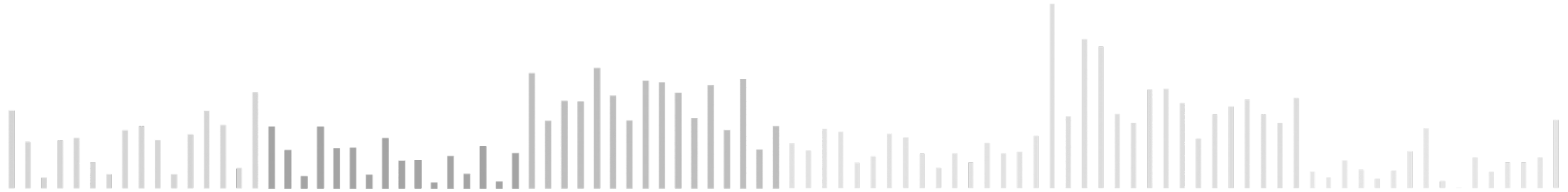


BIMM 143 – Class 18 Lab

Mutational Signatures in Cancer



Marcos Díaz-Gay

UC San Diego

 [marcos-diazg.github.io](https://github.com/marcos-diazg)

 [@mdiazgay](https://twitter.com/mdiazgay)

 [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...



Galicia, Spain

Barcelona, Spain

IDIBAPS
Institut
D'Investigacions
Biomèdiques
August Pi i Sunyer

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



... pivoting from civil engineering to computational biology



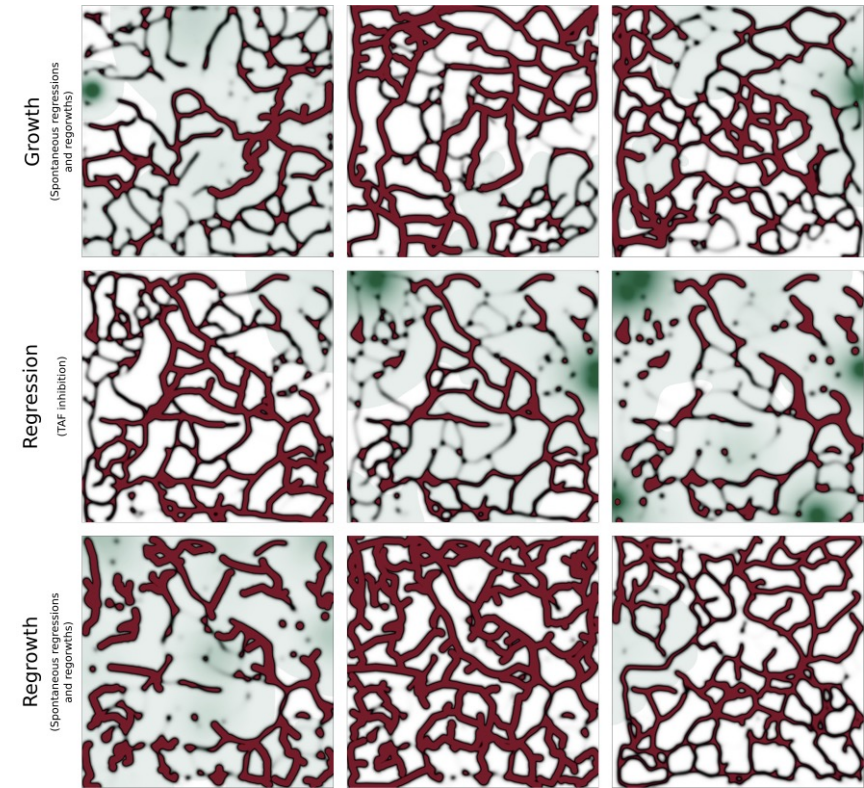
... pivoting from civil engineering to computational biology

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



... pivoting from civil engineering to computational biology

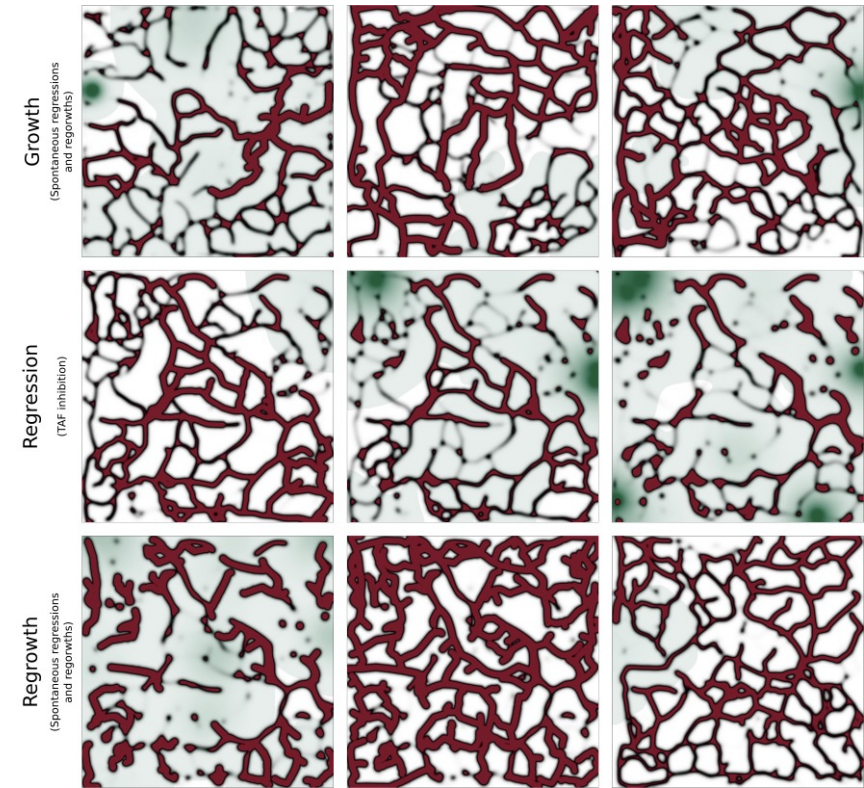
New paradigm of civil engineering
Apply the same mathematical
framework, but for cancer research



... 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

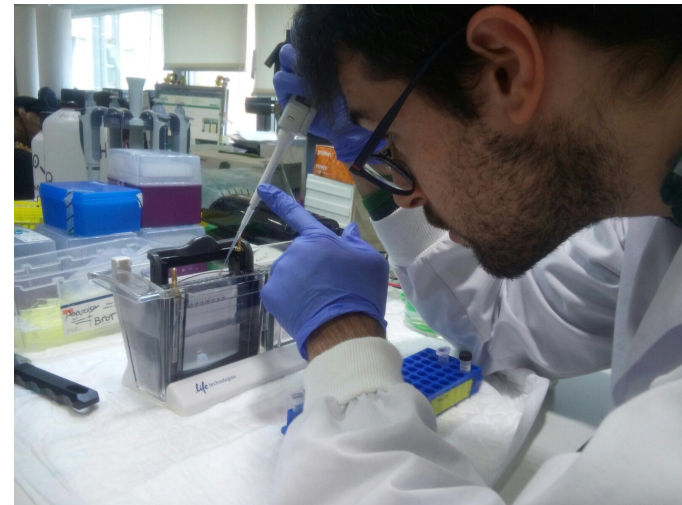
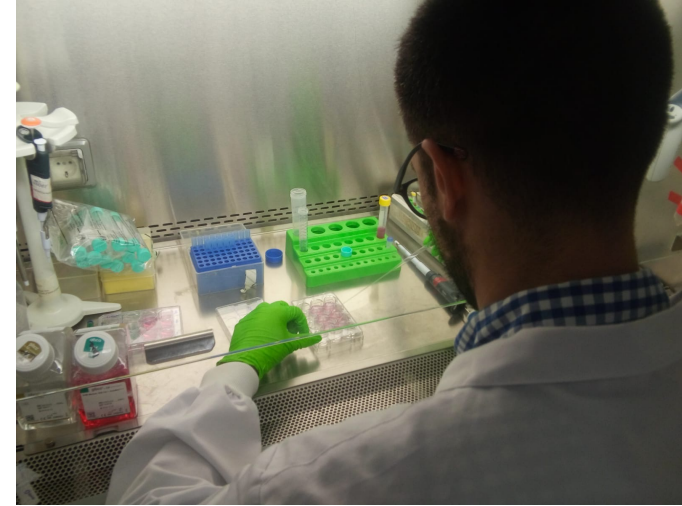


... pivoting from civil engineering to computational biology

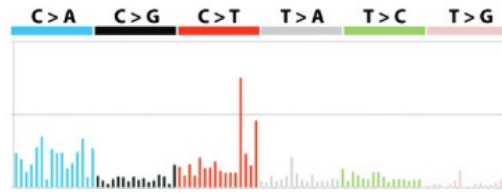
New paradigm of civil engineering
Apply the same mathematical
framework, but for cancer research

MSc in Biomedicine
The first step

PhD in Translational Medicine
A new biomedical researcher

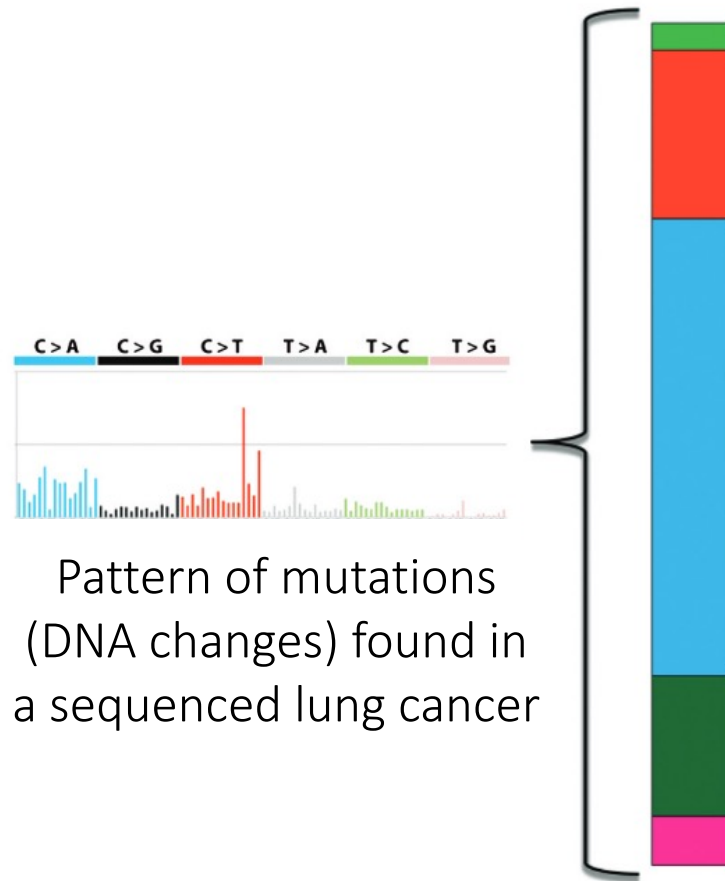


Spoiler alert!

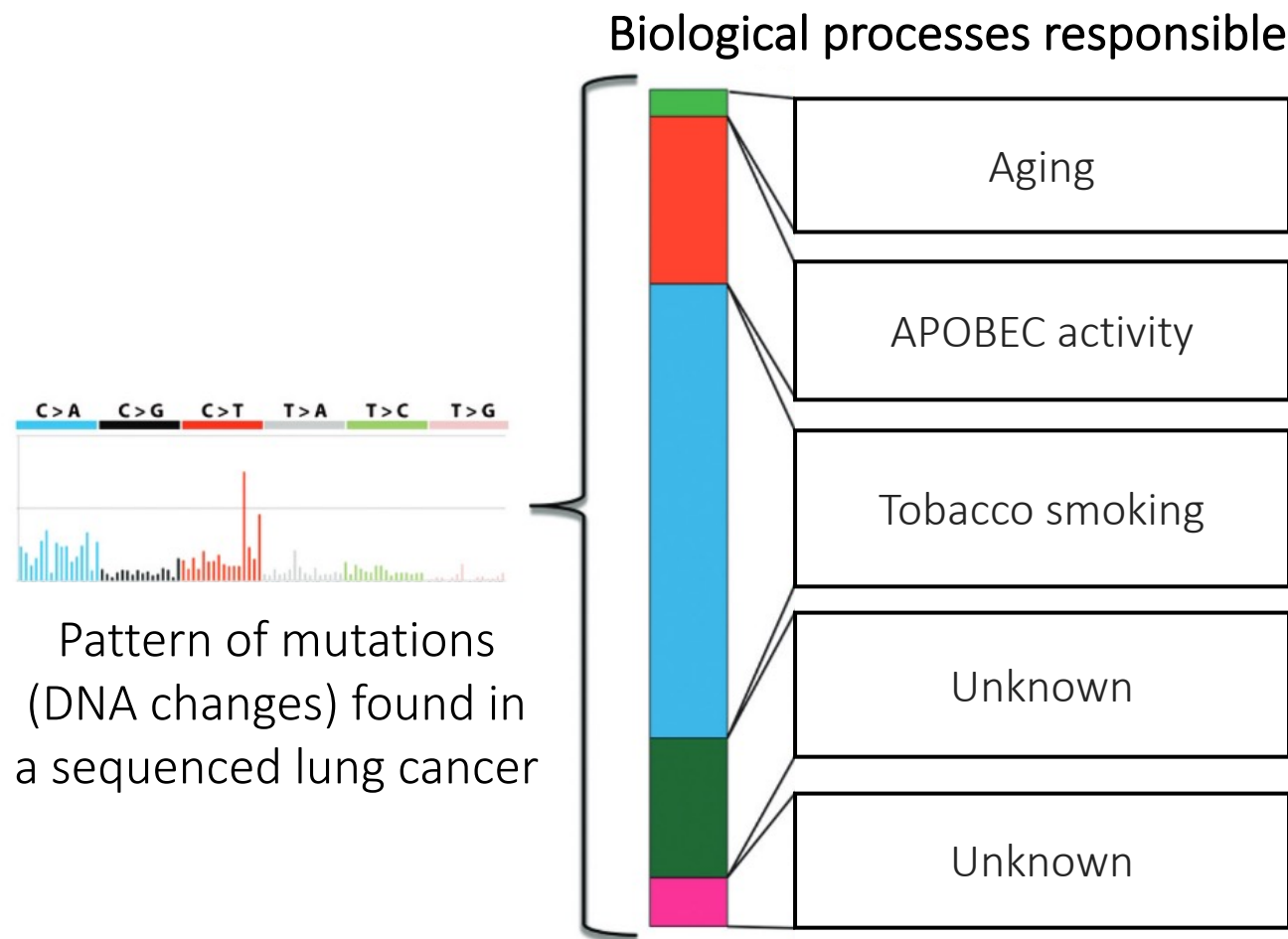


Pattern of mutations
(DNA changes) found in
a sequenced lung cancer

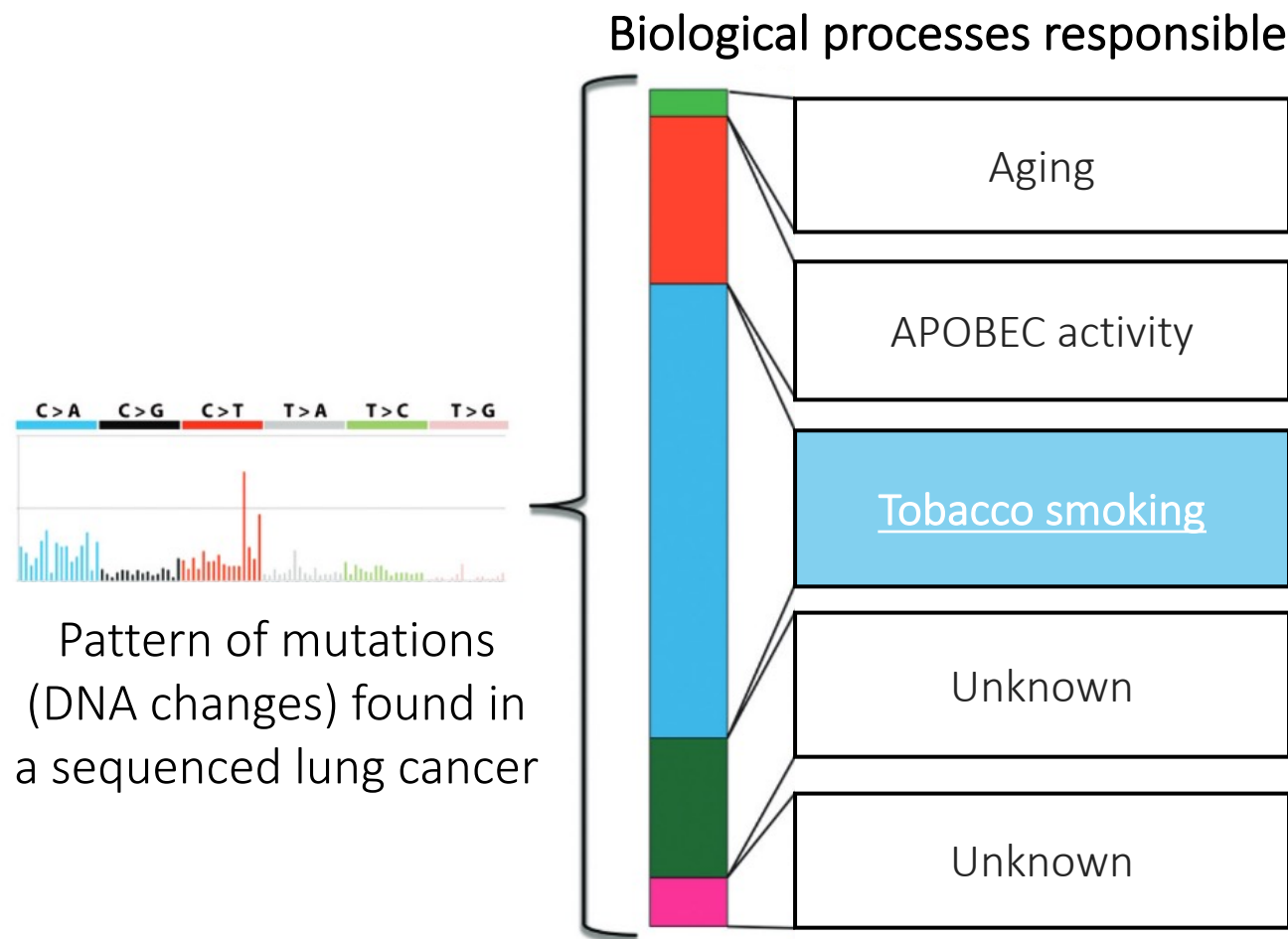
Spoiler alert!



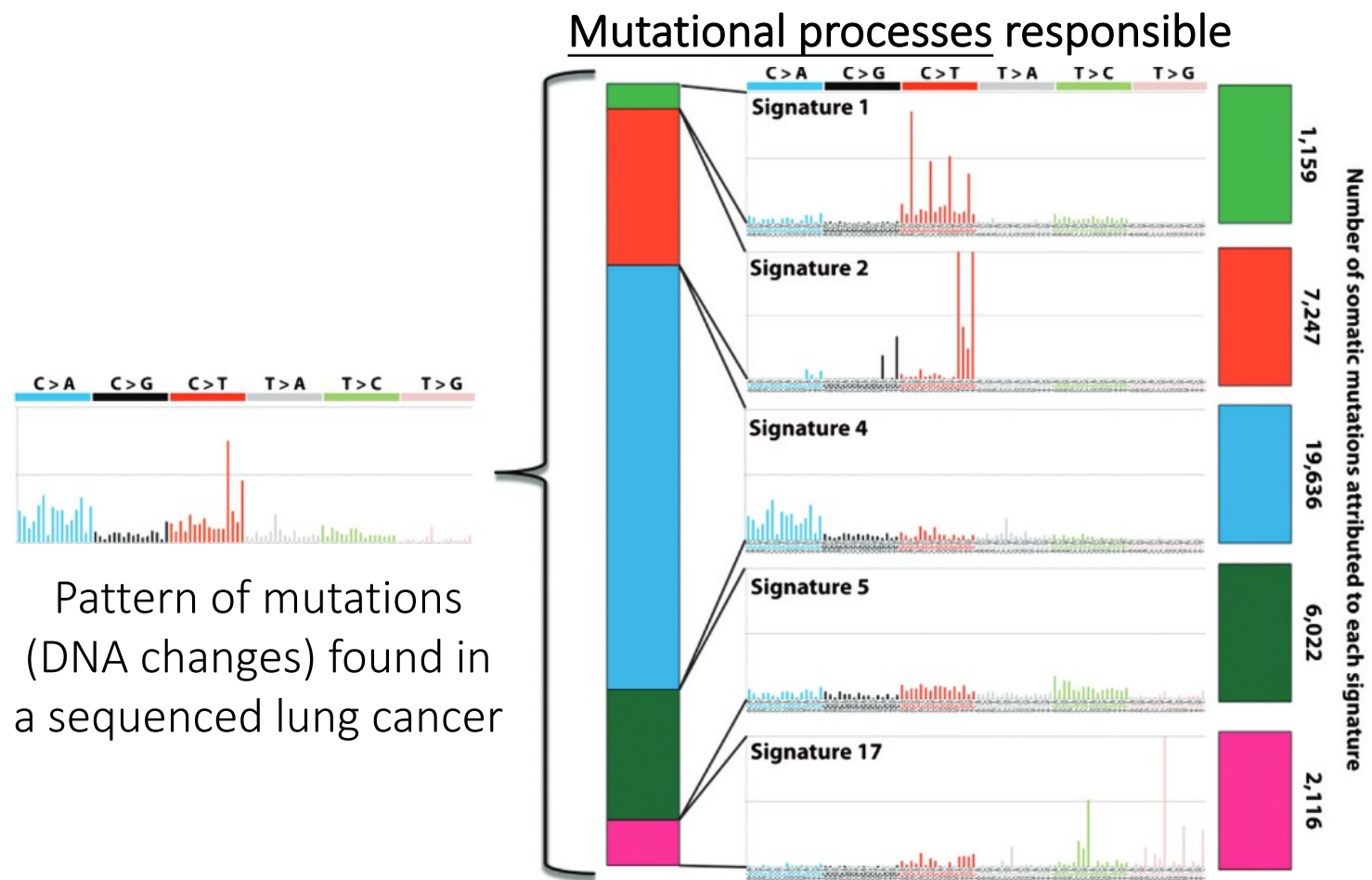
Spoiler alert!



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)

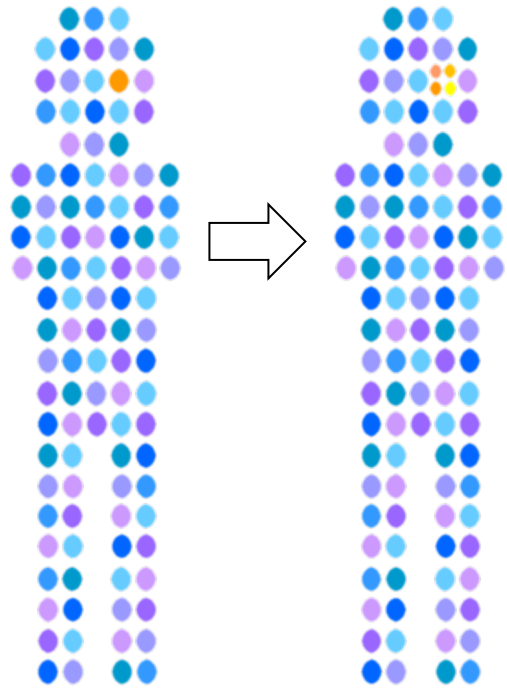
Cancer development

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).



Benign tumor

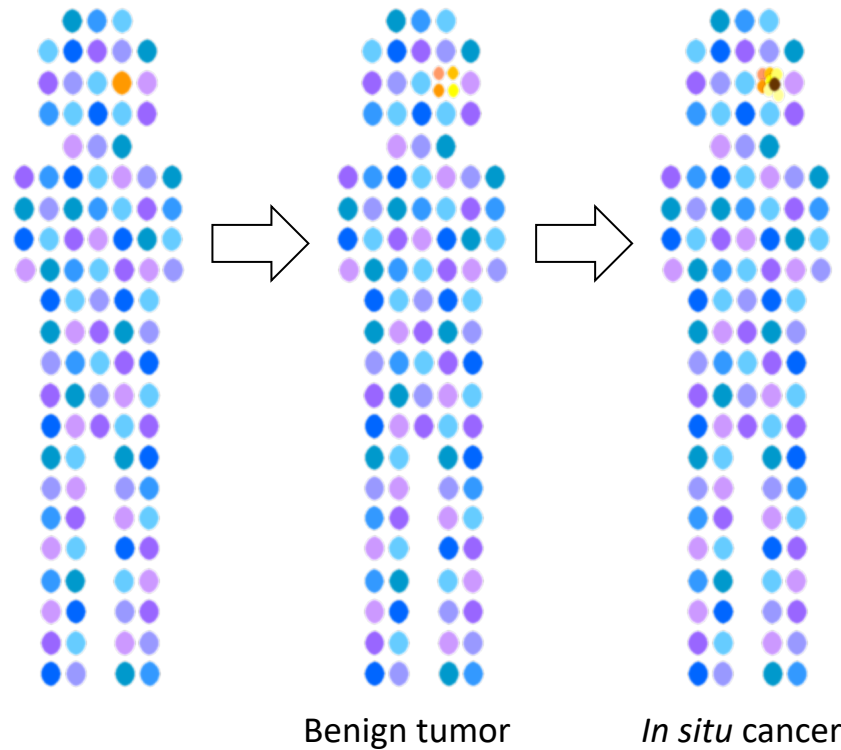
Moles (nevi) are an example of a benign tumor.

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.

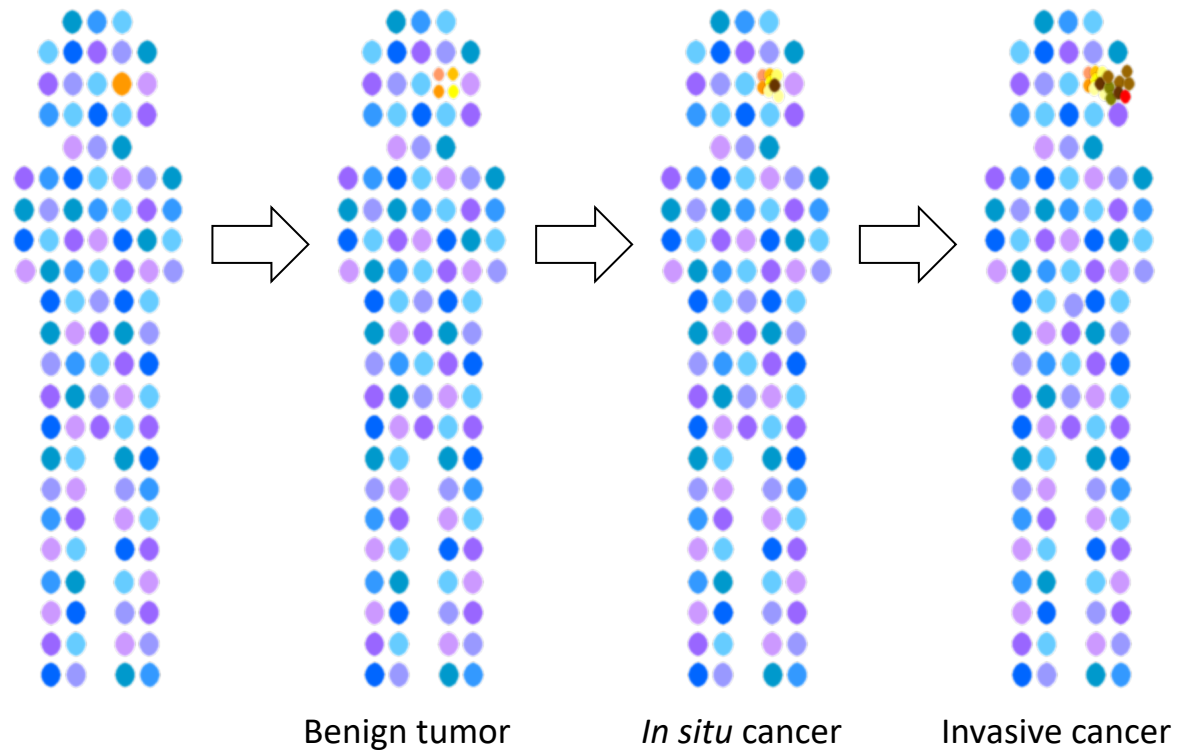


“in situ” means “in its natural place”.

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

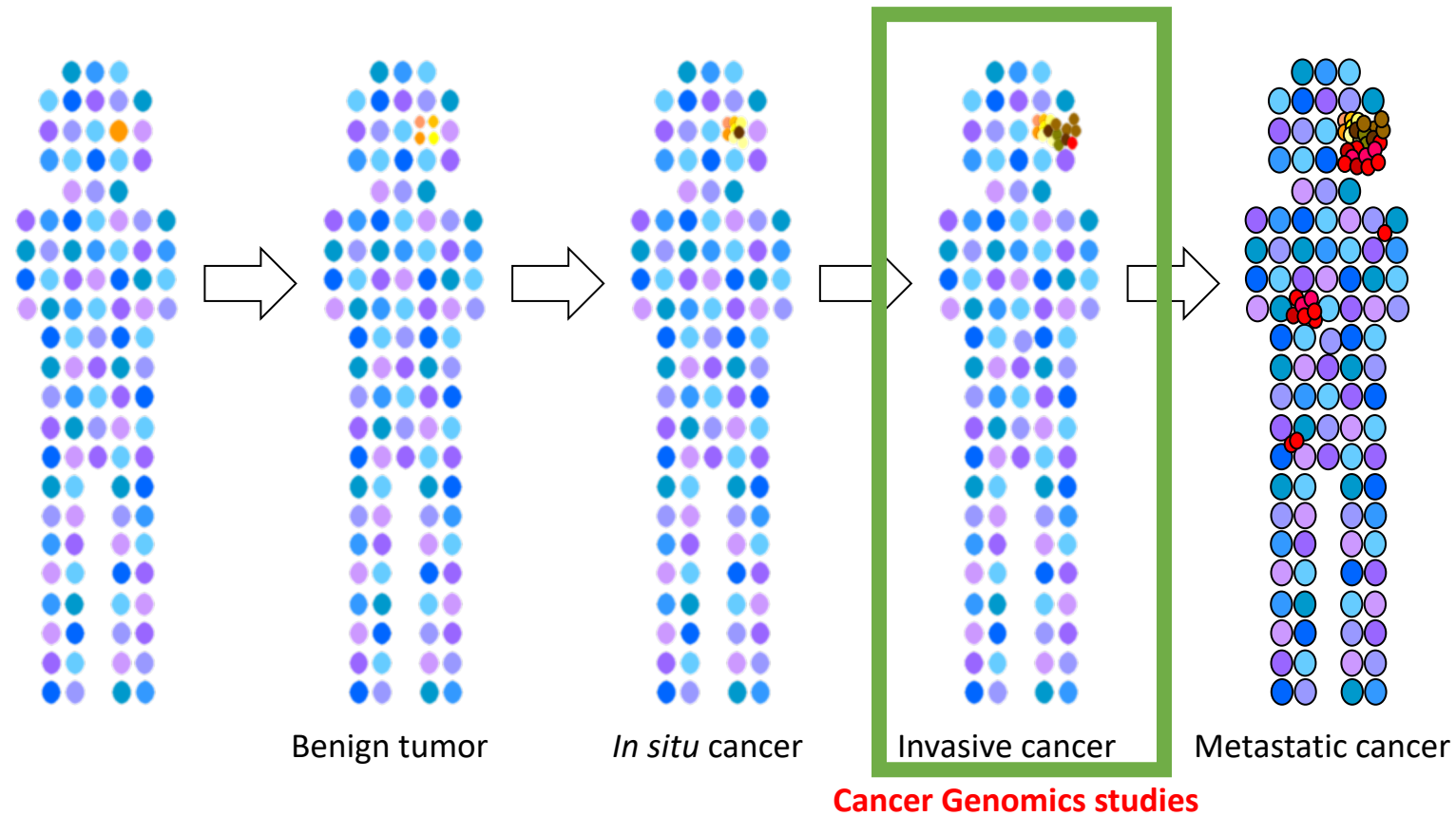
Cancer development

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

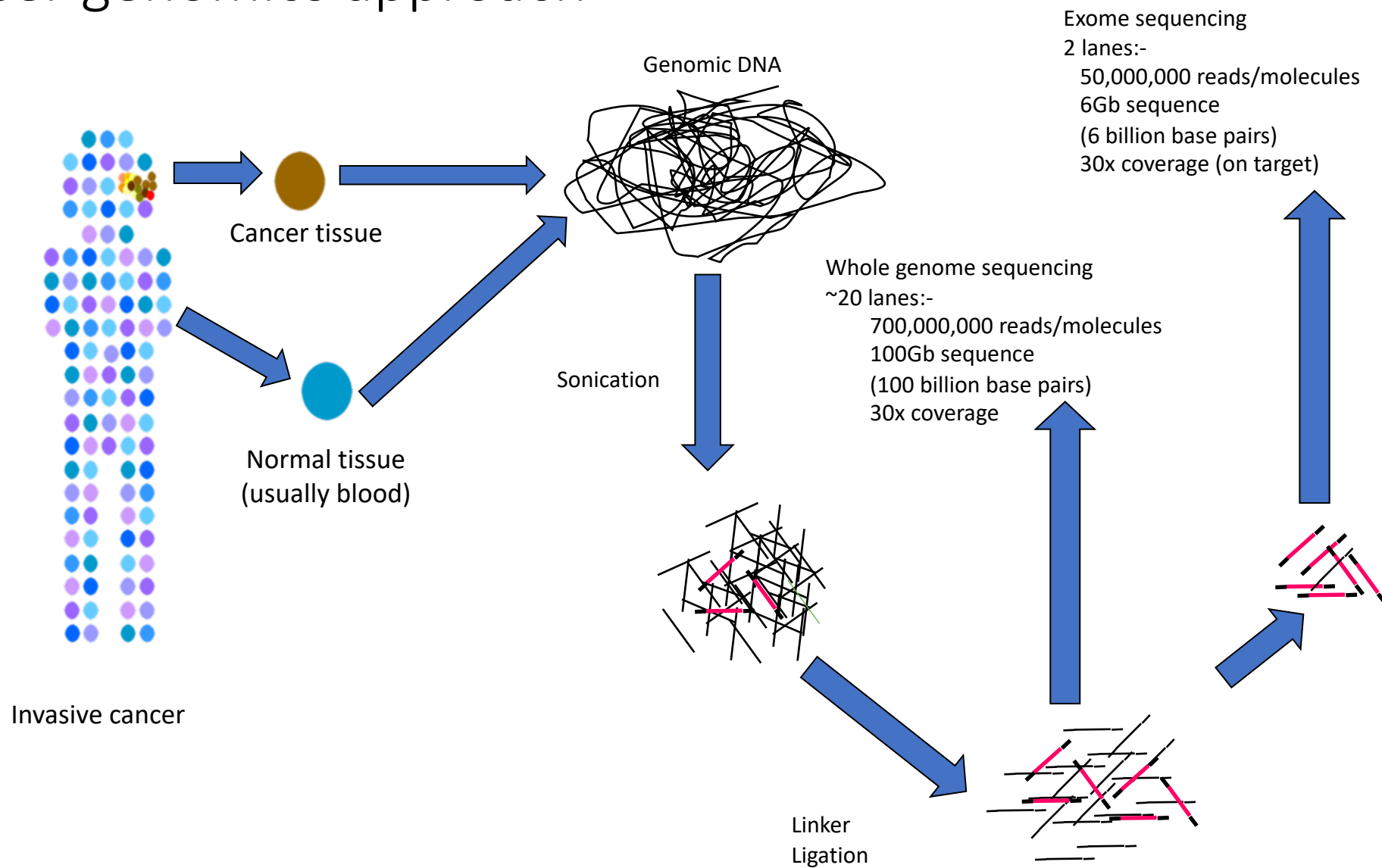


Cancer development

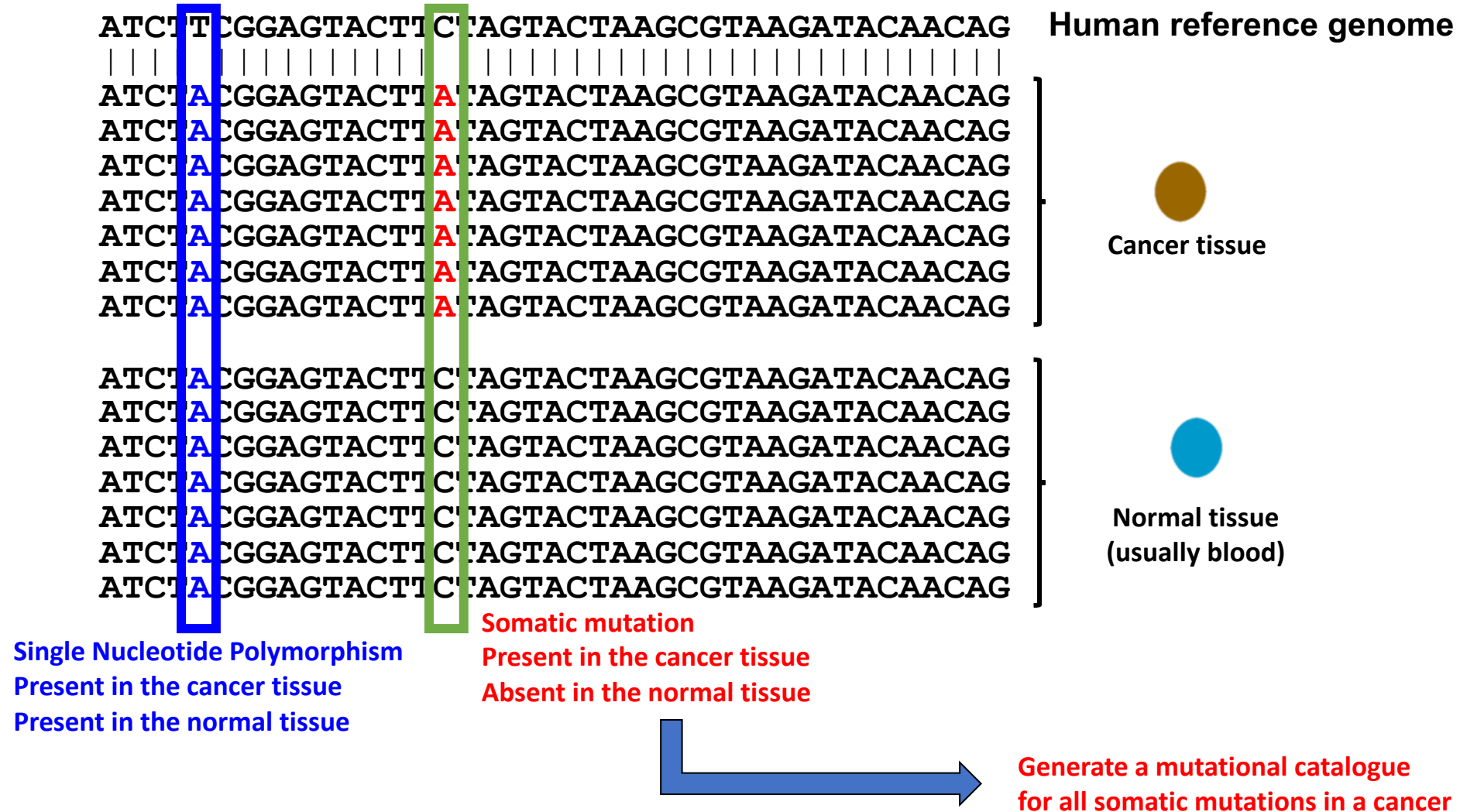
Metastatic cancer: the tumor has spread from the place where the cancer started to other parts of the body. Metastases are the cause of 90% of human cancer deaths.



Cancer genomics approach

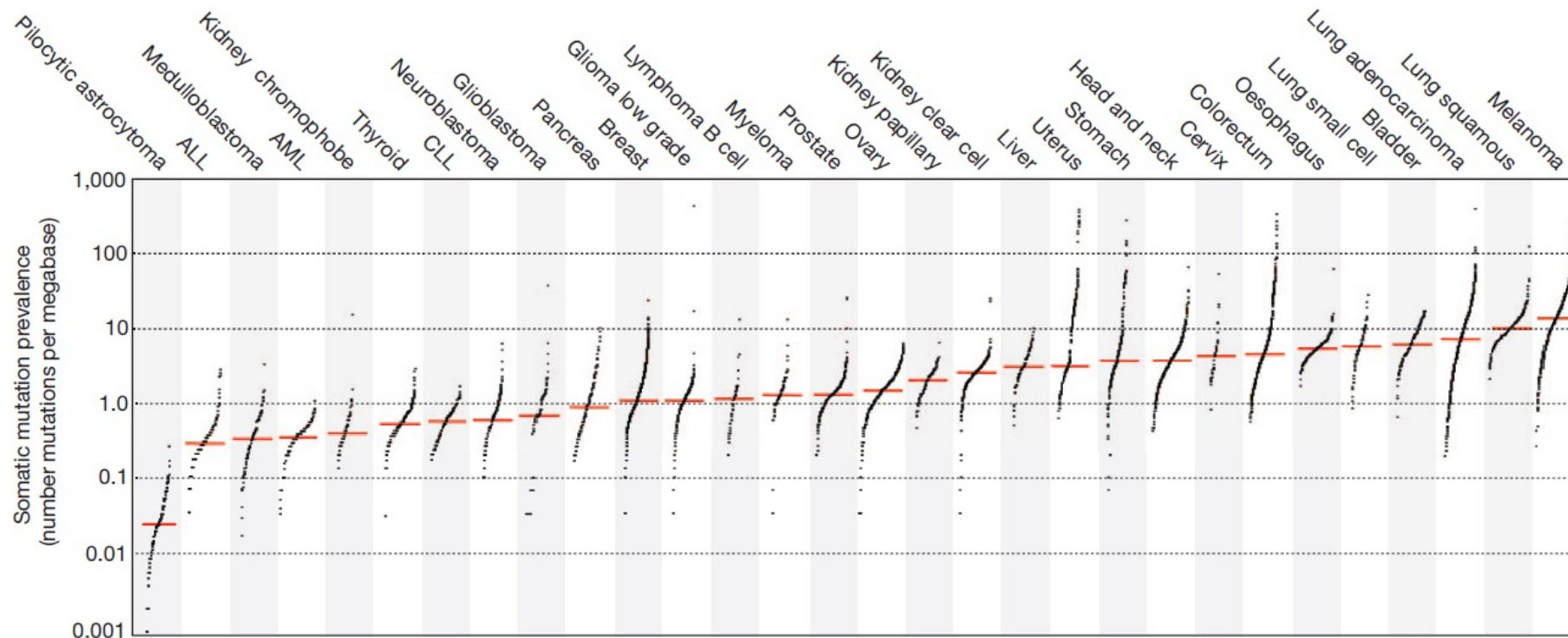


Cancer genomics approach



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

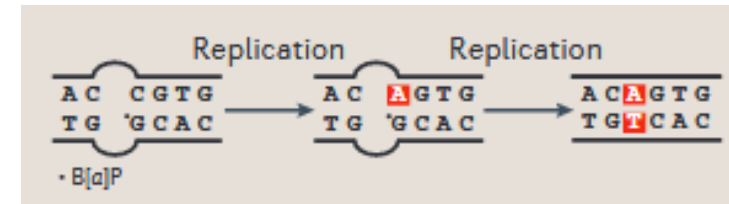
Mutational signature analysis using **MutationalPatterns**

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.

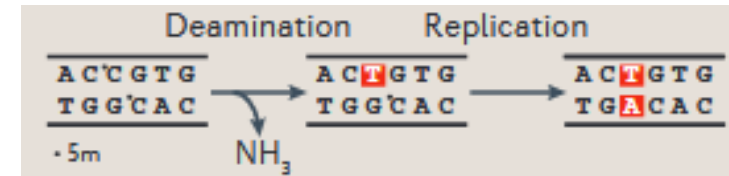
Environmental exposures

Tobacco smoking or chewing



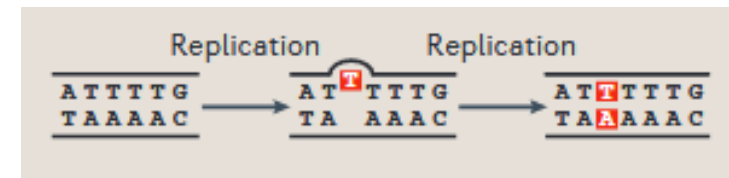
Normal cellular activities

Spontaneous deamination of methylated cytosines



Failure in DNA replication or repair

Aberrant mismatch repair pathway



Classification of base substitution mutations

.....ATCGGGAAT**C**GGACCCGATG.....
 ↓
.....ATCGGGAAT**T**GGACCCGATG.....

Classification of base substitution mutations

.....ATCGGGAA**TC**GGACCCGATG.....
 ↓
ATCGGGAA**TT**GGACCCGATG.....

Classification of base substitution mutations

.....ATCGGGAA**T****C**GGACCCGATG.....



.....ATCGGGAA**T****T**GGACCCGATG.....

.....ATCGGGAA**A****C**GGACCCGATG.....



.....ATCGGGAA**A****T**GGACCCGATG.....

Classification of base substitution mutations

.....ATCGGGAA**TC**GGACCCGATG.....



.....ATCGGGAA**TT**GGACCCGATG.....

.....ATCGGGAA**AC**GGACCCGATG.....



.....ATCGGGAA**AT**GGACCCGATG.....

.....ATCGGGAA**ACC**GGACCCGATG.....



.....ATCGGGAA**ATC**GGACCCGATG.....

Classification of base substitution mutations

C>T

C>A

C>G

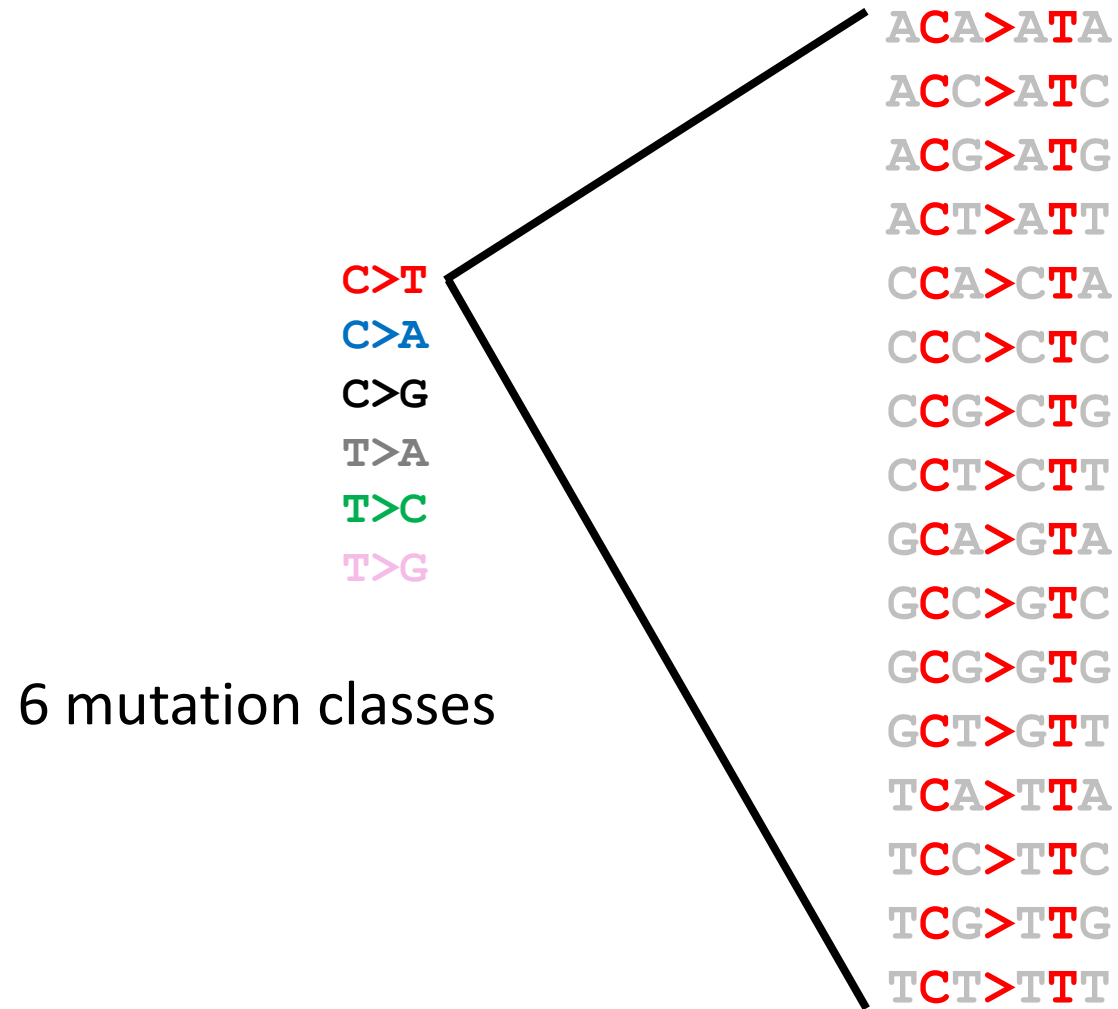
T>A

T>C

T>G

6 mutation classes

Classification of base substitution mutations



Classification of base substitution mutations

C>T

C>A

C>G

T>A

T>C

T>G

6 mutation classes

ACA>ATA
ACC>ATC
ACG>ATG
ACT>ATT
CCA>CTA
CCC>CTC
CCG>CTG
CCT>CTT
GCA>GTA
GCC>GTC
GCG>GTG
GCT>GTT
TCA>TTA
TCC>TTC
TCG>TGT
TCT>TTT

ACA>AAA
ACC>AAC
ACG>AAG
ACT>AAT
CCA>CAA
CCC>CAC
CCG>CAG
CCT>CAT
GCA>GAA
GCC>GAC
GCG>GAG
GCT>GAT
TCA>TAA
TCC>TAC
TCG>TAG
TCT>TAT

ACA>AGA
ACC>AGC
ACG>AGG
ACT>AGT
CCA>CGA
CCC>CGC
CCG>CGG
CCT>CGT
GCA>GGA
GCC>GGC
GCG>GGG
GCT>GGT
TCA>TGA
TCC>TGC
TCG>TGG
TCT>TGT

ATA>AAA
ATC>AAC
ATG>AAG
ATT>AAT
CTA>CAA
CTC>CAC
CTG>CAG
CTT>CAT
GTA>GAA
GTC>GAC
GTG>GAG
GTT>GAT
TTA>TAA
TTC>TAC
TTG>TAG
TTT>TAT

ATA>ACA
ATC>ACC
ATG>ACG
ATT>ACT
CTA>CCA
CTC>CCC
CTG>CCG
CTT>CCT
GTA>GCA
GTC>GCC
GTG>GCG
GTT>GCT
TTA>TCA
TTC>TCC
TTG>TCG
TTT>TCT

ATA>AGA
ATC>AGC
ATG>AGG
ATT>AGT
CTA>CGA
CTC>CGC
CTG>CGG
CTT>CGT
GTA>GGA
GTC>GGC
GTG>GGG
GTT>GGT
TTA>TGA
TTC>TGC
TTG>TGG
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

