pubmedKB API

2022-07-25

Li Peng-Hsuan (朋軒)

KB Overview

		pubmedKB-BERN	pubmedKB-PTC	pubmedKB-PTC-FT
Literature	Source	33M PubMed	32M PubMed	2M PubMed Central
	Type	Abstract	Abstract	Full text
	# PMIDs	4,348,135	10,772,202	1,674,383
Entity	# types	4	13	14
	# unique (type, name)s	1,772,821	2,718,110	2,530,623
	# unique (type, ID)s	289,629	598,074	648,891
Relation	# annotators	4	4	4
	# annotations	9,094,498	34,393,039	41,051,562

Entity Types

pubmedKB-BERN	pubmedKB-PTC	pubmedKB-PTC-FT
gene	Gene	Gene
variant	ProteinMutation DNAMutation SNP CopyNumberVariant	ProteinMutation DNAMutation SNP CopyNumberVariant
disease	Disease	Disease
chemical	Chemical	Chemical
-	Species	Species
-	CellLine	CellLine
-	RefSeq	RefSeq
-	GenomicRegion	GenomicRegion
-	Chromosome	Chromosome
-	Genus	-
-	-	DNAAcidChange
-	-	DomainMotif

Entity ID Systems

pubmedKB-BERN
BERN
MESH
CHEBI
OMIM
HGNC
MIM
Ensembl
IMGT/GENE-DB
miRBase

pubmedKB-PTC
MESH
NCBITaxID
tmVar
VariantGroup
CorrespondingGene
HGVS
RS#
CA#
OMIM
CVCL
CorrespondingSpecies
OriginalGene
Chr1~23
ChrX

pubmedKB-PTC-FT
MESH
CVCL
tmVar
VariantGroup
CorrespondingGene
HGVS
RS#
CA#
OMIM
CorrespondingSpecies
OriginalGene
Focus
Right
Left
Chr0~9,01~39,65,88,103
ChrX,Y,XX

Entity ID Search - Web API

NEN Query

Case sensitive (Y/N): N Max length difference: 1 Min similarity: 0.85 Max #name per query: 20 Max #alias per ID: 5

Query: V600E

show

Ready

Name: V600E Similarity: 100%

Type	ID	Frequency of (Type, ID, Name) in all texts	Top 5 aliases by frequency of (Type, ID, Alias) in all texts
ProteinMutation	tmVar:p|SUB|V|600|E	12,046	V600E (12,046) BRAFV600E (4,433) p.V600E (331) BrafV600E (97) p.Val600Glu (41)
ProteinMutation	HGVS:p.V600E	12,046	V600E (12,046) BRAFV600E (4,433) p.V600E (331) BrafV600E (97) p.Val600Glu (41)
ProteinMutation	CorrespondingGene:673	11,801	V600E (11,801) BRAFV600E (4,426) V600 (339) p.V600E (327) V600K (253)
ProteinMutation	RS#:113488022	11,617	V600E (11,617) BRAFV600E (4,433) p.V600E (283) V600 (274) V600E/K (159)
ProteinMutation	CA#:123643	11,617	V600E (11,617) BRAFV600E (4,412) p.V600E (283) BrafV600E (97) p.Val600Glu (38)
ProteinMutation	VariantGroup:0	11,460	V600E (11,460) V617F (8,378) T790M (5,599) BRAFV600E (4,395) C282Y (3,132)
ProteinMutation	VariantGroup:1	441	C282Y (1,209) H63D (1,184) T790M (1,145) Val/Val (812) L858R (796)
ProteinMutation	RS#:121913227	158	V600E (158) V600K (128) V600 (59) p.V600E (44) V600R (43)
ProteinMutation	CorrespondingSpecies:9606	110	T790M (123) V600E (110) BRAFV600E (93) D614G (86) G12C (60)
ProteinMutation	VariantGroup:2	103	Val/Val (366) Met/Met (264) Arg/Arg (243) H63D (227) C282Y (221)
ProteinMutation	VariantGroup:3	25	Arg/Arg (125) K103N (77) Y181C (73) Gln/Gln (71) M184V (70)
ProteinMutation	CorrespondingGene:109880	17	V600E (17)
ProteinMutation	CorrespondingGene:5594	16	K52R (19) E322K (17) serine/threonine (17) S6 (17) V600E (16)
ProteinMutation	RS#:121913378	14	V600E (14) V600M (6) V600L (4) V600 (4) V600R (3)
ProteinMutation	VariantGroup:4	13	K103N (65) Gln/Gln (51) M184V (45) L180M (41) Arg/Arg (40)
ProteinMutation	CorrespondingGene:3845	12	G12D (1,783) G12V (686) G13D (346) p.G12C (106) p.G13D (104)
ProteinMutation	CorrespondingGene:7015	11	V600E (11) p-E1A (5) R132H (4) K626A (4) K27M (3)
ProteinMutation	CorrespondingGene:22882	11	V600E (11) serine/threonine (9) A85K (5) N71R (2) S621A (2)

Relation Annotators

Annotator	Annotation	pubmedKB-BERN	pubmedKB-PTC	pubmedKB-PTC-FT
odds_ratio	OR, CI, p-value	V	V	V
rbert_cre	Relation	V	V	V
spacy_ore	Triplet	V	V	V
openie_ore	Triplet	V	V	V

Relation Annotator - Details

Annotator	Annotation		Explanation
odds_ratio	OR, CI, p-value	OR	Odds ratio extracted from the sentence
		CI	Confidence interval extracted from the sentence
		p-value	p-value statistics extracted from the sentence
rbert_cre	Relation	Relation: Cause-associated	VARIANT is a causal variant of DISEASE; or VARIANT is DISEASE-associated
		Relation: Appositive	VARIANT's association with DISEASE is presupposed
		Relation: In-patient	VARIANT occurs in some DISEASE patients
spacy_ore	Triplet	(subject, predicate, object)	The combined clause is entailed by the sentence
openie_ore	Triplet	(subject, predicate, object)	The combined clause is entailed by the sentence

Relation Annotation - Web API

REL Query

Entity A

Query by: Type & Name Type: ProteinMutation ID: tmVar:p|SUB|V|600|E Name: V600E

Entity B

Query by: Type & ID Type: Disease ID: MESH:D009369 Name: neoplasm

Query

Query by: Un-ordered pair (A, B) Max #rel per annotator: 20 PMID: 22494505 (ignored if query by entity or entity pair)

show

Ready

Head	Tail	Annotator	Annotation	PMID	Sentence
V600E [ProteinMutation] *CA#:123643* *CorrespondingGene:673* *HGVS:p.V600E* *RS#:113488022* *VariantGroup:0* *tmVar:p|SUB|V|600|E*	**tumors** [Disease] *MESH:D009369*	odds_ratio	OR: 1.87, CI: 1.01-3.47, p-value: x	20200438	ABSTRACT paragraph *The prevalence of BRAF V600E was 5.0% in this population, and the mutation was more likely to be found in tumors from cases who were of Ashkenazi Jewish descent [odds ratio (OR), 1.87; 95% confidence interval (95% CI), 1.01-3.47], female (OR, 1.97; P = 1.17-3.31), and older (73.8 years versus 70.3 years; P < 0.001).*
V600E [ProteinMutation] *HGVS:p.V600E* *tmVar:p|SUB|V|600|E*	**cancers** [Disease] *MESH:D009369*	rbert_cre	Cause-associated: 87.5%	35028596	ABSTRACT paragraph *The constitutive expression of wild-type BRAF (BRAFWT) and its mutant forms, especially V600E (BRAFV600E), has been linked to multiple cancers.*
V600E [ProteinMutation] *CA#:123643* *CorrespondingGene:673* *HGVS:p.V600E* *RS#:113488022* *VariantGroup:0* *tmVar:p|SUB|V|600|E*	**cancer** [Disease] *MESH:D009369*	spacy_ore	B-RafV600E, promotes, cancer cell behavior	22751131	ABSTRACT paragraph *The most prevalent mutation, B-RafV600E, promotes cancer cell behavior through mechanisms that are still incompletely defined.*
V600E [ProteinMutation] *CA#:123643* *CorrespondingGene:673* *HGVS:p.V600E* *RS#:113488022* *VariantGroup:0* *tmVar:p|SUB|V|600|E*	**tumors** [Disease] *MESH:D009369*	openie_ore	V600E mutation, was detected in, 25 of 27 tumors	23682579	ABSTRACT paragraph *The BRAF(V600E) mutation was detected in 25 of 27 tumors (92.6%).*

Head	Tail	Annotator	Annotation	PMID	Sentence
V600E [ProteinMutation] *CA#:123643* *CorrespondingGene:673* *HGVS:p.V600E* *RS#:113488022* *VariantGroup:0* *tmVar:p|SUB|V|600|E*	**tumours** [Disease] *MESH:D009369*	odds_ratio	OR: 12.33, CI: 4.21-41.54, p-value: 1.8x10-7	21625944	RESULTS paragraph *We observed a significant association of both the BRAF V600E mutation and MLH1 methylation with CIMP+ tumours (odds ratio (OR) = 12.33, 95% confidence interval (CI) = 4.21-41.54, P = 1.8 x 10-7 and OR = 15.5, 95% CI = 4.54-68.65, P = 2.8 x 10-7, respectively).*
V600E [ProteinMutation] *CA#:123643* *CorrespondingGene:673* *HGVS:p.V600E* *RS#:113488022* *VariantGroup:0* *tmVar:p|SUB|V|600|E*	**tumor** [Disease] *MESH:D009369*	rbert_cre	In-patient: 75.7%	23043499	CASE paragraph *Mutation analysis of the patient's tumor material demonstrated the presence of a BRAFV600E mutation and the patient subsequently initiated treatment with a BRAF inhibitor in September 2011.*
V600E [ProteinMutation] *CA#:123643* *CorrespondingGene:673* *HGVS:p.V600E* *RS#:113488022* *VariantGroup:0* *tmVar:p|SUB|V|600|E*	**tumor** [Disease] *MESH:D009369*	spacy_ore	the BRAFV600E inhibitor, not affect, tumor eradication	22735805	DISCUSS paragraph *Furthermore, using a tumor model in which tumor eradication is dependent on the activation of tumor antigen specific CD4+ and CD8+ cells and their production of both IFN-gamma and IL-2 demonstrated that the BRAFV600E inhibitor did not affect tumor eradication.*
V600E [ProteinMutation] *CA#:123643* *CorrespondingGene:673* *HGVS:p.V600E* *RS#:113488022* *VariantGroup:0* *tmVar:p|SUB|V|600|E*	**cancer** [Disease] *MESH:D009369*	openie_ore	crVAR V600E, is shared by, 17 cancer types	27977206	RESULTS paragraph *The crVAR V600E in BRAF protein (ENSP00000288602), which is a therapeutic target of metastatic melanoma, is shared by 17 cancer types.*

API Access

- Source code (programmable API + demo website)
  - https://github.com/jacobvsdanniel/pubmedkb_web
  - Document: https://github.com/jacobvsdanniel/pubmedkb_web/wiki
- Sample dataset (1/10 of pubmedKB-ptc)
  - https://drive.google.com/file/d/10_UmG_ozWSrvFB9vJ0TfY41WHCzswmhm/view?usp=sharing