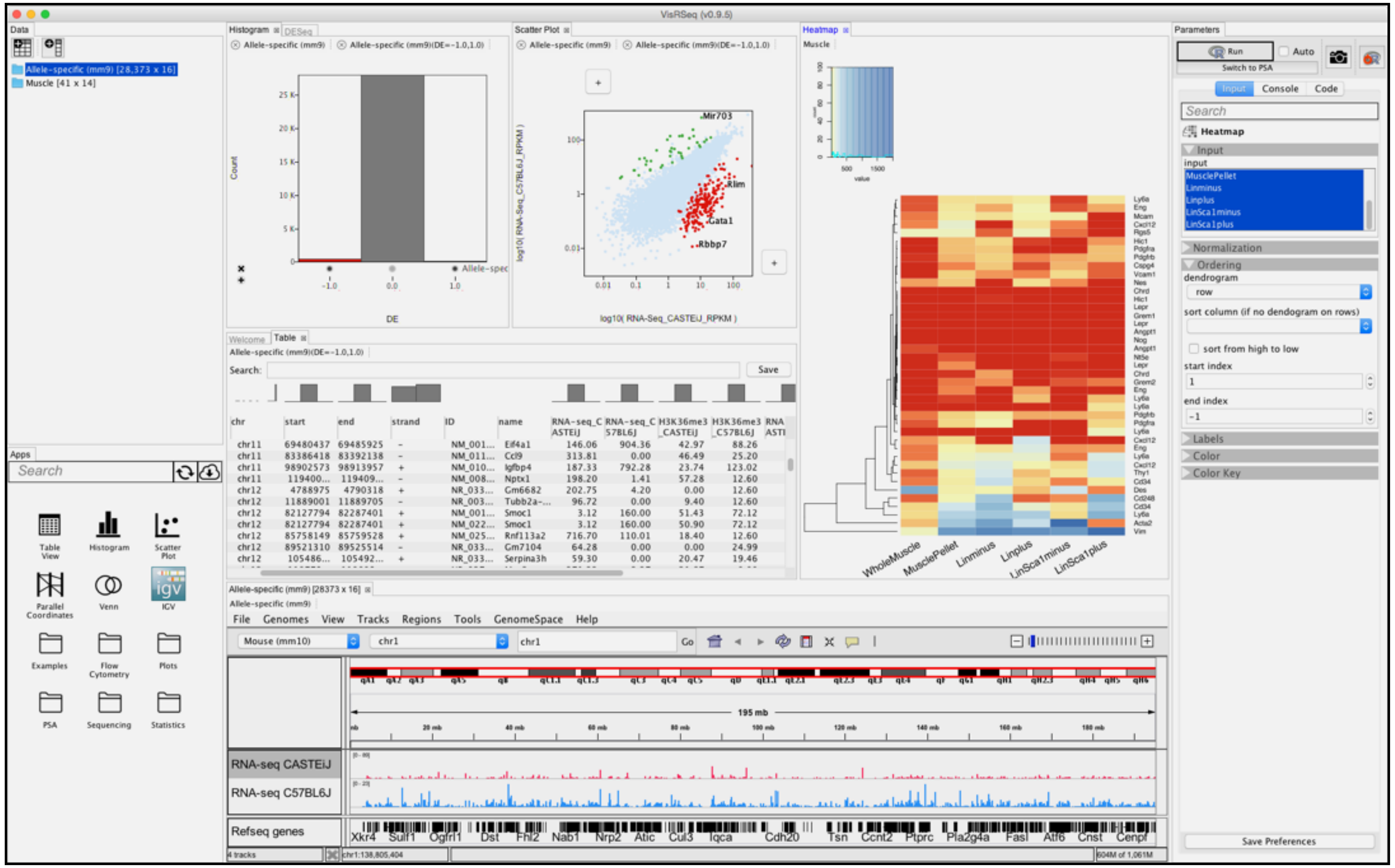
VisRseq: R-based visual analytics software for sequencing data

Hamid Younesy	*Graphics Usability and Visualization, Simon Fraser University, Canada*
Torsten Möller	*Visualization and Data Analysis Group, University of Vienna, Austria*
Matthew C. Lorincz	*Department of Medical Genetics, The University of British Columbia, Canada*
Mohammad M. Karimi	*Biomedical Research Centre, The University of British Columbia, Canada*
Steven J.M. Jones	*Genome Sciences Centre, BC Cancer Agency, Canada*

Overview

- introduction
- motivation
- VisRseq
- case study
- future works



Allele-specific (mm9)(DE=-1.0,1.0)

Search:

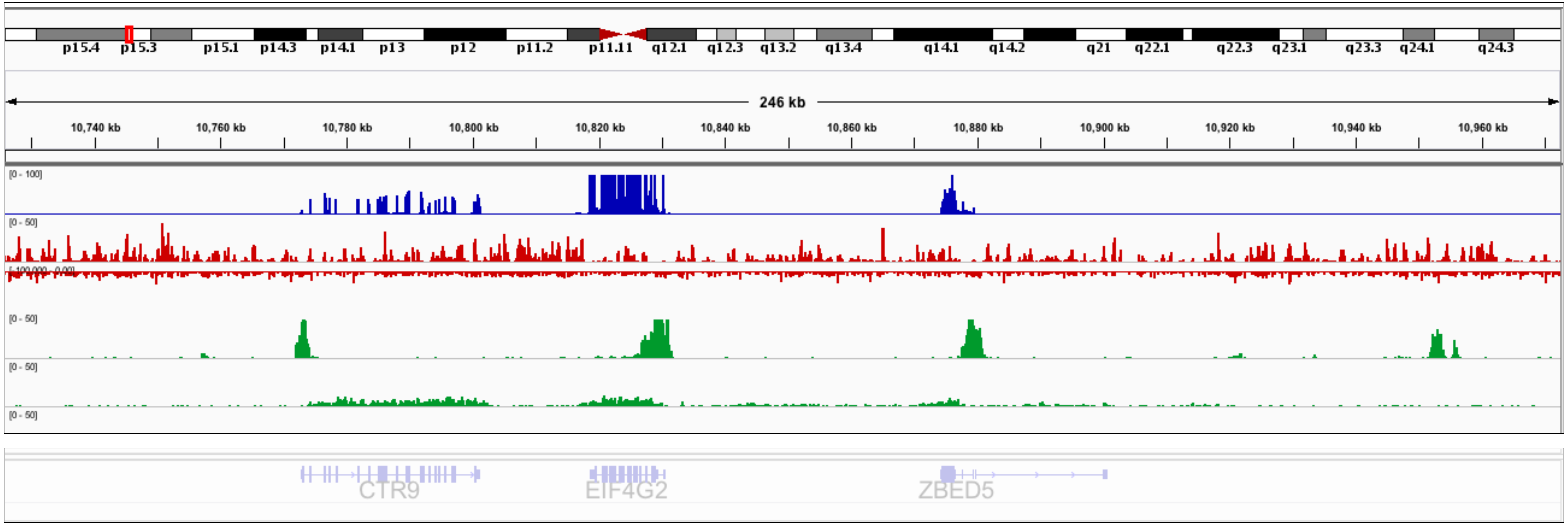
chr	start	end	strand	ID	name	RNA-seq_CASTEiJ	RNA-seq_C57BL6J	H3K36me3_CASTEiJ	H3K36me3_C57BL6J	RNA ASTI
chr11	69480437	69485925	-	NM_001...	Eif4a1	146.06	904.36	42.97	88.26	
chr11	83386418	83392138	-	NM_011...	Ccl9	313.81	0.00	46.49	25.20	
chr11	98902573	98913957	+	NM_010...	Igfbp4	187.33	792.28	23.74	123.02	
chr11	119400...	119409...	-	NM_008...	Nptx1	198.20	1.41	57.28	12.60	
chr12	4788975	4790318	+	NR_033...	Gm6682	202.75	4.20	0.00	12.60	
chr12	11889001	11889705	-	NR_003...	Tubb2a-...	96.72	0.00	9.40	12.60	
chr12	82127794	82287401	+	NM_001...	Smoc1	3.12	160.00	51.43	72.12	
chr12	82127794	82287401	+	NM_022...	Smoc1	3.12	160.00	50.90	72.12	
chr12	85758149	85759528	+	NM_025...	Rnf113a2	716.70	110.01	18.40	12.60	
chr12	89521310	89525514	-	NR_033...	Gm7104	64.28	0.00	0.00	24.99	
chr12	105486...	105492...	+	NR_033...	Serpina3h	59.30	0.00	20.47	19.46	

Allele-specific (mm9) [28373 x 16]

Allele-specific (mm9)

introduction

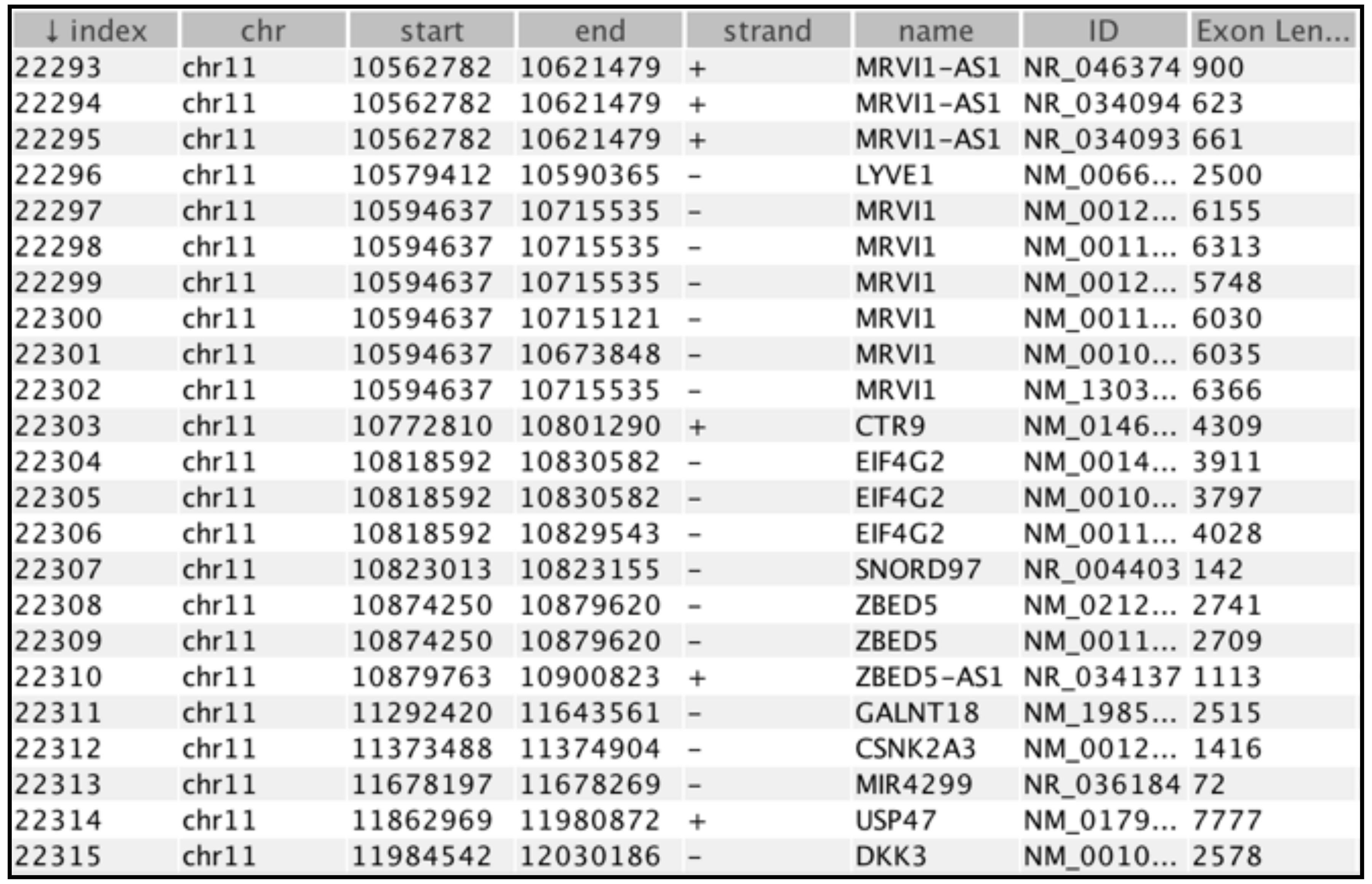
data: sequencing data



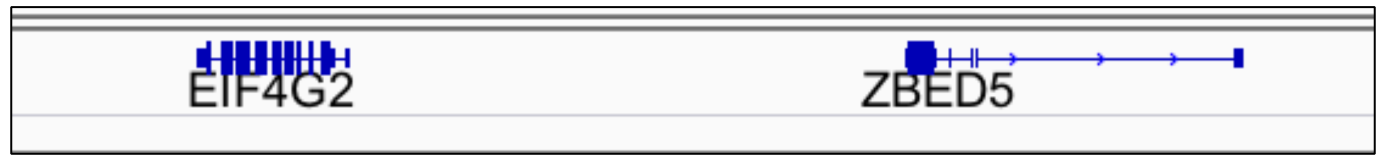
- numerical measures for biological properties per genomic location
- RNA-seq, ChIP-seq, Bis-seq, MeDIP, etc.

data: data table

regions

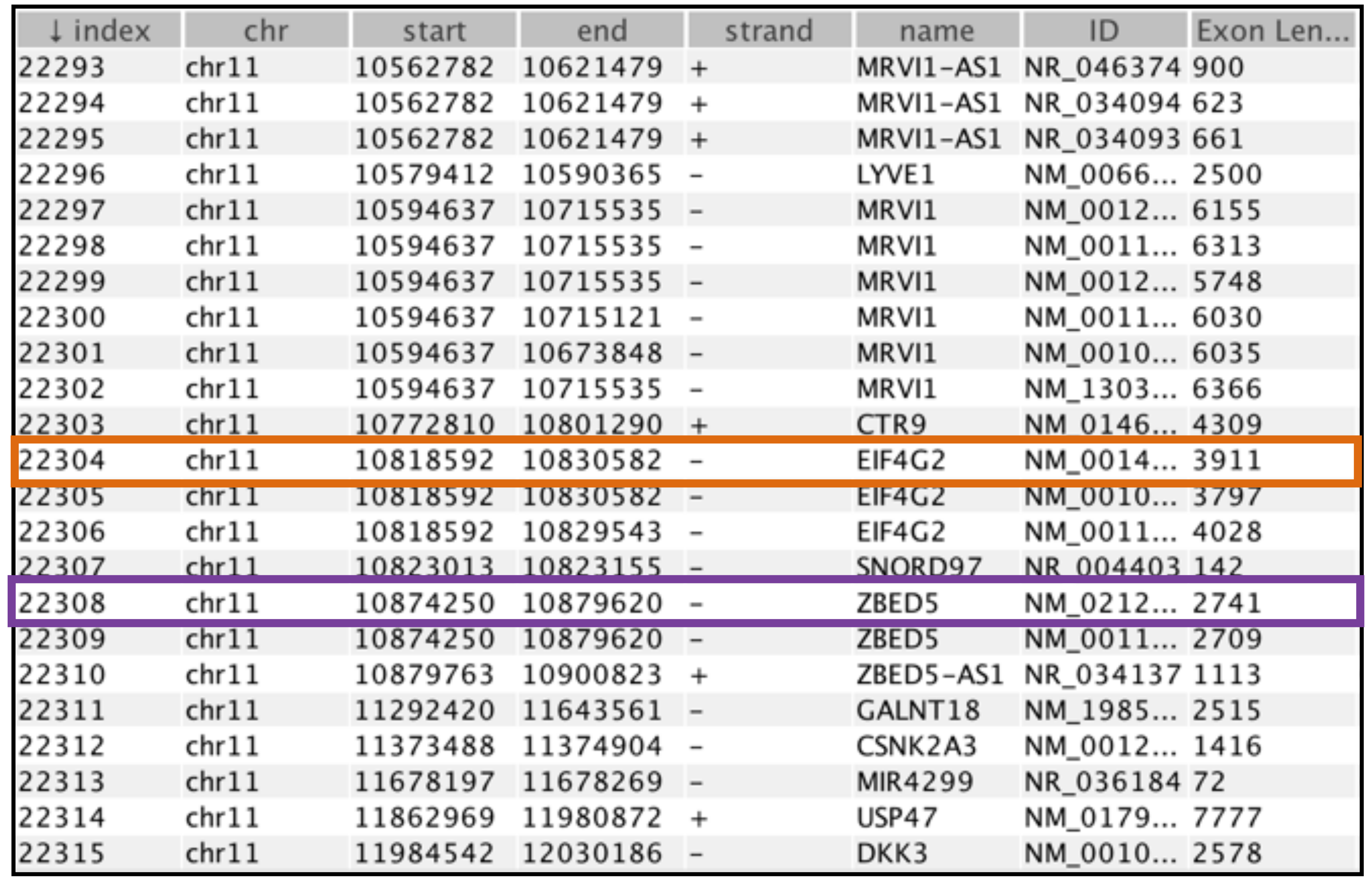


↓ index	chr	start	end	strand	name	ID	Exon Len...
22293	chr11	10562782	10621479	+	MRVI1-AS1	NR_046374	900
22294	chr11	10562782	10621479	+	MRVI1-AS1	NR_034094	623
22295	chr11	10562782	10621479	+	MRVI1-AS1	NR_034093	661
22296	chr11	10579412	10590365	-	LYVE1	NM_0066...	2500
22297	chr11	10594637	10715535	-	MRVI1	NM_0012...	6155
22298	chr11	10594637	10715535	-	MRVI1	NM_0011...	6313
22299	chr11	10594637	10715535	-	MRVI1	NM_0012...	5748
22300	chr11	10594637	10715121	-	MRVI1	NM_0011...	6030
22301	chr11	10594637	10673848	-	MRVI1	NM_0010...	6035
22302	chr11	10594637	10715535	-	MRVI1	NM_1303...	6366
22303	chr11	10772810	10801290	+	CTR9	NM_0146...	4309
22304	chr11	10818592	10830582	-	EIF4G2	NM_0014...	3911
22305	chr11	10818592	10830582	-	EIF4G2	NM_0010...	3797
22306	chr11	10818592	10829543	-	EIF4G2	NM_0011...	4028
22307	chr11	10823013	10823155	-	SNORD97	NR_004403	142
22308	chr11	10874250	10879620	-	ZBED5	NM_0212...	2741
22309	chr11	10874250	10879620	-	ZBED5	NM_0011...	2709
22310	chr11	10879763	10900823	+	ZBED5-AS1	NR_034137	1113
22311	chr11	11292420	11643561	-	GALNT18	NM_1985...	2515
22312	chr11	11373488	11374904	-	CSNK2A3	NM_0012...	1416
22313	chr11	11678197	11678269	-	MIR4299	NR_036184	72
22314	chr11	11862969	11980872	+	USP47	NM_0179...	7777
22315	chr11	11984542	12030186	-	DKK3	NM_0010...	2578

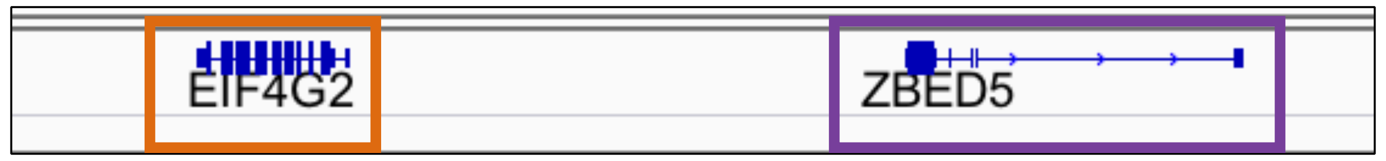


data: data table

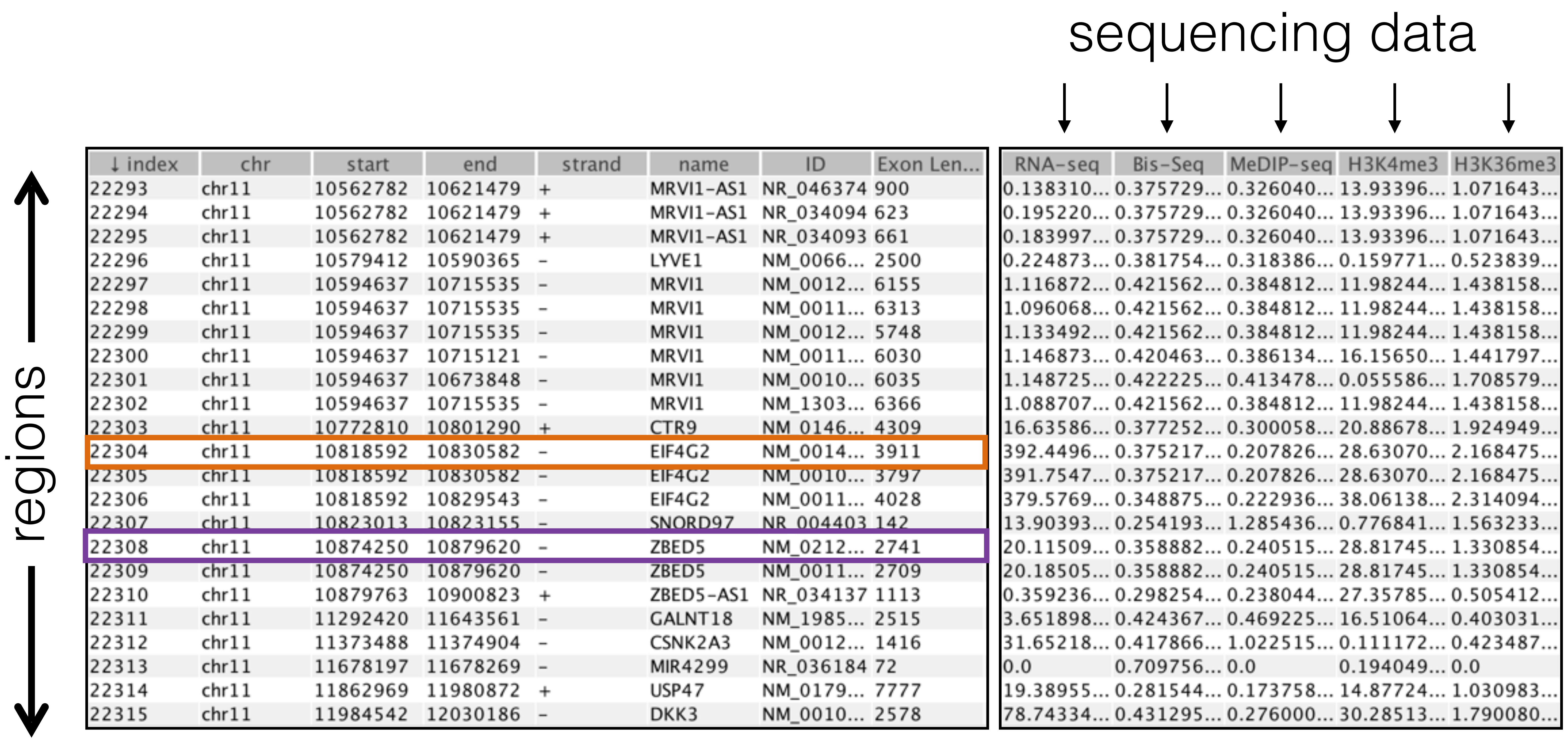
regions



↓ index	chr	start	end	strand	name	ID	Exon Len...
22293	chr11	10562782	10621479	+	MRVI1-AS1	NR_046374	900
22294	chr11	10562782	10621479	+	MRVI1-AS1	NR_034094	623
22295	chr11	10562782	10621479	+	MRVI1-AS1	NR_034093	661
22296	chr11	10579412	10590365	-	LYVE1	NM_0066...	2500
22297	chr11	10594637	10715535	-	MRVI1	NM_0012...	6155
22298	chr11	10594637	10715535	-	MRVI1	NM_0011...	6313
22299	chr11	10594637	10715535	-	MRVI1	NM_0012...	5748
22300	chr11	10594637	10715121	-	MRVI1	NM_0011...	6030
22301	chr11	10594637	10673848	-	MRVI1	NM_0010...	6035
22302	chr11	10594637	10715535	-	MRVI1	NM_1303...	6366
22303	chr11	10772810	10801290	+	CTR9	NM_0146...	4309
22304	chr11	10818592	10830582	-	EIF4G2	NM_0014...	3911
22305	chr11	10818592	10830582	-	EIF4G2	NM_0010...	3797
22306	chr11	10818592	10829543	-	EIF4G2	NM_0011...	4028
22307	chr11	10823013	10823155	-	SNORD97	NR_004403	142
22308	chr11	10874250	10879620	-	ZBED5	NM_0212...	2741
22309	chr11	10874250	10879620	-	ZBED5	NM_0011...	2709
22310	chr11	10879763	10900823	+	ZBED5-AS1	NR_034137	1113
22311	chr11	11292420	11643561	-	GALNT18	NM_1985...	2515
22312	chr11	11373488	11374904	-	CSNK2A3	NM_0012...	1416
22313	chr11	11678197	11678269	-	MIR4299	NR_036184	72
22314	chr11	11862969	11980872	+	USP47	NM_0179...	7777
22315	chr11	11984542	12030186	-	DKK3	NM_0010...	2578



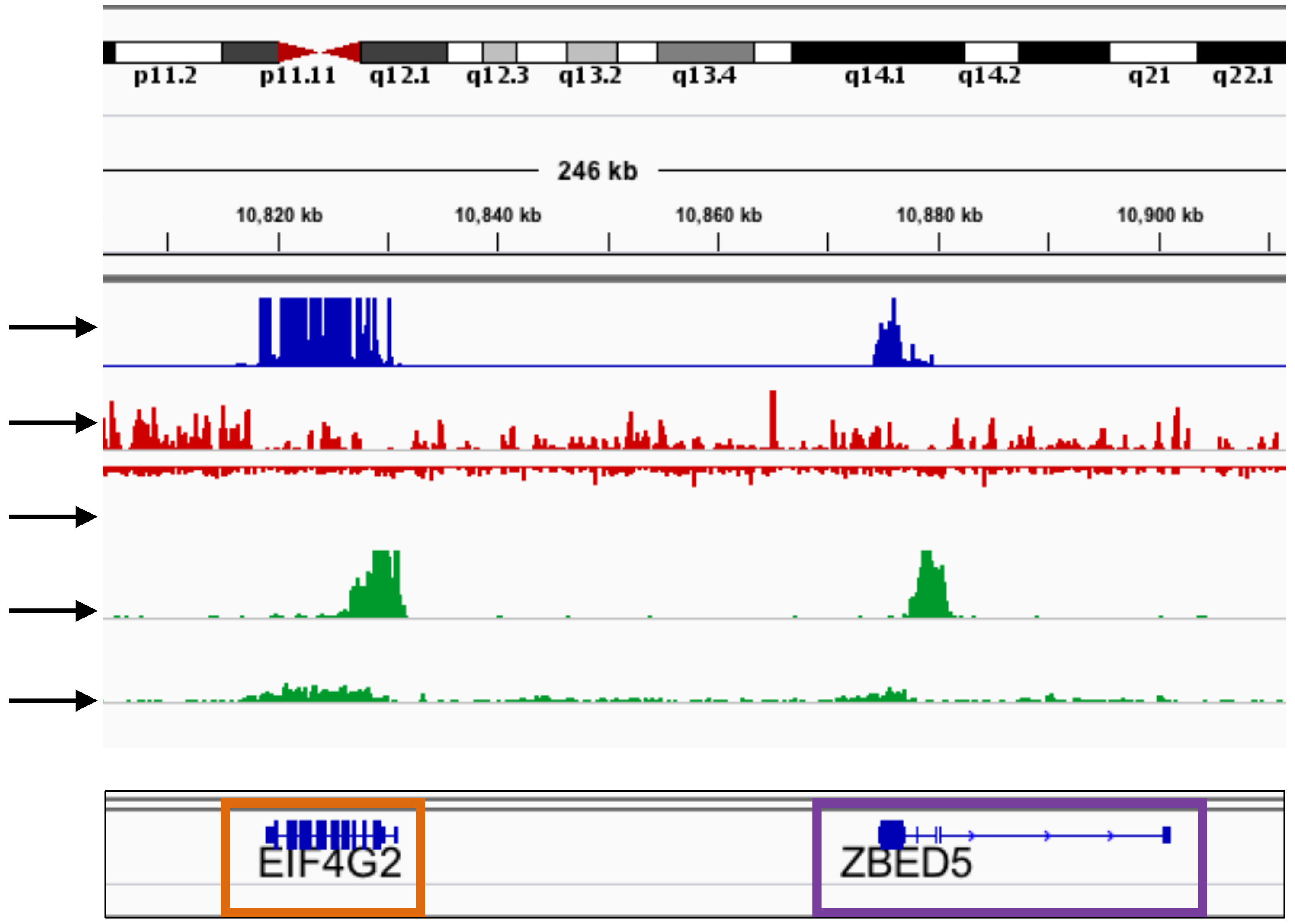
data: data table



sequencing data

regions

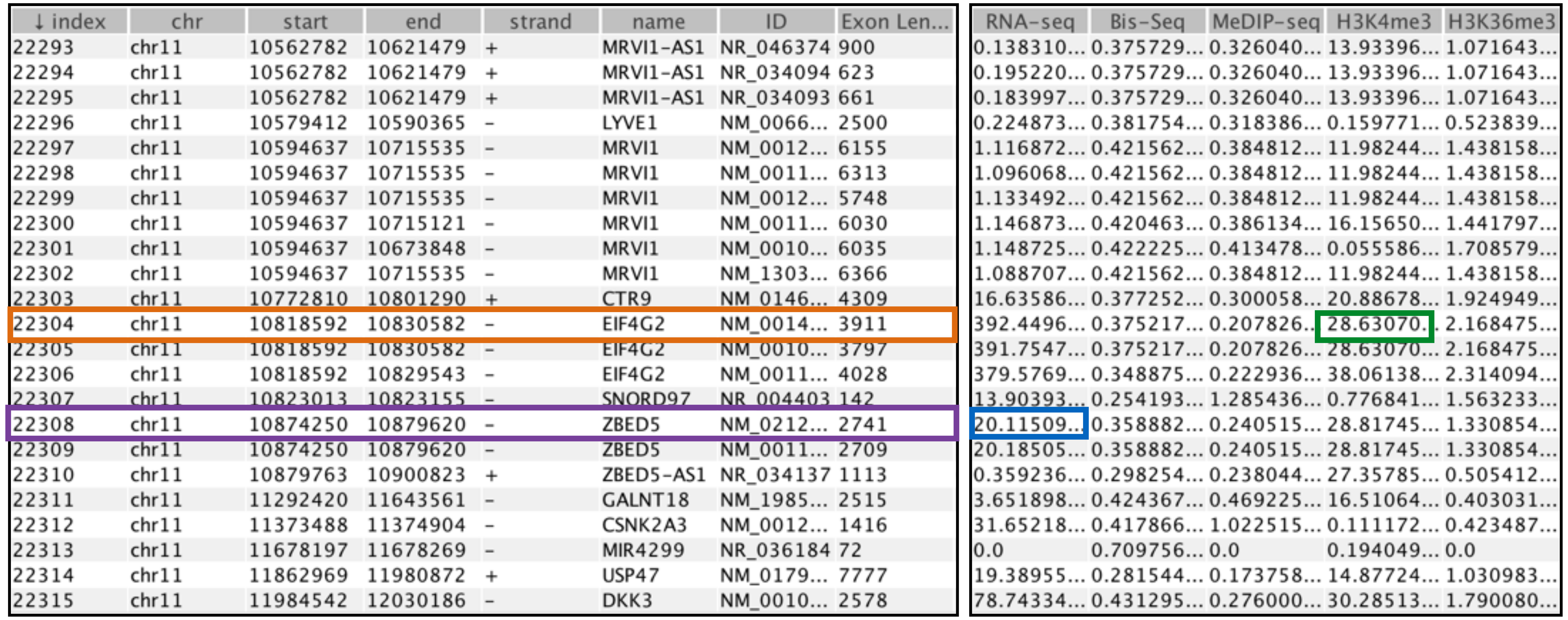
↓ index	chr	start	end	strand	name	ID	Exon Len...	RNA-seq	Bis-Seq	MeDIP-seq	H3K4me3	H3K36me3
22293	chr11	10562782	10621479	+	MRVI1-AS1	NR_046374	900	0.138310...	0.375729...	0.326040...	13.93396...	1.071643...
22294	chr11	10562782	10621479	+	MRVI1-AS1	NR_034094	623	0.195220...	0.375729...	0.326040...	13.93396...	1.071643...
22295	chr11	10562782	10621479	+	MRVI1-AS1	NR_034093	661	0.183997...	0.375729...	0.326040...	13.93396...	1.071643...
22296	chr11	10579412	10590365	-	LYVE1	NM_0066...	2500	0.224873...	0.381754...	0.318386...	0.159771...	0.523839...
22297	chr11	10594637	10715535	-	MRVI1	NM_0012...	6155	1.116872...	0.421562...	0.384812...	11.98244...	1.438158...
22298	chr11	10594637	10715535	-	MRVI1	NM_0011...	6313	1.096068...	0.421562...	0.384812...	11.98244...	1.438158...
22299	chr11	10594637	10715535	-	MRVI1	NM_0012...	5748	1.133492...	0.421562...	0.384812...	11.98244...	1.438158...
22300	chr11	10594637	10715121	-	MRVI1	NM_0011...	6030	1.146873...	0.420463...	0.386134...	16.15650...	1.441797...
22301	chr11	10594637	10673848	-	MRVI1	NM_0010...	6035	1.148725...	0.422225...	0.413478...	0.055586...	1.708579...
22302	chr11	10594637	10715535	-	MRVI1	NM_1303...	6366	1.088707...	0.421562...	0.384812...	11.98244...	1.438158...
22303	chr11	10772810	10801290	+	CTR9	NM_0146...	4309	16.63586...	0.377252...	0.300058...	20.88678...	1.924949...
22304	chr11	10818592	10830582	-	EIF4G2	NM_0014...	3911	392.4496...	0.375217...	0.207826...	28.63070...	2.168475...
22305	chr11	10818592	10830582	-	EIF4G2	NM_0010...	3797	391.7547...	0.375217...	0.207826...	28.63070...	2.168475...
22306	chr11	10818592	10829543	-	EIF4G2	NM_0011...	4028	379.5769...	0.348875...	0.222936...	38.06138...	2.314094...
22307	chr11	10823013	10823155	-	SNORD97	NR_004403	142	13.90393...	0.254193...	1.285436...	0.776841...	1.563233...
22308	chr11	10874250	10879620	-	ZBED5	NM_0212...	2741	20.11509...	0.358882...	0.240515...	28.81745...	1.330854...
22309	chr11	10874250	10879620	-	ZBED5	NM_0011...	2709	20.18505...	0.358882...	0.240515...	28.81745...	1.330854...
22310	chr11	10879763	10900823	+	ZBED5-AS1	NR_034137	1113	0.359236...	0.298254...	0.238044...	27.35785...	0.505412...
22311	chr11	11292420	11643561	-	GALNT18	NM_1985...	2515	3.651898...	0.424367...	0.469225...	16.51064...	0.403031...
22312	chr11	11373488	11374904	-	CSNK2A3	NM_0012...	1416	31.65218...	0.417866...	1.022515...	0.111172...	0.423487...
22313	chr11	11678197	11678269	-	MIR4299	NR_036184	72	0.0	0.709756...	0.0	0.194049...	0.0
22314	chr11	11862969	11980872	+	USP47	NM_0179...	7777	19.38955...	0.281544...	0.173758...	14.87724...	1.030983...
22315	chr11	11984542	12030186	-	DKK3	NM_0010...	2578	78.74334...	0.431295...	0.276000...	30.28513...	1.790080...



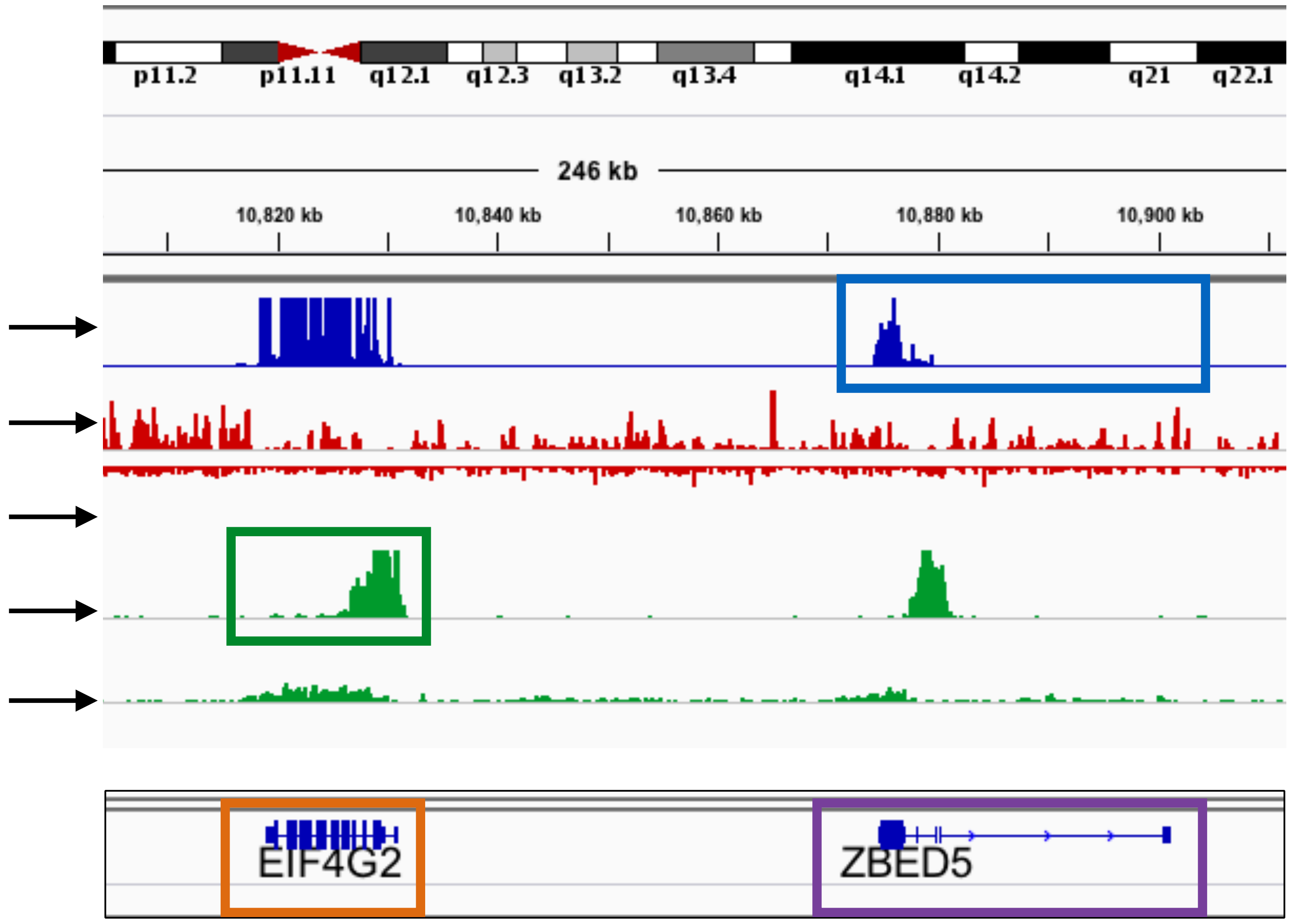
data: data table

sequencing data

regions



↓ index	chr	start	end	strand	name	ID	Exon Len...	RNA-seq	Bis-Seq	MeDIP-seq	H3K4me3	H3K36me3
22293	chr11	10562782	10621479	+	MRVI1-AS1	NR_046374	900	0.138310...	0.375729...	0.326040...	13.93396...	1.071643...
22294	chr11	10562782	10621479	+	MRVI1-AS1	NR_034094	623	0.195220...	0.375729...	0.326040...	13.93396...	1.071643...
22295	chr11	10562782	10621479	+	MRVI1-AS1	NR_034093	661	0.183997...	0.375729...	0.326040...	13.93396...	1.071643...
22296	chr11	10579412	10590365	-	LYVE1	NM_0066...	2500	0.224873...	0.381754...	0.318386...	0.159771...	0.523839...
22297	chr11	10594637	10715535	-	MRVI1	NM_0012...	6155	1.116872...	0.421562...	0.384812...	11.98244...	1.438158...
22298	chr11	10594637	10715535	-	MRVI1	NM_0011...	6313	1.096068...	0.421562...	0.384812...	11.98244...	1.438158...
22299	chr11	10594637	10715535	-	MRVI1	NM_0012...	5748	1.133492...	0.421562...	0.384812...	11.98244...	1.438158...
22300	chr11	10594637	10715121	-	MRVI1	NM_0011...	6030	1.146873...	0.420463...	0.386134...	16.15650...	1.441797...
22301	chr11	10594637	10673848	-	MRVI1	NM_0010...	6035	1.148725...	0.422225...	0.413478...	0.055586...	1.708579...
22302	chr11	10594637	10715535	-	MRVI1	NM_1303...	6366	1.088707...	0.421562...	0.384812...	11.98244...	1.438158...
22303	chr11	10772810	10801290	+	CTR9	NM_0146...	4309	16.63586...	0.377252...	0.300058...	20.88678...	1.924949...
22304	chr11	10818592	10830582	-	EIF4G2	NM_0014...	3911	392.4496...	0.375217...	0.207826...	28.63070...	2.168475...
22305	chr11	10818592	10830582	-	EIF4G2	NM_0010...	3797	391.7547...	0.375217...	0.207826...	28.63070...	2.168475...
22306	chr11	10818592	10829543	-	EIF4G2	NM_0011...	4028	379.5769...	0.348875...	0.222936...	38.06138...	2.314094...
22307	chr11	10823013	10823155	-	SNORD97	NR_004403	142	13.90393...	0.254193...	1.285436...	0.776841...	1.563233...
22308	chr11	10874250	10879620	-	ZBED5	NM_0212...	2741	20.11509...	0.358882...	0.240515...	28.81745...	1.330854...
22309	chr11	10874250	10879620	-	ZBED5	NM_0011...	2709	20.18505...	0.358882...	0.240515...	28.81745...	1.330854...
22310	chr11	10879763	10900823	+	ZBED5-AS1	NR_034137	1113	0.359236...	0.298254...	0.238044...	27.35785...	0.505412...
22311	chr11	11292420	11643561	-	GALNT18	NM_1985...	2515	3.651898...	0.424367...	0.469225...	16.51064...	0.403031...
22312	chr11	11373488	11374904	-	CSNK2A3	NM_0012...	1416	31.65218...	0.417866...	1.022515...	0.111172...	0.423487...
22313	chr11	11678197	11678269	-	MIR4299	NR_036184	72	0.0	0.709756...	0.0	0.194049...	0.0
22314	chr11	11862969	11980872	+	USP47	NM_0179...	7777	19.38955...	0.281544...	0.173758...	14.87724...	1.030983...
22315	chr11	11984542	12030186	-	DKK3	NM_0010...	2578	78.74334...	0.431295...	0.276000...	30.28513...	1.790080...



motivation

motivation

analysis software

motivation

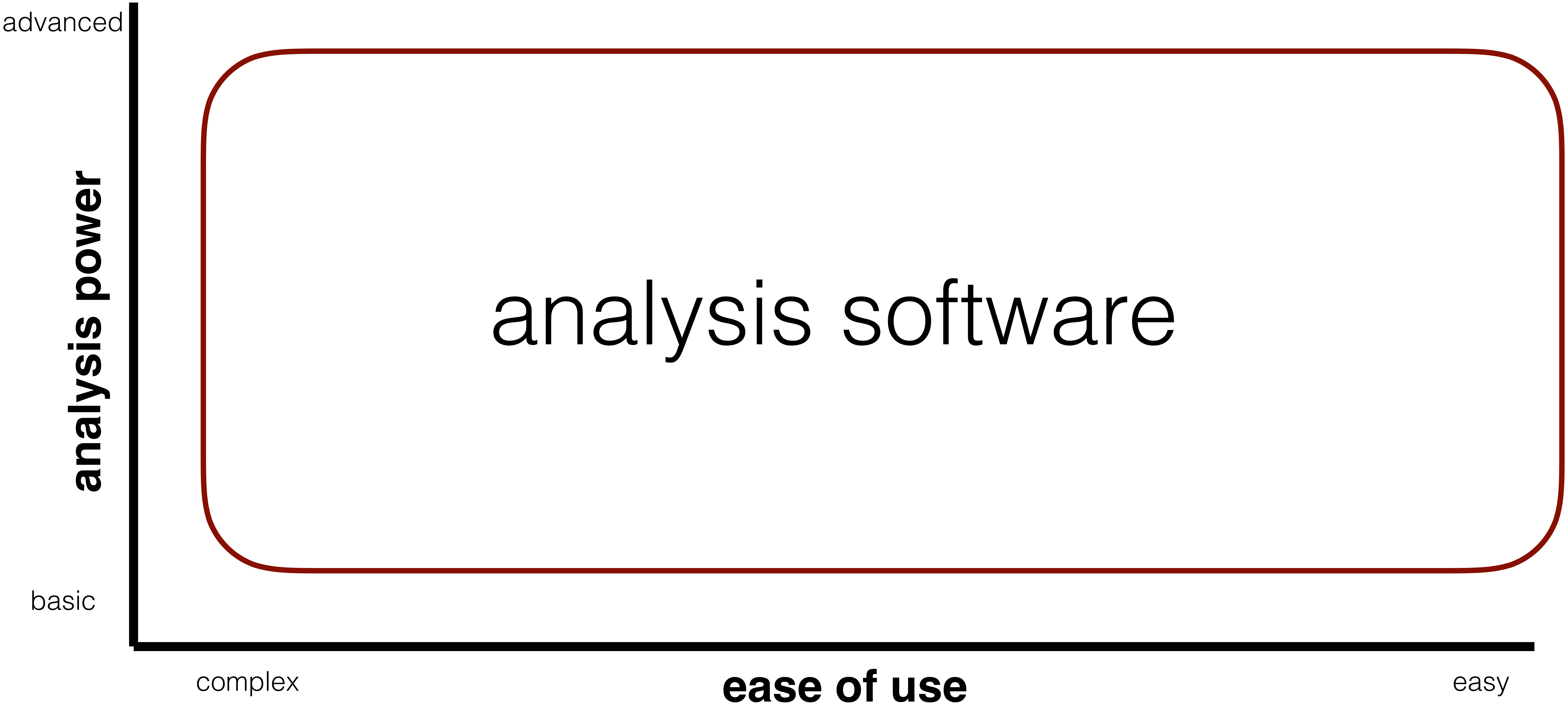
analysis software

complex

ease of use

easy

motivation



advanced

analysis power

basic

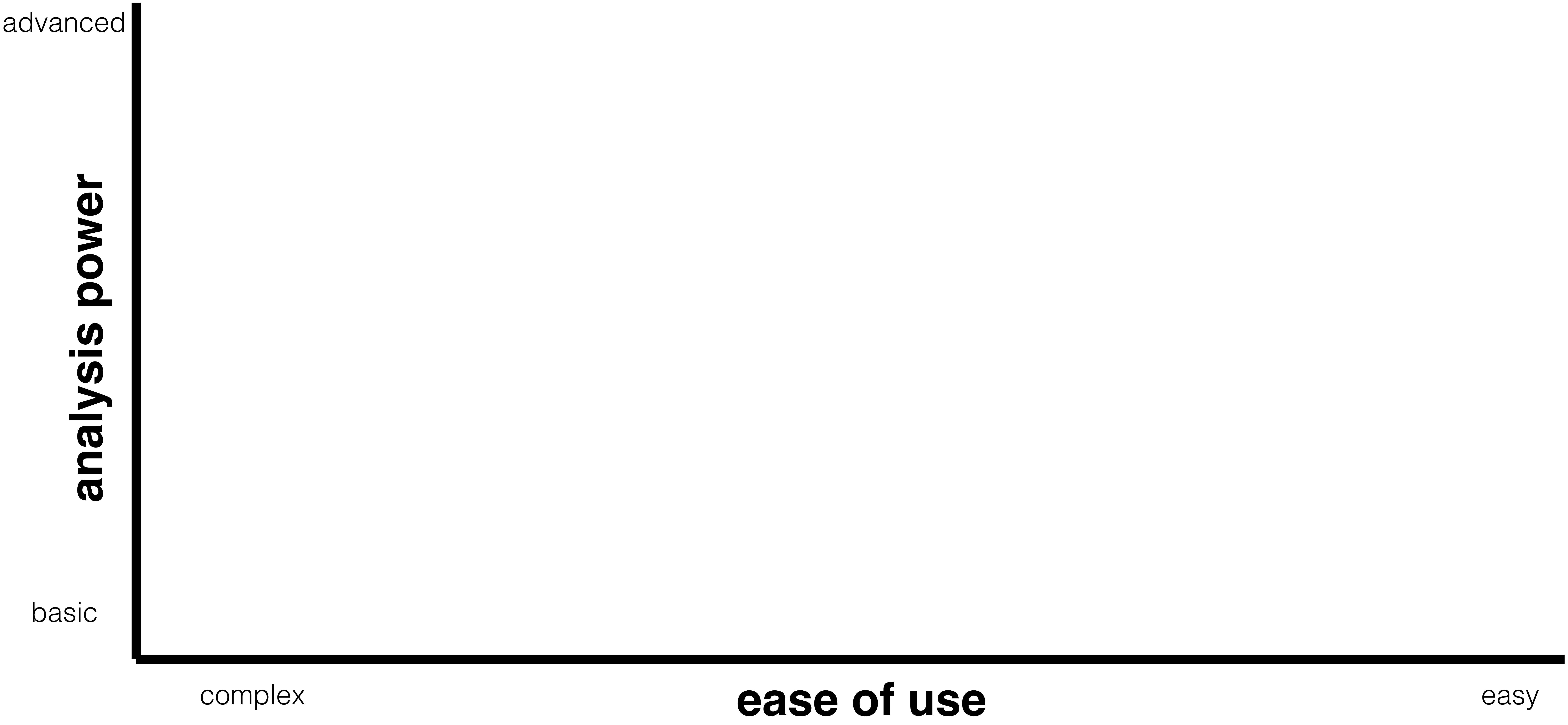
analysis software

complex

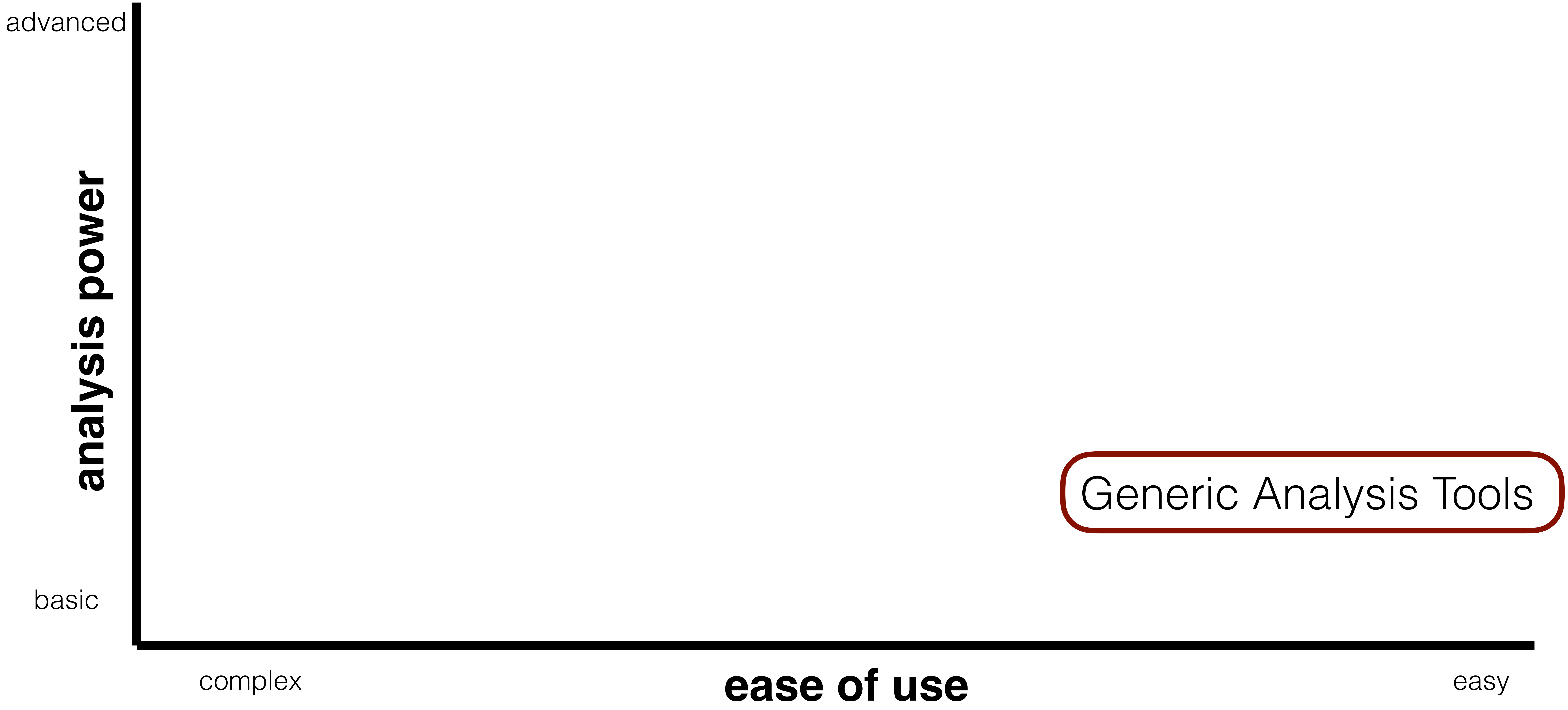
ease of use

easy

motivation



motivation



advanced

analysis power

basic

Generic Analysis Tools

complex

ease of use

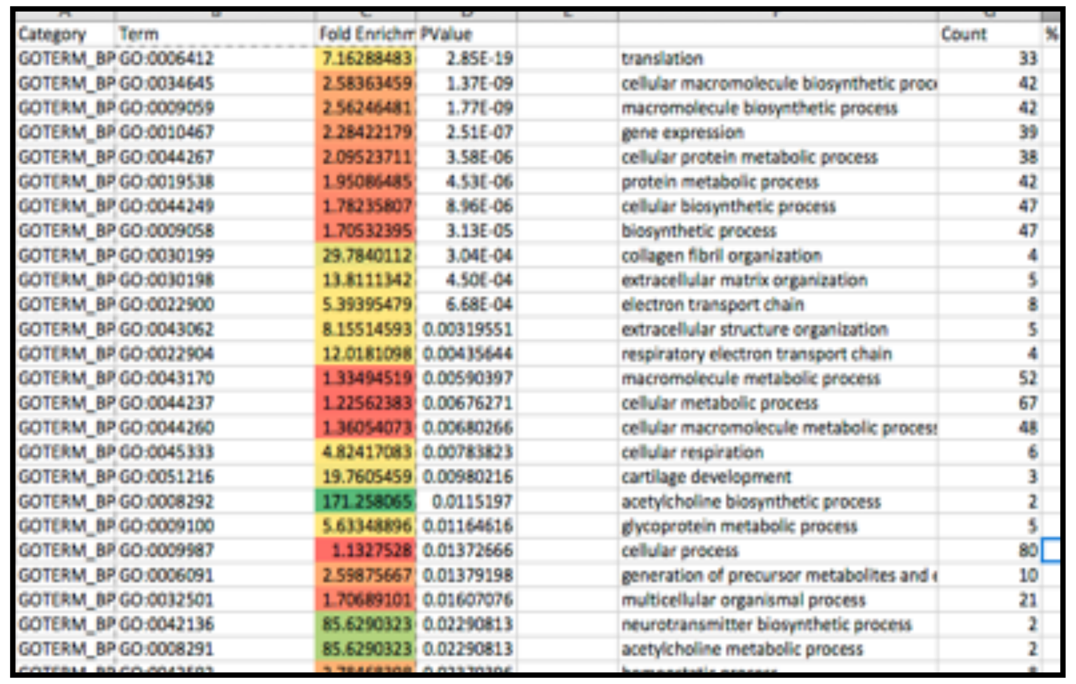
easy

motivation

advanced

analysis power

basic



Category	Term	Fold Enrichm	PValue			Count	%
GOTERM_BP	GO:0006412	7.16288483	2.85E-19		translation	33	
GOTERM_BP	GO:0034645	2.58363459	1.37E-09		cellular macromolecule biosynthetic proc	42	
GOTERM_BP	GO:0009059	2.56246481	1.77E-09		macromolecule biosynthetic process	42	
GOTERM_BP	GO:0010467	2.28422179	2.51E-07		gene expression	39	
GOTERM_BP	GO:0044267	2.09523711	3.58E-06		cellular protein metabolic process	38	
GOTERM_BP	GO:0019538	1.95086485	4.53E-06		protein metabolic process	42	
GOTERM_BP	GO:0044249	1.78235807	8.96E-06		cellular biosynthetic process	47	
GOTERM_BP	GO:0009058	1.70532395	3.13E-05		biosynthetic process	47	
GOTERM_BP	GO:0030199	29.7840112	3.04E-04		collagen fibril organization	4	
GOTERM_BP	GO:0030198	13.8111342	4.50E-04		extracellular matrix organization	5	
GOTERM_BP	GO:0022900	5.39395479	6.68E-04		electron transport chain	8	
GOTERM_BP	GO:0043062	8.15514593	0.00319551		extracellular structure organization	5	
GOTERM_BP	GO:0022904	12.0181098	0.00435644		respiratory electron transport chain	4	
GOTERM_BP	GO:0043170	1.33494519	0.00590397		macromolecule metabolic process	52	
GOTERM_BP	GO:0044237	1.22562383	0.00676271		cellular metabolic process	67	
GOTERM_BP	GO:0044260	1.36054073	0.00680266		cellular macromolecule metabolic proces	48	
GOTERM_BP	GO:0045333	4.82417083	0.00783823		cellular respiration	6	
GOTERM_BP	GO:0051216	19.7605459	0.00980216		cartilage development	3	
GOTERM_BP	GO:0008292	171.258065	0.0115197		acetylcholine biosynthetic process	2	
GOTERM_BP	GO:0009100	5.63348896	0.01164616		glycoprotein metabolic process	5	
GOTERM_BP	GO:0009987	1.1327528	0.01372666		cellular process	80	
GOTERM_BP	GO:0006091	2.59875667	0.01379198		generation of precursor metabolites and	10	
GOTERM_BP	GO:0032501	1.70689101	0.01607076		multicellular organismal process	21	
GOTERM_BP	GO:0042136	85.6290323	0.02290813		neurotransmitter biosynthetic process	2	
GOTERM_BP	GO:0008291	85.6290323	0.02290813		acetylcholine metabolic process	2	

Excel

Tableau

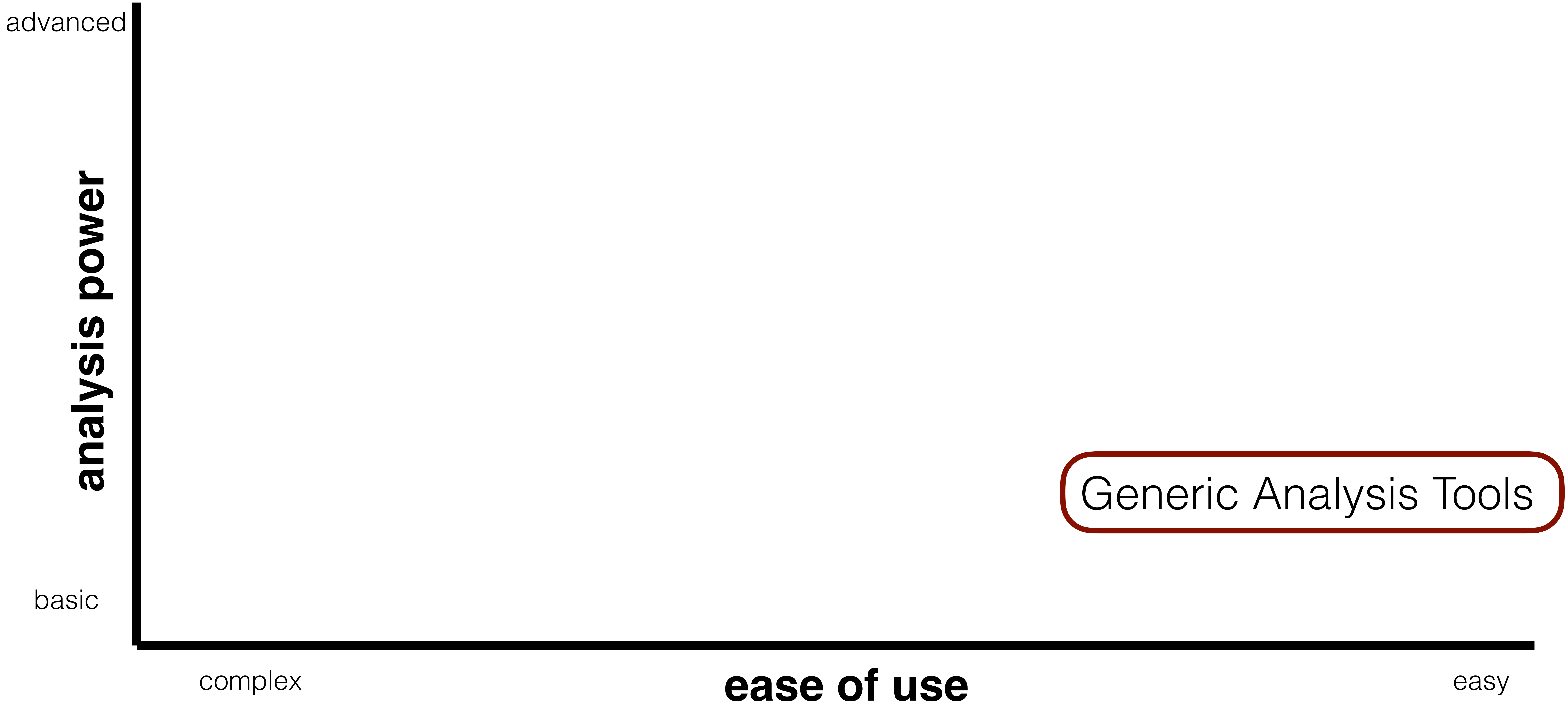
Generic Analysis Tools

complex

ease of use

easy

motivation



advanced

analysis power

basic

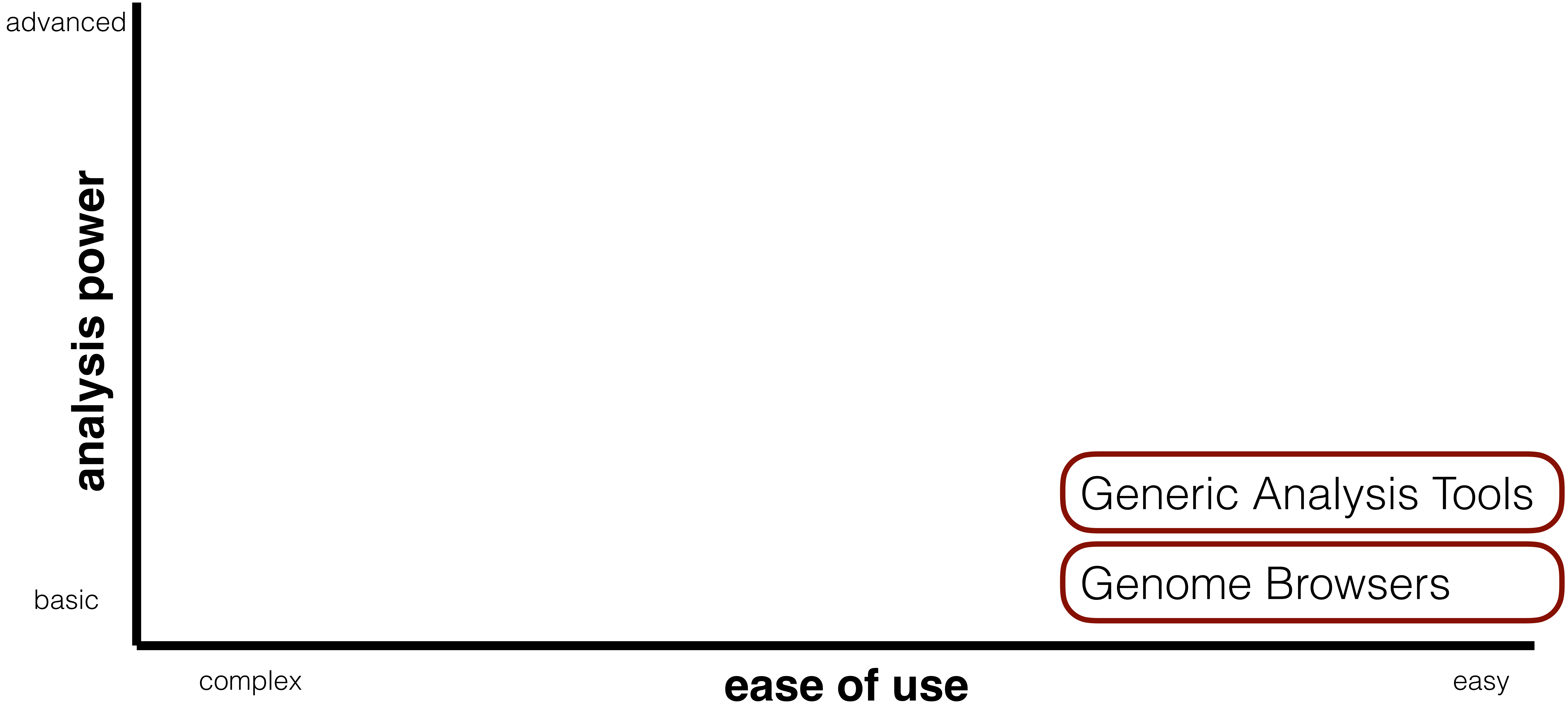
Generic Analysis Tools

complex

ease of use

easy

motivation

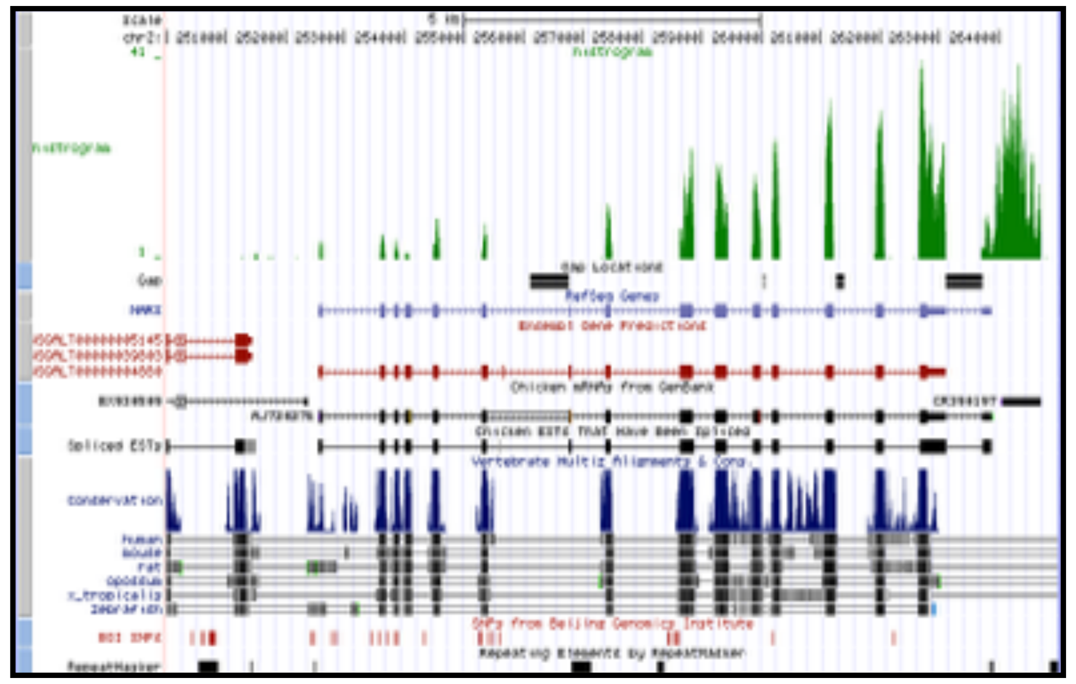

advanced
analysis power
basic
Generic Analysis Tools
Genome Browsers
complex
ease of use
easy

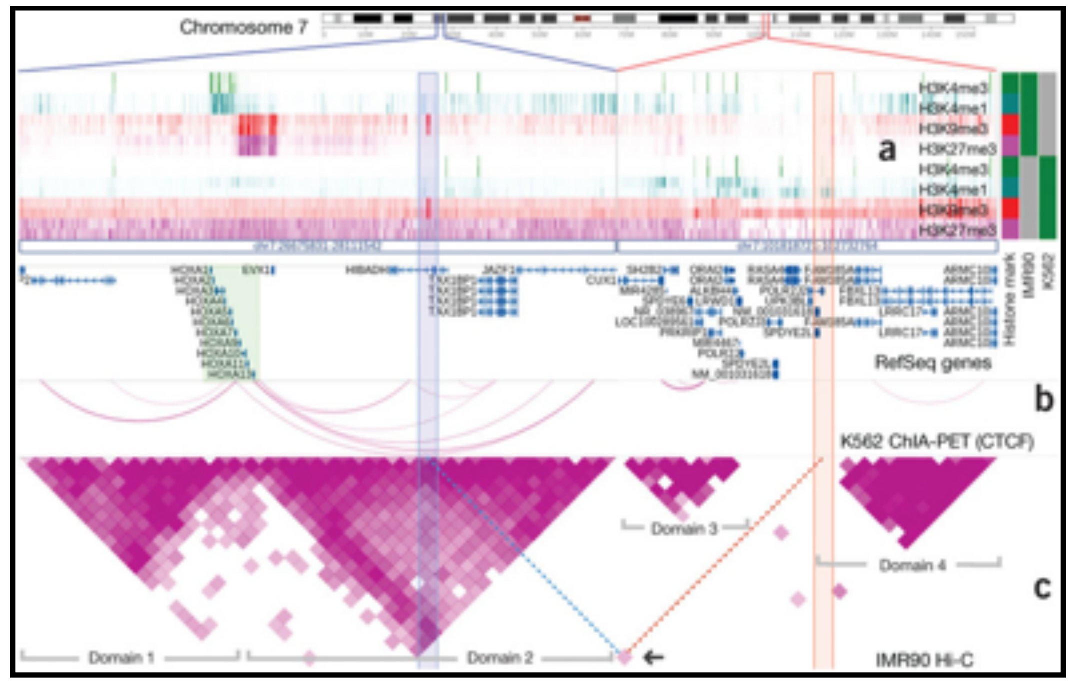
motivation

advanced

analysis power

basic


UCSC Genome Browser


WashU Epigenome Browser


Integrated Genome Viewer

Generic Analysis Tools

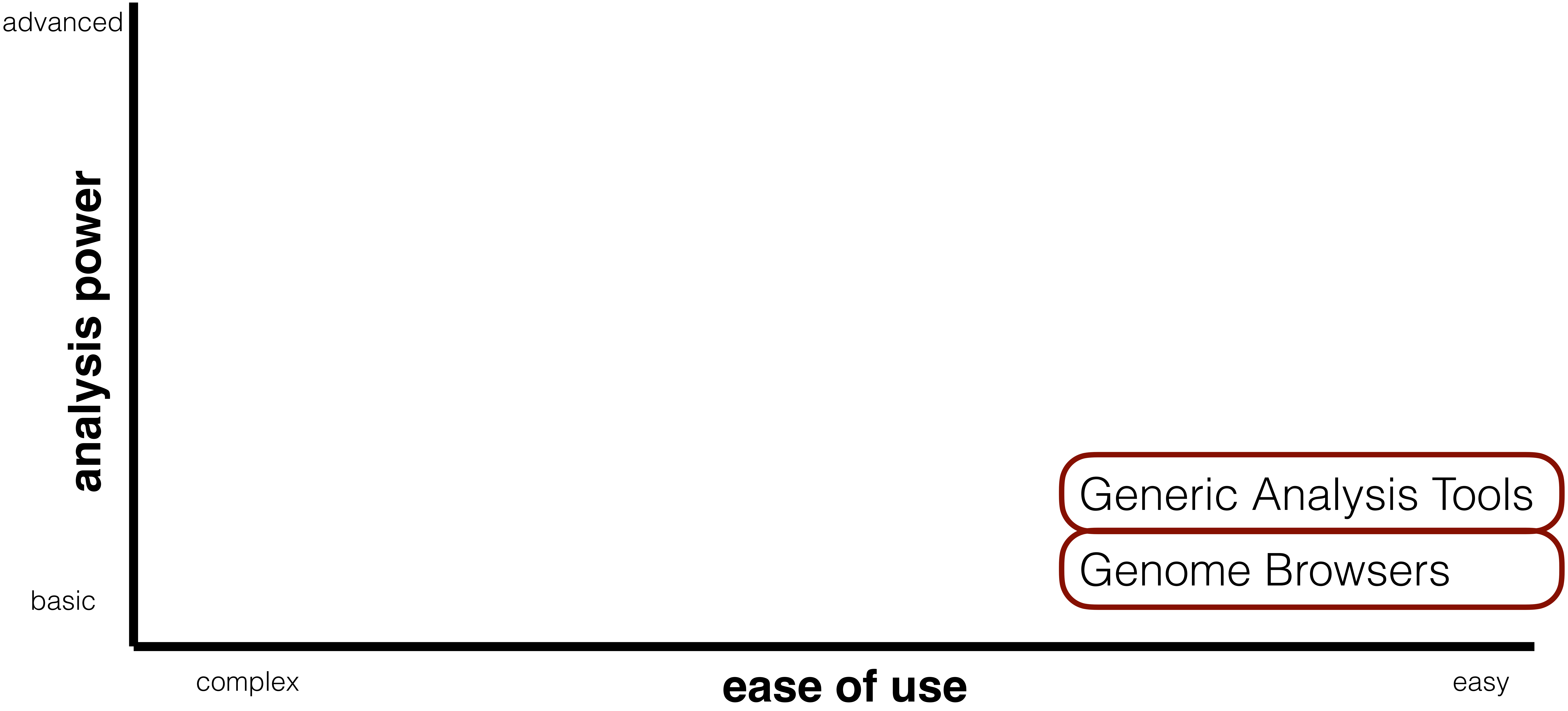
Genome Browsers

complex

ease of use

easy

motivation



advanced

analysis power

basic

Generic Analysis Tools

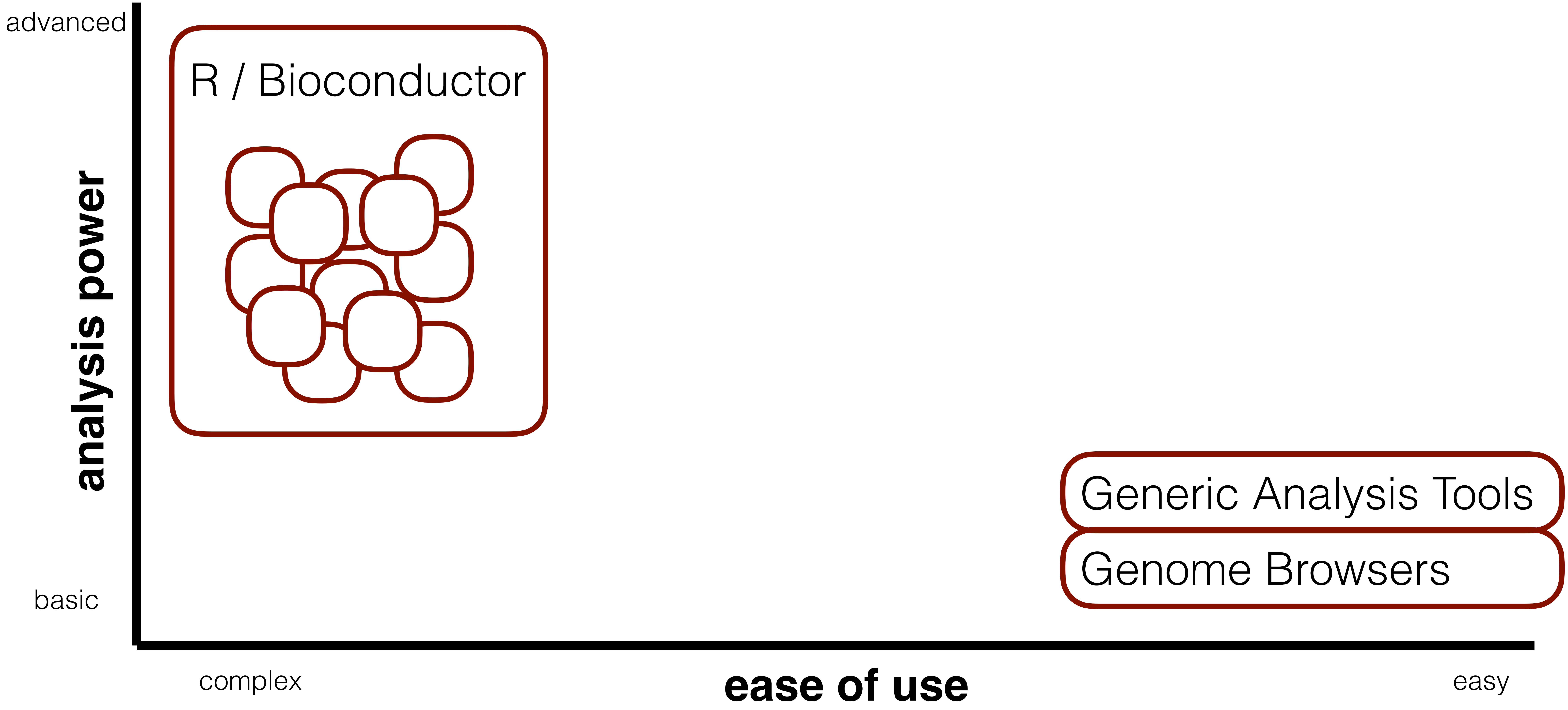
Genome Browsers

complex

ease of use

easy

motivation


advanced
analysis power
basic
R / Bioconductor
Generic Analysis Tools
Genome Browsers
complex
ease of use
easy

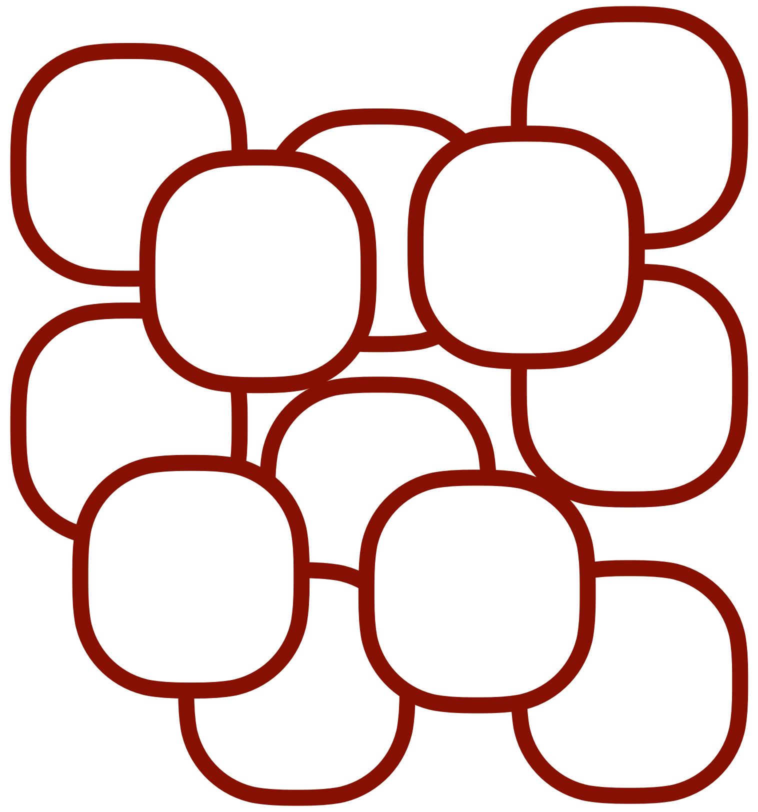
motivation

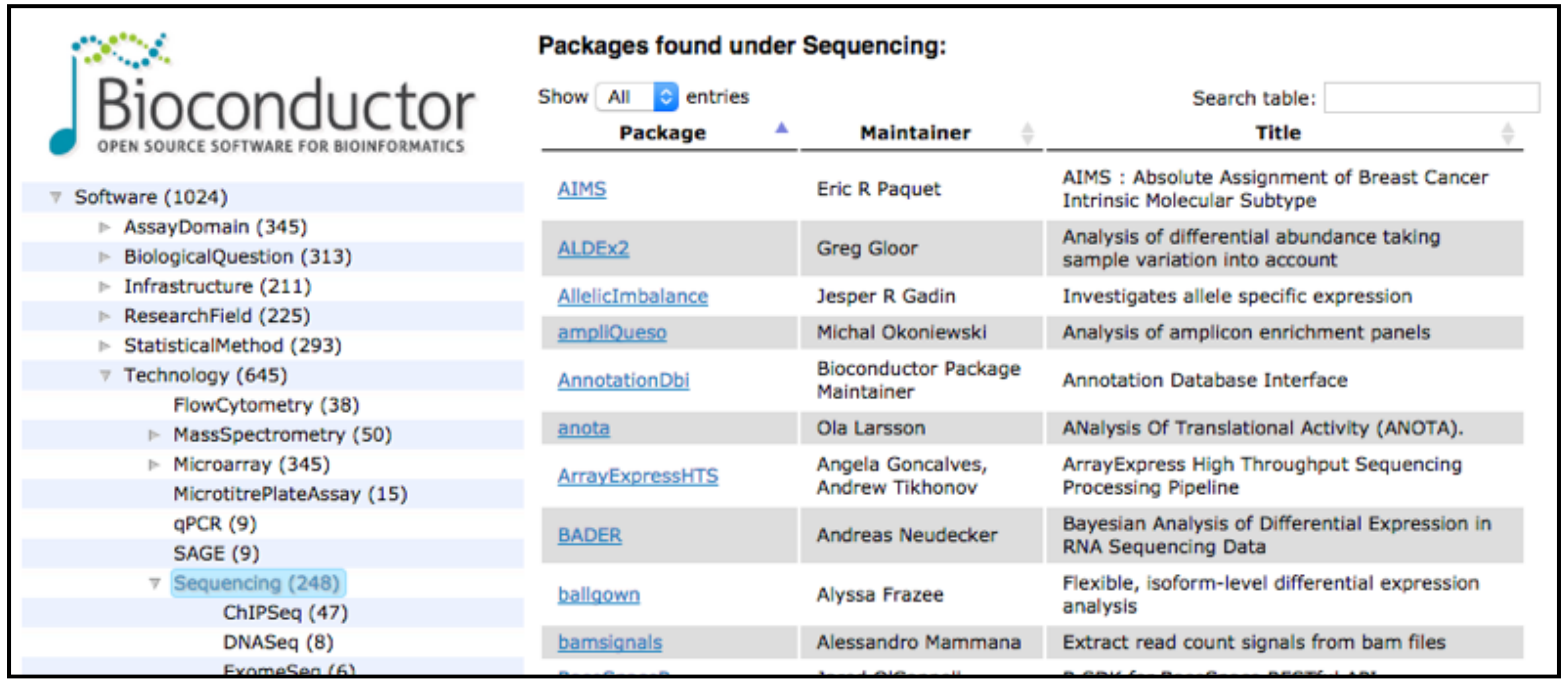
advanced

analysis power

basic

R / Bioconductor





Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

- Software (1024)
  - AssayDomain (345)
  - BiologicalQuestion (313)
  - Infrastructure (211)
  - ResearchField (225)
  - StatisticalMethod (293)
  - Technology (645)
    - FlowCytometry (38)
    - MassSpectrometry (50)
    - Microarray (345)
    - MicrotitrePlateAssay (15)
    - qPCR (9)
    - SAGE (9)
    - Sequencing (248)
      - ChIPSeq (47)
      - DNASeq (8)
      - ExomeSeq (6)

Packages found under Sequencing:

Show All entries

Search table:

Package	Maintainer	Title
AIMS	Eric R Paquet	AIMS : Absolute Assignment of Breast Cancer Intrinsic Molecular Subtype
ALDEx2	Greg Gloor	Analysis of differential abundance taking sample variation into account
AllelicImbalance	Jesper R Gadin	Investigates allele specific expression
ampliQueso	Michal Okoniewski	Analysis of amplicon enrichment panels
AnnotationDbi	Bioconductor Package Maintainer	Annotation Database Interface
anota	Ola Larsson	ANalysis Of Translational Activity (ANOTA).
ArrayExpressHTS	Angela Goncalves, Andrew Tikhonov	ArrayExpress High Throughput Sequencing Processing Pipeline
BADER	Andreas Neudecker	Bayesian Analysis of Differential Expression in RNA Sequencing Data
ballgown	Alyssa Frazee	Flexible, isoform-level differential expression analysis
bamsignals	Alessandro Mammana	Extract read count signals from bam files

Bioconductor Repository

Generic Analysis Tools

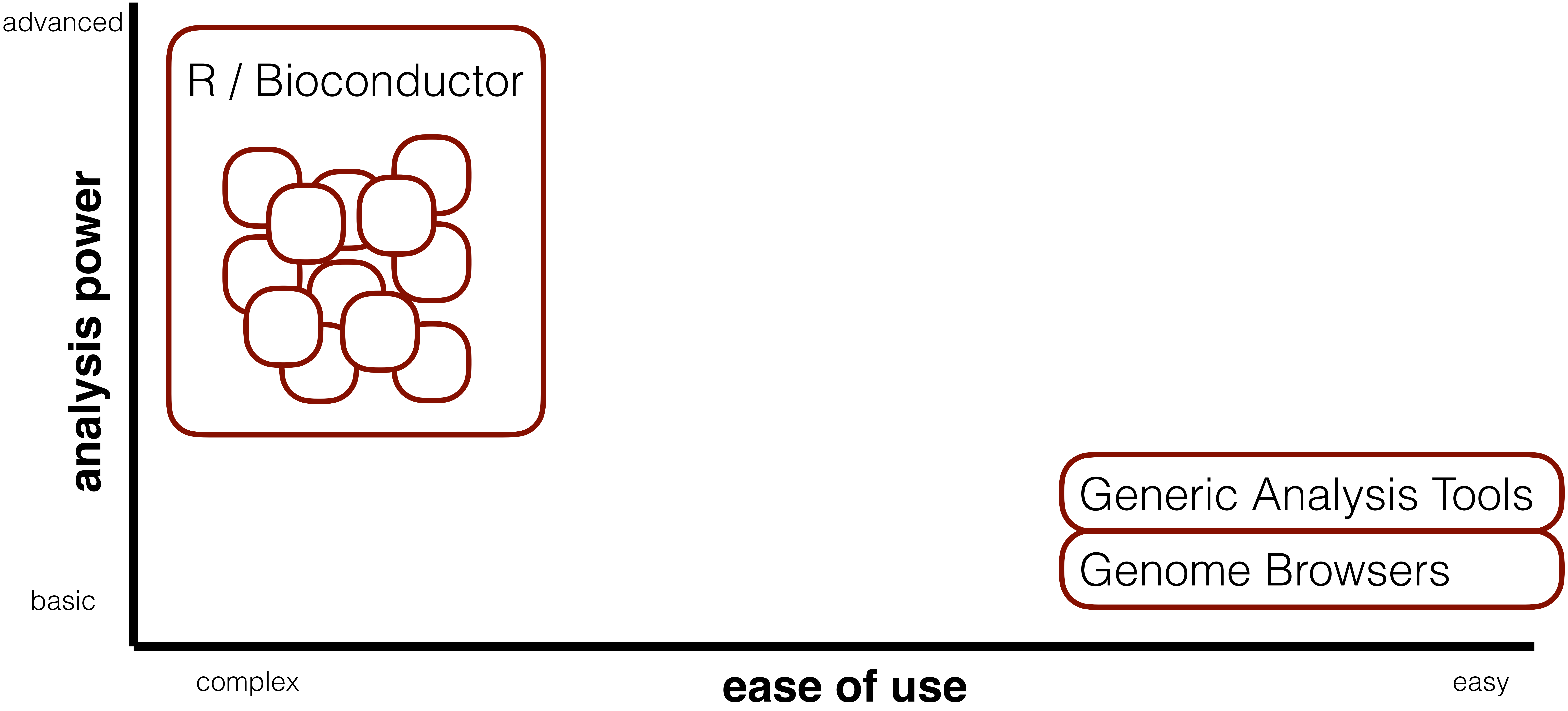
Genome Browsers

complex

ease of use

easy

motivation



advanced

analysis power

basic

R / Bioconductor

Generic Analysis Tools

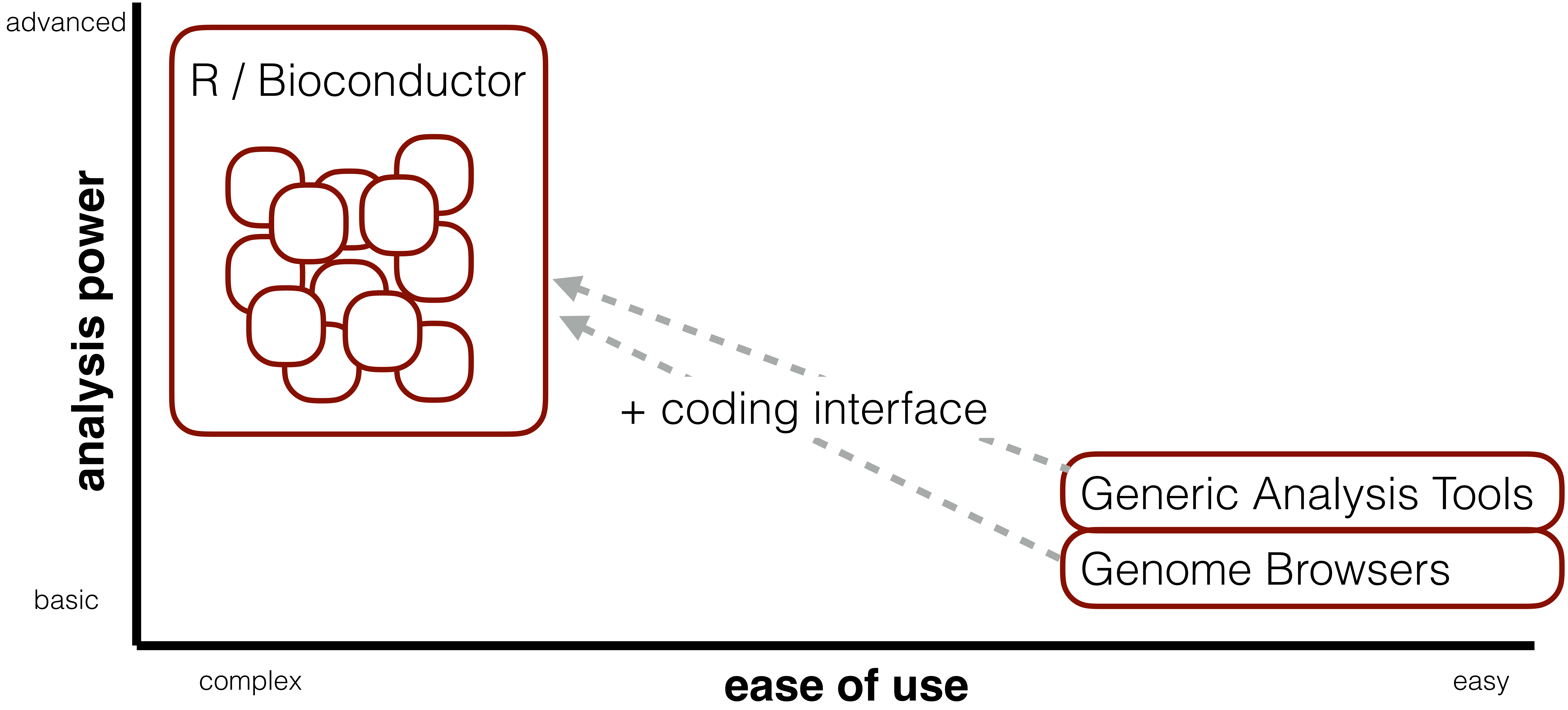
Genome Browsers

complex

ease of use

easy

motivation


advanced
basic
analysis power
complex
ease of use
easy
R / Bioconductor
+ coding interface
Generic Analysis Tools
Genome Browsers

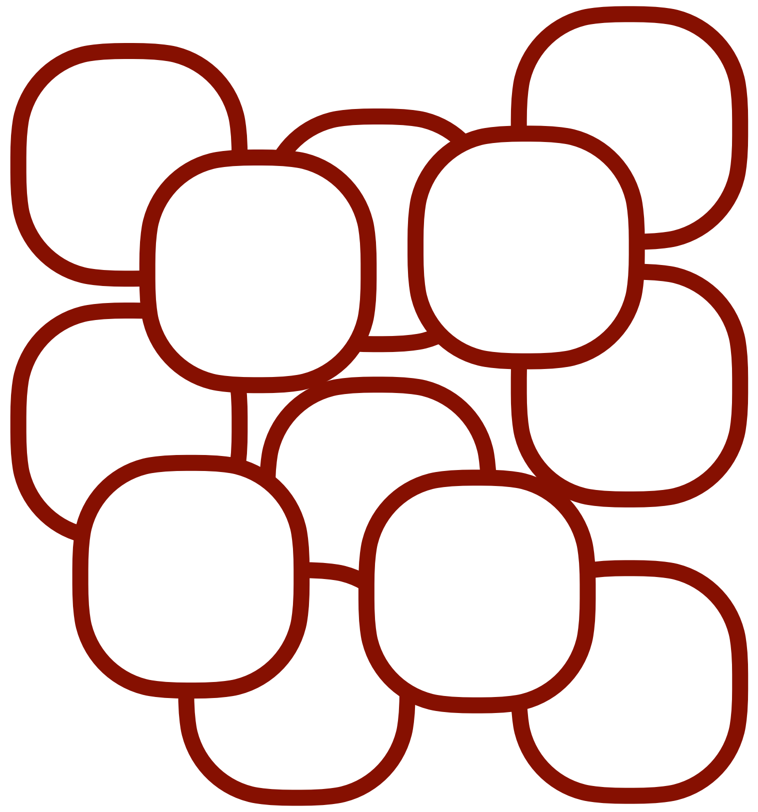
motivation

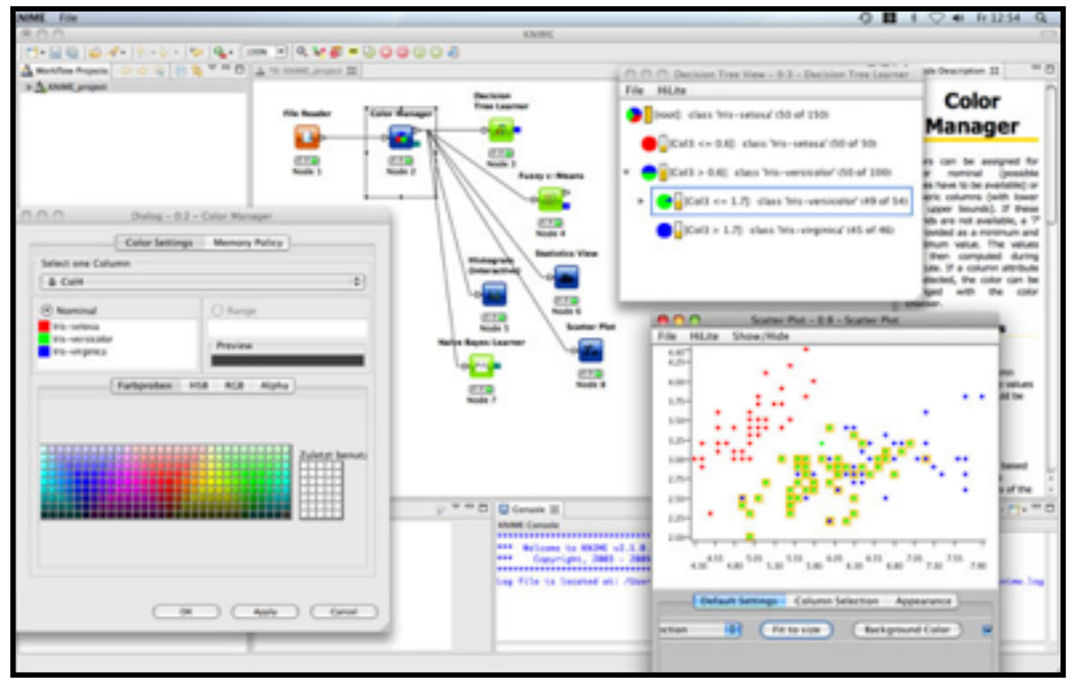
advanced

analysis power

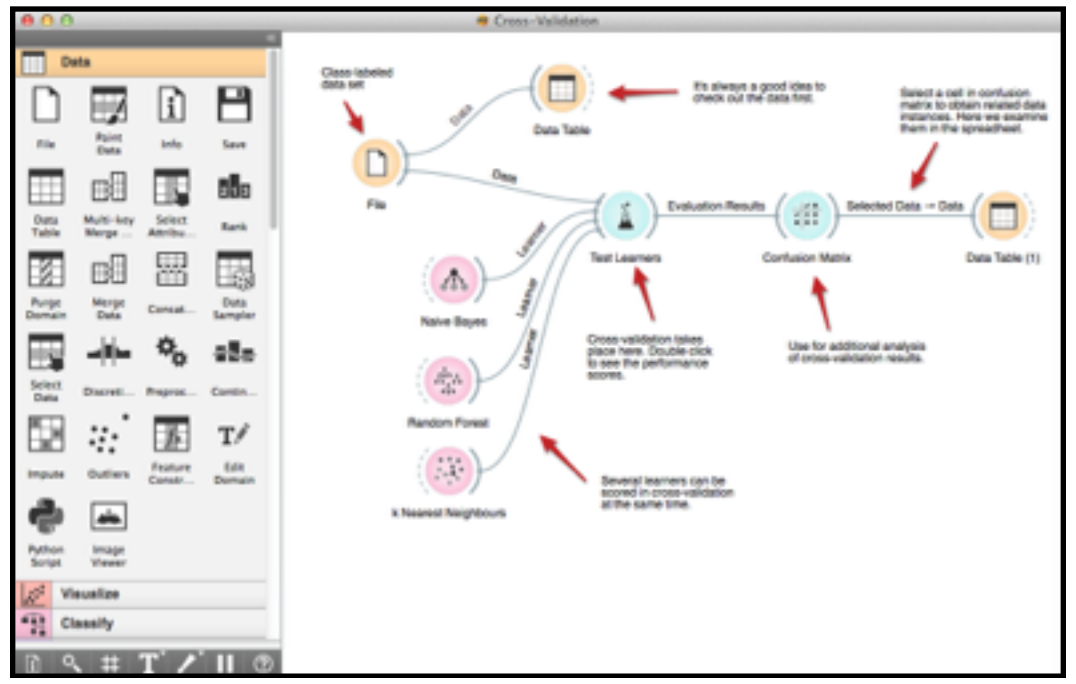
basic

R / Bioconductor





KNIME



Orange

+ coding interface

Generic Analysis Tools

Genome Browsers

complex

ease of use

easy

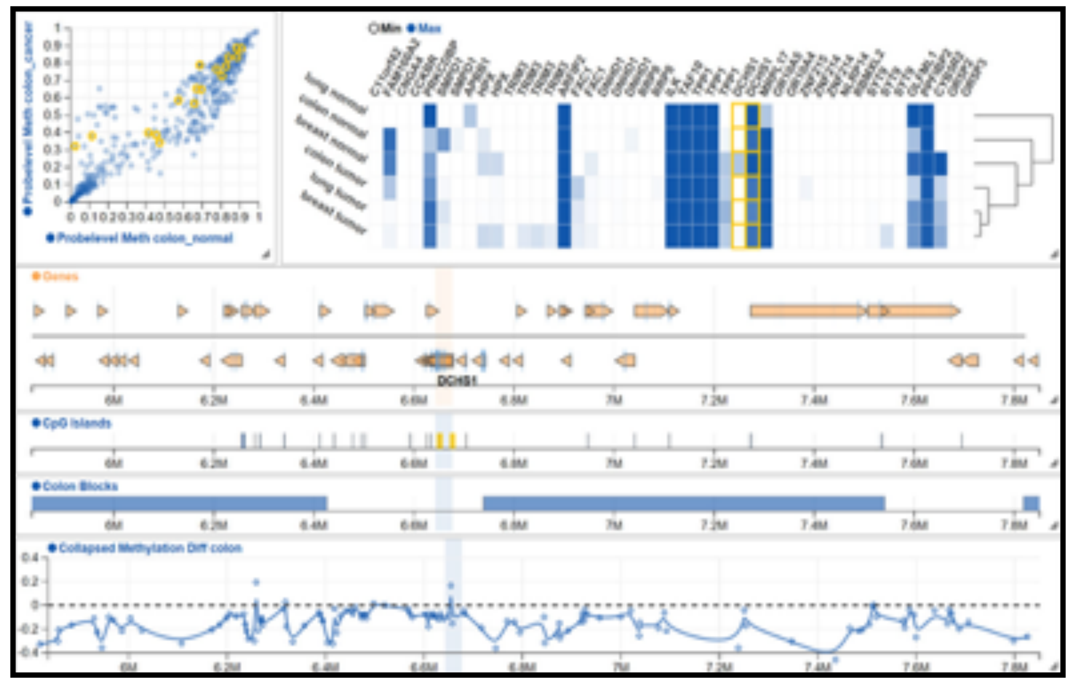
motivation

advanced

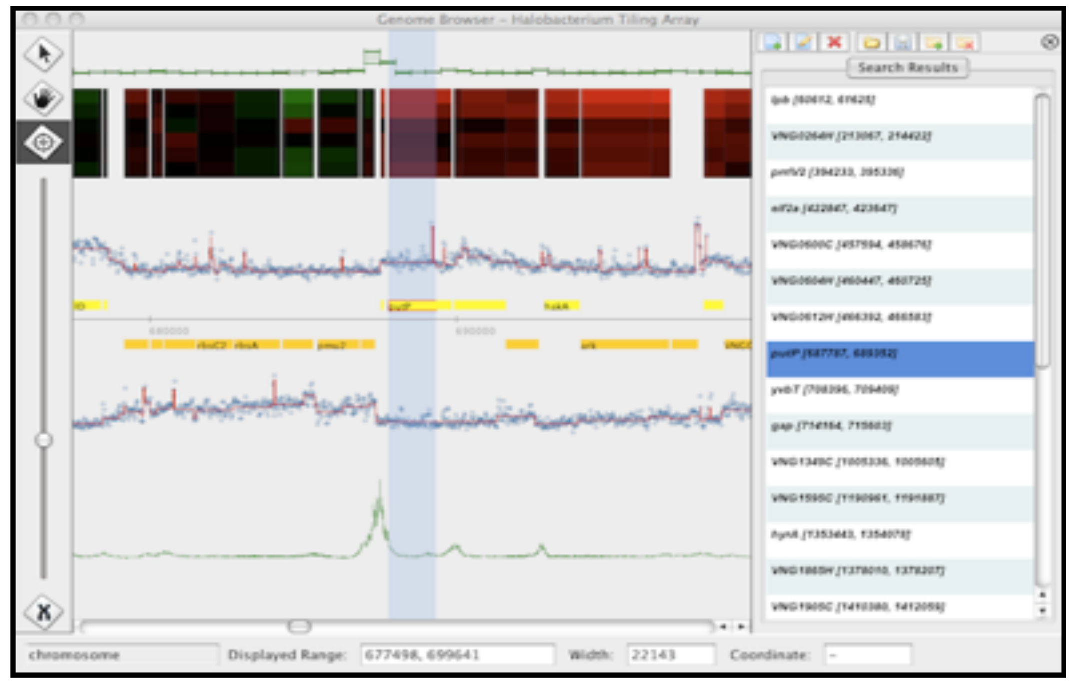
analysis power

basic

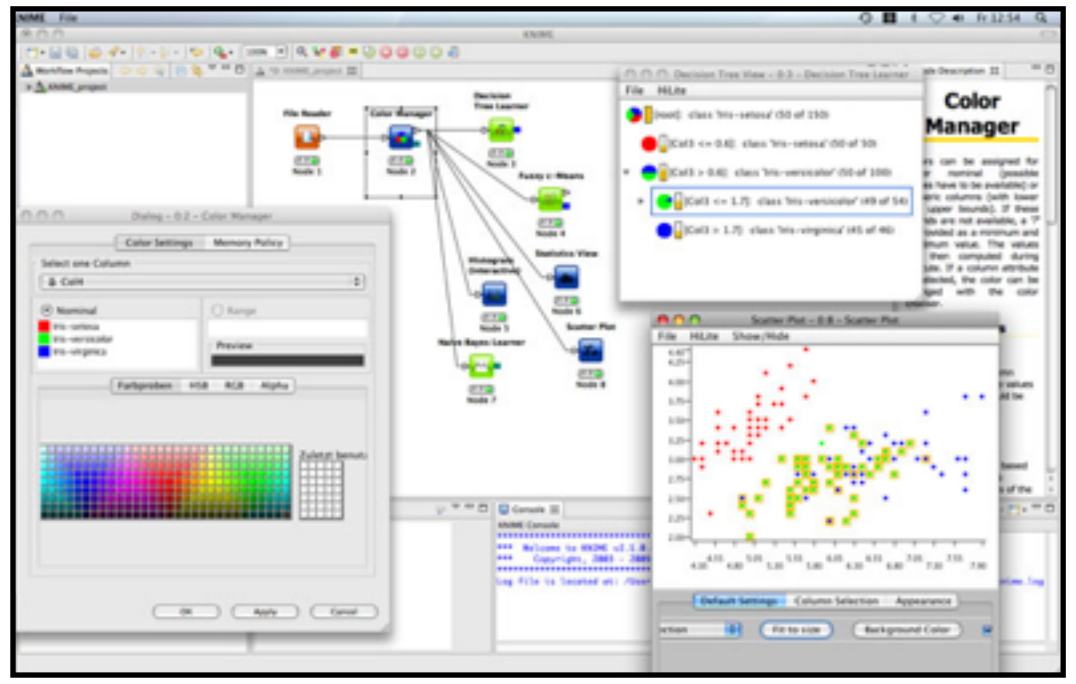
R / Bioconductor



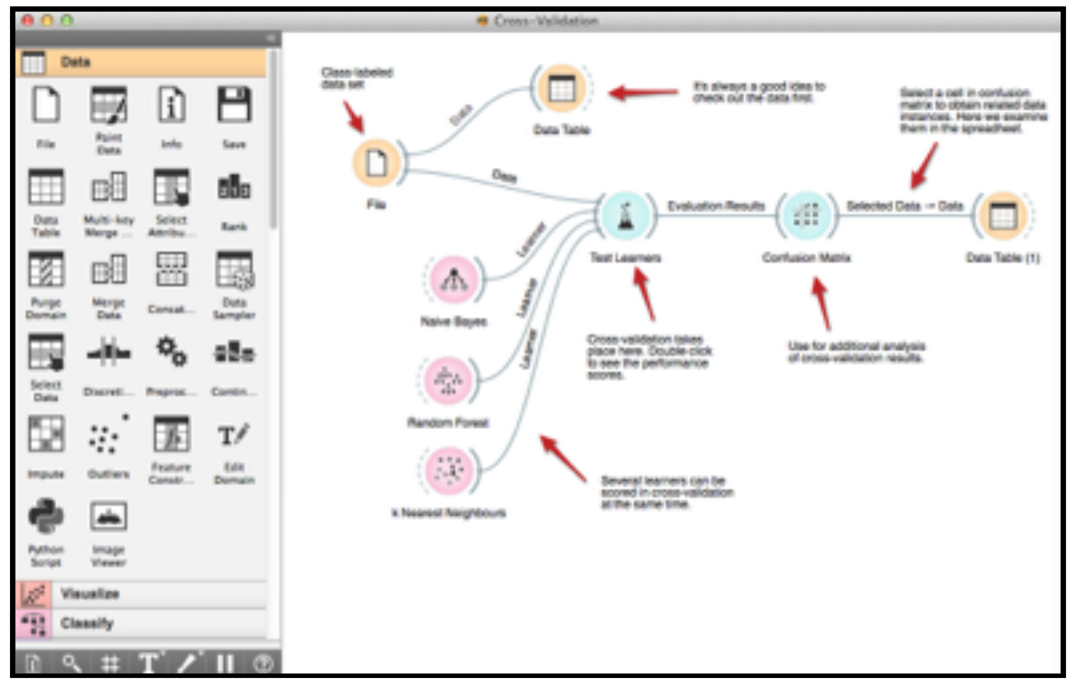
Epiviz



Gaggle Genome Browser



KNIME



Orange

+ coding interface

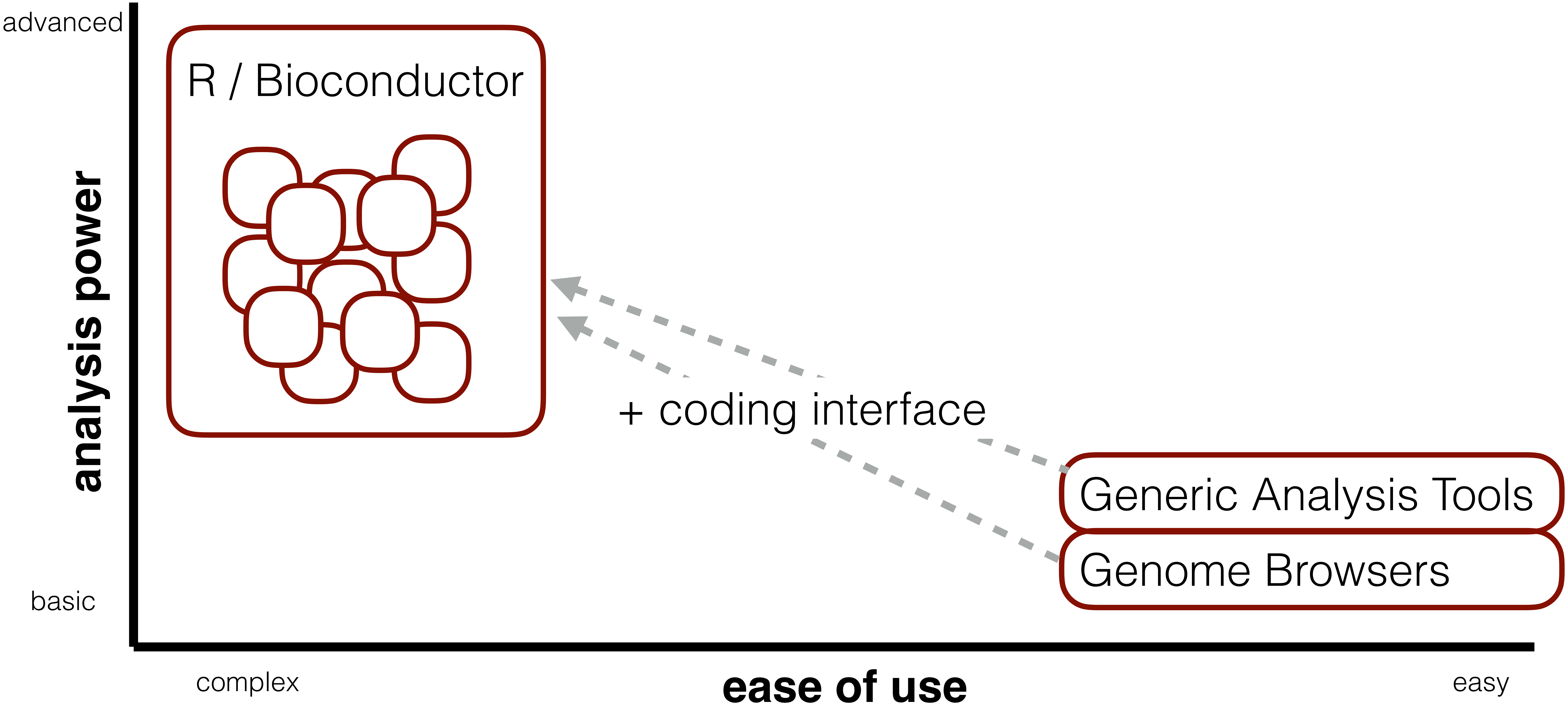
Generic Analysis Tools

Genome Browsers

ease of use

easy

motivation



advanced

analysis power

basic

R / Bioconductor

+ coding interface

Generic Analysis Tools

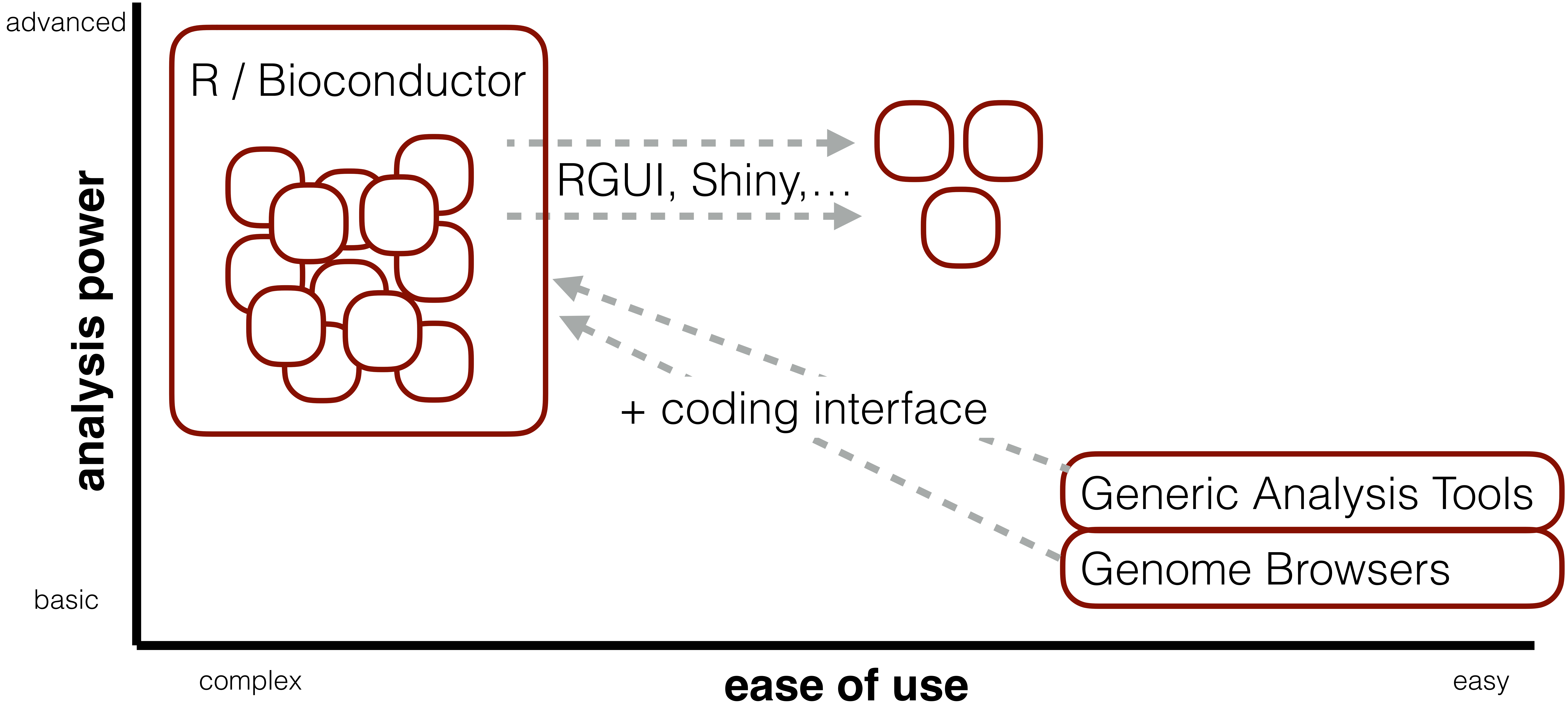
Genome Browsers

complex

ease of use

easy

motivation



advanced

basic

analysis power

R / Bioconductor

RGUI, Shiny,…

+ coding interface

Generic Analysis Tools

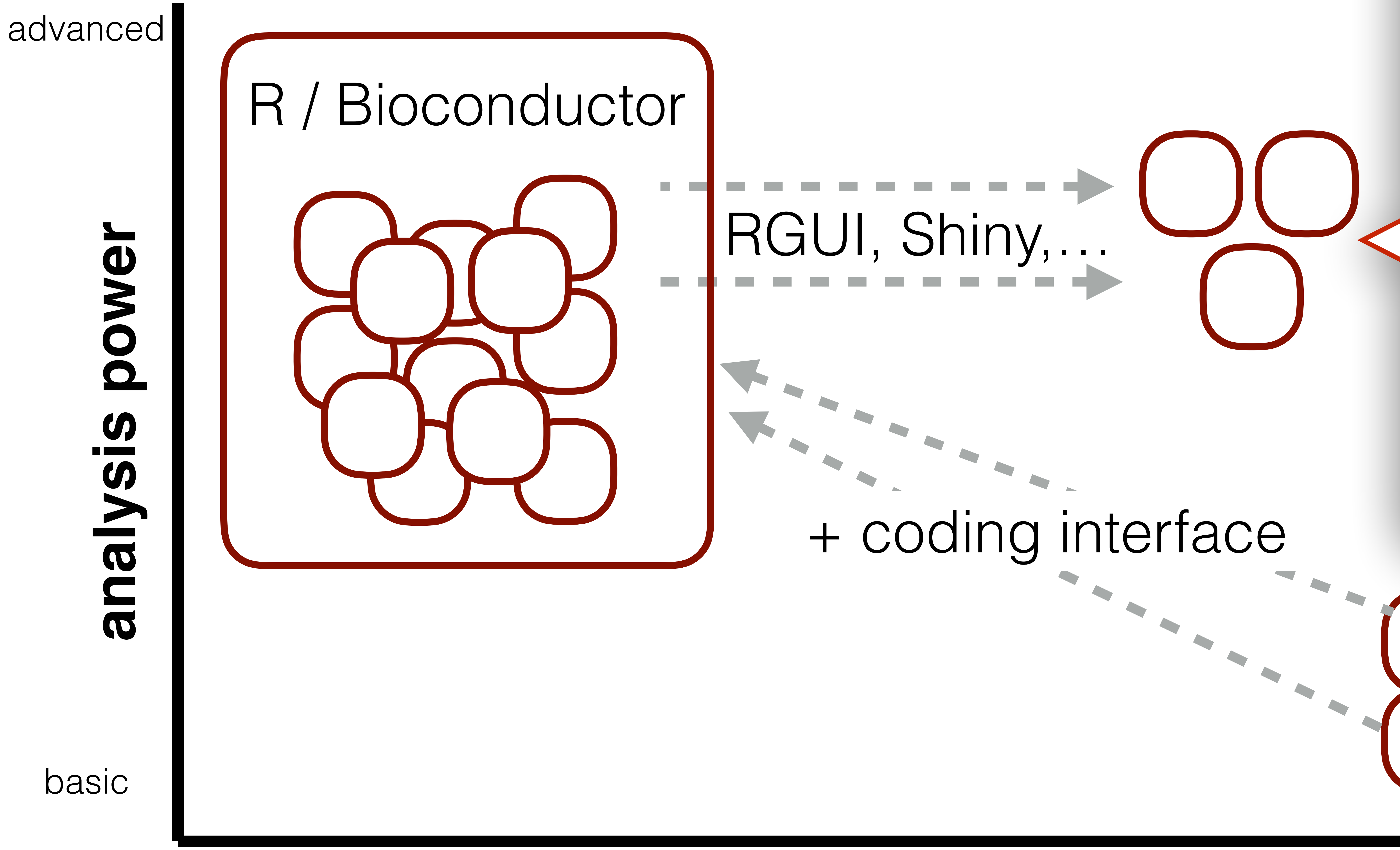
Genome Browsers

complex

ease of use

easy

motivation



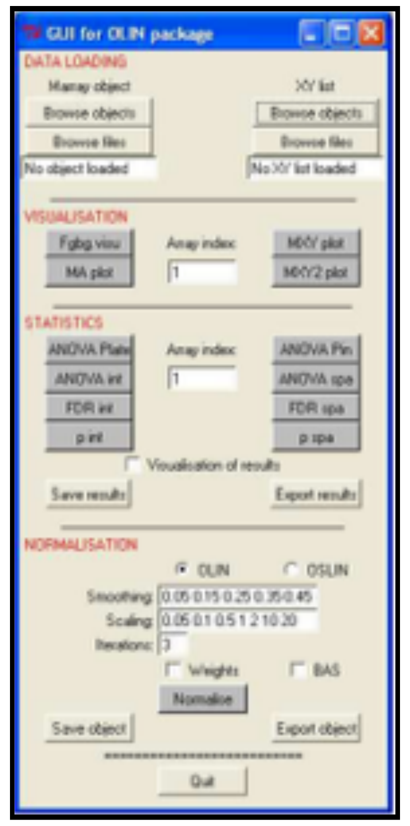
R / Bioconductor

RGUI, Shiny,…

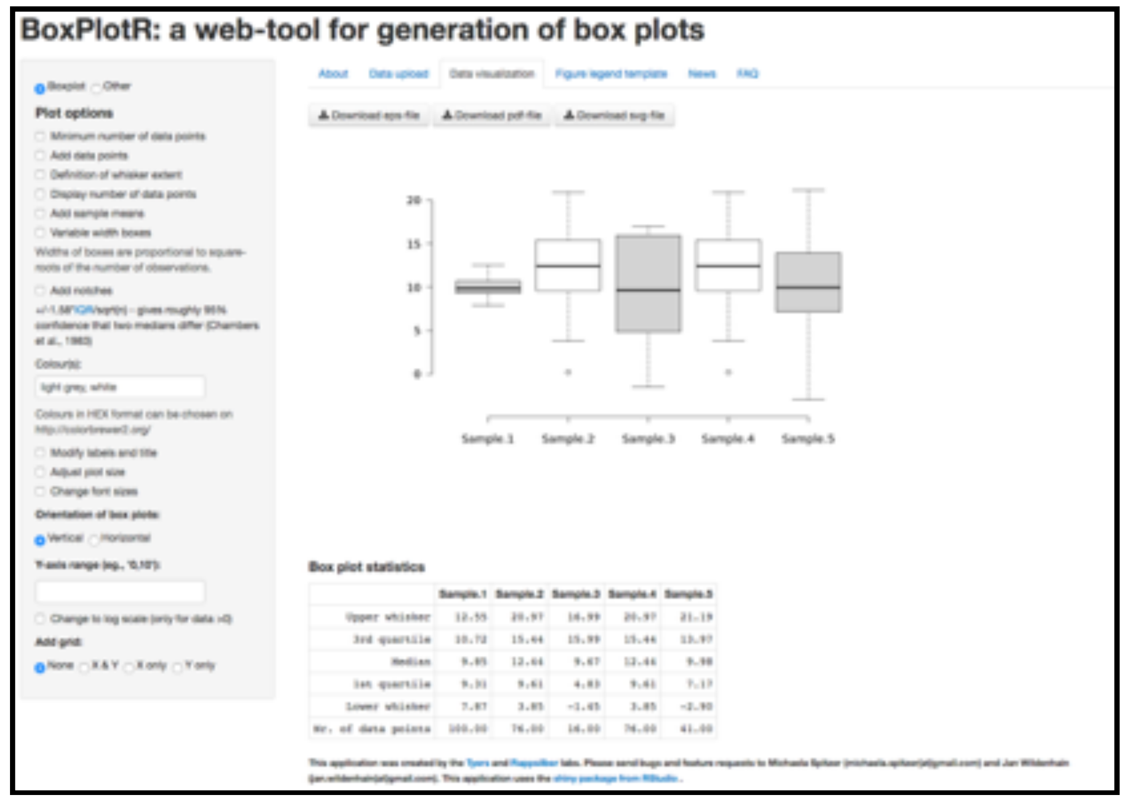
+ coding interface

Generic Analysis Tools

Genome Browsers



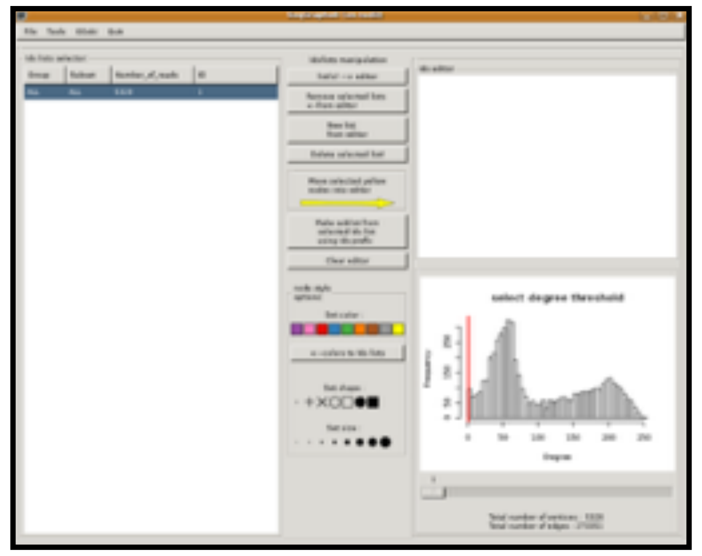
OLINgui



BoxPlotR: a web-tool for generation of box plots

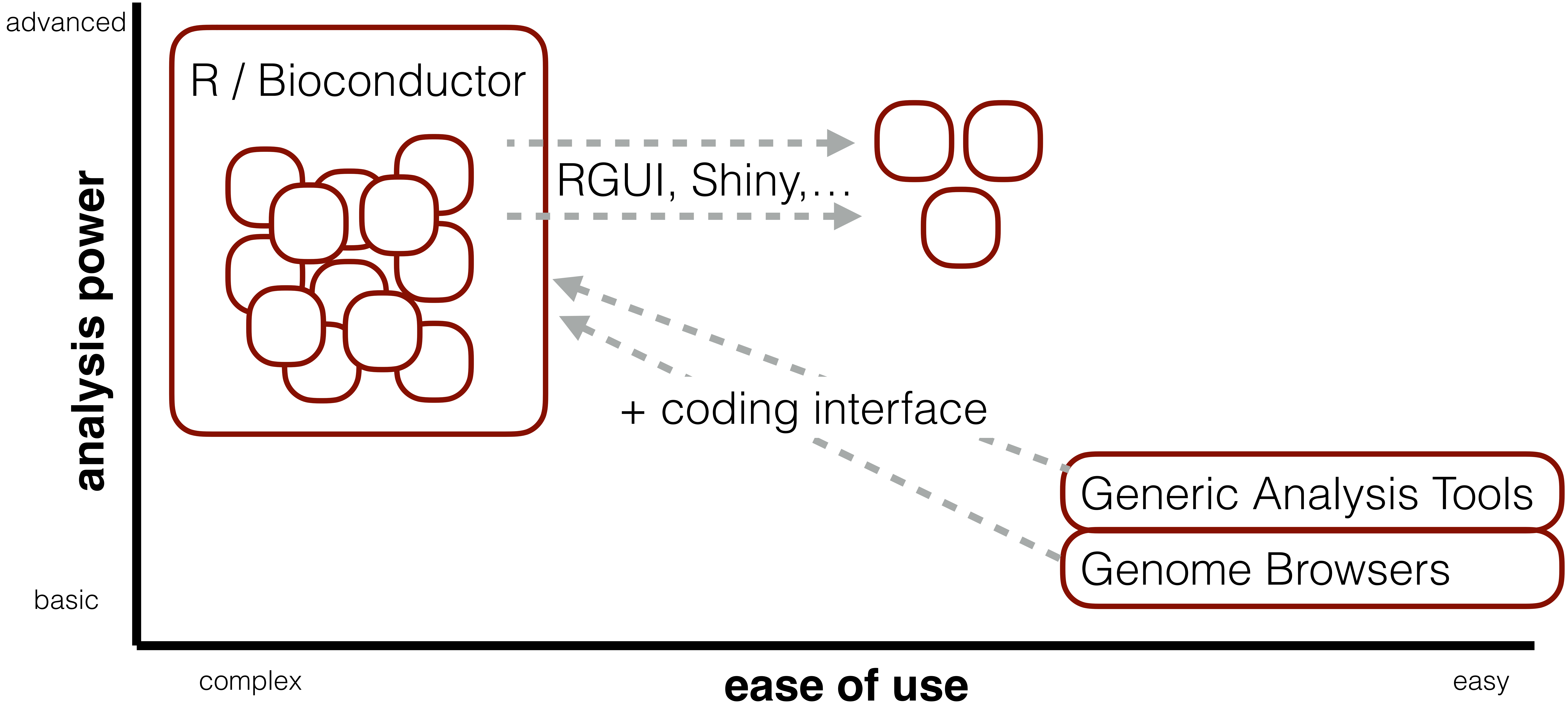
BoxPlotR

LimmaGUI

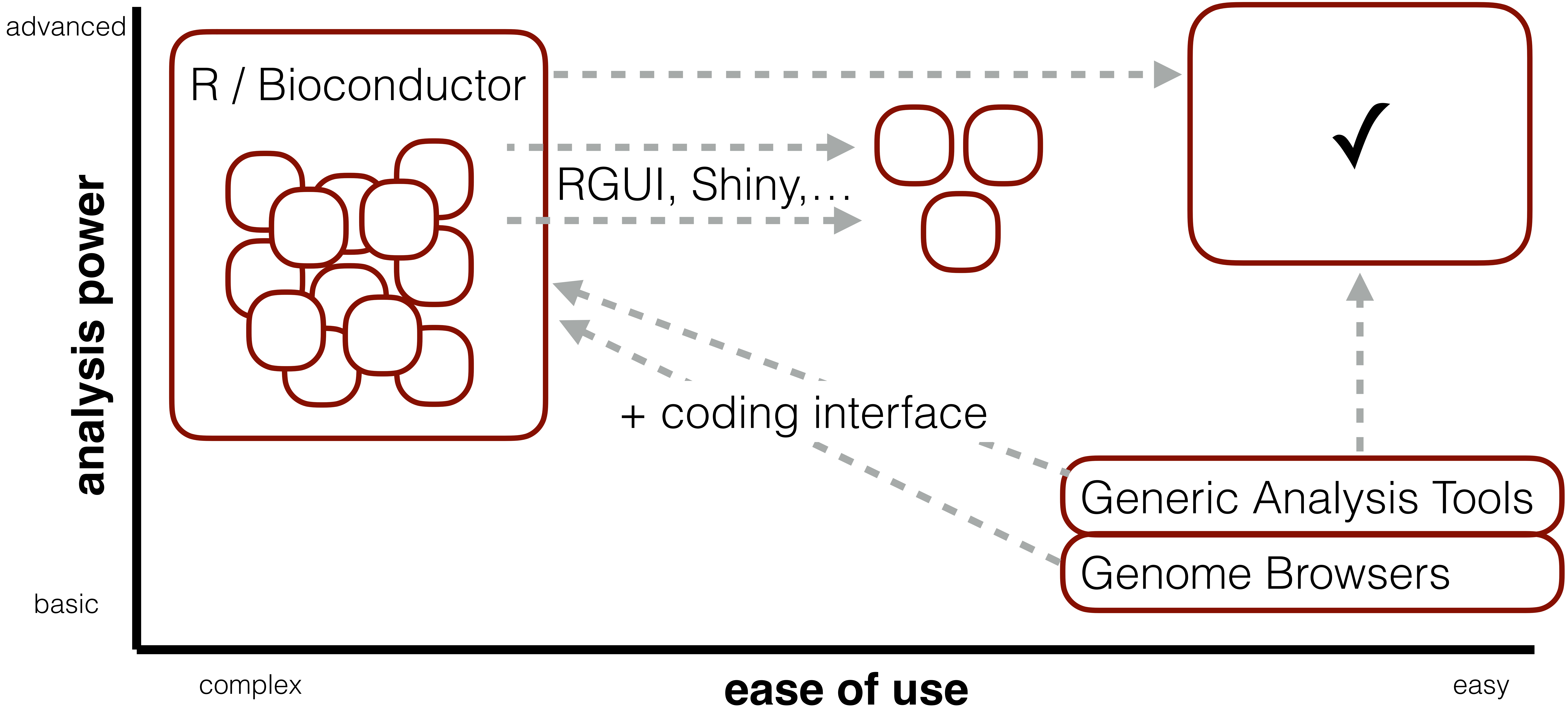


SeqGraphR

motivation

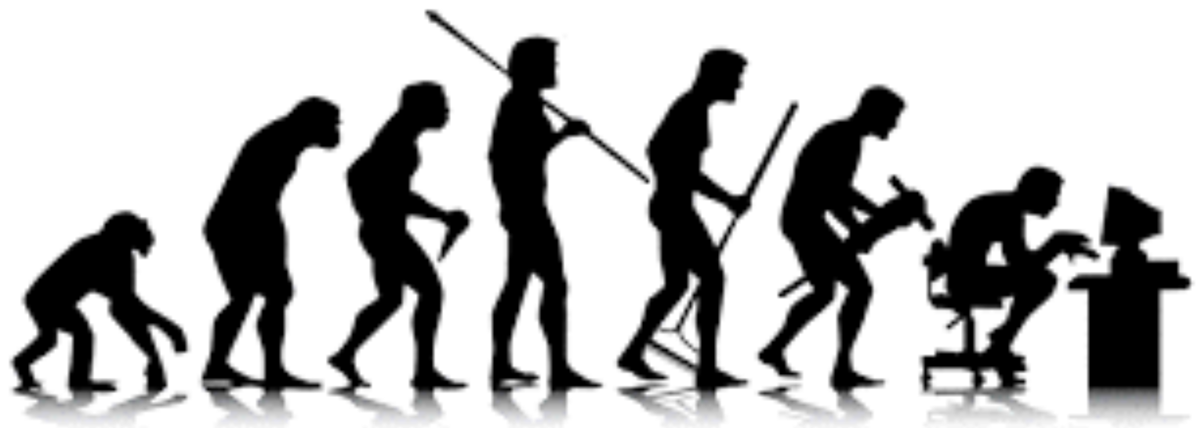

advanced
basic
analysis power
R / Bioconductor
RGUI, Shiny,…
+ coding interface
Generic Analysis Tools
Genome Browsers
complex
ease of use
easy

motivation


advanced
basic
analysis power
complex
ease of use
easy
R / Bioconductor
RGUI, Shiny,…
+ coding interface
✓
Generic Analysis Tools
Genome Browsers

VisRseq

design evolution



IGV plugin (2013)

workspace

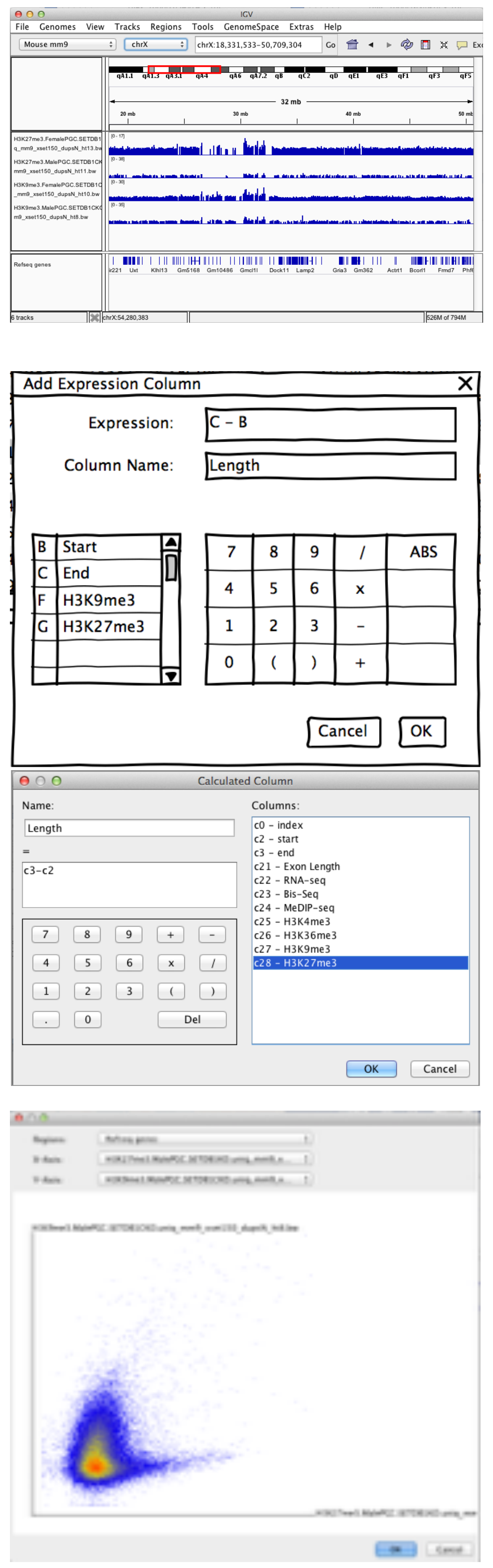
R integration

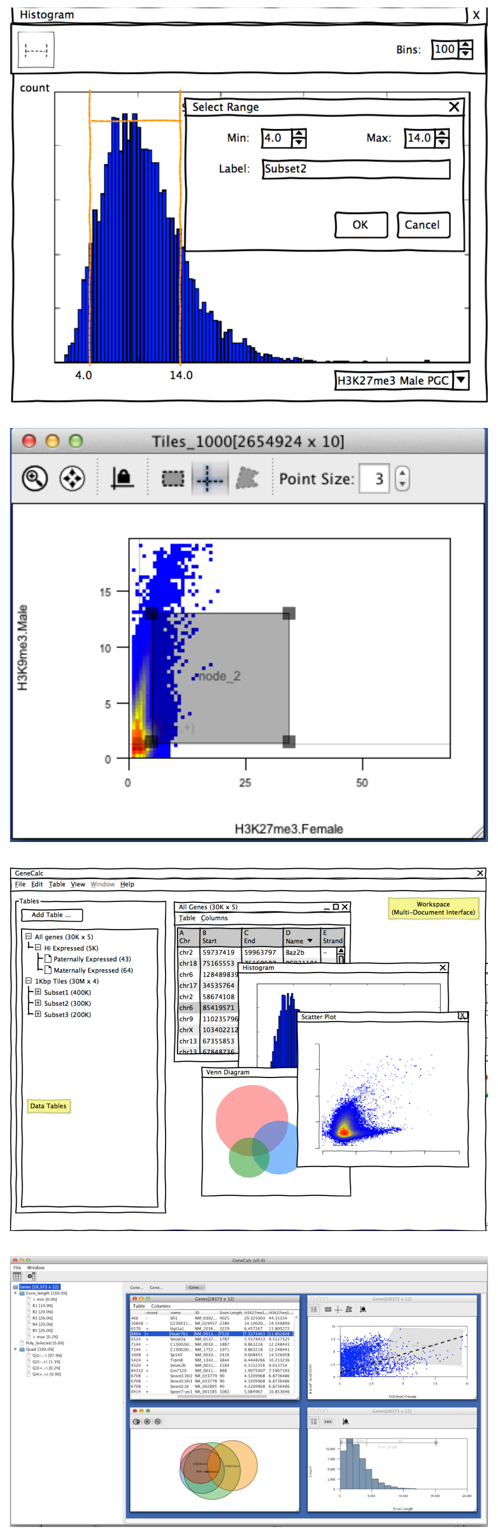
BioVis Poster (2014)

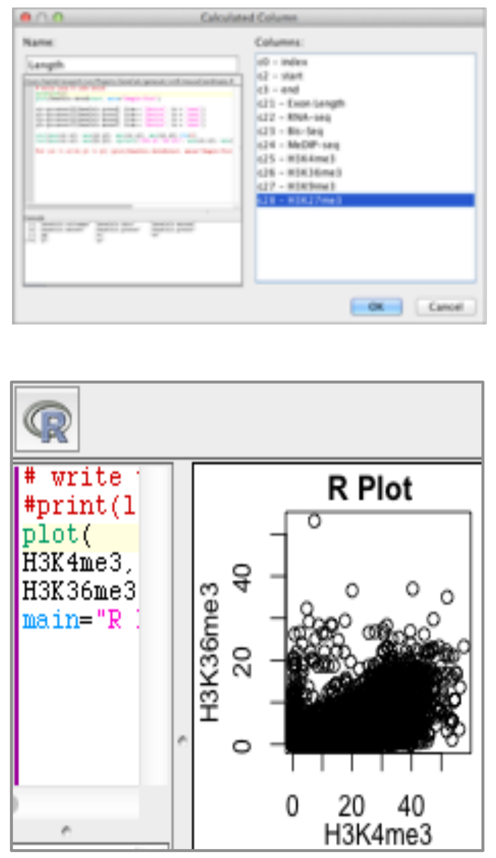
Redesigned Layout

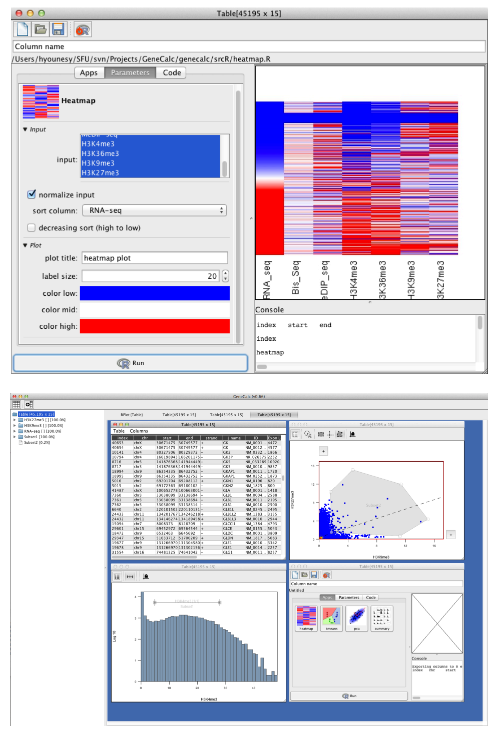
BMC Bioinf (2015)

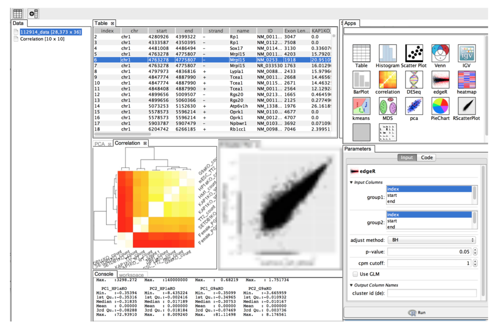
open beta

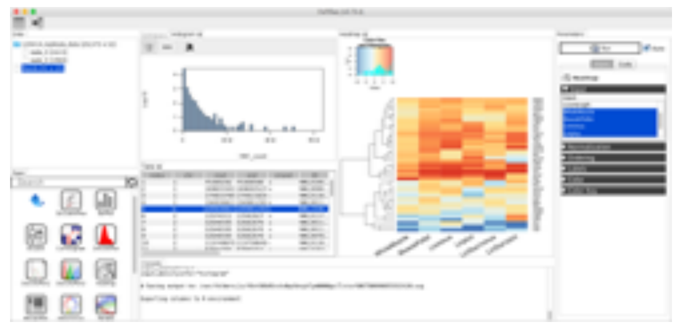






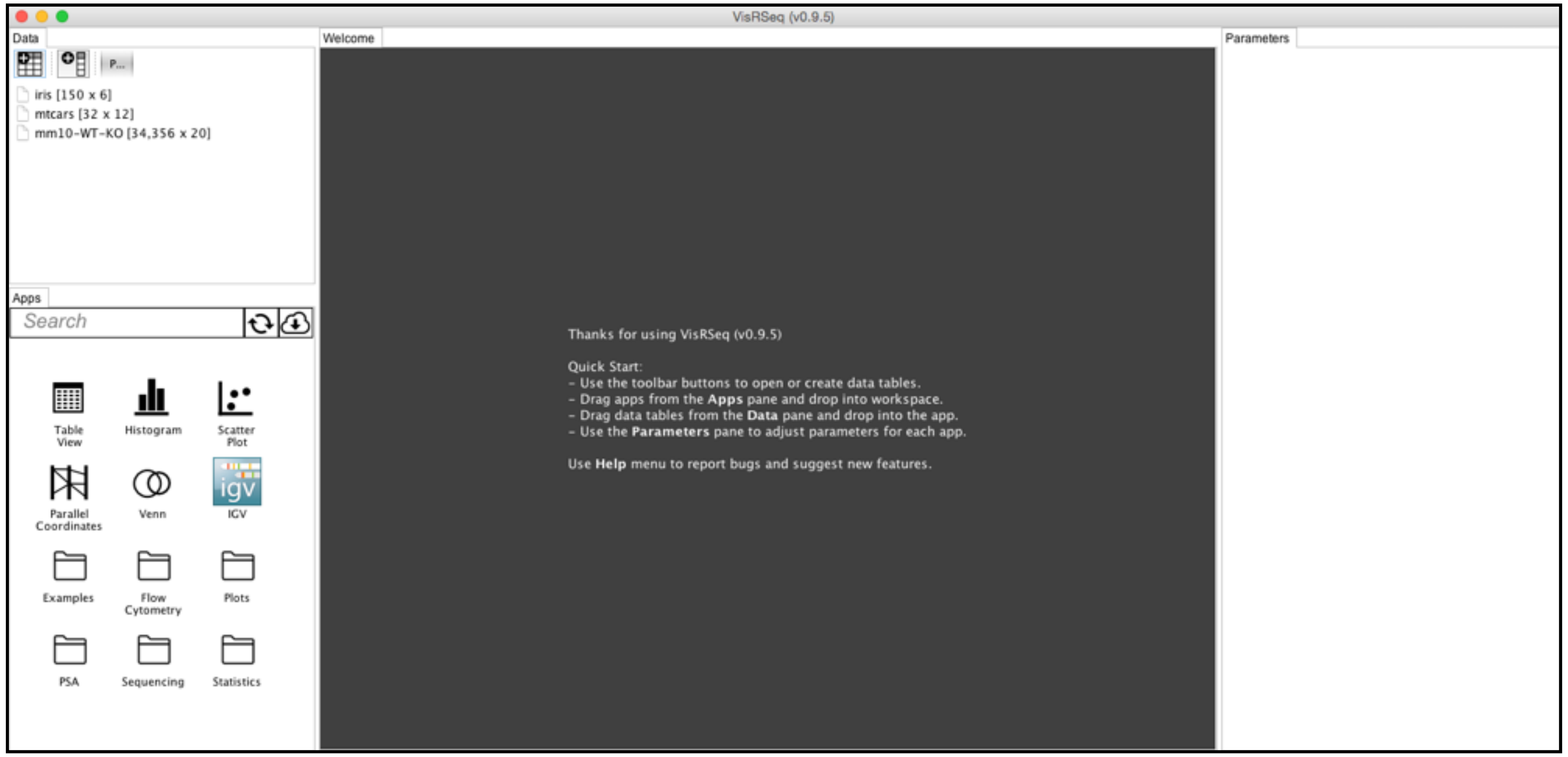






interface

- Installation Requirements:
  - Java SE 7
  - R
- Auto installs required packages
- Open Beta: http://visrseq.github.io

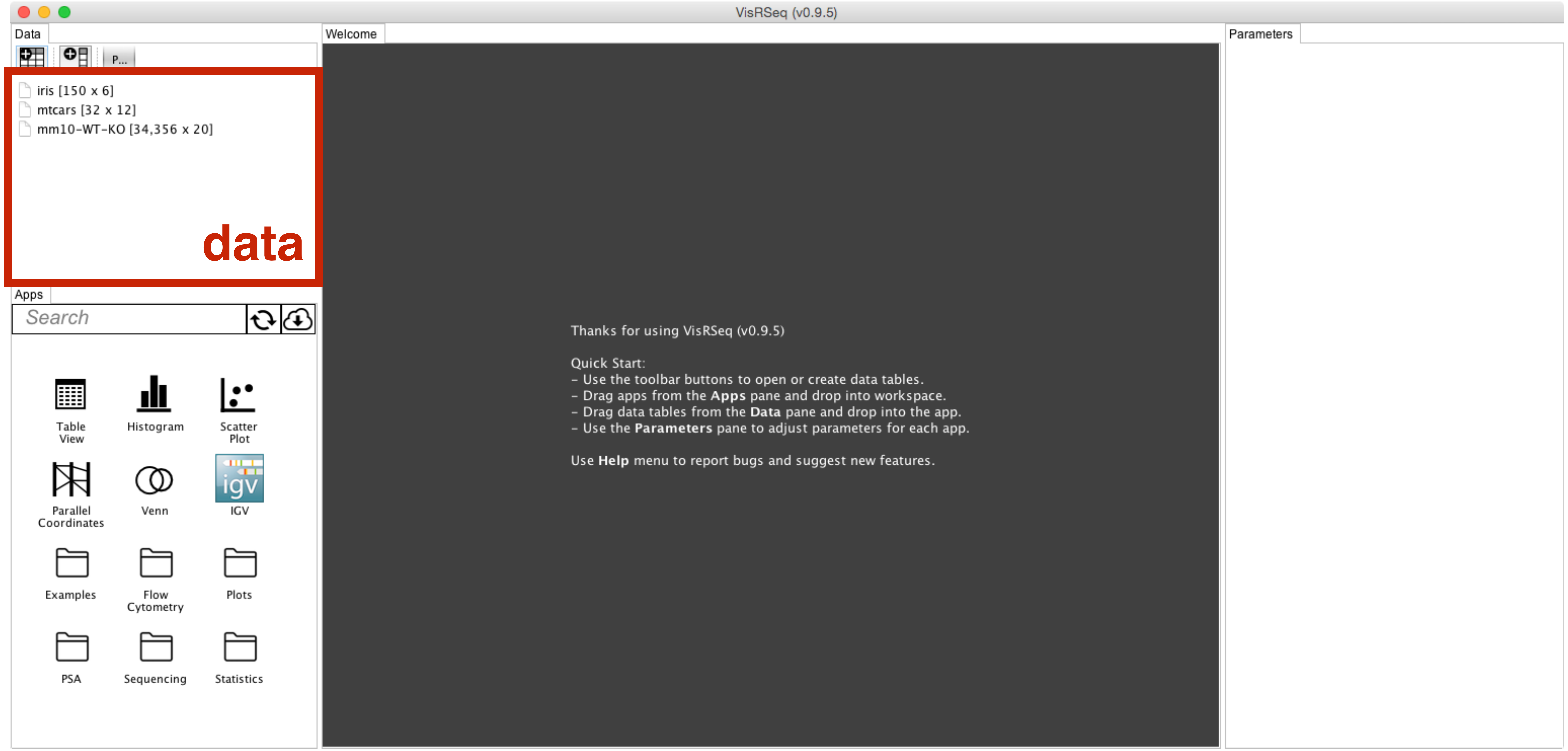


Thanks for using VisRSeq (v0.9.5)

Quick Start:
- Use the toolbar buttons to open or create data tables.
- Drag apps from the **Apps** pane and drop into workspace.
- Drag data tables from the **Data** pane and drop into the app.
- Use the **Parameters** pane to adjust parameters for each app.

Use **Help** menu to report bugs and suggest new features.

interface: data



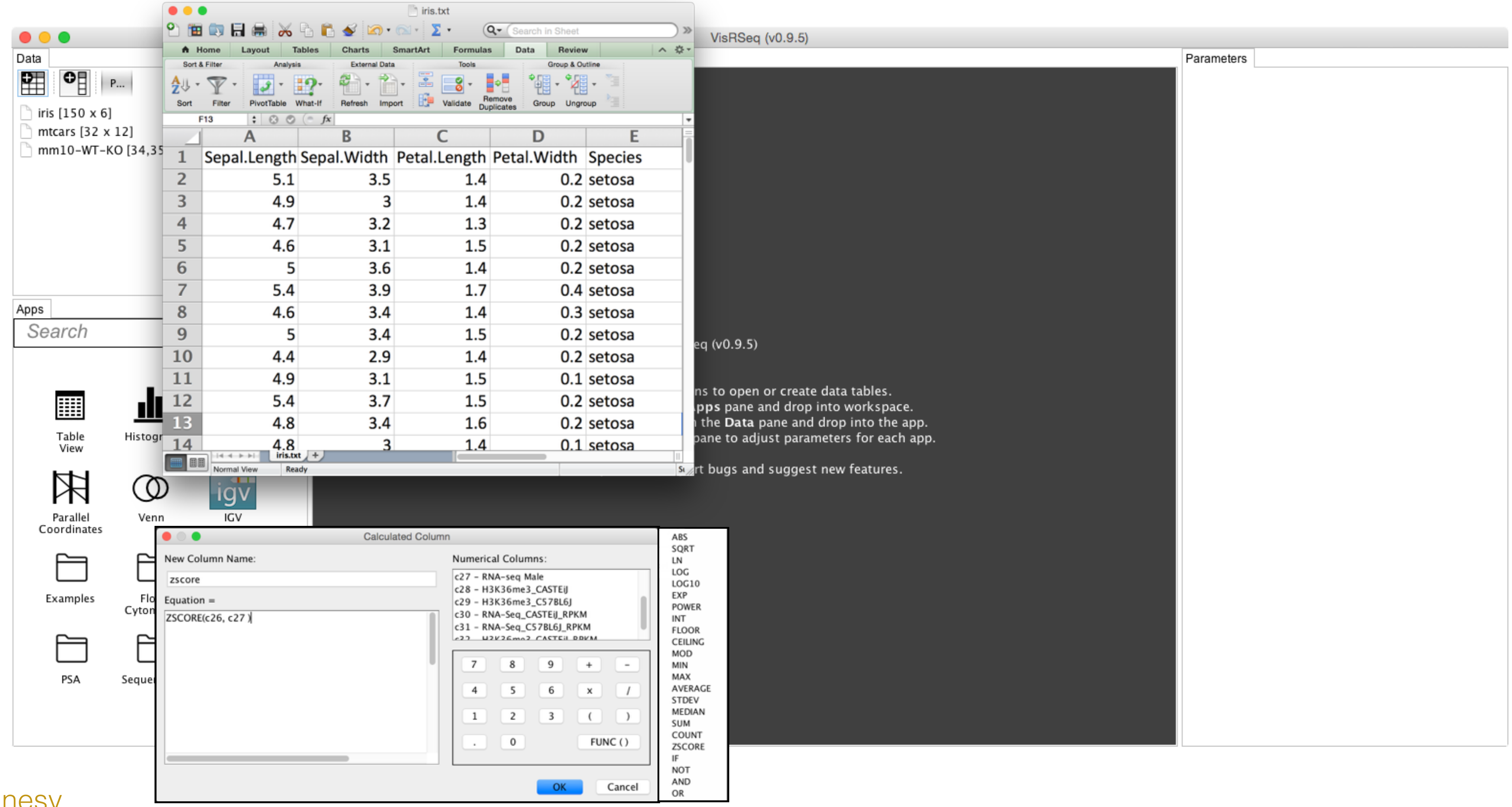
Thanks for using VisRSeq (v0.9.5)

Quick Start:
- Use the toolbar buttons to open or create data tables.
- Drag apps from the **Apps** pane and drop into workspace.
- Drag data tables from the **Data** pane and drop into the app.
- Use the **Parameters** pane to adjust parameters for each app.

Use **Help** menu to report bugs and suggest new features.

interface: data

Tab delimited Text



iris [150 x 6]
mtcars [32 x 12]
mm10-WT-KO [34,35

	A	B	C	D	E
1	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
2	5.1	3.5	1.4	0.2	setosa
3	4.9	3	1.4	0.2	setosa
4	4.7	3.2	1.3	0.2	setosa
5	4.6	3.1	1.5	0.2	setosa
6	5	3.6	1.4	0.2	setosa
7	5.4	3.9	1.7	0.4	setosa
8	4.6	3.4	1.4	0.3	setosa
9	5	3.4	1.5	0.2	setosa
10	4.4	2.9	1.4	0.2	setosa
11	4.9	3.1	1.5	0.1	setosa
12	5.4	3.7	1.5	0.2	setosa
13	4.8	3.4	1.6	0.2	setosa
14	4.8	3	1.4	0.1	setosa

eq (v0.9.5)

ns to open or create data tables.
pps pane and drop into workspace.
the **Data** pane and drop into the app.
pane to adjust parameters for each app.

rt bugs and suggest new features.

New Column Name: zscore

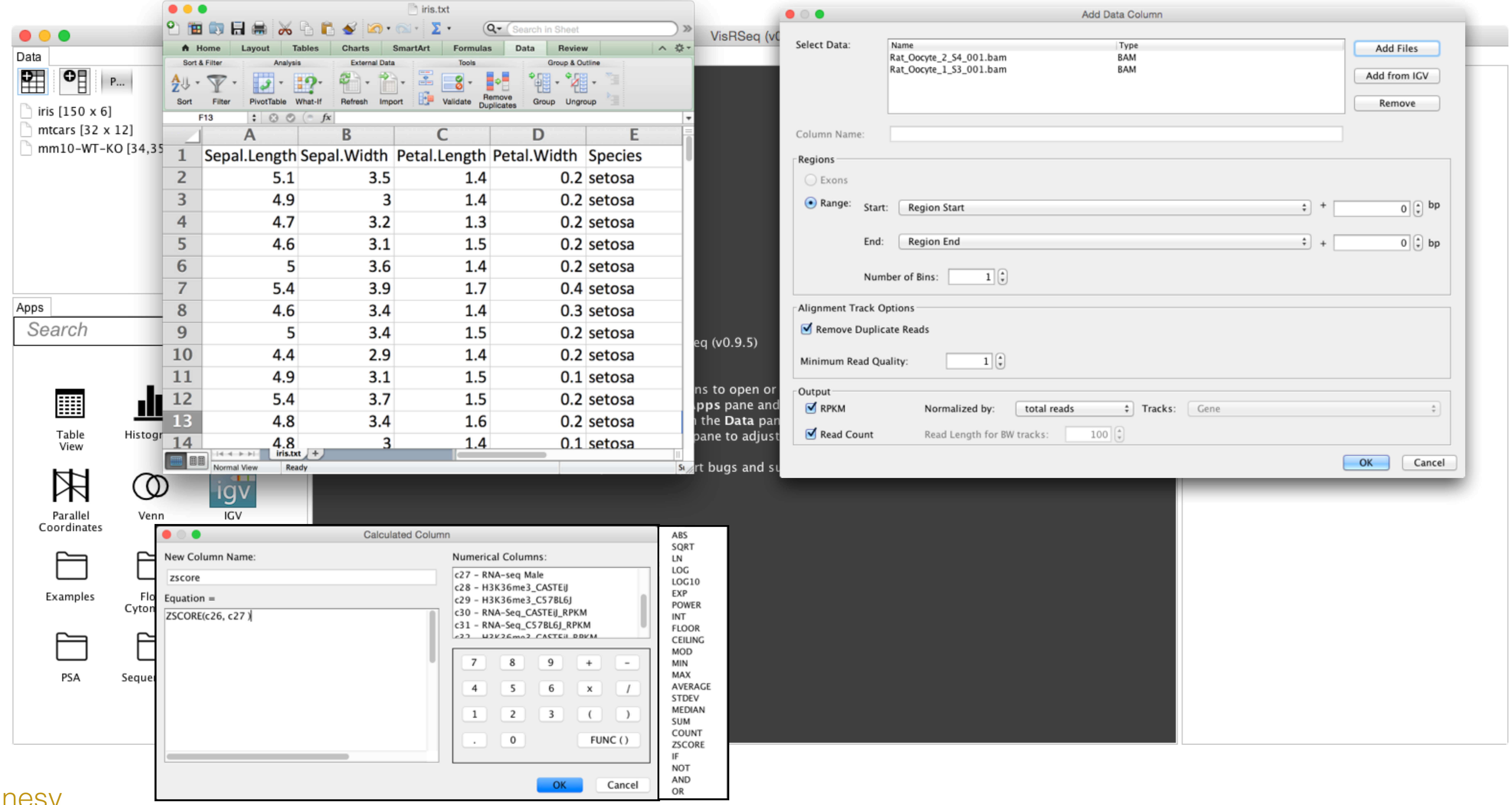
Equation = ZSCORE(c26, c27)

Numerical Columns:
c27 - RNA-seq Male
c28 - H3K36me3_CASTEiJ
c29 - H3K36me3_C57BL6J
c30 - RNA-Seq_CASTEiJ_RPKM
c31 - RNA-Seq_C57BL6J_RPKM

interface: data

Tab delimited Text

.BAM, .WIG



iris [150 x 6]
mtcars [32 x 12]
mm10-WT-KO [34,35

	A	B	C	D	E
1	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
2	5.1	3.5	1.4	0.2	setosa
3	4.9	3	1.4	0.2	setosa
4	4.7	3.2	1.3	0.2	setosa
5	4.6	3.1	1.5	0.2	setosa
6	5	3.6	1.4	0.2	setosa
7	5.4	3.9	1.7	0.4	setosa
8	4.6	3.4	1.4	0.3	setosa
9	5	3.4	1.5	0.2	setosa
10	4.4	2.9	1.4	0.2	setosa
11	4.9	3.1	1.5	0.1	setosa
12	5.4	3.7	1.5	0.2	setosa
13	4.8	3.4	1.6	0.2	setosa
14	4.8	3	1.4	0.1	setosa

Add Data Column

Select Data:

Name	Type
Rat_Oocyte_2_S4_001.bam	BAM
Rat_Oocyte_1_S3_001.bam	BAM

Column Name:

Regions
○ Exons
● Range: Start: Region Start + 0 bp
End: Region End + 0 bp
Number of Bins: 1

Alignment Track Options
☑ Remove Duplicate Reads
Minimum Read Quality: 1

Output
☑ RPKM Normalized by: total reads Tracks: Gene
☑ Read Count Read Length for BW tracks: 100

Calculated Column

New Column Name: zscore

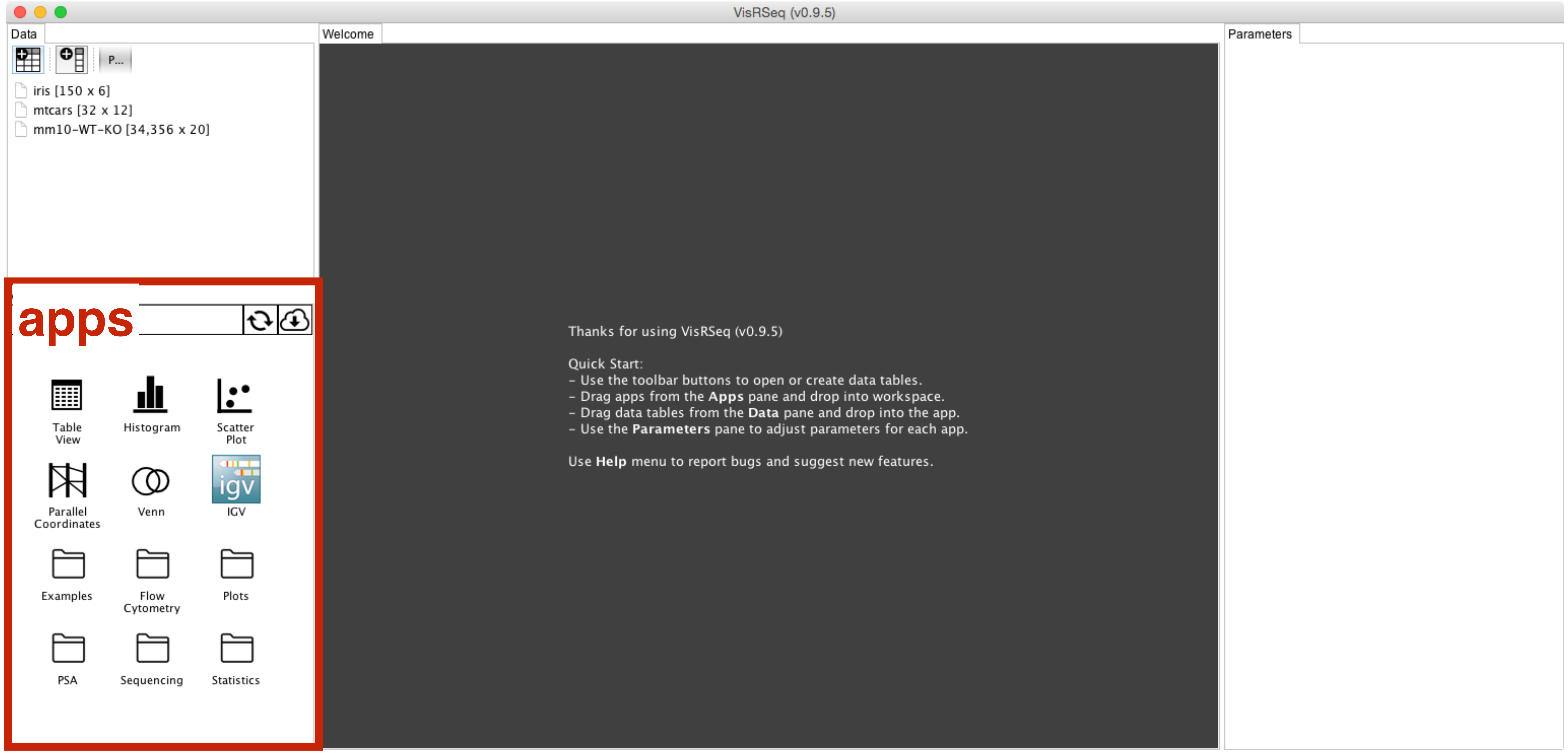
Equation = ZSCORE(c26, c27)

Numerical Columns:
c27 – RNA-seq Male
c28 – H3K36me3_CASTEiJ
c29 – H3K36me3_C57BL6J
c30 – RNA-Seq_CASTEiJ_RPKM
c31 – RNA-Seq_C57BL6J_RPKM

7 8 9 + -
4 5 6 x /
1 2 3 ( )
. 0

ABS
SQRT
LN
LOG
LOG10
EXP
POWER
INT
FLOOR
CEILING
MOD
MIN
MAX
AVERAGE
STDEV
MEDIAN
SUM
COUNT
ZSCORE
IF
NOT
AND
OR

interface: apps

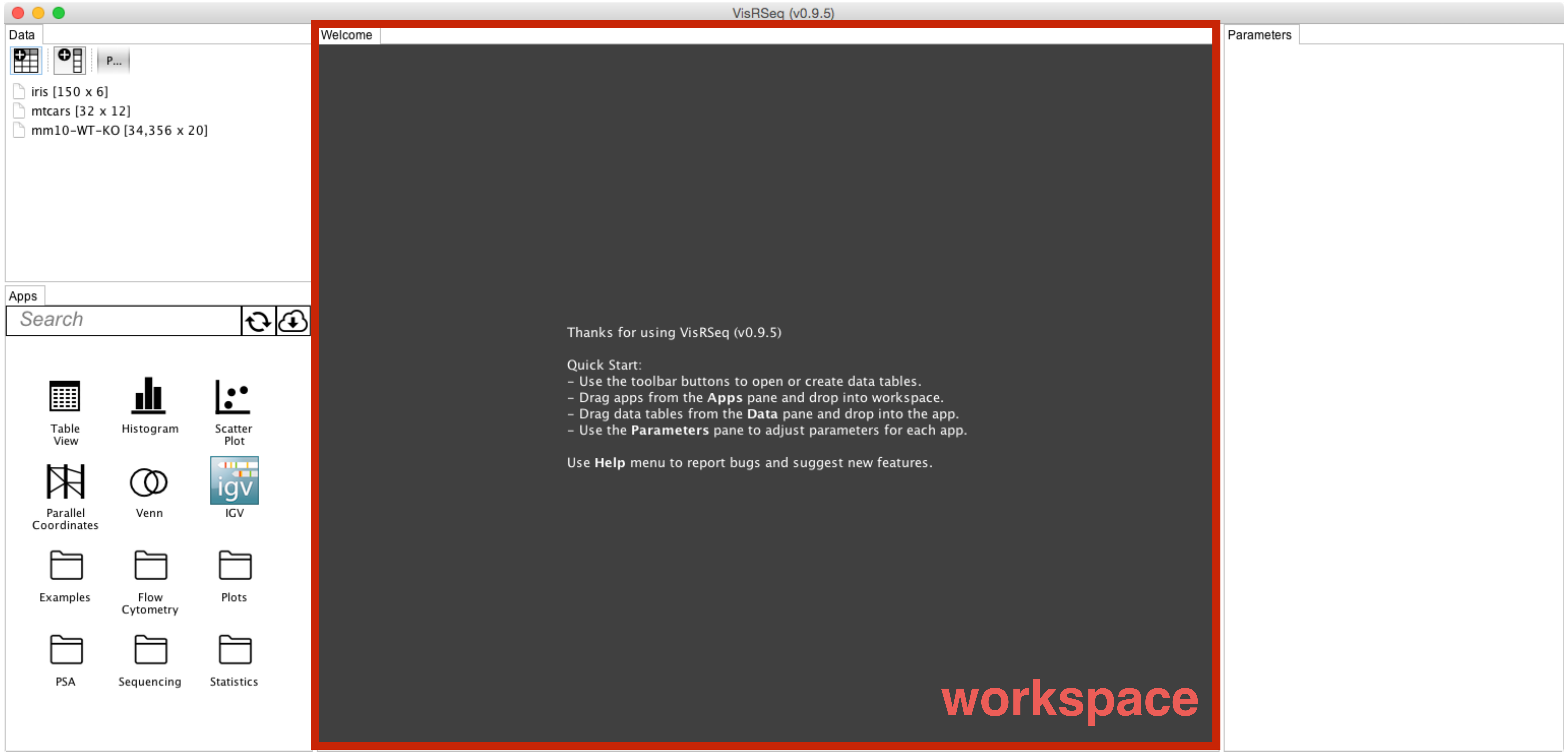


Thanks for using VisRSeq (v0.9.5)

Quick Start:
- Use the toolbar buttons to open or create data tables.
- Drag apps from the **Apps** pane and drop into workspace.
- Drag data tables from the **Data** pane and drop into the app.
- Use the **Parameters** pane to adjust parameters for each app.

Use **Help** menu to report bugs and suggest new features.

interface: workspace

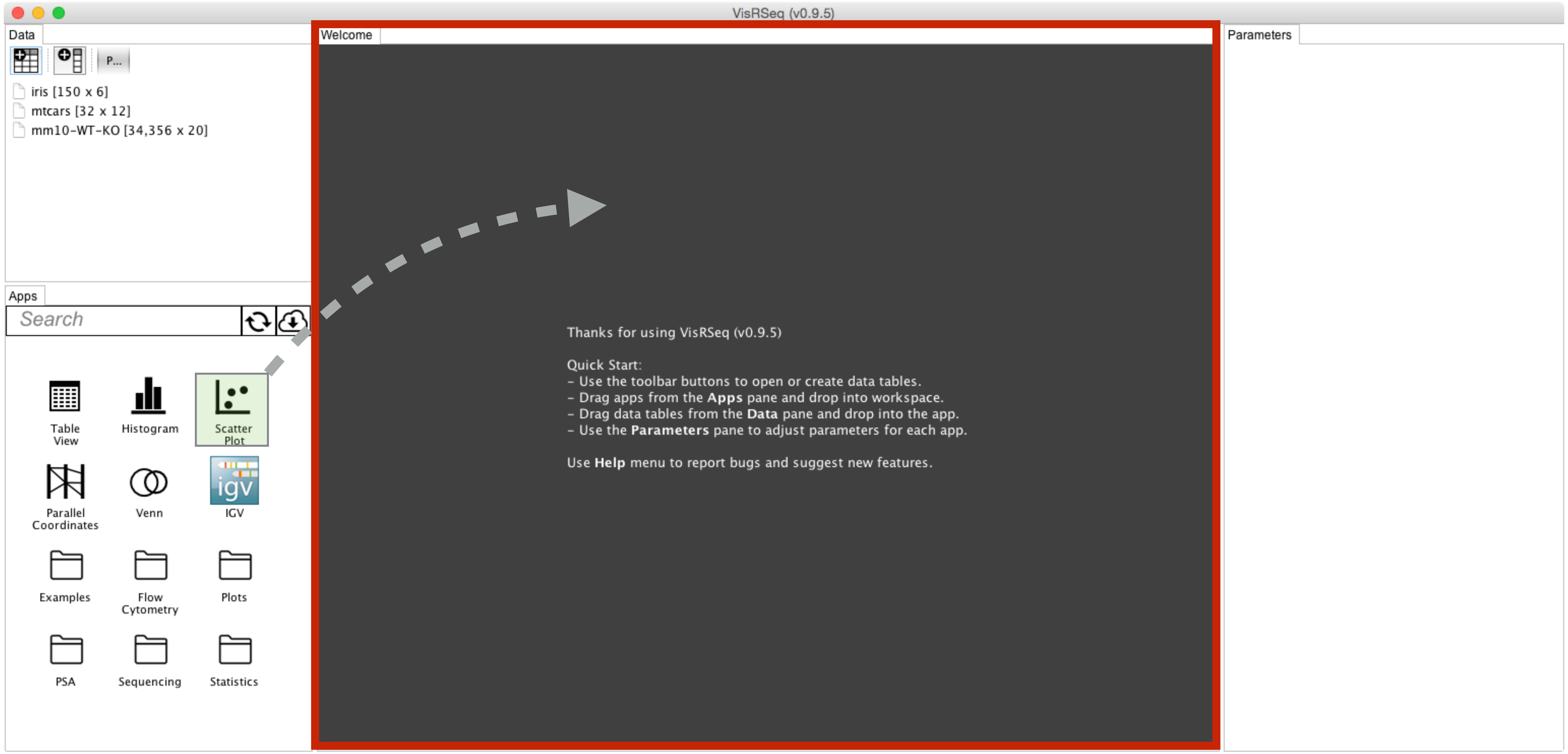


Thanks for using VisRSeq (v0.9.5)

Quick Start:
- Use the toolbar buttons to open or create data tables.
- Drag apps from the **Apps** pane and drop into workspace.
- Drag data tables from the **Data** pane and drop into the app.
- Use the **Parameters** pane to adjust parameters for each app.

Use **Help** menu to report bugs and suggest new features.

interface: workspace



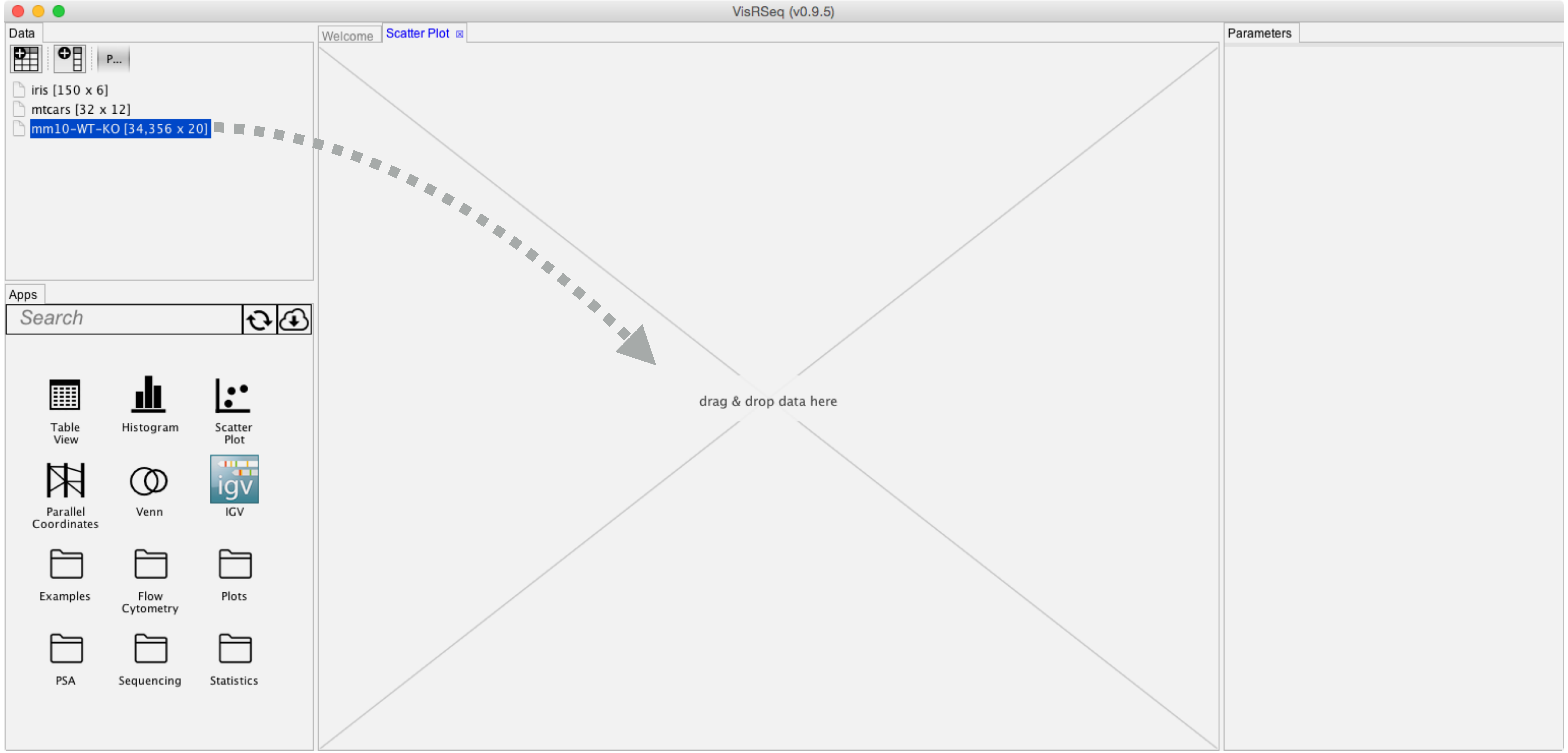
iris [150 x 6]
mtcars [32 x 12]
mm10-WT-KO [34,356 x 20]

Thanks for using VisRSeq (v0.9.5)

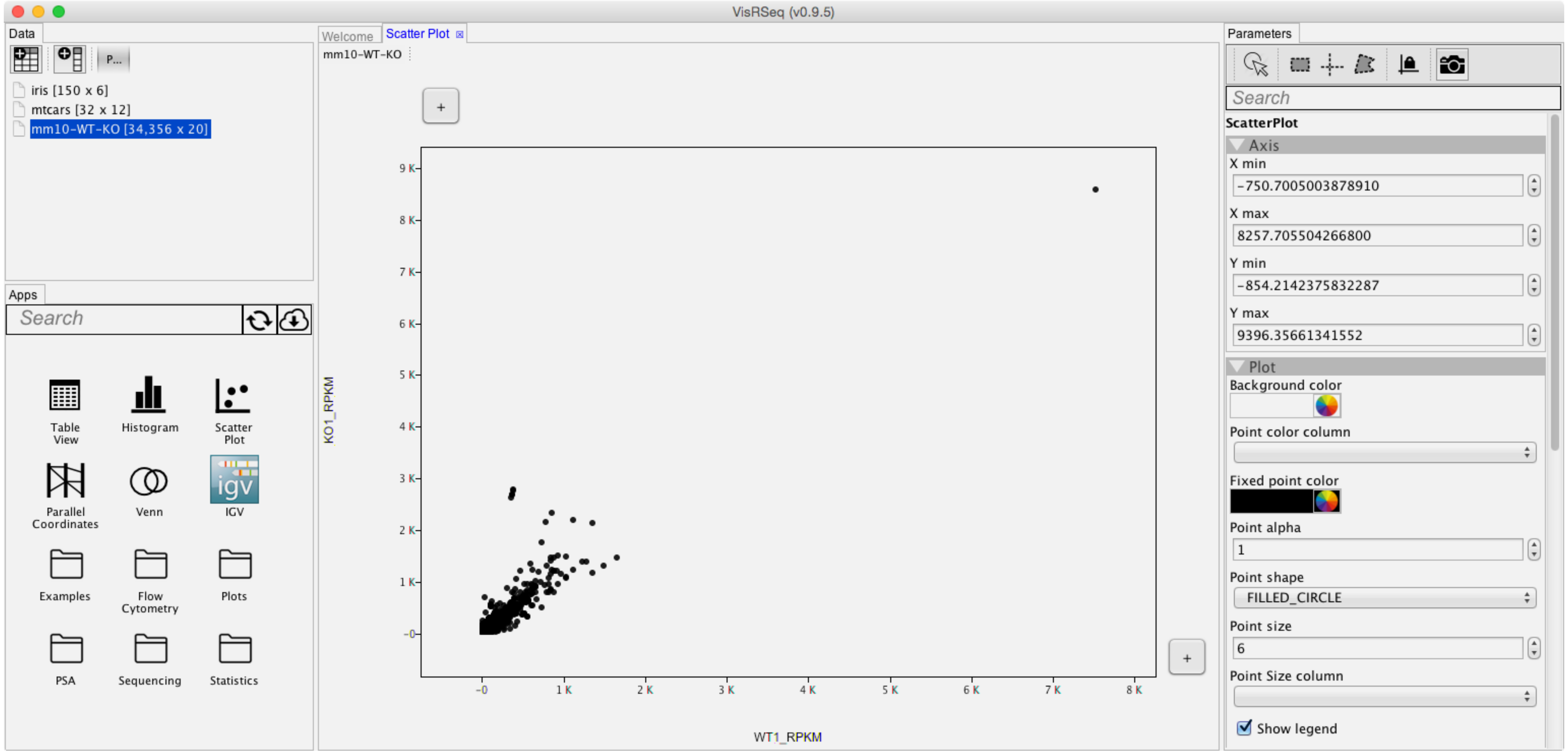
Quick Start:
- Use the toolbar buttons to open or create data tables.
- Drag apps from the **Apps** pane and drop into workspace.
- Drag data tables from the **Data** pane and drop into the app.
- Use the **Parameters** pane to adjust parameters for each app.

Use **Help** menu to report bugs and suggest new features.

interface: workspace



interface: workspace

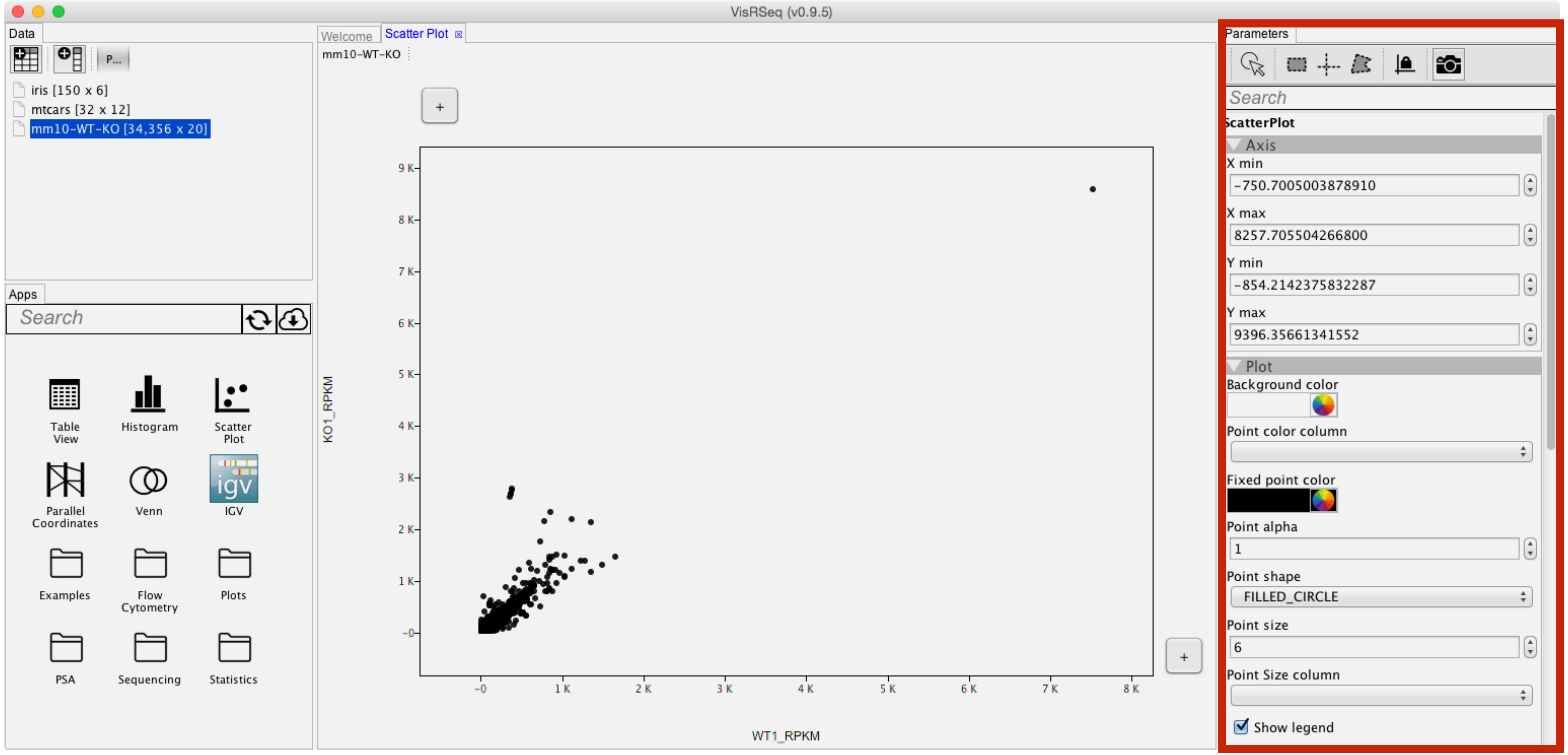


iris [150 x 6]
mtcars [32 x 12]
mm10-WT-KO [34,356 x 20]

X min: -750.7005003878910
X max: 8257.705504266800
Y min: -854.2142375832287
Y max: 9396.35661341552
Background color
Point color column
Fixed point color
Point alpha: 1
Point shape: FILLED_CIRCLE
Point size: 6
Point Size column
☑ Show legend

interface: parameters

parameters



ScatterPlot

Axis

X min: -750.7005003878910

X max: 8257.705504266800

Y min: -854.2142375832287

Y max: 9396.35661341552

Plot

Background color

Point color column

Fixed point color

Point alpha: 1

Point shape: FILLED_CIRCLE

Point size: 6

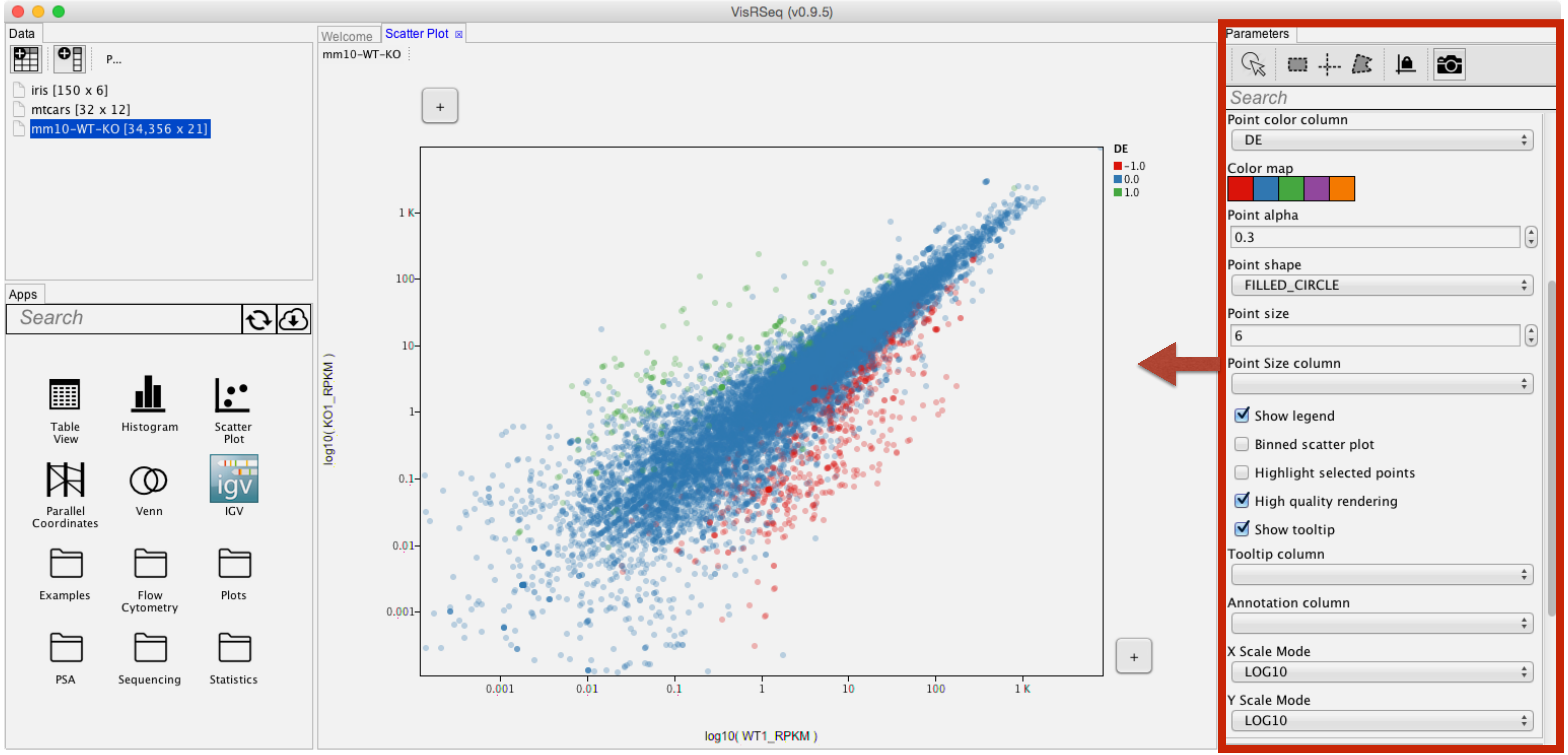
Point Size column

☑ Show legend

Hamid Younesy

interface: parameters

parameters



Point color column: DE

Color map: 🟥 🟦 🟩 🟪 🟧

Point alpha: 0.3

Point shape: FILLED_CIRCLE

Point size: 6

Point Size column:

☑ Show legend

☐ Binned scatter plot

☐ Highlight selected points

☑ High quality rendering

☑ Show tooltip

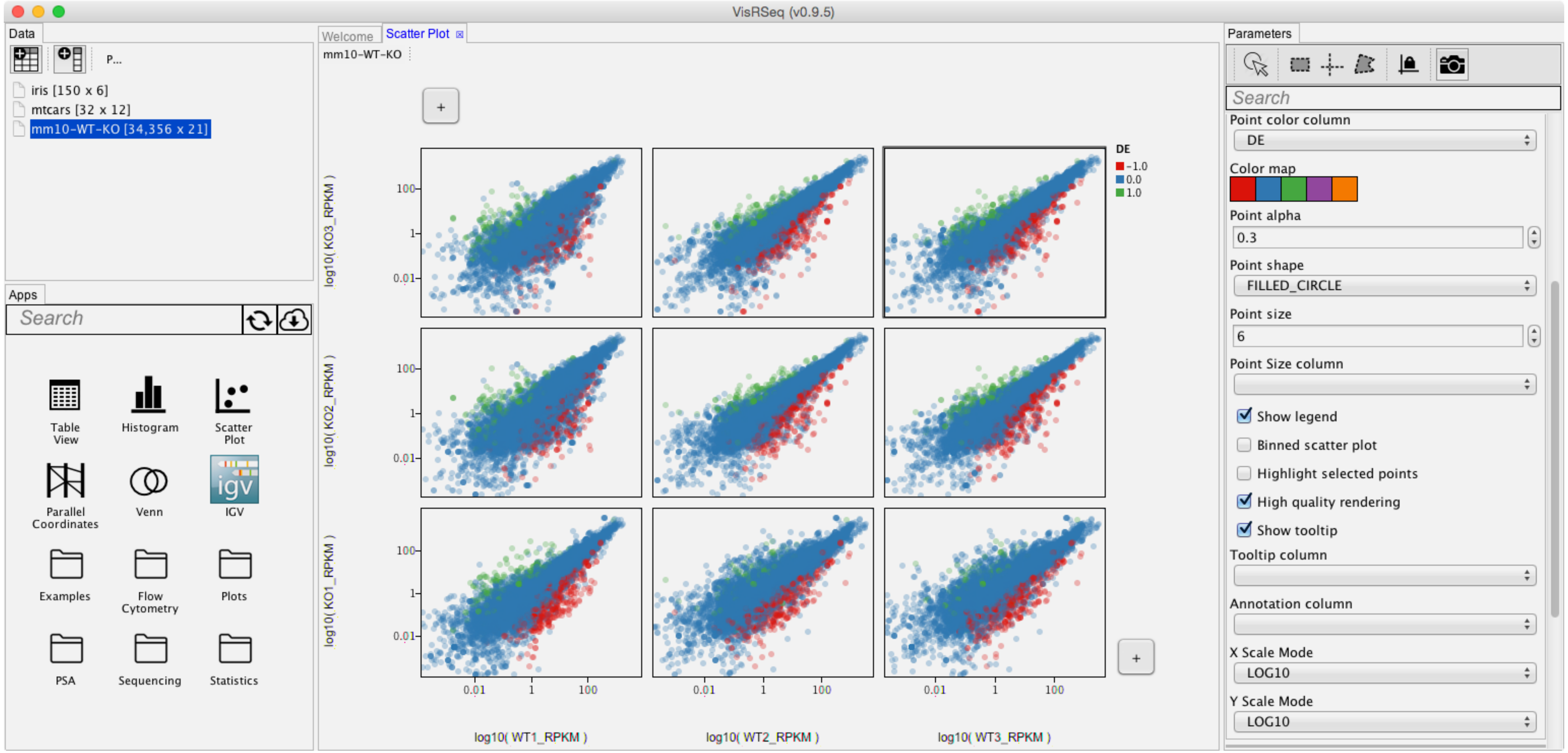
Tooltip column:

Annotation column:

X Scale Mode: LOG10

Y Scale Mode: LOG10

Apps: interactive

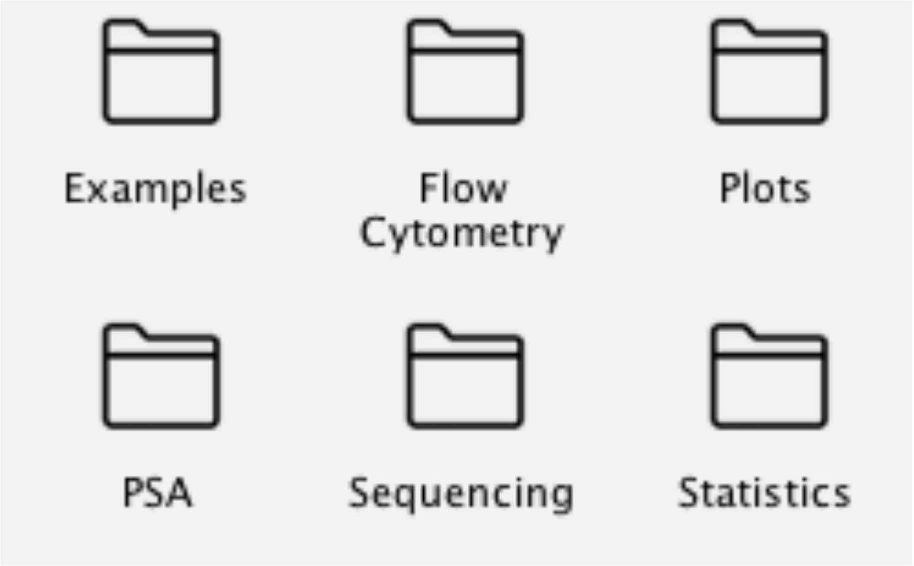


iris [150 x 6]
mtcars [32 x 12]
mm10-WT-KO [34,356 x 21]

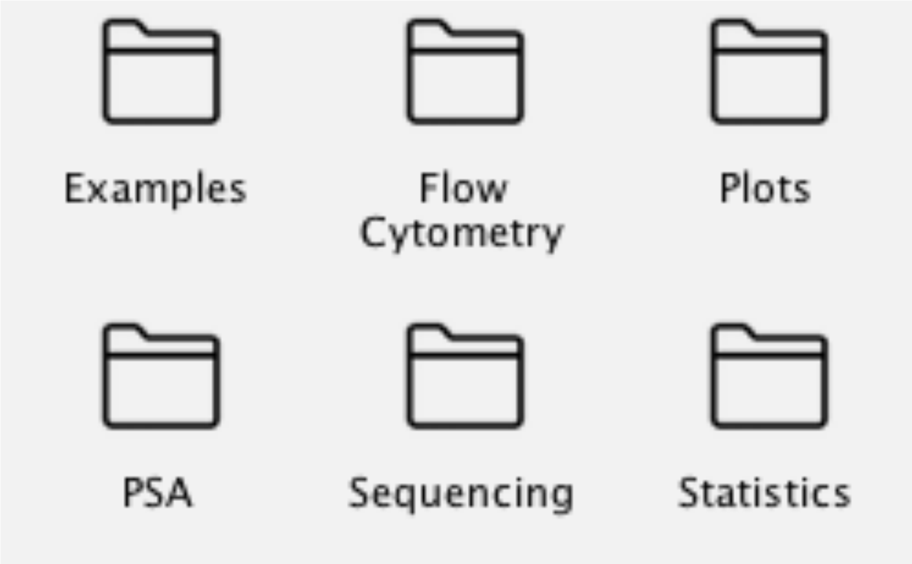
Point color column: DE
Color map: 🟥🟦🟩🟪🟧
Point alpha: 0.3
Point shape: FILLED_CIRCLE
Point size: 6
Point Size column:
☑ Show legend
☐ Binned scatter plot
☐ Highlight selected points
☑ High quality rendering
☑ Show tooltip
Tooltip column:
Annotation column:
X Scale Mode: LOG10
Y Scale Mode: LOG10

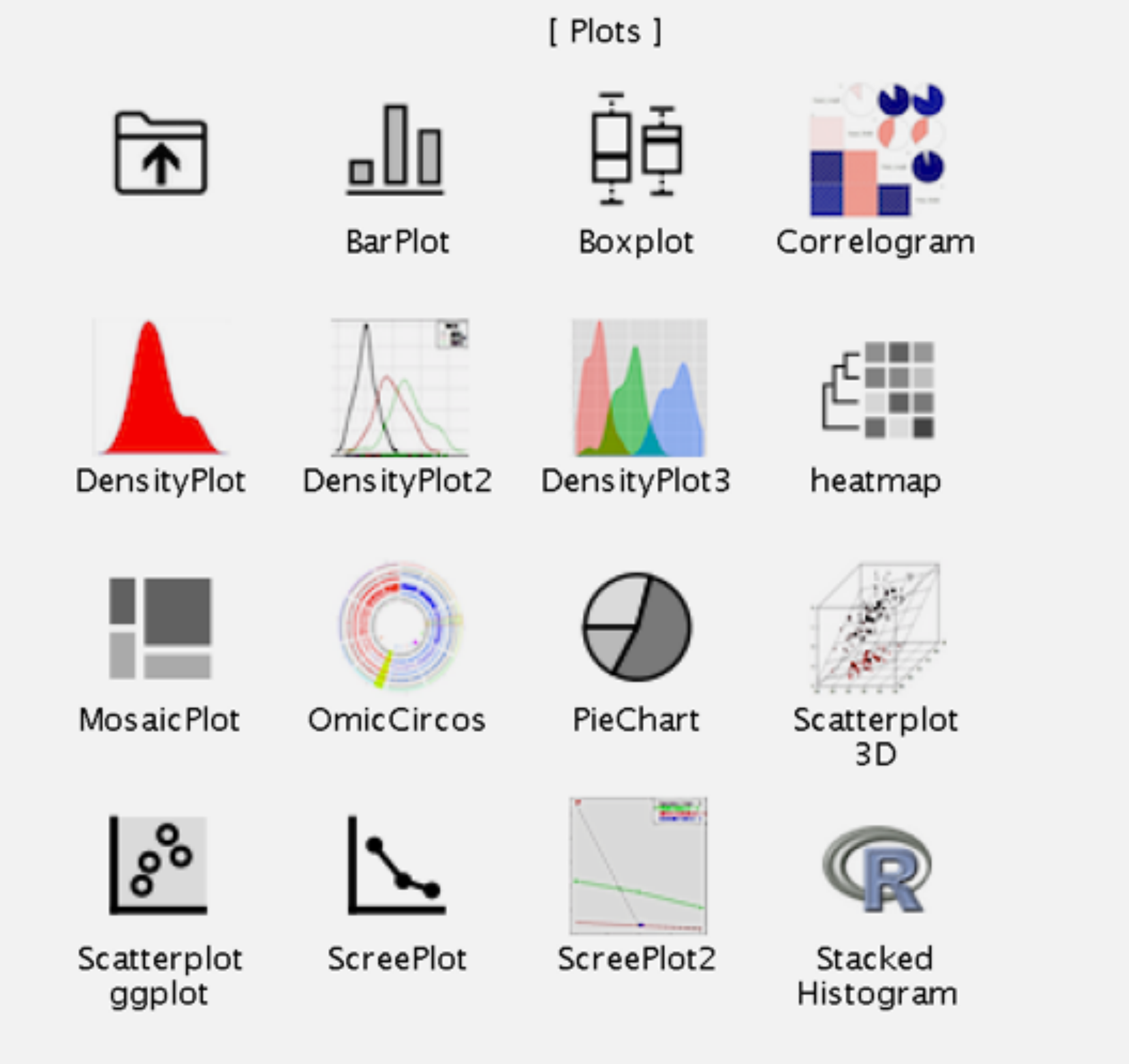
R - apps

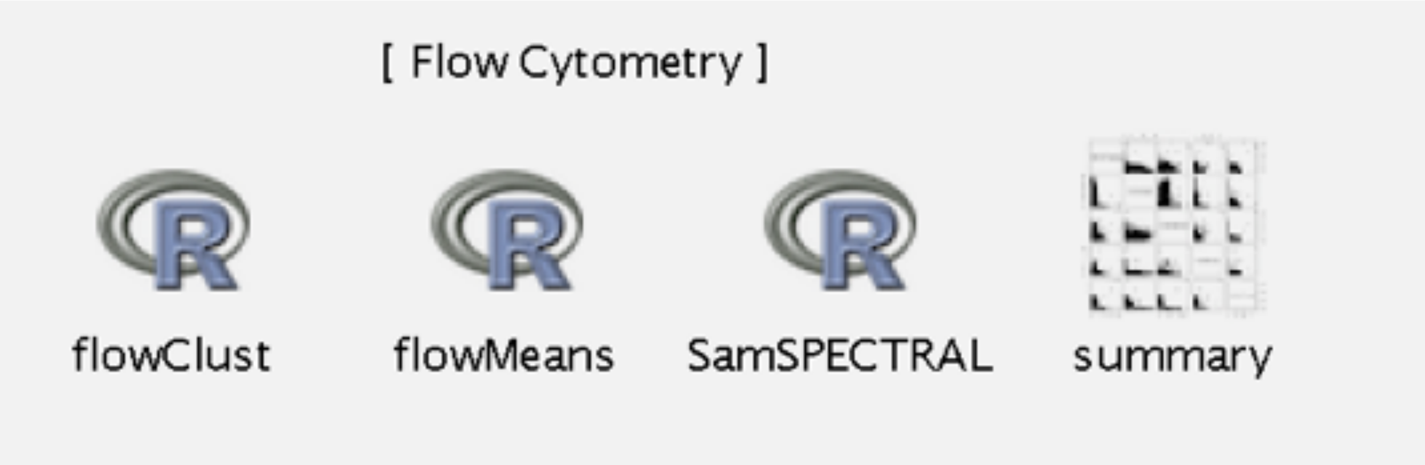
R Apps

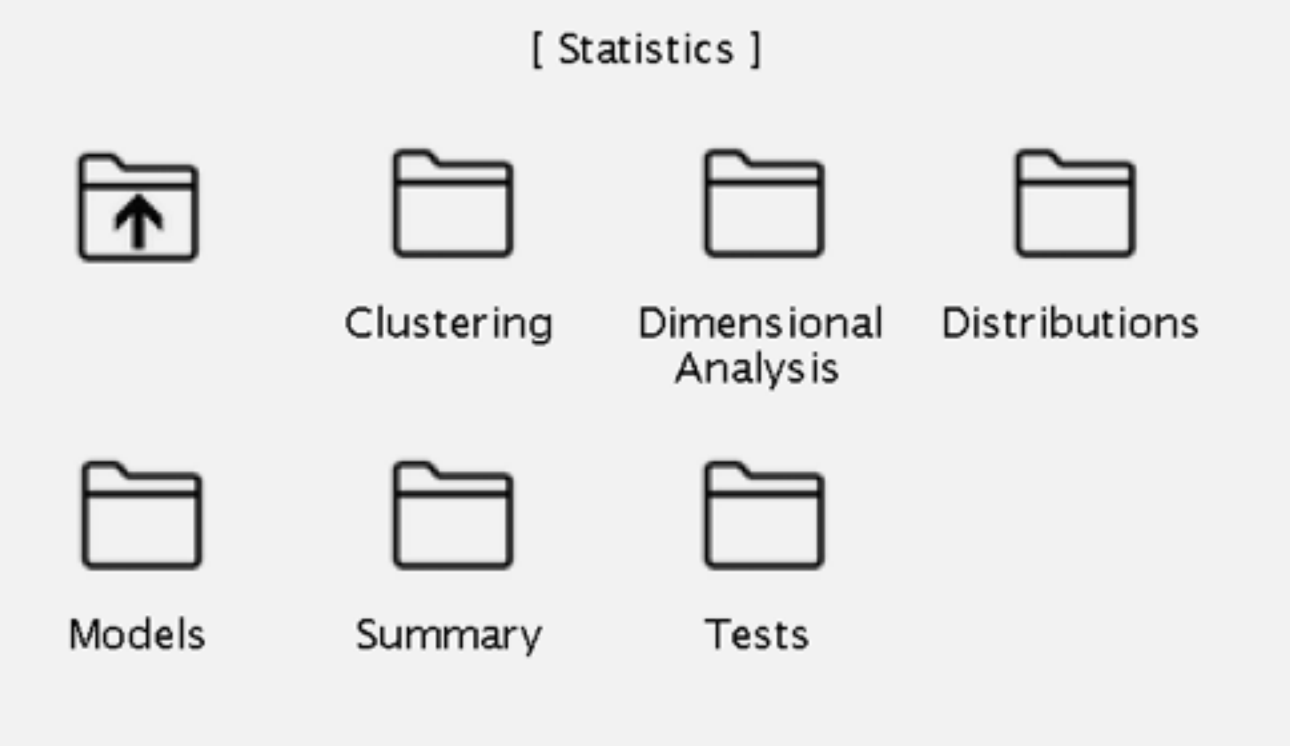


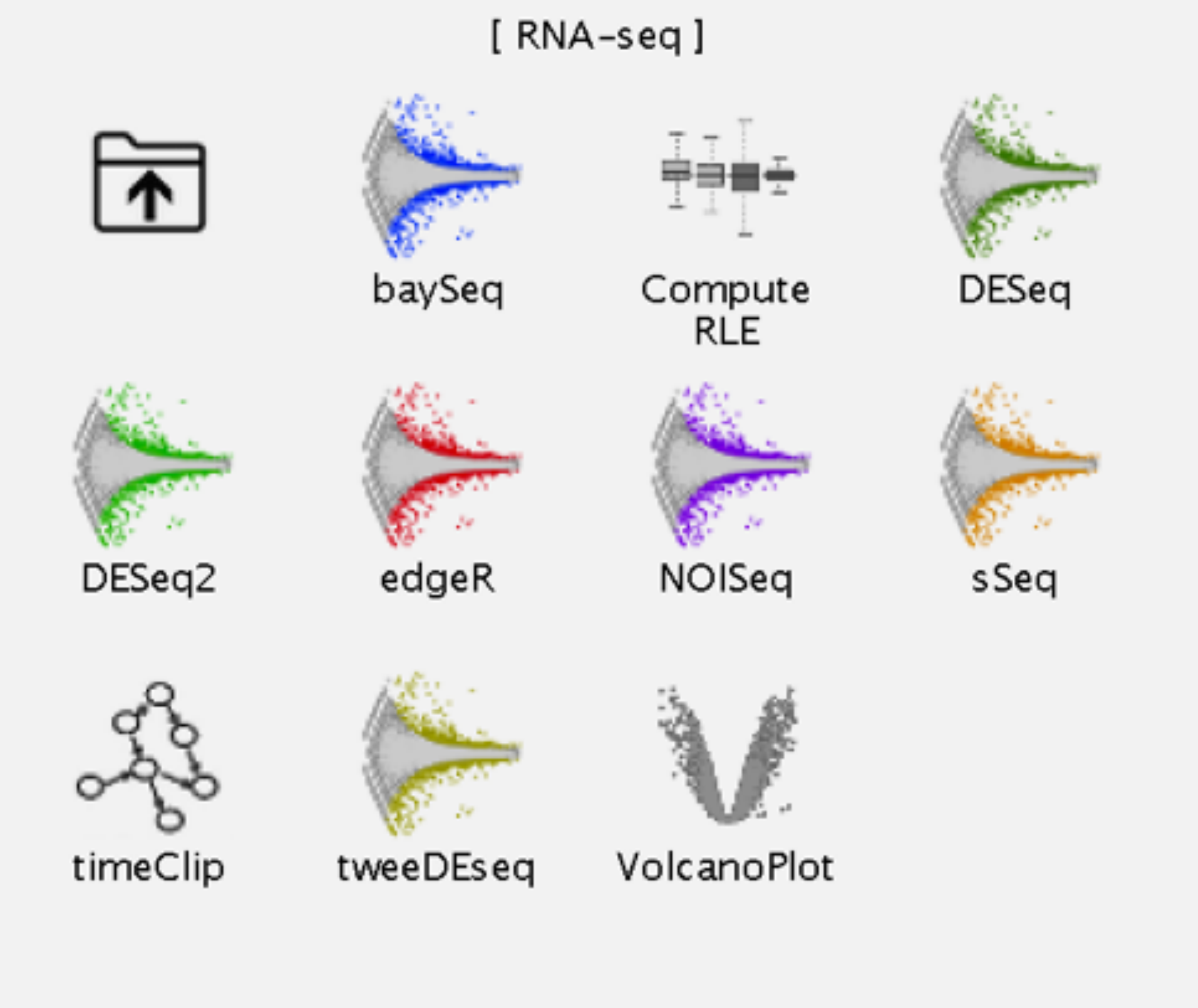
R Apps


Examples
Flow Cytometry
Plots
PSA
Sequencing
Statistics


[Plots]
BarPlot
Boxplot
Correlogram
DensityPlot
DensityPlot2
DensityPlot3
heatmap
MosaicPlot
OmicCircos
PieChart
Scatterplot 3D
Scatterplot ggplot
ScreePlot
ScreePlot2
Stacked Histogram


[Flow Cytometry]
flowClust
flowMeans
SamSPECTRAL
summary

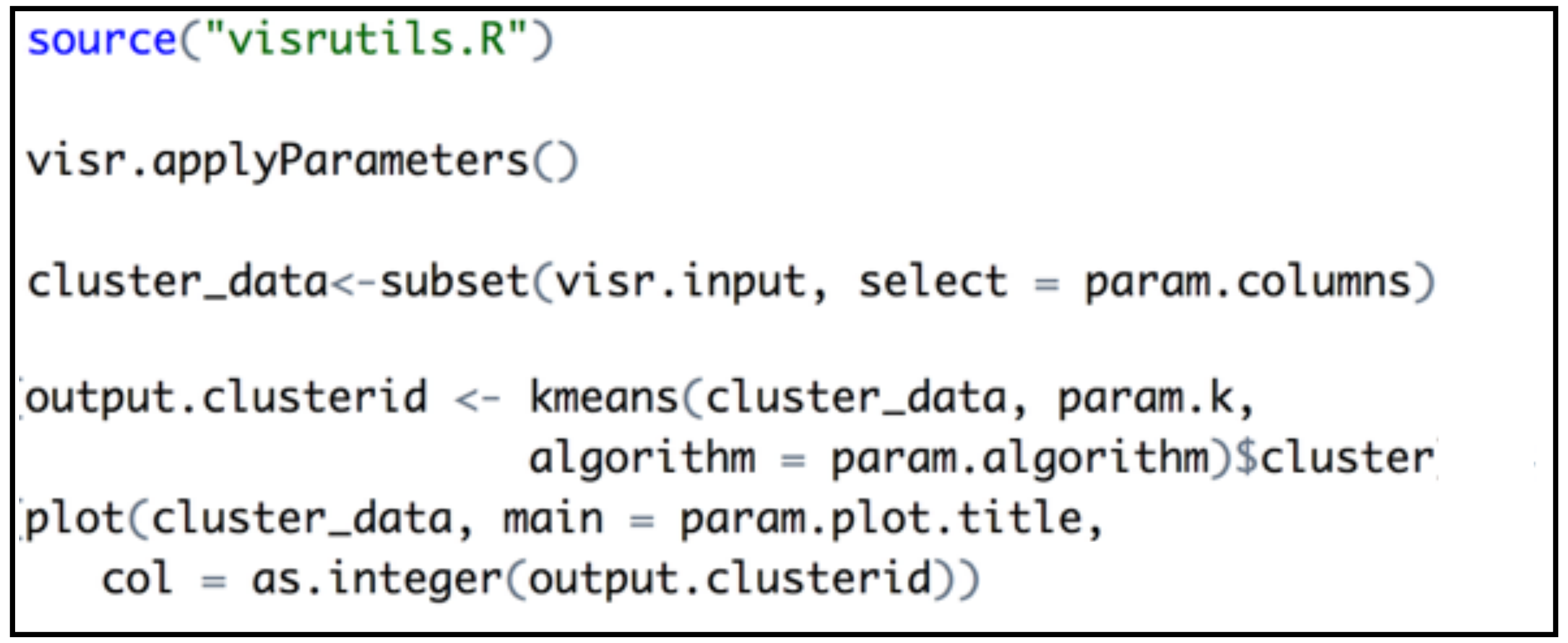

[Statistics]
Clustering
Dimensional Analysis
Distributions
Models
Summary
Tests


[RNA-seq]
baySeq
Compute RLE
DESeq
DESeq2
edgeR
NOISeq
sSeq
timeClip
tweeDEseq
VolcanoPlot

Anatomy of an R app

R code

kmeans.R

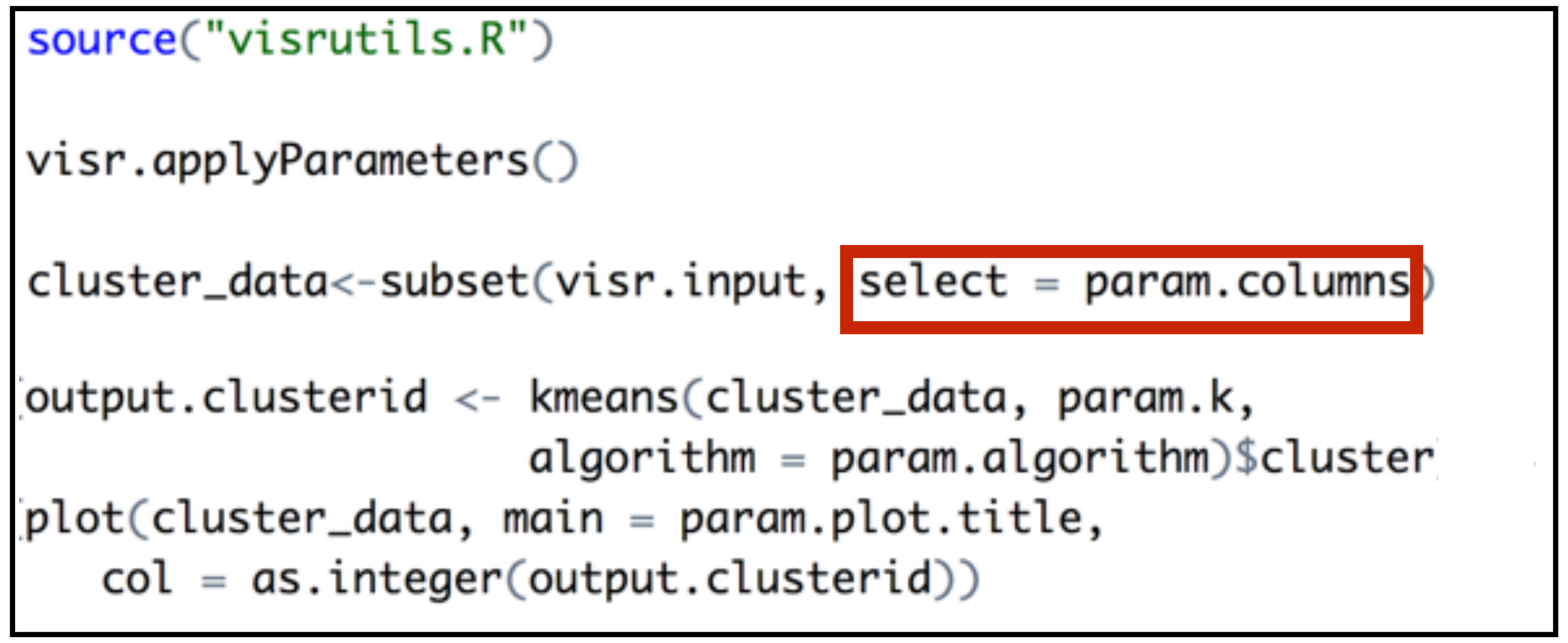


```r
source("visrutils.R")

visr.applyParameters()

cluster_data<-subset(visr.input, select = param.columns)

output.clusterid <- kmeans(cluster_data, param.k,
                           algorithm = param.algorithm)$cluster
plot(cluster_data, main = param.plot.title,
    col = as.integer(output.clusterid))
```

Anatomy of an R app

R code

kmeans.R



```
source("visrutils.R")

visr.applyParameters()

cluster_data<-subset(visr.input, select = param.columns)

output.clusterid <- kmeans(cluster_data, param.k,
                    algorithm = param.algorithm)$cluster
plot(cluster_data, main = param.plot.title,
   col = as.integer(output.clusterid))
```

parameters

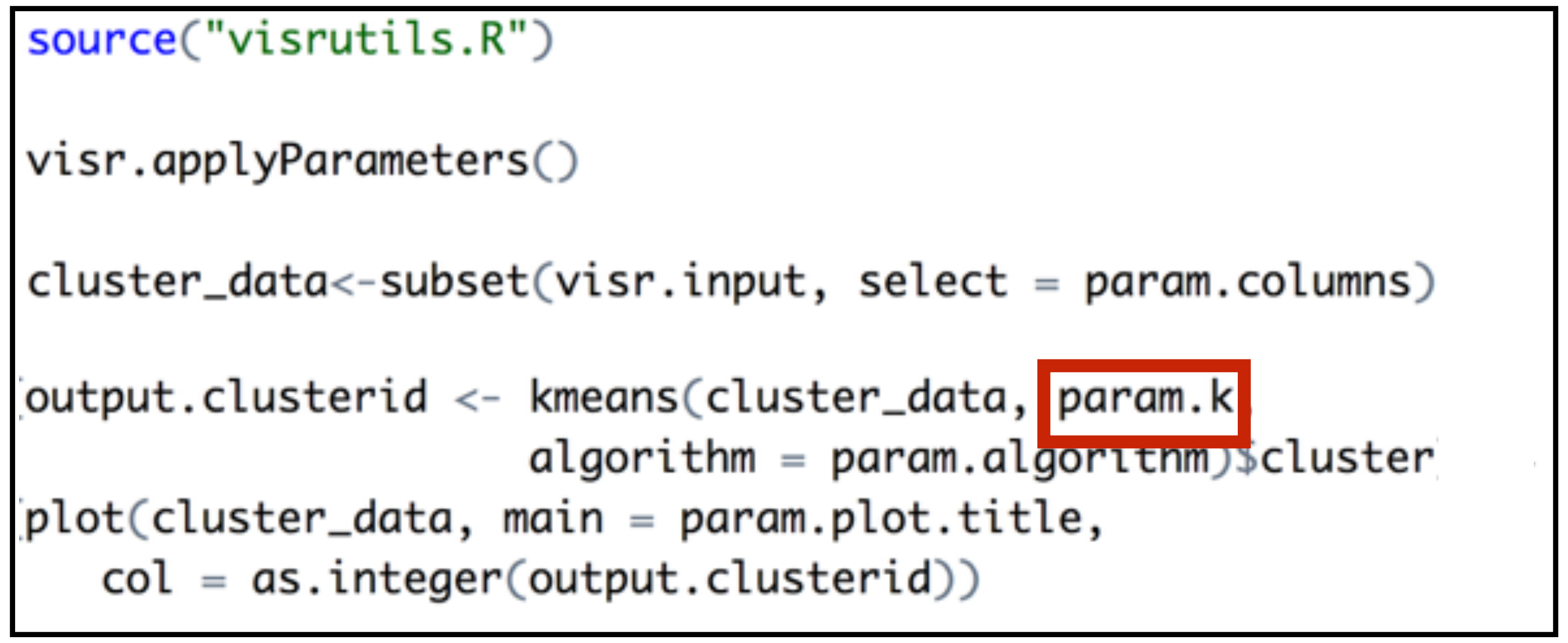
kmeans.json

```
{  "label":"kmeans", "categories":[
   { "variables":{
      "param.columns":
      "param.k":
      "param.algorithm":

      "param.plot.title":
      "output.clusterid":
   }}
]}
```

Anatomy of an R app

R code

kmeans.R



```r
source("visrutils.R")

visr.applyParameters()

cluster_data<-subset(visr.input, select = param.columns)

output.clusterid <- kmeans(cluster_data, param.k,
                    algorithm = param.algorithm)$cluster
plot(cluster_data, main = param.plot.title,
   col = as.integer(output.clusterid))
```

parameters

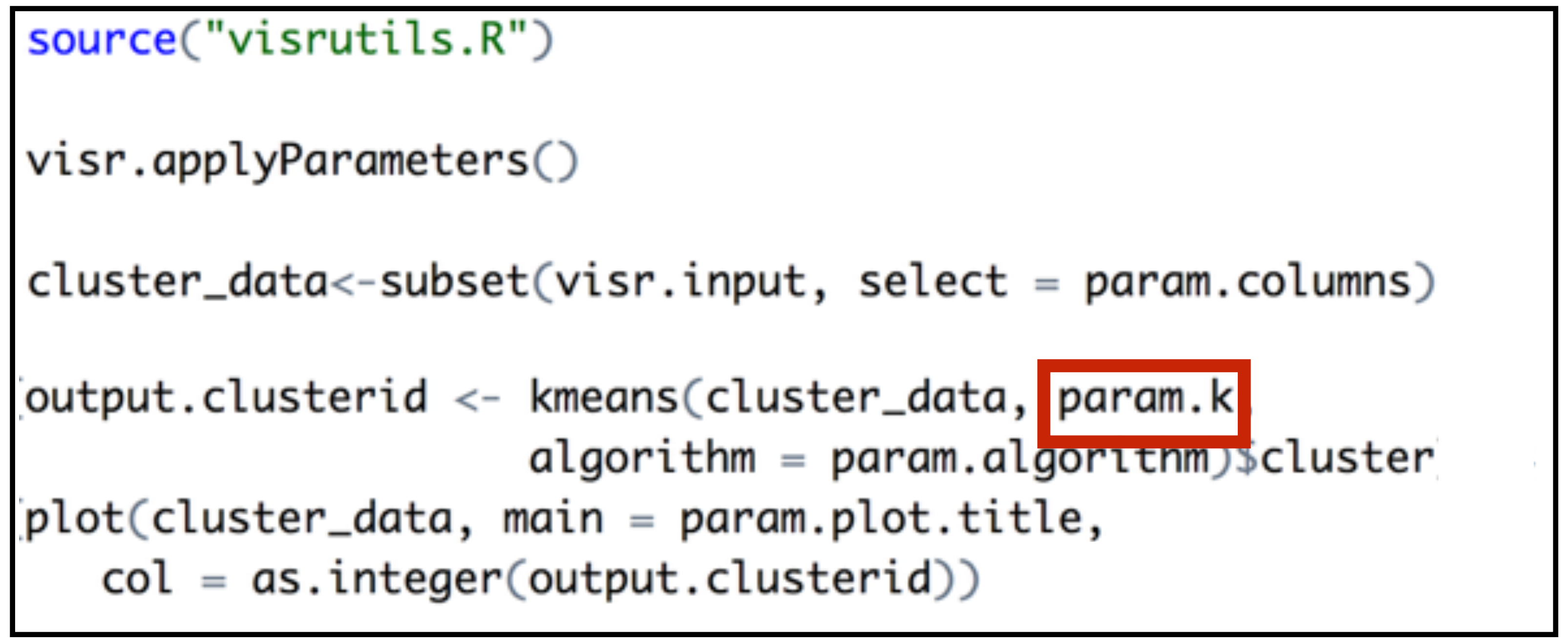
kmeans.json

```json
{  "label":"kmeans", "categories":[
   { "variables":{
      "param.columns":
      "param.k":
      "param.algorithm":

      "param.plot.title":
      "output.clusterid":
   }}
]}
```

Anatomy of an R app

R code

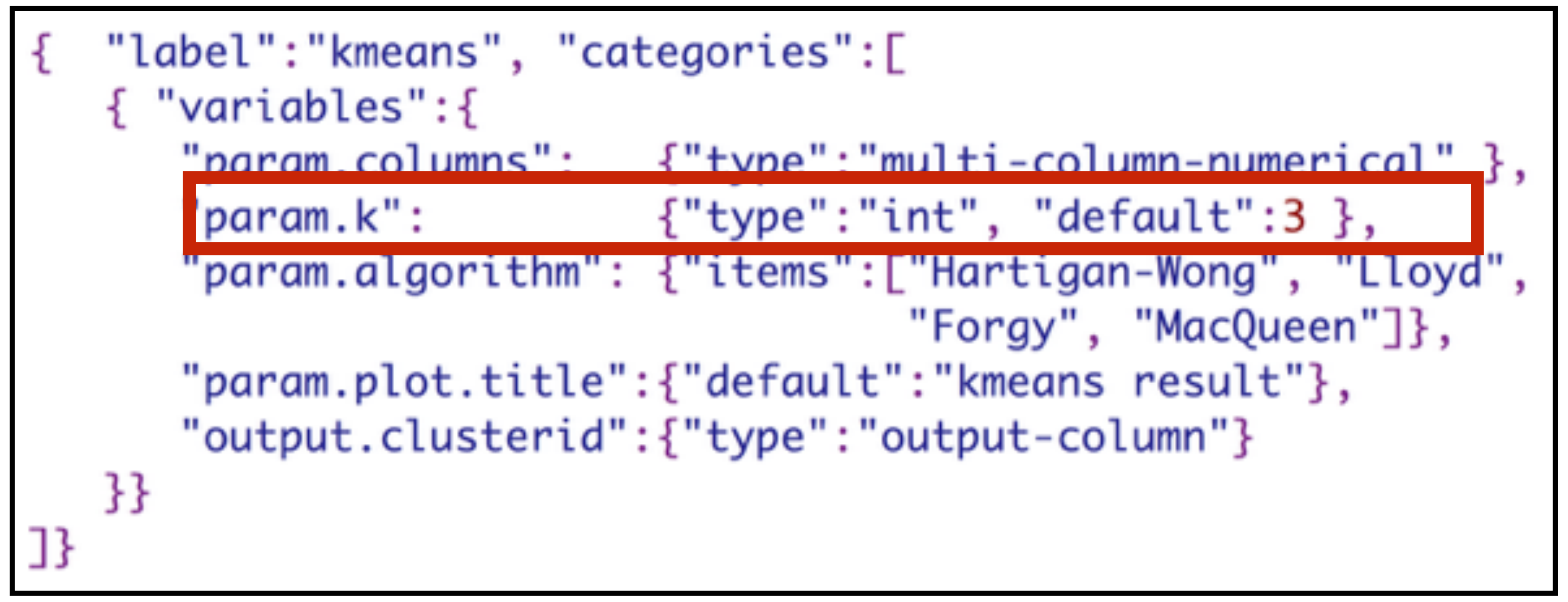
kmeans.R



```r
source("visrutils.R")

visr.applyParameters()

cluster_data<-subset(visr.input, select = param.columns)

output.clusterid <- kmeans(cluster_data, param.k,
                    algorithm = param.algorithm)$cluster
plot(cluster_data, main = param.plot.title,
   col = as.integer(output.clusterid))
```

parameters

kmeans.json

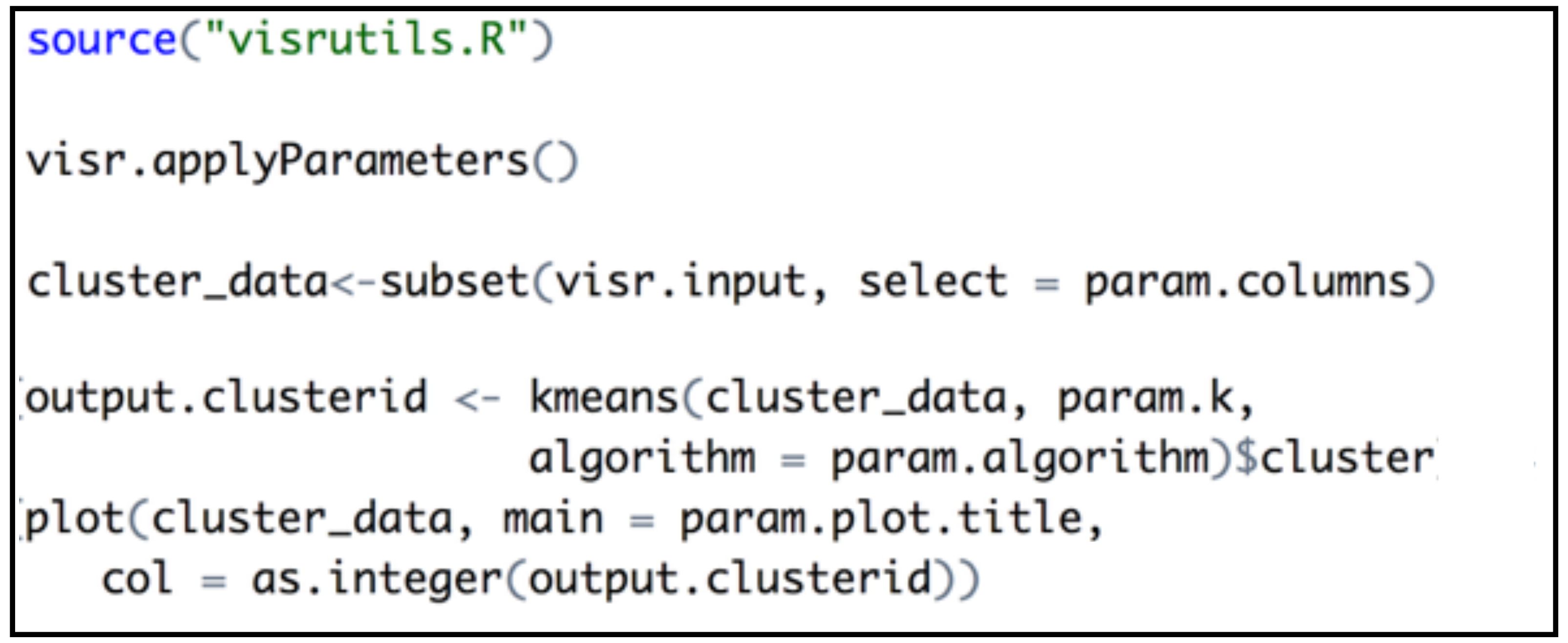


```json
{  "label":"kmeans", "categories":[
   { "variables":{
      "param.columns":    {"type":"multi-column-numerical" },
      "param.k":          {"type":"int", "default":3 },
      "param.algorithm": {"items":["Hartigan-Wong", "Lloyd",
                                   "Forgy", "MacQueen"]},
      "param.plot.title":{"default":"kmeans result"},
      "output.clusterid":{"type":"output-column"}
   }}
]}
```

Anatomy of an R app

R code

kmeans.R



```
source("visrutils.R")

visr.applyParameters()

cluster_data<-subset(visr.input, select = param.columns)

output.clusterid <- kmeans(cluster_data, param.k,
                    algorithm = param.algorithm)$cluster
plot(cluster_data, main = param.plot.title,
   col = as.integer(output.clusterid))
```

parameters

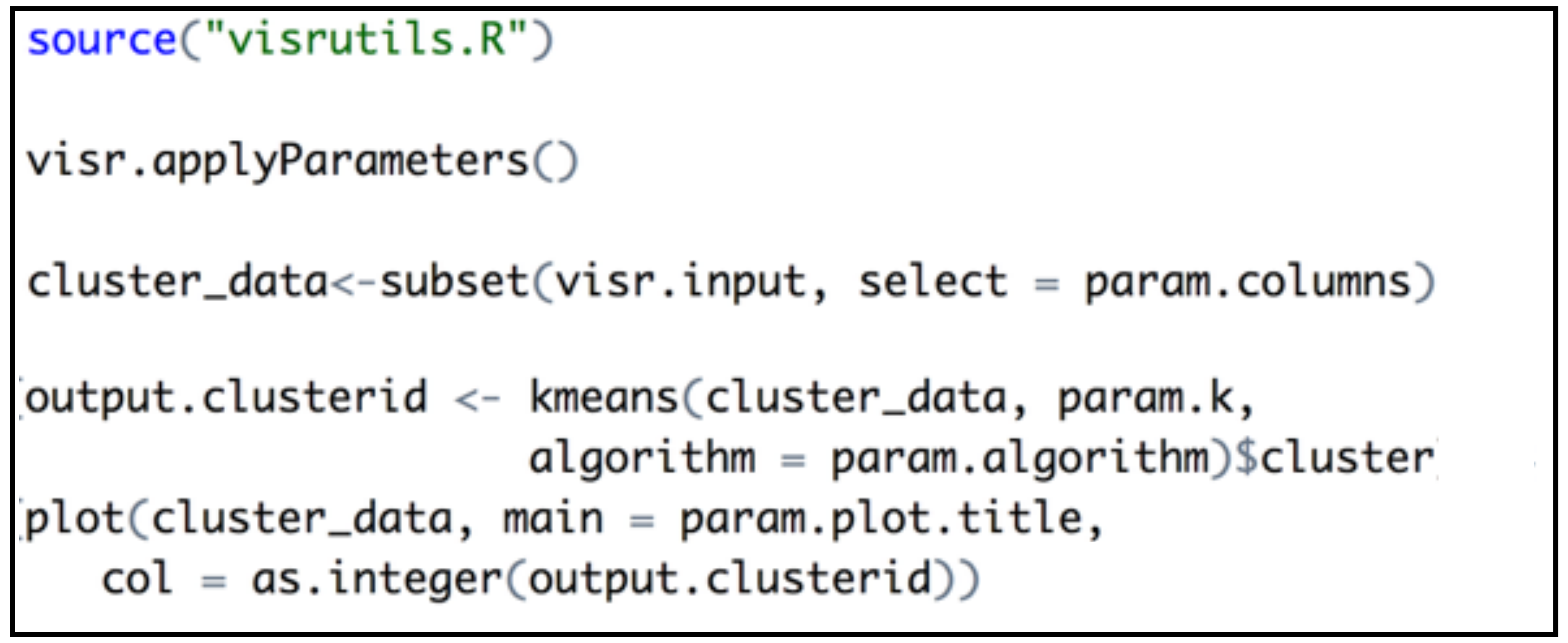
kmeans.json

```
{  "label":"kmeans", "categories":[
   { "variables":{
      "param.columns":   {"type":"multi-column-numerical" },
      "param.k":         {"type":"int", "default":3 },
      "param.algorithm": {"items":["Hartigan-Wong", "Lloyd",
                                   "Forgy", "MacQueen"]},
      "param.plot.title":{"default":"kmeans result"},
      "output.clusterid":{"type":"output-column"}
   }}
]}
```

Anatomy of an R app

R code

kmeans.R



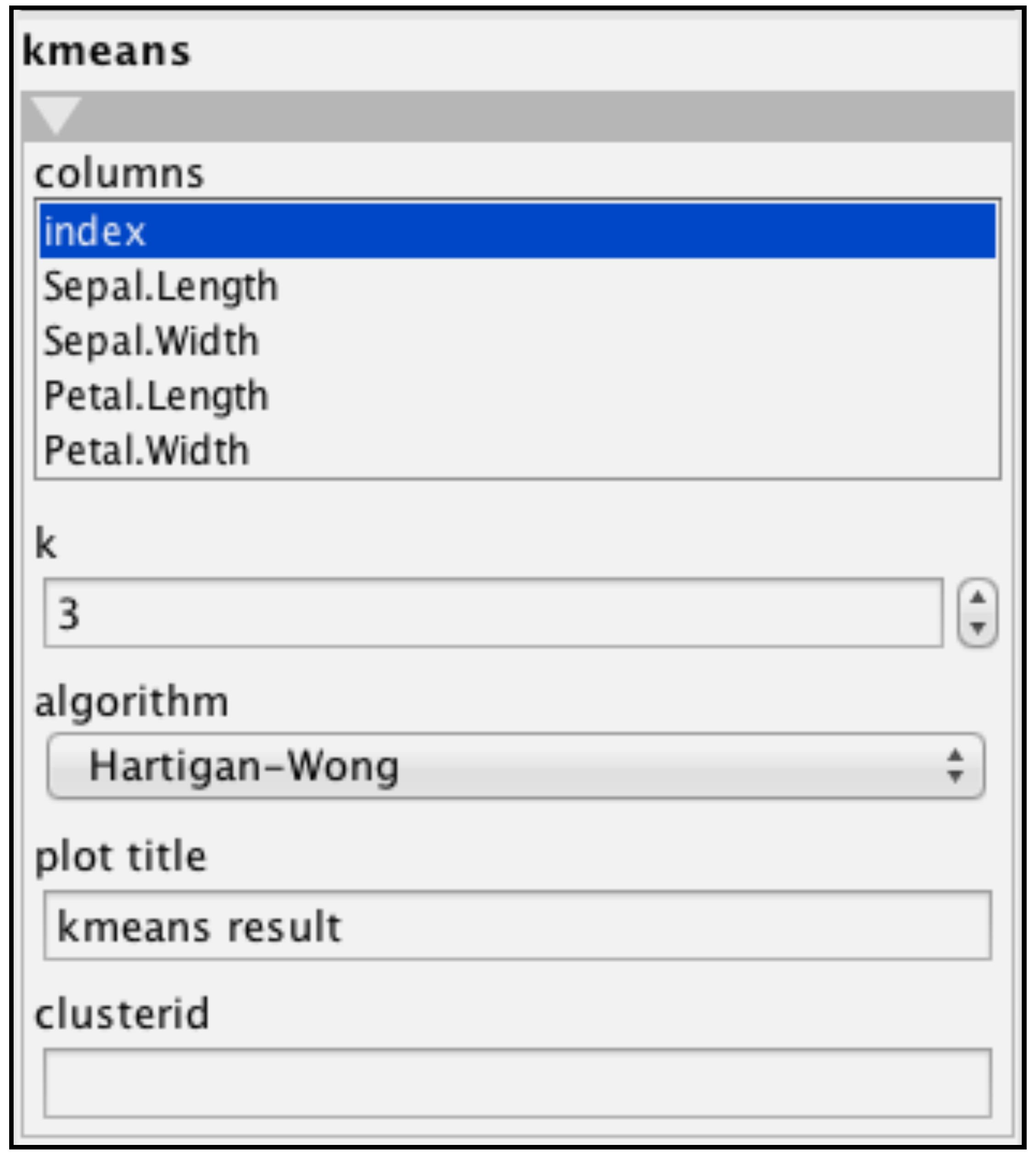
```
source("visrutils.R")

visr.applyParameters()

cluster_data<-subset(visr.input, select = param.columns)

output.clusterid <- kmeans(cluster_data, param.k,
                    algorithm = param.algorithm)$cluster
plot(cluster_data, main = param.plot.title,
   col = as.integer(output.clusterid))
```

parameters

kmeans.json

```
{  "label":"kmeans", "categories":[
   { "variables":{
      "param.columns":   {"type":"multi-column-numerical" },
      "param.k":         {"type":"int", "default":3 },
      "param.algorithm": {"items":["Hartigan-Wong", "Lloyd",
                                    "Forgy", "MacQueen"]},
      "param.plot.title":{"default":"kmeans result"},
      "output.clusterid":{"type":"output-column"}
   }}
]}
```

auto generated UI



kmeans

columns

- index
- Sepal.Length
- Sepal.Width
- Petal.Length
- Petal.Width

k

3

algorithm

Hartigan-Wong

plot title

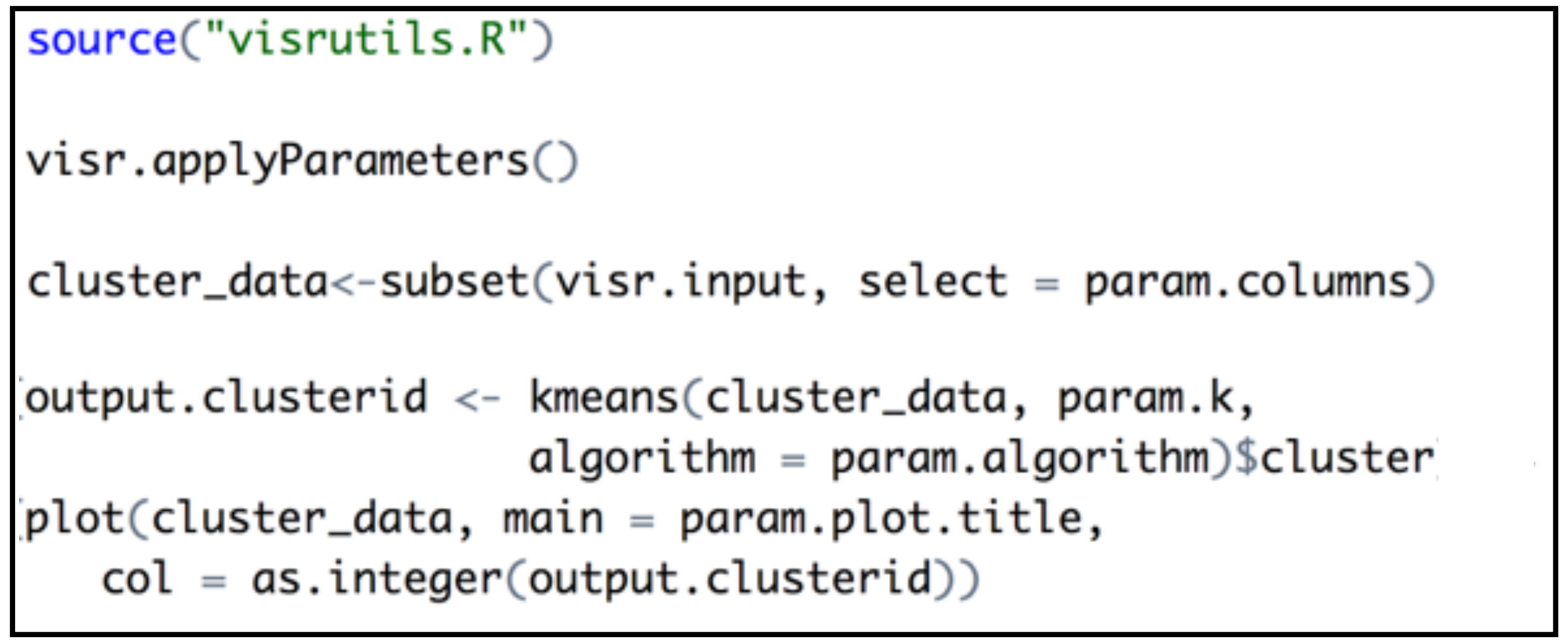
kmeans result

clusterid

Anatomy of an R app

R code

kmeans.R



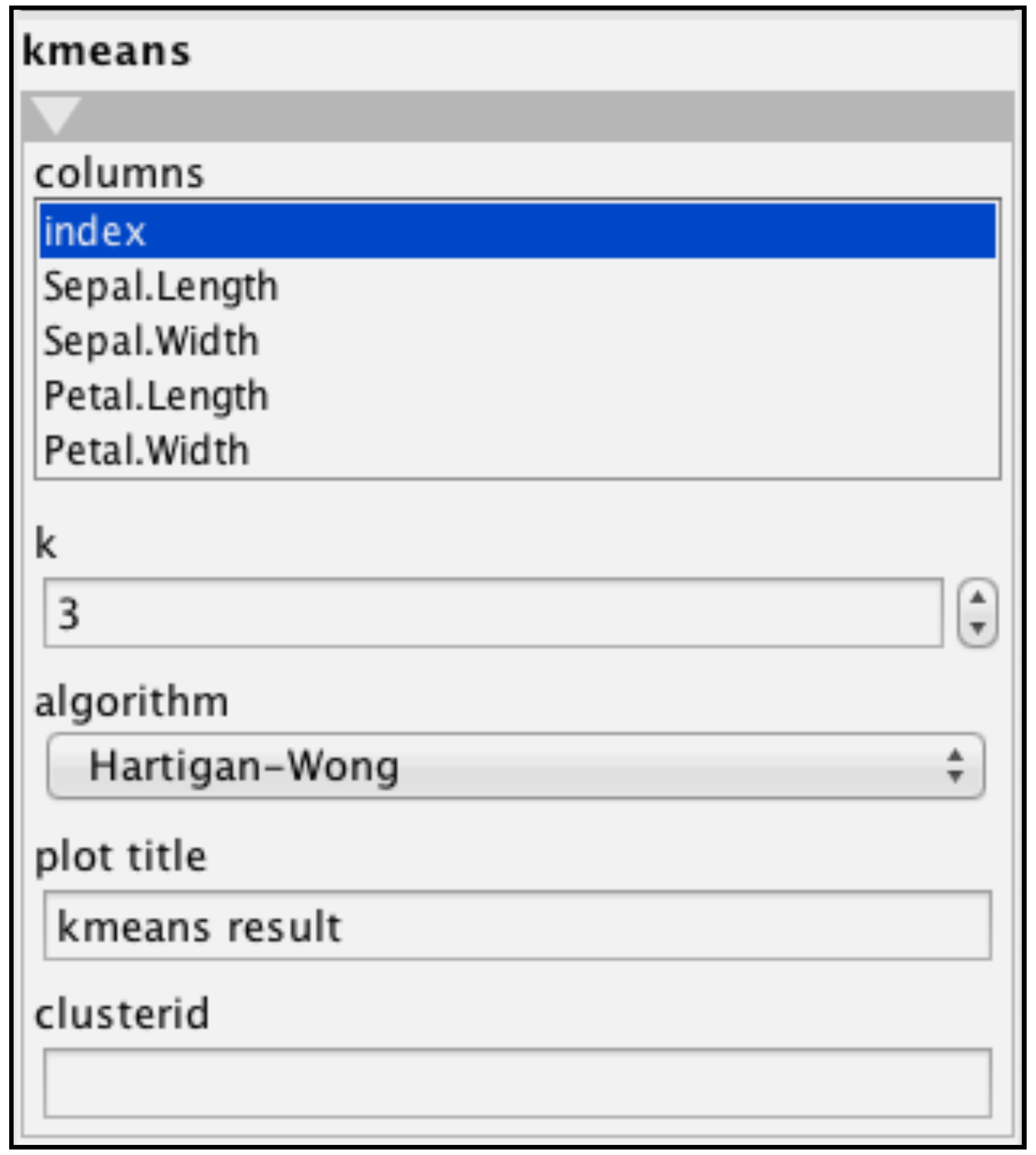
```
source("visrutils.R")

visr.applyParameters()

cluster_data<-subset(visr.input, select = param.columns)

output.clusterid <- kmeans(cluster_data, param.k,
                           algorithm = param.algorithm)$cluster
plot(cluster_data, main = param.plot.title,
    col = as.integer(output.clusterid))
```

parameters

kmeans.json

```
{  "label":"kmeans", "categories":[
   { "variables":{
      "param.columns":   {"type":"multi-column-numerical" },
      "param.k":         {"type":"int", "default":3 },
      "param.algorithm": {"items":["Hartigan-Wong", "Lloyd",
                                   "Forgy", "MacQueen"]},
      "param.plot.title":{"default":"kmeans result"},
      "output.clusterid":{"type":"output-column"}
   }}
]}
```

auto generated UI



kmeans

columns

- index
- Sepal.Length
- Sepal.Width
- Petal.Length
- Petal.Width

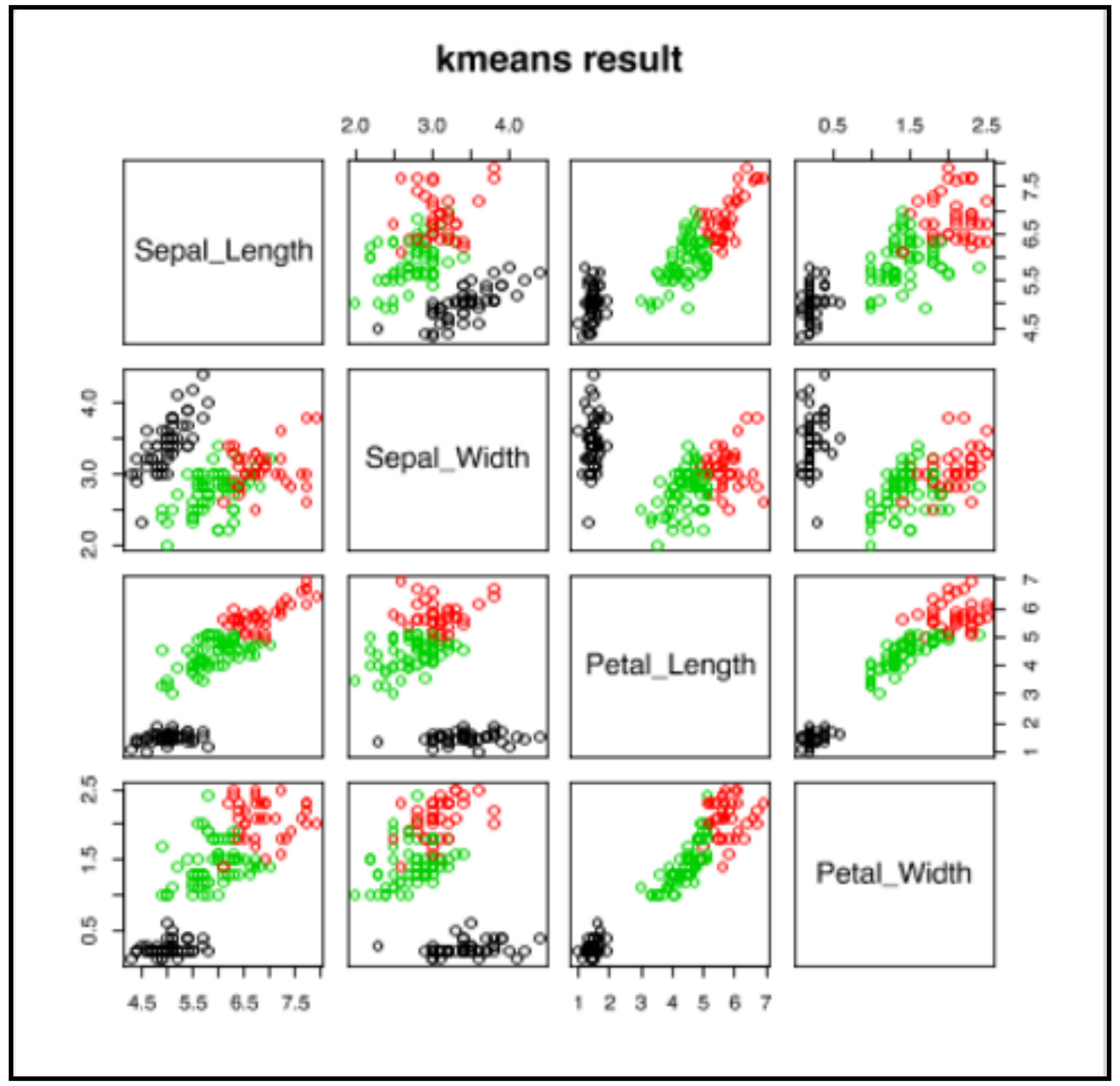
k: 3

algorithm: Hartigan-Wong

plot title: kmeans result

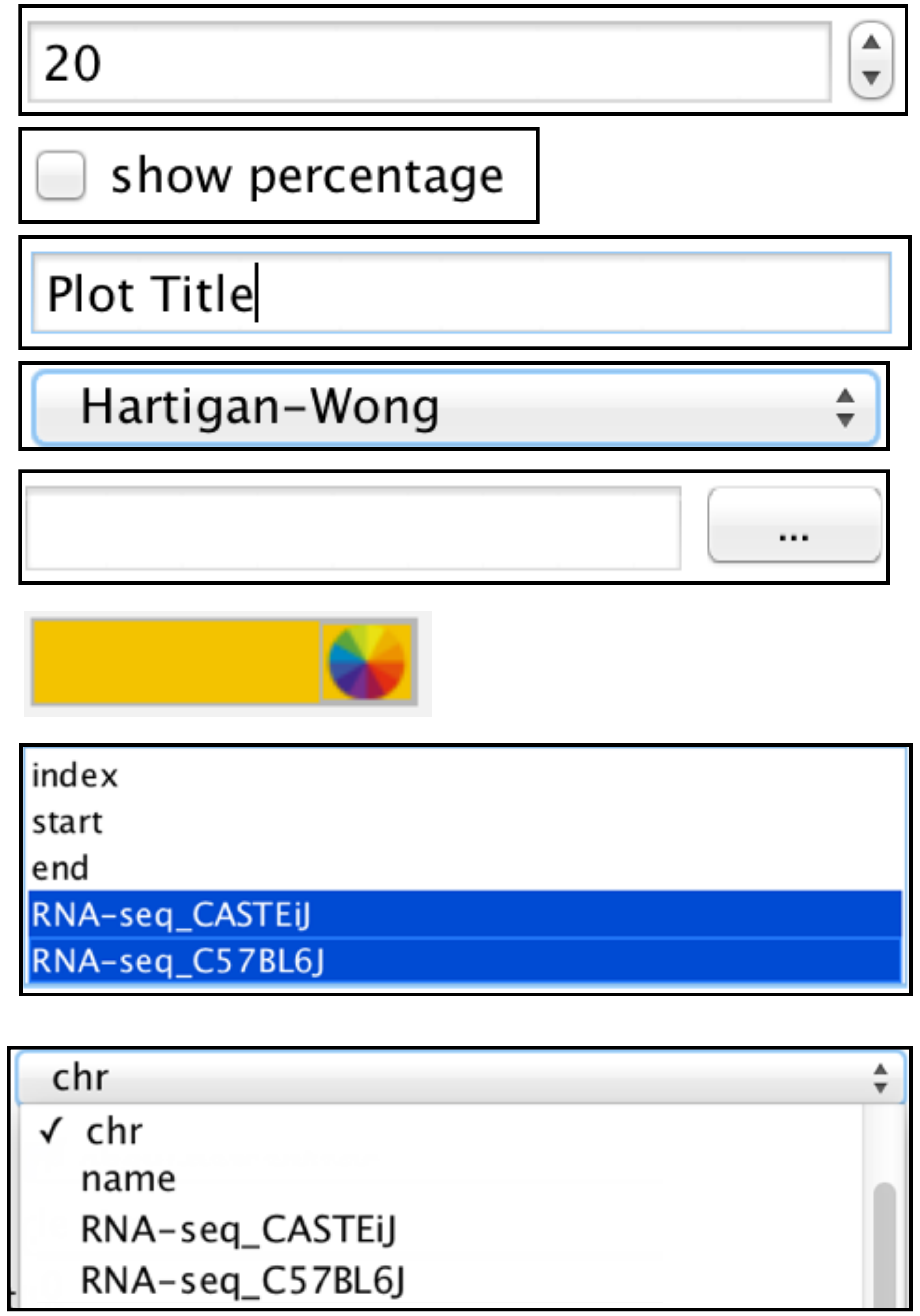
clusterid

output



creating R apps: variable types

variable type
int
double
boolean
string
string **with** items
filename
color
range-int
range-double
column
column-numerical
multi-column
multi-column-numerical
ouput-column
output-table


20

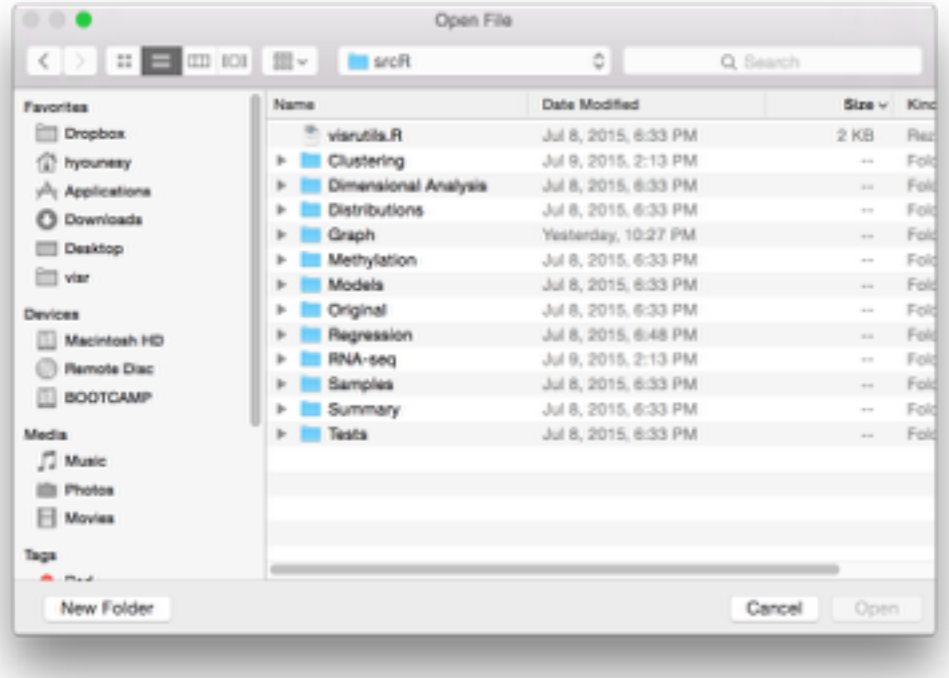
☐ show percentage

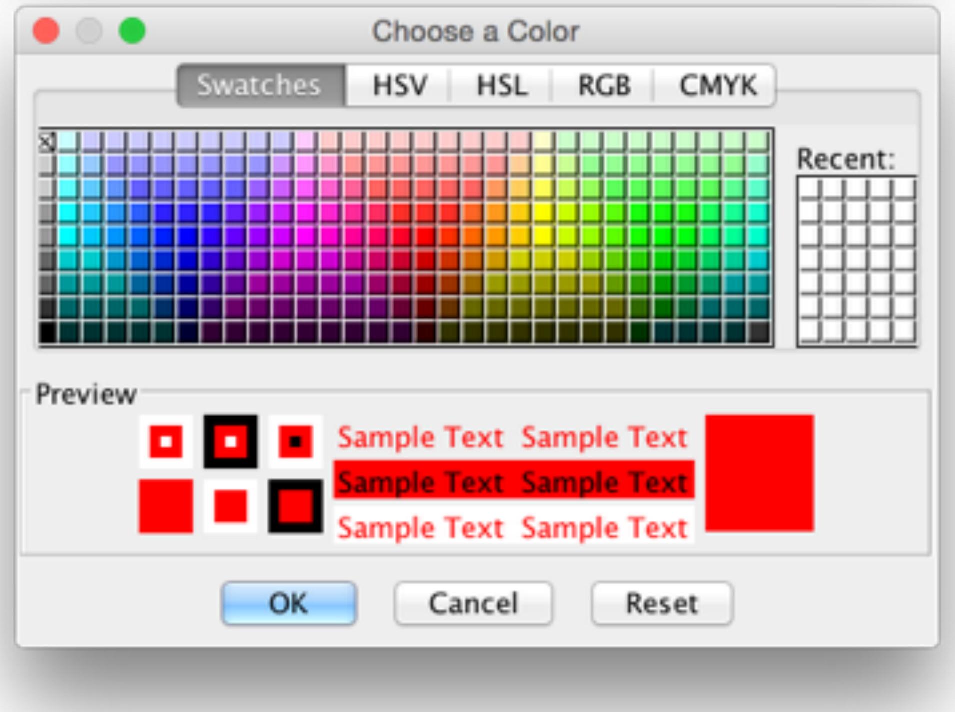
Plot Title

Hartigan-Wong

index
start
end
RNA-seq_CASTEiJ
RNA-seq_C57BL6J

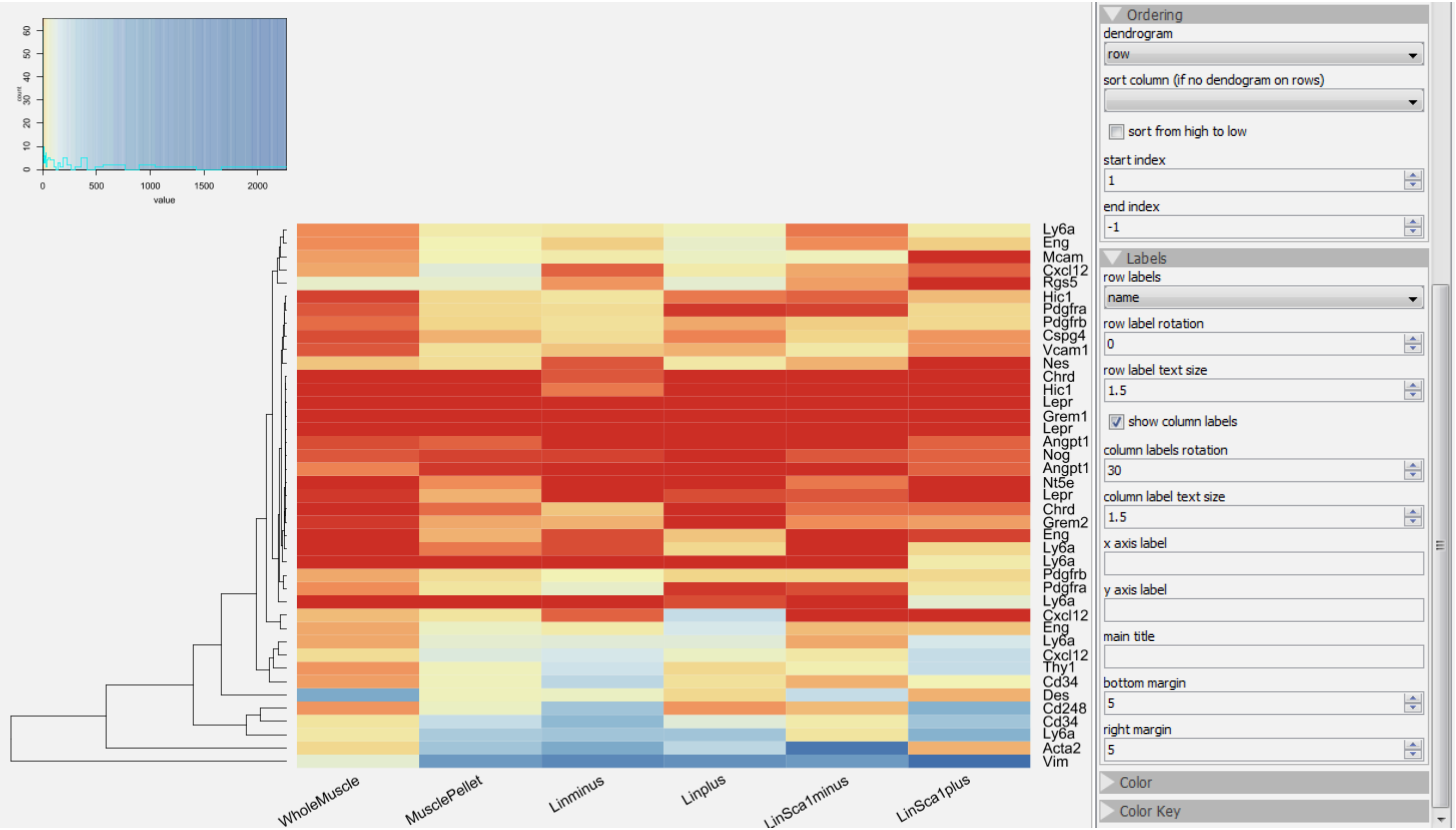
chr
✓ chr
name
RNA-seq_CASTEiJ
RNA-seq_C57BL6J





```
visr.message(message, type)
```

R – Apps example: heatmap



Ordering

dendrogram: row

sort column (if no dendogram on rows):

☐ sort from high to low

start index: 1

end index: -1

Labels

row labels: name

row label rotation: 0

row label text size: 1.5

☑ show column labels

column labels rotation: 30

column label text size: 1.5

x axis label:

y axis label:

main title:

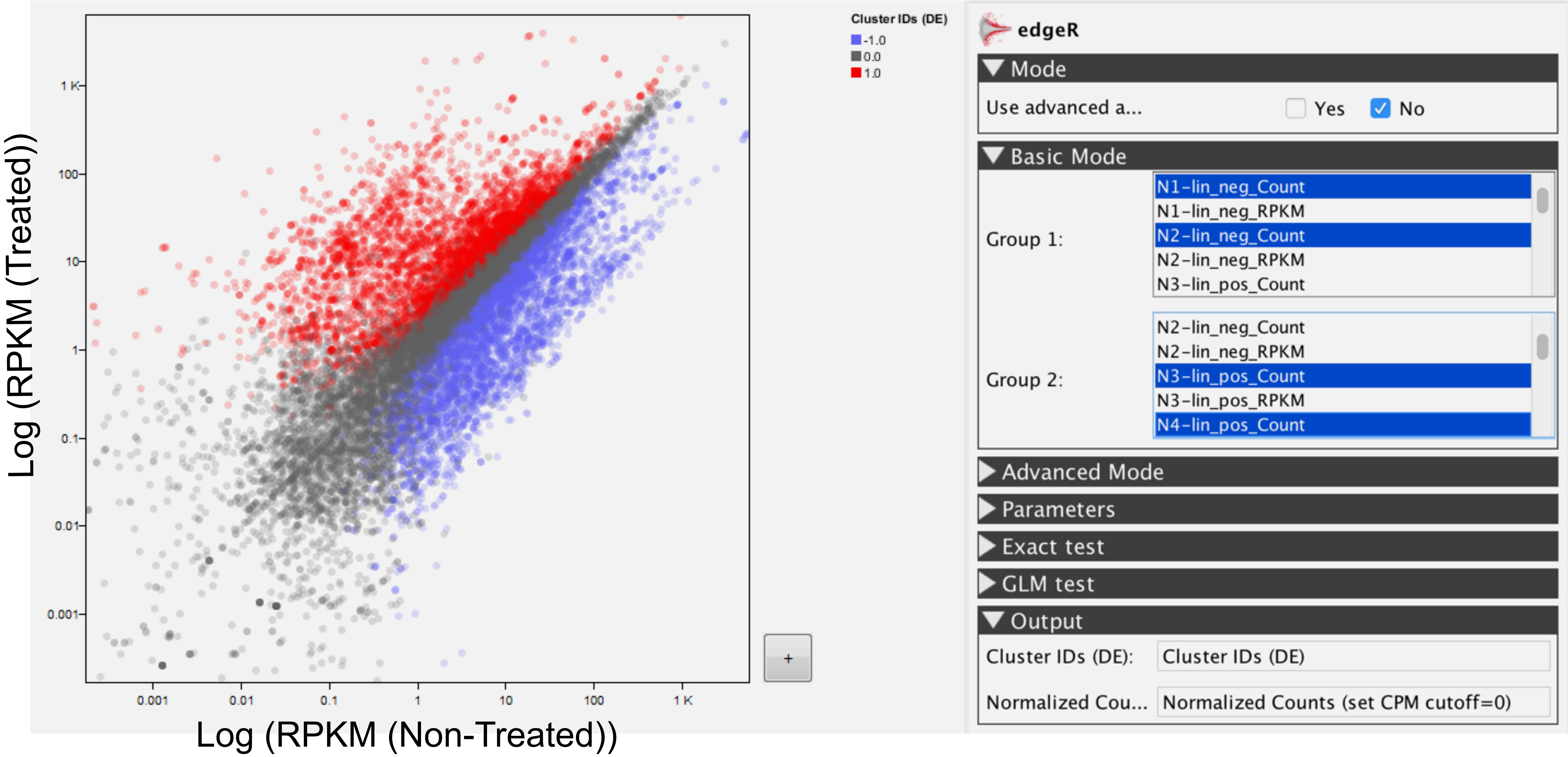
bottom margin: 5

right margin: 5

Color

Color Key

R – Apps example: edgeR



edgeR

Mode

Use advanced a... ☐ Yes ☑ No

Basic Mode

Group 1:
- N1–lin_neg_Count
- N1–lin_neg_RPKM
- N2–lin_neg_Count
- N2–lin_neg_RPKM
- N3–lin_pos_Count

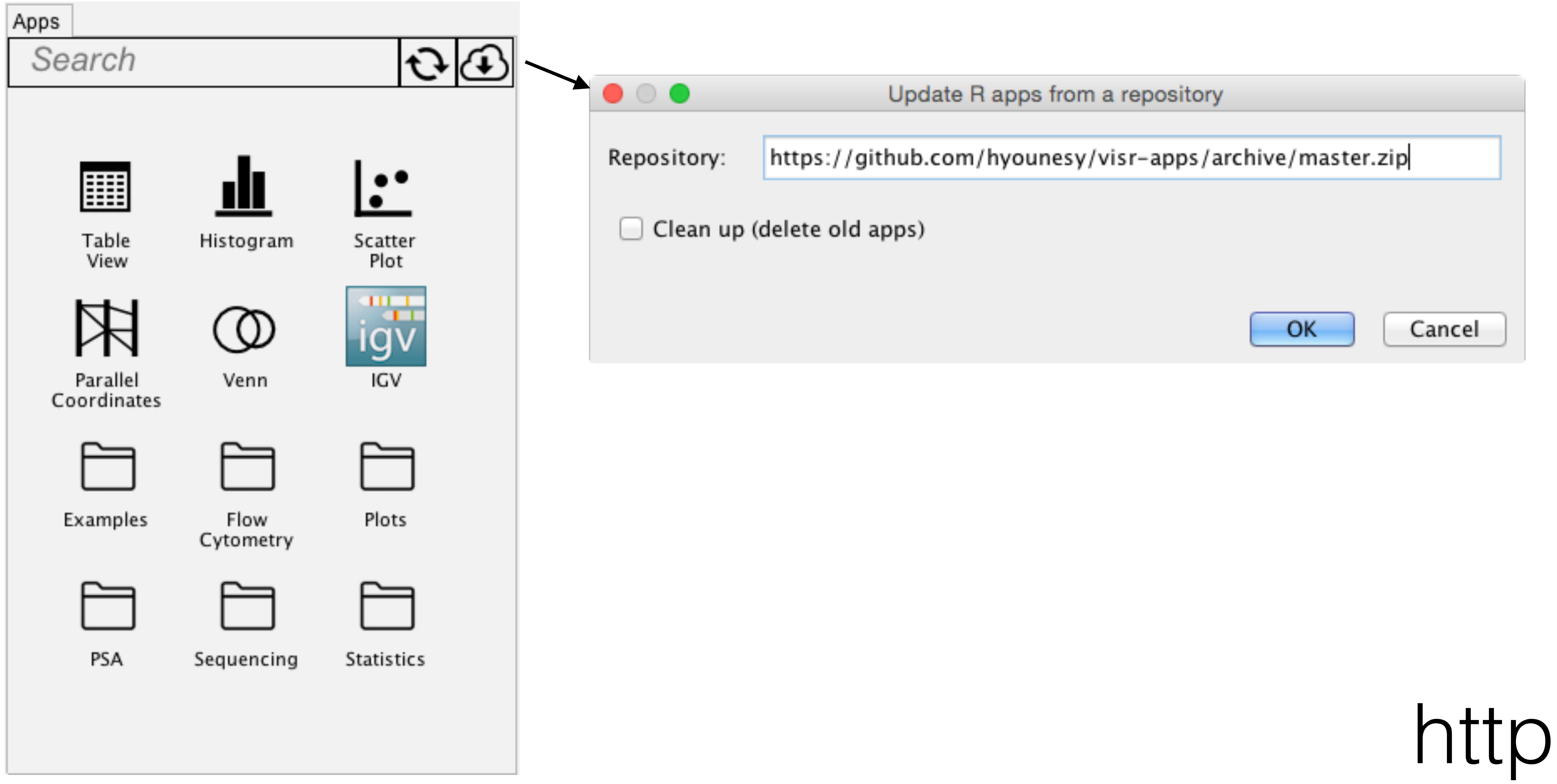
Group 2:
- N2–lin_neg_Count
- N2–lin_neg_RPKM
- N3–lin_pos_Count
- N3–lin_pos_RPKM
- N4–lin_pos_Count

Output

Cluster IDs (DE): Cluster IDs (DE)

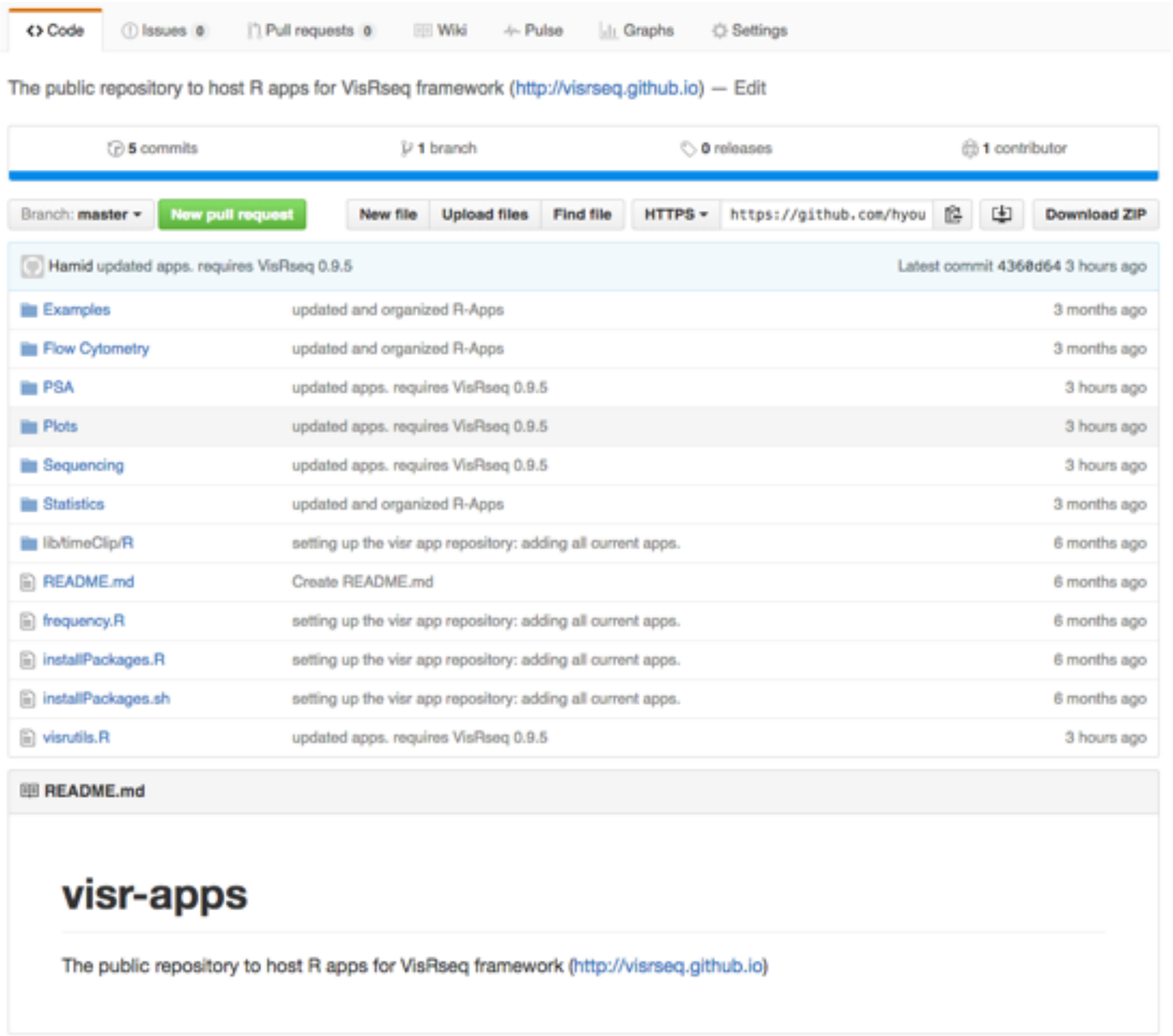
Normalized Cou... Normalized Counts (set CPM cutoff=0)

creating R apps: sharing



Repository: https://github.com/hyounesy/visr-apps/archive/master.zip

☐ Clean up (delete old apps)



The public repository to host R apps for VisRseq framework (http://visrseq.github.io) — Edit

Hamid updated apps. requires VisRseq 0.9.5 Latest commit 4360d64 3 hours ago

Examples	updated and organized R-Apps	3 months ago
Flow Cytometry	updated and organized R-Apps	3 months ago
PSA	updated apps. requires VisRseq 0.9.5	3 hours ago
Plots	updated apps. requires VisRseq 0.9.5	3 hours ago
Sequencing	updated apps. requires VisRseq 0.9.5	3 hours ago
Statistics	updated and organized R-Apps	3 months ago
lib/timeClip/R	setting up the visr app repository: adding all current apps.	6 months ago
README.md	Create README.md	6 months ago
frequency.R	setting up the visr app repository: adding all current apps.	6 months ago
installPackages.R	setting up the visr app repository: adding all current apps.	6 months ago
installPackages.sh	setting up the visr app repository: adding all current apps.	6 months ago
visrutils.R	updated apps. requires VisRseq 0.9.5	3 hours ago

README.md

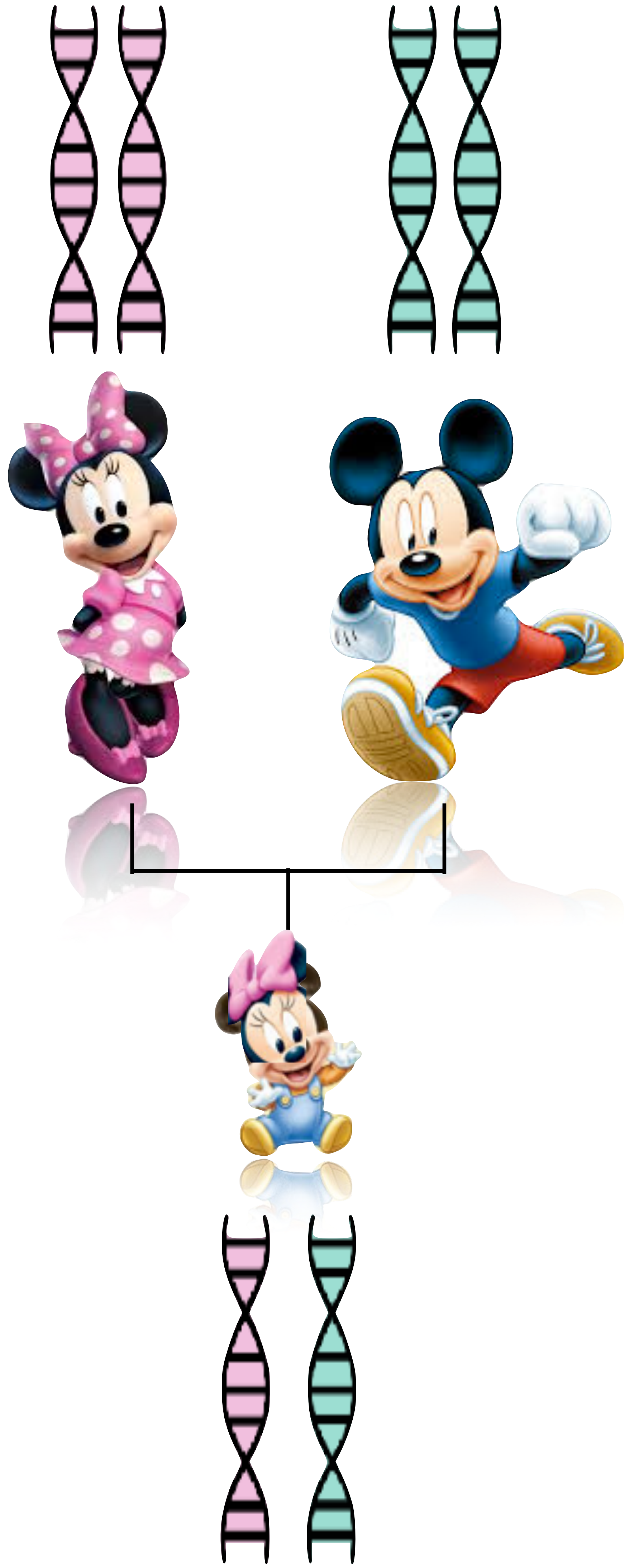
visr-apps

The public repository to host R apps for VisRseq framework (http://visrseq.github.io)

https://github.com/hyounesy/visr-apps/

case study

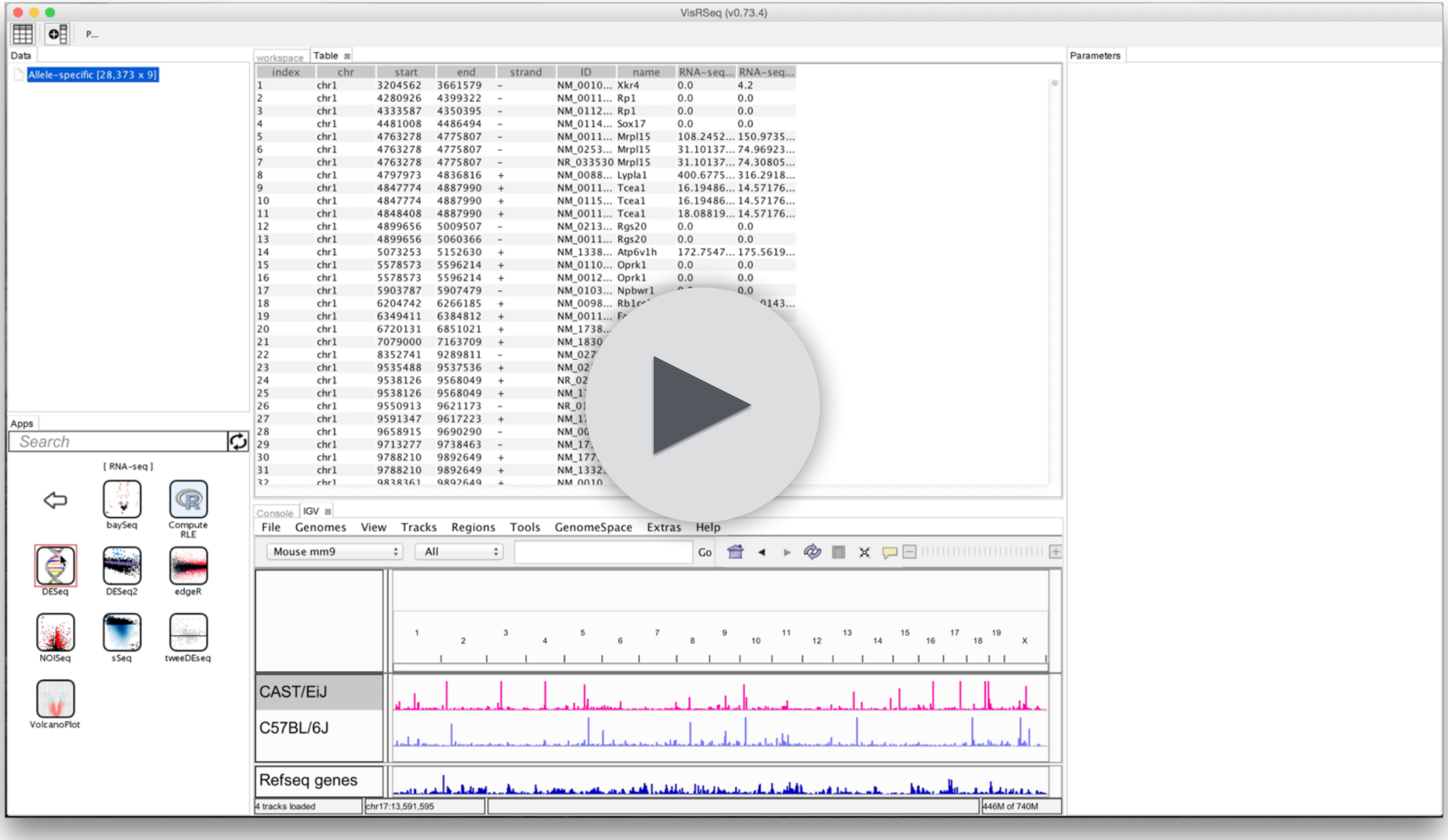
case study



- allele specific gene expression (RNA-seq)
- F1 hybrid mice: CAST + C57
- allelic RNA-seq created using ALEA[1]

[1] H. Younesy, et.al.: ALEA: a toolbox for allele-specific epigenomics analysis. Bioinformatics 30(8), 1172–1174 (2014)

case study: allele specific expression

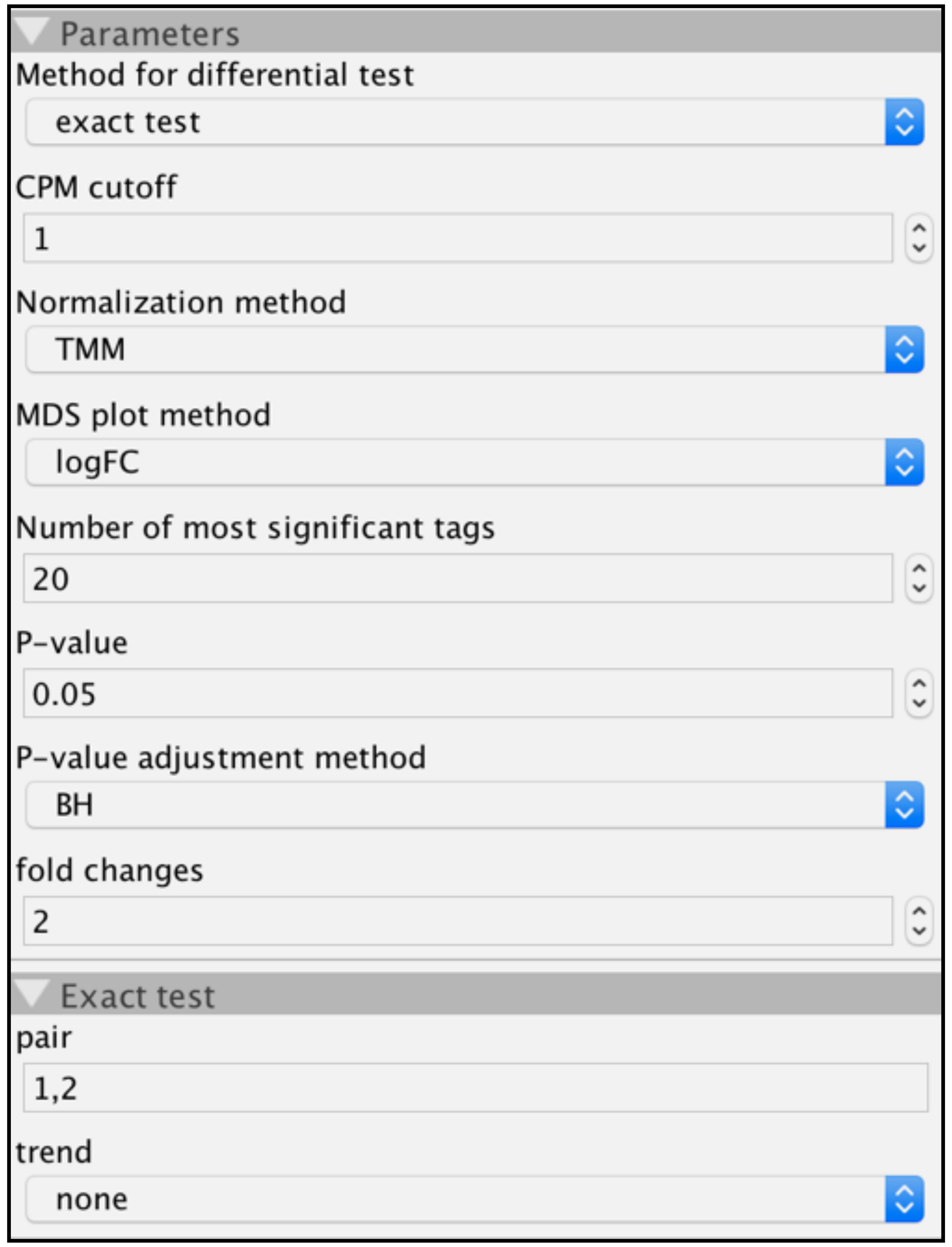


index	chr	start	end	strand	ID	name	RNA-seq...	RNA-seq...
1	chr1	3204562	3661579	-	NM_0010...	Xkr4	0.0	4.2
2	chr1	4280926	4399322	-	NM_0011...	Rp1	0.0	0.0
3	chr1	4333587	4350395	-	NM_0112...	Rp1	0.0	0.0
4	chr1	4481008	4486494	-	NM_0114...	Sox17	0.0	0.0
5	chr1	4763278	4775807	-	NM_0011...	Mrpl15	108.2452...	150.9735...
6	chr1	4763278	4775807	-	NM_0253...	Mrpl15	31.10137...	74.96923...
7	chr1	4763278	4775807	-	NR_033530	Mrpl15	31.10137...	74.30805...
8	chr1	4797973	4836816	+	NM_0088...	Lypla1	400.6775...	316.2918...
9	chr1	4847774	4887990	+	NM_0011...	Tcea1	16.19486...	14.57176...
10	chr1	4847774	4887990	+	NM_0115...	Tcea1	16.19486...	14.57176...
11	chr1	4848408	4887990	+	NM_0011...	Tcea1	18.08819...	14.57176...
12	chr1	4899656	5009507	-	NM_0213...	Rgs20	0.0	0.0
13	chr1	4899656	5060366	-	NM_0011...	Rgs20	0.0	0.0
14	chr1	5073253	5152630	+	NM_1338...	Atp6v1h	172.7547...	175.5619...
15	chr1	5578573	5596214	+	NM_0110...	Oprk1	0.0	0.0
16	chr1	5578573	5596214	+	NM_0012...	Oprk1	0.0	0.0
17	chr1	5903787	5907479	-	NM_0103...	Npbwr1	0.0	0.0
18	chr1	6204742	6266185	+	NM_0098...	Rb1c		0143...
19	chr1	6349411	6384812	+	NM_0011...	F		
20	chr1	6720131	6851021	+	NM_1738...			
21	chr1	7079000	7163709	+	NM_1830			
22	chr1	8352741	9289811	-	NM_027			
23	chr1	9535488	9537536	+	NM_02			
24	chr1	9538126	9568049	+	NR_02			
25	chr1	9538126	9568049	+	NM_1			
26	chr1	9550913	9621173	-	NR_01			
27	chr1	9591347	9617223	+	NM_1			
28	chr1	9658915	9690290	-	NM_00			
29	chr1	9713277	9738463	-	NM_17			
30	chr1	9788210	9892649	+	NM_177			
31	chr1	9788210	9892649	+	NM_1332.			
32	chr1	9838361	9892649	+	NM_0010			

parameter space exploration

parameter space exploration: motivation

what should be the parameter values?



Parameters

Method for differential test: exact test

CPM cutoff: 1

Normalization method: TMM

MDS plot method: logFC

Number of most significant tags: 20

P-value: 0.05

P-value adjustment method: BH

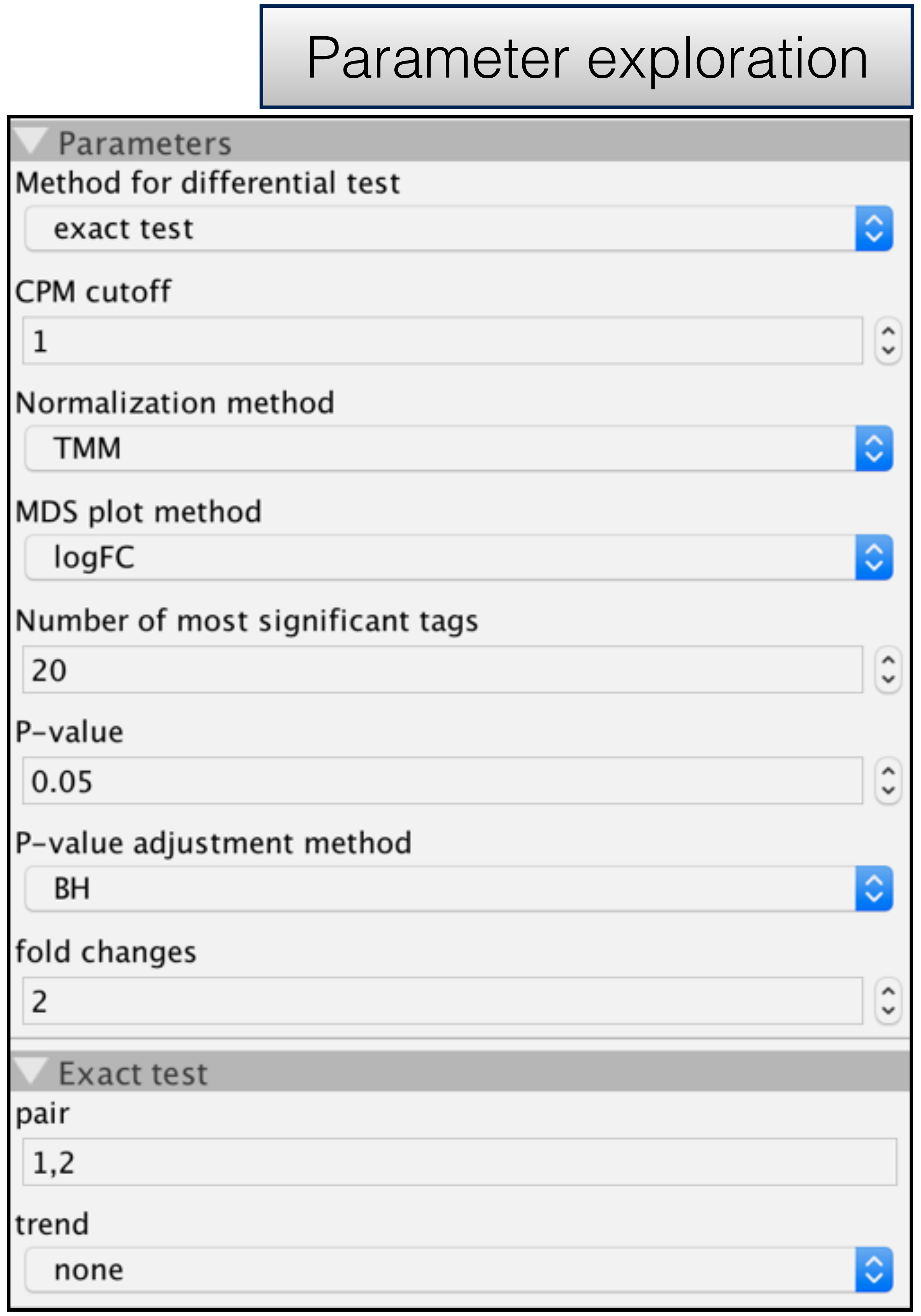
fold changes: 2

Exact test

pair: 1,2

trend: none

parameter space exploration: interface



Method for differential test: exact test

CPM cutoff: 1

Normalization method: TMM

MDS plot method: logFC

Number of most significant tags: 20

P-value: 0.05

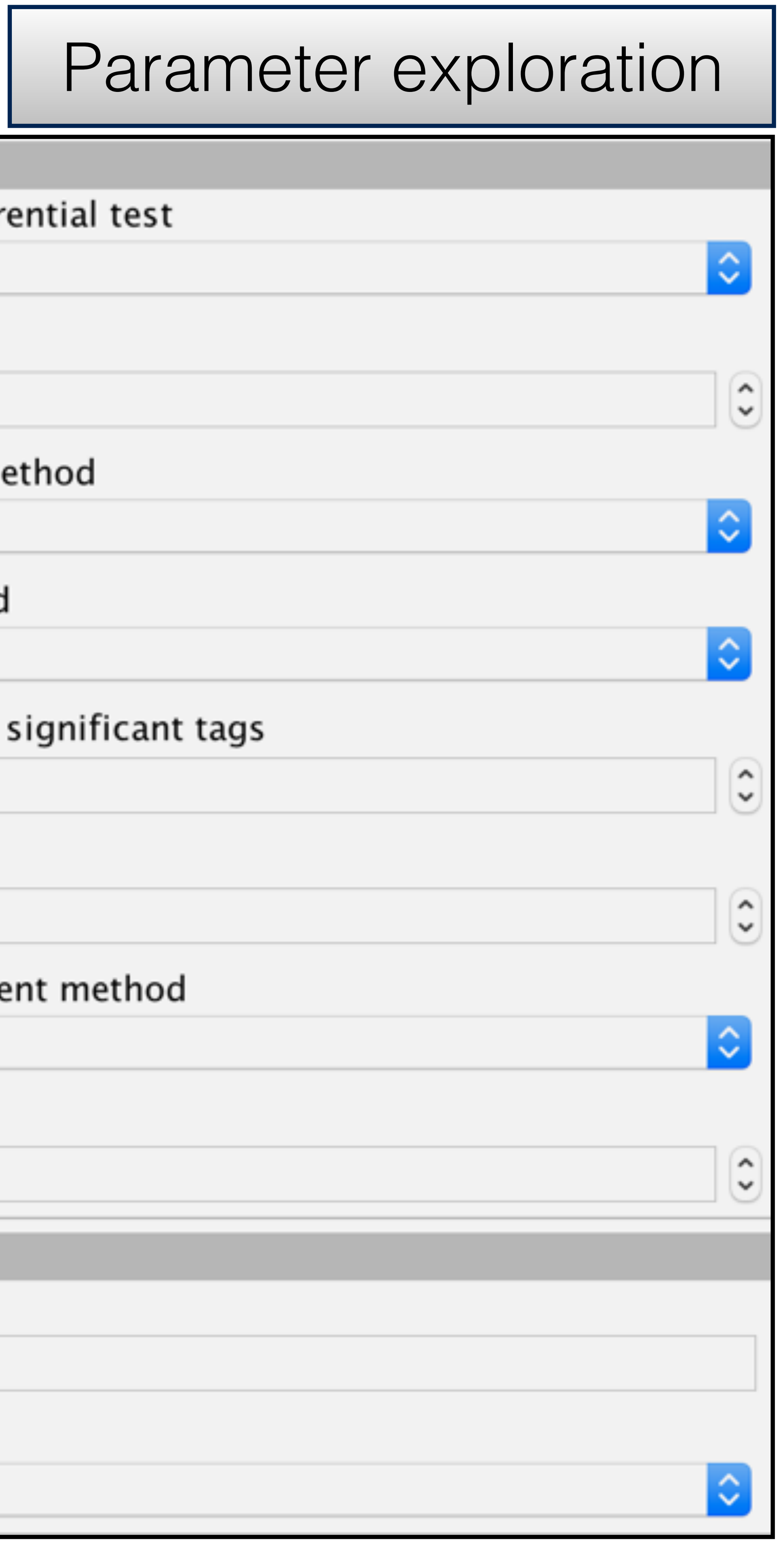
P-value adjustment method: BH

fold changes: 2

pair: 1,2

trend: none

parameter space exploration: interface



Parameter exploration

Parameters

Method for differential test
exact test

CPM cutoff
1

Normalization method
TMM

MDS plot method
logFC

Number of most significant tags
20

P-value
0.05

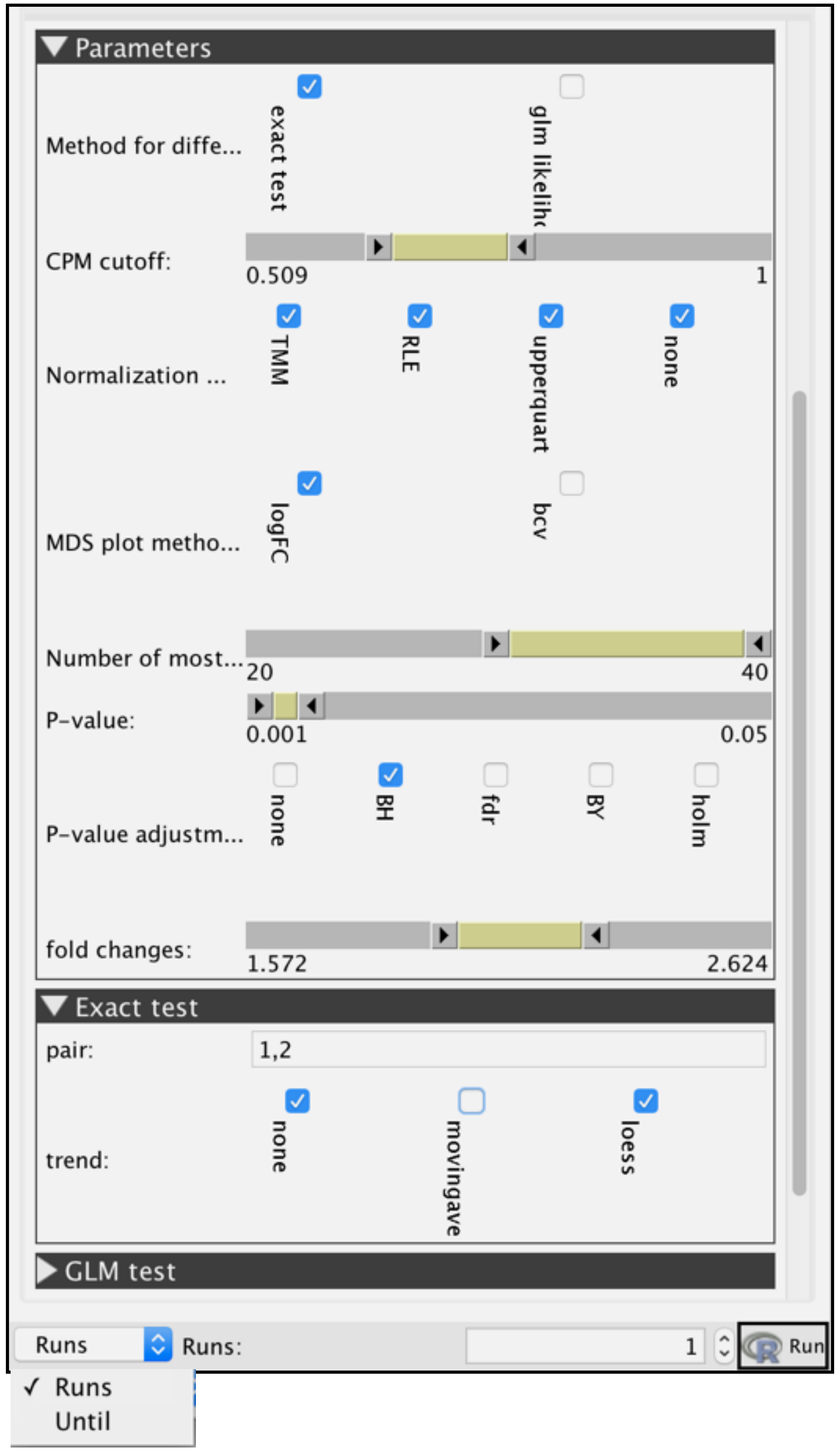
P-value adjustment method
BH

fold changes
2

Exact test

pair
1,2

trend
none



Parameters

Method for diffe... ☑ exact test ☐ glm likelihc

CPM cutoff: 0.509 – 1

Normalization ... ☑ TMM ☑ RLE ☑ upperquart ☑ none

MDS plot metho... ☑ logFC ☐ bcv

Number of most... 20 – 40

P-value: 0.001 – 0.05

P-value adjustm... ☐ none ☑ BH ☐ fdr ☐ BY ☐ holm

fold changes: 1.572 – 2.624

Exact test

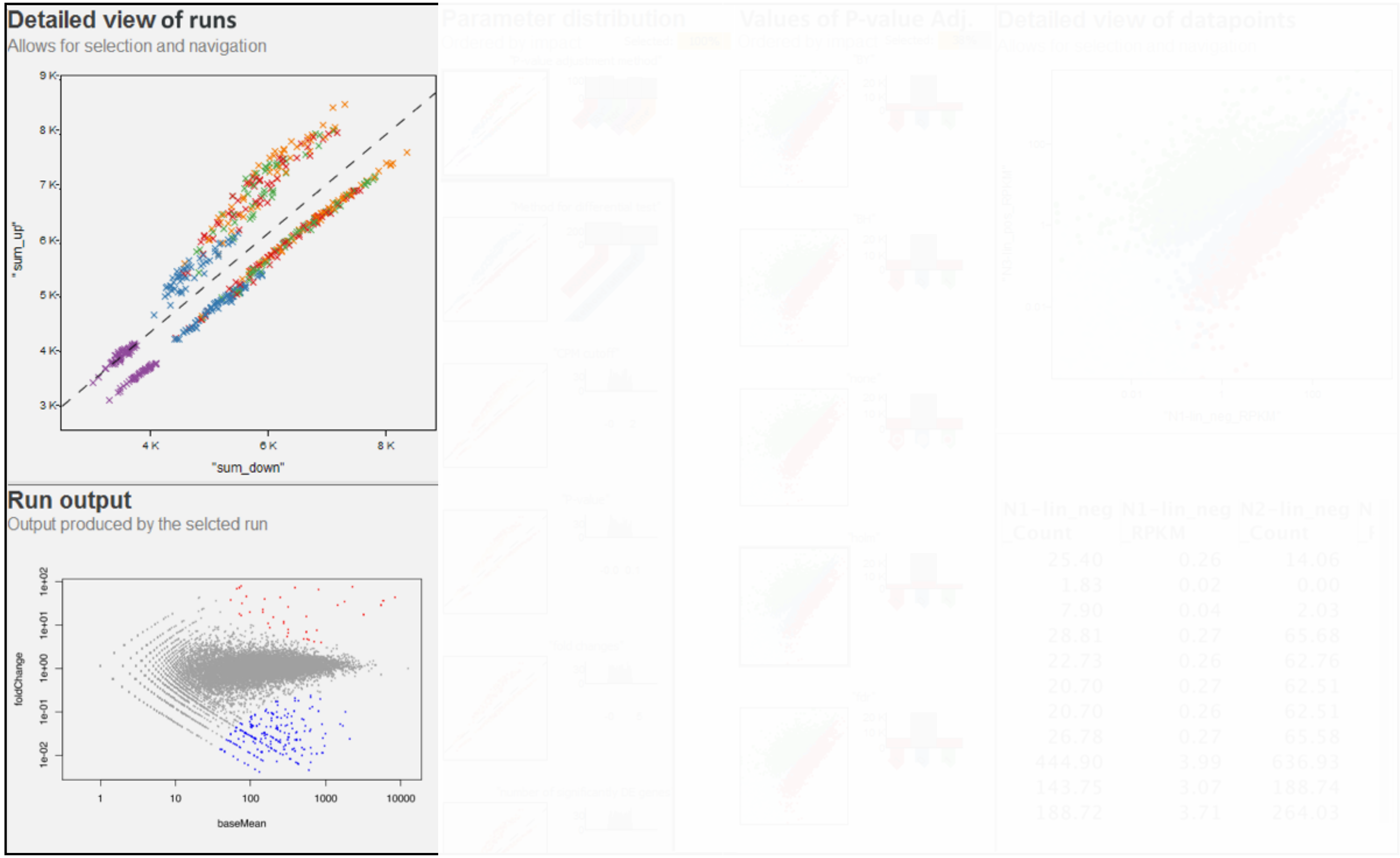
pair: 1,2

trend: ☑ none ☐ movingave ☑ loess

GLM test

Runs: 1

parameter space exploration: result



Detailed view of runs

Allows for selection and navigation

Run output

Output produced by the selcted run

Parameter distribution

Ordered by impact Selected: 100%

Values of P-value Adj.

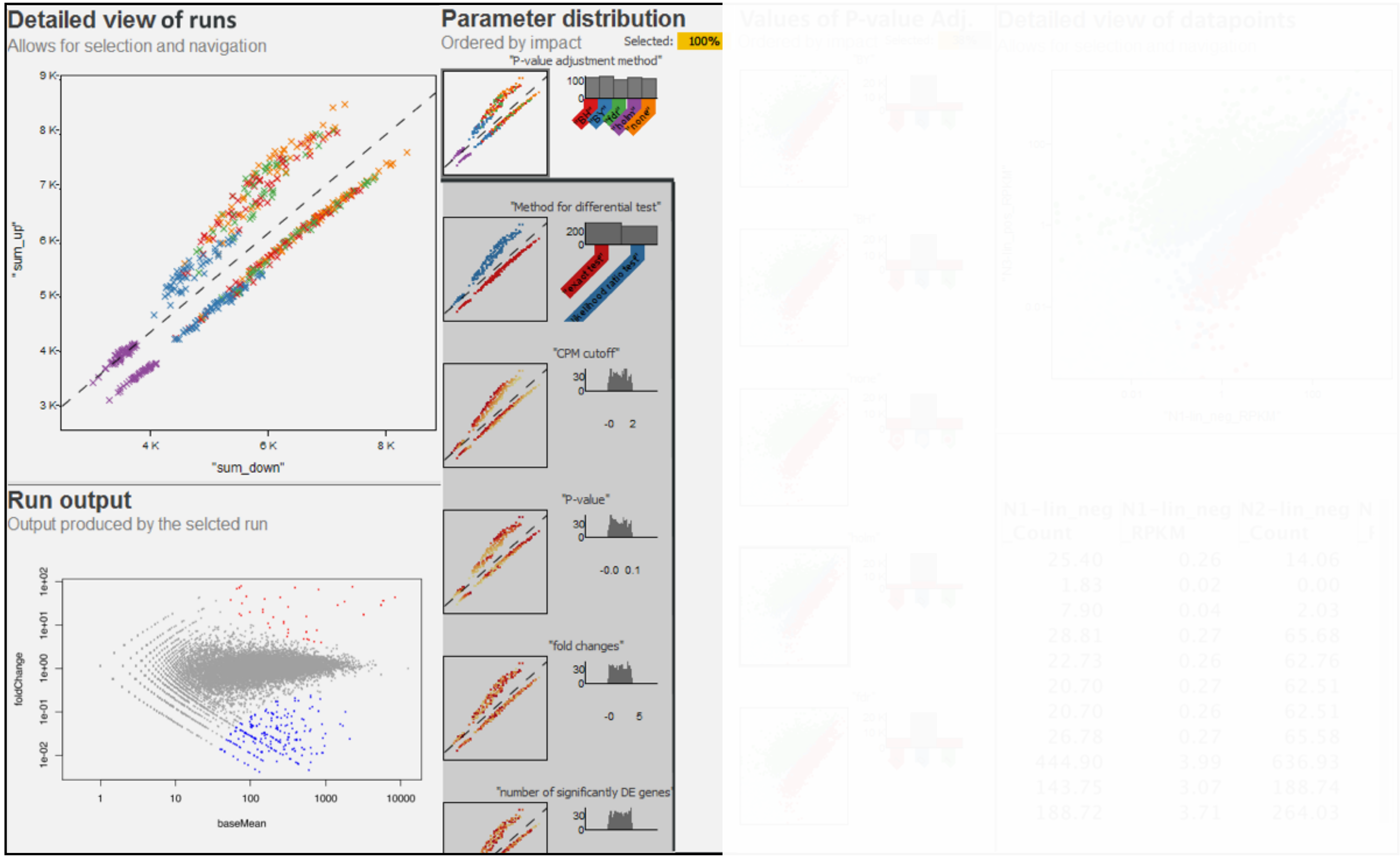
Ordered by impact Selected: 38%

Detailed view of datapoints

Allows for selection and navigation

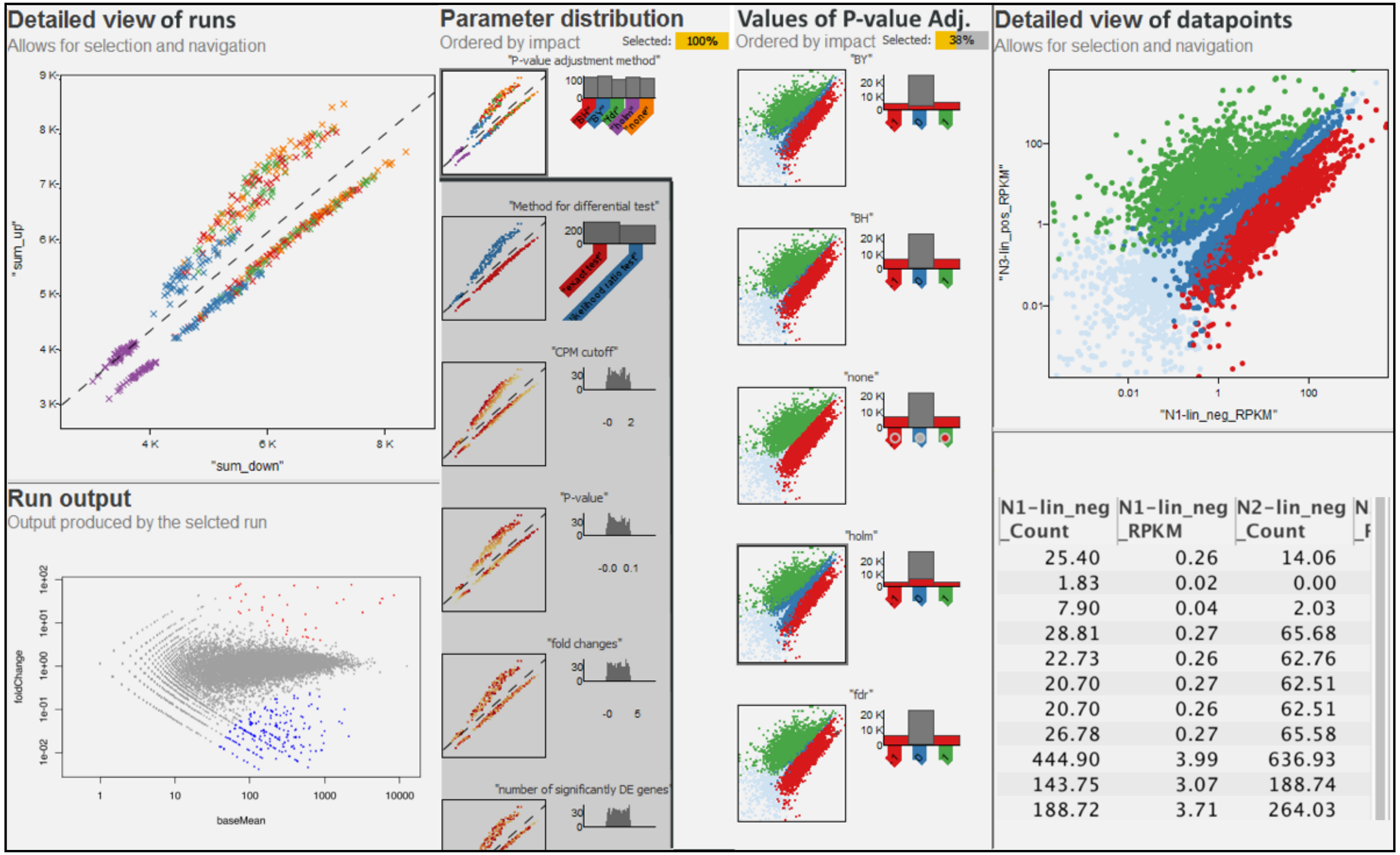
N1-lin_neg _Count	N1-lin_neg _RPKM	N2-lin_neg _Count
25.40	0.26	14.06
1.83	0.02	0.00
7.90	0.04	2.03
28.81	0.27	65.68
22.73	0.26	62.76
20.70	0.27	62.51
20.70	0.26	62.51
26.78	0.27	65.58
444.90	3.99	636.93
143.75	3.07	188.74
188.72	3.71	264.03

parameter space exploration: result



N1-lin_neg _Count	N1-lin_neg _RPKM	N2-lin_neg _Count
25.40	0.26	14.06
1.83	0.02	0.00
7.90	0.04	2.03
28.81	0.27	65.68
22.73	0.26	62.76
20.70	0.27	62.51
20.70	0.26	62.51
26.78	0.27	65.58
444.90	3.99	636.93
143.75	3.07	188.74
188.72	3.71	264.03

parameter space exploration: result



N1−lin_neg _Count	N1−lin_neg _RPKM	N2−lin_neg _Count
25.40	0.26	14.06
1.83	0.02	0.00
7.90	0.04	2.03
28.81	0.27	65.68
22.73	0.26	62.76
20.70	0.27	62.51
20.70	0.26	62.51
26.78	0.27	65.58
444.90	3.99	636.93
143.75	3.07	188.74
188.72	3.71	264.03

conclusions

limitations

- Scalability: few million rows and 30-40 columns of sequencing data
- cannot run R-apps in parallel
- desktop only

future work

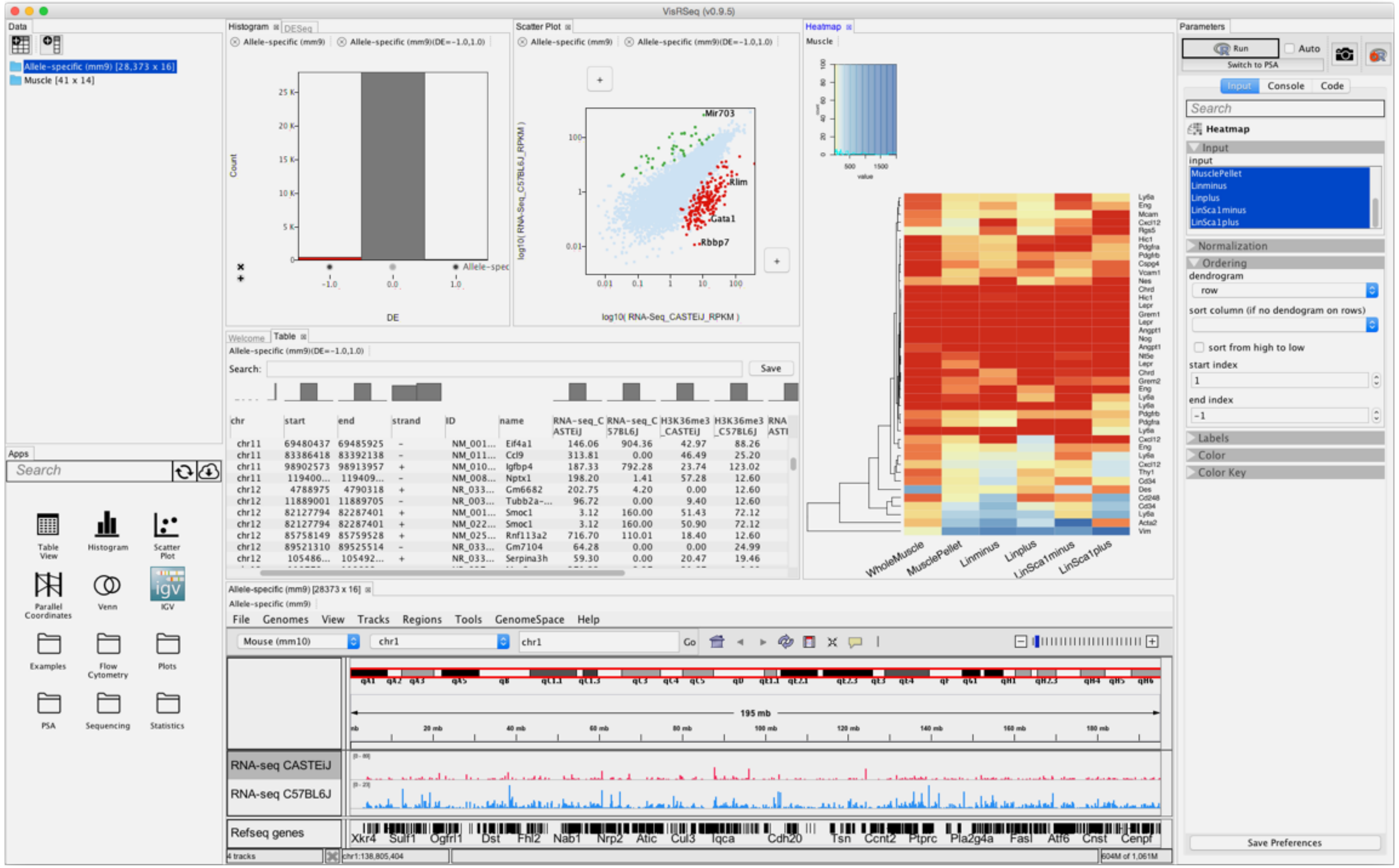
- general purpose framework: Apps for other domains (e.g. Proteomics, etc.)
- workflow designer : link several apps to create “macro” apps.
- reproducibility: provenance / parameter exploration

acknowledgement

- Developers: Xu Fan, Kathy Cheng, Joseph Poper
- BC Genome Sciences Centre
- Lorincz lab, Underhill lab, Rossi lab, Mager lab

SFU gruvi 650 JAHRE universität wien BC Cancer Agency CARE + RESEARCH UBC NSERC CRSNG

questions?



Allele-specific (mm9)(DE=-1.0,1.0)

Search:

chr	start	end	strand	ID	name	RNA-seq_CASTEiJ	RNA-seq_C57BL6J	H3K36me3_CASTEiJ	H3K36me3_C57BL6J	RNA ASTI
chr11	69480437	69485925	-	NM_001...	Eif4a1	146.06	904.36	42.97	88.26	
chr11	83386418	83392138	-	NM_011...	Ccl9	313.81	0.00	46.49	25.20	
chr11	98902573	98913957	+	NM_010...	Igfbp4	187.33	792.28	23.74	123.02	
chr11	119400...	119409...	-	NM_008...	Nptx1	198.20	1.41	57.28	12.60	
chr12	4788975	4790318	+	NR_033...	Gm6682	202.75	4.20	0.00	12.60	
chr12	11889001	11889705	-	NR_003...	Tubb2a-...	96.72	0.00	9.40	12.60	
chr12	82127794	82287401	+	NM_001...	Smoc1	3.12	160.00	51.43	72.12	
chr12	82127794	82287401	+	NM_022...	Smoc1	3.12	160.00	50.90	72.12	
chr12	85758149	85759528	+	NM_025...	Rnf113a2	716.70	110.01	18.40	12.60	
chr12	89521310	89525514	-	NR_033...	Gm7104	64.28	0.00	0.00	24.99	
chr12	105486...	105492...	+	NR_033...	Serpina3h	59.30	0.00	20.47	19.46	

Allele-specific (mm9) [28373 x 16]

Allele-specific (mm9)

visrseq.github.io

(please email for latest version)