TruSeq Custom Amplicon Low Inputキットを用いたターゲットリシーケンス – ドライ編 –

Mar 11, 2016

渡邊　大
イルミナ株式会社
テクニカルアプリケーションサイエンティスト

illumina®

本日の内容

- **BaseSpace cloud app, TruSeq Ampliconを用いたデータ解析**
  - シーケンスデータの参照
  - 変異の解析結果の参照

- **VariantStudioソフトウェアを用いたデータ解析**
  - データのインポート
  - 変異の絞り込み
  - アノテーション
  - データベースを利用した知見の入手

注：解析にはデモデータを使用しており、細かくは実際と異なる場合がございます。
データの参照と解析の流れをご案内する一例としてご参考下さい。

本日の内容

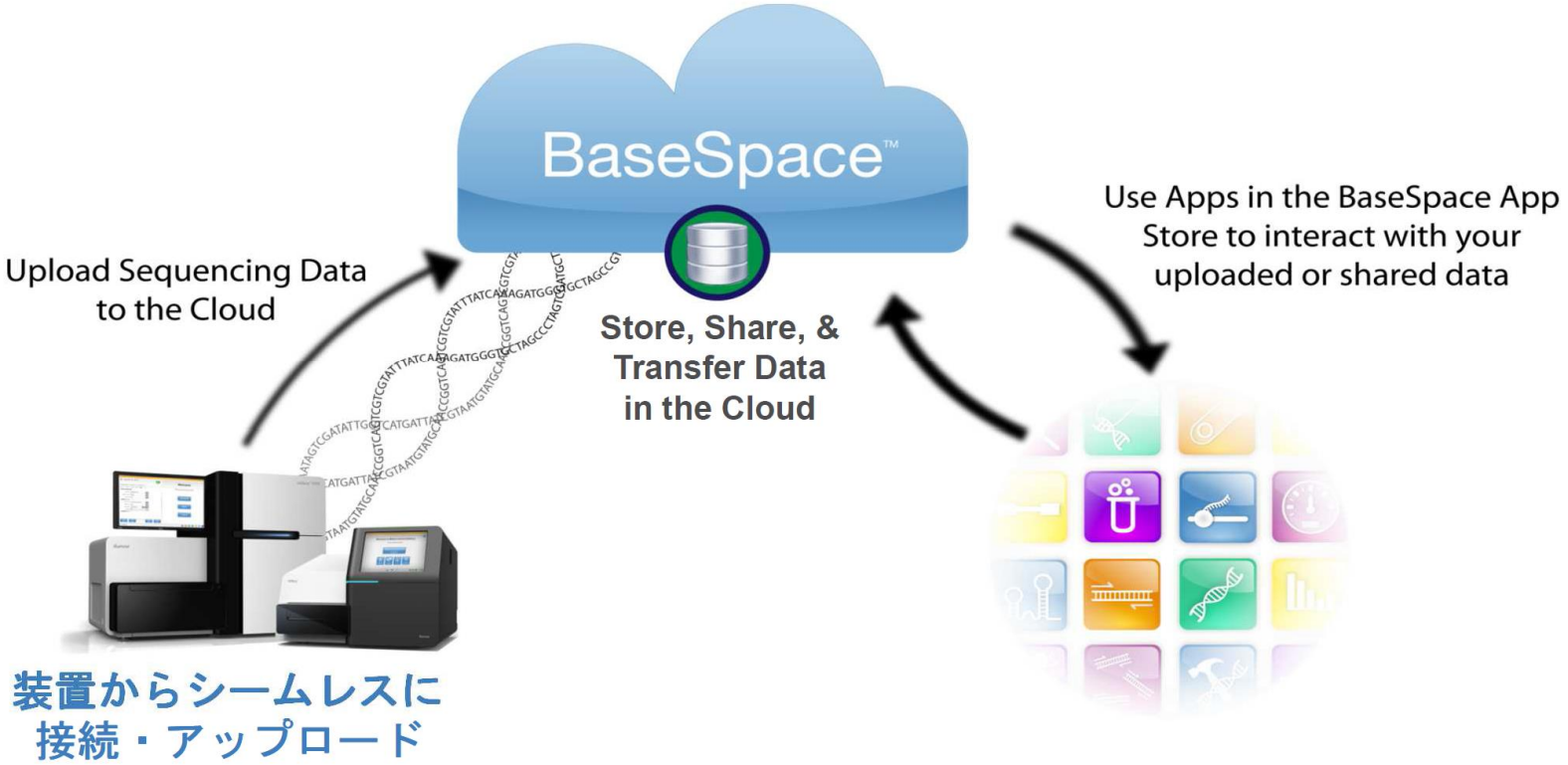
- **BaseSpace cloud app, TruSeq Ampliconを用いたデータ解析**
  - シーケンスデータの参照
  - 変異の解析結果の参照

- **VariantStudioソフトウェアを用いたデータ解析**
  - データのインポート
  - 変異の絞り込み
  - アノテーション
  - データベースを利用した知見の入手

注：解析にはデモデータを使用しており、細かくは実際と異なる場合がございます。
データの参照と解析の流れをご案内する一例としてご参考下さい。

BaseSpace cloudとは

Seamless upload, monitoring, storage, analysis, and collaboration



The Future: Cloud Solutions will be everywhere and specifically valuable to researchers

illumina®

BaseSpace appsとは

illumina

TruSeq Amplicon app

TruSeq Amplicon
Illumina, Inc. More info »
Free

CATEGORIES Resequencing, Targeted Sequencing, Variant Analysis

Description

The TruSeq Amplicon workflow enables streamlined analysis of Illumina TruSeq Amplicon panels. Alignment is performed using the banded Smith-Waterman algorithm in the targeted regions (specified in a manifest file). For variant calling, users can select GATK 1.6, Isaac Variant Caller, or the Illumina-developed Somatic Variant Caller. Both Ensembl and Refseq are available for annotation of the output files.

This App supports TruSeq Custom Amplicon analysis via import of custom manifests.

FOR RESEARCH USE ONLY.

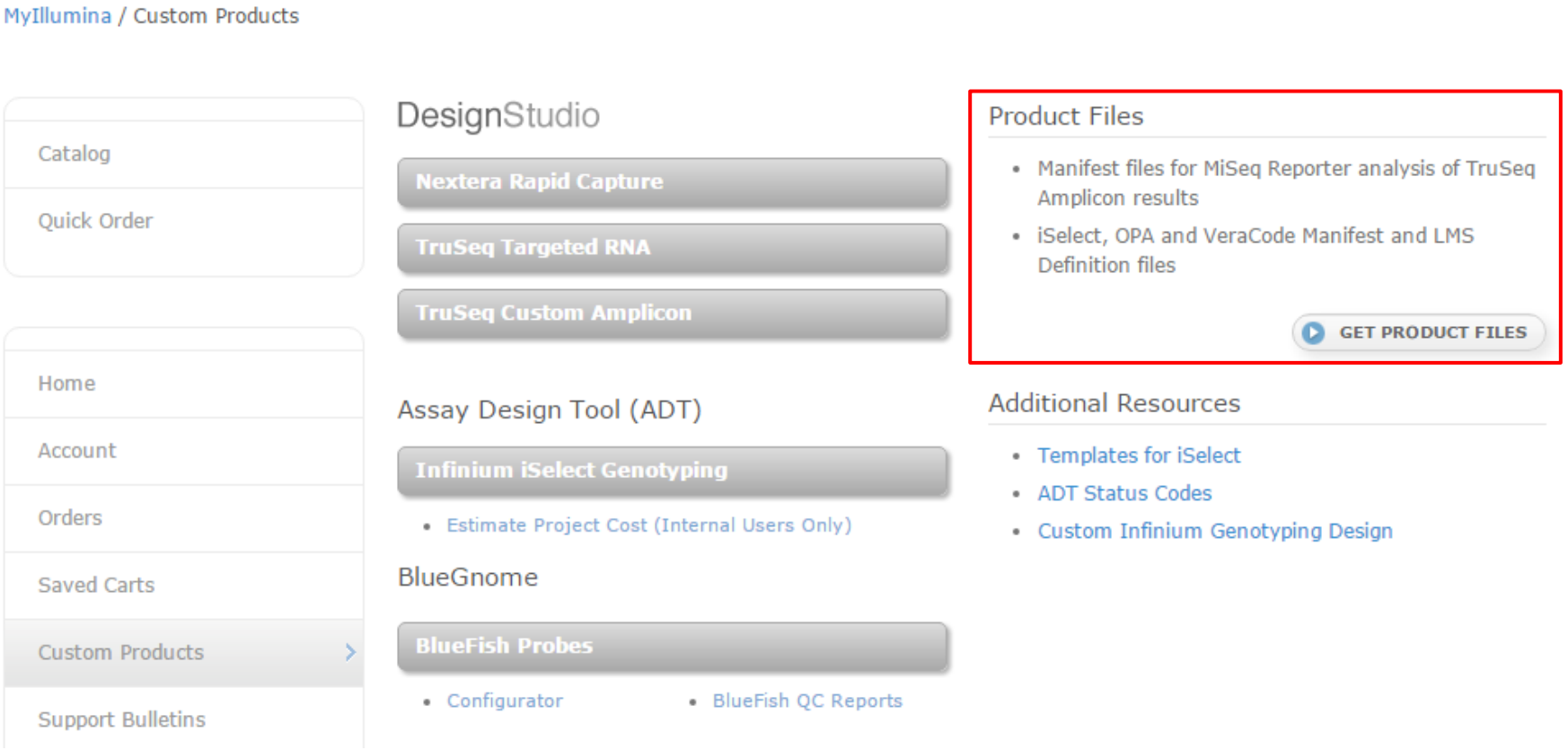
Current Limitations:

- This app supports Illumina TruSeq Custom Amplicon panels, and Illumina predesigned amplicon panels: the TruSeq Amplicon Cancer Panel, and the TruSight Myeloid Sequencing Panel.
- This app has been validated with Illumina MiSeq and NextSeq data only.

Technical Limitations:

- This app is designed for TruSeq Amplicon samples.
- Variants are only found in regions covered in the manifest.
- Reads must be at least 50 bases in length.
- No minimum number of reads, but use a reasonable input size to get your required coverage.
- App only supports running one manifest per analysis.
- Samples need to all be paired-end or all single-end; a warning is thrown if single-end.
- Samples all need to have the same read lengths

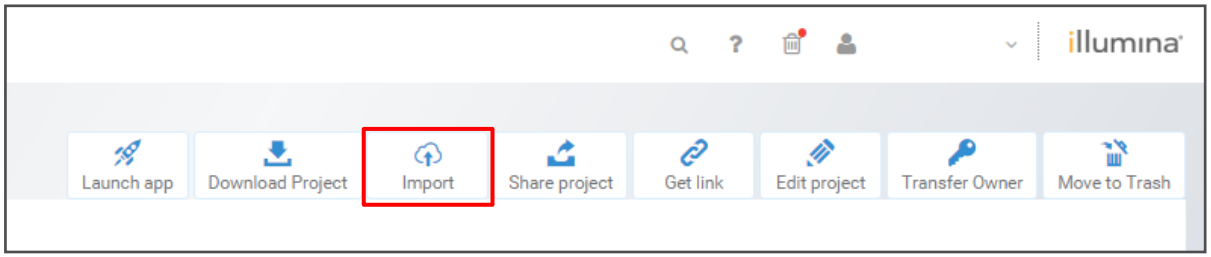
マニフェストファイルのダウンロード

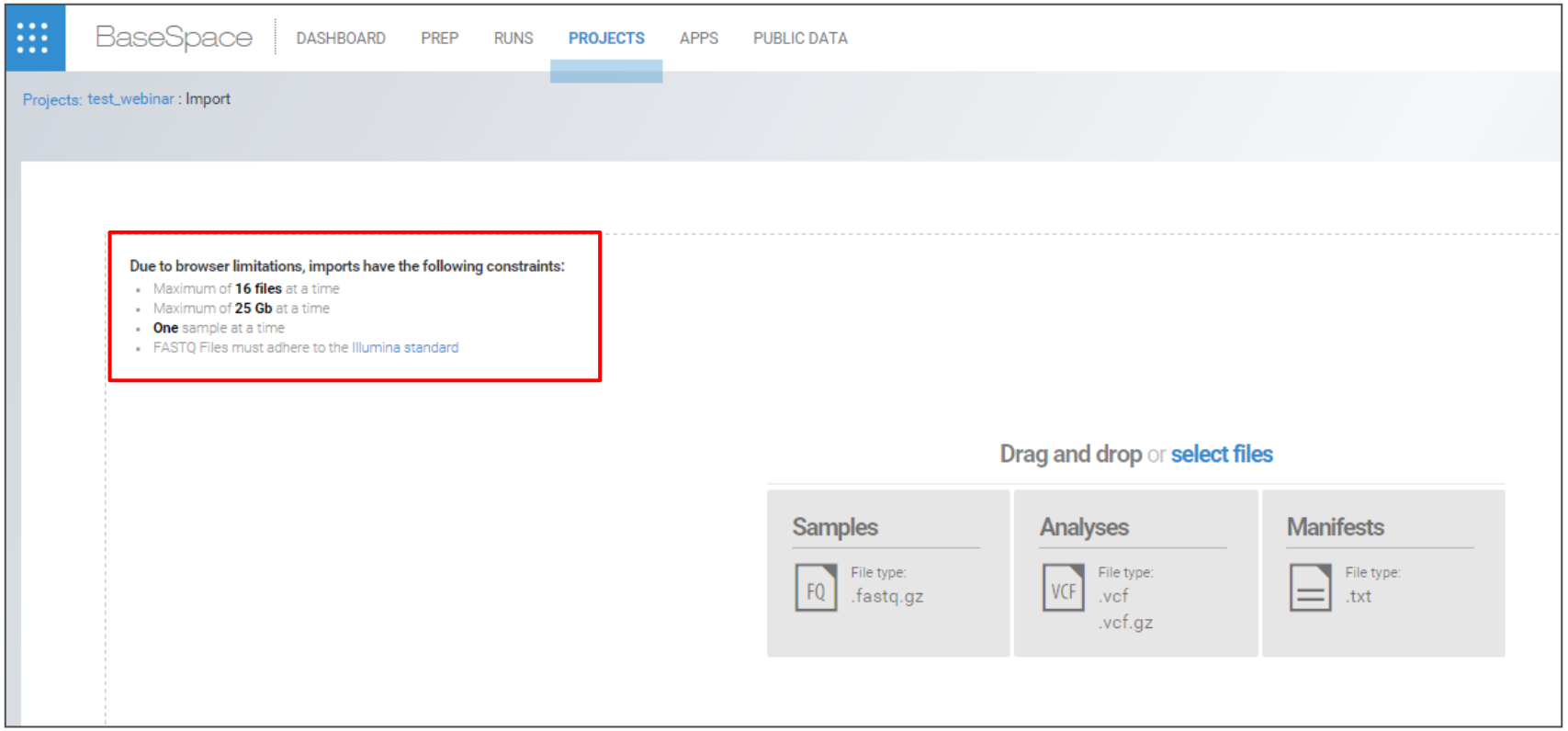


シーケンスの際に使用するマニフェストファイルはイルミナホームページからダウンロードし、BaseSpaceプロジェクト内にインポートする必要がある

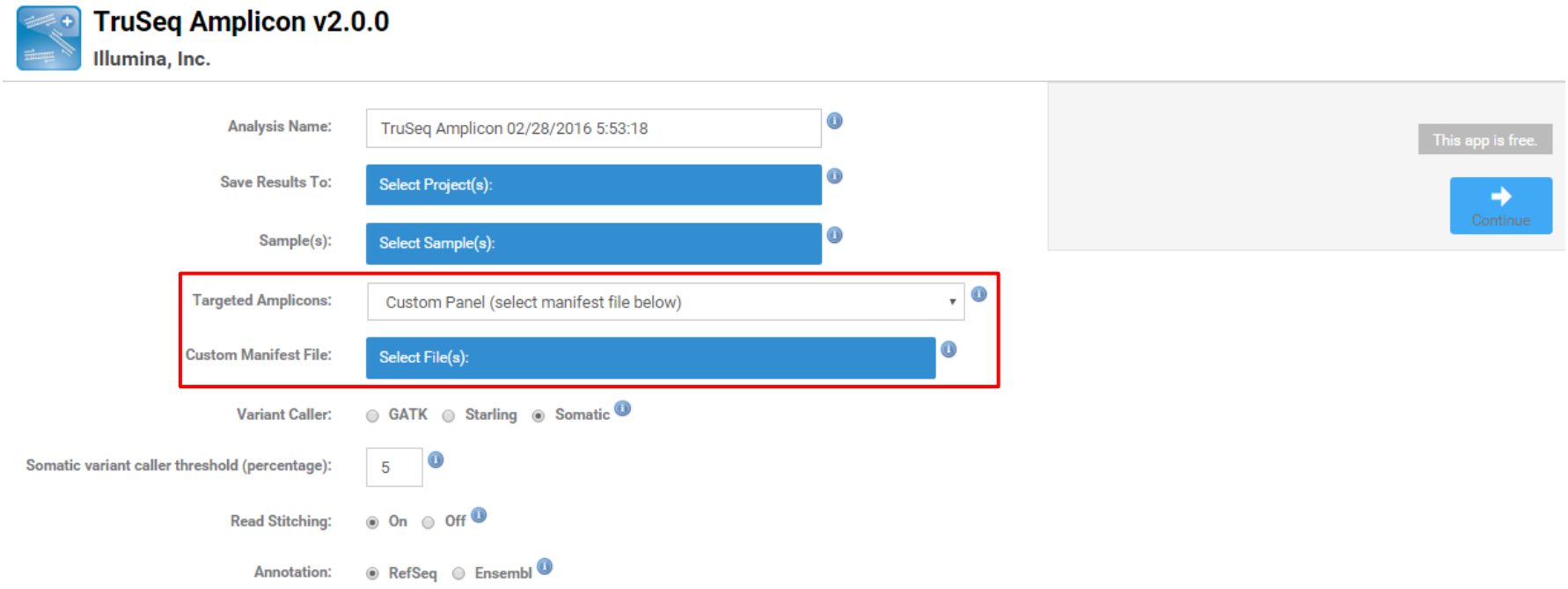
https://my.illumina.com/Custom/Index

BaseSpaceへのインポート



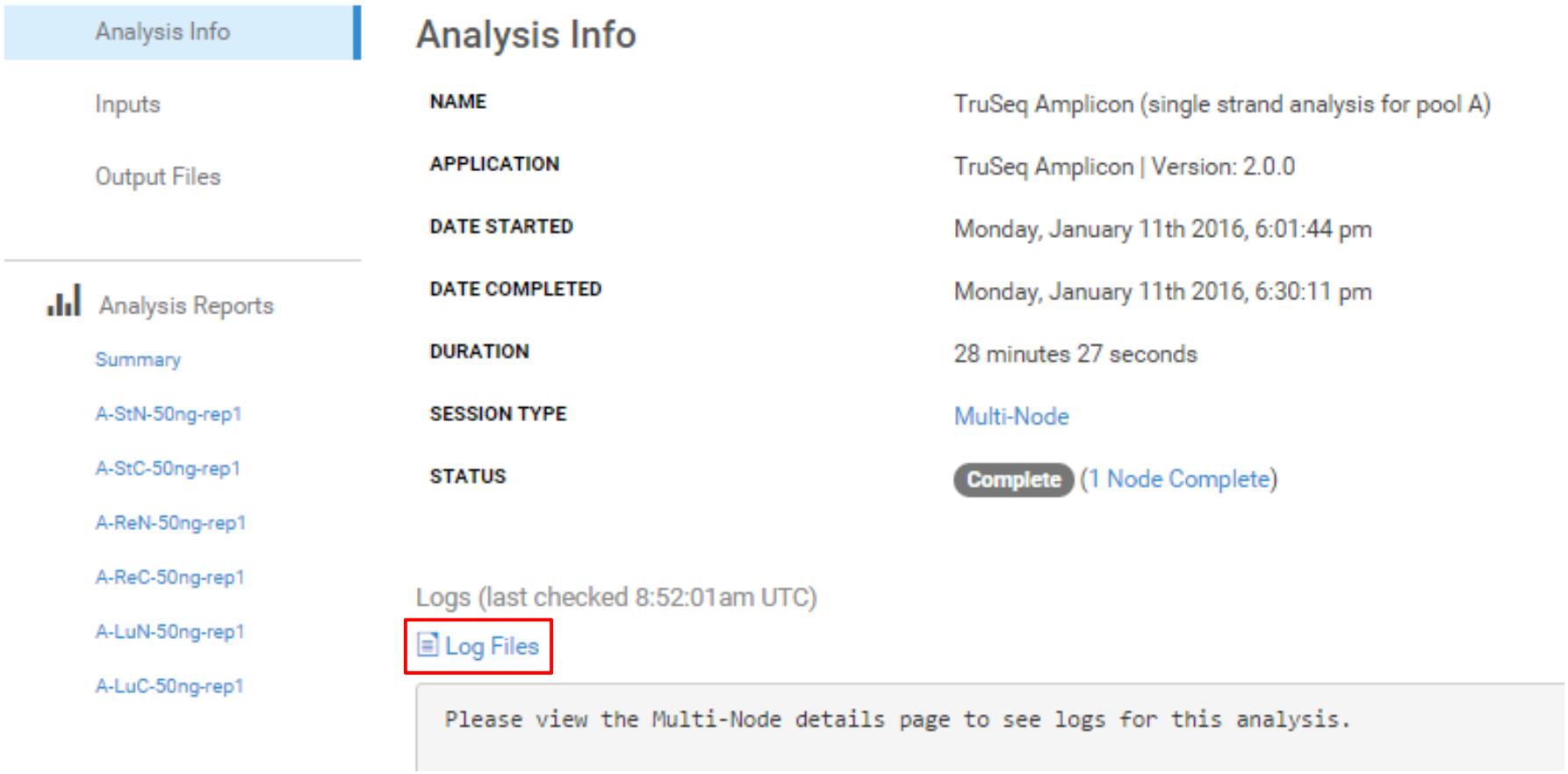


TruSeq Amplicon appへのデータインプット



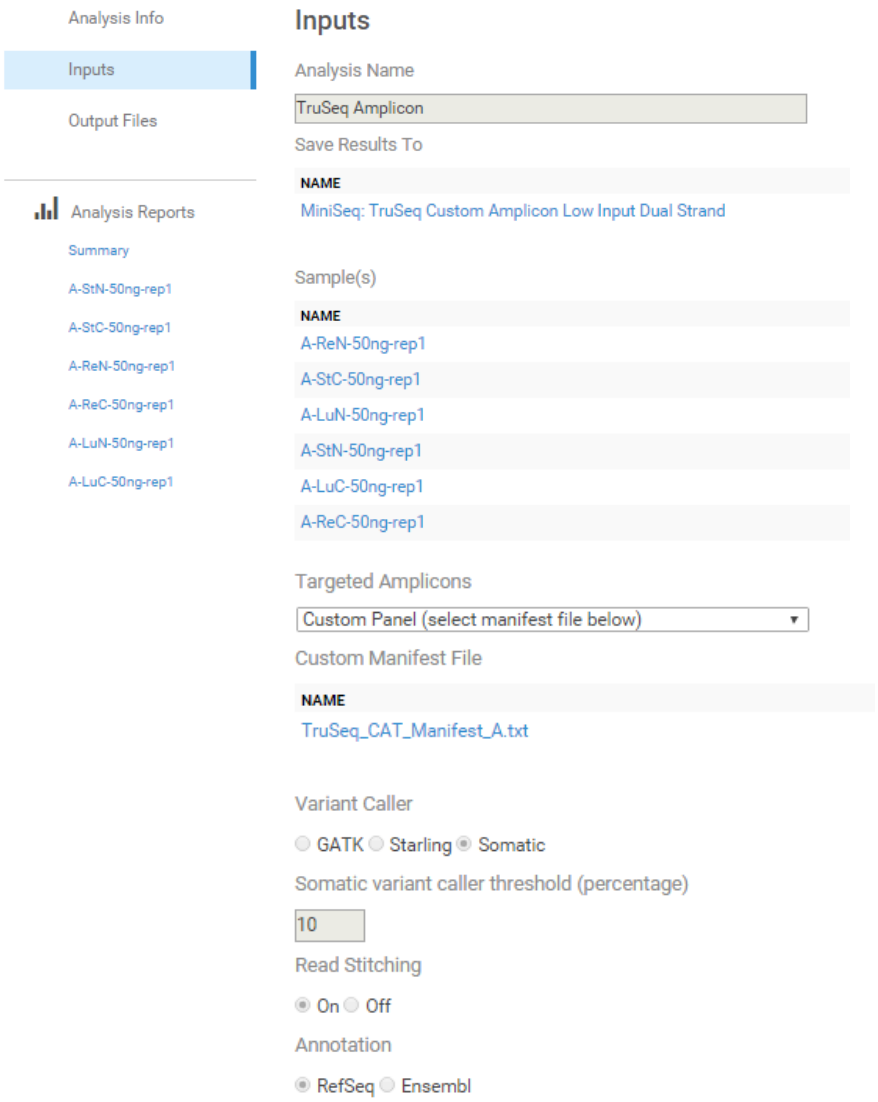
プロジェクト内にインポートしたマニフェストファイルを指定する

Analysis Infoタブ



解析ログはこのページから取得可能。テクニカルサポートからお願いする場合がございます

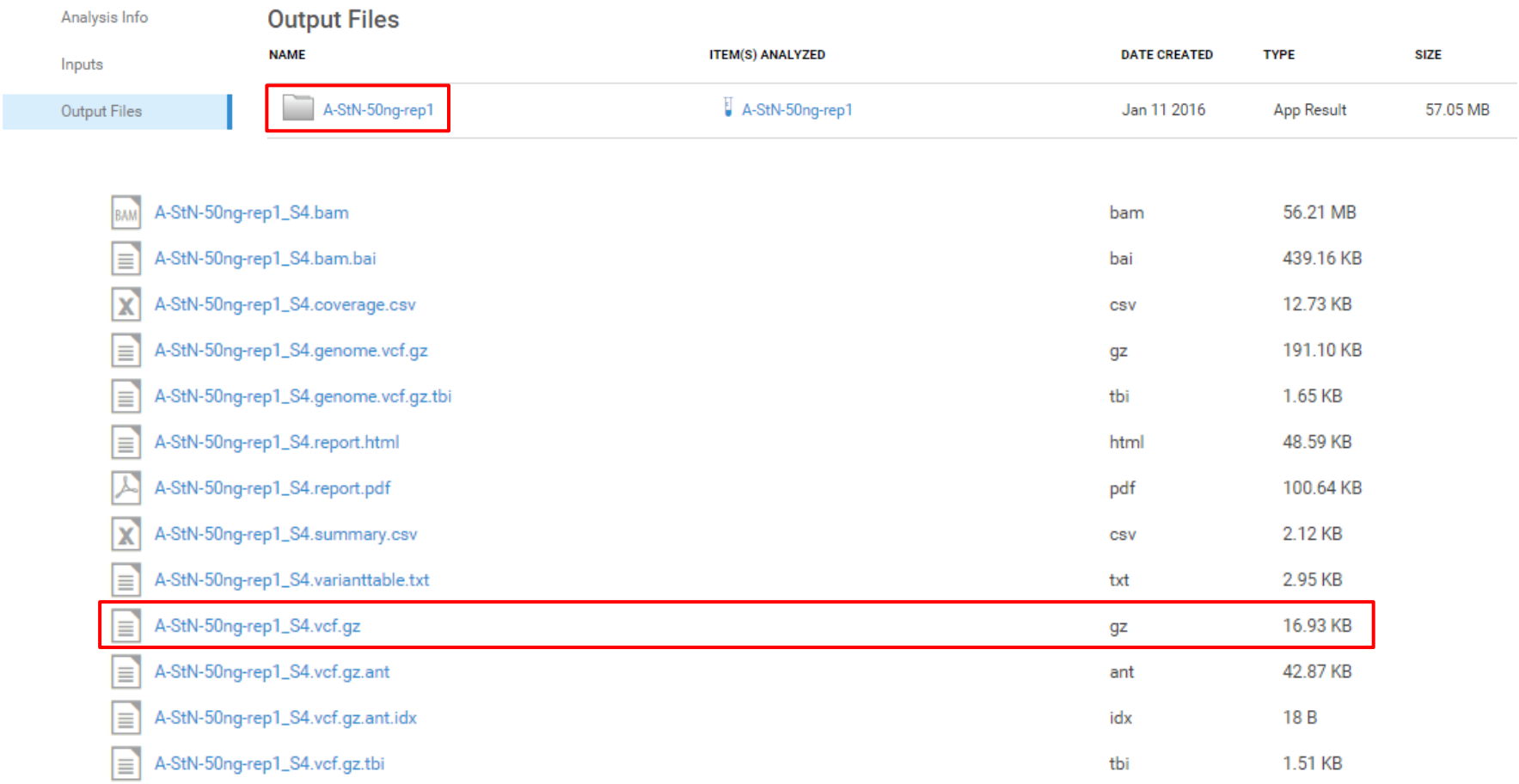
Inputsタブ



解析の条件を確認することができます

illumina

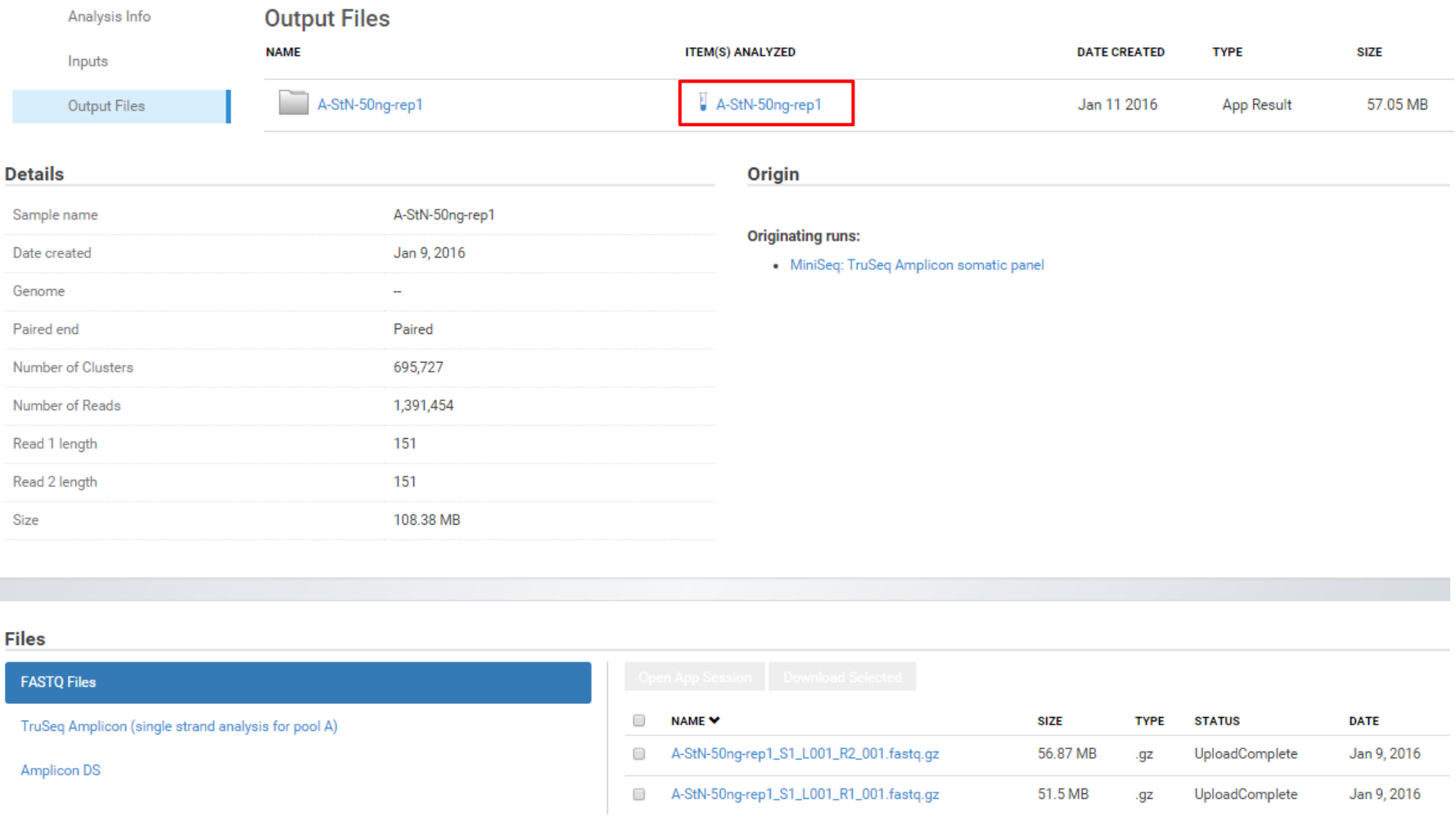
Output Filesタブ



下流の解析に使用するファイルの一覧が取得できます。.vcfファイルは後に使用します

illumina

Output Filesタブ



サンプルの詳細（fastqファイルなど）情報を取得することができます

Analysis Reports - Amplicon Summary

Amplicon Summary

Number of Amplicon Regions	Total Length of Amplicon Regions
125	18,423 bp

Read Level Statistics

Read	Total Aligned Reads	Percent Aligned Reads
1	616,943	88.68%
2	606,566	87.18%

Base Level Statistics

Read	Percent Q30	Total Aligned Bases	Percent Aligned Bases	Mismatch Rate
1	93.61%	92,277,982	92.14%	0.33%
2	89.10%	90,656,239	90.15%	0.34%

解析の結果はAnalysis Reportsタブから確認することができます

シーケンスとリファレンスゲノムへのマッピング結果

- リード数、マッピング率、クオリティスコアなど

Analysis Reports - Small Variant Summary

Small Variants Summary

	SNVs	Insertions	Deletions
Total Passing	24	0	1
Percent Found in dbSNP	62.50%	-	100.00%
Het/Hom Ratio	3	-	-
Ts/Tv Ratio	3	-	-

Variants by Sequence Context

	SNVs	Insertions	Deletions
Number in Genes	24	0	1
Number in Exons	13	0	0
Number in Coding Regions	11	0	0
Number in UTR Regions	2	0	0
Number in Splice Site Regions	1	0	0

Genes include exons, introns and UTR regions. Exons include coding and UTR regions. UTR regions include 5' and 3' UTR regions. Splice site regions include regions annotated as splice acceptor, splice donor, splice site or splice region.

SNV, insertion, deletionの結果

- 各変異の数
- ヘテロ／ホモ率
- Transition/Transversion率
- 変異の位置情報

Analysis Reports - Small Variant Summary

Variants by Consequence

	SNVs	Insertions	Deletions
Frameshifts	-	0	0
Non-synonymous	6	0	0
Synonymous	5	-	-
Stop Gained	0	0	0
Stop Lost	0	0	0

Variation consequences are calculated following the guidelines at
http://uswest.ensembl.org/info/genome/variation/predicted_data.html#consequences

SNV, insertion, deletionの結果

- 変異がもたらす結果のタイプ

Analysis Reports - Variants Table

Variants Table

Export (CSV)

Chr	Pos	Ref	Alt	Type	Context	Consequence	dbSNP	COSMIC	ClinVar	Qual	Alt Freq
chr1	1152...	G	A	SNV	Intergeni...	downstream_g...				100	0.393
chr1	1152...	G	A	SNV	Intergeni...	downstream_g...				100	0.395
chr2	2944...	G	A	SNV	Intron	splice_region_...	rs46226...			100	0.511
chr4	1808...	A	T	SNV	Intron,Int...	intron_variant,...				100	0.141
chr4	5514...	A	G	SNV	Coding	synonymous_v...	rs18737...	COSM143...		100	0.998
chr4	5559...	T	C	SNV	Intron	intron_variant	rs10086...			100	1.000
chr7	5525...	C	T	SNV	Coding,In...	synonymous_v...	rs22290...	COSM858...		100	0.494
chr9	2197...	C	T	SNV	Coding,In...	missense_varia...	rs37312...			100	0.320
chr10	8965...	C	T	SNV	Coding	synonymous_v...	rs15065...	COSM5168		100	0.505
chr10	8972...	C	T	SNV	Coding	synonymous_v...				100	0.573
chr10	8972...	A	G	SNV	Coding	synonymous_v...				100	0.576
chr10	8972...	T	C	SNV	3P_UTR	3_prime_UTR_...				100	0.450
chr11	534197	C	T	SNV	Intron,Int...	intron_variant,...	rs41258...			100	0.553
chr17	7577...	A	C	SNV	Intron	intron_variant	rs12951...			100	0.472
chr17	7577...	G	A	SNV	Intron	intron_variant	rs12947...			100	0.480
chr17	7578...	C	T	SNV	Intron	intron_variant	rs29094...			100	1.000
chr17	7579...	G	C	SNV	Coding	missense_varia...	rs10425...	COSM459...	non-path...	100	0.993

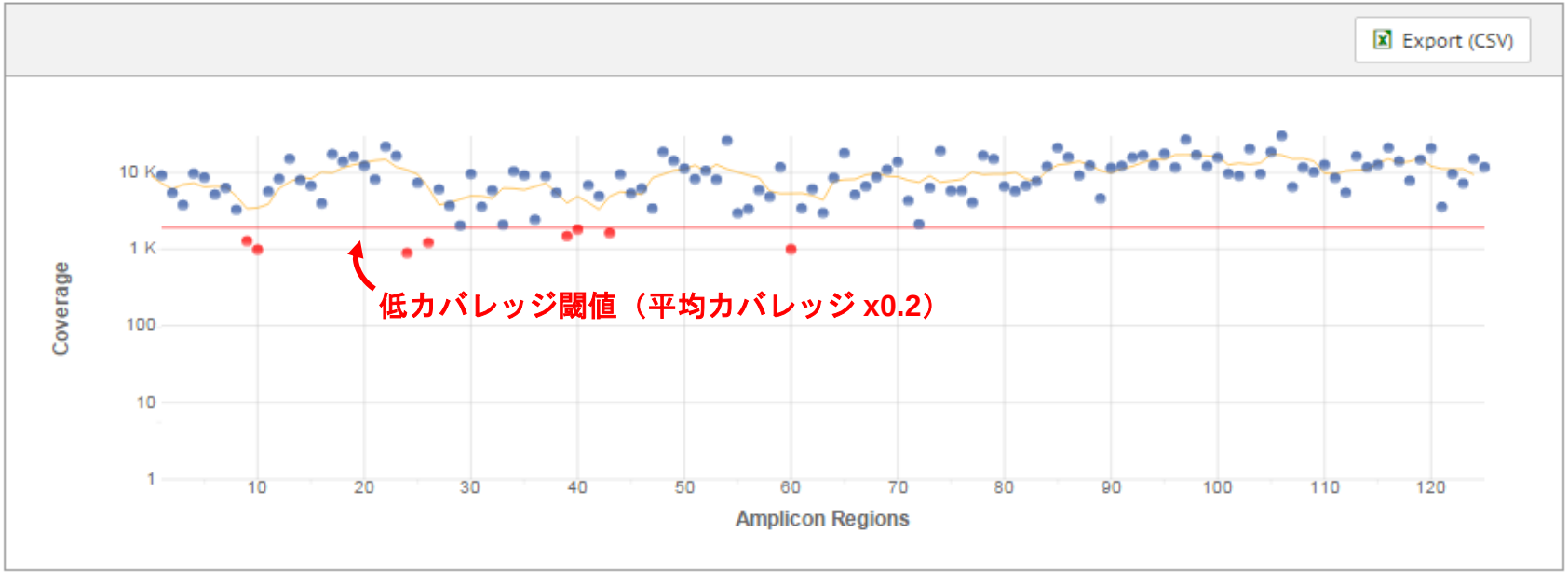
変異の詳細を確認でき、結果をダウンロードすることができます

Analysis Reports - Coverage Summary

Coverage Summary

Amplicon Mean Coverage	Uniformity of Coverage
9568.7	93.60% 80%がスペック

Coverage by Amplicon Region



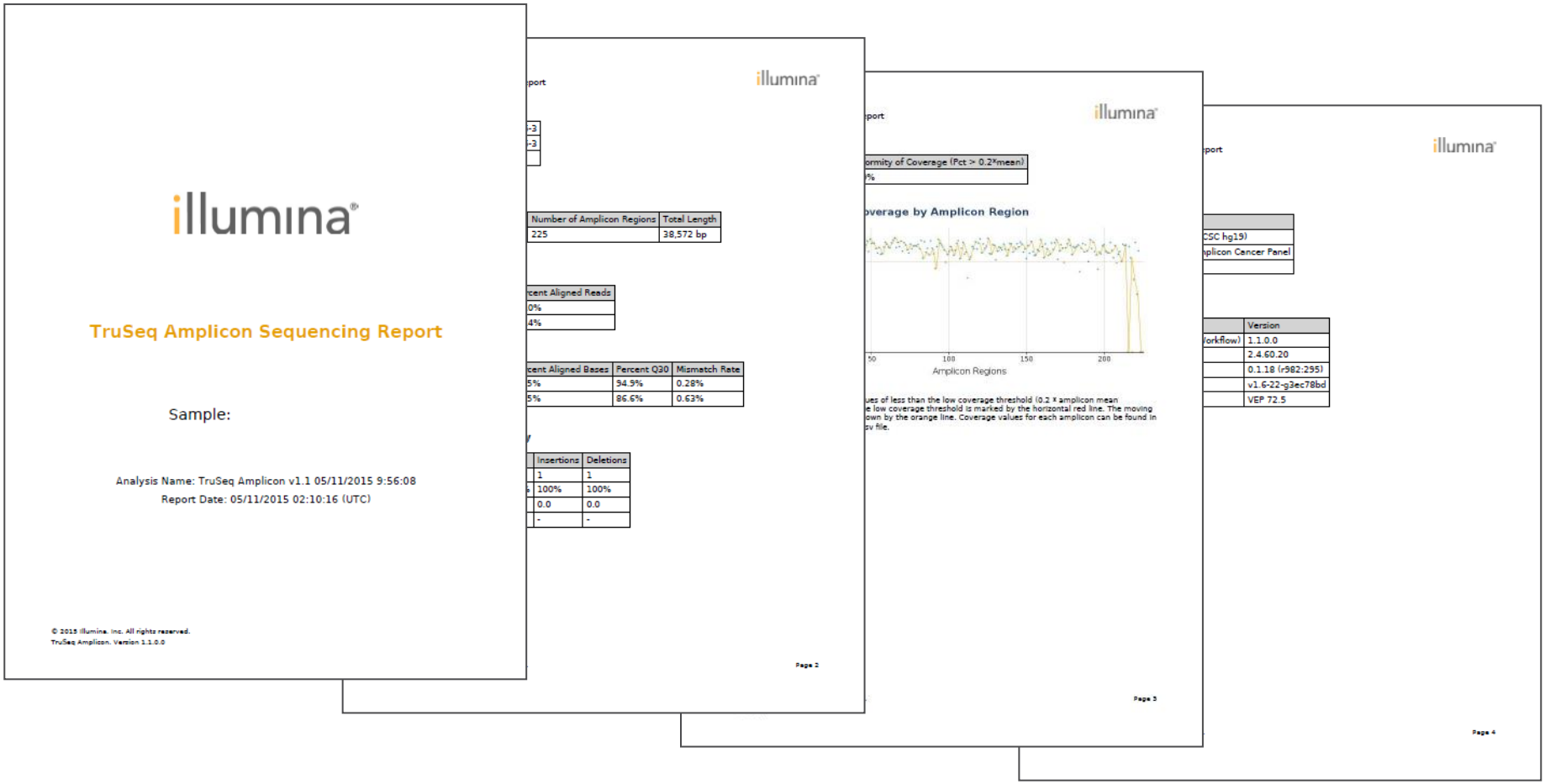
カバレッジ情報

- 平均カバレッジ、低カバレッジ変異
- マウスオンでどの変異か、個別のカバレッジが確認可能

Analysis Reports - PDF Summary Report

Results for Sample

PDF Summary Report



シーケンス結果はリンクからPDFで取得することができます

illumina

本日の内容

- BaseSpace cloud app, TruSeq Ampliconを用いたデータ解析
  - シーケンスデータの参照
  - 変異の解析結果の参照

- **VariantStudioソフトウェアを用いたデータ解析**
  - データのインポート
  - 変異の絞り込み
  - アノテーション
  - データベースを利用した知見の入手

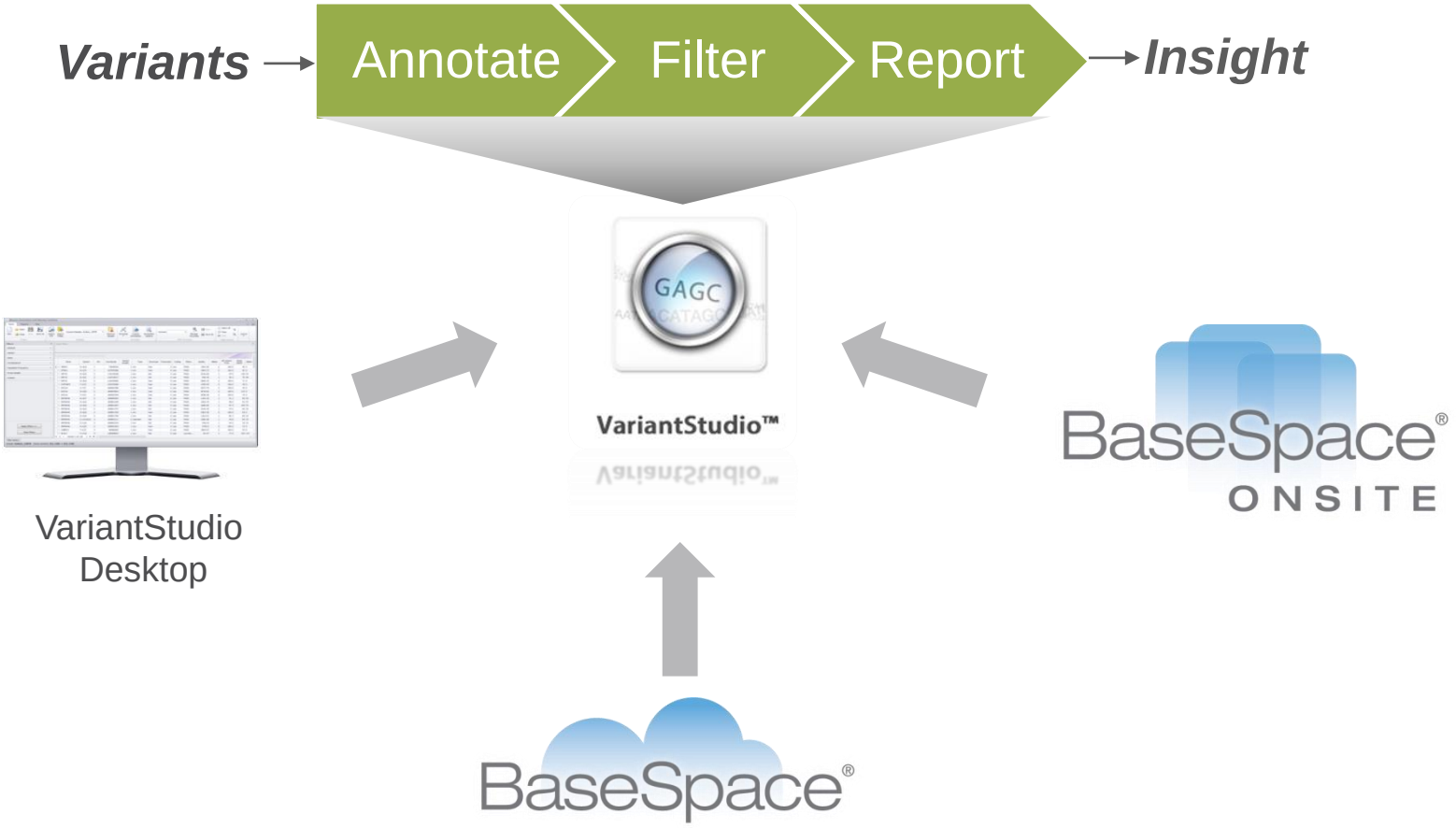
注：解析にはデモデータを使用しており、細かくは実際と異なる場合がございます。
データの参照と解析の流れをご案内する一例としてご参考下さい。

BaseSpace app：VariantStudioを用いたデータ解析

illumina

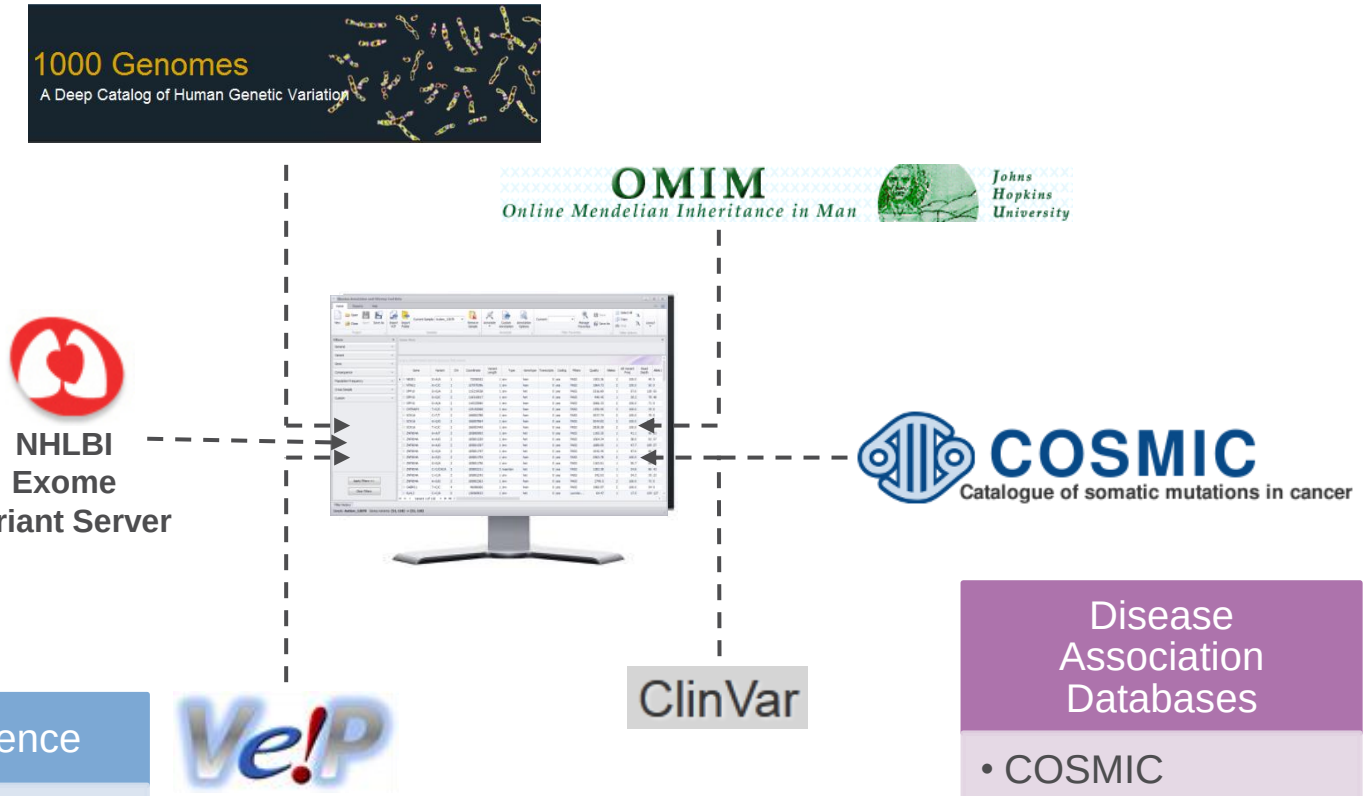
VariantStudioとは

- 変異解析のためのイルミナソフトウェア。
- リシーケンスappsで生成されるvcfファイルを読み込ませることで、コールした変異を**アノテーション**・**フィルター**し、**解析レポート**の形にするまでアシストします。



VariantStudioとは

- インターネット上の各種データベースの情報を簡便に統合し、知見を得ることができます



Population Variant Frequency

- 1000 Genomes
- Exome Variant Server

Transcript Consequence

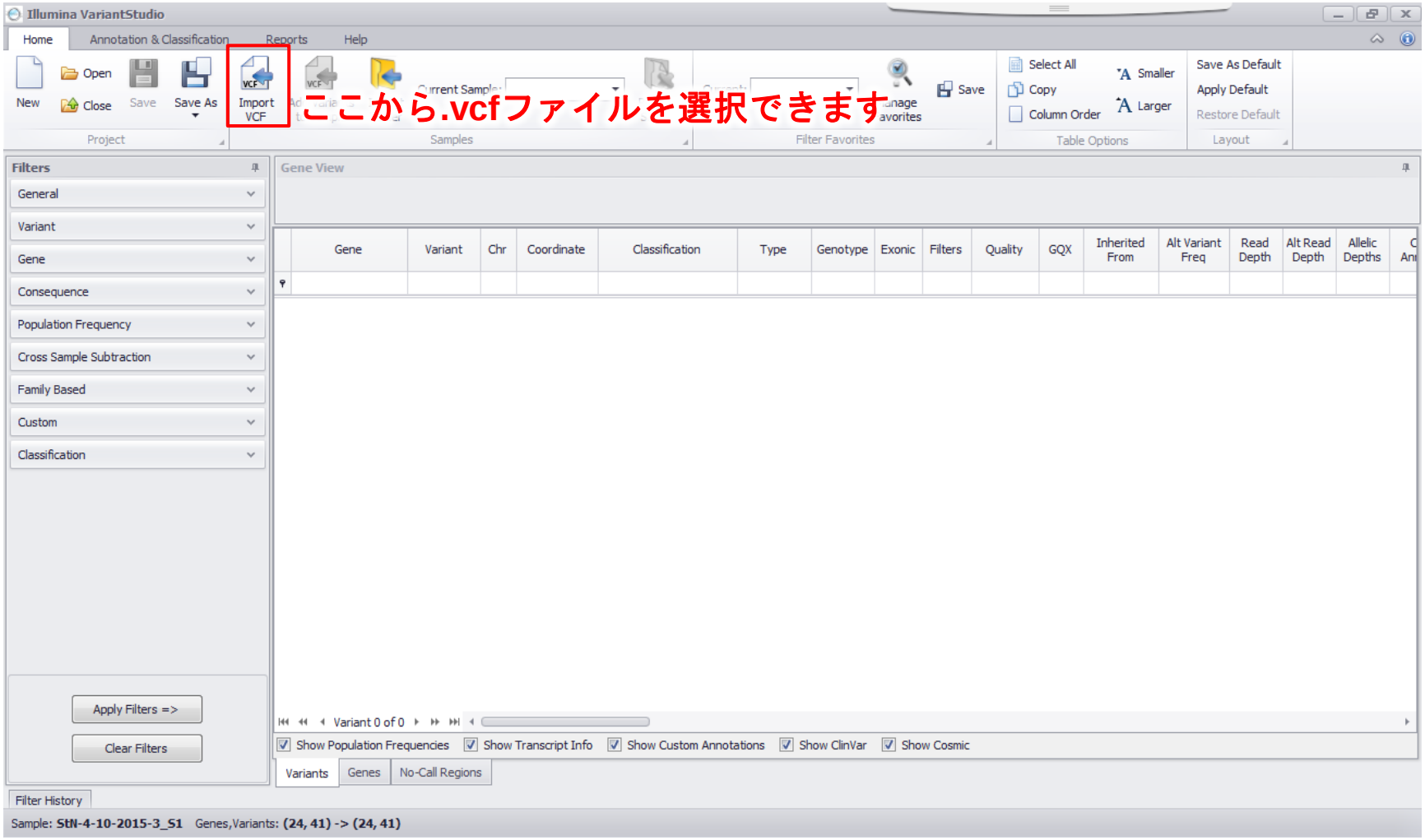
- Variant Effect Predictor
- Leverages databases such as RefSeq, PolyPhen, SIFT

Disease Association Databases

- COSMIC
- OMIM
- ClinVar

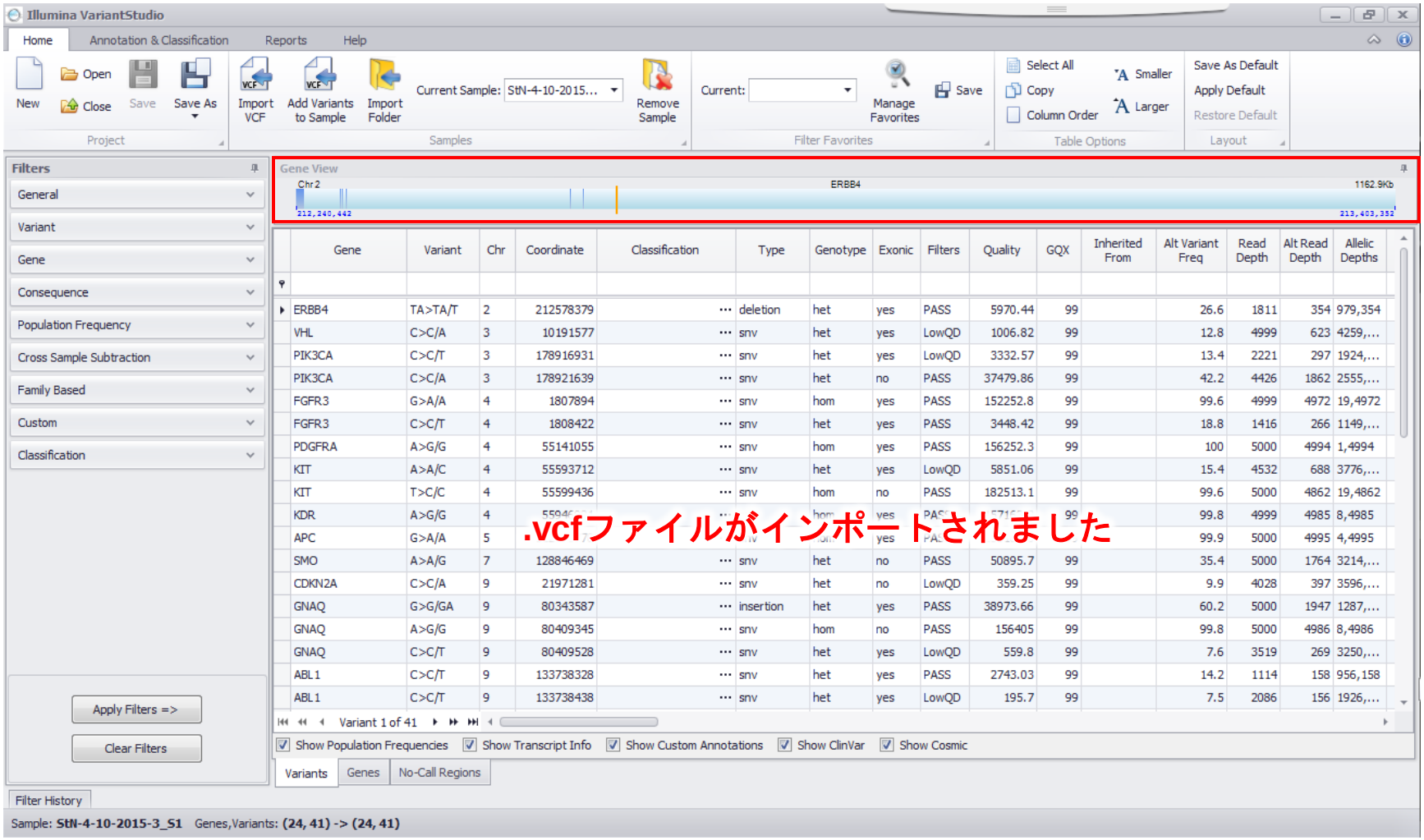
illumina

VariantStudioへのvcfファイルのインポート



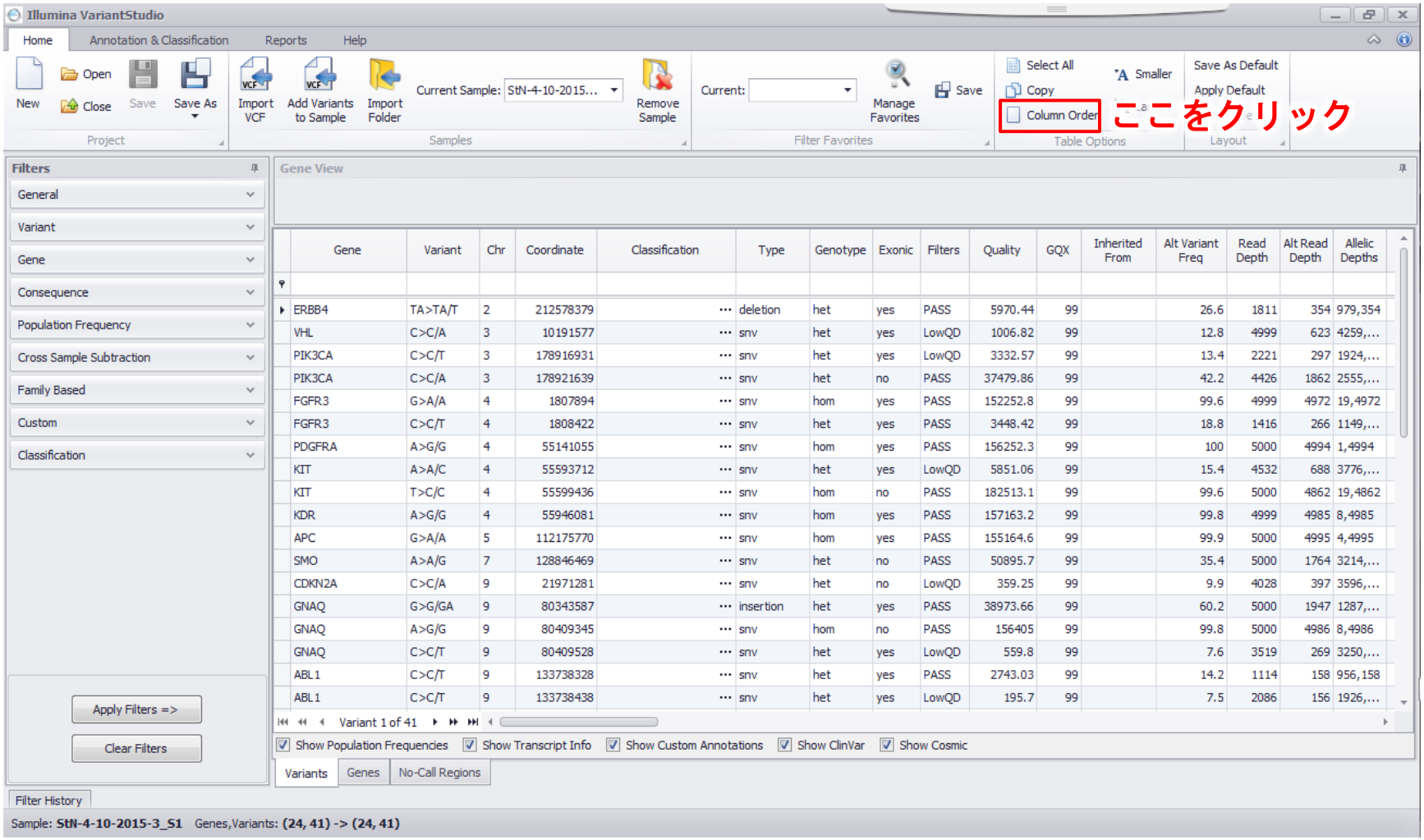
illumina

VariantStudioへのvcfファイルのインポート



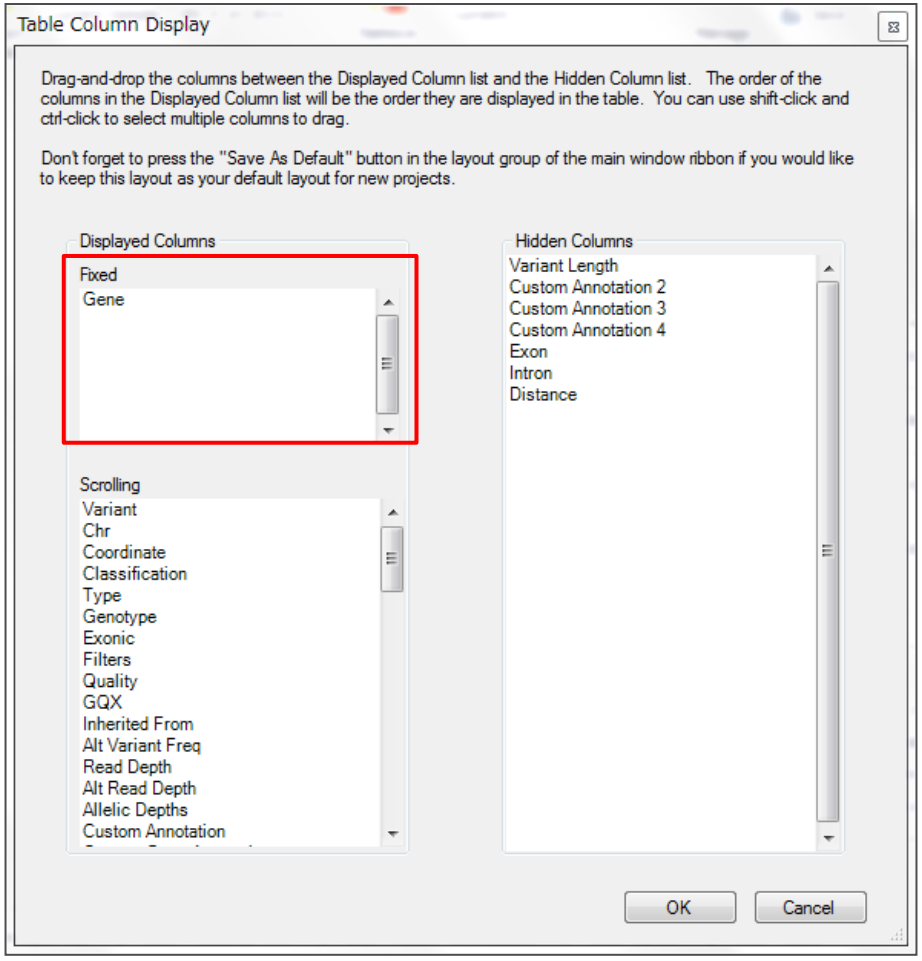
illumina

VariantStudioへのvcfファイルのインポート



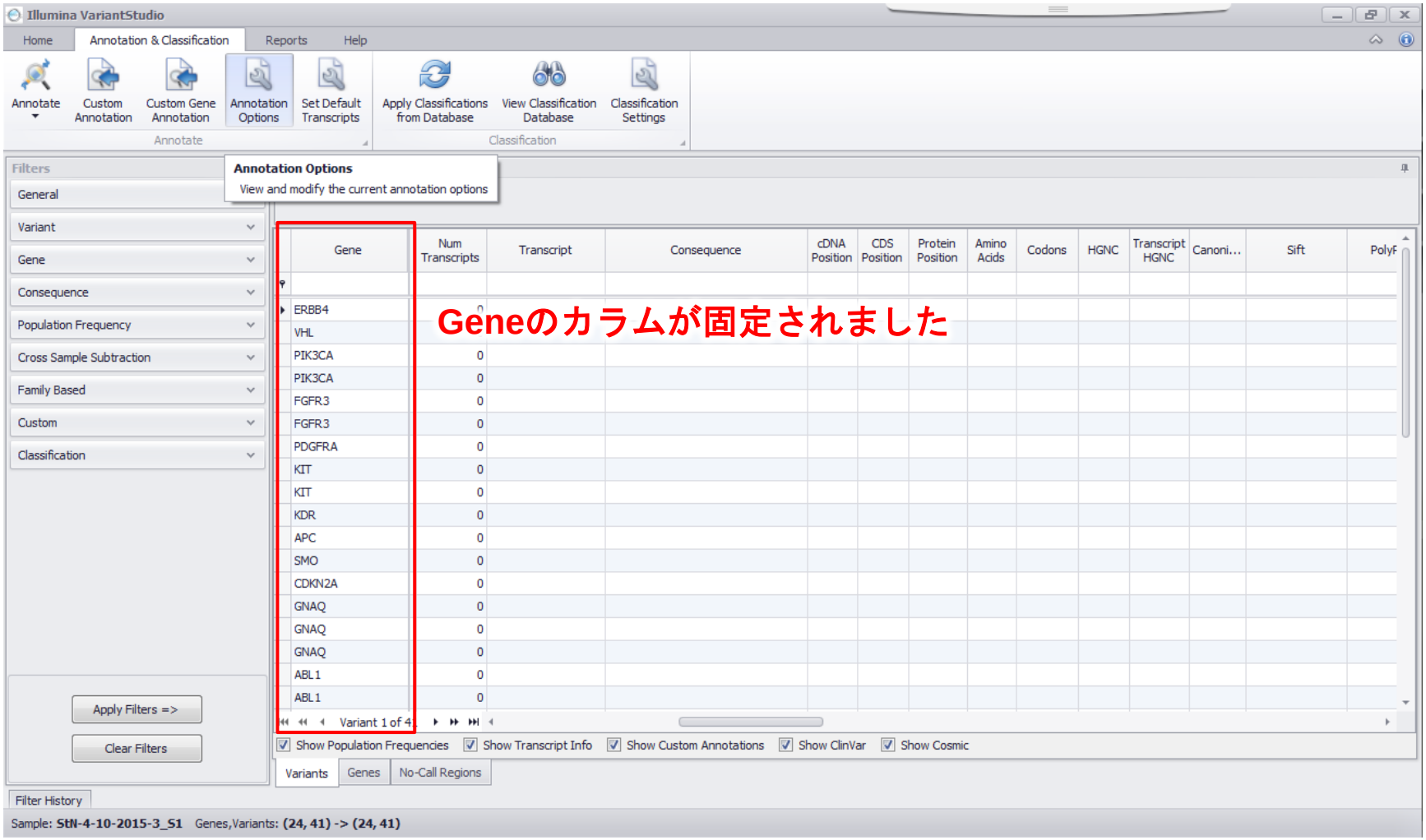
illumina

VariantStudioへのvcfファイルのインポート



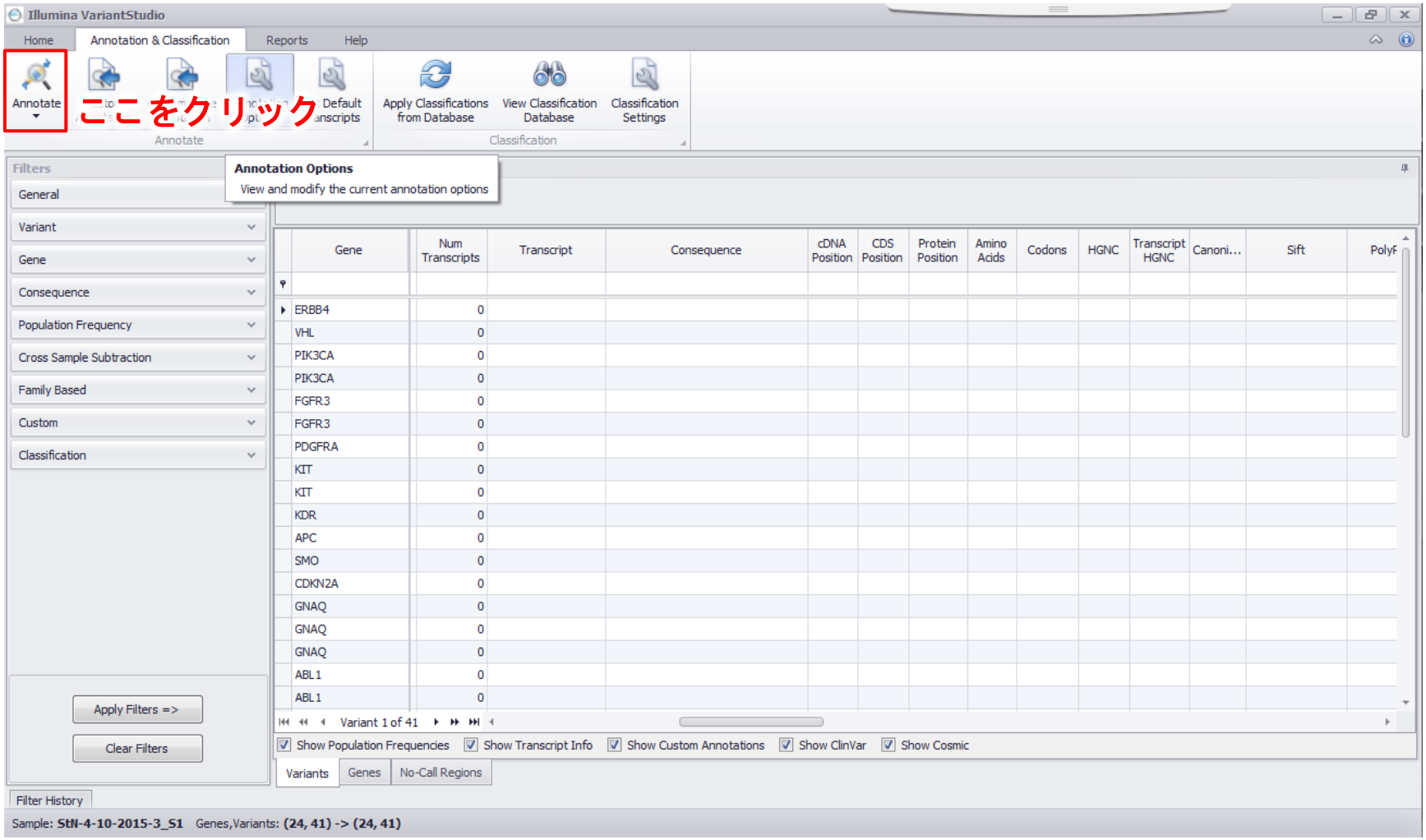
ドラッグで固定するカラムを選択することができます
‘Gene’の固定をお勧めします

Annotate: 変異のアノテーション



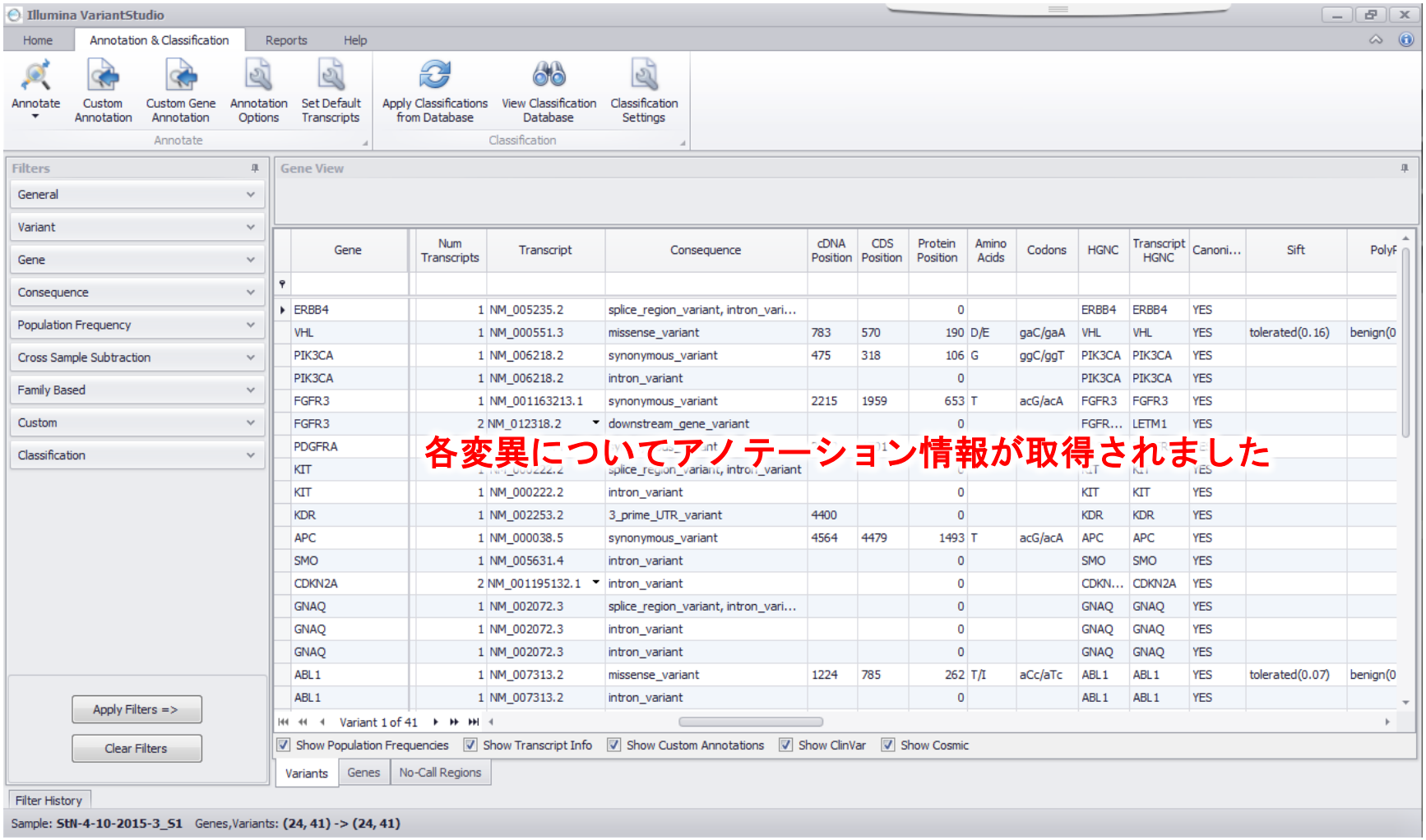
illumina

Annotate: 変異のアノテーション

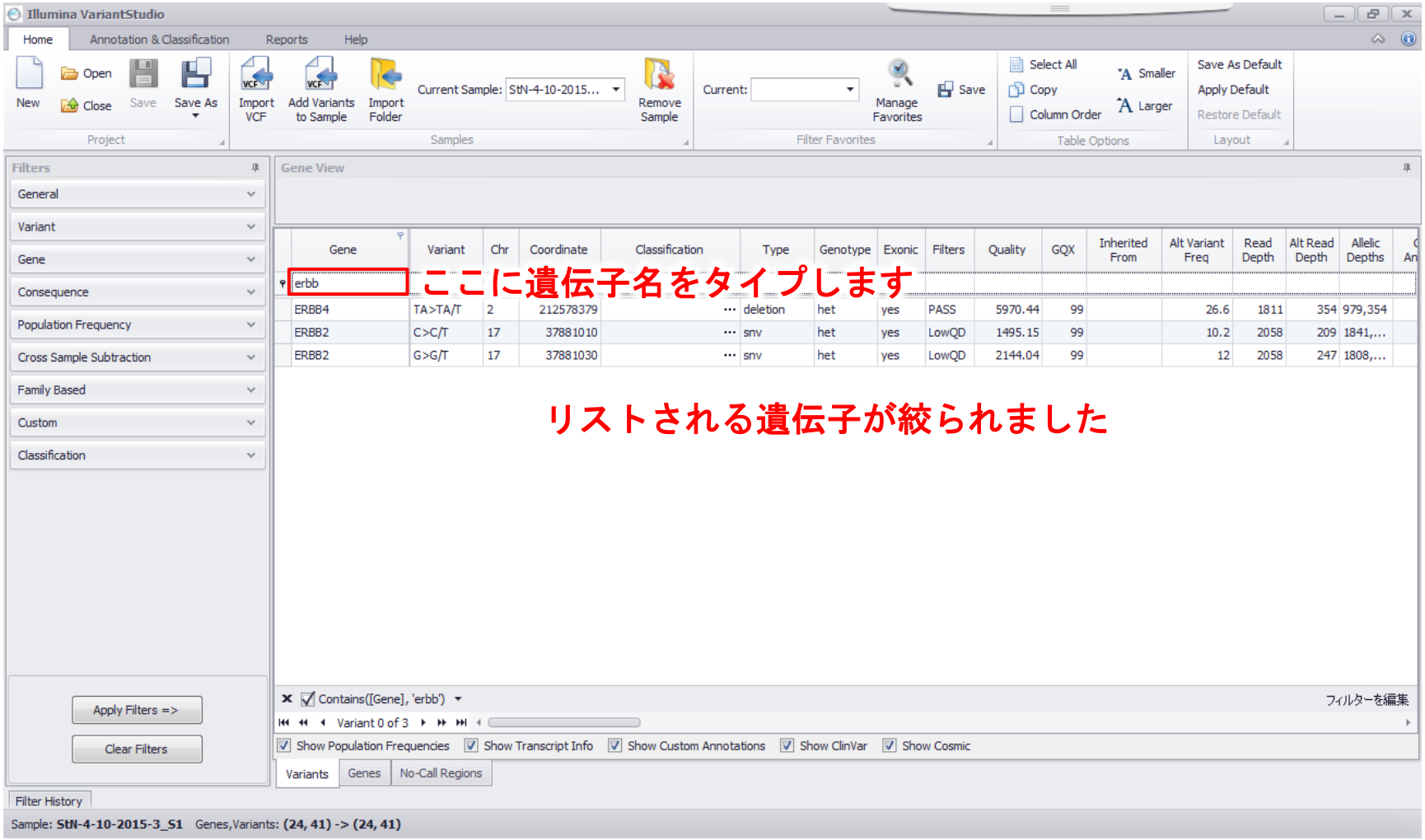


illumina

Annotate: 変異のアノテーション

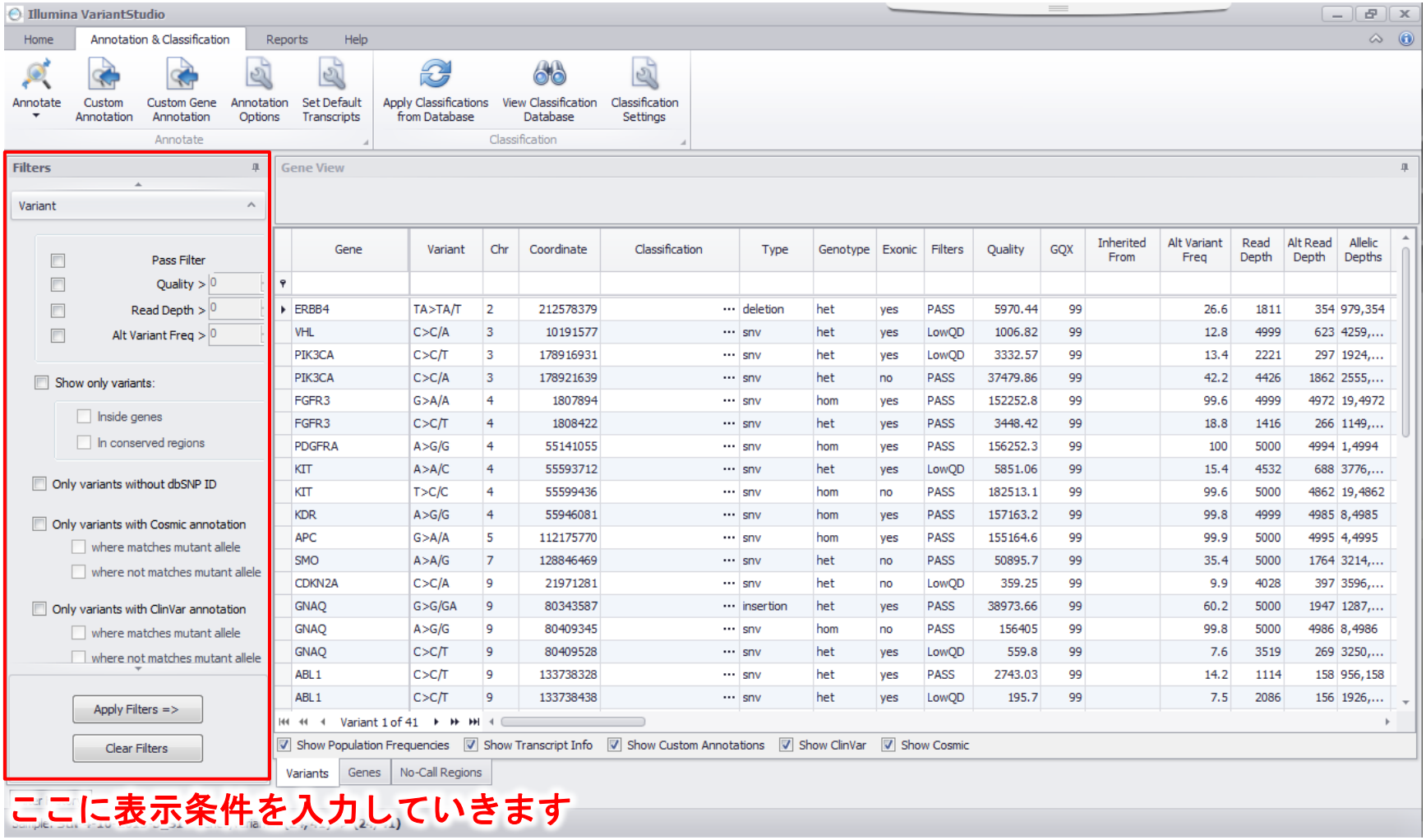


Filter: 変異の絞り込み（目的のキーワードがある場合）



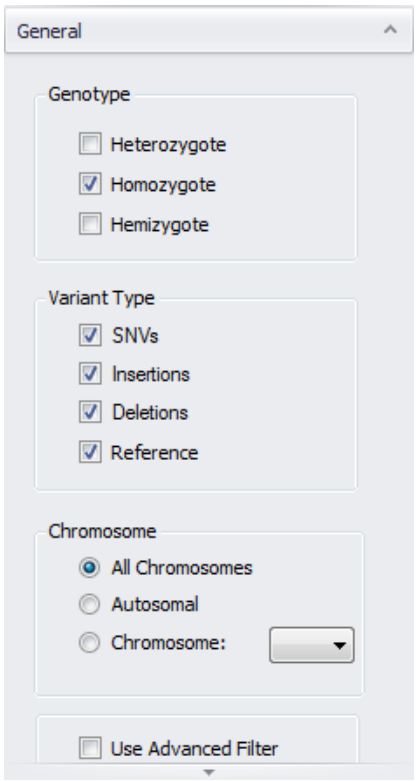
illumina

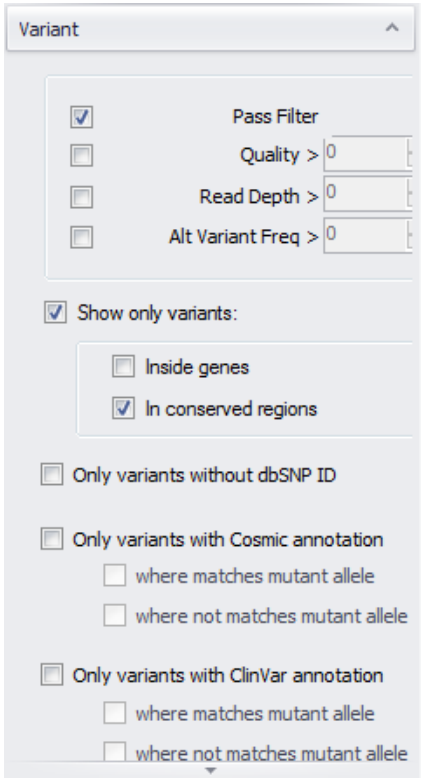
Filter: 変異の絞り込み（条件に合うものを探したい場合）

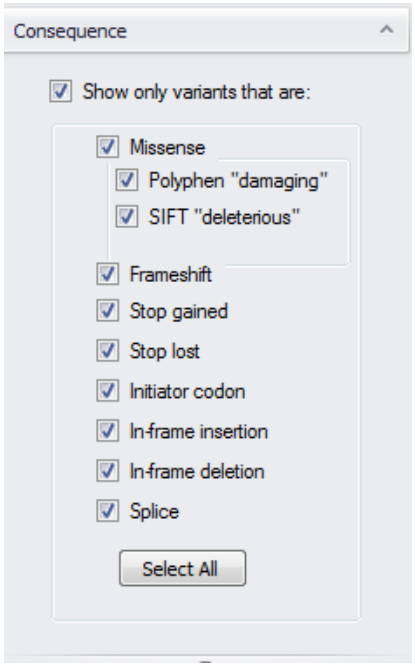


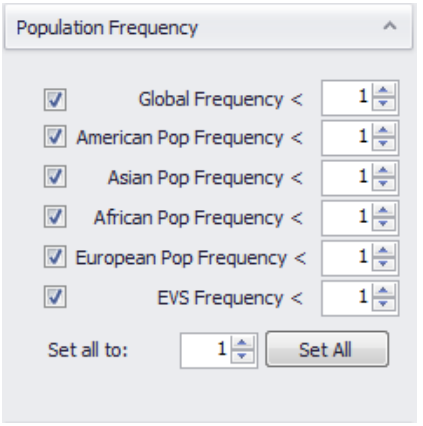
ここに表示条件を入力していきます

Filter: 変異の絞り込み（条件に合うものを探したい場合）



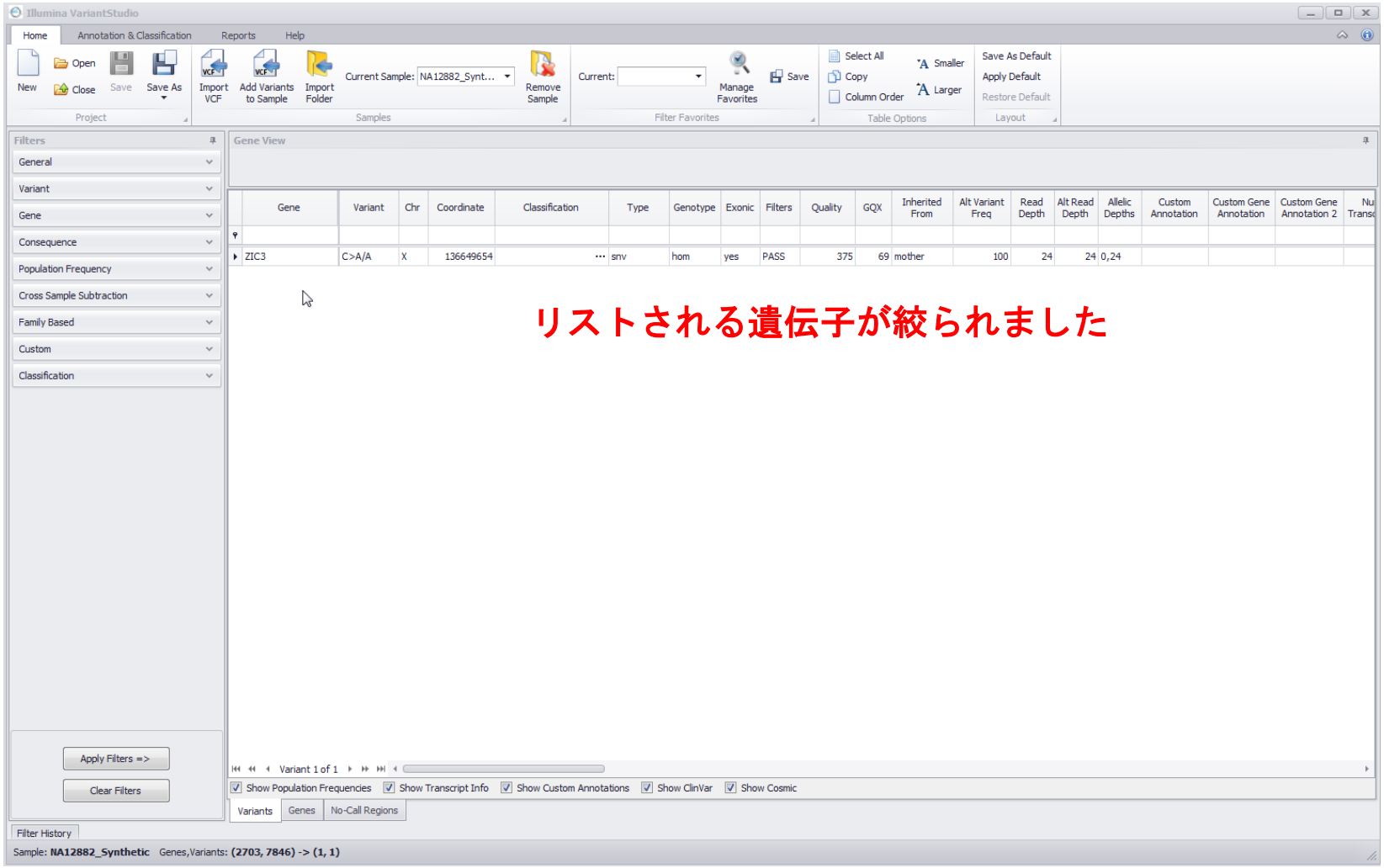






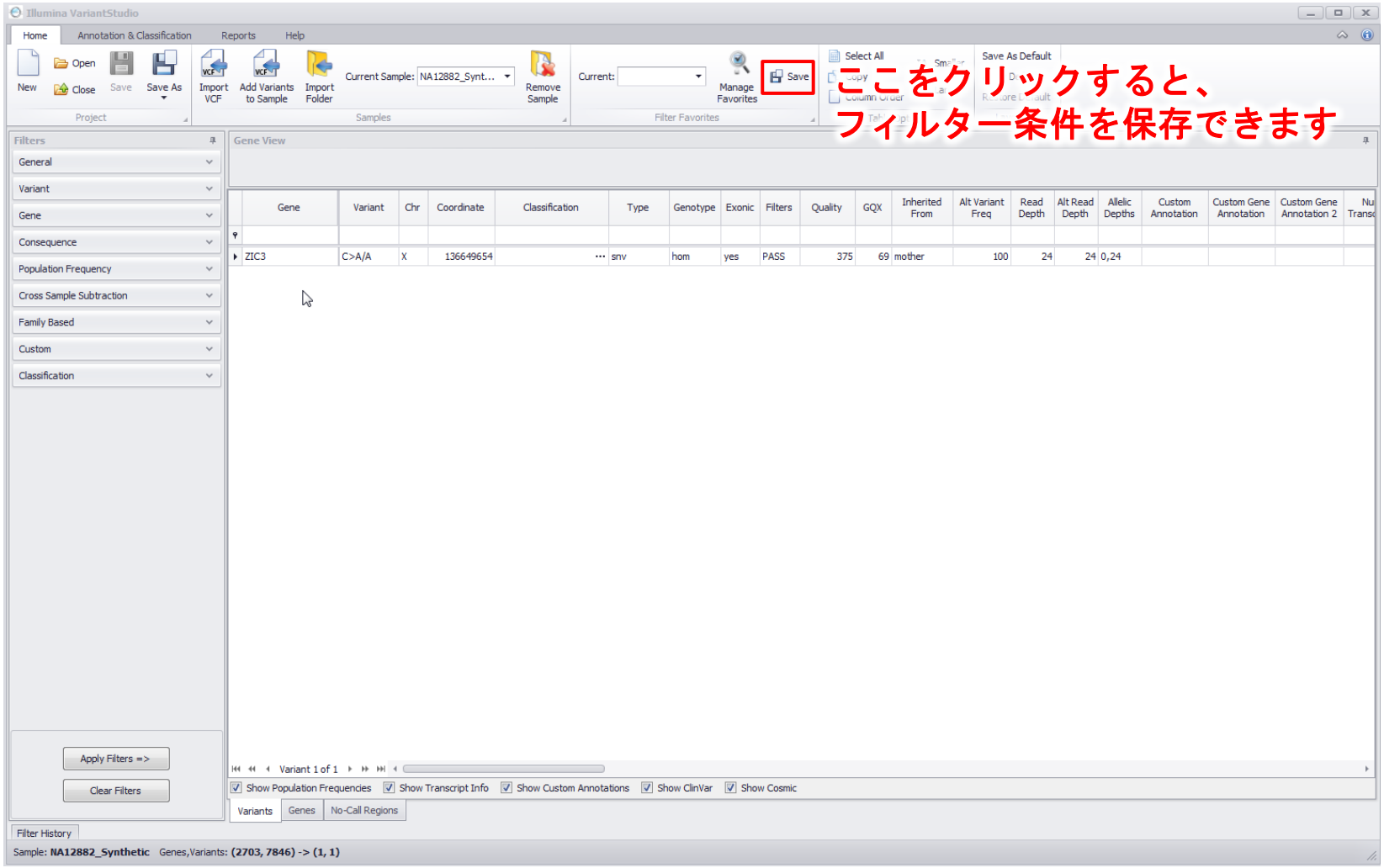
変異コールの情報、アノテーション情報から、任意の条件で変異を絞り込むことができます
条件設定が終わったら、'Apply Filter'をクリック

Filter: 変異の絞り込み（条件に合うものを探したい場合）



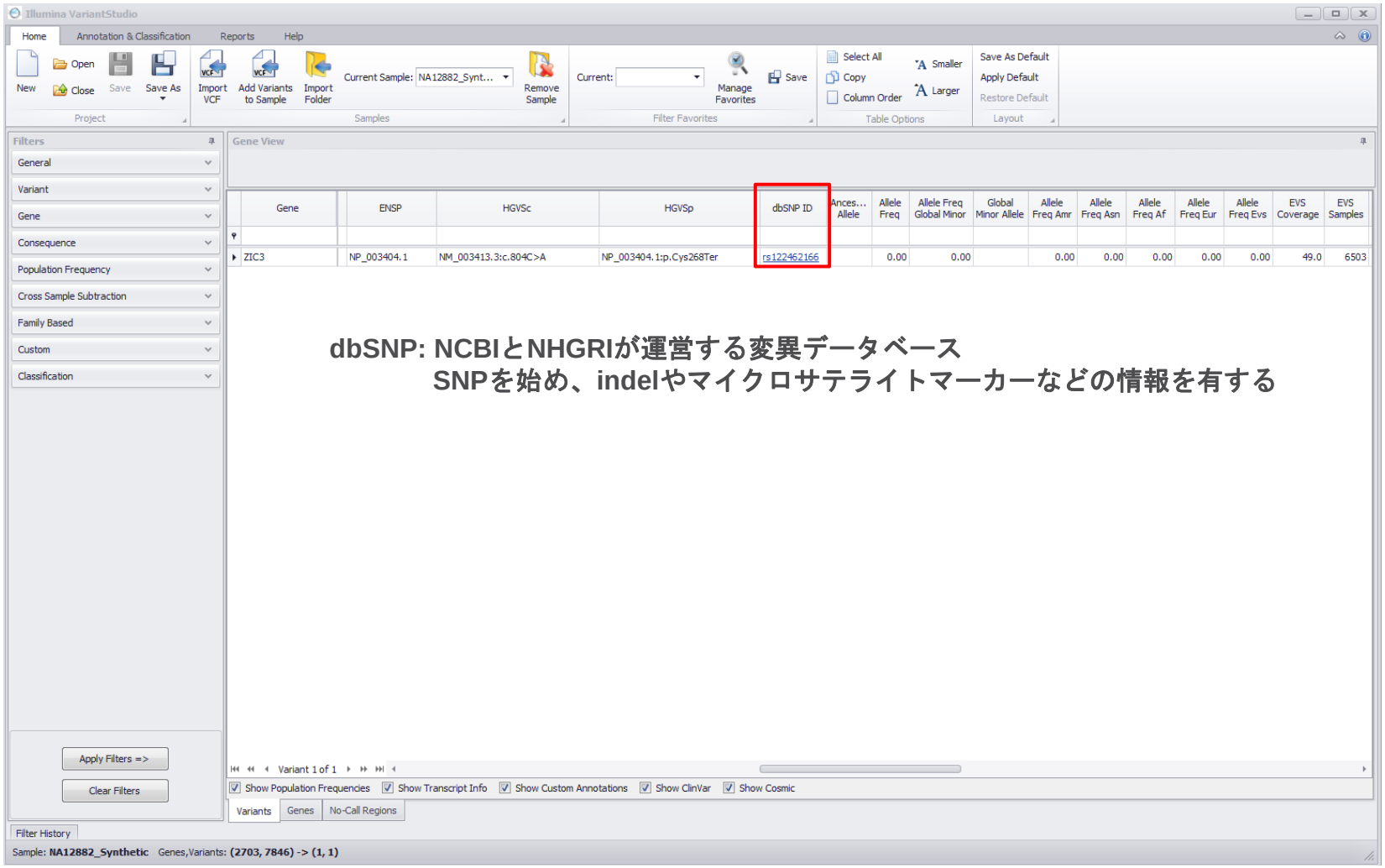
illumina®

Filter: 変異の絞り込み（条件に合うものを探したい場合）



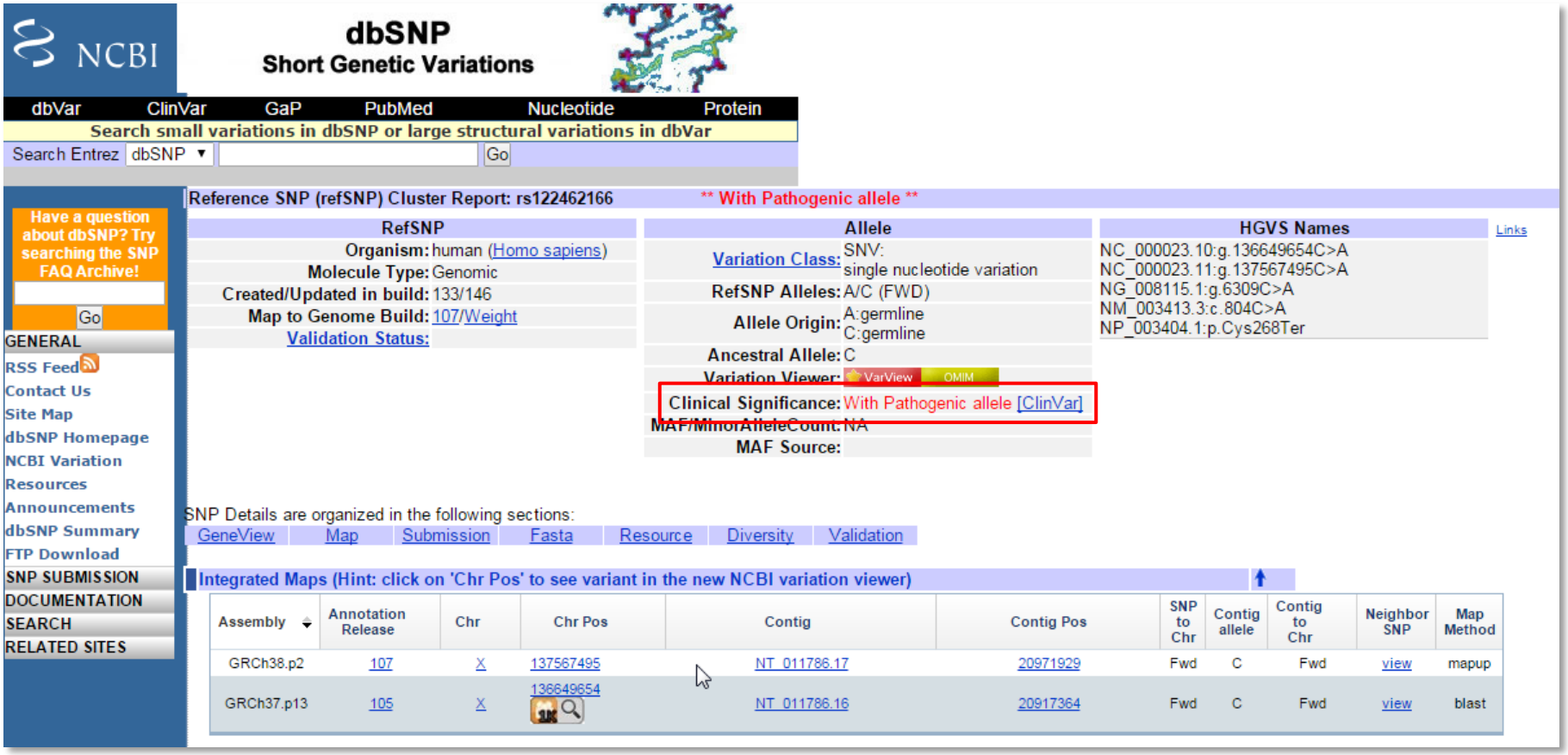
illumina®

変異の意義を知る: dbSNP



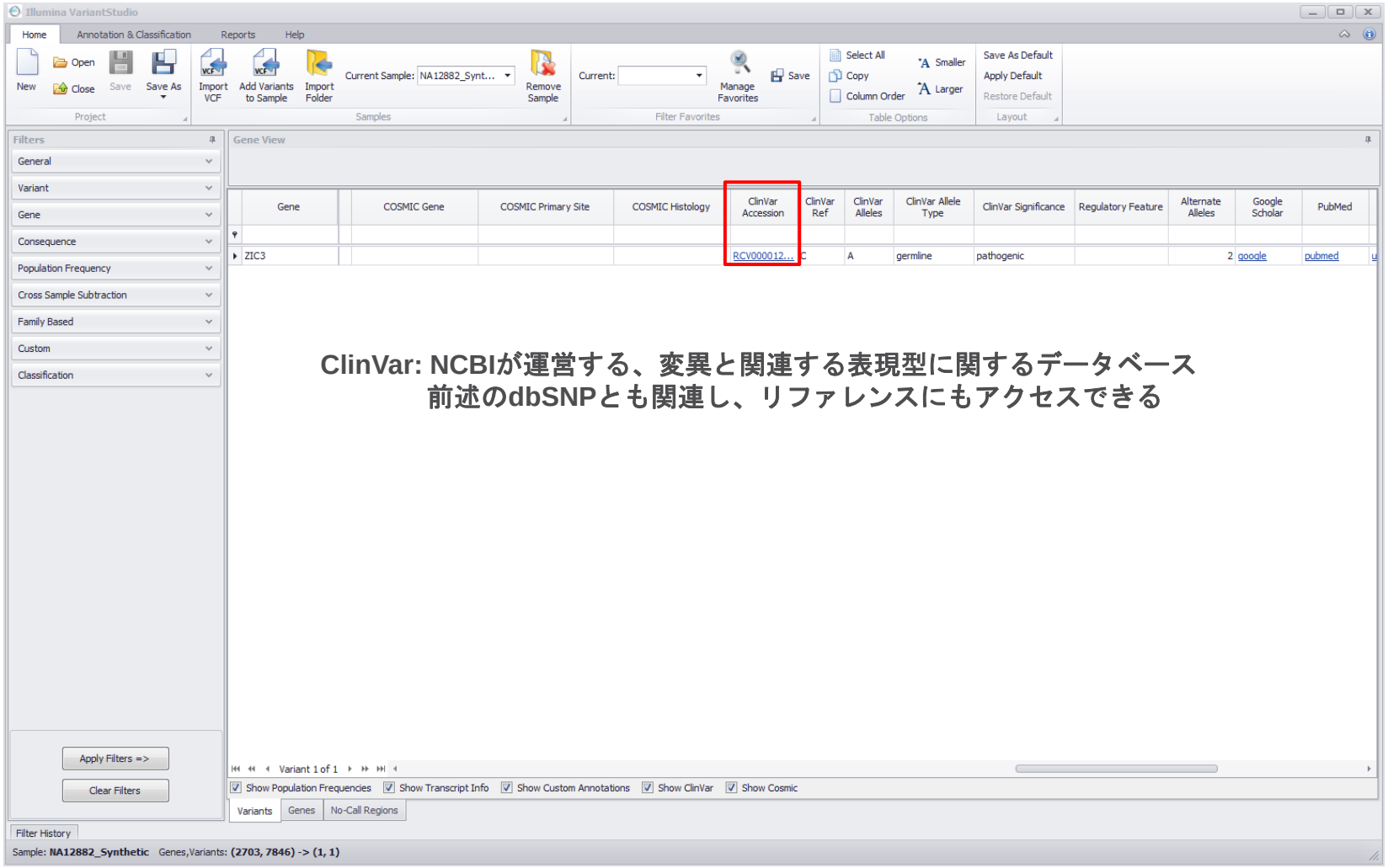
dbSNP: NCBIとNHGRIが運営する変異データベース
SNPを始め、indelやマイクロサテライトマーカーなどの情報を有する

変異の意義を知る: dbSNP



データベース内を参照すると、変異の発病性などについて知ることができる

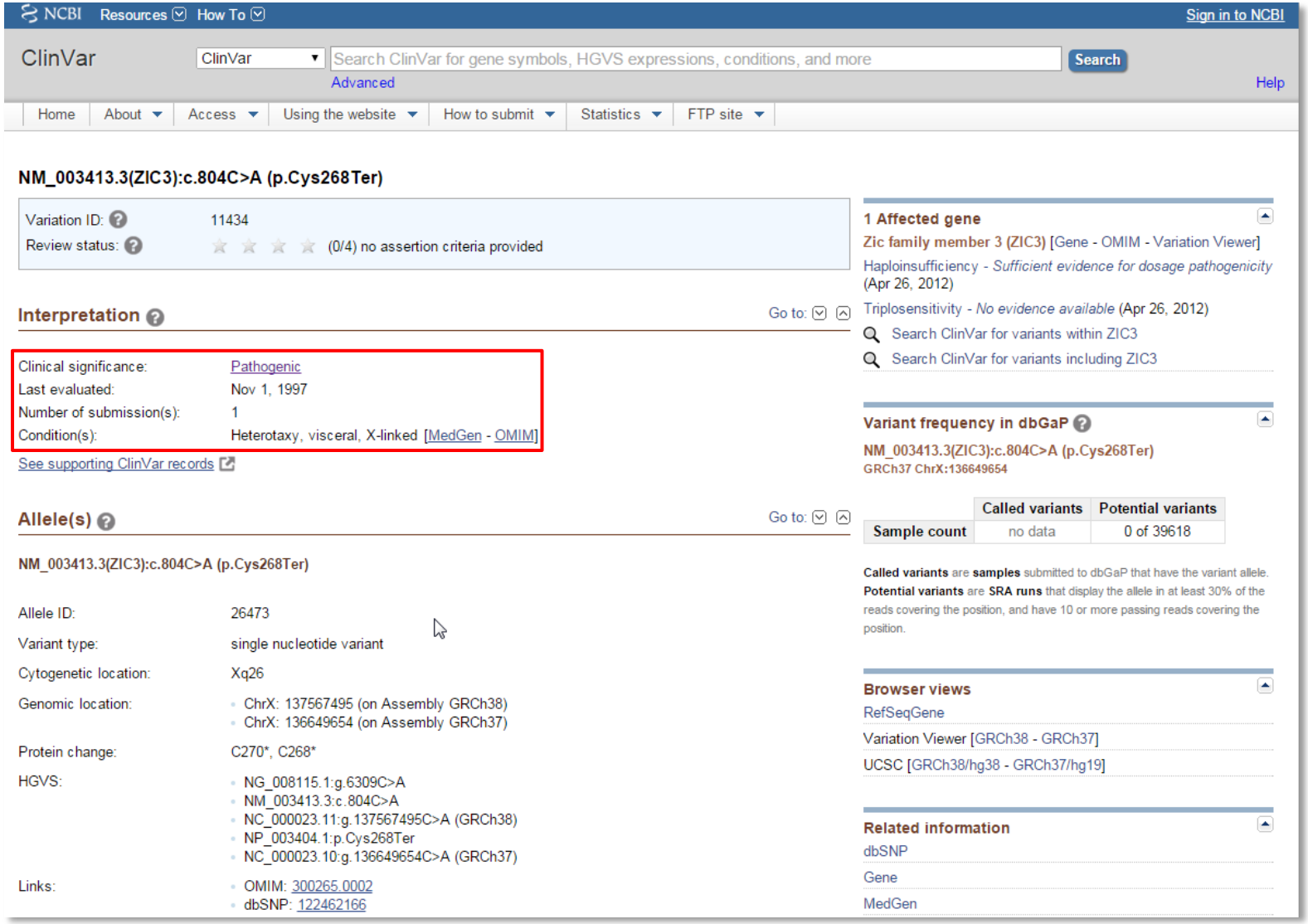
変異の意義を知る: ClinVar



ClinVar: NCBIが運営する、変異と関連する表現型に関するデータベース
前述のdbSNPとも関連し、リファレンスにもアクセスできる

illumina®

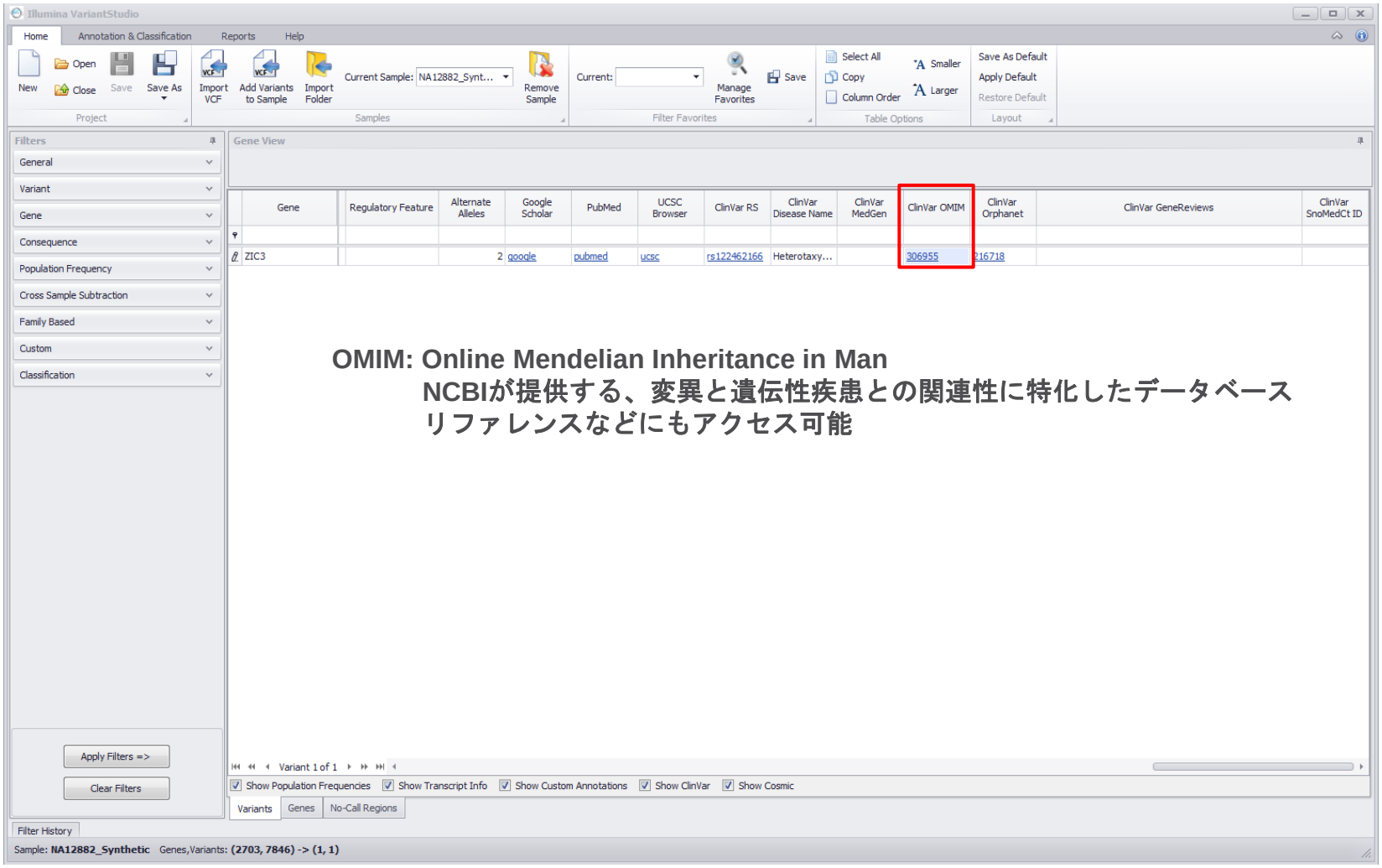
変異の意義を知る: ClinVar



変異の発病性や、表現型について情報を得ることができる

illumina

変異の意義を知る: OMIM



OMIM: Online Mendelian Inheritance in Man

NCBIが提供する、変異と遺伝性疾患との関連性に特化したデータベース

リファレンスなどにもアクセス可能

illumina®

変異の意義を知る: OMIM

ICD+

#306955

HETEROTAXY, VISCERAL, 1, X-LINKED; HTX1

Alternative titles; symbols

DEXTROCARDIA WITH OTHER CARDIAC MALFORMATIONS
LATERALITY, X-LINKED
SITUS INVERSUS, COMPLEX CARDIAC DEFECTS, AND SPLENIC DEFECTS, X-LINKED

Other entities represented in this entry:

CONGENITAL HEART DEFECTS, MULTIPLE TYPES, 1, X-LINKED, INCLUDED; CHTD1, INCLUDED

Phenotype-Gene Relationships

Location	Phenotype	Phenotype MIM number	Inheritance (in progress)	Phenotype mapping key	Gene/Locus	Gene/Locus MIM number
Xq26.3	Heterotaxy, visceral, 1, X-linked	306955	XLR	3	ZIC3	300265
Xq26.3	Congenital heart defects, nonsyndromic, 1, X-linked	306955	XLR	3	ZIC3	300265

Clinical Synopsis | Phenotypic Series

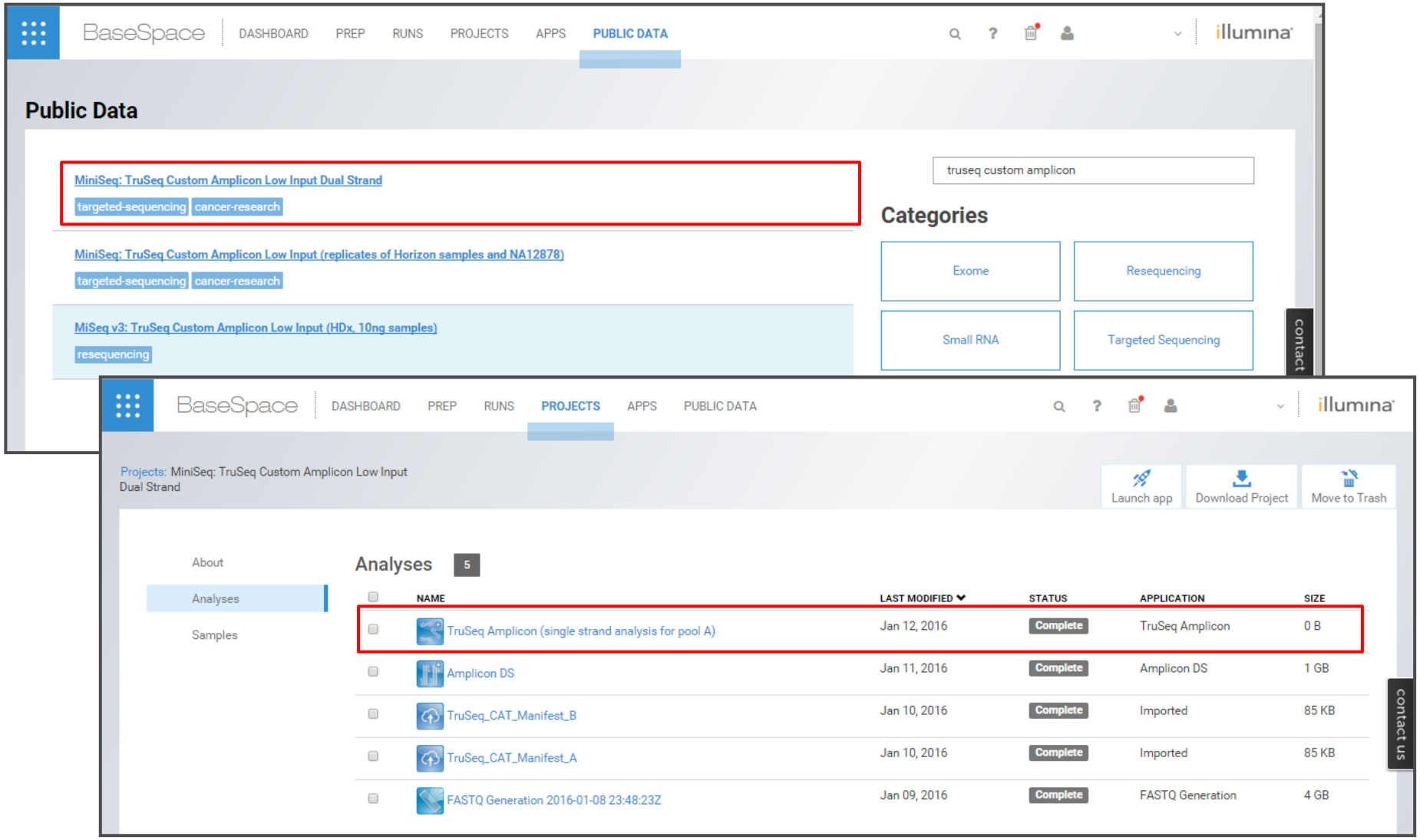
TEXT

A number sign (#) is used with this entry because X-linked heterotaxy-1 (HTX1) and multiple types of congenital heart defects-1 (CHTD1) are caused by mutation in the ZIC3 gene (300265) on chromosome Xq26.

Mutation in the ZIC3 gene can also cause VACTERL with or without hydrocephalus (VACTERLX; 314390), a disorder with overlapping features.

変異が関与する遺伝性疾患について知ることができる

Public dataの取得



本日の内容

- **BaseSpace cloud app, TruSeq Ampliconを用いたデータ解析**
  - シーケンスデータの参照
  - 変異の解析結果の参照

- **VariantStudioソフトウェアを用いたデータ解析**
  - データのインポート
  - 変異の絞り込み
  - アノテーション
  - データベースを利用した知見の入手

注：解析にはデモデータを使用しており、細かくは実際と異なる場合がございます。
データの参照と解析の流れをご案内する一例としてご参考下さい。