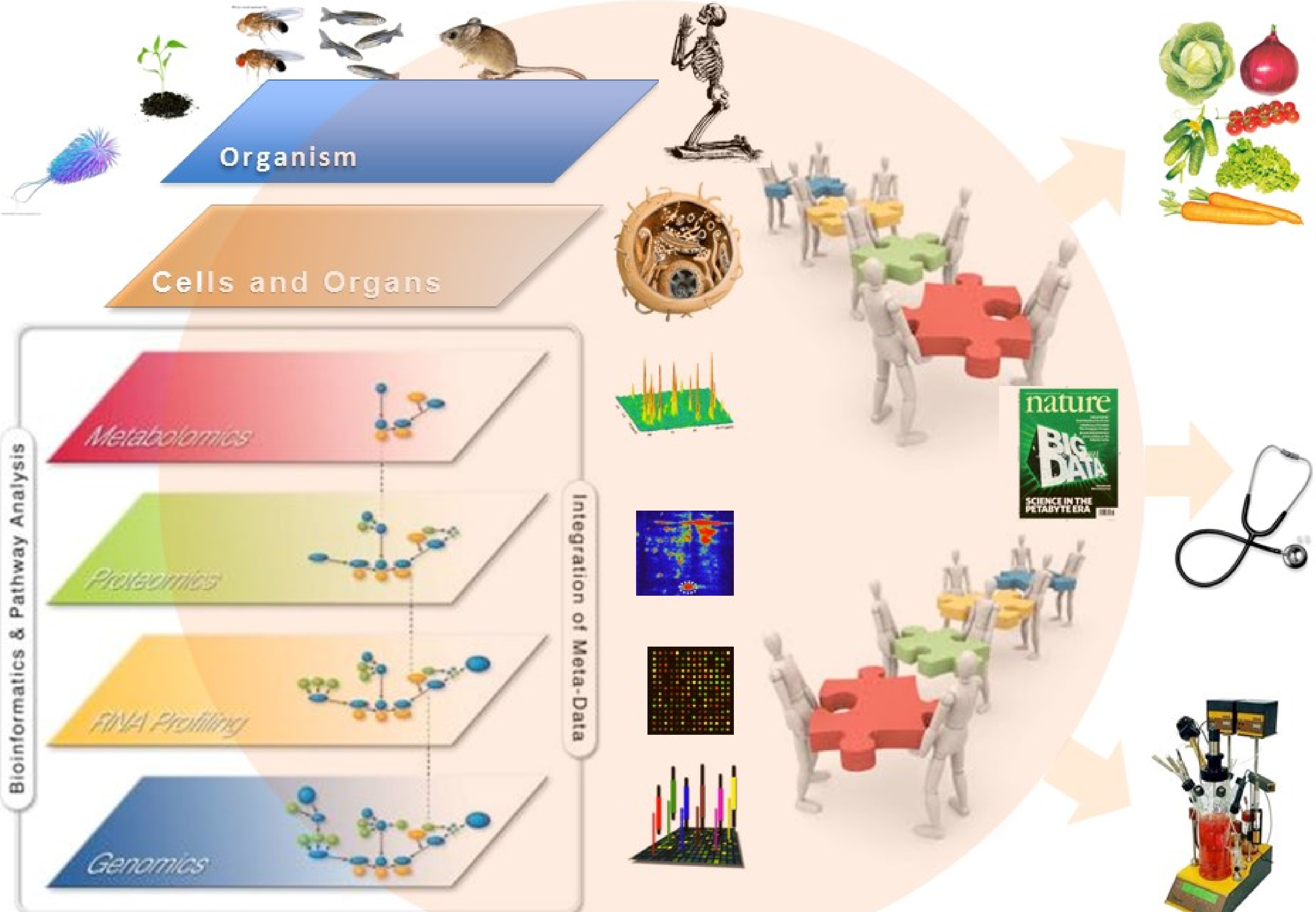
NGS applications and methods

Notis Argiriou, PhD
Senior Scientist
Institute of Applied Biosciences | CERTH

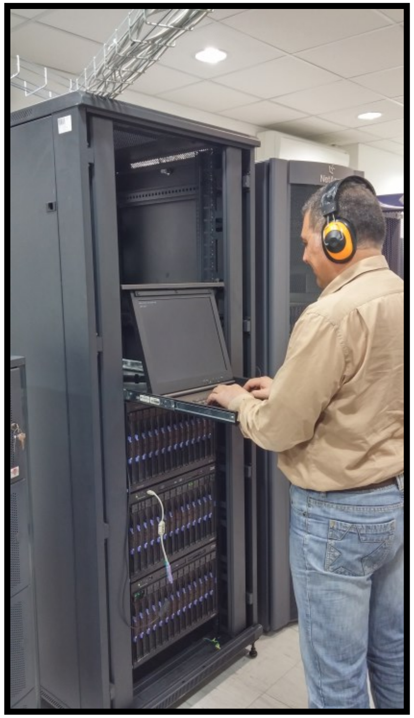
Life Science data: Multi-omics, multi-technology, multi organism, multi dimensional

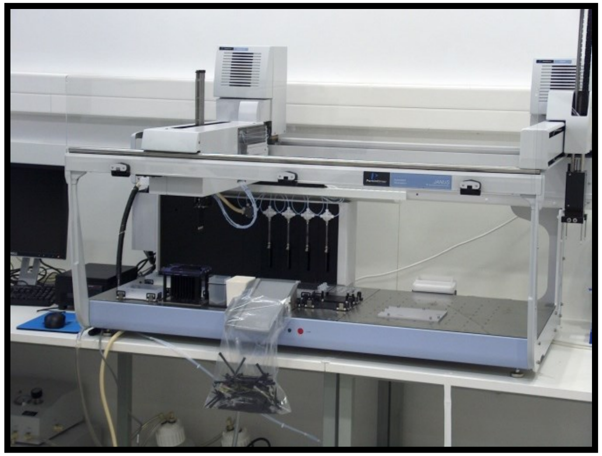


http://www.elixir-europe.org

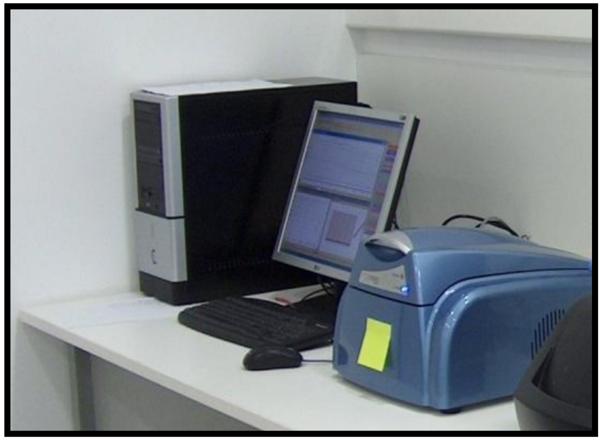
Genomic Infrastructures

Infrastructure– Expertise





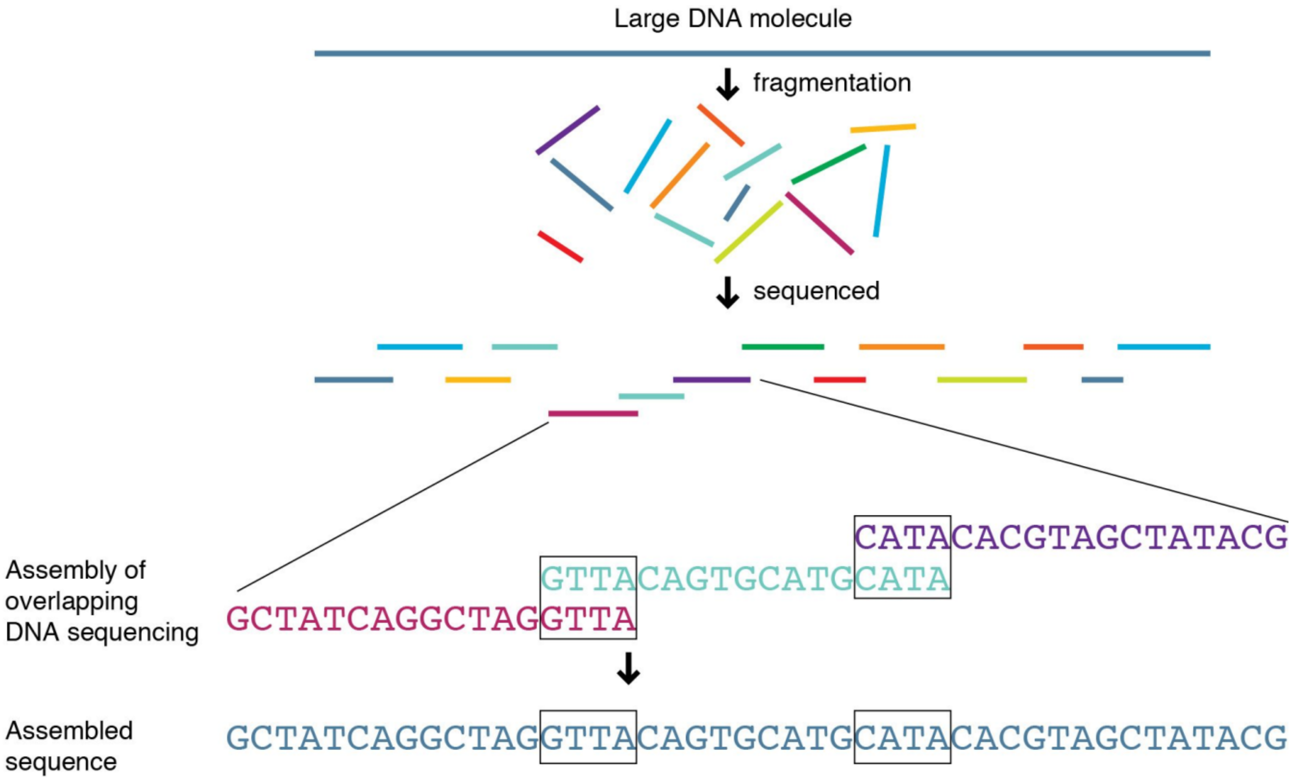
Collaborate.
Reduce costs.
Get results.



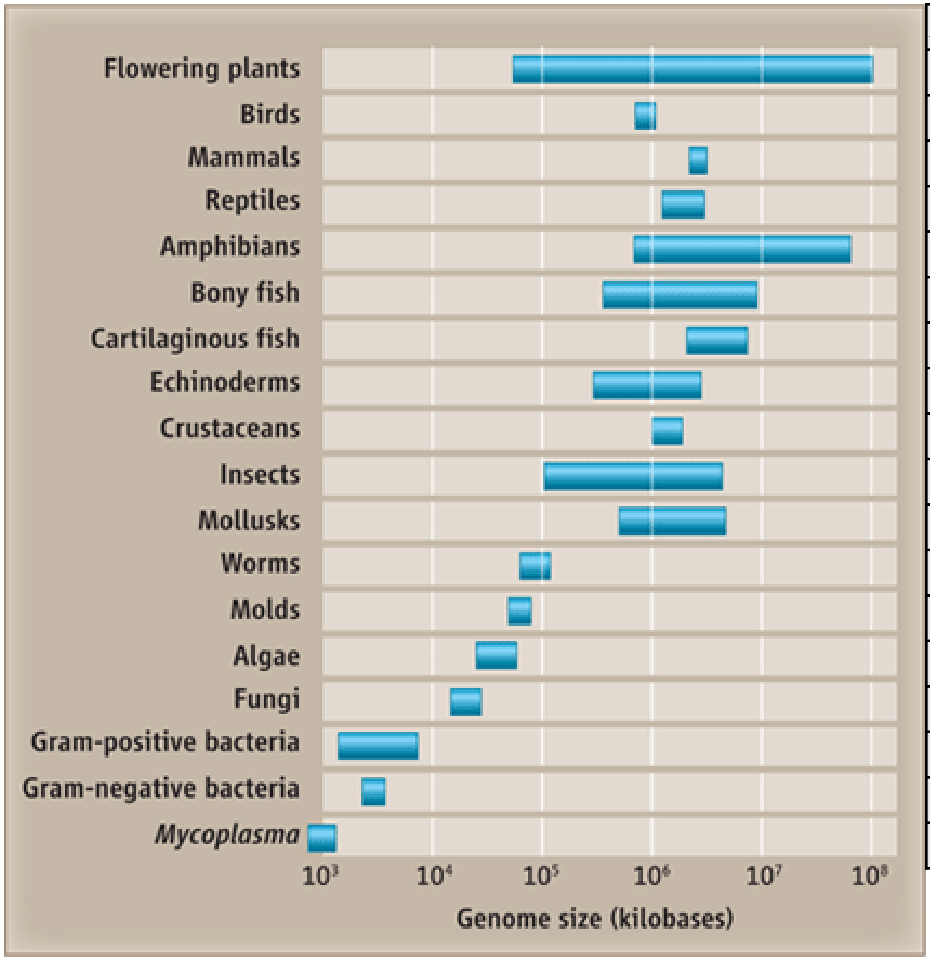
DNA Sequencing

- Whole Genome Seq
- Targeted DNA-Seq
- Exome Sequencing
- ChIP-Seq
- Methylome Seq
- ATAC-Seq
- Metagenome Seq

Whole Genome Seq

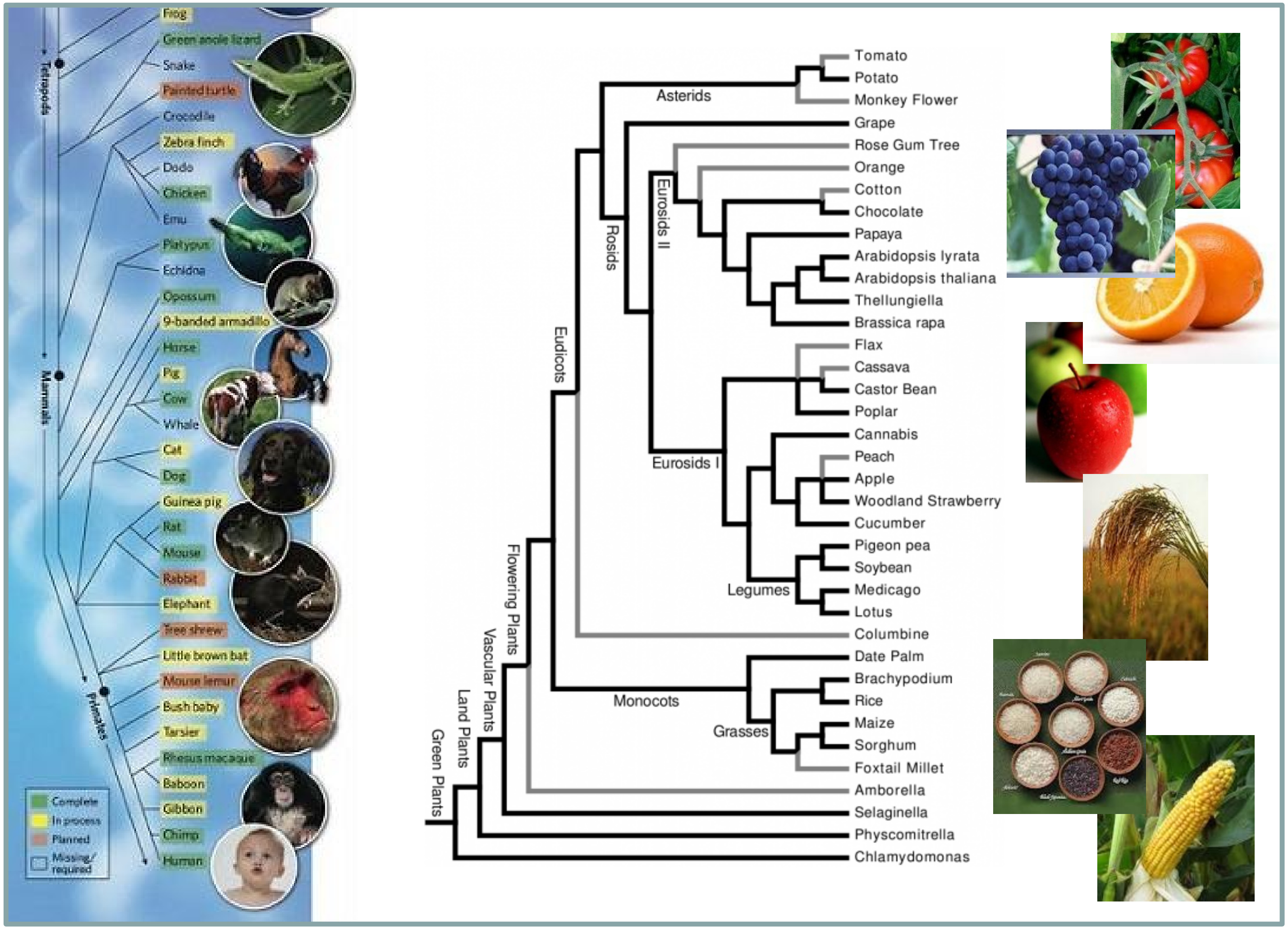


Genome size

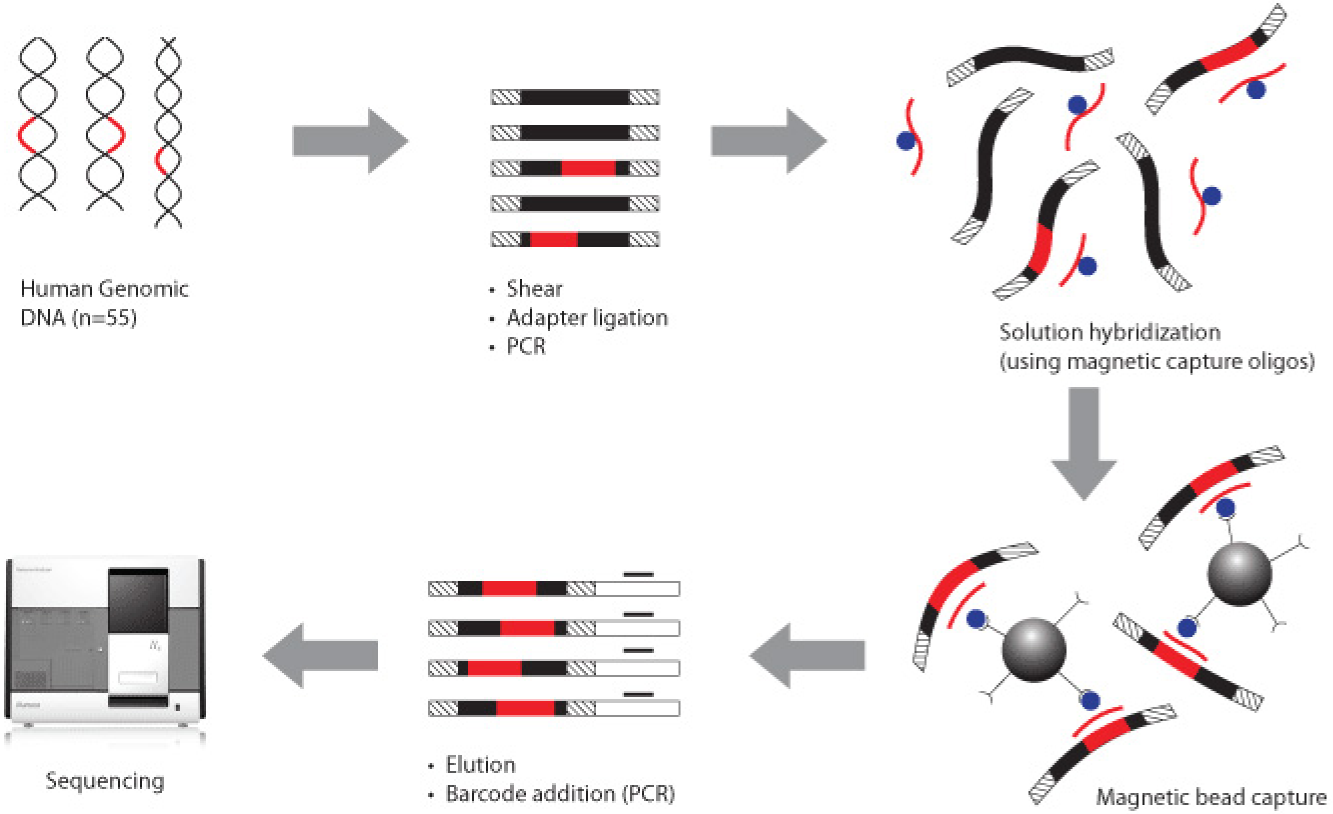


Size	**40X (2x150bp)**	**100X (2x150bp)**
100Mb- 10Gb	14 mln - 1.4e10	35 min - 3.5e10

Sequenced Genomes



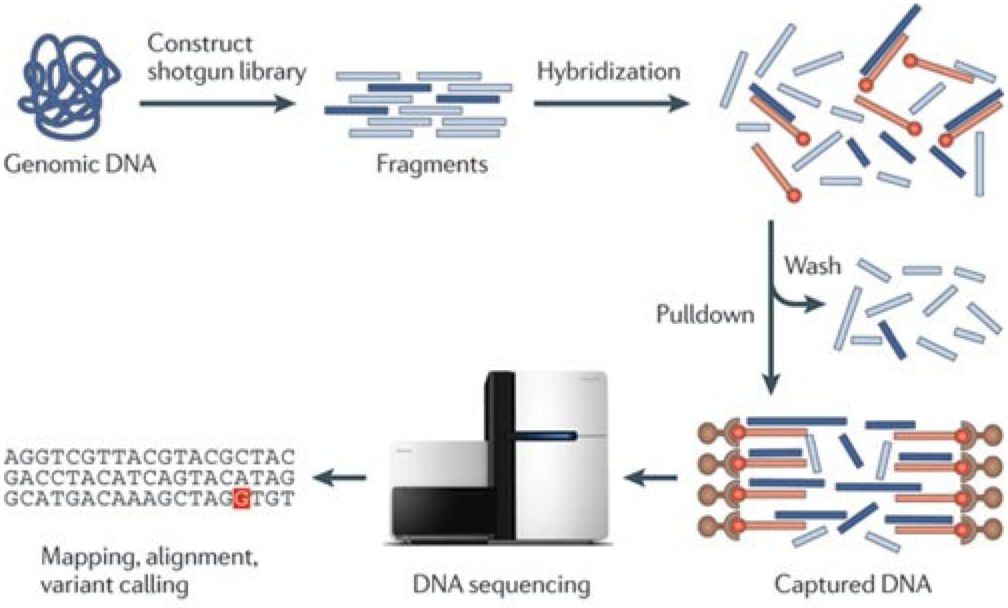
Targeted Genome Seq



- Cardio
- Onco
- DNA Virus

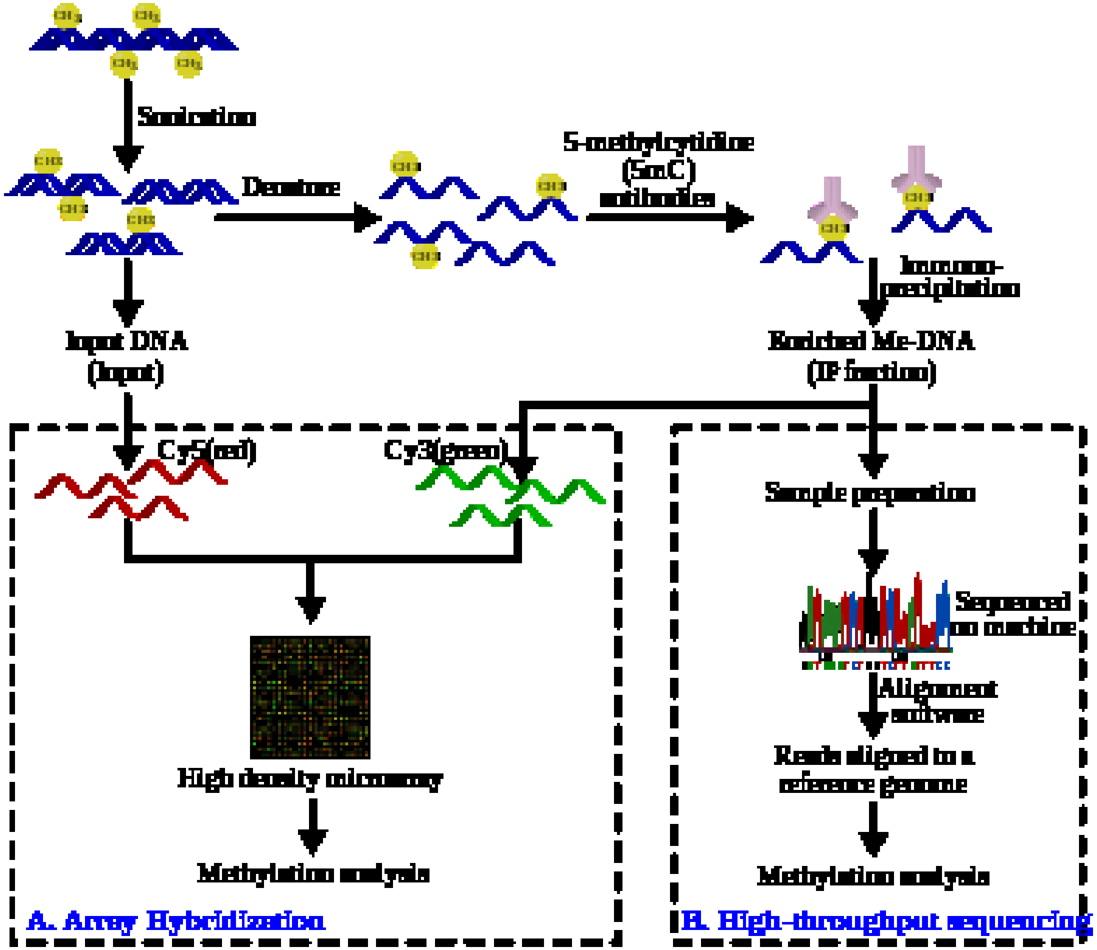
etc

Exome Sequencing

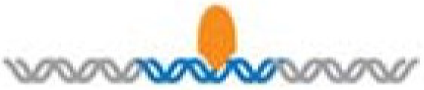


Nature Reviews | Genetics

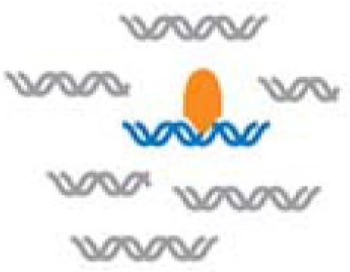
Methylome Seq



ChIP-Seq



1. Cross-link bound proteins to DNA.



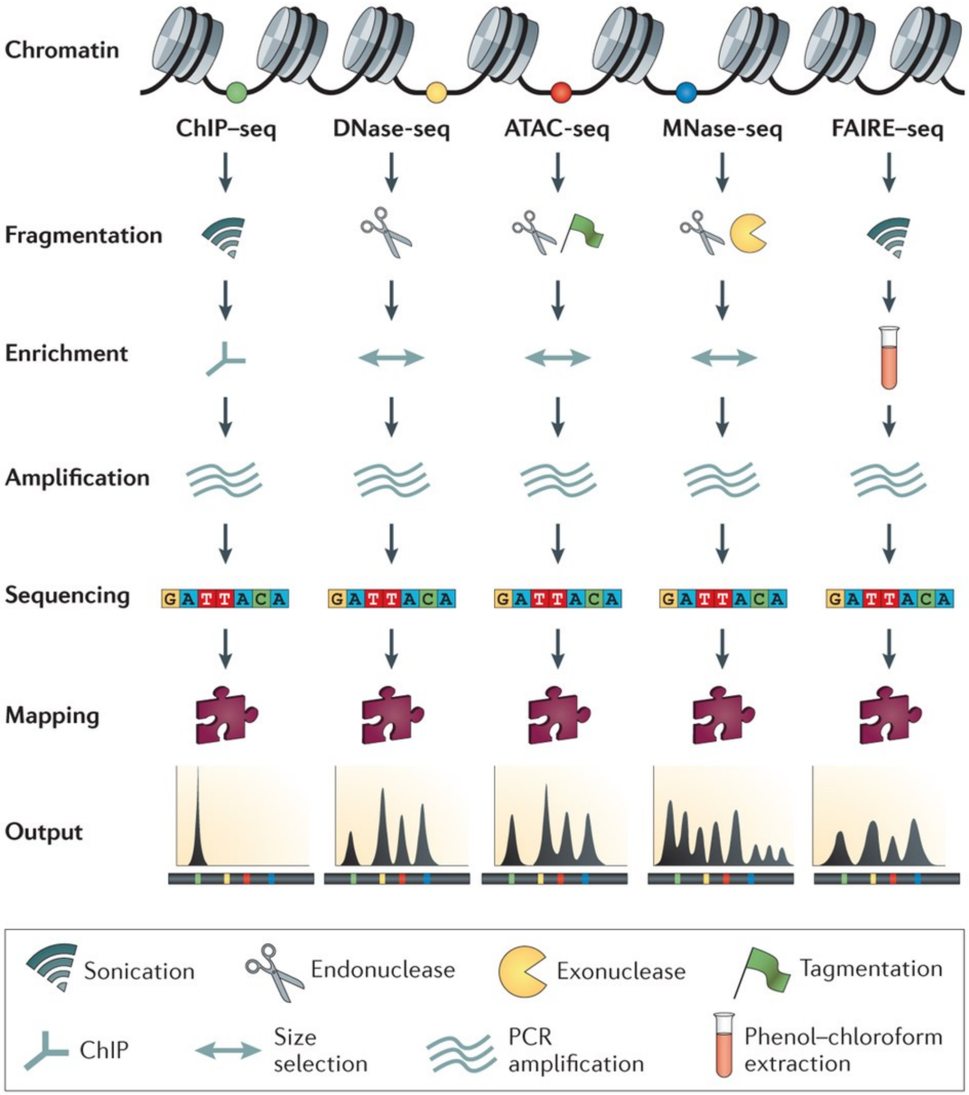
2. Isolate chromatin and shear DNA.

3. Precipitate chromatin with protein-specific antibody.

4. Reverse cross-link and digest protein.

5. Ligate P1 and P2 adaptors to construct fragment library.

Targeted Sequencing

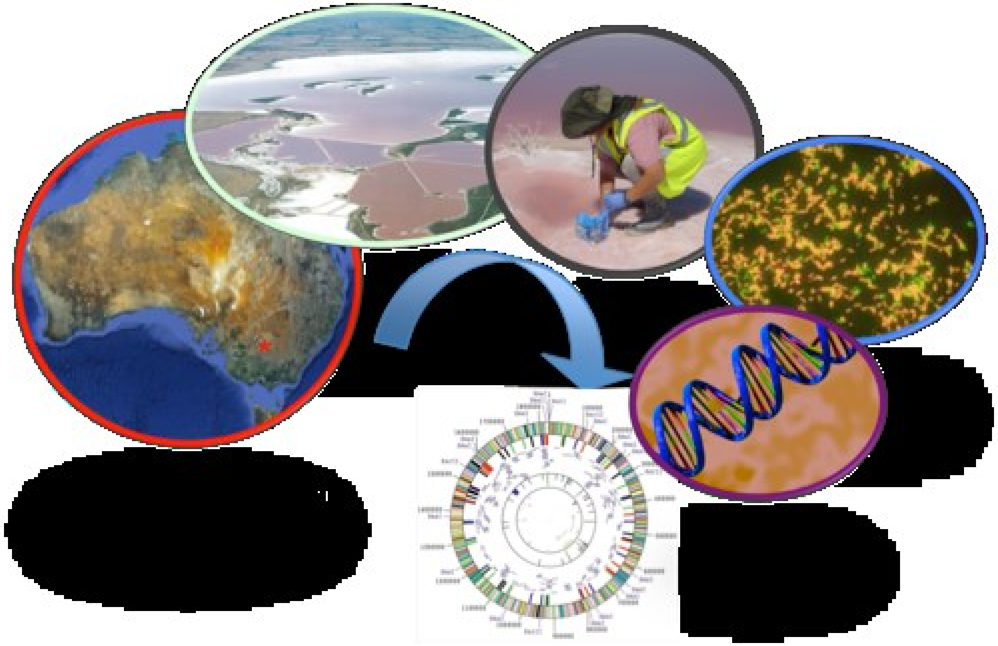


Nature Reviews | Genetics

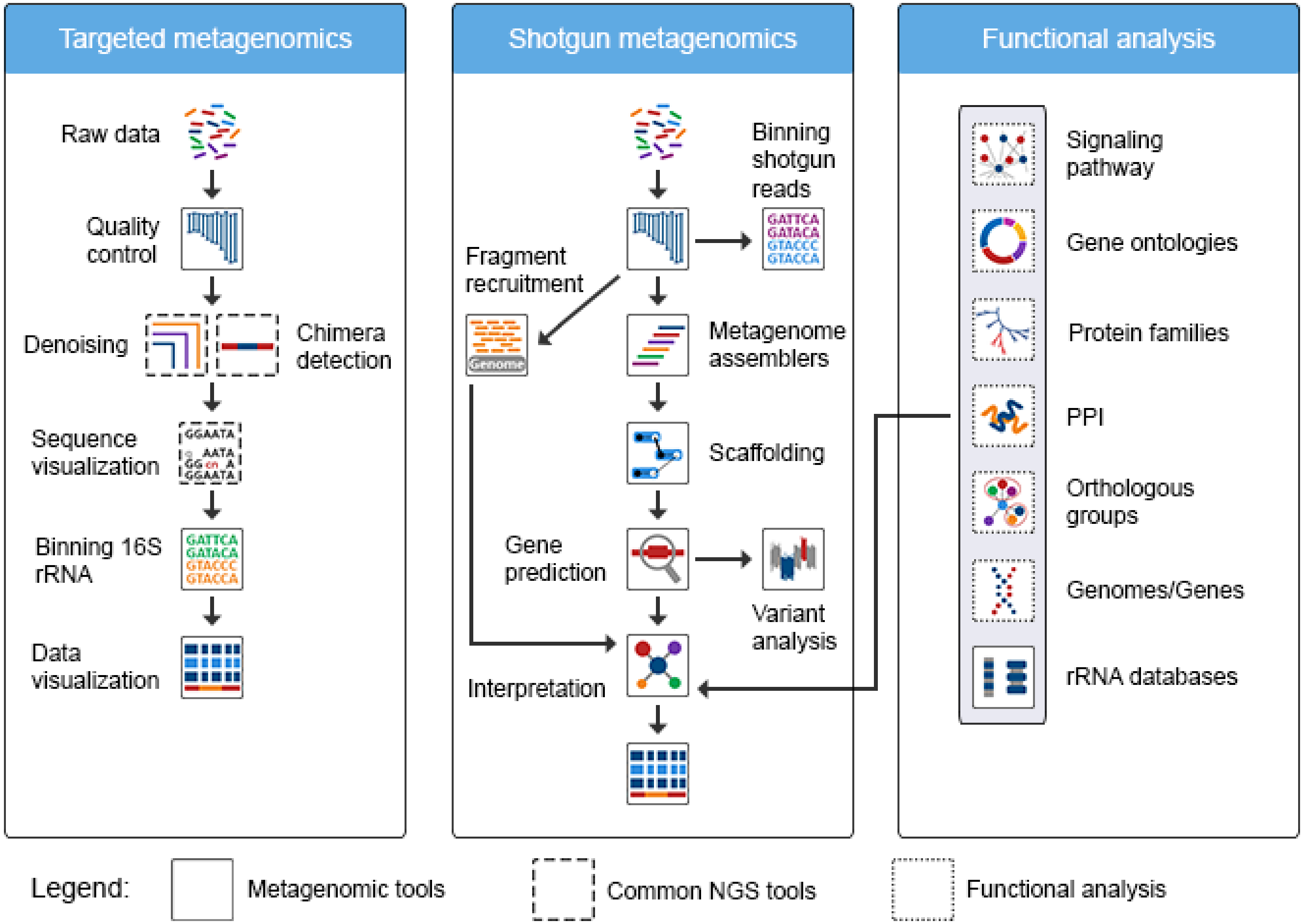
Some Methods

- **FAIRE-Seq** (Formaldehyde-Assisted Isolation of Regulatory Elements): determine the sequences in the genome associated with regulatory activity. Formaldehyde cross-linking is more efficient in nucleosome-bound DNA than it is in nucleosome-depleted regions of the genome. This method then segregates the non cross-linked DNA that is usually found in open chromatin, which is then sequenced. The protocol consists of cross linking, phenol extraction and sequencing the DNA in aqueous phase.
- **ATAC-seq** (Assay for Transposase-Accessible Chromatin). Used to study chromatin (complex DNA structure), namely its accessible sites. ATAC-seq employs a mutated hyperactive transposase. Transposons are believed to incorporate preferentially into genomic regions free of nucleosomes (nucleosome-free regions).

Metagenomics



Metagenomics



Targeted metagenomics

Raw data

Quality control

Denoising

Chimera detection

Sequence visualization

Binning 16S rRNA

Data visualization

Shotgun metagenomics

Binning shotgun reads

Fragment recruitment

Metagenome assemblers

Scaffolding

Gene prediction

Variant analysis

Interpretation

Functional analysis

Signaling pathway

Gene ontologies

Protein families

PPI

Orthologous groups

Genomes/Genes

rRNA databases

Legend:

Metagenomic tools

Common NGS tools

Functional analysis

RNA Sequencing

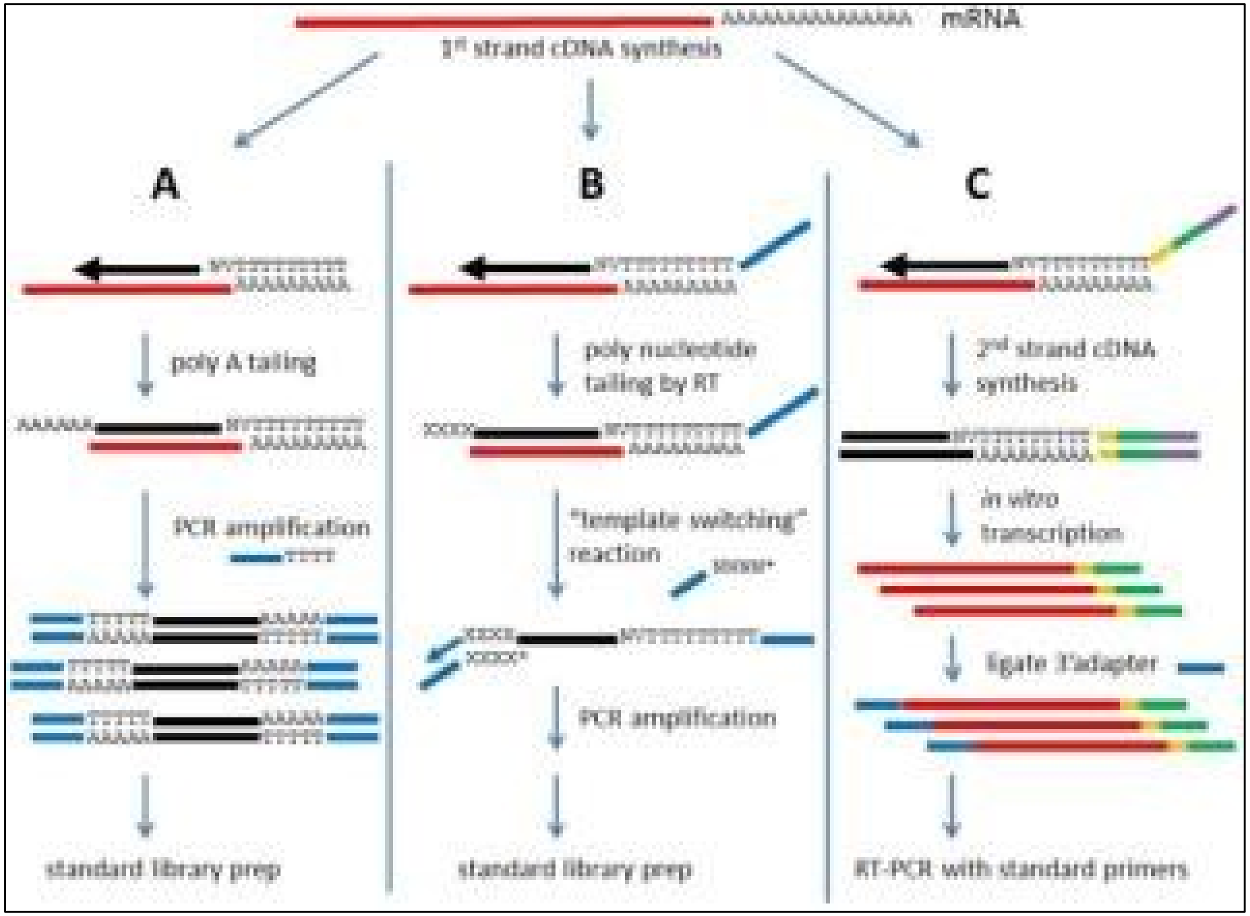
RNA-Seq

Targeted RNA-Seq

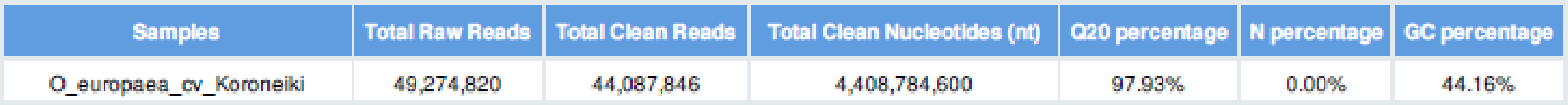
- Ribosome profiling (Ribo-Seq)
- RIP-seq/CLIP-seq

miRNA-Seq

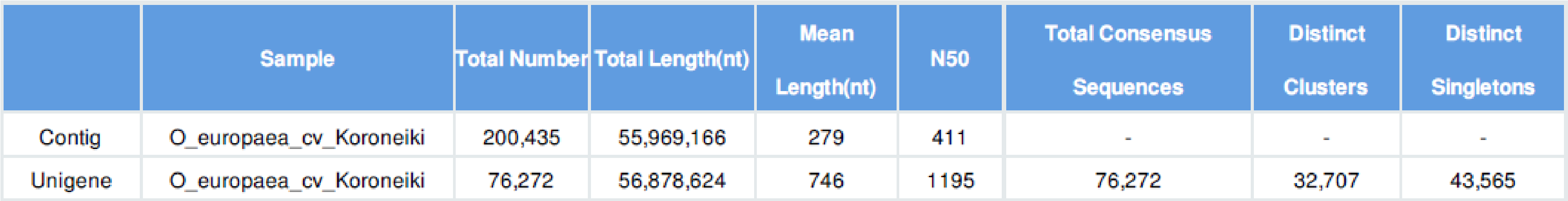
RNA-Seq


AAAAAAAAAAAAAAA
mRNA
1st strand cDNA synthesis
A
B
C
AAAAAAAAA
poly A tailing
AAAAAA
AAAAAAAAA
PCR amplification
TTTTT
TTTTT
AAAAA
AAAAA
TTTTT
standard library prep
AAAAAAAAA
poly nucleotide tailing by RT
XXXX
AAAAAAAAA
"template switching" reaction
XXXX*
XXXX
XXXX*
PCR amplification
standard library prep
AAAAAAAAA
2nd strand cDNA synthesis
AAAAAAAAA
in vitro transcription
ligate 3'adapter
RT-PCR with standard primers

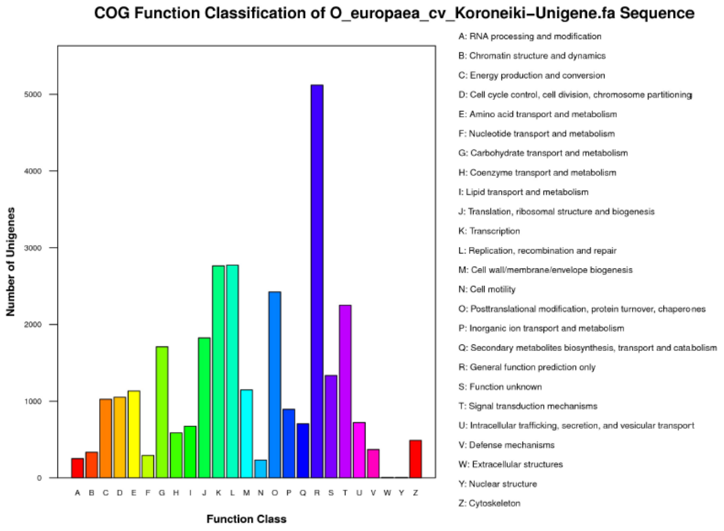
RNA-Seq of *Olea europea* cv Koroneiki



Samples	Total Raw Reads	Total Clean Reads	Total Clean Nucleotides (nt)	Q20 percentage	N percentage	GC percentage
O_europaea_cv_Koroneiki	49,274,820	44,087,846	4,408,784,600	97.93%	0.00%	44.16%



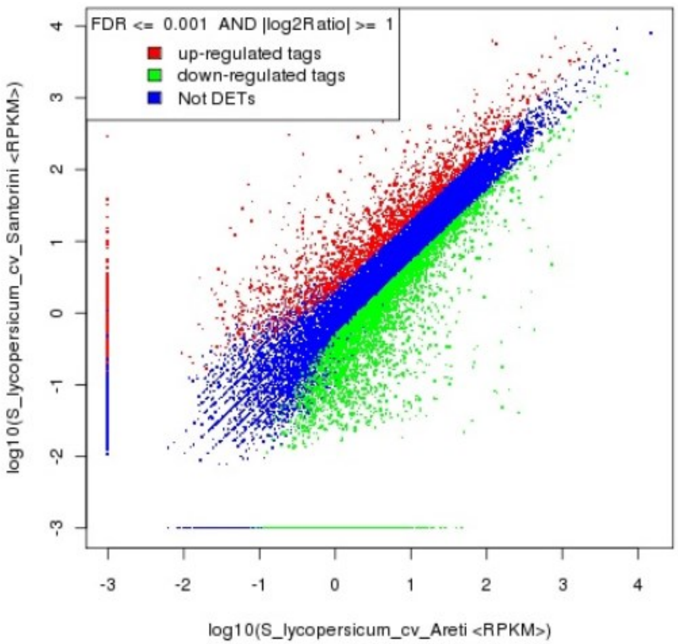
	Sample	Total Number	Total Length(nt)	Mean Length(nt)	N50	Total Consensus Sequences	Distinct Clusters	Distinct Singletons
Contig	O_europaea_cv_Koroneiki	200,435	55,969,166	279	411	-	-	-
Unigene	O_europaea_cv_Koroneiki	76,272	56,878,624	746	1195	76,272	32,707	43,565

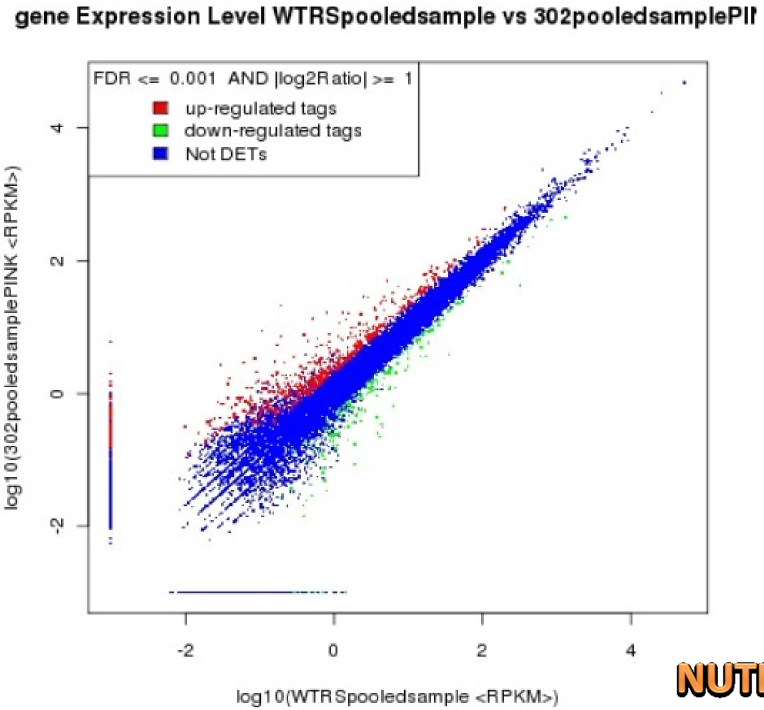


BIOHYDROX-T

RNA-Seq of tomatoes

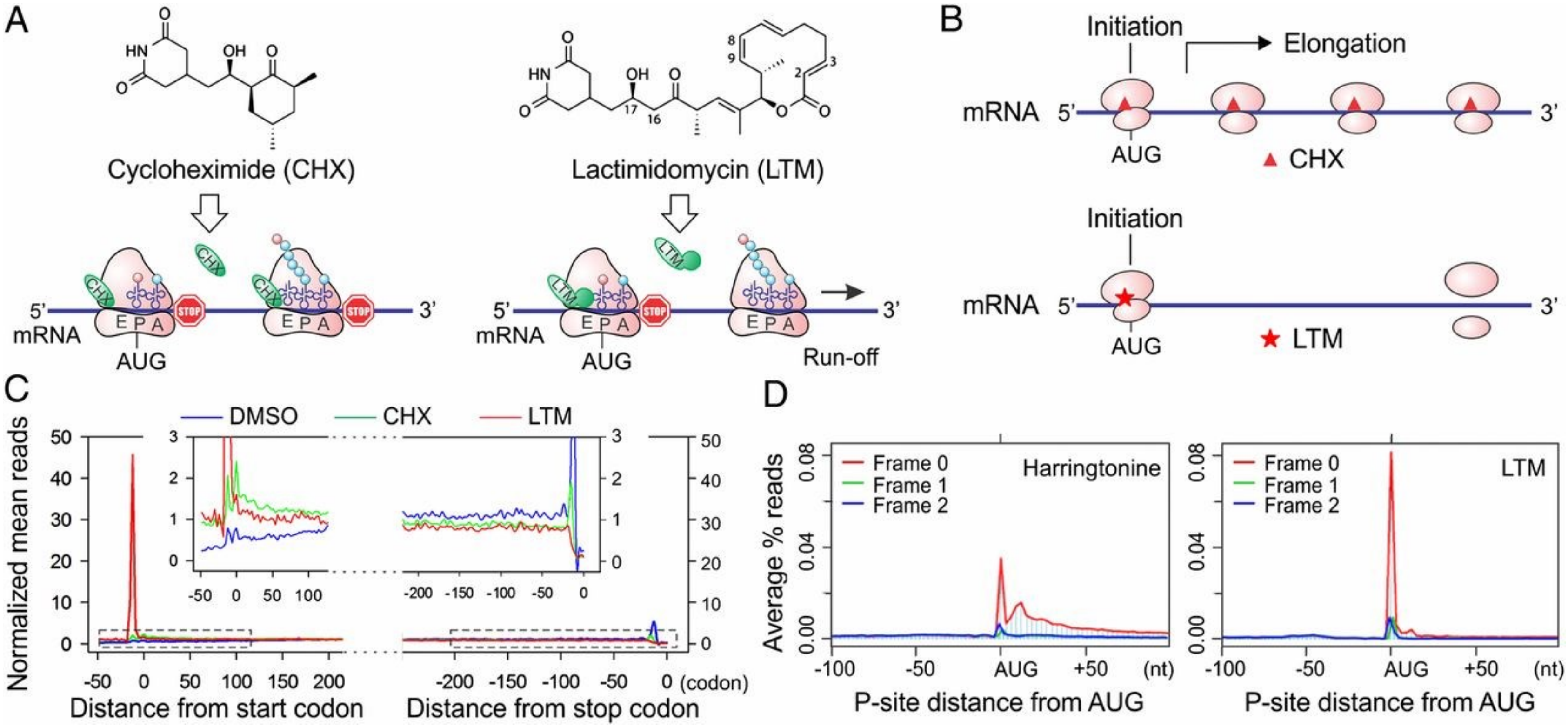
PDO Tomato of Santorini vs cv Areti





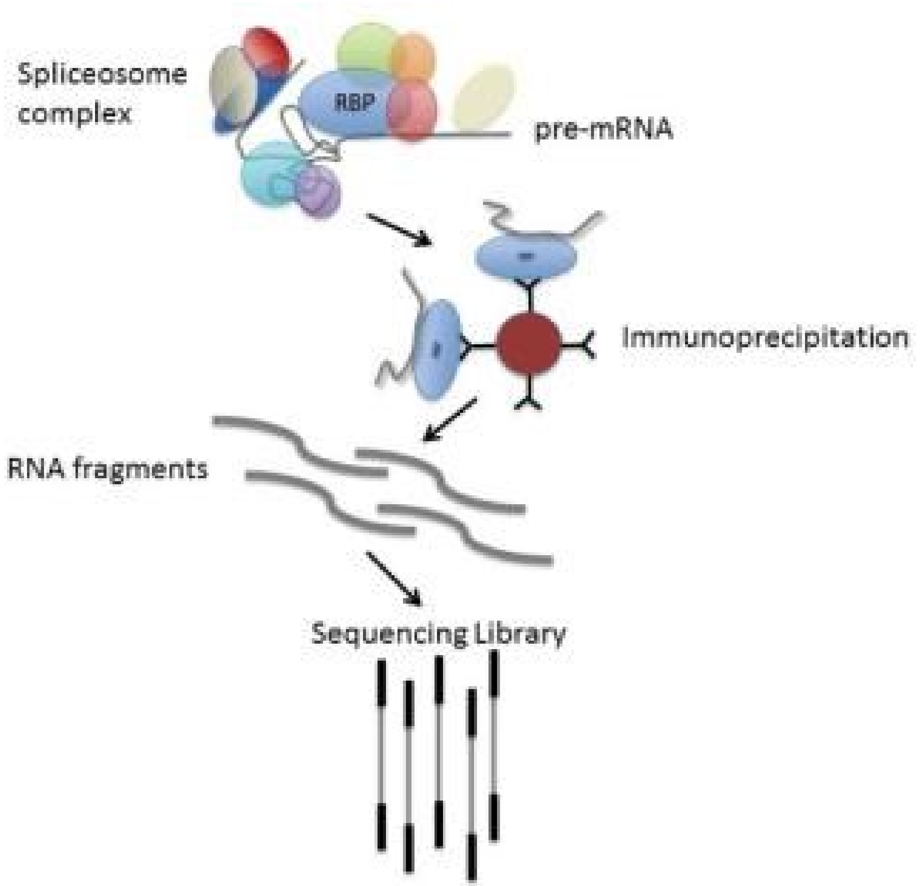
RNA-Seq of *Solanum eleagnifolium*

Ribo-Seq



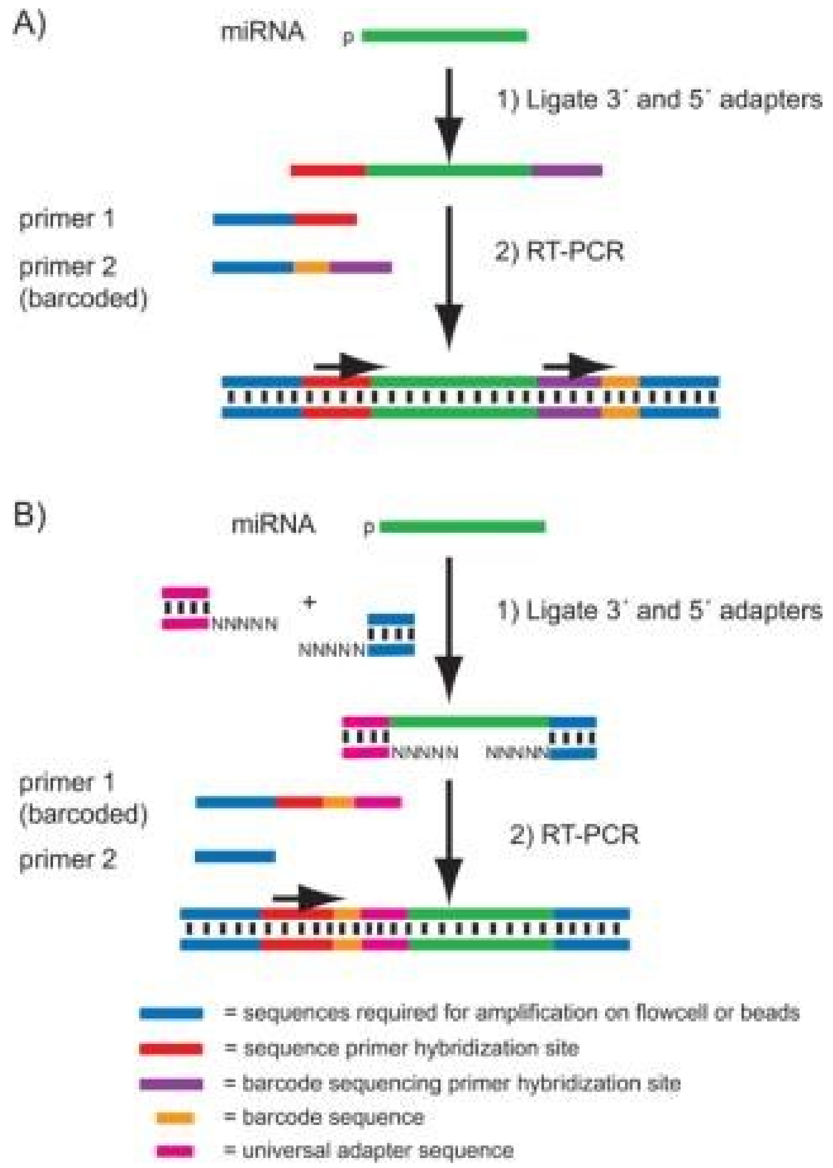
Ribosome profiling is a recently developed high-throughput sequencing technique that captures approximately 30 bp long ribosome-protected mRNA fragments during translation

RIP-seq



RNA immunoprecipitation (RIP-seq) done by targeting RNA binding proteins (RBPs).

miRNA-Seq


A)
miRNA
p
1) Ligate 3´ and 5´ adapters
primer 1
primer 2
(barcoded)
2) RT-PCR
B)
miRNA
p
NNNNN
+
NNNNN
1) Ligate 3´ and 5´ adapters
NNNNN
NNNNN
primer 1
(barcoded)
primer 2
2) RT-PCR
🟦 = sequences required for amplification on flowcell or beads
🟥 = sequence primer hybridization site
🟪 = barcode sequencing primer hybridization site
🟧 = barcode sequence
= universal adapter sequence

Notis Argiriou

THANK YOU