

# BIMM 143 – Class 18 Lab

## Mutational Signatures in Cancer



Marcos Díaz-Gay

**UC San Diego**

@ [marcos-diazg.github.io](https://marcos-diazg.github.io)  
Twitter: [@mdiagay](https://twitter.com/mdiagay)  
Alexandrov Lab

# A long journey to San Diego and the States...



UNIVERSIDADE  
DA CORUÑA  
Galicia, Spain



# A long journey to San Diego and the States...



A long journey to San Diego and the States...



... pivoting from civil engineering to computational biology



... pivoting from civil engineering to computational biology

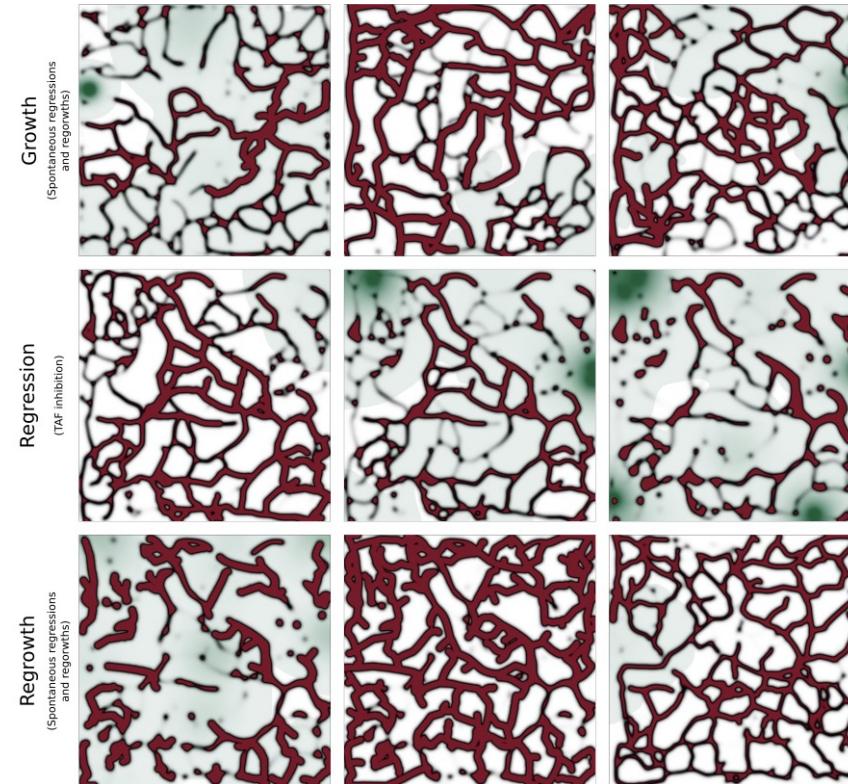
New paradigm of civil engineering

Apply the same mathematical  
framework, but for cancer research



# ... pivoting from civil engineering to computational biology

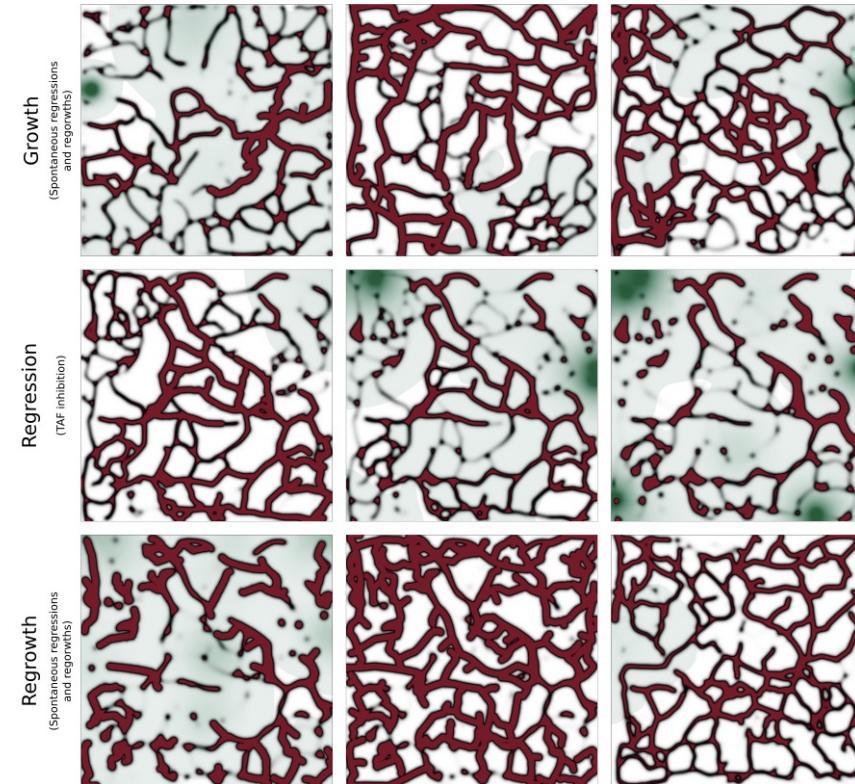
New paradigm of civil engineering  
Apply the same mathematical  
framework, but for cancer research



# ... pivoting from civil engineering to computational biology

New paradigm of civil engineering  
Apply the same mathematical  
framework, but for cancer research

MSc in Biomedicine  
The first step

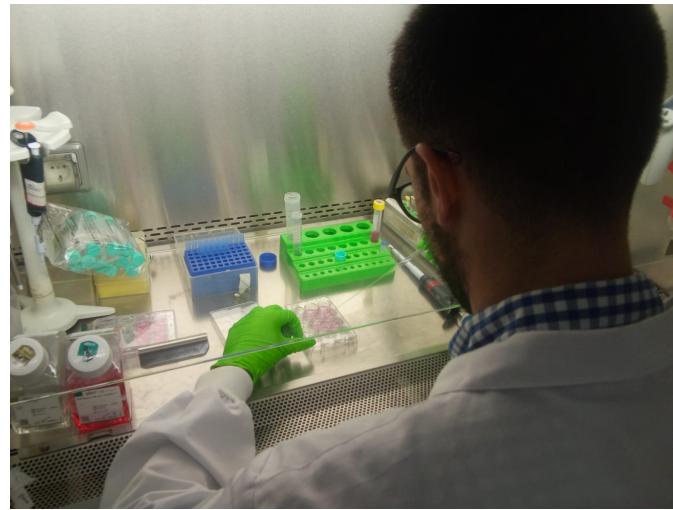


... pivoting from civil engineering to computational biology

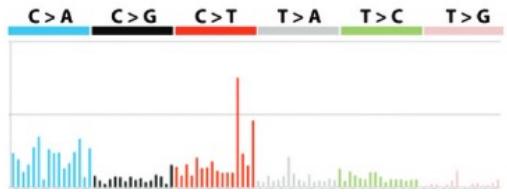
New paradigm of civil engineering  
Apply the same mathematical  
framework, but for cancer research

MSc in Biomedicine  
The first step

PhD in Translational Medicine  
A new biomedical researcher

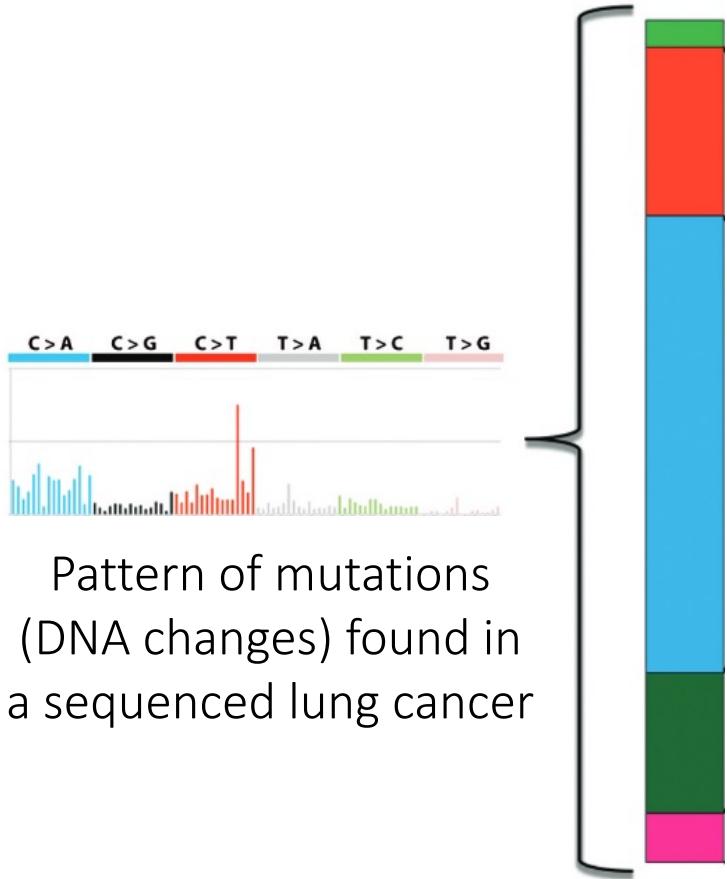


# Spoiler alert!

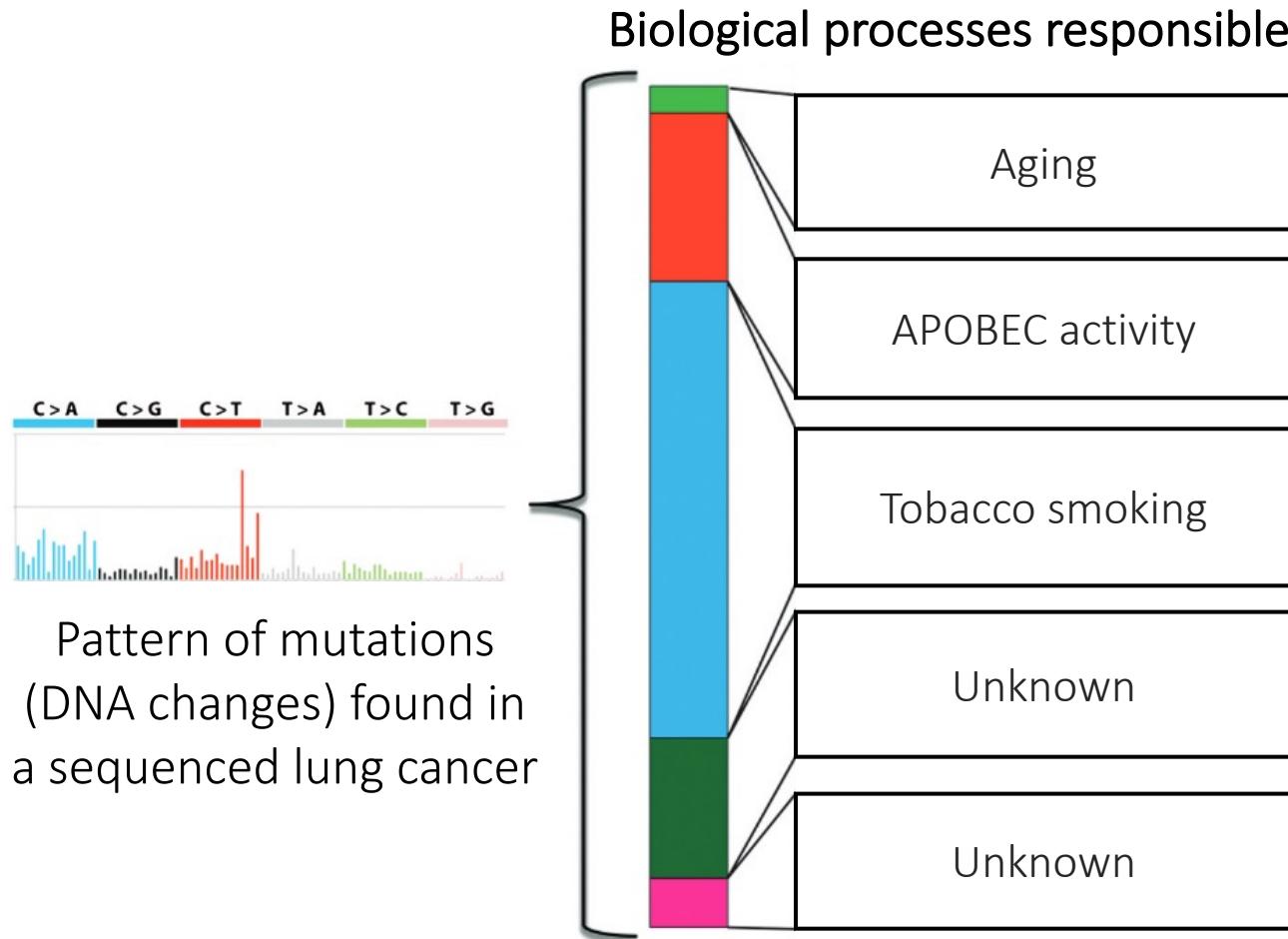


Pattern of mutations  
(DNA changes) found in  
a sequenced lung cancer

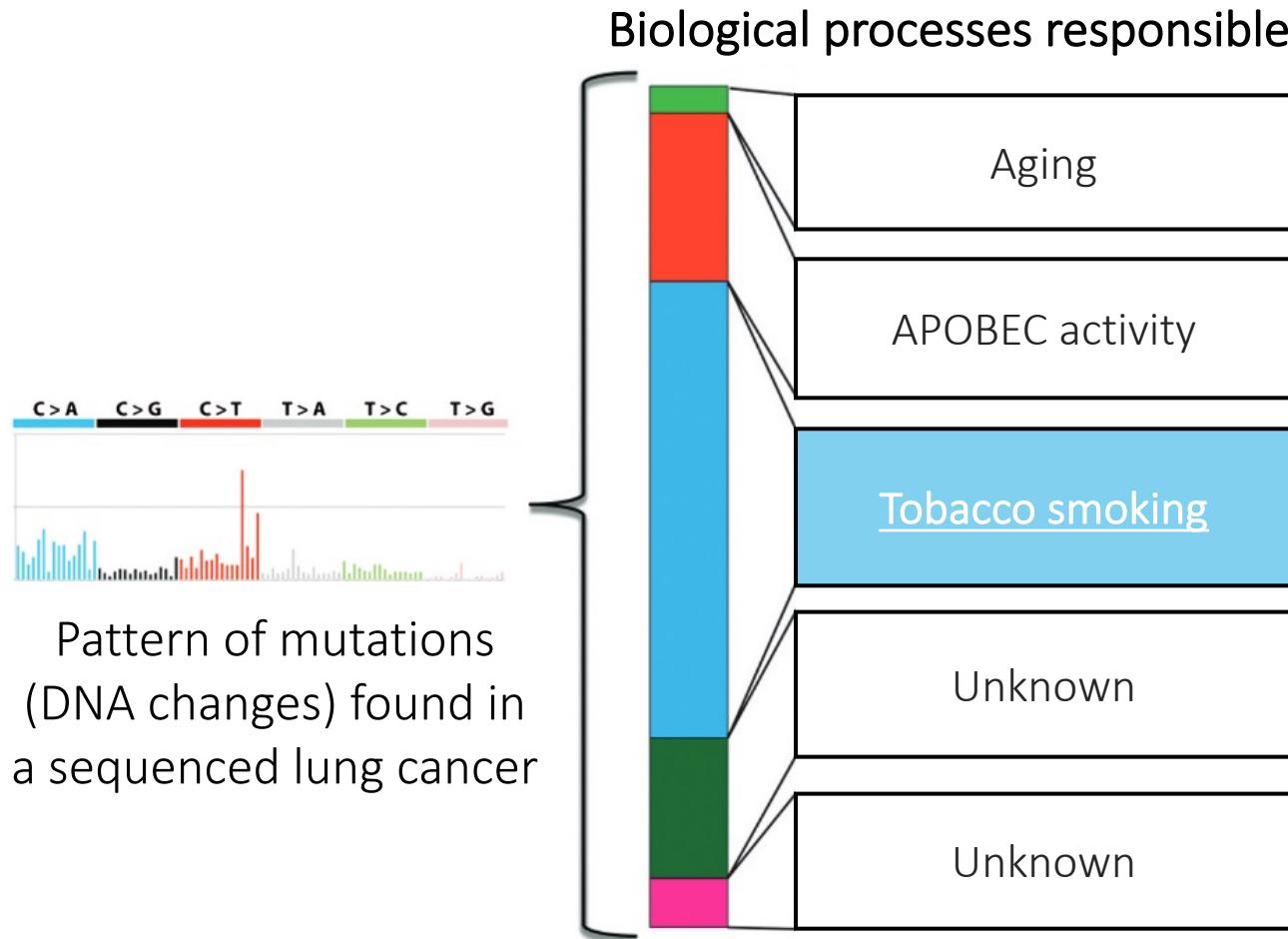
# Spoiler alert!



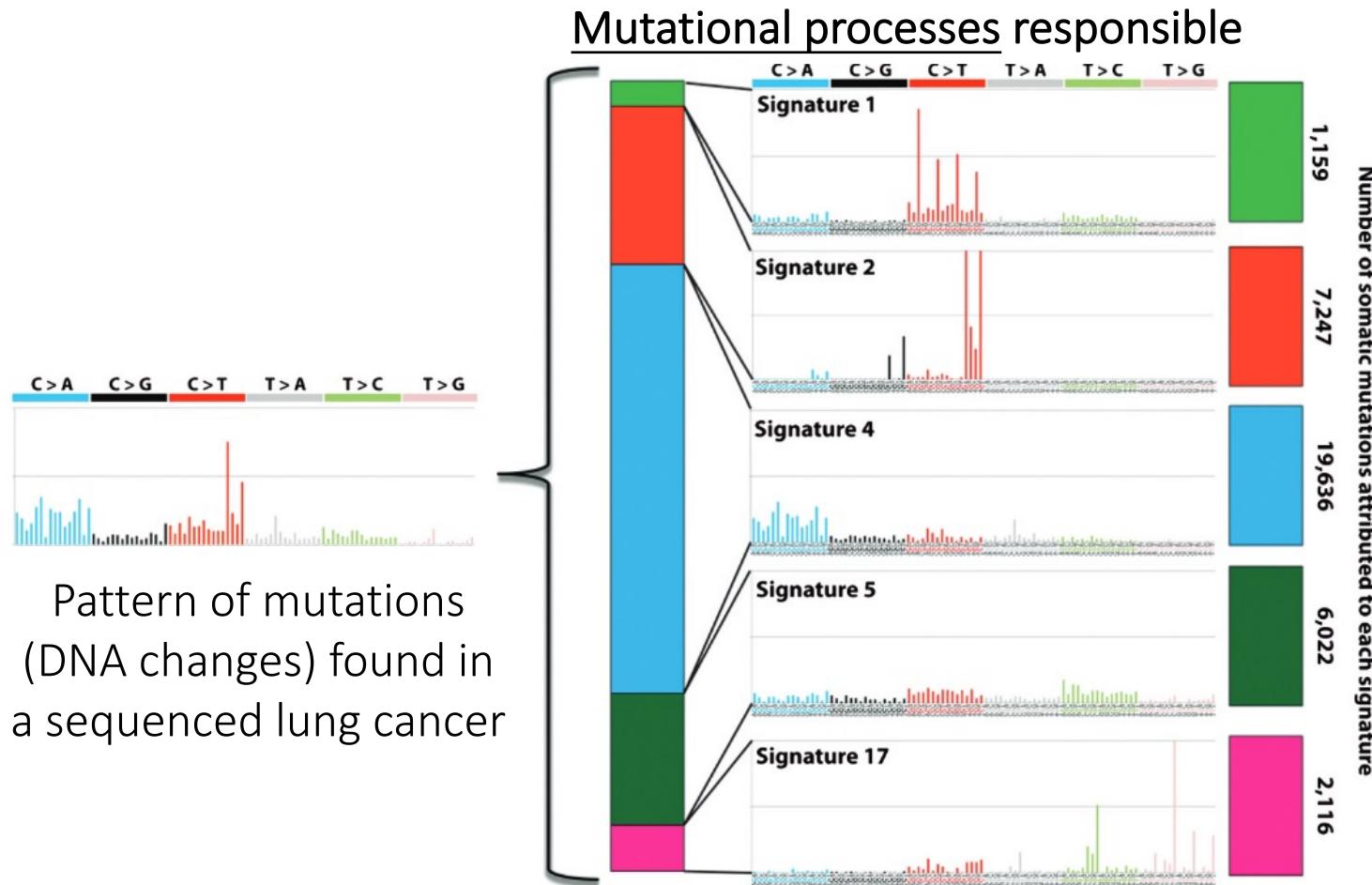
# Spoiler alert!



# Spoiler alert!



# Spoiler alert!



# Today's agenda

Basics of cancer genomics: genomic sequencing data and somatic mutations identification

Exploring and obtaining tumor mutation data from cBioPortal

Characterization of the patterns of mutations in cancer

Mutational matrix generation using **Maftools**

Exploration of the biological processes generating mutations in different cancer types

Mutational signature analysis using **MutationalPatterns**

# Today's agenda

## Basics of cancer genomics: genomic sequencing data and somatic mutations identification

Exploring and obtaining tumor mutation data from cBioPortal

Characterization of the patterns of mutations in cancer

Mutational matrix generation using Maftools

Exploration of the biological processes generating mutations in different cancer types

Mutational signature analysis using MutationalPatterns

# Overview of cancer

- All cancers originate from a single cell that starts to behave abnormally, dividing uncontrollably and invading adjacent tissues
- The reason that this single cell begins to behave abnormally is because of acquired changes in its genome known as somatic mutations
- Cancer is a disease of the genome and the most common human genetics disease

# Types of mutations

- DNA molecules in our cells are targeted by diverse mutagenic processes that can occur in:
  - **germ** cells, contributing to species evolution
  - or in **somatic** cells, accumulating with age and contributing to diseases, especially cancer
- Recent mutation rate studies of tumors have focused on deciphering the **somatically acquired changes** in the DNA of cancer cells to advance our understanding of the relations among mutagenic exposures, DNA damage and repair, and outcomes (such as cancer and uncontrolled cell growth)

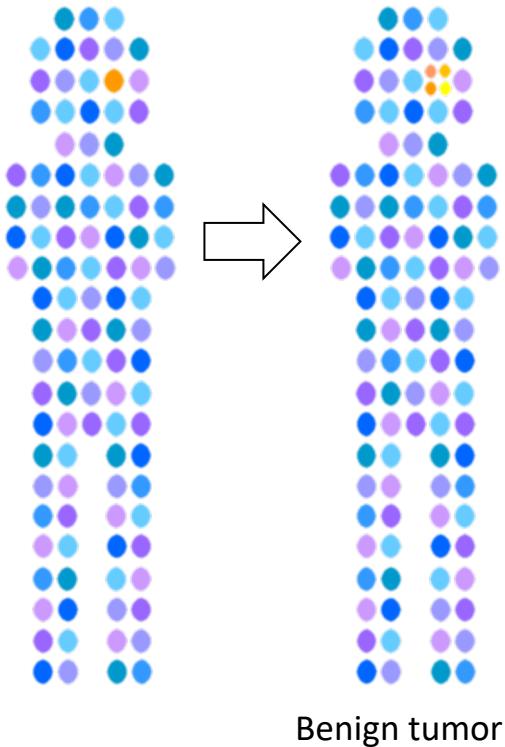
# Cancer development

A cancer arises when a single cell acquires somatic mutations and begins to behave abnormally.  
(dividing when it should be quiescent)



# Cancer development

**Benign tumor:** a cell has evaded some controls on growth giving rise to a ‘clonal mass’, however they lack many of the aggressive characteristics of more advanced cancer (i.e. unlimited invasive growth).



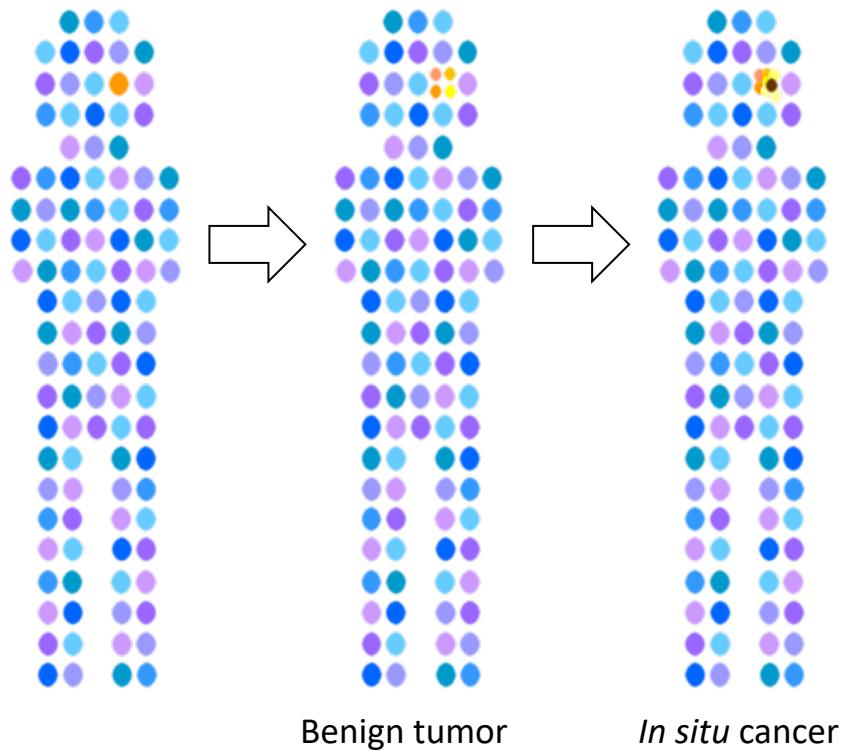
Moles (nevi) are an example of a benign tumor.

82% of nevi have a mutation of the known cancer gene *BRAF*.

*BRAF* mutations are thought to be the initiating event in melanoma.

# Cancer development

*In situ* cancer: the tumor has evaded controls on cell division and grows in a disorderly fashion.

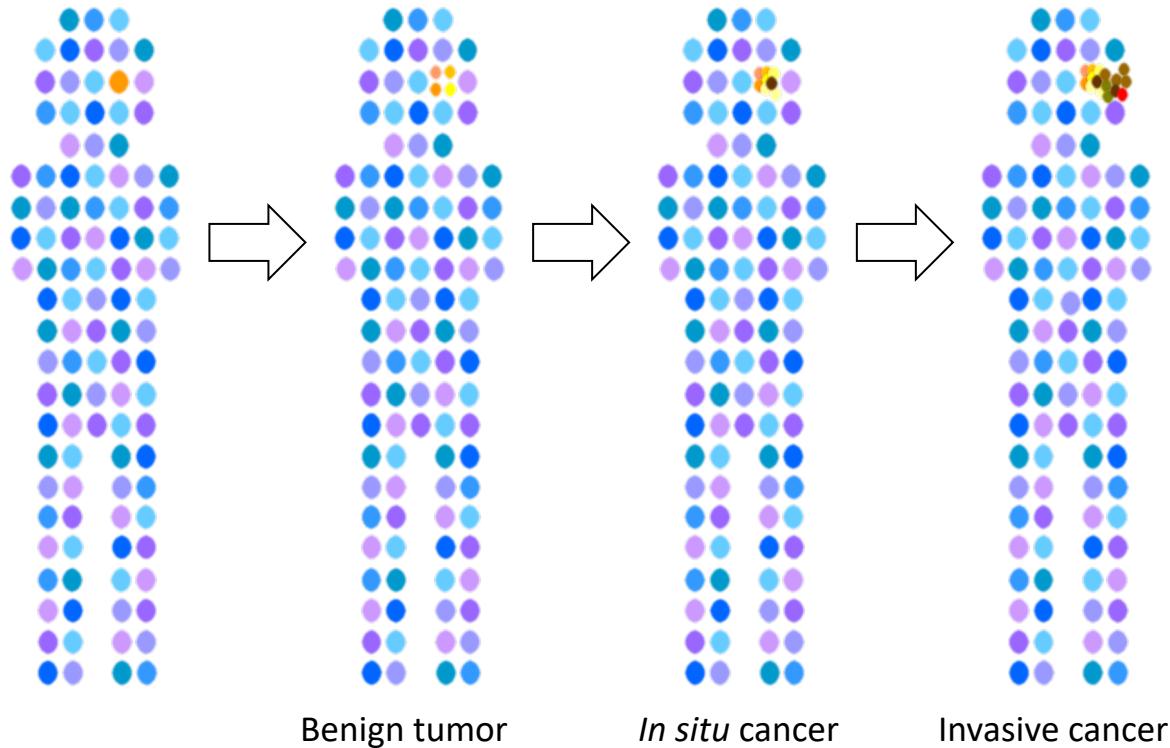


"in situ" means "in its natural place".

The tumor cells are still confined to the site where they originated.

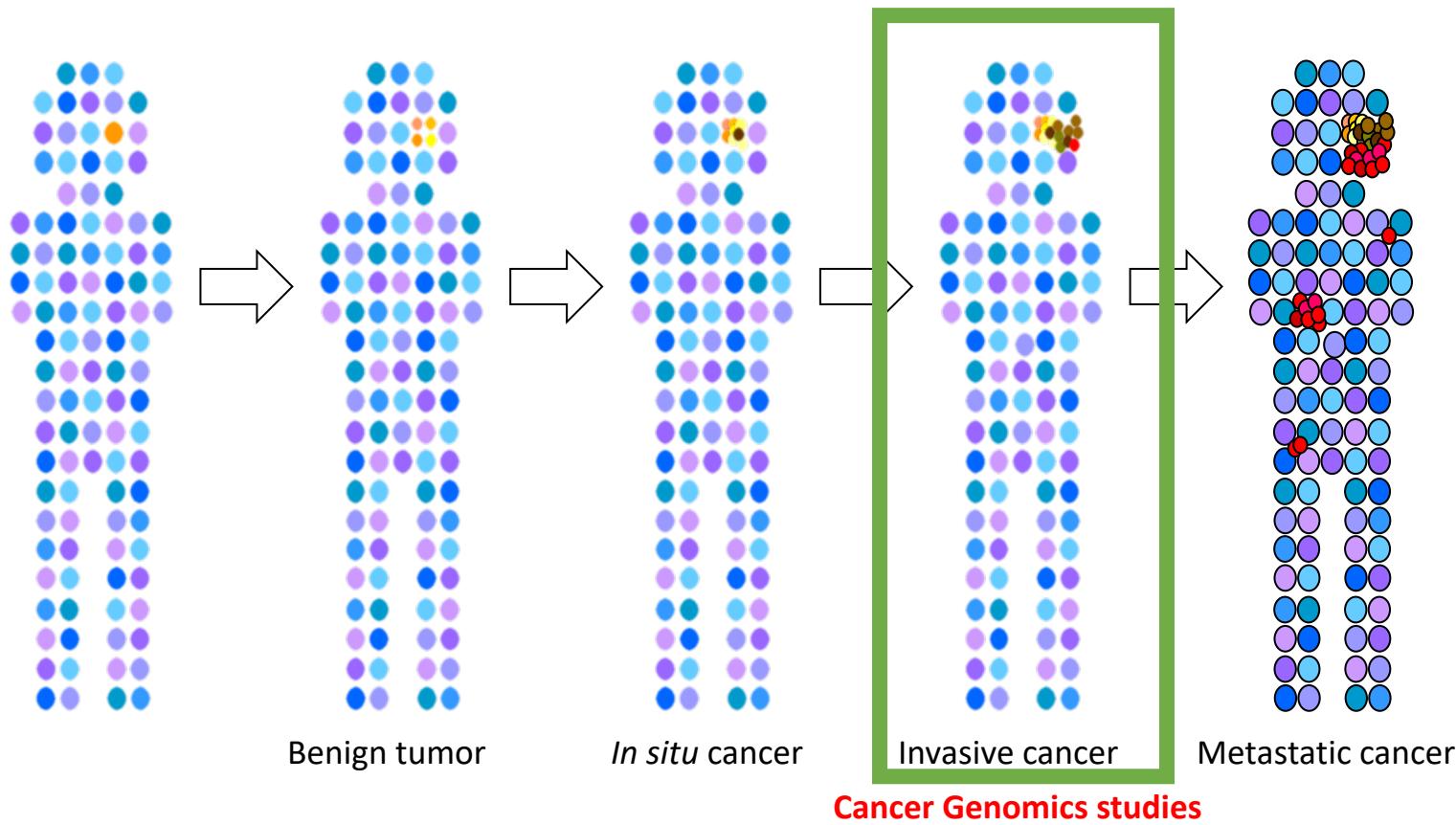
# Cancer development

Invasive cancer: the tumor has spread beyond the layer of tissue in which it developed and is growing into surrounding, healthy tissues.

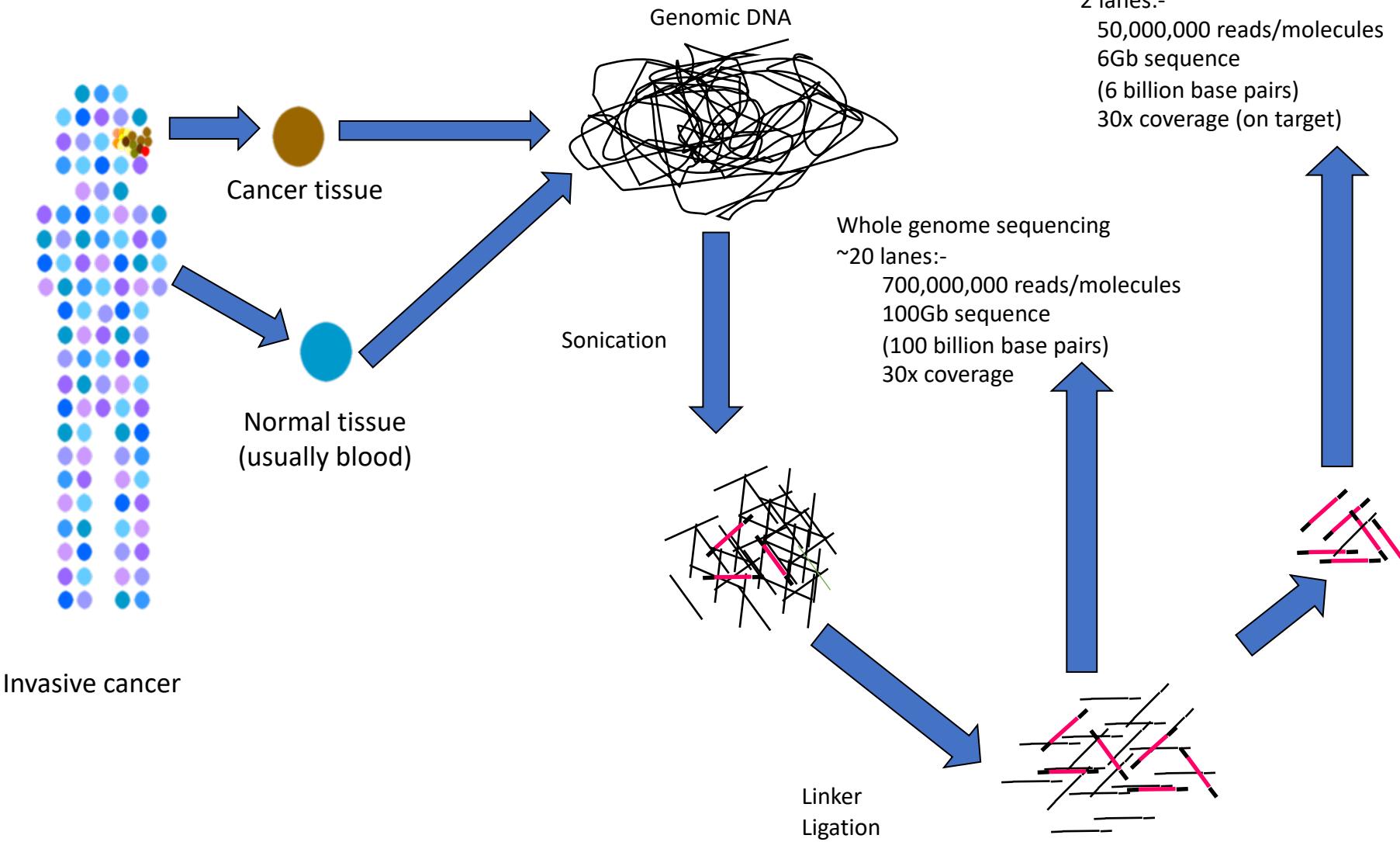


# Cancer development

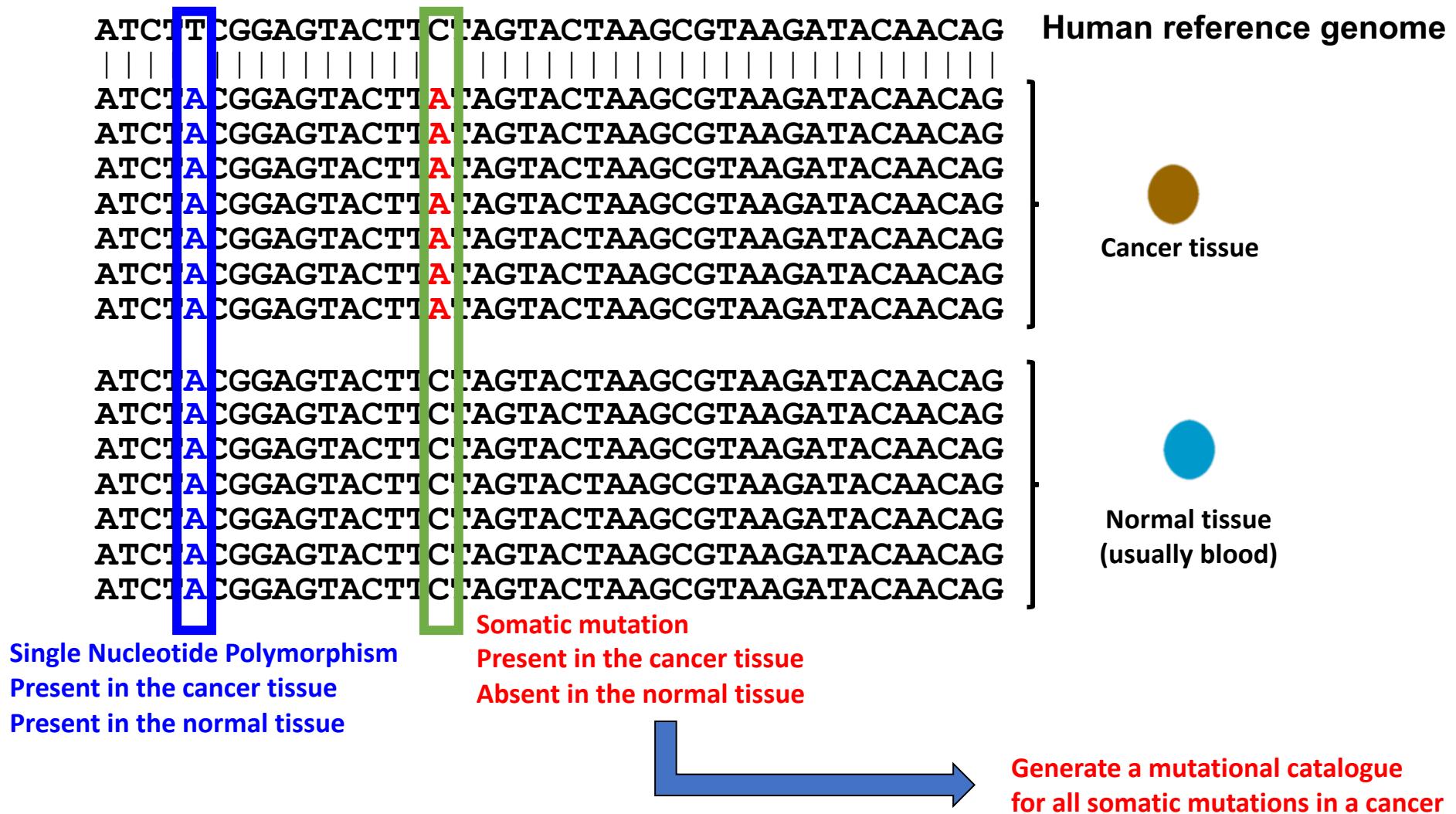
Metastatic cancer: the tumor has spread from the place where the cancer started to other parts of the body. Metastases are the cause of 90% of human cancer deaths.



# Cancer genomics approach

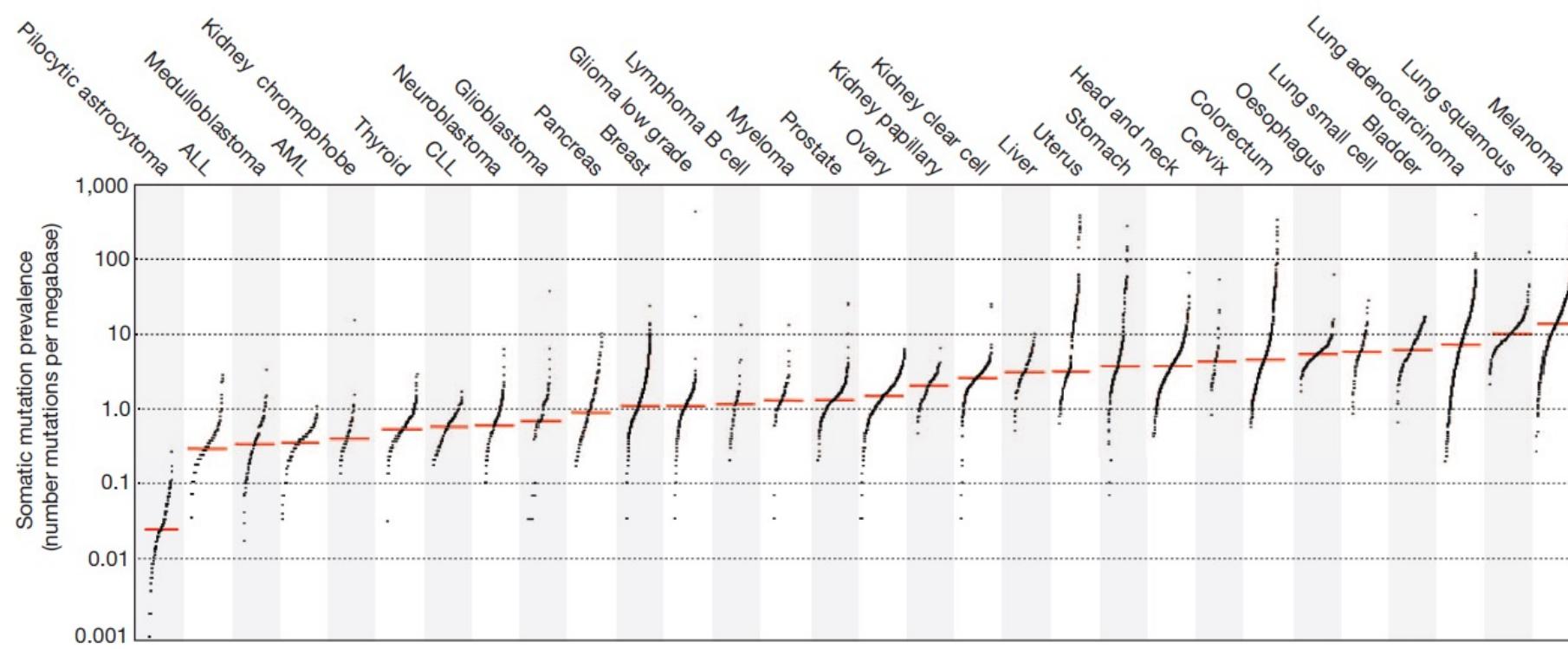


# Cancer genomics approach



# Somatic mutations in cancer

- The burden of somatic mutations is highly variable among different cancer types
- The most mutated cancer types (lung and skin cancers) are associated with well-known environmental mutagens (tobacco smoking and UV light exposure, respectively)



**Statistical analysis to identify 1 to 10 driver mutations.**

# Today's agenda

Basics of cancer genomics: genomic sequencing data and somatic mutations identification

Exploring and obtaining tumor mutation data from cBioPortal

Characterization of the patterns of mutations in cancer

Mutational matrix generation using **Maftools**

Exploration of the biological processes generating mutations in different cancer types

Mutational signature analysis using **MutationalPatterns**

# Today's agenda

Basics of cancer genomics: genomic sequencing data and somatic mutations identification

Exploring and obtaining tumor mutation data from cBioPortal

## Characterization of the patterns of mutations in cancer

Mutational matrix generation using Maftools

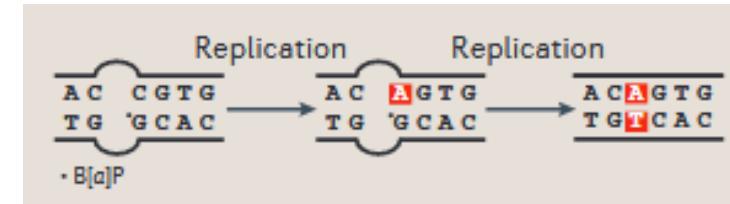
Exploration of the biological processes generating mutations in different cancer types

Mutational signature analysis using MutationalPatterns

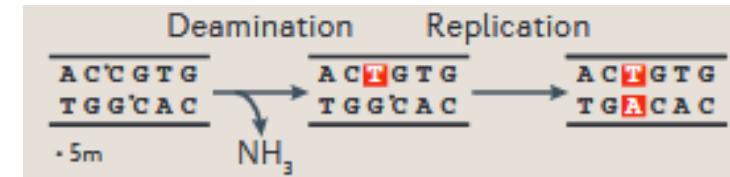
# Mutational processes

Cancer genomes accumulate a large number of somatic mutations resulting from various endogenous and exogenous causes, including mutations triggered by carcinogenic exposures, normal cellular activities, and cancer-related aberrations of the DNA maintenance machinery.

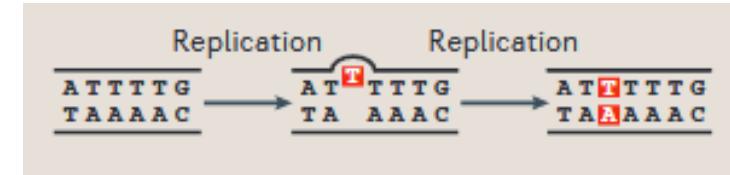
Environmental exposures  
Tobacco smoking or chewing



Normal cellular activities  
Spontaneous deamination of methylated cytosines



Failure in DNA replication or repair  
Aberrant mismatch repair pathway



# Classification of base substitution mutations

.....ATCGGGAAAT**C**GGACCCGATG.....  
                        ↓  
.....ATCGGGAAAT**T**GGACCCGATG.....

# Classification of base substitution mutations

.....ATCGGGAA**TCG**GACCCGATG.....  
                        ↓  
.....ATCGGGAA**TTG**GACCCGATG.....

# Classification of base substitution mutations

.....ATCGGGAA**TCG**GACCCGATG.....  
                        ↓  
.....ATCGGGAA**TTG**GACCCGATG.....

.....ATCGGGAA**ACG**GACCCGATG.....  
                        ↓  
.....ATCGGGAA**ATG**GACCCGATG.....

# Classification of base substitution mutations

.....ATCGGGAA**TCG**GACCCGATG.....



.....ATCGGGAA**TTG**GACCCGATG.....

.....ATCGGGAA**ACG**GACCCGATG.....



.....ATCGGGAA**ATG**GACCCGATG.....

.....ATCGGGAA**ACC**GACCCGATG.....



.....ATCGGGAA**ATC**GACCCGATG.....

# Classification of base substitution mutations

C>T

C>A

C>G

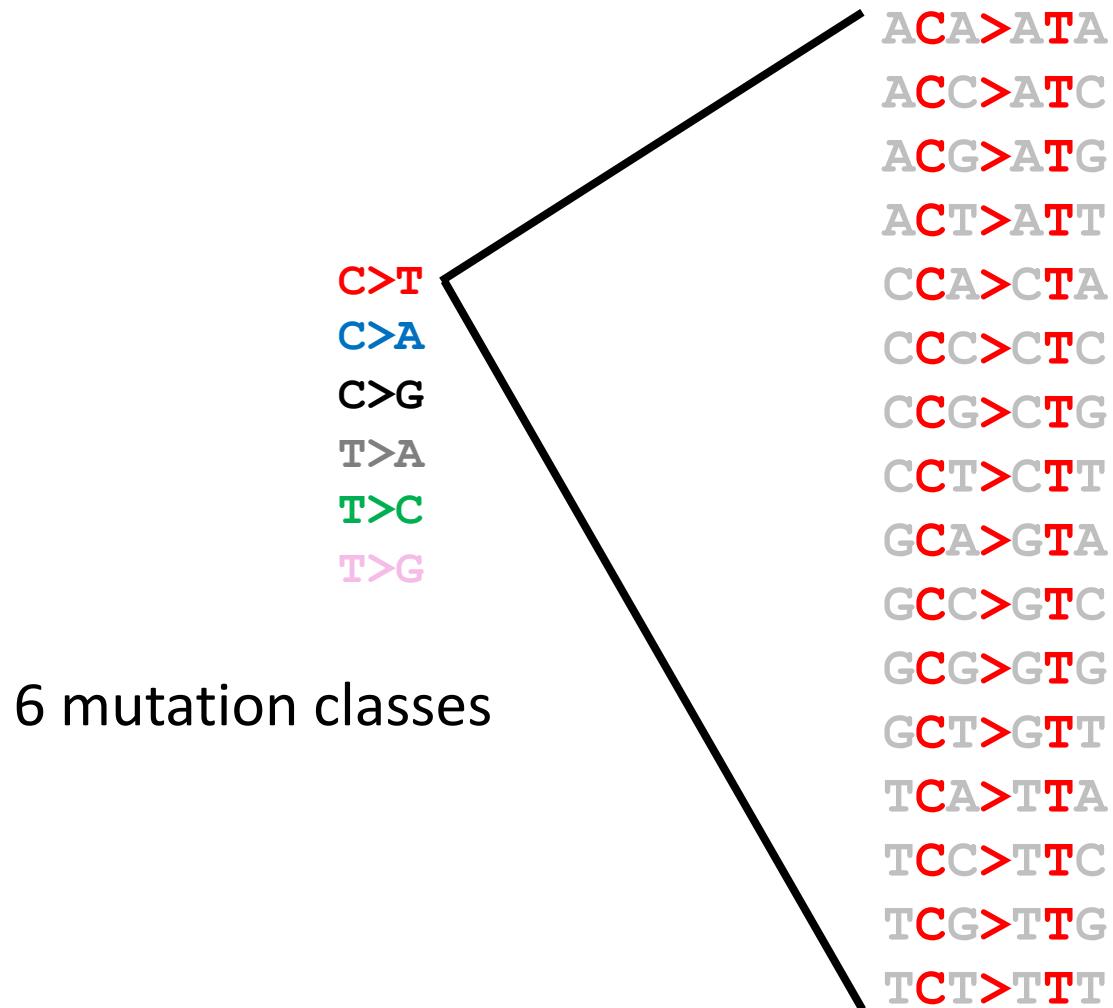
T>A

T>C

T>G

6 mutation classes

# Classification of base substitution mutations



# Classification of base substitution mutations

6 mutation classes

C>T

C>A

C>G

T>A

T>C

T>G

ACA>ATA	ATA>AAA
ACC>ATC	ATC>AAC
ACG>ATG	ATG>AAG
ACT>ATT	ATT>AAT
CCA>CTA	CTA>CAA
CCC>CTC	CTC>CAC
CCG>TG	CTG>CAG
CTT>CTT	CTT>CAT
GCA>CTA	GTA>GAA
GCC>TC	GTC>GAC
GCG>GTG	GTG>GAG
GCT>CTT	GTT>GAT
TCA>TTA	TTA>TAA
TCC>ITC	TTC>TAC
TCG>TG	TTC>TAG
TCT>ITT	TTT>TAT
ACA>AAA	ATA>ACA
ACC>AAC	ATC>ACC
ACG>ACG	ATG>ACG
ACT>ACT	ATT>ACT
CCA>CAA	CTA>CCA
CCC>CAC	CTC>CCC
CCG>AG	CTG>CCG
CCT>CAT	CTT>CCT
GCA>AA	GTA>GCA
GCC>GAC	GTC>GCC
GCG>GAG	GTG>GCG
GCT>CAT	GTT>GCT
TCA>TAA	TTA>TCA
TCC>TAC	TTC>TCC
TCG>TAG	TTG>TCG
TCT>TAT	TTT>TCT
ACA>AGA	ATA>AGA
ACC>GC	ATC>AGC
ACG>GG	ATG>AGG
ACT>GT	ATT>AGT
CCA>CGA	CTA>CGA
CCC>GC	CTC>CGC
CCG>GG	CTG>CGG
CCT>CGT	CTT>CGT
GCA>GA	GTA>GGA
GCC>GC	GTC>GGC
GCG>GG	GTG>GGG
GCT>GT	GTT>GGT
TCA>GA	TTA>TGA
TCC>GC	TTC>TGC
TCG>GG	TTG>TGG
TCT>GT	TTT>TGT

96 mutation classes

# Patterns of mutations are defined by base substitutions and context

Six classes of single-base mutations

Reported by pyrimidine

Adding 5' and 3' adjacent bases

96 possibilities considering context

