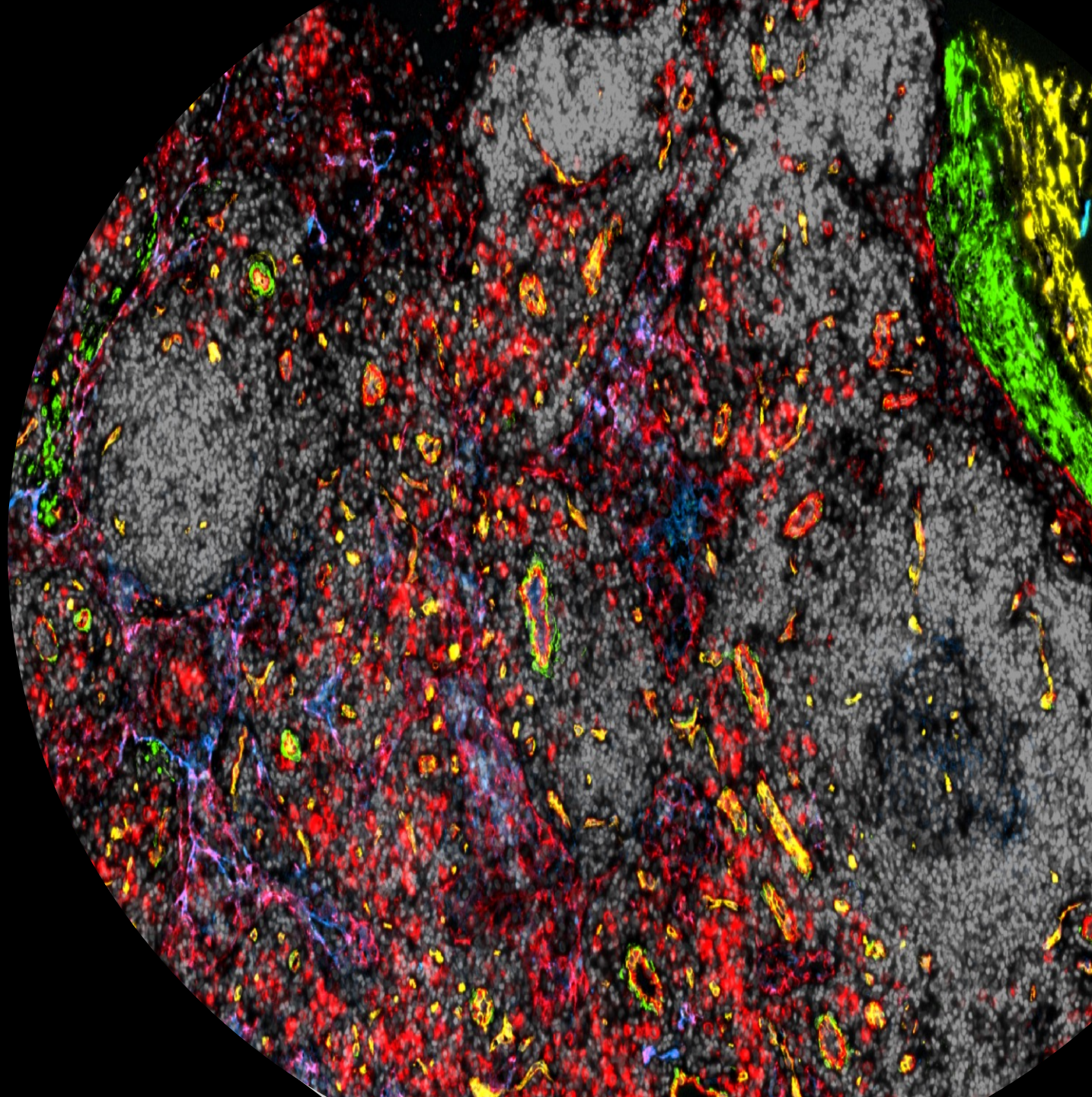
Cancer Biology I

Part-II

Fall semester 2024
Week 8



AGENDA

Nov 5th: Cancer genomics- mutations

Nov 11th: Cancer genomics-copy number alteration, heterogeneity, evolution
(recording)

Nov 18th: Cancer Epigenetics- chromatin 3D structure, cell plasticity

Nov 25th: – Major signaling pathways leading to cancer

Dec 2th: Cancer Therapies – chemo and targeted therapies

Dec 9th: Introduction to immunotherapies –

Dec 16th: Exam
(if it conflicts with another exam it is possible to do the exam on Dec 18th)

Dec 18th: discussion of exam questions and career development discussion towards a PhD or not…

Exercise

- Two TAs will help in the discussion of the paper

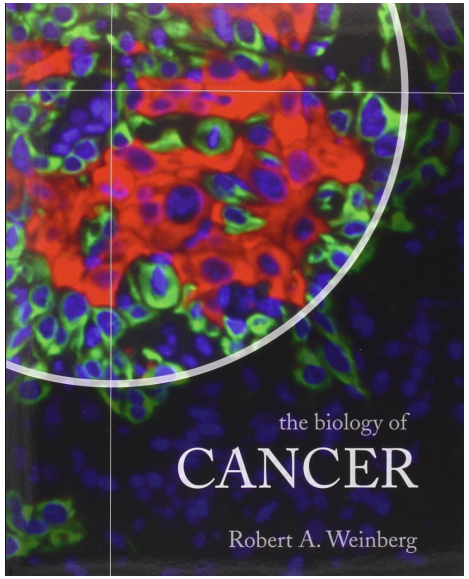
Marie Denise Rumpler
Irmak Kaysudu

- Wednesday, Nov 20th and Wednesday, Dec 11:
- I will post two short-mock exam sessions with questions from the previous years

Exam Information

- Exam questions will be in English, but you're allowed to answer in French (*if you do, please use capital letters*)
- *Day to be confirmed:* Dec 16 or 18th , 2 - 4 PM (*room will be communicated*)
- ~10 open questions
  - 10 points for each question
  - 2 questions will be on the papers that you will read during the exercises
- The exam will count for 50% of your final grade

If you want to read on specific topics:

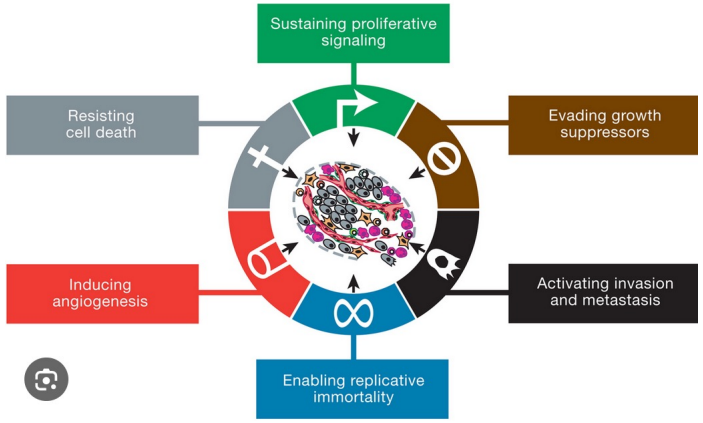


Some information if this book

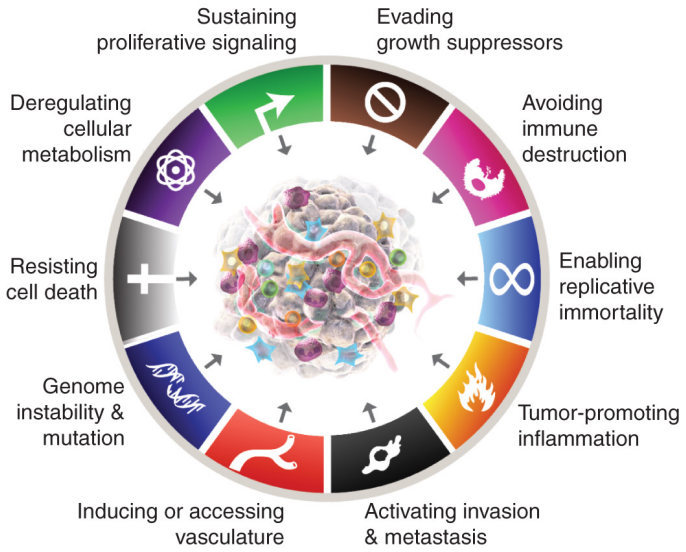
Send an email:

I can send you recent Reviews written by prominent cancer scientists in the fields

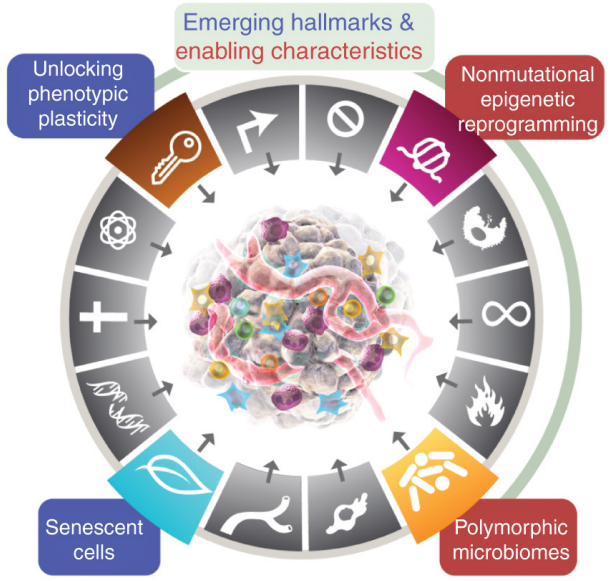
Cancer is a complex disease that involves changes in many different biological aspects



Hanahan and Weinberg, Cell 2000



Hanahan and Weinberg, Cell 2011



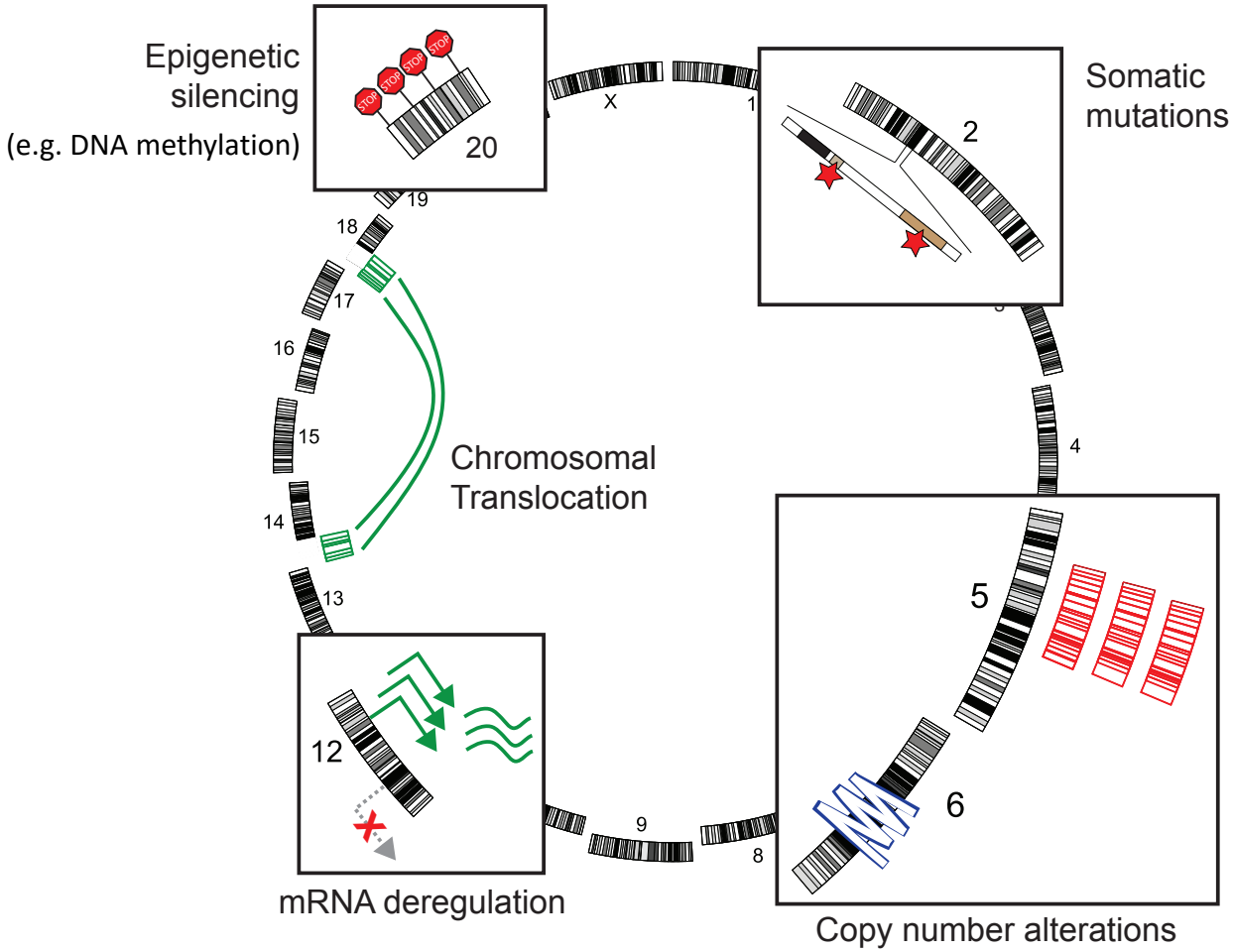
Hanahan D. Cancer Discovery 2022

1. Cancer is a disease of the genome

Cancer genomics: studies the acquisition of alterations in the genome (i.e. in our DNA) that can cause cancer development

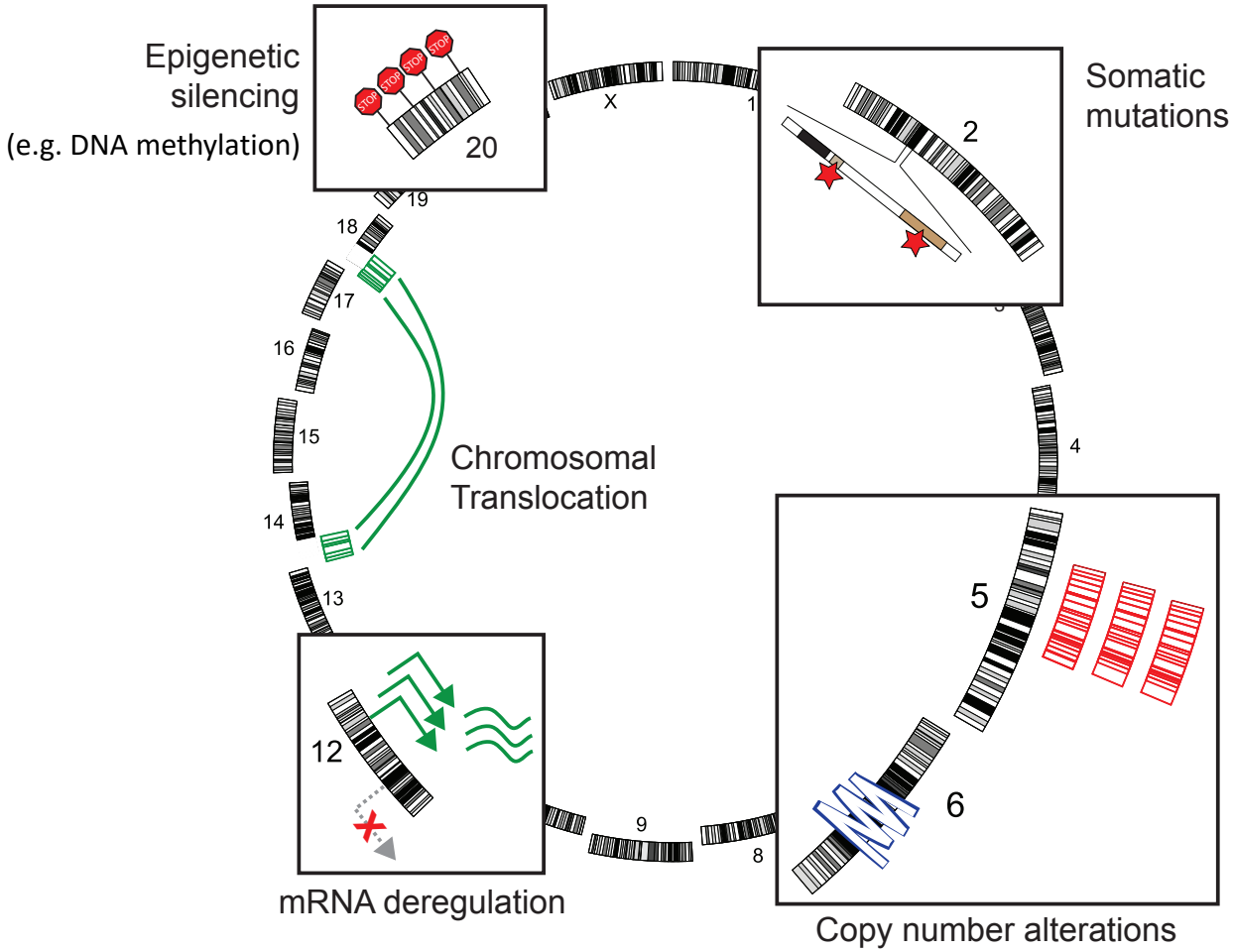
What are they?

Cancer Genomic Alterations



What are they?

Cancer Genomic Alterations



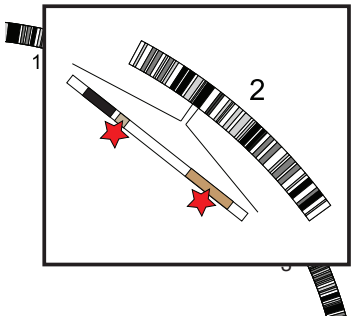
You like my dogs

↓

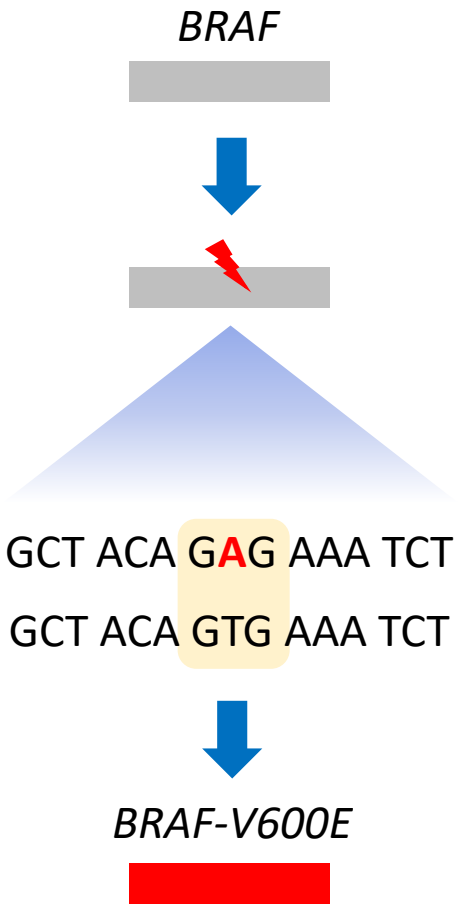
You like my dots

Somatic Mutations

- Single nucleotide changes



Somatic mutations



Somatic Mutations

- Single nucleotide changes
  - Missense Mutations
    - GAG = Valine
    - GTG = Glutamate
  - Silent Mutations
    - TCT = Serine
    - TCC = Serine
  - Nonsense Mutations
    - TAC = Tyrosine
    - TAG = **Stop Codon!**

	T			C			A			G		
T	TTT	Phe	F	TCT	Ser	S	TAT	Tyr	Y	TGT	Cys	C
	TTC	Phe	F	TCC	Ser	S	TAC	Tyr	Y	TGC	Cys	C
	TTA	Leu	L	TCA	Ser	S	TAA	stop	*	TGA	stop	*
	TTG	Leu	L	TCG	Ser	S	TAG	stop	*	TGG	Trp	W
C	CTT	Leu	L	CCT	Pro	P	CAT	His	H	CGT	Arg	R
	CTC	Leu	L	CCC	Pro	P	CAC	His	H	CGC	Arg	R
	CTA	Leu	L	CCA	Pro	P	CAA	Gln	Q	CGA	Arg	R
	CTG	Leu	L	CCG	Pro	P	CAG	Gln	Q	CGG	Arg	R
A	ATT	Ile	I	ACT	Thr	T	AAT	Asn	N	AGT	Ser	S
	ATC	Ile	I	ACC	Thr	T	AAC	Asn	N	AGC	Ser	S
	ATA	Ile	I	ACA	Thr	T	AAA	Lys	K	AGA	Arg	R
	ATG	Met	M	ACG	Thr	T	AAG	Lys	K	AGG	Arg	R
G	GTT	Val	V	GCT	Ala	A	GAT	Asp	D	GGT	Gly	G
	GTC	Val	V	GCC	Ala	A	GAC	Asp	D	GGC	Gly	G
	GTA	Val	V	GCA	Ala	A	GAA	Glu	E	GGA	Gly	G
	GTG	Val	V	GCG	Ala	A	GAG	Glu	E	GGG	Gly	G

Somatic Mutations

- Frame-shift mutations: insertion or deletion that change the reading frame
- **Deletion:** deletion of 1 or more nucleotides

ACC AGC TGC ACT
Thr Ser Cys Thr

ACC AGC TGA CT
Thr Ser stop

- **Insertion**: Addition 1 or more extra-nucleotides

ACC AGC TGC ACT
Thr Ser Cys Thr

ACC AGC TGC CAC CT
Thr Ser Cys His

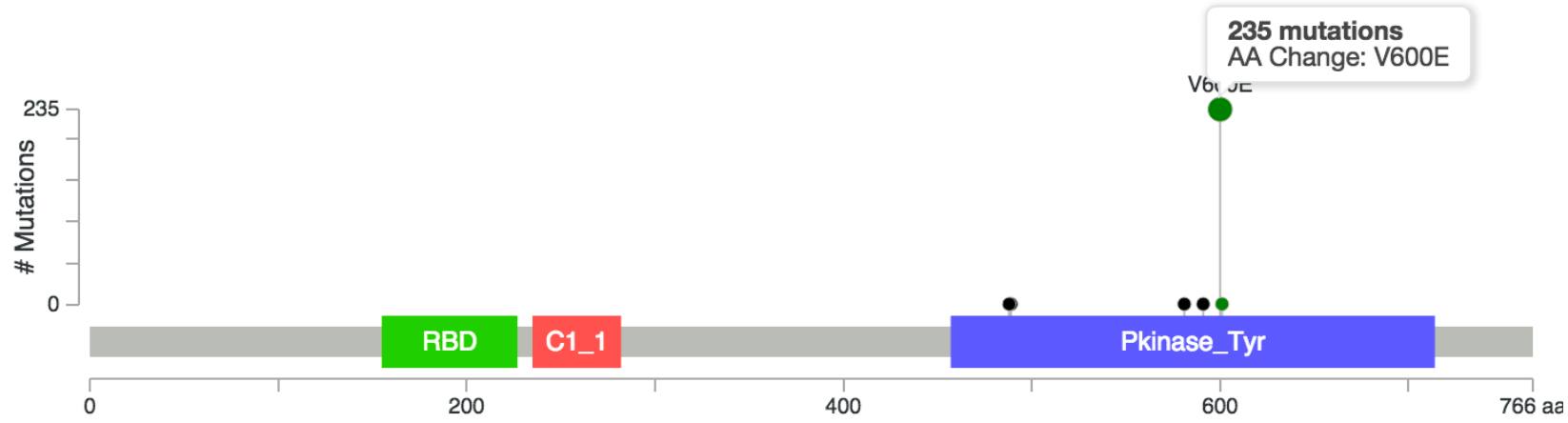
HOTSPOT mutations

(activating an oncogene)

BRAF V600E mutations in Thyroid Carcinoma (399 patients)

GTG = Valine (V)

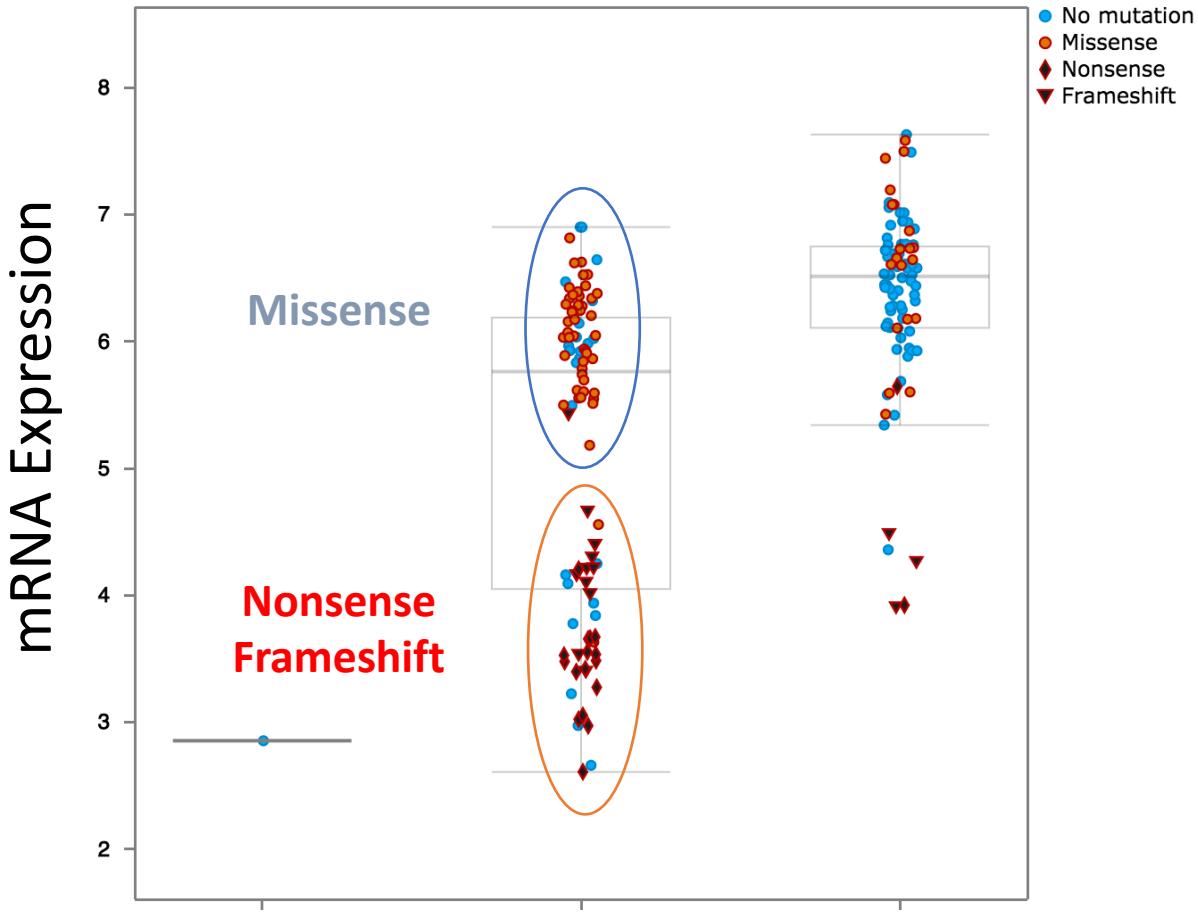
GAG = Glutamate (E)



Truncating Mutations

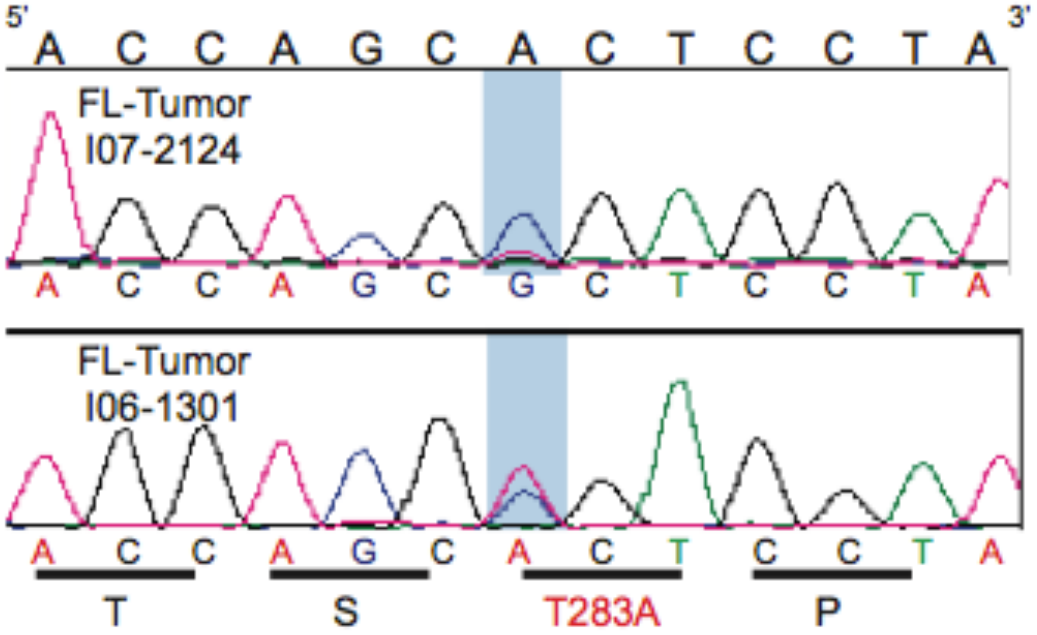
(inactivating a tumor suppressor)

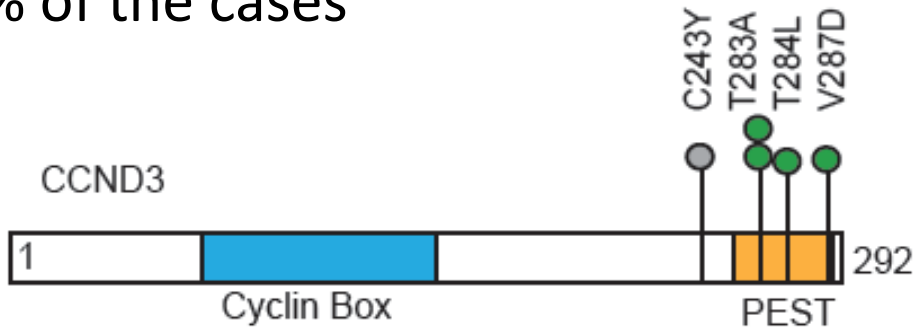
- **TP53 mutations in Colorectal cancer**

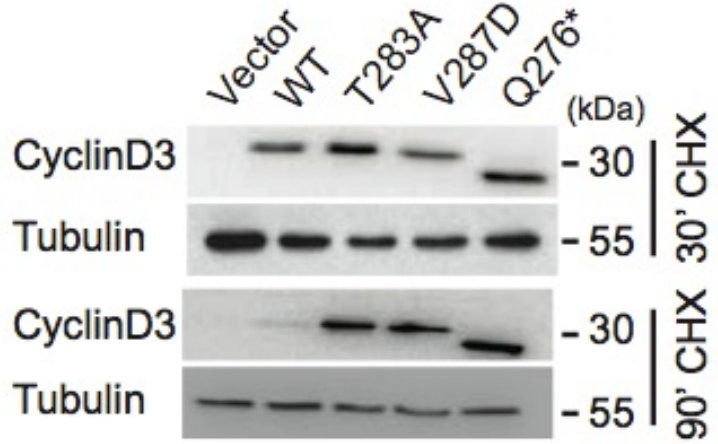


Truncating Mutations
(activating an oncogene)

In Lymphoma mutation in CyclinD3 occurs in ~10% of the cases





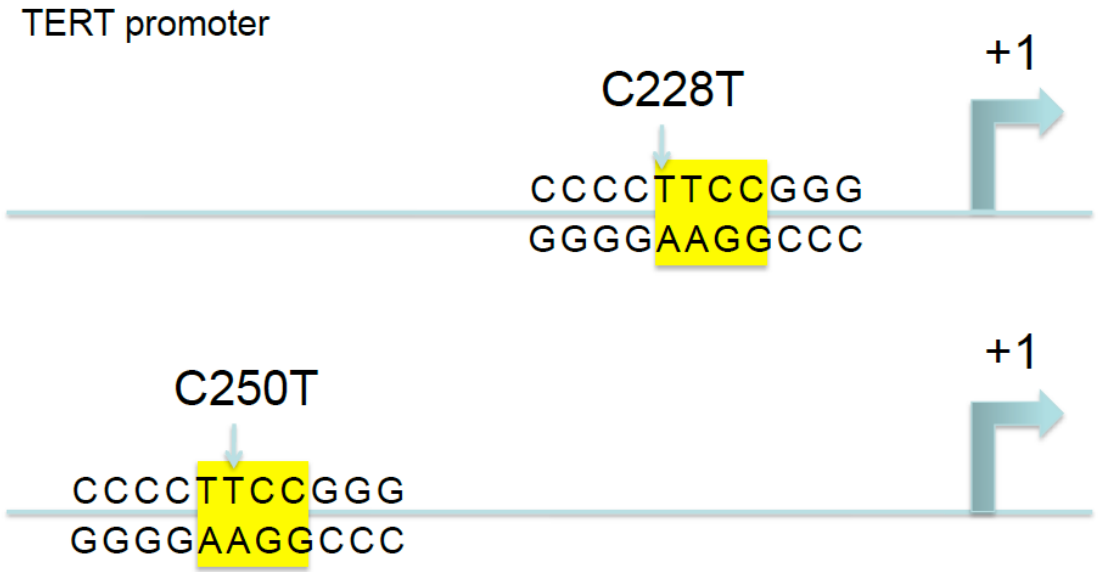


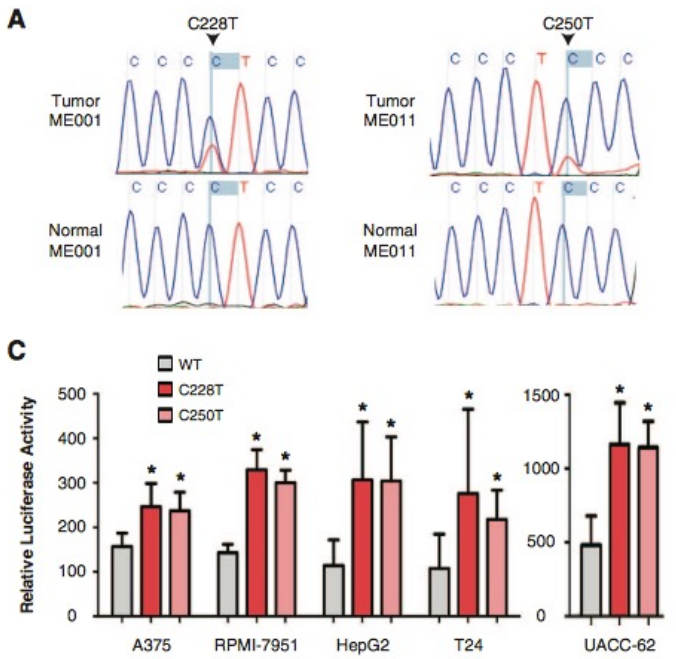
Non-coding Mutations

Highly Recurrent *TERT* Promoter Mutations in Human Melanoma

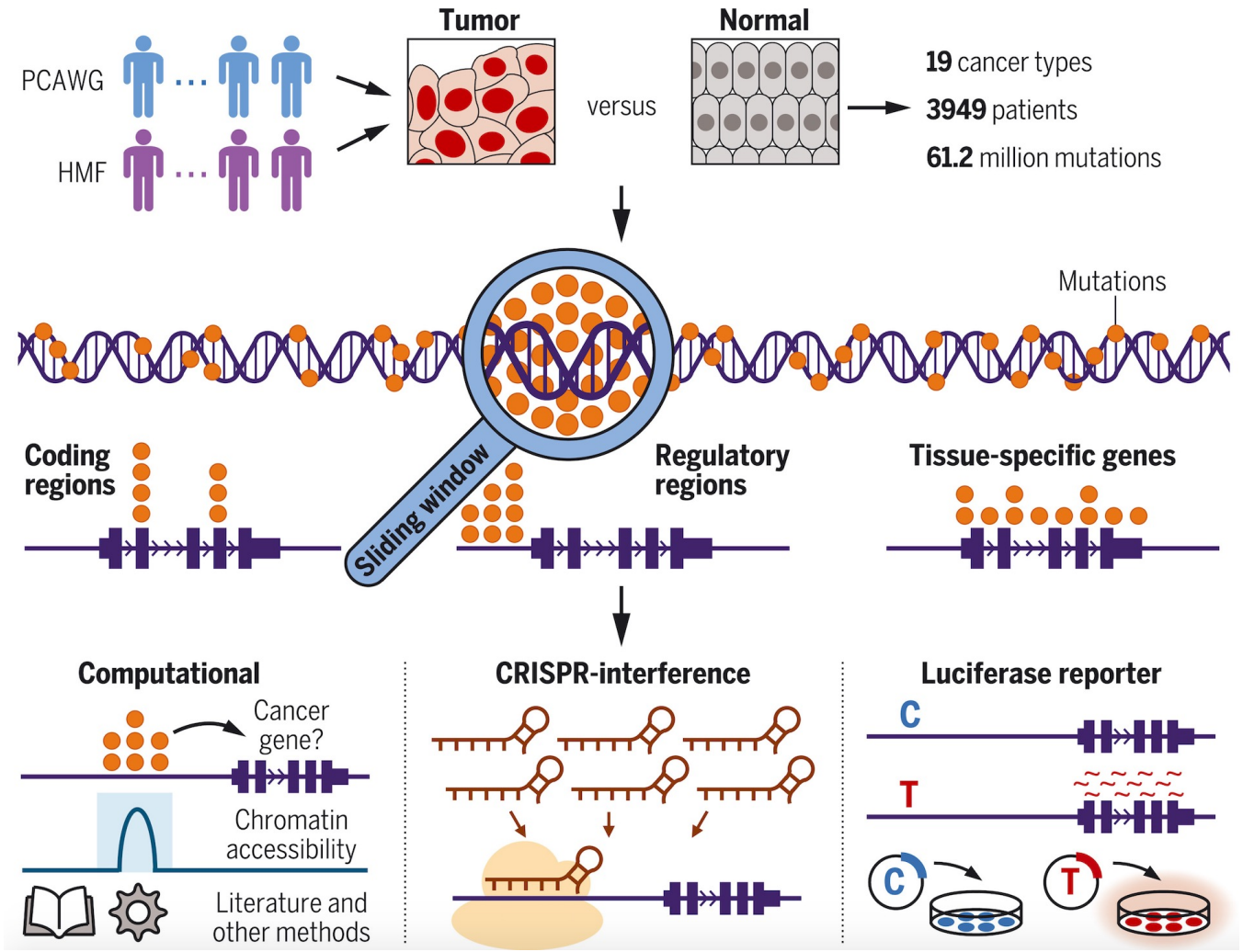
Franklin W. Huang,[1,2,3]* Eran Hodis,[1,3,4]* Mary Jue Xu,[1,3,4] Gregory V. Kryukov,[1] Lynda Chin,[5,6] Levi A. Garraway[1,2,3]†

(Science, 2013)



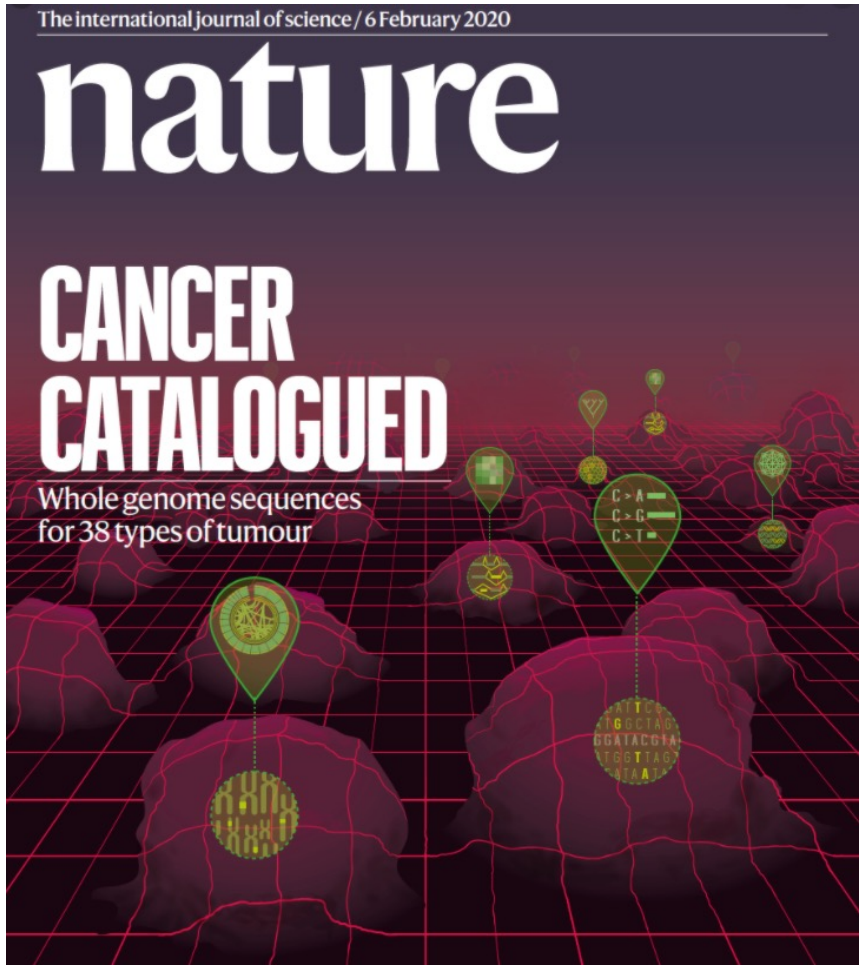


Non-coding Mutations



(Dietlein et al. Science 2022)

Pan-Cancer Analyses of Whole Genomes



Article | Open Access | Published: 05 February 2020

Analyses of non-coding somatic drivers in 2,658 cancer whole genomes

Esther Rheinbay, Morten Muhlig Nielsen, [...] PCAWG Consortium

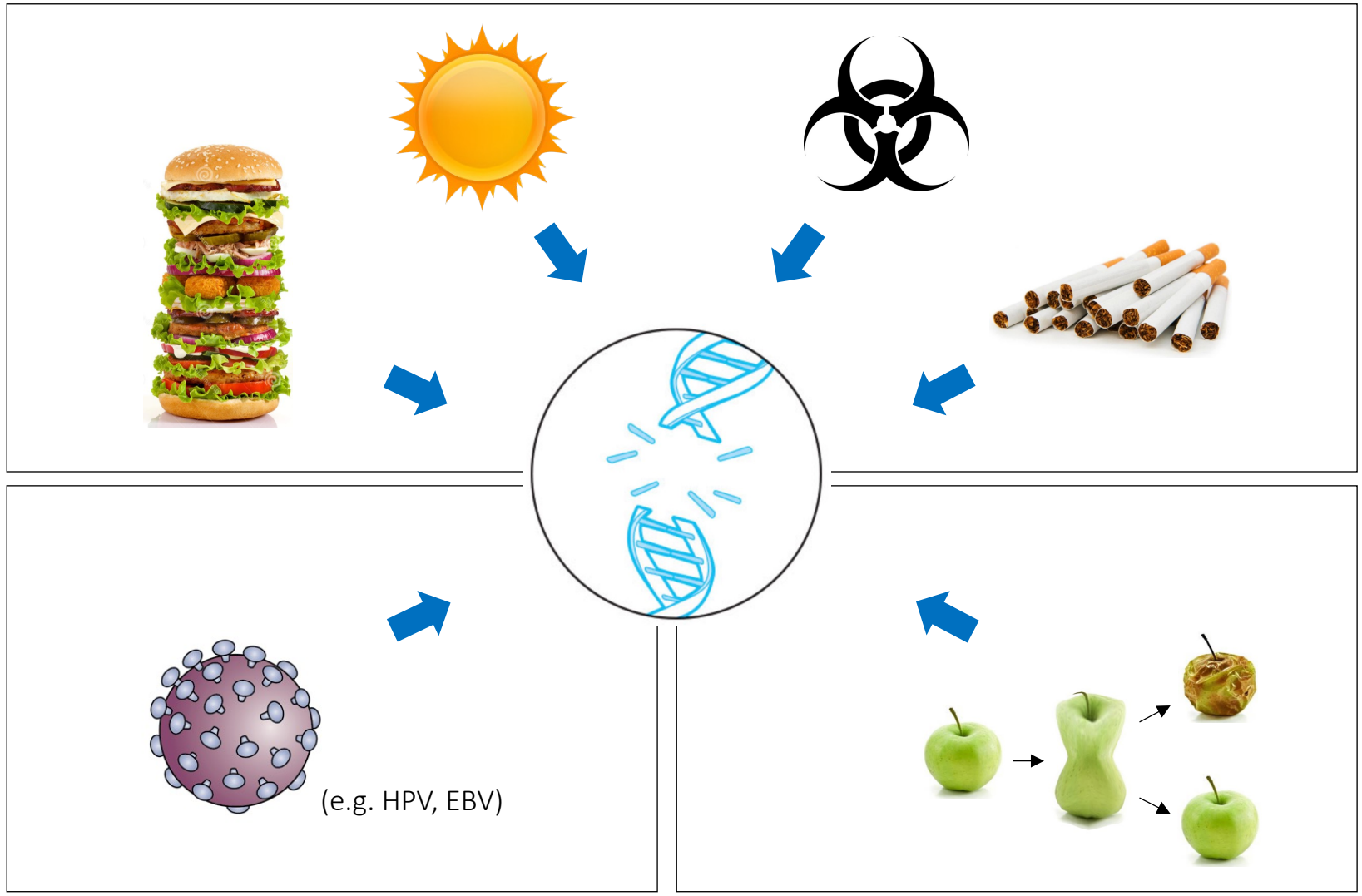
Nature **578**, 102–111(2020) | Cite this article

Paucity of non-coding drivers in cancer

Why?

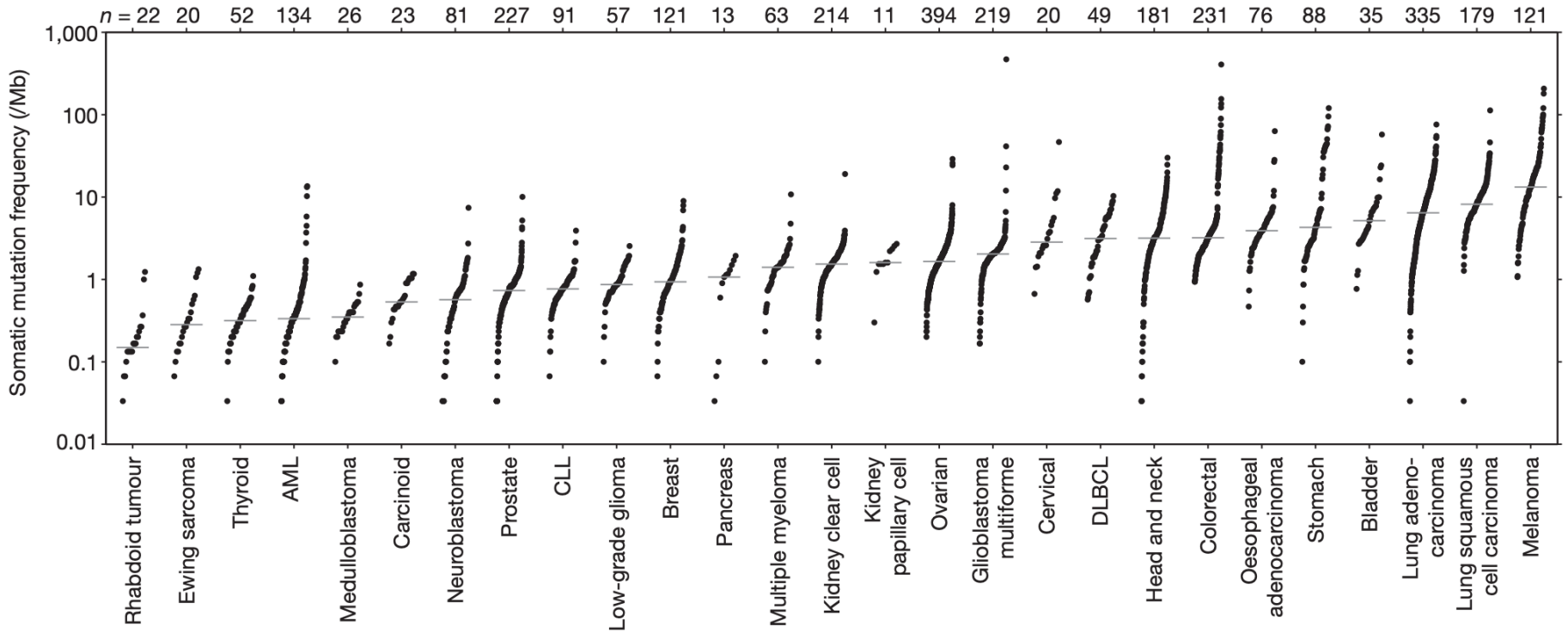
How do alteration emerge?

How do *mutations* emerge?



What do you get after sequencing 1000 tumors?

of mutations per Mb

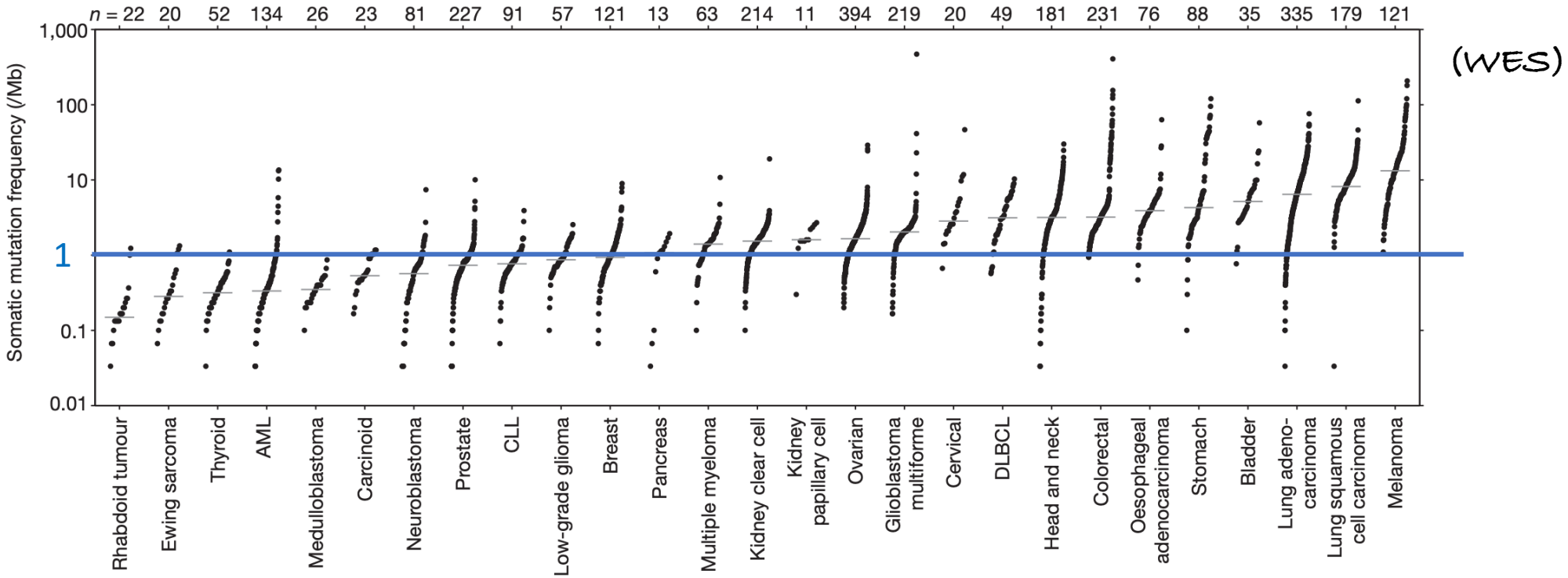


(Whole exome seq.)

(Lawrence et al, Nature 2013)

What do you get after sequencing 1000 tumors?

of mutations per Mb



(WES)

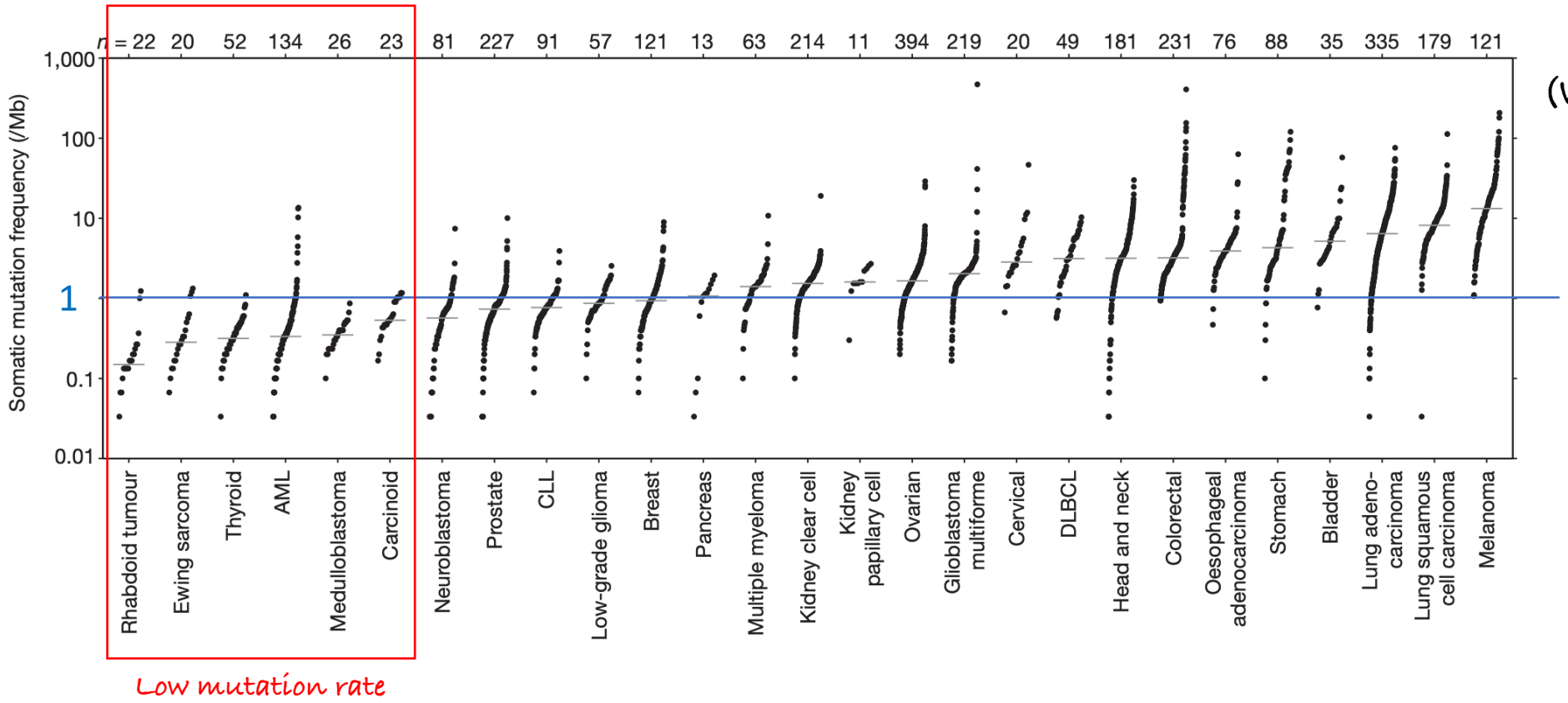
1 → 1/Mb * 3000Mb ~ 3000 mutations

2% are coding sequences → ~60 mutations in gene sequences

(Lawrence et al, Nature 2013)

What do you get after all this sequencing?

Young patients / pediatric diseases



(WES)

# of mutations per Mb

Low mutation rate

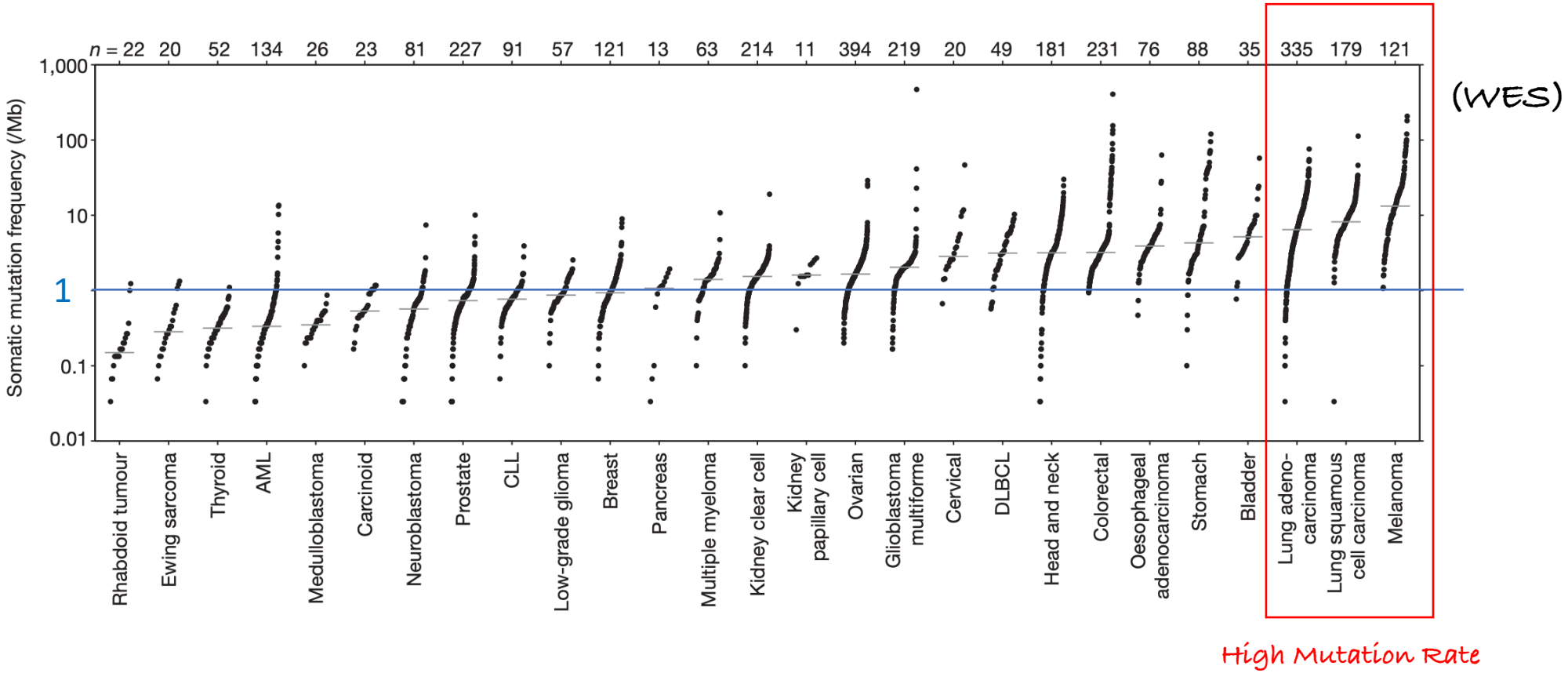
1 → 1/Mb * 3000Mb ~ 3000 mutations

2% are coding sequences → ~60 mutations in gene sequences

(Lawrence et al, Nature 2013)

What do you get after all this sequencing?

of mutations per Mb



(WES)

High Mutation Rate

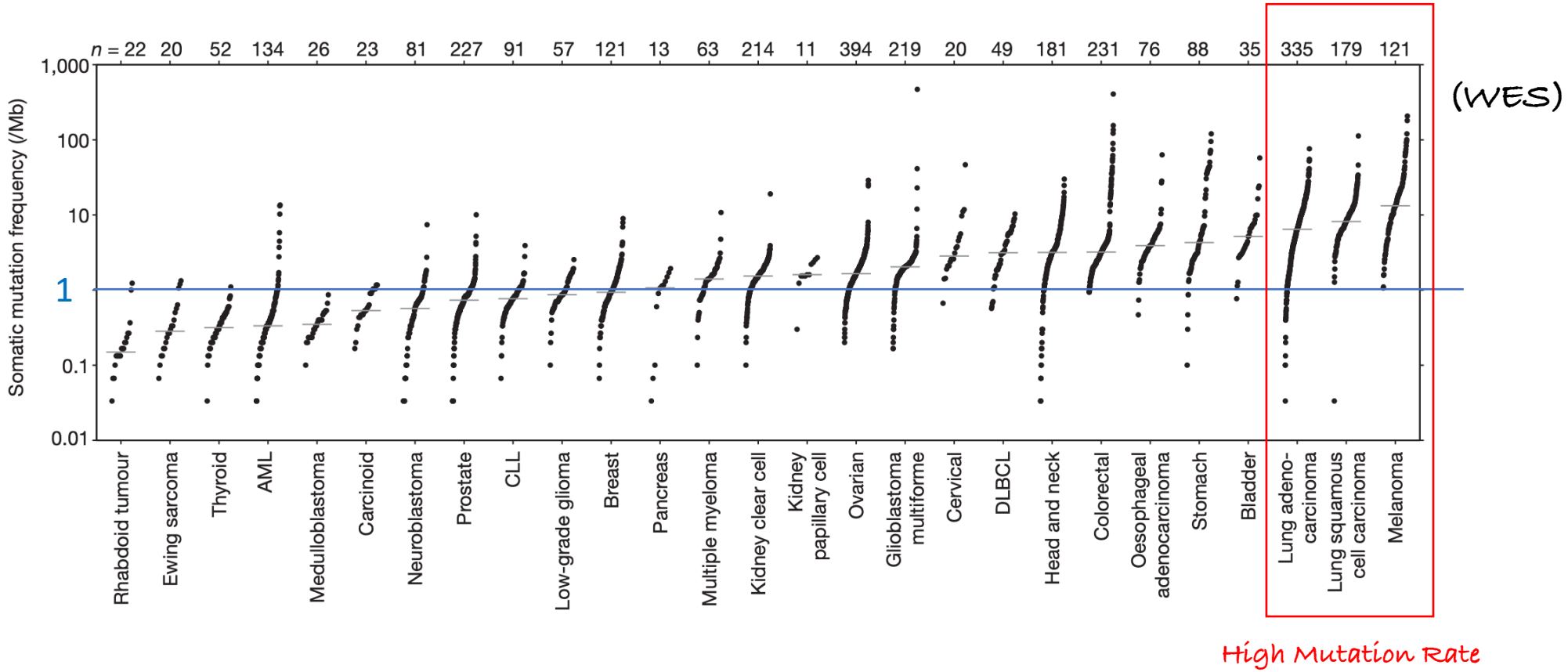
1 → 1/Mb * 3000Mb ~ 3000 mutations

2% are coding sequences → ~60 mutations in gene sequences

(Lawrence et al, Nature 2013)

What do you get after all this sequencing?

Exogenous Mutagens



(WES)

# of mutations per Mb

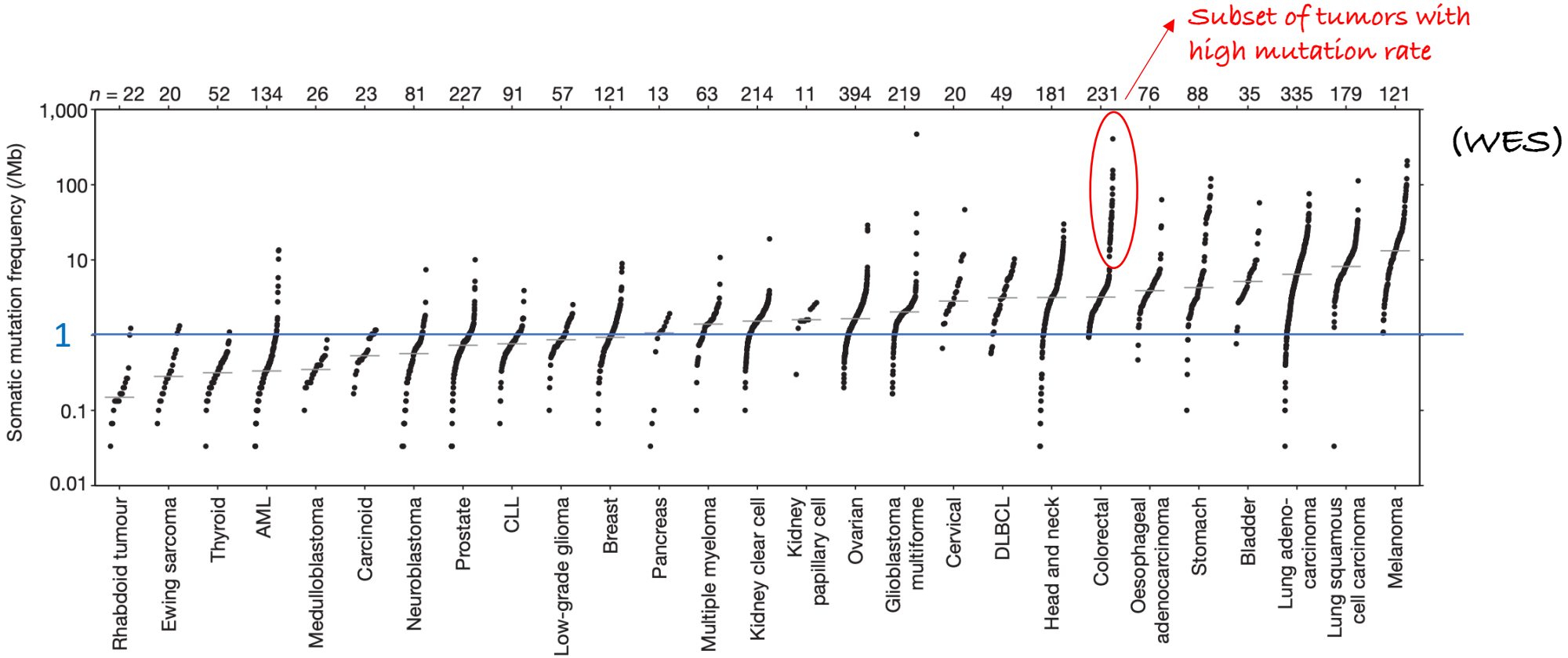
High Mutation Rate

1 → 1/Mb * 3000Mb ~ 3000 mutations

2% are coding sequences → ~60 mutations in gene sequences

(Lawrence et al, Nature 2013)

What do you get after all this sequencing?



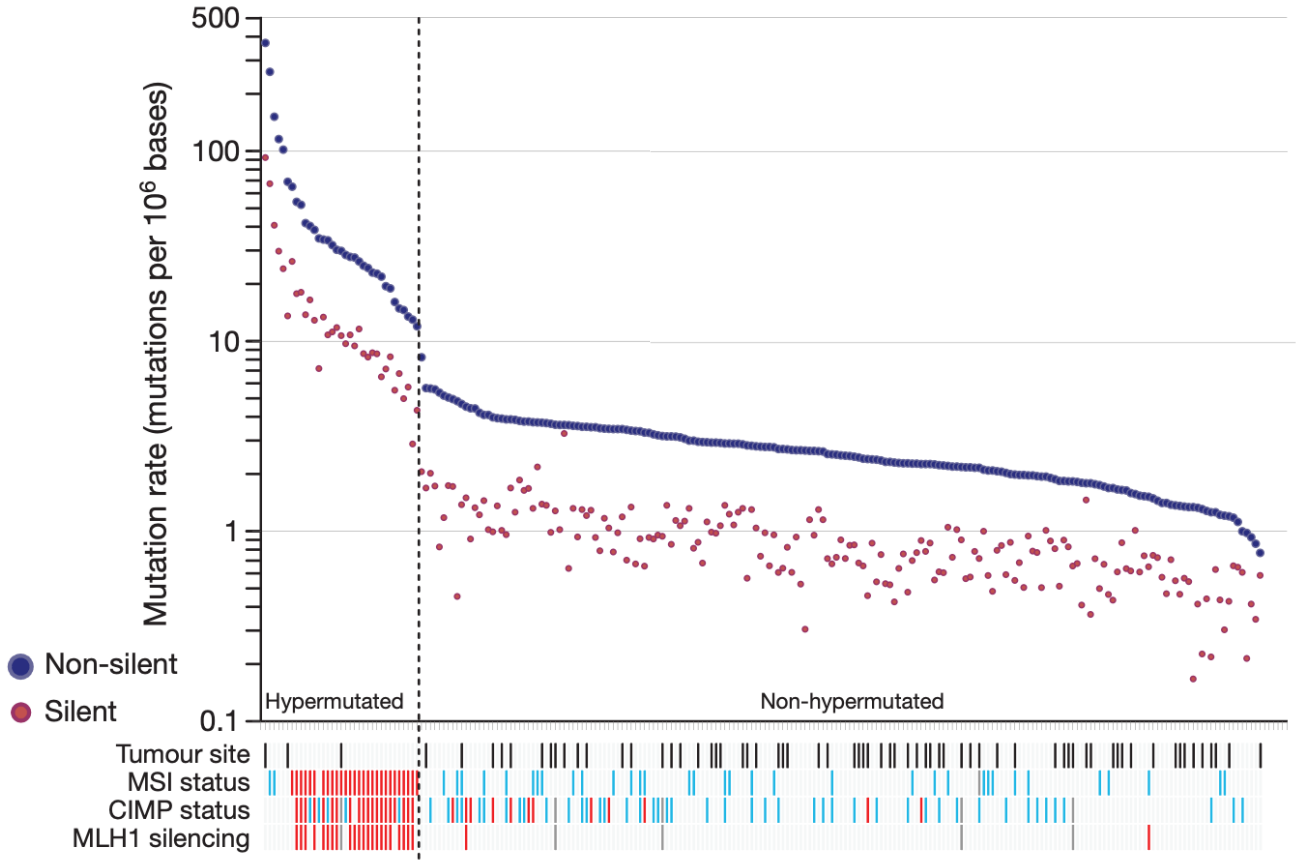
# of mutations per Mb

1 → 1/Mb * 3000Mb ~ 3000 mutations

2% are coding sequences → ~60 mutations in gene sequences

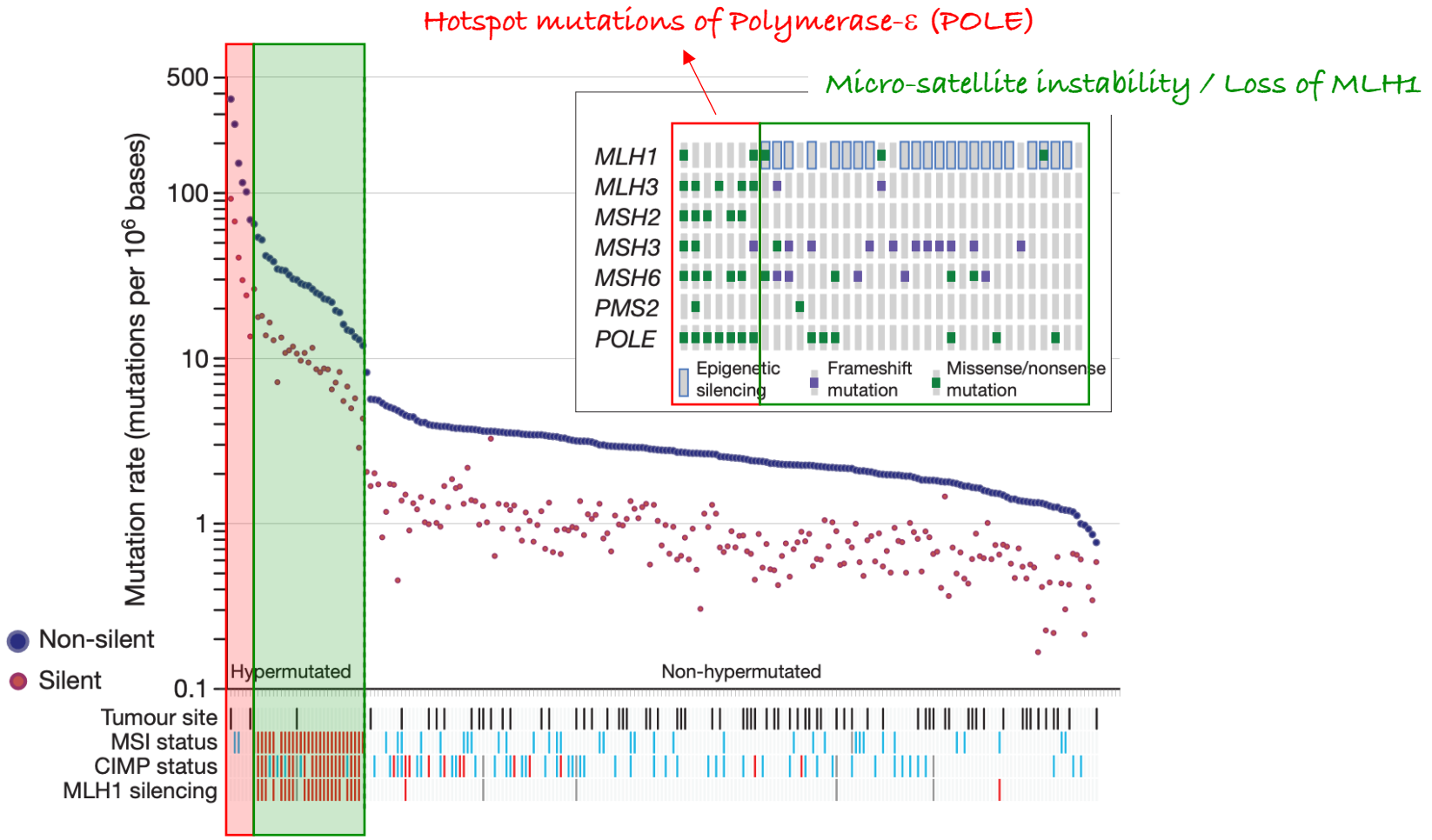
(Lawrence et al, Nature 2013)

Heterogenous mutation rates in the same tumor type

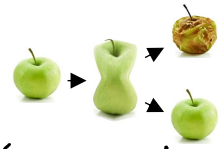


(TCGA, Nature 2012)

Heterogenous mutation rates in the same tumor type

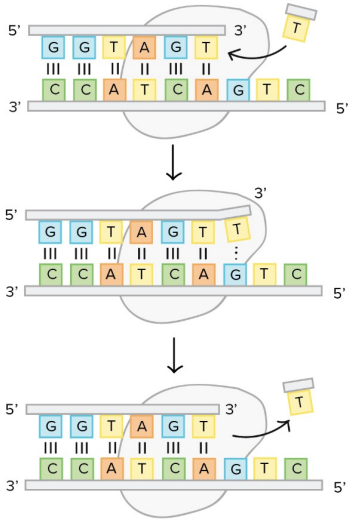


(TCGA, Nature 2012)

DNA repair defects  lead to high mutation rates

(not only in colon cancer)

Proofreading

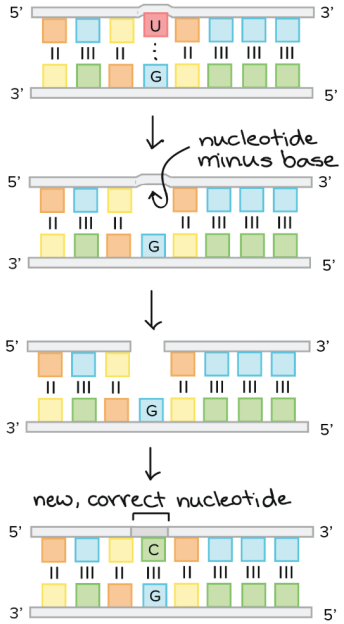


Polymerase adds an incorrect nucleotide to the new strand of DNA.

Polymerase detects that bases are mispaired.

Polymerase uses 3 '→ 5' exonuclease activity to remove incorrect nucleotide.

Base excision repair



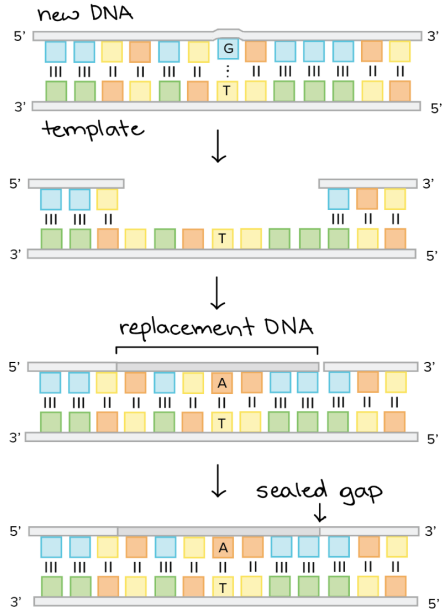
Deamination converts a cytosine base into a uracil.

The uracil is detected and removed, leaving a base-less nucleotide.

The base-less nucleotide is removed, leaving a small hole in the DNA backbone.

The hole is filled with the right base by a DNA polymerase, and the gap is sealed by a ligase.

Mismatch repair

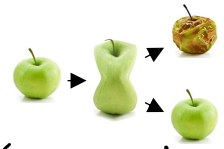


A mismatch is detected in newly synthesized DNA.

The new DNA strand is cut, and the mispaired nucleotide and its neighbors are removed.

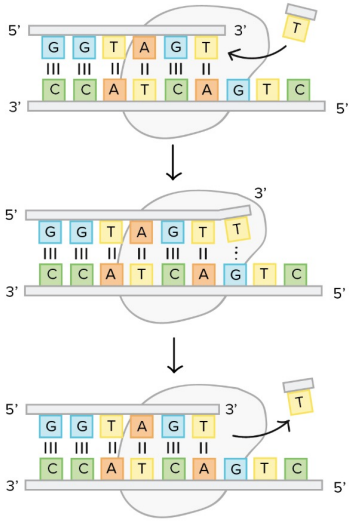
The missing patch is replaced with correct nucleotides by a DNA polymerase.

A DNA ligase seals the gap in the DNA backbone.

DNA repair defects  lead to high mutation rates

(not only in colon cancer)

Proofreading

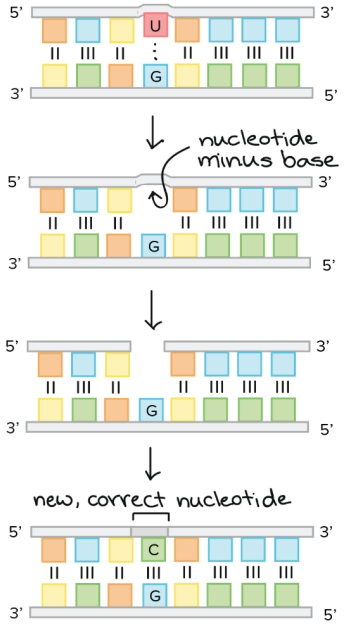


Polymerase adds an incorrect nucleotide to the new strand of DNA.

Polymerase detects that bases are mispaired.

Polymerase uses 3 '→ 5' exonuclease activity to remove incorrect nucleotide.

Base excision repair



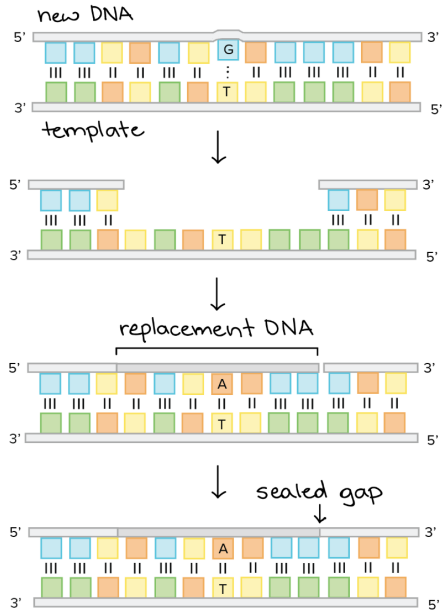
Deamination converts a cytosine base into a uracil.

The uracil is detected and removed, leaving a base-less nucleotide.

The base-less nucleotide is removed, leaving a small hole in the DNA backbone.

The hole is filled with the right base by a DNA polymerase, and the gap is sealed by a ligase.

Mismatch repair



A mismatch is detected in newly synthesized DNA.

The new DNA strand is cut, and the mispaired nucleotide and its neighbors are removed.

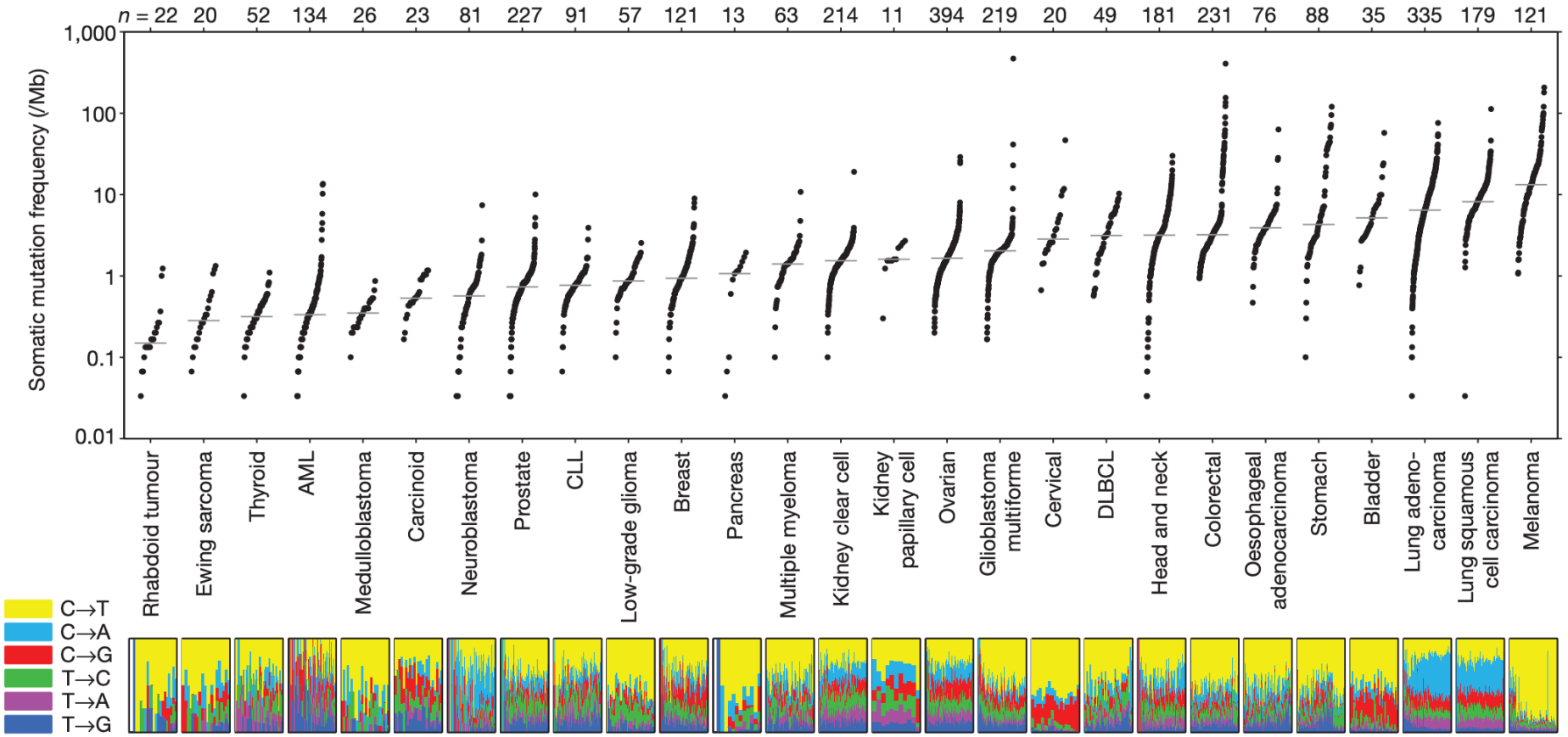
The missing patch is replaced with correct nucleotides by a DNA polymerase.

A DNA ligase seals the gap in the DNA backbone.

Mutations have different origins; can we trace them?

Mutational Signatures

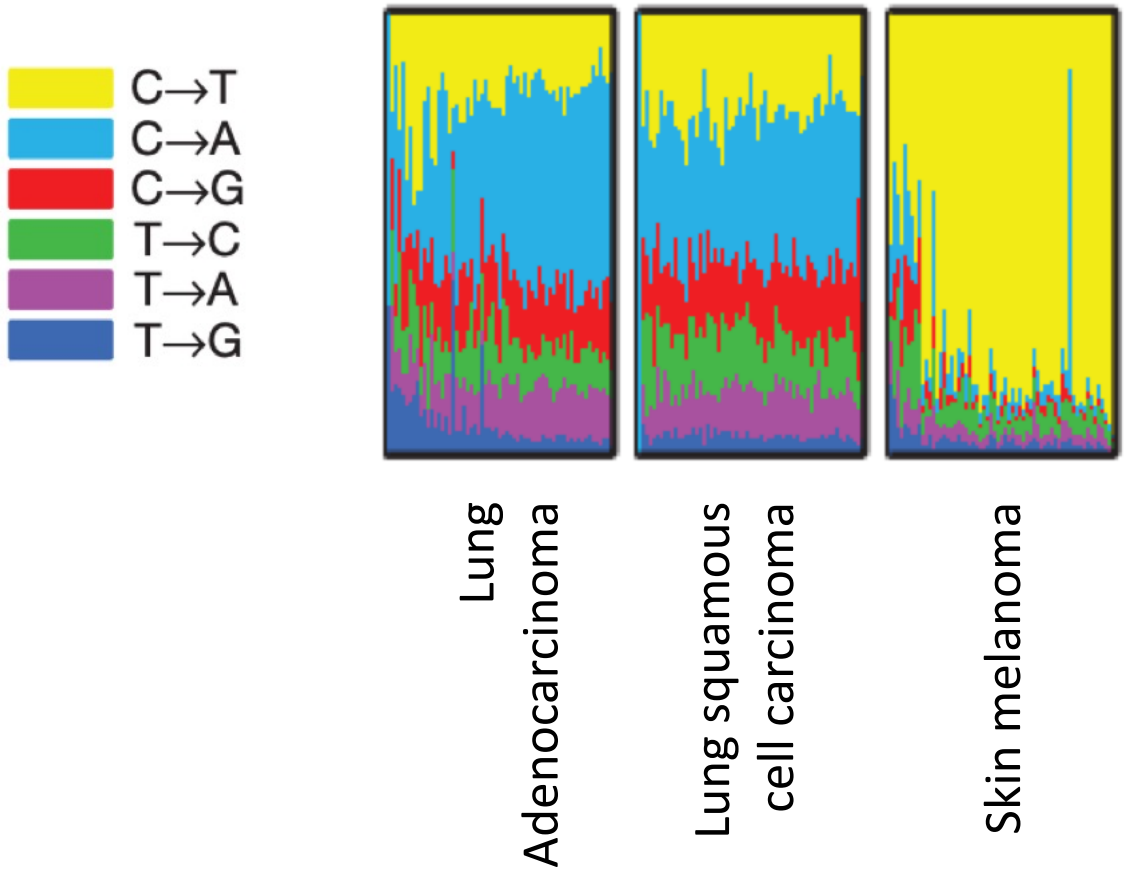
of mutations per Mb



(WES)

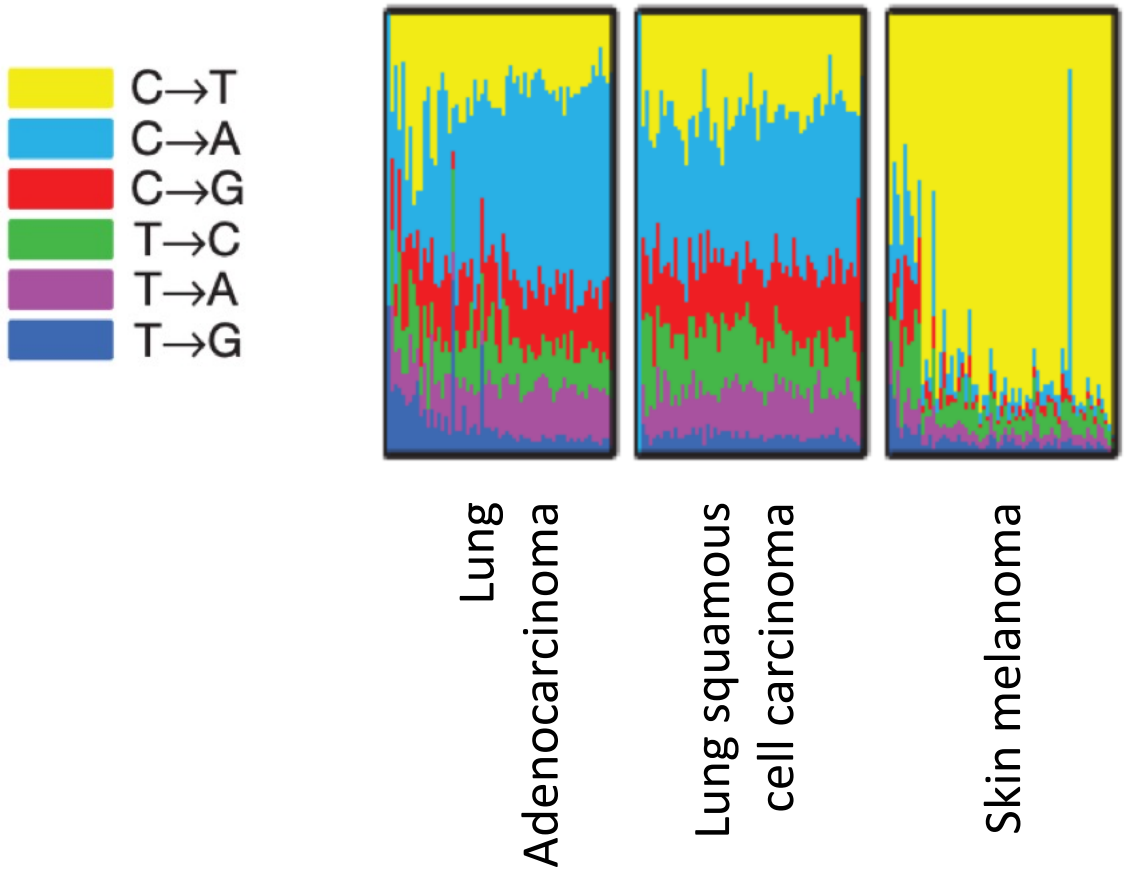
(Lawrence et al, Nature 2013)

Mutational Signatures



(Lawrence et al, Nature 2013)

Mutational Signatures



C → A

C → T

(Lawrence et al, Nature 2013)

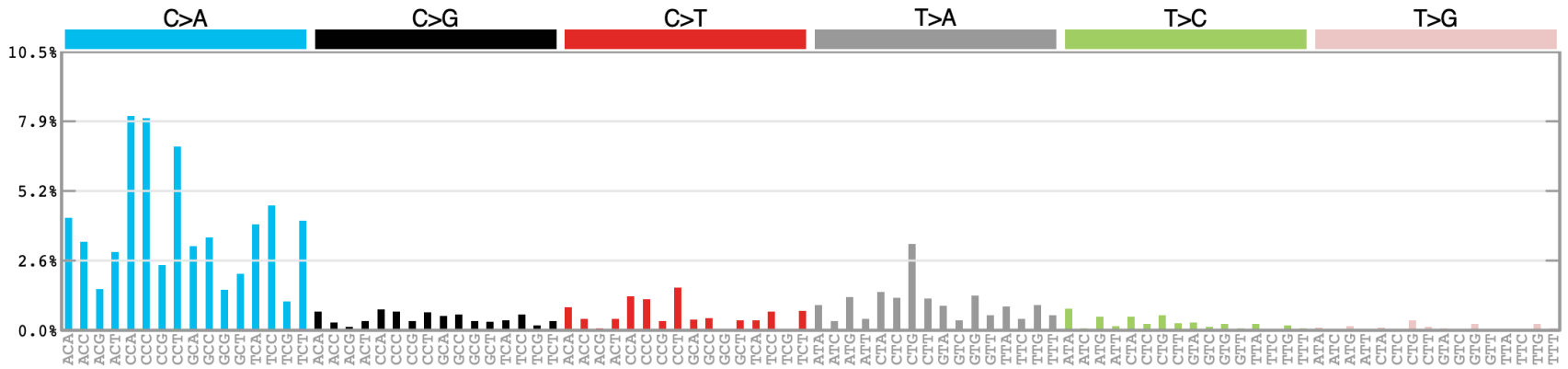
Mutational Signatures

C > A

Which nucleotide is before the mutated one

? C ?
v
? A ?

Which nucleotide is after the mutated one



Mutational Signatures

Spontaneous De-amination (defective base excision repair)

Tobacco smoking

(Alexandrov et al., Nature 2013)

Do all mutations provide a selective advantage?

i.e all mutations are bad and cause cancer?

Driver vs Passenger mutations

Mutational Processes

Driver mutations

Passenger mutations

Oncogenes /
Tumor Suppressors

Any kind of gene

Hallmarks of Cancer
(e.g. uncontrolled proliferation, escape cell death, etc.)

No phenotype

If you have a driver, usually you also have a passenger

Driver vs Passenger mutations

Mutational Processes

Driver mutations

Passenger mutations

Oncogenes /
Tumor Suppressors

Any kind of gene

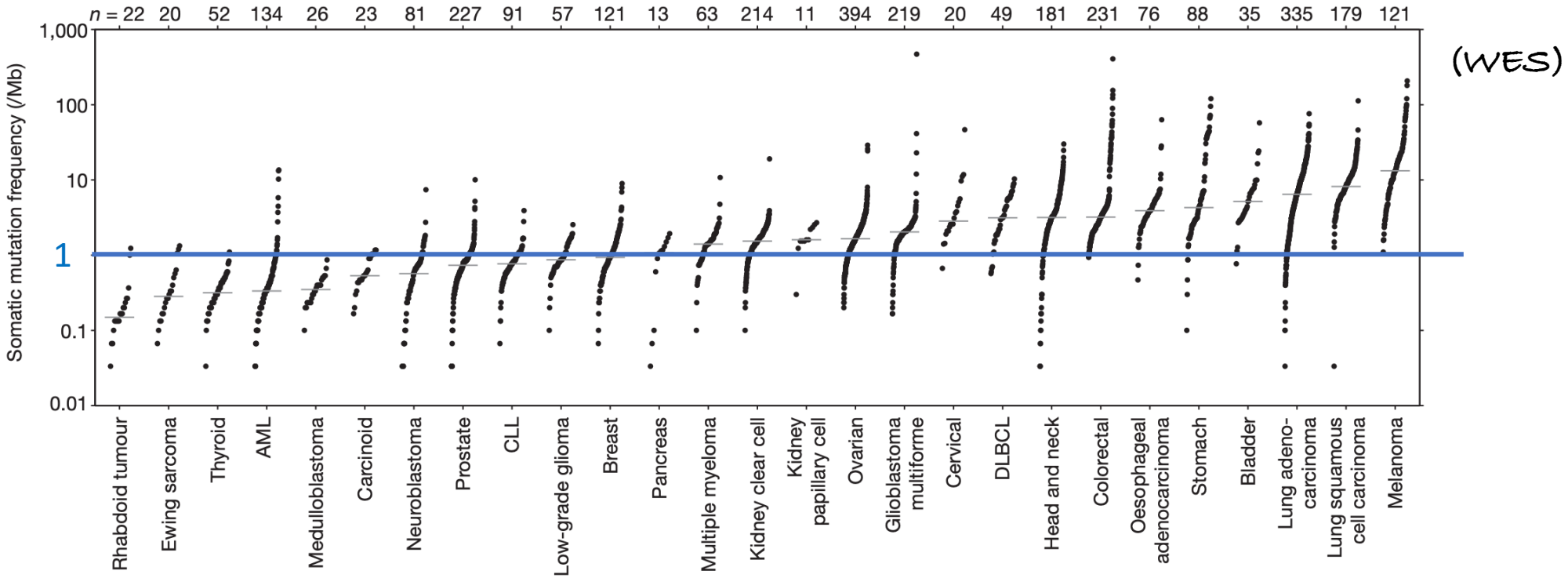
Hallmarks of Cancer
(e.g. uncontrolled proliferation, escape cell death, etc.)

No phenotype

Often you have many more passengers than drivers!

What do you get after sequencing 1000 tumors?

of mutations per Mb



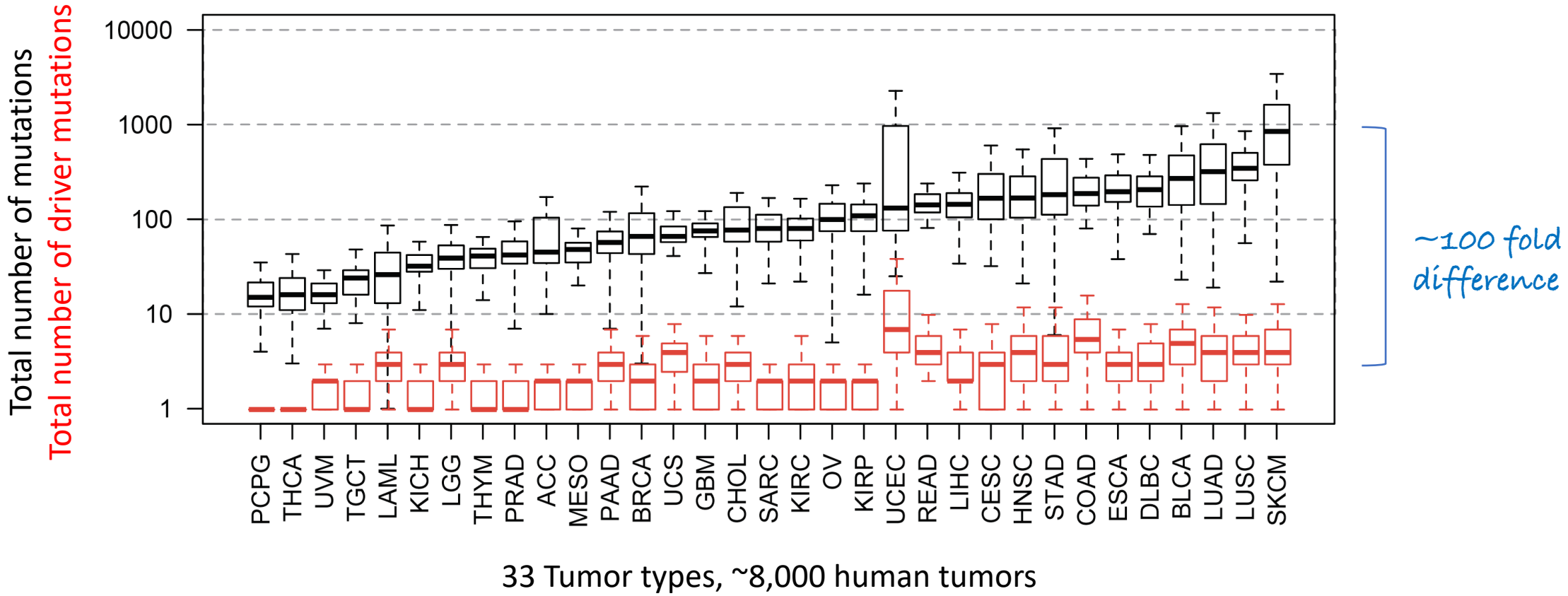
(WES)

1 → 1/Mb * 3000Mb ~ 3000 mutations

2% are coding sequences → ~60 mutations in gene sequences

(Lawrence et al, Nature 2013)

Driver vs. passenger alterations



Driver vs Passenger mutations

Mutational Processes

Driver mutations

Passenger mutations

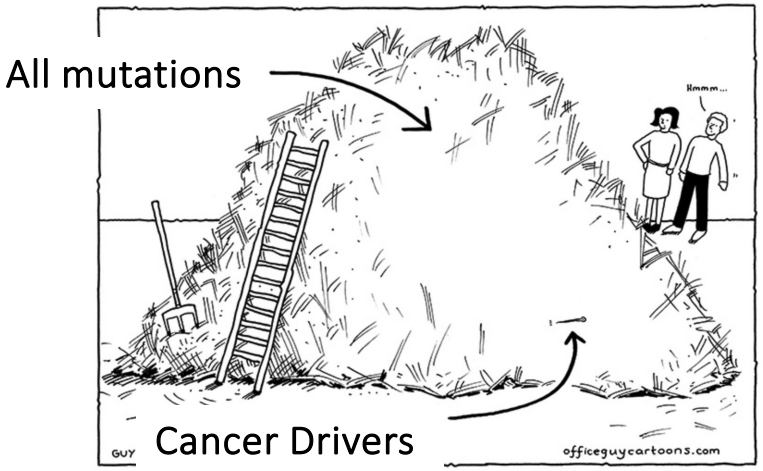
Oncogenes /
Tumor Suppressors

Any kind of gene

Hallmarks of Cancer
(e.g. uncontrolled proliferation, escape cell death, etc.)

No phenotype

How do we identify cancer drivers?



How do we identify driver mutations?

- “driver alterations” vs. “passenger alterations”

How do we identify driver mutations?

- “driver alterations” vs. “passenger alterations”

- Design an experiment to prove a given alteration can induce cancer
  - Cell lines / Mouse models / tumor organoids / etc.

How do we identify driver mutations?

- “driver alterations” vs. “passenger alterations”

- Design an experiment to prove a given alteration can induce cancer
  - Cell lines / Mouse models / tumor organoids / etc.

- *However, it is unfeasible to test >10,000 mutations!*

In the genomics era, we need strategies to prioritize

How do we identify driver mutations?

- “driver alterations” vs. “passenger alterations”

- **Main idea**: *A cancer driver gene is a gene that is mutated more frequently than <u>expected</u> in a large tumor cohort*

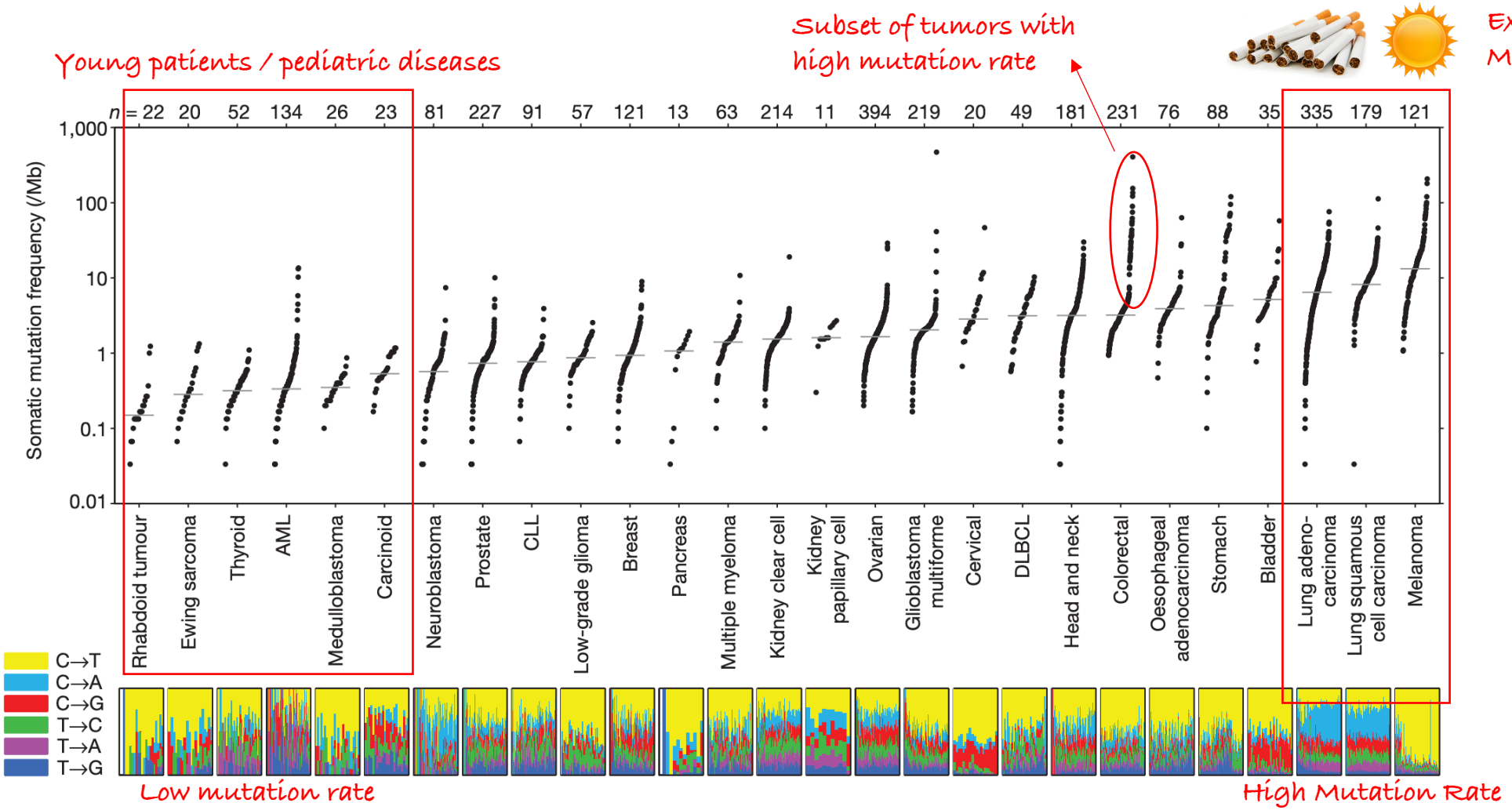
How do we identify driver mutations?

- “driver alterations” vs. “passenger alterations”

- **Main idea**: *A cancer driver gene is a gene that is mutated more frequently than expected in a large tumor cohort*

  - Count the number of mutations observed in each gene in the cohort
  - Determine the expected background mutation rate (BMR)
  - Estimate significance of observed vs. expected number of mutations

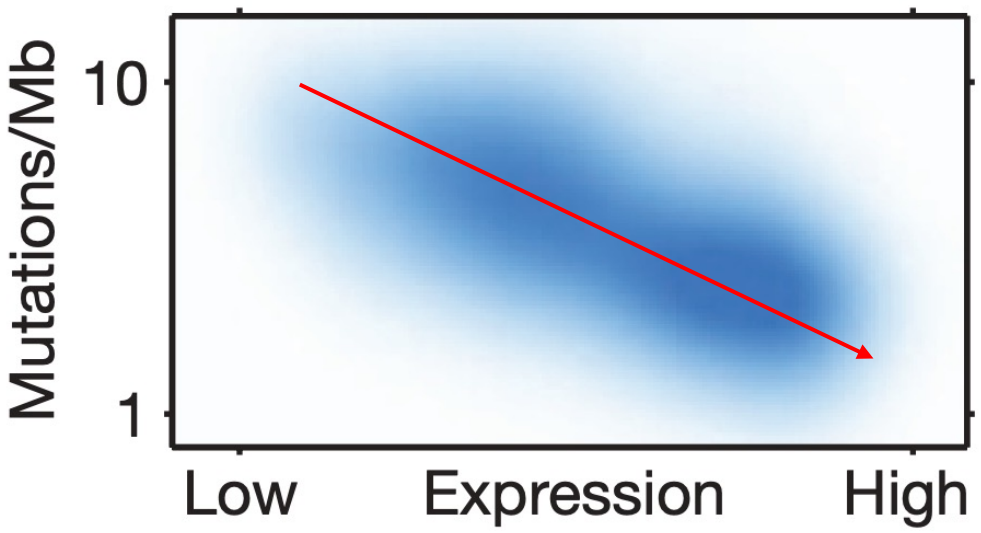
What does influence the BMR?



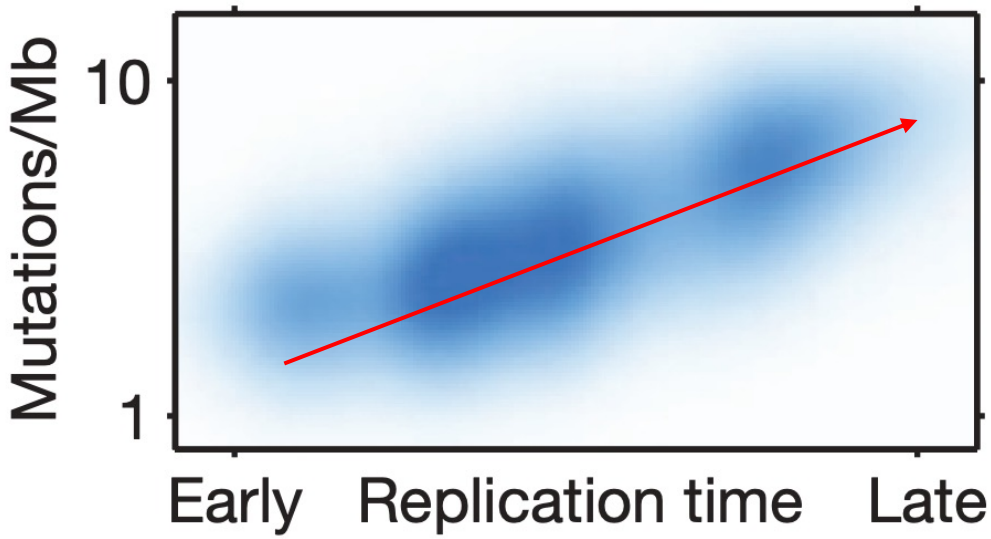
(Lawrence et al, Nature 2013)

What does influence the BMR?

Gene expression



Replication Time



What does influence the BMR?

- **Sample specific features**
  - Tissue-type
  - Impact of specific alterations
- **Underlying mutational processes**
  - (e.g. UV-light or tobacco consumption)
- **Regional genome properties:**
  - Gene expression
  - Replication time
  - Heterochromatin vs. euchromatin

What does influence the BMR?

- **Sample specific features**
  - Tissue-type
  - Impact of specific alterations
- **Underlying mutational processes**
  - (e.g. UV-light or tobacco consumption)
- **Regional genome properties:**
  - Gene expression
  - Replication time
  - Heterochromatin vs. euchromatin

BMR often inferred from silent and non-coding mutations in regions categorized based on covariates

Estimating Coding Drivers in 2020

Review Article | Published: 10 August 2020

A compendium of mutational cancer driver genes

Francisco Martínez-Jiménez, Ferran Muiños, Inés Sentís, Jordi Deu-Pons, Iker Reyes-Salazar, Claudia Arnedo-Pac, Loris Mularoni, Oriol Pich, Jose Bonet, Hanna Kranas, Abel Gonzalez-Perez ✉ & Nuria Lopez-Bigas ✉

Nature Reviews Cancer **20**, 555–572(2020) | Cite this article

Estimating Coding Drivers in 2020

Review Article | Published: 10 August 2020

A compendium of mutational cancer driver genes

Francisco Martínez-Jiménez, Ferran Muiños, Inés Sentís, Jordi Deu-Pons, Iker Reyes-Salazar, Claudia Arnedo-Pac, Loris Mularoni, Oriol Pich, Jose Bonet, Hanna Kranas, Abel Gonzalez-Perez ✉ & Nuria Lopez-Bigas ✉

Nature Reviews Cancer **20**, 555–572(2020) | Cite this article

What is the evidence of selection?

Estimating Coding Drivers in 2020

Review Article | Published: 10 August 2020

A compendium of mutational cancer driver genes

Francisco Martínez-Jiménez, Ferran Muiños, Inés Sentís, Jordi Deu-Pons, Iker Reyes-Salazar, Claudia Arnedo-Pac, Loris Mularoni, Oriol Pich, Jose Bonet, Hanna Kranas, Abel Gonzalez-Perez ✉ & Nuria Lopez-Bigas ✉

Nature Reviews Cancer **20**, 555–572(2020) | Cite this article

Evidence of selection:

- Recurrence (it requires to estimate BMR)

Estimating Coding Drivers in 2020

Review Article | Published: 10 August 2020

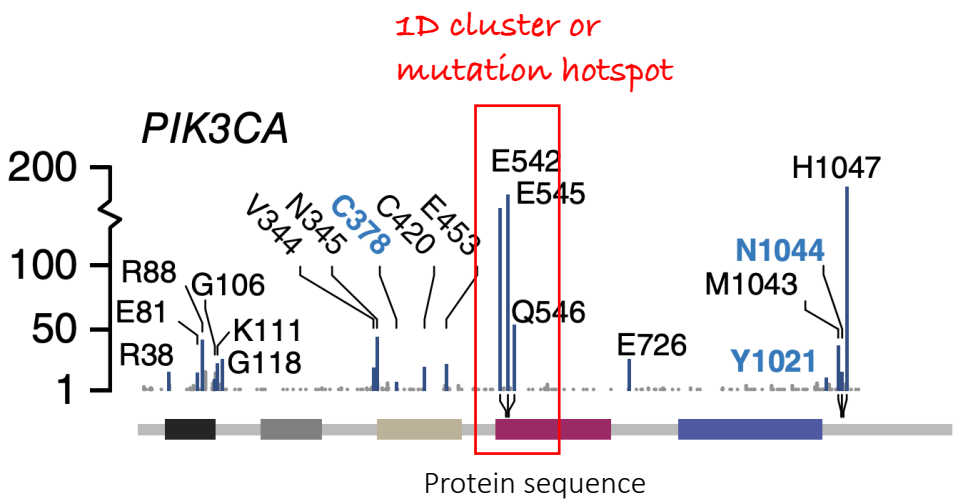
A compendium of mutational cancer driver genes

Francisco Martínez-Jiménez, Ferran Muiños, Inés Sentís, Jordi Deu-Pons, Iker Reyes-Salazar, Claudia Arnedo-Pac, Loris Mularoni, Oriol Pich, Jose Bonet, Hanna Kranas, Abel Gonzalez-Perez ✉ & Nuria Lopez-Bigas ✉

Nature Reviews Cancer **20**, 555–572(2020) | Cite this article

Evidence of selection:

- Recurrence (it requires to estimate BMR)
- Distribution of mutations (1D and 3D clusters)



(Chang et al. Nat. Biotech 2016)

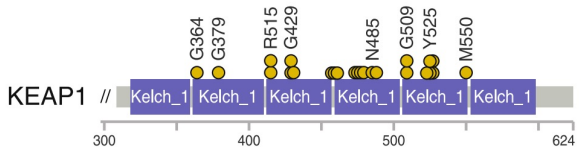
Estimating Coding Drivers in 2020

Review Article | Published: 10 August 2020

A compendium of mutational cancer driver genes

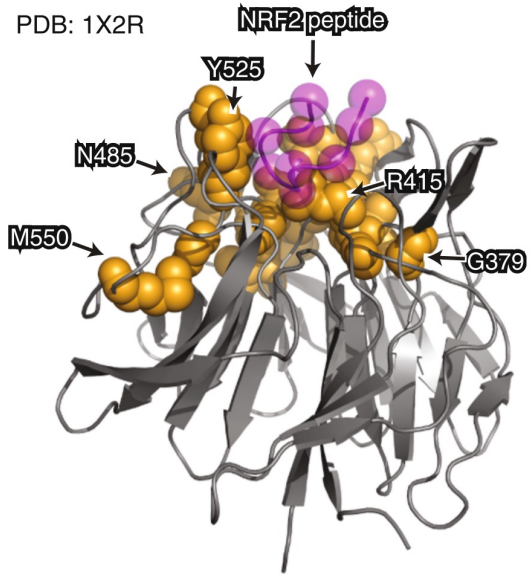
Francisco Martínez-Jiménez, Ferran Muiños, Inés Sentís, Jordi Deu-Pons, Iker Reyes-Salazar, Claudia Arnedo-Pac, Loris Mularoni, Oriol Pich, Jose Bonet, Hanna Kranas, Abel Gonzalez-Perez ✉ & Nuria Lopez-Bigas ✉

Nature Reviews Cancer **20**, 555–572(2020) | Cite this article



Evidence of selection:

- Recurrence (it requires to estimate BMR)
- Distribution of mutations (1D and 3D clusters)



(Gao et al. Gen. Biology 2017)

Estimating Coding Drivers in 2020

Review Article | Published: 10 August 2020

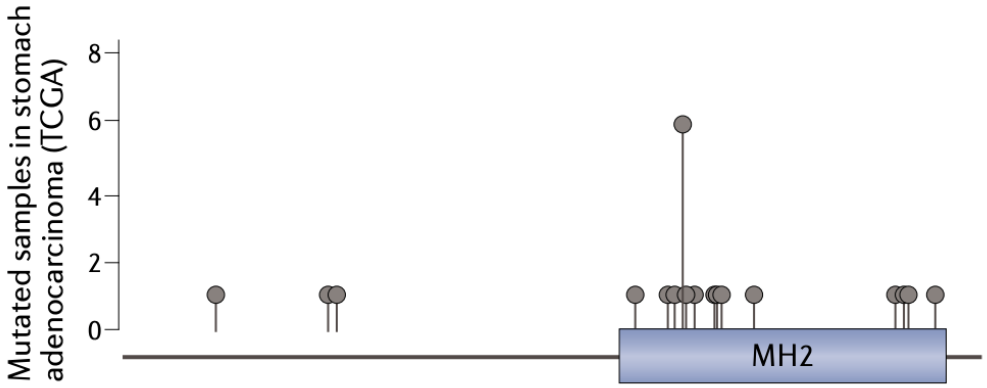
A compendium of mutational cancer driver genes

Francisco Martínez-Jiménez, Ferran Muiños, Inés Sentís, Jordi Deu-Pons, Iker Reyes-Salazar, Claudia Arnedo-Pac, Loris Mularoni, Oriol Pich, Jose Bonet, Hanna Kranas, Abel Gonzalez-Perez ✉ & Nuria Lopez-Bigas ✉

Nature Reviews Cancer **20**, 555–572(2020) | Cite this article

Evidence of selection:

- Recurrence (it requires to estimate BMR)
- Distribution of mutations (1D and 3D clusters)
- Mutations in functional domains



Gene	Tumour type	PFAM domain
SMAD4	Stomach adeno.	MH2 domain

Estimating Coding Drivers in 2020

Review Article | Published: 10 August 2020

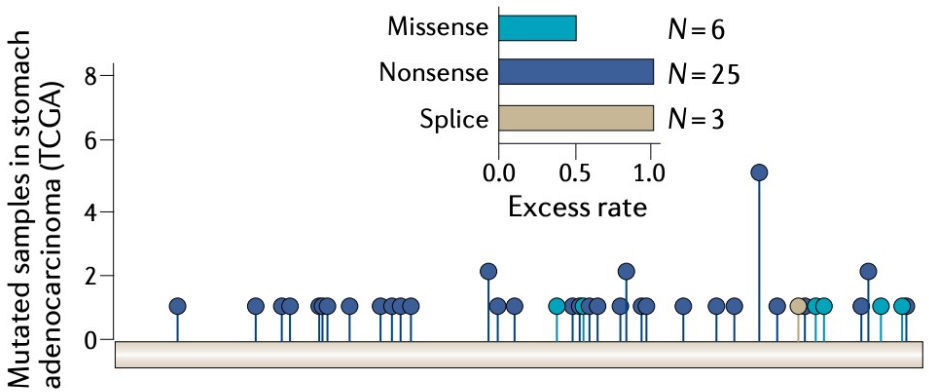
A compendium of mutational cancer driver genes

Francisco Martínez-Jiménez, Ferran Muiños, Inés Sentís, Jordi Deu-Pons, Iker Reyes-Salazar, Claudia Arnedo-Pac, Loris Mularoni, Oriol Pich, Jose Bonet, Hanna Kranas, Abel Gonzalez-Perez ✉ & Nuria Lopez-Bigas ✉

Nature Reviews Cancer **20**, 555–572(2020) | Cite this article

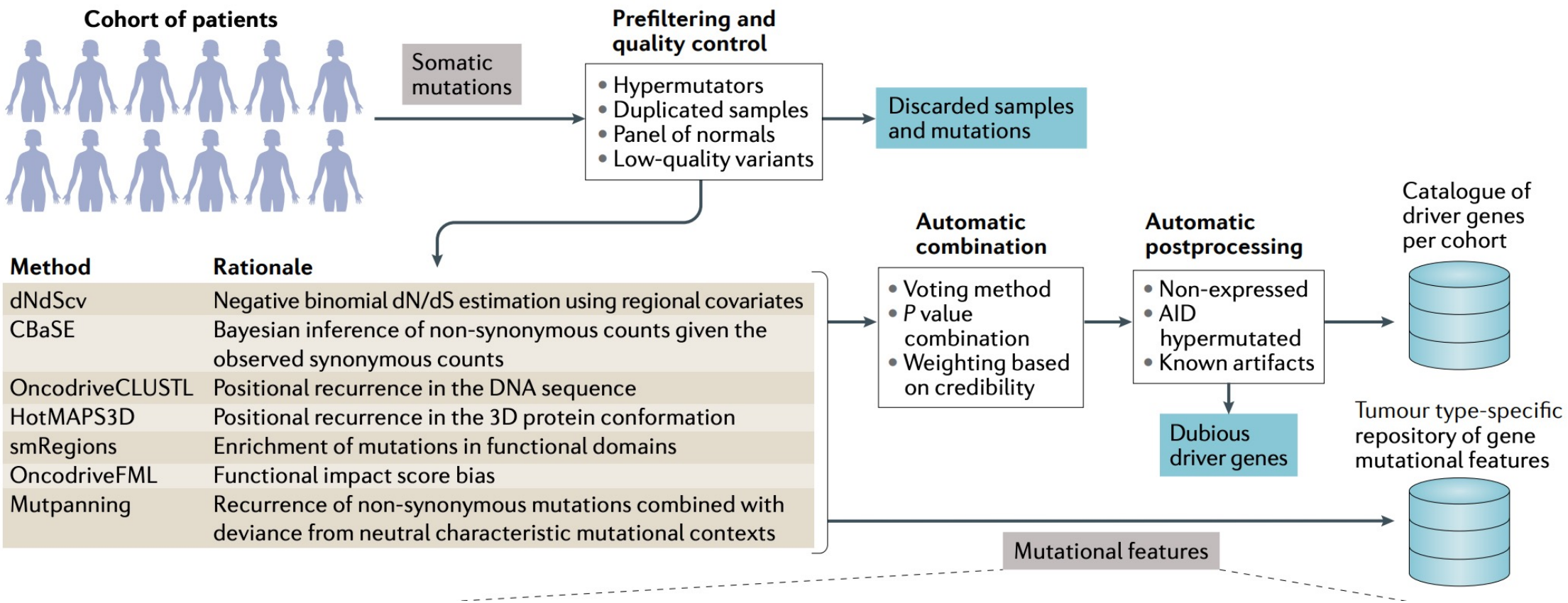
Evidence of selection:

- Recurrence (it requires to estimate BMR)
- Distribution of mutations (1D and 3D clusters)
- Mutations in functional domains
- *Functional impact bias*
  - *(also evolutionary conservation)*



Gene	Tumour type	Consequence type	Excess rate
ARID1A	Stomach adeno.	Missense	0.50
ARID1A	Stomach adeno.	Nonsense	0.99
ARID1A	Stomach adeno.	Splice	0.99

Estimating Coding Drivers in 2020



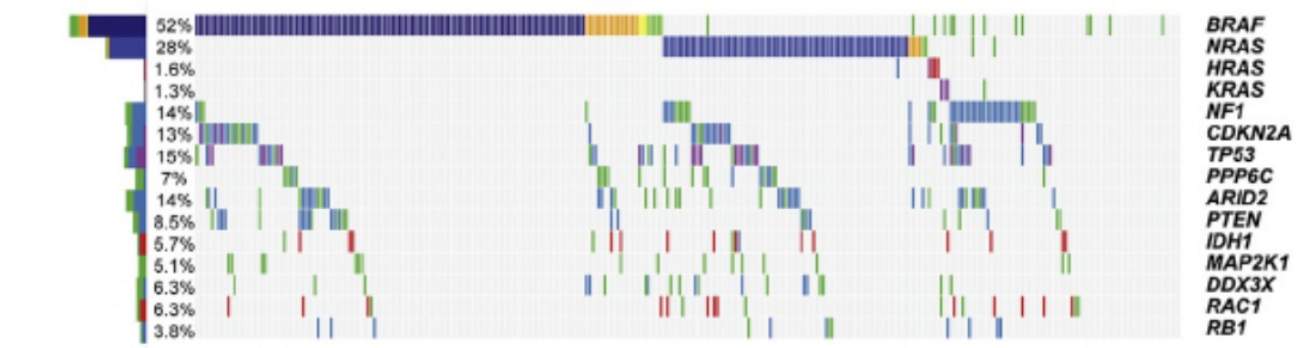
Method	Rationale
dNdScv	Negative binomial dN/dS estimation using regional covariates
CBaSE	Bayesian inference of non-synonymous counts given the observed synonymous counts
OncodriveCLUSTL	Positional recurrence in the DNA sequence
HotMAPS3D	Positional recurrence in the 3D protein conformation
smRegions	Enrichment of mutations in functional domains
OncodriveFML	Functional impact score bias
Mutpanning	Recurrence of non-synonymous mutations combined with deviance from neutral characteristic mutational contexts

(Martinez-Jimenez et al. Nat. Rev. Cancer, 2020)

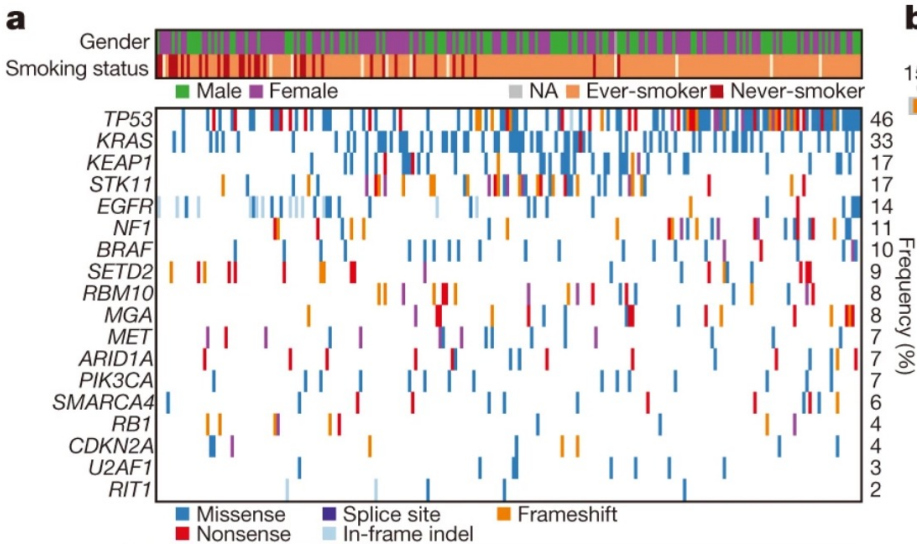
What did we learn about cancer from all these mutation analyses?

Cancer heterogeneity between different tumor types

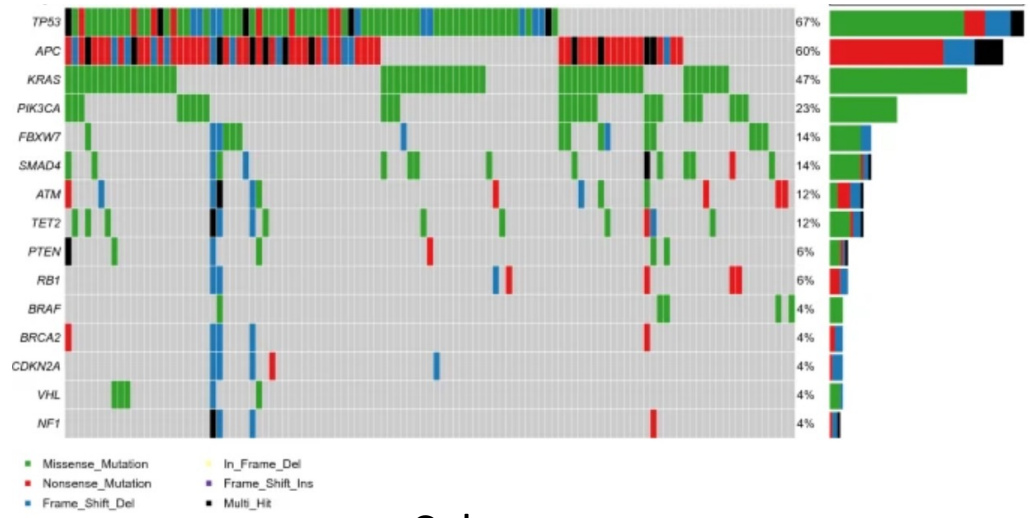
What are the most frequent mutated genes in each tissue?



Melanoma



Lung cancer

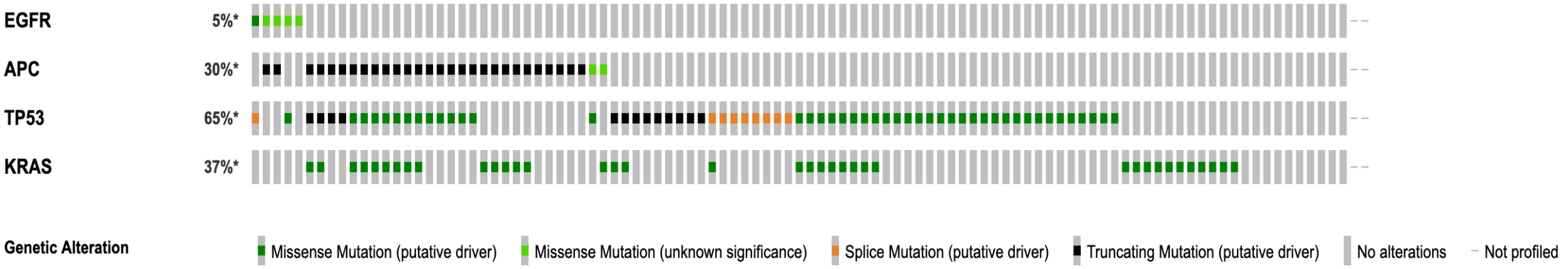


Colon cancer

Cancer inter-patient heterogeneity

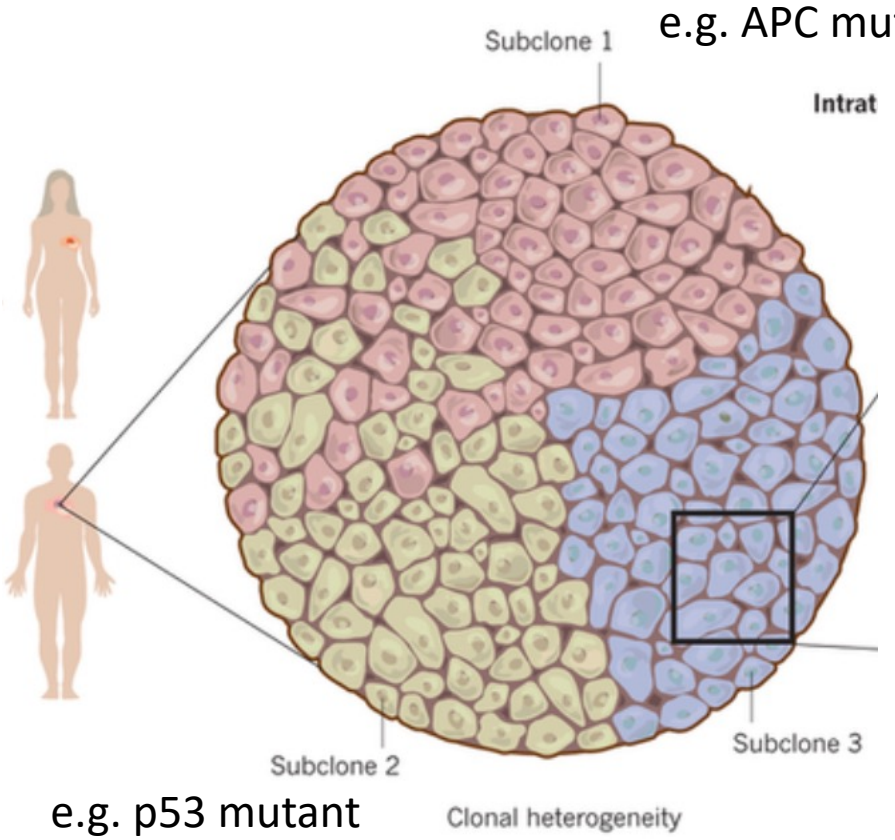
Same tumor: different mutation profile in each patient

103 patients colon cancer patient: each bar in this plot represents a patient



Cancer intra tumor heterogeneity

The same patient in different areas of the tumor have different mutations

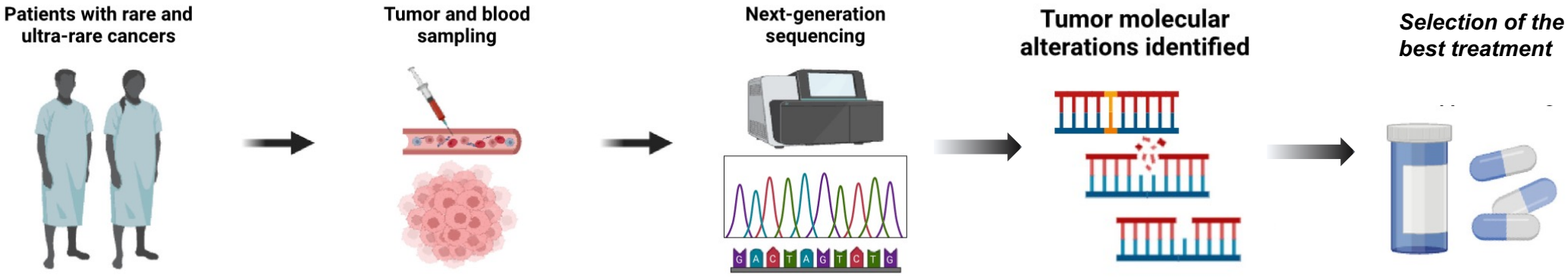


e.g. APC mutant

e.g. KRAS mutant

e.g. p53 mutant

Personalize oncology

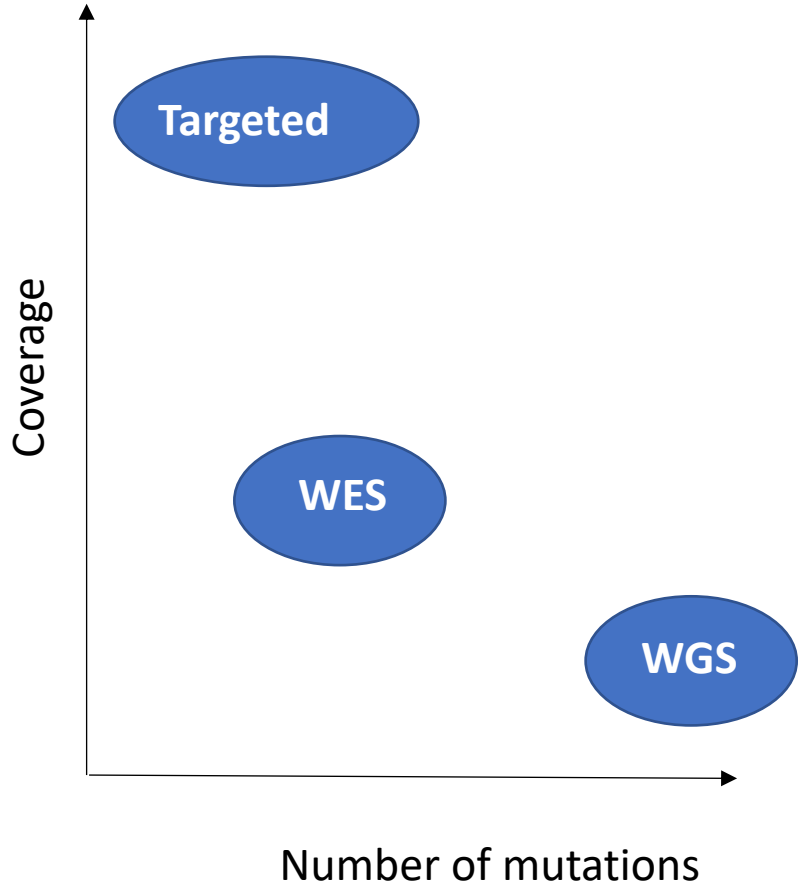


What type of sequencing?

Cancer Genomics

- Sequencing the blood and tumor biopsy
- **Targeted sequencing –** select a limited number of genes that will be sequenced (e.g. the most frequent mutated genes)
- **Whole-exome sequencing – sequence the coding genome**
- **Whole-genome sequencing- sequence coding and non-coding regions**

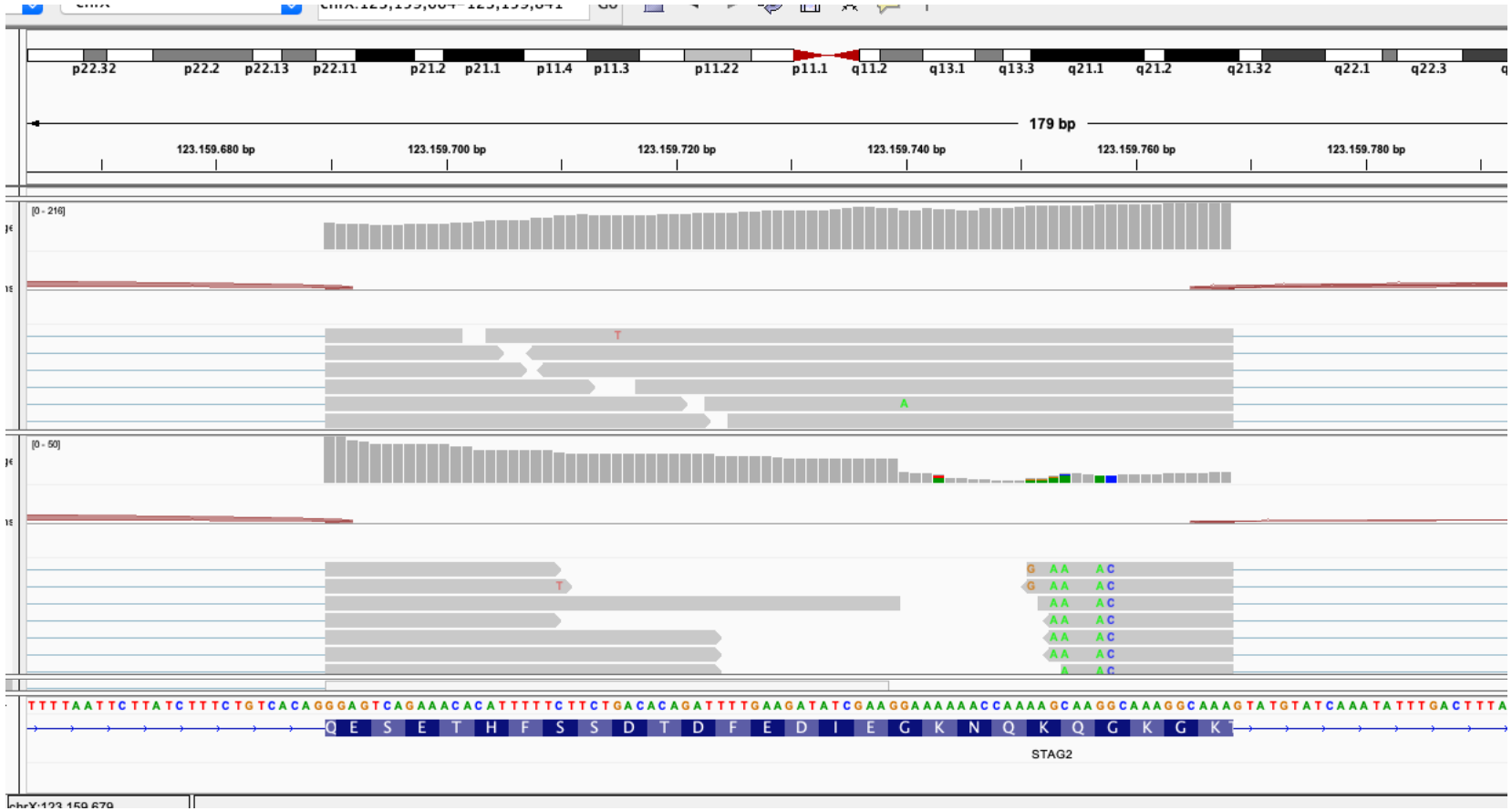
Sequencing a tumor



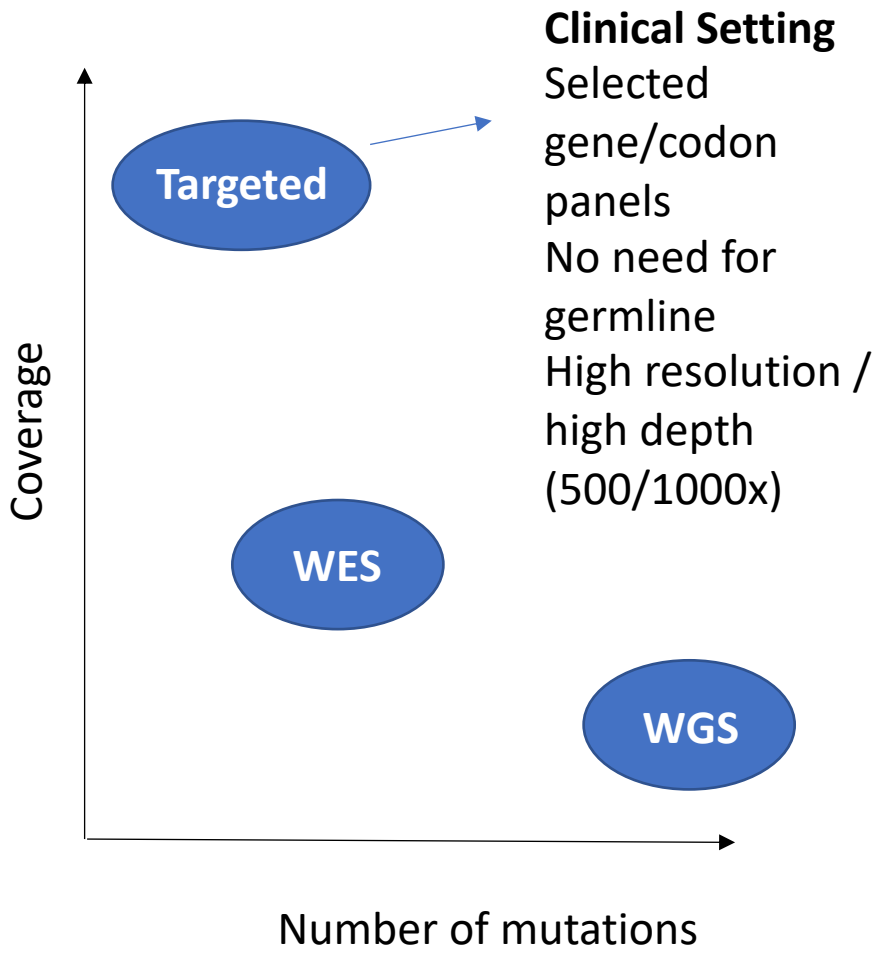
Coverage versus number of mutations

Wild-type

Mutated



Targeted Sequencing



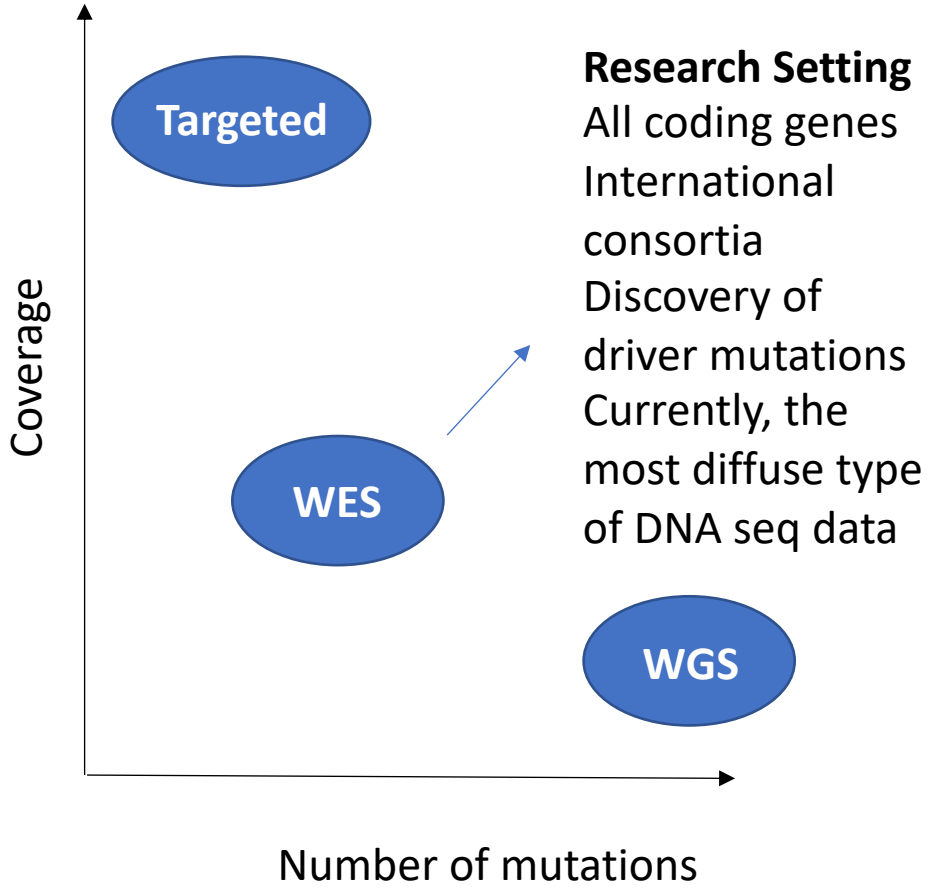
Clinical Setting

Selected gene/codon panels

No need for germline

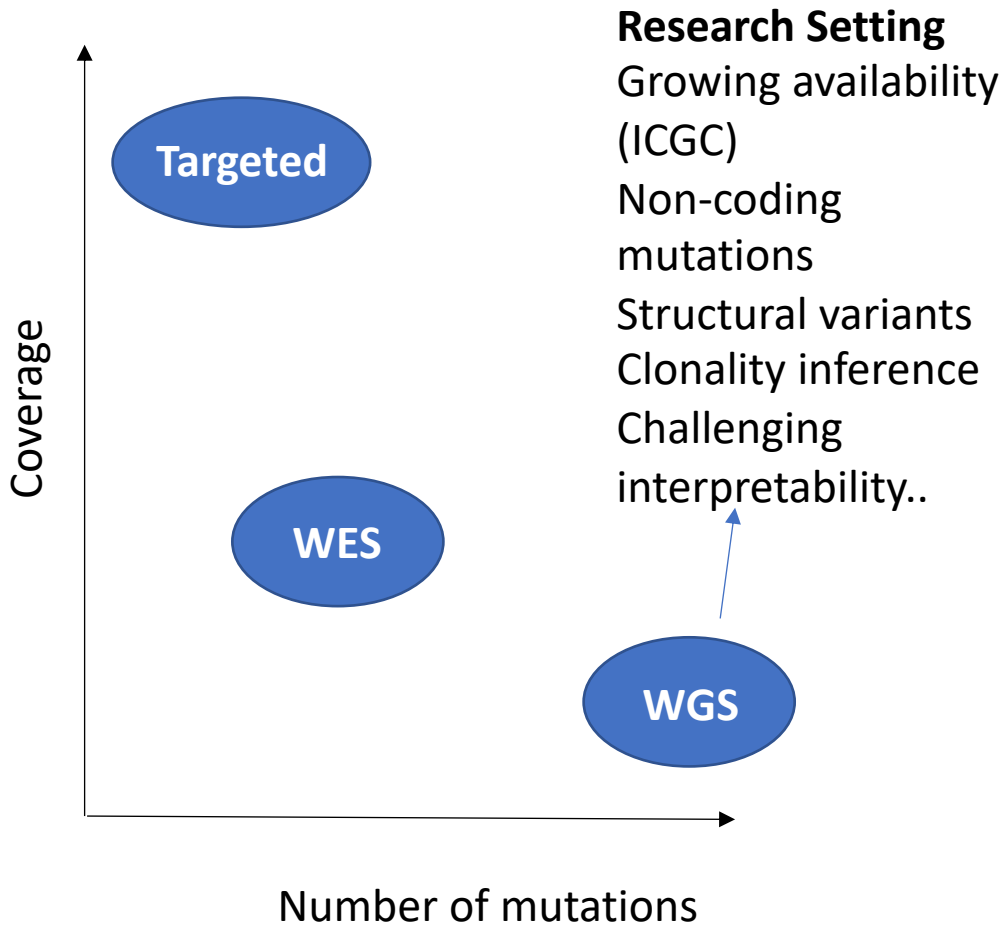
High resolution / high depth (500/1000x)

Whole Exome Sequencing



Research Setting
All coding genes
International consortia
Discovery of driver mutations
Currently, the most diffuse type of DNA seq data

Whole genome sequencing



Research Setting

Growing availability (ICGC)

Non-coding mutations

Structural variants

Clonality inference

Challenging interpretability..

Personalize oncology

To characterize each tumor based on their genomic profile

To select the best therapeutic options

Clinic: Establish the molecular tumor board where data each patient is discuss and several parameters need to be taken in consideration

Exercise

Exercise: 1.15h-2h30 Prepare the paper presentation
2.30h-3h30 presentation of the papers

https://www.nature.com/articles/s41586-020-1965-x

The pdf is in Moodle with a list of questions

Each group presents 1 figure, as it was organized in the previous exercise session.

Address some of the questions