NGS変異データの三次解析

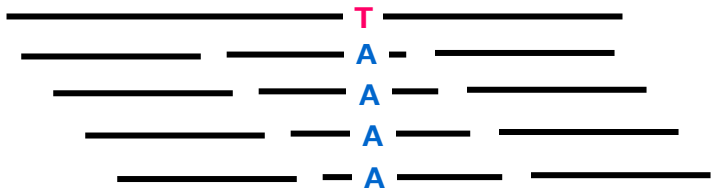
フィルジェン株式会社　バイオインフォマティクス部
(biosupport@filgen.jp)

はじめに

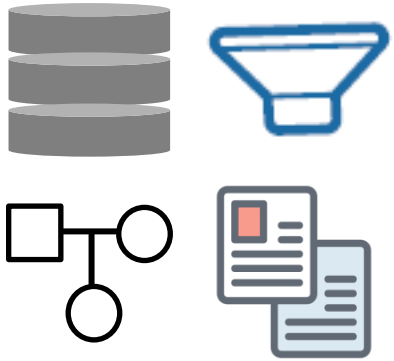
シークエンス → **二次解析** → **三次解析**

✓ ベースコール

✓ リードマッピング
✓ バリアントコール



✓ アノテーション付け / フィルタリング
✓ トリオ解析 / 集団解析
✓ 各種データのビジュアライゼーション
✓ レポート作成 …など



- 次世代シークエンサーによる変異解析では、データ作成の基本となる二次解析に加え、おもにデータの解釈を行うための三次解析を行うことが必要

- 三次解析では、変異情報データベースのような外部データリソースや、サンプルの表現型情報・家系情報などが必要

Golden Helix ソフトウェア製品ラインナップ

Filgen® biosciences & nanosciences

SNP & VARIATION SUITE

- GWAS & SNP Analysis
- Large-N DNA-Seq Analysis
- Genomic Prediction
- Copy Number Analysis
- RNA-Seq Analysis

varSEQ®

- Cancer & Rare Disease Diagnostics
- ACMG & AMP Guidelines Workflow
- CNV Calling
- Clinical Reporting
- High-throughput NGS Testing

- Golden Helix社では、遺伝統計解析ソフトウェア「SNP & Variation Suite」と、クリニカルシークエンス解析ソフトウェア「VarSeq®」の2種類のソフトウェアパッケージを販売
- 医学・生物学研究や、家畜や作物の品種改良などの農学研究、さらに疾患の診断や最適な治療オプションの決定における医療分野、遺伝学的解析などの教育現場などで利用される

VarSeq®

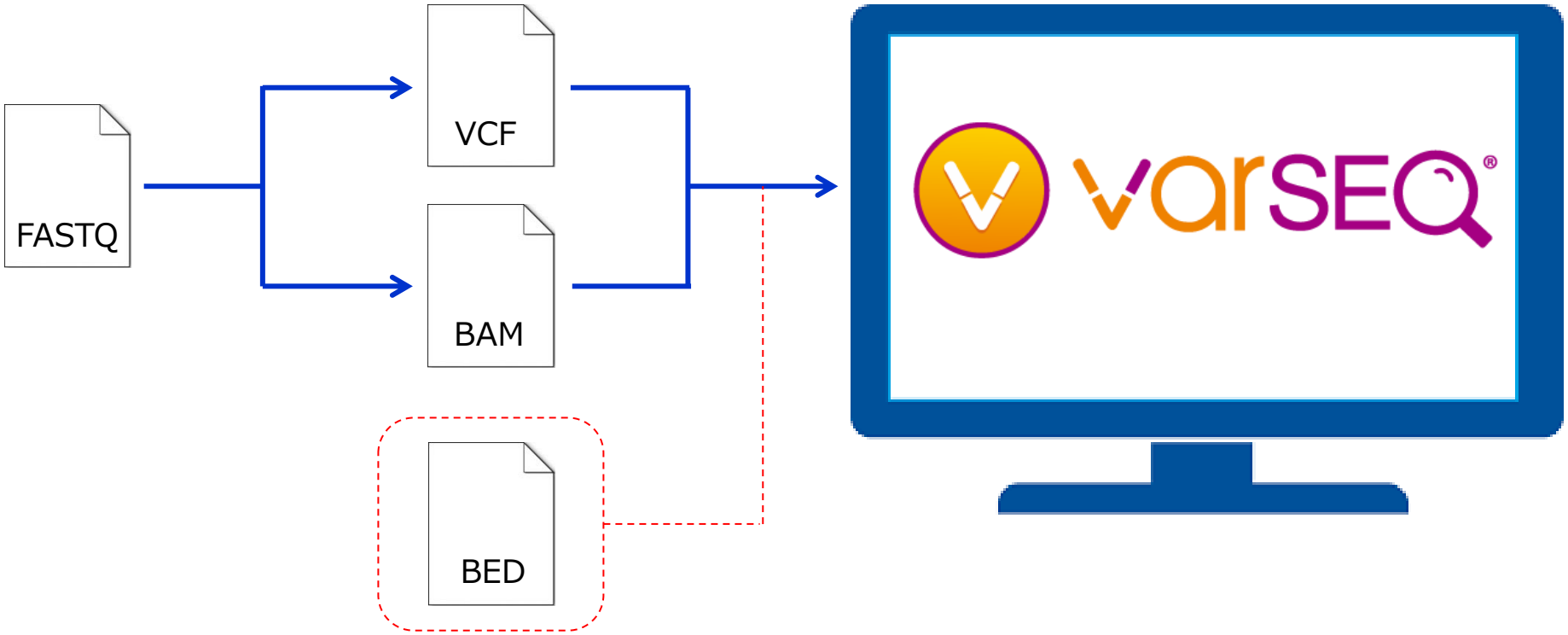
- キュレーションされた様々なデータリソースを使用し、変異データへアノテーション付けを実行
  - RefSeq Genes
  - dbSNP
  - ClinVar
  - CIVic
  - ICGC / TCGA
  - PhrmaGKB
  - BRCA Exchange
  - 1000 Genomes
  - NHLBI 6500 Exomes
  - ExAC Variant
  - gnomAD Exomes/Genomes
  - dbNSFP (SIFT, PolyPhen…)
  - 各種遺伝子パネルのターゲットデータ

  …など
- VCFファイルに含まれる変異データから、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
- ゲノムブラウザーを搭載し、BAMファイルデータや各種アノテーションデータをグラフ表示
- 有償アドオンによる機能拡張で、CNVコールやレポート作成、パイプライン機能などが利用可能

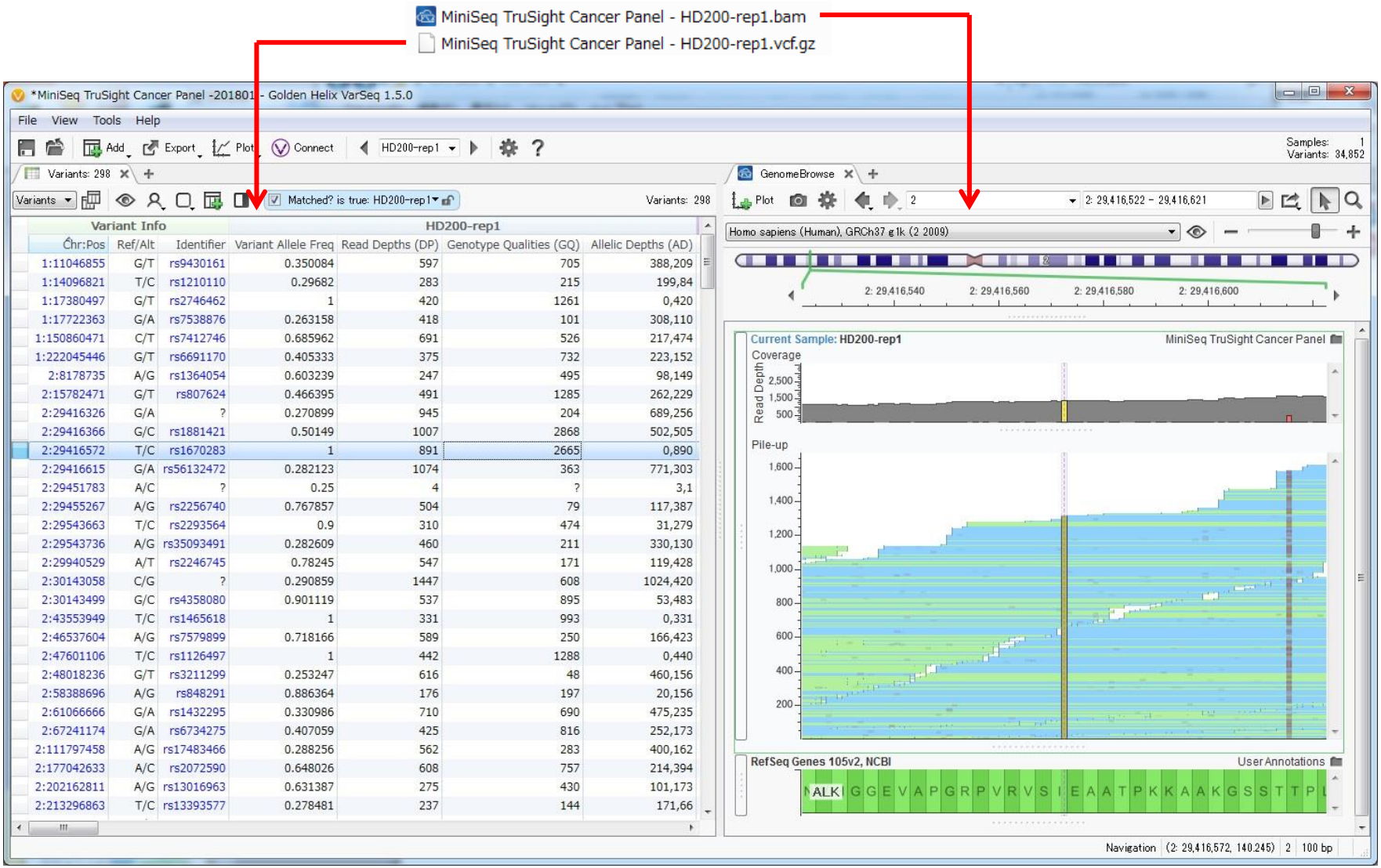
解析使用ファイル

- VarSeq®では、二次解析によって取得した、サンプルごとの変異データ（VCFファイル）が必須で、オプションでリード配列のマッピングデータ（BAMファイル）を使用
- 遺伝子パネルやエクソーム解析の場合は、オプションでゲノム上のターゲット領域データ（BEDファイル）も使用可能

*一部のパネルのターゲット領域データは、ソフトウェア上からダウンロードが可能



VarSeq®メイン画面



	Variant Info			HD200-rep1			
	Chr:Pos	Ref/Alt	Identifier	Variant Allele Freq	Read Depths (DP)	Genotype Qualities (GQ)	Allelic Depths (AD)
	1:11046855	G/T	rs9430161	0.350084	597	705	388,209
	1:14096821	T/C	rs1210110	0.29682	283	215	199,84
	1:17380497	G/T	rs2746462	1	420	1261	0,420
	1:17722363	G/A	rs7538876	0.263158	418	101	308,110
	1:150860471	C/T	rs7412746	0.685962	691	526	217,474
	1:222045446	G/T	rs6691170	0.405333	375	732	223,152
	2:8178735	A/G	rs1364054	0.603239	247	495	98,149
	2:15782471	G/T	rs807624	0.466395	491	1285	262,229
	2:29416326	G/A	?	0.270899	945	204	689,256
	2:29416366	G/C	rs1881421	0.50149	1007	2868	502,505
	2:29416572	T/C	rs1670283	1	891	2665	0,890
	2:29416615	G/A	rs56132472	0.282123	1074	363	771,303
	2:29451783	A/C	?	0.25	4	?	3,1
	2:29455267	A/G	rs2256740	0.767857	504	79	117,387
	2:29543663	T/C	rs2293564	0.9	310	474	31,279
	2:29543736	A/G	rs35093491	0.282609	460	211	330,130
	2:29940529	A/T	rs2246745	0.78245	547	171	119,428
	2:30143058	C/G	?	0.290859	1447	608	1024,420
	2:30143499	G/C	rs4358080	0.901119	537	895	53,483
	2:43553949	T/C	rs1465618	1	331	993	0,331
	2:46537604	A/G	rs7579899	0.718166	589	250	166,423
	2:47601106	T/C	rs1126497	1	442	1288	0,440
	2:48018236	G/T	rs3211299	0.253247	616	48	460,156
	2:58388696	A/G	rs848291	0.886364	176	197	20,156
	2:61066666	G/A	rs1432295	0.330986	710	690	475,235
	2:67241174	G/A	rs6734275	0.407059	425	816	252,173
	2:111797458	A/G	rs17483466	0.288256	562	283	400,162
	2:177042633	A/C	rs2072590	0.648026	608	757	214,394
	2:202162811	A/G	rs13016963	0.631387	275	430	101,173
	2:213296863	T/C	rs13393577	0.278481	237	144	171,66

- VCFファイルとBAMファイルのインポートを行うと、それぞれ変異テーブルとゲノムブラウザーで表示される

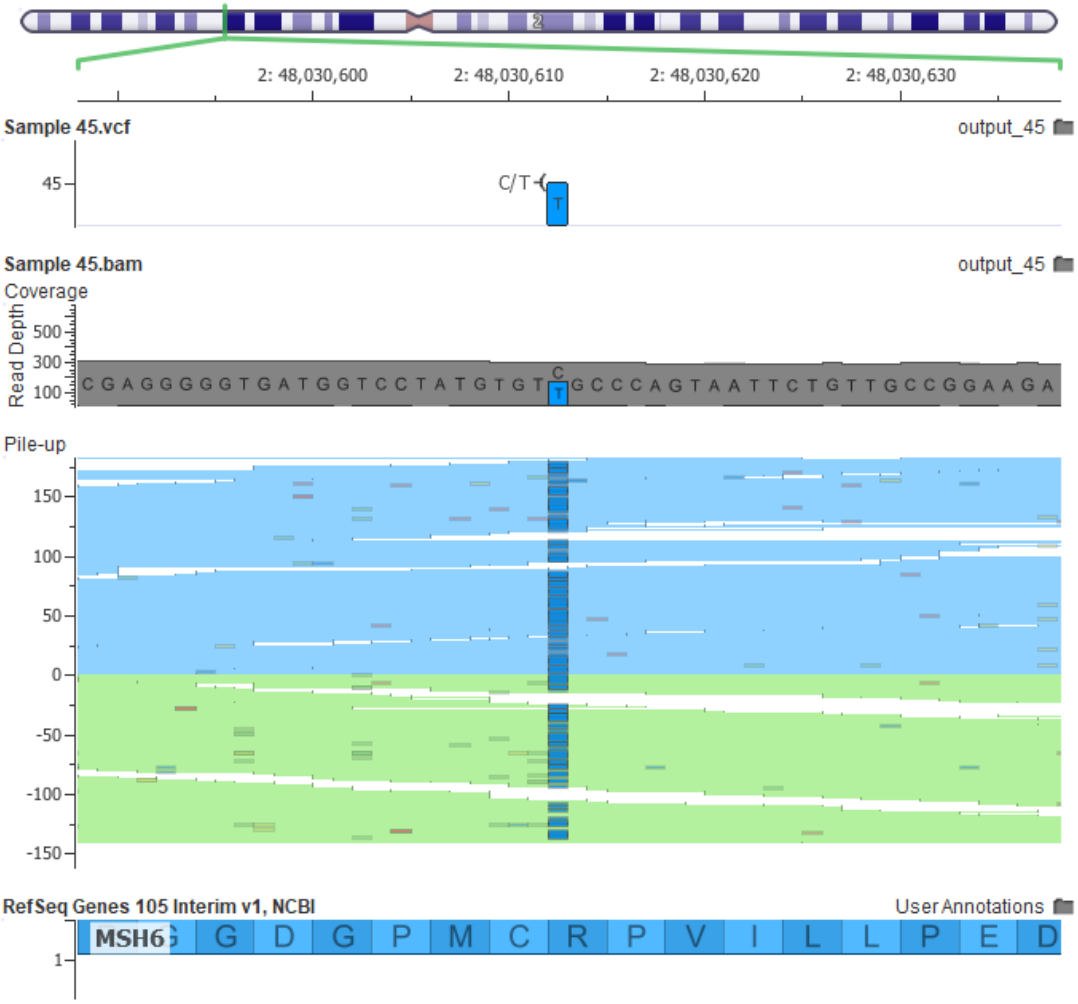
NGS三次解析

クオリティーチェック
- Ti/Tv、Homo/Hetero Ratio、Call Rate
- 親縁チェック、性別チェック
- 遺伝子パネルのターゲットカバレッジ計算

アノテーション
- 遺伝子アノテーション
- 臨床情報データベース
- アレル頻度データベース
- 機能予測アノテーション
- 疾患別の関連遺伝子ランキング

サンプル間比較
- トリオ解析
- 腫瘍-正常サンプルのペア解析
- 集団サンプル解析

ビジュアライゼーション
- ゲノムブラウザー
- 各種データの集計グラフ



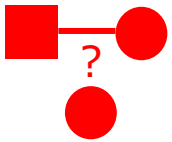
クオリティーチェック

Sample Info				Sample Statistics Tr				
Samples	Mother	Father	Affected?	Ti/Tv	Coding Ti/Tv	Het Ratio	Hom Ratio	Hemi Ratio
HG02024	HG02025	HG02026	Affected	2.26071017826337	2.99105754276827	0.601626096970178	0.388419962972548	0.011525632371044
HG02025	?	?	Unaffected	2.23023689970388	2.9485542621104	0.598907407739644	0.392103820471111	0.0100378550027808
HG02026	?	?	Unaffected	2.20932700711699	2.92923823749066	0.600705075493815	0.390026550300991	0.0101737320013297
HG02046	HG02025	HG02026	Affected	2.2631533365089	2.86044762083882	0.592015342971244	0.399912443909379	0.00914023524270784

Ti/Tv、Homo/Hetero Ratioなど

Sample Statistics Trio Analysis 2AF					
Mother Kinship Coefficient	Father Kinship Coefficient	Mother/Father Kinship Coefficient	Kinship Errors	Chrom X Het Ratio	Inferred Gender
0.238881427623149	0.235499422105968	0.181719117775259	?	0.592132505175983	Female
?	?	?	?	0.579927338782925	Female
?	?	?	?	0.198375084631009	Male
0.153511836790346	0.150444258172674	0.181719117775259	Specified mother and father are not related to proband.	0.602579132473622	Female

親縁チェック



性別チェック

- サンプル情報テーブルにて、Ti/Tv、Homo/Hetero Ratio、親縁チェック、性別チェックなどの結果を確認

クオリティーチェック

Filgen® biosciences & nanosciences

カバレッジテーブル

Coverage Region Info		HD200-rep1									
Region	Name	Counted Bases	Mean Depth	Min Depth	Max Depth	% 1x	% 20x	% 100x	% 500x	Mean MQ	Reads in Region
1:10385471-10385472	rs17401966.chr1.10385470.10385472	2	503	502	504	100	100	100	100	60	506
1:11046855-11046856	rs9430161.chr1.11046854.11046856	2	623	620	626	100	100	100	100	60	627
1:14096821-14096822	rs1210110.chr1.14096820.14096822	2	287	285	289	100	100	100	0	60	289
1:14804874-14804875	rs7555566.chr1.14804873.14804875	2	573	573	573	100	100	100	100	60	574
1:17345376-17345454	SDHB.chr1.17345375.17345454	79	634.367	536	721	100	100	100	100	60	1012
1:17349103-17349226	SDHB.chr1.17349102.17349226	124	368.605	238	451	100	100	100	0	60	666
1:17350468-17350570	SDHB.chr1.17350467.17350570	103	476.282	276	709	100	100	100	33.9806	60	1002
1:17354244-17354361	SDHB.chr1.17354243.17354361	118	592.28	326	681	100	100	100	83.0508	60	1007
1:17355095-17355232	SDHB.chr1.17355094.17355232	138	809.688	556	1006	100	100	100	100	60	1591
1:17359555-17359641	SDHB.chr1.17359554.17359641	87	630.885	531	723	100	100	100	100	60	1022
1:17371256-17371384	SDHB.chr1.17371255.17371384	129	505.016	340	623	100	100	100	55.0388	60	959
1:17380443-17380515	SDHB.chr1.17380442.17380515	73	572.397	443	727	100	100	100	82.1918	60	1111
1:17722363-17722364	rs7538876.chr1.17722362.17722364	2	433	433	433	100	100	100	0	60	433
1:45794978-45795110	MUTYH.chr1.45794977.45795110	133	373.053	145	473	100	100	100	0	60	661
1:45796188-45796230	MUTYH.chr1.45796187.45796230	43	511.953	427	544	100	100	100	79.0698	60	618

サンプル情報テーブル

Sample Info		Coverage Statistics TruSight Cancer Amplicon Design 2013-02-05, Illumina								
Samples	Affected?	Sample Mean Depth	Sample Mean Forward Depth	Sample Mean Reverse Depth	Sample Mean Filtered Depth	Sample % 1x	Sample % 20x	Sample % 100x	Sample % 500x	Sample Errors
HD200-rep1	Affected	695.039	322.597	372.441	1.27973	99.9988	99.6033	98.5985	69.7414	?
HD200-rep2	Affected	746.986	347.222	399.764	1.44163	99.9988	99.6977	98.7399	75.4424	?
HD701-rep1	Affected	705.099	330.94	374.159	1.33995	99.9673	99.3353	97.9315	71.9847	?
HD701-rep2	Affected	557.102	261.195	295.907	0.988303	99.9271	99.0557	96.7308	55.9456	?
NA12878-rep1	Affected	356.256	171.281	184.975	0.790086	99.8507	98.6276	93.1217	21.4599	?
NA12878-rep2	Affected	383.213	184.191	199.022	0.824232	99.7833	98.7244	93.9457	26.8968	?

- 遺伝子パネル解析では、各ターゲットキャプチャー領域ごとのリードカバレッジと、サンプルごとのカバレッジデータを確認
- 計算にはBAMファイルと、ターゲット領域データ（BEDファイルなど）が必要

アノテーション一覧

Filgen® biosciences & nanosciences

Assemblies
- GRCh 37 (hg19)
- GRCh 38 (hg38)

Genes
- RefSeq
- Ensembl

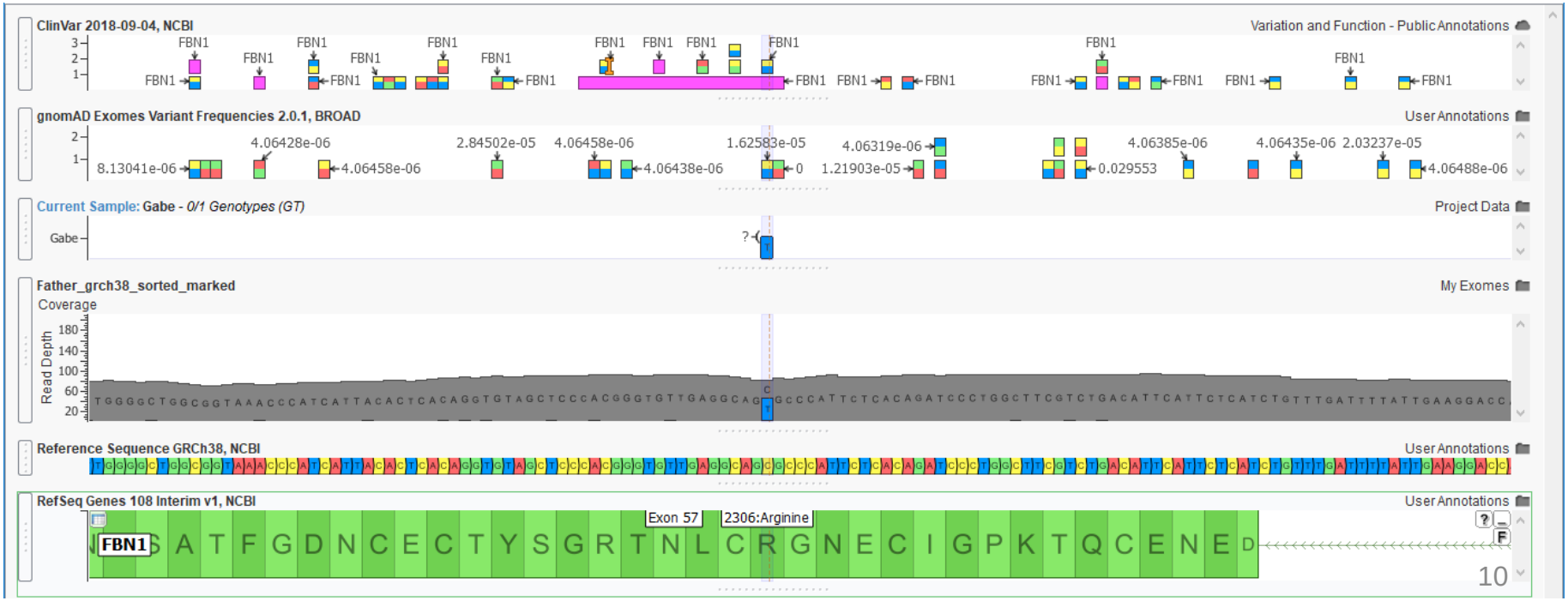
Clinical Annotations
- ClinVar
- OMIM
- PharmaGKB
- BRCA Exchange

Cancer Annotations
＊赤字は有償アドオン
- COSMIC
- CIViC
- Cancer Hotspots
- ICGC Simple Somatic Mutations
- TCGA Variants
- MSK Impact
- Precision Medicine Knowledgebase (PMKB)

Functional Annotations
- dbNSFP (SIFT, PolyPhen, MutationTaster...)
- REVEL Functional Prediction
- CADD

Population Catalogs
- dbSNP
- 1000 Genomes
- NHLBI 6500 Exomes
- ExAC Variant
- gnomAD Exomes/Genomes
- TOPMed
- UK10K

Targeted Panels
- TruSight
- Ion AmpliSeq



アノテーション付加

Variant Info		HD200-rep1		RefSeq Genes 105 Interim v1, NCBI			ClinVar 2019-07-01, NCBI				
Chr:Pos	Ref/Alt	VAF	DP	Gene Names	EffectCombined	HGVS p. (Clinically Relevant)	Classification	Review Status	Conditions	In COSMIC?	MutationIDUni…
2:29416326	G/A	0.270899	945	ALK	Missense	NP_004295.2:p.Pro1543Ser	?	?	?	True	COSM2941442
2:48018236	G/T	0.253247	616	MSH6	Missense	NP_000170.1:p.Ser144Ile	Benign	(3 Stars) Reviewed …	not provided,Lynch syndro…	True	COSM4988649
3:10138019	G/A	0.305785	847	FANCD2,FA…	Missense	NP_149075.2:p.Asp1350Asn,?	?	?	?	True	COSM2915697,…
4:55604693	C/A	0.34	300	KIT	Other	NP_000213.1:p.Ser967=	?	?	?	True	COSM3301473
5:112179431	C/T	0.280665	482	APC	Missense	NP_000029.2:p.Arg2714Cys	Uncertain Significance	(1 Stars) Criteria Pr…	Hereditary cancer-predispo…	True	COSM2991126
8:145739087	T/A	0.270751	1014	RECQL4	Missense	NP_004251.3:p.Thr690Ser	Uncertain Significance	(1 Stars) Criteria Pr…	Baller-Gerold syndrome	True	COSM2871905
9:98239971	C/T	0.298957	865	PTCH1	Missense	NP_000255.2:p.Cys454Tyr	?	?	?	True	COSM2733335,…
9:98278975	C/T	0.259594	1331	PTCH1	Missense	?	?	?	?	True	COSM199000,C…
9:100437823	A/G	0.285714	280	XPA	Other	NP_000371.1:p.Ile240=	?	?	?	True	COSM2793140
10:43604493	C/T	0.273738	1170	RET	Missense	NP_066124.1:p.Arg360Trp	?	?	?	True	COSM33011,C…
13:32912750	G/T	0.336232	345	BRCA2	Missense	NP_000050.2:p.Asp1420Tyr	Benign	(3 Stars) Reviewed …	not provided,Hereditary br…	True	COSM3736087,…
13:103515036	G/A	0.269939	978	BIVM-ERC…	Missense	NP_001191354.1:p.Ala967Thr,N…	?	?	?	True	COSM696103
15:89849327	G/A	0.2875	1200	FANCI	Missense	NP_001106849.1:p.Glu1147Lys	?	?	?	True	COSM2015328,…
16:2134450	G/A	0.286654	1064	TSC2	Other	NP_000539.2:p.Arg1409=	?	?	?	True	COSM4656201,…
16:3658433	C/A	0.28629	744	SLX4	Missense	NP_115820.2:p.Arg178Ile	?	?	?	True	COSM2919796
17:41234451	G/A	0.241218	427	BRCA1	LoF	NP_009225.1:p.Arg1443Ter	Pathogenic	(3 Stars) Reviewed …	Breast-ovarian cancer, fami…	True	COSM979730,C…

Variants: 16

Variants | In COSMIC? is true: HD200-rep1

17:41234451 - G/A (1bp sub)

Chr:Pos: 17:41234451

rs41293455

ClinVar Assessments 2019-07-01, NCBI

	1	2	3	4	5	6	7	8	9
Ref/Alt	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A
Variant ID	17675	17675	17675	17675	17675	17675	17675	17675	17675
Assessment Source	Counsyl	Michigan Medical Genetics Laboratories,University of Michigan	Baylor Miraca Genetics Laboratories,	Genetic Services Laboratory, University of Chicago	Fulgent Genetics	Integrated Genetics/Laboratory Corporation of America	OMIM	Laboratory for Molecular Medicine,Partners HealthCare Personalized Medicine	Genome Diagnostics Laboratory,University Medical Center Utrecht
Date	2014-10-02	2016-04-21	2017-02-23	2016-11-29	2017-05-18	2016-08-15	1994-12-01	2016-10-28	2014-10-08
Classification	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic
Conditions	Breast-Ovarian Cancer, Familial 1	Not Provided	Familial Cancer Of Breast	{Breast-Ovarian Cancer, Familial, 1}	Not Provided	Not Provided	Breast-Ovarian Cancer, Familial, Susceptibility	Not Provided	Not Provided

- アノテーション情報は変異データテーブルにそのまま追加され、データベースへのハイパーリンクもつく
- 変異データテーブル上の任意のデータをクリックすると、アノテーションの詳細情報も閲覧可能

アノテーション例

RefSeq

RefSeq Genes 105 Interim v3.1, NCBI						
Gene Names	Sequence Ontology (Combined)	Effect (Combined)	Nof4PredictedSplicingDisruptedCombined	Predicted Splicing Disrupted (Combined)	HGVS c. (Clinically Relevant)	HGVS p. (Clinically Relevant)
TM4SF4	3_prime_UTR_variant	Other	?	?	NM_004617.3:c.*37C>T	?
WWTR1	frameshift_variant	LoF	?	?	NM_015472.4:c.1199_1200insTTAA	NP_056287.1:p.?
COMMD2	synonymous_variant	Other	?	?	NM_016094.3:c.567T>C	NP_057178.2:p.Asn189=
COMMD2	missense_variant	Missense	?	?	NM_016094.3:c.337A>C	NP_057178.2:p.Ile113Leu
RNF13	intron_variant	Other	0 of 4 Predicted Splicing Disrupted	?	NM_007282.4:c.322-11_322-10delTT	?
RNF13	synonymous_variant	Other	1 of 4 Predicted Splicing Disrupted	NNSplice	NM_007282.4:c.789C>T	NP_009213.1:p.His263=
PFN2	intron_variant	Other	?	?	NM_002628.4:c.325+69A>G	?

ClinVar

ClinVar 2020-04-02, NCBI						
Ref/Alt	Variant ID	Classification	Clinical Significance	Review Status	RSID	Gene Names
A/G	189481	Pathogenic	Pathogenic	(1 Stars) Criteria Provided, Single Submitter	rs786204926	PTEN
G/A	377376	Likely Pathogenic	Pathogenic/Likely Pathogenic	(2 Stars) Criteria Provided, Multiple Submitters, No Conflicts	rs1057520208	PTEN
T/G	428239	Uncertain Significance	Uncertain Significance	(1 Stars) Criteria Provided, Single Submitter	rs1114167658	PTEN
T/C	821476	Likely Benign	Likely Benign	(1 Stars) Criteria Provided, Single Submitter	?	PTEN
G/A	584752	Uncertain Significance	Uncertain Significance	(1 Stars) Criteria Provided, Single Submitter	rs1564829780	PTEN
C/T	591438	Uncertain Significance	Uncertain Significance	(0 Stars) No Assertion Criteria Provided	rs1564829783	PTEN
A/G	135909	Likely Benign	Likely Benign	(2 Stars) Criteria Provided, Multiple Submitters, No Conflicts	rs587780710	PTEN

COSMIC

COSMIC Mutations 89, GHI										
In COSMIC?	Mutation ID	Mutation CDS	Mutation AA	Gene Name	Pubmed ID	Sample Count	Samples	Sample Primary Sites	Sample Tissue Types	Oncotree Tissue Type
True	COSM110753	c.1930C>T	p.Q644*	TET2	21828143,219…	2	1554939,1432…	Haematopoietic and ly…	Lymphoid,CNS/Brain	Lymphoid,CNS/Brain
True	COSM1673411	c.2605C>T	p.Q869*	CUX1	23856246,247…	2	1998442,2301…	Large intestine,Large in…	Bowel,Bowel	Bowel
True	COSM7409795	c.3564G>A	p.W1188*	CUX1	30304655	1	2776816	Haematopoietic and ly…	Lymphoid	Lymphoid
True	COSM476	c.1799T>A	p.V600E	BRAF	19919912,206…	28376	1513212,2579…	Skin,Thyroid,Skin,Skin,…	Skin,Thyroid,Skin,Skin…	Thyroid,Skin,Bowel,CN…
True	COSM1738109	c.482-1G>A	p.?	RAD21	23955599	1	2088799	Haematopoietic and ly…	Lymphoid	Lymphoid
True	COSM1140132,COSM532	c.38G>A,c.38G>A	p.G13D,p.G13D	KRAS	19679400,284…	5730	1774048,1774…	Large intestine,Haemat…	Bowel,Lymphoid,Bow…	Bowel,Lymphoid,Lung,…
True	COSM1135366,COSM521	c.35G>A,c.35G>A	p.G12D,p.G12D	KRAS	20926413,235…	15340	1511227,1511…	Large intestine,Oesoph…	Bowel,Esophagus/Sto…	Bowel,Pancreas,Lung,B…

アノテーション例

1000 Genomes

1kG Phase3 – Variant Frequencies 5a with Genotype Counts, GHI									
Ref/Alt	Identifier	Allele Counts	Allele Frequencies	AMR – Allele Counts	AMR – Allele Frequencies	AFR – Allele Counts	AFR – Allele Frequencies	SAS – Allele Counts	SAS – Allele Frequencies
A/G	rs1412829	923	0.184305	131	0.188761	18	0.0136157	265	0.270961
G/A	rs4977756	3564	0.711661	550	0.792507	899	0.68003	710	0.725971
G/A	rs357564	1987	0.396765	282	0.40634	298	0.225416	461	0.47137
?	?	?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?	?	?
A/G	rs965513	4005	0.79972	484	0.697406	1181	0.893343	795	0.812883

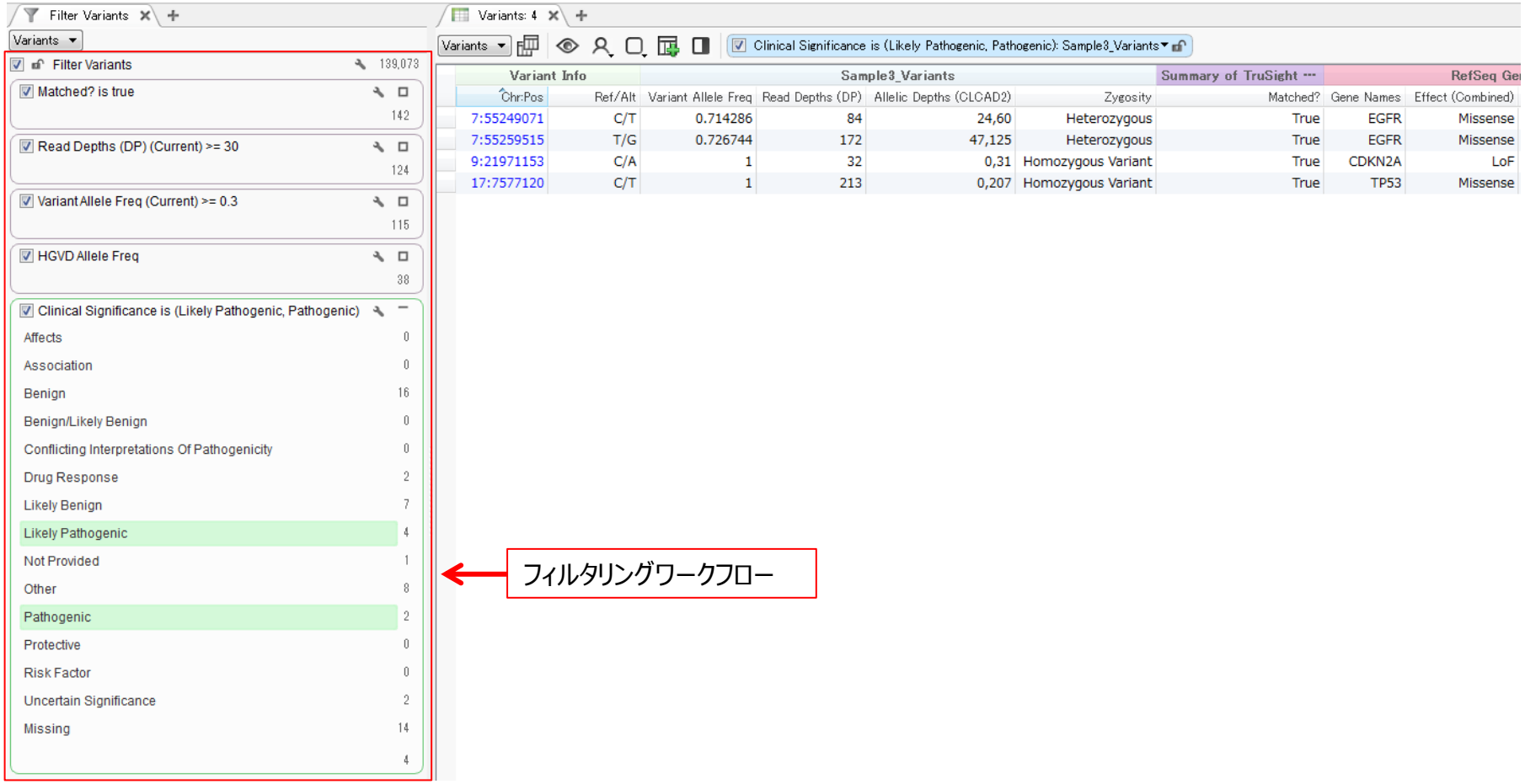
dbNSFP

dbNSFP Functional Prediction Voting						
N of 6 Predicted Tolerated	N of 6 Predicted Damaging	SIFT Pred (C)	Polyphen2 HVAR Pred (C)	MutationTaster Pred (C)	MutationAssessor Pred (C)	FATHMM Pred (C)
0 of 6 Predicted as Tolerated	6 of 6 Predicted as Damaging	Damaging	Possibly damaging	Damaging	Predicted functional (medium)	Damaging
1 of 6 Predicted as Tolerated	5 of 6 Predicted as Damaging	Damaging	Probably damaging	Damaging	Predicted functional (medium)	Tolerated
0 of 6 Predicted as Tolerated	6 of 6 Predicted as Damaging	Damaging	Possibly damaging	Damaging	Predicted functional (medium)	Damaging
0 of 6 Predicted as Tolerated	6 of 6 Predicted as Damaging	Damaging	Probably damaging	Damaging	Predicted functional (medium)	Damaging
2 of 6 Predicted as Tolerated	4 of 6 Predicted as Damaging	Damaging	Possibly damaging	Damaging	Predicted non-functional (neutral)	Tolerated

PhoRank

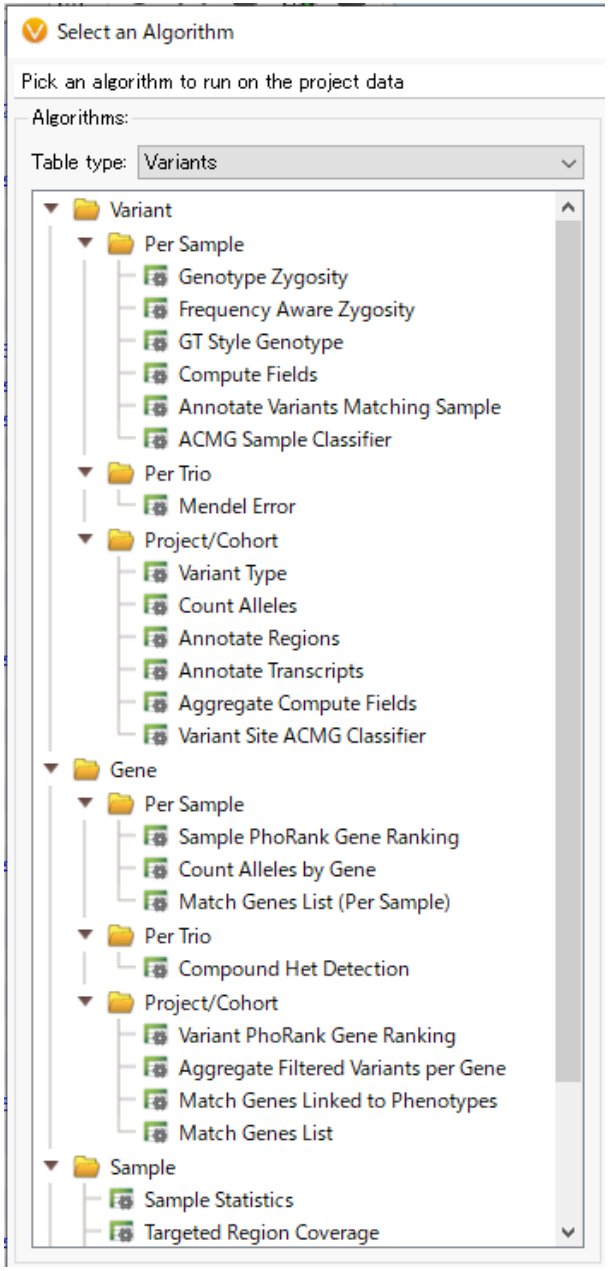
Variant Gene Info	Muscular Dystrophy PhoRank		
Gene Names	Gene Rank	Gene Score	Path
DYSF	0.984652	0.746612	DYSF / HP:0…
B3GALNT2	0.965761	0.742633	B3GALNT2 / …
PLEC	0.955136	0.738786	PLEC / HP:0…
DMD	0.955136	0.738786	DMD / HP:00…
TTN	0.943329	0.734573	TTN / HP:00…

フィルタリングワークフロー



- 付加したアノテーション情報などを用いて、変異データに対する任意のフィルタリングワークフローを作成
- ワークフローの各項目ごとに、フィルタリング結果の変異データテーブルを確認可能で、項目を変更するとフィルタリング結果の変異データテーブルも自動的にアップデートされる

サンプル比較用の解析アルゴリズム



Genotype Zygosity

- 各変異ごとの接合性情報（Homozygous, Heterozygousなど）によるアノテーション付けの実行

Mendel Error

- 各変異ごとのメンデル遺伝情報（親からの遺伝、De Novo変異など）によるアノテーション付けの実行

Count Alleles

- 全サンプルあるいはサンプルグループごとに、プロジェクトデータ内のアレル数や頻度情報などのアノテーション付けの実行

Compound Het Detection

- 複合ヘテロ接合体の検出

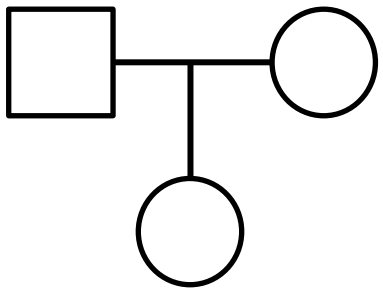
トリオ解析

Variant Info		Proband (HG02024)					Mother (HG02025)					Father (HG02026)				
Chr:Pos	Ref/Alt	VAF	DP	Mendel Error	Inherited From	Zygosity	VAF	DP	Mendel Error	Inherited From	Zygosity	VAF	DP	Mendel Error	Inherited From	Zygosity
1:1225553	C/T	0.666667	12	Transmitted	Father	Heterozygous	?	?	?	?	?	0.483871	31	?	?	Heterozygous
1:1225558	T/C	0.214286	14	de Novo Allele	?	Heterozygous	?	?	?	?	?	?	?	?	?	?
1:1225563	-/CC	0.25	12	de Novo Allele	?	Heterozygous	?	?	?	?	?	?	?	?	?	?
1:1225579	G/C	0.5	24	Transmitted	Mother	Heterozygous	1	73	?	?	Homozygous Variant	0.438596	57	?	?	Heterozygous
1:1225612	G/A	1	71	Transmitted	Both	Homozygous Variant	0.398496	133	?	?	Heterozygous	0.472222	109	?	?	Heterozygous
1:1225641	C/A	1	92	Transmitted	Both	Homozygous Variant	0.378698	169	?	?	Heterozygous	1	127	?	?	Homozygous Variant
1:1225959	C/G	0.4375	144	Transmitted	Father	Heterozygous	0.461279	298	?	?	Heterozygous	1	241	?	?	Homozygous Variant
1:1226102	G/A	?	?	Untransmitted	?	?	0.433526	173	?	?	Heterozygous	0.439716	143	?	?	Heterozygous
1:1226221	A/G	1	28	Transmitted	Both	Homozygous Variant	1	70	?	?	Homozygous Variant	1	70	?	?	Homozygous Variant

Mendel Error (Current) is de Novo Allele

MIE	10
Transmitted	273
Untransmitted	1
de Novo Allele	152
Missing	0
0 0	152

✓ de Novo変異



Zygosity (Current) is Homozygous Variant

Heterozygous	415
Homozygous Variant	20
Reference	1
Missing	0
0 0	20

Zygosity (Mother) is Heterozygous

Heterozygous	10
Homozygous Variant	4
Reference	0
Missing	6
0 0	10

Zygosity (Father) is Heterozygous

Heterozygous	5
Homozygous Variant	2
Reference	0
Missing	3
0 0	5

✓ 劣性ホモ変異

Zygosity (Current) is Heterozygous

Heterozygous	186
Homozygous Variant	124
Reference	0
Missing	0
	186

Zygosity (Mother) is Heterozygous

Heterozygous	109
Homozygous Variant	19
Reference	2
Missing	56
	109

Zygosity (Father) is (Reference, missing)

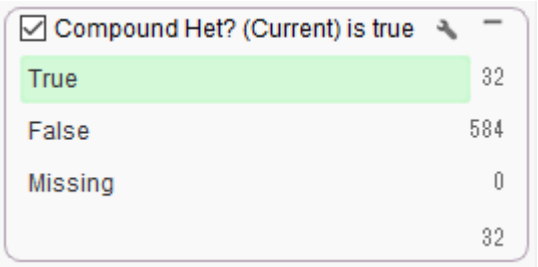
Heterozygous	69
Homozygous Variant	10
Reference	1
Missing	29
	30

✓ 優性ヘテロ変異

トリオ解析

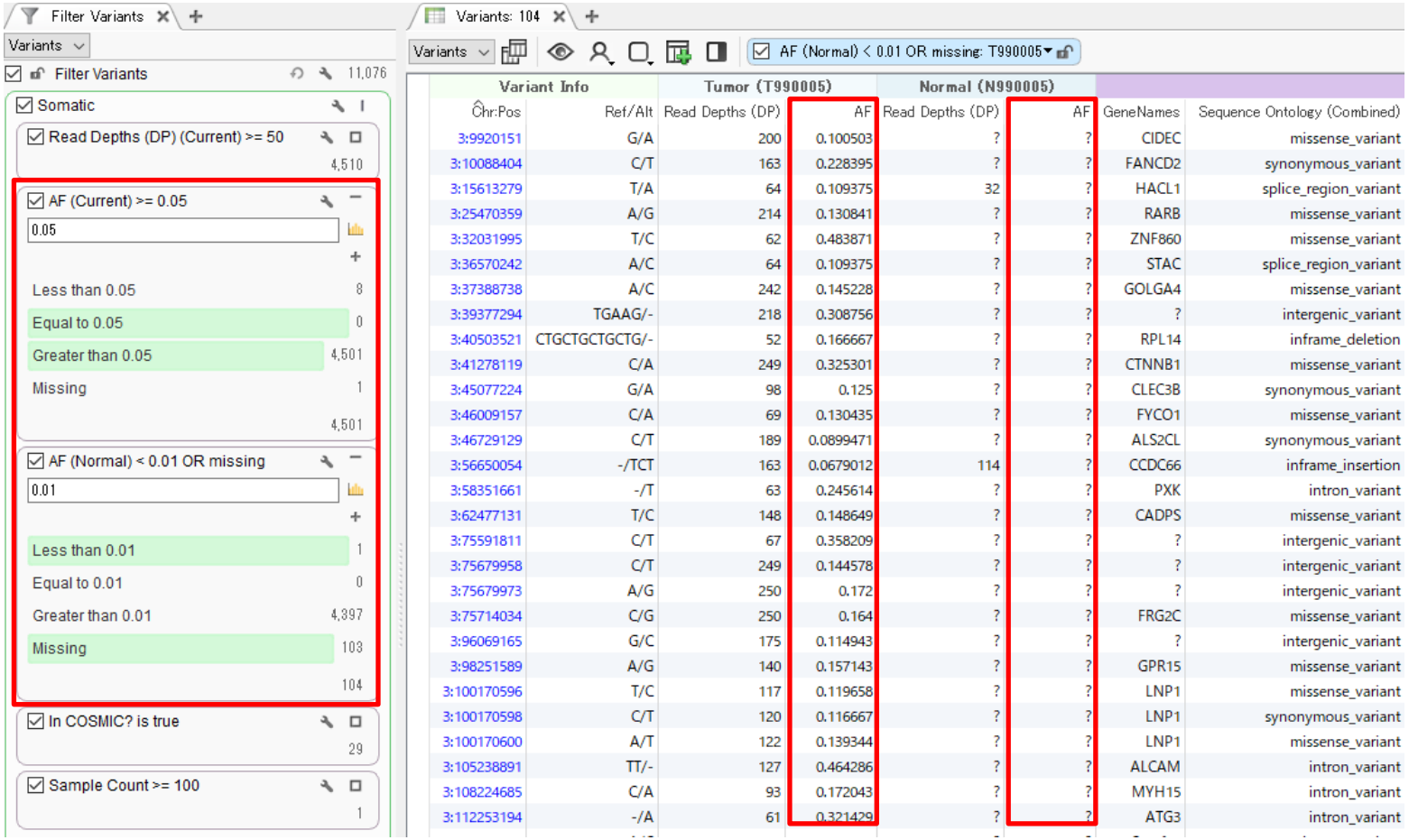
✓ 複合ヘテロ変異

Gene	Compound Het Variants for Proband ···		Compound Het Genes for Proband (HG02024)				
Gene Names	Compound Het?	Inherited From	Has Compound Het?	Inherited from Father	Inherited from Mother	Inherited Total	Hets In Both Parents
BPIFB3	False	NA	False	0	1	1	0
BPIFB4	False	NA	False	0	2	2	0
BPIFB4	False	NA	False	0	2	2	0
RALY	True	Father	True	1	1	2	0
RALY	True	Mother	True	1	1	2	0
FAM83C	False	NA	False	1	0	1	0



Variant ···	Compound He···	Variant Info		Proband (HG02024)				Mother (HG02025)				Father (HG02026)			
Gene Names	HasCompoundHet	Chr:Pos	Ref/Alt	VAF	DP	Inherited From	Zygosity	VAF	DP	Inherited From	Zygosity	VAF	DP	Inherited From	Zygosity
ZNF862	True	20:32664865	-/CAG	0.341463	41	Father	Heterozygous	?	?	?	?	0.4375	48	?	Heterozygous
SCN11A	True	20:32664955	G/T	0.504425	113	Mother	Heterozygous	0.414286	140	?	Heterozygous	?	?	?	?
SAMD14	True														
RP1L1	True														
RALY	True														
PCDHGB2	True														
PCDHGB1	True														

腫瘍-正常サンプル解析



Variant Info		Tumor (T990005)		Normal (N990005)			
Chr:Pos	Ref/Alt	Read Depths (DP)	AF	Read Depths (DP)	AF	GeneNames	Sequence Ontology (Combined)
3:9920151	G/A	200	0.100503	?	?	CIDEC	missense_variant
3:10088404	C/T	163	0.228395	?	?	FANCD2	synonymous_variant
3:15613279	T/A	64	0.109375	32	?	HACL1	splice_region_variant
3:25470359	A/G	214	0.130841	?	?	RARB	missense_variant
3:32031995	T/C	62	0.483871	?	?	ZNF860	missense_variant
3:36570242	A/C	64	0.109375	?	?	STAC	splice_region_variant
3:37388738	A/C	242	0.145228	?	?	GOLGA4	missense_variant
3:39377294	TGAAG/-	218	0.308756	?	?	?	intergenic_variant
3:40503521	CTGCTGCTGCTG/-	52	0.166667	?	?	RPL14	inframe_deletion
3:41278119	C/A	249	0.325301	?	?	CTNNB1	missense_variant
3:45077224	G/A	98	0.125	?	?	CLEC3B	synonymous_variant
3:46009157	C/A	69	0.130435	?	?	FYCO1	missense_variant
3:46729129	C/T	189	0.0899471	?	?	ALS2CL	synonymous_variant
3:56650054	-/TCT	163	0.0679012	114	?	CCDC66	inframe_insertion
3:58351661	-/T	63	0.245614	?	?	PXK	intron_variant
3:62477131	T/C	148	0.148649	?	?	CADPS	missense_variant
3:75591811	C/T	67	0.358209	?	?	?	intergenic_variant
3:75679958	C/T	249	0.144578	?	?	?	intergenic_variant
3:75679973	A/G	250	0.172	?	?	?	intergenic_variant
3:75714034	C/G	250	0.164	?	?	FRG2C	missense_variant
3:96069165	G/C	175	0.114943	?	?	?	intergenic_variant
3:98251589	A/G	140	0.157143	?	?	GPR15	missense_variant
3:100170596	T/C	117	0.119658	?	?	LNP1	missense_variant
3:100170598	C/T	120	0.116667	?	?	LNP1	synonymous_variant
3:100170600	A/T	122	0.139344	?	?	LNP1	missense_variant
3:105238891	TT/-	127	0.464286	?	?	ALCAM	intron_variant
3:108224685	C/A	93	0.172043	?	?	MYH15	intron_variant
3:112253194	-/A	61	0.321429	?	?	ATG3	intron_variant

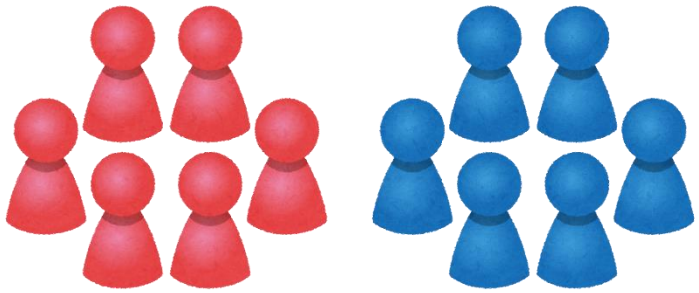
- 腫瘍-正常細胞の比較解析では、各サンプル内のアレル頻度データを比較し、腫瘍サンプル特異的な変異の抽出を行う

集団サンプル解析

Count Alleles									
Group2 - Allele Frequencies	Group2 - # Alleles	Group2 - # Het	Group2 - # HomoVar	Group2 - # Samples	Group1 - Allele Frequencies	Group1 - # Alleles	Group1 - # Het	Group1 - # HomoVar	Group1 - # Samples
0.5	16	8	0	8	0	14	0	0	0
0.666667	12	4	4	8	0.5	12	4	2	6
0.727273	11	3	5	8	0.583333	12	5	2	7
0.666667	12	4	4	8	0.5	12	4	2	6
0.5	16	8	0	8	0.5	14	7	0	7
1	8	0	8	8	1	7	0	7	7
0.5	16	8	0	8	0	14	0	0	0
0.5	16	8	0	8	0	14	0	0	0
0.5	14	5	2	7	0.416667	12	3	2	5
0.666667	12	3	5	8	0.583333	12	5	2	7

サンプルグループ1

サンプルグループ2



- 全サンプル、またはユーザー定義のサンプルグループごとに、各変異のアレル数や頻度、変異をもつサンプル数を計算し、フィルタリングに用いることが可能
- サンプル数は、ホモとヘテロそれぞれの数をカウント

ゲノムブラウザー

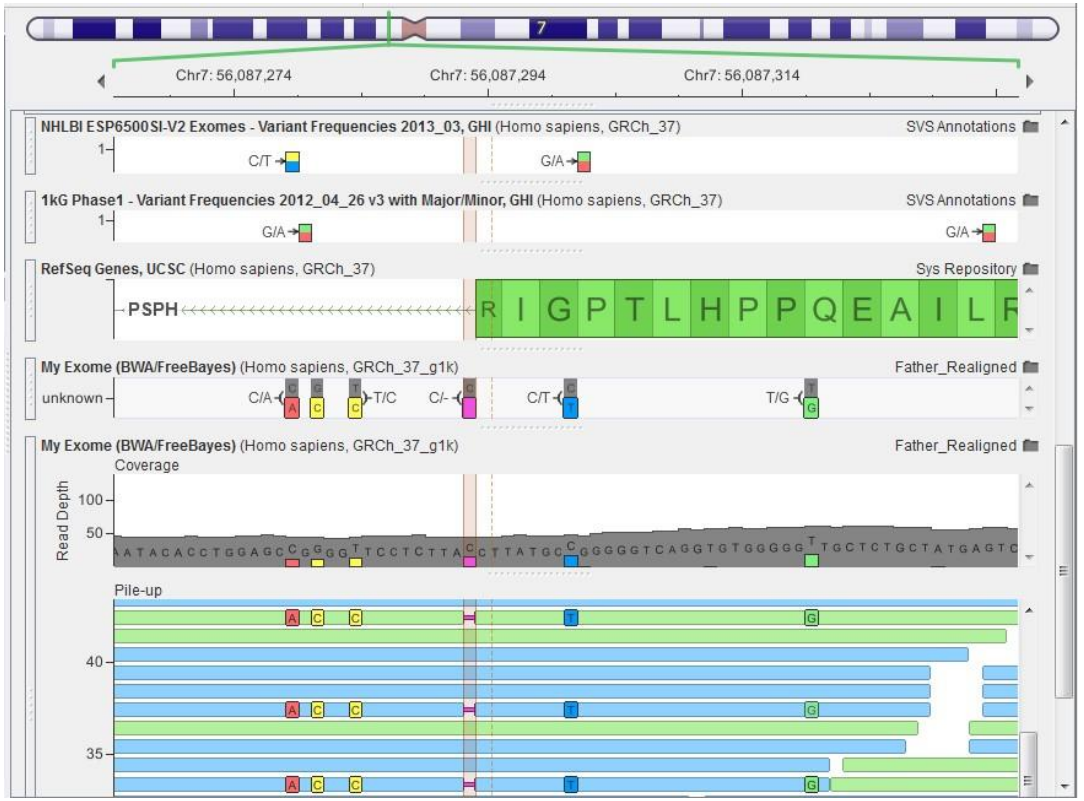
- VCFファイルに加えて、BAMファイルの情報もプロットを行うことができ、シークエンスデータの目視チェックなどに活用
- アノテーション用データベースも同時にプロットし、アノテーションの詳細情報などの確認が可能

✓ データベース

RefSeq Genes　ClinVar　dbSNP

✓ サンプルデータ

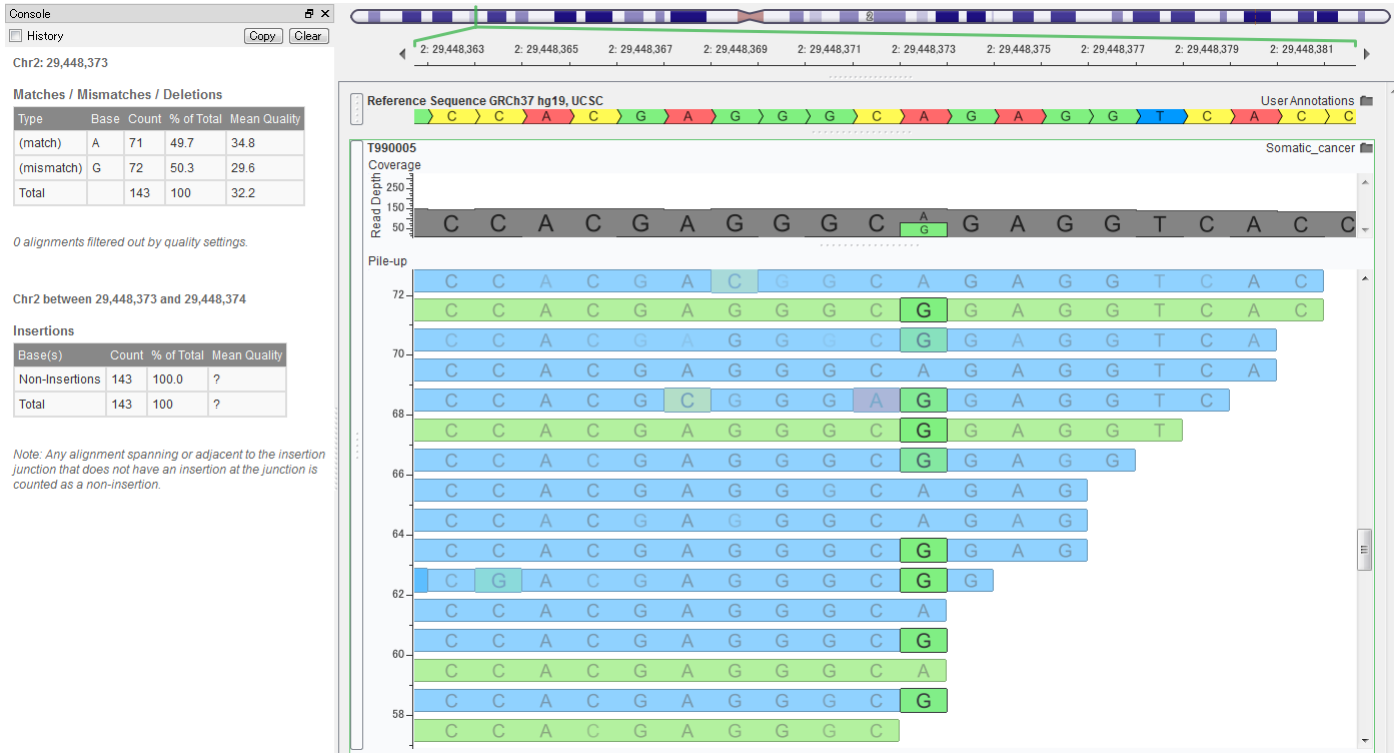
VCF　BAM



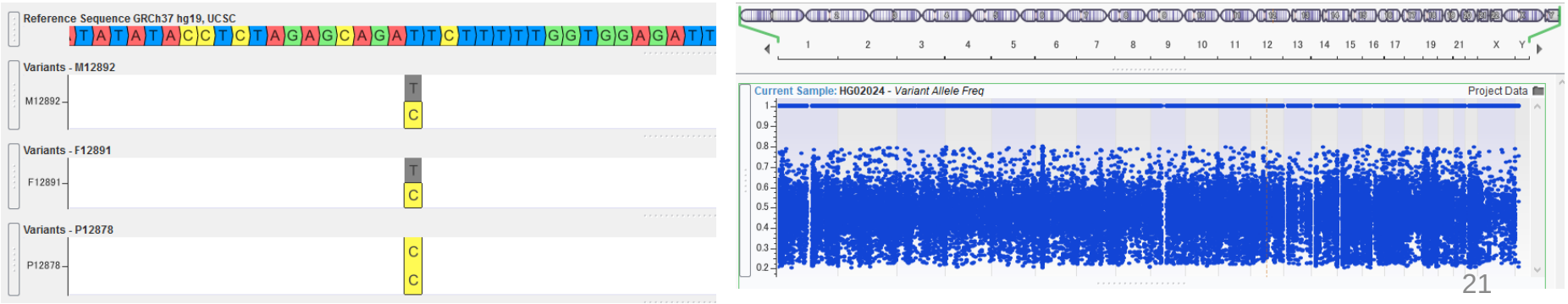
ゲノムブラウザー

Filgen® biosciences & nanosciences

✓ BAMファイルプロット

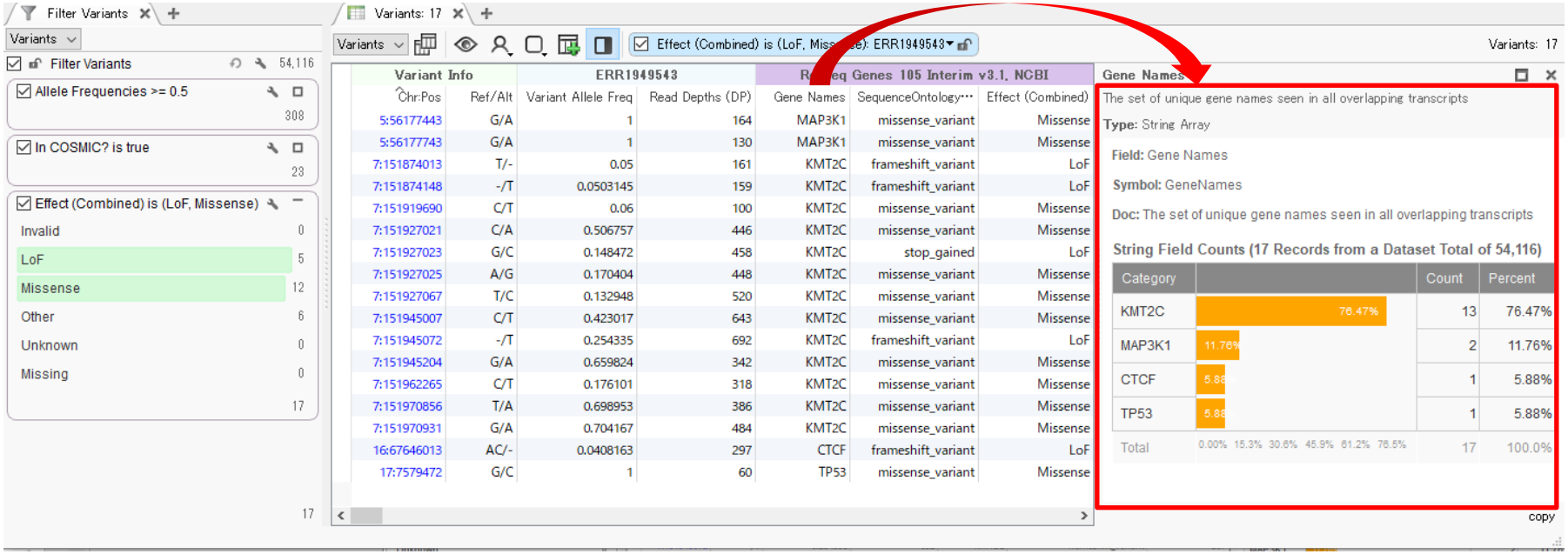


✓ VCFファイルプロット



データ集計グラフ

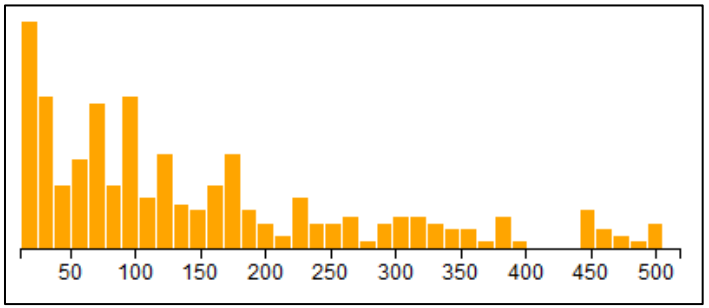
- 変異データテーブル上の各フィールドのデータは、ワンクリックで集計グラフの作成が可能



✓ Sequence Ontology

Category		Count	Percent
intergenic_variant	52.27%	161	52.27%
intron_variant	25.97%	80	25.97%
3_prime_UTR_variant	10.71%	33	10.71%
missense_variant	5.52%	17	5.52%
synonymous_variant	2.9%	9	2.92%
frameshift_variant	1.3%	4	1.30%
splice_region_variant	0.6%	2	0.65%
stop_gained	0.3%	1	0.32%
5_prime_UTR_variant	0.3%	1	0.32%

✓ Read Depth



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: biosupport@filgen.jp