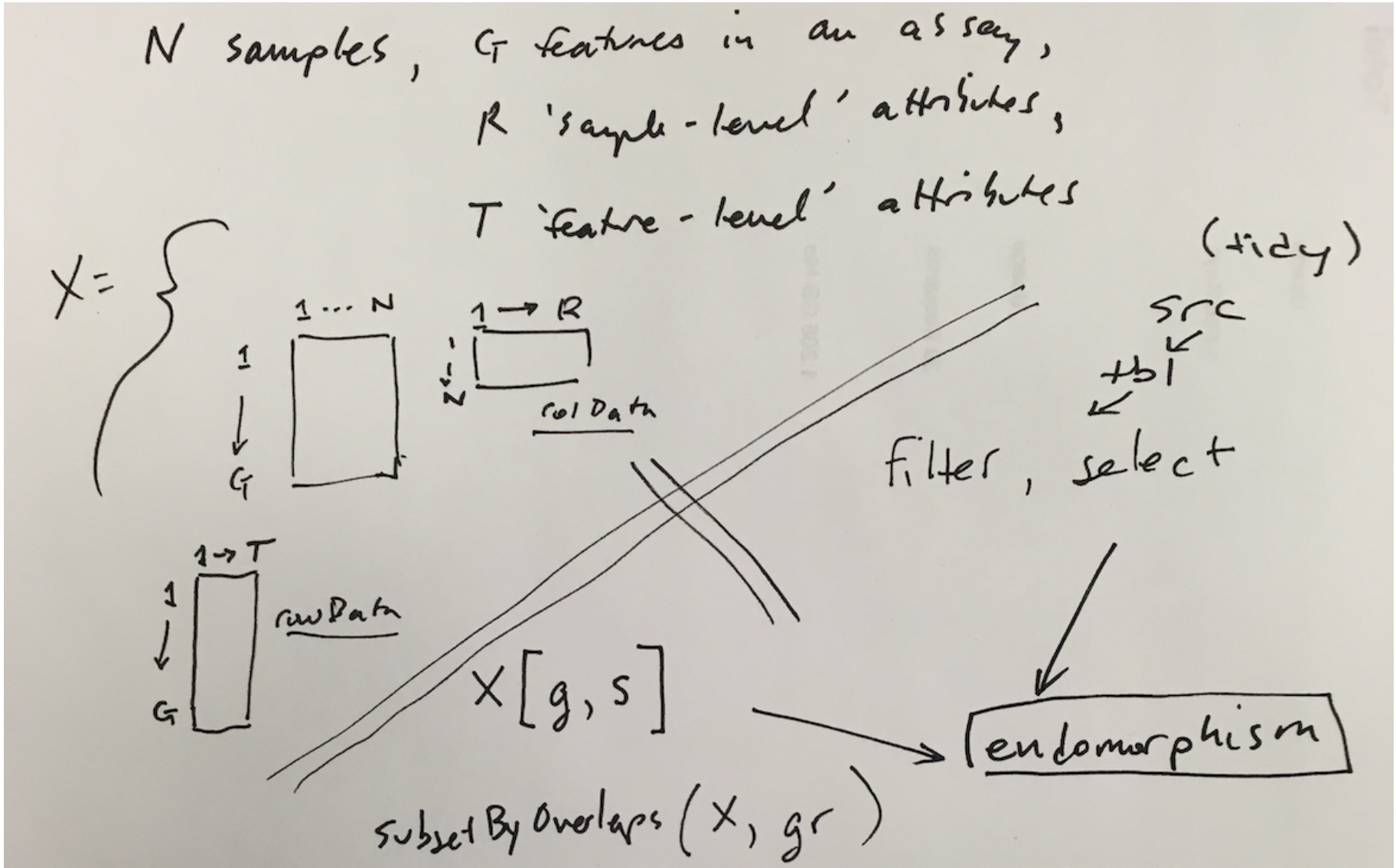
Recent developments in MultiAssayExperiment

Vince Carey stvjc@channing.harvard.edu, reflecting efforts of
Marcel Ramos mramos09@gmail.com,
Lucas Schiffer schiffer.lucas@gmail.com,
Levi Waldron lwaldron.research@gmail.com,
Martin Morgan martin.morgan@roswellpark.org, and others

Rich containers, “tidy” processes



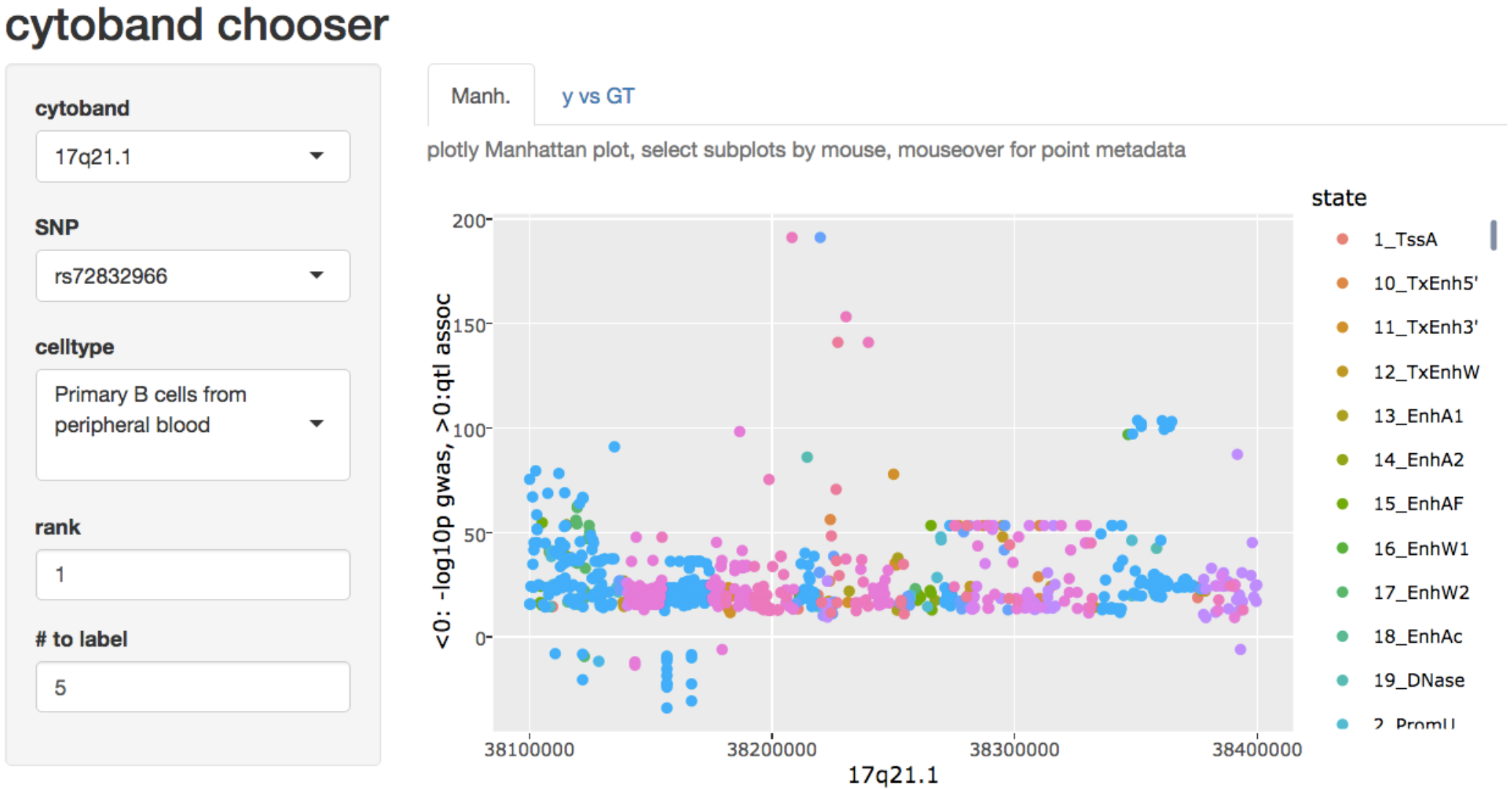
Prototype: YRI Coriell cell lines with RNA-seq, 450K DNA Meth, 1.5M DHS scores, and all 1000 Genome VCF

```
library(MultiAssayExperiment)
API()
# GEUVADIS -- ExperimentHub?
library(geuvPack); data(geuFPKM)
# Gilad 450k
library(yriMulti); data(banovichSE)
# Degner DnaseI hyp. QTL
library(dsQTL); data(DHStop5_hg19)
# 1000 genomes in S3 VCF (in cloud)
library(ldblock); st = stack1kg()
el = ExperimentList(list(YRIexp=geuFPKM,
   YRImeth=banovichSE, YRIdhs = DHStop5_hg19,
   YRIsnp=st))
YRImult = MultiAssayExperiment(el,
   pData=colData(geuFPKM))
```

Show method has hints

```
> YRImult
A MultiAssayExperiment object of 4 listed
 experiments with user-defined names and respective classes.
 Containing an ExperimentList class object of length 4:
 [1] YRIexp: RangedSummarizedExperiment with 23722 rows and 462 columns
 [2] YRImeth: RangedSummarizedExperiment with 329469 rows and 43 columns
 [3] YRIdhs: RangedSummarizedExperiment with 1465442 rows and 50 columns
 [4] YRIsnp: VcfStack with 22 rows and 445 columns
To access:
 experiments() - to obtain the ExperimentList instance
 pData() - for the primary/phenotype DataFrame
 sampleMap() - for the sample availability DataFrame
 metadata() - for the metadata object of ANY class
See also: subsetByAssay(), subsetByRow(), subsetByColumn()
```

library(gQTLstats); example(tqbrowser,ask=FALSE)



AWS S3-resident representations of TCGA

MAEO URLs

File Edit View Insert Format Data Tools Add-ons Help Last edit was made on September 12 by Levi Waldron

Comments Share

MAEO (MultiAssayExperiment Object) URLs

	A	B	C	D	E
1	**MAEO (MultiAssayExperiment Object) URLs**				
2					
3	**Requirement:**	https://bioconductor.org/packages/MultiAssayExperiment/			
4	**Bug reports:**	https://github.com/vjcitn/MultiAssayExperiment/issues			
5	**Tiny URL**	http://tinyurl.com/MAEOurls			
6	**Google Link**	https://docs.google.com/spreadsheets/d/1Ih64DDS5mqDlYFzDyCY9HAUnxvI1b6hapKP_akFuNPY/edit?usp=sharing			
7	**Scripts**	https://github.com/waldronlab/MultiAssayExperiment-TCGA			
8		https://github.com/waldronlab/MultiAssayExperiment-CCLE			
9	**Cohort #**	**Cohort Name**	**Dataset Name**	**MAEO Name**	**MAEO URL**
10	1	Adrenocortical carcinoma	ACC	accMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/accMAEO.rds
11	2	Bladder Urothelial Carcinoma	BLCA	blcaMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/blcaMAEO.rds
12	3	Breast invasive carcinoma	BRCA	brcaMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/brcaMAEO.rds
13	4	Cervical squamous cell carcinoma and endocervical adenocarcinoma	CESC	cescMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/cescMAEO.rds
14	5	Cholangiocarcinoma	CHOL	cholMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/cholMAEO.rds
15	6	Colorectal adenocarcinoma	COADREAD	N/A	N/A
16	7	Colon adenocarcinoma	COAD	coadMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/coadMAEO.rds
17	8	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma	DLBC	dlbcMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/dlbcMAEO.rds
18	9	Esophageal carcinoma	ESCA	escaMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/escaMAEO.rds
19	10	FFPE Pilot Phase II	FPPP	N/A	N/A
20	11	Glioma	GBMLGG	N/A	N/A
21	12	Glioblastoma multiforme	GBM	gbmMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/gbmMAEO.rds
22	13	Head and Neck squamous cell carcinoma	HNSC	hnscMAEO.rds	http://s3.amazonaws.com/multiassayexperiments/hnscMAEO.rds

Conclusions

- MAEs fairly easy to understand, construct
- Hybrid in-memory/out-of-memory representations are coordinated
- Deployed against TCGA and CCLE
- Supplement with annotation components defining feature and sample relationships
- Can get large; HDF5 or other back end schemes under study
- Multi-omic SIG should communicate on strategies and solutions