Bioc Technical Advisory Board Minutes

6 May 2021

Attending: Vince Carey, Charlotte Soneson, Laurent Gatto, Stephanie Hicks, Levi Waldron, Aaron Lun, Wolfgang Huber, Shila Ghazanfar, Rafael Irizarry, Hector Corrada Bravo, Aedin Culhane, Robert Gentleman, Lori Shepherd, Nitesh Turaga, Jennifer Wokaty, Hervé Pagès, Marcel Ramos
Regrets: Mike Love, Martin Morgan, Kasper Hansen

:02 - :07) Greatest Hits April 2021
- The core team has done a great job:
    - weathering recent [knitr modifications](#)
    - ingesting/reviewing/deprecating a very large number of packages for 3.13
    - enhancing BiocCheck to improve capacity of developers to meet standards
    - discovering unwitting abuse of the distribution system, negotiating solutions at user end, and engineering BiocManager to reduce risk going forward
    - solving dozens of developer-initiated git problems
    - engineering/documenting a multi-platform build system that can be rapidly replicated by a new core member
    - paving the way for effective use of Bioc in GCP and Azure
    - contributing to developer-targeted capabilities for scalable genomics analysis via efficient matrix representations and advanced approaches to parallelization
- Nominations for Bioconductor Community Awards opened (and the CAB created a page on the Bioconductor website: https://bioconductor.org/about/awards/). Deadline for nominations May 31st.
- April developer forum dedicated to release 3.13 (https://www.youtube.com/watch?v=UWsBcpvsQU8)
- https://code.bioconductor.org/ (by Mike Smith) allows browsing and searching the code of all current Bioconductor software packages. See [slack post](#).
- BiocParallel - discussion on [job specific seeds](#)
- Matrix methods call/lightning talks: [slides](#), [recording](#)
- Updates to the [TAB](#) website. It now includes photos (to edit, please submit a [pull request](#)).
- 3 letters of intent to CZI EOSS Cycle 4 accepted for full submission (Aedin "Diversity/Web site/Training", Vince "Genetics", Vince "Diversity follow-on")
- BioC2021 (Aug 4-6): 126 submissions received. Accepted 48 short talks, 21 package demos, 12 long workshops, 1 birds-of-a-feather, 43 posters. Scheduled from 6am-8pm Pacific time (incl. 2 long mid-day breaks) to accommodate speakers from all timezones. 45 min dedicated to a "Meet the TAB" session (prel. Aug 5, 8:45am Pacific time) - suggestions for structure/topics to discuss?
- The BioC2021 committee has asked the CAB to establish a GDPR/data privacy policy for Bioconductor

:07 - :10) Getting ready for TAB term renewals, new nominations

The voting procedure is currently described as follows in the governance document:

Members, including renewing members, are nominated through a process that includes public solicitation of nominees (Month 1); nomination by current board members (Month 2) and supermajority approval by current board members (Month 3). The solicitation and outcome of the nomination process are intended to be transparent. Discussion is conducted private to the Technical Advisory Board. Voting is private and coordinated by the Secretary.

The executive committee proposes to amend the governance document with the following section:

If the number of nominees achieving a supermajority approval exceeds the number of available positions in the Board (N), nominees will be ranked by the fraction of approval votes and the top N nominees will be elected. Ties that would lead to the number of elected nominees exceeding N will be resolved by a voting round in which each TAB member casts a vote for exactly one of these nominees, and the nominees are ranked by the number of supporting votes.

The proposal was approved by the TAB (11 votes; 11 'yes', 0 'no', 0 'abstain')

Open call for nomination of new members/re-candidature of members whose terms expire (nomination form). The terms of Vince Carey, Levi Waldron, Martin Morgan, Wolfgang Huber, Kasper Hansen and Rafael Irizarry expire on August 1, 2021.

:10 - :15) BiocManager and GitHub Actions (setup-R action)
- R 4.x slack thread. Release = 4.0, devel = 4.2.
- Discuss BiocManager changes in detail (reinstallation should be requested, will not occur by default)

:15 - :20) Status of Release 3.13: Lori Shepherd
- Release schedule
- RELEASE_3_12 branch frozen earlier this week
- James McDonald assisting with annotations
- Cross-checking packages deprecated in 3.12
- On schedule for release in two weeks

:20 - :25) CAB update: Aedin
- First meeting for new members: Daniela Cassol, Katerina Kechris, Estefania Mancini. Kevin Rue-Albrecht, Mike Smith. The CAB has now 15 members. Updated CAB website.
- Governance Issues. Revised doc. CAB voted to amend the CAB governance documents

- - Officers will serve 1 year but may be reelected. Included a clause that requires "Executive officers should have one or more years of regular member experience within the previous 3 years."
    - Approved ranked voting for officer elections
    - There was a discrepancy about founding members term length in the governance docs. Approved motion that founding members may opt to serve a 2 or 3 year term and that the CAB may exceed the recommended number of members during the year 2022.
    - Procedural change: Working groups reports were limiting the time for discussion in CAB meetings. Voted to create an Addendum document to the agenda for working groups reports. This will be approved during the meeting. Items that working groups wish to report or require a vote on will be added to the agenda.
    - Governance doc items needing clarification will be passed to the governance docs committee: 1) clarify regarding the schedule of the "3-month process" for approving new officers. 2) approval process when amending governance docs.
  - New working group - Developer Training to be lead by Mike Smith, Kevin Rue-Albrecht, Lori Shepherd
  - CAB approved proposal for new developer mentorship program. [Version 0.1](#)
  - CZI EOSS application EOSS4-000000020, PI, Aedin accepted. Submitting full application,working with Vince, Levi, CAB and Teaching Committee (Laurent/Charlotte)
  - Launched [Bioconductor Awards 2021](#). Invited Simone to present the awards. Criteria for nomination
    - Outstanding new package contributions
    - Outstanding support site contributions
    - Outstanding documentation contributions
    - Outstanding community engagement and outreach
    - Outstanding long-term contributions to the project
    - Outstanding contributions enabling computational analysis
    - Outstanding contributions enabling the use of new informatic technologies within the project
  - Events
    - H3 Africa, April 28-30 included two Bioc workshops
      - Public data resources and Bioconductor (Sean, Levi, Ben)
      - Bioconductor and Epidemiology for Bioinformaticians (Chloe, Levi)
    - BioC2021: 4 confirmed sponsors ($25K in total), and a number of sponsors who have either contacted us, or displayed interest in supporting BioC2021.
    - [useR](#) 5-9th July, 2021. CAB panel session - volunteers?
    - Community of Bioinformatics Software Developers (CDSB, Spanish/Latin America Bioconductor group) 2021 meeting August 9-13 ([website](#))

:25 - :45) General technical discussion
  - A) Articulation of the service commitment
    - i) git management of ~2000 software packages -- extensive human intervention required for keys, changes of access/maintainer … increase self-service -- IAM?

- ○ ii) high-availability distribution of release binaries (CloudFront-like service) … expensive, Microsoft may absorb, but cheaper pay-as-you-go (e.g., AWS LightSail) should be fully understood and used; Robert G also talking to AWS. Comments:
  - ■ Do a better job of controlling the costs.
  - ■ Users need to access previous releases - to what extent?
  - ■ Packages or Hub data? - Most of it is package downloads
- ○ iii) acceptable cadence of fully synchronized devel build reports -- clearly we have survived with many interruptions to daily build. What would the implications be of changing to weekly or some other frequency?
  - ■ From a developer's perspective, daily builds are very helpful (at least if an update was pushed).
  - ■ Move towards checking the impact on the ecosystem as a whole? Induces complexity.
- ○ iv) Annotation packages -- ingestion, build, distribution
- ○ v) Experiment packages -- ingestion, build, distribution
- ● B) Planning technical evolution
  - ○ i) Docker-based builder configuration
    - ■ a) Step 1: Linux. Need multistage container build or docker compose to keep a light execution image but smooth connection to a heavier build/wordprocessing image
    - ■ b) Prospects for "Windows containers"
  - ○ ii) Role of GitHub Actions
    - ■ a) For developers to get real-time checks on their changes and the impacts of changes
    - ■ b) For build system
  - ○ iii) Kubernetes-based repository generation -- assumes installable images available via BiocManager, but could build from source. Nitesh has these running on GCP and Azure Kubernetes Systems (AKS)

Comments:
- ● Annotation packages special case, not in any automated builds - added directly to repository.
- ● If you run BiocManager::install, it doesn't check the version, and installs lots of packages automatically. Could it check the version?
  - ○ devtools::install_github() does check against commit id to avoid re-downloading, unless force = TRUE, is there an analog for BiocManager::install (just by version number then?)
    - ■ BiocManager 1.30.13 implements version checking and will only re-download a package when force = TRUE.
    - ■ Version numbers are our currency of choice rather than commit IDs which work better on GitHub.
  - ○ Have been discussing this internally. Consider caching packages.
- ● Another suggestion: Docker images (can be run on Kubernetes). Build only Linux packages daily on Kubernetes, Windows/macOS weekly.

- ○ Windows problems are often the most painful to fix - may need an on-demand system to debug on Windows. Interaction with Microsoft may help here.

:45 - :50) Infrastructure committee concept (Levi)
- Mandate: advise and steer the implementation of Bioconductor infrastructure
- Topics:
    - ○ 1. data resource management (including ExperimentHub/AnnotationHub)
    - ○ 2. build system
    - ○ 3. Bioconductor identity (IAM?)
    - ○ 4. website
- Membership:
    - ○ those directly involved in implementation, advisors who are expert on technical implementation, and involved developers/users.
    - ○ project manager
- More frequent meeting (~weekly, rotating topics)
- Volunteers: Lori Shepherd, Nitesh Turaga, Herve Pages, Vince Carey, Jennifer Wokaty, Levi Waldron
- Reach out to: Azure representative (Erdal), Sean Davis, Leo Collado Torres, ...

:50 - :60) Open floor
- 2021-04-01 minutes approved
- Discussion on maximum allowed number of maintainers for a package.
    - ○ Git provides tools for 'social coding'/collaboration even if only one person has rights to push to the Bioconductor git server.
    - ○ Avoid Bioconductor becoming an arbiter in dispute situations.
    - ○ Continue with the rule of one maintainer per package. Summarize reasoning and solutions for collaborative projects in a document for developers. Make sure that procedure for switching maintainers is transparent (include link to procedure).
    - ○ Mention in the upcoming developer training.