



**Emerging Approaches for Tumor Analyses in Epidemiological Studies Workshop**

---

# **Anthology of unusual patterns of somatic mutations in cancer genomes**

Ludmil B. Alexandrov, PhD

Associate Professor of Bioengineering  
Associate Professor of Cellular and Molecular Medicine

[L2alexandrov@health.ucsd.edu](mailto:L2alexandrov@health.ucsd.edu)

January 19, 2023



**UC San Diego**  
MOORES CANCER CENTER

# Faculty Disclosure -- Conflict of Interest

- Chief scientific officer, compensated consultant, and equity holder in io9, LLC.
- Inventor of a US Patent 10,776,718 for source identification by nonnegative matrix factorization.
- Inventor of U.S. provisional patent applications:
  - Clustered mutations for the treatment of cancer (63/289,601).
  - Artificial intelligence architecture for predicting cancer biomarker (63/269,033).
  - Methods and systems for detecting homologous recombination deficiency in cancer therapies (63/366,392).
  - Drugs for treating head and neck cancers with chromosome 9p loss (63/367,846).
  - Genetically-defined immune-checkpoint inhibitor resistance in aggressive precursors of HPV- head and neck squamous cancer (63/412,835).
- Spouse employed by Hologic, Inc. a publicly traded medical technology company.

# Anthology Outline

## Stories of The Past

- Mutational signatures as a machine learning approach that allows detecting the *unusual patterns of somatic mutations*.
- Utilizing mutational signatures for developing cancer prevention strategies.
- Utilizing mutational signatures for understanding failed DNA repair and targeted cancer treatment.

## Anecdotes of The Present

- Utilizing clustered mutations for understanding cancer development and evolution.
- The repertoire of copy-number signatures in human cancer.
- A novel machine learning approach for detecting homologous recombination deficiency.

## Dreams of The Future

- Beyond genomics: Utilizing AI for addressing inequalities of cancer diagnosis

HRD







# Stories of The Past



# Stories of a simpler, but not so distant, past

- Mutational signatures as a machine learning approach that allows detecting the *unusual patterns of somatic mutations*.
- Utilizing mutational signatures for developing cancer prevention strategies.
- Utilizing mutational signatures for understanding failed DNA repair and targeted cancer treatment.

Mutational signatures as a machine learning approach that allows detecting the *unusual patterns of somatic mutations*



# Somatic Mutations, Mutational Signatures, and Human Cancer

---

- **Somatic mutations** accumulate daily in every cell of the human body. These mutations originate from lifestyle choices, defective cellular machineries, and even from normal cellular processes.
- **Cancer risk** is strongly affected by mutagenesis. Lifestyle choices can cause somatic mutations and significantly affect the risk for developing cancer. For example, from 105 patients with lung squamous cell carcinomas only 1 has never smoked.
- **Mutational signatures** analysis is a machine learning approach that allows detecting the *unusual patterns of somatic mutations* generated by different mutagenic processes from DNA sequencing data.

# Human cancers and their origins

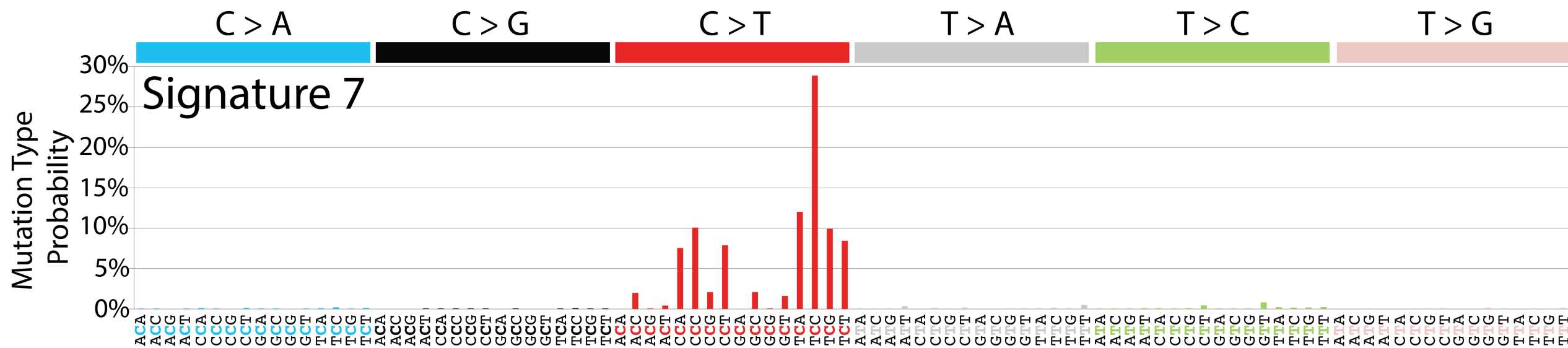


Skin Cancer



Majority caused by UV-light exposure

Predominately **C>T** somatic mutations



**Mutational signature: a molecular fingerprint found in a cancer cell**



# Human cancers and their origins

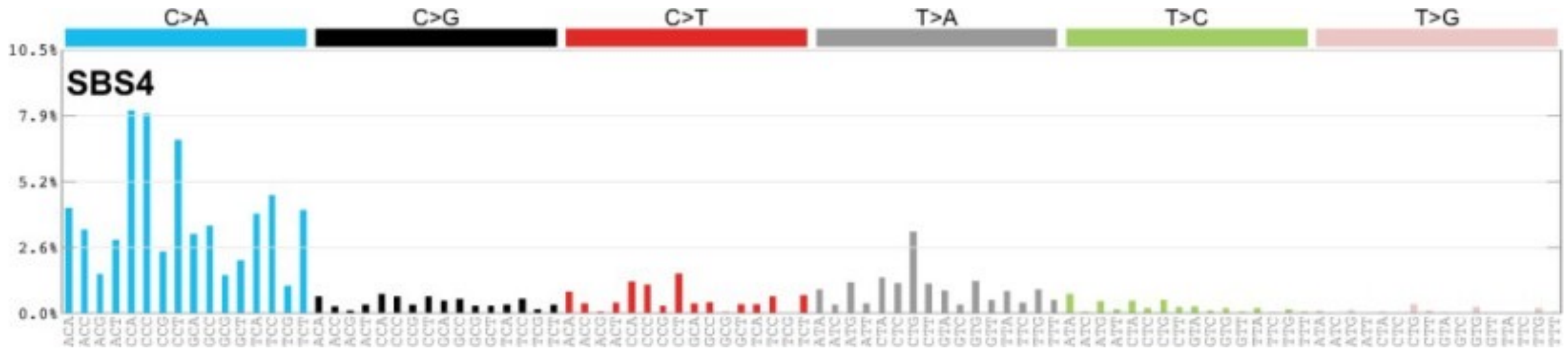


Lung Cancer

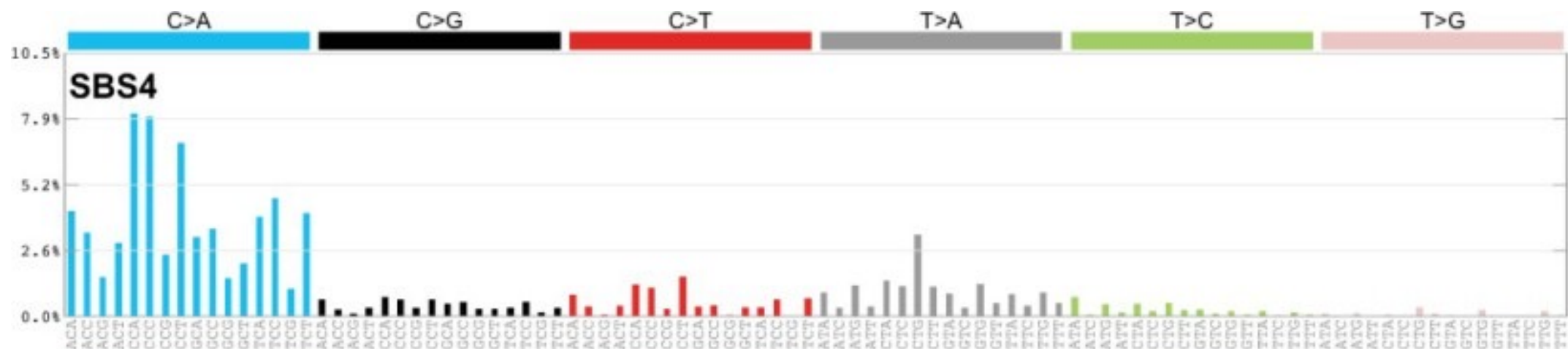


~80% caused by tobacco smoking

Predominately **C>A** somatic mutations



# Quantifying the mutations in a tobacco smoker

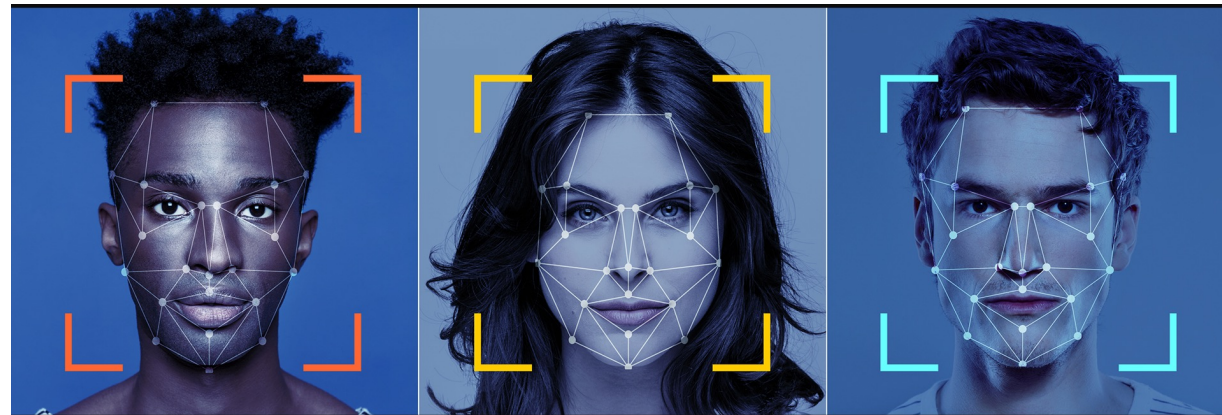




# How do we identify mutational signatures?

## A Suite of Computational Tools

- SigProfilerMatrixGenerator
- SigProfilerMatrixGenerator2
- SigProfilerPlotting
- SigProfilerPlotting2
- SigProfilerSimulator
- SigProfilerExtractor
- SigProfilerClusters
- SigProfilerTopography
- SigProfilerAssignment
- fastNMF



**Develop and utilize state-of-the-art artificial intelligence algorithms for pattern recognition**

RESEARCH ARTICLE

**Nonnegative/Binary matrix factorization with a D-Wave quantum annealer**

Daniel O'Malley<sup>1,2\*</sup>, Velimir V. Vesselinov<sup>1</sup>, Boian S. Alexandrov<sup>3</sup>, Ludmil B. Alexandrov<sup>4,5</sup>

**Development of next-generation of algorithms for quantum computer**

# Mutation Signatures in Human Cancer



## ***Proposed Etiology:***

APOBEC activity  
 Deamination of 5' methycytosine  
 Reactive oxygen species  
 Polymerase  $\eta$  activity  
 POLE mutation

Defective DNA MMR  
 Defective DNA BER  
 Defective DNA HRR

Temozolomide treatment  
 Platinum treatment  
 Azathioprine treatment

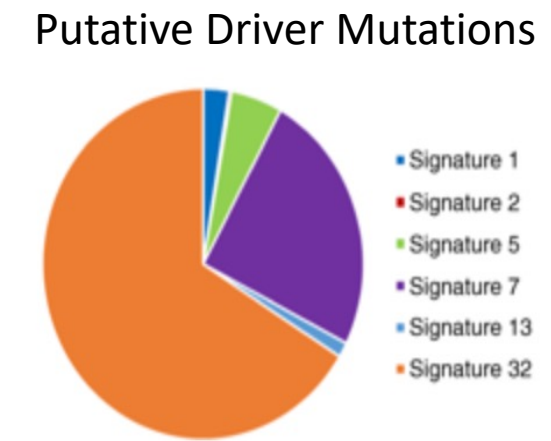
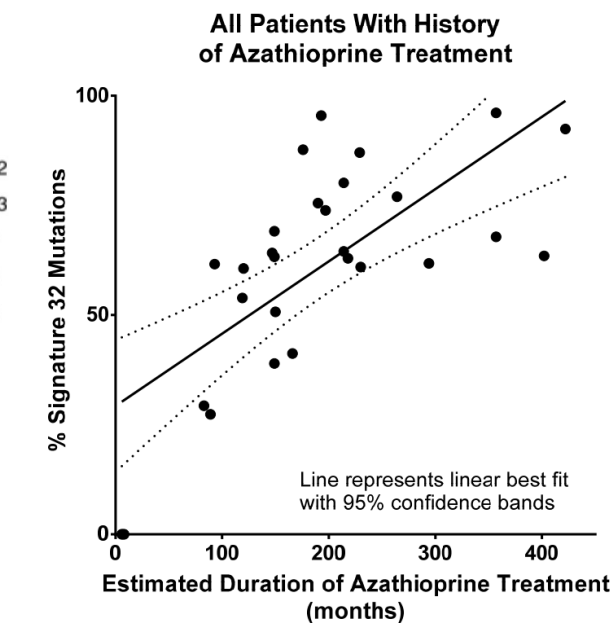
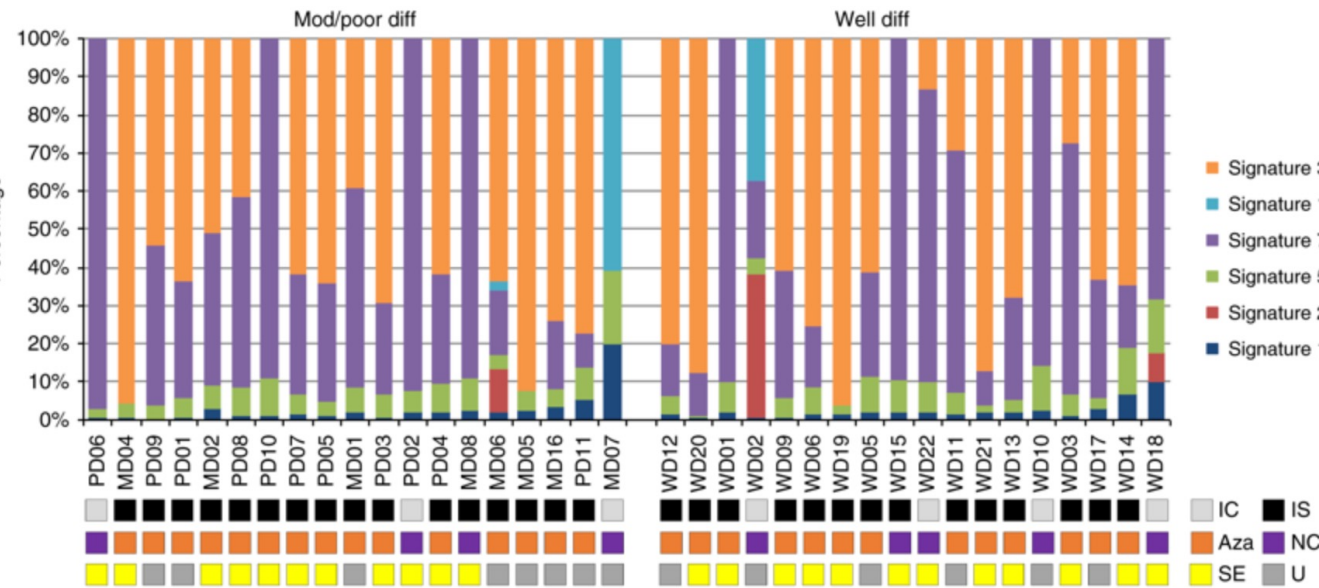
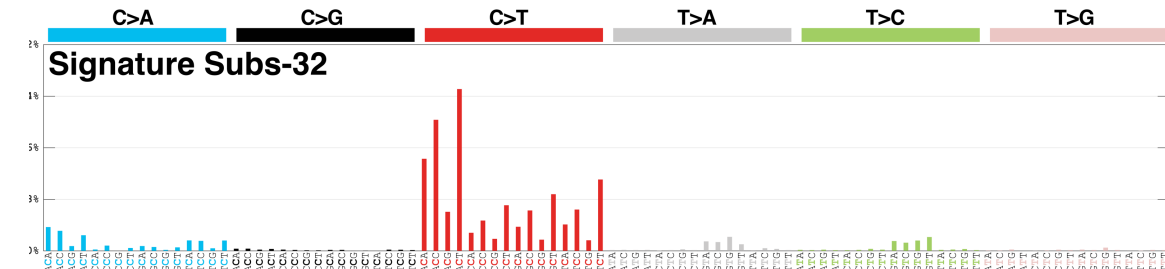
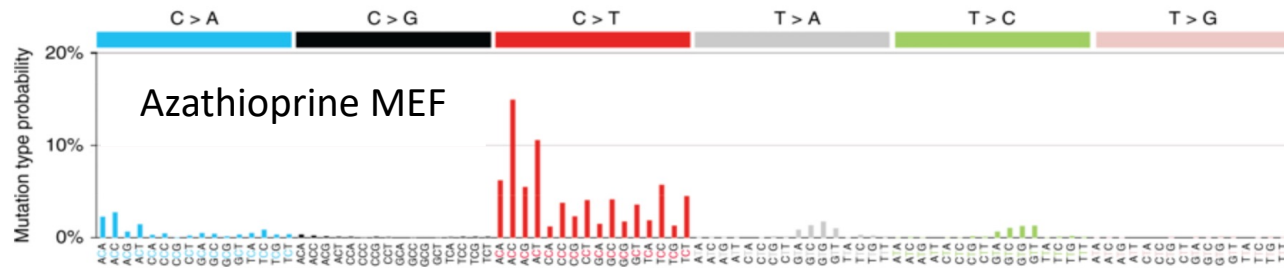
Haloalkane exposure  
 Aristolochic acid exposure  
 Aflatoxin exposure  
 Ultraviolet light exposure  
 Tobacco smoking

# Utilizing mutational signatures for developing cancer prevention strategies

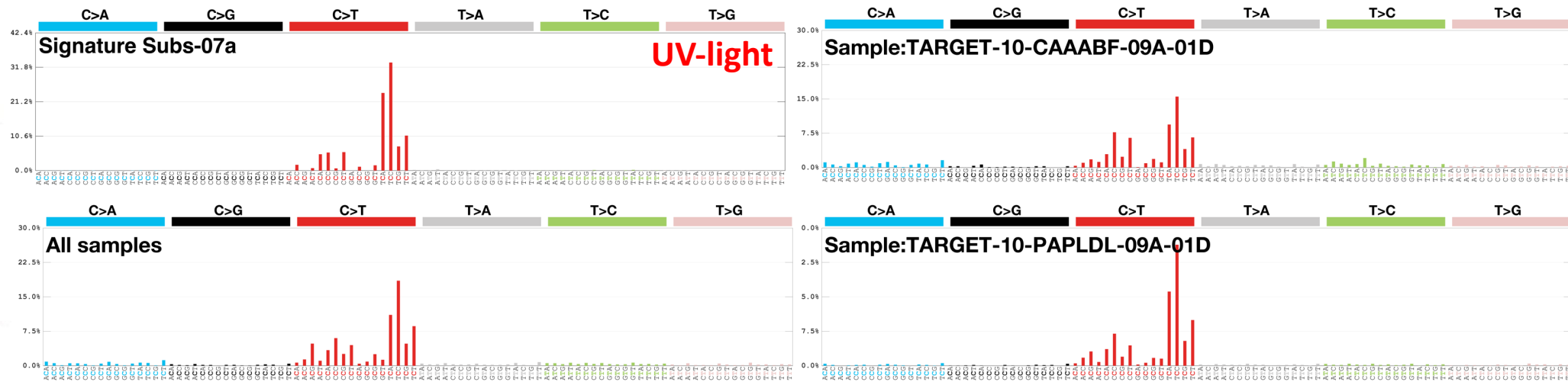
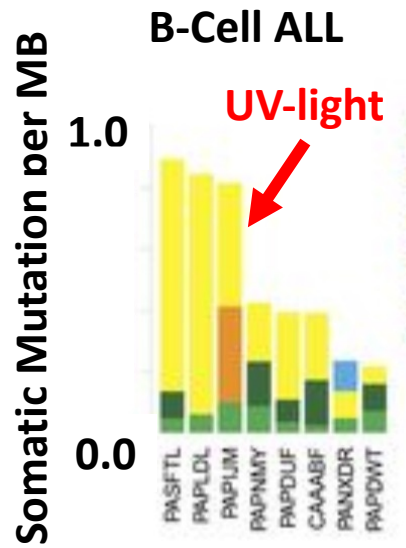


# (Somewhat) unexpected carcinogens: Azathioprine

Azathioprine, sold under the brand name Imuran among others, is an immunosuppressive medication. Azathioprine is on the World Health Organization's List of Essential Medicines, the most effective and safe medicines needed in a health system. **Epidemiological studies by International Agency for Research on Cancer have provided "sufficient" evidence of azathioprine carcinogenicity in humans (Group 1), although the methodology of past studies and the possible underlying mechanisms are questioned.**



# Known carcinogen in unexpected cancer types: **UV-light**



**Similarity extends to strand bias, dinucleotide, and indel patterns. Confirmed in three other cohorts. Signature found only in white Caucasian children. Much lower mutation burden compared to skin cancer.**




[Cancer Causes & Control](#)

October 2017, Volume 28, Issue 10, pp 1075–1083 | [Cite as](#)

Residential exposure to ultraviolet light and risk of precursor B-cell acute lymphoblastic leukemia: assessing the role of individual risk factors, the ESCALE and ESTELLE studies

Authors

[Authors and affiliations](#)

Astrid Coste , Denis Hémon, Laurent Orsi, Mathieu Boniol, Jean-François Doré, Laure Faure, Jacqueline Clavel,

Stéphanie Goujon

# Mutational Signatures for Discovery of Germline Predisposition Syndromes

nature  
genetics

Analysis | Published: 09 November 2015

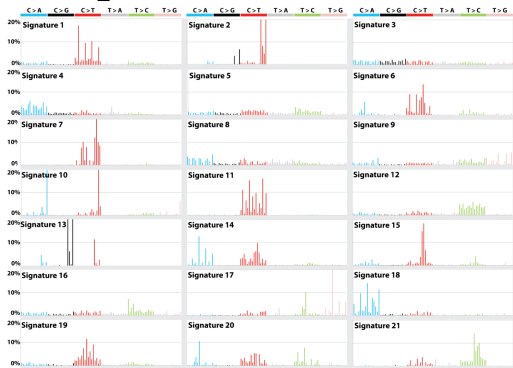
Clock-like mutational processes in human somatic cells

**9 novel mutational signatures identified**

**Signature 30** discovered; etiology/cause unknown

Alexandrov *et al.*, Nature Genetics 2015

2013



First map of 21 mutational signatures identified in human cancer  
Alexandrov *et al.*, Nature 2013

2015

2017

RESEARCH

CANCER

Use of CRISPR-modified human stem cell organoids to study the origin of mutational signatures in cancer

**Signature 30** functionally associated with failure BER due to defective *NTHL1*.

Drost *et al.*, Science 2017

## Cancer Cell

Mutational Signature Analysis Reveals *NTHL1* Deficiency to Cause a Multi-tumor Phenotype

**Signature 30** found in 29 tumors from 7 organs in 17 families. *NTHL1* germline deficiency (SNPs & indels) found in all but one patient.

Grolleman *et al.*, Cancer Cell 2019

2019

**Cancer Cell main conclusions:**  
“Mutational signature analyses can assist to identify germline DNA repair defects.”

“This study illustrates the power of mutational signature analysis in defining tumor phenotypes in rare cancer predisposition syndromes and provides **proof-of-principle for recognizing new patients with cancer syndromes based on tumor sequence data.**”

Utilizing mutational signatures for understanding failed DNA repair and targeted cancer treatment



# Mutational signatures associated with failed DNA repair

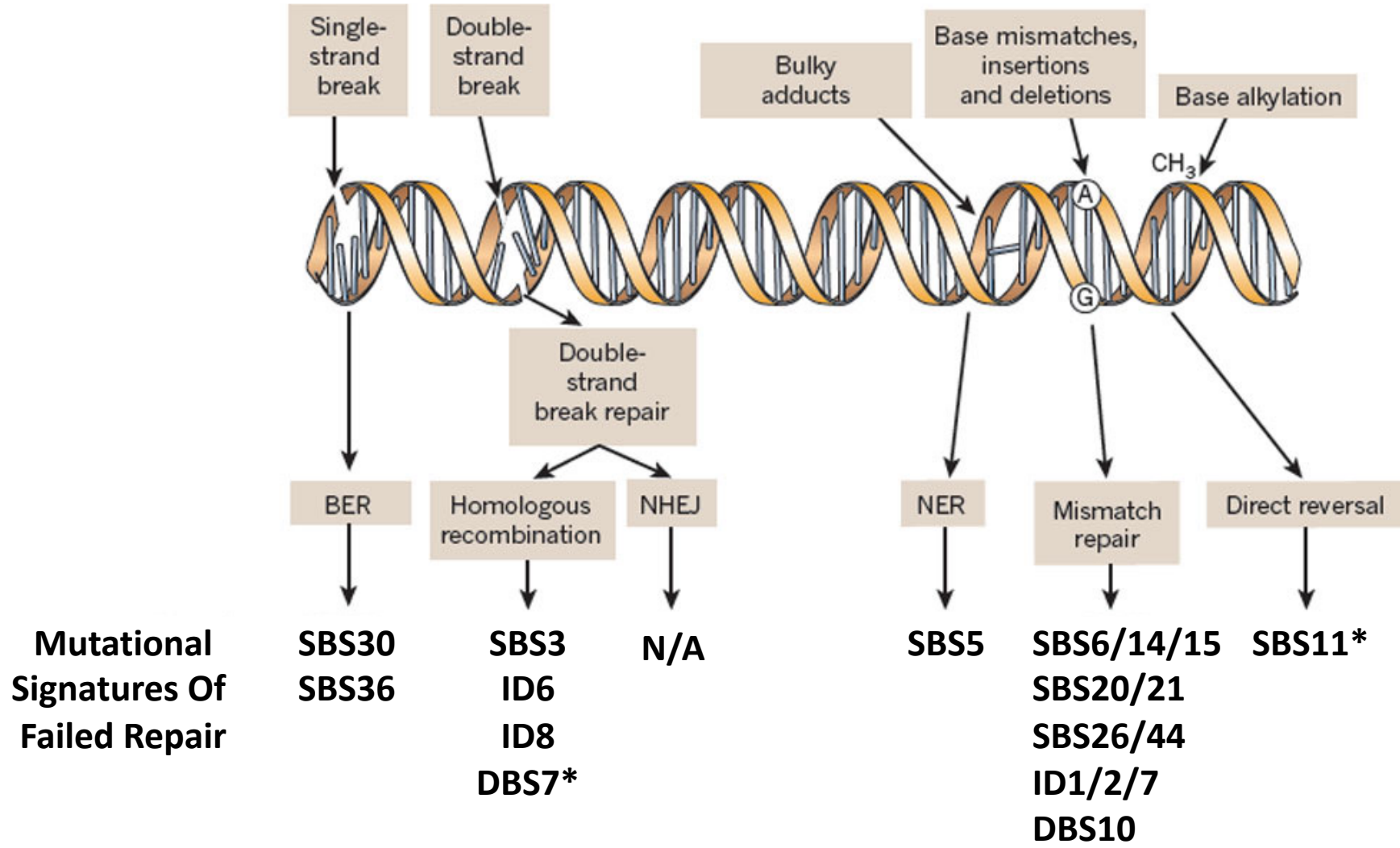


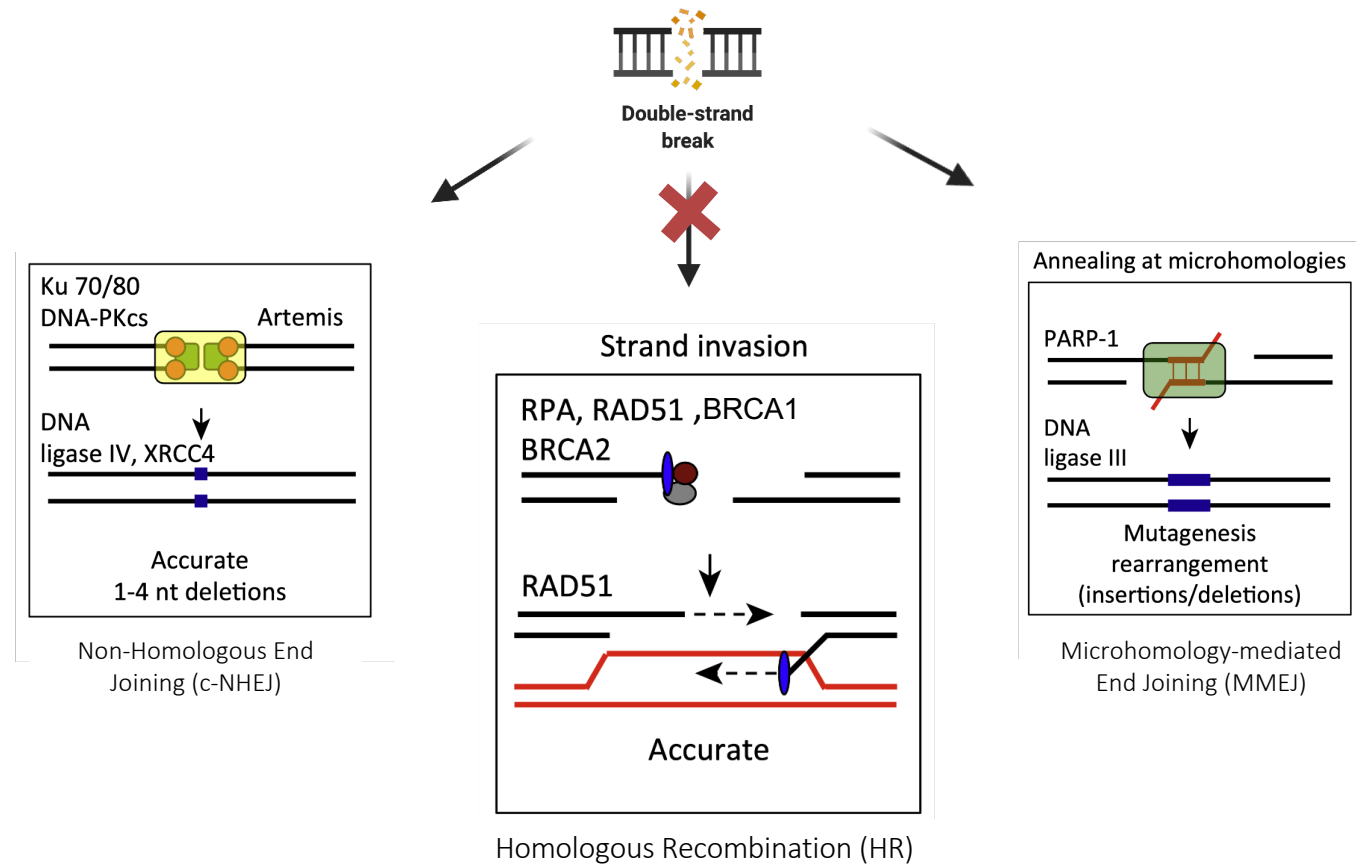
Figure adapted from Lord & Ashworth, Nature 2012  
Alexandrov et al., 2020, Nature

\*Proposed based on limited evidence

# Mutational Signatures with Known Predictive Power

A Mutational signatures useful in analysis						B Underlying mutational process	C Relevant genes	D Predisposition syndrome	E Proposed therapy choice
CS-3	CS-8	MH-indels	RS-3	RS-5	HRD index	Homologous Recombination Repair Deficiency	<i>BRCA1, BRCA2, RAD51C, PALB2</i>	Hereditary Breast and Ovarian Cancer Syndrome	PARP inhibition <sup>32-34</sup> , Platinum-based chemotherapy <sup>35-37</sup>
CS-6	CS-15	CS-20	CS-26	STR-indels		Mismatch Repair Deficiency	<i>MLH1, MSH2, MSH6, PMS1, PMS2</i>	Lynch, CMMRD, BMMR-D, HNPCC	PD1-immunotherapy <sup>48-49,52</sup>
CS-5	CS-8	TSB-sign				Nucleotide Excision Repair Deficiency	<i>ERCC1, ERCC2, XPC</i>	Xeroderma Pigmentosum	Cisplatin <sup>63-65</sup>
CS-18	CS-30	TSB-sign	C>A*	G>T*	C>T*	Base excision Repair Deficiency	<i>MUTYH, OGG1</i> <i>NTHL1, SMUG1</i>	MAP NAP	
CS-10	STR-indels					Deficient DNA polymerase proofreading activity	<i>POLE, POLD1</i>	PPAP	PD1-immunotherapy <sup>48-49,52</sup>
?						Non-Homologous End Joining Deficiency		Nijmegen Breakage Syndrome	
CS-2	CS-13	Kataegis				APOBEC Over-activity	<i>APOBEC1, APOBEC3A, APOBEC3B</i>		Tamoxifen Resistance <sup>70,71</sup>

# Utilizing signatures for detecting homologous recombination deficiency (HRD)



# Utilizing signatures for detecting homologous recombination deficiency (HRD)



FDA Approved drugs for treating advanced-stage ovarian as well as metastatic breast and prostate cancers

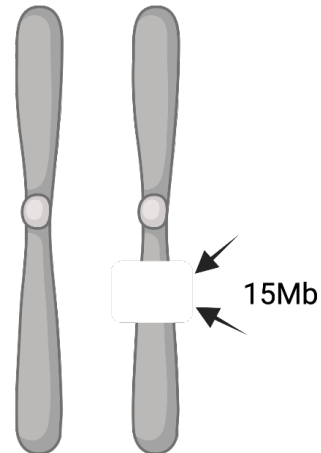
## PARPi leverages synthetic lethality to target HRD cancer cells





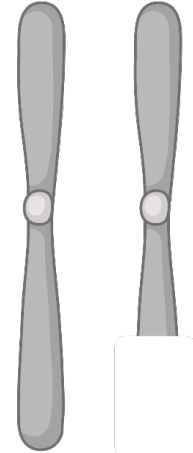
## Examples of HRD Diagnostic Tests

Laboratory Name	Test Name	HRD Status determined by	Genes Assessed	List Price
Foundation Medicine	 FOUNDATIONONE® CDx	BRCA1/BRCA2-positive or LOH $\geq$ 16 %	324 genes, including <i>BRCA1</i> and <i>BRCA2</i>	\$5,800
Myriad	 MYRIAD myChoice® CDx	BRCA1/BRCA2-positive or GIS $\geq$ 42	2 genes: <i>BRCA1</i> , <i>BRCA2</i>	\$4,040



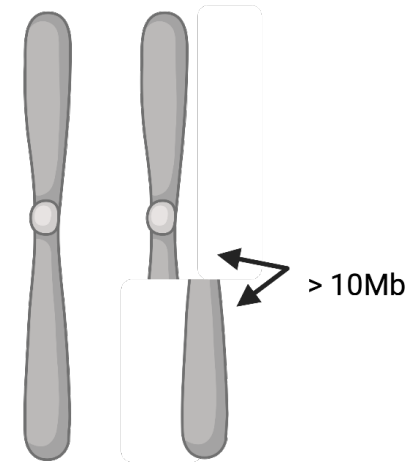
**Loss of Heterozygosity (LOH)**

loss of one normal copy of a gene or a group of genes



**Telomeric Allelic Imbalance**

unequal number of parental and maternal alleles at the telomeres

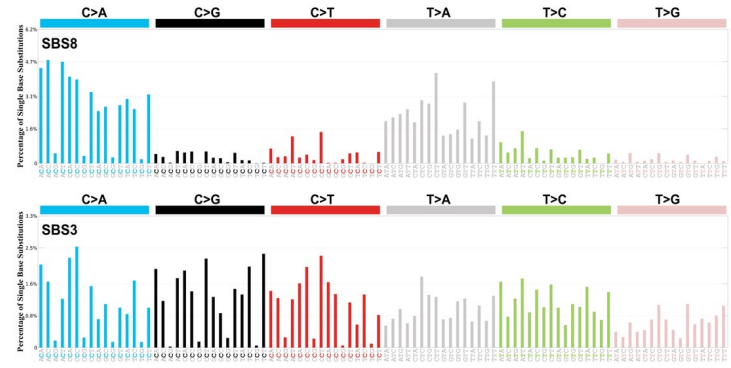


**Large State Transitions (LST)**

chromosomal break between adjacent regions of at least 10 Mb

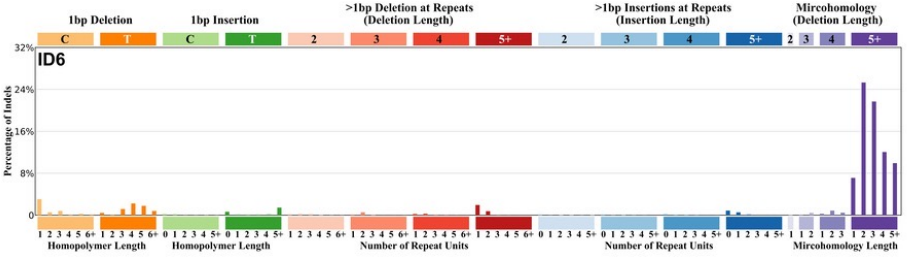
# Mutational signatures/Genomic footprint of HRD

## Single base substitutions



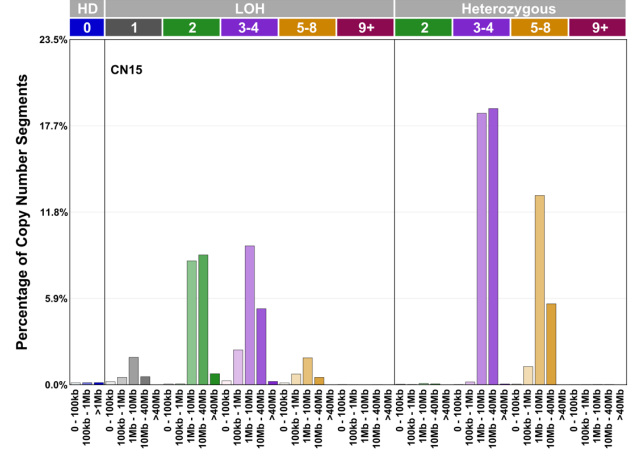
Alexandrov et al. *Nature* 2013

## Microhomology-mediated deletions



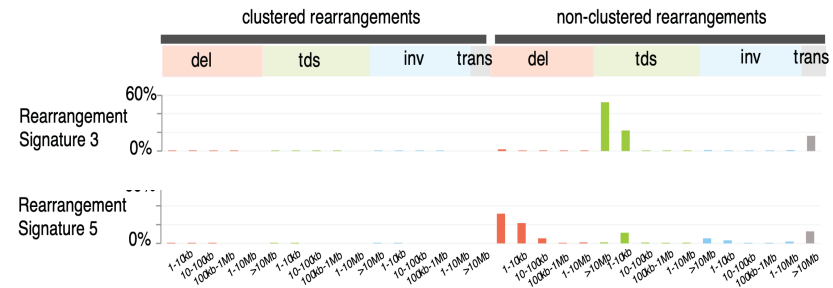
Alexandrov et al. *Nature* 2020

## Copy Number Alterations



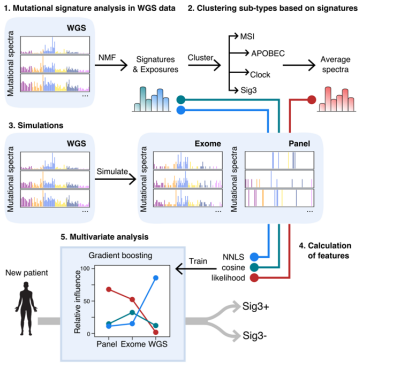
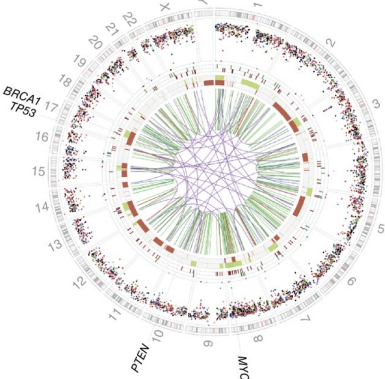
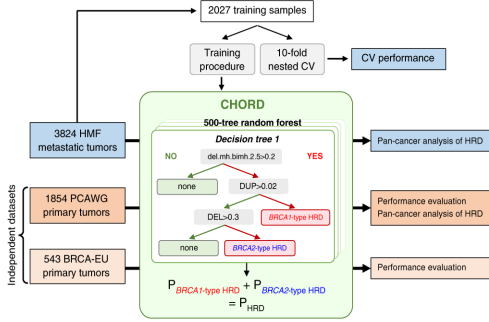
Steele et al. *Nature*, 2022

## Structural Variations

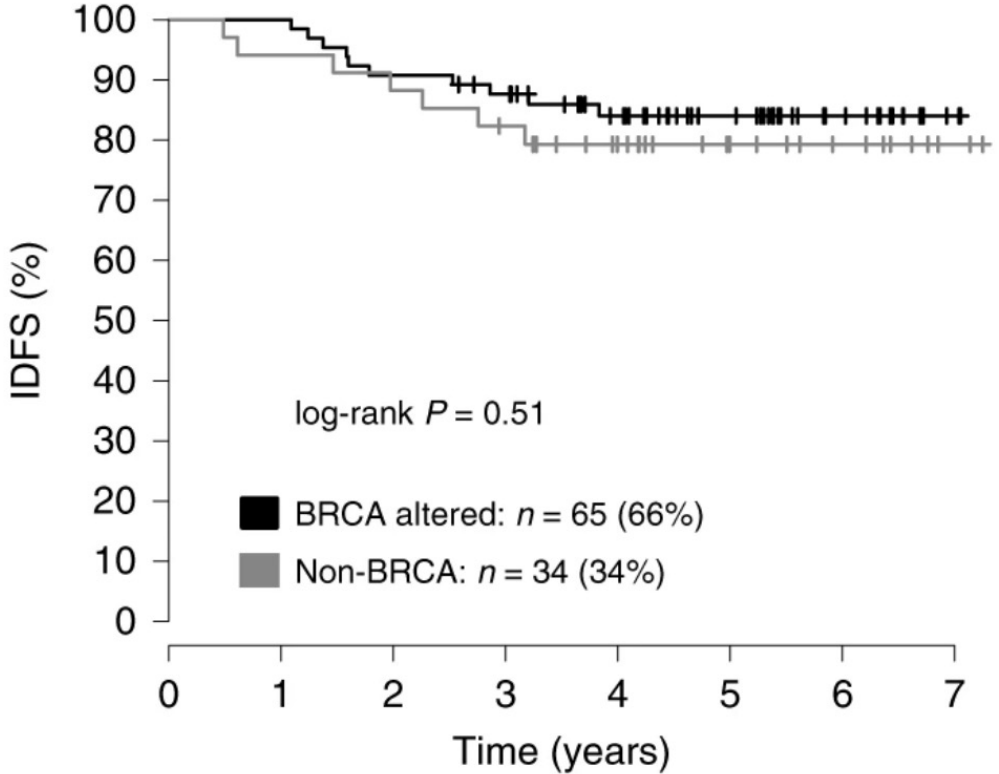
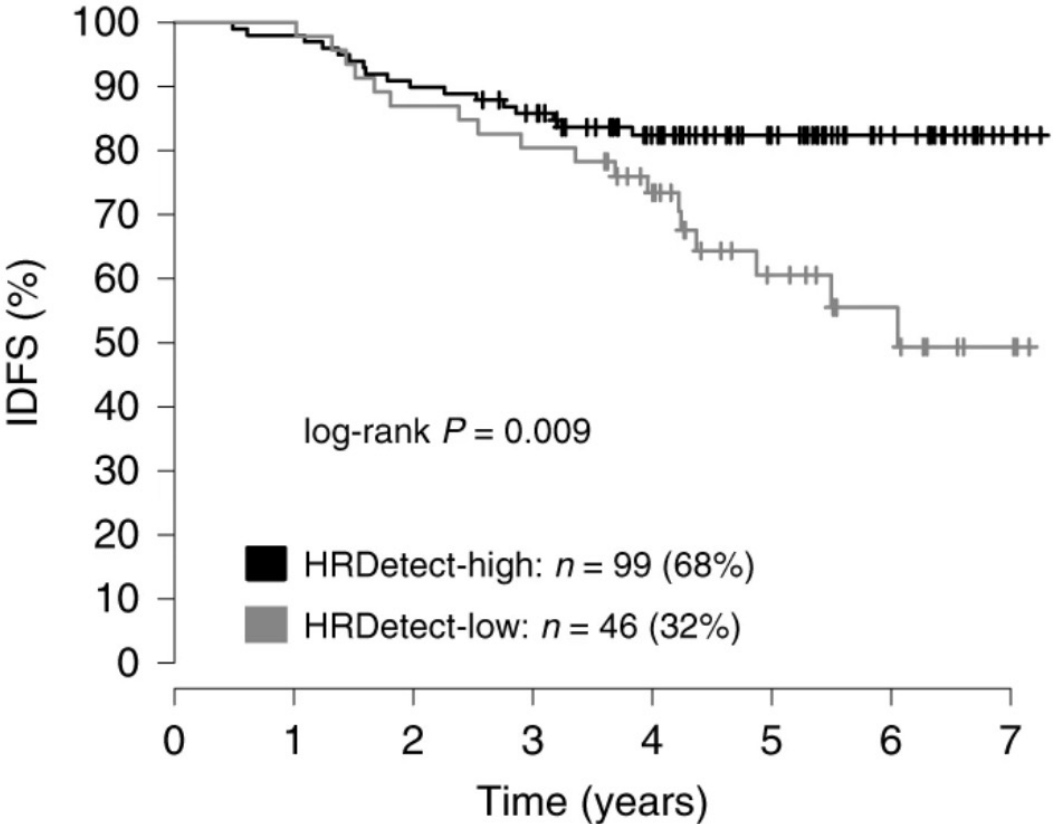


Nik-Zainal et al. *Nature* 2016

# HRD Prediction tools that use mutational signatures or mutational patterns

Tools	 <p style="text-align: center;"><b>SigMA</b></p> <p style="text-align: center;">Gulhan, D. <i>et al. Nat Genet</i> (2019)</p>	 <p style="text-align: center;"><b>HRDetect</b></p> <p style="text-align: center;">Davies, H. <i>et al. Nat Med</i> (2017)</p>	 <p style="text-align: center;"><b>CHORD</b></p> <p style="text-align: center;">Nguyen, L. <i>et al. Nat Commun</i> (2020)</p>
Features	SBS3	SBS3, SBS8, Microhomology-mediated deletions, RS3, RS5	SBS, ID, and SV mutational patterns
Sequencing Platform	WGS, WES, Panels	WGS	WGS
Advantages & Limitations	<p>Method can be applied to WGS, WES, and panel data.</p> <p>SBS3 is flat and method can be used only for highly mutated panels (~25% breast cancers).</p>	<p>Whole-genome sequencing is expensive approach especially at high-coverage. In many cases it requires fresh cancer tissues, and it is not commonly used in clinical practice.</p> <p>HRDetect &amp; CHORD can detect ~50% more samples that will respond to PARPi when compared myChoice CDx.</p>	

# Example of applying an academic HRD tool to a breast cancer cohort



IDFS: invasive disease-free survival





# Anecdotes of The Present



# A brief look at the present with glimpses of the future

- Utilizing clustered mutations for understanding cancer development and evolution.
- The repertoire of copy-number signatures in human cancer.
- A novel machine learning approach for detecting homologous recombination deficiency.

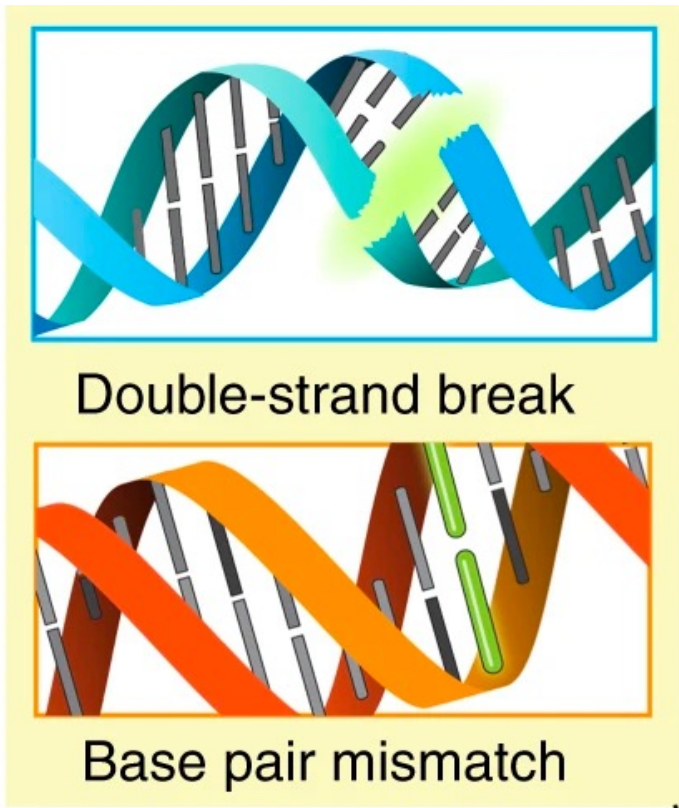
# Utilizing clustered mutations for understanding cancer development and evolution



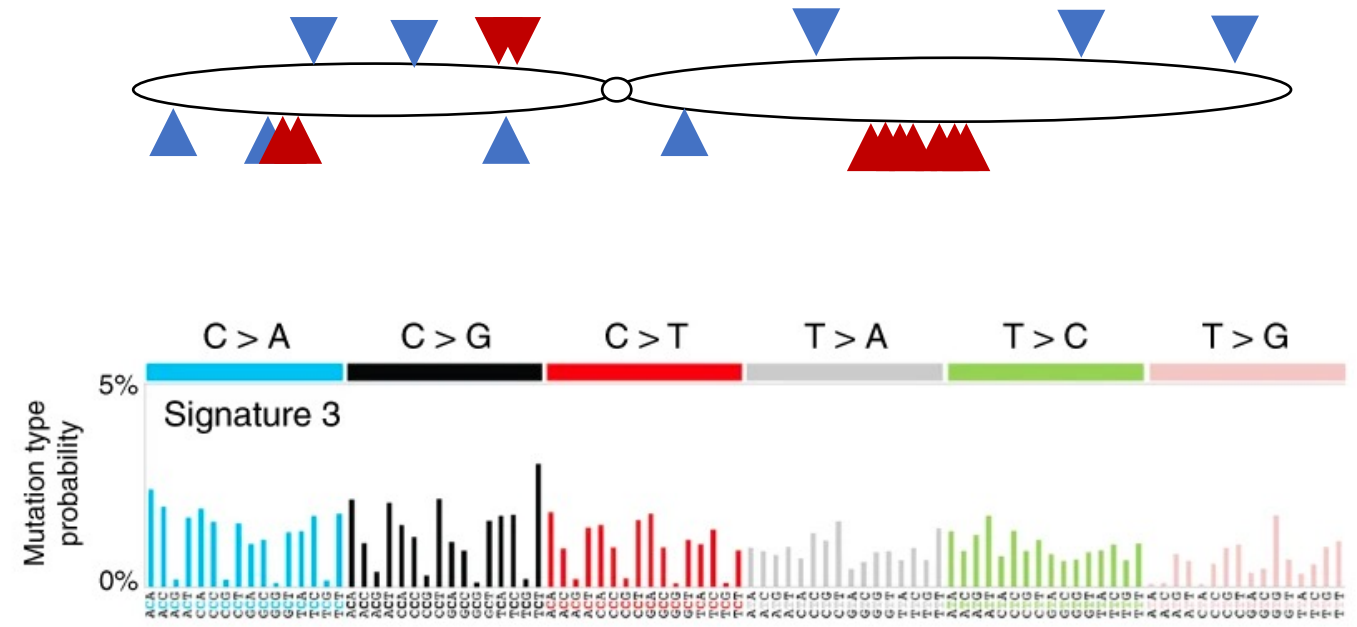
Erik Bergstrom





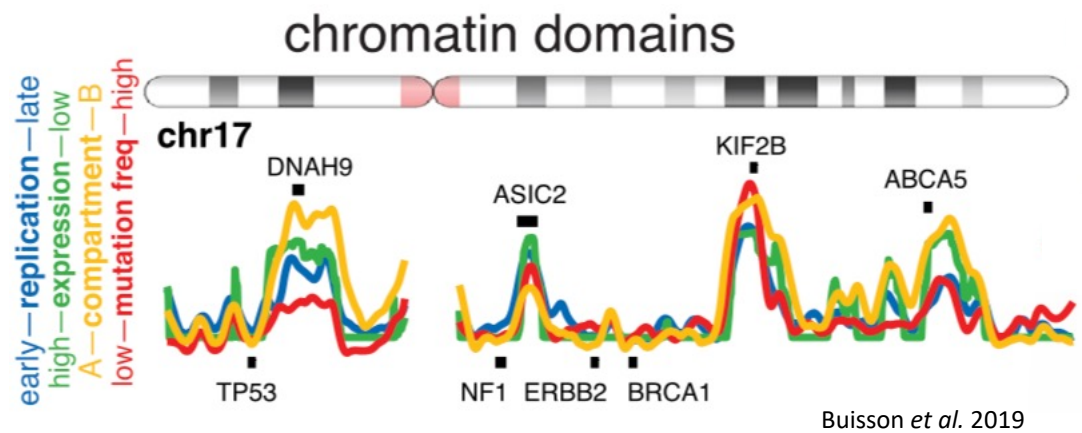


*Mutations occur as single, independent events randomly across the genome*

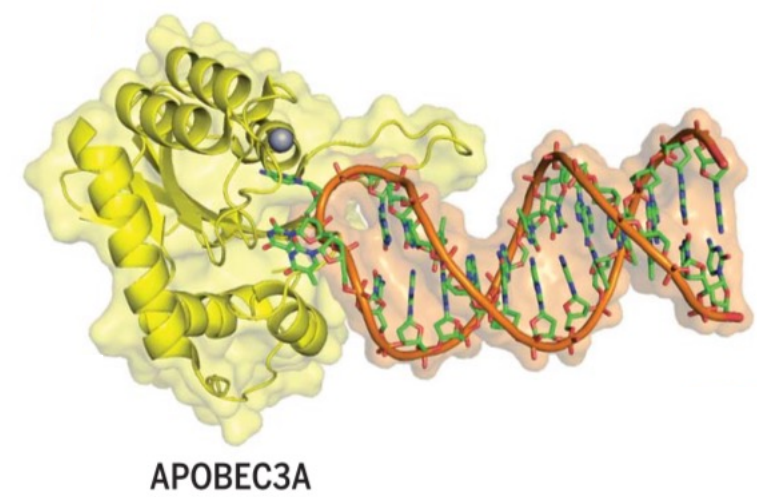
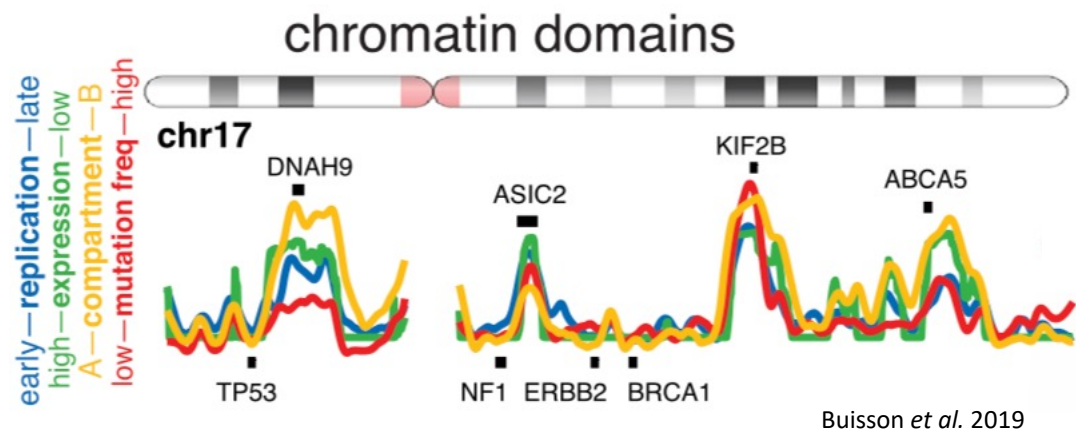


DNA repair pathway alterations

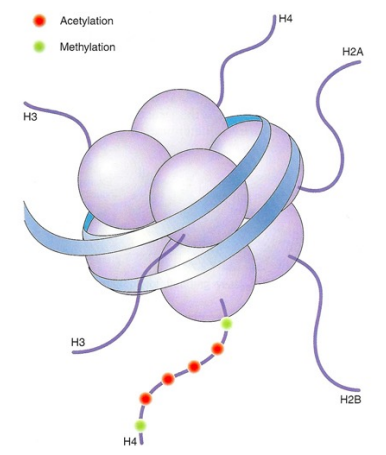
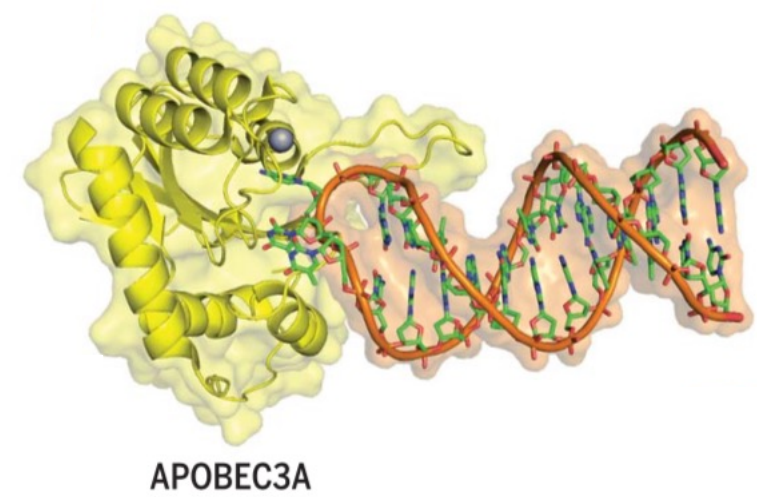
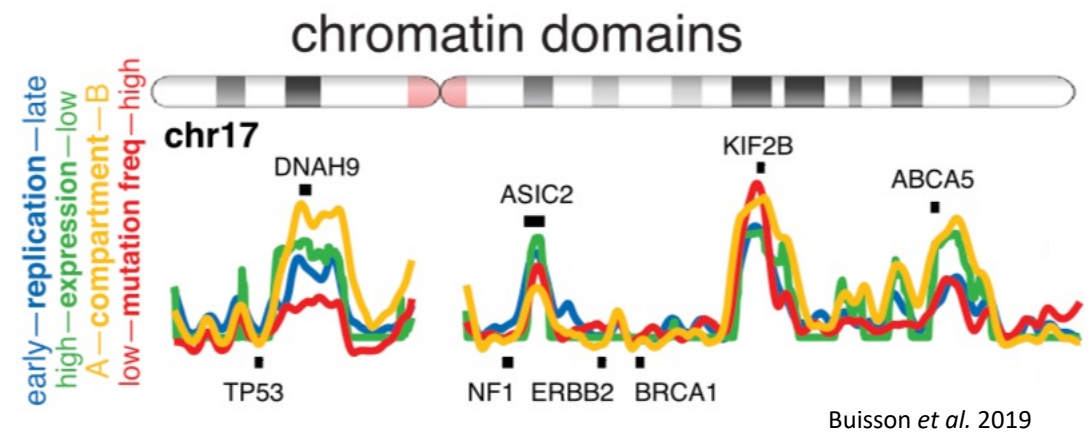
# Mutation rate is dependent on a range of genomic features



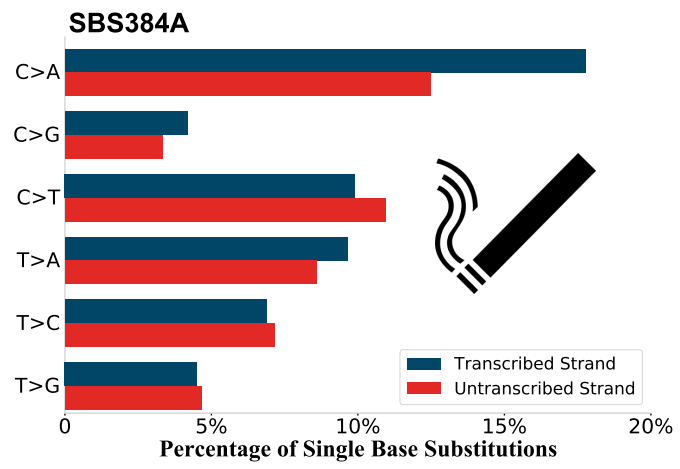
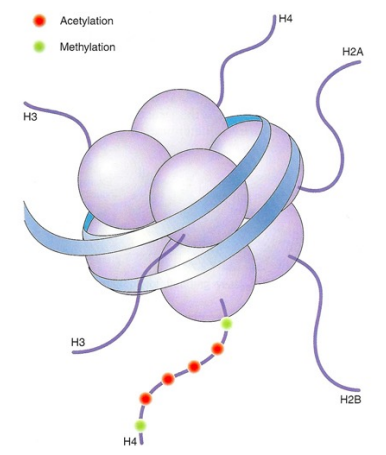
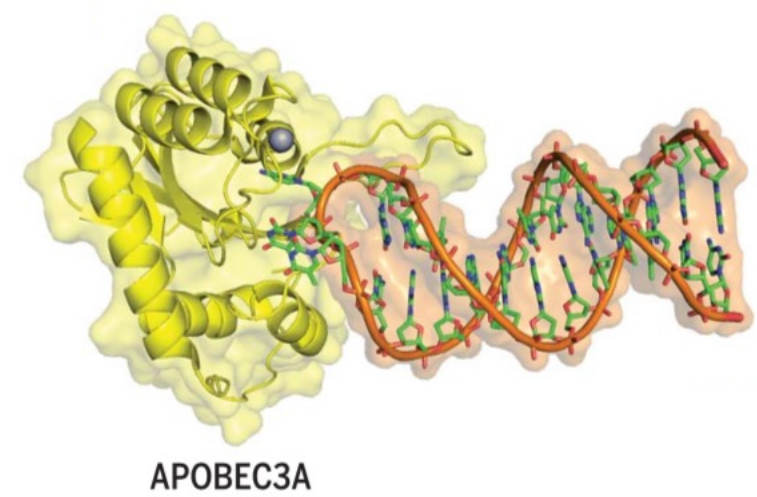
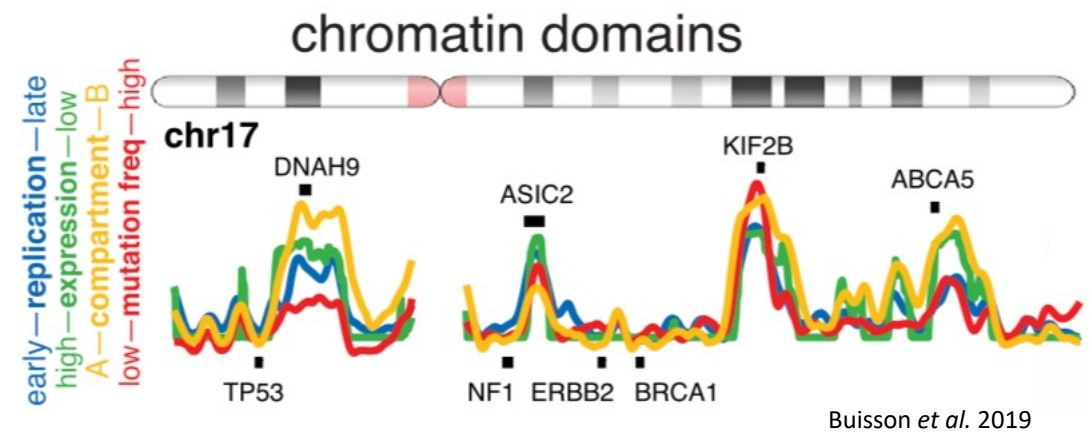
# Mutation rate is dependent on a range of genomic features



# Mutation rate is dependent on a range of genomic features

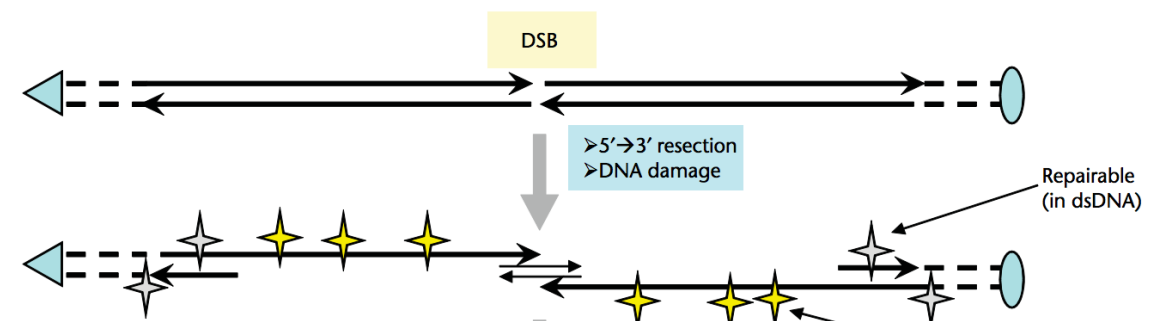
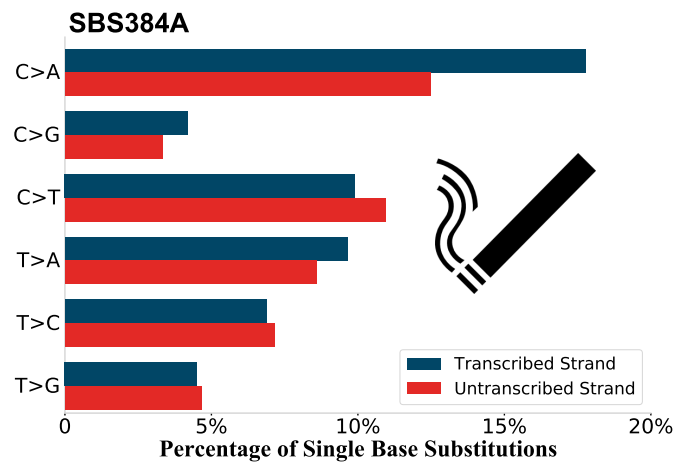
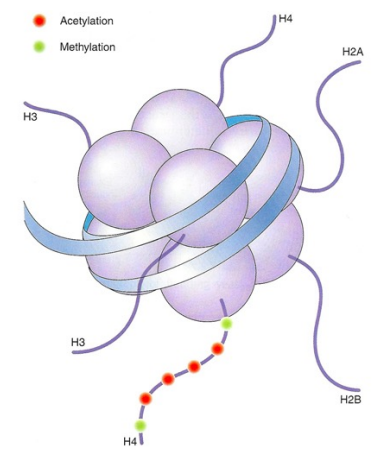
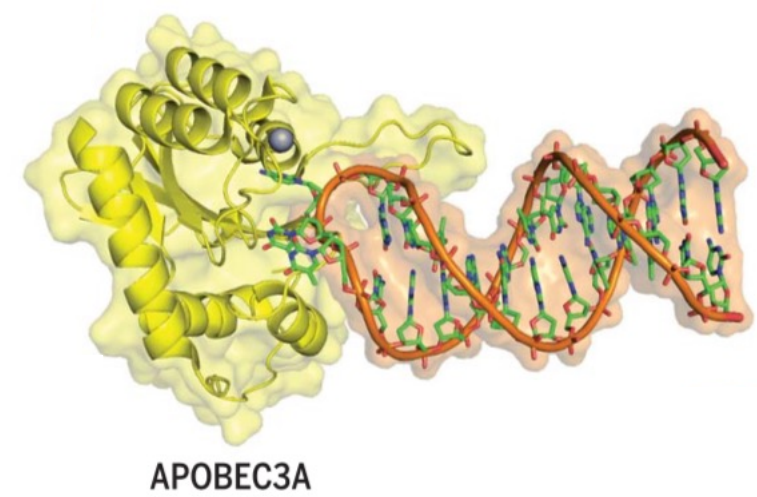
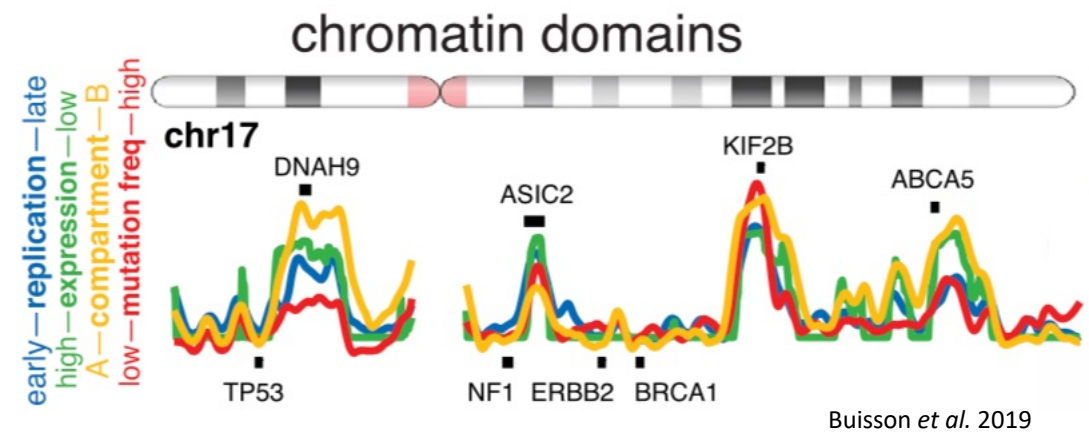


# Mutation rate is dependent on a range of genomic features

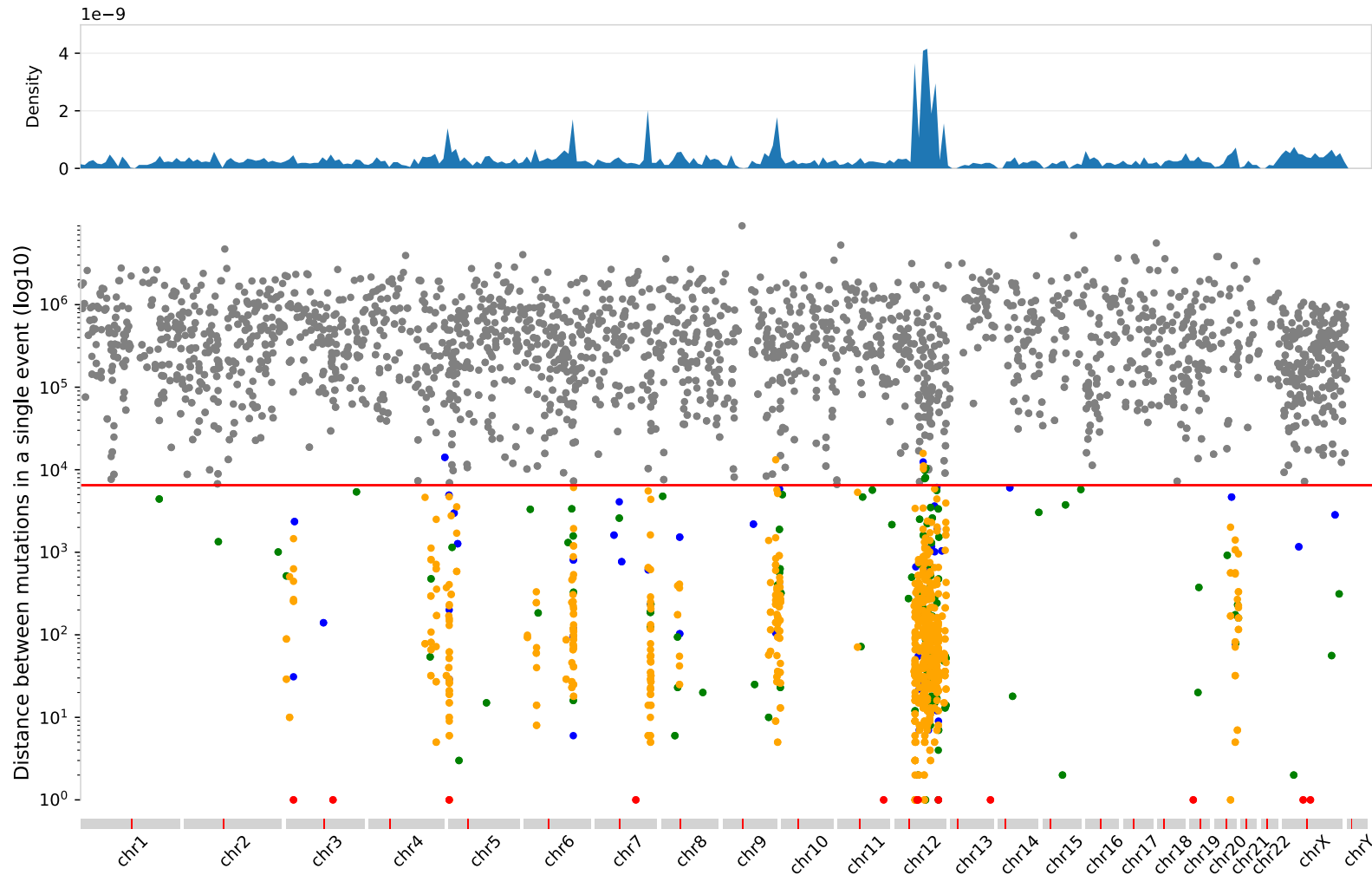




# Mutation rate is dependent on a range of genomic features



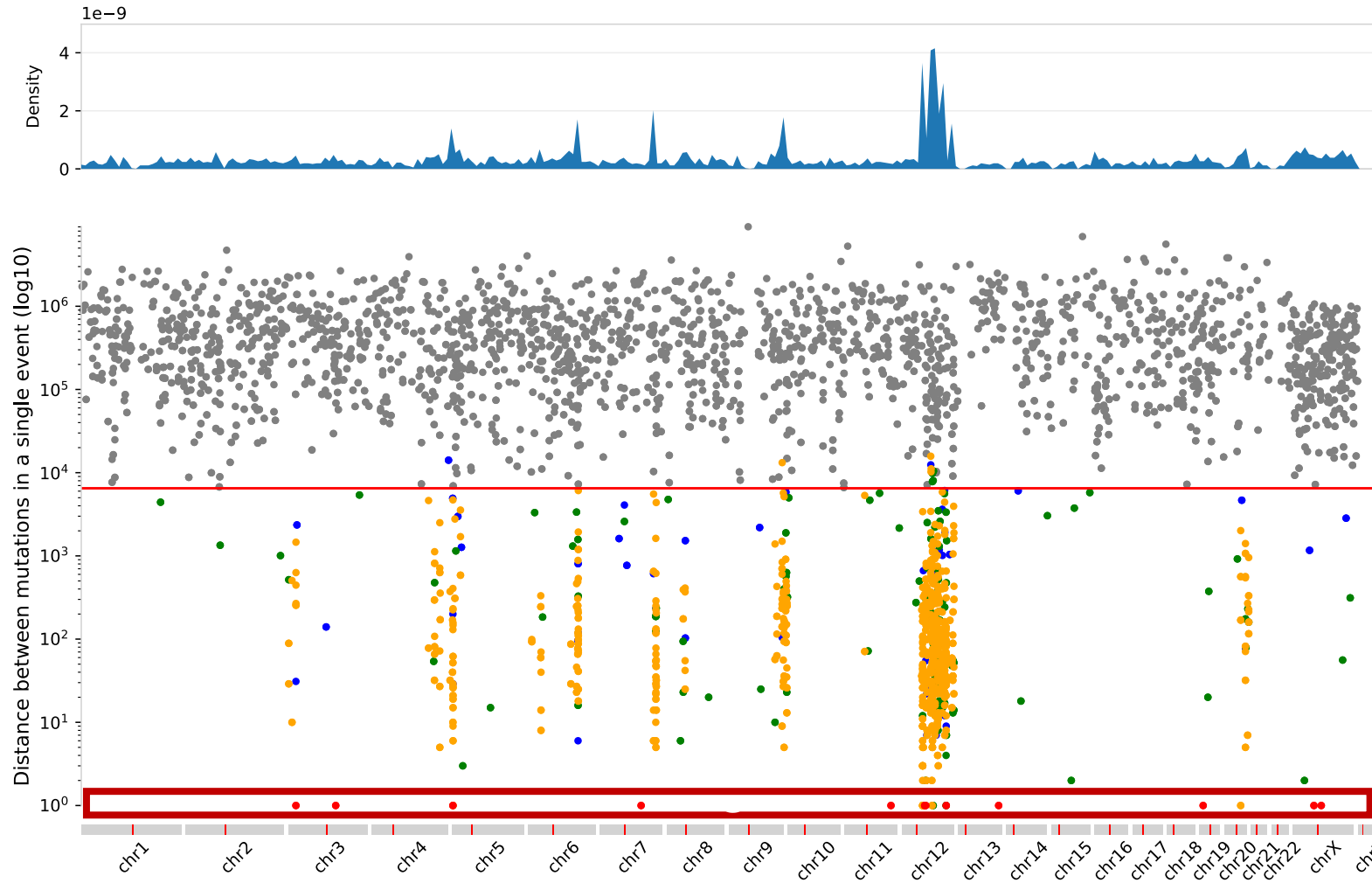
## Clustered mutations - SARC-US\_SP121828



Bergstrom *et al.*, Nature 2022

# Classification of clustered mutations

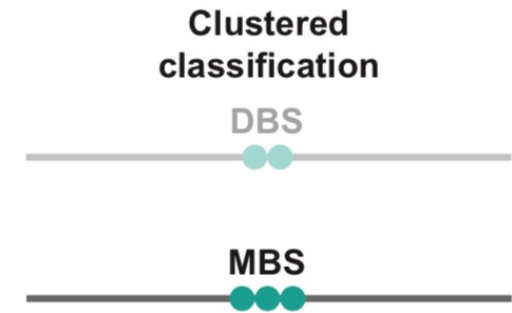
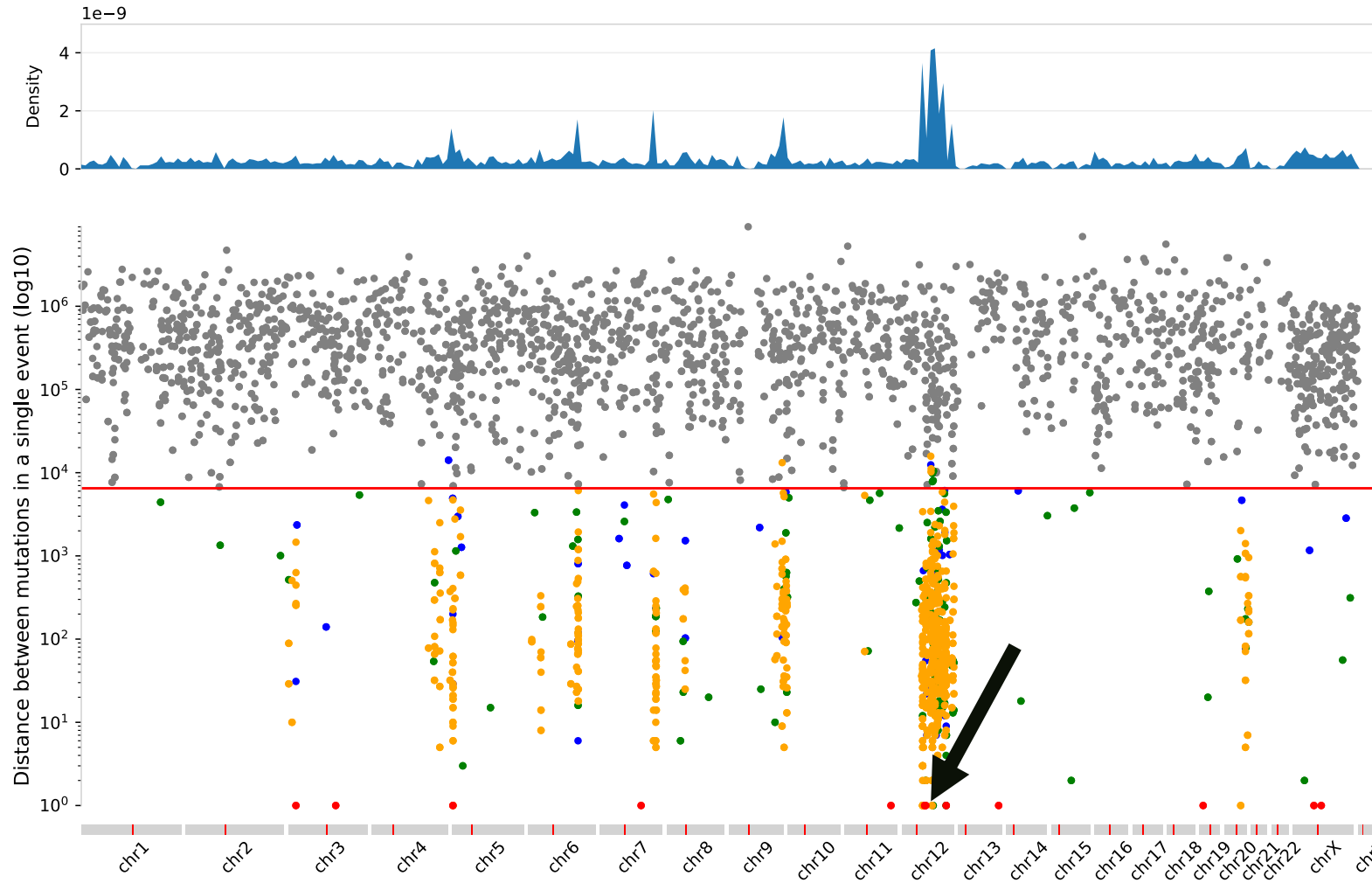
## Clustered mutations - SARC-US\_SP121828



Bergstrom *et al.*, Nature 2022

# Classification of clustered mutations

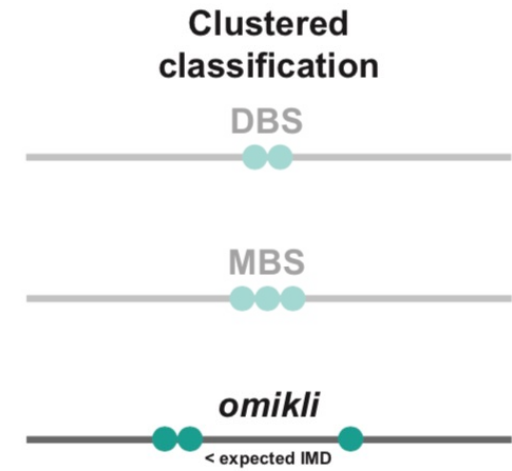
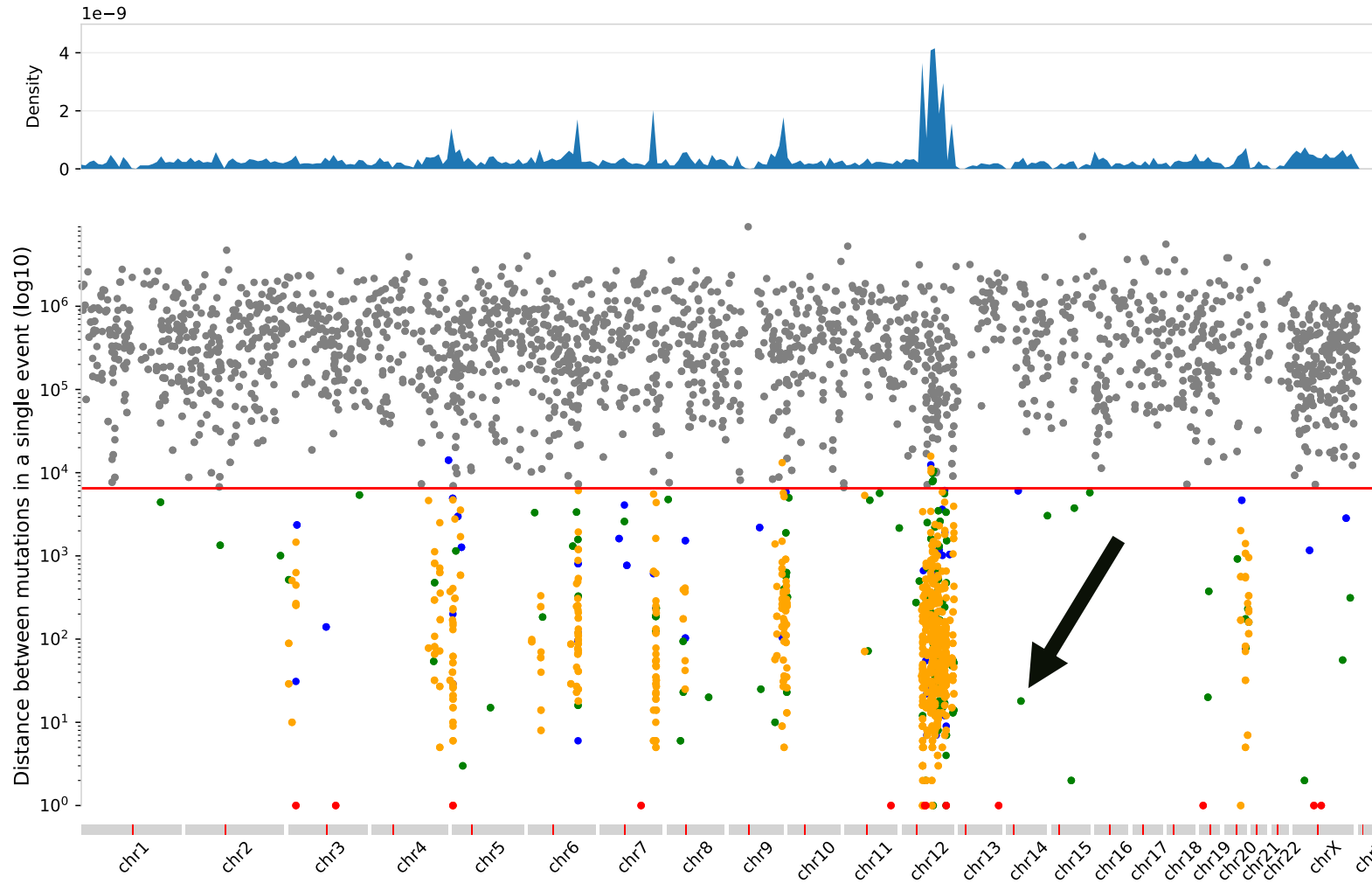
Clustered mutations - SARC-US\_SP121828



Bergstrom *et al.*, Nature 2022

# Classification of clustered mutations

## Clustered mutations - SARC-US\_SP121828

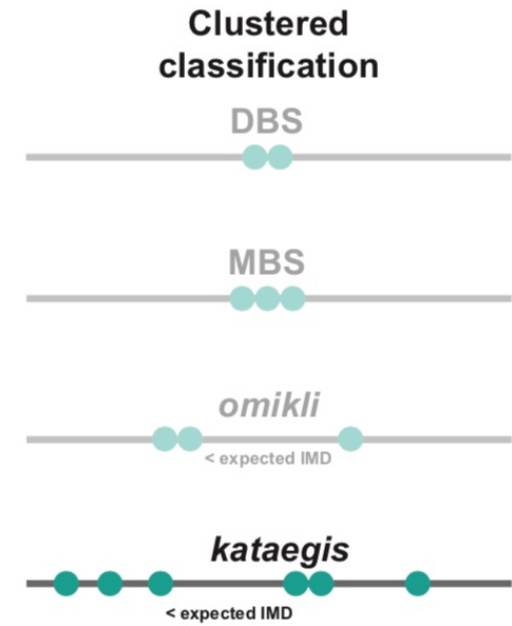
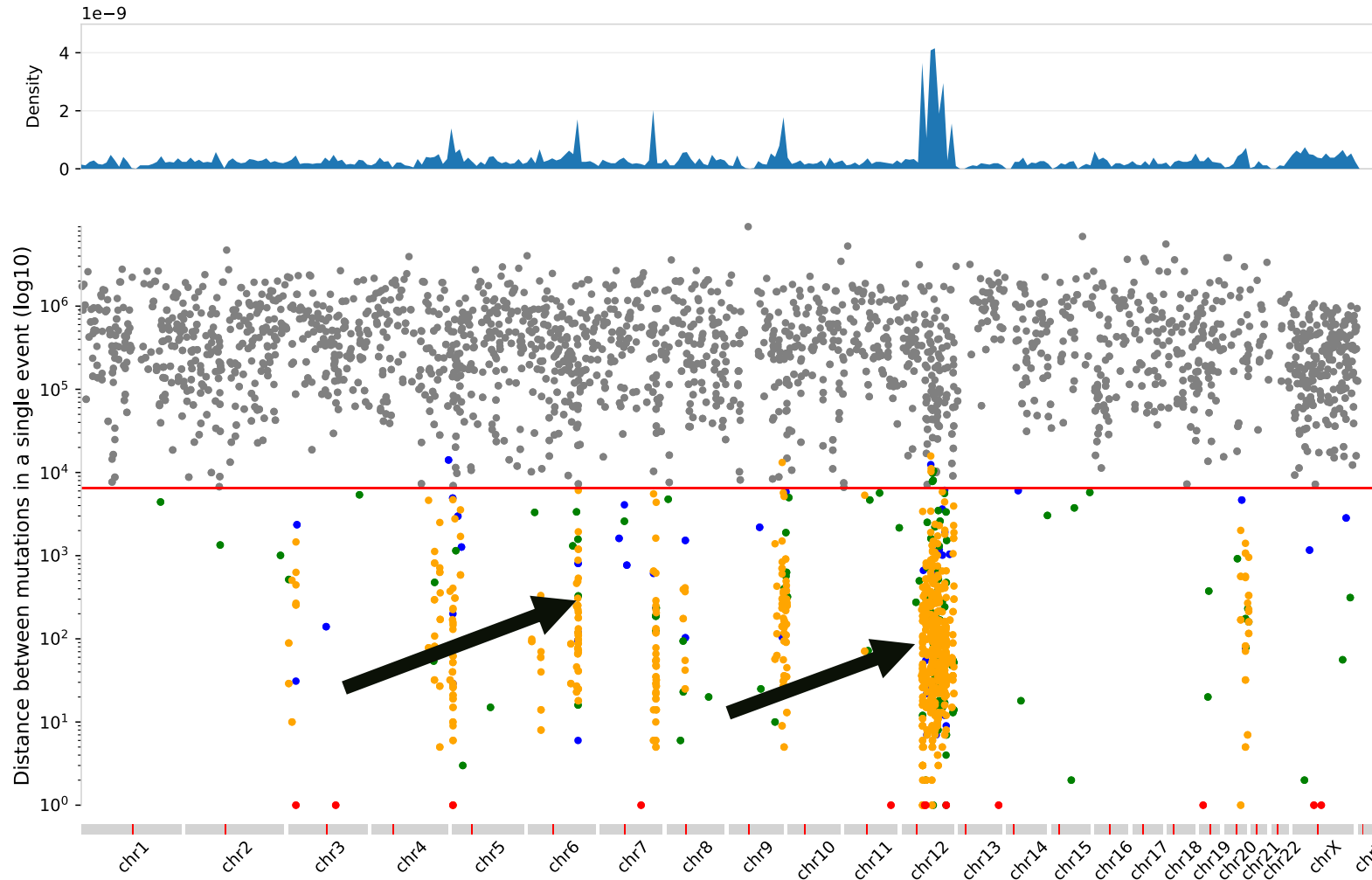


Bergstrom *et al.*, Nature 2022



# Classification of clustered mutations

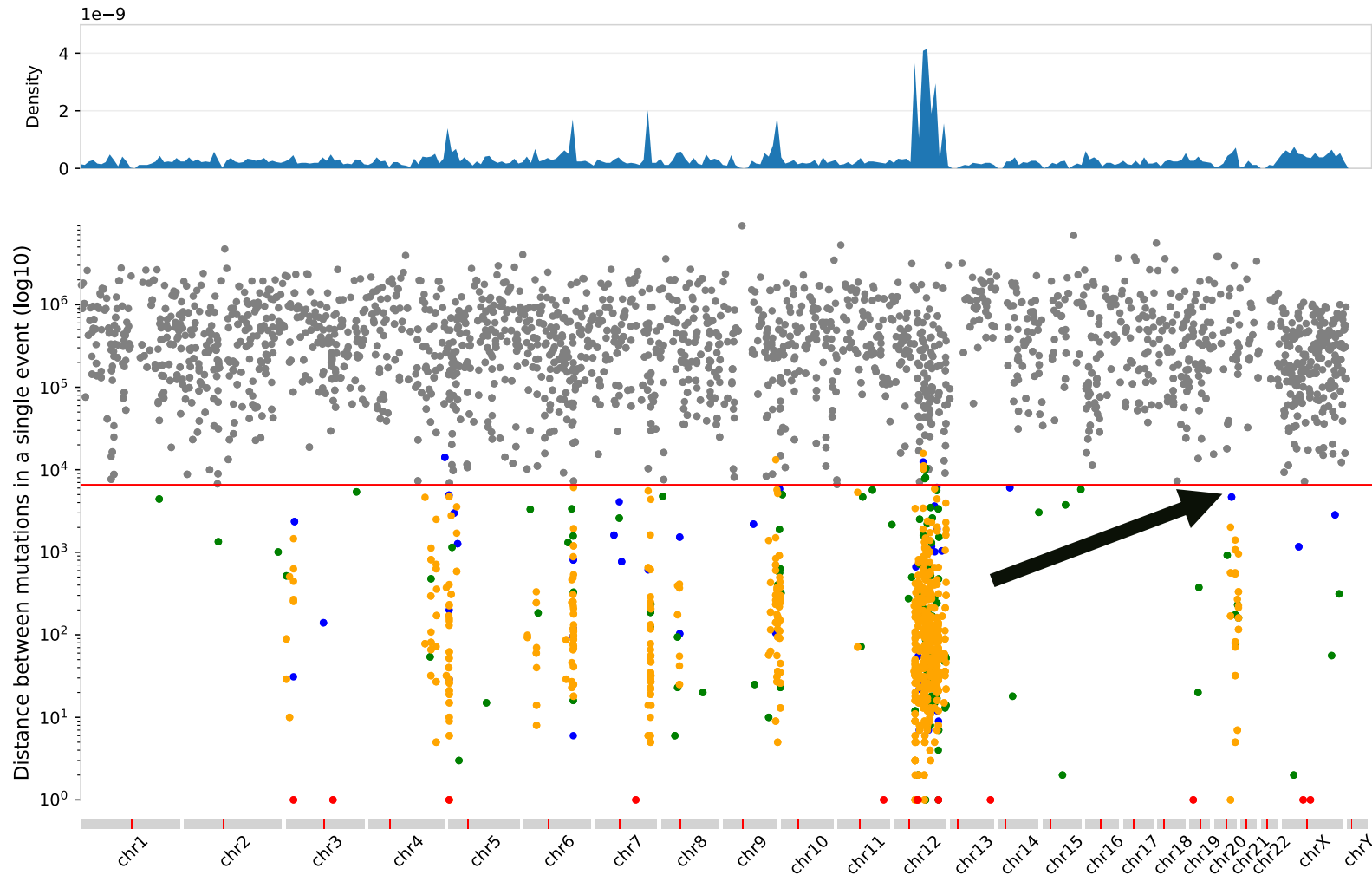
## Clustered mutations - SARC-US\_SP121828



Bergstrom *et al.*, Nature 2022

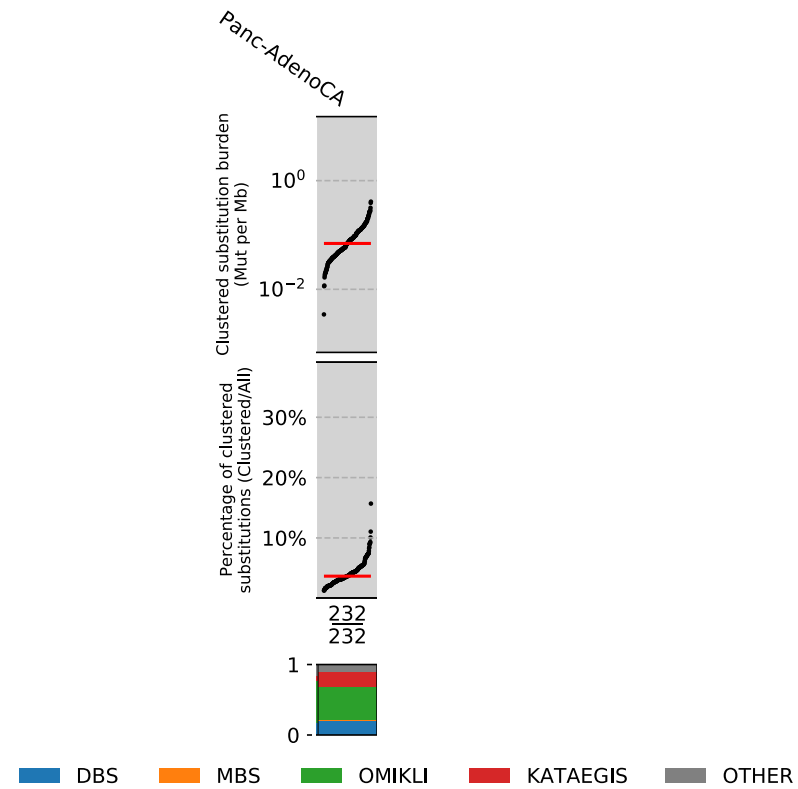
# Classification of clustered mutations

## Clustered mutations - SARC-US\_SP121828



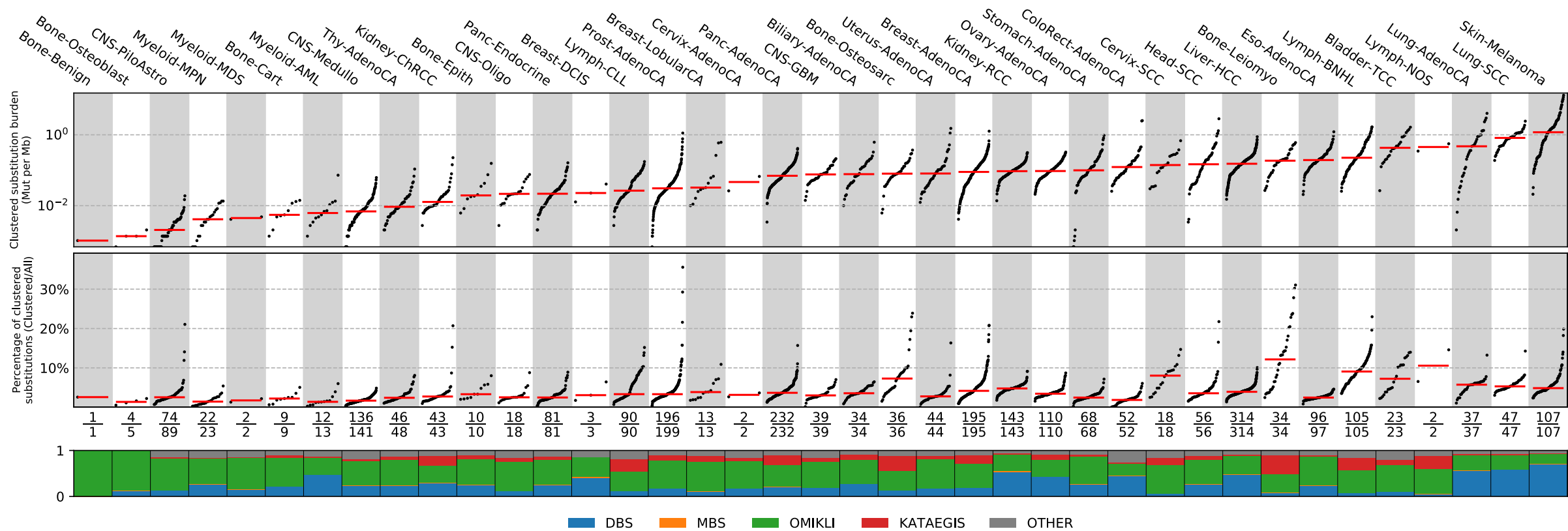
Bergstrom *et al.*, Nature 2022

# The landscape of clustered mutations across human cancer



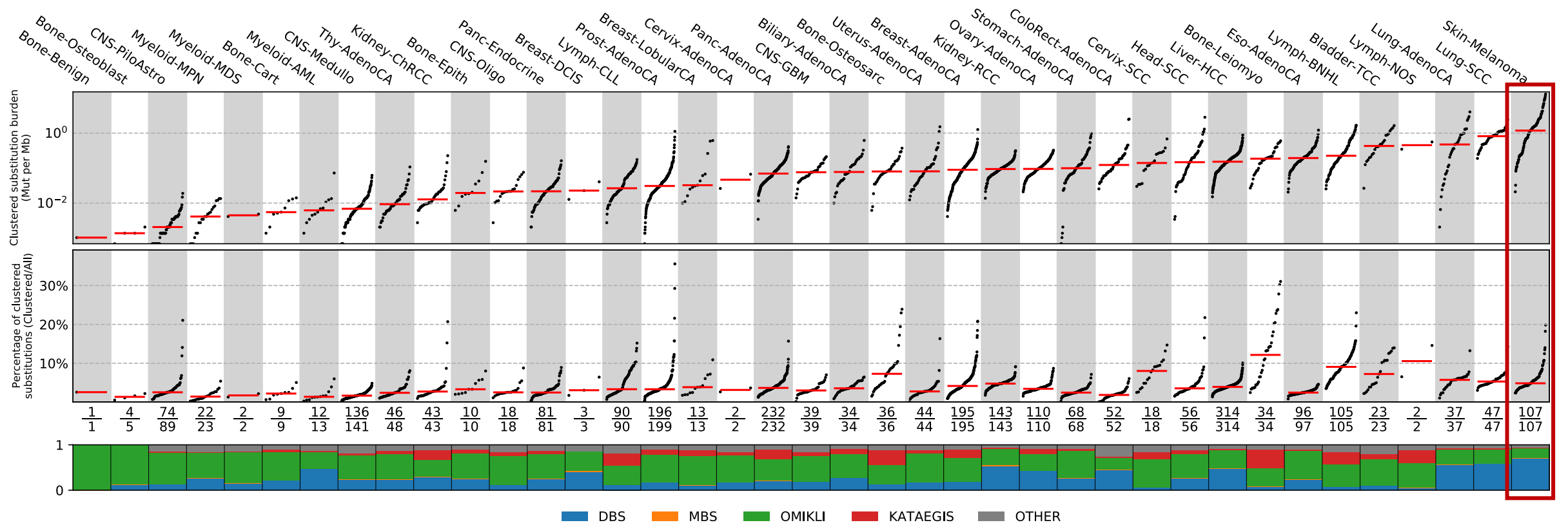
Bergstrom *et al.*, Nature 2022

# The landscape of clustered mutations across human cancer



Bergstrom *et al.*, Nature 2022

# The landscape of clustered mutations across human cancer

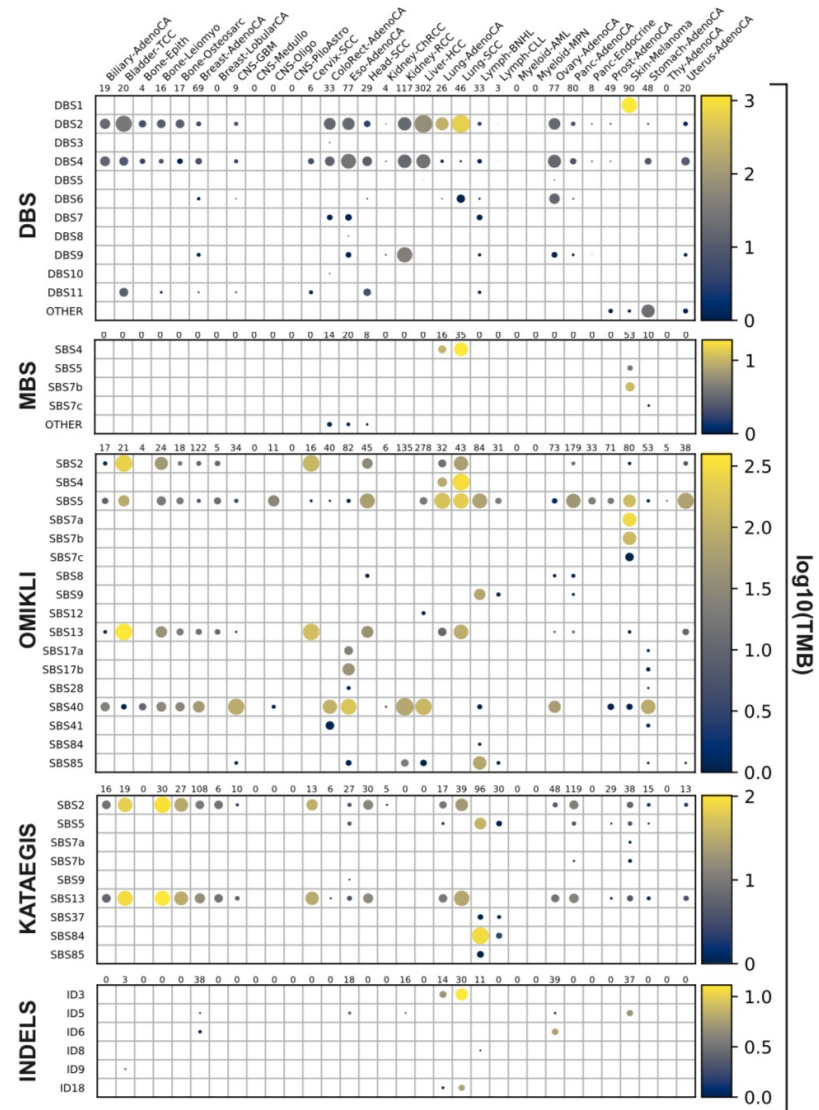


Bergstrom *et al.*, Nature 2022



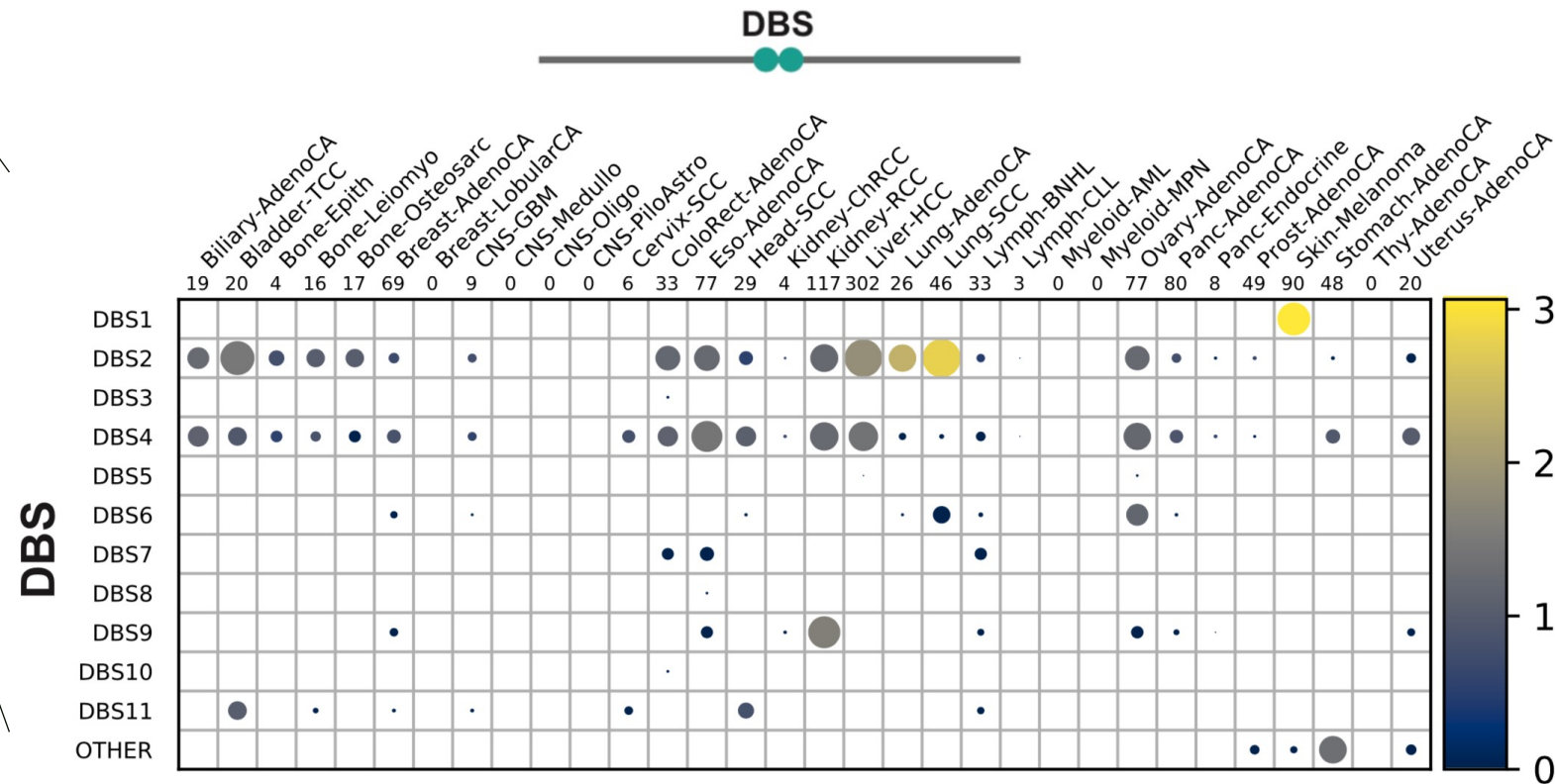
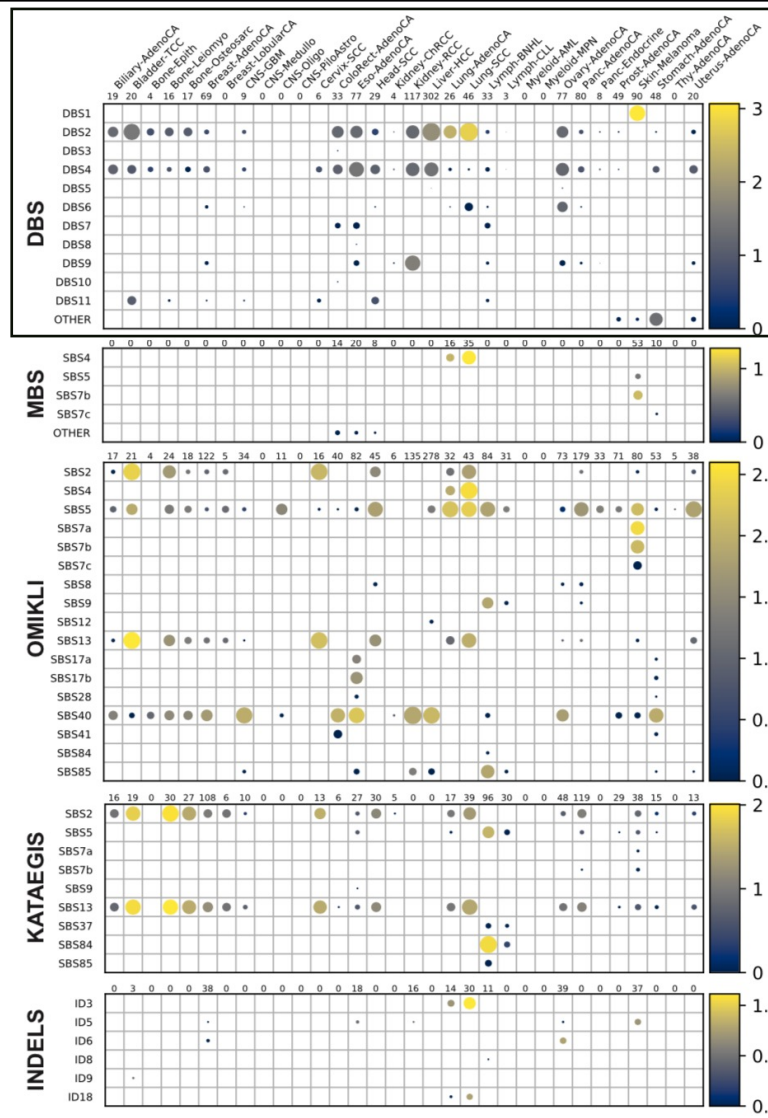


# Mutational processes underlying clustered events



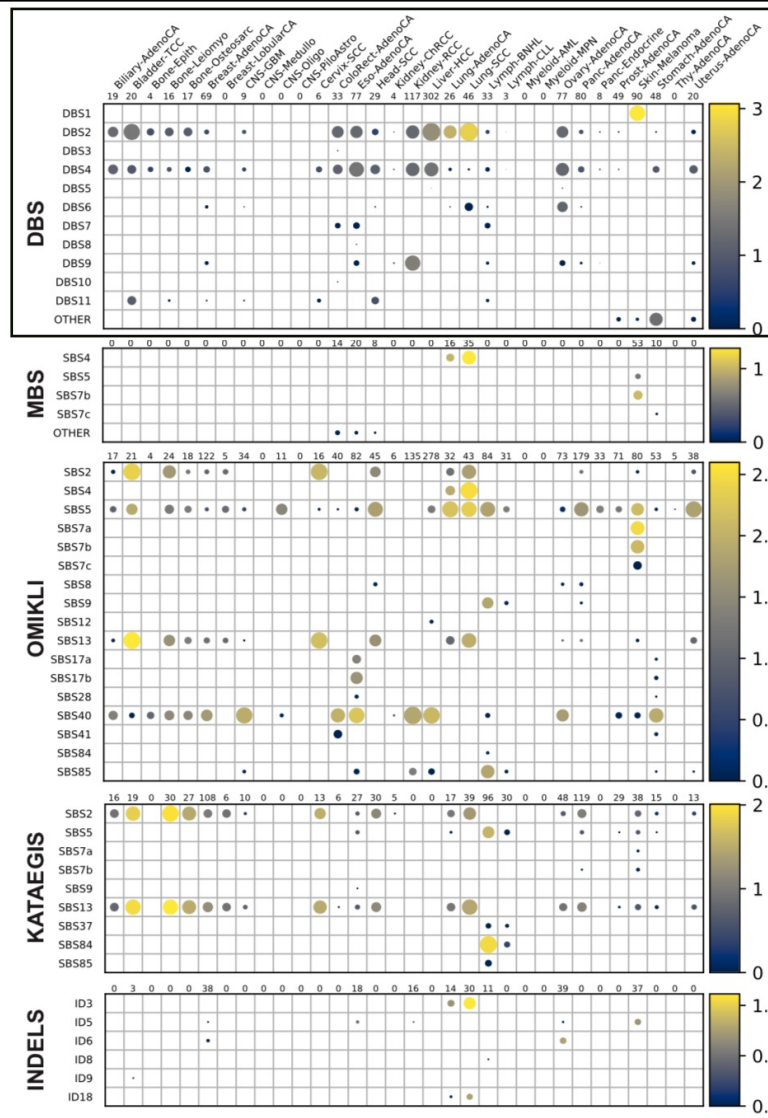
Bergstrom *et al.*, Nature 2022

# Mutational processes underlying clustered events



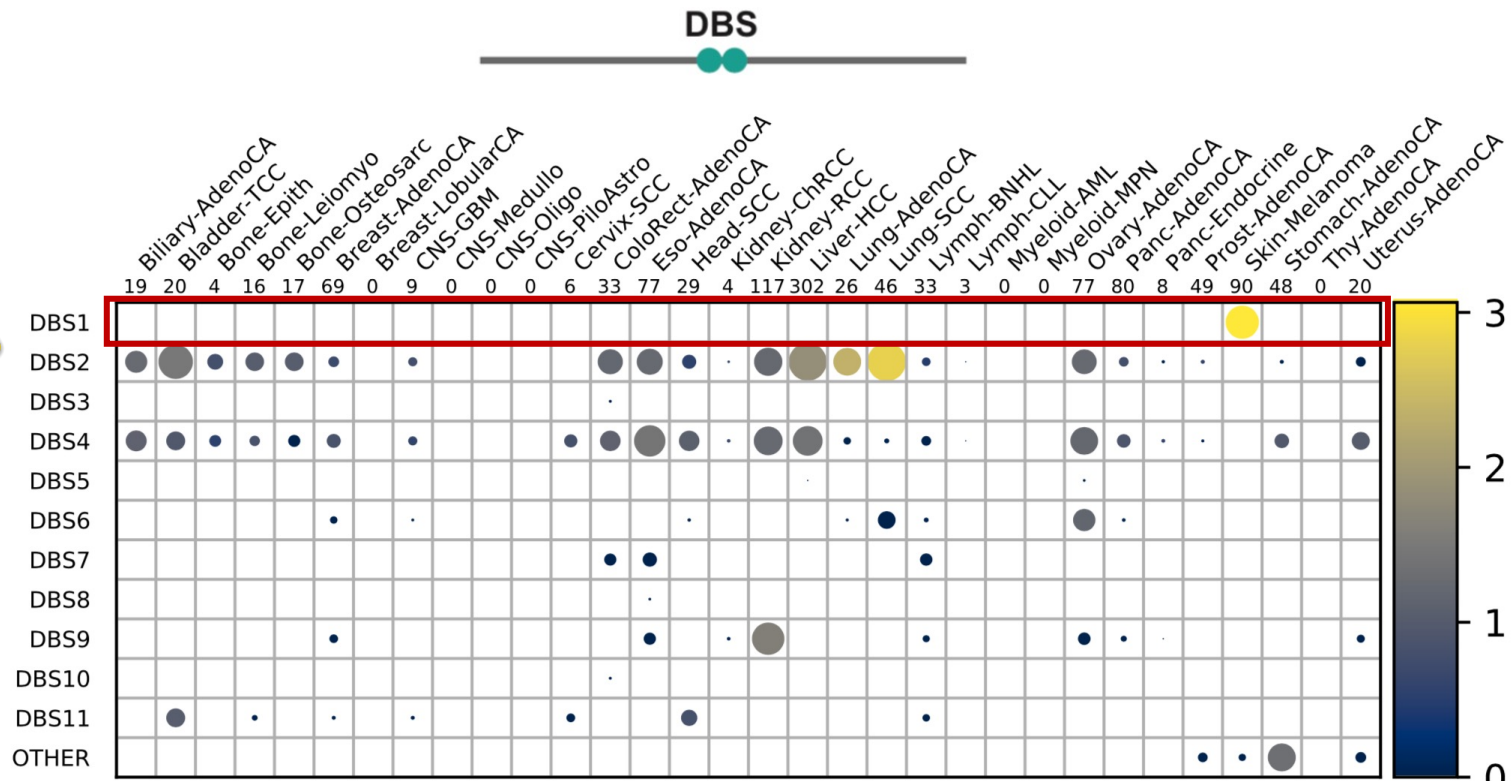
Bergstrom *et al.*, Nature 2022

# Mutational processes underlying clustered events



**DBS**

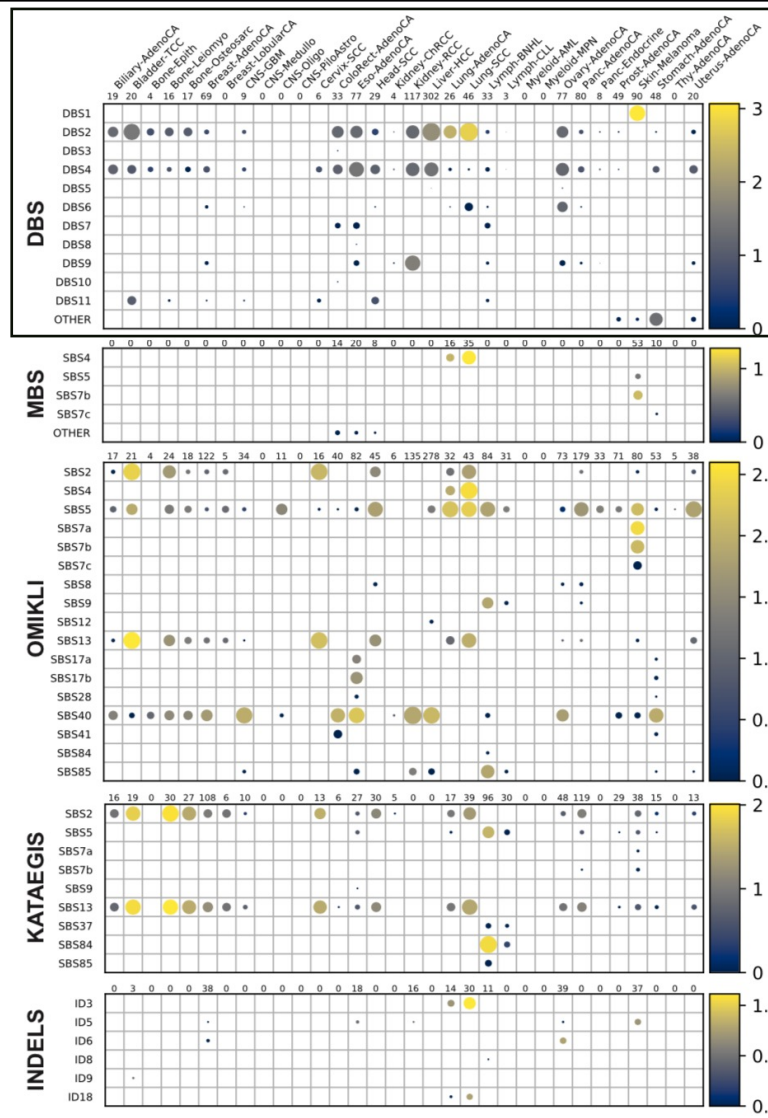
$\log_{10}(\text{TMB})$



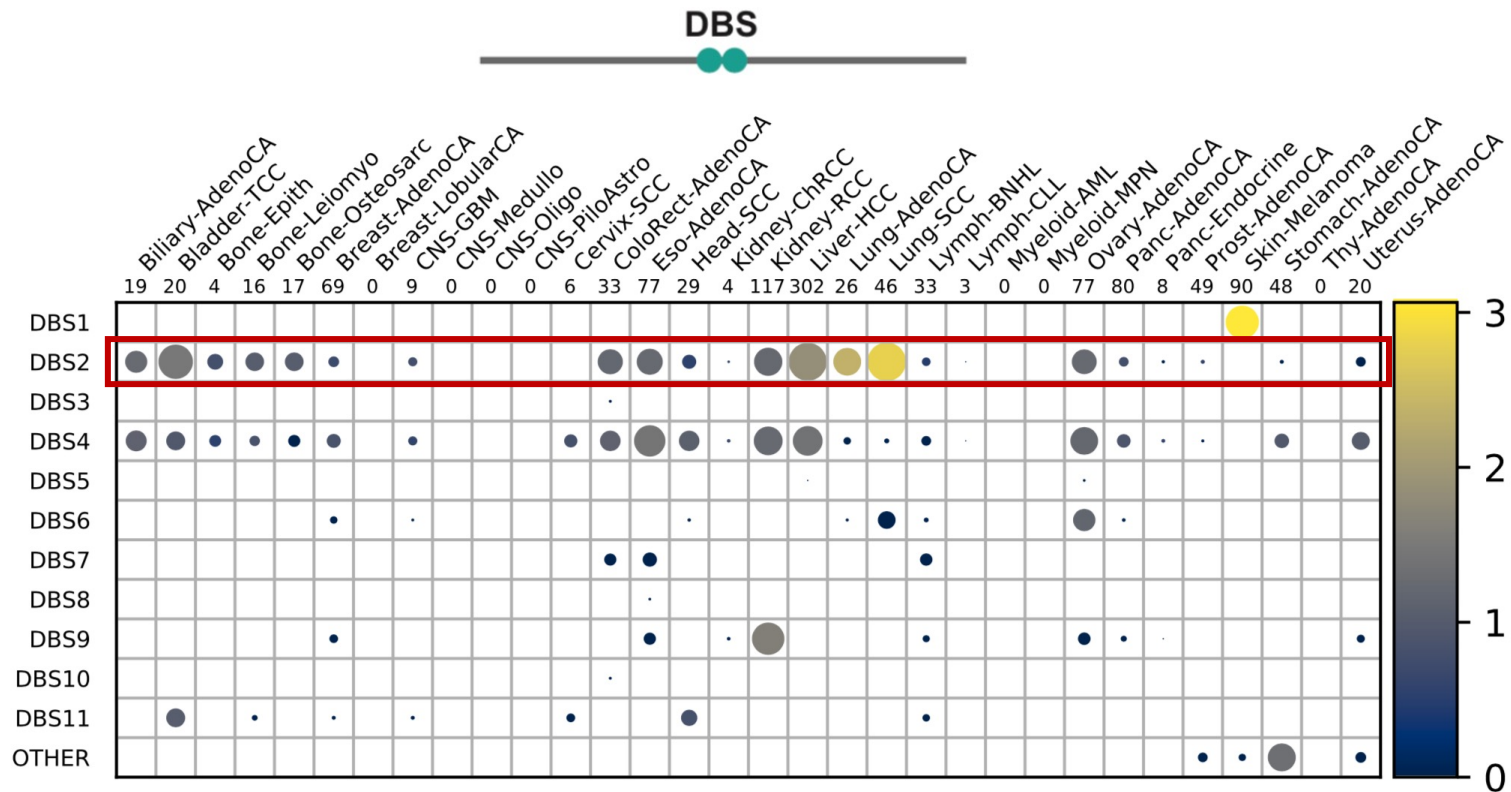
Bergstrom *et al.*, Nature 2022



# Mutational processes underlying clustered events



**DBS**

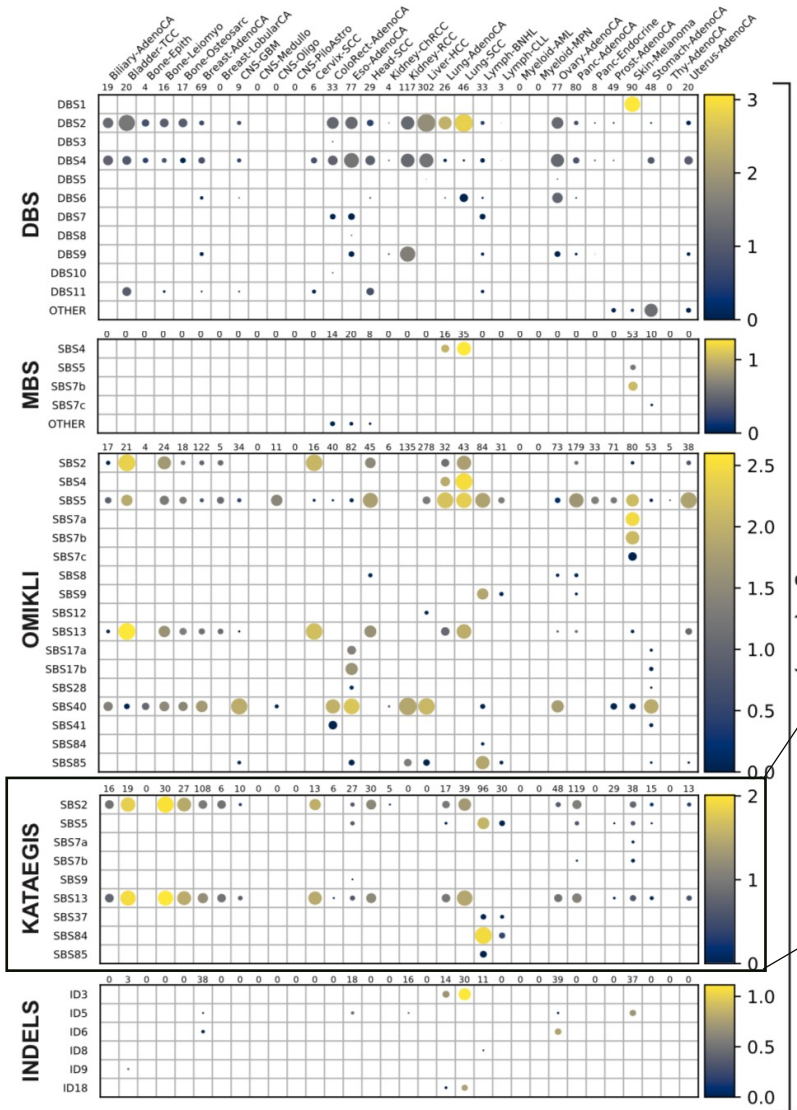


Bergstrom *et al.*, Nature 2022

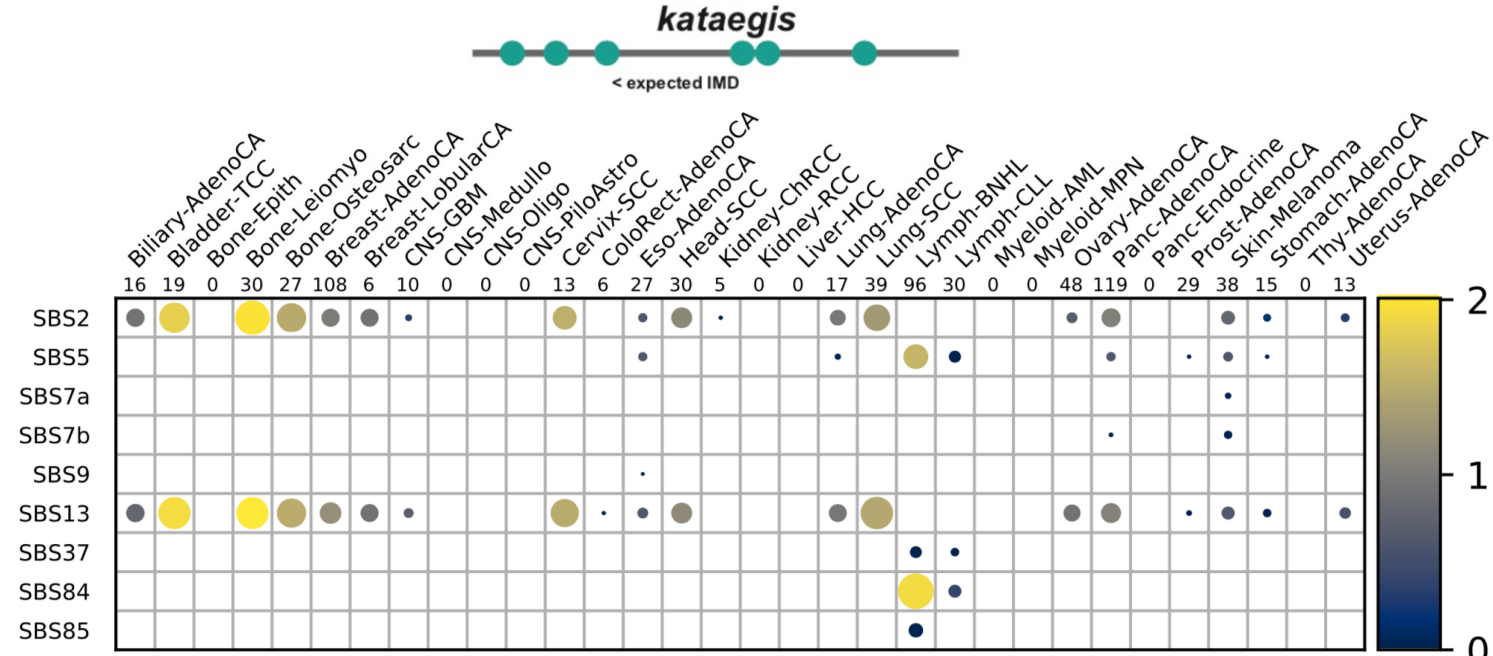




# Mutational processes underlying clustered events

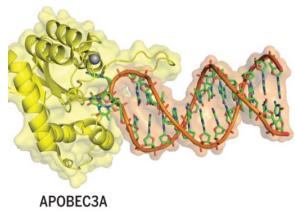
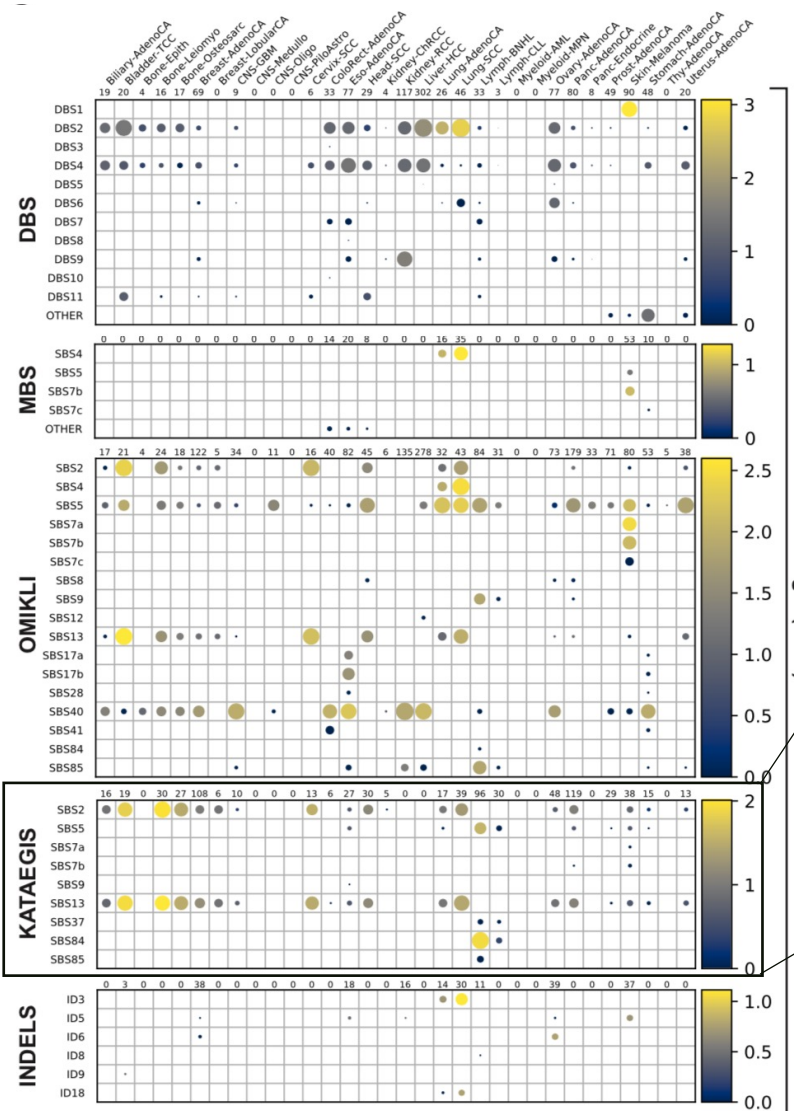


## KATAEGIS

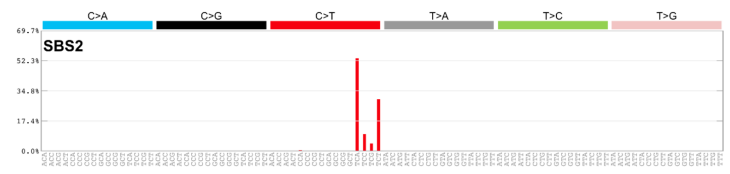
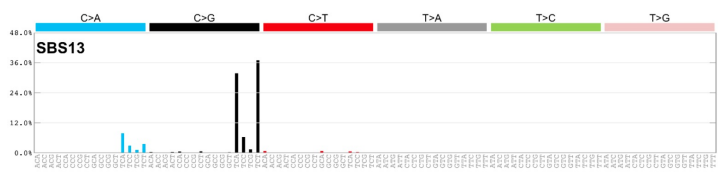
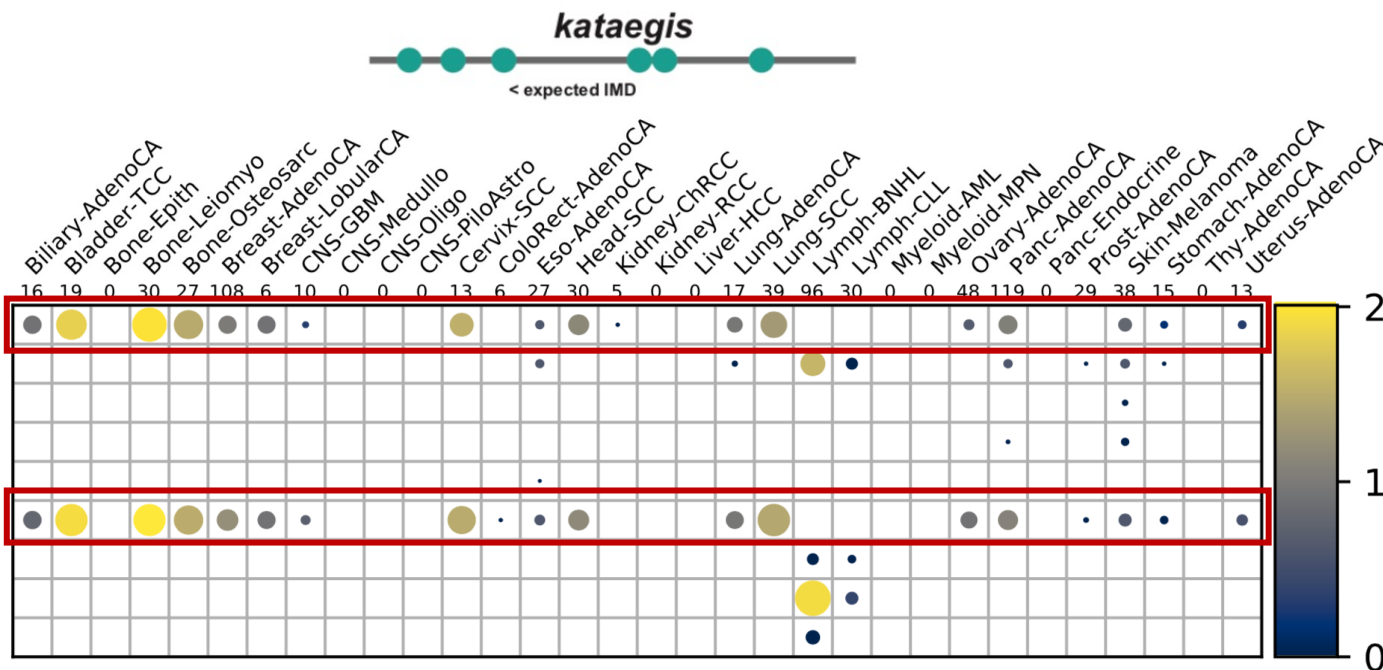


Bergstrom *et al.*, Nature 2022

# Mutational processes underlying clustered events



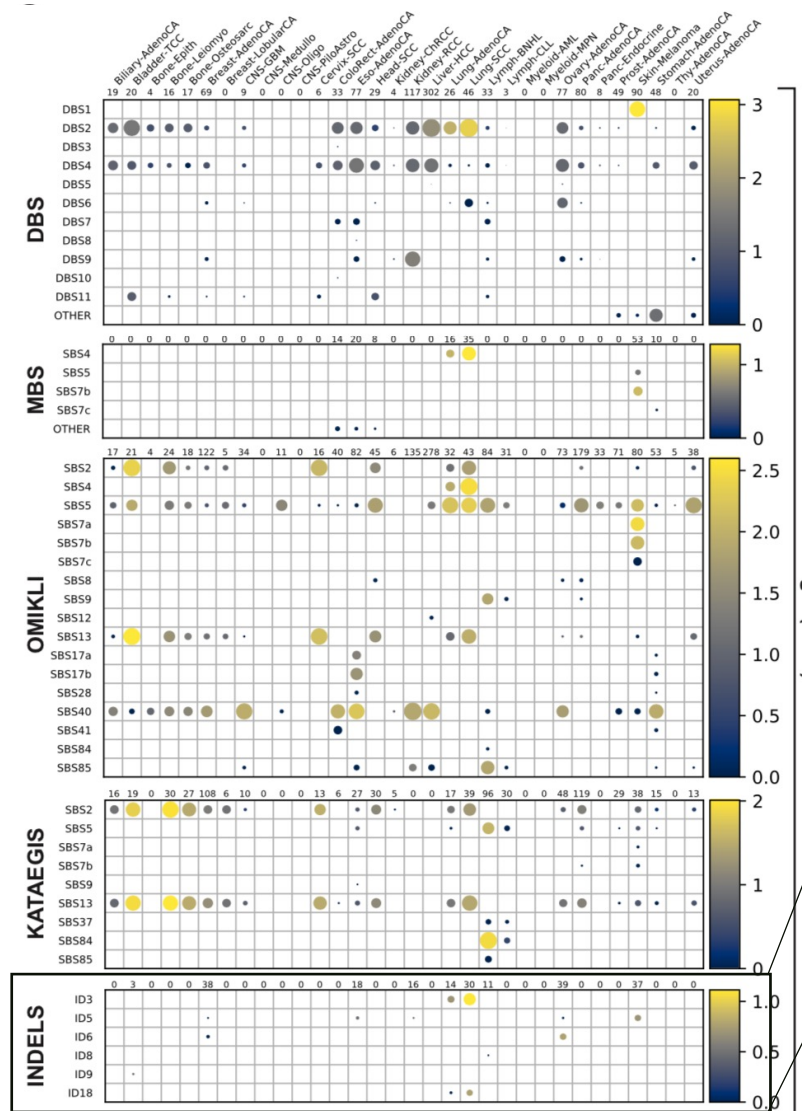
**KATAEGIS**



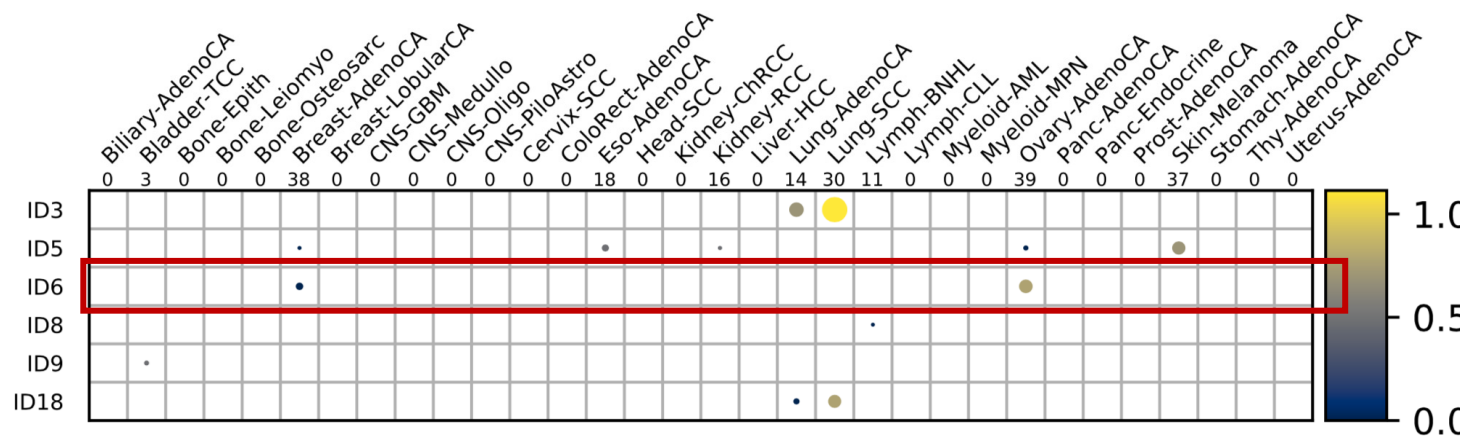
Bergstrom *et al.*, Nature 2022



# Mutational processes underlying clustered events

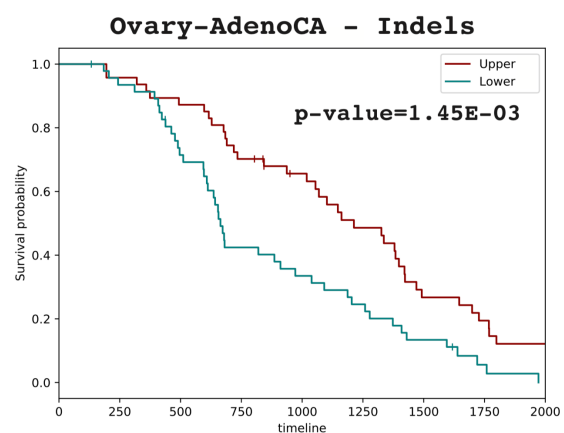
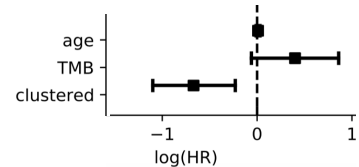


INDELS



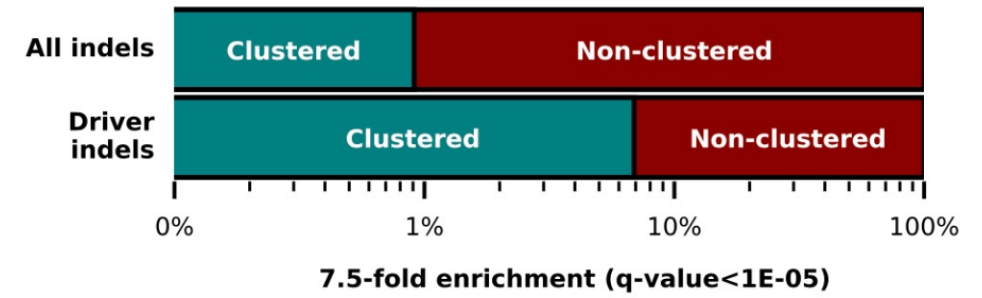
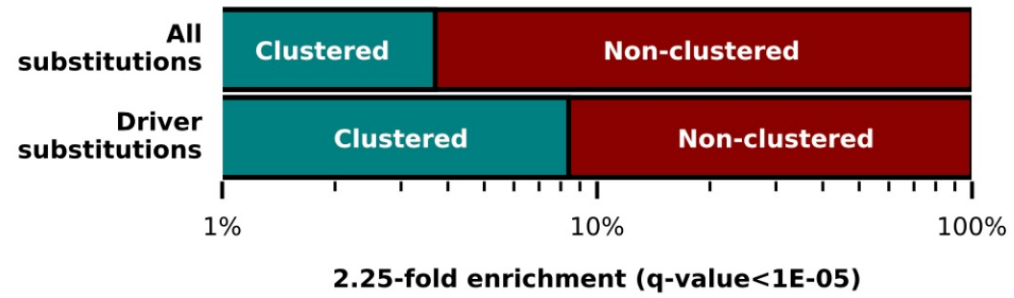
HRD

Ovary-AdenoCA  
Indels  
(q=2.5E-03)



Bergstrom *et al.*, Nature 2022

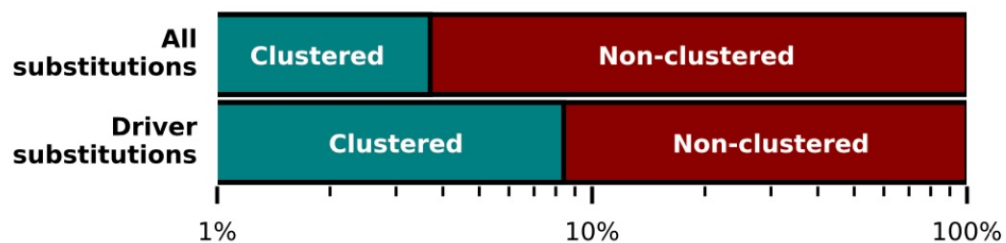
# Panorama of clustered driver mutations in human cancer



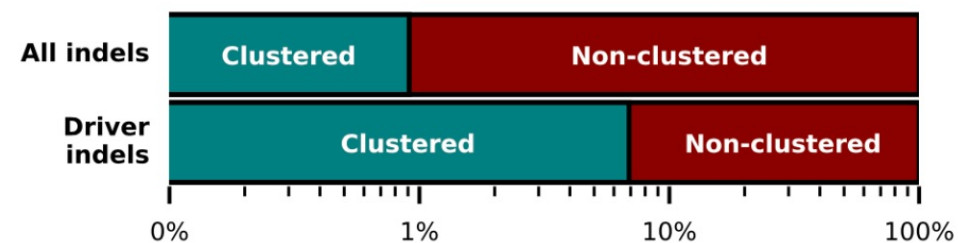
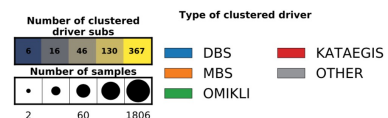
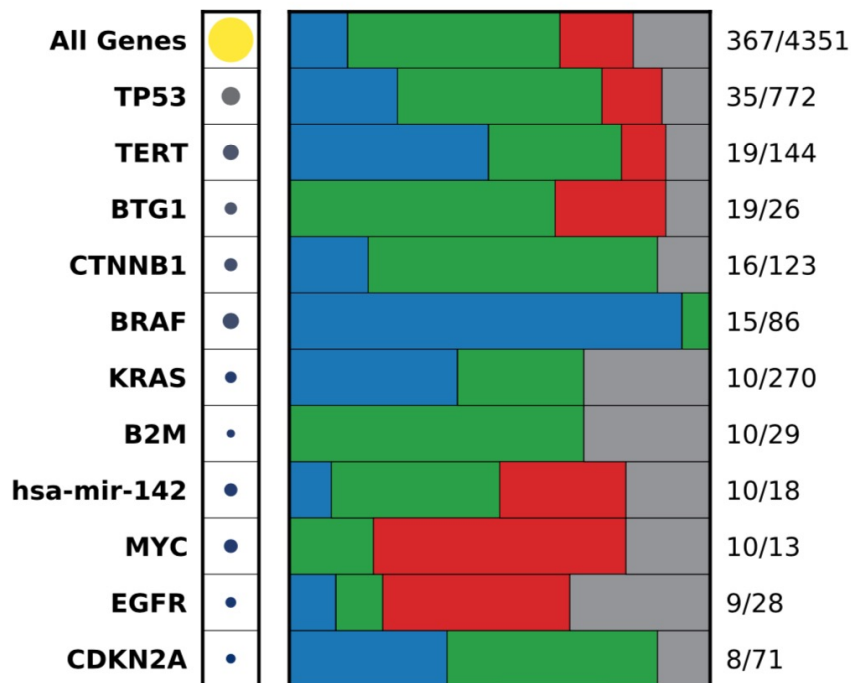
Bergstrom *et al.*, Nature 2022



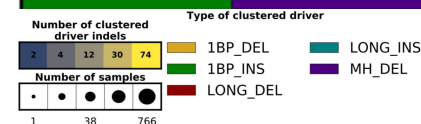
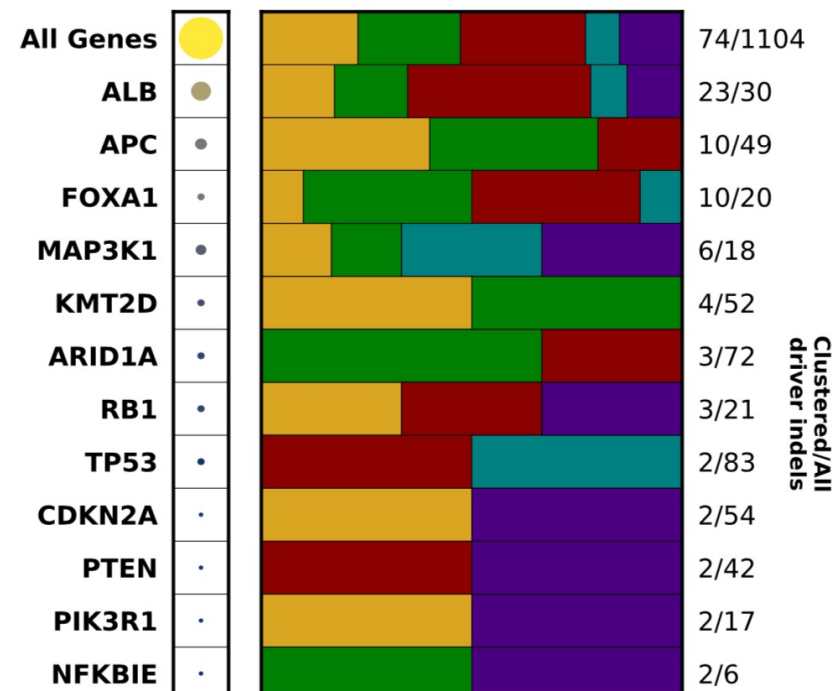
# Panorama of clustered driver mutations in human cancer



2.25-fold enrichment (q-value < 1E-05)

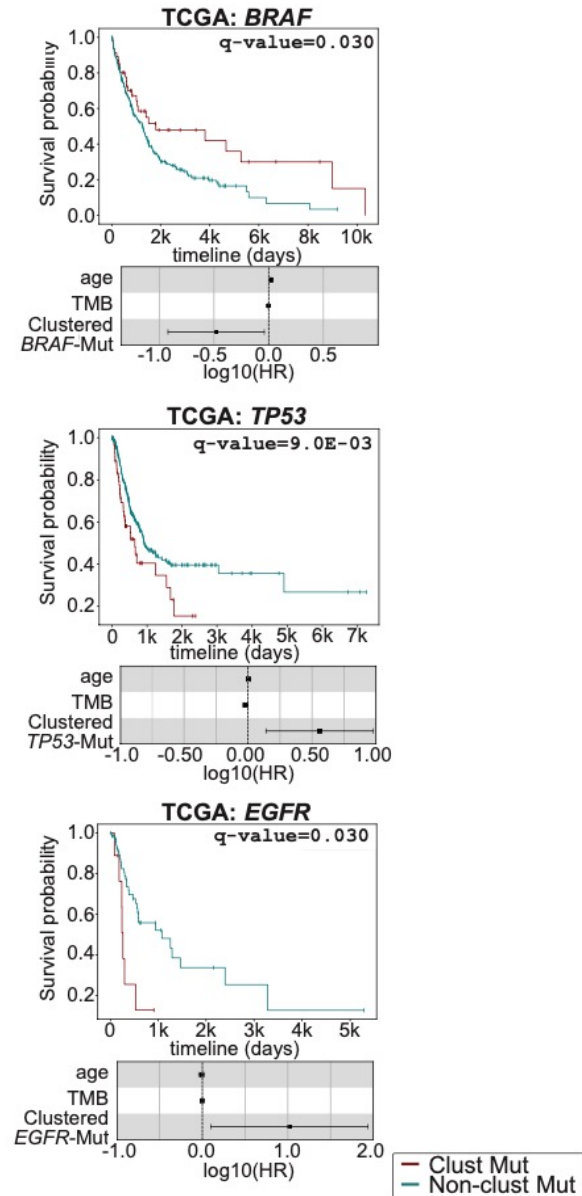


7.5-fold enrichment (q-value < 1E-05)



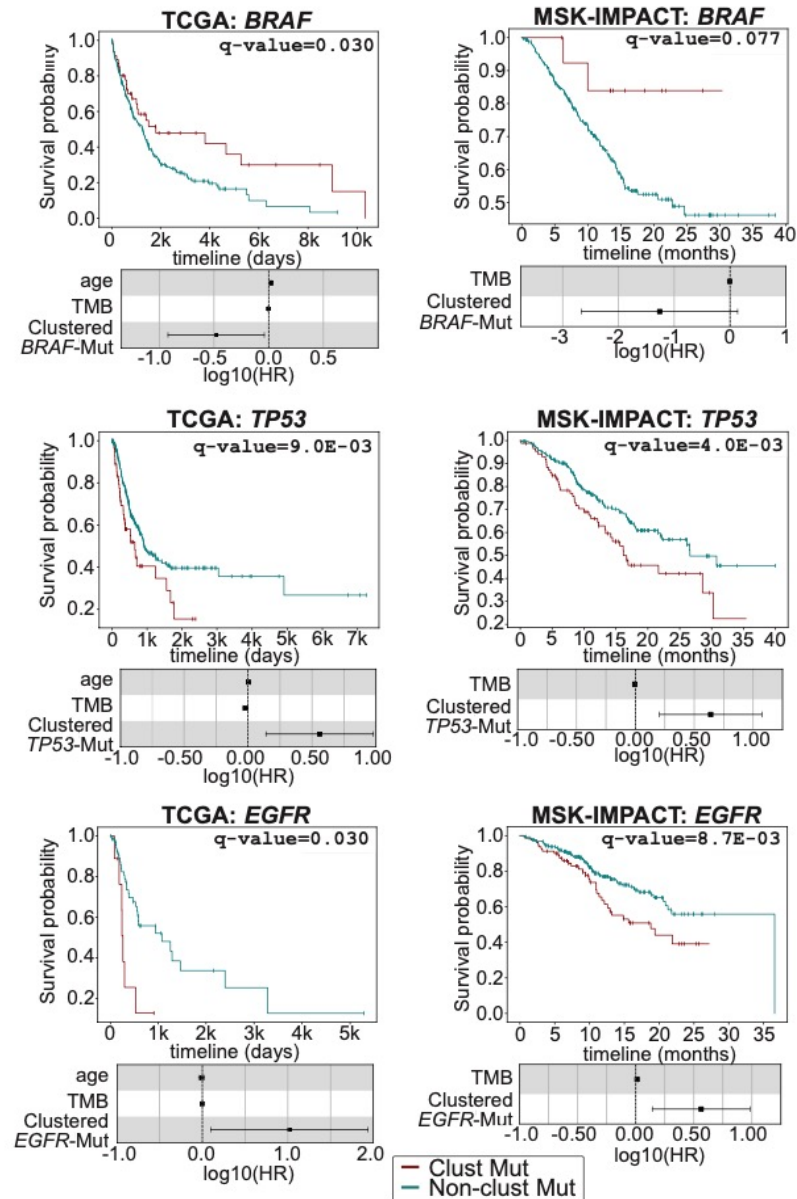
Bergstrom *et al.*, Nature 2022

# Clustered mutations in driver genes serve as a prognostic biomarker



Bergstrom *et al.*, Nature 2022

# Clustered mutations in driver genes serve as a prognostic biomarker



Bergstrom *et al.*, Nature 2022

# The repertoire of copy-number signatures in human cancer

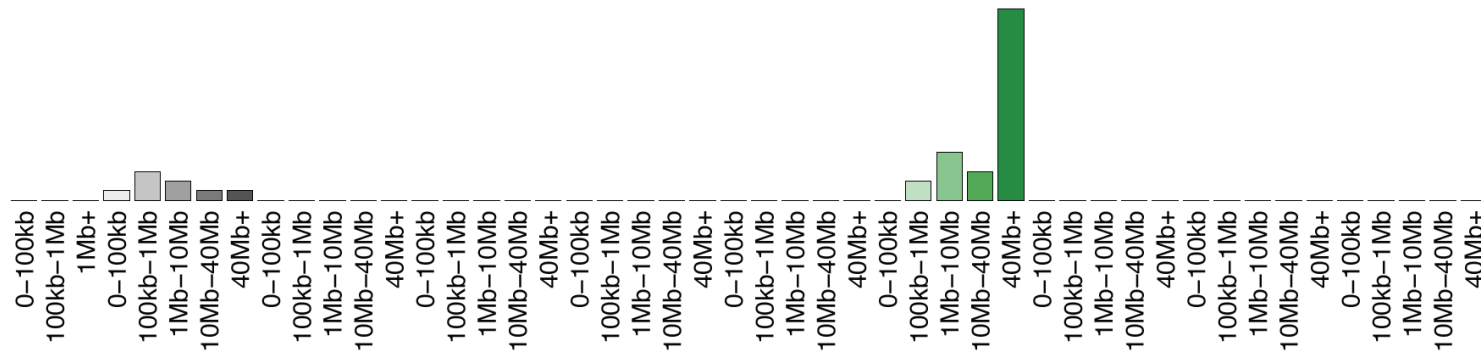
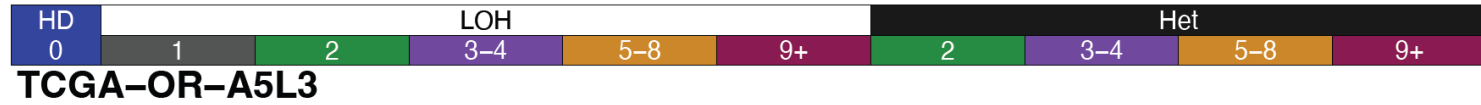
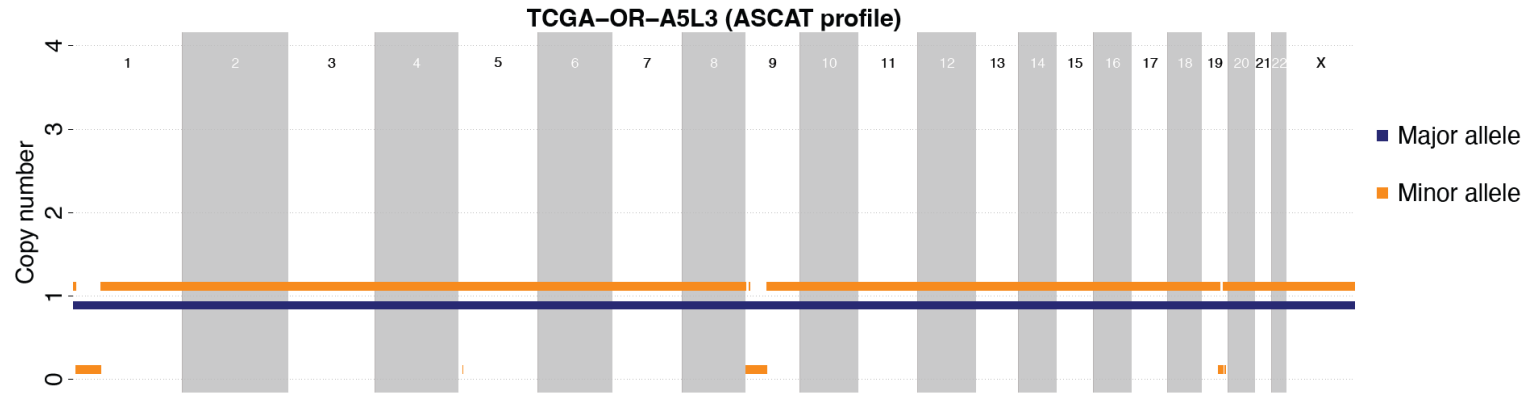


Chris Steele



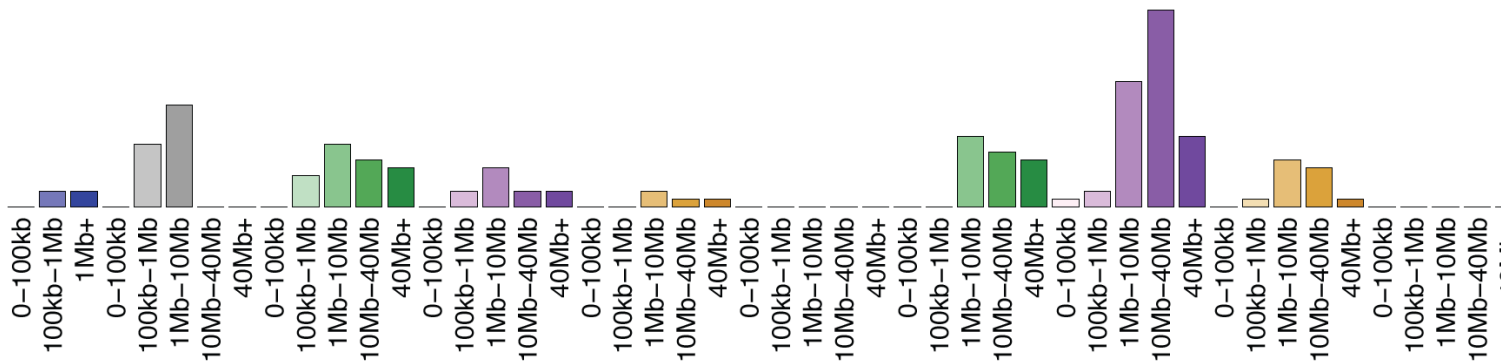
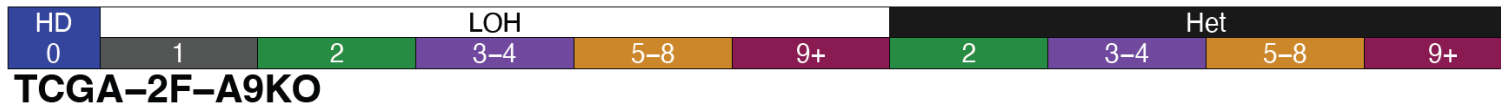
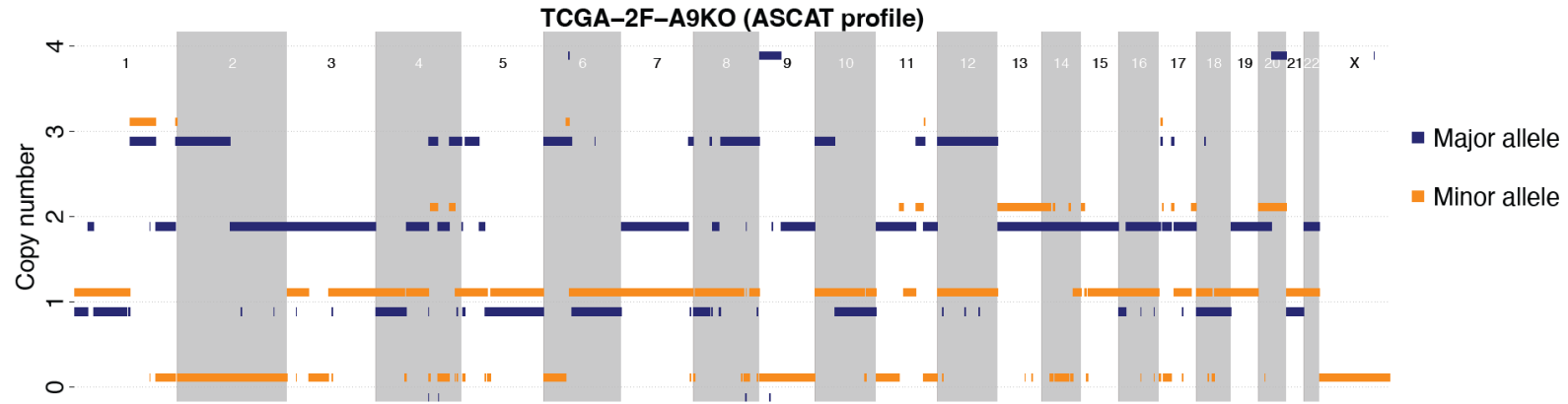


# From copy-number profiles to summarized copy-number patterns (example 1)



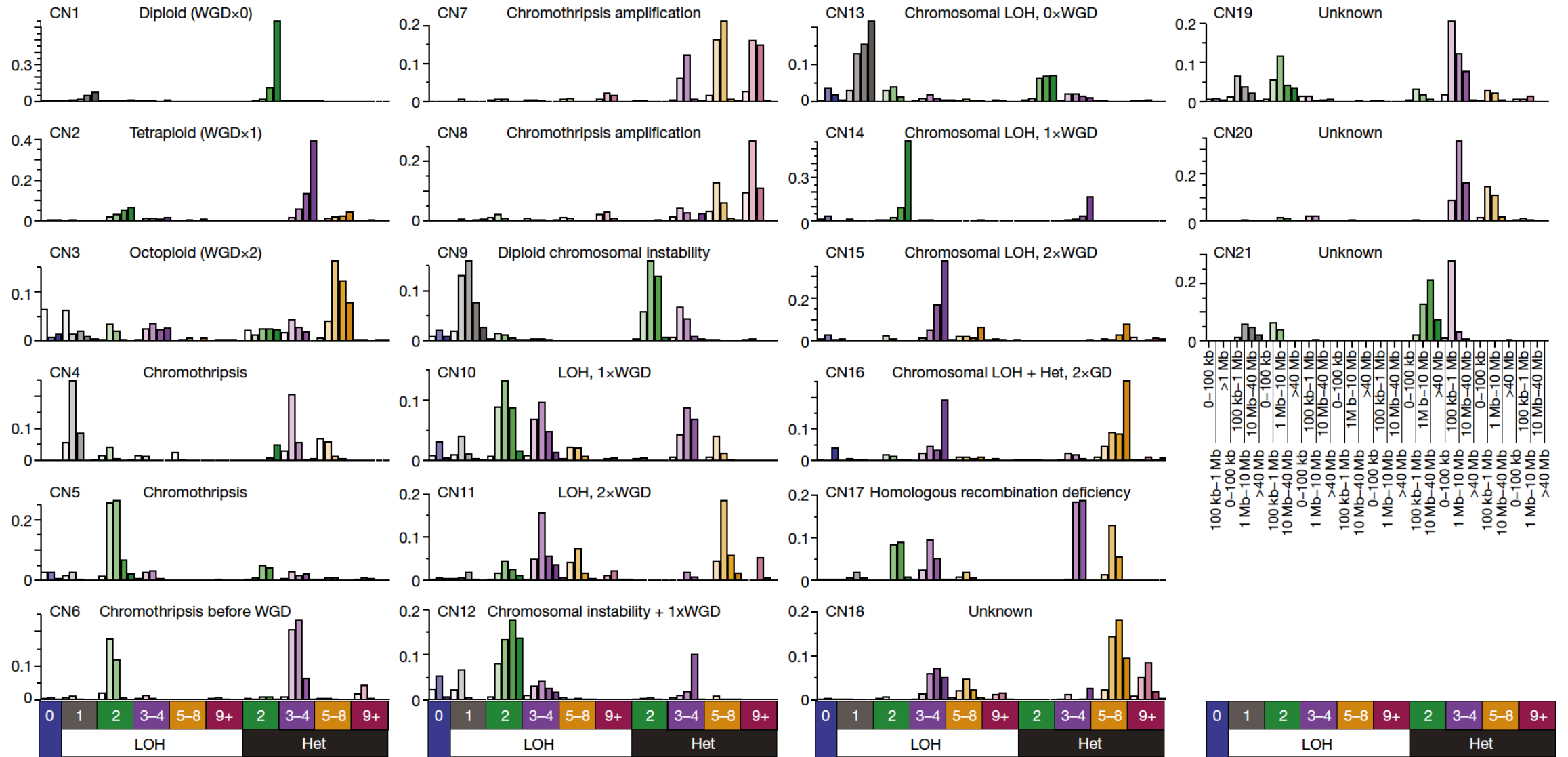
Steele et al., Nature, 2022

# From copy-number profiles to summarized copy-number patterns (example 2)



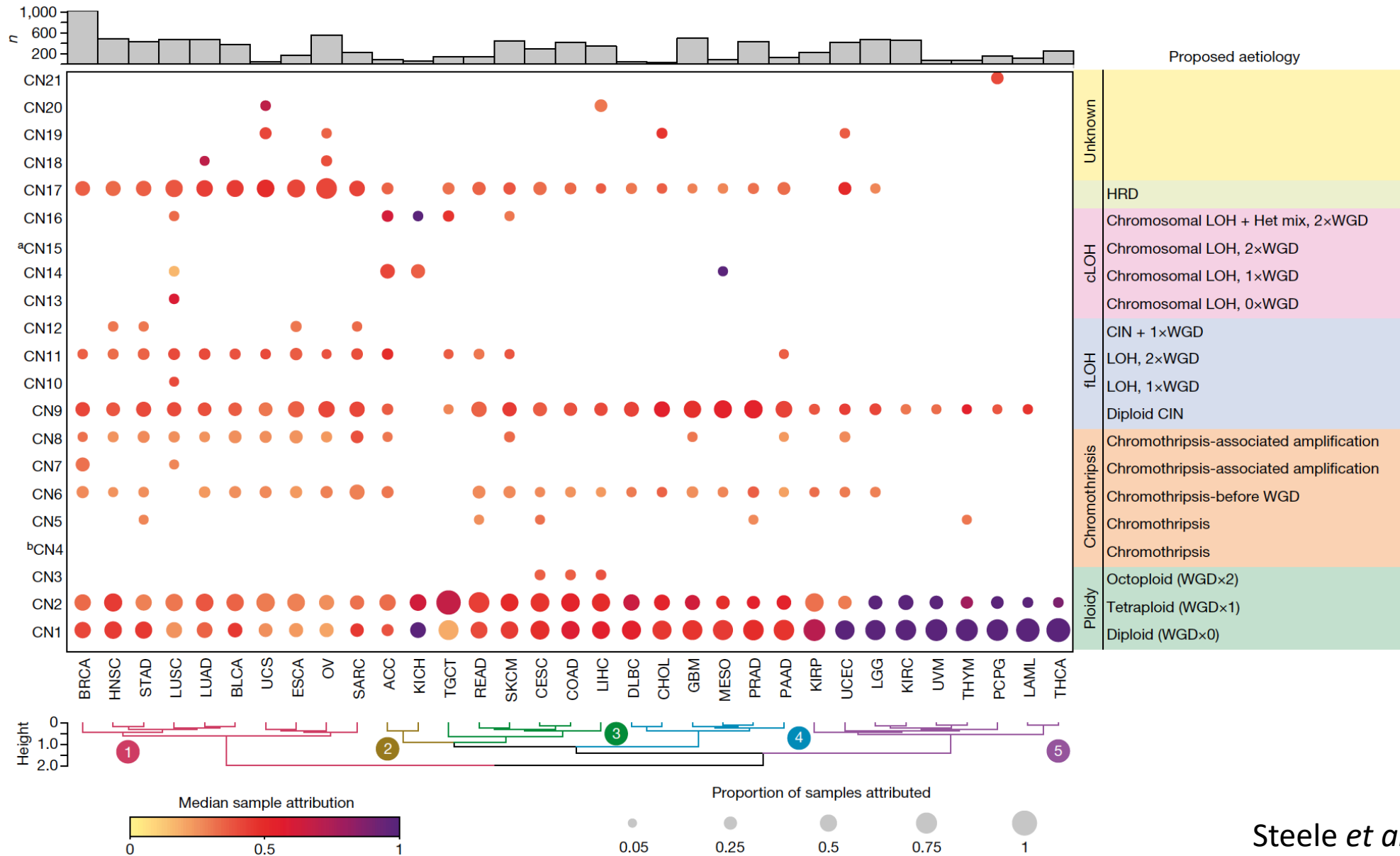
Steele et al., Nature, 2022

# From copy-number profiles to copy-number mutational signatures



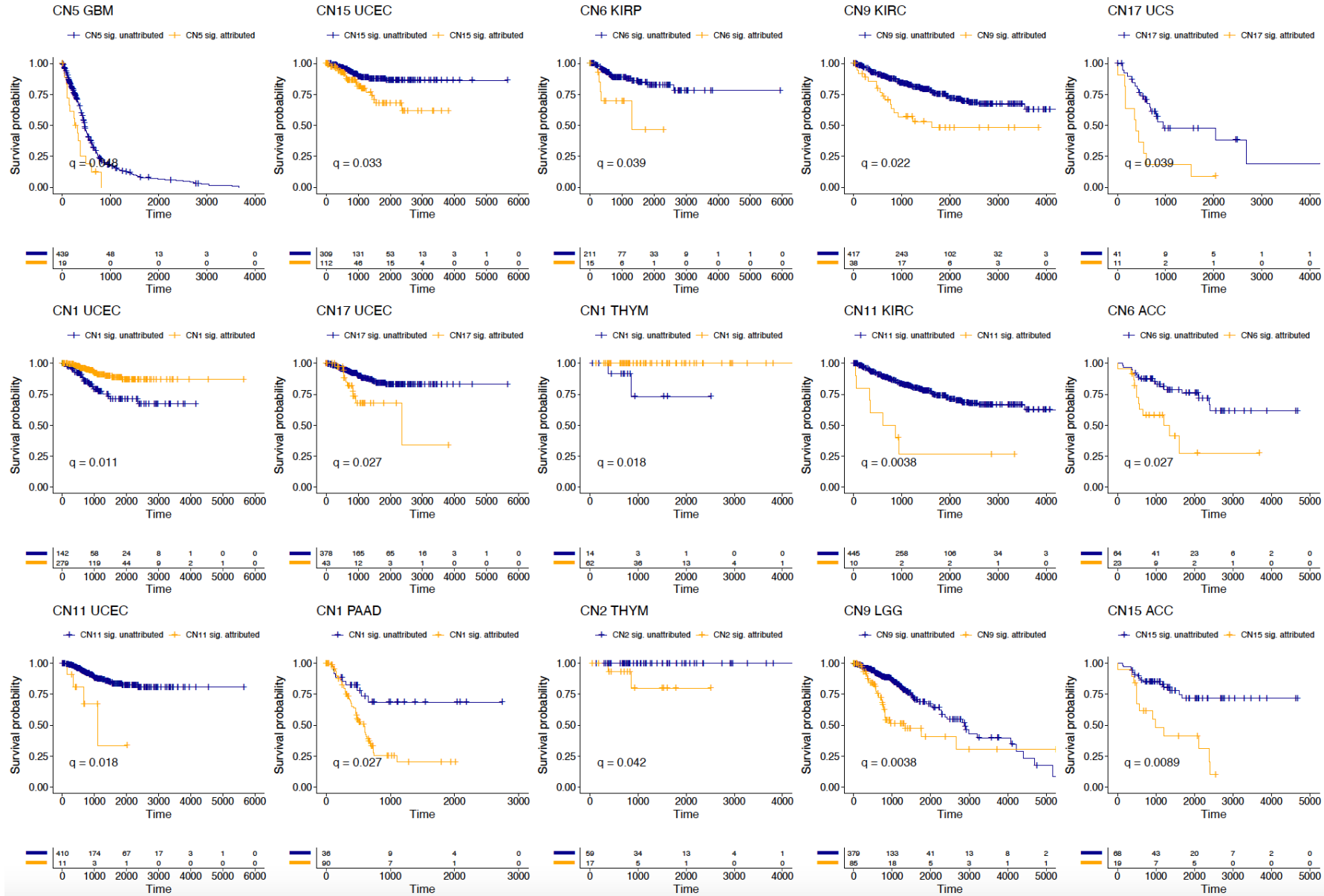
Steele et al., Nature, 2022

# Copy-number mutational signatures across human and their etiologies



Steele *et al.*, *Nature*, 2022

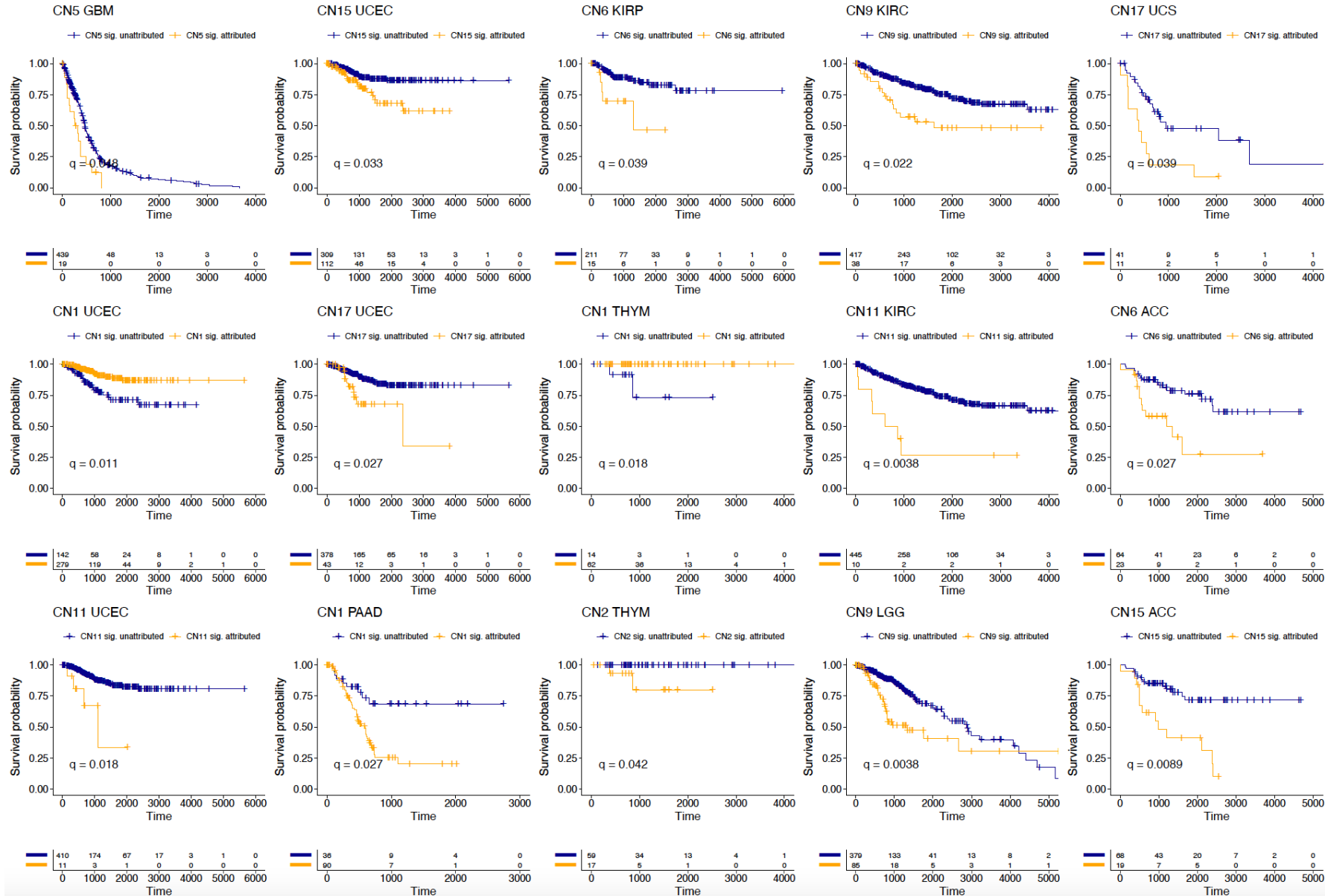
# Clinical utility of copy-number mutational signatures as prognostic biomarkers



Steele et al., Nature, 2022



# Clinical utility of copy-number mutational signatures as prognostic biomarkers



In contrast to other types of mutational signatures can be robustly detected from multiple platforms:

- Whole-genome sequencing
- Whole-exome sequencing
- Reduced-representation bisulfite sequencing
- Single-cell DNA sequencing
- SNP6 microarrays

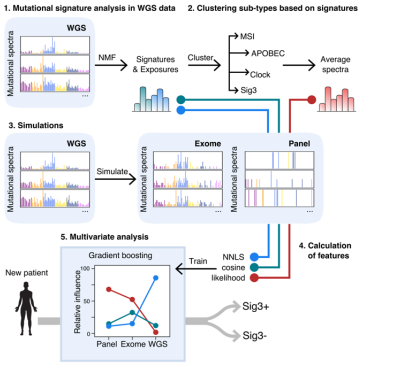
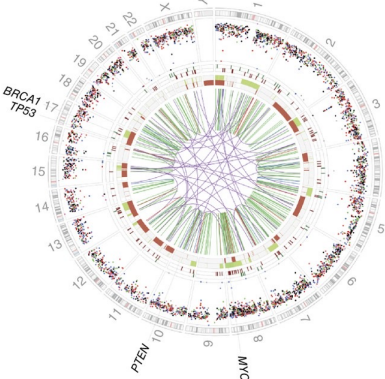
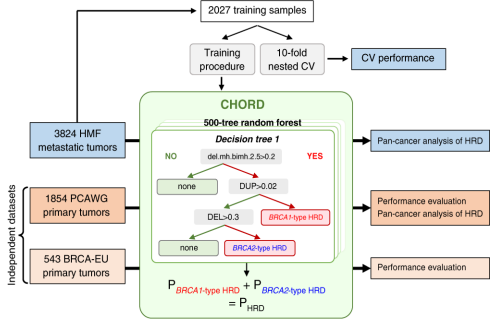
Steele *et al.*, *Nature*, 2022

# A novel machine learning approach for detecting homologous recombination deficiency



Ammal Abbasi

# HRD Prediction tools that use mutational signatures or mutational patterns

Tools	 <p style="text-align: center;"><b>SigMA</b></p> <p style="text-align: center;">Gulhan, D. <i>et al. Nat Genet</i> (2019)</p>	 <p style="text-align: center;"><b>HRDetect</b></p> <p style="text-align: center;">Davies, H. <i>et al. Nat Med</i> (2017)</p>	 <p style="text-align: center;"><b>CHORD</b></p> <p style="text-align: center;">Nguyen, L. <i>et al. Nat Commun</i> (2020)</p>
Features	SBS3	SBS3, SBS8, Microhomology-mediated deletions, RS3, RS5	SBS, ID, and SV mutational patterns
Sequencing Platform	WGS, WES, Panels	WGS	WGS
Advantages & Limitations	<p>Method can be applied to WGS, WES, and panel data.</p> <p>SBS3 is flat and method can be used only for highly mutated panels (~15% breast cancers).</p>	<p>Whole-genome sequencing is expensive approach especially at high-coverage. In many cases it requires fresh cancer tissues, and it is not commonly used in clinical practice.</p> <p>HRDetect &amp; CHORD can detect ~50% more samples that will respond to PARPi when compared myChoice CDx.</p>	

# Training iHRD with breast cancer samples

**Training dataset:** 234 genomically quiescent **whole-genome sequenced breast cancer samples** used as homologous recombination proficient (HRP). 77 BRCA1 or BRCA2 deficient **whole-genome sequenced breast cancer samples** used as homologous recombination deficient (HRD).

**Testing dataset (WGS):** 370 whole-genome sequenced breast cancer samples (77 HRD & 293 HRP).

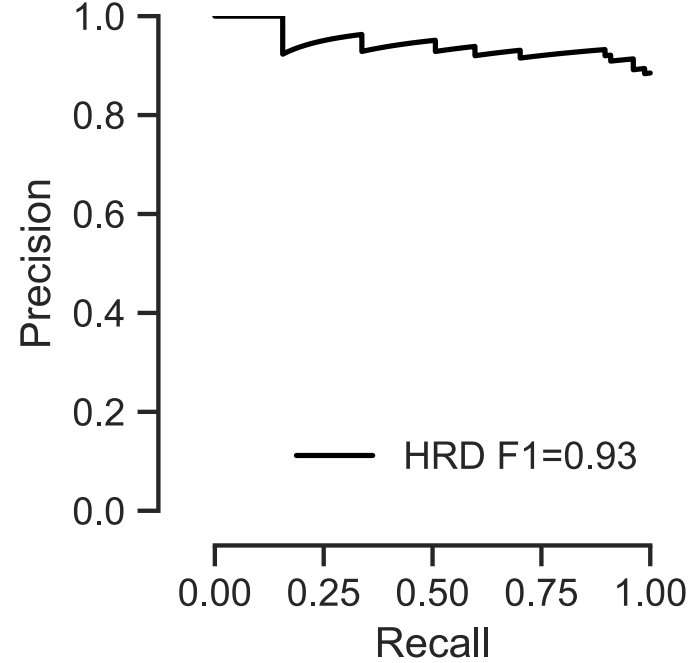
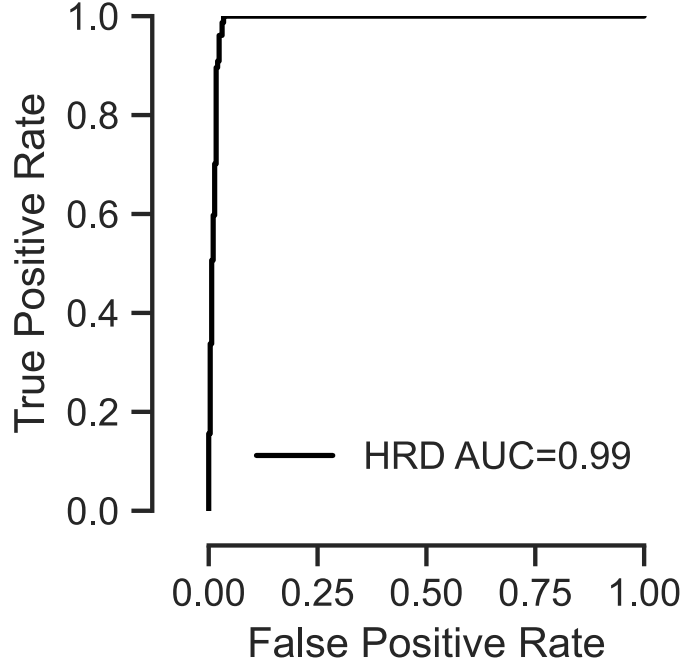
**Independent validation dataset (WGS):** 237 whole-genome sequenced triple-negative breast cancer samples (95 HRD & 142 HRP).

**Independent validation dataset (WES):** TCGA breast and ovarian whole-exome sequenced samples with consensus HRD and HRP status.

# Classifier performance on **test dataset** (370 WGS samples)

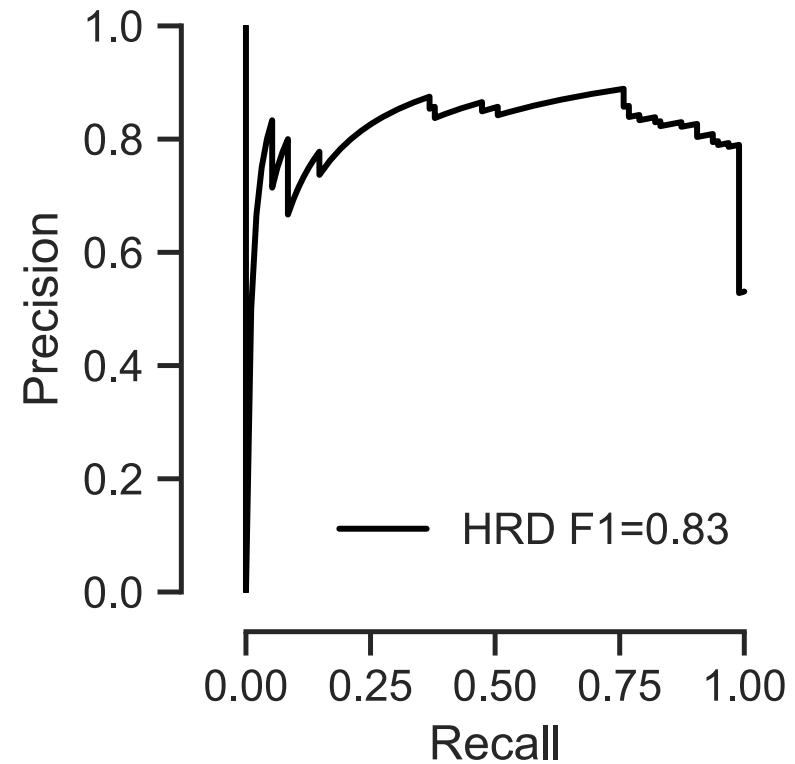
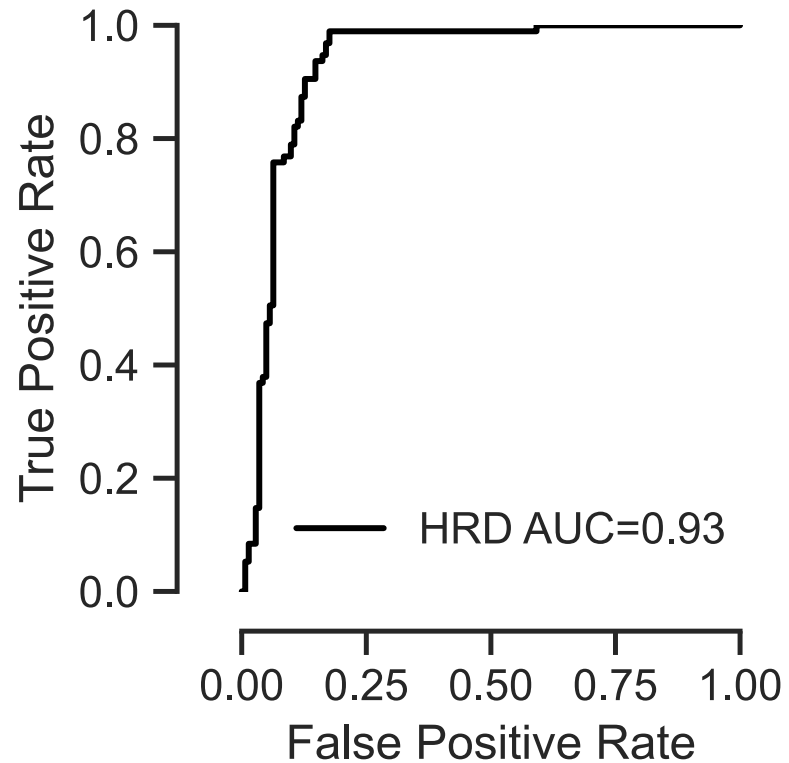
True labels	HRP	284	9
	HRD	2	75
		HRP	HRD

Predicted labels

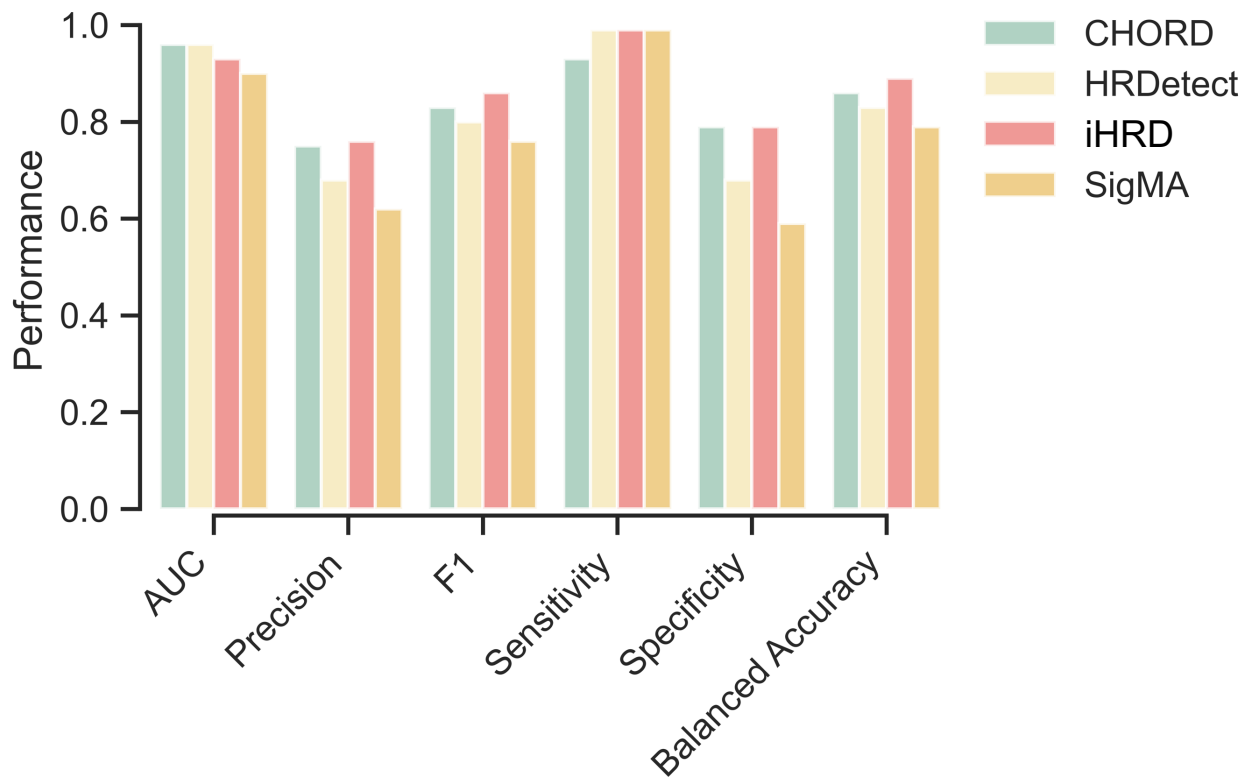




# Classifier performance on validation dataset (237 WGS TNBC)



# Model performance on validation **WGS dataset** across HRD genomic tools



**iHRD**

True labels	HRP	113	29
	HRD	1	94
		HRP	HRD
		Predicted labels	

**SigMA**

True labels	HRP	87	53
	HRD	2	93
		HRP	HRD
		Predicted labels	

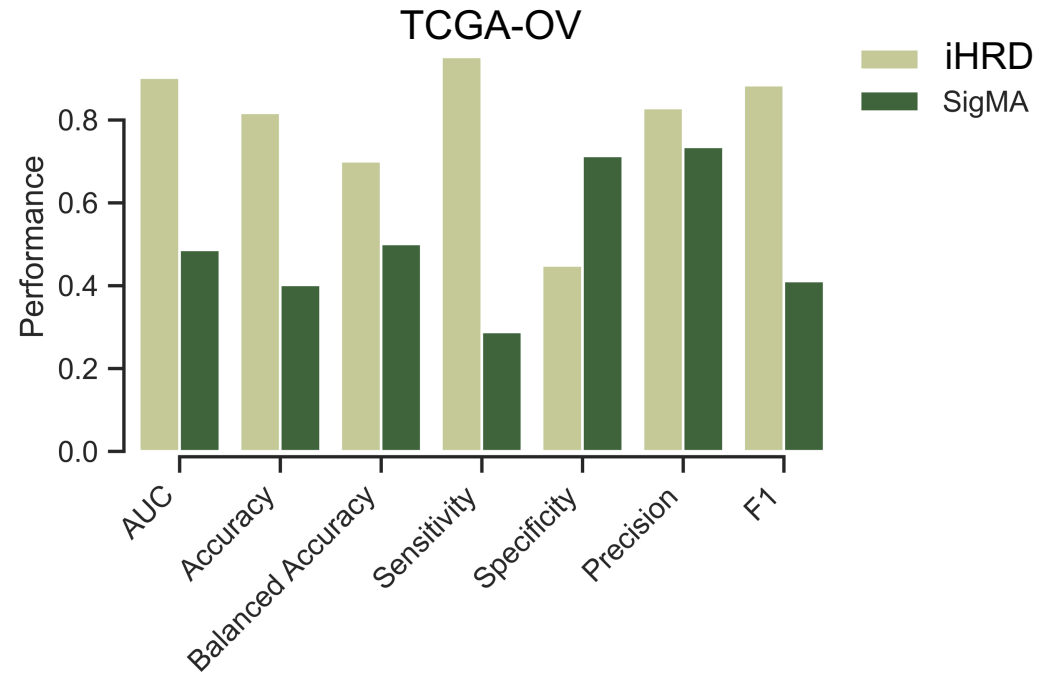
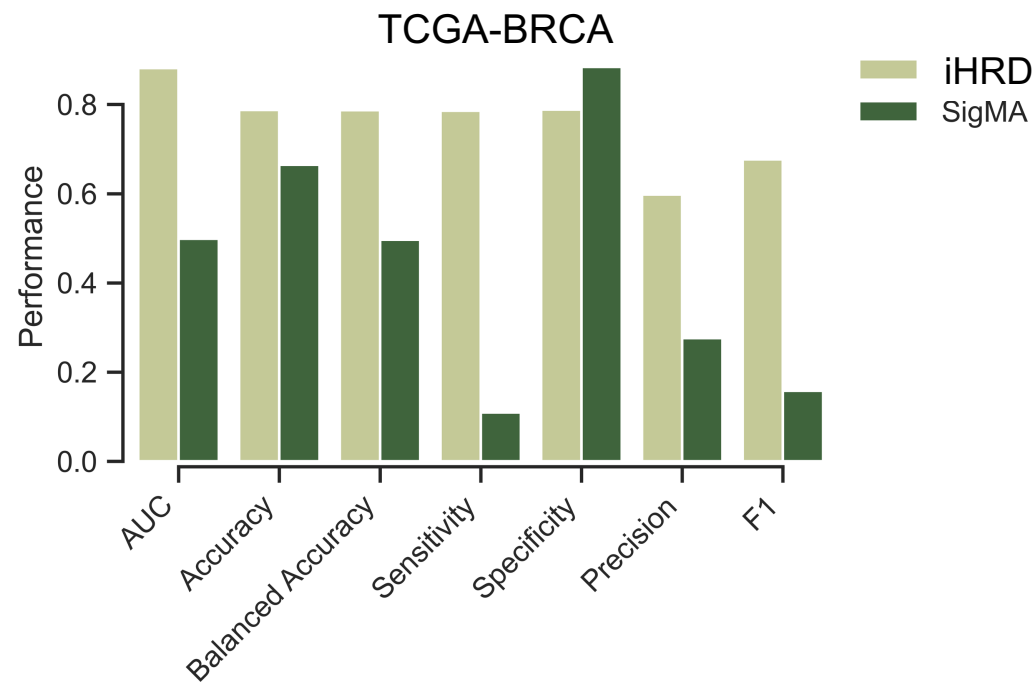
**CHORD**

True labels	HRP	105	35
	HRD	1	94
		HRP	HRD
		Predicted labels	

**HRDetect**

True labels	HRP	95	45
	HRD	1	94
		HRP	HRD
		Predicted labels	

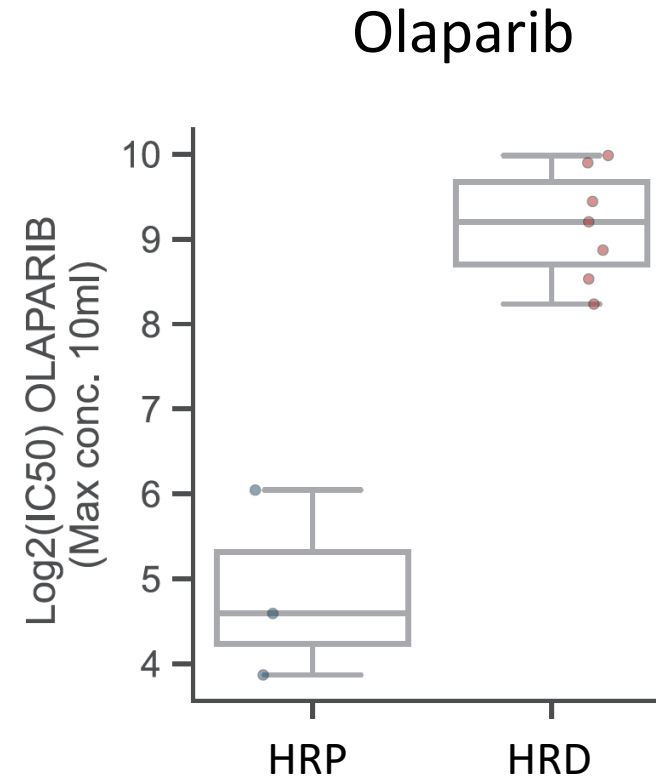
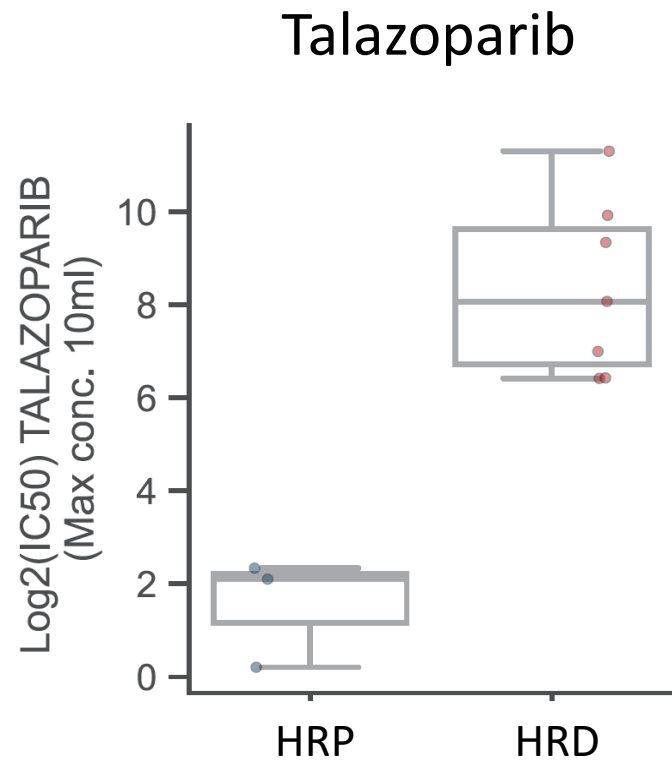
# Model performance on validation WES dataset across HRD genomic tools



iHRD uses model trained on whole-genome sequenced breast cancers.

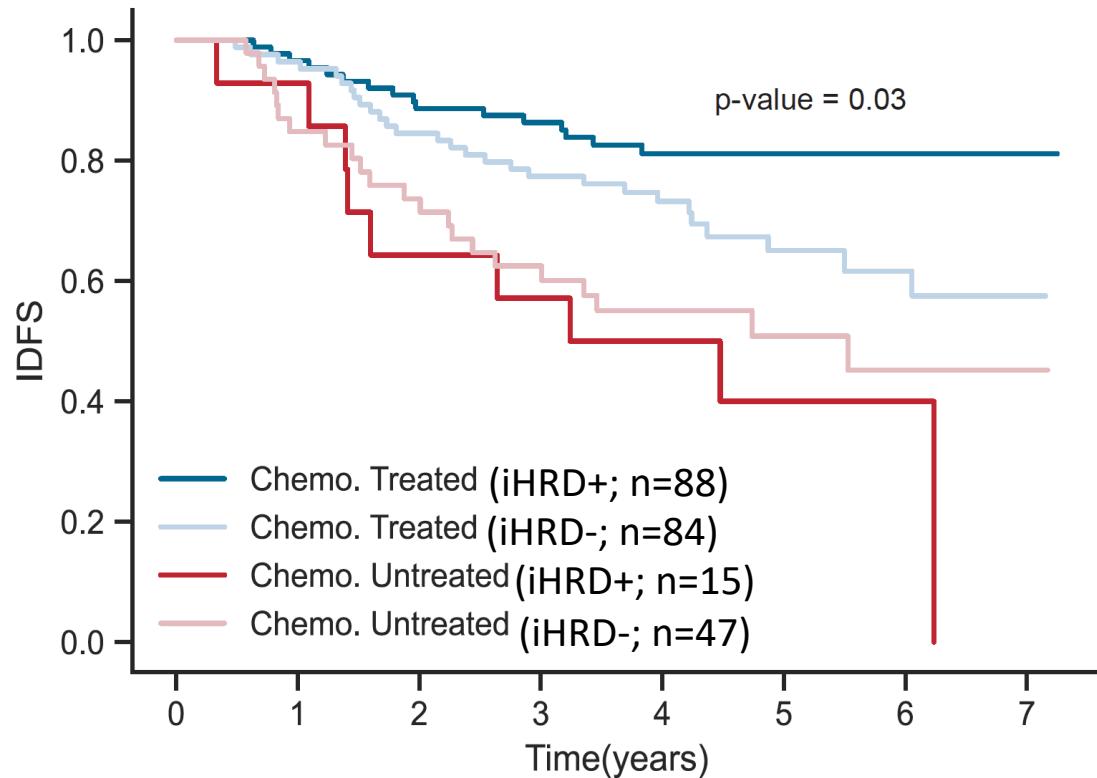
SigMA uses tissue-specific models trained on whole-exome sequencing data.

# Applying iHRD to exome sequenced cell lines with PARPi response

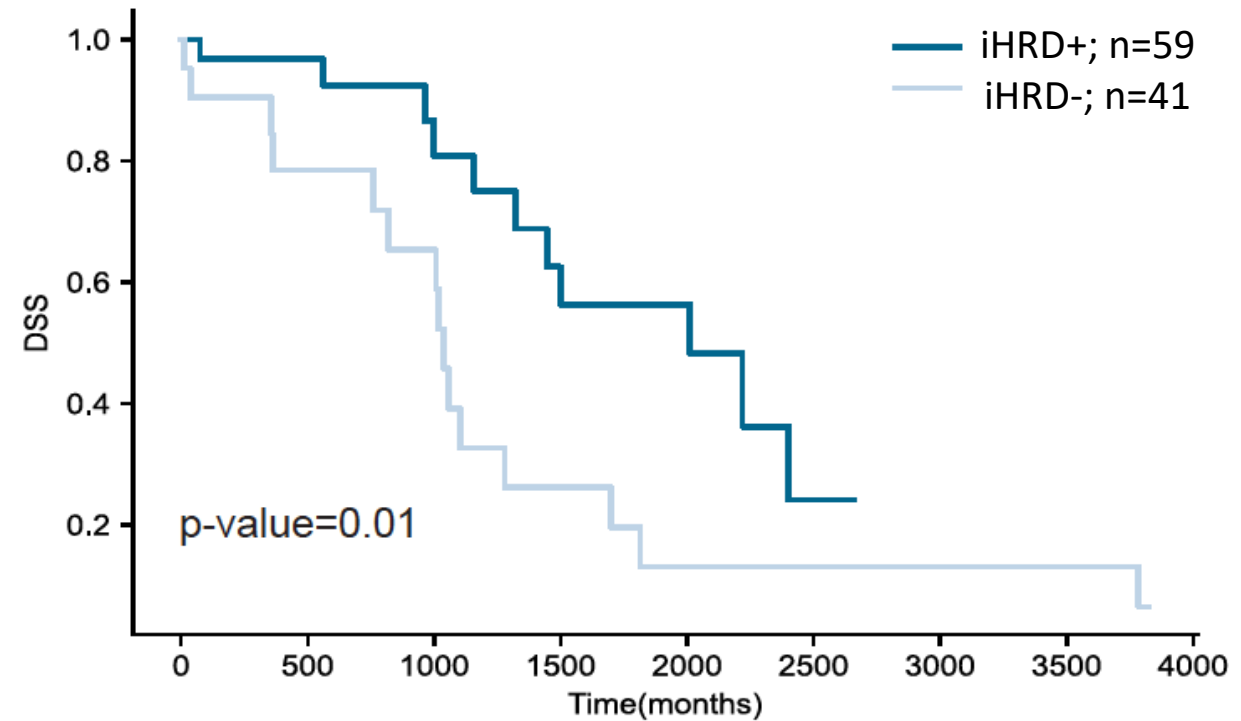


# Applying iHRD to exome sequenced retrospective clinical cohorts

## Breast Cancers & Platinum Therapy



## Ovarian Cancers & Platinum Therapy





# Ongoing iHRD work

- Applying to a breast cancer clinical cohort with known response to PARPi
- Applying to a prostate cancer clinical cohort with known response to PARPi
- Applying to a uterine sarcoma clinical cohort with known response to PARPi
- Working on extending its applicability to panel sequencing data

# Dreams of The Future

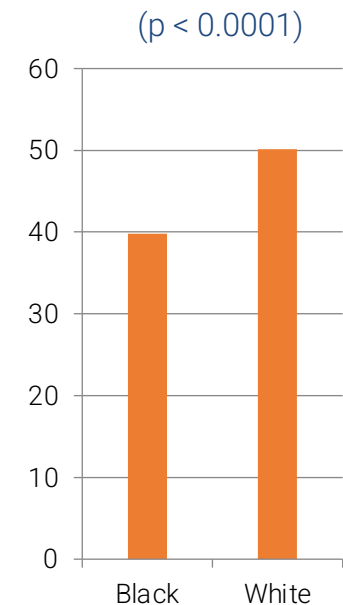
# Beyond genomics: Utilizing AI for addressing inequalities of cancer diagnosis



Erik Bergstrom

- NGS profiling is not available to all patients in the US and access outside the US is very limited
- Adoption of proven companion diagnostics is low due to cost:
  - Lung cancer biomarkers testing was first approved by the FDA in 2004
  - Recent data show that NGS testing rates in the US for the 5 SOC biomarkers is <50% overall.

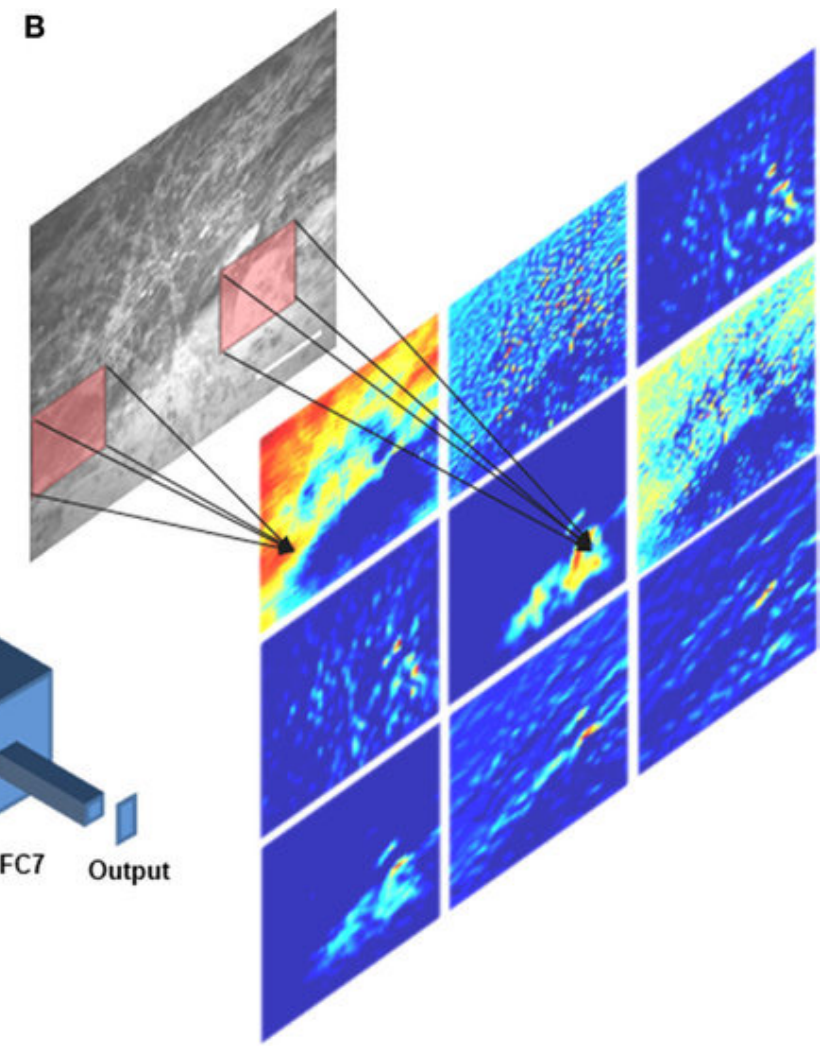
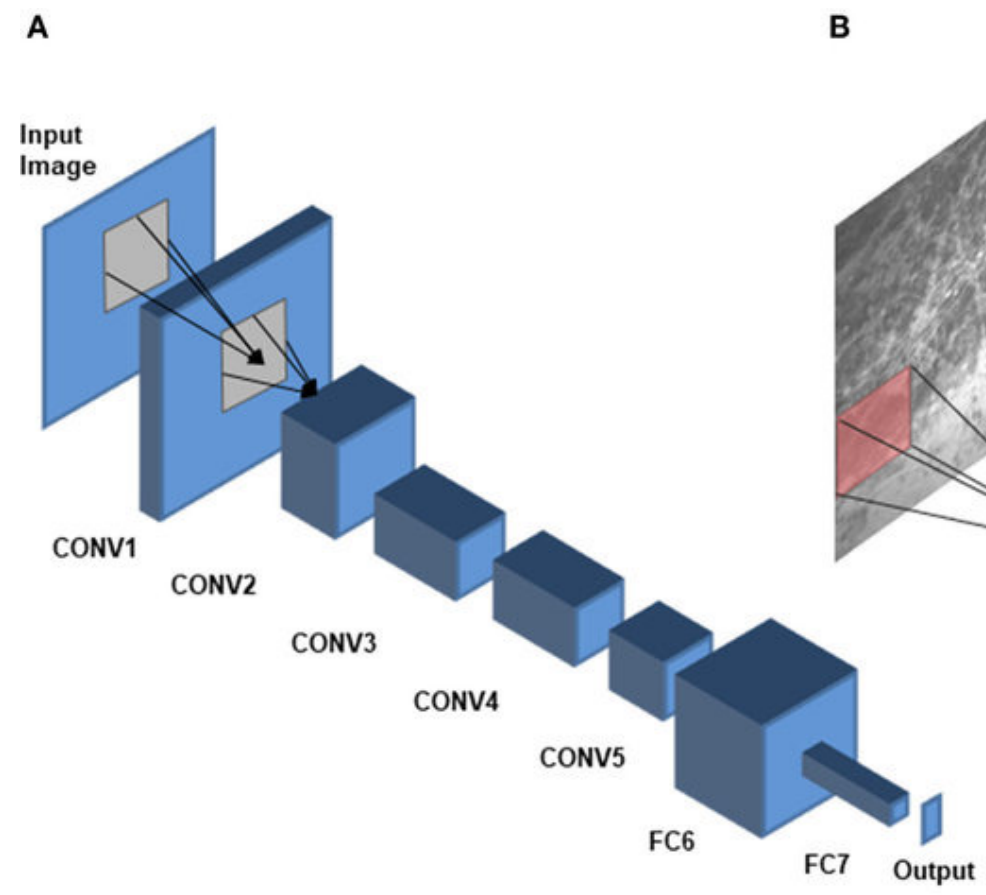
Roberts, N *et al* on behalf of the MYLUNG Consortium™ Collaborators: The US Oncology Network & Sponsors. ASCO Meeting, June 2021



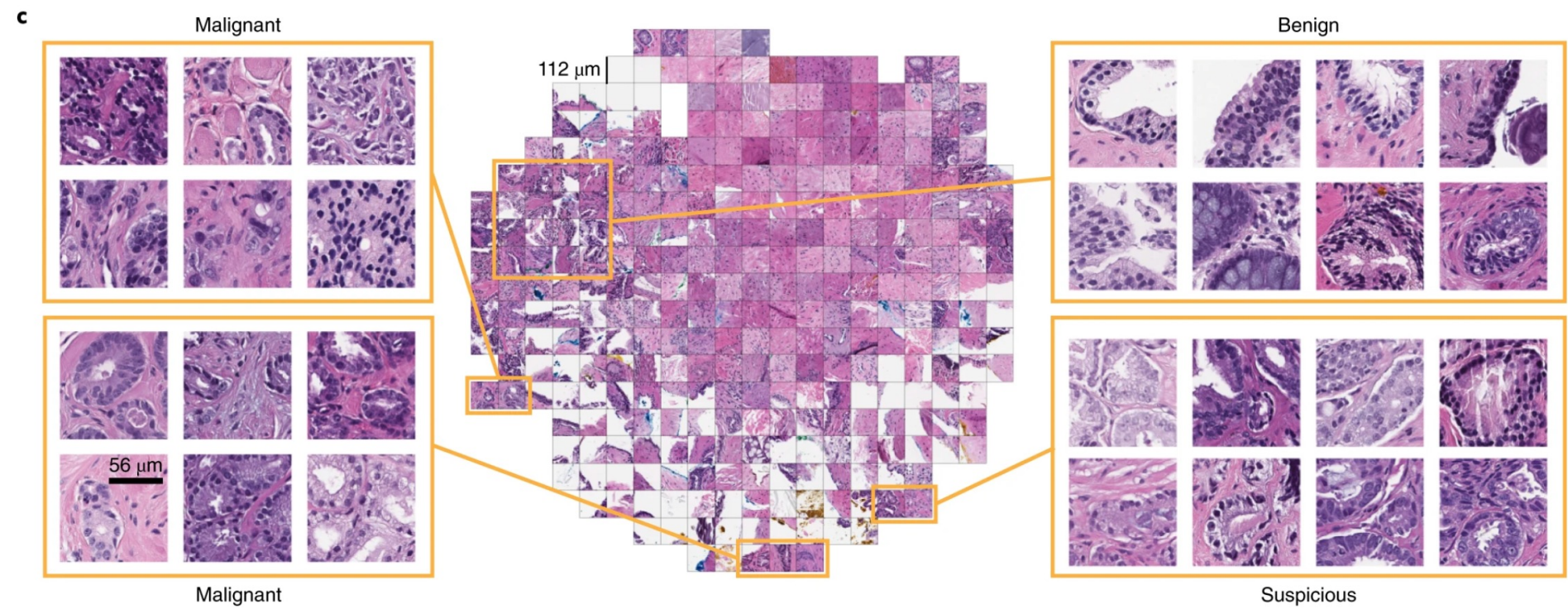
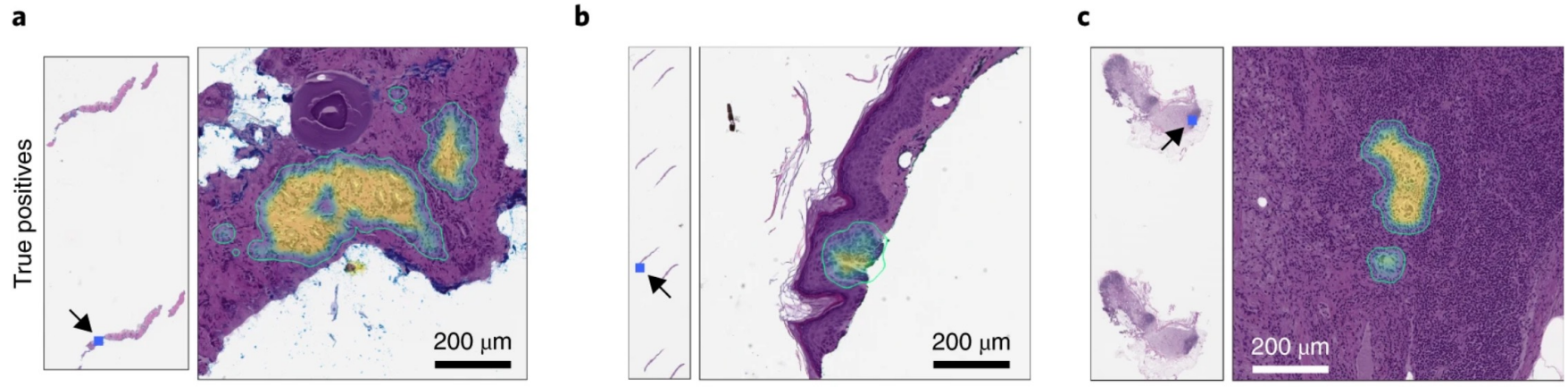
\*\*Bruno et al.; Roberts et al.,  
Lung Cancer, ASCO, June 2021

# Convolutional neural network to automatically extract image features

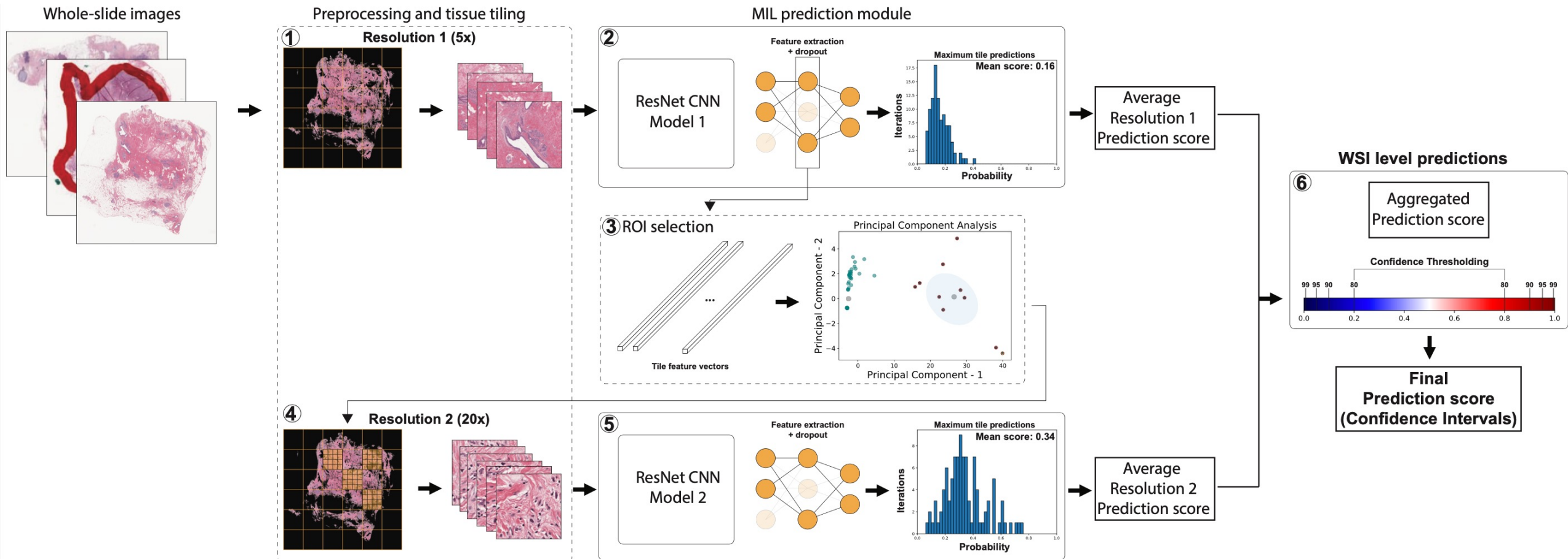
FFPE and frozen







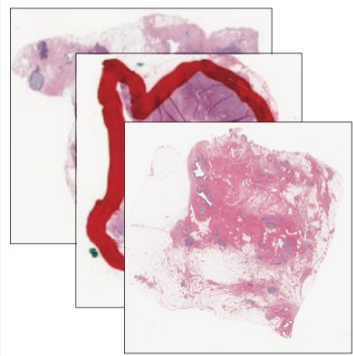
# Convolutional neural network to detect biomarkers from histopathological slides



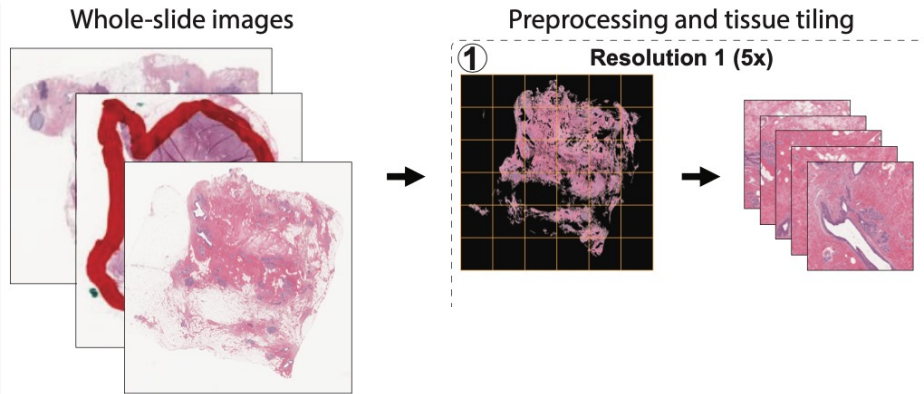
# *Convolutional neural network to detect biomarkers from histopathological slides*

---

Whole-slide images

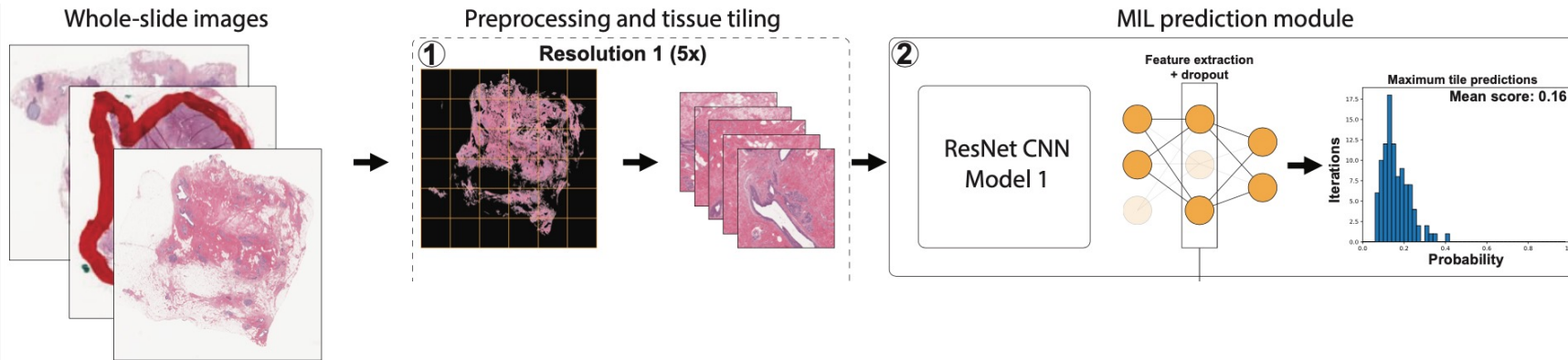


# Convolutional neural network to detect biomarkers from histopathological slides

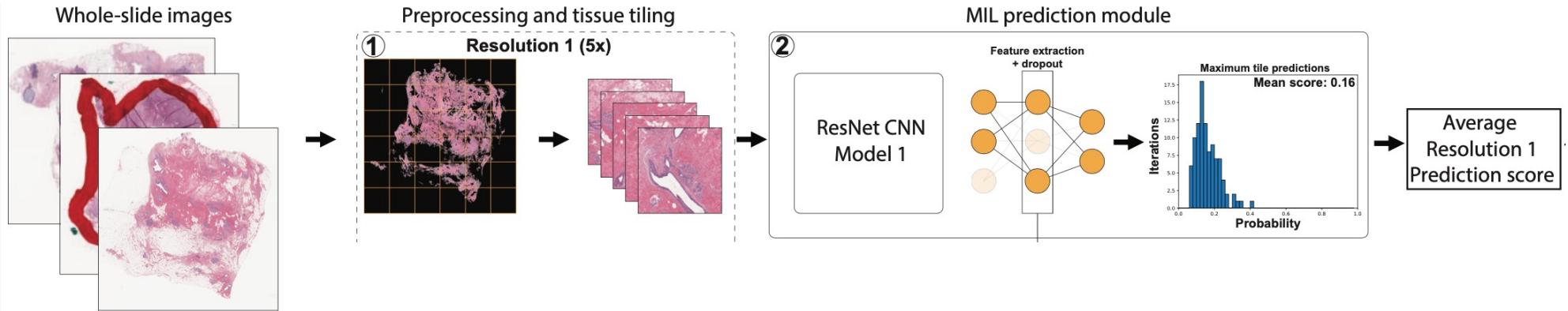




# Convolutional neural network to detect biomarkers from histopathological slides

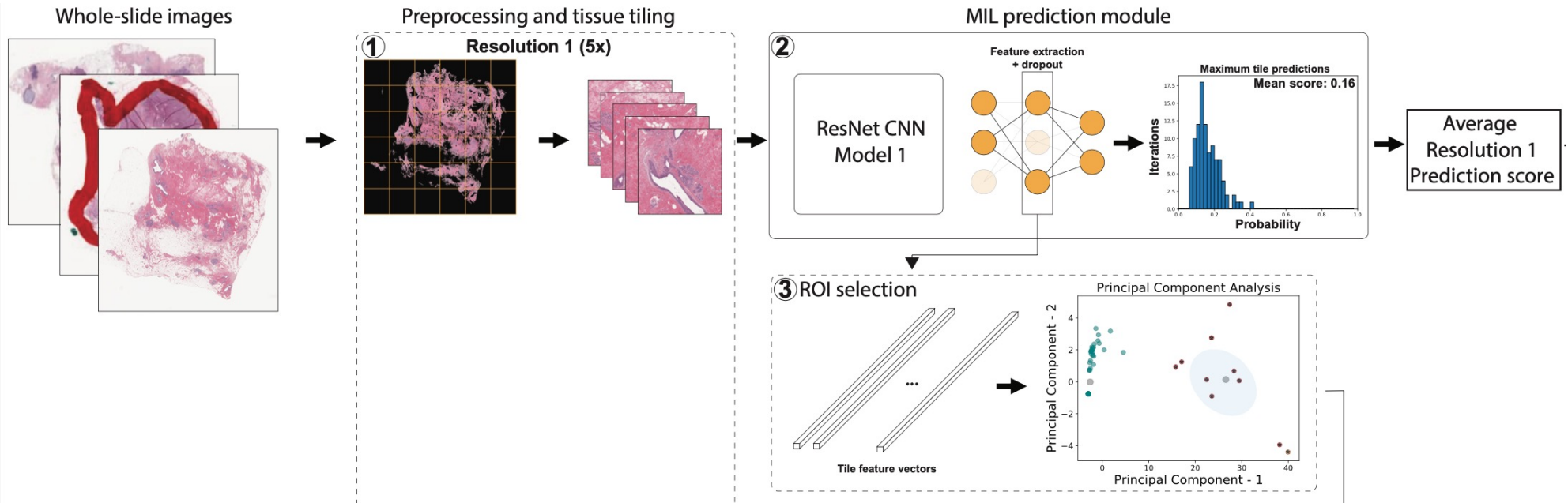


# Convolutional neural network to detect biomarkers from histopathological slides

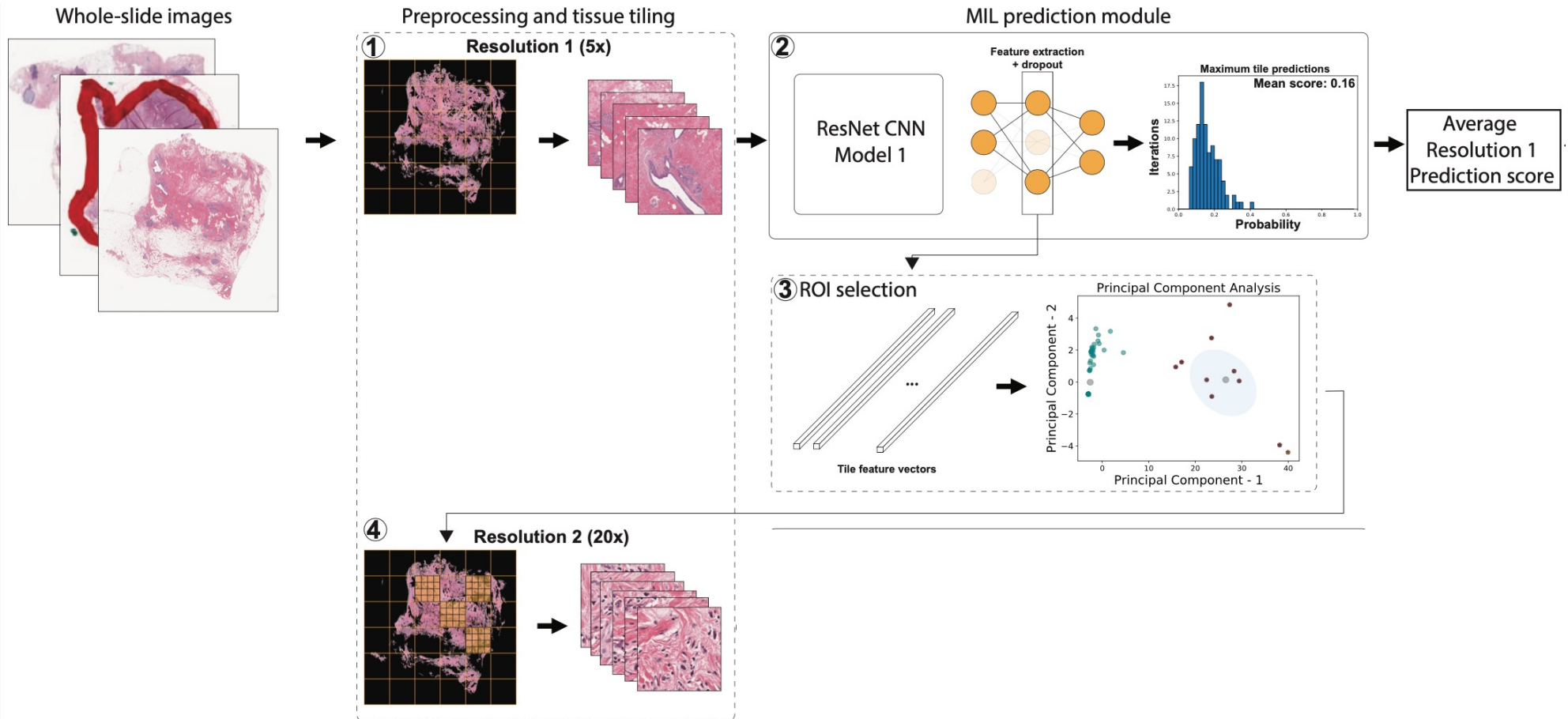




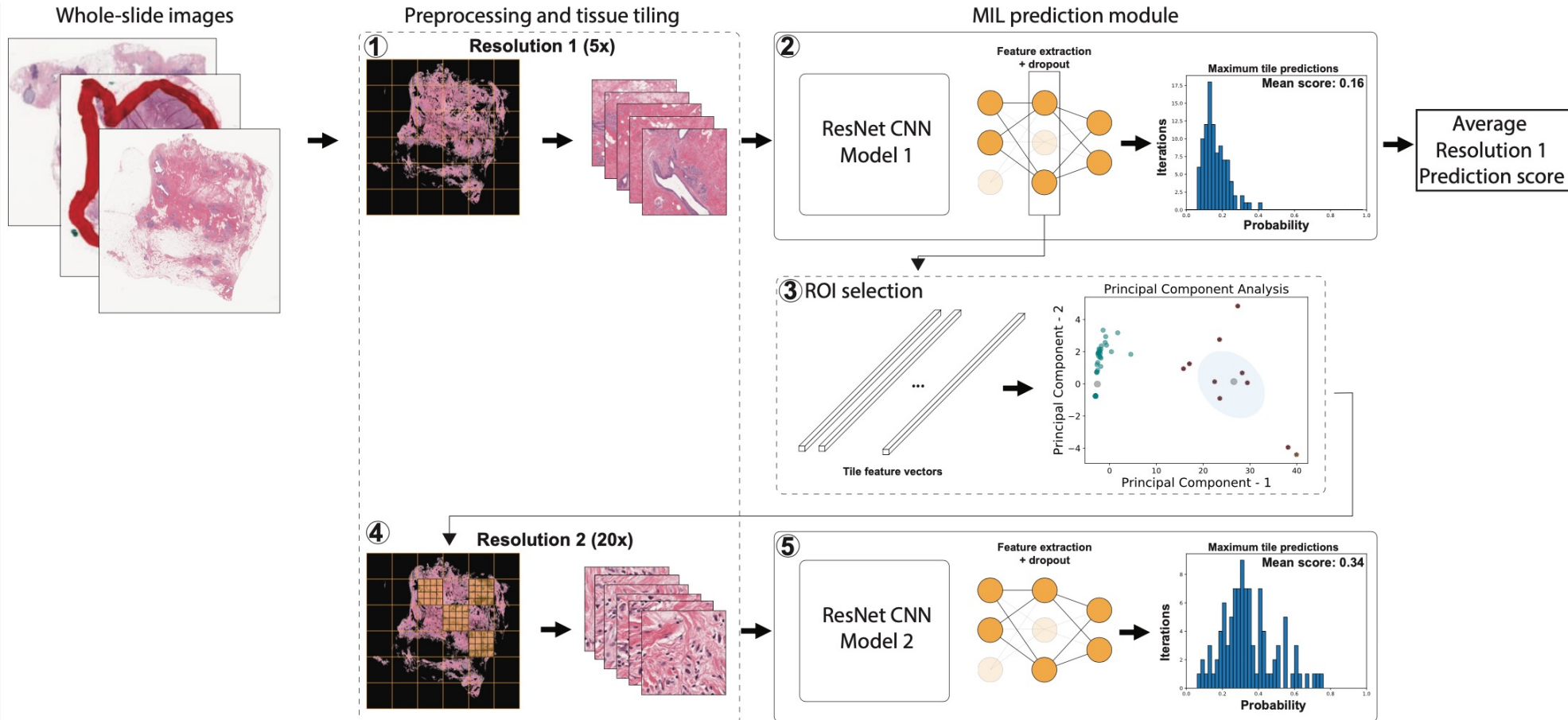
# Convolutional neural network to detect biomarkers from histopathological slides



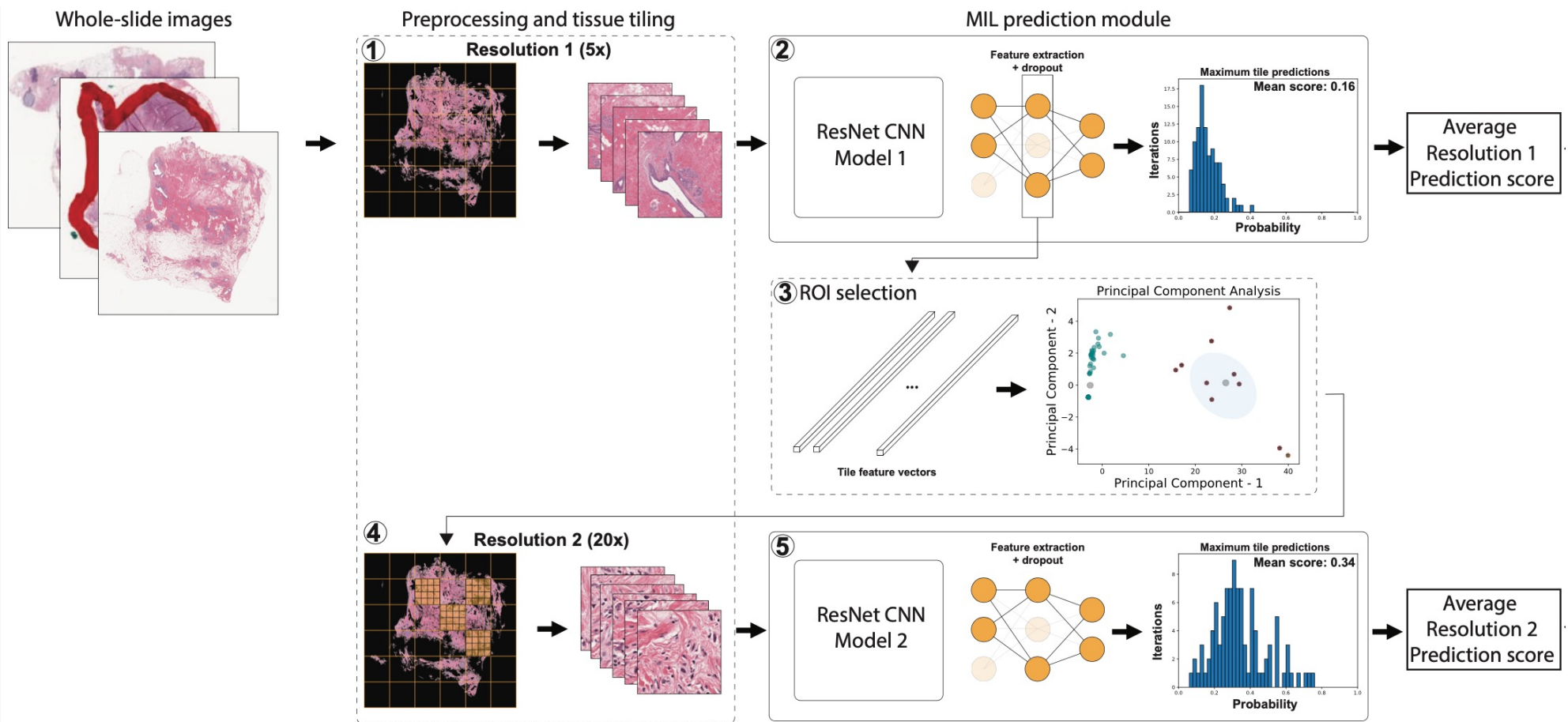
# Convolutional neural network to detect biomarkers from histopathological slides



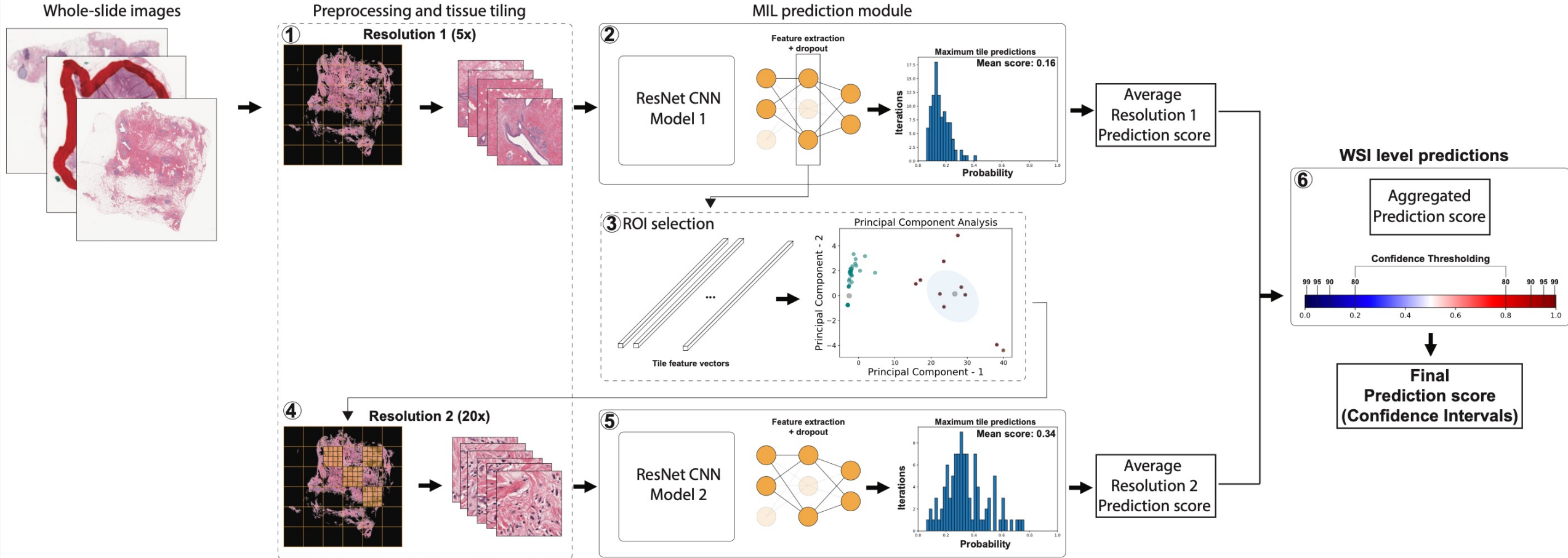
# Convolutional neural network to detect biomarkers from histopathological slides



# Convolutional neural network to detect biomarkers from histopathological slides

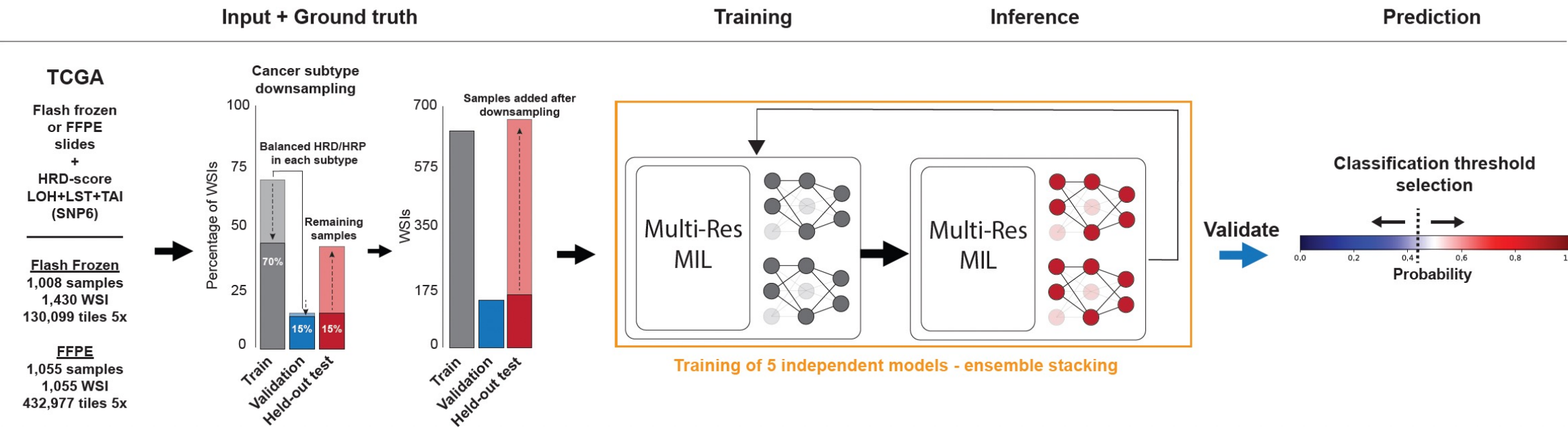


# Convolutional neural network to detect biomarkers from histopathological slides

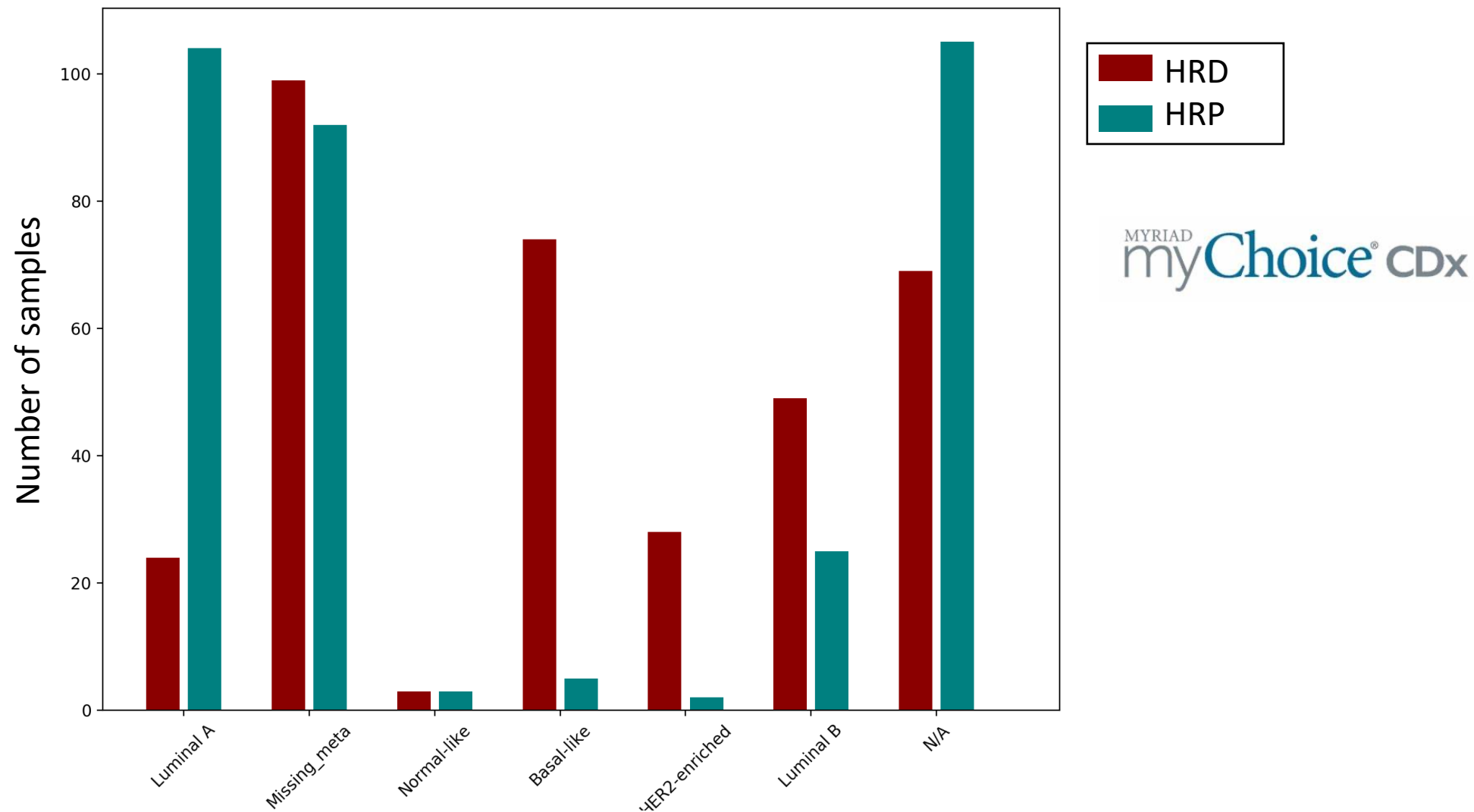




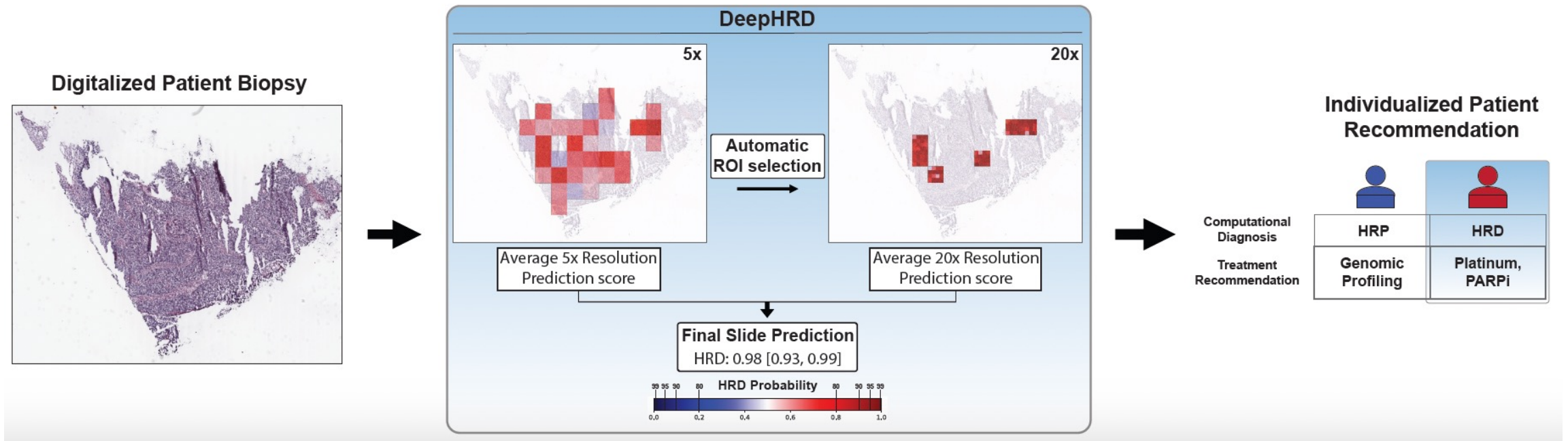
# Training a **DeepHRD** prediction model for detecting HRD from digital H&E slides



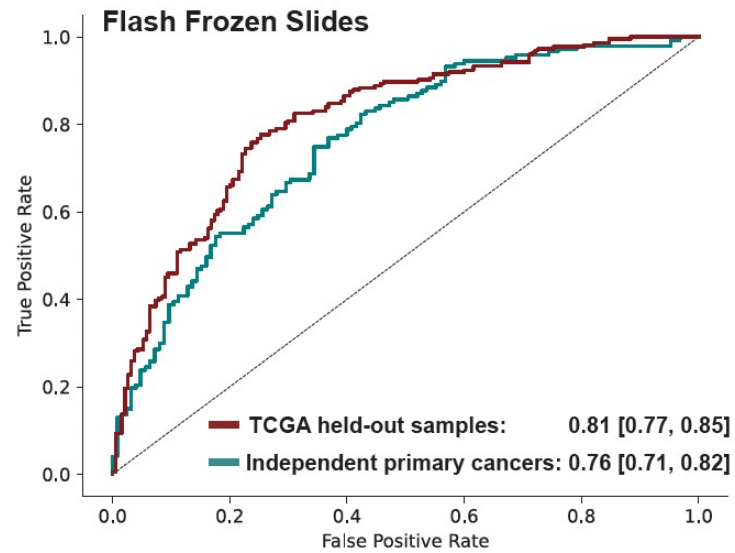




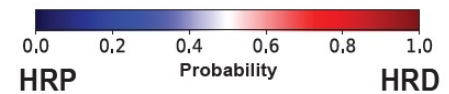
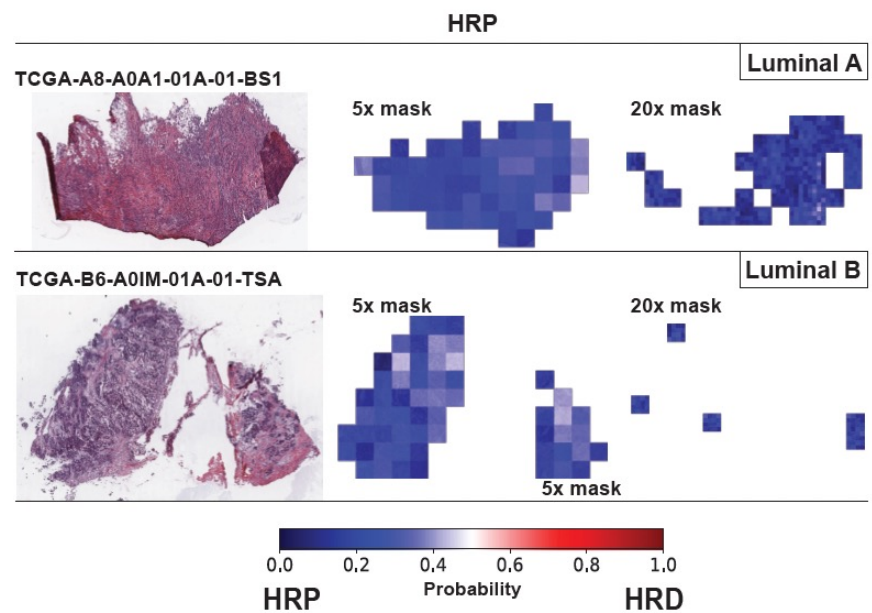
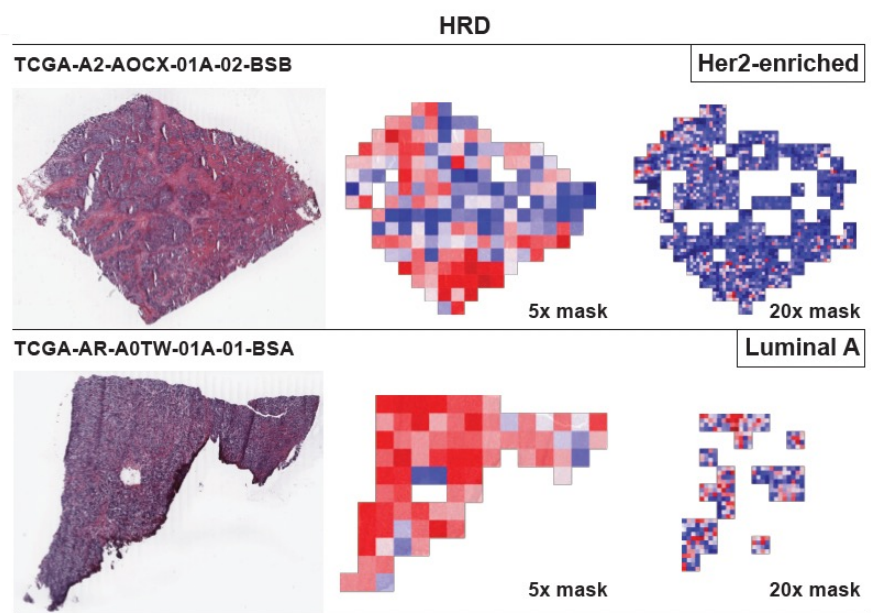
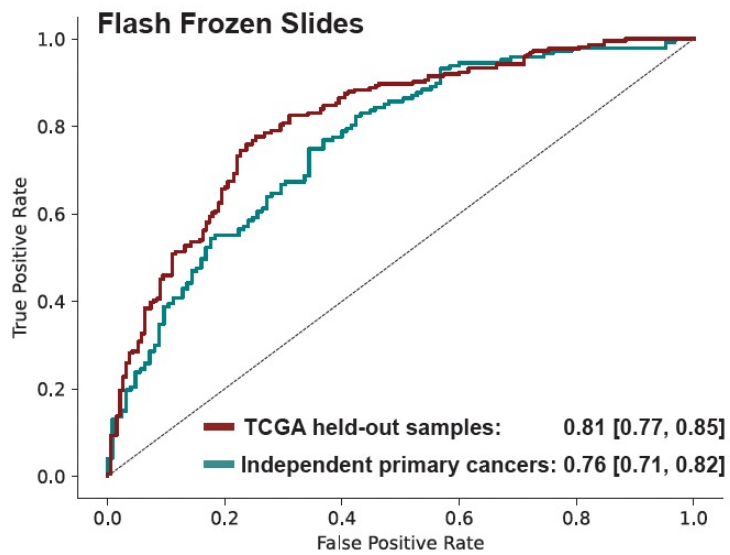
# Applying a trained DeepHRD model for predicting HRD from digital H&E slides



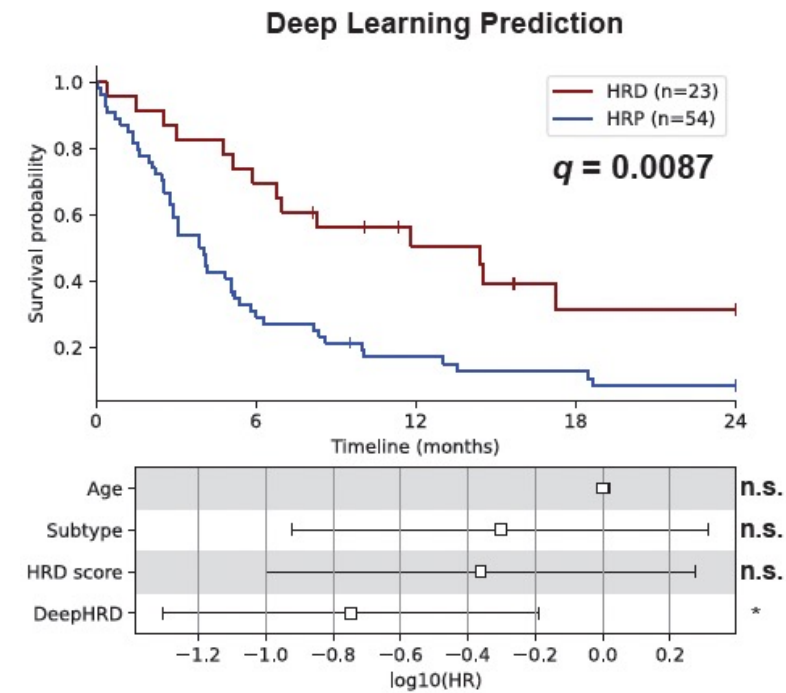
# Comparing *DeepHRD* detection power to the score of a molecular test in *primary breast cancer*



# Comparing *DeepHRD* detection power to the score of a molecular test in *primary breast cancer*



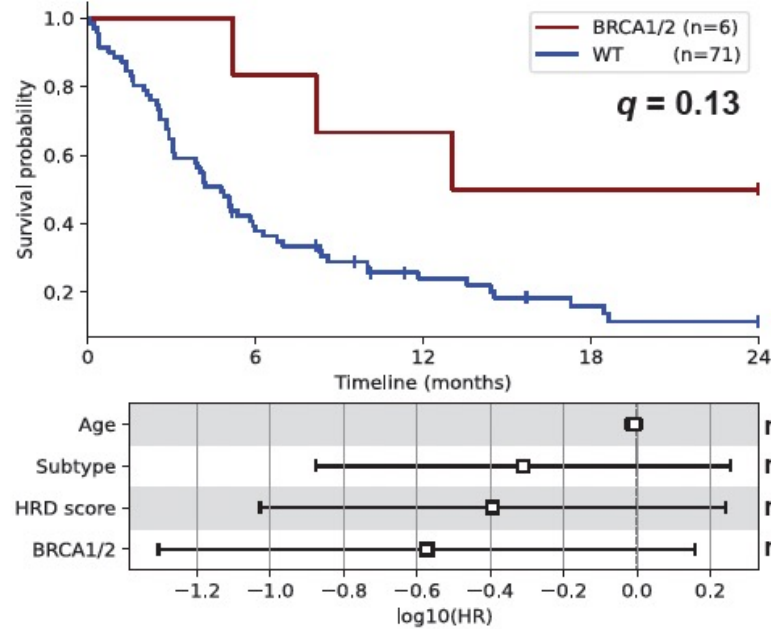
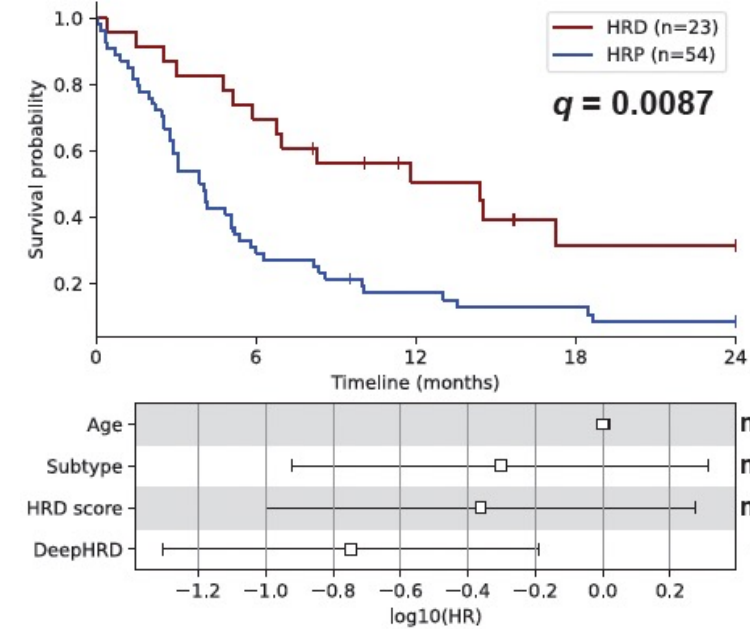
### Platinum-treated Metastatic Breast Cancers



Platinum-treated Metastatic Breast Cancers

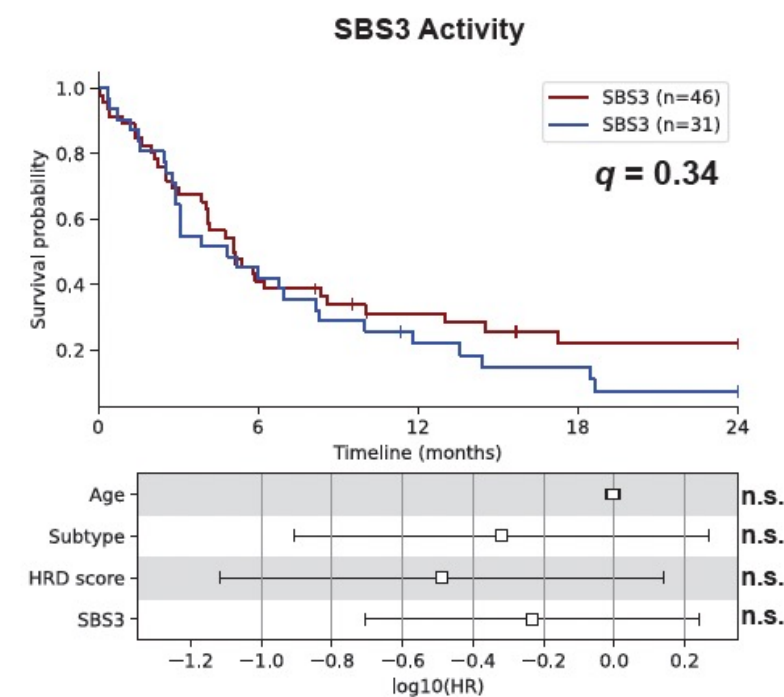
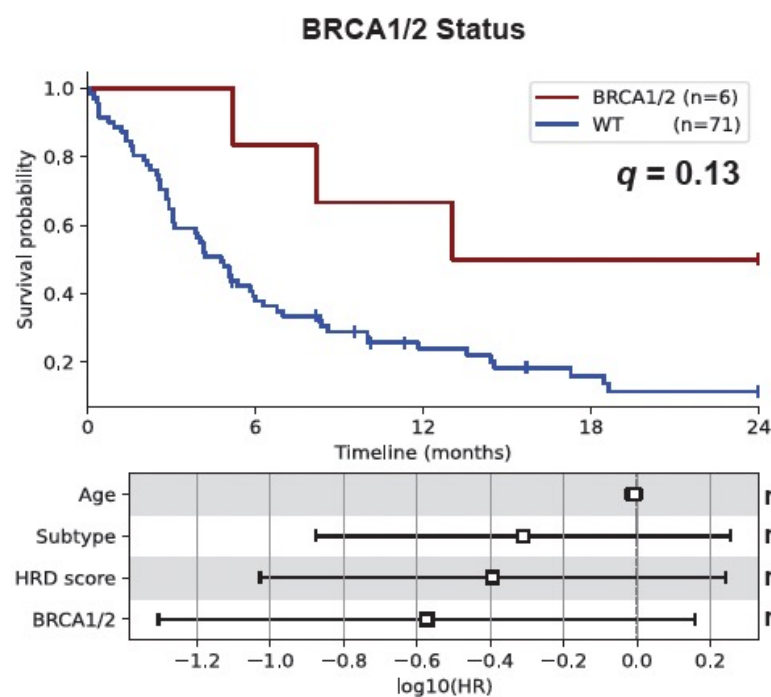
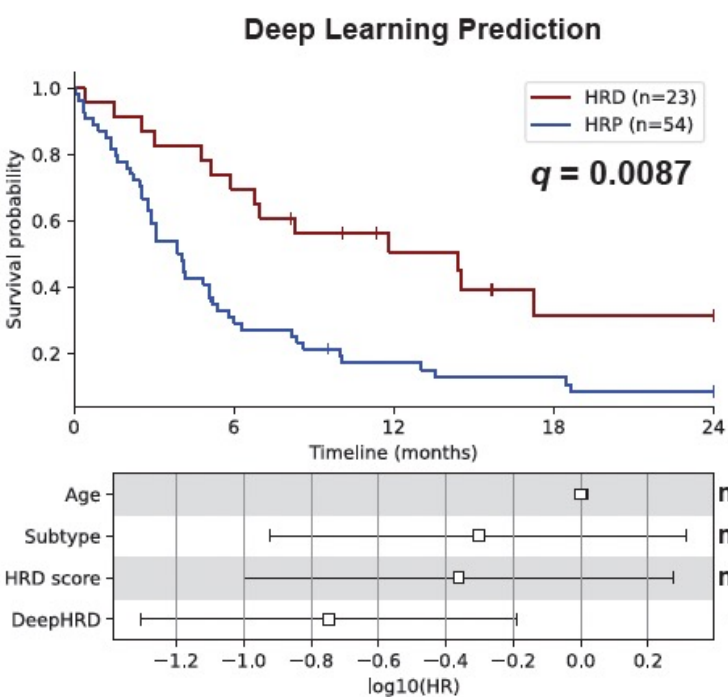
Deep Learning Prediction

BRCA1/2 Status

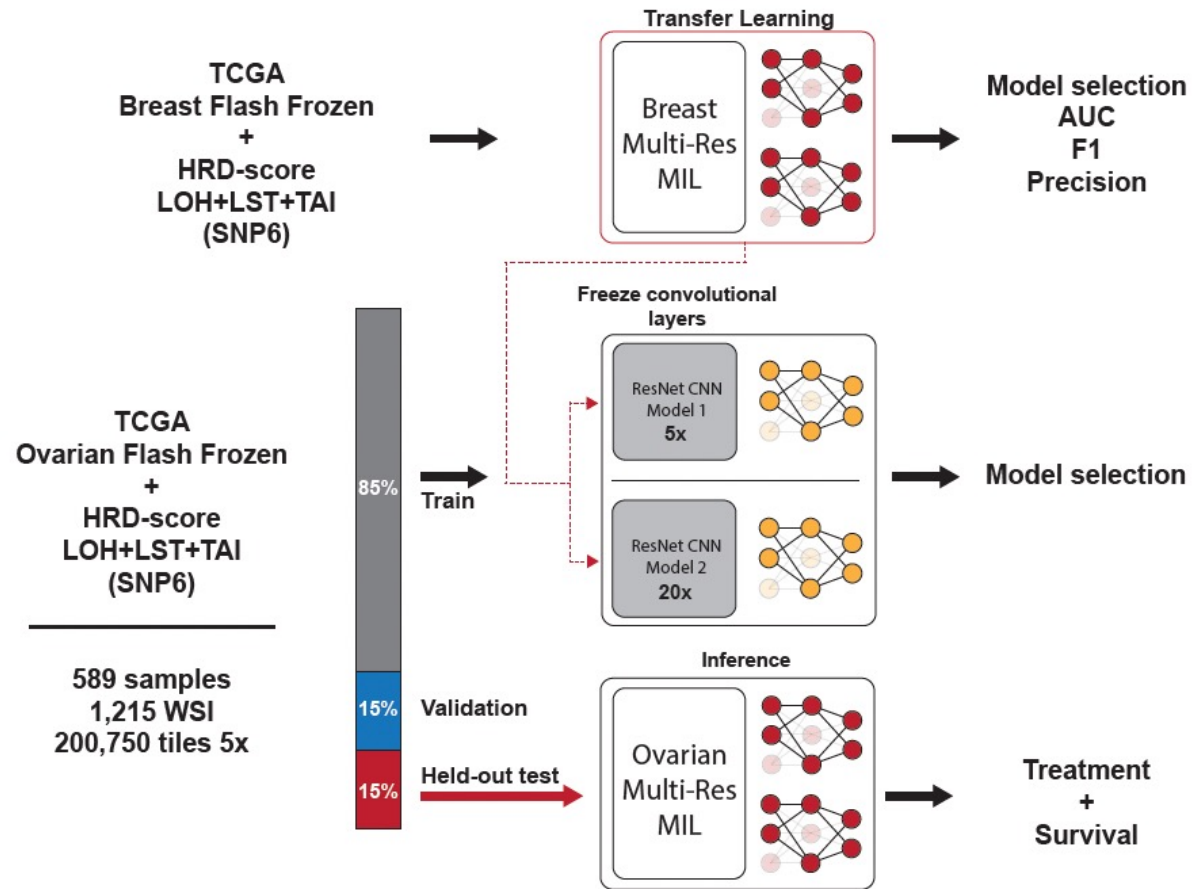




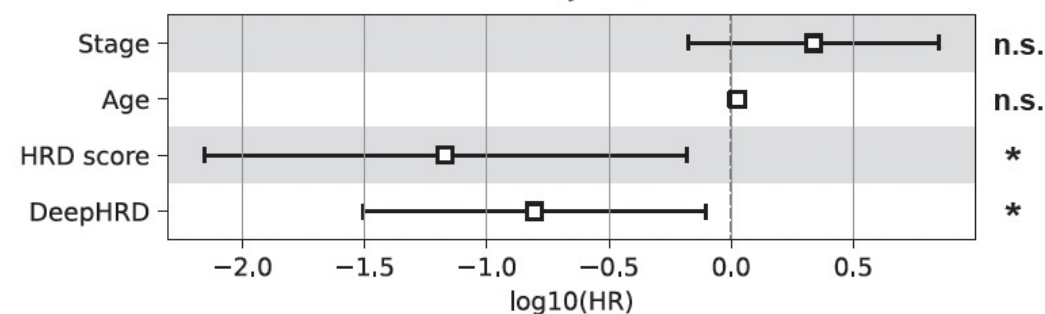
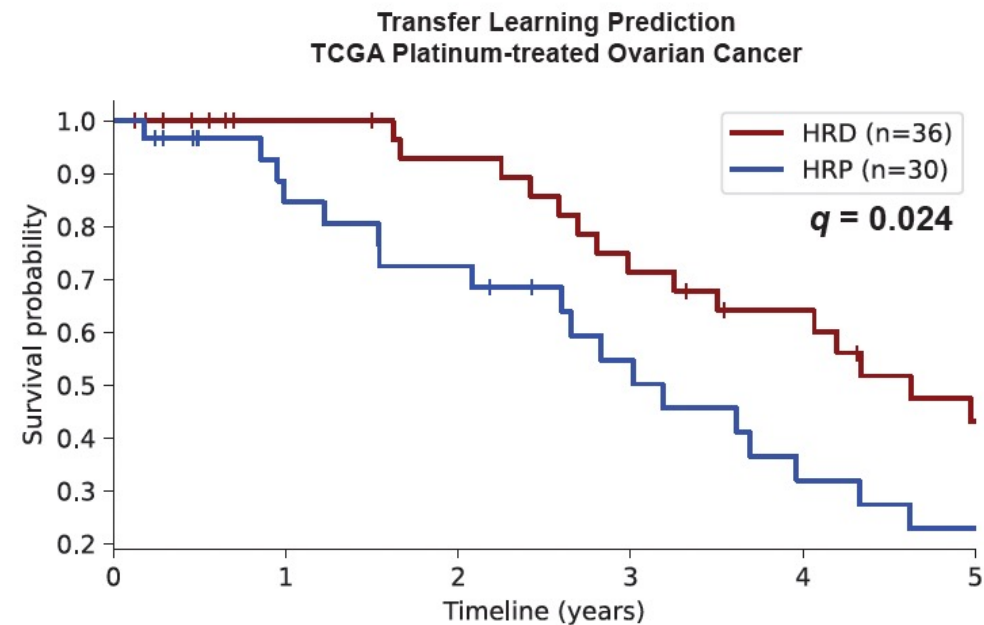
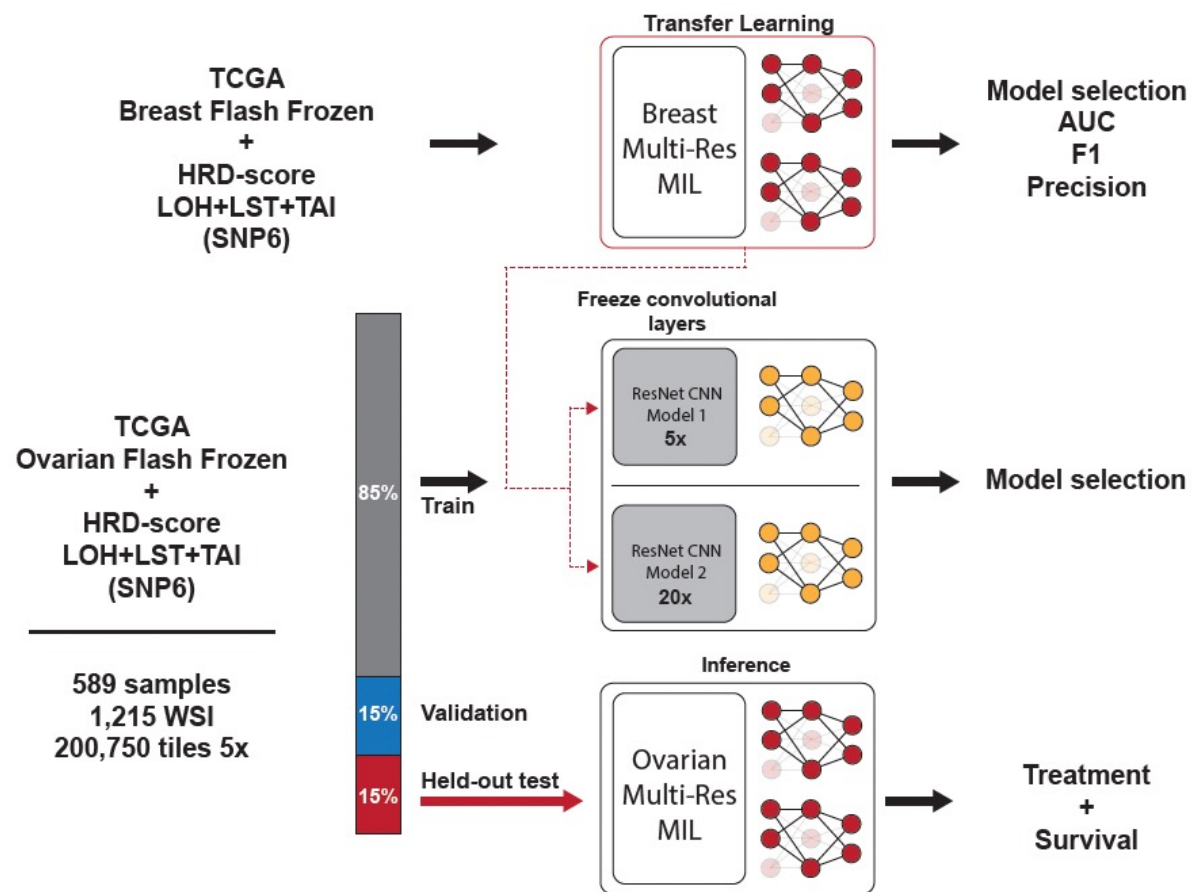
### Platinum-treated Metastatic Breast Cancers



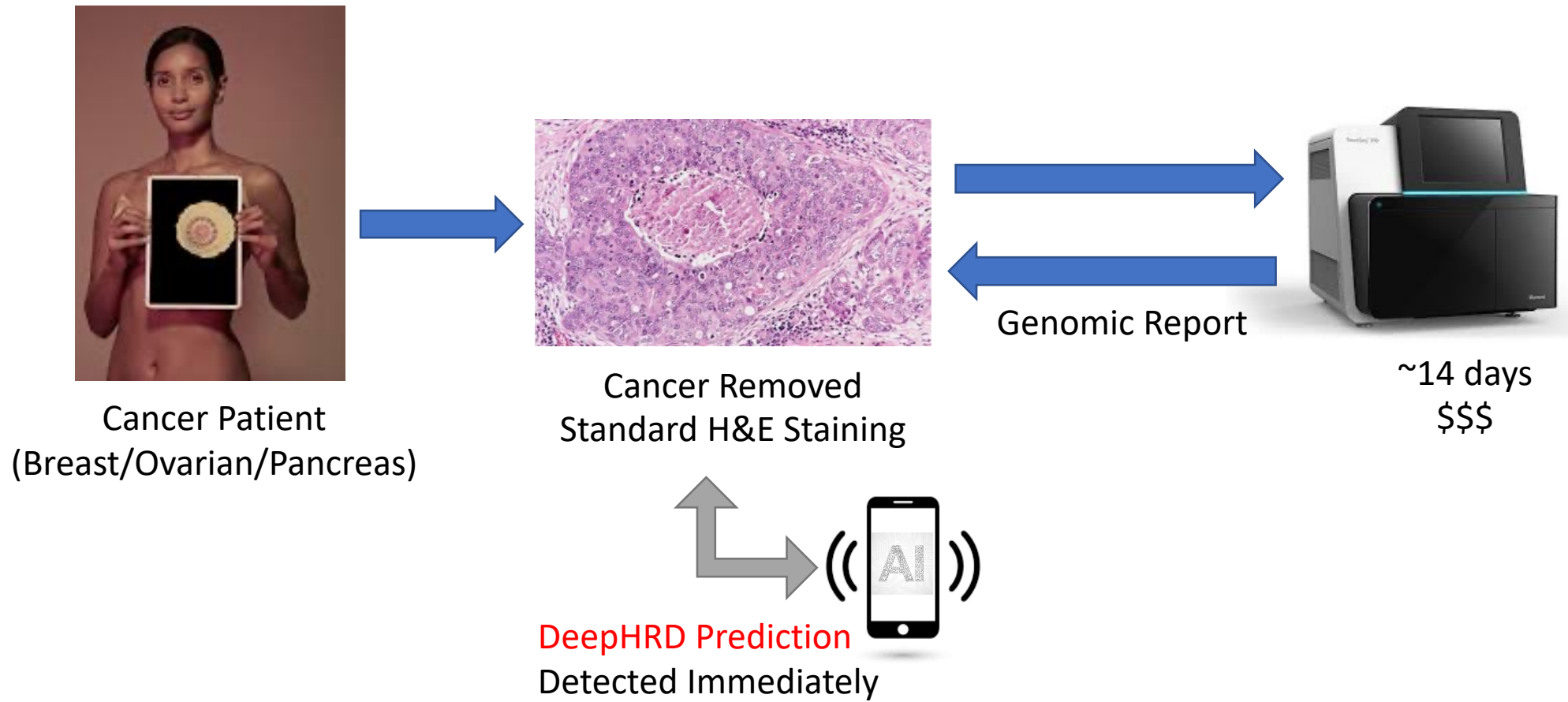
# DeepHRD *transfer learning* predicts *platinum response* in ovarian cancer



# DeepHRD transfer learning predicts platinum response in ovarian cancer



# Looking towards the future for detecting HRD score and other predictive biomarkers



# Summary

# Summary

## Stories of The Past

- Mutational signatures as a machine learning approach that allows detecting the *unusual patterns of somatic mutations*.
- Utilizing mutational signatures for developing cancer prevention strategies.
- Utilizing mutational signatures for understanding failed DNA repair and targeted cancer treatment.

## Anecdotes of The Present

- Utilizing clustered mutations for understanding cancer development and evolution.
- The repertoire of copy-number signatures in human cancer.
- A novel machine learning approach for detecting homologous recombination deficiency.

## Dreams of The Future

- Beyond genomics: Utilizing AI for addressing inequalities of cancer diagnosis

HRD





## Alexandrov Lab Members



NATIONAL  
CANCER  
INSTITUTE



the David & Lucile  
Packard  
FOUNDATION



CANCER  
RESEARCH  
UK

GRAND  
CHALLENGE  
MUTOGRAPHS

## Sanger Institute

Mike Stratton  
Peter Campbell  
Sam Behjati  
Inigo Martincorena  
Mia Petljak

## University of California, San Diego

Scott Lippman  
Jill Mesirov  
Michael Karin  
J. Silvio Gutkind  
David Brenner  
Catriona Jamieson

## Collaborators

Steve Rozen  
Reuben Harris  
Jiri Zavadil  
David Wedge  
Elli Papaemmanuil  
Jaegil Kim  
Stefan Pfister  
Sean Grimmond  
Peter Lichter  
Jessica Zucman-Rossi  
Gad Getz  
Tatsuhiko Shibata  
Reiner Siebert  
Margherita Bignami  
Graham Mann  
Maria Teresa Landi  
David Phillips  
Roger Woodgate  
Young Seok Ju  
Moritz Gerstung  
Faiyaz Notta  
Jan Korbel  
Jinghui Zhang  
Elias Campo  
Xose Puente  
Matthew Meyerson  
Lincoln Stein  
Niccolò Bolli  
Steven Gallinger  
Rebecca Fitzgerald





# Questions?