

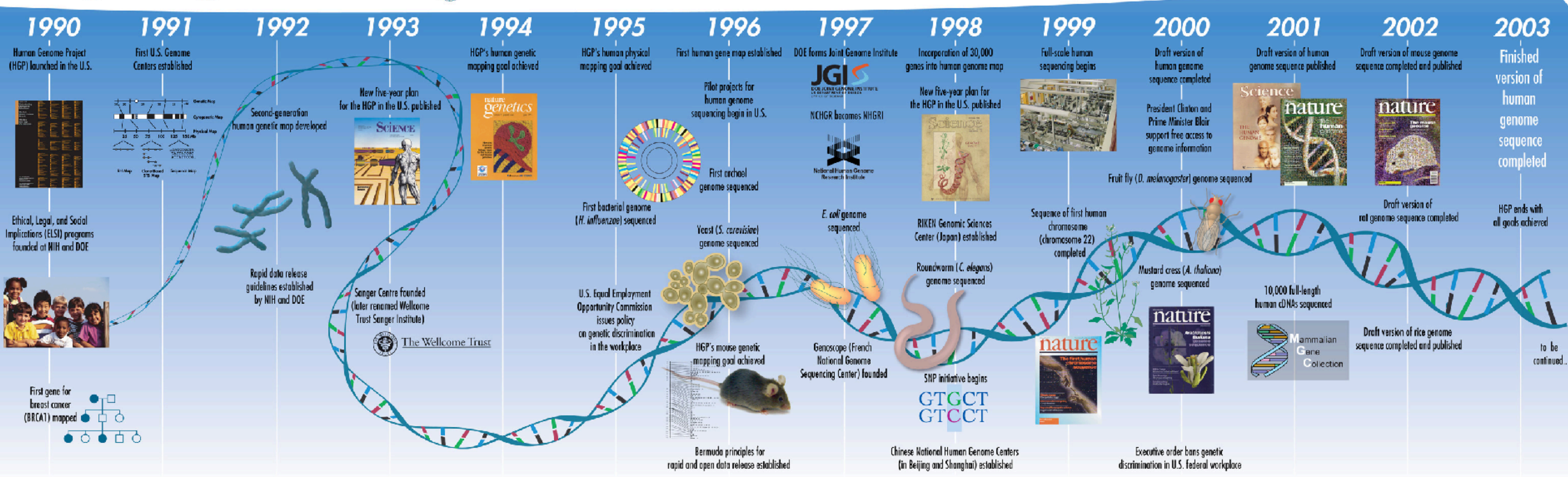
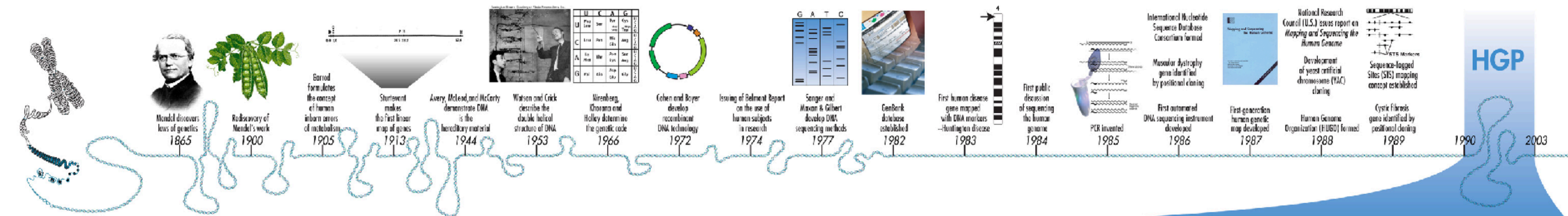


Bioinformatics for cancer omics data

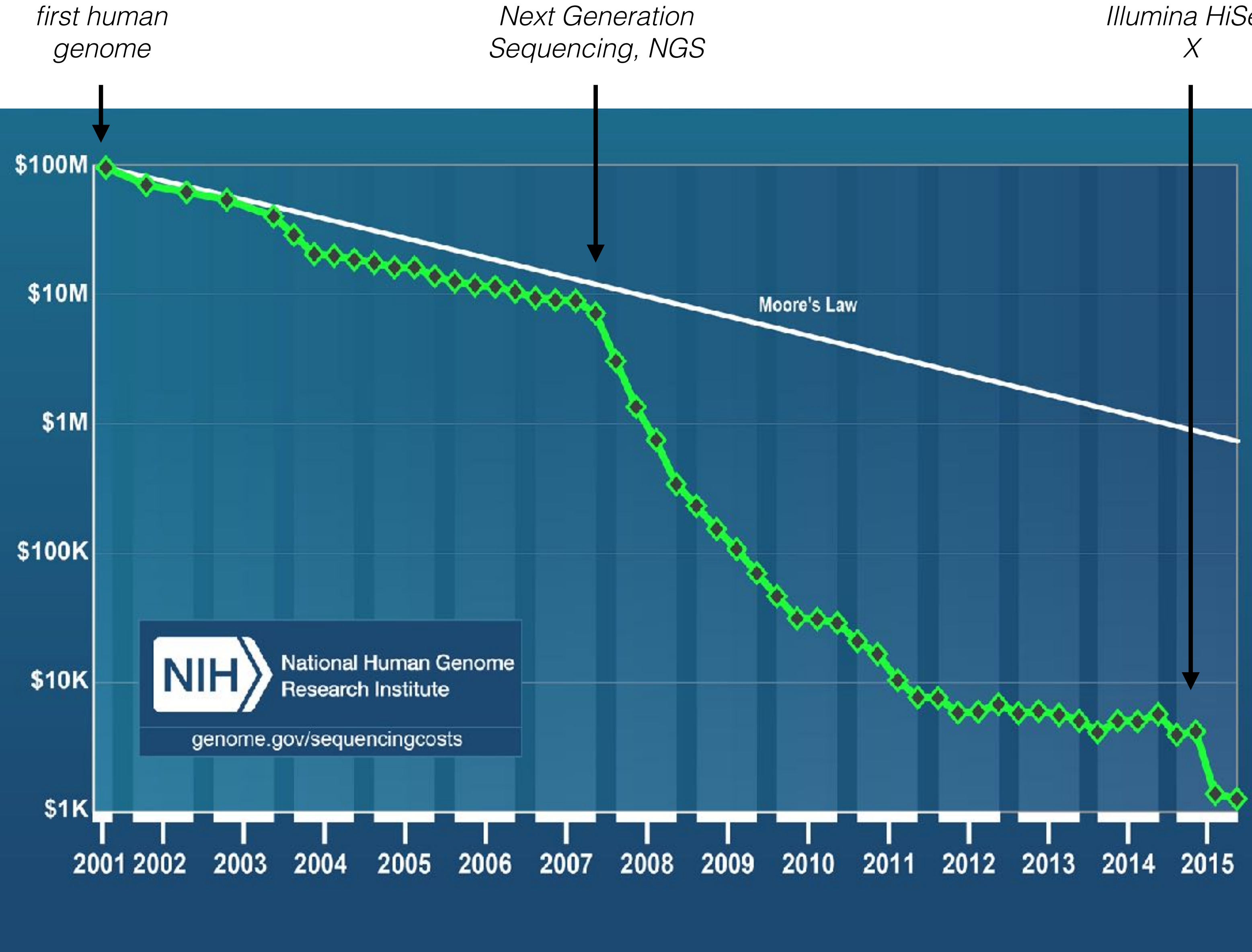
Matteo Benelli, PhD
Hospital of Prato
Prato, Italy

106° Congresso Nazionale SIF
18 Settembre 2020

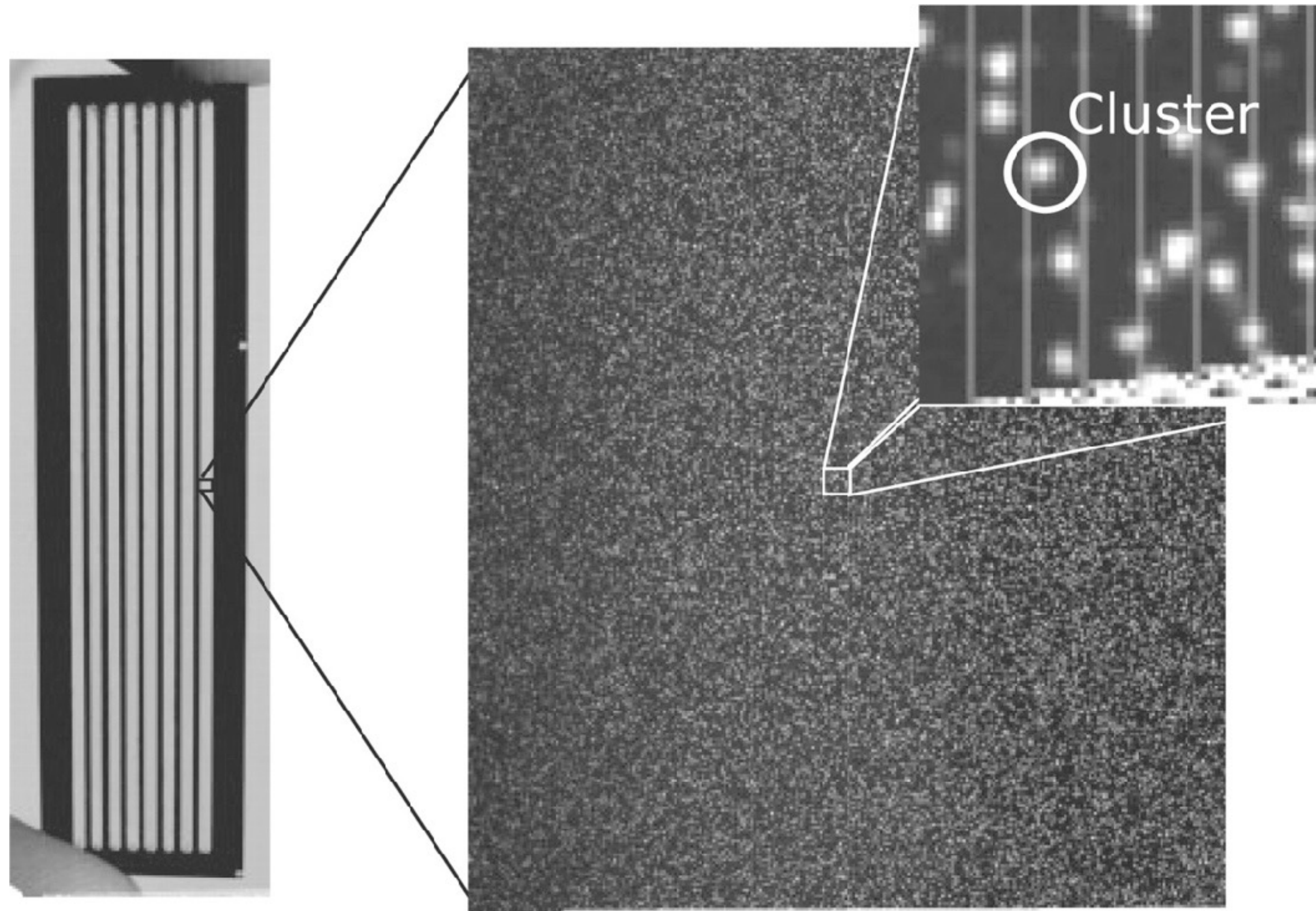
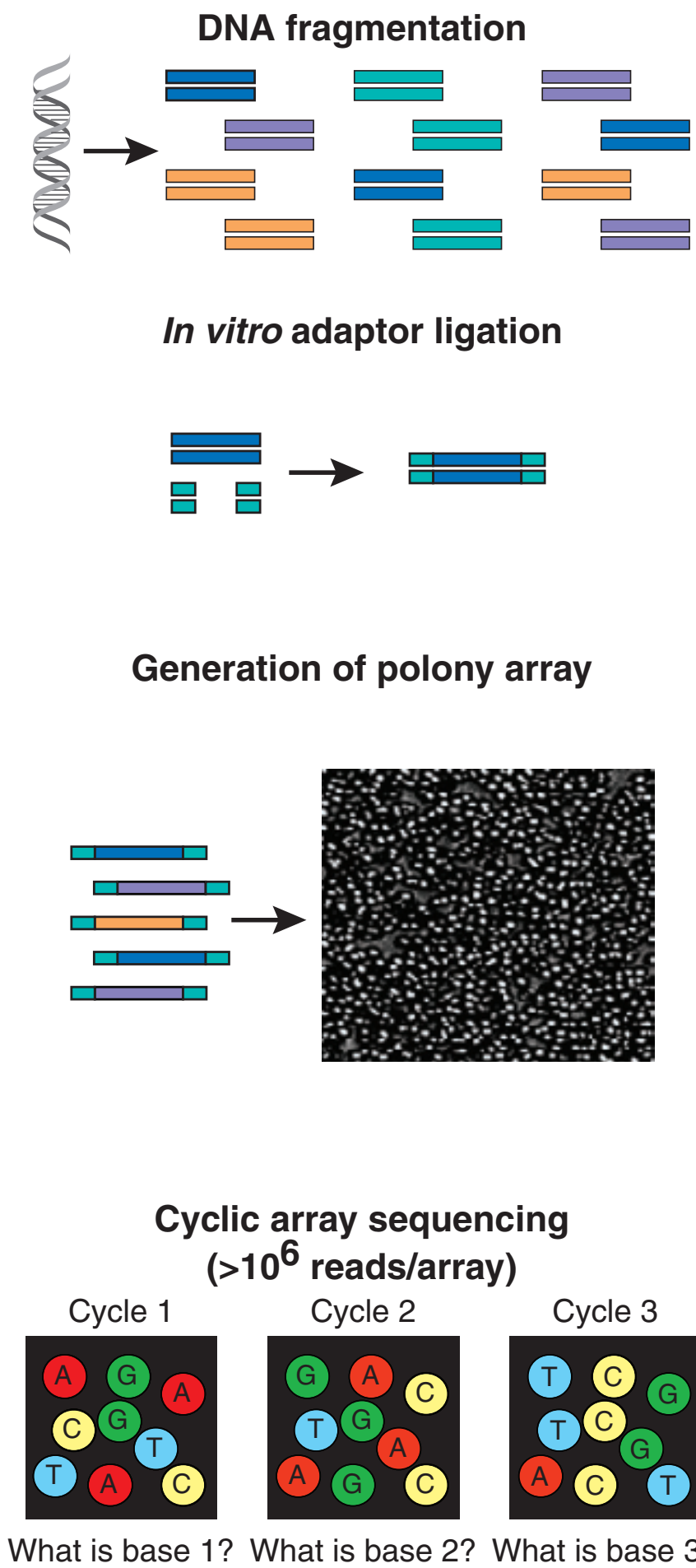
The human genome project



Enabling technologies: high throughput sequencing

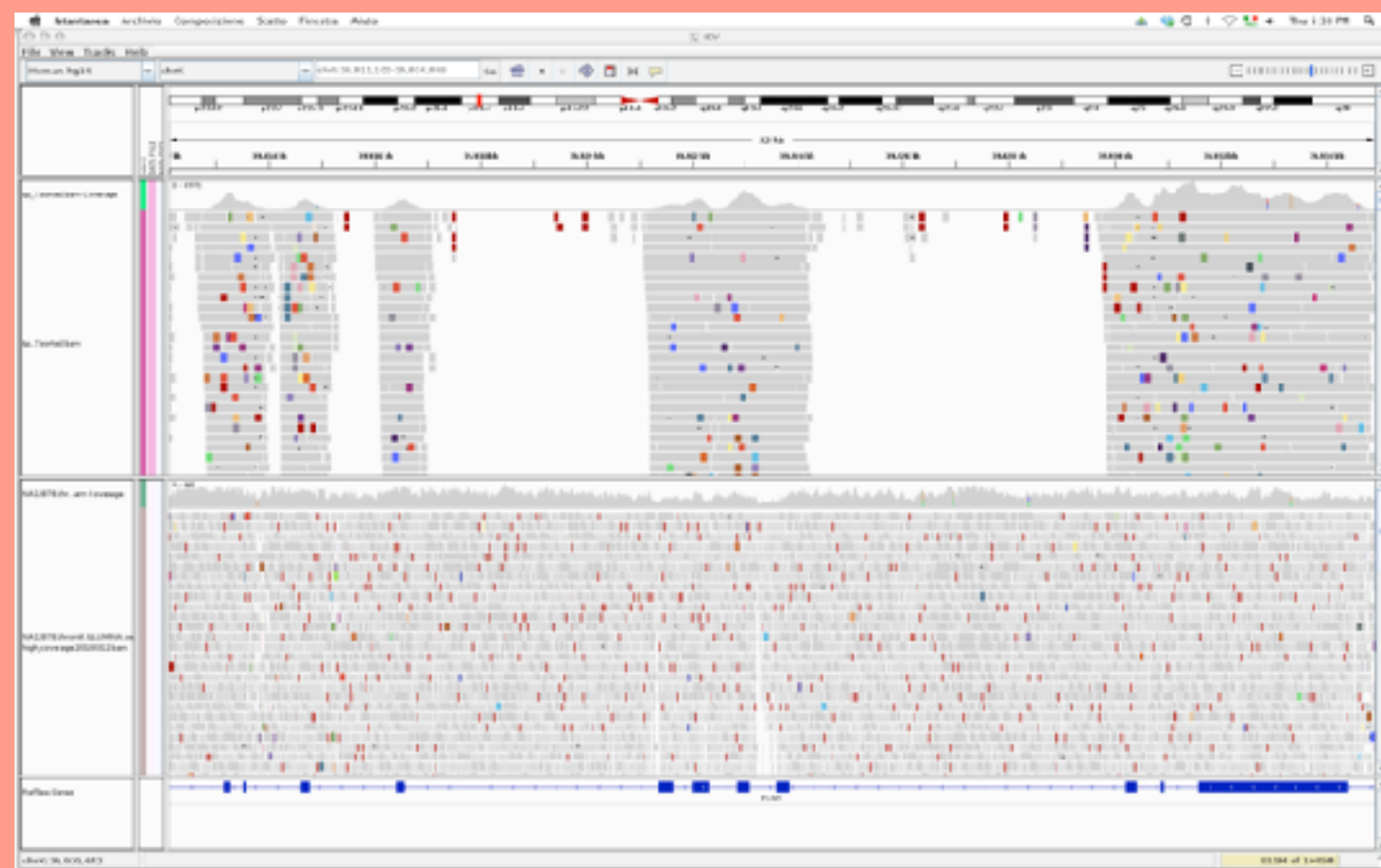


High throughput sequencing

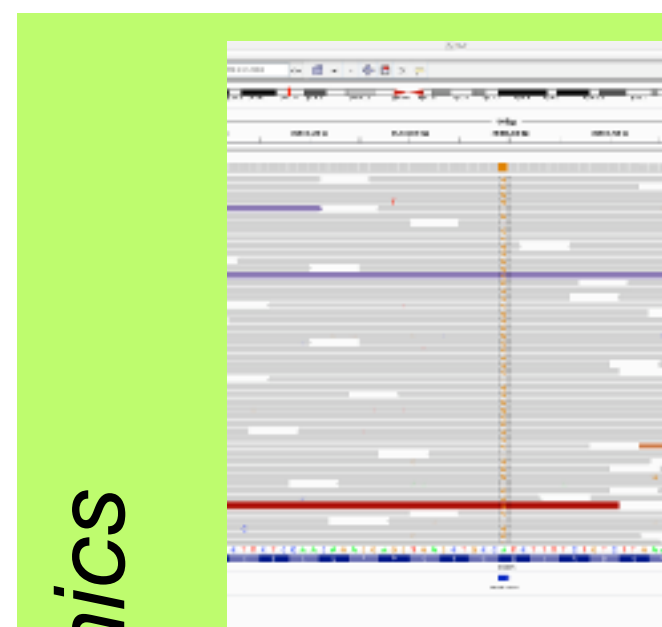


1 cluster ~ 1 sequence read

High throughput sequencing: applications

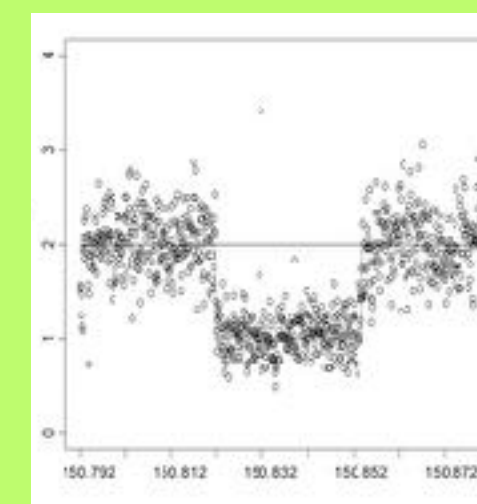


mapped reads



SNV and indels detection
frequentistic, probabilistic,
split read

genomics

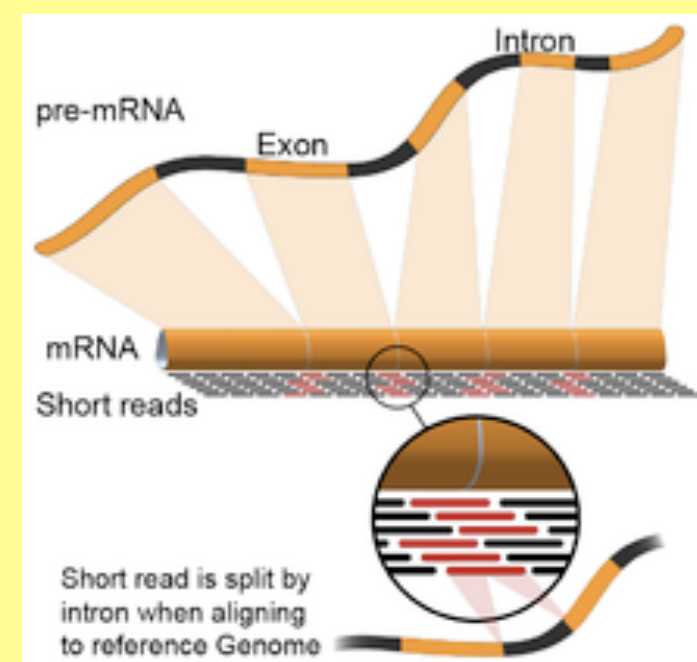


SV detection
(CNV, translocation, inversion)
read count, split read,
paired-end

transcriptomics

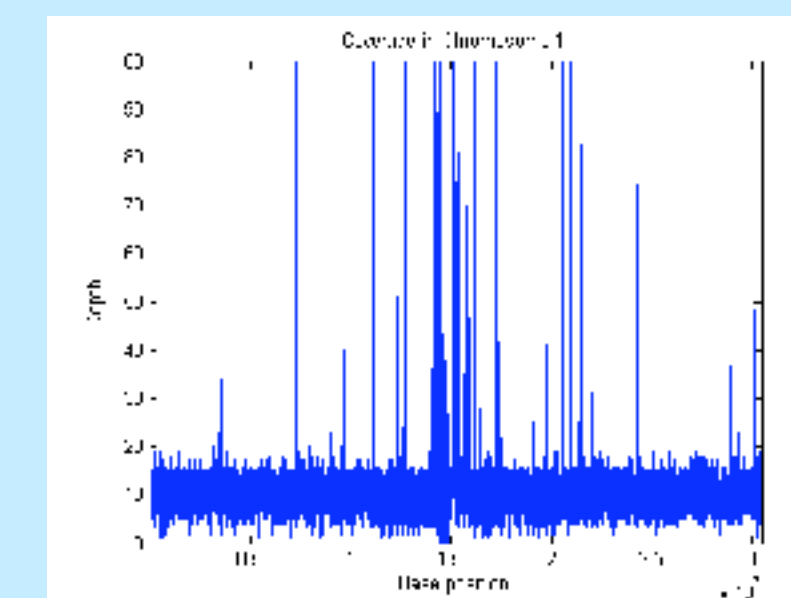
RNA seq

Gene expression,
transcriptome
assembly, gene
fusions, alternative
splicing



epigenomics

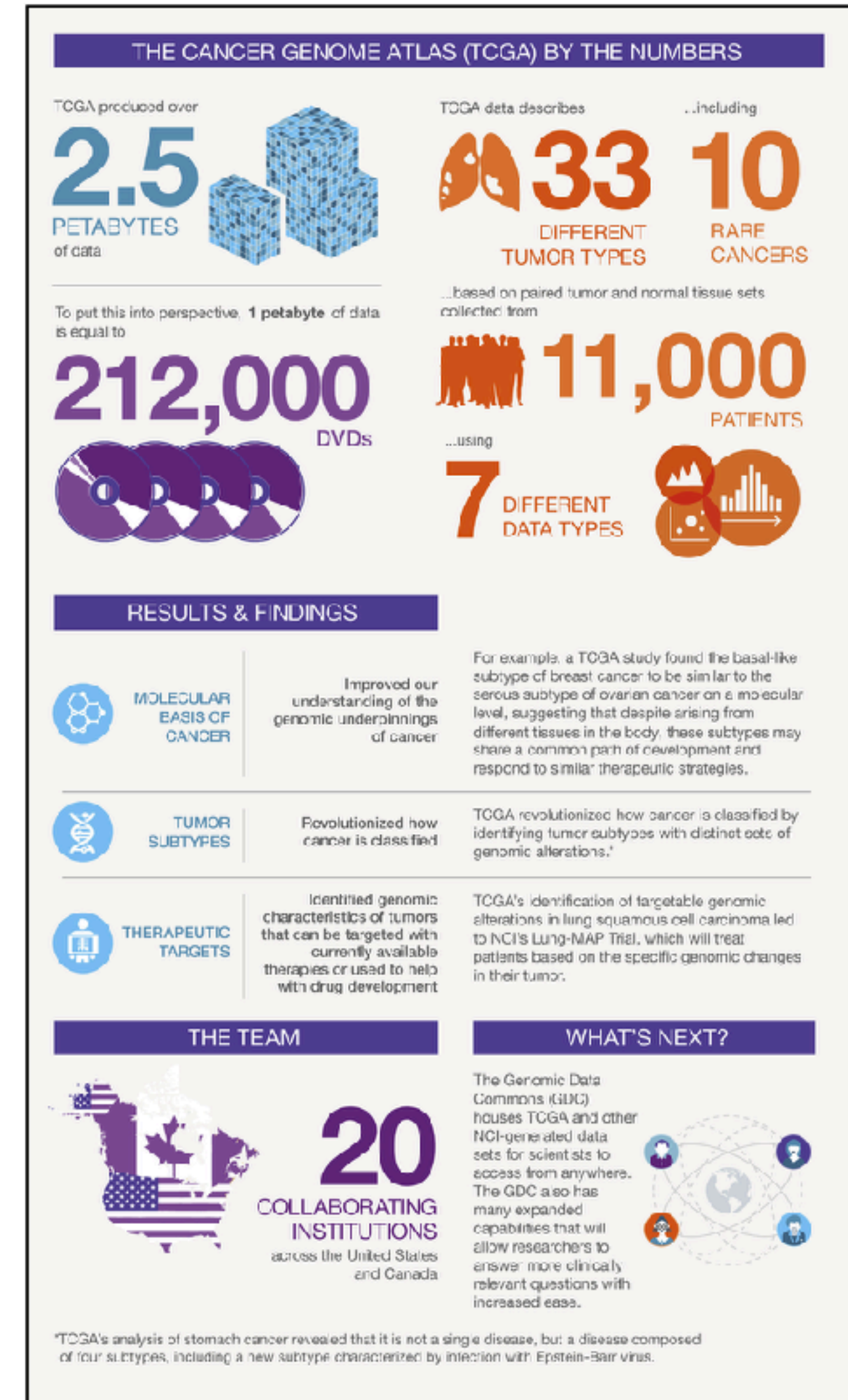
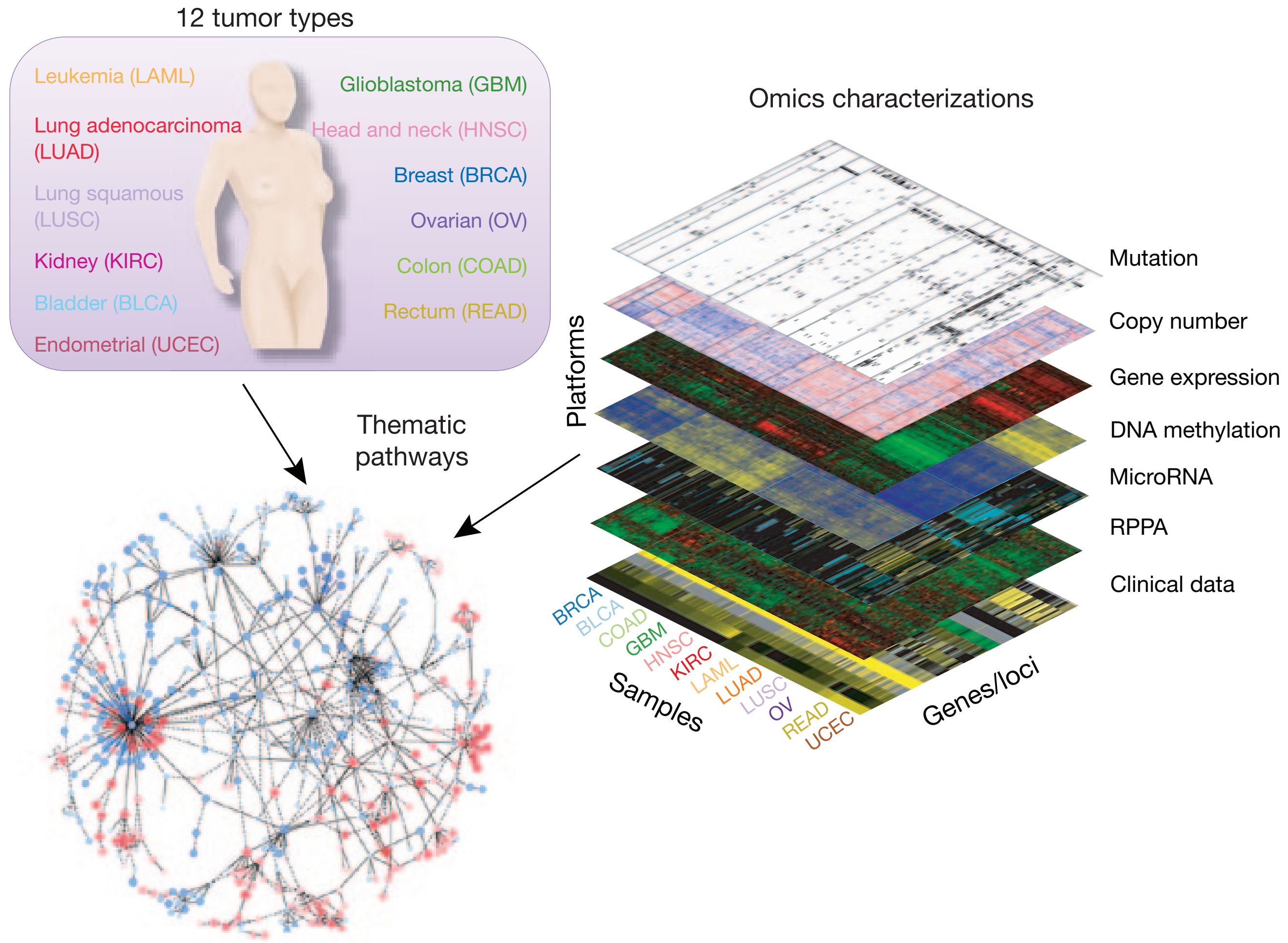
ChIP seq, MeDIP seq
peak finder
DNAmeth-seq



-omics and precision cancer medicine: TCGA initiative

2013

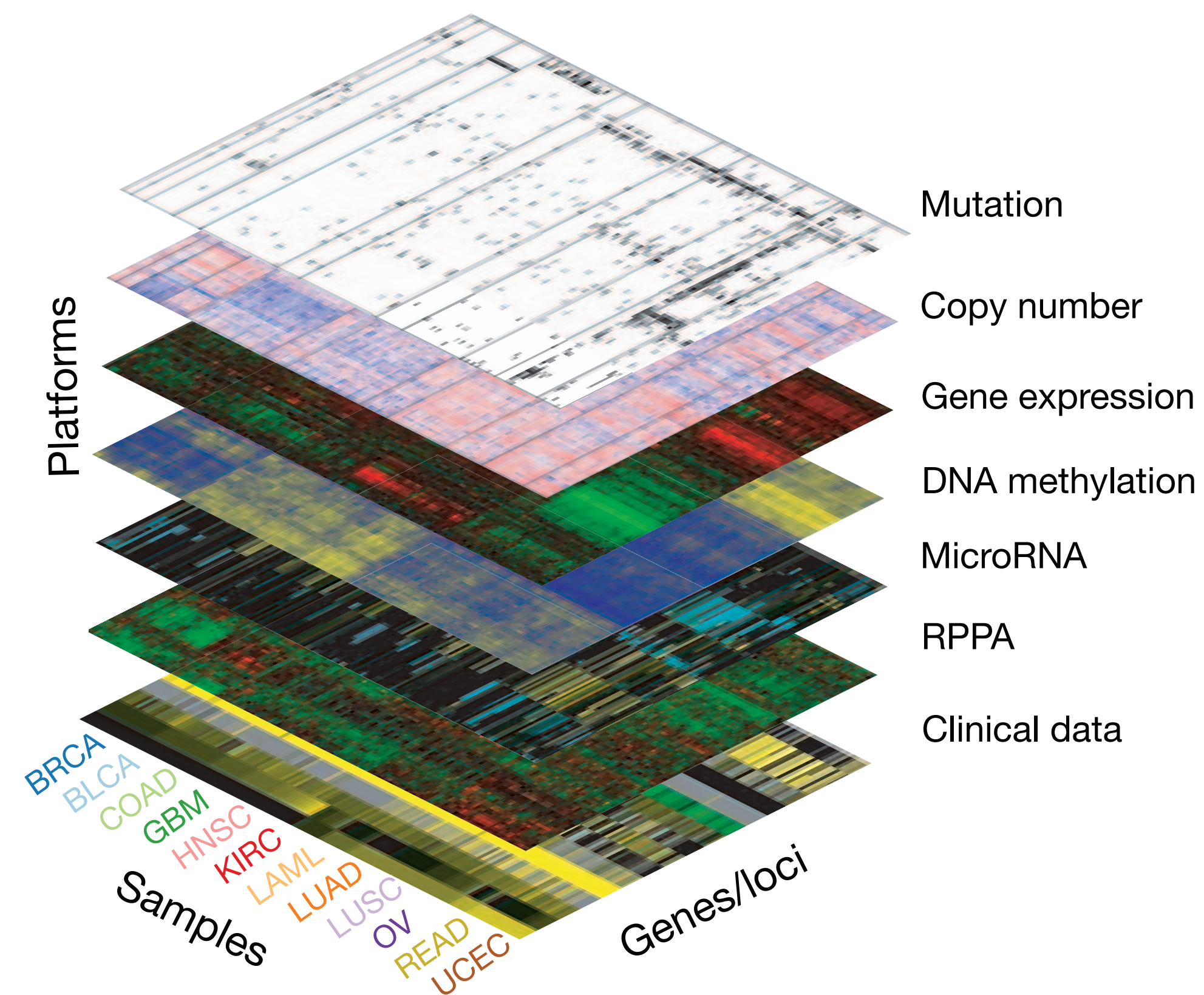
2019



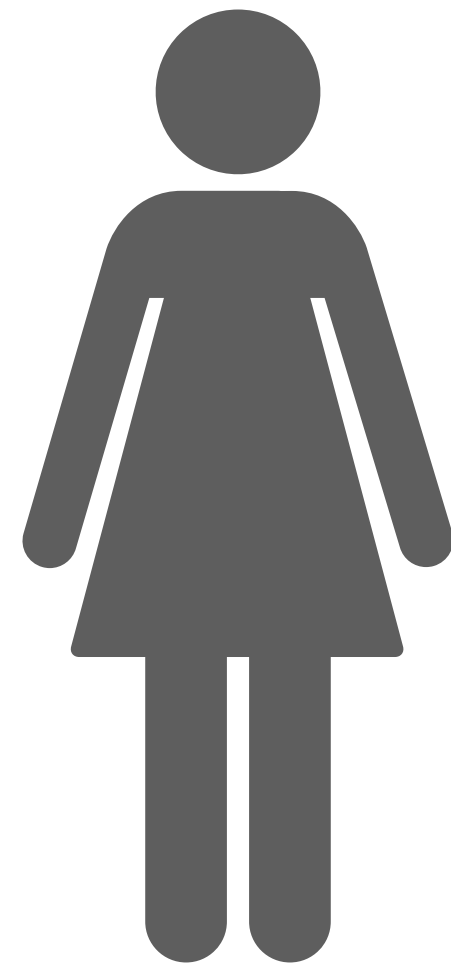
TCGA Res Network, *Nat Genet* (2013)
 Hutter, Zenklusen, *Cell* (2019)

Bioinformatics in cancer research

- Development of computational/statistical methods for the analysis of specific omics data
 - eg., aligner of sequences, identification of differentially expressed genes, etc.
- Development/set-up/application of data analysis pipelines
- Development/application of methods for the integrative analysis of multiple omics layers
- Modeling of cancer initiation, evolution and progression
- Identification of prognostic/predictive biomarkers (outcome, response/resistance to drugs)
-



Data scientist



*Bioinformatics specialist /
computational biologist*

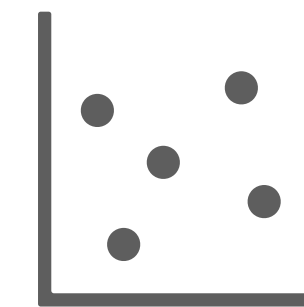
Big data



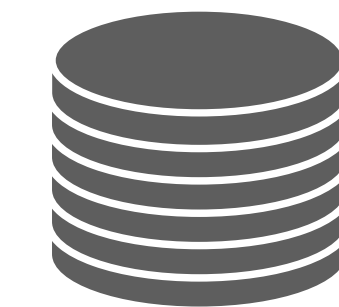
Programming



Visualisation



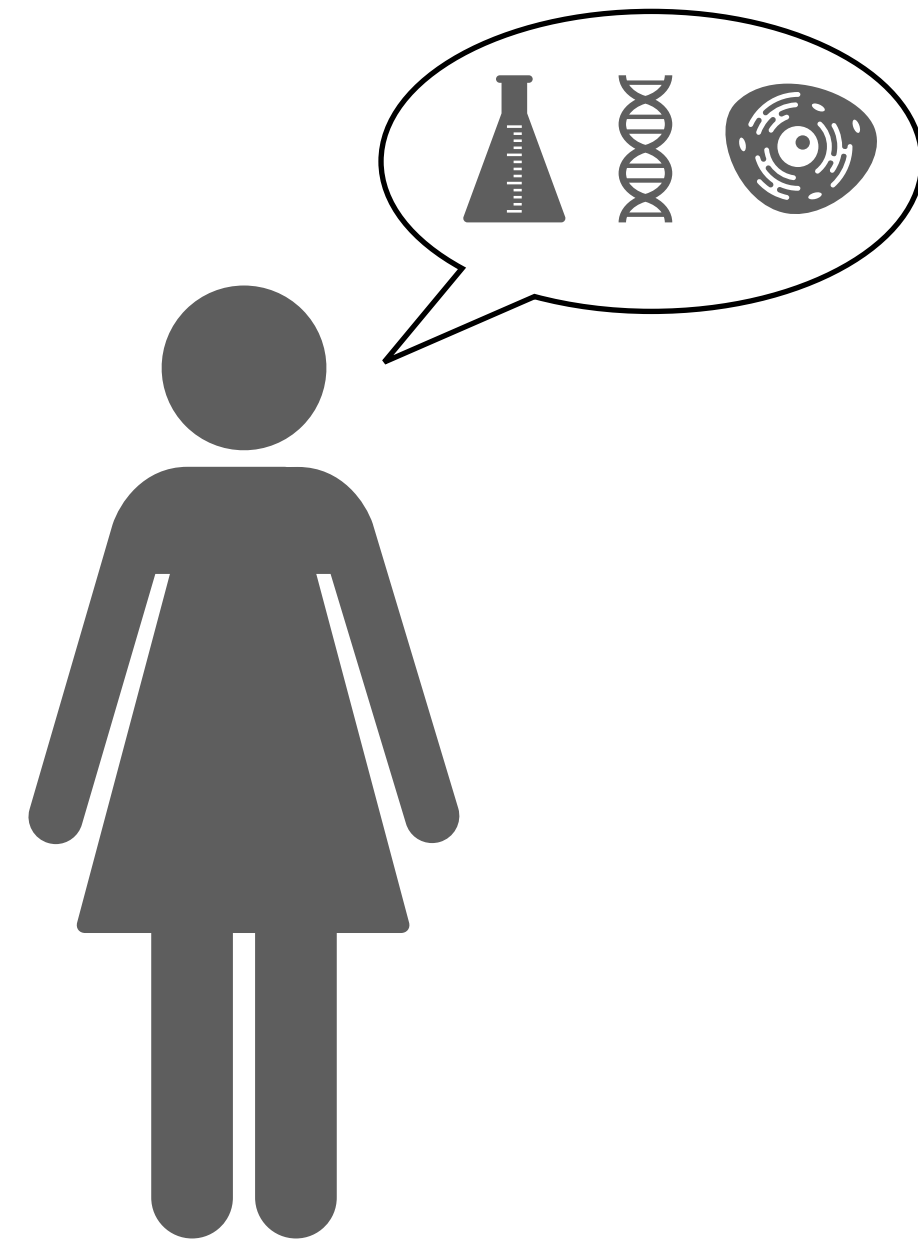
Statistics



Data handling

Omics data

Data scientist

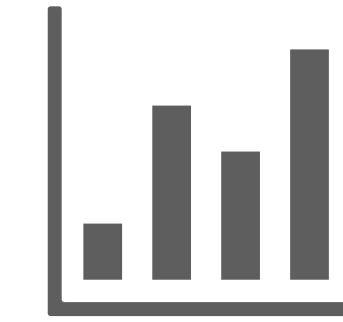


*Bioinformatics specialist /
computational biologist*

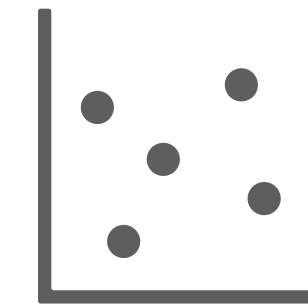
Big data



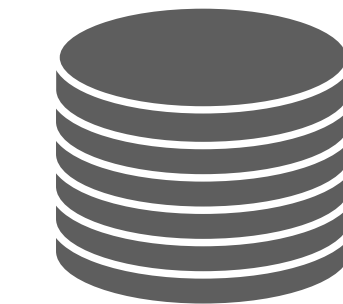
Programming



Visualisation



Statistics

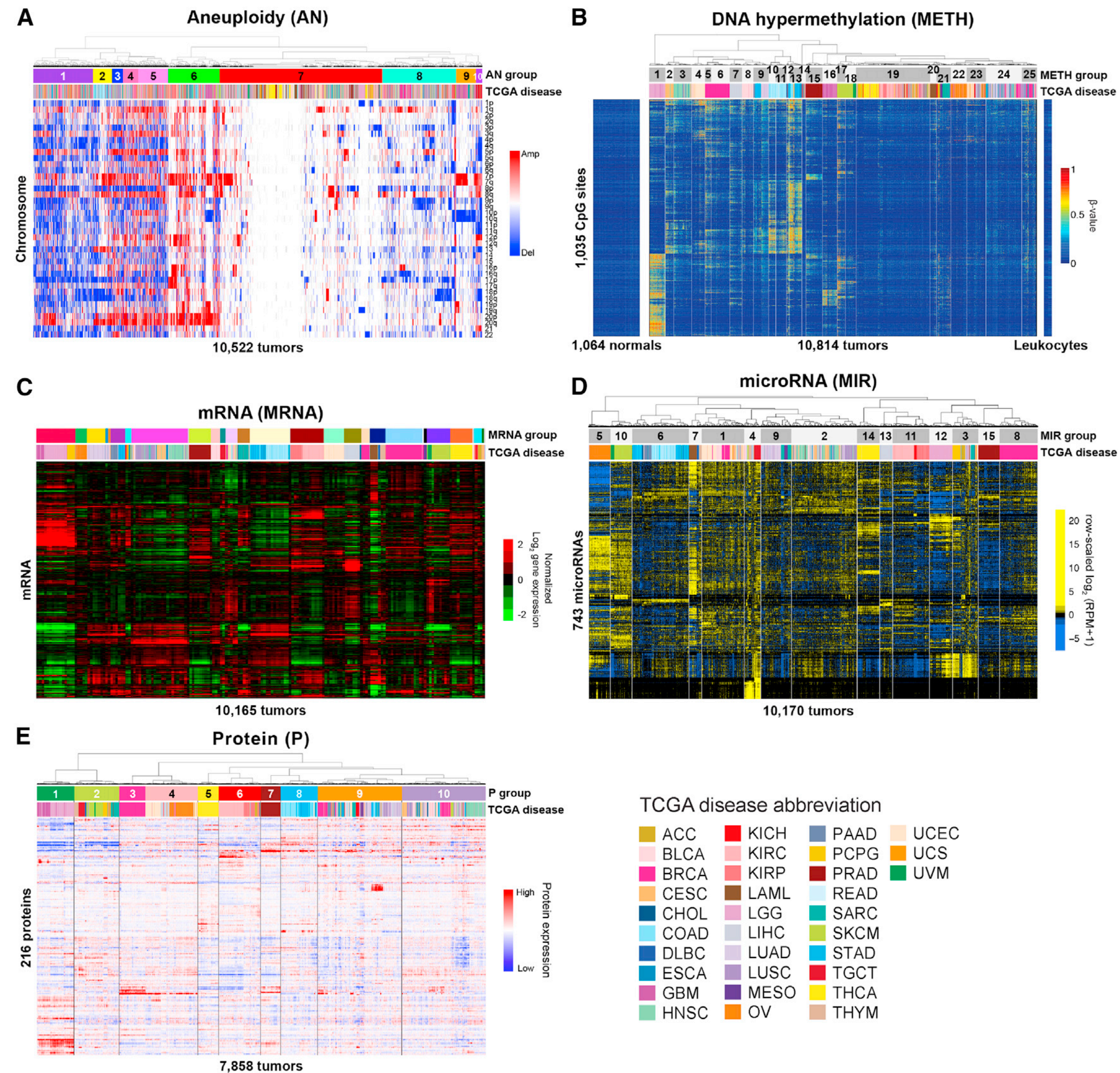


Data handling

Omics data

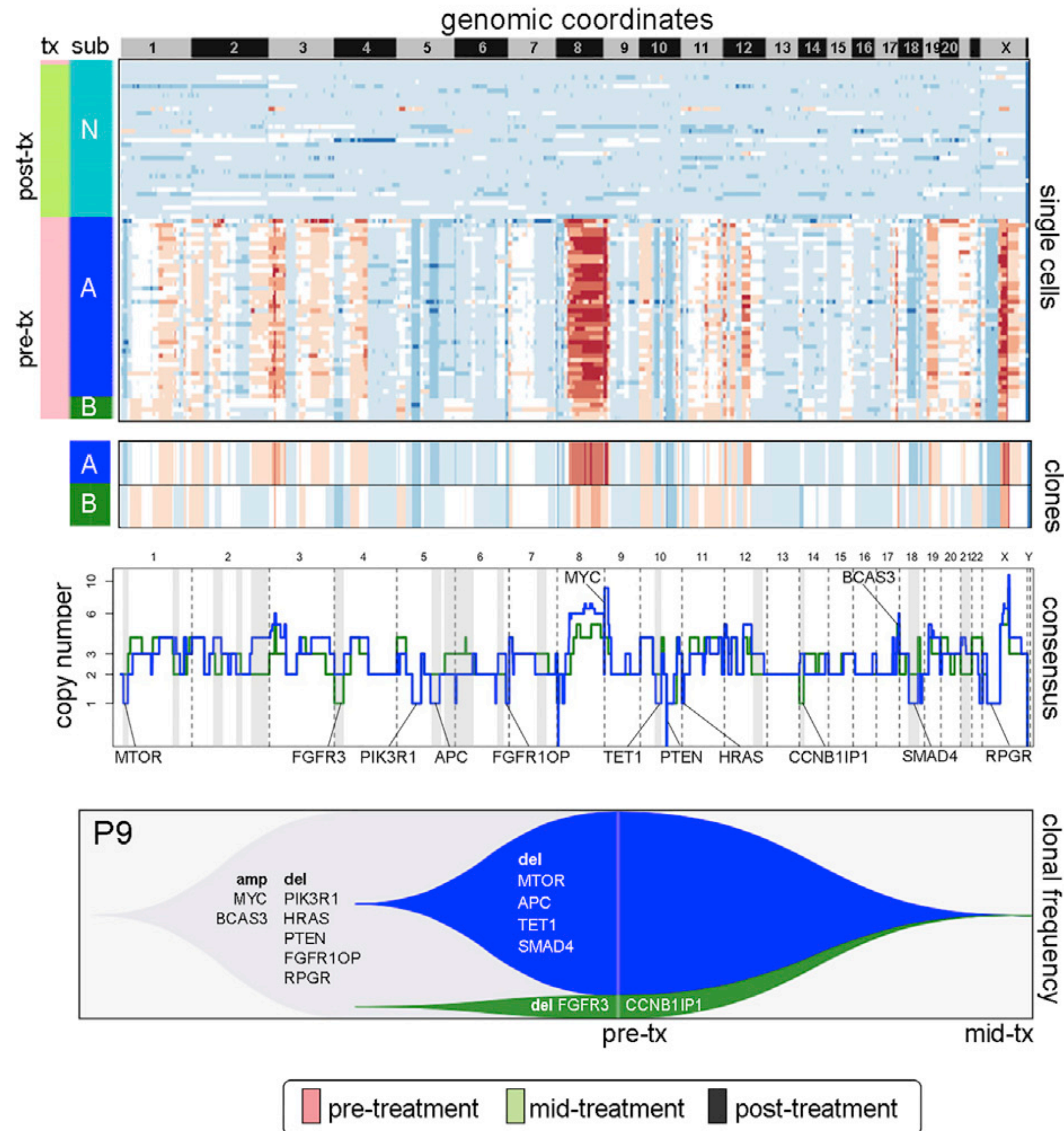
a personal (non-exhaustive) selection of notable works

Cell-of-origin patterns across multiple tumour types by different molecular layers

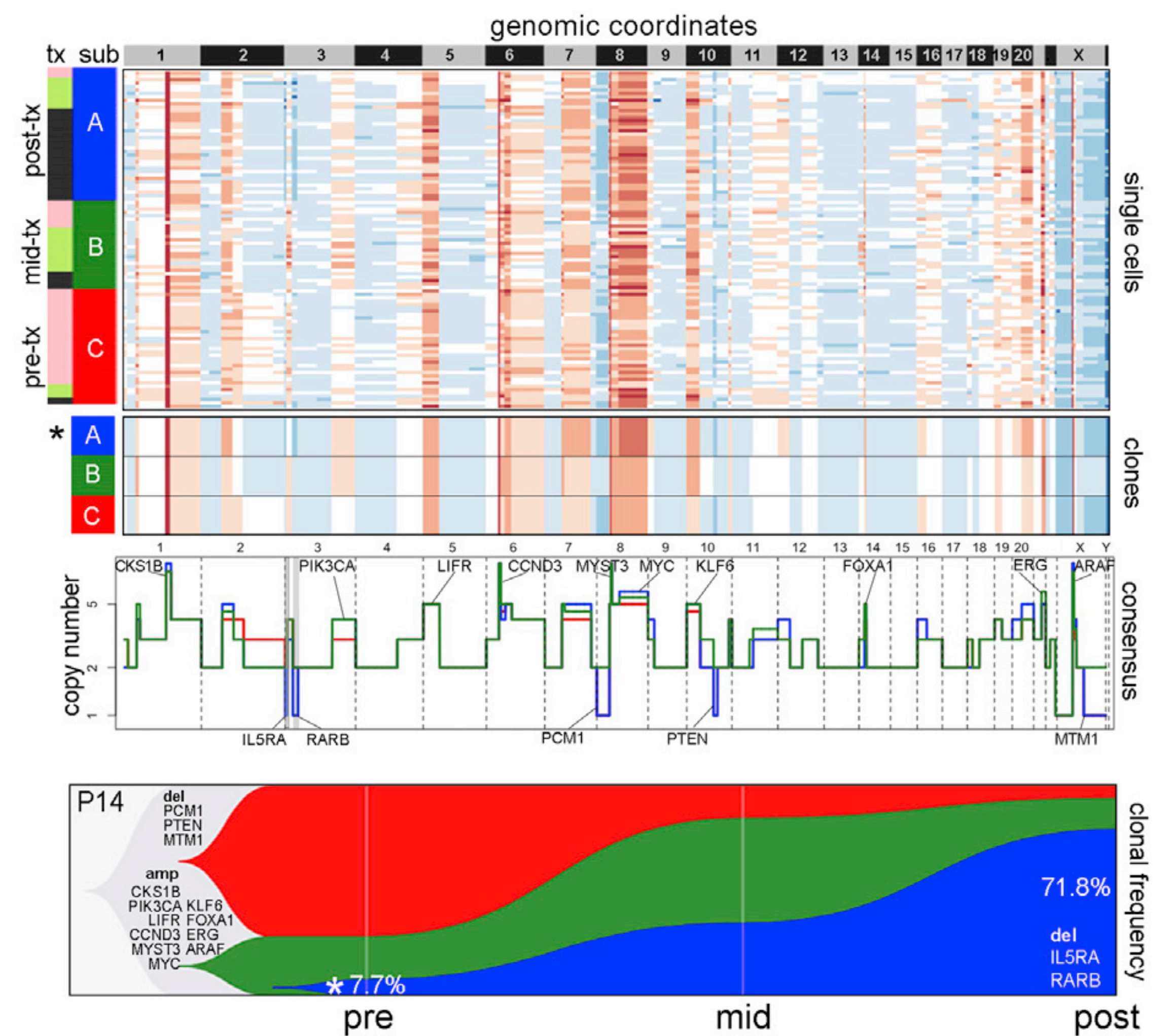


Chemoresistance evolution in breast cancer by single-cell sequencing

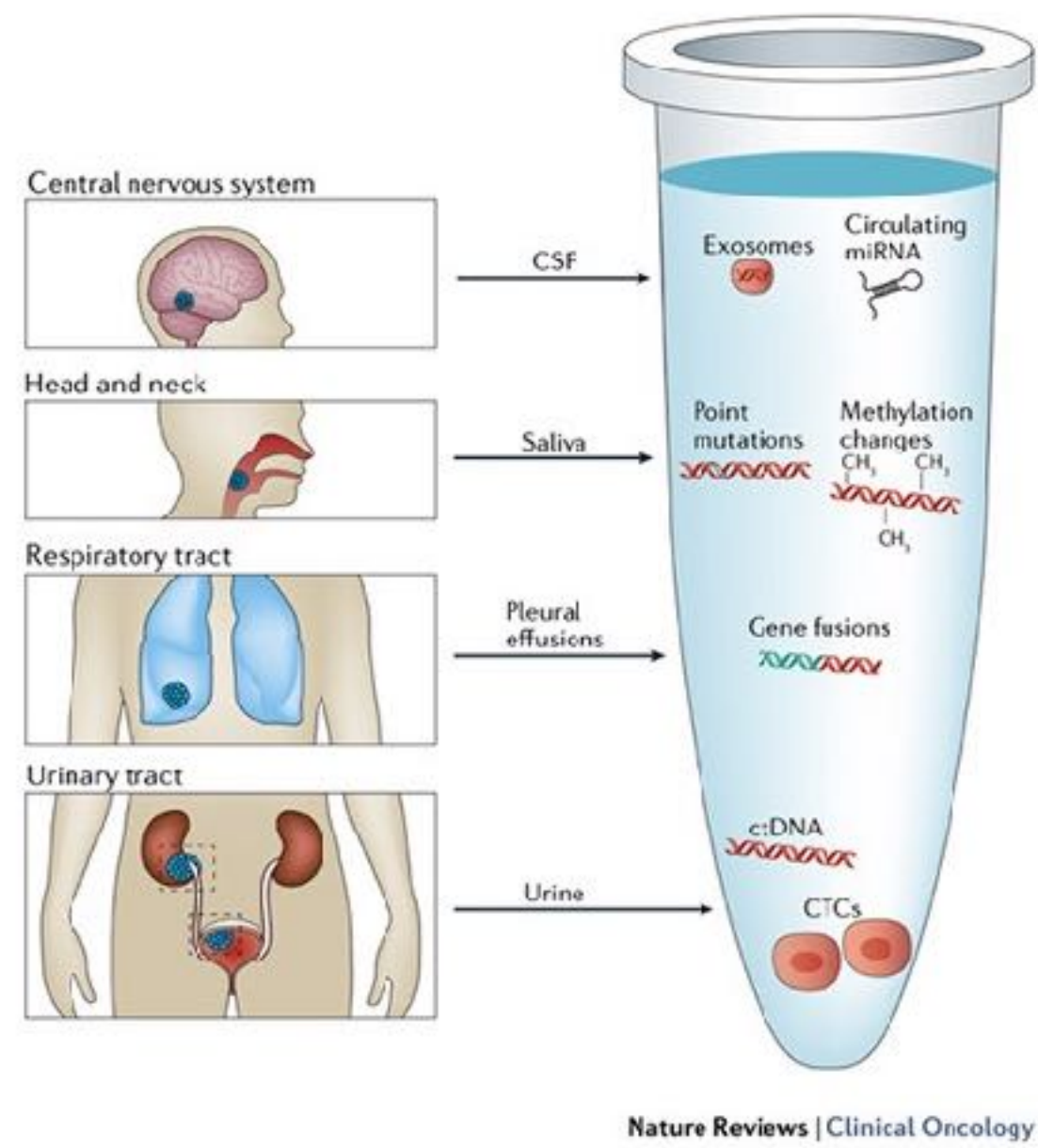
Response



Resistance (clonal expansion)



Liquid biopsy analysis for non-invasive detection and monitoring of cancer

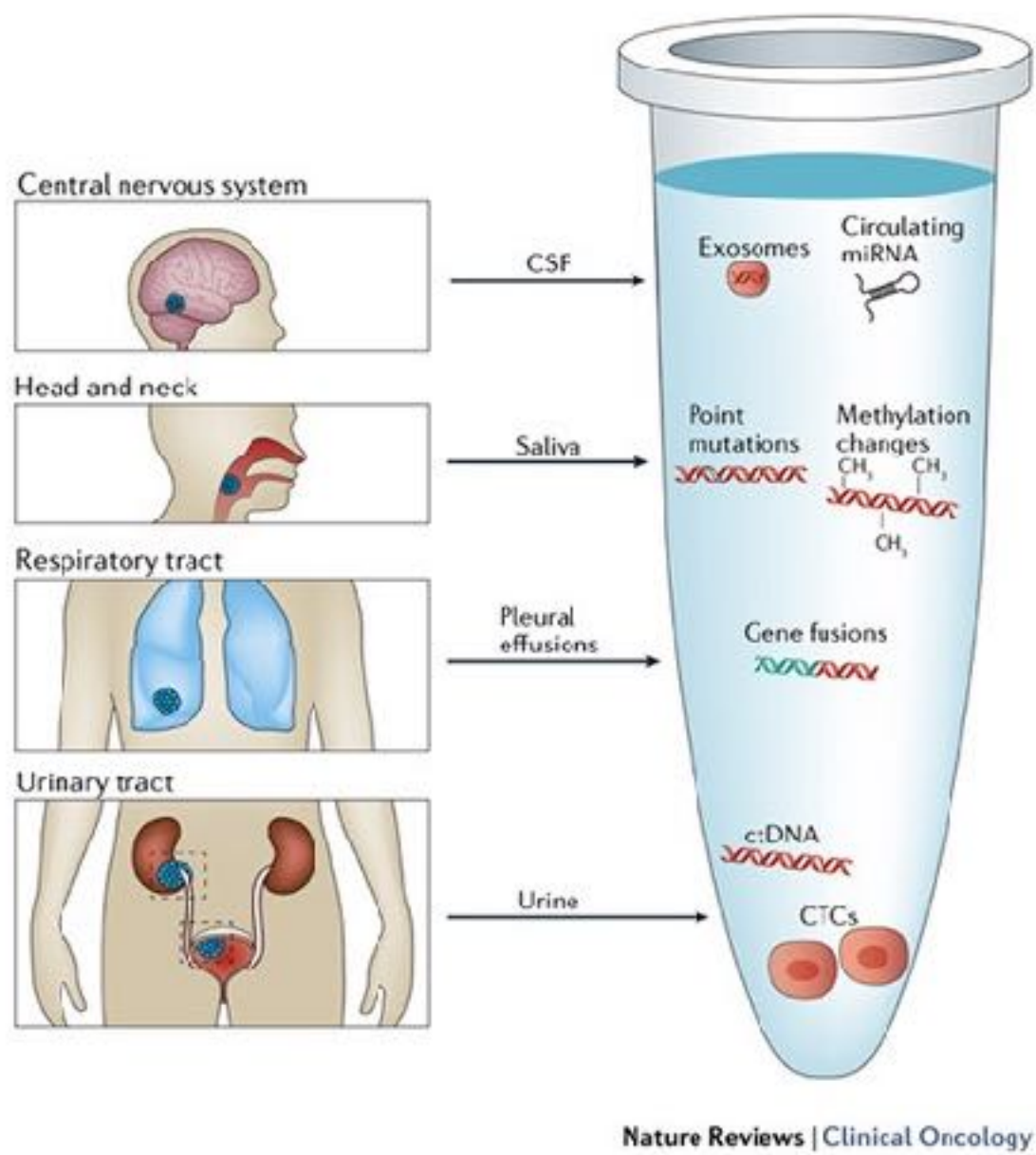


Siravegna et al., *Nat Rev Clin Oncol* (2017)

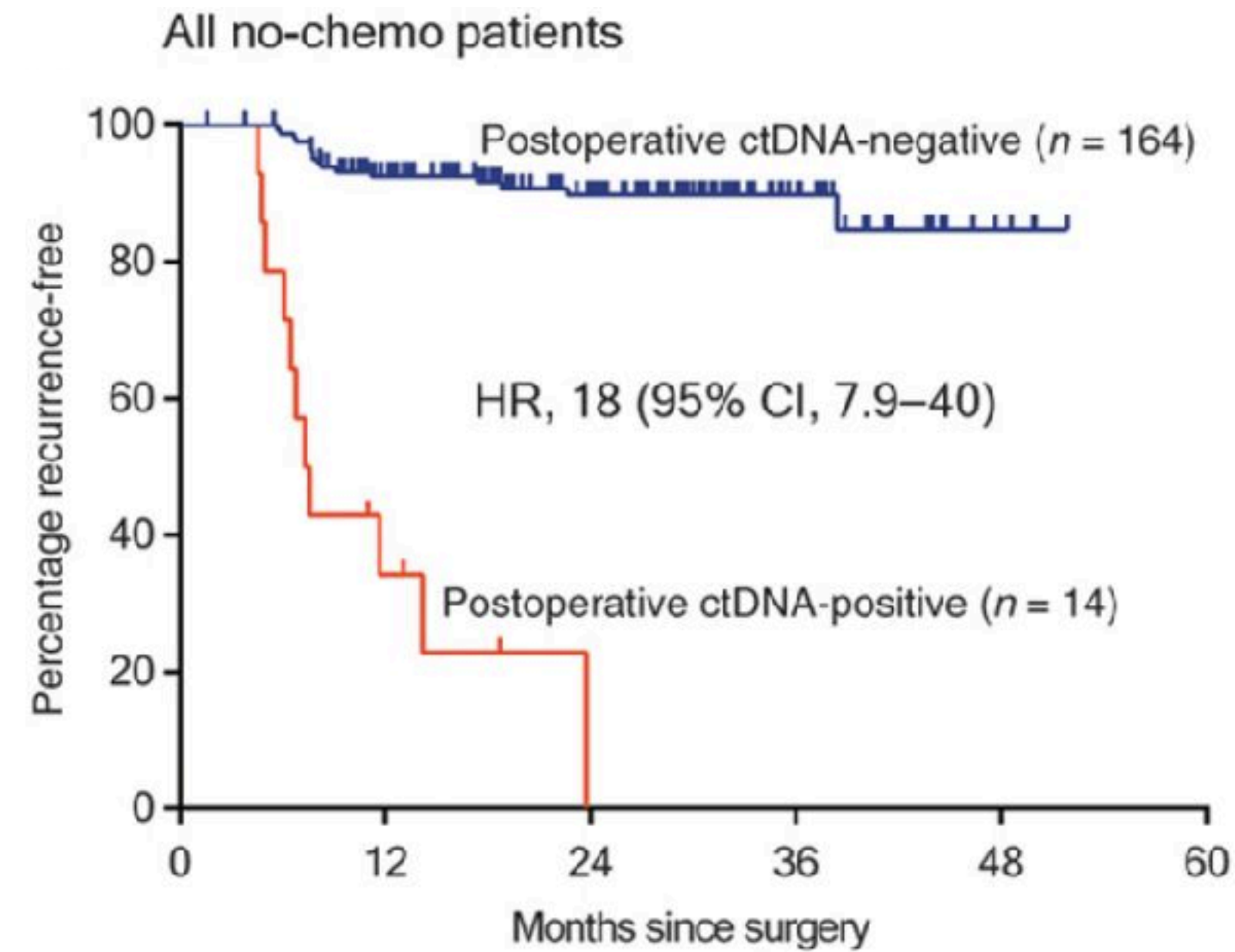
Tie et al., *Sci Transl Med* (2016)

Chen et al., *Nat Commun* (2020)

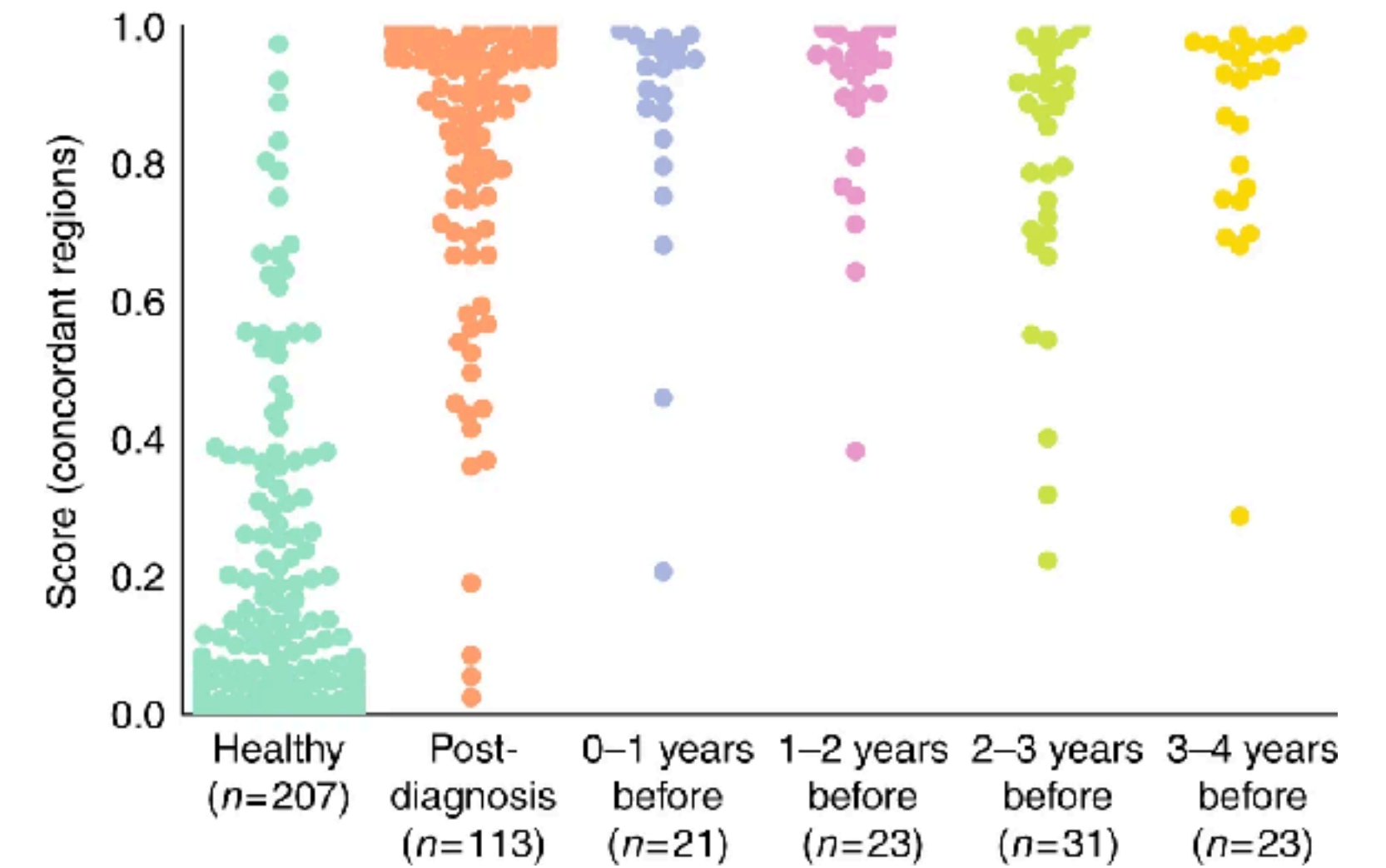
Liquid biopsy analysis for non-invasive detection and monitoring of cancer



Detection of minimal residual disease



Disease detection **four** years before conventional diagnosis



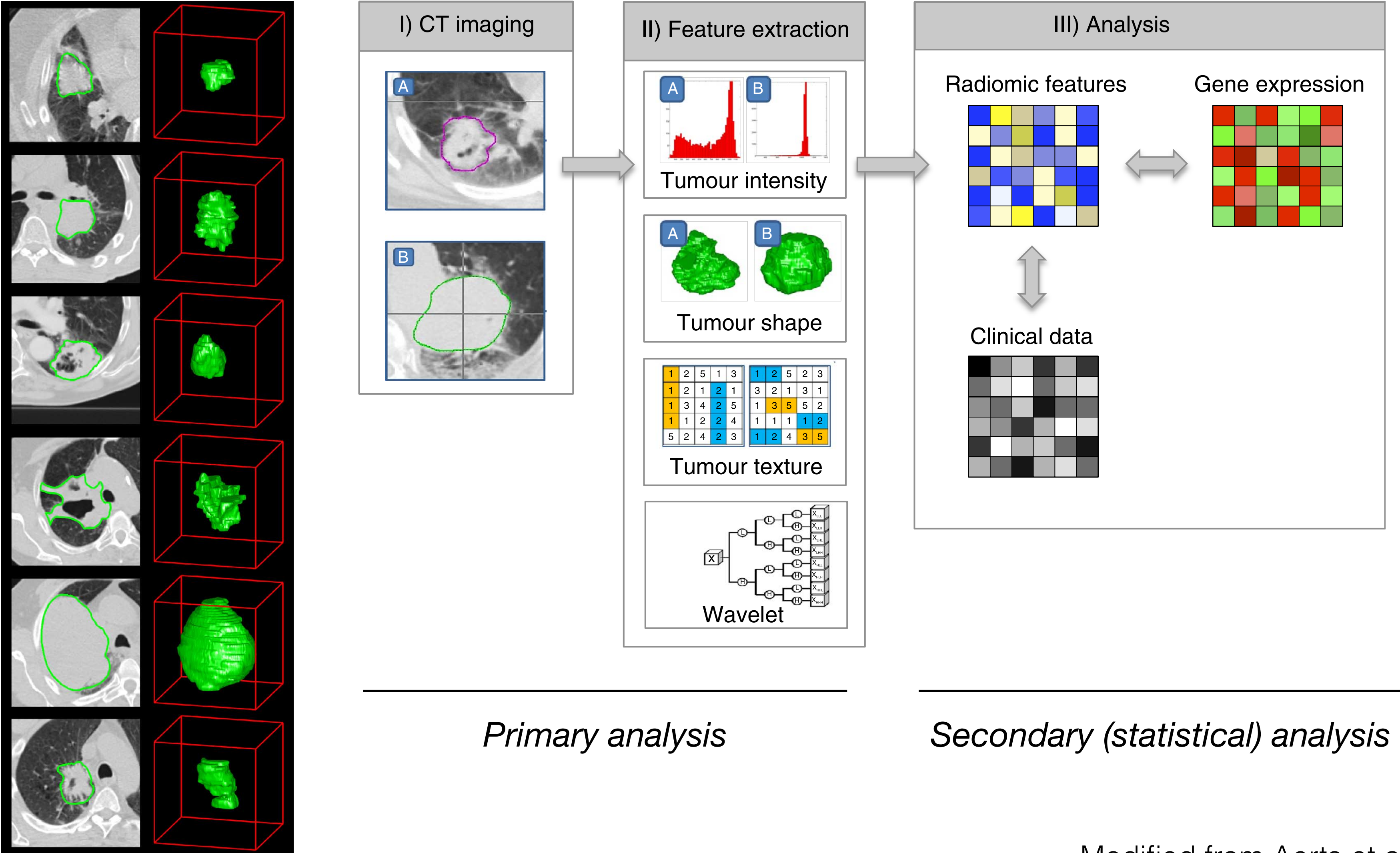
Siravegna et al., *Nat Rev Clin Oncol* (2017)

Tie et al., *Sci Transl Med* (2016)

Chen et al., *Nat Commun* (2020)

Radiomics

- Quantitative analysis of diagnostic images from CT, MRI and PET techniques



RadAR: Radiomics Analysis with R

- RadAR is a package for R to perform *secondary* analysis of radiomic features
- Manually curated dictionary of radiomic features to facilitate interpretation
- Detailed step-by-step tutorial
- freely available under MIT license at <https://github.com/cgplab/RadAR>

Import

- 3DSlicer
 - PyRadiomics
 - LifeX
 - Generic table
- } manually curated dictionary of features

Pre-processing

- Scaling
- Normalization
- Outliers removal
- Filtering by image type
- Filtering by feature type

Statistics

- Feature selection by different strategies
- Unsupervised analysis by hierarchical clustering
- Differential radiomics
- Signature development and validation

Visualization

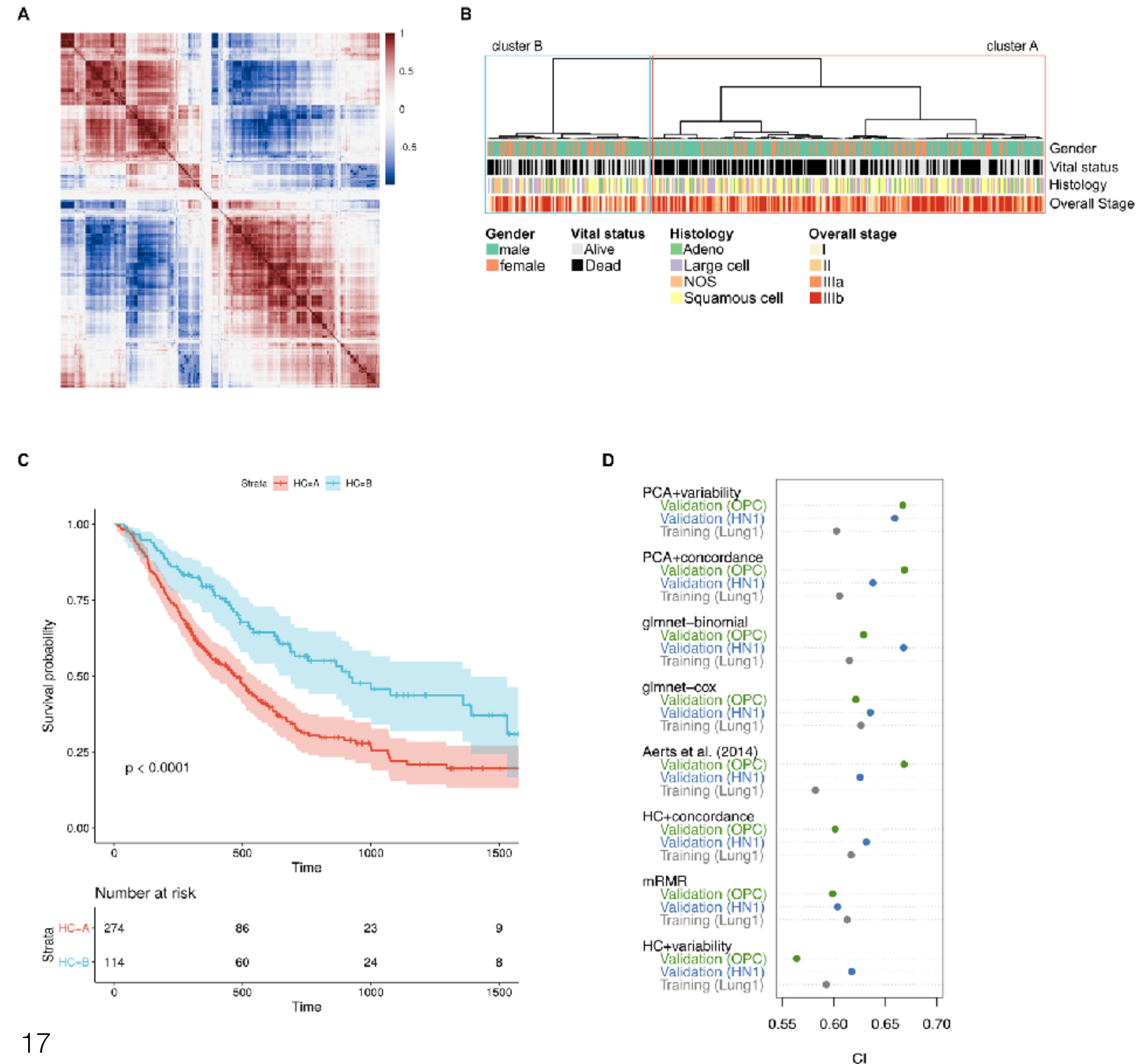
- heat map / correlation plot of correlation matrix
- heat map + clustering
- radiomic feature distribution across conditions

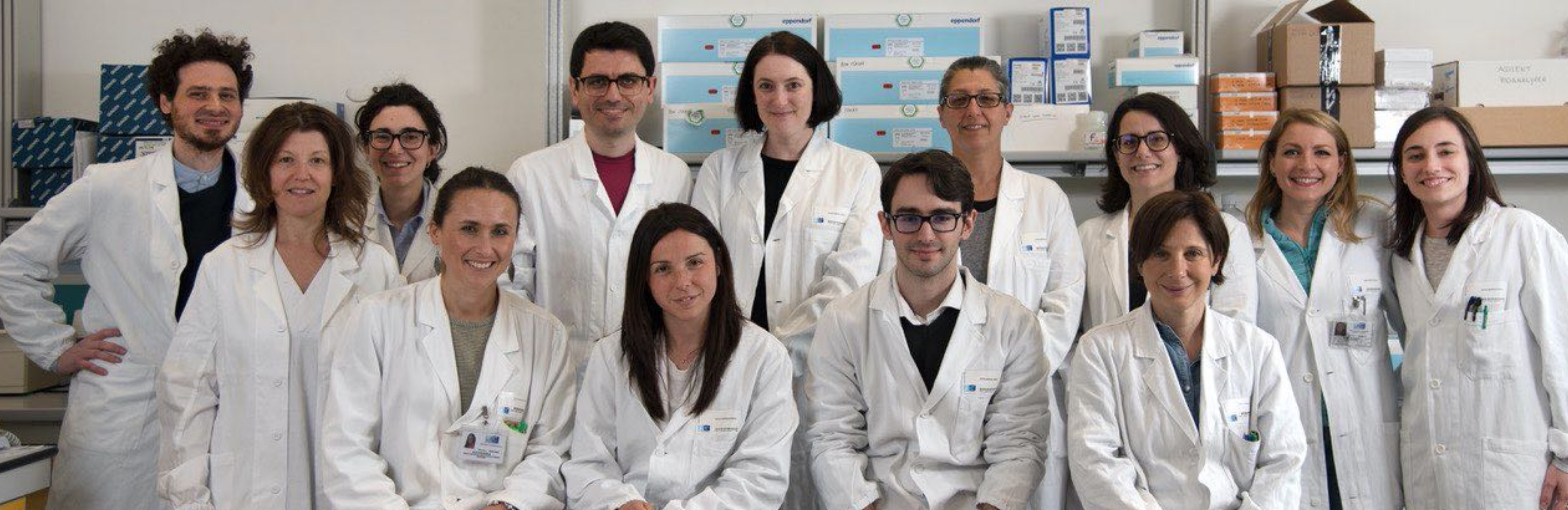
Output

- Top features
- Plots
- RadAR R objects

RadAR: Radiomics Analysis with R

- Radar was tested on the radiomic profiles of more than 850 cancer patients from publicly available datasets from The Cancer Imaging Archive (TCIA)
- RadAR was able to recapitulate expected results based on original findings
- Useful tool to train/assist medical physicists in radiomics analyses





matteo.benelli@uslcentro.toscana.it



FONDAZIONE
CR FIRENZE



Ministero della Salute

