TruSight One

「高効率・高感度な臨床研究を可能にする

TruSight Oneシーケンスパネル～ドライ編～」

Dec 19, 2014

長井 陽子
イルミナ株式会社
シーケンシングスペシャリスト

illumina®

TruSight Oneシーケンスパネル

サンプル調製から解析までの一連の工程をサポート

BaseSpace®

ライブラリー調製

TruSight One
シーケンスパネル

2014年12月5日実施
ウェビナー参照

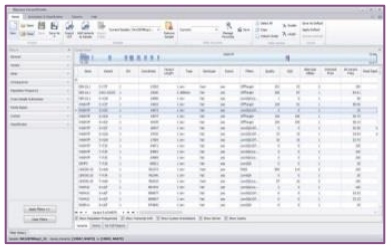
シーケンス

アライメント
& 変異コール

Enrichment

アノテーション

VariantStudio



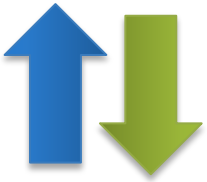
本ウェビナー内容

illumina®

イルミナ VariantStudioの流れ

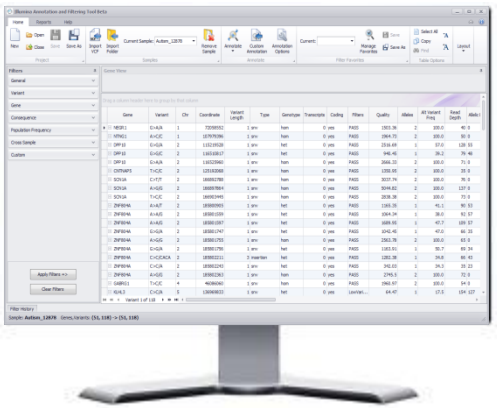
バリアントコールデータを入力、生物学的意味付けを出力

VariantStudio
アノテーションデータベース



VCF または gVCF ファイルの入力

VCF：バリアントコールフォーマット



VariantStudio™

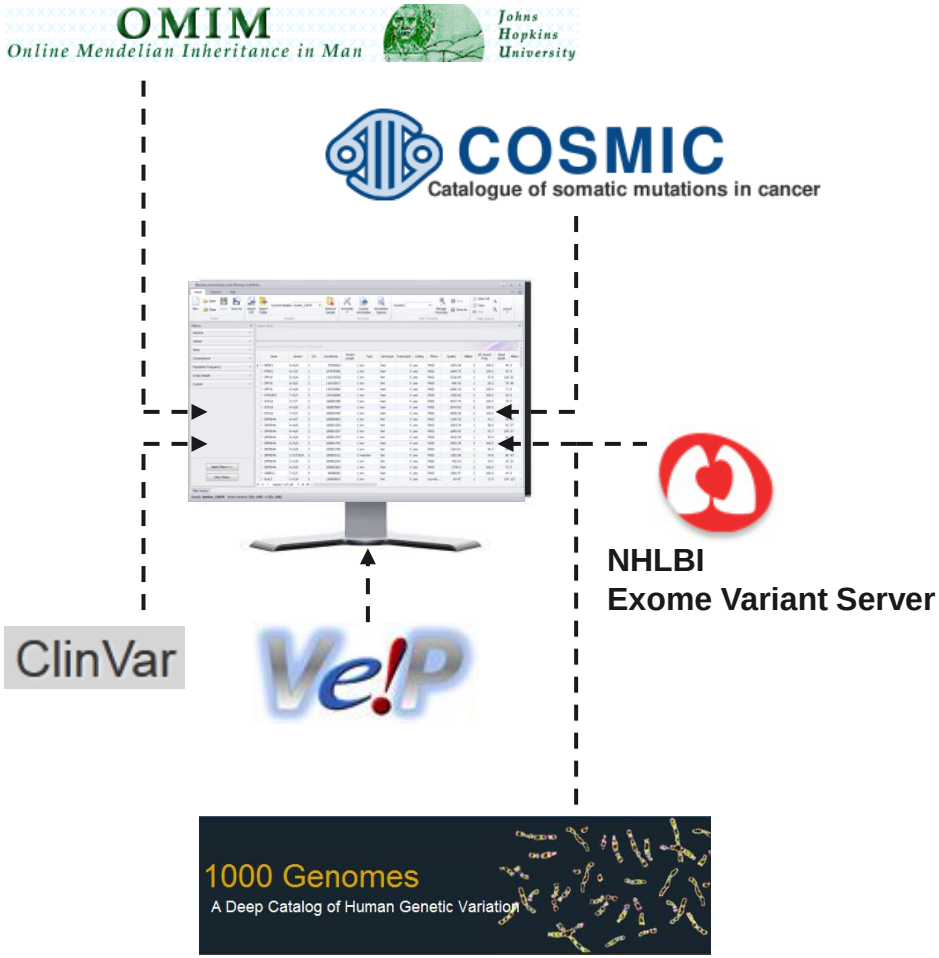
解釈した変異をレポートとして出力

illumina®

アノテーション機能

様々なアノテーション情報を活用した生物学的意味付け

- **提供するアノテーション情報**
  - 変異レベル
  - 遺伝子レベル
- **アノテーションに含まれる項目**
  - 転写産物名
  - 関連する疾患や表現型
  - 各種人類集団におけるアリル頻度
  - 遺伝子機能への影響
  - 文献検索
  - データベースへのリンク
- **外部データ**
  - テキストファイルを読み込み

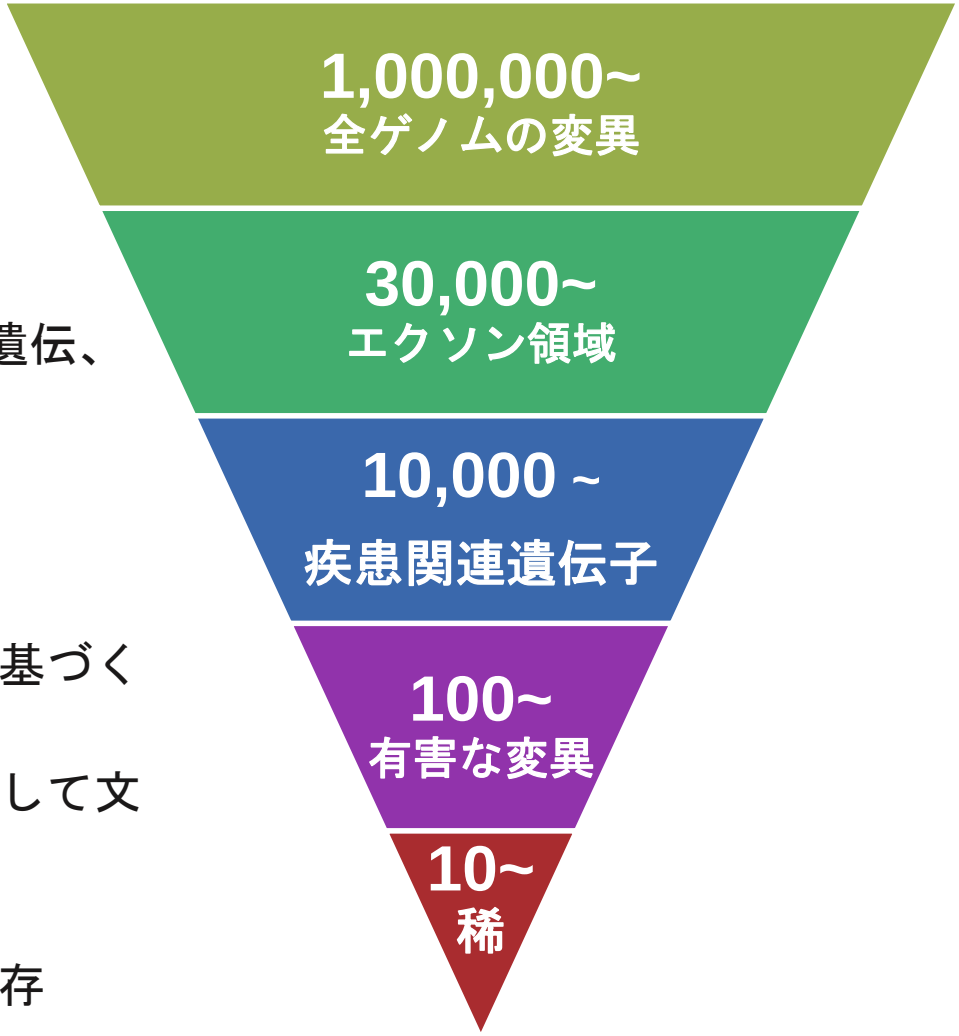


illumina

VariantStudioによるバリアントの絞り込み

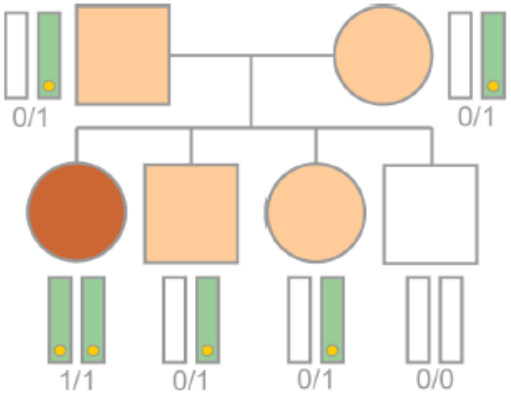
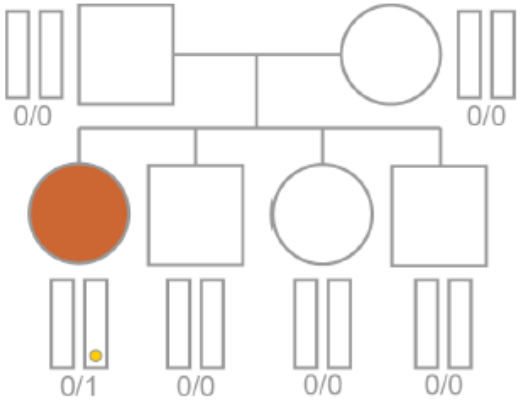
シーケンシングより得られた膨大な変異から疾患原因となる変異を同定

- **様々な研究に対応**
  - 腫瘍-正常ペアの体細胞変異解析
  - 先天性疾患の疾患原因変異解析
  - 親子トリオなどの家系解析
  - X連鎖性劣性遺伝、常染色体劣性遺伝、常染色体優性遺伝、*de novo* 変異
- **柔軟で簡単な絞り込み機能**
  - 遺伝子名、疾患名、アレル頻度に基づく簡単な絞り込み
  - 変異テーブルの全てのカラムに対して文字検索や並べ替えが可能
  - 絞り込み条件の履歴を記憶
  - 複雑な絞込条件をお気に入りに保存



illumina

家系解析の研究デザイン例

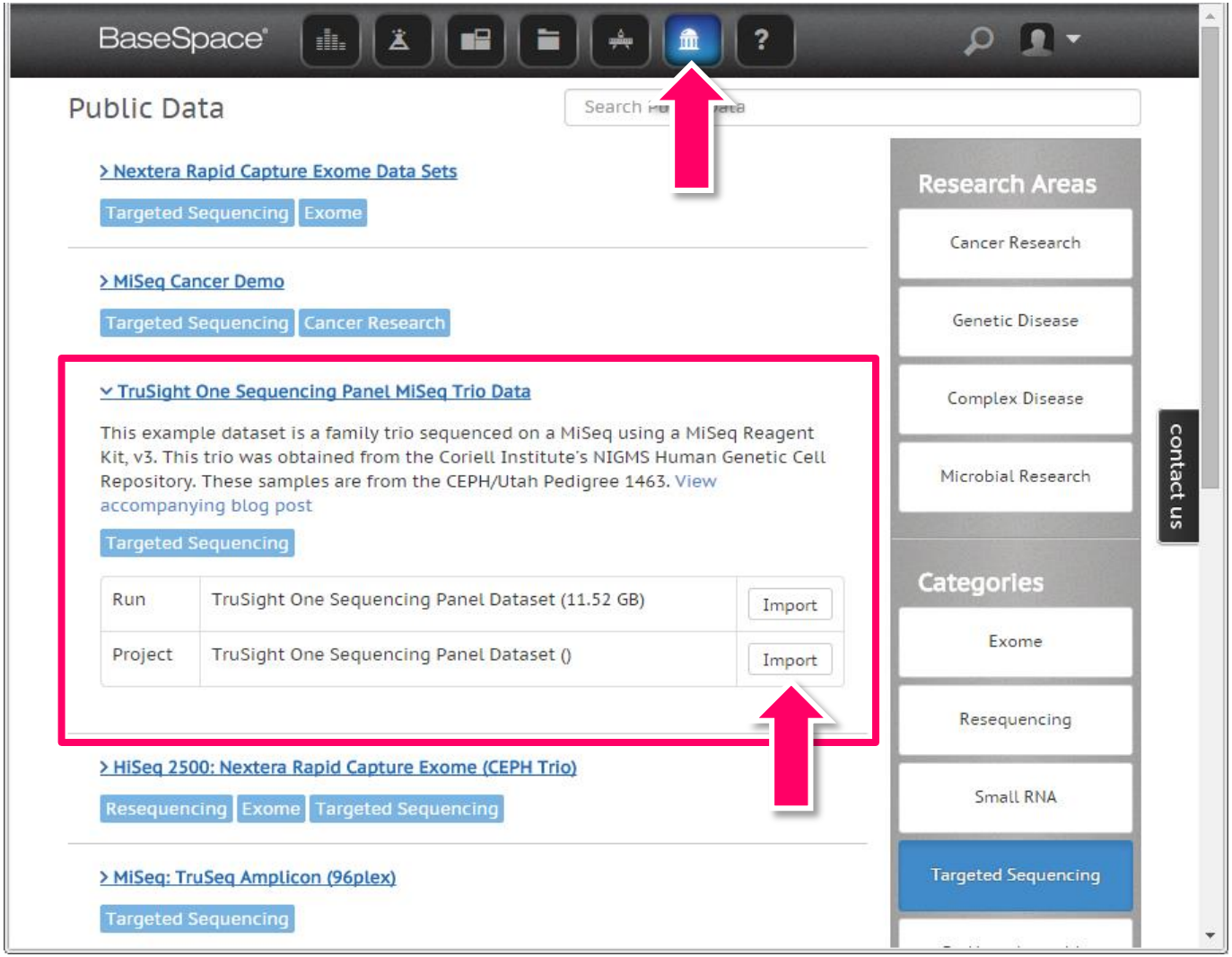
患者情報	 **家族内集積がある遺伝性疾患**	 **家族内集積なし孤発性疾患**
仮説	常染色体劣性遺伝	*De Novo* 変異
研究デザイン	スクリーニング：家族サンプル 検証：独立した複数家族	スクリーニング：家族サンプル 検証：独立した複数発症者

デスクトップ版とクラウド版VariantStudioの違い

	Illumina VariantStudio 2.2	BaseSpace VariantStudio 2.2
特徴	デスクトップ版	クラウド版
ソフトウェア起動	ローカルPCより起動	BaseSpaceより起動
VCF読み込み	ローカルPCに保存されたVCFファイル	BaseSpaceにアップロード/共有/インポートされたVCFファイル
プロジェクト保存	ローカルPCに保存	ローカルPCに保存
プロジェクト読み込み	PCに保存されたプロジェクト	PCに保存されたプロジェクト

本ウェビナーでは、BaseSpaceパブリックデータのTruSightOne シーケンスパネルVCFファイルに独自に変異を導入したデモデータを用いて解析をします。

BaseSpaceからパブリックデータの取得方法

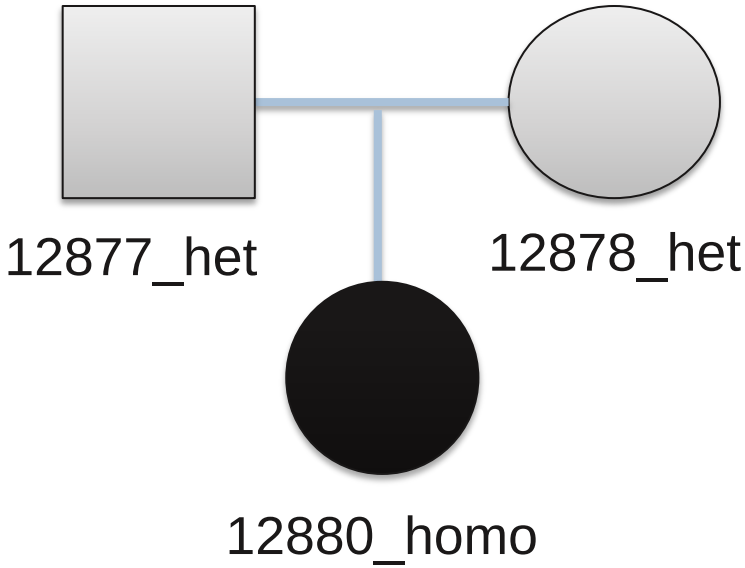


illumina

本ウェビナー デモ用親子トリオデータについて

デモデータ

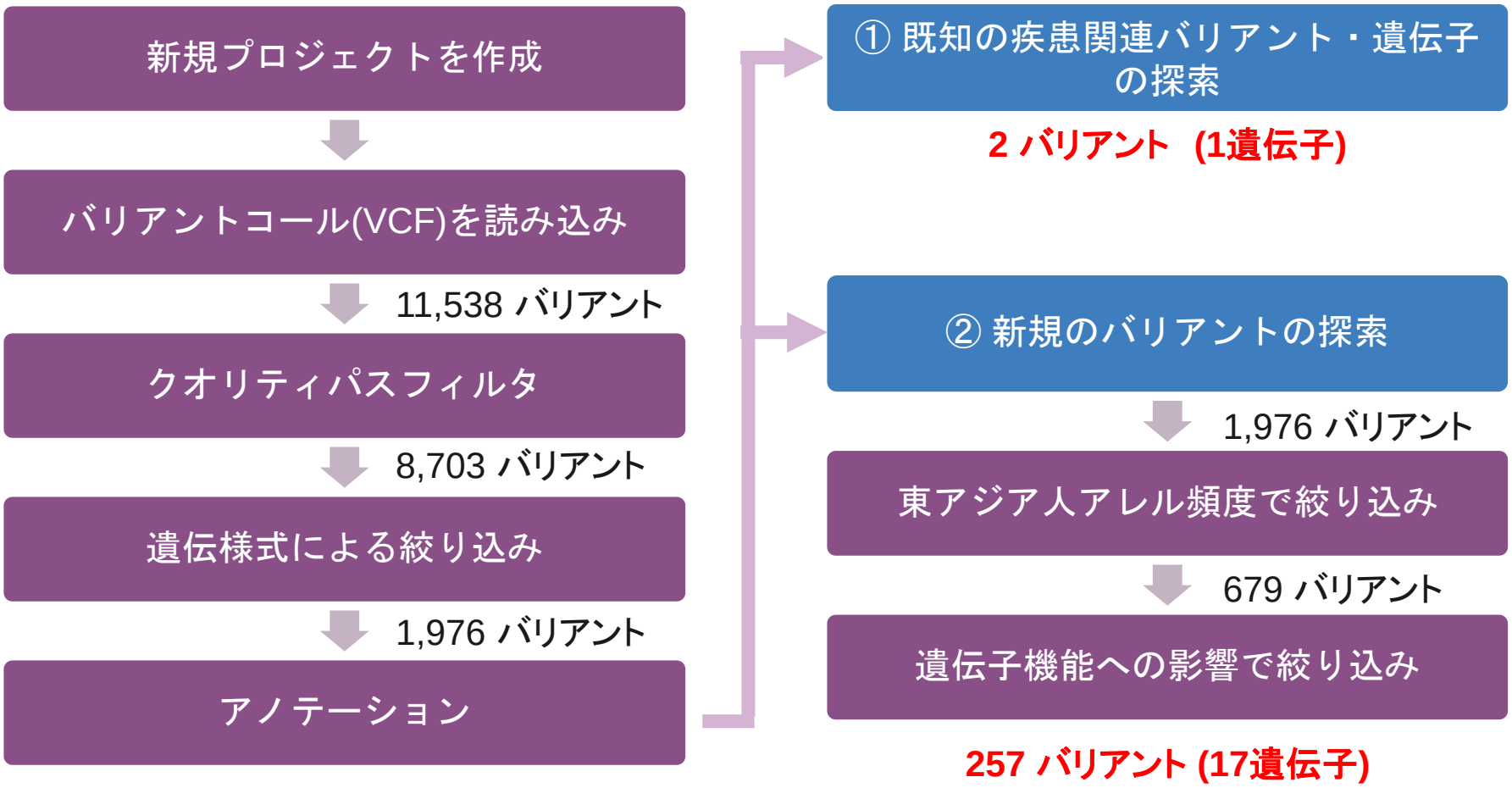
両親にヘテロ、子にホモの先天性疾患原因変異を導入



トリオ解析の前提

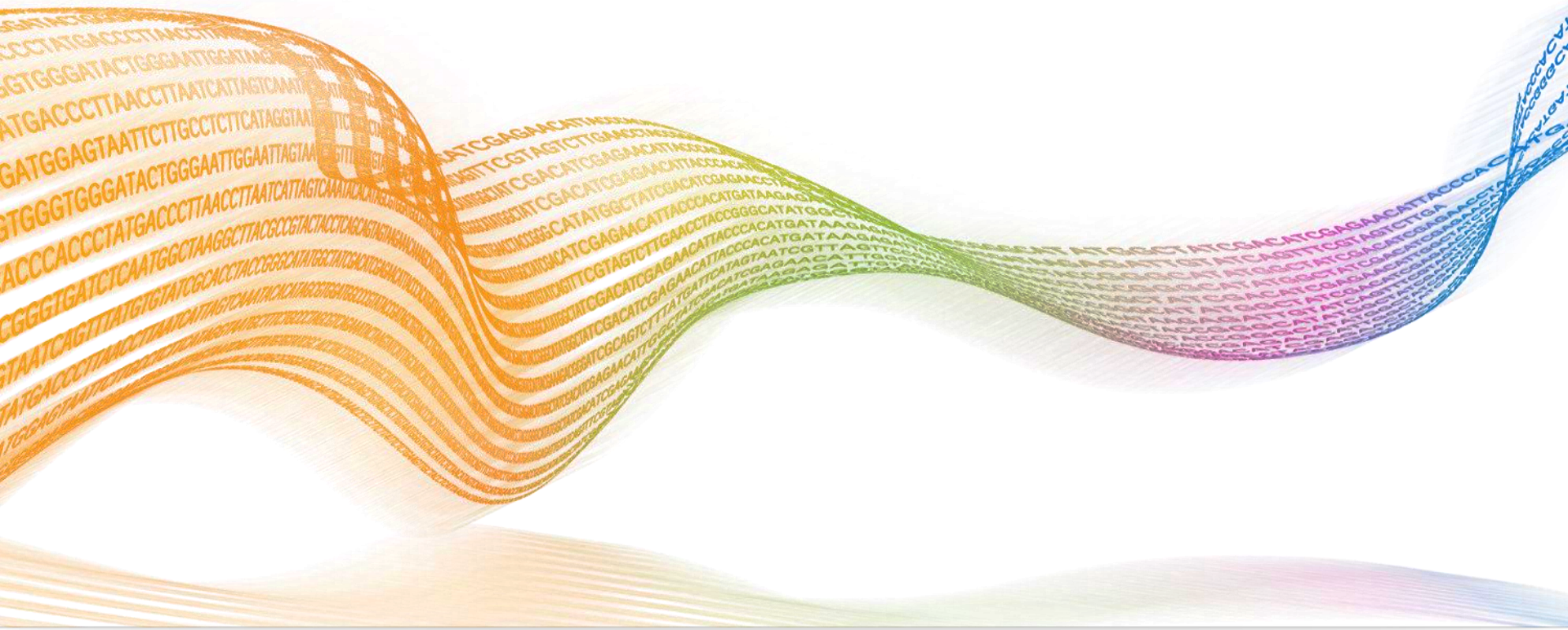
1. 疾患は常染色体劣性遺伝と仮定。
2. 希少疾患で、原因変異はアレル頻度が低いと仮定。
3. 遺伝子上の変異で遺伝子機能に影響を与えると仮定。

VariantStudio解析ワークフロー例



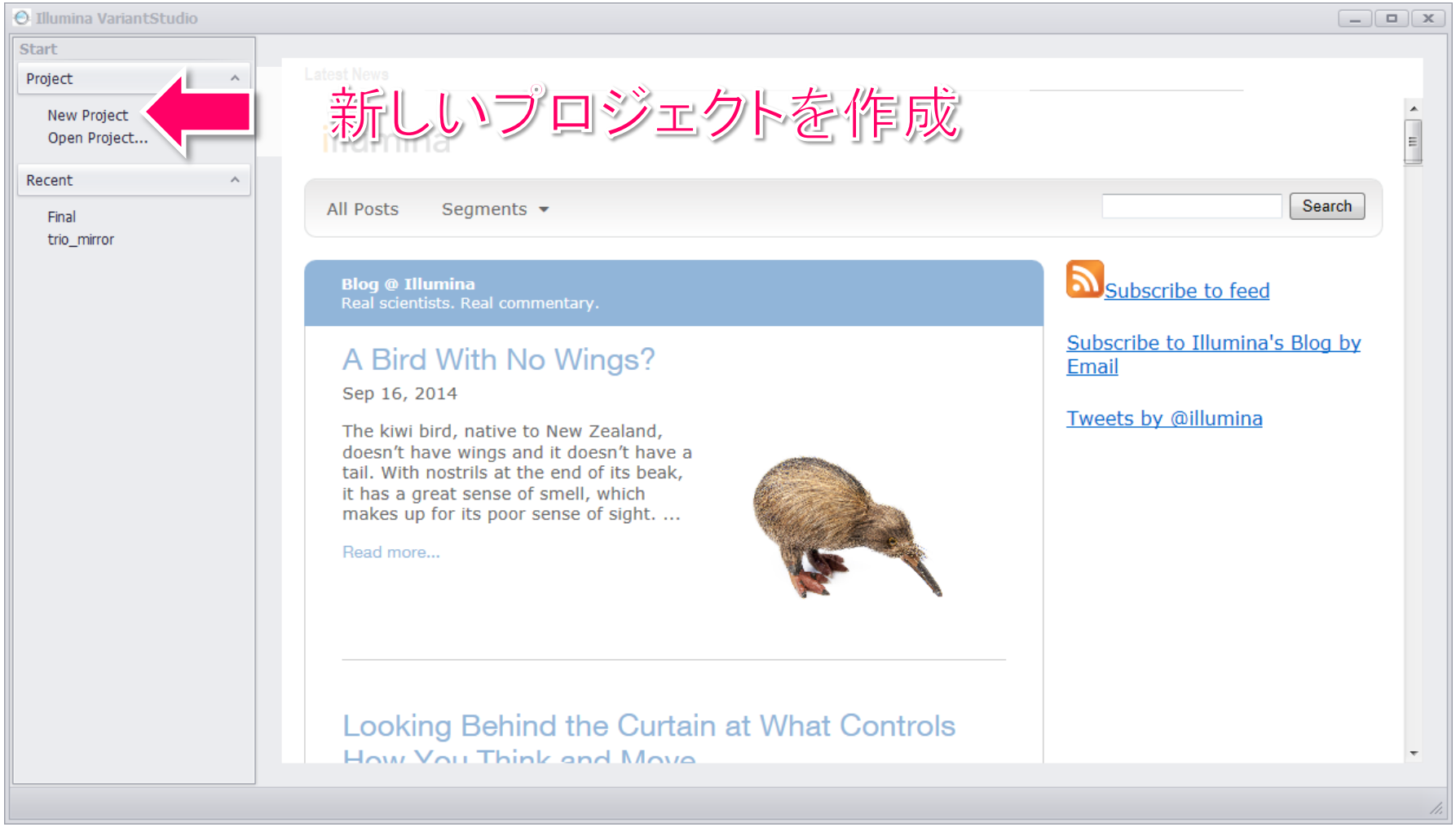
illumina

VariantStudioを使ってみよう

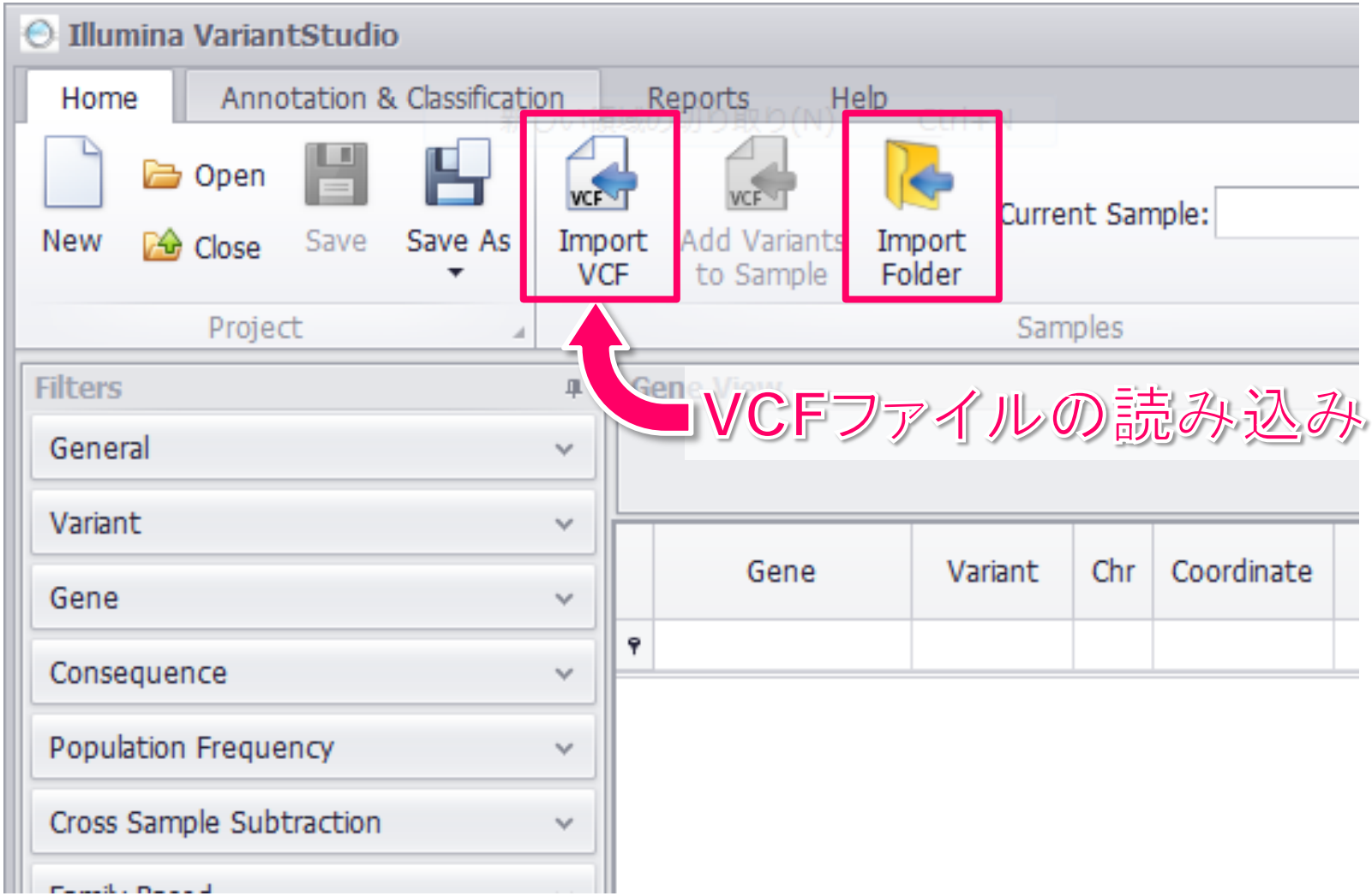


illumina®

1. プロジェクトを作成する

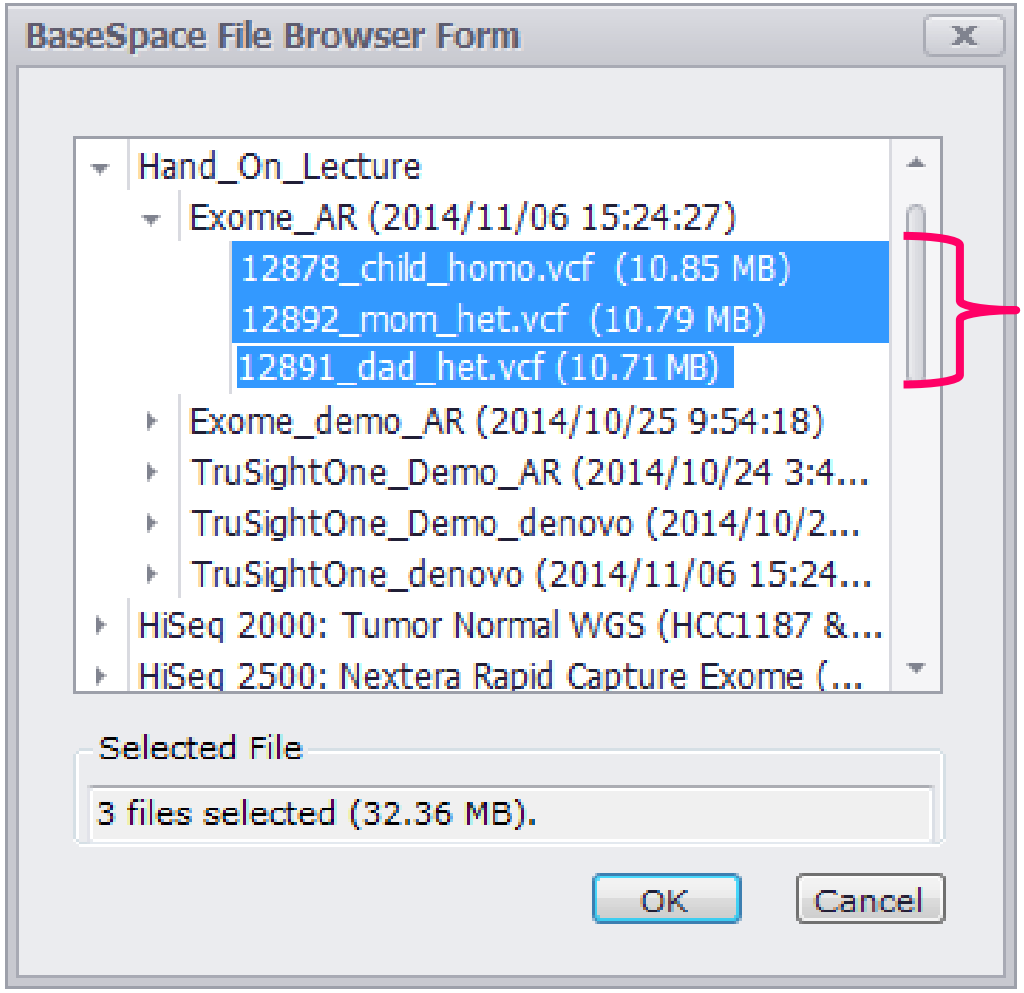


2. サンプルVCFファイルの読み込み



illumina

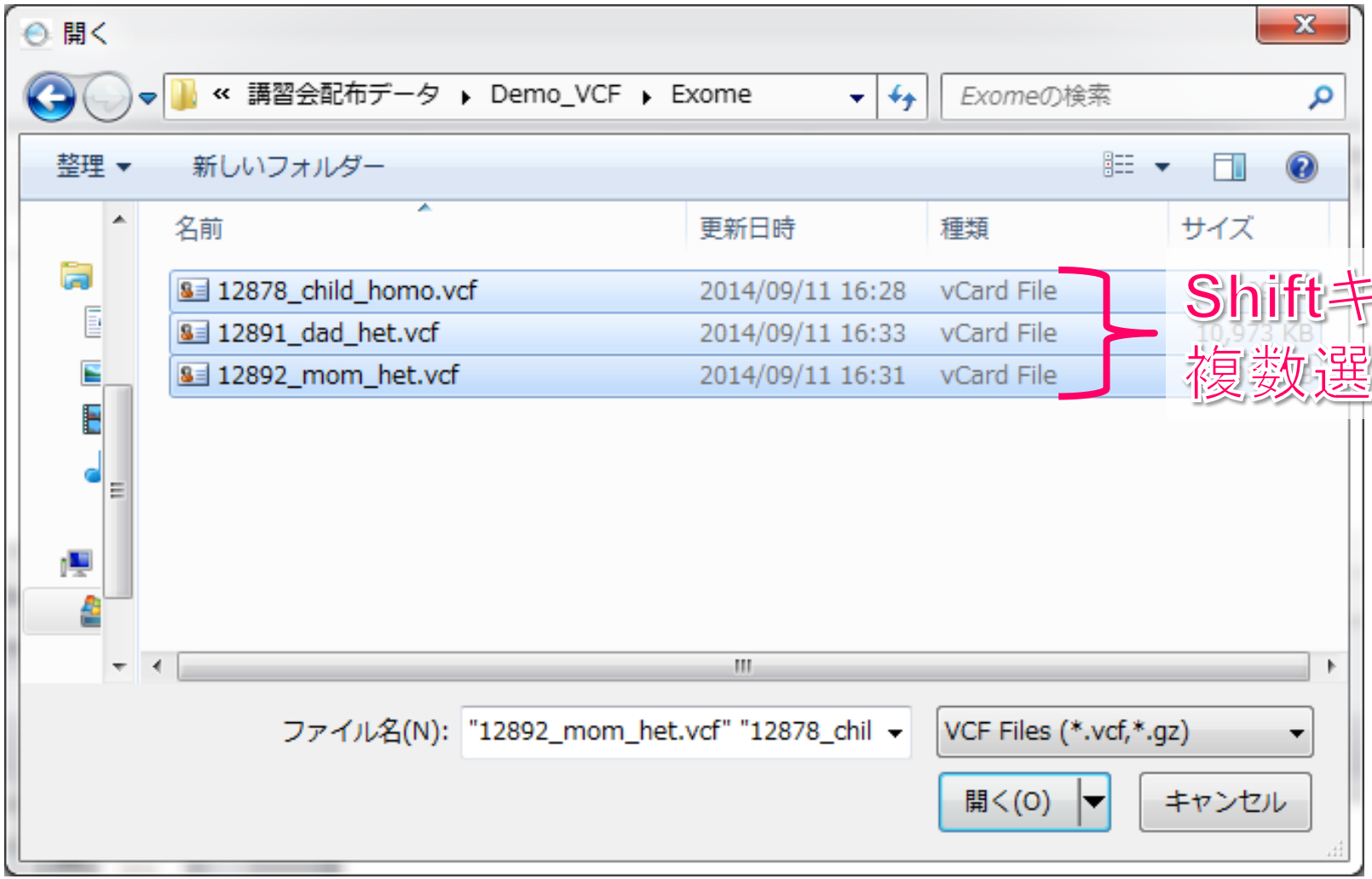
クラウド版：BaseSpaceからVCFファイルの読み込み



Shiftキーで
複数選択可

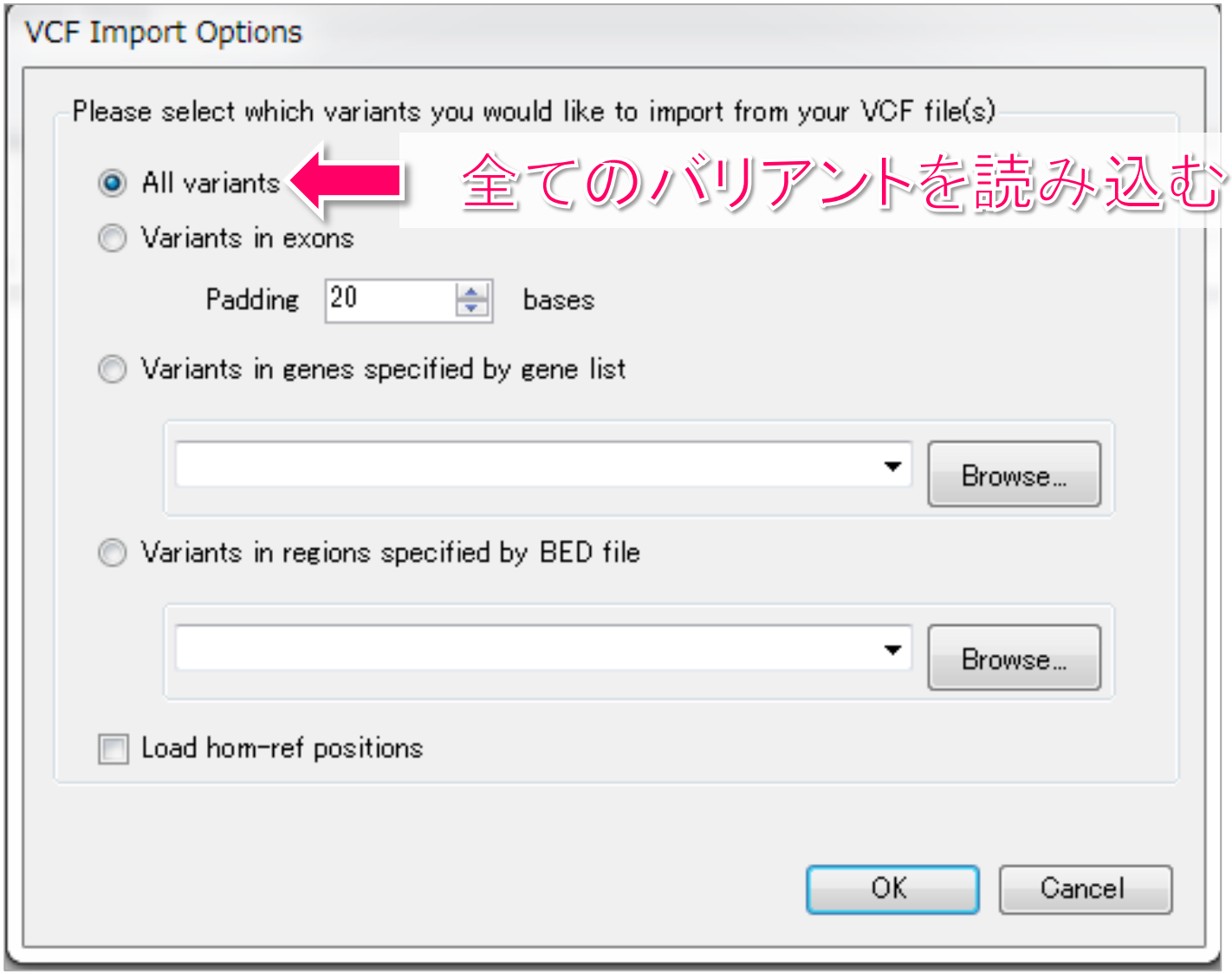
illumina

デスクトップ版：VCFファイルの読み込み



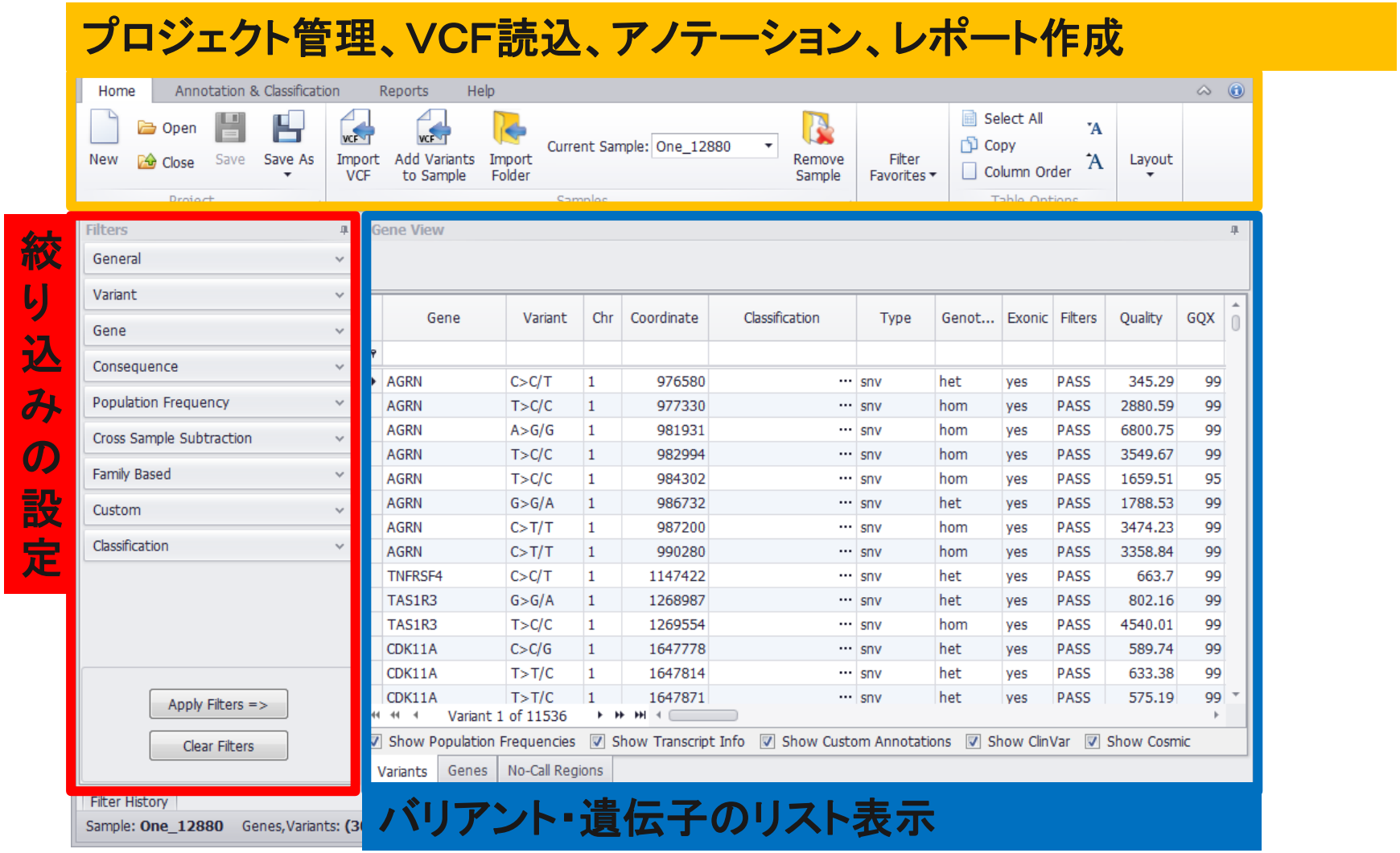
illumina

2. サンプルVCFファイルの読み込み（オプション）



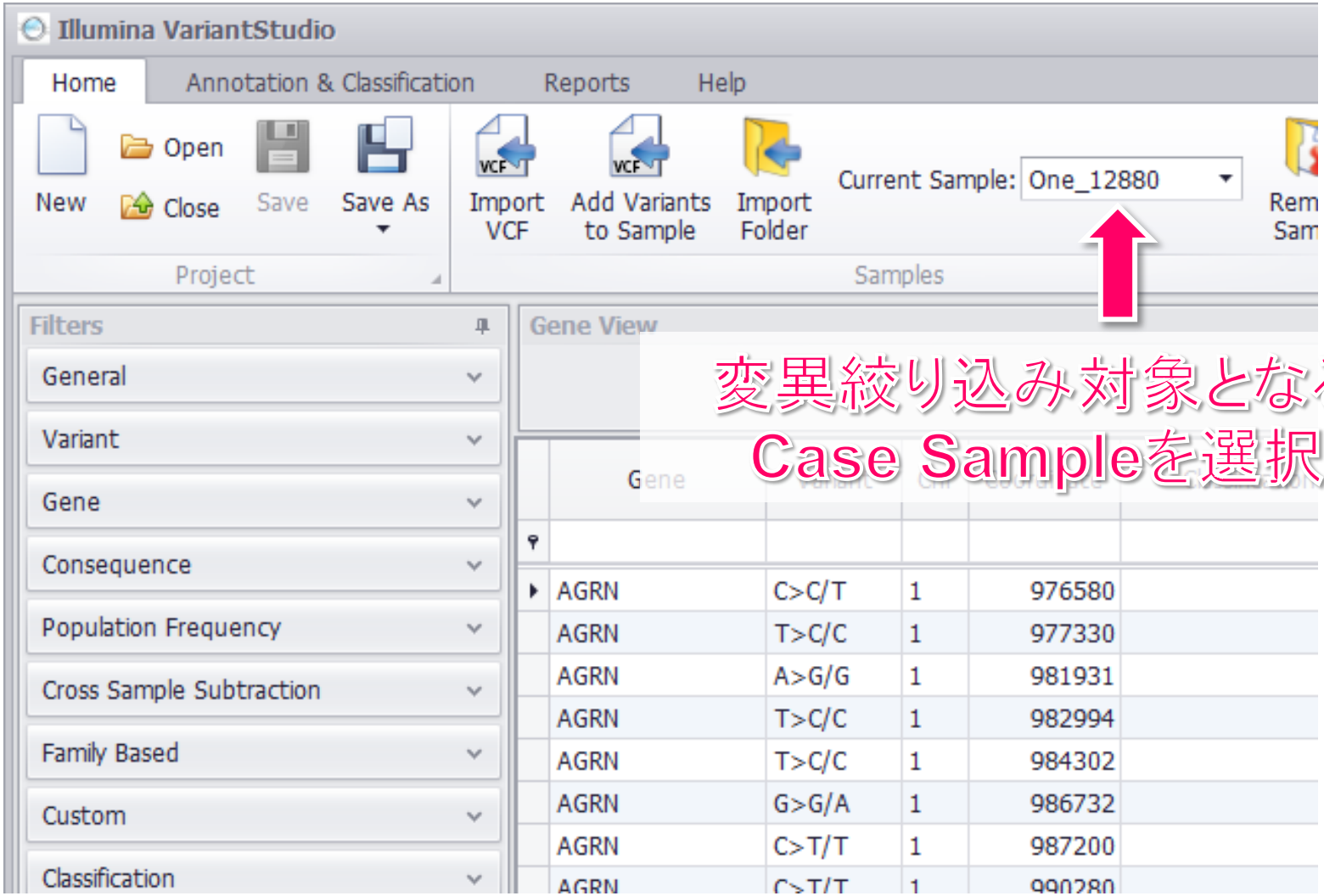
illumina

2. 画面の見方



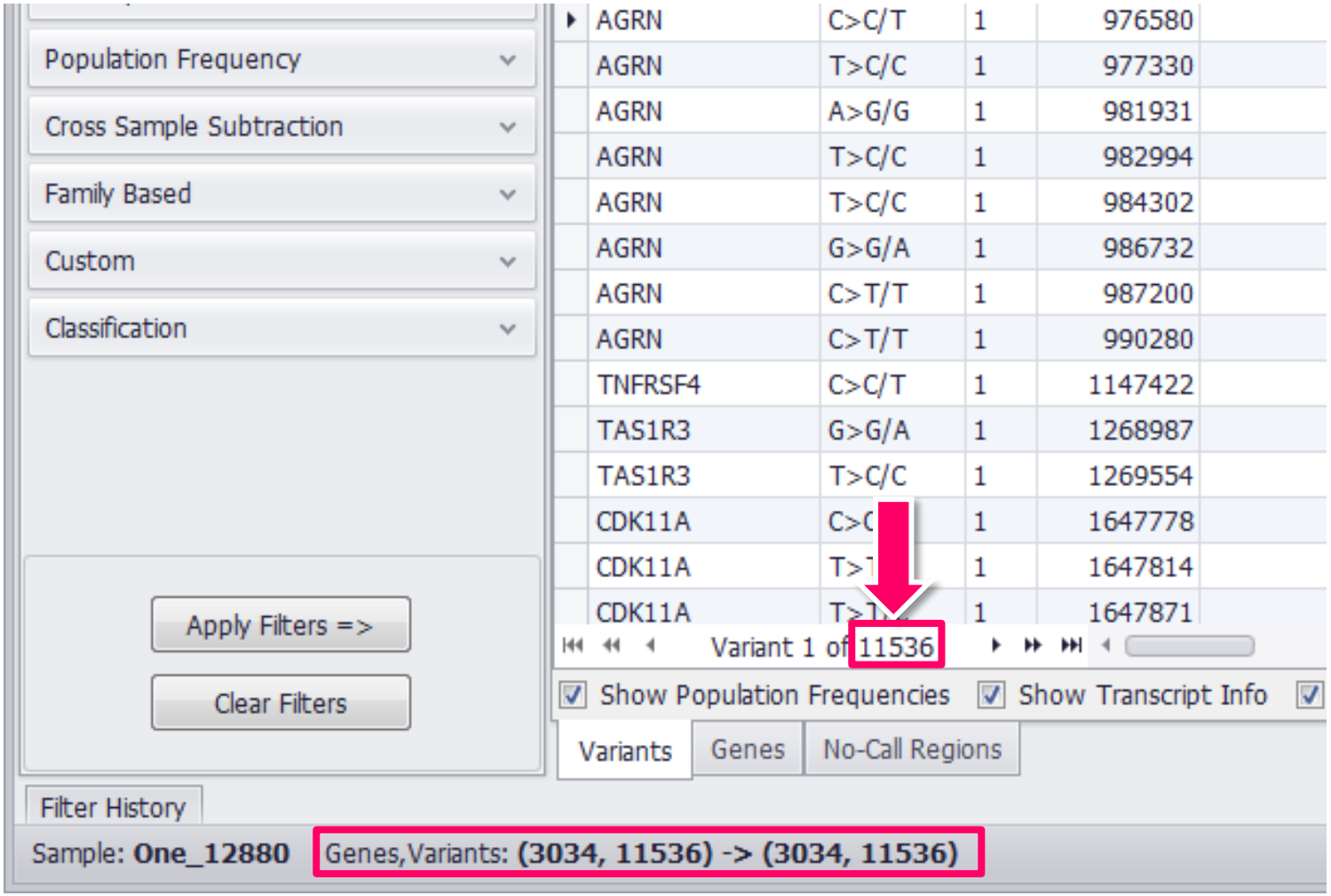
illumina

2. サンプルごとのバリアントリストを表示



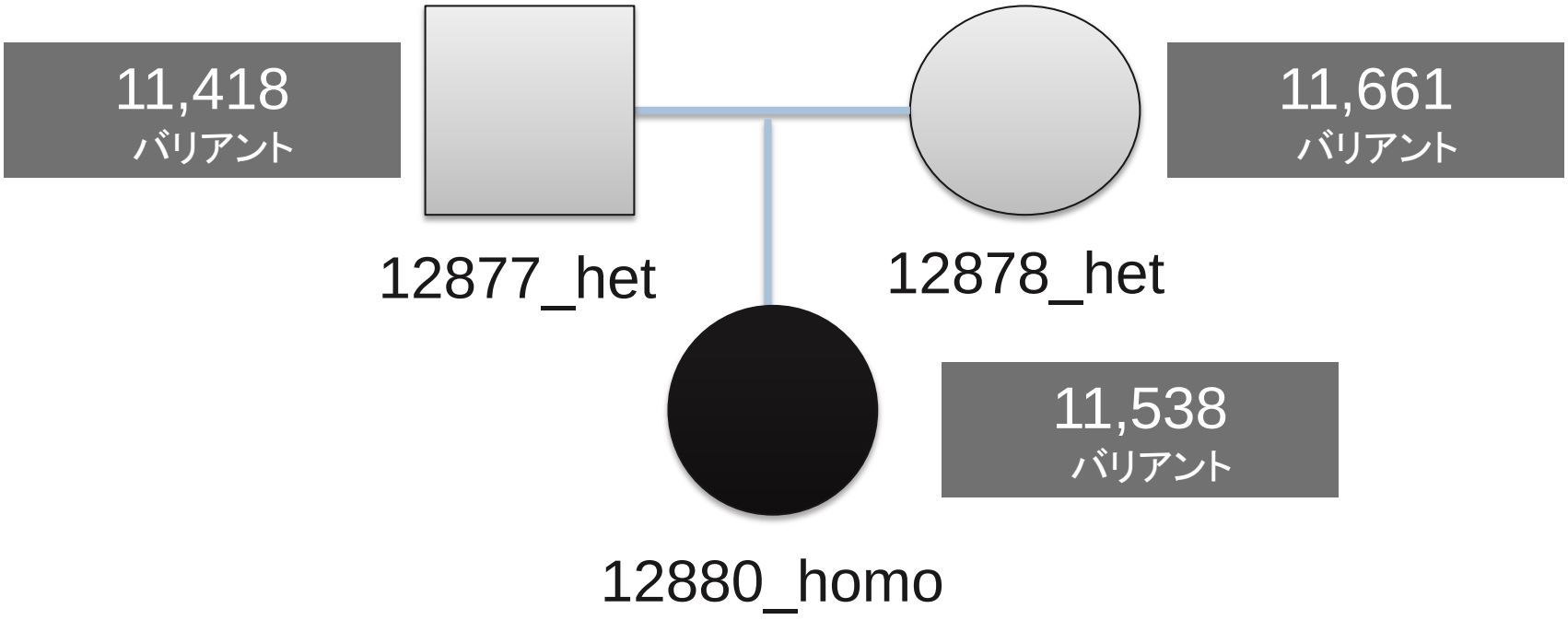
illumina

2. サンプルごとのバリアントリストを表示



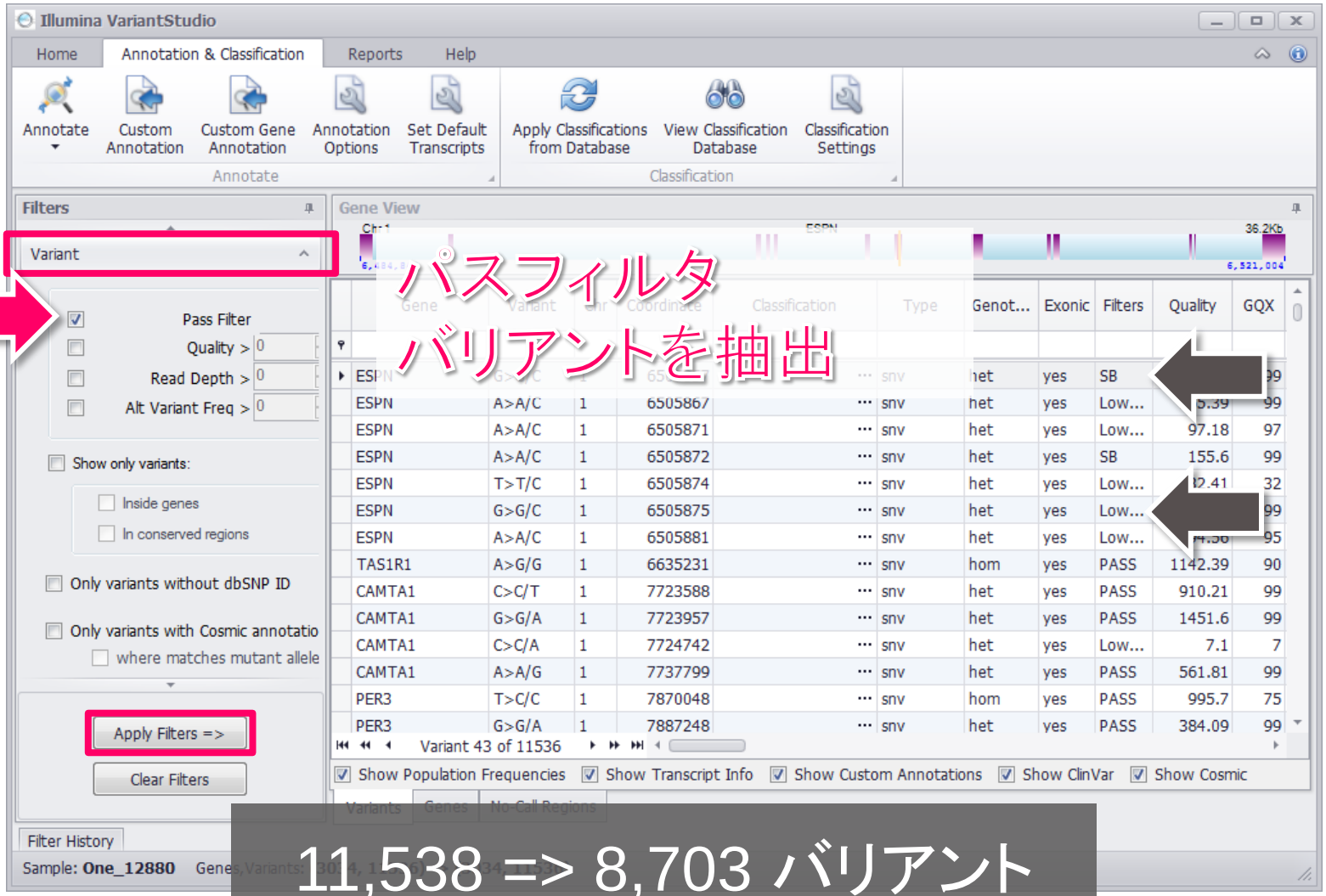
サンプルごとの遺伝子数、バリアント数

デモトリオサンプルのバリアント数



Caseが有する11,538個のバリアントから
疾患原因候補遺伝子の絞り込みを行う

3. クオリティが良いバリアントの絞り込み



illumina

4. 遺伝様式での絞り込み（常染色体劣性遺伝）

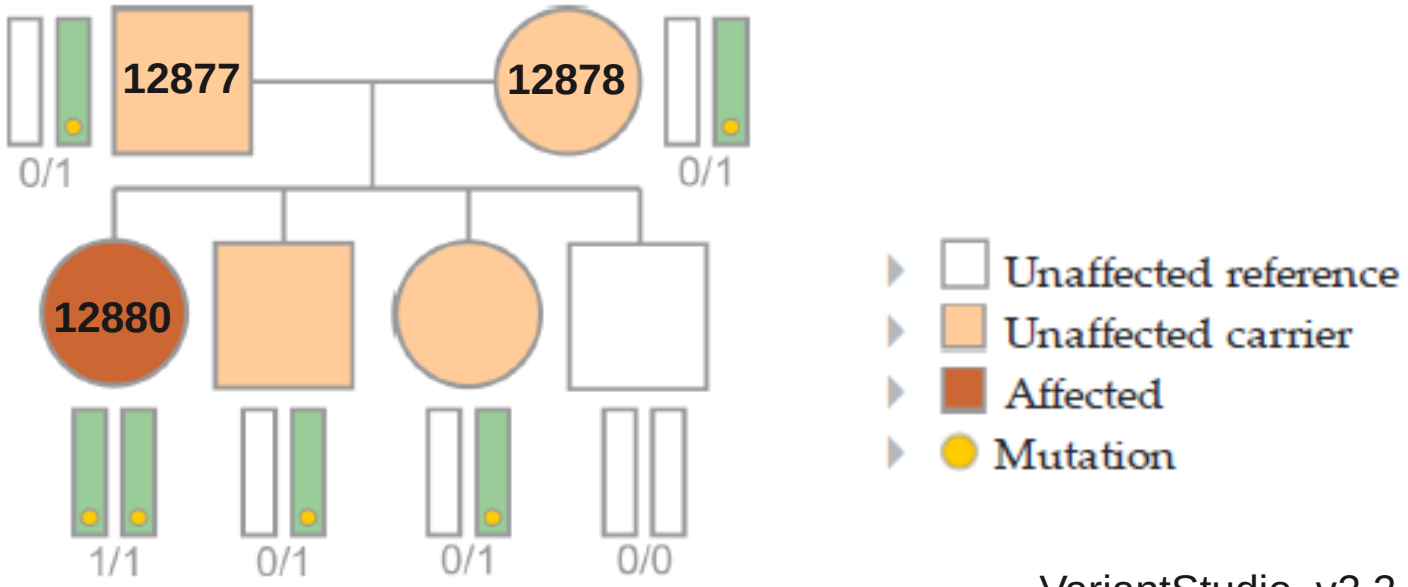
Autosomal Recessive Transmission Workflow

There are two possibilities for recessive transmission.

1 A single gene contains a variant that is:

- Heterozygous (0/1) in the mother
- Heterozygous (0/1) in the father
- Homozygous (1/1) in the affected children

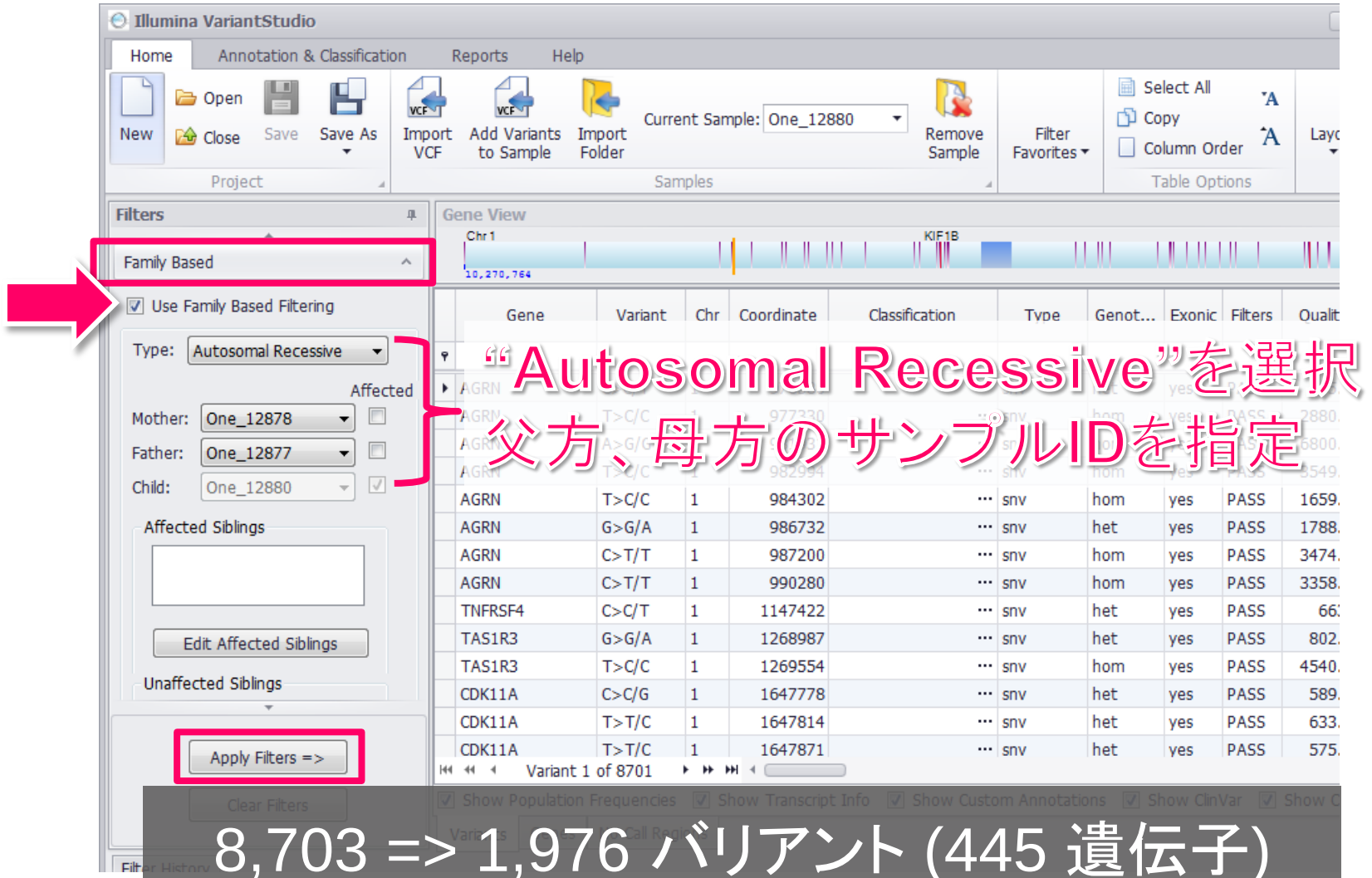
Figure 27 Autosomal Recessive Transmission Logic #1



VariantStudio v2.2
Software User Guide

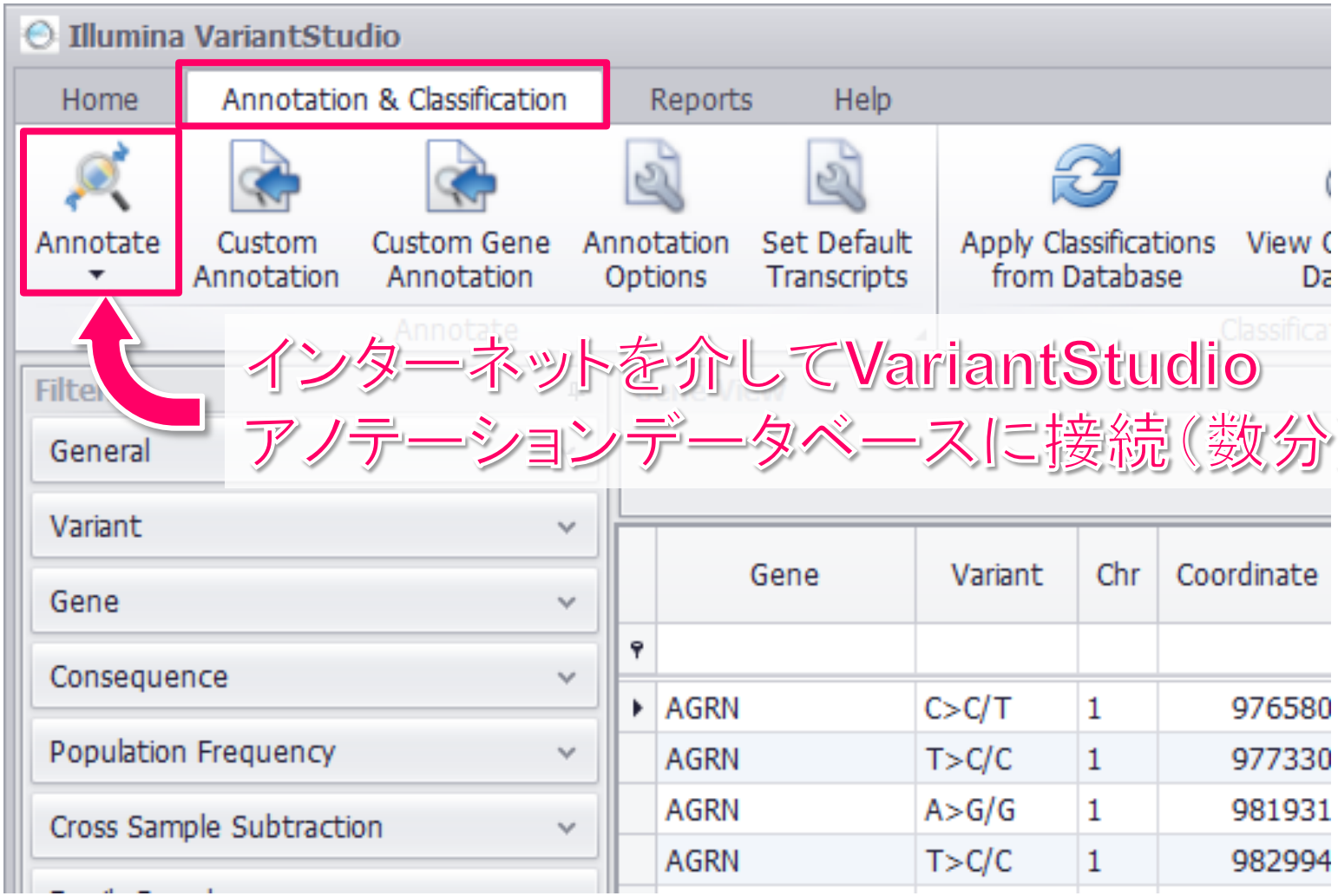
illumina

4. 遺伝様式での絞り込み（常染色体劣性遺伝）



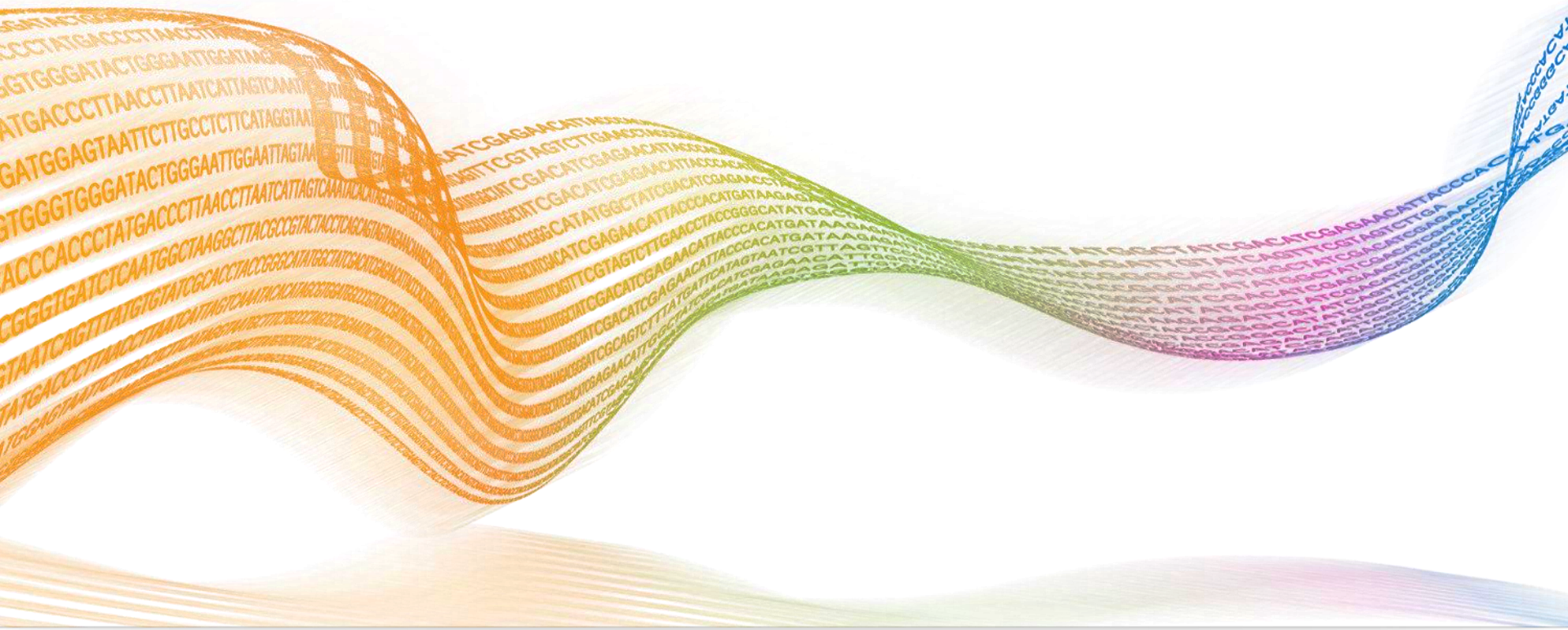
illumina

5. バリアントへのアノテーション（意味付け）

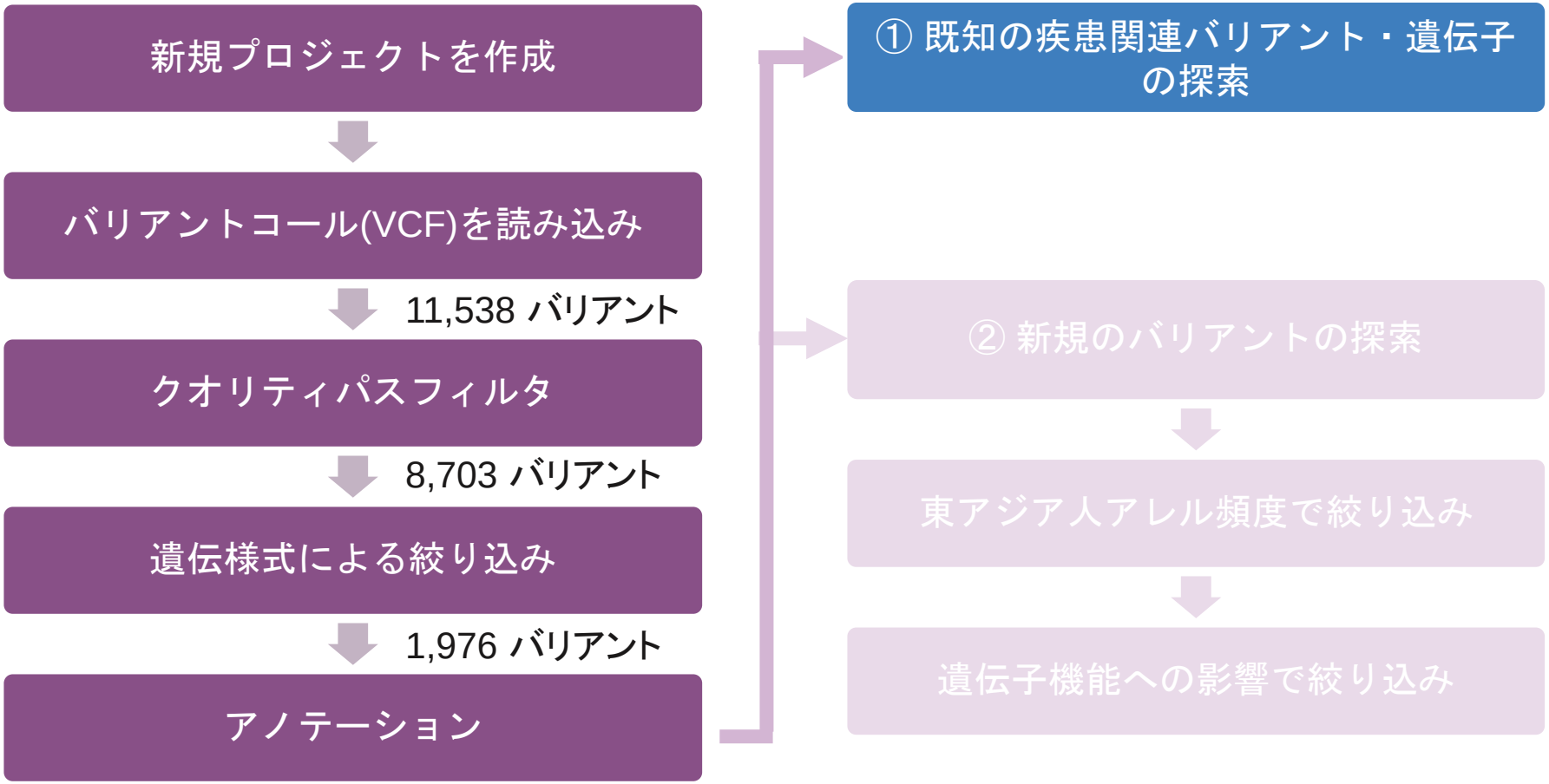


illumina

既知の疾患関連バリアント・遺伝子の探索

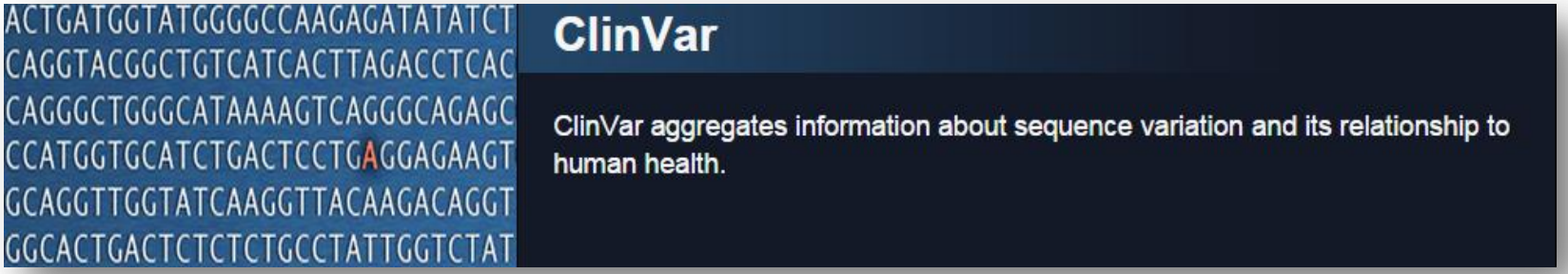


既知の疾患関連バリアント・遺伝子の探索



illumina

ClinVarを用いた疾患関連バリアントの探索

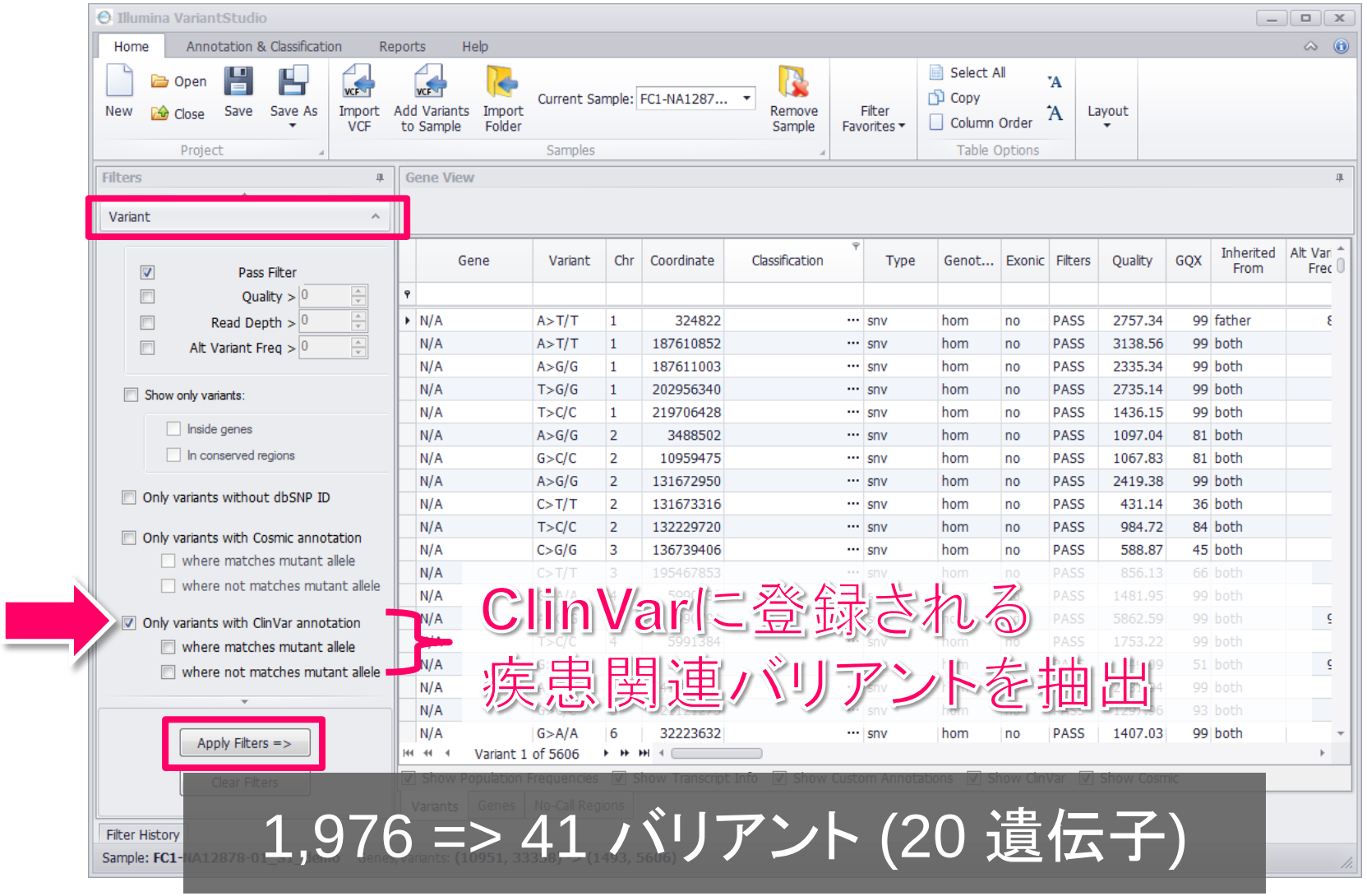


NIHのデータベース：http://www.ncbi.nlm.nih.gov/clinvar/

- 疾患と関連するバリアントのデータベース
- CLIA承認あるいはISO1589のラボで実施された遺伝子研究として報告された臨床的に有意なバリアント
- 研究プロジェクトの一部として、ヒトで同定されたバリアント
- 論文から抽出した効果的なバリアント（GWASは含まない）

illumina

ClinVarに登録されるバリアントの絞り込み



illumina

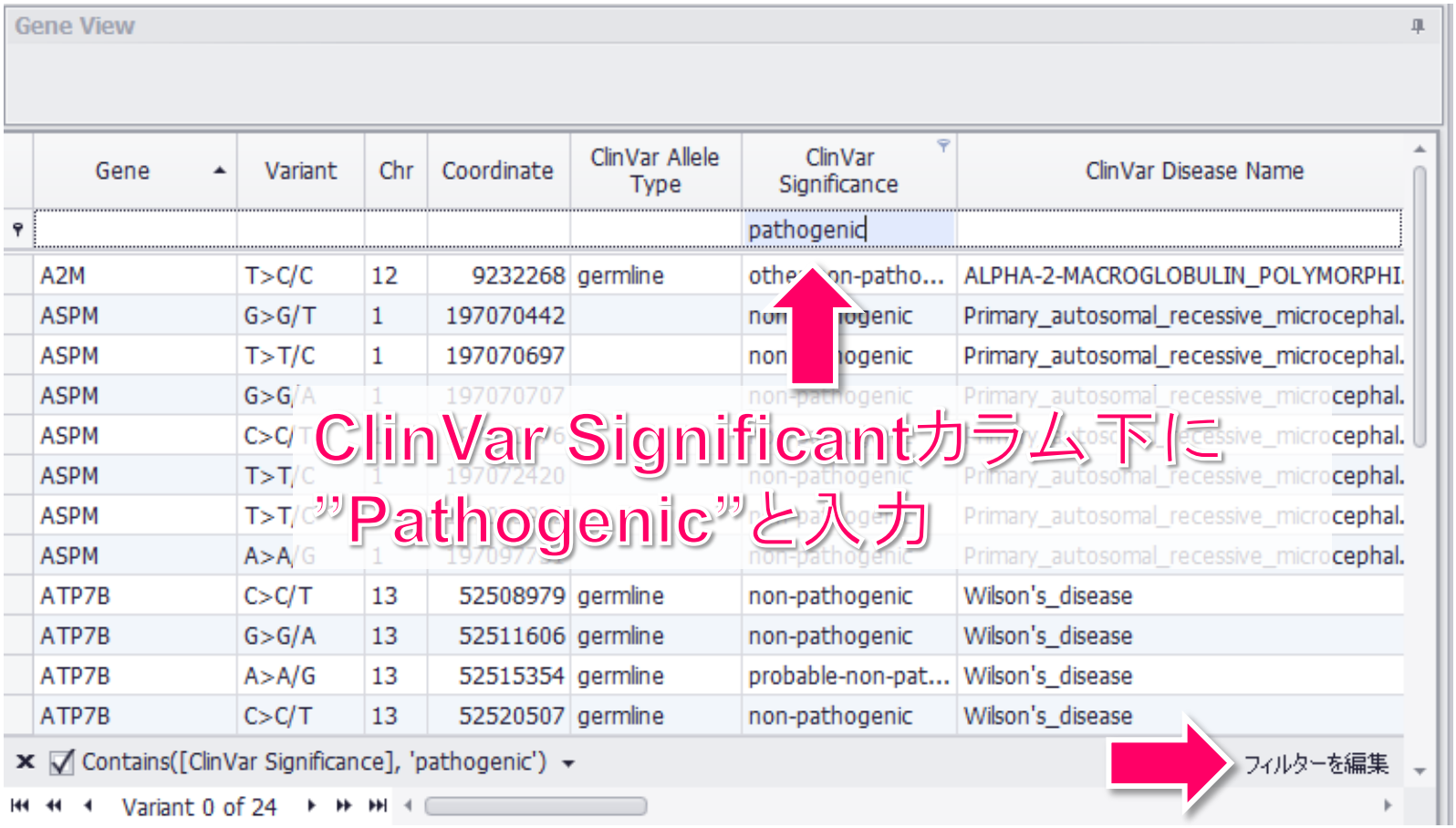
ClinVarの定義によるバリアントの分類

ASN.1 terms	ClinVar and VCF
0 – unknown	Uncertain significance
1 – untested	not provided (includes the cases where data are not available or unknown)
2 - non-pathogenic	Benign
3 - probable-non-pathogenic	Likely benign
4 - probable-pathogenic	Likely pathogenic
5 – pathogenic	Pathogenic
6 - drug-response	drug response
7 – histocompatibility	histocompatibility
255 - other	other
	confers sensitivity
	risk factor
	association
	protective

http://www.ncbi.nlm.nih.gov/clinvar/docs/clinsig/

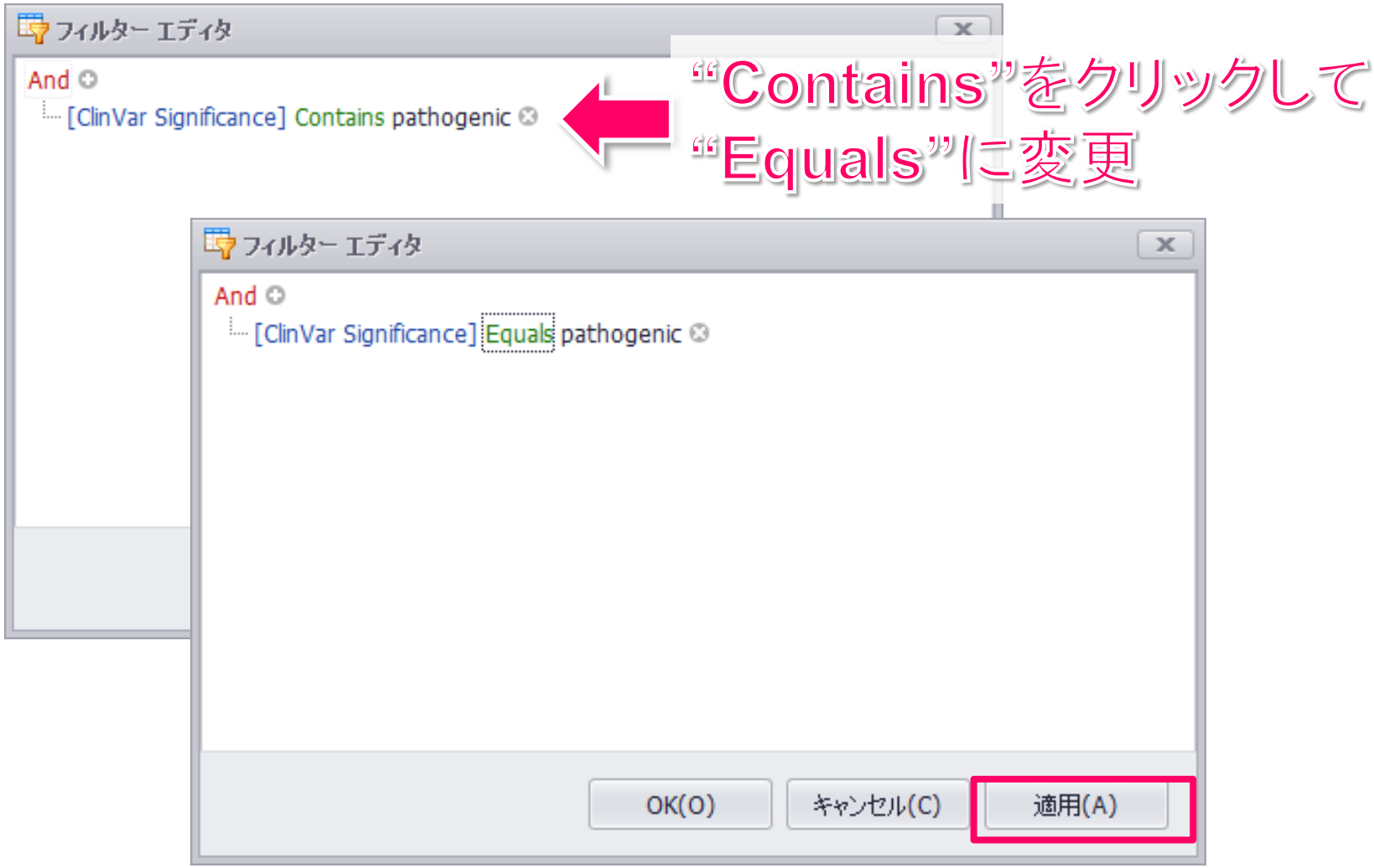
- ClinVarの分類のうち”Pathogenic“（病原性の）バリアントを絞り込む

Pathogenicなバリアントの絞り込み



illumina

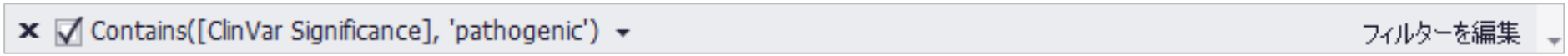
Pathogenicなバリアントの絞り込み



illumina

Pathogenicなバリアントの絞り込み結果

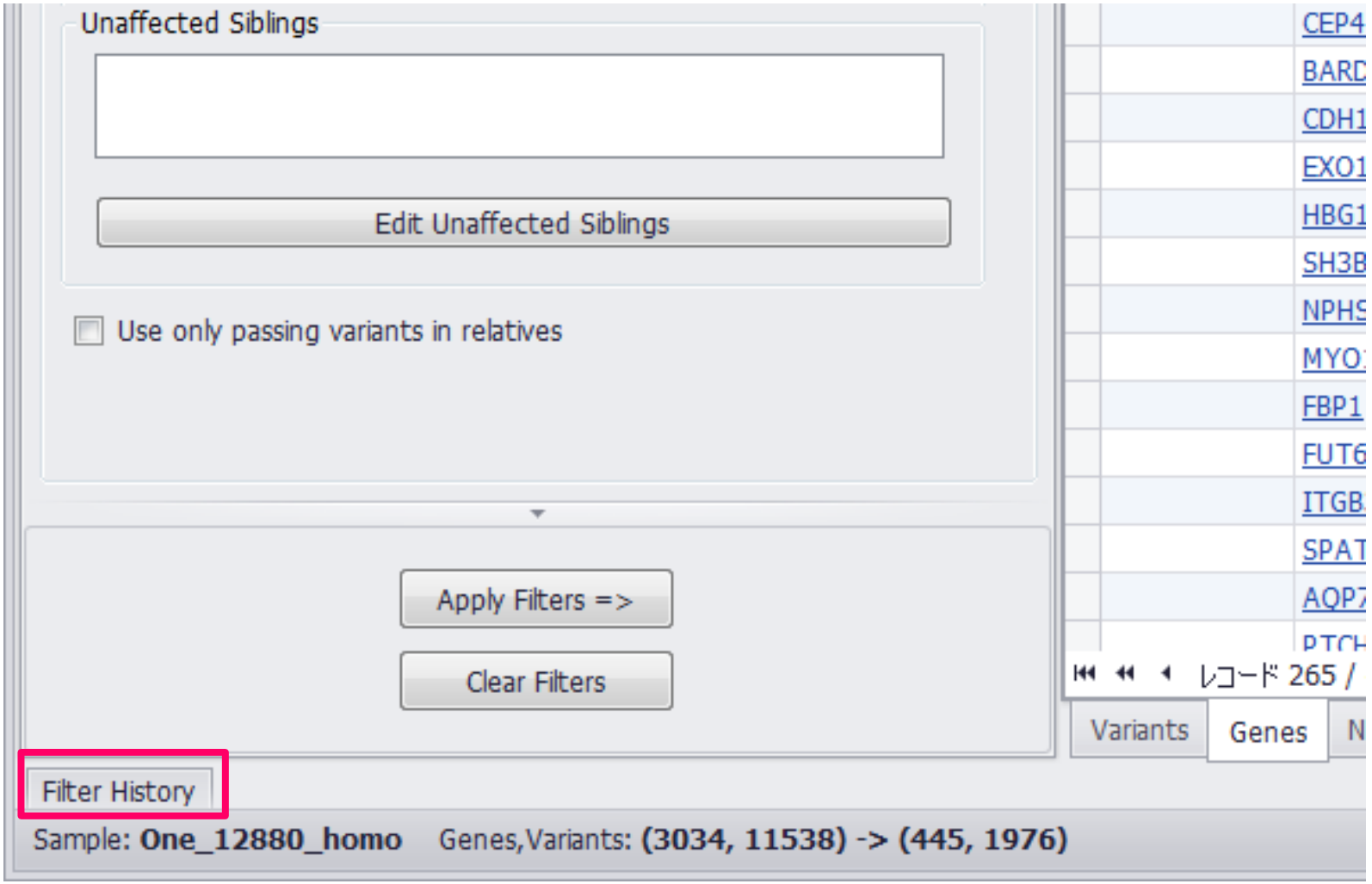
Gene	Variant	Chr	Coordinate	ClinVar Significance	ClinVar Disease Name	Inherited From
DHODH	G>G/A	16	72050942	pathogenic	Miller_syndrome	father
DHODH	C>C/T	16	72057435	pathogenic	Miller_syndrome	mother



フィルタのキャンセルの仕方

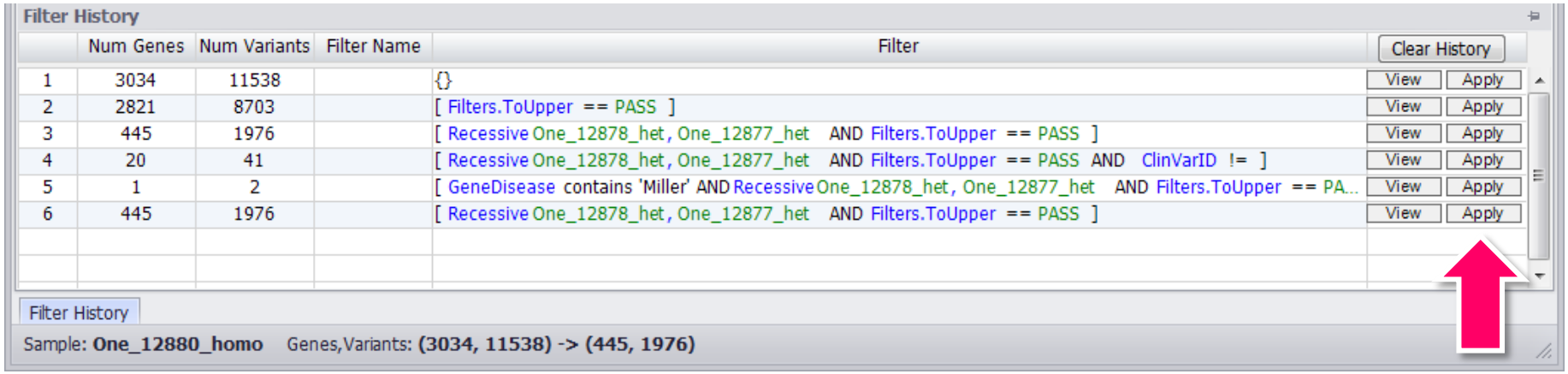
41 => 2 バリアント (1 遺伝子)の絞り込みに成功

絞り込み履歴から過去の絞り込み条件を見る



illumina

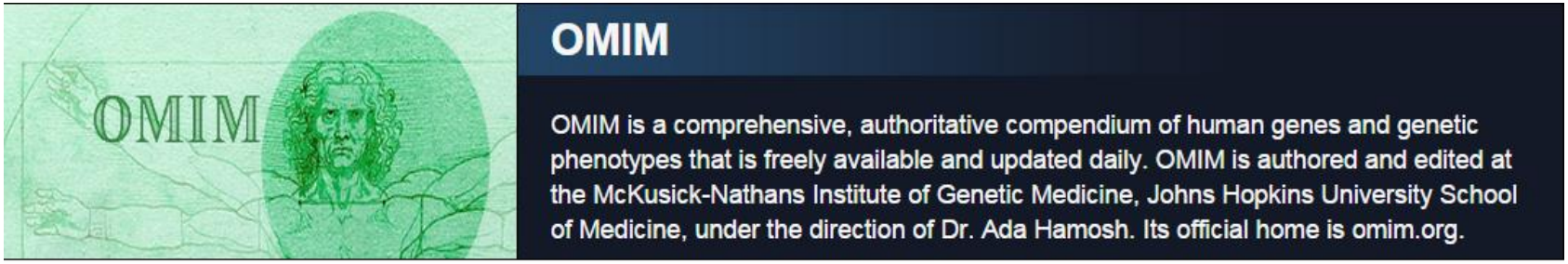
履歴の一覧から戻りたい絞り込み条件を選択



Applyで過去の条件に戻る

3つ目レコードの「Apply」をクリックすることで
アノテーション後の条件に戻る

OMIMを用いた疾患関連遺伝子上のバリアントの探索



NCBIのデータベース：http://www.ncbi.nlm.nih.gov/omim

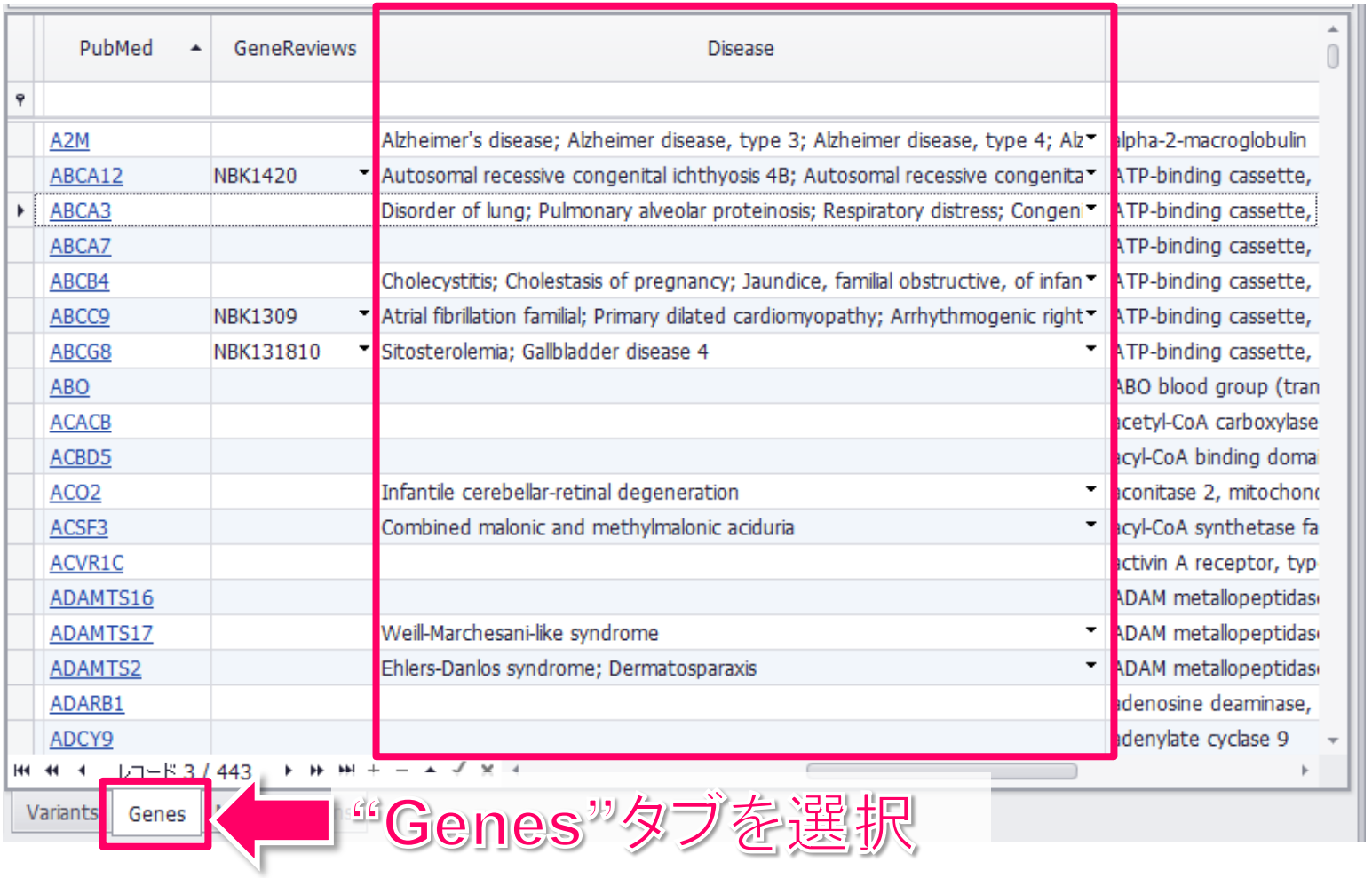
Number of Entries in OMIM (Updated October 10th, 2014)：

Prefix	Autosomal	X Linked	Y Linked	Mitochondrial	Totals
* Gene description	13,963	683	48	35	14,729
+ Gene and phenotype, combined	88	2	0	2	92
# Phenotype description, molecular basis known	3,930	287	4	28	4,249
% Phenotype description or locus, molecular basis unknown	1,542	134	5	0	1,681
Other, mainly phenotypes with suspected mendelian basis	1,730	113	2	0	1,845
Totals	21,253	1,219	59	65	22,596

- 疾患や表現型と関連する遺伝子やゲノム情報のデータベース
- メンデル疾患、多因子疾患にも対応

遺伝子と関連する疾患名の表示

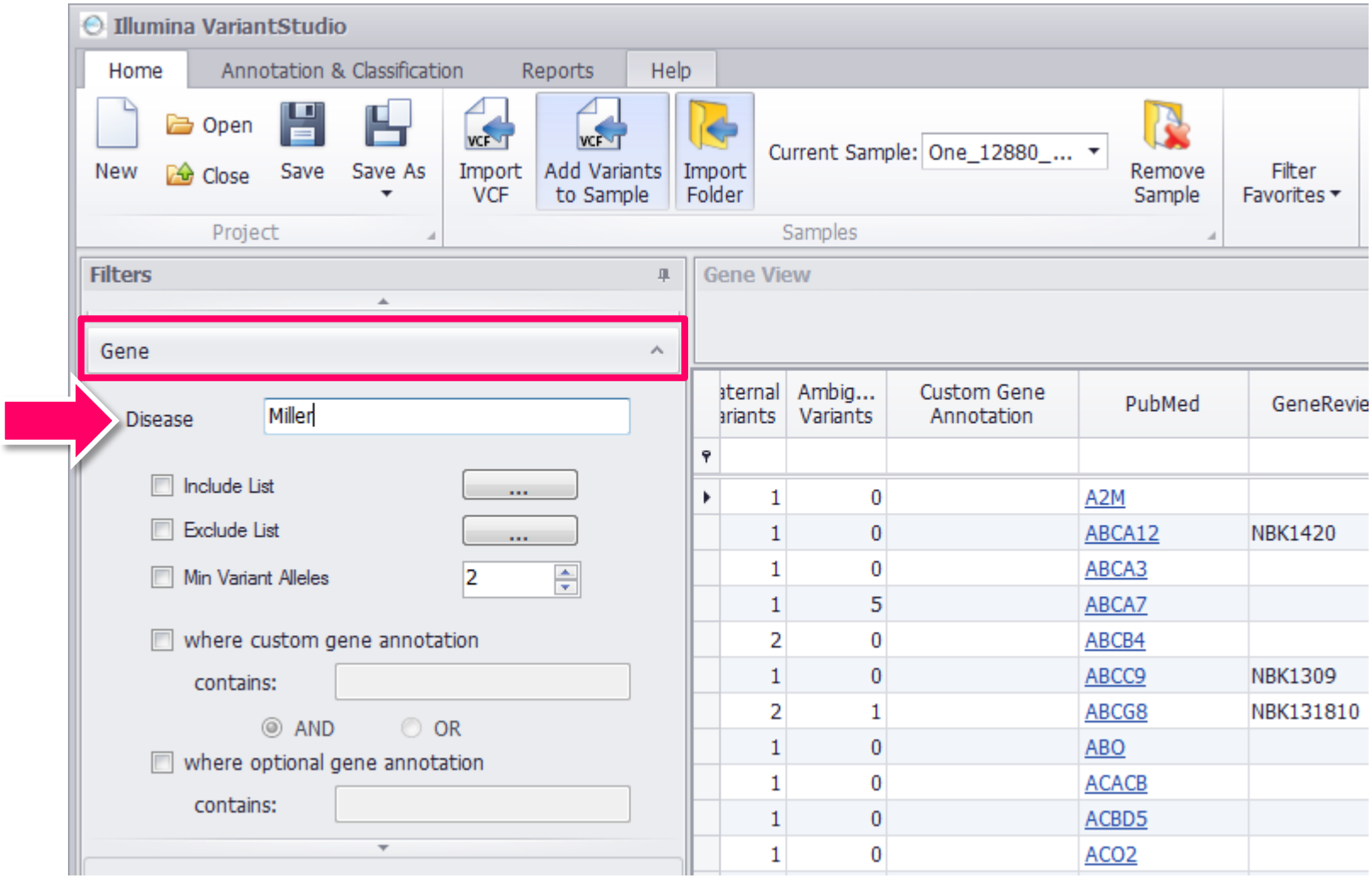
OMIM由来アノテーション



“Genes”タブを選択

illumina

疾患名の入力による疾患関連遺伝子の絞り込み



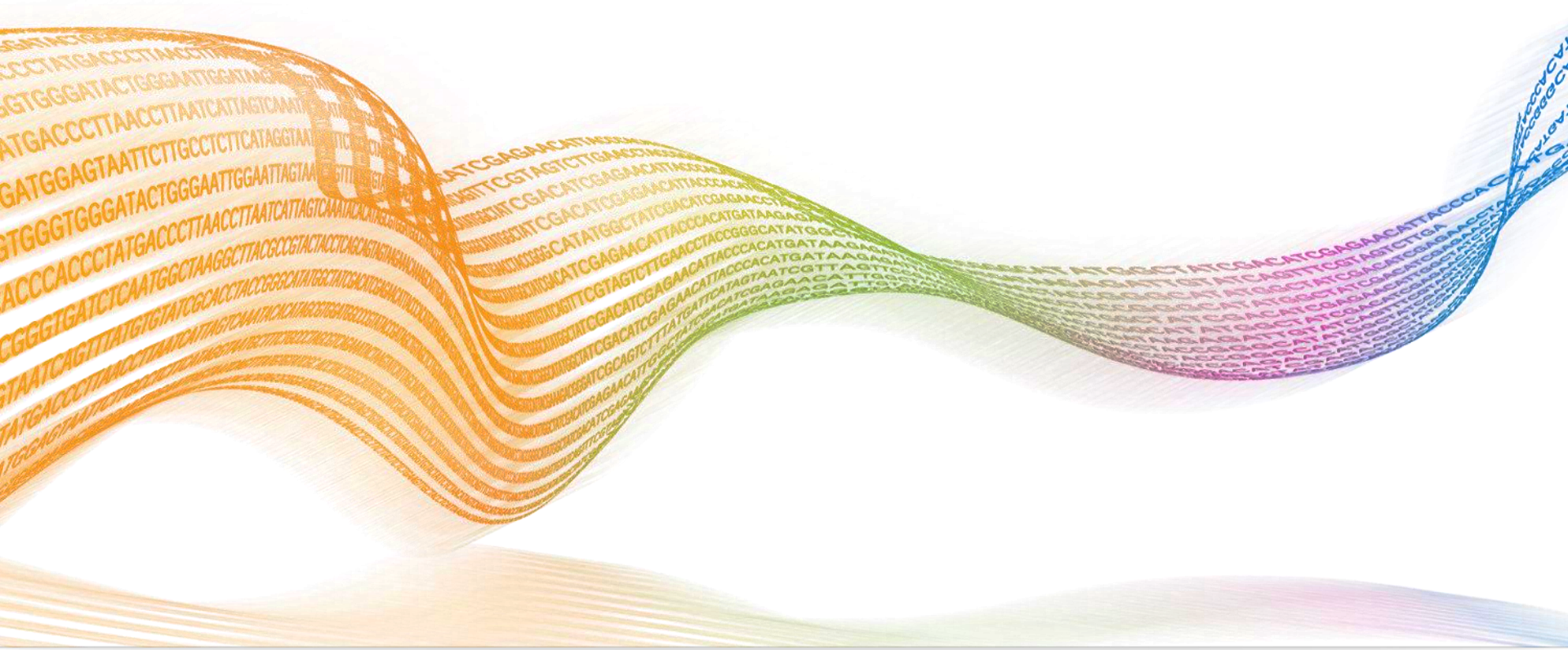
illumina

既知の疾患関連遺伝子の絞り込み結果

Gene	Variant	Chr	Coordinate	ClinVar Significance	ClinVar Disease Name	Inherited From
DHODH	G>G/A	16	72050942	pathogenic	Miller_syndrome	father
DHODH	C>C/T	16	72057435	pathogenic	Miller_syndrome	mother

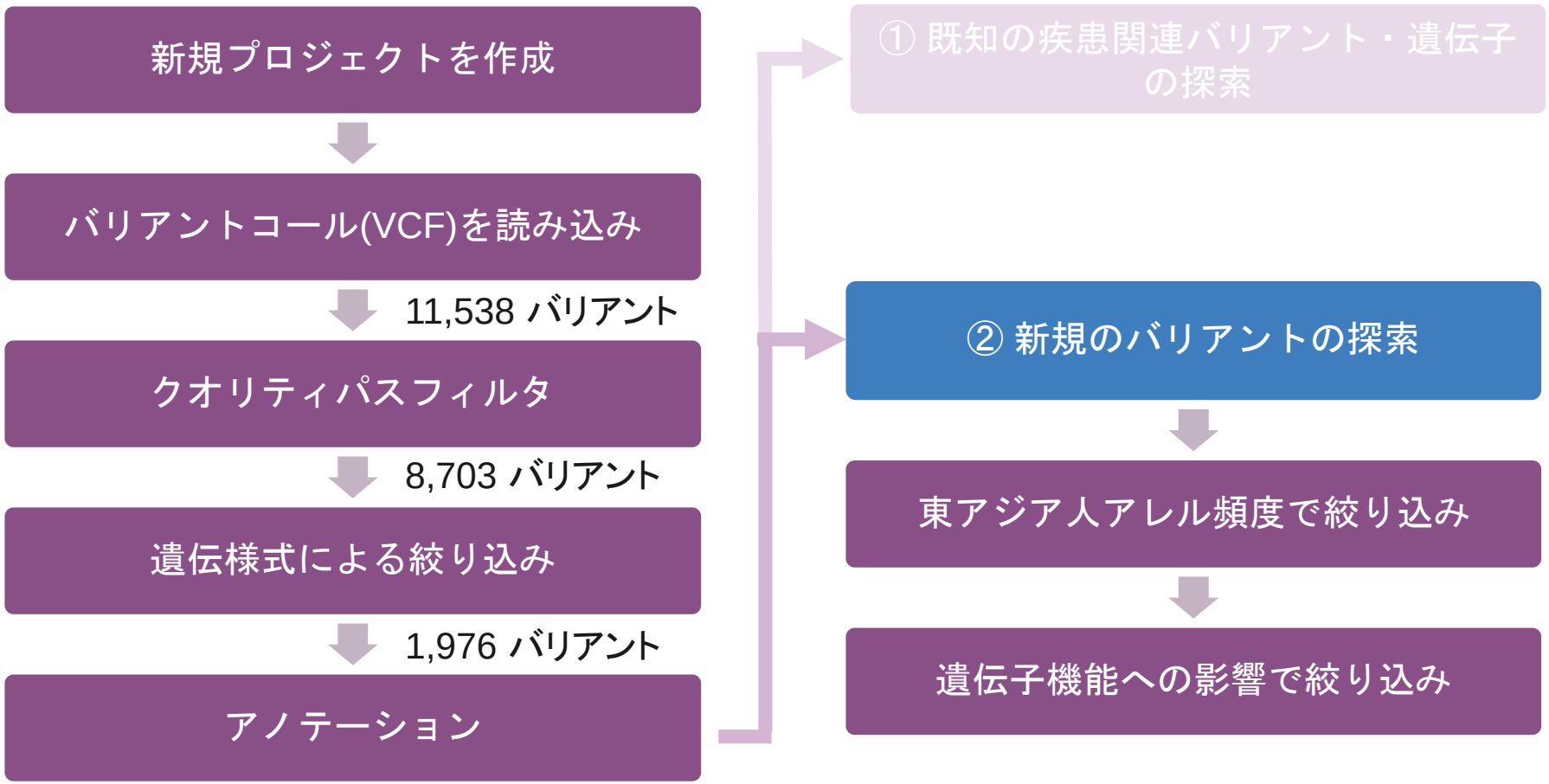
1,976 => 2 バリアント (1 遺伝子)の絞り込みに成功

既知情報のない新規バリアントの探索

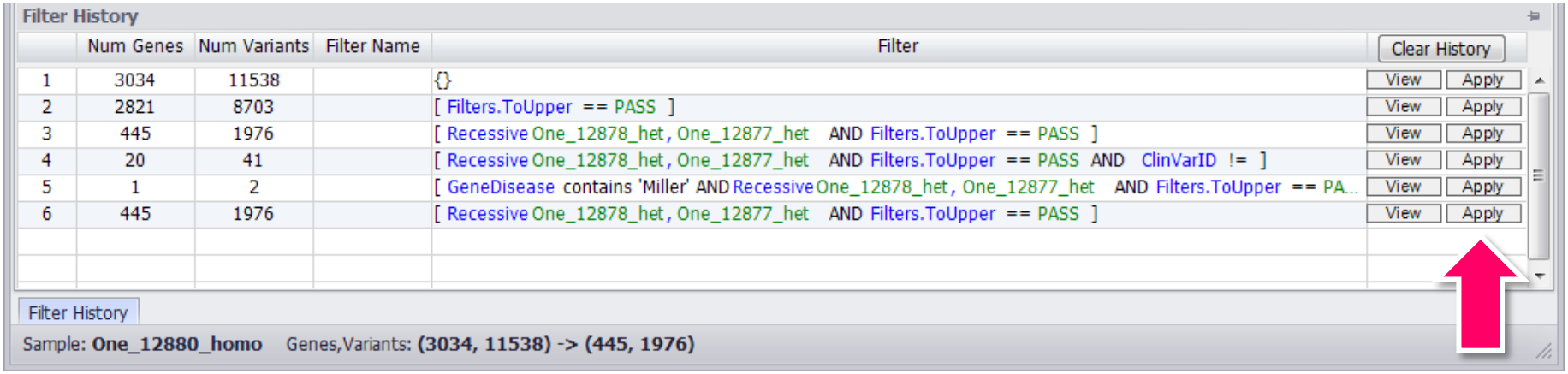


illumina®

② 既知情報のない新規バリアントの探索



履歴の一覧から戻りたい絞り込み条件を選択

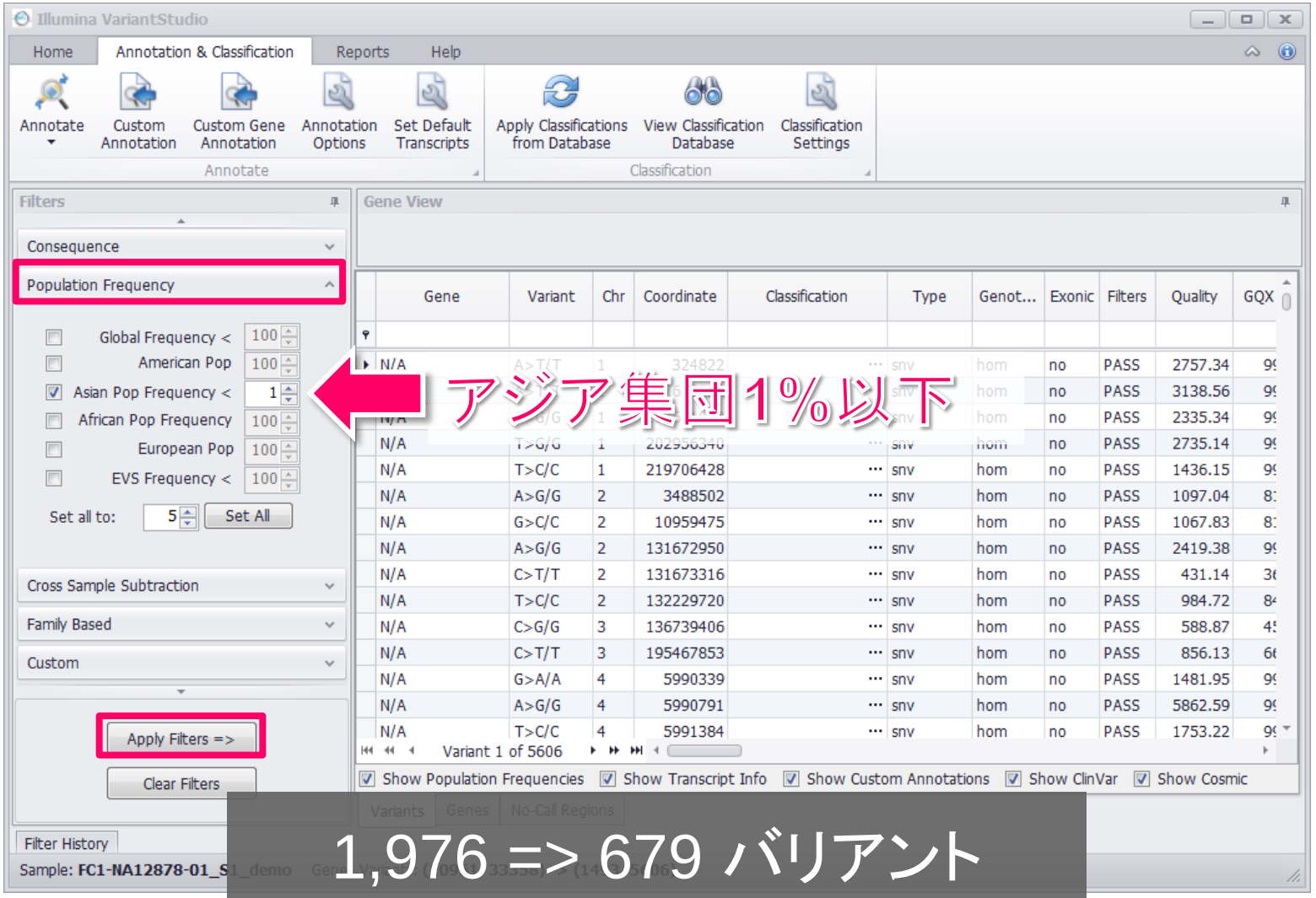


Applyで過去の条件に戻る

3つ目レコードの「Apply」をクリックすることで
アノテーション後の条件に戻る

illumina

6. アリル頻度が低いバリアントの絞り込み



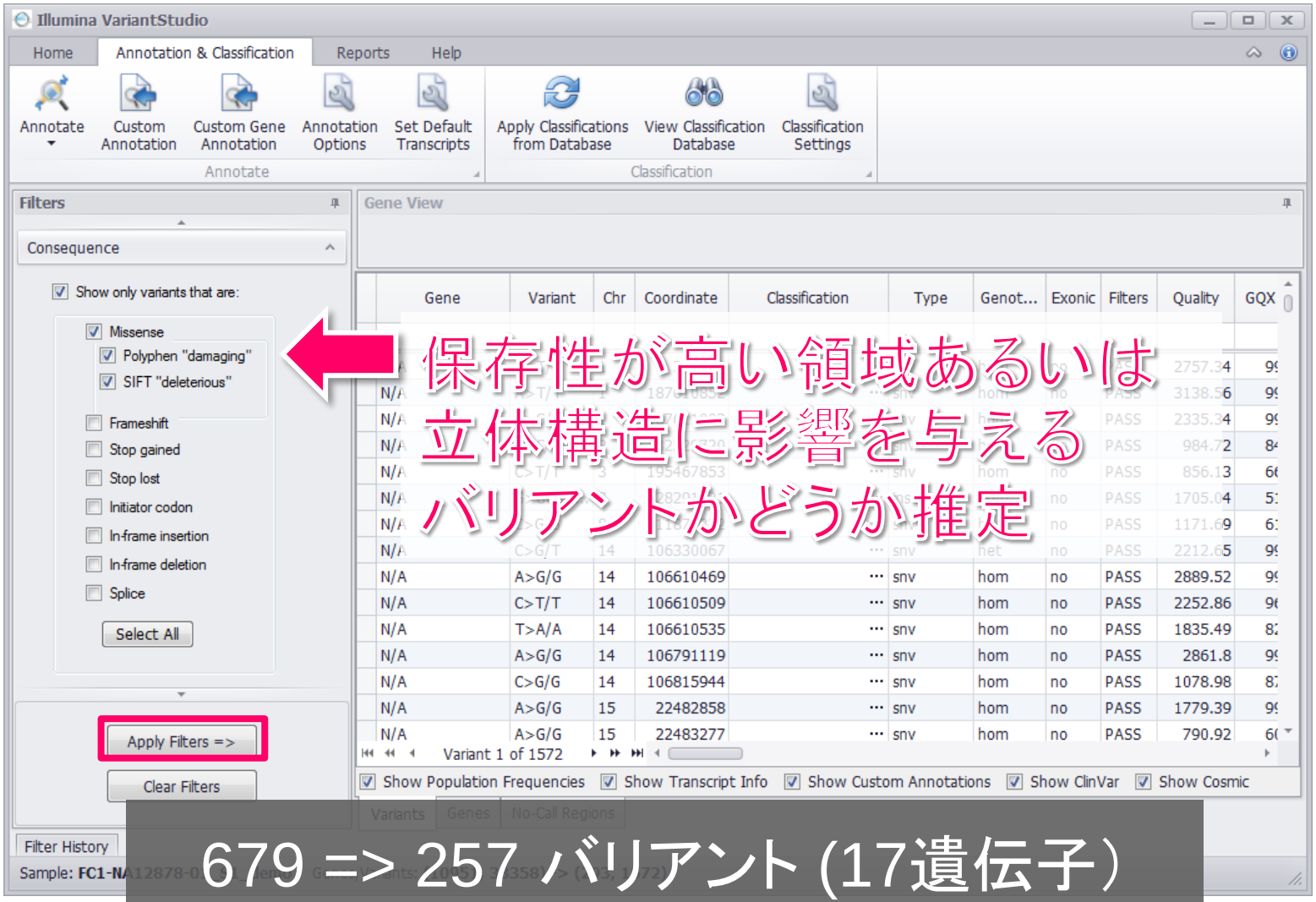
illumina

7. 遺伝子機能に影響を与えるバリアントの絞り込み

- Polyphen: タンパク質の立体構造に影響を与えるインパクトが強いほど（疎水性など）Damagingとする。

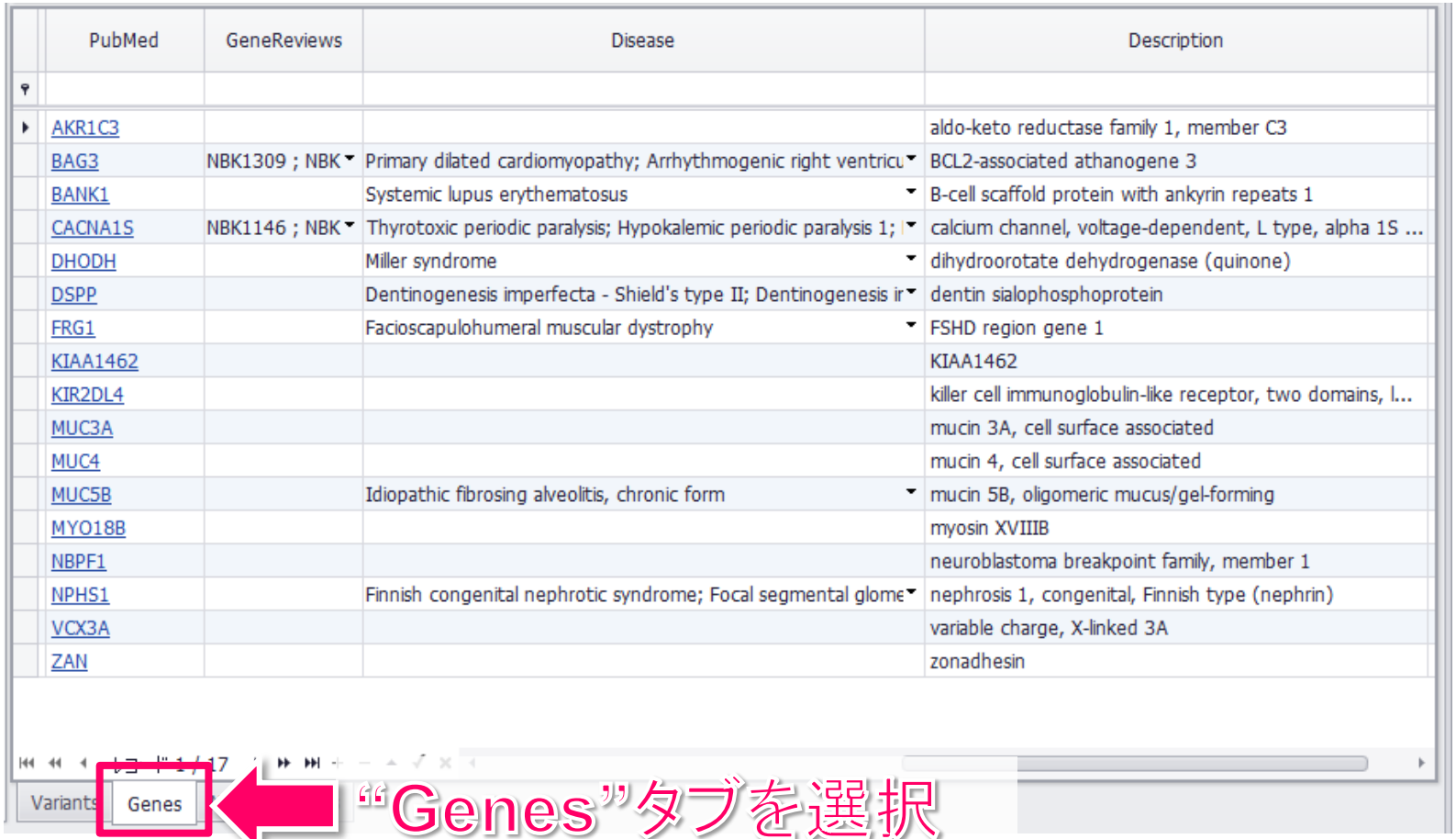
- SIFT: タンパク質ファミリーで進化的に保存性が高いアミノ酸ほどDeleteriousとする。

7. 遺伝子機能に影響を与えるバリアントの絞り込み



illumina

7. 絞り込んだ遺伝子に関連する疾患等の情報

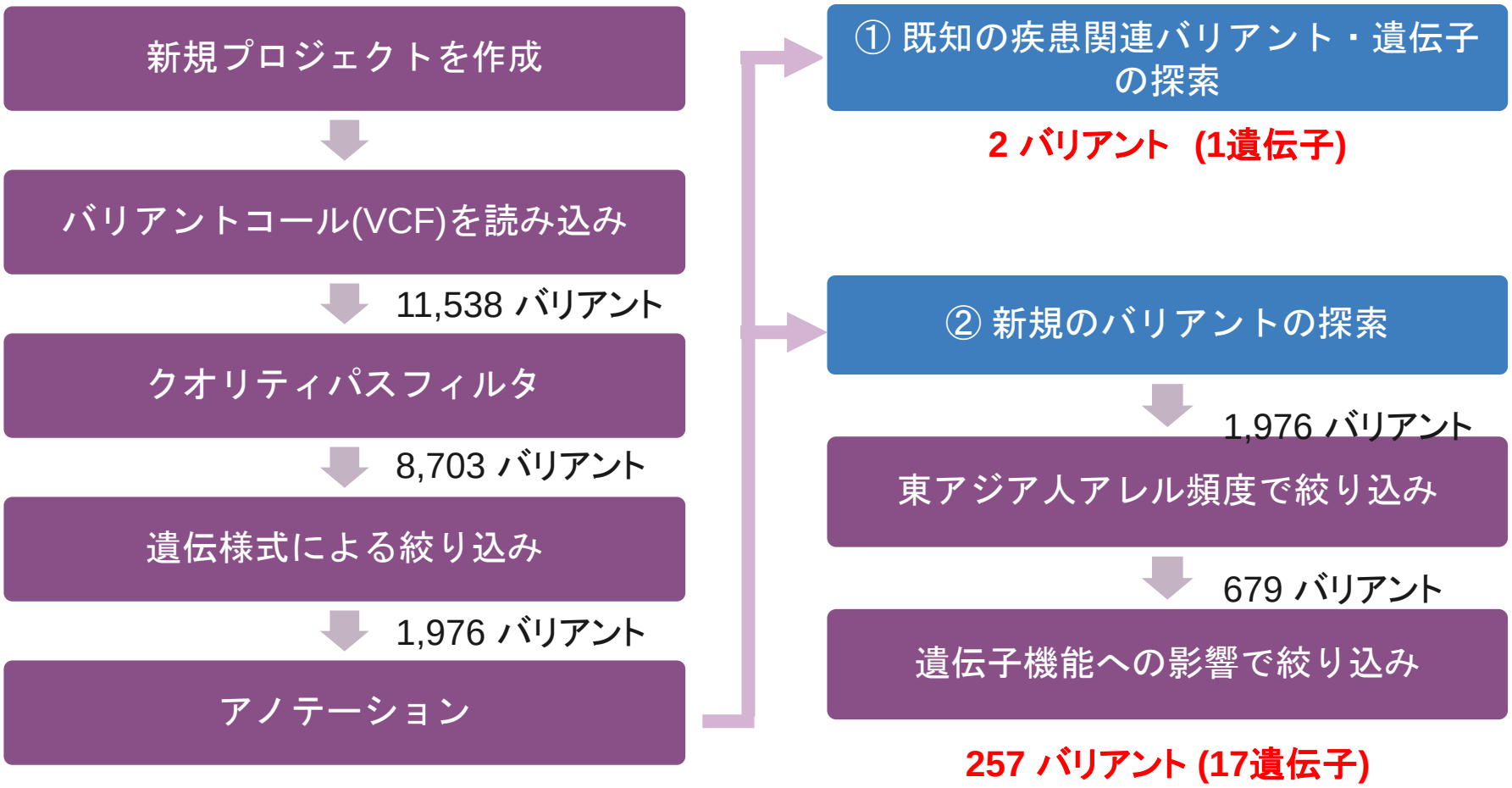


illumina

8. 原因バリアントの絞り込みと検証例

1. **過去の文献や知見を調査**
   - 注目する遺伝子と関連する遺伝性疾患と類似性はあるか
   - 代謝経路やシグナル伝達経路が疾患と関与ありそうか
   - ノックアウトマウスの表現型が疾患と関連ありそうか
2. **技術的な検証**
   - サンガー法シーケンサーで検証
3. **再現性の検証**
   - 症例数を増やして再現性があるか検証
4. **生化学的実験による検証**
   - モデル生物を用いた遺伝子機能解析
   - *in vitro* での遺伝子機能解析

VariantStudio解析ワークフローまとめ



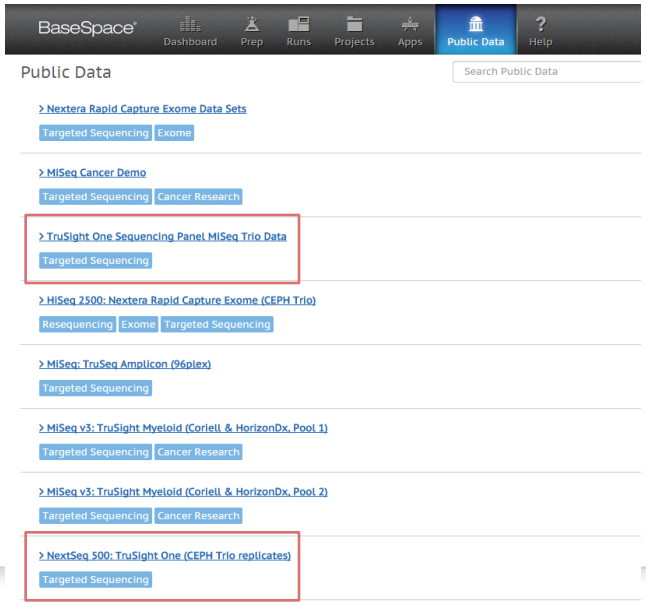
illumina

参考文献一覧

- VEP v2.8 (Ensembl)
  - http://asia.ensembl.org/info/docs/tools/vep/index.html
  - McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. **Bioinformatics** 26(16):2069-70(2010)
- PolyPhen
  - http://genetics.bwh.harvard.edu/pph2/
  - Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. **Nat Methods** 7(4):248-249 (2010).
- SIFT
  - http://sift.jcvi.org/
  - Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. **Nat Protoc**. 2009;4(7):1073-81.
- ClinVar Version September 5, 2013 (NCBI)
  - www.ncbi.nlm.nih.gov/clinvar/
  - Landrum MJ *et al.* ClinVar: public archive of relationships among sequence variation and human phenotype. **Nucleic Acids Res.** 2014 Jan 1;42(1):D980-5.

TruSight One その他 資料

- TruSight One User Guideなど
  http://support.illumina.com/sequencing/sequencing_kits/trusight_one_kit/documentation.html
- TruSight One Data Sheetなど
  http://support.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet_trusight_one_panel.pdf
- TruSight One 遺伝子一覧
  http://products.illumina.com/content/dam/illumina-marketing/documents/products/gene_lists/gene_list_trusight_one.zip
- TruSight One BaseSpace Public data
  https://basespace.illumina.com/datacentral



illumina®

ご清聴ありがとうございました!

本日セッション終了後のご質問は、
techsupport@illumina.comにお問い合わせください

テクニカルサポート直通のフリーダイヤルも
ご利用くださいませ。
0800-111-5011