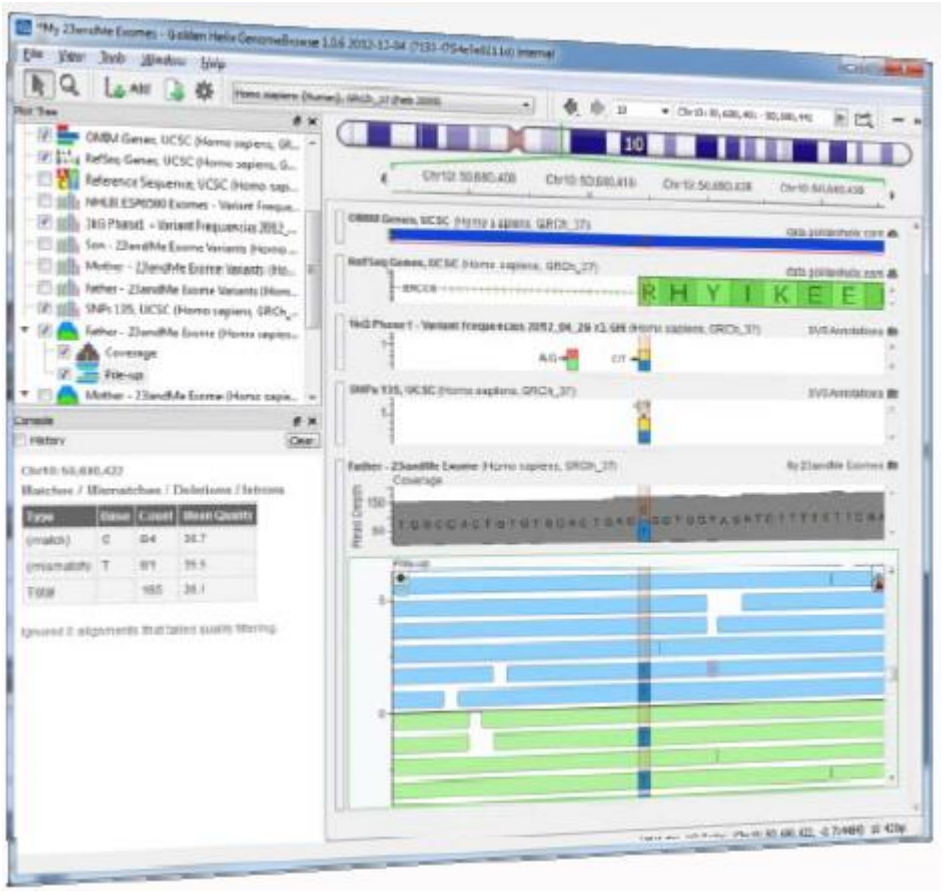
無料ゲノムブラウザーを用いた遺伝子変異解析

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



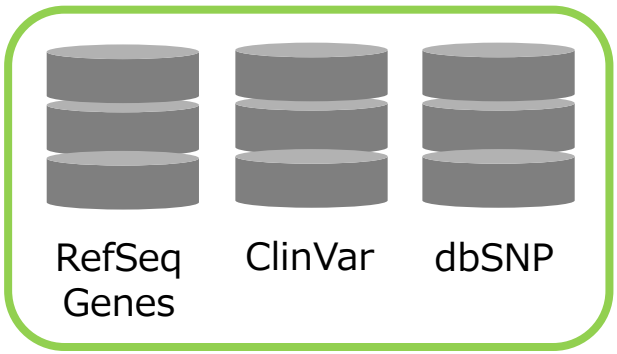
- 次世代シークエンサーより得られたシークエンスデータのクオリティチェックやビジュアライゼーションには、ゲノムブラウザーが用いられる

- Golden Helix社がアカデミックフリーで配布している高機能なゲノムブラウザーGenomeBrowse®を使用することにより、シークエンスデータや遺伝子変異データのクオリティチェックやビジュアライゼーション、さらに変異の各種データベースへの照合を容易に実行することができる

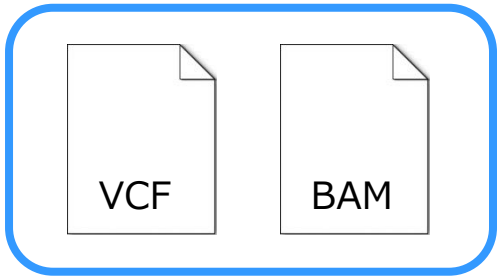
GenomeBrowse®

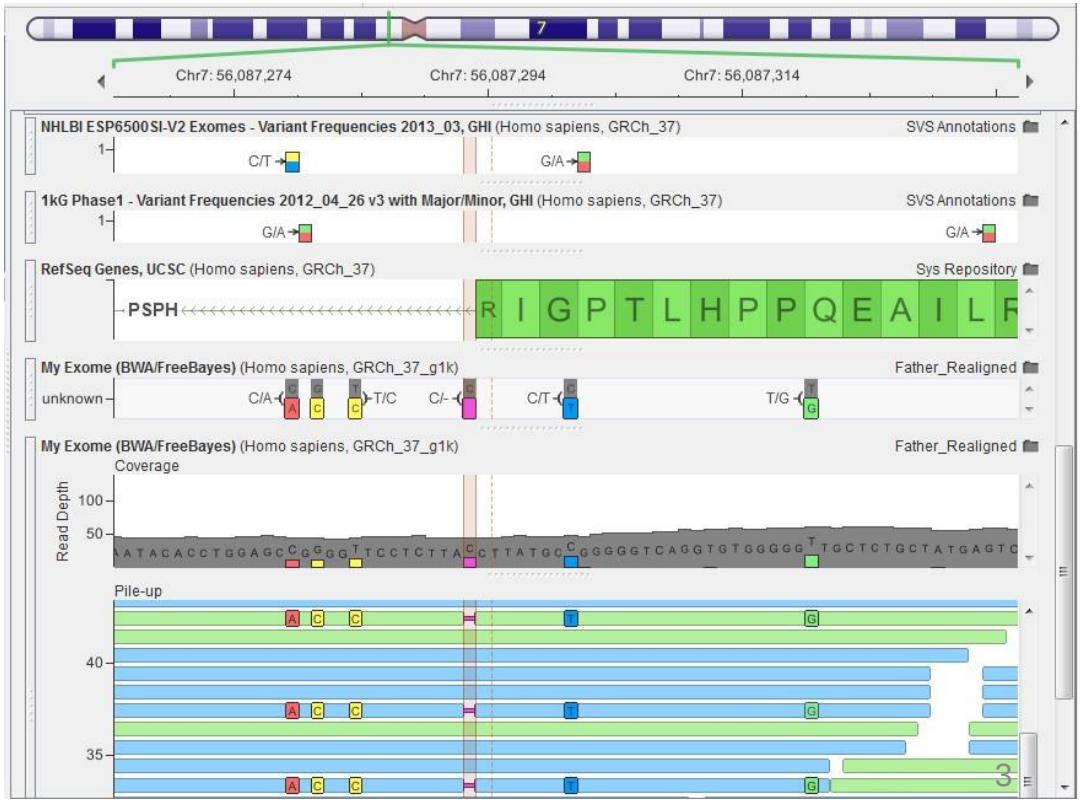
- Golden Helix社が無料で配布しているゲノムブラウザーソフトウェア
- 次世代シークエンサーのデータ解析パイプラインより出力されたVCFファイルやBAMファイルを、ドラッグアンドドロップで表示可能
- データベースマネージャーより、キュレーション済みの様々なデータベースをワンクリックでダウンロードし、アノテーションとして表示が可能

✓ データベース

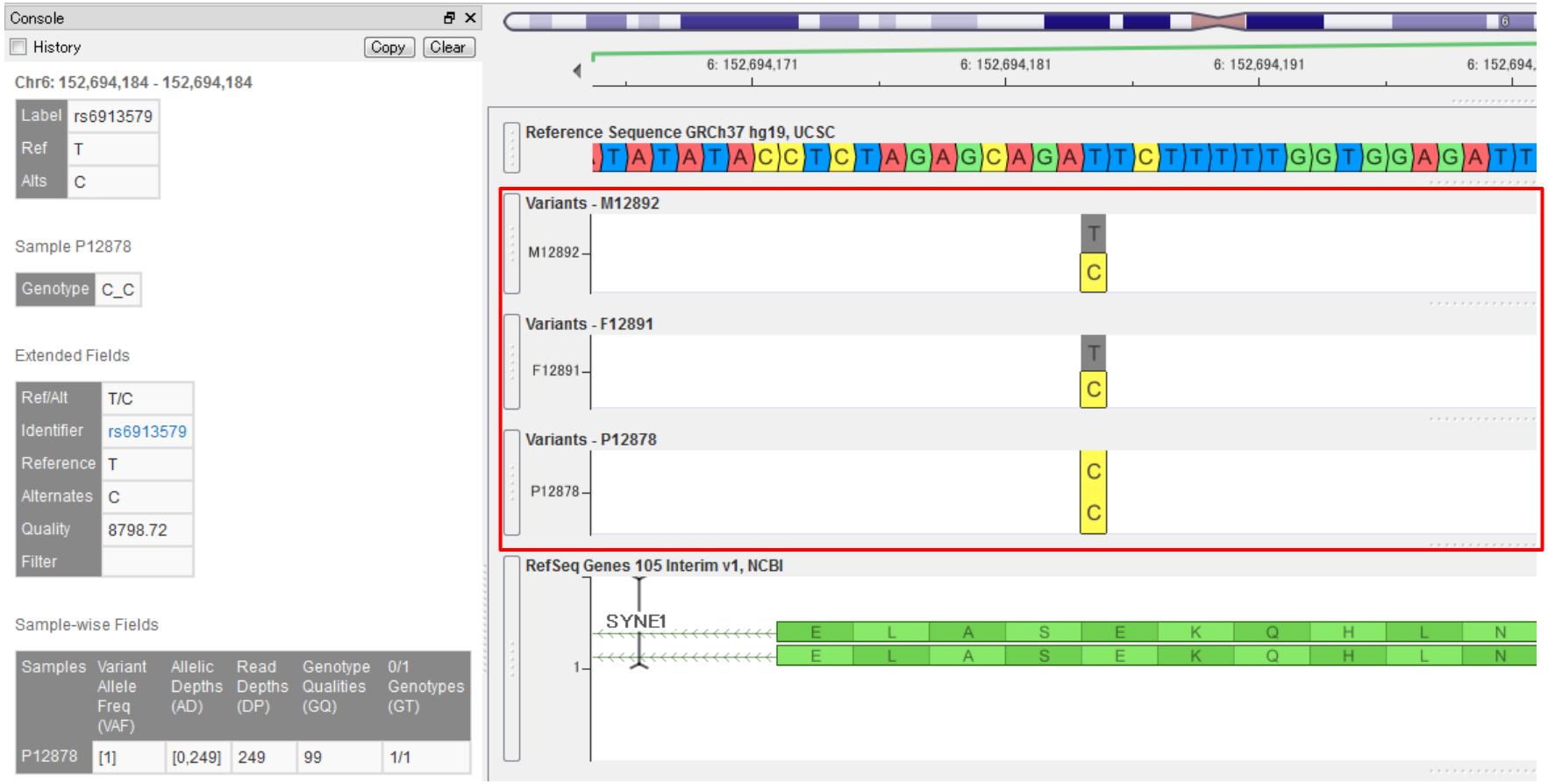


✓ サンプルデータ



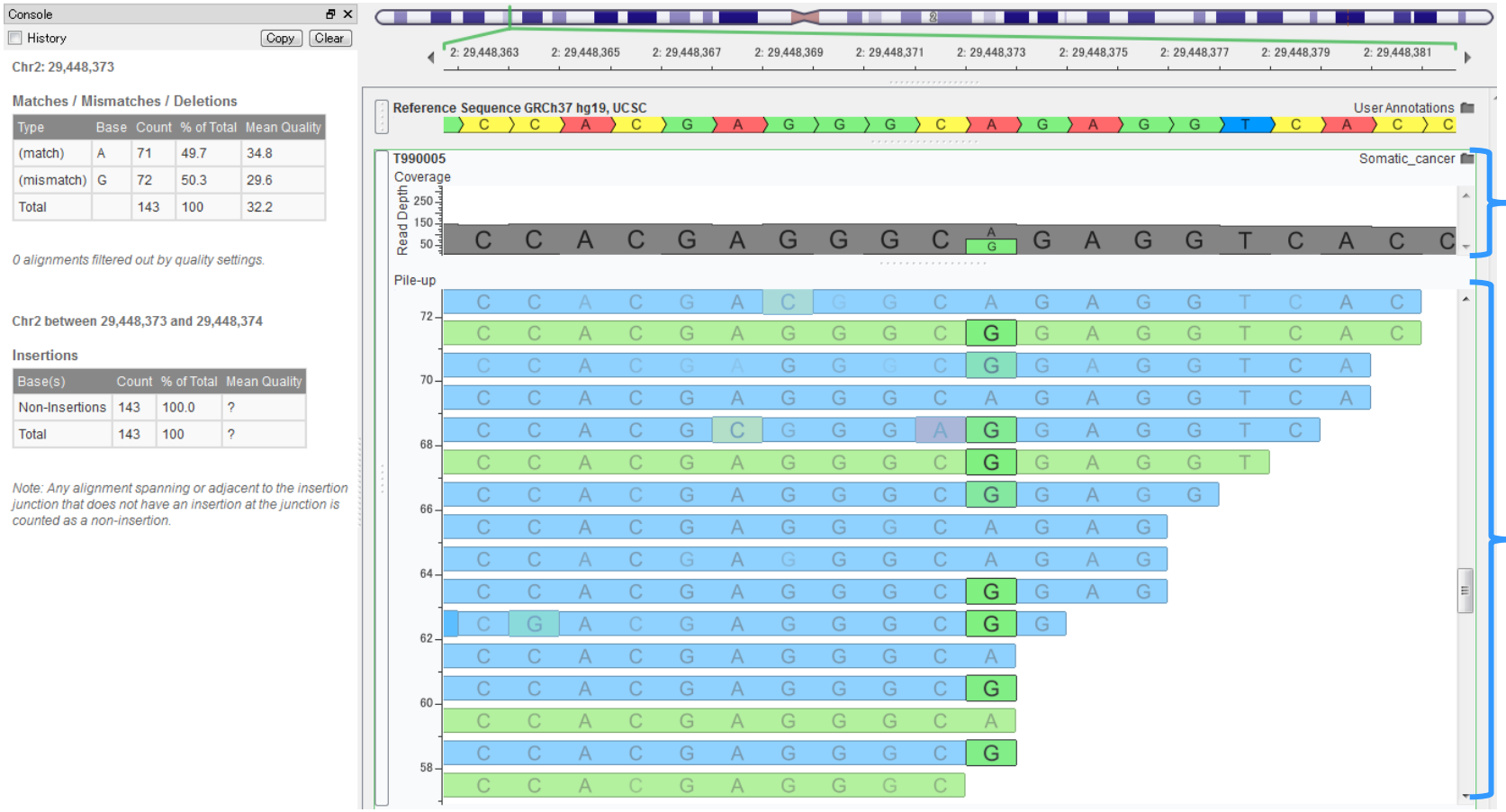


VCFファイルプロット



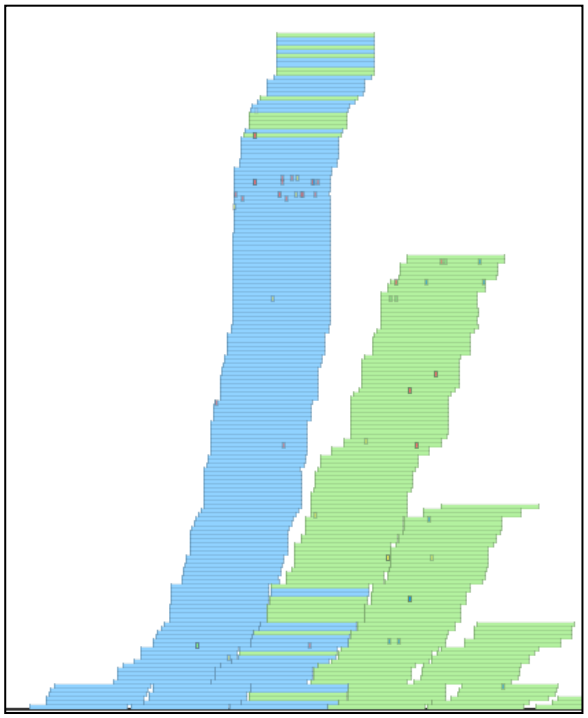
- VCFファイルに含まれるアリル情報をプロットし、またコンソールにてアリル頻度やリード深度などの詳細情報を確認
- 複数のVCFファイルを並べてプロットし、サンプル間の比較を行うことも可能

BAMファイルプロット

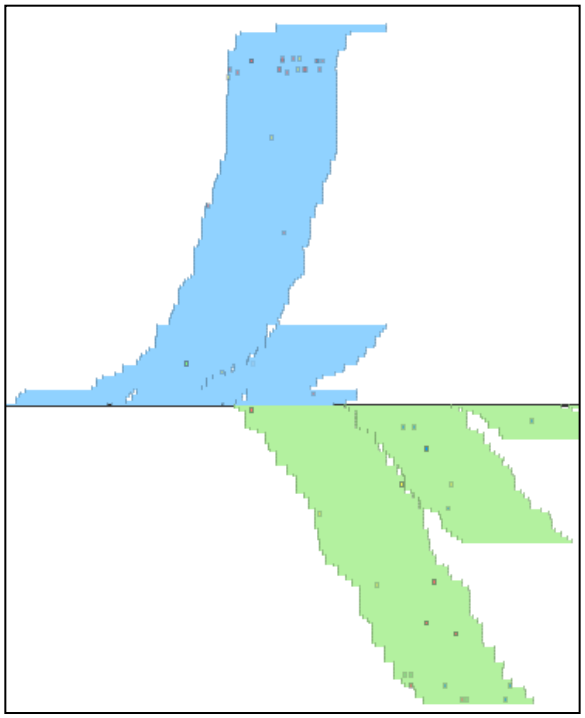


- BAMファイルプロットでは、アライメントされた全リード配列をそのまま表示するPile-Upグラフと、各アリルの頻度をまとめたCoverageグラフの2種類が表示される
- Coverageグラフをクリックすると、クリックしたゲノム上の位置にアライメントされた全リード配列のアリル別の集計データが表示される

BAMファイルプロット

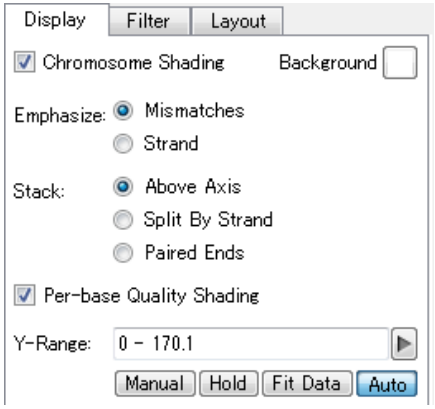


通常プロット



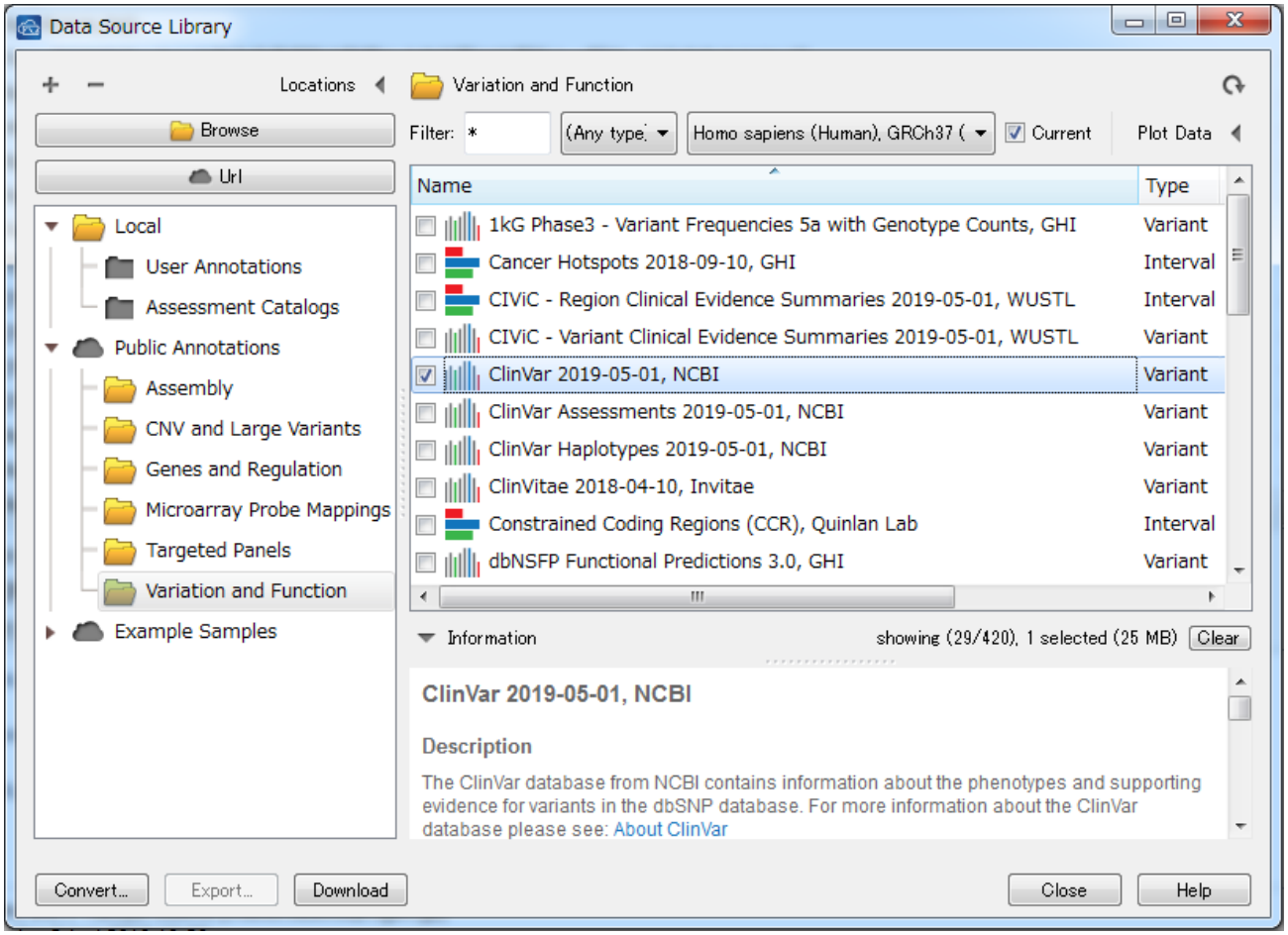
Split by Strandプロット

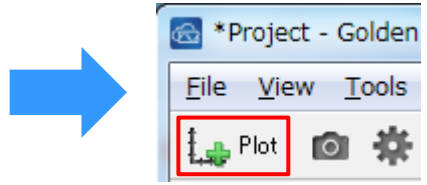
Paired Endsプロット



- Pile-Upグラフは、リード配列の表示形式を3タイプに切り替えることが可能
- 変異などの参照ゲノム配列とのミスマッチや、リード配列のクオリティーに基づいた表示を行うことで、変異やシークエンス結果のクオリティーチェックに活用できる

Data Source Library





ダウンロードしたデータベースを自由に選択し、ゲノムブラウザーにプロット

- 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、ゲノムブラウザーにプロット
- 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

アノテーションリソース

Filgen® biosciences & nanosciences

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

Genes
- RefSeq
- Ensembl

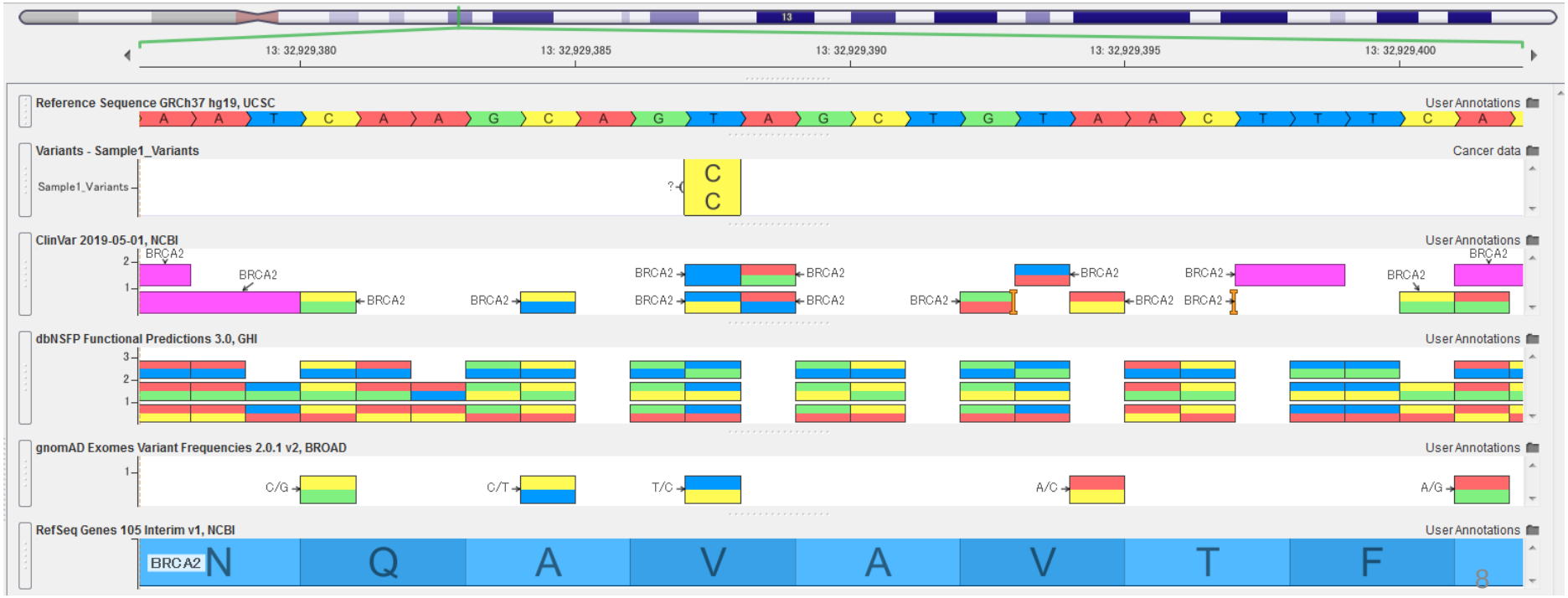
Clinical Annotations
- ClinVar
- GWAS Catalog

Cancer Annotations
- CIViC
- Cancer Hotspots
- ICGC Simple Somatic Mutations
- TCGA Variants
- MSK Impact

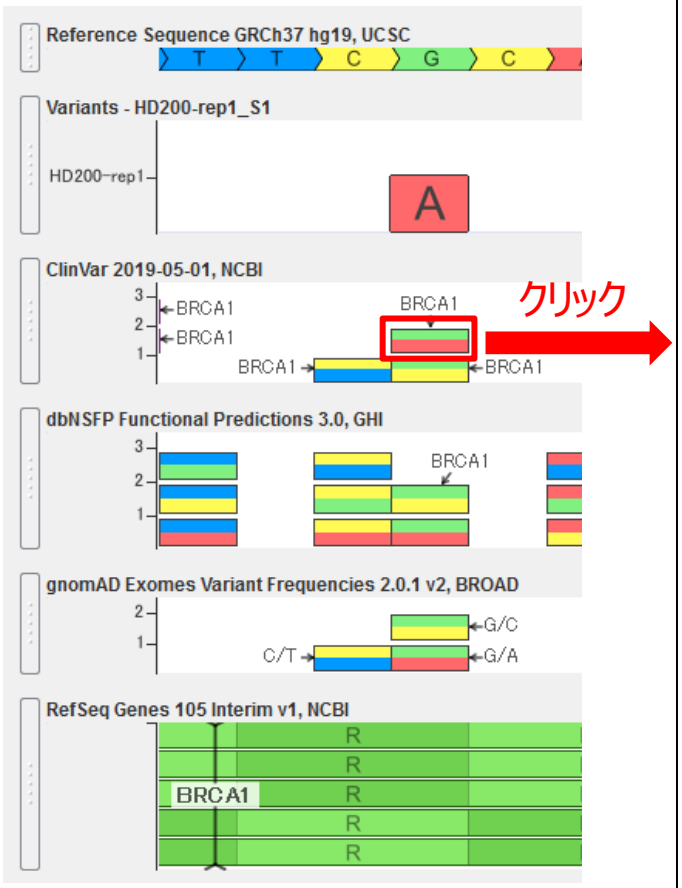
Functional Annotations
- dbNSFP (SIFT, PolyPhen, MutationTaster…)

Population Catalogs
- dbSNP
- 1000 Genome
- NHLBI 6500 Exomes
- ExAC Variant
- gnomAD Exomes
- TOPMed

Targeted Panels
- TruSight
- Ion AmpliSeq



アノテーションの詳細



クリック

Chr17: 41,234,451 - 41,234,451

NCBI, UCSC, Ensembl

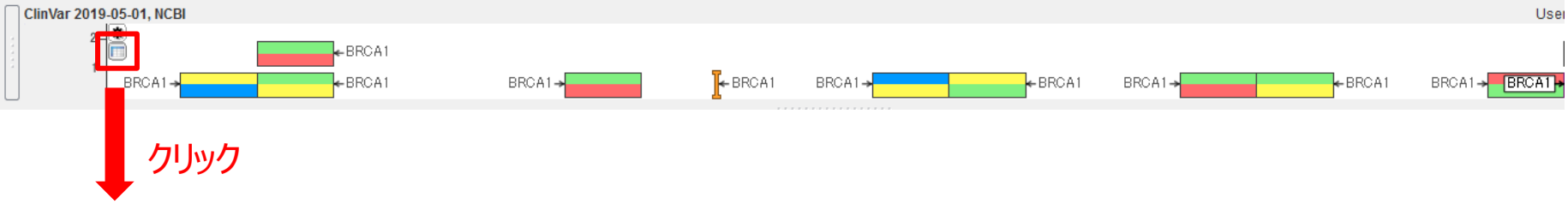
Normalized Alleles	G/A

Extended Fields

Ref/Alt	G/A
Variant ID	17675
Classification	Pathogenic
Clinical Significance	Pathogenic
Aggregate of Interpretations from Submissions	Pathogenic (32)
Review Status	(3 Stars) Reviewed By Expert Panel
RSID	rs41293455
Gene Names	BRCA1
Gene IDs	672
HGVS g. Name	NC_000017.10:g.41234451G>A
HGVS g. Name (GRCh38)	NC_000017.11:g.43082434G>A
HGVS c. Name	NM_007300.3:c.4327C>T
HGVS p. Name	NP_009231.2:p.Arg1443Ter
Per Condition Record Accessions	RCV000019244, RCV000048523, RCV000131880, RCV000159989, RCV000235131, RCV000239083, RCV000515235, RCV000735445, RCV000763000
Conditions	Breast-ovarian cancer, familial 1, Hereditary breast and ovarian cancer syndrome, Hereditary cancer-predisposing syndrome, Familial cancer of breast, not provided, not specified, Breast-ovarian cancer, familial 1;Familial cancer of breast;Pancreatic cancer 4, Breast and/or ovarian cancer, Breast-ovarian cancer, familial 1;FANCONI ANEMIA, COMPLEMENTATION GROUP S;Familial cancer of breast;Pancreatic cancer 4

- ゲノムブラウザー上でプロットをクリックすると、コンソールに各プロットの詳細が表示される

Feature List



Feature List

ClinVar 2019-05-01, NCBI — Read: All / Zoom

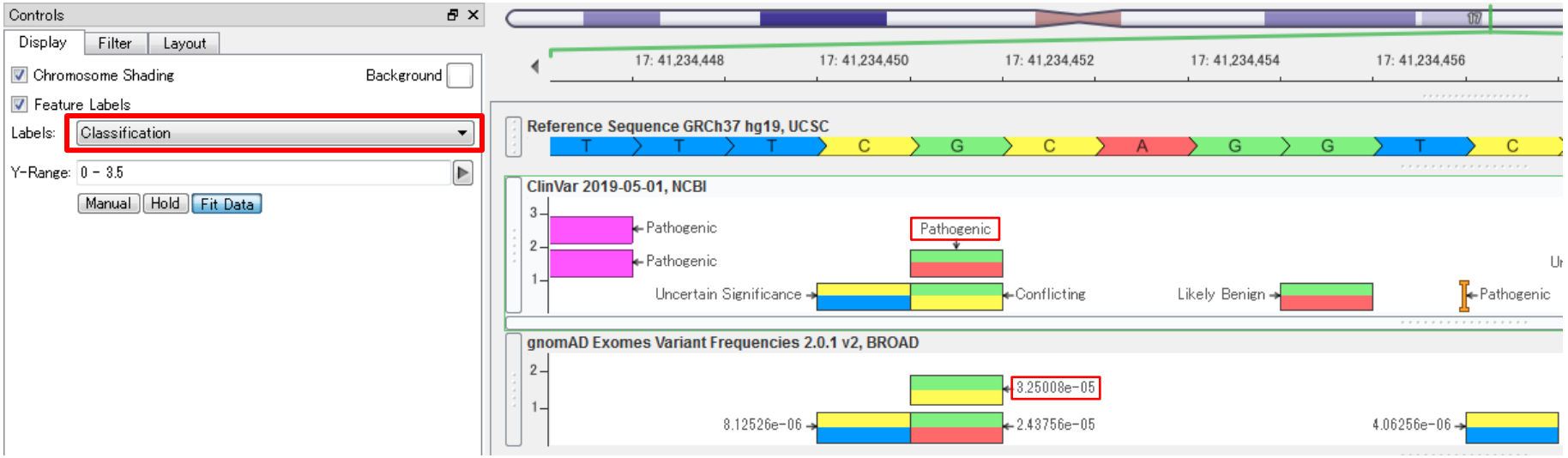
	Chr	Start	Stop	Ref/Alt	Variant ID	Classification	Clinical Significance	Aggregate of Interpretations from Submissions	Review Status	RSID	Gene Names	Gene IDs
1	17	41234450	41234450	C/T	55169	Uncertain Sig…	Uncertain Significa…	Uncertain significance (1)	(0 Stars) No …	rs4986849	BRCA1	672
2	17	41234451	41234451	G/C	17676	Conflicting	Conflicting Interpr…	Benign(1),Likely Benign(6),Pathogenic(1),Uncert…	(1 Star) Crite…	rs41293455	BRCA1	672
3	17	41234451	41234451	G/A	17675	Pathogenic	Pathogenic	Pathogenic (32)	(3 Stars) Rev…	rs41293455	BRCA1	672
4	17	41234455	41234455	G/A	388613	Likely Benign	Likely Benign	Likely benign (2)	(2 Stars) Crit…	rs1057523168	BRCA1	672
5	17	41234457	41234456	-/C	55168	Pathogenic	Pathogenic	Pathogenic (4)	(3 Stars) Rev…	rs80357748	BRCA1	672
6	17	41234459	41234459	T/C	185580	Uncertain Sig…	Uncertain Significa…	Uncertain significance (2)	(2 Stars) Crit…	rs786202288	BRCA1	672
7	17	41234460	41234460	C/G	433710	Uncertain Sig…	Uncertain Significa…	Uncertain significance (1)	(1 Stars) Crit…	rs1555584070	BRCA1	672
8	17	41234463	41234463	G/A	216668	Conflicting	Conflicting Interpr…	Likely Benign(1),Uncertain Significance(2)	(1 Star) Crite…	rs781260818	BRCA1	672
9	17	41234464	41234464	G/C	55166	Likely Benign	Likely Benign	Likely benign (6)	(3 Stars) Rev…	rs80356856	BRCA1	672
10	17	41234467	41234467	A/G	427282	Likely Benign	Likely Benign	Likely benign (1)	(3 Stars) Rev…	rs1131692092	BRCA1	672
11	17	41234468	41234468	G/-	431261	Pathogenic	Pathogenic	Pathogenic (2)	(3 Stars) Rev…	rs1135401874	BRCA1	672
12	17	41234468	41234468	G/A	531421	Uncertain Sig…	Uncertain Significa…	Uncertain significance (1)	(1 Stars) Crit…	rs1555584083	BRCA1	672
13	17	41234469	41234469	A/-	266468	Pathogenic	Pathogenic	Pathogenic (2)	(3 Stars) Rev…	rs886040223	BRCA1	672

13 features read (complete)

(Chr17: 41,234,449 – 41,234,468)

- 複数のプロットの詳細をまとめて確認したい場合は、各リソースのFeature Listアイコンをクリックすると、テーブル形式で表示が可能
- テーブル上の任意の行をクリックすると、自動的にゲノムブラウザー上の該当位置にジャンプする

ラベル表示の変更



- 各リソースのプロットラベルは、コンソールより表示フィールドの選択が可能
- その他、プロットするデータの選択フィルターや、レイアウトの編集など、ゲノムブラウザー上での表示設定を、任意に変更することが可能

デモンストレーション

解析サンプル：腫瘍細胞と正常細胞のペアサンプル

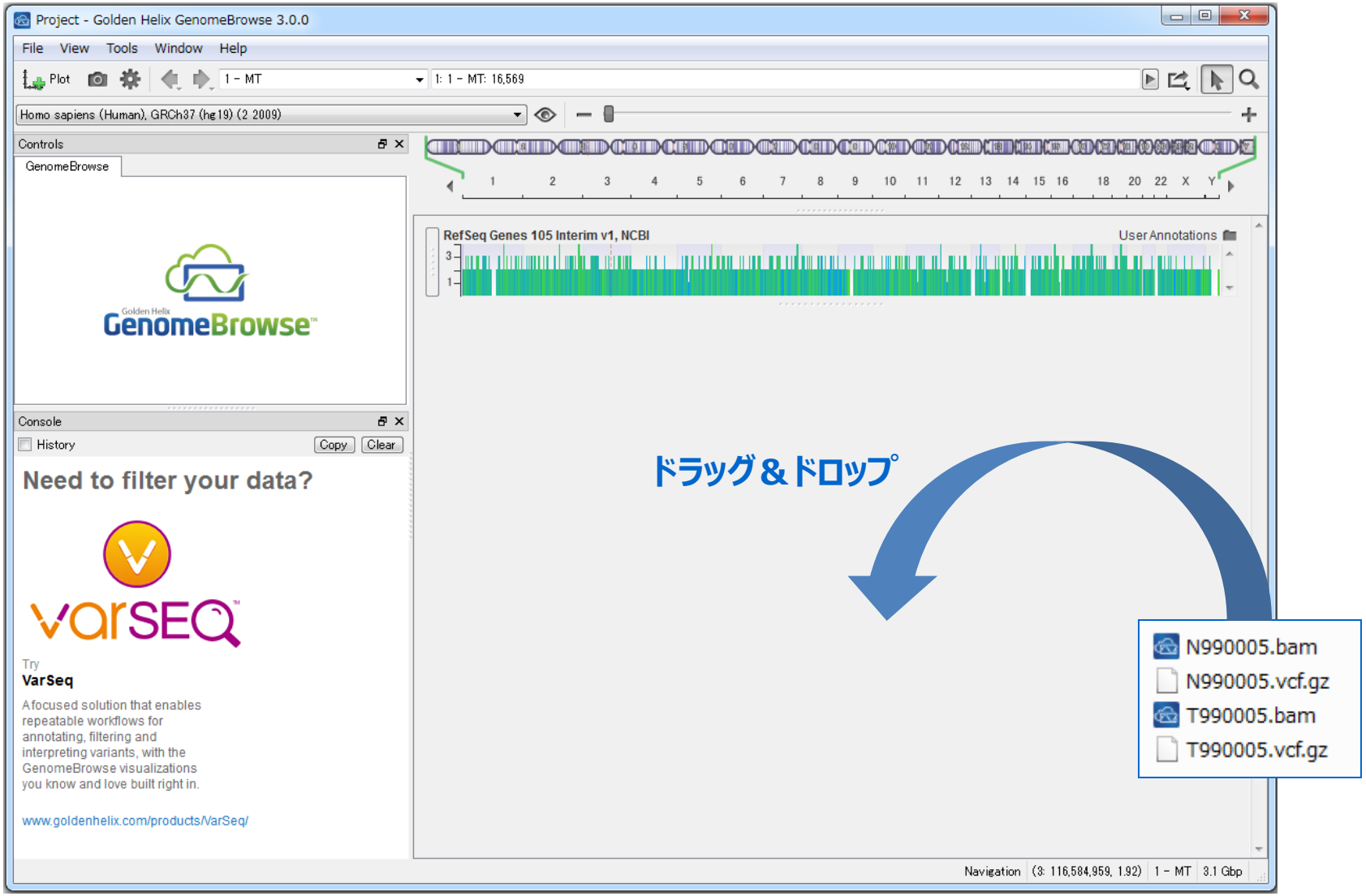
解析に用いるデータファイル：各サンプルのVCFファイル、BAMファイル

解析の目的：がん遺伝子における体細胞変異の抽出と、臨床アノテーションの確認

胃腺がん患者の腫瘍細胞サンプルと正常細胞サンプルのゲノムシークエンスより取得したVCFファイルとBAMファイルを使用し、特定のがん遺伝子の体細胞変異の抽出とクオリティチェックを行い、さらにデータベースに照合することで、病原性の有無や抗がん剤情報といった、臨床アノテーションの確認を行う。

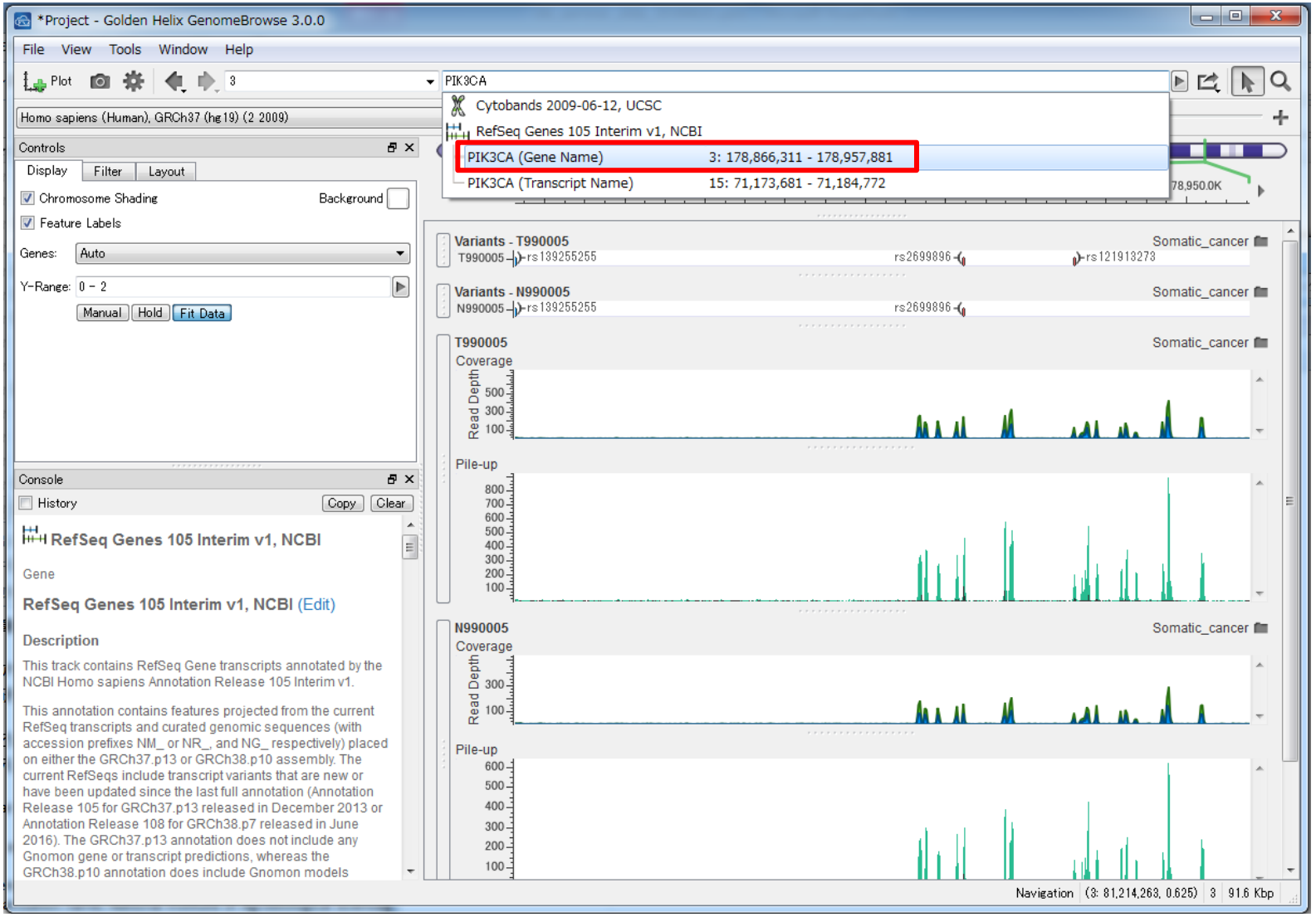
デモンストレーション

手順1：VCF, BAMファイルの取り込み



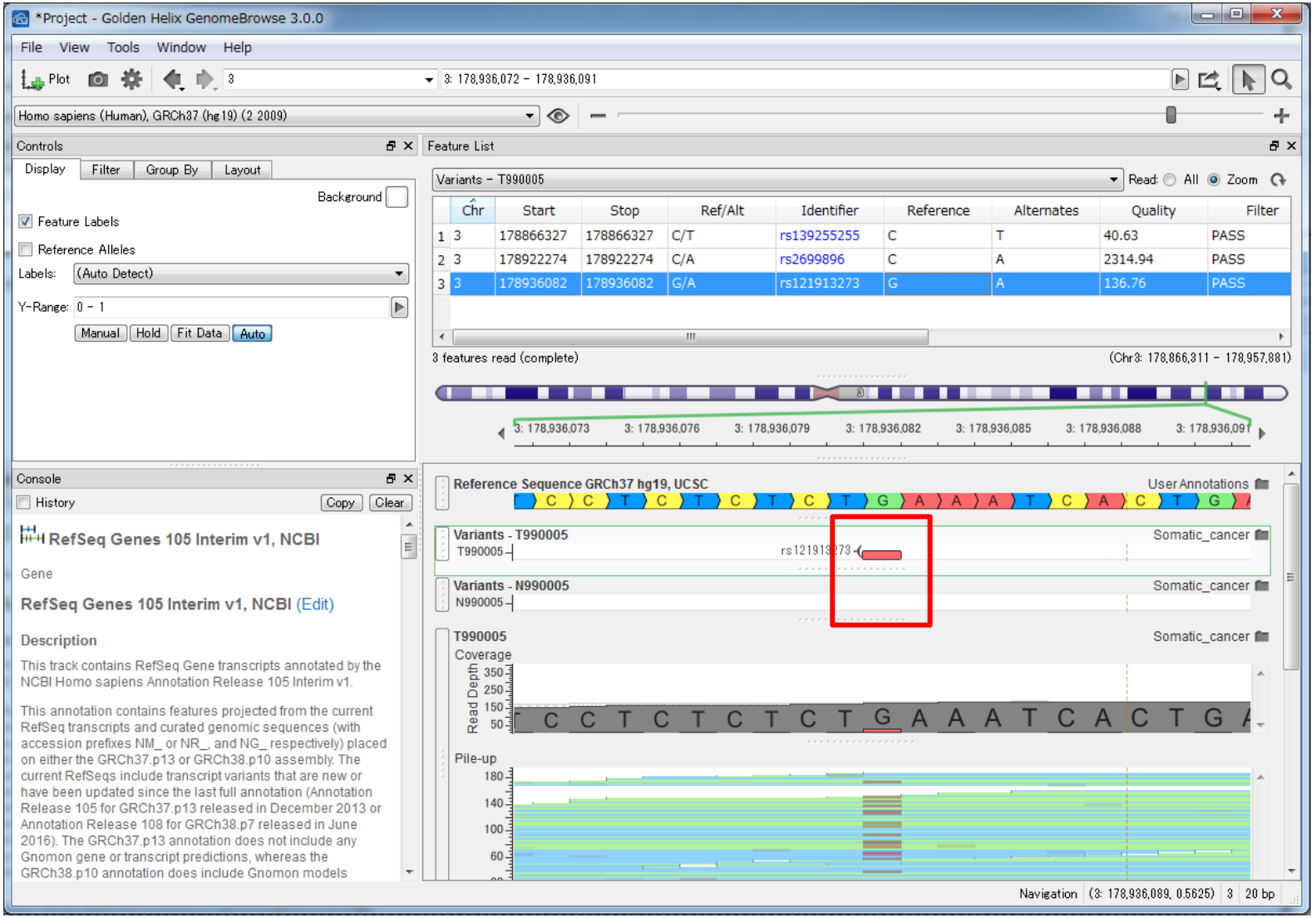
デモンストレーション

手順2：ターゲットとする遺伝子名の検索



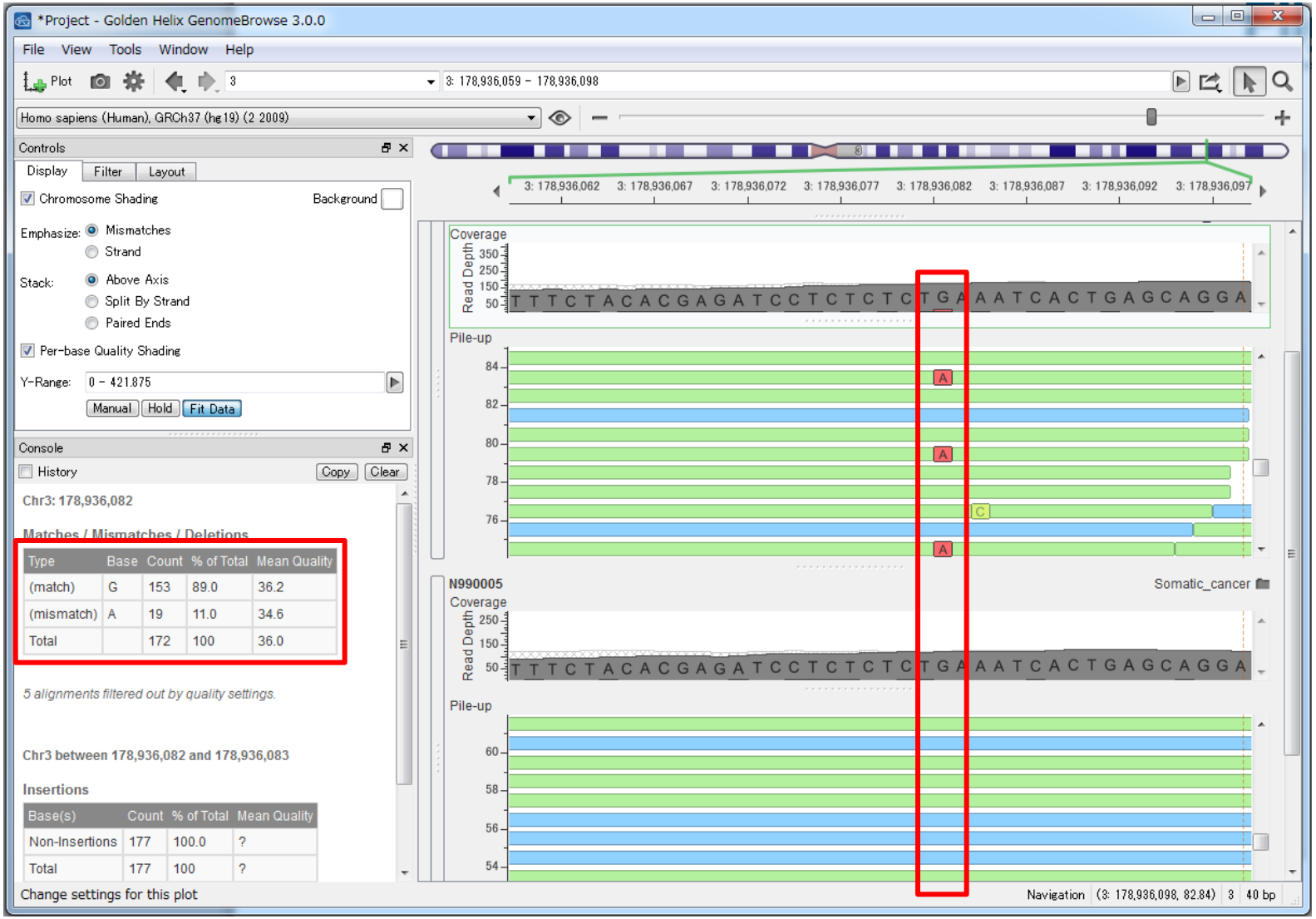
デモンストレーション

手順3：腫瘍サンプルと正常サンプルの比較による、体細胞変異の探索



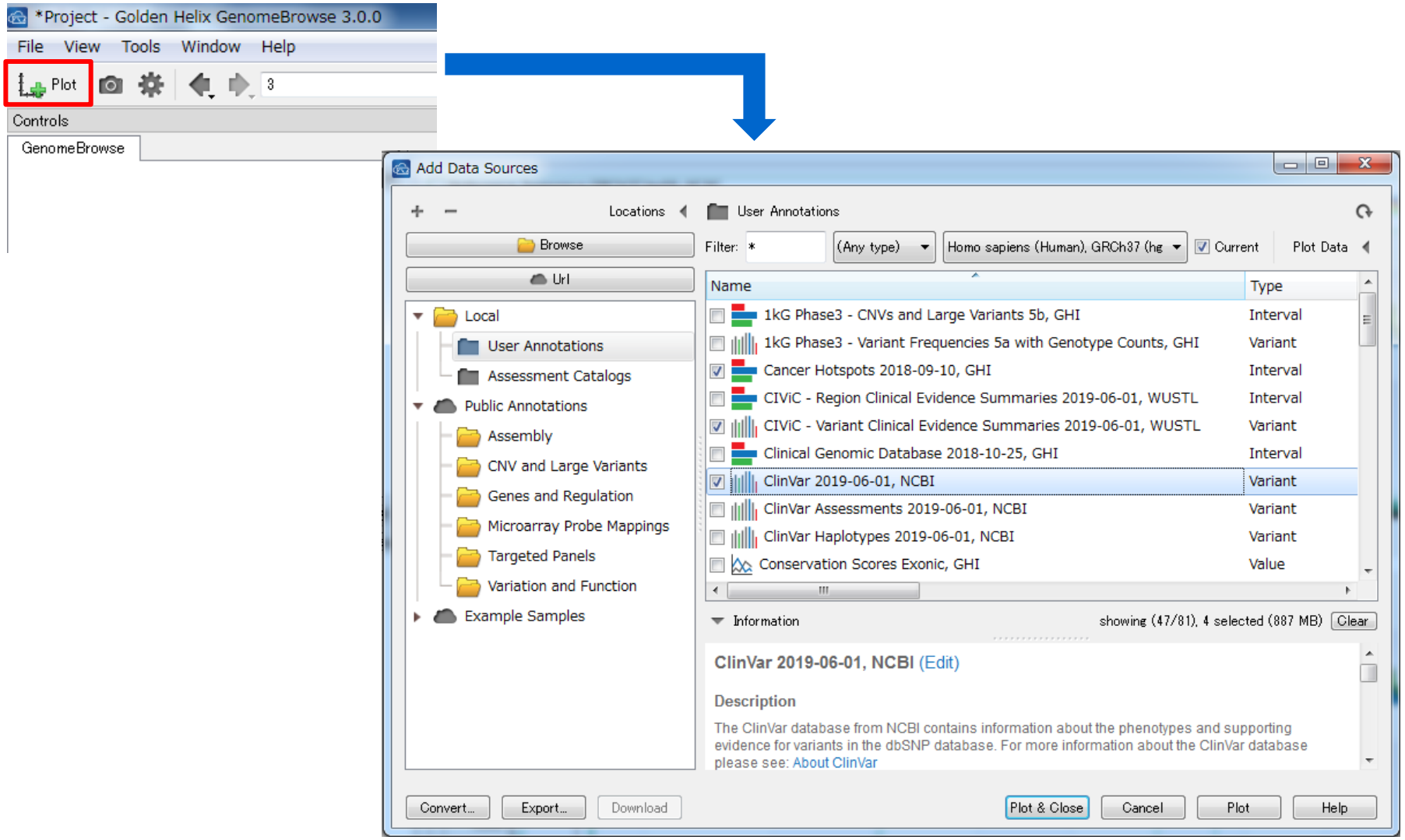
デモンストレーション

手順4：BAMファイルプロットによる、検出変異のクオリティー確認



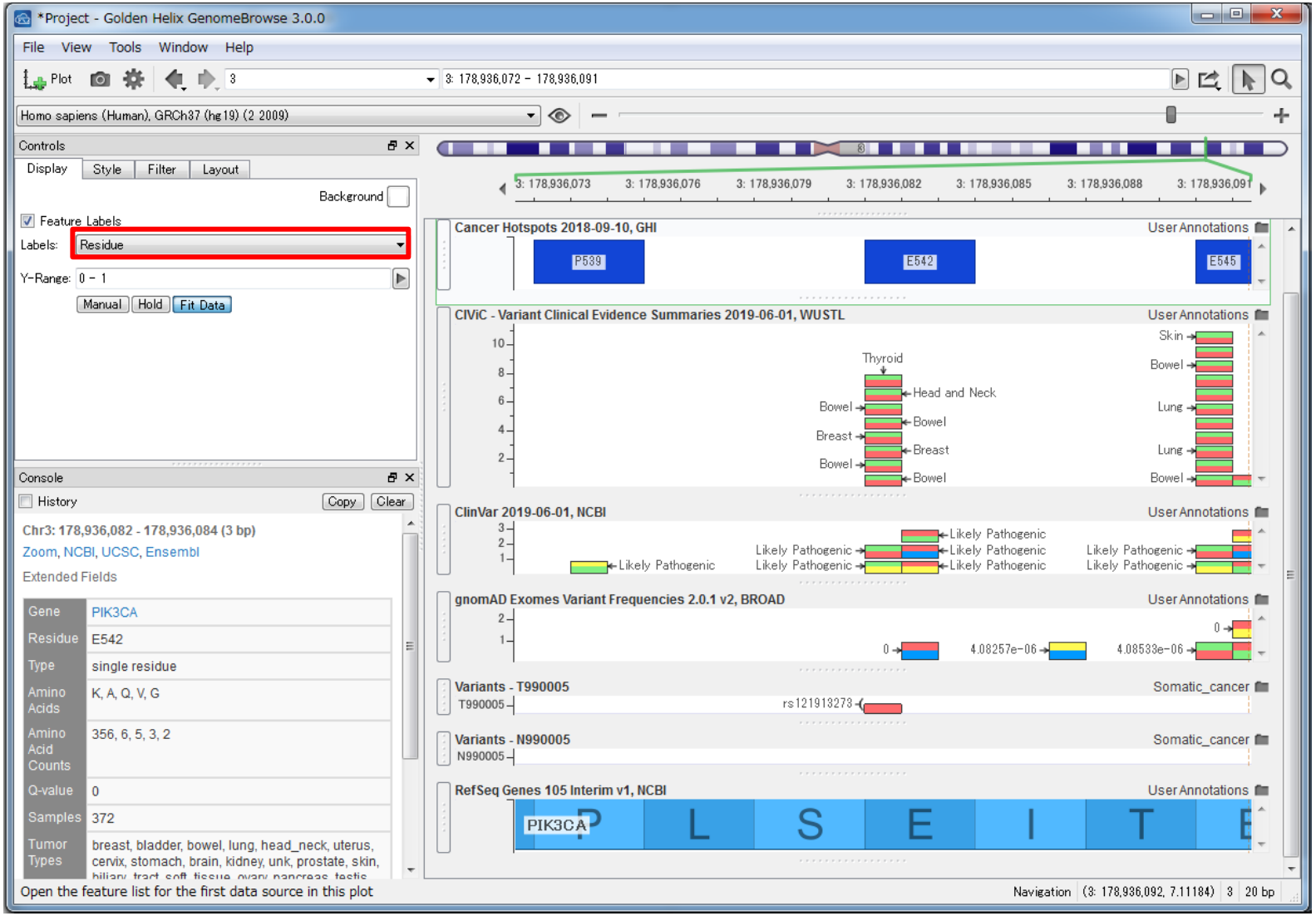
デモンストレーション

手順5：アノテーションリソースのプロット



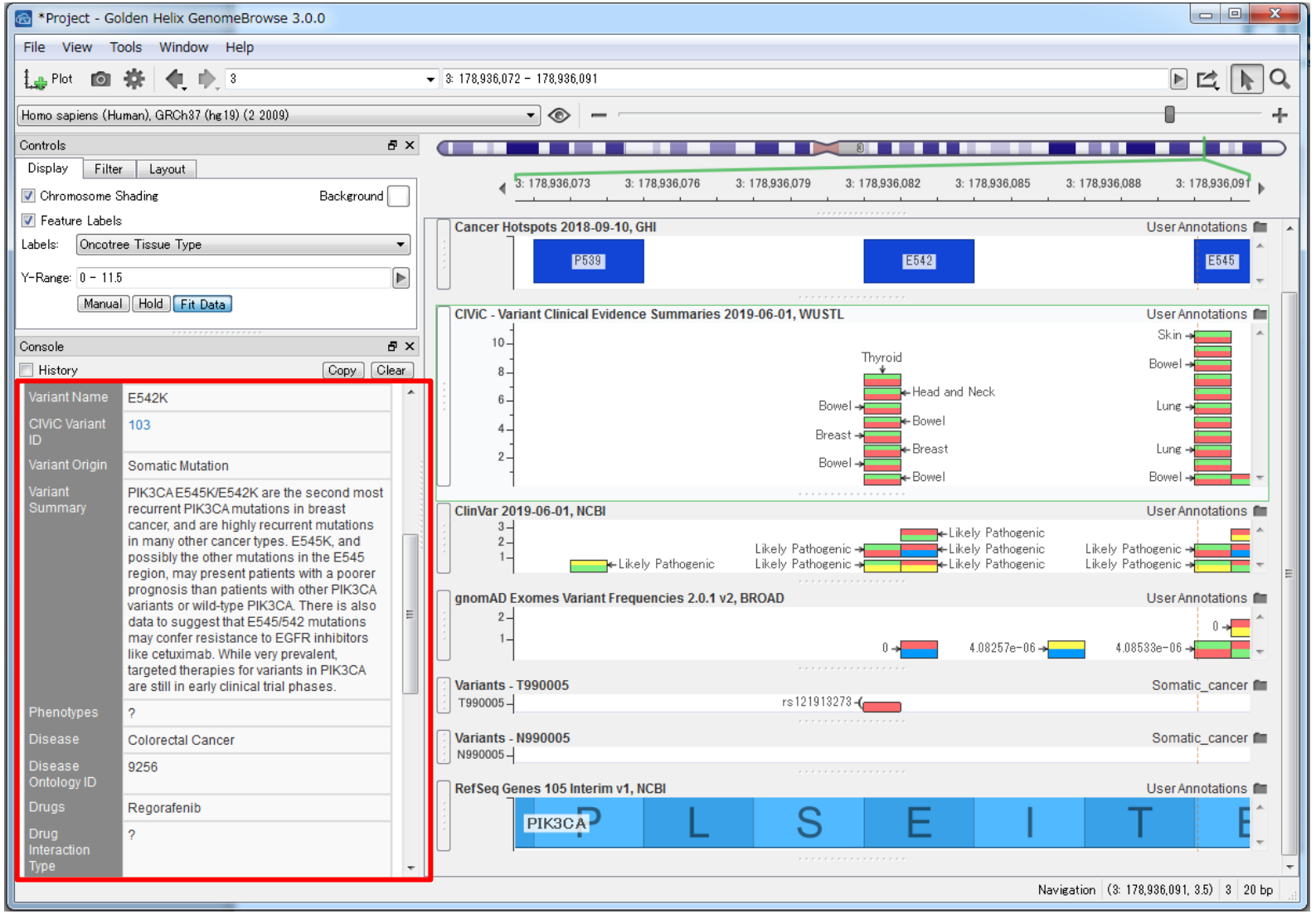
デモンストレーション

手順6：アノテーションリソースのプロットラベルを変更



デモンストレーション

手順7：体細胞変異の臨床アノテーションを確認



GenomeBrowse®ダウンロード：

http://goldenhelix.com/products/GenomeBrowse/index.html#download

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