



The Elizabeth H.  
and James S. McDonnell III

**McDONNELL  
GENOME INSTITUTE**  
at Washington University

## GenViz Module 5: Variant annotation and interpretation

Malachi Griffith, Obi Griffith, Zachary Skidmore  
Genomic Data Visualization and Interpretation  
September 11-15, 2017  
Berlin



## Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).

### You are free to:

**Share** — copy and redistribute the material in any medium or format

**Adapt** — remix, transform, and build upon the material for any purpose, even commercially.



The licensor cannot revoke these freedoms as long as you follow the license terms.

### Under the following terms:



**Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.



**ShareAlike** — If you remix, transform, or build upon the material, you must distribute your contributions under the [same license](#) as the original.

**No additional restrictions** — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.

# Learning objectives of the course

---

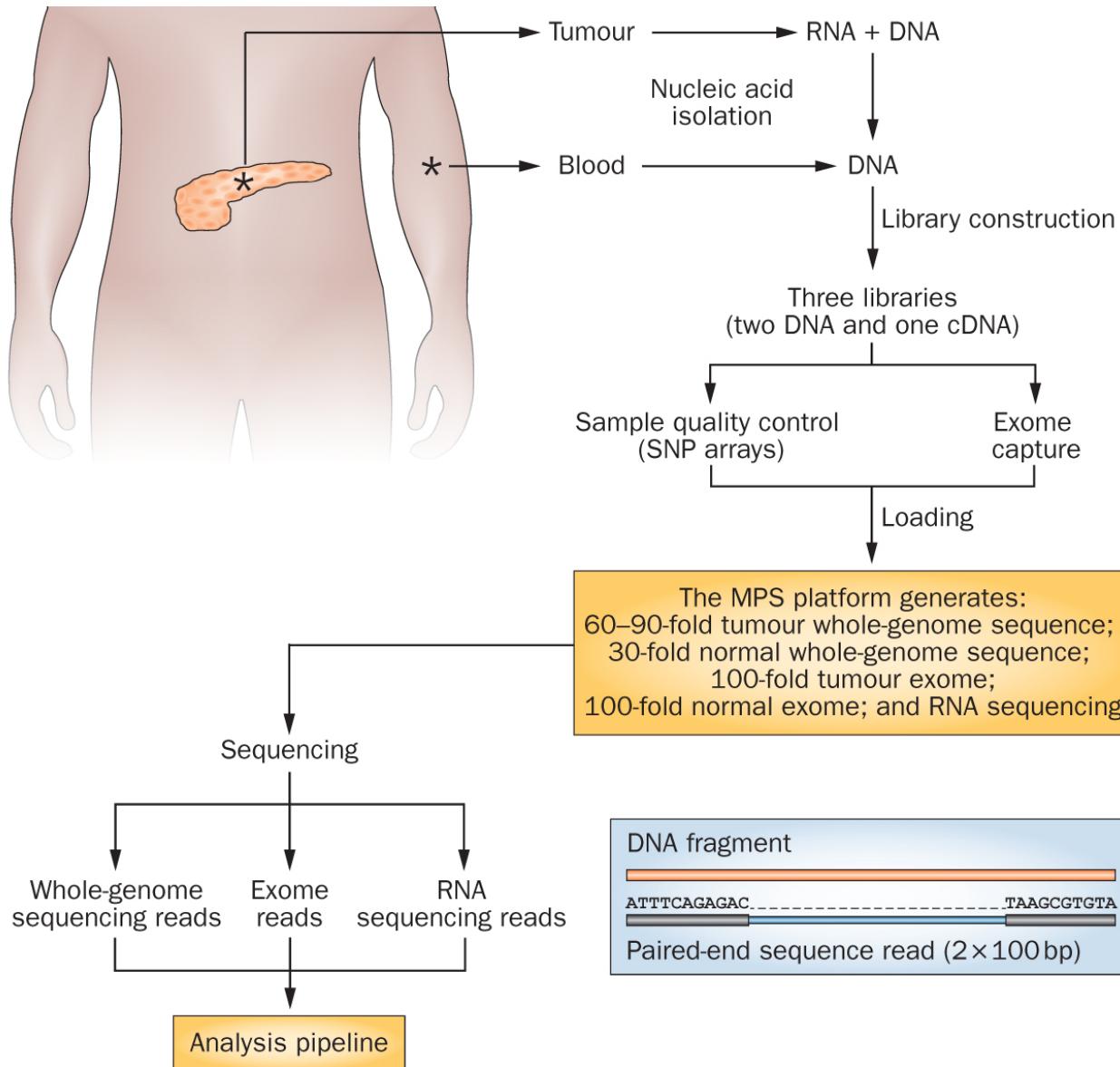
- **Module 1: Introduction to genomic data visualization and interpretation**
- **Module 2: Using R for genomic data visualization and interpretation**
- **Module 3: Introduction to GenVisR**
- **Module 4: Expression profiling, visualization, and interpretation**
- **Module 5: Variant annotation and interpretation**
- **Module 6: Q & A, discussion, integrated assignments, and working with your own data**
- **Tutorials**
  - Provide working examples of data visualization and interpretation
  - Self contained, self explanatory, portable

# Learning objectives of module 5

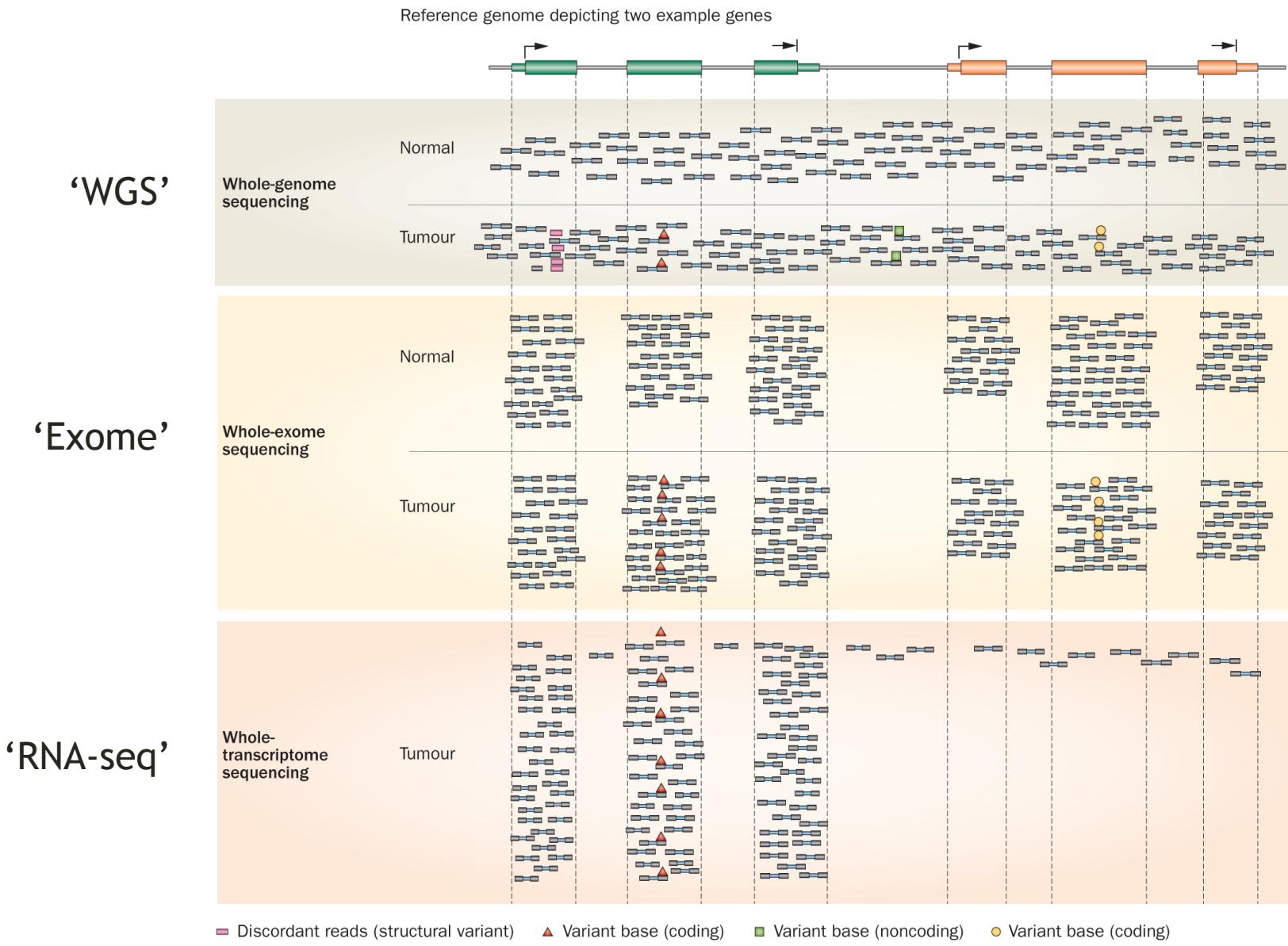
---

- Variant annotation and interpretation
  - Types of variation
  - Key concepts for interpreting variants
  - Variant resources
    - Human disease
    - Cancer
    - Non-human
  - Functional characterization
  - Clinical relevance

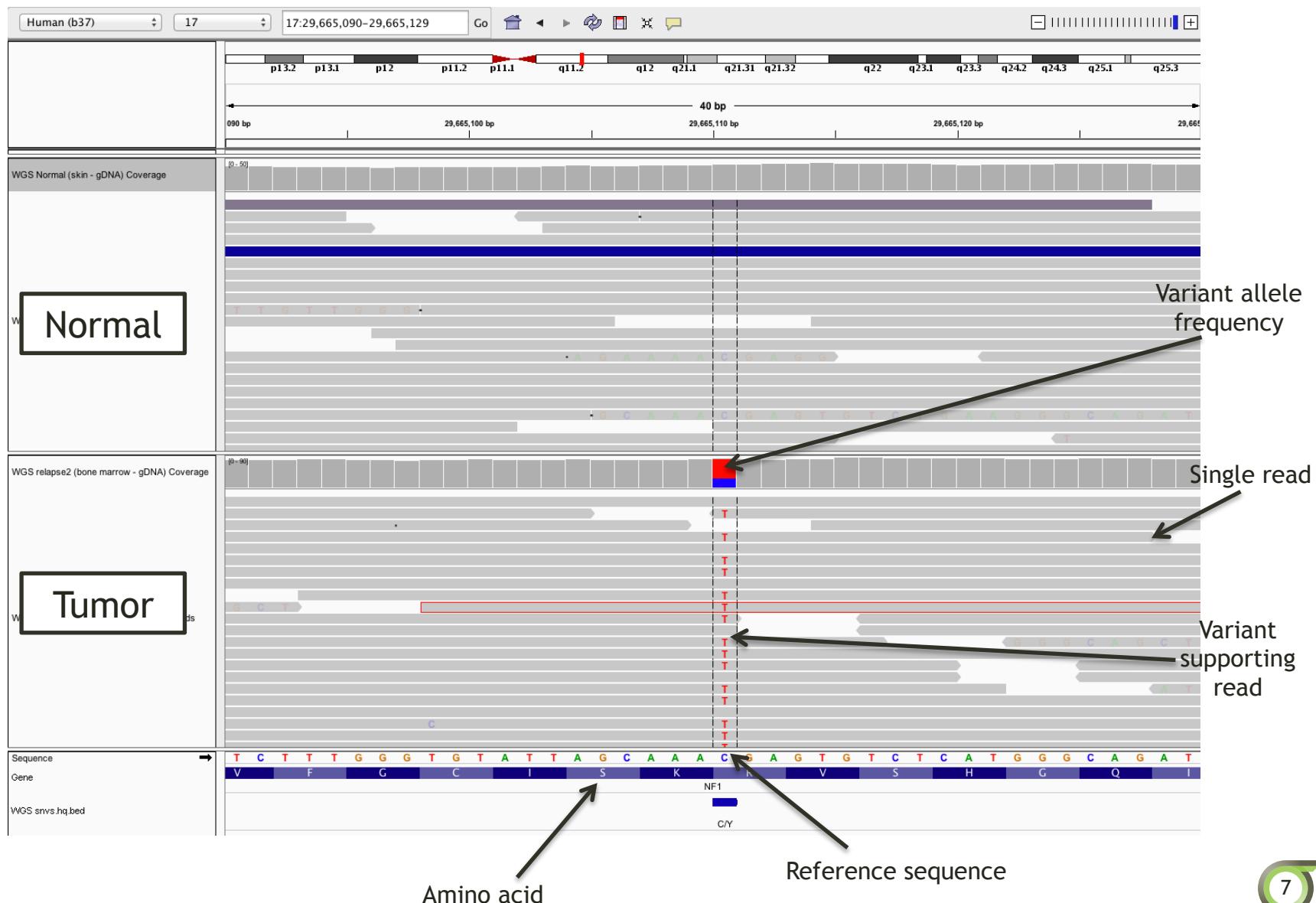
# Cancer genomics research has exploded with the rapid advances in DNA sequencing technologies



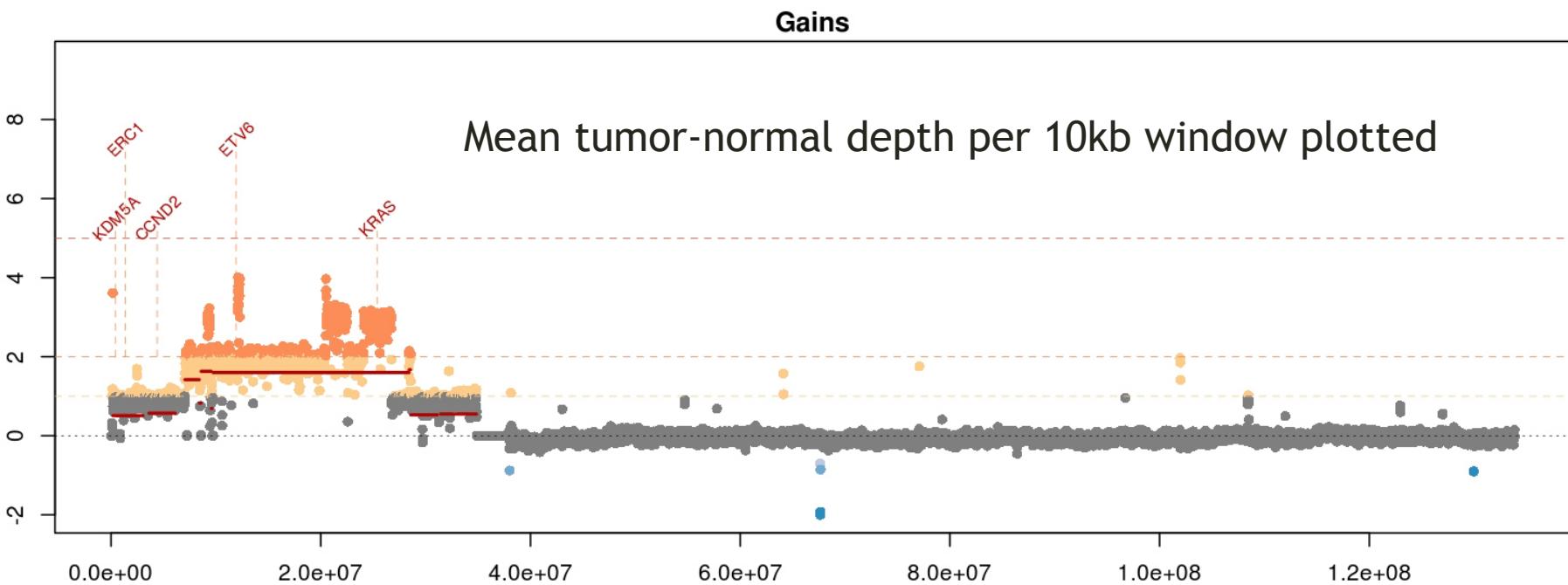
# Whole genome, exome and transcriptome sequencing allows us to detect and confirm many different variant types



# Single nucleotide variants (SNVs) and insertions/deletions (indels) appear as short alignment discrepancies from reference genome



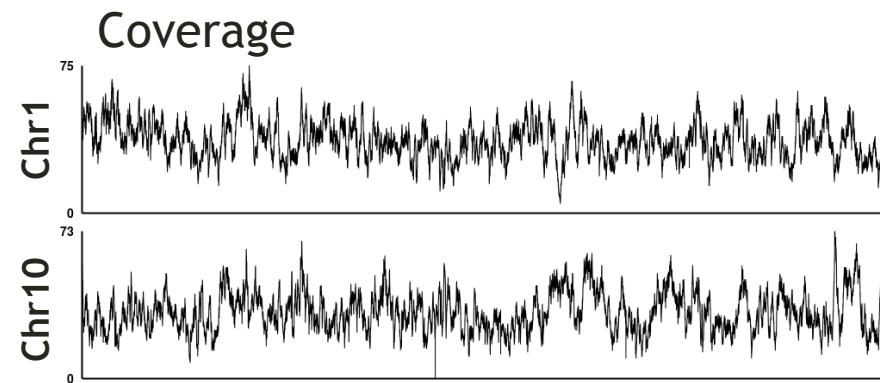
# Copy number variants (CNVs) appear as deviations from in alignment “depth” or “coverage”



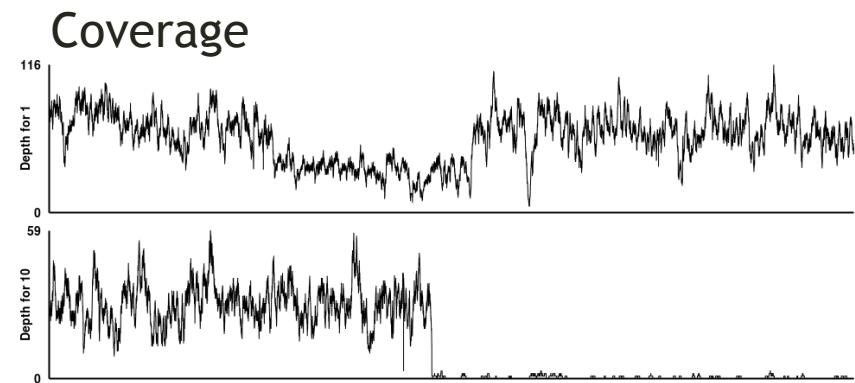
KRAS amplification in a metastatic breast cancer

# Structural variants (SVs) can be identified using a combination of coverage and discordant read alignments

Normal



Tumor



Chr1

1:168243587 1:168284587

Chr10

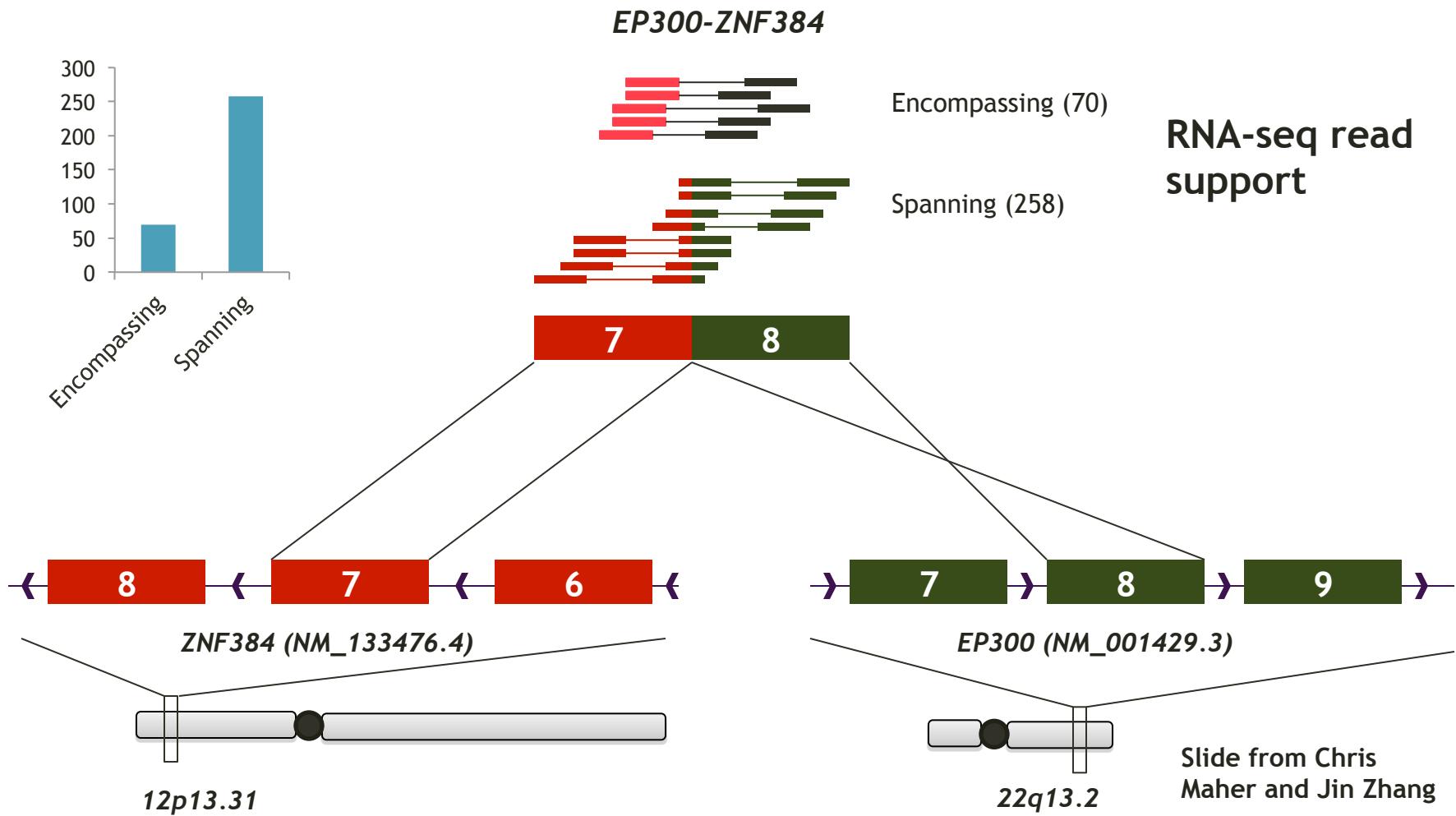
10:104273357 10:104314357 1:168243587 10:104314357

Discordant read support

A Chr1-Chr10 (TBX19-SUFU) unbalanced translocation identified in an adult acute lymphocytic leukemia.

Discordant read support

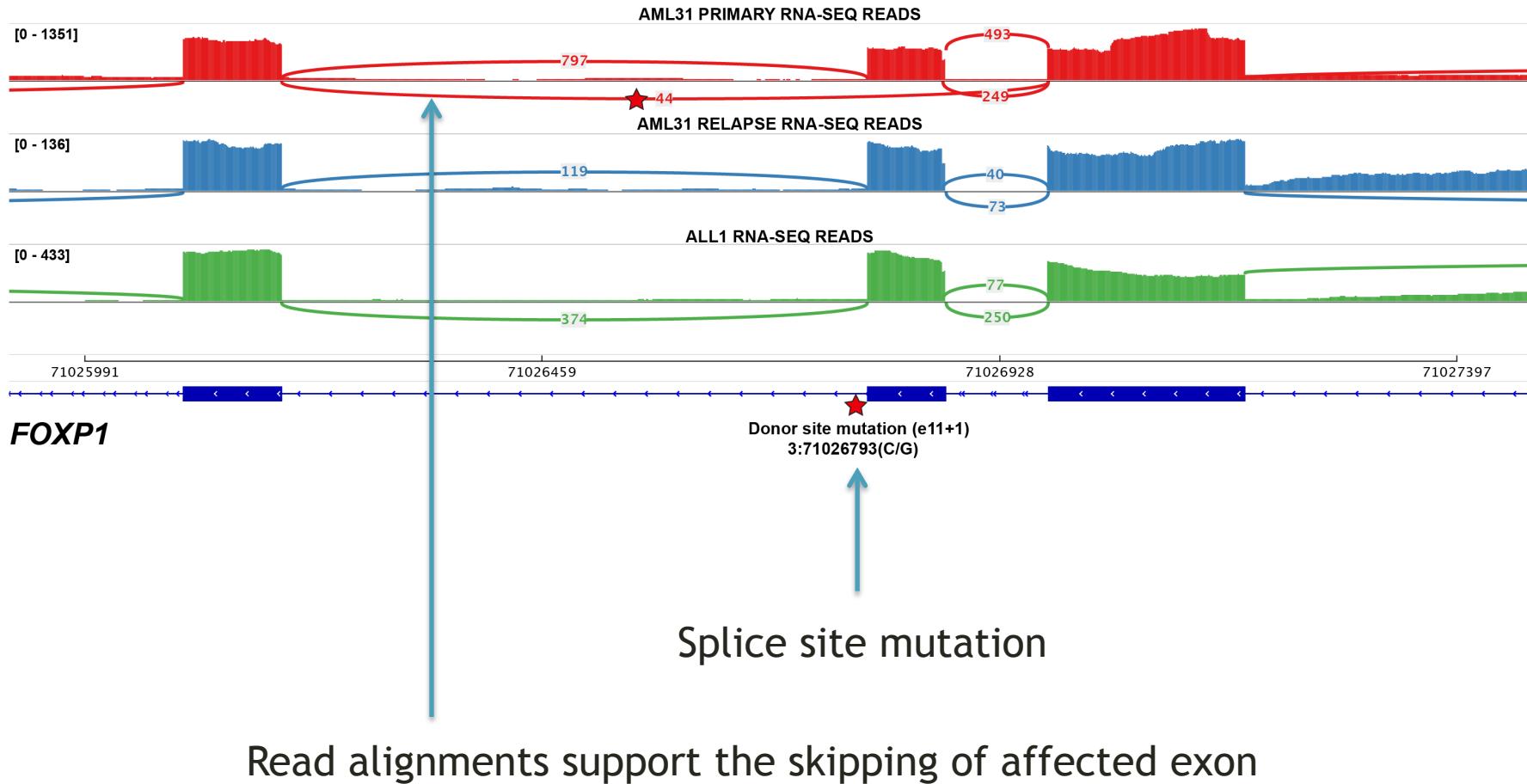
# Expressed gene fusions can be identified by discordant read alignments spanning known exons from RNA-seq data



Exons 1-8 of EP300 fused to exons 7-10 of ZNF384 in head-to-tail fashion.

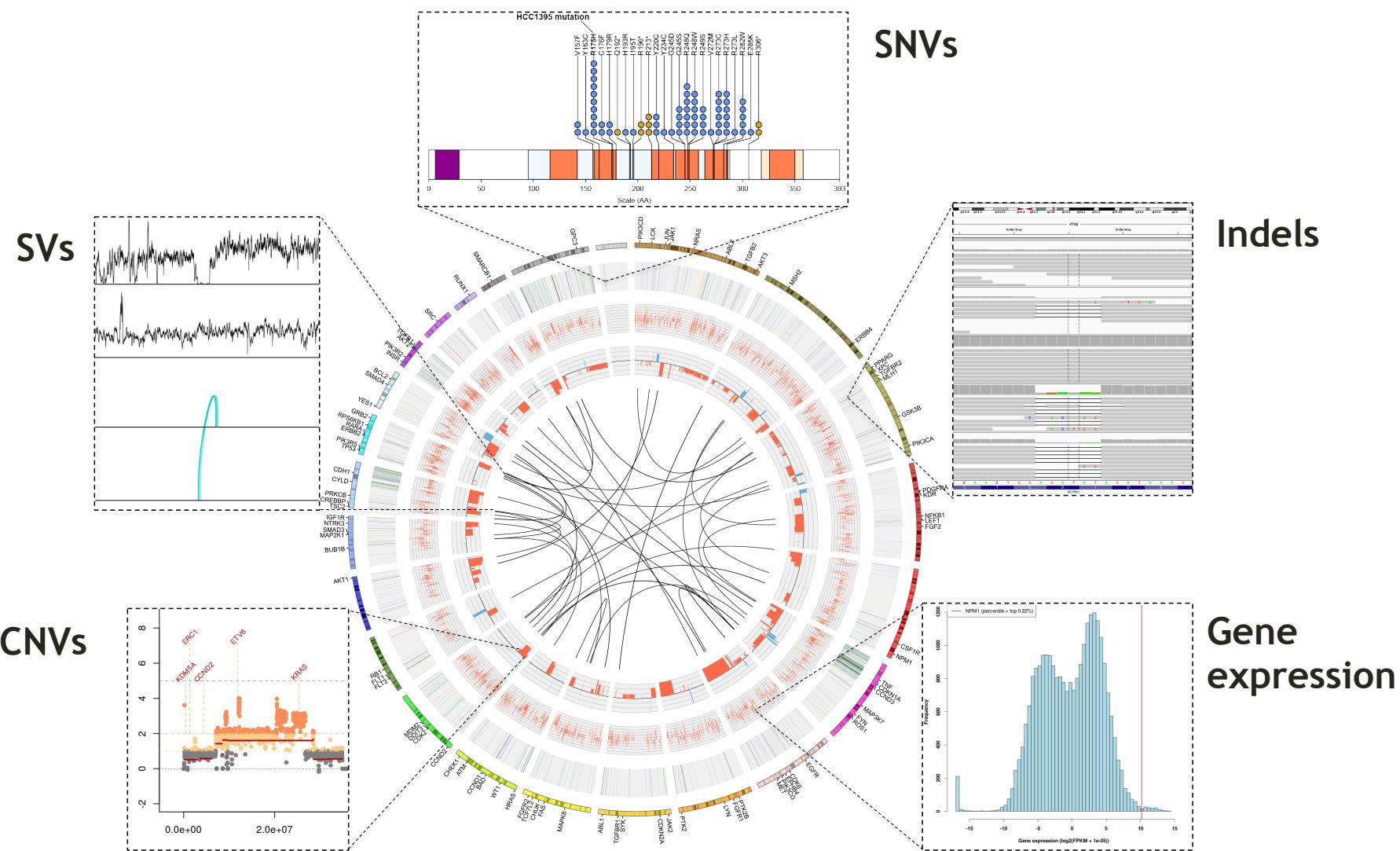
# RNA-seq can also reveal the splicing consequence of somatic mutations detected in WGS

## A. Sub-clonal somatic splicing event in *FOXP1* observed in the primary tumor but cleared in relapse

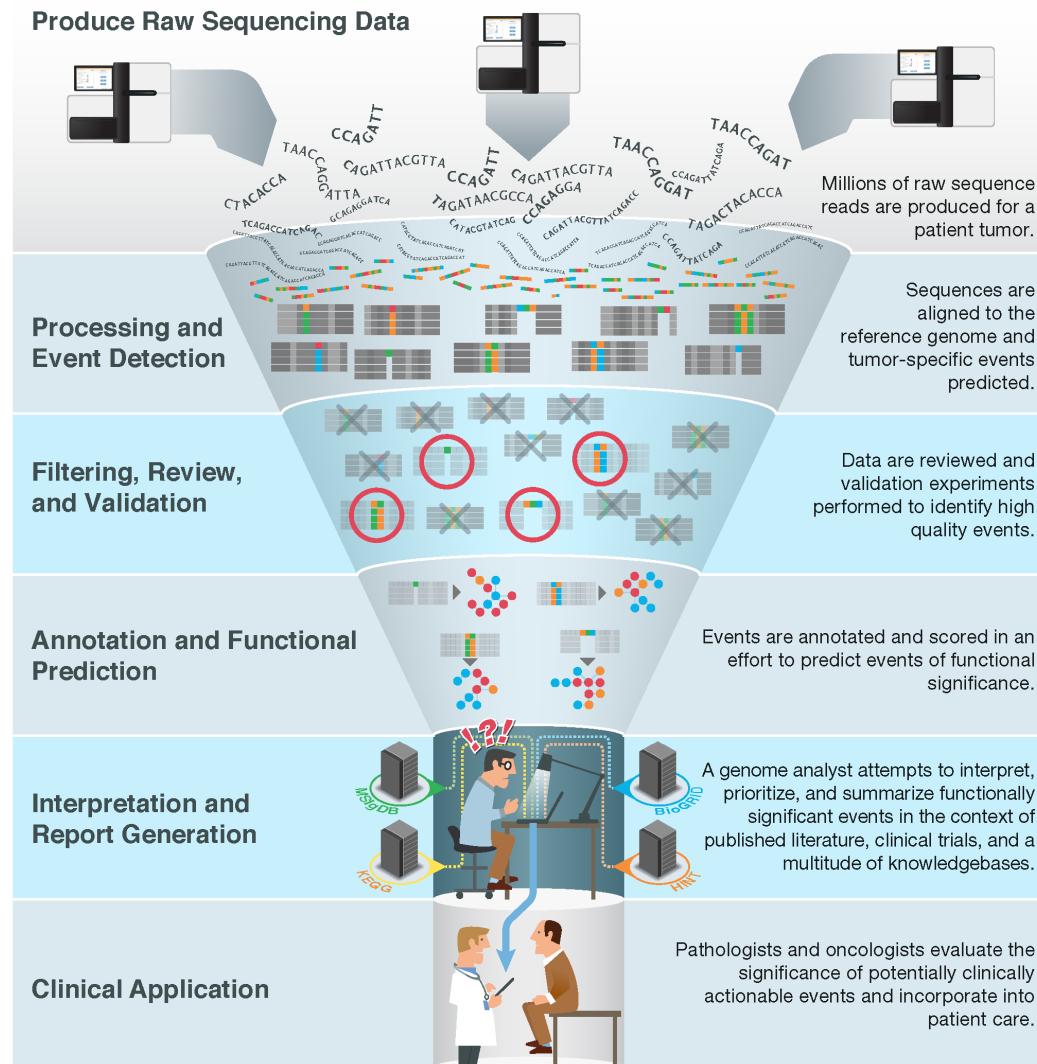


We created '[RegTools](#)' to help characterize these kinds of events

# Tumor genome analysis will typically reveal dozens to thousands of alterations of multiple types

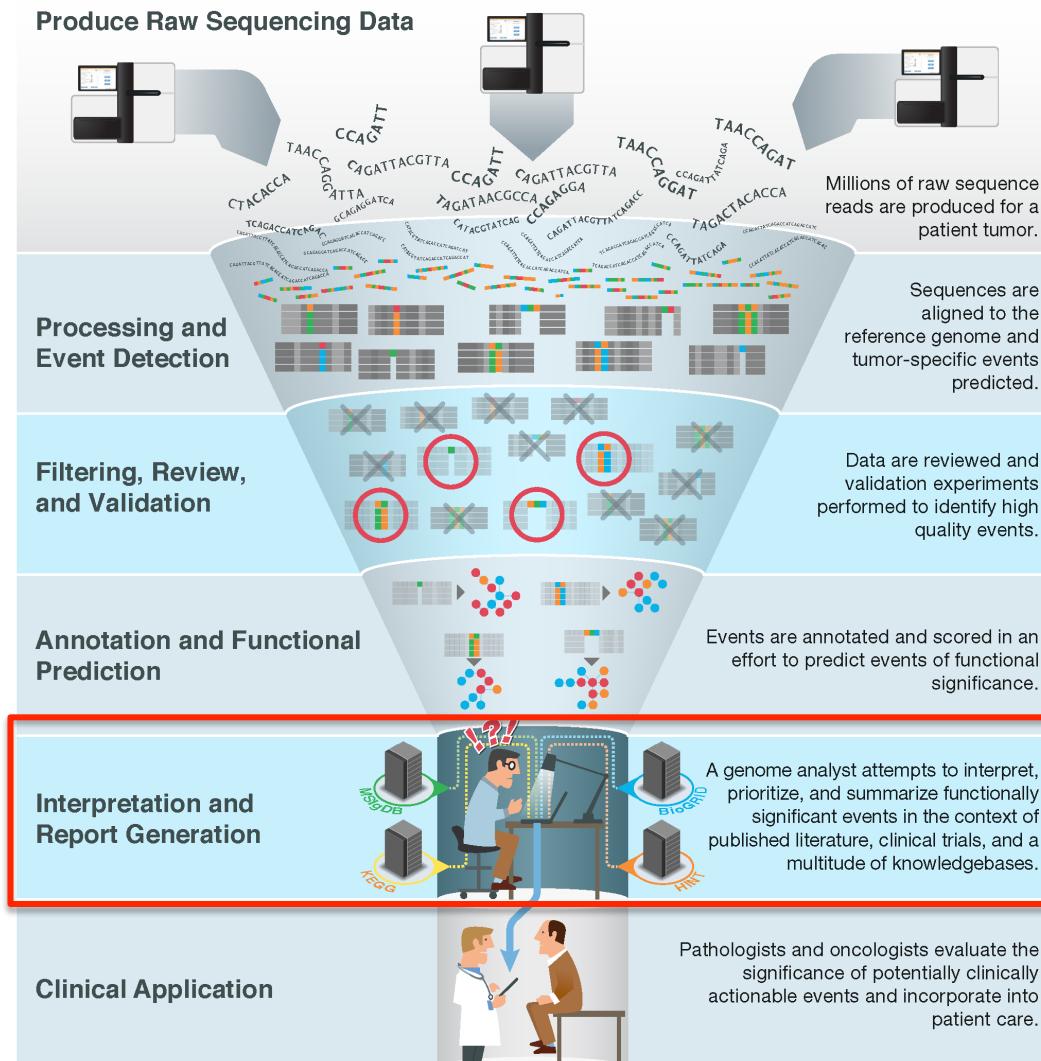


# Despite remaining challenges, most aspects of sequencing and analysis have been standardized and streamlined



Good BM, Ainscough BJ, McMichael JF, Su Al†, Griffith OL†. 2014. Genome Biology. 15(8):438.

# Interpretation of genomic alterations in the context of the clinical relevance remains the major bottleneck



Good BM, Ainscough BJ, McMichael JF, Su Alt†, Griffith OL†. 2014. Genome Biology. 15(8):438.

# Variant analysis/interpretation starts with a raw variant list (VCF file)

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	H_TU-GTB15-3685	H_TU-GTB15-M1501867	
1	1026106	.	G	T	.	PASS	NT=ref;QSS=18;QSS_NT=18;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1216591	.	G	A	.	PASS	NT=ref;QSS=120;QSS_NT=108;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1249123	.	G	T	.	PASS	NT=ref;QSS=16;QSS_NT=16;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1262394	.	G	T	.	PASS	NT=ref;QSS=34;QSS_NT=34;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1326886	.	C	T	.	PASS	NT=ref;QSS=199;QSS_NT=157;SGT=CC->CT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1391597	.	T	C	.	PASS	NT=ref;QSS=32;QSS_NT=32;SGT=TT->CT;TQSS=2;TQSS_NT=2		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1904481	.	G	T	.	PASS	NT=ref;QSS=24;QSS_NT=24;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1912142	.	G	T	.	PASS	NT=ref;QSS=33;QSS_NT=33;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	1919717	.	G	A	.	PASS	NT=ref;QSS=17;QSS_NT=17;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	2319028	.	C	T	.	PASS	NT=ref;QSS=76;QSS_NT=76;SGT=CC->CT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	2333646	.	G	T	.	PASS	NT=ref;QSS=26;QSS_NT=26;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	3328555	.	G	T	.	PASS	NT=ref;QSS=20;QSS_NT=20;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	3350384	.	G	A	.	PASS	NT=ref;QSS=33;QSS_NT=33;SGT=GG->AG;TQSS=2;TQSS_NT=2		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	3388456	.	C	T	.	PASS	NT=ref;QSS=55;QSS_NT=55;SGT=CC->CT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	3662615	.	G	T	.	PASS	NT=ref;QSS=18;QSS_NT=18;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	3774072	.	G	T	.	PASS	NT=ref;QSS=21;QSS_NT=21;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	6021727	.	G	A	.	PASS	NT=ref;QSS=16;QSS_NT=16;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	6271112	.	G	T	.	PASS	NT=ref;QSS=52;QSS_NT=52;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	6278217	.	G	T	.	PASS	NT=ref;QSS=30;QSS_NT=30;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	6609812	.	G	A	.	PASS	NT=ref;QSS=74;QSS_NT=74;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	9338624	.	G	A	.	PASS	NT=ref;QSS=15;QSS_NT=15;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	10678477	.	G	T	.	PASS	NT=ref;QSS=26;QSS_NT=26;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	10720178	.	G	T	.	PASS	NT=ref;QSS=33;QSS_NT=33;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11140620	.	A	C	.	PASS	NT=ref;QSS=20;QSS_NT=20;SGT=AA->AC;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11194363	.	G	T	.	PASS	NT=ref;QSS=19;QSS_NT=19;SGT=GG->GT;TQSS=2;TQSS_NT=2		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11294450	.	C	T	.	PASS	NT=ref;QSS=35;QSS_NT=35;SGT=CC->CT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11561899	.	G	A	.	PASS	NT=ref;QSS=32;QSS_NT=32;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11595041	.	G	A	.	PASS	NT=ref;QSS=137;QSS_NT=105;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11735264	.	G	T	.	PASS	NT=ref;QSS=170;QSS_NT=122;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11852226	.	G	T	.	PASS	NT=ref;QSS=39;QSS_NT=39;SGT=GG->GT;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	11855448	.	G	A	.	PASS	NT=ref;QSS=32;QSS_NT=32;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		
1	12198424	.	G	A	.	PASS	NT=ref;QSS=24;QSS_NT=24;SGT=GG->AG;TQSS=1;TQSS_NT=1		GT:AD:BQ:SS:DP:FDP:SDP:SUBDP:AU:CU:GU: TU:FT		

Details of the VCF file format: [hts-specs](#), [VCF-v4.2.pdf](#)

**How do we interpret these variants?**

# Important concepts in the interpretation of variants

---

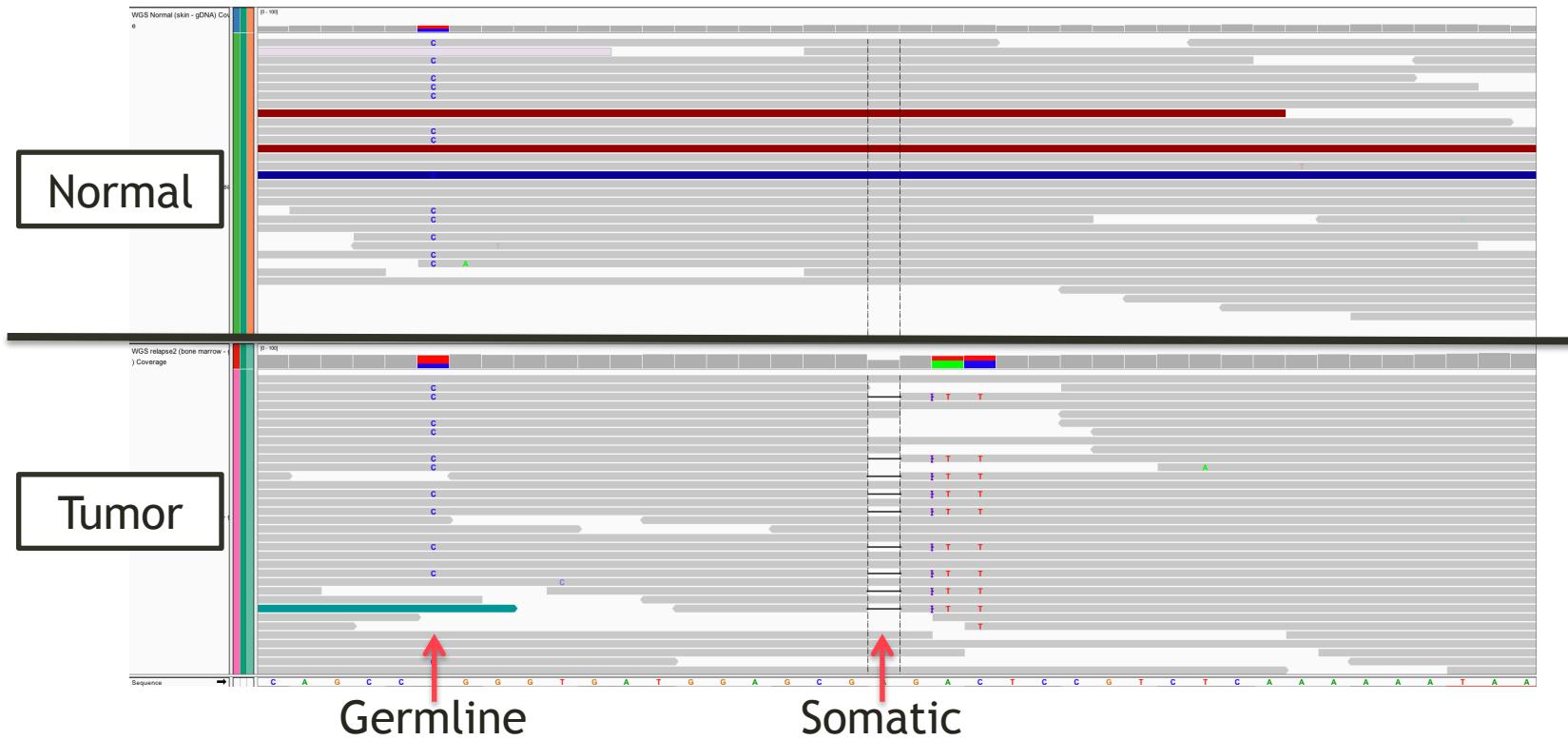
- False positives vs. true positives
- Somatic mutation vs. germline mutation vs. germline polymorphism
- Gain-of-function (activating) vs. loss-of-function (inactivating)
- Deleterious vs. tolerated
- Recurrent vs. random
- Driver vs. passenger
- Dominant clone vs. sub-clonal
- Regulatory vs. coding
- Relevant to cancer biology
  - ‘canonical’ variants, ‘hotspot’ variants, ‘cancer genes’
- Clinically relevant / ‘actionable’
  - ‘Druggable’
  - Predictive, prognostic, diagnostic, predisposing

# Somatic mutation vs. germline mutation vs. germline polymorphism

- Somatic mutations are best distinguished by adequate sequencing of a matched normal
  - Affected and unaffected family members may help to determine origin of a germline mutation
- Comparison of variants to variant databases can also help to classify variants as:
  - Germline polymorphisms
    - [1000 genomes](#)
    - [Exome sequencing project](#) (~6,500 individuals)
    - [ExAC, Exome Aggregation Consortium](#) (~60,000 individuals)
    - [gnomAD browser](#) (123,136 WXS and 15,496 WGS)
  - Germline mutations
    - [OMIM](#), [HGMD](#), [PharmGKB](#), [ClinVar](#)
    - [ACMG guidelines](#)
    - [Gemini](#)
  - Somatic mutations
    - By inference if the mutation is not a common polymorphism (often a weak inference) or is a classic hotspot mutation

# Somatic versus germline demonstration

- Germline mutations
  - Present in egg or sperm
    - All cells of affected offspring
  - Heritable
  - Cause of familial cancers
- Somatic mutations
  - Occur in non-germline tissues
    - Only tumor cells (breast, lung, blood, etc.)
  - Non-heritable
  - Cause of sporadic cancers

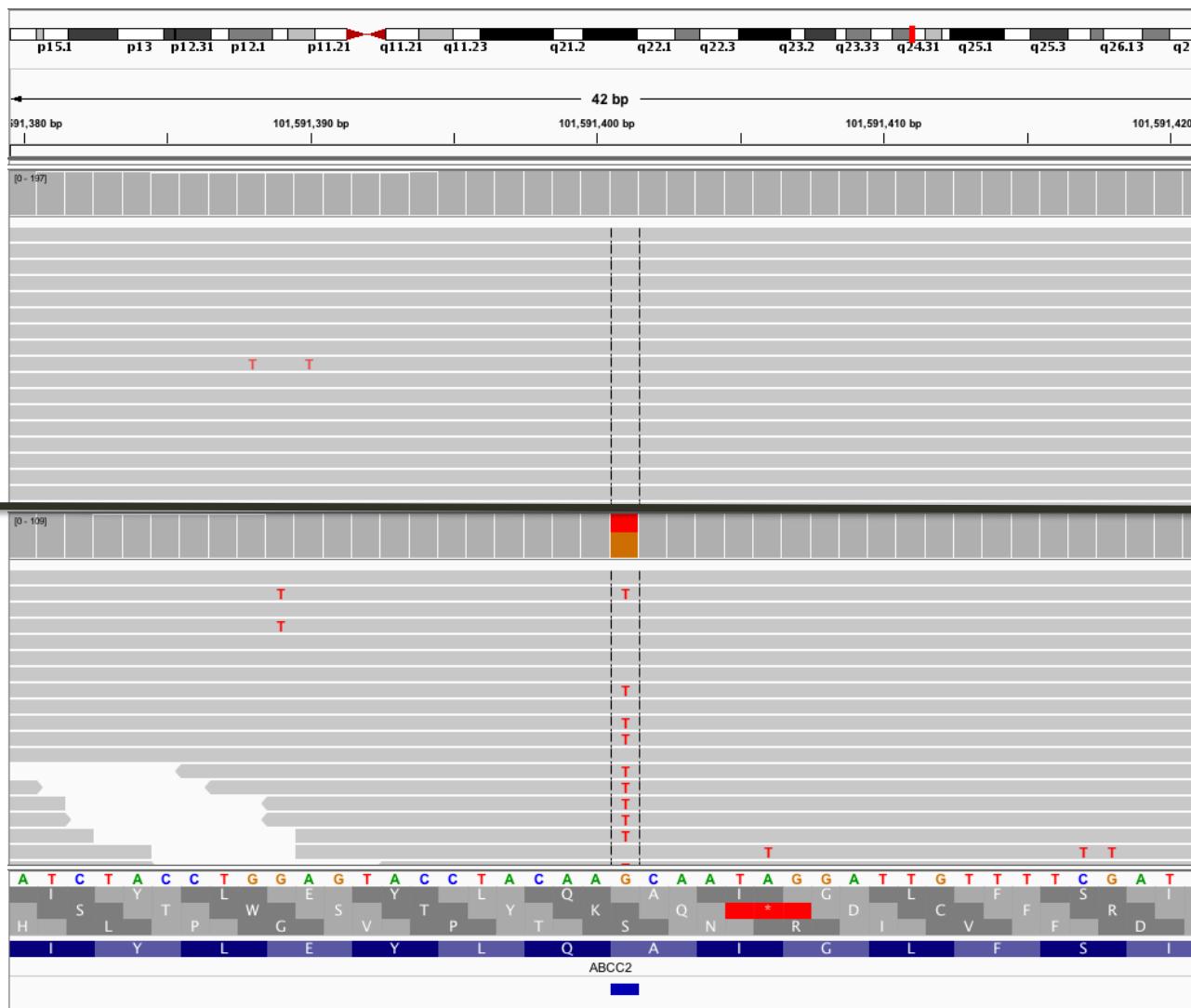


# False positives

---

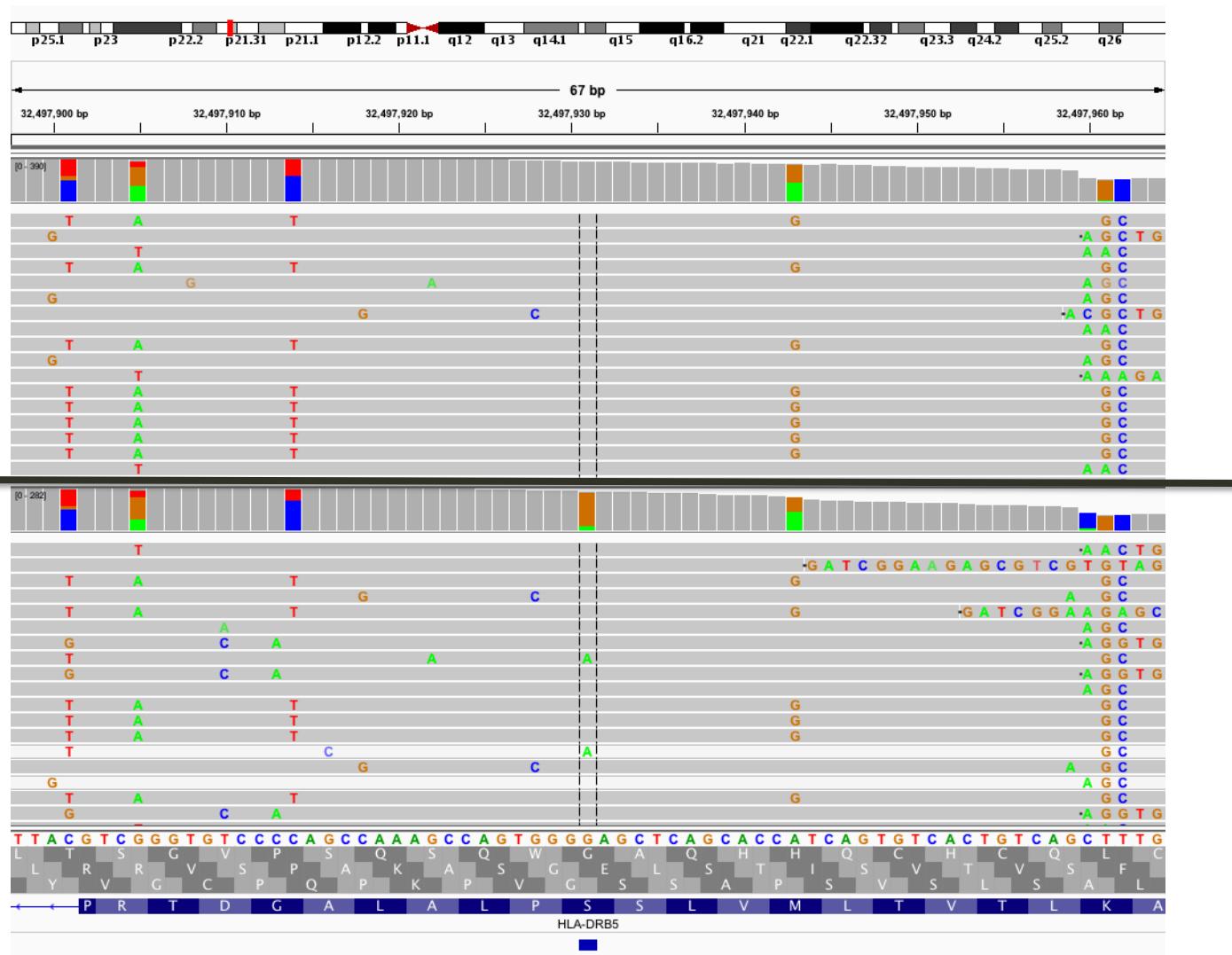
- Use IGV to examine alignments for artifacts
  - Use an intersection of variant callers
  - “panel of normals” analysis
- 
- Useful resources
    - [Discussion of variant callers](#)
    - [Optimizing tumor genome analysis](#)
    - [Hands-on IGV tutorial](#)

# Example of a high quality somatic variant



This G/T variant was independently called by 5 somatic variant callers

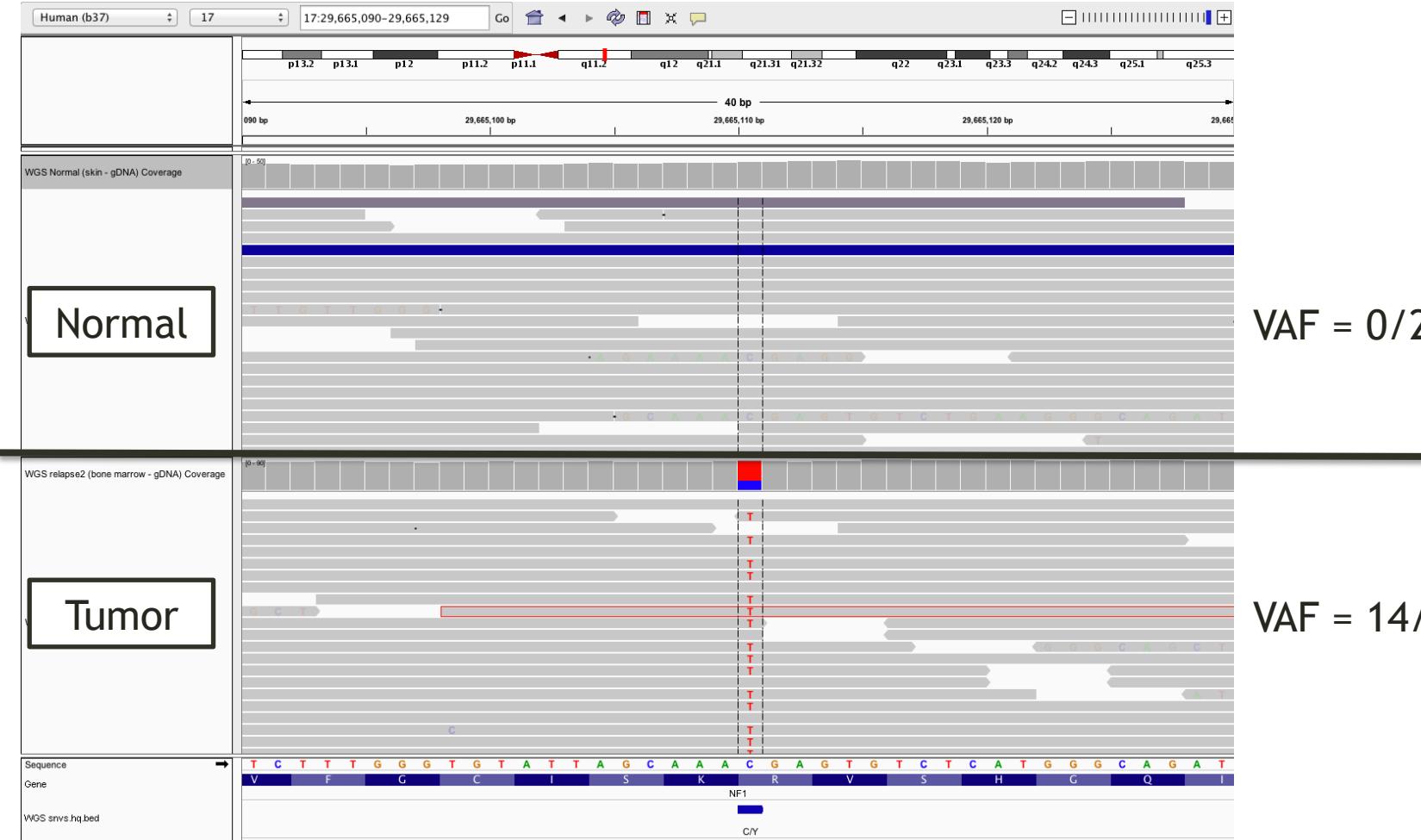
# Example of a low quality somatic variant



This G/A variant was called by only 1 of 5 somatic variant callers

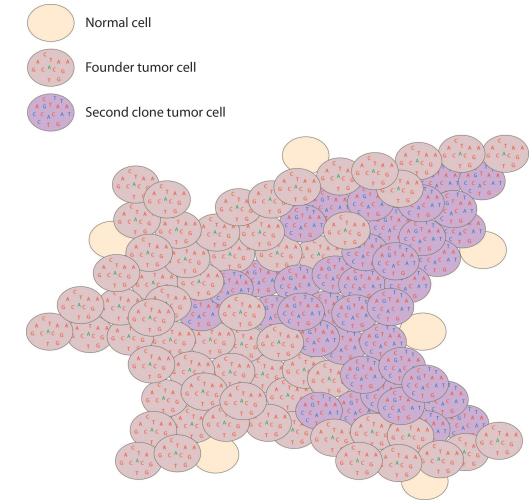
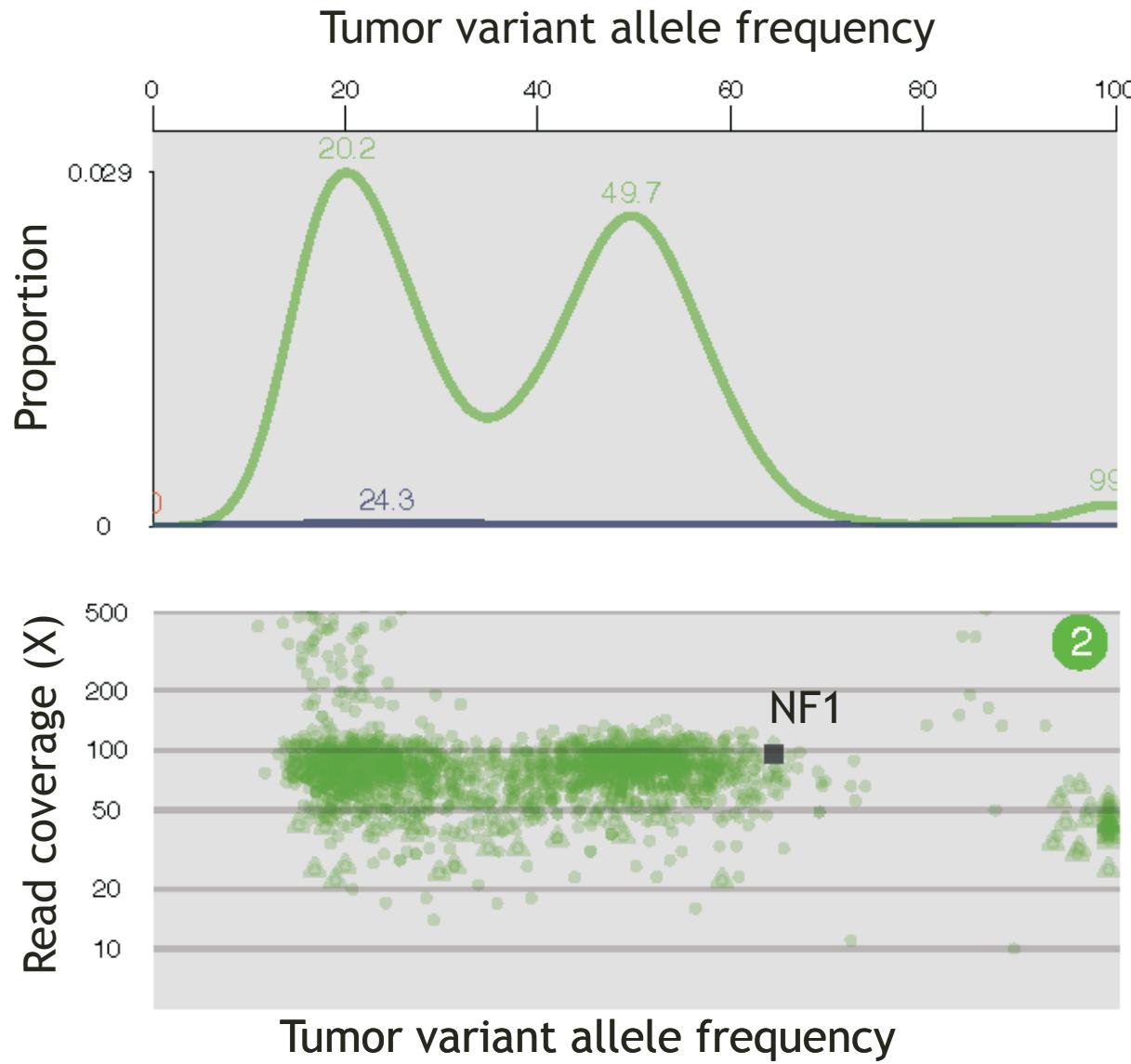
# Variant allele frequency (VAF)

VAF = Variant reads / Total reads



A heterozygous variant is expected to have VAF = 50%. Often not true due to sample purity, tumor heterogeneity, sampling error, alignment issues, copy number variation, etc.

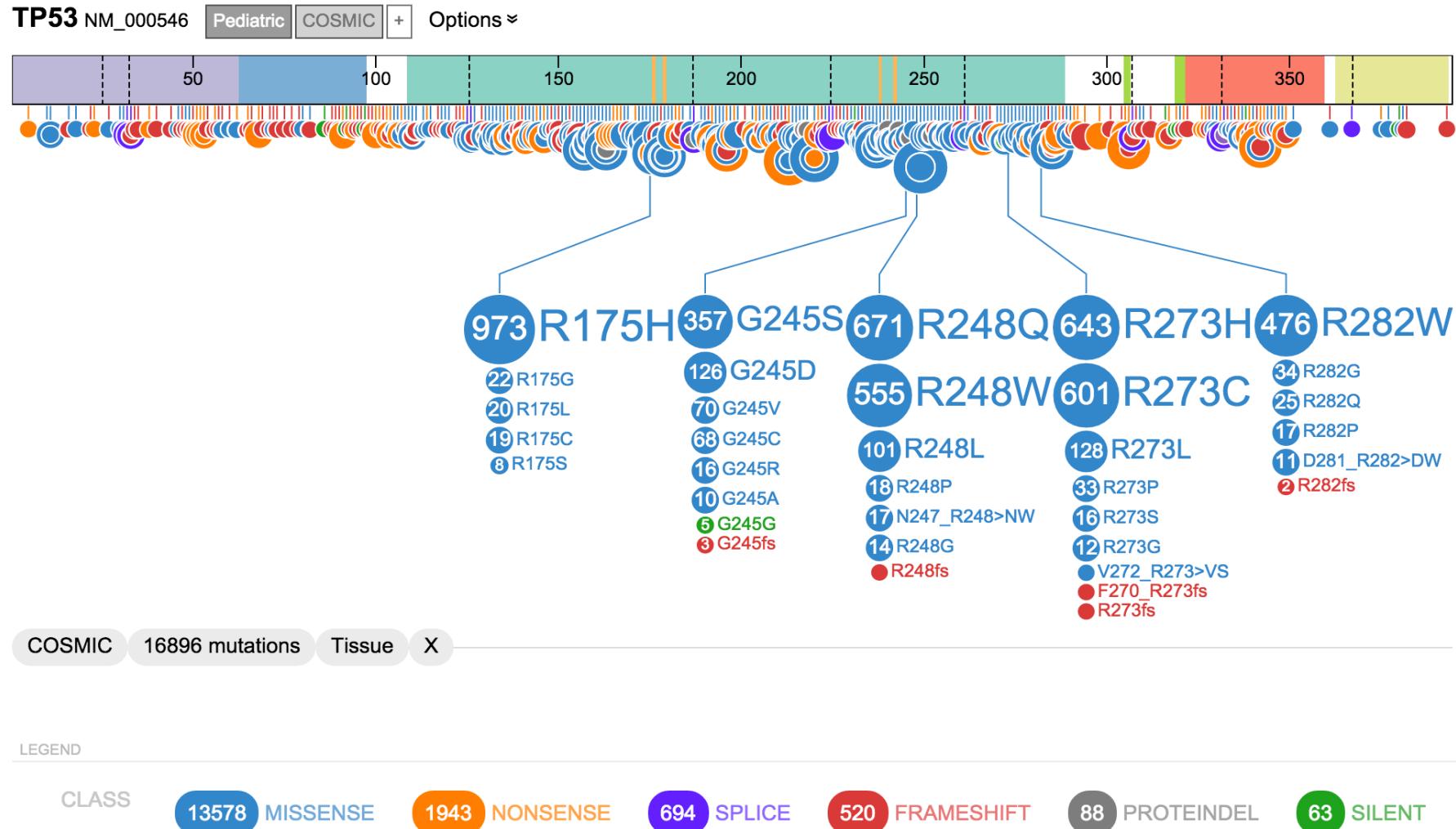
# Dominant clone vs. sub-clonal (and driver vs. passenger)



# Gain-of-function vs. Loss-of-function

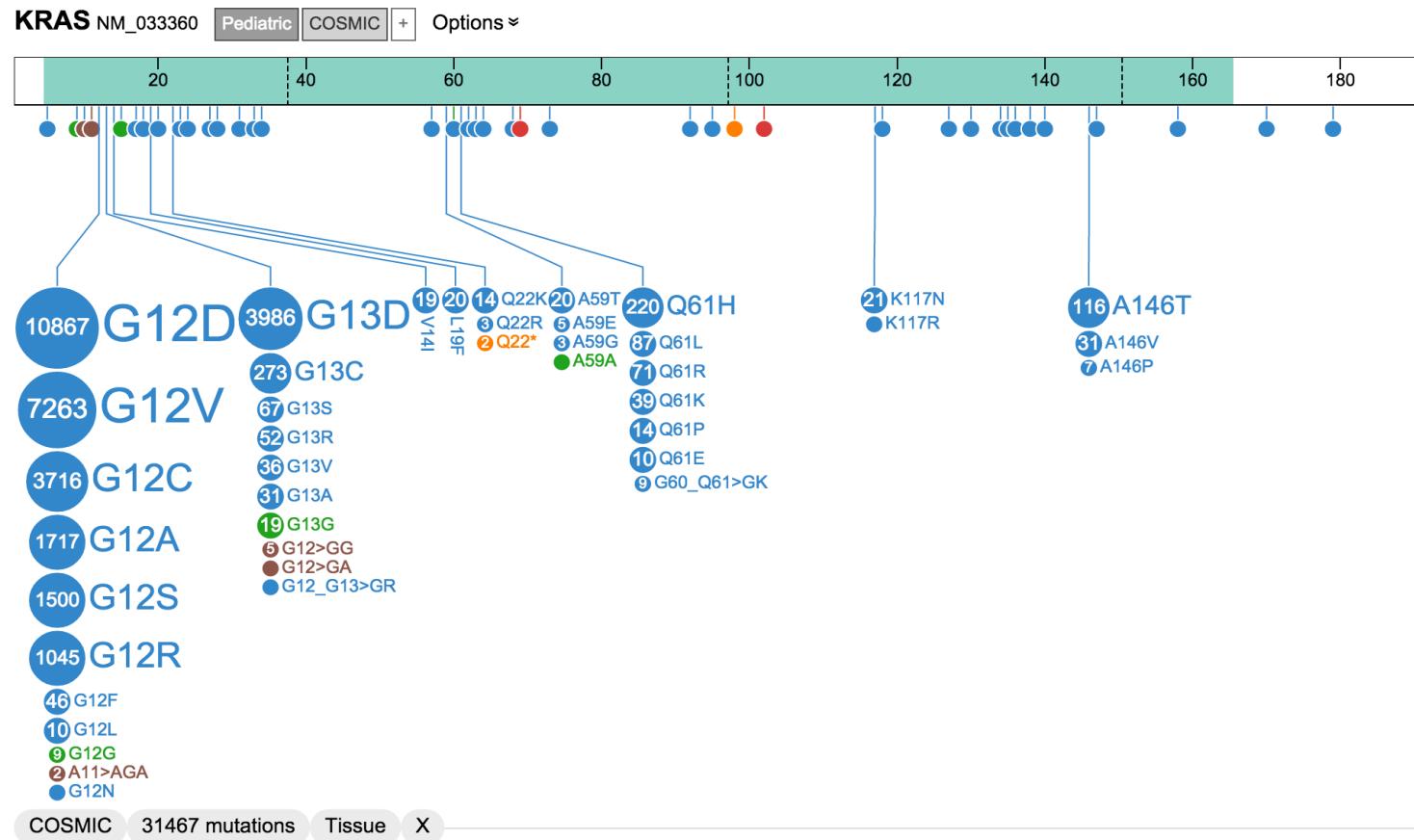
- Gain-of-function and loss-of-function generally refer to a change in specific function of a gene product (protein or RNA) leading to a phenotypic effect
  - Loss-of-function implies that a normal gene function is no longer possible
    - Many random mutations can lead to inactivation
    - Manifests in mutation data as scattered across the length of the gene and may be frameshift, nonsense, and splice site mutations
  - Gain-of-function implies an increased or new gene function
    - There are far fewer ways to create an activating mutation
    - These tend to have dominant phenotypes
    - Manifest in mutation data as “hotspots”
    - Usually missense mutations
- The pattern of **\*recurrence\*** can be a powerful hint for distinguishing activating vs. inactivating mutations.

# Gain-of-function vs. Loss-of-function



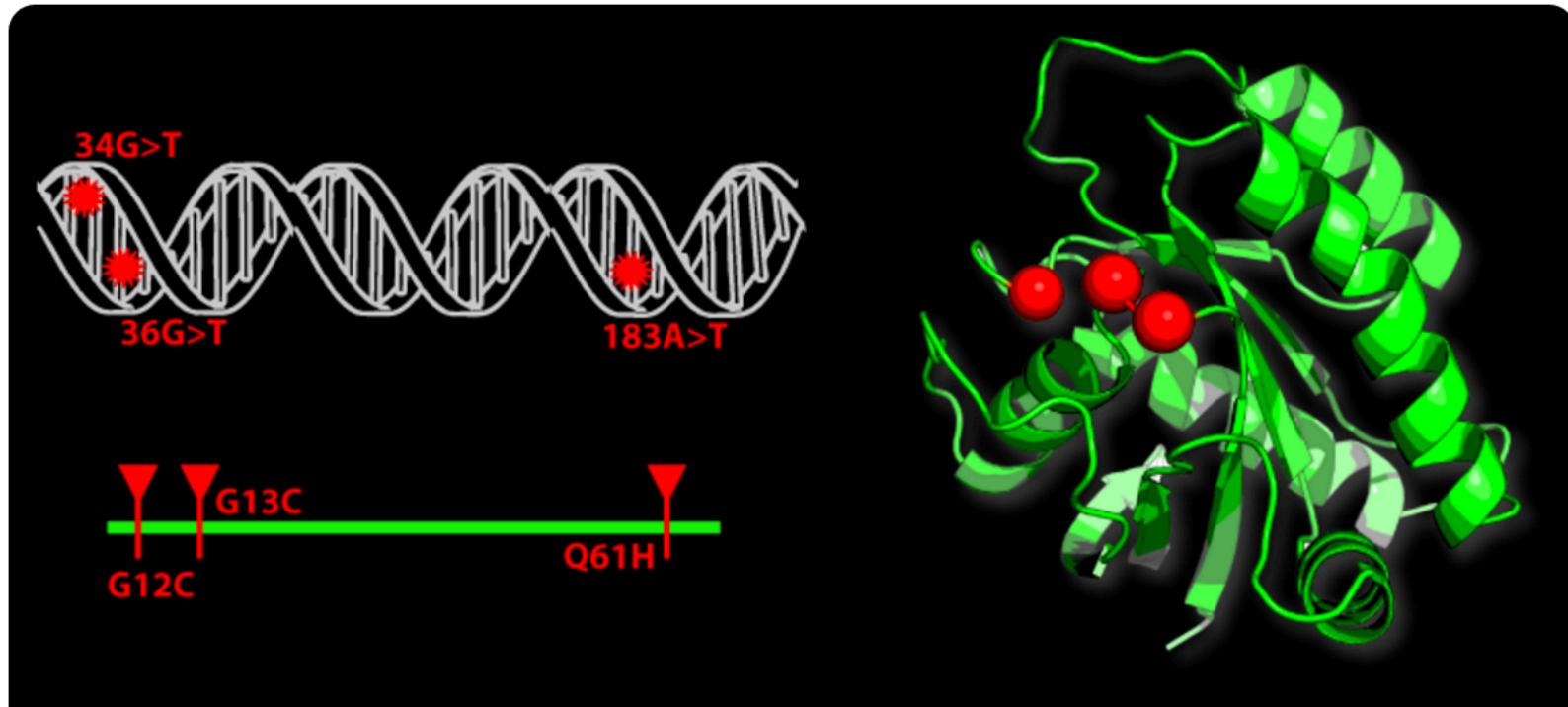
TP53: Recurrence pattern and mutation types suggest loss-of-function

# Gain-of-function vs. Loss-of-function



KRAS: Recurrence pattern and mutation types suggest gain-of-function

# Gain-of-function “hotspots” may be missed in 2D but seen in 3D models of protein structure



From: [Mutation3D](#) website.

For a more detailed discussion of relevant tools:

[Finding Mutation Hotspot At Level Of Amino Acid By Spatial Proximity In Protein Structures](#)

# Recurrent vs. random/background mutation

---

- Recurrent mutation of a gene or pathway can imply functional relevance of that gene/pathway
- Particular somatic mutations that are recurrent (e.g. BRAF V600E) can imply an activating mutation
- One of the most basic (but fruitful) goals of large scale tumor genome sequencing projects has been to look for these patterns of recurrence
  - Since random/background mutation rates can be very high in some tumors, determining whether an observed level/pattern of recurrence is **significant** is important
  - Overall mutation burden, gene size, pathway size, systematic artifacts, and other factors complicate this goal

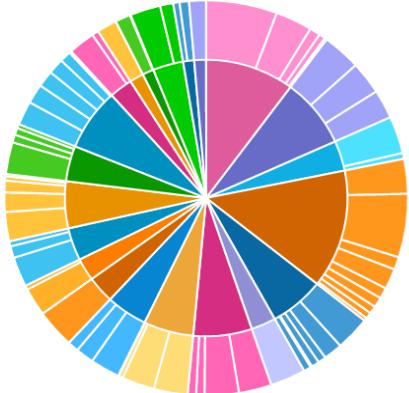
# Human disease variation resources and initiatives

---

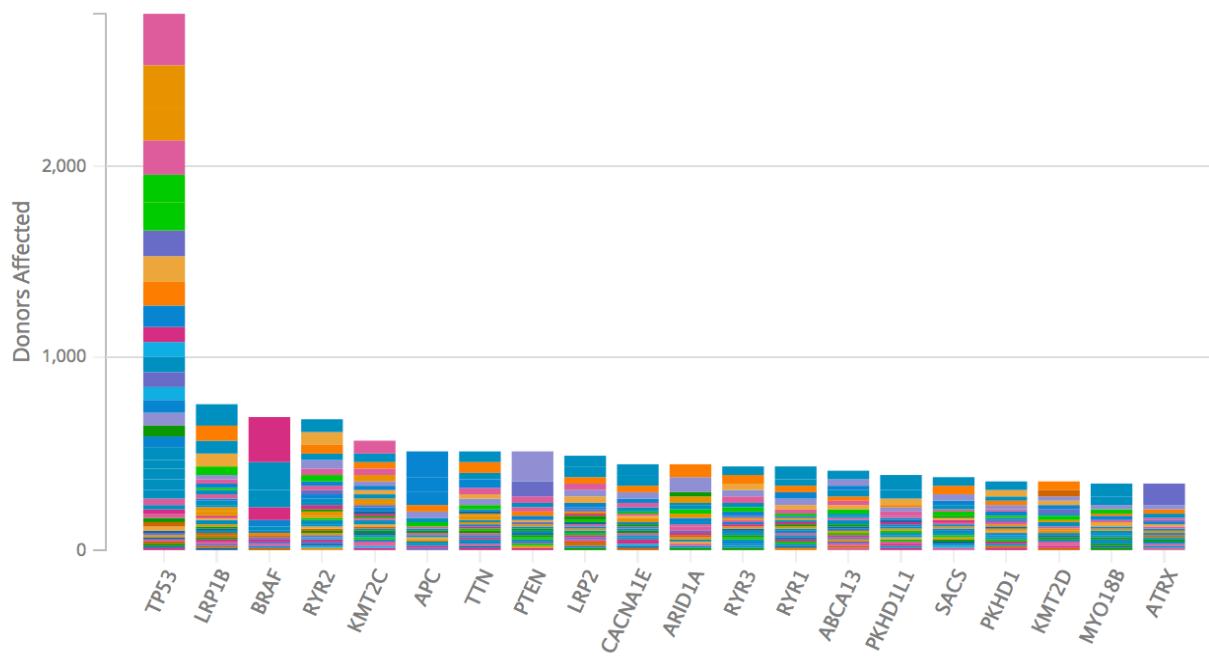
- [Gemini](#)
- [OMIM](#)
- [HGMD](#)
- [dbSNP](#)
- [ExaC](#)
- [GNOMAD](#)
- [Varsome](#)
- [Allele Registry](#)
- [MyVariant.info](#)
- [Ensembl](#)
- [UCSC](#)
- Many more ...

# Cancer variation resources: TCGA and ICGC have sequenced 10,000s of exomes, 1,000s of whole genomes

Donor Distribution  
19,305 Donors across 70 Projects



Top 20 Mutated Genes with High Functional Impact SSMs  
10,648 Unique SSM-Tested Donors



<https://dcc.icgc.org/>

# Exploring cancer mutation data portals

---

- Many, many resources
- Some of the resources that we find most useful
  - [Genomic data commons](#)
  - [ICGC data portal](#)
  - [cBioPortal](#) ([OncoPrinter](#) & [MutationMapper](#))
  - [Cosmic](#)
  - [TCGA data portal](#)
  - [St. Jude pediatric cancer portal](#) ([ProteinPaint](#))
- We track (blog about) these resources as they develop (on BioStars [here](#)).

# Non-human variation resources

---

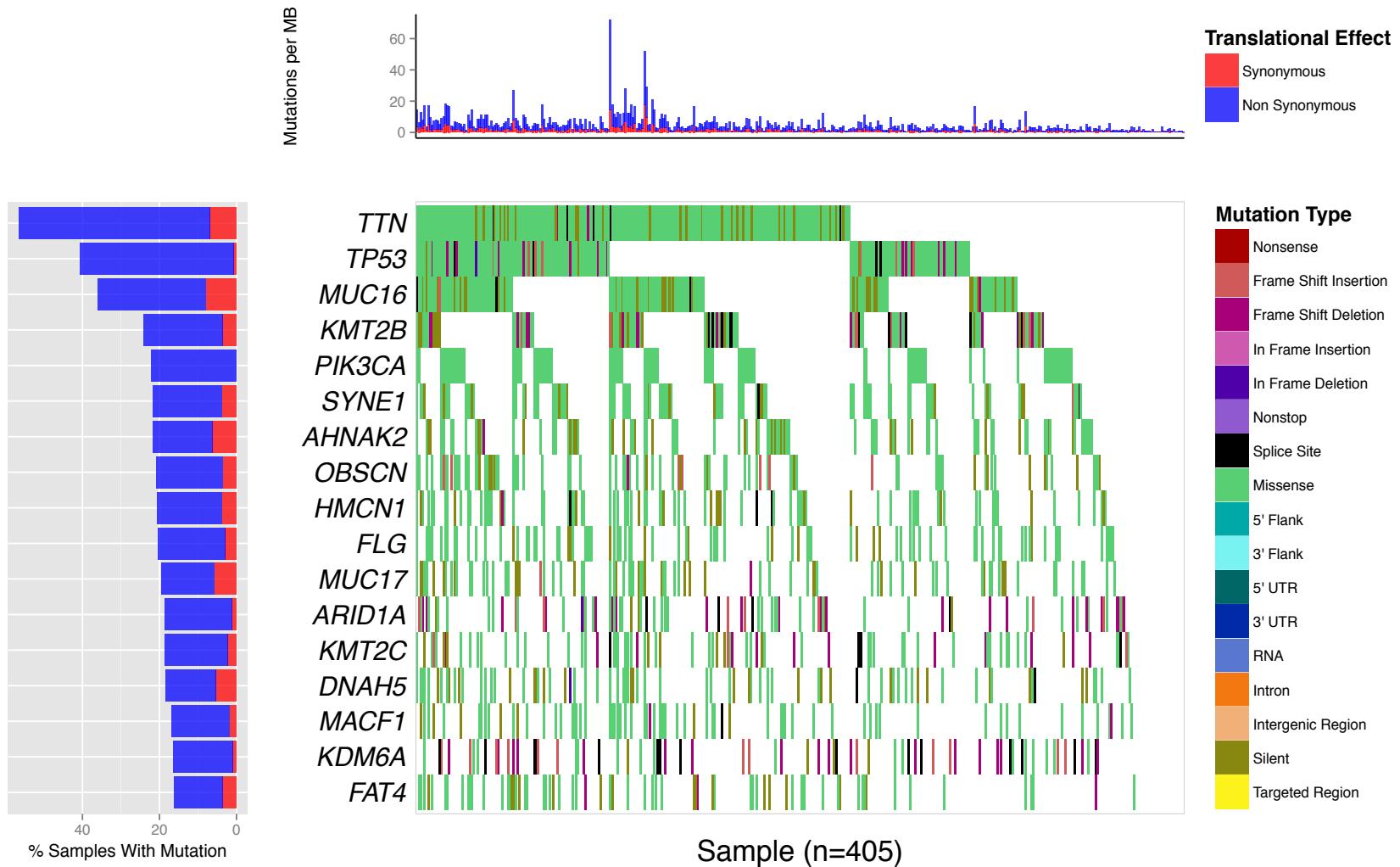
- [European Variation Archive \(EVA\)](#)
- [Online Mendelian Inheritance in Animals \(OMIA\)](#)
- Species specific resources
  - [FlyBase](#)
  - [Mouse Genome Informatics](#)
  - [Rat Genome Database](#)
  - [Zebrafish Information Network](#)

# Deleterious vs. tolerated (functional vs. non-functional)

---

- Many tools/resources attempt to classify variants as “deleterious” vs. “tolerated” (aka “benign”)
  - E.g. Sift, PolyPhen, Condel, [CADD](#), etc.
- The goal of these tools is often confused with predicting whether the mutation is gain-of-function/activating
  - Not the same thing...
  - Mostly driven by sequence conservation (though CADD considers a more complex set of features)
- Variant effect annotators
  - VEP, snpEff, ANNOVAR, VAAST

# A ‘waterfall’ plot is one way to visualize the pattern of recurrence in a cohort



<https://github.com/griffithlab/GenVisR>

# Functional relevance to cancer biology?

- Many useful resources that are gene centric
  - [Databases of tumor suppressors and oncogenes](#)
  - If the gene where a variant occurs is relevant to a disease or phenotype, then we turn our attention to the significance of the specific variant
- Does the variant have established functional relevance?
  - [DoCM - the Database of Curated Mutations](#) is a resource that aggregates specific variants with documented relevance to cancer
  - [Cancer Hotspots](#)
  - Others

# Clinically relevant?

---

- A gene or variant can be relevant to disease biology or a specific phenotype but not have any established clinical relevance
  - Arguably this is mostly the case...
  - The majority of genes and variants when observed do not lead to a clinical action
- Types of clinical relevance
  - Predictive (“druggable”)
  - Prognostic
  - Diagnostic
  - Predisposing

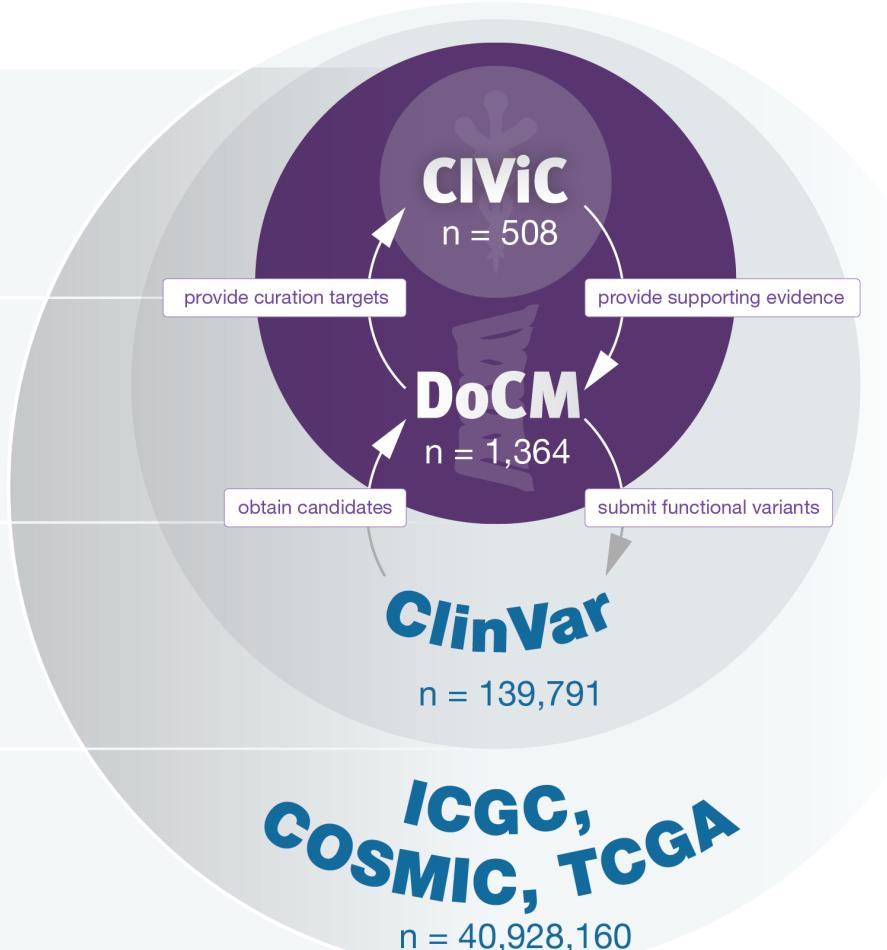
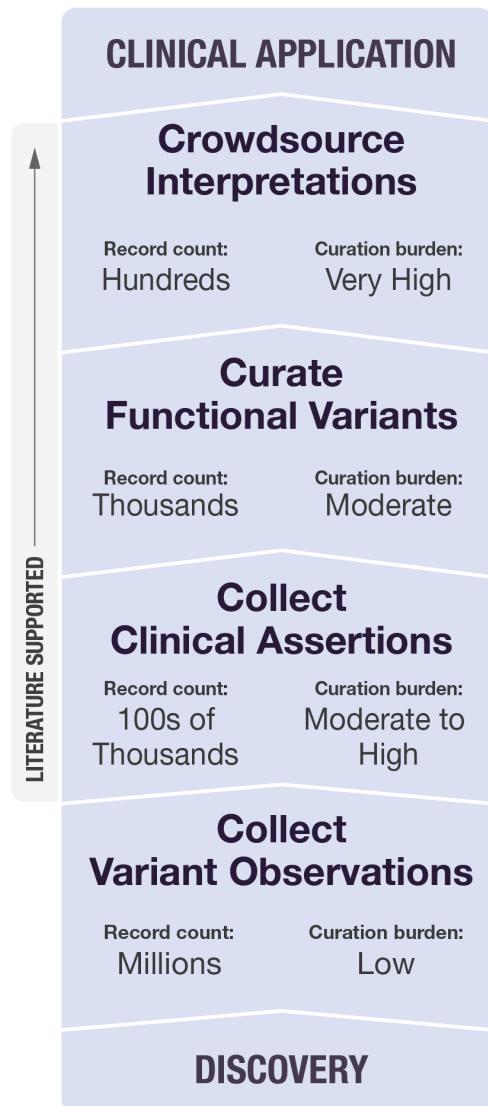
# Providing current and comprehensive interpretations of clinical significance of variants is a non-trivial task

## GENOMIC ALTERATIONS

GENE ALTERATION	INTERPRETATION
● PIK3CA H1047R	Mutations in PIK3CA have been reported in 26% to 33% of breast cancer cases (COSMIC, Jun 2012 and Kalinsky et al., 2009; 19671852). Activating mutations in PIK3CA, such as the one seen here, may predict sensitivity to inhibitors of PI3 kinase or its downstream signaling pathway (the PI3K/Akt/mTOR pathway) (Huang et al., 2007; 18079394). The mTOR inhibitors temsirolimus and everolimus have been tested in several clinical trials in breast cancer, and have been approved by the FDA for use in other tumor types. Inhibitors of PI3K and Akt are currently in clinical trials in breast cancer, alone or in combination with other therapies. PIK3CA mutations may play a role in resistance to hormonal therapy in ER+ breast cancers (Miller et al., 2011; 22114931). Activating mutations in PIK3CA may also confer resistance to anti-Her2 therapies (Chakrabarty et al., 2010; 20581867, Kataoka et al., 2010; 19633047, Wang et al., 2011; 21676217); combined inhibition of Her2 and the PI3K pathway may be required in tumors with ERBB2 amplification and PIK3CA mutation, though this remains an area of active investigation.
● CCND1 amplification	CCND1 amplification has been reported in approximately 10-15% of invasive breast cancers, more frequently in BRCA-negative cancers (Elsheikh et al., 2008; 17653856, Bane et al., 2011; 21327470). There are no approved therapies that directly target the protein product of CCND1 (Cyclin D1); however, CCND1 amplification may predict sensitivity to inhibitors of Cdk4 and Cdk6, which are currently under investigation in clinical trials. Overexpression of Cyclin D1 has also been associated with resistance to endocrine therapy in breast cancer (reviewed in Lange et al., 2011; 21613412; Musgrove and Sutherland, 2009; 19701242, Butt et al., 2005; 16113099).
● CDH1 E167*	CDH1 mutations are present in approximately 17% of breast cancers, and more often in luminal type cancers (COSMIC, Jun 2012, Hollestelle et al., 2010; 19593635). Loss of the E-cadherin protein, which is encoded by the CDH1 gene, has been associated with poor prognosis in triple negative breast cancer (Kashiwagi et al., 2010; 20551954, Tang et al., 2011; 21519872). Presently, there are no targeted therapies to address loss of CDH1/E-cadherin.

- Interpretations are typically produced by paid curators with no provenance and no mechanism for feedback
- This effort would be enhanced by a public domain effort

# Resources to help interpret clinical relevance



# Resources providing clinical interpretations of mutations are the most lacking

- [ClinVar](#)
- [ClinGen](#)
- [LOVD](#)
- [CIViC](#) (cancer specific)
- Many ad hoc “databases” at academic centers and hospitals
- Commercial software solutions, pharmaceutical companies, etc.

# CIViC was created to fill gaps in these resources

- [www.civicdb.org](http://www.civicdb.org)
- Information on the clinical impact of many cancer variants is scattered throughout the published literature
  - This makes collecting that information both time and labor-intensive
  - Yet we are all repeating each others efforts...
- CIViC could act as a centralized forum for curation, interpretation, and debate
- Existing resources do not facilitate computational access
  - No APIs, not open source, not open access, etc.
- Goal: facilitate automated report generation on clinical variants

The screenshot illustrates the CIViC BETA interface with various features highlighted:

- Site navigation**: Located at the top left.
- Edit content**: A green box highlights the "Edit content" link in the top navigation bar.
- Gene-level interpretation**: A blue box highlights the gene summary for **FLT3**, which includes a detailed description of its role in hematopoiesis and its association with AML, along with sources from Vardiman et al. (2009) and Strelakova et al. (2005).
- Gene variant navigation**: A blue box highlights the variant navigation section for **FLT3**, showing options like **DBS5**, **DBS5H**, **DBS5H/Y**, **ITD**, **MUTATION**, and **OVEREXPRESSION**.
- Sequence ontology**: A blue box highlights the **T227M** variant entry, which is categorized as a **TKD MUTATION**.
- Evidence records**: A blue box highlights the evidence record for **EID190**, which details a meta-analysis showing reduced survival for patients with FLT3-ITD.
- User activity/ attribution**: A green box highlights the "Submitted by" and "Accepted by" fields in the evidence record.
- Evidence record details**: A blue box highlights the detailed evidence record for EID190, including fields for Evidence Level (A-Validated), Evidence Type (Prognostic), Evidence Direction (Supports), Clinical Significance (Poor Outcome), Variant Origin (Somatic Mutation), Disease (Acute Myeloid Leukemia), Drug (N/A), Citation (Poni et al., 2014, Ann. Hematol.), and Trust Rating (★★★☆☆).
- Disclaimer**: A red box highlights the disclaimer at the bottom of the page, stating: "Disclaimer: This resource is intended for purely research purposes. It should not be used for emergencies or medical or professional advice."

Annotations on the right side point to specific UI elements:

- Sign In/ notifications**: Points to the "Sign In" button in the top right.
- "Talk page" (comments)**: Points to the "Gene Talk" tab in the top right.
- Imported gene information**: Points to the "MyGene.info" link in the gene summary panel.
- Variant-level interpretation**: Points to the variant summary panel for **ITD**.
- Variant coordinates**: Points to the reference build information in the variant summary panel.
- Data download/ table legend**: Points to the "Get Data" and "Help" buttons in the evidence table header.
- Suggested revision notice**: Points to the "0" badge in the "Evidence Talk" tab.
- Disease ontology**: Points to the "Disease" field in the evidence record details.
- Primary literature source**: Points to the "Citation" field in the evidence record details.
- Extensive documentation**: Points to the "Glossary of Terms", "API Documentation", "Data Releases", "Presentation Graphics", and "Contact" links at the bottom.
- CC attribute license**: Points to the "Creative Commons Attribution 4.0 International License" link at the bottom.

## Features of CIViC

- Browse evidence records for variant-drug-disease relationships as well as diagnostic and prognostic biomarkers
- Curation/ crowdsourcing interface
- Expert review
- Structured data model
- Use of existing ontologies
- JSON API to automate report generation