

# Complexity and Peculiarities of Biomolecular Data in Agrifood Research

Alessandra Stella

Consiglio Nazionale delle Ricerche

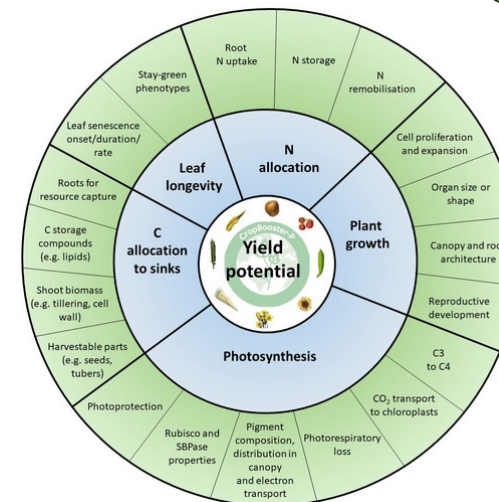
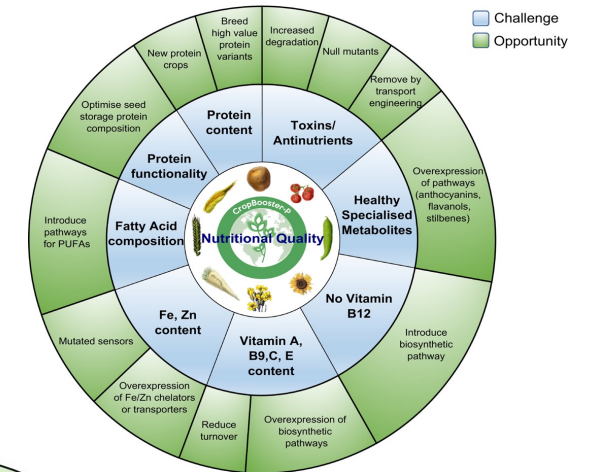
# The goals of biomolecular research in agriculture

Overall sustainability  
Economic + Environmental

Food security, safety and quality

Protein transition

Climate change mitigation



# Complexity of agricultural traits

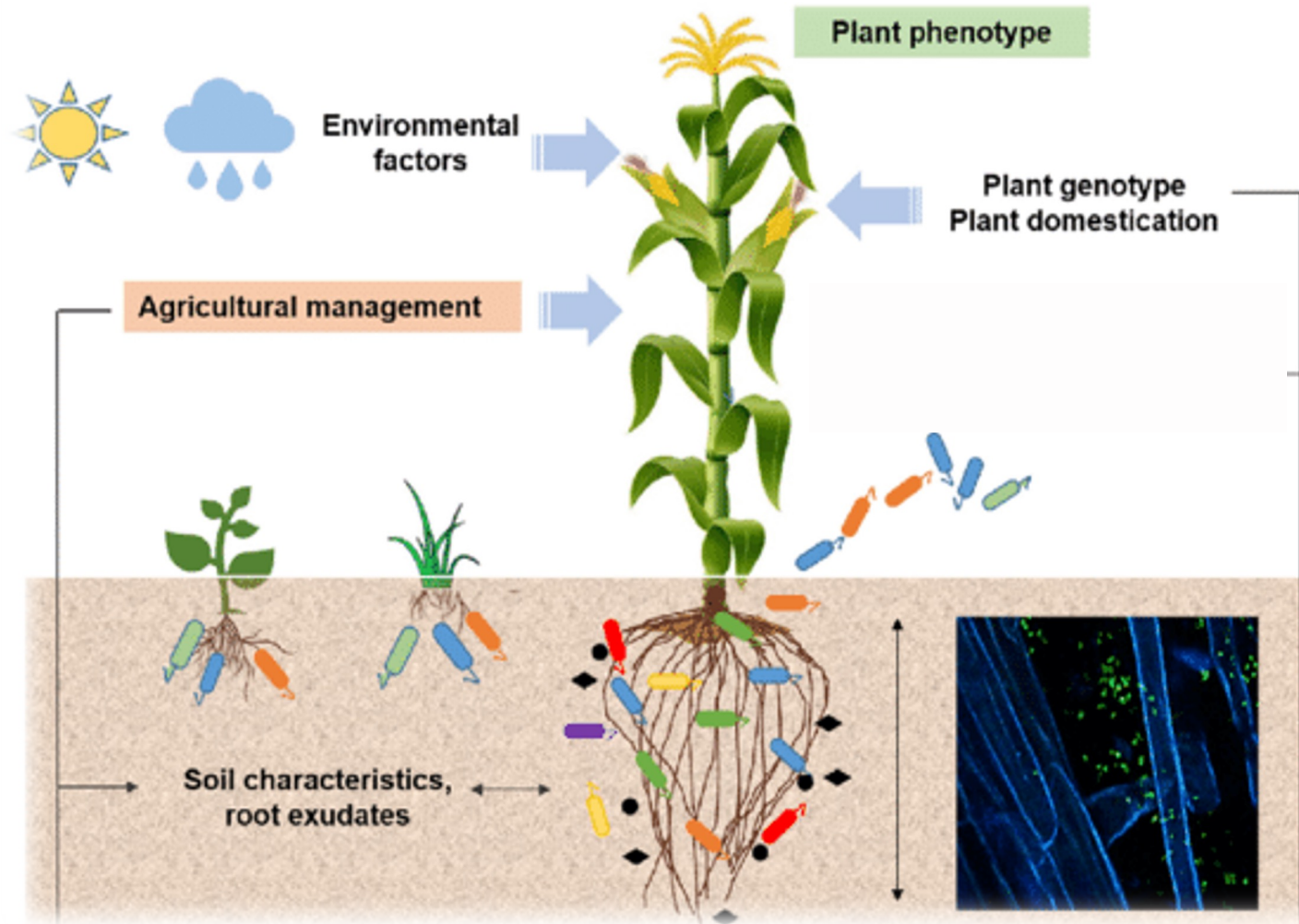
## Interactions

Among loci

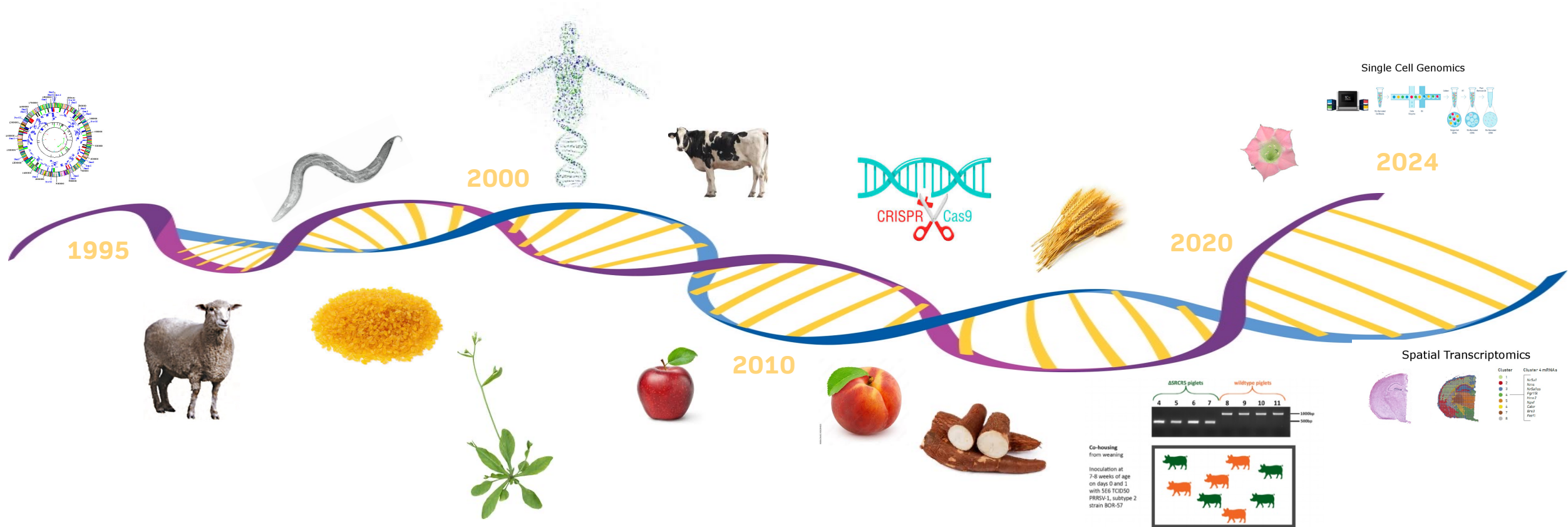
With biological factors

Between genotypes and environment/management

With microbiota



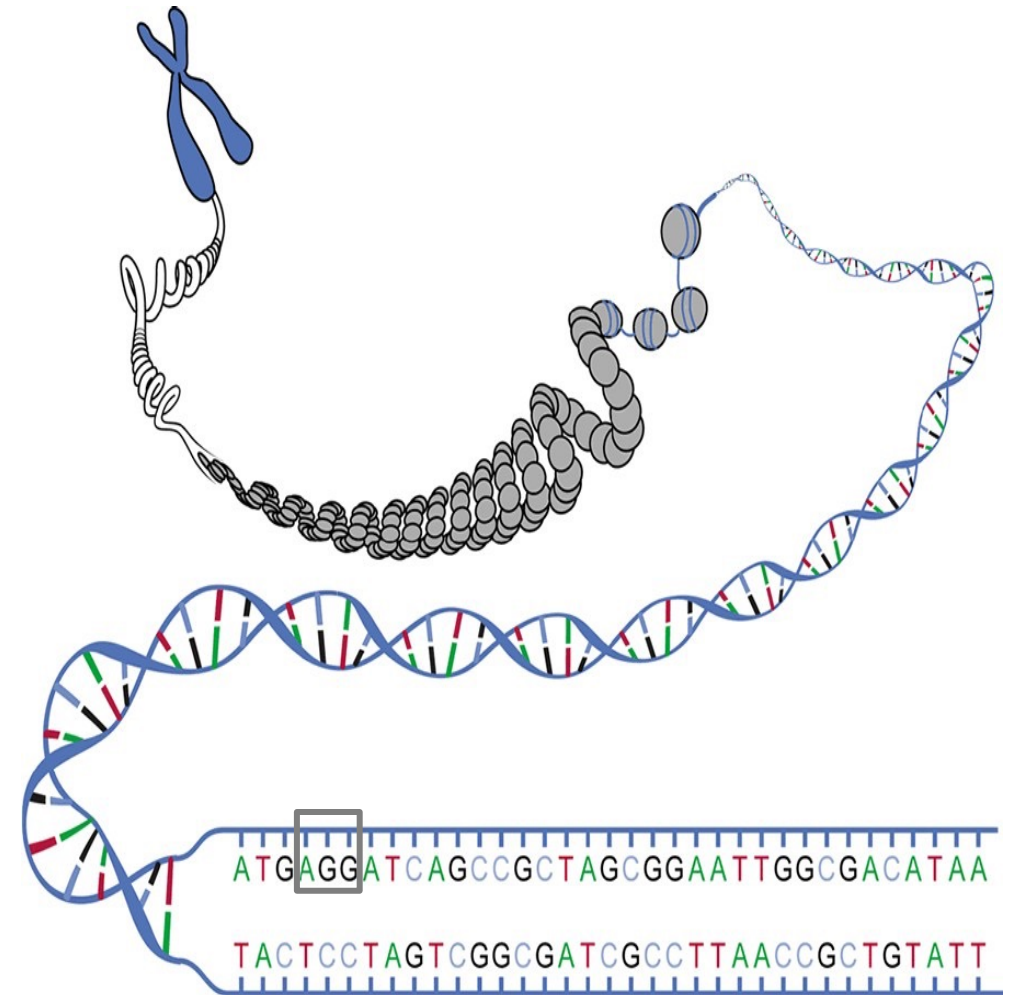
# Landmarks in Agrobiotech Research



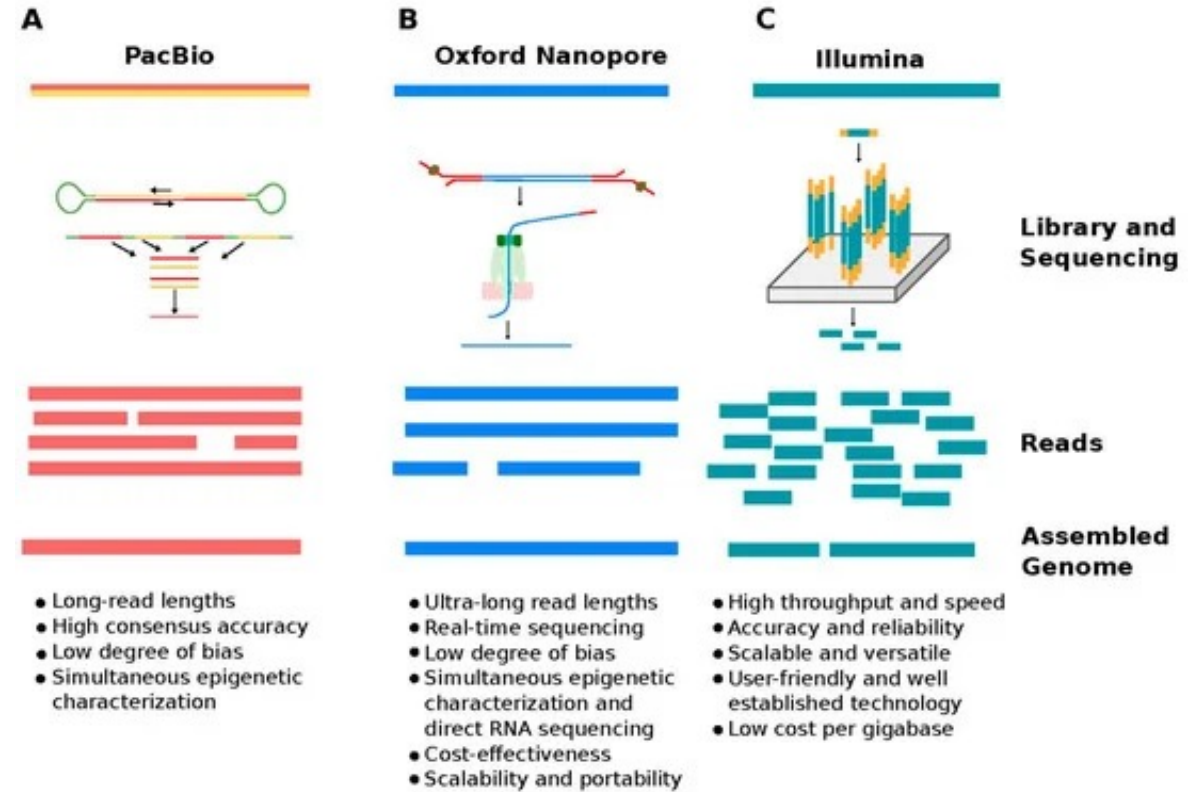
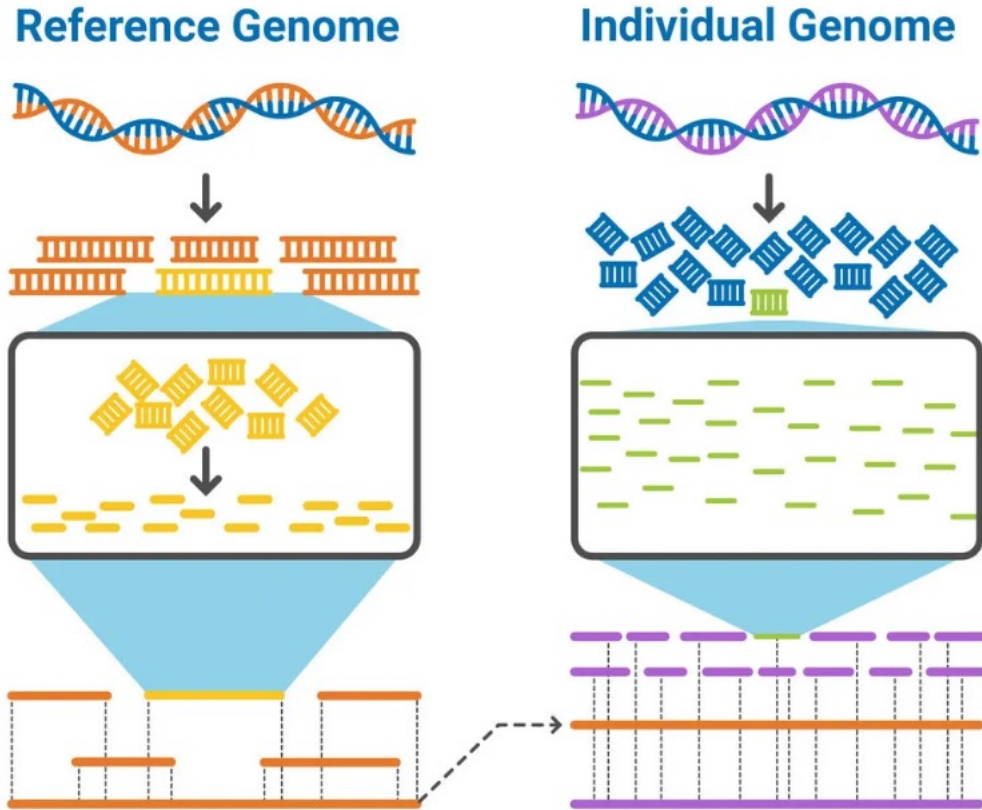
# Big Data and Genetics and Genomics

Technological advancements have enabled scientists to generate, analyse and data at a pace exponentially faster than 20 years ago

DNA technologies (and physics) led the data revolution

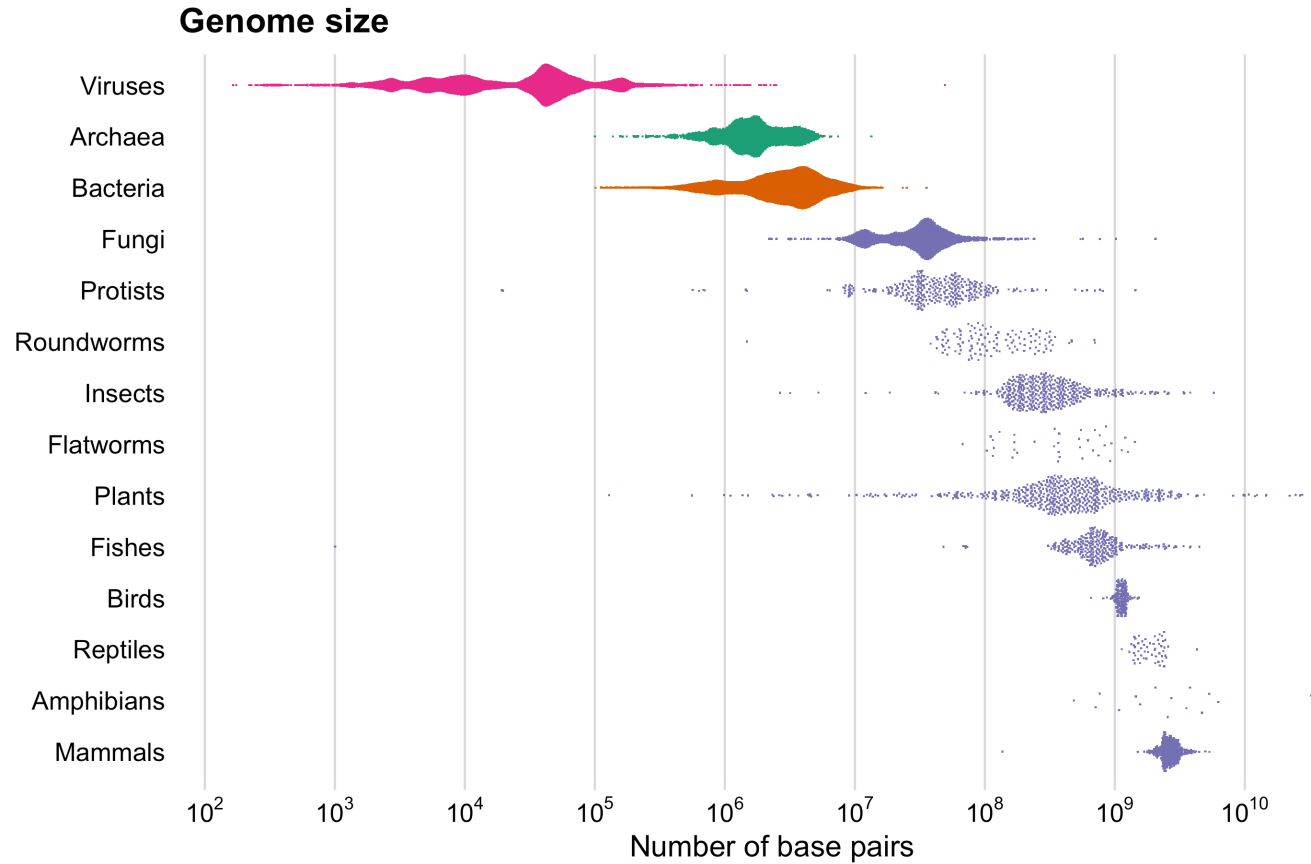


# How genome sequencing works

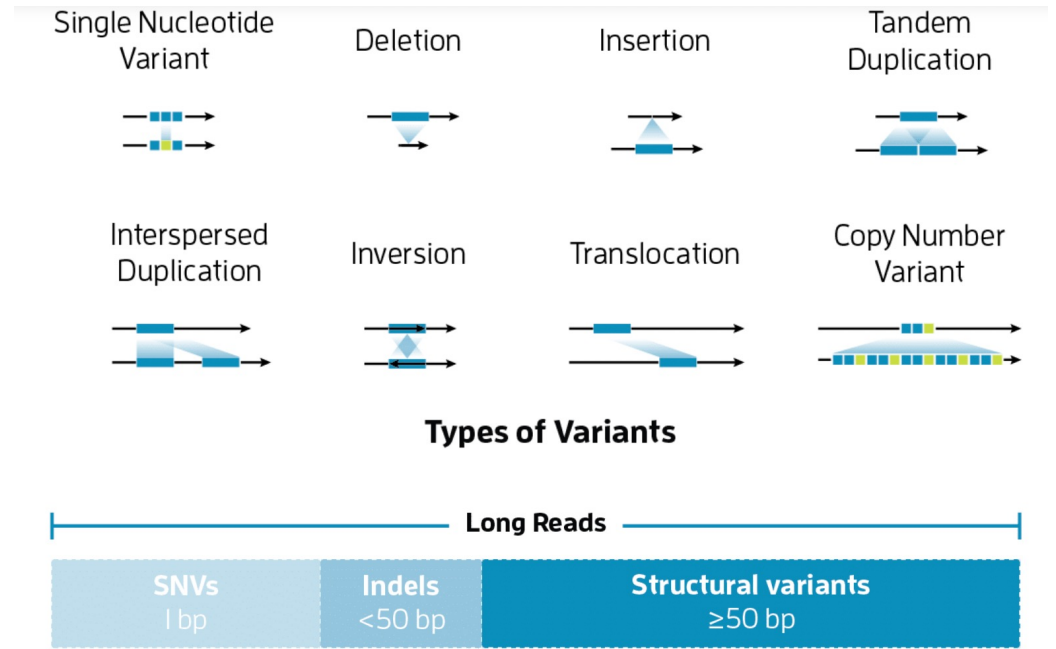


Ermini e Driguez, 2024

# It's all about diversity



Data source: National Center for Biotechnology Information

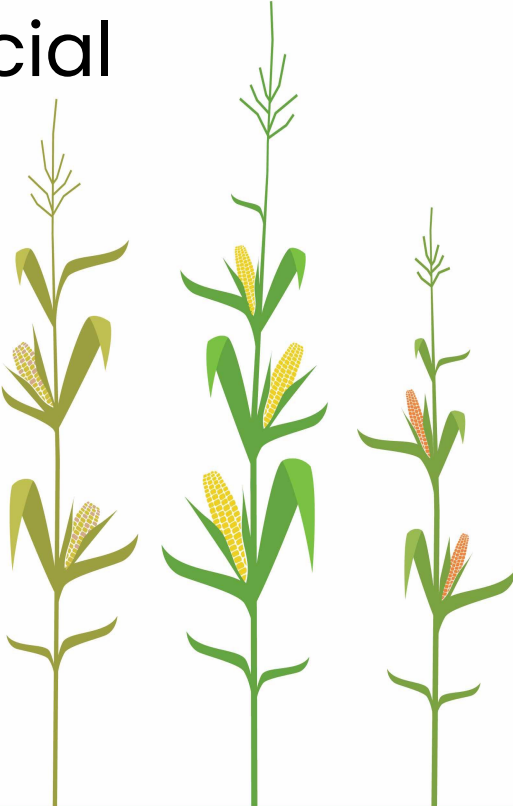


# It's all about diversity

Genetic variation is crucial

Within and across:  
Species, subspecies,  
varieties, breeds....

Induced or created

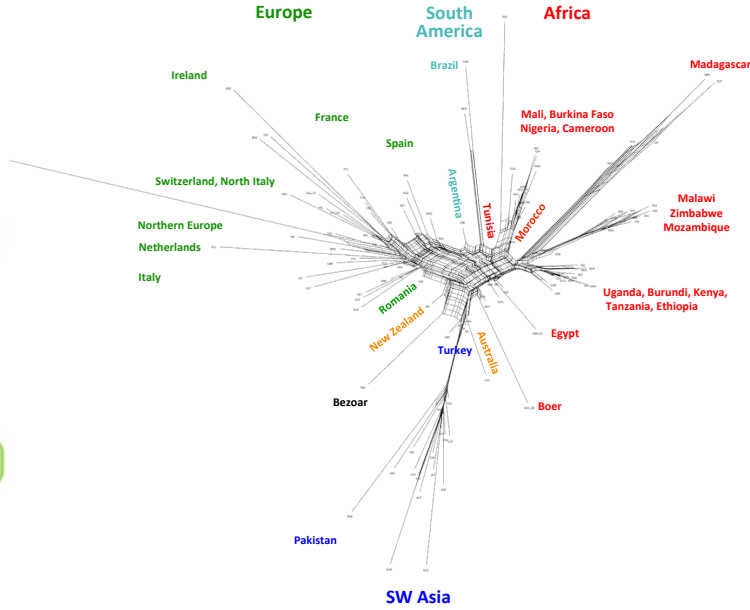
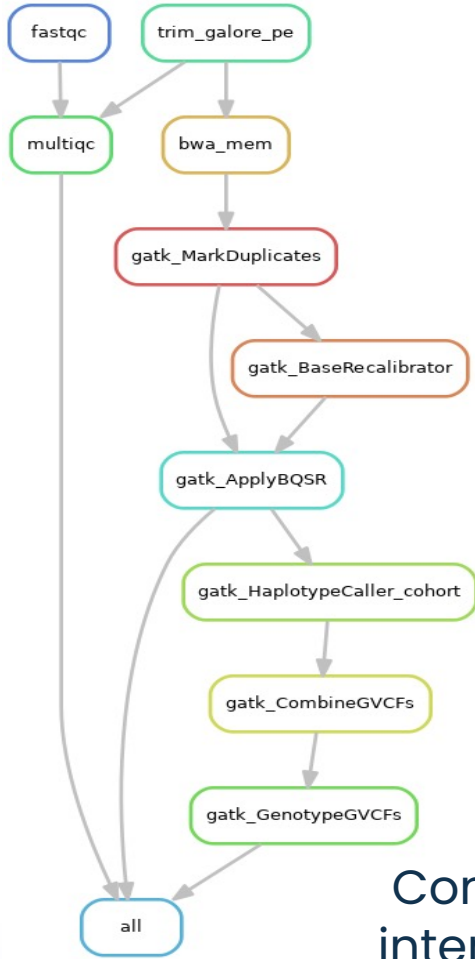


Within a species, individual plants carry different genetic information, which shapes the distinct characteristics of each individual

**Why genetic diversity is crucial:**

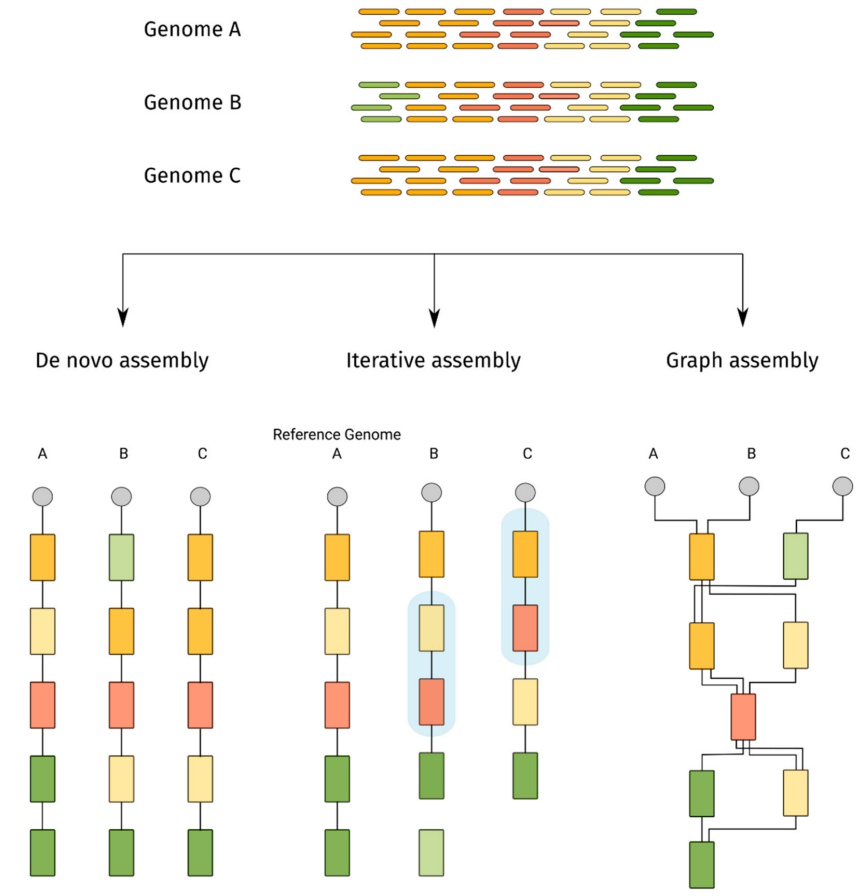
- ADAPTATION**  
Genetic variation allows plants to adapt to a range of environments
- RESILIENCE**  
Diversity helps plant species to withstand pests, diseases, and varying climate conditions
- COMPLEX TRAITS**  
In breeding, genetic diversity acts as the building blocks for enhancing complex traits

# Exploring diversity: from Hapmap to Pangenomes

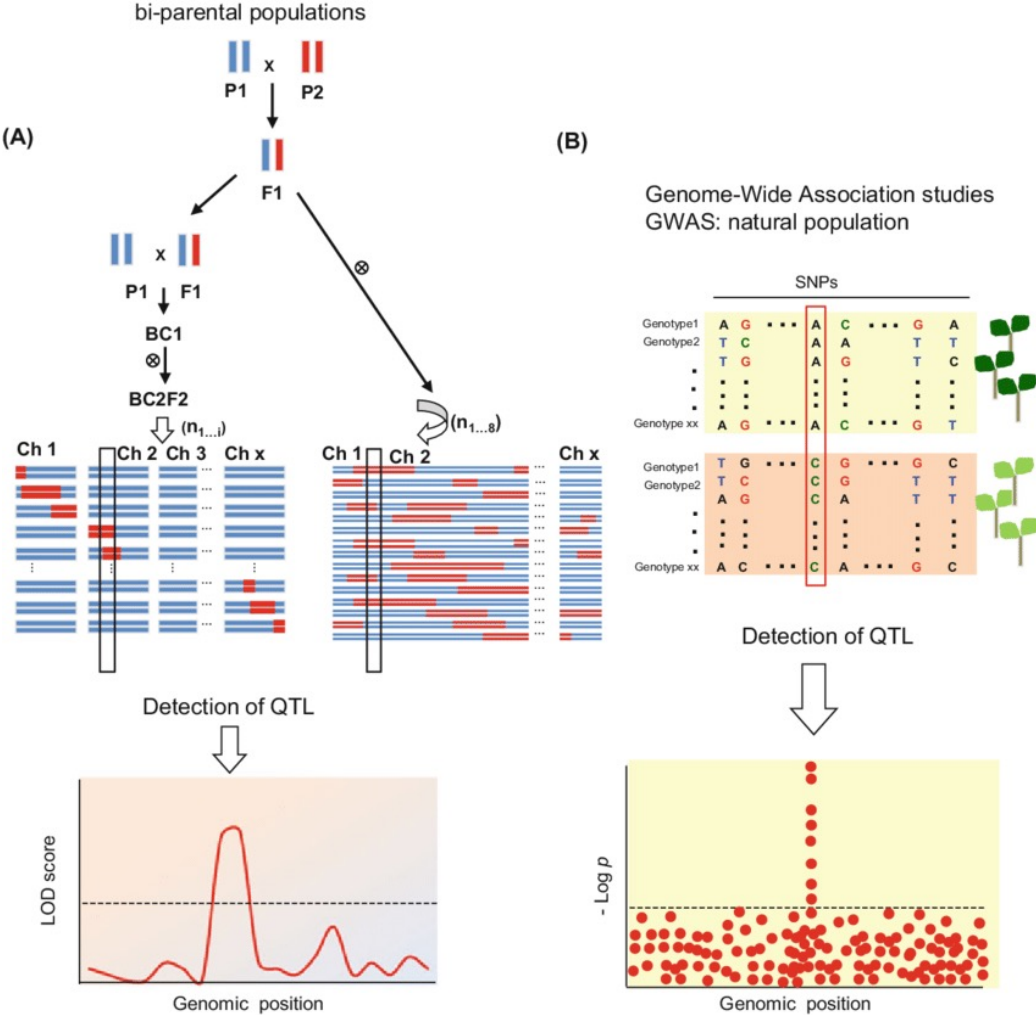
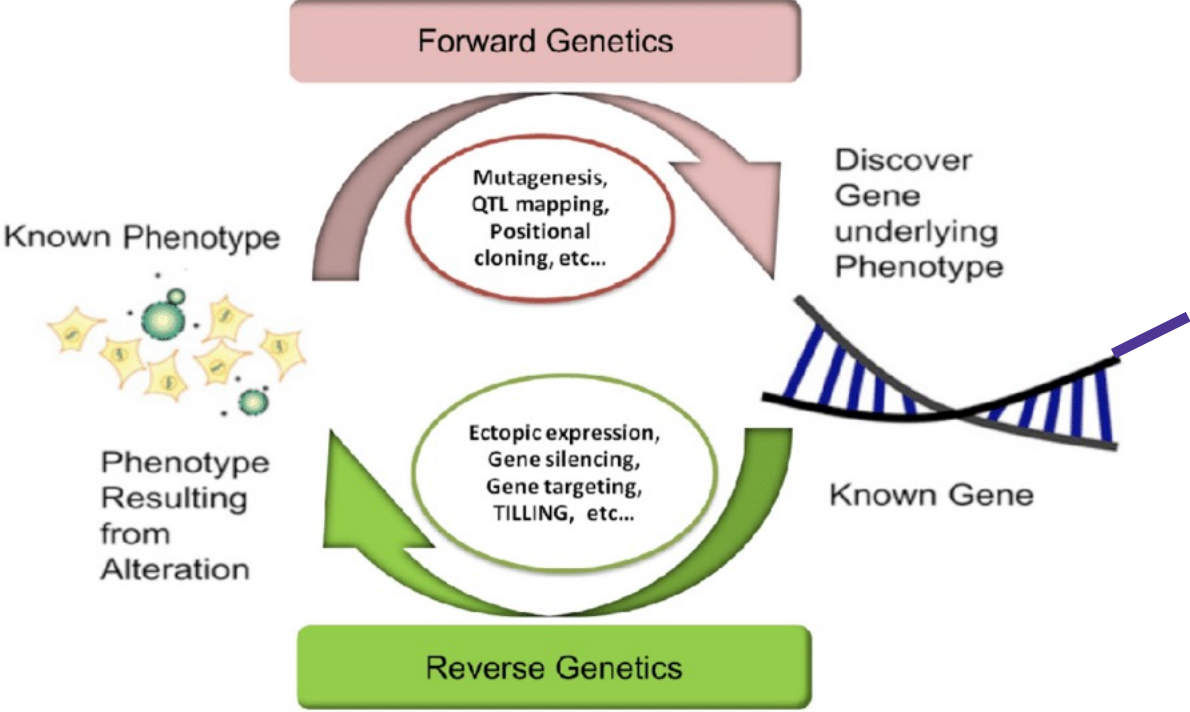


Constructing a pangenome is a resource-intensive process, requiring petabyte-scale storage, thousands of CPU cores, and terabytes of RAM

## Pangenome assembly methods



# Finding causative variants or QTLs

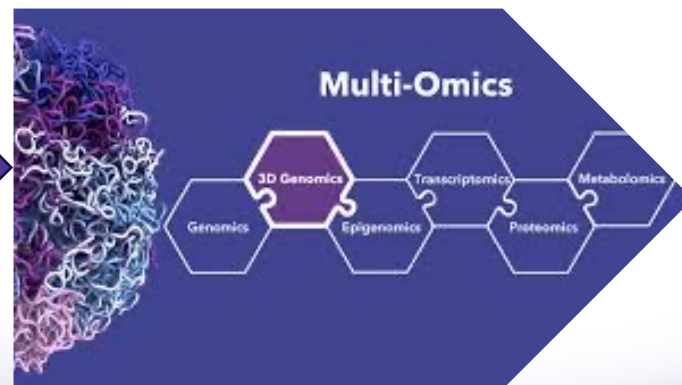
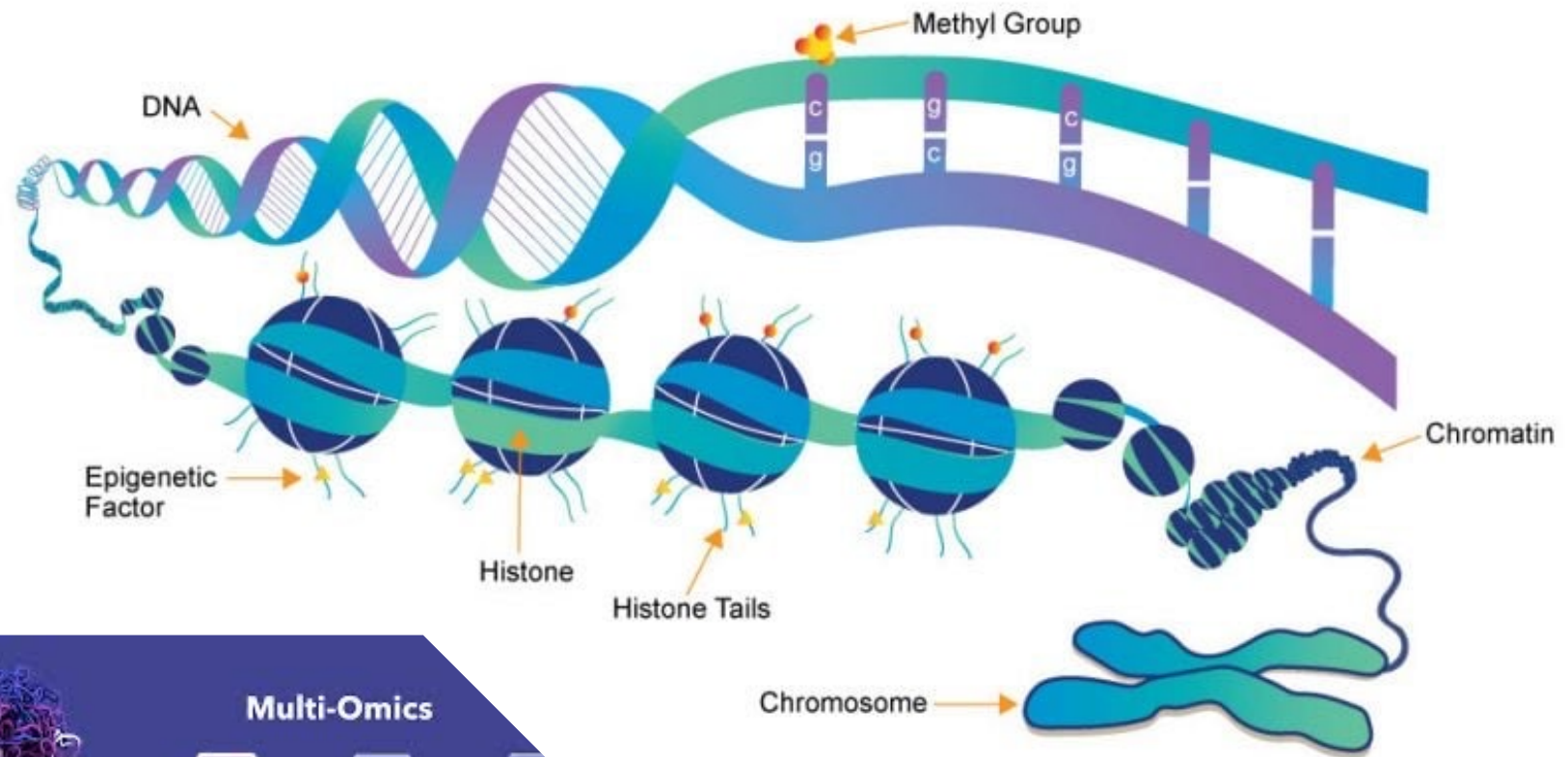


Alseekh et al 2018

# Exploring molecular mechanisms: Epigenetics

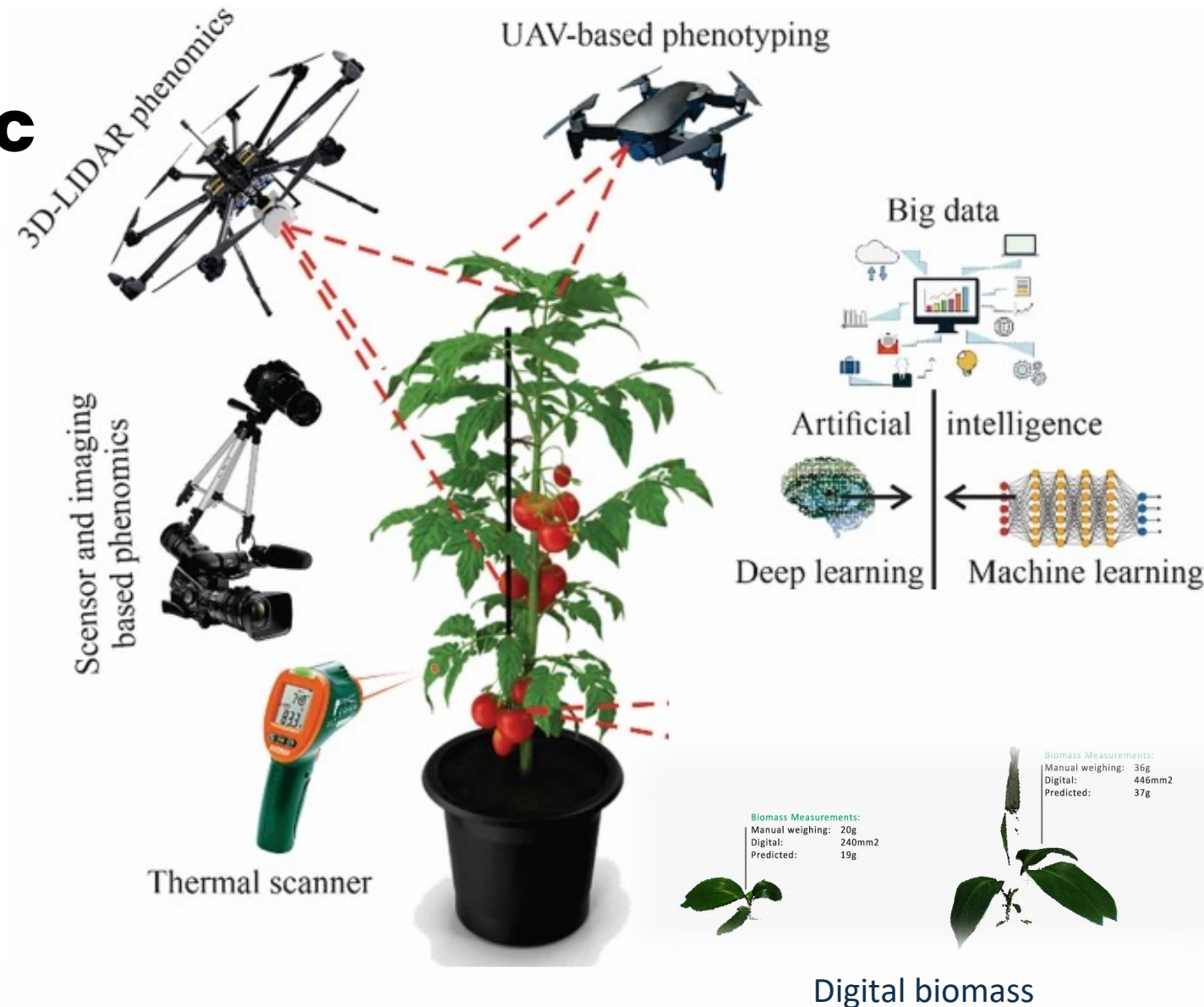
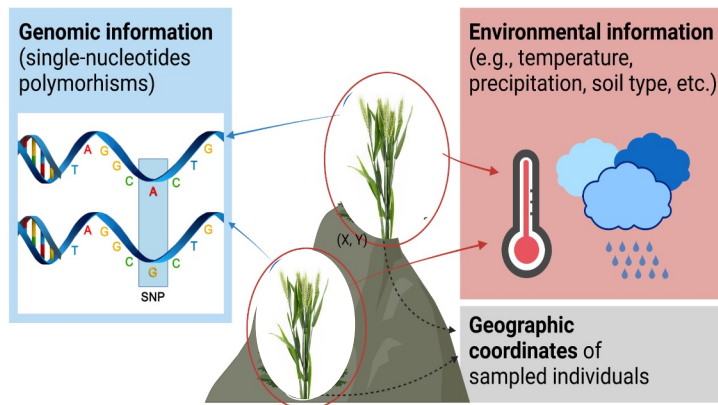
Environmental effects

Transgenerational transmission

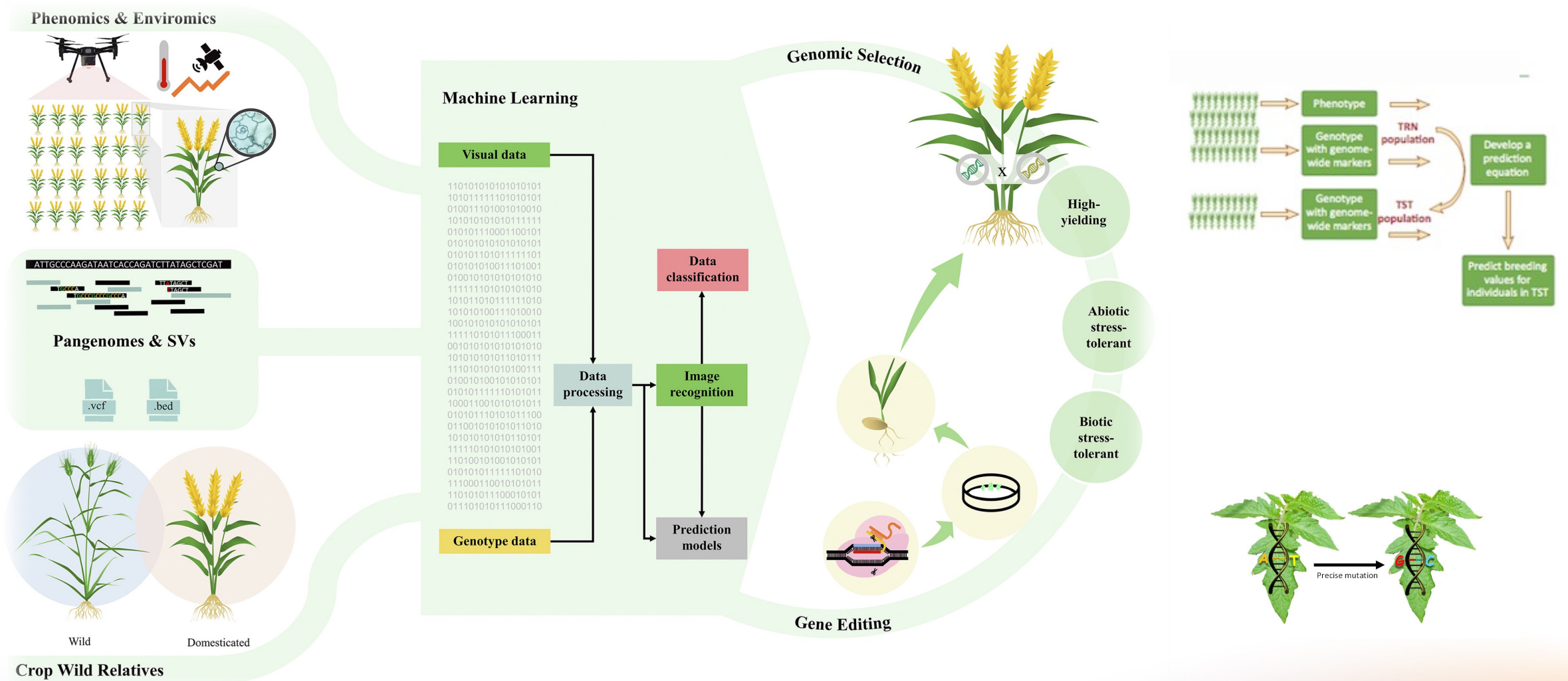


# Exploring phenotypic variation

## Phenomics



# Designing plant and livestock for the future



Current Opinion in Plant Biology

Saad et al. 2022

# Dealing with data in agro-biotech

## STORAGE

International Archives

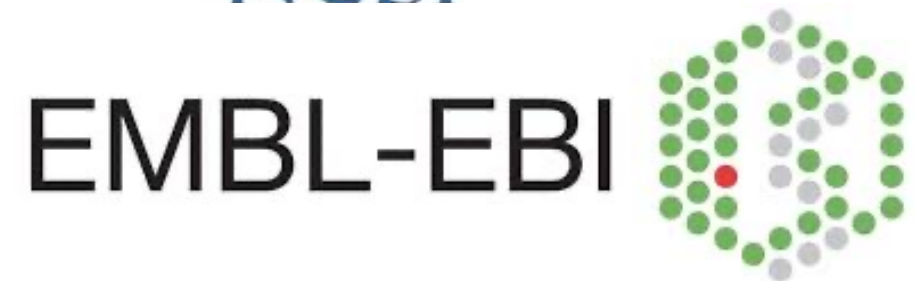
Novel methods for data representation

## MEMORY and COMPUTING

ML Methods

Trans-disciplinary approaches

STILL A MAJOR LIMITATION



CONFERENZA GARR 2024  
**Navigare la complessità**



**agritech**

National Center for  
Technology in Agriculture

<https://agritechcenter.it>



<https://ibba.cnr.it>

Alessandra.Stella@cnr.it