VariantGrid: Drag and Drop Variant Analysis

David M. Lawrence[1,4], Jinghua Feng[1,4], Liam C. McIntyre[3,4], Karin S. Kassahn[3,4], Hamish S. Scott[1,2,4], Joel Geoghegan[1,4], Andreas W. Schreiber[1,4]

[1] ACRF Cancer Genomics Facility, Centre for Cancer Biology, SA Pathology, Adelaide, South Australia
[2] Genetics and Molecular Pathology and Centre for Cancer Biology, SA Pathology, Adelaide, South Australia, Australia
[3] Genetic and Molecular Pathology, SA Pathology, Women's and Children's Hospital, North Adelaide, South Australia, Australia
[4] School of Biological Sciences, University of Adelaide, South Australia

Variant Database

High throughput sequencing and variant calling has provided an unprecedented level of insight into inherited disorders and cancer.

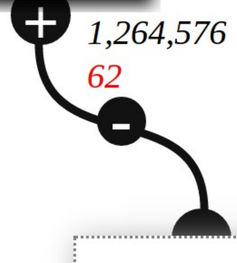
We have built a variant database to manage this data, with a user-friendly web based interface specifically designed to rapidly search through millions of variants. Users chain together nodes representing sources or filters of variants, allowing the creation of custom pipelines.

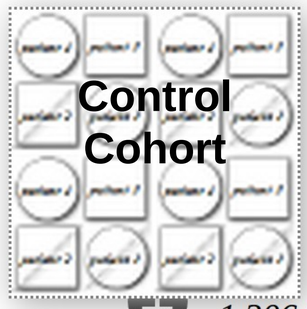
This node represents a patient sample.

The icon is square as the patient is male

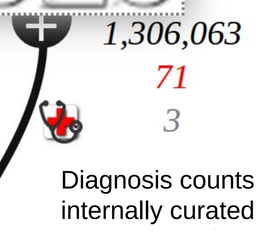
Black numbers show the total variant count.

Red numbers are variants listed as pathological in ClinVar



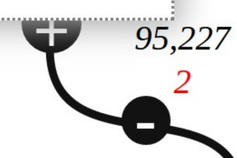


A cohort represents multiple patient samples. This node can be configured to show variants which occur a minimum or maximum number of times across samples.

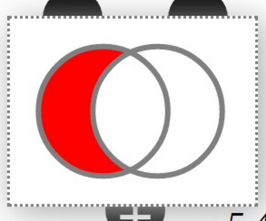


Diagnosis counts shows internally curated variants reported previously.

Filter variants to exon capture regions using a .bed file provided by manufacturer.

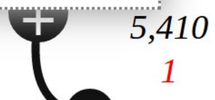


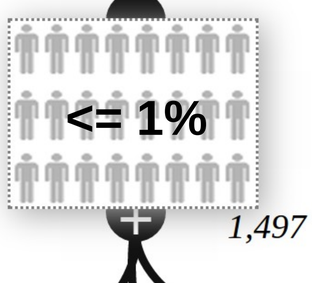
Bottom connector is output, top connector is input



Members of “control cohort” do not have the same disorder as the patient, so we can remove their variants from consideration.

This is done with a Venn Filter configured to take variants in the left parent but not in the right.

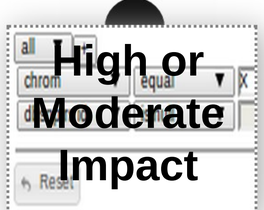




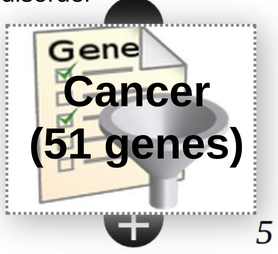
This disorder is rare, so we can remove all variants that appear at greater than 1% frequency in the 1000 genomes, Exac or ESP projects.

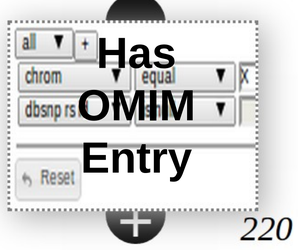
Filter to genes which have entries in public disease databases.

Filter using computational predictions

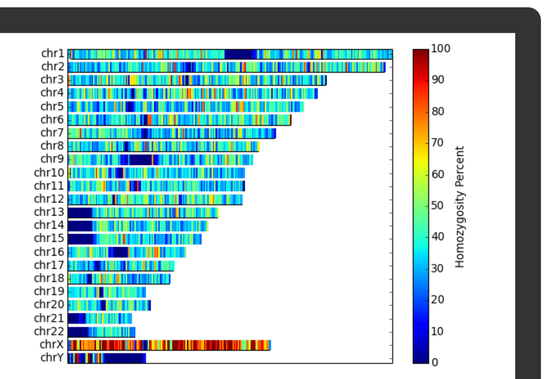


Filter to a list of genes associated with the disorder



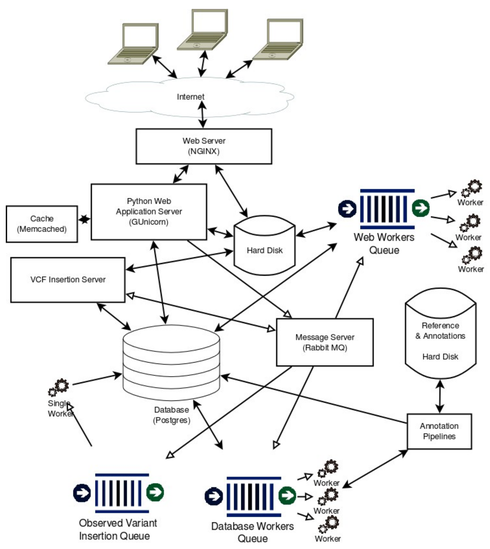


We are constantly improving VariantGrid with the goal of helping research and diagnostics better manage and search through their variants. Contact us to be notified when it will be available for external users.



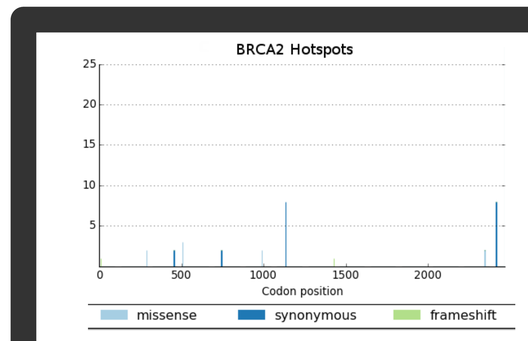
VariantGrid includes graph generation abilities to help visualise your data.

The figure above shows the homozygosity of variants across the genome. Can you tell the sex of this patient?

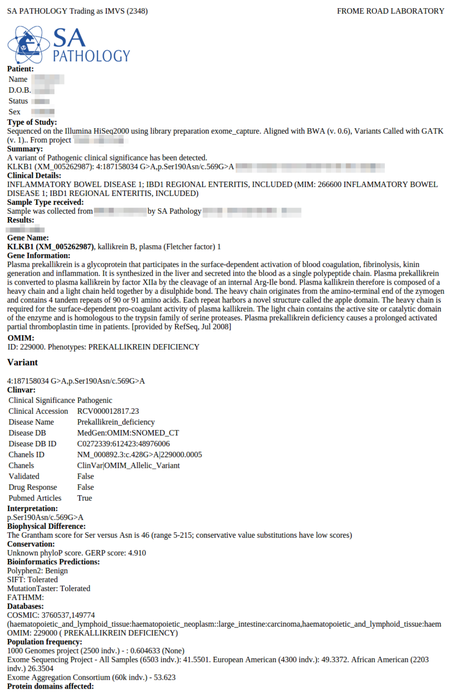


Behind the scenes is a powerful distributed computational back-end.

This allows rapid uploading, annotation and storage, and efficient searching and analysis of your data.



Collect patients and samples into cohorts and pedigrees to allow multi-sample analysis.



SA PATHOLOGY Trading as IMVS (2348) FROME ROAD LABORATORY

SA PATHOLOGY

Patient:
Name
D.O.B.
Status
Sex
Type of Study:
Sequenced on the Illumina HiSeq2000 using library preparation exome_capture. Aligned with BWA (v. 0.6), Variants Called with GATK (v. 1).. From project
Summary:
A variant of Pathogenic clinical significance has been detected.
KLKB1 (XM_005262987): 4:187158034 G>A,p.Ser190Asn/c.569G>A
Clinical Details:
INFLAMMATORY BOWEL DISEASE 1; IBD1 REGIONAL ENTERITIS, INCLUDED (MIM: 266600 INFLAMMATORY BOWEL DISEASE 1; IBD1 REGIONAL ENTERITIS, INCLUDED)
Sample Type received:
Sample was collected from by SA Pathology
Results:
Gene Name:
KLKB1 (XM_005262987), kallikrein B, plasma (Fletcher factor) 1
Gene Information:
Plasma prekallikrein is a glycoprotein that participates in the surface-dependent activation of blood coagulation, fibrinolysis, kinin generation and inflammation. It is synthesized in the liver and secreted into the blood as a single polypeptide chain. Plasma prekallikrein is converted to plasma kallikrein by factor XIIa by the cleavage of an internal Arg-Ile bond. Plasma kallikrein therefore is composed of a heavy chain and a light chain held together by a disulphide bond. The heavy chain originates from the amino-terminal end of the zymogen and contains 4 tandem repeats of 90 or 91 amino acids. Each repeat harbors a novel structure called the apple domain. The heavy chain is required for the surface-dependent pro-coagulant activity of plasma kallikrein. The light chain contains the active site or catalytic domain of the enzyme and is homologous to the trypsin family of serine proteases. Plasma prekallikrein deficiency causes a prolonged activated partial thromboplastin time in patients. [provided by RefSeq, Jul 2008]
OMIM:
ID: 229000. Phenotypes: PREKALLIKREIN DEFICIENCY

Variant

4:187158034 G>A,p.Ser190Asn/c.569G>A
Clinvar:

Clinical Significance	Pathogenic
Clinical Accession	RCV000012817.23
Disease Name	Prekallikrein_deficiency
Disease DB	MedGen:OMIM:SNOMED_CT
Disease DB ID	C0272339:612423:48976006
Chanels ID	NM_000892.3:c.428G>A|229000.0005
Chanels	ClinVar|OMIM_Allelic_Variant
Validated	False
Drug Response	False
Pubmed Articles	True

Interpretation:
p.Ser190Asn/c.569G>A
Biophysical Difference:
The Grantham score for Ser versus Asn is 46 (range 5-215; conservative value substitutions have low scores)
Conservation:
Unknown phyloP score. GERP score: 4.910
Bioinformatics Predictions:
Polyphen2: Benign
SIFT: Tolerated
MutationTaster: Tolerated
FATHMM:
Databases:
COSMIC: 3760537,149774
(haematopoietic_and_lymphoid_tissue:haematopoietic_neoplasm::large_intestine:carcinoma,haematopoietic_and_lymphoid_tissue:haem
OMIM: 229000 (PREKALLIKREIN DEFICIENCY)
Population frequency:
1000 Genomes project (2500 indv.) - : 0.604633 (None)
Exome Sequencing Project - All Samples (6503 indv.): 41.5501. European American (4300 indv.): 49.3372. African American (2203 indv.) 26.3504
Exome Aggregation Consortium (60k indv.) - 53.623
Protein domains affected:

Working closely with Genetics and Molecular Pathology diagnostics, we have built reporting, patient management and variant curation features.

ID	chrc	position	ref	alt	dbsnp rs id	gene symbol	snpeff transcript id	snpeff amino acid change	snpeff coc	snpeff effect	snpi	snpeff impact	low complexity region	segment dupl	tier	esp maf ea	esp maf aa	esp maf all	af 1kg	aa 1kg	exac alt frec	sift	polyph	mutation ta	fathmm	phylop	gerp	cadd score	next prot	pathway from uniprotkb	tissue specificity from uniprotkl	omim phenotypes	gene biotype
468267	1	161643798	T	C	rs1050501	FCGR2B	ENST00000367962	p.Ile232Thr/c.695T>C	aTt/aCt	missense_variant		MODERATE	false	true	1				0.185903	T	10.656	T	B	T	T		2.85	10.18	transmembrane_region:Transm		Is the most broadly distributed f	MALARIA, SUSCEPTIBILITY TO :: SYSTEMIC LUPUS ERYTHEM	protein_coding
2578420	1	115236057	G	A	rs17602729	AMPD1	ENST00000520113	p.Gln45*/c.133C>T	Caa/Taa	stop_gained		HIGH	false	false	1	13.0961	2.3604	9.4586	0.038139	G	8.703			T		1.143	2.84	29.9		Purine metabolism; IMP biosyn	Three isoforms are present in n	MYOPATHY DUE TO MYOADENYLATE DEAMINASE DEFICIENC	protein_coding
2578418	1	115231254	G	A	rs61752479	AMPD1	ENST00000520113	p.Pro81Leu/c.242C>T	cCg/cTg	missense_variant		MODERATE	false	false	1	13.3023	2.3377	9.5879	0.0389377	G	8.836	D	D	T	T	2.804	5.62	19.67		Purine metabolism; IMP biosyn	Three isoforms are present in n	MYOPATHY DUE TO MYOADENYLATE DEAMINASE DEFICIENC	protein_coding
1083852	1	161599693	T	C	rs448740	FCGR3B	ENST00000531221	p.Asn101Ser/c.302A>G	aAc/aGc	missense_variant		MODERATE	false	true	1				0.532947	C	58.778	T		T	T				topological_domain:Extracellula		Expressed specifically by polym		protein_coding
2555209	1	46655645	C	T	rs74374973	POMGNT1	ENST00000371986	p.Asp556Asn/c.1666G>A	Gac/Aac	missense_variant		MODERATE	false	false	1	1.186	0.227	0.8611	0.00519169	C	0.893	T	P	D	D	2.882	6.06	36	topological_domain:Lumenal	Protein modification; protein gly	Constitutively expressed. An ac	MUSCULAR DYSTROPHY-DYSTROGLYCANOPATHY (LIMB-GIR	protein_coding
523211	10	112839575	A	G	rs553668	ADRA2A	ENST00000280155	c.*427A>G	427	3_prime_UTR_variant		MODIFIER	false	false	3				0.670527	A												
2669305	10	99509251	G	T	rs35077384	ZFYVE27	ENST00000356257	p.Gly191Val/c.572G>T	gGc/gTc	missense_variant		MODERATE	false	false	1	0.7674	7.2628	2.9679	0.0261581	G	1.095	D	B	D	T	1.436	4.88	14.71	transmembrane_region:Transm			SPASTIC PARAPLEGIA 33, AUTOSOMAL DOMINANT; SPG33	protein_coding
219756	11	18290859	C	T	rs1136743	SAA1	ENST00000356524	p.Ala70Val/c.209C>T	gCc/gTc	missense_variant		MODERATE	false	true	1						48.657	T	B	T	T						Expressed by the liver; secrete		protein_coding
2742206	11	126162843	C	T	rs8177374	TIRAP	ENST00000392678	p.Ser180Leu/c.539C>T	tCg/tTg	missense_variant		MODERATE	false	false	1	15.8446	3.0441	11.5095	0.0858626	C	12.416		B	T	T	2.672	4.71	10.15	domain:TIR		Highly expressed in liver, kidne	BACTEREMIA, SUSCEPTIBILITY TO, 1; BACTS1 :: MALARIA, SU	protein_coding
2752907	12	10271087	A	C	rs16910526	CLEC7A	ENST00000304084	p.Tyr238*/c.714T>G	taT/taG	stop_gained		HIGH	false	false	1	7.8953	2.6328	6.1126	0.0409345	A	6.307	T		D				14.03	disulfide_bond|domain:C-type_		Highly expressed in peripheral	ASPERGILLOSIS, SUSCEPTIBILITY TO :: CANDIDIASIS, FAMILI	protein_coding
1400779	12	14993439	C	T	rs11276	ART4	ENST00000228936	p.Asp265Asn/c.793G>A	Gac/Aac	missense_variant		MODERATE	false	false	1	39.5349	27.9165	35.599	0.292732	T	34.853	T	B	T	T						Expressed in spleen and T-cell	ADP-RIBOSYLTRANSFERASE 4; ART4	protein_coding
2878913	15	28230318	C	T	rs1800407	OCA2	ENST00000354638	p.Arg419Gln/c.1256G>A	cGg/cAg	missense_variant		MODERATE	false	false	1	7.7791	1.2256	5.559	0.0253594	C	4.412	T	P	T	D	1.384	4.44	21.8	topological_domain:Cytoplasmi			SKIN/HAIR/EYE PIGMENTATION, VARIATION IN, 1; SHEP1 :: AL	protein_coding
2030356	16	69745145	G	A	rs1800566	NQO1	ENST00000320623	p.Pro187Ser/c.559C>T	Cct/Tct	missense_variant		MODERATE	false	false	1	19.8721	19.4722	19.7368	0.288938	G	24.466	T	D	T	T	2.696	5.41	23.7					protein_coding
2914057	16	3820723	T	C		CREBBP	ENST00000262367	p.Thr910Ala/c.2728A>G	Acc/Gcc	missense_variant		MODERATE	false	false	1	0.3372	0.0228	0.2309	0.000399361	T	0.233	T	B	T	D		2.19	11.32				RUBINSTEIN-TAYBI SYNDROME 1; RSTS1	protein_coding
325586	17	12915009	G	A	rs4792311	ELAC2	ENST00000338034	p.Ser217Leu/c.650C>T	tCg/tTg	missense_variant		MODERATE	false	false	1	29.6744	23.468	27.5719	0.214457	G	27.104	T	B	T	T	1.372	3.23				Widely expressed. Highly expre	COMBINED OXIDATIVE PHOSPHORYLATION DEFICIENCY 17; C	protein_coding

Variants and annotations columns from the grid. Users can customise and sort their own columns (30 columns shown, over 100 available)

variantgrid.com

Centre for Cancer Biology · SA PATHOLOGY · UniSA · The University of Adelaide · Australian Cancer Research Foundation · Government of South Australia · SA Health