



インハウス型NGS変異データ管理&解釈支援ワークベンチ

Geneticist Assistant

Geneticist Assistant (ジェネティシストアシスタント) は、メイヨークリニックのLaboratory Medicine,

Information Technology and Health Science Research Departmentsとのコラボレーションで開発された、エキソームシーケンスや遺伝子パネルシーケンスから得られる変異データをインハウスで統合的に管理できるデータベース・ワークベンチです。サンプル単位の変異リストをVCF (Variant Call Format) ファイルとBAMファイル (オプション) でRunごとにデータベースにアップロードし、Runやサンプルの変異情報をパネルや患者情報と紐づけて統合的に管理できます。

Variants of 722305.variants.filter:														
ID	Chr	Ref	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	Coverage	HGVSP	Protein	Panel	HGVSC	Coding	Times Observed Per Panel
7	5	112162854	rs2229992	Likely Deleterious	APC	12	synonymous	0.5	69	p.Tyr486>	DIMP	c.1458T>C	10	10
8	5	112164561	rs351771	Benign	APC	14	synonymous	0.5	69	p.Ala545>	DIMP	c.1635G>A	10	10
9	10	88635779	rs11528010	Likely Deleterious	BMPRIA	3	missense	1	99	p.Pro27Irr	DIMP	c.AC>A	5	5
14	14	75513883	rs125981	Benign	MLH3	2	missense	1	55	p.Asn628Arg	DIMP	c.2476A>G	11	11
15	17	7579472	rs1042522	Deleterious	TP53	4	missense	1	46	p.Pro72Arg	DIMP	c.215C>G	9	9
16	17	63533768	rs1133683	Deleterious	AXIN2	6	synonymous	0.5	50	p.Pro462>	DIMP	c.1386C>T	8	8
17	17	63533789	rs9915936	Likely Benign	AXIN2	6	synonymous	0.5	54	p.Pro455>	DIMP	c.1365A>G	9	9
18	17	6354591	rs2240308	Likely Benign	AXIN2	2	missense	1	111	p.Pro305Ser	DIMP	c.148C>T	8	8
38	14	75513828	rs125981	Unknown	MLH3	2	missense	1	55	p.Pro64Leu	DIMP	c.2531C>T	6	6
45	2	48010488	rs1042821	Unknown	MSH6	1	missense	1	64	p.Gly39Glu	DIMP	c.116G>A	3	3

Variants of 722305.variants.filter:														
ID	Chr	Ref	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	Coverage	HGVSP	Protein	Panel	HGVSC	Coding	Times Observed Per Panel
5	5	112162854	rs2229992	Likely Deleterious	APC	12	synonymous	0.5	69	p.Tyr486>	DIMP	c.1458T>C	10	10
5	5	112164561	rs351771	Benign	APC	14	synonymous	0.5	69	p.Ala545>	DIMP	c.1635G>A	10	10
10	10	88635779	rs11528010	Likely Deleterious	BMPRIA	3	missense	1	99	p.Pro27Irr	DIMP	c.AC>A	5	5
14	14	75513883	rs125981	Benign	MLH3	2	missense	1	55	p.Asn628Arg	DIMP	c.2476A>G	11	11
17	17	7579472	rs1042522	Deleterious	TP53	4	missense	1	46	p.Pro72Arg	DIMP	c.215C>G	9	9
17	17	63533768	rs1133683	Deleterious	AXIN2	6	synonymous	0.5	50	p.Pro462>	DIMP	c.1386C>T	8	8
17	17	63533789	rs9915936	Likely Benign	AXIN2	6	synonymous	0.5	54	p.Pro455>	DIMP	c.1365A>G	9	9
18	17	6354591	rs2240308	Likely Benign	AXIN2	2	missense	1	111	p.Pro305Ser	DIMP	c.148C>T	8	8
38	14	75513828	rs125981	Unknown	MLH3	2	missense	1	55	p.Pro64Leu	DIMP	c.2531C>T	6	6
45	2	48010488	rs1042821	Unknown	MSH6	1	missense	1	64	p.Gly39Glu	DIMP	c.116G>A	3	3

ヒストリカルデータベースの構築

Geneticist Assistantはみつかった全てのVariantについて、決定したVariantの病原性を記録します。これにより、同じVariantについて二重に調べる無駄を省き、コストを削減しつつ迅速な分析が可能になります。このデータベースを使うことで、病原性判定が必要な膨大な数のVariantをいくつかの新規Variantへ素早く減らすことができます。

C:\Users\soft\Desktop\GA\references\Human 37\ESP55005_V2_SS137.vcf														
Chromosome	ID	Ref	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	Coverage	HGVSP	Protein	Panel	HGVSC	Coding	Times Observed Per Panel
Chromosome	17	7579472	rs1042522	TP53	4	missense	1	46	p.Pro72Arg	DIMP	c.215C>G	9	9	9
Chromosome	17	63533768	rs1133683	AXIN2	6	synonymous	0.5	50	p.Pro462>	DIMP	c.1386C>T	8	8	8
Chromosome	17	63533789	rs9915936	AXIN2	6	synonymous	0.5	54	p.Pro455>	DIMP	c.1365A>G	9	9	9
Chromosome	17	6354591	rs2240308	AXIN2	2	missense	1	111	p.Pro305Ser	DIMP	c.148C>T	8	8	8
Chromosome	14	75513828	rs125981	MLH3	2	missense	1	55	p.Pro64Leu	DIMP	c.2531C>T	6	6	6
Chromosome	2	48010488	rs1042821	MSH6	1	missense	1	64	p.Gly39Glu	DIMP	c.116G>A	3	3	3

C:\Users\soft\Desktop\GA\references\Human 37\clinvar_00_latest.vcf														
Chromosome	ID	Ref	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	Coverage	HGVSP	Protein	Panel	HGVSC	Coding	Times Observed Per Panel
Chromosome	17	7579472	rs1042522	TP53	4	missense	1	46	p.Pro72Arg	DIMP	c.215C>G	9	9	9
Chromosome	17	63533768	rs1133683	AXIN2	6	synonymous	0.5	50	p.Pro462>	DIMP	c.1386C>T	8	8	8
Chromosome	17	63533789	rs9915936	AXIN2	6	synonymous	0.5	54	p.Pro455>	DIMP	c.1365A>G	9	9	9
Chromosome	17	6354591	rs2240308	AXIN2	2	missense	1	111	p.Pro305Ser	DIMP	c.148C>T	8	8	8
Chromosome	14	75513828	rs125981	MLH3	2	missense	1	55	p.Pro64Leu	DIMP	c.2531C>T	6	6	6
Chromosome	2	48010488	rs1042821	MSH6	1	missense	1	64	p.Gly39Glu	DIMP	c.116G>A	3	3	3

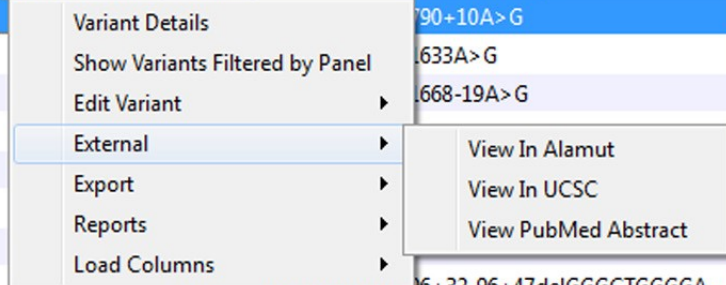
C:\Users\soft\Desktop\GA\references\Human 37\clinvar_02_4_variant.vcf														
Chromosome	ID	Ref	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	Coverage	HGVSP	Protein	Panel	HGVSC	Coding	Times Observed Per Panel
Chromosome	17	7579472	rs1042522	TP53	4	missense	1	46	p.Pro72Arg	DIMP	c.215C>G	9	9	9
Chromosome	17	63533768	rs1133683	AXIN2	6	synonymous	0.5	50	p.Pro462>	DIMP	c.1386C>T	8	8	8
Chromosome	17	63533789	rs9915936	AXIN2	6	synonymous	0.5	54	p.Pro455>	DIMP	c.1365A>G	9	9	9
Chromosome	17	6354591	rs2240308	AXIN2	2	missense	1	111	p.Pro305Ser	DIMP	c.148C>T	8	8	8
Chromosome	14	75513828	rs125981	MLH3	2	missense	1	55	p.Pro64Leu	DIMP	c.2531C>T	6	6	6
Chromosome	2	48010488	rs1042821	MSH6	1	missense	1	64	p.Gly39Glu	DIMP	c.116G>A	3	3	3

C:\Users\soft\Desktop\GA\references\Human 37\clinvar_02_4_variant.vcf														
Chromosome	ID	Ref	Pathogenicity	Gene	Exon Number	Type	Variant Frequency	Coverage	HGVSP	Protein	Panel	HGVSC	Coding	Times Observed Per Panel
Chromosome	17	7579472	rs1042522	TP53	4	missense	1	46	p.Pro72Arg	DIMP	c.215C>G	9	9	9
Chromosome	17	63533768	rs1133683	AXIN2	6	synonymous	0.5	50	p.Pro462>	DIMP	c.1386C>T	8	8	8
Chromosome	17	63533789	rs9915936	AXIN2	6	synonymous	0.5	54	p.Pro455>	DIMP	c.1365A>G	9	9	9
Chromosome	17	6354591	rs2240308	AXIN2	2	missense	1	111	p.Pro305Ser	DIMP	c.148C>T	8	8	8
Chromosome	14	75513828	rs125981	MLH3	2	missense	1	55	p.Pro64Leu	DIMP	c.2531C>T	6	6	6
Chromosome	2	48010488	rs1042821	MSH6	1	missense	1	64	p.Gly39Glu	DIMP	c.116G>A	3	3	3

Variant病原性判定情報

- 機能予測: SIFT/PolyPhen2/LRT/Mutation Taster/FATHMM/CADD/Mutatin Assessor
- 疾患関連性: ClinVar/OMIM/COSMIC/CiViC/Alamut*/LOVD(Leiden Open Variant Database)
- Conservation scores: PhyloP/GERP++/phastCons/SiPhy
- 集団頻度: 1000 Genomes/Exome Variant Server/ExAC

Variants of '800463.igv-sorted Output Mutation Report1_filtered': *Filters Applied						
ID	Chr : ChrPos	Rs	Pathogenicity	Pathogenicity Status	Gene	HGVS Coding
28	3:37056045	rs182733777	Unassigned			90+10A>G
29	3:37081751	rs267607840	Unassigned			633A>G
6	3:37083740	rs9876116	Benign			668-19A>G
13	14:75505016	rs175075	Benign			
14	14:75513883	rs175081	Benign			
31	14:75514489	rs28756986	Unassigned			
15	17:7579472	rs1042522	Deleterious			
1443	17:7579669	rs17878362	Unassigned			6+32_96+47delGGGCTGGGGA...
18	17:63554591	rs2240308	Likely Deleterious	Confirmed	AXIN2	NM_004655.3:c.148C>T
19	18:48584856	rs386387676	Likely Benign		SMAD4	NM_005359.5:c.904+45_904+46insTT



外部データベースとの連携

Alamut*やUCSCゲノムブラウザ、LOVDデータベースなど外部ツールからさらなる情報を簡単な操作で検索できます（*Alamutは別途ライセンスが必要です。）。

Family Comparison of 4 Samples:									
AF	Cov	AF	Cov	AF	Cov	AF	Cov	ID	Chr:O
0.84	166	0.437	119	1.000	121	0.984	127	1640	1:114443
0.527	93	0.448	58	1.000	56	0.984	64	1745	1:1108307
0.409	22	0.316	19	1.000	7	1.000	11	3167	2:206911
0.636	11	0.667	24	0.773	22	0.800	10	3306	2:111384
0.471	121	0.516	85	1.000	97	1.000	79	3382	2:21055790
0.485	130	0.426	122	1.000	135	0.992	118	3390	2:21055792
0.466	103	0.469	113	0.987	76	0.959	74	3396	2:210558162
0.374	187	0.379	145	1.000	169	1.000	130	3400	2:21055960
0.425	80	0.594	64	0.982	55	1.000	59	3428	2:211456637
0.436	78	0.597	62	0.982	57	1.000	59	3432	2:211456639
0.423	184	0.500	182	0.981	157	1.000	136	3440	2:211481257
0.625	72	0.519	79	1.000	86	1.000	82	3594	2:159636316

Table Filters	
ID	Type does not contain synonymous
Chromosome	
Chromosome Position	
Chr : ChrPos	Coverage >= 300
Rs	Variant Frequency > 0.40
Ref	Frequency Observed Per Panel < 0.10
Ref AA	Coverage >= 100
Alt	10000p1_AF < 0.05
Alt AA	
Type	
Coverage	
Pathogenicity	
Pathogenicity Status	
Artifact Type	
Variant Frequency	
Zygosity	
Read Balance	
Gene	
Gene Strand	
Exon Number	
Transcript	
Protein	
Coding Base	
Codon Position	
AA Position	
HGVS Genomic	
HGVS Coding	
HGVS Protein	
Variant Comment	
Times Observed Per Run	
Delete All	
Apply Filters	
Clear Filters	
Save Filters	
Load Filters	

サンプル比較解析

Variant CallやVariant頻度を複数サンプル間で比較できます。別々の解析パイプラインから出力されたVariantデータも、VCFファイルをインポートすることにより、トリオ解析など家族メンバー間の比較解析が可能です。

- 研究用のみ使用できます。診断目的およびその手続き上での使用はできません。
- ここに掲載の内容、ソフトウェア仕様は予告無く変更されることがあります。（2019年7月現在）
- 記載の社名および製品名は、SoftGenetics社または各社の商標または登録商標です。

カスタムフィルタリングオプション

Variantリストを始めとするGeneticist Assistant内のデータテーブルは、データフィールドを組み合わせることでフィルタリングすることが可能です。データフィールドはドラッグアンドドロップでフィルタリングに使用でき、複数のフィルターを組み合わせられます。フィルターの組み合わせは保存しておき別の解析に使用できます。

輸入・販売元 バイオアップロード合同会社

- オフィス 栃木県足利市伊勢町3丁目4-8
- TEL/FAX 0284-22-4213/4223
- E-mail info@bio-upload.com
- URL <https://www.bio-upload.com/>

開発元 米国SoftGenetics, LLC.

- Address 100 Oakwood Avenue, Suite 350, State College, PA 16803, USA
- URL <http://www.softgenetics.com/index.php>

最小システム要件 (Geneticist Assistant Version: 1.8)

サーバー/クライアント

- OS: Windows 64 bit OS: 7, 8.1, 10または64bit Linux (Ubuntu12.04以降推奨) / Windows 64 bit OS: 7, 8.1, 10
- CPU: 2コア以上
- 動作メモリ: 4GB RAM以上/8GB RAM以上
- 必要HDD容量: 100GB以上の空き容量(SSD推奨) / 250GB以上の空き容量