Bioc Technical Advisory Board Minutes

1 July 2021

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

:03 - :10 Welcome and information
- Previous meeting minutes approved.
- Next TAB call (Aug 5) will overlap with the 'Meet the TAB session' at BioC2021. Consider an August joint CAB/TAB meeting in the CAB time slot (Aug 12, 10AM Eastern time)?
- TAB members: get in touch with Lori to be registered for the conference.
- DOI minting via direct membership in datacite.org: 2000 Euros/year. Grant or Foundation expense? It was not explicitly budgeted in grant. DOIs minted to date were through a connection at NCI. Seems grant-related (part of the package deliverable), check back with NCI.

:10 - :17 Meet the TAB session at BioC2021
- suggestion 1 - 'Bioc under the hood' (what is behind the surface that most people see). Coordinate with the 'meet the core team' and CAB sessions.
- suggestion 2 - an overview of the projects and working groups driven by TAB members the past year + presentation of TAB members.
- Only Q&A can be difficult given that many attendees have little advance knowledge about the purpose and tasks of the boards.
- Encourage conference participants to suggest high-level discussion topics (in advance). Email in advance with short description of the board. Some form of 'reward' for questions?
- Prepare some questions ourselves to get started.
- 'Mechanistic' aspects of how the TAB works, how nominations work, voting, working groups, TAB activities, e.g. show an agenda to offer concreteness. Point out that minutes are on the website.
- Breakout rooms.

:17 - :20 Other updates
- New TAB membership review
  - Lori Kern new member.
  - Martin Morgan leaves the board.
- Conference workshops are in good shape.
- "Scalability" meeting in september (like the matrix meeting)
  - Heng Li tweet on HDF5: https://twitter.com/lh3lh3/status/1410442786185351168

- Translation initiative -- [BiocManager po-translations](#) branch; [#documentation-translation](#) slack channel

:20 - :30 CAB update
- Identified that there is a need to establish a more comprehensive system for messaging users (e.g. to solicit nominations to CAB/TAB, awards). Will discuss in future meetings. Also how to make the CAB/TAB less intimidating for newer or junior members.
- Discussed developer forum/quarterly lighting talks proposal. Suggested: should there be a website page dedicated to the developer forum (and other events) with short descriptions and important links.
- Discussed the recent posts on reforming the administration of CRAN's package ecosystem (https://github.com/MilesMcBain/cran_change.org/blob/main/letter.md)
  - Report of the discussion: It is good to be aware and Bioconductor already takes some procedures and precautions to mitigate the concerns raised about CRAN.
  - Bioconductor open review and encourages conversations on submitted GitHub issues.
  - Package guidelines and recommendations publicly available (package review committee hopes new updates also will make them easier to find and more public facing).
  - Packages are deprecated before official removal and after automatic failure notifications from the daily builder and individual email outreach from core team member.
  - Source code for all packages are available for view. Just not for direct editing.
  - In progress of governance revamp.
  - We publicly post minutes of TAB/CAB on the website.
- Working group updates:
  - Publication (Susan) - Conference F1000. Japanese F1000 channels. Contact F1000 to see if they will support multiple languages.
  - Webpages with Bioconductor in different languages. Translate BiocManager in other languages. Page from Bioconductor.org linking to resource/white paper Kozo/Leo? New Working group - Multilingual Bioconductor (Sticker Jo).
  - Bioc2021 Awards. Submissions closed June 1, voting closed June 18th. Jo created updated awards sticker and a sticker to celebrate 20 years of Bioconductor.
  - Package Review Working Group. Will open package reviews to community reviewers in the next few weeks. Kevin has made a first draft of bookdown version of the package guidelines. Eventually this will transition into the Bioconductor organization owned [GitHub repository](#). Interested members of the TAB contact Lori. Expanded list of automated checks?
  - Privacy Policy Working Group (Katerina) needs a list of administrators of Bioconductor groups that collect personal information so they can create an inventory of data that is collected: website (Lori), support forum (Lori), Slack (maybe Marcel/Levi), Bioc-devel (Lori), conference registration contact information (Erica, Matt (Asia), Wolfgang/Simone/Davide Risso (Euro)). Mike

added a draft privacy statement text to https://code.bioconductor.org/about.html - emphasizes a need for a consistent privacy message across Bioc.
- ○ Events/Calendar working group. Discussed the minimum requirements for a Bioconductor community event that is added to the calendar. Currently, "Having a Bioconductor focus or being of interest to Bioconductor users/developers" is the only requirement for community events. Are there any other requirements? Discussed role for Bioconductor's CoC compliance - will discuss with CoC committee.
- ○ Education (Kevin). Discussed if there should be a Carpentries section or webpage be added to an Education/training section on the Bioconductor website. Should there be a form for people to apply for Carpentries training?
- ○ BioC 2021. BioC2021 scholarships (free registration) and caregiver scholarships are open. Rolling application, deadline July 1st. Please tweet/advertise BioC2021 caregiver scholarships up to $300 for child or elder care during meeting.
- ○ CZI HCA / Bioconductor Seed Network Symposium. Time: June 15, 9am U.S. Pacific.
- ○ Terra / AnVIL R / Bioconductor Popup Workshops.
- ○ BioC2021 Meet the CAB Aug 6th 17:00 - 18:00 (GMT-4 / Boston / EST)
- ○ BioC2021 BiocAward Ceremony Aug 5th 15:30 - 16:00 (GMT-4 / Boston / EST)

:30 - :35 Comments on r-universe / organizations (Martin)
- ● Strange definition of 'Bioconductor' (repos under www.github.com/Bioconductor); no recognition of devel and release branches; installation instructions don't differentiate Bioconductor versions.
- ● GitHub Actions build-on-demand; an app, rather than file system, behind `contrib.url()` -- scope for metrics, throttling, etc. 'Meta' take-home? -- we're doing things differently, and this consistently causes problems.
- ● Re-envision BiocManager::install() -- narrow scope to secured, managed installation of Bioc packages & dependencies?

:35 - :50 Relevant topics for the TAB to work on - consolidate into achievable goals (smaller scale) and form small groups (can include non-TAB members) who agree on periodic reporting to TAB.
- ● Technical genomics/genetics software:
  - ○ Rafa smallcount effort, glmpca on all 1.3 million neurons with SparseArray or RleArray
  - ○ Spatial transcriptomics working group activities
  - ○ *Hub API review
  - ○ Interfacing to OpenGWAS at UK Bristol
  - ○ Also needed: identify core packages (e.g., S4Vectors), define quality metrics (test coverage etc), milestones for future development
- ● Technical platform software:
  - ○ AnnotationHub/ExperimentHub APIs
  - ○ build system

- ○ single package builder (SPB)
  - ○ gitolite, GitContribution; potential for gitea or other avenues to improve "git self-service"
- Community priorities:
  - ○ Initiate community-based package review (guidelines ported, a little more needed)
  - ○ (Kozo) Hub table-browsing app served globally (shinyapps.io)
  - ○ Example: web site exposure of assets
    - ■ annotationhub.bioconductor.org API
    - ■ enumeration of "projects" (GEO, SRA, ENCODE, TCGA, etc.) with convenient access/curation should be prominent
    - ■ enumeration of datatypes and tasks supported
    - ■ "platforms" that can be used (Google Cloud, AMI, Azure)
    - ■ developer/contributor "assets": what would it be like to create a world map for bioc similar to https://www.encodeproject.org/help/project-overview/?
- Compute platforms for users:
  - ○ recount3 featured workspace on AnVIL
  - ○ first steps on Microsoft Genomics Data Lake
  - ○ dust off the OSCA-book Azure Machine Image
  - ○ costs on Azure
- Infrastructure committee:
  - ○ docker-based builder for Linux
  - ○ specifying system requirements at package level
  - ○ vision of "project dashboard"
- legacy working group issues, project for package reviews

Comments:
- Some things are critical to the project: build system, SPB - technical knowledge limited to a few people. Upgrading.
- Other things are more about "let's make a great project even better" (less directly critical).
- Are there issues related to the build system that affect developers' ability to develop their packages? E.g. cadence of builds (release currently less frequent). Many developers are to some extent managing their own builds via GitHub Actions. If frequency is reduced, the build system is no longer a 'development tool'.
  - ○ 3 levels after a change: (1) local system (simple, easy and fast), (2) GitHub Actions (-> pkgdown website), (3) Bioconductor build system (most important, validated 'finished product', cross-platform). Any change on the Bioc side wouldn't directly impact the development work.
  - ○ Many developers rely on the build system for testing - path from git push to generation of build report (including expected turn-around time) not always clear.
- Working groups can include non-board members - not clear from the current GitHub repo structure (make more visible on the website, easier for people to join). Stephanie and Lori to work on.

- Each working group to come up with a few tangible deliverables (could be discussed in the TAB or another suitable body).
- Galaxy community organization - anyone from the community could apply to launch a working group/propose a topic. Not all working groups have to be led by a TAB/CAB member.
- Make a list of working groups (including non-active/proposed ones in need of a working group leader).