Genomic Vocabularies

Ontologies incorporated:

- **HGNC (HUGO Gene Nomenclature Committee)**
- **ClinVar**
- **CiVIC (Clinical Interpretation of Variants in Cancer)**
- **NCIt (National Cancer Institute Thesaurus)**
- **CGI (Cancer Genome Interpreter)**
- **CAP (College of American Pathologists Cancer Checklists)**
- **JAX CKB (The Clinical Knowledgebase by The Jackson Laboratory)**

Ontologies provided HGVS expressions which were split on refseq, version of refseq, sequence type and locations.

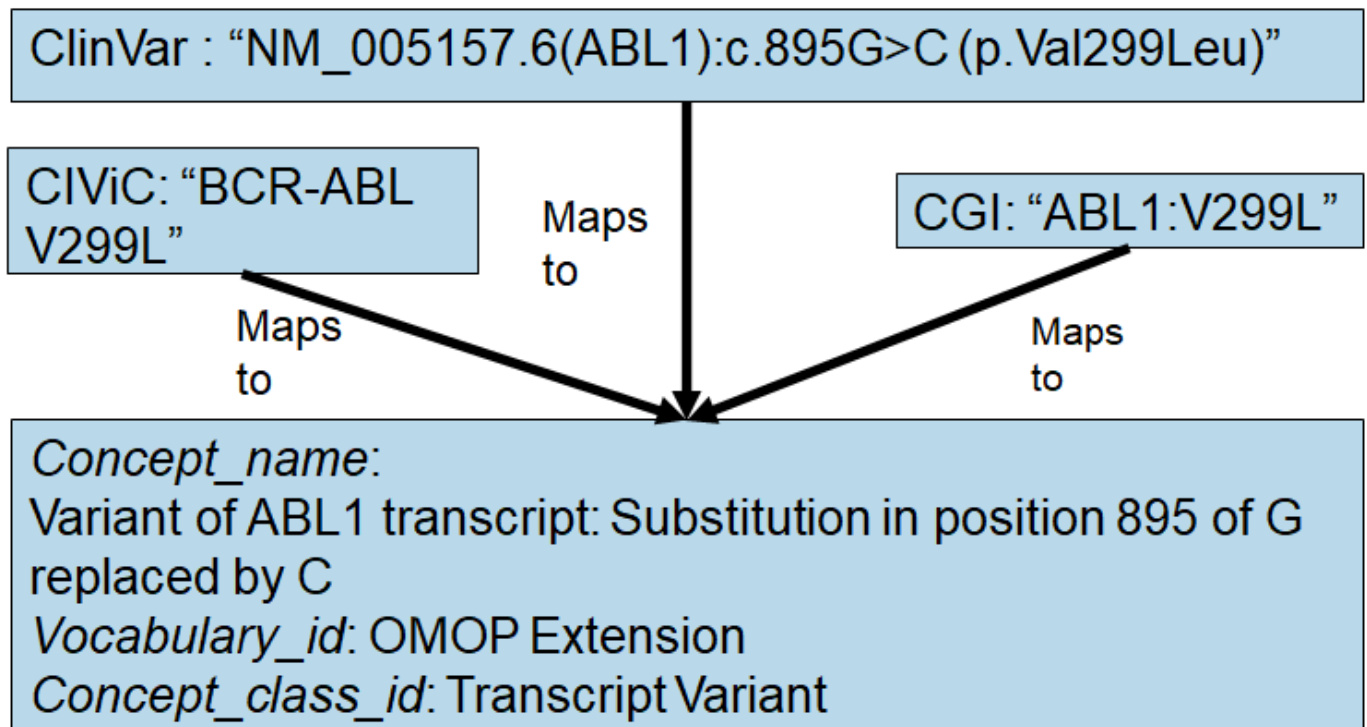| concept_name	vocabulary_id	concept_code	gene	refseq	version	seqtype	variant
ABL1:V299L	CGI	ABL1:V299L	ABL1	ABL1		c	895G>C
BCR-ABL V299L	CIViC	1231	ABL1	ENST00000318560	.5	c	895G>C
BCR-ABL V299L	CIViC	1231	ABL1	NM_005157	.5	c	895G>C
NM_005157.6(ABL1):c.895G>C (p. Val299Leu)	ClinVar	362996	ABL1	NM_005157	.6	c	895G>C
V299L	JAX	3783	ABL1	ABL1		c	895G>C

We developed a simple consolidation approach of single nucleotide polymorphism, insertions, deletions and duplications based on gene symbol, sequence type (g, c, p) and locations.

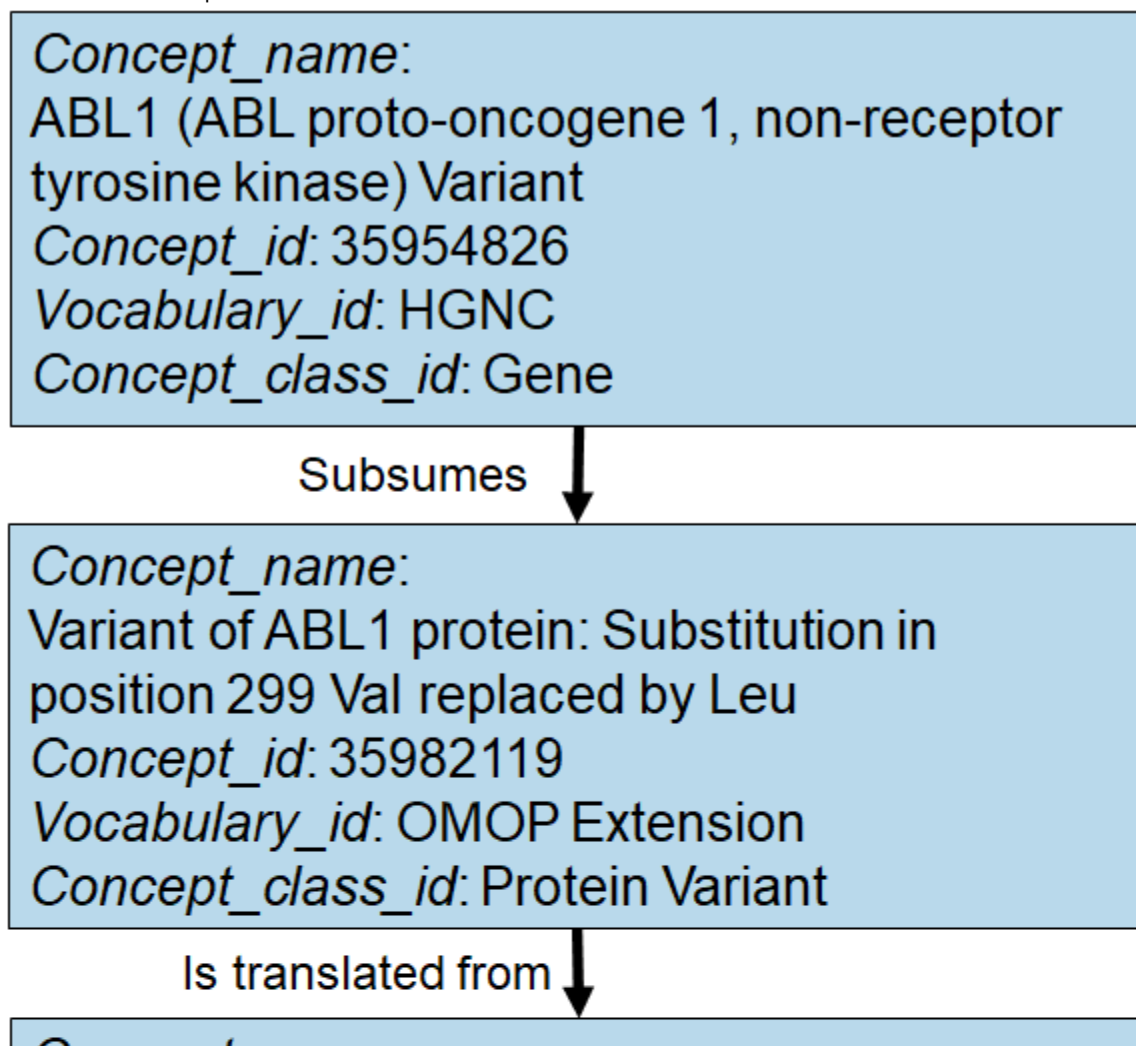| concept_name	vocabulary_id	concept_code	gene	seqtype	variant
ABL1:V299L	CGI	ABL1:V299L	ABL1	c	895G>C
BCR-ABL V299L	CIViC	1231	ABL1	c	895G>C
BCR-ABL V299L	CIViC	1231	ABL1	c	895G>C
NM_005157.6(ABL1):c.895G>C (p.Val299Leu)	ClinVar	362996	ABL1	c	895G>C
V299L	JAX	3783	ABL1	c	895G>C

Genomic, Transcript and Protein variants were created as Standard concept in vocabulary OMOP Extension, where *concept_code* represented as HGVS expression and *concept_name* created from HGVS expression as human readable format.

Concept_id	35982287
Concept_name	Variant of ABL1 transcript: Substitution in position 895 of G replaced by C
Domain_id	Measurement
Vocabulary_id	OMOP Extension
Concept_class_id	Transcript Variant
Standard_concept	S
Concept_code	NM_005157:c.895G>C
Valid_start_date	22/09/2020
Valid_end_date	12/31/2099
Invalid_reason	

All terms from ontologies are mapped to Standard concepts

ClinVar : "NM_005157.6(ABL1):c.895G>C (p.Val299Leu)"

CIViC: "BCR-ABL V299L"

Maps to

CGI: "ABL1:V299L"

Maps to

Maps to

Concept_name:
Variant of ABL1 transcript: Substitution in position 895 of G replaced by C
Vocabulary_id: OMOP Extension
Concept_class_id: Transcript Variant

Standard Variants are placed into ancestor

Concept_name:
ABL1 (ABL proto-oncogene 1, non-receptor tyrosine kinase) Variant
Concept_id: 35954826
Vocabulary_id: HGNC
Concept_class_id: Gene

Subsumes

Concept_name:
Variant of ABL1 protein: Substitution in position 299 Val replaced by Leu
Concept_id: 35982119
Vocabulary_id: OMOP Extension
Concept_class_id: Protein Variant

Is translated from

Concept_name:

Concept_name:
Variant of ABL1 transcript: Substitution in
position 895 of G replaced by C
Concept_id: 35982287
Vocabulary_id: OMOP Extension
Concept_class_id: Transcript Variant

Is transcribed from ↓

Concept_name:
Variant in Chromosome 9 Substitution in
position 133747588 of G replaced by C
Concept_id: 35981735
Vocabulary_id: OMOP Extension
Concept_class_id: Genomic Variant