Bioinformatics for (clinical) interpretation of cancer genomics

David Tamborero, PhD
Senior researcher
Karolinska Institutet, SciLifeLab

> *Trained as engineer (~2004)*
> *PhD in Medicine (HC-UB – cardiology department) (~2009)*
> *Master in **bioinformatics** & molecular biology (~2013)*

IRB BARCELONA
INSTITUTE FOR RESEARCH IN BIOMEDICINE

Barcelona

upf. Universitat Pompeu Fabra Barcelona

FIMM
Institute for Molecular Medicine Finland
Nordic EMBL Partnership for Molecular Medicine

Helsinki

Stockholm

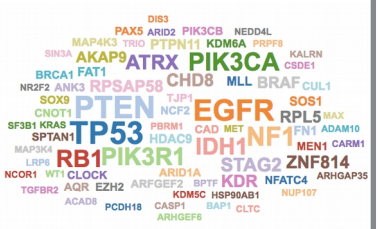
SciLifeLab

KAROLINSKA INSTITUTET ANNO 1810
Karolinska Institutet

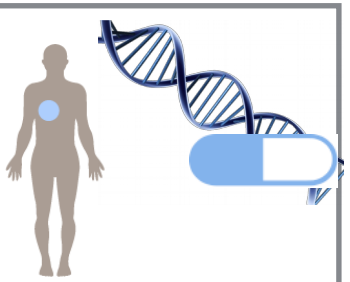
LehtiöLab

(Bioinformatics) study of tumor genomes

Understanding mutational processes



Finding drivers of cancer

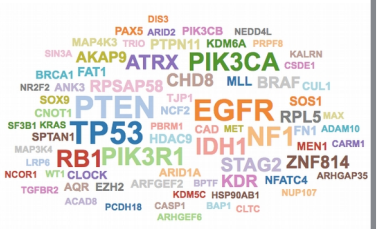


Precision cancer medicine

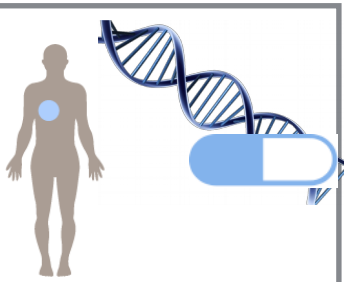
Bioinformatics enables testing hypotheses at an **unprecedent scale** by using computational methods across **large datasets**

(Bioinformatics) study of tumor genomes

Understanding mutational processes



Finding drivers of cancer



Precision cancer medicine

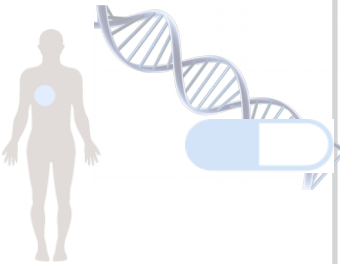
findings in one supports the other!

Bioinformatics enables testing hypotheses at an **unprecedent scale** by using computational methods across **large datasets**

(Bioinformatics) study of tumor genomes

Understanding mutational processes

Finding drivers of cancer

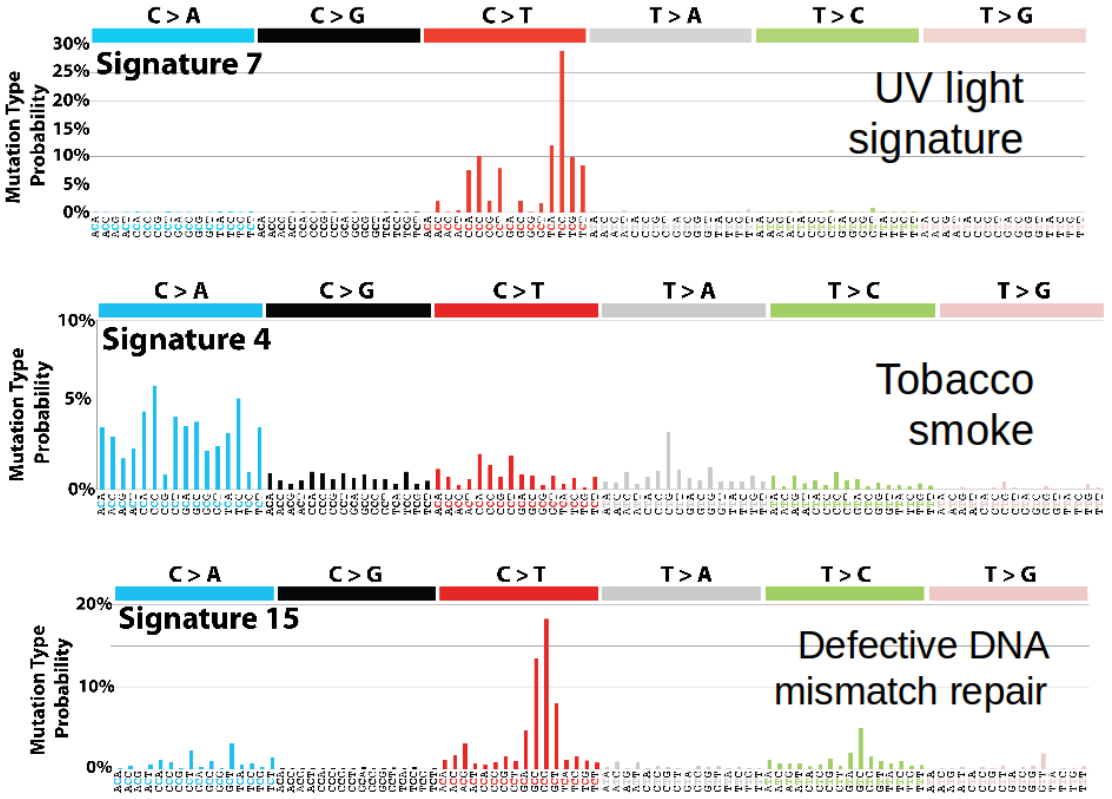


Precision cancer medicine

Marc Rosenthal

Identifying patterns in the mutations observed across cancer cohorts provides knowledge about the biology of the (tumor) cell
(e.g. ***mutational processes****)*

Processes occurring ~at the whole **genome scale**:

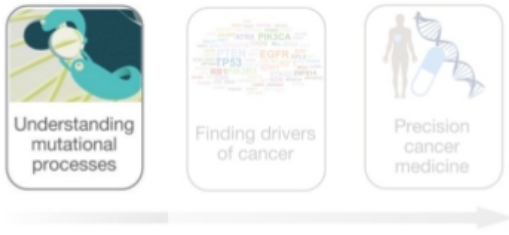


COSMIC website

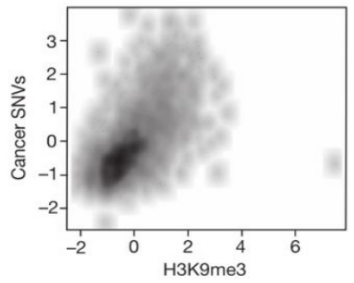
Tumor mutation signatures

Obtained by de-convoluting **unique, distinct patterns** of mutations across tumor cohorts

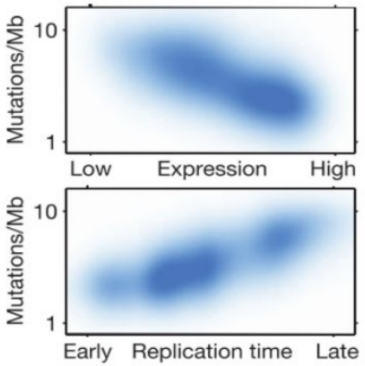
They are associated to specific **biological processes/ environmental "aggressions"** acting through the **~entire genome**



Processes constrained at the **megabase scale:**



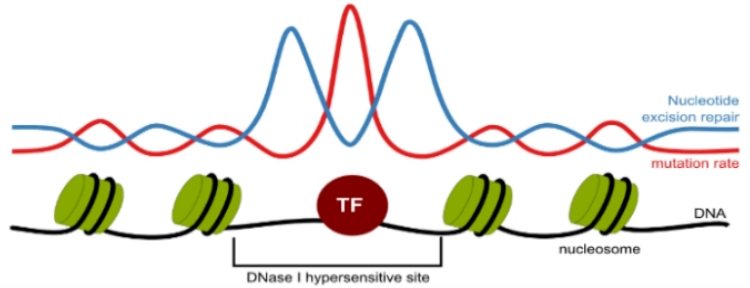
Schuster-Böckler & Lehner. *Nature*. 2012



Lawrence *et al. Nature*. 2013

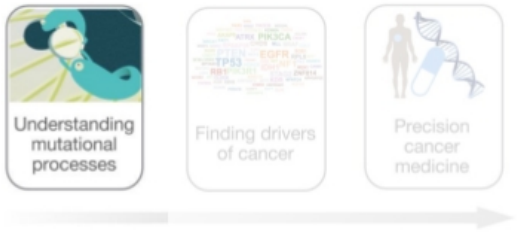
More mutations in regions with **closed-chromatin** conformations and **late-replicated** regions

Processes constrained to **(few) nucleotides** scale:

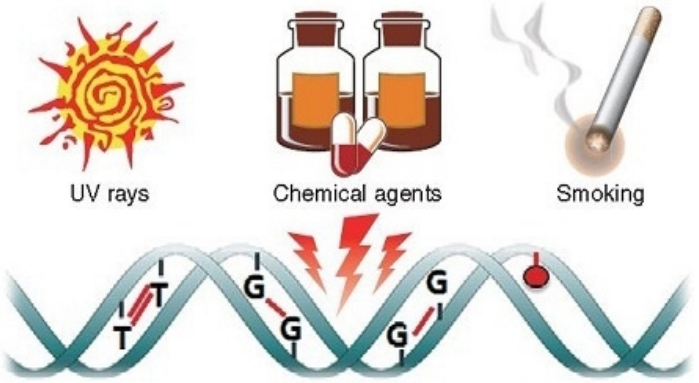


Sabarinathan et al. Nature 2016

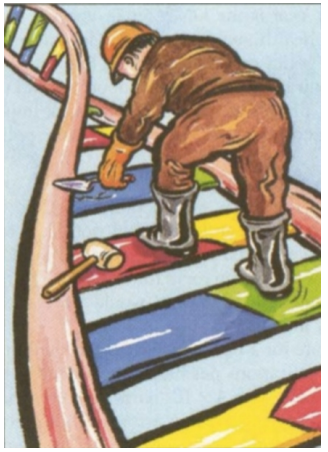
More mutations in **transcription factor binding sites** of melanocytes *(due to interference with NER proteins)*



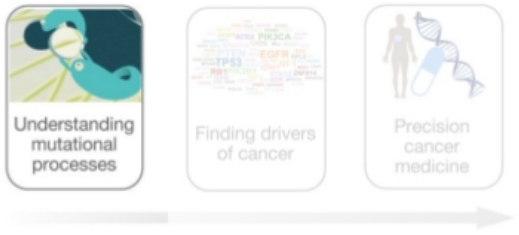
Bioinformatics analyses of large data sets has contributed to a better understanding on how the **mutagenic processes** vs. **mechanisms of DNA repair** shapes the landscape of (somatic) mutations in (tumor) cells



Mutagenic processes (intrinsic/extrinsic)



Mechanisms of DNA repair

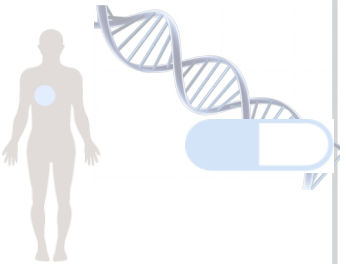


Adapted from Lopez-Bigas

(Bioinformatics) study of tumor genomes

Understanding mutational processes

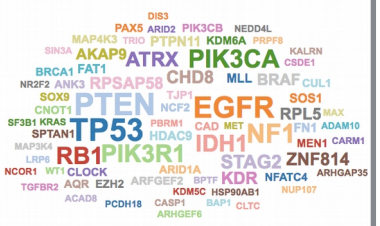
Finding drivers of cancer



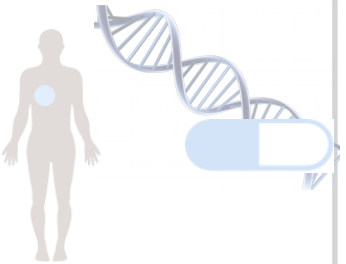
Precision cancer medicine

(Bioinformatics) study of tumor genomes

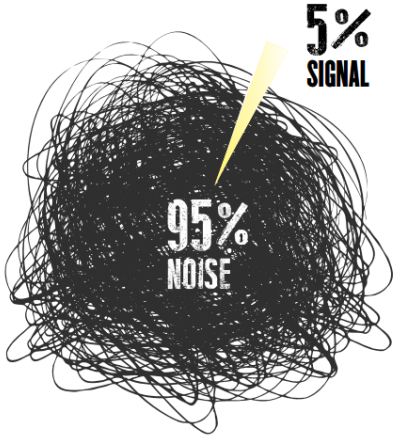
Understanding mutational processes



Finding drivers of cancer



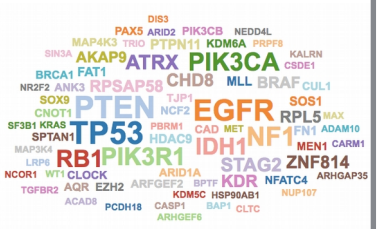
Precision cancer medicine



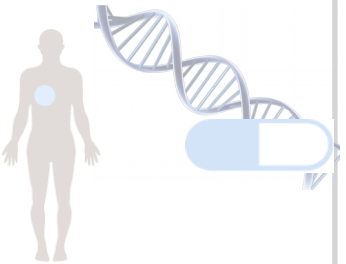
Most of the tumor mutations are a **consequence** of the tumorigenic processes, while only a minority are **causing** the disease

(Bioinformatics) study of tumor genomes

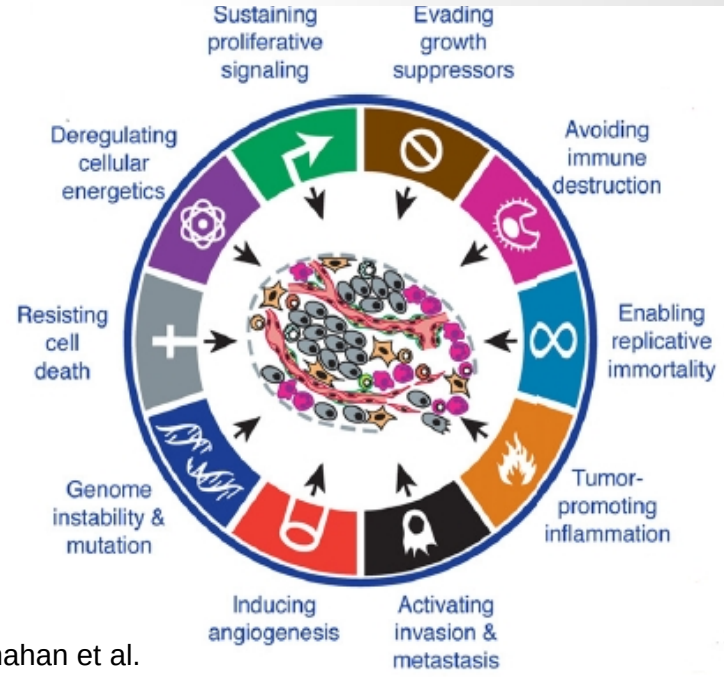
Understanding mutational processes



Finding drivers of cancer



Precision cancer medicine

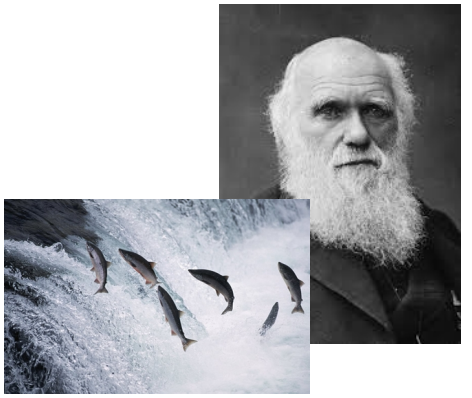


Hanahan et al.

How <u>bionformatic</u> analyses can support the identification of molecular alterations that **<u>drive</u>** the cancer hallmarks **?**

(vs passengers)

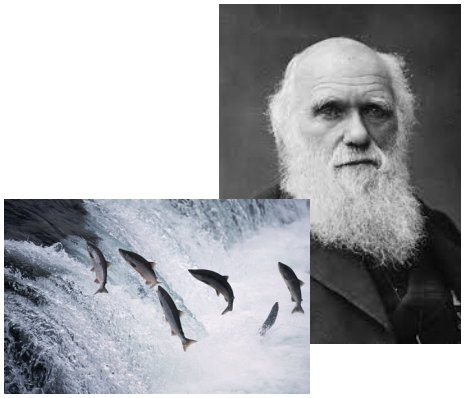
Cancer is an evolutionary process



Mutations are fixed during rounds of clonal expansion **if** they confer a **selective advantage** to the tumor cells (*~Darwinian process*)

Bioinfo analyses aim to detect the **signals of positive selection** in the mutations that occur in **driver genes** → i.e. those mutations that deviate from a **random* distribution**

Cancer is an evolutionary process

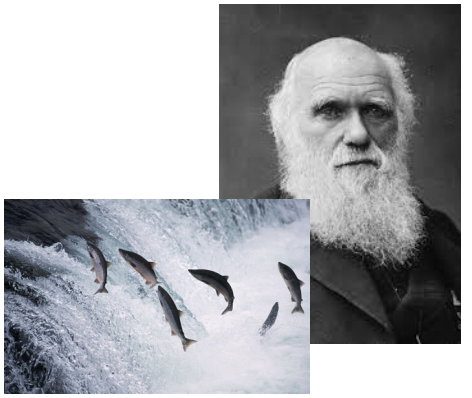


Mutations are fixed during rounds of clonal expansion **if** they confer a **selective advantage** to the tumor cells (*~Darwinian process*)

Bioinfo analyses aim to detect the **signals of positive selection** in the mutations that occur in **driver genes** → i.e. those mutations that deviate from a **random* distribution**

random = as estimated by a **background model*

Cancer is an evolutionary process



Mutations are fixed during rounds of clonal expansion **if** they confer a **selective advantage** to the tumor cells (*~Darwinian process*)

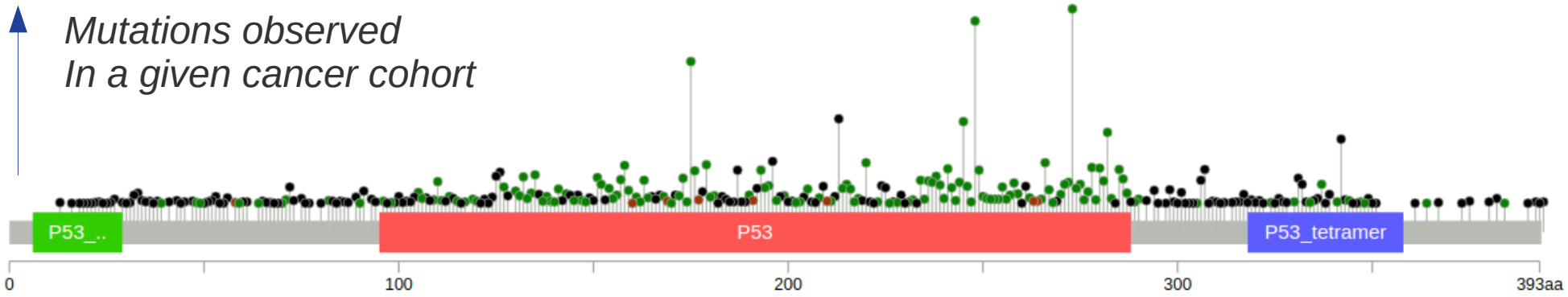
Bioinfo analyses aim to detect the **signals of positive selection** in the mutations that occur in **driver genes** → i.e. those mutations that deviate from a **random* distribution**

random = as estimated by a **background model*

Based on the knowledge about the 'random pattern of mutations' in tumor cells

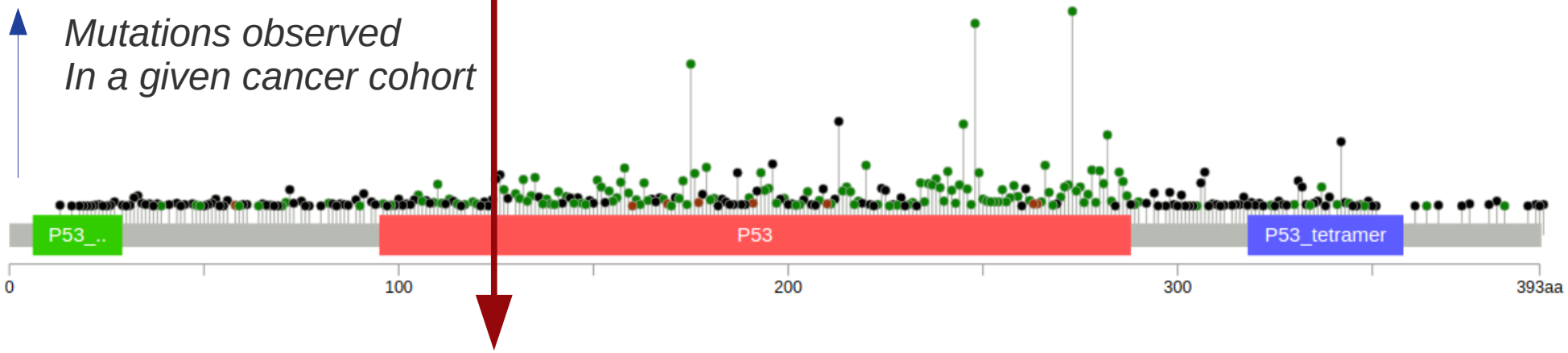
Hotspots: protein sites accumulating more somatic mutations than expected by the background processes

Hotspots point out mutated protein sites that are **positively selected** by **tumor cells** (of a given cancer type)

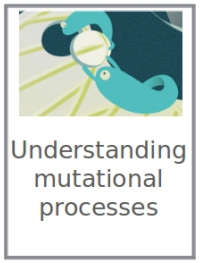


Hotspots: protein sites accumulating more somatic mutations than expected by the background processes

Hotspots point out mutated protein sites that are **positively selected** by **tumor cells** (of a given cancer type)

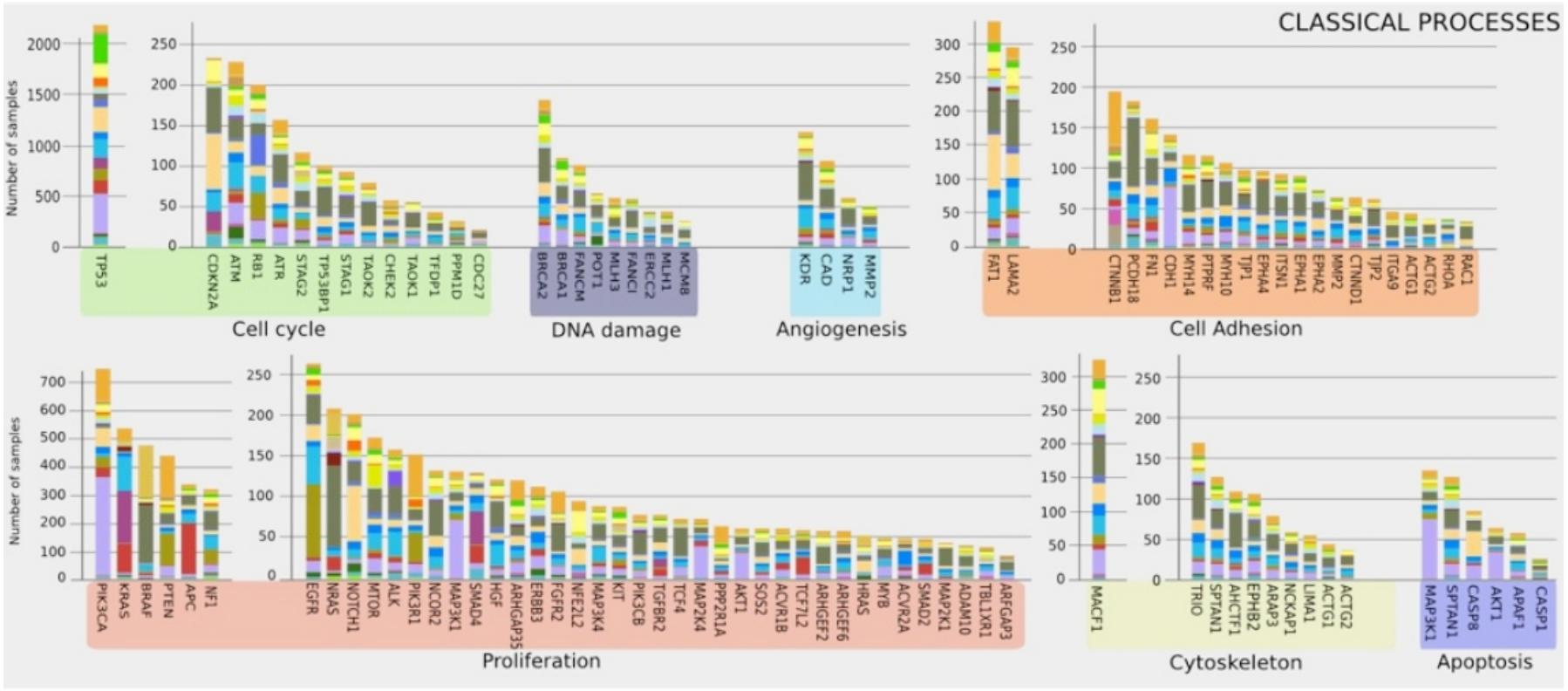


Whether a specific protein site accumulates more mutations than ***expected by chance**** *is determined by the* ***mutational processes*** *in that* ***genomic region/cancer type***



*(*mutation count is 'only' an observation – hotspot is a statistical claim)*

Catalogs of cancer genes



Perez, Tamborero et al., Cancer Cell 2016

(Bioinformatics) study of tumor genomes

During last years, **bio**informatics analyses leveraged the large amount of available (molecular) cancer data to **provide/support** new insights about the mechanisms of the tumor cell, such as:

> mutational processes

> catalogs of driver genes

> gene mechanisms of action

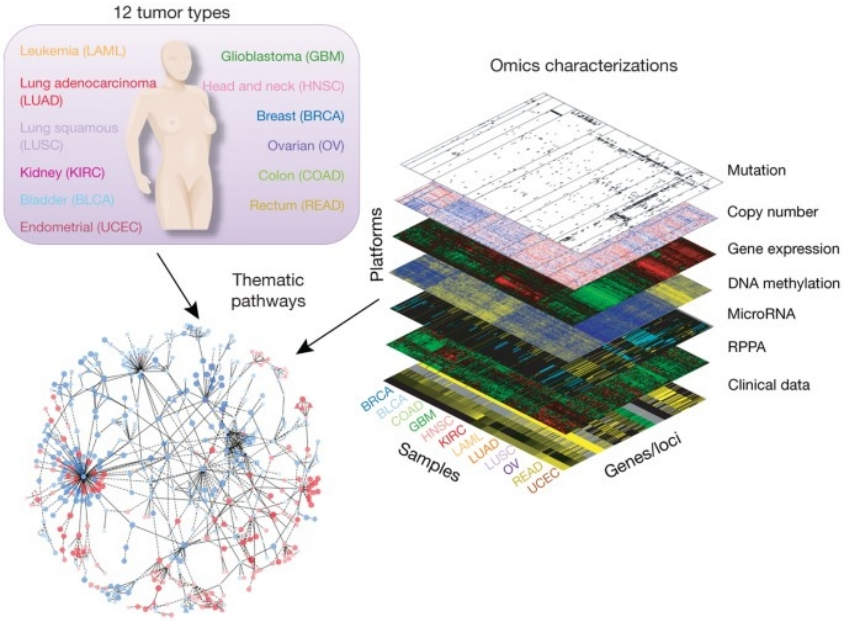
> molecular subtypes of cancer

> tumor evolution in time and space

> interactions with the TME

> identification of tumor vulnerabilities for drug interventions

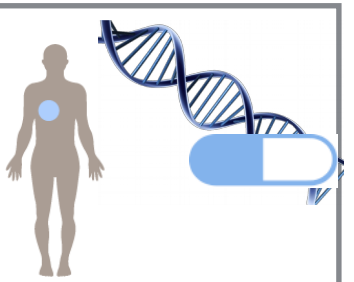
> ...



(Bioinformatics) study of tumor genomes

Understanding mutational processes

Finding drivers of cancer

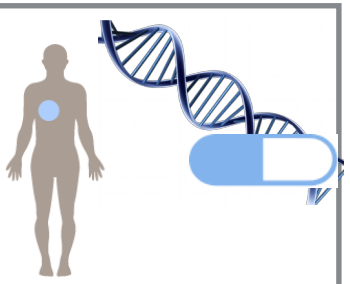


Precision cancer medicine

(Bioinformatics) study of tumor genomes

Understanding mutational processes

Finding drivers of cancer



Precision cancer medicine

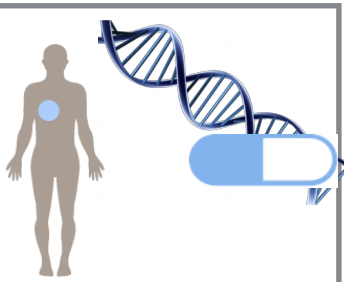
Note that the objective here is not to generate new knowledge but to use the already available knowledge to inform a (prospective) n=1 clinical intervention

From cohorts to individual tumors

(Bioinformatics) study of tumor genomes

Understanding mutational processes

Finding drivers of cancer



Precision cancer medicine

Given the genomic alterations observed in a given patient's tumor:

- which ones are more likely to **drive** that tumor?

- which ones may be therapeutically **actionable**?