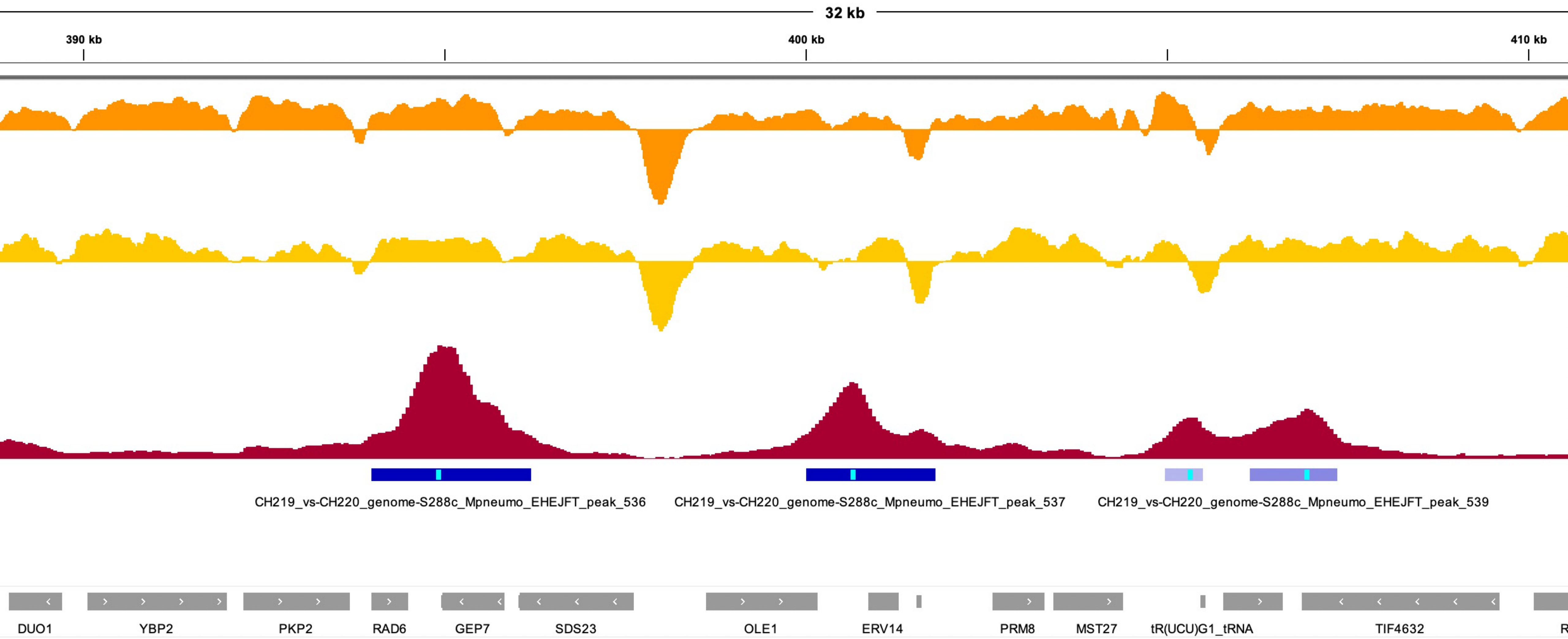
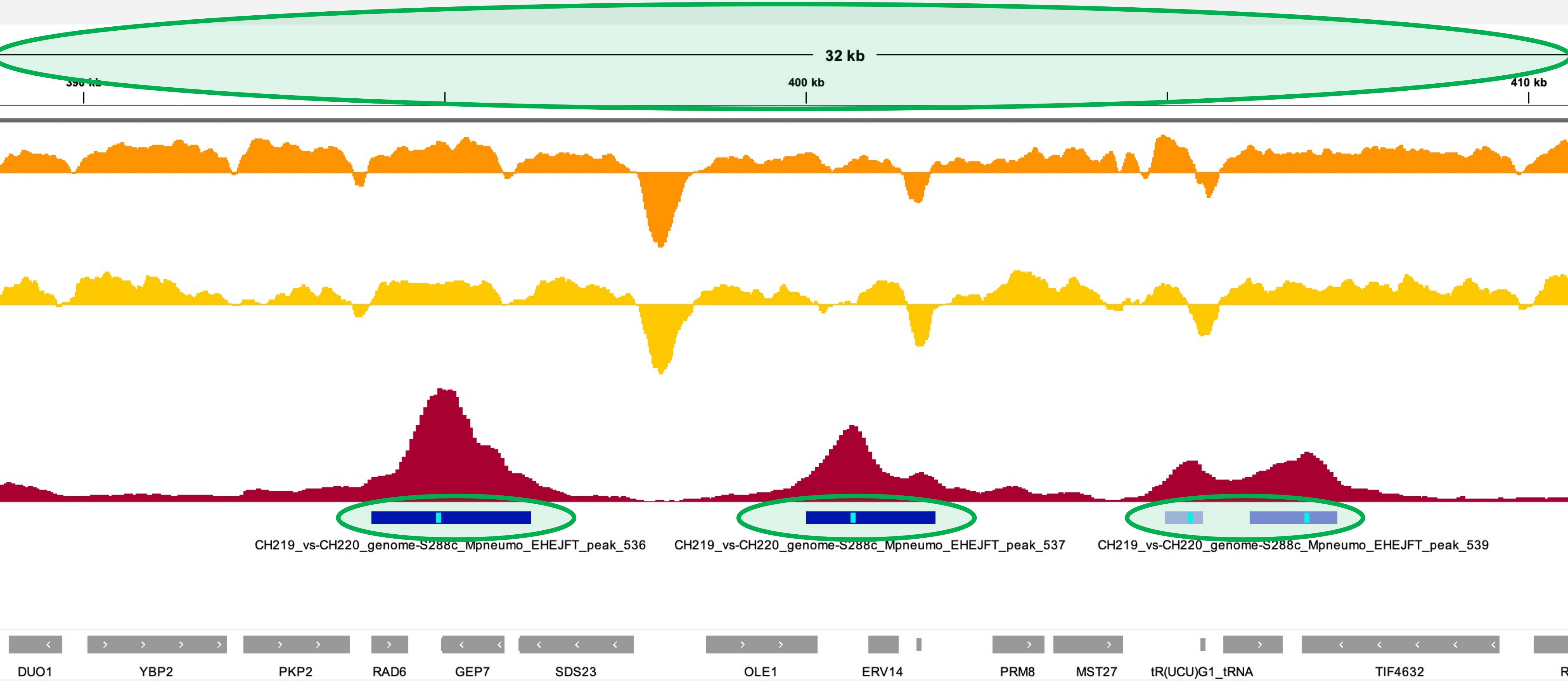
R 201

Epigenomics Data Analysis
Jacques Serizay
Physalia 2023

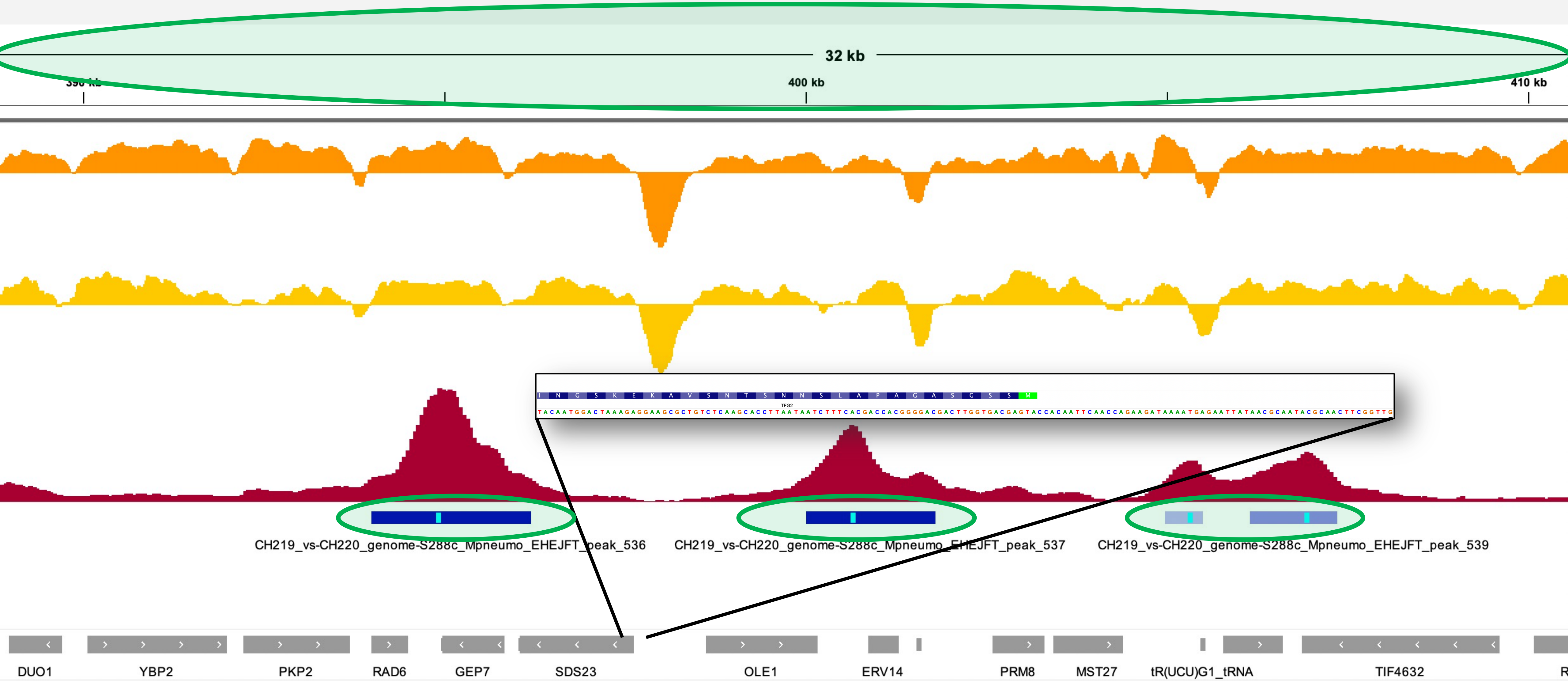
Epigenomics in a browser



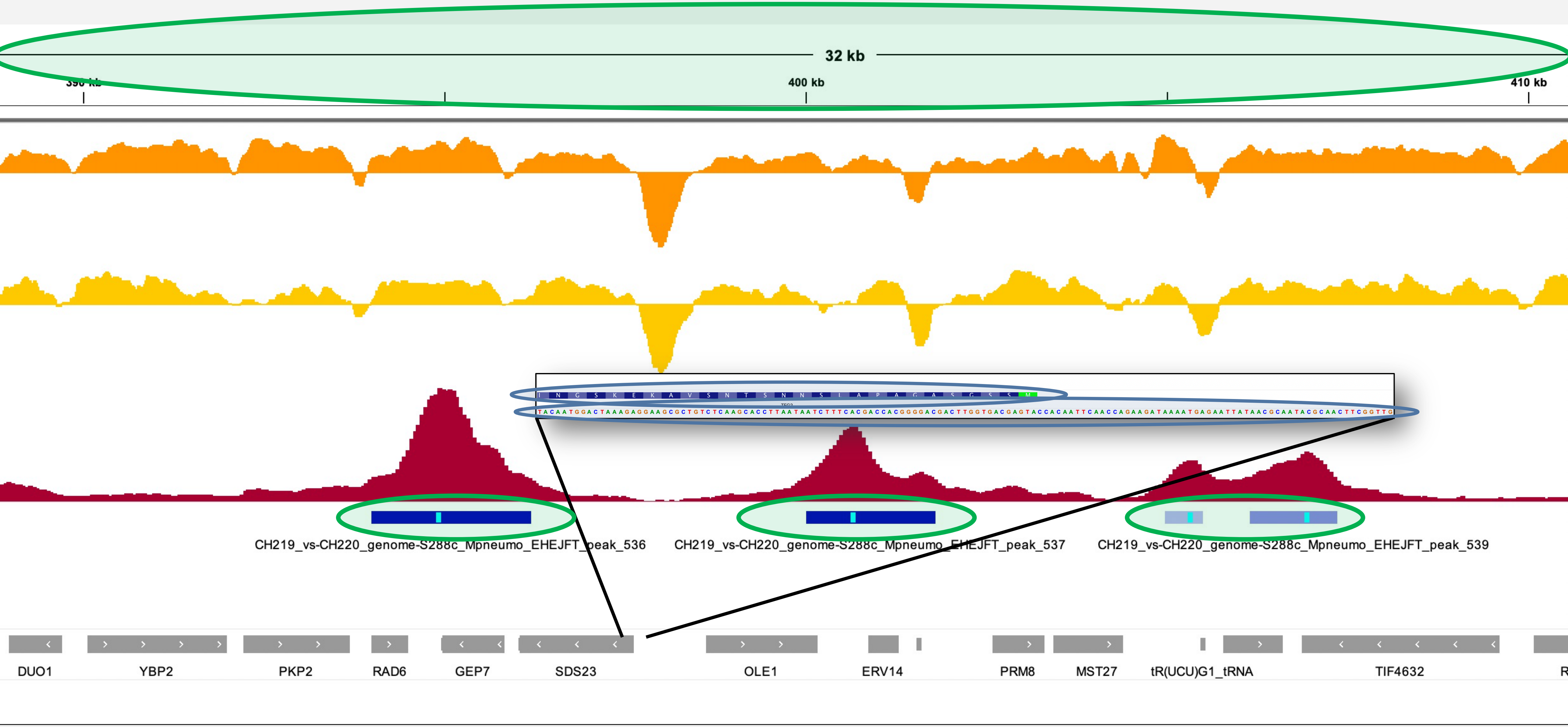
Epigenomics in a browser



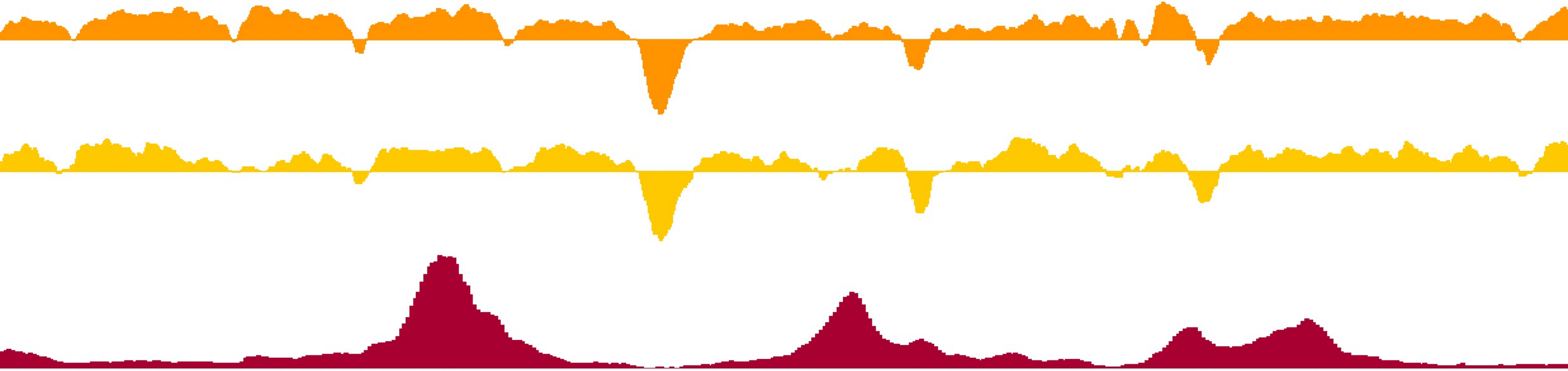
Epigenomics in a browser



Epigenomics in a browser



Epigenomics in a browser



Epigenomics in a browser

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format

```
I          2          5          0.153096
I          5          7          0.459288
I          7          9          0.612384
I          9          11         0.76548
I          11         15         0.918576
I          15         16         1.07167
I          16         17         1.37786
I          17         30         1.68406
```

Epigenomics in a browser

- Genomic tracks are generally stored as bigwig files.

- bigwig files store long numerical vectors in a binarized format

```
I          2          5          0.153096
I          5          7          0.459288
I          7          9          0.612384
I          9          11         0.76548
I          11         15         0.918576
I          15         16         1.07167
I          16         17         1.37786
I          17         30         1.68406
```

```
�&XHA'\0�eJ]R�@6���9�0|]<pc<�d?�f?�,@@,-@~_@K�`���?�V�@53d�ej�A��i�B���IJ�II�h
                                                                               III��IV`VjVIQVII|�VIIIIP�X     a
                                                                                                                   XI
,�
XII
�   sXIII
       ��XIV
  XVXVIbwMitoOx�=�    �U�������{��so�J�2KD��V2U�d(���2RQ
���BÙJ�
�4(2
����'��[i����JM���~X�w��9"��Y���Jp<�����4ں|YQ��4NL[e��X_���[ub�S8O'����i��S|$mÓ3|"�m�O�mM����N�i�ū�$�,k�!�:X�Y��NO�[&xn��\���i[Y94�8O�nH�F����
[f��ﮯi�bڜg2�]�'8;mm���s|��H�ۏF&�f�,�Jp����N��s�$�n�n��{��9��=�ڮ����~s�6�'�����4�ω    .e�&%��������9�J��w1�]����ᄉ-��zl뙚w�f�:3���N����`gsḺKc�-
��f���c�-Z4;�����I��|ō̄%��������`/s<���/���m���c?s<�}��=>x�9�i=��8x�9���|k���/x�9^l���-
3{n�W��8�l_b�ɔ�P�78&��8�1�����9m����ڹ�o%82mc�|;�[X���M�Vse���6���  86$?������{�ۼm����]�_
```

Epigenomics in a browser

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package

```
> library(rtracklayer)

> import('….bw')

GRanges object with 6243328 ranges and 1 metadata column:
            seqnames    ranges strand |     score
               <Rle> <IRanges>  <Rle> | <numeric>
        [1]        I       3-5      * |  0.153096
        [2]        I       6-7      * |  0.459288
        [3]        I       8-9      * |  0.612384
        [4]        I     10-11      * |  0.765480
        [5]        I     12-15      * |  0.918576
        ...      ...       ...    ... .       ...
  [6243324]     Mito     85775      * |   9.79815
  [6243325]     Mito     85776      * |   8.26719
  [6243326]     Mito     85777      * |   6.43003
  [6243327]     Mito     85778      * |   5.66455
  [6243328]     Mito     85779      * |   3.21502
  -------
  seqinfo: 17 sequences from an unspecified genome
```

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `import()` from the `rtracklayer` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

```
b b b b k k e e f a a a a a a a a g g g
```

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

b b b b k k e e f a a a a a a a a g g g

4 b

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `import()` from the `rtracklayer` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

b b b b **k k** e e f a a a a a a a a g g g

4 b

2 k

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

b b b b k k **e e** f a a a a a a a a g g g

4 b
2 k
2 e

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

```
b b b b k k e e f a a a a a a a a g g g
```

```
4 b
2 k
2 e
1 f
```

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

```
b b b b k k e e f a a a a a a a a g g g
```

```
4 b
2 k
2 e
1 f
8 a
```

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

b b b b k k e e f a a a a a a a a **g g g**

4	**b**
2	**k**
2	**e**
1	**f**
8	**a**
3	**g**

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

b b b b k k e e f a a a a a a a a **g g g**

Run-values: b k e f a g

Run-lengths: 4 2 2 1 8 3

Run-length encoding vectors

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

b b b b k k e e f a a a a a a a a **g g g**

Run-values: b k e f a g
Run-lengths: 4 2 2 1 8 3

12 alpha-numeric values instead of 20 alphabetic values

Epigenomics in a browser

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

```
> library(rtracklayer)

> import('….bw')

GRanges object with 6243328 ranges and 1 metadata column:
            seqnames    ranges strand |     score
               <Rle> <IRanges>  <Rle> | <numeric>
        [1]        I       3-5      * |  0.153096
        [2]        I       6-7      * |  0.459288
        [3]        I       8-9      * |  0.612384
        [4]        I     10-11      * |  0.765480
        [5]        I     12-15      * |  0.918576
        ...      ...       ...    ... .       ...
  [6243324]     Mito     85775      * |   9.79815
  [6243325]     Mito     85776      * |   8.26719
  [6243326]     Mito     85777      * |   6.43003
  [6243327]     Mito     85778      * |   5.66455
  [6243328]     Mito     85779      * |   3.21502
  -------
  seqinfo: 17 sequences from an unspecified genome
```

Epigenomics in a browser

- Genomic tracks are generally stored as bigwig files.
- bigwig files store long numerical vectors in a binarized format.
- In R, bigwig files can be imported with `` `import()` `` from the `` `rtracklayer` `` package.
- bigwig files can be imported as **numerical vectors**, stored as **Run-length encoding vectors**.

```
> import('….bw', as = 'RLE')

RleList of length 17
$I
numeric-Rle of length 230218 with 104639 runs
  Lengths:         2         3         2         2         2 ...         2        30         1      1084
  Values :  0.000000  0.153096  0.459288  0.612384  0.765480 ...  0.612384  0.459288  0.306192  0.000000

$II
numeric-Rle of length 813184 with 424729 runs
  Lengths:         2         1         1         2         1 ...          6         1         5         2
  Values :  0.153096  0.306192  0.612384  0.918576  1.071670 ...   0.459288  0.306192  0.153096  0.000000

...
<15 more elements>
```

SummarizedExperiment

An NGS analysis workflow typically involves:

1. Defining features of interest (e.g. gene annotations for RNA-seq, or accessibility peaks for ATAC-seq)
2. Counting reads overlapping with each feature
3. Performing differential analysis
4. Extracting results

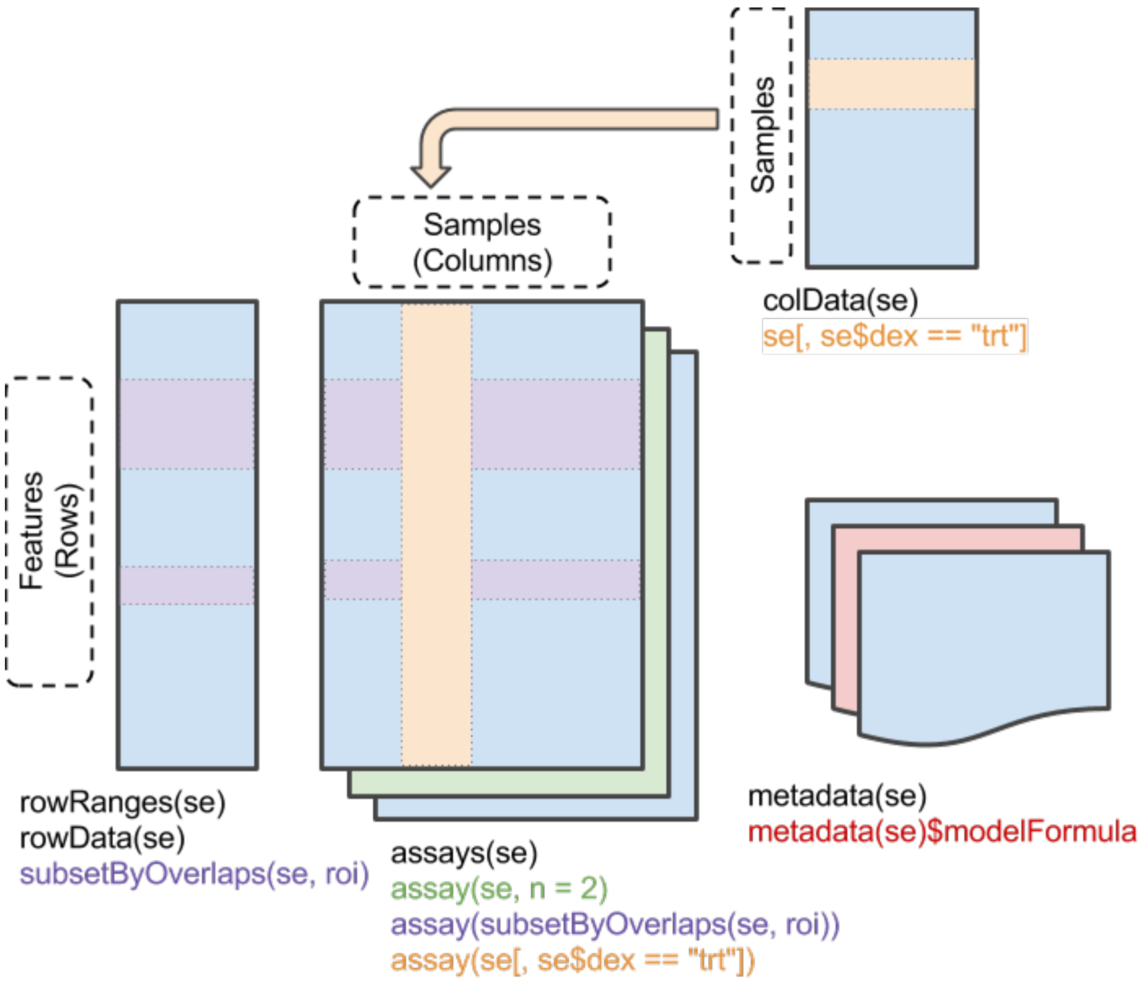
SummarizedExperiment

Published: 29 January 2015

Orchestrating high-throughput genomic analysis with Bioconductor

Wolfgang Huber, Vincent J Carey, Robert Gentleman, Simon Anders, Marc Carlson, Benilton S Carvalho, Hector Corrada Bravo, Sean Davis, Laurent Gatto, Thomas Girke, Raphael Gottardo, Florian Hahne, Kasper D Hansen, Rafael A Irizarry, Michael Lawrence, Michael I Love, James MacDonald, Valerie Obenchain, Andrzej K Oleś, Hervé Pagès, Alejandro Reyes, Paul Shannon, Gordon K Smyth, Dan Tenenbaum, Levi Waldron & Martin Morgan -Show fewer authors

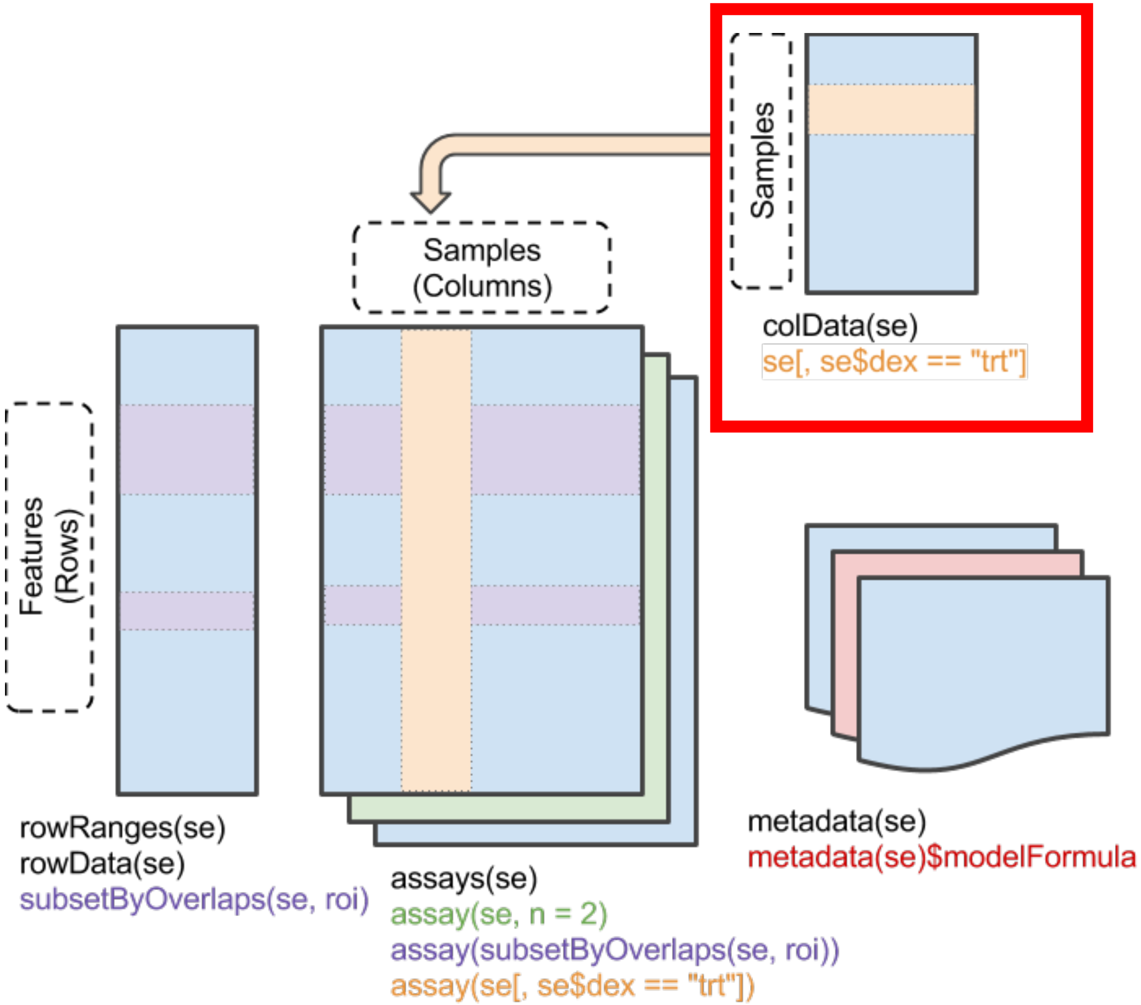
Nature Methods **12**, 115–121(2015) | Cite this article



Huber et al., Nat. Meth. 2015

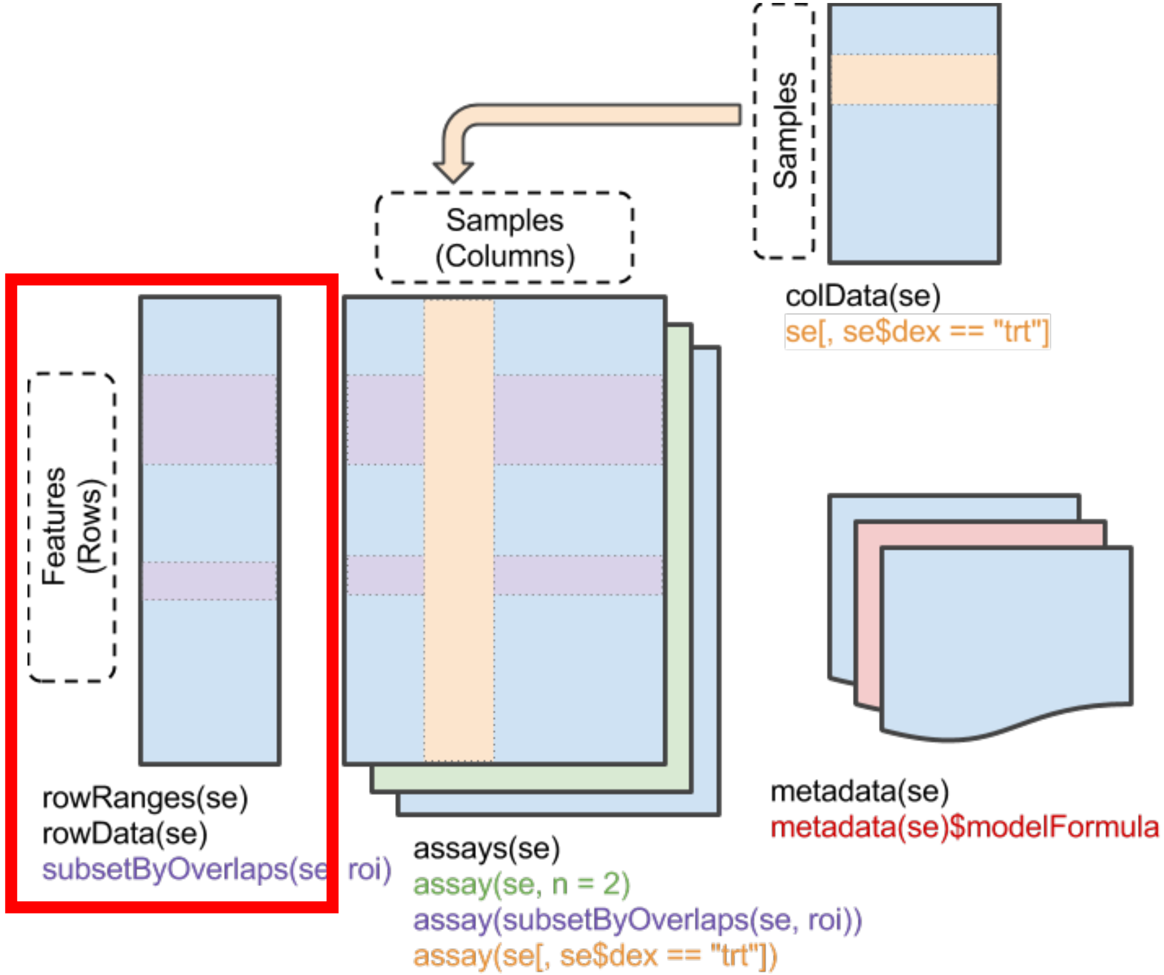
SummarizedExperiment

- colData(): Annotations on each column, as a DataFrame.
    - E.g., description of each sample



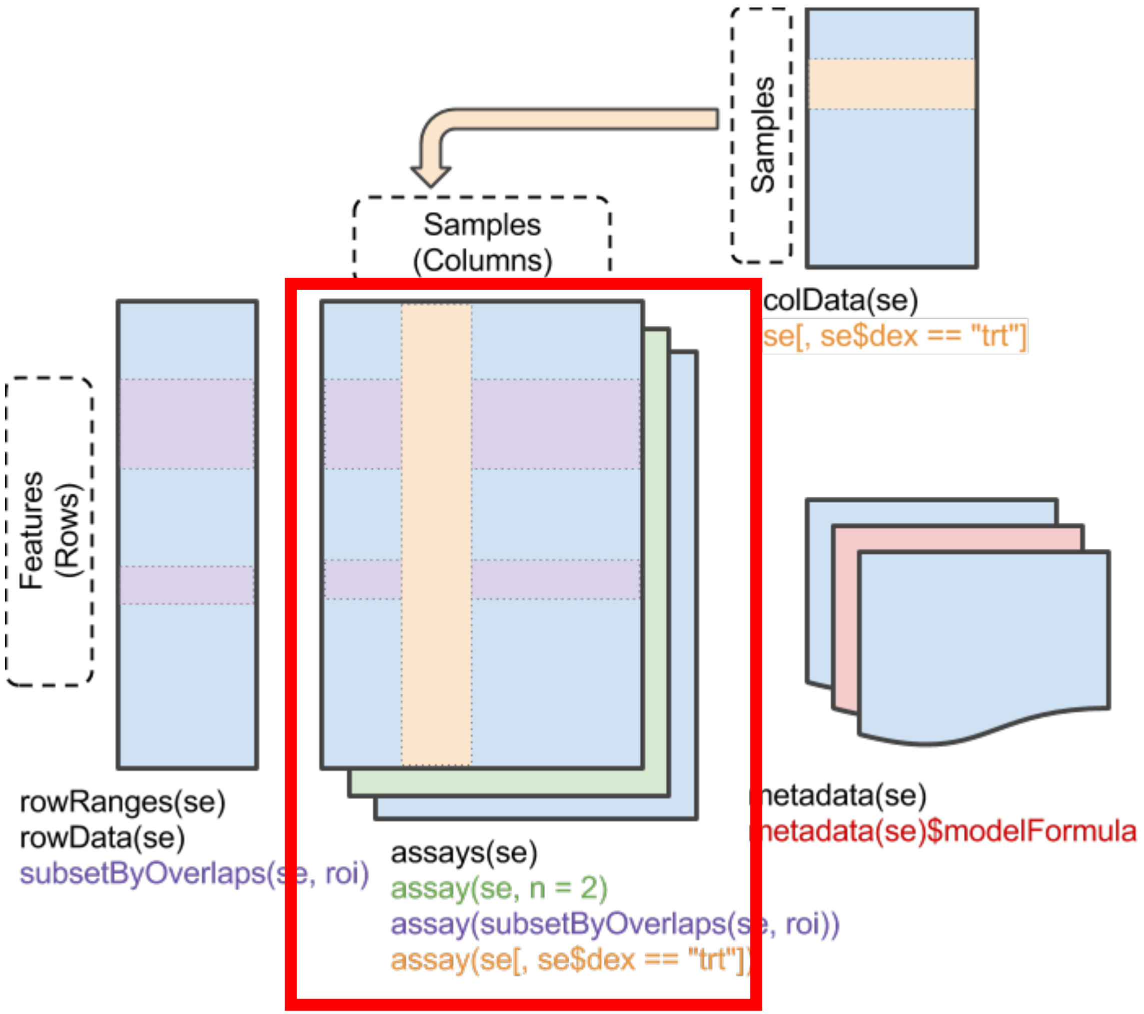
SummarizedExperiment

- colData(): Annotations on each column, as a DataFrame.
    - E.g., description of each sample
- rowData/rowRanges(): Annotations on each row.
    - E.g., coordinates of gene / exons /peaks in transcripts / etc.



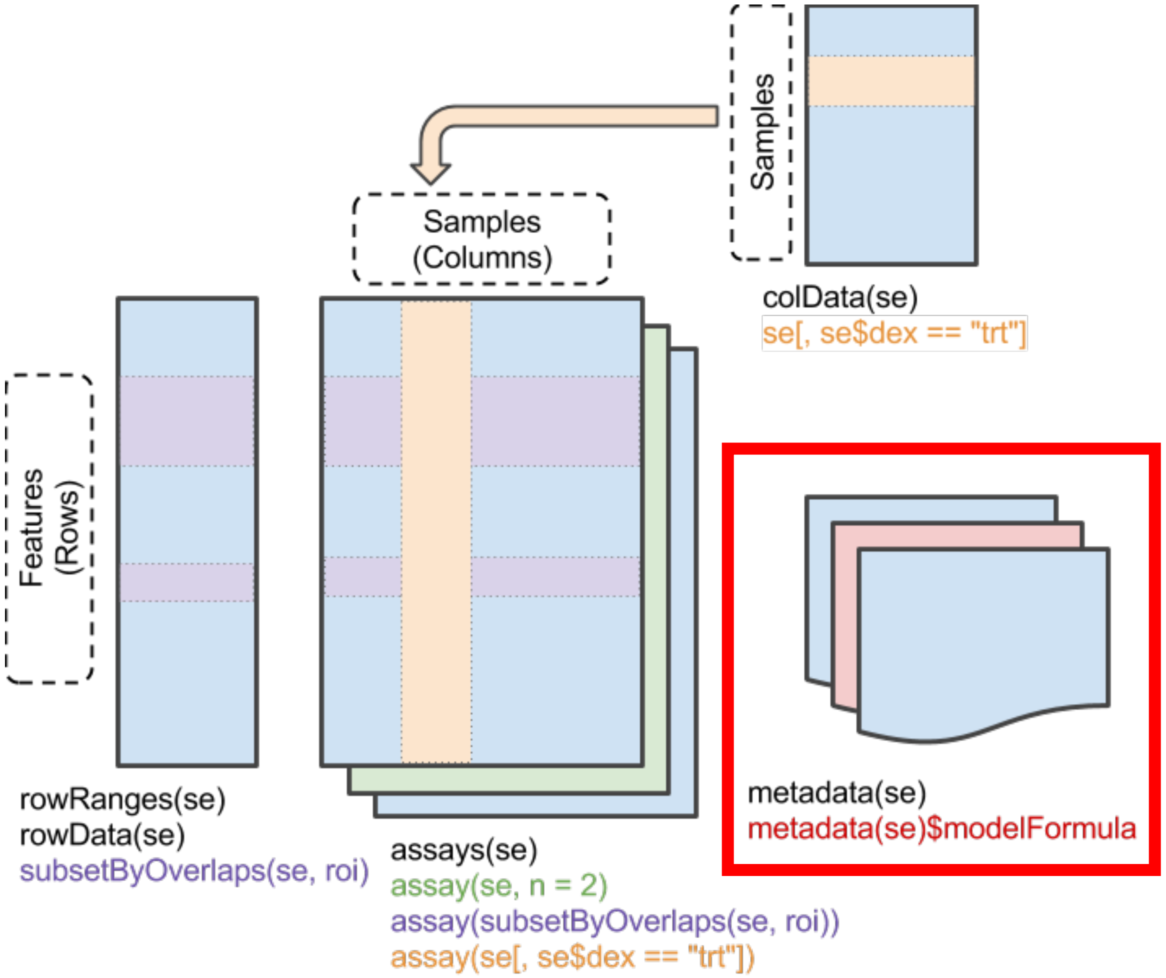
SummarizedExperiment

- colData(): Annotations on each column, as a DataFrame.
  - E.g., description of each sample
- rowData/rowRanges(): Annotations on each row.
  - E.g., coordinates of gene / exons /peaks in transcripts / etc.
- assay(), assays(): A matrix-like or list of matrix-like objects of identical dimension
  - rows: refer to **rowRanges**: genes, genomic coordinates, etc.
  - columns: refer to **colData**: samples, cells, etc.
  - Implements dim(), dimnames() and 2-dimensional [,]
  - **Can be several assays!!!**



SummarizedExperiment

- colData(): Annotations on each column, as a DataFrame.
    - E.g., description of each sample
- rowData/rowRanges(): Annotations on each row.
    - E.g., coordinates of gene / exons /peaks in transcripts / etc.
- assay(), assays(): A matrix-like or list of matrix-like objects of identical dimension
    - rows: refer to **rowRanges**: genes, genomic coordinates, etc.
    - columns: refer to **colData**: samples, cells, etc.
    - Implements dim(), dimnames() and 2-dimensional [,]
    - **Can be several assays!!!**
- metadata(): List of unstructured metadata describing the overall content of the object.



SummarizedExperiment

SummarizedExperiment is a key data structure in Bioconductor, reused by or modified in:

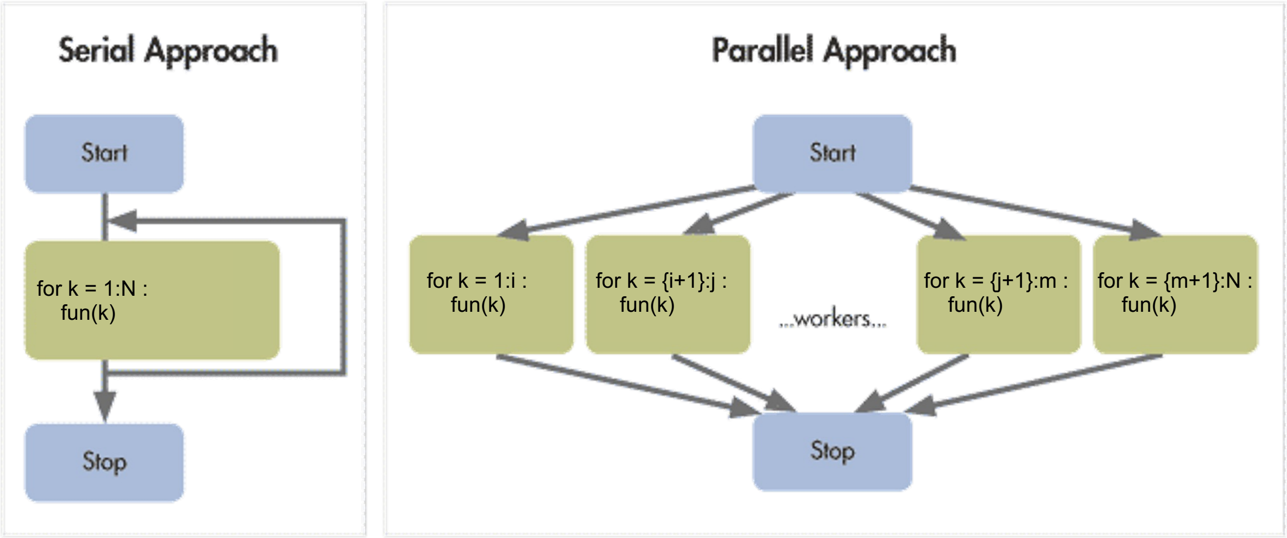
- clusterExperiment
- SingleCellExperiment
- LoomExperiment
- MsExperiment
- TreeSummarizedExperiment
- SpatialExperiment
- And 100s more…

Depends On Me	AffiXcan, airway, alabaster.se, AllelicImbalance, ASpediaFI, atena, bambu, BDMMAcorrect, benchmarkfdrData2019, BiocSklearn, BioPlex, BiSeq, bnbc, bodymapRat, bsseq, CAGEfightR, celaref, celldex, clusterExperiment, compartmap, CoreGx, coseq, csaw, CSSQ, curatedAdipoChIP, curatedAdipoRNA, curatedMetagenomicData, DaMiRseq, deco, deepSNV, DeMixT, DESeq2, DEXSeq, DiffBind, diffcoexp, diffHic, divergence, DMCFB, DMCHMM, ENmix, EnrichmentBrowser, epigenomix, evaluomeR, EventPointer, exomePeak2, ExperimentSubset, ExpressionAtlas, extraChIPs, FEAST, fission, FRASER, GenomicAlignments, GenomicFiles, GenomicSuperSignature, GRmetrics, GSEABenchmarkeR, HDCytoData, HelloRanges, hermes, HiCDOC, HighlyReplicatedRNASeq, hipathia, HMP16SData, InTAD, InteractionSet, IntEREst, iSEE, iSEEhex, iSEEhub, ISLET, isomiRs, ivygapSE, lefser, lipidr, LoomExperiment, Macarron, made4, MatrixQCvis, MBASED, MetaGxOvarian, MetaGxPancreas, methrix, methylPipe, MethylSeqData, MetNet, mia, miaViz, MicrobiomeBenchmarkData, microbiomeDataSets, microRNAome, minfi, miRmine, moanin, MouseGastrulationData, MouseThymusAgeing, mpra, MultiAssayExperiment, NADfinder, NBAMSeq, NewWave, ObMiTi, OUTRIDER, padma, parathyroidSE, PDATK, phenomis, PhIPData, profileplyr, qmtools, qsvaR, recount, recount3, RegEnrich, REMP, restfulSE, restfulSEData, ROCpAI, rqt, runibic, sampleClassifierData, Scale4C, scAnnotatR, scGPS, scMultiome, scone, screenCounter, scTreeViz, SDAMS, SeqGate, SEtools, SGSeq, signatureSearch, SingleCellExperiment, singleCellTK, SingleR, soGGi, spatialDmelxsim, spqn, spqnData, sRACIPE, ssPATHS, stageR, SummarizedBenchmark, survtype, TENxIO, tidySummarizedExperiment, timecoursedata, TissueEnrich, TNBC.CMS, TREG, tuberculosis, UMI4Cats, VanillaICE, VariantAnnotation, VariantExperiment, velociraptor, weitrix, yamss, zinbwave

Imports Me	ADAM, ADImpute, aggregateBioVar, airpart, alabaster.spatial, ALDEx2, alpine, ANCOMBC, animalcules, anota2seq, APAlyzer, apeglm, APL, appreci8R, ASICS, ASURAT, ATACseqTFEA, AUCell, autonomics, awst, barcodetrackR, BASiCS, BASiCStan, batchelor, BayesSpace, bayNorm, BBCAnalyzer, beer, benchdamic, bigPint, BiocOncoTK, BioNERO, biosigner, biotmle, biovizBase, biscuiteer, BiSeq, blacksheepr, BloodCancerMultiOmics2017, BloodGen3Module, brgedata, BRGenomics, BUMHMM, BUScorrect, BUSseq, CAEN, CAGEr, CATALYST, CBEA, cBioPortalData, ccfindR, celda, CelliD, CellMixS, CellTrails, censcyt, Cepo, CeTF, CHETAH, ChIPpeakAnno, ChromSCape, chromVAR, CiteFuse, CLLmethylation, clustifyr, cmapR, CNVfilteR, CNVRanger, CoGAPS, comapr, combi, condiments, consensusDE, consICA, CopyNumberPlots, corral, COSMIC.67, countsimQC, CTSV, curatedTCGAData, cydar, CyTOFpower, cytofQC, cytoKernel, cytomapper, cytoviewer, DAMEfinder, dasper, debCAM, debrowser, decompTumor2Sig, DEFormats, DEGreport, DELocal, deltaCaptureC, DEP, DEScan2, DESpace, destiny, DEWSeq, diffcyt, DifferentialRegulation, diffUTR, Dino, DiscoRhythm, distinct, dittoSeq, DMRcate, DominoEffect, doppelgangR, doseR, DropletUtils, Dune, easierData, easyRNASeq, eisaR, ELMER, emtdata, ensemblVEP, epialleleR, epigraHMM, EpiMix, epimutacions, epistack, epivizrData, erma, escape, escheR, EWCE, ExpHunterSuite, FCBF, fcScan, FeatSeekR, FieldEffectCrc, FindIT2, fishpond, FLAMES, FlowSorted.Blood.EPIC, FlowSorted.CordBloodCombined.450k, fluentGenomics, FuseSOM, GARS, gCrisprTools, gemma.R, GeneTonic, genomicInstability, GeoTcgaData, getDEE2, ggbio, ggspavis, Glimma, glmGamPoi, glmSparseNet, GNET2, GRaNIE, GreyListChIP, gscreend, GSE13015, GSVA, gwasurvivr, GWENA, HiContacts, HMP2Data, HTSeqGenie, HumanTranscriptomeCompendium, hummingbird, iasva, icetea, ideal, IFAA, IgGeneUsage, IHWpaper, ILoReg, imcRtools, infercnv, INSPEcT, InterMineR, IntOMICS, iSEEu, IsoformSwitchAnalyzeR, LACE, LineagePulse, lineagespot, lionessR, lisaClust, MADSEQ, MAI, mariner, marr, MAST, mastR, mbkmeans, MBQN, mCSEA, MEAL, MEAT, MEB, MerfishData, MetaboAnnotation, metabolomicsWorkbenchR, MetaGxBreast, MetaNeighbor, MetaScope, metaseqR2, MethReg, MethylAid, methylscaper, methylumi, miaSim, MicrobiotaProcess, midasHLA, miloR, MinimumDistance, miRLAB, miRSM, missMethyl, MLInterfaces, MLSeq, monaLisa, MoonlightR, motifbreakR, motifmatchr, MPRAnalyze, MsExperiment, MsFeatures, msgbsR, MSPrep, msqrob2, MuData, MultiDataSet, multiOmicsViz, mumosa, muscat, musicatk, MWASTools, NanoMethViz, Nebulosa, NetActivity, netSmooth, nnSVG, NoRCE, NormalyzerDE, NxtIRFcore, oligoClasses, omicRexposome, OmicsLonDA, omicsPrint, omicsViewer, oncomix, ontoProc, ORFik, OVESEG, PAIRADISE, pairedGSEA, pairkat, pcaExplorer, peco, PharmacoGx, phemd, phenopath, PhosR, pipeComp, planttfhunter, pmp, POMA, POWSC, proActiv, proDA, psichomics, pulsedSilac, PureCN, QFeatures, qsmooth, quantiseqr, R453Plus1Toolbox, RadioGx, RaggedExperiment, RareVariantVis, RcisTarget, receptLoss, regionReport, regsplice, rgsepd, rifi, rifiComparative, Rmmquant, RNAAgeCalc, RNAsense, RnaSeqSampleSize, roar, RolDE, ropls, rScudo, RTCGAToolbox, RTN, satuRn, SBGNview, SC3, SCArray, SCArray.sat, scater, scBFA, scCB2, scDblFinder, scDD, scDDboost, scds, scFeatures, scHOT, scmap, scMerge, scMET, scmeth, SCnorm, scoreInvHap, scp, scPipe, scran, scReClassify, scRepertoire, scRNAseq, scruff, scry, scTensor, scTGIF, scuttle, scviR, sechm, segmenter, seqCAT, sesame, SGCP, sigFeature, signifinder, SigsPack, SimBu, simpleSeg, SingleCellMultiModal, singscore, SingscoreAMLMutations, slalom, slingshot, snapcount, SNPhood, Spaniel, spaSim, SpatialCPie, spatialDE, SpatialExperiment, SpatialFeatureExperiment, spatialHeatmap, spatialLIBD, spatzie, SPIAT, spicyR, splatter, SpliceWiz, SplicingFactory, SpotClean, srnadiff, sSNAPPY, standR, stJoincount, struct, StructuralVariantAnnotation, supersigs, SVMDO, switchde, systemPipeR, systemPipeTools, TabulaMurisSenisData, TBSignatureProfiler, TCGAbiolinks, TCGAutils, TCGAWorkflow, TCGAWorkflowData, TCseq, TEKRABber, tenXplore, tidybulk, tidySingleCellExperiment, TOAST, tomoda, ToxicoGx, tradeSeq, TrajectoryUtils, transformGamPoi, traviz, TreeSummarizedExperiment, Trendy, tricycle, TSCAN, tscR, TTMap, TVTB, tximeta, UCell, VAExprs, VariantFiltering, VDJdive, vidger, Voyager, wpm, xcms, zellkonverter, zFPKM

BiocParallel: executing code in parallel

- Genomic analyses require heavy resources.
- Generally, this benefits from parallelization.



BiocParallel: executing code in parallel

- Genomic analyses require heavy resources.
- Generally, this benefits from parallelization.
- BiocParallel is a Bioconductor package designed to :
  - Reduce the complexity to parallelize tasks, a problem faced by the end-user.
  - Expose parallelization options in functions to the end-users, a problem faced by the package developer.

BiocParallel: executing code in parallel

- Genomic analyses require heavy resources.
- Generally, this benefits from parallelization.
- BiocParallel is a Bioconductor package designed to :
    - Reduce the complexity to parallelize tasks, a problem faced by the end-user.
    - Expose parallelization options in functions to the end-users, a problem faced by the package developer.

BiocParallel aims to provide a unified interface to existing parallel infrastructures.

BiocParallel: executing code in parallel

To enable parallelization in Bioconductor, one has to **register** a parallelization strategy:

```
> library(BiocParallel)

> BiocParallel::registered()

$MulticoreParam
class: MulticoreParam
  bpisup: FALSE; bpnworkers: 8; bptasks: 0; bpjobname: BPJOB
  bplog: FALSE; bpthreshold: INFO; bpstopOnError: TRUE
  bpRNGseed: ; bptimeout: NA; bpprogressbar: FALSE
  bpexportglobals: TRUE; bpexportvariables: FALSE; bpforceGC: FALSE
  bpfallback: TRUE
  bplogdir: NA
  bpresultdir: NA
  cluster type: FORK

$SerialParam
class: SerialParam
  bpisup: FALSE; bpnworkers: 1; bptasks: 0; bpjobname: BPJOB
  bplog: FALSE; bpthreshold: INFO; bpstopOnError: TRUE
  bpRNGseed: ; bptimeout: NA; bpprogressbar: FALSE
  bpexportglobals: FALSE; bpexportvariables: FALSE; bpforceGC: FALSE
  bpfallback: FALSE
  bplogdir: NA
  bpresultdir: NA
```

BiocParallel: executing code in parallel

To enable parallelization in Bioconductor, one has to **register** a parallelization strategy:

```
> register(MulticoreParam(workers = 4, progressbar = TRUE), default = TRUE)

> bpparam()

$MulticoreParam
class: MulticoreParam
  bpisup: FALSE; bpnworkers: 4; bptasks: 2147483647; bpjobname: BPJOB
  bplog: FALSE; bpthreshold: INFO; bpstopOnError: TRUE
  bpRNGseed: ; bptimeout: NA; bpprogressbar: TRUE
  bpexportglobals: TRUE; bpexportvariables: FALSE; bpforceGC: FALSE
  bpfallback: TRUE
  bplogdir: NA
  bpresultdir: NA
  cluster type: FORK
```

BiocParallel: executing code in parallel

To **execute** with parallelization in Bioconductor, one has to use the BPPARAM argument (when supported):

```
> t0 <- Sys.time()
> lapply(1:3, \(x) {
    print(x)
    Sys.sleep(1)
    Sys.time() - t0
})

[1] 1
Time difference of 1.005459 secs
[1] 2
Time difference of 2.01272 secs
[1] 3
Time difference of 3.019525 secs
```

BiocParallel: executing code in parallel

To **execute** with parallelization in Bioconductor, one has to use the BPPARAM argument (when supported):

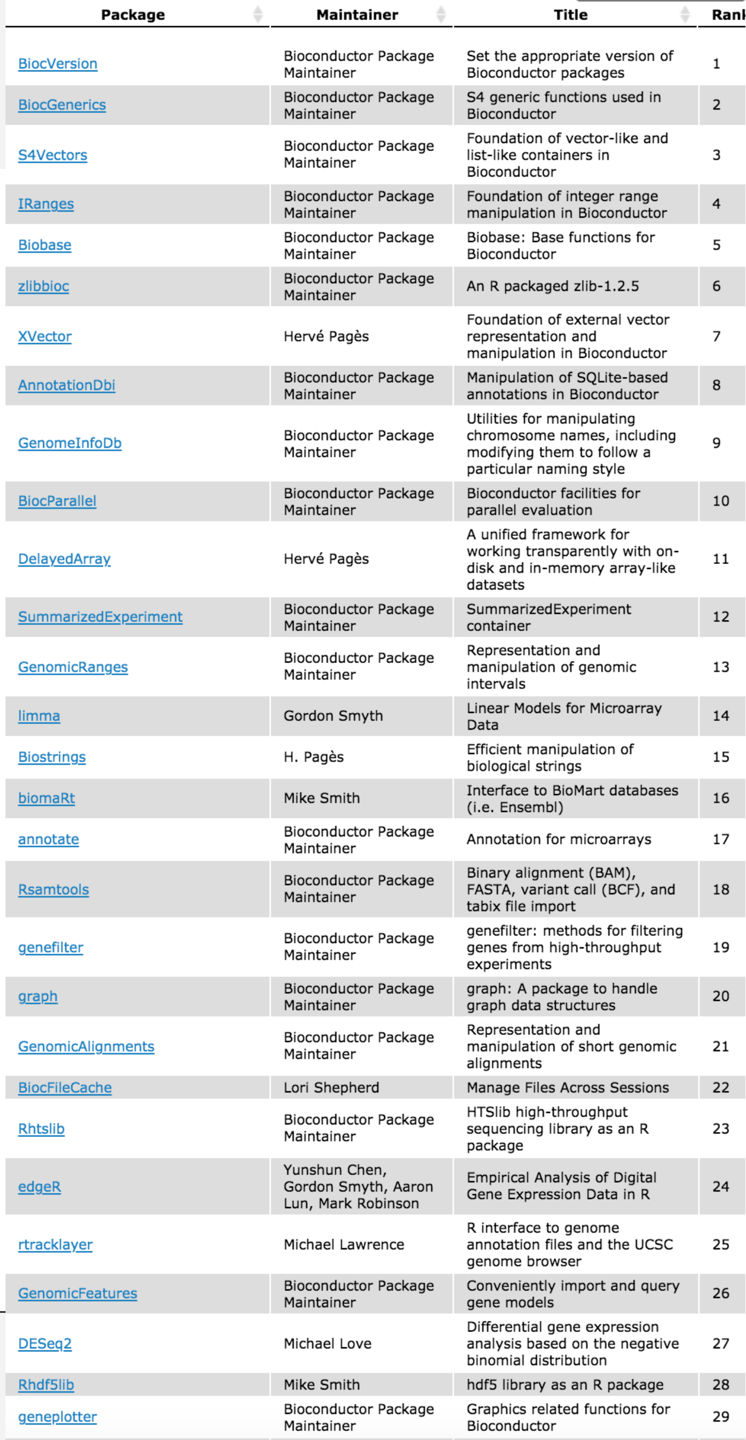
```
> t0 <- Sys.time()
> lapply(1:3, \(x) {
    print(x)
    Sys.sleep(1)
    Sys.time() - t0
})

[1] 1
Time difference of 1.005459 secs
[1] 2
Time difference of 2.01272 secs
[1] 3
Time difference of 3.019525 secs

> t0 <- Sys.time()
> bplapply(BPPARAM = bpparam(), 1:3, \(x) {
    print(x)
    Sys.sleep(1)
    Sys.time() - t0
})

[1] 1
Time difference of 1.015662 secs
[2] 2
Time difference of 1.018448 secs
[3] 3
Time difference of 1.021177 secs
```

Additional packages in Bioconductor

Rsamtools: interacting with BAM files

GenomicAlignements: counting BAM files over Granges

Many others….



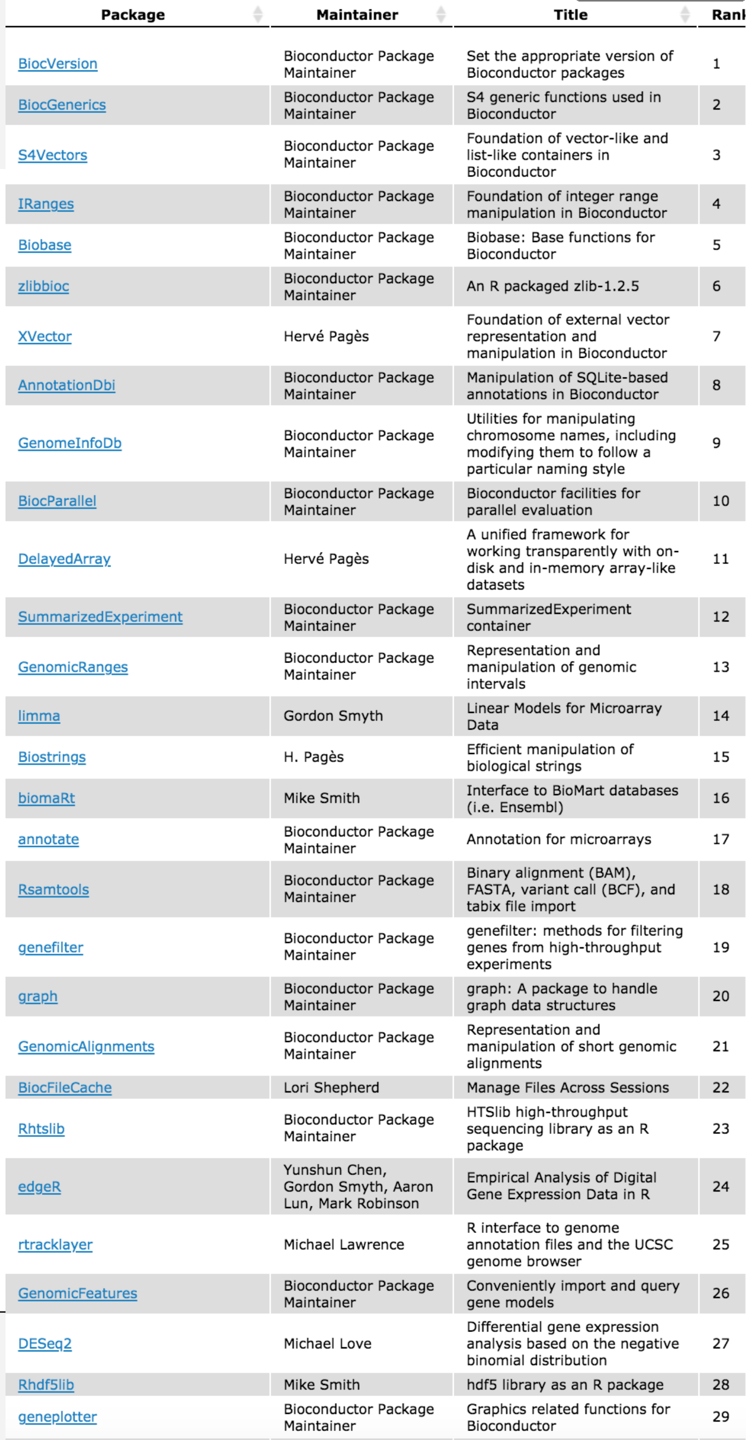
	Package	Maintainer	Title	Ran
	BiocVersion	Bioconductor Package Maintainer	Set the appropriate version of Bioconductor packages	1
	BiocGenerics	Bioconductor Package Maintainer	S4 generic functions used in Bioconductor	2
*	S4Vectors	Bioconductor Package Maintainer	Foundation of vector-like and list-like containers in Bioconductor	3
*	IRanges	Bioconductor Package Maintainer	Foundation of integer range manipulation in Bioconductor	4
	Biobase	Bioconductor Package Maintainer	Biobase: Base functions for Bioconductor	5
	zlibbioc	Bioconductor Package Maintainer	An R packaged zlib-1.2.5	6
	XVector	Hervé Pagès	Foundation of external vector representation and manipulation in Bioconductor	7
	AnnotationDbi	Bioconductor Package Maintainer	Manipulation of SQLite-based annotations in Bioconductor	8
*	GenomeInfoDb	Bioconductor Package Maintainer	Utilities for manipulating chromosome names, including modifying them to follow a particular naming style	9
*	BiocParallel	Bioconductor Package Maintainer	Bioconductor facilities for parallel evaluation	10
	DelayedArray	Hervé Pagès	A unified framework for working transparently with on-disk and in-memory array-like datasets	11
*	SummarizedExperiment	Bioconductor Package Maintainer	SummarizedExperiment container	12
*	GenomicRanges	Bioconductor Package Maintainer	Representation and manipulation of genomic intervals	13
	limma	Gordon Smyth	Linear Models for Microarray Data	14
*	Biostrings	H. Pagès	Efficient manipulation of biological strings	15
	biomaRt	Mike Smith	Interface to BioMart databases (i.e. Ensembl)	16
	annotate	Bioconductor Package Maintainer	Annotation for microarrays	17
*	Rsamtools	Bioconductor Package Maintainer	Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import	18
	genefilter	Bioconductor Package Maintainer	genefilter: methods for filtering genes from high-throughput experiments	19
	graph	Bioconductor Package Maintainer	graph: A package to handle graph data structures	20
*	GenomicAlignments	Bioconductor Package Maintainer	Representation and manipulation of short genomic alignments	21
	BiocFileCache	Lori Shepherd	Manage Files Across Sessions	22
	Rhtslib	Bioconductor Package Maintainer	HTSlib high-throughput sequencing library as an R package	23
	edgeR	Yunshun Chen, Gordon Smyth, Aaron Lun, Mark Robinson	Empirical Analysis of Digital Gene Expression Data in R	24
*	rtracklayer	Michael Lawrence	R interface to genome annotation files and the UCSC genome browser	25
*	GenomicFeatures	Bioconductor Package Maintainer	Conveniently import and query gene models	26
*	DESeq2	Michael Love	Differential gene expression analysis based on the negative binomial distribution	27
	Rhdf5lib	Mike Smith	hdf5 library as an R package	28
	geneplotter	Bioconductor Package Maintainer	Graphics related functions for Bioconductor	29

Additional packages in Bioconductor

Rsamtools: interacting with BAM files

GenomicAlignements: counting BAM files over Granges

Many others….



	Package	Maintainer	Title	Ranl
	BiocVersion	Bioconductor Package Maintainer	Set the appropriate version of Bioconductor packages	1
	BiocGenerics	Bioconductor Package Maintainer	S4 generic functions used in Bioconductor	2
*	S4Vectors	Bioconductor Package Maintainer	Foundation of vector-like and list-like containers in Bioconductor	3
*	IRanges	Bioconductor Package Maintainer	Foundation of integer range manipulation in Bioconductor	4
	Biobase	Bioconductor Package Maintainer	Biobase: Base functions for Bioconductor	5
	zlibbioc	Bioconductor Package Maintainer	An R packaged zlib-1.2.5	6
	XVector	Hervé Pagès	Foundation of external vector representation and manipulation in Bioconductor	7
*	AnnotationDbi	Bioconductor Package Maintainer	Manipulation of SQLite-based annotations in Bioconductor	8
*	GenomeInfoDb	Bioconductor Package Maintainer	Utilities for manipulating chromosome names, including modifying them to follow a particular naming style	9
*	BiocParallel	Bioconductor Package Maintainer	Bioconductor facilities for parallel evaluation	10
	DelayedArray	Hervé Pagès	A unified framework for working transparently with on-disk and in-memory array-like datasets	11
*	SummarizedExperiment	Bioconductor Package Maintainer	SummarizedExperiment container	12
*	GenomicRanges	Bioconductor Package Maintainer	Representation and manipulation of genomic intervals	13
	limma	Gordon Smyth	Linear Models for Microarray Data	14
*	Biostrings	H. Pagès	Efficient manipulation of biological strings	15
*	biomaRt	Mike Smith	Interface to BioMart databases (i.e. Ensembl)	16
	annotate	Bioconductor Package Maintainer	Annotation for microarrays	17
*	Rsamtools	Bioconductor Package Maintainer	Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import	18
	genefilter	Bioconductor Package Maintainer	genefilter: methods for filtering genes from high-throughput experiments	19
	graph	Bioconductor Package Maintainer	graph: A package to handle graph data structures	20
*	GenomicAlignments	Bioconductor Package Maintainer	Representation and manipulation of short genomic alignments	21
	BiocFileCache	Lori Shepherd	Manage Files Across Sessions	22
	Rhtslib	Bioconductor Package Maintainer	HTSlib high-throughput sequencing library as an R package	23
	edgeR	Yunshun Chen, Gordon Smyth, Aaron Lun, Mark Robinson	Empirical Analysis of Digital Gene Expression Data in R	24
*	rtracklayer	Michael Lawrence	R interface to genome annotation files and the UCSC genome browser	25
*	GenomicFeatures	Bioconductor Package Maintainer	Conveniently import and query gene models	26
*	DESeq2	Michael Love	Differential gene expression analysis based on the negative binomial distribution	27
	Rhdf5lib	Mike Smith	hdf5 library as an R package	28
	geneplotter	Bioconductor Package Maintainer	Graphics related functions for Bioconductor	29