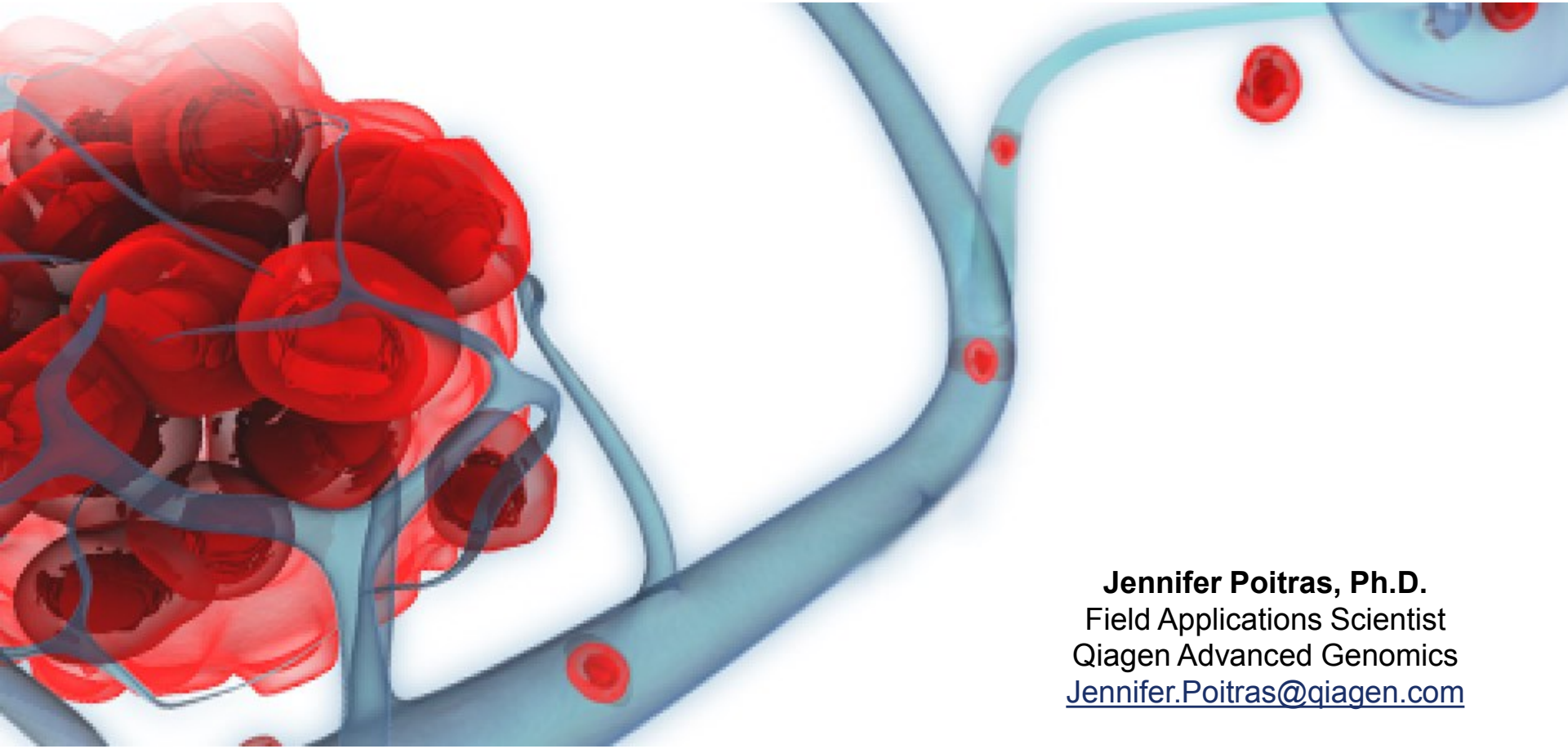
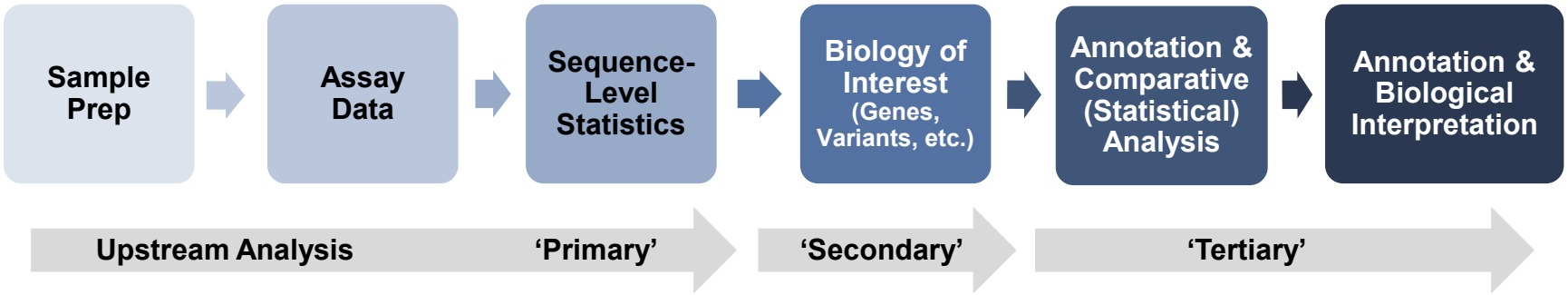


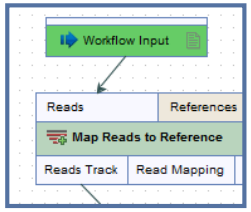
Jennifer Poitras, Ph.D.
Field Applications Scientist
Qiagen Advanced Genomics
Jennifer.Poitras@qiagen.com

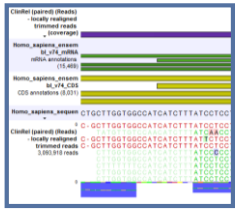
CLC Biomedical Genomics Workbench General Overview

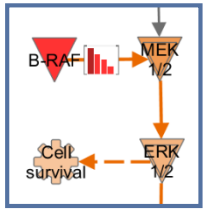
QIAGEN Sample to Insight

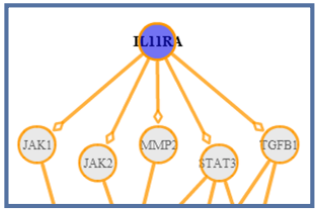


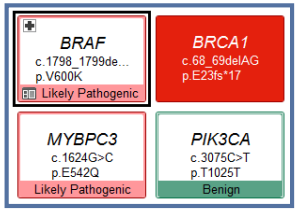
Sample











Insight

QIAGEN

clcbio A QIAGEN Company

Gx Bx

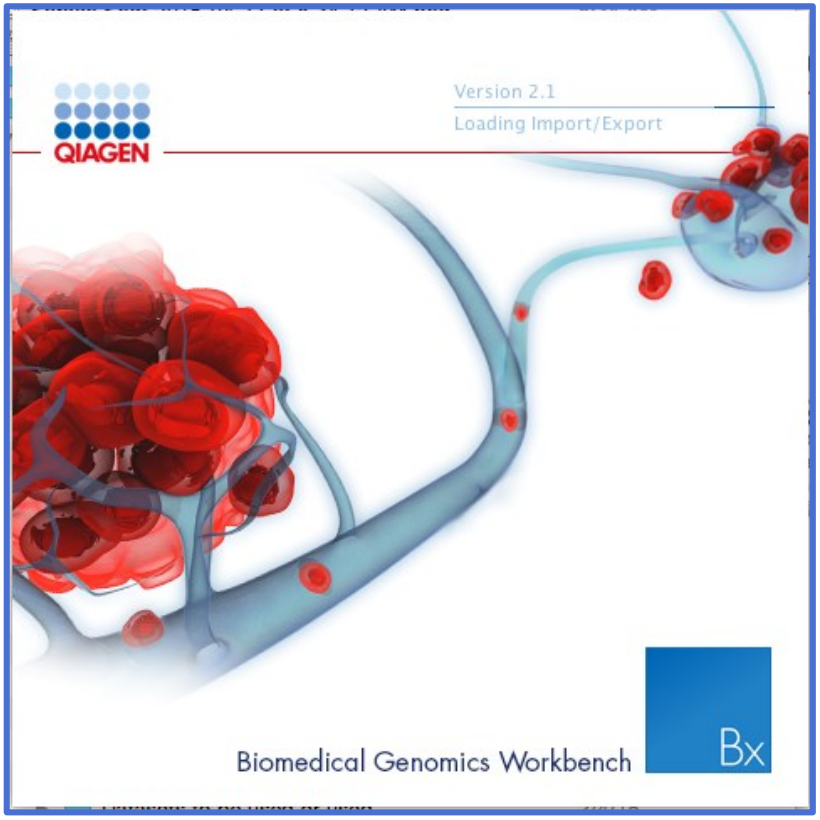
INGENUITY PATHWAY ANALYSIS

INGENUITY VARIANT ANALYSIS

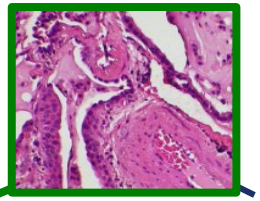
QIAGEN Clinical Insight

BIOBASE BIOLOGICAL DATABASES

Fast and Easy Analysis



- **Accurate and trustworthy results**
  - ✓ Whole Genome Sequencing
  - ✓ Whole Exome Sequencing
  - ✓ Targeted or Whole Transcriptome Sequencing
  - ✓ ChIP-Seq data
- **Intuitive and easy-in-use**
  - ✓ Comprehensive end-to-end analysis workflows for single samples or cohort studies
  - ✓ One-click analysis of QIAGEN GeneRead DNASeq Amplicon Panels
  - ✓ Streamlined integration with Ingenuity Pathway Analysis (IPA) & Ingenuity Variant Analysis
- **Flexible & customizable**
  - ✓ Ready-to-use workflows can be customized
  - ✓ Build your own workflows



RNA

DNA

Sequencing

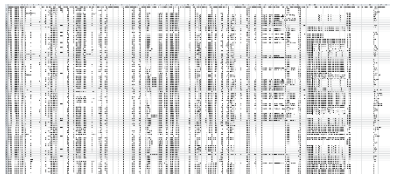
Bx

.fastq
and many other file types

Data Analysis

RNA-seq

Variants



VCF file

Interpretation

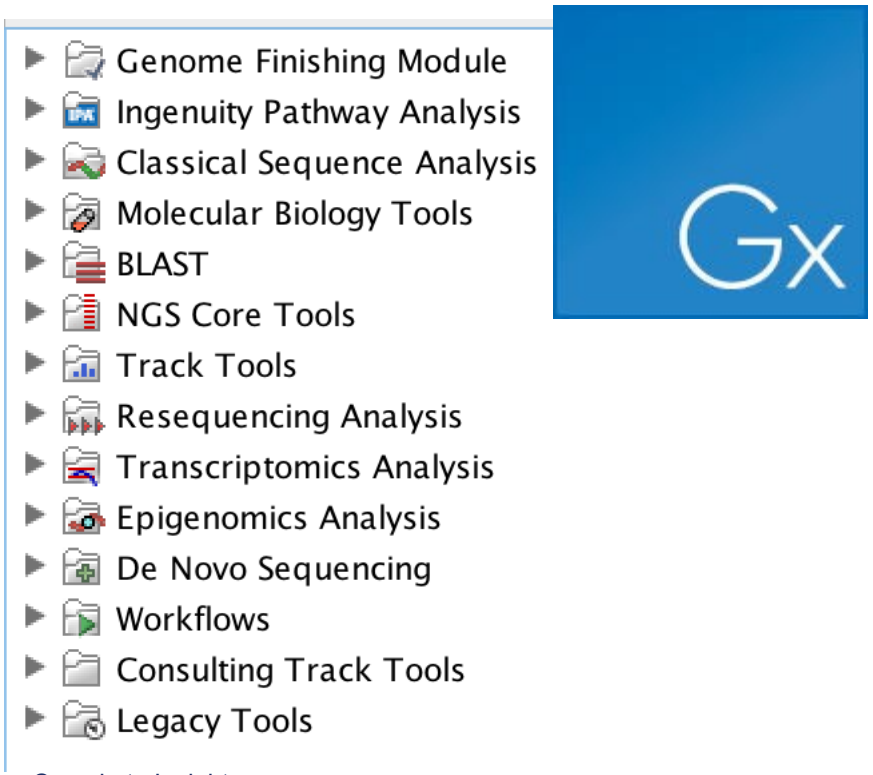
INGENUITY PATHWAY ANALYSIS

INGENUITY VARIANT ANALYSIS

CLC Genomics Workbench vs. Biomedical Workbench

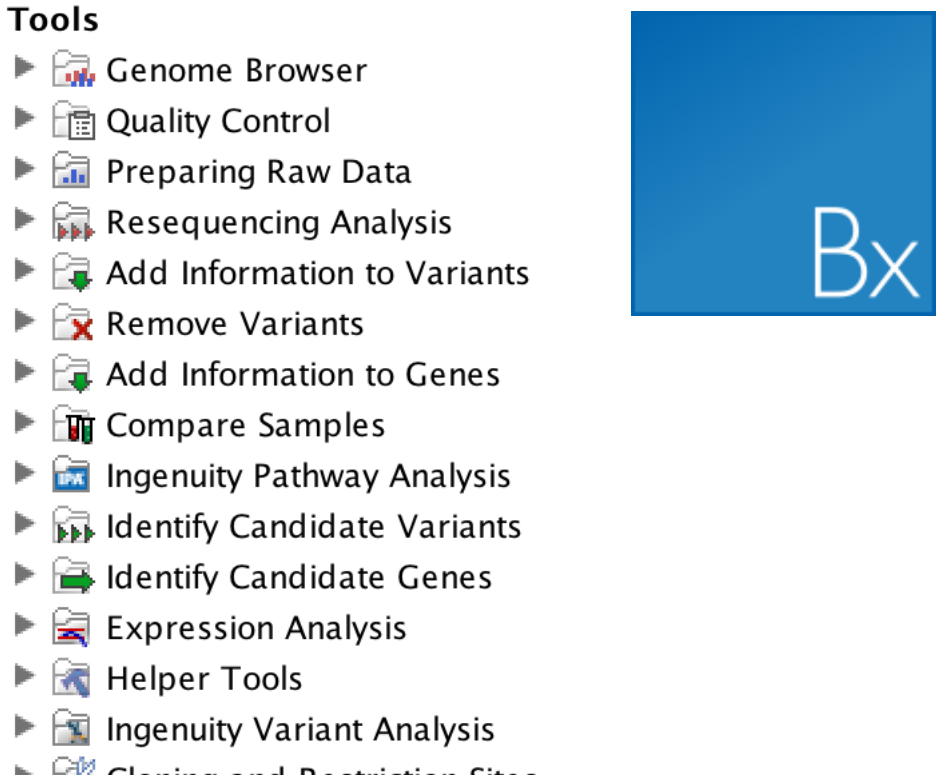
Genomics Workbench

- Analysis of NGS data on ALL organisms
- Reference data management for model organisms: *C. elegans, S. cervisiae,* etc.
- Supports microbial and non-mammalian NGS analysis
- Supports de novo assembly



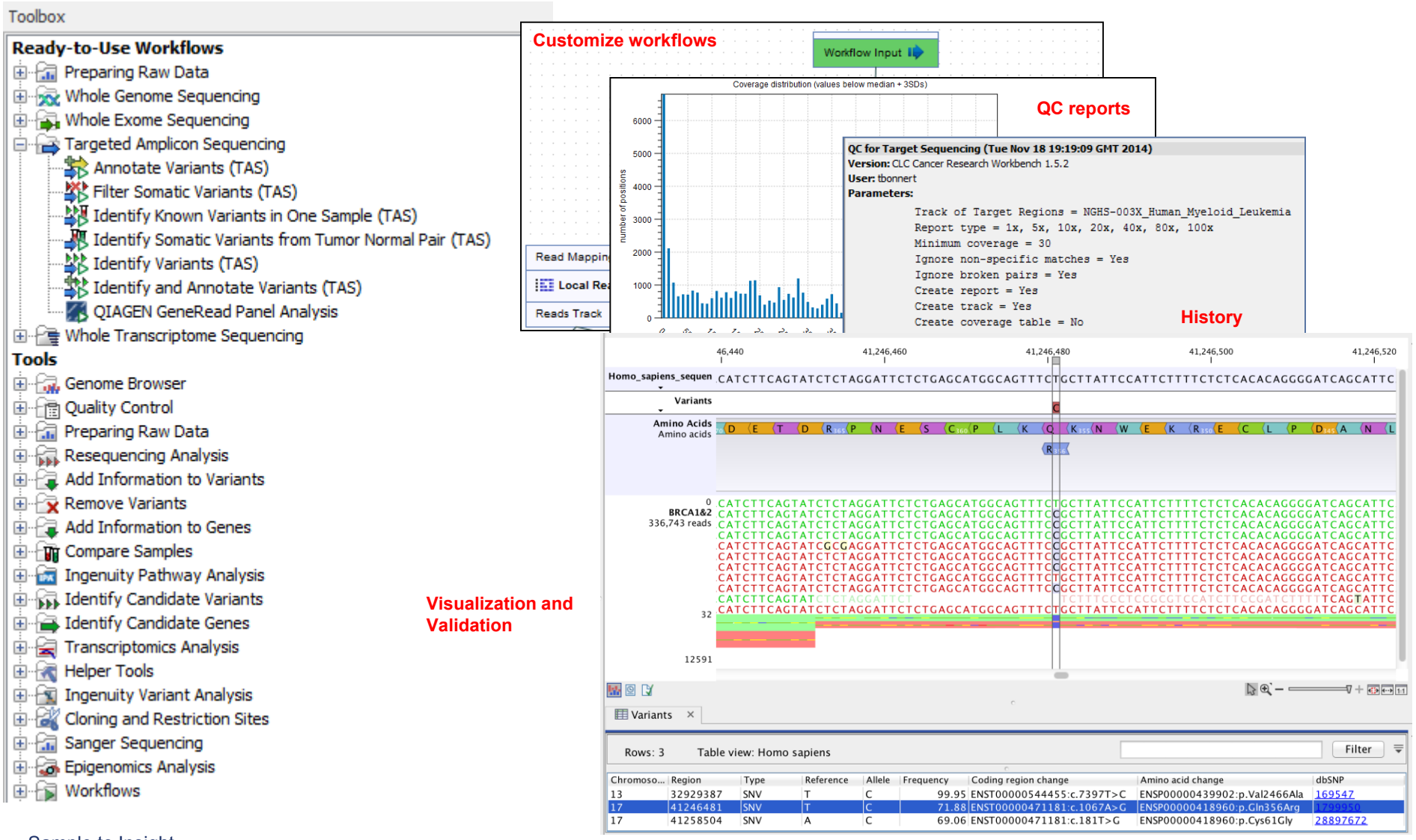
Biomedical Workbench

- Tools most relevant to clinical applications on human data
- Automation and Reference Management geared towards for Human research
  - Human, Mouse,Rat
- Ready-to-use work flows



Biomedical Genomics Workbench – Complex Tasks, Simply Done

Streamlined workflows and a rich toolbox to efficiently process data



In CLC Workbenches you can..

1. QC and preprocess NGS data (RNA-Seq, smRNA, and DNAseq reads)

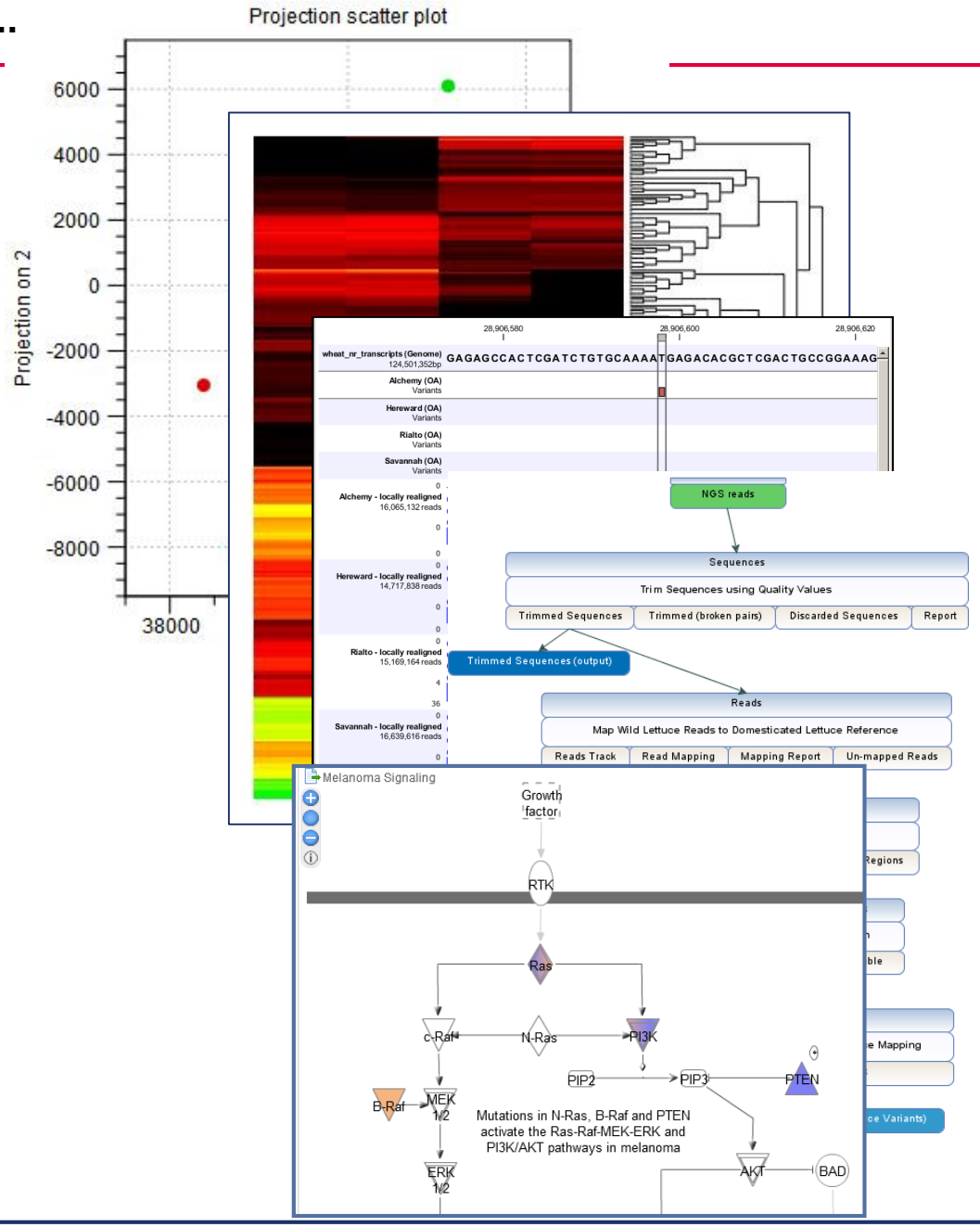
1. Differential expression and statistical analysis for RNA-Seq and smRNA

2. Generate, annotate, and compare high-confidence variant calls

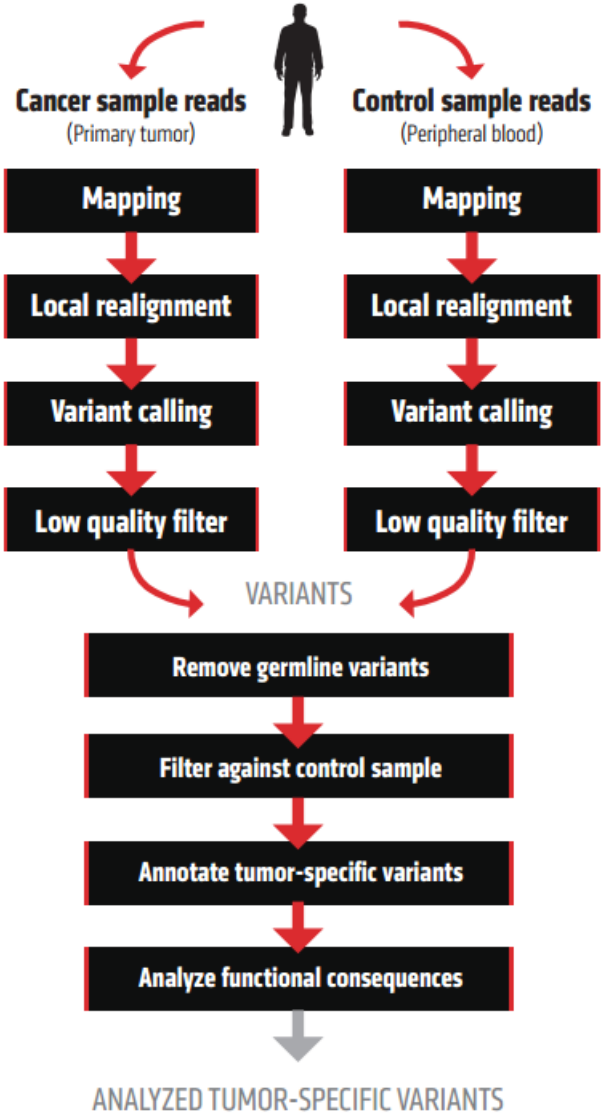
1. Facilitate analysis with interactive visualization

2. Construct automated workflows in user friendly interface

1. Modules [**plugins**] available for additional or custom functionality



Cancer research applications (WGS/WES/RNAseq)



- Matched Tumor/Normal
- Tumor alone
- Tumor vs controls

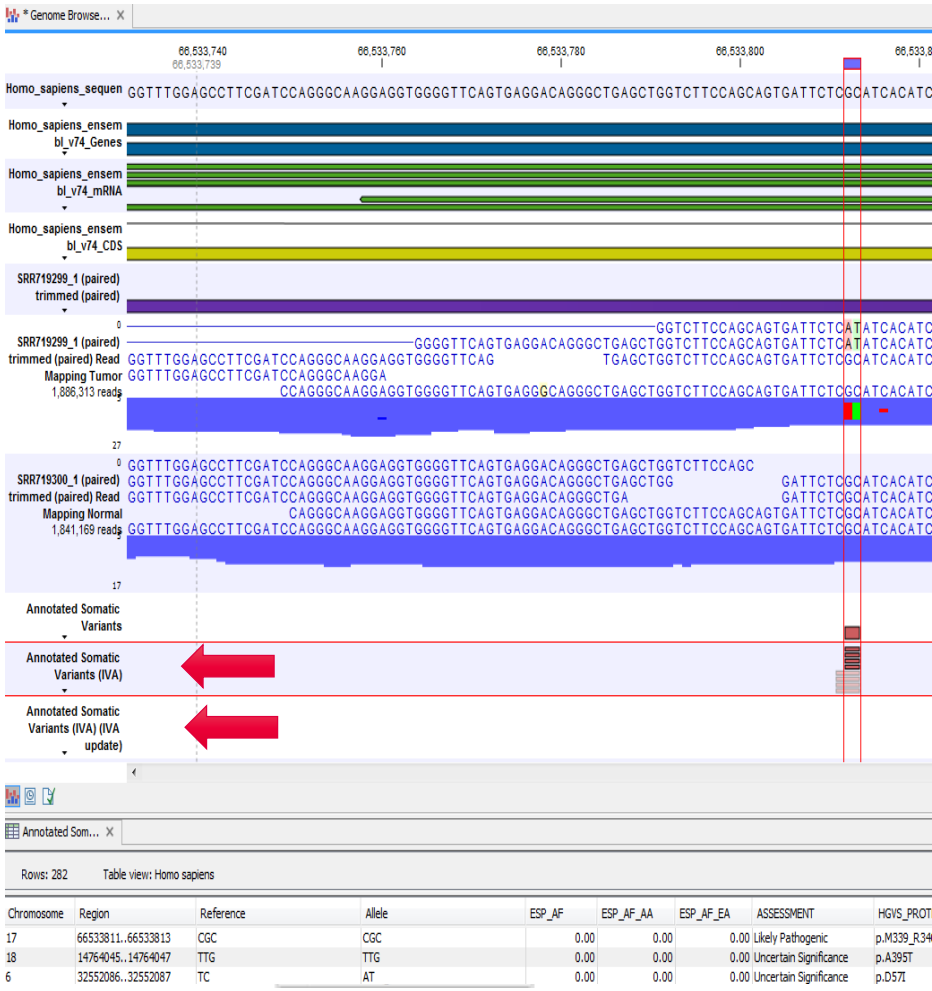
QIAGEN Bioinformatics Products Streamline Integration

Ingenuity Variant Analysis integration: two ways!

Biomedical Genomics Workbench & Server

INGENUITY VARIANT ANALYSIS

- Ready to use workflows + Ingenuity Variants Analysis plugin
- Make it one click by leveraging workflow customization!
- Click on the link to see results in IVA
- Edit filtering cascade in IVA and send the result back in BxWB
- Use tracks to compare results from different filter cascades in BxWB



Biomedical Genomics Workbench – Complex Tasks, Simply Done

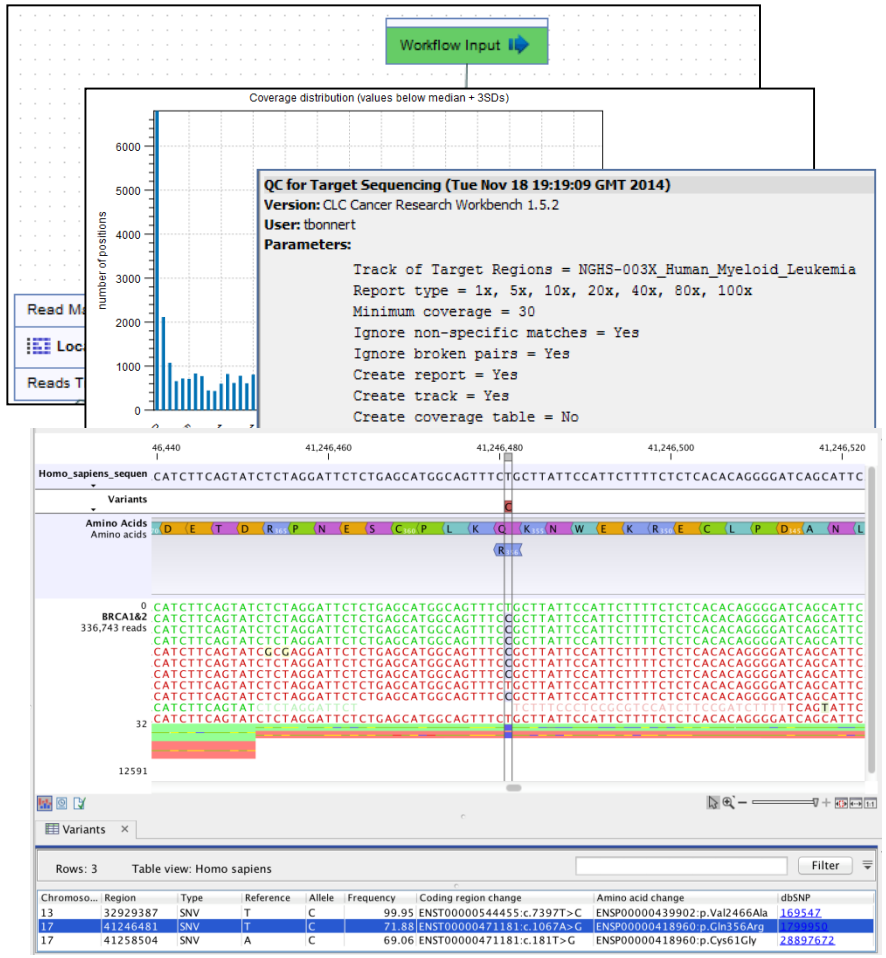
Streamlined workflows and a rich toolbox to efficiently process data

Ready-to-Use Workflows

- Preparing Raw Data
- Whole Genome Sequencing
- Whole Exome Sequencing
  - General Workflows (WES)
  - Somatic Cancer (WES)
  - Hereditary Disease (WES)
    - Filter Causal Variants (WES-HD)
    - Identify Causal Inherited Variants in Family of Four (WES)
    - Identify Causal Inherited Variants in Trio (WES)
    - Identify Rare Disease Causing Mutations in Family of Four (WES)
    - Identify Rare Disease Causing Mutations in Trio (WES)
    - Identify Variants (WES-HD)
    - Identify and Annotate Variants (WES-HD)
    - Identify and Interpret Causal Variants in Family of Four using IVA (WES)
    - Identify and Interpret Causal Variants in Trio using IVA (WES)
- Targeted Amplicon Sequencing
- Whole Transcriptome Sequencing

Tools

- Genome Browser
- Quality Control
- Preparing Raw Data
- Resequencing Analysis
- Add Information to Variants
- Remove Variants
- Add Information to Genes
- Compare Samples
- Identify Candidate Variants
- Identify Candidate Genes
- Expression Analysis
- Helper Tools
- Ingenuity Variant Analysis
- Cloning and Restriction Sites
- Sanger Sequencing
- Epigenomics Analysis



Customize workflows

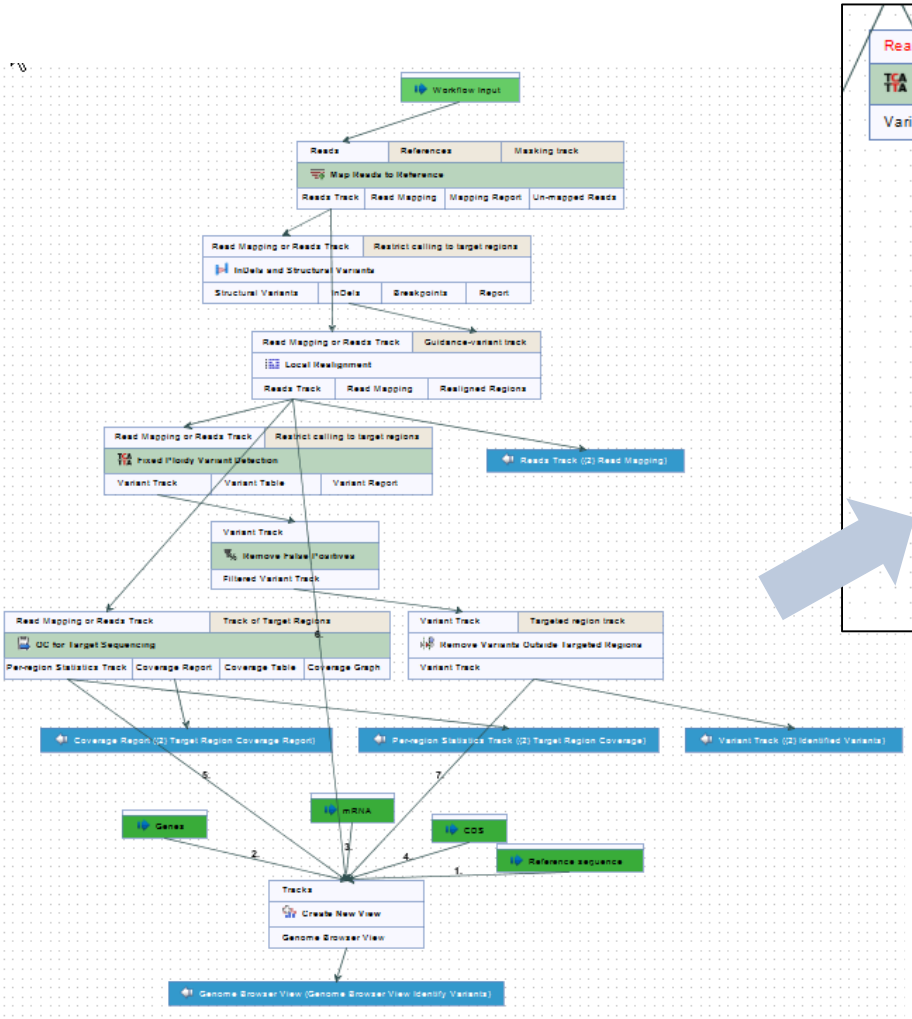
QC reports

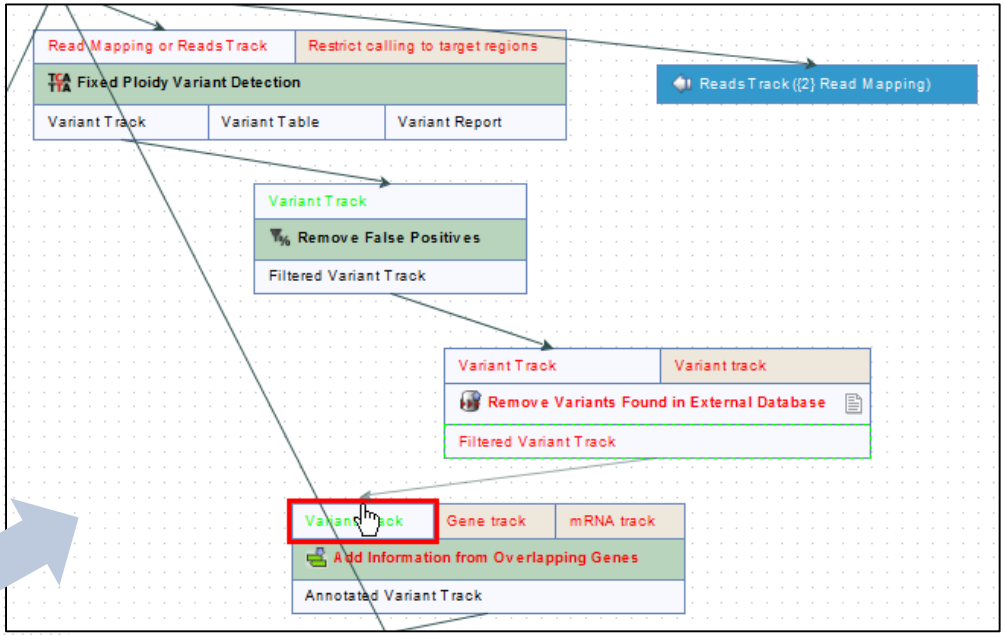
History

Visualization and Validation

Ready to Use Workflows

Very easy customization of ready-to-use workflows

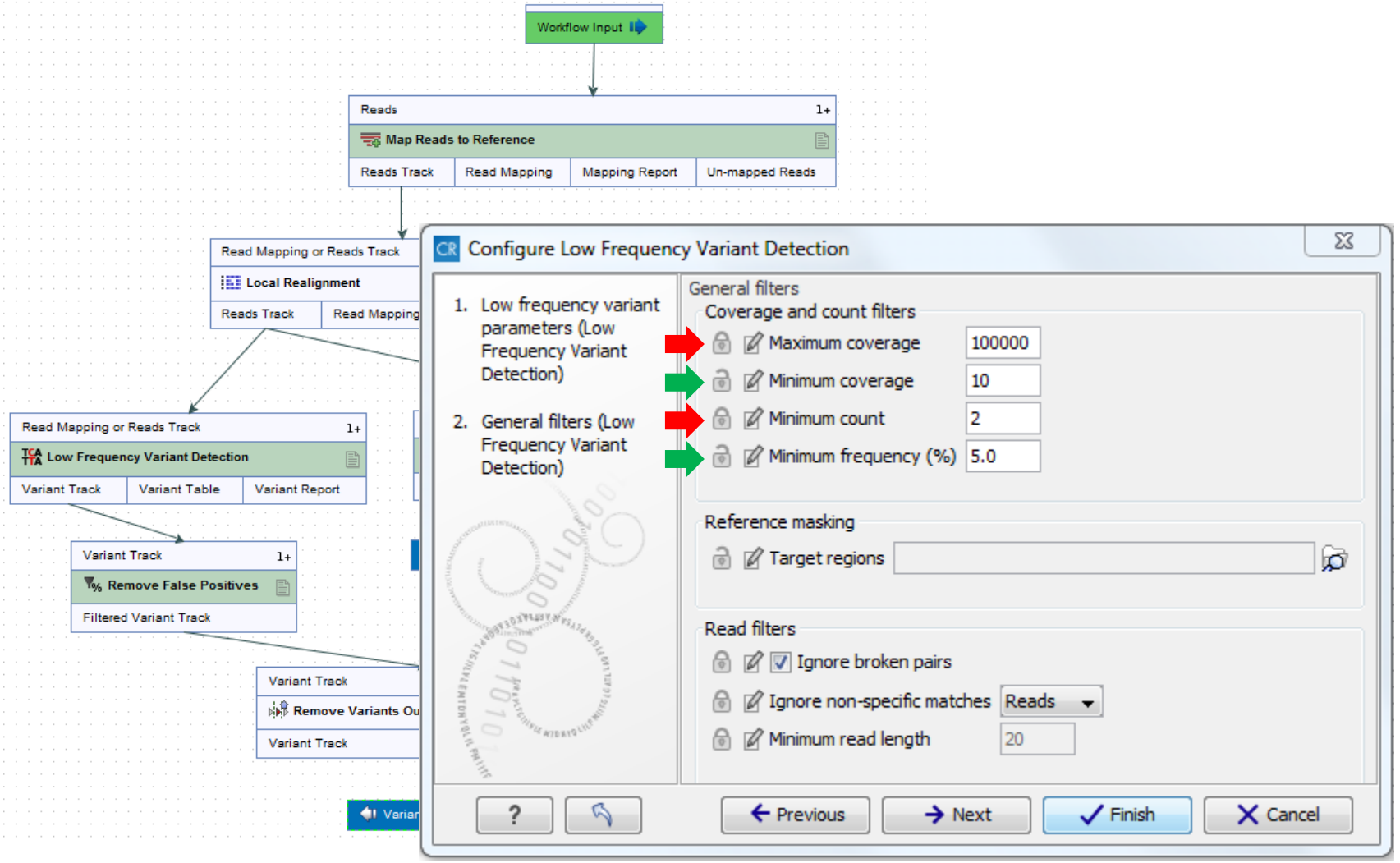




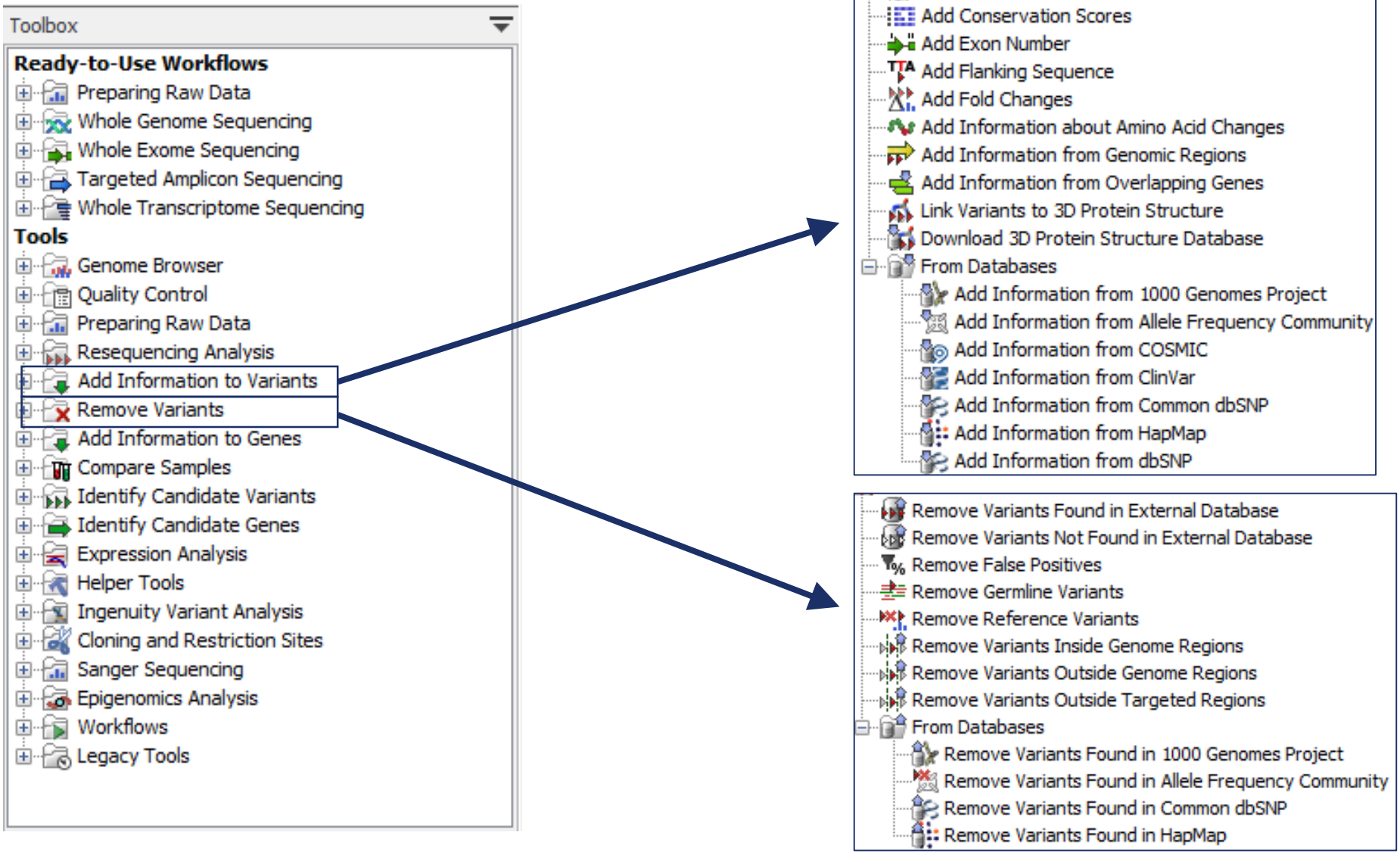
- Add new tool with drag & drop
- Remove tools
- Add new outputs (e.g. all large deletions)
- Change parameters and lock them down

Flexible workflows tailored for end-users

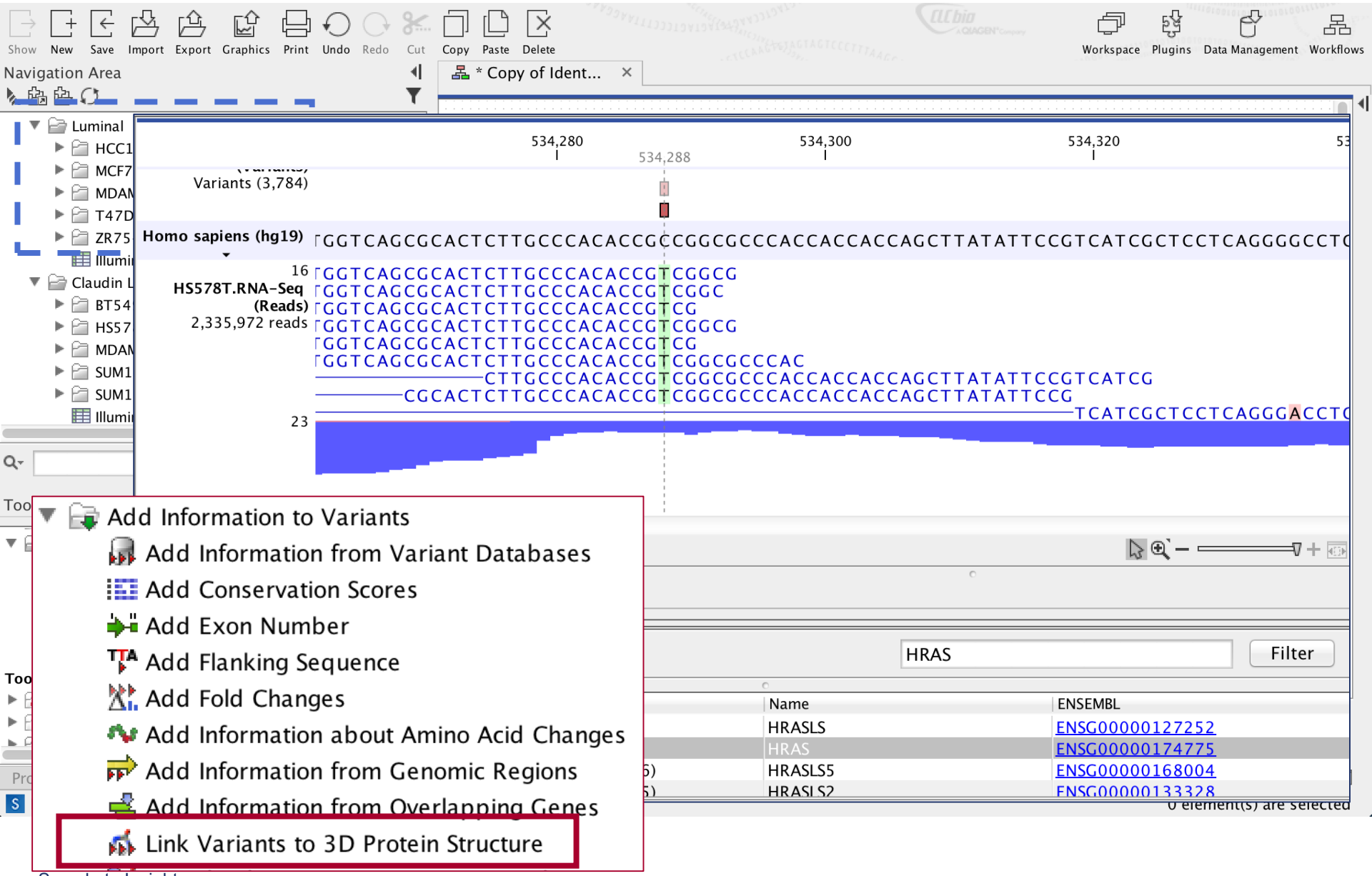
Lock key parameters of workflow to standardise processing



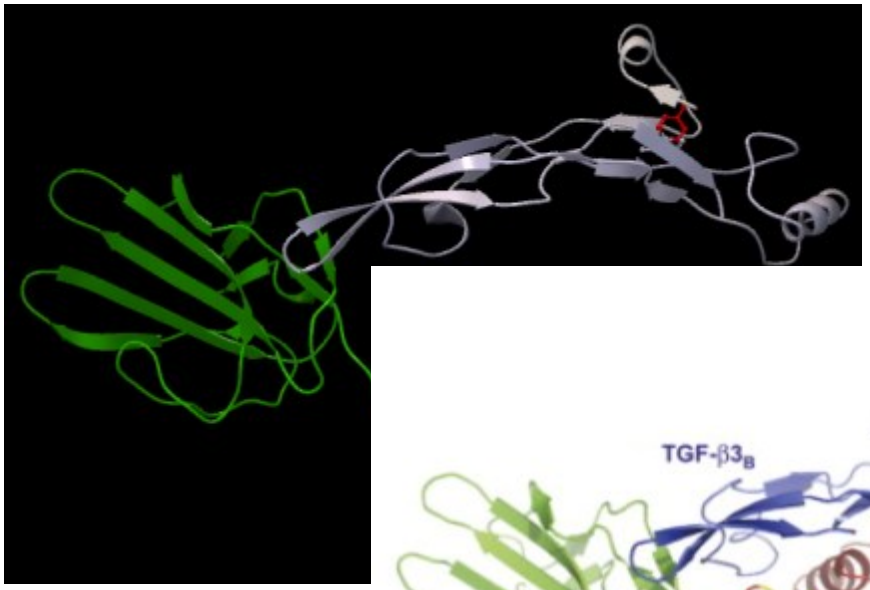
Powerful, task-driven, user-friendly tools

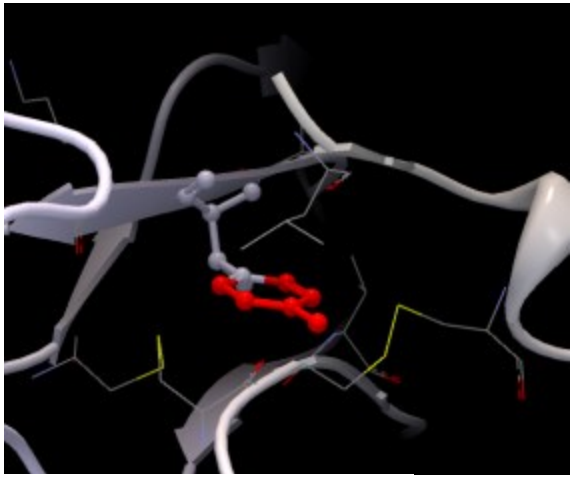


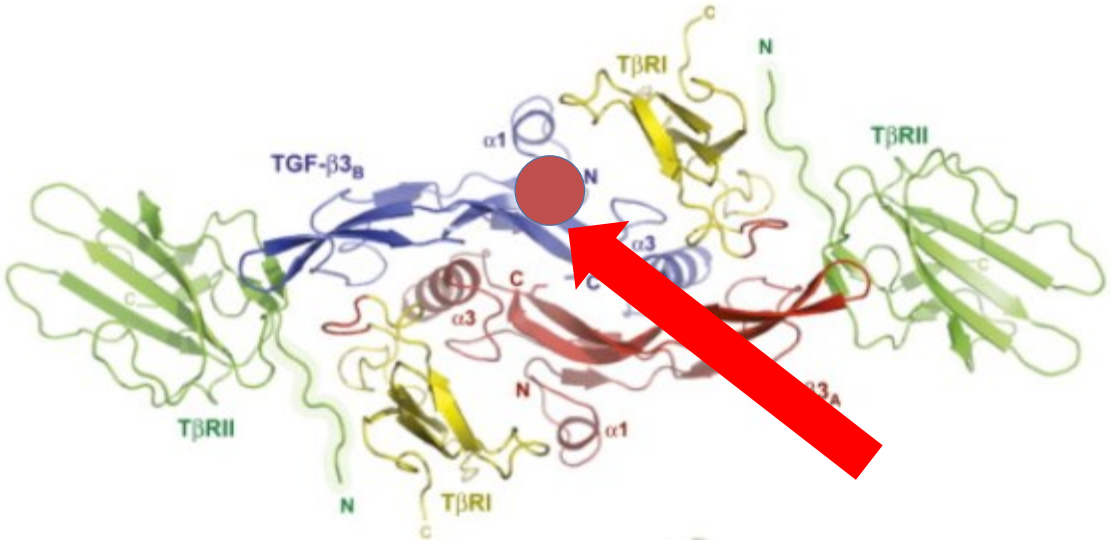
Biomedical Workbench: Identify Genetic Variants



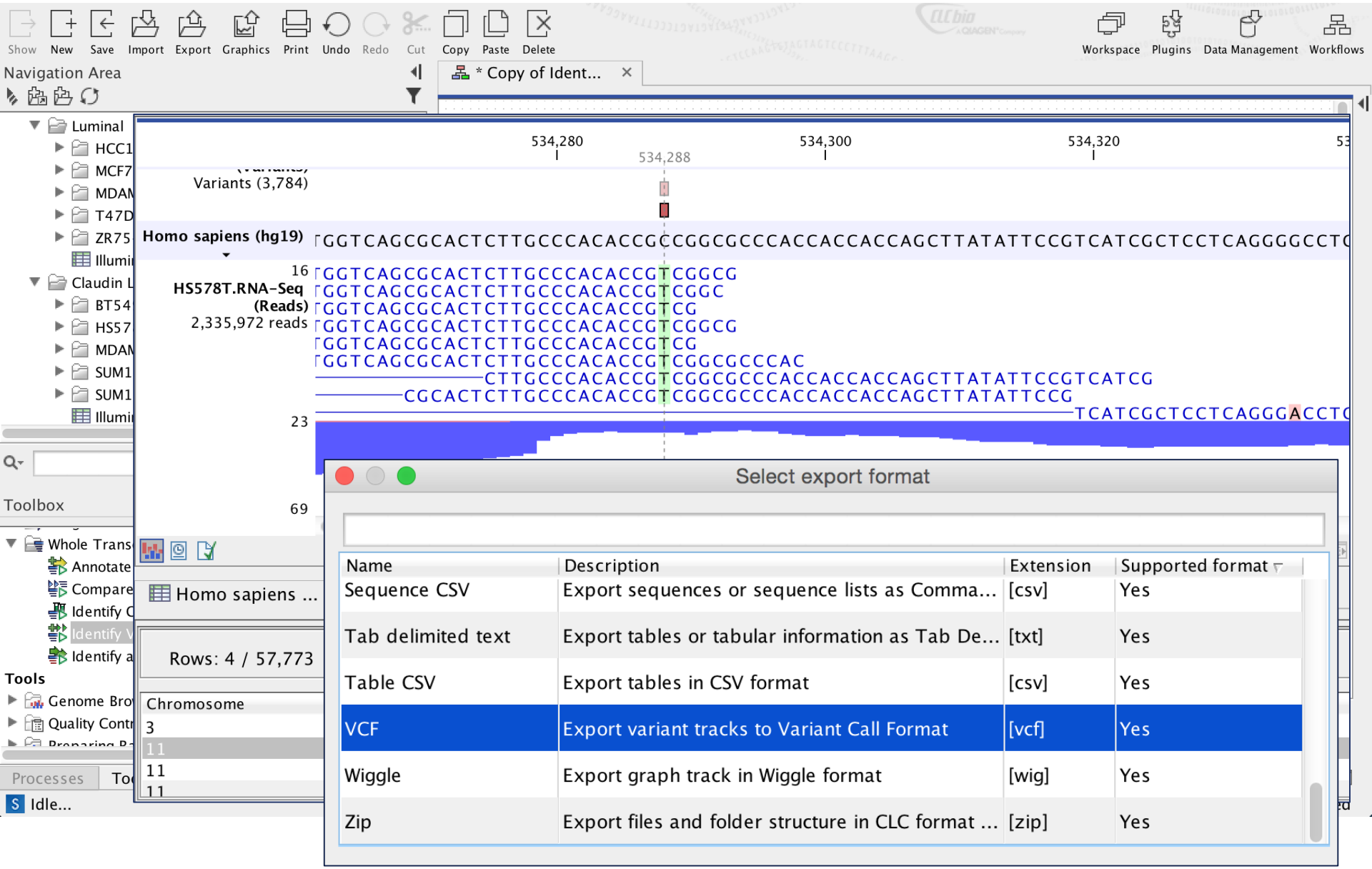
Mutation (in red) disrupt TGF-b3/TbR1 binding



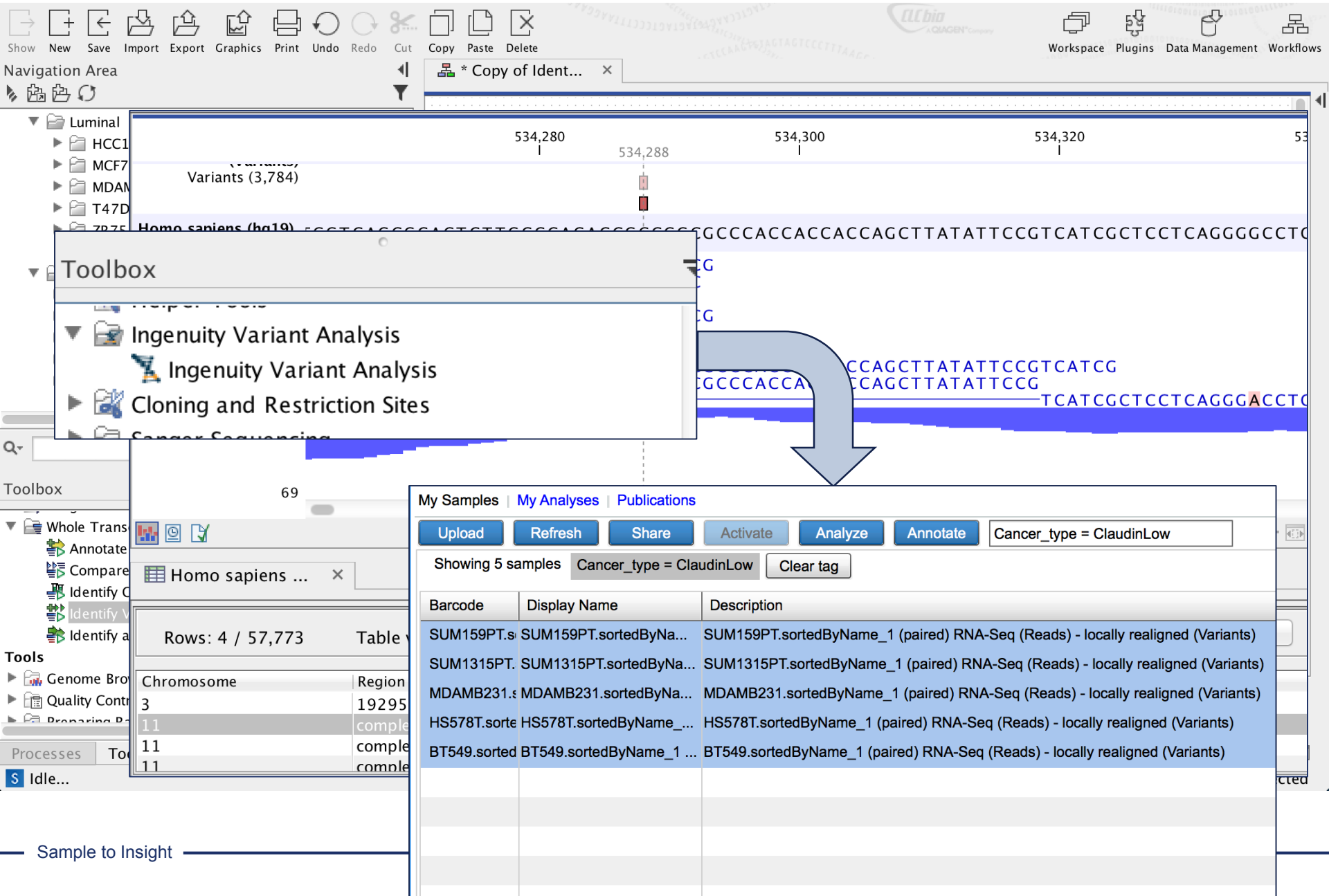




Biomedical Workbench: Identify Genetic Variants

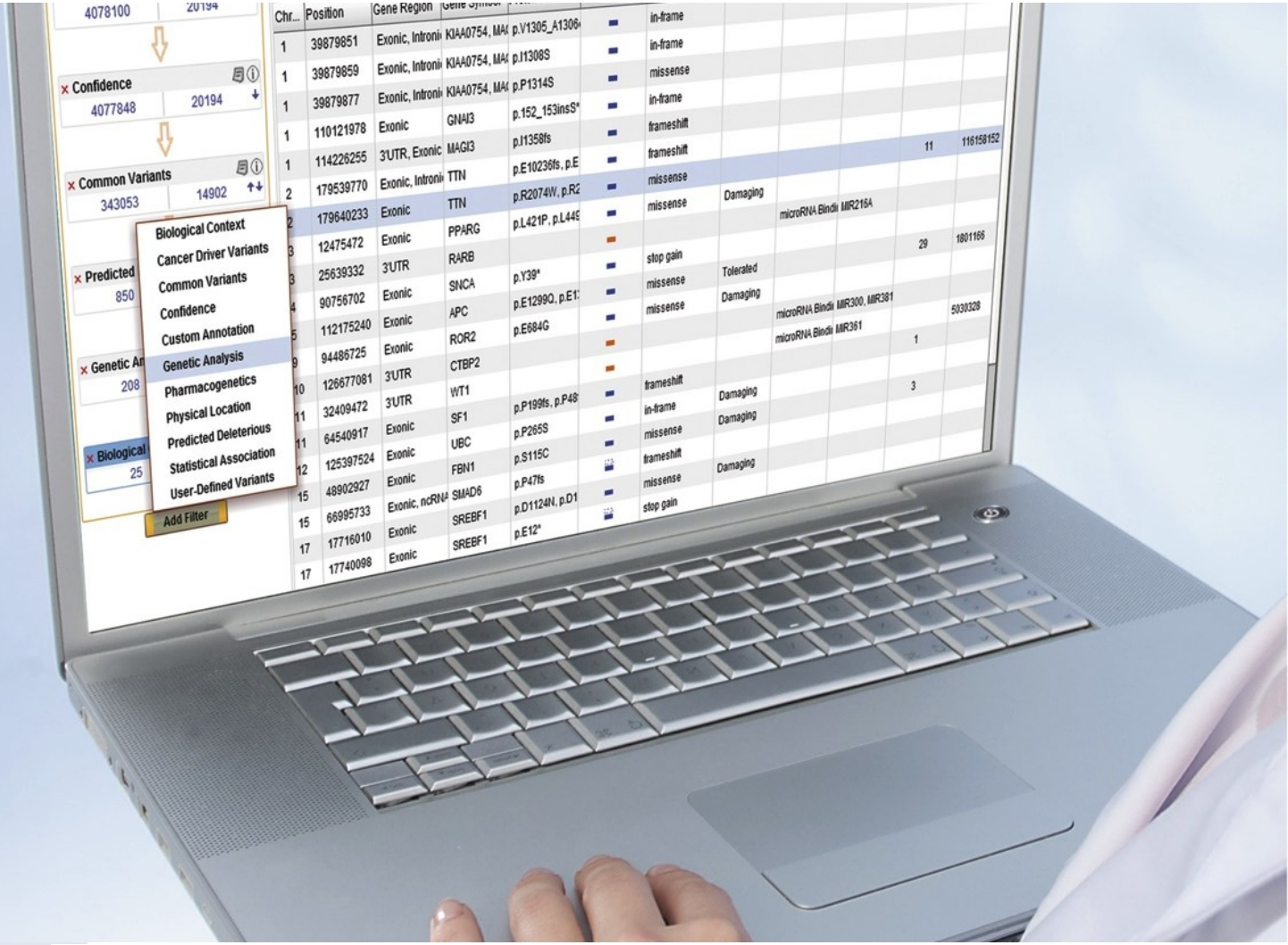


Biomedical Workbench: Integrates with Ingenuity Products



1. QC and preprocess NGS data (RNA-Seq, smRNA, and genomic reads)
2. Differential expression and statistical analysis for RNA-Seq and smRNA
3. Generate, annotate, and compare high-confidence variant calls
4. Facilitate analysis with interactive visualization
5. Construct automated workflows in user friendly interface
6. Modules available for additional or custom functionality
7. Push processed data to Ingenuity and identify pathways affected by genetic variants/expression changes.

QUESTIONS?



HIPAA Health Insurance Portability and Accountability Act

U.S.-EU SAFE HARBOR

TRUSTe CERTIFIED PRIVACY

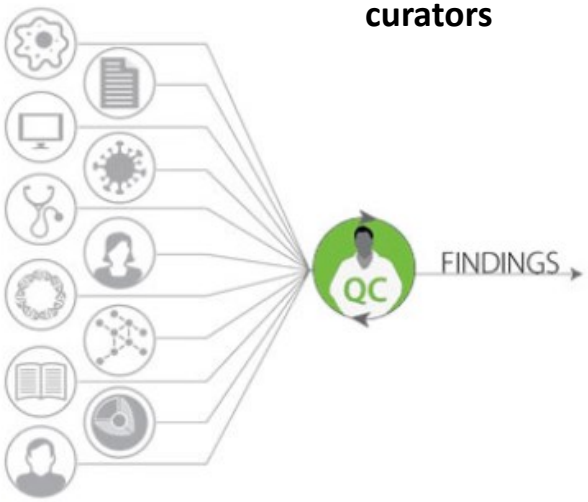
Ingenuity® Variant Analysis™ Overview

A knowledgebase (KB) that's 15 years in the making

Unprecedented Access to Literature Knowledge

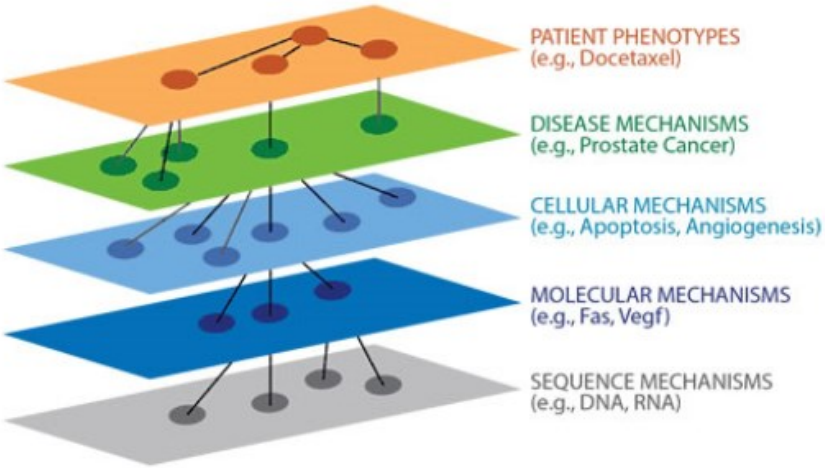
Literature findings

MD/PhD level curators



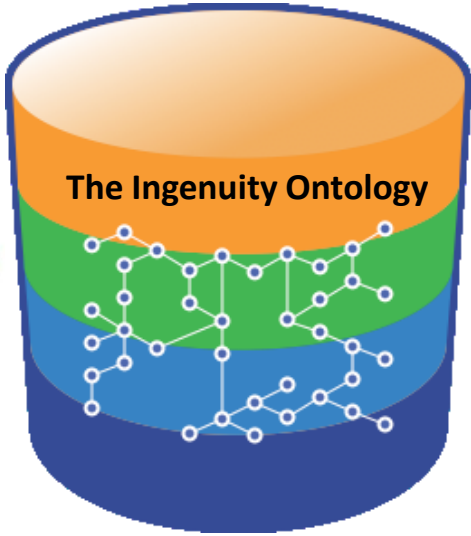
Content Acquisition

Biomedical Ontology



Ingenuity Ontology

The Ingenuity Knowledge Base



Powered By INGENUITY® SYSTEMS

Variant Analysis Content

Quality, Context, Coverage, and Timeliness of Content (ca.1/2014)

COSMIC
Catalogue of somatic mutations in cancer

1000 Genomes
A Deep Catalog of Human Genetic Variation

Mouse Genome Database

The Cancer Genome Atlas
Understanding genomics to improve cancer care

OMIM
Online Mendelian Inheritance in Man
Johns Hopkins University

VISTA

NHLBI Exome Sequencing Project (ESP)
Exome Variant Server

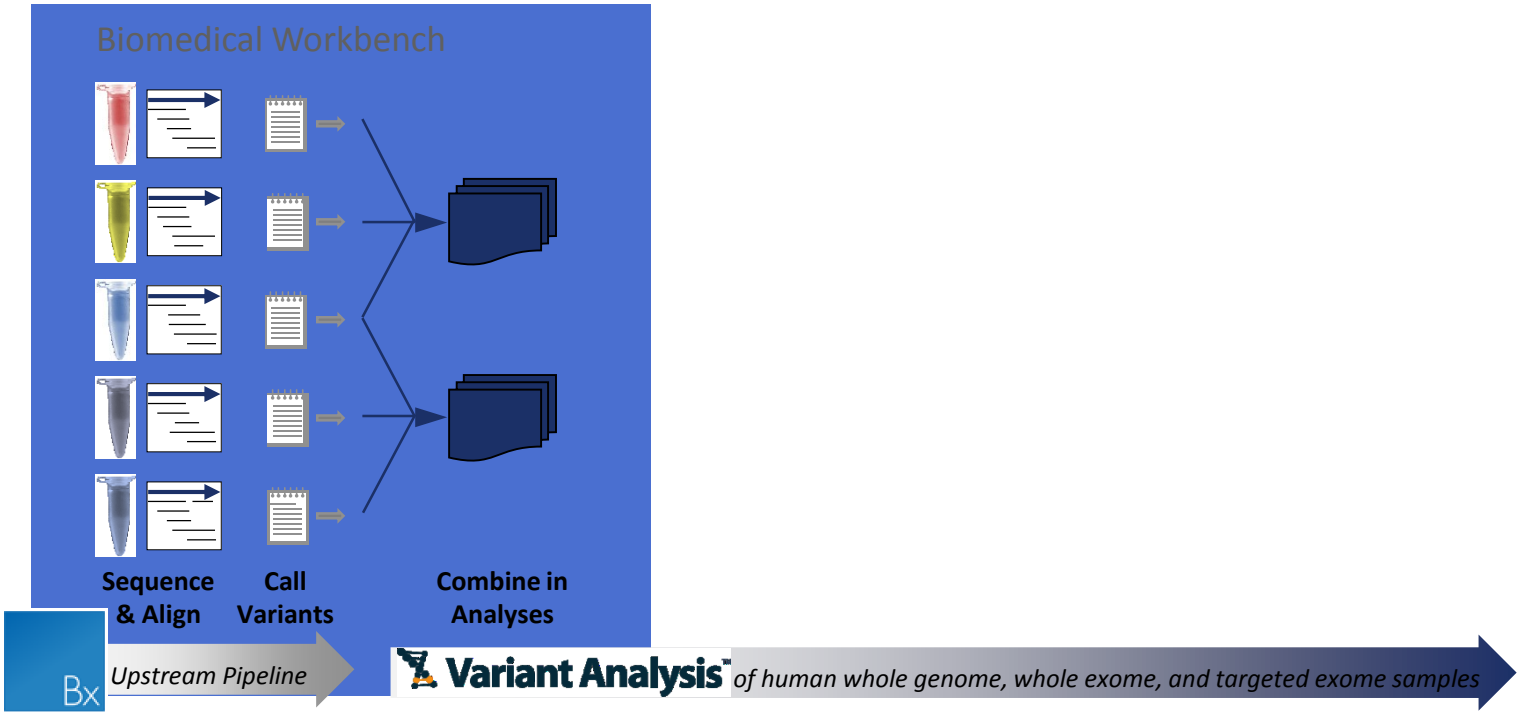
Sorting Intolerant From Tolerant

PolyPhen-2 prediction of functional effects of human nsSNPs

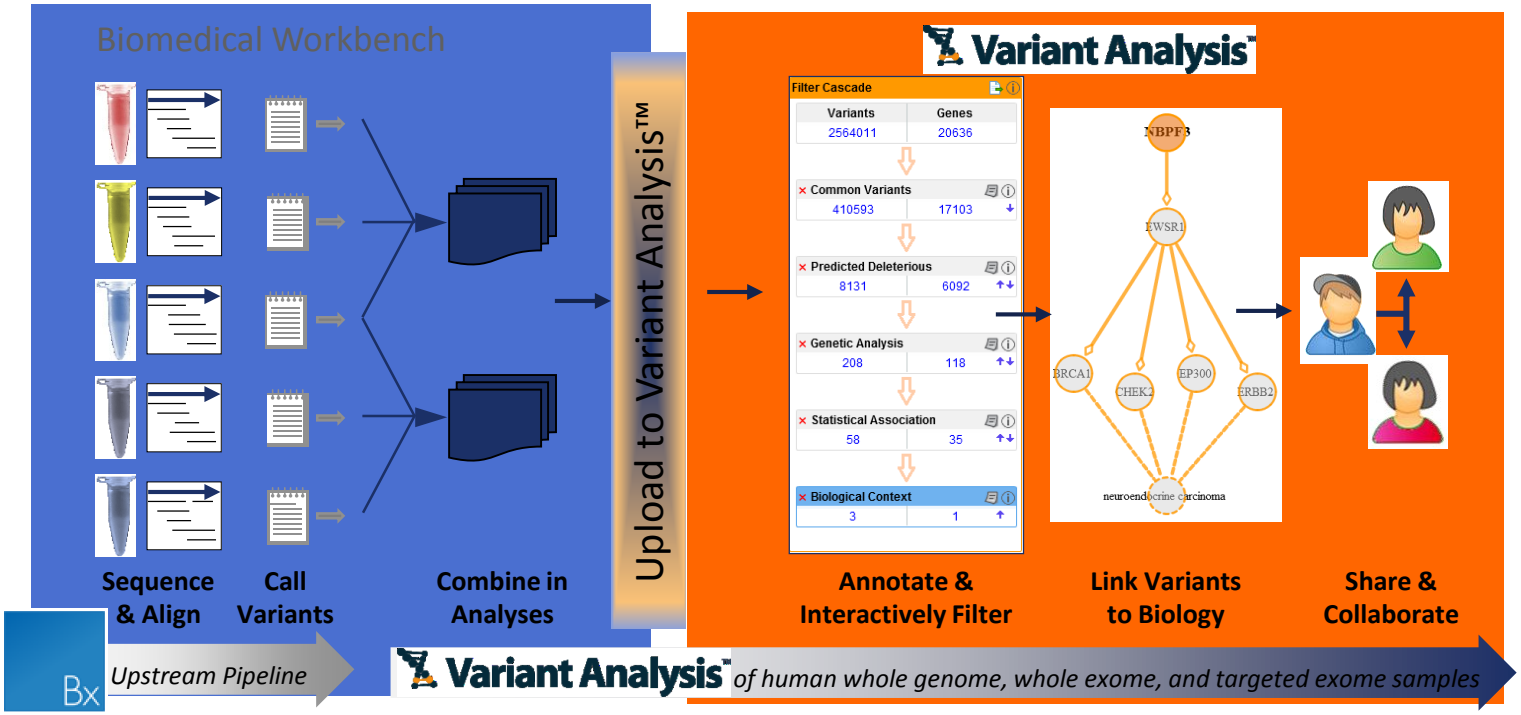
HGMD®
The Human Gene Mutation Database
Cardiff

Additionally

- 349,748+ Ph.D./M.D. expert-curated human phenotype-associated mutation findings
- ~3M+ manual literature findings
- 21,458+ curated disease models
- 185,310+ curated pharmacogenetic (PGx) findings

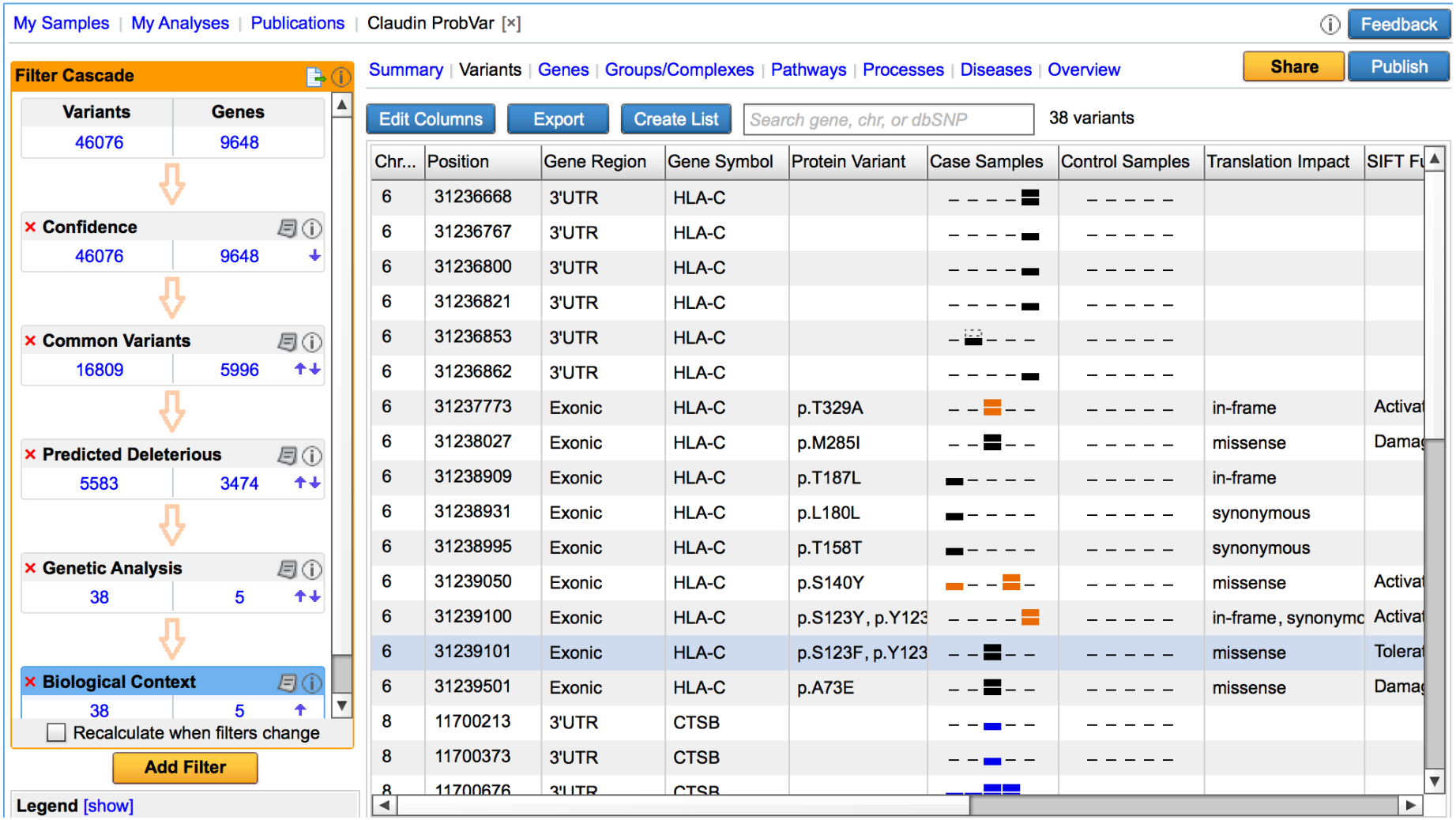


- Identify 'short' list of most plausible variants
  - Causal
  - Biomarkers
- Construct mechanistic hypothesis based on supported biological relationships
- Share with colleagues/ potential collaborators

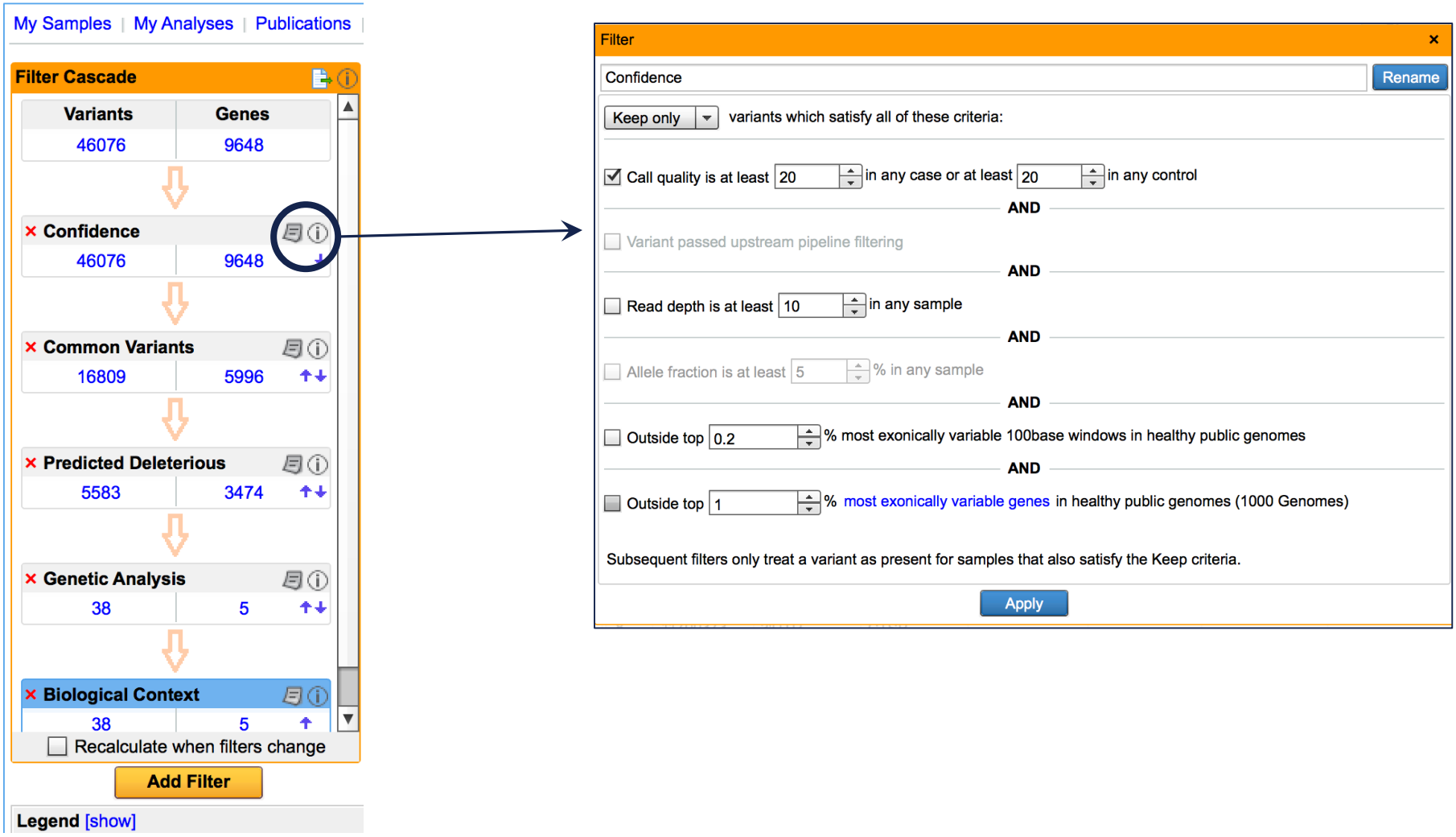


- Identify 'short' list of most plausible variants
  - Causal
  - Biomarkers
- Construct mechanistic hypothesis based on supported biological relationships
- Share with colleagues/ potential collaborators

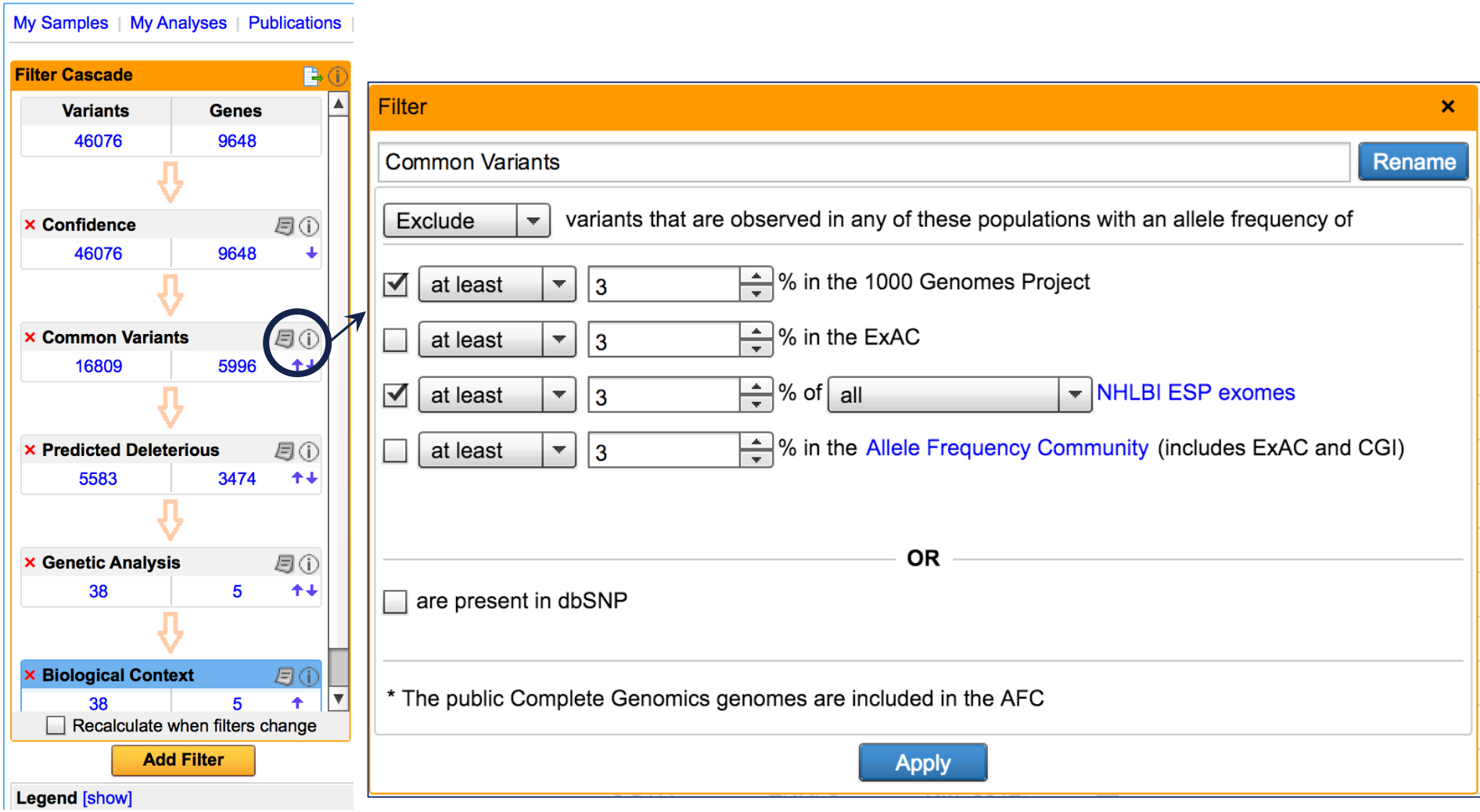
Results: With Default Settings



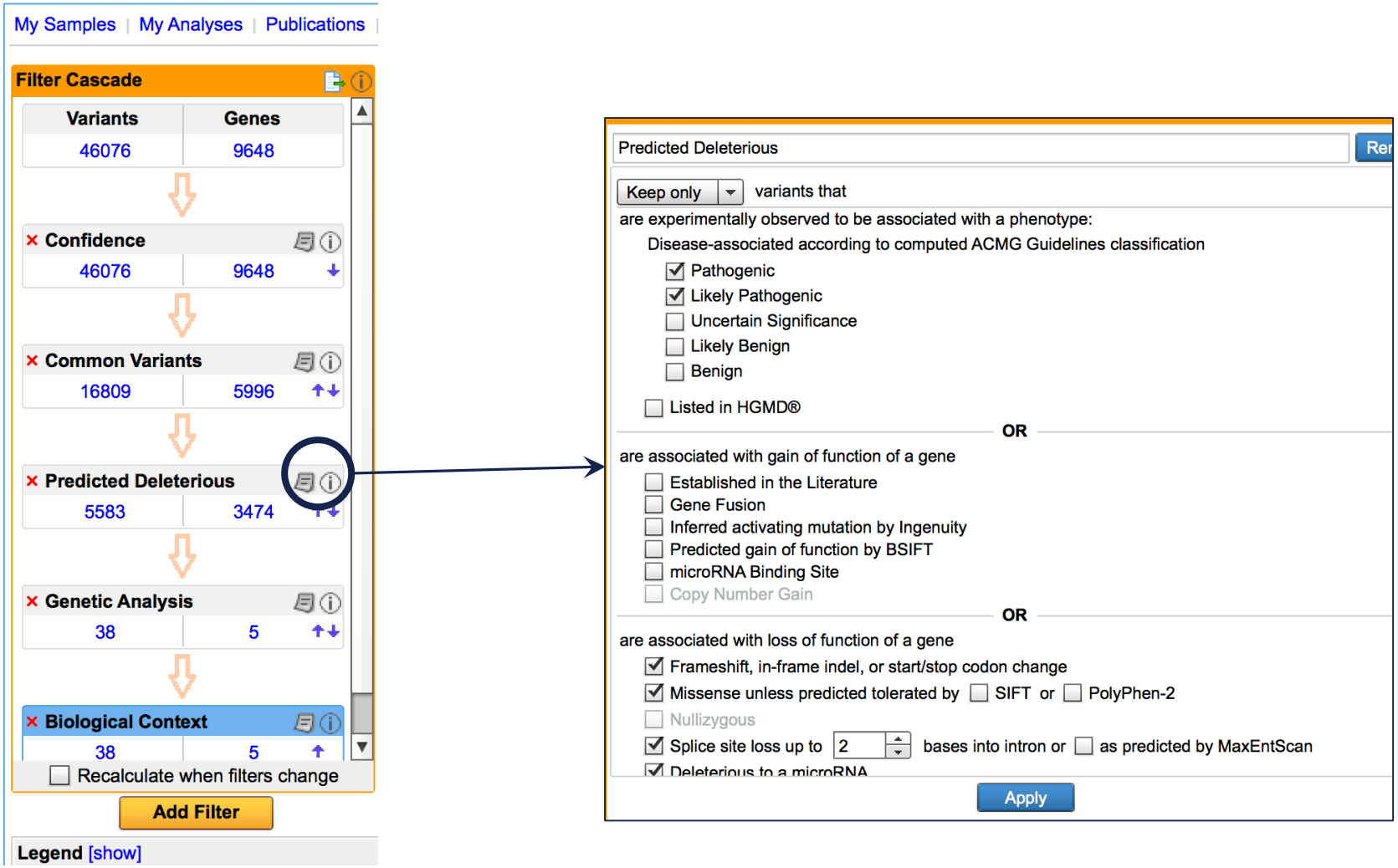
Filter Cascade: Confidence



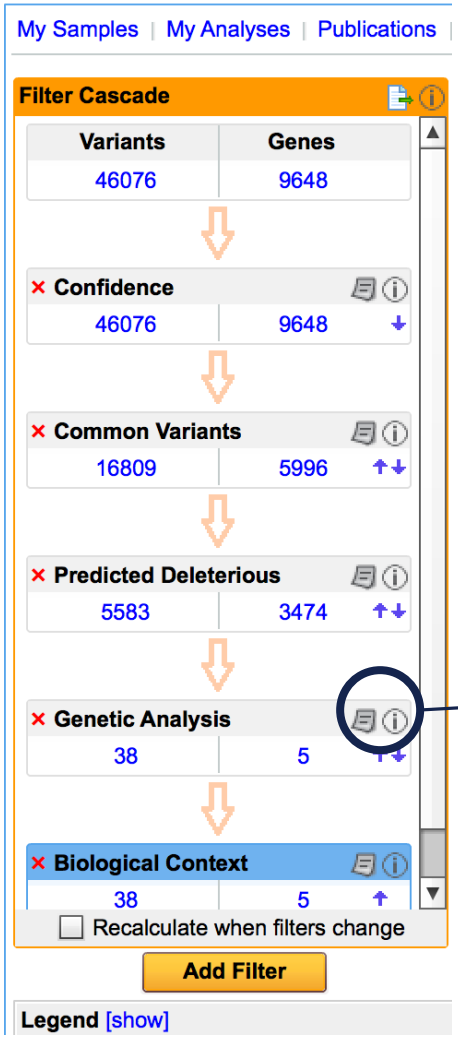
Filter Cascade: Common Variants

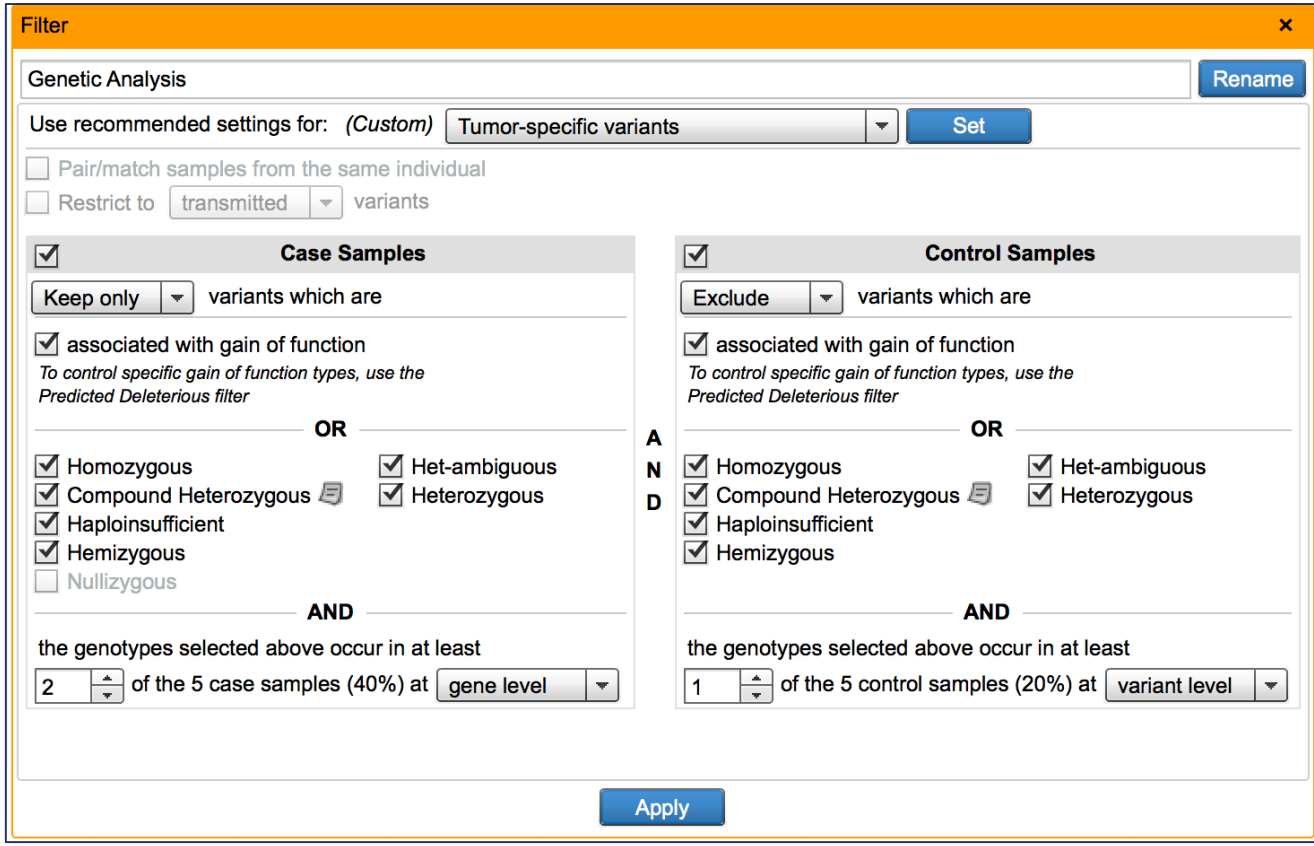


Filter Cascade: Predicted Deleterious

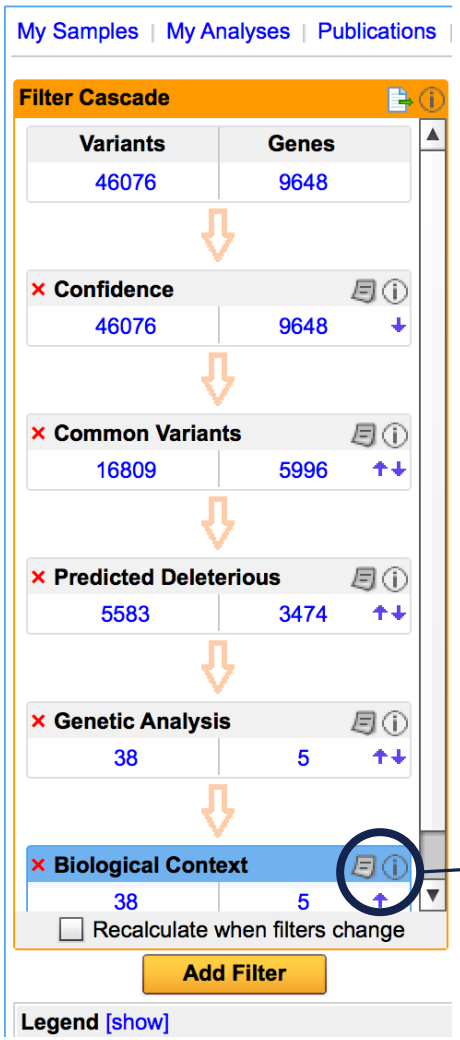


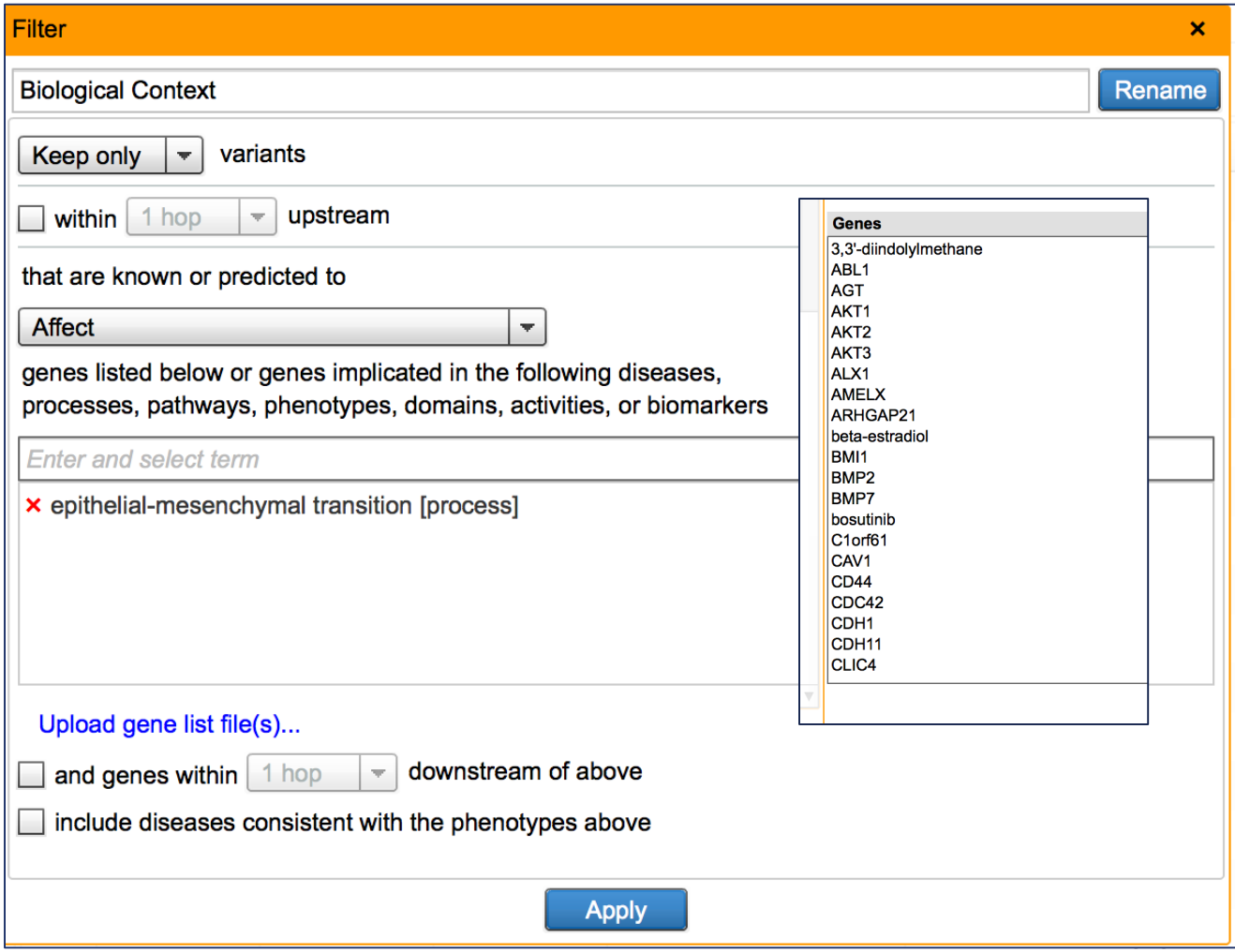
Filter Cascade: Genetic Analysis



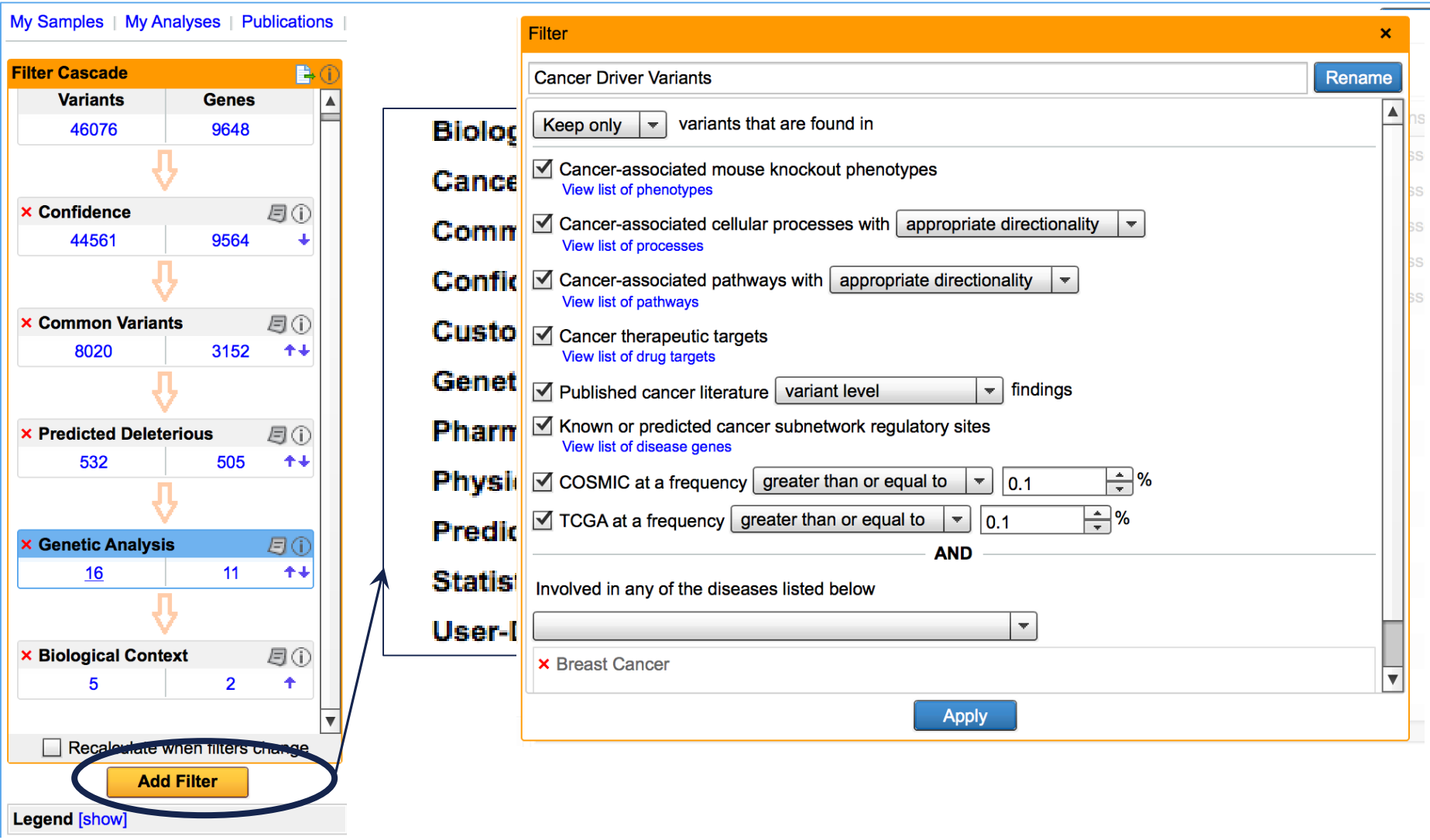


Filter Cascade: Biological Context

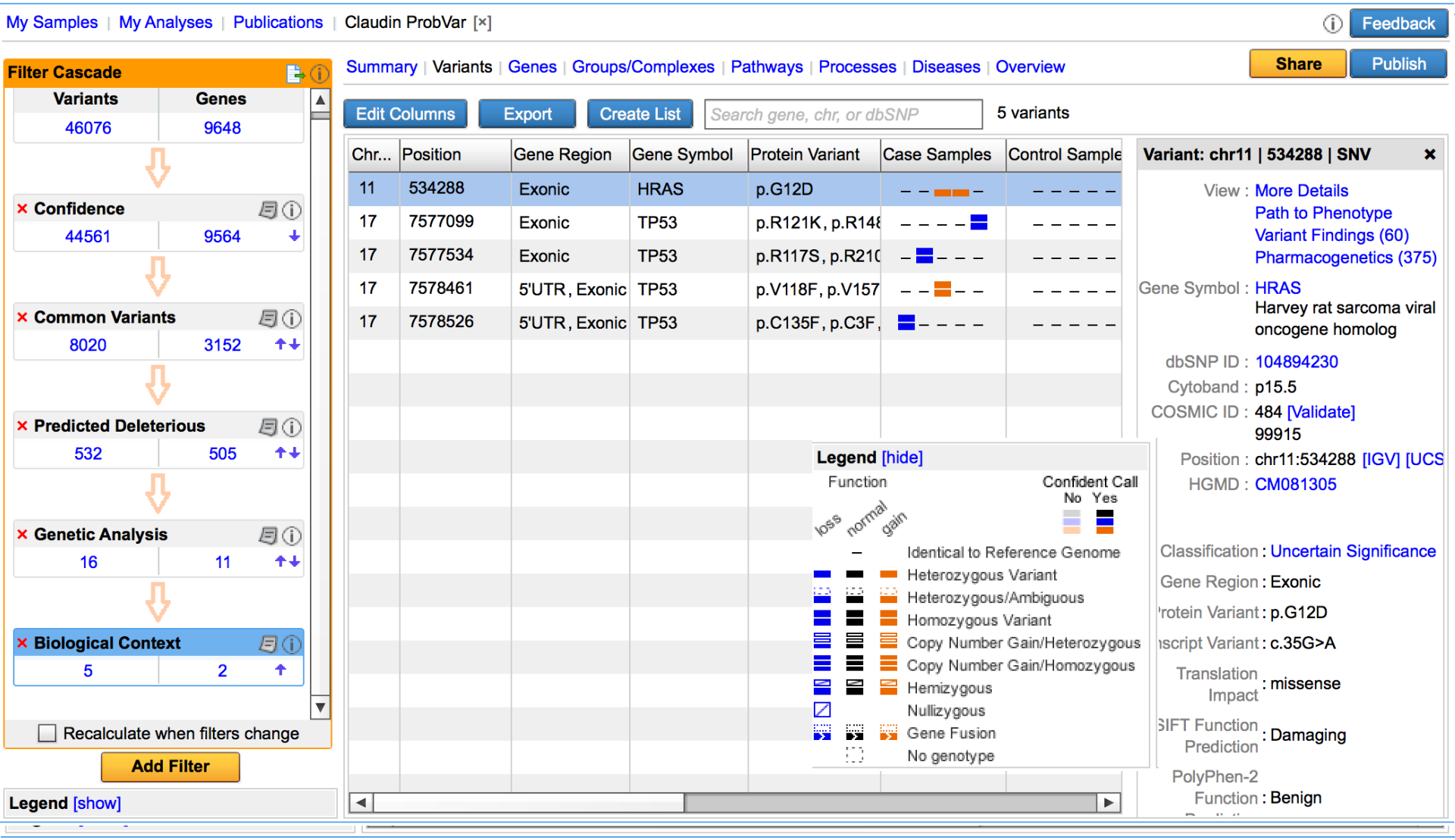




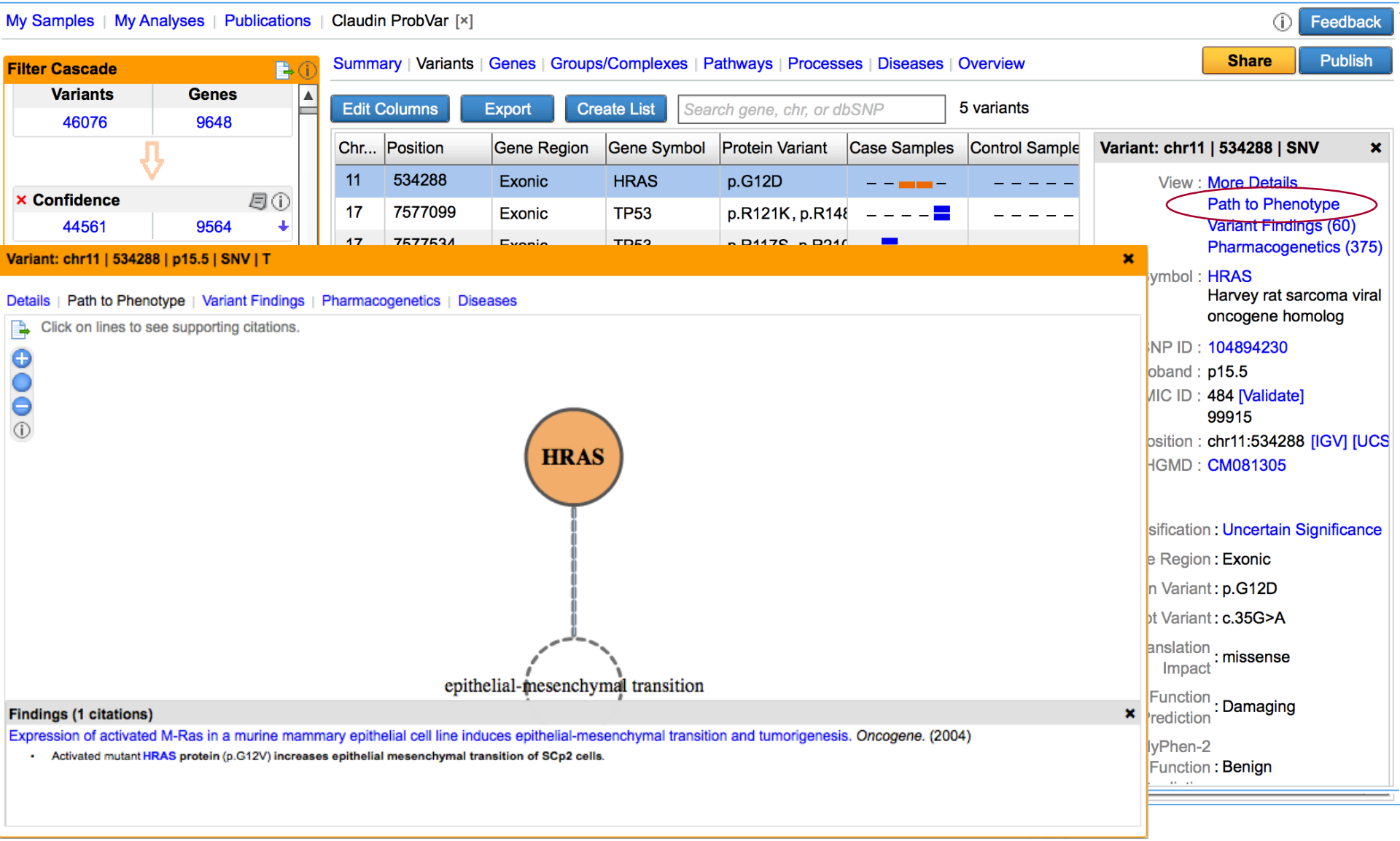
Filter Cascade: Add/Rearrange Filters



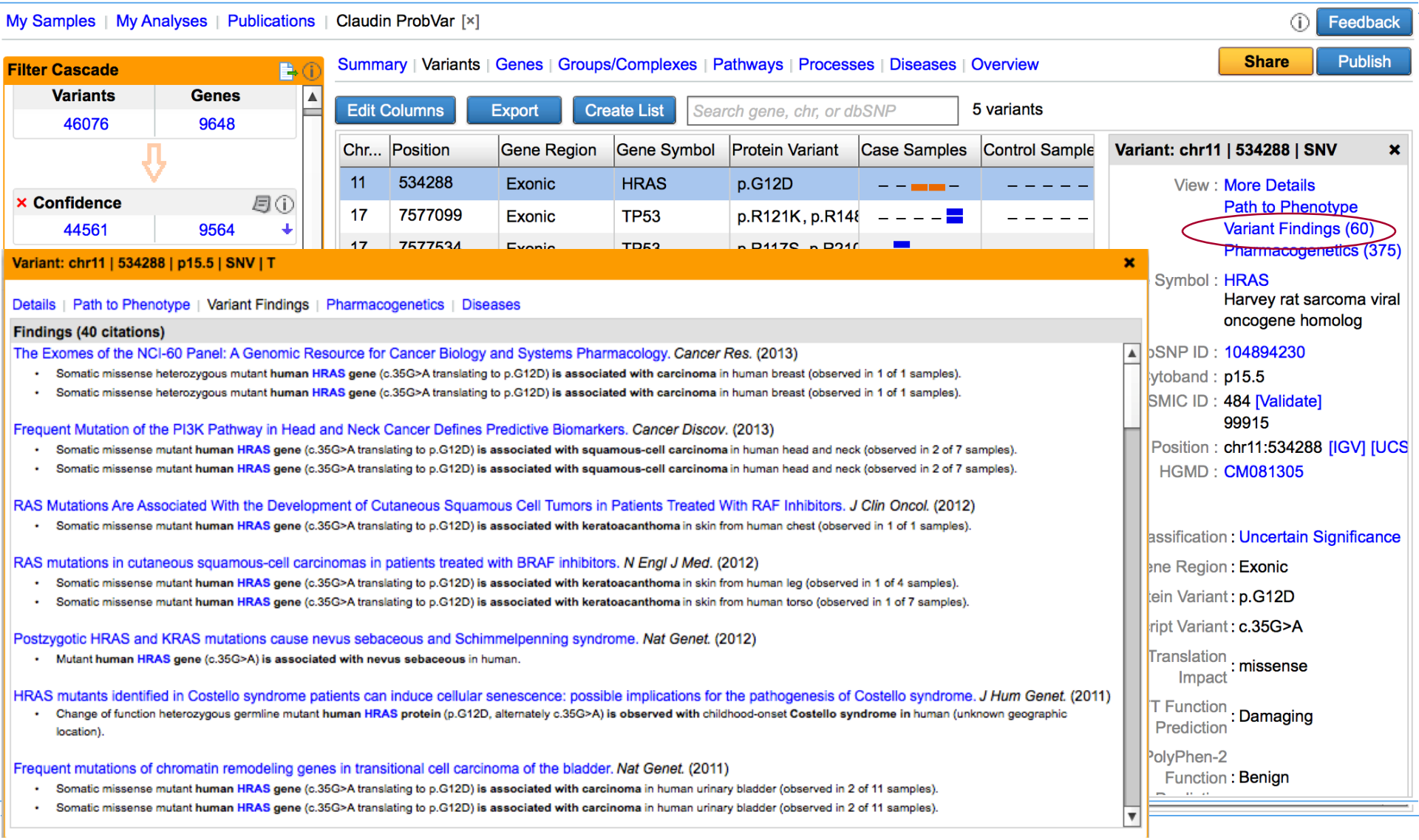
Results: The Short List



Results: The Short List



Results: The Short List



Results: The Short List

My Samples | My Analyses | Publications | Claudin ProbVar [×]

Feedback

Filter Cascade

Variants	Genes
46076	9648
Confidence	
44561	9564
Common Variants	
8020	3152
Predicted Deleterious	
532	505
Genetic Analysis	

Summary | Variants | Genes | Groups/Complexes | Pathways | Processes | Diseases | Overview

Share | Publish

Edit Columns | Export | Create List | Search gene, chr, or dbSNP | 5 variants

Chr...	Position	Gene Region	Gene Symbol	Protein Variant	Case Samples	Control Sample
11	534288	Exonic	HRAS	p.G12D		
17	7577099	Exonic	TP53	p.R121K, p.R148		
17	7577534	Exonic	TP53	p.R117S, p.R210		
17	7578461	5'UTR, Exonic	TP53	p.V118F, p.V157		
17	7578526	5'UTR, Exonic	TP53	p.C135F, p.C3F,		

Variant: chr11 | 534288 | SNV

View : More Details
Path to Phenotype
Variant Findings (60)
Pharmacogenetics (375)

Gene Symbol : HRAS
Harvey rat sarcoma viral oncogene homolog

dbSNP ID : 104894230
Cytoband : p15.5
COSMIC ID : 484 [Validate]
99915
Position : chr11:534288 [IGV] [UCS
HGMD : CM081305

Classification : Uncertain Significance

HGMD® Professional 2014.1

Gene Mutation Reference Batch Advanced | Statistics Information Support | Home

HGMD accession	Disease/phenotype	Gene symbol	Codon change	Amino acid change	Codon number	Feedback
CM081305	Costello syndrome	HRAS	GGC-GAC	Gly-Asp	12	Feedback

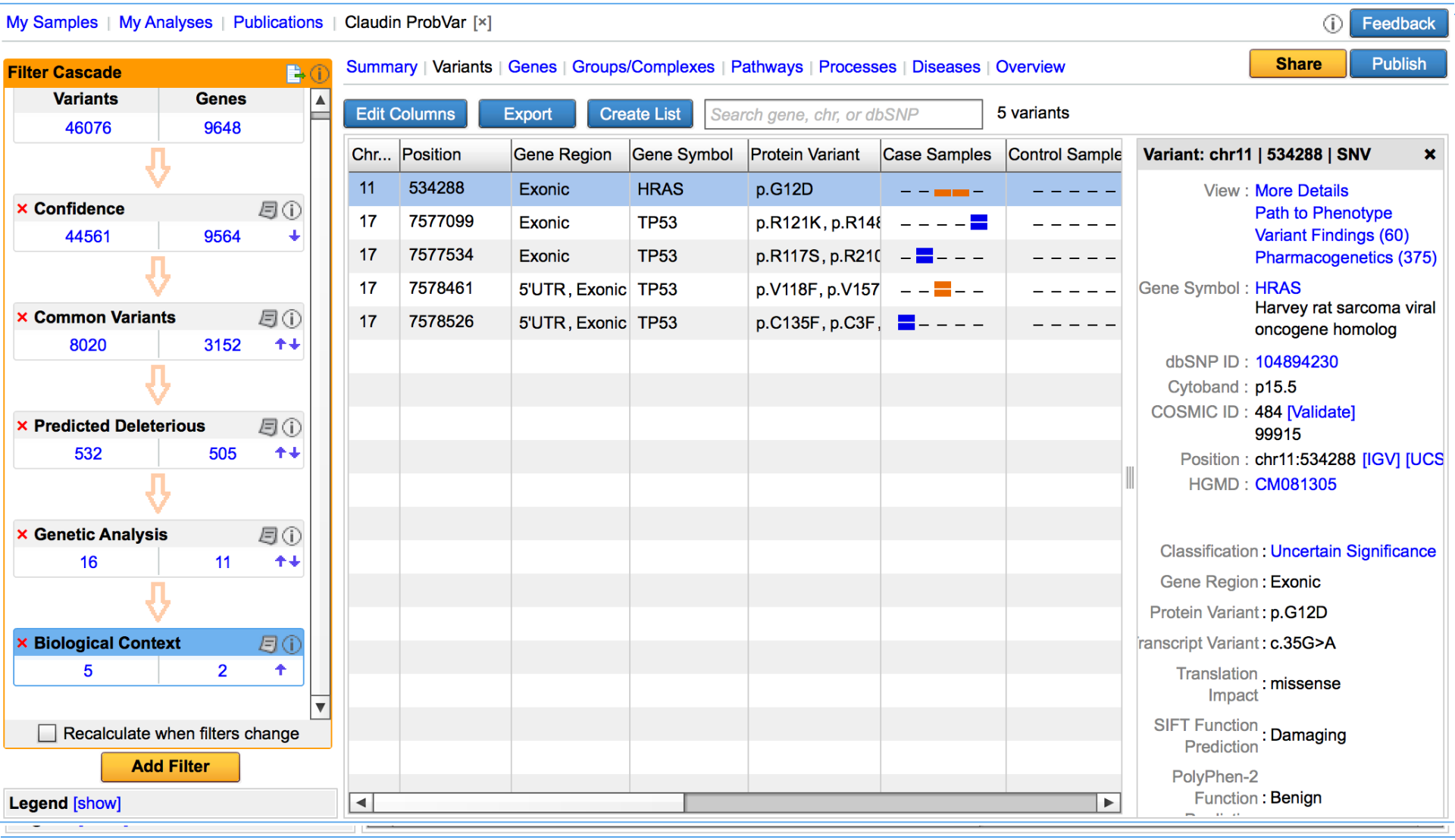
The G12D substitution exhibits a shift in polarity from non-polar to negatively charged and displays a decrease in Kyte-Doolittle hydrophobicity from -0.4 to -3.5. Approximately 1.95% of missense mutations in HGMD are Gly-Asp. The mutation occurs 178 amino acids from the end of the protein.

Literature citation	Citation type	Notes
1. Lo (2008) *J Med Genet* **45:** 167 PubMed: 18039947	Primary literature report	None
2. Niihori (2011) *J Hum Genet* **56:** 707 PubMed: 21850009	Functional characterisation	None
3. Cheng (2012) *PLoS Comput Biol* **8:** e1002738 PubMed: 23093928	Functional characterisation	Systemic impact quantified as a combination of protein stability change and pathway perturbation.

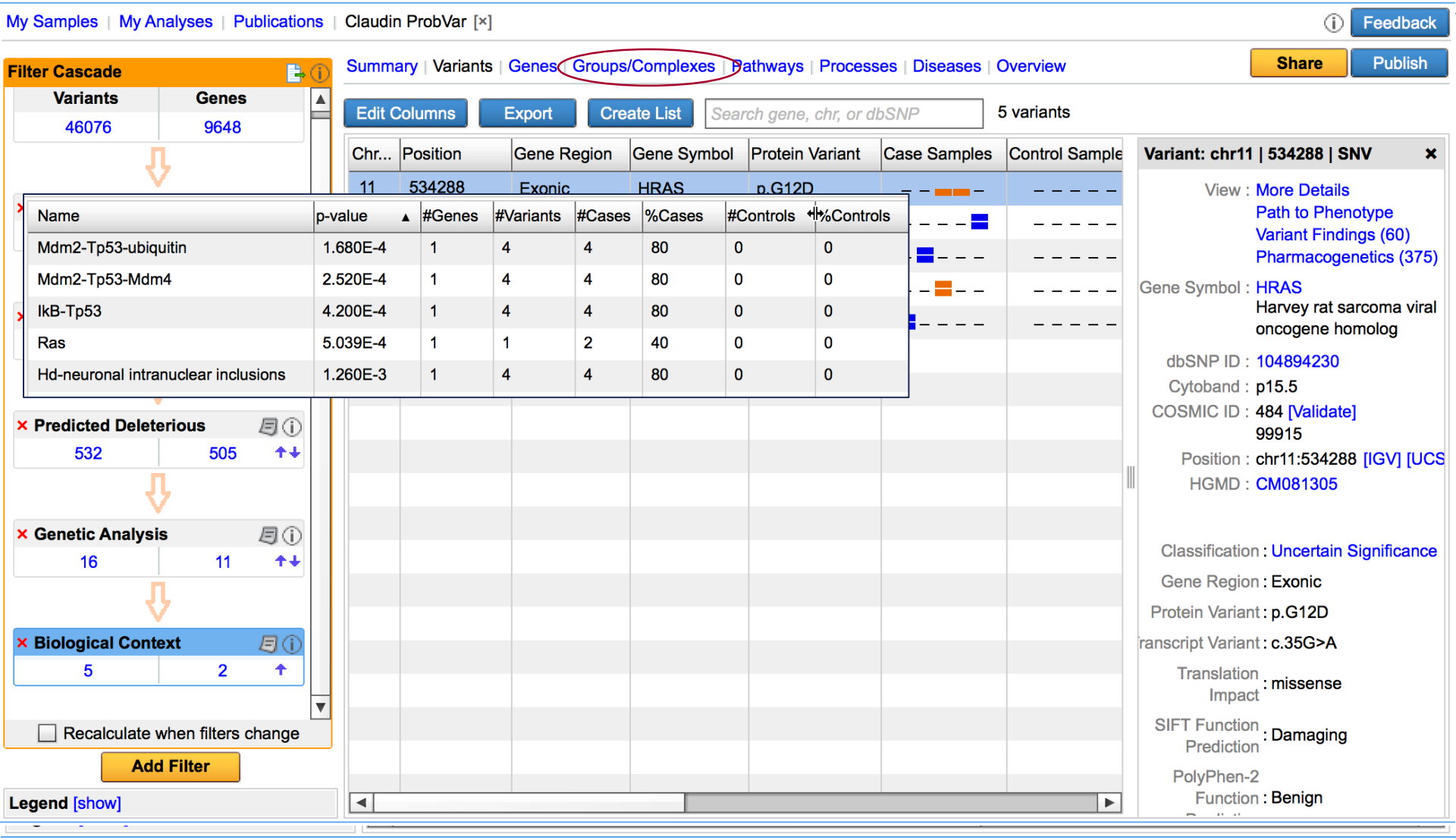
Extra information

Coding strand genomic sequence (GRCh37.3) CGGAATATAAGCTGGTGGTGGTGGGCGCCG(G/A)CGGTGTGGGCAAGAGTGCGCTGACCATCCA

Results: The Short List



Results: The Short List



Results: The Short List

My Samples | My Analyses | Publications | Claudin ProbVar [×]

Feedback

Filter Cascade

Variants	Genes
46076	9648

Confidence: 44561 | 9564

Common Variants: 8020

Predicted Deleteri: 532

Genetic Analysis: 16

Biological Context: 5

Recalculate whe

Add Fi

Legend [show]

Summary | Variants | Genes | Groups/Complexes | Pathways | Processes | Diseases | Overview

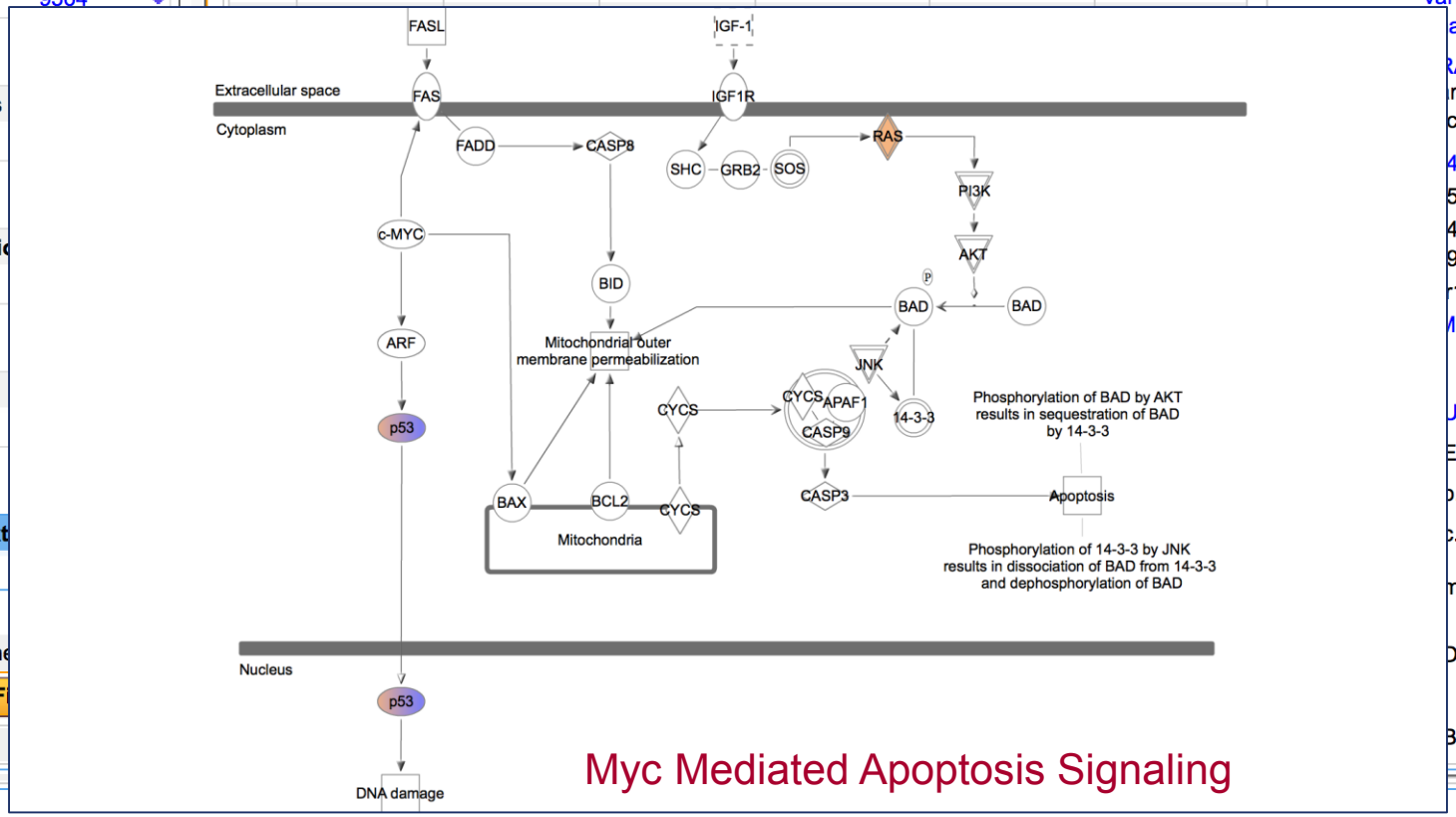
Share | Publish

Edit Columns | Export | Create List | Search gene, chr, or dbSNP | 5 variants

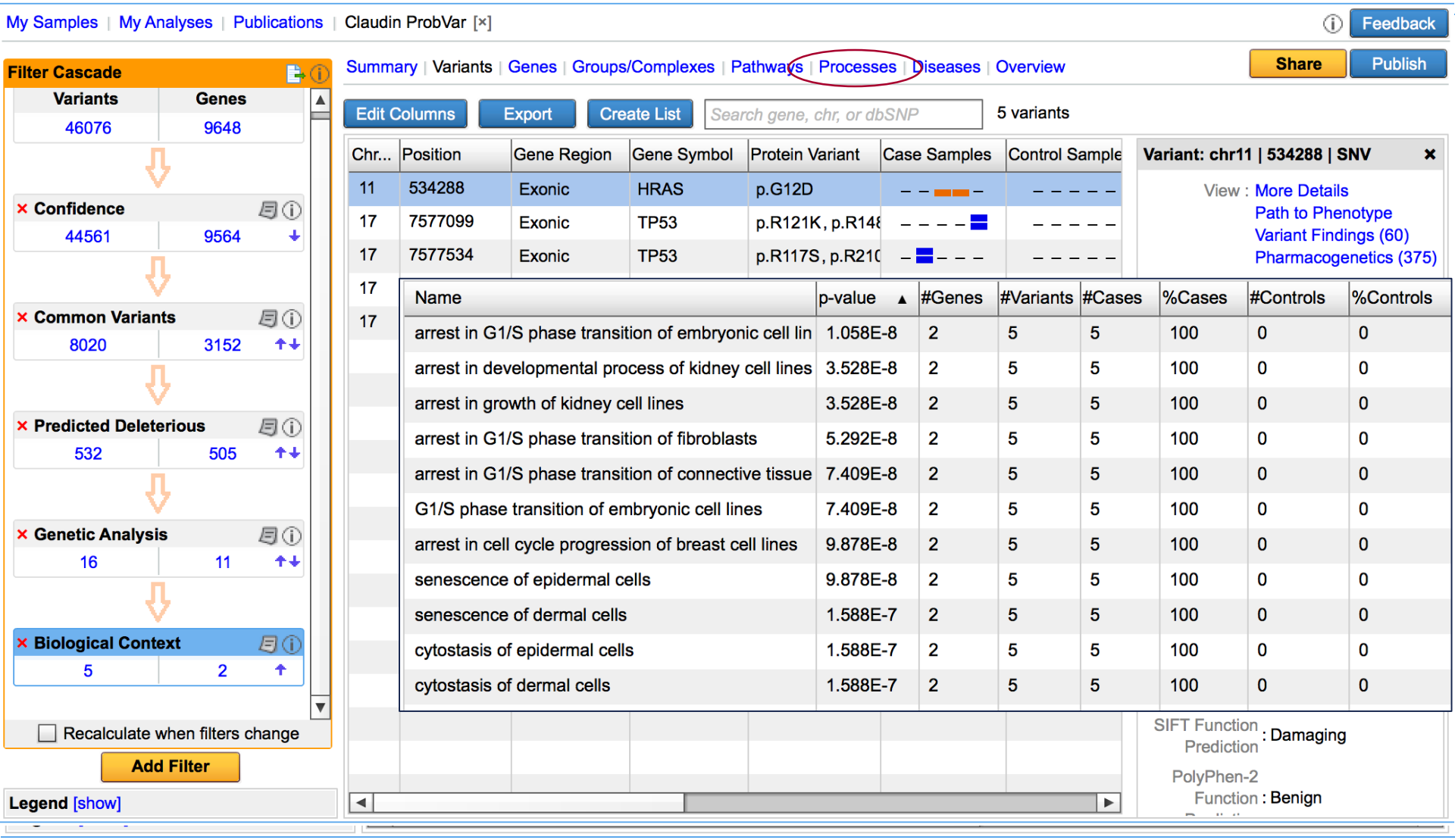
Chr...	Position	Gene Region	Gene Symbol	Protein Variant	Case Samples	Control Sample
11	534288	Exonic	HRAS	p.G12D		
17	7577099	Exonic	TP53	p.R121K, p.R148		

Variant: chr11 | 534288 | SNV

View : More Details
Path to Phenotype
Variant Findings (60)
armacogenetics (375)
RAS
rvey rat sarcoma viral
cogene homolog
4894230
5.5
4 [Validate]
915
r11:534288 [IGV] [UCS
M081305
Uncertain Significance
Exonic
.G12D
.35G>A
missense
Damaging
Benign

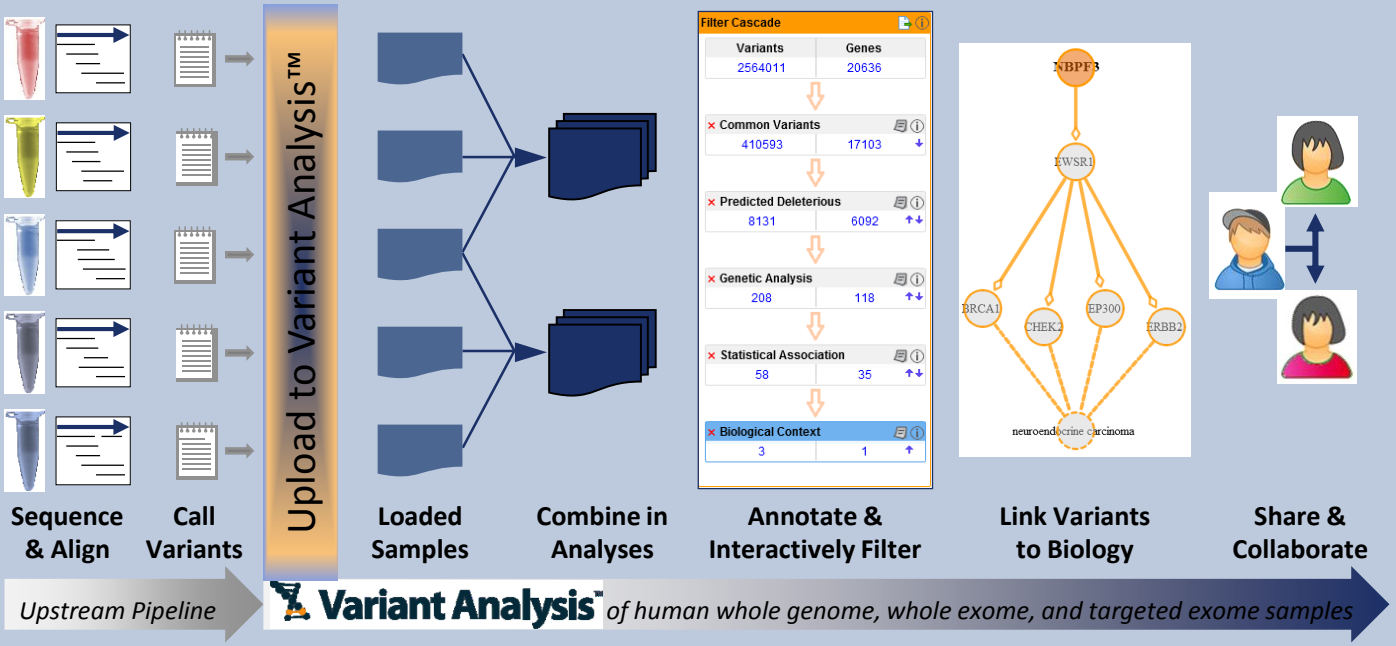


Results: The Short List

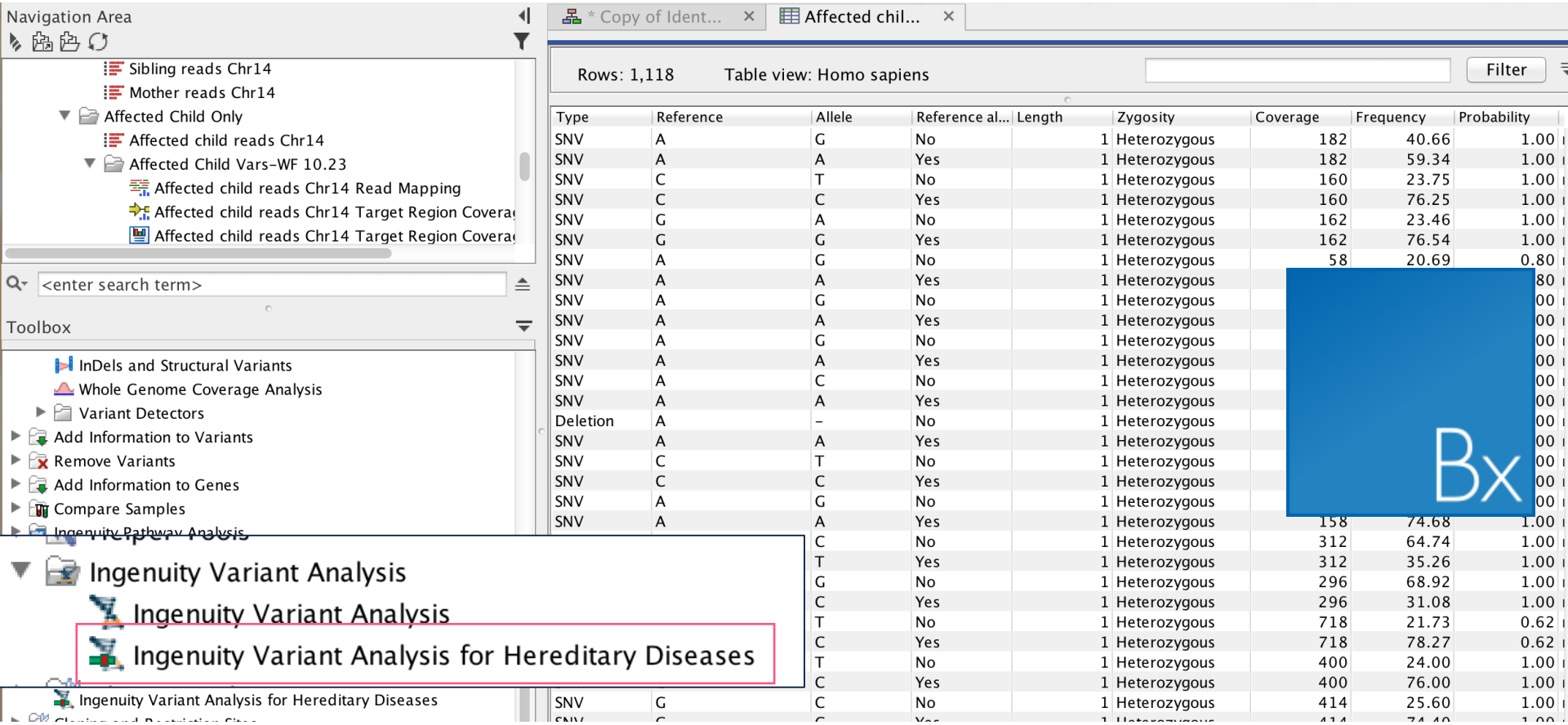


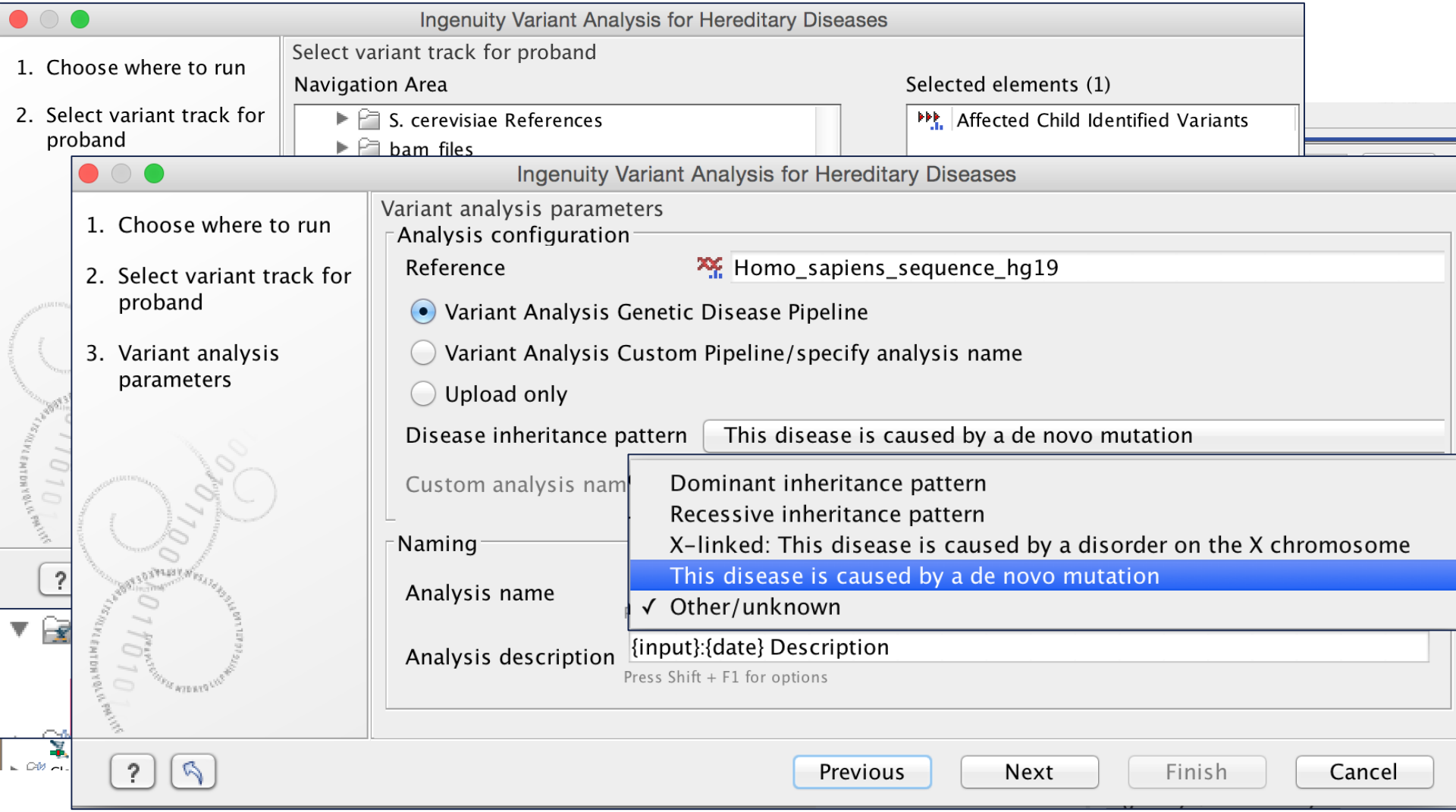
Variant Analysis- Available as separate software

Can add it to your current fastq->vcf pipeline



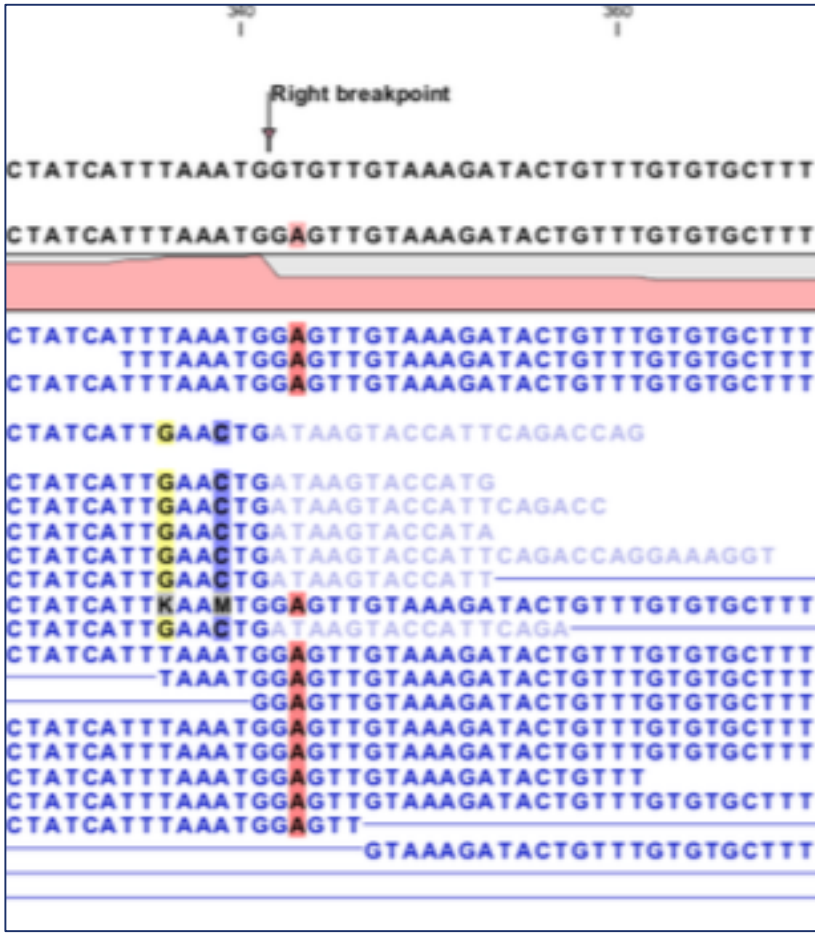
- Identify 'short' list of most plausible variants
  - Causal
  - Biomarkers
- Construct mechanistic hypothesis based on supported biological relationships
- Share with colleagues/ potential collaborators


Navigation Area
Sibling reads Chr14
Mother reads Chr14
Affected Child Only
Affected child reads Chr14
Affected Child Vars-WF 10.23
Affected child reads Chr14 Read Mapping
Affected child reads Chr14 Target Region Coverag
Affected child reads Chr14 Target Region Coverag
<enter search term>
Toolbox
InDels and Structural Variants
Whole Genome Coverage Analysis
Variant Detectors
Add Information to Variants
Remove Variants
Add Information to Genes
Compare Samples
Ingenuity Pathway Analysis
Ingenuity Variant Analysis
Ingenuity Variant Analysis
Ingenuity Variant Analysis for Hereditary Diseases
Ingenuity Variant Analysis for Hereditary Diseases
* Copy of Ident...
Affected chil...
Rows: 1,118
Table view: Homo sapiens
Filter
Type | Reference | Allele | Reference al... | Length | Zygosity | Coverage | Frequency | Probability
SNV | A | G | No | 1 | Heterozygous | 182 | 40.66 | 1.00
SNV | A | A | Yes | 1 | Heterozygous | 182 | 59.34 | 1.00
SNV | C | T | No | 1 | Heterozygous | 160 | 23.75 | 1.00
SNV | C | C | Yes | 1 | Heterozygous | 160 | 76.25 | 1.00
SNV | G | A | No | 1 | Heterozygous | 162 | 23.46 | 1.00
SNV | G | G | Yes | 1 | Heterozygous | 162 | 76.54 | 1.00
SNV | A | G | No | 1 | Heterozygous | 58 | 20.69 | 0.80
SNV | A | A | Yes | 1 | Heterozygous
SNV | A | G | No | 1 | Heterozygous
SNV | A | A | Yes | 1 | Heterozygous
SNV | A | G | No | 1 | Heterozygous
SNV | A | A | Yes | 1 | Heterozygous
SNV | A | C | No | 1 | Heterozygous
SNV | A | A | Yes | 1 | Heterozygous
Deletion | A | - | No | 1 | Heterozygous
SNV | A | A | Yes | 1 | Heterozygous
SNV | C | T | No | 1 | Heterozygous
SNV | C | C | Yes | 1 | Heterozygous
SNV | A | G | No | 1 | Heterozygous
SNV | A | A | Yes | 1 | Heterozygous | 158 | 74.68 | 1.00
C | No | 1 | Heterozygous | 312 | 64.74 | 1.00
T | Yes | 1 | Heterozygous | 312 | 35.26 | 1.00
G | No | 1 | Heterozygous | 296 | 68.92 | 1.00
C | Yes | 1 | Heterozygous | 296 | 31.08 | 1.00
T | No | 1 | Heterozygous | 718 | 21.73 | 0.62
C | Yes | 1 | Heterozygous | 718 | 78.27 | 0.62
T | No | 1 | Heterozygous | 400 | 24.00 | 1.00
C | Yes | 1 | Heterozygous | 400 | 76.00 | 1.00
SNV | G | C | No | 1 | Heterozygous | 414 | 25.60 | 1.00
Bx


Ingenuity Variant Analysis for Hereditary Diseases
1. Choose where to run
2. Select variant track for proband
Select variant track for proband
Navigation Area
S. cerevisiae References
bam_files
Selected elements (1)
Affected Child Identified Variants
Ingenuity Variant Analysis for Hereditary Diseases
1. Choose where to run
2. Select variant track for proband
3. Variant analysis parameters
Variant analysis parameters
Analysis configuration
Reference
Homo_sapiens_sequence_hg19
Variant Analysis Genetic Disease Pipeline
Variant Analysis Custom Pipeline/specify analysis name
Upload only
Disease inheritance pattern
This disease is caused by a de novo mutation
Custom analysis nam
Dominant inheritance pattern
Recessive inheritance pattern
X-linked: This disease is caused by a disorder on the X chromosome
This disease is caused by a de novo mutation
Other/unknown
Naming
Analysis name
Analysis description
{input}:{date} Description
Press Shift + F1 for options
Previous
Next
Finish
Cancel

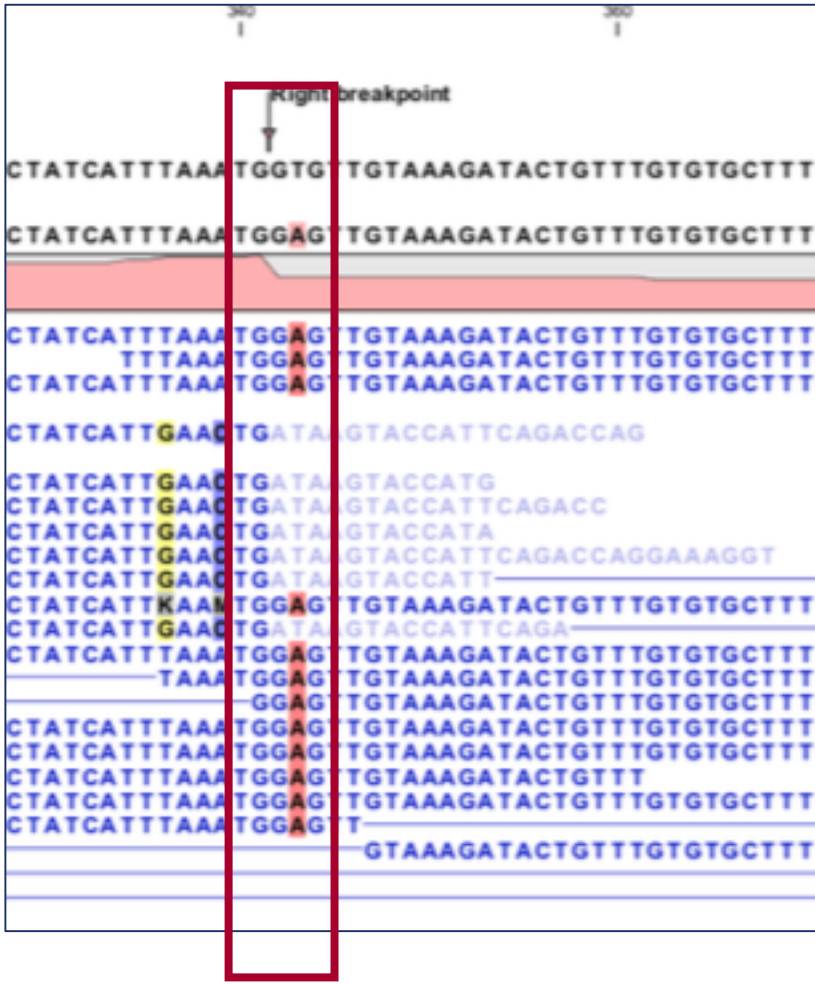
Indel and Structural Variants

- Identify insertions, deletions [Indels], inversions, translocations and tandem duplications [Structural] from read mappings.
- Relies on unmapped read ends
- The Algorithm
  - First identifies positions with an excess of reads with left (or right) unaligned ends.
  - Determines consensus sequences of the unaligned ends
  - Maps the consensus sequences to the reference sequence around other positions with unaligned ends.
  - Structural variant is called when mappings are in accordance with a 'signature' of a structural variant.



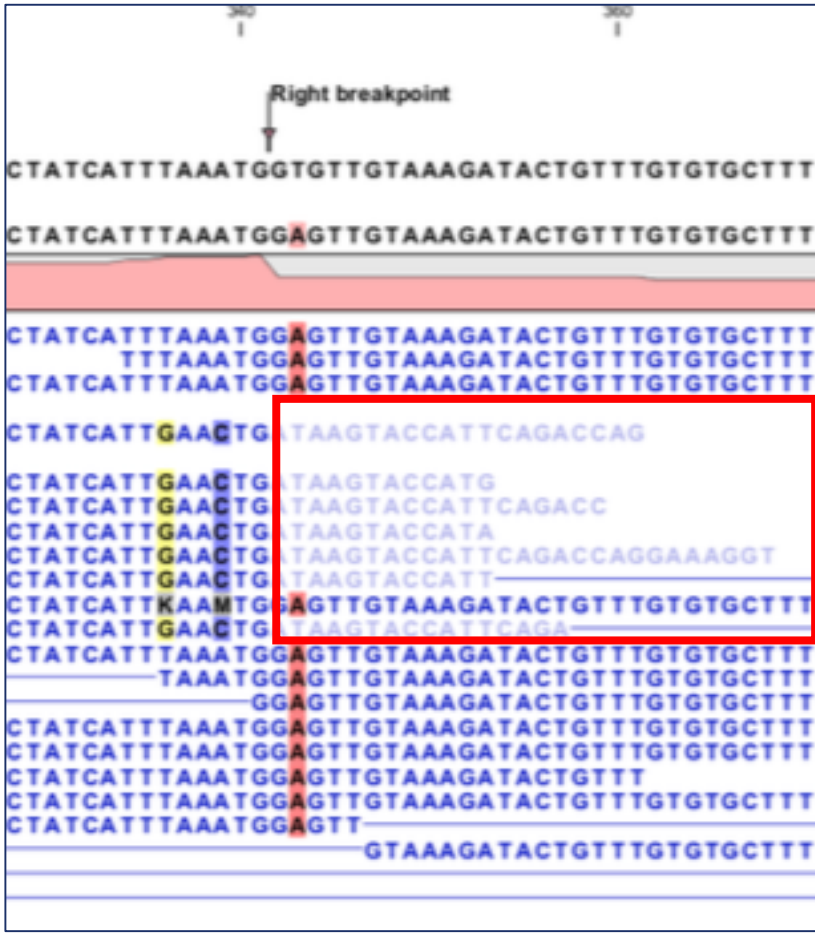
Indel and Structural Variants

- Identify <u>insertions, deletions [Indels], inversions, translocations and tandem duplications [Structural] from read mappings.</u>
- Relies on unmapped read ends
- The Algorithm
  - First identifies positions with an excess of reads with left (or right) unaligned ends.
    - Determines consensus sequences of the unaligned ends
    - Maps the consensus sequences to the reference sequence around other positions with unaligned ends.
    - Structural variant is called when mappings are in accordance with a 'signature' of a structural variant.



Indel and Structural Variants

- Identify insertions, deletions [Indels], inversions, translocations and tandem duplications [Structural] from read mappings.
- Relies on unmapped read ends
- The Algorithm
  - First identifies positions with an excess of reads with left (or right) unaligned ends.
  - Determines consensus sequences of the unaligned ends
  - Maps the consensus sequences to the reference sequence around other positions with unaligned ends.
  - Structural variant is called when mappings are in accordance with a 'signature' of a structural variant.



Indel and Structural Variants

- Identify insertions, deletions [Indels], inversions, translocations and tandem duplications [Structural] from read mappings.

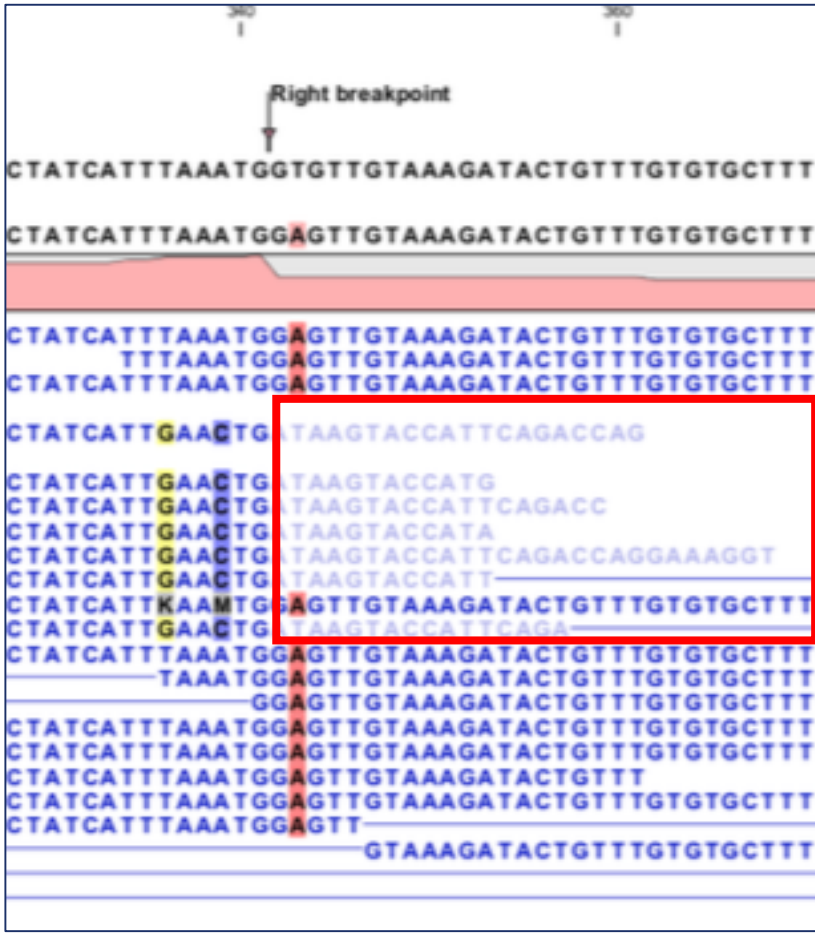
- Relies on unmapped read ends

- The Algorithm
  - First identifies positions with an excess of reads with left (or right) unaligned ends.

  - Determines consensus sequences of the unaligned ends

  - **Maps** the consensus sequences to the reference sequence around other positions with unaligned ends.

  - Structural variant is called when mappings are in accordance with a 'signature' of a structural variant.



Indel and Structural Variants

- Identify insertions, deletions [Indels], inversions, translocations and tandem duplications [Structural] from read mappings.

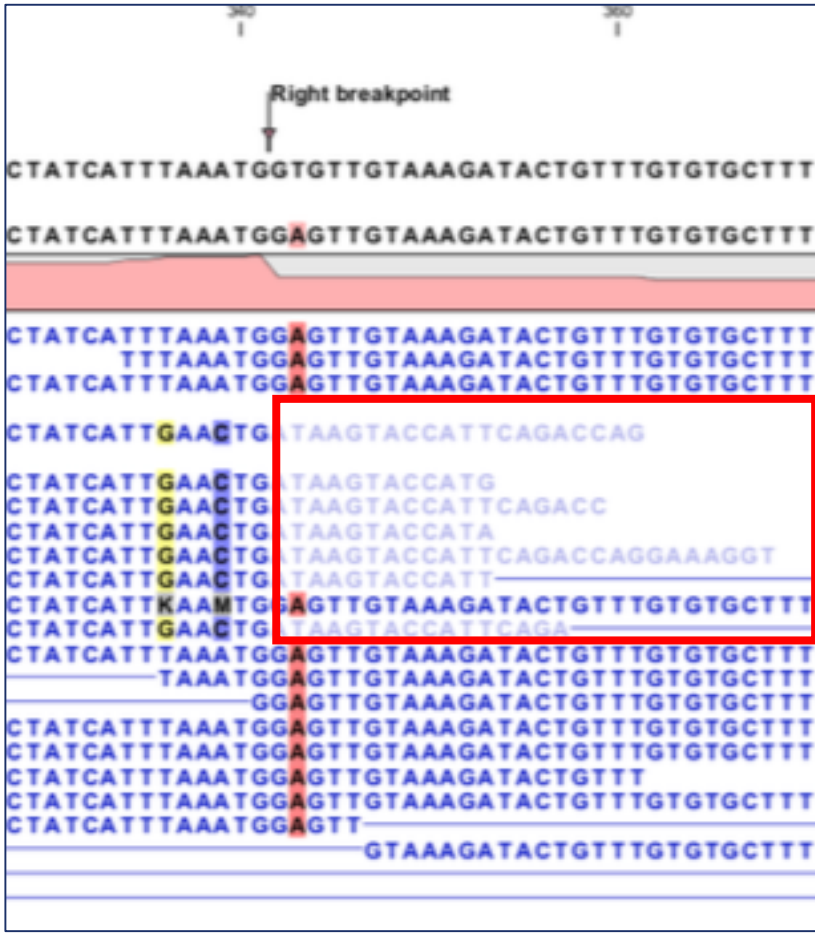
- Relies on unmapped read ends

- The Algorithm
  - First identifies positions with an excess of reads with left (or right) unaligned ends.

  - Determines consensus sequences of the unaligned ends

  - Maps the consensus sequences to the reference sequence around other positions with unaligned ends.

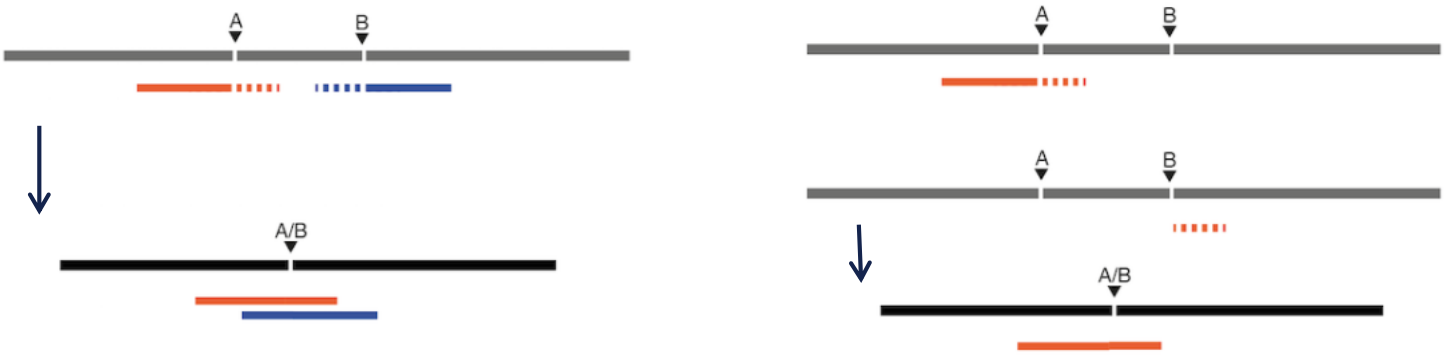
  - Structural variant is called when mappings are in accordance with a **'signature' of a structural variant**.



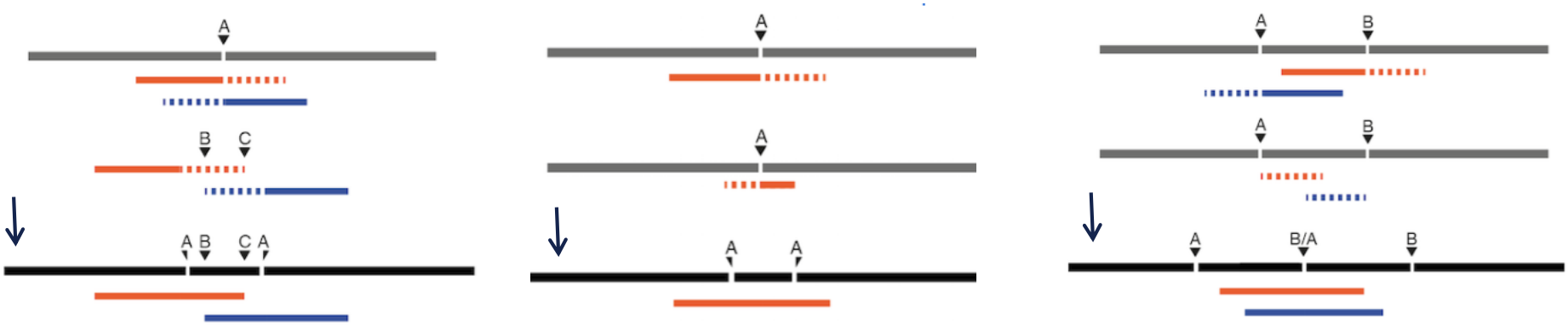
Indel and Structural Variants- The Algorithm

Structural variant signatures:

Deletions: Unaligned sequence in between breakpoints



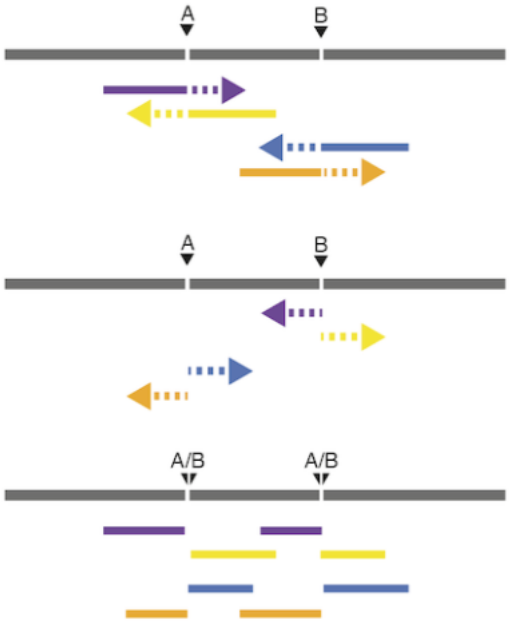
Insertions: Unaligned sequence on either side of a breakpoint



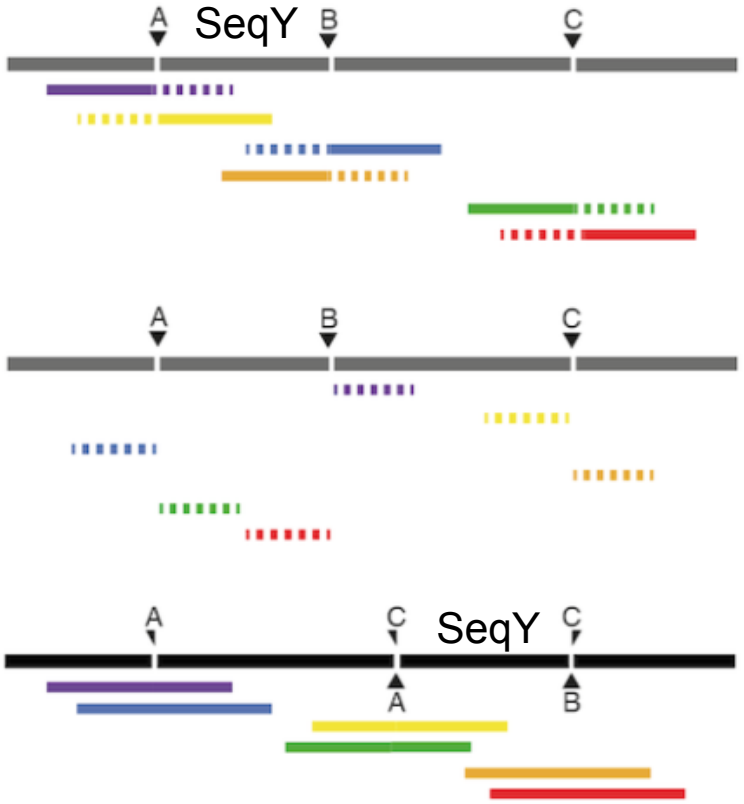
Indel and Structural Variants- The Algorithm

Structural variant signatures:

Inversions



Translocation: Deletion of SeqY between breakpoints A and B and insertion of SeqY at C.



Replacement