

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITÉ SAAD DAHLAB- BLIDA 1
FACULTE DES SCIENCES DE LA NATURE ET DE LA VIE
DEPARTEMENT DE BIOLOGIE



Mémoire de fin d'études

En vue de l'obtention du Diplôme de master dans le domaine

Des sciences de la Nature et de la vie

Filière sciences biologiques

Option : Génétique

Thème :

**Identification faciale par machine learning à partir de quelques
gènes cible in silico**

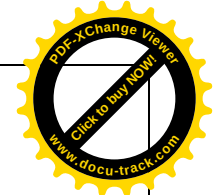
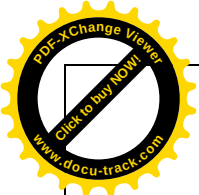
Date de soutenance : 08/07/2025

Présentés par : DJAHMOUN MAROUA et SADALLAH SALMA

Devant le jury :

Nom	Grade	Qualité
HAMZI W.	MCA/USDB1	Présidente
BELKHITER S.	MCA/USDB1	Examinatrice
MOHAMED SAID R.	MCA/USDB1	Promoteur

Promotion 2024-2025



Remerciement :

***En premier lieu, nous n'aurons pas pu faire ce travail sans bénédiction et puissance d'Allah
sounhanho wa taala.***

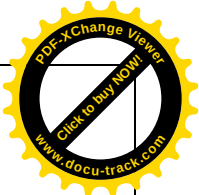
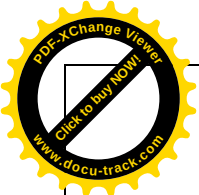
***Nous tenons tout particulièrement à remercier Monsieur MOHAMED SAID Ramdane, notre
Promoteur, pour son accompagnement tout au long de ce travail. Sa disponibilité, ses conseils
éclairés et sa bienveillance ont largement contribué à la réussite de ce mémoire***

Nous adressons également nos sincères remerciements aux membres du jury

Madame HAMZI de nous avoir honoré pour présider notre jury de soutenance et à

Madame BELKHITER d'avoir bien voulu et acceptée d'examiner et d'évaluer notre travail.

***Nos remerciements les plus sincères à tous nos enseignants du département de biologie et de la
faculté SNV***



Dédicaces de Maroua :

Il est difficile de trouver les mots juste pour exprimer toute la gratitude que je ressens. Ce travail est l'aboutissement d'un parcours jalonné de doutes, d'efforts, mais surtout de belles rencontres et de soutiens inestimables.

A tous ceux qui sont chers, ceux à qui je dois mon succès :

A notre « seigneur, Dieu tout puissant », merci dieu de m'avoir donné la vie, la foi, les prières exaucées pour m'y arriver.

A celle dont ses paroles m'accompagnent depuis que j'étais petite, la premier Femme Forte, courageuse qui a souffert sans me laisser souffrir, celle qui m'a appris à être une fille forte et qui cherche toujours à achever ses objectifs à tout prix ; celle qui m'a encouragé à poursuivre mes rêves même les obstacles semblaient insurmontables, Mon ange gardien ton amour et ta présence dans ma vie et le fait d'être ma Mère continueront d'éclairer mon chemin

A celui qui m'a dit un jour que je suis la prune de ses yeux, et qui voyait dans mon plus simple succès la plus grande réalisation, à l'homme qui fait toujours de son mieux pour être à mes côtés et me faire sentir qu'il est mon soutien constant malgré toutes les épreuves, celui qui a fait de moi ce que je suis aujourd'hui car il est mon Père, je ferai de mon mieux pour que tu sois fier de moi

A ma chère sœur Safaa « ullah yrhmha », l'âme de mon âme, ma source d'inspiration dans ma vie ; celle qui m'a apporté le véritable sens de la fraternité et les moments de bonheur, tu resteras toujours dans mon cœur et j'espère que tu es fière de ta grande sœur, ce mémoire te rend hommage à toi qui as tant compté pour moi

A mes chers frères Amine et Riad, je n'oublierai jamais vos encouragements permanents et soutiens moraux, votre présence dans mon cœur est un trésor précieux

A ma grande mère Zineb, ma deuxième maman, Merci de m'avoir élevée et aimée comme l'un de tes propres enfants, c'est ta prière pour moi qui m'a aidée à arriver jusqu'ici

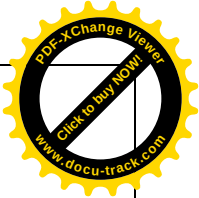
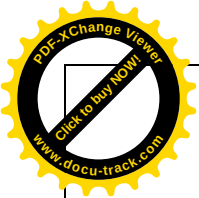
A ma tante Naima merci pour tes encouragements et ta présence

A mon chère témoin des étapes de ma vie, dans mon succès, joie, tristesse : Othmane merci pour l'énergie positive que tu apportes dans mon existence

A mes chères amies : Achouak, Hiba, Selma, Ines, Maroua votre amitié inestimable a illuminé mon parcours et enrichi cette expérience

A ma Binôme Sadallah Salma merci d'être la meilleure version de toi-même, ta contribution a été essentielle à cette expérience

Enfin je dédie ce travail à moi-même, c'est le temps d'y être !



Dédicaces de Salma :

J'ai l'honneur de dédier ce modeste travail réalisé grâce à l'aide d'Allah le tout puissant

À mes chers parents, Pour leur amour inconditionnel, leur soutien constant et leurs sacrifices sans limites, les piliers de ma vie, ceux qui ont cru en moi même lorsque moi, je doutais.

Votre amour, vos sacrifices, vos prières silencieuses et votre patience infinie ont été les fondations de mon parcours. Merci d'être tout simplement vous. Rien n'aurait été possible sans vous

À ma grande sœur Nessrine, pour sa présence réconfortante, ses encouragements et ses conseils toujours justes. Et à ses enfants, Alaa et Amir, mes petits trésors, mes amours, qui illuminent ma vie de leur innocence et de leur tendresse.

À mes frères, Youcef et Mohamed, pour leur affection sincère, leur complicité et leur soutien discret qui m'a portée dans les moments de fatigue

À ma chère grand-mère Houria, dont les prières, la tendresse et la sagesse m'accompagnent à chaque étape de ma vie

À Bouchra, mon amie et ma sœur de cœur, présente dans les joies comme dans les épreuves, merci pour ton écoute, ton soutien et ta loyauté

À Achouak, une amie précieuse, douce et lumineuse, dont l'amitié m'est chère et dont les encouragements m'ont souvent redonné l'élan nécessaire

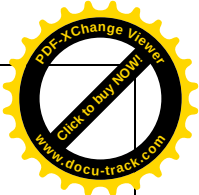
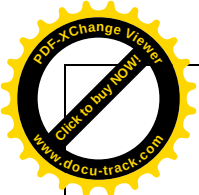
À Maroua, mon amie et ma binôme de mémoire, avec qui j'ai partagé cette belle aventure pleine de défis. Merci pour ta bienveillance, ton sérieux, ta patience et ton amitié précieuse.

À mes cousines Hayet, Khadidja et aya, pour les rires partagés, l'affection sincère et les souvenirs inoubliables, et à toute la famille Sadallah

Et à moi-même,

pour avoir cru, persévéré, et tenu bon face aux doutes et aux obstacles. Pour les longues nuits de travail, les efforts silencieux, les sacrifices et le courage de ne jamais abandonner. Aujourd'hui, je suis fière du chemin parcouru.

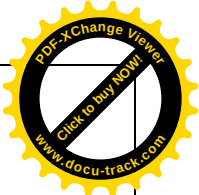
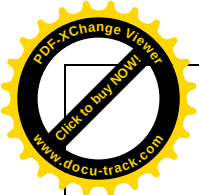
Ce mémoire est bien plus qu'un simple document : c'est un morceau de moi, soutenu par tout l'amour et la force des miens.



Résume :

La morphologie faciale, un trait complexe sous contrôle de la génétique. Mendel propose que les **traits** soient définis par des "éléments héréditaires" appelés **gènes**. Les gènes apparaissent sous différentes versions ou expressions, qu'on désigne par **allèles**. Les **allèles dominants** sont exprimés au détriment des **allèles récessifs**. Les allèles récessifs ne s'expriment que lorsqu'ils sont présents tous les deux. De nombreuses recherches ont identifié plusieurs gènes associés à divers caractères faciaux. Cette présente étude va essayer à partir d'une trace biologique de prédire et reconnaître le faciès d'une personne en utilisant son profil génétique. Ceci dit c'est à partir d'un ensemble restreint de gènes clés impliqués dans la pigmentation (comme OCA2, HERC2, TYR) et la morphologie faciale, que nous essayons de développer un système intelligent capable de prédire les traits du visage humain. Un système intelligent a été développé pour interpréter ces données génétiques et produire des prédictions visuelles, L'implémentation de ce système a été réalisée en langage Python, à l'aide de bibliothèques spécialisées en science des données et en apprentissage automatique. Ce système basé sur la pigmentation des yeux particulièrement, est un essai pour les autres traits faciaux qui par modélisation 3D de la morphologie faciale va donc nous donner la prédiction d'un profil inconnu en faciès connu. Les résultats obtenus, bien qu'exploratoires et limités à quelques traits, démontrent la faisabilité technique de cette approche et constituent une preuve de concept solide. Cette étude ouvre ainsi des perspectives prometteuses pour le développement d'outils plus complets d'identification faciale à visée biométrique, médico-légale ou anthropologique.

Mots-clés : morphologie faciale, génétique, allèles, pigmentation, OCA2, HERC2, TYR, profil génétique, prédiction, Python, modélisation 3D, biométrie



Abstract :

Facial morphology, a complex trait controlled by genetics. Mendel proposed that traits are defined by “hereditary elements” called genes. Genes appear in different versions or expressions, known as alleles. Dominant alleles are expressed at the expense of recessive alleles. Recessive alleles are only expressed when both are present. Numerous studies have identified several genes associated with various facial characteristics. This study will attempt to predict and recognize a person's facial features using their genetic profile based on biological traces. That said, we are attempting to develop an intelligent system capable of predicting human facial features based on a limited set of key genes involved in pigmentation (such as OCA2, HERC2, TYR) and facial morphology. An intelligent system has been developed to interpret this genetic data and produce visual predictions.

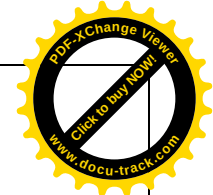
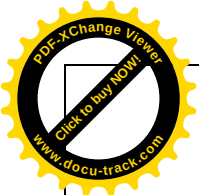
The system was implemented in Python, using libraries specialized in data science and machine learning. This system, based specifically on eye pigmentation, is a trial for other facial features which, through 3D modeling of facial morphology, will give us the prediction of an unknown profile in a known face. The results obtained, although exploratory and limited to a few features, demonstrate the technical feasibility of this approach and constitute solid proof of concept. This study thus opens up promising prospects for the development of more comprehensive facial identification tools for biometric, forensic, or anthropological purposes.

Keywords : facial morphology, genetics, alleles, pigmentation, OCA2, HERC2, TYR, prediction, Python, 3D modeling, biometrics

ملخص

تشكل الوجه، سمة معقدة تخضع للسيطرة الجينية. يقترح مندل أن السمات يتم تحديدها بواسطة "عناصر وراثية" تسمى الجينات. تظهر الجينات في أشكال أو تعبيرات مختلفة، يشار إليها باسم الأليلات. الأليلات السائدة تظهر على حساب الأليلات المتنحية. الأليلات المتنحية لا تظهر إلا عندما تكون موجودة كلاهما. حددت العديد من الأبحاث عدة جينات مرتبطة بسمات وجهية مختلفة. ستحاول هذه الدراسة، انطلاقاً من أثر بيولوجي، التنبؤ بالملامح الوجهية لشخص ما والتعرف عليها باستخدام ملفه الجيني. ومع ذلك، فإننا نحاول تطوير نظام ذكي قادر على التنبؤ بملامح الوجه البشري انطلاقاً من مجموعة محدودة من الجينات الرئيسية المرتبطة بالتصبغ) مثل OCA2 و HERC2 و TYR ومورفولوجيا الوجه. تم تطوير نظام ذكي لتفسير هذه البيانات الجينية وإنتاج تنبؤات بصرية. تم تنفيذ هذا النظام بلغة Python ، باستخدام مكتبات متخصصة في علم البيانات والتعلم الآلي. هذا النظام، الذي يعتمد بشكل خاص على تصبغ العيون، هو تجربة لسمات الوجه الأخرى التي ستمنحنا، من خلال النمذجة ثلاثية الأبعاد لمورفولوجيا الوجه، توقعات لملامح غير معروفة في وجه معروف. النتائج التي تم الحصول عليها، على الرغم من أنها استكشافية ومحدودة ببعض السمات، تثبت الجدوى التقنية لهذه المقاربة وتشكل دليلاً قوياً على صحة المفهوم. تفتح هذه الدراسة آفاقاً واعدة لتطوير أدوات أكثر شمولاً للتعرف على الوجه لأغراض بيومترية أو طبية قانونية أو أنثروبولوجية.

الكلمات المفتاحية: تشكل الوجه، الجينات، الأليلات، التصبغ، OCA2، HERC2، TYR، التنبؤ، بايثون، النمذجة ثلاثية الأبعاد، التعرف البيومتري.



Liste des figures :

Figure 01 : Représentation des Allèles : Homozygotes (BB, bb) et Hétérozygote (Bb).....	04
Figure 02 : Analyse des variations SNP chez quatre individus : comparaison des allèles maternels et paternels.....	06
Figure 03 : Association génétique avec des traits faciaux régionalisés dans les populations normales.....	12
Figure 04 : Structure de chromosome Y.....	15
Figure 05 : : Structure du gène SRY.....	16
Figure 06 : Structure des gènes OCA2 et HERC2 sur le chromosome 15q12–q13.1. Deux SNP clés pour la prédiction de la couleur des yeux, rs1800407 et rs12913832, inclus dans le panel IrisPlex, sont également représentés.....	19
Figure 07 : : La génétique du nez : Comment nos gènes façonnent notre profil nasal.....	20
Figure 08 : Vue à l'intérieur de l'œil.....	22
Figure 09 : Analyse morphométrique des yeux selon la position, la forme et les dimensions oculaires en lien avec la morphologie faciale.....	23
Figure 10 : Exemples de stimuli montrant deux identités et les sept niveaux de taille des lèvres...	25
Figure 11 : Organisation génomique des gènes OCA2 et HERC2.....	27
Figure 12 : Effet phénotypique du SNP rs12913832.....	29
Figure 13 : Organisation génomique des principaux gènes impliqués dans la couleur des yeux autres que HERC2/OCA2.....	30
Figure 14 : représentation des 41 SNP de système HIrisplex-s pour la prédiction de couleur de peau cheveux et yeux d'un individu.....	38
Figure 15 : Importation des bibliothèques nécessaires.....	44
Figure 16 : Fonction pour appliquer une image d'arrière-plan.....	45
Figure 17 : Appliquer l'image de fond.....	45
Figure 18 : Données génétiques réelles.....	46
Figure 19 : Encodage des variables catégorielles.....	46
Figure 20 : Entraînement du modèle Random Forest.....	47
Figure 21 : Interface utilisateur Streamlit	48
Figure 22 : Interface utilisateur Streamlit	48
Figure 23 : l'importance relative de plusieurs SNPs dans la prédiction de la couleur des yeux, selon le modèle Random Forest.....	50

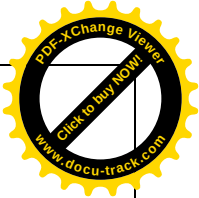
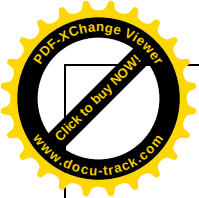
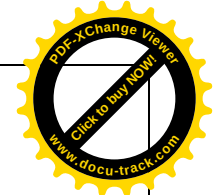
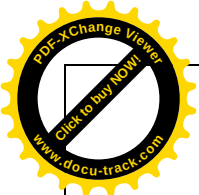


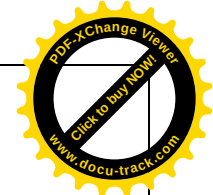
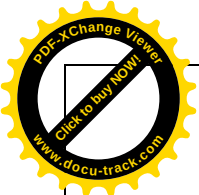
Figure 24 : Entrée des données génétiques pour la prédiction de couleur yeux marron.....	51
Figure 25 : Interface de prédiction de la couleur des yeux marron.....	51
Figure 26 : Entrée des données génétiques pour la prédiction de couleur yeux bleu.....	52
Figure 27 : Interface de prédiction de la couleur des yeux bleu.....	53
Figure 28 : Entrée des données génétiques pour la prédiction de couleur yeux vert.....	54
Figure 29 : Interface de prédiction de la couleur des yeux verte	55
Figure 30 : Entrée des données génétiques pour la prédiction de couleur yeux noisette.....	56
Figure 31 : Interface de prédiction de la couleur des yeux noisette.....	56



Liste des tableaux :

Tableau I : Liste des gènes et des SNP associés à une variation normale classés par position chromosomique.....	07
Tableau II : Caractéristiques embryonnaires qui contribuent au développement du visage.....	13
Tableau III : Les autres gènes du déterminisme du sexe.....	14
Tableau IV : représentation des SNP du système Hirisplex pour la prédication de couleur de peau..	17
Tableau V : Propriétés des SNP d'index dans les régions montrant une association significative à l'échelle du génome avec les traits faciaux quantitatifs.....	19
Tableau VI : Principaux gènes impliqués dans la forme des yeux humains.....	20
Tableau VII : les gènes utilisés dans la pigmentation des yeux.....	42

Liste des abréviations :



ADN : Acide Désoxyribonucléique

ARNr : Acide Ribonucléique Ribosomique

ARNt : Acide Ribonucléique de transfert

ARNm : Acide Ribonucléique messenger

SNP : Polymorphisme Mononucléotidique

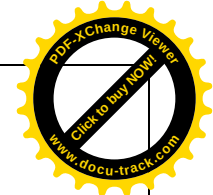
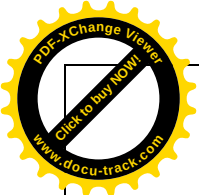
NCBI : National Center for Biotechnology Information

NHGRI : National Human Genome Research Institute

GWAS : Genome-Wide Association Study

3D : trois Dimensions

Table des Matières :



Remerciments

Dédicaces

Résumé

Liste des figures

Liste des tableaux

Liste d'abréviation

Introduction 01

Partie bibliographique :

1. généralité : Quelques définitions

1.1 L'ADN : L'acide désoxyribonucléique 03

1.2 Gène 03

1.3 allèle 03

1.4 génotype 04

1.5 Phénotype 04

1.6 Polymorphisme 05

1.7 SNP : polymorphisme nucléotidique simple 05

2. Généralité sur les caractères faciaux 07

3. Les caractères faciaux 13

3.1 le sexe 13

3.1.1 la détermination de sexe 13

3.1.1.1 Le chromosome Y et leur structure 14

3.1.1.2 La région pseudo autosomale : PAR 15

3.1.1.3 SRY : la région Y déterminant le sexe 15

3.2 La couleur de la peau 16

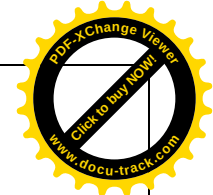
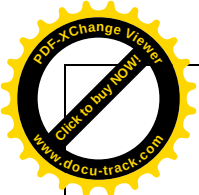
3.3 La couleur des yeux 17

3.3.1 La base physique de la couleur des yeux 17

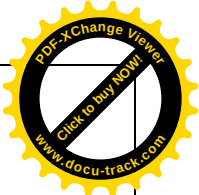
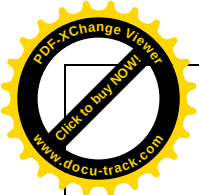
3.3.2 La base génétique du caractère de la couleur des yeux 17

3.3.2.1 Gènes impliqués dans la pigmentation des yeux 17

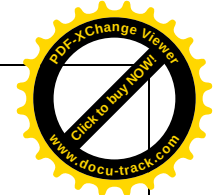
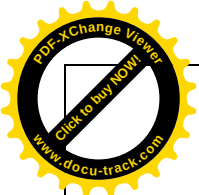
3.3.2.2 Rôle des gènes OCA2 et HERC2 dans la couleur des yeux 18



3.4	Forme du nez	19
3.5	Forme du yeux	20
3.5.1	la morphologie de l'œil	21
3.5.2	Impact de la forme du visage sur la morphologie des yeux	22
3.6	Forme de bouche	23
3.6.1	Définition morphologique de la forme de la bouche	23
3.6.2	Bases génétiques de la forme de la bouche	24
3.6.2.1	Influence génétique et les gènes impliqués	24
3.6.2.2	Influence des facteurs environnementaux	25
4.	La couleur yeux comme caractère phénotypique de l'identification faciale	26
4.1	Les bases génétiques de la couleur des yeux	26
4.1.1	Les bases génétiques de la couleur des yeux : les polymorphismes OCA2 et HERC2	26
4.1.1.1	Gène OCA2	27
4.1.1.2	Gène HERC2	28
4.1.1.3	Modèle pour la régulation du gène OCA2	28
4.1.2	Autres gènes liés à la couleur des yeux	29
4.1.2.1	Gène TYR	29
4.1.2.2	Gène TYRP1 et DCT	31
4.1.2.3	Gène SLC45A2	31
4.1.2.4	Gène LC24A4	32
4.1.2.5	Gène SLC24A5	32
4.1.2.6	Gène ASIP	33
4.1.2.7	Gène IRF4	34
5.LA PROGRAMMATION :	Codage et programmation	35
5.1	Analyse bioinformatique	36
5.1.1	GWAS: pour génome -Wide association study	36
5.1.2	Meta analyse	36
5.1.3	Le système HIrisPlex-S	36
5.2	Programmation python	38
	Matériel et méthode	39



1. Analyse bioinformatique	40
1.1 Sources de données	40
1.1.1 Données génétiques	40
1.2 Les gènes sélectionnés	41
2. Analyse informatique	42
2.1 Utilisation de Python pour le l'analyse génétique et facial	42
2.2 Méthodologie de programmation	42
2.2.1 Importation des bibliothèques	42
2.2.2 Définir un fond d'écran	43
2.2.3 Données génétique	44
2.2.4 Encodage des données	45
2.2.5 Préparation pour le modèle	46
2.2.6 Création du modèle Random Forest	46
2.2.7 Création de l'interface avec Streamlit	46
Les résultats	48
1. Prédiction de la couleur des yeux marron	49
2. Prédiction de la couleur des yeux bleus	51
3. Prédiction de la couleur des yeux vert	52
4. Prédiction de la couleur des yeux noisette	54
Discussion	57
Conclusion	59

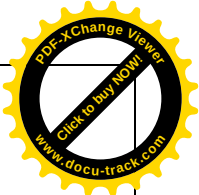
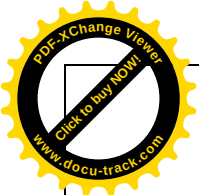


Introduction :

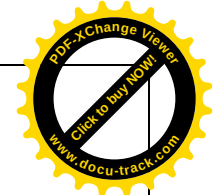
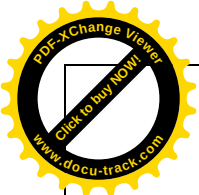
L'identification faciale représente un enjeu de premier ordre dans des disciplines telles que la biométrie, la médecine légale et la sécurité. Traditionnellement fondée sur des techniques d'analyse d'images ou de reconnaissance faciale, cette approche connaît une évolution significative grâce aux avancées en génétique moléculaire et en intelligence artificielle. L'émergence du phénotypage basé sur l'ADN, qui permet d'anticiper certains traits physiques à partir de l'information génétique marque une avancée majeure vers de nouvelles approches d'identification, notamment la prédiction automatisée de caractéristiques faciales à partir de gènes ciblés.

Certains traits du visage, tels que la couleur des yeux, la forme de la bouche ou la structure nasale, sont influencés par des polymorphismes génétiques bien identifiés. Des gènes comme OCA2, HERC2, TYR, SLC24A4 ou encore FREM1 jouent un rôle dans la détermination de caractéristiques pigmentaires et morphologiques. En intégrant ces données génétiques dans des modèles de machine learning, il devient possible d'établir des prédictions phénotypiques faciales à partir d'un simple échantillon d'ADN.

Ce mémoire s'inscrit dans une démarche exploratoire visant à évaluer le potentiel de cette approche. Il se concentre sur une sélection de gènes clés, impliqués dans les processus de pigmentation et de morphologie faciale. Dans cette perspective, nous développons un système intelligent capable d'estimer certains traits faciaux à partir de profils génétiques. Cette approche illustre la convergence croissante entre les avancées en génomique et les outils de l'apprentissage automatique appliqués à l'identification humaine.



Synthèse bibliographique :



1.GÉNÉRALITE : Quelques définitions

1.1- L'ADN : L'Acide DésoxyriboNucléique (ADN) :

L'acide désoxyribonucléique est une molécule hélicoïdale bicaténaire constituée de monomères appelés nucléotides. Chaque nucléotide est composé d'un sucre (désoxyribose), d'un groupement phosphate et de l'une des quatre bases azotées : adénine, thymine, cytosine ou guanine. L'ADN contient les instructions génétiques essentielles au développement, au fonctionnement, à la croissance et à la reproduction de tous les organismes vivants connus ainsi que de nombreux virus. Sa structure unique en double hélice permet une réplication fidèle et l'emmagasinement de l'information héréditaire (Brooker ,2020).

1.2 –Le Gène :

Dans le cadre de la biologie moléculaire contemporaine, la définition du gène dépasse largement l'idée d'une simple séquence codante pour une protéine. En effet, un gène peut être défini comme une unité fonctionnelle de l'hérédité, localisée sur l'ADN et composée d'une séquence spécifique de nucléotides. Cette séquence contient l'information nécessaire à la synthèse d'un produit fonctionnel, qui peut être soit une protéine, soit un ARN non codant aux rôles essentiels, tels que les ARN ribosomiques (ARNr) ou les ARN de transfert (ARNt). De plus, un gène comprend souvent des régions régulatrices comme les promoteurs ou les enhancers qui modulent son expression. Sur le plan structurel, la définition moderne inclut les exons, parties codantes du gène, ainsi que les introns, segments non codants mais transcrits. Enfin, certains gènes peuvent se chevaucher ou donner lieu à plusieurs produits différents selon les mécanismes d'épissage alternatif, ce qui illustre la complexité de l'expression génique. (Alberts et al ,2022)

1.3 – L'Allèle (Fig. 01) :

Selon Hartwell, un allèle désigne une variante spécifique d'un gène, localisée à un locus spécifique sur un chromosome. Chaque individu diploïde possède deux allèles pour un même gène, hérités respectivement de chaque parent. Ces allèles peuvent être identiques (homozygotes) ou différents (hétérozygotes). Les différences entre allèles résultent de variations dans la séquence d'ADN, pouvant entraîner des différences dans la fonction ou l'expression du gène, ce qui se traduit par des variations phénotypiques observables. Certains allèles sont dominants et s'expriment pleinement dans le phénotype, tandis que d'autres sont récessifs et ne se manifestent que lorsqu'ils sont présents en deux exemplaires. L'étude des allèles revêt une importance capitale en génétique. Elle permet de saisir les mécanismes de la transmission des traits héréditaires, de comprendre la diversité génétique au sein des populations et d'identifier la base moléculaire des maladies génétiques. (Hartwell et al ,2018)

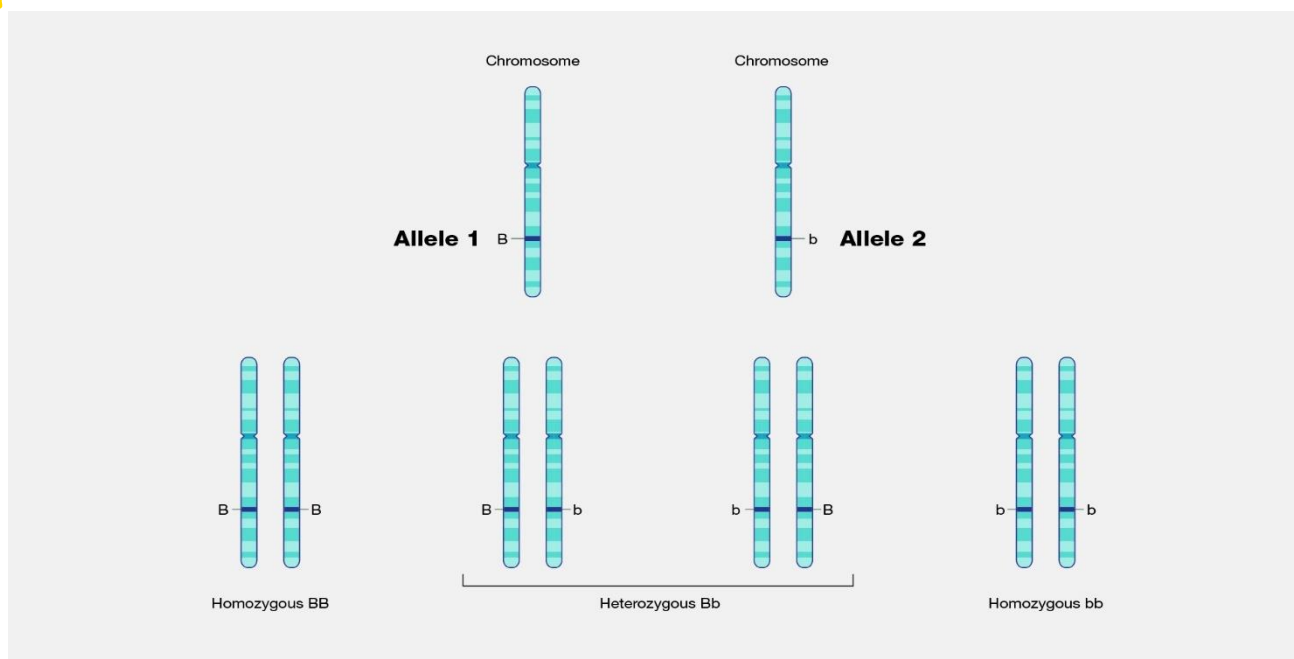


Fig. 01 : Représentation des Allèles : Homozygotes (BB, bb) et Hétérozygote (Bb) (NIH, 2025).

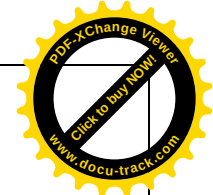
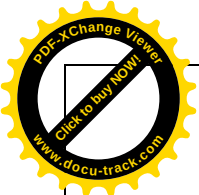
1.4 –Le Génotype :

Le génotype correspond à l'ensemble des allèles portés par un individu, qu'ils s'expriment ou non dans son phénotype. Il constitue la base héréditaire de tous les caractères d'un organisme. (Alberts et al ,2002)

Selon David Adams de l'NHGRI (National Human Genome Research Institute) le génotype est l'ensemble des séquences d'ADN (allèles) d'un individu à des loci spécifiques, responsables de traits héréditaires. Il peut influencer directement ou indirectement l'expression phénotypique, mais son effet dépend aussi d'interactions génétiques et environnementales. (David et al ,2025)

1.5 –Le phénotype :

Selon Griffiths le phénotype est l'expression observable du génome d'un individu, influencé à la fois par l'environnement et le génotype, et il est essentiel dans la recherche génétique. (Griffiths et al ,2007). Le phénotype regroupe l'ensemble des caractéristiques observables d'un organisme qu'ils soient morphologiques, physiologiques, biochimiques ou comportementaux. Il résulte de l'interaction complexe entre le génotype de l'individu et les influences de son environnement. (Norton ,2018)



1.6 –Le Polymorphisme :

Dans l'ouvrage Human Molecular Genetics de Tom Strachan et Andrew Read (5e édition, 2018), le polymorphisme désigne l'existence de plusieurs variantes (ou allèles) d'un même gène ou d'une séquence d'acide désoxyribonucléique au sein d'une population, présentes à une fréquence significative (généralement supérieure à 1 %). Ces variations peuvent être considérées comme neutres, bénéfiques ou délétères, et influencer, dans certains cas, des traits observables ou la prédisposition à certaines pathologies, le polymorphisme est un concept scientifique majeur qui permet de mieux comprendre la diversité génétique, l'évolution et les études d'association génétique. **(Strachan et al ,2018)**

1.7-Le SNP : polymorphisme nucléotidique simple (fig.02) :

Les polymorphismes nucléotidiques simples (SNP) sont des variations de la séquence d'ADN qui se produisent lorsqu'un seul nucléotide du génome diffère dans les chromosomes appariés. Certains SNP dans la région codante modifient la séquence d'acides aminés d'une protéine, tandis que d'autres dans la région codante n'ont aucune incidence sur la séquence protéique. Les SNP situés en dehors de la région codante peuvent également affecter la liaison des facteurs de transcription, l'épissage des gènes ou la dégradation de l'ARNm. Qu'ils aient ou non un impact sur la fonction biologique des produits géniques, les SNP sont très utiles comme marqueurs pour examiner le déséquilibre de liaison et explorer les polymorphismes génétiques dans la recherche en génétique des populations et en sciences médicales .Les connaissances sur les SNP devraient aider à identifier les gènes de susceptibilité aux maladies complexes, les gènes de susceptibilité aux médicaments et les gènes d'histocompatibilité pour les transfusions ou les transplantations.**(Wakui et al ,2013)**

Individual 1
Maternal ...CGATATTCC**T**ATCGAATGTC...
Paternal ...CGATATTCC**C**ATCGAATGTC...

Individual 2
Maternal ...CGATATTCC**C**ATCGAATGTC...
Paternal ...CGATATTCC**C**ATCGAATGTC...

Individual 3
Maternal ...CGATATTCC**T**ATCGAATGTC...
Paternal ...CGATATTCC**T**ATCGAATGTC...

Individual 4
Maternal ...CGATATTCC**C**ATCGAATGTC...
Paternal ...CGATATTCC**T**ATCGAATGTC...

Fig. 02 : Analyse des variations SNP chez quatre individus : comparaison des allèles maternels et paternels. (National Human Genome Research Institute. 2025)

2- Généralité sur les caractères faciaux :

La morphologie faciale humaine est l'une des caractéristiques biologiques visibles les plus importantes. Comprendre les bases génétiques qui sous-tendent les traits de forme du visage a des implications importantes en génétique des populations, en biologie du développement et en sciences médico-légales. (Liu et al ,2020)

Le visage humain englobe un éventail captivant de caractéristiques physiques qui présentent des variations notables selon les régions géographiques.il comprend des organes sensoriels complexes et présente des variations remarquables de traits faciaux. Ces traits complexes et délicats subissent une régulation génétique précise au cours du développement facial, ce qui donne naissance à nos visages hautement héréditaires, uniques et reconnaissables, qui jouent un rôle essentiel dans la communication et l'identification physique (tableau I). (Du et al ,2024)

Tableau I : Liste des gènes et des SNP associés à une variation normale classés par position chromosomique (GWAS) (Richmond et al ,2018)

Trait du visage	Phénotype facial	Position chr	SNP	Les gènes
Ouverture de l'aile	Zone subtile et proéminente autour des ailes et entrée plus large dans les narines	4q31.3 (1548283 66)	rs99958 21	DCHS2
Allométrie	Allométrie (taille et forme de mise à l'échelle), distance canthale interne, forme et largeur globales du visage	15 (8502994 5), (8506109 5)	rs12908 400, rs12909 111	
Taille du centroïde	Taille du centroïde, hauteur et largeur du visage, mais indépendantes de la taille du centroïde.	3q25.33	rs79909 949	SCHIP
Proéminence du menton	Proéminence centralisée du menton s'étendant jusqu'à la lèvre inférieure associée à une récession mandibulaire alignée avec les commissures.	1q31.3 (1973290 41)	rs28211 16	ASPM

Proéminence du menton	Proéminence du bouton du menton associée à une récession latérale.	7q21.3 (96124975)	rs10238953	DLX6, DYNC1L1
Proéminence et forme du menton	Forme et protubérance du menton	2q12	rs3827760	EDAR
Forme des yeux	Courbure de la paupière	2:176 820 065	rs970797	HOXD1-MTX2
Forme des yeux	Longueur de la queue de l'œil	6:169 699 889	rs3736712	WDR27
Largeur des yeux	Distance entre les globes oculaires et le nasion	2q35	rs974448	PAX3
Largeur des yeux	Distance entre les globes oculaires	2q35 (222713558)	rs16863422 rs12694574 rs974448	PAX3
Largeur des yeux	Distance entre les globes oculaires	3q28 (191032117)	rs17447439	TP63
Largeur des yeux	Distance inter-canthi	Xq13.2 (72289467)	rs11093404	PABP1-C1L2A, HADC8
Largeur et profondeur des yeux	Hauteur, largeur et profondeur des yeux		rs7560738	TMEM163
Largeur et profondeur des yeux	Distance entre les globes oculaires et le nasion	10q24.3 (105800390)	rs805722	COL17A1
Hauteur du visage	Hauteur du milieu du visage	6q26	rs9456748	PARC 2
Hauteur/profondeur du visage	Gonion droit vers exR et enR	2:19595772	rs7567283	OSR1-WDR35
Largeur du visage	Distance inter tragi	14q21.1 (38038468)	rs17106852	PAX9, SLC25A2,

				MIPOL1, FOXA1
Largeur du visage	Distance inter tragi	20q12 (38904203)	rs6129564	MAFB
Profondeur du visage	Tragus gauche au nasion	11q22.1 (101394765)	rs12786942	TRPC6
Front	Proéminence centralisée du front avec une verticale Dépression au-dessus des orbites	1p12 (119762175)	rs72691108	TBX15
Front	Partie centrale récessive du front avec proéminence latérale	6q23.2 (133615646)	rs5880172	RPS12, EYA4
Lèvre (supérieure)	Hauteur centrale de la lèvre supérieure	9p22	rs72713618	FREM1
Proéminence des lèvres	Lèvres proéminentes, rétrusives latéralement aux lèvres supérieures et inférieures ; léger rétrécissement du pli nasogénien aux parois nasales latérales avec une petite proéminence supérieurement.	2q31.1 (177111819)	rs970797	Cluster HOXD
Pli mental	Proéminence du pli mental, avec des effets rétrusifs subtils sur le pli labio-mandibulaire.	2p21 (42181679)	rs6740960	PKDCC
Prédominance de la nation	Proéminence et position verticale du nasion	2q35	rs7559271	PAX3
Nasion, yeux et proéminence du zygoma	Position du nasion par rapport au zygoma et aux globes oculaires	5q35.1 (171061069)	rs6555969	C5orf50
Pont nasal	Petite dépression au-dessus de la pointe du nez, associée à de petites dépressions au milieu des parois nasales latérales. Proéminence accrue de l'arête du nez	1p32.1 (60997570)	rs4916068	Intergénique
Pont nasal	Associé à de petites crêtes transversales sur le pont du nez, proéminentes à mi-chemin et moins proéminentes sur les ailes du nez	3q27.1 (184333169)	rs58022575	EPHB3, DVL3
Nez proéminent	De l'aile gauche au bout du nez	14q11.2 (21365801)	rs21365801	ZNF219, CHD8
Nez proéminent	Angle labial nasal et proéminence du nez	17:67,537,404	rs2193054	SOX9
Nez proéminent	Pointe du nez récessive avec une largeur accrue le long de la paroi latérale du nez.	17q24.3 (69139583)	rs11655006	BC039327 /C ASC17

Nez proéminent	Proéminent au-dessus de la pointe du nez associé à une zone localisée de récession juste au-dessus des ailes	19q13.11 (3429099 5)	rs28710 4	KCTD15
Nez proéminent	Inclinaison de la columelle, protrusion du nez, angle de la pointe du nez	4q31 (1552353 92)	rs12644 248	DCHS2
Proéminence et largeur du nez	Largeur du nez, hauteur du nez et ailes gauche et droite au prn	1p36.32	rs46483 79	PRDM16
Largeur du nez	Extrémités latérales proéminentes du nez associé à des ailes alaires légèrement réduites	17q24.3 (7003647 9)	rs58218 92	SOX9
Largeur du nez	Largeur et hauteur du nez	1p36.23-p33 (3251376)	rs46483 79	PRDM16
Largeur du nez	Largeur des ailes	20:37,426 ,155	rs22064 37	DHX35
Largeur du nez	Largeur de l'aile du nez	20p11 (2204157 7)	rs92783 3	PAX1
Largeur du nez	Largeur inter alaire	20p11.22 (2163254 5)	rs24243 99	PAX1
Largeur du nez	Proéminence au niveau du nasion et largeur de la pointe/des ailes du nez avec largeur réduite le long des parois latérales du nez.	2q36.1 (2230390 52)	rs10176 525	PAX3
Largeur du nez	Largeur du pont du nez	6p21 (4532965 6)	rs18529 85	SUPT3H/RUN X2
Largeur des yeux	Largeur intercanthale	1p13.3 (1102187 61)	rs61968 6	GSTM2, GNAI3, ALX3
Largeur du nez	Dépression nasale associée à un élargissement des parois nasales latérales en position supérieure. Légère dépression sous-nasale également.	6p21.1 (4468184 0)	rs22783 3	SUPT3H
Largeur du nez	Largeur de l'aile du nez	7p13 (4213139 0)	rs17640 804	GL13
Philtrum	Dépression du philtrum avec piliers philtraux proéminents ; ale réduit et sous-nasal proéminent.	3q21.3 (1281062 67)	rs29775 62	RAB7A, ACAD9
Hauteur du profil facial supérieur	Schéma squelettique /profil mandibulaire		rs11642 644	MBTPS1

Proéminence du profil facial supérieur	Schéma squelettique/profil mandibulaire		rs20451 45	PCDH15
---	---	--	---------------	--------

Le complexe craniofacial est initialement modulé par l'expression précise des gènes embryonnaires et les interactions moléculaires médiées par des voies complexes. À mesure que l'homme grandit, les hormones et les facteurs biomécaniques affectent également de nombreuses parties du visage. L'incapacité à résumer systématiquement les variations faciales a entravé la découverte des déterminants et des corrélats de la forme du visage. Contrairement aux technologies génomiques, le phénotypage systématique et complet a pris du retard. Cela est particulièrement vrai dans le contexte de traits multipartites tels que le visage humain. Dans les études d'association pangénomique (GWAS) typiques, les phénotypes sont aujourd'hui résumés comme des variables univariées, ce qui est intrinsèquement limitatif pour les traits multivariés, qui, par définition, ne peuvent être exprimés par des variables uniques. **(Claes.P et al ,2014)**

La surface du visage est facilement visible et identifiable avec une relation étroite avec les structures cartilagineuses et squelettiques sous-jacentes, Les différences de taille relative **(Fig.03)**, de forme et de disposition spatiale (verticale, horizontale et en profondeur) entre les différentes caractéristiques du visage (par exemple, les yeux, le nez, les lèvres, etc.) rendent chaque visage humain unique, bien que des individus étroitement apparentés tels que les jumeaux monozygotes aient des structures faciales très similaires. **(Richmond et al ,2018).**

Gene association with regionalised facial features in normal populations

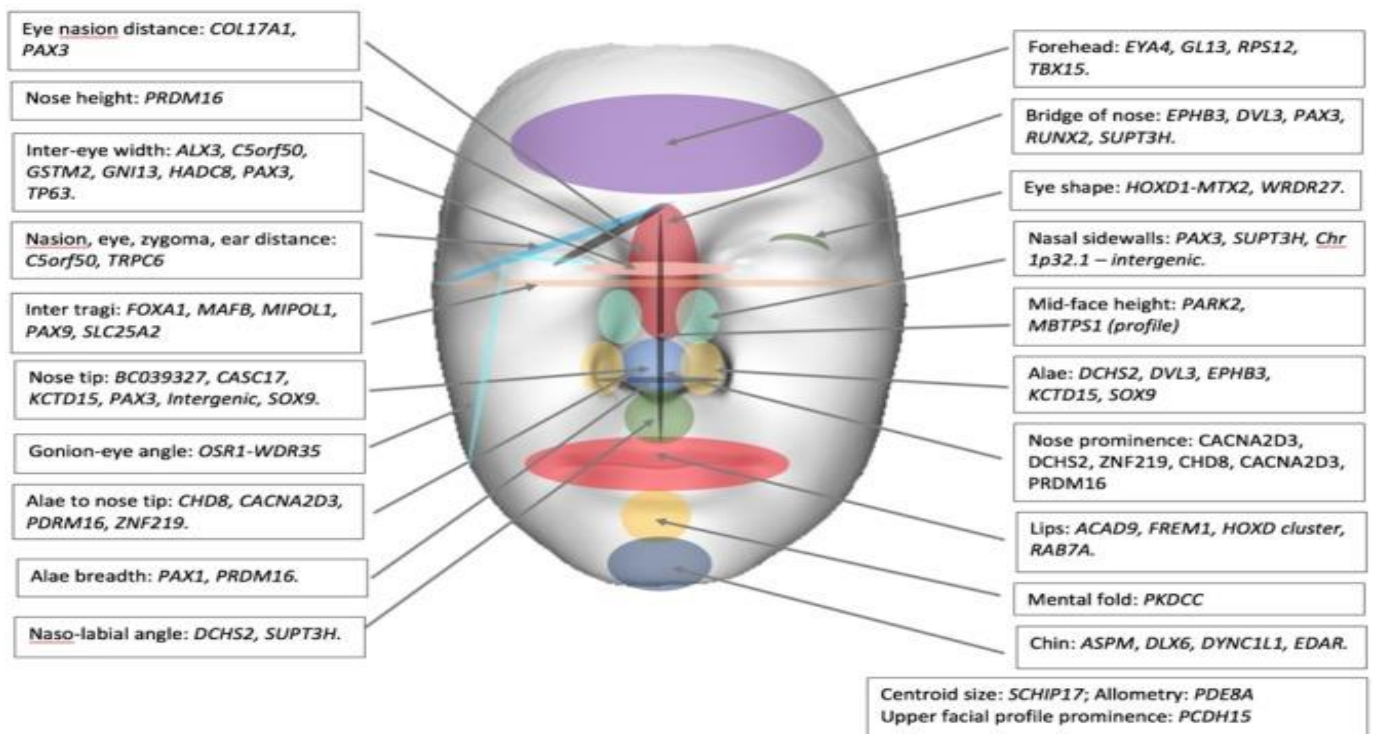


Fig.03 : Association génétique avec des traits faciaux régionalisés dans les populations normales. (Richmond et al ,2018)

Bien que les bases génétiques de la morphologie faciale humaine soient encore limitées, de nombreuses variantes génétiques associées aux traits faciaux ont été signalées ces dernières années. (Huang et al ,2021), il est intéressant de noter que la morphologie faciale est influencée par des facteurs génétiques et environnementaux et leurs interactions complexes qui représentent le plus grand défi de la biologie moderne (Djordjevic et al ,2016).

Le développement du visage implique une série complexe et coordonnée d'événements embryonnaires. Les caractéristiques reconnaissables du visage humain se développent vers la 4e semaine de gestation et sont étroitement liées aux cellules de la crête neurale crânienne. Les processus constitutifs du développement facial sont répertoriés (dans le Tableau II), et la séquence embryonnaire humaine des événements peut être visualisée, ce qui facilite la compréhension du mouvement des processus faciaux suivi de leur fusion (Richmond et al ,2018).

Tableau II : Caractéristiques embryonnaires qui contribuent au développement du visage. (Richmond et al ,2018).

Processus faciaux de développement	Traits du visage
Frontonasal	Front, paupières supérieures, conjonctive
Nasale médiale	Nez, lèvre supérieure/philtrum, prémaxillaire, incisives supérieures
Nasale latérale	Ailes et base du nez
Maxillaire	Paupières inférieures, joues, parties latérales de la lèvre supérieure, maxillaire, canines, prémolaires et molaires
Mandibulaire	Lèvre inférieure entière, mâchoire inférieure (mandibule, y compris les dents)

3. Les caractères faciaux : les caractères faciaux sont regroupés de la façon suivante :

3.1 Le sexe :

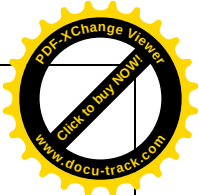
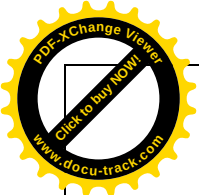
Il existe 3 niveaux différents pour définir le sexe :

- Le niveau génétique associé à la présence, chez l'homme, du chromosome Y (caryotype 46, XY) et chez la femme son absence (46, XX). On parle alors de sexe génétique.
- Le niveau gonadique associé à la présence chez l'homme de testicules et d'ovaires chez la femme. On parle de sexe gonadique.
- Le niveau phénotypique selon l'aspect des organes génitaux externes. On parle de sexe phénotypique.

Ces 3 niveaux sont les étapes d'une cascade d'opérations que le sexe génétique va influencer l'apparition d'un type de gonades, et ces mêmes gonades vont ensuite sécréter des hormones qui vont permettre une évolution des organes génitaux externes et internes vers le phénotype féminin ou masculin. **(Vialard et Fellous ,2002)**

3.1.1 la détermination de sexe : (tableau III)

D'après Vilain, la détermination du sexe chez les mammifères correspond à la détermination, chez le mâle, du testicule. Plus d'une dizaine de gènes impliqués dans le développement précoce de la gonade



ont été décrits chez l'homme, C'est donc la présence ou l'absence de chromosome Y qui détermine le sexe (respectivement mâle ou femelle). (Vilain et al ,1992)

Tableau III : Les autres gènes du déterminisme du sexe (Stéphanie B ,2019)

Gène	Localisation chromosomique	Tissu ou le gène s'exprime	Le rôle de sa protéine	Caractéristiques de sa protéine
WT1	Sur le bras long du chromosome 11	Cellules de sertoli Cellules de la granulosa	Essentiel à la morphogénèse de la crête génitale Régulation de la transcription du gène SRY	Les isoformes de protéines Ont la capacité de se lier à l'ADN
SOX9	Sur le bras long du chromosome 17	Crêtes génitales des deux sexes Cellule de sertoli	Activation de l'expression du gène de l'AMH	Contient 509 aa Un domaine de liaison à l'ADN
DAX1	Sur le bras long du chromosome 9	Crêtes génitales	Inhibition de l'action de SF1	Contient 470 aa Région de liaison à l'ADN Région de liaison pour les ligands des récepteurs nucléaires

3.1.1.1 Le chromosome Y et sa structure (Fig. 04) :

D'après kuroki et fukami, le chromosome Y est un génome haploïde propre aux hommes, Il se transmet facilement à la génération suivante sans être réparé par recombinaison, même en cas d'altération structurelle génomique majeure. En revanche, le génome du chromosome Y est essentiellement une région transmise uniquement de père en fils, reflétant une hérédité masculine intergénérationnelle, avec une taille de génome d'environ 60 Mb. La moitié est une région euchromatinienne contenant plusieurs gènes, tandis que l'autre moitié est une région hétérochromatine constituée uniquement de séquences répétées, La région euchromatinienne est divisée en régions pseudo-autosomiques (PAR), situées aux deux extrémités du chromosome Y et dont la séquence génomique est supposée presque identique à celle du chromosome X, et en région mâle-spécifique du chromosome Y (MSY) (Kuroki, al ,2023)

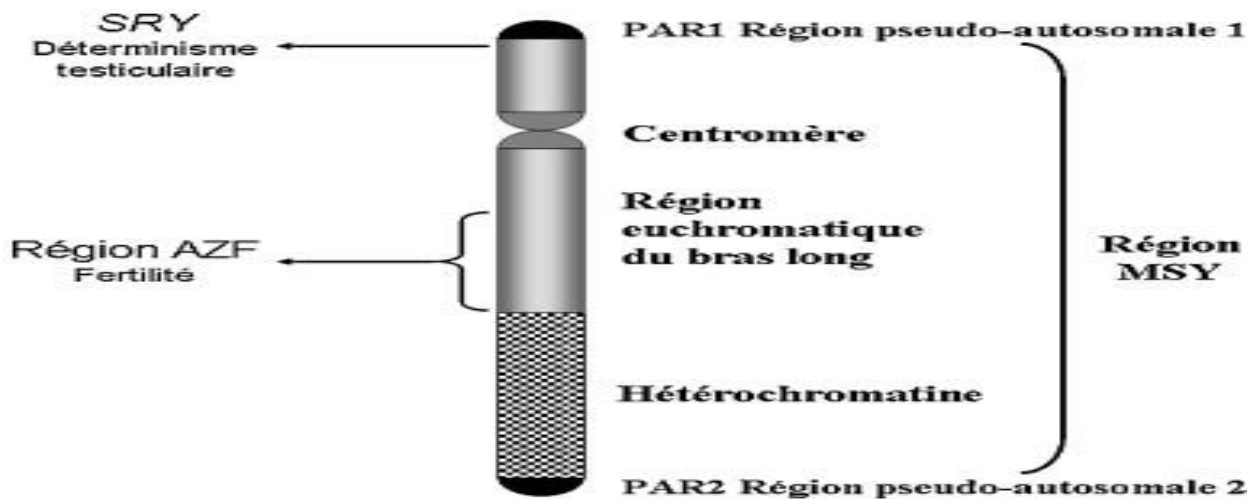


Fig.04 : Structure de chromosome Y (Ravel et Siffroi ,2009)

3.1.1.2 La région pseudo autosomale : PAR

PAR1 est situé à l'extrémité du bras court et mesure 2,6 mégabases, PAR2 est situé à l'extrémité du bras long et mesure 320 kilobases. Bien que les chromosomes X et Y soient très divergents, les PAR son homologue. Les PAR du chromosome Y subissent un appariement homologue et une recombinaison avec les PAR du chromosome X pendant la méiose. Cet appariement est considéré comme important pour faciliter l'orientation et la ségrégation correctes des chromosomes sexuels pendant la méiose (**Punjani et al ,2020**)

3.1.1.3 SRY : la région Y déterminant le sexe :

Selon les informations adoptées de la plateforme NCBI, aussi connu sous le nom de TDF ; TDY ; SRXX1 ; SRXY1, ce gène est sans intron (mono-exonique) et il est exprimé uniquement dans le testicule embryonnaire, Il code pour une protéine composée de 204 acides aminés, Cette protéine le facteur déterminant du testicule (TDF), répartis en 3 régions dont la centrale est une boîte HMG (High Mobility Group) ayant une haute affinité avec l'ADN (**Vialard et Fellous ,2002**), Ce domaine HMG occupe une position centrale sur le gène SRY (**Fig. 05**). Il a un rôle crucial dans l'activité de SRY (**Leroux ,1999**).

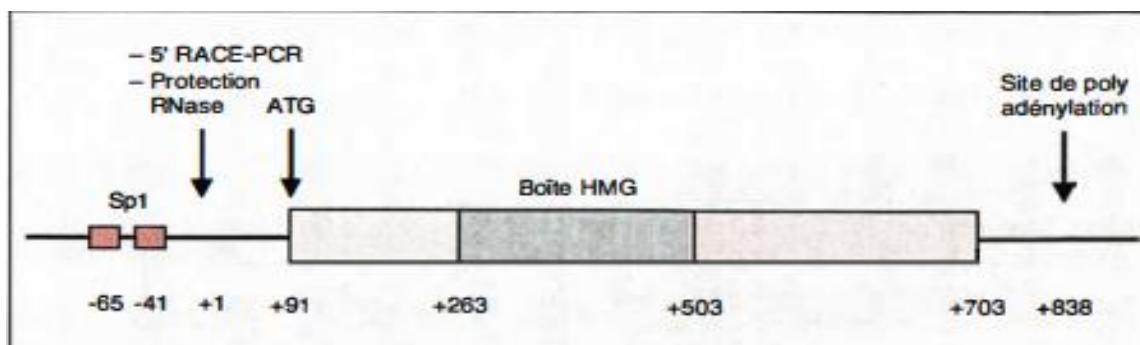


Fig.05 : Structure du gène SRY (Vilain et al ,1992)

3.2 La couleur de la peau :

La couleur de la peau humaine va de la peau pâle ou blanche jusqu'à la peau colorée (brun léger, brun marron et noir). Elle résulte d'un mélange des chromophores rouges (oxyhémoglobine), bleus (hémoglobine désoxygénée), jaune-orangé (carotène, un pigment exogène) et bruns (mélanine). La mélanine est un constituant majeur de la couleur de la peau. La présence ou l'absence de mélanine dans les mélanosomes des mélanocytes et dans les kératinocytes sont responsables de la pigmentation épidermique, tandis que la mélanine dans les macrophages et les mélanocytes du derme est responsable de la pigmentation dans le derme. (Ortonne ,2012).

La synthèse de mélanine régulé par le récepteur de la mélanocortine 1 (MC1R) est une protéine réceptrice transmembranaire couplée à la protéine G présente dans les mélanocytes de la couche basale de l'épithélium. Le MC1R interagit avec son ligand endogène, l'alpha-hormone stimulant les mélanocytes (αMSH), pour activer la voie de signalisation de l'AMPc et l'expression de MITF (Su et al ,2023)

Il existe deux types de mélanine : la phéomélanine rouge et l'eumélanine noire. Les mutations génétiques entraînant une perte de fonction sont associées à une augmentation de la production de phéomélanine, ce qui entraîne un éclaircissement de la peau et des cheveux. (NCBI).

Le gène SLC45A2 (tableau IV), code une protéine de transport qui intervient dans la synthèse de la mélanine (MATP). Cette protéine est exprimée dans un pourcentage élevé de lignées cellulaires de mélanome, qui assure la synthèse de mélanine par le biais de la tyrosinase et/ou du transport de protons vers les mélanosomes. Le gène MATP est situé en 5p13.2 et, lorsqu'il est muté, il est responsable de l'albinisme oculocutané de type 4 (OCA4) (Fracasso et al ,2013)

Tableau IV : représentation des SNP du système HIRISplex pour la prédiction de couleur de peau (<https://hirisplex.erasmusmc.nl/> site web 1)

Gene	SNP	Allèle	Phénotype
SLC45A2	Rs28777	C	Peau plus claire
SLC45A2	Rs16891982	C	Peau clair

3.3 La couleur des yeux :

La pigmentation des yeux constitue un trait phénotypique visible, dont la variation est largement influencée par plusieurs facteurs

3.3.1 La base physique de la couleur des yeux :

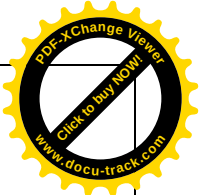
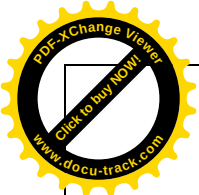
La couleur des yeux est déterminée par un polymorphisme et un trait polygénique. Le brun est la couleur des yeux la plus répandue dans le monde, représentant environ 79 %, le bleu environ (8 à 10 %), le noisette (5 %) et le vert (2 %). Les couleurs rares sont le gris et le rouge/violet. Différents facteurs interviennent dans la détermination de la couleur des yeux. Les deux facteurs les plus importants sont la pigmentation de l'iris et la façon dont la lumière est diffusée par l'iris. L'expression génétique détermine la pigmentation de l'iris et la quantité de mélanine présente dans l'œil, c'est-à-dire le nombre de sous-unités de mélanine qui identifie la couleur des yeux. (Dorgaleleh et al ,2020)

3.3.2 La base génétique du caractère de la couleur des yeux :

La couleur des yeux est un caractère héréditaire complexe, déterminé par plusieurs gènes qui contrôlent la quantité, la répartition et le type de mélanine (pigment) dans l'iris. Depuis longtemps, la communauté scientifique considérait que la couleur des yeux était déterminée par un seul gène portant deux allèles (bleu et brun). Cependant, des études récentes ont révélé que ce processus est en réalité le résultat de l'interaction de plus de 16 gènes.

3.3.2.1 Gènes impliqués dans la pigmentation des yeux :

La couleur de l'iris est fortement héréditaire, De précédentes études d'association à l'échelle du génome (GWAS) ont identifié divers polymorphismes mononucléotidiques (SNP) dans et autour de 10 gènes significativement associés à la couleur des yeux, soulignant la nature polygénique d'un caractère qui dans le passé, était supposé être génétiquement simple. L'influence génétique la plus forte sur la couleur des yeux est exercée par les gènes HERC2 (chromosome 15q13.1) et OCA2 (chromosome



15q11.2–q12) voisins, où il a été démontré qu'un SNP intronique dans HERC2 avait un effet amplificateur à longue distance, interagissant avec le promoteur d'OCA2 et fonctionnant comme un commutateur moléculaire entre la pigmentation claire et la pigmentation foncée. Les connaissances génétiques disponibles précédemment permettent de prédire avec précision la couleur des yeux bleus et bruns, par exemple avec un système de test ADN basé sur six SNP de six gènes, dont HERC2 et OCA2, qui a été utilisé dans des applications anthropologiques et médico-légales. Cependant, la couleur des yeux non bleus et non bruns peut être prédite génétiquement avec une précision considérablement plus faible, probablement parce que les SNP prédictifs et les gènes responsables sont inconnus.

En outre, la variance phénotypique de la couleur des yeux non expliquée précédemment par les études d'association pangénomique variait entre 26 % pour une échelle bleue par rapport à une échelle brune et 50 % pour une échelle utilisant trois catégories d'une étude à l'autre. Cela illustre la présence probable de gènes encore inconnus responsables de l'absence de précision génétique de prédiction et de l'absence d'héritabilité de la couleur des yeux humains (Simcoe et al ,2021)

3.3.2.2 Rôle des gènes OCA2 et HERC2 dans la couleur des yeux :

Les gènes OCA2 et HERC2 (**Fig. 06**) jouent un rôle central dans l'hérédité de la couleur des yeux, en particulier pour les yeux bleus et bruns. Des polymorphismes génétiques dans la région régulatrice et codante de OCA2 sont associés à la variation de la pigmentation des yeux, des cheveux et de la peau. Ces découvertes ont permis de mieux comprendre la base génétique de la pigmentation humaine et ont attiré l'attention sur leurs applications potentielles, telles que les enquêtes médico-légales, les recherches historiques et anthropologiques.

1. Le SNP rs12913832 du gène HERC2, localisé dans l'intron 86, joue un rôle déterminant dans l'étude de la prédisposition à la couleur des yeux, notamment pour les yeux de couleur bleue. Cette prédisposition est régulée par le gène OCA2, ce qui permet d'expliquer les variations de pigmentation oculaire. Il est inclus dans le panel IrisPlex, un système d'analyse qui permet de prédire la couleur des yeux avec une précision d'environ 94 % pour les yeux bleus ou marron en Europe. Cependant, la précision est réduite (73 %) pour les couleurs intermédiaires ou dans les populations génétiquement mixtes. Ce syndrome d'épissage spécifique (SNP) exerce une influence sur l'expression du gène OCA2, qui code pour la protéine P. Cette dernière est impliquée dans le transport de protéines vers les mélanosomes, les organites responsables de la pigmentation. (Suarez et al ,2021)

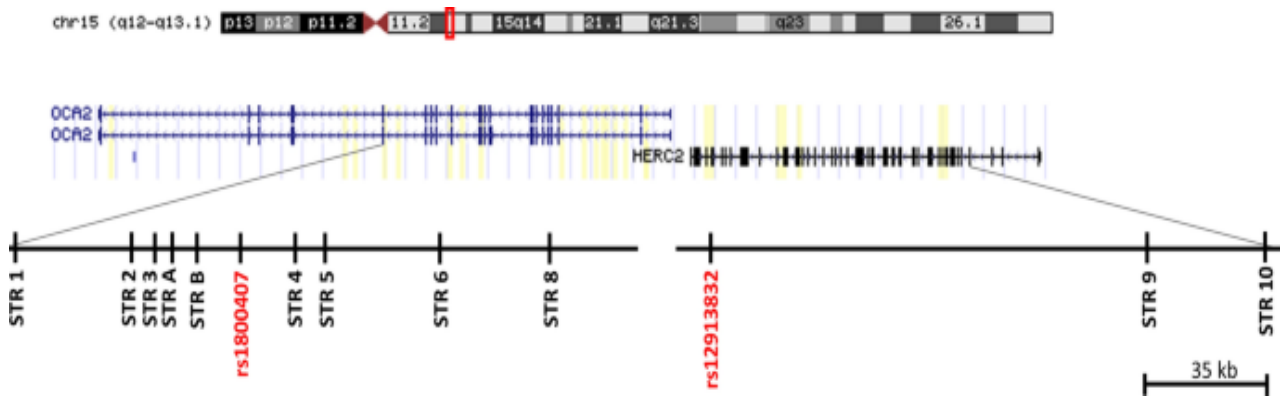


Fig.06 : Structure des gènes OCA2 et HERC2 sur le chromosome 15q12–q13.1. Deux SNP clés pour la prédiction de la couleur des yeux, rs1800407 et rs12913832, inclus dans le panel IrisPlex, sont également représentés. (Suarez et al ,2021)

3.4 Forme du nez :

Le nez constitue la partie la plus proéminente du profil et occupe la position la plus visible du visage. Il a une grande influence sur l'apparence et le profil du visage (**Fig. 07**). L'équilibre et l'harmonie du nez sont donc très importants. Les nez existent en différentes tailles et formes : retroussés ou droits, avec ou sans bosse. De même, en profil, le dos du nez peut être classé comme : droit, convexe ou concave (**Jankowska et al ,2021**). Le génotype (**Tableau V**), ainsi que les facteurs épigénétiques et environnementaux déterminent le phénotype de l'organisme, mais le génotype est établi avant la caractérisation du phénotype. Les caractéristiques nasales comme la largeur, la hauteur et la proéminence du nez ont une forte composante génétique (**Gupta et al ,2020**).

Tableau V : Propriétés des SNP d'index dans les régions montrant une association significative à l'échelle du génome avec les traits faciaux quantitatifs. (Adhikari et al ,2016).

Gene	Région chromosomique	Allèles	SNP	Trait associé
EDAR	2q12	A>G	rs3827760	Protrusion du menton
PAX3	2q35	A>G	rs7559271	Position nationale
DCHS2	4q31	G>A	rs2045323	Inclinaison de la columelle
DCHS2	4q31	G>A	rs2045323	Saillie du nez
DCHS2	4q31	G>A	rs2045323	Angle de la pointe du nez
RUNX2	6p21	C>T	rs1852985	Largeur du pont du nez
GLI3	7p13	C>T	rs17640804	Largeur de l'aile du nez

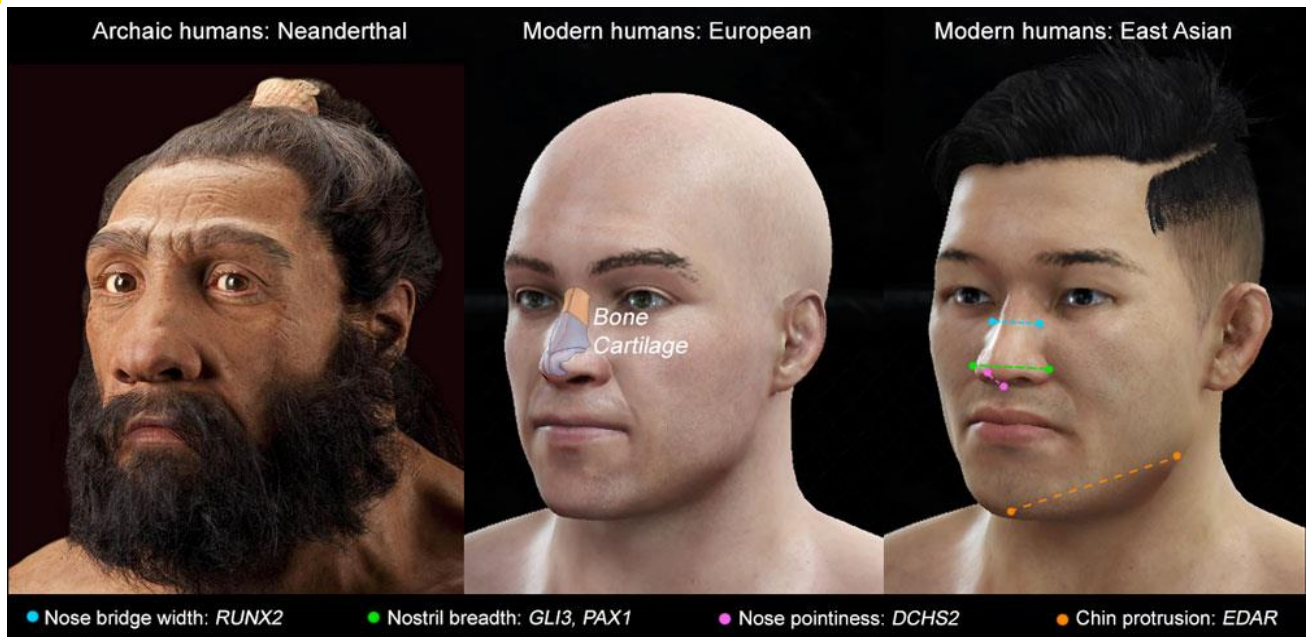


Fig. 07 : La génétique du nez : Comment nos gènes façonnent notre profil nasal (Jalinière ,2016 sites web 2)

3.5 La forme des yeux :

La forme des yeux est une caractéristique morphologique influencée par des facteurs génétiques multiples. Des gènes tels que PAX6, FOXE3 et BMP4 (**tableau VI**) sont impliqués dans le développement embryonnaire des structures faciales, y compris l'orbite et les paupières. Ces variations génétiques contribuent à la diversité interindividuelle des formes oculaires, qui peuvent être quantifiées par des outils de morphométrie faciale. Des recherches récentes ont montré que la morphologie oculaire présente un critère facial significatif pour l'identification assistée par intelligence artificielle, notamment dans les systèmes de reconnaissance faciale (**Claes et al ,2014**)

Tableau VI : Principaux gènes impliqués dans la forme des yeux humains (Claes et al ,2014)

Gène	Rôle	Impact sur la morphologie oculaire
PAX6	Gène maître du développement oculaire	Responsable du développement des structures de l'œil, y compris la cornée, l'iris et la rétine. Une mutation peut provoquer des malformations majeures.
FOXE3	Régule le développement de la lentille et des tissus oculaires	Associé à des anomalies de la lentille, mais peut aussi influencer la forme générale du globe oculaire.
EPHA2	Gène de signalisation cellulaire	Influence la forme et la taille de la lentille et joue un rôle dans la morphogenèse oculaire.
BMP4	Acteur de croissance impliqué dans la morphogenèse craniofaciale	Participe au modelage des tissus périoculaires.

ALX1 ALX3 ALX4	/ / 	Implication dans le développement craniofacial	Des mutations sont associées à des différences dans la forme du visage, qui peuvent indirectement modifier la forme des yeux.
TFAP2A		Facteur de transcription du développement facial	Régule l'expression de plusieurs gènes impliqués dans la morphogenèse de l'œil.
ZIC2		Régule le développement médian du visage	Une variation dans ce gène peut influencer l'espacement et la position des yeux (hypertélorisme ou hypotélorisme).

3.5.1 la morphologie de l'œil :

La morphologie des yeux (**Fig. 08**) fait référence à la structure physique et à la forme des différentes parties de l'œil. Cela inclut les aspects visibles et internes qui contribuent à la fonction et à l'apparence de l'œil (**Garrity et al ,2022**).

L'orbite : est une cavité osseuse qui contient le globe oculaire, les muscles, les nerfs, les vaisseaux sanguins et les structures qui sécrètent et drainent les larmes. Chacune des deux cavités orbitaires, en forme de poire, est composée de plusieurs os.

Sclère : Le revêtement extérieur de l'œil consiste en une couche blanche relativement dure appelée sclère (ou blanc de l'œil).

La cornée : La lumière pénètre dans l'œil à travers la cornée, la couche incurvée, transparente, située devant l'iris et la pupille. La cornée a une fonction protectrice ; elle contribue de plus à la convergence des rayons lumineux sur la rétine qui se trouve dans la partie postérieure de l'œil.

La pupille : les rayons lumineux traversent la pupille, la zone noire au centre de l'iris.

L'iris : partie colorée et circulaire qui entoure la pupille) contrôle la quantité de lumière qui pénètre dans l'œil.

Le Cristallin : est de forme convexe et flexible. Il devient moins flexible avec l'âge, ce qui peut entraîner des problèmes de vision tels que la presbytie, est une lentille transparente située derrière l'iris. Il permet la mise au point des objets en ajustant sa forme.

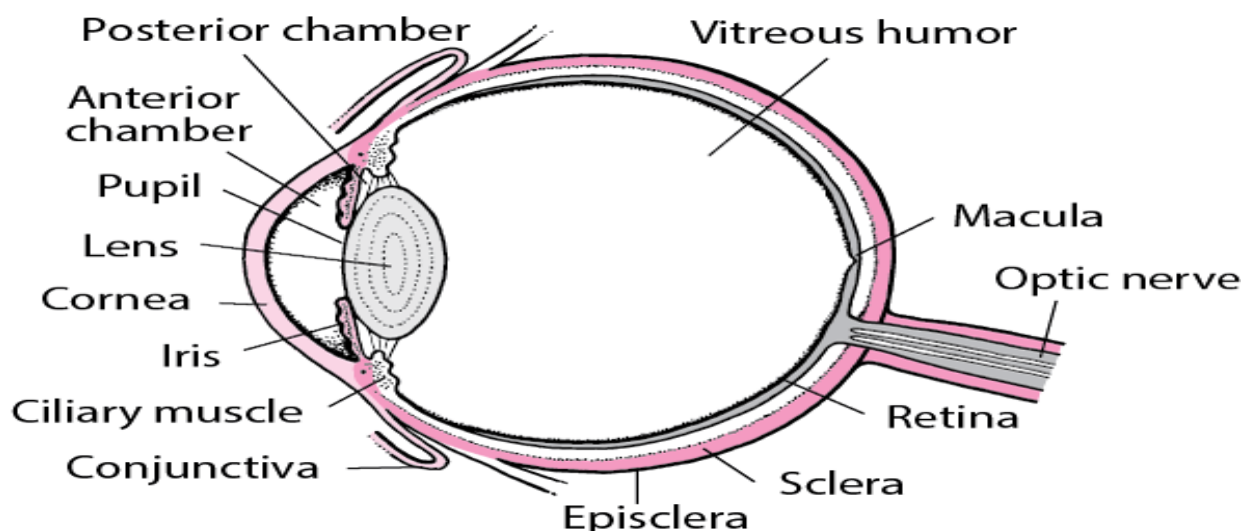


Fig. 08 : Vue à l'intérieur de l'œil (Garrity et al ,2022)

3.5.2 Impact de la forme du visage sur la morphologie des yeux :

Plusieurs paramètres morphologiques permettent d'analyser la relation entre les yeux et la forme du visage, (**fig. 09**) Parmi eux :

L'indice oculaire : constitue une mesure importante. Il reflète la taille et la forme des yeux en rapport avec la dimension globale du visage. Cet indice est particulièrement utile pour comparer la proportionnalité des yeux selon les types de visages (longs, courts, larges, etc.)

Indice oculaire = $(\text{Largeur de l'œil} / \text{Hauteur de l'œil}) \times 100$: Indice < 60 : œil rond

Indice entre 60 et 80 : œil de forme intermédiaire

Indice > 80 : œil étroit

La position des yeux : déterminée par leur distance par rapport à la largeur totale du visage, varie selon les morphotypes. Par exemple, les visages longs présentent généralement une position oculaire légèrement différente de celle observée chez les individus au visage plus court.

La distance inter-oculaire : c'est-à-dire la distance entre les coins internes des yeux, a également été mesurée et comparée entre différentes formes faciales. Les résultats révèlent des variations notables, ce qui suggère un lien entre la morphologie générale du visage et l'espacement des yeux.

La divergence oculaire : l'angle et l'orientation des yeux (plus ouverts ou plus inclinés vers le haut ou le bas) montre également des tendances spécifiques associées aux différentes structures faciales. (Jang et al ,2020)

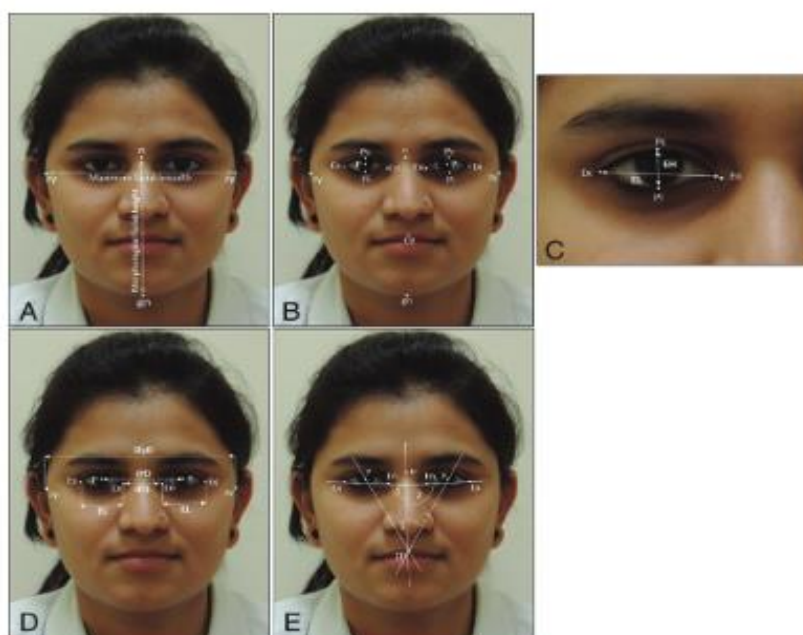


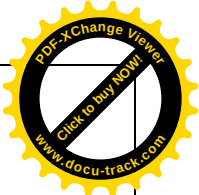
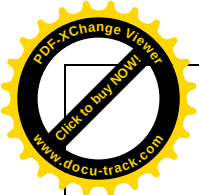
Fig. 09 : Analyse morphométrique des yeux selon la position, la forme et les dimensions oculaires en lien avec la morphologie faciale (Jang et al ,2020)

3.6- Forme de bouche :

La morphologie de la bouche, à l'instar d'autres caractéristiques faciales, est déterminée par des facteurs génétiques qui interviennent dans le développement du système nerveux central. Des études de génétique d'association à l'échelle du génome (GWAS) ont permis d'identifier plusieurs loci associés à la largeur des lèvres, à la courbure de l'arc de Cupidon et à la projection labiale. Des gènes tels que PAX9, MSX1, FGFR2 et IRF6 jouent un rôle clé dans le modelage des structures oro-faciales pendant l'embryogenèse. Ces variations morphologiques peuvent être quantifiées par des outils de morphométrie tridimensionnelle et utilisées dans des modèles de prédiction faciale à partir de l'ADN. Ainsi, il apparaît que la bouche joue un rôle crucial dans le cadre de l'identification biométrique par reconnaissance faciale. (Claes et al ,2018)

3.6.1 Définition morphologique de la forme de la bouche :

La forme de la bouche est un caractère facial clé qui englobe plusieurs dimensions anatomiques, telles que la largeur labiale, la hauteur des lèvres, la forme de l'arc de Cupidon, la courbure des commissures labiales et la projection des lèvres. Ces éléments contribuent à l'apparence générale de la région bucco-labiale et varient largement d'une personne à l'autre. Morphologiquement, la bouche est composée de



deux structures principales : la lèvre supérieure, souvent caractérisée par une forme en « arc » plus ou moins marquée, et la lèvre inférieure, généralement plus pleine et arrondie. La classification morphologique de la bouche peut se faire selon plusieurs critères, notamment la symétrie, le volume (fin, moyen ou épais), l'épaisseur, ou encore la position des commissures (relevées, neutres ou tombantes). Ces traits sont utilisés en anthropologie, en médecine esthétique, en orthodontie, et plus récemment en génétique faciale pour mieux comprendre les variations interindividuelles et les associations avec certains variants génétiques. La forme de la bouche joue également un rôle essentiel dans les expressions faciales, la parole et la perception sociale. (Kau et al ,2010).

3.6.2 Bases génétiques de la forme de la bouche :

La forme de la bouche, incluant sa taille, la largeur de ses lèvres, la courbure de l'arc de Cupidon et la projection labiale, constitue un caractère phénotypique d'une grande complexité, influencé par plusieurs gènes. Ces gènes interviennent dans le développement du système stomatognathique durant l'embryogenèse et modulent la morphologie des tissus mous et osseux autour de la cavité buccale. (Paternoster et al ,2012)

3.6.2.1 Influence génétique et les gènes impliqués :

La morphologie de la bouche, incluant les lèvres, les commissures labiales et le philtrum, est significativement déterminée par des facteurs génétiques. Le développement facial est le résultat d'une série d'interactions complexes entre gènes et signaux moléculaires durant l'embryogenèse. Plusieurs loci génétiques ont été identifiés comme associés à la morphologie labiale à travers des études d'association à l'échelle du génome (GWAS).

Parmi les principaux gènes impliqués dans ce processus, IRF6 (Interferon Regulatory Factor 6) joue un rôle central dans la fusion des structures faciales et est bien connu pour son implication dans les cas de fente labio-palatine. Il a été démontré que les variants de ce gène peuvent influencer la forme des lèvres, même en l'absence de pathologie. En outre, le gène PAX3 joue un rôle dans la détermination de la morphologie faciale générale, notamment en ce qui concerne les proportions des lèvres. Des études scientifiques ont permis d'identifier la présence de plusieurs gènes dans des populations européennes et asiatiques. Ces gènes, dont les noms sont TBX15, VAX1, PRDM16 et COL17A1, semblent jouer un rôle dans la variation naturelle de la forme de la bouche.

Les travaux de Claes et al. (2018) ont permis de modéliser la forme tridimensionnelle du visage à partir de données génétiques. Ils révèlent une influence spécifique de certains polymorphismes (SNPs) sur la courbure, l'épaisseur et la position des lèvres. Des disparités interethniques notables ont en outre été

constatées, mettant en exergue la diversité génétique faciale au sein des populations humaines. (Claes et al ,2018)

3.6.2.2 Influence des facteurs environnementaux :

La morphologie de la bouche n'est pas exclusivement dictée par le génome. Des facteurs environnementaux prénatals (comme les carences nutritionnelles (**fig.10**), les infections maternelles ou l'exposition à des substances tératogènes telles que l'alcool ou certains médicaments) peuvent perturber le développement normal du visage. Par ailleurs, des influences postnatales telles que les habitudes orales (suction prolongée du pouce, respiration buccale, posture de la langue) ou encore les fonctions oro-faciales (mastication, déglutition, phonation) peuvent moduler la croissance des structures péri-orales, affectant ainsi la forme finale des lèvres et des maxillaires. (White et al ,2021)

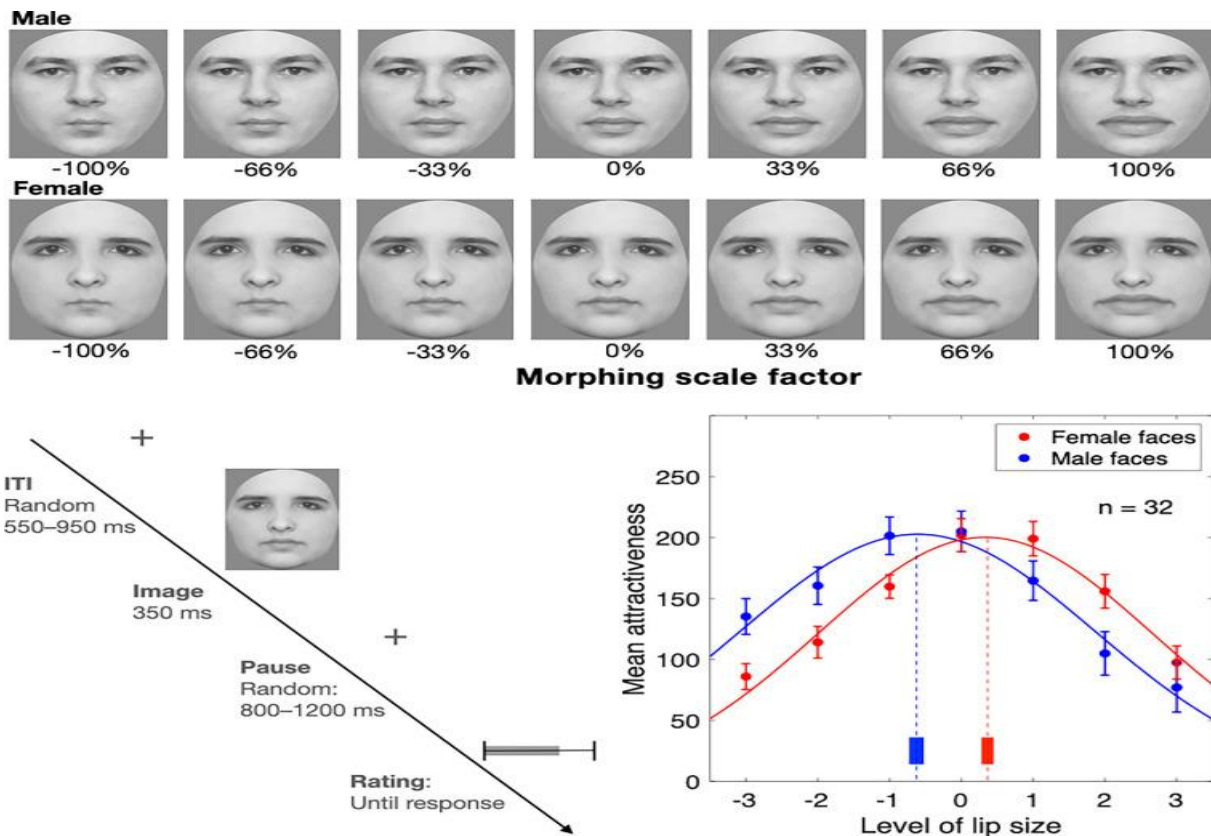


Fig.10 : Exemples de stimuli montrant deux identités et les sept niveaux de taille des lèvres. (David et al ,2025)

4. La couleur yeux comme caractère phénotypique de l'identification faciale :

La pigmentation de l'iris constitue un trait phénotypique manifeste, qui trouve son origine dans des altérations génétiques affectant divers gènes impliqués dans la synthèse et la distribution de la mélanine, tels que OCA2 et HERC2. En sa qualité de caractéristique faciale stable et aisément discernable, elle représente un indicateur adéquat pour l'identification des individus, en particulier dans le contexte des approches biométriques et de la prédiction faciale à partir des données ADN. L'intégration de cette technologie dans les systèmes d'identification assistée par intelligence artificielle permet de renforcer la précision et la fiabilité de la reconnaissance faciale. **(Chaitanya et al ,2018)**

4.1- Les bases génétiques de la couleur des yeux :

La couleur des yeux est principalement déterminée par des variations génétiques dans les gènes OCA2 et HERC2, qui régulent la production de mélanine dans l'iris. Une mutation courante, rs12913832, influence fortement cette pigmentation. D'autres gènes comme SLC24A4 ou TYR contribuent également à la diversité des couleurs observées. **(Chaitanya et al ,2018)**

4.1.1 Les bases génétiques de la couleur des yeux : les polymorphismes OCA2 et HERC2 (fig.11) :

La couleur des yeux présente une pigmentation variable en fonction des différences de quantité, de type et de distribution de la mélanine obtenue par le processus de mélano-genèse dans les mélanocytes , Pour comprendre l'étendue des variations phénotypiques dans les traits pigmentaires, il faut prendre en compte la régulation de la mélano-genèse, car ce processus en plusieurs étapes est sous contrôle génétique impliquant de multiples gènes, qui affectent globalement le résultat de la couleur des yeux chez l'homme. Dans le passé, OCA2 était considéré comme le principal gène impliqué dans la couleur des yeux. Cependant, il a été démontré que le gène HERC2, situé très près de l'OCA2, joue un rôle essentiel dans la détermination de la couleur de l'iris, en influençant l'expression de l'OCA2.

Les interactions géniques sont souvent impliquées dans la détermination de traits complexes ; ainsi, la pertinence d'un seul gène sur des caractéristiques phénotypiques complexes ne peut pas être déterminée de manière simple en raison des effets épistatiques entre différents gènes, même si l'interaction entre les gènes HERC2 et OCA2 ne peut pas être comparée à un effet épistatique classique. En effet, HERC2 influence le gène OCA2 au niveau de l'ADN en modulant son activité transcriptionnelle, contrairement à d'autres gènes dont l'interaction dans la détermination de la couleur des yeux est différente. En effet, outre les principaux déterminants de la couleur des yeux HERC2/OCA2, avec plus de 50 SNP liés à la couleur des yeux, d'autres gènes se sont révélés impliqués dans la couleur des yeux. Avec HERC2 et OCA2, ces gènes sont les meilleurs prédictors de la couleur des yeux bleus, bruns et intermédiaires,

sachant que le phénotypage de la couleur des yeux a suivi plusieurs propositions, notamment clair vs foncé et bleu vs brun. Ce n'est que récemment qu'une autre catégorisation a été proposée : bleu (comprenant le bleu clair à foncé), intermédiaire (comprenant le vert et le noisette) et marron (comprenant le marron clair à foncé) (**Brancato et al ,2023**)

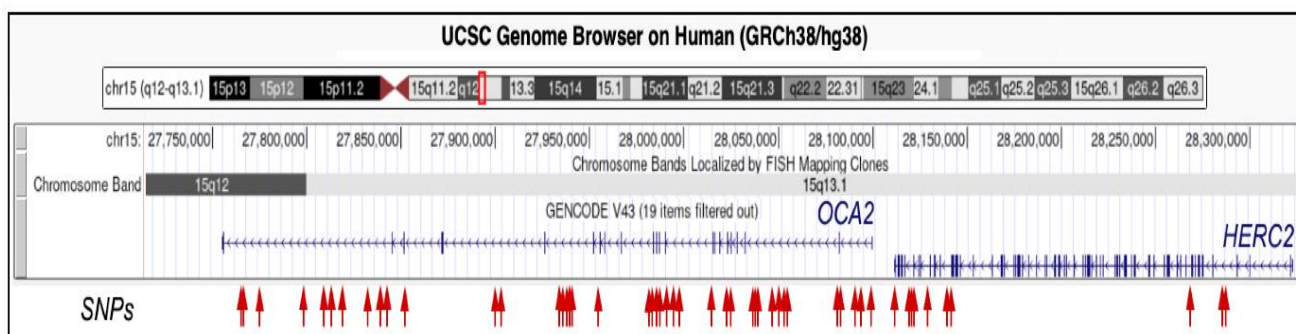


Fig.11 : Organisation génomique des gènes OCA2 et HERC2 (Brancato et al ,2023)

4.1.1.1 –Le Gène OCA2 :

Le gène OCA2 (NM_000275.3), localisé sur le chromosome 15q13.1, s'étend sur une région d'environ 345 kilobases et comprend 24 exons. Le sujet code une protéine composée de 838 acides aminés, désignée par le nom de protéine P. Cette protéine présente une homologie avec le gène pink-eyed dilution chez la souris. Cette protéine membranaire intégrale des mélanosomes contient 12 domaines transmembranaires et participe à la régulation du pH intra-mélanosomal, ainsi qu'au transport et à la stabilisation de protéines clés telles que la tyrosinase. Par conséquent, toute modification de la protéine P peut influencer sur la pigmentation par altération de la disponibilité ou de la fonction de la tyrosinase. Cette fonction, préservée tout au long de l'évolution, est validée par l'existence de mutations du gène responsable de l'albinisme oculocutané de type 2 (OCA2). Cette pathologie, classée comme autosomique récessive, se manifeste par une production insuffisante ou nulle de mélanine au sein des mélanocytes cutanés, capillaires et oculaires.

La variabilité génétique d'OCA2 est fortement corrélée à la couleur des yeux, avec une contribution estimée à environ 74 % de la variance phénotypique entre les yeux bleus et marron. Ce gène est donc reconnu comme l'un des principaux déterminants génétiques de la pigmentation de l'iris. À ce jour, 44 allèles variants non pathogènes ont été identifiés dans OCA2, avec des fréquences variables selon les populations, ce qui explique la diversité ethnique observée dans les phénotypes pigmentaires. La

majorité de ces polymorphismes (37 sur 44) se trouvent dans les régions introniques, tandis que six sont situés dans des régions codantes, dont l'un représente une substitution synonyme.

Certains SNP introniques, tels que rs7495174 (A/G), rs4778241 (anciennement rs6497268, A/C) et rs4778138 (anciennement rs11855019, A/G), localisés dans l'intron 1 du gène, sont particulièrement associés à la distinction yeux bleus/marron. Par ailleurs, deux SNP codants, rs1800401 (G/A → Arg305Trp) et rs1800407 (C/T → Arg419Gln), ont été corrélés respectivement à la pigmentation des yeux marron et aux nuances vert/noisette. Des effets épistatiques ont également été mis en évidence, notamment entre rs1800407 (OCA2) et rs12913832 (HERC2), ce dernier agissant comme un régulateur transcriptionnel majeur d'OCA2. En tenant compte également du SNP rs1129038 (HERC2), fortement en liaison déséquilibrée avec rs12913832, la précision des prédictions de couleur intermédiaire de l'iris s'en trouve renforcée (**Balding et al ,2023**).

4.1.1.2-Le Gène HERC2 :

Le gène HERC2 (NM_004667), situé à proximité du gène OCA2 sur le chromosome 15q13.1, s'étend sur une région d'environ 211 kilobases et comprend un total de 98 exons. Cette protéine, de nature ligase E3 et appartenant à la famille HERC, se distingue par la présence d'un domaine RCC1 et d'un domaine catalytique HECT. En plus de son rôle en ubiquitination et homéostasie cellulaire, des variants dans HERC2 sont associés à des pathologies neurologiques, telles que le syndrome d'Angelman, l'autisme et certaines formes de déficience intellectuelle. Bien que ses fonctions régulatrices soient encore partiellement élucidées, le SNP rs12913832 situé dans l'intron 86 est aujourd'hui reconnu comme le principal marqueur génétique de la couleur des yeux, en raison de son association forte avec le niveau d'expression de OCA2. (**Cook et al ,2012**)

Des études expérimentales ont révélé que rs12913832 agit comme un enhancer à distance, modulant la mise en boucle chromatinienne entre la région régulatrice de HERC2 et le promoteur d'OCA2, ce processus est facilité par l'intégration de facteurs de transcription spécifiques, tels que HLTF, LEF1 et MITF, conduisant à une augmentation de l'expression d'OCA2. L'allèle « C » (brun) favorise une conformation ouverte de la chromatine, induisant une expression accrue d'OCA2. En revanche, l'allèle « T » (bleu) inhibe ce processus de repliement, ce qui se traduit par une diminution de la pigmentation de l'iris. (**Salvo et al ,2023**)

4.1.1.3 Modèle pour la régulation du gène OCA2 :

Couleurs des yeux (**Fig.12**), représentatives liées aux trois génotypes du SNP rs12913832. Le génotype homozygote A/A est lié à la couleur brune plus foncée (images du haut). Au contraire, le génotype homozygote G/G est lié à la couleur bleue plus claire des yeux (images du bas). Le génotype

hétérozygote A/G est lié à une couleur intermédiaire des yeux, notamment par l'iris avec une quantité moindre de mélanine (respectant le génotype A/A) déterminant une intensité moindre de la couleur brune ou vert-hazel



Fig.12 : Effet phénotypique du SNP rs12913832 : A/G sur la couleur des yeux. (Brancato et al ,2023)

4.1.2 -Autres gènes liés à la couleur des yeux :

En plus de OCA2 et HERC2, plusieurs autres gènes participent à la variation de la couleur des yeux (**fig.13**). Parmi eux, SLC24A4, TYR, IRF4 et SLC45A2 jouent un rôle dans la régulation de la mélanogenèse ou dans le transport des ions liés à la pigmentation. Bien que leur influence soit moins dominante, ces gènes contribuent à affiner les nuances intermédiaires comme le vert ou le noisette.

4.1.2.1 –Le Gène TYR :

Le gène TYR, situé sur le chromosome 11q14.3, est crucial pour la pigmentation humaine en codant pour la protéine tyrosinase impliquée dans la biosynthèse de la mélanine. Les mutations de ce gène sont liées à l'albinisme OCA1, entraînant une production réduite de mélanine. Le TYR comprend 5 exons et code une glycoprotéine liée à la membrane mélanosomale. Les SNP du locus TYR, tels que rs1126809 : G/A et rs1393350 : G/A, sont associés à des variations de pigmentation, notamment pour la couleur des yeux, des cheveux et de la peau. Cependant, bien que certains SNP soient liés à des caractéristiques spécifiques (yeux bleus, cheveux blonds, peau sensible au soleil), aucune sélection positive ou homozygotie étendue de l'haplotype n'a été identifiée. Ces polymorphismes influencent la

diversité de pigmentation sans montrer de corrélations claires avec des phénotypes spécifiques. (Brancato et al ,2023).

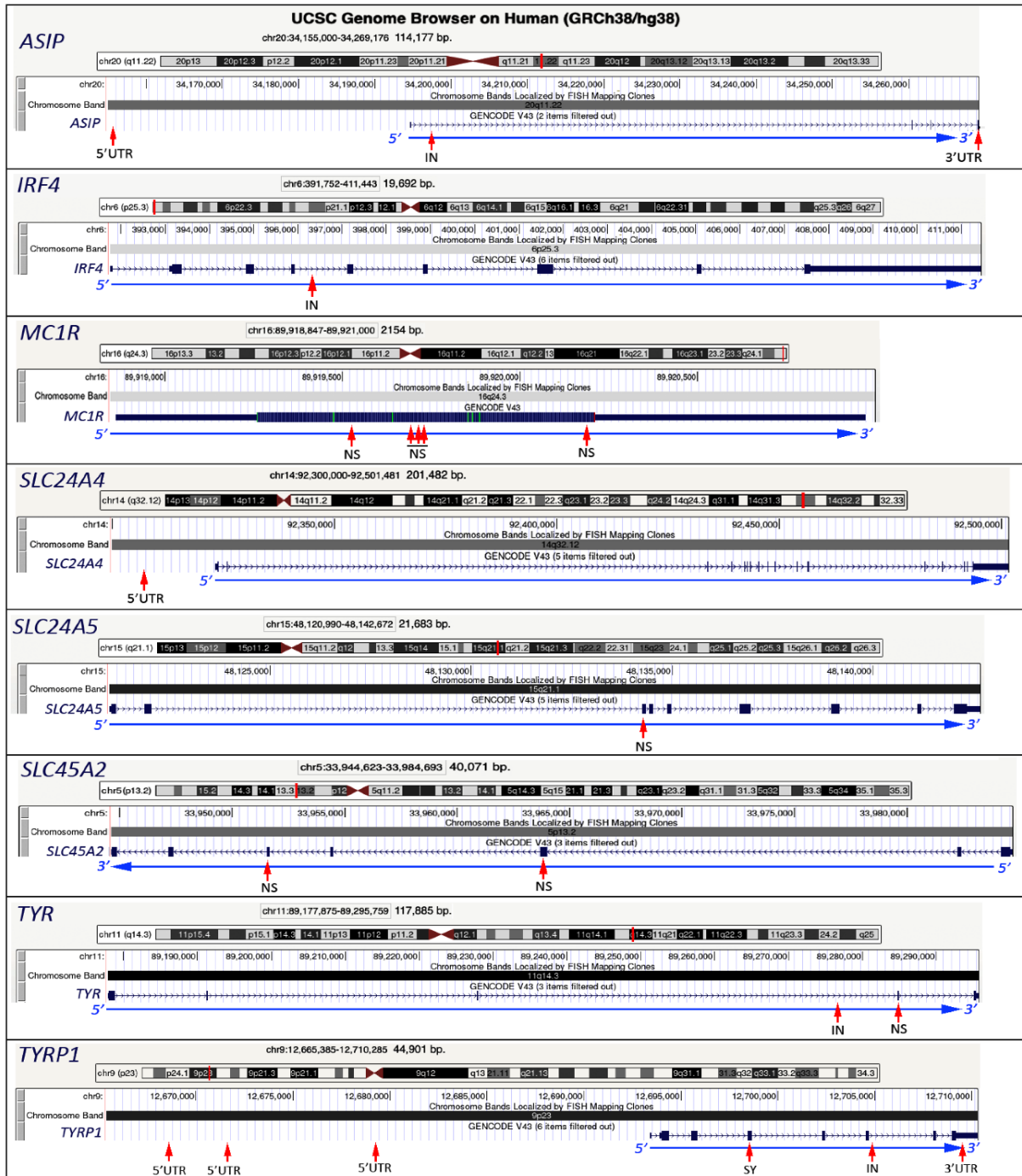
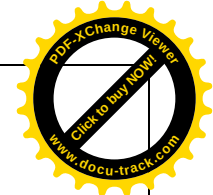
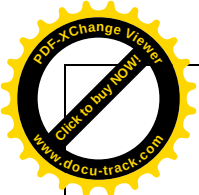


Fig.13 : Organisation génomique des principaux gènes impliqués dans la couleur des yeux autres que HERC2/OCA2. (<http://genome.ucsc.edu> site web 3 ,2023)



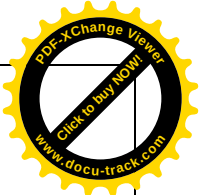
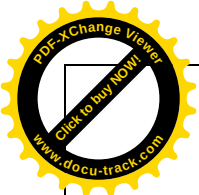
4.1.2.2 –Le Gène TYRP1 et DCT :

Le gène TYRP1, situé sur la bande chromosomique 9p23, est impliqué dans les différences de pigmentation et est associé à l'albinisme oculocutané (OCA3). Il code une enzyme tyrosinase participant à la biosynthèse de la mélanine et est essentiel à la production d'eumélanine. Des variations génétiques telles que rs1408799 : T/C et rs683 : C/A jouent un rôle dans la prédiction de la couleur des yeux, tandis que d'autres variants comme rs62538956 : T/C et rs35866166 : T/C, influencent la production d'eumélanine. Une régulation négative de TYRP1 entraîne une réduction de la pigmentation et peut conduire à des yeux bleus. Ces découvertes mettent en lumière l'importance du gène TYRP1 dans la pigmentation cutanée et des yeux, ainsi que ses liens avec l'albinisme oculocutané. **(Brancato et al ,2023)**

Alors que TYR catalyse l'étape initiale clé de la production de mélanine, TYRP1 et DCT (*TYRP2*) agissent lors d'étapes ultérieures, influençant la quantité et la qualité des mélanines et stabilisant l'enzyme TYR. Ils ont également une certaine fonction dans la survie des mélanocytes après le stress du rayonnement ultraviolet (UVR) et dans le maintien des structures mélanomanes. La famille de gènes TYRP a évolué par duplication génique récurrente à partir d'un TYR ancestral commun au début de l'évolution, donnant lieu à TYRP1 et à DCT. Chez l'homme, des mutations de TYRP1 à l'origine de l'OCA3 sont fréquemment observées en Afrique du Sud. Une partie de ce phénotype est la peau rouge au bronze, les cheveux roux et les irides bleus. Dans une découverte surprenante, il a été récemment rapporté qu'un changement d'acides aminés non synonyme, Arg93Cys, dans la protéine TYRP1 est un déterminant majeur du phénotype blond des cheveux. Cet allèle a une fréquence de 0,26 sur les îles Salomon et n'est pas vu en dehors de l'Océanie, de sorte qu'il représente un événement évolutif très récent, l'allèle agissant de manière récessive. De même, une très haute fréquence d'une mutation *OCA2* spécifique à la population provoquant l'expression d'un phénotype d'albinisme complet a été trouvée dans les îles Polynésiennes. Les deux cas de dérive génétique résultant d'effets de la population retrouvés (encadré 1). Dans la population européenne, *TYRP1* rs1408799-A a été associé à la couleur des yeux bleus, bien que d'autres SNP soient également associés aux yeux bleus, de sorte que le SNP causal n'a pas encore été reconnu. **(Sturm et Duffy ,2012)**

4.1.2.3 –Le Gène SLC45A2 :

Au cours des dernières années, le gène SLC45A2, également connu sous le nom de MATP, a été associé à l'albinisme OCA4 avec des mutations affectant la pigmentation. La protéine MATP, découverte comme l'antigène AIM1 dans le mélanome humain, est régulée par MITF et présente des polymorphismes pertinents pour la pigmentation des yeux. Les allèles rs26722 et rs16891982 montrent



des distributions de fréquence distinctes et des effets sur la production de mélanine. L'allèle G favorise la production d'eumélanine, tandis que l'allèle C réduit la fonctionnalité de la protéine et affecte le trafic intracellulaire des protéines mélanosomales. Ces différences ont des conséquences sur l'absorption des rayons UVB selon les latitudes, influençant la synthèse de la vitamine D3. **(Brancato et al ,2023)**

Selon Soejima et Koda, Le gène SLC45A2, également connu sous le nom de MATP, joue un rôle important dans la synthèse de la mélanine. Le polymorphisme rs16891982 (Leu374Phe) est fortement associé à une couleur plus claire de la peau, des cheveux et des yeux chez les populations européennes. Le variant 374Phe (allèle C) montre une fonction réduite et serait impliqué dans la régulation du pH des mélanosomes ainsi que dans le trafic des protéines, contribuant ainsi à une production moindre de mélanine. **(Soejima et Koda ,2007)**

4.1.2.4 –Le Gène SLC24A4 :

SLC24A4 (solute carrier family 24 member 4) est situé en 14q32.12 est un gène qui appartient à la même famille que SLC24A5, contient 17 exons couvrant 178 kb, et code un membre de la famille des protéines échangeuses sodium/calcium dépendantes du potassium. Les variantes génétiques de SLC24A4 contribuent à la variation de la pigmentation de la peau, des cheveux et des yeux. En particulier, le gène rs12896399 : G/T a été initialement associé à la couleur des yeux, et la plus grande différence de fréquence d'allèle pour le gène rs12896399 a été observée entre les individus aux yeux bleus et les individus aux yeux verts. L'effet de masquage du SNP HERC2 rs12913832 sur rs12896399 dans la détermination de la couleur des yeux bleus a été révélé. Il a été démontré que les génotypes GG, GT et TT de rs12896399 sont strictement liés à rs12913832. En effet, parmi ces génotypes, rs12896399 :TT a été classé avec une "forte probabilité" d'avoir les yeux bleus lorsqu'il est associé au génotype GG dans rs12913832. L'allèle A dans rs12913832 dans HERC2 masque cet effet, et, indépendamment du génotype dans rs12896399, les individus sont classés dans le groupe de "faible probabilité" d'avoir les yeux bleus. **(Brancato et al ,2023)**

Selon Liu, ce SNP présente une association significative avec la variation de la couleur des yeux, mais son effet est modulé par le polymorphisme rs12913832 situé dans le gène HERC2. En effet, ce dernier peut masquer l'effet du rs12896399, révélant ainsi une interaction épistatique entre HERC2 et SLC24A4. Les génotypes GG, GT et TT de rs12896399 sont fortement liés à rs12913832. Ainsi, le génotype TT dans SLC24A4, associé à GG dans HERC2, est corrélé à une forte probabilité d'avoir les yeux bleus, tandis que la présence de l'allèle A dans rs12913832 masque cet effet, réduisant

significativement la probabilité de cette caractéristique pigmentaire, quel que soit le génotype du rs12896399. (Liu et al ,2010)

4.1.2.5 –Le Gène SLC24A5 :

Le gène SLC24A5 (solute carrier family 24 member 5), situé sur le bras long du chromosome 15q21.1, comporte 9 exons s'étendant sur 22 kb et code la protéine NCKX5, un transporteur d'ions jouant un rôle essentiel dans le contrôle du pH des mélanosomes. Des polymorphismes nucléotidiques non synonymes de ce gène ont été associés à la variation de la couleur de la pigmentation dans des populations homogènes. Le SNP rs1426654 (A/G), localisé dans le troisième exon, entraîne une substitution de la thréonine par une alanine en position 111 de la protéine (Thr111Ala), ce qui réduit l'efficacité du transport ionique et de la production de phéomélanine. Le génotype homozygote rs1426654 : A est fortement associé aux yeux bleus et sa fréquence varie considérablement selon les populations. Ainsi, l'allèle A (Thr111) est quasi fixé chez les Européens (98–100 %), tandis que l'allèle ancestral G (Ala111) est largement répandu chez les Africains (93–100 %). En 2014, ce SNP a été identifié comme le premier marqueur permettant de distinguer les traits noirs des traits non-noirs, soulignant ainsi l'importance du gène SLC24A5 dans les applications médico-légales. Ce polymorphisme présente également un profil spécifique selon les populations, similaire à celui du variant p. Leu374Phe du gène SLC45A2, avec un allèle dérivé A fixé chez les Européens. (Brancato et al ,2023)

4.1.2.6 –Le Gène ASIP :

Étant donné le rôle que joue l'ASIP (NM_001672) dans le passage de la production d'eumélanine à celle de phaeomélanine, il s'agit de l'un des nombreux candidats solides pour expliquer la variation interindividuelle de la pigmentation chez l'homme. L'ASIP est situé sur 20q11.22, avec 4 exons qui couvrent une région de 28 kb, code pour la protéine de signalisation Agouti, une protéine de 132 acides aminés, et est exprimé dans plusieurs tissus. Elle joue un rôle dans la régulation de la pigmentation en inhibant la synthèse de l'eumélanine et en induisant la production de phéomélanine. Dans l'ASIP, le SNP rs6058017 se trouve dans le 3' UTR du gène, 25 pb en aval du codon STOP, et correspond à une substitution de l'adénine par une guanine (A/G). Le mécanisme d'action proposé pour le SNP ASIP rs6058017 : G implique une stabilité réduite de l'ARNm et une dégradation prématurée de la transcription, exerçant moins d'antagonisme envers MC1R, qui se liera à l' α -Melanocyte-stimulating hormone (α -MSH), conduisant à la production d'eumélanine, à savoir les yeux foncés. L'allèle A du polymorphisme susmentionné est présent à des fréquences élevées (>75 %) dans la population

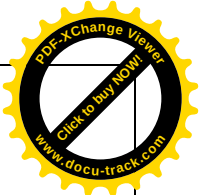
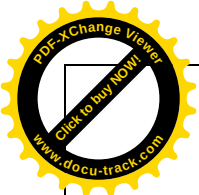
européenne, et l'allèle G (l'allèle ancestral) est présent dans les populations d'Asie de l'Est (plus de 60 %) et afro-américaines (40 %). **(Brancato et al ,2023)**

D'après Bonilla, l'allèle ancestral G est plus fréquent chez les populations d'Afrique et d'Asie de l'Est, tandis que l'allèle dérivé A prédomine en Europe (>75 %). L'allèle G est associé à une expression réduite de ASIP, permettant une activation accrue de MC1R par l' α -MSH, ce qui conduit à une production plus importante d'eumélanine, et donc à une pigmentation plus foncée, notamment au niveau des yeux. Ce mécanisme implique une dégradation plus rapide de l'ARNm de ASIP, réduisant son effet antagoniste sur MC1R. Ces différences de fréquence reflètent probablement des adaptations évolutives à des environnements lumineux variables. **(Bonilla et al ,2005)**

4.1.2.7 –Le Gène IRF4 :

Le gène IRF4 (interferon regulatory factor 4), situé sur le chromosome 6p25.3, s'étend sur 20 kb et comprend 9 exons. Initialement associé aux cellules pigmentaires, il a ensuite été impliqué dans la pigmentation humaine. En 2013, il a été démontré que le régulateur des mélanocytes MITF active l'expression du gène IRF4, à condition de la présence du facteur de transcription TFAP2A. Un variant génétique chevauchant le site de liaison de TFAP2A diminue cette liaison, réduisant ainsi l'expression d'IRF4. Praetorius et al, ont précisé ce mécanisme en montrant que l'allèle T du polymorphisme rs12203592, localisé dans un enhancer spécifique aux mélanocytes situés dans l'intron 4 d'IRF4, perturbe la fixation de TFAP2A, entraînant une baisse de l'expression du gène. Or, comme IRF4 et MITF coopèrent pour activer l'expression de la tyrosinase (TYR), enzyme clé de la synthèse de mélanine, ce variant entraîne une diminution de la pigmentation via une réduction des niveaux de TYR. Ce SNP est ainsi associé à la couleur des cheveux, de la peau et des yeux chez les populations européennes. **(Praetorius et al ,2013)**

En 2015, il a été prouvé que cet enhancer interagit avec le promoteur d'IRF4 via une boucle chromatinienne dépendante de l'allèle, renforçant ainsi l'activation spécifique de la transcription d'IRF4. Le rs12203592 a d'abord été identifié comme un déterminant de la couleur des cheveux chez les femmes européennes dans une étude GWAS, avant d'être également associé à une peau plus claire, une faible capacité de bronzage et des yeux clairs ou bleus. Toutefois, bien qu'il soit inclus dans l'IrisPlex, aucune association n'a été trouvée entre rs12203592 et la couleur des yeux chez les



individus porteurs du génotype rs12913832 :GG, car les associations observées étaient catégoriques mais non quantitatives. **(Brancato et al ,2023)**

5- LA PROGRAMMATION : Codage et programmation :

La programmation fait souvent référence au fait d'analyser des problèmes et de les résoudre à l'aide d'algorithmes, puis de traduire ceux-ci dans l'un des nombreux langages de programmation existants, et les algorithmes sont des méthodes de résolution de problèmes à l'aide de différentes étapes de traitement d'un certain nombre de données.

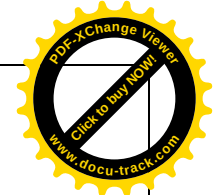
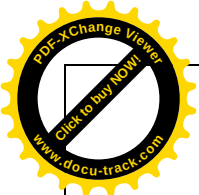
De manière générale, lorsqu'on parle de codage, on fait référence au fait de traduire en langage de programmation un algorithme existant. Bref, la tâche de codage est plus une opération de traduction que de raisonnement en soi.

Que l'on parle de coder ou de programmer en milieu scolaire, l'important sur le plan pédagogique est ceci : plus la tâche implique une grande portion d'analyse, plus on vise l'utilisation de la métacognition et le développement d'aptitudes en résolution de problèmes

La robotique est l'une des façons d'appliquer des notions de programmation. Dans un contexte pédagogique, elle permet de concrétiser certaines notions qui seraient souvent beaucoup plus abstraites si elles étaient enseignées de manière formelle ou traditionnelle. De plus, certaines applications robotiques utilisées en milieu scolaire ne nécessitent peu ou pas de notions approfondies en programmation. **(Desjardins et al ,2018)**

Les différents langages de programmation : Les plus connus et utilisés sont les suivants :

- Java
- C
- C++
- C#
- Objective-C
- Python
- Visual Basic .NET
- HP
- JavaScript



5.1 - Analyse bioinformatique :

La bio-informatique est Ensemble de méthodes, de logiciels et d'applications en ligne qui permettent de gérer, manipuler, et analyser des données biologiques.

5.1.1 - GWAS: pour génome -Wide association study:

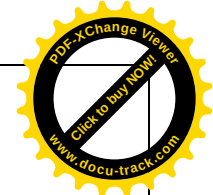
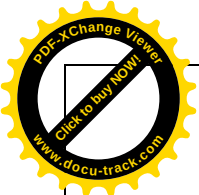
Les études d'association à l'échelle du génome (GWAS) visent à identifier les associations entre les génotypes et les phénotypes en testant les différences dans la fréquence des allèles des variantes génétiques entre des individus qui sont similaires sur le plan ancestral mais qui diffèrent sur le plan phénotypique. Les études d'association pangénomique peuvent prendre en compte des variantes du nombre de copies ou des variations de séquence dans le génome humain, bien que les variantes génétiques les plus couramment étudiées dans les études d'association pangénomique soient les polymorphismes mononucléotidiques (SNP).

Les études d'association pangénomiques font généralement état de blocs de SNP corrélés qui présentent tous une association statistiquement significative avec le trait étudié, connu sous le nom de loci de risque génomique. Après 15 ans de GWAS¹, de nombreux loci de risque génomique répliqués ont été associés à des maladies et à des traits. **(Uffelmann et al ,2021)**

5.1.2- Meta analyse :

La méta-analyse est une démarche de recherche scientifique qui repose sur un protocole précis. Correctement réalisée, elle offre une évaluation objective des données de la littérature sur un sujet donné et contribue à l'avancement des connaissances. Considérées en général de très haut niveau de preuve, les méta-analyses sont souvent utilisées pour établir des recommandations sur la pratique clinique. La faiblesse méthodologique de certaines d'entre elles et les difficultés éventuelles d'interprétation du résultat global conduisent à conseiller aux lecteurs de conserver leur sens critique. **(Coulm ,2022).**

La méta-analyse est une méthode permettant de réaliser un tel travail en combinant les résultats de plusieurs études pour faire une synthèse objective selon un protocole précis et ainsi reproductible. Elle permet aussi de quantifier le résultat global pour l'ensemble des études considérées. Elle répond à une méthode précise tant pour la recherche, la sélection, la présentation et l'analyse des études disponibles pour une question donnée **(Patrick , 2010 Site web 4)**



5.1.3 -Le système HirisPlex-S :

Le système HirisPlex-S (**Fig.14**), représente une avancée majeure dans le domaine du phénotypage génétique prédictif appliqué à la médecine légale. Cet outil combine 41 SNPs répartis sur plusieurs gènes clés impliqués dans la pigmentation (tels que *HERC2*, *SLC24A5*, *MC1R*, *TYR* et *IRF4*), permettant ainsi de prédire simultanément la couleur des yeux, des cheveux et de la peau à partir d'un simple échantillon d'ADN, même s'il est dégradé ou ancien. Validé sur des échantillons modernes et archéologiques, HirisPlex-S a démontré un haut niveau de précision (jusqu'à 90 % pour les yeux, 80 % pour les cheveux et 75 % pour la peau), ce qui en fait un outil de choix pour les enquêtes médico-légales et l'identification de restes humains. Au-delà de son intérêt en criminalistique, ce système s'impose comme une solution efficace pour reconstituer l'apparence d'individus historiques ou non identifiés, en l'absence d'indices visuels ou de correspondances ADN dans les bases de données récemment traditionnelles ce qui est connu sous le nom de « test ADN ». (**Chaitanya et al ,2018**)

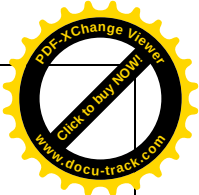
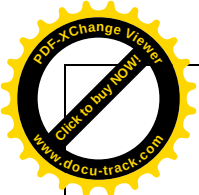
Gène	SNP	Allèle	Nombre d'allèles
1 MC1R	rs312262906	UN	0 1 2 NA
2 MC1R	rs11547464	UN	0 1 2 NA
3 MC1R	rs885479	T	0 1 2 NA
4 MC1R	rs1805008	T	0 1 2 NA
5 MC1R	rs1805005	T	0 1 2 NA
6 MC1R	rs1805006	UN	0 1 2 NA
7 MC1R	rs1805007	T	0 1 2 NA
8 TUBB3	rs1805009	C	0 1 2 NA
9 MC1R	rs201326893	UN	0 1 2 NA
10 MC1R	rs2228479	UN	0 1 2 NA
11 MC1R	rs1110400	C	0 1 2 NA
12 SLC45A2	rs28777	C	0 1 2 NA
13 SLC45A2	rs16891982	C	0 1 2 NA
14 KITLG	rs12821256	G	0 1 2 NA
15 LOC105374875	rs4959270	UN	0 1 2 NA
16 IRF4	rs12203592	T	0 1 2 NA
17 TYR	rs1042602	T	0 1 2 NA
18 OCA2	rs1800407	UN	0 1 2 NA
19 SLC24A4	rs2402130	G	0 1 2 NA
20 HERC2	rs12913832	T	0 1 2 NA
21 PIGU	rs2378249	C	0 1 2 NA
22 LOC105370627	rs12896399	T	0 1 2 NA
23 TYR	rs1393350	T	0 1 2 NA
24 TYRP1	rs683	G	0 1 2 NA
25 ANKRD11	rs3114908	T	0 1 2 NA
26 OCA2	rs1800414	C	0 1 2 NA
27 BNC2	rs10756819	G	0 1 2 NA
28 HERC2	rs2238289	C	0 1 2 NA
29 SLC24A4	rs17128291	C	0 1 2 NA
30 HERC2	rs6497292	C	0 1 2 NA
31 HERC2	rs1129038	G	0 1 2 NA
32 HERC2	rs1667394	C	0 1 2 NA
33 TYR	rs1126809	UN	0 1 2 NA
34 OCA2	rs1470608	UN	0 1 2 NA
35 SLC24A5	rs1426654	G	0 1 2 NA
36 ASIP	rs6119471	C	0 1 2 NA
37 OCA2	rs1545397	T	0 1 2 NA
38 RASSEMBLEMENT	rs6059655	T	0 1 2 NA
39 OCA2	rs12441727	UN	0 1 2 NA
40 MC1R	rs3212355	UN	0 1 2 NA
41 DEF8	rs8051733	C	0 1 2 NA

Fig. 14 : représentation des 41 SNP de système Hirisplex-s pour la prédiction de couleur de peau cheveux et yeux d'un individu (<https://hirisplex.erasmusmc.nl/> Site web 1)

5.2- Programmation python :

Le langage de programmation Python a été créé en 1989 par Guido van Rossum, aux Pays-Bas. Le nom Python vient d'un hommage à la série télévisée Monty Python's Flying Circus dont G. van Rossum est fan.

La première version publique de ce langage a été publiée en 1991. La dernière version de Python est la version 3. Plus précisément, la version 3.11 a été publiée en octobre 2022.

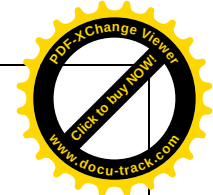
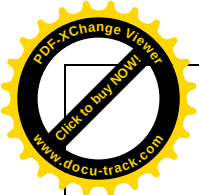


Ce langage de programmation présente de nombreuses caractéristiques intéressantes :

Il est multiplateforme c'est-à-dire qu'il fonctionne sur de nombreux systèmes d'exploitation : Windows, Mac OS X, Linux, Android, iOS, depuis les mini-ordinateurs Raspberry Pi jusqu'aux supercalculateurs.

C'est un langage de haut niveau. Il demande relativement peu de connaissance sur le fonctionnement d'un ordinateur pour être utilisé. C'est un langage interprété. Un script Python n'a pas besoin d'être compilé pour être exécuté, contrairement à des langages comme le C ou le C++. • Il est orienté objet c'est-à-dire qu'il est possible de concevoir en Python des entités qui miment celles du monde réel (une molécule d'ADN, une protéine, un atome, etc.) avec un certain nombre de règles de fonctionnement et d'interactions. C'est le langage de programmation le plus utilisé au monde (voir les classements TIOBE 3 et IEEE Spectrum 4).

Enfin, il est très utilisé en bio-informatique, chimie-informatique et plus généralement en analyse de données. **(Fuchs et Poulain ,2024 Site web 5)**



Matériel et méthode :

L'objectif de notre problématique est de prédire et créer une image de profil facial inconnu à partir de trace biologique. Pour cela nous avons rassemblé une panoplie d'analyses et de codage pour essayer de parvenir à une réponse eu égard à notre objectif.

Pour la réalisation de ce projet d'identification faciale à partir de données génétiques, nous avons opté pour le choix d'un seul caractère physique héréditaire en premier lieu : c'est la couleur des yeux, et le choix de ce caractère est justifié par la richesse non seulement dans le volume de données, mais également dans la précision et la variété de ses dernières.

En second lieu, la pigmentation oculaire a été sélectionnée comme un trait phénotypique d'intérêt et ceci pour plusieurs raisons biologiques et techniques. D'abord il convient de noter que la couleur des yeux constitue un caractère visuellement observable et aisément quantifiable. Cette propriété le positionne comme un sujet d'étude particulièrement pertinent pour la prédiction phénotypique, fondée sur l'analyse des données génétiques. Ensuite, la pigmentation des yeux est fortement influencée par des variants génétiques bien identifiés, notamment dans les gènes *HERC2*, *OCA2*, *TYR*, *SLC45A2*, *SLC24A4* et *IRF4* ; Ces derniers ont subi différentes modifications dans leur séquence par plusieurs moyens : soit de nombreuses mutations ou bien des SNP, ce qui nous donne différentes expressions de ses allèles permettant ainsi de nous donner de nombreux phénotypes différents. Les rôles de ces variants dans la biogenèse de la mélanine sont bien établis. Cette association génotype-phénotype relativement simple permet une modélisation efficace par des algorithmes d'apprentissage automatique.

On a commencé notre travail par la création d'une nouvelle DATA7 spécifique à partir des bases de données. Cette dernière est limitée par quelques gènes, et centrée sur les gènes associés à la couleur des yeux, dans le but de relier l'information génétique aux phénotypes, et ceci afin de développer un modèle prédictif (en Python) pour estimer la couleur des yeux.

1-Analyse bioinformatique :

Il s'agit d'une étude qui combine entre la génétique et la bio-informatique, Elle vise à explorer la relation entre certains polymorphismes nucléotidiques simples (SNPs) et les traits faciaux, en utilisant des modèles d'apprentissage automatique pour prédire certains aspects phénotypiques à partir de génotypes spécifiques.

1.1 -Sources de données :

Une source de données est base d'où proviennent les informations que l'on utilise pour faire notre étude.

1.1.1- Données génétiques :

Les données SNP utilisées sont issues de bases reconnues :

- HIrisPlex-S, IrisPlex : Les bases de données IrisPlex (couleur des yeux) et HIrisPlex-S (couleur de peau, cheveux et les yeux) ont été utilisé dans cette étude en raison de leur validation scientifique et de leur pertinence dans le domaine de la prédiction phénotypique à partir de l'ADN. Ces systèmes permettent d'établir un lien direct entre des polymorphismes nucléotidiques simples (SNPs) spécifiques et les caractéristiques visibles du visage, telles que la couleur des yeux, de la peau et des cheveux. Leur efficacité a été démontrée dans de nombreuses études avec une précision de prédiction atteignant jusqu'à 94% pour la couleur des yeux. En outre, la compatibilité de ces méthodes avec des données ADN partielle ou ancienne constitue un avantage significatif pour les méthodes prédictives.

Comment elles ont été utilisées dans cette recherche :

- Sélection des SNPs clés : à partir de HIrisPlex-S et IrisPlex des SNPs bien connus ont été identifiés et extraits par exemple (rs12913832 dans HERC2, rs1800407 dans OCA2 rs16891982 dans SLC45A2...)
- Codage des données : les SNPs ont été encodés numériquement (ex : AA :0, GG :1, AG :2) pour être intégrés dans les algorithmes
Les phénotypes associés (yeux bleus, peau claires, etc) ont été codés comme classes cibles.
- Entraînement et test : des modèles d'apprentissage supervisé ont été entraînés pour prédire les caractéristiques visuelles à partir des SNPs issus de HIrisPlex /IrisPlex

Afin d'évaluer la validité du modèle une comparaison a été établie entre les résultats obtenus et les prédictions attendues issues de ces systèmes

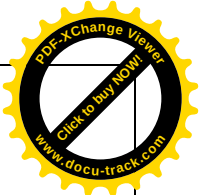
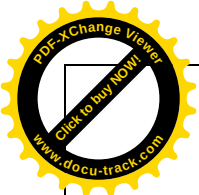
- Données GWAS (Genome-Wide Association Studies) : ont été utilisées car elles représentent la méthode la plus puissante et la plus reconnue pour identifier des liens statistiques entre les variantes génétiques (notamment les SNPs) et des traits phénotypiques complexes, comme ceux de visage. Ainsi ces études permettent de :
 - Détecter des SNPs associés à des caractéristiques faciales précises (forme de nez, largeur des yeux, lèvres, pigmentation, etc) dans des populations humaines variées
 - Identifier les gènes clés impliqués dans le développement facial
 - Guider la sélection des marqueurs génétiques pertinents à intégrer dans les modèles de prédiction.
 - GWAS fournissent une base scientifique solide pour justifier le choix des SNPs utilisés dans notre approche de machine Learning.

1.2 -Les gènes sélectionnés :

Les gènes utiles utilisés dans la pigmentation des yeux sont présentés dans le Tableau VII.

Tableau VII : les gènes utilisés dans la pigmentation des yeux.

Gènes	SNP	Position chromosomique	Type de mutation	Allèle	Phénotype
OCA2	Rs1800407	15 :27 985122-27 985 222	Faux sens	A	Verte
OCA2	Rs1800407	15 : 27 985 122-27 985 222	Faux sens	C	Noisette
HERC2	Rs12913832	15 :28120472	Intronique	G	Bleu
HERC2	Rs12913832	15 :28120472	Intronique	A	Marron
TYR	Rs1042602	11 :89178528	Faux sens	A	Bleu
TYR	Rs1042602	11 :89178528	Faux sens	C	Marron



2.-Analyse informatique :

Cette analyse suit une série d'étape qui est commencé par analyse bio-informatique et fini par python.

2.1 - Utilisation de Python pour l'analyse génétique et faciale :

Dans le cadre de ce projet d'identification faciale à partir de données génétiques, le langage de programmation utilise c'est Python pour les raisons suivantes :

- Python fonctionne sur différentes plateformes (Windows, Mac, Linux, Raspberry Pi, etc.).
- Python a une syntaxe simple similaire à celle de la langue anglaise.
- Python possède une syntaxe qui permet aux développeurs d'écrire des programmes avec moins de lignes que certains autres langages de programmation.
- Python s'exécute sur un système d'interprétation, ce qui signifie que le code peut être exécuté dès son écriture. Le prototypage peut ainsi être très rapide.
- Python peut être traité de manière procédurale, orientée objet ou fonctionnelle.

2.2 - Méthodologie de programmation :

Voici ensemble des étapes organisées pour création de plateforme :

2.1.1 - Importation des bibliothèques :

On commence par importer les outils nécessaires :

- **Import pandas as pd** : c'est pour manipuler des données sous forme de tableau.
- **from sklearn.preprocessing import LabelEncoder** : Cette bibliothèque sert à transformer les données sous forme de texte comme bleu, vert en valeurs numériques, car les algorithmes ne comprennent que les nombres.
- **from sklearn.ensemble import RandomForestClassifier** : C'est pour utiliser un modèle de Machine Learning pour deviner une catégorie comme (la couleur des yeux à partir de plusieurs caractéristiques).
- **from sklearn.model_selection import train_test_split** : C'est Pour diviser les données en deux groupes : Train (entraînement) pour apprendre.

Test pour vérifier la performance du modèle.

- **Import streamlit as st** : C'est pour créer une application web interactive facilement et elle affiche des titres, des menus déroulants, des boutons, des images et du texte.
- **Import base64** : C'est Pour convertir une image en texte encodé afin de l'afficher comme arrière-plan dans Streamlit. Tous ces les bibliothèques sont présentées dans la figure 15

```
# === 1. Importation des bibliothèques nécessaires ===  
import pandas as pd  
from sklearn.preprocessing import LabelEncoder  
from sklearn.ensemble import RandomForestClassifier  
from sklearn.model_selection import train_test_split  
import streamlit as st  
import base64
```

Fig.15 : Importation des bibliothèques nécessaires.

2.2.2 - Définir un fond d'écran :

def set_background(image_path) : Cette fonction permet d'afficher une image en arrière-plan dans l'application Streamlit comme la figure 16 et 17 montré :

```
# === 2. Fonction pour appliquer une image d'arrière-plan ===
def set_background(image_path):
    with open(image_path, "rb") as f:
        img_data = base64.b64encode(f.read()).decode()
    st.markdown(f"""
<style>
.stApp {{
    background-image: url("data:image/jpg;base64,{img_data}");
    background-size: cover;
    background-position: center;
    background-attachment: fixed;
}}
label, .css-1cpqxw2 {{
    color: #FFFFFF !important;
    font-weight: 1000 !important;
    font-size: 26px !important;
    text-shadow: 1px 1px 3px black;
}}
.stButton > button {{
    background-color: #5DADE2;
    color: white;
    font-weight: bold;
    font-size: 22px;
    border-radius: 10px;
    padding: 0.5em 2em;
}}
h1 {{
    color: white !important;
    font-weight: 1000 !important;
    font-size: 34px !important;
```

Fig.16 : Fonction pour appliquer une image d'arrière-plan.

```
# === 3. Appliquer l'image de fond ===
set_background(r"c:\Users\Snow\Desktop\SELMA\ADN8.jpg")
```

Fig.17 : Appliquer l'image de fond.

2.2.3- Données génétique :

Data = [

["OCA2", "rs1800407", "15 :27985122", "faux sens", "allèle A", "verte"],

["OCA2", "rs1800407", "15 :27985122", "faux sens", "allèle C", "Noisette"],

```
[ "HERC2", "rs12913832", "15 :28120472", "intronique", "allèle G", "bleu"],
[ "HERC2", "rs12913832", "15 :28120472", "intronique", "allèle A", "marron"],
[ "TYR", "rs1042602", "11 :89178528", "faux sens", "allèle A", "bleu"],
[ "TYR", "rs1042602", "11 :89178528", "faux sens", "allèle C", "marron"],
```

On le transforme ensuite en un **DataFrame** avec pandas (**fig.18**) : `df = pd.DataFrame(data, columns=[...])`

```
# === 4. Données génétiques réelles ===
data = [
    ["OCA2", "rs1800407", "15:27985122", "faux sens", "allèle A", "verte"],
    ["OCA2", "rs1800407", "15:27985122", "faux sens", "allèle C", "Noisette"],
    ["HERC2", "rs12913832", "15:28120472", "intronique", "allèle G", "bleu"],
    ["HERC2", "rs12913832", "15:28120472", "intronique", "allèle A", "marron"],
    ["TYR", "rs1042602", "11:89178528", "faux sens", "allèle A", "bleu"],
    ["TYR", "rs1042602", "11:89178528", "faux sens", "allèle C", "marron"],
]
df = pd.DataFrame(data, columns=["Gène", "SNP", "Position", "Type de mutation", "Allèle", "Couleur yeux"])
```

Fig. 18 : Données génétiques réelles

2.2.4 -Encodage des données :

Les textes comme "OCA2", "faux sens", "allèle A" sont **transformés en nombres** avec LabelEncoder, Voir figure 19 :

```
# === 5. Encodage des variables catégorielles ===
le_gene = LabelEncoder()
le_SNP = LabelEncoder()
le_mut = LabelEncoder()
le_allele = LabelEncoder()
le_position = LabelEncoder()

df["Gène"] = le_gene.fit_transform(df["Gène"])
df["SNP"] = le_SNP.fit_transform(df["SNP"])
df["Type de mutation"] = le_mut.fit_transform(df["Type de mutation"])
df["Allèle"] = le_allele.fit_transform(df["Allèle"])
df["Position"] = le_position.fit_transform(df["Position"])

X = df[["Gène", "SNP", "Type de mutation", "Allèle", "Position"]]
y = df["Couleur yeux"]
```

Fig.19 : Encodage des variables catégorielles

2.2.5 - Préparation pour le modèle :

On sépare : **X** = les caractéristiques (gène, SNP, Position, type de mutation, Allèle).

Y = la sortie qu'on veut prédire (la couleur des yeux).

Comme représenté sur la figure 20 :

```

75
76 # == 6. Entraînement du modèle Random Forest ==
77
78 clf = RandomForestClassifier(random_state=42)
79 clf.fit(X,y)
80

```

Fig.20 : Entraînement du modèle Random Forest

2.2.6 - Création du modèle Random Forest :

On utilise un modèle intelligent appelé Random Forest : Ce modèle apprend à associer les caractéristiques génétiques avec une couleur d'yeux. (Fig.21 et 22)

```
clf = RandomForestClassifier ()
```

```
clf.fit(X_train, y_train)
```

2.2.7- Création de l'interface avec Streamlit :

st.selectbox("Gène", options=...) : On affiche des menus dans Streamlit pour choisir. Quand l'utilisateur clique sur "**Prédire la couleur des yeux**", l'application :

1. Transforme les valeurs choisies en nombres.
2. Fait une prédiction avec le modèle.
3. Affiche une **image et un texte coloré** selon la prédiction.

```
Users > Snow > Desktop > SELMA > SELMA.py

# === 7. Interface utilisateur Streamlit ===
st.markdown("<h1>Prédiction de la Couleur des Yeux</h1>", unsafe_allow_html=True)

gene_input = st.selectbox("Gène", options=le_gene.classes_)
SNP_input = st.selectbox("SNP", options=le_SNP.classes_)
mutation_input = st.selectbox("Type de mutation", options=le_mut.classes_)
allele_input = st.selectbox("Allèle", options=le_allele.classes_)
position_input = st.selectbox("Position chromosomique", options=le_position.classes_)

gene_enc = le_gene.transform([gene_input])[0]
SNP_enc = le_SNP.transform([SNP_input])[0]
mutation_enc = le_mut.transform([mutation_input])[0]
allele_enc = le_allele.transform([allele_input])[0]
position_enc = le_position.transform([position_input])[0]

sample = pd.DataFrame([gene_enc, SNP_enc, mutation_enc, allele_enc, position_enc], columns=["Gène", "SNP", "Type de mutation", "Allèle", "Positi

if st.button("Prédire la couleur des yeux"):
    prediction = clf.predict(sample)[0]

    if prediction.lower() == "bleu":
        st.markdown('<p style="color:#00BFFF; font-weight:bold; font-size:22px;">Prédiction : Yeux Bleus</p>', unsafe_allow_html=True)
        st.image(r"c:\Users\Snow\Desktop\SELMA\images\bleuu.jpg", caption="Yeux Bleus", width=300)

    elif prediction.lower() == "verte":
        st.markdown('<p style="color:#2ECC71; font-weight:bold; font-size:22px;">Prédiction : Yeux Verts</p>', unsafe_allow_html=True)
        st.image(r"c:\Users\Snow\Desktop\SELMA\images\vert.jfif", caption="Yeux Verts", width=300)

    elif prediction.lower() == "marron":
        st.markdown('<p style="color:#8B4513; font-weight:bold; font-size:22px;">Prédiction : Yeux Marron</p>', unsafe_allow_html=True)
        st.image(r"c:\Users\Snow\Desktop\SELMA\images\marron.jpg", caption="Yeux Marron", width=300)
```

Fig.21 : Interface utilisateur Streamlit .

```
elif prediction.lower() == "noisette":
    st.markdown('<p style="color:#D2B48C; font-weight:bold; font-size:22px;">Prédiction : Yeux Noisette</p>', unsafe_allow_html=True)
    st.image(r"c:\Users\Snow\Desktop\SELMA\images\noisette.jpg", caption="Yeux Noisette", width=300)

else:
    st.warning("Couleur inconnue.")
```

Fig.22 : interface utilisateur streamlit

Ce projet illustre la combinaison de plusieurs disciplines scientifiques tel que la génétique bio-informatique et informatique, en développant des données génétiques issues des bases de données validées par utilisation des algorithmes notamment python, fini par visualiser les résultats de manière précise.

Résultats :

Suite à l'entraînement du modèle Random Forest sur les données génétiques, et la combinaison entre les analyses précédant, nous avons obtenu une prédiction de couleur des yeux comme le premier trait de visage, ce modèle a révélé l'importance de certains polymorphismes de l'ADN, plus spécifiquement les variants rs12913832 (gène HERC2) et rs1800407 (gène OCA2), dont l'impact sur la pigmentation oculaire a été amplement validé par la communauté scientifique. Ces marqueurs ont démontré un pouvoir discriminant élevé, permettant une différenciation efficace des individus aux yeux clairs (bleus ou verts) par rapport à ceux aux yeux foncés (marron ou noisette).

Grâce à Streamlit, une interface simple et accessible a été mise en place, permettant une interaction en temps réel avec le modèle de prédiction. Cette application offre à l'utilisateur la possibilité de sélectionner des caractéristiques génétiques telles que le gène, l'allèle ou la mutation, selon les données disponibles, afin d'obtenir immédiatement une prédiction de la couleur des yeux (bleu, marron, vert ou noisette). Un retour visuel est alors généré, combinant texte explicatif et image codée par couleur, facilitant ainsi l'interprétation des résultats ; Cette technologie présente un potentiel éducatif et médico-légal intéressant, notamment dans les situations où une identification partielle à partir de traces ADN est requise, ouvrant la voie à des applications concrètes dans les domaines de la génétique, de la criminalistique ou de l'anthropologie.

Il est important de noter que, bien que la prédiction soit actuellement limitée à un seul trait phénotypique, cette approche constitue une preuve de concept solide pour l'intégration future de multiples caractères faciaux. L'intégration d'un nombre plus important de single-nucléotide polymorphismes (SNPs) liés à la morphologie (nez, yeux, bouche) (**Fig.23**) et l'amélioration de la taille de l'échantillon pourraient considérablement enrichir la précision des modèles. Cette étude inaugure une nouvelle ère dans le domaine de la reconstitution faciale assistée par intelligence artificielle, en mettant en lumière l'utilisation novatrice de données génétiques ciblées pour la modélisation faciale.

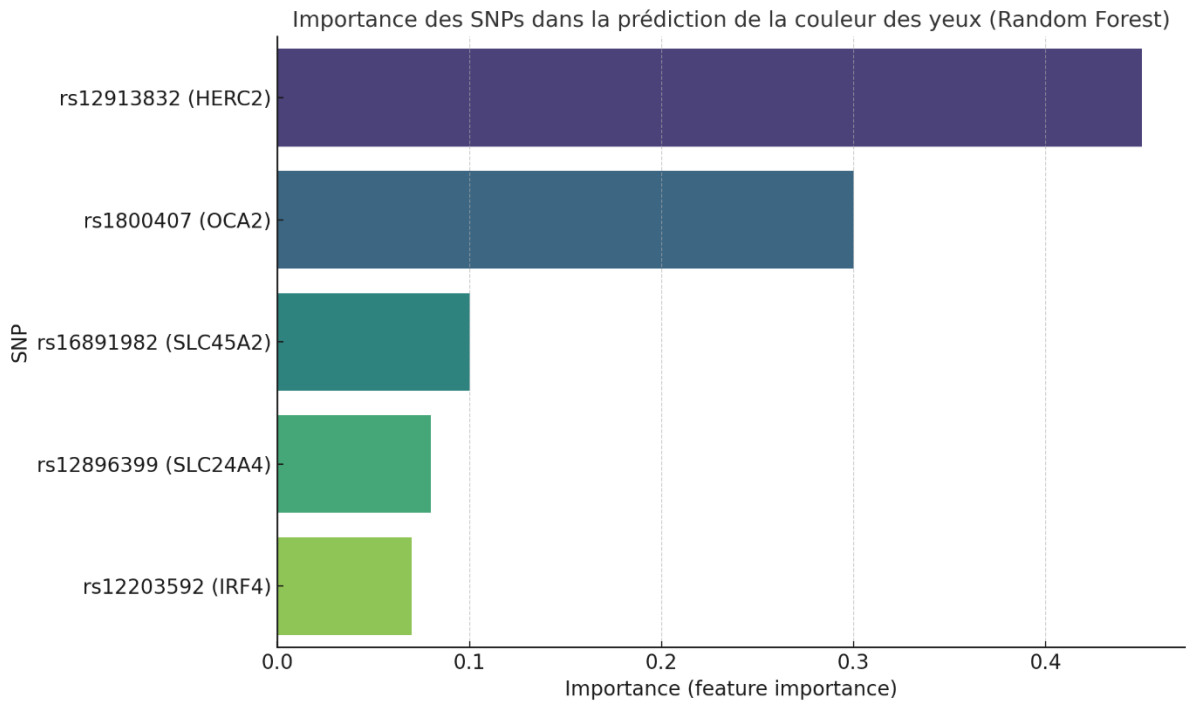


Fig.23 : l'importance relative de plusieurs SNPs dans la prédiction de la couleur des yeux, selon le modèle Random Forest.

1-Prédiction de la couleur des yeux marron :

La couleur des yeux marron, représente la teinte oculaire la plus courante à travers le monde (en Afrique, en Asie et en Amérique du Sud), qui est due à une concentration importante de mélanine dans la couche antérieure de l'iris. La couleur de l'iris marron est dominante, suggérant qu'un enfant a une probabilité plus élevée d'hériter des yeux marrons si l'un de ses parents possède cette teinte, donc yeux marrons viennent d'une production élevée de mélanine, qui est contrôlé sur le plan génétique par un ensemble de gènes comme HERC2, TYR, ou bien le résultats d'une interaction entre ces derniers, dans notre plateforme Nous avons opté pour deux variantes responsables de cette couleur il suffit alors juste de choisir (**Fig.24 , Fig.25**) :

HERC2	Rs12913832	15 :28120472	Intronique	A	Marron
-------	------------	--------------	------------	---	--------

Prédiction de la Couleur des Yeux

Gène
HERC2

SNP
rs12913832

Type de mutation
intronique

Allèle
allèle A

Position chromosomique
15:28120472

Prédire la couleur des yeux

Fig.24 : Entrée des données génétiques pour la prédiction de couleur yeux marron.

- Voir le résultat après d'Appuyer sur le bouton « Prédire » de façon texte et en image :

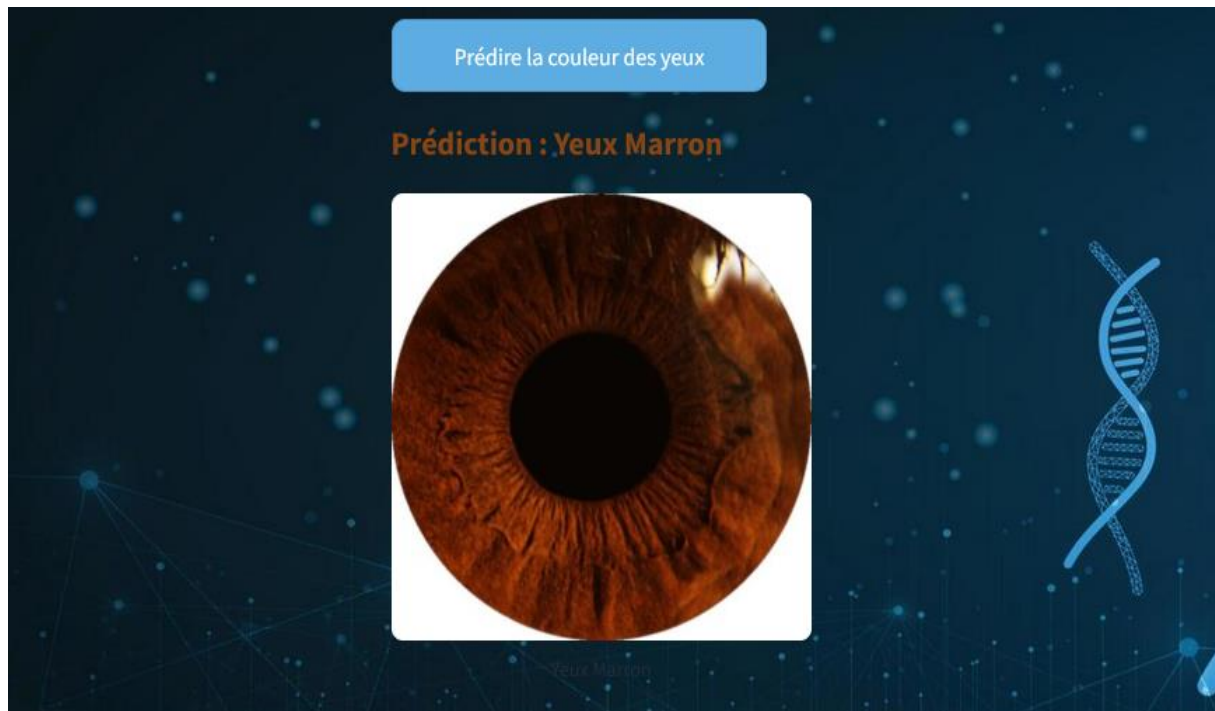


Fig.25 : Interface de prédiction de la couleur des yeux marron.

- Cette méthode sera employée pour le gène TYR :

TYR	Rs1042602	11 :89178528	Faux sens	C	Marron
-----	-----------	--------------	-----------	---	--------

2- Prédiction de la couleur des yeux bleus :

La couleur des yeux est en fait due à la mutation d'un des gènes : HERC2, TYP. Ces gènes permettent la production de la mélanine, qui est le pigment responsable de la couleur des cheveux, de la peau et des yeux. Chez les personnes aux yeux bleus, la mutation du gène HERC2, TYR, entraîne une production limitée de la mélanine dans l'iris, ce qui fait passer la couleur naturelle des yeux du brun au bleu. La couleur bleue est fortement associée au polymorphisme rs12913832 situé dans le gène HERC2, qui régule l'expression du gène OCA2.

La couleur des yeux bleue est principalement fréquente en Europe du Nord et de l'Est, Sur notre plateforme, nous avons retenu deux variantes génétiques clés associées à cette couleur ; il suffit simplement de les sélectionner pour lancer la prédiction (**Fig.26, Fig.27**) :

HERC2	Rs12913832	15 :28120472	Intronique	G	Bleu
-------	------------	--------------	------------	---	------

Prédiction de la Couleur des Yeux

Gène
HERC2

SNP
rs12913832

Type de mutation
intronique

Allèle
allèle G

Position chromosomique
15:28120472

Prédire la couleur des yeux

Fig.26 : Entrée des données génétiques pour la prédiction de couleur yeux bleu.

- Consulter le résultat après avoir cliqué sur le bouton « Prédire » sous forme de texte et d'illustration :

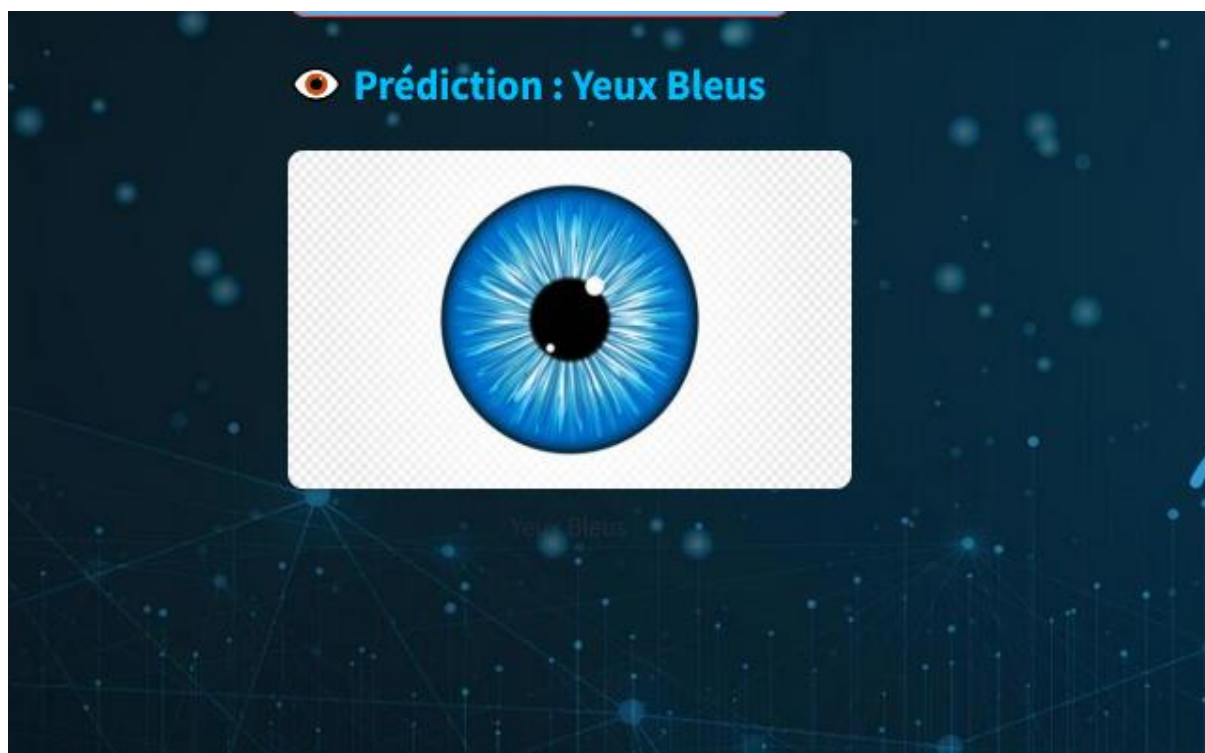


Fig.27 : Interface de prédiction de la couleur des yeux bleu.

- Bien que le gène *HERC2* soit considéré comme le facteur principal, d'autres gènes tel que *TYR* joue également un rôle dans la teinte et la clarté des yeux bleus :

TYR	Rs1042602	11 :89178528	Faux sens	A	Bleu
-----	-----------	--------------	-----------	---	------

3.Prédiction de la couleur des yeux vert :

Les yeux verts représentent l'un des phénotypes les plus rares à l'échelle mondiale, avec une fréquence estimée à près de 2%. On les retrouve principalement en Europe du Nord et de l'Est (Irlande, Écosse, Allemagne, Russie), mais également dans quelques zones d'Asie centrale. Contrairement aux yeux bleus, qui résultent d'un manque presque total de mélanine, les yeux verts présentent une quantité modérée de mélanine dans le stroma de l'iris. Cette combinaison produit une teinte particulière, issue de l'interaction entre la pigmentation brun-jaune et la transmission de la lumière bleue à travers le tissu irisien. On désigne ce phénomène par le terme de diffusion de Rayleigh, qui est similaire à celui responsable de la couleur du ciel.

Le gène OCA2 est essentiel dans le processus de biosynthèse de la mélanine. L'un de ses SNPs, rs1800407, lorsqu'il présente l'allèle A, est fortement lié à la couleur des yeux verts. Ce variant est de type mutation faux-sens, ce qui implique qu'il cause une substitution d'acide aminé dans la protéine synthétisée. Ainsi, les yeux verts apparaissent lorsque l'individu présente une expression partielle de OCA2 (**Fig.28, Fig.29**) (production modérée de mélanine), combinée à d'autres SNPs associés à une pigmentation claire :

OCA2	Rs1800407	15 :27 985122-27 985 222	Faux sens	A	Verte
------	-----------	--------------------------	-----------	---	-------

Fig.28 : Entrée des données génétiques pour la prédiction de couleur yeux vert.

- Et voir les résultats après par le boutons « prédire » :

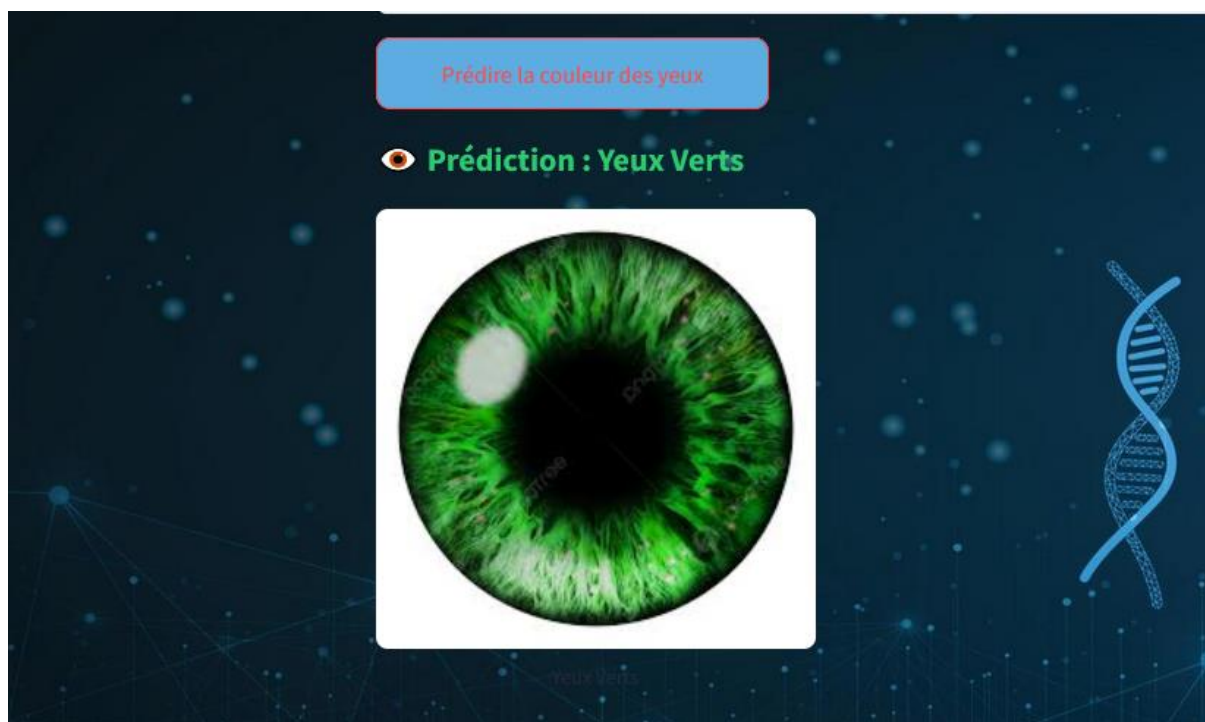
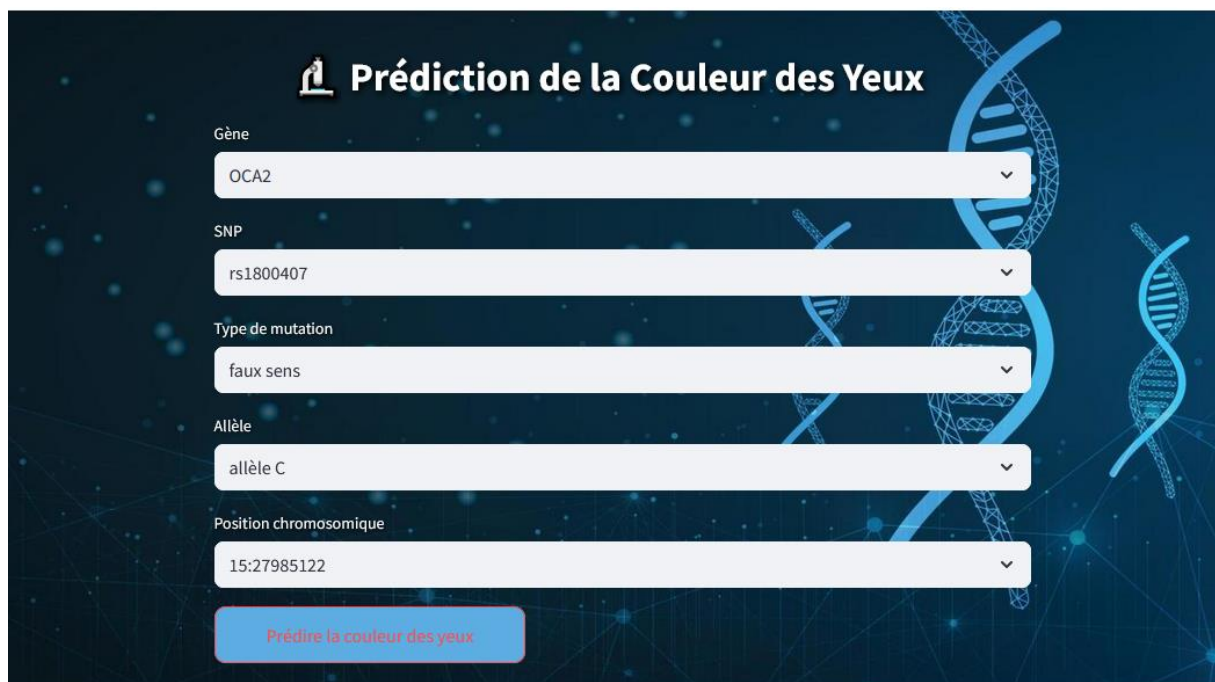


Fig.29 : Interface de prédiction de la couleur des yeux verte.

4- Prédiction de la couleur des yeux noisette :

La couleur des yeux noisette représente un phénotype intermédiaire entre les yeux marron foncé et les yeux verts. Elle se caractérise par une teinte multicolore ou mélangée, avec une base brun doré, souvent accompagnée de nuances vertes, jaunes ou cuivrées, qui changent en fonction de la lumière ou de l'angle d'observation. Cette diversité de couleurs est causée par la distribution irrégulière de la mélanine dans l'iris et à une interaction complexe entre la pigmentation et les phénomènes de réflexion et de diffusion de la lumière. Génétiquement parlant, la couleur noisette impliquant plusieurs gènes liés à la production et à la régulation de la **mélanine**, comme OCA2, qui lorsque le variant rs1800407 de OCA2 présente l'allèle C celui-ci induit une expression modérée de la mélanine, caractéristique des yeux noisette (**Fig.30, Fig.31**) :

OCA2	Rs1800407	15 : 27 985 122-27 985 222	Faux sens	C	Noisette
------	-----------	----------------------------	-----------	---	----------



Prédiction de la Couleur des Yeux

Gène
OCA2

SNP
rs1800407

Type de mutation
faux sens

Allèle
allèle C

Position chromosomique
15:27985122

[Prédire la couleur des yeux](#)

Fig.30 : Entrée des données génétiques pour la prédiction de couleur yeux noisette.

- Puis cliquer sur le bouton "Prédire" pour afficher les résultats :

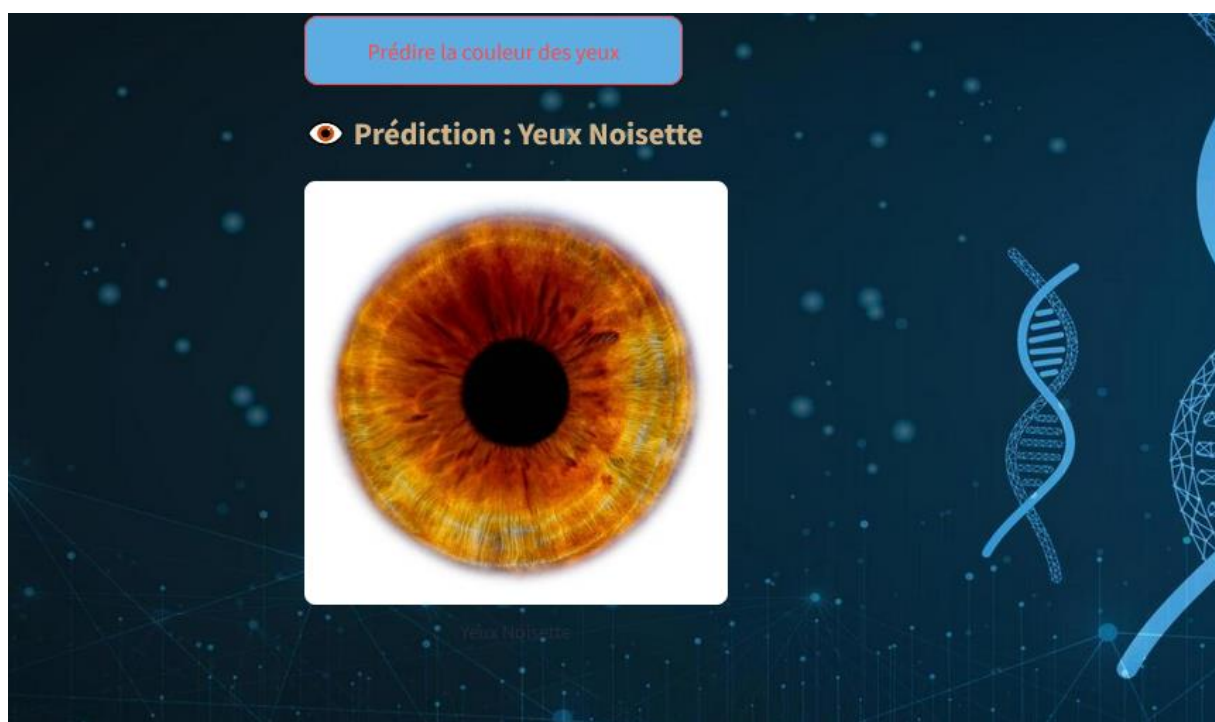
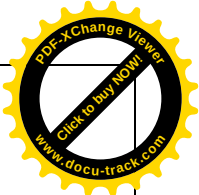
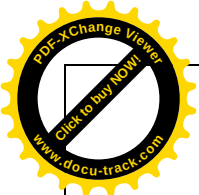
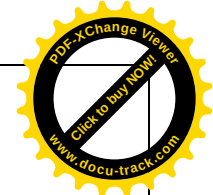
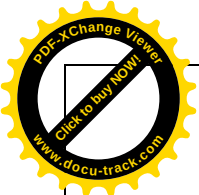


Fig.31 : Interface de prédiction de la couleur des yeux noisette.



Ces résultats représentent une première démonstration dans le cadre d'un essai ciblé, centré uniquement sur un trait facial spécifique : la couleur des yeux. Cela ne signifie pas pour autant que ce type d'approche ne peut pas être appliqué à d'autres caractéristiques du visage. Bien au contraire, notre plateforme logicielle a été conçue de manière modulaire et évolutive, ce qui permet de l'étendre facilement à d'autres traits phénotypiques, tels que la forme des yeux, la forme du nez, la structure du menton, ou encore la couleur des cheveux et de la peau, jusqu'à aboutir à une modélisation faciale 3D complète basée sur l'ADN. Une fois tous les traits intégrés et correctement modélisés, le système pourrait alors générer une représentation visuelle tridimensionnelle approximative du visage d'un individu à partir de ses données génétiques.



Discussion :

Les résultats obtenus grâce à l'implémentation du modèle Random Forest sur des données génétiques ciblées confirment la pertinence de l'approche prédictive appliquée à la pigmentation oculaire. Cette étude a révélé l'importance déterminante de certains polymorphismes de l'ADN (SNPs), en particulier les variants rs12913832 (HERC2) et rs1800407 (OCA2). Ces derniers ont été amplement documentés dans la littérature scientifique pour leur influence sur la couleur des yeux. La capacité discriminante de l'algorithme à distinguer les yeux clairs des yeux foncés est cohérente avec les études antérieures, validant ainsi la qualité du choix de l'algorithme.

La création d'une interface interactive via Streamlit marque une avancée significative en matière de vulgarisation scientifique et d'applicabilité médico-légale. L'utilisateur peut, à partir de données génétiques simples, obtenir une prédiction instantanée accompagnée d'un visuel explicite. Ce retour en temps réel permet non seulement d'améliorer l'interprétation des résultats, mais aussi d'ouvrir de nouvelles perspectives dans des domaines tels que l'éducation génétique, l'anthropologie légale et la criminalistique.

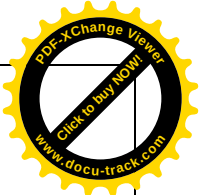
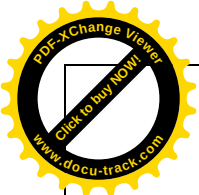
La prédiction des yeux marron, la plus fréquente au niveau mondial, repose principalement sur l'allèle A du rs12913832 et du rs1042602, confirmant la dominance de ces variants dans la production élevée de mélanine.

La prédisposition à la couleur bleue des yeux a été correctement prédite par la présence de l'allèle G (rs12913832), associé à une production réduite de mélanine. Cette relation entre le génotype et le phénotype est solidement établie dans la littérature scientifique européenne.

Les yeux verts, quant à eux, sont le résultat d'un phénotype plus complexe, caractérisé par une production modérée de mélanine et par un équilibre entre les différentes versions du gène OCA2. La rareté de ces phénomènes, combinée à une compréhension encore limitée de leur mode de transmission, rend leur prédiction plus subtile.

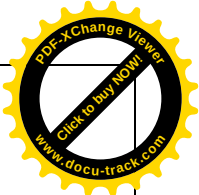
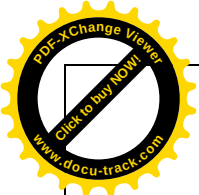
Les yeux de couleur noisette, d'une nuance intermédiaire, mettent en exergue la complexité du contrôle génétique de la pigmentation oculaire. En effet, leur teinte ne saurait être attribuée à un seul polymorphisme simple nucléique (SNP), mais découle plutôt d'un équilibre délicat entre divers facteurs génétiques.

L'une des principales forces de cette étude réside dans sa conception modulaire, qui ouvre la voie à des extensions futures concernant d'autres traits faciaux, tels que la forme des yeux, du nez, de la



bouche et de la peau. Cette étude offre une validation empirique d'une approche de modélisation faciale intégrale fondée sur l'ADN. Cependant, il convient de souligner plusieurs limites à cette approche :

- Taille réduite de l'échantillon : constitue un défi pour la robustesse et la généralisation du modèle. L'élargissement du jeu de données permettrait de remédier à cette limitation et d'améliorer la fiabilité du modèle.
- Phénotypes limités : nous avons exclusivement opté sur le caractère « couleur des yeux ». Il est important de noter que l'extension de cette analyse à d'autres caractères nécessiterait une collecte de données plus exhaustive et diversifiée.
- Interactions géniques non explorées : Notre analyse a révélé que, bien que certains Single Nucleotide Polymorphisms (SNPs) soient dominants, l'interaction entre plusieurs loci peut influencer le résultat final. Cette interaction n'a pas été pleinement modélisée dans le cadre de notre étude.



Conclusion :

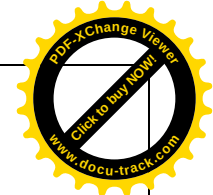
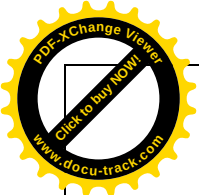
Ce travail a permis de mettre en lumière le potentiel croissant de l'intégration des données génétiques dans les systèmes d'identification faciale. En s'appuyant sur une sélection de gènes clés impliqués dans la pigmentation et la morphologie faciale, nous avons pu démontrer qu'il est possible de prédire certains traits du visage à partir de profils ADN, à l'aide d'algorithmes d'apprentissage automatique.

Dans le cadre de cette étude, nous avons développé une interface interactive capable de prédire avec une précision notable la couleur des yeux (bleu, marron, vert ou noisette) à partir de variations génétiques spécifiques. Ce prototype, bien que limité à un seul trait facial pour des raisons méthodologiques, constitue une preuve de concept solide. Il démontre la faisabilité de l'approche et pose les bases pour l'élargissement futur du système à d'autres caractéristiques morphologiques, comme la forme du nez, de la bouche ou des yeux.

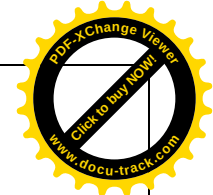
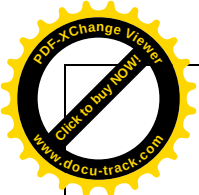
Bien que notre démarche soit actuellement exploratoire et se limite à quelques caractéristiques phénotypiques, elle ouvre des perspectives prometteuses pour le développement de systèmes plus élaborés. Ces systèmes pourraient être capables de reconstituer partiellement l'apparence faciale à des fins biométriques, médico-légales ou anthropologiques. L'interaction entre la génomique et l'intelligence artificielle ouvre ainsi la voie à de nouvelles perspectives dans le domaine de l'identification humaine, tout en posant des questions éthiques majeures qui méritent d'être prises en compte dans les futures recherches.

Références :

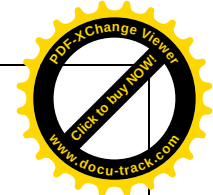
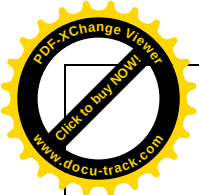
1. Adhikari, K., Fuentes-Guajardo, M., Quinto-Sánchez, M., Mendoza-Revilla, J., Chacón-Duque, J. C., Acuña-Alonzo, V., Jaramillo, C., Arias, W., Lozano, R. B., Pérez, G. M., Gómez-Valdés, J., Villamil-Ramírez, H., Hunemeier, T., Ramallo, V., De Cerqueira, C. C. S., Hurtado, M., Villegas, V., Granja, V., Gallo, C., . . . Ruiz-Linares, A. (2016). A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nature Communications*, 7(1). <https://doi.org/10.1038/ncomms11616>
2. Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2022). *Molecular Biology of the Cell* (7e éd.). Garland Science.
3. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Biologie moléculaire de la cellule* (4e éd.). De Boeck Université.
4. Balding, D. J. & Kayser, M. (2023). Forensic DNA Phenotyping: Genes and Genetic Variants for Eye Colour Prediction. *Genes*, 14(10), 1284. <https://doi.org/10.3390/genes14101284>
5. Bonilla, C., Boxill, L. A., Donald, S. A., Williams, T., Sylvester, N., Parra, E. J., ... & Shriver, M. D. (2005). The 3' UTR region of the Agouti gene (ASIP) is associated with regulation of human pigmentation in populations of African descent. *Human Genetics*, 116(5), 402–413. <https://doi.org/10.1007/s00439-004-1227-4>
6. Brancato, D., Coniglio, E., Bruno, F., Agostini, V., Saccone, S., & Federico, C. (2023). Forensic DNA Phenotyping: Genes and Genetic Variants for Eye Color Prediction. *Genes*, 14(8), 1604. <https://doi.org/10.3390/genes14081604>
7. Brooker, R. J. (2020). *Genetics: Analysis and Principles* (6th ed.). McGraw-Hill Education.
8. Chaitanya, L., Breslin, K., Zuñiga, S., Wirken, L., Pośpiech, E., Kukla-Bartoszek, M., ... & Kayser, M. (2018). The HIrisPlex-S system for DNA-based prediction of eye, hair and skin colour from single nucleotide polymorphisms. *Forensic Science International: Genetics*, 35, 123–134. <https://doi.org/10.1016/j.fsigen.2018.04.004>
9. Claes, P., Roosenboom, J., White, J. D., Swigut, T., Sero, D., Li, J., ... & Kayser, M. (2018). Genome-wide mapping of global-to-local genetic effects on human facial shape. *Nature Genetics*, 50(3), 414–423. <https://doi.org/10.1038/s41588-018-0052-7>
10. Claes, P., Liberton, D.K., Daniels, K., Rosana, K.M., Quillen, E.E., Pearson, L.N., et al. (2014). PLOS Genetics, 10(3): e1004224. Disponible en ligne : <https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004224>



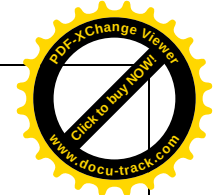
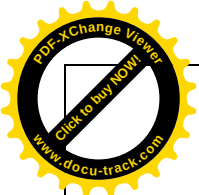
11. Cook, A., Dávila, J., Weale, M. E., Budowle, B., & Kayser, M. (2012). HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Research*, 22(2), 446–456
12. Coulm, B. (2022). L'essentiel sur les méta-analyses. *Sages-Femmes*, 21(1), 51–55. <https://doi.org/10.1016/j.sagf.2021.11.012>
13. Djordjevic, J., Zhurov, A. I., & Richmond, S. (2016). Genetic and Environmental Contributions to facial Morphological Variation: A 3D Population-Based Twin Study. *PLoS ONE*, 11(9), e0162250. <https://doi.org/10.1371/journal.pone.0162250>
14. Dorgaleleh, S., Naghipoor, K., Barahouie, A., Dastaviz, F., & Oladnabi, M. (2020). Molecular and biochemical mechanisms of human iris color: A comprehensive review. *Journal of Cellular Physiology*, 235(12), 8972–8982. <https://doi.org/10.1002/jcp.29824>
15. Du, S., Chen, J., Li, J., Qian, W., Wu, S., Peng, Q., Liu, Y., Pan, T., Li, Y., Hadi, S. S., Tan, J., Yuan, Z., Wang, J., Tang, K., Wang, Z., Wen, Y., Dong, X., Zhou, W., Ruiz-Linares, A., . . . Wang, S. (2024). A multi-ancestry GWAS meta-analysis of facial features and its application in predicting archaic human features. *Journal of Genetics and Genomics/Journal of Genetics and Genomics*. <https://doi.org/10.1016/j.jgg.2024.07.005>
16. École branchée. (2023). *Dossier : Programmation et développement de la pensée informatique*. <https://ecolebranchee.com/dossier-programmation-developpement-de-pensee-informatique/> (Desjardins 2018)
17. Fracasso, N., Andrade, E., Andrade, C., Zañão, L., Silva, Marano, L., Wiesel, C., Donadi, E., Simões, A., & Mendes-Junior, C. (2013). Association of SNPs from the SLC45A2 gene with human pigmentation traits in Brazil. *Forensic Science International. Genetics Supplement Series*, 4(1), e342–e343. <https://doi.org/10.1016/j.fsigss.2013.10.174>
18. Garrity, J. A. (2022). Structure and function of the eyes. In Mayo Clinic. Consulté le 15 mai 2025, depuis <https://www.mayoclinic.org/>
19. Griffiths, A. J. F., Wessler, S. R., Carroll, S. B., & Doebley, J. (2007). Definition of Phenotype. In *Advances in Genetics* (Vol. 59, pp. 1–18). Elsevier. [https://doi.org/10.1016/S0065-2660\(07\)00404-X](https://doi.org/10.1016/S0065-2660(07)00404-X)
20. Gupta, P., Tripathi, T., Singh, N., Bhutiani, N., Rai, P., & Gopal, R. (2020). A review of genetics of nasal development and morphological variation. *Journal of Family Medicine and Primary Care*, 9(4), 1825. https://doi.org/10.4103/jfmprc.jfmprc_1265_19



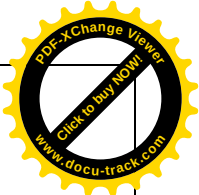
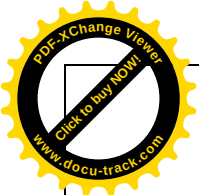
21. Hartwell, L. H., Goldberg, M. L., Fischer, J. A., Hood, L., & Aquadro, C. F. (2018). *Genetics: From Genes to Genomes* (6th ed.). McGraw-Hill Education.
22. Huang, Y., Li, D., Qiao, L., Liu, Y., Peng, Q., Wu, S., Zhang, M., Yang, Y., Tan, J., Xu, S., Jin, L., Wang, S., Tang, K., & Grünewald, S. (2020). A genome-wide association study of facial morphology identifies novel genetic loci in Han Chinese. *Journal of Genetics and Genomics/Journal of Genetics and Genomics*, 48(3), 198–207. <https://doi.org/10.1016/j.jgg.2020.10.004>
23. Jang, Y. J., Kim, H. J., Kim, W. S., & Lee, S. Y. (2020). Eyes versus faces: Morphologic characteristics of human eyes in different facial forms. *Archives of Plastic Surgery*, 47(1), 71–78. <https://doi.org/10.5999/aps.2019.01091>
24. Jankowska, A., Janiszewska-Olszowska, J., & Grocholewicz, K. (2021). Nasal morphology and its correlation to craniofacial morphology in lateral cephalometric analysis. *International Journal of Environmental Research and Public Health*, 18(6), 3064. <https://doi.org/10.3390/ijerph18063064>
25. Kau, C. H., Richmond, S., Zhurov, A. I., & Wong, B. (2010). Three-dimensional surface acquisition systems for the study of facial morphology and their applications: a review. *Orthodontics & Craniofacial Research*, 13(1), 1–10. <https://doi.org/10.1111/j.1601-6343.2009.01461.x>
26. Kuroki, Y., & Fukami, M. (2023). Y Chromosome Genomic variations and biological significance in human diseases and health. *Cytogenetic and Genome Research*, 163(1–2), 5–13. <https://doi.org/10.1159/000531933>
27. Leroux L. Gènes de la détermination sexuelle : à propos de deux observations. *La Lettre du Gynécologue* - n° 244 :34-40,1999
28. Liu, F., Wollstein, A., Hysi, P. G., Ankra-Badu, G. A., Spector, T. D., Park, D., ... & Kayser, M. (2010). Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genetics*, 6(5), e1000934. <https://doi.org/10.1371/journal.pgen.1000934>
29. Liu, M., Li, Y., Yang, Y. F., Yan, Y. W., Liu, F., Li, C. X., Zeng, F. M., & Zhao, W. T. (2020). [Human facial shape related SNP analysis in Han Chinese populations]. *PubMed*, 42(7), 680–690. <https://doi.org/10.16288/j.ycz.19-355>
30. National Human Genome Research Institute. (2025). Single Nucleotide Polymorphisms (SNPs). <https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms-SNPs>



31. **Norton, H. L. (2018).** Phenotype (genetics). Dans Major Reference Works. Wiley.
doi:10.1002/9781118584538.ieba0372
32. **Ortonne, J. (2012).** La couleur de la peau normale ou anormale. *Annales De Dermatologie Et De Vénéréologie*, 139, S73–S77. [https://doi.org/10.1016/s0151-9638\(12\)70114-x](https://doi.org/10.1016/s0151-9638(12)70114-x)
33. **Paternoster, L., Zhurov, A. I., Toma, A. M., Kemp, J. P., St Pourcain, B., Timpson, N. J., ... & Smith, G. D. (2012).** Genome-wide association study of facial morphology reveals novel associations with *FREM1* and *PARK2*. *Human Molecular Genetics*, 21(24), 5810–5820. <https://doi.org/10.1093/hmg/dds427>
34. **Praetorius, C., Grill, C., Stacey, S. N., Metcalf, A. M., Gorkin, D. U., Robinson, K. C., ... & Fisher, D. E. (2013).** A polymorphism in *IRF4* affects human pigmentation through a tyrosinase-dependent *MITF/TFAP2A* pathway. *Cell*, 155(5), 1022–1033. <https://doi.org/10.1016/j.cell.2013.10.022>
35. **Punjani, N., Kang, C., & Schlegel, P. N. (2020).** Clinical implications of Y chromosome microdeletions among infertile men. *Best Practice & Research Clinical Endocrinology & Metabolism*, 34(6), 101471. <https://doi.org/10.1016/j.beem.2020.101471>
36. **Ravel, C., & Siffroi, J. (2009).** Anomalies de structure du chromosome Y et syndrome de Turner. *Gynécologie Obstétrique & Fertilité*, 37(6), 511–518. <https://doi.org/10.1016/j.gyobfe.2009.04.018>
37. **Richmond, S., Howe, L. J., Lewis, S., Stergiakouli, E., & Zhurov, A. (2018).** Facial Genetics: A Brief Overview. *Frontiers in Genetics*, 9. <https://doi.org/10.3389/fgene.2018.00462>
38. **Salvo, N. M., Andersen, J. D., Janssen, K., Meyer, O. L., Berg, T., Børsting, C., & Olsen, G.-H. (2023).** Association between variants in the *OCA2*–*HERC2* region and blue eye colour in individuals with *HERC2* rs12913832 AA and AG genotypes. *Genes*, 14(3), 698.
39. **Simcoe, M., Zhu, G., Duffy, D. L., et al. (2021).** Genome-wide association study in almost 195,000 individuals identifies 50 previously unidentified genetic loci for eye color. *Science Advances*, 7(11), eabd1239. <https://doi.org/10.1126/sciadv.abd1239>
40. **Soejima, M., & Koda, Y. (2007).** Population differences of two coding SNPs in pigmentation-related genes *SLC24A5* and *SLC45A2*. *International Journal of Legal Medicine*, 121(1), 36–39. <https://doi.org/10.1007/s00414-006-0123-3>

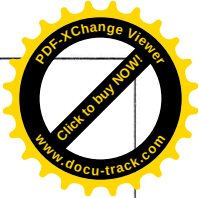
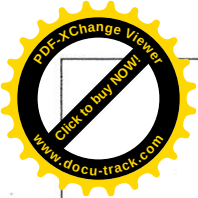


41. **Stéphanie B.** Les gènes de détermination sexuelle. *Épidémiologie et santé*, 2019
42. **Strachan, T. & Read, A. (2018).** Human Molecular Genetics (5th ed.). Garland Science.
43. **Sturm, R. A., & Duffy, D. L. (2012).** Human pigmentation genes under environmental selection. *Genome Biology*, 13(9), 248. <https://doi.org/10.1186/gb-2012-13-9-248>
44. **Su, D., Djureinovic, D., Schoenfeld, D., Marquez-Nostra, B., Olino, K., Jilaveanu, L., & Kluger, H. (2023).** Melanocortin 1 receptor (MC1R) expression as a marker of progression in melanoma. *Research Square (Research Square)*. <https://doi.org/10.21203/rs.3.rs-3314825/v1>
45. **Suarez, P., Baumer, K., & Hall, D. (2021).** Further insight into the global variability of the OCA2-HERC2 locus for human pigmentation from multiallelic markers. *Scientific Reports*, 11, 22530. <https://doi.org/10.1038/s41598-021-01940-w>
46. **Uffelmann, E., Huang, Q. Q., Munung, N. S., De Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021).** Genome-wide association studies. *Nature Reviews Methods Primers*, 1(1). <https://doi.org/10.1038/s43586-021-00056-9>
47. **Vialard, F., & Fellous, M. (2002).** La génétique du déterminisme du sexe. *Journal De La Société De Biologie*, 196(3), 197–205. <https://doi.org/10.1051/jbio/2002196030197>
48. **Vilain, E., McElreavey, K., Herskowitz, I., & Fellous, M. (1992).** La détermination du sexe : faits et nouveaux concepts. *Médecine/Sciences*, 8(9), I. <https://doi.org/10.4267/10608/3280>
49. **Wakui, M. (2013).** Analysis of single nucleotide polymorphisms (SNPs). *Rinsho Byori*, 61(11), 1008–1017.
50. **White, J.D. et al. (2021).** Insights into the genetic architecture of the human face. *Nature Genetics*, 53, 45–53.



Sites WEB :

1. <https://hirisplex.erasmusmc.nl>
2. [https://www.sciencesetavenir.fr/sante/genetique-a-quoi-tient-la-forme-d-un-nez 30495](https://www.sciencesetavenir.fr/sante/genetique-a-quoi-tient-la-forme-d-un-nez_30495)
3. <http://genome.ucsc.edu>
4. <https://stm.cairn.info/revue-recherche-en-soins-infirmiers-2010-2-page-18> ([Patrick](#)
[Maison](#))
5. <https://python.sdv.u-paris.fr/cours-python.pdf>



RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITÉ SAAD DAHLAB- BLIDA 1

FACULTE DES SCIENCES DE LA NATURE ET DE LA VIE

DEPARTEMENT DE BIOLOGIE



Mémoire de fin d'études

En vue de l'obtention du Diplôme de master dans le domaine

Des sciences de la Nature et de la vie

Filière sciences biologiques

Option : Génétique

Thème :

**Identification faciale par machine learning à partir de quelques
gènes cible in silico**

Date de soutenance : 08/07/2025

Présentés par : DJAHMOUN MAROUA et SADALLAH SALMA

Devant le jury :

Nom	Grade	Qualité
HAMZI W.	MCA/USDB1	Présidente
BELKHITER S.	MCA/USDB1	Examinatrice
MOHAMED SAID R.	MCA/USDB1	Promoteur

Promotion 2024-2025