

Parent Class	Field	Required/ Optional	Ontology Identifier	Definition	Guidance	Examples	Deprecated ID	Version Tracking		Description/ Guidance	Editor Notes
								Label	ID		
organism field	organism	Required	OBID:0100026	A material entity that is an individual living system, such as animal, plant, bacteria or virus, that is capable of replicating or reproducing, growth and maintenance in the right environment. An organism may be unicellular or made up, like humans, of many billions of cells divided into specialized tissues and organs.		Severe acute respiratory syndrome coronavirus 2		0.0.0	0.0.0		
reference genome accession field	reference accession	Required	OBID:0002885	An identifier that specifies an individual sequence record in a public sequence repository.	Use the versioned accession number here.	NC_045512.2		0.0.0	0.0.0	0.0.0	
	reference database name	Required	OBID:0002883	An identifier of a biological or bioinformatics database.	The name of the reference database may need to be entered manually if not automatically recorded as a parameter within the tool.	RefSeq		0.0.0	0.0.0	0.0.0	
	original mutation description	Required	GENEPIO:0101030	A sequence datum which describes a mutation, written in the user's own notation.		C3563S, 12-bp deletion in E at position 54647, D614del, D to G substitution at amino acid 614		0.0.0	0.0.0		
	nucleotide sequence mutation (symbol)	Optional	GENEPIO:0100402	A sequence datum which describes a nucleotide sequence change as compared to a reference sequence, represented by a specified notation system.	Use HGVS syntax; c. (coding) g.(genomic), or r. (ribosomal) mutations should be included here	NC_045512.2:c.1850_1851insGT		0.0.0	0.0.0	0.0.0	
	nucleotide position	Required	GENEPIO:0101033	A sequence datum which is the 1-based start position of a sequence feature for a given nucleotide sequence.		351		0.0.0	0.0.0		
	amino acid sequence mutation (symbol)	Required	GENEPIO:0100406	A sequence datum which describes an amino acid sequence change as compared to a reference sequence, represented by a specified notation system.	Use HGVS syntax; p. (protein) mutations should be included here; use 3 letter code preferred for amino acids. For multiple amino acids that are correlated with a functional impact as a group, input a comma-separated list of mutations. It does not need to be known whether all mutations in the group are individually correlated with the functional impact.	YP_009724390.1:p.Cys3563Ser, YP_009724390.1:p.Asp614Gly		0.0.0	0.0.0	0.0.0	
	amino acid mutation alias	Optional		For SARS-CoV-2. The alias is an amino acid name indexed from the start of each nsp in Orf1ab, as opposed to from the start of the gene as is the default.	Use HGVS syntax; p. (protein) mutations should be included here; use 3 letter code preferred for amino acids	YP_009724390.1:p.Cys300Ser		0.0.0		0.0.0	
gene name field	gene name	Required	OBID:0002878	The name of a gene, (typically) assigned by a person and/or according to a naming scheme. It may contain white space characters and is typically more intuitive and readable than a gene symbol. It (typically) may be used to identify similar genes in different species and to derive a gene symbol.	The long name for the gene/protein. Gene names or gene product names are acceptable. SARS-CoV-2 gene names are from GENEPIO.	GENEPIO_0100154: Spike gene (SARS-CoV-2)		0.0.0	0.0.0	0.0.0	
	gene symbol	Optional	OBID:0002877	The short name of a gene; a single word that does not contain white space characters. It is typically derived from the gene name.	Acceptable values may represent gene symbols or gene product symbols.	S gene, N gene		0.0.0	0.0.0	0.0.0	
	protein name	Optional	MINT	The name of a protein, (typically) assigned by a person and/or according to a naming scheme. It may contain white space characters and is typically more intuitive and readable than a protein symbol. It (typically) may be used to identify similar proteins in different species and to derive a protein symbol.	Protein name taken from the Protein Ontology.	Spike protein		0.0.0		0.0.0	
	protein symbol	Optional	MINT	The short name of a protein; a single word that does not contain white space characters. It is typically derived from the protein name.		S protein		0.0.0			
	assay	Optional	OBID:0000070	A planned process that has the objective to produce information about a material entity (the evaluant) by examining it.				0.0.0	0.0.0		
	mutation functional effect category	Required	MINT	The functional category of the mutation impact.		dimerization		0.0.0			
	mutation functional effect description	Required	GENEPIO:0101036	A proposition which is a freetext description of the functional effect of a mutation.		This mutation demonstrates that glutathionylation (in vivo oxidative post-translational modification of cysteine) of Mpro/NSP5 residue C300 is responsible for controlling the dimerization of the protein, where the monomer's protease activity is hampered by the collapsing of the active site pocket.		0.0.0	0.0.0		
data authors field	author	Optional	dcterms:creator	The last name of the first author of the paper or other resource containing the evidence for the mutation impact.	The surname of the publication's first author.	Davis		0.0.0	0.0.0	0.0.0	
date field	publication year	Optional	dcterms:date	The publication year of the evidence source.		2021		0.0.0	0.0.0		
publication ID field	URL	Required	dcterms:source	A link to the literature evidence source.	If the evidence is found in supplemental material, still link to the paper itself.	https://www.biorxiv.org/content/10.1101/2021.01.22.427749v1		0.0.0	0.0.0	0.0.0	
publication ID field	DOI	Optional	dbXref	The DOI of the literature evidence source.		doi:10.1101/2021.01.22.427749		0.0.0	0.0.0		
publication ID field	PMID	Optional	dbXref	The PMID of the literature evidence source.		pmid:34755213		0.0.0	0.0.0		
	peer-review status	Optional		Review stage for the literature evidence document.		yes, no		0.0.0			
contributor acknowledgement	curator ORCID	Required	dcterms:contributor	The ORCID of the curator who created the annotation record.		0000-0002-0548-891X		0.0.0	0.0.0		
	mutation functional annotation resource	Optional	dcterms:seeAlso	The resource containing the functional annotation for the mutation.	guidance: URL to a source? (Revisit after emailing Paul for examples, to decide on names vs urls, if url, can use dbxref to website, dcterms:source > dcterms:seeAlso, gather more use cases!)	Pokay, SARS-CoV-2 Harvest Variants		0.0.0	0.0.0		