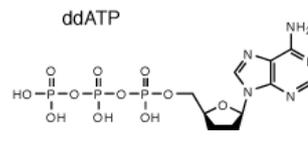
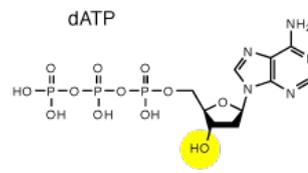


Partie 0 : Techniques et concepts/ à lire avant le TD.

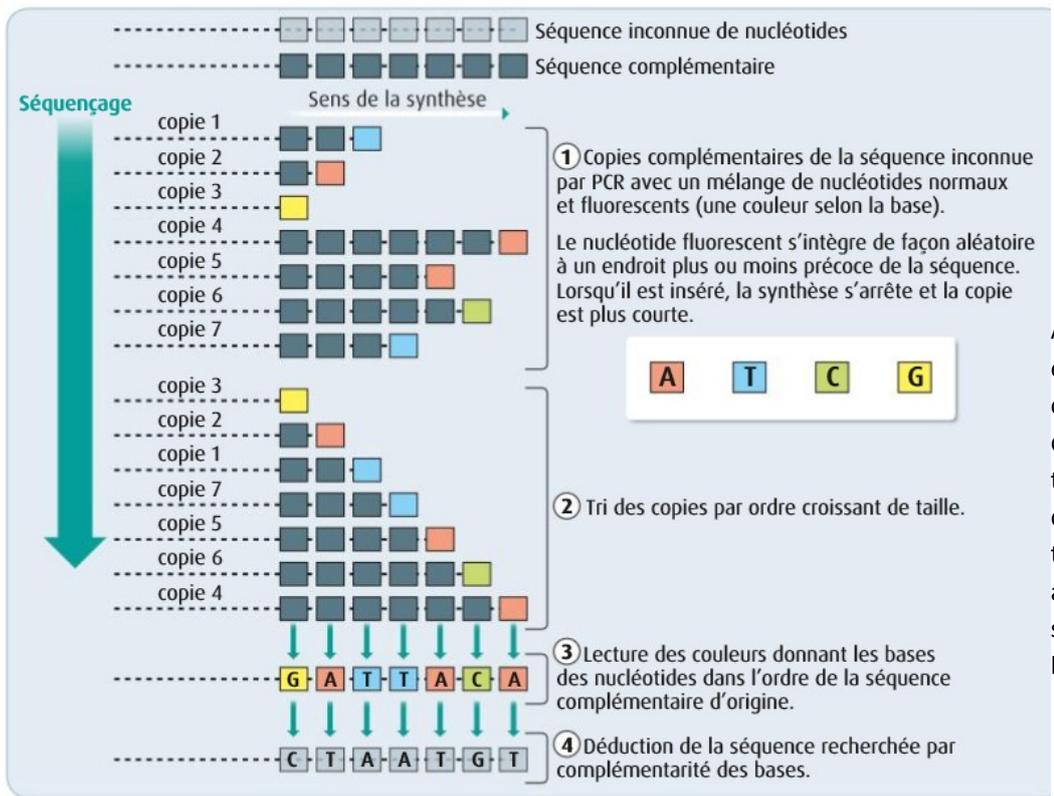
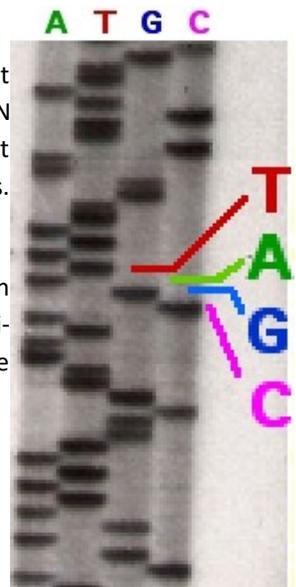
Le but de cette partie est de voir ou de revoir des technologies et des concepts indispensables à la réalisation du TD d'aujourd'hui.

Doc.1 : Le début du séquençage : La Méthode SANGER : En 1977, Frederick SANGER met au point

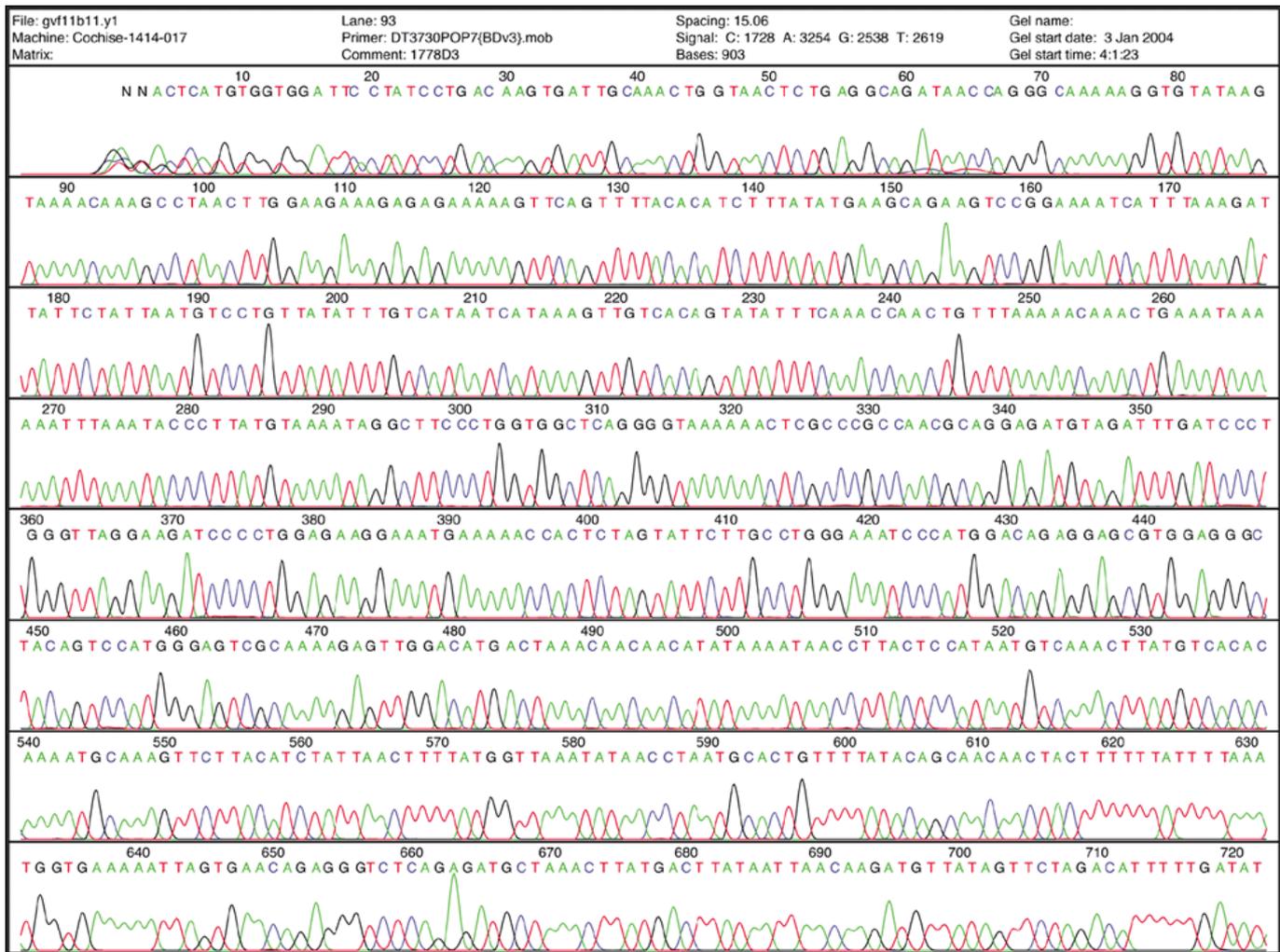
une méthode de séquençage d'ADN par synthèse enzymatique : l'ADN polymérase se sert du brin à séquencer comme modèle et on introduit des nucléotides dXTP normaux ET des nucléotides ddXTP fluorescents. Les ddXTP provoquent l'arrêt de la répllication.



Dans sa forme initiale, cette méthode utilisait 4 bacs différents, chacun possédant un seul ddXTP radioactif. On obtenait le type de résultat ci-contre après avoir fait migrer sur gel les extraits d'ADN obtenus (plus le morceau va long, plus il est de petite taille...)

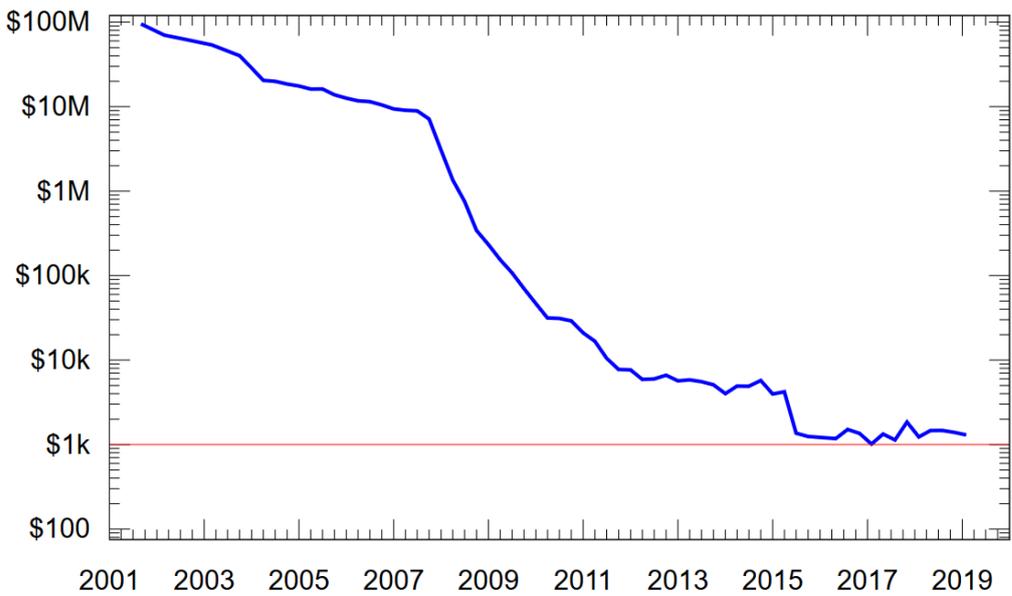


A présent, on mélange tout et les ddXTP sont fluorescents dans 4 longueurs d'ondes différentes. On obtient alors après migration des morceaux dupliqués de tailles variables, et après analyse, un spectre d'absorption qui nous permet de lire la séquence.



Résultat d'un SANGER moderne(spectres) et son interprétation(séquence).

Doc. 2 : Coût de séquençage d'un génome humain (USD)



Organismes	Taille du génome	Date du premier séquençage complet
Bactériophage MS2 RNA (virus)	3600 bases	1976 (le premier génome complet connu)
<i>Haemophilus influenzae</i> (bactérie)	1,8 millions de paires de base	1995
Levure de bière (champignon)	12 millions de paires de base	1996
Humain, <i>Homo sapiens</i>	3 milliards de paires de base	2004

Doc.3 : Mutations et apparitions de nouveaux allèles.

Notre but est l'étude des parentés, les **mutations touchant des cellules somatiques n'ont donc aucun intérêt** pour nous (retrouvez la raison de ce fait). Nous allons donc nous intéresser aux **mutations touchant des cellules germinales** (à l'origine des gamètes). Ces mutations peuvent en effet se transmettre aux descendants. L'apparition d'un nouvel allèle chez l'enfant dont les parents ne le possède pas se nomme mutation *de novo*.

Rq : Une mutation très précoce d'une cellule embryonnaire lors de embryogenèse peut engendrer l'apparition d'un nouvel allèle si la cellule d'origine est l'ascendante des cellules germinales...

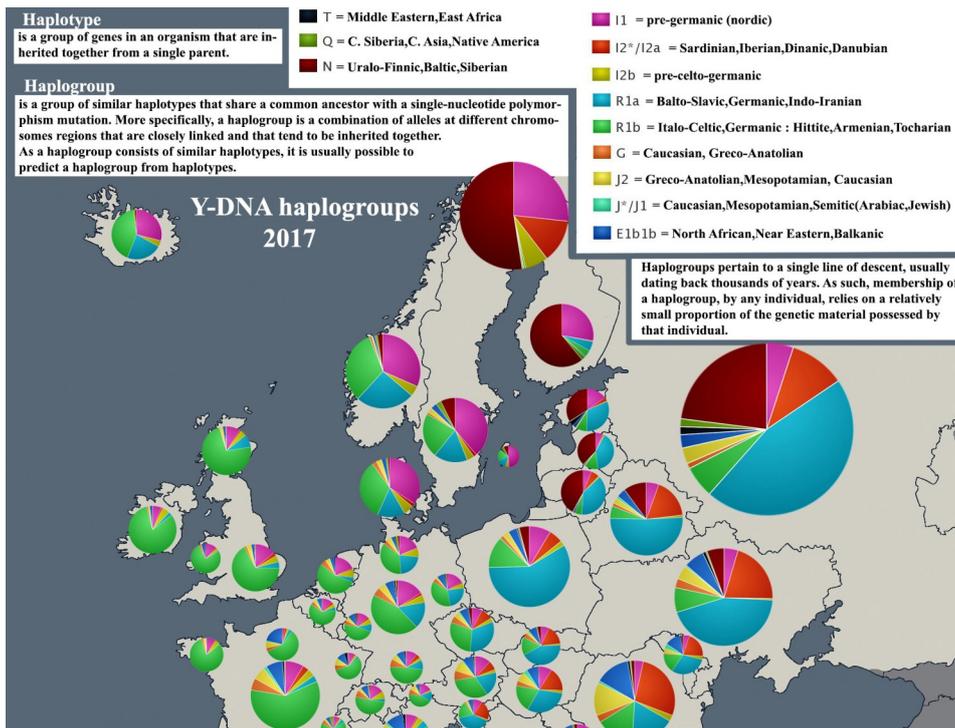
Doc.4 : Notion de SNP et d'haplogroupes:

4a : Les SNP : Il s'agit des Single Nucleotide polymorphisme, on s'est rendu compte que les **différences de séquences d'ADN** reposaient sur quelques nucléotides bien précis. Chez l'humain, seul **0,1% des nucléotides sont différents entre tous les allèles des gènes**. La variation doit être située à un endroit spécifique du génome et apparaître sur une proportion supérieure à 1% de la population pour être caractérisée comme SNP. Les SNP peuvent se retrouver au sein de **régions codantes** de gènes (exon), de **régions non codantes de gènes** (intron), ou de **régions intergéniques**, entre les gènes.

Dans le cas où les SNP se retrouvent au sein des régions codantes, celles-ci ne vont pas obligatoirement modifier la séquence d'acide aminé de la protéine produite, et ce, grâce à la redondance du code génétique. Les SNP qui se retrouvent dans des régions non codantes peuvent avoir des conséquences sur l'épissage, les facteurs de transcription, ou sur les séquences d'ARN non codant.

Tableau de répartition des SNP connus sur les chromosomes humains

Chromosome	Nombre de SNP	Chromosome	Nombre de SNP	Chromosome	Nombre de SNP
		8	4962	16	5 771
1	16759	9	5790	17	6392
2	748	10	6014	18	2682
3	10112	11	6931	19	7664
4	6995	12	7375	20	5381
5	9146	13	2847	21	3478
6	13888	14	5827	22	5400
7	12389	15	4343	X	3 253
Total	177 594			Y	63



4b : Notions d'haplogroupes : Un haplogroupe réunit les humains qui partagent un ensemble de mêmes SNP. Ce partage montre une origine commune. Ainsi, l'étude des haplogroupes montrent comment des groupes de population se sont déplacés sur la terre. Les haplogroupes définissent donc aussi une zone géographique. Les haplogroupes plus anciens sont donc souvent plus grands et répandus que les haplogroupes jeunes, issus des anciens.

NB : convention d'écriture
L'haplogroupe L3 est caractérisé par le SNP : C769T = C est le nucléotide des individus de l'haplogroupe L3 en position 769, T est le nucléotide en position 769 des individus n'appartenant pas à l'haplogroupe. (illustration : répartition des haplogroupe en Europe, source : eupedia.com)

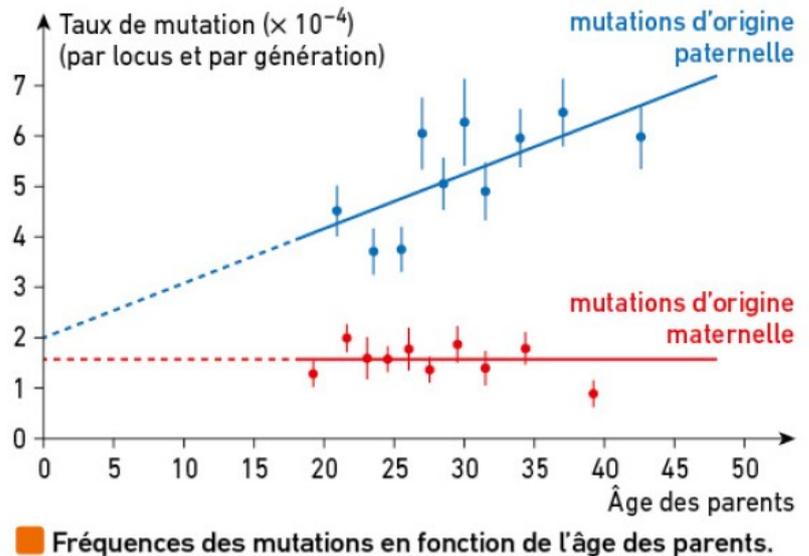
4c : travail des haplogroupes sur les chromosomes Y et/ou sur l'ADN mitochondrial.

L'étude des haplogroupes basés sur les SNP du chromosome Y permet de réunir des hommes, qui partagent un ancêtre dans la pure lignée paternelle. Le chromosome Y est toujours transmis du père au fils.

Les mitochondries possèdent un chromosome circulaire. Seules les mères donne ce chromosome à leurs descendants(ceux de l'ovocyte perdurent dans la cellule œuf alors que le spermatozoïde ne lègue pas les siens, ils restent DEHORS!!!). Un haplogroupe ADNmt réunit femmes et hommes qui partagent un ancêtre dans la pure lignée maternelle. L'ADNmt est toujours transmis de la mère à ses enfants.

Doc.5 : Fréquence des mutations transmises en fonction de l'âge des parents.

A chaque génération, des mutations sont transmises à la descendance. Cette constante permet de dater la séparation de deux lignées. Plus il y a de différences dans leurs génomes, plus la séparation est ancienne : c'est la notion d'horloge moléculaire.



Doc en plus, intéressant, à propos, mais pas utile ici : Des empreintes pour distinguer des individus sur la base de leur ADN

L'empreinte génétique repose sur le fait suivant : bien que deux humains aient une large majorité de leur patrimoine génétique identique, un **certain ensemble de séquences dans leur ADN reste spécifique à chaque individu (en raison du polymorphisme)**. Ce sont ces séquences spécifiques d'un individu que l'analyse d'empreinte génétique permet de comparer. des **portions qui ne codent aucune protéine**. Ce sont certaines d'entre elles, appelées les **microsatellites et minisatellites**, qui sont très variables selon les individus et permettent donc d'établir les empreintes génétiques.

Les microsatellites et minisatellites sont des séquences de nucléotides composées de répétitions de séquences plus petites. Il y a :

- les séquences répétées en tandem courtes, appelées aussi **microsatellites** ou STR, pour Short Tandem Repeats, en anglais. La plupart des séquences répétées sont des répétitions de quatre nucléotides, mais les longueurs de deux à cinq bases sont aussi étudiées. Exemple : CTGGCTGGCTGGCTGGCTGGCTGG
- les **minisatellites**, ou VNTR, pour « Variable Number Tandem Repeats ». Les séquences répétées sont des répétitions de 10 à 100 nucléotides. On regroupe parfois ces méthodes sous le nom « Multiple Loci VNTR Analysis » (MLVA).

En France et aux États-Unis, on utilise couramment treize loci (régions de séquence répétée) pour une identification. Et comme chaque locus est composé d'une certaine séquence de répétition (Par exemple, pour trois locus A, B et C, indépendants, et pour lesquels il existe plusieurs versions (A1, A2, A3, et B1, B2, etc.) , on peut dire que Probabilité (A1, B2, C1)=Probabilité (A1) x Probabilité (B2,) x Probabilité (c1)....Ainsi, pour 13 locus, la probabilité d'avoir deux séquences identiques pour deux individus différents non apparentés est estimée à 1 chance sur 10^{18} . De plus, chaque individu possède 2 séquences différentes.

→ demandez moi des nouvelles de Michael Jackson en cours et on en reparlera....

Partie 1 : Construction de l'histoire humaine et des liens entre individus : utilisation des allèles liés à un caractère monogénique.

Notre patrimoine génétique subit donc des **mutations** ayant plus ou moins d'influence sur nos **phénotypes**. Notre environnement engendre des pressions sur ces phénotypes en favorisant le maintien de certains au détriment d'autres dans les générations successives, c'est le phénomène de **sélection naturelle** abordé en seconde. Ces facteurs rythment nos évolutions et leurs études permettent de retracer une histoire de l'humanité et les liens qui existent entre les peuples.

On cherche, à travers un exemple unique, à retracer l'histoire de l'apparition et de l'expansion d'un nouveau phénotype et d'en déduire des liens entre les humains.

Consigne : Expliquer l'origine de cette diversité allélique et représenter les liens de parentés entre 3 peuples.

Aide à la résolution :

Travail sur une famille :

- ✓ 1- Repérer les SNP des allèles du gène étudié en fonction du phénotype de l'individu
- ✓ 2- Construire une explication de l'apparition de ces deux phénotypes.
- ✓ 3- Expliquer les liens de parenté entre les individus en fonction de cette explication.

Travail sur les patrimoines des peuples :

- ✓ 4- Construire un arbre phylogénétique « approximatif » des peuples endémiques anglais(Europ LP), thaïlandais(Eur LNP), centrafricains(Afr LNP) et marocains(Afr LP).

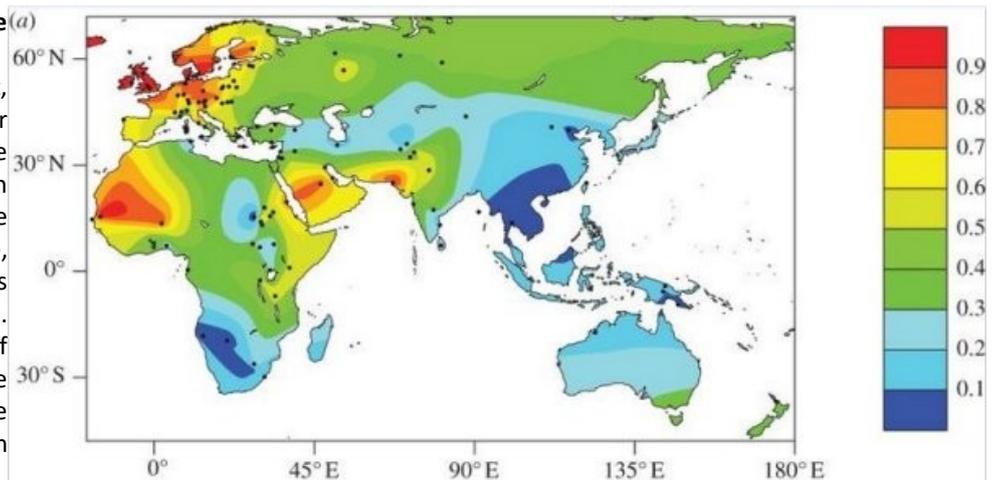
Ressources :

- Un ordinateur, Anagène et une banque de séquences.

- un ensemble documentaire et un ordinateur équipé d'un accès internet et d'un navigateur web

Doc 1 : Lactase persistance dans le monde.

Le lactose, principal glucide du lait, est un disaccharide formé par l'union d'une molécule de glucose et de galactose. Son absorption nécessite au préalable une hydrolyse réalisée par la lactase, enzyme de la bordure en brosse des cellules de l'épithélium intestinal. Les enfants produisent tous (sauf maladie) de la lactase mais une partie des adultes perd cette production (Lactase Non Persistence)



La fréquence du phénotype Lactase Persistence à l'échelle mondiale est estimée à 35% mais varie considérablement suivant les populations. Les plus fortes fréquences sur le continent européen sont observées dans le nord-ouest de l'Europe, en particulier dans les îles britanniques et la Scandinavie où elles varient entre 89% et 96%. On constate un déclin du phénotype LP du nord au sud et de l'ouest à l'est de l'Europe. Dans l'est de l'Asie, la fréquence est très faible. En Afrique, la distribution est hétérogène avec de forts contrastes entre populations voisines.

Doc 2 : Génétique de la lactase.

Le fait de pouvoir digérer du lactose est un phénotype mono-génétique. Nous nous intéresserons donc à ce seul gène de la lactase.

La base de données vous permet de visualiser : (dans le dossier Candia)

- les allèles de personnes LP et LNP d'une même famille → [Famille-LP-LNP.edi](#)

- Les ARNm présents dans les cellules intestinales des personnes LP et LNP → [ARNM-LP-LNP.edi](#)

- Les séquences régulatrices du gène de la lactase dans la famille étudiée → [Reg-Famille-LCT.edi](#)

- Les séquences régulatrices d'individus européens du néolithique (-8000ans)-> [REG pop et neolithique.edi](#)

- Les séquences régulatrices d'individus humains et d'un chimpanzé actuels → [Alleles LCT HumainsEtPan.edi](#)

Doc 3. : Des avantages à digérer le lait à l'âge adulte.

Le **premier avantage** peut résider dans l'apport énergétique du lait. On estime que la production de lait par une vache préhistorique devait être entre 400 et 600 kg suite à une gestation. Après avoir soustrait la quantité de lait nécessaire au jeune veau, il reste 150 à 200kg. Ceci est presque équivalent à l'apport énergétique obtenu à partir de la viande d'une vache. Cet apport énergétique qui nécessite de digérer le lactose pouvait être particulièrement important dans les périodes de disette entre les périodes de récolte des cultures céréalières et donc favoriser la survie des personnes LP.

Un **deuxième avantage** sélectif fourni par le lait est en rapport avec l'assimilation du calcium. La vitamine D favorise l'absorption intestinale du calcium. Dans les régions nordiques où le rayonnement UV est faible pendant plusieurs mois de l'année (voir dossier pigmentation de la peau) la production cutanée de précurseurs de la vitamine D, sous l'action des UV, est réduite. Si l'apport alimentaire des éleveurs du néolithique fournissait peu de vitamine D, il en résultait des risques de rachitisme. Le lait en apportant le calcium et un peu de vitamine D pouvait contribuer à l'éviter. L'avantage sélectif fourni par la consommation du lait selon cette hypothèse n'est pas en relation directe avec la capacité à digérer du lactose, puisque ce sont d'autres constituants du lait qui sont en jeu. Il est toutefois en relation indirecte puisque les personnes LNP ne pouvant consommer du lait à cause de leur incapacité de digérer le lactose, ne bénéficiaient pas de l'apport vitaminique et en calcium du lait.

Une **autre explication** est en rapport avec le climat. Dans les régions où sévit la sécheresse, le lait représente une source d'eau non polluée. Les personnes LNP ne pouvaient bénéficier de cette source d'eau et au contraire, les diarrhées en cas de consommation de lait pouvaient entraîner une déshydratation pouvant être mortelle.

Doc 4. : Quel est le premier des deux ???

Si l'on veut, au sein d'une population hétérogène à propos d'un gène, déterminer quels sont les allèles ancestraux, il faut comparer les séquences des différents allèles et celle d'un lointain cousin possédant le même gène.

La séquence identique à celle du lointain cousin sera la plus ancienne(ancestrale) et les autres issues de mutations plus récentes...

Partie 2 : Construction de l'histoire humaine et des liens entre individus : utilisation des haplogroupes.

Rabbie-Régis a fait réaliser un test ADN pour découvrir ses origines doutant toujours que ses parents soient vraiment ses parents tant ils sont « relous », c'est de son âge !!! Il choisit l'option « double étude » du laboratoire : L'option propose de déterminer ses haplogroupes selon son ADN mitochondrial et son chromosome Y. Il reçoit les résultats : il fait partie de l'haplogroupe L3.

Mirtille et Léa, ses sœurs, ont commandé sur internet un kit de test ADN pour déterminer ses origines, le site ne leur propose pas l'option « double étude » mais elles le commandent quand même !!!

Les résultats obtenus se basent sur les analyses de l'ADN mitochondrial.

NB : le travail proposé ici n'est qu'une version très restreinte de l'analyse réelle qui est faite à partir des tests ADN, qui porte sur un volume de données beaucoup important. C'est pourquoi des résultats réels proposeraient une appartenance à une population plus précise que ce que nous ferons.

Consigne : A partir des documents ci-dessous, exposez les origines les plus probables de la fratrie et expliquez pourquoi le labo ne leur propose pas la « double étude » .

Et surtout, que dire de leur filiation ?

Aide à la résolution :

- ✓ 1- Repérer les SNP que ces « frangin(e)s » possèdent
- ✓ 2- Comparer avec les données fournies.

Ressources :

- Un ordinateur, Anagène et une banque de séquences.

ATTENTION : Les SNP relatifs à l'apparition des haplogroupes SNP N, R P, JT et HV ont été omis(parce que je ne les ai pas trouvés...)

- un ensemble documentaire et un ordinateur équipé d'un accès internet et d'un navigateur web

Doc 1 :

Document C : SNP spécifiques pour quelques haplogroupes sur l'ADN mitochondrial

Haplogroupe	SNP Spécifiques
M	G10400A
A	T1736C
H	C2706T
L3	C769T

NB : convention d'écriture : L'haplogroupe L3 est caractérisé par le SNP : C769T = C est le nucléotide des individus de l'haplogroupe L3 en position 769, T est le nucléotide en position 769 des individus n'appartenant pas à l'haplogroupe

Doc 2 : Encadrement législatif en France :

Le recours aux tests ADN pour déterminer ses origines s'est démocratisé ces dernières années aux États-Unis, succès relayé et amplifié par des célébrités. Ce succès s'est propagé dans d'autres pays. En France, on trouve facilement sur Internet des vidéos de vidéastes plus ou moins connus qui en ont aussi utilisé.

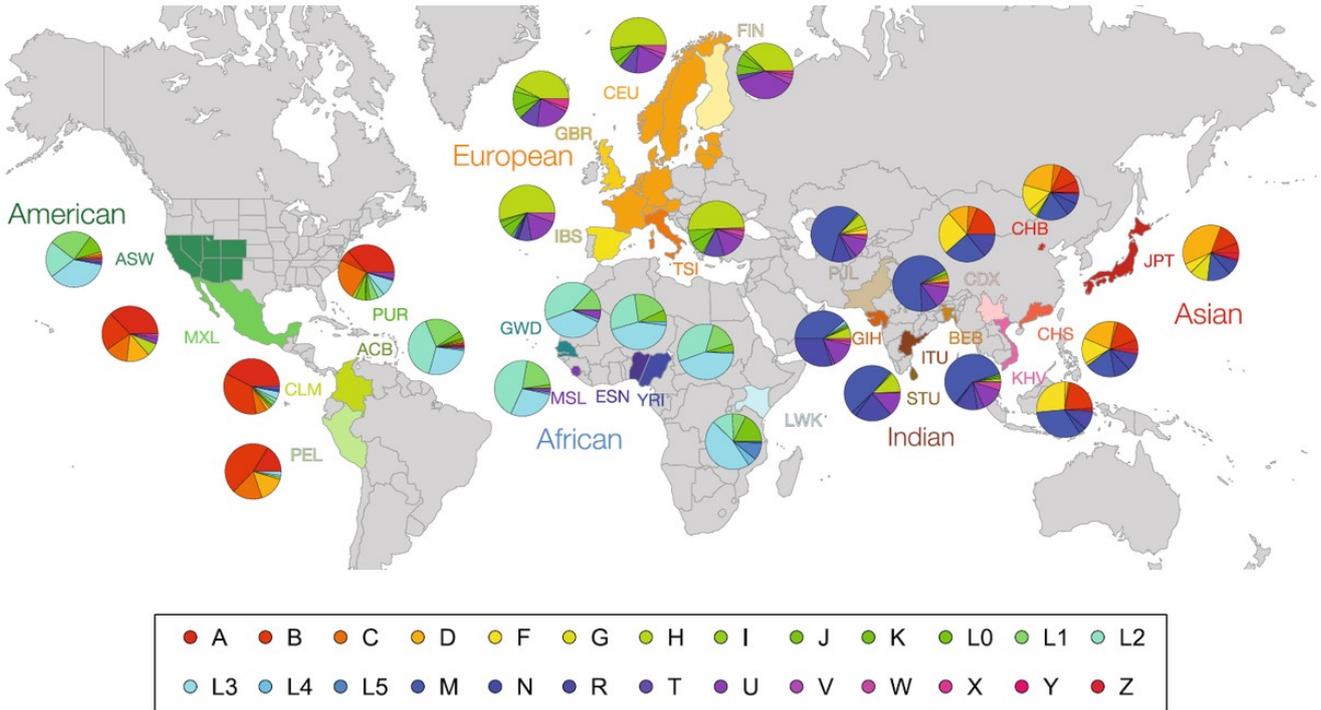
Ces tests ADN posent un certain nombre de questions, qu'il pourrait être intéressant d'aborder avec les élèves :

- ces tests sont illégaux en France (3 750 euros d'amende), bien qu'il ne soit guère compliqué d'en commander sur Internet. Les lois de bioéthique de 1994 ont été établies et l'on peut lire l'article 16-10 du Code civil : « l'examen des caractéristiques génétiques d'une personne ne peut être entrepris qu'à des fins médicales ou de recherche scientifique ».
- les compagnies qui mettent en vente ces kits lorgnent sur le marché de la santé, en faisant le lien entre allèles et risque de maladie. Se pose alors la question d'un individu qui reçoit cette information en dehors de tout cadre médical, sans accompagnement pour la comprendre et en cerner les limites (par exemple influence de l'environnement) dans les maladies multifactorielles.
- ces mêmes compagnies, au fur et à mesure qu'elles ont des clients, remplissent leurs bases de données et pourraient les monnayer, par exemple à des organismes de recherche médicale publics ou privés.

- la fiabilité de l'interprétation données sur les origines dépend du nombre de SNP analysés et de la base de données utilisé. Ainsi deux compagnies ou la même compagnie à deux temps différents ne vont pas fournir le même résultat.

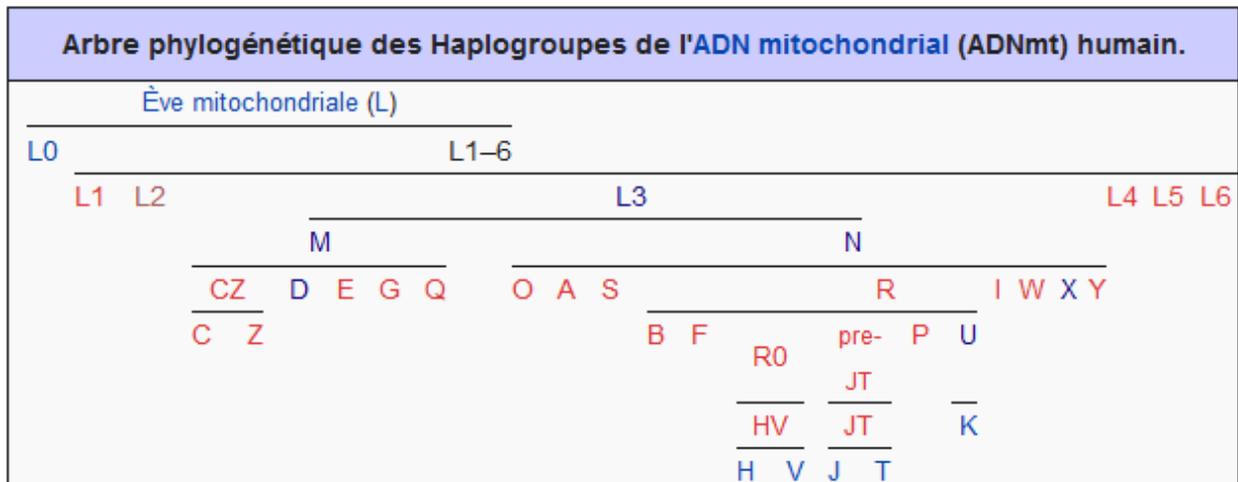
Doc3 : Répartition mondiale des principaux haplogroupes.

A



Distribution de l'haplogroupe d'ADNmt humain contemporain, basée sur l'analyse de 2 054 individus de 26 populations.(a) Diagrammes à secteurs sur la carte.

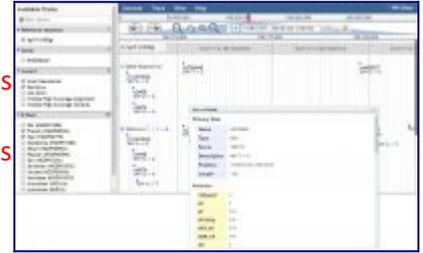
Doc 4 : L'étude des SNP des ADN mitochondriaux permet de tracer une généalogie des haplogroupes.



En plus mais pas très utile.

Un outil en ligne permet de visualiser et de comparer des séquences anciennes et actuelles, et des SNP sont signalés.

Un clic donne accès à des informations sur chaque variant, y compris des fréquences d'allèles.



Available Tracks

- filter tracks
- Reference sequence 1
 - hg19 (1000g)
- Genes 1
 - ENSEMBL67
- Ancient 5
 - Altai Neandertal
 - Denisova
 - Ust-Ishim
 - Vindija High Coverage Alignment
 - Vindija High Coverage Variants
- B-Team 14
 - Dai (HGDP01308)
 - French (HGDP00533)
 - Han (HGDP00775)
 - Mandenka (HGDP01286)
 - Mbuti (HGDP00982)
 - Papuan (HGDP00546)
 - San (HGDP01036)
 - Sardinian (HGDP01076)
 - Yoruba (HGDP00934)

Genome Track View Help

0 50,000,000 100,000,000 150,000,000

24,891,500 24,892,000

hg19 (1000g) reference Zoom in to see sequence Zoom in to see sequence

SNV G -> C
rs58955784
SNV A -> G

Han (HGDP00775) 24105
SNV C -> T

Mandenka (HGDP01286)
SNV C -> T

Mbuti (HGDP00982)

1- choisir le chromosome à étudier

2- cocher la référence

3- choisir les séquences populations à analyser (les SNP vont être présentés 4-)

5- zoomer assez sur les séquences pour visualiser les SNP....

Malheureusement, il n'y a pas de données pour l'ADN mitochondrial....

Partie 3 : Construction de l'histoire humaine et des liens entre individus : connaissance des individus disparus.

Depuis plus d'un siècle, l'humain se passionne pour l'étude de ses origines. Longtemps imaginée comme une lignée d'individus, l'installation du genre Homo est à présent reconnu comme un phénomène complexe buissonnant dont les racines se trouvent dans l'Est africain.

On cherche à expliquer une partie de la diversification du genre Homo à travers l'étude comparée de quelques fossiles.

Lire le doc1 :

Consigne : A quelle espèce appartient la phalange retrouvée ? Fort de votre conclusion, construire un arbre phylogénétique des individus fossiles : H. néanderthalien, H.sapiens, l'individu à la phalange et celui à la molaire.

Aide à la résolution :

- ✓ 1-Proposer une ou des hypothèses répondant à la première partie de la consigne.
- ✓ 2-Proposer une stratégie de résolution
- ✓ 3- Mettre en œuvre la stratégie.
- ✓ 4- Construire un arbre phylogénétique.

Ressources :

- Un ordinateur, Anagène et une banque de séquences.

- un ensemble documentaire et un ordinateur équipé d'un accès internet et d'un navigateur web



Doc 1 : la découverte de fossiles dans la grotte de Denisova

La grotte de Denisova se trouve dans le massif de l'Altaï en Russie, à proximité des frontières avec la Chine et la Mongolie. Elle est étudiée depuis les années 1970 et a livré de nombreux artefacts attestant de son occupation depuis 280 000 ans.

En 2008 une phalange a été découverte dans des sédiments datés de 50 000 à 30 000 ans.

À cette époque, dans cette région, on connaît au moins deux espèces contemporaines appartenant au genre Homo : Homo sapiens (comme l'homme de Cro Magnon, qui est en fait notre espèce) et Homo neanderthalensis. La grotte de Denisova a livré d'autres

restes humains, comme une molaire. De l'ADN mitochondrial a pu être extrait de cette molaire.

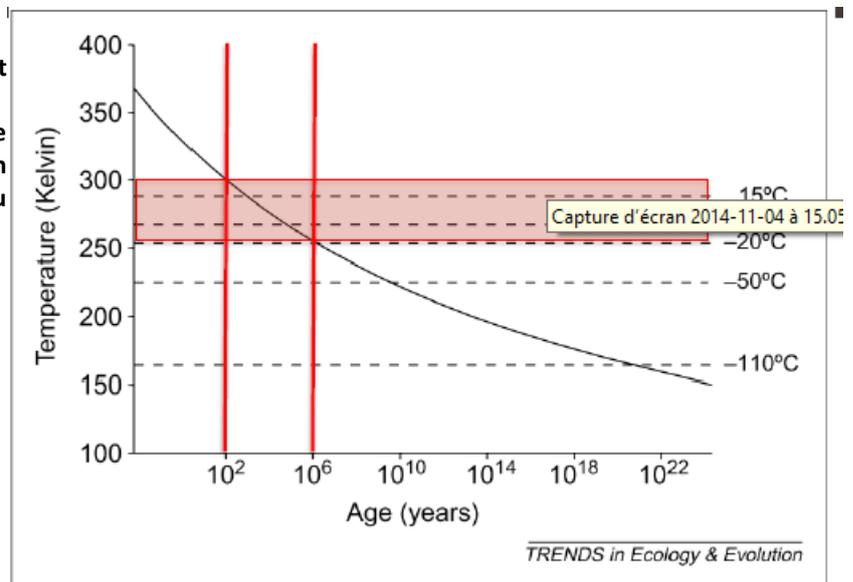
On se demande à quelle espèce elle appartient.

Doc. 2: Limites de conservation de l'ADN.

Deux problèmes se posent lorsqu'on veut utiliser des études ADN en paléontologie :

- on est confronté à de l'ADN parasite ancien (bactéries) ou récent (contamination lors de la sorite par des micro-organismes ou par les manipulateurs...)

- L'ADN se dégrade avec le temps.



Temps de survie d'un ADN de 100 pb en fonction de la température (si le seul dommage est la dépurination de l'ADN)

Doc 3 : On peut se poser la question du % de différence nécessaire et suffisant pour déclarer qu'il s'agit d'espèces différentes.

Pour s'en faire une idée, on peut imaginer de comparer

- des séquences appartenant à des espèces différentes : on pourra ici utiliser néanderthal et sapiens
- des séquences appartenant à des individus appartenant à la même espèce.

D'autres ADN néanderthaliens ont été séquencés, par exemple :

voir la source : [neanderthalensis les cottes z4 1514 mt](#)

Rq : ATTENTION à comparer avec discontinuité pour ne pas lire plus de différence qu'il y en a en réalité !!



une fois la comparaison faite, l'icône permet d'afficher des informations liées à au traitement effectué.

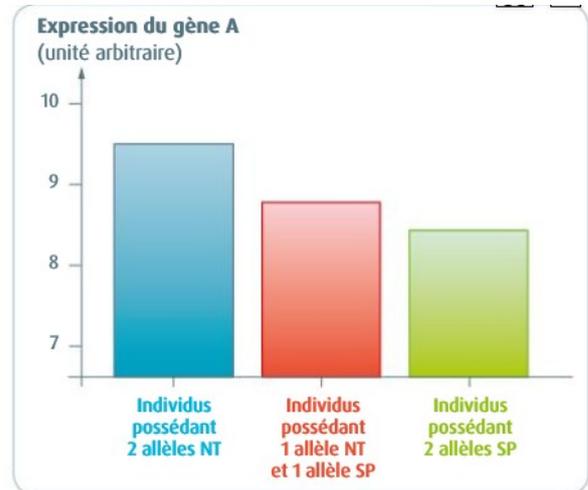
En plus :

Un allèle néandertalien sélectionné

Une séquence d'ADN d'origine néandertalienne a été identifiée dans le génome de certaines populations actuelles. Cela s'explique par une hybridation entre des ancêtres *Homo sapiens* et néandertaliens. Cette séquence d'ADN est un allèle d'un gène appelé A, régulant l'expression des récepteurs TLR6, rendant la réponse immunitaire plus performante.

Allèles	Origine de l'allèle	Répartition géographique	Fréquence moyenne dans les populations
V	<i>H. sapiens</i>	Partout dans le monde	39-88 %
VI, VIII et IX	<i>H. sapiens</i>	Essentiellement en Afrique	< 1 %
III	néandertalienne	Essentiellement hors d'Afrique	11-51 %

▲ 1. Fréquence et origine de 5 allèles du gène A dans les populations actuelles. L'analyse de la répartition et de la fréquence des allèles est effectuée sur 1000 génomes de populations réparties dans le monde.



▲ 2. Expression du gène A en fonction des allèles de ce gène. Chaque individu possède 2 exemplaires du gène de régulation. NT est l'allèle d'origine néandertalienne et SP est un des allèles d'origine *Homo sapiens*.

QUESTION D'après les documents, relevez l'argument permettant d'affirmer que la présence de cet allèle néandertalien est le résultat d'une sélection et proposez une explication à l'avantage sélectif de cet allèle chez les *Homo sapiens*.