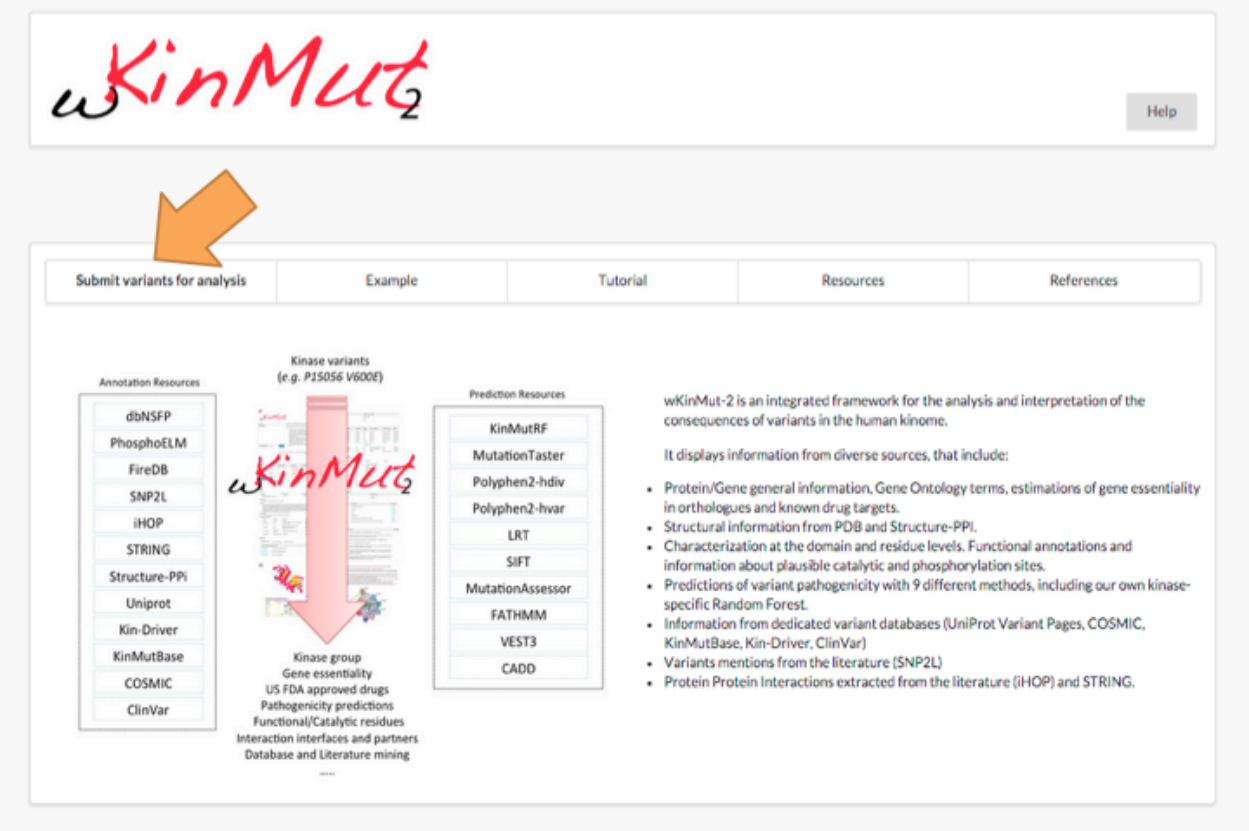


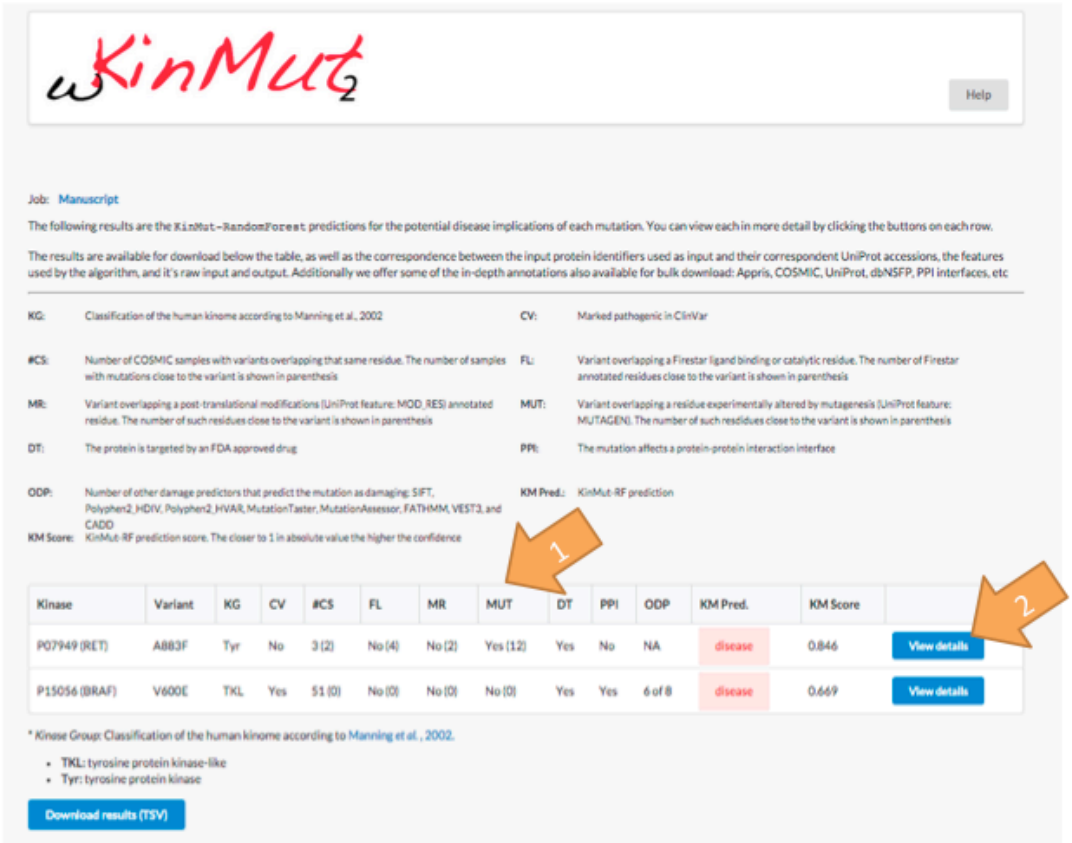
Kinmut2.bioinfo.cnio.es presents a welcome home page that includes link to useful examples, tutorials, a summary of the resources and the references to our KinMut related publications.
To begin the analysis of variants, please, click on “Submit variants for analysis” (an orange arrow will identify clickable areas in this tutorial).

(1) The input to wKinMut-2 are single point missense variants affecting human protein kinases. A variant can be defined in a very simple format using the accession number from UniProt, the wild-type and mutant amino acids and the position in the protein. For example, the well studied change from Valine (V) to Glutamate (E) in position 600 of the B-Raf proto-oncogene (UniProt accession number: P15056) would be encoded as "P15056 V600E". Multiple variants can be analysed at a time using wKinMut-2, for that the input should contain one variant per line.

Please, note that the server displays information for an alternative amino acid at a time. Non-standard amino acids (B,Z) will be decomposed into separate instances of their standard counterparts. For example, "P07949 A883B" would be read internally as "P07949 A883D" and "P07949 A883N". wKinMut-2 focuses on the analysis of missense variants. Synonymous and truncating variants will be excluded from the analysis.

(2) After introducing the variant, user can optionally name the experiment.

(3) Press the "Submit" button to start running the analysis. Depending on the load of the server and the number of variants submitted for analysis, some calculations might take some time to be ready. Please, be patient.

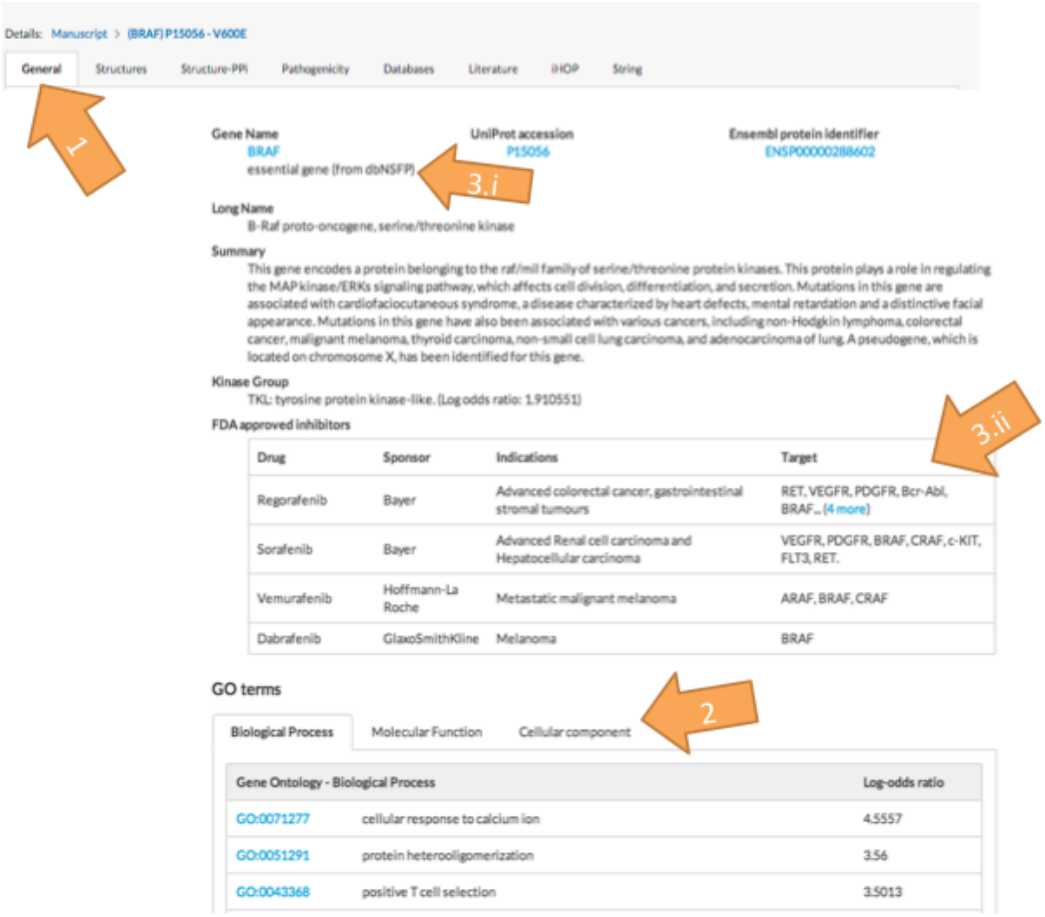


(1) After the calculation for the submitted variants has finished, the user is presented with a table that summarizes these results. The table includes information to start the prioritization of mutations for further analysis including: KG: The classification of the kinase in the context of the human kinome according to Manning's taxonomy; CV: The mutation has been classified as pathogenic in ClinVar; #CS: The number of COSMIC (somatic mutations in cancer) samples with variants in the same residue; FL: The variant sits on a residue that Firestar considers relevant for ligand binding; MR: Whether the residue is annotated as subject of translational modification (MOD_RES) in UniProt; MUT: Whether the residue is annotated as subject to mutagenesis (MUTAGEN) in UniProt, under the assumption that targeted experiments would focus on functionally relevant residues. The number in parenthesis reflects the amount of residues in the proximity that fulfill these criteria; DT: Whether the protein is targeted by a FDA approved drug; PPI: The residue is involved in a protein-protein interface; ODP: wKinMut-2 calculates the consequences of variants according to 8 external classifiers: SIFT, Polyphen2-HDIV, Polyphen2-HVAR, MutationTaster, MutationAssessor, FATHMM, VEST, and CADD. The ODP (Other Disease Predictors) reflects how many of these classify the variant as pathogenic; "KM Pred." and "KM Score": In addition, the variants are evaluated with a random forest (KinMutRF) developed ad-hoc for the study of human protein kinases. Variants are classified disease or neutral and a reliability index (ranging from 0 to 1) will assess the confidence in the prediction. The closer to 1 in absolute value the higher the confidence.

(2) Once the variants of interest have been identified, users can obtain additional information by clicking on "View details" button.

(1) Some users might find themselves interested in having a report of the predictions. These can be downloaded directly from the server as a tab-separated (TSV) file.

(2) Additional files used for the prediction or containing supplemental information can be downloaded as well.



(1) The general tab will describe basic features of the kinase where the variant of interest occurred.

(2) Information present includes the gene name and the description from UniProt, the protein identifier in Ensembl and the classification in kinase groups as defined by Manning and collaborators. In addition, as a proxy to understand the cellular role of the protein, we list GeneOntology annotations grouped by sub-ontology (i.e., Molecular Function, Cellular Compartment and Biological Process).

(3) And we provide information about:

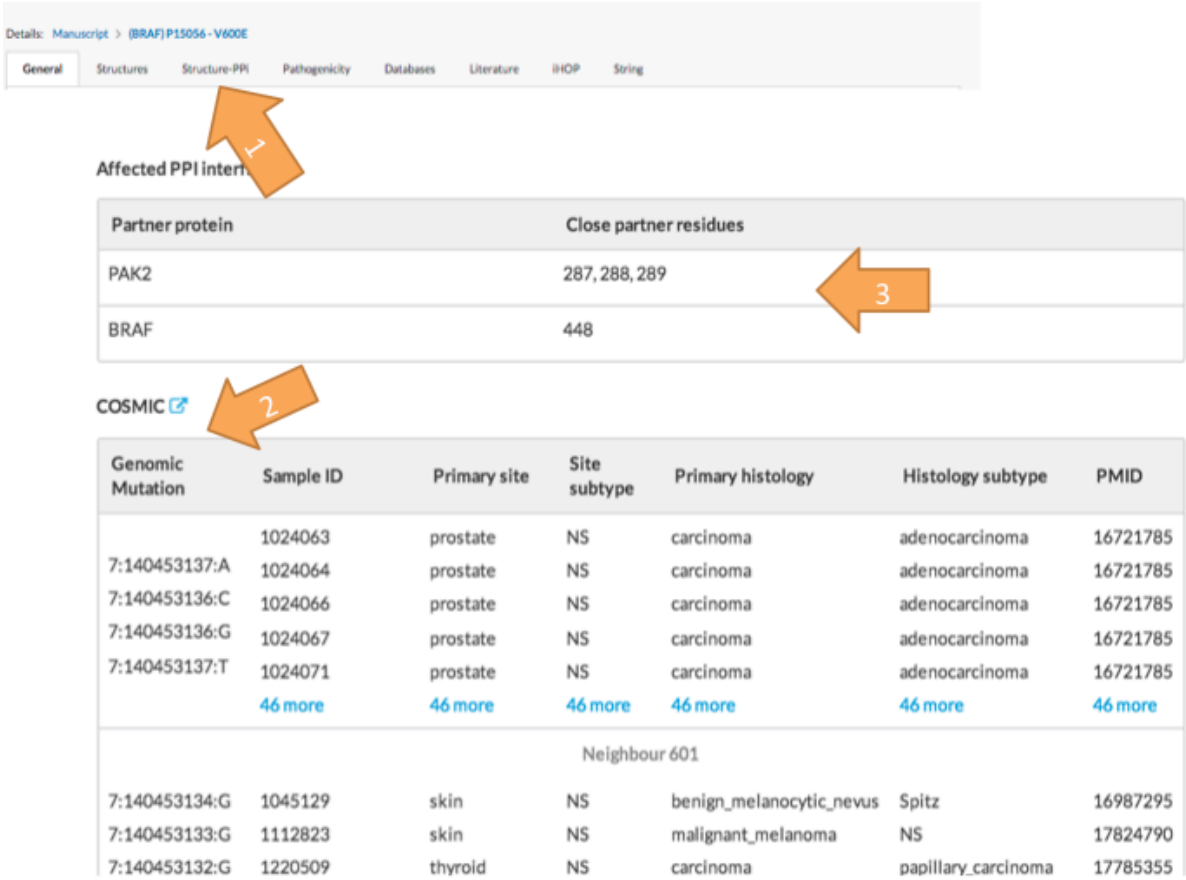
i) Essential or non-essential phenotype-changing of the homologous gene in mouse based on the information collected by dbNSFP [Liu et al. 2013] from the Mouse Genome Informatics database (https://sites.google.com/site/jpopgen/dbNSFP).

ii) US FDA [Jänne et al. 2009] approved protein kinase inhibitors (http://www.brimr.org/PKI/PKIs.htm).

(1) The Structure tab represents the variants with respect to the protein structures of the kinases.

(2) A Jmol representation of the structures and the variants helps understand the functional consequences of the latter.

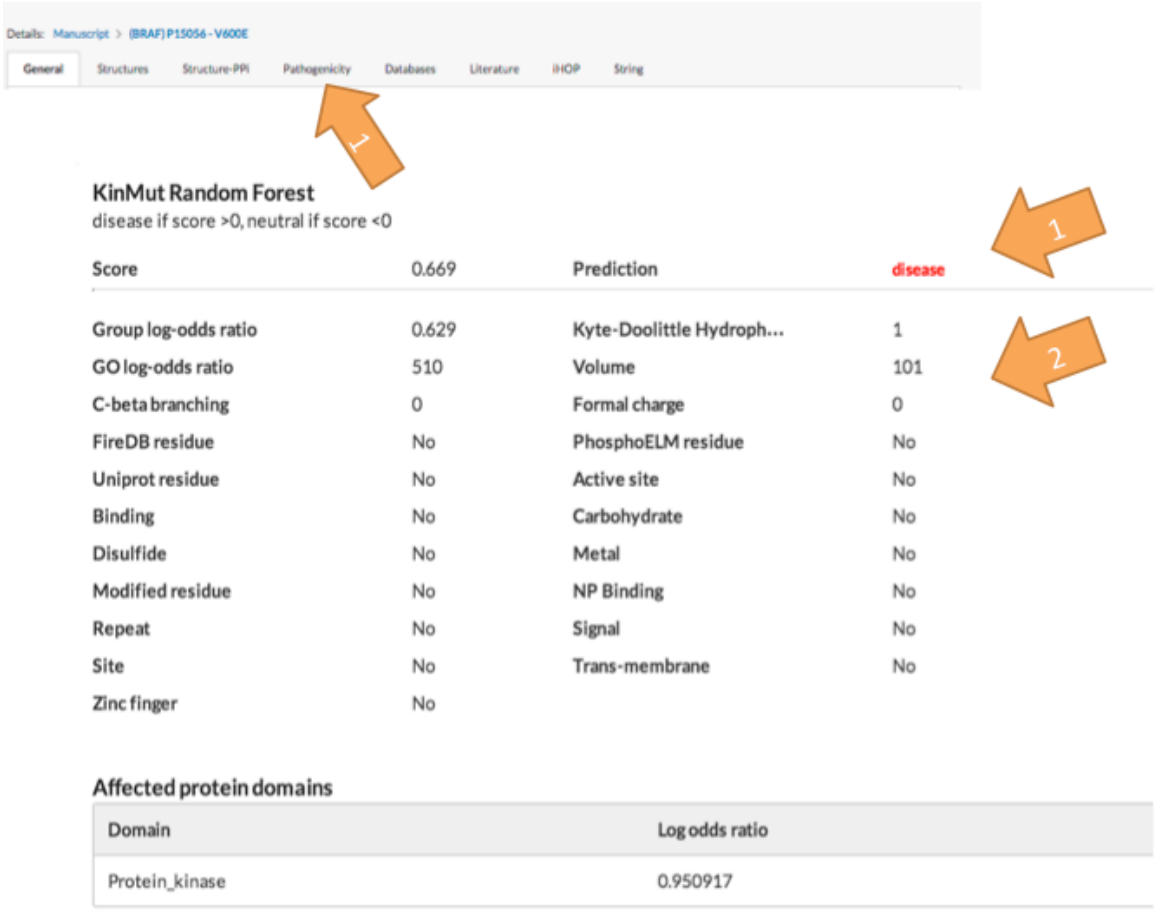
(3) Domain information is also provided to elucidate the potential impact of the variants on the function of the kinases.



(1) Structure-PPi is a system to facilitate the analysis of variation in the context of protein complexes.

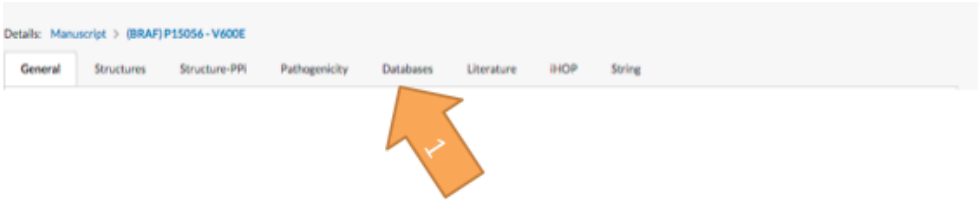
(2) The system combines information from protein structures with functional annotations from a number of relevant databases and reports protein features (e.g., functional domains, known somatic variation in different types of cancer, UniProt annotations from missense variants, ligand binding residues, catalytic sites) that overlap the variant's 'direct matches' or their 'neighbors' in close physical proximity. These are defined by being within 5 angstroms spatial distance or adjacent in the sequence if no PDB covers that area.

(3) When variants affect the interfaces of protein complexes (when the variant is at a distance of less than 8 angstroms from a residue in the partner protein), Structure-PPi also reports the partner proteins, and the residues in those proteins that are in close proximity to the variant.



(1) wKinMut-2 implements an ad-hoc method specific to the protein kinase superfamily. We termed this new methodology KinMutRF. KinMutRF classifies variants as neutral or disease-associated. A score ranging from 0 to 1 provides a measure of the reliability of the prediction.

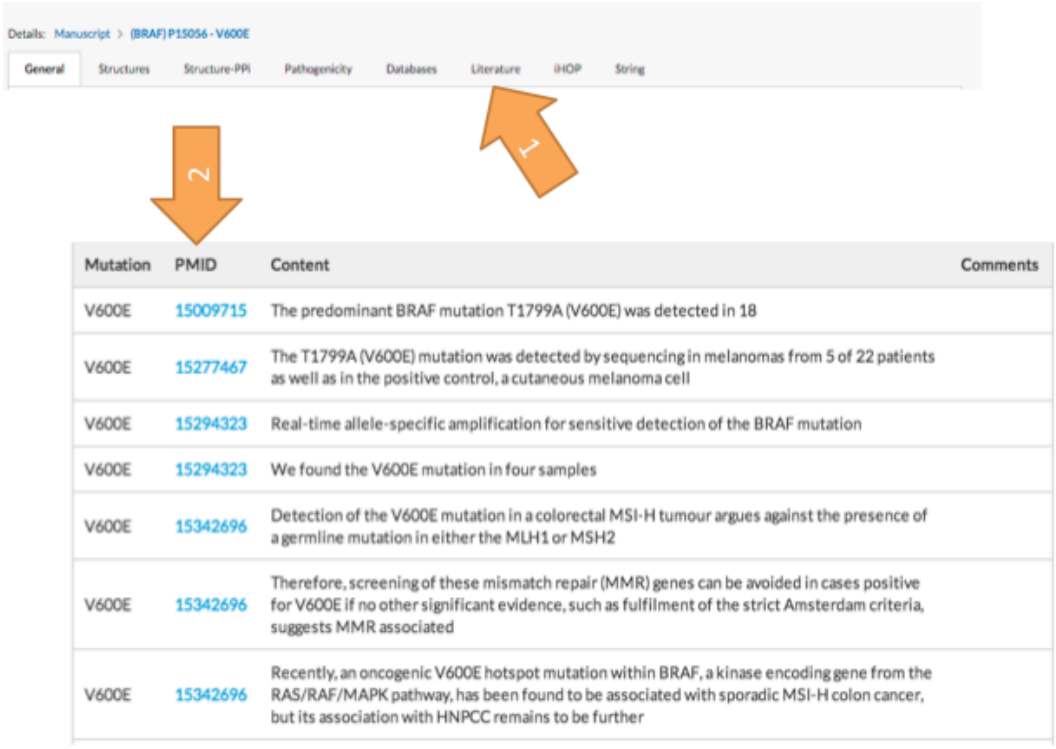
(2) The method relies on a random forest classifier consisting of 26 decision trees that evaluate a number of sequence-derived features that characterize variants affecting human protein kinases at different levels: a) at the gene level, including membership to a Kinbase group and Gene Ontology categories; b) at the domain level, using PFAM domains; and c) at the residue level, involved amino acids types, changes in biochemical properties, functional annotations from UniProt, Phospho.ELM and FireDB. These features are provided to guide the interpretation of the pathogenicity predictions and to help draw hypotheses on the plausible biological mechanisms by which the pathogenicity arose.



Mutation	Description	Database
V600E	somatic mutation in cancer	cosmic
V600E	passenger	greenman
V600E	driver	greenman
V600D	Polymorphism	uniprot
V600E	Disease: Colorectal cancer (CRC) [MIM:114500]	uniprot
BRAF_HUMAN.V600E	Freq. 192% Activating: Colorectal carcinoma, melanoma and many others	kindriver

(1) wKinMut-2 incorporates relevant information from the UniProt Variant Pages, KinMutBase, Kin-Driver, COSMIC and ClinVar. The information is intended to facilitate a digested contextual framework for the interpretation of consequences of the variants. Of particular interest, any experimental evidence relating variants and disease.

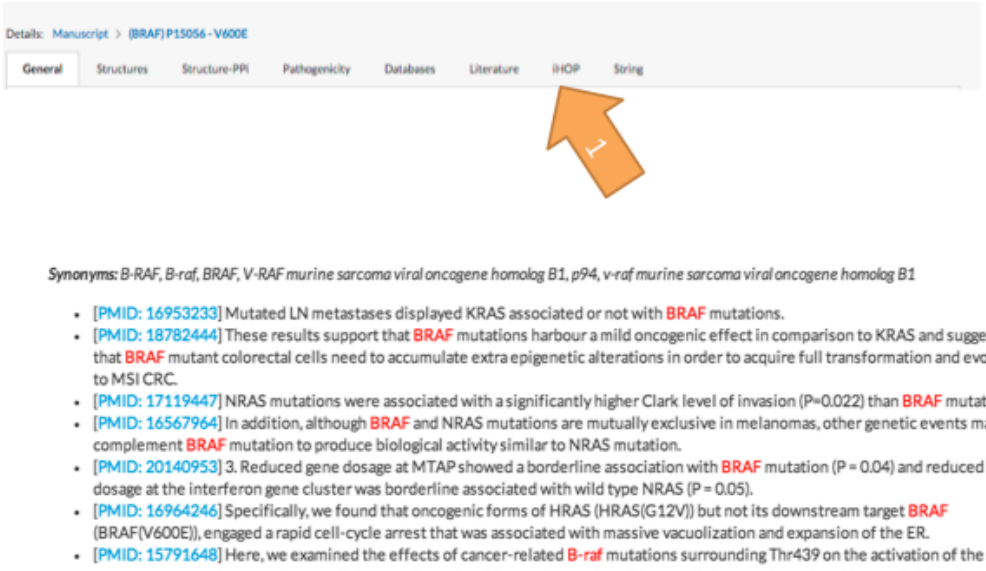
Disclaimer: Users of previous versions of the tool would notice that SAAPdb is not included in the current implementation, as its authors have discontinued it.



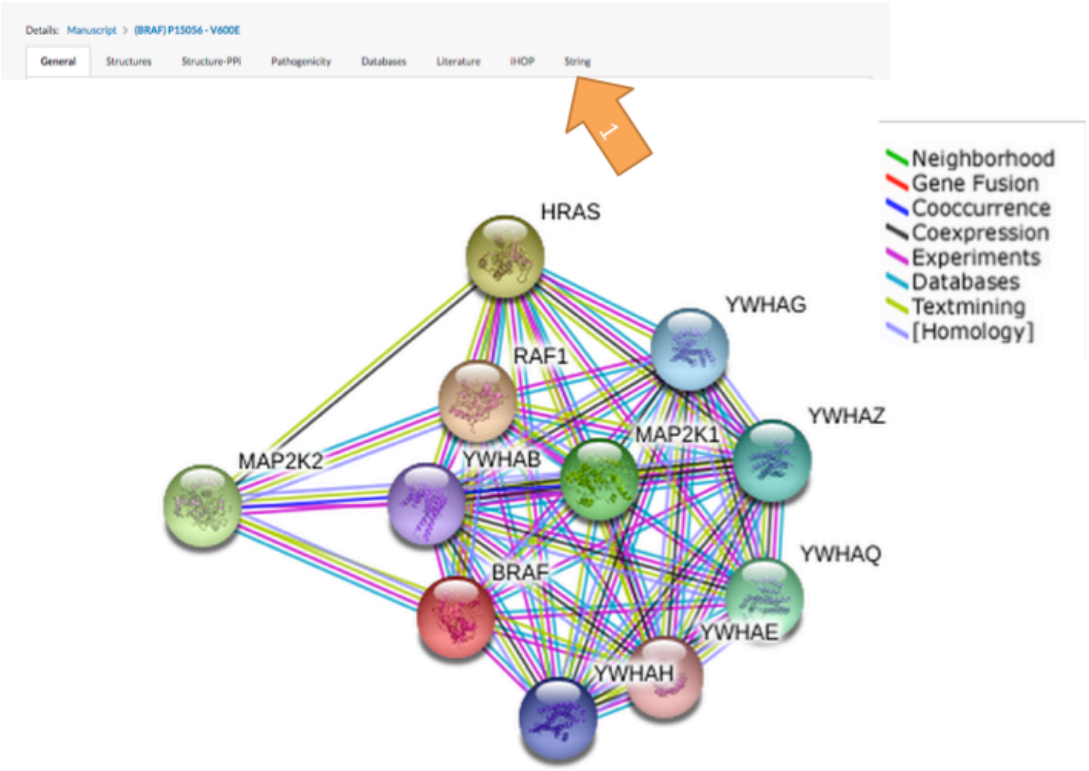
Mutation	PMID	Content	Comments
V600E	15009715	The predominant BRAF mutation T1799A (V600E) was detected in 18	
V600E	15277467	The T1799A (V600E) mutation was detected by sequencing in melanomas from 5 of 22 patients as well as in the positive control, a cutaneous melanoma cell	
V600E	15294323	Real-time allele-specific amplification for sensitive detection of the BRAF mutation	
V600E	15294323	We found the V600E mutation in four samples	
V600E	15342696	Detection of the V600E mutation in a colorectal MSI-H tumour argues against the presence of a germline mutation in either the MLH1 or MSH2	
V600E	15342696	Therefore, screening of these mismatch repair (MMR) genes can be avoided in cases positive for V600E if no other significant evidence, such as fulfilment of the strict Amsterdam criteria, suggests MMR associated	
V600E	15342696	Recently, an oncogenic V600E hotspot mutation within BRAF, a kinase encoding gene from the RAS/RAF/MAPK pathway, has been found to be associated with sporadic MSI-H colon cancer, but its association with HNPCC remains to be further	

(1) In an attempt to complement the information provided by the databases, wKinMut-2 provides information mined directly from the literature (Pubmed abstract and full-texts) with SNP2L. It is often the case that relevant information about the experimental conditions, the patients in the cohort, etc, can be found in the literature although it is missing in the databases due to the particular constrains of its design. The full text articles should provide a deeper understanding of these individual peculiarities.

(2) The system provides links to the original publications (PMIDs) and displays the specific sentences where the variants were mentioned.



(1) Similarly, wKinMut-2 sources literature co-mentions from iHOP. Literature co-mentions constitute a good proxy to interactions. In addition to the links to the original articles (PMIDs), the specific sentences in the literature are displayed. They are intended to provide contextual information that can facilitate the interpretation of the consequences of the pathogenic variation.



(1) STRING is a resource that stores known and predicted protein interactions from different sources including genomic context, high-throughput experiments, coexpression and text-mining of the literature. STRING interactions include direct (physical) and indirect (functional) association.