



Méthodes d'exploration du génome



Solena LE SCOUARNEC, PhD

Equipe I : Génétique humaine

L'unité de recherche de l'institut du thorax

Inserm UMR 1087 / CNRS UMR 6291

Nantes, France



Human Genetics - J-J. SCHOTT / R. REDON



Ion Channels & Cardiopathies - F. CHARPENTIER / M. DE WAARD



Vascular & Pulmonary diseases - G. LOIRAND / V. SAUZEAU



Cardiometabolic diseases - B. CARIOU / C. LE MAY

TEAM I - Human Genetics

Jean-Jacques Schott

Richard Redon



Methods in biostatistics and
bioinformatics

C. DINA, A. GAIGNARD, P. LINDENBAUM AND R. REDON

Cardiac arrhythmia

JULIEN BARC AND VINCENT PROBST

Congenital heart disease

ALBAN BARUTEAU AND JEAN-JACQUES SCHOTT

Cardiac valve disease

R. CAPOULADE, S. LE SCOURNEC, J. MEROT,
J.J. SCHOTT, AND T. LE TOURNEAU

Intracranial aneurysm

ROMAIN BOURCIER AND RICHARD REDON

Inherited erythrocytosis

BETTY GARDIE

Developmental disorders

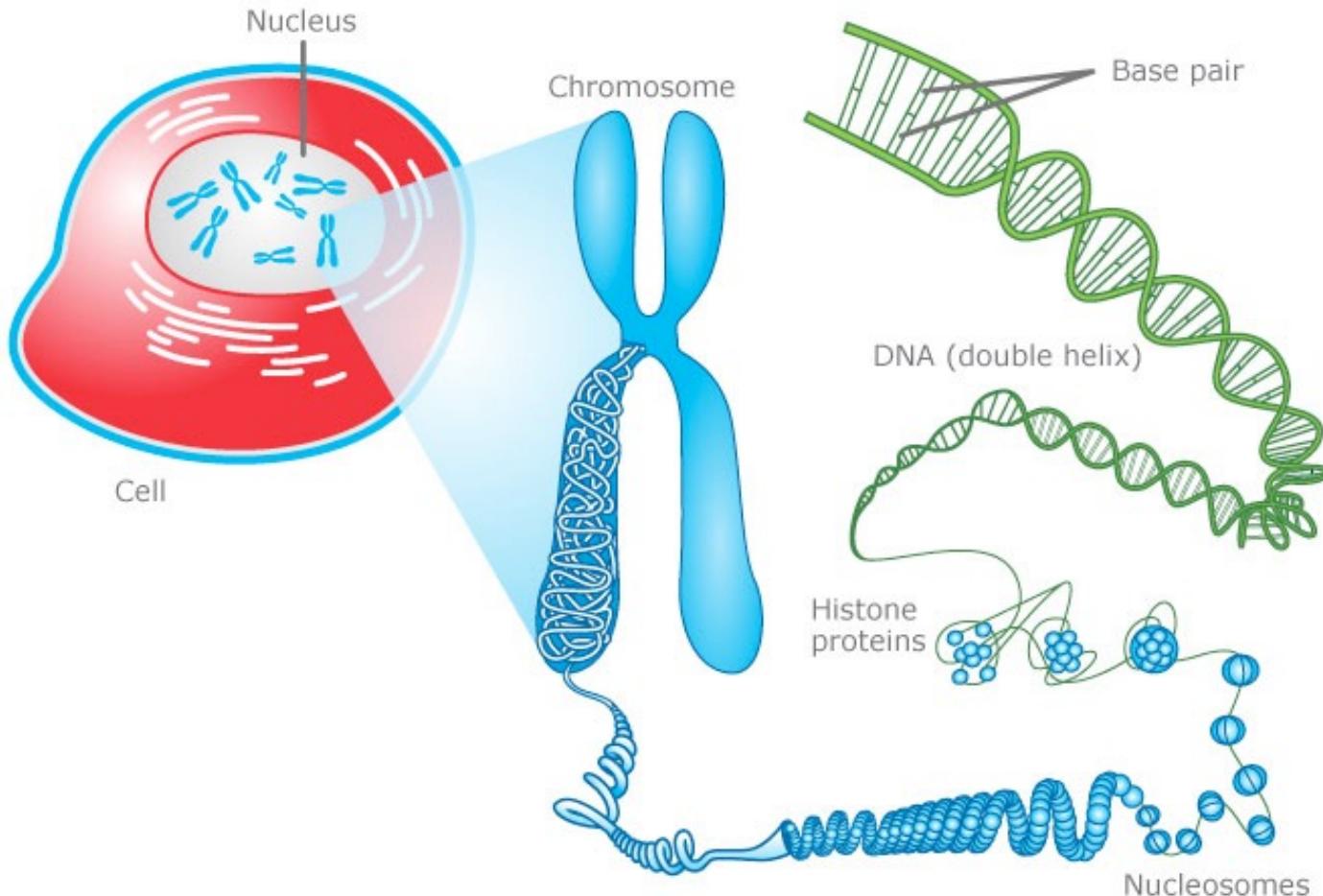
STÉPHANE BÉZIEAU

Plan

- Introduction
- Variation du génome humain
- Méthodes d'identification de gènes
 - ✓ Maladies monogéniques
 - ✓ Maladies complexes

Plan

- Introduction
- Variation du génome humain
- Méthodes d'identification de gènes
 - ✓ Maladies monogéniques
 - ✓ Maladies complexes



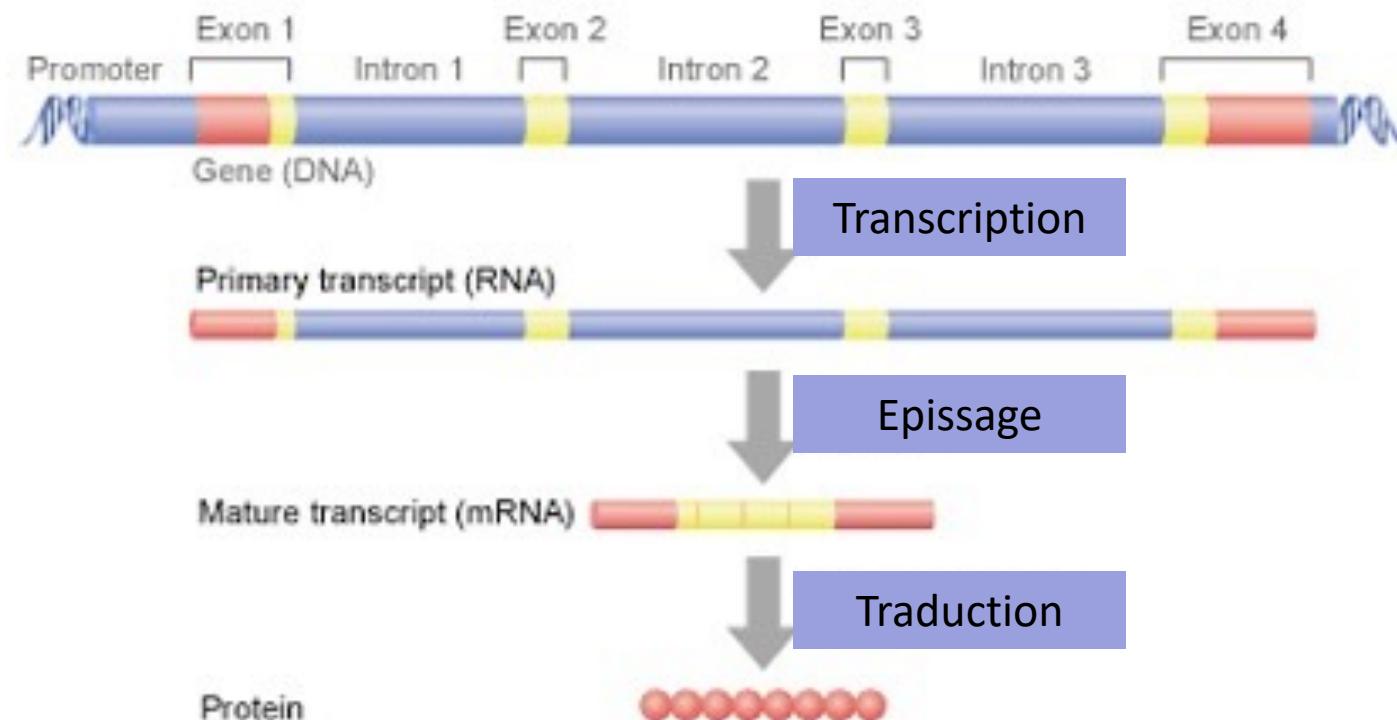
- Génome : ensemble du matériel génétique d'un organisme
- Chaque organisme possède son propre génome
- Génome humain : >3 milliards de paires de bases

Gène

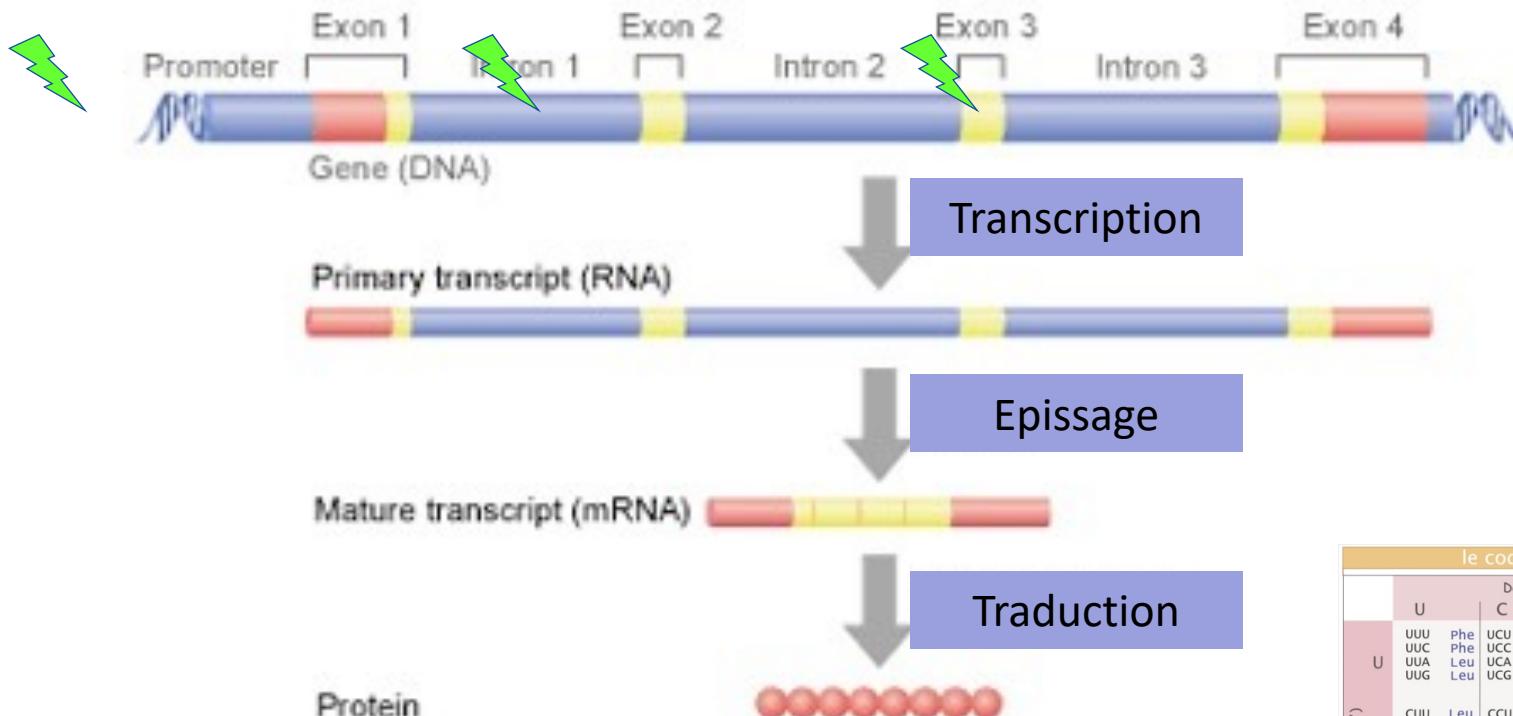
2 copies de chaque gène

20 000 – 25 000 gènes dans le génome humain

Exons / introns



Mutations génétiques



le code génétique												
	Deuxième lettre			Troisième lettre (côté 3')								
	U	C	A	G	CGU	UGC	UCA	UAG	UCG	UCC	UAC	UAA
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGC	Cys	U	C		
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C	A		
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop	A	G		
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp	G			
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U	C		
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C	A		
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	G			
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg				
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U	C		
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C	A		
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	G			
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg				
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U	C		
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C	A		
	GUA	Val	GCA	Ala	GAA	Glu	GGG	Gly				
	GUG	Val	GCG	Ala	GAG	Glu						

codon d'initiation codon de terminaison

Maladies cardiovasculaires

Maladies monogéniques

Arythmies cardiaques : syndrome du QT long, syndrome de Brugada...

Cardiomyopathies : hypertrophiques, dilatées...

Valvulopathies

Syndrome de Marfan

Hypercholestérolémie familiale

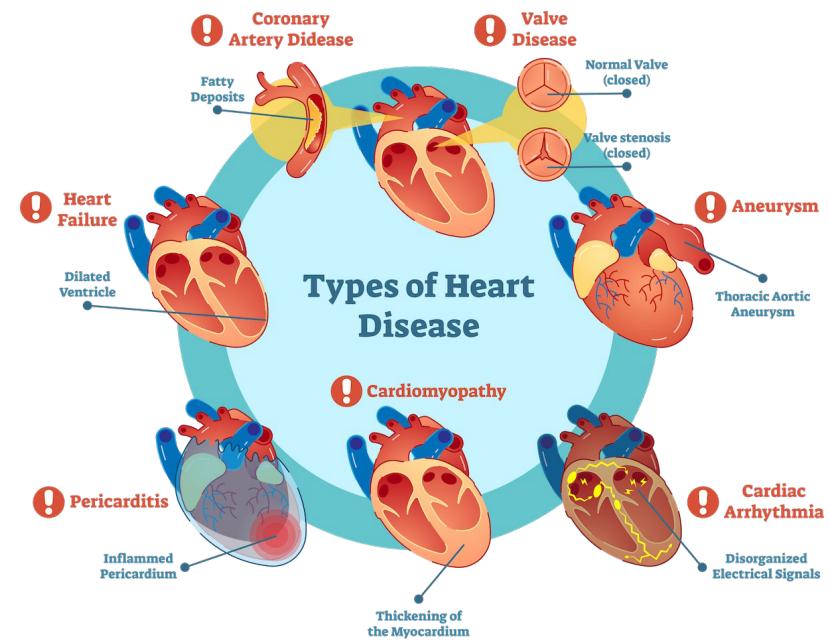
Maladies complexes

Hypertension artérielle

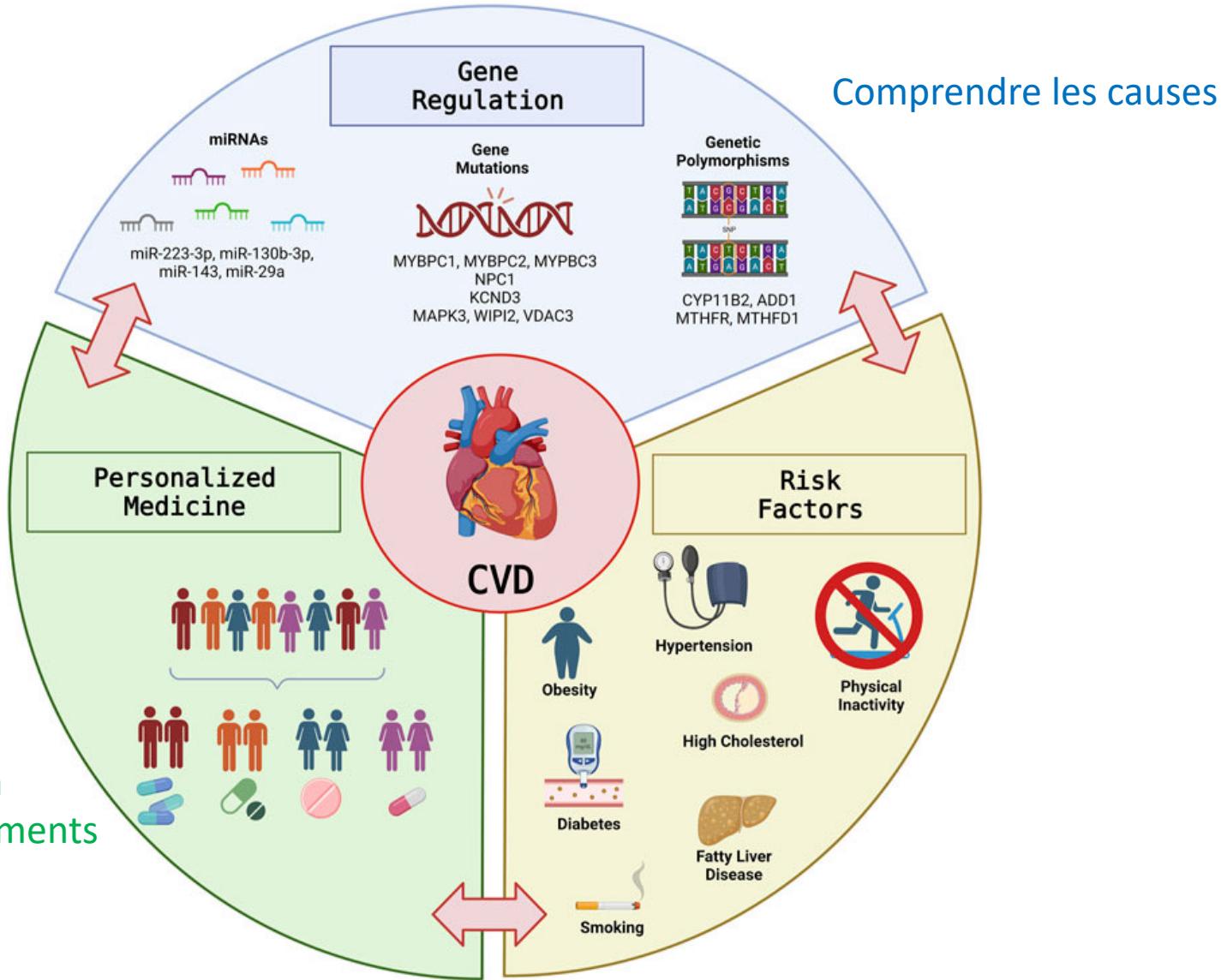
Maladie coronarienne

Fibrillation auriculaire

Asthme



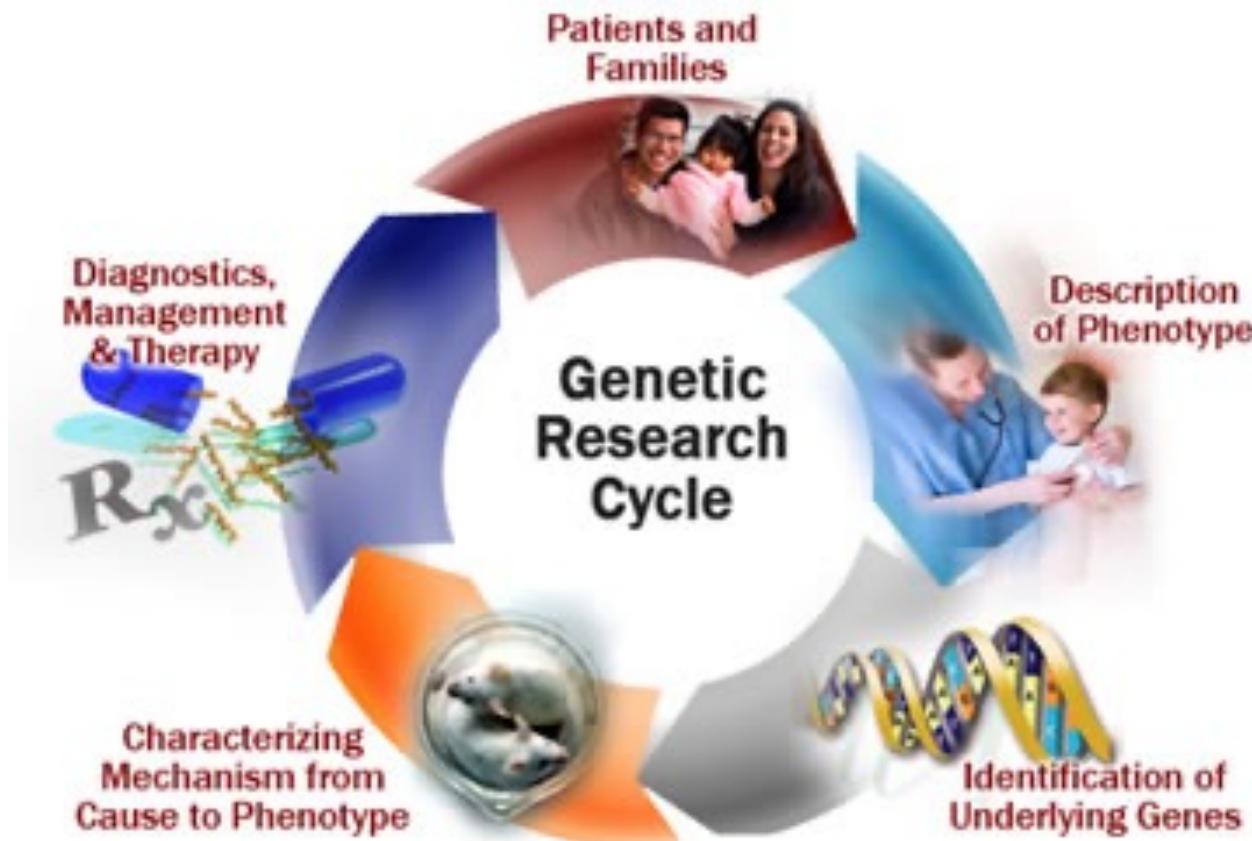
Pourquoi la génomique?

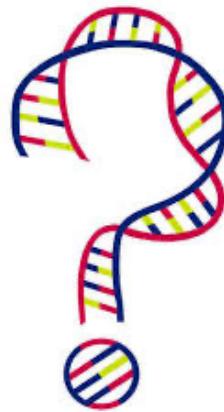


Pourquoi la génomique?

La génétique moderne a transformé de nombreuses disciplines

Discipline récente, en plein essor, et en constante évolution





La séquence du génome humain
est-elle connue ?

Projet génome humain

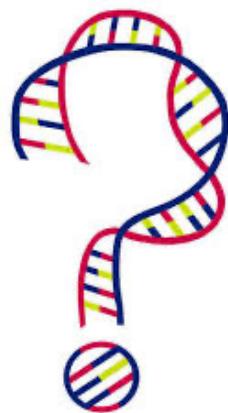


HUMAN GENOMICS

The complete sequence of a human genome

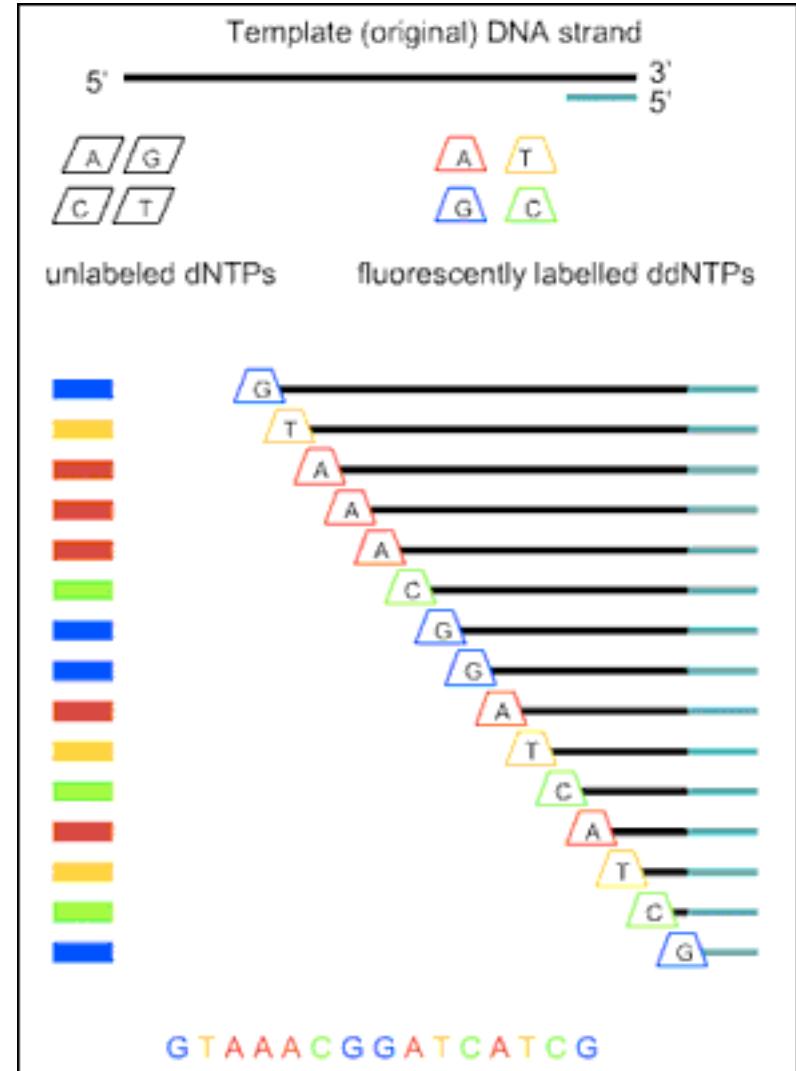
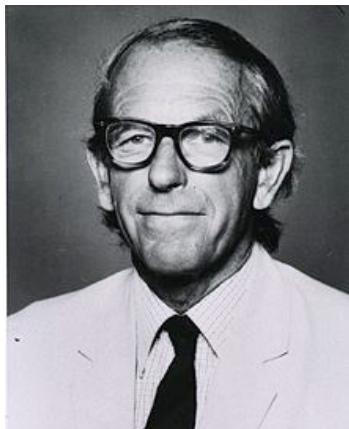
Sergey Nurk^{1†}, Sergey Koren^{1†}, Arang Rhee^{1†}, Mikko Rautiainen^{1†}, Andrey V. Bzikadze², Alla Mikheenko³, Mitchell R. Vollger⁴, Nicolas Altemose⁵, Lev Uralsky^{6,7}, Ariel Gershman⁸, Sergey Aganezov^{9†}, Savannah J. Hoyt¹⁰, Mark Diekhans¹¹, Glennis A. Logsdon⁴, Michael Alonge⁹, Stylianos E. Antonarakis¹², Matthew Borchers¹³, Gerard G. Bouffard¹⁴, Shelise Y. Brooks¹⁴, Gina V. Caldas¹⁵, Nae-Chyun Chen⁹, Haoyu Cheng^{16,17}, Chen-Shan Chin¹⁸, William Chow¹⁹, Leonardo G. de Lima¹³, Philip C. Dishuck⁴, Richard Durbin^{19,20}, Tatiana Dvorkina³, Ian T. Fiddes²¹, Giulio Formenti^{22,23}, Robert S. Fulton²⁴, Arkarachai Fungtammasan¹⁸, Erik Garrison^{11,25}, Patrick G. S. Grady¹⁰, Tina A. Graves-Lindsay²⁶, Ira M. Hall²⁷, Nancy F. Hansen²⁸, Gabrielle A. Hartley¹⁰, Marina Haukness¹¹, Kerstin Howe¹⁹, Michael W. Hunkapiller²⁹, Chirag Jain^{1,30}, Miten Jain¹¹, Erich D. Jarvis^{22,23}, Peter Kerpeljiev³¹, Melanie Kirsche⁹, Mikhail Kolmogorov³², Jonas Korlach²⁹, Milinn Kremitzki²⁶, Heng Li^{16,17}, Valerie V. Maduro³³, Tobias Marschall³⁴, Ann M. McCartney¹, Jennifer McDaniel³⁵, Danny E. Miller^{4,36}, James C. Mullikin^{14,28}, Eugene W. Myers³⁷, Nathan D. Olson³⁵, Benedict Paten¹¹, Paul Peluso²⁹, Pavel A. Pevzner³², David Porubsky⁴, Tamara Potapova¹³, Evgeny I. Rogaev^{6,7,38,39}, Jeffrey A. Rosenfeld⁴⁰, Steven L. Salzberg^{9,41}, Valerie A. Schneider⁴², Fritz J. Sedlazeck⁴³, Kishwar Shafin¹¹, Colin J. Shew⁴⁴, Alaina Shumate⁴¹, Ying Sims¹⁹, Arian F. A. Smit⁴⁵, Daniela C. Soto⁴⁴, Ivan Sovic^{29,46}, Jessica M. Storer⁴⁵, Aaron Streets^{5,47}, Beth A. Sullivan⁴⁸, Françoise Thibaud-Nissen⁴², James Torrance¹⁹, Justin Wagner³⁵, Brian P. Walenz¹, Aaron Wenger²⁹, Jonathan M. D. Wood¹⁹, Chunlin Xiao⁴², Stephanie M. Yan⁴⁹, Alice C. Young¹⁴, Samantha Zarate⁹, Urvashi Surti⁵⁰, Rajiv C. McCoy⁴⁹, Megan Y. Dennis⁴⁴, Ivan A. Alexandrov^{3,7,51}, Jennifer L. Gerton^{13,52}, Rachel J. O'Neill¹⁰, Winston Timp^{8,41}, Justin M. Zook³⁵, Michael C. Schatz^{9,49}, Evan E. Eichler^{4,53*}, Karen H. Miga^{11,54*}, Adam M. Phillippy^{1*}

Since its initial release in 2000, the human reference genome has covered only the euchromatic fraction of the genome, leaving important heterochromatic regions unfinished. Addressing the remaining 8% of the genome, the Telomere-to-Telomere (T2T) Consortium presents a complete 3.055 billion-base pair sequence of a human genome, T2T-CHM13, that includes gapless assemblies for all chromosomes except Y, corrects errors in the prior references, and introduces nearly 200 million base pairs of sequence containing 1956 gene predictions, 99 of which are predicted to be protein coding. The completed regions include all centromeric satellite arrays, recent segmental duplications, and the short arms of all five acrocentric chromosomes, unlocking these complex regions of the genome to variational and functional studies.



Quelles sont les différentes méthodes de
séquençage de l'ADN ?

Séquençage Sanger



Nouvelles technologies de séquençage

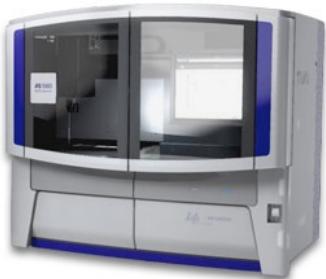
2005-2013



Pyrosequencing

life
technologies™

ThermoFisher
SCIENTIFIC



*Sequencing
by ligation*

2006-

illumina®



*Sequencing
by
synthesis*

2011-

PB

PACIFIC
BIOSCIENCES™

*Single molecule
via dye labels
Long reads*

2015-

Oxford
NANOPORE
Technologies™

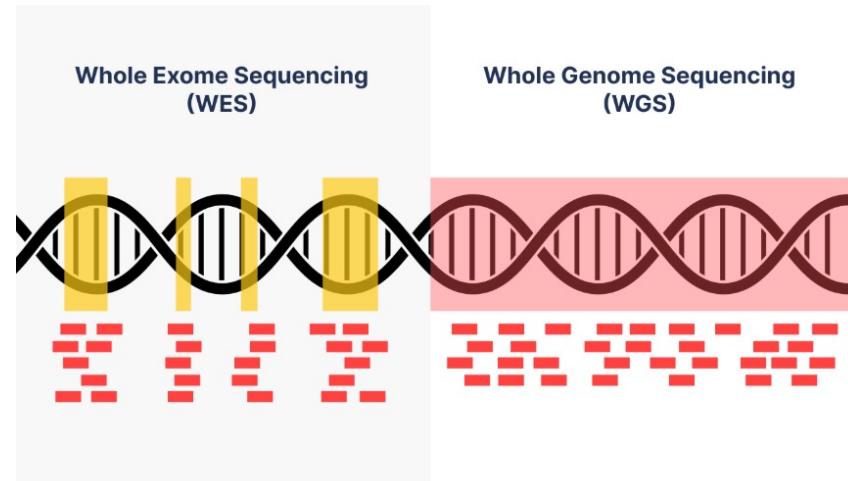


*Electronic nanopore sensing
No dye labels*

Nouvelles technologies de séquençage

Next-generation sequencing (NGS) = Séquençage massif = Séquençage haut-débit

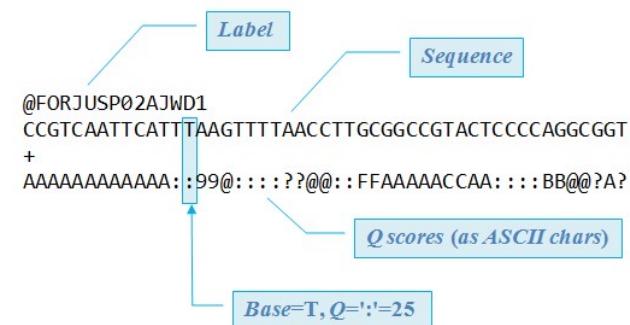
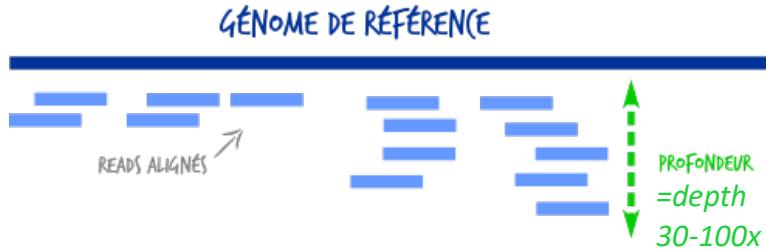
- Séquençage de génomes
- Séquençage ciblé (panels de gènes, exomes)
- Transcriptome
- Epigénétique



Technologies

- Short-read (Illumina) ~50-300 pb
- Long-read (ONT, PacBio) 10-100 kb

} Substitutions / indels
} Variants structuraux / assemblage *de novo*



Séquençage de génomes personnels

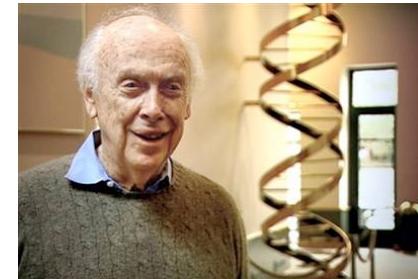
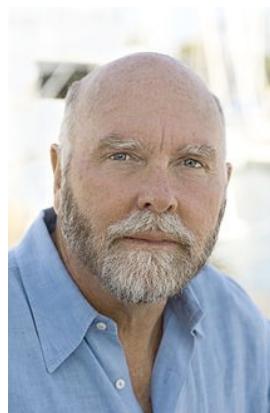
QUICKER, SMALLER, CHEAPER

Sanger

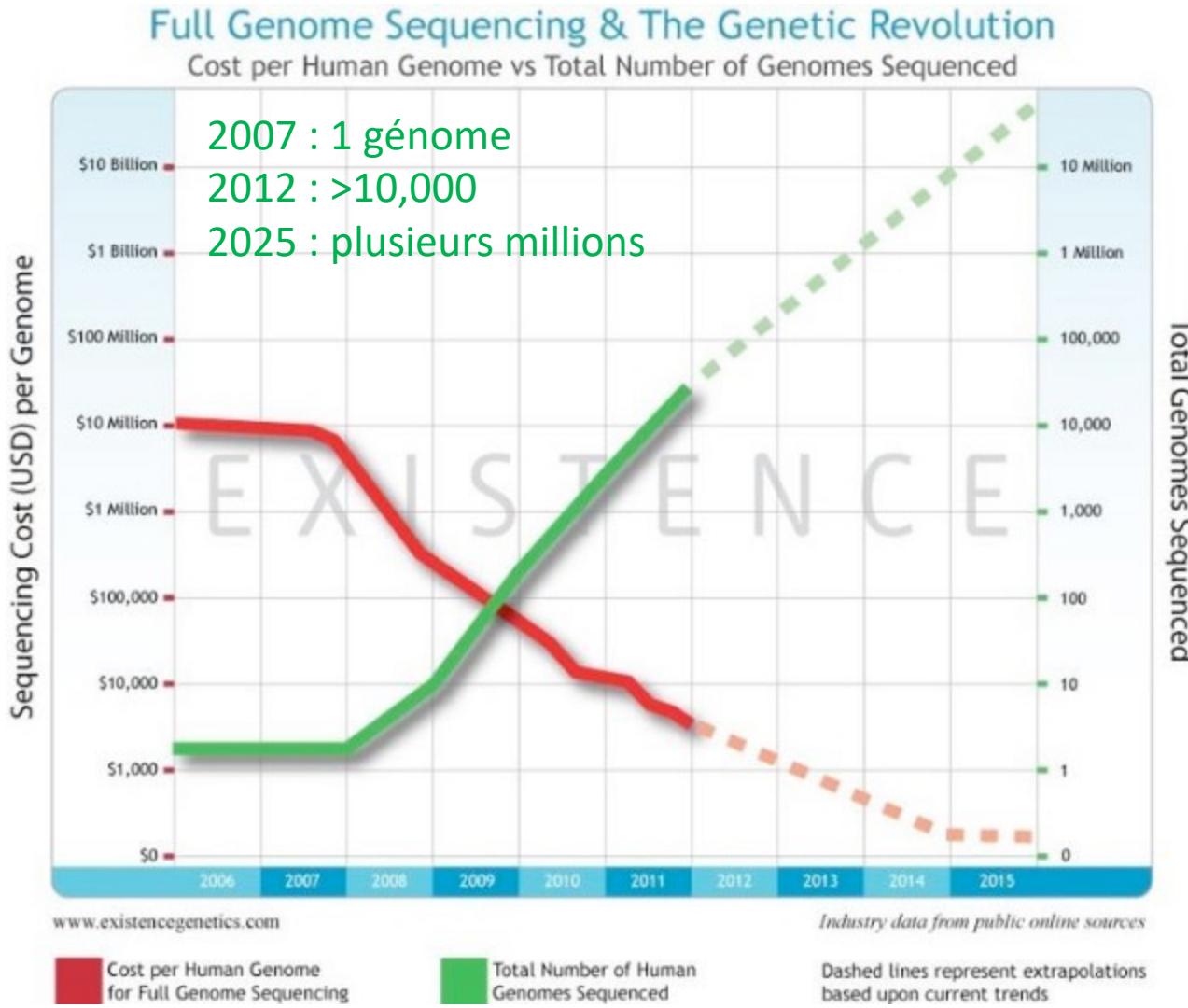
Sanger

454

Genome sequenced (publication year)	HGP (2003)	Venter (2007)	Watson (2008)
Time taken (start to finish)	13 years	4 years	4.5 months
Number of scientists listed as authors	> 2,800	31	27
Cost of sequencing (start to finish)	\$2.7 billion	\$100 million	< \$1.5 million
Coverage	8-10 ×	7.5 ×	7.4 ×
Number of institutes involved	16	5	2
Number of countries involved	6	3	1

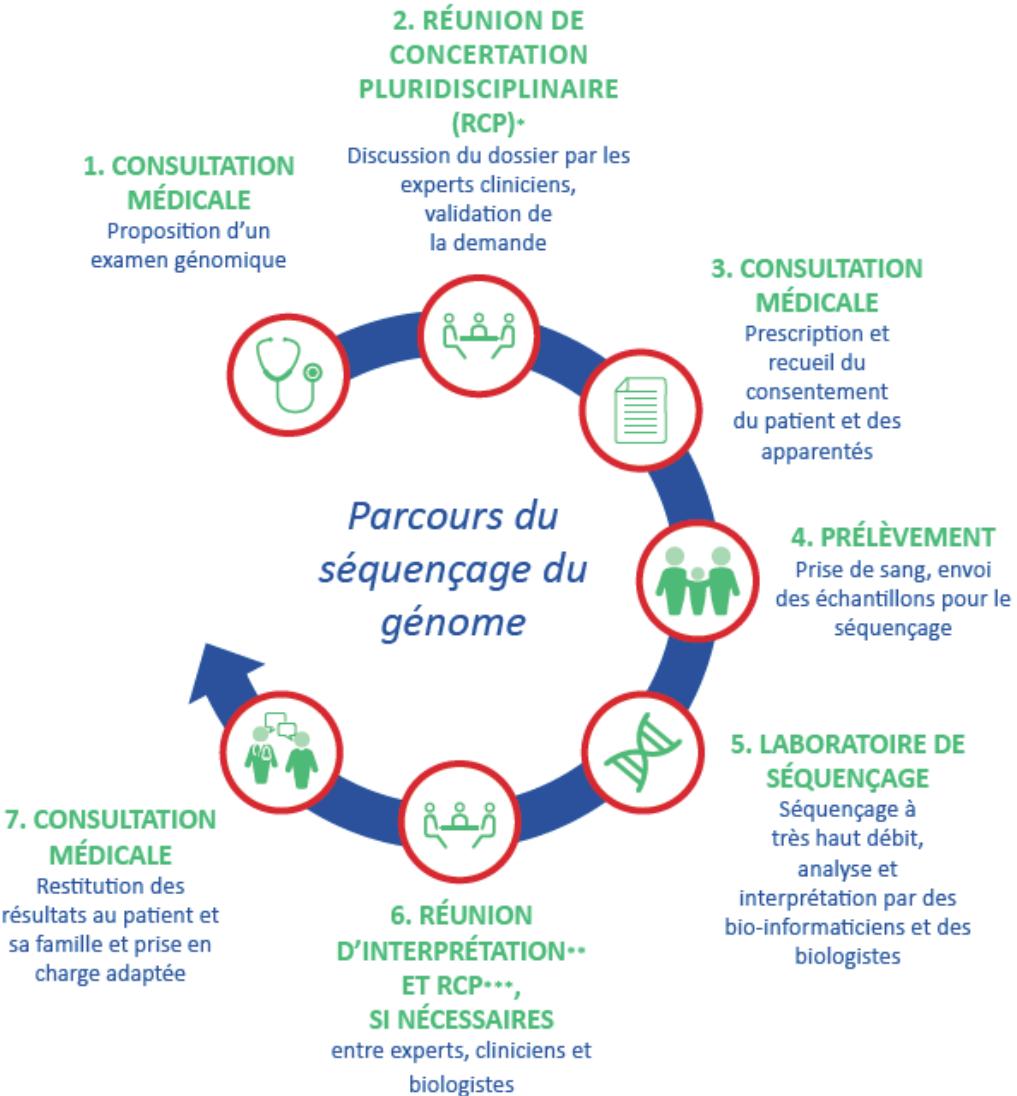


Séquençage de génomes personnels



Séquençage de génomes personnels

FRANCE MEDECINE GENOMIQUE 2025



Séquençage de génomes personnels

United Kingdom
Genomics England 2012-
100,000 Genomes: rare disease, cancer
£350M (USD\$485M)
Scottish Genomes £6M (USD\$8M)
Welsh Genomics for Precision Medicine
£6.8M (USD\$9M)
Northern Ireland Genomic Medicine
Centre £3.3M (USD\$4.6M)

Switzerland
Swiss Personalized Health Network 2017-2020
Infrastructure
CHF68M (USD69M)

France
Genomic Medicine Plan 2016-2025
Rare disease, cancer, diabetes. €870M
(USD\$799M)

Estonia
Estonian Genome Project 2000 –
Infrastructure and population-based cohort
2017: €5M for 100,000 individuals

Netherlands
RADICON-NL 2016-2025
Rare disease
Health Research Infrastructure

Finland
National Genome Strategy 2015-2020
Infrastructure
€50M (\$USD 59M)

Denmark
Genome Denmark 2012-
DK 86M (USD\$13.5M)
FarGen 2011- 2017
DK 10M (USD\$1.6M)
Infrastructure, population-based cohort, pathogen project

Turkey
Turkish Genome Project 2017-2023
Infrastructure, clinical and population-based cohorts

Japan
Japan Genomic Medicine Program, 2015-
Infrastructure, clinical and population-based cohorts, drug discovery
JPY10.2B (USD\$90.05M)

China Precision Medicine Initiative
100,000,000 genomes
CNY60 billion (USD\$9.2 billion)

United States of America
National Human Genome Research Institute 2007-
Infrastructure and clinical cohorts
USD\$427M
All of Us 2016-2025
Population cohort
USD\$500M (first two years)

Brazil 2015-
Brazil Initiative on Precision Medicine
Infrastructure, disease and population cohorts

Saudi Arabia
Saudi Human Genome Program, 2013-
Infrastructure, clinical cohorts and population-based cohorts
SAR300M (USD\$80M)

Qatar
Qatar Genome 2015-
Infrastructure, population cohort

Australia
Australian Genomics 2016-2021
Infrastructure, rare disease and cancer
AUD\$125M (USD\$95M)
Genomics Health Futures Mission 2018-2028
AUD\$500M (USD\$372M)

Le génome humain

- 3,3 milliards de paires de bases
- 23 paires de chromosomes
- 20-25 000 gènes et ~6 isoformes par gène
- ~1-2 % du génome ⇔ **exome** (régions codantes)



1 génome
2 jours
<1000 euros
150 Go

If you spoke
one 'letter' of DNA
per second

A, T, C, G, T, A, A, C, G, T...

24 hours a day,

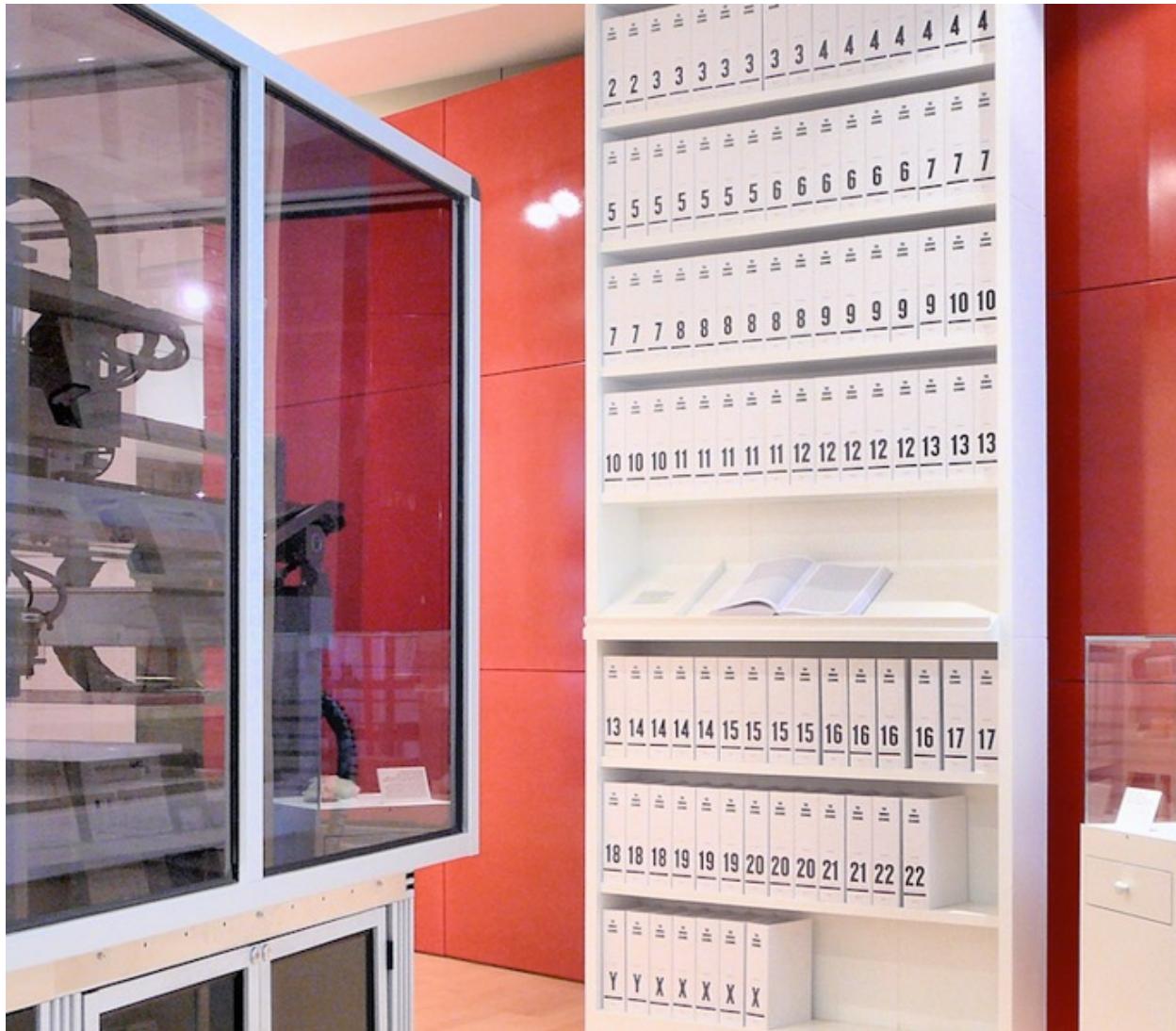
it would take about

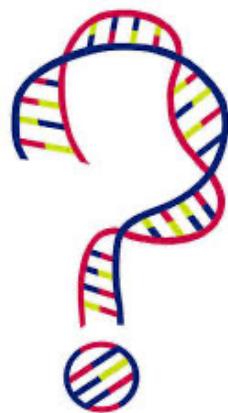
100 years to recite the entire
human genome



Plan

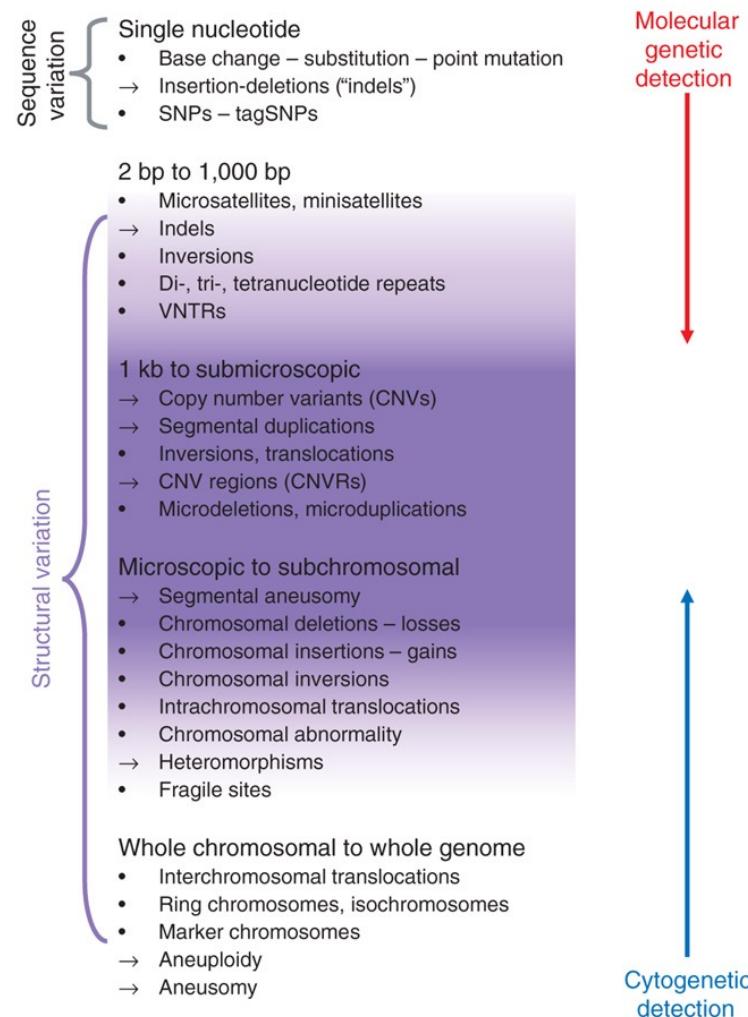
- Introduction
- Variation du génome humain
- Méthodes d'identification de gènes
 - ✓ Maladies monogéniques
 - ✓ Maladies complexes





Quels sont les différents types de variants
du génome humain ?

Types de variations



Variation du génome

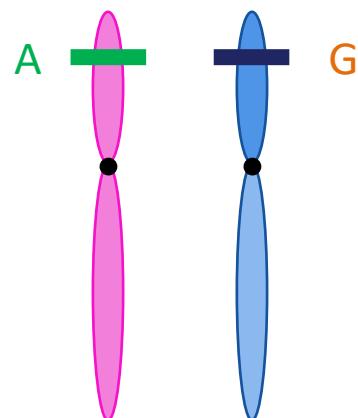
- Deux génomes humains sont identiques à 99,9%
- Différences dans la séquence du génome = **variations génétiques**
- Chaque variation génétique est caractérisée par ≥ 2 **allèles**

Allèle

Version variable d'un même locus génétique

Exemple : à une position donnée, allèle A ou allèle G

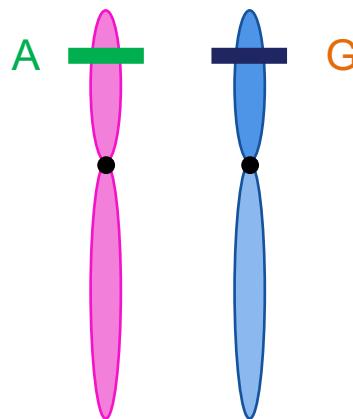
CGATGCTA[A/G]GCATGCAT



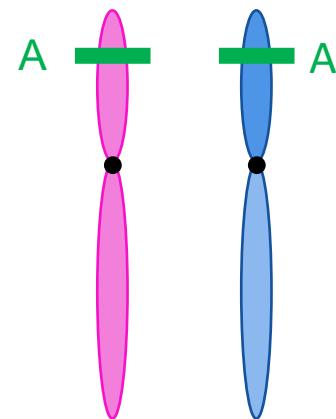
allèle maternel : A allèle paternel : G

Génotype

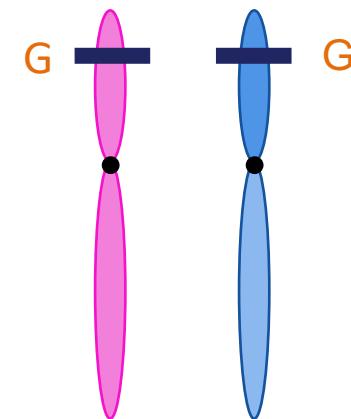
Composition allélique pour une variation génétique donnée



Génotype individu 1 : A/G



Génotype individu 2 : A/A



Génotype individu 3 : G/G

Si les 2 allèles sont différents,
on parle d'**hétérozygotie**.

Si les 2 allèles sont identiques,
on parle d'**homozygotie**.

Variation du génome

- Deux génomes humains sont identiques à 99,9%
- Différences dans la séquence du génome = variations génétiques
- Chaque variation génétique est caractérisée par ≥ 2 allèles
- Chaque variation peut être fréquente, rare ou privée
(notion de **fréquence**)

Exemple de variant fréquent



Génome de référence : CGATGCTA**A**GCATGCAT

SNP
Polymorphisme

Variant : CGATGCTA**G**GCATGCAT



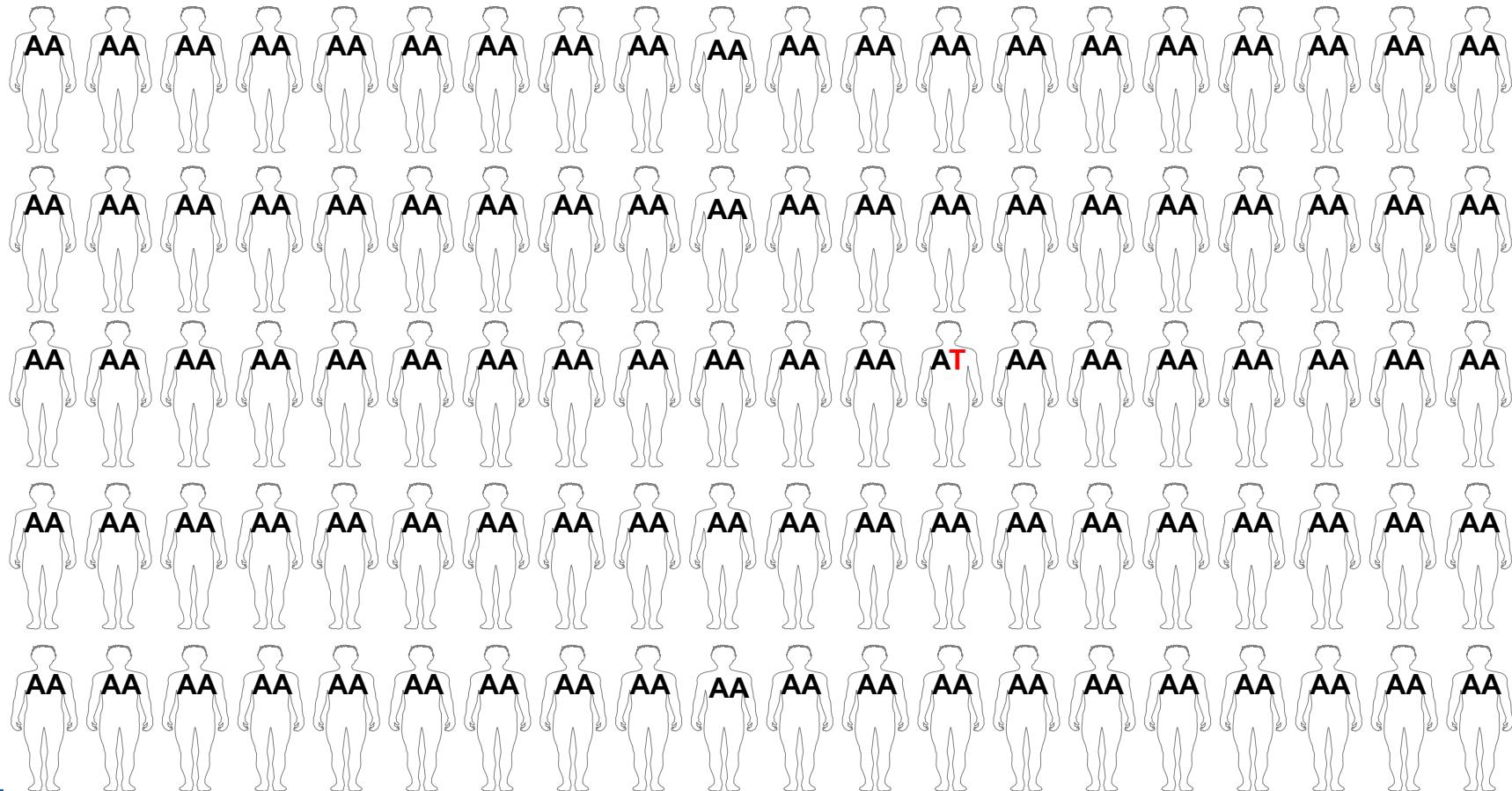
60 **G** / 200 allèles
=30 %

Exemple de variant rare (<1%)

SNV
Variant

Génome de référence : CGATGCTA**A**GCATGC**A**T

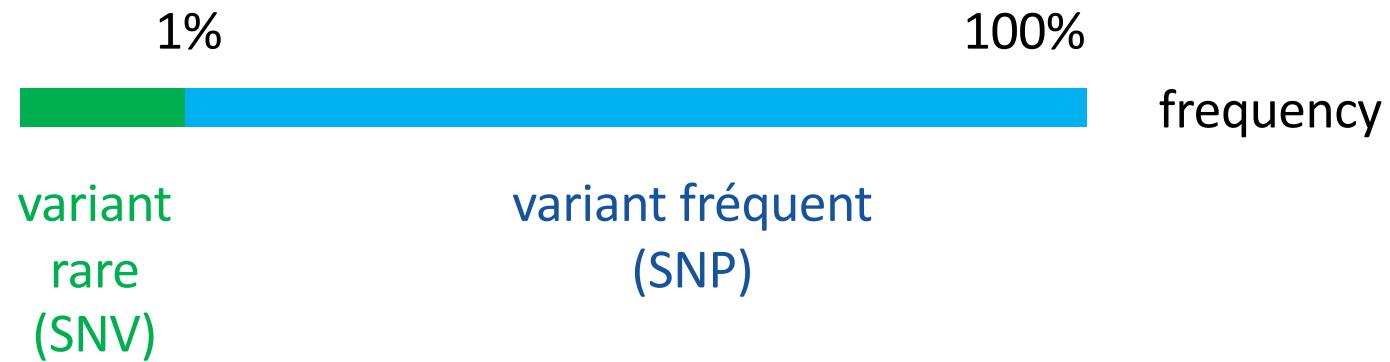
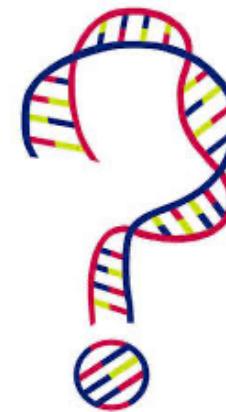
Variant : CGATGCTA**G**GCATGC**T**T



1 **T** / 200 allèles
=0,5 %

Fréquence des variants : SNV ou SNP ?

- 5%
- 0.8%



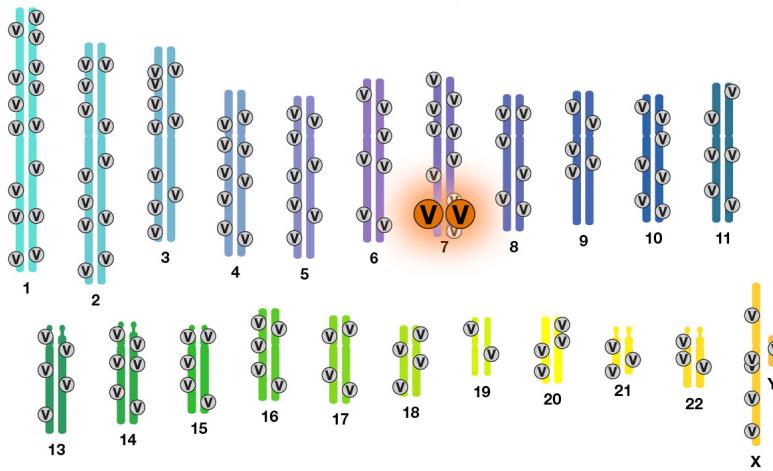
Variation du génome

- Deux génomes humains sont identiques à 99,9%
- Différences dans la séquence du génome = variations génétiques
- Chaque variation génétique est caractérisée par ≥ 2 allèles
- Chaque variation peut être fréquente, rare ou privée (notion de fréquence)
- Chaque variation, selon sa localisation, peut être bénigne, contribuer à des traits phénotypiques, prédisposer aux maladies, ou être responsable de maladies (notion **d'effet**)

Effet des variants génétiques

Effets forts

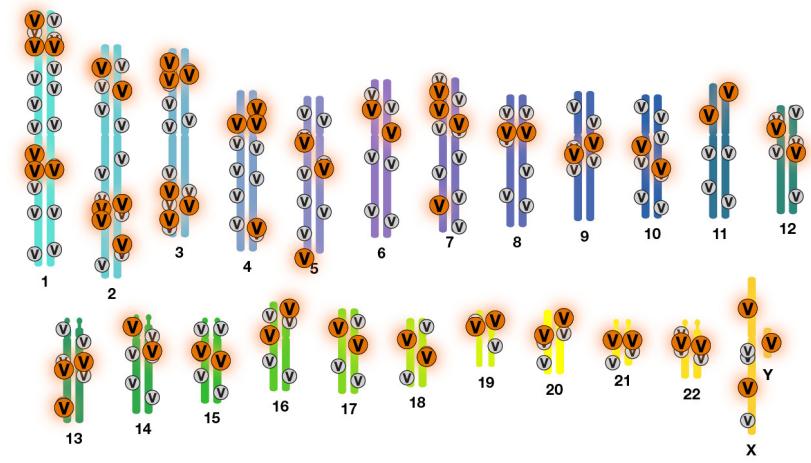
Modèle monogénique



Etude de familles ou cas isolés

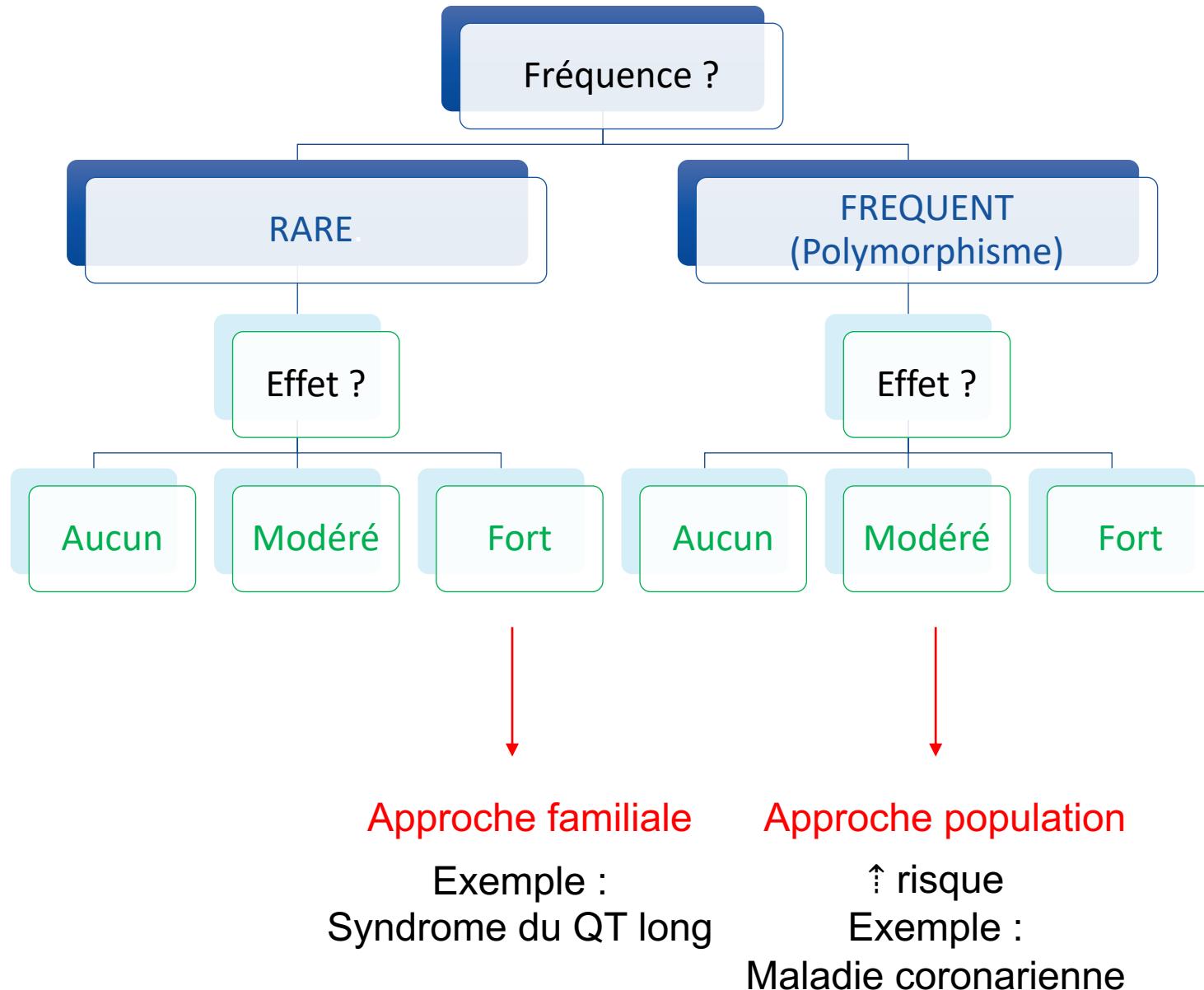
Effets faibles

Modèle polygénique

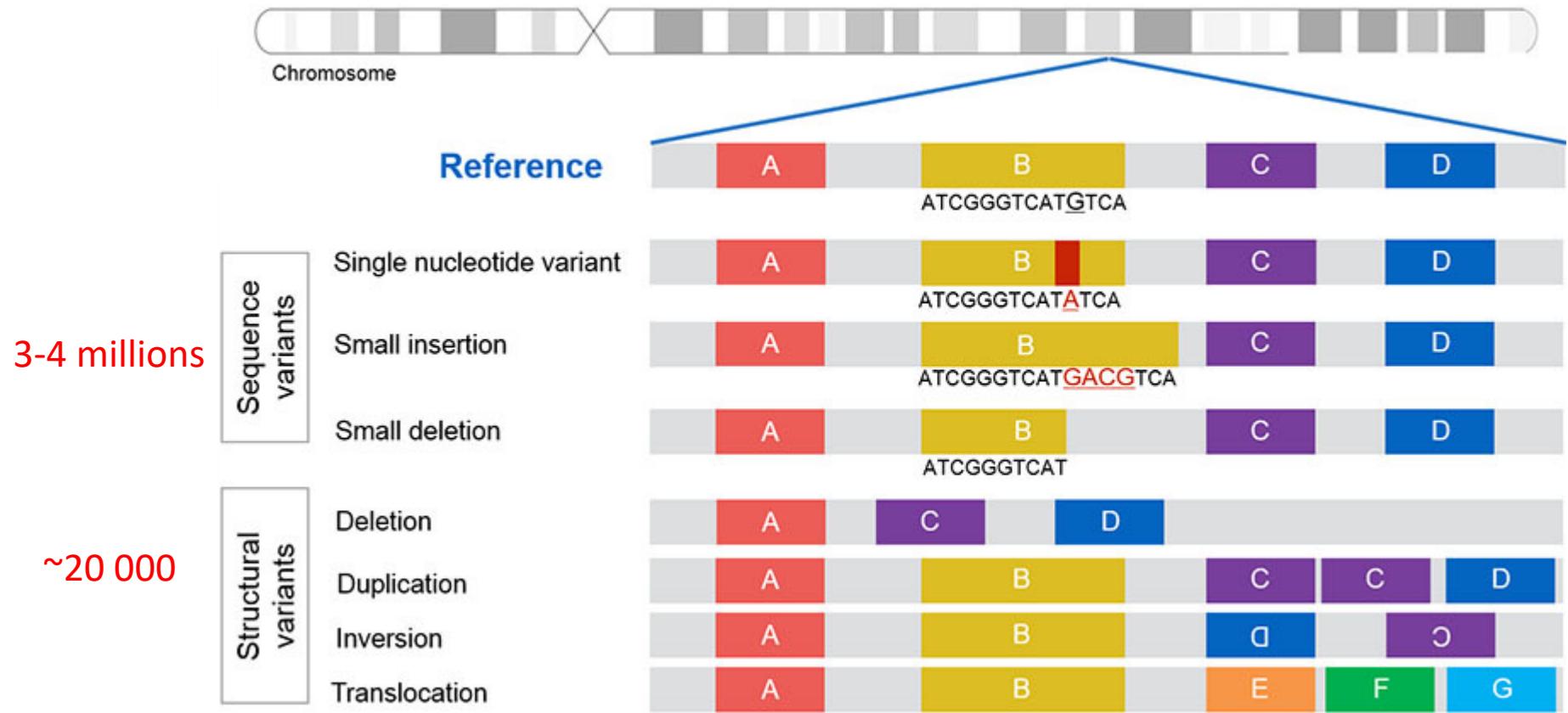


Etude de populations

Effet des variants génétiques



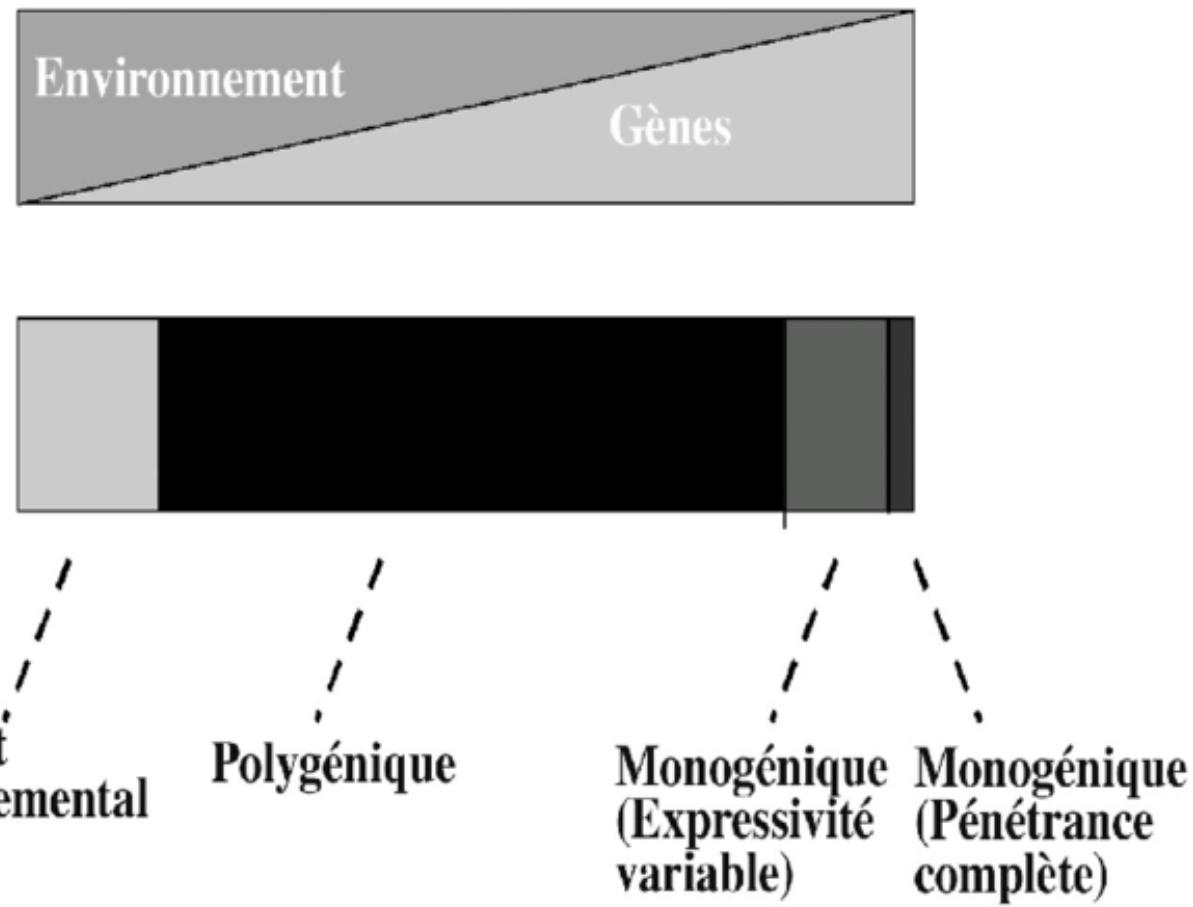
Variabilité du génome



Plan

- Introduction
- Variation du génome humain
- Méthodes d'identification de gènes
 - ✓ Maladies monogéniques
 - ✓ Maladies complexes

Maladies humaines



Génétique moléculaire

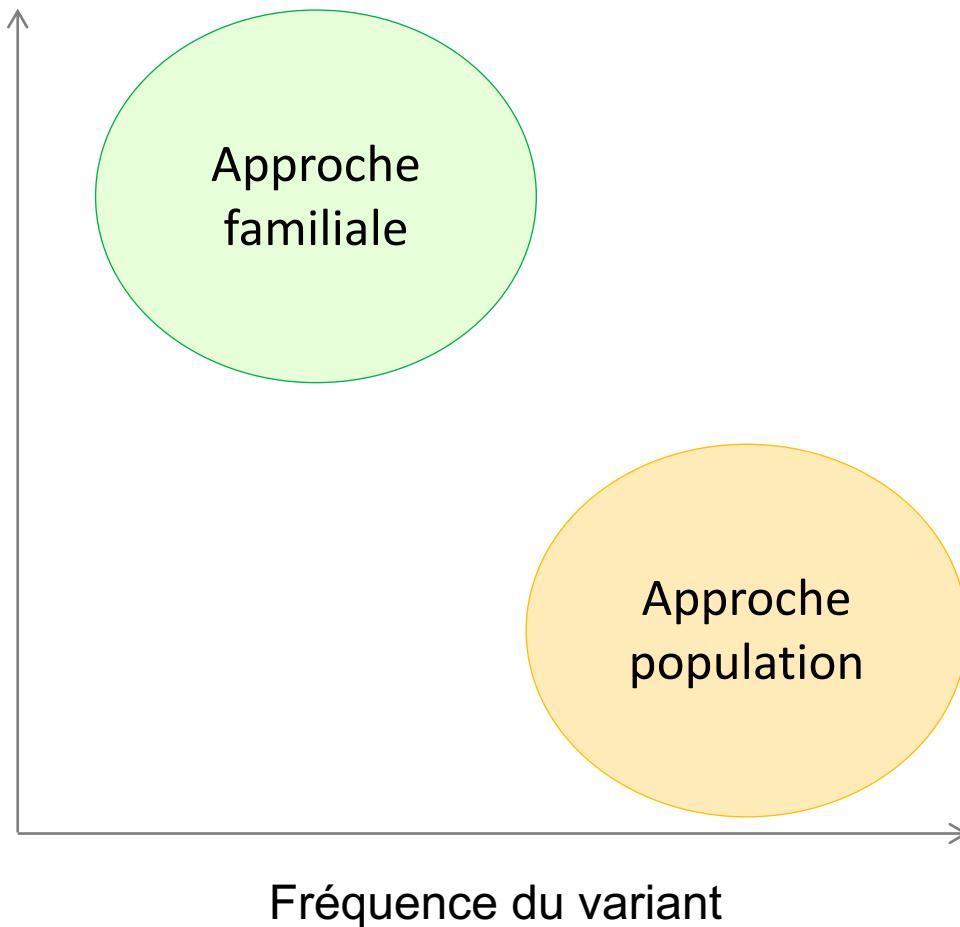
- ⇒ Gènes de susceptibilité
- ⇒ Nouvelles connaissances physiopathologiques
- ⇒ Nouveaux traitements
- ⇒ Médecine personnalisée

Implications éthiques et légales !

Ex : 23andMe

Stratégies génétiques

Effet du variant



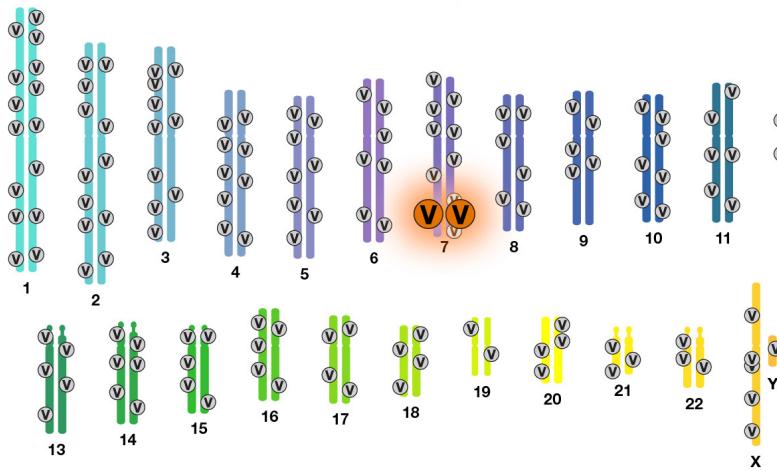
Plan

- Introduction
- Variation du génome humain
- Méthodes d'identification de gènes
 - ✓ Maladies monogéniques
 - ✓ Maladies complexes

Maladies monogéniques

Effets forts

Modèle monogénique

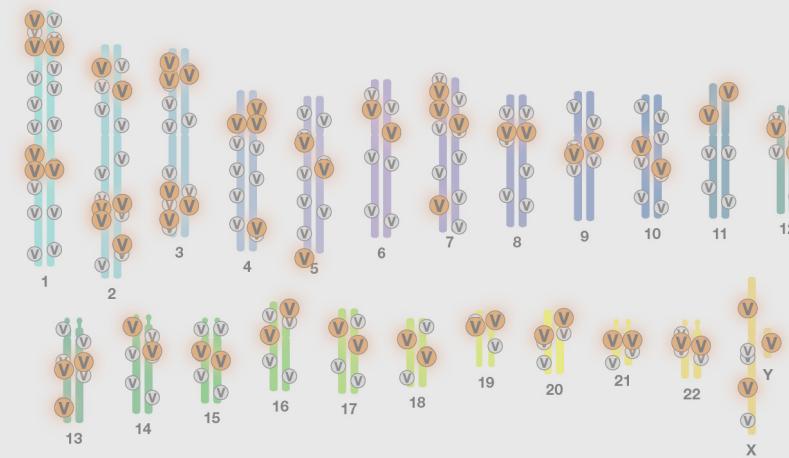


Etude de familles ou cas isolés

- Un gène
- Mutations rares avec un effet fort
- Agrégation familiale (souvent)
- Age d'apparition souvent précoce

Effets faibles

Modèle polygénique



Etude de populations

Les modes de transmission: autosomique dominant

Un des 2 parents est atteint

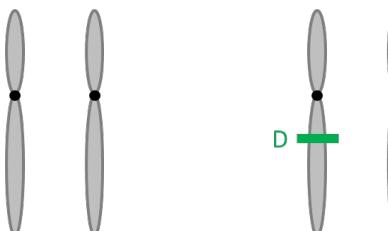
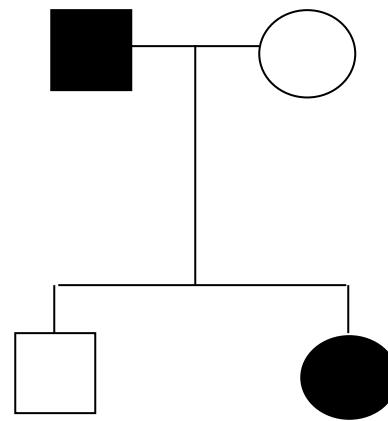
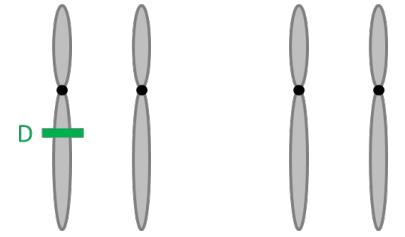
⇒ Un enfant sur 2 est atteint

Mais parfois dans une famille, on observe:

-la mutation chez un individu sain (**« non pénétrance »**)

-la mutation est absente chez un individu atteint (**« phénomécopie »**)

-une variabilité dans les phénotypes observés: **expressivité variable** (fonction d'autres gènes, de l'environnement?!)



Les modes de transmission: autosomique récessif

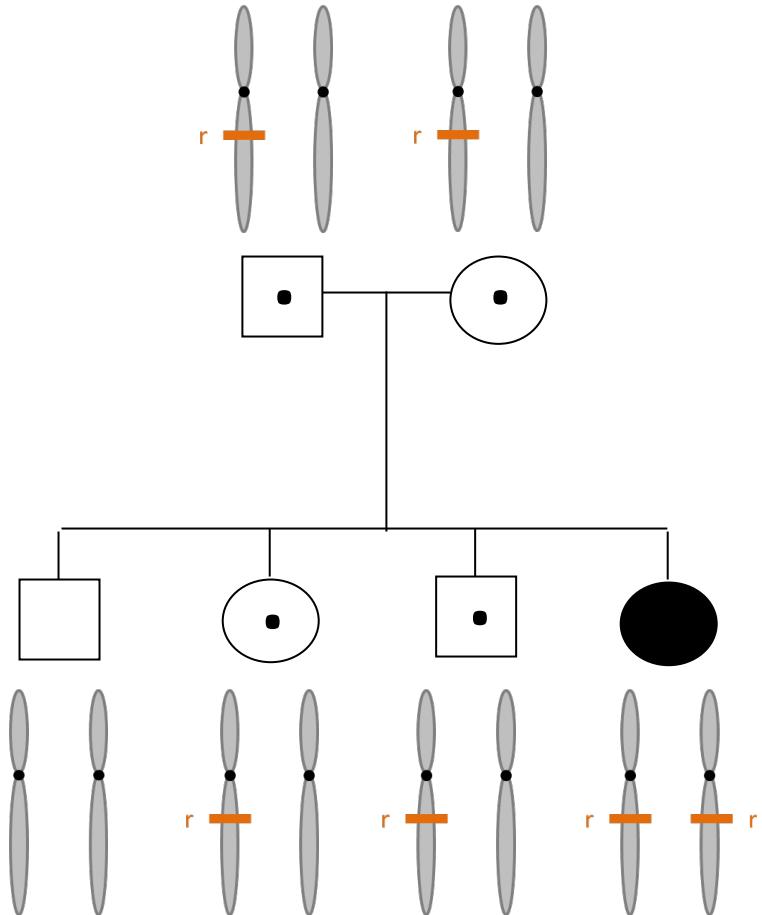
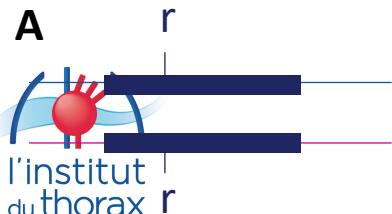
Les parents ne sont pas atteints

⇒ Un enfant sur 4 est atteint

⇒ Un enfant sur 2 est porteur sain

-si les variants paternels/maternels sont identiques, il y a donc homozygotie chez l'enfant atteint (**A**). Risque \geq avec consanguinité.

-si les 2 variants sont distincts, mais situés dans un même gène, alors l'enfant atteint est **hétérozygote composite** pour le gène (**B**), ce qui a le même effet qu'un variant récessif homozygote (les 2 copies du gène sont altérées)



Modes de transmission

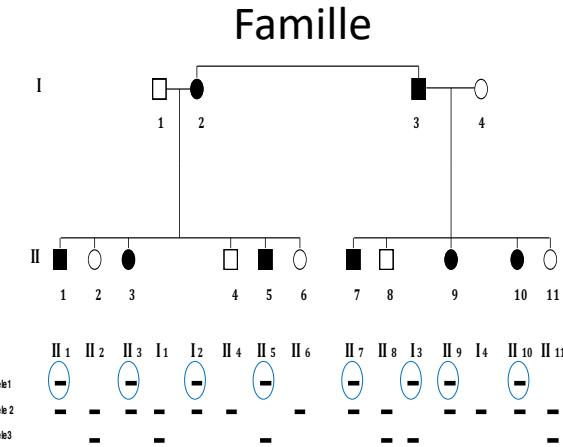
Mode de transmission	Localisation du locus responsable du trait	Principales caractéristiques	Exemples de phénotypes
Autosomique dominant	Chromosome 1-22	<ul style="list-style-type: none"> -hommes et femmes sont atteints avec une fréquence équivalente -les enfants atteints ont un parent atteint 	Huntington, syndrome de Marfan, arythmies, cardiomyopathies
Autosomique récessif	Chromosome 1-22	<ul style="list-style-type: none"> -hommes et femmes sont atteints avec une fréquence équivalente -les enfants atteints n'ont pas de parent atteint -peut « sauter » une ou plusieurs générations 	Mucoviscidose, drépanocytose
Dominant lié à l'X	Chromosome X	<ul style="list-style-type: none"> -les femmes atteintes ont 50% d'enfants atteints -les hommes sont plus sévèrement atteints -toutes les filles de pères atteints sont atteintes 	Syndrome de Rett
Récessif lié à l'X	Chromosome X	<ul style="list-style-type: none"> -beaucoup plus d'hommes atteints que de femmes -les hommes atteints ont une mère saine (porteuse) -le trait n'est jamais transmis de père à fils -peut « sauter » une ou plusieurs générations 	Hémophilie
Lié à l'Y	Chromosome Y	<ul style="list-style-type: none"> -tous les hommes atteints passent le trait à leur fils -les femmes ne sont pas atteintes 	Stérilité
Mitochondrial	Pas chromosomique ADN mitochondrial	<ul style="list-style-type: none"> -tous les enfants d'une femme atteinte sont atteints -les hommes atteints ne transmettent pas le trait 	Neuropathie optique de Leber

Du malade au gène et du gène au malade

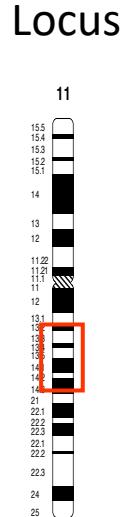
Malade (cas index)



Enquête familiale

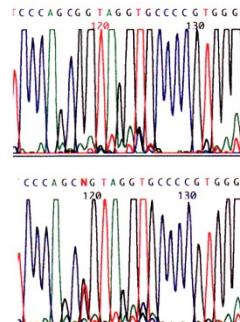


Génotypage Analyse de liaison

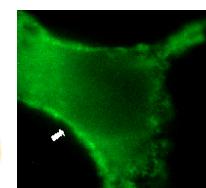
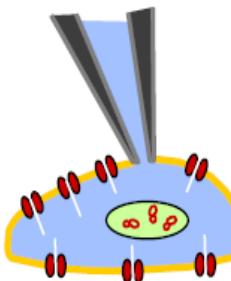


Séquençage

Gène/Mutation



Fonction



Malade Diagnostic moléculaire Prise en charge clinique

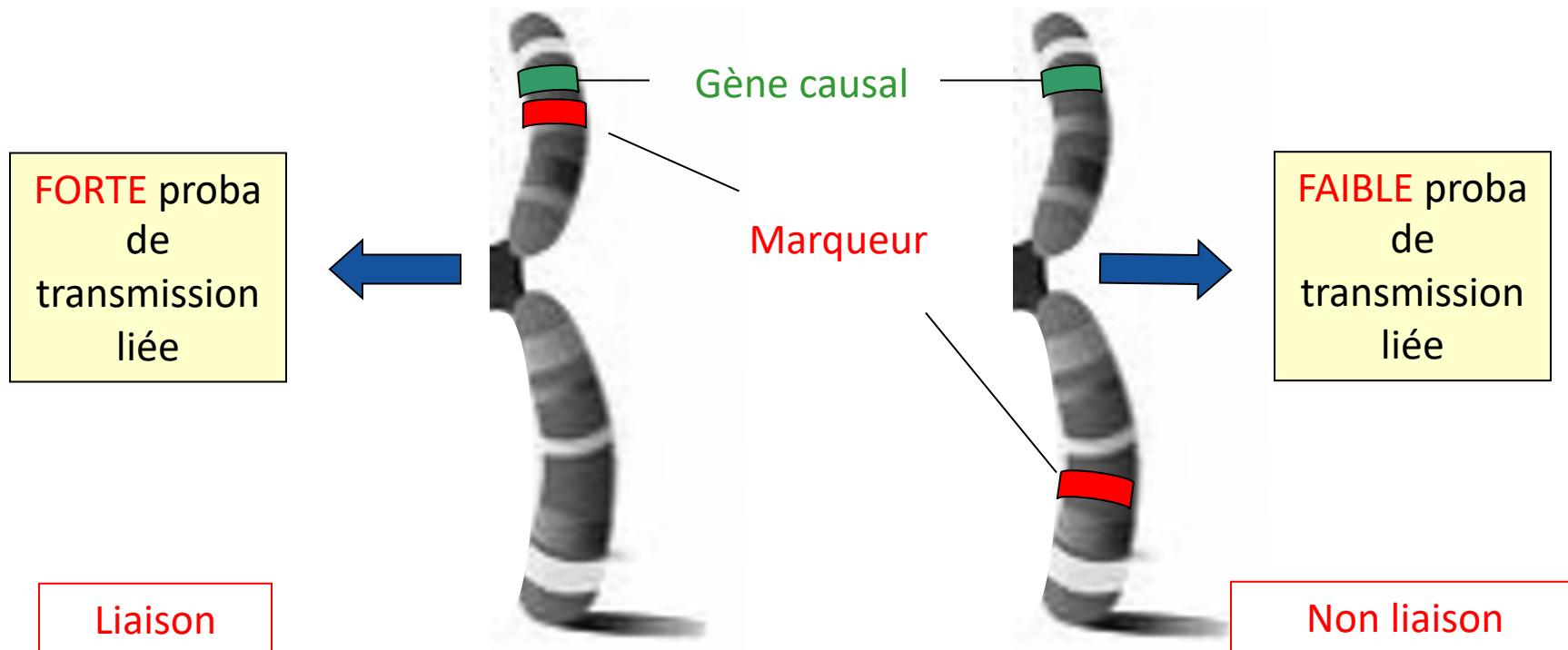
Analyse de liaison

Corréler la transmission d'un génotype avec
un phénotype

L'analyse de liaison génétique

= clonage positionnel

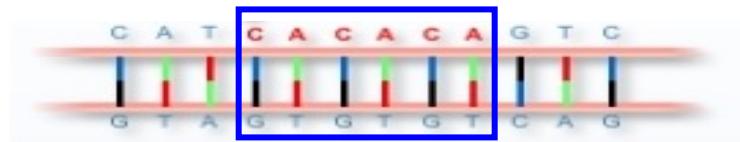
= « linkage »



Lod Score > 3

Génotypage de marqueurs

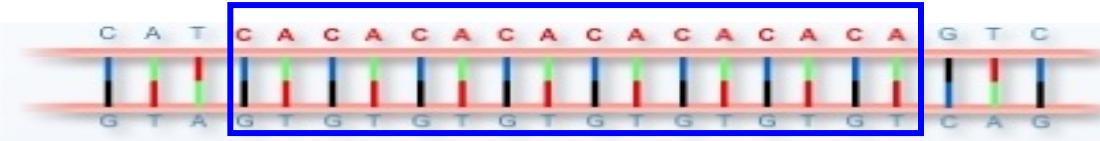
Allèle 4



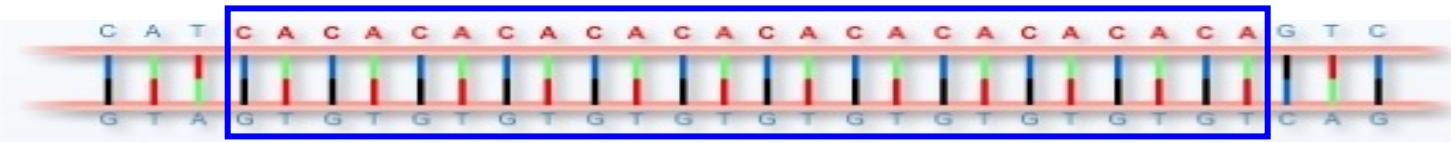
Allèle 3



Allèle 2

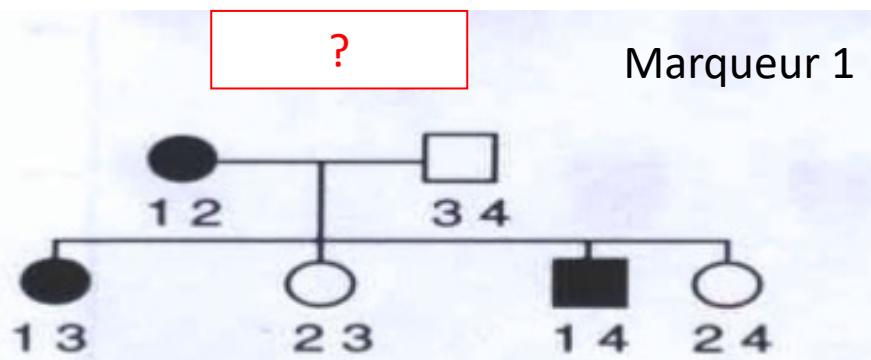


Allèle 1



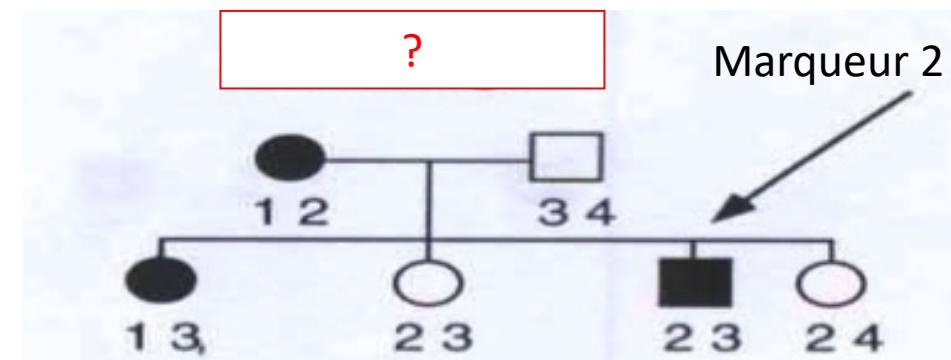
?

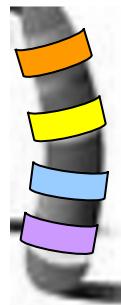
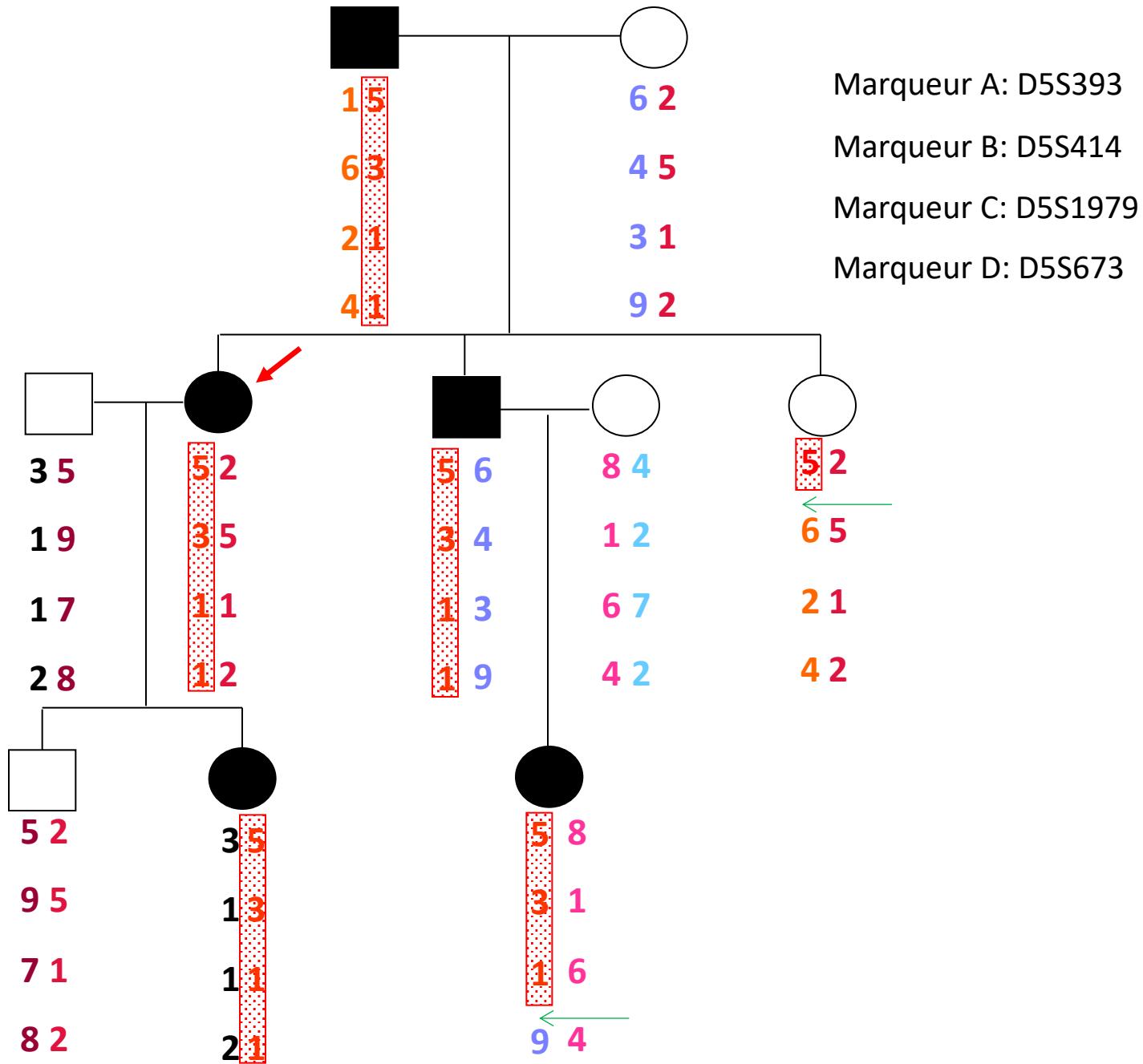
Marqueur 1



?

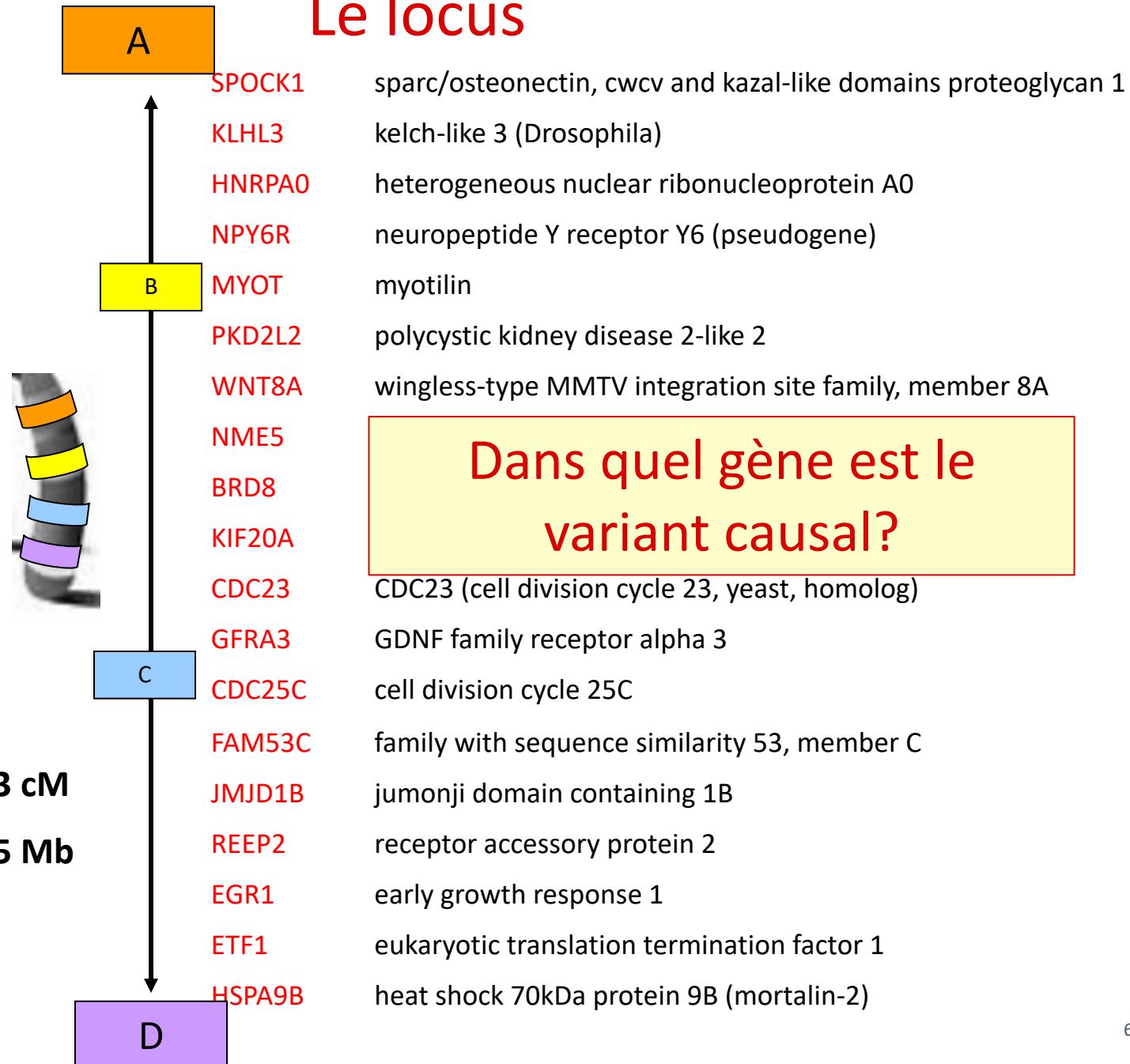
Marqueur 2





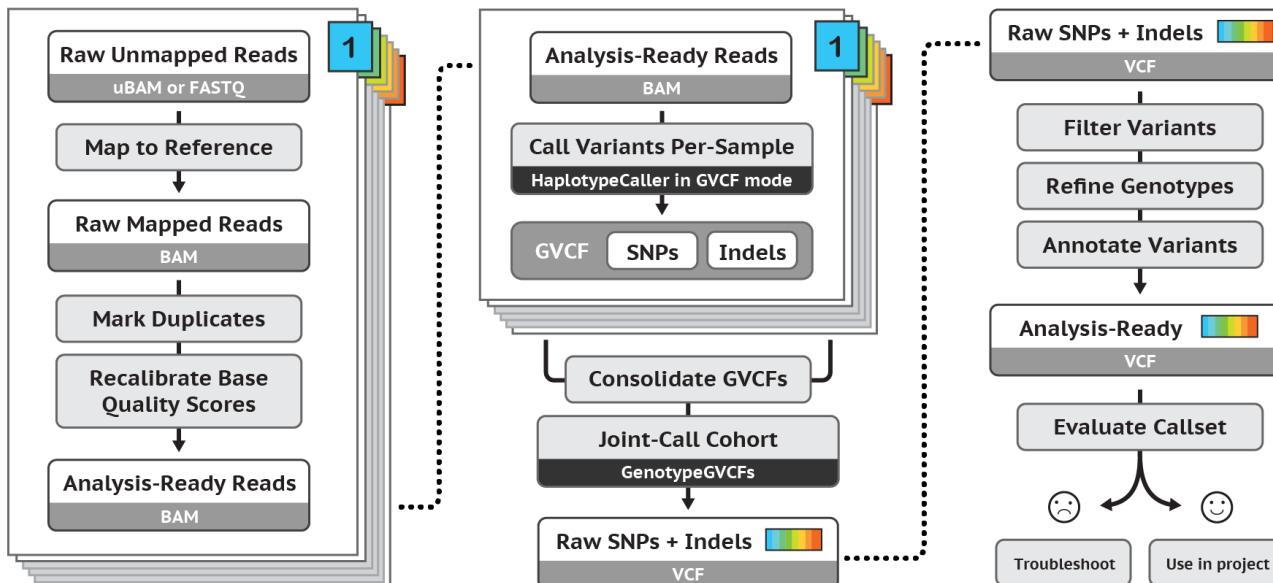
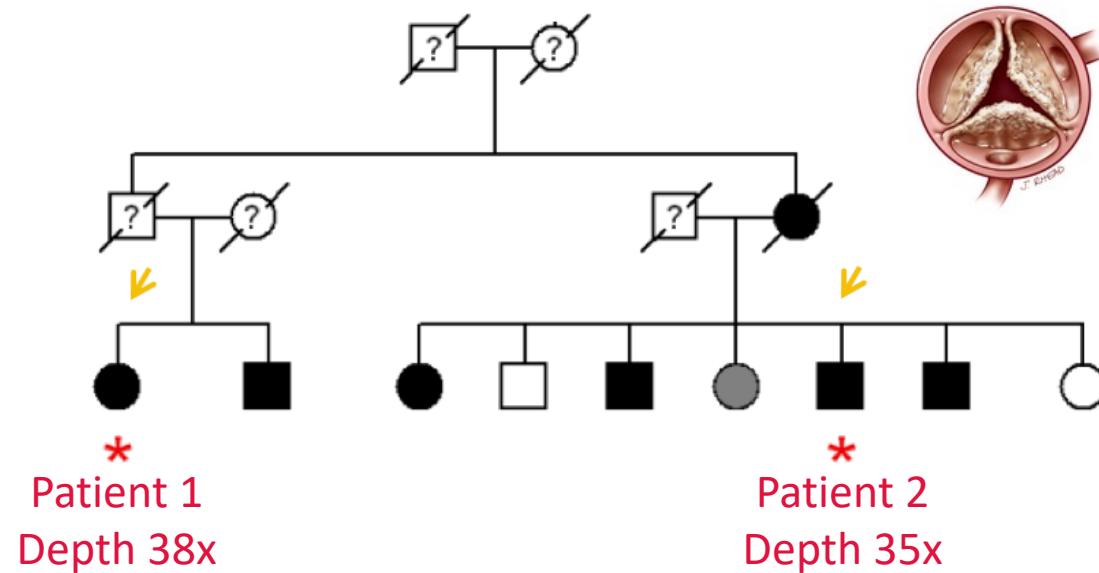
Le locus

Marqueur A: D5S393
Marqueur B: D5S414
Marqueur C: D5S1979
Marqueur D: D5S673



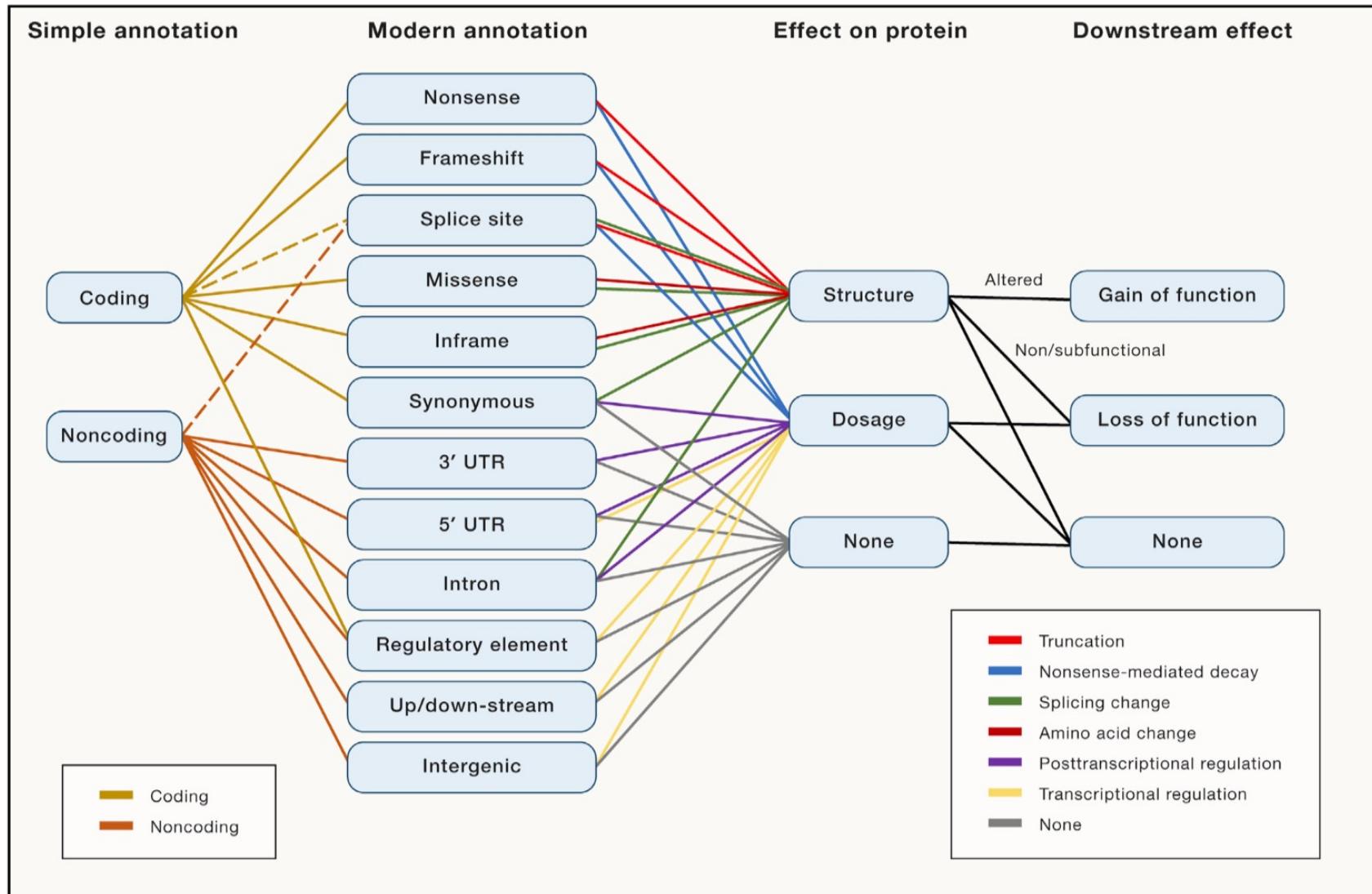
Dans quel gène est le variant causal?

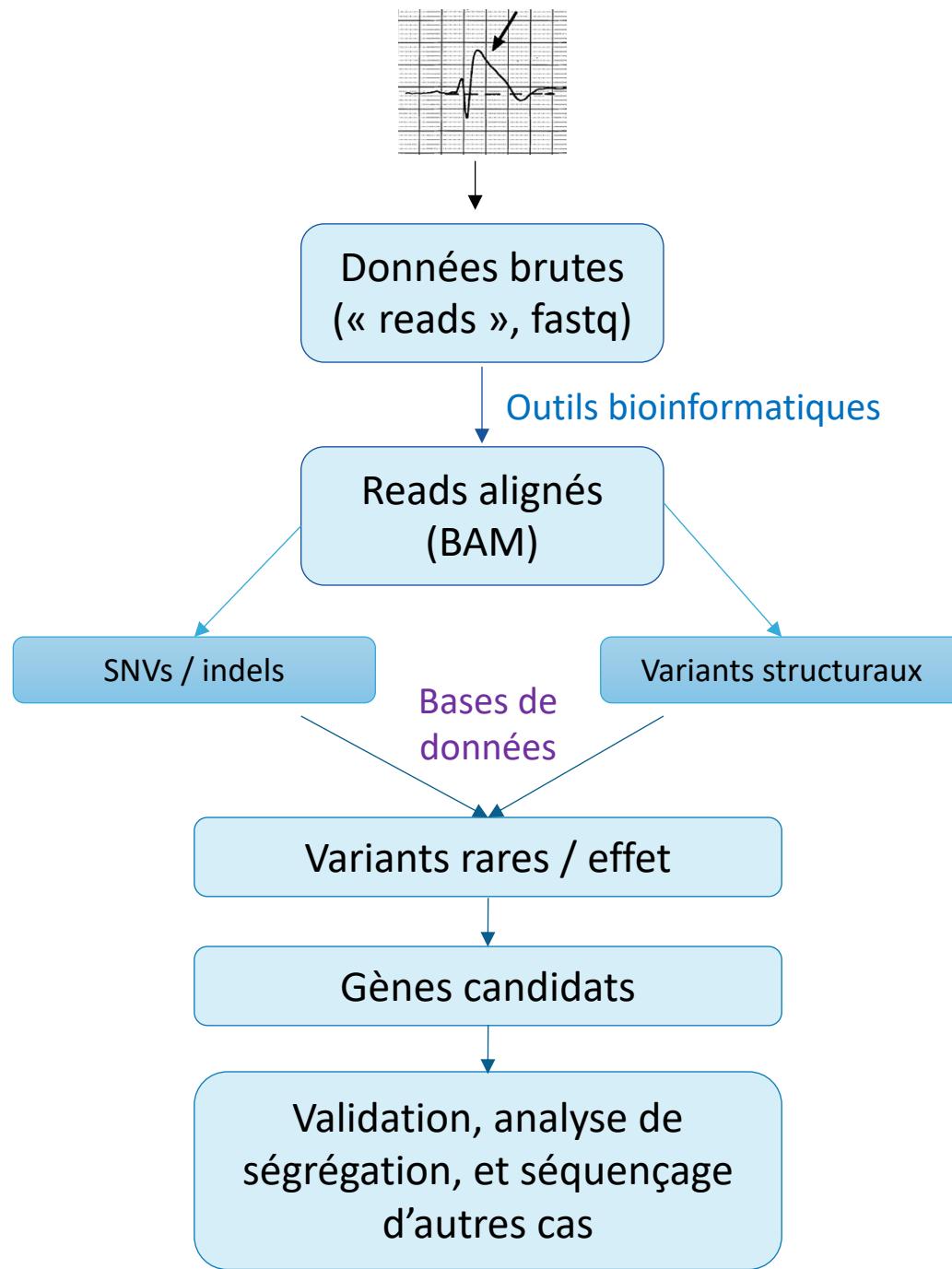
Analyse des données de séquençage



Analyses
bioinformatiques

Déterminer l'effet des variations : un challenge





Interprétation des données de séquençage

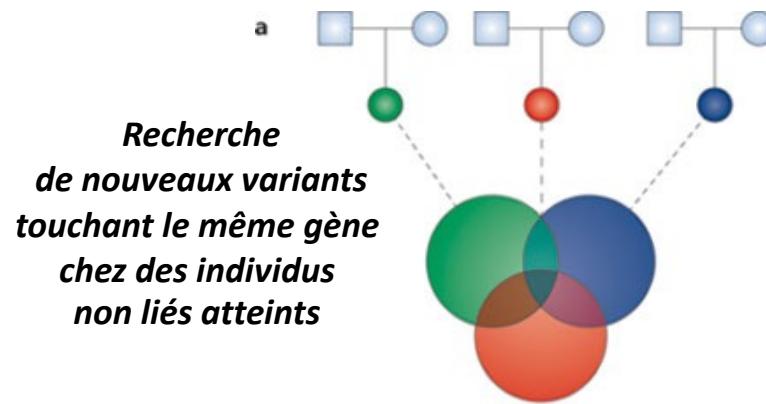
	Patient 1	Patient 2
# total substitutions + indels	4 978 101	4 867 920
# variants rares	58 468	52 972
# variants rares + conséquence (exons / épissage)	450	438
# variants rares + faux-sens	276	266
# variants rares + stop	8	12
# variants rares + frameshift	44	38
# variants rares + conséquence partagés	55 (dont 8 stop/frameshift)	



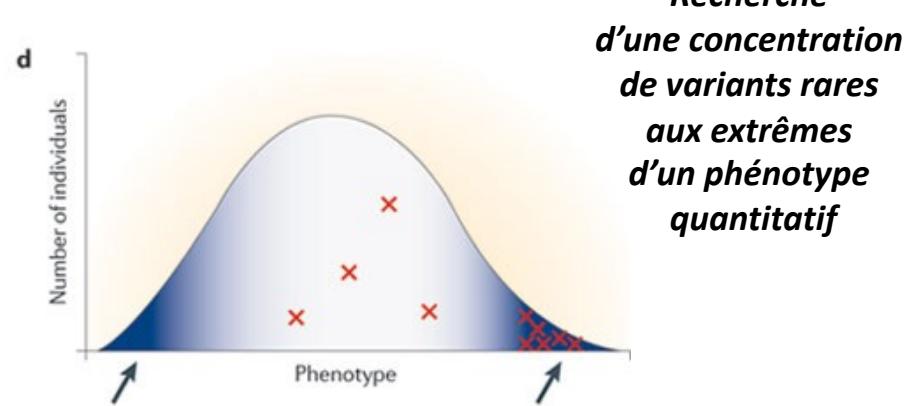
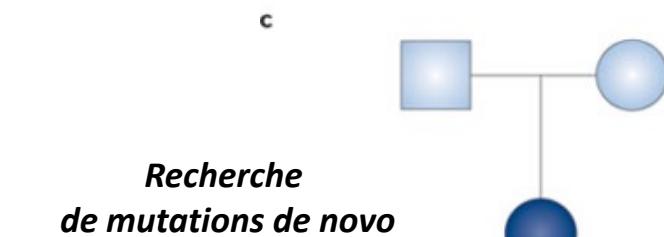
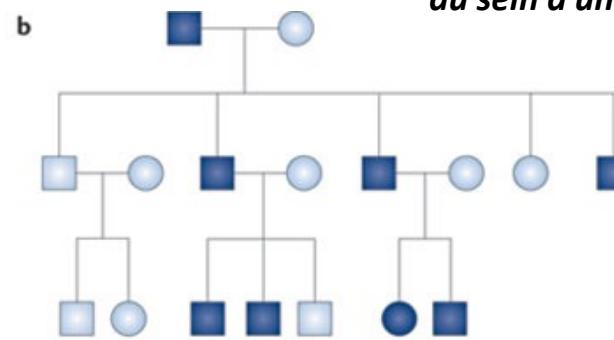
« Grands » gènes (exemple de la *TTN* : un variant rare faux-sens chez un individu sur 2)

Take-home message : la plupart des variants rares sont bénins

Découverte par séquençage de variants rares causant des maladies



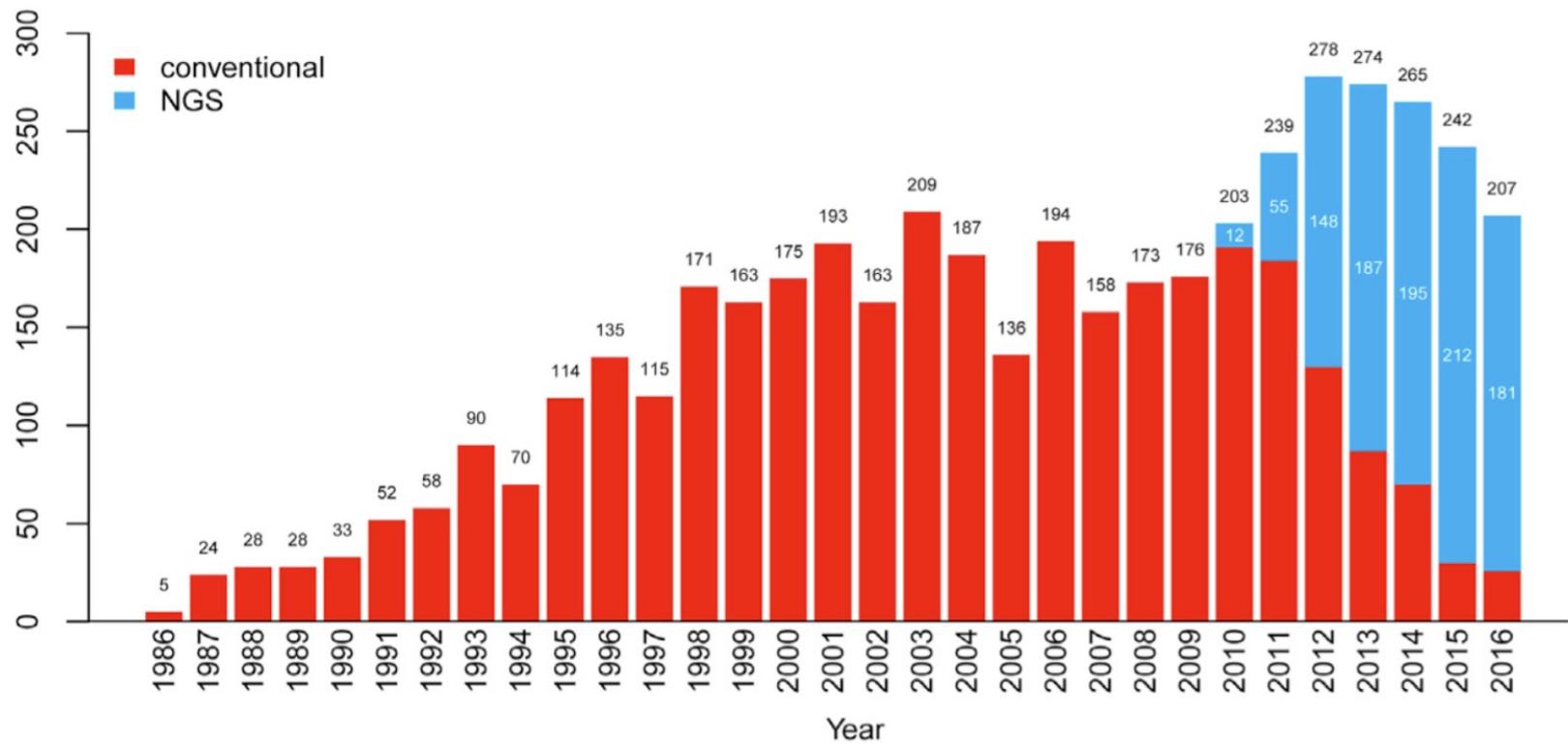
Recherche de nouveaux variants partagés entre individus atteints au sein d'une même famille



Nature Reviews | Genetics

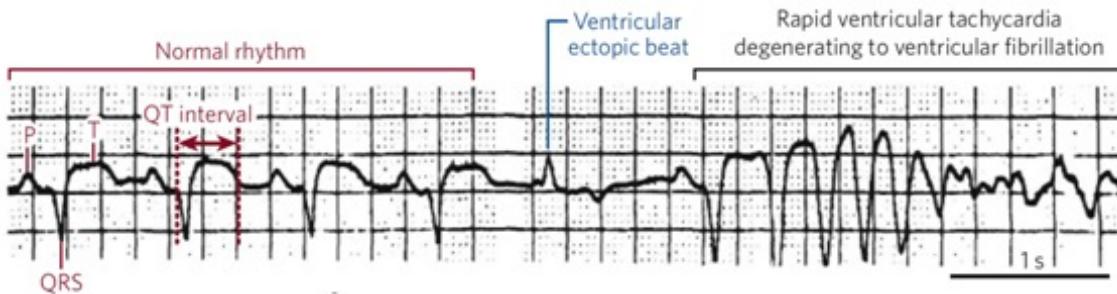
Bamshad et al, Nat Rev Genet 12:745, 2011⁶⁵

Identification de nouveaux gènes (maladies rares)

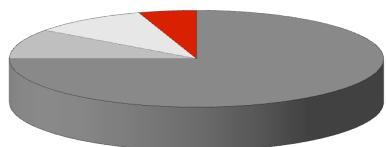


Arythmies cardiaques

Mort subite cardiaque



Fibrillation ventriculaire : mécanisme dans 80% des cas de MSC



Chugh SS, 2000

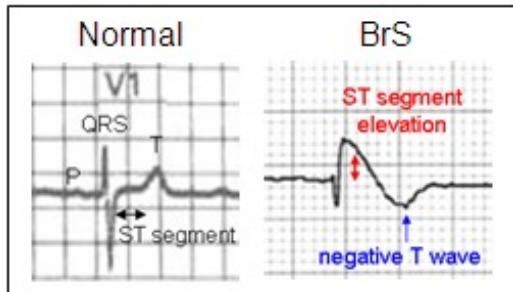
- Maladie coronarienne (75%)
- Cardiomyopathie hypertrophique (10%)
- Cardiomyopathie dilatée (10%)
- Cœur structurellement normal (5%)

Fibrillation ventriculaire

- ⇒ **Syndrome de Brugada**
- ⇒ **Syndrome de repolarisation précoce**
- ⇒ **Syndrome du QT long**
- ⇒ ...

Syndrome de Brugada

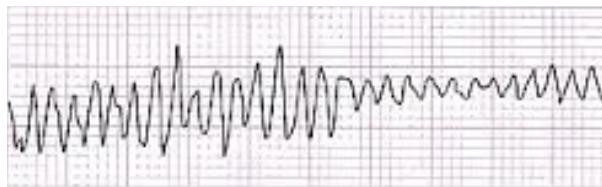
Elévation du segment ST (> 2 mm) dans les dérivations précordiales droites



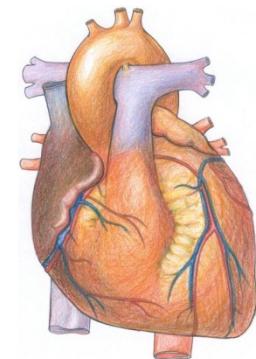
Brugada et al, 1992

Chen et al, 1998

Susceptibilité aux arythmies ventriculaires



Cœur structurellement normal (à la description du syndrome)



- Modèle : transmission autosomique dominante ou complexe
- Mutations SCN5A dans ~ 15-25% des cas
- Faible pénétrance
- Mutations dans plus de 20 autres gènes

➤ La majorité des cas sont encore inexplicables => autres études

Séquençage d'exome – syndrome de Brugada

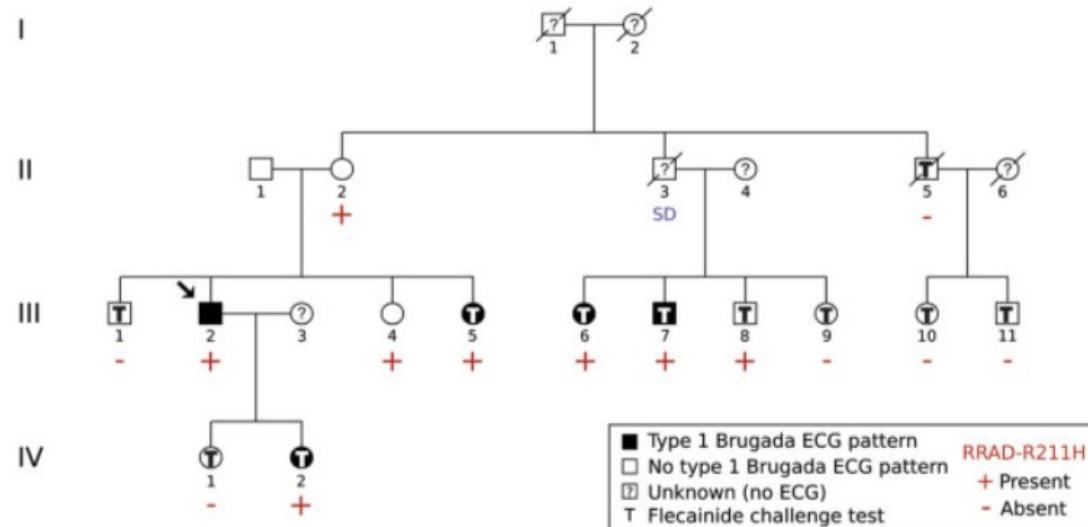


European Heart Journal (2019) 0, 1–14
doi:10.1093/eurheart/ehz308

BASIC SCIENCE
Arrhythmia/electrophysiology

RRAD mutation causes electrical and cytoskeletal defects in cardiomyocytes derived from a familial case of Brugada syndrome

- Identification de RRAD-R211H dans une forme familiale
- 3 autres variants identifiés chez des cas sporadiques
- Investigations fonctionnelles iPSC-CMs



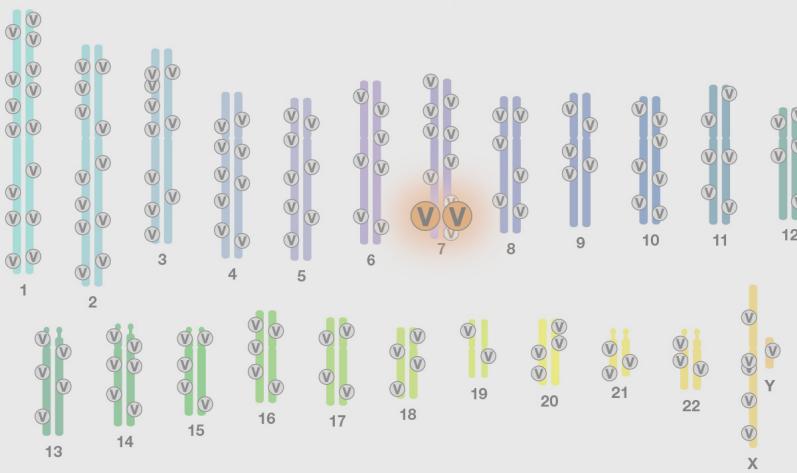
Exome sequencing (patient III:2)	
Coding +/- 8 bp*	22,562
Functional**	10,412
Rare in ExAC***	217
In shared genomic regions	9
Validated and shared by the 5 BrS patients	6 missense variants (<i>RRAD</i> , <i>FHOD1</i> , <i>H6PD</i> , <i>SNAP23</i> , two in <i>SPTBN5</i>)
Predicted deleterious (SIFT and PolyPhen)	<i>RRAD</i> (R211H)

Plan

- Introduction
- Variation du génome humain
- Méthodes d'identification de gènes
 - ✓ Maladies monogéniques
 - ✓ Maladies complexes

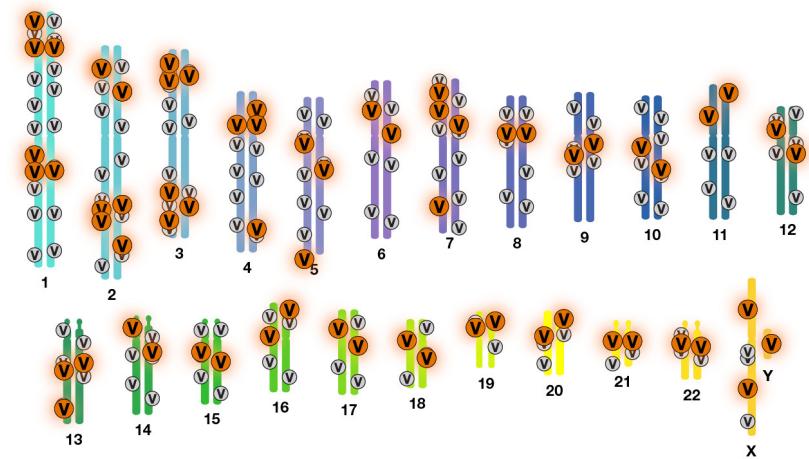
Maladies polygéniques

Effets forts
Modèle monogénique



Etude de familles ou cas isolés

Effets faibles
Modèle polygénique



Etude de populations

Etude d'association

Corréler la présence d'un génotype avec un phénotype

Hypothèse : l'héritabilité des maladies communes est liée à des variants génétiques fréquents

Principe de l'étude d'association

Marqueur 1

Cas 35 %

Témoins 15 %

Marqueur 2

Cas 20 %

Témoins 20 %

Association

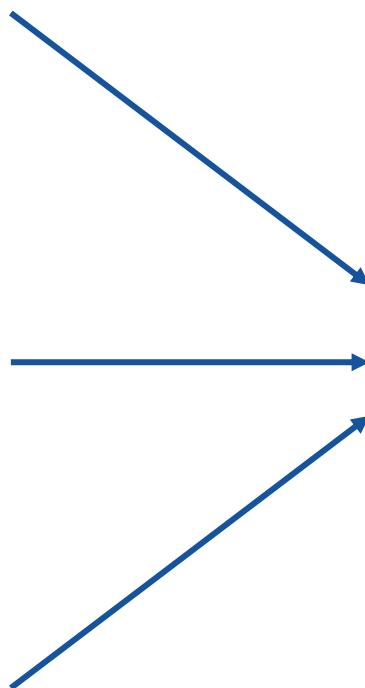
Pas d'association

Etudes d'association chez l'homme

Meilleure compréhension de la variation du génome humain (HapMap, 1000G)

Avancées dans les technologies de génotypage (SNP)

Grandes bio-collections de patients

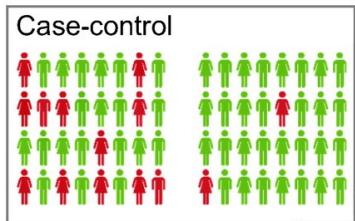


Genome-wide association studies
= (GWAS)



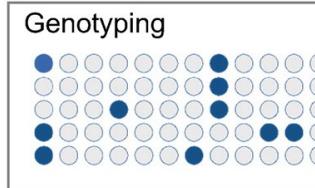
Etapes d'une GWAS

A

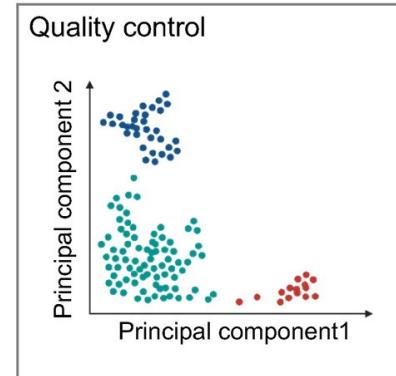


Quantitative trait

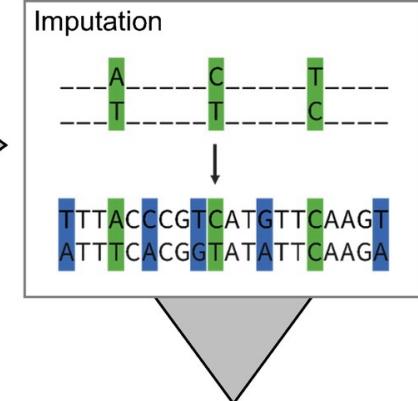
B



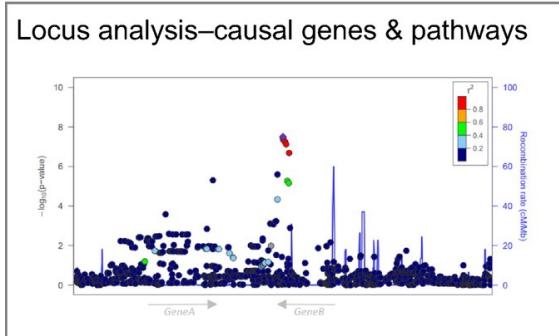
C



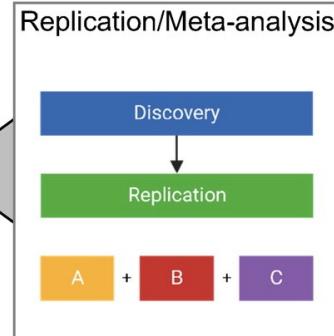
D



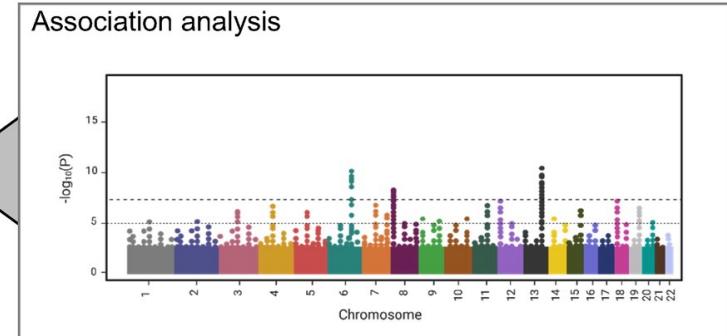
G



F

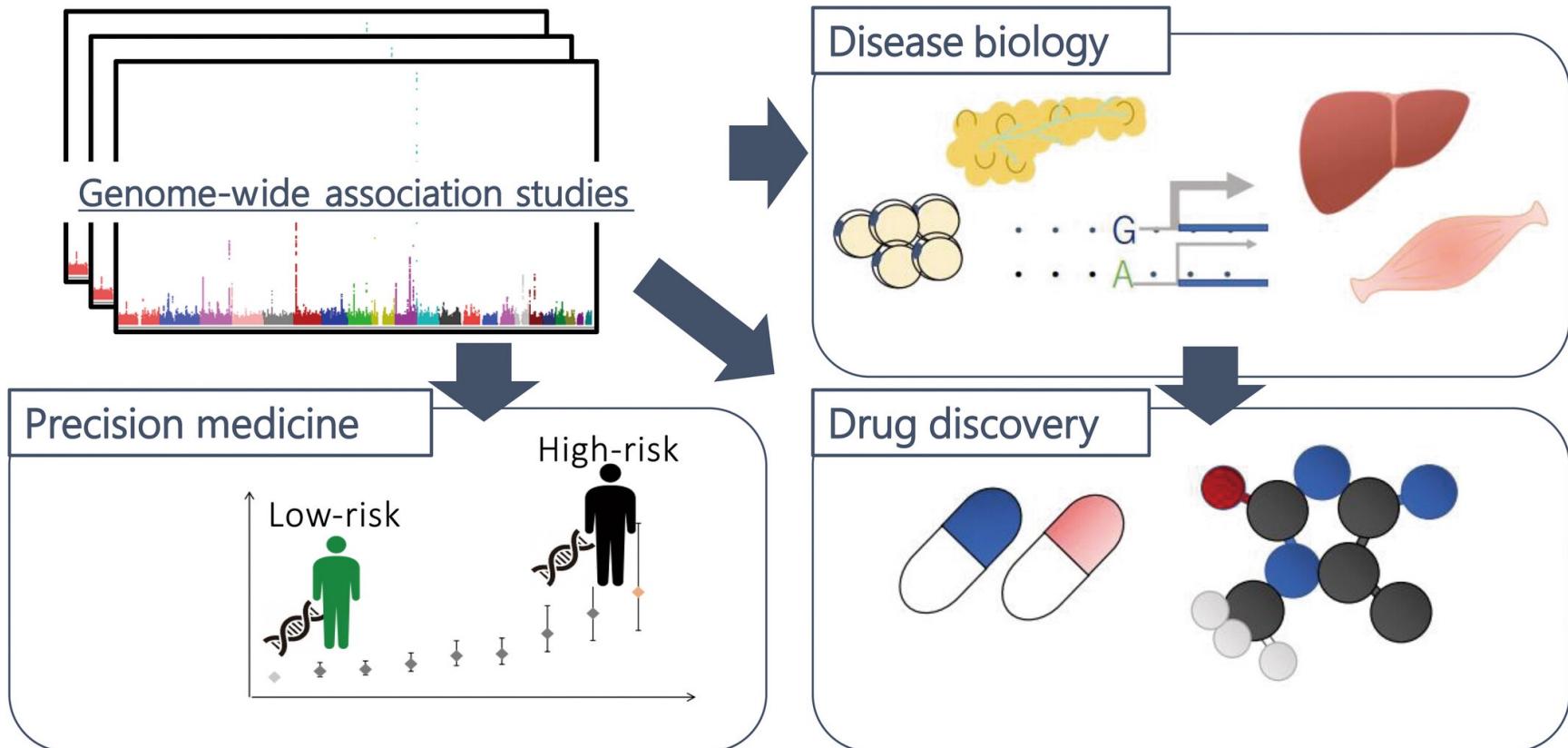


E





Perspectives



Stratégies génétiques

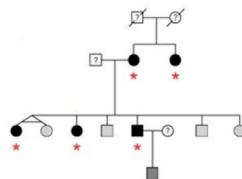
Effet du variant

Fort

Faible

Approche familiale

Analyse de liaison
Séquençage



Séquençage de génomes

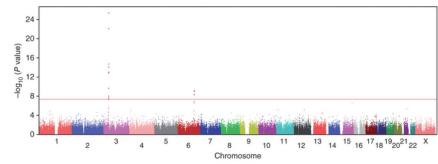
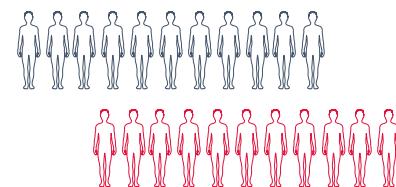
Rare

Elevée

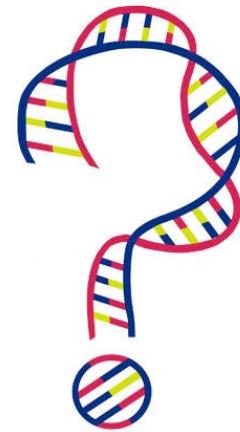
Fréquence du variant

Approche population

Etude d'association
(GWAS)



Questions



umr1087.univ-nantes.fr



L'unité de recherche de l'institut du thorax
Inserm UMR 1087 / CNRS UMR 6291
Nantes, France



Daniel Buren et Patrick Bouchain, Les Anneaux, Quai des Antilles, Nantes, création pérenne Estuaire 2007 © Martin Argyroglo/LVAN