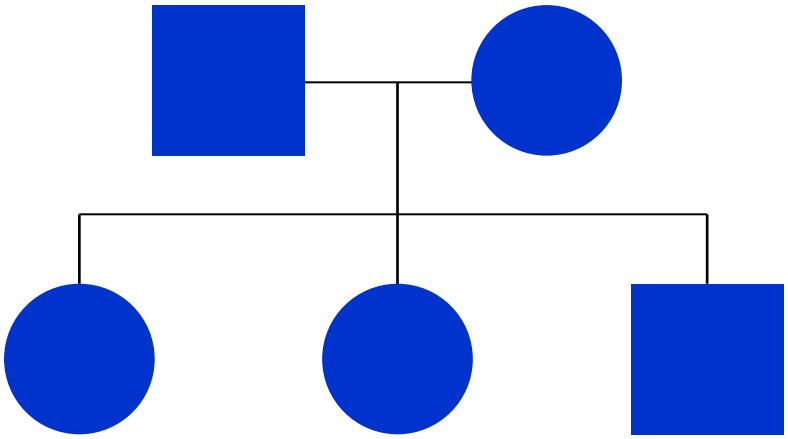
家系サンプルを用いた遺伝性疾患のエクソーム解析

フィルジェン株式会社　バイオサイエンス部
(biosupport@filgen.jp)

はじめに

- 遺伝性疾患の責任遺伝子を決定するために全エクソームシークエンスを用いる場合、患者サンプルの遺伝子変異データを基に、両親や近親者などの家系サンプルの変異データも利用して、疾患の候補遺伝子の絞り込みを行う。

- Golden Helix社VarSeq®では、サンプルの家系情報に基づいたトリオ解析や、遺伝子と疾患の関連のランク付けを行う解析アルゴリズム、また人種ごとのアレル頻度情報のデータベースなどを使用し、膨大な変異データの中から、疾患の原因となる変異を迅速に特定できる。



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 Diagnostics
- Gene Testing & Rare Disease Diagnosis
- CNV Calling
- Clinical Reporting
- High-throughput NGS Testing

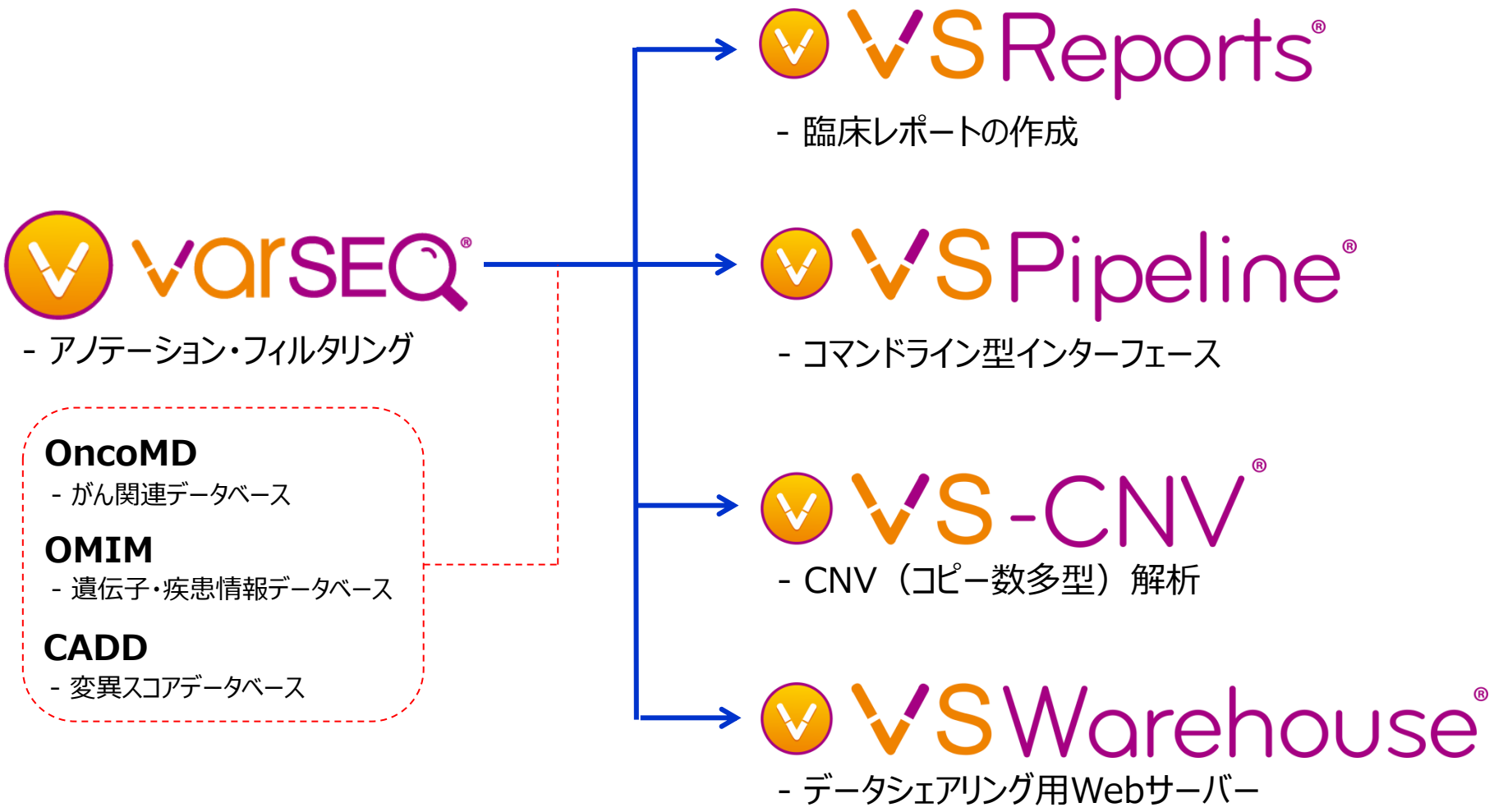
- Golden Helix社では、遺伝統計解析ソフトウェア「SNP & Variation Suite」と、遺伝子診断支援ソフトウェア「VarSeq®」の2種類のソフトウェアパッケージを販売
- 医学・生物学研究や、家畜や作物の品種改良などの農学研究、さらに疾患の診断や最適な治療オプションの決定における医療分野、遺伝学的解析などの教育現場などで利用される

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
  - dbSNP
  - RefSeq Genes
  - COSMIC
  - 1000 Genome
  - NHLBI 6500 Exomes
  - ExAC Variant
  - gnomAD Exomes
  - SIFT and PolyPhen
  - dbNSFP Functional Predictions
  - ClinVar
  - CIVic
  - ICGC Simple Somatic Mutation
  - 各種遺伝子パネルのターゲットデータ

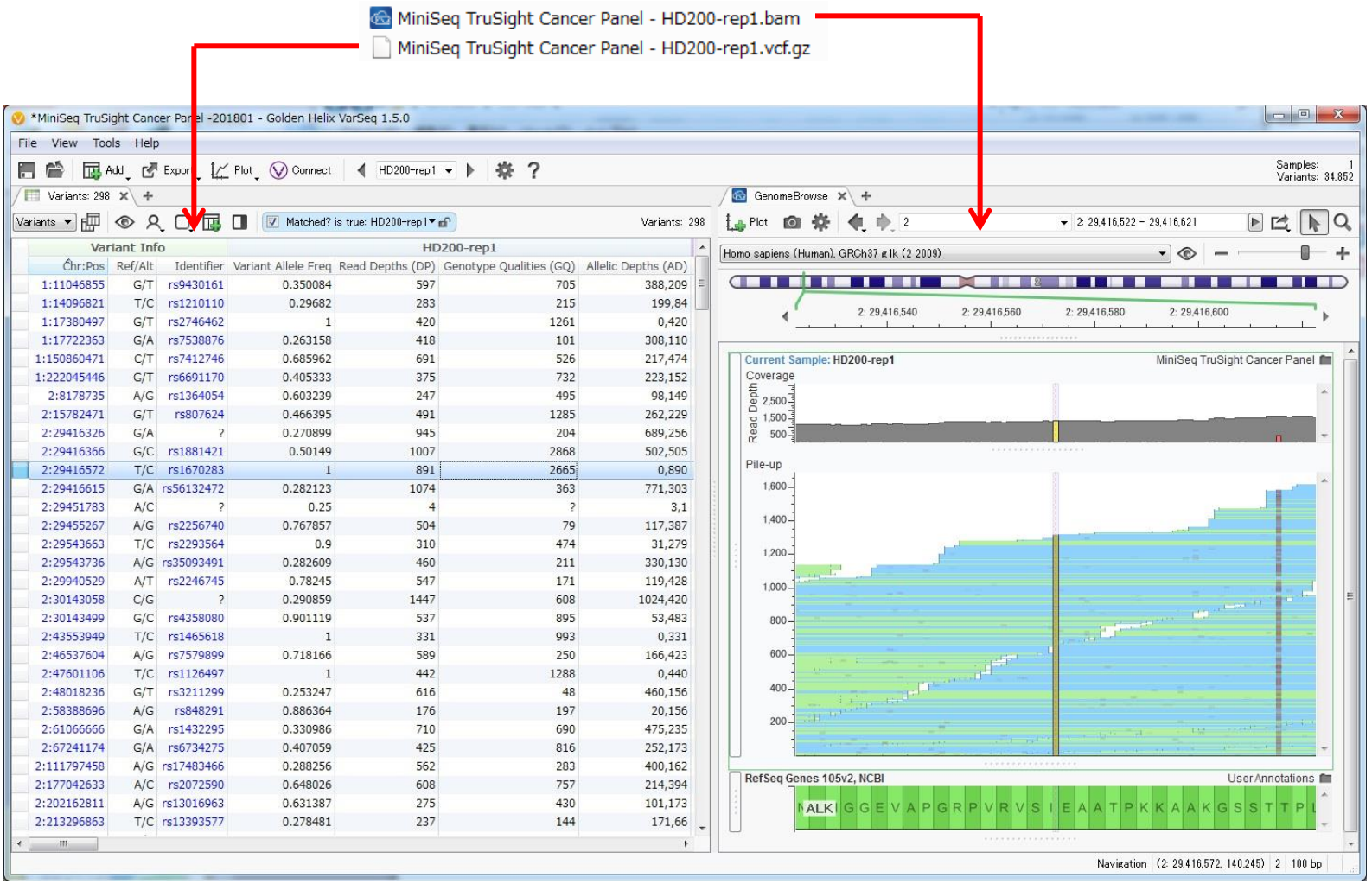
  …など
- VCFファイルに含まれる変異データから、任意の検索条件でデータのフィルタリングを行うワークフローを作成
- HGVDなどの独自定義ファイルもアノテーションデータとして利用可能
- カバレッジ計算やトリオ解析、表現型情報に基づく遺伝子ランキングなどの解析アルゴリズムを搭載
- ゲノムブラウザーを搭載し、BAMファイルデータや各種アノテーションデータをグラフ表示
- 無償提供のビューワーソフトウェアが利用でき、解析結果を容易にシェアすることが可能

有償アドオン



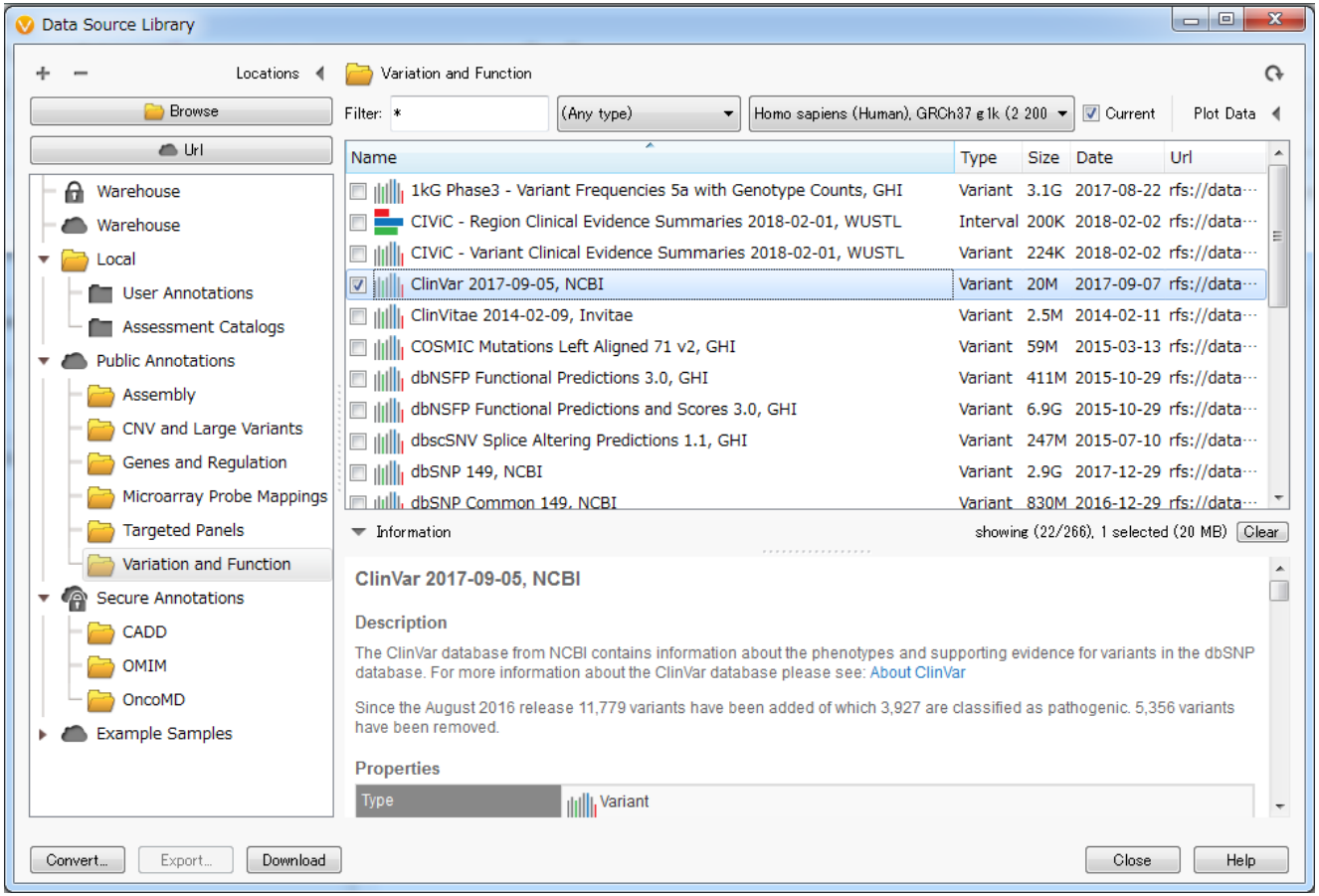
- 別途有償のアドオンを追加することで、新たなアノテーションリソースや、新たな機能の使用が可能になる

使用するデータ



- 次世代シークエンサーの変異解析パイプラインで作成したVCFファイル（変異データ）とBAMファイル（アライメントデータ）を使用して、各種データ解析を実行

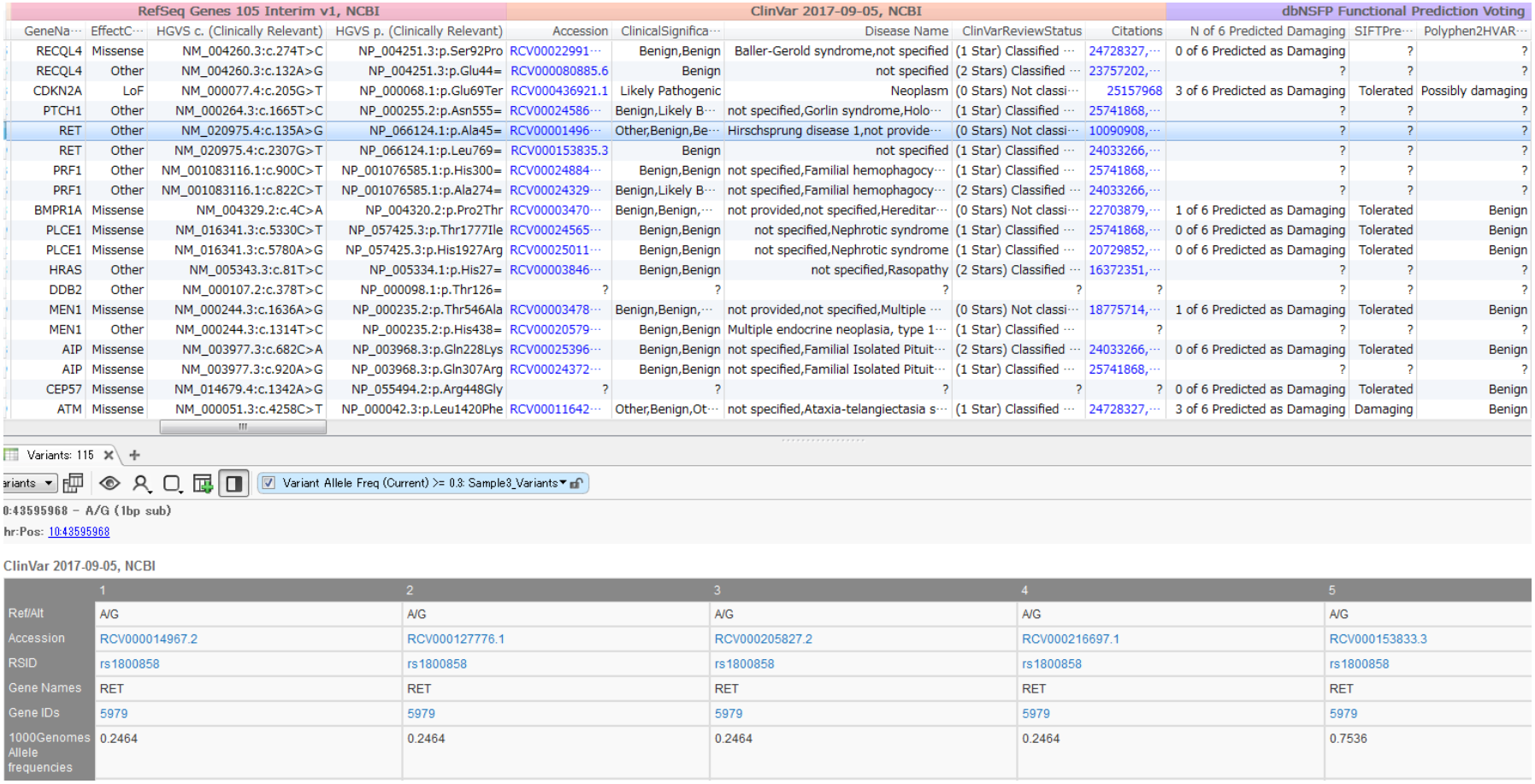
アノテーションリソース



- Golden Helix社によってデータの精査・メンテナンスが行われている各種アノテーションリソースを、Data Source Libraryより自由にダウンロードし、変異データへのアノテーション付けに使用可能

- 日本人変異データベースHGVDなどのカスタムアノテーションデータをインポートし、アノテーションリソースとして使用が可能

アノテーションデータ

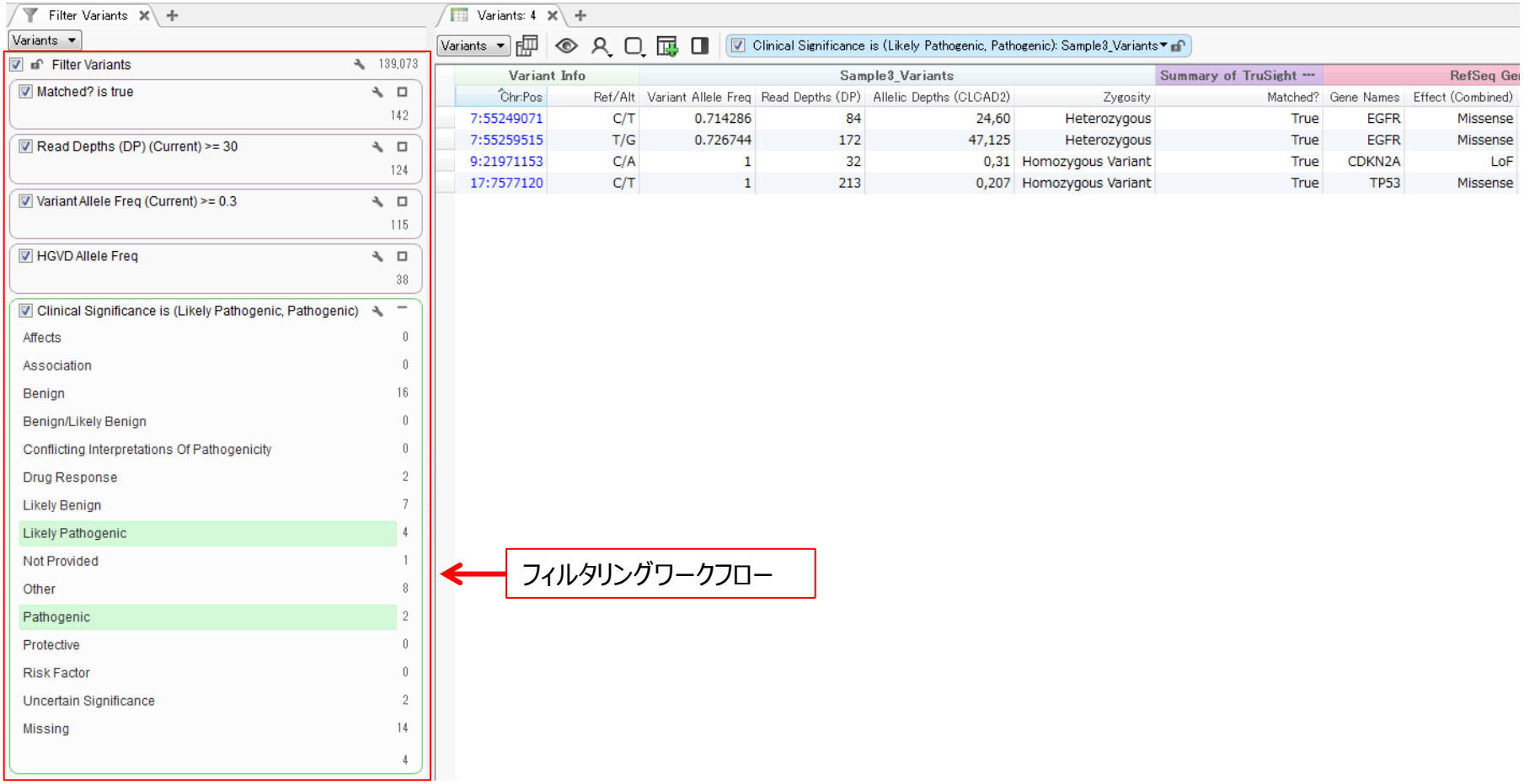


RefSeq Genes 105 Interim v1, NCBI				ClinVar 2017-09-05, NCBI					dbNSFP Functional Prediction Voting		
GeneNa…	EffectC…	HGVS c. (Clinically Relevant)	HGVS p. (Clinically Relevant)	Accession	ClinicalSignifica…	Disease Name	ClinVarReviewStatus	Citations	N of 6 Predicted Damaging	SIFTPre…	Polyphen2HVAR…
RECQL4	Missense	NM_004260.3:c.274T>C	NP_004251.3:p.Ser92Pro	RCV00022991…	Benign,Benign	Baller-Gerold syndrome,not specified	(1 Star) Classified …	24728327,…	0 of 6 Predicted as Damaging	?	?
RECQL4	Other	NM_004260.3:c.132A>G	NP_004251.3:p.Glu44=	RCV000080885.6	Benign	not specified	(2 Stars) Classified …	23757202,…	?	?	?
CDKN2A	LoF	NM_000077.4:c.205G>T	NP_000068.1:p.Glu69Ter	RCV000436921.1	Likely Pathogenic	Neoplasm	(0 Stars) Not classi…	25157968	3 of 6 Predicted as Damaging	Tolerated	Possibly damaging
PTCH1	Other	NM_000264.3:c.1665T>C	NP_000255.2:p.Asn555=	RCV00024586…	Benign,Likely B…	not specified,Gorlin syndrome,Holo…	(1 Star) Classified …	25741868,…	?	?	?
RET	Other	NM_020975.4:c.135A>G	NP_066124.1:p.Ala45=	RCV00001496…	Other,Benign,Be…	Hirschsprung disease 1,not provide…	(0 Stars) Not classi…	10090908,…	?	?	?
RET	Other	NM_020975.4:c.2307G>T	NP_066124.1:p.Leu769=	RCV000153835.3	Benign	not specified	(1 Star) Classified …	24033266,…	?	?	?
PRF1	Other	NM_001083116.1:c.900C>T	NP_001076585.1:p.His300=	RCV00024884…	Benign,Benign	not specified,Familial hemophagocy…	(1 Star) Classified …	25741868,…	?	?	?
PRF1	Other	NM_001083116.1:c.822C>T	NP_001076585.1:p.Ala274=	RCV00024329…	Benign,Likely B…	not specified,Familial hemophagocy…	(2 Stars) Classified …	24033266,…	?	?	?
BMPR1A	Missense	NM_004329.2:c.4C>A	NP_004320.2:p.Pro2Thr	RCV00003470…	Benign,Benign,…	not provided,not specified,Hereditar…	(0 Stars) Not classi…	22703879,…	1 of 6 Predicted as Damaging	Tolerated	Benign
PLCE1	Missense	NM_016341.3:c.5330C>T	NP_057425.3:p.Thr1777Ile	RCV00024565…	Benign,Benign	not specified,Nephrotic syndrome	(1 Star) Classified …	25741868,…	0 of 6 Predicted as Damaging	Tolerated	Benign
PLCE1	Missense	NM_016341.3:c.5780A>G	NP_057425.3:p.His1927Arg	RCV00025011…	Benign,Benign	not specified,Nephrotic syndrome	(1 Star) Classified …	20729852,…	0 of 6 Predicted as Damaging	Tolerated	Benign
HRAS	Other	NM_005343.3:c.81T>C	NP_005334.1:p.His27=	RCV00003846…	Benign,Benign	not specified,Rasopathy	(2 Stars) Classified …	16372351,…	?	?	?
DDB2	Other	NM_000107.2:c.378T>C	NP_000098.1:p.Thr126=	?	?	?	?	?	?	?	?
MEN1	Missense	NM_000244.3:c.1636A>G	NP_000235.2:p.Thr546Ala	RCV00003478…	Benign,Benign,…	not provided,not specified,Multiple …	(0 Stars) Not classi…	18775714,…	1 of 6 Predicted as Damaging	Tolerated	Benign
MEN1	Other	NM_000244.3:c.1314T>C	NP_000235.2:p.His438=	RCV00020579…	Benign,Benign	Multiple endocrine neoplasia, type 1…	(1 Star) Classified …	?	?	?	?
AIP	Missense	NM_003977.3:c.682C>A	NP_003968.3:p.Gln228Lys	RCV00025396…	Benign,Benign	not specified,Familial Isolated Pituit…	(2 Stars) Classified …	24033266,…	0 of 6 Predicted as Damaging	Tolerated	Benign
AIP	Missense	NM_003977.3:c.920A>G	NP_003968.3:p.Gln307Arg	RCV00024372…	Benign,Benign	not specified,Familial Isolated Pituit…	(1 Star) Classified …	25741868,…	?	?	?
CEP57	Missense	NM_014679.4:c.1342A>G	NP_055494.2:p.Arg448Gly	?	?	?	?	?	0 of 6 Predicted as Damaging	Tolerated	Benign
ATM	Missense	NM_000051.3:c.4258C>T	NP_000042.3:p.Leu1420Phe	RCV00011642…	Other,Benign,Ot…	not specified,Ataxia-telangiectasia s…	(1 Star) Classified …	24728327,…	3 of 6 Predicted as Damaging	Damaging	Benign

	1	2	3	4	5
Ref/Alt	A/G	A/G	A/G	A/G	A/G
Accession	RCV000014967.2	RCV000127776.1	RCV000205827.2	RCV000216697.1	RCV000153833.3
RSID	rs1800858	rs1800858	rs1800858	rs1800858	rs1800858
Gene Names	RET	RET	RET	RET	RET
Gene IDs	5979	5979	5979	5979	5979
1000Genomes Allele frequencies	0.2464	0.2464	0.2464	0.2464	0.7536

- 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

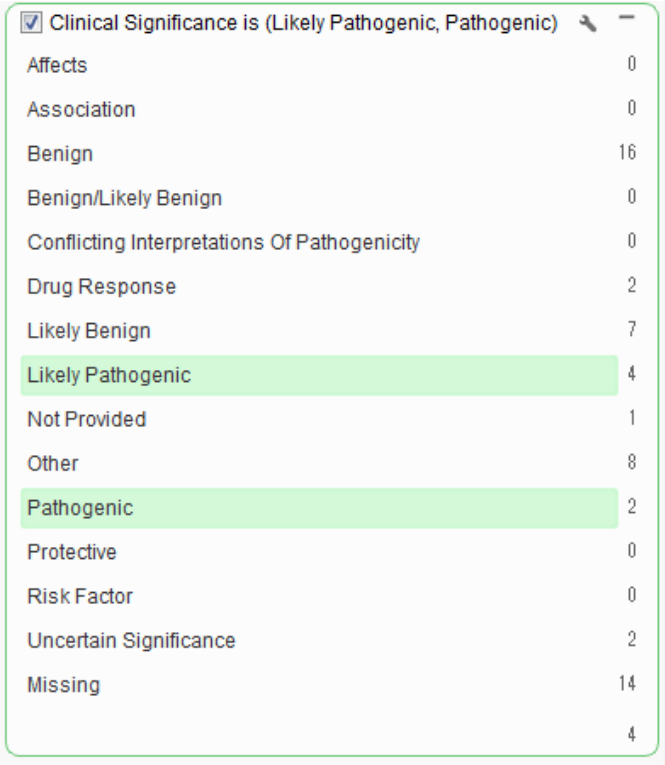
ワークフロー作成と結果の確認



- 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

ClinVar

ClinVar 2017-09-05, NCBI							
Ref/Alt	Accession	Gene Names	HGVS g. Name	Clinical Significance	MedGen	Disease Name	ClinVar Review Status
T/-	RCV000007567.3	CFTR	NC_000007.13:g.117171108delT	Pathogenic	C0010674	Cystic fibrosis	(0 Stars) Not classified by submitter
G/A	RCV000114996.3	ZPR1,APOA5	NC_000011.9:g.116660686G>A	Risk Factor	C2676231	Hypertriglyceridemia, susc...	(0 Stars) Not classified by submitter
A/-	RCV000029281.1	ABCC9	NC_000012.11:g.21958999delA	Uncertain Significance	C0878544	Cardiomyopathy	(1 Star) Classified by single submit...
G/A	RCV000337303.1	TBX3	NC_000012.11:g.115117337G>A	Likely Benign	C1866994	Ulnar-mammary syndrome	(1 Star) Classified by single submit...



- NCBIで提供している、ヒト疾患関連の変異情報をまとめたデータベース
- 遺伝子名や疾患名、疾患に対する重要度、データの信憑性などのアノテーションが、サンプルの変異データに付加される
- ワークフローを用いて、疾患に関連する変異の迅速な検索が可能になる

OMIM

OMIM Variants 2017-04-01, GHI

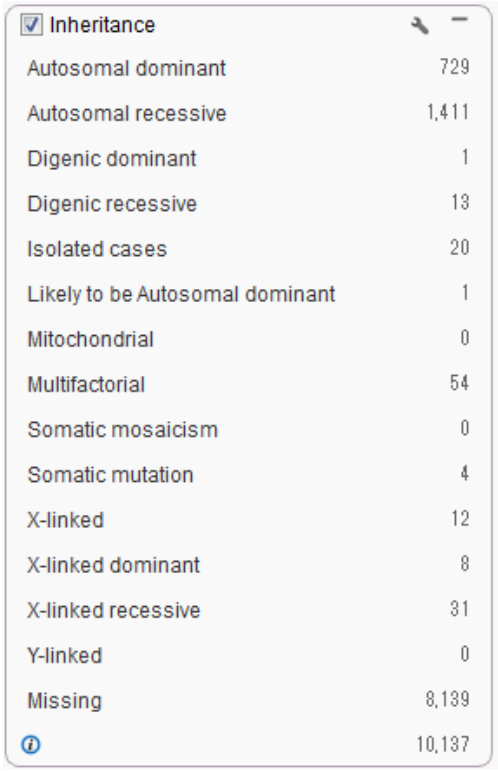
Ref/Alt	Phenotype	Gene Name	GeneOMIMID	Entrez Gene ID	PubMed ID	HasPubMedID	Name	dbSNP	Description	References
A/C	INSULIN ...	HNF1A	142410	6927	12788852,...	True	HNF1A, IL...	rs1169288	<p><a hr...	1. Babaya ...
G/C	CODON 7...	TP53	191170	7157	11403041,...	True	TP53, PR...	rs1042522	<p><a hr...	1. Aaltone...

OMIM Genes 2017-04-01, GHI

Gene Name	OMIM ID	PubMed ID	Title	Description	Gene Status	Disorders
HNF1A	142410	12788852,1707...	HNF1 HOMEO...	?	Confirmed	Diabetes mellit...
TP53	191170	11403041,8673...	TUMOR PROTE...	<p>The transc...	Confirmed	Adrenal cortica...

OMIM Phenotypes 2017-04-01, GHI

Gene Names	Cytogenetic Locations	Entrez Gene IDs	GeneOMIMIDs	Inheritance	OMIM ID	PubMed ID	HasPubMedID	Title	Alternative Title(s)	Description	References
GPD2,NE...	2q24.1,2q32,2q36,3p...	2820,4760,366...	138430,601...	Autosomal dominant,Multifactorial,A...	125853,14...	17726085,...	True,True,Tr...	DIABETES...	DIABETES MELLITU...	<p>Mole...	16. Elbein ...
RAD54L,C...	1p32,2q33,2q34-q35,...	8438,841,580,...	603615,601...	Autosomal dominant,Somatic mutati...	114480,11...	19330027,...	True,True,Tr...	BREAST C...	BREAST CANCER F...	<p>Breas...	9. Anzick ...



- おもにヒトの遺伝性疾患と遺伝子に関連した情報をまとめたデータベースで、使用には別途ライセンスの購入が必要
- 表現型レベルと遺伝子レベル、変異レベルの3種類のアノテーションデータがサンプルの変異データに付加される
- 表現型の遺伝様式に基づいた、変異データのフィルタリングが可能

がん関連変異情報データベース

- がんに関連した変異情報のみに絞ったデータベース
- CIViC, COSMIC, ICGCの3種類のデータベースを利用可能
- がんのタイプや遺伝子名などの他に、データベースごとに様々な詳細情報がアノテーションとして利用できる

CIViC - Region Clinical Evidence Summaries 2017-08-01, WUSTL

Gene Name	Representative Transcript	Variant Type	Disease	Disease Ontology ID	Drugs	Clinical Significance	Evidence Direction	Evidence Level	Trust Rating
APC,APC	ENST00000457016.1,ENST0000...	MUTATION,...	Colon Carcin...	1520,9256	JW55,G007-LK	Sensitivity,Sensitivity	Supports,Supports	D - Preclinical,D - P...	3 out of 5 Stars,4 out ...
PTCH1,PTCH1	ENST00000331920.6,ENST0000...	MUTATION,L...	Brain Medull...	0060105,0060105	Vismodegib,...	Sensitivity,Sensitivity	Supports,Supports	B - Clinical,B - Clini...	4 out of 5 Stars,2 out ...
TP53,TP53,T...	ENST00000269305.4,ENST0000...	DELETERIOU...	Head And N...	5520,5520,3748,7061,0...	Chemothera...	Poor Outcome,Poor ...	Supports,Does Not S...	B - Clinical,B - Clini...	3 out of 5 Stars,3 out ...

COSMIC Mutations Left Aligned 71 v2, GHI

Ref/Alt	Mutation ID	Mutation CDS	Mutation AA	Gene Name	Transcript ID	Gene CDS Length	HGNC ID	Primary Site	Mutation Description	Mutation Zygosity
A/C	430522	c.79A>C	p.I27L	HNF1A	ENST0000025...	1896	11621	Prostate (2),...	Substitution - Missense	Heterozygous (1)
G/C,G/C,...	250061,376...	c.215C>G,c....	p.P72R,p.P...	TP53,TP5...	ENST0000026...	1182,1182,1041,1...	11998,?,?,?,?	Upper aerod...	Substitution - Missense,Substitution - Misse...	Homozygous Varia...

ICGC Simple Somatic Mutations 22, GHI

Ref/Alt	Identifier	AffectedDonorsForAllProjects	Project Count	Project ID	Primary Site	Affected Donors	Total Samples	Affected Donor Frequency
G/C	MU151094	1	1	COAD-US	Colorectal	1	216	0.00463
A/G	MU156543	1	1	COAD-US	Colorectal	1	216	0.00463
T/C	MU3888690	1	1	THCA-SA	HeadAndNeck	1	129	0.00775
A/G	MU112255	1	1	COAD-US	Colorectal	1	216	0.00463

OncoMD

- MedGenome社が作成している、がん関連の様々な情報をまとめたデータベースで、使用には別途ライセンスの購入が必要
- がん関連遺伝子や変異などの機能情報の他に、承認薬や臨床試験情報、論文情報や機能情報など、様々な情報をアノテーションとして使用できる
- 有償アドオン「VSReports」を使用することで、効果をもつ承認薬や、現在進行中の臨床試験情報をまとめたレポートの作成が可能

OncoMD Drugs Targeting Mutation

Gene Symbol	Drug	Generic Name	Response Rate	ReportedSample…	Response (Summary)	Response Category (Summary)	Affected Domain (Summary)	Target Effect (Summary)	Overall Survival (Summary)
EGFR,EGFR…	Iressa,Tarceva,G…	Gefitinib,Er…	0%,0%,0%…	18,2,2,32	Progressive disea…	Nonresponder (18),Nonre…	Kinase (18),Kinase (2),…	ND (18),ND (2),ND …	ND (18),59 Months (2)…
EGFR,EGFR…	Caprelsa,Iressa,…	Vandetanib…	0%,64%,52…	3,337,23,7,2	Progressive disea…	Nonresponder (3),ND (7),…	Kinase (3),ND (2),Kinas…	ND (3),ND (284),Re…	38.5 Months (1),4.1 M…

OncoMD Clinical Trials

Gene Symbol	Cancer Type	Country	Drugs	Inclusion Criterion	Status	Trial Number
EGFR,EGFR…	Pancreatic …	Canada,Chi…	erlotinib hy…	EGFR MUTATION,E…	Recruiting,…	NCT01013…
EGFR,EGFR…	Pancreatic …	Canada,Chi…	erlotinib hy…	EGFR MUTATION,E…	Recruiting,…	NCT01013…
?	?	?	?	?	?	?
?	?	?	?	?	?	?

OncoMD Studies with Variant

Gene Symbol	PubMed ID	Study Type	Title	Sample Count
EGFR,EGFR…	?,15604253…	No Study T…	?,Mutations of th…	1,1,1,5,1,1,…
EGFR,EGFR…	?,22815900…	No Study T…	?,Loss of activati…	1,5,2,6,13,…
CDKN2A,C…	7478613,7…	NA,NA,NA,…	Mechanism of in…	4,1,1,1,1,1,…
TP53,TP53,…	10212000,…	NA,NA,Tar…	Mutation of p53 …	1,3,1,1,1,1,…

OncoMD Variant Summary

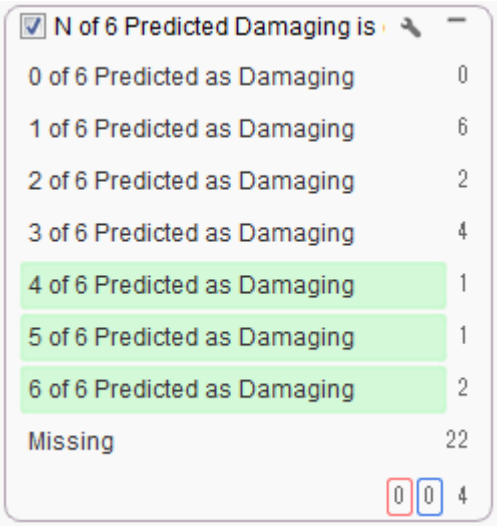
Gene Symbol	cDNA Mutation	AA Mutation	Observed Cancers	Study Count	Sample Count
EGFR	c.2369C > T	p.T790M	Lung cancer (1043),N…	165	1057
EGFR	c.2573T > G	p.L858R	Lung cancer (9247),N…	358	9275
CDKN2A	c.205G > T	p.E69X	Lung cancer (7),Panc…	16	20
TP53,TP53	c.42091G >…	p.R273H,p.R273H	Stomach (gastric) ca…	1,366	1,702

OncoMD Functional Validation of Variant

PubMed ID	Cancer Type	Biochemical Assay	Cellbased Assay	Effect on Protein Function	Mechanism of Action	MOA Validation	Comments
17877814,…	Non Small …	ND,Western blo…	No hyper phos…	None,None,Gain-of-fu…	ND,ND,ND,ND	ND,Crystallo…	ND,ND,Cell…
16187797,…	Lung aden…	Western blot ana…	>Colony form…	Gain-of-function,Gain-…	ND,ND,ND,ND,N…	ND,ND,ND,N…	>NIH-3T3 …
?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?

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)	FATHMM MKL Coding Pred (C)
0 of 6 Predicted as Tolerated	6 of 6 Predicted as Damaging	Damaging	Possibly damaging	Damaging	Predicted functional (medium)	Damaging	Damaging
1 of 6 Predicted as Tolerated	5 of 6 Predicted as Damaging	Damaging	Probably damaging	Damaging	Predicted functional (medium)	Tolerated	Damaging
0 of 6 Predicted as Tolerated	6 of 6 Predicted as Damaging	Damaging	Possibly damaging	Damaging	Predicted functional (medium)	Damaging	Damaging
0 of 6 Predicted as Tolerated	6 of 6 Predicted as Damaging	Damaging	Probably damaging	Damaging	Predicted functional (medium)	Damaging	Damaging
2 of 6 Predicted as Tolerated	4 of 6 Predicted as Damaging	Damaging	Possibly damaging	Damaging	Predicted non-functional (neutral)	Tolerated	Damaging
2 of 6 Predicted as Tolerated	4 of 6 Predicted as Damaging	Damaging	Probably damaging	Damaging	Predicted non-functional (low)	Tolerated	Damaging
2 of 6 Predicted as Tolerated	4 of 6 Predicted as Damaging	Damaging	Possibly damaging	Damaging	Predicted non-functional (low)	Tolerated	Damaging
1 of 6 Predicted as Tolerated	5 of 6 Predicted as Damaging	Damaging	Probably damaging	Damaging	Predicted non-functional (low)	Damaging	Damaging



- コンピュータプログラムで、変異の病原性を予測したデータベース
- SIFT, Polyphen2, MutationTaster, MutationAssessor, FATHMM, FATHMM MKLの6種類のプログラムによる予測結果を同時に評価が可能
- 予測値のスコアも確認することができる

アレル頻度情報データベース

- 特定の集団内のアレル頻度情報のデータベース
- ExAC, gnomAD, NHLBI ESP6500, 1000 Genomes Phase3の4種類のデータベースを利用可能
- アレル頻度はヨーロッパ、アジア、アフリカなどのカテゴリーごとに計算されており、これらアレル頻度データを使用することで、集団内に低頻度で存在する変異などを検索可能

ExAC Variant Frequencies 0.3, BROAD					gnomAD Exomes Variant Frequencies 2.0.1 v2, BROAD				
Ref/Alt	Identifier	Filter	Alt Allele Freq (AF)	Alt Allele Counts (AC)	Ref/Alt	Filter	Alt Allele Prob (RF)	Alt Allele Freq (AF)	Ashkenazi Jewish Allele Count (AC_ASJ)
G/A	rs72975710	PASS	0.0001978	24	G/A	PASS	0.95295	0.000199914	0
C/T	rs72996036	PASS	3.295e-05	4	C/T	PASS	0.953006	2.4375e-05	0
A/G	rs421016	VQSRTrancheSNP99.60to99.80	0.003155	383	A/G	PASS	0.11371	0.00130657	26
G/A	rs73035708	PASS	0.0001977	24	G/A	PASS	0.945595	0.000138133	0
G/T	rs72914988	PASS	0.001466	178	G/T	PASS	0.954732	0.00167319	21
C/T	rs73477443	PASS	8.242e-06	1	C/T	PASS	0.892885	2.47519e-05	2
?	?	?	?	?	C/A	RF	0.0074095	4.10826e-06	0
?	?	?	?	?	A/G	PASS	0.95369	4.06121e-06	0

NHLBI ESP6500SI-V2-SSA137 Exomes Variant Frequencies 0.0.30, GHI							
Ref/Alt	Identifier	All AAF	European American AAF	African American AAF	All MAF	All HomoVar GTC	All Het GTC
G/A	rs72975710	0.000461326	0	0.00136178	0.000461326	0	6
C/T	rs72996036	7.68876e-05	0	0.000226963	7.68876e-05	0	1
A/G	rs421016	0.00030755	0.000465116	0	0.00030755	0	4
G/A	rs73035708	0.000615101	0	0.00181571	0.000615101	0	8
G/T	rs72914988	0.00284484	0.000930233	0.00658193	0.00284484	0	37

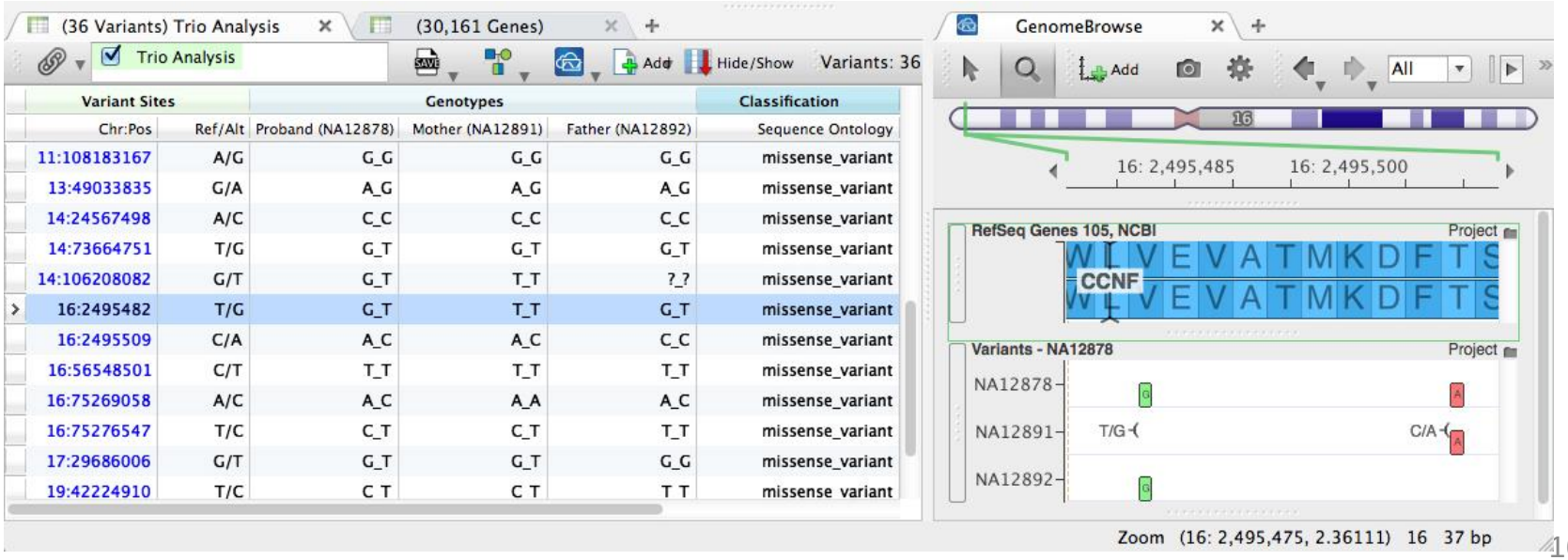
1kG Phase3 - Variant Frequencies 5b, GHI						
Ref/Alt	Identifier	All Indiv Freq	European Allele Freq (EUR_AF)	African/African American Allele Freq (AFR_AF)	American Allele Freq (AMR_AF)	South Asian Allele Freq (SAS_AF)
G/A	rs72975710	0.000798722	0	0.003	0	0
C/T	rs72996036	0.000399361	0	0.0008	0.0014	0
A/G	rs421016	0.00339457	0.0119	0.0015	0	0.002
G/A	rs73035708	0.000998403	0	0.0038	0	0
G/T	rs72914988	0.00319489	0.002	0.0061	0.0086	0
C/T	rs73477443	0.000199681	0	0.0008	0	0
C/A	rs73297817	0.000199681	0	0.0008	0	0

トリオ解析

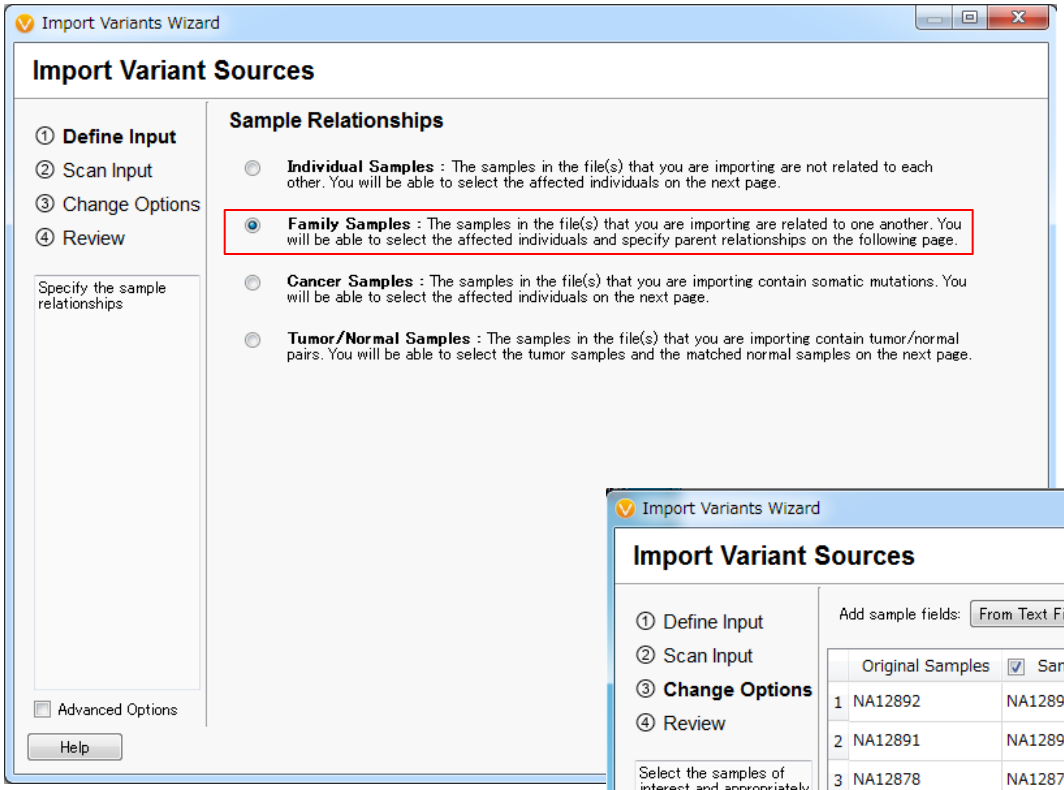
- VarSeq®では、外部データリソースによる変異データへのアノテーション付け以外に、サンプルの内部的なデータ（変異情報やリード深度情報など）を使用する解析アルゴリズムが搭載されている

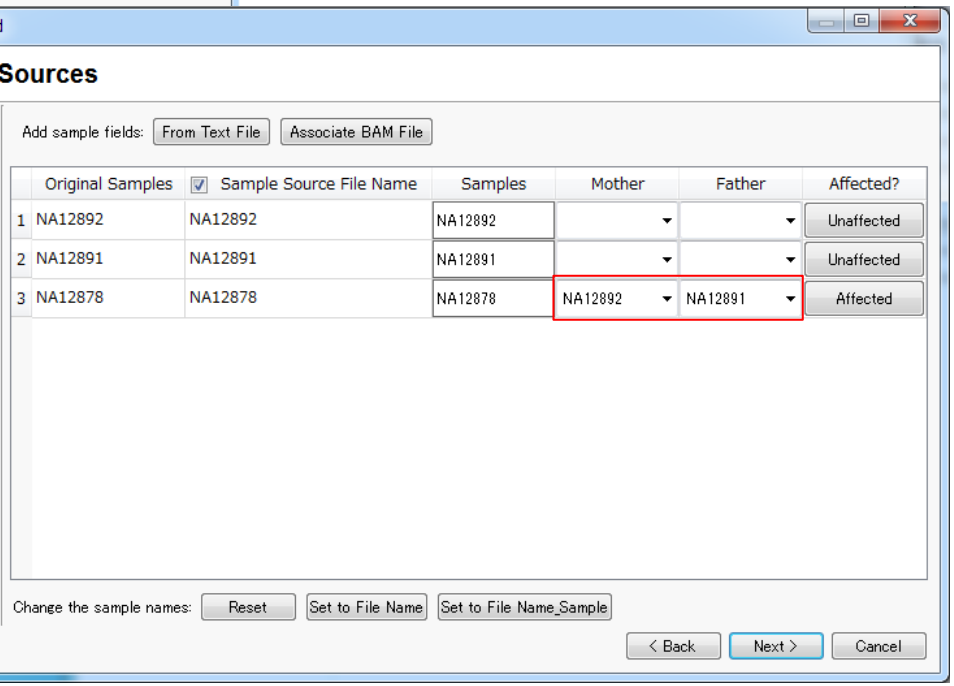
- これらの中には、親子サンプルのデータを比較して遺伝様式などを調べるトリオ解析のアルゴリズムが含まれており、遺伝性疾患の研究に活用できる

  - De novo mutation
  - Compound Heterozygous
  - Autosomal dominant
  - Autosomal recessive
  - X-Linked



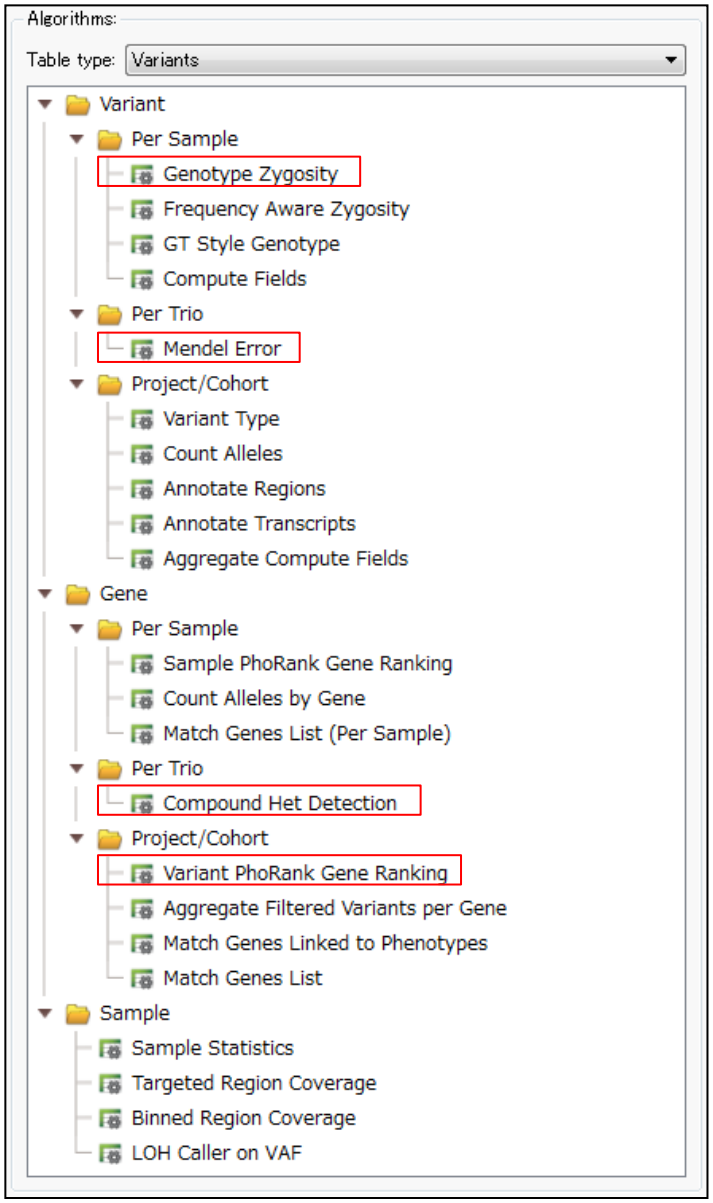
サンプルデータのインポート





- VCFファイルのインポート時に、親子関係情報を指定

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

Compound Het Detection

- 複合ヘテロ接合体の検出

Variant PhoRank Gene Ranking

- ユーザー指定のヒト表現型名と関連する遺伝子のランキング付け

Genotype Zygosity

Variant Info			Proband (NA12878)					
Chr:Pos	Ref/Alt	Identifier	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Genotype Qualities (GQ)	0/1 Genotypes (GT)	Zygosity
1:11188155	G/A	?	0.321951	139,66	205	99	0/1	Heterozygous
1:16456013	T/C	?	0.483333	31,29	60	86.59	0/1	Heterozygous
1:17085564	A/G	rs3851921	0.346535	66,35	101	99	0/1	Heterozygous
1:17431438	G/A	?	0.386364	27,17	44	75.78	0/1	Heterozygous
1:21036222	T/C	rs79606717	0.354839	40,22	62	99	0/1	Heterozygous
1:22142474	T/G	rs202212506	0.43617	53,41	95	99	0/1	Heterozygous
1:22915753	T/C	rs606002	1	0,49	49	99	1/1	Homozygous Variant
1:23111504	T/G	rs202178471	0.3625	51,29	80	99	0/1	Heterozygous
1:26371682	T/C	?	0.545455	40,48	89	99	0/1	Heterozygous
1:28209268	A/G	?	0.323529	92,44	138	99	0/1	Heterozygous
1:28800380	C/A	rs61785974	0.508772	84,87	171	99	0/1	Heterozygous
1:29638016	A/G	rs201177009	0.422535	41,30	72	65.77	0/1	Heterozygous
1:33058676	C/A	?	0.423077	45,33	78	84.71	0/1	Heterozygous
1:38435268	T/G	?	0.393939	80,52	132	62.95	0/1	Heterozygous
1:39776026	A/G	?	0.382022	55,34	89	66.21	0/1	Heterozygous
1:39934315	T/G	rs200234738	0.416667	42,30	72	88.75	0/1	Heterozygous
1:45808994	G/C	?	0.377273	137,83	221	99	0/1	Heterozygous
1:46195375	T/C	rs28375469	1	0,93	93	99	1/1	Homozygous Variant
1:46870761	C/A	rs324420	1	0,95	95	99	1/1	Homozygous Variant

- Homozygous, Heterozygous, Hemizygousなどの接合性情報を、各変異ごとにテーブルに追加する
- 子供がHomo、両親がHeteroというフィルター条件にて、劣性ホモ変異を検出するワークフローを作成が可能

Mendel Error

Variant Info		Proband (NA12878)				Mother (NA12892)				Father (NA12891)			
Chr:Pos	Ref/Alt	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Mendel Error	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Mendel Error	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Mendel Error
1:879516	T/C	0.444444	30,24	54	Transmitted	?	?,?	?	?	0.410714	33,23	56	?
1:883899	T/G	0.545455	35,42	77	Transmitted	0.505263	47,48	95	?	?	?,?	?	?
1:888659	T/C	1	0,36	36	Transmitted	1	0,54	54	?	1	0,49	49	?
1:900560	A/C	0.484848	17,16	34	Transmitted	?	?,?	?	?	0.666667	15,30	45	?
1:949608	G/A	0.492537	34,33	68	Transmitted	0.46	27,23	50	?	0.373134	42,25	67	?
1:1225707	C/A	0.325581	29,14	43	de Novo Allele	?	?,?	?	?	?	?,?	?	?
1:1392552	C/G	0.652174	16,30	48	Transmitted	0.447917	53,43	98	?	0.661538	22,43	65	?
1:1423281	G/A	0.40625	38,26	64	Transmitted	0.535211	33,38	71	?	?	?,?	?	?
1:1423286	C/G	0.47619	44,40	84	Transmitted	?	?,?	?	?	0.423729	68,50	118	?
1:1562677	T/G	0.531915	22,25	47	de Novo Allele	?	?,?	?	?	?	?,?	?	?
1:1650845	G/A	0.979866	3,146	149	MIE	?	?,?	?	?	0.935484	8,116	124	?
1:1686040	G/T	1	0,103	103	Transmitted	0.526316	18,20	38	?	0.47191	47,42	89	?
1:1756864	T/C	0.751004	62,187	250	Transmitted	0.548	113,137	250	?	0.704	74,176	250	?
1:1900107	-/CTC	0.392857	119,77	197	Transmitted	?	?,?	?	?	0.728261	50,134	185	?
1:1900232	T/C	0.546296	49,59	109	Transmitted	?	?,?	?	?	0.988506	1,86	88	?
1:2444414	G/A	0.536424	70,81	153	Transmitted	0.534884	20,23	44	?	0.507576	65,67	132	?
1:2494330	G/A	0.508772	28,29	57	Transmitted	?	?,?	?	?	0.395349	52,34	86	?
1:2523042	C/G	0.378378	69,42	111	Transmitted	?	?,?	?	?	0.382716	100,62	162	?
1:2938408	AGA/-	0.358025	52,29	81	Transmitted	?	?,?	?	?	0.472222	38,34	72	?
1:2938924	T/G	0.532258	29,33	62	Transmitted	0.68	8,17	25	?	?	?,?	?	?
1:3328358	T/C	0.992481	1,132	133	MIE	?	?,?	?	?	1	0,136	136	?
1:3328659	C/T	0.458716	59,50	110	de Novo Allele	?	?,?	?	?	?	?,?	?	?
1:3389727	C/T	0.451613	34,28	62	Transmitted	1	0,94	94	?	?	?,?	?	?
1:3397062	G/A	0.487805	21,20	42	Transmitted	?	?,?	?	?	0.511628	21,22	44	?
1:3424388	G/A	0.484536	50,47	97	Transmitted	?	?,?	?	?	0.45614	62,52	114	?

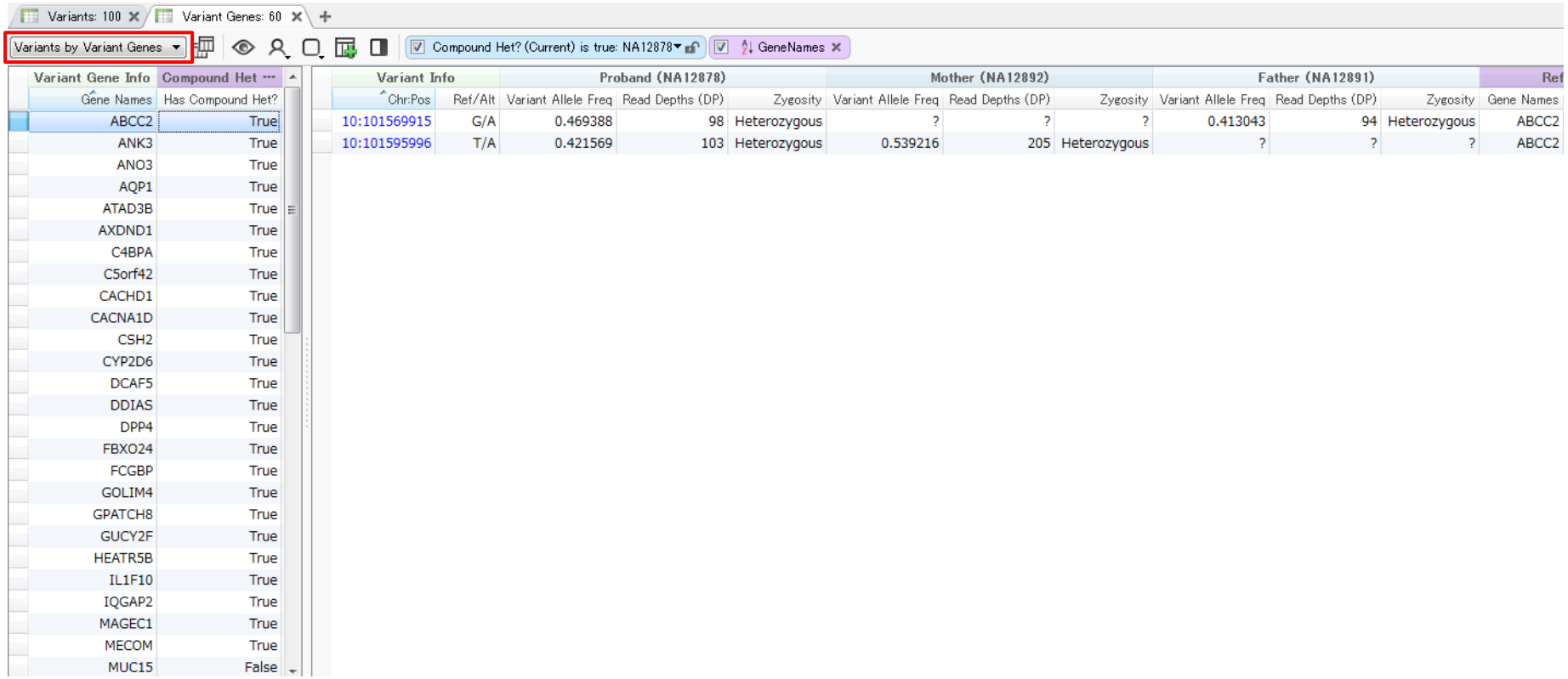
- 親子のサンプルを比較し、各変異ごとに、Transmitted, de Novo Alleleなどの情報をテーブルに追加する

Compound Het Detection

Proband (NA12878)	Mother (NA12892)	Father (NA12891)	Compound Het Variants f---		Compound Het Genes for Proband (NA12878)						RefSe
Genotypes	Genotypes	Genotypes	Compound Het?	Inherited From	Has Compound Het?	Inherited from Father	Inherited from Mother	Inherited Total	Hets In Both Parents	Second Smallest Freq	GeneNames
A/G	?	A/G	True	Father	True	1	1	2	0	0.00119808	PLB1
C/T	C/T	?	True	Mother	True	1	1	2	0	0.00119808	PLB1
G/T	?	G/T	True	Father	True	1	1	2	0	?	HEATR5B
C/G	C/G	?	True	Mother	True	1	1	2	0	?	HEATR5B
A/G	A/G	?	True	Mother	True	1	1	2	0	0.00199681	IL1F10
C/G	?	C/G	True	Father	True	1	1	2	0	0.00199681	IL1F10
A/C	?	A/C	True	Father	True	1	2	3	0	0.000199681	DPP4
C/G	C/G	?	True	Mother	True	1	2	3	0	0.000199681	DPP4
C/T	C/T	?	True	Mother	True	1	2	3	0	0.000199681	DPP4
C/G	C/G	?	True	Mother	True	1	1	2	0	0.000798722	TRAK1
A/C	?	A/C	True	Father	True	1	1	2	0	0.000798722	TRAK1
G/T	?	G/T	True	Father	True	1	1	2	1	0.000399361	CACNA1D
A/G	A/G	?	True	Mother	True	1	1	2	1	0.000399361	CACNA1D
-/TCC	?	-/TCC	True	Father	True	1	1	2	0	0.0267572	GOLIM4
C/T	C/T	?	True	Mother	True	1	1	2	0	0.0267572	GOLIM4
C/G	?	C/G	True	Father	True	1	1	2	0	0.000998403	MECOM
C/T	C/T	?	True	Mother	True	1	1	2	0	0.000998403	MECOM
G/T	G/T	?	True	Mother	True	1	1	2	0	0.019369	C5orf42
C/T	?	C/T	True	Father	True	1	1	2	0	0.019369	C5orf42
A/C	?	A/C	True	Father	True	1	1	2	0	0.00539137	IQGAP2
C/T	C/T	?	True	Mother	True	1	1	2	0	0.00539137	IQGAP2

- 1遺伝子内に、両親からそれぞれ別の変異を受け継いでいる遺伝子を検出する
- 検出時に、1000 GenomeやNHLBI 6500 Exome Projectなどの、民族ごとのアレル頻度データを指定し、低頻度変異のみを用いることも可能

Compound Het Detection

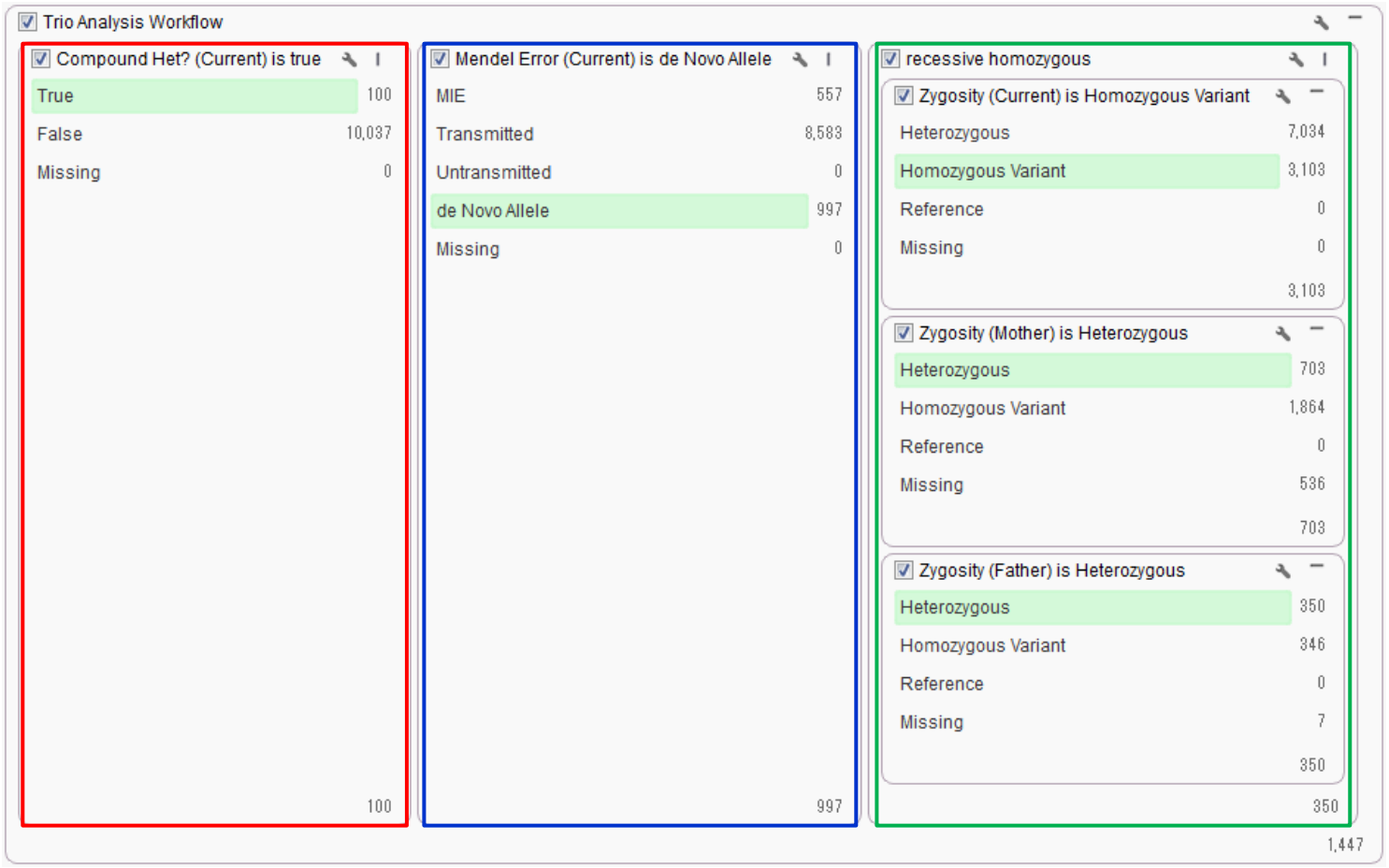


Variant Gene Info	Compound Het	Variant Info		Proband (NA12878)			Mother (NA12892)			Father (NA12891)			Ref
Gene Names	Has Compound Het?	Chr:Pos	Ref/Alt	Variant Allele Freq	Read Depths (DP)	Zygosity	Variant Allele Freq	Read Depths (DP)	Zygosity	Variant Allele Freq	Read Depths (DP)	Zygosity	Gene Names
ABCC2	True	10:101569915	G/A	0.469388	98	Heterozygous	?	?	?	0.413043	94	Heterozygous	ABCC2
ANK3	True	10:101595996	T/A	0.421569	103	Heterozygous	0.539216	205	Heterozygous	?	?	?	ABCC2
ANO3	True												
AQP1	True												
ATAD3B	True												
AXDND1	True												
C4BPA	True												
C5orf42	True												
CACHD1	True												
CACNA1D	True												
CSH2	True												
CYP2D6	True												
DCAF5	True												
DDIAS	True												
DPP4	True												
FBXO24	True												
FCGBP	True												
GOLIM4	True												
GPATCH8	True												
GUCY2F	True												
HEATR5B	True												
IL1F10	True												
IQGAP2	True												
MAGEC1	True												
MECOM	True												
MUC15	False												

- 遺伝子-変異の対応テーブルに表示を切り替え、遺伝子ごとに変異を確認することが可能

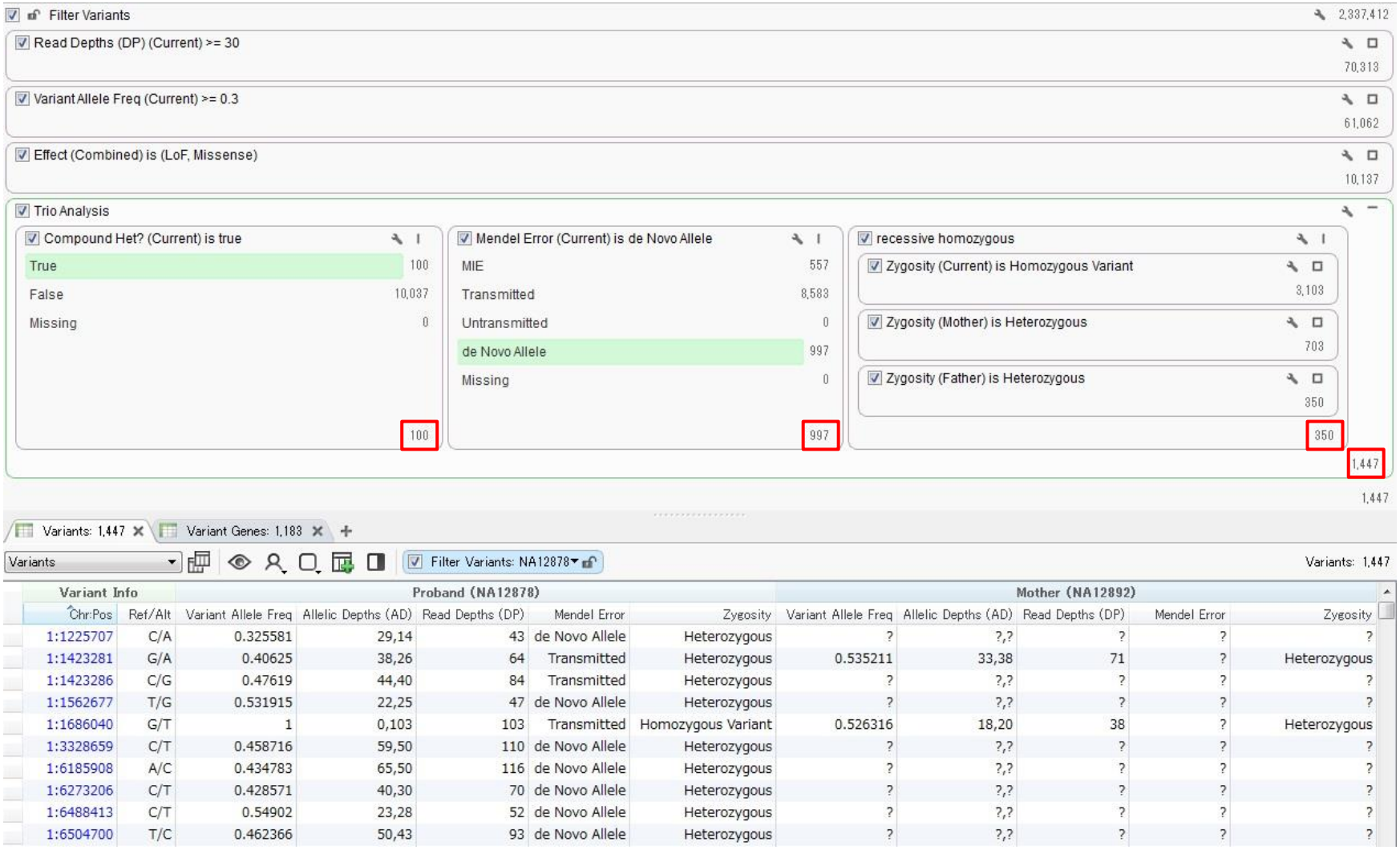
トリオ解析ワークフロー作成例

複合ヘテロ接合体 **De Novo変異** **劣性ホモ**



- 検索条件を分岐させ、複合ヘテロ接合体、De Novo変異、劣性ホモ変異などを同時に検出するワークフローを作成

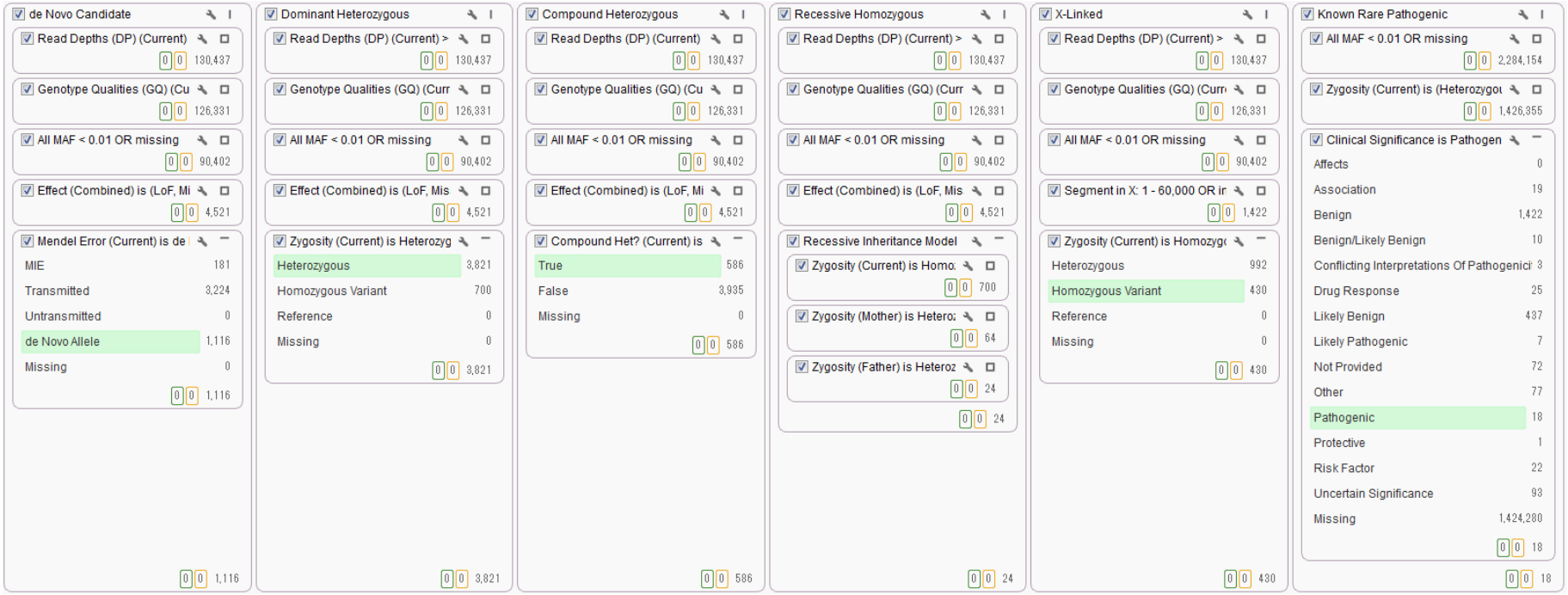
トリオ解析ワークフロー作成例



Variant Info		Proband (NA12878)					Mother (NA12892)				
Chr:Pos	Ref/Alt	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Mendel Error	Zygosity	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Mendel Error	Zygosity
1:1225707	C/A	0.325581	29,14	43	de Novo Allele	Heterozygous	?	?,?	?	?	?
1:1423281	G/A	0.40625	38,26	64	Transmitted	Heterozygous	0.535211	33,38	71	?	Heterozygous
1:1423286	C/G	0.47619	44,40	84	Transmitted	Heterozygous	?	?,?	?	?	?
1:1562677	T/G	0.531915	22,25	47	de Novo Allele	Heterozygous	?	?,?	?	?	?
1:1686040	G/T	1	0,103	103	Transmitted	Homozygous Variant	0.526316	18,20	38	?	Heterozygous
1:3328659	C/T	0.458716	59,50	110	de Novo Allele	Heterozygous	?	?,?	?	?	?
1:6185908	A/C	0.434783	65,50	116	de Novo Allele	Heterozygous	?	?,?	?	?	?
1:6273206	C/T	0.428571	40,30	70	de Novo Allele	Heterozygous	?	?,?	?	?	?
1:6488413	C/T	0.54902	23,28	52	de Novo Allele	Heterozygous	?	?,?	?	?	?
1:6504700	T/C	0.462366	50,43	93	de Novo Allele	Heterozygous	?	?,?	?	?	?

- 各フィルター結果の数字をクリックすることで、フィルター結果の変異データテーブルを確認

レディーメイドのトリオ解析ワークフロー



- 自分でワークフローを作成する以外に、レディーメイドのトリオ解析用ワークフローを使用し、De Novo変異、優性ヘテロ、劣性ホモ、複合ヘテロ接合体、X連鎖、既知の病原変異の同時検出が可能
- このワークフローをテンプレートにして、カスタマイズしたワークフローを作成することも可能

候補遺伝子リスト

Variants: 1,447 | Variant Genes: 1,188

Variants by Variant Genes | Filter Variants: NA12878

Variant Genes: 1,188
Variants: 1,447

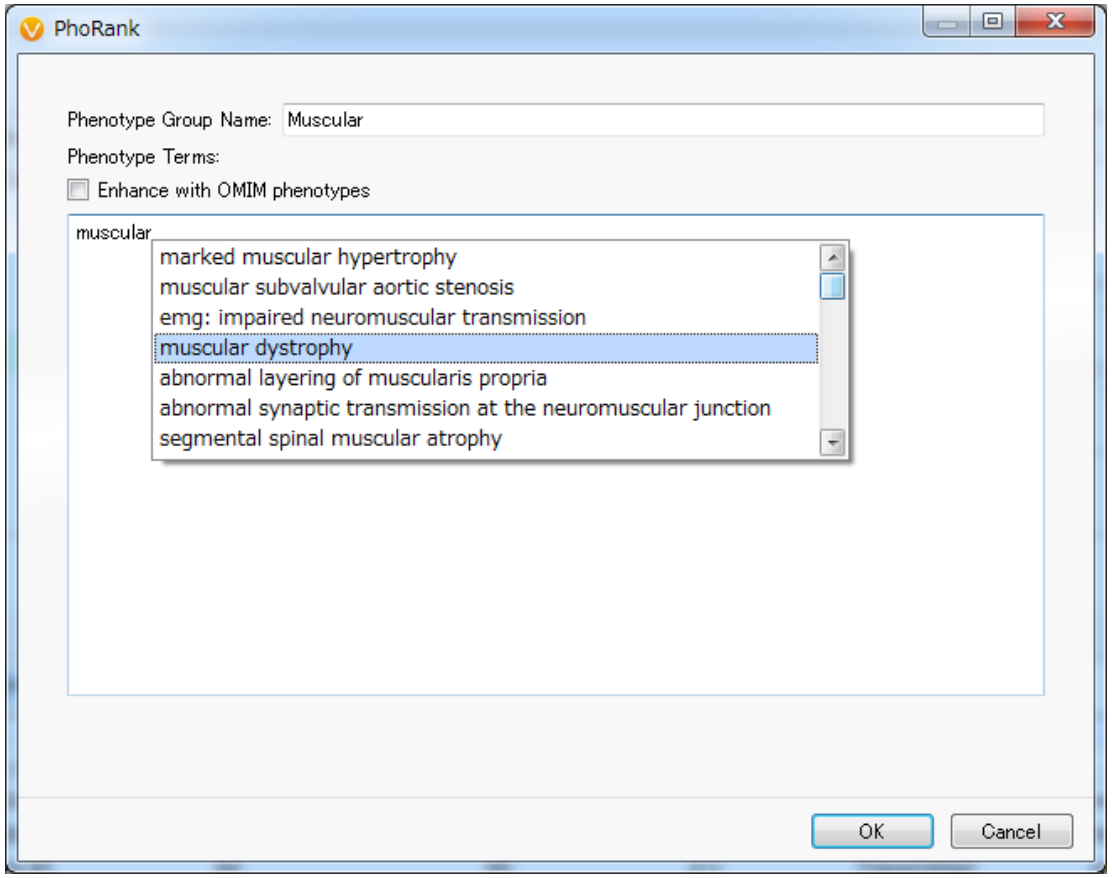
Variant Gene Info	Compound Het	Variant Info						Proband (NA12878)				Mother (NA12892)	
Gene Names	Has Compound Het?	Chr:Pos	Ref/Alt	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Mendel Error	Zygosity	Variant Allele Freq	Allelic Depths (AD)			
A2ML1	False	1:1225707	C/A	0.325581	29,14	43	de Novo Allele	Heterozygous	?	?,?			
ABCA6	False	1:1423281	G/A	0.40625	38,26	64	Transmitted	Heterozygous	0.535211	33,38			
ABCA13	False	1:1423286	C/G	0.47619	44,40	84	Transmitted	Heterozygous	?	?,?			
ABCC2	True	1:1562677	T/G	0.531915	22,25	47	de Novo Allele	Heterozygous	?	?,?			
ABCC3	False	1:1686040	G/T	1	0,103	103	Transmitted	Homozygous Variant	0.526316	18,20			
ABCC5	False	1:3328659	C/T	0.458716	59,50	110	de Novo Allele	Heterozygous	?	?,?			
ACRBP	False	1:6185908	A/C	0.434783	65,50	116	de Novo Allele	Heterozygous	?	?,?			
ACSL4	False	1:6273206	C/T	0.428571	40,30	70	de Novo Allele	Heterozygous	?	?,?			
ACSM4	False	1:6488413	C/T	0.54902	23,28	52	de Novo Allele	Heterozygous	?	?,?			
ADAMTS3	False	1:6504700	T/C	0.462366	50,43	93	de Novo Allele	Heterozygous	?	?,?			
ADAMTS14	False	1:7723498	C/G	0.445161	86,69	155	de Novo Allele	Heterozygous	?	?,?			
ADAMTS18	False	1:8073391	A/G	0.48951	73,70	143	de Novo Allele	Heterozygous	?	?,?			
ADCY7	False	1:9085065	G/T	0.309091	38,17	55	de Novo Allele	Heterozygous	?	?,?			
ADCY10	False	1:9171516	G/C	0.411765	50,35	85	de Novo Allele	Heterozygous	?	?,?			
ADGRB3	False	1:9778858	A/G	0.38806	41,26	67	de Novo Allele	Heterozygous	?	?,?			
ADGRD1	False	1:11188155	G/A	0.321951	139,66	205	de Novo Allele	Heterozygous	?	?,?			
ADGRE2	False	1:16456013	T/C	0.483333	31,29	60	de Novo Allele	Heterozygous	?	?,?			
ADGRF2	False	1:17085564	A/G	0.346535	66,35	101	de Novo Allele	Heterozygous	?	?,?			
ADGRL4	False	1:17431438	G/A	0.386364	27,17	44	de Novo Allele	Heterozygous	?	?,?			
ADGRV1	False	1:21036222	T/C	0.354839	40,22	62	de Novo Allele	Heterozygous	?	?,?			
ADNP2	False	1:22142474	T/G	0.43617	53,41	95	de Novo Allele	Heterozygous	?	?,?			
ADORA3	False	1:22915753	T/C	1	0,49	49	Transmitted	Homozygous Variant	0.514286	17,18			

候補遺伝子リスト

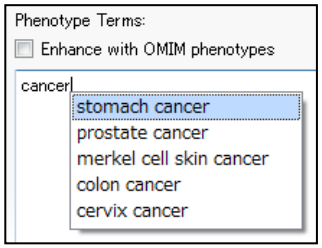
疾患との関連のランキング付け

- 遺伝子-変異の対応テーブルより、疾患の候補遺伝子リストを確認する
- 全エクソームシークエンスの場合、変異のフィルタリングを行っても候補遺伝子数が膨大になることがあり、Variant PhoRank Gene Rankingで遺伝子のランキング付けを行う

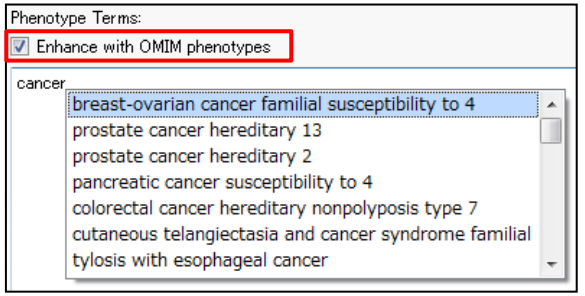
Variant PhoRank Gene Ranking



OMIMなし

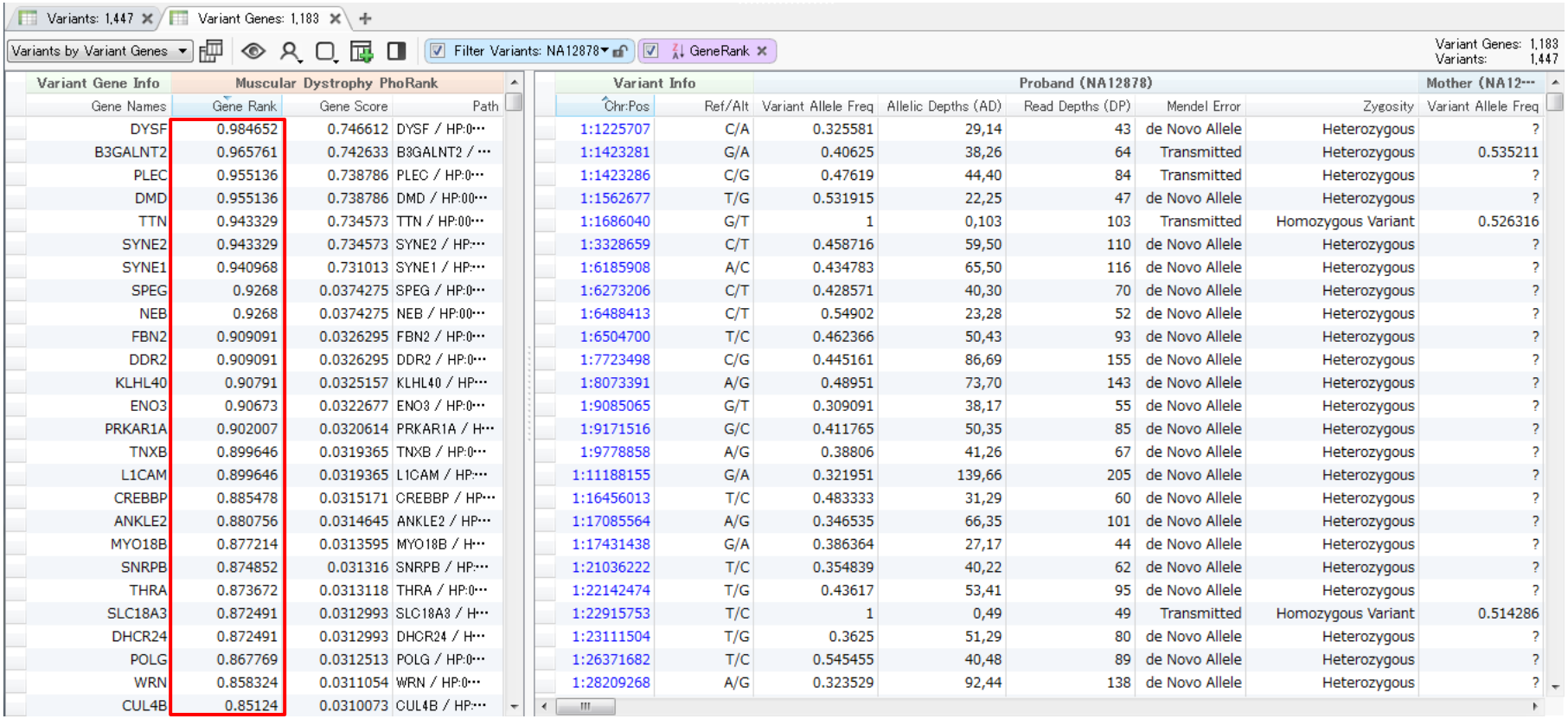


OMIMあり



- 任意のGO (Gene Ontology) あるいはHPO (Human Phenotype Ontology)のキーワードを入力すると、検索可能なTermの候補リストが表示される
- OMIMアドオンのライセンスを所持している場合、Termの候補リストを拡張するオプションが使用可能になる

Variant PhoRank Gene Ranking



Variant Gene Info / Muscular Dystrophy PhoRank

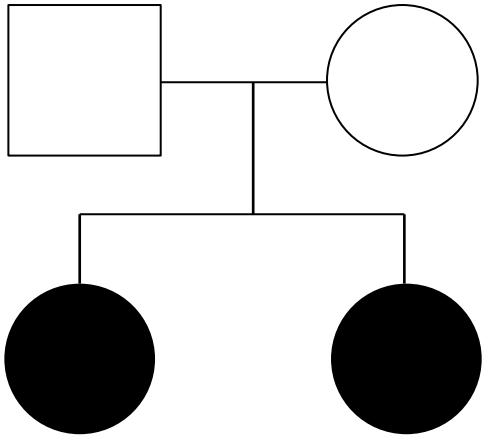
Gene Names	Gene Rank	Gene Score	Path
DYSF	0.984652	0.746612	DYSF / HP:0…
B3GALNT2	0.965761	0.742633	B3GALNT2 / …
PLEC	0.955136	0.738786	PLEC / HP:0…
DMD	0.955136	0.738786	DMD / HP:00…
TTN	0.943329	0.734573	TTN / HP:00…
SYNE2	0.943329	0.734573	SYNE2 / HP:…
SYNE1	0.940968	0.731013	SYNE1 / HP:…
SPEG	0.9268	0.0374275	SPEG / HP:0…
NEB	0.9268	0.0374275	NEB / HP:00…
FBN2	0.909091	0.0326295	FBN2 / HP:0…
DDR2	0.909091	0.0326295	DDR2 / HP:0…
KLHL40	0.90791	0.0325157	KLHL40 / HP:…
ENO3	0.90673	0.0322677	ENO3 / HP:0…
PRKAR1A	0.902007	0.0320614	PRKAR1A / H…
TNXB	0.899646	0.0319365	TNXB / HP:0…
L1CAM	0.899646	0.0319365	L1CAM / HP:…
CREBBP	0.885478	0.0315171	CREBBP / HP:…
ANKLE2	0.880756	0.0314645	ANKLE2 / HP:…
MYO18B	0.877214	0.0313595	MYO18B / H…
SNRPB	0.874852	0.031316	SNRPB / HP:…
THRA	0.873672	0.0313118	THRA / HP:0…
SLC18A3	0.872491	0.0312993	SLC18A3 / H…
DHCR24	0.872491	0.0312993	DHCR24 / H…
POLG	0.867769	0.0312513	POLG / HP:0…
WRN	0.858324	0.0311054	WRN / HP:0…
CUL4B	0.85124	0.0310073	CUL4B / HP:…

Variant Info / Proband (NA12878) / Mother (NA12…

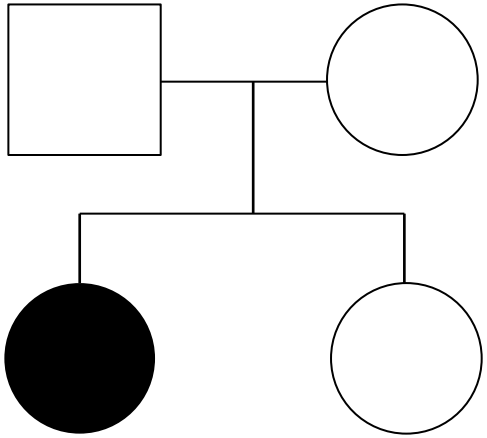
Chr:Pos	Ref/Alt	Variant Allele Freq	Allelic Depths (AD)	Read Depths (DP)	Mendel Error	Zygosity	Variant Allele Freq
1:1225707	C/A	0.325581	29,14	43	de Novo Allele	Heterozygous	?
1:1423281	G/A	0.40625	38,26	64	Transmitted	Heterozygous	0.535211
1:1423286	C/G	0.47619	44,40	84	Transmitted	Heterozygous	?
1:1562677	T/G	0.531915	22,25	47	de Novo Allele	Heterozygous	?
1:1686040	G/T	1	0,103	103	Transmitted	Homozygous Variant	0.526316
1:3328659	C/T	0.458716	59,50	110	de Novo Allele	Heterozygous	?
1:6185908	A/C	0.434783	65,50	116	de Novo Allele	Heterozygous	?
1:6273206	C/T	0.428571	40,30	70	de Novo Allele	Heterozygous	?
1:6488413	C/T	0.54902	23,28	52	de Novo Allele	Heterozygous	?
1:6504700	T/C	0.462366	50,43	93	de Novo Allele	Heterozygous	?
1:7723498	C/G	0.445161	86,69	155	de Novo Allele	Heterozygous	?
1:8073391	A/G	0.48951	73,70	143	de Novo Allele	Heterozygous	?
1:9085065	G/T	0.309091	38,17	55	de Novo Allele	Heterozygous	?
1:9171516	G/C	0.411765	50,35	85	de Novo Allele	Heterozygous	?
1:9778858	A/G	0.38806	41,26	67	de Novo Allele	Heterozygous	?
1:11188155	G/A	0.321951	139,66	205	de Novo Allele	Heterozygous	?
1:16456013	T/C	0.483333	31,29	60	de Novo Allele	Heterozygous	?
1:17085564	A/G	0.346535	66,35	101	de Novo Allele	Heterozygous	?
1:17431438	G/A	0.386364	27,17	44	de Novo Allele	Heterozygous	?
1:21036222	T/C	0.354839	40,22	62	de Novo Allele	Heterozygous	?
1:22142474	T/G	0.43617	53,41	95	de Novo Allele	Heterozygous	?
1:22915753	T/C	1	0,49	49	Transmitted	Homozygous Variant	0.514286
1:23111504	T/G	0.3625	51,29	80	de Novo Allele	Heterozygous	?
1:26371682	T/C	0.545455	40,48	89	de Novo Allele	Heterozygous	?
1:28209268	A/G	0.323529	92,44	138	de Novo Allele	Heterozygous	?

- 遺伝子ごとに、指定したキーワードの疾患との関連のランキングやスコアが表示される

トリオ以外のワークフロー作成例



- 両方の患者で共通に検出されたDe Novo変異や複合ヘテロ接合体を検出



- 健常者で検出されたDe Novo変異を差し引いたり、Heterozygousアリルをもつサンプル数を制限して複合ヘテロ接合体を検出

お問い合わせ先：フィルジェン株式会社

TEL: 052-624-4388 (9:00～17：00)

FAX: 052-624-4389

E-mail: biosupport@filgen.jp