

Bioc Technical Advisory Board Minutes

4 April 2024

Present: Vince Carey, Laurent Gatto, Wolfgang Huber, Henrik Bengtsson, Helena Crowell, Lori Kern, Robert Shear, Levi Waldron, Alexandru Mahmoud, Erdal Cosgun, Jen Wokaty, Marcel Ramos, Stephanie Hicks, Brian Schilder, Maria Doyle, Hervé Pagès, Kasper Hansen, Ludwig Geistlinger, Davide Risso

Apologies: Michael Love, Charlotte Soneson, Rafael Irizarry, Sean Davis

:03 - :05 Previous [minutes](#) approved

:05 - :08 Follow-ups from previous meetings

- mac build reliability has improved.
- Conference thinking
 - Call for nominations for Bioc Awards
 - Workshop, community/developer day
 - Need volunteers from TAB

:08 - :12 Governance

- Reminder of 70% attendance requirement for TAB members - the seat will be reopened for nomination if not met.
- TAB nomination period will be moved closer to conference date.

:12 - :17 Core package maintenance

- Transition of Biostrings maintenance to Aidan Lakshman - [slides](#) summarizing plans to robustify access to UCSC and Ensembl services, and separate alignment functions from Biostrings to a separate package.
 - Opportunity for managed evolution of core infrastructure packages in general.
 - Need a change review/approval process.
 - Automated evaluation of package state using GitHub actions and other relevant tools.
 - Factoring of Biostrings and rtracklayer.
 - Feature request from Kasper: make additional source-specific columns available in Tx.db packages - will file GH issue.

:17 - :25 gypsum/ArtifactDb with initial example of scRNAseq

After:

```
library(scRNAseq)
ss = BachMammaryData(legacy=TRUE)
ss
```

```
saveObject(ss, path="gyp3") # alabaster.se
```

we have

```
%vjc> tree gyp3
gyp3
├── assays
│   ├── 0
│   │   ├── matrix.h5
│   │   └── OBJECT
│   └── names.json
├── column_data
│   ├── basic_columns.h5
│   └── OBJECT
├── OBJECT
├── row_data
│   ├── basic_columns.h5
│   └── OBJECT
├── row_ranges
│   ├── concatenated
│   │   ├── OBJECT
│   │   ├── ranges.h5
│   │   └── sequence_information
│   │       ├── info.h5
│   │       └── OBJECT
│   └── OBJECT
└── partitions.h5
```

7 directories, 14 files

```
%vjc> h5ls gyp3/assays/0/matrix.h5/compressed_sparse_matrix
data                Dataset {61219942}
indices             Dataset {61219942}
indptr              Dataset {25807}
shape                Dataset {2}
```

:25 - :40 Cloud working group notes

- Bioconductor Hubs storage account security updates.
 - Have implemented Storage Defender for automated malware scanning for ExperimentHub/AnnotationHub/packages. Max file size 2GB, monthly max scan size limit 5TB.
- GitHub actions and [speculative plans](#) for GitHub in build/check/distribute process.
- Strategic plan updates (Thank you Sean Davis!).
- BioC 2024: “Birds-of-a-feather” session proposal submitted.
- Alex Mahmoud investigating GitHub Actions for building Bioconductor packages.

- Henrik highlighted the importance of discussing and coordinating with r-universe (Jeroen Ooms) and not reinventing what they've done. Is an rOpenSci project and funded by the R Consortium ISC grants.
- R-universe YouTube presentation from Bioc Devel Forum:
<https://www.youtube.com/watch?v=JefyrHI63-0>.

:40 - :50 CZI grant funding

- Decisions currently embargoed.
- CZI meeting in Boston in June.
- What are our options for declined projects? NumFocus process is moving forward.