News from the Bioconductor Project

by Bioconductor Core Team

The Bioconductor project provides tools for the analysis and comprehension of high-throughput genomic data. The 1211 software packages available in Bioconductor can be viewed at http://bioconductor.org/packages/. Navigate packages using 'biocViews' terms and title search. Each package has an html page with a description, links to vignettes, reference manuals, and usage statistics. Start using Bioconductor version 3.3 by installing R 3.3.1 and evaluating the commands

```
source("https://bioconductor.org/biocLite.R")
biocLite()
```

Install additional packages and dependencies, e.g., **AnnotationHub**, with

```
source("https://bioconductor.org/biocLite.R")
biocLite("AnnotationHub")
```

Continued availability of Bioconductor Docker and Amazon images provides a very effective on-ramp for power users to rapidly obtain access to standardized and scalable computing environments.

Bioconductor 3.3 Release Highlights

Bioconductor 3.3 was released on 4 April, 2016. It is compatible with R 3.3 and consists of 1211 software packages, 293 experiment data packages, and 916 up-to-date annotation packages. There are 107 new software packages and many updates and improvements to existing packages. The release announcement includes descriptions of new packages and updated NEWS files provided by package maintainers.

Our collection of microarray, transcriptome and organism-specific *annotation packages* use the 'select' interface (keys, columns, keytypes) to access static information on gene annotations (**org.*** packages) and gene models (**TxDb.*** packages); these augment packages for querying web-based resources. The **AnnotationHub** continues to complement our traditional offerings with diverse whole genome annotations from Ensembl, ENCODE, dbSNP, UCSC, and elsewhere; example uses are described in the AnnotationHub How-To vignette.

User support

The Bioconductor project web site helps orient users and developers to the project. It includes essential information for software installation, detailed landing pages for each package (e.g., https://bioconductor.org/packages/GenomicRanges) including links to current manuals and vignettes, extensive training material, and links to the current literature. A recent innovation has been the development of the Bioconductor F1000 publishing channel for academic publication of work flows and other extended software use cases.

The project support site is a question-and-answer forum where users can easily search for existing solutions or pose specific questions about use of Bioconductor packages. The support site is quite active, with expert responses often within a matter of hours. It is very helpful, when asking about error messages, to ensure that your Bioconductor installation is correct (using BiocInstaller::biocValid()) and current (include the output of sessionInfo() in your question), that the question includes code chunks that someone else can evaluate to reproduce the problem (e.g., using code or data from example pages of package manuals), and that the error message and traceback() output are included.

Bioconductor holds an annual user conference each summer, this year in conjunction with UseR! 2016. Conference resources (talks and workshops) are available.

Developer support

A very natural progression in the R and Bioconductor community is from user to package developer, transforming your knowledge and domain expertise into software that others can use. The Bioconductor web site includes developer resources to help this transition. The Bioconductor developer mailing list provides a forum dedicated to developer-related questions.

New packages are now submitted to Bioconductor using an open review model. Prospective authors develop their package and, when ready, open an issue on the public Contributions github repository. Packages are then built and checked across Linux, Mac, and Windows platforms for conformance to R (R CMD check) and Bioconductor (using the **BiocCheck** package) standards. Once the package is in good shape, a member of the Bioconductor core team performs a preview of the package. The preview identifies technical issues that are not easy to detect automatically.

A key strength of the Bioconductor project is the use of well-defined objects (especially from the **GenomicRanges** infrastructure) to represent data; this encourages software re-use and enables end-user interoperability between packages. For this reason, the technical review often leads to suggestions for data representations and interfaces that use Bioconductor objects rather than general-purpose containers such as a data.frame.

Forthcoming activities

Forthcoming Bioconductor events include an Asian workshop workshop and developer meeting (3-4 November, Brisbane, Australia) and European developer conference (6-7 December, Basel, Switzerland) developer conferences, as well as global training opportunities.

The next Bioconductor release will occur in October, 2016.

Bioconductor Core Team
Biostatistics and Bioinformatics
Roswell Park Cancer Institute, Buffalo, NY
USA maintainer@bioconductor.org