



Genomics and transcriptomics: from raw data to users

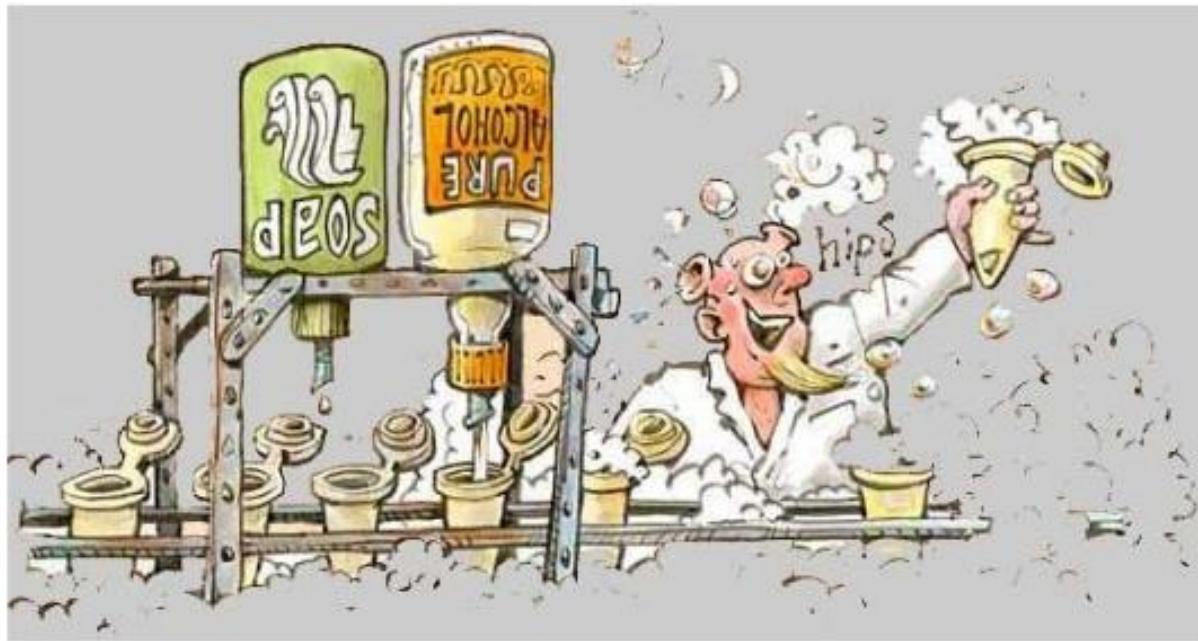
Chiara Colantuono
chiara.colantuono@szn.it

Extracting....

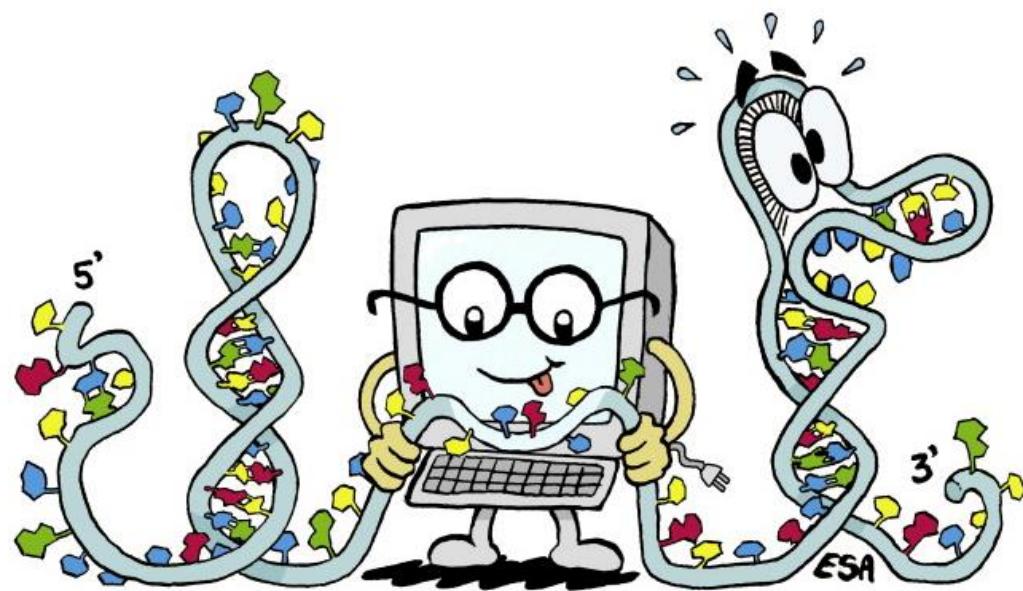
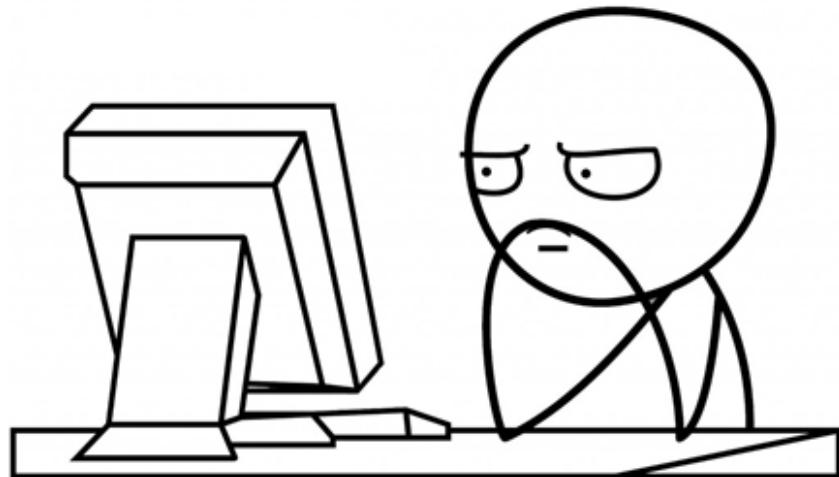


- Species
- Organisms
- Conditions
- Tissues
-

- DNA
- mRNA
- Total RNA
- Single cell
-



Waiting...



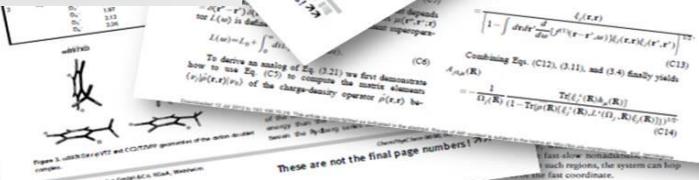


to exchange com
field-drive class
as

fish the address
well as well as
the feasible as the
TSH metric
version of the
programs. At
days, the TSH
available softw
and Newton-X

such as those
who introduced only several
of the main variables in
the model and the
mainly composed by
related approaches. A series
extensions were proposed during
method.²¹ Among them,
Ling et al.,²² and Miller's contribu-
tion²³ on Landau-Zener
with nondiagonal perturba-
tions.²⁴ Another important
made by Kato et al., who intro-
duced the TSH metric
where the driving point was
set correctly.²⁵ In addition,
the authors proposed the
method the TSH method to an ar-
ticle.²⁶ At the same time, Kato
and the model approach also
introduced.²⁷ The concept of
TSH metric was also
introduced.²⁸ The TSH
metric has been explored in recent
the 1990s, the exten-

²¹ C. L. Yu, J. H. Kim, and S. H. Kim, "A new method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).
²² Y. Ling, Z. X. Wang, and J. H. Liu, "A new method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).
²³ R. Miller, "A numerical method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).
²⁴ R. Miller, "A numerical method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).
²⁵ T. Kato, T. Ueda, and T. Yamamoto, "A new method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).
²⁶ T. Kato, T. Ueda, and T. Yamamoto, "A new method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).
²⁷ T. Kato, T. Ueda, and T. Yamamoto, "A new method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).
²⁸ T. Kato, T. Ueda, and T. Yamamoto, "A new method for solving the Schrödinger equation with a time-dependent potential," *J. Korean Phys. Soc.*, **33**, 103–107 (1998).



Sequencing results

You've
G



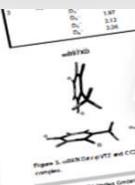
With the address
as well as with
ments and nota-
metals, the fea-
the TSH meth-
available soft-
and Newton-X

such as those
who introduced many of the
tions available only several
years ago. In this paper we review
the current understanding of the
TSH mechanism and the
various models proposed by
researchers. A series
of extensions were proposed during
the 1970s. Among them,
Ling and Miller's contribu-
tion was particularly important.
They proposed a model involving a
coupling between the TSH receptor
and the TSH-binding globulin (TBG)
in the serum. Another important
model was proposed by Kuroda et al., who intro-
duced the TBG-binding point was
not coupled to the receptor. In addition,
they proposed a model involving a
coupling between the TSH receptor and
the TSH-binding globulin (TBG) in the serum.
The concept of max-
imal TSH binding was also proposed.
It has been explored in recent
years.

As a result of the
introduction of the
TSH receptor and the
TBG-binding point, the
various models proposed by
researchers have been
modified and improved.

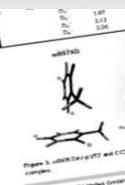
In the 1970s, another
important model was proposed by
Kuroda et al., who intro-
duced the TBG-binding point was
not coupled to the receptor. In addition,
they proposed a model involving a
coupling between the TSH receptor and
the TSH-binding globulin (TBG) in the serum.
The concept of max-
imal TSH binding was also proposed.
It has been explored in recent
years.

These are not the final page numbers!

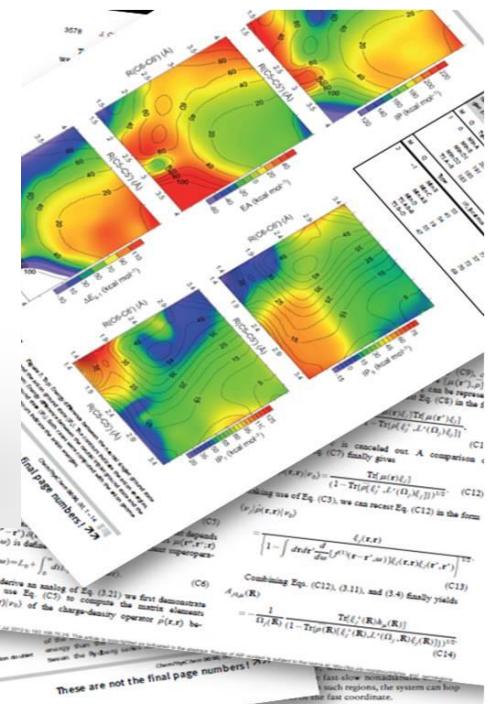


These are not the final page numbers!

These are not the final page numbers!



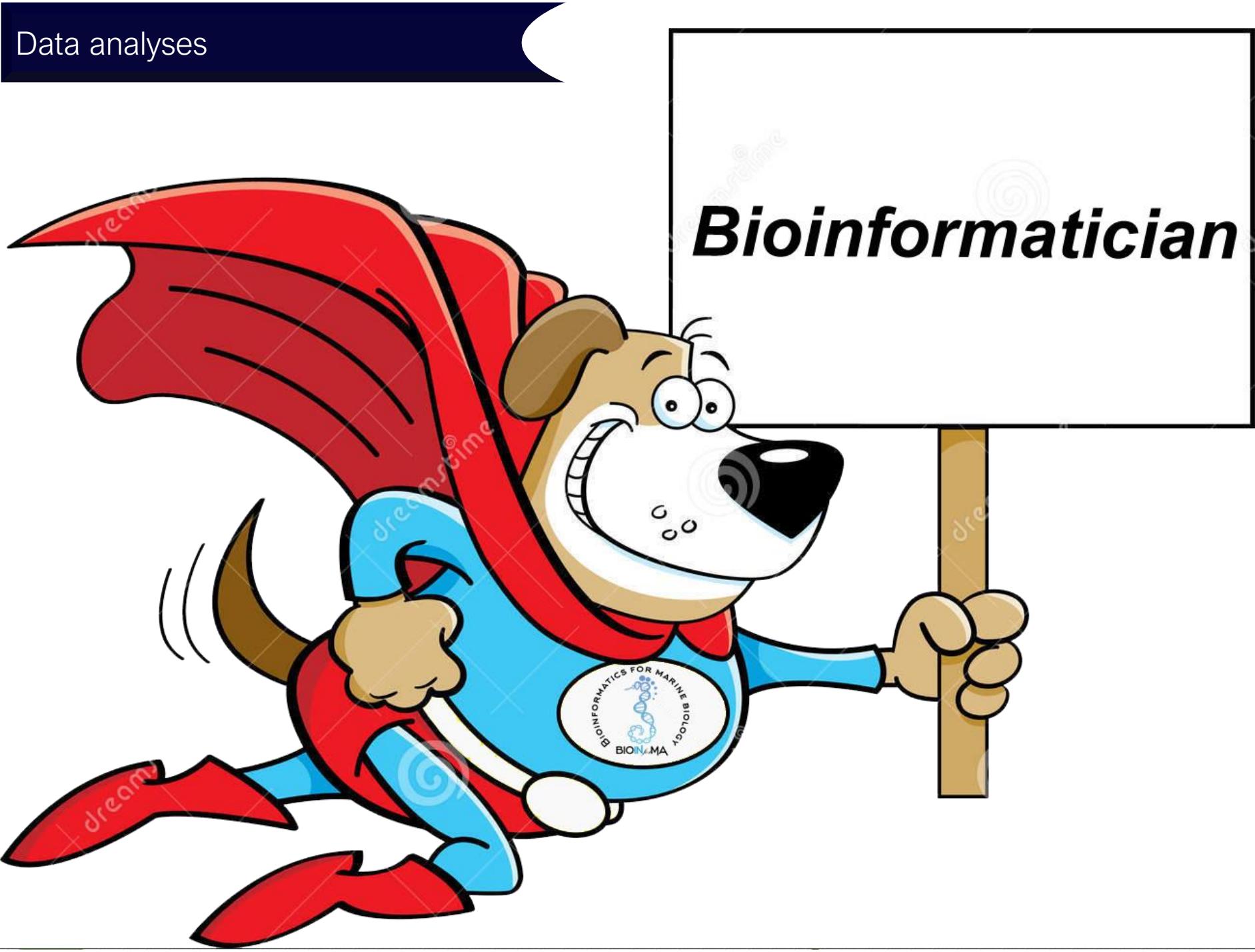
These are not the final page numbers!



Sequencing results

```
AAAAA!EEEEEEEEE<EEEEEEEEEEEEEEEEEEEEEEAEE<EEEEEEEEE<EEE<E<E/EEEEEEEEE<AEE/E  
@NS500352:219:HTTJGAFXX:1:11101:15623:1046 1:N:0:GAGTGG  
CCAGCNCAGGAAGCCGAGGGCGATGAGAAGAAGGGCAGCAAGCAGGTGGGAAGAAGAGGAAGACAAGAGGCCACAGCCTCG  
+  
AAAAA!EEEEEEEEE<EEEEEEEEE<EEEEEEEEE<AEE<EEE/EEEEEEEEE<EEE<A<EEEAAAE  
@NS500352:219:HTTJGAFXX:1:11101:20326:1046 1:N:0:GAGTGG  
GACCTNCGGTGACCATTCGTGATAACACCGTGCCTTCAACGCTGCCAACAGCCTGCCACATGCAACCATTCTGCC  
+  
AAAAA!EEEEEEEEE<EEEEEEEEE<AEE<EEE<EEE//AA/<E/EEEEEAEE/EEAEE/<6E/EEAE<EEA  
@NS500352:219:HTTJGAFXX:1:11101:14871:1046 1:N:0:GAGTGG  
TAAGTNCGCATCGCTGATAACACACGGTAGGGCGAGCTGATCCAACCATTTCTATTTTGTGCAATTGCTTTGCA  
+  
AAAAA!EEEEEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EEEEEEEEE<EEE<EEE<EEE  
@NS500352:219:HTTJGAFXX:1:11101:6926:1046 1:N:0:GAGTGG  
CTGTIANCGAAGCTGCCGCTGCAATACCCATATTAAACGGAACCCGGCTAAGCTTAAAGTAGAAGTGACAACCCCTTTAGT  
+  
AAAAA!EEAEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EEEEEEEEE<EEE<EEE<EEE  
@NS500352:219:HTTJGAFXX:1:11101:4621:1046 1:N:0:GAGTGG  
CCGCANGACGCTGCCAACGTAACATTCAACGGAGCAACGCAATGACCAAGGTAAGGAGTATCTGGCGACCCCCAATCT  
+  
AAAAA!EEAEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EEEEEEEEE<EEE<EEE<EEE  
@NS500352:219:HTTJGAFXX:1:11101:22884:1046 1:N:0:GAGTGG  
TCAAANCATCTGGCAATGCCAGTAACGAGCGCACTCAGCTGATAACGGAGATTACAGCGTGAAAATCAAGAGCTTAAACTGG  
+  
AAAAA!EEAEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EEAEE<EEAEE<EEAEE<EEAEE<EE  
@NS500352:219:HTTJGAFXX:1:11101:13661:1046 1:N:0:GAGTGG  
GATCGNCTCGCCCGGGCGGTGACCGCTCCGTTCAACGGCGGGCGGACGGGAAAGACCCATGCGGCCGAACCGTGCCT  
+  
AAAAA!EEEEEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EE</EE/E6E<EEE<AE/A<EEE  
@NS500352:219:HTTJGAFXX:1:11101:11372:1047 1:N:0:GAGTGG  
ATTGCNTCTCTTATCCGCTGCTGGCGCCCTCGCCGTGACGCTCCGTTGACCGGAAATGCTGG  
+  
AAAAA!EE/EEAEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EEAEE<EEAEE<EEAEE<EE  
@NS500352:219:HTTJGAFXX:1:11101:12828:1047 1:N:0:GAGTGG  
CCTGTNTTCCCTCAGCACCGATAATAACTGCGATCAAAGAAGCTTCAGTCACGCCATGCTAA  
+  
AAAAA!EEEEEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EEAEE<EEAEE<EEAEE<EE  
@NS500352:219:HTTJGAFXX:1:11101:21498:1047 1:N:0:GAGTGG  
GCGAANAGCAGGGTGGTGTGACGCCCTCGCTGACCGGGACCGGAACTGAAAGTGGCGTAAT  
+  
AAAAA!EEAEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EEAEE<EEAEE<EEAEE<EE  
@NS500352:219:HTTJGAFXX:1:11101:2615:1048 1:N:0:GAGTGG  
AAATTNCCTCCCTCTCGTATTGTCAAAATAGTGTACTCCCTCATCAACAGATTGTTTATTTA  
+  
AAA/A!/EEAEEAE/EE/EEE/<E/EEE/E/EEEAE/EEE/<6/EEAEEAE<EEE///E/AE//  
@NS500352:219:HTTJGAFXX:1:11101:26202:1048 1:N:0:GAGTGG  
GTCCGNGGTCAGCTGGCCCTCGTGTGGTGGACTGATACCGACGGCGCAGCATGCTTCCAGTG  
+  
AAAAA!EEAEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EE<EE<EE<EE<EE<EE  
@NS500352:219:HTTJGAFXX:1:11101:11420:1048 1:N:0:GAGTGG  
ATGATNTCGGCTACGCCCTCACTGTAGTTGGCTCTGGAGATACCGACCCAGGAAGACCAAGCTTCTCATTAGGCC  
+  
AAAAA!EEEEEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EE<EE<EE<EE<EE<EE  
@NS500352:219:HTTJGAFXX:1:11101:16386:1049 1:N:0:GAGTGG  
AAATTNCACGAGCAGGAACATTGAGATTTAACTCTAAACTAGGAGCAGTGATTGATGTCACGTGTTGACACGTG  
+  
<AAAAA!EEEEEEEEE<EEEEEEEEE<AEE<EEE<EEE<EEEAE/EE<EE<EE<EE<EE<EE  
@NS500352:219:HTTJGAFXX:1:11101:1201:1049 1:N:0:GAGTGG  
GTTTGNNTCCCATCAACTTACAAAATAGTAAATGGGTATGGGTTAAGGGTGTGCTTTATTTGGGTACAAATTGATAAACTGA  
+  
--More-- (0%)
```

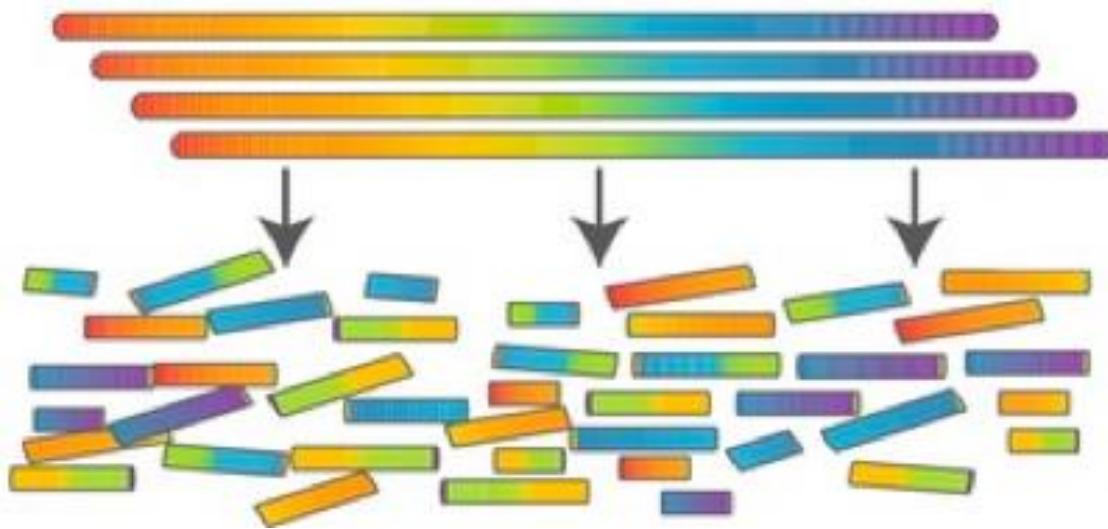




Reads UTRs
RNaseq HeatMap HPC
Genoma Annotation GOenrich
Variant Function Genes
Transposons Adaptor
Transcriptome Assembly
Clean Sequencing mRNA
Structure Promotors
Repeats

RNaseq Reads UTRs
Genome HeatMap HPC
Annotation GOenrich
Function Genes
Transposons Adaptor
Variant Transcriptome Assembly
Clean Sequencing mRNA
Structure Promotors mRNA
Repeats

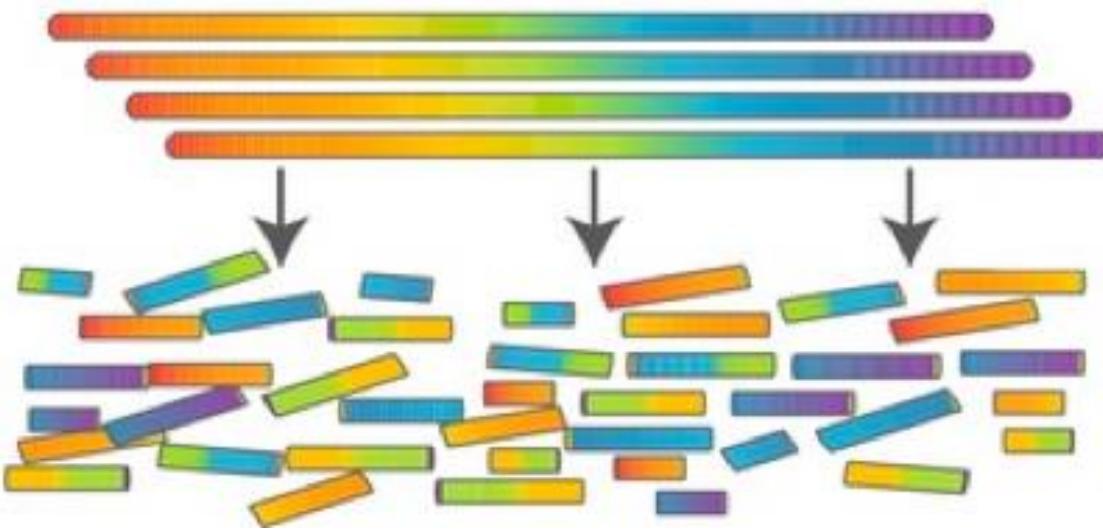
Reads Assembly



Genome/Transcriptome

Sequenced reads

Reads Assembly

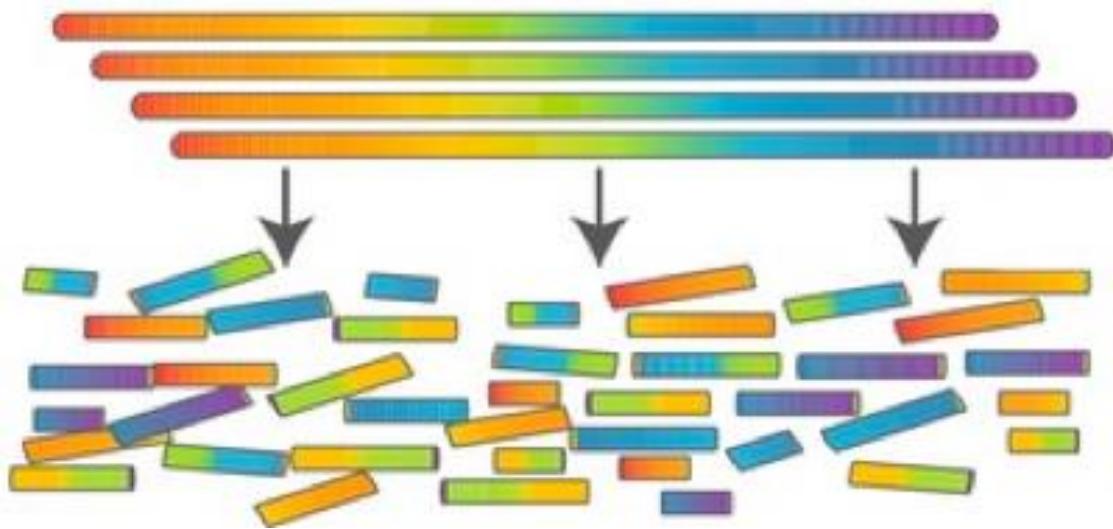


Genome/Transcriptome

Sequenced reads

Technology	Read length	Yield (Reads per Run)
Roche 454	700	~700 thousand
Illumina HiSeq	300	~300 billion
SOLiD	100	~200 billion
Ion Torrent	200	~60 billion
PacBio RS II	14,000	~47 thousand

Reads Assembly

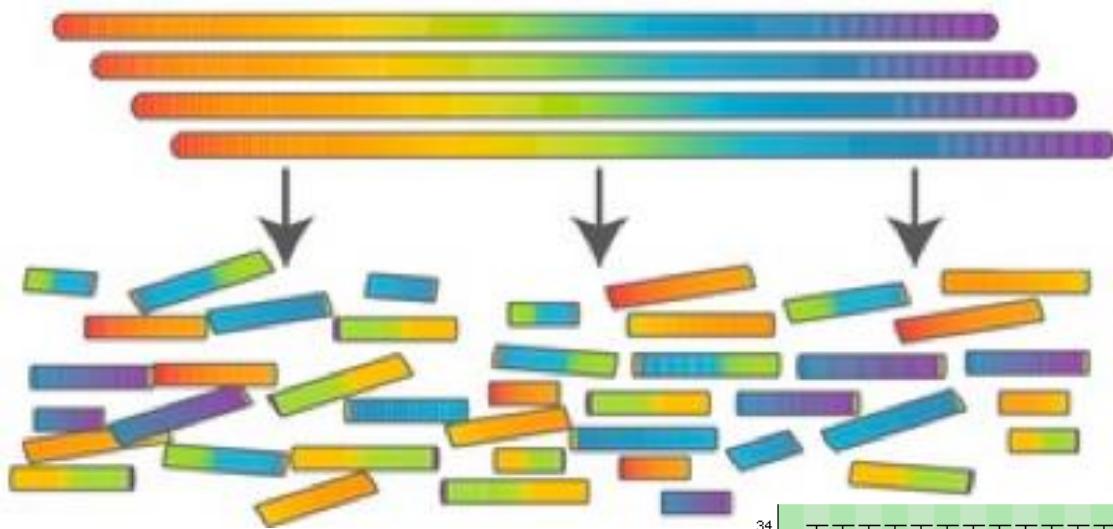


Genome/Transcriptome

Sequenced reads

- single-end
- paired-end R1
- paired-end R2
- mate-pair R1
- mate-pair R2

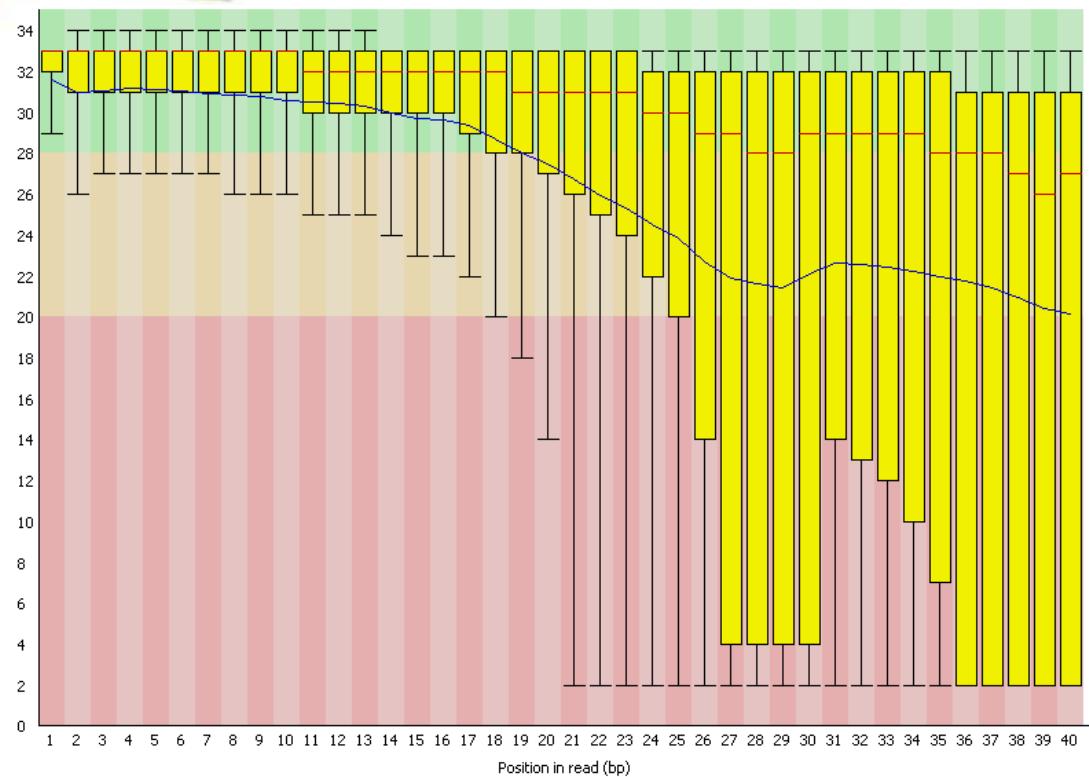
Reads Assembly



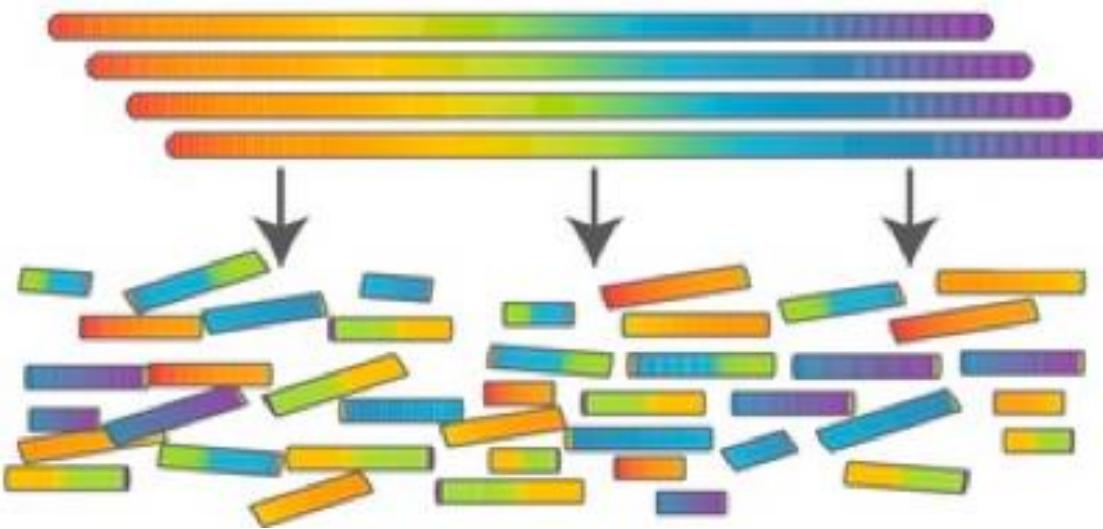
Genome/Transcriptome

Sequenced reads

Quality check



Reads Assembly



Genome/Transcriptome

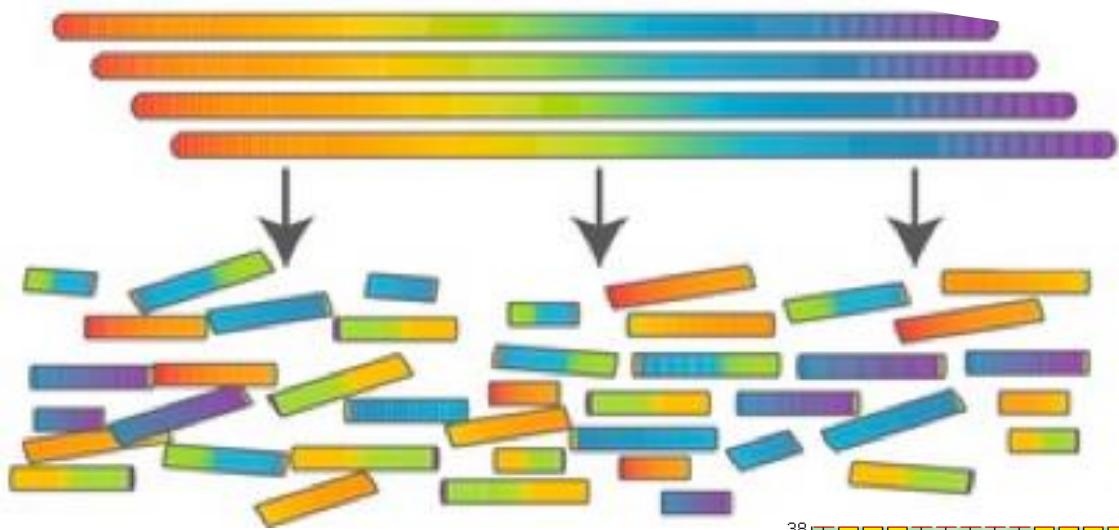
Sequenced reads

Quality check

Adaptors/low quality
bases trimming



Reads Assembly

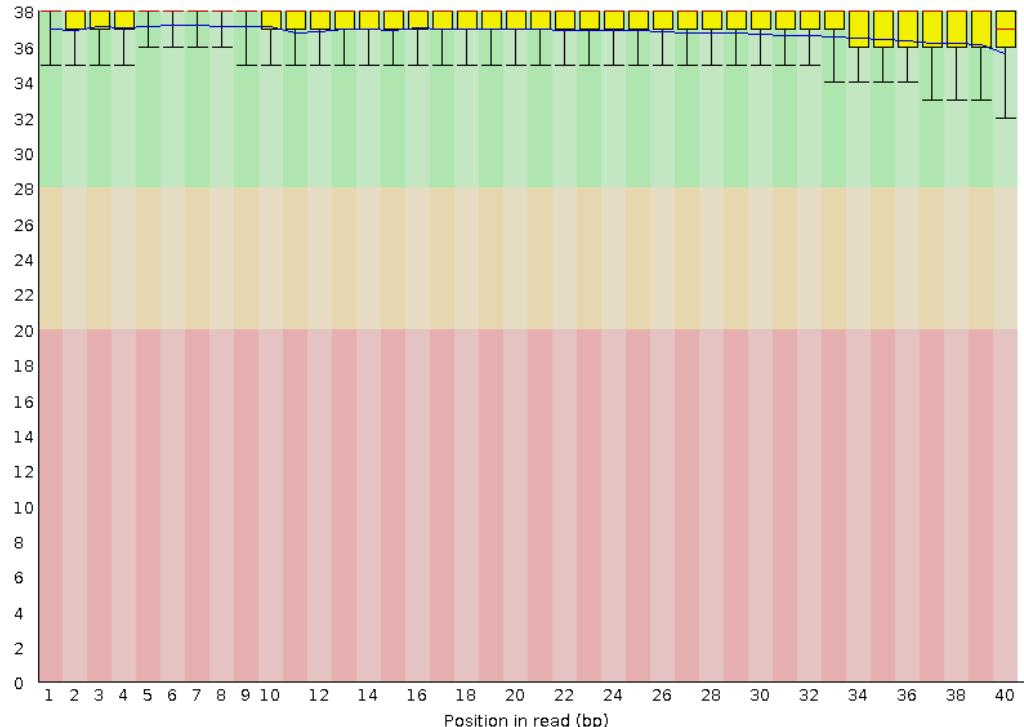


Genome/Transcriptome

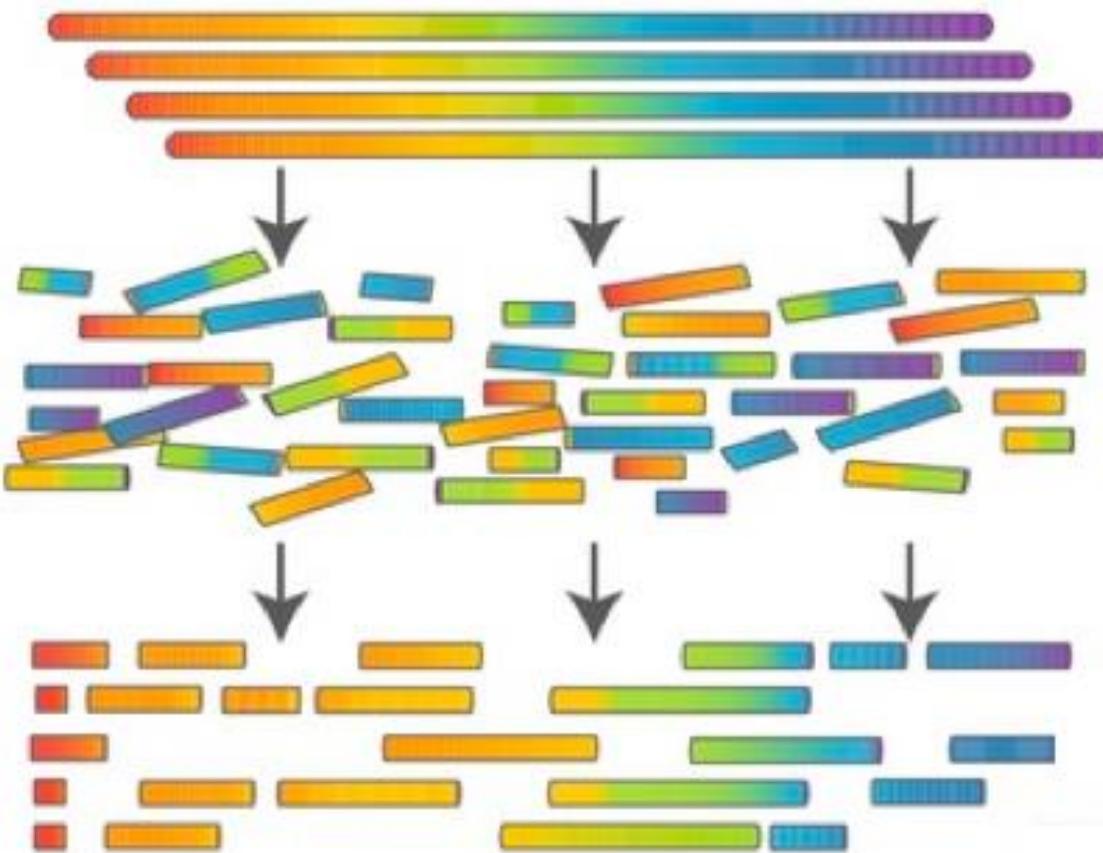
Sequenced reads

Quality check

Adaptors/low quality bases trimming



Reads Assembly



Genome/Transcriptome

Sequenced reads

Path reconstruction by overlap

Consensus

ATGTTCCGATTAGGAAACCTATCTGTAACTGTTCAATTCAAGTAAAAGGGAGGAA



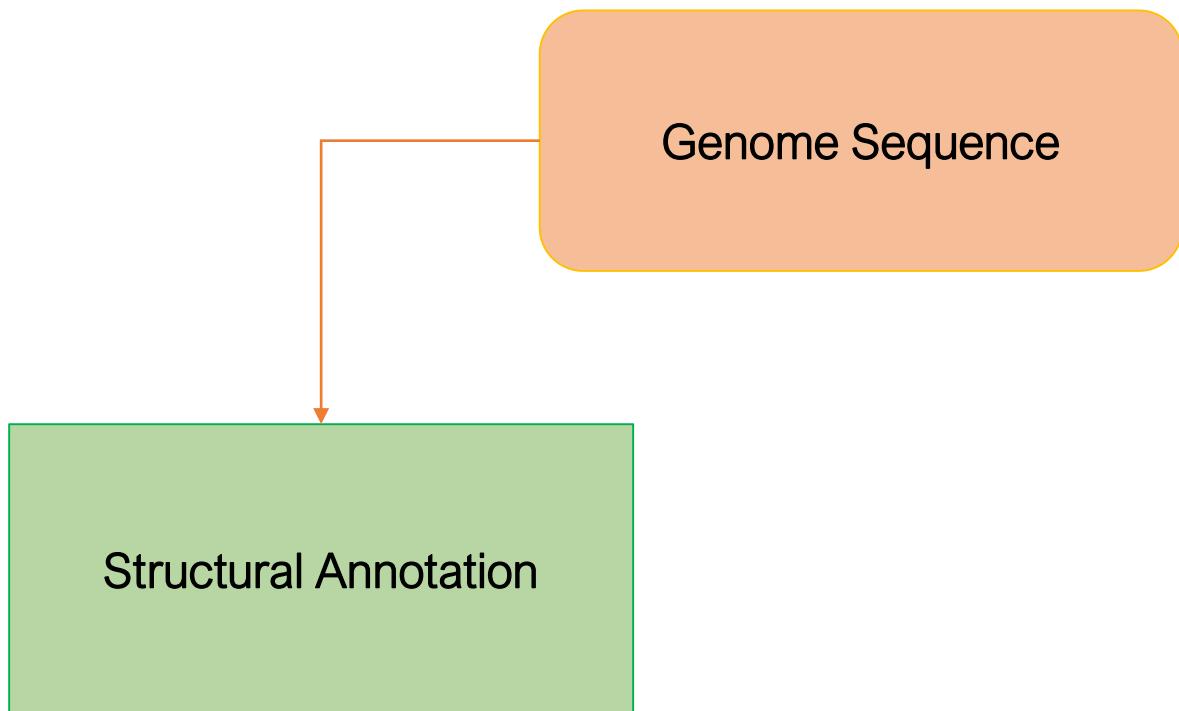
SPADES Assembler 3.0 BETA
ALGORITHMIC BIOLOGY LAB



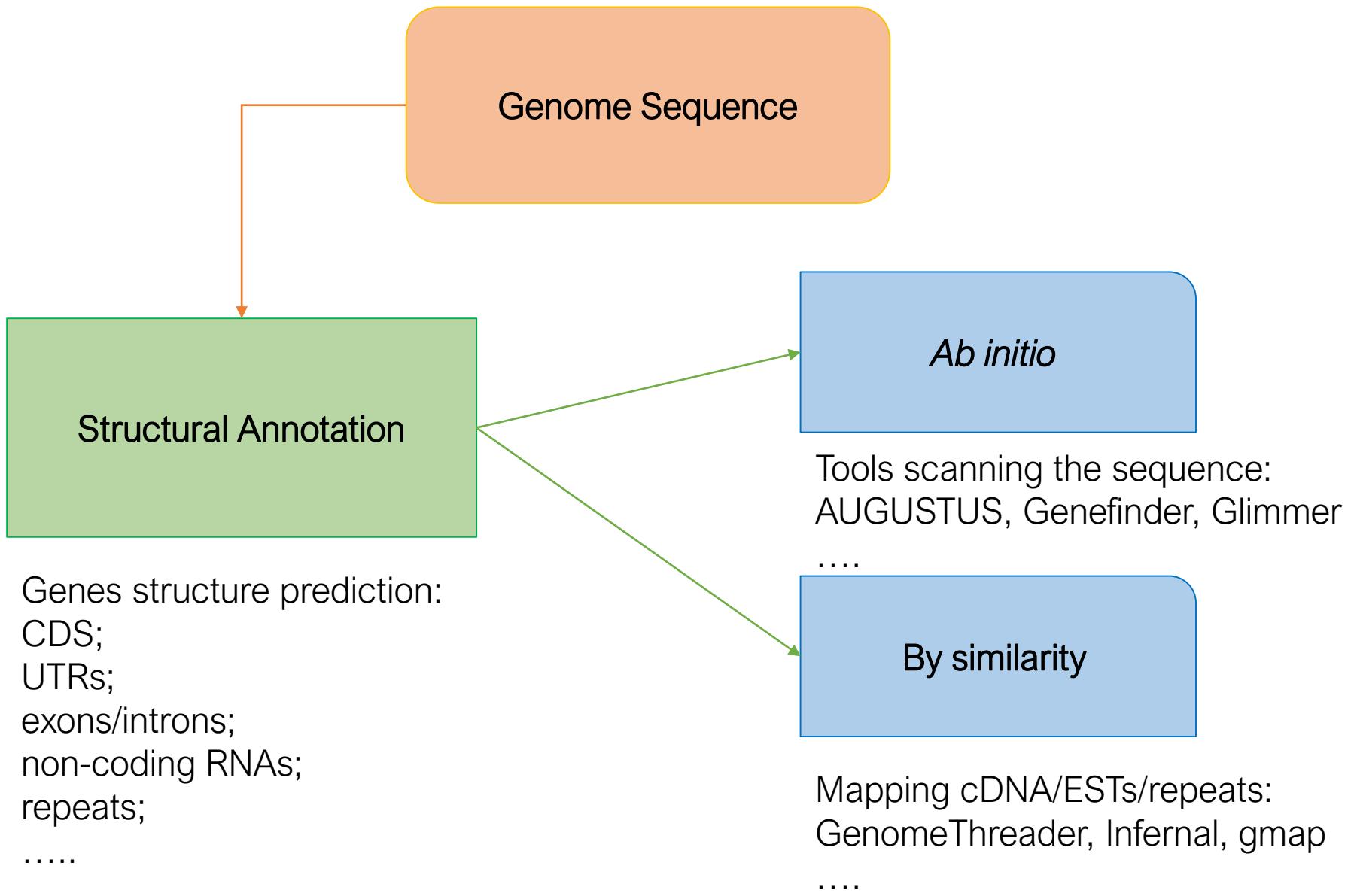
Velvet /
oases

AB~~X~~SS v

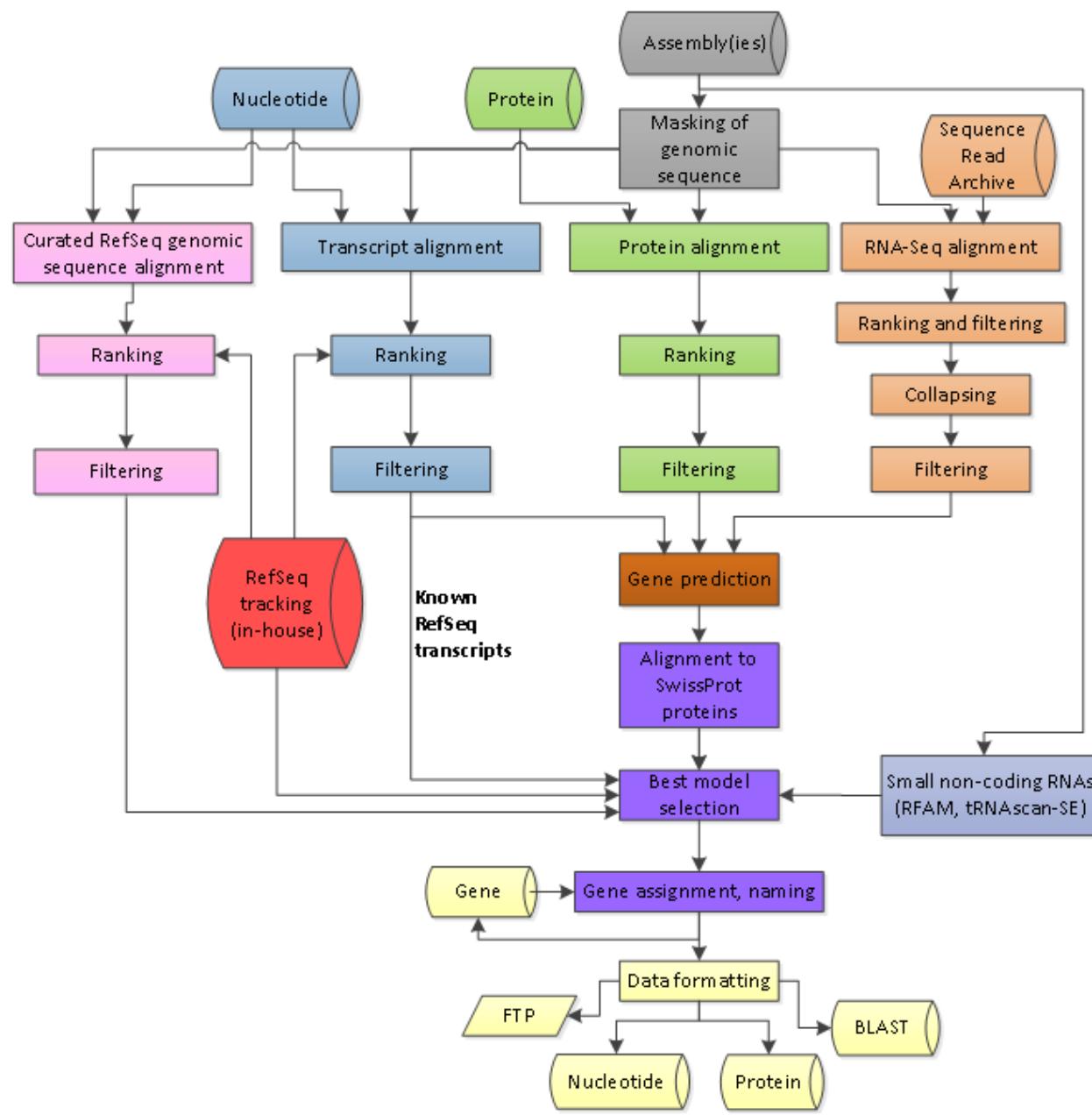
RNAseq Reads UTRs
HeatMap HPC
Genoma GOenrich
Annotation Genes
Function Adaptor
Variant Transposons
Transcriptome Assembly
Clean Sequencing
Structure mRNA
Promotors mRNA
Repeats



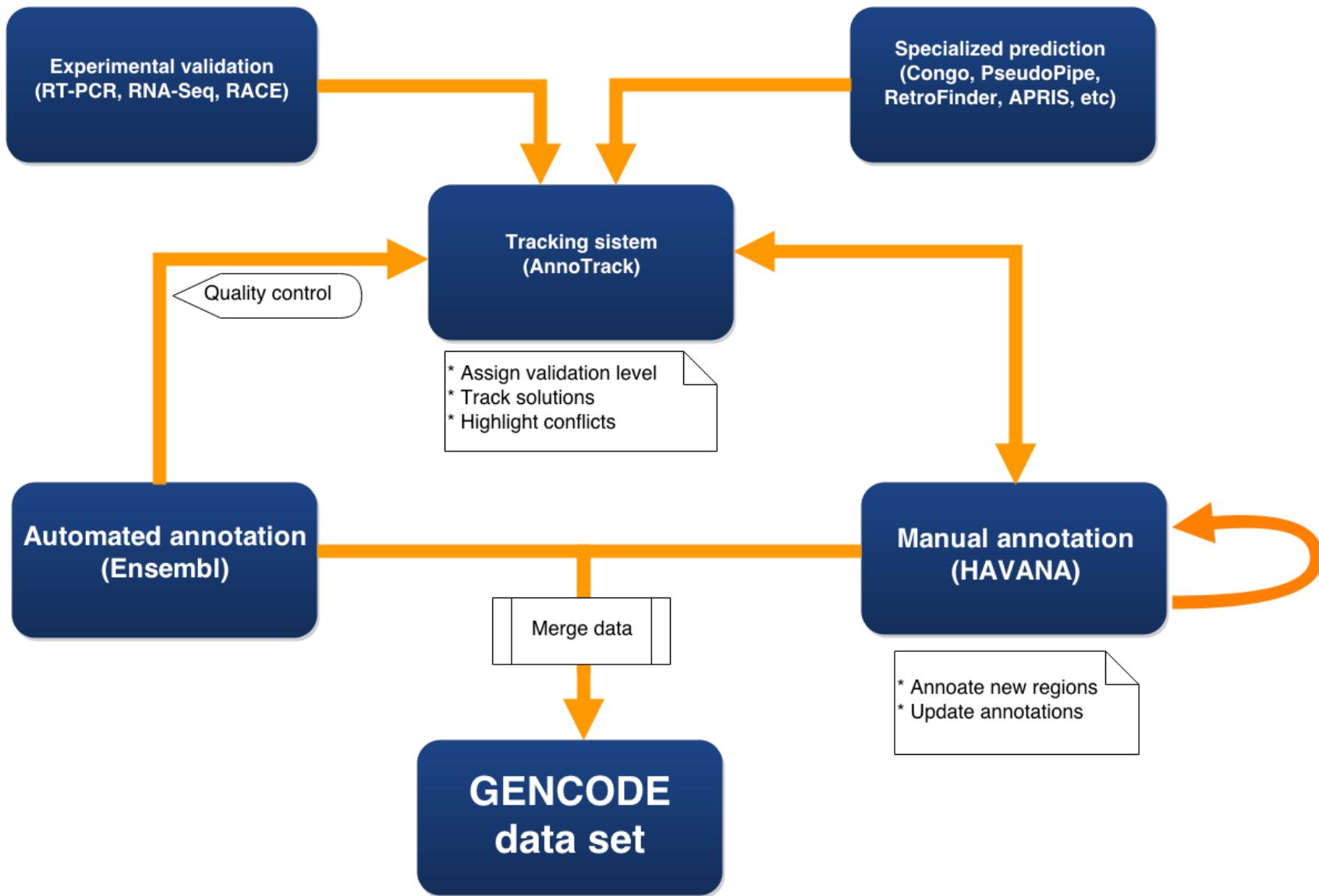
Genes structure prediction:
CDS;
UTRs;
exons/introns;
non-coding RNAs;
repeats;
.....



Genome Annotation: the RefSeq pipeline



Genome Annotation: the GENCODE pipeline



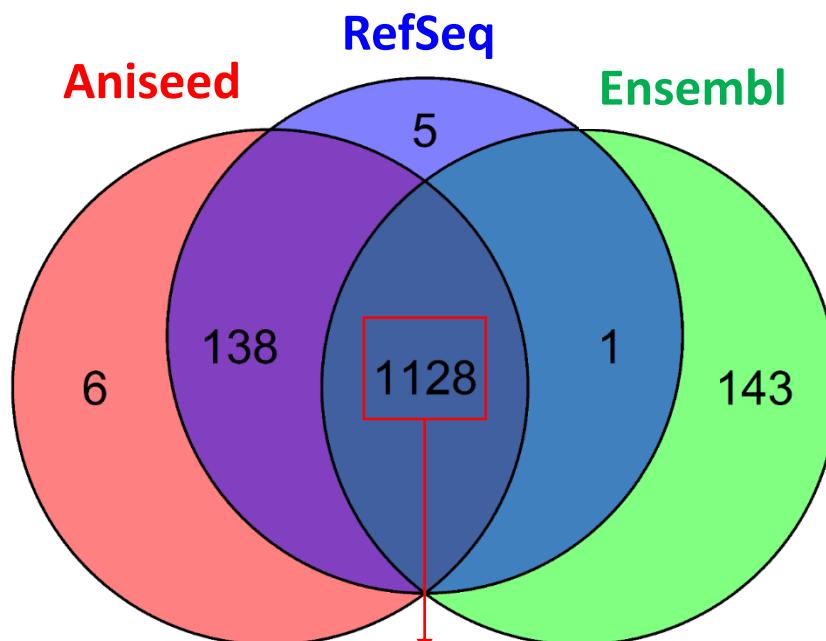
Available resources:
the example of *C. robusta*

Ciona robusta



	ANISEED	REFSEQ	ENSEMBL
ANISEED	1272	1266	1128
REFSEQ	1266	1272	1129
ENSEMBL	1128	1129	1272

Number of genome scaffolds



Number of scaffolds
in common among
all the assemblies

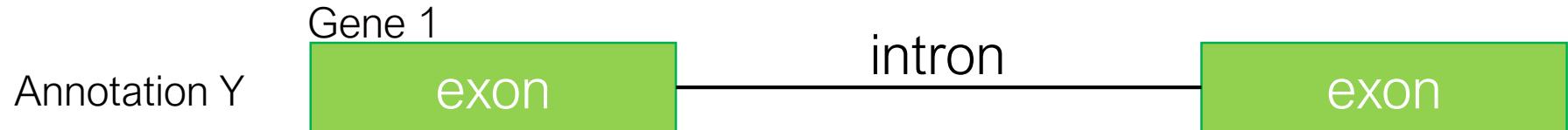
ANISEED SPECIFIC	REFSEQ SPECIFIC	ENSEMBL SPECIFIC
chr4	chr4	chr1
chr9	chr9	chr2
chr11	chr11	chr3
chr12	chr12	chr4
chr13	chr13	chr5
MIT		chr6
		chr7
		chr8
		chr9
		chr10
		chr11
		chr12
		chr13
		chr14
		129 SCAFFOLDS

The differences among the scaffolds are mainly inversions or a different number of undefined nucleotides (Ns).

Annotation comparison

Identical exons start and end

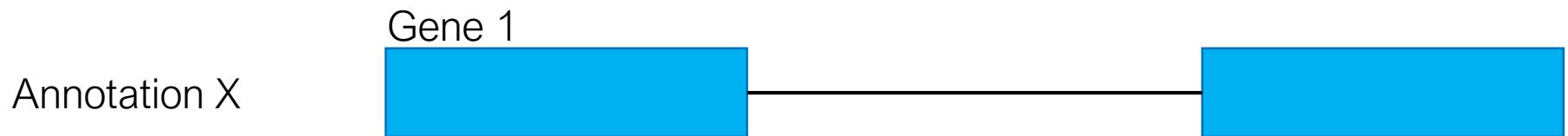
Genome sequence



Annotation comparison

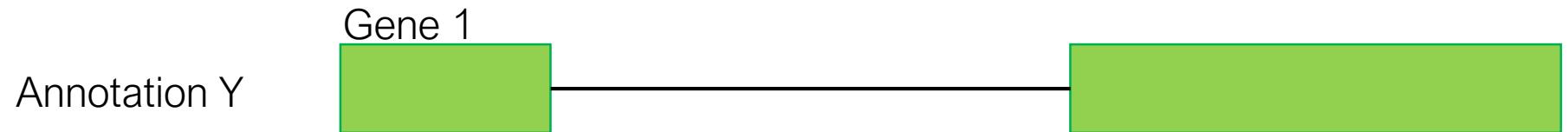
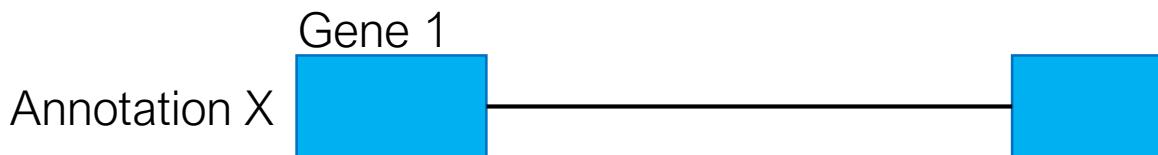
Genes overlapping but with a different gene structure

Genome sequence



Genes overlapping but with a different gene structure

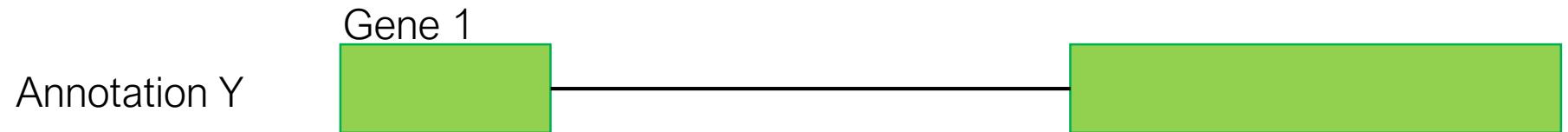
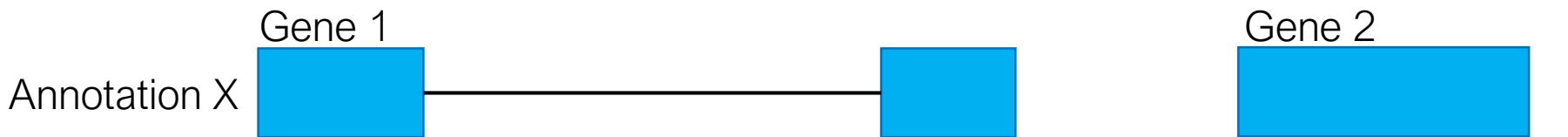
Genome sequence



Annotation comparison

Genes overlapping but with a different gene structure

Genome sequence



Annotation comparison

Not overlapping genes,
gene annotation specific

Genome sequence

Gene 1



Annotation X

Gene 2



Annotation Y

Collecting marine data...

2

Acanthaster planci

<http://www.ncbi.nlm.nih.gov/>
<http://marinegenomics.oist.jp/>



Patiria miniata

<http://www.echinobase.org/>
<http://www.ncbi.nlm.nih.gov/>

2

2

Branchiostoma floridae

<http://www.ncbi.nlm.nih.gov/>
<http://genome.jgi.doe.gov/>



Phaeodactylum tricornutum

<http://protists.ensembl.org/>
<http://genome.jgi.doe.gov/>
<http://www.ncbi.nlm.nih.gov/>

3

3

Ciona robusta

<http://www.ensembl.org/>
<http://www.aniseed.cnrs.fr/aniseed/>
<http://www.ncbi.nlm.nih.gov>



Octopus bimaculoides

<http://www.metazome.net/>
<https://groups.oist.jp/>
<http://www.ncbi.nlm.nih.gov/>

3

3

Ciona savignyi

<http://www.ensembl.org/>
<http://www.aniseed.cnrs.fr/aniseed/>
<http://www.ncbi.nlm.nih.gov>



Saccoglossus kowalevskii

<http://www.ncbi.nlm.nih.gov>
<https://groups.oist.jp/>

2

3

Gasterosteus aculeatus

<http://www.ncbi.nlm.nih.gov>
<http://www.ensembl.org>
<http://sticklebrowser.stanford.edu/>



Strongylocentrotus purpuratus

<http://www.echinobase.org/>
<https://www.hgsc.bcm.edu/>
<http://www.ncbi.nlm.nih.gov/>

3

GENOMA: GENOMES FOR MARINE BIOLOGY



<http://bioinfo.szn.it/genoma/>

GENOMA: GENOMES FOR MARINE BIOLOGY

Species	NCBI_latest_versions	ENSEMBL_latest_versions	JGI_latest_versions	OTHERS_latest_versions
<i>Acanthaster planci</i>	GCF_001949145.1	NA	NA	MarineGenomics: oki-cotsv1
<i>Branchiostoma floridae</i>	GCF_00003815.1_V2	NA	Branchiostoma_floridae_v2.0	NA
<i>Ciona robusta</i>	GCF_000224145.2_KH	GCA_000224145.1_KH	Assembly v2.0	Aniseed: KH12
<i>Ciona savignyi</i>	GCA_000149265.1	CSAV 2.0	NA	Aniseed: ENS81
<i>Gasterosteus aculeatus</i>	GCA_000180675.1	BROAD S1	NA	UCSC: gasAcu1
<i>Octopus bimaculoides</i>	GCF_001194135.1_V2.0	PRJNA270931	Obimaculoides_280	NA
<i>Patiria miniata</i>	GCA_000285935.1	NA	NA	Echinobase: V2
<i>Phaeodactylum tricornutum</i>	GCF_000150955.2	ASM15095v2	Phatr2	NA
<i>Ptychoderma flava</i>	NA	NA	NA	OIST: 1426034866.unique_0.7_1e-10
<i>Saccoglossus kowalevskii</i>	GCF_000003605.2	NA	NA	OIST: V3
<i>Strongylocentrotus purpuratus</i>	GCF_000002235.4_Spur_4.2	Spur_3.1	NA	Echinobase: v4.2
<i>Zostera marina</i>	GCA_001185155.1_Zosma_marinav.2.1	NA	NA	NA

[back to the GENOMA home page](#)



***Ciona robusta* genome portal**

Genome information overview:

there are three different genome assemblies available:

- GCF_000224145.3 from NCBI;
- KH.92 from ENSEMBL;
- KHGene.2012 from Aniseed.

All the assemblies consist of 14 chromosomes plus the mitochondrion and 1257 unplaced scaffolds, having in total 1272 sequences. We compared the sequences of the assemblies, checking how many sequences are completely identical. The number of genomic elements in common (identical sequence) among the assemblies is reported in the following table:

	NCBI	ENSEMBL	ANISEED
NCBI	1272	1129	1266
ENSEMBL	1129	1272	1128
ANISEED	1166	1128	1272

The GENOMA platform

Gene content overview:

Search: Show 50 entries

Type	Gene ID	Chromosome	Start	End	Strand	Annotation
RefSeq	XM_018815685.1	Chr01	2407	20500	-	regulator of G-protein signaling 22-like
ANISEED	KH.C1.286.v1.R.ND1-1	Chr01	2417	20402	-	ORF:[<1..4215] Homolog:[cDNA FLJ40080 fis clone TESTI2001795 highly similar to Homo sapiens regulator of G-protein signalling 22 (RGS22) mRNA]
ANISEED	KH.C1.286.v1.A.SL2-1	Chr01	2417	20413	-	ORF:[12..4226] Homolog:[cDNA FLJ40080 fis clone TESTI2001795 highly similar to Homo sapiens regulator of G-protein signalling 22 (RGS22) mRNA]
ENSEMBL	ENSCING00000016550	Chr01	4219	20394	-	n/a
ANISEED	KH.C1.560.v1.C.ND1-1	Chr01	14261	14503	+	ORF:[*90..245]
RefSeq	XM_002123150.2	Chr01	23507	52048	+	ileal sodium/bile acid cotransporter-like
ANISEED	KH.C1.907.v1.A.SL3-1	Chr01	23513	34214	+	ORF:[1..588] Homolog:[Solute carrier family 10 member 6]
ENSEMBL	ENSCING00000016539	Chr01	23518	34155	+	n/a
ANISEED	KH.C1.907.v1.R.ND1-1	Chr01	23530	34214	+	ORF:[<1..571] Homolog:[Solute carrier family 10 member 6]
ANISEED	KH.C1.907.v2.B.ND2-1	Chr01	24117	34214	+	ORF:[<1..965] Homolog:[similar to solute carrier family 10 member 2]
RefSeq	XM_018815677.1	Chr01	33654	60443	-	brefeldin A-inhibited guanine nucleotide-exchange protein 2 transcript variant X3
RefSeq	XM_018815674.1	Chr01	36633	60443	-	brefeldin A-inhibited guanine nucleotide-exchange protein 2 transcript variant X2
RefSeq	XM_009863811.2	Chr01	36633	60443	-	brefeldin A-inhibited guanine nucleotide-exchange protein 2 transcript variant X1

The GENOMA platform



Download information:

Source	Genome Assembly (FASTA-big file)	Gene Annotation (GFF3)	Gene Annotation (CSV)	Gene Sequences (FASTA)	mRNA Sequences (FASTA)	Protein Sequences (FASTA)	INTERPRO results (CSV)
NCBI	GCF_000224145.3	GCF_000224145.3	download	download	download	download	download
ENSEMBL	KH.92	KH.92	download	download	download	download	download
ANISEED	KHGene.2012	KHGene.2012	download	download	download	download	download

Genome Assembly: FASTA file of all the chromosome/scaffolds/contigs;

Gene Annotation: GFF file containing information about all the annotated genes, mRNAs, UTRs, CDSs;

Gene Annotation: CSV file containing information ONLY about the annotated genes (Gene ID, chromosome, start, end, strand, functional annotation);

Gene Sequences: FASTA file of all the gene sequences;

mRNA Sequences: FASTA file of all the mRNA sequences;

Protein Sequences: FASTA file of all the protein sequences;

INTERPRO results: CSV file containing the INTERPROSCAN results (Gene ID, Interpro, PFAM, Panther, Pathways).

[back to the GENOMA home page](#)



Ciona robusta genome portal

Browse the genome through the [JBrowse](#) interface

Available Tracks

filter tracks

- Ciona robusta - ANISEED KH2012 annotation
- Ciona robusta - ENSEMBL KH.92 annotation
- Ciona robusta - REFSEQ GCF_000224145.3 annotation
- Ciona robusta ESTs

Reference sequence

REFSEQ GCF_000224145.3 assembly

Genome **Track** **View** **Help**

0,000 4,000,000 5,000,000 6,000,000 7,000,000 8,000,000 9,000,000 10,000

NC_020166.2 NC_020166.2:..10041005 (10.04 Mb) Go

3,750,000 5,000,000 6,250,000 7,500,000 8,750,000 10,000

C. robusta - REFSEQ GCF_000224145.3 assembly

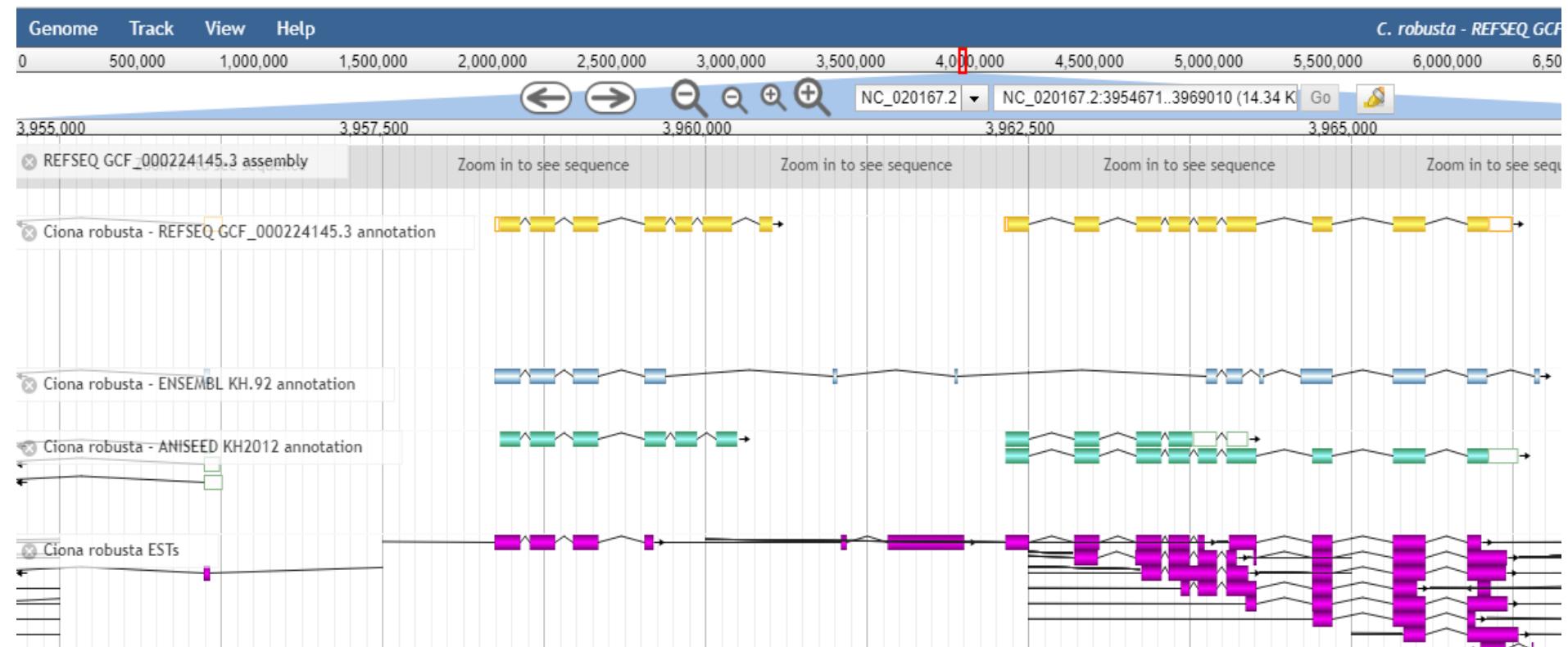
C. robusta - REFSEQ GCF_000224145.3 annotation (feature density)

C. robusta - ENSEMBL KH.92 annotation (feature density)

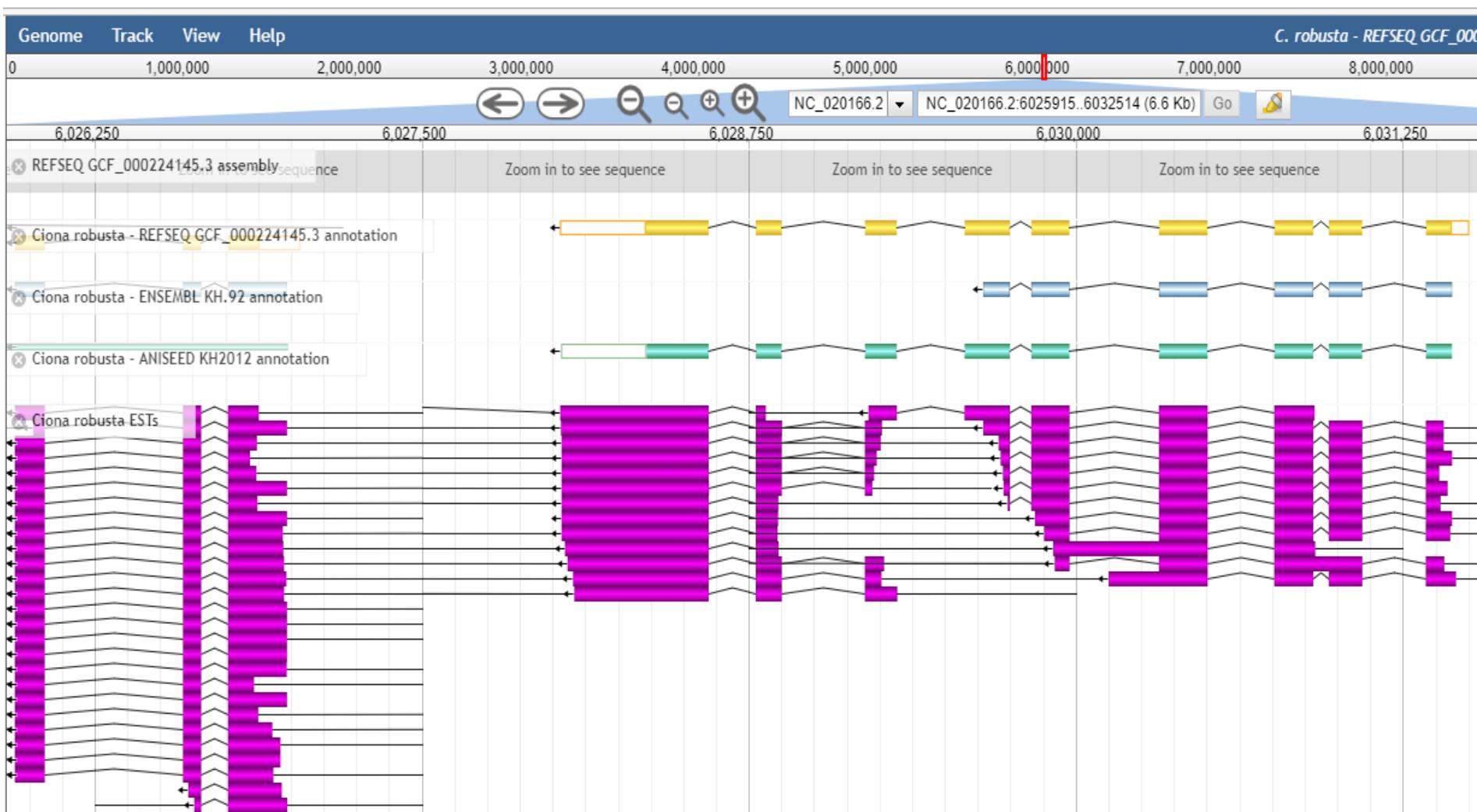
C. robusta - ANISEED KH2012 annotation (feature density)

C. robusta ESTs (feature density)

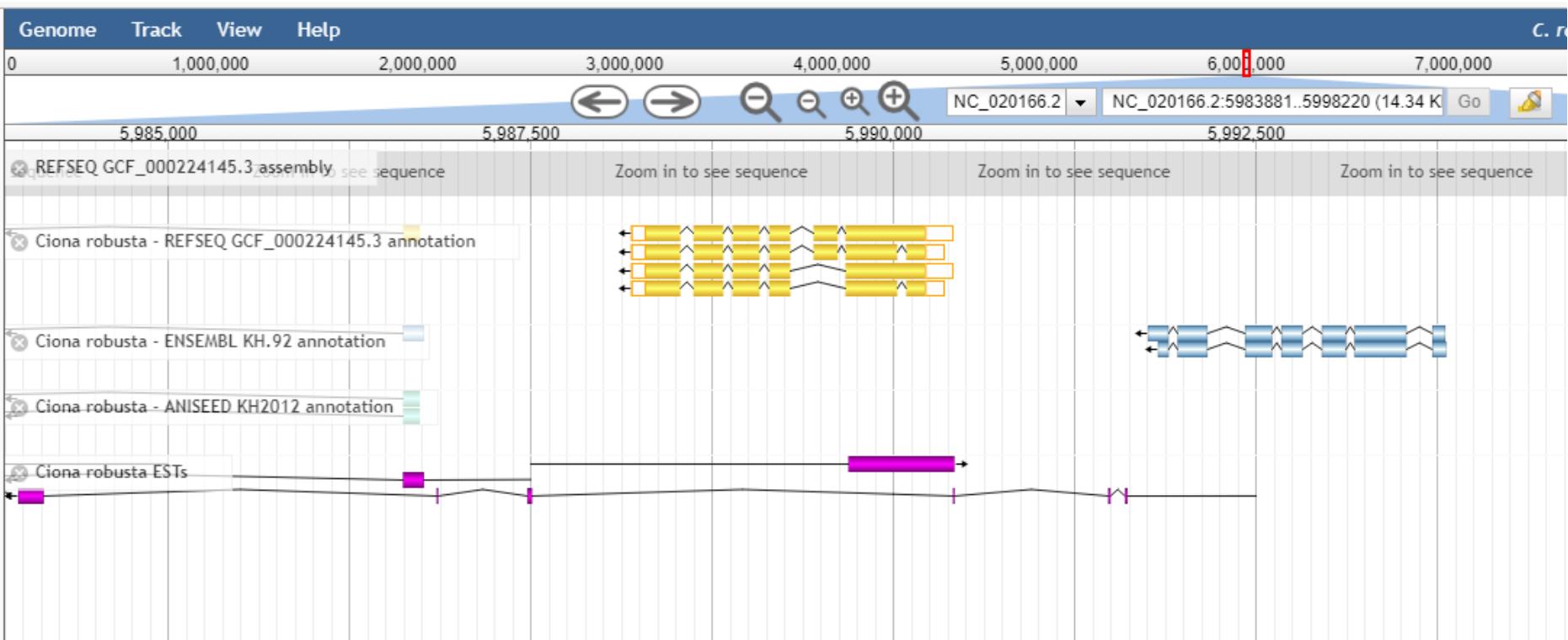
The GENOMA platform



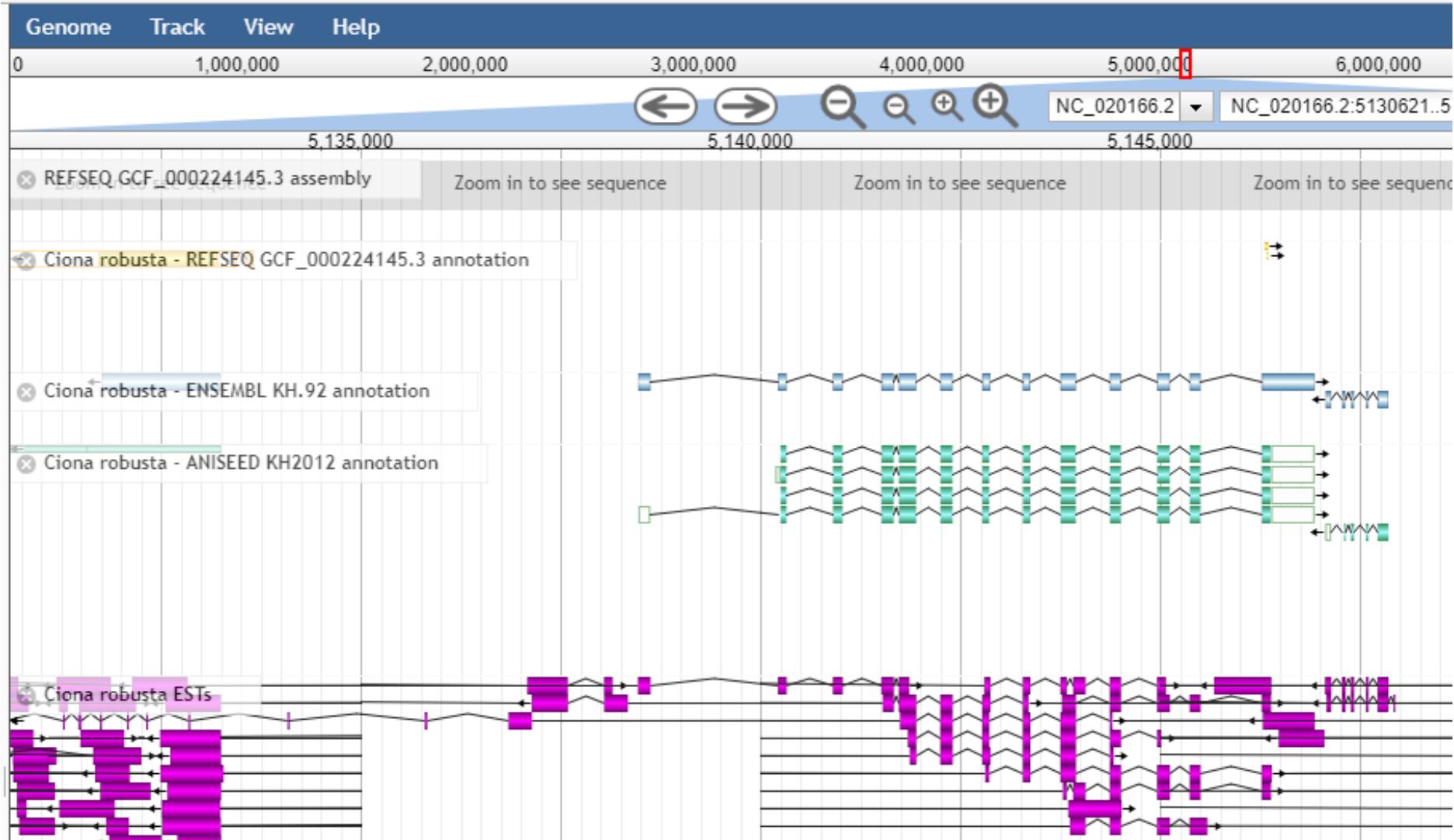
The GENOMA platform



The GENOMA platform



The GENOMA platform



[back to the GENOMA home page](#)



***Ciona robusta* genome portal**

Browse the genome through the [JBrowse](#) interface

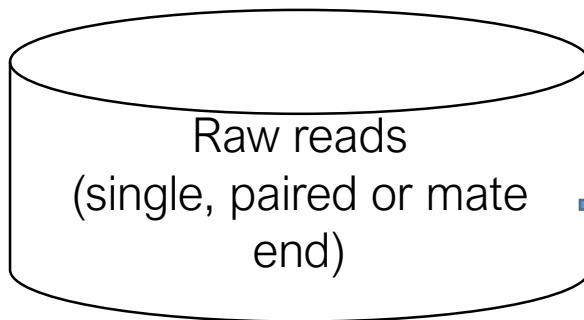
Search for specific elements through the [Query page](#)



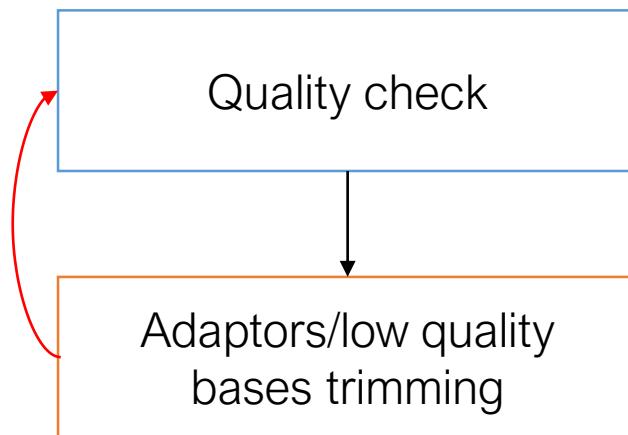
RNaseq Reads UTRs
Genoma HeatMap HPC
Variant Annotation GOenrich
Function Genes
Transposons Adaptor
Transcriptome Assembly
Clean Sequencing mRNA
Structure Promotors
Repeats

Experimental design
(conditions, stages...)

Sequencing

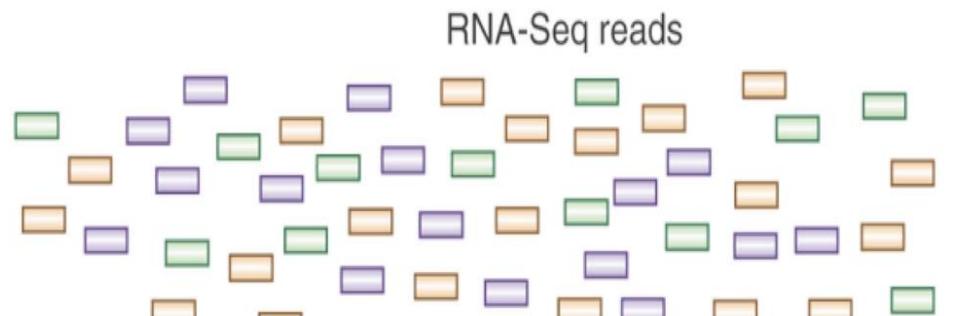


Data pre-processing

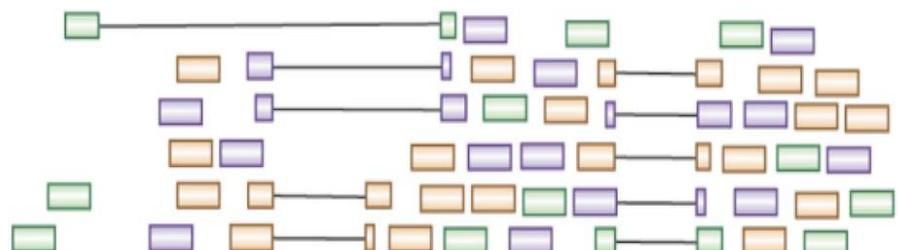


Data analyses

Reads mapping



RNA-Seq reads



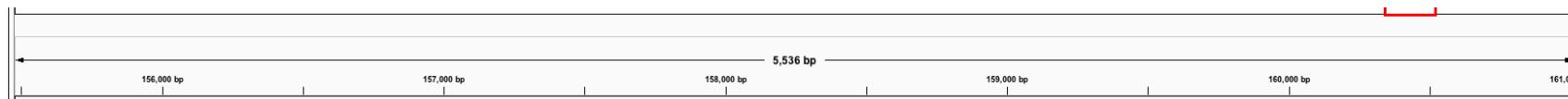
Genome

Differentially Expressed Genes

Data analyses

Reads mapping

Genome



Cond. 1



Cond. 2



Annotation

Data analyses

Reads mapping

Count and quantification



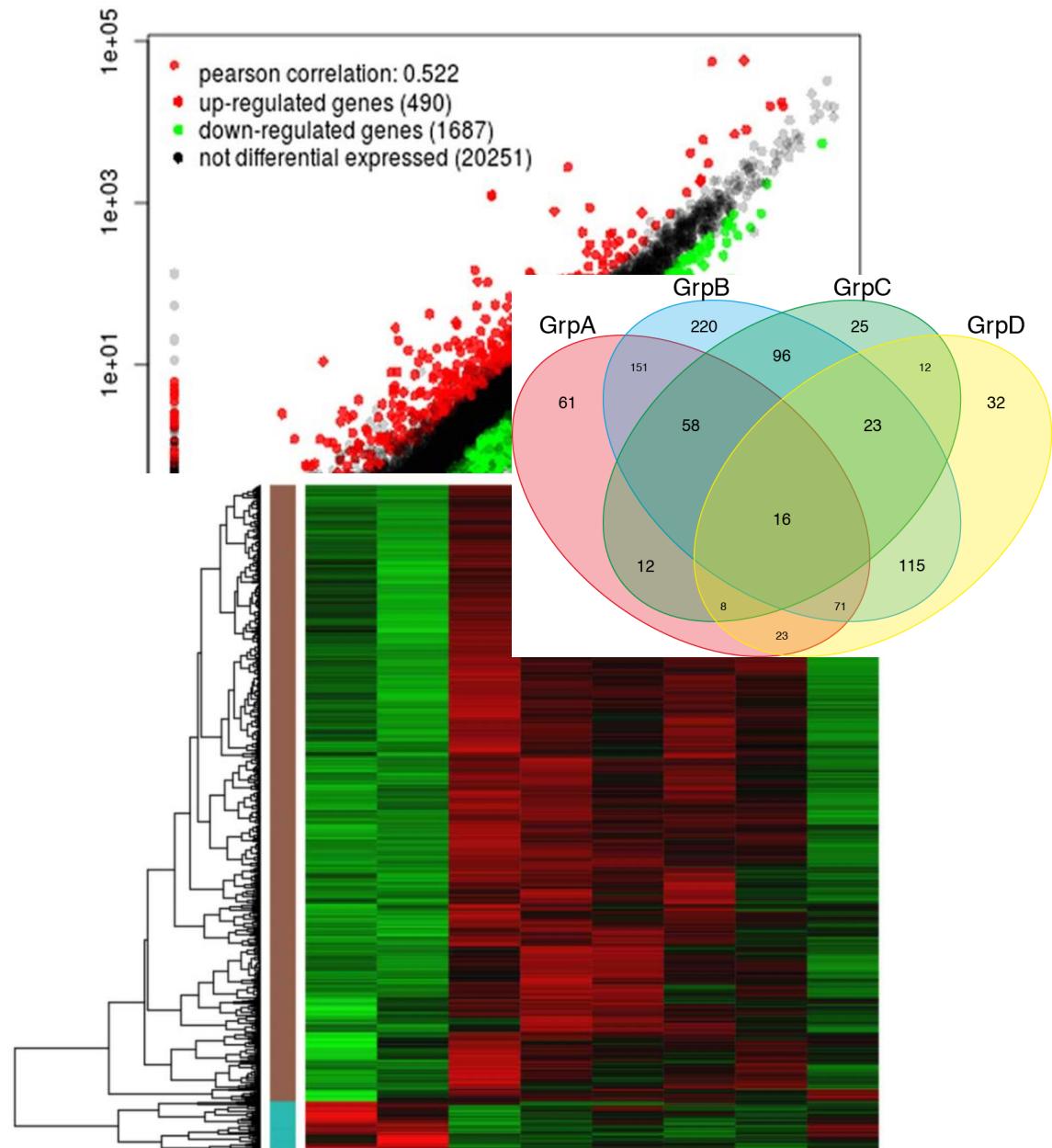
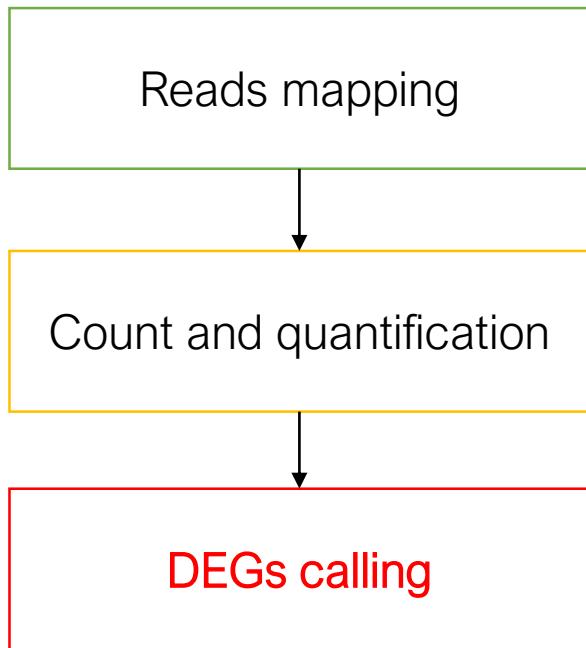
DESeq
edgeR
Cuffdiff

...

Geneid	C1	C2	C4	C5	C7	C8
AT1G01010.1	201	182	147	133	229	217
AT1G01020.1	224	218	229	206	187	167
AT1G01020.2	218	216	226	204	186	166
AT1G01030.1	75	94	47	49	69	64
AT1G01040.1	1017	1107	902	866	888	1083
AT1G01040.2	1001	1088	890	855	873	1072
AT1G01046.1	39	50	28	34	34	27
AT1G01050.1	763	810	732	685	695	805
AT1G01060.1	1442	1503	17	16	6716	9911
AT1G01060.2	1442	1505	17	16	6718	9913
AT1G01060.3	1357	1427	15	15	6321	9431
AT1G01060.4	1442	1503	17	16	6716	9911
AT1G01060.5	1376	1441	16	15	6401	9519
AT1G01070.1	192	160	117	104	73	93
AT1G01070.2	192	160	117	104	73	93
AT1G01073.1	0	0	0	0	0	1
AT1G01080.1	323	324	357	310	231	237
AT1G01080.2	324	324	357	310	231	237
AT1G01090.1	1735	1790	1752	1542	1698	1823
AT1G01100.1	1398	1562	1602	1464	1381	1559
AT1G01100.2	1398	1563	1603	1463	1381	1558
AT1G01100.3	1398	1563	1603	1463	1381	1558
AT1G01100.4	1398	1563	1603	1463	1381	1558

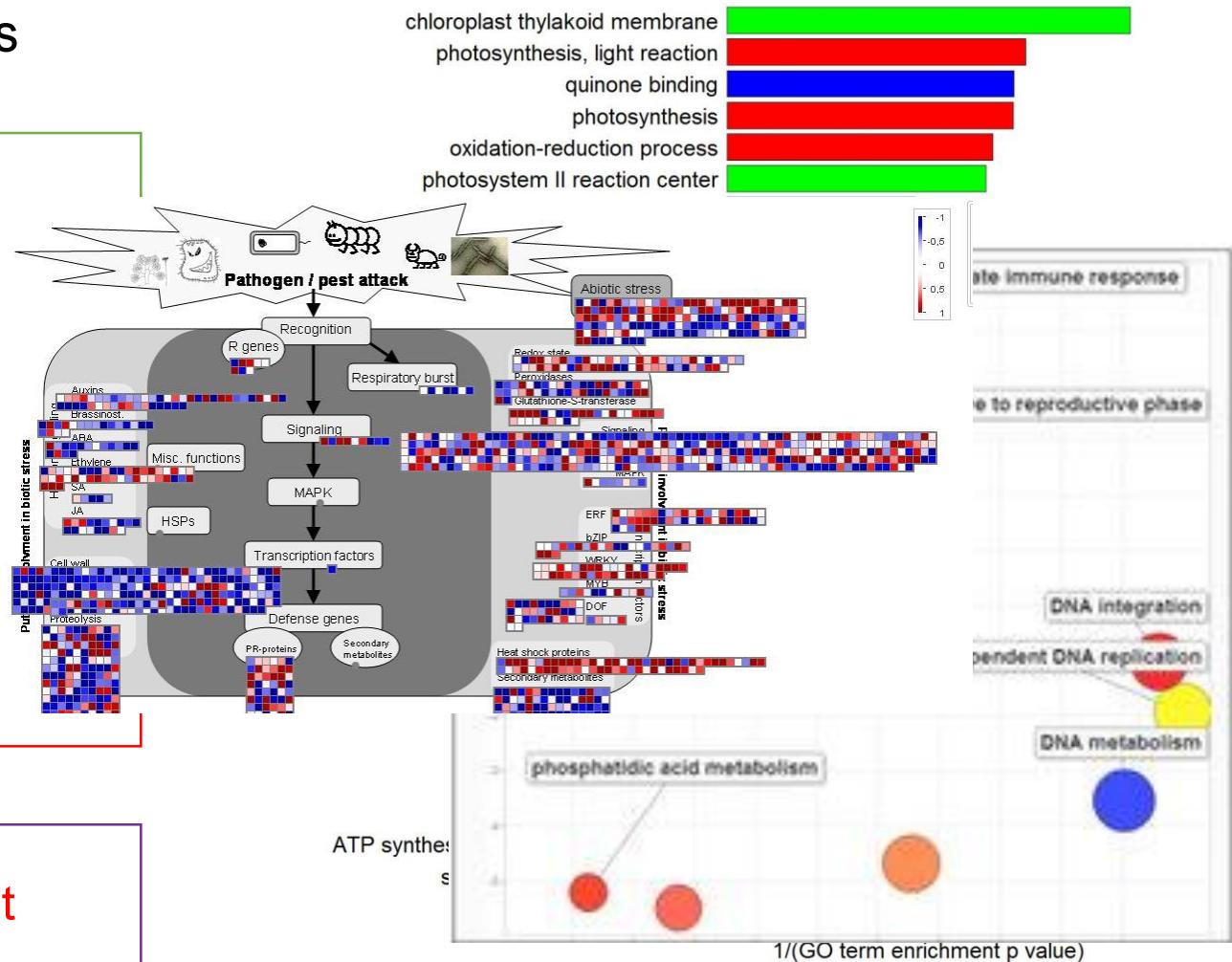
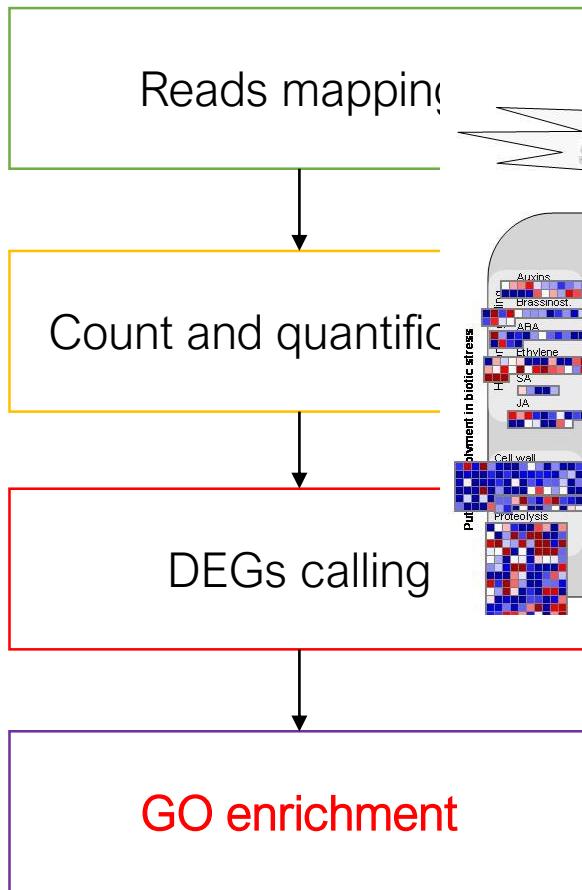
Differentially Expressed Genes

Data analyses



Differentially Expressed Genes

Data analyses



Genomics

Genome Composition (GC/CpG etc content)
Gene annotation structure quality check (via GFF)
Gene-mRNA-protein sequences extraction from GFF
GFF/GTF/BED files parsing and manipulation
SAM file parsing and manipulation
Genome sequence extraction
Reads quality check and cleaning
Reads mapping

Transcriptomics

Reads quality check and cleaning
Reads/ESTs mapping
Reads count and RPKM calculation
DEGs calling
SAM file parsing and manipulation

...available soon

Assembly & clustering
Genome masking
Preliminary genome gene structure annotation
Preliminary genome repeats annotation
Variant calling

<https://bioinfo.szn.it>

<https://ticketing.bioinfo.szn.it>

bioinforma@szn.it



THANK YOU
500000°
MUCH!