

The Elizabeth H.  
and James S. McDonnell III

**McDONNELL  
GENOME INSTITUTE**  
at Washington University



**Washington**  
University in St. Louis  

---

SCHOOL OF MEDICINE

## PMBIO Module 07

### Clinical. Clinical Applications

Malachi Griffith, Obi Griffith, Zachary Skidmore, Huiming Xia  
Introduction to bioinformatics for DNA and RNA sequence analysis (IBDR01)

29 October - 2 November, 2018  
Glasgow



## 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 module 07: Clinical

---

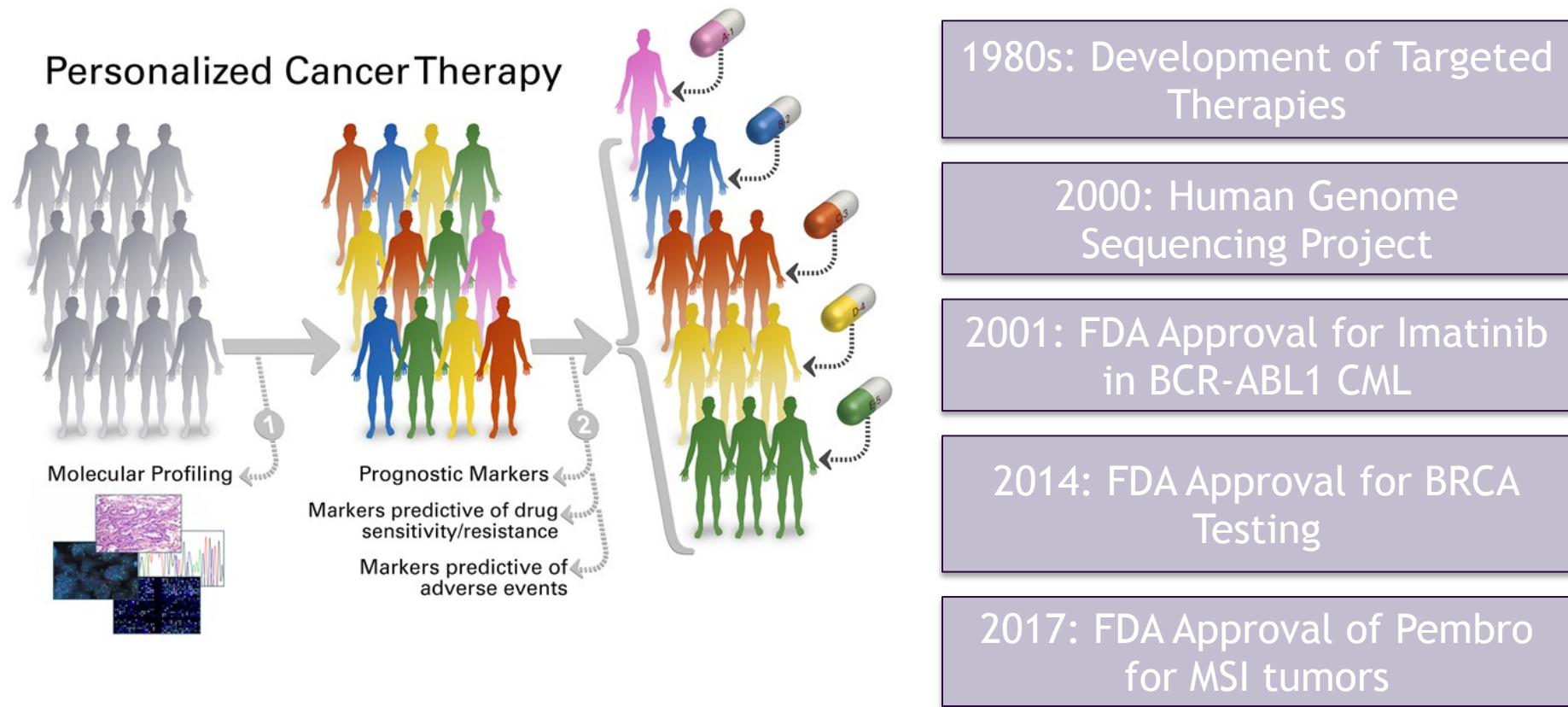
- **Key concepts:** Clinical interpretation of variants, variant interpretation types (predisposing, predictive, predisposing, diagnostic), variant interpretation guidelines, clinical applications of passenger somatic variants (early detection, minimal residual disease tracking, personalized vaccines).
- Prioritize germline variants to identify putative de novo and clinically relevant mutations
- Germline pathogenicity assessment (demonstrate concepts with some example variants using ClinGen Variant Curation Interface)
- Biological function or oncogenicity (demonstrate concepts by uploading a filtered VCF into CRAVAT)
- Somatic clinical actionability (demonstrate concepts with some example variants using CIViC)

# Strategies to bring genomics information to bear for as many cancer patients as possible

---

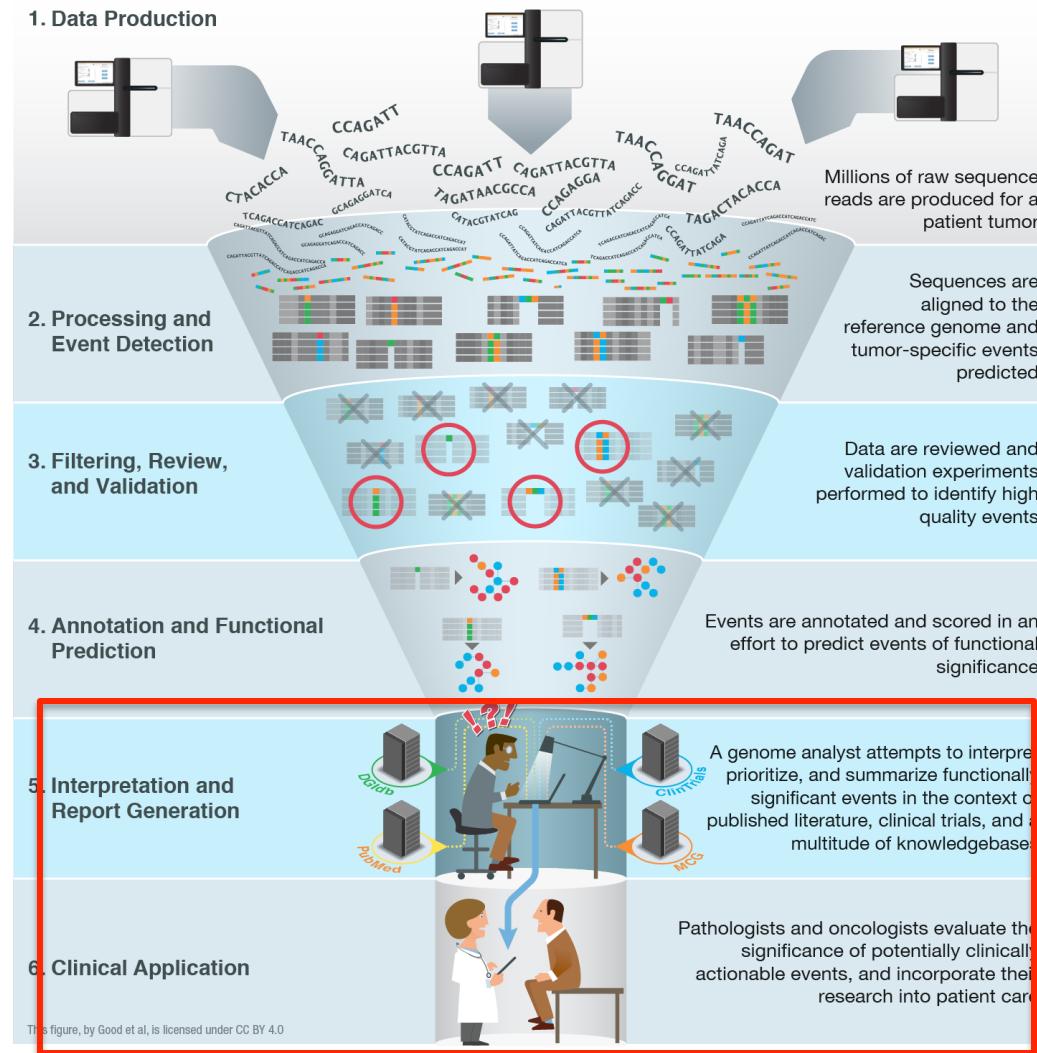
1. Precision medicine targeting of driver mutations
2. Leveraging passenger variants
  - a. Tracking minimal residual disease
  - b. Identifying neoepitopes
    - Predicting response to immunotherapy
    - Developing personalized vaccines

# Precision medicine targeting of driver mutations



BRAF -> V600E -> Melanoma -> Predictive -> Vemurafenib  
ERBB2 -> Amplification -> Breast -> Predictive -> Trastuzumab  
EGFR -> L858R -> Lung -> Predictive -> Erlotinib  
ALK -> Fusions -> Lung -> Predictive -> Crizotinib  
EWSR1-FLI1 -> Fusions -> Ewing Sarcoma -> Diagnostic  
DNAJB1-PRKACA -> Fusions -> fHCC -> Diagnostic  
VHL -> Loss of function mutations -> Kidney -> Predisposing  
... an increasingly long tail of rare but clinically relevant variants

# Problem: Clinical interpretation of genomic alterations remains a major bottleneck for realizing precision medicine



# Clinical interpretations of variants are currently created in information silos with limited use of standards and poor interoperability

## 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
- Interpretations are stored in hundreds of private databases
- This effort would be enhanced by an open public domain effort

# 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

---

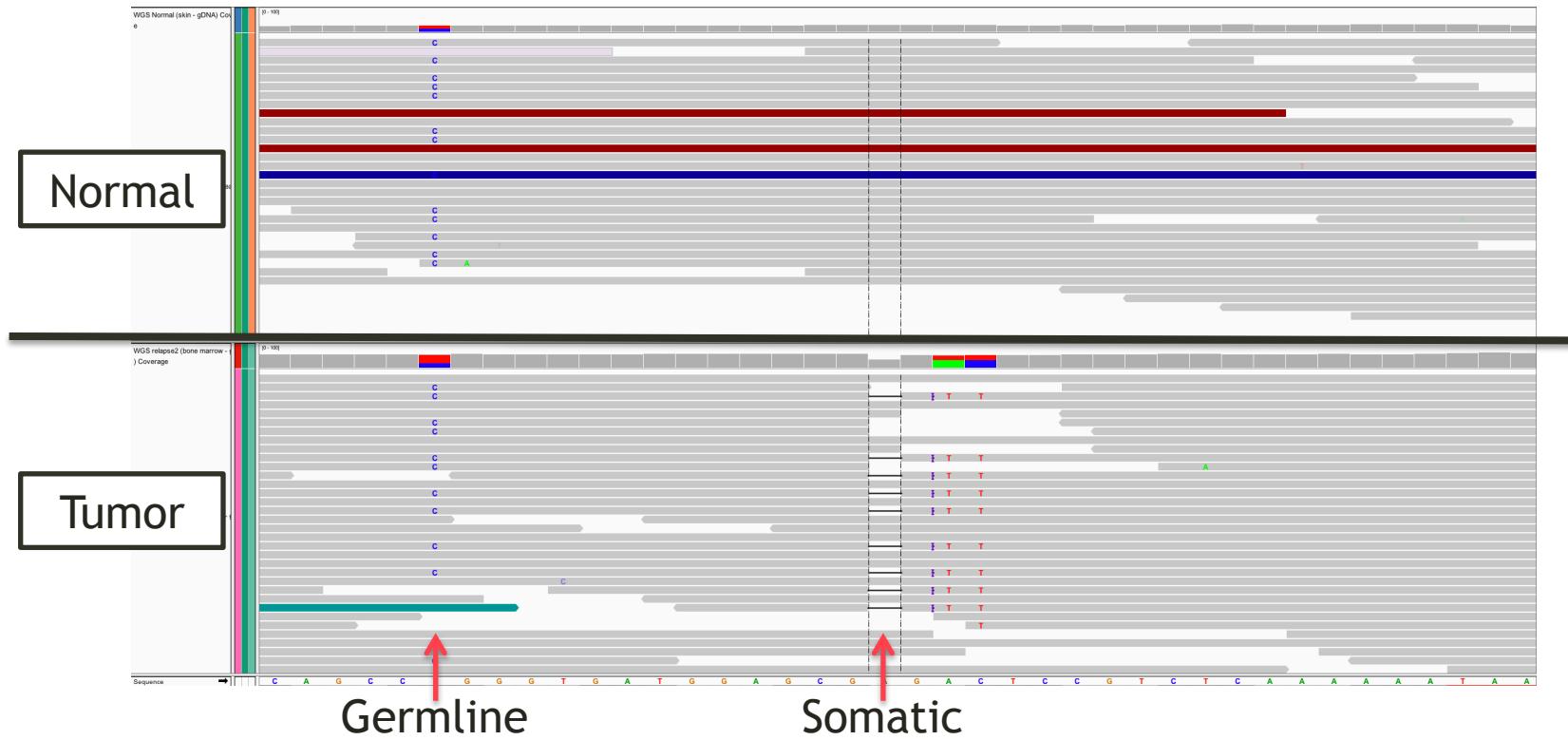
- False positives vs. true positives
- Somatic mutation vs. germline mutation vs. germline polymorphism
- Gain-of-function (activating) vs. loss-of-function (inactivating)
- Deleterious vs. tolerated
- Recurrent vs. random
- Driver vs. passenger
- Dominant clone vs. sub-clonal
- Regulatory vs. coding
- 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

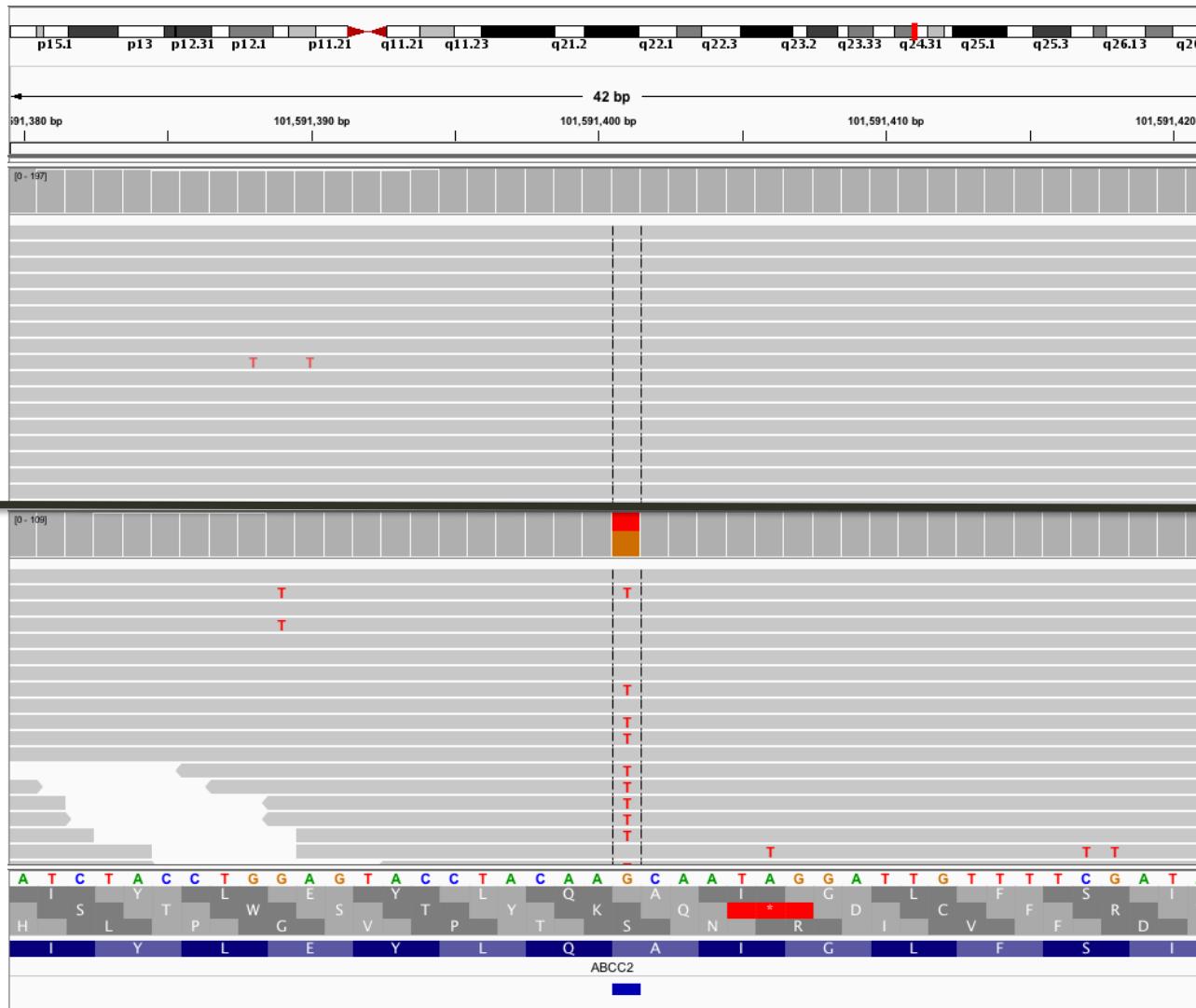


# False positives

---

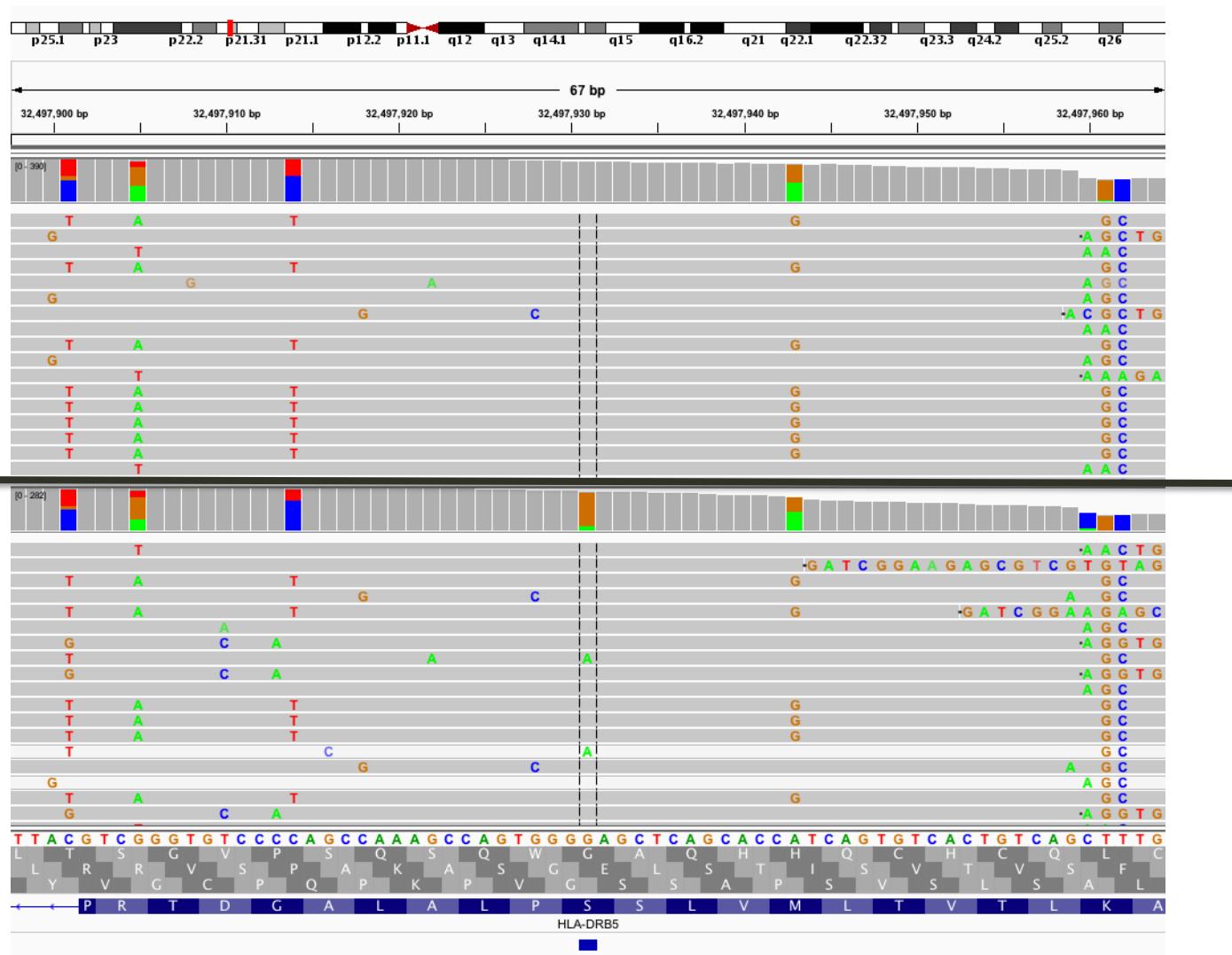
- Use IGV to examine alignments for artifacts
  - Use an intersection of variant callers
  - “panel of normals” analysis
- 
- Useful resources
    - [Discussion of variant callers](#)
    - [Optimizing tumor genome analysis](#)
    - [Hands-on IGV tutorial](#)

# Example of a high quality somatic variant



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

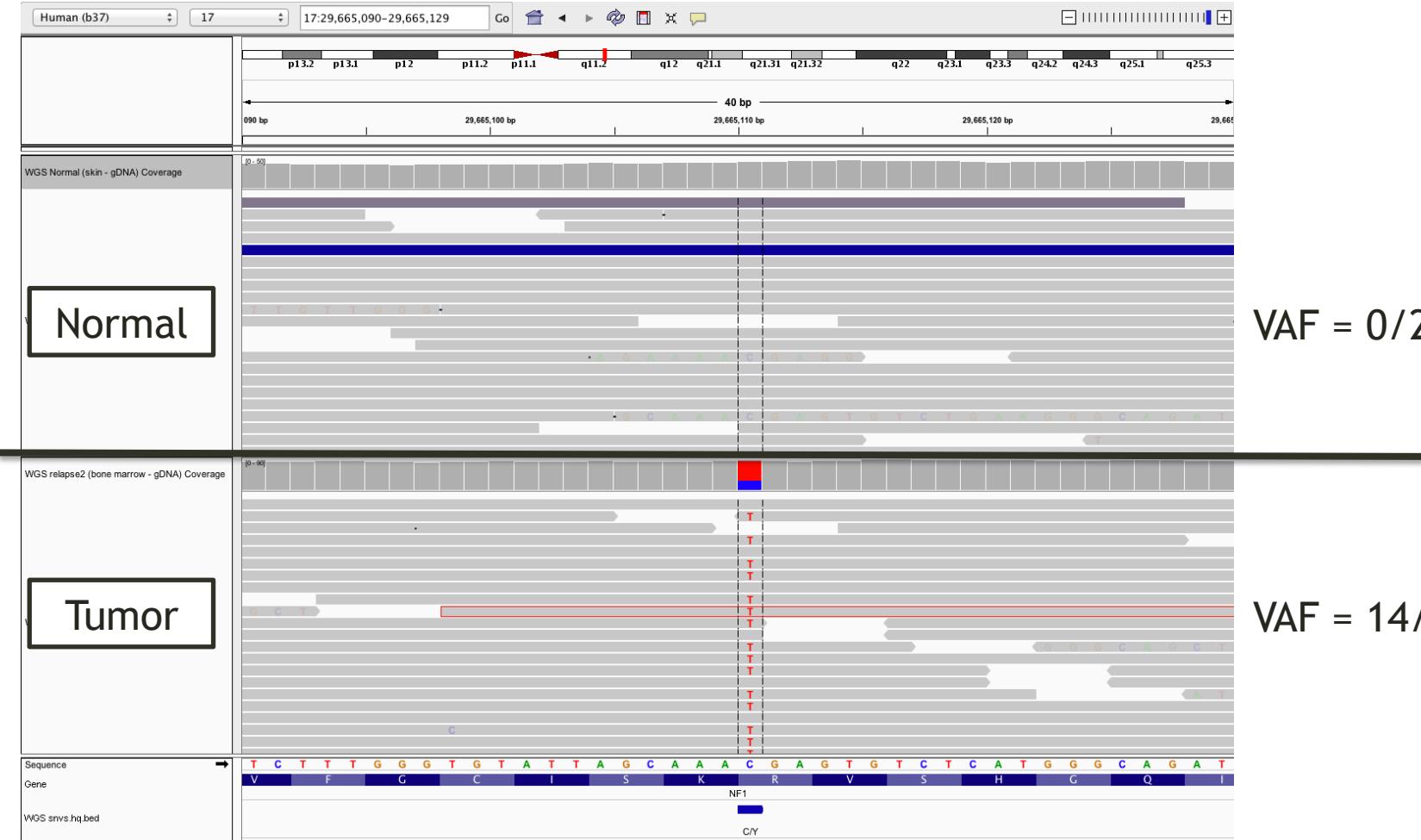
# Example of a low quality somatic variant



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

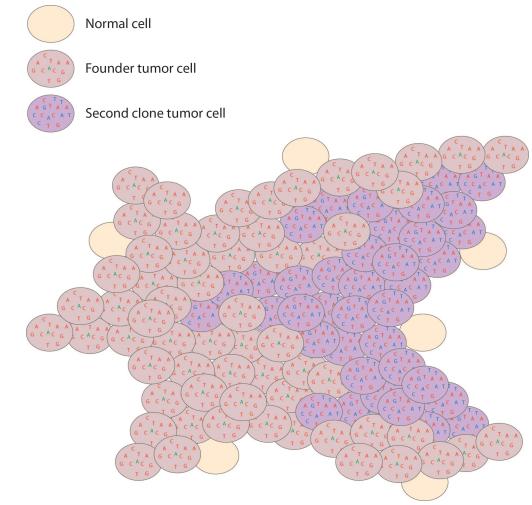
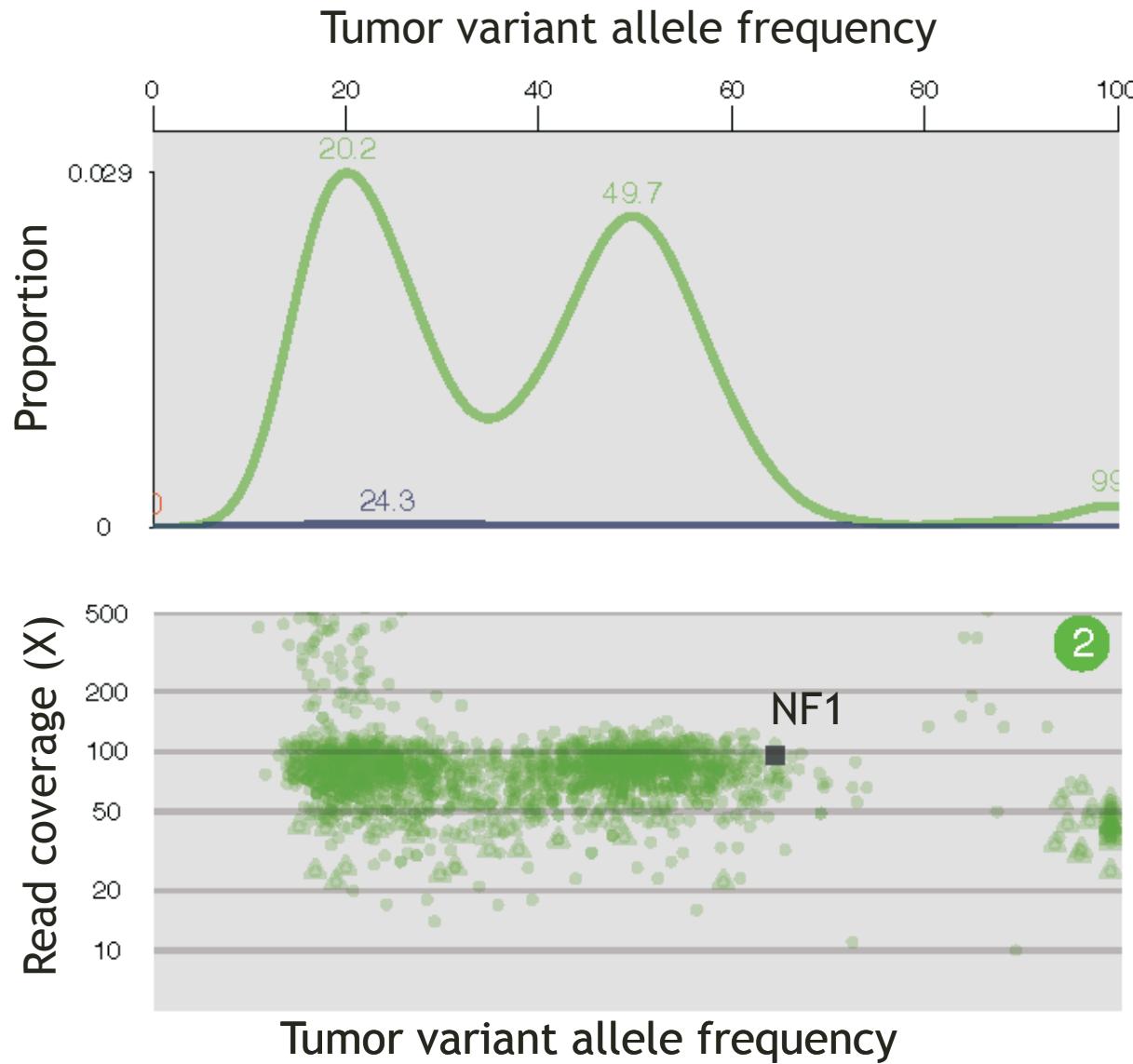
# Variant allele frequency (VAF)

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.

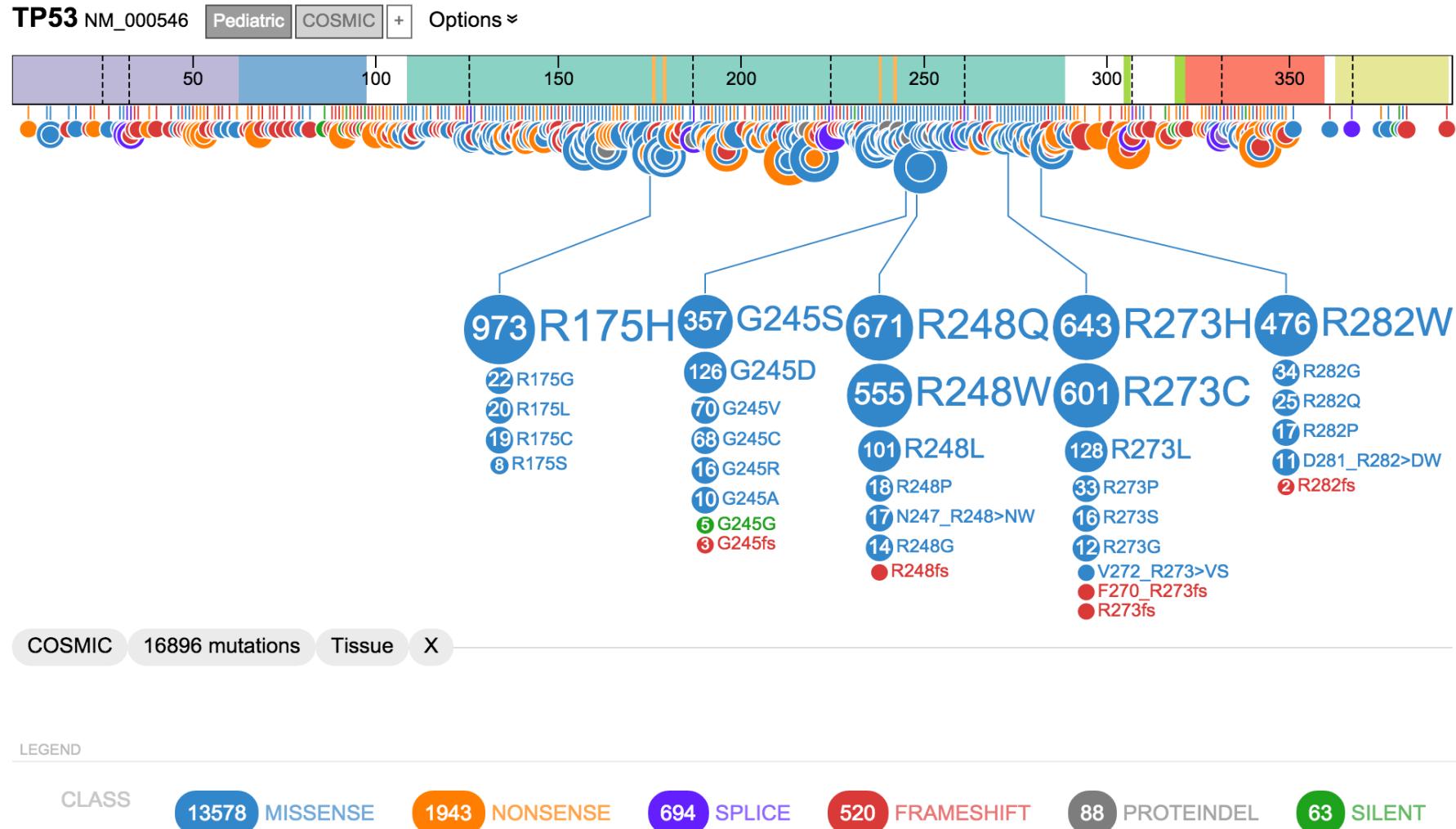
# Dominant clone vs. sub-clonal (and driver vs. passenger)



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

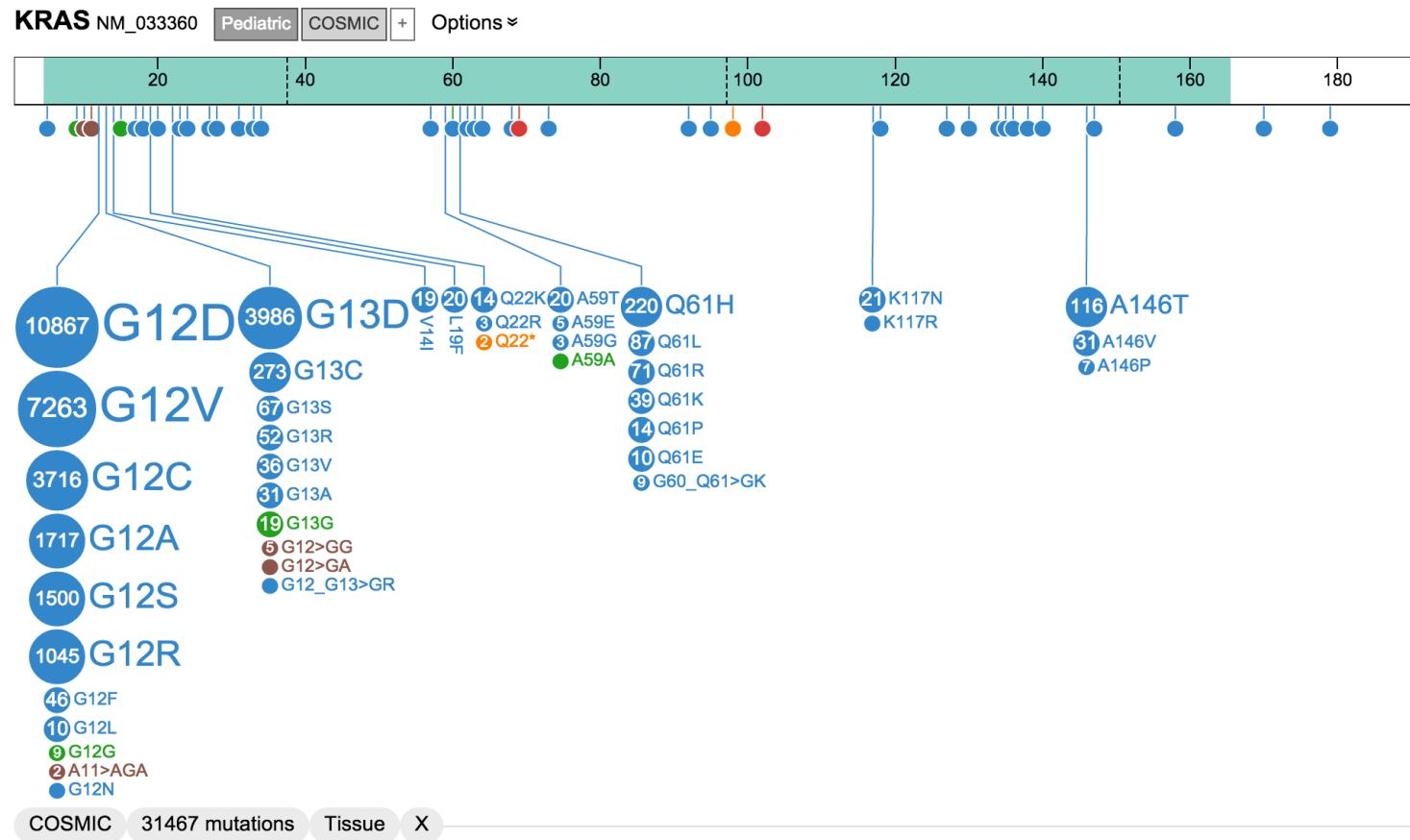
- Gain-of-function and loss-of-function generally refer to a change in specific function of a gene product (protein or RNA) leading to a phenotypic effect
  - Loss-of-function implies that a normal gene function is no longer possible
    - Many random mutations can lead to inactivation
    - Manifests 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



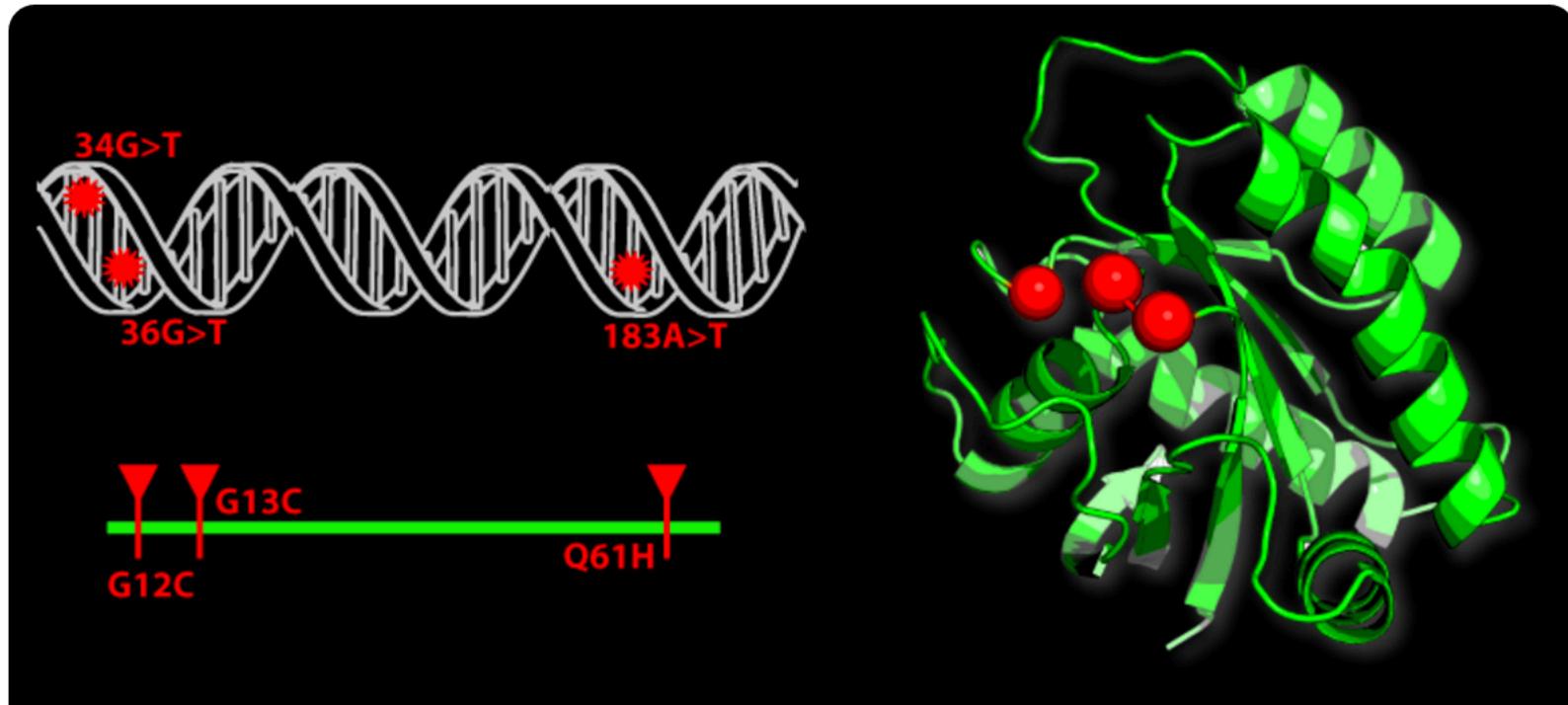
TP53: Recurrence pattern and mutation types suggest loss-of-function

# 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

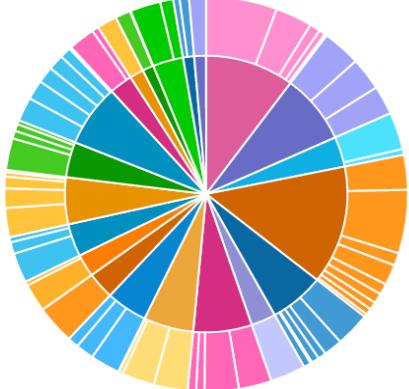
---

- 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

So far, TCGA and ICGC have sequenced 1000s of exomes, 100s of whole genomes, and surveyed the landscape of mutated genes for 40 cancer types

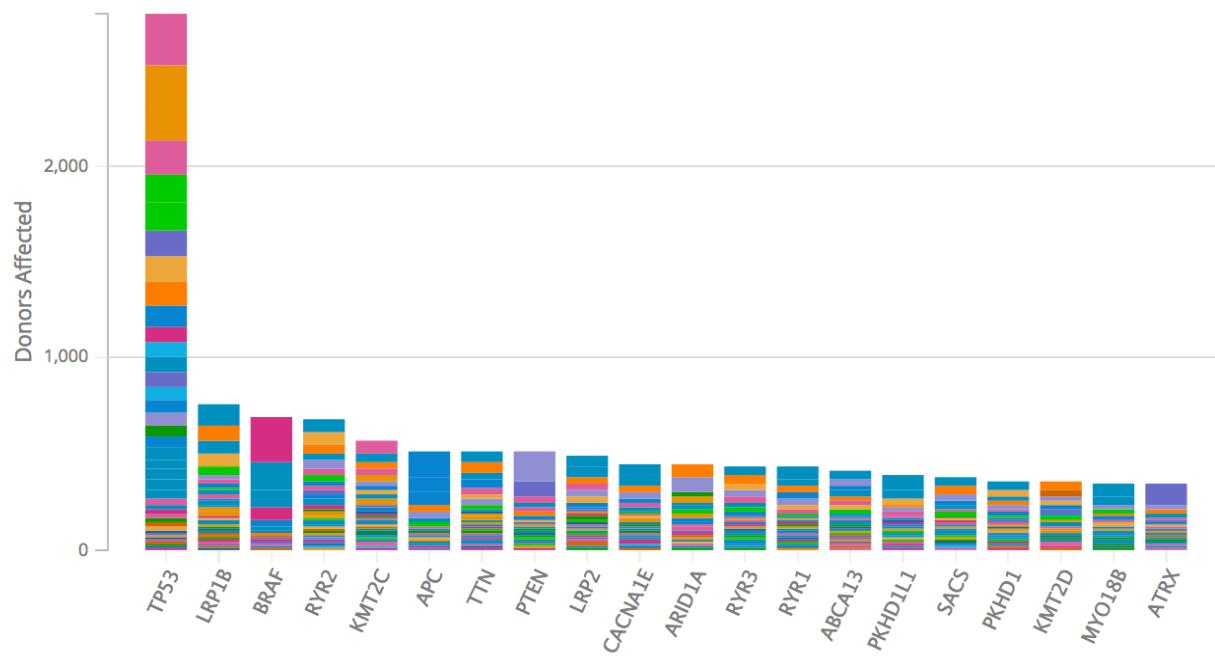
Donor Distribution

19,305 Donors across 70 Projects



Top 20 Mutated Genes with High Functional Impact SSMs

10,648 Unique SSM-Tested Donors



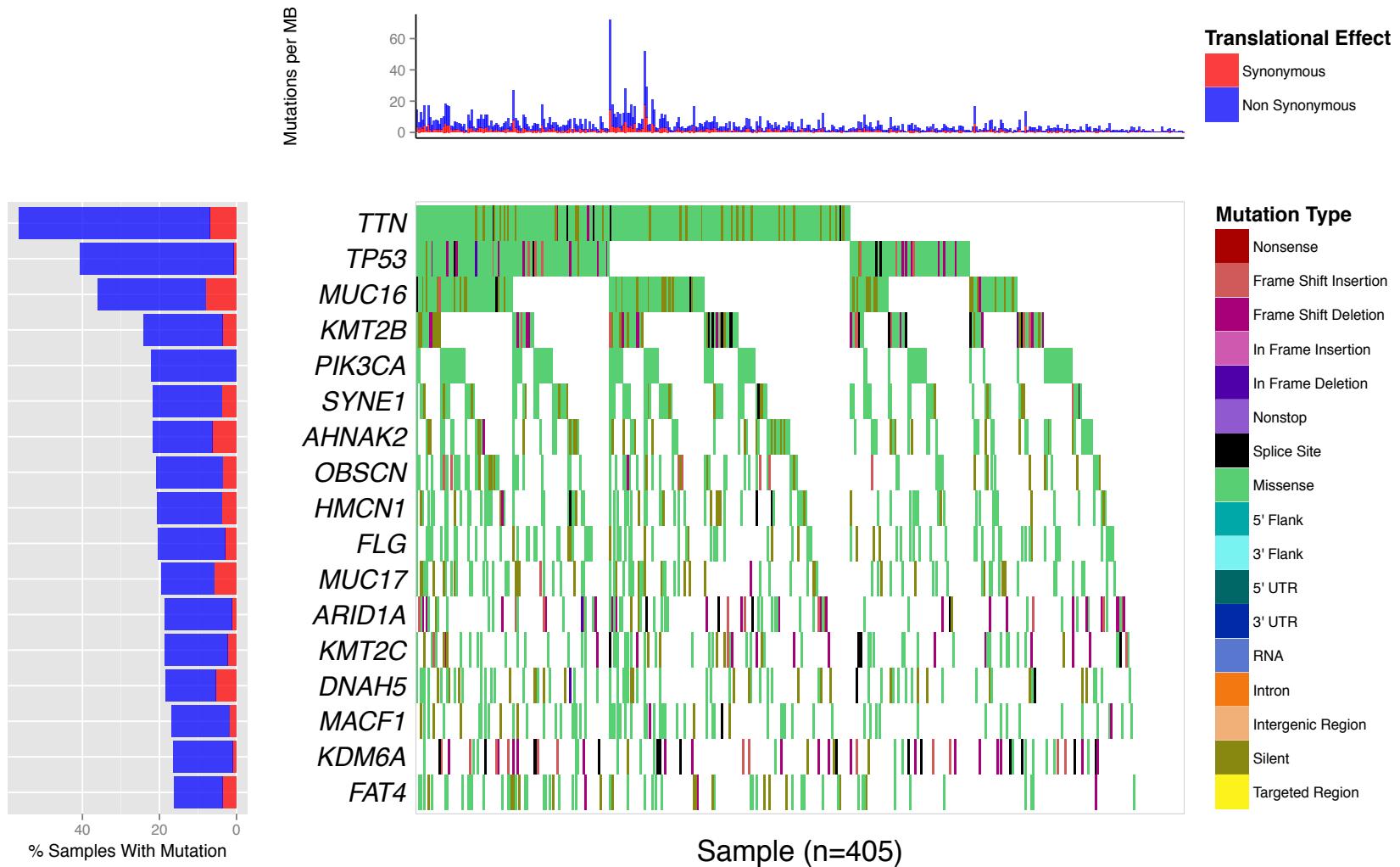
<https://dcc.icgc.org/>

# 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](#)).

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



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

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

---

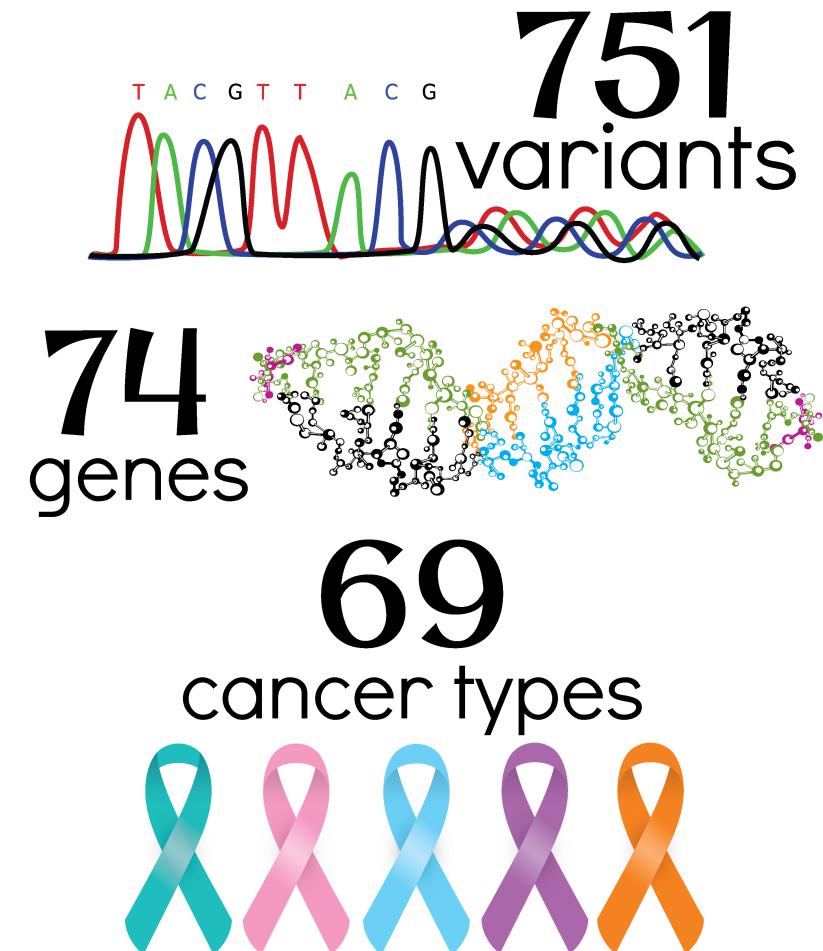
- Many tools/resources attempt to classify variants as “deleterious” vs. “tolerated” (aka “benign”)
  - E.g. Sift, PolyPhen, Condel, [CADD](#), etc.
- The goal of these tools is often confused with predicting whether the mutation is gain-of-function/activating
  - Not the same thing...
  - Mostly driven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

# 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 cancer, 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



- Criteria for inclusion
  - Confirmed to be recurrently mutated in one or more cancer types
  - Functional evidence
    - Studies elucidating function in cell lines/mouse models



<http://docm.info>

# Clinically relevant?

---

- A gene or variant can be relevant to cancer biology 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

# We created CIViC to address this need - an open knowledgebase and curation system for clinical interpretation of variants in cancer

### Data and Knowledge Production

Millions of raw sequence reads are produced for a patient tumor.

Sequences are aligned to the reference genome and tumor-specific events predicted.

Data are reviewed and validation experiments performed to identify high quality events.

Events are annotated and scored in an effort to predict events of functional significance.

### CIViC Curation

Crowdsourced curation efforts, moderated by experts in oncology and bioinformatics, help to build a knowledge-base of clinical interpretations of variants in cancer, describing the therapeutic, prognostic, diagnostic, and predisposing relevance of inherited and somatic variants of all types. Anyone may sign up to be a curator, add evidence, suggest changes to records, and discuss ongoing curation efforts.

### CIViC Curation Cycle

### Add New Evidence

### Review and Discuss Edits

### Research Gene, Variant, & Evidence Summaries

### A genome analyst uses CIViC's summaries to interpret and prioritize functionally significant events in the context of published literature, clinical trials, and linked knowledgebases.

### Pathologists and oncologists review analysts' reports to help evaluate the significance of potentially clinically actionable events and incorporate into patient care.

[www.civicdb.org](http://www.civicdb.org)

# CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer

Malachi Griffith, Nicholas C Spies, Kilannin Krysiak, Joshua F McMichael, Adam C Coffman, Arpad M Danos, Benjamin J Ainscough, Cody A Ramirez, Damian T Rieke, Lynzey Kujan, Erica K Barnell, Alex H Wagner, Zachary L Skidmore, Amber Wollam, Connor J Liu, Martin R Jones, Rachel L Bilski, Robert Lesurf, Yan-Yang Feng, Nakul M Shah, Melika Bonakdar, Lee Trani, Matthew Matlock, Avinash Ramu, Katie M Campbell, Gregory C Spies, Aaron P Graubert, Karthik Gangavarapu, James M Eldred, David E Larson, Jason R Walker, Benjamin M Good, Chunlei Wu, Andrew I Su, Rodrigo Dienstmann, Adam A Margolin, David Tamborero, Nuria Lopez-Bigas, Steven J M Jones, Ron Bose, David H Spencer, Lukas D Wartman, Richard K Wilson, Elaine R Mardis & Obi L Griffith [Show fewer authors]

[Affiliations](#) | [Contributions](#) | [Corresponding authors](#)

*Nature Genetics* 49, 170–174 (2017) | doi:10.1038/ng.3774

Published online 31 January 2017

# All content is freely available for all uses at

## [www.civicdb.org](http://www.civicdb.org)

The screenshot shows the CIViC website homepage with a dark purple background featuring a faint circular DNA helix pattern. At the top right is a navigation bar with links: About, Participate, Community, Help, FAQ, and Sign In/Sign Up. Below the navigation is a search bar with the placeholder "Go to Genes & Variants" and a "Go!" button. Underneath the search bar are three buttons: BROWSE, SEARCH, and ACTIVITY.

**CIViC**  
CLINICAL INTERPRETATIONS OF  
VARIANTS IN CANCER

 Discover supported clinical interpretations of mutations related to cancer.

 Participate with colleagues to add variants and support for cancer-related mutations.

**The Precision Medicine Revolution**  
Precision medicine refers to the use of prevention and treatment strategies that are tailored to the unique features of each individual and their disease. In the context of cancer this might involve the identification of specific mutations shown to predict response to a targeted therapy. The biomedical literature describing these associations is large and growing rapidly. Currently these interpretations exist largely in private or encumbered databases resulting in extensive repetition of effort.

**CIViC's Role in Precision Medicine**  
Realizing precision medicine will require this information to be centralized, debated and interpreted for application in the clinic. **CIViC is an open access, open source, community-driven web resource for Clinical Interpretation of Variants in Cancer.** Our goal is to enable precision medicine by providing an educational forum for dissemination of knowledge and active discussion of the clinical significance of cancer genome alterations. For more details refer to the 2017 [CIViC publication](#) in Nature Genetics.

# ClinGen Variant Curation Interface (VCI)

The screenshot shows the ClinGen Variant Curation Interface (VCI) interface. At the top, there is a navigation bar with links for "Help", "New Variant Curation", "New Gene Curation", a home icon, and "Logout Malachi Griffith". The main content area displays a variant record:

**NM\_005902.3(SMAD3):c.-28C>T**  
This interpretation is not yet associated with a disease or mode of inheritance

**View Summary**

Variant ID Sources	Variant Genomic Context	My Interpretation
ClinVar VariationID: 139214 dbSNP ID: rs144374592	UCSC [GRCh38/hg38]   GRCh37/hg19 Variation Viewer [GRCh38]   [GRCh37] Ensembl Browser [GRCh38]   [GRCh37]	Disease: Not associated Calculated Pathogenicity: None Modified Pathogenicity: None Status: In Progress Last Edited: 2018 Feb 15, 2:26 pm

Below the table are buttons for various interpretation categories: BA1, BS1, BS2, BS3, BS4, BP1, BP2, BP3, BP4, BP5, BP6, BP7, PP1, PP2, PP3, PP4, PP5, PM1, PM2, PM3, PM4, PM5, PM6, PS1, PS2, PS3, PS4, PSV1.

**Variant Interpretation Record**

**Benign** No criteria met      **Pathogenic** No criteria met      Calculated Pathogenicity: None

Basic Information   Population   Predictors   Experimental   Case/Segregation   Gene-centric

**Genomic**  
NC\_000015.10:g.67066127C>T (GRCh38)  
NC\_000015.9:g.67358465C>T (GRCh37)

**Overall ClinVar Interpretation** See data in ClinVar  
Review status: Criteria provided, multiple submitters, no conflicts   Last evaluated: Jun 14, 2016  
Clinical significance: Benign   Number of submission(s): 3

**Interpretations Submitted to ClinVar** See data in ClinVar  
Clinical significance (Last evaluated)   Review Status (Assertion method)   Condition(s) (Mode of inheritance)   Submitter - Study name   Submission accession

Benign (Oct 21, 2013)	criteria provided, single submitter GeneDx Variant Classification (06012015)	not specified [MedGen]	GeneDx	SCV000171758.11
-----------------------	---	---------------------------	--------	-----------------

CIViC and the VCI are distinguished by their emphasis on the curation process, provenance, adherence to documented guidelines, openness, etc.

# The CIViC curation cycle promotes quality, provenance and an up-to-date knowledgebase

NCBI Resources How To

PubMed Advanced

Format: Abstract Send to

Cancer Discov. 2013 Feb;3(2):224-37. doi: 10.1158/2159-8290.CD-12-0349. Epub 2012 Dec 7.

Activating HER2 mutations in HER2 gene amplification negative breast cancer.

Bose R<sup>1</sup>, Kavuri SM, Searleman AC, Shen W, Shen D, Koboldt DC, Monsey J, Goel N, Aronson AB, Li S, Ma CX, Ding L, Mardis ER, Ellis MJ.

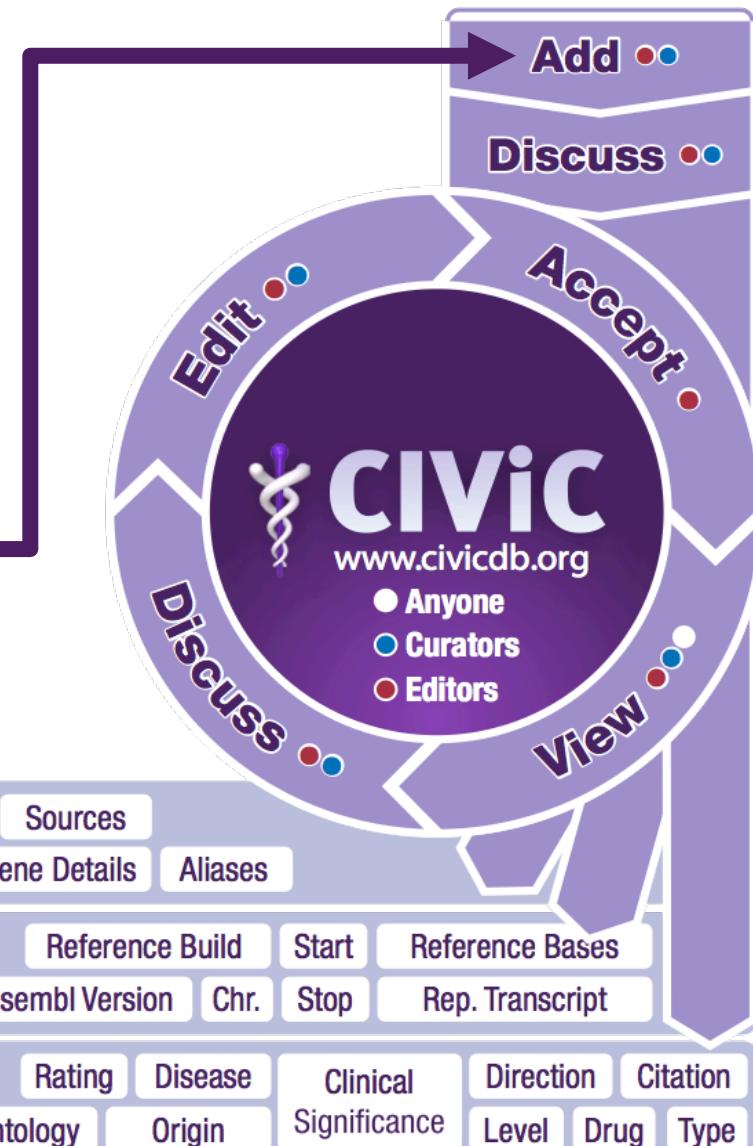
Author information

Abstract

Data from 8 breast cancer genome-sequencing projects identified 25 patients with HER2 somatic mutations in cancers lacking HER2 gene amplification. To determine the phenotype of these mutations, we functionally characterized 13 HER2 mutations using *in vitro* kinase assays, protein structure analysis, cell culture, and xenograft experiments. Seven of these mutations are activating mutations, including G309A.

**Evidence Record**

Rating	Disease	Clinical Significance	Direction	Citation	
Statement	Disease Ontology	Origin	Level	Drug	Type



## How is quality maintained?

- Content creation is completely transparent
- Anyone can become a curator but this role has limited powers
- All content must be reviewed by a site editor or domain expert
- Users can not accept their own contributions
- Problems can be identified by comment or flag

# How can you participate? Visit [www.civicdb.org](http://www.civicdb.org) to join as a curator

EVIDENCE EID352

Submitted by Accepted by Last Modified by Last Reviewed by

Patients with SF3B1 mutations had a statistically significant longer overall survival as well as event free survival. Both before and after adjustment for age, karyotype and sex.

Evidence Level: <b>B - Clinical</b>	Disease: Myelodysplastic Syndrome
Evidence Type: Prognostic	Drug: N/A
Evidence Direction: Supports	Citation: Papayannidis et al., 2011, N. Engl. J. Med.
Clinical Significance: Better Outcome	Pubmed ID: <a href="#">21995386</a>
Variant Origin: Somatic Mutation	Trust Rating: ★★★★☆

EID352 Revisions EID352 Comments EID352 Log

According to Figure 2 of this source, only five of the SF3B1 mutations observed were actually K666N. There does not seem to be statistical support for the prognostic claim for this specific mutation. In the paper (Figure 2), the survival analysis compared patients with any mutation in SF3B1 to patients that were wild type. We should probably move this evidence to a new MUTATION variant where we list examples of the mutations they combined together.

Posted by 5 months ago

Note that I moved this variant over to MUTATION.

Posted by 5 months ago

In PMID26842708 Alsaif et. al use RNASeq to examine the splicing effect of SF3B1-K666 and SF3B1-R625 v SF3B1-WT in susceptible splice junctions in primary uveal melanoma. They demonstrate that these mutations mimic neither knockdown nor over-expression of SF3B1 but "change of function" mutations that alter the spliceosomes ability to efficiently recognize the splice branchpoint (BP) in a subset of splice junctions. Hotspot mutations in SF3B1 are spatially related and all occur in HEAT domains. It is unclear why K700 mutations predominate in haematological malignancy and not uveal melanoma. Altogether, literature to date suggests loss of conserved Arg or Lys within the HEAT domains of SF3B1 define a particular class of mutation leading to spliceosome dysfunction and should not merely be classified as "mutation".

Discuss or flag  
for review

Suggest  
publications

EVIDENCE EID1528

Submitted by Accepted by Last Modified by Last Reviewed by

A 25-year-old Caucasian male presented with a second relapse of B-ALL. RNA sequencing revealed FLT3 overexpression (also confirmed by IHC). The patient was treated with 22 days of sunitinib, permitting the patient to undergo a bone marrow transplant. A bone marrow biopsy showed complete clearance by variant allele frequency and the patient remains in remission 4 years post matched-donor allograft.

Evidence Level: <b>C - Case Study</b>	Disease: Adult B-Lymphoblastic Leukemia
Evidence Type: Predictive	Drug: Sunitinib
Evidence Direction: Supports	Citation: N/A
Clinical Significance: Sensitivity	Pubmed ID: <a href="#">27181063</a>
Variant Origin: Somatic Mutation	Trust Rating: ★★★★☆

## EDIT EVIDENCE ITEM EID1528

Complete your edits, then click the 'Submit Revision for Review' button.

\* Variant Origin: Somatic Mutation

\* Pubmed ID: 27181063

Citation: Griffith et al., 2016, Exp. Hematol.

\* Disease: Adult B-Lymphoblastic Leukemia

Disease Ontology ID: DOID:0000392

1 total revisions

Re... Submitted by	Status	Created
17...	applied	7 months ago

Revision #1732 applied

Description

- DELETIONS + INSERTIONS

In patients with non-small cell lung cancer harboring EML4-ALK fusion, the C1156Y variant has been shown to confer resistance to crizotinib. A 28 year old patient with T4N3M1 stage IV adenocarcinoma harboring an EML4-ALK variant 1 fusion was treated with crizotinib after failing conventional therapy. The patient achieved a partial

Suggest  
revisions

Add new  
content

Go to Genes & Variants Go! BROWSE SEARCH ACTIVITY ADD SUGGEST

**Suggest Source**

If you have a source publication that you believe to contain research relevant to CIVIC's curation efforts, the community would appreciate your input! Please specify as many details as you can using the form below (only the Pubmed ID and Comment are required). Your suggested source will be placed in the [source curation queue](#), and you will be given a suggestion credit if your suggestion results in a new Evidence Item added to the CIVIC database.

\* Pubmed ID:  Citation:  PubMed ID for the publication associated with the evidence statement (e.g. 23463675)

Gene Entrez Name:  Entrez ID:  Variant Name:  Disease:

**Curation Tools**

Source Suggestions 90 total items (showing 81)

Status	Submitter	Citation	Gene	Variant	Disease	Comment	Actions
new	RodrigoDienemann	2016, Nat. Med.	MET		Glioblastoma Multiforme	case report of response...	+ ○ ○ ○ C
new		Aster et al., 2002, Am. J...	BCL2		Follicular Lymphoma	BCL2 rearrangements in ...	+ ○ ○ ○ C
new		Aung et al., 2016, Cold ...	ERBB2	L755S	Colorectal Adenocarcinoma	A case report indicating...	+ ○ ○ ○ C

FLT3 Variants

D835 D835H D835H/Y ITD MUTATION OV

VARIANT OVEREXPRESSION ↗

Last Modified by Last Reviewed by

This Variant does not have any variants. Participate Community Help FAQ (113)

Add a Summary

Variant Type: N/A HGVS Expression: None specified.

ACTIVITY ADD SUGGEST

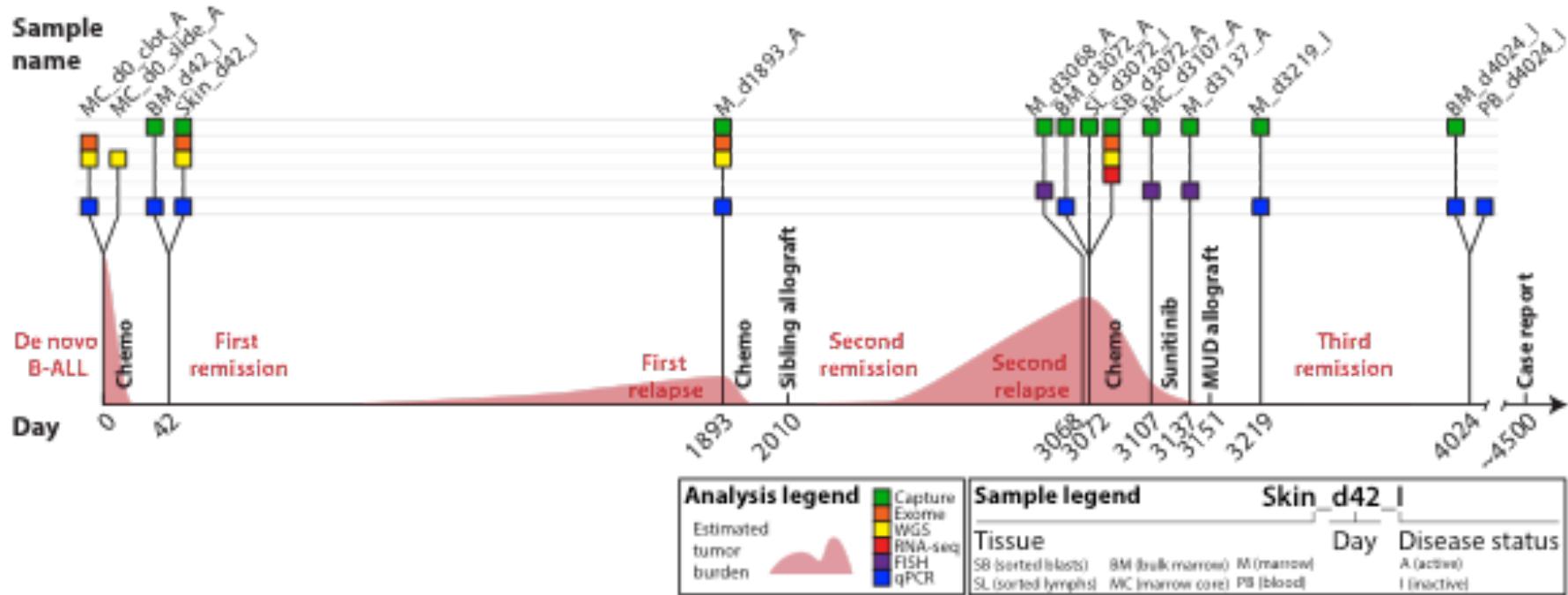
# Strategies to bring genomics information to bear for as many cancer patients as possible

---

1. Precision medicine targeting of driver mutations
  
2. Leveraging passenger variants
  - a. Tracking minimal residual disease
  - b. Identifying neoepitopes
    - Predicting response to immunotherapy
    - Developing personalized vaccines

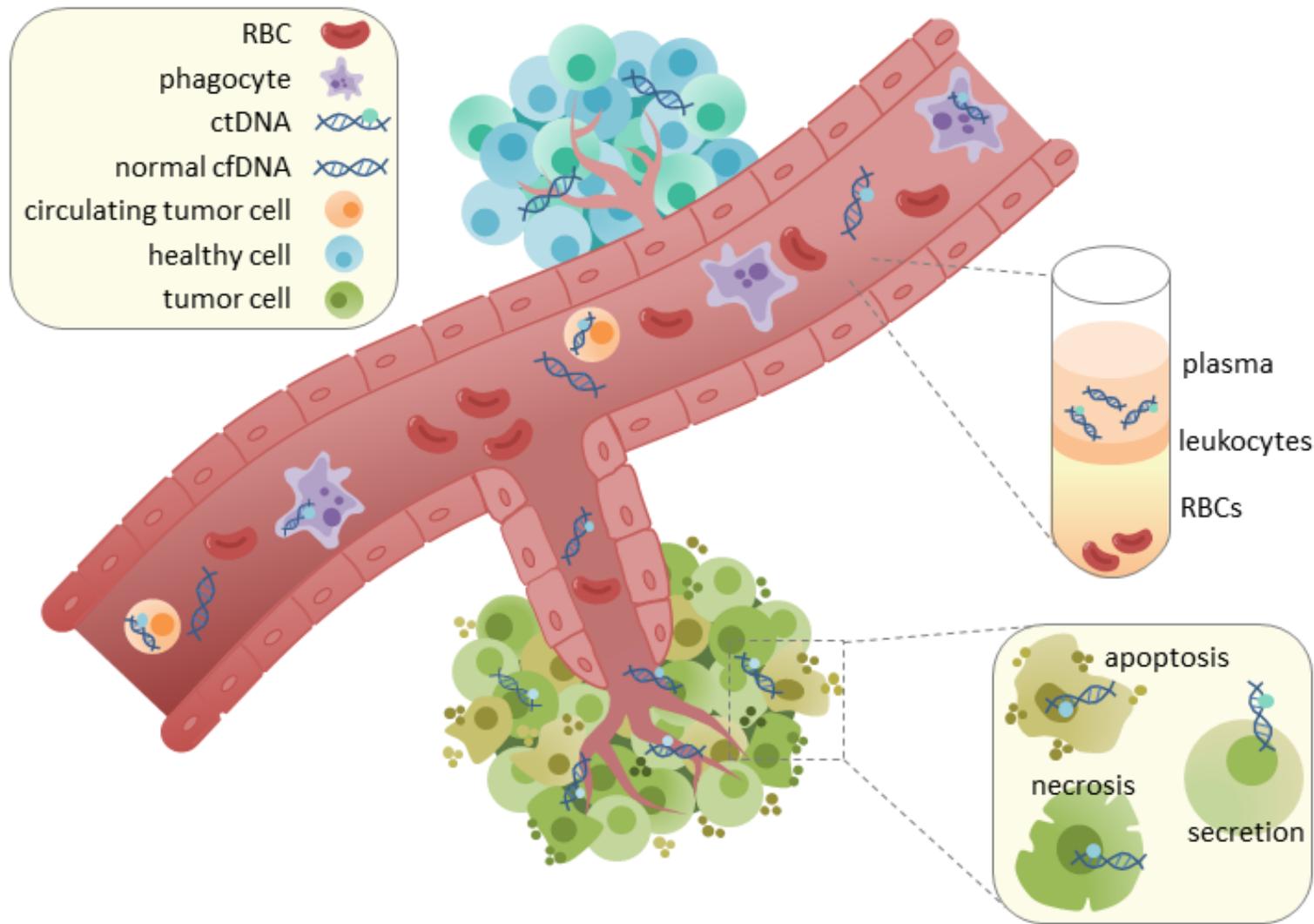
In the example leukemia (ALL) case we performed comprehensive disease tracking using bone marrow biopsies

A

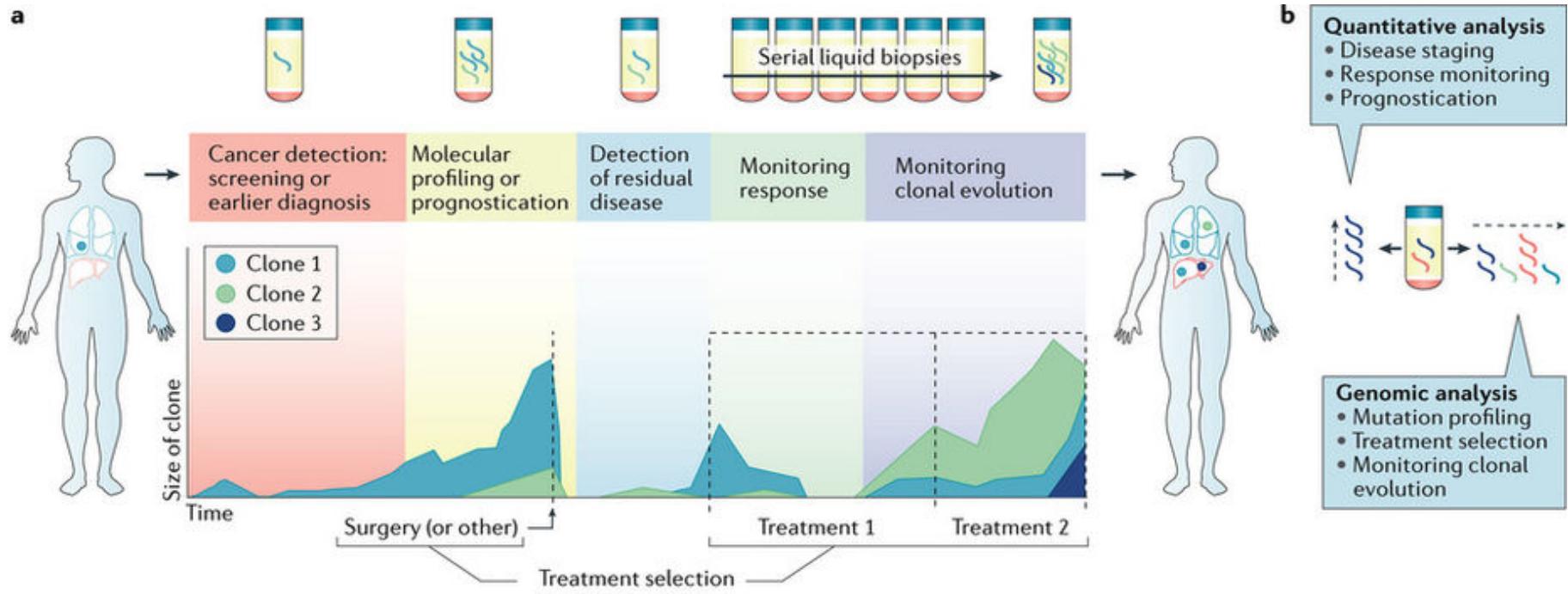


We ultimately profiled 18 samples from nine time points (bone marrow biopsies) throughout disease progression to better understand the evolution of this tumor

# Circulating tumor DNA (ctDNA) could allow generalized tracking in any cancer type

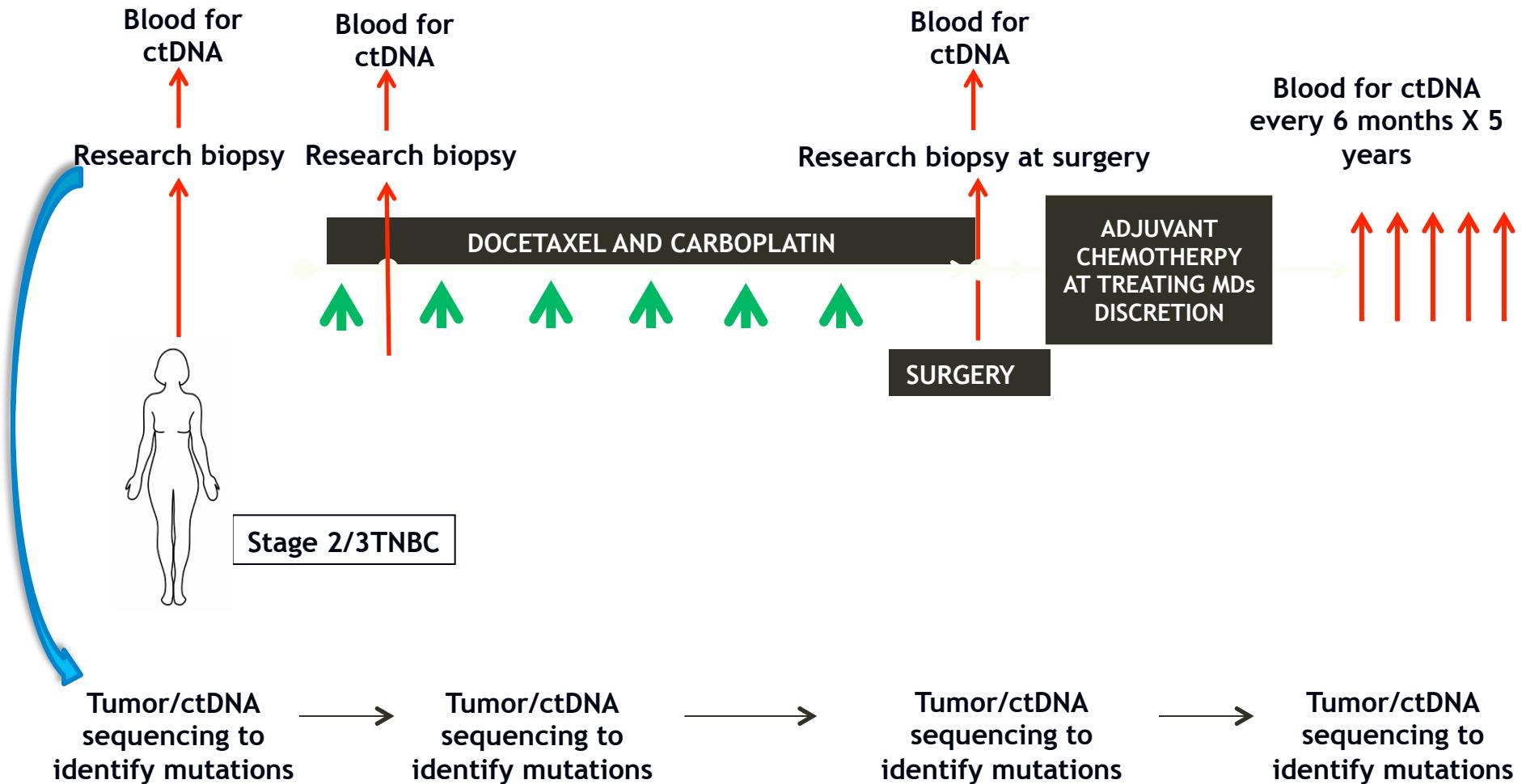


# ctDNA could allow early detection, risk assessment, monitoring of response during therapy, and detection of residual disease after therapy



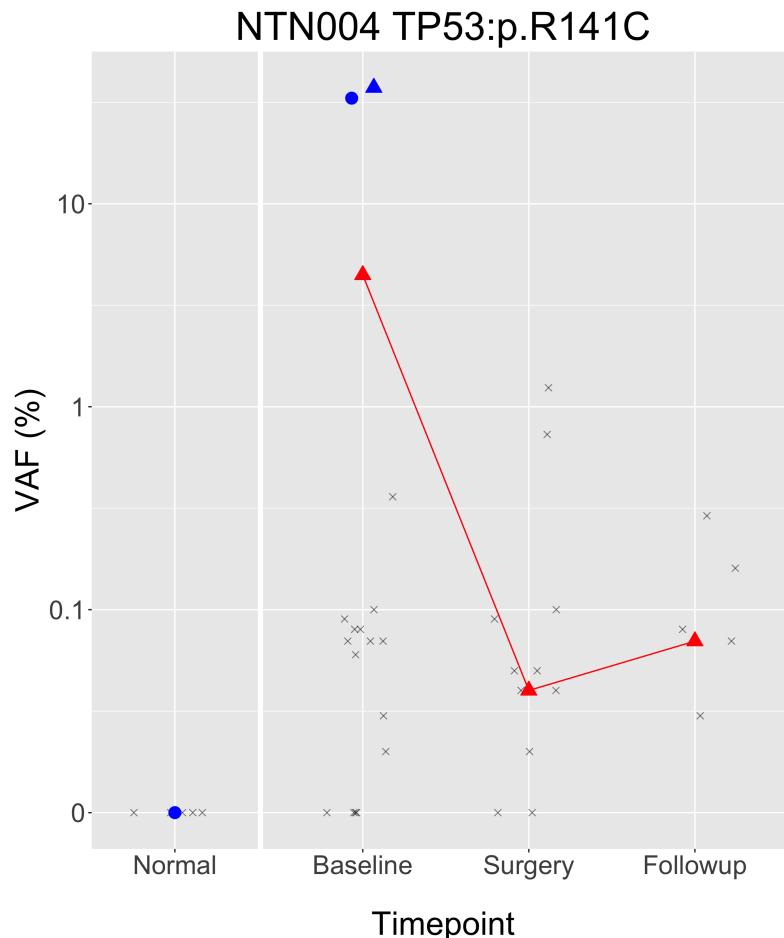
Nature Reviews | Cancer

# Preliminary ctDNA data in triple negative breast cancer

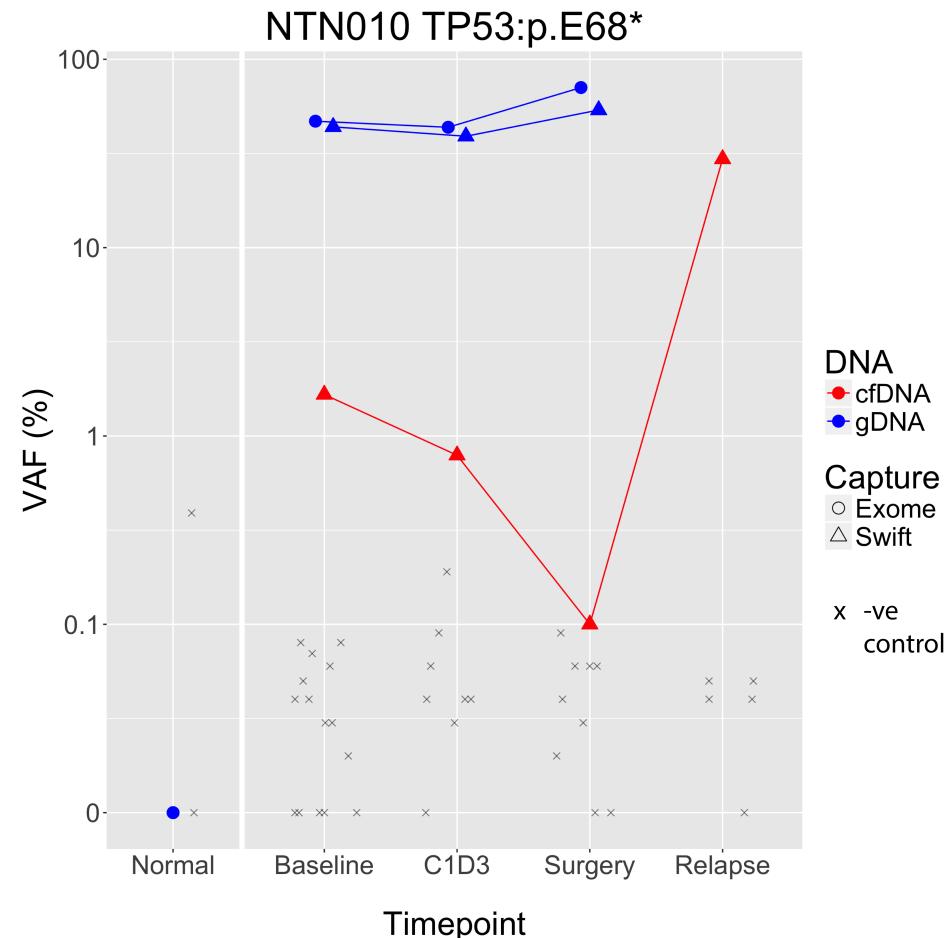


# Breast ctDNA preliminary data is promising

pCR Example



Non pCR Example



Exome sequencing of tumor tissue to discover somatic variants, followed by ultradeep targeted assay of cfDNA isolated from plasma