

News from Bioconductor

by *Bioconductor Core Developers*

Software

Bioconductor version 3.18 was released on Oct 25 2023. The system is compatible with R 4.3. See [the release announcement](#) for full details. Noteworthy additions since our last report to the R Journal include

- a [BSgenome package for the telomere-to-telomere build](#) of the human genome (Aganezov et al. (2022)),
- the [SparseArray](#) package for overcoming the limit on the number of non-zero elements allowable in a sparse Matrix instance,
- [BiocBook](#), a package to facilitate the creation of package-based, versioned online books authored in Quarto,
- a pair of Annotation packages with the the AlphaMissense (Cheng et al. (2023)) pathogenicity scores for coding variants in the human genome for builds [hg19](#) and [hg38](#).

See the [release announcement](#) for full details. The growth of the package repository is greatly aided by a group of committed and energetic reviewers. All reviews are conducted in github issues streams at contributions.bioconductor.org/issues.

Infrastructure

National Science Foundation ACCESS Award BIR190004 provides significant compute resources in the [Jetstream2 academic cloud](#) along with storage provided by the [Open Storage Network](#).

These resources form the basis for the Galaxy/Kubernetes-backed [Bioconductor workshop platform](#) originally known as [Orchestra](#). Workshop submissions are now accepted through a [Shiny app](#) made available at the platform site. The “[BuildABiocWorkshop](#)” template has been updated with GitHub Actions, and now uses the GitHub Container Registry ([ghcr.io](#))

At present, Bioconductor 3.18 packages are tested regularly on Ubuntu 22.04, macOS 13.6 (arm64), macOS 12.7 (x86_64), and Windows Server 2022 Datacenter. Testing of packages in the devel branch includes an arm64 Linux platform (openEuler 22.03) thanks to efforts of Martin Grigorov and Yikun Jiang.

Docker container updates

An active effort to revamp the Bioconductor Docker stack is in progress, maintaining backwards compatibility, but featuring a number of new capabilities. Notably, all containers are now published both on [DockerHub](#) as well as the [GitHub Container Registry](#) (GHCR), so any container previously pulled as `bioconductor/bioconductor_docker` for example, can now also be pulled from GHCR as `ghcr.io/bioconductor/bioconductor_docker`.

Additionally, the traditional rstudio-based containers, previously published under the `bioconductor/bioconductor_docker` name, are now also available under the `bioconductor/bioconductor` name, eliminating the need to type the `_docker` suffix. Moreover, release tags can still be used as “`RELEASE_3_18`” as before, but the simpler “`3.18`” tag now suffices. The latest rstudio-based container can thus now be pulled as `docker pull bioconductor/bioconductor:3.18` and will be identical to `bioconductor/bioconductor_docker:RELEASE_3_18`.

Bioconductor now has containers built on top of different flavors of rocker such as `bioconductor/r-ver` container, a slimmer container with R but not RStudio, and `bioconductor/ml-verse`, featuring tidyverse and some GPU drivers pre-installed. These and more container flavors can be found at [DockerHub](#) and [GitHub](#).

Bioc2u alpha release

Ucar and Eddelbuettel (2021) discuss motivations and methods for distributing pre-compiled R packages via Linux system package managers. Bioconductor has recently

undergone an effort to make a full package repository of Bioconductor packages and their dependencies available via apt on Ubuntu systems. Building on top of previous work from the Debian [r-pkg-team](#) and the [r2u project](#), the Bioc2u repository currently offers 3708 packages for the Bioconductor 3.18 release for Ubuntu Jammy. More information can be found at <https://github.com/bioconductor/bioc2u>, and alpha testers are welcome to join the #bioc2u channel on the Bioconductor [Community Slack](#).

Developer support

Bioconductor is developing containers similar to the Bioconductor Linux build machines. [BBS containers](#) are configured like the build machines and aim to provide a comparable experience for developers to troubleshoot issues observed on the linux build machines. The 3.18 BBS container is available for testing with

```
docker pull ghcr.io/bioconductor/bioconductor_salt:jammy-bioc-3.18-r-4.3.2
```

Future work on the BBS containers will focus on testing the container's performance in comparison to the linux build machine, building the devel container, and incorporating the container in a GitHub Action Workflow.

User support

Thanks to support from the Chan-Zuckerberg Initiative Essential Open Source Software for Science program, the web site at bioconductor.org has been extensively revised.

Partnering with Outreachy

Bioconductor mentored three interns in the May - August 2023 Outreachy cohort. [Outreachy](#) partners with open source and open science organizations to create paid open source internships to individuals underrepresented in technology. The organization, which recently celebrated surpassing 1000 interns, funded the interns for three Bioconductor-mentored projects through their general fund. Interns are selected based on the contributions they make to projects as part of their final application.

Atrayee Samanta, an undergraduate student at IEST Shibpur, India, curated microbiome studies for [BugSigDB](#), a comprehensive database of published microbial signatures. Daena Rys, a computer science student from Cameroon, worked on issues within the [miaverse](#), an ecosystem based on (Tree)SummarizedExperiment for microbiome bioinformatics. Sonali Kumari, an IGDTUW student from New Delhi, India, converted Sweave vignettes in Bioconductor packages to R Markdown for [Sweave2Rmd](#). You can read more about their experiences on the Bioconductor blog at [Our Journey as Outreachy Interns with Bioconductor](#).

Outreachy will also fund three internships with Bioconductor for the December 2023 - March 2024 Outreachy cohort. Chioma Onyido, Ester Afuape, and Peace Sandy of Nigeria will curate microbiome studies for BugSigDB.

References

- Aganezov, Sergey, Stephanie M. Yan, Daniela C. Soto, Melanie Kirsche, Samantha Zarate, Pavel Avdeyev, Dylan J. Taylor, et al. 2022. "A Complete Reference Genome Improves Analysis of Human Genetic Variation." *Science* 376. <https://doi.org/10.1126/science.abl3533>.
- Cheng, Jun, Guido Novati, Joshua Pan, Clare Bycroft, Akvilė Žemgulytė, Taylor Applebaum, Alexander Pritzel, et al. 2023. "Accurate Proteome-Wide Missense Variant Effect Prediction with AlphaMissense." *Science* 381 (September). <https://doi.org/10.1126/science.adg7492>.
- Ucar, İñaki, and Dirk Eddelbuettel. 2021. "Binary r Packages for Linux: Past, Present and Future." <https://arxiv.org/abs/2103.08069>.

Bioconductor Core Developers

Massachusetts, New York, Seattle
bioconductorcoreteam@gmail.com