DATE DE MISE A JOUR: 25mars2024

GLEAVES v3.4.39

Nom de colonne GLEAVES (nom par défaut)	Description	panneau GLEAVES	dossier pris en exemple	chr:position pris en exemple	exemple	mise en forme GLEAVES ? (data)	annotation spécifique du variant ou du/des gènes ?	quel outil prépare cette annotation du VCF	source : annotation du VCF ou autre fichier géré par GLEAVES	colonne ou champ VCF	fichier d'annotation utilisé par l'application	fichier d'annotation source (web)	fichier source modifié par SeqOIA	notes
Sample ID	Sample ID	SNV	1695E	NA	● 1690 1691 1692	oui	variant	HaplotypeCaller	VCF	nom de tous les échantillons décrits après le tag "FORMAT" dans la ligne de titre				chaque ligne du panneau "liste de variants" décrit les génotypes lus pour tous les individus (du dossier). La colonne "Sample ID" permet de lire les identifiants des individus intégrés dans l'analyse du dossier. un symbole permet de décrire le genre des individus (rond: femme carré: homme); si le symbole est rempli en noir, l'individu est atteint par une maladie. les identifiants des individus sont reconnus par l'utilitaire d'appel de variants, et disponible dans le fichier VCF.
Chromosome	Chromosome	SNV	1695E	3:43605720	3	non	variant	HaplotypeCaller	VCF	colonne CHROM				
Position	Position (GRCh38)	SNV	1695E	3:43605720	43605720	non	variant	HaplotypeCaller	VCF	colonne POS				
Chr:Pos	Chromosome:Position (GRCh38)	SNV	1695E	3:43605720	3:43605720	non	variant	HaplotypeCaller	VCF					
HGVS.c	Variant using HGVS notation (DNA level)	SNV	1695E	3:43605720	c.132dupA	non	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	10ème élément de la liste décrite par le champ ANN de SNPEff
HGVS.p	If variant is coding, this field describes the variant using HGVS notation (Protein level).	SNV	1695E	3:43605720	p.Asp45fs	non	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	11ème élément de la liste décrite par le champ ANN de SNPEff
Position/hgvs	Chromosome:Position (GRCh38)/HGVS.c et hgvs.p	SNV	1695E	1:1623412	1:1623412 hgvs c.: c.2T>C Hgvs p.: p.Met1?	oui	variant	non applicable	VCF	non applicable (mise en forme seule	non applicable (mise en forme seulement)	non applicable (mise en forme seulement)	non applicable (mise en forme seule	concaténation des champs « Chromosome », « Position », « HGVS.c » et « HGVS.p » + ajout d'hyperliens vers les pages web d'annotations du variant de MobiDetails, Varsome, UCSC_SNP, UCSC_CNV, gnomAD r4, deCAF, SpliceAI, variant validator, Ensembl, VEP, All of Us, et TOPMed
Gene Name	Common gene name (HGNC); the closest gene is used when the variant is intergenic	SNV	1695E	1:1623412	MIB2	oui	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	4ème élément de la liste décrite par le champ ANN de SNPEff + ajout d'hyperliens vers les pages web d'annotations du gène: HGNC, Genecards, Ensembl, gnomAD, Decipher, Gtex, MobiDetails, Google, PanelApp
QUAL	Phred-scaled quality score for the assertion made in ALT. i.e.−10log10prob(call in ALT iswrong). If ALT is '.' (no variant) then this is−10log10prob(variant), and if ALT is not '.' this is−10log10prob(no variant). High QUAL scores indicate high confidence calls.	SNV	1695E	1:1623412	6882.73	non	variant	HaplotypeCaller	VCF	colonne QUAL	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/			
Genotype	Genotype	SNV	1695E	1:1623412	C/C (lu au survol)	oui	variant	HaplotypeCaller	VCF	colonne REF colonne ALT ##FORMAT=<ID=GT>				dans GLEAVES, 1 ligne est spécifique d'1 allele ALT (et 1 seul) - dans le cas de variants multi-allélique, les autres variants ALT sont décrits sur d'autres lignes.
Recurrence	Allele frequencies in gLeaves from cohort (MR or Cancer, AF") and preindication ("Preind AF"), using SeqOIA genome (WGS-C) data. Nhomalt is the corresponding number of homozygote individuals carrying the variant.	SNV	1695E	1:1623412	cohort: [AF: 0.17(119/708) Nhomalt: 0] preind: [AF: 0.95(84/88) Nhomalt: 40]	oui	variant	GLEAVES	Worker Recurrence de GLEAVES	non applicable	non applicable	non applicable	oui	extrait de la BD interne de récurrence de SeqOIA
SV	SV types intersecting this SNV	SNV	1695E	3:43605720	DUP (1)	oui	variant	GLEAVES	Worker Import de GLEAVES	non applicable	NA			annotation déduite par GLEAVES, avec un hyperlien sur la valeur (hyperlien vers l'onglet « SV », la liste de SV affichées dans le panneau SV est réduite aux SV décrits par la valeur)
Genotype Quality	(GQ) Genotype Quality	SNV	1695E	1:1623412	99	non	variant	HaplotypeCaller	VCF	##FORMAT=<ID=GQ				
Allele count	(AC) Allele count in genotypes, for each ALT allele	SNV	1695E	1:1623412	6	non	variant	HaplotypeCaller	VCF	##INFO=<ID=AC				
Total number of alleles	(AN) Total number of alleles in called genotypes	SNV	1695E	1:1623412	6	non	variant	HaplotypeCaller	VCF	##INFO=<ID=AN				
Allele Frequency	(AF) Allele Frequency, for each ALT allele	SNV	1695E	1:1623412	1	non	variant	HaplotypeCaller	VCF	##INFO=<ID=AF				
DP (Read depth)	(DP) Approximate read depth (reads with MQ=255 or with bad mates are filtered)	SNV	1695E	1:1623412	72	oui	variant	HaplotypeCaller	VCF	##INFO=<ID=DP				
OMIM id	Online Mendelian Inheritance in Man IDentifiant	SNV	1695E	1:1623412	611141	oui	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"omim_id"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	OMIM, genemap2.txt /path du fichier source	oui	
Annotation	(a.k.a. effect). Annotated using Sequence Ontology terms	SNV	1695E	1:1623412	start_lost	oui	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	2ème élément de la liste décrite par le champ ANN de SNPEff
Annotation Impact	A simple estimation of putative impact / deleteriousness : {HIGH, MODERATE, LOW, MODIFIER}	SNV	1695E	1:1623412	HIGH	non	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	3ème élément de la liste décrite par le champ ANN de SNPEff
Inheritance mode	OMIM inheritance mode	SNV	1695E	3:43605720	Autosomal recessive	oui	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"inheritances"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	OMIM, genemap2.txt /path du fichier source sur le web ?		
Diseases	OMIM disease name Gene symbol followed by a list of OMIM disease names; otherwise, (if not avalable), the character "-" is used (example, the MIB2 gene will display "MIB2 -")	SNV	1695E	3:43605720	ANO10 Spinocerebellar ataxia, autosomal recessive 10 (613728),	oui	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"diseases"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	OMIM, genemap2.txt /path du fichier source sur le web ?		
HPOs	List of HPO terms	SNV	1695E	3:43605720	Hyperreflexia (0001347) Hyperreflexia in upper limbs (0007350)	oui	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"hpos"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	http://purl.obolibrary.org/obo/hp/hpoa/genes_to_phenotype.txt		un clic sur le nom du gène affiche la liste des termes HPO
PanelApp	Genomics England PanelApp is a publicly-available knowledgebase. It provides an opportunity for the standardisation of gene panels, and a consensus on which genes have sufficient evidence for disease association. • publication: The National Genomics Research and Healthcare Knowledgebase v5, Genomics England. doi:10.6084/m9.figshare.4530893.v5. 2019 • annotations source: PanelApp API https://panelapp.genomicsengland.co.uk/api/v1 /?format=json • annotations version: v1 (2022-04-26)	SNV	1695E	3:43605720	ANO10 Ataxia and cerebellar anomalies - narrow panel ANO10 Hereditary ataxia ANO10 Hereditary ataxia with onset in adulthood ANO10 Likely inborn error of metabolism - targeted testing not possible ANO10 Mitochondrial disorders ANO10 Possible mitochondrial disorder - nuclear genes ANO10 Severe Paediatric Disorders ANO10 Undiagnosed metabolic disorders ANO10 Childhood onset dystonia, chorea or related movement disorder ANO10 Adult onset neurodegenerative disorder ANO10 Intellectual disability - microarray and sequencing	oui	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"panelapp"."name" "panelapp"."inheritance" "panelapp"."score"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl		oui	le survol sur le nom du gene indique le Mode Of Inheritance (MOI); le code couleur utilisé suit la convention du niveau de confiance en l'information et établi sur la valeur de "panelapp"."score"
Feature type	Which type of feature is in the 'Feature ID' field (e.g. transcript, motif, miRNA, etc.)	SNV	1695E	1:1623412	transcript	non	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	6ème élément de la liste décrite par le champ ANN de SNPEff
Gene ID	ENSEMBL Gene ID	SNV	1695E	1:1623412	ENSG00000197530	oui	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	5ème élément de la liste décrite par le champ ANN de SNPEff
GnomAD pLI score	(GnomAD Gene Constraint). pLI is the probability of belonging in the haploinsufficient class. The pLI score is intended to measure how constrained or intolerant a gene is to lof (loss-of-function). A pLI score closer to 1 indicates that the gene cannot tolerate protein truncating variation (nonsense, splice acceptor and splice donor variation). The gnomAD team recommends transcripts with a pLI >= 0.9 for the set of transcripts extremely intolerant to truncating variants. pLI is based on the idea that transcripts can be classified into three categories: null: heterozygous or homozygous protein truncating variation is completely tolerated recessive: heterozygous variants are tolerated but homozygous variants are not haploinsufficient: heterozygous variants are not tolerated publication: Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B. and Tukiainen, T., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536(7616), pp.285-291. Source Annotations: gnomAD, 'pLoF Metrics by Gene TSV' version: gnomAD v2.1.1 data set	SNV	1695E	1:1623412	MIB2:2.8024e-7	oui	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	non applicable	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl			
GnomAD Obs/Exp lof	GnomAD recommends now using the observed / expected score (Obs/Exp lof, described also as (oe) metric) instead of the probability of pLI score (Loss-of-function Intolerant score, developed with ExAC). It is very important to note that the scale of oe is very different from that of pLI; in particular low oe values are indicative of strong intolerance. In addition, while pLI incorporated the uncertainty around low counts (i.e a gene with low expected count could not have a high pLI), oe does not. Therefore, the oe metric comes with a 90% CI. It is important to consider the confidence interval when using oe. Since pLI > 0.9 is widely used in research and clinical interpretation of Mendelian cases, we suggest using the upper bound of the oe confidence interval < 0.35 if a hard threshold is needed. (the upper bound of the oe confidence interval is named LOEUF Score, and stands for the Loss Of function observed/Expected Upper bound Fraction.). for more details: https://macarthurlab.org/2018/10 /17/gnomad-v2-1/ publication: Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B. and Tukiainen, T., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature, 536(7616), pp.285-291. Source Annotations: gnomAD, 'pLoF Metrics by Gene TSV' version: gnomAD v2.1.1 data set	SNV	1695E	1:1623412	MIB2:0.42234 [0.294 – 0.62]	oui	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	non applicable	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl			
Feature ID	Depending on the annotation, this may be: Transcript ID (preferably using version number), Motif ID, miRNA, ChipSeq peak, Histone mark, etc. Note: Some features may not have ID (e.g. histone marks from custom Chip-Seq experiments may not have a unique ID)	SNV	1695E	1:1623412	ENST00000379458.8	oui	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	7ème élément de la liste décrite par le champ ANN de SNPEff
Transcript biotype	Whenever possible, a description of ENSEMBL biotypes (https://m.ensembl.org/info/genome/genebuild /biotypes.html)	SNV	1695E	1:1623412	protein_coding	non	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	8ème élément de la liste décrite par le champ ANN de SNPEff
LOF (#Transcript)	Number of transcripts in gene affected by the predicted loss of function (LOF)	SNV	1695E	1:1623412	17	non	gene							
LOF (%Transcript)	Percent of transcripts in gene affected by the predicted LoF loss of function (LOF)	SNV	1695E	1:1623412	12%	oui	gene							
NMD (#Transcript)	Number of transcripts in gene affected by the predicted NMD	SNV	1695E	2:74528458	2	non	gene							
NMD (%Transcript)	Percent of transcripts in gene affected by the predicted NMD	SNV	1695E	2:74528458	100%	oui	gene							
clinvar ALLELEID	The ClinVar Allele ID	SNV	1695E	3:43605720	171737	oui	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/GRCh38/index/clinvar/clinvar_20240107			
clinvar CLNSIG	Clinical significance for this single variant	SNV	1695E	3:43605720	Pathogenic/Likely_pathogenic	non	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/GRCh38/index/clinvar/clinvar_20240107			
clinvar CLNSIGCONF	ClinVar SigConf	SNV	1695E	3:43605720	-	non	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/GRCh38/index/clinvar/clinvar_20240107			
ClinVar review status	ClinVar SigConf ClinVar review status for the Variation ID	SNV	1695E	3:43605720	★★ criteria provided multiple submitters no conflicts	oui	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/GRCh38/index/clinvar/clinvar_20240107			
Mitomap	Mitomap	SNV		27057:52:00	-	oui	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/GRCh38/index/mitomap/mitomap_20240125.csv	https://www.mitomap.org/cgi-bin/disease.cgi		
Cosmic Coding ID	The cosmic mutation id	SNV	1695E	1:1623412	COSM7229018 COSM7229017	non	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	non applicable	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl			
Cosmic Non Coding ID	The cosmic non coding variant id	SNV	1695E	11:60896912	COSN8310244	non	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	non applicable	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl			
Census	Is a cancer gene census from COSMIC	SNV	1695E	1:2221609	false true	non	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	non applicable	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	cosmic, cancer_gene_census.csv / path sur le web ?		annotation "Cancer Gene Census" pour les 2 "Gene ID", décrits dans le même ordre
Tier	Tier classification from COSMIC	SNV	1695E	1:2221609	tier 2	non	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"tier"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	cosmic, cancer_gene_census.csv / path sur le web ?		
Band	Cytoband location	SNV	1695E	1:2221609	36.33	non	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"band"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	cosmic, cancer_gene_census.csv / path sur le web ?		
Role	COSMIC gene role	SNV	1695E	1:2221609	oncogene	non	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"role"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	cosmic, cancer_gene_census.csv / path sur le web ?		
Tumor type	COSMIC somatic type	SNV		1:2221609	melanoma	non	gene	GLEAVES	fichier jsonl (généré par geneset aggregator)	"tumor_type"	/data/annotations/Human/GRCh38/index/geneset/geneset_annotation.v.2.0.3.jsonl	cosmic, cancer_gene_census.csv / path sur le web ?		
Mutect2	Listed as somatic variation by Mutect2	SNV	385CB	1:11858349	0.5	non	variant	VCF	Mutect2					applicable pour des prescriptions en génétique somatique (et pour lesquel l'outil Mutect2 est spécifiquement exécuté)
dbSNP RS	dbSNP ID (i.e. rs number)	SNV	1695E	1:1623412	rs12755088	oui	variant	VCF	snpsift	dbsnp_RS	/data/annotations/Human/GRCh38/DB/custom/dbsnp/20180418/All_20180418_chr_{wildcards.index}.vcf.gz		oui	
GnomAD Genome/Exomes	dbSNP ID (i.e. rs number) Gnomad Genome/Exome	SNV	1695E	1:1623412	Genomes: AC : 123604 nhomalt : 53931 Exomes: AC : 191120 nhomalt : 84133	oui	variant	VCF	snpsift	genomes:non_par,AC0,AS_VQSR,In	/data/annotations/Human/GRCh38/DB/custom/gnomad/3.0/gnomad.genomes.r3.0.sites.chr{wildcards.index}.vcf.gz		**oui**	
										exome: non_par,AC0,InbreedingCoe	/data/annotations/Human/GRCh38/DB/public/gnomad/2.1.1/gnomad.exomes.r2.1.1.sites.liftover_grch38.vcf.bgz		**non**	
1000Genomes	1000Genome Allele Frequency for the AFR, EAS, EUR and SAS populations	SNV	1695E	1:1623412	EAS EUR AFR AMR SAS 0.5516 0.9364 0.7057 0.8674 0.8599	oui	variant			EAS_AF,EUR_AF,AFR_AF,AMR_AF	/data/annotations/Human/GRCh38/DB/custom/1000g/phase3/1000GENOMES-phase_3_chr_{wildcards.index}.vcf.gz		oui	
										"name": "kg_",				
CADD phred score	CADD (Combined Annotation Dependent Depletion) : PMID : 24487276 CADD integrates 63 annotations (including SIFT, PolyPhen, GERP++, PhastCons, PhyloP and dbscSNV) into one metric (raw score) for each possible substitution of the human genome (8.6x109). Phred score (-10log10(raw score rank/total)) indicates percentile including variant of interest. Phred score equal to 20 indicates the variants belongs to the 1% most deleterious substitutions that you can do to the human genome.	SNV	1695E	1:1623412	2.343	non	variant	VCF	snpsift	CADD_phred	/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
CADD raw	CADD raw	SNV	1695E	1:1623412	0.08465	non	variant	VCF	snpsift	CADD_raw	/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
FATHMM prediction	FATHMM (Functional Analysis through Hidden Markov Models) : PMID : 23033316 FATHMM is an algorithms based on protein sequence conservation analysis. D = Damaging T = Tolerated	SNV	1695E	1:1623412	T	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
dbNSFP GERP++_NR	Neutral Rate (NR) estimates the total number of neutral substitution/site captured by the phylogenic tree after multiple sequence alignment removing species gapped at that position.	SNV	1695E	1:1623412	1.5	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
dbNSFP GERP++_RS	Rejected substitutions (RS) score is defined as the number of substitutions expected under neutrality minus the number of substitutions observed at the position. Positive scores represent a substitution deficit (i.e., fewer substitutions than the average neutral site) and thus indicate that a site may be under evolutionary constraint. Negative scores indicate that a site is probably evolving neutrally.	SNV	1695E	1:1623412	-2.47	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
LRT prediction	LRT (Likelihood Ratio Test) : PMID: 19602639 LRT compares genomics data set of 32 vertebrate species. D = Deleterious N = Neutral U = Unknown Prediction is not solely determined by the score.	SNV	1695E	1:1623412	U	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
MetaLR prediction	MetaLR (Logistic Regression) : PMID : 25552646 MetaLR is a logistic regression based metapredictor for nonsynonymous SNVs, which incorporates 10 scores (SIFT, PolyPhen-2 HDIV, PolyPhen-2 HVAR, GERP++, MutationTaster, Mutation Assessor, FATHMM, LRT, SiPhy, PhyloP) and the maximum frequency observed in the 1000 Genomes populations. D = Damaging T = Tolerated	SNV	1695E	1:1623412	T	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
MetaSVM prediction	MetaSVM (Support Vector Machine) : PMID : 25552646 MetaSVM is a support vector machine based metapredictor for nonsynonymous SNVs, which incorporates 10 scores (SIFT, PolyPhen-2 HDIV, PolyPhen-2 HVAR, GERP++, MutationTaster, Mutation Assessor, FATHMM, LRT, SiPhy, PhyloP) and the maximum frequency observed in the 1000 Genomes populations. D = Damaging T = Tolerated	SNV	1695E	1:1623412	T	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
MutationTaster prediction	Mutation Taster : PMID: 20676075 MutationTaster integrates data about evolutionary conservation (amino acid : 10 species ; nucleotide : PhyloP, PhastCons), splice-site changes (NNSplice), loss of protein features and changes that might affect the amount of mRNA. Naive Bayes classifier (> 390 000 known disease mutations HGMD vs > 6,800,000 harmless SNPs and Indel polymorphisms from the 1000 Genomes Project) A = Disease causing automatic : known to be deleterious (dbSNP/TGP/ClinVar/HGMD) D = Disease causing P = Polymorphism automatic : known to be harmless (dbSNP/TGP/ClinVar/HGMD) N = Polymorphism	SNV	1695E	1:1623412	P	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
Polyphen2 HDIV prediction	PolyPhen-2 : PMID: 20354512 PolyPhen-2 predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. Naive Bayes classifier (5564 Mendelian disease variants vs. 7539 divergence from close mammalian homologs of human proteins (>=95% sequence identity)). D = Probably damaging, P = Possibly damaging, B = Benign	SNV	1695E	1:153304947	P	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
Polyphen2 HVAR prediction	PolyPhen-2 : PMID: 20354512 PolyPhen-2 predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. Naive Bayes classifier (22196 human variants associated with some disease (except cancer mutations) or loss of activity/function vs. 21119 common (minor allele frequency >1%) human polymorphism with no reported association with a disease or other effect). D = Probably damaging, P = Possibly damaging, B = Benign	SNV	1695E	1:153304947	P	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
Provean prediction	PROVEAN (PROtein Variation Effect ANalyzer) : PMID: 23056405 PROVEAN predicts the functional impact for all classes of protein sequence variations not only single amino acid substitutions but also insertions, deletions, and multiple substitutions. D = Damaging N = Neutral	SNV	1695E	1:1623412	N	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			

SIFT prediction	SIFT (Sorting Intolerant From Tolerant) : PMID: 12824425 SIFT is based on the premise that protein evolution is correlated with protein function. D = Damaging T = Tolerated	SNV	1695E	1:1623412	T,D	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
Alpha Missense	Calibrated AlphaMissense pathogenicity scores (ranging between 0 and 1), which can be interpreted as the predicted probability of a variant being clinically pathogenic. Classification of the protein_variant into one of three discrete categories: 'likely_benign','likely_pathogenic', or ambiguous'. These are derived using the following thresholds: 'likely_benign' if alphamissense_pathogenicity < 0.34; 'likely_pathogenic' if alphamissense_pathogenicity > 0.564; and 'ambiguous' otherwise.	SNV	375E4	1:3497541	likely_pathogenic (0.7273)	oui	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/GRCh38/index/AlphaMissense/version_20230919			
Revel	REVEL (rare exome variant ensemble learner), an ensemble method for predicting the pathogenicity of missense variants on the basis of individual tools: MutPred, FATHMM, VEST, PolyPhen, SIFT, PROVEAN, MutationAssessor, MutationTaster, LRT, GERP, SiPhy, phyloP, and phastCons. The REVEL score for an individual missense variant can range from 0 to 1, with higher scores reflecting greater likelihood that the variant is disease-causing. Publication: Ioannidis, N. M. et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am. J. Hum. Genet. 99, 877–885 (2016).	SNV	375E4	1:3022425	0.739	oui	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/GRCh38/index/REVEL/version_20230503			
SpliceAI prediction	SpliceAI (Illumina): PMID: 30661751 Delta score of a variant, defined as the maximum of (DS_AG, DS_AL, DS_DG, DS_DL), ranges from 0 to 1 and can be interpreted as the probability of the variant being splice-altering. In the paper, a detailed characterization is provided for 0.2 (high recall), 0.5 (recommended), and 0.8 (high precision) cutoffs. Delta position conveys information about the location where splicing changes relative to the variant position (positive values are downstream of the variant, negative values are upstream).	SNV	1695E	1:1623412	AG: 0.03 AL: 0.06 DG: 0 DL: 0	oui	variant	GLEAVES	Worker Import de GLEAVES	non applicable	/data/annotations/Human/hg38/index/spliceAI/illumina_ensembl/version_illumina+ensembl-20200313	non applicable	oui	
PhastCons	PhastCons (PHylogenetic Analysis with Space/Time models CONServation) : PMID : 16024819 PhastCons conservation score based on the multiple alignments of 100 vertebrate genomes (including human). The larger the score, the more conserved the site. Scores range from 0 to 1.	SNV	1695E	1:153304947	0.977	non	variant	VCF	snpsift	a faire vs	/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
QD (Variant Quality by Depth)	Variant Confidence/Quality by Depth	SNV	1695E	1:1623412	32.31	non	variant	HaplotypeCaller	VCF	##INFO=<ID=QD				
OLD_MULTIALLELIC	Original chr:pos:ref:alt encoding	SNV	1695E	1:15274	1:15274:A/T/G	non	variant	Vt	VCF	##INFO=<ID=OLD_MULTIALLELIC,				
Downsampling	Were any of the samples downsampled?	SNV	-	-	-	non	variant	HaplotypeCaller	VCF	DS				
Fisher's exact test	Phred-scaled p-value using Fisher's exact test to detect strand bias	SNV	1695E	1:153304947	1.079	non	variant	HaplotypeCaller	VCF	##FORMAT=<ID=SB				
Symmetric Odds Ratio	Symmetric Odds Ratio of 2x2 contingency table to detect strand bias	SNV	1695E	1:153304947	0.968	non	variant	HaplotypeCaller	VCF	##INFO=<ID=SOR				
dbSCSNV AdaBoost prediction score	dbscSNV (Single Nucleotide Variants within splicing consensus regions database) : PMID: 25416802 dbscSNV ada-boost score predicts probability for a Single Nucleotide Variants within splicing consensus regions (−3 to +8 at the 5' splice site and −12 to 2 at the 3' splice site) to affect splicing. This tool combines analysis of 8 scores (Position Weight Matrix, MaxEntScan, NNSplice, GeneSplicer, Human Splicing Finder, NetGene2, GENSCAN, SplicePredictor). Ranges 0 to 1. The larger the score the higher probability the scSNV will affect splicing. The suggested cutoff for a binary prediction affecting splicing vs. not affecting splicing is 0.6.	SNV	1695E	1:153304947	0.9988	non	variant	VCF	snpsift	ADA_SCORE	/data/annotations/Human/GRCh38/DB/custom/dbscsnv/1.1/hg38_dbscsnv11_chr_{wildcards.index}.vcf.gz		oui	
										"name": "dbscsnv_",				
MetaSVM score	MetaSVM (Support Vector Machine) : PMID : 25552646 MetaSVM is a support vector machine based metapredictor for nonsynonymous SNVs, which incorporates 10 scores (SIFT, PolyPhen-2 HDIV, PolyPhen-2 HVAR, GERP++, MutationTaster, Mutation Assessor, FATHMM, LRT, SiPhy, PhyloP) and the maximum frequency observed in the 1000 Genomes populations. score > 0.5 : Damaging score < 0.5 : Tolerated	SNV	1695E	1:153304947	-1.0585	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
Polyphen2 HVAR score from dbNSFP	PolyPhen-2 : PMID: 20354512 PolyPhen-2 predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. Naive Bayes classifier (22196 human variants associated with some disease (except cancer mutations) or loss of activity/function vs. 21119 common (minor allele frequency >1%) human polymorphism with no reported association with a disease or other effect). [0.909 ; 1] : Probably damaging, [0.447 ; 0.908] : Possibly damaging, [0 ; 0.446] : Benign	SNV	1695E	1:153304947	0.6	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
Polyphen2 HDIV score	PolyPhen-2 predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations. Naive Bayes classifier (5564 Mendelian disease variants vs. 7539 divergence from close mammalian homologs of human proteins (>=95% sequence identity)). [0.957 ; 1] : Probably damaging, [0.454 ; 0.956] : Possibly damaging, [0 ; 0.452] : Benign	SNV	1695E	1:153304947	0.911	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
MutationTaster score	Mutation Taster : PMID: 20676075 MutationTaster integrates data about evolutionary conservation (amino acid : 10 species ; nucleotide : PhyloP, PhastCons), splice-site changes (NNSplice), loss of protein features and changes that might affect the amount of mRNA. Naive Bayes classifier (> 390 000 known disease mutations HGMD vs > 6,800,000 harmless SNPs and Indel polymorphisms from the 1000 Genomes Project) > 0.5 : Disease causing < 0.5 : Polymorphism	SNV	1695E	1:153304947	0.427564	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
FATHMM score	FATHMM (Functional Analysis through Hidden Markov Models) : PMID : 23033316 FATHMM is an algorithms based on protein sequence conservation analysis. score ≤ -1.5 : Damaging score > -1.5 : Tolerated	SNV	1695E	1:153304947	2.6	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
PROVEAN score	PROVEAN (PROtein Variation Effect ANalyzer) : PMID: 23056405 PROVEAN predicts the functional impact for all classes of protein sequence variations not only single amino acid substitutions but also insertions, deletions, and multiple substitutions. ≤ -2.5 : Damaging > -2.5 : Neutral	SNV	1695E	1:153304947	-1.88	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
MetaLR score	MetaLR (Logistic Regression) : PMID : 25552646 MetaLR is a logistic regression based metapredictor for nonsynonymous SNVs, which incorporates 10 scores (SIFT, PolyPhen-2 HDIV, PolyPhen-2 HVAR, GERP++, MutationTaster, Mutation Assessor, FATHMM, LRT, SiPhy, PhyloP) and the maximum frequency observed in the 1000 Genomes populations. score > 0.5 : Damaging score < 0.5 : Tolerated	SNV	1695E	1:153304947	0.0217	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
LRT score	LRT (Likelihood Ratio Test) : PMID: 19602639 LRT compares genomics data set of 32 vertebrate species.	SNV	1695E	1:153304947	0.000123	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
SIFT score	SIFT (Sorting Intolerant From Tolerant) : PMID: 12824425 SIFT is based on the premise that protein evolution is correlated with protein function. < 0.05 : Damaging > 0.05 : Tolerated	SNV	1695E	1:153304947	0.134	non	variant	VCF	snpsift		/data/annotations/Human/hg38/DB/public/dbnsfp/4.0a/dbNSFP4.0a.txt.gz			
					0.065									
dbSCSNV RandomForest prediction score	dbscSNV (Single Nucleotide Variants within splicing consensus regions database) : PMID: 25416802 dbscSNV random-forest score predicts probability for a Single Nucleotide Variants within splicing consensus regions (−3 to +8 at the 5' splice site and −12 to 2 at the 3' splice site) to affect splicing. This tool combines analysis of 8 scores (Position Weight Matrix, MaxEntScan, NNSplice, GeneSplicer, Human Splicing Finder, NetGene2, GENSCAN, SplicePredictor). Ranges 0 to 1. The larger the score the higher probability the scSNV will affect splicing. The suggested cutoff for a binary prediction affecting splicing vs. not affecting splicing is 0.6.	SNV	1695E	1:153304947	0.324	non	variant	VCF	snpsift	RF_SCORE	/data/annotations/Human/GRCh38/DB/custom/dbscsnv/1.1/hg38_dbscsnv11_chr_{wildcards.index}.vcf.gz			
										"name": "dbscsnv_",				
RMS Mapping Quality	RMS Mapping Quality	SNV	1695E	1:153304947	60	non	variant	HaplotypeCaller	VCF	##INFO=<ID=MQ				
Z-score, MQRankSum	Z-score From Wilcoxon rank sum test of Alt vs. Ref read mapping qualities	SNV	1695E	1:153304947	0	non	variant	HaplotypeCaller	VCF	##INFO=<ID=MQRankSum				
Z-score, ReadPosRankSum	Z-score from Wilcoxon rank sum test of Alt vs. Ref read position bias	SNV	1695E	1:153304947	0.166	non	variant	HaplotypeCaller	VCF	##INFO=<ID=ReadPosRankSum				
Phred-scaled p-value	Phred-scaled p-value for exact test of excess heterozygosity	SNV	1695E	1:153304947	3.0103	non	variant	HaplotypeCaller	VCF	##INFO=<ID=ExcessHet				
MLE AC	Maximum likelihood expectation (MLE) for the allele counts (not necessarily the same as the AC), for each ALT allele, in the same order as listed	SNV	1695E	1:153304947	1	non	variant	HaplotypeCaller	VCF	##INFO=<ID=MLEAC				
MLE AF	Maximum likelihood expectation (MLE) for the allele frequency (not necessarily the same as the AF), for each ALT allele, in the same order as listed	SNV	1695E	1:153304947	0.167	non	variant	HaplotypeCaller	VCF	##INFO=<ID=MLEAF				
AD REF (/Total)	Allelic depths for the REF allele	SNV	1695E	1:1623412	0(/72)	oui	variant	HaplotypeCaller	VCF	##FORMAT=<ID=AD				ici, le nombre de reads supportant l'allele REF est considéré (valeur extraite du champ "AD"). Pour la mise en forme, la valeur de la colonne "DP (Read depth)" est utilisée pour décrire le nombre total de reads alignés à cette position.
VAF REF	Allelic Frequency for the REF Allele	SNV	1695E	1:1623412	0	oui	variant	GLEAVES	GLEAVES					
AD ALT (/Total)	Allelic depths for the ALT allele	SNV	1695E	1:1623412	72(/72)	oui	variant	HaplotypeCaller	VCF	##FORMAT=<ID=AD				ici, le nombre de reads supportant l'allele ALT est considéré (valeur extraite du champ "AD"). Pour la mise en forme, la valeur de la colonne "DP (Read depth)" est utilisée pour décrire le nombre total de reads alignés à cette position.
VAF ALT	Allelic Frequency for the ALT Allele	SNV	1695E	1:1623412	1	oui	variant	GLEAVES	GLEAVES					
Phred-scaled likelihoods	(PL) Normalized, Phred-scaled likelihoods for genotypes 0/0, 0/1, and 1/1 respectively	SNV	1695E	1:1623412	2352,217,0	non	variant	HaplotypeCaller	VCF	##FORMAT=<ID=PL				
Error/Warning/Info	Error, warning or informative message that can affect the annotation accuracy	SNV	1695E	1:25801816	Transcript : multiple stop codons	oui	variant	SNPEff	VCF	##INFO=<ID=ANN>	/data/annotations/Human/GRCh38/index/snpEff/v4.3t/GRCh38.92/	https://pcingola.github.io/SnpEff/download/	non	16ème élément de la liste décrite par le champ ANN de SNPEff
Mes commentaires	Mes commentaires	SNV	1695E	1:13684	P	oui	variant	GLEAVES	GLEAVES					j'ajoute un commentaire avec une classification "Pathogène" sur ce variant
Autres commentaires	Autres commentaires	SNV	1695E	1:13684	P VSI	oui	variant	GLEAVES	GLEAVES					ici, les classifications du variant, également détecté et commentés dans d'autres dossiers, sont décrites. (un clic sur l'icone, en forme d' "oeil" permet d'avoir accès au commentaire de chaque dossier)