病的バリアントの分類・解釈の自動化

フィルジェン株式会社　バイオインフォマティクス部
(support@filgen.jp)

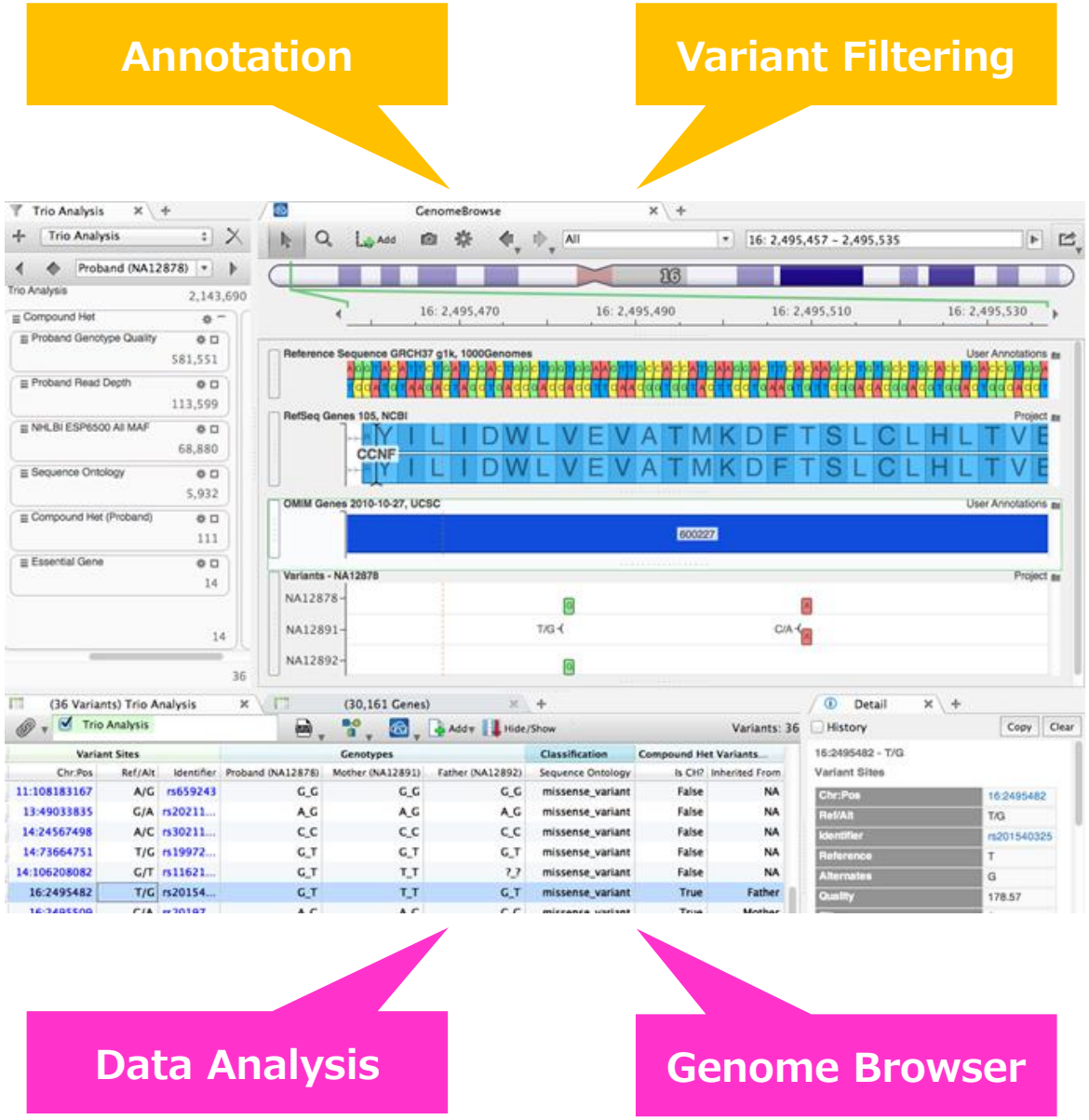
はじめに

- 次世代シークエンサーで検出されたバリアントのリストより、病的なバリアントを抽出するには、文献やデータベース検索を駆使しながら、バリアントごとの病原性を評価する必要があり、非常に手間がかかります。

- Golden Helix社VarSeq®の有償アドオンであるVSClinical ACMGでは、ACMGガイドラインに基づき、生殖細胞系列バリアントの病原性を自動で評価することで、遺伝学的検査の結果の解釈を、効率的に行うことが可能です。

VarSeq®

Filgen® biosciences & nanosciences



■ 様々なデータベースを用いて、バリアントデータ（VCFファイル）へアノテーション付けを実行

- RefSeq Genes
- dbSNP
- ClinVar
- OMIM
- CIViC
- ICGC / TCGA
- MSK Impact
- Orphanet
- BRCA Exchange
- PMKB
- dbNSFP
- REVEL
- CADD
- 1000 Genomes
- NHLBI 6500 Exomes
- ExAC Variant
- gnomAD Exomes/Genomes
- GenomeAsia 100K
- 各種遺伝子パネルのターゲットデータ

...など

■ アノテーション付けされたバリアントデータより、任意の検索条件でデータのフィルタリングを行うワークフローを作成

■ カバレッジ計算やトリオ解析、表現型情報に基づく遺伝子ランキングなどの解析アルゴリズムを搭載

■ ゲノムブラウザーにより、サンプルのバリアントデータやリードアライメントデータ（BAM/CRAMファイル）、また各種アノテーションを可視化

■ 有償アドオンによる機能拡張で、CNV検出や臨床的意義の自動評価、パイプライン機能などが利用可能

VSClinical

VSClinical (ACMG)

- メンデル遺伝病における生殖細胞系列バリアントを、ACMGガイドラインの33種類の評価基準に基づいて分類し、病原性（Pathogenic）や良性（Benign）の判定を行う
- ガイドラインのうち18種類の評価基準については、バリアントのアレル頻度、機能予測、臨床エビデンスデータベースなどを用いて、自動分類を実行
- 専用の分類用ツールを実行することで、VCFファイルに含まれる全バリアントに対して一括で評価を行い、評価結果に基づきバリアントのフィルタリングが可能

VSClinical (AMP)

- がんにおける体細胞バリアントの臨床的意義の評価に使用
- 各種バイオマーカー（SNV, InDels, CNV, 融合遺伝子, TMB/MSIなど）をAMPガイドラインのエビデンスレベルで分類し、がんの治療薬や臨床試験情報を含めたレポートを作成
- 主要ながんにおけるバイオマーカー情報などを収録した、専用の知識ベースGolden Helix CancerKBが利用可能
- 体細胞バリアントの腫瘍原性（Oncogenicity）をスコアで評価

ACMGガイドライン

■ **病原性の評価方法**

- 一般集団内のアレル頻度
- 既知の臨床エビデンス
- 遺伝子機能への影響
- バイオインフォマティクス解析による機能予測
- 家族歴

■ **評価項目と結果**

- 33種類の評価項目
- 評価スコアの合計より、バリアントの病原性を5段階に分類
  - ➢ Pathogenic
  - ➢ Likely Pathogenic
  - ➢ VUS (Variant of Uncertain Significance)
  - ➢ Likely Benign
  - ➢ Benign

	Benign		Pathogenic			
	Strong	Supporting	Supporting	Moderate	Strong	Very Strong
Population Data	MAF is too high for disorder *BA1/BS1* OR observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

VSClinicalによる自動評価

#CHROM	POS	ID	REF	ALT
1	569492	rs6594033	T	C
4	4927684	rs1264288	A	G
10	4586828	rs1465531	G	A

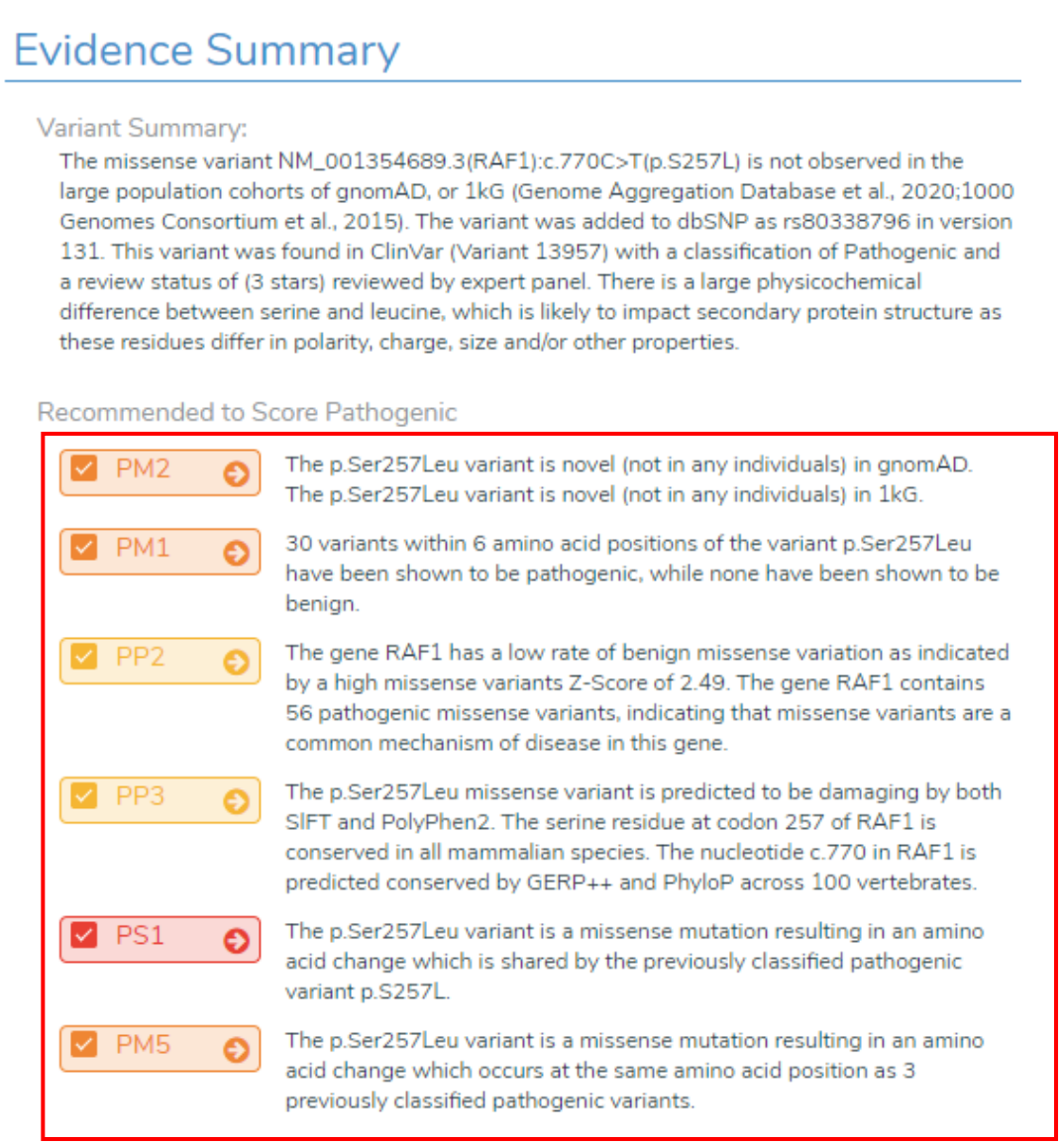
バリアントリスト

Variant 1	Benign
Variant 2	Pathogenic
Variant 3	Likely Pathogenic

評価結果

- VSClinicalでは、バリアントのリストを読み込むだけで、自動でバリアントの評価を実行
- バリアントリストはVCFファイルまたは、バリアント名の手動入力に対応

評価結果の出力データ



検出された評価項目

- VSClinicalでバリアントの評価を行うと、バリアントごとに評価項目が自動で検出される
- 評価項目のサマリーテキストでは、アノテーションデータベースの照合結果などをまとめたエビデンス情報なども確認が可能

評価結果の出力データ

■ **評価結果の出力データ**

- VSClinicalでは、パスした評価項目のスコア、および病原性の分類結果が表示される
- 各評価項目の強さや、サマリーなども確認が可能

評価項目の集計データ

Scoring System:

ACMG Variant Classification (Richards et al. 2015)

To change the classification system close the evaluation and edit the Project Options...

Scored Criteria by Strength:

Pathogenic	Very Strong		x0
	Strong	PS1	x1
	Moderate	PM2, PM1, PM5	x3
	Supporting	PP2, PP3	x2
Benign	Supporting		x0
	Strong		x0
	Stand Alone		x0

ACMG Classification:

Pathogenic

Rule Pathogenic (iii): 1 Strong AND ≥3 Moderate, OR 2 Moderate AND ≥2 Supporting, OR 1 Moderate AND ≥4 Supporting

バリアントの病原性評価の分類

Recommended Criteria:

- Perform functional assay to determine the effect of the variant in the gene
- Establish the state of the variant in the parents

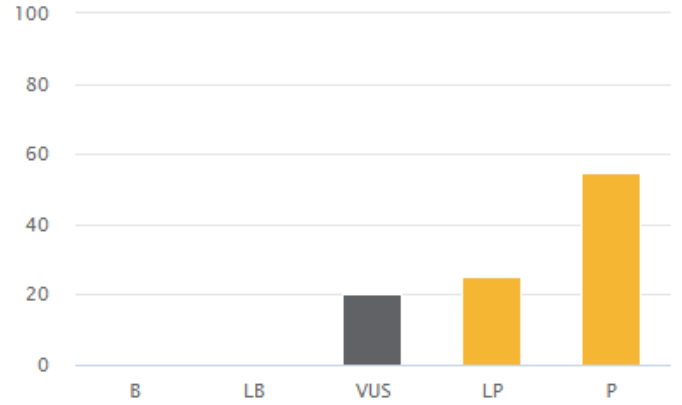
> ACMG Criteria Summary

Golden Helix Classification Prediction

Probability of Pathogenic given Scored Criteria:

Status: Probability of 88.7% predicting classification of Pathogenic

Probability for Each Classification:



A probability model was developed that takes the number of criteria scored at each evidence level and computes a probability of each of the five classifications. It was modeled and trained to agree with ACMG classification rules when provided non-conflicting criteria.

評価に用いるデータベース

■ **遺伝子データベース**

- RefSeq

■ **アレル頻度データベース**

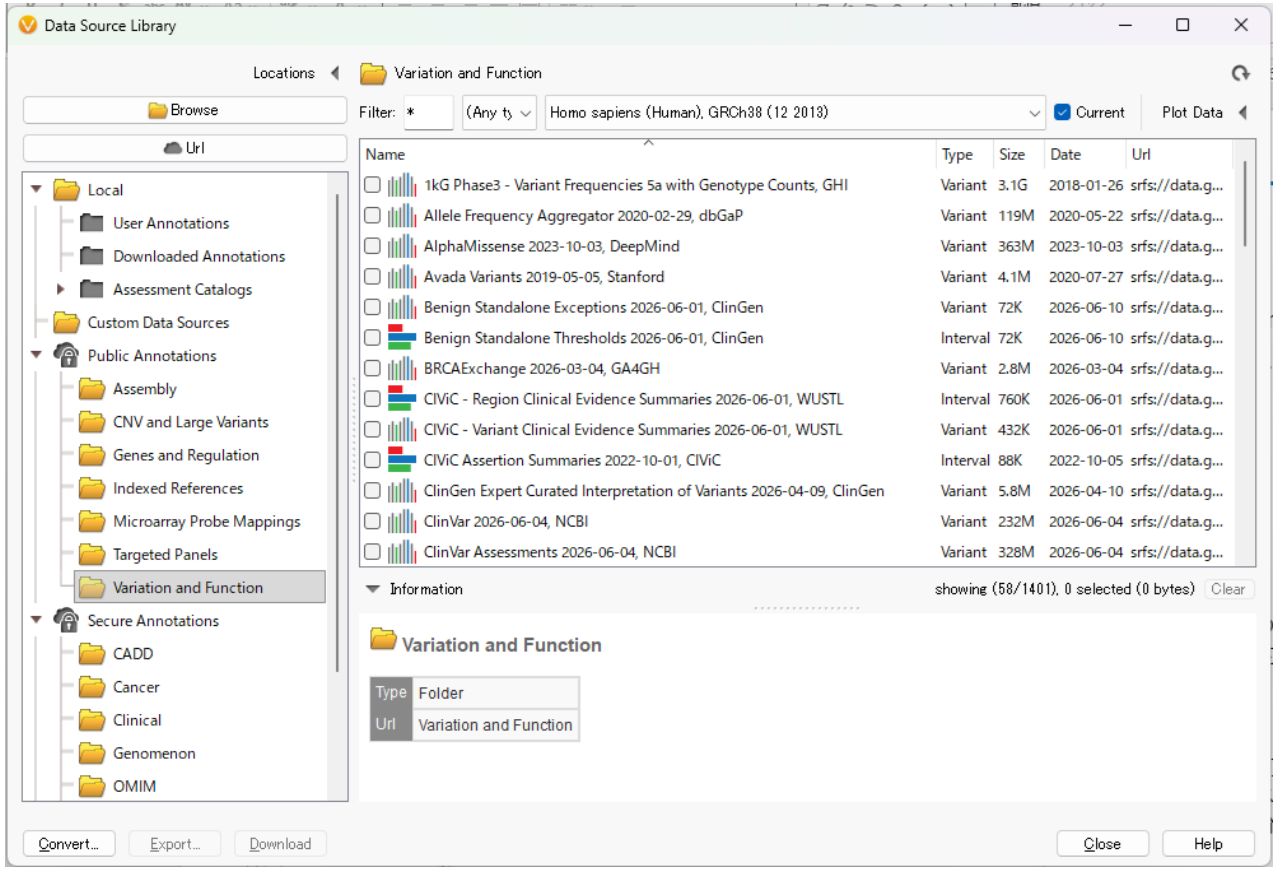
- 1000Genome
- gnomAD

■ **疾患ゲノムデータベース**

- ClinVar
- ClinGen
- OMIM
- Orphanet

■ **機能予測データベース**

- SIFT
- Polyphen2
- Conservation Score
- CADD
- GeneSplicer …など

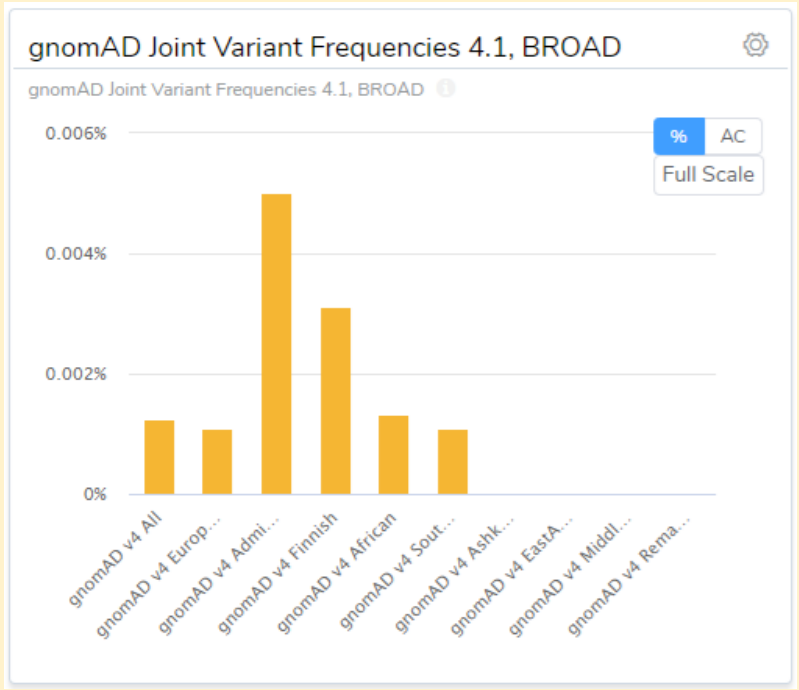


■ 評価を行うのに必要なアノテーションデータベースは、最新のものが自動的にダウンロードされ使用される

■ ソフトウェアのアノテーションライブラリーより、任意のアノテーションデータベースをダウンロードし、アノテーション付けに使用することも可能

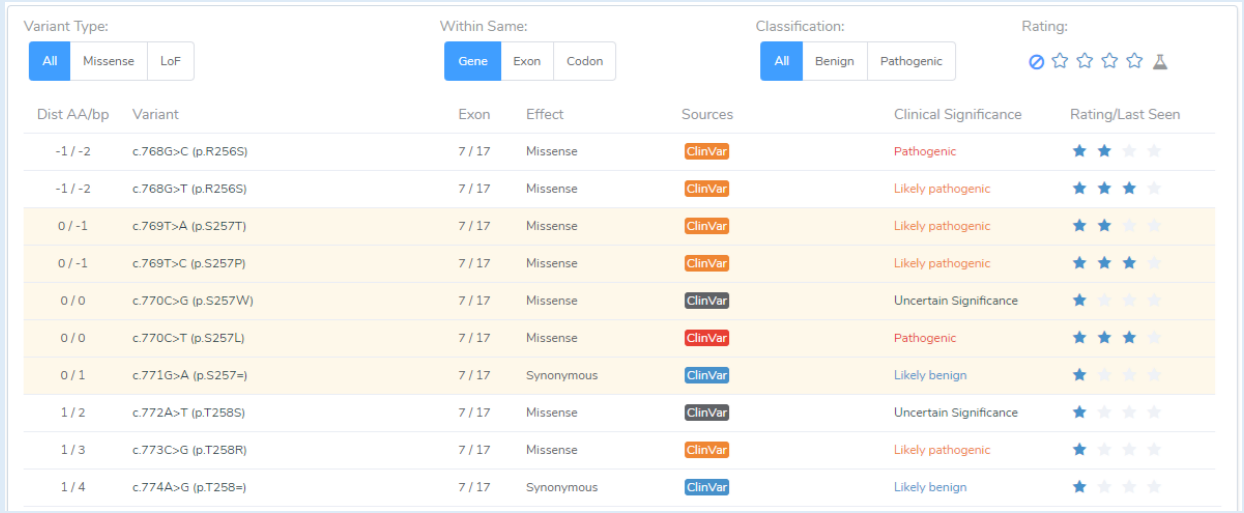
評価項目

一般集団内のアレル頻度による評価



- gnomAD, 1000 Genomesのアレル頻度データベースをもとに、バリアントの一般集団内のアレル頻度を確認
- 集団内のアレル頻度が高いバリアントは良性、データベースに存在しないか、または頻度の低いものは病的と評価する

遺伝子上の位置情報を用いた評価

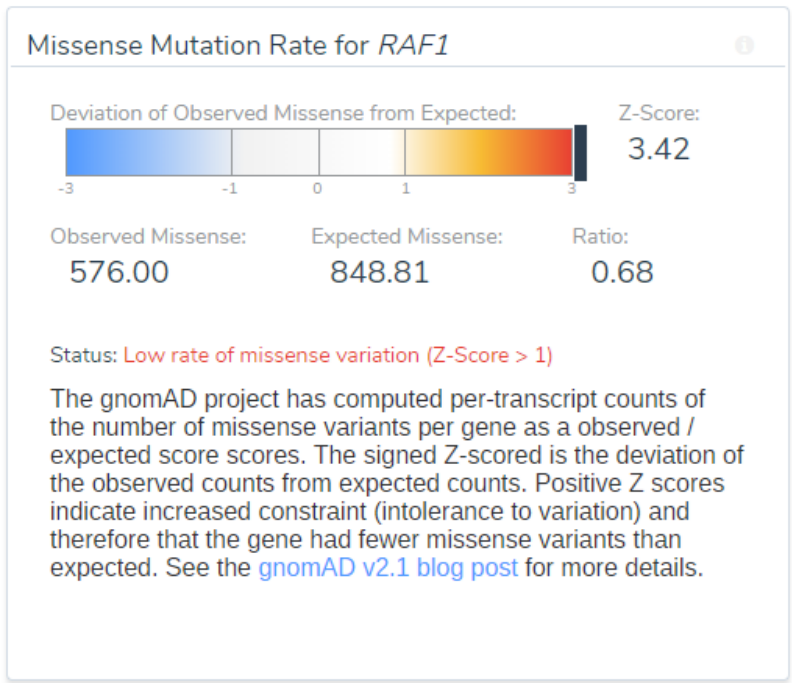


Dist AA/bp	Variant	Exon	Effect	Sources	Clinical Significance	Rating/Last Seen
-1 / -2	c.768G>C (p.R256S)	7 / 17	Missense	ClinVar	Pathogenic	★★☆☆
-1 / -2	c.768G>T (p.R256S)	7 / 17	Missense	ClinVar	Likely pathogenic	★★★☆
0 / -1	c.769T>A (p.S257T)	7 / 17	Missense	ClinVar	Likely pathogenic	★★☆☆
0 / -1	c.769T>C (p.S257P)	7 / 17	Missense	ClinVar	Likely pathogenic	★★★☆
0 / 0	c.770C>G (p.S257W)	7 / 17	Missense	ClinVar	Uncertain Significance	★☆☆☆
0 / 0	c.770C>T (p.S257L)	7 / 17	Missense	ClinVar	Pathogenic	★★★☆
0 / 1	c.771G>A (p.S257=)	7 / 17	Synonymous	ClinVar	Likely benign	★☆☆☆
1 / 2	c.772A>T (p.T258S)	7 / 17	Missense	ClinVar	Uncertain Significance	★☆☆☆
1 / 3	c.773C>G (p.T258R)	7 / 17	Missense	ClinVar	Likely pathogenic	★☆☆☆
1 / 4	c.774A>G (p.T258=)	7 / 17	Synonymous	ClinVar	Likely benign	★☆☆☆

- バリアントが、遺伝子のホットスポットや機能ドメイン上に存在するかどうかや、近傍に良性バリアントが存在するかどうかを確認
- ホットスポット上に存在する場合や、同じコドン領域に良性のバリアントが他に存在しない場合、病的と評価する

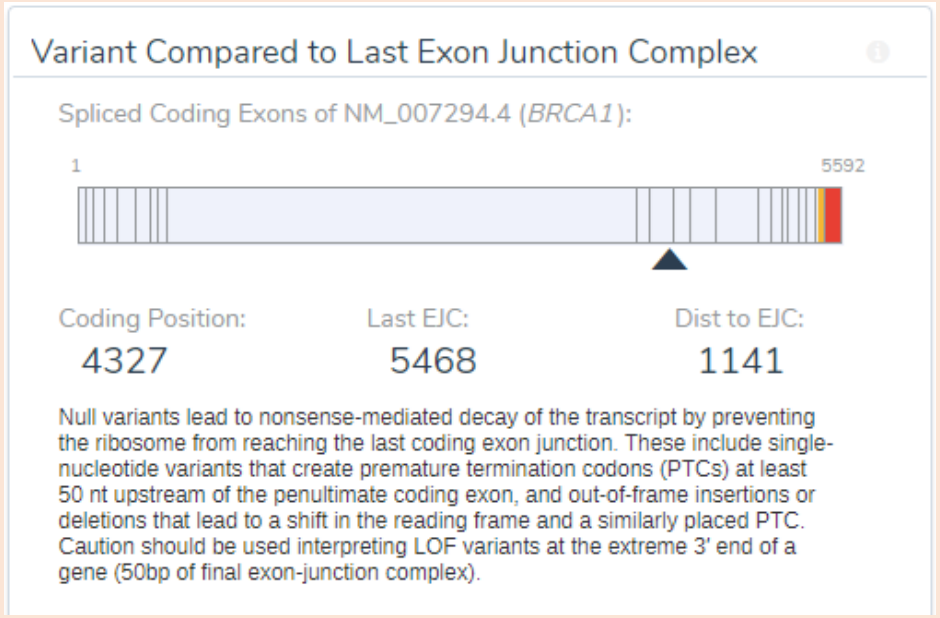
評価項目

ミスセンスバリアントの評価



- データベース登録されている良性のミスセンスバリアントの割合や、コンピュータ予測によって、ミスセンスバリアントの病原性を評価したスコアが計算される
- 良性のミスセンスバリアントが多く、なおかつタンパク質の短縮を引き起こすバリアントが疾患の原因だと考えられる場合は、ミスセンスバリアントは良性と判定される

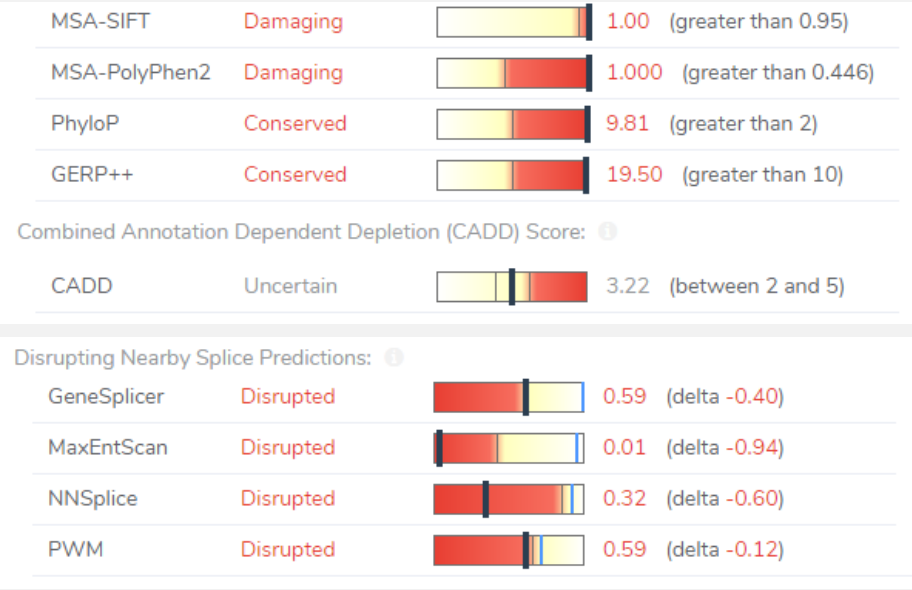
ナンセンスバリアントの評価



- 同じ遺伝子上で、評価対象のバリアントより下流に、病原性のナンセンスバリアントが多数報告されている場合は、強い病原性をもつと判定されるが、バリアントが遺伝子の最後のエクソン（または最後から2番目のエクソンの最終50bp以内）に位置する場合は、良性とする
- 遺伝子ごとに、ナンセンスバリアントの実測値／期待値のスコアが計算されるので、この値に基づいて評価を行うことも可能

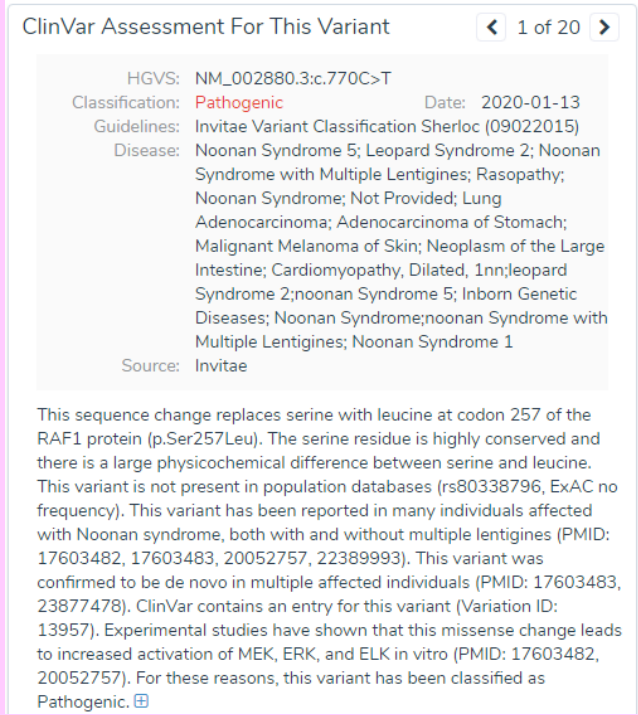
評価項目

コンピュータ予測による評価



- 生物種間の保存度や、SIFT, PolyPhen2などの複数のコンピュータ予測用ツールによって、バリアントの有害性を評価する
- スプライシング異常を引き起こすバリアントについては、GeneSplicerなどのスプライス部位予測用ツールが使用される

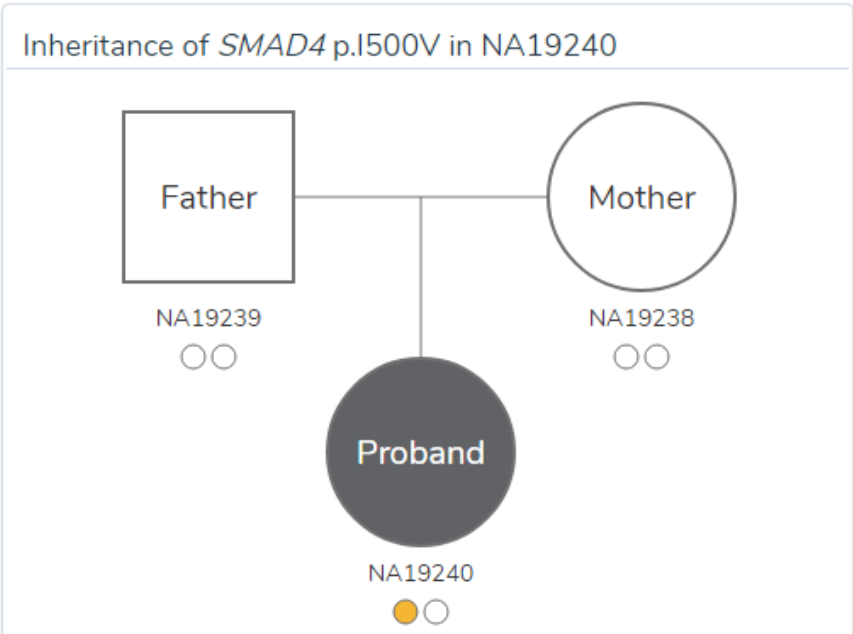
既知の臨床エビデンスによる評価



- データベースで病的バリアントと登録されているものと同じアミノ酸置換を引き起こすバリアントや、遺伝子上の同じ位置に存在し、異なるアミノ酸置換を引き起こすバリアントにを評価する
- データベース登録されているレコード情報や、出典論文情報なども確認が可能

評価項目

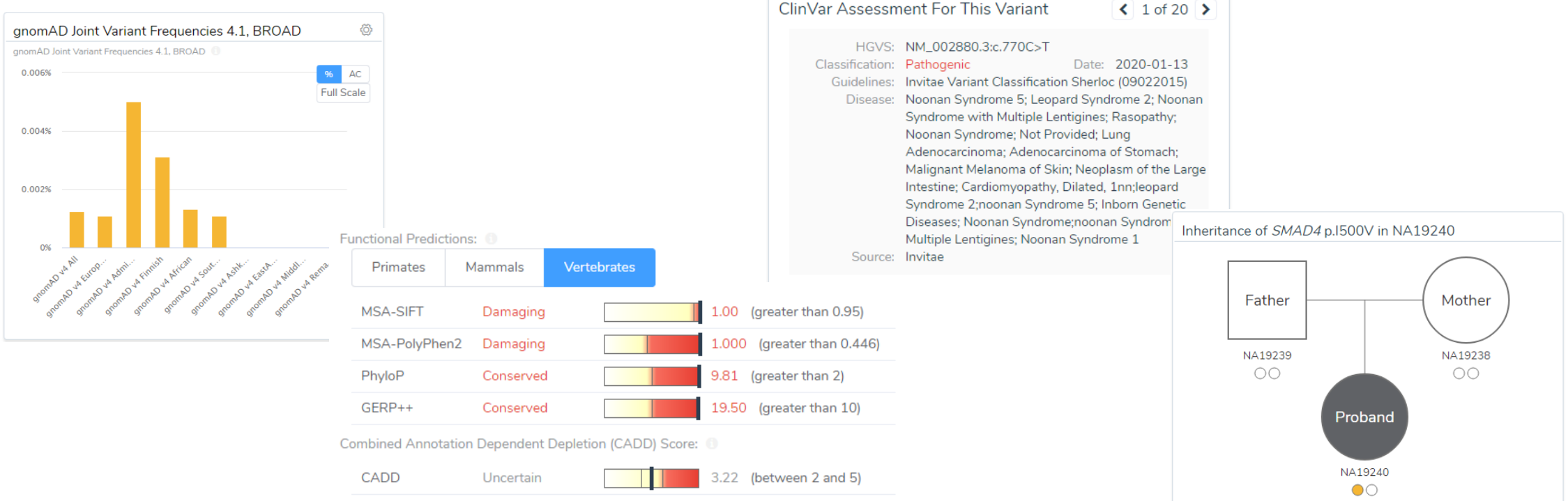
家族歴を利用した評価



- 両親には存在せず、罹患者のみに存在するバリアント（de Novo バリアント）の場合は、病的と評価される
- 両親がバリアントをもたないことを確認済みの場合は、病原性が強いと判定されるが、未確認の場合は中程度の評価となる

評価ワークフロー

評価対象のバリアントを入力



各種データベースと自動的に照合

評価スコアを算出

評価の実行方法

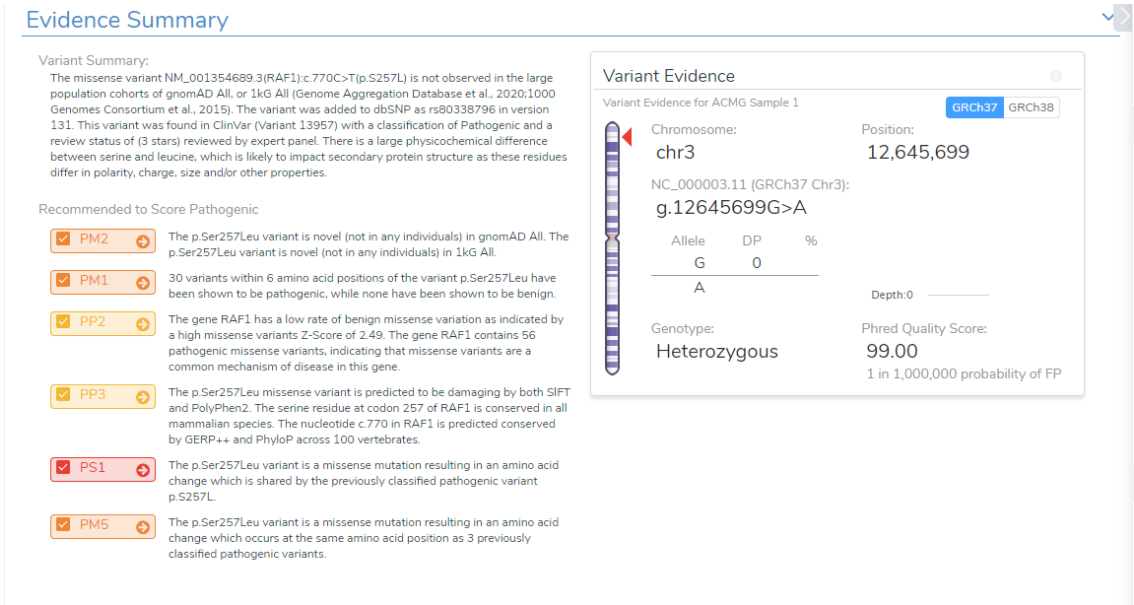
■ ACMG Sample Classifier

- 全バリアント（VCFファイル）へ、評価結果と関連データを自動でアノテーション付け
- 詳細データの確認や評価スコアの調整などは不可

Chr:Pos	Ref/Alt	Zygosity	DP	Gene Name	HGVS pDot	Sequence Ontology	Classification	ACMG Classification Criteria
7:107824678	A/G	Homozygous Variant	2	NRCAM	?	splice_region_variant	Benign	BA1,BS2,BP4,BP7
7:117144425	-/A	Heterozygous	15	CFTR	?	splice_region_variant	Likely Benign	PM2,BP4,BP7
7:117176569	GATT/-	Heterozygous	4	CFTR	?	splice_region_variant	Benign	BA1,BP7,BP6
7:117180144	T/-	Heterozygous	5	CFTR	?	splice_region_variant	Likely Benign	PM2,BP4,BP7
7:117199533	G/A	Homozygous Variant	27	CFTR	NP_000483.3:p.Val470Met	missense_variant	Benign	BA1,BS2,BP6
7:117227860	G/A	Heterozygous	12	CFTR	NP_000483.3:p.Gly551Asp	missense_variant	Likely Pathogenic	PM1,PS1,PM5,PP3
7:117242922	G/-	Heterozygous	8	CFTR	?	splice_region_variant	Pathogenic	PM2,PVS1,PP3

■ VSClinical Variant Dashbord

- バリアントフィルタリングによって抽出された、少数のバリアントのみに対して評価を行う
- 評価に用いるデータベースや文献などで根拠を確認しながらの評価が可能
- 評価スコアの調整なども可能



ACMG Sample Classifer

■ **ACMG Sample Classifierから実行する場合**

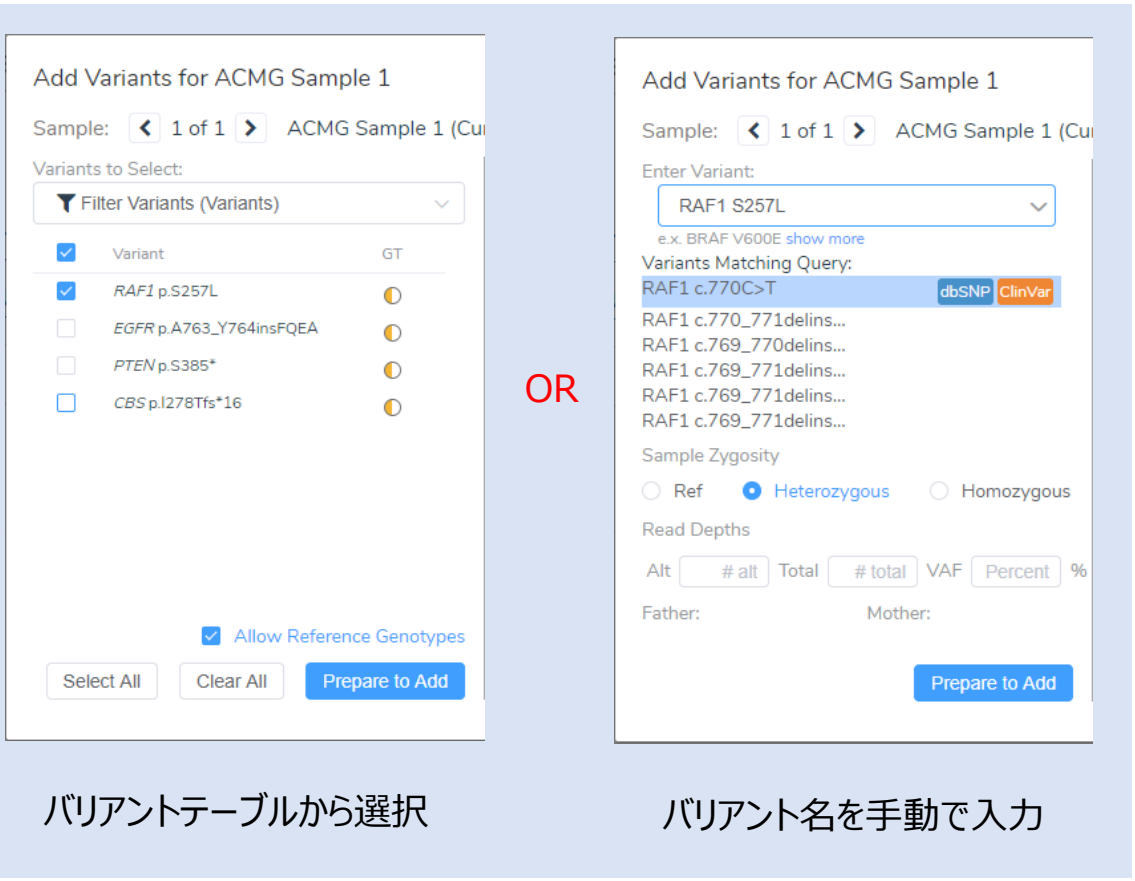
- ACMG Sample Classifierを実行すると、バリアントリストにACMG評価結果の情報がアノテーションとして付加される
- 全バリアントより、任意の評価結果（例：Pathogenicのみ）のバリアントのフィルタリングが可能

Variant Info		HD200-rep1			
Chr:Pos	Ref/Alt	Variant Allele Fraction	Read Depths (DP)	Genotype Qualities (GQ)	Zygosity
17:36097775	G/A	0.780488	41	35	Heterozygous
17:36098040	G/A	0.589091	275	574	Heterozygous
17:36101156	T/C	0.560897	312	747	Heterozygous
17:41234451	G/A	0.241218	427	?	Heterozygous
17:47436514	G/A	0.287129	101	81	Heterozygous
17:56773834	T/C	0.282609	92	69	Heterozygous
17:59761706	A/G	0.666667	84	125	Heterozygous
17:59820068	G/A	0.684211	57	91	Heterozygous
17:59820170	C/A	0.69281	153	131	Heterozygous
17:59857275	C/A	0.725	80	62	Heterozygous

ACMG Sample Classifierによるアノテーション付け

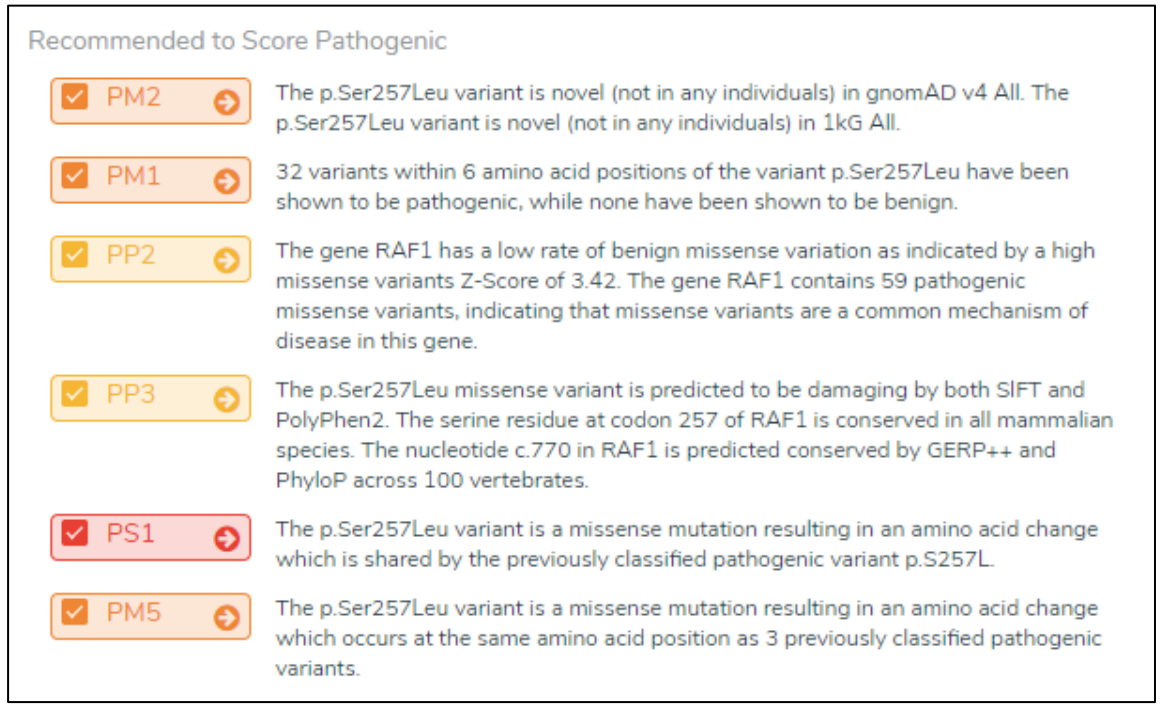
Variant Info		HD200-rep1				ACMG Sample Classifier for HD200-rep1				
Chr:Pos	Ref/Alt	Variant Allele Fraction	Read Depths (DP)	Genotype Qualities (GQ)	Zygosity	Gene Name	Gene Inheritance	Transcript Name	ACMG Classification Criteria	Classification
17:36097775	G/A	0.780488	41	35	Heterozygous	HNF1B	Dominant	NM_000458.4	BA1,BS2,BP7	Benign
17:36098040	G/A	0.589091	275	574	Heterozygous	HNF1B	Dominant	NM_000458.4	BA1,BS2,BP7,BP6	Benign
17:36101156	T/C	0.560897	312	747	Heterozygous	HNF1B	Dominant	NM_000458.4	BA1,BS2,BP7	Benign
17:41234451	G/A	0.241218	427	?	Heterozygous	BRCA1	Dominant	NM_007294.4	PM2,PVS1,PP5	Pathogenic
17:47436514	G/A	0.287129	101	81	Heterozygous	ZNF652	Default (Recessive)	NM_001145365.3	BA1,BS2,BP7	Benign
17:56773834	T/C	0.282609	92	69	Heterozygous	RAD51C	Recessive	NM_058216.3	BA1,BS2,BP7,BP6	Benign
17:59761706	A/G	0.666667	84	125	Heterozygous	BRIP1	Dominant	NM_032043.3	BA1,BS2,BP7,BP6	Benign
17:59820068	G/A	0.684211	57	91	Heterozygous	BRIP1	Dominant	NM_032043.3	BA1,BS2,BP7,BP6	Benign
17:59820170	C/A	0.69281	153	131	Heterozygous	BRIP1	Dominant	NM_032043.3	BA1,BS2,BP7,BP6	Benign
17:59857275	C/A	0.725	80	62	Heterozygous	BRIP1	Dominant	NM_032043.3	BA1,BS2,BP7	Benign

VSClinical Variant Dashbord



バリアントテーブルから選択

バリアント名を手動で入力

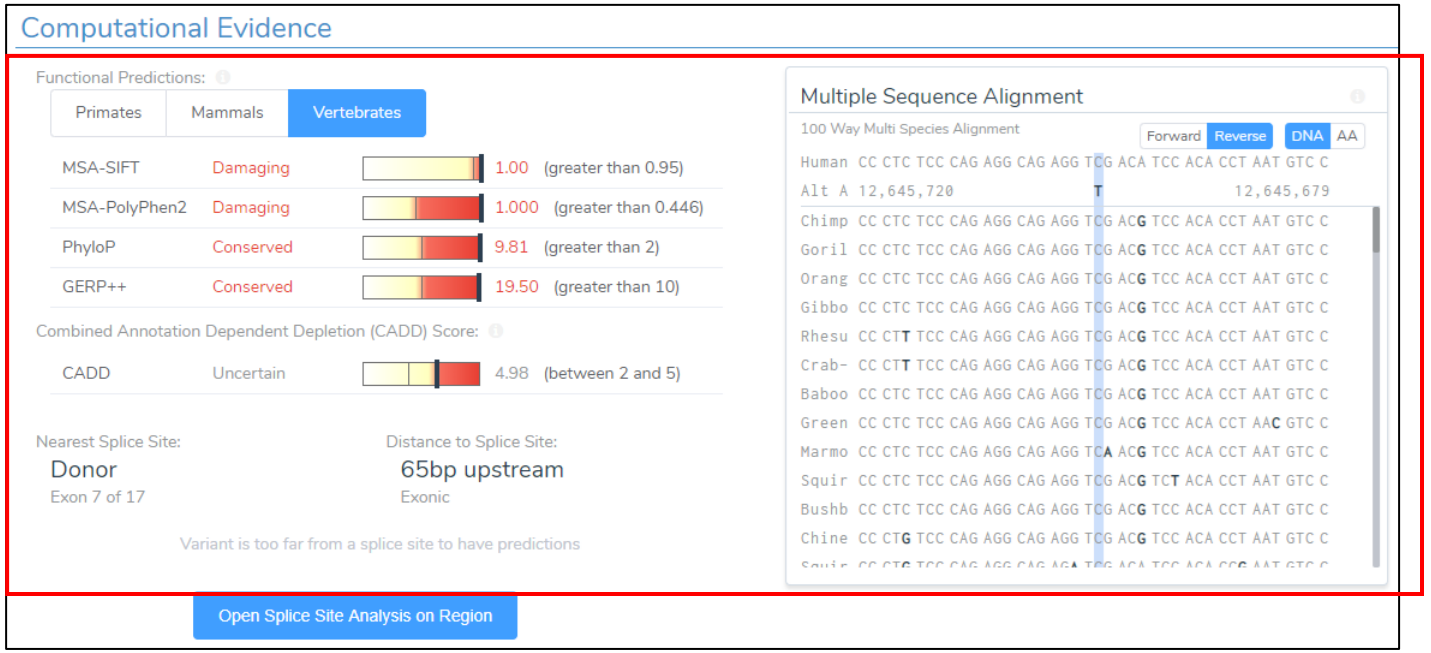


バリアントに該当する評価項目と結果を自動検出

■ **VSClinical Variant Dashbordから実行する場合**

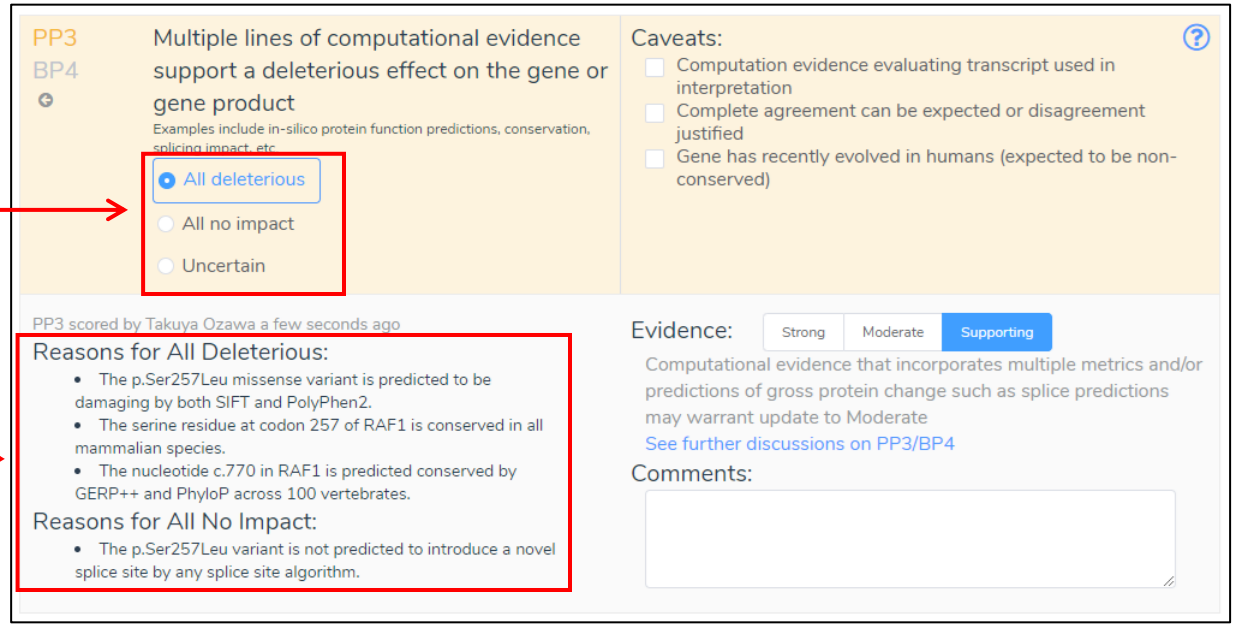
- 評価に用いるバリアントは、VCFファイルよりインポートしたバリアントテーブルより選択するか、キーボードでバリアント名を直接入力することも可能
- バリアントを指定すると、該当する評価項目が自動で検出され、評価結果のテキストとともに表示される

VSClinical Variant Dashbord



← データベースの照合結果

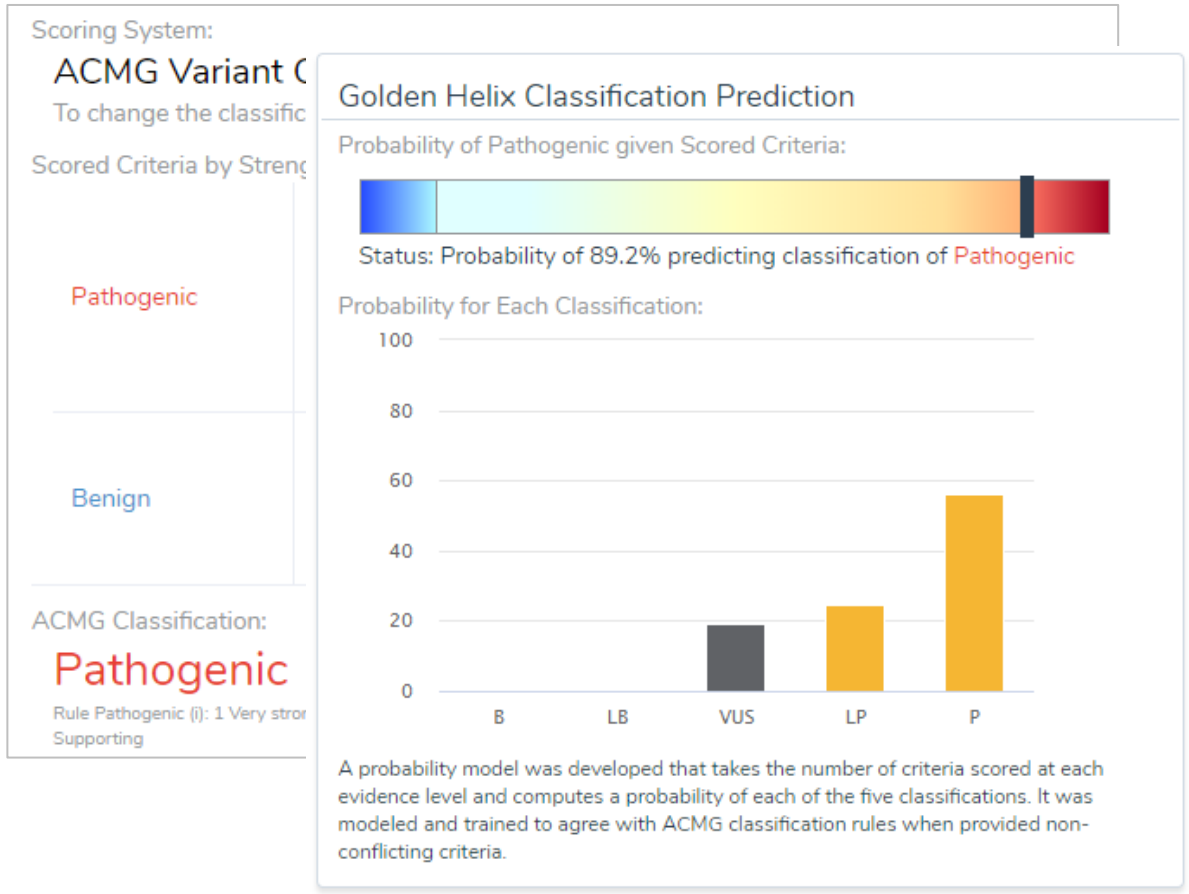
評価結果 →



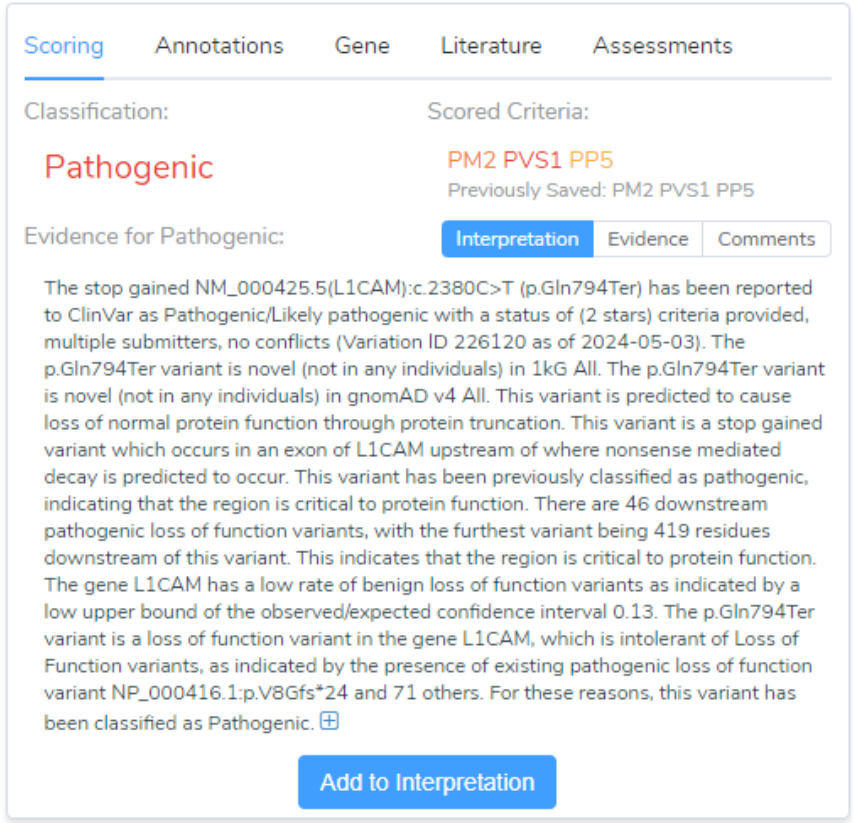
評価の根拠 →

- 各評価項目をクリックすると、データベースとの照合結果と評価結果を確認可能
- 評価を下した根拠などの情報も、同時に表示される
- 評価結果は手動で編集することも可能

VSClinical Variant Dashbord



評価結果サマリー



レポート用テキスト

- 評価結果の編集の完了後、各項目が病原性の可能性を示唆するスコアや良性の可能性を示唆するスコア、最終的な判定結果をまとめたサマリーを表示

- 同時にレポートに記載される臨床的解釈のテキストも出力

お問い合わせ先：フィルジェン株式会社

TEL: 052-624-4388 (9:00～18：00)

FAX: 052-624-4389

E-mail: support@filgen.jp