

# Interpretation of DNA variation in the context of cancer

*Dr. Sigve Nakken*

*Centre for Cancer Cell Reprogramming (CanCell)*

*Dept. of Tumor Biology*

*Institute for Cancer Research*

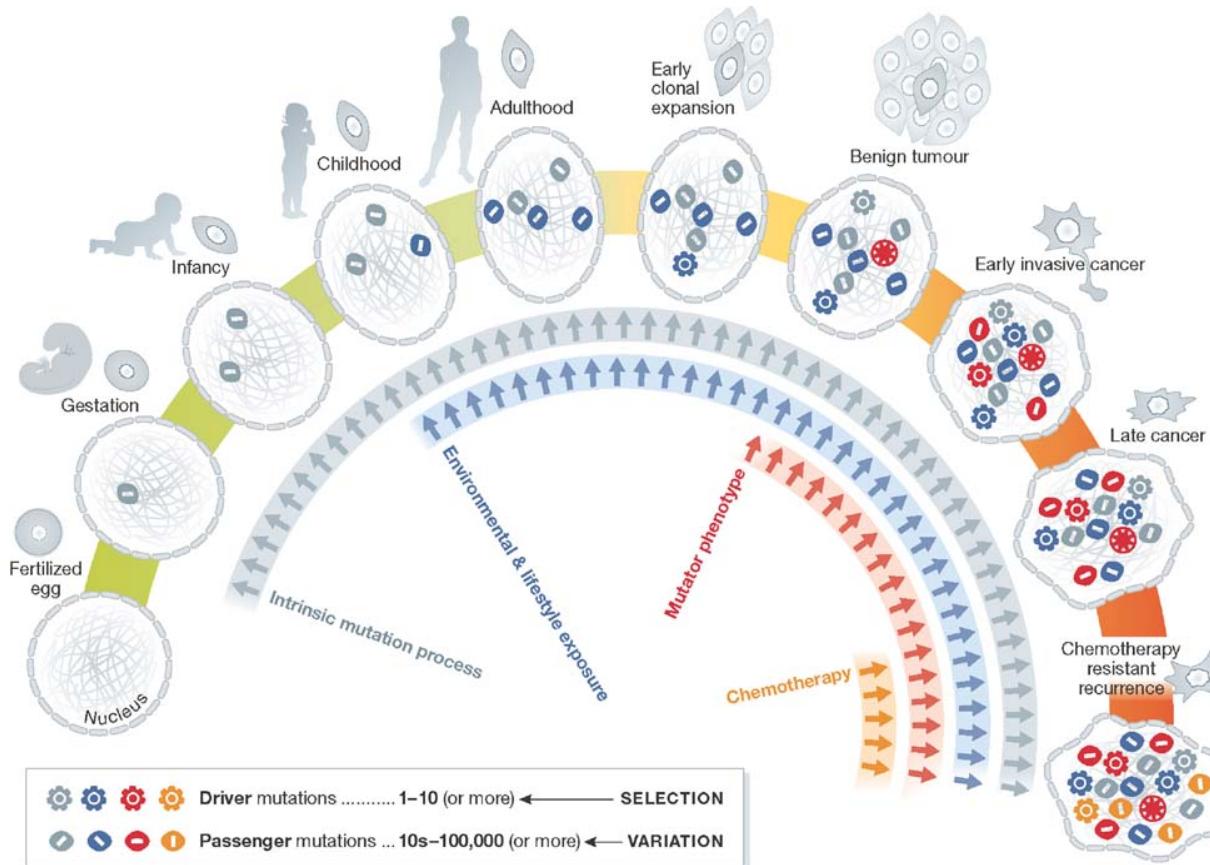
*Oslo, Norway*

*IN-BIOS9000/5000, Fall 2020*

***PS. At 1530, Torbjørn will give some info  
regarding exam and course evaluation***

# Why cancer?

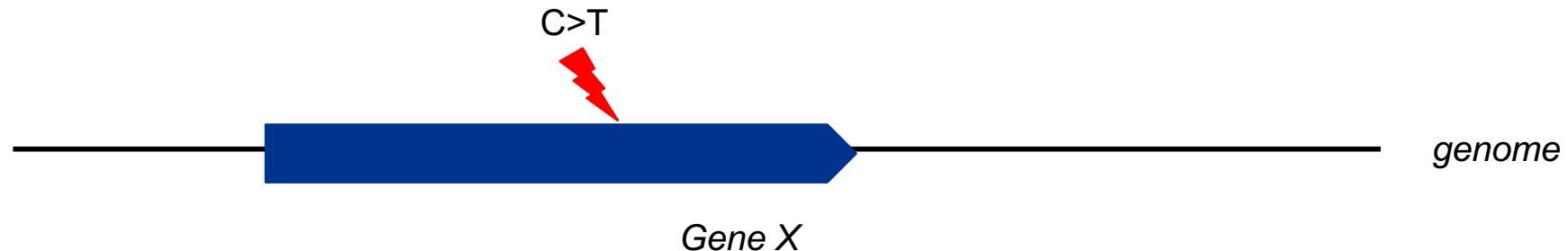
# Cancer – a disease of the genome



Stratton MR, EMBO Molecular Medicine (2013)

# Why is the variant interpretation framework so important?

# Importance of variant interpretation (I)

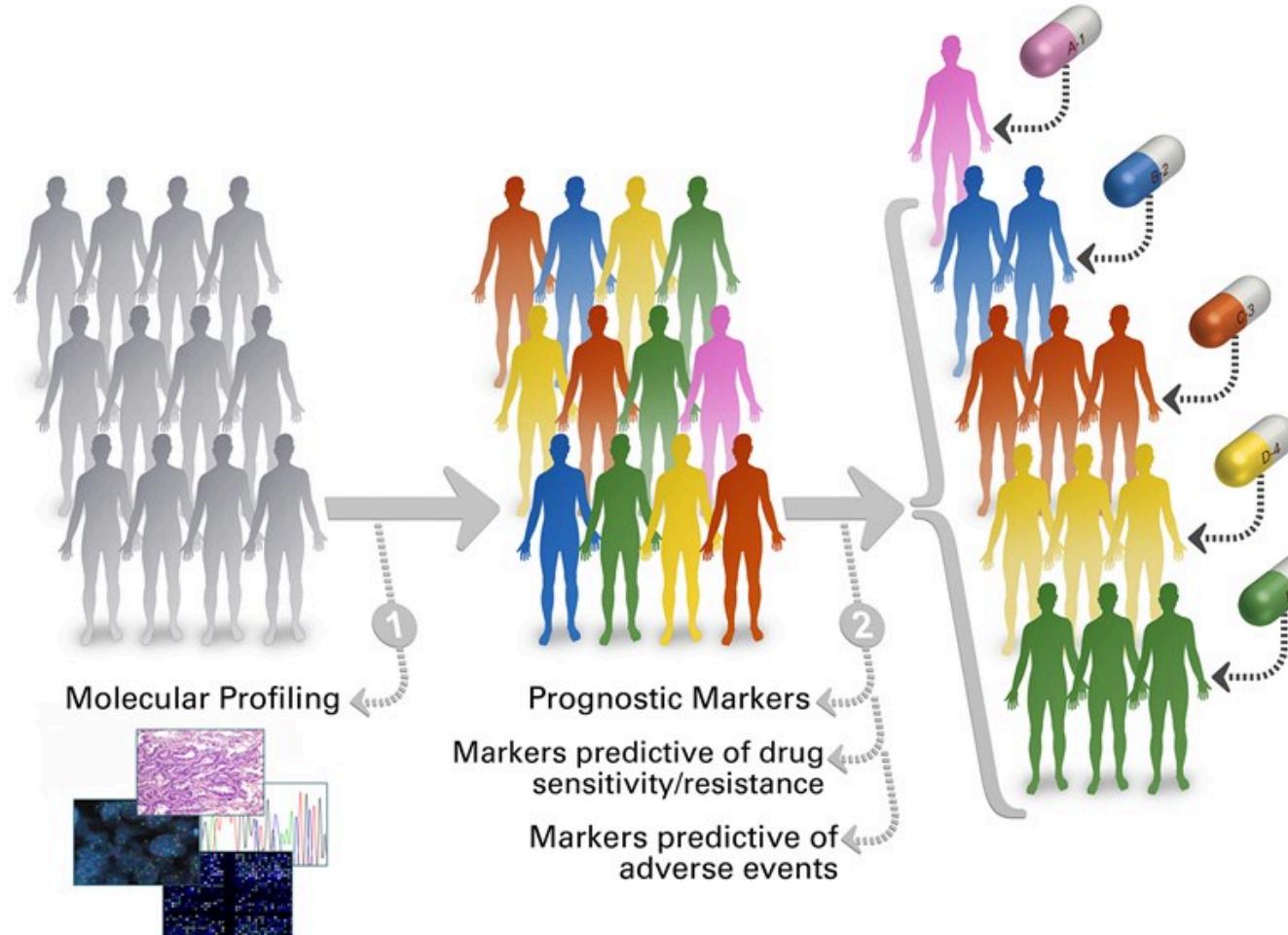


1. Which genes are affected by variants?
2. For a given gene variant, what is the consequence for the encoded protein?

# Importance of variant interpretation (II)

1. Somatic variants: clinical significance
  - **Treatment options** / diagnosis / prognosis
2. Germline variants: cancer predisposition
  - Surveillance / diagnosis / prognosis / **treatment options**
3. Somatic (and germline) variants
  - **Origins and causes of tumors**

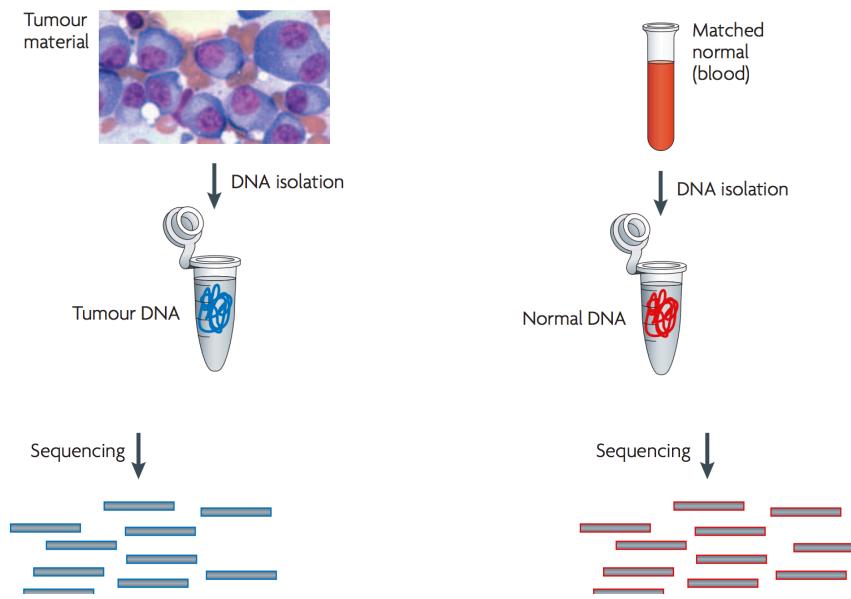
# Precision cancer medicine



# Lecture learning outcomes

- The importance of variant interpretation
- Tumor-only vs tumor-control sequencing design – impact on variant interpretation
- Determination of variant consequence – not straightforward
- Reporting tumor aberrations
  - – Personal Cancer Genome Reporter (PCGR)
- Reporting germline variants
  - – Cancer Predisposition Sequencing Reporter (CPSR)

# Cancer genome sequencing – ideal setup

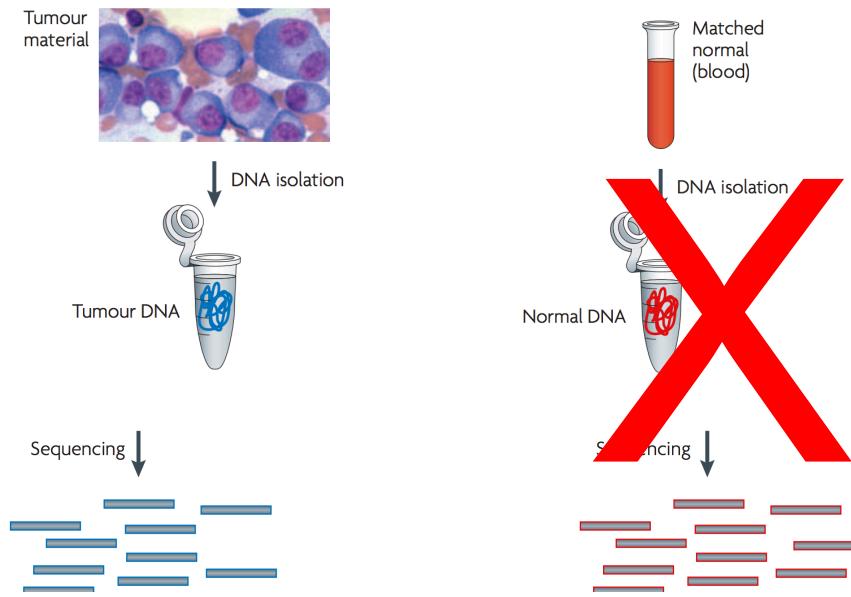


# Cancer genome sequencing: ideal setup



- Sequence tumor
- Sequence normal
- Align sequence reads to reference genome
- Somatic variants are found exclusively in the tumor

# Cancer genome sequencing: tumor-only

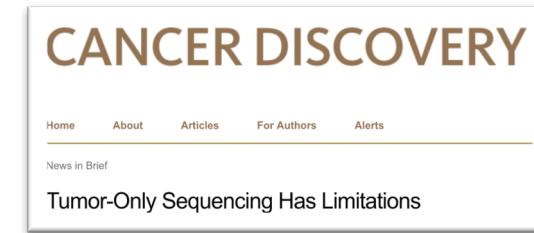
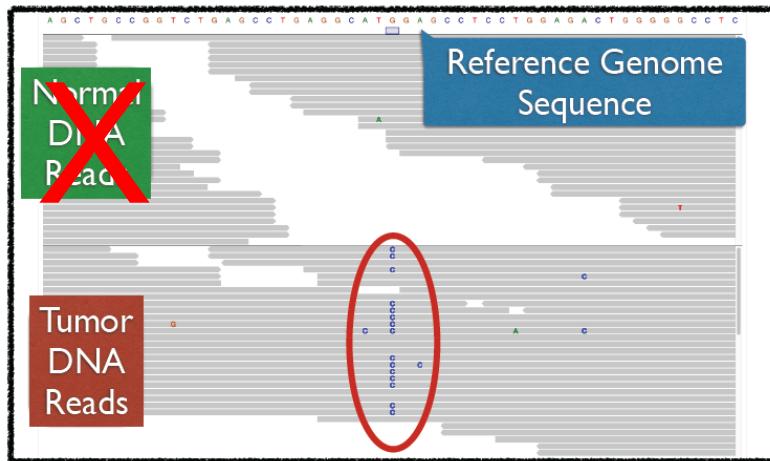


# Cancer genome sequencing: tumor-only



- Sequence tumor
- Align sequence reads to reference genome
- Variants will be a mix of both germline variants and somatic variants
- “Identify” somatic events by exclusion of known germline variants found in public databases

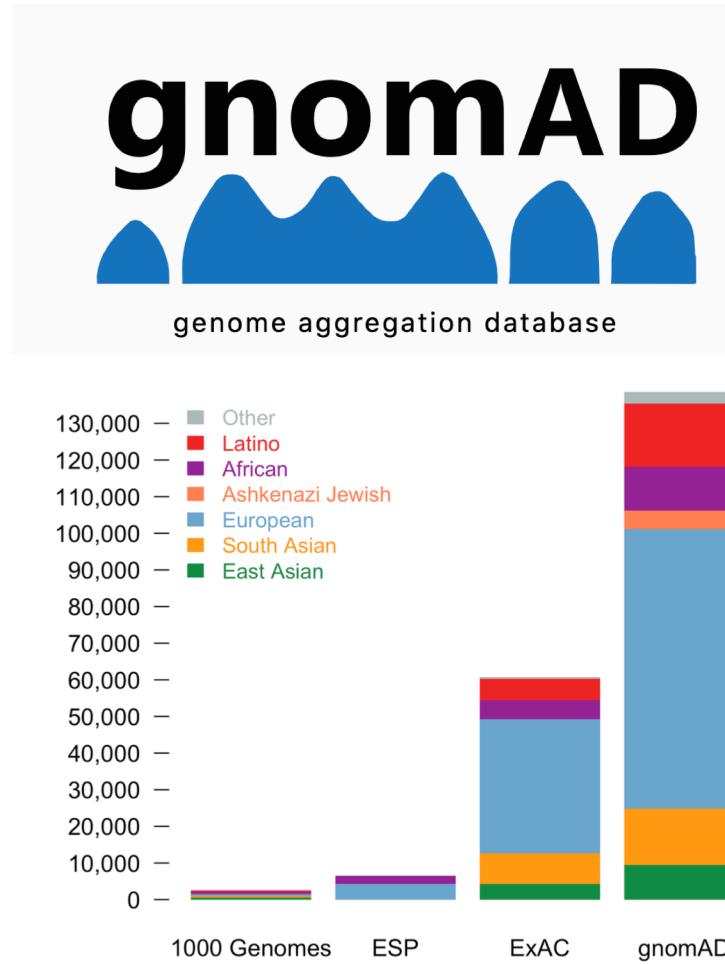
# Cancer genome sequencing: tumor-only



- Each individual is estimated to carry an extensive set of rare variants (i.e. *singletons*)
- Ethnic subpopulations are under-represented in germline variant databases

# Germline variant database: gnomAD

- genome Aggregation Database
- Harmonizes germline variant both exome and genome sequencing data from a wide variety of large-scale sequencing projects
- Freely available to the scientific community
- ~125,000 WES samples
- ~16,000 WGS samples



# Germline variant database: norgene

## Norwegian Germline variants browser



[Explore the Norwegian Germline variations database](#)

Norwegian Cancer Genomics Consortium's database of normal variation in the Norwegian population. This database currently contains 1 547 121 individual variants coming from 1590 normal chromosomes of cancer patients. Genome build hg19/GRCh37.

[Enter](#)

Based on vcf-miner from Mayo Clinic  
The funding was provided by the [Center for Individualized Medicine](#) at Mayo Clinic.  
[Terms and Conditions of Use](#)

*norgene.no*

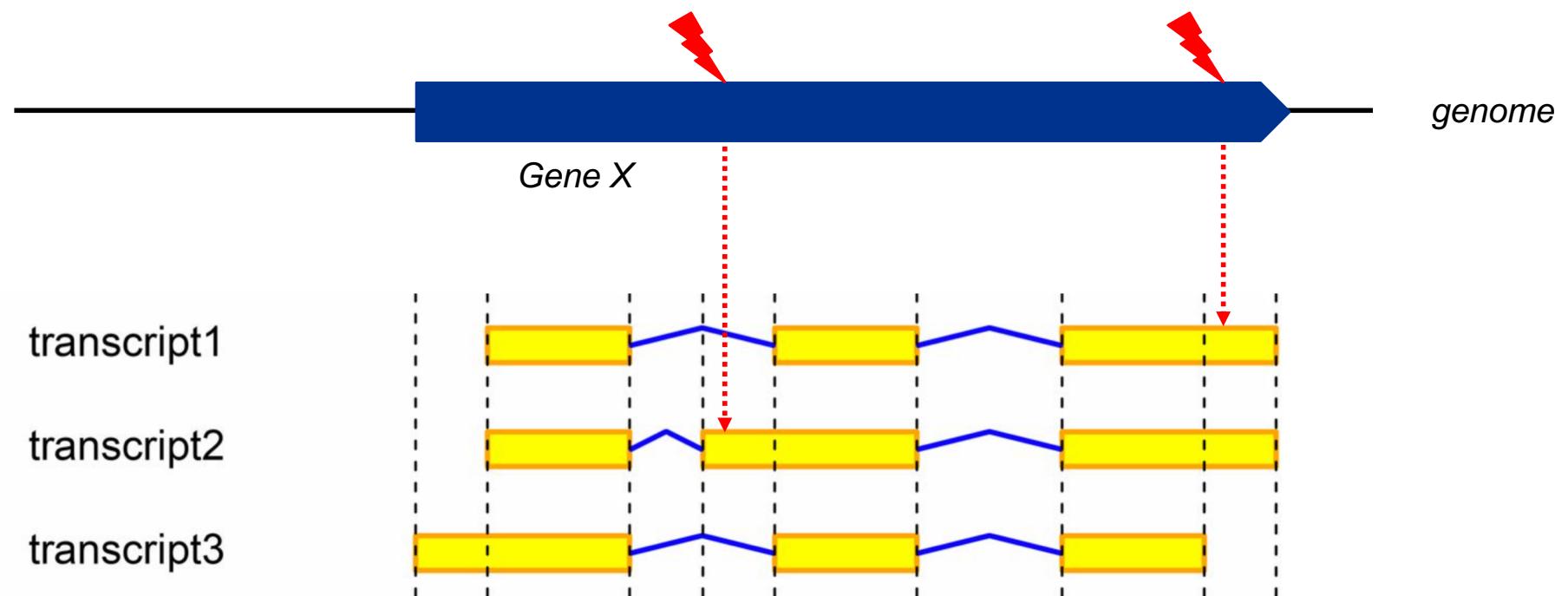
# Variants have been found – now what?

ACTG**C**CTACGTCTACCGTCGACTTCAAATCG**C**TTAACCCGTACTCCCATTGCTACTGC  
ATCTCGGGTTAACTCGACGTTT**T**CATGCATGTGTGCACCCCAATATATATGCA**A**CTT  
TTGTGCACCTCTGTCA CGCGCGAGTTGGCACTGTCGCCCTGTGTGCATGTGCACTGT  
CTC**T**CGCTGCACTGCCTACGTCTACCGTCGACTTCAAATCG**C**TTAACCCGTACTCCC  
ATGCTACTGCATCTCGGGTTAACTCGACGTTT**G**CATGCATGTGTGCACCCCAATATA  
TATGCA**A**CTTTGTGCACCTCTGTCA CGCGCGAGTTGGCACTGTCGCCCTGTGTGCA  
TGTGCACTGTCTC**T**CGAGTTT**G**CATGCATGTGTGCACCTCTGTTACGTCT



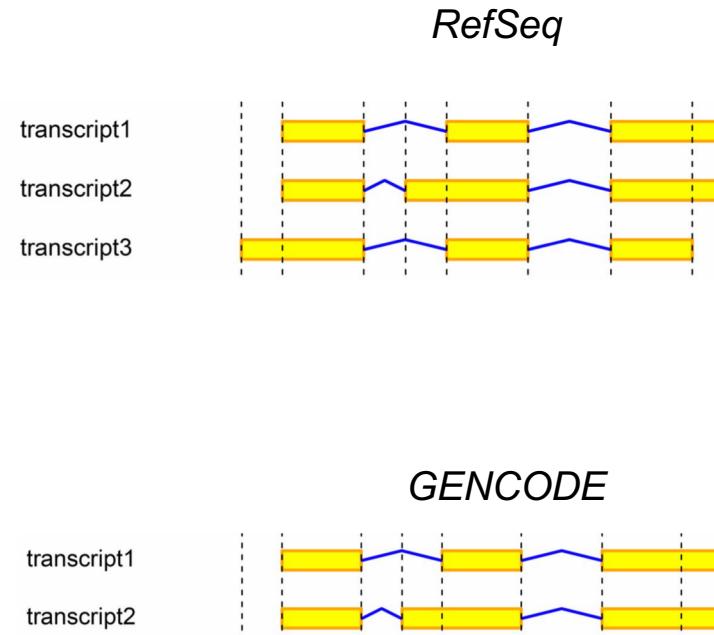
# Variant interpretation challenges (I)

- A gene consists of multiple transcript isoforms



# Variant interpretation challenges (II)

- Several transcript databases
  - RefSeq
  - Ensembl
  - GENCODE
- Choice of transcript database impacts variant consequence/annotation
- Frequent strategy: Report variant consequence in most commonly expressed isoform (i.e. ***principal*** isoform)



# Variant calling vs. interpretation

Biopsy

DNA-seq (exome)

- Single nucleotide variants
- Insertions/deletions
- Copy number aberrations
- (Rearrangements)



- *Oncogene/tumor suppressor gene mutations*
- *Contribution of mutational signatures*
- *Microsatellite instability status*
- *BRCA*ness status
- *Mutational burden*
- *Clinical trial matcher*
- *Predisposing germline variants*

2. *Molecular profiling – genomics*



3. *Measures to inform precision therapy*

# Variant calling vs. interpretation

## Core NGS bioinformatics

- Robust sequencing pipelines
- Variant calling

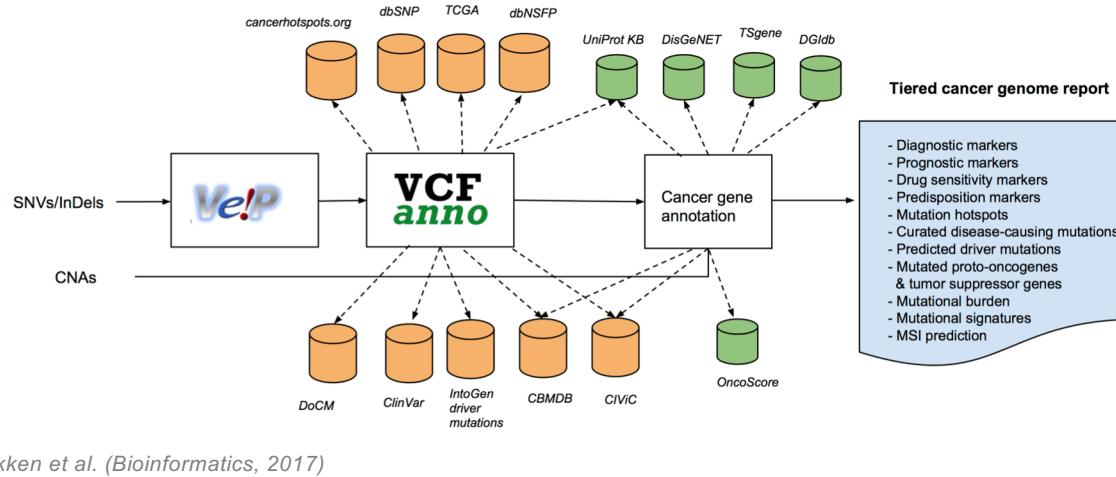
## Translational bioinformatics

- Personal Cancer Genome Reporter
- Cancer Predisposition Sequencing Reporter

2. *Molecular profiling – genomics*

3. *Measures to inform precision therapy*

# Personal Cancer Genome Reporter (PCGR)



- What is it?
- Knowledge resources
- Input data
- Output data

# Personal Cancer Genome Reporter (I)

- In brief: A reporting engine for clinical interpretation of tumor genomes

Nakken et al., Bioinformatics, 2017

Cancer Genome Report by PCGR   SNVs and InDels ▾   TMB and MSI   Mutational signatures   Clinical trials   Settings & Docs   TCGA-BR-8078-01A | Esophagus/Stomach | Tumor-Control | WES

### SNVs/InDels

Overview

2379 Total variants   1665 SNVs   714 InDels   0 TIER 1 variants   2 TIER 2 variants

#### Variants per tier

- TIER 1 : 0
- TIER 2 : 2
- TIER 3 : 77
- TIER 4 : 1395
- NONCODING : 905

#### Allelic support plot

The prioritization of SNV and InDels found in the tumor sample is done according to a four-tiered structure, adopting the joint consensus recommendation by AMP/ACMG Li et al., 2017.

- TIER 1:** Variants of strong clinical significance - constitutes variants linked to predictive, prognostic, or diagnostic biomarkers in the CIVIC database and the Cancer Biomarkers Database that are
  - Found within the same tumor type/class as specified by the user, AND
  - Of strong clinical evidence (i.e. part of guidelines, validated or discovered in late clinical trials [CIVIC evidence levels A/B])
  - overlap between variants in the tumor sample and reported biomarkers must occur at the

#### Global variant datatable - filters

- Filtering on sequencing depth/variant allelic fraction depends on input provided by user
- Filtering performed here will only apply to the datatable and not any other visualizations presented in this page

Tier   Consequence   Call confidence   Allelic fraction tumor   Sequencing depth tumor

#### Global variant datatable

SYMBOL	CONSEQUENCE	PROTEIN_CHANGE	VARIANT_CLASS	TIER	GENOMIC_CHANGE
1 KRAS	missense_variant	p.Gly13Asp	SNV	TIER 2	12:g.25245347C>T
2 ERBB3	missense_variant	p.Val104Met	SNV	TIER 2	12:g.56085070G>A
3 BRCA2	frameshift_variant	p.Glu2981ArgfsTer37	insertion	TIER 3	13:g.32379495C>CA
4 BRCA2	frameshift_variant	p.Thr3085AsnfsTer26	insertion	TIER 3	13:g.32380135G>GA
5 BRAF	frameshift_variant	p.Pro403LeufsTer8	deletion	TIER 3	7:g.140783126AG>A
6 ARID1A	frameshift_variant	p.Gln1519ArgfsTer8	deletion	TIER 3	1:g.26774776GC>G

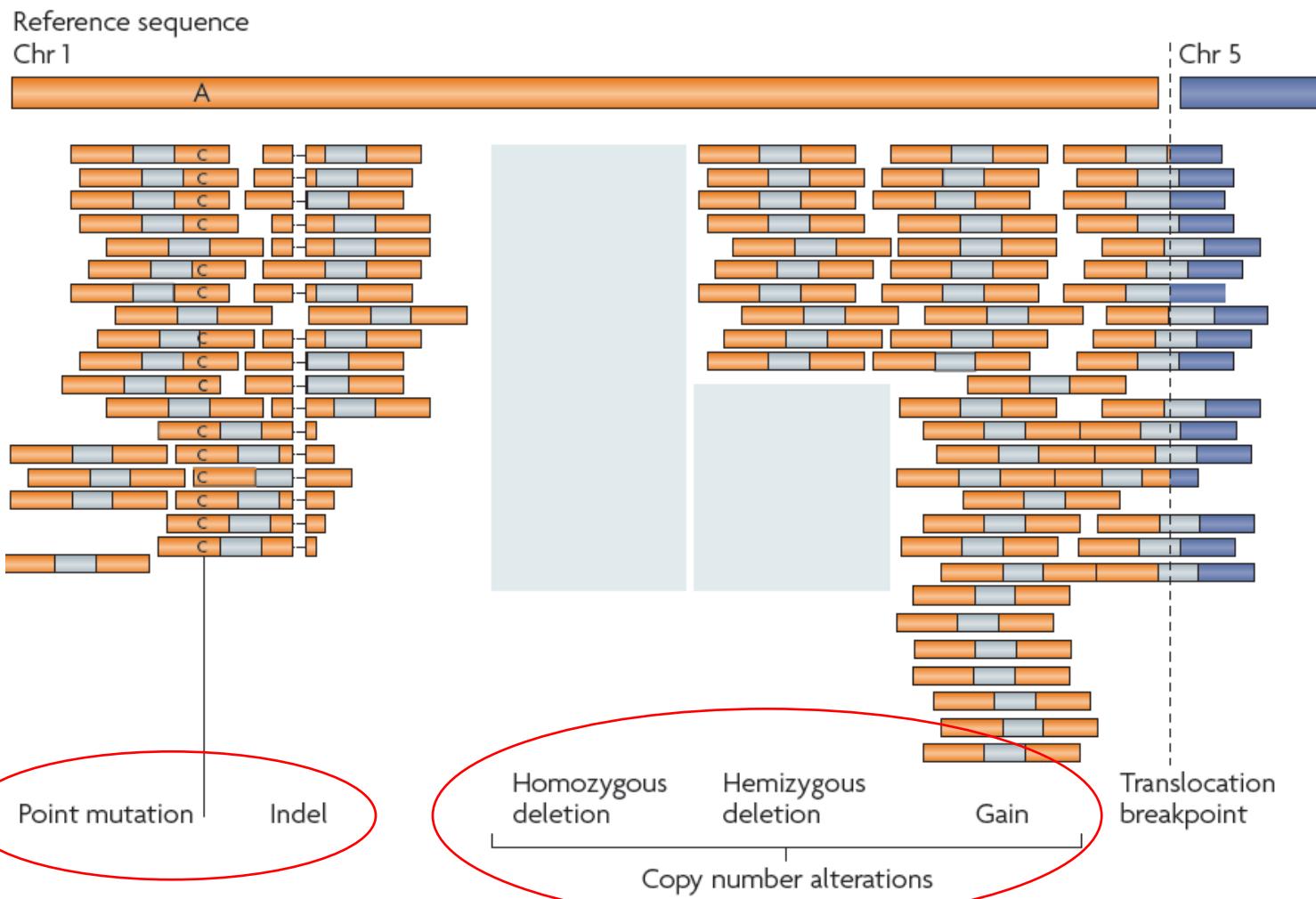
Showing 1 to 8 of 2,379 entries   Previous   1   2   3   4   5   ...   298   Next

# Personal Cancer Genome Reporter (II)

- Targeted cancer drugs
- Known biomarkers for prognosis and diagnosis
- Known biomarkers for drug sensitivity/resistance
- Gene-tumor type associations
- Mutational hotspots
- Predicted driver mutations
- Signaling pathways
- Proto-oncogenes/tumor suppressors
- Prediction of variant effect on protein function



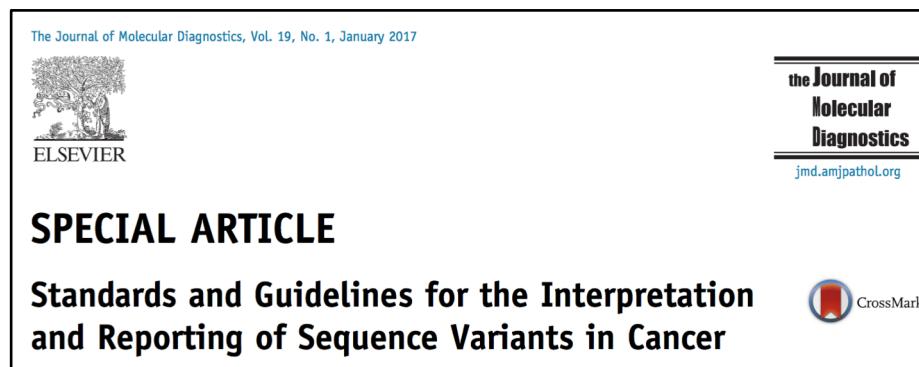
# Personal Cancer Genome Reporter (III)



Meyerson et al. Nat Rev Genet 2010

# Interpretation concept: tier structure

- SNVs/InDels
  - *Tier 1: Variants of strong clinical significance*
  - *Tier 2: Variants of potential clinical significance*
  - Tier 3: Other cancer gene mutations
  - Tier 4: Other coding mutations
  - Tier 5: Non-coding mutations
- CNAs
  - Amplifications and homozygous deletions (as defined by the user)
  - Aberrations of **strong and potential** clinical significance



# Interpretation concept: evidence item

- Each actionable variant associated with one or more evidence item
- A piece of information that has been manually curated from trustable medical literature about a variant or genomic 'event' that has implications in cancer diagnosis (*aka* molecular classification), prognosis, or predictive response to therapy
  - Therapeutic context
  - Evidence level
  - Tumor type
  - Evidence type (Prognostic, Predictive etc.)
  - Etc.

**Copy number aberrations as biomarkers**

A total of 1 aberrations are associated with clinical evidence items in the database for clinical interpretations of variants in cancer, CIViC, with the following number of evidence items:

- Predictive: 23 evidence items
- Prognostic: 0 evidence items
- Diagnostic: 0 evidence items

Aberrations of strong clinical significance

Predictive biomarkers      Prognostic biomarkers      Diagnostic biomarkers

Cancer type      Therapeutic context  
Trastuzumab

Clinical significance      Log-ratio  
1.78

Evidence level

The table below lists all variant-evidence item associations:

CSV      Excel      Search:

SYMBOL	CANCER_TYPE	CNA_TYPE	EVIDENCE_LEVEL	CLINICAL_SIGNIFICANCE
+ 1 ERBB2	Her2-receptor Positive Breast Cancer	gain	A: Validated	Sensitivity
+ 22 ERBB2	Her2-receptor Positive Breast Cancer	gain	B: Clinical evidence	Sensitivity
+ 23 ERBB2	Her2-receptor Positive Breast Cancer	gain	B: Clinical evidence	Sensitivity

Showing 1 to 3 of 3 entries (filtered from 23 total entries)

Previous      1      Next

# Personal Cancer Genome Reporter (IV)

Cancer Genome Report by PCGR   SNVs and InDels ▾   TMB and MSI   Mutational signatures   Clinical trials   Settings & Docs   TCGA-BR-8078-01A | Esophagus/Stomach | Tumor-Control | WES   

**SIGNATURES**   **1665**      **MMR deficiency**   **99.3**   **0**

Mutational Signatures (SBS)   SNVs eligible for analysis   Most dominant aetiology   Accuracy of signature fitting (%)   High confident kataegis events

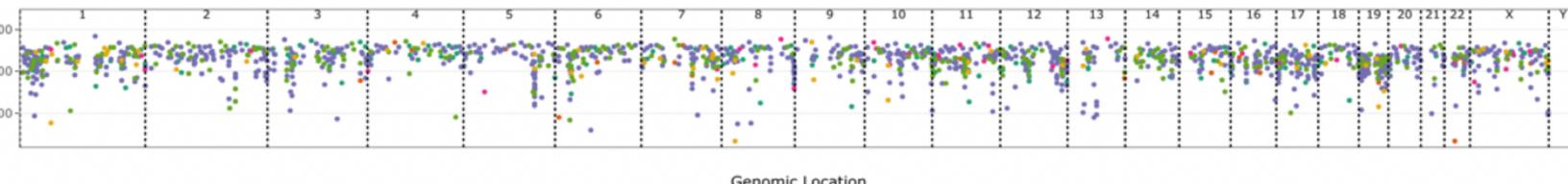
**Mutational signatures - aetiology contributions**



signature_id	contribution	group	aetiology	comments
3	17.3%	MMR deficiency	Defective DNA mismatch repair.	SBS44 is one of seven mutational signatures associated with defective DNA mismatch repair (MSI) and is often found in the same samples as other MSI associated signatures: SBS6, SBS14, SBS15, SBS20, SBS21, and SBS26.
4	11.8%	MMR deficiency	Defective DNA mismatch repair.	SBS15 is one of seven mutational signatures associated with defective DNA mismatch repair (MSI) and is often found in the same samples as other MSI associated signatures: SBS6, SBS14, SBS20, SBS21, SBS26, and SBS44.

Showing 1 to 10 of 11 entries   Previous   1   2   Next

**Mutational context frequency**   **Genomic distribution - rainfall**   **Kataegis events**



# Personal Cancer Genome Reporter (IV)

Cancer Genome Report by PCGR   SNVs and InDels ▾   sCNA ▾   TMB and MSI   Mutational signatures   Clinical trials   Settings & Docs   TCGA-14-0866-01B | CNS/Brain | Tumor-Control | WES   

**sCNA**

Overview

**7**  
Copy number gains



**50**  
Copy number losses



**1**  
TIER 1 biomarkers



**6**  
TIER 2 biomarkers



**Copy number segments - filters**

The following user-defined thresholds determine copy number aberrations shown here:

- Copy number amplifications** : Log(2) ratio  $\geq 0.4$
- Homozygous deletions** : Log(2) ratio  $\leq -0.4$

A total of **57** unfiltered aberration segments satisfied the above criteria.

- A total of **57** copy number segments satisfy the current filtering criteria.

**Log-ratio**  


**Cytoband**

**Event type**

**Copy number segments**

CSV   Excel   Search:

SEGMENT	SEGMENT_LENGTH_MB	CYTOBAND	LOG_R	EVENT_TYPE
<a href="#">chr7:54942675-55577616</a>	0.63494	chr7:p11.2	<b>3.471</b>	focal
<a href="#">chr7:24039878-24040383</a>	0.0005	chr7:p15.3	<b>2.207</b>	focal
<a href="#">chr7:705284-24035127</a>	23.32984	chr7:p22.3 - p15.3	<b>0.476</b>	broad
<a href="#">chr7:24126318-52647610</a>	28.52129	chr7:p15.3 - p12.1	<b>0.465</b>	broad
<a href="#">chr7:82610187-158385118</a>	75.77493	chr7:q21.11 - q36.3	<b>0.462</b>	broad
<a href="#">chr7:63112265-82605569</a>	19.4933	chr7:q11.21 - q21.11	<b>0.453</b>	focal
<a href="#">chr20:455764-62219837</a>	61.76407	chr20:p13 - q13.33	<b>0.452</b>	broad

Showing 1 to 10 of 57 entries   Previous   1 2 3 4 5 6 Next

**Key findings**

- Proto-oncogenes subject to amplifications: 28**
- Tumor suppressor genes subject to homozygous deletions: 19**
- Other drug targets subject to amplification: 18**

**Documentation**

Somatic copy number aberrations identified in the tumor sample are classified into **two main tiers**:

- TIER 1: Aberrations of strong clinical significance** - constitutes amplified/lost genes linked to predictive, prognostic, or diagnostic biomarkers in the [Civic database](#) and the [Cancer Biomarkers Database](#) that are
  - Found within the same tumor type/class as specified by the user, **AND**
  - Of strong clinical evidence (i.e. part of guidelines, validated or discovered in late clinical trials ([Civic evidence levels A/B](#)))
- TIER 2: Aberrations of potential clinical significance** - constitutes amplified/lost genes linked to predictive, prognostic, or diagnostic biomarkers in the [Civic database](#) and the [Cancer Biomarkers Database](#) that are either
  - Of strong clinical evidence in other tumor types/classes than the one specified by the user, **OR**
  - Of weak clinical evidence (early trials, case reports etc. ([Civic evidence levels C/D/E](#)))) in the same tumor type/class as specified by the user

Included in the report is also a complete list of [all oncogenes subject to amplifications](#), [tumor suppressor genes subject to homozygous deletions](#), and [other drug targets subject to amplification](#).

# Personal Cancer Genome Reporter (IV)

Cancer Genome Report by PCGR   SNVs and InDels ▾   sCNA ▾   TMB and MSI   Mutational signatures   Clinical trials   Settings & Docs   TCGA-BR-8078-01A | Esophagus/Stomach | Tumor-Control | WES  

**Clinical trials (Beta)**

Molecularly targeted trials	116 Not yet recruiting	437 Recruiting	10 Enrolling by invitation	67 Active, not recruiting	1 Unknown status
-----------------------------	---------------------------	-------------------	-------------------------------	------------------------------	---------------------

**Molecularly targeted trials - filters**

(e.g. *inclusion/exclusion criteria*) attempts to highlight the presence of established molecular biomarkers in cancer and relevant therapeutic contexts.

Condition (cancer subtype)	Phase
<input type="text"/>	<input type="text"/>
Status	Gender
<input type="text"/>	<input type="checkbox"/> All <input type="checkbox"/> Female <input type="checkbox"/> Male
Drug(s)	Minimum age
<input type="text"/>	
Drug target(s)	Maximum age
<input type="text"/>	
Therapeutic context mentions (text-mined)	Metastases mentions (text-mined)
<input type="text"/>	<input type="text"/>
Biomarker mentions (text-mined)	
<input type="text"/>	

**Molecularly targeted trials**

CSV   Excel   Search:

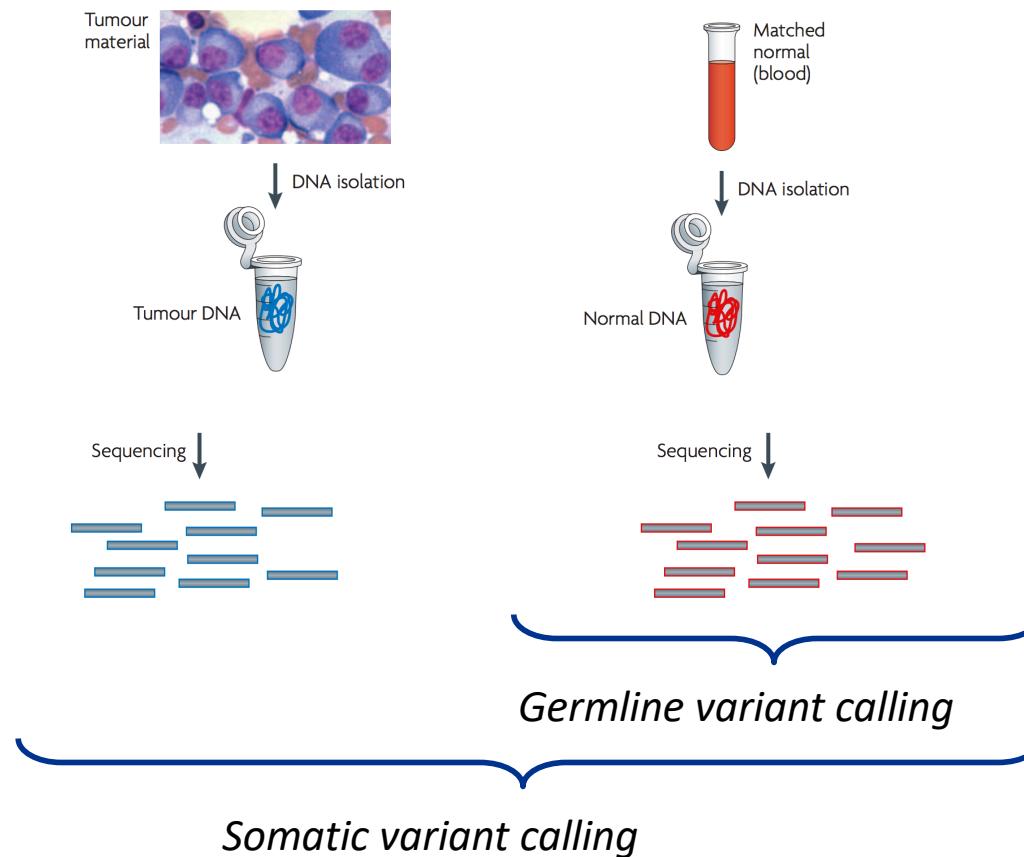
NCT_ID	TITLE	OVERALL_STATUS	CONDITION	KEYWORD	INTERVENTION	PHASE	START_DATE
464 NCT02734004	A Phase I/II Study of MEDI4736 in Combination With Olaparib in Patients With Advanced Solid Tumors.	Active, not recruiting	Malignant Gastric Neoplasm	ER Positive, HER2 Negative, HR deficiency/PARPi, Immunotherapy, PR Positive, Radiotherapy	Bevacizumab, Durvalumab, Olaparib	1.5	2016-03-17

PRIMARY\_COMPLETION\_DATE 2022-08-05  
CONDITION\_RAW Malignant Gastric Neoplasm  
INTERVENTION\_RAW Bevacizumab, Durvalumab, Olaparib  
INTERVENTION\_TARGET CD274, PARP1, PARP2, PARP3, VEGFA  
BIOMARKER\_INDEX ATM mutation, BARD1 mutation, BRCA1 mutation, BRCA2 mutation, BRIP1 mutation, CDK12 mutation, CHEK1 mutation, HER2 gene mutation, HER2 mutation, HER2 negative  
METASTASES\_INDEX Bone Metastases|Brain Metastases  
GENDER All  
MINIMUM\_AGE 18  
MAXIMUM\_AGE 100

Showing 1 to 1 of 1 entries (filtered from 644 total entries)

Previous   1   Next

# Cancer genome sequencing



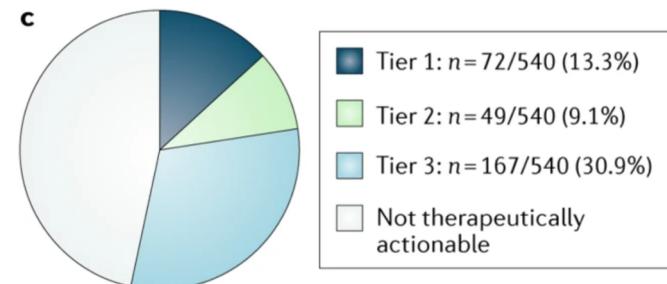
# Cancer Predisposition Interpretation (I)

- Goal: Identify pathogenic variants (germline) conferring increased risk of tumor development
- Why important?
  - Implement surveillance and risk-reducing interventions
  - May impact type of surgery (radical /conservative)
  - Targeted therapy implications
    - BRCA (PARP)



Review Article | Published: 19 February 2019

## Therapeutic implications of germline genetic findings in cancer



*Clinical actionability - TCGA*

# Cancer Predisposition Sequencing Reporter

- Flexible tool for interpretation of sequencing screens for cancer predisposition
- Utilizes Docker technology for software encapsulation
- Tier structure
  - Pathogenic
  - Likely pathogenic
  - Unclassified variants
  - Likely Benign
  - Benign
- Automated pathogenicity classification
  - Predicted loss-of-function
  - population allele frequency
  - ++
- Incidental findings can also be reported



<https://github.com/sigven/cpsr>

# Cancer Predisposition Sequencing Reporter

- Flexible tool for interpretation of sequencing screens for cancer predisposition
- Utilizes Docker technology for software encapsulation
- Tier structure
  - Pathogenic
  - Likely pathogenic
  - Unclassified variants
  - Likely Benign
  - Benign
- Automated pathogenicity classification
  - Predicted loss-of-function
  - population allele frequency
  - ++
- Incidental findings can also be reported

**Class 5 - Pathogenic variants**

A total of n = 5 variants are registered with a *Pathogenic* clinical significance in ClinVar.  
A total of n = 4 *non-ClinVar* variants (i.e. not registered in ClinVar) are classified with a *Pathogenic* significance by CPSR (ACMG criteria - based on population frequency and variant effect).

ClinVar    Non-ClinVar

Consequence   
Genotype  heterozygous  
Gene  POLD1 POLE

CPSR classification (ACMG criteria codes)   
CPSR pathogenicity score   
MAF gnomAD (Non-Finnish European non-cancer subset)

CSV Excel Search:

SYMBOL	SOURCE	CONSEQUENCE	PROTEIN_CHANGE	GENOTYPE	GENE_NAME
1	POLE	Other	frameshift_variant	p.Lys1170AsnfsTer49	heterozygous DNA polymerase epsilon, catalytic subunit
4	POLD1	Other	frameshift_variant	p.Arg180GlyfsTer3	heterozygous DNA polymerase delta 1, catalytic subunit

Showing 1 to 2 of 2 entries (filtered from 4 total entries) Previous  Next

<https://github.com/sigven/cpsr>

# Cancer Predisposition Sequencing Reporter

- Flexible tool for interpretation of sequencing screens for cancer predisposition
- Utilizes Docker technology for software encapsulation
- Tier structure
  - Pathogenic
  - Likely pathogenic
  - Unclassified variants
  - Likely Benign
  - Benign
- Automated pathogenicity classification
  - Predicted loss-of-function
  - population allele frequency
  - ++
- Incidental findings can also be reported

**Genomic biomarkers**

- Variants (class 4/5) in the query sample that overlap with reported clinical biomarkers from the [database for clinical interpretations of variants in cancer, CIVC](#) are considered. Note that several variants in the query can overlap the same existing biomarker, given that biomarkers are reported at different resolutions (variant/gene level). Total number of clinical evidence items that coincide with query variants:
  - Predisposing: 1 evidence items
  - Predictive: 2 evidence items
  - Prognostic: 0 evidence items
  - Diagnostic: 0 evidence items

Predisposing      Predictive      Prognostic      Diagnostic

Cancer type      Gene

Clinical significance      Biomarker mapping

Evidence level      Therapeutic context

The table below lists all variant-evidence item associations:

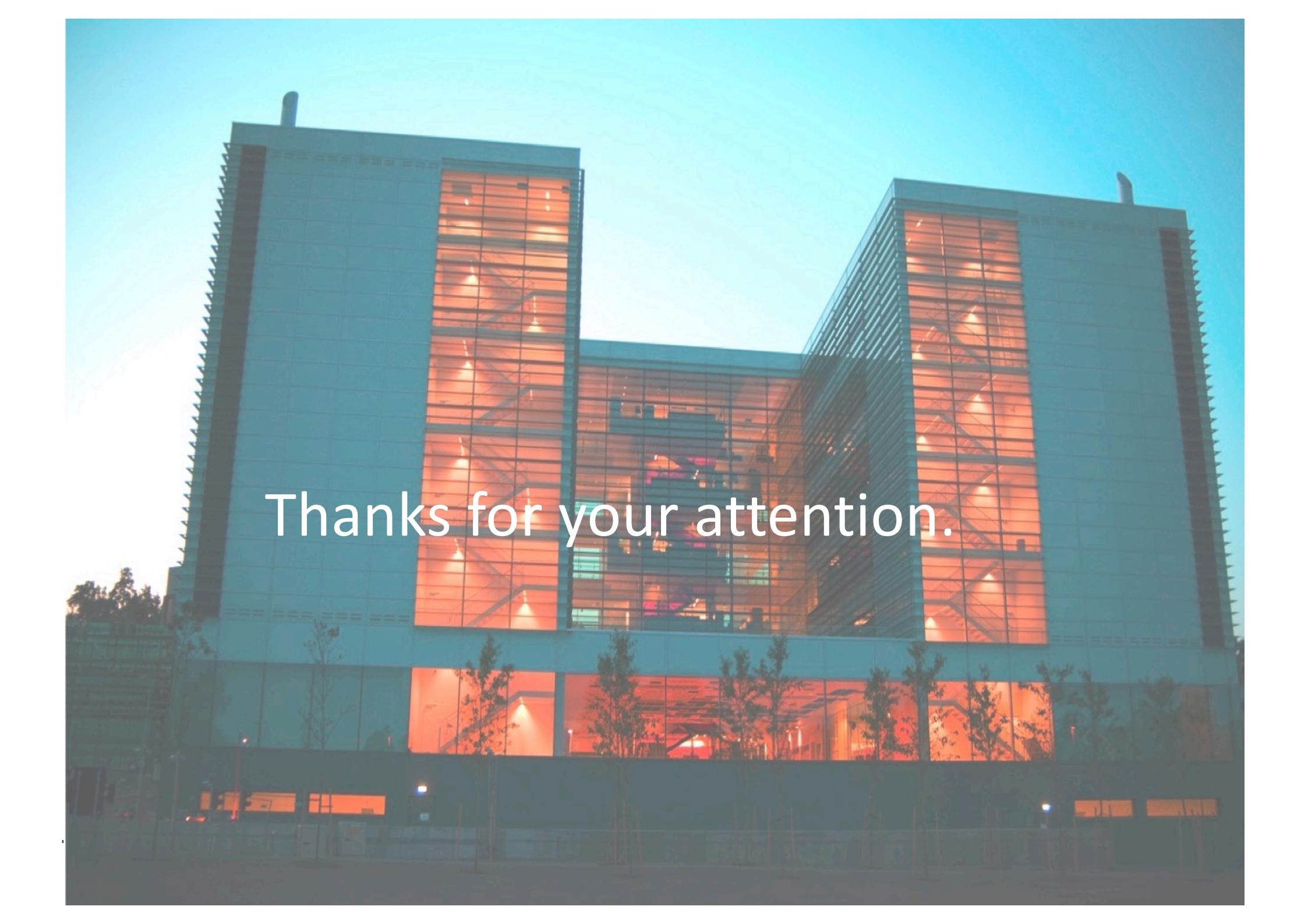
CSV      Excel      Search:

SYMBOL	GENE_NAME	CANCER_TYPE	CLINICAL_SIGNIFICANCE	EVIDENCE_LEVEL
1 NF1	neurofibromin 1	Plexiform Neurofibroma	Sensitivity/Response	B: Clinical evidence
2 POLE	DNA polymerase epsilon, catalytic subunit	Glioblastoma Multiforme	Sensitivity/Response	C: Case study

Showing 1 to 2 of 2 entries      Previous            Next

# Summary

- Clinical variant interpretation is critical for implementation of precision cancer medicine
- Sequencing design will impact interpretation
- Choice of transcript database will affect variant consequence annotation
- An increasing number of resources are being erected to facilitate clinical interpretation of cancer genomes
- Key concepts: tier structure & evidence items
- Interpretation of the germline background of cancer patients adds an additional dimension for clinical interpretation

A photograph of a modern architectural complex at dusk or night. The building features a grid-like facade with large glass windows that are brightly lit from within, casting a warm orange glow. The sky is a clear, pale blue. In the foreground, there are some small trees and a dark, flat surface.

Thanks for your attention.