'Y humain. *médecine/sciences*, 20(1), 17-18. <https://doi.org/10.1051/medsci/200420117>
- Gudbjartsson, D. F., Walters, G. B., Thorleifsson, G., Stefansson, H., Halldorsson, B. V., Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., & Steinberg, S. (2008). Many sequence variants affecting diversity of adult human height. *Nature genetics*, 40(5), 609-615.
- Gumpinger, A. C., Roqueiro, D., Grimm, D. G., & Borgwardt, K. M. (2018). Methods and Tools in Genome-wide Association Studies. In L. von Stechow & A. Santos Delgado

- (Éds.), *Computational Cell Biology : Methods and Protocols* (p. 93-136). Springer.
https://doi.org/10.1007/978-1-4939-8618-7_5
- Guo, M. H., Hirschhorn, J. N., & Dauber, A. (2018). Insights and Implications of Genome-Wide Association Studies of Height. *The Journal of Clinical Endocrinology & Metabolism*, *103*(9), 3155-3168. <https://doi.org/10.1210/jc.2018-01126>
- Gusmão, L., Brión, M., & Gomes, I. (2008). Chapter 30 The human Y chromosome male-specific polymorphisms and forensic genetics. In M. J. Bogusz (Éd.), *Handbook of Analytical Separations* (Vol. 6, p. 969-1000). Elsevier Science B.V.
[https://doi.org/10.1016/S1567-7192\(06\)06034-7](https://doi.org/10.1016/S1567-7192(06)06034-7)
- He, M., Xu, M., Zhang, B., Liang, J., Chen, P., Lee, J.-Y., Johnson, T. A., Li, H., Yang, X., Dai, J., Liang, L., Gui, L., Qi, Q., Huang, J., Li, Y., Adair, L. S., Aung, T., Cai, Q., Cheng, C.-Y., ... Qi, L. (2015). Meta-analysis of genome-wide association studies of adult height in East Asians identifies 17 novel loci. *Human Molecular Genetics*, *24*(6), 1791-1800.
<https://doi.org/10.1093/hmg/ddu583>
- Karimian, E., Chagin, A. S., & Sävendahl, L. (2012). Genetic Regulation of the Growth Plate. *Frontiers in Endocrinology*, *2*, 113. <https://doi.org/10.3389/fendo.2011.00113>
- KEYSER, C., & PETKOVSKI, E. (2006). Utilisation des SNP pour l'identification humaine. *Spectra analyse*, *35*(249), 24-29.
- Lamoril, J., Ameziane, N., Deybach, J.-C., Bouizegarène, P., & Bogard, M. (2008). Notions de génétique moléculaire pour comprendre l'hérédité. *Immuno-analyse & Biologie Spécialisée*, *23*(6), 331-352. <https://doi.org/10.1016/j.immbio.2008.10.005>
- Leroux, I. (1999). *Gènes de la détermination sexuelle : À propos de deux observations*. 7

- Mahner, M., & Kary, M. (1997). What Exactly Are Genomes, Genotypes and Phenotypes? And What About Phenomes? *Journal of Theoretical Biology*, 186(1), 55-63.
<https://doi.org/10.1006/jtbi.1996.0335>
- Marouli, E., Graff, M., Medina-Gomez, C., Lo, K. S., Wood, A. R., Kjaer, T. R., Fine, R. S., Lu, Y., Schurmann, C., Highland, H. M., Rieger, S., Thorleifsson, G., Justice, A. E., Lamparter, D., Stirrups, K. E., Turcot, V., Young, K. L., Winkler, T. W., Esko, T., ... Lettre, G. (2017). Rare and low-frequency coding variants alter human adult height. *Nature*, 542(7640), 186-190. <https://doi.org/10.1038/nature21039>
- Meng, H. T., Lan, Q., & Zhu, B. F. (2019). DNA Molecular Identification of Human Phenotypic Characteristics—New Progress Over the Past Five Years. *Fa Yi Xue Za Zhi*, 35(5), 512-518. <https://doi.org/10.12116/j.issn.1004-5619.2019.05.002>
- Mettrie, R. de la., Saint-Léger, Didier., Loussouarn, Geneviève., Garcel, A.-Lise., Porter, Crystal., & Langaney, André. (2007). Shape Variability and Classification of Human Hair : A Worldwide Approach. *Human Biology*, 79(3), 265-281.
<https://doi.org/10.1353/hub.2007.0045>
- Muthuirulan, P., & Capellini, T. D. (2019). Complex Phenotypes : Mechanisms Underlying Variation in Human Stature. *Current Osteoporosis Reports*, 17(5), 301-323.
<https://doi.org/10.1007/s11914-019-00527-9>
- Poulat, F., Barbara, P. D. S., Desclozeaux, M., Soullier, S., Moniot, B., Bonneaud, N., Boizet, B., & Berta, P. (1997). The human testis determining factor SRY binds a nuclear factor containing PDZ protein interaction domains. *Journal of Biological Chemistry*, 272(11), 7167-7172. <https://doi.org/10.1074/jbc.272.11.7167>

Références bibliographique

- Paus, R., & Cotsarelis, G. (1999). The Biology of Hair Follicles. *New England Journal of Medicine*, 341(7), 491-497. <https://doi.org/10.1056/NEJM199908123410706>
- Pham-Hoai, E., Pussiau, A., & Curci, J.-Y. (2017). ADN, une perpétuelle avancée. *Revue Francophone des Laboratoires*, 2017(489), 67-76. [https://doi.org/10.1016/S1773-035X\(17\)30031-X](https://doi.org/10.1016/S1773-035X(17)30031-X)
- Pośpiech, E., Karłowska-Pik, J., Marcińska, M., Abidi, S., Andersen, J. D., Berge, M. van den, Carracedo, Á., Eduardoff, M., Freire-Aradas, A., Morling, N., Sijen, T., Skowron, M., Söchtig, J., Syndercombe-Court, D., Weiler, N., Schneider, P. M., Ballard, D., Børsting, C., Parson, W., ... Branicki, W. (2015). Evaluation of the predictive capacity of DNA variants associated with straight hair in Europeans. *Forensic Science International: Genetics*, 19, 280-288. <https://doi.org/10.1016/j.fsigen.2015.09.004>
- Pugeat, M. (2021). Déterminisme sexuel, différenciation gonadique et identité du genre. *In Analysis*, 5(1), 54-61. <https://doi.org/10.1016/j.inan.2020.11.005>
- Ravel, C., Chantot-Bastaraud, S., McElreavey, K., & Siffroi, J.-P. (2006). Polymorphismes du chromosome Y et fertilité masculine. *Gynécologie Obstétrique & Fertilité*, 34(10), 885-893. <https://doi.org/10.1016/j.gyobfe.2006.07.032>
- Ravel, C., & Siffroi, J.-P. (2009). Chromosome Y et spermatogenèse. *Gynécologie Obstétrique & Fertilité*, 37(11), 901-907. <https://doi.org/10.1016/j.gyobfe.2009.09.009>
- Shastry, B. S. (2002). SNP alleles in human disease and evolution. *Journal of Human Genetics*, 47(11), 561-566. <https://doi.org/10.1007/s100380200086>
- Schneider, M. R., Schmidt-Ullrich, R., & Paus, R. (2009). The Hair Follicle as a Dynamic Miniorgan. *Current Biology*, 19(3), R132-R142. <https://doi.org/10.1016/j.cub.2008.12.005>

Références bibliographique

Sicot, G. (2006). *Étude du codage dans l'ADN* [These de doctorat, Télécom Bretagne].

<https://www.theses.fr/2006TELB0015>

Tozzo, P., Politi, C., Delicati, A., Gabbin, A., & Caenazzo, L. (2021). External visible characteristics prediction through SNPs analysis in the forensic setting : A review.

Frontiers in Bioscience-Landmark, 26(10), 828-850. <https://doi.org/10.52586/4991>

Visscher, P. M. (2008). Sizing up human height variation. *Nature Genetics*, 40(5), 489-490.

<https://doi.org/10.1038/ng0508-489>

Wu, S., Tan, J., Yang, Y., Peng, Q., Zhang, M., Li, J., Lu, D., Liu, Y., Lou, H., Feng, Q., Lu, Y., Guan, Y., Zhang, Z., Jiao, Y., Sabeti, P., Krutmann, J., Tang, K., Jin, L., Xu, S., & Wang, S. (2016). Genome-wide scans reveal variants at EDAR predominantly

affecting hair straightness in Han Chinese and Uyghur populations. *Human Genetics*, 135(11), 1279-1286. <https://doi.org/10.1007/s00439-016-1718-y>

Yang, T.-L., Guo, Y., Zhang, L.-S., Tian, Q., Yan, H., Guo, Y.-F., & Deng, H.-W. (2010).

HMGA2 Is Confirmed To Be Associated with Human Adult Height : Replication Between HMGA2 and Height. *Annals of Human Genetics*, 74(1), 11-16.

<https://doi.org/10.1111/j.1469-1809.2009.00555>

