

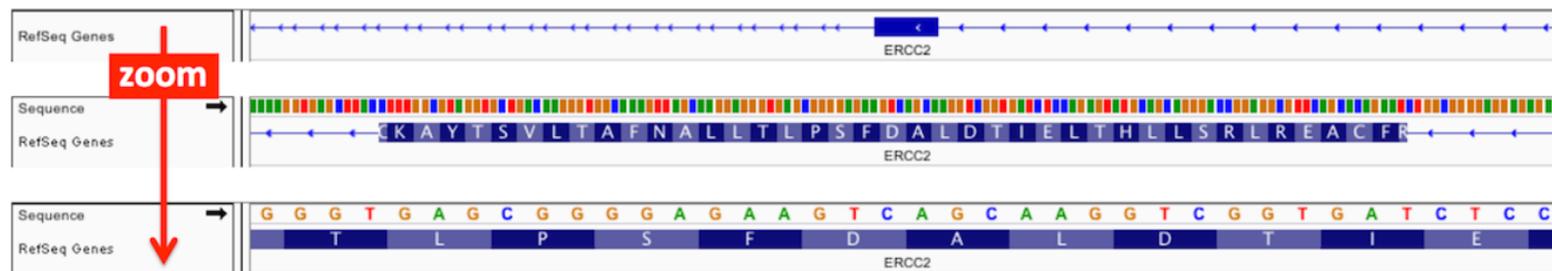
# An introduction to the cancer genome and mutational processes in cancer



Karolinska  
Institutet

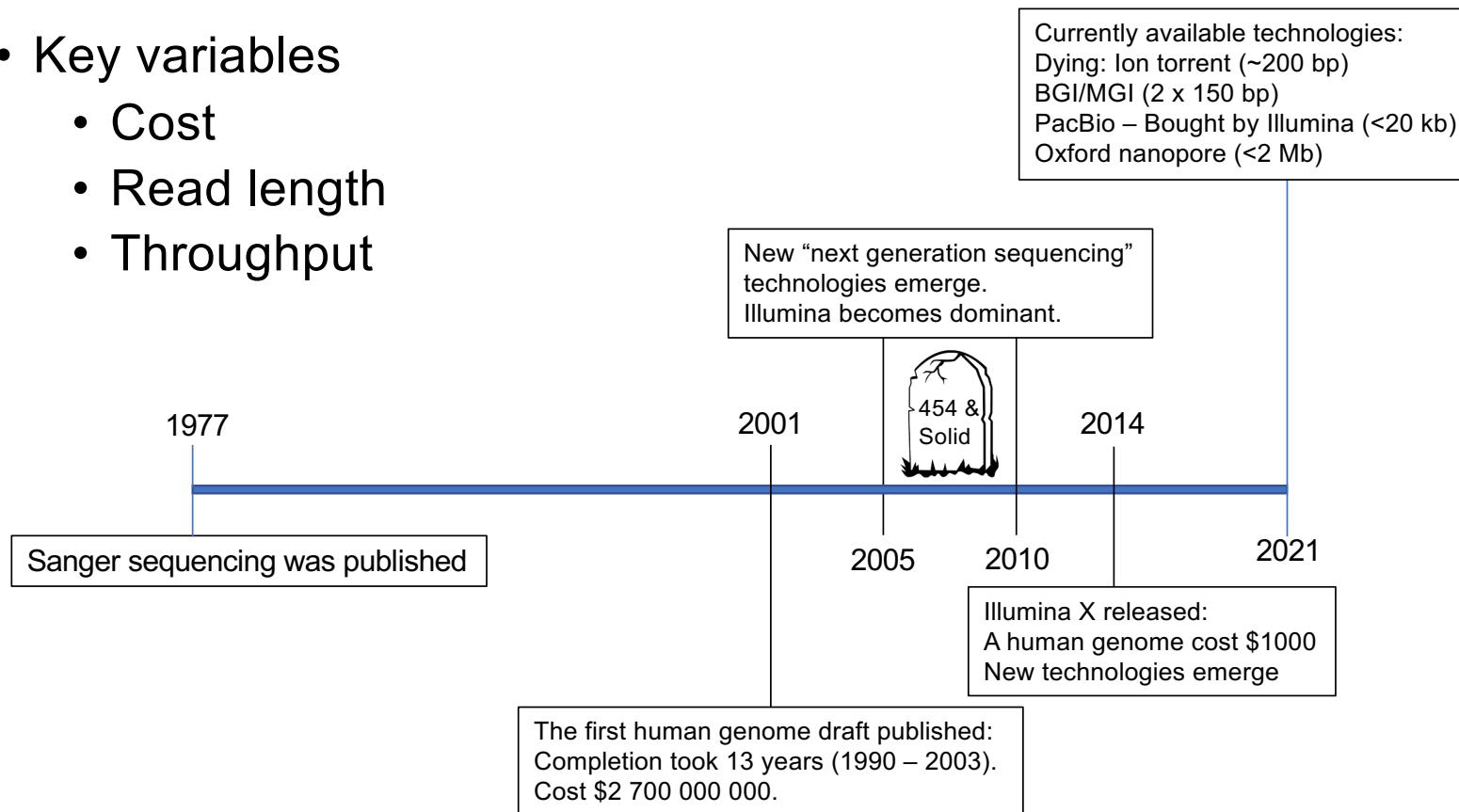
# The Sequencing revolution

- Sequencing = to determine the order of nucleotides in a DNA molecule.

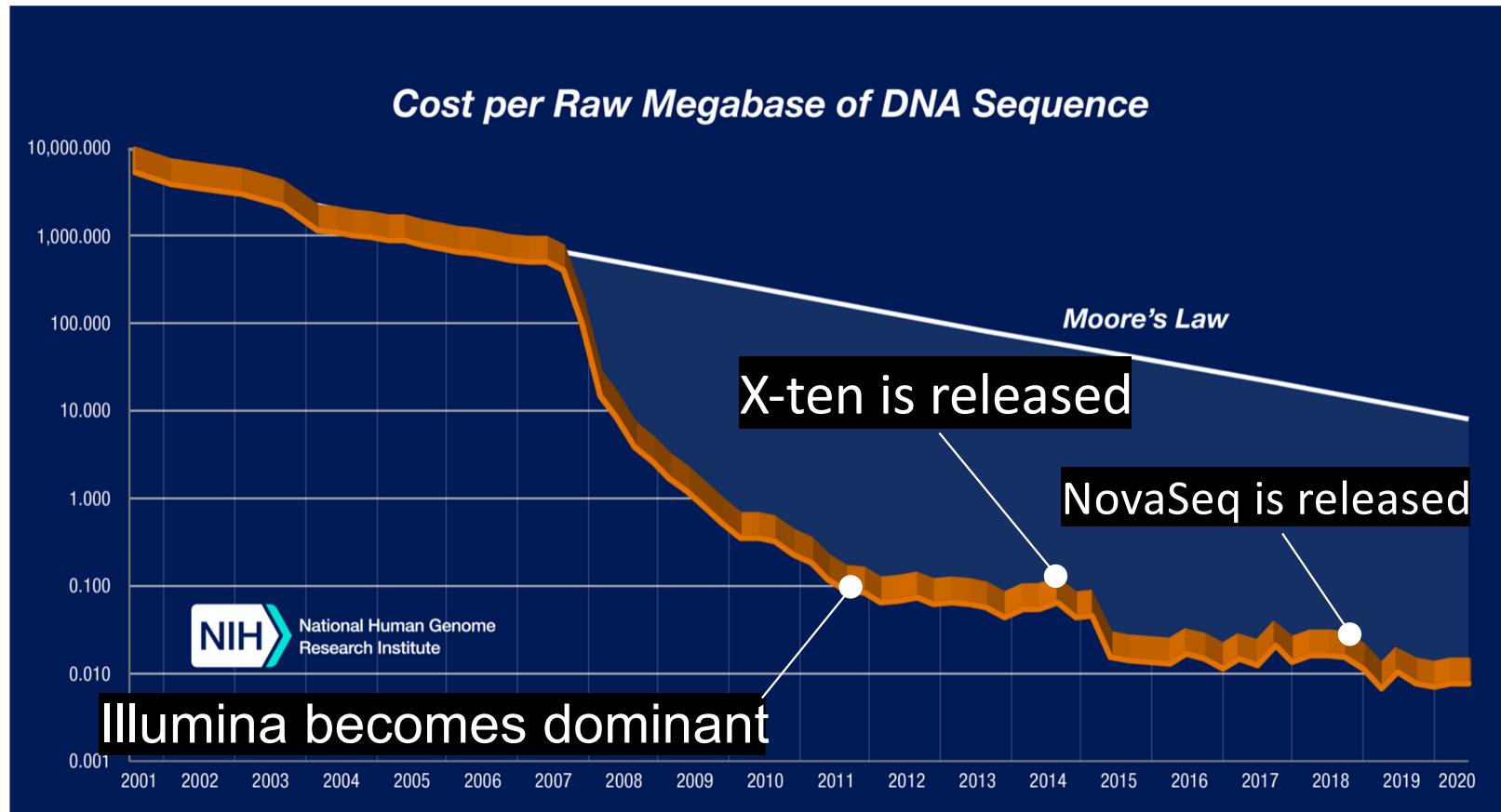


# Multiple companies on the market – one dominant

- Key variables
  - Cost
  - Read length
  - Throughput



## Lack of competition ...

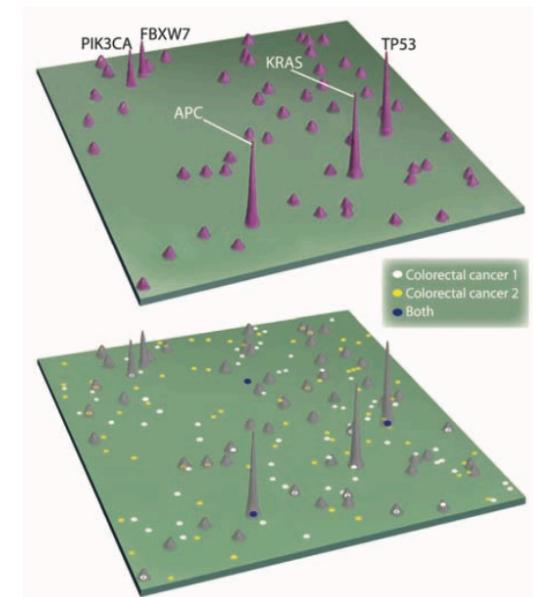


# The first set of genome data

- Exome-wide sequencing
- Whole genome amplification x 2
- Sanger sequencing
- 11 breast- and colorectal cancers
- The first take on “the long tail”



**The Genomic Landscapes of Human Breast and Colorectal Cancers**  
 Laura D. Wood, *et al.*  
*Science* **318**, 1108 (2007);  
 DOI: 10.1126/science.1145720



**The Consensus Coding Sequences of Human Breast and Colorectal Cancers**  
 Tobias Sjöblom, *et al.*  
*Science* **314**, 268 (2006);  
 DOI: 10.1126/science.1133427

# International consortia and other initiatives

---

- ICGC
  - International Cancer Genome Consortium
  - WGS + RNAseq on ~3000 cases
- TCGA
  - The Cancer Genome Atlas
  - WES + RNAseq + Methylation profiling etc on ~15000 cases
- Memorial Sloan Kettering
  - Panel-sequencing of >50.000 locally advanced/advanced cases
- The Hartwig Foundation
  - WGS + RNAseq on ~5000 cases in a clinical trial context

WGS = whole genome sequencing

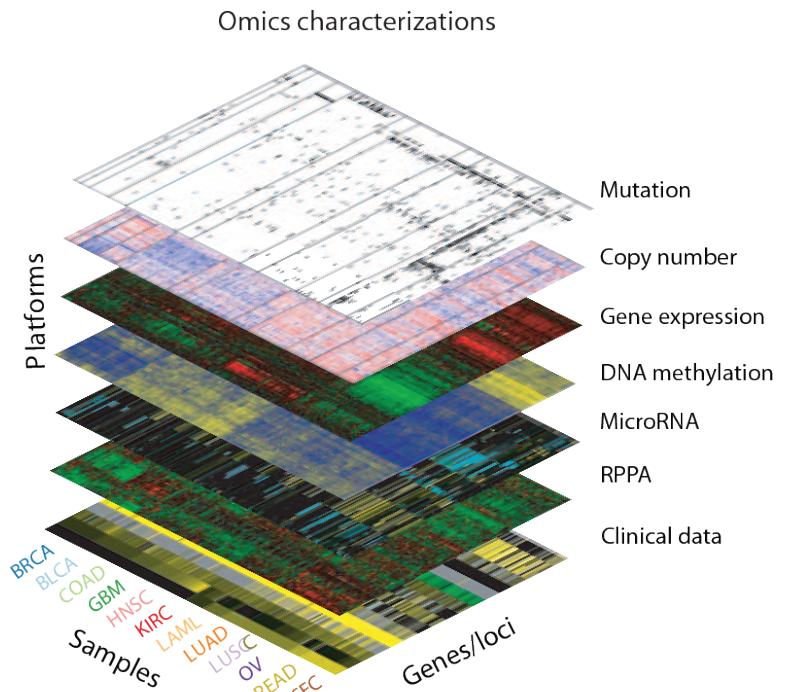
WES/WEX = whole exome sequencing

RNAseq = RNA sequencing

Targeted/panel sequencing = sequencing of selected parts of the human genome

# International consortia and other initiatives

- Course focus
  - DNA sequencing
  - RNA sequencing
  - Meaningful interpretation of single samples
- Complex data
  - E.g methylation
  - Need large cohorts for meaningful interpretation



The Cancer Genome Atlas Pan-Cancer analysis project, Nature Genetics 2013

## Landmark papers

---

- 2007-ish to 2012 new knowledge accumulates from publications on small cohorts of cancers
- 2012 and onward, key papers for different types of cancers from the TCGA
  - Significantly mutated genes
  - Cancer subtypes defined by integrated analysis of different data types
  - Pathway analysis

# Breast cancer – an example

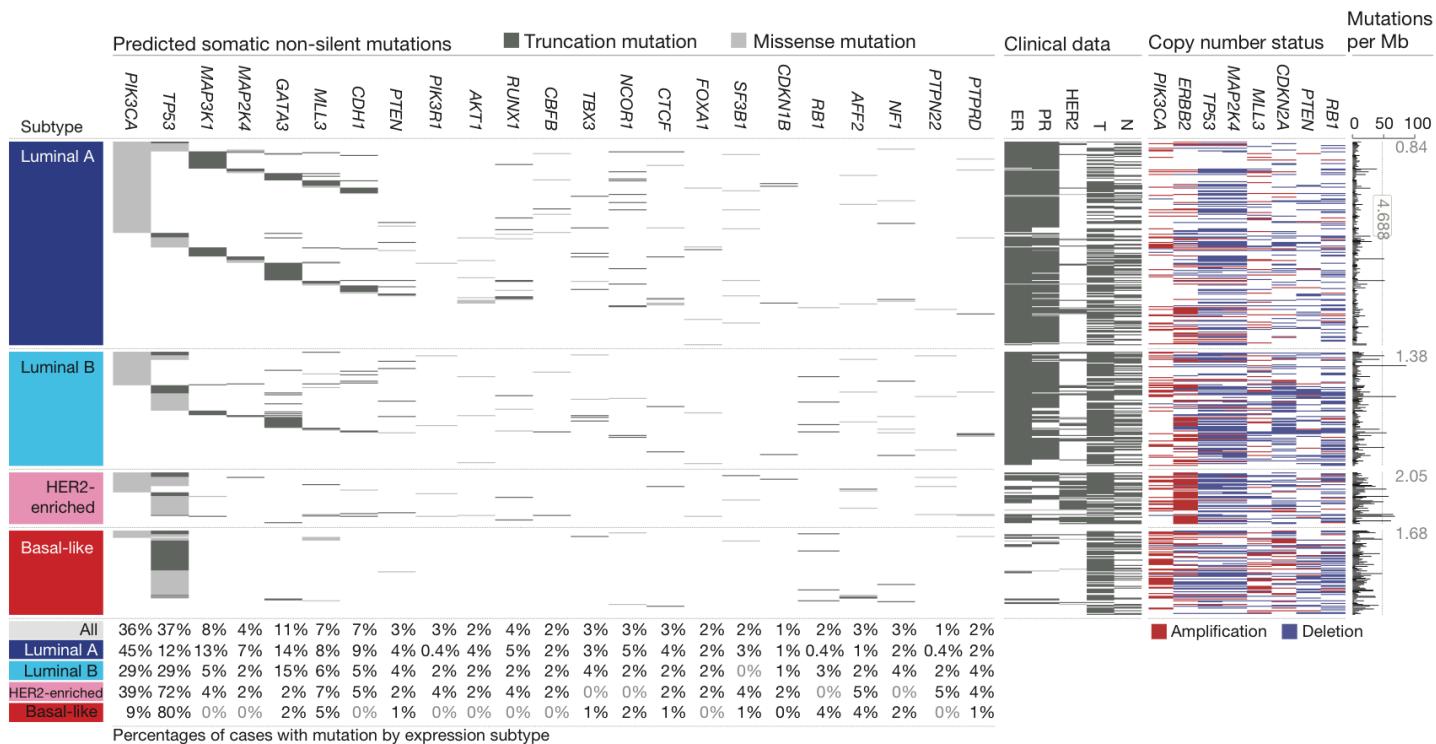
## ARTICLE

doi:10.1038/nature11412

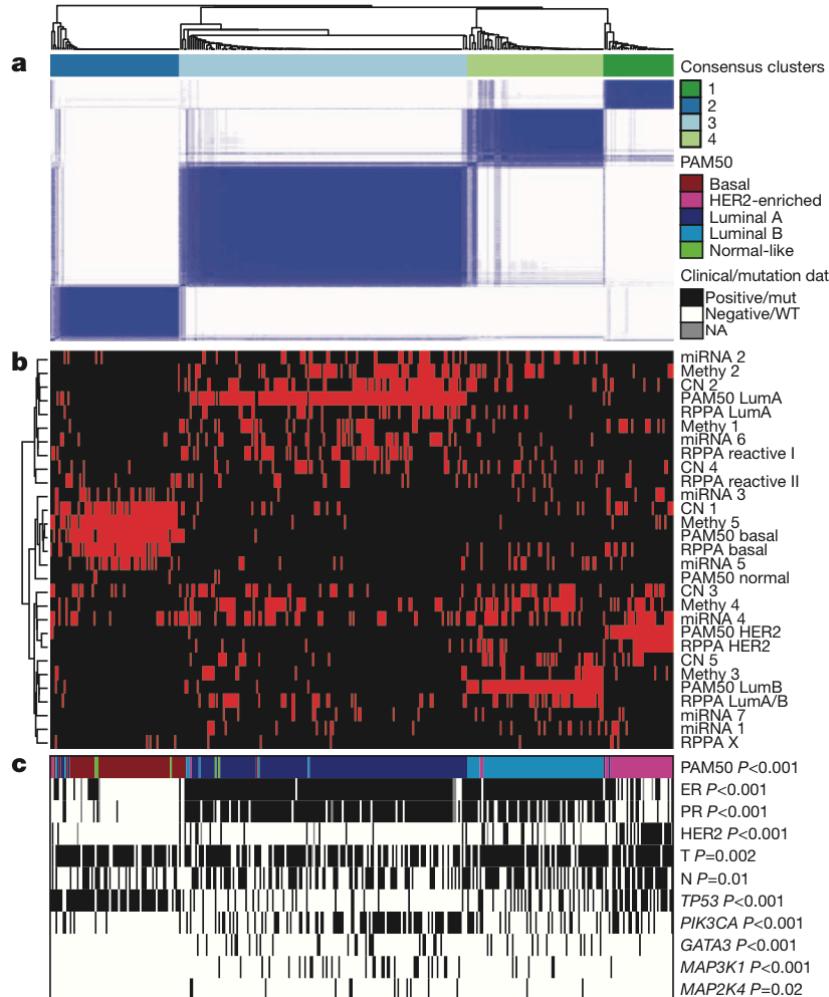
### Comprehensive molecular portraits of human breast tumours

The Cancer Genome Atlas Network\*

- ~500 patients with multiple data types



# Breast cancer – an example



- Cancer subtype
- NCI dictionary definition:
  - Describes the smaller groups that a type of cancer can be divided into, based on certain characteristics of the cancer cells.
  - These characteristics include how the cancer cells look under a microscope and whether there are certain substances in or on the cells or certain changes to the DNA of the cells.
- For me:
  - a fraction of cancers of a certain histological origin with a distinct molecular phenotype.
    - Example: Breast cancer subtypes (Basal, HER2, Luminal A and Luminal B).

# Two early great reviews in 2013



REVIEW

## Cancer Genome Landscapes

Bert Vogelstein, Nickolas Papadopoulos, Victor E. Velculescu, Shabin Zhou,  
Luis A. Diaz Jr., Kenneth W. Kinzler\*

- Insights into mutational processes causing cancer
- Structural variant patterns
- New categories of significantly mutated genes
- The long tail of driver variants
- Heterogeneity (within the same tumor, between tumors of the same diagnosis, between cancers of different diagnoses)
- ...
- The main take-homes summarized in a couple of slides

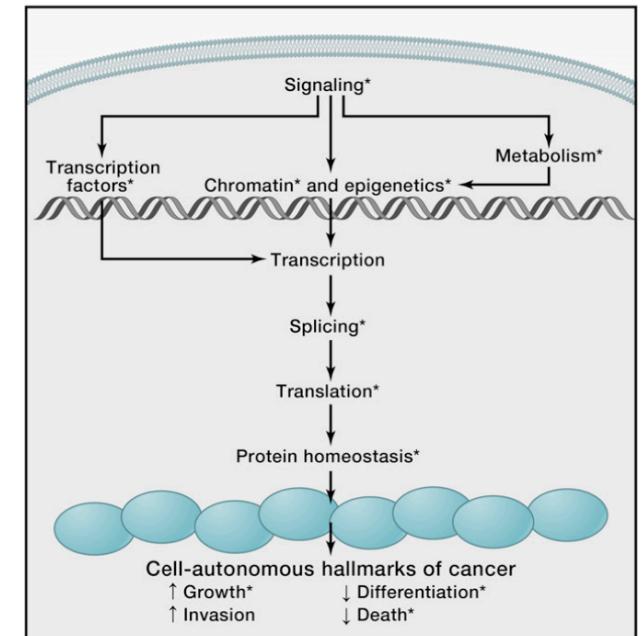


Leading Edge  
Review

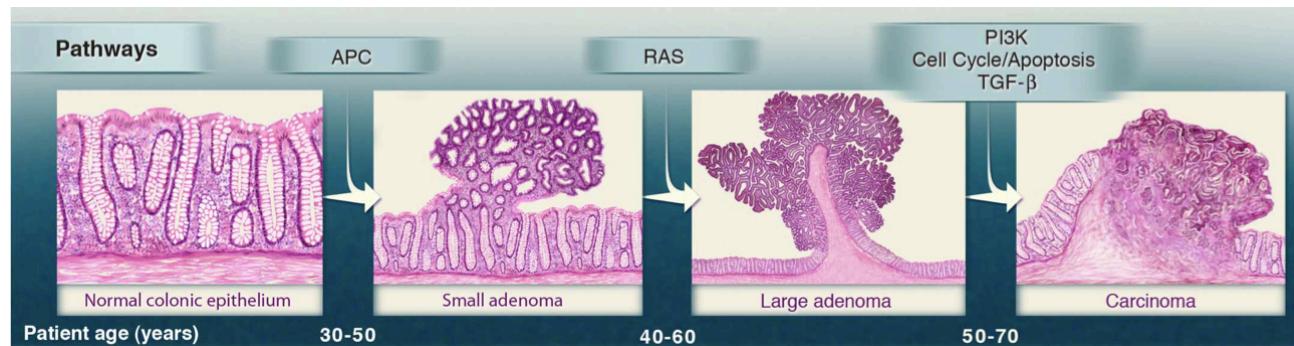
Cell

## Lessons from the Cancer Genome

Levi A. Garraway<sup>1,2,4</sup> and Eric S. Lander<sup>3,4,5,\*</sup>



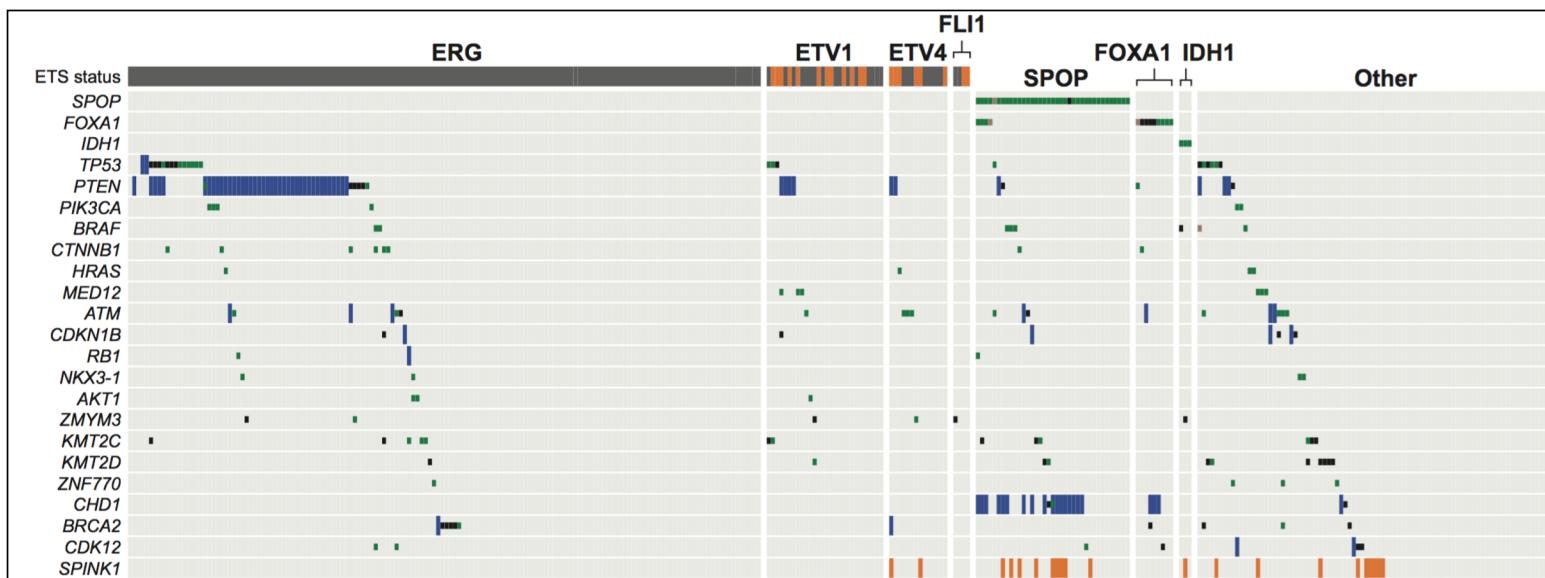
# Cancer development



- Occurs through sequential somatic mutations
- Takes decades to develop
- Majority of deaths are due to late detection (after 90% of cancers lifespan)
- Claimed that mortality can be reduced with >75% if improvements are made in early detection/prevention (more on this later during cfDNA)

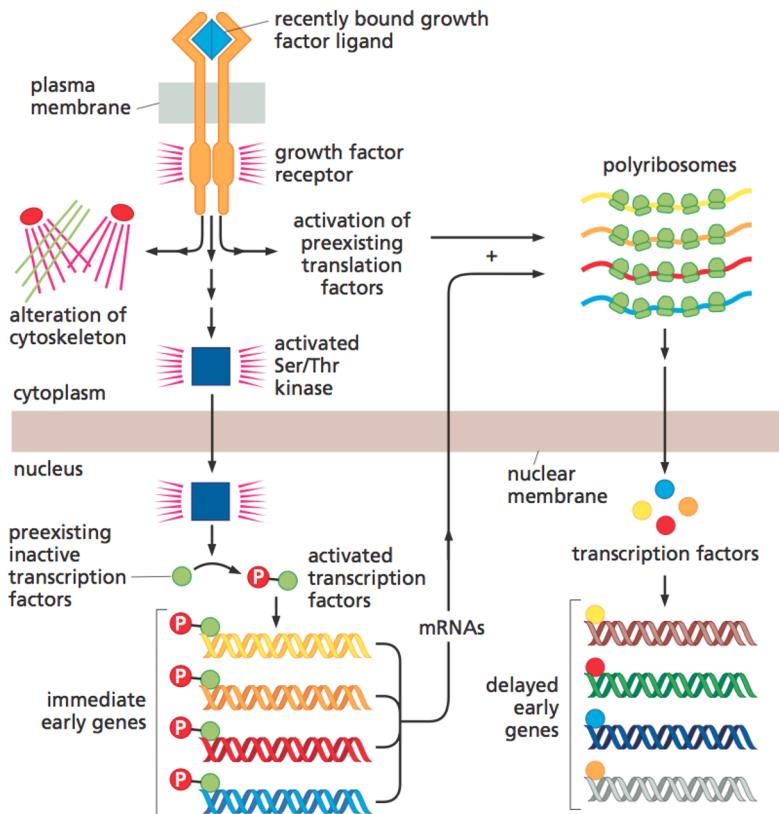
# Cancer – heterogeneous in its nature

- Localized prostate cancer – intra disease heterogeneity
- Subtypes, defined by the main driver detected in precursor lesions



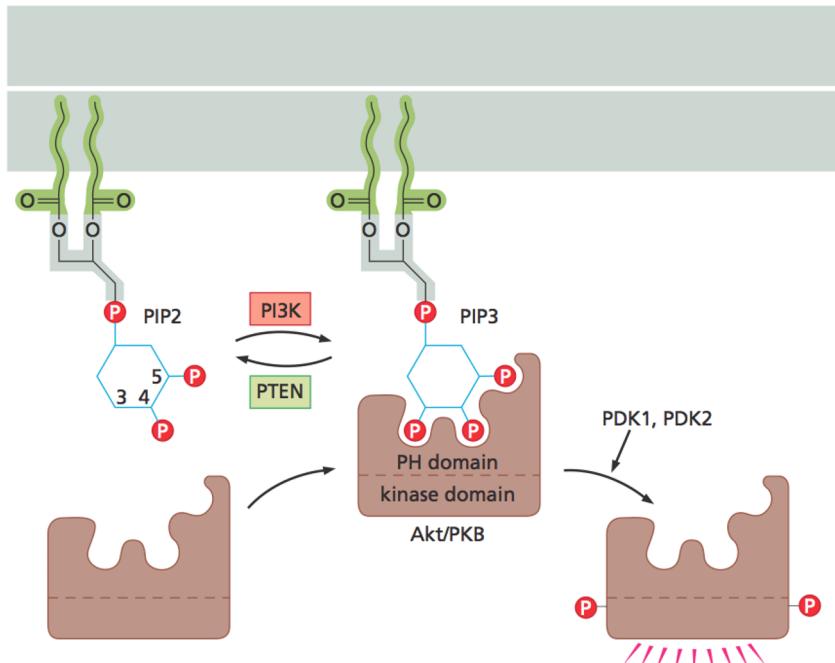
The molecular taxonomy of primary prostate cancer, Cell, 2015

# Unifying pathways



Biology of Cancer, Weinberg

# Unifying pathways – an example

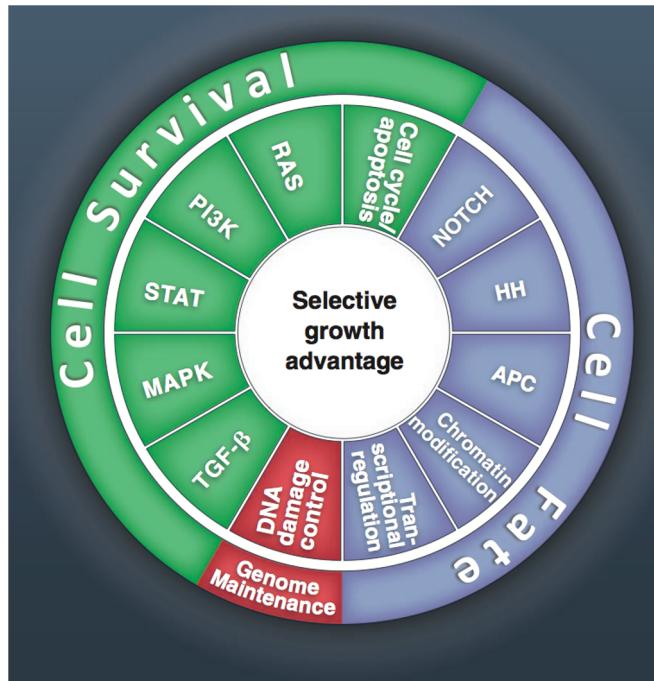


**PI3K** = Oncogene. A gene when activated by mutation gives a growth advantage

**PTEN** = Tumor suppressor gene. A gene when inactivated by mutation gives a growth advantage

# Unifying pathways

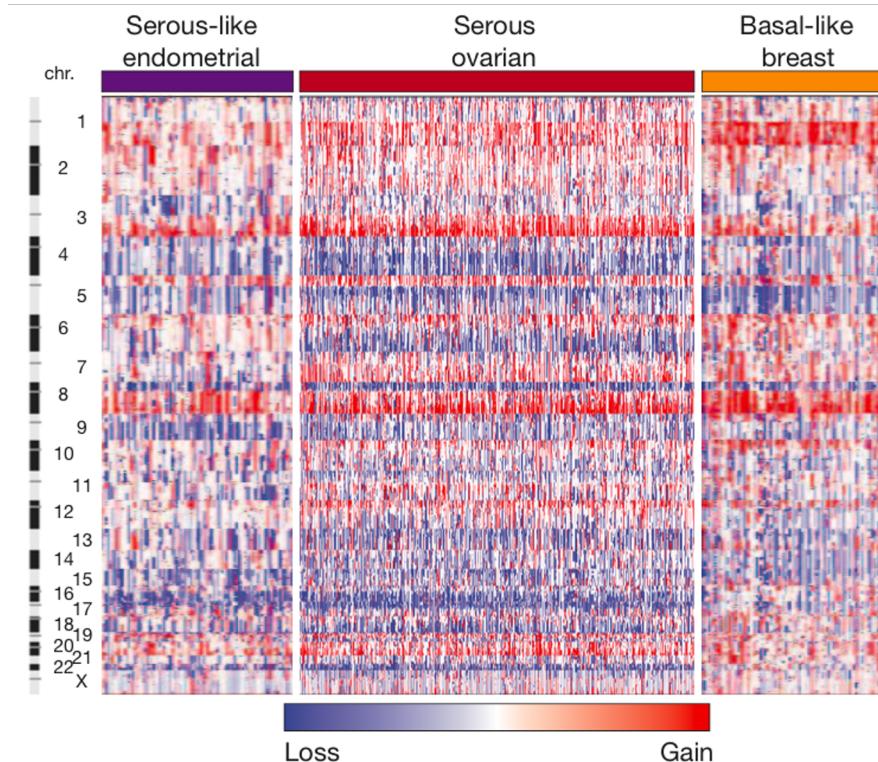
- Five to eight drivers mutated in a limited number of pathways for each solid tumors
  - Cause selective growth advantages



Cancer Genome Landscapes, Science 2013

# Inter disease homogeneity

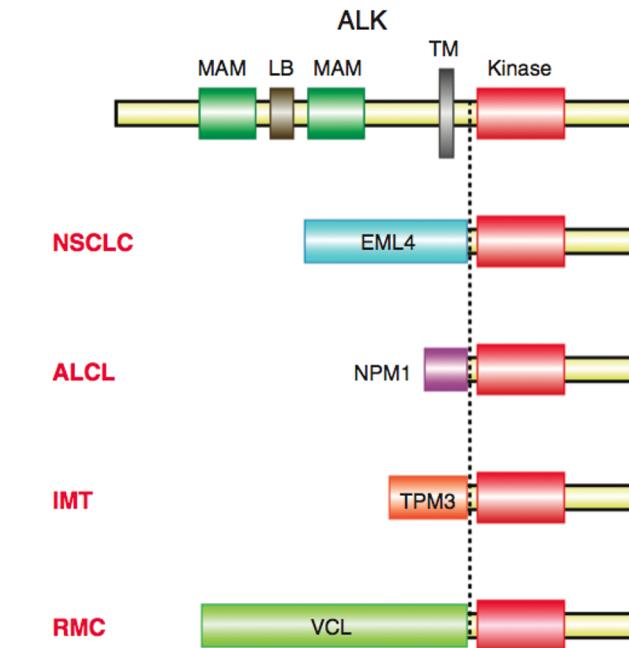
- Almost all cancer drivers are mutated in all cancers BUT at different frequencies
- TP53-mutated cancer signature the same in multiple organs
- Similarities suggest pan-cancer treatment possibilities
- Stage-shift from organ-based to genomic-based treatment rationales?



Integrated genomic characterization of endometrial carcinoma, TCGA, Nature 2013

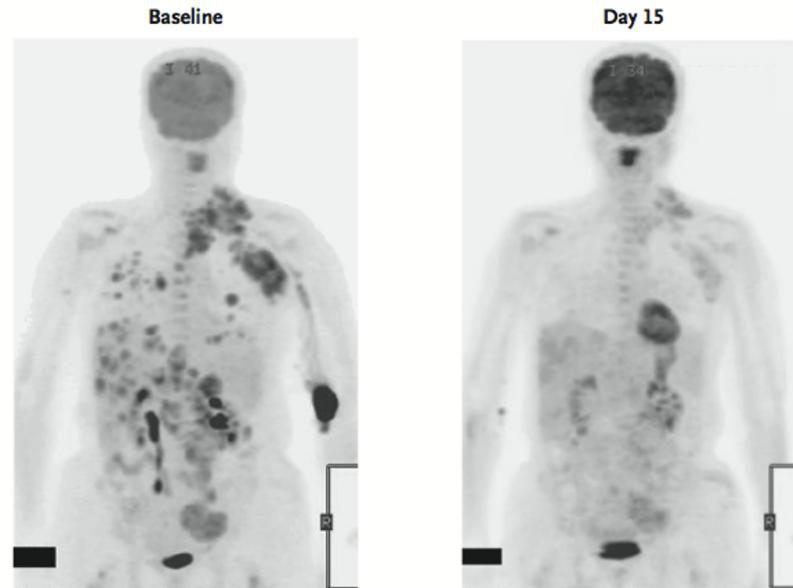
# Inter disease homogeneity

- ALK (Anaplastic lymphoma kinase)
  - Protein tyrosine kinase acting as driver in multiple cancers of different origin = Alkomas
  - Inhibitors are effective across a range of different histologies



ALKoma: A Cancer Subtype with a Shared Target, CD, 2012

# Inter disease heterogeneity

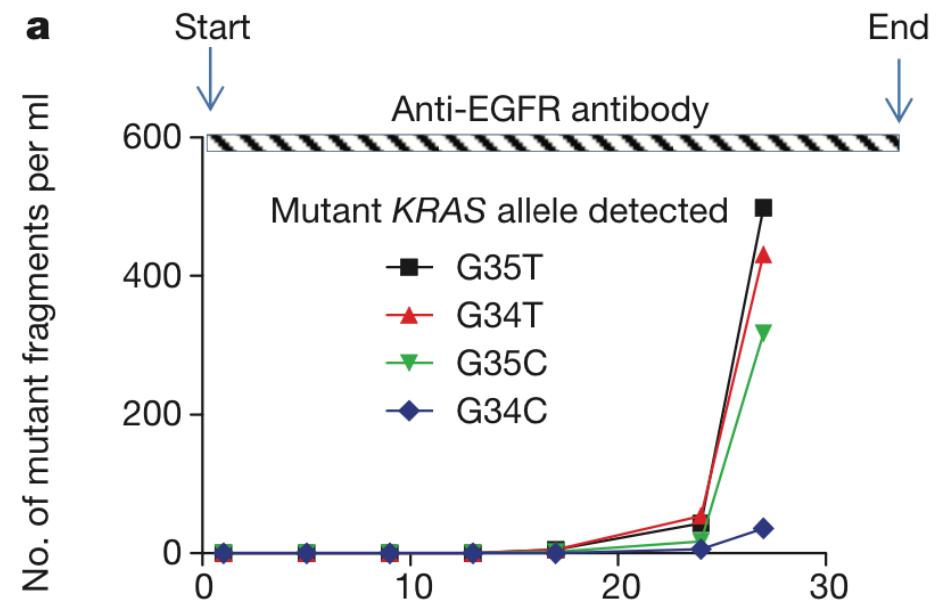


- Vemurafenib is highly effective in BRAF V600E mutated melanoma but not colorectal cancer
- Each drug – diagnosis combination need to be evaluated in a clinical trial

Inhibition of Mutated, Activated BRAF in Metastatic Melanoma, NEJM, 2010  
PLX4032 in metastatic colon cancer patients with mutant BRAF tumors, JCM 2010

# Why targeted mono-therapy fail in metastatic disease

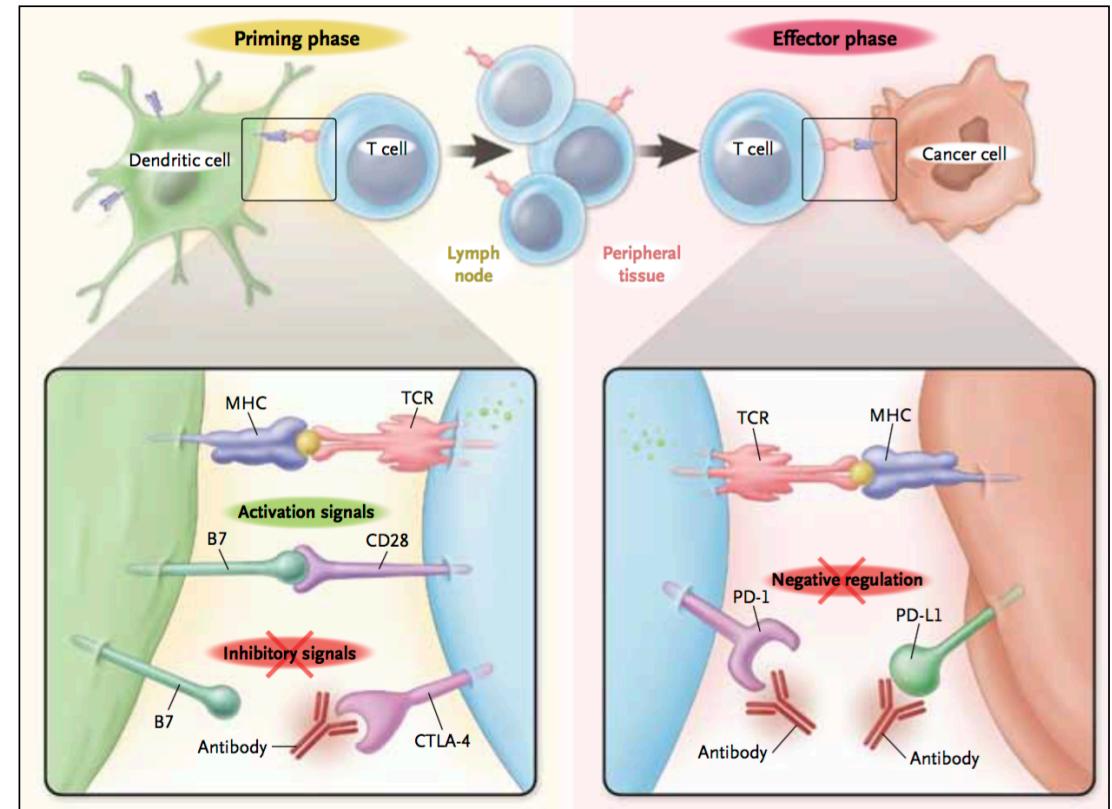
- Metastatic patients basically always contain resistant clones to targeted monotherapy
  - Inevitable due to random errors during cell-division
  - Combinations the way forward hitting on orthogonal functions.
    - Side effects ...



The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers, Nature, 2012  
 Evolutionary dynamics of cancer in response to targeted combination therapy, ELife, 2013

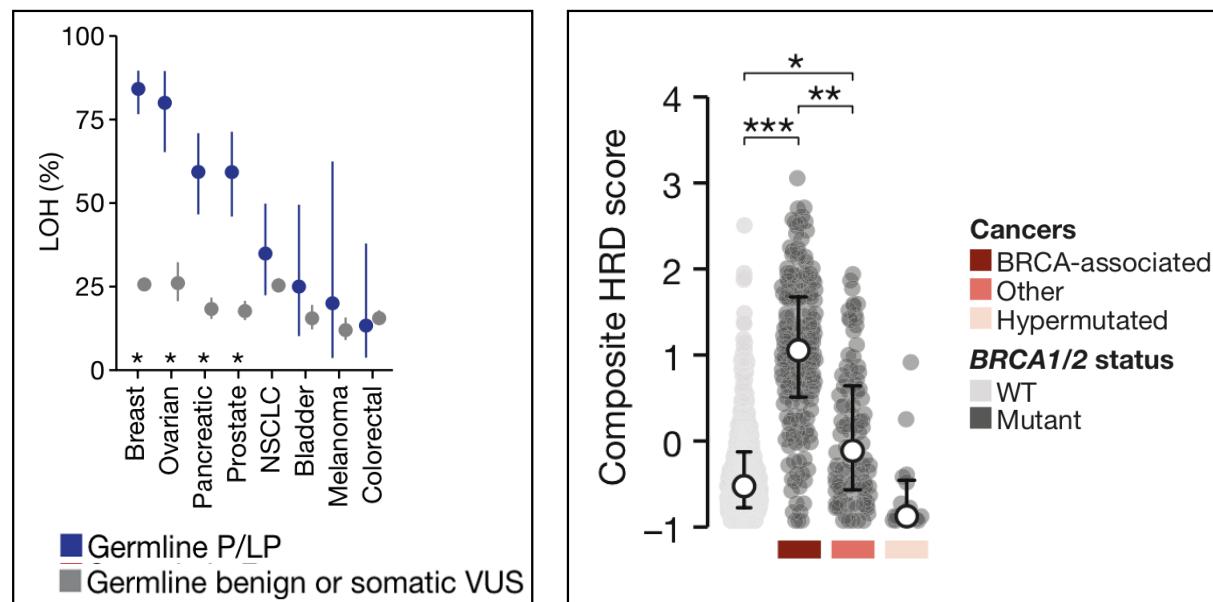
# Why targeted mono-therapy fail in metastatic disease

- Immune-modulating antibodies seem to be an exception

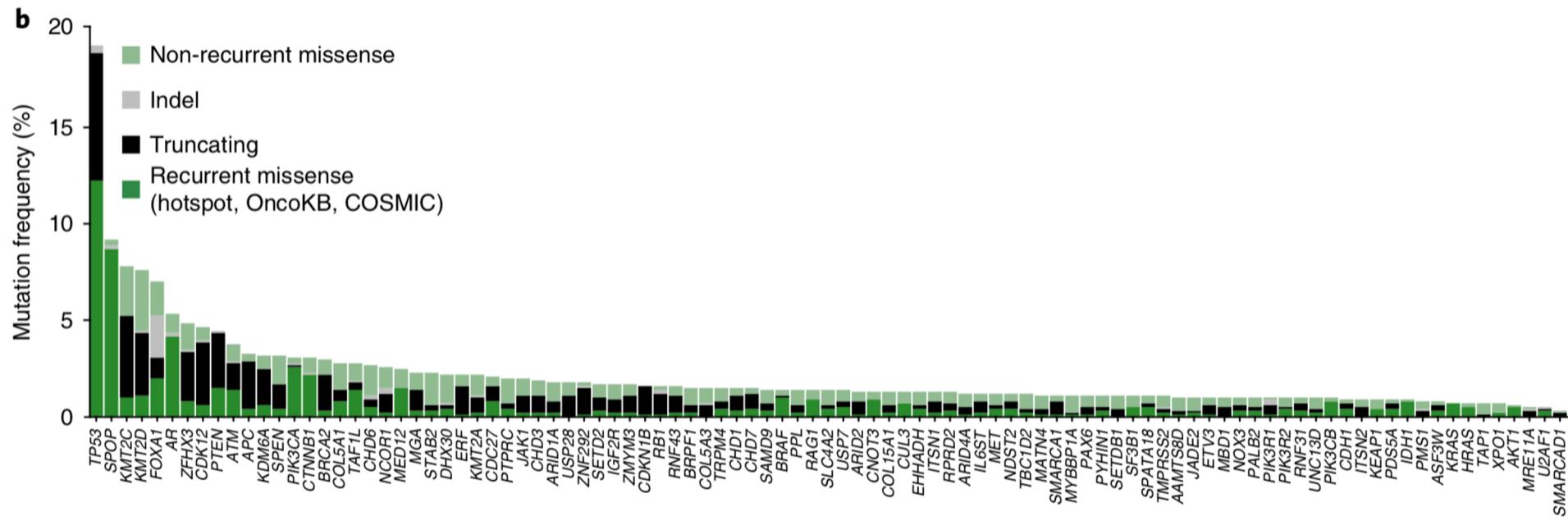


# Inter disease heterogeneity

- BRCA1/2 only relevant in cancer with increased risk of heritable cancer
  - Selective pressure for biallelic inactivation
  - Sensitivity to PARP inhibition

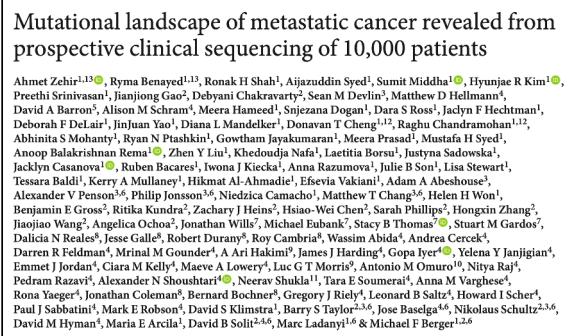


# The long tail – a prostate cancer example

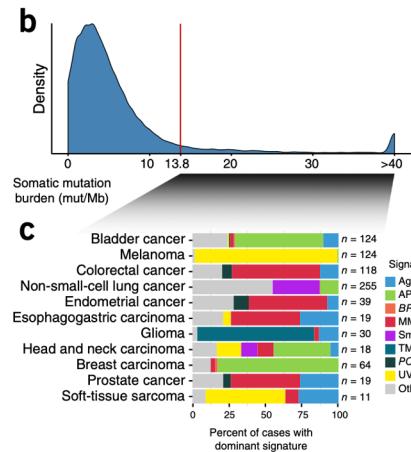
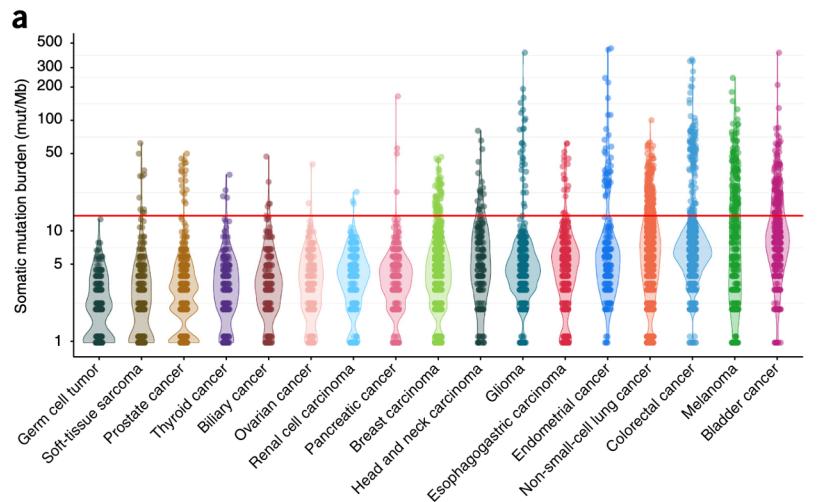


The long tail of oncogenic drivers in prostate cancer, Nature Genetics, 2018

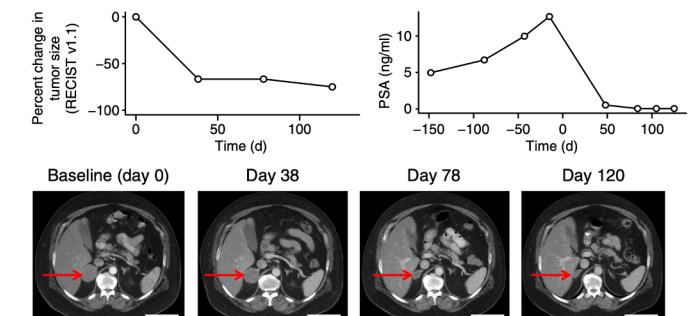
# The long tail – a cancer phenotype example



- ~400 genes interrogated in 10,000 patients
- Detection of relevant variants in 9% with low tumor purity
- 11% enrolled in genetically matched clinical trials
- Detection of unexpected gene fusions (e.g. BRAF in prostate cancer)
- Detection of hypermutated cancers



Metastatic prostate cancer with MSI+ treated with immunomodulators



Sum(the long tail of actionability) – can be a large fraction of patients

---

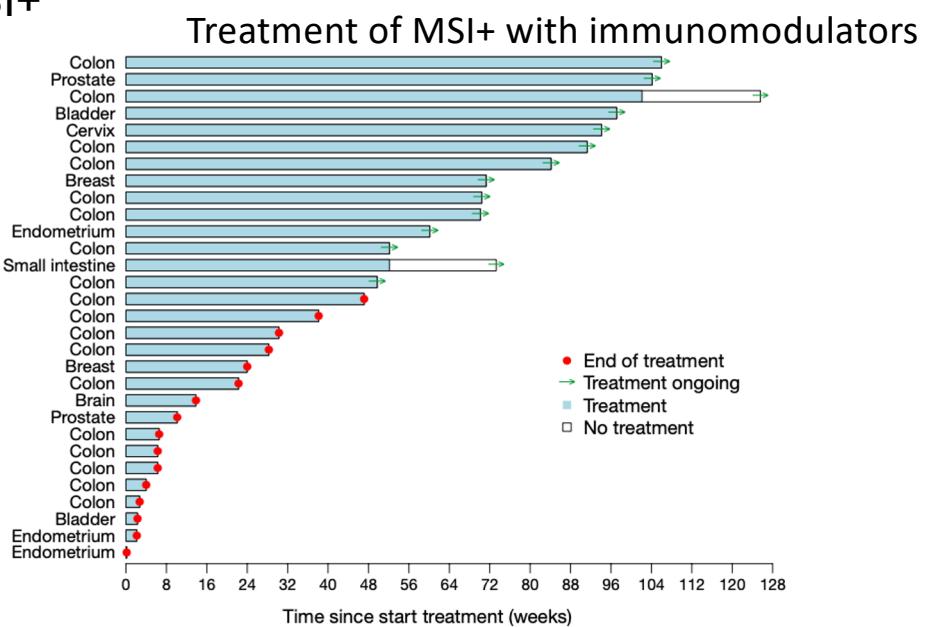
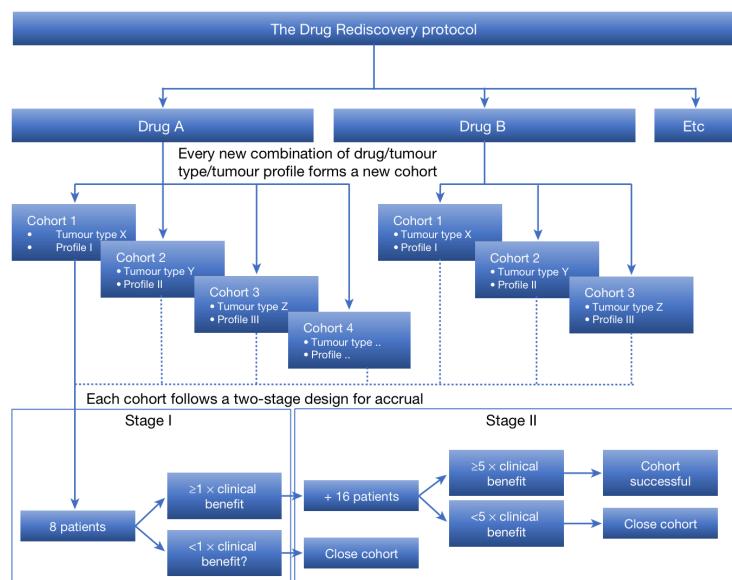
- Metastatic prostate cancer
  - Increased neoantigens on the cell surface
    - Potential drug: Immunomodulator e.g. pembrolizumab
    - MSI+/TMB-H/Tandem duplication phenotype: ~10%
  - Homologous recombination deficiency
    - Potential drug: Parp inhibitor e.g. Niraparib
    - BRCA-complex genes only: ~10%

- Summary

- Two cancers of the same histological origin may have no common somatic denominator.
- Not total chaos – a limited number of pathways are affected.
- Cancers with different histological origin but common drivers may have genomic similarities providing a rationale for drug repurposing.
- Drug repurposing does not always work - the concept of pan-can treatment needs to be evaluated for each drug target combination x histological origin.
- The concept of the long tail in terms of both
  - Mutated driver genes
  - Cancer phenotypes

## Drug repurposing – being evaluated in the DRUP trial

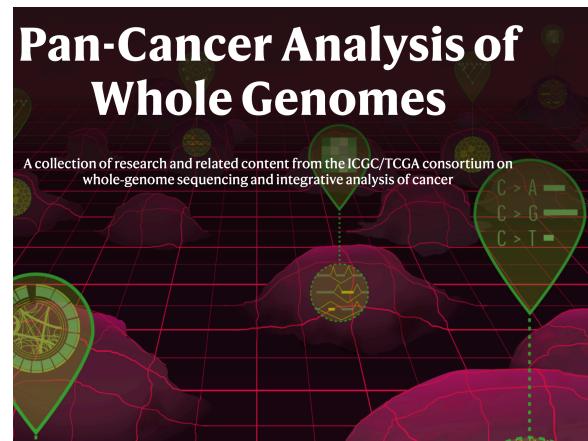
- Sponsored by the Hartwig Foundation
- The Center for Personalized Cancer Treatment
  - 45 hospitals in the Netherlands, 219 patients reported
- Structured evaluation of off-target drugs in collaboration with drug companies
- First paper with a focus on pan-cancer MSI+



The Drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs, Nature 2019

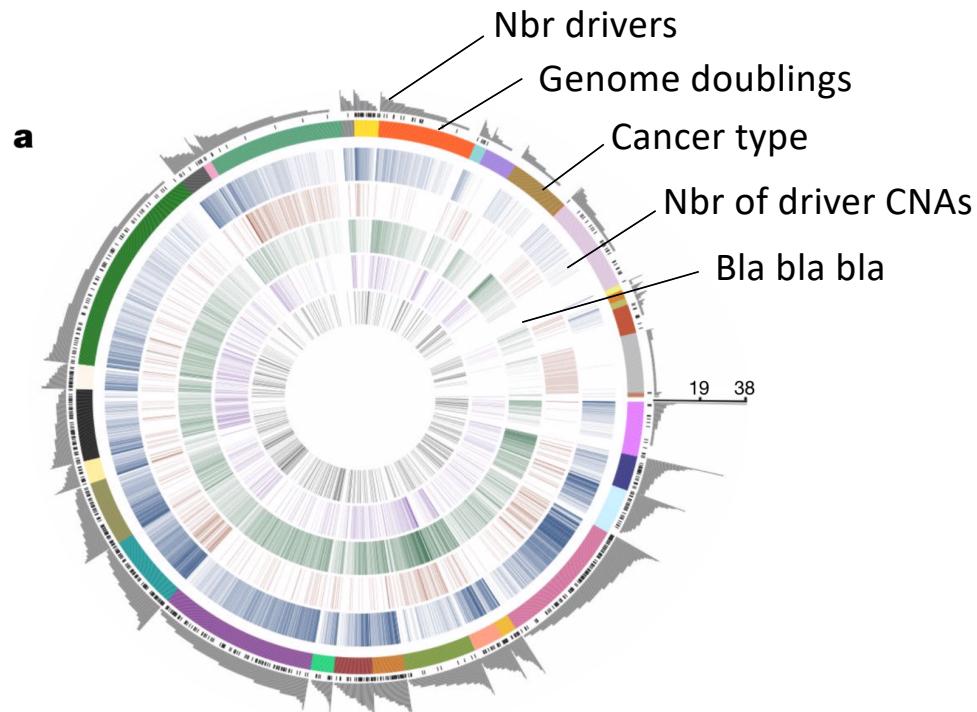
## Flagship paper from the ICGC

- Whole genome sequencing (wgs) of 2,605 primary tumours and 173 metastases or local recurrences
- RNA-sequencing data were available for 1,222 donors
- Mean age: 56 years
- Largest data set so far with “no compromises”

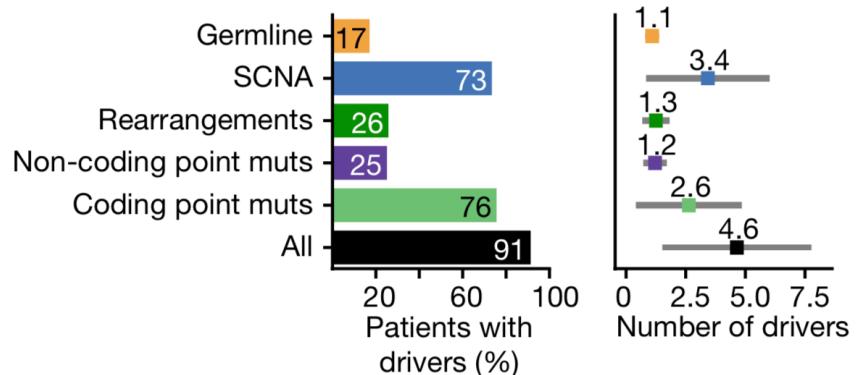


# Flagship paper from the ICGC – driver distribution

Circos plot: The most overused plot in genomics

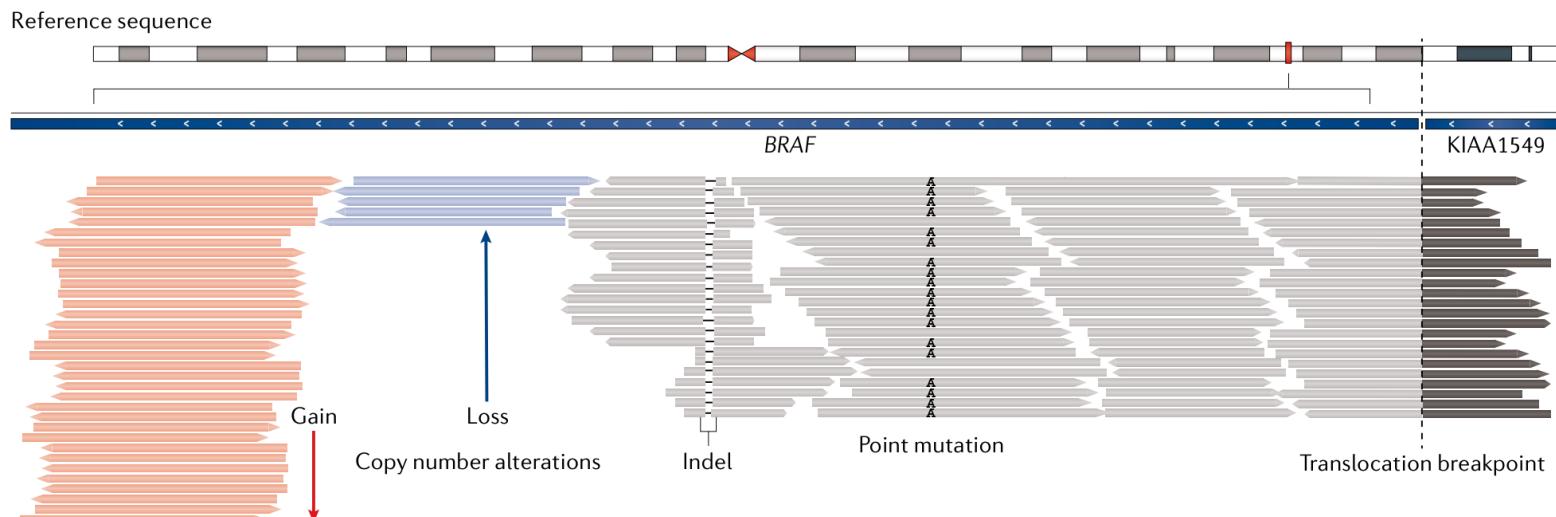


Keep it simple!



## Flagship paper from the ICGC – driver distribution

- Description of all types of drivers
- Distribution and number per case
- Individual types discussed later in the course in conjunction to practical sessions



Clinical cancer genomic profiling, Nat Rev Gen 2021

## Flagship paper from the ICGC – driver distribution

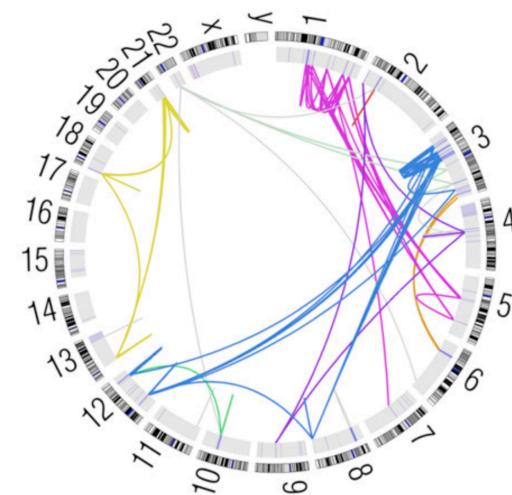
---

- Driver types overrepresented in some cancer types, examples:
  - Breast cancer:
    - Structural variants, mean nbr: 6.4, sd:  $\pm$  3.7
    - Point mutations, mean nbr: 2.2, sd  $\pm$  1.3
    - $P < 1 \times 10^{-16}$ , Mann–Whitney *U*-test)
  - Colorectal cancer
    - Structural variants, mean nbr: 2.4, sd:  $\pm$  1.4
    - Point mutations, mean nbr: 7.4, sd  $\pm$  7.0
    - $P < 1 \times 10^{-16}$ , Mann–Whitney *U*-test)

## Flagship paper from the ICGC – frequency of driver-causing mechanisms

---

- Overview of mechanisms that cause somatic alterations
  - Chromoplexy
    - Repair of co-occurring double-stranded DNA breaks are glued together by the DNA repair machinery to create shuffled chains of rearrangements
    - 17.8% of all cases
  - Kataegis
  - Chromotripsy
  - etc
- Known from before ...



Mutational signature paper from the ICGC – no individual lecture



- Somatic mutations are caused by both exogenous and endogenous processes.
    - DNA repair, smoking, sunlight, ageing, chemotherapy etc
  - Mathematical methods have been developed to investigate mutation data to determine the mutational signatures from a cancer genome and its origin.

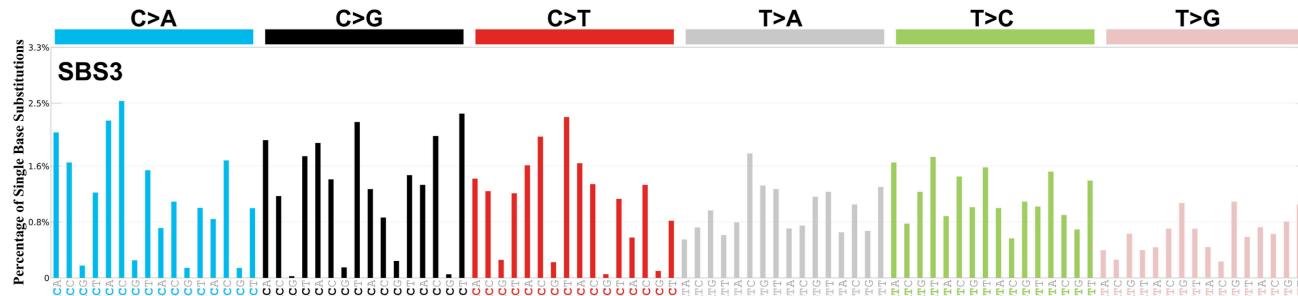
Statistically significant mutational signature x Cancer type association x Known explaining mutagenic process?  
Associated with treatment outcome?

The repertoire of mutational signatures in human cancer, Nature 2020

# Mutational signature paper from the ICGC



- Single base substitutions (SBS), an example
    - Input data, possible mutations: C>A, C>G, C>T, T>A, T>C and T>G
  - Account for 5' and 3' base, leads to 96 possibilities.
  - Account for transcribed or untranscribed strand
    - 192 possibilites

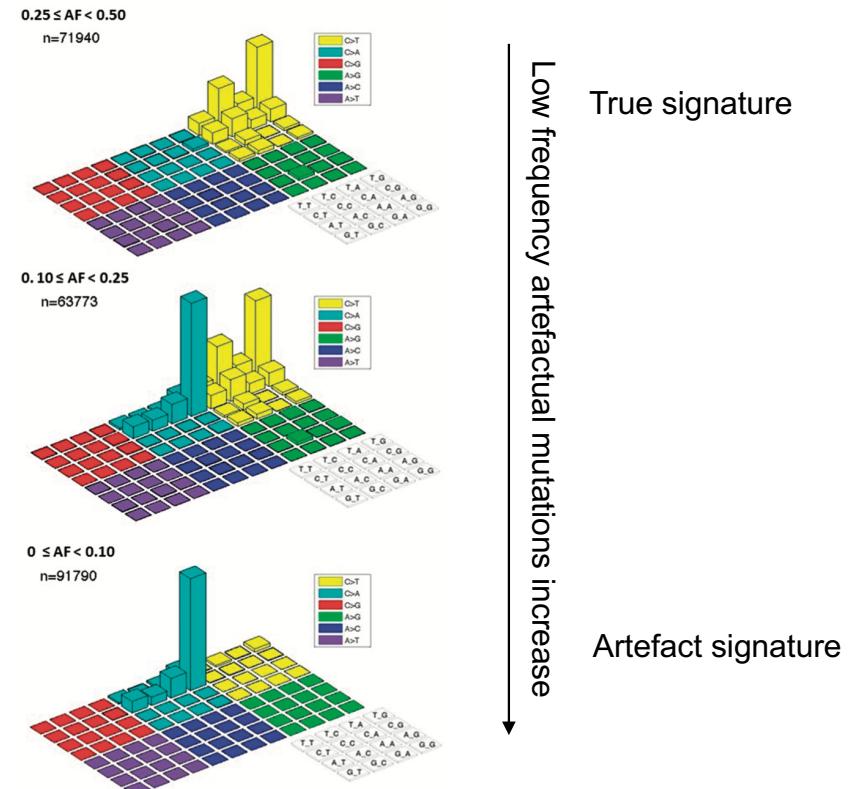


SBS3: signature associated with homologous recombination deficiency and sensitivity to parp inhibitors and carboplatin.

The repertoire of mutational signatures in human cancer, Nature 2020

## Mutational signature paper from the ICGC

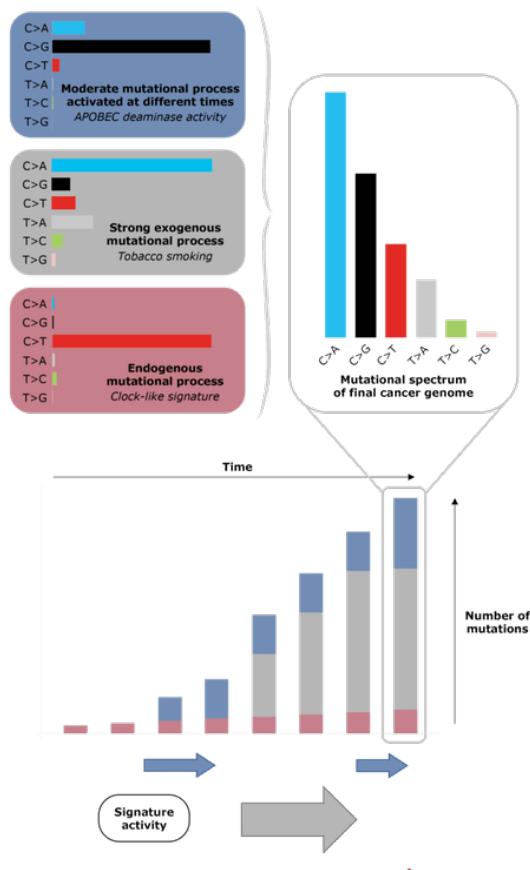
- At the Broad Institute noise variants were detected in the TCGA whole exome data
- Another way of displaying the 96 options →
- Similarly done for
  - Indels
  - Double base substitutions
  - Vary flanking bases
  - Etc ...



Discovery and characterization of artefactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation, NAR, 2013

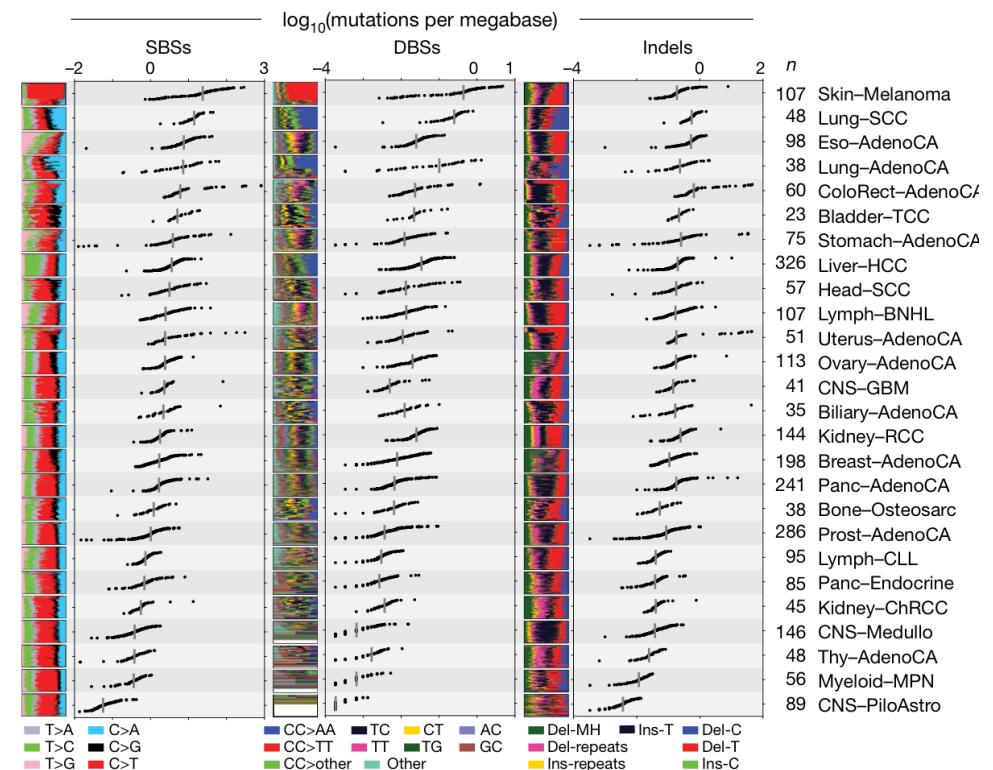
# Mutational signature paper from the ICGC

Identify the mutational causing processes for an individual cancer



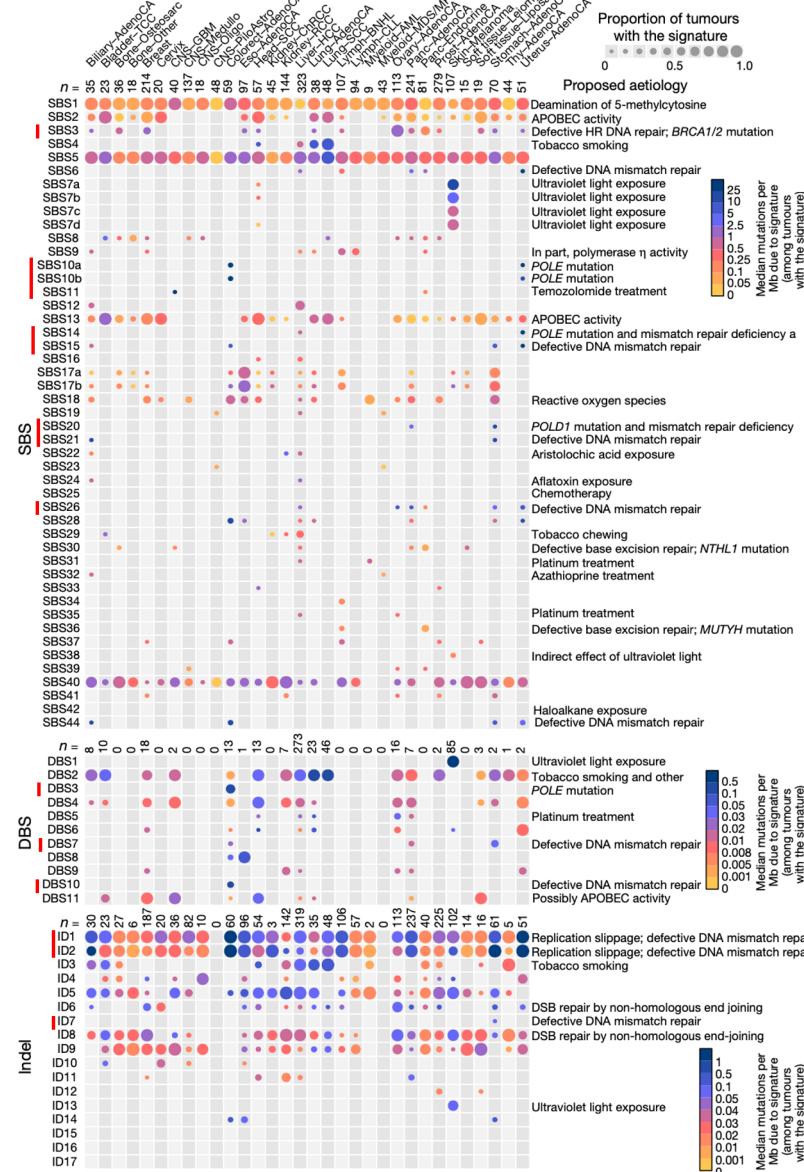
<https://cancer.sanger.ac.uk/signatures/>

Reflected in the mutation rates of different cancers



The repertoire of mutational signatures in human cancer, Nature 2020

# Mutational signatures



Treatment relevant mutational signatures, e.g. homologous recombination deficiency (SBS3).

An argument for WGS/WES.

Panel sequencing is often 5x – 50x smaller than WES

Hard to robustly assess mutational signatures.

If time allows you will get to upload mutations to an online portal to see if you can determine the relevant signatures for some cancers.

Questions?

---