

# Utilisation de l'information génétique en biologie : applications et implications scientifiques et sociétales

UE de Bioinformatique : L2 2B2M - BCP - BOPE

**Maxime Bonhomme**

Laboratoire de Recherche en Sciences Végétales

## L'information génétique

variabilité génétique  
marqueurs moléculaires

## Utilisation de l'information génétique

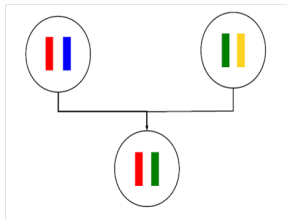
DNA fingerprinting  
association génotype / phénotype, génétique humaine  
génétique quantitative et agronomie  
génétique des populations, phylogéographie



## Qu'est-ce que l'information génétique ?

L'information génétique est l'information héréditaire portée par l'ensemble des gènes des chromosomes d'un individu.

### Hérédité mendélienne



- **locus** = segment d'ADN précisément situé dans le génome (cela peut être un gène, mais pas forcément).
- **locus polymorphe** = locus qui présente au moins deux états alléliques.
- **allèle** = version d'un gène (polymorphisme de séquence).
- **génotype** = chez un organisme diploïde, composition allélique d'un individu à un locus donné.

## Variabilité inter-individuelle de l'information génétique

**Qui dit "information génétique" dit "hérédité" mais aussi "variabilité" de l'information.**

### Mutations de l'ADN

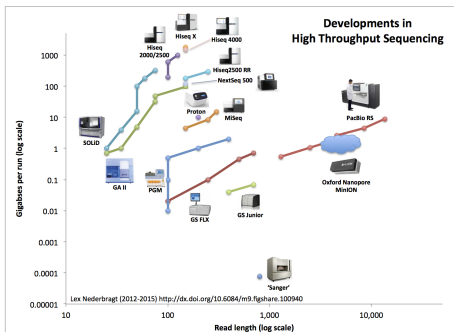
- **substitutions** :  $10^{-8}$  à  $10^{-9}$  nouvelles mutations par nucléotide (base) par génération.
- ou :  $10^{-4}$  à  $10^{-6}$  nouvelles mutations par copie de gène par génération.
- **indels** :  $\simeq 10^{-9}$  insertion/délétion d'une base par nucléotide par génération.
- réarrangements génomiques/chromosomiques : variables suivant les organismes.
- **polyploïdisation** : très fréquente chez les végétaux (beaucoup chez les espèces cultivées).

## Marqueurs moléculaires

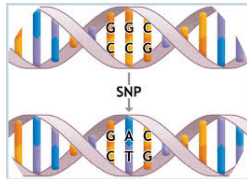
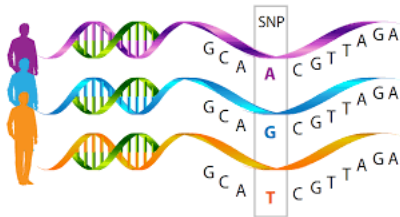
Comment capturer l'information de variabilité génétique des individus ?

### Outils de biologie moléculaire

- amplification ciblée de l'ADN (PCR)
- séquençage ou génotypage (enzymes de restrictions, séparation de fragments de tailles variables,...)
- aujourd'hui : apport des Nouvelles Technologies de Séquençage à haut-débit  
==> révolution en biologie ! donc en génétique...



## Marqueurs moléculaires

Polymorphismes de séquence : les **Single Nucleotide Polymorphisms (SNP)**

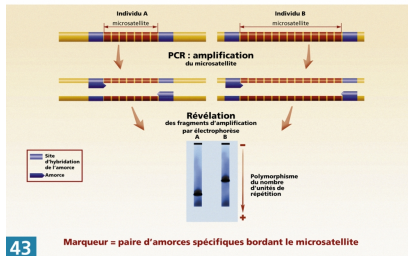
des millions de SNPs dans un génome !

## Marqueurs moléculaires

## Polymorphismes de longueur (insertions/délétions) : exemple des microsatellites (ou Short Tandem Repeats)



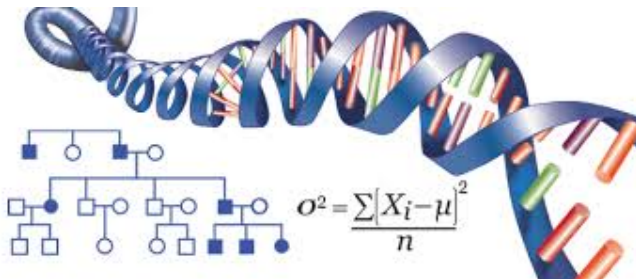
## Les marqueurs microsatellites



moins fréquents que les SNPs, mais plus polymorphes (beaucoup d'allèles)



# Utilisation de l'information génétique en biologie : implications scientifiques et sociétales





## DNA fingerprinting

### Criminologie

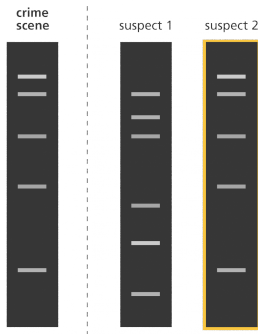


### Abby's Lab



## DNA fingerprinting

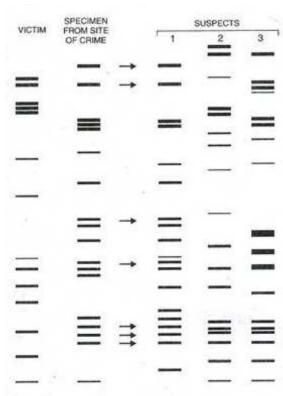
## Criminologie



- suspect 2 le plus probable, mais pas à 100% certains! (des bandes supplémentaires pourraient ne pas correspondre)
- utilisation actuelle de dizaines de marqueurs microsatellites pour affiner la probabilité de "matching"

## DNA fingerprinting

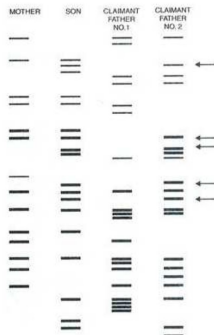
## Criminologie



Identification of a criminal through DNA fingerprinting. Suspect number one is real culprit, as its VNTR bands are matching with specimen from site of crime.

## DNA fingerprinting

### Tests de paternité/maternité





DNA fingerprints of mother under consideration, the child and two claimant fathers.  
Claimant father No. 2 proved to be the real/biological father.



### Médecine légale (identification)

## DNA fingerprinting

### Commerce illégal

**Bio Bulletin**  **BUSHMEAT TRACKED BY BARCODE**  
 SEPTEMBER 21, 2009  SCIENCE BULLETINS **AFRICA AND SOUTH AMERICA**

From each bushmeat sample, scientists at the American Museum of Natural History sequenced a fragment of DNA called a “barcode.” All individuals in a species have a similar barcode.

 **Black-fronted duiker**  
*Cephalophus nigrifrons* 

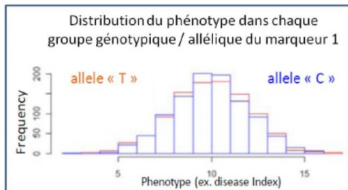
# association génotype / phénotype, génétique humaine



## Association génotype / phénotype

## Principe

individu (haploïde)	phénotype (ex: symptôme)	marqueur 1
1	11	C
2	9	C
3	10	C
4	9.5	C
5	10.3	T
6	9.8	T
..	..	.....
..	..	.....

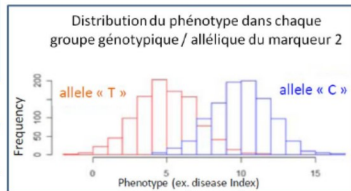


→ Les moyennes sont très proches → pas de différences phénotypiques dues au génotype/allèle → pas d'association → marqueur non causal

## Association génotype / phénotype

## Principe

individu (haploïde)	phénotype (ex: symptôme)	marqueur 2
1	11	C
2	9	C
3	10	C
4	9.5	C
5	5	T
6	4	T
..	..	..
..	..	..



→ Les moyennes sont significativement différentes → association phénotype/génotype  
→ marqueur causal (?)

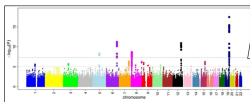
## Association génotype / phénotype

## Génétique médicale

- identifier des variants (mutations) associés à des pathologies humaines : "Genome-Wide Association Study"

*Nature* 2009 August 6, 460(7256): 744–747. doi:10.1038/nature08186.

Common variants conferring risk of schizophrenia

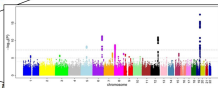


**A genome-wide association study identifies novel risk loci for type 2 diabetes**

Robert Slade<sup>1,2,4</sup>, Ghislain Rocheau<sup>1,2</sup>, John Ross<sup>4</sup>, Christian Dina<sup>4</sup>, Liuhang Shen<sup>1</sup>, David Serru<sup>1</sup>, Philippe Boutin<sup>1</sup>, Daniel Vincent<sup>1</sup>, Alexandre Boullet<sup>1</sup>, Sany Hadjilov<sup>1</sup>, Besenly Balkas<sup>1</sup>, Barbara Heude<sup>1</sup>, Guillaume Charpentier<sup>1</sup>, Thomas J. Hudson<sup>1,3</sup>, Alexandre Montpetit<sup>1</sup>, Aubrey V. Polyakovsky<sup>1</sup>, Marc Frenck<sup>1,4,5</sup>, Barry I. Posner<sup>1,2</sup>, David J. Balding<sup>1,2</sup>, David Meyre<sup>1</sup>, Constantin Polychronakos<sup>1,2</sup> & Philippe Froguel<sup>1,4</sup>

*Nat. Genet.* 2015 April; 47(4): 375–380. doi:10.1038/ng.324

Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer



*Nat. Genet.* 2012 December; 44(12): 1340–1346. doi:10.1038/ng.2462

High density genetic mapping identifies new susceptibility loci for rheumatoid arthritis

*Nat. Genet.* 2009 June; 41(6): 657–665. doi:10.1038/ng.388.

Genome-wide and fine-resolution association analysis of malaria in West Africa

## Association génotype / phénotype

## Génétique personnelle (prédictive)



## Healthy genome used to predict disease risk in later life

A study by a Stanford University scientist, published in the Lancet, is the first to use the full genome to glean information on somebody's future wellbeing



Detailed genetic analyses will allow doctors to give more targeted health advice and inform them which drugs will work best for particular patients. Photograph: AP

For the first time, doctors have used the genetic profile of an apparently healthy middle-aged man to predict his risk of developing dozens of diseases in later life.

Dr Stephen Quake, a 40-year-old scientist at Stanford University in California, was found to be carrying a rare genetic mutation that can cause a sudden and fatal heart attack, and other genes that boosted his risk of becoming diabetic and obese to more than 50%.

Article d'avril 2010

# Association génotype / phénotype

## Génétique personnelle (prédictive)

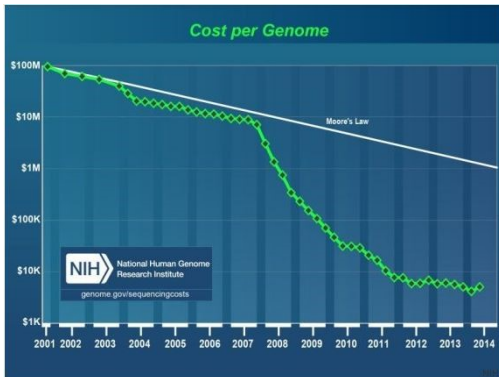
### USA

---

- [23andMe](#) (adoption, deep ancestry, ethnicity, exome sequencing, genealogy, health)
- [Counsyl](#) (inherited cancer gene screening, invasive pre-natal screening, preconception screening)
- [DNA LifeStyle Coach](#) (diet, fitness, wellness, dental and skincare, telomere length)
- [Fitness Genes](#) (fitness, health and nutrition)
- [Full Genomes Corporation](#) (whole genome sequencing)
- [Guardiome](#) (whole genome sequencing)
- [Gene by Gene](#) (research, exome sequencing, whole genome sequencing) (an amalgamation of the companies formerly known as DNA DTC and DNA Traits)
- [Genelex](#) (paternity, pharmacogenetics)
- [Genomic Express](#) (ancestry, nutrition, pharmacogenetics, sports, traits)
- [Healthspek PGT](#) (pharmacogenetics)
- [Helix](#) (to launch in 2016)
- [Interleukin Genetics](#) (health)
- [InVitaе](#) (health)
- [Kailos Genetics](#) (health, pharmacogenetics)
- [LifeNome](#) (allergies, dieting and fitness, nutrition, preventive well-being, skin care, sleep and rejuvenation)
- [Stanford Sports Genetics](#) (sports)
- [Sure Genomics](#) (health, whole genome sequencing)
- [Ubiome](#) (microbiome sequencing)
- [Veritas Genetics](#) (whole genome sequencing, BRCA, prenatal testing )

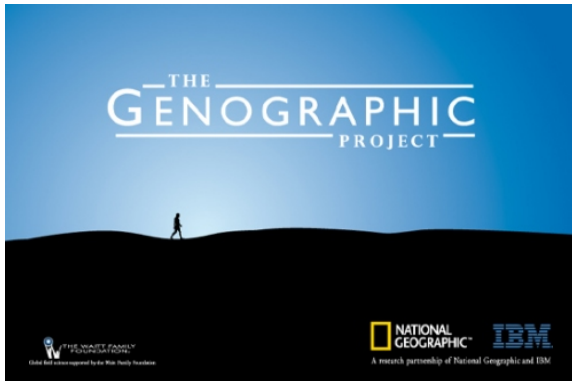


## Séquencer son propre génome



Déjà en 2014, la société Illumina affirmait être en mesure de décrypter les 3 milliards de paires qui composent le génome humain pour 1000 dollars. La différence, c'est que Veritas Genetics propose également une interprétation des résultats, afin de connaître les prédispositions génétiques d'une personne.

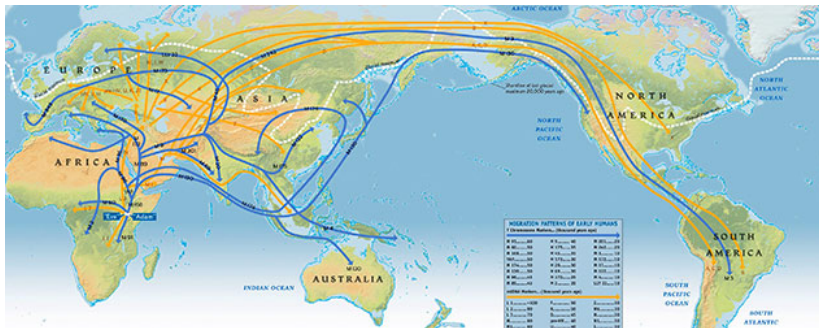
## Histoire génétique personnelle



- <https://genographic.nationalgeographic.com/>
- On achète un kit (99.95 US dollars) pour prélèvement salivaire ==> analyse ==> information DNA dans une base de données accessible.



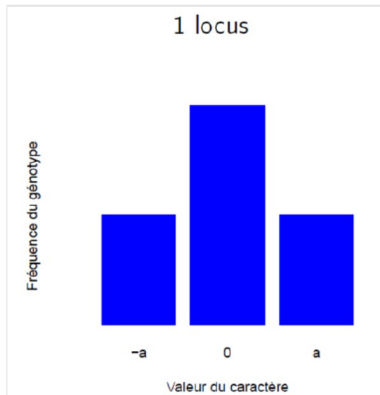
## Histoire génétique personnelle





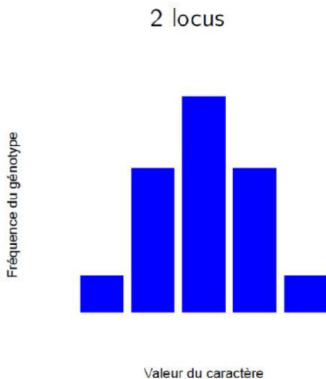
## Déterminisme génétique des caractères

Effet de plusieurs gènes sur un caractère



## Déterminisme génétique des caractères

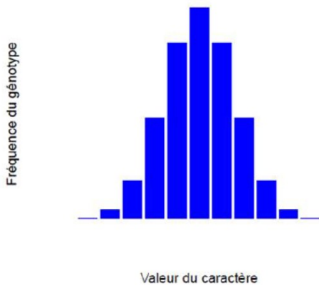
Effet de plusieurs gènes sur un caractère



## Déterminisme génétique des caractères

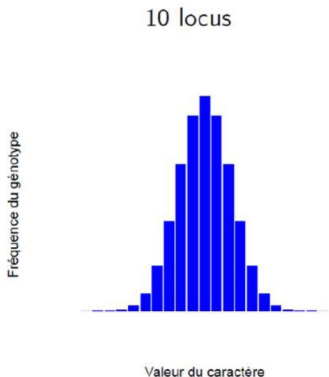
Effet de plusieurs gènes sur un caractère

5 locus



## Déterminisme génétique des caractères

Effet de plusieurs gènes sur un caractère

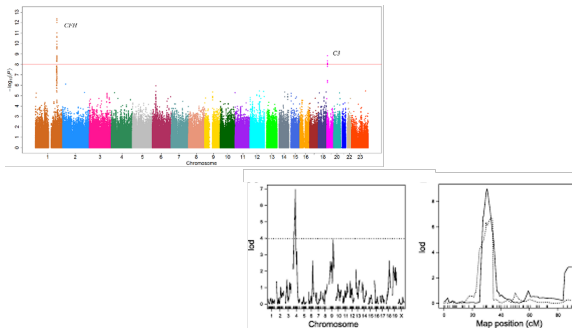


c'est le modèle polygénique infinitésimal (Fischer, 1918)

## Déterminisme génétique des caractères

### Identification des gènes

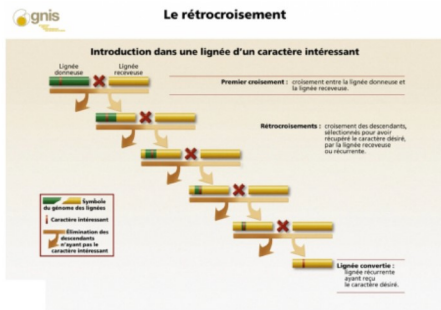
- analyses Quantitative Trait Loci (QTL)
- analyses d'association (GWAS)



- ==> sélection et amélioration génétique des animaux et des plantes, sur la base des allèles aux gènes d'intérêts identifiés

# Sélection et amélioration génétique des animaux et des plantes

## Sélection assistée par marqueurs



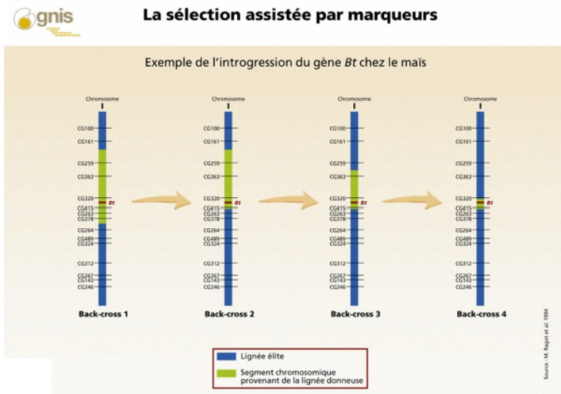
Les [rétrocroisements](#) sont utilisés pour réaliser l'introgression d'un [gène](#) dans une [variété élite](#). Mais en dépit d'un grand nombre de rétrocroisements, il reste toujours dans la lignée receveuse des segments chromosomiques du parent donneur autour du gène d'intérêt. Les [marqueurs](#) moléculaires sont souvent utilisés pour conduire l'introgression car ils permettent de réduire les temps de sélection. En effet, par sélection classique, un minimum de 7 back-cross est nécessaire afin d'obtenir un retour vers le parent récurrent de 97 %. Avec l'aide des marqueurs moléculaires, quatre rétrocroisements suffisent pour arriver au même résultat car on peut à chaque [génération](#) choisir les plantes ayant recombiné le plus petit segment chromosomique.

Les marqueurs moléculaires sont ainsi beaucoup utilisés pour les conversions de [lignées](#) pour un [transgène](#). A chaque génération, les plantes ayant récupéré le transgène sont sélectionnées sur la base de caractérisation à l'aide de marqueurs moléculaires. **Source: GNIS**



## Sélection et amélioration génétique des animaux et des plantes

## Sélection assistée par marqueurs



croisements assistés par marqueurs pour intégrer les allèles d'intérêt dans le génome d'une variété (plante), race (animaux d'élevages)

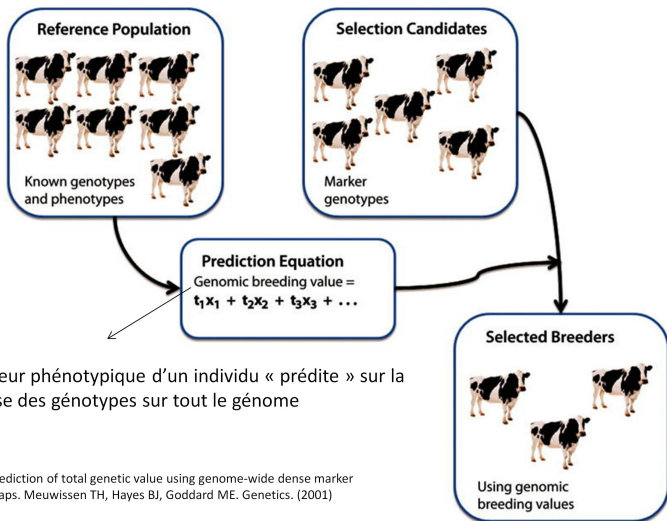
## Sélection et amélioration génétique des animaux et des plantes

### Avancée récente : sélection génomique

- bénéficie de l'apport des technologies de séquençage et génotypage haut-débit, qui permettent d'obtenir le génotype d'un individu à des milliers de marqueurs moléculaires (ex : SNP).
- on s'intéresse à prédire l'effet de milliers de marqueurs SNP sur la valeur phénotypique d'un individu.
- les généticiens/agronomes/sélectionneurs utilisent des modèles linéaires généralisés pour estimer l'effet de chaque SNP sur le phénotype.
- le but est ensuite de prédire la valeur phénotypique future d'un individu à partir de son génotypage.

# Sélection et amélioration génétique des animaux et des plantes

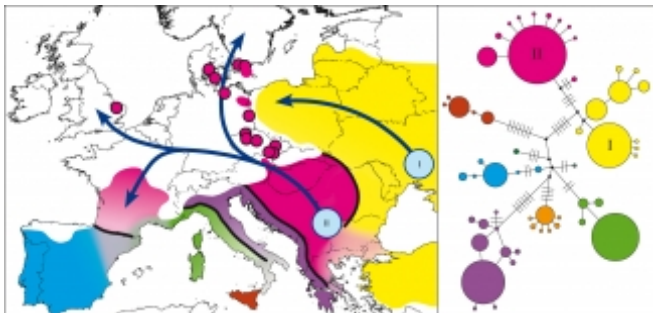
## Avancée récente : sélection génomique



# génétique des populations, phylogéographie

## Histoire démographique des populations

Retracer les événements migratoires



# Histoire démographique des populations

## Estimer la taille des populations

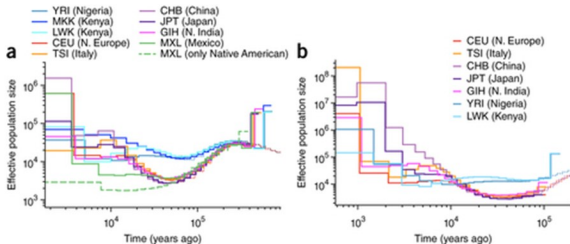
Molecular Ecology (2008) 17, 1009–1019

doi: 10.1111/j.1365-294X.2007.03645.x

### Origin and number of founders in an introduced insular primate: estimation from nuclear genetic data

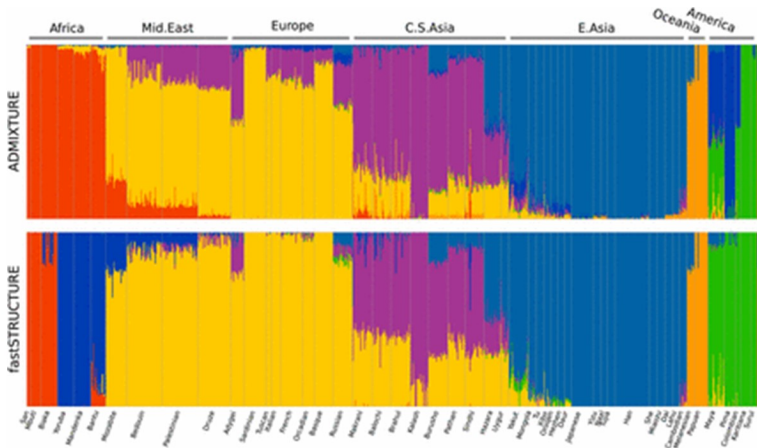
M. BONHOMME,\* A. BLANCHER,† S. CUARTERO,\* L. CHIKHI\* and B. CROUAU-ROY\*†

\*Université Paul Sabatier, 118 route de Narbonne, Laboratoire UMR CNRS 5174 'Evolution et Diversité Biologique' EDB, Bat1VR3b2, Toulouse 31062 cedex 9, France, †Faculté de Médecine de Rangueil, Bat A2, 133 route de Narbonne, Toulouse 31062 cedex 9, France



# Histoire démographique des populations

Estimer la structure des populations







## Conclusion

... une grande variété d'applications à partir de données génétiques...  
faites de la bioinformatique et de la génétique !!