# Bioc Technical Advisory Board Minutes

2 September 2021

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

:02 - :06 Prologue
- Should we actively seek financial contributions/collaboration with industry partners? E.g. via a Bioconductor Consortium with annual corporate memberships. Should enhanced funding be sought from other organizations (e.g., the Moore Foundation)?
  - Additional staff/salary support would be helpful.
  - Value proposition important, what would Bioconductor give back to the funders. E.g., R consortium - ensures continuity for the codebase. Can we take advantage of funding already provided to the R consortium?
  - Reach out to organizations that benefit from Bioconductor.
  - HubMAP RFA ([https://commonfund.nih.gov/hubmap/funding](https://commonfund.nih.gov/hubmap/funding))
- Suppose we were able to assemble $1 million for the foundation, to be distributed in a granting process. What would you apply for? What should the foundation fund?
  - Education. Hire educators and establish programs for more accessible teaching of use and development of Bioconductor.
    - Much of the "training" happens by users googling and finding relevant material on the internet.
    - "Getting started" tutorials on the website and continually improved documentation of the core packages and workflows (instead of unintegrated one-off teaching material that captures the state of the system at one point in time).
    - Community members active on the forum, Stackoverflow, & other relevant fora/social media helping people who have posted questions.
  - Sabbaticals. Buy time of major code contributors for higher-risk development - maybe exploring other languages or computing paradigms.
  - Environment. Enlarging and hardening the build infrastructure.
  - Innovation. Fund developers to assist investigators in implementing new tools for new assay types.
  - More professional/up-to-date looking website.

:06 - :10 Information
- [Previous meeting minutes](#) (July) were approved.
- Joint CAB/TAB meeting suggested for October 14, 10AM Eastern time (coordinate with CAB).
- DOI minting supported by NCI, has resumed for new software packages.

- La Piana governance support project initiated, contract signed.

: 11 - :20 "Greatest hits"
- Kasper's and Dirk's assistance on GHMM
  https://community-bioc.slack.com/archives/C35G93GJH/p1629917057046500
- Mentoring initiative - 7 mentors (Kevin, Aedin, Mike S to review applications), interest from several mentees, If any TAB members wish to join mentorship program, sign up on the form.
- Community reviewers onboarded (2 at the moment, 3 in process of training, 15 more volunteers to eventually process).
- Working groups - Stephanie and Lori in the process of building a more formal website for working groups (classified as active/inactive), community members can suggest working groups.
  - spatial - Stephanie
  - motif - Spencer Nystrom
  - twitterbot (Levi) - project for package-submission-following (tweet new packages submitted, with biocViews) - e.g. preempt objections to package naming
- EOSS funded proposals (https://chanzuckerberg.com/eoss/proposals/?cycle=4):
  - "Adapting limma and edgeR for single-cell and proteomics" (Gordon Smyth).
  - "Bioconductor - high quality training and support for a worldwide community" (Aedin Culhane).
  - "Bioconductor: Sustaining a worldwide community of genome data scientists" (Vince Carey; Diversity and Inclusiveness follow-on to the BBS cycle 2 project).

:20 - :30 Discussion of the "Deep stack"
- "bioc stack" - foundational package such as GenomicRanges, Biostrings, …
- "deep stack" - BBS, SPB, Hub API, and associated technical tools.
  - How much reliability/engineering is needed, how do we measure, who runs covr, etc. … this applies to the bioc stack as well to the "deep stack".
  - Translating vignettes to Rmd.
  - Movement of *Hub to Msoft Genomic Data Lake; *Hub API modifications ready; creating a testing script to be used at various international locations to measure throughput of the AWS and Azure endpoints.
    - New work: Engage an API expert (Quan, thanks to Levi) to help a) transition existing API to OpenAPI/swagger, b) move service to Msoft? c) distribute responsibilities for *Hub contribution maintenance, so that contributors can make their own version upgrades to their contributions.
    - Hub content component: BoF at BioC2021 - not clear who owns what resources, what resources are missing, etc. Aim to construct a 'map' - what's missing, what do we want to bring in? (Mike). Avoid the core team having to add new resources/content. Relationship between ExperimentData packages and ExperimentHub packages - should all be brought into the Hub?
    - Specs for an ExperimentHub API

- - - Containerization talk from Rstudio -- Kevin Ushey has been contacted. Get input from RStudio's roadmap for helping R users in this field (containers, build system, binaries).
  - https://pins.rstudio.com - similarities to BiocFileCache? (Laurent)

:30 - :40 Update from CAB
- Outcomes of BioC2021 conference?
  - Consensus that the banner around slides should have been minimized during talks, and also in published YouTube videos.
- Setting up a structure to begin to formalize events related to CZI grant; reach out to ROpenSci, Galaxy (community manager, Carpentries membership).
- Looking at Crowdin and Transifex for translation of key Bioconductor documents (including CoC).
- Following BioC2021, Matt has created a FAQ for people wishing to join the CAB or become involved in the community. It will be posted on the CAB page on the website.
- Events
  - BioC Asia 2021 (November 1~4)
    - Website source code has been moved to the Bioconductor GitHub organization https://github.com/bioconductor/biocasia2021.
    - Requests for invited speakers are currently underway.
    - The website and "call for abstracts" will start from August 22.
  - New procedure for adding events to the calendar
    - Instruction video ▶ Bioconductor calendar and event group
    - Updating http://bioconductor.org/help/events/ with the consent of the move from email to the GitHub issue form.
- Publications Working Group (Susan)
  - Blog highlighting the most influential papers from the channel at was created for the conference https://blog.f1000.com/2021/08/04/best-of-the-bioconductor-gateway/

:40 - :60 Open floor
- Planning for next release date (late October) is underway.
- Johannes & Laurent propose a mass spectrometry/proteomics/metabolomics working group. Identify community members and priorities, promote usage of dedicated slack channels.
- Summit meeting on scalability proposed
  - Make better use of parallelization (lower level).
  - Aaron is rewriting a lot of the basic single-cell analysis in C++ - figuring out how to wrap it in an R package. https://github.com/ltla/tatami
- Suggestion for a working group to take over further development and maintenance of the OSCA book.
- miaverse (https://microbiome.github.io/) - planning book about microbiome analysis. Basic book infrastructure is quite robust by now. rebook package: https://github.com/ltla/rebook

- Any issues with e.g. basilisk and M1 mac? Relying on conda.
- Working group on containers (part of infrastructure working group?). Containers need testing regularly. Make it an intrinsic part of people's workflows - make it an integrated part of the BBS (Linux, possibly Windows).
    - Issue with the current bioc-devel container: rstudio doesn't support plotting.
- Is there a central location for developers to "give up maintenance" for a package? Currently the bioc-devel mailing list is typically used. E.g. debian lists packages that are "up for adoption".