



**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE MINISTERE DE
L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE UNIVERSITÉ
ABOU BEKR BELKAID – TLEMCEN**



**FACULTÉ :SNV
Département de Biologie
Filière :Biologie moléculaire et cellulaire (M1)**

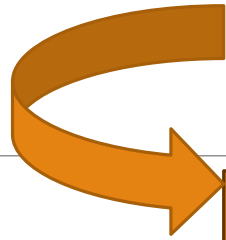
Genome Browsers , identification des SNP « ensembl »

Présenté Par : Dr SALMI Takwa

Plan de travail

- ❖ Structure et Organisation des Banques
- ❖ Formats (séquences)
- ❖ Les banques génomiques
- ❖ Exemple de manipulation sur Ensembl , comment identifier des SNP ?

Structure et organisation des banques

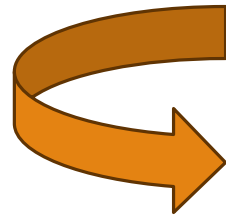


Fichier et formats

Est-ce que toutes les banques utilisent les formats pour présenter les informations ?



Harmonisation !



Harmoniser le classement des informations
Utiliser un langage commun pour échanger des informations entre toutes les bases

Genbank (NCBI)

EMBL

Formats (séquences)

1

Format Genbank

2

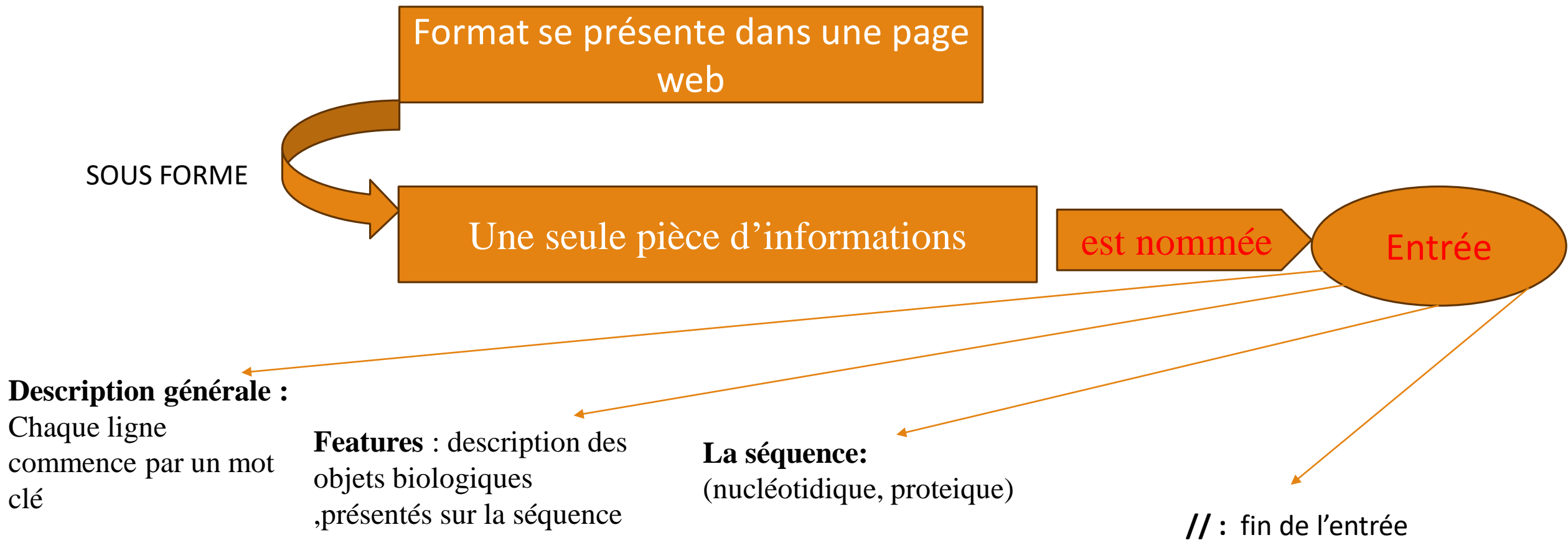
Format EMBL

3

Format FAST


Les séquences (formats) sont stockées en général sous forme de fichiers texte

Le format correspond à l'ensemble des règles :



Format Genbank

Format Genbank (NCBI, nucléotides)

GenBank 

Homo sapiens BRCA1 DNA repair associated (BRCA1), RefSeqGene (LRG_292) on chromosome 17

NCBI Reference Sequence: NG_005905.2

[FASTA](#) [Graphics](#)

Go to: 

```

LOCUS       NG_005905                193689 bp    DNA     linear   PRI 09-OCT-2023
DEFINITION  Homo sapiens BRCA1 DNA repair associated (BRCA1), RefSeqGene
            (LRG_292) on chromosome 17.
ACCESSION   NG_005905
VERSION     NG_005905.2
KEYWORDS    RefSeq; RefSeqGene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 193689)
AUTHORS    Orban TI and Olah E.
TITLE      Emerging roles of BRCA1 alternative splicing
JOURNAL    Mol Pathol 56 (4), 191-197 (2003)
PUBMED    12890739
REMARK     Review article
REFERENCE   2  (bases 1 to 193689)
AUTHORS    Orban TI and Olah E.
TITLE      Expression profiles of BRCA1 splice variants in asynchronous and in
            G1/S synchronized tumor cell lines
JOURNAL    Biochem Biophys Res Commun 280 (1), 32-38 (2001)
PUBMED    11162473
REFERENCE   3  (bases 1 to 193689)
AUTHORS    Paterson JM

```

Forme EMBL

Forme EMBL /EBI: <https://www.ebi.ac.uk/>

```
ID      13      standard; DNA; HTG; 84195 BP.
XX
AC      chromosome:GRCh37:13:32889611:32973805:1
DT      12-NOV-2011
DE      Homo sapiens chromosome 13 GRCh37 partial sequence 32889611..32973805
DE      annotated by Ensembl
OS      Homo sapiens (human)
OC      Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia;
...
OC      Hominoidea; Hominidae; Homininae; Homo.
XX
CC      This sequence was annotated by the Ensembl system. Please visit the Ensembl
CC      web site, http://www.ensembl.org/ for more information.
XX
FH      Key          Location/Qualifiers
FT      gene          1..84195
FT                      /gene=ENSG00000139618
FT                      /locus_tag="BRCA2"
FT                      /note="breast cancer 2, early onset [Source:HGNC
FT                      Symbol;Acc:1101]"
FT      mRNA          join(1..194,949..1054,3604..3852,9603..9711,10628..10677,
FT                      10769..10809,11026..11140,13970..14019,15446..15557,
...
FT                      /gene="ENSG00000139618"
FT                      /note="transcript_id=ENST00000380152"
...
FT                      /translation="MPIGSKERPTFFEIFKTRCNKADLGPISLWFEELSSEAPPYNSE
FT                      PAEESSEHKNNNYEPNLFKTPQRKPSYNQLASTPIIFKEQGLTLPLYQSPVKELDKFKLD
```

ID : Identificateur, c'est le nom de l'entrée contenant la séquence. Cette ligne a la structure suivante : nom de l'entrée ; classe de la donnée ; molécule ; division ; longueur. Le nom est suivi de l'indication de la classe de donnée, puis du type de molécule ADN, ARN ou ADNc (XXX si l'entrée n'a pas été annotée) ; ensuite la division à laquelle l'entrée appartient et enfin la longueur de la séquence en paires de bases (bp).

AC : Numéro d'accession de l'entrée qui ne varie pas au cours des versions successives de la banque. Il peut y avoir plusieurs numéros d'accessions pour une même entrée. En effet lorsque deux entrées sont fusionnées en une seule, un nouveau numéro peut être attribué à la nouvelle entrée et ceux provenant des ex-entrées indépendantes sont conservés

Forme FASTA

Format FAST(commence par l'indice plus grand)

Homo sapiens chromosome 17, GRCh38.p14 Assemblée primaire

Séquence de référence NCBI : NC_000017.11

[Graphiques](#) [GenBank](#)

>NC_000017.11:c43170327-43044295 Homo sapiens chromosome 17, GRCh38.p14 Assemblée primaire

```
CTTTTGTCCC GCCCTT CCTCTGACTGTGTCTTGATTTCCTATTCTGAGAGGCTATTGCTCAGCGGTTTC
CGTGGAACAGTAAAGCGTGGGAATTACAGATAAAATTTAAACTGTGGAACCCCTTTCCTCGGCTGCCGCC
AAGGTGTTCCGGTCTTCCGAGGAAGCTAAGGCCGCGTTGGGGTGAGACCCTCACTTCATCCGGTGAGTAG
CACCGCGTCCGGCAGCCCCAGCCCCACACTCGCCCCGCGCTATGGCCTCCGTCTCCCAGCTTGCCTGCATC
TACTCTGCCCTCATTCTGCACGGACTATGAGGTGACCTTTACGGAGGATAAGATCAATGCCCTTATTAAG
CAGCCAGTGTAAATATTGAACTTTTTGGCCTGGCTTGTTTGCAAAGGTCCTGGCCAACGTCAACATTGG
GAGCCACATCTGCAGTGTAGAGGGGGGGAAAAAACGTGACTGCGCGTCGTGAGCTCGCTGAGACGTTCT
GGACGGGGGACAGGCCGTGGGGTTTCTCAGATAACTGGGCCCTGGGCTCAGGAGGCCTGGCACCCCTCTGC
TCTGGGTTAAGGTAGAAGAGCCCCGGGAAAGGGACAGGGGCCAAGGGATGCTCCGGGGGACGGGCGGGG
GAAAGTGAATTTCCGAAGCTAGGCAGATGGGTATTCTTATGCGAGGGGCGGGGGCGGAACCTGGAGAGGCA
TAAGGCGTTGTGAACCCCCGGGGAAGGGGGCAGTTTGTAGGTCTCGAGGGAAGCACTAAGGATCAGGTT
GGGGGCACAGTGTGTCCGAGGAGGAATCCTCCTGATAGGAACTGGAATGTGCCTTGAAGGGGACACCATG
TGTATAAGAACATCAGCTGGTCGCCGGGGATGGTGGCTTACGCCTGTATTCTAGCACTTTTGGGAGGCCA
AGGCGGATGGATCACGAGGTGAGGAGTTCGAGACCAGCCTGACCATCGTGGTCAAACCCCGTCTCTACTA
AAAATACAAAAATTAGCCGGGCGTGGTGGCGCGCCGCCAGCTACTCAGGAGCTGAGGCAGGAGAATCGCTT
GAACCCAGGAGGCGGAGGTTGCAGTGAGCCGAGATCGCGCCATTGCACTCCAGCCTGGGTGGGCAGAACGA
```

FASTA

Homo sapiens chromosome 17, GRCh38.p14 Assemblée primaire

Séquence de référence NCBI : NC_000017.11

[Graphiques](#) [GenBank](#)

```
>NC_000017.11:c43170327-43044295 Homo sapiens chromosome 17, GRCh38.p14 Assemblée primaire
CTTTGTCCCGCCCTTCTCTGACTGTGTCTTGATTTCTATTCTGAGAGGCTATTGCTCAGCGGTTTC
CGTGGCAACAGTAAAGCGTGGGAATTACAGATAAATTAACACTGTGGAACCCCTTTCCTCGGCTGCCGCC
AAGGTGTTCCGGTCTTCCGAGGAAGCTAAGGCCGCGTTGGGGTGAGACCCTCACTTCATCCGGTGAGTAG
CACCGGTCCGGCAGCCCCAGCCCCACACTCGCCCGCGCTATGGCTCCGTCTCCAGCTTGCCTGCATC
TACTCTGCCCTCATTCTGCACGGACTATGAGGTGACCTTACGGAGGATAAGATCAATGCCCTTATTAAG
CAGCCAGTGTAATATTGAACTTTTTGGCCTGGCTTGTGGCAAAGTCTGGCCAACGTCAACATTGG
GAGCCACATCTGCAGTGTAGAGGGGGGAAAAAACGTGACTGCGCGTCGTGAGCTCGCTGAGACGTTCT
GGACGGGGGACAGGCCGTGGGGTTTCTCAGATAACTGGGCCCTGGGCTCAGGAGGCCTGGCACCTCTGC
TCTGGGTTAAGGTAGAAGAGCCCCGGGAAAGGGACAGGGGCCAAGGGATGCTCCGGGGGACGGGCGGGG
GAAAGTGAATTTCCGAAGCTAGGCAGATGGGTATTCTTATGCGAGGGGGCGGGGGCGGAACCTGGAGAGGCA
TAAGGCGTTGTGAACCCCCGGGGAAGGGGGCAGTTTGTAGGTCTCGAGGGAAGCACTAAGGATCAGGT
GGGGGCACAGTGTGTCGAGGAGGAATCCTCCTGATAGGAACTGGAATGTGCCTTGAAGGGGACACCATG
TGTATAAGAACATCAGCTGGTCGCCGGGGATGGTGGCTTACGCCTGTATTCCTAGCACTTTTGGGAGGCCA
AGGCGGATGGATCACGAGGTCAGGAGTTCGAGACCAGCCTGACCATCGTGGTAAAACCCCGTCTCTACTA
AAAATACAAAATTAGCCGGGCGTGGTGGCGCGCCGCCAGCTACTCAGGAGCTGAGGCAGGAGAATCGCTT
GAACCCAGGAGGCGGAGGTTGCAGTGAGCCGAGATCGCGCCATTGCACTCCAGCCTGGGTGGGCAGAACGA
```

Exemples des navigateurs génomiques

Genome Browsers= Navigateurs génomiques

Ensembl « EBI »

UCSC Genome Browser

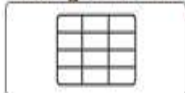
NCBI –MAP viewer



Visionneuse de données génomiques

GDV prend en charge l'exploration et l'analyse d' assemblages de génomes eucaryotes annotés NCBI et sélectionnés non annotés NCBI. Actuellement, des assemblages de plus de 2 420 organismes sont disponibles.

Changer de vue

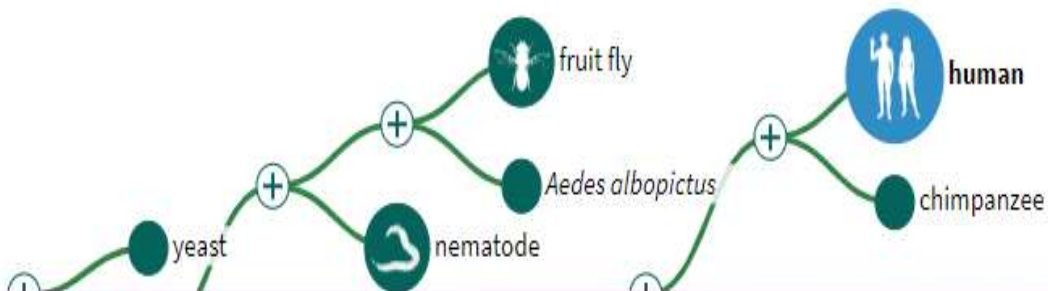


Rechercher des organismes

Homo sapiens (human)

Pour afficher plus d'organismes dans l'arborescence, cliquez sur les nœuds comportant des signes « + ». Appuyez et maintenez le « + » pour développer et révéler tous les sous-groupes. Ou recherchez un organisme en utilisant le champ de recherche ci-dessus.

Nouveau! Cliquez sur **Changer de vue** en haut pour voir une autre façon de naviguer dans les génomes.



Homo sapiens (humain)



Rechercher dans le génome

Location, gene or phenotype



Exemples: TP53,chr17 : 7667000-7689000,Réparation de l'ADN

Assemblée

GRCh38.p14



Browse/Select Species

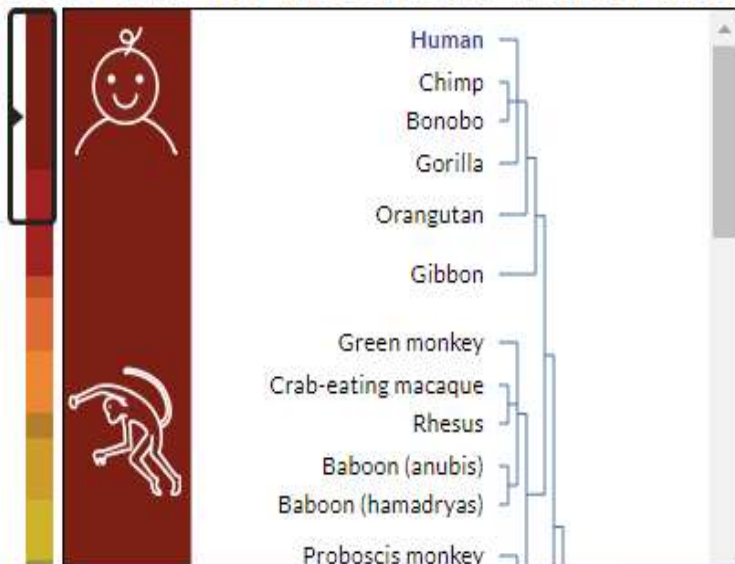
POPULAR SPECIES



Search through thousands of genome browsers

[Unable to find a genome? Send us a request.](#)

UCSC SPECIES TREE AND CONNECTED ASSEMBLY HUBS



Find Position

Human Assembly



Position/Search Term

Current position: chr8:39,771,328-39,786,309

Human Genome Browser - hg19 assembly

[view sequences](#)

The February 2009 human reference sequence (GRCh37) was produced by the **Genome Reference Consortium**. For more information about this assembly, see [GRCh37](#) in the NCBI Assembly database.

Sample position queries

A genome position can be specified by the accession number of a sequenced genomic clone, an mRNA or EST or STS marker, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the human genome. See the [User's Guide](#) for more information.

Request:

Genome Browser Response:

chr7

Displays all of chromosome 7

chrUn_gl000212

Displays all of the unplaced contig gl000212



Homo sapiens
 (Graphic courtesy of CBSE)

Outils

[Tous les outils](#)

BioMart >

Exportez des ensembles de données personnalisés depuis Ensembl avec cet outil d'exploration de données

BLAST/BLAT >

Recherchez dans nos génomes votre séquence d'ADN ou de protéine

Prédicteur d'effet de variante >

Analysez vos propres variantes et prédiriez les conséquences fonctionnelles des variantes connues et inconnues

Recherche

 pour

par exemple [BRCA2](#) ou [rat 5:62797383-63627669](#) ou [rs699](#) ou [maladie coronarienne](#)

Ensembl est un navigateur **génomique pour les génomes vertébrés** qui prend en charge la recherche en génomique comparative, évolution, variation de séquence et régulation transcriptionnelle. Ensembl annote les gènes, calcule plusieurs alignements, prédit la fonction de régulation et collecte des données sur les maladies. Les outils Ensembl incluent BLAST, BLAT, BioMart et le Variant Effect Predictor (VEP) pour toutes les espèces prises en charge.

Ensembl version 110 (juillet 2023)

- L'assemblage du génome humain a été mis à jour avec la dernière version du correctif GRCh38.p14
- Annotation réglementaire Une annotation des régions ouvertes de chromatine et des promoteurs chez le porc, le poulet, le saumon de l'Atlantique, le turbot et le bar européen a été ajoutée (une collaboration avec les consortiums GENE-SWitCH et AQUA-FAANG)
- Trois souches de rats ont été ajoutées : SHR/Utx RGD_8142385, WKY/Bbb RGD_1581635 et SHRSP/BbbUtx.
- Attributs de transcription canoniques MANE et Ensembl à ajouter en tant que balises dans les fichiers GFF3

[Plus d'actualités](#) sur notre blog

Ensembl libération rapide

Nouveaux assemblages avec annotation des gènes et des protéines toutes les deux semaines .

Remarque : les espèces qui existent déjà sur ce site

Tous les génomes



Races de porcs

Génome de référence porcine et 12 races supplémentaires

Génomes préférés



Humain

GRCh38.p14

[Vous utilisez toujours GRCh37 ?](#)



Souris

Genome Browsers« Ensembl »

<https://www.ensembl.org/index.html>

Banque de données spécialiste

génomique

Initiée en 1999

Projet en collaboration entre l'European Bioinformatics Institute (EBI) et le Wellcome Trust Sanger Institute (WTSI)

Restrict category to:

Gene	229
Transcript	324
GeneTree	164
ProbeFeature	12

Restrict species to:

Human	37
Mouse	64
Zebrafish	3
African green monkey	3
African savanna elephant	3
Algerian mouse	4
... 158 more species ...	

Per page:

10	25	50	100
----	----	----	-----

Layout:

Standard	Table
----------	-------

729 results match IDO1

[IDO1 \(Human Gene\)](#)
ENSG00000131203 8:39902275-39928790:1
 Indoleamine 2,3-dioxygenase 1 [Source:HGNC Symbol;Acc:HGNC:6059]

INDOLEAMINE 2,3-DIOXYGENASE; **IDO1** [*147435] (MIM gene record; description: INDOLEAMINE 2,3-DIOXYGENASE; **IDO1**;;INDO;;IDO,) is an external reference matched to Gene ENSG00000131203
[Variant table](#) • [Phenotypes](#) • [Location](#) • [External Refs.](#) • [Regulation](#) • [Orthologues](#) • [Gene tree](#)

[IDO1-201 \(Human Transcript\)](#)
ENST00000253513 8:39913869-39928444:1
 Indoleamine 2,3-dioxygenase 1 [Source:HGNC Symbol;Acc:HGNC:6059].
[Location](#) • [External Refs.](#) • [cDNA seq.](#) • [Exons](#) • [Variant table](#) • [Protein seq.](#) • [Population](#) • [Protein summary](#)

[IDO1-205 \(Human Transcript\)](#)
ENST00000521480 8:39922552-39923931:1
 Indoleamine 2,3-dioxygenase 1 [Source:HGNC Symbol;Acc:HGNC:6059].
[Location](#) • [External Refs.](#) • [cDNA seq.](#) • [Exons](#) • [Variant table](#) • [Population](#)

[IDO1-206 \(Human Transcript\)](#)
ENST00000521636 8:39917885-39922911:1
 Indoleamine 2,3-dioxygenase 1 [Source:HGNC Symbol;Acc:HGNC:6059].
[Location](#) • [External Refs.](#) • [cDNA seq.](#) • [Exons](#) • [Variant table](#) • [Population](#)

[IDO1-207 \(Human Transcript\)](#)
ENST00000522495 8:39902375-39928426:1
 Indoleamine 2,3-dioxygenase 1 [Source:HGNC Symbol;Acc:HGNC:6059].
[Location](#) • [External Refs.](#) • [cDNA seq.](#) • [Exons](#) • [Variant table](#) • [Protein seq.](#) • [Population](#) • [Protein summary](#)

- Transcript-based displays**
- Summary
 - Sequence
 - Exons**
 - cDNA
 - Protein
 - Protein Information
 - Protein summary
 - Domains & features
 - Variants
 - PDB 3D protein model
 - AlphaFold predicted model
 - Genetic Variation
 - Variant table
 - Variant image
 - Haplotypes
 - Population comparison
 - Comparison image
 - External References
 - General identifiers
 - Oligo probes
 - Supporting evidence
 - ID History
 - Transcript history
 - Protein history

Transcript: ENST00000518237.6 IDO1-202

Description indoleamine 2,3-dioxygenase 1 [Source:HGNC Symbol;Acc:HGNC:6059]

Gene Synonyms IDO, INDO

Location [Chromosome 8: 39,913,891-39,928,790](#) forward strand.

About this transcript This transcript has [10 exons](#), is annotated with [168 domains and features](#), is associated with [4718 variant alleles](#) and maps to [415 oligo probes](#).

Gene This transcript is a product of gene [ENSG00000131203.13](#) Hide transcript table

Show/hide columns (1 hidden)

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags		
ENST00000518237.6	IDO1-202	1849	403aa	Protein coding	CCDS47847	A0A348GSI3 P14902	NM_002164.6	MANE Select	Ensembl Canonical	GENCODE basic
ENST00000522495.5	IDO1-207	1709	403aa	Protein coding	CCDS47847	A0A348GSI3 P14902	-			GENCODE basic APPRIS
ENST00000519154.5	IDO1-204	1138	180aa	Protein coding		A0A140T9Z2	-			TSL:5 CDS 3' inc
ENST00000518804.5	IDO1-203	612	78aa	Protein coding		ESRIX2	-			TSL:4 CDS 3' inc
ENST00000522840.1	IDO1-208	572	47aa	Protein coding		ESRH36	-			TSL:4 CDS 3' inc
ENST00000253513.11	IDO1-201	1525	170aa	Nonsense mediated decay		J3KN03	-			TSL:5
ENST00000521636.1	IDO1-206	700	No protein	Retained intron		-	-			TSL:3
ENST00000521480.1	IDO1-205	563	No protein	Retained intron		-	-			TSL:2
ENST00000523779.1	IDO1-209	511	No protein	Retained intron		-	-			TSL:2

Exons ?

Download sequence

- Configure this page
- Custom tracks
- Export data
- Share this page
- Download sequence

- Résumé
- Séquence
 - Exons
 - ADNc
 - Protéine
- Informations sur les protéines
 - Résumé des protéines
 - Domaines et fonctionnalités
 - Variants
 - Modèle de protéine 3D PDB
 - Modèle prédit AlphaFold
- Variation génétique
 - Tableau des variantes
 - Image variante
 - Haplotypes
 - Comparaison de la population
 - Image de comparaison
- Références externes
 - Identifiants généraux
 - Sondes oligo
 - Preuve à l'appui
- Historique d'identification
 - Historique des transcriptions
 - Histoire des protéines

- Configurer cette page
- Pistes personnalisées
- Exporter des données
- Partager cette page
- Marquer cette page

Description Réparation de l'ADN BRCA1 associée [Source : Symbole HGNC ; Acc : [HGNC : 1100](#)]

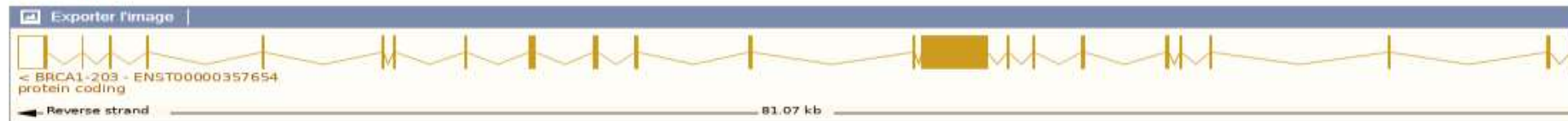
Synonymes de gènes BRCC1, FANCS, PPP1R53, RNF53

Emplacement [Chromosome 17 : 43 044 295-43 125 364](#) brin inversé.

A propos de cette transcription Ce transcrit comporte **23 exons**, est annoté avec **120 domaines et caractéristiques**, est associé à **31 848 allèles variants** et correspond à **1 133 sondes oligo**.

Gène Cette transcription est un produit du gène [ENSG0000012048.25](#) [Afficher le tableau de transcription](#)

Résumé



Statistiques Exons : 23, Exons codants : 22, Longueur de transcription : 7 088 bps, Longueur de traduction : 1 863 résidus

CRNIÈRE Ce relevé de notes MANE Select contient [ENSP00000350283](#) et correspond à [NM_007294.4](#) et [NP_009225.1](#)

Uniprot Ce transcrit correspond aux identifiants Uniprot suivants : [P38398](#)

CCDS Cette transcription fait partie de l'ensemble Human CCDS : [CCDS11453](#)

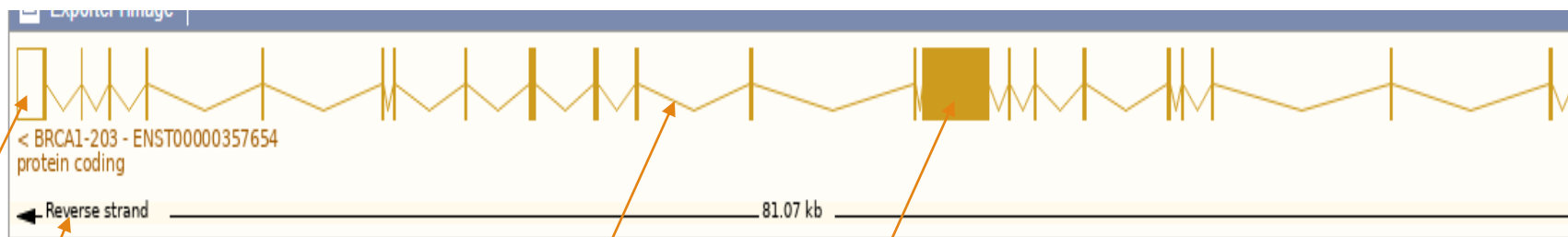
Niveau de prise en charge des transcriptions (TSL) [TSL : 1](#)

Version ENST00000357654.9

Taper Codage des protéines

Méthode d'annotation Transcription où la transcription du gène Ensembl et l'annotation manuelle de La Havane ont la même séquence, pour chaque paire de bases. Voir [article](#).

Gène de base GENCODE Cette transcription fait partie de l'ensemble de gènes [de base Gencode](#).



Région 3' UTR

Sens du gène

Intron

Partie codante de l'exon

Exons


Exons/Introns [Séquence traduite](#) [Séquence d'accompagnement](#) [Séquence d'introns](#) [UTR](#)

Variantes [3 premiers UTR](#) [5 premiers UTR](#) [Séquence de codage](#) [Changement de cadre](#) [Suppression Inframe](#)

[Insertion Inframe](#) [Faux-sens](#) [Variante modifiant les protéines](#) [Accepteur d'épissure](#) [Donateur d'épissage](#)

[Région d'épissure](#) [Commencez perdu](#) [Début retenu](#) [Arrêt gagné](#) [Arrêt retenu](#) [Synonyme](#)

Balisateur chargé

Montrer **Tous** entrées 

Non.	Exon/Intron	Commencer	Fin	Phase de démarrage	Phase finale	Longueur	Séquence
	Séquence amont de 5'					agcgcgggaattacagataaattaaactgogactgogcggcgtgagctc
1	ENSE00001852567	43 125 364	43 125 271	-	-	94	G C T G A G A C T T C C T G G A C G G G G G A C A G G C T G T G G G T T T C T C A G G A G G C C T T C A C C C T C T G C T C T G G G T A A A G
	Introns 1-2	43 125 270	43 124 116			1 155	gtagtagagtcccgggaaaggaca.....tatatgtttttctaagtgttaaag
2	ENSE00003559512	43 124 115	43 124 017	-	2	99	T T C A T T G G A A C A G A A A G A A A T G G A T T T A T C T G C T C T C A T T A A T G C T A T G C A G A A A T C T T A G A G T G T C C C A T C
	Intron2-3	43 124 016	43 115 780			8 237	gtaagtcagcacaagagtgtattaa.....tttttctccccctaccctgctag
3	ENSE00003510592	43 115 779	43 115 726	2	2	54	T C T G G A G T T G A T C A A G G A A C C T G T C T C C A C A A A G T G
	Intron3-4	43 115 725	43 106 554			9 192	gtaagtttgaatgtttatgtggct.....ttctttctttctttataatttatag

Identification des variants génétiques

Définition : variants :

versions des gènes créés par des mutations au sein des individus d'une même espèce , chaque individu étant génétiquement différent des autres individus

Mutation ponctuelle isolée

Polymorphisme d'un seul nucléotide (polymorphisme nucléotidique).

Variation stable de la séquence d'ADN génomique

portant sur une seule base

Beaucoup de SNP n'ont pas d'implications fonctionnelles mais ils définissent un locus unique dans le génome et sont polymorphes.

SNP (pour *single nucleotide polymorphism*)

rs4275724

Mutation ponctuelle

Sans modification du nombre de (désoxy)nucléotides

Avec modification du nombre de (désoxy)nucléotides

SUBSTITUTION ou INVERSION

DÉLÉTION ou INSERTION

Pas de changement
d'acide aminé :
mutation
silencieuse

Apparition
d'un codon stop
et formation
d'une protéine tronquée
non fonctionnelle :
mutation non sens

Changement
d'acide aminé :
mutation
faux sens

Décalage du cadre
de lecture :
modification
des acides aminés
en aval de la mutation

Pas de changement
de phénotype

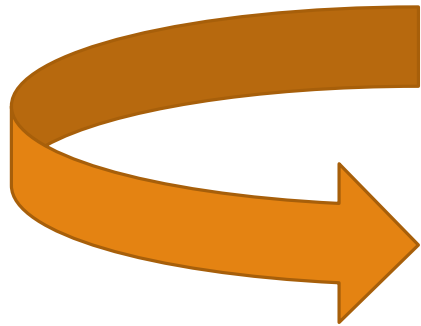
Changement
de phénotype

EXONS

Identifier toutes les variantes génétiques d'un gène

Pour identifier toutes les variantes génétiques d'un gène :

- il faut cliquer sur l'option variantes (qui se trouve) dans la barre , à gauche
- après avoir identifié le gène dans la banque de données Ensembl



Identifier toutes les variantes génétiques d'un gène

Residue	Variant ID	Conseq. Type	Source	Evidence	Alleles	Ambig. code	Residues	Codons	SIFT	Poly-Phen	CADD	REVEL	MetaLR	Mutation Assessor
10	rs746614436	missense variant	dbSNP	gnomAD ExAC AD	A/G	R	T, A	ACA, GCA	0.39	0.027	0	0.024	0.071	0.144
12	rs780742846	synonymous variant	dbSNP	gnomAD ExAC AD	T/C	Y	S	AGT, AGC	-	-	-	-	-	-
12	COSV53715809	coding sequence variant	COSMIC		COSMIC_MUTATION	-	-	-	-	-	-	-	-	-
13	rs1198318159	missense variant	dbSNP	gnomAD	A/G	R	K, R	AAA, AGA	0.2	0	10	0.082	0.068	0.088
14	rs1022734986	missense variant	dbSNP	gnomAD	A/C/T	H	E, A	GAG, GCG	0.06	0.4	15	0.108	0.107	0.555
14	rs1022734986	missense variant	dbSNP	gnomAD	A/C/T	H	E, V	GAG, GTG	0.04	0.322	16	0.156	0.153	0.752
15	rs1459177933	missense variant	dbSNP	gnomAD	T/A	W	Y, N	TAC, AAC	0	0.975	20	0.192	0.401	0.923
16	rs966695918	synonymous variant	dbSNP	gnomAD ExAC AD	T/C	Y	H	CAT, CAC	-	-	-	-	-	-
17	COSV99503172	coding sequence variant	COSMIC		COSMIC_MUTATION	-	-	-	-	-	-	-	-	-
17	COSV105020612	coding sequence variant	COSMIC		COSMIC_MUTATION	-	-	-	-	-	-	-	-	-
19	rs1179317746	synonymous variant	dbSNP	gnomAD	A/G	R	E	GAA, ...	-	-	-	-	-	-

Identification de la séquence nucléotidique du SNP :

Pour chercher la position du SNP dans la séquence nucléotidique il faut :

Aller vers la banque de données généralistes NCBI (<https://www.ncbi.nlm.nih.gov>)

The image shows the homepage of the National Center for Biotechnology Information (NCBI). At the top, there is a navigation bar with the NIH logo and the text "National Library of Medicine National Center for Biotechnology Information". A search bar is located in the top right corner with a "Search" button. Below the navigation bar, there is a "Welcome to NCBI" section with a brief description of the center's mission. To the left, there is a vertical menu with various categories such as "NCBI Home", "Resource List (A-Z)", "All Resources", "Chemicals & Bioassays", "Data & Software", "DNA & RNA", "Domains & Structures", "Genes & Expression", "Genetics & Medicine", "Genomes & Maps", "Homology", "Literature", "Proteins", "Sequence Analysis", "Taxonomy", and "Training & Tutorials". In the center, there are six main service tiles: "Submit" (Deposit data or manuscripts into NCBI databases), "Download" (Transfer NCBI data to your computer), "Learn" (Find help documents, attend a class or watch a tutorial), "Develop" (Use NCBI APIs and code libraries to build applications), "Analyze" (Identify an NCBI tool for your data analysis task), and "Research" (Explore NCBI research and collaborative projects). To the right, there is a "Popular Resources" section listing PubMed, Bookshelf, PubMed Central, BLAST, Nucleotide, Genome, SNP, Gene, Protein, and PubChem. At the bottom right, there is a "NCBI News & Blog" section with a link to "Search, Download, and Visualize Human RNA-Seq Gene Expression Data in NCBI's Gene Expression Omnibus (GEO)". The date "19 Apr 2023" is visible in the bottom right corner.

Chercher un SNP (rs256324)

Chercher la position du SNP dans séquence nucléotidique



NIH National Library of Medicine
National Center for Biotechnology Information

Log in

dbSNP [Help](#)

[Create alert](#) [Advanced](#)

Display Settings:

rs35099072 [*Homo sapiens*]

1.

Variant type: SNV
Alleles: G>A [\[Show Flanks\]](#)
Chromosome: 8:39918134 (GRCh38)
8:39775653 (GRCh37)
Canonical SPDI: NC_000008.11:39918133:G:A
Gene: IDO1 ([Varview](#))
Functional Consequence: coding_sequence_variant,missense_variant
Validated: by frequency,by alfa,by cluster
MAF: A=0.000385/81 ([ALFA](#))
A=0./0 (ALSPAC)
A=0.00027/1 (TWINSUK)
[...more](#)

HGVS: NC_000008.11:g.39918134G>A, NC_000008.10:g.39775653G>A, NG_028155.1:g.9326G>A, NM_002164.6:c.230G>A, NM_002164.5:c.230G>A, NP_002155.1:p.Arg77His

[PubMed](#)

Search details

rs35099072[All Fields]

[See more...](#)

Recent activity

[Turn Off](#) [Clear](#)

rs35099072 (1) SNP

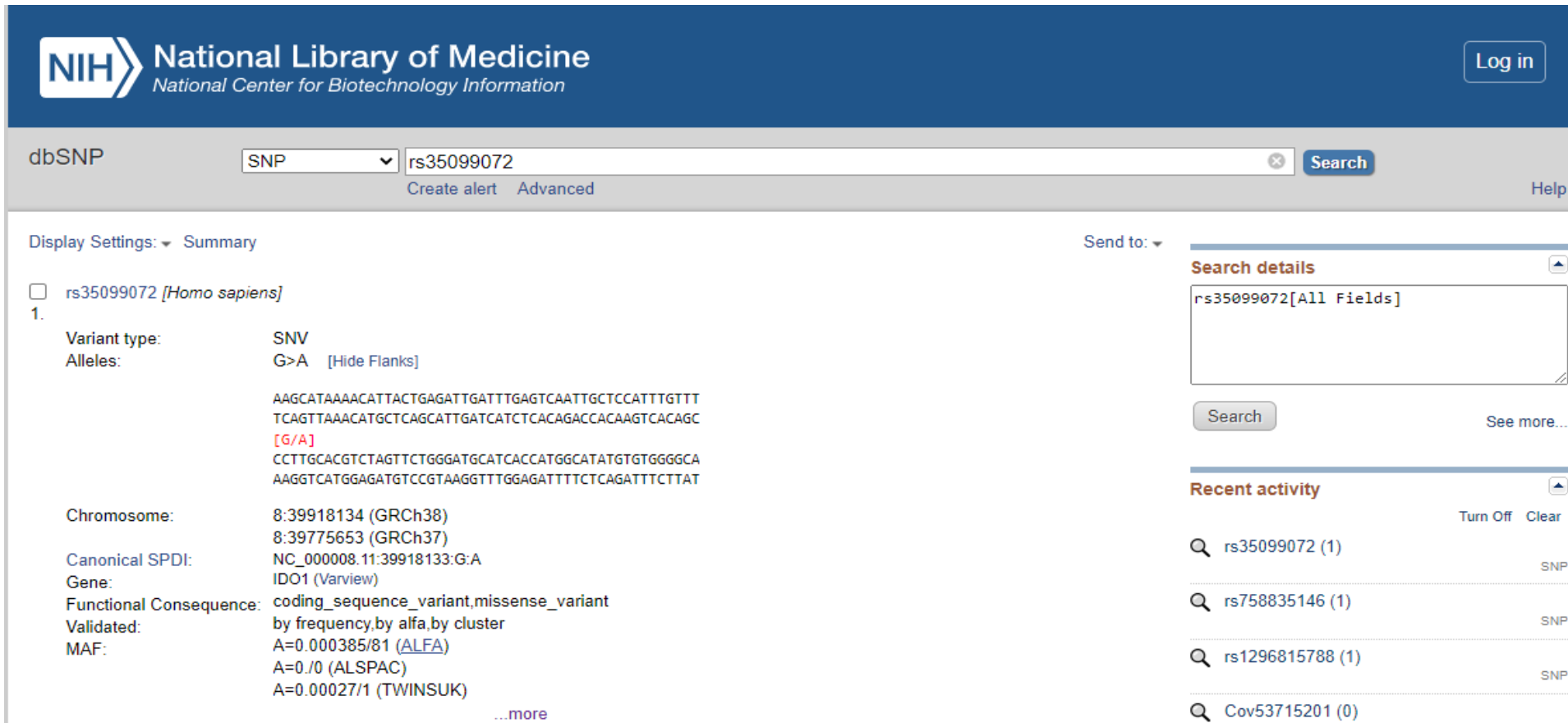
rs758835146 (1) SNP

rs1296815788 (1) SNP

Cov53715201 (0) SNP

Chercher la position du SNP dans séquence nucléotidique

cliquer sur : Hide FLANKS (afficher les | flanks)



NIH National Library of Medicine
National Center for Biotechnology Information

Log in

dbSNP SNP rs35099072 Search

Create alert Advanced Help

Display Settings: Summary

Send to:

rs35099072 [*Homo sapiens*]

1.

Variant type: SNV
Alleles: G>A [Hide Flanks]

AAGCATAAAACATTACTGAGATTGATTTGAGTCAATTGCTCCATTGTTT
TCAGTTAAACATGCTCAGCATTGATCATCTCACAGACCACAAGTCACAGC
[G/A]
CCTTGACAGTCTAGTTCTGGGATGCATCACCATGGCATATGTGTGGGGCA
AAGGTCATGGAGATGTCGGTAAGGTTTGGAGATTTTCTCAGATTTCTTAT

Chromosome: 8:39918134 (GRCh38)
8:39775653 (GRCh37)

Canonical SPDI: NC_000008.11:39918133:G:A
Gene: IDO1 (Varview)

Functional Consequence: coding_sequence_variant,missense_variant

Validated: by frequency,by alfa,by cluster

MAF: A=0.000385/81 (ALFA)
A=0./0 (ALSPAC)
A=0.00027/1 (TWINSUK)

...more

Search details

rs35099072[All Fields]

Search See more...

Recent activity

Turn Off Clear

rs35099072 (1) SNP

rs758835146 (1) SNP

rs1296815788 (1) SNP

Cov53715201 (0)

Comment choisir des SNP ?

Une recherche bibliographique :

- Les polymorphismes répétés dans la maladie
- Les polymorphismes dans les populations

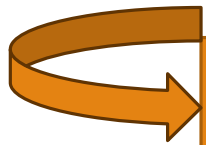
Vérifier la fréquence allélique (MAF): entre (0,02-0,05):

- inférieure à 0,02: rare
- supérieure à 0,05 :très abondante

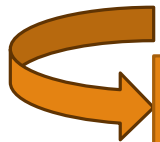
SNP exoniques

SNP faux sens

++++ SNP: sur le même chromosome



L'effet des polymorphismes sur la fonction de la protéine



Etude haplotype, déséquilibre de liaison

Merci

