

Chapitre 5. Du génome aux génomes : génomique des populations

Version détaillée

Introduction à la bioinformatique (SSV3U15, L2 SV AMU)

Jacques van Helden

Aix-Marseille Université

orcid.org/0000-0002-8799-8584

1. Variations génétiques au sein des populations humaines
2. Retracer les origines géographiques d'un individu dans son ADN
3. Du génome aux génomes : génomique des populations
4. Retracer l'évolution de l'espèce humaine dans l'ADN

Variations génétiques au sein des populations humaines

Rappel: typologie des mutations

- **Mutations ponctuelles** : un ou quelques nucléotides
 - Substitution : remplacement d'un résidu (une lettre) par un autre
 - Délétion : suppression d'un fragment de la molécule
 - Insertion : ajout d'un fragment de molécule
- **Variations structurales** : résultent de réarrangements chromosomiques, et affectent des régions plus larges
 - Délétions
 - Insertions
 - Translocations
 - Variations du nombre de copies

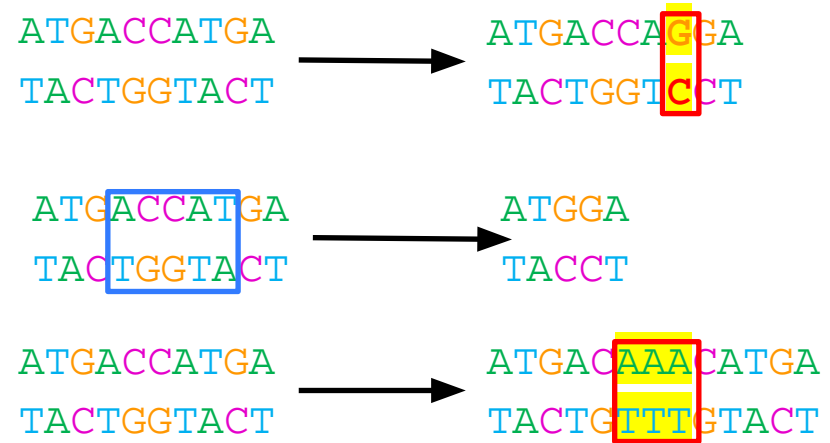
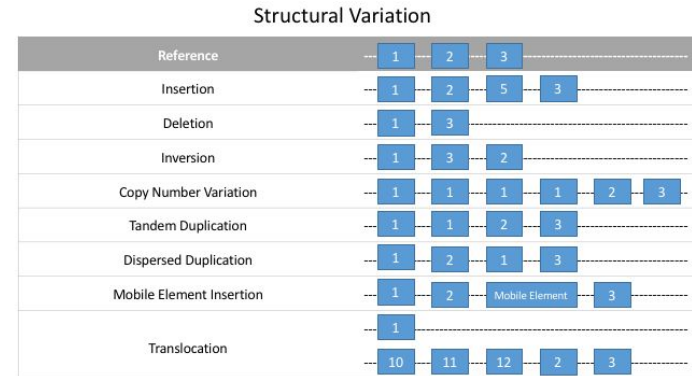


Figure 1: Depiction of different types of structural variants compared to the reference genome. Each different number represents a different gene.

Les **marqueurs génétiques** sont des petits fragments de génome qui présentent des variations interindividuelles.

- SNP: nucléotide variable d'un individu à l'autre. Ce type de variation est appelée « **SNP** » pour « **Single-Nucleotide Polymorphism** » (prononcez « **snip** »).

Type 1 CAGTGCGAGGG

Type 2 CAGTGCGGGGG

- **Microsatellite**: fragment de séquence génomique comportant une courte séquence (2 à 5 résidus) répétée (typiquement 5 à 50 fois selon les cas). Le nombre de répétitions varie souvent d'un individu à l'autre.

Type 1 CAGCAGCAGCAGCAGCAG

Type 2 CAGCAGCAGCAGCAGCAGCAG

Type 3 CAGCAGCAGCAGCAGCAGCAGCAG

Utilisation des marqueurs génétiques

- Les microsatellites sont utilisés pour l'identification génétique de personnes (police scientifique, tests de paternité, ...).
- Recherche en santé / soin : certains marqueurs sont associés à des maladies
- Génomique des populations : analyse des variations génétiques au sein des populations humaines, diversification des groupes humains.
- Marqueurs associés ou non à des traits phénotypiques.

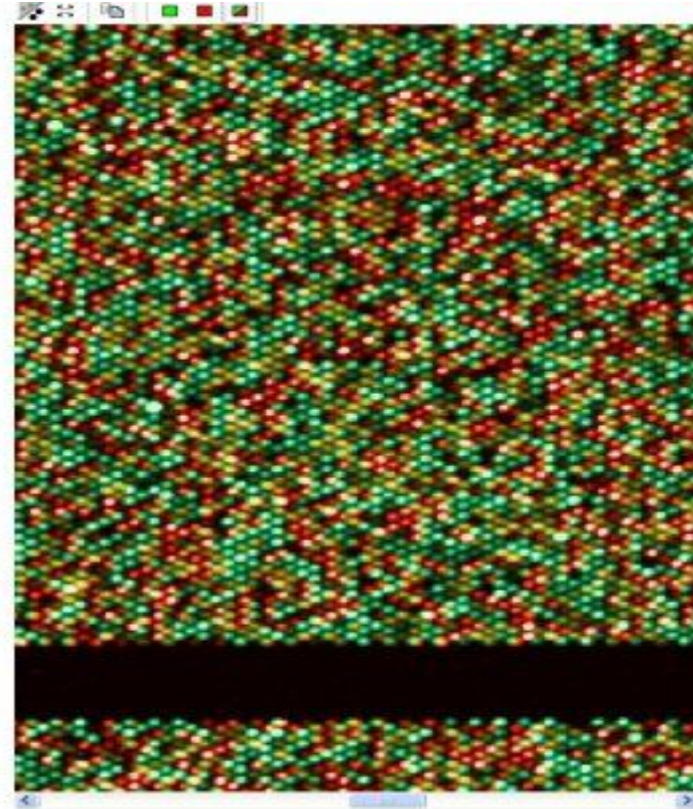
Profilage génomique par biopuces – le code-barre d'un génome individuel

- Depuis le début des années 2000, on a développé des **biopuces à SNP** pour le génotypage individuel.
 - Chaque point lumineux correspond à un SNP
 - La couleur indique le variant particulier (A, C, G, T) présent chez l'individu considéré.
 - Sur une seule biopuce ("microarray"), on peut caractériser de 500.000 à 6.000.000 SNP différents (le nombre a augmenté d'année en année).
- **Profil génomique** individuel : combinaison d'allèles spécifiques à un individu ("code barre").
- Usage
 - identification de personnes
 - caractéristiques phénotypiques
 - facteurs génétiques de risques de maladies
- Génotypage **partiel, mais peu coûteux** par rapport au séquençage d'un génome complet.

Biopuce



Détail



Définitions

- **Génotype** : ensemble des caractères génétiques d'un être vivant, qu'ils se traduisent ou non dans son phénotype (ensemble des caractères physiques et biologiques d'un individu).
Source : dictionnaire Larousse
- **Génotypage** : Fait de déterminer le génotype ou une fraction représentative du génotype d'un individu. *Source: dictionnaire Larousse.*

Méthodes de génotypage

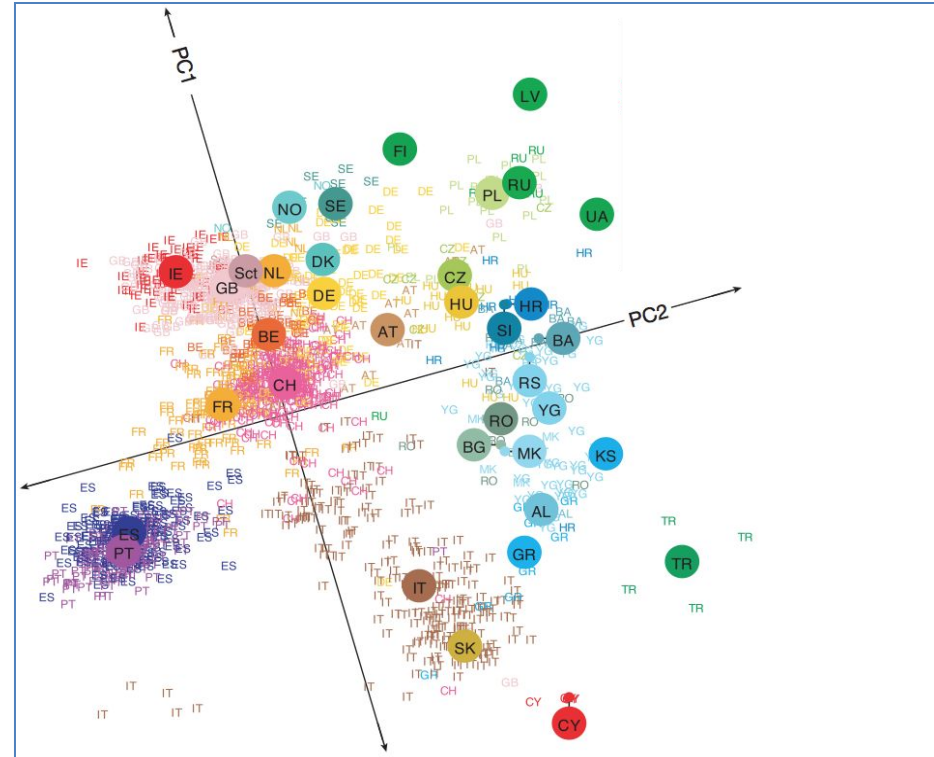
- **Biopuces à SNPs**
 - Il ne s'agit pas de séquençage à proprement parler, mais de caractériser les SNPs d'un individu
 - Les biopuces actuelles permettent de caractériser 500.000 à 6 millions de SNPs
 - Coût: ~80 \$ pour un génotypage à 600.000 SNPs (juillet 2026)
- **Séquençage génomique**
 - **Ciblé** : 1 gène ou un panel de gènes associés à une maladie
 - **Exome** : l'ensemble des exons
 - **Génome complet** (whole genome sequencing, WS) : Coût: ~200 US\$ / génome (2026)

Retracer les origines géographiques d'un individu d'après son ADN

Quand les gènes révèlent nos origines géographiques

Etude publiée en 2008 par Novembre et ses collaborateurs.

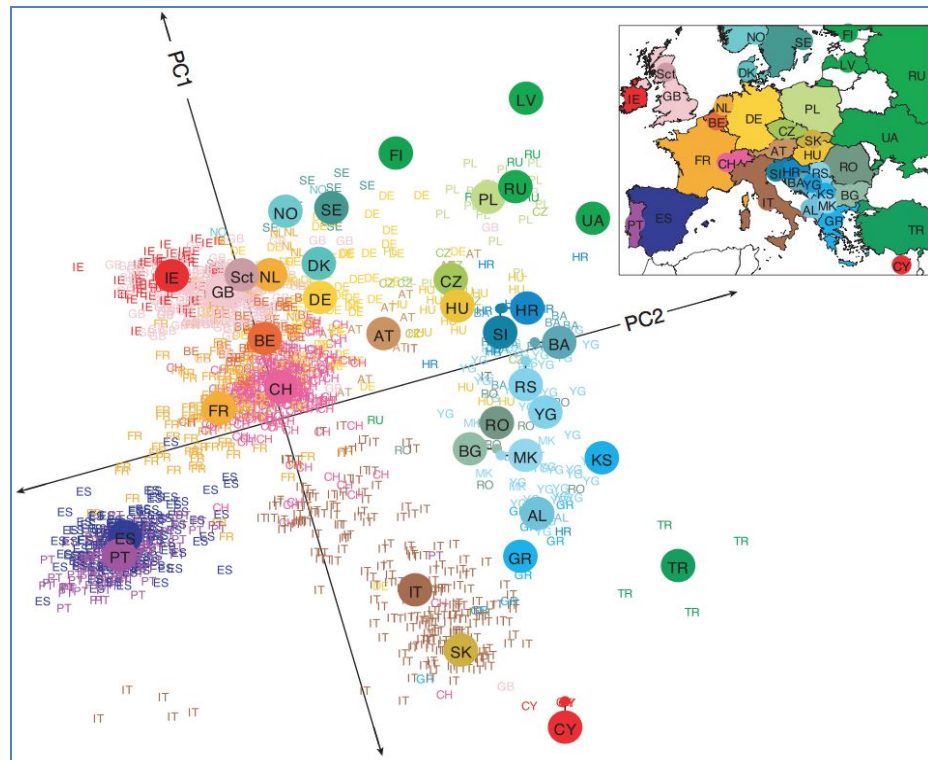
- Les chercheurs disposent du profil génomique (biopuces à 500.000 SNPs) de 3192 individus européens.
- Ils sélectionnent le **sous-ensemble des individus dont les 4 grands-parents vivaient au même endroit**, ce qui donne 1,387 individus.
- Ils utilisent l'**analyse des composante principales** (méthode statistique classique) pour positionner chaque individu sur une carte en 2 dimensions, en tenant compte uniquement des informations génétiques.
- Les **positions sur la carte** sont déterminées uniquement par la **similarité génétique entre individus** : les génotypes similaires sont proches, les génotypes plus différents sont plus éloignés.
- Dans un deuxième temps, on assigne à chaque point une couleur et des lettres qui indiquent son appartenance géographique.



Quand les gènes révèlent nos origines géographiques

Comparaison entre carte génétique et carte géographique

- Les individus de même origine géographique sont génétiquement proches
- La **carte génétique reproduit de façon impressionnante la carte géographique** (d'où le titre de l'article publié : *"Les gènes sont le miroir de la géographie en Europe"*).
- Les auteurs estiment que les profils génomiques de ce type permettent de déterminer l'origine géographique d'un individu européen avec une précision de 400 km, sur simple base de son profil génomique.
- Ils soulignent l'utilité de cette technique en criminalistique.
- Dans une réponse publiée dans le même journal, Vincent Detours souligne les risques de dérive de cette même méthode, si elle venait à être utilisée à des fins d'épuration ethnique.



Novembre et al. Genes mirror geography within Europe. Nature (2008)

<https://doi.org/doi:10.1038/nature07331>

Detours, V. Editorial comment should accompany hot papers online. Nature 455, 861

(2008). <https://doi.org/10.1038/455861b>

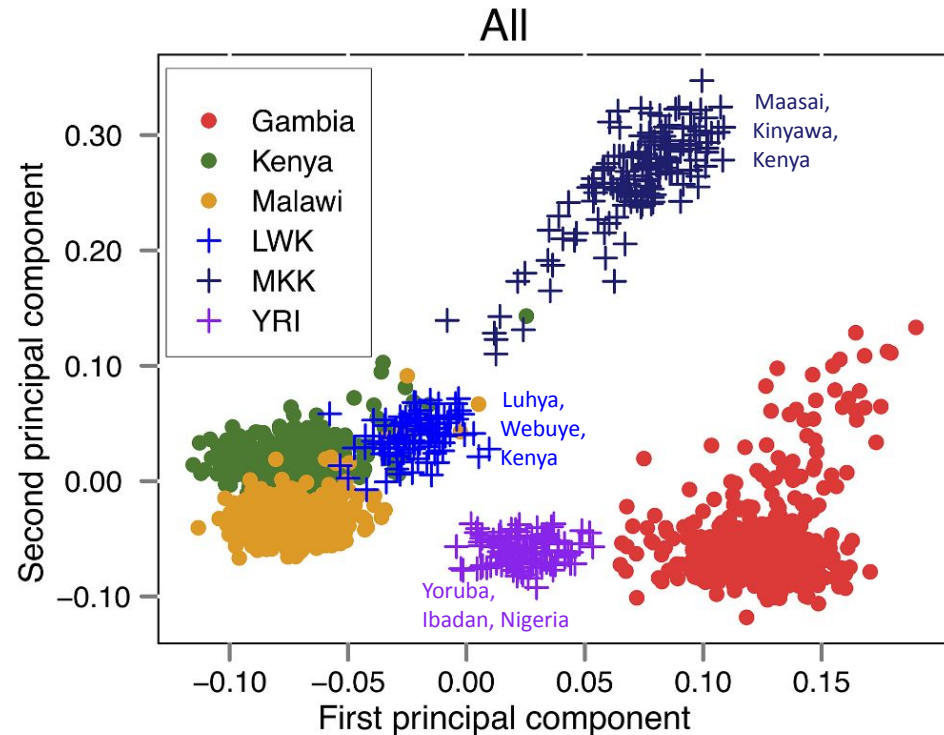
Stratification de populations africaines dans une étude sur la susceptibilité à la malaria

Profils génomiques d'individus africains dans HapMap

- Les profils génomiques obtenus par biopuces à SNP révèlent la structuration génétique de populations humaines.
- Les graphiques ci-contre montrent une projection sur 2 dimensions (sur base d'une méthode appelée analyse des composantes principales, ACP), où chaque point correspond à un individu, et les distances entre points sont (approximativement) proportionnelles au nombre de différences génétiques.
- On colore ensuite les individus selon leur origine géographique (cette information n'est pas utilisée pour le calcul des distances).
 - Nationalité (500 personnes par pays)
 - Groupes ethniques (Luhya, Massai, Yorubas)

Observation

- On observe que cette analyse relativement simple permet de distinguer la nationalité et le groupe ethnique au sein des populations africaines.

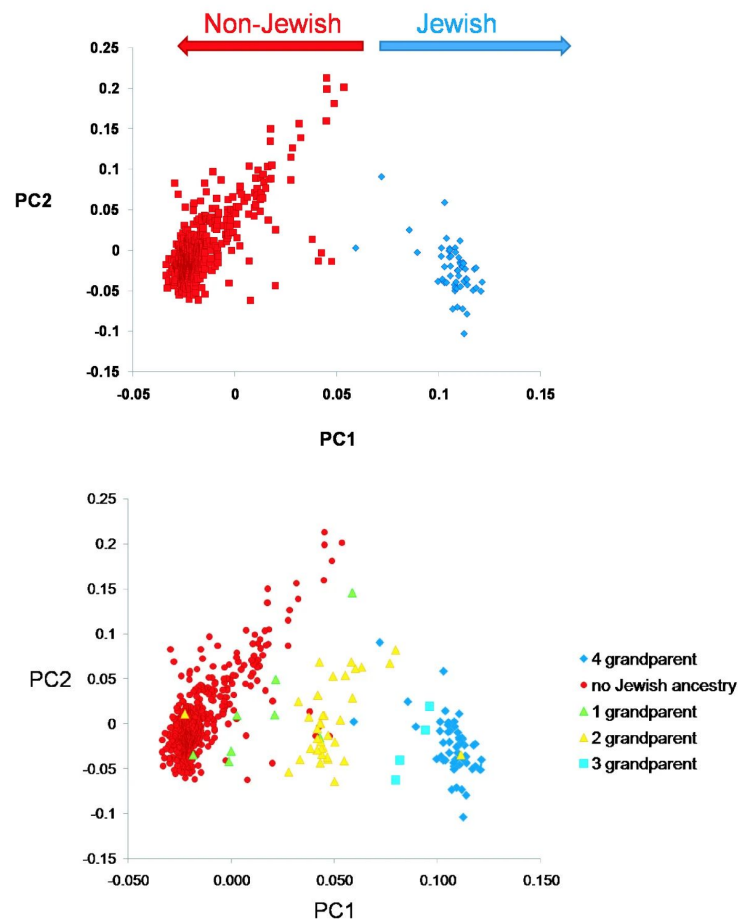


Principal components analysis (PCA) of the African populations from Hapmap 3 (LWK=Luhya in Webuye, Kenya, 90 individuals; MKK=Maasai in Kinyawa, Kenya, 143 individuals; YRI=Yoruba in Ibadan, Nigeria, 113 individuals) with 500 randomly selected control samples from each of the three study cohorts.

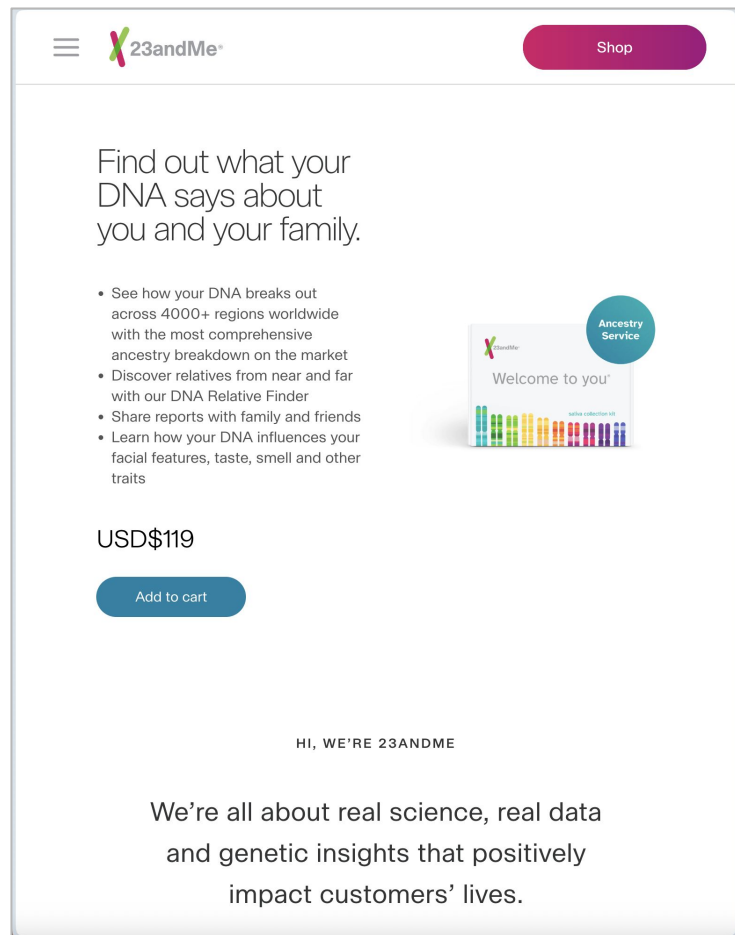
Le profilage génomique permet de discriminer les populations juives des non-juives

Profilage génomique d'individus vivant aux Etats-Unis

Une analyse des profils génomiques permet de distinguer, dans un échantillon américain, les personnes juives des non-juives (figure du haut), voire de distinguer le nombre de juifs parmi leurs grands-parents (figure du bas).



Need et al. A genome-wide genetic signature of Jewish ancestry perfectly separates individuals with and without full Jewish ancestry in a large random sample of European Americans. *Genome Biol* (2009) vol. 10 (1) pp. R7.
<https://doi.org/10.1186/gb-2009-10-1-r7>



The screenshot shows the 23andMe website interface. At the top left is the 23andMe logo with a hamburger menu icon. At the top right is a purple 'Shop' button. The main content area features a large heading 'Find out what your DNA says about you and your family.' followed by a bulleted list of features: 'See how your DNA breaks out across 4000+ regions worldwide with the most comprehensive ancestry breakdown on the market', 'Discover relatives from near and far with our DNA Relative Finder', 'Share reports with family and friends', and 'Learn how your DNA influences your facial features, taste, smell and other traits'. To the right of the text is a product card for the 'Ancestry Service' which says 'Welcome to you!' and shows a colorful bar chart. Below the product card is the price 'USD\$119' and a blue 'Add to cart' button. At the bottom, there is a greeting 'HI, WE'RE 23ANDME' and a closing statement: 'We're all about real science, real data and genetic insights that positively impact customers' lives.'

Find out what your DNA says about you and your family.

- See how your DNA breaks out across 4000+ regions worldwide with the most comprehensive ancestry breakdown on the market
- Discover relatives from near and far with our DNA Relative Finder
- Share reports with family and friends
- Learn how your DNA influences your facial features, taste, smell and other traits

USD\$119

Add to cart

HI, WE'RE 23ANDME

We're all about real science, real data and genetic insights that positively impact customers' lives.

La compagnie 23andme

- Fondée en 2007
- Financée par Google, puis des investissements croissants, notamment big pharma
- Propose différents services liés au génotypage.
 - origines géographiques
 - recherche de parents
 - réseaux sociaux basés sur les similarités génétiques
 - santé (facteurs génétiques de risque)
- Le modèle s'est étendu, et d'autres compagnies proposent des services similaires
- Note : pour des personnes résidant en France, le profilage génomique n'est pas autorisé, sauf dans un cadre très réglementé :
 - diagnostic médical, avec l'obligation de passer par un médecin généticien
 - enquêtes judiciaires

Génomique “récréative” : recherche des origines géographiques

Ancestry Breakdown

Dig deeper into your ancestry.

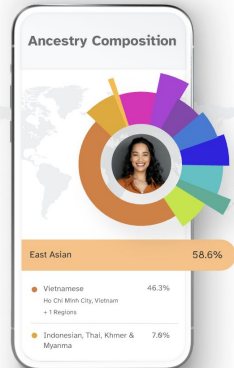
It's the most complete genetic breakdown on the market, and the most comprehensive portrait of you yet.

- **Ancestry Composition**

Discover where in the world your DNA is from across 4000+ regions — in some cases, down to the county level.

- **Ancestry Detail Report**

[See all regions](#)



- L'affichage de 23andme met un fort accent sur la recherche des origines géographiques.
- Cette question suscite un intérêt assez marqué aux Etats-Unis, vraisemblablement du fait de l'histoire de ce pays, qui repose sur plusieurs vagues de migrations (colonisation, commerce d'esclaves, ...).

Family and Friends

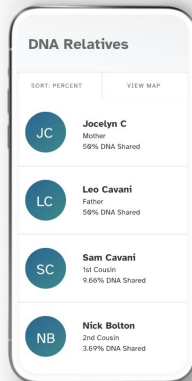
Find your people.
Open your world.

Discover people who share your DNA. From close family members to distant ones, you'll be amazed by the way your DNA relatives connect you to the world.

- **DNA Relative Finder (optional)**

Find your matches. Compare ancestries and traits. Message relatives directly to better understand your family connection.

- **Family Tree**



Ancestral History

Trace your path back thousands of years.

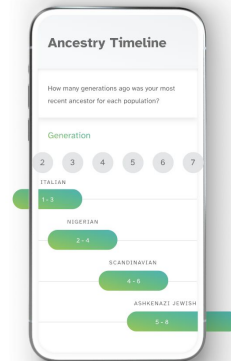
Travel back in time to gain a clearer picture of where you came from, where your ancestors lived and when they lived there.

- **Ancestry Timeline**

Learn when your most recent ancestors from each population lived, going back more than 8 generations.

- **Haplogroups**

- **Neanderthal DNA**

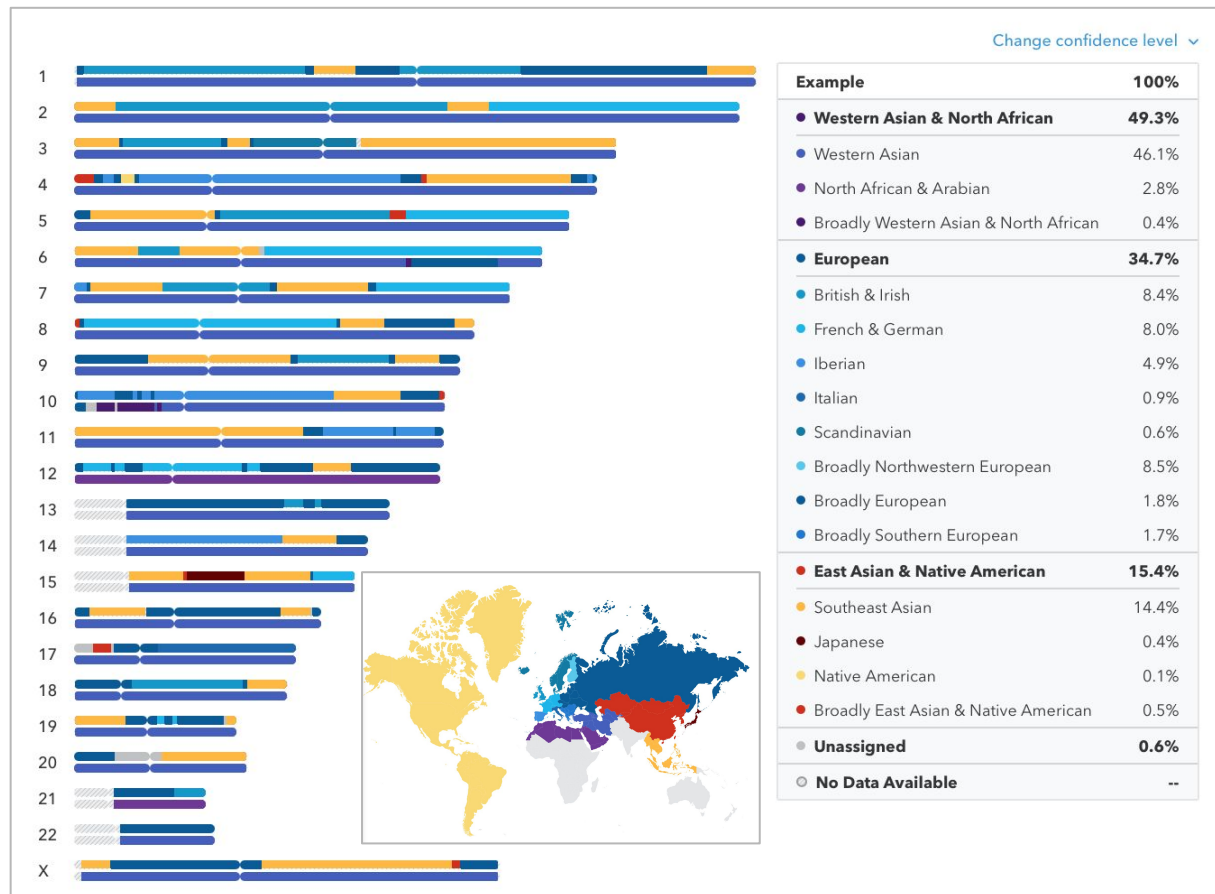


Coloriage chromosomique

Coloriage chromosomique (chromosome painting) consiste à identifier les origines géographiques de chaque morceau de chromosome d'un individu. .

Il permet ensuite de lister des pourcentages du génome provenant de différentes régions de la planète.

Noter le mélange de morceaux chromosomiques d'origines différentes.



Du génome aux génomes : génomique des populations

Concepts-clés

- Les projets initiaux de séquençage portaient sur « le » **génom de référence** pour une espèce, ce qui consistait à séquencer le génome d'un individu particulier, ou d'un petit nombre d'individus.
- **Génomique personnelle** : séquençage du génome complet d'un individu
- **Génomique des populations** : séquençage de milliers, voire de millions d'individus;
- Objectifs
 - étudier la diversité génétique de l'espèce humaine (génomique des populations),
 - identifier des gènes associés à des pathologies.

Quelques dates-clés

- 2001: génome humain de référence ("brouillon")
- 2004: génome humain de référence (complet sauf les télomères)
- 2007: séquençage massivement parallèle ("next generation sequencing")
- 2008-2015 : projet « **1000 génomes** » (au final, 2500)
 - www.internationalgenome.org
 - But: caractériser la diversité génétique des populations humaines
- 2016 : plan "France Médecine Génomique 2025" (FMG2025)
 - **Annonce** : 12 plateformes de séquençage, capacité de 200.000 génomes complets / an.
 - **En pratique**, 2 plateformes de séquençage créées, et 45.000 génomes séquencés (2026).
 - **Soin** : diagnostic individuel (**cancers + maladies rares**)
 - **Recherche** : découverte de gènes associés à ces maladies
- 2018 : Plan européen "**One plus million genomes**" ([1+MG](#))
 - Regroupement de plusieurs plans nationaux de séquençage de génomes.
- 2020: génome de référence "**de télomère à télomère**" (T2T)
- 2022 : **European Genome Data Infrastructure (GDI)**
 - Partage des données génomiques humaines
 - Focus: cancer et maladies rares

- **Site variable / variation génétique** : position ou région génomique où l'on observe des différences de séquence entre les individus d'une population.
- **Allèle** : l'une des modalités d'une variation génétique. Par exemple, un SNP peut présenter un A chez 84% (**allèle majoritaire**) des individus, et un T chez 16% (**allèle minoritaire**).
- **Variant** : terme utilisé dans 2 sens selon le contexte (ambiguïté)
 - différence par rapport à une séquence de référence
 - site ou événement de variation
- **Variant commun** : variant dont au moins 1 allèle minoritaire atteint une fréquence $\geq 5\%$ dans une population. Note: ce seuil est arbitraire.
- **Variant rare** : fréquence allélique $\leq 1\%$ dans toutes les populations étudiées. Seuil arbitraire.
- **Variant individuel / privé** : variant présent chez un seul individu unique ou une seule famille.
- **Haplotype** : groupe contigu d'allèles présent dans une même région chromosomique et généralement transmis ensemble d'une génération à la suivante.

Taux de mutations par division cellulaire

- Certaines erreurs de réplication échappent aux mécanismes de relecture et de réparation de l'ADN et deviennent des mutations permanentes, avec un taux de 1 à 3 mutations ponctuelles par division cellulaire.
- Le génotype d'un individu varie donc (très légèrement) d'une cellule à l'autre.
- La plupart de ces mutations n'ont pas d'effet détectable sur le fonctionnement cellulaire.
- Certaines peuvent perturber le comportement des cellules, et provoquer par exemple des cancers.

- Ginsburg and Willard. Genomic and personalized medicine: foundations and applications. Transl Res (2009) vol. 154 (6) pp. 277-87.
- Boulouri, H. (2010). Personal Genomics and Personalized Medicine, 1st edn. Imperial College Press.
- Topol et al. The genomics gold rush. JAMA (2007) vol. 298 (2) pp. 218-21.

Définitions: génome de référence, génome individuel et génome moyen

Génome de référence

- Le génome humain de référence est le premier qui a été séquencé par le projet public (2001 pour la version brouillon)
- Il s'agit d'un génome composite, obtenu en séquençant différents morceaux de chromosomes à partir d'une série de personnes différentes, afin de préserver en partie leur anonymat.

Génome individuel

- Séquence génomique d'un individu particulier

Génomique des populations

- Depuis 2008 un grand nombre de projets de séquençage populationnel ont démarré, qui visent à obtenir la séquence complète du génome d'un grand nombre d'individus.
- Chaque génome individuel est ensuite comparé au génome de référence pour détecter des différences (polymorphisme génétique, "variants").

Allèles majoritaire / minoritaire

- A chaque position génomique faisant l'objet d'une variation, on calcule la fréquence de chaque allèle et on définit **l'allèle majoritaire** et **le ou les allèles minoritaires** (dans certains cas on observe plus de 2 allèles au sein d'une population).

Attention : l'allèle majoritaire n'est pas forcément l'allèle de référence.

- L'allèle de référence a été établi arbitrairement sur base d'un petit nombre d'individus. Les allèles de référence dépendent des particularités de ces individus, et n'ont aucune raison d'être majoritaires dans la population (nous avons chacun certains allèles majoritaires et certains allèles minoritaires à différentes positions de notre génome).

Génome moyen d'une population

- Calculé en retenant l'allèle majoritaire à chaque position.
- Les allèles du génome moyen peuvent différer d'une population à l'autre.

Exemple de fichier de description de variants

Chaque ligne caractérise une variation

Certaines sont connues préalablement à l'étude → on indique leur identifiant.

Les informations incluent une indication de la fréquence de l'allèle alternatif dans l'échantillon analysé.

CHROM	POS	Identifiant (si variation connue) ID	Allèle de référence REF	Allèle alternatif ALT	QUAL	FILTER	INFO	Fréq. allèle alternatif
22	16050408	.	T	C	.	PASS	DP=2009;AF=0.030;CB=BI,BC;AFR_R2=0.655	
22	16050612	.	C	G	.	PASS	DP=2636;AF=0.012;CB=UM,BI,BC;EUR_R2=0.691;AFR_R2=0.718	
22	16050678	.	C	T	.	PASS	DP=2950;AF=0.010;CB=BI,BC;EUR_R2=0.501;AFR_R2=0.686	
22	16050822	.	G	A	.	PASS	DP=2779;AF=0.041;CB=UM,BI;EUR_R2=0.209;AFR_R2=0.286	
22	16050933	.	G	A	.	PASS	DP=2792;AF=0.108;CB=UM,BI,BC;EUR_R2=0.45;AFR_R2=0.641	
22	16050967	.	C	G	.	PASS	DP=1045;AF=0.030;CB=BI,BC	
22	16050994	rs7288968	G	C	.	PASS	DP=1232;AF=0.010;CB=BI,BC;AFR_R2=0.618	
22	16051107	rs6518357	C	A	.	PASS	DP=1993;AF=0.020;CB=BI,BC;AFR_R2=0.955	
22	16051209	rs7292503	A	G	.	PASS	DP=2179;AF=0.010;CB=BI,NCBI;AFR_R2=0.79	
22	16051249	rs62224609	T	C	.	PASS	DP=3184;AF=0.063;CB=UM,BI,BC,NCBI;EUR_R2=0.851;AFR_R2=0.593	
22	16051255	rs113903952	C	T	.	PASS	DP=3426;AF=0.010;CB=BI,BC,NCBI;EUR_R2=0.618;AFR_R2=0.923	
22	16051295	rs76439996	C	G	.	PASS	DP=1901;AF=0.010;CB=BI,BC;AFR_R2=0.967	
22	16051347	rs62224610	G	C	.	PASS	DP=2918;AF=0.231;CB=UM,BI,BC;EUR_R2=0.561;AFR_R2=0.667	

Grands projets de génomique (volet populationnel + médecine génomique)

Carte des projets génomiques en 2019

- Initiatives répertoriées dans 14 pays
- France: objectif initial 200.000 génomes / an
- UK : objectif initial 100.000 génomes, atteint en 2018
- Chine : objectif 100.000.000 (!) génomes

Etat des lieux 2026

- Plan France Médecine Génomique : 45.000 génomes séquencés fin 2025
- Royaume-Uni : objectif 500.000 en 2030
- 90 initiatives de génomique populationnelle dans 70 pays
- Utilisation d'intelligences artificielles pour interpréter les résultat et extraire de nouvelles connaissances

Genome Of Europe (genomeofeurope.eu)

- 27 pays, , 45M€
- Cible : 100.000 génomes (15.000 finis)
- But : caractériser la diversité génétique des citoyens Européens, et fournir un "génomme de référence" pour la médecine personnalisée

- Stark, Z. et al. Integrating Genomics into Healthcare: A Global Responsibility. Am J Hum Genet 104, 13–20 (2019). <https://doi.org/10.1016/j.ajhg.2018.11.014>
- Hazel, J.W., Scheurwater, M., de Ruijter, A. et al. Global patterns and gaps in AI policies of national genomics initiatives. npj Digit. Public Health 1, 38 (2026). doi.org/10.1038/s44482-026-00043-5

2019

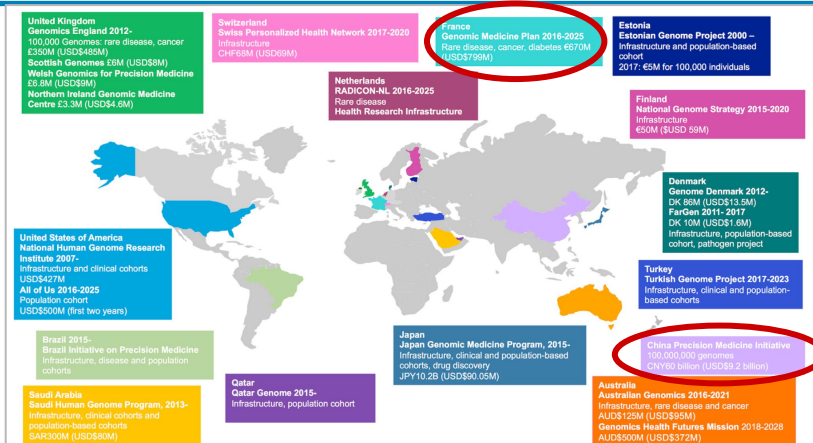
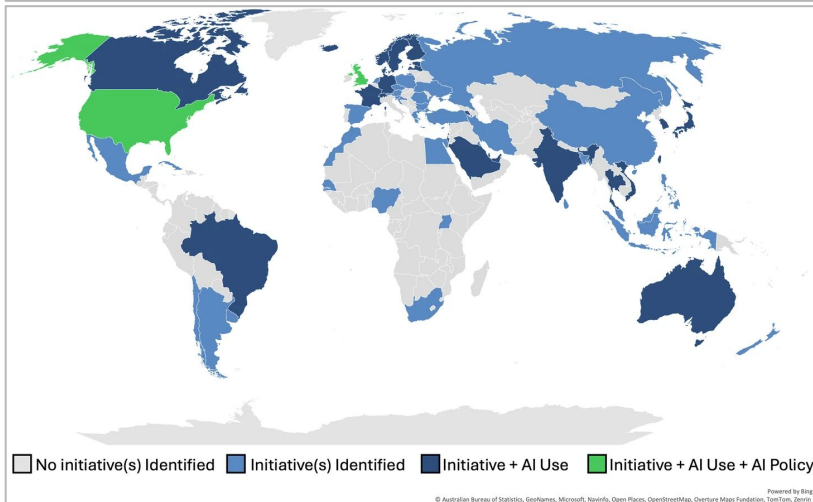


Figure 1. Map of Currently Active Government-Funded National Genomic-Medicine Initiatives

2026



The Genome Aggregation Database (gnomAD)

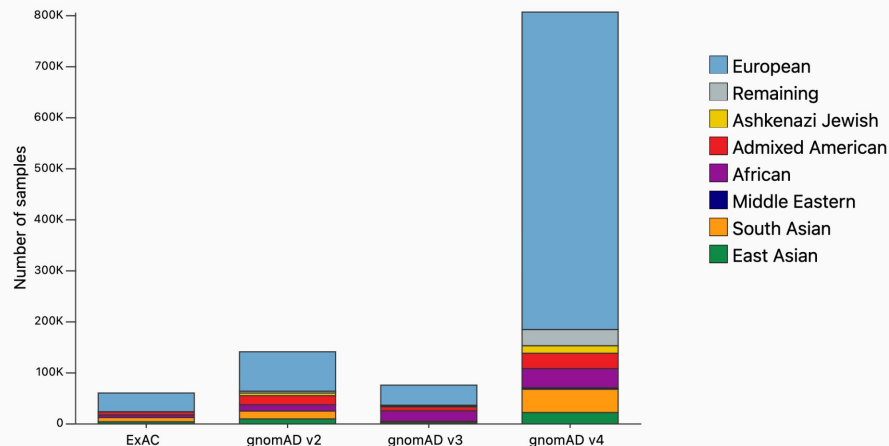
- Agrégation des données de différents projets de génomique populationnelle
- dernière version (v4): 800.000 individus
- Séquençage d'exomes (90%) et de génomes complets (10%)
- Sur-représentation d'individus d'origine européenne → représentation incomplète de la diversité génétique des populations humaines

Exome : ensemble des exons d'un organisme.

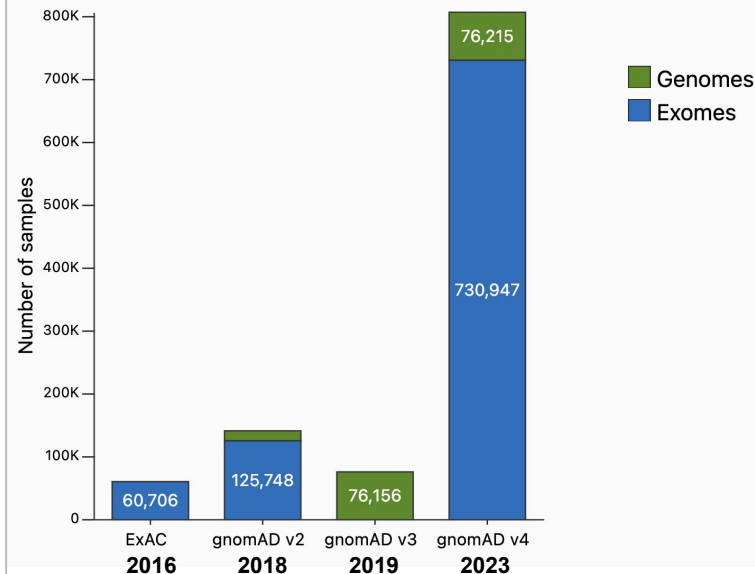
Pour des raisons de coût, certains projets de séquençage se limitent à l'exome, qui représente 2% de la longueur du génome humain.

Porjet ExAc (2016) : Exome Aggregation Consortium

Per genetic ancestry group count of samples in gnomAD releases



Sample size across major ExAC/gnomAD releases



The Genome Aggregation Database (gnomAD) : répartition des mutations

Short variants

- **Total SNPs: 786,500,648**
- Total InDels: 122,583,462
- Variant type* counts
 - Synonymous: 9,643,254
 - Missense: 16,412,219
 - Nonsense: 726,924
 - Frameshift: 1,186,588
 - Canonical splice site: 542,514

**This is only a subset of commonly asked for variant types from the dataset.*

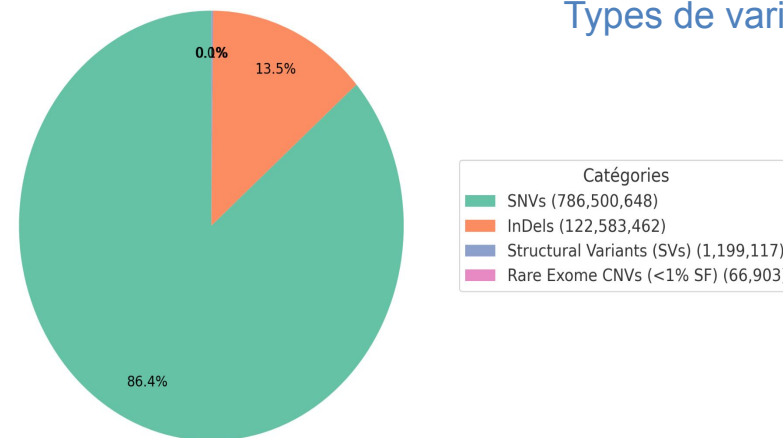
Structural variants

- 1,199,117 genome structural variants
 - 627,947 Deletions
 - 258,882 Duplications
 - 711 CNVs
 - 296,184 Insertions
 - 2,185 Inversions
 - 13,116 Complex
 - 92 Canonical reciprocal translocations
- 66,903 rare exome Copy Number Variations (<1% site frequency (SF))
 - 30,877 Deletions
 - 36,026 Duplications

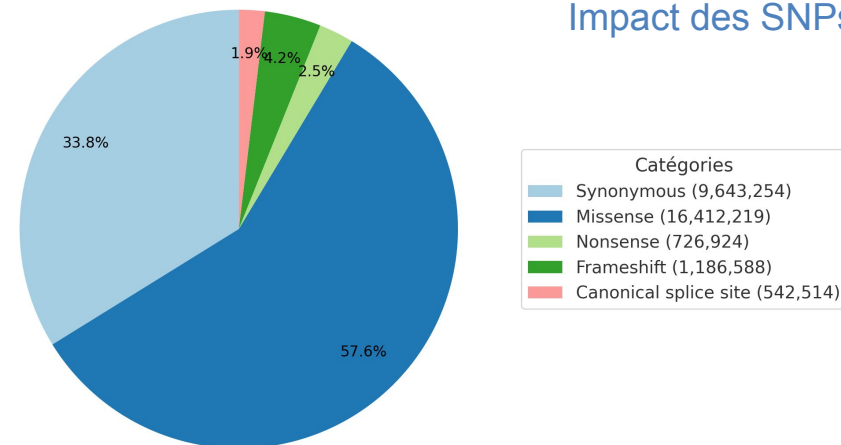
Average number of variants per person

- SNPs per person (*coming soon*)
- 1 rare (<1% SF) coding CNV per individual
- 11,844 SVs per genome

Types de variants



Impact des SNPs



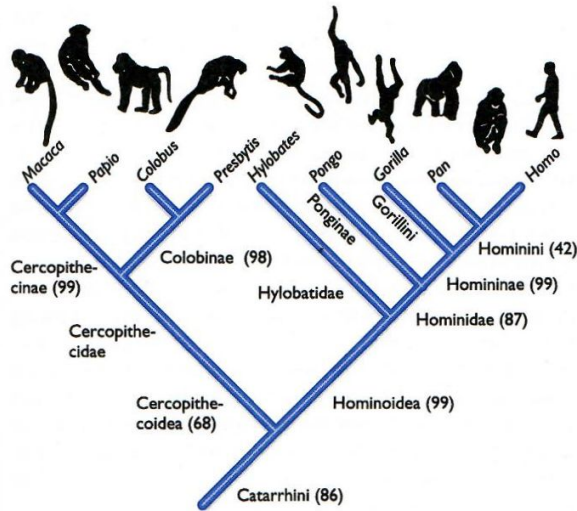
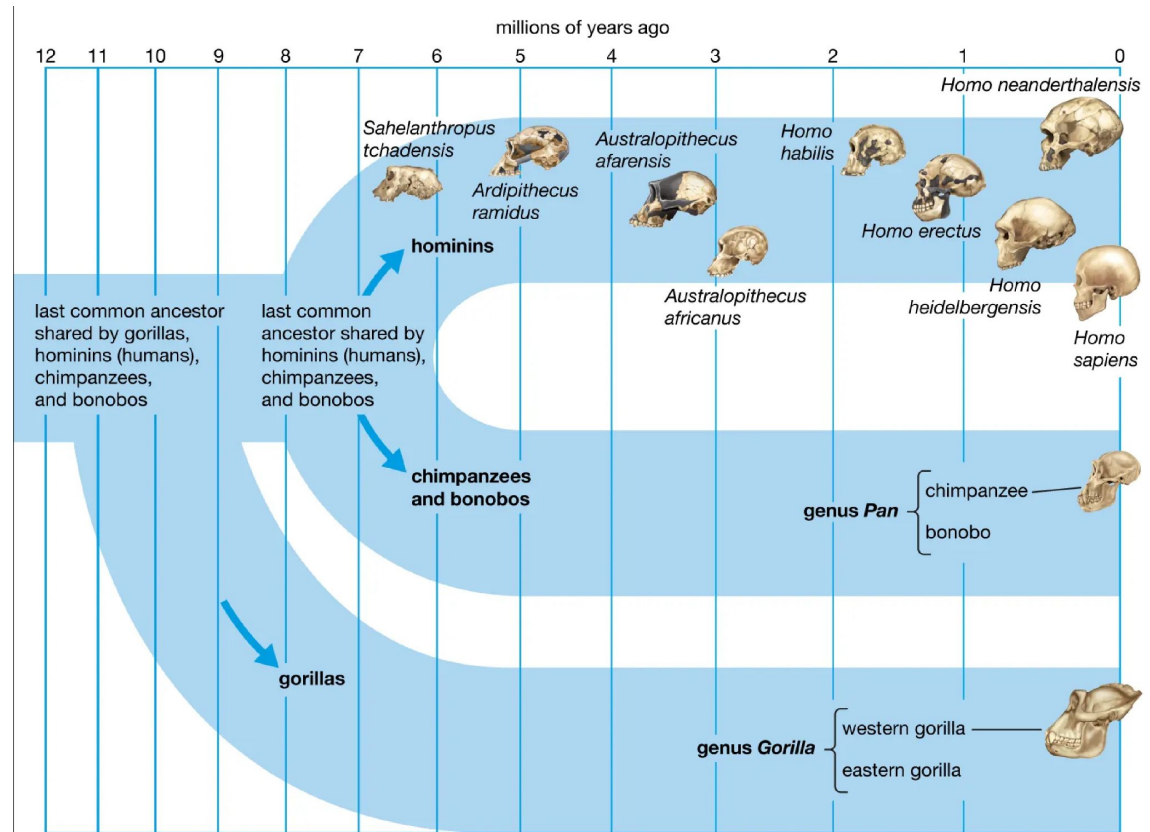
Retracer l'évolution de l'espèce humaine à partir de séquences d'ADN

Evolution des homininés

- **Hominina** = la lignée humaine (Homo, Australopithèque, ...)
- **Hominini** = humain + chimpanzé + bonobo
- **Homininae** = Hominini + gorille
- **Hominidae** = Homininae + orang-outang
- **Hominoidae** = Hominidae + gibbon

Divergences

- Hominini vs gorille : 9-11 Ma
- Humain vs chimpanzé: 6-8 Ma



© Encyclopædia Britannica, Inc.



Migrations des hominins

Les premières sorties d'Afrique (Out of Africa 1)

- La répartition géographique et la datation des fossiles indiquent plusieurs vagues de migration des différentes espèces d'hominins depuis l'Afrique vers l'Europe et l'Asie. On regroupe ces migrations sous le nom de **première sortie d'Afrique** ("Out of Africa 1", flèches plus claires pour *Homo erectus* sur la figure du haut). Les premières traces d'*Homo erectus* hors d'Afrique datent de 1.8 millions d'années, mais on pense que d'autres sorties se sont ensuite produites régulièrement.

Premières traces d'*Homo sapiens* (en Afrique)

- ~200.000 ans : premières traces fossiles de l'espèce *Homo sapiens* en Afrique. Cette espèce est donc apparue après la première sortie d'Afrique.

Deuxième sortie d'Afrique (Out of Africa 2)

- ~80.000 ans : **deuxième sortie d'Afrique** ("Out of Africa 2", flèches bleues sur la carte du haut), migration d'*Homo sapiens* vers l'Europe, puis l'Asie, puis l'Océanie.
- Le passage au continent américain (via le détroit de Béring) est ultérieur (15 à 35Ma).
- L'Homme de Neandertal (aire de répartition en orange sur la figure du bas) s'est éteint il y a ~30.000 ans.
- Homo sapiens* et *Homo neanderthalis* ont donc co-existé en Eurasie pendant plusieurs dizaines de milliers d'années

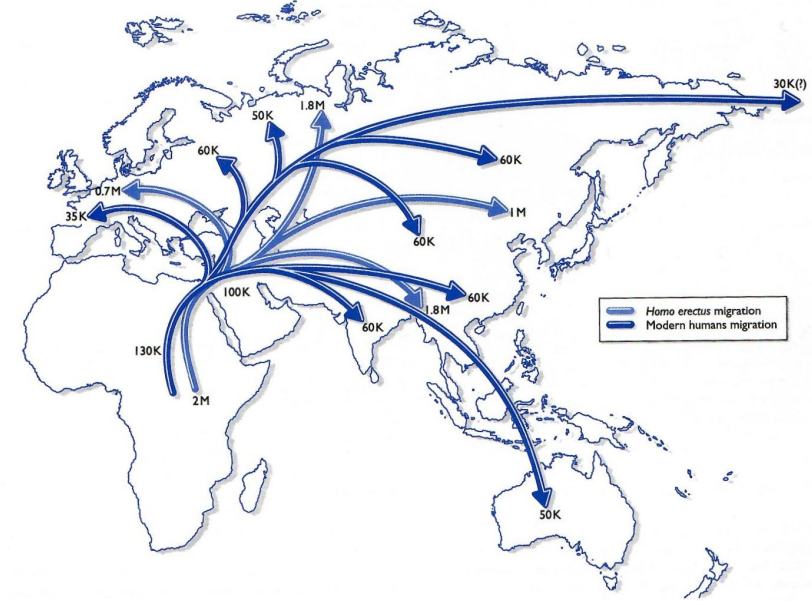
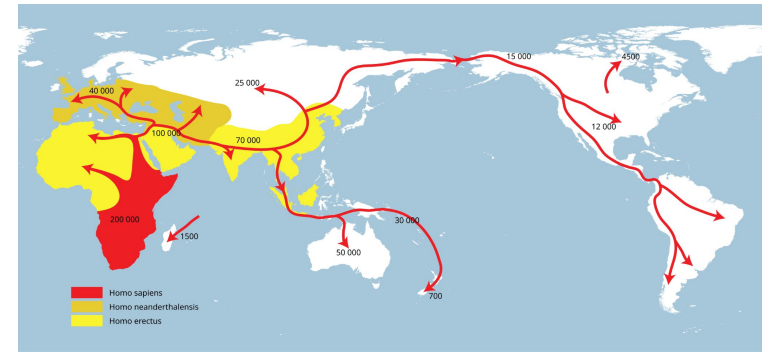


FIGURE 28.2 Two migrations: If the single, recent-origin model is correct, then the original expansion of *Homo erectus* from Africa into the rest of the Old World would have been followed much later by a similar expansion of modern people. This presentation is

certain to be an oversimplification because it implies two discrete events. In fact, multiple population movements must have occurred at different times and in different places. (K, thousand years ago; M, million years ago).



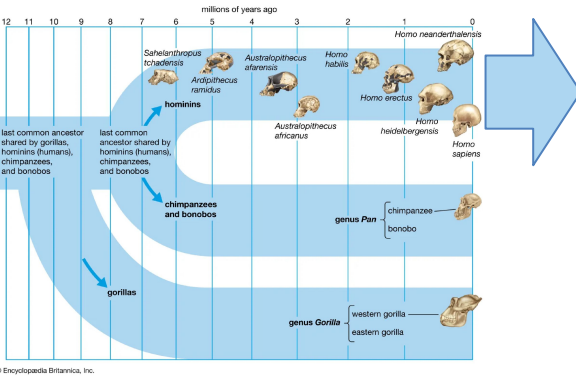
Groucutt, H. S. et al. Rethinking the dispersal of *Homo sapiens* out of Africa. *Evolutionary Anthropology* 24, 149–164 (2015). doi.org/10.1002/evan.21455

Green, R. E. et al. A Draft Sequence of the Neandertal Genome. *Science* 328, 710–722 (2010). <https://doi.org/10.1126/science.1188021>

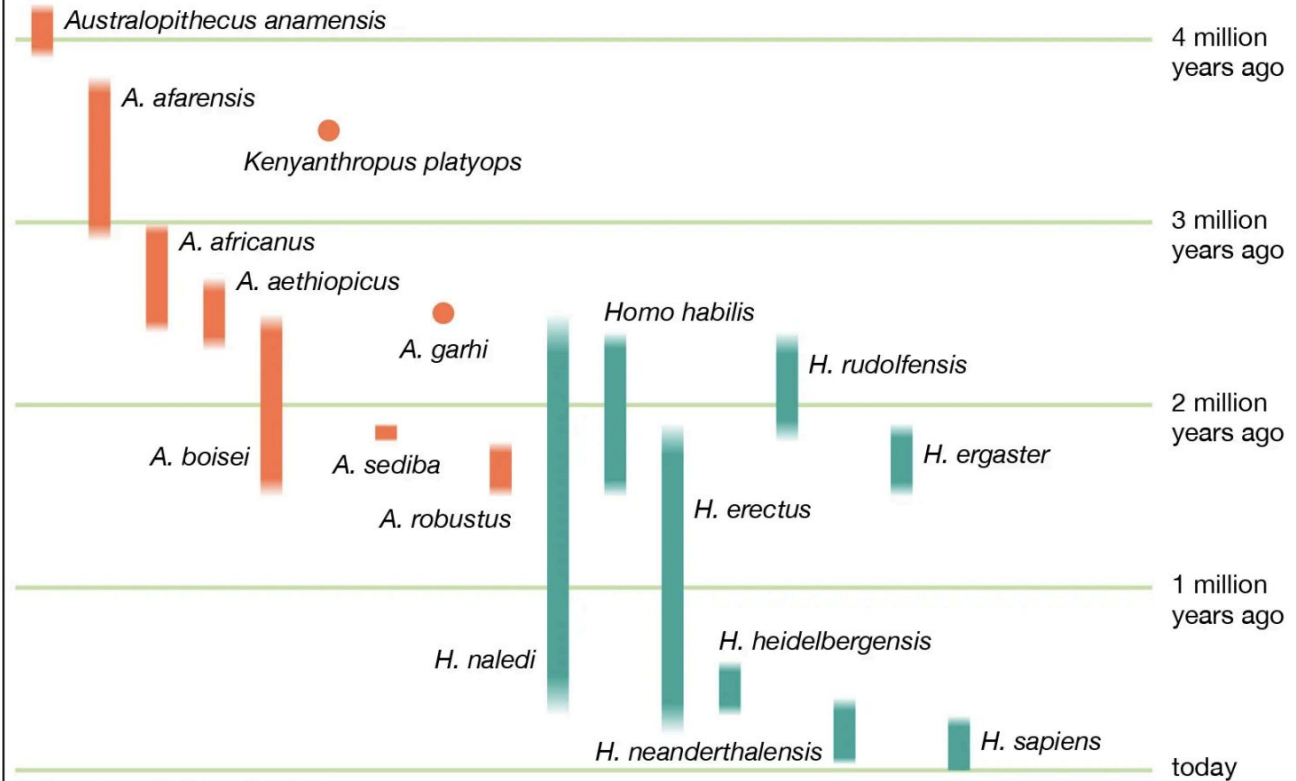
Richard Edward Green. Neanderthal Genome Project: Insights into Human Evolution (2012). <https://youtu.be/QS8bukolJTw>

Evolution de la lignée humaine

- Co-existence de plusieurs genres (*Australopithecus* et *Homo*)
- Co-existence d'*Homo sapiens* et *Homo neanderthalensis* pendant une longue période



Timeline of hominin evolution



© Encyclopædia Britannica, Inc.

Introgression de gènes néandertaliens dans le génome de l'homme moderne

- En 2010, Svante Pääbo et son équipe séquencent l'ADN de plusieurs espèces humaines fossiles, et comparent chaque région génomique à 5 génomes de référence représentatifs de différentes populations humaines actuelles.
- Cette étude révèle la présence de gènes d'origine Neandertal dans les génomes d'*Homo sapiens*, qui couvrent 1 à 4% du génome des génomes d'Européens ou d'Asiatiques actuels.
- On n'en trouve pas dans les génomes d'Africains.
- On en déduit qu'ils résultent de croisements entre Néandertaliens et Hommes modernes, qui ont eu lieu dans la région qui correspond actuellement au Moyen Orient (flèche rouge sur la figure du bas), peu après la deuxième sortie d'Afrique, et ont précédé l'expansion d'*Homo sapiens* sur l'ensemble du continent eurasien.
- Ces croisements ont provoqué une **introgression** (transfert de gènes entre espèces proches) de Neandertal à *Homo sapiens*.

Échantillons ayant servi aux premiers séquençages génomiques de Neandertal

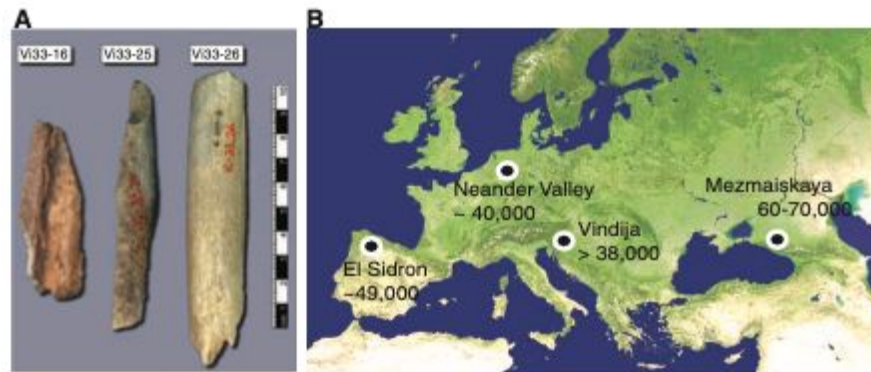
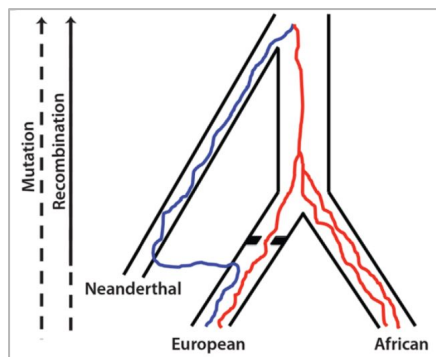
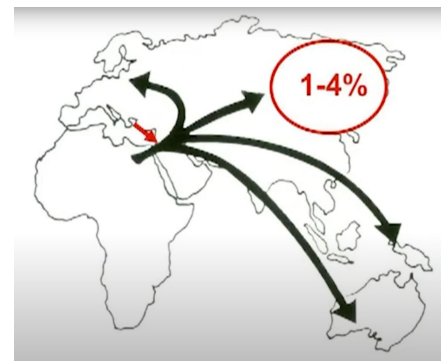


Fig. 1. Samples and sites from which DNA was retrieved. (A) The three bones from Vindija from which Neandertal DNA was sequenced. (B) Map showing the four archaeological sites from which bones were used and their approximate dates (years B.P.).

Introgression de gènes de Neandertal dans le génome d'*Homo sapiens*



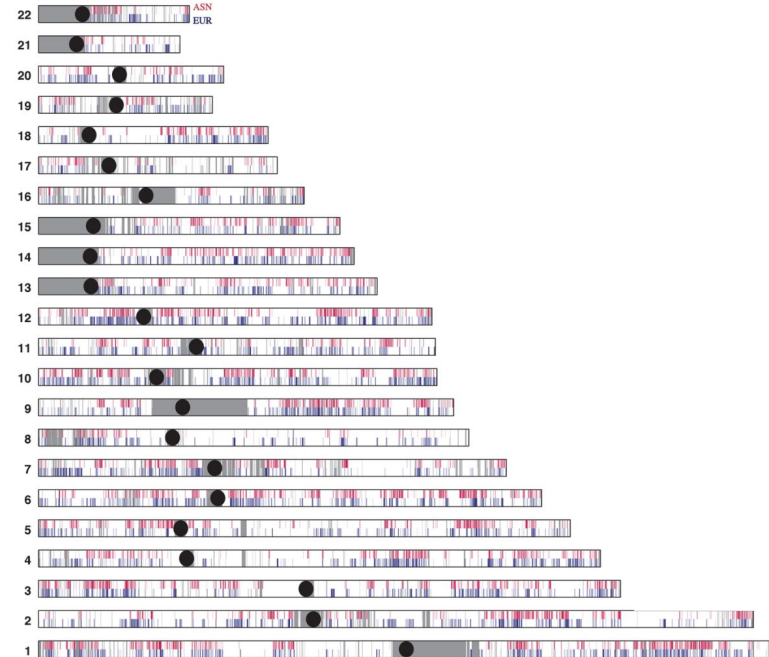
Green, R. E. et al. A Draft Sequence of the Neandertal Genome. Science 328, 710–722 (2010). <https://doi.org/10.1126/science.1188021>

Un excellent exposé par le premier auteur de cette étude : Richard Edward Green. Neanderthal Genome Project: Insights into Human Evolution (2012).

<https://youtu.be/QS8bukoLJTw>

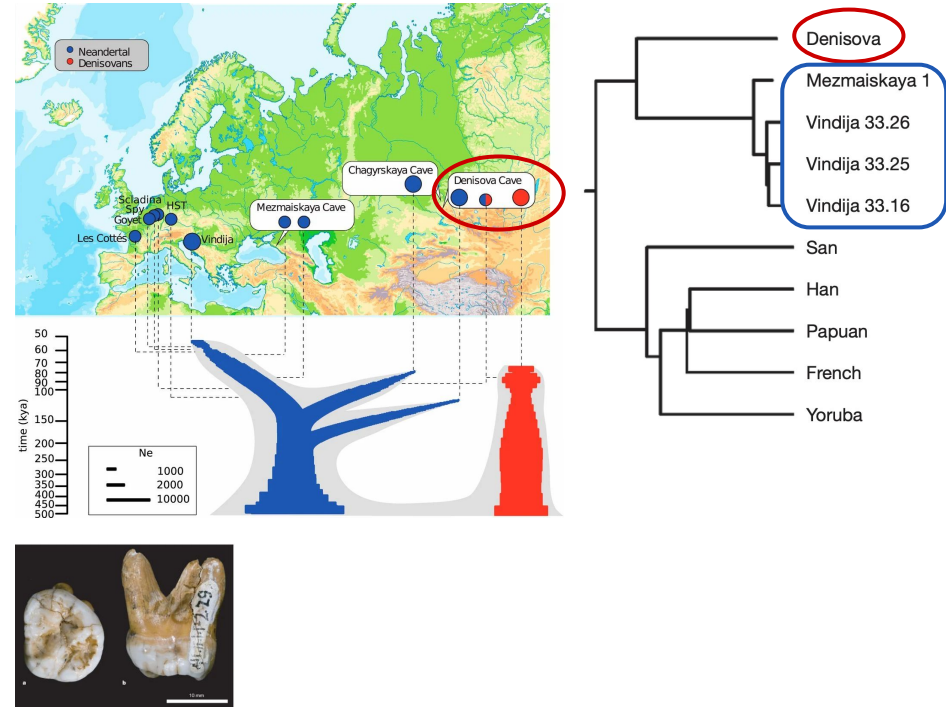
Introgression de gènes néandertaliens dans le génome de l'homme moderne

- Chaque ligne représente un chromosome humain
- Barres verticales : régions génomiques de Neandertal trouvées dans des humains modernes **asiatiques (rouge)** ou **européens (bleu)**
- Certaines régions génomiques se trouvent dans les deux groupes, d'autres spécifiquement dans l'un ou l'autre
- Zones grisées : séquençage de trop faible qualité



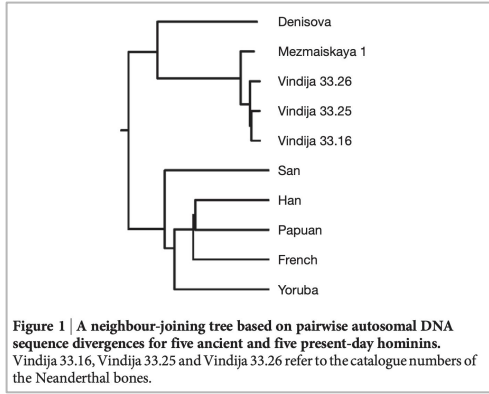
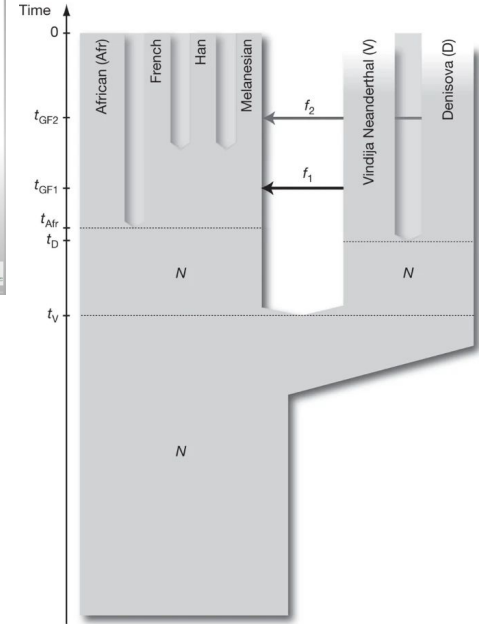
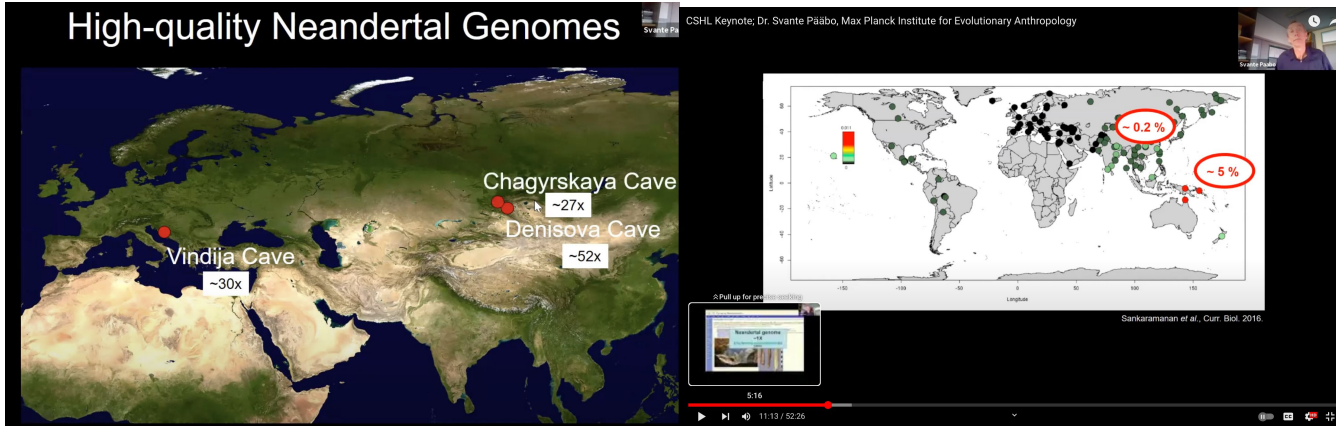
Introgression de gènes de l'homme de Denisova dans le génome d'*Homo sapiens*

- L'équipe de Svante Pääbo a également séquencé l'ADN d'échantillons fossiles de la grotte de Denisova, en Sibérie.
- Cette étude a révélé
 - une introgression *Homo denisovens* → *Homo sapiens*
 - un croisement entre *Homo denisovens* et *Homo neandertalis*



- Reich, D. et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468, 1053–1060 (2010). doi.org/10.1038/nature09710
- Mafessoni, F. et al. A high-coverage Neanderthal genome from Chagyrskaya Cave. *Proc. Natl. Acad. Sci. U.S.A.* 117, 15132–15136 (2020). doi.org/10.1073/pnas.2004944117
- Svante Pääbo. Archaic Genomics. https://youtu.be/726Wwh_P9U4?t=399

Introgression de gènes de l'homme de Denisova dans le génome d'*Homo sapiens*



Reich, D. et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468, 1053–1060 (2010). <https://doi.org/10.1038/nature09710>

Svante Pääbo, Archaic Genomes CSHL Keynote (2020). https://youtu.be/726Wwh_P9U4

Richard Edward Green. Neanderthal Genome Project: Insights into Human Evolution (2012). <https://youtu.be/QS8bukoLITw>

Fille de Neandertal et de Denisova

Cette carte génétique montre le résultat du séquençage d'un spécimen trouvé dans la grotte de Denisova.

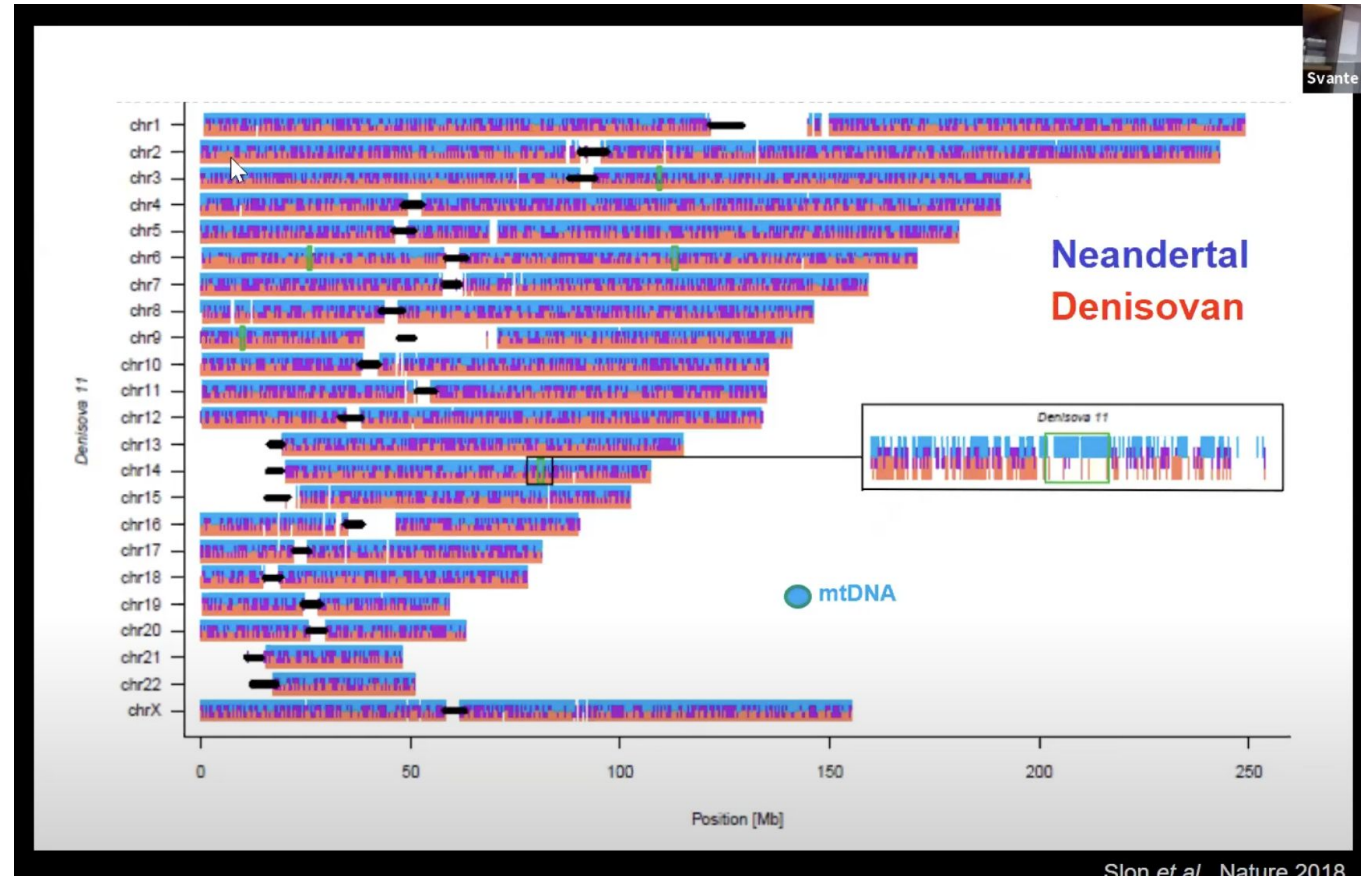
La coloration indique les morceaux de chromosome qui sont génétiquement plus proches du génome de Denisova (rouge) ou Neandertal (bleu).

Deux chromosomes X → il s'agissait d'une femme

On constate que chaque paire de chromosome comporte un exemplaire Neandertal et un Denisova.

Le chromosome mitochondrial (mtDNA) est de type Neandertal.

On conclut que cette femme avait une mère Neandertal et un père Denisovien.



Chronologie des migrations et des échanges génétiques entre groupes humains

Flèches bleues épaisses: expansion initiale

Flèches bleues continues fines: migrations préhistoriques et brassage génétique (croisements entre individus des différents groupes).

Flèches bleues pointillées: migrations et brassages génétiques historiques.

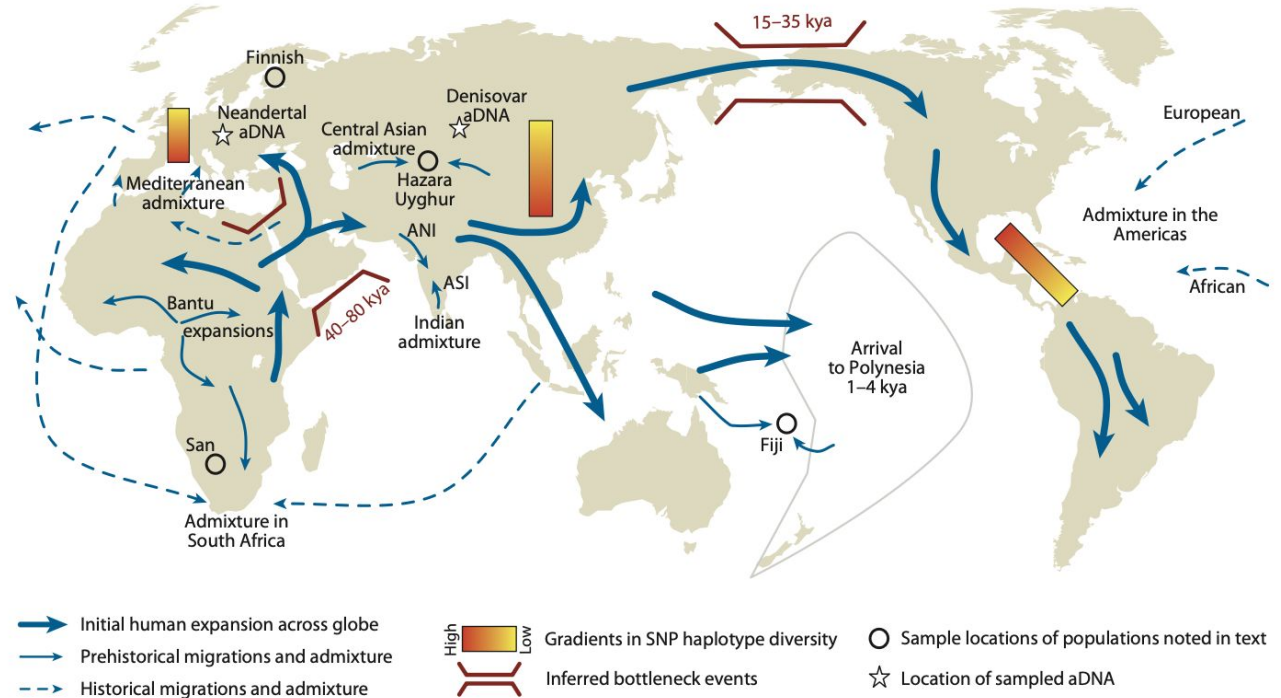


Figure 7

A schematic of human demographic history, highlighting hypotheses investigated by recent single-nucleotide polymorphism (SNP) studies discussed in this review. Numbers indicate the estimated number of years before present at which migrations took place across continents. The distinction between prehistorical and historical is in some cases approximate—the timing and duration of these events is an ongoing area of study. Abbreviations: aDNA, ancestral DNA; ANI, Ancestral North Indian; ASI, Ancestral South Indian.

Conclusions du chapitre

Génomique des populations

- Repose sur le génotypage d'un grand nombre d'individus, soit par profils génétiques basés sur des biopuces, soit par séquençage génomique. Les profils génomiques permettent de **retracer l'origine d'un individu**, sur base de la similarité de son génome avec celui d'autres individus issus de la même région. Cette analyse peut être menée de façon plus précise au niveau de chaque région chromosomique. On observe généralement des **origines diverses chez un même individu**, qui reflètent l'histoire des migrations de ses ancêtres.

Paléogénomique

- On a également séquencé les génomes d'espèces éteintes (Homme de Neanderdal, de Denisova) à partir d'os fossiles. Ces études ont démontré deux introgressions vers Homo sapiens.
 - Neandertal : 1 à 4% du génome d'Homo sapiens d'Européens et d'Asiatiques
 - Denisova : 5% dans les génomes de Mélanésien, et faible présence dans les génomes d'Asiatiques
 - Ces introgression ont eu lieu après la deuxième sortie d'Afrique, et on ne trouve donc pas ces gènes dans les génomes d'Africains.
- La génomique des populations et la paléogénomique permettent d'établir de nouveaux scénarios concernant les migrations historiques et préhistoriques des populations humaines, ou de conforter certaines hypothèses qui avaient été formulées sur base de données paléontologiques.