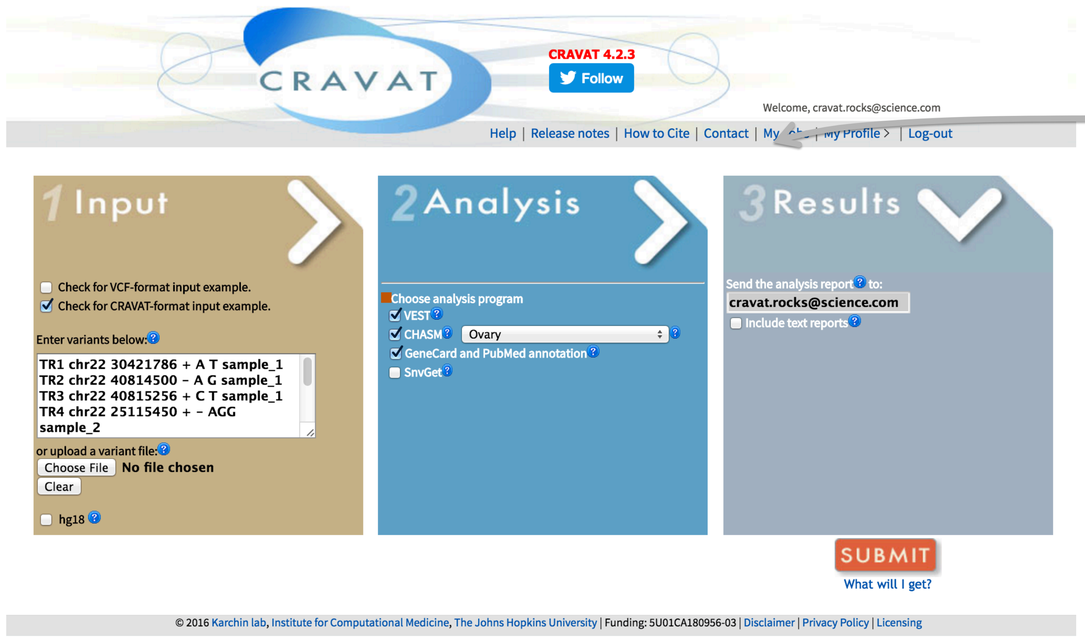
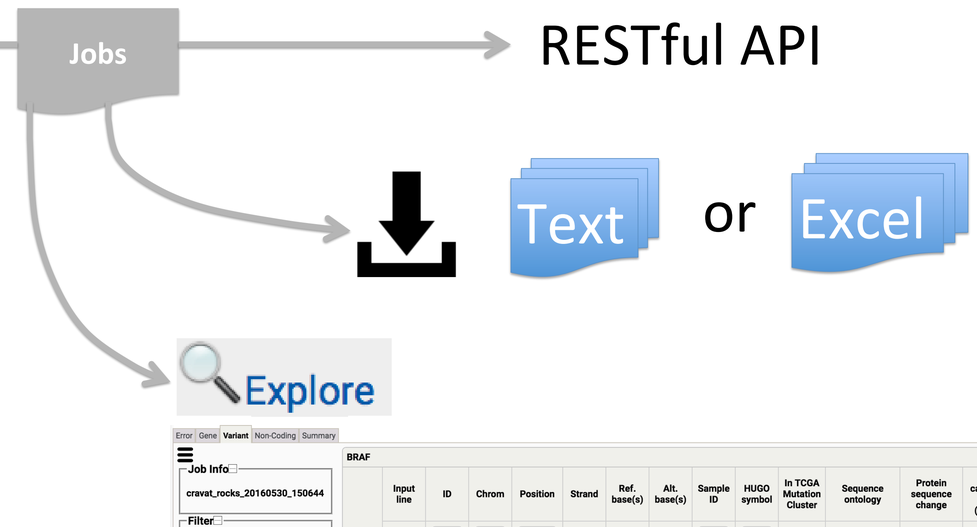
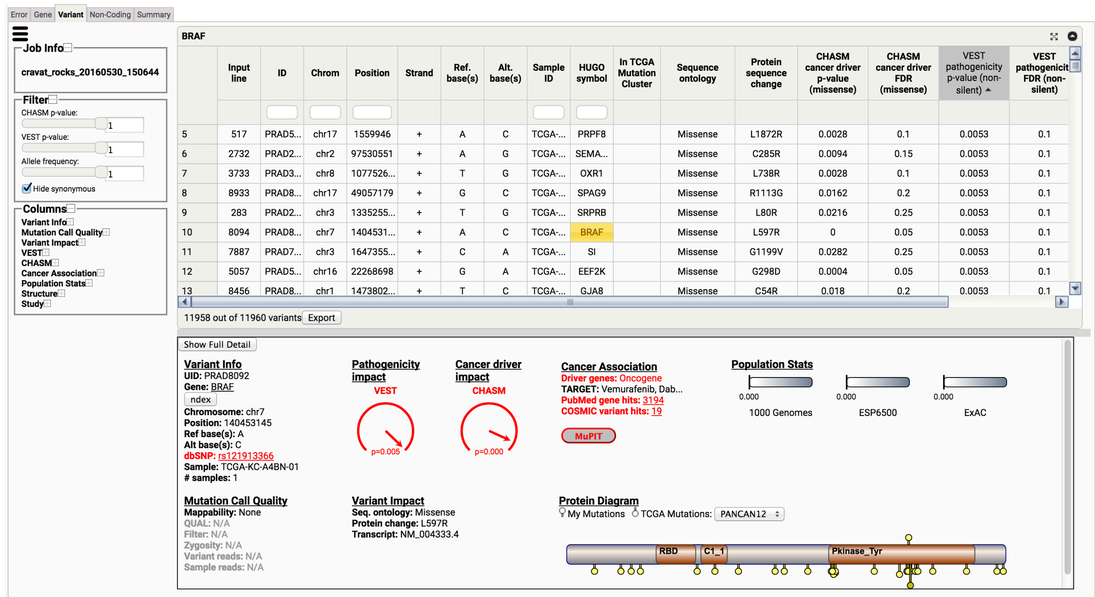
CRAVAT 4.2: informatics tools for high-throughput analysis of exome variants

Collin Tokheim[1,*], Christopher Douville[1,*], David L. Masica[1,*], Rick Kim[3,*], Rohit Bhattacharya[2], Noushin Niknafs[1], Derek Gygax[3], Kyle Moad[3], Michael Ryan[3] and Rachel Karchin[1,4]

[1]Department of Biomedical Engineering [2]Department of Computer Science, Institute for Computational Medicine, Johns Hopkins University, Baltimore, MD, USA
[3]In Silico Solutions, Fairfax, VA, USA [4]Department of Oncology, Johns Hopkins Medical Institutions, Baltimore MD, USA. *Authors contributed equally.







Cancer-Related Analysis of VAriants Toolkit

- ✧ Comprehensive and easy-to-use web interface for assessing and prioritizing genes and variants important in cancer.
- ✧ Scores and annotates all coding exonic non-silent mutation types, including missense, in-frame and frame-shift insertions and deletions, splice variants and stop mutations.
- ✧ Analyze one variant or tens of millions, fast results are quickly emailed to users in spreadsheet and machine-friendly format.
- ✧ Create an account to store your projects online.
- ✧ Create an account to access the new graphical interactive results browser.

Important new features

- ✧ Scoring for all non-silent mutation types.
- ✧ Cancer mutation hotspot regions identified by new HotMAPS algorithm.
- ✧ Docker version, useful for protected data, allows analysis on local computers or in the cloud.
- ✧ Interactive results viewer includes: lollipop diagrams, population MAF widget, dynamic sorting and columns, and pathway enrichment viewer.

Broadened reach

- ✧ CRAVAT is now available in GALAXY.
- ✧ Integration with NDEx, Xena, IGV, and Trinity.
- ✧ Follow on twitter @CravatMupitTool.

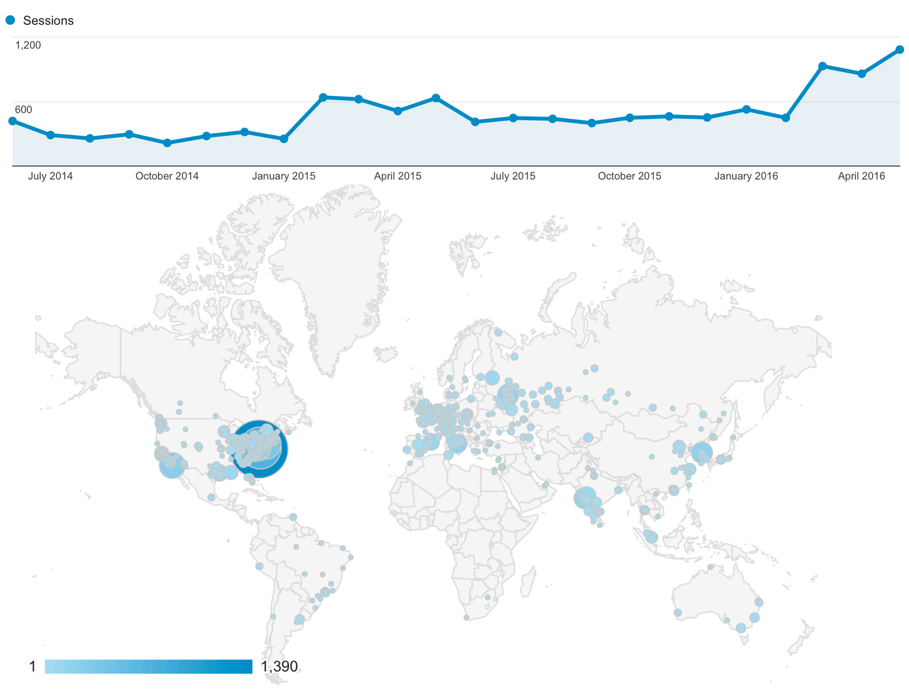
Future development

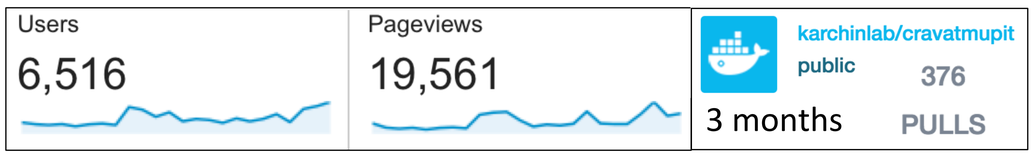
- ✧ New study overview summary page coming soon!
- ✧ Machine learning prediction of cancer driver genes from user's data
- ✧ Annotation for non-coding variants.
- ✧ Availability on NCI cloud pilot platforms.
- ✧ Support for genome assembly hg38.

Publications this year

Tokheim C. *Exome-scale discovery of hotspot mutation regions*. ***Cancer Research*** (2016)
Douville C. *Assessing the pathogenicity of indels*. ***Human Mutation*** (2015)

CRAVAT usage past two years

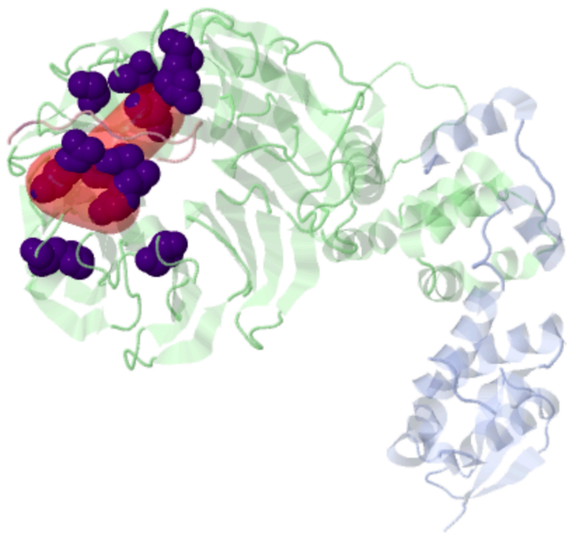




NIH ITCR 5U01CA180956 & U24CA204817

JOHNS HOPKINS MEDICINE | INSTITUTE for COMPUTATIONAL MEDICINE | In Silico SOLUTIONS

HotMAPS algorithm

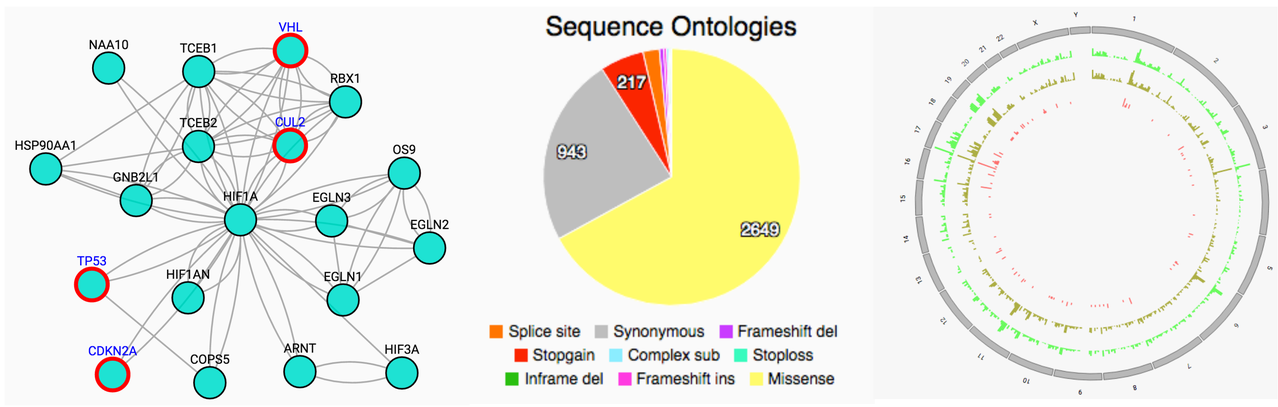


Name	Color	Chain IDs
F-BOX/WD_REPEAT_PROTEIN_7		B
S-PHASE_KINASE-ASSOCIATED_PROTEIN_1A		A
CYCLINE_C-TERMINAL_DEGRON		C

TCGA Mutations

	TCGA Mutations		TCGA 3D Mutation Hot Regions	
Tissue	Color	# Mutated Residues in Structure	Color	Number of TCGA Hot Regions
BLCA		13		1
BRCA		9		
CESC		7		1

New summary page widgets



CRAVAT integration