tous semblables, tous différents

PO17421

Explorer avec les élèves l'unicité et la diversité de l'ADN

Login: usr_m_bibcmu
Password: geneve14

http://education.expasy.org/cours/PO17421/

SIB
Swiss Institute of Bioinformatics

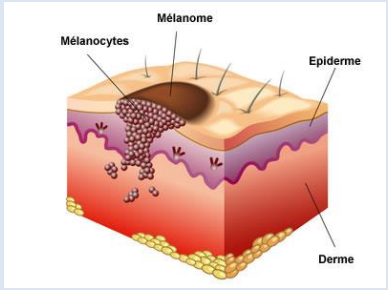
François.Lombard@unige.ch
&
Marie-Claude.Blatter@sib.swiss

(1) Variations inter-espèces




(2) Variations intra-espèce (humain)

(3) Variations intra-individu (somatiques)



Activités

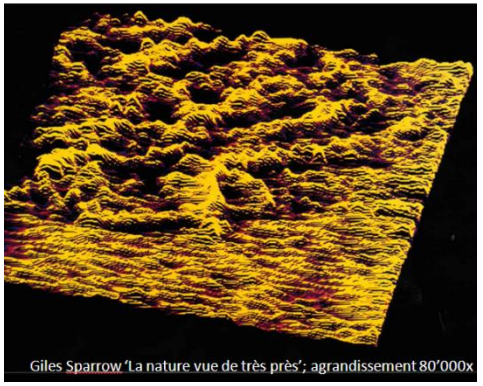
- 1.1 ‘Voir’ l’unicité/diversité en comparant des séquences de protéines (1) (Align@UniProt)
- 1.2 ‘Voir’ l’unicité/diversité en comparant des séquences de protéines (2) (Align@UniProt)

- 2.1 Comparer des séquences ADN de mitochondrie (Néanderthal, ‘mon génome’, chimpanzé, etc…) (Align@UniProt)
- 2.2 Comparer les séquence ADN de mitochondrie de 60 individus (Align@UniProt)
- 2.3 Comparer une séquence ADN de mitochondrie avec le génome de référence et repérer les SNP (BLAT@UCSC) *(expert)*
- 2.3 bis Comparer la séquence ADN de mitochondrie de ‘mon génome’ avec le génome de référence et repérer les SNP (BLAT@UCSC) *(expert)*
- 2.4 Comparer une séquence ADN de mitochondrie humain avec le génome différentes espèces (BLAT@UCSC); ‘Voir’ l’unicité/diversité
- 2.5 Annotation des variants dans UniProtKB/Swiss-Prot (TAS2R38, CFTR, BRCA1)
- 2.6 ‘Voir’ les variations génétiques de *Arabidopsis Thaliana*
- 2.7 Une famille, un arbre généalogique, une maladie génétique: exemple complet

- 3.1 ‘Trouver’ une variation somatique associée avec le mélanome (Align@UniProt)
- 3.2 Traitement personnalisé: exemple BRAF
- Variations somatiques chez les plantes….

http://education.expasy.org/cours/PO17421/

ADN vu au microscope électronique à balayage



Chromosome

ADN

Séquence ADN

Séquence protéine

Swiss Institute of Bioinformatics

De l'ADN aux protéines…

Chromosome 10

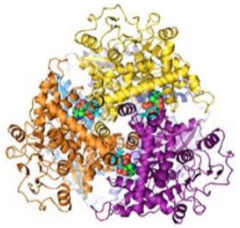
ggatatcggtgag

*

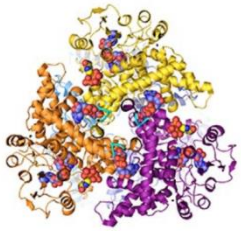
ggatattggtgag

Référence

Mlle XXXX



?



Génotype & Phénotype

Association statistique (GWAS)

Analyse fonctionnelle de la protéine (expérimentale)

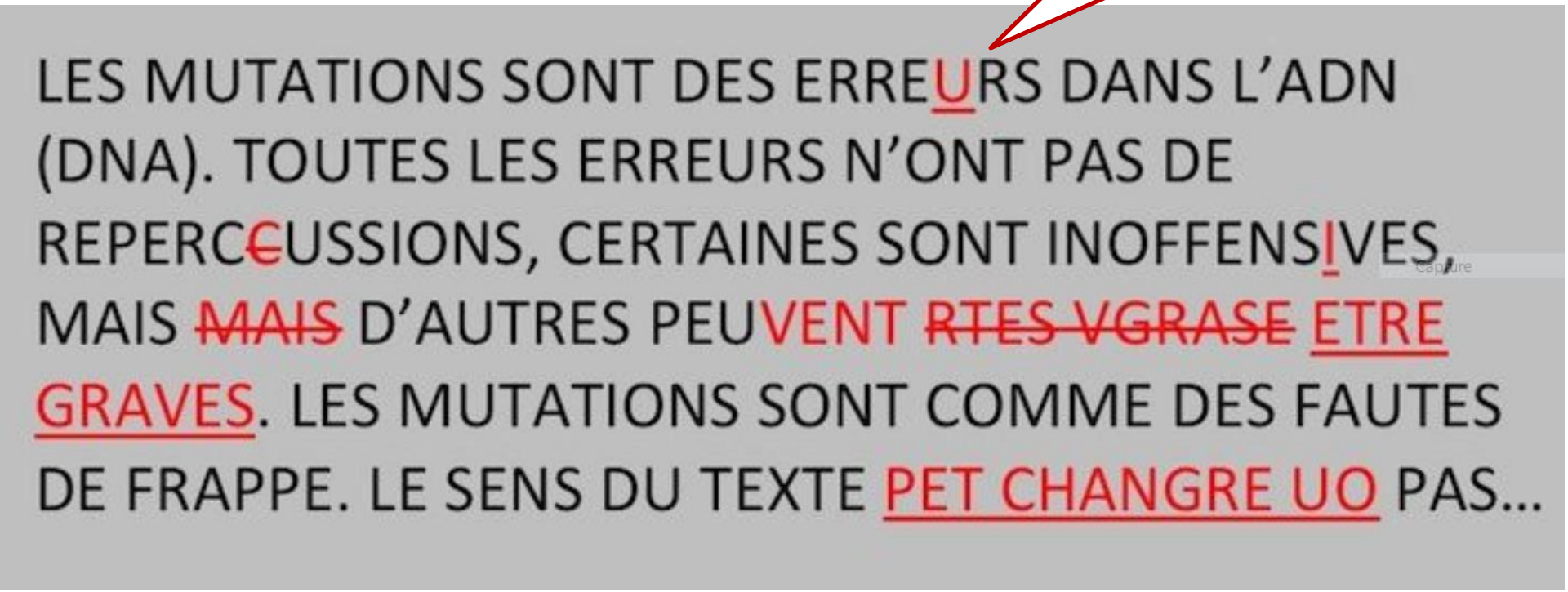
Génétique des population (fréquence)

- Les techniques de séquençage de l'ADN:
  - NGS (Next Generation Sequencing)
  - 10 x coverage (minimum)
  - Validation des variants avec la bonne vieille méthode Sanger….

Sanger Validation of NextGen Sequencing Variants: "Of over 5,800 NGS-derived variants, 19 were not validated by Sanger data. " (PMC4878677)

https://m.simplyscience.ch/le-monde-des-genes/articles/le-monde-

Variations par rapport à une séquence de référence

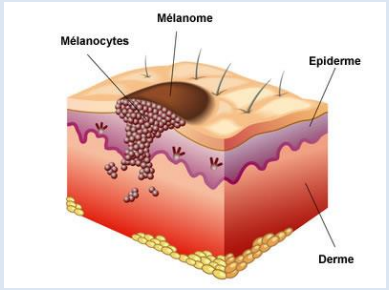


(1) Variations inter-espèces




(2) Variations intra-espèce (humain)

(3) Variations intra-individu (somatiques)



Activité 1.1

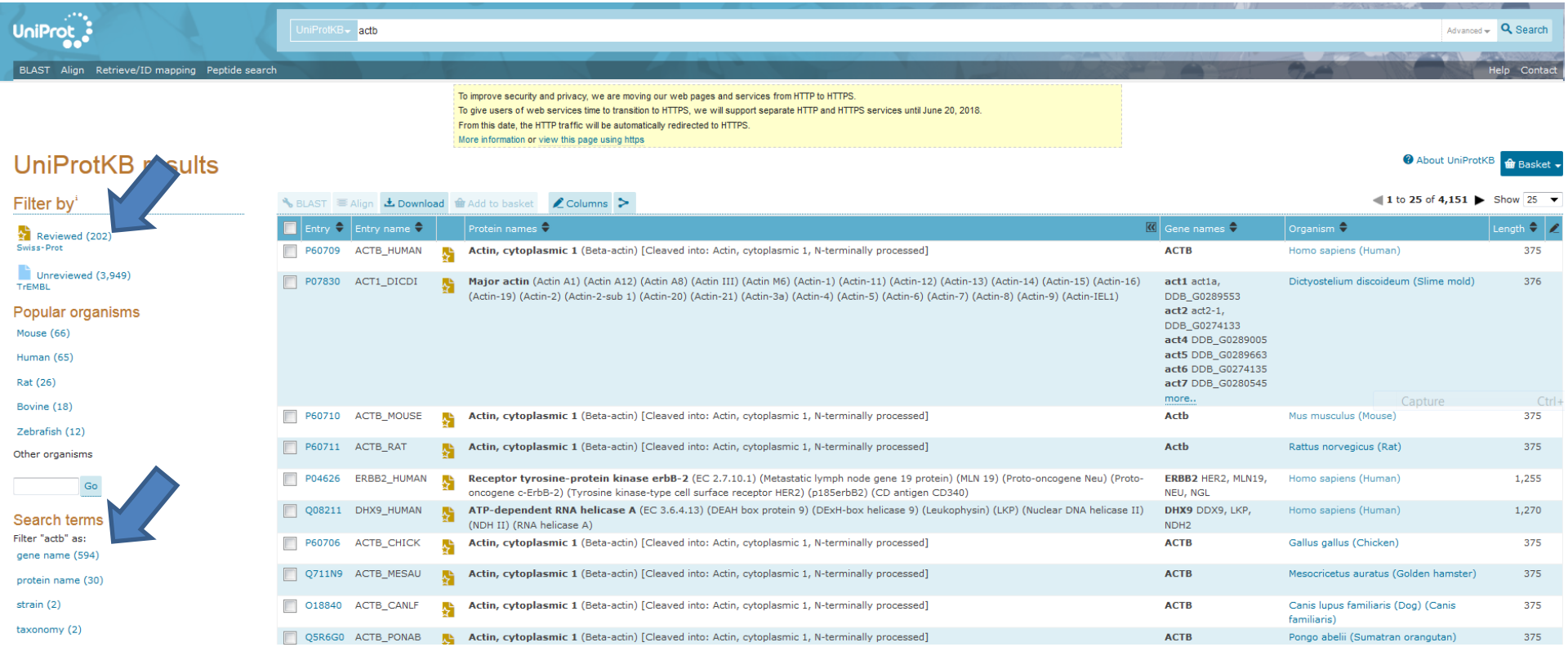
‘Voir’ l’unicité/diversité en comparant des séquences de protéines (1) (Align@UniProt)

‘Voir’ l’évolution en comparant des séquences de protéines….

Aligner les séquences d’une même protéine chez différentes espèces
Choisir différentes protéines; garder les même espèces
www.uniprot.org, align

Depuis le site www.uniprot.org:

Taper un nom de gène (ACTB)
Cliquer sur ‘Reviewed’ et ‘Gene name’

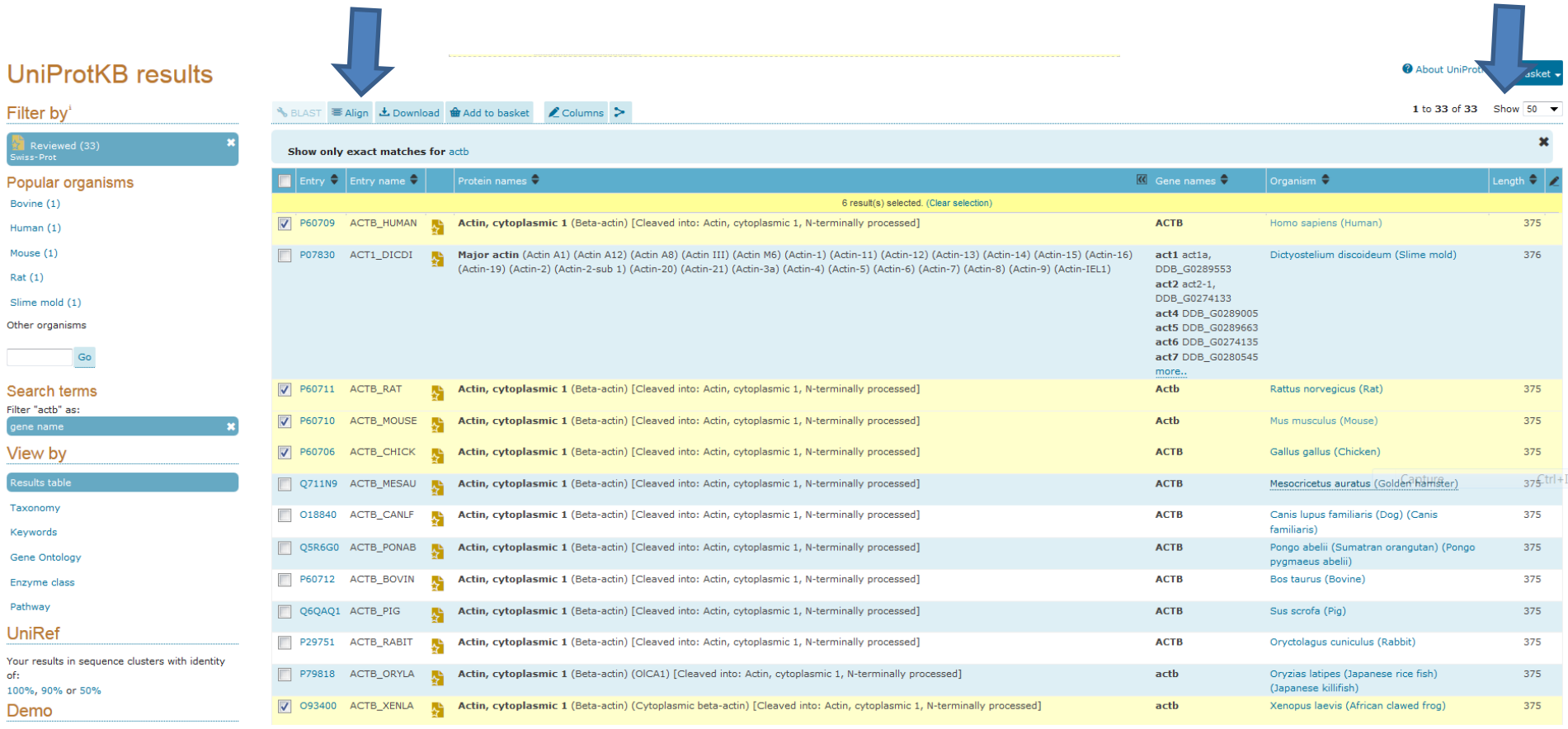


Sélectionner HUMAN, PANTR (Chimp), MOUSE, RAT, XENLA, CHICKEN

Liste des codes des espèces utilisée dans UniProtKB: h[ttp://www.uniprot.org/docs/speclist](http://www.uniprot.org/docs/speclist)

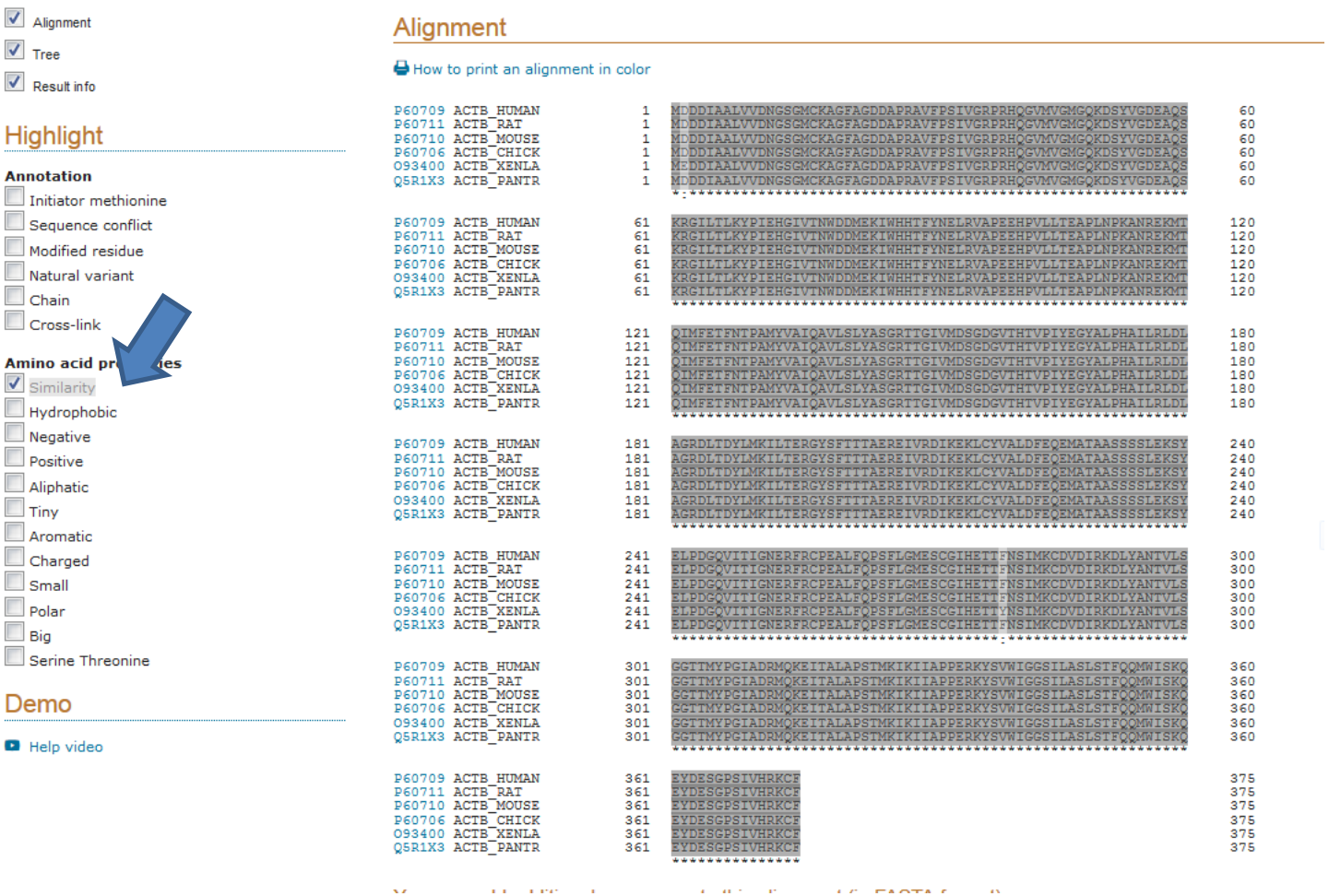
Cliquer sur 'Align'

Show 100

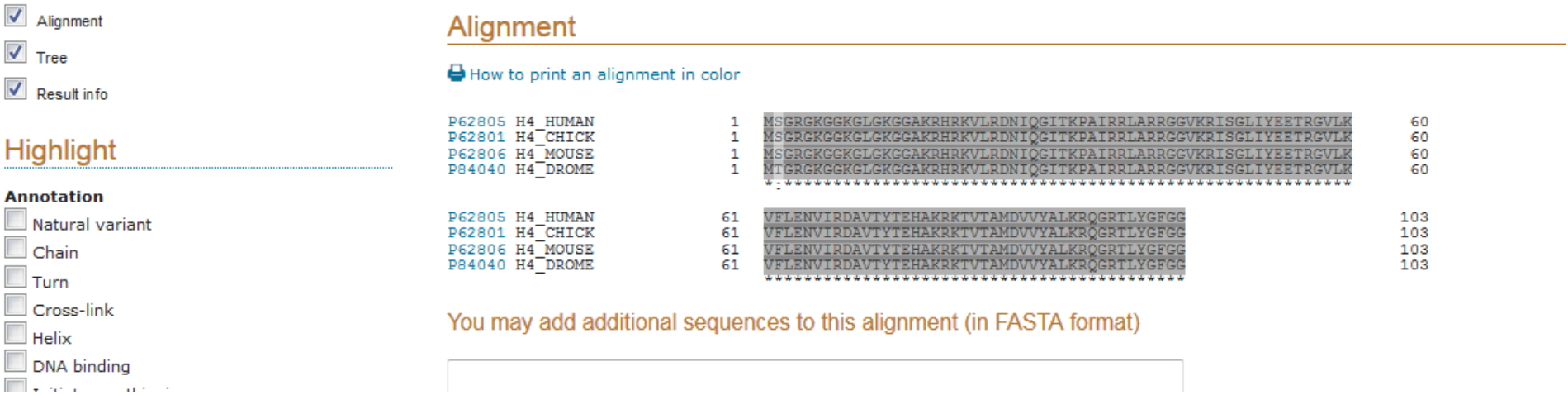


ACTB

Cliquer sur ‘Similarity’



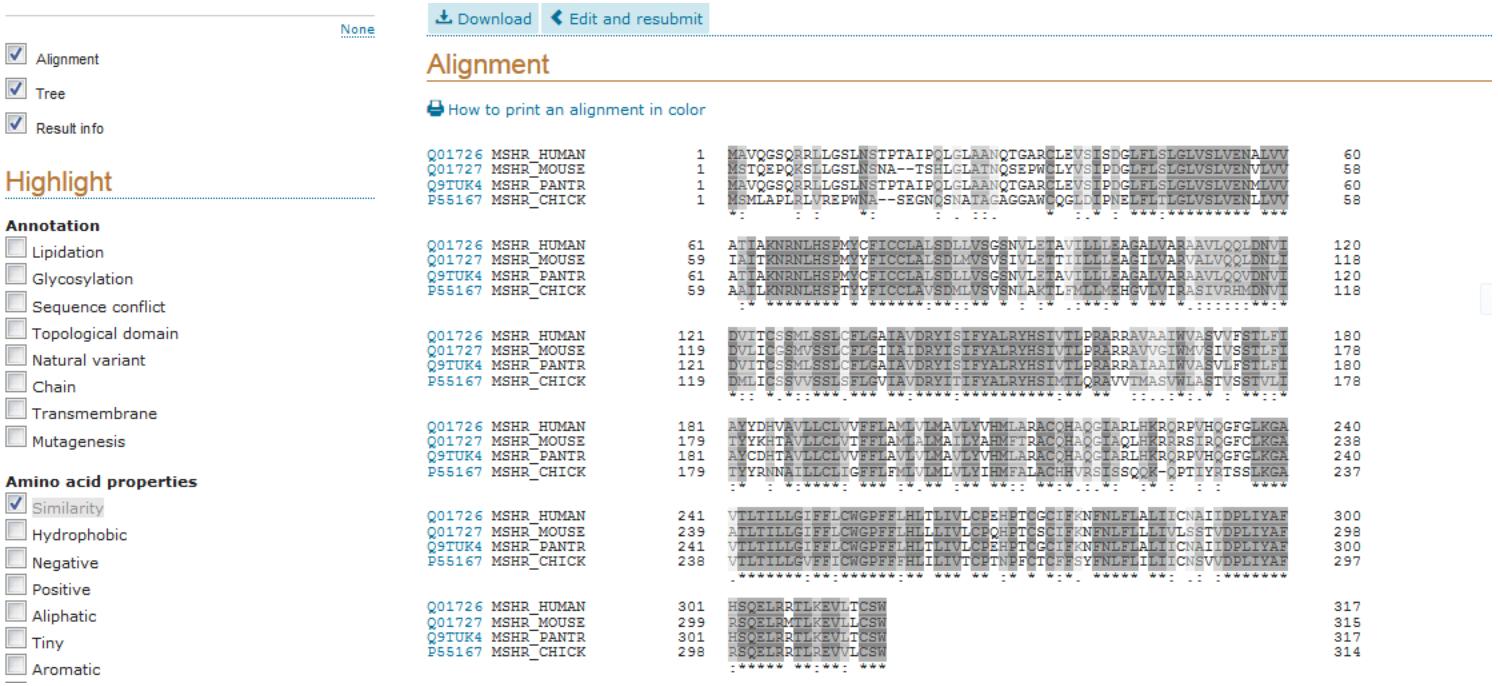
H4



Les gènes évoluent à des vitesse différentes: le gène H4 (un ‘petit’ gène) est particulièrement conservé (Histone H4: beaucoup d’acides aminés chargés positivement (en vert) pour interagir avec l’ADN chargé négativement)

```
P62805 H4_HUMAN      1  MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK  60
P62801 H4_CHICK      1  MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK  60
P62806 H4_MOUSE      1  MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK  60
P84040 H4_DROME      1  MTGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGVLK  60
                        *:************************************************************

P62805 H4_HUMAN     61  VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG                  103
P62801 H4_CHICK     61  VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG                  103
P62806 H4_MOUSE     61  VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG                  103
P84040 H4_DROME     61  VFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG                  103
                        *****************************************
```

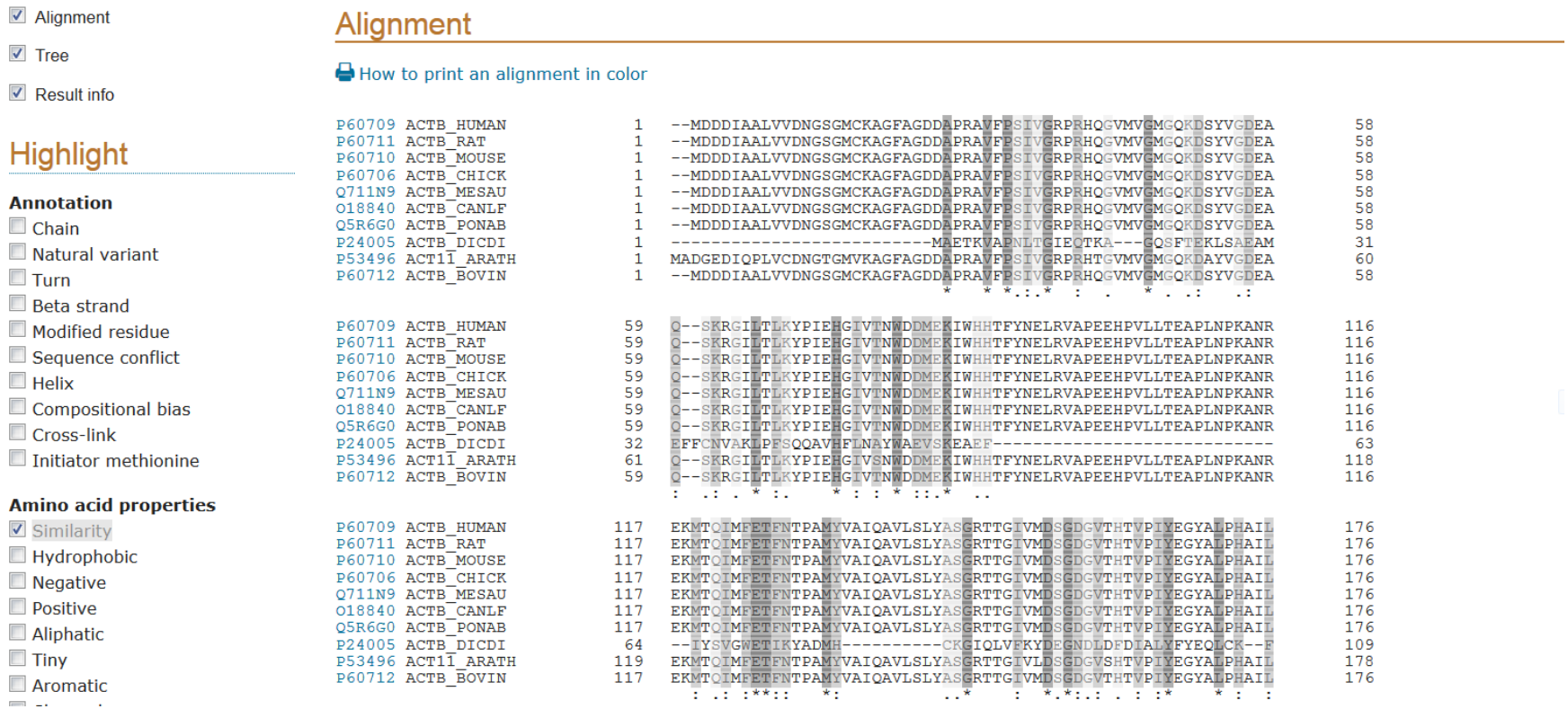
MC1R



Les gènes évoluent à des vitesse différentes: MC1R évolue rapidement….

ACTB

!!!! Dans ce cas, l'alignement est très mauvais, et ne peut pas être interprété, car les séquences sont trop différentes (voir peut-être fausses (pas orthologues pour ARATH (ACTN11) …)



Activité 1.2

‘Voir’ l’unicité/diversité en comparant des séquences de protéines (2) (Align@UniProt)

‘Voir’ l’évolution en comparant des séquences de protéines….

Autre approche en partant des sets de séquences de protéines disponibles ici:
http://education.expasy.org/bioinformatique/Sequences/Liste_prot_evol.html

‘Voir’ l’évolution en comparant des séquences de protéines….

Sélectionner un set de séquence (format fasta)
http://education.expasy.org/bioinformatique/Sequences/Liste_prot_evol.html

Choisir les séquences en fonction des espèces
Liste des codes des espèces utilisée dans UniProtKB: http://www.uniprot.org/docs/speclist

Proposition: ajouter au fur et à mesure les séquences de protéines d’espèces de plus en plus éloignées du point de vue de l’évolution …

Coller les séquences sur le site www.uniprot.org/align

Cliquer sur ‘Run Align’

Depuis la page de résultats, à gauche: cliquer sur ‘Similarity’
Enlever manuellement les séquences qui posent problème (trop différentes).

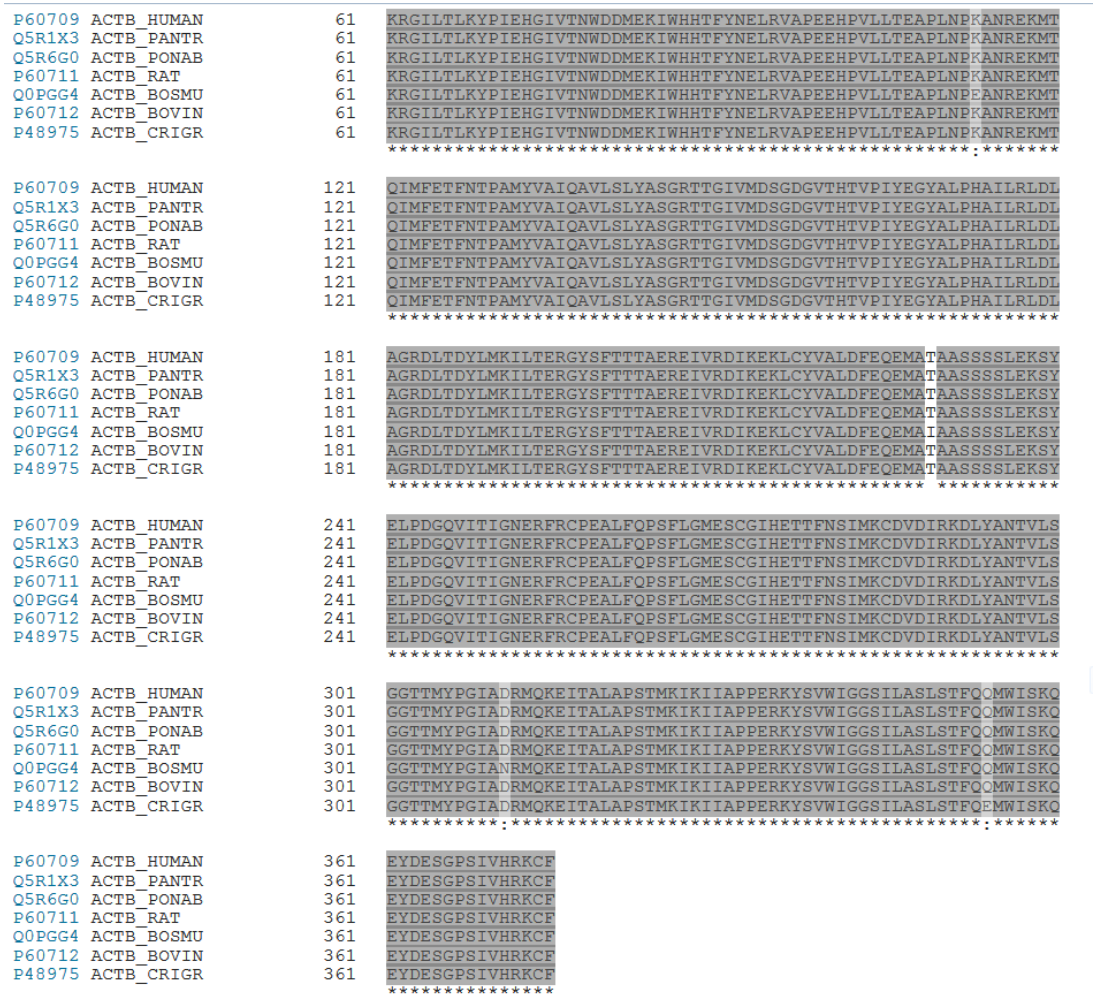
ACTB

primates

```
P60709 ACTB_HUMAN          1   MDDDIAALVVDNGSGMCKAGFAGDDAPRAVFPSIVGRPRHQGVMVGMGQKDSYVGDEAQS
Q5R1X3 ACTB_PANTR          1   MDDDIAALVVDNGSGMCKAGFAGDDAPRAVFPSIVGRPRHQGVMVGMGQKDSYVGDEAQS
Q5R6G0 ACTB_PONAB          1   MDDDIAALVVDNGSGMCKAGFAGDDAPRAVFPSIVGRPRHQGVMVGMGQKDSYVGDEAQS
                               ************************************************************

P60709 ACTB_HUMAN         61   KRGILTLKYPIEHGIVTNWDDMEKIWHHTFYNELRVAPEEHPVLLTEAPLNPKANREKMT
Q5R1X3 ACTB_PANTR         61   KRGILTLKYPIEHGIVTNWDDMEKIWHHTFYNELRVAPEEHPVLLTEAPLNPKANREKMT
Q5R6G0 ACTB_PONAB         61   KRGILTLKYPIEHGIVTNWDDMEKIWHHTFYNELRVAPEEHPVLLTEAPLNPKANREKMT
                               ************************************************************

P60709 ACTB_HUMAN        121   QIMFETFNTPAMYVAIQAVLSLYASGRTTGIVMDSGDGVTHTVPIYEGYALPHAILRLDL
Q5R1X3 ACTB_PANTR        121   QIMFETFNTPAMYVAIQAVLSLYASGRTTGIVMDSGDGVTHTVPIYEGYALPHAILRLDL
Q5R6G0 ACTB_PONAB        121   QIMFETFNTPAMYVAIQAVLSLYASGRTTGIVMDSGDGVTHTVPIYEGYALPHAILRLDL
                               ************************************************************

P60709 ACTB_HUMAN        181   AGRDLTDYLMKILTERGYSFTTTAEREIVRDIKEKLCYVALDFEQEMATAASSSSLEKSY
Q5R1X3 ACTB_PANTR        181   AGRDLTDYLMKILTERGYSFTTTAEREIVRDIKEKLCYVALDFEQEMATAASSSSLEKSY
Q5R6G0 ACTB_PONAB        181   AGRDLTDYLMKILTERGYSFTTTAEREIVRDIKEKLCYVALDFEQEMATAASSSSLEKSY
                               ************************************************************

P60709 ACTB_HUMAN        241   ELPDGQVITIGNERFRCPEALFQPSFLGMESCGIHETTFNSIMKCDVDIRKDLYANTVLS
Q5R1X3 ACTB_PANTR        241   ELPDGQVITIGNERFRCPEALFQPSFLGMESCGIHETTFNSIMKCDVDIRKDLYANTVLS
Q5R6G0 ACTB_PONAB        241   ELPDGQVITIGNERFRCPEALFQPSFLGMESCGIHETTFNSIMKCDVDIRKDLYANTVLS
                               ************************************************************

P60709 ACTB_HUMAN        301   GGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQQMWISKQ
Q5R1X3 ACTB_PANTR        301   GGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQQMWISKQ
Q5R6G0 ACTB_PONAB        301   GGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQQMWISKQ
                               ************************************************************

P60709 ACTB_HUMAN        361   EYDESGPSIVHRKCF
Q5R1X3 ACTB_PANTR        361   EYDESGPSIVHRKCF
Q5R6G0 ACTB_PONAB        361   EYDESGPSIVHRKCF
                               ***************
```

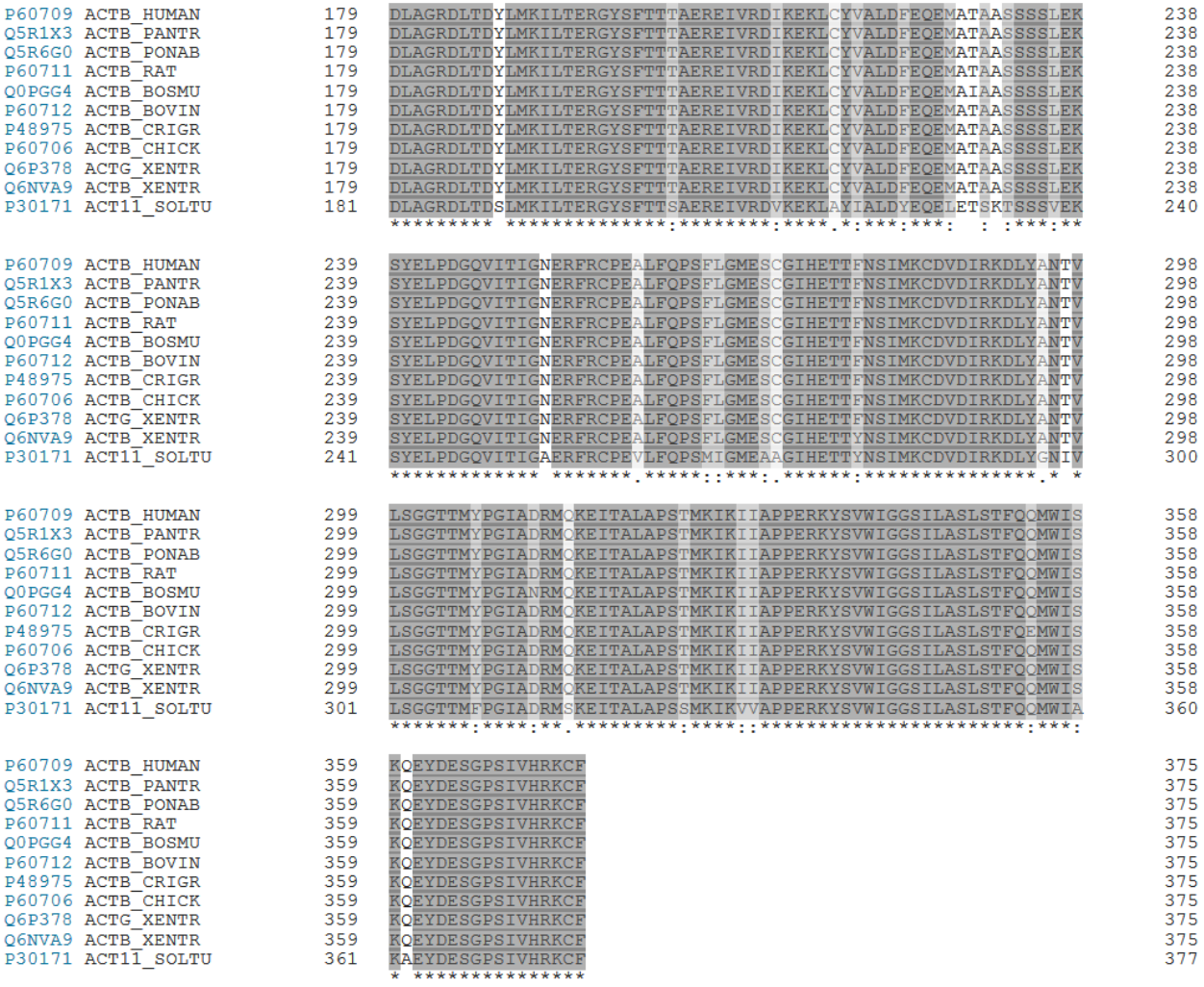
ACTB

mammifères



ACTB

Vertébrés + plante

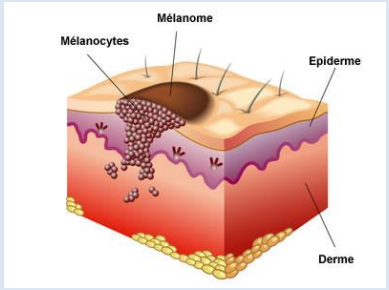


(1) Variations inter-espèces




(2) Variations intra-espèce (humain)

(3) Variations intra-individu (somatiques)



(2) Variation intra-espèces

- **Génome de référence**
- Variations génétiques

Génomes humains ‘en ligne’

- **Genome Reference Consortium (GRCh)**- public (13 individus) (génomes nucléaire + mitochondrial)

- **Craig Venter** – Celera (génome nucléaire)

- **Néanderthal** (génome mitochondrial)

En 2001, les revues Nature et Science publiaient une première analyse de la séquence du génome humain.

Genome Reference Consortium (GRCh)

"As they are often assembled from the sequencing of DNA from a number of donors, reference genomes do not accurately represent the set of genes of any single person.

Instead a reference provides a haploid mosaic of different DNA sequences from each donor. For example, *GRCh38*, the Genome Reference Consortium human genome (build 38) is derived from **thirteen anonymous volunteers** from Buffalo, New York.

The ABO blood group system differs among humans, but the human reference genome contains only an O allele (although the other alleles are annotated)."

Wikipedia

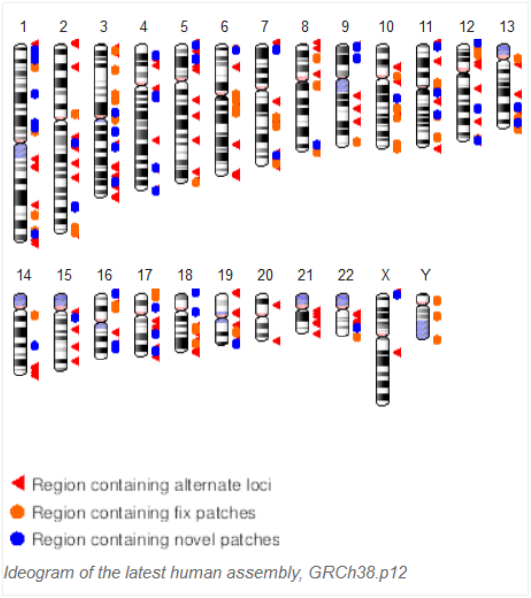
Genome Reference Consortium

GRC Home | Data | Help | Report an Issue | Contact Us | Credits | Curators Only

Human Overview | Human Genome Issues | Human Assembly Data

Human Genome Overview

Information about the continuing improvement of the human genome



Ideogram of the latest human assembly, GRCh38.p12

The GRC is working hard to provide the best possible reference assembly for human. Weas do this by both generating multiple representations (alternate loci) for regions that are too complex to be represented by a single path. Additionally, we are releasing regional fixes known as patches. This allows users who are interested in a specific locus to get an improved representation without affecting users who need chromosome coordinate stability.

Download data:

- GRCh38.p12 (latest minor release) FTP
- GRCh38 (latest major release) FTP
- Genomic regions under review FTP
- Current Tiling Path Files (TPFs)

Transitioning to GRCh38? Try the NCBI Remapping Service, which uses the same assembly-assembly alignments used by the GRC.

Next assembly update
The GRC will review data for the next assembly update in summer 2018.

GRC News

GRCh38.p11: Clinically Relevant Updates to SLC39A4 — *Sep 13, 2017*

GRCh38.p11: Update to GCNT2 — *Jul 11, 2017*

see all

Resolved Human Issues

HG-2087 — *Jan 19, 2018*

HuRef contig ABBA01054858.1 corrects an approximately 3kb deletion in NC_000017.11 component AC026954.14.

HG-2179 — *Jan 19, 2018*

CHM13 component LDOC03008379.1 (GCA_001015385.3) agrees with reference component AP002981.2, suggesting this is

see all

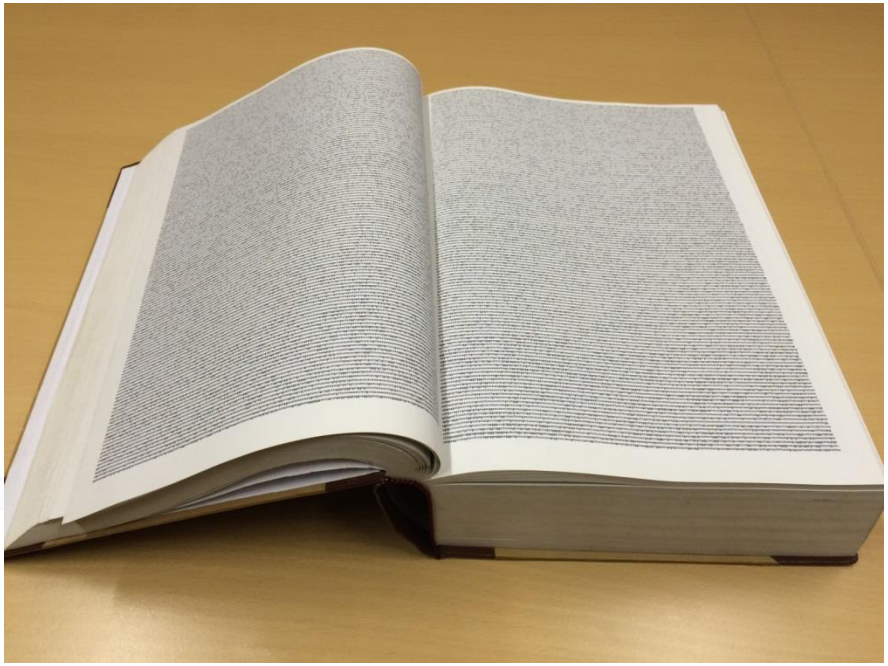
https://www.ncbi.nlm.nih.gov/grc/human

Chromosome 21

le plus petit chromosome humain

~ 1.6 cm d'ADN

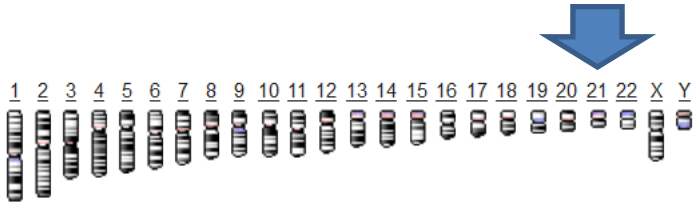
46'709'983 nucléotides



1470 pages

248 nucléotides par ligne

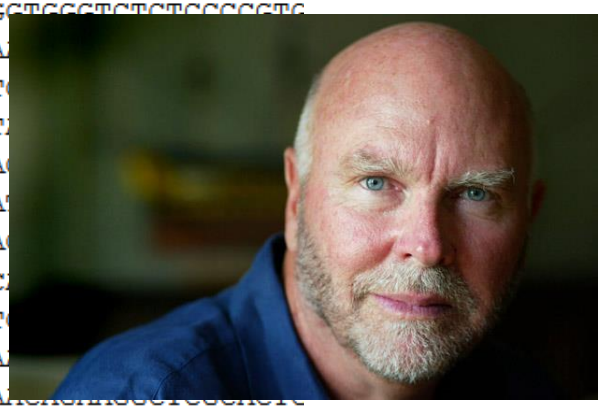
31'775 nucléotides par page



Tous les chromosomes: 3'096'649'726 nucléotides

SIB
Swiss Institute of Bioinformatics

Genome Craig Venter



TCCACACGCTCCTTGCGGCCTCATGGGTGTAGGGTCCAGCCCCACAGGGTCG
GCAGAGACGAGAGAGTGTAGAAATAAAGACACAAGACAAAGAGATTAAAAAA
GGACCACTACCACCAATGCGCGGAGACCGGTAGTGGCCCCGAATGTCTGGCT
TACAAGGCAAAAGGGGCAGGGTAAAGAGTGTGAGTCATCTCCAATGATAGAT
TGTCCACTGGACAGGGGGCCCTTCCCTGCCTGGCAGCCGAGGCAGAGAGGGA
ATAGCTTACGCCATTATTTTTGTATATTAGAGACGTTTAGTACTTTCACTAA
AAGGCAGAGCCAGGTGCACAGGATGGAACATGAAGGAGGACTAGGAGCGTGA
ACAGGGAGACGGTTAGGCCTCCGGATAACTGCGGGCAGGTCTGACTGATGTC
GGAGGAGCAGAGTCTTCTCTAAACTCCCCCGGGGAAAGGGAGCCCCTCCTTT
GGGTGTTTTTCCTTGACACTTACGCTACCGCTAGACCACGGTCCGCTTGGCA
GCTGGCATCACCGCTAGACCAAGGAGCCCTCTAGTGGCCTTGTCCGGGCATA
TGTCTTCTGGTCACTCCTCACTATGTCCCCTCAGCTCCTATCTCTGTATGGCCTGGTTTTTCCTAGTTT
TGATTATAGAGCGAGGATTGTTATAATATTGGAATAAAGAGTAATTGCTACAAACTAATGATTAATGAT
TTCATATATAATCATATCTAAGATCTATATCTGGTGTAACTATTTTTATTTTATATTTTATTATACTGG
ACAGCTCGTGTCCTCAGTCTCTTGCCTCGGCACCTGGGTGGCTTGCCGCCCACAATGGGCAGCTTATTC
TCAGGGAAGGCCTTTGTCTCCACACCTGTGGGTGAAGACCATCGGGATGCTTTGCCTTCAACAGGCAAG
CCAACAATTCACCTTCACTTCCCTCCCTCCAGGAACACCAGCTCCCAGCTCAGAGTCATCGGCCTCGCT
ACAGGGACGTCACACTACCCGCTCTGTGGGGGGCATCGTGTGGTCTGGACTTGCTGAGCAGAAAGTAGC
GCTGCCCTCAACACCTCCCTAGAGCATCTGCGAGCCGAACACCTGGGGCCCACGCCTCCGGCACGTCTA
GGACCCAGTGGTCCATCCCTTCCCAAGCACAAGGCAAGTGGCTACCTCAGTCCCTTCCTCCACGAAGAA
GAGGCACGATGCCTAGTGCTGTAGGTCCCATGTTATTTGGGAAGCAACTTTTGCCCTATTTGGAAGTGC

http://www.ncbi.nlm.nih.gov/nuccore/71514639?report=fasta

This is part of the sequence of Craig Venter chromosome 11
(GenBank database; 3’852’046 bp over 135'006'516 bp)

Neanderthal: ADN mitochondrial

GATCACAGGTCTATCACCCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTTCGTCTGGGGG
GTGTGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCTATGTCGCAGTATCTGTCTTTGATTC
CTGCCCCATTCCATTATTTATCGCACCTACGTTCAATATTACAGGCGAGCATACTTACTGAAGTGTGTTA
ATTAATTAATGCTTGTAGGACATAATAATAACGACTAAATGTCTGCACAGCTGCTTTCCACACAGACATC
ATAACAAAAAATTTCCACCAAACCCCCCCTCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGCCA
AACCCCAAAAACAAAGAACCCTAACACCAGCCTAACCAGATTTCAAATTTTATCTTTTGGCGGTATACAC
TTTTAACAGTCACCCCCTAACTAACACATTATTTTCCCCTCCCACTCCCATACTACTAATCTCATCAATA
CAACCCCCGCCCATCCTACCCAGCACACACCGCTGCTAACCCCATACCCCGAGCCAACCAAACCCCAAAG
ACACCCCCCACAGTTTATGTAGCTTACCTCCTCAAAGCAATACACTGAAAATGTTTAGACGGGCTCACAT
CACCCCATAAACAAATAGGTTTGGTCCTAGCCTTTCTATTAGCTCTTAGTAAGATTACACATGCAAGCAT
CCCCATTCCAGTGAGTTCACCCTCTAAATCACCACGATCAAAAGGGACAAGCATCAAGCACGCAACAATG
CAGCTCAAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATAAGCCTTTAGCAATAAACGA
AAGTTTAACTAAGCTATACTAACCCCAGGGTTGGTCAATTTCGTGCCAGCCACCGCGGTCACACGATTAA
CCCAAGTCAATAGAAGCCGGCGTAAAGAGTGTTTTAGATCACCCCCTCCCCAATAAAGCTAAAACTCACC
TGAGTTGTAAAAAACTCCAGTTGACACAAAATAAACTACGAAAGTGGCTTTAACATATCTGAACACACAA
TAGCTAAGACCCAAACTGGGATTAGATACCCCACTATGCTTAGCCCTAAACCTCAACAGTTAAATCAACA
AAACTGCTCGCCAGAACACTACGAGCCACAGCTTAAAACTCAAAGGACCTGGCGGTGCTTCATATCCCTC
TAGAGGAGCCTGTTCTGTAATCGATAAACCCCGATCAACCTCACCACCTCTTGCTCAGCCTATATACCGC
CATCTTCAGCAAACCCTGATGAAGGCTACAAAGTAAGCGCAAGTACCCACGTAAAGACGTTAGGTCAAGG
TGTAGCCCATGAGGTGGCAAGAAATGGGCTACATTTTCTACCCCAGAAAACTACGATAGCCCTTATGAAA
CCTAAGGGTCGAAGGTGGATTTAGCAGTAAACTGAGAGTAGAGTGCTTAGTTGAACAGGGCCCTGAAGCG
CGTACACACCGCCCGTCACCCTCCTCAAGTATACTTCAAAGGACATTTAACTAAAACCCCTACGCATTTA
TATAGAGGAGACAAGTCGTAACATGGTAAGTGTACTGGAAAGTGCACTTGGACGAACCAGAGTGTAGCTT
AACACAAAGCACCCAACTTACACTTAGGAGATTTCAACTTAACTTGACCGCTCTGAGCTAAACCTAGCCC
CAAACCCACTCCACCTTACTACCAAACAACCTTAGCCAAACCATTTACCCAAATAAAGTATAGGCGATAG
AAATTGAAACCTGGCGCAATAGATGTAGTACCGCAAGGGAAAGATGAAAAATTATAACCAAGCATAATAT
AGCAAGGACTAACCCCTATACCTTCTGCATAATGAATTAACTAGAAATAACTTTGCAAGGAGAGCCAAAG
CTAAGACCCCCGAAACCAGACGAGCTACCTAAGAACAGCTAAAAGAGCACACCCGTCTATGTAGCAAAAT
AGTGGGAAGATTTATAGGTAGAGGCGACAAACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAGATAGAA

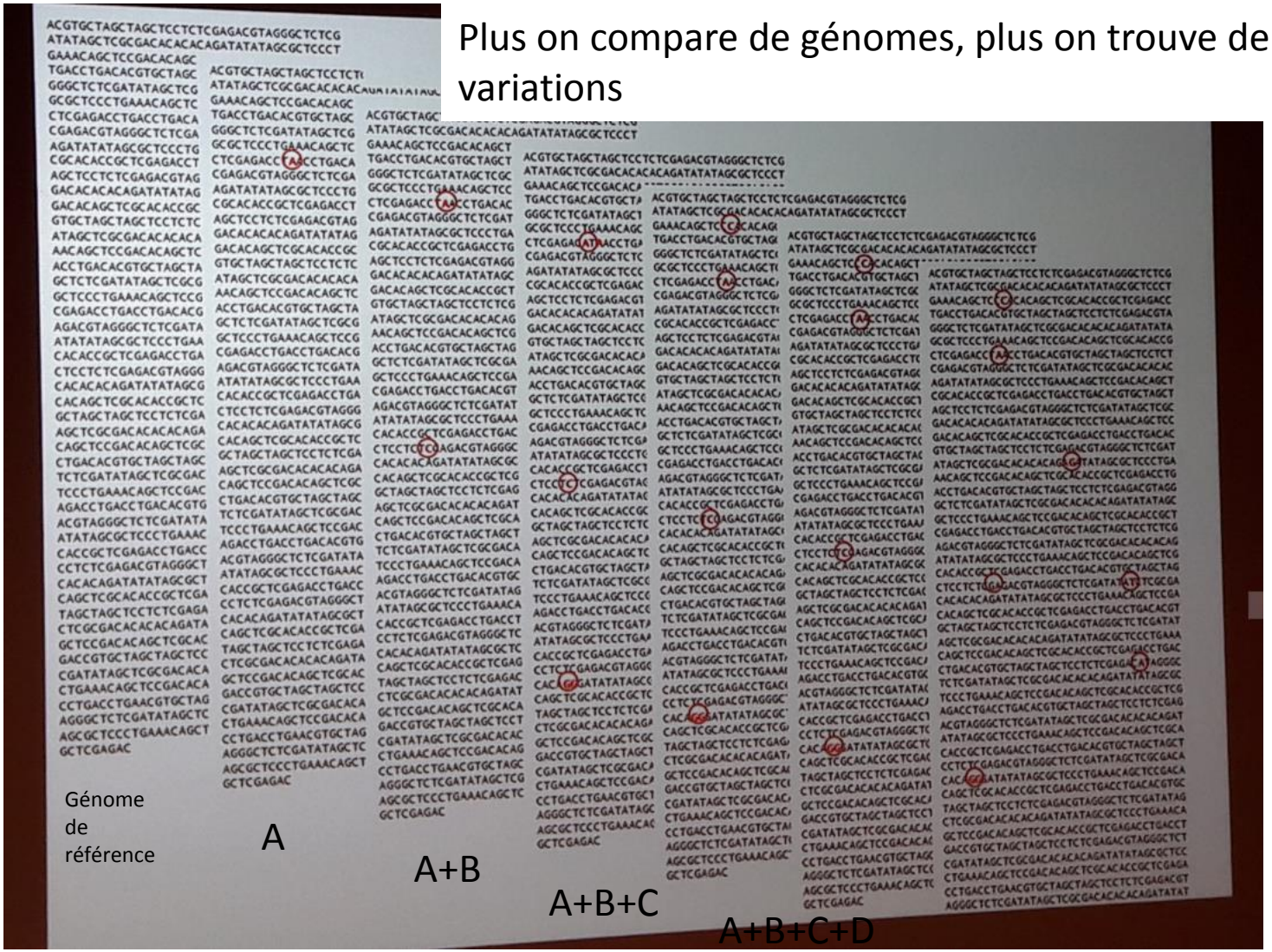
(2) Variation intra-espèces

- Génome de référence
- **Variations génétiques**

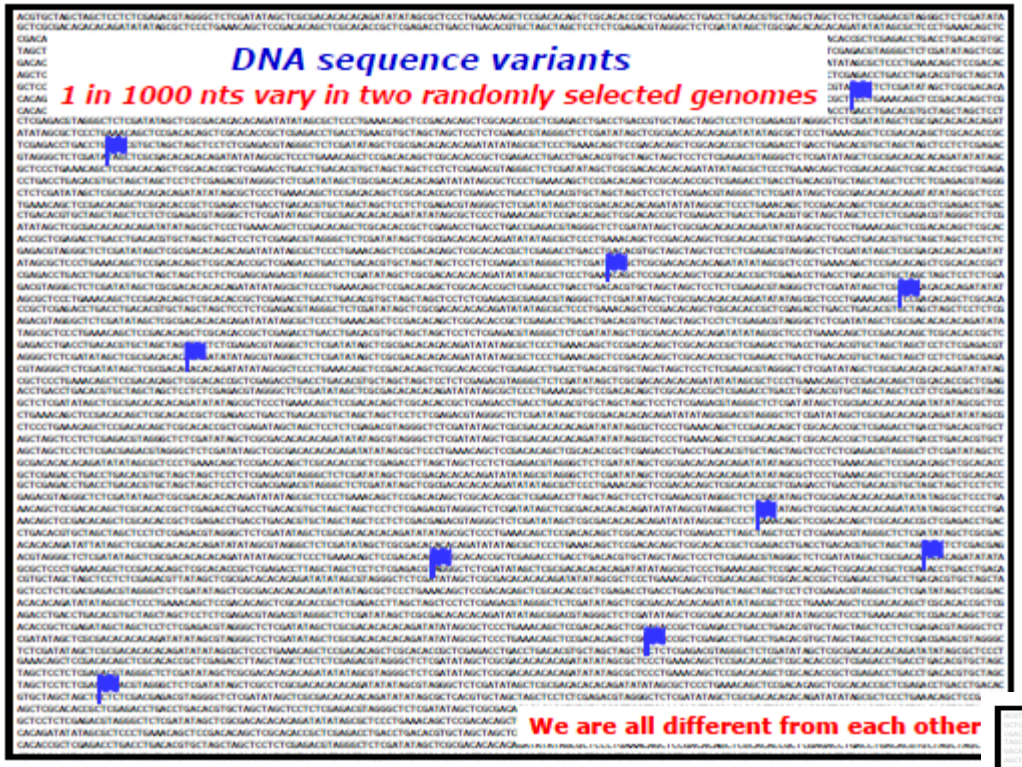
- Toutes les séquences des génomes humains ne sont pas disponibles en ligne (seulement les génomes de référence)

- Toutes les variations sont rapportées au génome de référence

Plus on compare de génomes, plus on trouve de variations



S. Antonarakis, nov 2017



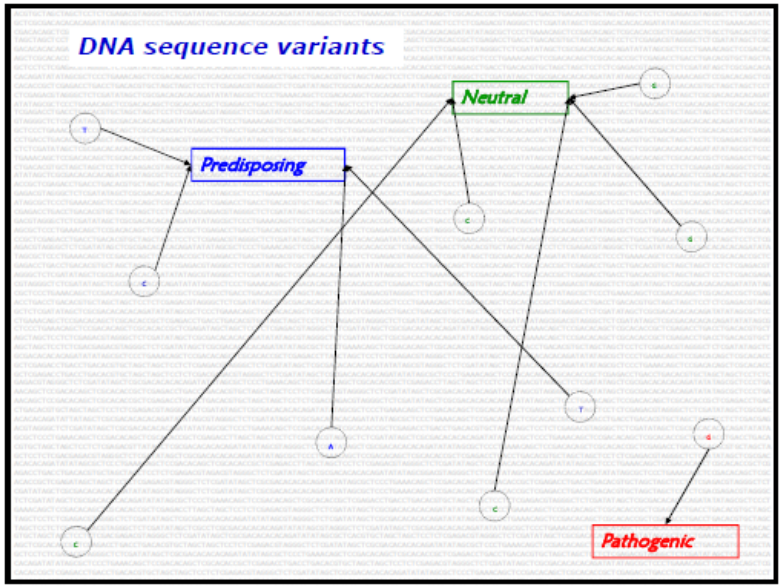
Quelle(s) variation(s), quel(s) effet(s)?

tous différents…

~ 3.3 millions de variations entre 2 individus

~10 millions de variations dans la population

- neutres (la majorité)
- **associées** à un phénotype particulier…
- **associées** à une prédisposition **pathogéniques**…
- variants of unknown significance (**VUS**)



S.E. Antonarakis, UNIGE

- Taux de mutation à chaque génération: $1/10^8$
- En moyenne un nouveau-né porte entre 50 et 100 ‘nouvelles mutations’
- Ces mutations sont importantes: possibilité d’évolution et d’adaptation (i.e. résistance au HIV)
- Un prix à payer: nouvelles pathologies

S. Antonarakis, nov 2017

- Homozygote
  - Si les deux locus ont chacun un allèle mutant différent, on parle d'« hétérozygote *composite* »

- Hétérozygote

- A mutation affecting only one allele is called *heterozygous*. A *homozygous* mutation is the presence of the identical mutation on both alleles of a specific gene. However, when both alleles of a gene harbor mutations, but the mutations are different, these mutations are called *compound heterozygous*.

Clinical significance value (n=5)

GWAS, génétique des populations, analyse fonctionnelle

Clinical significance value	Guidance for use in ClinVar SCV records
Benign	As recommended by ACMG/AMP for variants interpreted for Mendelian disorders.
Likely benign	As recommended by ACMG/AMP for variants interpreted for Mendelian disorders.
Uncertain significance	As recommended by ACMG/AMP for variants interpreted for Mendelian disorders.
Likely pathogenic	As recommended by ACMG/AMP for variants interpreted for Mendelian disorders.
Pathogenic	As recommended by ACMG/AMP for variants interpreted for Mendelian disorders. Variants that have low penetrance may be submitted as "Pathogenic"; please also include information about the penetrance in a "Comment on clinical significance".
drug response	A general term for a variant that affects a drug response, not a disease. We anticipate adding more specific drug response terms based on a recommendation by CPIC.
association	For variants identified in a GWAS study and further interpreted for their clinical significance.
risk factor	For variants that are interpreted not to cause a disorder but to increase the risk.
protective	For variants that decrease the risk of a disorder, including infections.
Affects	For variants that cause a non-disease phenotype, such as lactose intolerance.

90%

(VUS)

90%

https://www.ncbi.nlm.nih.gov/**clinvar**/docs/clinsig/

- More than 10,000 monogenic inherited disorders have been identified, affecting millions of people worldwide.
- The global prevalence of all single gene diseases at birth is approximately 10/1000.
- In Canada, it has been estimated that taken together, monogenic diseases may account for upto 40% of the work of hospital based paediatric practice (Scriver, 1995).

http://www.who.int/genomics/public/geneticdiseases/en/index2.html

Fatal familial insomnia (prion)

"If you have a genetic risk that you believe is predicting disease but isn't, you can end up doing drastic things."


NEWS FEATURE

SEEING DEADLY MUTATIONS IN A NEW LIGHT

How one of the largest genome resources in the world has quietly been changing scientists' understanding of human genetics.

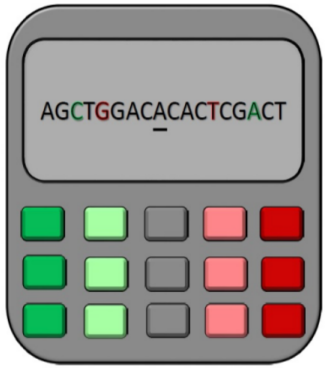
BY ERIKA CHECK HAYDEN

http://education.expasy.org/cours/PO17421/publications/

Bioinformatique

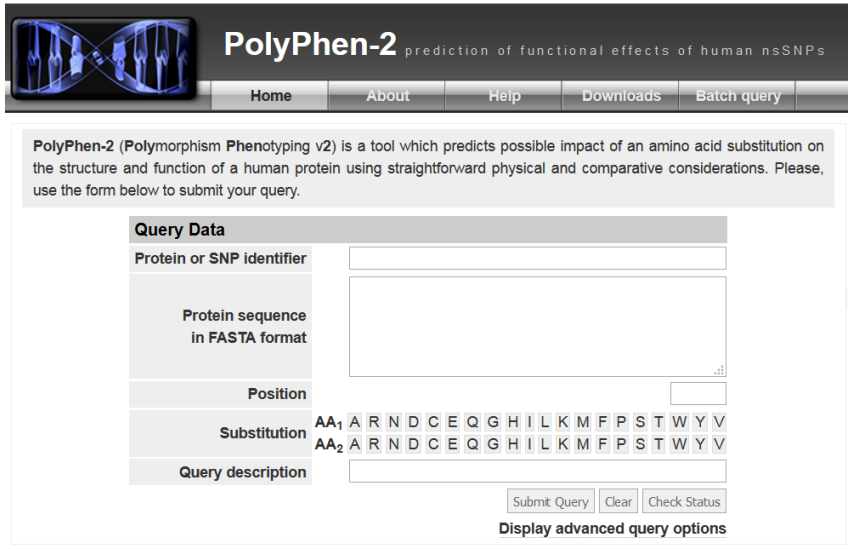
Analyse des variants…

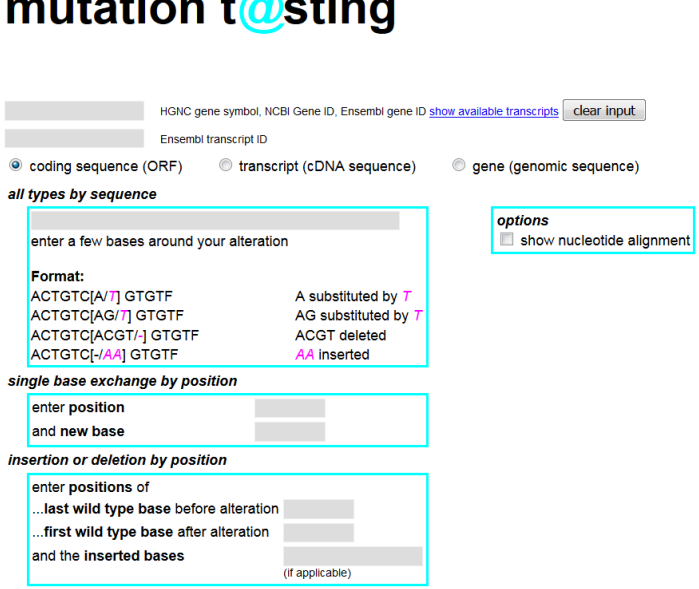
- Outils de prédiction
  - PolyPhen
  - Mutalyzer
  - Mutation taster
  - Pathogenicity calculator (aide à la décision)
- Banques de données
  - dbSNP / dbVar / ClinVar
  - ExAC (Exome Aggregation Consortium; 60'000 personnes)
  - UniProtKB/Swiss-Prot
  - …



WHAT IS THE CLINGEN PATHOGENICITY CALCULATOR?

The shift from genetic testing of individual genes to exome and genome sequencing has been accompanied by new challenges in genome interpretation. The American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) have published **Standards and Guidelines for the Interpretation of Sequence Variants**. To enable wide application of the ACMG/AMP and similar guidelines and the development of collective knowledge by the community, ClinGen has developed the ClinGen Pathogenicity Calculator. By automating the formal reasoning, the Calculator eliminates errors in rule application and makes it possible to automatically calculate provisional conclusions based on latest evidence. Moreover, the Calculator makes reasoning explicit by





Analyse manuelle
(@UniProtKB/Swiss-Prot)

rs4240199 *[Homo sapiens]*

1.

GGAATATACCATTCTGTTAGGACTT[A/G]GGGGAAAAAAGATCTTTTAGGAGGT

Chromosome:	2:85322745
Gene:	TGOLN2 (GeneView)
Functional Consequence:	intron variant,missense,stop gained
Validated:	by 1000G,by cluster,by frequency
Global MAF:	A=0.2210/1107
HGVS:	NC_000002.11:g.85549868A>G, NC_000002.12:g.85322745A>G, NG_030379.1:g.10552C, NG_030379.1:g.10552C>T, NM_001206840.1:c.1363C, NM_001206840.1:c.1363C>T, NM_001206841.1:c.1322C, NM_001206841.1:c.1322C>T, NM_001206844.1:c.1135-4T>C, NM_006464.3:c.1309-4T>C, NP_001193769.1:p.Gln455, NP_001193769.1:p.Gln455Ter, NP_001193770.1:p.Pro441, NP_001193770.1:p.Pro441Leu

O43493; JBC 1998

- **dbSNP**: short genetic variations (single nucleotide polymorphisms) present in more than 1% of the population (-> not disease associated)
- **dbVar**: structural variation — insertions, deletions, duplications, inversions, mobile elements, and translocations

- **ClinVar**: aggregates information about genomic variation and its relationship with human health

!!! Financement des banques de données !!!

Due to the lapse in government funding, the information on this web site may not be up to date, transactions submitted via the web site may not be processed, and the agency may not be able to respond to inquiries until appropriations are enacted. Updates regarding government operating status and resumption of normal operations can be found at opm.gov.

ClinVar

ClinVar aggregates information about genomic variation and its relationship to human health.

Using ClinVar

- About ClinVar
- Data Dictionary
- Downloads/FTP site
- FAQ
- Contact Us
- RSS feed/What's new?
- Factsheet

Tools

- ACMG Recommendations for Reporting of Incidental Findings
- ClinVar Submission Portal
- Submissions
- Variation Viewer
- Clinical Remapping - Between assemblies and RefSeqGenes
- RefSeqGene/LRG

Related Sites

- ClinGen
- GeneReviews®
- GTR®
- MedGen
- OMIM®
- Variation

Les variations ‘missense’ & UniProtKB/Swiss-Prot

- http://www.uniprot.org/docs/humsavar

  30’182 disease variants, 39’800 polymorphisms, 7’600 unclassified

CFTR	P13569	VAR_000101	p.Ser13Phe	Disease	rs397508635	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000102	p.Arg31Cys	Polymorphism	rs1800073	-
CFTR	P13569	VAR_000103	p.Arg31Leu	Disease	rs149353983	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000104	p.Ser42Phe	Disease	rs143456784	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000105	p.Asp44Gly	Disease	rs1800074	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000106	p.Asp44Val	Polymorphism	rs1800074	-
CFTR	P13569	VAR_000107	p.Ser50Tyr	Disease	rs397508220	Congenital bilateral absence of the vas deferens (CBAVD) [MIM:277180]
CFTR	P13569	VAR_000108	p.Trp57Gly	Disease	rs397508272	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000109	p.Pro67Leu	Disease	rs368505753	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000110	p.Arg74Trp	Disease	rs115545701	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000111	p.Arg75Gln	Polymorphism	rs1800076	-
CFTR	P13569	VAR_000112	p.Gly85Glu	Disease	rs75961395	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000113	p.Phe87Leu	Disease	rs397508403	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000114	p.Gly91Arg	Disease	rs121908750	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000115	p.Glu92Lys	Disease	rs121908751	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000116	p.Gln98Arg	Disease	rs397508464	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000117	p.Ile105Ser	Disease	-	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000118	p.Tyr109Cys	Disease	rs121909031	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000119	p.Asp110His	Disease	rs113993958	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000120	p.Pro111Leu	Disease	rs140502196	Congenital bilateral absence of the vas deferens (CBAVD) [MIM:277180]
CFTR	P13569	VAR_000121	p.Arg117Cys	Disease	rs77834169	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000122	p.Arg117His	Disease	rs78655421	Congenital bilateral absence of the vas deferens (CBAVD) [MIM:277180]
CFTR	P13569	VAR_000122	p.Arg117His	Disease	rs78655421	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000123	p.Arg117Leu	Disease	rs78655421	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000124	p.Arg117Pro	Disease	rs78655421	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000125	p.Ala120Thr	Disease	rs201958172	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000126	p.His139Arg	Disease	rs76371115	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000127	p.Ala141Asp	Disease	rs397508700	Cystic fibrosis (CF) [MIM:219700]

Les données sont en constante évolution

- apoE: génotype 4/4 associé avec un risque d'hyperlipémie -> 1993

  Depuis 1993: associé avec un risque d'Alzheimer familial précoce !

*"About 60% of the AD patients in the Caucasian population are APOE ε4 carriers and this reduces the age of onset of the disease in a gene dose-dependent manner by as much as 7 to 9 years per allele copy" (*Science, (1993), PMID: 8346443)

- TASR38: goût amer (PTC)

  Depuis quelques années: associé avec un risque de développer un cancer colorectal et des sinusites… (http://education.expasy.org/cours/PO17421/publication/)

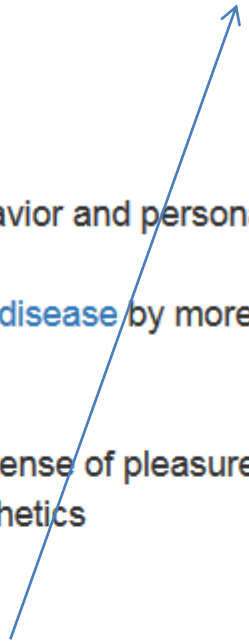
- Variant of 'unknown significance' in 2016: en 2020 ?

Cependant, il a été trouvé le variant *BRCA1* c.2002C>T / p.Leu668Phe dont la répercussion fonctionnelle n'est pas clairement établie. Ce variant a été référencé dans la base de données internationale du BIC (http://research.nhgri.nih.gov/bic/), et est classé comme «unknown clinical significance». Selon les principaux algorithmes de prédiction ainsi qu'une récente publication de test fonctionnel *in vitro* ce variant serait prédit comme neutre (P. Bouwman et al., Cancer Discovery, 2013). Ce variant ne peut pas être utilisé dans le cadre d'un dépistage informatif pour les autres membres de la famille tant que sa nature pathogène ou polymorphique n'a pas été confirmée.

- Autres sites intéressants…

SNPedia
https://www.snpedia.com/

http://www.chromosomewalk.ch/chromosome/chromosome-15/



Popular [edit]

- rs53576 in the oxytocin receptor influences social behavior and personality
- rs1815739 muscle performance
- rs7412 and rs429358 can raise the risk of Alzheimer's disease by more than 10x
- rs6152 can influence baldness
- rs333 resistance to HIV
- rs1800497 in a dopamine receptor may influence the sense of pleasure
- rs1805007 determines red hair and sensitivity to anesthetics
- rs9939609 triggers obesity and type-2 diabetes
- rs662799 prevents weight gain from high fat diets
- rs7495174 green eye color and rs12913832 for blue eye color
- rs7903146 in 3% of the population greatly increases the risk of type-2 diabetes
- rs12255372 linked to type-2 diabetes and breast cancer
- rs1799971 makes alcohol cravings stronger
- rs17822931 determines earwax, sweating and body odor
- rs4680 varied cognitive effects
- rs1333049 coronary heart disease
- rs1051730 and rs3750344 nicotine dependence
- rs4988235 lactose intolerance

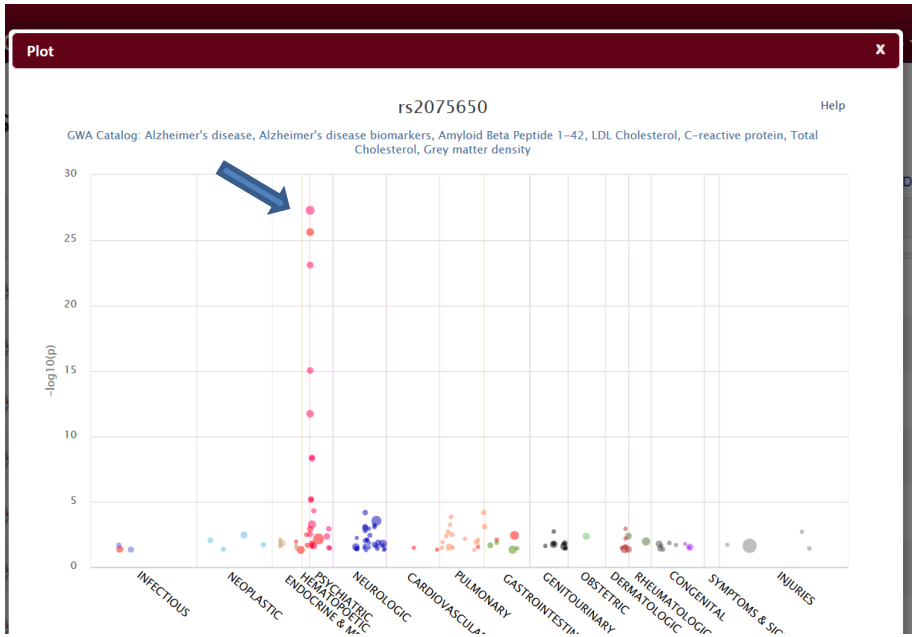
View all 107819 snps in SNPedia.

GWAS phenotype

https://phewascatalog.org/phewas

Genome wide associate studies (GWAS) are a common method used in associating single nucleotide polymorphisms (SNPs) to a disease or trait under study.
For a case/control phenotype, you're basically looking to see if there are more risk alleles in the case population than the control population (i.e. log (p)) (PMC3154648)

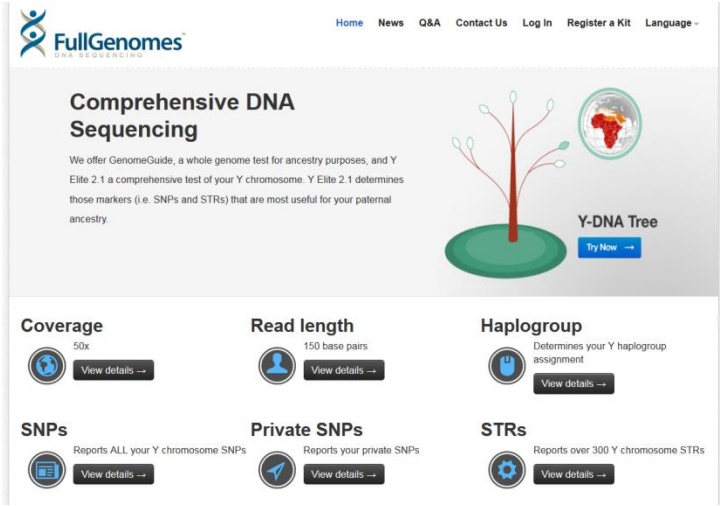
La variation rs2075650 est retrouvée plus fréquemment chez des personnes avec Alzheimer: cela ne signifie pas qu'elle est la 'cause' du problème



Séquencer mon génome ?

Mon génome…. (celui d'un collègue)

- https://www.fullgenomes.com/ (juin 2017)
- 800 dollards, 20x coverage



- Après 6 mois: répertoire de 50 Gb
  - Variants autosomes
  - Variants chromosome Y
  - Séquence ADN mitochondrie (http://education.expasy.org/cours/PO17421/)

```
# The following table is formatted as tab separated values.
#GeneId GeneName        BioType Bases affected (DOWNSTREAM)     Total score (DOWNSTREAM)        Length (DOWNSTREAM)     Bases affected (EXON)   Total
xon_loss_variant)       Count (5_prime_UTR_variant)     Count (disruptive_inframe_deletion)     Count (disruptive_inframe_deletion+synonymous_variant)
etained_variant)        Count (synonymous_variant)      Count (upstream_gene_variant)
ENSG00000000003 TSPAN6  protein_coding  11      0       9772    2       0       2968    10      0       11322   8       0       8354    6       0
ENSG00000000005 TNMD    protein_coding  21      0       7354    1       0       1610    19      0       15084   18      0       13474   19      0
ENSG00000000419 DPM1    protein_coding  20      0       6281    1       0       1207    5       0       23689   5       0       22482   0       0
ENSG00000000457 SCYL3   protein_coding  29      0       14488   2       0       6876    86      0       44637   84      0       37761   64      0
ENSG00000000460 C1orf112        protein_coding  49      0       27580   7       0       6354    531     0       191977  526     0       185623  0
ENSG00000000938 FGR     protein_coding  15      0       13473   0       0       3474    23      0       23214   23      0       19740   10      0
ENSG00000000971 CFH     protein_coding  56      0       20002   14      0       8144    299     0       95627   294     0       87483   0       0
ENSG00000001036 FUCA2   protein_coding  39      0       12134   1       0       3119    2       0       16880   28      0       13761   0       0
ENSG00000001084 GCLC    protein_coding  31      0       26933   9       0       8463    144     0       119630  138     0       111167  0       0
ENSG00000001167 NFYA    protein_coding  4       0       7026    0       0       3811    0       0       27032   0       0       23221   2       0
ENSG00000001460 STPG1   protein_coding  67      0       30736   15      0       8514    120     0       59936   118     0       51422   52      0
ENSG00000001461 NIPAL3  protein_coding  44      0       24375   1       0       9404    181     0       57183   173     0       47779   41      0
ENSG00000001497 LAS1L   protein_coding  4       0       10001   1       0       5589    7       0       22194   6       0       16605   0       0
ENSG00000001561 ENPP4   protein_coding  26      0       5000    1       0       4651    50      0       16707   40      0       12056   10      0
ENSG00000001617 SEMA3F  protein_coding  25      0       19214   2       0       4826    46      0       34031   43      0       29205   41      0
ENSG00000001626 CFTR    protein_coding  11      0       31738   3       0       8099    90      0       250188  89      0       242089  63      0
ENSG00000001629 ANKIB1  protein_coding  13      0       28547   1       0       7130    58      0       155151  55      0       148021  28      0
ENSG00000001630 CYP51A1 protein_coding  4       0       11543   1       0       4074    13      0       30802   12      0       26728   4       0
ENSG00000001631 KRIT1   protein_coding  17      0       26192   0       0       6204    29      0       47198   29      0       40994   25      0
ENSG00000002016 RAD52   protein_coding  38      0       19102   4       0       4190    221     0       77977   224     0       73787   9       0
ENSG00000002079 MYH16   pseudogene      32      0       31626   6       0       6387    89      0       72337   84      0       65950   0       0
ENSG00000002330 BAD     protein_coding  10      0       10478   0       0       1708    0       0       14875   0       0       13167   0       0
ENSG00000002549 LAP3    protein_coding  55      0       17220   2       0       3785    122     0       30781   118     0       26996   3       0
ENSG00000002586 CD99    protein_coding  99      0       23222   10      0       3284    203     0       50131   197     0       46847   200     0
ENSG00000002587 HS3ST1  protein_coding  10      0       10000   0       0       8166    62      0       36616   55      0       28450   0       0
ENSG00000002726 AOC1    protein_coding  25      0       9287    3       0       3824    97      0       36878   88      0       33054   10      0
ENSG00000002745 WNT16   protein_coding  9       0       6646    2       0       3261    37      0       15738   34      0       12477   0       0
ENSG00000002746 HECW1   protein_coding  123     0       40823   8       0       13965   972     0       453403  960     0       439438  280     0
ENSG00000002822 MAD1L1  protein_coding  133     0       57916   8       0       6955    1359    0       417450  1357    0       410495  1356    0
ENSG00000002834 LASP1   protein_coding  11      0       16641   8       0       7061    141     0       51912   131     0       44851   125     0
ENSG00000002919 SNX11   protein_coding  55      0       15174   3       0       3606    64      0       19718   68      0       16112   27      0
ENSG00000002933 TMEM176A        protein_coding  17      0       7831    15      0       3479    1       0       4718    17      0       1239    14
ENSG00000003056 M6PR    protein_coding  3       0       10122   1       0       3777    6       0       9593    3       0       5816    3       0
ENSG00000003096 KLHL13  protein_coding  3       0       10759   0       0       6983    132     0       219528  132     0       212545  4       0
ENSG00000003137 CYP26B1 protein_coding  8       0       11042   0       0       5015    14      0       18801   12      0       13786   0       0
ENSG00000003147 ICA1    protein_coding  173     0       35468   23      0       4891    487     0       149504  484     0       144613  264     0
ENSG00000003249 DBNDD1  protein_coding  0       0       6615    2       0       3522    42      0       15264   49      0       11742   0       0
ENSG00000003393 ALS2    protein_coding  119     0       38684   8       0       10381   254     0       80636   251     0       70255   168     0
ENSG00000003400 CASP10  protein_coding  40      0       38437   3       0       7598    47      0       46526   46      0       38928   13      0
ENSG00000003402 CFLAR   protein_coding  46      0       45114   1       0       22528   107     0       60584   53      0       38056   51      0
ENSG00000003436 TFPI    protein_coding  15      0       23787   12      0       11714   30      0       101531  30      0       89817   4       0
```

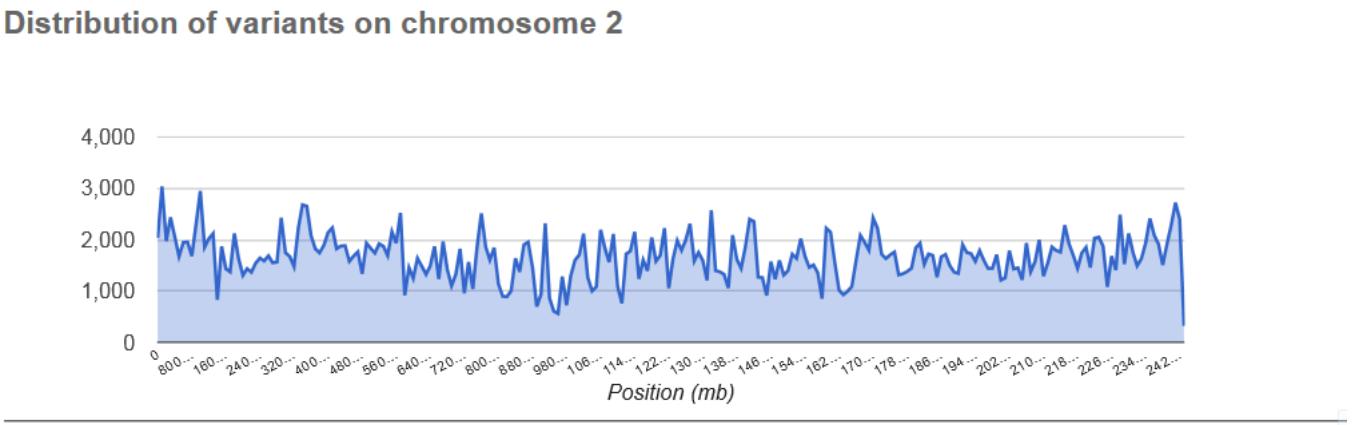
Mon génome…. (celui d'un collègue)

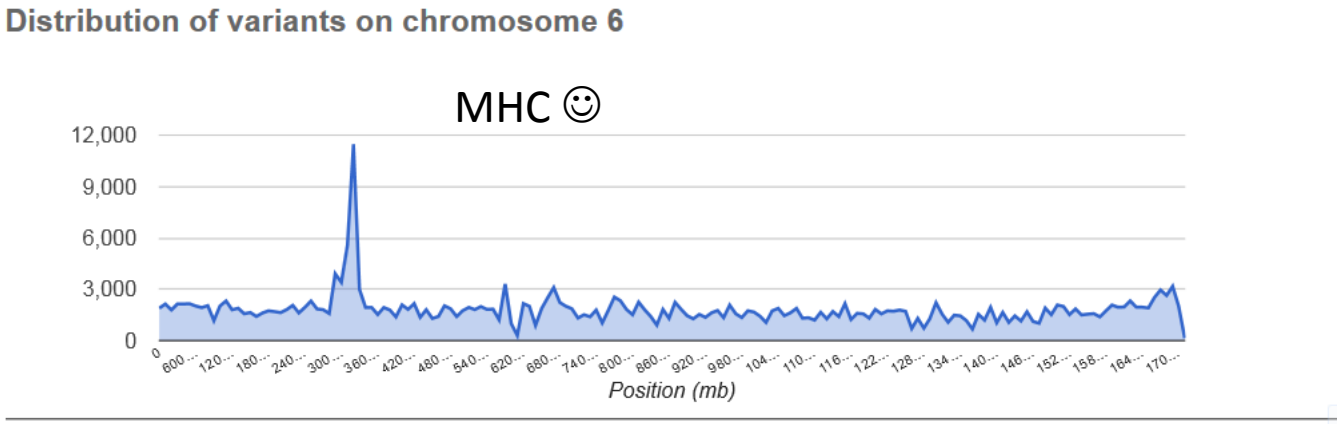
- 10 millions de SNPs, 5 millions 'validés'
- 28.8% des variants sont 'nouveaux'
- 13'000 SNPs codants, 50 % homozygotes, 50 % hétérozygotes

General statistics

Lines of input read	
Variants processed	5022516
Variants filtered out	0
Novel / existing variants	1446690 (28.8) / 3575826 (71.2)
Overlapped genes	57286
Overlapped transcripts	195446
Overlapped regulatory features	105340

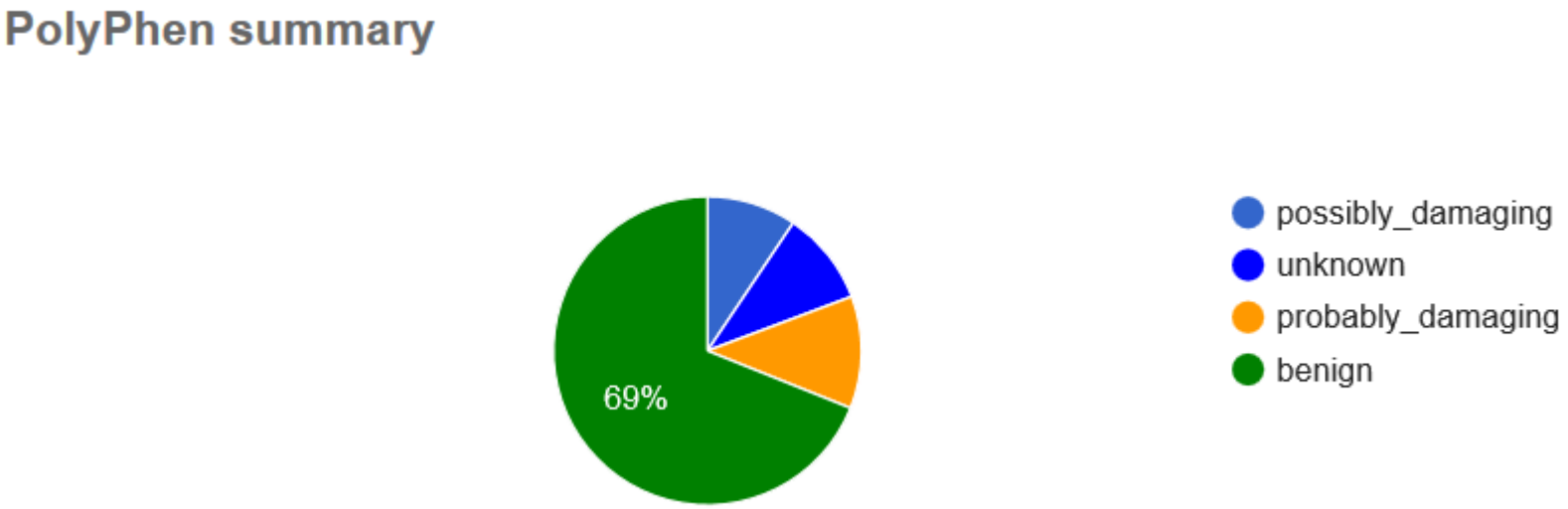
Mon génome…. (celui d'un collègue)





file:///C:/TEMP/XY57G.vep_stats-1.html

- http://genetics.bwh.harvard.edu/pph2/
- a tool which predicts possible impact of an amino acid substitution on the structure and function of a human protein



Activité 2.1

Comparer des génomes de mitochondrie (Néanderthal, ‘mon génome’, chimpanzé, etc…)
(Align@UniProt)

Comprendre la notion de génome de référence
Comprendre la notion de variation génétique

Séquences ADN mitochondrial: http://education.expasy.org/cours/PO17421/
‘mon génome/collègue’, Néanderthal, Génome de référence

Depuis le site
http://education.expasy.org/cours/PO17421

Récupérer les séquences ADN (format fasta)
Copier / coller les séquences : http://www.uniprot.org/align/

Cliquer sur ‘Similarity’

Comparaison ADN collègue (MT) / MT_Reference
99.89% identité

Alignment

How to print an alignment in color

```
MT_Reference      1 GATCACAGGTCTATCACCCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTT   60
MT_collegue       1 GATCACAGGTCTATCACCCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTT   60
                    ************************************************************

MT_Reference     61 CGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCTATGTC  120
MT_collegue      61 CGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCTATGTC  120
                    ************************************************************

MT_Reference    121 GCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATTTATCGCACCTACGTTCAATATT  180
MT_collegue     121 GCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATTTATCGCACCTACGTTCAATATT  180
                    ************************************************************

MT_Reference    181 ACAGGCGAACATACTTACTAAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATA  240
MT_collegue     181 ACAGGCGAACATACTTACTAAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATA  240
                    ************************************************************

MT_Reference    241 ACAATTGAATGTCTGCACAGCCACTTTCCACACAGACATCATAACAAAAAATTTCCACCA  300
MT_collegue     241 ACAATTGAATGTCTGCACAGCCGCTTTCCACACAGACATCATAACAAAAAATTTCCACCA  300
                    ********************.***************************************

MT_Reference    301 AACCCCCCCTCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGCCAAACCCCAAAA  360
MT_collegue     301 AACCCCCCCCCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGCCAAACCCCAAAA  360
                    *********  ************************************************

MT_Reference    361 ACAAAGAACCCTAACACCAGCCTAACCAGATTTCAAATTTTATCTTTTGGCGGTATGCAC  420
MT_collegue     361 ACAAAGAACCCTAACACCAGCCTAACCAGATTTCAAATTTTATCTTTTGGCGGTATGCAC  420
                    ************************************************************

MT_Reference    421 TTTTAACAGTCACCCCCCAACTAACACATTATTTTCCCCTCCCACTCCCATACTACTAAT  480
MT_collegue     421 TTTTAACAGTCACCCCCCAACTAACACATTATTTTCCCCTCCCACTCCCATACTACTAAT  480
                    ************************************************************

MT_Reference    481 CTCATCAATACAACCCCCGCCCATCCTACCCAGCACACACACACCGCTGCTAACCCCATA  540
MT_collegue     481 CTCATCAATACAACCCCCGCCCATCCTACCCAGCACACACACACCGCTGCTAACCCCATA  540
                    ************************************************************

MT_Reference    541 CCCCGAACCAACCAAACCCCAAAGACACCCCCCACAGTTTATGTAGCTTACCTCCTCAAA  600
MT_collegue     541 CCCCGAACCAACCAAACCCCAAAGACACCCCCCACAGTTTATGTAGCTTACCTCCTCAAA  600
                    ************************************************************
```

Comparaison ADN collègue (MT) / Néanderthal
98.721% identité

Alignment

How to print an alignment in color

```
MT          1 GATCACAGGTCTATCACCCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTT 60
NC_011137.1 1 GATCACAGGTCTATCACCCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTT 60
              ************************************************************

MT         61 CGTCTGGGGGGTATGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCTATGTC 120
NC_011137.1 61 CGTCTGGGGGGTGTGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCTATGTC 120
               ************.***********************************************

MT        121 GCAGTATCTGTCTTTGATTCCTGCCTCATCCTATTATTTATCGCACCTACGTTCAATATT 180
NC_011137.1 121 GCAGTATCTGTCTTTGATTCCTGCCCCATTCCATTATTTATCGCACCTACGTTCAATATT 180
               ************************* *** * ****************************

MT        181 ACAGGCGAACATACTTACTAAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATA 240
NC_011137.1 181 ACAGGCGAGCATACTTACTGAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAATA 240
               ********.**********.*******************************************

MT        241 ACAATTGAATGTCTGCACAGCCGCTTTCCACACAGACATCATAACAAAAAATTTCCACCA 300
NC_011137.1 241 ACGACTAAATGTCTGCACAGCTGCTTTCCACACAGACATCATAACAAAAAATTTCCACCA 300
               **.* *.*************** *******************************************

MT        301 AACCCCCCCCCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGCCAAACCCCAAAA 360
NC_011137.1 301 AACCCCCCCCTCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGCCAAACCCCAAAA 360
               ********** ************************************************

MT        361 ACAAAGAACCCTAACACCAGCCTAACCAGATTTCAAATTTTATCTTTTGGCGGTATGCAC 420
NC_011137.1 361 ACAAAGAACCCTAACACCAGCCTAACCAGATTTCAAATTTTATCTTTTGGCGGTATACAC 420
               ************************************************************.***

MT        421 TTTTAACAGTCACCCCCCAACTAACACATTATTTTCCCCTCCCACTCCCATACTACTAAT 480
NC_011137.1 421 TTTTAACAGTCACCCCCTAACTAACACATTATTTTCCCCTCCCACTCCCATACTACTAAT 480
               ***************** *******************************************

MT        481 CTCATCAATACAACCCCCGCCCATCCTACCCAGCACACACACCGCTGCTAACCCCATA 540
NC_011137.1 481 CTCATCAATACAACCCCCGCCCATCCTACCCAG----CACACACCGCTGCTAACCCCATA 536
               ********************************    ************************

MT        541 CCCCGAACCAACCAAACCCCAAAGACACCCCCCACAGTTTATGTAGCTTACCTCCTCAAA 600
NC_011137.1 537 CCCCGAGCCAACCAAACCCCAAAGACACCCCCCACAGTTTATGTAGCTTACCTCCTCAAA 596
               *******.************************************************************
```

Activité 2.2

Comparer les séquence ADN de mitochondrie
de 60 individus (Align@UniProt)

Comprendre la notion de génome de référence
Comprendre la notion de variation génétique

Séquences ADN mitochondrial (60 séquences anonymes / set Experiment@al)

Génome mitochondrie

- Le nombre de mitochondries par cellule est régulé par l'activité cellulaire (100 -> 1’000’000).
- 5 à 10 copies du génome dans une mitochondrie.
- Nombre de variations intra-individu = ???

Depuis le site
http://education.expasy.org/cours/PO17421

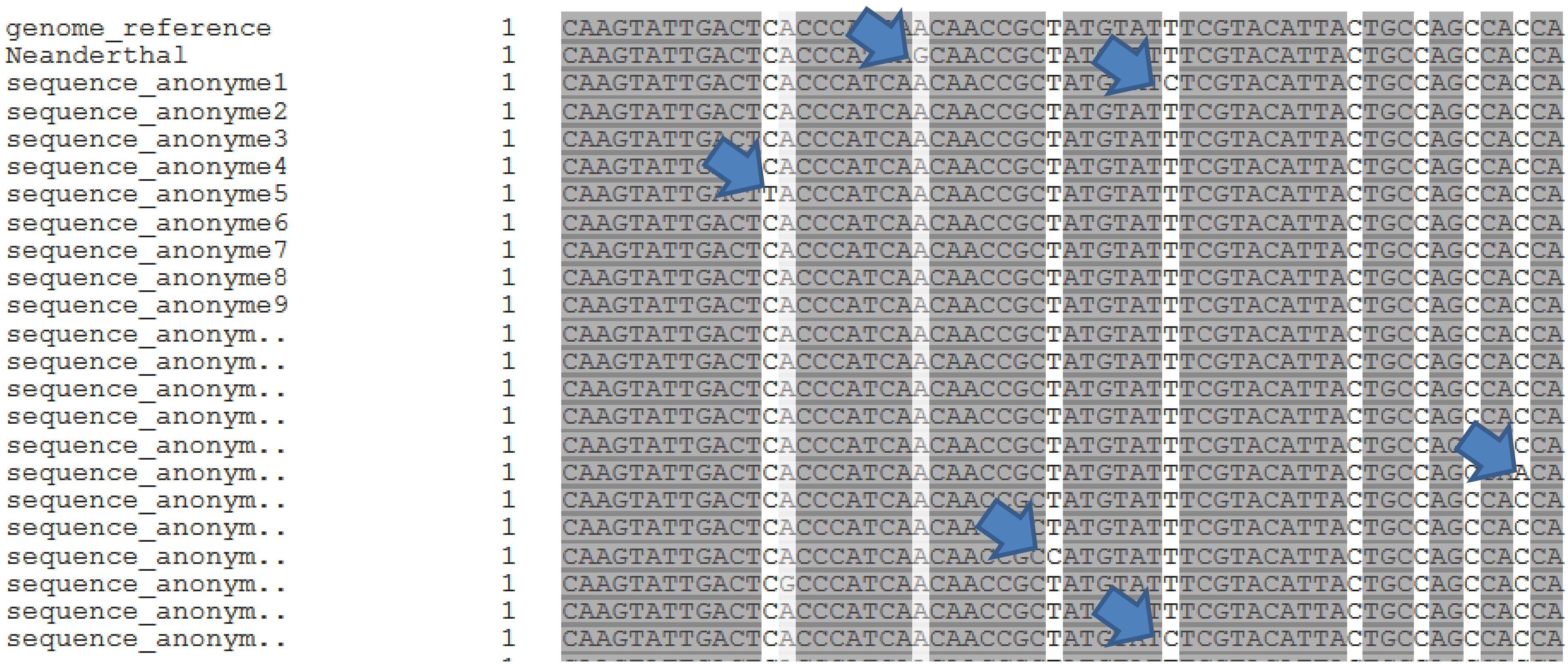
Récupérer les séquences ADN de mitochondries de 60 personnes (format fasta)
Copier / coller les séquences : http://www.uniprot.org/align/

Cliquer sur ‘Similarity’

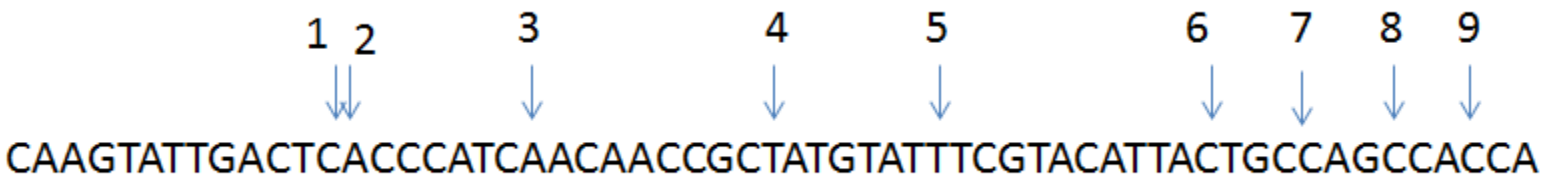
Repérer les différences….

Alignment

How to print an alignment in color



9 différences sur les 60 premiers nucléotides;
une seule différence est présente chez 2 individus



9 différences sur les 60 premiers nucléotides

Activité 2.3

Comparer une séquence ADN de mitochondrie avec le génome de référence et repérer les SNP (BLAT@UCSC)
(expert)

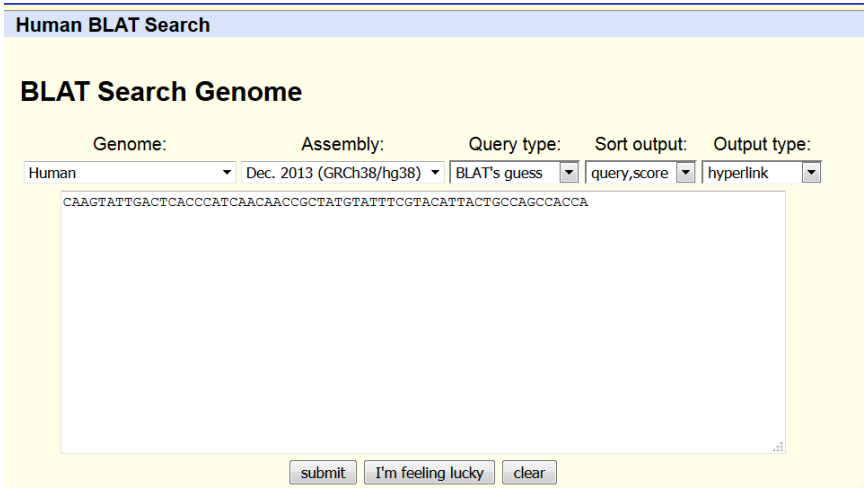
‘Voir’ les différences par rapport au génome de référence...
Notion de variations génétiques et de ‘nouvelles variations’

BLAT @ UCSC
Banque de données dbSNP

BLAT @UCSC

Aller sur le site BLAT UCSC:

http://genome.ucsc.edu/cgi-bin/hgBlat?command=start
http://genome.ucsc.edu/cgi-bin/hgBlat?hgsid=653344189_KGLWGb4JX7tbta73c2kVaMzByjD7&command=start



Copier / Coller la séquence

'CAAGTATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACCA'

Cliquer sur 'Submit'

BLAT @UCSC

BLAT Search Results

Go back to chrM:16057-16116 on the Genome Browser.

Custom track name: blat YourSeq Custom track description: blat on YourSeq

Build a custom track with these results

ACTIONS	QUERY	SCORE	START	END	QSIZE	IDENTITY	CHRO	STRAND	START	END	SPAN
browser details	YourSeq	60	1	60	60	100.0%	M	+	16057	16116	60

Explication du résultat:

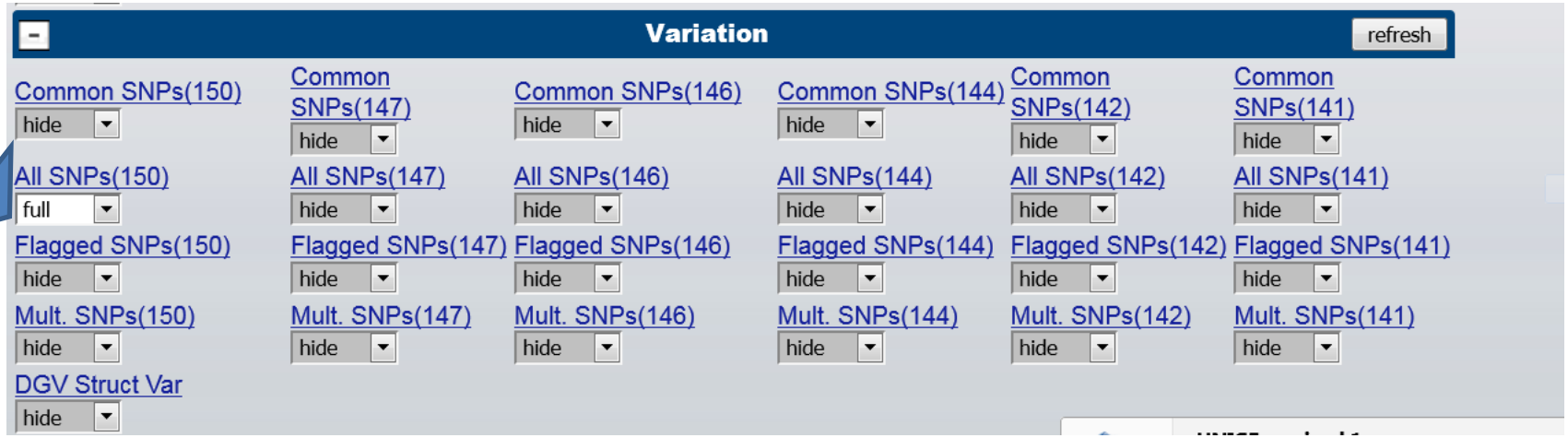
La séquence ADN (60 nucléotides) est 100 % identique à la séquence du génome mitochondrial de référence.

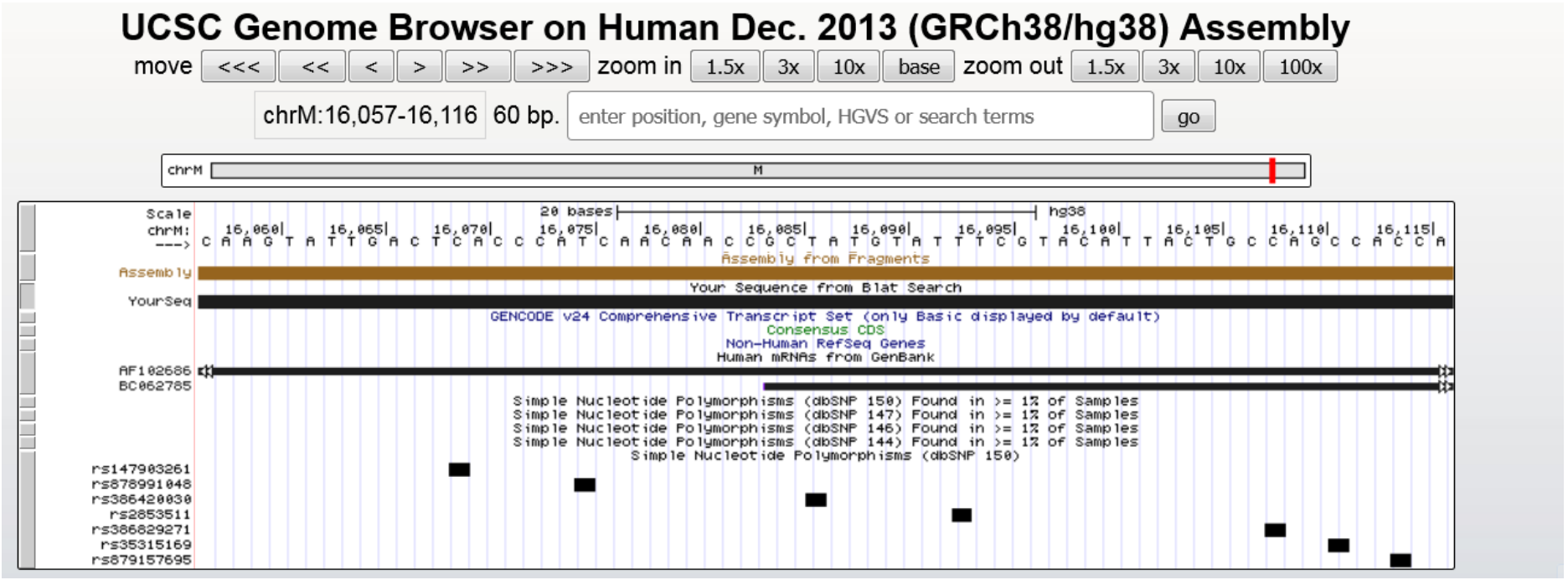
Elle est située sur le génome de la mitochondrie (M) en position 16’057-16’116

Cliquer sur ‘browser’

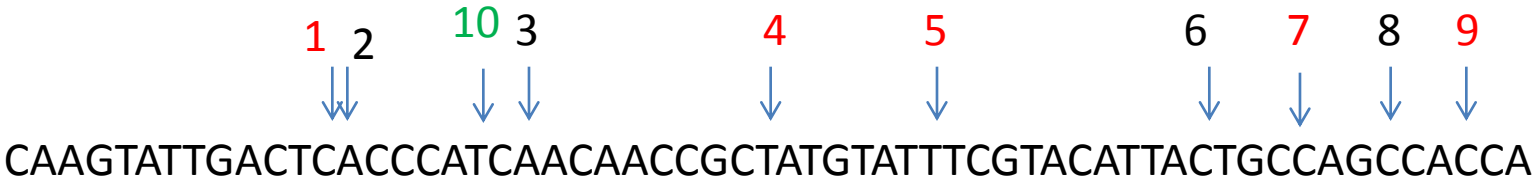
BLAT @UCSC

Choisir les options: variation all SNPs(150) full





Lister la liste des no AC des SNPs: i.e. rs147903261



1: rs147903261
2: pas dans dbSNP -> ‘nouvelle variation’
3: pas dans dbSNP; présent seulement dans l’ADN de référence de Neanderthal
4: rs386420030
5: rs2853511
6: pas dans dbSNP -> ‘nouvelle variation’
7: rs386829271
8: rs35315169
9: rs879157
10: rs878991048: ce polymorphisme est retrouvé dans plus de 1% de la population, mais n’est pas retrouvé parmi les 60 séquences

Information sur les variants: dbSNP@NCBI

Aller sur le site dbSNP: https://www.ncbi.nlm.nih.gov/SNP/

Copier / coller le no AC du SNP (i.e. rs147903261)

La banque de données dbSNP contient des informations sur les SNPs présents à plus de 1% dans la population (et qui ont été soumis et validés….)

1

rs147903261 *[Homo sapiens]*

1.

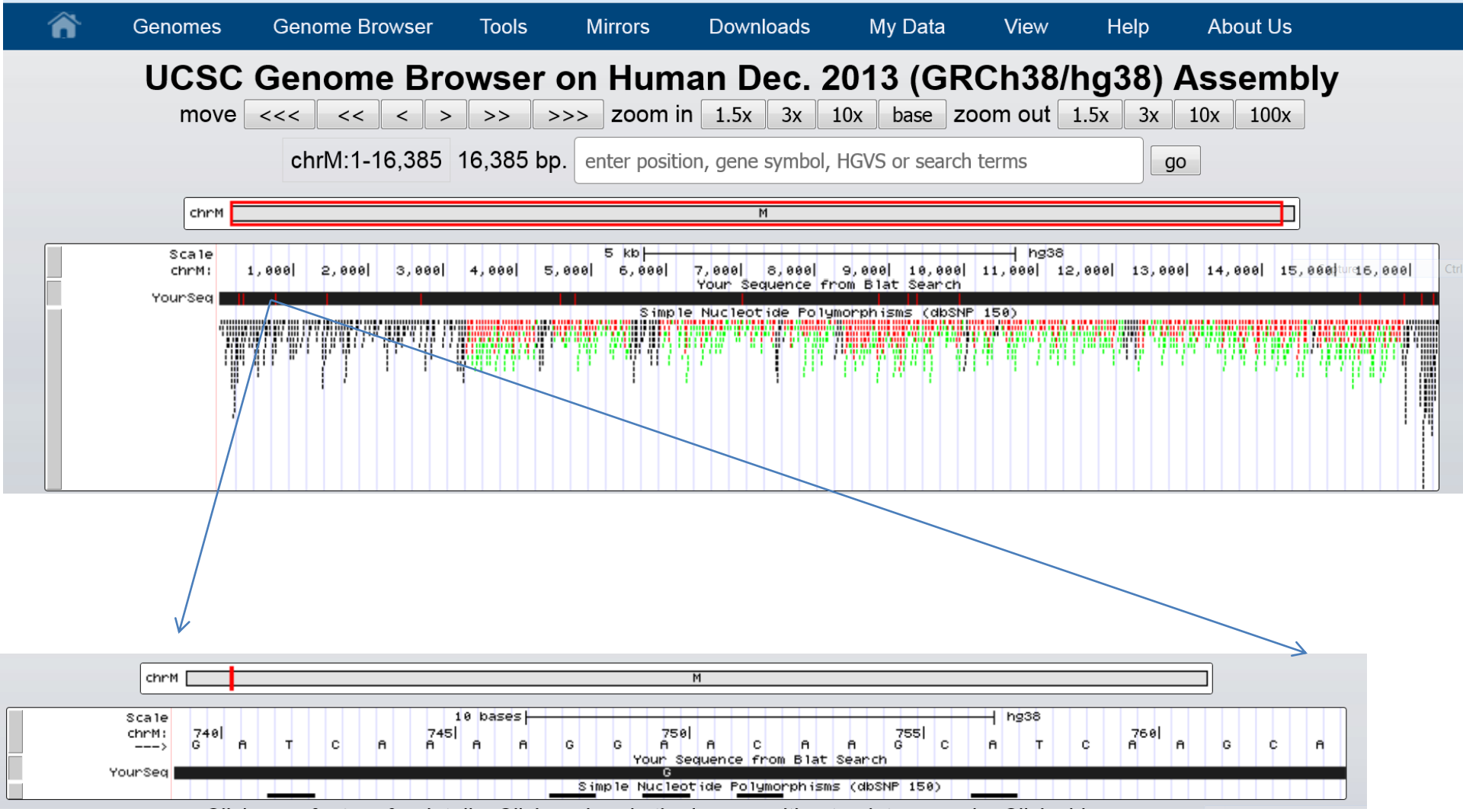
TTTGGGTACCACCCAAGTATTGACT[C/T]ACCCATCAACAACCGCTATGTATTT

Chromosome:	MT:16069
Validated:	by cluster
HGVS:	NC_012920.1:m.16069C>T

Activité 2.3 bis

Comparer une séquence ADN de mitochondrie de ‘mon génome’ avec le génome de référence et repérer les SNP (BLAT@UCSC) *(expert)*

Mon génome…. (celui d'un collègue)



Rechercher les SNP connus dans la séquence ADN de 'mon génome mitochondrial'

Activité 2.4

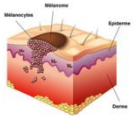
Comparer une séquence ADN de mitochondrie humain avec le génome différentes espèces (BLAT@UCSC)

Analyser les variations entre espèces pour une même séquence d'ADN
Faire un BLAT@USCS; tester différents génomes

(1) Variations inter-espèces

(2) Variations intra-espèce (humain)

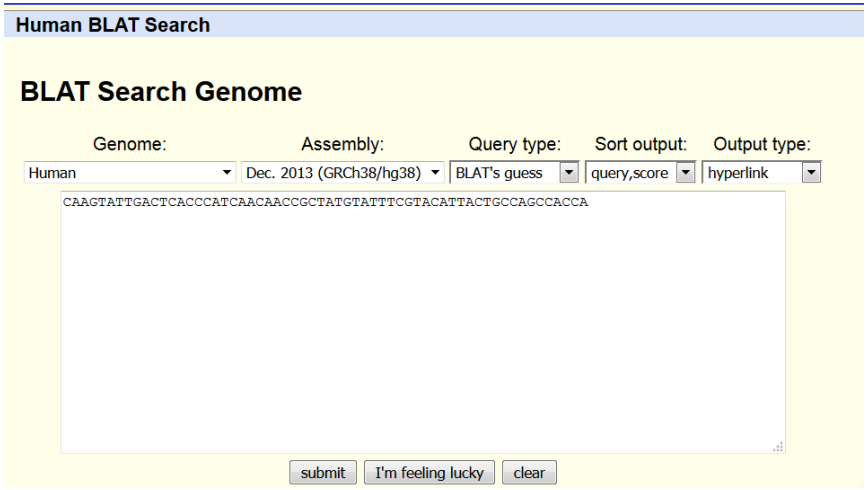
(3) Variations intra-individu (cancer)



BLAT @UCSC

Aller sur le site BLAT UCSC:

http://genome.ucsc.edu/cgi-bin/hgBlat?command=start
http://genome.ucsc.edu/cgi-bin/hgBlat?hgsid=653344189_KGLWGb4JX7tbta73c2kVaMzByjD7&command=start



Copier / Coller la séquence

'CAAGTATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACCA'

Sélection un génome de votre choix

Cliquer sur 'Submit'

‘CAAGTATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCACC’

espèces	% identité
Homo sapiens (M)	100
Chimp (M)	100
Bonobo (M)	100
Gorilla	No match
mouse	No match
rat	No match
chick	No match
xenopus	No match
D. Melanogaster	No match

>Gene Histone H4 GCRh38

TCCCGCCATT TCCTGGGGCT TGGAGGAGGG GTTAAAGGAG CGGACTGTAG GCGTCACATT
TCCCGCCTGC GCGCTTTTCA GTCTCAGTGT CCGCTGGAGG TGGGGGCAGG GGTAACGTAG
ATATATAAAG ATCGGTTTCC TATTCTCTCA CTTGCTCTTG GTTCACTTCT TGGGAAGTCA
TGTCTGGACG TGGTAAGGGC GGGAAGGGTT TGGGTAAGGG GGGTGCCAAG CGCCACCGCA
AGGTGTTGCG TGACAACATC CAGGGCATCA CCAAGCCGGC CATCCGGCGT CTGGCCCGGC
GTGGCGGTGT GAAGCGGATC TCTGGTCTGA TCTACGAGGA GACTCGCGGG GTGCTCAAGG
TGTTTTTGGA GAACGTGATC CGTGACGCTG TCACCTATAC GGAGCACGCC AAGCGCAAGA
CAGTCACTGC CATGGACGTG GTCTACGCGC TTAAGCGCCA GGGACGCACC CTTTATGGCT
TTGGCGGTTA AGGTTGCTGA TTTCTCCACA GCTTGCATTT CTGAACCAAA GGCCCTTTTC
AGGGCCGCCC AACTAAACAA AAGAAGAGCT GTATCCATTA AGTCAAGAAG CTCAATGTGT
AATTAAGATG AATGATACTG AGCTGACATC CTAAAAAGGA AAGATTAGGG GAACTCCAAG
TTTGCCCTCC

espèces	% identité	chromosome
Homo sapiens	100	6
Chimp	98.9	6
Bonobo	98.9	6
Gorilla	97.5	6
mouse	84.3	13
rat	86.9	17
chick	87.4	1
xenopus	78.7	3
D. Melanogaster	89.1	3R

210-492 /670

```
>Gene ACTN1 (oartie; EXON, intronl) 780 nucléotides
     ttccactcac ttctagcgct ccctttgttc ctcttgccca catattccca tctctgtcCA
     GGGACTGACC ACAGCCCATG AGCAGTTCAA GGCCACCCTC CCTGATGCCG ACAAGGAGCG
     CCTGGCCATC CTGGGCATCC ACAATGAGGT GTCCAAGATT GTCCAGACCT ACCACGTCAA
     TATGGCGGGC ACCAACCCCT ACACAACCAT CACGCCTCAG GAGATCAATG GCAAATGGGA
     CCACgtgagt tgaagggcat gggccgagcc attgtaagtt tcataaaggc agggattttt
     gtccatttta attgattttt gtcaatttca tttaaaattg tcatatagta aacactcatt
     gtttgtccaa atgtcaaatg actgggtttc cagaacaggt tgacaccttt actcctgttc
     tgttactgac cggcccttcc ccttctgggt cctccatttc ctcatctggc acatggacag
     atggctgacc catcctgaaa gtcgccttgt tgttcctgcc ccagGTGCGG CAGCTGGTGC
     CTCGGAGGGA CCAAGCTCTG ACGGAGGAGC ATGCCCGACA GCAGCACAAT GAGAGGCTAC
     GCAAGCAGTT TGGAGCCCAG GCCAATGTCA TCGGGCCCTG GATCCAGACC AAGATGGAGg
     tgggtcccat gctgggaagc agtgctgggg ctcccacccc tccccccCga cactatccta
     gccagccagg ctgcctgtct ccctgggtgg gacccatgga ggactgaggc ccggagccca
```

espèces	% identité	chromosome
Homo sapiens	100	14
Chimp	99.4	14
Bonobo	99.4	14
Gorilla	99.0	14
mouse	87.8	8
rat	87.2	6
chick	87.6	5
xenopus	---	
D. Melanogaster	No match	

59 538 / 780

295 320 / 780

Conservation d'une séquence ADN du chromosome Y….

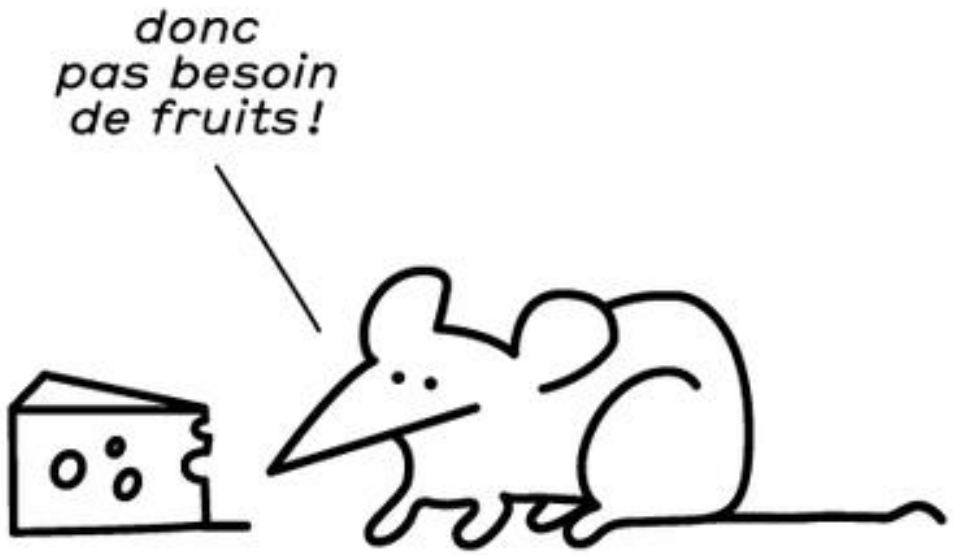
```
>Gene SRY
      ATGCAATCAT ATGCTTCTGC TATGTTAAGC GTATTCAACA GCGATGATTA CAGTCCAGCT
      GTGCAAGAGA ATATTCCCGC TCTCCGGAGA AGCTCTTCCT TCCTTTGCAC TGAAAGCTGT
      AACTCTAAGT ATCAGTGTGA AACGGGAGAA AACAGTAAAG GCAACGTCCA GGATAGAGTG
      AAGCGACCCA TGAACGCATT CATCGTGTGG TCTCGCGATC AGAGGCGCAA GATGGCTCTA
      GAGAATCCCA GAATGCGAAA CTCAGAGATC AGCAAGCAGC TGGGATACCA GTGGAAAATG
      CTTACTGAAG CCGAAAAATG GCCATTCTTC CAGGAGGCAC AGAAATTACA GGCCATGCAC
      AGAGAGAAAT ACCCGAATTA TAAGTATCGA CCTCGTCGGA AGGCGAAGAT GCTGCCGAAG
      AATTGCAGTT TGCTTCCCGC AGATCCCGCT TCGGTACTCT GCAGCGAAGT GCAACTGGAC
      AACAGGTTGT ACAGGGATGA CTGTACGAAA GCCACACACT CAAGAATGGA GCACCAGCTA
      GGCCACTTAC CGCCCATCAA CGCAGCCAGC TCACCGCAGC AACGGGACCG CTACAGCCAC
      TGGACAAAGC TG
```

espèces	% identité	chromosome	coverage
Homo sapiens	100	y	1 - 612 / 612
Chimp	98.7	Y	1 - 612 / 612
Bonobo	75.9%	17	236- 268 / 612
Gorilla	94.0%	16	178 - 210 /612
mouse	85.0%	Y	242-374 /612
rat			
chick			
xenopus			
D. Melanogaster			

GULO et le besoin de vitamine C

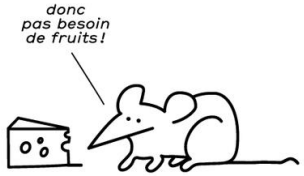
http://www.chromosomewalk.ch/chromosome/chromosome-8/

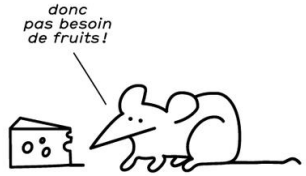
La large majorité des mammifères élaborent eux-mêmes leur propre vitamine C, comme la souris et le chien. Comment avons-nous perdu cette faculté? Parce que nous ne produisons plus la **protéine** responsable de sa fabrication: Gulo. (Gulo est un pseudogène chez l'humain)



>Gene (mRNA) GULO mouse

GCCTTTCTTGGTACCTGTGGCTAAACTTCCAGTCCCGTTCTGCCTGAGGTAACCCAGAGCCGAGGTTGCC
TGACCACTGCATCTGCTGCTGCCCAGGGCTTTGTTCAACTTCCTGTGGGAACGCTTCAAGTCAAGTCGTC
TGACCCTGTTGATCGCTGGAATCATGGTCCATGGGTACAAAGGGGTCCAGTTCCAAAACTGGGCGAAGAC
CTATGGCTGCAGTCCAGAGATGTACTACCAGCCCACATCAGTGGGGGAGGTCAGAGAGGTGCTGGCCCTG
GCCCGGCAGCAGAACAAGAAAGTGAAGGTGGTGGGTGGCGGCCACTCGCCTTCAGACATCGCCTGCACCG
ATGGCTTCATGATTCACATGGGCAAGATGAACCGGGTTCTCCAGGTGGACAAGGAGAAGAAGCAGGTCAC
AGTGGAAGCCGGTATCCTCCTGACTGACCTGCACCCACAGCTGGACAAGCATGGCCTGGCCCTGTCTAAT
CTGGGAGCCGTGTCTGATGTGACGGTTGGTGGCGTCATTGGGTCTGGAACACATAACACCGGGATCAAGC
ACGGTATCCTGGCCACCCAGGTGGTGGCCCTGACCCTGATGAAGGCTGATGGAACAGTTCTGGAATGTTC
TGAGTCAAGTAATGCAGATGTGTTCCAGGCTGCAAGGGTGCACCTGGGCTGCCTGGGTGTTATCCTCACT
GTCACCCTGCAGTGTGTGCCACAGTTCCACCTTCTGGAGACATCCTTTCCTTCGACCCTCAAGGAGGTCC
TTGACAACCTGGACAGCCACCTGAAGAAGTCTGAGTACTTCCGCTTCCTCTGGTTTCCTCACAGTGAGAA
CGTCAGCATCATCTACCAAGATCACACCAACAAGGAGCCCTCCTCTGCATCTAACTGGTTTTGGGACTAT
GCCATTGGGTTCTACCTCCTGGAATTCTTGCTCTGGACCAGCACCTACCTGCCACGCCTCGTGGGCTGGA
TCAACCGCTTCTTCTTCTGGCTGCTGTTCAACTGCAAGAAGGAGAGCAGCAACCTCAGCCACAAGATCTT
CTCCTACGAGTGTCGCTTCAAGCAGCATGTCCAAGACTGGGCCATCCCCAGGGAGAAGACCAAGGAGGCC
CTGCTGGAGCTAAAGGCCATGCTGGAGGCCCACCCCAAGGTGGTAGCCCACTACCCCGTGGAGGTGCGCT
TCACCCGAGGTGATGACATCCTGCTGAGCCCGTGCTTCCAGAGGGACAGCTGCTACATGAACATCATTAT
GTACAGGCCCTATGGGAAGGATGTGCCTCGGTTGGATTACTGGCTGGCCTATGAGACCATCATGAAGAAG
TTTGGAGGCAGGCCCCACTGGGCAAAGGCCCACAATTGCACCAGGAAGGACTTTGAGAAAATGTACCCCG
CCTTTCACAAGTTCTGTGACATCCGCGAGAAGCTGGACCCCACTGGAATGTTCTTGAATTCGTACCTGGA
AAAGGTTTTCTACTAAAGTAGAAGCAGAAGCAAGCTGCCCTGACCCCTCACTTCTGCTGCCTTGGGGGGG
GGGGGGCAGGGCAGTGTCTCACAGGCACAGTGGGAGACGACCTCTCCTGAACACAAGGAGGGGCTGGGCT
CTGGCCGAGGCCTCTGCCTTCTTTATCATCTTCATAACACCCCAGGCAAGAAGTGGCCTCTCACTCAAAT
TCCTGTTTGCATCTCCACGAGCCATACATAAACTACAATCATCTCAGGAAAAGGGGTTCCCCTTGCATCA
TATCTGTCCAGGCTAAGGATTTGGTCCTTCTAGGTTCTACTGGTCCACCAAGTATAGAGAGATCCCTGGG
GCCTGCAGTTTTCCCTCCCTCTTCAGAAGGGATCTCTTGGCAACAGAGGTAGCATGAGGCATGCTCTGCT
TACTTTTATCCTTAAAGGCCTTTCAGATGCCCAGAGTCTGTCTGTTGGTCCTGAGCAAGCCATCTTCCAG
ATGGGTCCACGTGGCCTTCTGACTGCCATGGCCTGGCCCTCACAGTGTCTCTTTCGGGTGGTGTTTAGAG
TGGAATTTGCCTCGTCTTCTTAACCAGTTCCTGTTAGATCCCTGTGTTTTCTCCCTTCACCTCAGAGACA
ATTCTTTGGGCTGGGATCTCGCCGTGTCCCTGGGTTTCCCTGGGTCTTGGTTTCATCTTTCTCTTCACAG
AGATGATTTCAGTTTATTTGTGGCCTTTCTGGAATGTTCCTTGGAGAGCCAATATGTTCCAGGTACTTTG
TCAGAATTAAAAGACTCTGAAAGAT





espèces	% identité	chromosome	coverage
Homo sapiens	87.4%	8	392 -1267 /2265
Chimp	87.1%	8	392 -1267 /2265
Bonobo			
Gorilla			
mouse	100	14	1-2265 /2265
rat	91.9%	15	17-2253 /2265
chick	86.8%	3	279 1331 2265

GULO chez l'homme (et les primates) est un 'pseudogène': la séquence ADN 'existe' mais aucune protéine n'est produite

Activité 2.5

L’annotation des variants (missense; changement d’acide aminé) dans UniProtKB/Swiss-Prot

TASR38 / PTC

http://education.expasy.org/bioinformatique/gout_amer.html

>nonsensible1

atgttgactc taactcgcat ccgcactgtg tcctatgaag tcaggagtac atttctgttt
atttcagtcc tggagtttgc agtggggttt ctgaccaatg ccttcgtttt cttggtgaat
ttttgggatg tagtgaagag gcaggcactg agcaacagtg attgtgtgct gctgtgtctc
agcatcagcc ggcttttcct gcatggactg ctgttcctga gtgctatcca gcttacccac
ttccagaagt tgagtgaacc actgaaccac agctaccaag ccatcatcat gctatggatg
attgcaaacc aagccaacct ctggcttgct gcctgcctca gcctgcttta ctgctccaag
ctcatccgtt tctctcacac cttcctgatc tgcttggcaa gctgggtctc caggaagatc
tcccagatgc tcctgggtat tattctttgc tcctgcatct gcactgtcct ctgtgtttgg
tgctttttta gcagacctca cttcacagtc acaactgtgc tattcatgaa taacaatacc
aggctcaact ggcagaataa agatctcaat ttatttttatt cctttctctt ctgctatctg
tggtctgtgc ctccttttct attgtttctg gtttcttctg ggatgctgac tgtctccctg
ggaaaggcaca tgaggacaat gaaggtctat accagaaact ctcgtgaccc cagcctggag
gcccacatta aagccctcaa gtctcttgtc tccttttttct gcttctttgt gatatcatcc
tgtgttgcct tcatctctgt gcccctactg attctgtggc gcgacaaaat aggggtgatg
gtttgtgttg ggataatggc agcttgtccc tctgggcatg cagccatcct gatctcaggc
aatgccaagt tgaggagagc tgtgatgacc attctgctct gggctcagag cagcctgaag
gtaagagccg accacaaggc agattcccgg acactgtgtt ga

>nonsensible2

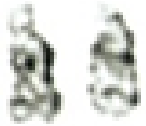
atgttgactc taactcgcat ccgcactgtg tcctatgaag tcaggagtac atttctgttc
atttcagtcc tggagtttgc agtggggttt ctgaccaatg ccttcgtttt cttggtgaat
ttttgggatg tagtgaagag gcaggcactg agcaacagtg attgtgtgct gctgtgtctc
agcatcagcc ggcttttcct gcatggactg ctgttcctga gtgctatcca gcttacccac
ttccagaagt tgagtgaacc actgaaccac agctaccaag ccatcatcat gctatggatg
attgcaaacc aagccaacct ctggcttgct gcctgcctca gcctgcttta ctgctccaag
ctcatccgtt tctctcacac cttcctgatc tgcttggcaa gctgggtctc caggaagatc
tcccagatgc tcctgggtat tattctttgc tcctgcatct gcactgtcct ctgtgtttgg
tgctttttta gcagacctca cttcacagtc acaactgtgc tattcatgaa taacaatacc
aggctcaact ggcagaataa agatctcaat ttatttttatt cctttctctt ctgctatctg
tggtctgtgc ctccttttct attgtttctg gtttcttctg ggatgctgac tgtctccctg
ggaaaggcaca tgaggacaat gaaggtctat accagaaact ctcgtgaccc cagcctggag
gcccacatta aagccctcaa gtctcttgtc tccttttttct gcttctttgt gatatcatcc
tgtgttgcct tcatctctgt gcccctactg attctgtggc gcgacaaaat aggggtgatg
gtttgtgttg ggataatggc agcttgtccc tctgggcatg cagccatcct gatctcaggc
aatgccaagt tgaggagagc tgtgatgacc attctgctct gggctcagag cagcctgaag
gtcagagccg accacaaggc agattcccgg acactgtgct ga

>sensible1

atgttgactc taactcgcat ccgcactgtg tcctatgaag tcaggagtac atttctgttc
atttcagtcc tggagtttgc agtggggttt ctgaccaatg ccttcgtttt cttggtgaat
ttttgggatg tagtgaagag gcagccactg agcaacagtg attgtgtgct gctgtgtctc
agcatcagcc ggcttttcct gcatggactg ctgttcctga gtgctatcca gcttacccac
ttccagaagt tgagtgaacc actgaaccac agctaccaag ccatcatcat gctatggatg
attgcaaacc aagccaacct ctggcttgct gcctgcctca gcctgcttta ctgctccaag
ctcatccgtt tctctcacac cttcctgatc tgcttggcaa gctgggtctc caggaagatc
tcccagatgc tcctgggtat tattctttgc tcctgcatct gcactgtcct ctgtgtttgg
tgctttttta gcagacctca cttcacagtc acaactgtgc tattcatgaa taacaatacc
aggctcaact ggcagaataa agatctcaat ttatttttatt cctttctctt ctgctatctg
tggtctgtgc ctccttttct attgtttctg gtttcttctg ggatgctgac tgtctccctg
ggaaaggcaca tgaggacaat gaaggtctat accagaaact ctcgtgaccc cagcctggag
gcccacatta aagccctcaa gtctcttgtc tccttttttct gcttctttgt gatatcatcc
tgtgctgcct tcatctctgt gcccctactg attctgtggc gcgacaaaat aggggtgatg
gtttgtgttg ggataatggc agcttgtccc tctgggcatg cagccatcct gatctcaggc
aatgccaagt tgaggagagc tgtgatgacc attctgctct gggctcagag cagcctgaag
gtcagagccg accacaaggc agattcccgg acactgtgct ga

>sensible2

atgttgactc taactcgcat ccgcactgtg tcctatgaag tcaggagtac atttctgttc
atttcagtcc tggagtttgc agtggggttt ctgaccaatg ccttcgtttt cttggtgaat
ttttgggatg tagtgaagag gcagccactg agcaacagtg attgtgtgct gctgtgtctc
agcatcagcc ggcttttcct gcatggactg ctgttcctga gtgctatcca gcttacccac
ttccagaagt tgagtgaacc actgaaccac agctaccaag ccatcatcat gctatggatg
attgcaaacc aagccaacct ctggcttgct gcctgcctca gcctgcttta ctgctccaag
ctcatccgtt tctctcacac cttcctgatc tgcttggcaa gctgggtctc caggaagatc
tcccagatgc tcctgggtat tattctttgc tcctgcatct gcactgtcct ctgtgtttgg
tgctttttta gcagacctca cttcacagtc acaactgtgc tattcatgaa taacaatacc
aggctcaact ggcagaataa agatctcaat ttatttttatt cctttctctt ctgctatctg
tggtctgtgc ctccttttct attgtttctg gtttcttctg ggatgctgac tgtctccctg
ggaaaggcaca tgaggacaat gaaggtctat accagaaact ctcgtgaccc cagcctggag
gcccacatta aagccctcaa gtctcttgtc tccttttttct gcttctttgt gatatcatcc
tgtgctgcct tcatctctgt gcccctactg attctgtggc gcgacaaaat aggggtgatg
gtttgtgttg ggataatggc agcttgtccc tctgggcatg cagccgtcct gatctcaggc
aatgccaagt tgaggagagc tgtgatgacc attctgctct gggctcagag cagcctgaag
gtaagagccg accacaaggc agattcccgg acactgtgct ga



Position 145

Sensible: cag **c**ca ctg

Non sensible: cag **g**ca ctg

Position 785

Sensible: tgt g**c**t gcc

Non sensible: tgt g**t**t gcc

Position 886

Sensible: gcc **g**tc ctg

Non sensible: gcc **a**tc ctg

Natural variant[i] (VAR_017860)	49	A → P 1 Publication Corresponds to variant dbSNP:rs713598 Ensembl, ClinVar.
Natural variant[i] (VAR_017861)	262	A → V 5 Publications Corresponds to variant dbSNP:rs1726866 Ensembl, ClinVar.
Natural variant[i] (VAR_017862)	296	I → V 1 Publication Corresponds to variant dbSNP:rs10246939 Ensembl, ClinVar.

An Alternative Laboratory Designed to Address Ethical Concerns Associated with Traditional *TAS2R38* Student Genotyping[S]

From the †Department of the History of Science, Harvard University, Cambridge, Massachusetts 02138, ‡Department of Biological Sciences, Wellesley College, Wellesley, Massachusetts 02481

The AVI variant of T2R38 is associated with an increased risk of both colorectal cancer and Pseudomonas aeruginosa-associated sinus infection and T2R38 variants have been implicated in off-target drug responses.

Polymorphism

Variations in TAS2R38 are associated with the ability to taste phenylthiocarbamide (PTC tasting) [MIM:171200]; also called thiourea tasting. The ability to taste the substance PTC and a number of related substances is genetically controlled. Genetic studies have demonstrated complex inheritance for this trait. For some people (and some chimpanzees also), the chemical PTC tastes very bitter. For others, it is tasteless. Actually, substantial variation in taste sensitivity exists in human. Five haplotypes arising from three coding SNPs in the TAS2R38 gene are associated with distinct phenotypes of PTC taste sensitivity.

Natural variant

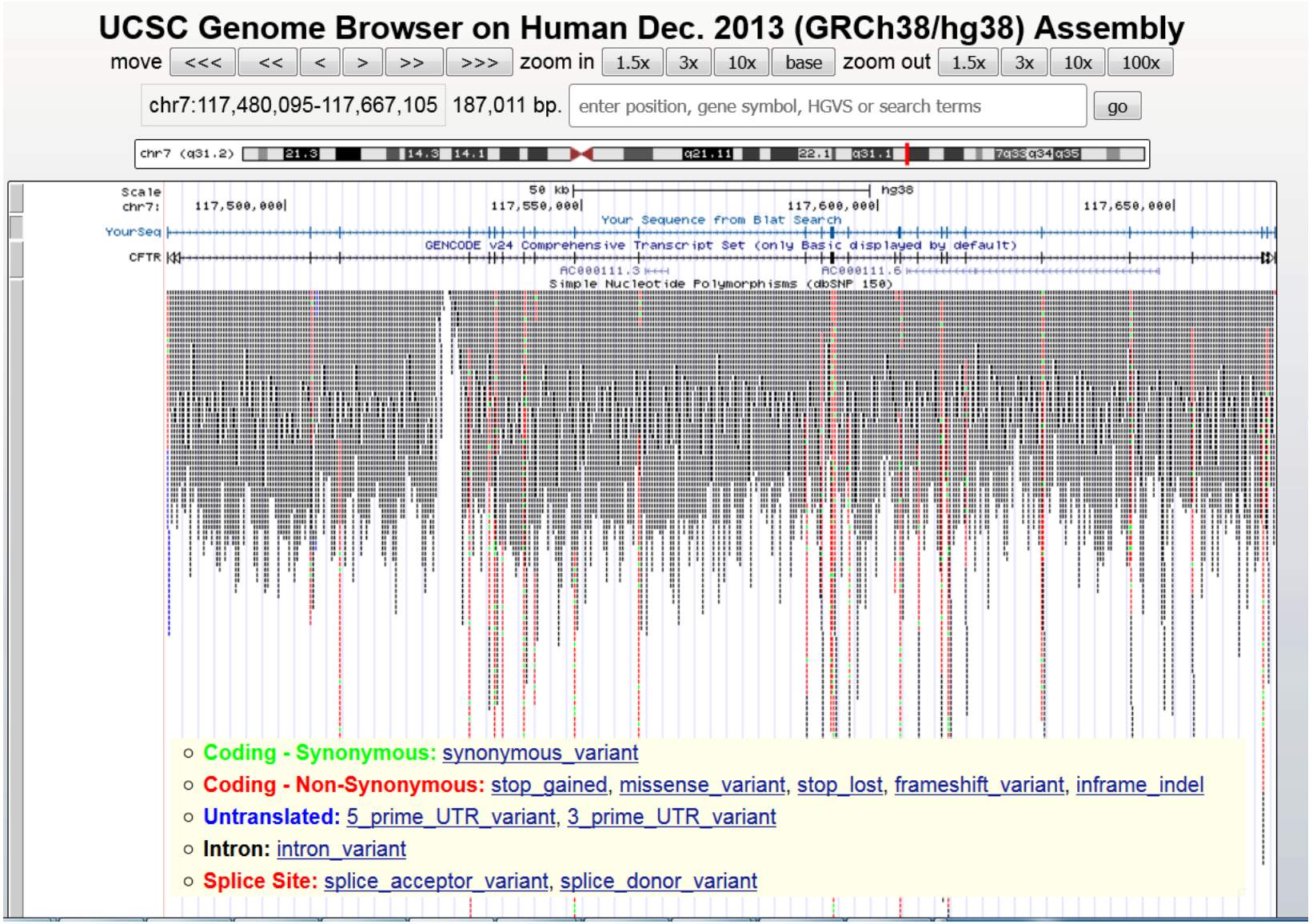
Feature key	Position(s)	Description Actions	Graphical view	Length
Natural variant (VAR_017860)	49	A → P 1 Publication Corresponds to variant dbSNP:rs713598 Ensembl, ClinVar.		1
Natural variant (VAR_017861)	262	A → V 5 Publications Corresponds to variant dbSNP:rs1726866 Ensembl, ClinVar.		1
Natural variant (VAR_017862)	296	I → V 1 Publication Corresponds to variant dbSNP:rs10246939 Ensembl, ClinVar.		1

http://www.uniprot.org/uniprot/P59533#sequences

CFTR / muscoviscidose

Affiche parking
Mont-Blanc,
(déc 2017)

Séquence CFTR_HUMAN: http://www.uniprot.org/uniprot/P13569.fasta
Blat@UCSC



CFTR @ UniProt

http://www.uniprot.org/uniprot/P13569#pathology_and_biotech

Pathology & Biotech

Involvement in disease

Cystic fibrosis (CF) 64 Publications

The disease is caused by mutations affecting the gene represented in this entry.
Disease description: A common generalized disorder of the exocrine glands which impairs clearance of secretions in a variety of organs. It is chara
bronchopulmonary disease (with recurrent respiratory infections), pancreatic insufficiency (which leads to malabsorption and growth retardation) a
the most common genetic disease in Caucasians, with a prevalence of about 1 in 2'000 live births. Inheritance is autosomal recessive.
See also OMIM:219700

Feature key	Position(s)	Description	Actions
Natural variant (VAR_000101)	13	S → F in CF. 1 Publication Corresponds to variant dbSNP:rs397508635	Ensembl.
Natural variant (VAR_000103)	31	R → L in CF. 1 Publication Corresponds to variant dbSNP:rs149353983	Ensembl.
Natural variant (VAR_000104)	42	S → F in CF. 1 Publication Corresponds to variant dbSNP:rs143456784	Ensembl, ClinVar.
Natural variant (VAR_000105)	44	D → G in CF. Corresponds to variant dbSNP:rs1800074	Ensembl.
Natural variant (VAR_000108)	57	W → G in CF. 1 Publication Corresponds to variant dbSNP:rs397508272	Ensembl.
Natural variant (VAR_000109)	67	P → L in CF. Corresponds to variant dbSNP:rs368505753	Ensembl, ClinVar.
Natural variant (VAR_000110)	74	R → W in CF. Corresponds to variant dbSNP:rs115545701	Ensembl, ClinVar.
Natural variant (VAR_000112)	85	G → E in CF. 1 Publication Corresponds to variant dbSNP:rs75961395	Ensembl, ClinVar.
Natural variant (VAR_000113)	87	F → L in CF. 1 Publication Corresponds to variant dbSNP:rs397508403	Ensembl.
Natural variant (VAR_000114)	91	G → R in CF. Corresponds to variant dbSNP:rs121908750	Ensembl, ClinVar.
Natural variant (VAR_000115)	92	E → K in CF. 2 Publications Corresponds to variant dbSNP:rs121908751	Ensembl, ClinVar.
Natural variant (VAR_000116)	98	Q → R in CF. 1 Publication Corresponds to variant dbSNP:rs397508464	Ensembl, ClinVar.
Natural variant (VAR_000117)	105	I → S in CF.	
Natural variant (VAR_000118)	109	Y → C in CF. 1 Publication Corresponds to variant dbSNP:rs121909031	Ensembl, ClinVar.
Natural variant (VAR_000119)	110	D → H in CF. Corresponds to variant dbSNP:rs113993958	Ensembl, ClinVar.
Natural variant (VAR_000121)	117	R → C in CF. 1 Publication Corresponds to variant dbSNP:rs77834169	Ensembl, ClinVar.
Natural variant (VAR_000122)	117	R → H in CF and CBAVD; decreases single channel conductance; promotes rapid return to the closed state of the channel. 2 Publications Corresponds to variant dbSNP:rs78655421	Ensembl, ClinVar.

UniprotKB/Swiss-Prot annote les variations ‘missense’, rarement les délétions. La célèbre variation DeltaF508 n’est donc pas annotée…

http://www.uniprot.org/docs/humsavar

CFTR	P13569	VAR_000101	p.Ser13Phe	Disease	rs397508635	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000102	p.Arg31Cys	Polymorphism	rs1800073	-
CFTR	P13569	VAR_000103	p.Arg31Leu	Disease	rs149353983	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000104	p.Ser42Phe	Disease	rs143456784	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000105	p.Asp44Gly	Disease	rs1800074	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000106	p.Asp44Val	Polymorphism	rs1800074	-
CFTR	P13569	VAR_000107	p.Ser50Tyr	Disease	rs397508220	Congenital bilateral absence of the vas deferens (CBAVD) [MIM:277180]
CFTR	P13569	VAR_000108	p.Trp57Gly	Disease	rs397508272	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000109	p.Pro67Leu	Disease	rs368505753	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000110	p.Arg74Trp	Disease	rs115545701	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000111	p.Arg75Gln	Polymorphism	rs1800076	-
CFTR	P13569	VAR_000112	p.Gly85Glu	Disease	rs75961395	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000113	p.Phe87Leu	Disease	rs397508403	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000114	p.Gly91Arg	Disease	rs121908750	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000115	p.Glu92Lys	Disease	rs121908751	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000116	p.Gln98Arg	Disease	rs397508464	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000117	p.Ile105Ser	Disease	-	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000118	p.Tyr109Cys	Disease	rs121909031	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000119	p.Asp110His	Disease	rs113993958	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000120	p.Pro111Leu	Disease	rs140502196	Congenital bilateral absence of the vas deferens (CBAVD) [MIM:277180]
CFTR	P13569	VAR_000121	p.Arg117Cys	Disease	rs77834169	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000122	p.Arg117His	Disease	rs78655421	Congenital bilateral absence of the vas deferens (CBAVD) [MIM:277180]
CFTR	P13569	VAR_000122	p.Arg117His	Disease	rs78655421	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000123	p.Arg117Leu	Disease	rs78655421	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000124	p.Arg117Pro	Disease	rs78655421	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000125	p.Ala120Thr	Disease	rs201958172	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000126	p.His139Arg	Disease	rs76371115	Cystic fibrosis (CF) [MIM:219700]
CFTR	P13569	VAR_000127	p.Ala141Asp	Disease	rs397508700	Cystic fibrosis (CF) [MIM:219700]

Information sur les variants: dbSNP@NCBI

Aller sur le site dbSNP: https://www.ncbi.nlm.nih.gov/SNP/

Copier / coller le no AC du SNP (i.e. rs397508635)

!!!! Encore beaucoup de désaccords sur la pathogénicité d'un variant entre les différentes banques de données. Les informations dans dbSNP sont annotées par la personne qui a soumis le variant….Plusieurs initiatives pour mettre en commun.

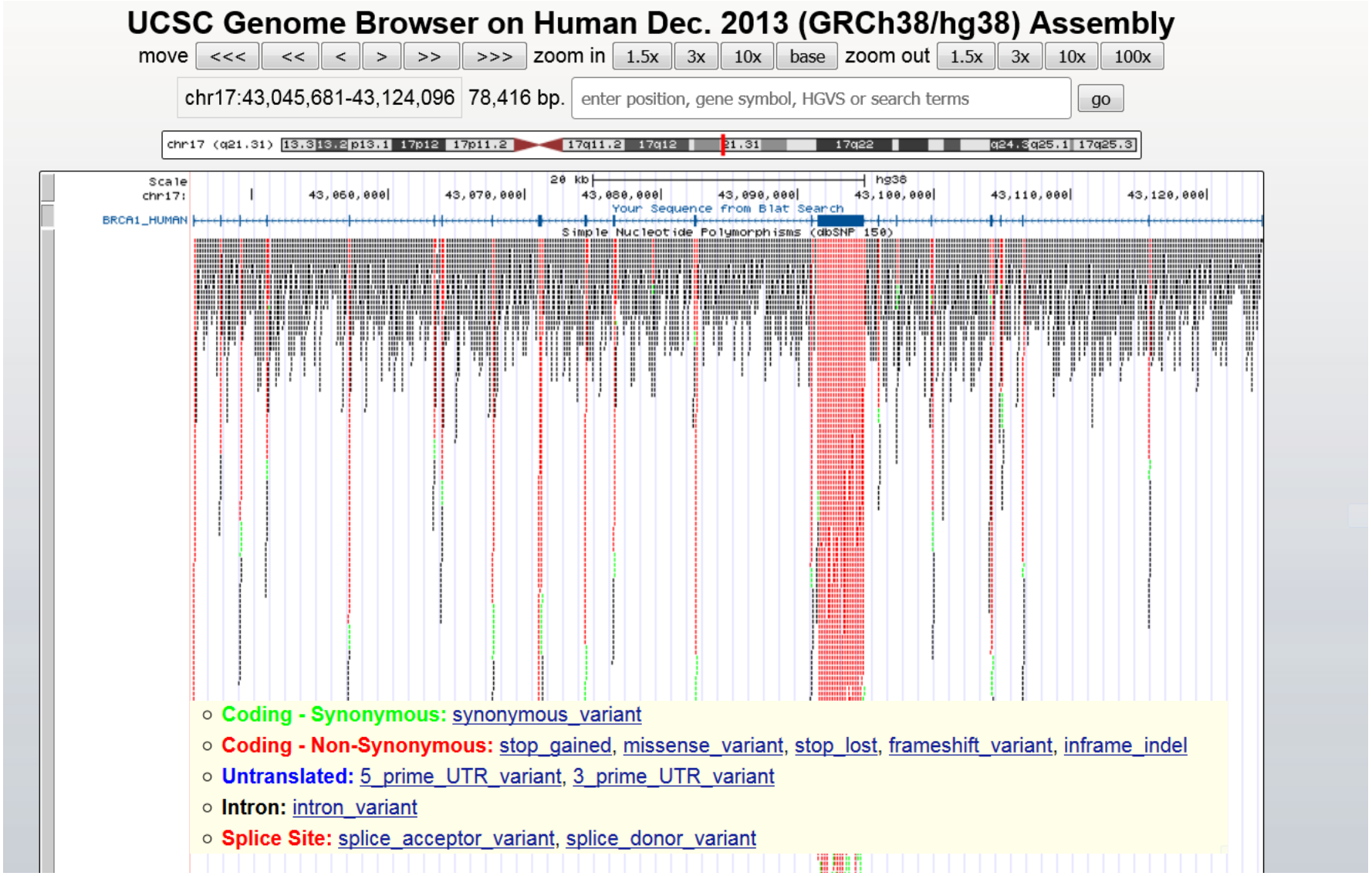
BRCA1 / cancer

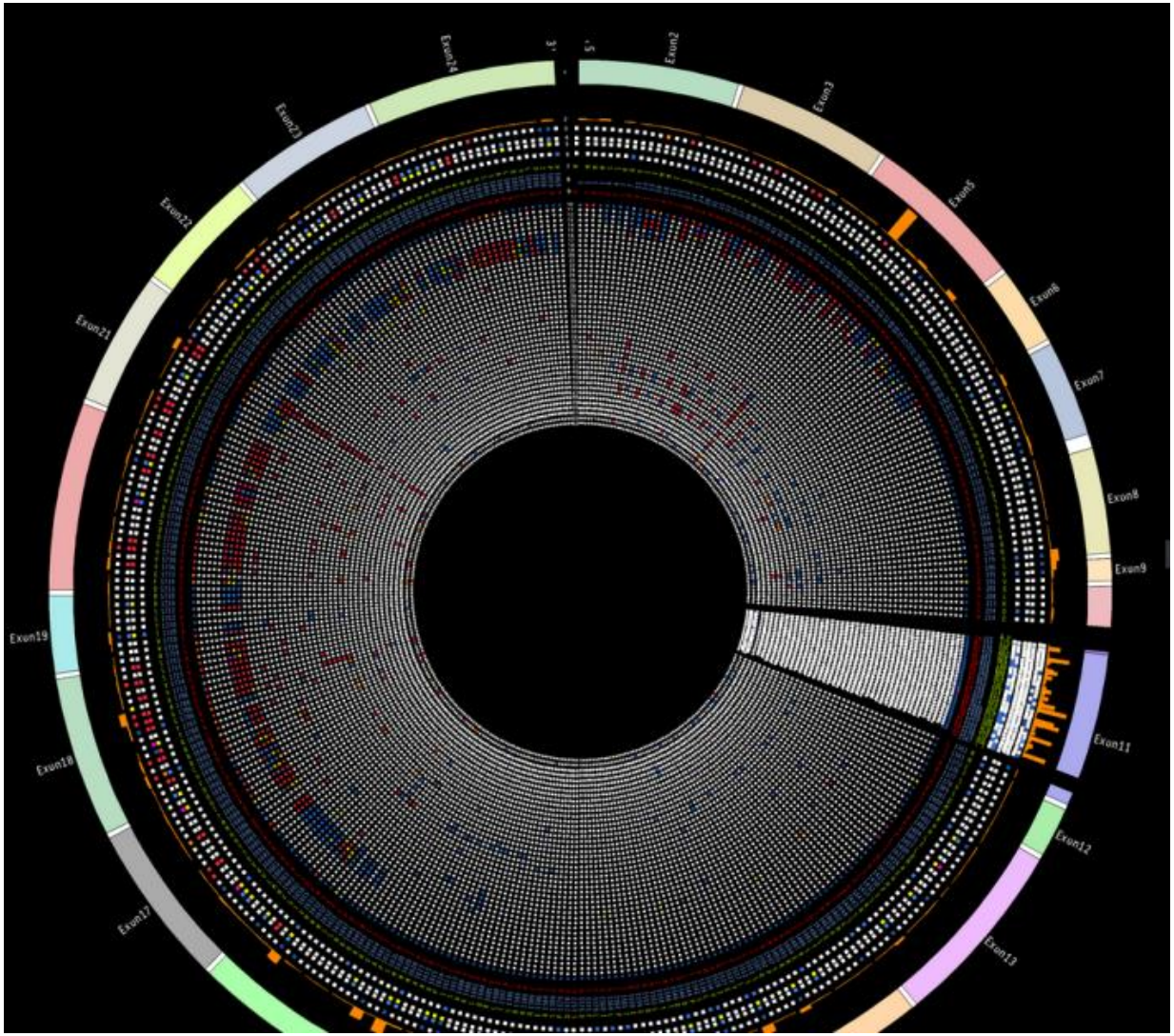
Plays a central role in DNA repair by facilitating cellular responses to DNA damage.

Mutations in BRCA1 are thought to be responsible for 45% of inherited breast cancer.

Mutations génétiques + somatiques

Séquence BRCA1_HUMAN: http://www.uniprot.org/uniprot/P38398.fasta
BLAT @ UCSC; browser





https://research.nhgri.nih.gov/projects/bic/circos/BIC_Circos_BRCA1no11-zoomer.shtml

http://www.uniprot.org/uniprot/P38398#pathology_and_biotech

Pathology & Biotech

Involvement in disease

Breast cancer (BC) 22 Publications

Disease susceptibility is associated with variations affecting the gene represented in this entry. Mutations in BRCA1 are thought to be responsible Moreover, BRCA1 carriers have a 4-fold increased risk of colon cancer, whereas male carriers face a 3-fold increased risk of prostate cancer. Cells repair by homologous recombination.

Disease description: A common malignancy originating from breast epithelial tissue. Breast neoplasms can be distinguished by their histologic pat the most common type. Breast cancer is etiologically and genetically heterogeneous. Important genetic factors have been indicated by familial oc Mutations at more than one locus can be involved in different families or even in the same case.

See also OMIM:114480

Feature key	Position(s)	Description	Actions
Natural variant (VAR_070458)	4	S → F in BC; unknown pathological significance. 1 Publication Corresponds to variant dbSNP:rs786203152	Ensembl, ClinVar.
Natural variant (VAR_020679)	10	E → K in BC and BROVCA1. 1 Publication	
Natural variant (VAR_063899)	18	M → T in BC; unknown pathological significance. 2 Publications Corresponds to variant dbSNP:rs80356929	Ensembl, ClinVar.
Natural variant (VAR_007756)	22	L → S in BC. 1 Publication Corresponds to variant dbSNP:rs80357438	Ensembl, ClinVar.
Natural variant (VAR_020680)	23	E → K in BC and BROVCA1. 1 Publication	
Natural variant (VAR_070459)	45	K → Q in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication Corresponds to variant dbSNP:rs769650474	Ensembl, ClinVar.
Natural variant (VAR_007757)	61	C → G in BC and ovarian cancer; no interaction with BAP1. 7 Publications Corresponds to variant dbSNP:rs28897672	Ensembl, ClinVar.
Natural variant (VAR_007758)	64	C → G in BC; no interaction with BAP1. 4 Publications Corresponds to variant dbSNP:rs80357064	Ensembl, ClinVar.
Natural variant (VAR_070460)	67	D → Y in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication Corresponds to variant dbSNP:rs80357102	Ensembl, ClinVar.
Natural variant (VAR_020681)	71	R → K in BC; unknown pathological significance. 1 Publication Corresponds to variant dbSNP:rs80356913	Ensembl, ClinVar.
Natural variant (VAR_070462)	132	N → K in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication Corresponds to variant dbSNP:rs80357413	Ensembl, ClinVar.
Natural variant (VAR_070463)	142	P → H in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication Corresponds to variant dbSNP:rs55971303	Ensembl, ClinVar.

http://www.uniprot.org/docs/humsavar

BRCA1	P38398	VAR_007761	p.Val271Met	Disease	rs80357244	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007762	p.Gln356Arg	Polymorphism	rs1799950	-
BRCA1	P38398	VAR_007764	p.Ile379Met	Polymorphism	rs56128296	-
BRCA1	P38398	VAR_007765	p.Phe461Leu	Disease	rs56046357	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007766	p.Tyr465Asp	Disease	rs397508869	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007767	p.Arg507Ile	Unclassified	rs80357224	-
BRCA1	P38398	VAR_007768	p.Gly552Val	Disease	rs397508893	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007769	p.Asp693Asn	Polymorphism	rs4986850	-
BRCA1	P38398	VAR_007770	p.Val772Ala	Polymorphism	rs80357467	-
BRCA1	P38398	VAR_007771	p.Lys820Glu	Polymorphism	rs56082113	-
BRCA1	P38398	VAR_007772	p.Thr826Lys	Unclassified	rs28897683	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007773	p.Arg841Trp	Unclassified	rs1800709	Breast-ovarian cancer, familial, 1 (BROVCA1) [MIM:604370]
BRCA1	P38398	VAR_007774	p.Pro871Leu	Polymorphism	rs799917	-
BRCA1	P38398	VAR_007775	p.Leu892Ser	Disease	rs397508994	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007776	p.Gly960Asp	Disease	rs397509022	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007777	p.Met1008Ile	Polymorphism	rs1800704	-
BRCA1	P38398	VAR_007778	p.Thr1025Ile	Disease	rs397509034	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007779	p.Glu1038Gly	Polymorphism	rs16941	-
BRCA1	P38398	VAR_007780	p.Ser1040Asn	Polymorphism	rs4986852	-
BRCA1	P38398	VAR_007781	p.Val1047Ala	Disease	rs397509037	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007782	p.Pro1150Ser	Disease	rs80357272	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007783	p.Lys1183Arg	Polymorphism	rs16942	-
BRCA1	P38398	VAR_007784	p.Glu1219Asp	Unclassified	rs80356876	-
BRCA1	P38398	VAR_007785	p.Arg1347Gly	Polymorphism	rs28897689	-
BRCA1	P38398	VAR_007786	p.Ser1431Pro	Polymorphism	-	-
BRCA1	P38398	VAR_007787	p.Arg1443Gly	Unclassified	rs41293455	Breast cancer (BC) [MIM:114480]
BRCA1	P38398	VAR_007788	p.Ser1512Ile	Polymorphism	rs1800744	-
BRCA1	P38398	VAR_007789	p.Thr1561Ile	Unclassified	rs56158747	-

Actualité › Société › Santé

Cancer du sein: "l'effet Angelina Jolie" fait doubler le nombre de consultations

Par **Cécile Casciano**, publié le 23/09/2014 à 18:05, mis à jour le 24/09/2014 à 02:00

Après la médiatisation de la double mastectomie préventive de l'actrice américaine, le dépistage de femmes porteuses de gènes augmentant les risques de contracter un cancer du sein a doublé, selon une étude canadienne menée en Grande-Bretagne.

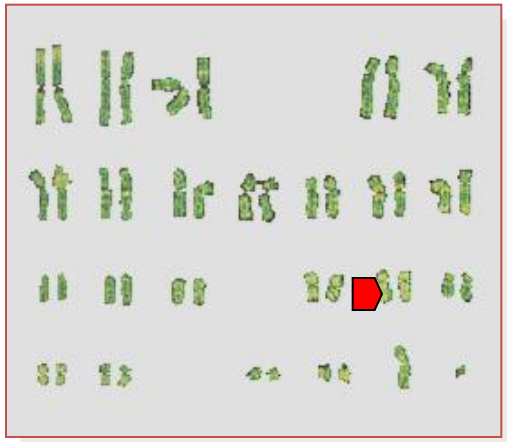
"L'effet Angelina Jolie" pourrait bien sauver des vies. L'actrice avait surpris, en mai 2013, en annonçant avoir subi une double mastectomie préventive dans une tribune intitulée *My medical choice* publiée par le *New York Times*. Porteuse du gène mutant BRCA1, ses risques de contracter un cancer du sein étaient de 87% avant l'opération. Aujourd'hui l'intervention a permis de réduire ce risque à 5%.

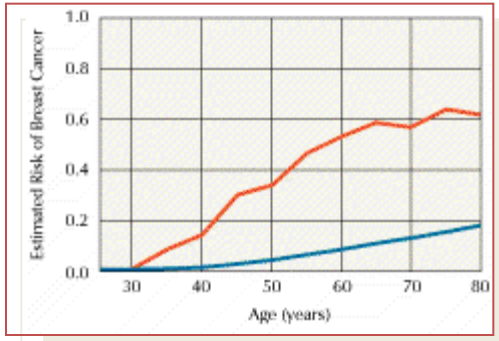
en Grande-Bretagne, le nombre de rendez-vous pour des dépistages génétiques à grimpé de 90% en six mois. Un afflux de consultations très bénéfique, soulignent les chercheurs, puisque deux fois plus de femmes porteuses du gène BRCA1 ou BRCA2, facteurs de risques pour le cancer du sein mais aussi des ovaires, ont été identifiées durant cette période.

L'express

Breast Cancer

BRCA1 gene





S. Antonarakis

CCTCTGCTCTGGGTAAAGTTAGTGGAGTCCCGGGAAAGGGACAGGGGGCCCAAGTGATGCTCTGGGGTAC
TGGCGTGGGAGAGTGGATTTCCGAAGCTGACAGATGGTTCATTGGAACAGAAAGAAATGGATTTATCTGC
TCTTCGCGTTGAAGAAGTACAAAATGTCATTAATGCTATGCAGAAAATCTTAGAGTGTCCCATCTGTCTG
GATCAAGGAACCTGTCTCCACAAAGTGTGACCACATATTTTGCAAATTTTGCATGCTGAAACTTC
CAGAAGAAAGGGCCTTCACAGTGTCCTTTATGAGCCTACAAGAAAGTACGAGATTCAGTCAACTT
AGAGCTATTGAAAACCATTTGTGCTTTTCAGCTTGACACAGGTTTGGAGTATGCAAACAGCTATA
GCAAAAAAGGAAAATAACTCTCCTGAACATCTAAAAGATGAAGTTTCTATCATCCAAAGTATGGG
CTACAGAAACCGTGCCAAAAGACTTCTACAGAGTGAACCCGAAAATCCTTCCTTGGAAACCAGTCTCAGT
GTCCAACTCTCTAACCTTGGAACTGTGAAACTCTGAGCACAAAGCAGGGATACAACCTCAAAAGACGT
CTGTCTACATTGAATTGGGATCTGATTCTTCTGAAGATACCGTTAATAAGGCAACTTATTGCAGTGTGGG
AGATCAAGAATTGTTACAAATCACCCCTCAAGGAACCAGGGATGAAATCAGTTTGGACTCTGCAAAAAAG
GCTGCTTGTGAATTTTCTGAGACGGATGTAACAAATACTGAACATCATCAACCCAGTAATAATGATTTGA
ACACCACTGAGAAGCGTGCAGCTGAGAGGCATCCAGAAAAGTATCAGGGTAGTTCTGTTTCAAACTTGCA
TGTGGAGCCATGTGGCACAAATACTCATGCCAGCTCATTACAGCATGAGAACAGCAGTTTATTACTCACT
AAAGACAGAATGAATGTAGAAAAGGCTGAATTCTGTAATAAAAGCAAACAGCCTGGCTTAGCAAGGAGCC
AACATAACAGATGGGCTGGAAGTAAGGAAACATGTAATGATAGGCGGACTCCCAGCACAGAAAAAAAGGT
AGATCTGAATGCTGATCCCCTGTGTGAGAGAAAAGAATGGAATAAGCAGAAACTGCCATGCTCAGAGAAT
CCTAGAGATACTGAAGATGTTCCTTGGATAACACTAAATAGCAGCATTCAGAAAGTTAATGAGTGGTTTT
CCAGAAGTGATGAACTGTTAGGTTCTGATGACTCACATGATGGGGAGTCTGAATCAAATGCCAAAGTAGC
TGATGTATTGGACGTTCTAAATGAGGTAGATGAATATTCTGGTTCTTCAGAGAAAATAGACTTACTGGCC
AGTGATCCTCATGAGGCTTTAATATGTAAAAGTGAAAGAGTTCACTCCAAATCAGTAGAGAGTAATATTG
AAGACAAAATATTTGGGAAAACCTATCGGAAGAAGGCAAGCCTCCCCAACTTAAGCCATGTAACTGAAAA
TCTAATTATAGGAGCATTTGTTACTGAGCCACAGATAATACAAGAGCGTCCCCTCACAAATAAATTAAAG
CGTAAAAGGAGACCTACATCAGGCCTTCATCCTGAGGATTTTATCAAGAAAGCAGATTTGGCAGTTCAAA
AGACTCCTGAAATGATAAATCAGGGAACTAACCAAACGGAGCAGAATGGTCAAGTGATGAATATTACTAA
TAGTGGTCATGAGAATAAAACAAAAGGTGATTCTATTCAGAATGAGAAAAATCCTAACCCAATAGAATCA
CTCGAAAAAGAATCTGCTTTCAAAACGAAAGCTGAACCTATAAGCAGCAGTATAAGCAATATGGAACTCG
AATTAAATATCCACAATTCAAAAGCACCTAAAAAGAATAGGCTGAGGAGGAAGTCTTCTACCAGGCATAT
TCATGCGCTTGAACTAGTAGTCAGTAGAAATCTAAGCCCACCTAATTGTACTGAATTGCAAATTGATAGT
TGTTCTAGCAGTGAAGAGATAAAGAAAAAAAAGTACAACCAAATGCCAGTCAGGCACAGCAGAAACCTAC
AACTCATGGAAGGTAAAGAACCTGCAACTGGAGCCAAGAAGAGTAACAAGCCAAATGAACAGACAAGTAA
AAGACATGACAGCGATACTTTCCCAGAGCTGAAGTTAACAAATGCACCTGGTTCTTTTACTAAGTGTTCA
AATACCAGTGAACTTAAAGAATTTGTCAATCCTAGCCTTCCAAGAGAAGAAAAAGAAGAGAAACTAGAAA
CAGTTAAAGTGTCTAATAATGCTGAAGACCCCAAAGATCTCATGTTAAGTGGAGAAAGGGTTTTGCAAAC
TGAAAGATCTGTAGAGAGTAGCAGTATTTCATTGGTACCTGGTACTGATTATGGCACTCAGGAAAGTATC
TCGTTACTGGAAGTTAGCACTCTAGGGAAGGCAAAAACAGAACCAAATAAATGTGTGAGTCAGTGTGCAG
CATTTGAAAACCCCAAGGGACTAATTCATGGTTGTTCCAAAGATAATAGAAATGACACAGAAGGCTTTAA
GTATCCATTGGGACATGAAGTTAACCACAGTCGGGAAACAAGCATAGAAATGGAAGAAAGTGAACTTGAT
GCTCAGTATTTGCAGAATACATTCAAGGTTTCAAAGCGCCAGTCATTTGCTCCGTTTTCAAATCCA....

Délétion: Dominant, une allèle suffit

Mutation c.187_188delAG occurs due to a 2-bp (AG) deletion of the normal sequence TTA GAG of codons 22-23 in exon2 of BRCA1

SAIN

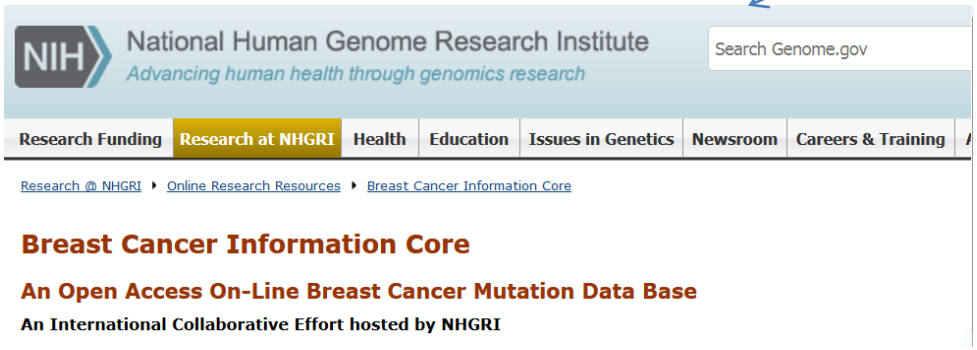
```
                                        **
TACAAAATGTCATTAATGCTATGCAGAAAATCTTAGAGTGTCCCATCTG
```

CANCER

```
TACAAAATGTCATTAATGCTATGCAGAAAATCTTAGTGTCCCATCTG
```

Exemple d'information donnée au patient….

Cependant, il a été trouvé le variant *BRCA1* c.2002C>T / p.Leu668Phe dont la répercussion fonctionnelle n'est pas clairement établie. Ce variant a été référencé dans la base de données internationale du BIC (http://research.nhgri.nih.gov/bic/), et est classé comme «unknown clinical significance». Selon les principaux algorithmes de prédiction ainsi qu'une récente publication de test fonctionnel *in vitro* ce variant serait prédit comme neutre (P. Bouwman et al., Cancer Discovery, 2013). Ce variant ne peut pas être utilisé dans le cadre d'un dépistage informatif pour les autres membres de la famille tant que sa nature pathogène ou polymorphique n'a pas été confirmée.



https://www.ncbi.nlm.nih.gov/clinvar/variation/37441/
http://www.uniprot.org/uniprot/P38398#pathology_and_biotech

Natural variant (VAR_007766)	465	Y → D in BC. 1 Publication Corresponds to variant dbSNP:rs397508869
Natural variant (VAR_007768)	552	G → V in BC. 1 Publication Corresponds to variant dbSNP:rs397508893
Natural variant (VAR_070473)	668	L → F in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication dbSNP:rs80357250
Natural variant (VAR_070474)	695	D → N in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication dbSNP:rs28897681
Natural variant (VAR_020683)	749	D → Y in BC. 1 Publication Corresponds to variant dbSNP:rs80357114
Natural variant (VAR_070475)	798	P → L in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication dbSNP:rs876660005
Natural variant (VAR_070476)	810	N → Y in BC; unknown pathological significance; functionally neutral in vitro. 1 Publication dbSNP:rs28897682

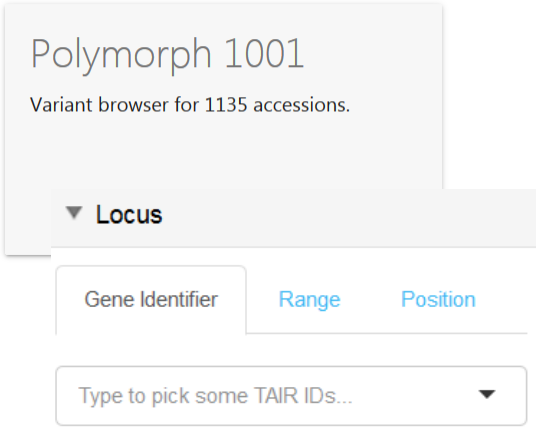
Activité 2.6

‘Voir’ les variations génétiques de *Arabidopsis Thaliana*

Génome Arapidopsis Thaliana Columbia
1135 génomes séquencés…

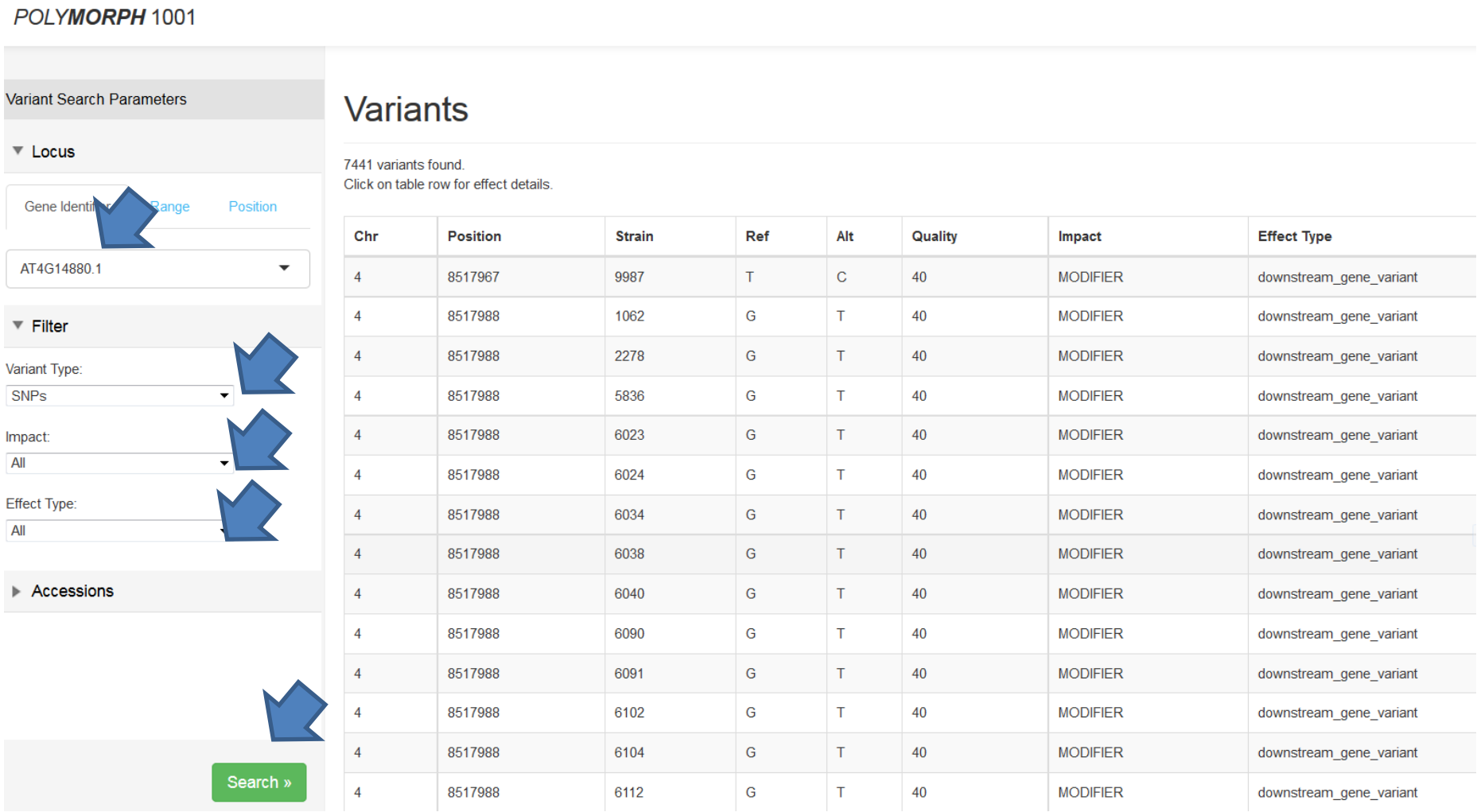
Sur le site www.1001genomes.org
Tools -> GO

Cliquer sur Polymorphism 1001



Choisir un locus (TAIR Ids: taper AT….)

Exemple avec le locus AT4G14880.1



Et Mendel dans tout ça….

Journal List › Genetics › v.189(1); 2011 Sep › PMC3176118

GENETICS

Information for Authors Editorial Board Submit a Manuscript

Genetics. 2011 Sep; 189(1): 3–10.
doi: 10.1534/genetics.111.132118

PMCID: PMC3176118

Mendel's Genes: Toward a Full Molecular Characterization

James B. Reid and John J. Ross[1]

Table 1

Seven characters of *P. sativum* examined by Mendel and a summary of the genes, phenotypes, and presumed mutations involved

Trait	Dominant phenotype	Recessive phenotype	Symbol group	Linkage group	Cloned	Gene function	Molecular nature of mutation
Seed shape	Round	Wrinkled	*R*	V	Yes	Starch branching enzyme 1	0.8-kb insertion
Stem length	Tall	Dwarf	*LE*	III	Yes	GA 3-oxidase1	G-to-A substitution
Cotyledon color	Yellow	Green	*I*	I	Yes	Stay-green gene	6-bp insertion
Seed coat/flower color	Purple	White	*A*	II	Yes	bHLH transcription factor	G-to-A at splice site
Pod color	Green	Yellow	*GP*	V	No	Chloroplast structure in pod wall	Unknown
Pod form	Inflated	Constricted	*V*?	III	No	Sclerenchyma formation in pods	Unknown
Position of flowers	Axial	Terminal	*FA*	IV	No	Meristem function	Unknown

References are given in the text.

Bio Tremplin: http://tecfa-bio-news.blogspot.ch/2013/11/de-mendel-et-ses-pois-verts-aux-risques.html

Activité 2.7

Exemple complet d'une l'analyse génétique (famille)

Remarque: Il s'agit d'un exemple très particulier d'une famille norvégienne. Cet exemple doit être utilisé pour expliquer les variations génétiques et non pas le diabète de type I

De l'ADN aux protéines…

https://www.ncbi.nlm.nih.gov/nuccore/AH012037

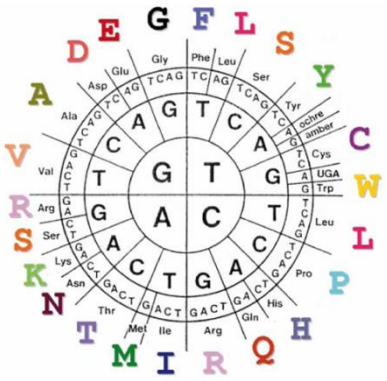
~ 20'000 gènes humains

2 % du 'texte' génome humain

(en rouge: le gène de l'insuline sur le chromosome 11)

De l'ADN aux protéines…

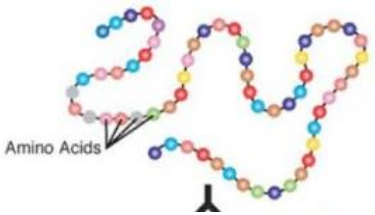
AAC ACA TGT CAC AGA ATG TCG TTC GAG CCG

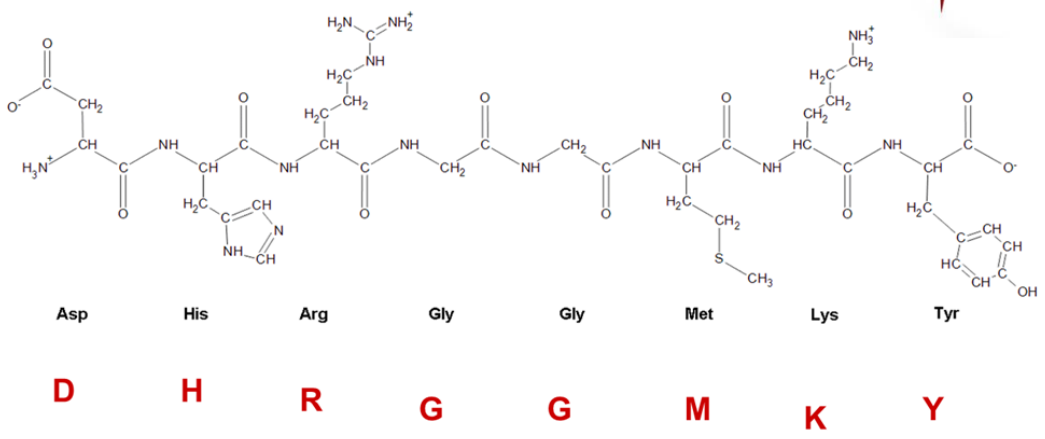


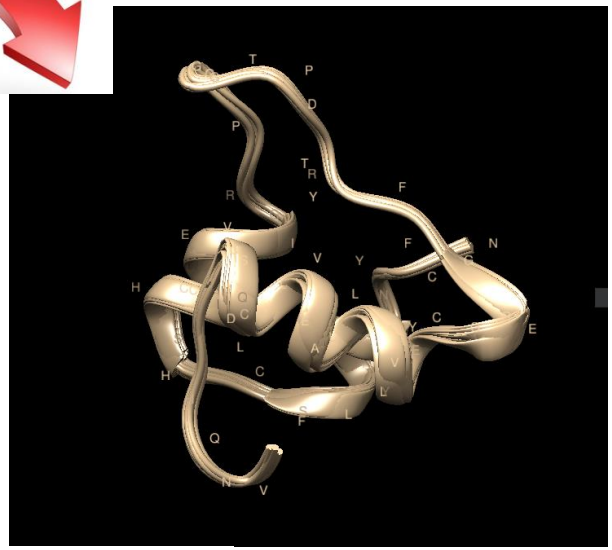
Code génétique

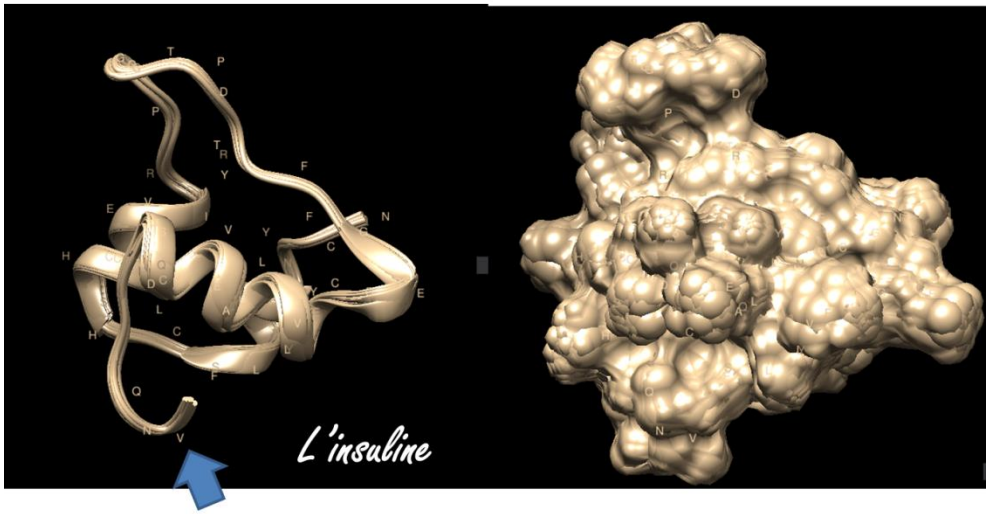
N T C H R M S F E P

Traduction d'un bout de séquence d'ADN en un bout de séquence de protéine
en utilisant le code génétique









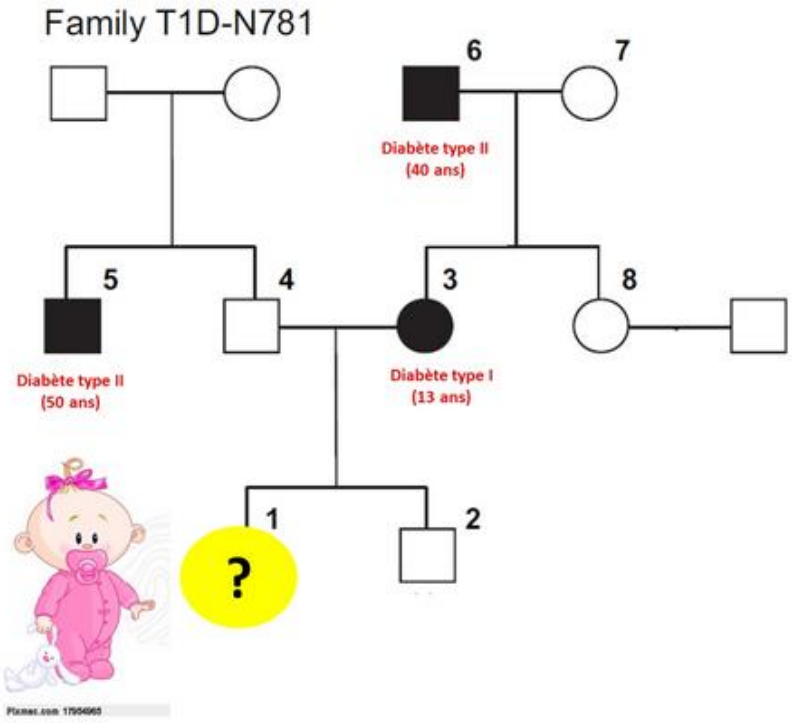
MALWMRLLPL LALLALWGPD PAAAFVNQHL CGSHLVEALY LVCGERGFFY TPKTRREAED
LQVGQVELGG GPGAGSLQPL ALEGSLQKRG IVEQCCTSIC SLYQLENYCN

Concentration max d'insuline dans le sang:

430 10^{-12} mole/L, soit ~ 2600 x 10^{11} molécules/L

En 2008, des chercheurs ont étudié une famille norvégienne dont plusieurs membres sont diabétiques (type I et type II).
Toutes les personnes diabétiques type I de cette famille sont porteuses de la même variation dans le gène qui code pour l'insuline (Molven et al., 2008).

Voici le **pédigré de la famille** (phénotype et liens de parenté):



http://education.expasy.org/bioinformatique/journee_diabete_3.html

Afin de savoir si le bébé est porteur de la mutation pathogénique, les chercheurs ont extrait l'ADN de 8 des membres de cette famille et ont séquencé une partie du gène qui code pour l'insuline.

```
>1
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggggaacgaggcttcttctacacacccaagacctgccgggaggcagaggacc
>2
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggggaacgaggcttcttctacacacccaagacccgccgggaggcagaggacc
>3
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggggaacgaggcttcttctacacacccaagacctgccgggaggcagaggacc
>4
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggggaacgaggcttcttctacacacccaagacccgccgggaggcagaggacc
>5
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggggaacgaggcttcttctacacacccaagacccgccgggaggcagaggacc
>6
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggggaacgaggcttcttctacacacccaagacccgccgggaggcagaggacc
>7
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggagaacgaggcttcttctacacacccaagacccgccgggaggcagaggacc
>8
cagccgcagcctttgtgaaccaacacctgtgcggctcacacctggtggaagctctctacc
tagtgtgcggggaacgaggcttcttctacacacccaagacccgccgggaggcagaggacc
```

Approche papier crayon:
...bandelettes de papier avec les séquences ADN à analyser manuellement afin de bien comprendre le principe de la comparaison de séquences, et de l'alignement:

8 membres de la famille - 4 séquences ADN - une allèle

8 membres de la famille - 1 séquence ADN - 2 allèles

8 membres de la famille - 2 séquences ADN - 2 allèles (plus difficile)

Approche bioinformatique:
Construire un alignement des 8 séquences à l'aide d'un outil bioinformatique et identifier la variation commune aux personnes diabétiques de type I

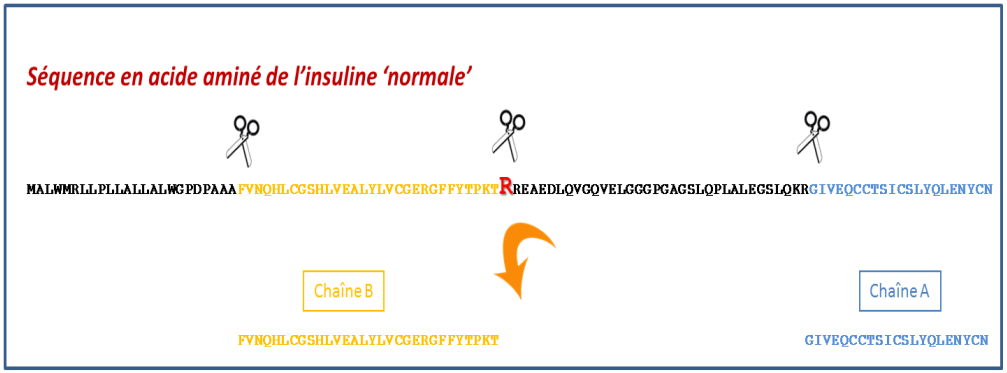
* Copier/Coller les 8 séquences (inclue la ligne '>1') dans l'outil d'alignement
* Cliquer sur l'icône 'Run align'
* Sur la page des résultats, colonne de droite 'Highlight': sélectionner 'Similarity'

....informations médicales sur la famille

Le sujet (**1**) avec la mutation R55C (hétérozygote) a présenté un diabète de type I à l'âge de 10 ans. Elle avait un taux de glucose dans le sang de 17.6 mmol/l.
Sa maman (**3**) a développé un diabète de type I à l'âge de 13 ans. Elle est sous traitement d'insuline (...). Elle est aussi hétérozygote pour la mutation R55C.
Le grand-père maternel (**6**) a été diagnostiqué diabétique de type II à l'âge de 40 ans. Il est traité à l'insuline (...). Ni lui ni sa femme (en bonne santé) ne sont porteurs de la mutation R55C, ce qui suggère qu'il s'agit d'une mutation de novo qui pourraient avoir eu lieu par exemple dans les cellules germinales d'un des 2 grands-parents (...).
Les patients **(1)** et **(3)** porteurs de la mutation R55C ont des taux de C-peptide quasi normaux, ce qui suggère qu'ils sont quand même capables de fabriquer de l'insuline.
Les scientifiques ne comprennent pas pourquoi les patients **(1)** et **(3)** ont besoin de s'injecter de l'insuline à des doses aussi élevées (...) (Molven et al., 2008).

....voici la séquence du gène de l'insuline et la liste des variations (en rouge) connues du gène de l'insuline; beaucoup de variations ne sont pas associées avec un diabète.

La mutation c -> t dans le gène de l'insuline conduit au remplacement de l'acide aminé R (arginine; codon cgc) par l'acide aminé C (cystéine; codon cgt) en position 55:
ce changement empêche la protéine insuline d'être 'coupée', un processus qui est essentiel pour que l'insuline puisse être fonctionnelle (Molven et al., 2008).
L'insuline est coupée par une enzyme appelée 'protéase' (insulin protease , insulinase).
Le site de coupure reconnu par l'insulinase est très spécifique: un changement dans la séquence en acide aminé du site de coupure (comme celui induit par la mutation R55C), empêche la protéase de faire son 'travail'.

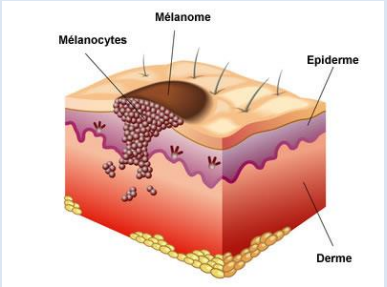


(1) Variations inter-espèces




(2) Variations intra-espèce (humain)

(3) Variations intra-individu (somatiques)

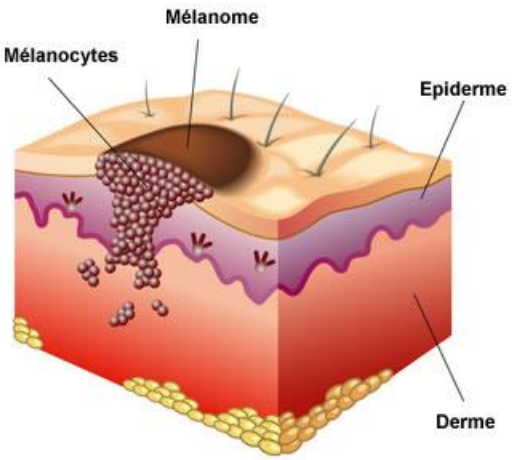


Le mélanome: cancer de la peau …et médecine personnalisée

- Le **mélanome** est un **cancer de la peau**, conséquence de la **prolifération anormale des cellules** appelées ‘mélanocytes’.

- Les cellules de la tumeur dérivent toutes d’une cellule ‘initiatrice’ qui a acquis certaines caractéristiques lui permettant de **se diviser indéfiniment** (perte de contrôle).

- Mutation (somatique)



Résultat de la recherche bioinformatique (statistique)

Cutaneous melanoma ✖

250 mutational cancer drivers have been detected in 2 *Cutaneous melanoma* (CM) projects. The most mutated drivers are: *BRAF*, *NRAS*, *TP53*, *CDKN2A*, *PTEN*, etc.

Table 2. Genes most frequently mutated in melanoma. Data from COSMIC database, (accessed on 20 November 2011).

Gene	Frequency (%)	Gene	Frequency (%)
BRAF	45	GNAQ	8
CDKN2A	29	CTNNB1	6
NRAS	19	NF2	5
TP53	17	PDGFRA	4
PTEN	17	PIK3CA	2
STK11	10	HRAS	2
FGFR2	9	KRAS	2
KIT	8	GNA11	2

Mutational cancer driver genes: 250

Cloud | Plot | Table

Downloads

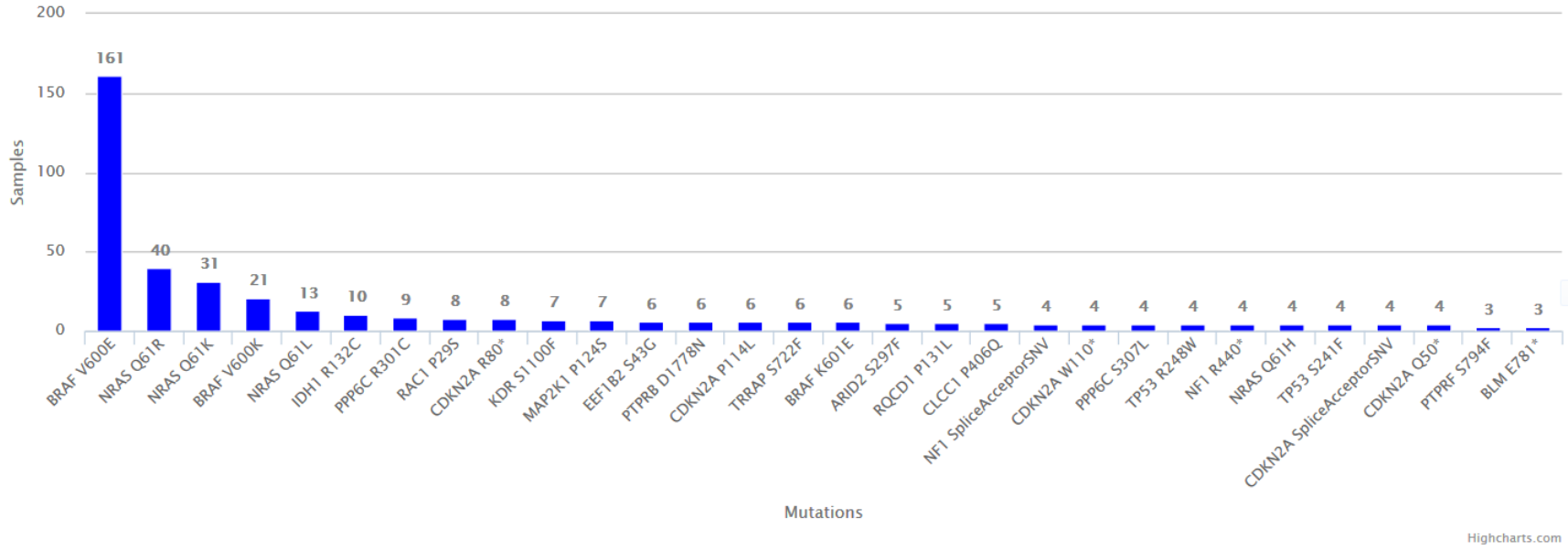
This driver cloud represents the most recurrently mutated **cancer driver genes in Cutaneous melanoma**. The size of the gene symbol is relative to the count of samples with PAMs.

https://www.intogen.org/search?cancer=CM

Catalog of driver mutations

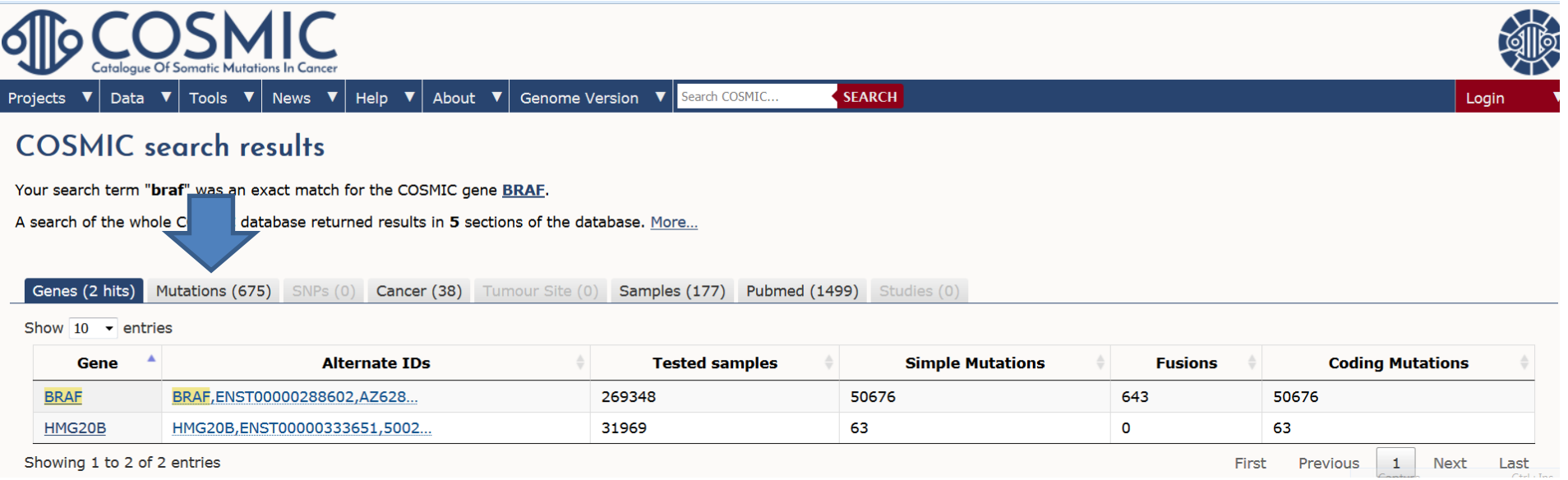
Plot | Table

Downloads



This plot shows the top 30 driver or known mutations

COSMIC
http://cancer.sanger.ac.uk/cosmic/search?q=braf



COSMIC
Catalogue Of Somatic Mutations In Cancer

Projects | Data | Tools | News | Help | About | Genome Version | Search COSMIC... SEARCH | Login

COSMIC search results

Your search term "**braf**" was an exact match for the COSMIC gene **BRAF**.

A search of the whole COSMIC database returned results in **5** sections of the database. More...

Genes (2 hits) | Mutations (675) | SNPs (0) | Cancer (38) | Tumour Site (0) | Samples (177) | Pubmed (1499) | Studies (0)

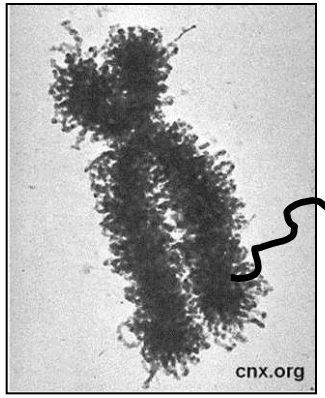
Show 10 entries

Gene	Syntax	Alternate IDs	Recurrence
BRAF	c.1799T>A	BRAF c.1799T>A...	26609
BRAF	c.?	BRAF c.?...	20618
BRAF	c.1798_1799GT>AA	BRAF c.1798_1799GT>AA...	693
BRAF	c.?	BRAF c.?...	373
BRAF	c.?	BRAF c.?...	366
BRAF	c.?	BRAF c.?...	127
BRAF	c.1801A>G	BRAF c.1801A>G...	109
BRAF	c.1798_1799GT>AG	BRAF c.1798_1799GT>AG...	108
BRAF	c.1799_1800TG>AA	BRAF c.1799_1800TG>AA...	87
BRAF	c.1781A>G	BRAF c.1781A>G...	80

Showing 1 to 10 of 675 entries

First Previous 1 2 3 4 5 ... 68 Next Last

Mutation somatique



chromosome 7

gène BRAF

variation génétique
T -> A (V600E)
retrouvée dans ~50 % des cas de mélanome

diagnostic + choix du traitement

TTTTAAAAACCATGAAAATCCATACATGCGTGTACACAC
CACACACACACACACACACCCTCCTTAAAAAAGAGT
ATTTAGGCCAGGGTGAAAACAAGATTTCTTTATGCAA
GGGAGGTTGTCTCCTACTTTAAGGCTAGGAAAACTAA
CCTTTCCTCCAAATTCCAAAATTAAATATCTTTGAT
AATAAAAAAGCATCCTATCTTAAAAAGGAGCCAGAGT
TTTAGACCAAAACAAAATGTATGAGTCCAAACTTTCACACGTCAA
TTTGTTAATGAGCTCACAATGATTCATATACTACTCTACTTTTATGTATGTTTA
GCCAGACATGGTGGCTCATGCCTGTAATCCCAGCATTTTGGGAGGCCAAGGCAGGCAGATCA
CTGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTATTAAAAACACAAAAATTAGCCAG
GCATGGTGGCAGGCTCCTGTCATCCTGGCTACTAGGGAGGCTGAGGCAGGAGAATCACTTGAACTTAGGA
GGCGGAGGTTGTAGTGAGCCAAGACTGTGCCACTGCACTCCAGCCTGGGTGACAGAGTGAGACTCCATCT
CAAAAAATATATATATATAAAATTCATATAATTTATACTTTATATATAATGTATATAATTTATATATAAC
ATAATTTATATACAATATATAACATATATATTATATATAACATATATATTATATATAACATATATATATT
ATATATAACATATATAATTTATATATATTATATATATAAAAGTTTCCATAGTAAAAAGTTTTTTAAAAAA
ACAGTTGTTCTAAAATTTAGGTGGTCATCAAAATTGTTATATGTTTAATTCTTTGAGATAGTACATTAAT
TTGGTAGTTTGTTTTACAGTGCAGACGATTATATGCCTTAAGTAAGCTTCATGATACAGTAACATCTTAA
ATAAACTGGCACAGAATATACATGCCAATATACAGATTTTGAATACAAATTTAAATTACATAGCTTTCTG
ATTTTATAATTTAACTTAATGCCAAGTTCTTTTTTGTGACTATTACATCAATTACATTTTTTTTTTTTTT
TTGAGACAGAGTCTCGCTCTGTCGCCCAGGCTGCAGTGCAGTGGCGTGATCTCAGCTCACTGCAAGCTCC
ACCTCCCGGGCTCACGCCATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGTCTGCCACCA
CGCCTGGCTAATTTTGTTTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCTCCTTGTTAGCCAAGATGGT
CTCGATCTCCTGACCTCATGATTCGCCCGCCTCAGCCTCCCAAAGTGCTAGGATTACAGGTGTGAGCCAT
TTTTTTTTTTTTTTTTTTTTGAGACAGGGTCTTGCTCTGCCACCCAGGCTTGTGCAATGGTGCGACCTAG
AATCACTGCAAACTCCACTTCACAGGTTCAAGGGATTCTCATGCCTCAGCTTCCAGAGCAGCTGGGATTA
CAGGCATGTGCCAGCAAGCCCGGCTAATTTTTTTGTATTGTTAGTAAAGACGGGGTTTTAGTATGTTCCT
CAGGCTGGTCTCGAACTCCTTACCTCAAGTGACCCGCCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAG
GTGTGAGCCACTGCAACTGGCCTAATTACAATATTGTTTTAAAGAAAAAAAAATTGTTATGATCTGCTTA

http://www.ncbi.nlm.nih.gov/nuccore/568815591?report=fasta

Activité 3.1

Comprendre comment les scientifiques font l'association entre un variant et un phénotype.

Alignement multiple @ UniProt

Récupérer les séquence ADN du gène BRAF de différents patients.
http://education.expasy.org/cours/PO17421/

Pour chaque patient, il y a l'ADN de cellules normales et l'ADN de cellules de mélanomes….

Copier/coller les séquences dans www.uniprot.org/align

Cliquer éventuellement sur 'Similarity'

Alignement multiple

```
1_normal          gtttgaactaggtgattttggtctagctacagtgaaatctcgatggagtgggtcccat
2_normal          gtttgaactaggtgattttggtctagctacagtgaaatctcgatggagtgggtcccat
3_normal          gtttgaactaggtgattttggtctagctacagtgaaatctcgatggagtgggtcccat
Patient_normal    gtttgaaataggtgattttggtctagctacagtgaaatctcgatggagtgggtcccat
1_melanome        gtttgaactaggtgattttggtctagctacagagaaatctcgatggagtgggtcccat
2_melanome        gtttgaaataggtgattttggtctagctacagtgaaatctcgatggagtgggtcccat
3_melanome        gtttgaaataggtgattttggtctagctacagagaaatctcgatggagtgggtcccat
Patient_melanome  gtttgaaataggtgattttggtctagctacagagaaatctcgatggagtgggtcccat
                  *******.************************:**************************
```

Question

Est-ce que le mélanome du patient possède la variation génétique retrouvée dans 60 % des mélanomes ?

Un variant ‘pathogenic’ ne devrait pas se retrouver dans la population saine….

Activité 3.2

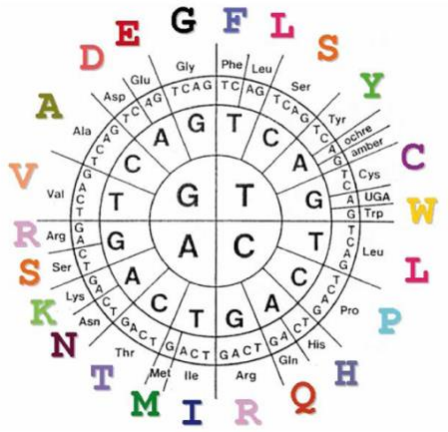
Comprendre le lien entre une analyse génétique (mutation somatique) et la prescription d'un médicament

ADN
Gène BRAF

Mutation

t -> a

…ggt gat ttt ggt cta gct aca gag aaa tct cga tgg…



Code génétique

Protéine
B-Raf

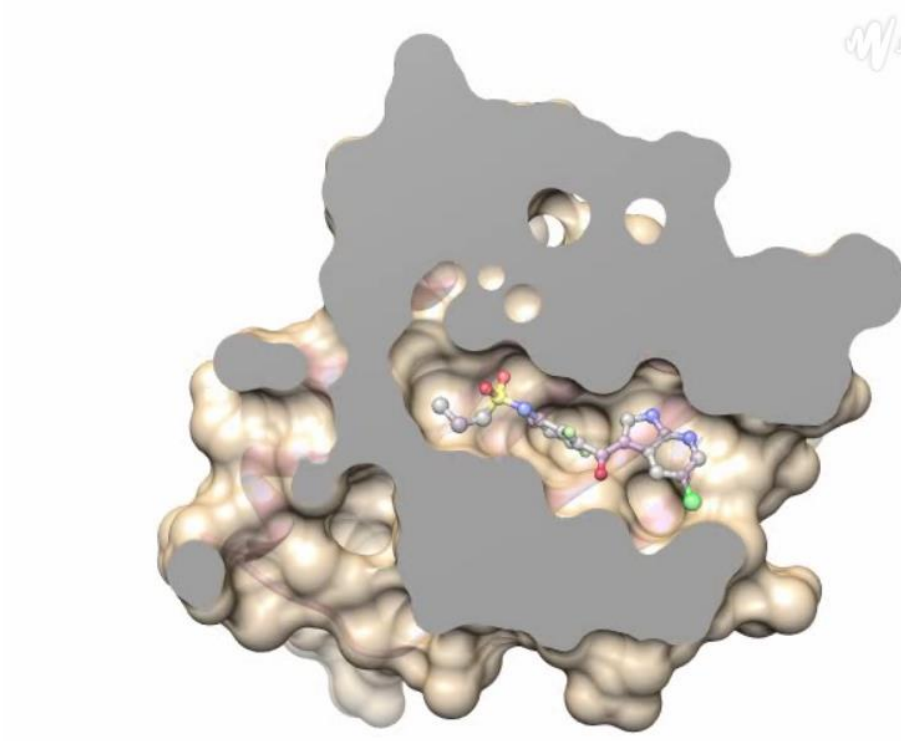
… G D F G L A T E K S R W …

Mutation

V -> E

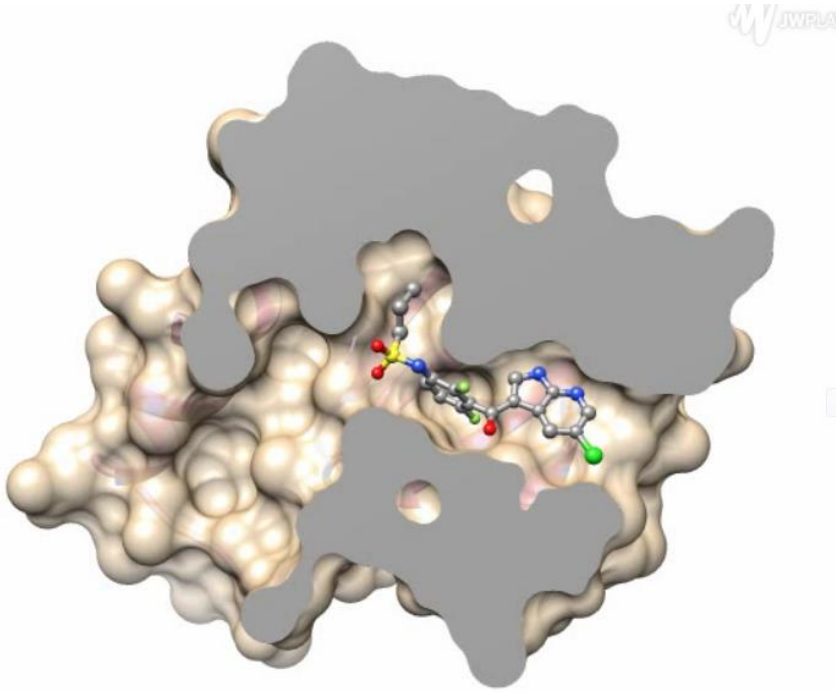
Changement d'un acide aminé (V600E)
-> Changement de la forme
-> Changement de la fonction
-> Cible spécifique d'un médicament

Protéine B-RAF



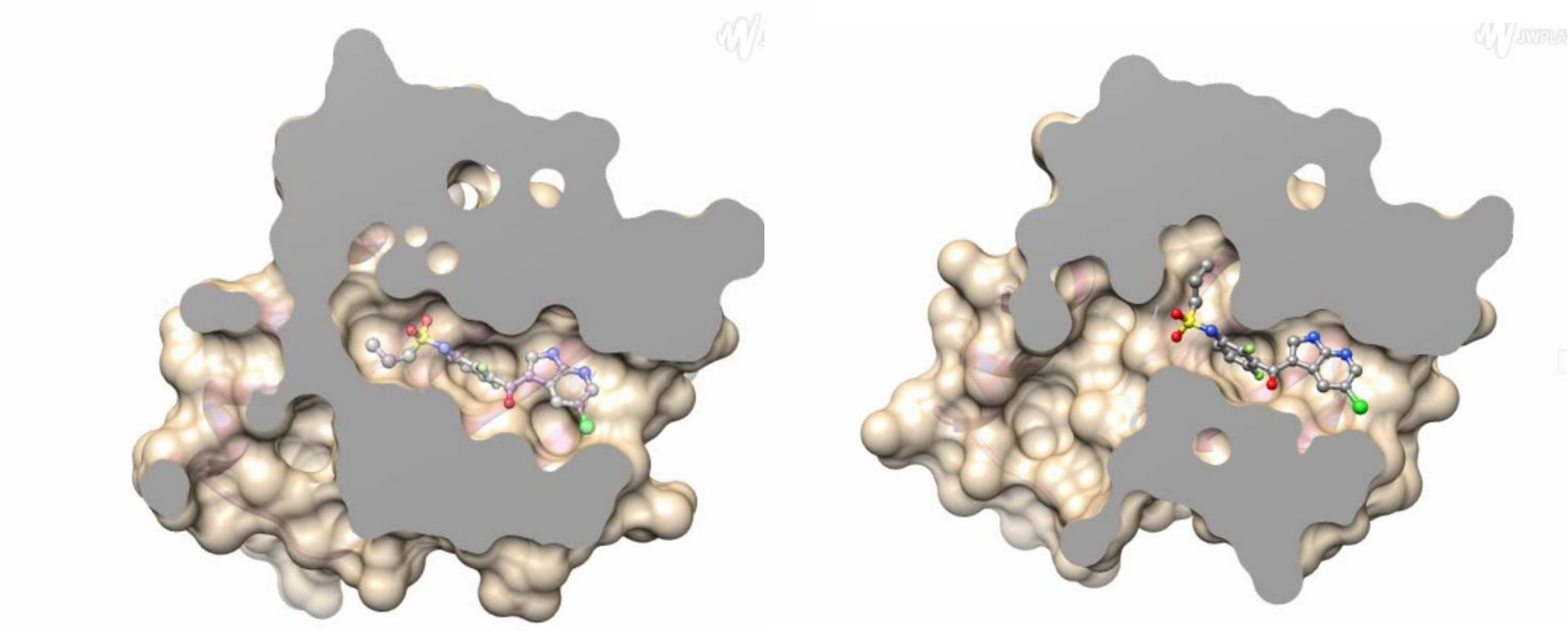
Contrôle la division des cellules

Protéine B-RAF mutée



Contrôle la division des cellules

Modélisation de la transition conformationnelle entre les formes inactives et actives de BRAF: http://www.atelier-drug-design.ch/videos.php

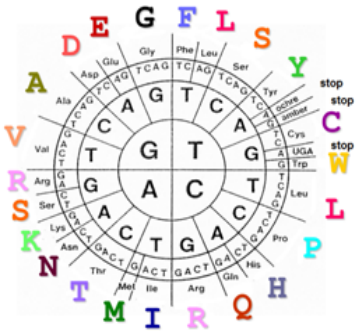


activité pédagogique

Séquence ADN

… ggt gat ttt ggt cta gct aca gtg aaa tct cga tgg …

Code génétique



ggt -> G

gat -> D

Séquence protéine

Protéine

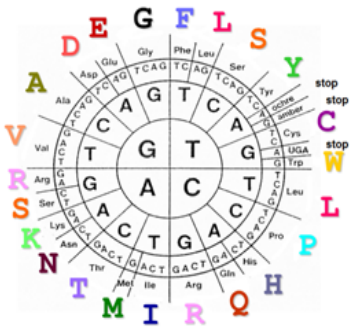
Quel médicament cible votre protéine ?

activité pédagogique

Séquence ADN

… ggt gat ttt ggt cta gct aca gag aaa tct cga tgg …

Code génétique



ggt -> G

gat -> D

Séquence protéine

Protéine

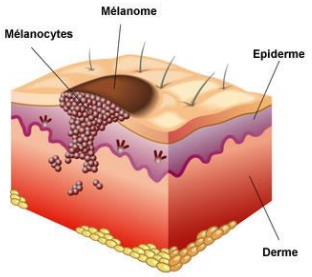
Quel médicament cible votre protéine ?

(1) Variations inter-espèces




(2) Variations intra-espèce (humain)

(3) Variations intra-individu (somatiques)



Low rate of somatic mutations in a long-lived oak tree

http://dx.doi.org/10.1101/149203

http://education.expasy.org/cours/PO17421/publication

Séquençage du génome du chêne de l'Université de Lausanne

Le temps passe, l'ADN du chêne reste

BOTANIQUE Une équipe lausannoise a pour la première fois séquencé le génome d'un chêne, celui du «chêne de Napoléon», qui pousse sur le campus de Dorigny. Surprise, son ADN est resté étonnamment stable durant ses 239 ans de vie. De quoi questionner les biologistes

FABIEN GOUBET
@fabiengoubet

C'est un vieux monsieur, devant lequel on peut passer sans lui prêter la moindre attention. Un arbre qui détient les clés d'un mystère qui plane sur la botanique depuis des décennies. Du haut de ses 30 mètres, 239 ans d'histoire nous contemplent, aurait pu dire Napoléon Bonaparte… d'après lequel a justement été baptisé ce chêne majestueux.

Nommé le «chêne de Napoléon», cet arbre planté sur le campus de l'Université de Lausanne (Unil) en a vu d'autres. A commencer par Napoléon Bonaparte donc, de passage à Saint-Sulpice pour une revue de ses troupes en 1800. Légende ou pas, c'est en tout cas ainsi que fut baptisé le vénérable sylvestre.

Malgré tous les changements dont fut témoin cet arbre, il est une chose qui est demeurée chez lui étonnamment stable: son génome. Telle est du moins la conclusion d'une très belle étude pluridisciplinaire tout juste parue dans la revue Nature Plants, étude qui pourrait amener les botanistes à revoir la manière dont ces honorables végétaux traversent les siècles.

http://education.expasy.org/cours/PO17421/publication

Dernières nouvelles…

http://www.bbc.com/news/world-europe-42805485

Boris Johnson 'is descendant' of mummified Basel woman

By Imogen Foulkes
BBC News, Basel

6 hours ago Share

- The woman is the great-great-great-great-great-great-great grandmother of UK Foreign Secretary Boris Johnson. The body was uncovered in 1975 while renovations were being done on Basel's Barfüsser Church.
- Now the scientists and the historians were sure (99.8% probability): the mummy was none other than Anna Catharina Bischoff. Born in Basel in 1719, she died there in 1787.

HUMAN GENOMICS

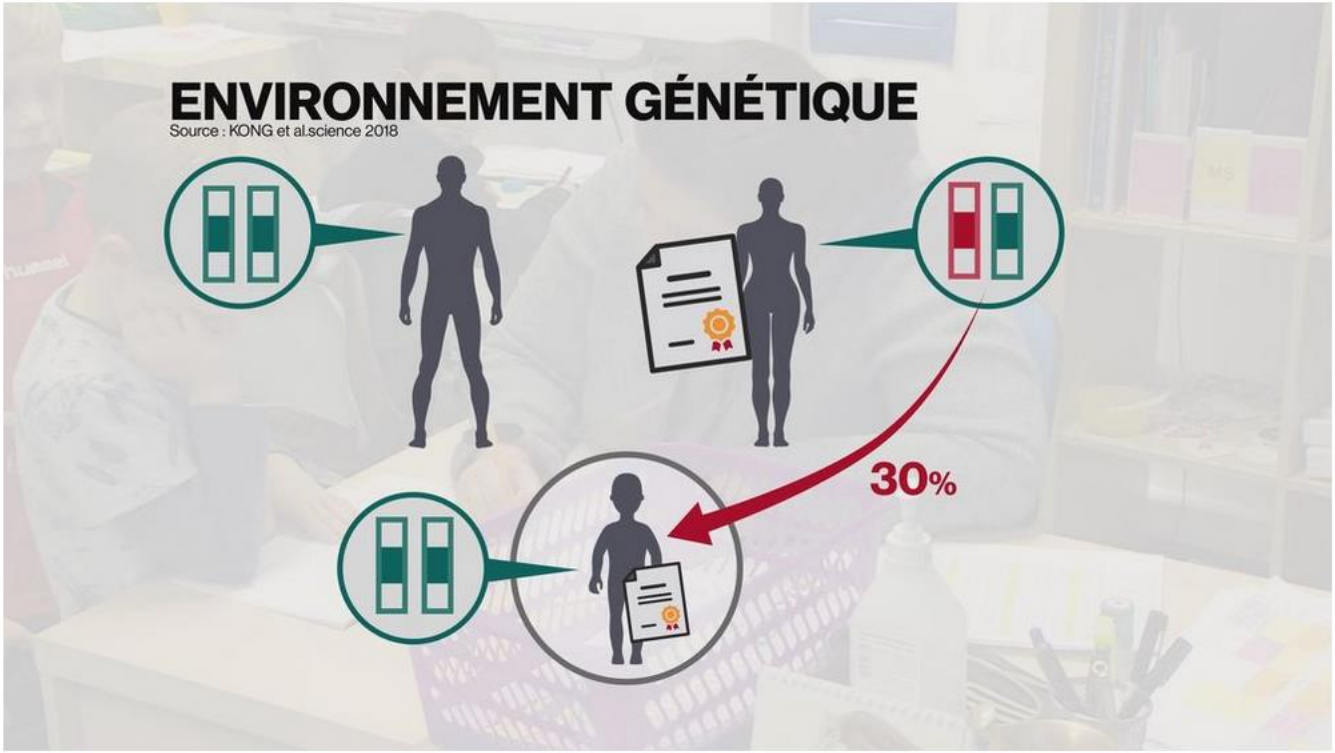
Kong et al., Science 359, 424–428 (2018)

The nature of nurture: Effects of parental genotypes

Augustine Kong,[1,2,3]* Gudmar Thorleifsson,[1] Michael L. Frigge,[1]
Bjarni J. Vilhjalmsson,[4,5] Alexander I. Young,[1,2,6] Thorgeir E. Thorgeirsson,[1]
Stefania Benonisdottir,[1] Asmundur Oddsson,[1] Bjarni V. Halldorsson,[1] Gisli Masson,[1]
Daniel F. Gudbjartsson,[1,3] Agnar Helgason,[1,7] Gyda Bjornsdottir,[1]
Unnur Thorsteinsdottir,[1,8] Kari Stefansson[1,8]*

Sequence variants in the parental genomes that are not transmitted to a child (the proband) are often ignored in genetic studies. Here we show that nontransmitted alleles can affect a child through their impacts on the parents and other relatives, a phenomenon we call "genetic nurture." Using results from a meta-analysis of educational attainment, we find that the polygenic score computed for the nontransmitted alleles of 21,637 probands with at least one parent genotyped has an estimated effect on the educational attainment of the proband that is 29.9% ($P = 1.6 \times 10^{-14}$) of that of the transmitted polygenic score. Genetic nurturing effects of this polygenic score extend to other traits. Paternal and maternal polygenic scores have similar effects on educational attainment, but mothers contribute more than fathers to nutrition- and heath-related traits.

http://education.expasy.org/cours/PO17421/publication
https://www.rts.ch/info/sciences-tech/medecine/9279297-la-reussite-scolaire-depend-de-l-environnement-social-et-genetique.html



Cet exemple montre les séquences d'ADN transmises des parents à l'enfant. La séquence rouge serait corrélée à la réussite scolaire de la mère. L'enfant n'a pas reçu cet ADN. Pourtant, au contact de sa mère, il grandit dans un environnement favorable qui augmente quand même ses chances à l'école. [- RTS]

Our analyses implicitly assume that direct genetic effects and genetic nurturing effects are additive, but interactive effects could certainly exist and further complicate the interpretation of observed effects. Moreover, alleles other than those in the parents can also have an effect; for example, the genetic makeup of the population of the probands could also be an important environmental contributor to their phenotypes.

Nature. 1988 Jun 30;333(6176):816-8.

Human basophil degranulation triggered by very dilute antiserum against IgE.

Davenas E[1], Beauvais F, Amara J, Oberbaum M, Robinzon B, Miadonna A, Tedeschi A, Pomeranz B, Fortner P, Belon P, et al.

Author information

Abstract

When human polymorphonuclear basophils, a type of white blood cell with antibodies of the immunoglobulin E (IgE) type on its surface, are exposed to anti-IgE antibodies, they release histamine from their intracellular granules and change their staining properties. The latter can be demonstrated at dilutions of anti-IgE that range from 1 x 10(2) to 1 x 10(120); over that range, there are successive peaks of degranulation from 40 to 60% of the basophils, despite the calculated absence of any anti-IgE molecules at the highest dilutions. Since dilutions need to be accompanied by vigorous shaking for the effects to be observed, transmission of the biological information could be related to the molecular organization of water.

PMID: 2455231 DOI: 10.1038/333816a0

Nature. 1994 Aug 4;370(6488):322.

Memory of water revisited.

Benveniste J, Ducot B, Spira A.

Comment on

Human basophil degranulation is not triggered by very dilute antiserum against human IgE. [Nature. 1993]

PMID: 8047128 DOI: 10.1038/370322a0

Réponse à un médicament

Voir aussi activité 3.2

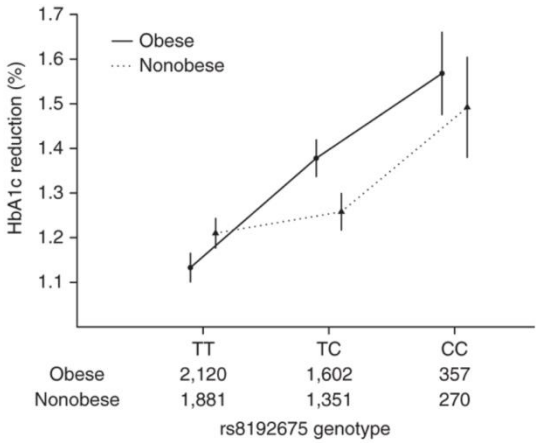
et

http://www.chromosomewalk.ch/chromosome/chromosome-10/

Réponse à un médicament

NATURE GENETICS | LETTER
日本語要約

Variation in the glucose transporter gene *SLC2A2* is associated with glycemic response to metformin



Séquence ADN

… ggggatgcaatagtagtagtggtactgtagaggatgaagtagatgggtgca …

Séquence du gène SLC2A2 / GLUT2 qui est impliqué dans le transport du glucose. **Les patients avec la variation 't -> c' répondent mieux au traitement Metformin** (Glucophage, Stagid, …), un antidiabétique oral de la famille des biguanides normoglycémiant utilisé dans le traitement du diabète de type 2

"The normal dose of metformin used to treat patients with diabetes is between 500mg and 2000mg. We have found that overweight people who carry two copies of the genetic variant c respond much better to metformin, equivalent to receiving an extra 550mg of the drug."

Chromosome 3: SLC2A2 (rs8192675**)**

Nature Genetics 48, 1055–1059 (2016)

Table 1 Currently available antidiabetic drugs and their associated candidate genes involved in efficacy/toxicity

Class	Common medical representatives	Mechanism of action	Candidate genes involved in pharmacotherapy	Ref.
Biguanide	Metformin	AMP-kinase activation	*SLC22A1, SLC22A2, SLC22A3, SLC47A1, SLC47A2*	[28-39]
Sulfonylureas	Gliburide, gliclazide, Glimepiride, glipizide	Inhibition of KATP channel on plasma membrane of β-cells	*KCNJ11, ABCC8, CYP2C9, TCF7L2*	[8,10,48-91]
Thiazolidinediones	Pioglitazone, rosiglitazone	Activates PPAR-γ	*PPAR-γ, ADIPOQ, TNF-α, LEP, CYP2C8*	[92-131]
Meglitinides	Nateglinide, repaglinide	Inhibition of KATP channel on Plasma Membrane of β-cells	*SLCOB1, CYP2C8, KCNQ1, SLC30A8, KCNJ11, TCF7L2*	[78,106,132-144]
DPP-4 inhibitors	Alogliptin, linagliptin, saxagliptin, sitagliptin, vildagliptin	Inhibits DPP-4, Affect GLP-1 receptor pathway	Possibly *TCF7L2*	[145-148]
α-glucosidase inhibitors	Acarbose, miglitol, voglibose	Inhibits intestinal α-glucosidase	Yet to identify?	[10]
SGLT-2 inhibitors	Canagliflozin, dapagliflozin, empagliflozin	Inhibits SGLT2 transporters in kidney	Yet to identify?	[10]
GLP-1 agonist	Exenatide, liraglutide	Activate GLP-1 receptor	Yet to identify?	[10]

DPP-4: Dipeptidyl peptidase-4; SGLT-2: Sodium glucose transporter-2; GLP-1: Glucagon like peptide-1; KATP: ATP-sensitive potassium channel; PPARγ: Peroxisome proliferator-activated receptor γ.

Type 2 diabetes

About **415 million people** are suffering from T2DM (642 millions in 2040)

About **70 genetic loci** have been identified to be associated with T2DM

Over the last decade, the number of available oral antidiabetic drugs (OADs) has considerably increased.

However, clinical treatment of T2DM patients has become more complex due to different degrees of therapeutic outcomes.

*Personalized differences during OADs therapeutics have been linked with numerous variants related to **drug-transporters**, **drug-targets**, **drug catabolizing enzymes** and T2DM risk genes. In addition, CYP gene encoding **Cytochrome P450 enzymes** also play a crucial role with respect to metabolism of drugs.*

REVIEW

Pharmacogenetic studies update in type 2 diabetes mellitus

Août 2016

LE TEMPS

Septembre 2015

Médecine personnalisée, la révolution

> Santé Un programme a été lancé en juin aux Etats-Unis pour faire progresser les thérapies ciblées sur le profil génétique

> Les grandes pharmaceutiques, en particulier Roche et Novartis, sont à la pointe

Willy Boder

Début juillet a commencé, dans 2400 hôpitaux américains, une vaste opération de dépistage de patients atteints d'une forme ou d'une autre de cancer.

Contrairement aux études cliniques traditionnelles, l'objectif n'est pas de recruter des malades souffrant d'un même type de cancer, du sein, des poumons, ou de la prostate par exemple, pour tester l'efficacité d'un seul nouveau médicament. Il s'agit, au contraire, de rassembler les patients selon le profil génétique de leur tumeur, toutes catégories confondues.

Les médecins administreront une vingtaine de médicaments, dont certains proviennent de groupes pharmaceutiques suisses, qui sont déjà sur le marché ou encore en développement. Une patiente souffrant d'un cancer du sein avancé recevra par exemple une thérapie normalement destinée à un patient souffrant d'un cancer du poumon.

Cette nouvelle approche des programmes d'essais cliniques via le profil génétique des tumeurs et des patients, désignée par le terme «essais en corbeille» est suivie d'un œil bienveillant par la Food and Drug Administration (FDA), chargée du contrôle des médicaments aux Etats-Unis. Cette vaste étude s'intègre dans la tendance scientifique générale suivie actuellement par tous les grands groupes pharmaceutiques, à savoir rendre chaque médicament plus efficace en le modulant selon le profil génétique de la maladie dont souffre exactement tel ou tel patient.

La méthode d'essais cliniques en «corbeille», menée aux Etats-Unis sous le contrôle de l'Institut national du cancer (NCI), vise à élargir et à affiner ce qu'on appelle la médecine personnalisée, ou la médecine de précision, selon le terme utilisé par Barack Obama. Le président des Etats-Unis est persuadé de pouvoir faire progresser rapidement la recherche scientifique de cette manière et tente d'obtenir un financement par le Congrès de ces nouvelles formes de thérapie.

Joe Jimenez, patron de Novartis, deuxième entreprise au monde, derrière Roche, dans la mise à disposition de médicaments contre le cancer, estime qu'un quart des médicaments sont actuellement gaspillés de différentes manières. L'une des principales causes de ce gaspillage, qui coûte très cher aux systèmes de santé, provient du manque de connaissances scientifiques précises sur le profil génétique de telle ou telle maladie à mettre en corrélation avec l'ADN du patient. Le médecin est dès lors contraint de tester plusieurs médicaments sur un patient avant de trouver celui qui est le plus efficace.

Cette approche empirique se produit pour de très nombreuses maladies, de l'hypertension aux maladies infectieuses, en passant par des affections très graves. Pour les maladies où le pronostic vital est engagé, comme certaines formes de cancer, ce tâtonnement, synonyme de perte de temps, peut conduire à la mort.

Selon les spécialistes, la médecine de précision a aussi pour avantage de réduire les coûts de la santé. Les médicaments de ce type associés à un diagnostic (biomarqueur ou test spécifique d'ADN accompagnant le traitement) sont, certes, nettement plus chers que les produits thérapeutiques traditionnels, mais ils évitent le tâtonnement médical et la facturation aux caisses maladie de médicaments inefficaces.

> Rassembler les patients selon le profil génétique de leur tumeur, toutes catégories confondues

Le programme du NCI, qui touche 2400 hôpitaux et 1000 patients sélectionnés au sein d'un groupe de 3000 malades du cancer, est doté d'un budget de 30 à 40 millions de dollars. Le critère de succès de la thérapie sera principalement basé sur une réduction rapide de la tumeur d'au moins un tiers.

GETTY IMAGES

La médecine personnalisée combine profil génétique et médicaments thérapeutiques de manière ciblée. ARCHIVES

Depuis le premier séquençage d'un génome humain, en 2003, l'analyse des maladies, en particulier le cancer, repose de manière de plus en plus précise sur la découverte de mutations génétiques et l'activation ou la désactivation de protéines, au sein même ou à la surface des cellules. Ces mécanismes provoquent la prolifération des cellules cancéreuses dans l'organe touché, puis sous forme de métastases.

Les cancers ne sont désormais plus considérés comme différentes formes d'une même maladie, mais comme une multitude de maladies ayant des caractéristiques et un profil génétique propre. Le cancer évolue différemment selon chaque patient, ce qui entraîne, si le mécanisme génétique qui dysfonctionne peut être ciblé et corrigé, la prescription d'un médicament spécifique et une approche thérapeutique personnalisée.

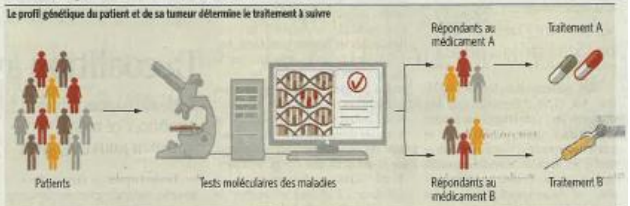
«Grâce la médecine personnalisée nous obtenons immédiatement des taux de réponse inédits aux traitements. Il y a vraiment un changement de paradigme dans ce domaine scientifique», s'enthousiasme Dietmar Berger, responsable du développement de Roche en oncologie. Il cite le cas du médicament Alectinib, contre le cancer du poumon au stade avancé, en phase d'homologation. Les métastases dans le cerveau se réduisent rapidement avec un taux de réponse jusqu'à 70% d'une durée jusqu'à onze mois.

Roche place aussi beaucoup d'espoir dans Atezolizumab, un médicament, associé à la présence de la protéine PD-L1, qui fait appel à la stimulation du système immunitaire pour détruire les cellules cancéreuses. Ce mécanisme, identique chez certains patients spécialement diagnostiqués, peut être appliqué dans la lutte contre les cancers du poumon, de la prostate, du sein et du rein. «70% de nos médicaments en phase clinique II et III en oncologie ont désormais un test diagnostic spécifique associé», constate Dietmar Berger.

«On assiste réellement à une révolution dans le traitement de ces maladies, confirme Severin Schwan, patron de Roche. Auparavant, un cancer métastasé menait le plus souvent à la mort. Aujourd'hui, dans de nombreux cas, grâce à la médecine personnalisée, on peut prolonger la vie durant des années, et peut-être guérir de ce type de cancer.»

Comment ça marche

Le profil génétique du patient et de sa tumeur détermine le traitement à suivre



La majorité des thérapies contre le cancer à l'étude sont associées à un ou plusieurs tests génétiques développés en parallèle avec l'élaboration finale de la substance active contenue dans le médicament. Le taux d'efficacité de la thérapie, ainsi adaptée au profil précis du patient et de sa tumeur, peut dépasser 70% et augmente fortement ses chances de survie à la suite d'un cancer métastasé. SOURCE: PFIZER