

# Intro to Bioconductor package development

[bit.ly/selin-monbug](https://bit.ly/selin-monbug)

Selin Jessa

MonBUG

March 14<sup>th</sup>, 2018

Why package your code  
Why Bioconductor  
Developing for Bioconductor  
Bioconductor submission  
chromswitch

Why package your code

Why Bioconductor

Developing for Bioconductor

Bioconductor submission

chromswitch

An exploratory workflow might look something like this...

Update functions.R

> source('functions.R')

Update analysis.R

# Why package your code?

- Portability

- Make code more accessible to yourself (in other analyses, on other machines, etc) and share it with others

- Reproducibility

- R package development integrates well with versioning

- Robustness

- Software/package development best practices encourage good habits

# Resources/tools for R package development

- R packages book, by Hadley Wickam
  - <http://r-pkgs.had.co.nz/> (free online)
- devtools
  - R package to facilitate and automate package development technicalities
- usethis
  - Refactoring of devtools, new and improved functionality  
<http://usethis.r-lib.org/articles/articles/usethis-setup.html>

# MVP (minimum viable package)

```
> library(usethis)
Warning message:
package 'usethis' was built under R version 3.4.3
> tmp <- file.path(tempdir(), "myUtils")
> create_package("myUtils")
Changing active project to myUtils
✓ Creating 'R/'
✓ Creating 'man/'
✓ Writing 'DESCRIPTION'
✓ Writing 'NAMESPACE'
✓ Changing working directory to 'myUtils'
> █
```

<https://uoftcoders.github.io/studyGroup/lessons/r/packages/lesson/>

<https://hilaryparker.com/2014/04/29/writing-an-r-package-from-scratch/>

# MVP (minimum viable package)

```
> library(usethis)
Warning message:
package 'usethis' was built under R version 3.4.3
> tmp <- file.path(tempdir(), "myUtils")
> create_package("myUtils")
Changing active project to myUtils
✓ Creating 'R/'
✓ Creating 'man/'
✓ Writing 'DESCRIPTION'
✓ Writing 'NAMESPACE'
✓ Changing working directory to 'myUtils'
>
> use_package("dplyr")
✓ Adding 'dplyr' to Imports field in DESCRIPTION
• Refer to functions with `dplyr::fun()`
> █
```

<https://uoftcoders.github.io/studyGroup/lessons/r/packages/lesson/>

<https://hilaryparker.com/2014/04/29/writing-an-r-package-from-scratch/>



# Package development/research workflow

Update & install  
myUtils package

`> library(myUtils)`

Update analysis.R

Why package your code

**Why Bioconductor**

Developing for Bioconductor

Bioconductor submission

chromswitch

# Why submit to Bioconductor

- Bioinformatics-focused package repository
- Existing “infrastructure” for NGS data and metadata
- Open, in-depth, technical review
- Respected in the community
- Dedicated mailing list for developers and support site for Bioconductor users

Why package your code

Why Bioconductor

**Developing for Bioconductor**

Bioconductor submission

chromswitch

# How to develop for Bioconductor

- Study the open reviews of other packages
  - <https://github.com/Bioconductor/Contributions/issues>
- Re-use existing Bioconductor methods and classes
  - Established tools for data import and storage  
<https://bioconductor.org/developers/how-to/commonMethodsAndClasses/>
- Ensure comprehensive documentation (vignette & exported functions)
  - Live code in vignette and examples
  - roxygen2 for object documentation <http://r-pkgs.had.co.nz/man.html>

# How to develop for Bioconductor

- Write unit tests for robust code
  - <https://bioconductor.org/developers/how-to/unitTesting-guidelines/>
  - <http://r-pkgs.had.co.nz/tests.html>
- Continuous integration
  - <https://docs.travis-ci.com/user/languages/r/>
- Incorporate Bioconductor technical guidelines ASAP
  - BiocCheck: run it early and often  
<http://master.bioconductor.org/packages/devel/bioc/vignettes/BiocCheck/inst/doc/BiocCheck.html>
  - Running R CMD BiocCheck after R CMD check in Travis CI:  
<https://support.bioconductor.org/p/82646/#106919>

# Common Bioconductor reviewer critiques

- Not using existing Bioconductor methods/classes
- For loops instead of vectorized code
- Documentation lacking
- Lack of arguments validation
- No handling of edge cases

Why package your code

Why Bioconductor

Developing for Bioconductor

**Bioconductor submission**

chromswitch



# Before Bioconductor submission

- Package must be in a GitHub repository
- Use git throughout the development process:
  - Easy to set up with devtools/useethis
  - Version control
- Use GitHub throughout the development process:
  - Regular diagnostics with Travis CI
  - Report of test coverage

# Bioconductor submission process

- Submit via an issue on GitHub: <https://github.com/Bioconductor/Contributions>



bioc-issue-bot commented on Sep 22, 2017

Collaborator



Hi @sjessa


Thanks for submitting your package. We are taking a quick look at it and you will hear back from us soon.

The DESCRIPTION file for this package is:

```
Package: chromswitch
Title: An R package to detect chromatin state switches from epigenomic data
Version: 0.99.0
Date: 2017-09-20
Authors@R: c(person("Selin", "Jessa", email = "selinjessa@gmail.com", role = c("aut", "cre")),
  person(c("Claudia", "L."), "Kleinman", role = "aut"))
Description: chromswitch implements a flexible method to detect chromatin state
  switches between samples in two biological conditions in a specific genomic
  region of interest given epigenomic data like ChIP-seq peaks or previously
  learned chromatin state segmentations.
```

# Bioconductor submission process

- Iterate until the build passes with no errors or warnings



bioc-issue-bot commented on Sep 22, 2017 Collaborator + 😊

Dear Package contributor,



This is the automated single package builder at [bioconductor.org](http://bioconductor.org).

Your package has been built on Linux, Mac, and Windows.

On one or more platforms, the build results were: "skipped, ERROR".  
This may mean there is a problem with the package that you need to fix.  
Or it may mean that there is a problem with the build system itself.

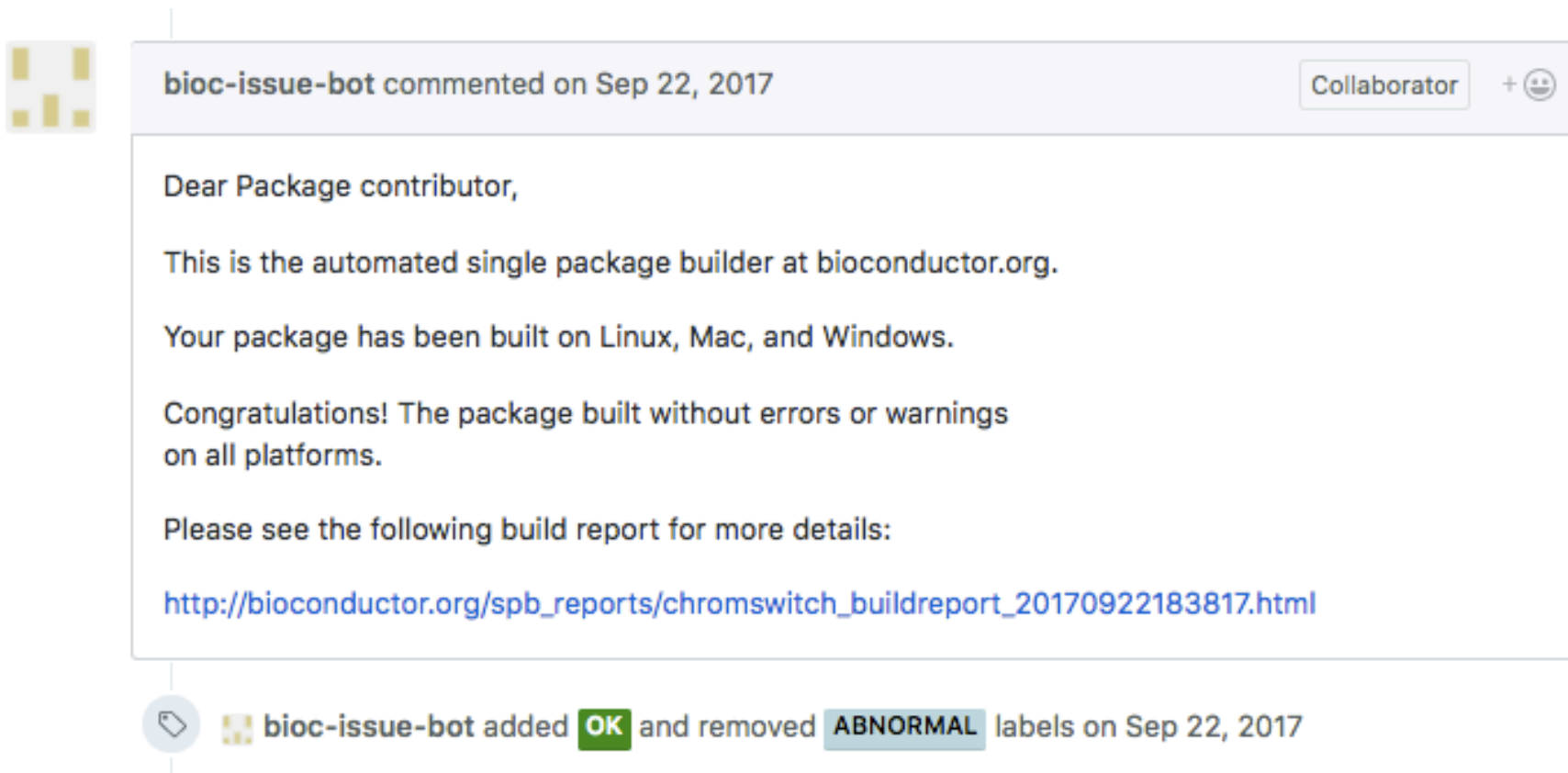
Please see the following build report for more details:

[http://bioconductor.org/spb\\_reports/chromswitch\\_buildreport\\_20170922095812.html](http://bioconductor.org/spb_reports/chromswitch_buildreport_20170922095812.html)

  bioc-issue-bot added the **ERROR** label on Sep 22, 2017

# Bioconductor submission process

- Iterate until the build passes with no errors or warnings



# Bioconductor submission process

- After R CMD BiocCheck passes, review by Bioconductor team member
- Technical, not scientific
- Review of functionality, class implementation, code style, consistency, documentation
- After implementing changes, respond point-by-point
- After acceptance, package is added to Bioconductor dev version
- Release in April and October (next deadline: April 9)

Why package your code

Why Bioconductor

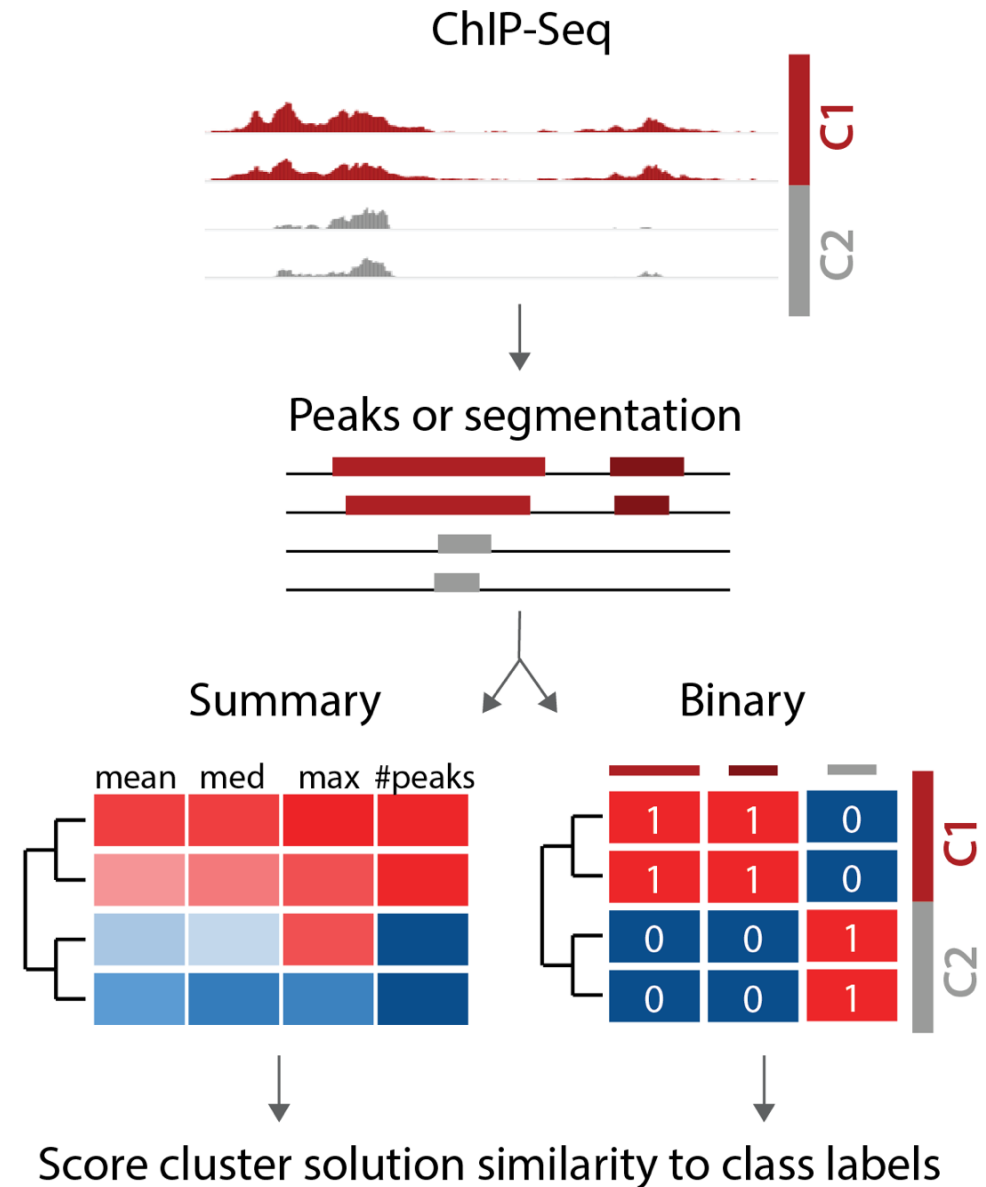
Developing for Bioconductor

Bioconductor submission

chromswitch

# chromswitch

- R/Bioconductor package to detect chromatin state switches  
<https://bioconductor.org/packages/chromswitch>
- Selin Jessa, Claudia L Kleinman (2018). chromswitch: a flexible method to detect chromatin state switches, *Bioinformatics*:  
<https://doi.org/10.1093/bioinformatics/bty075>



# Complete analysis available online

- 1 Introduction
- 2 Preliminaries
- 3 Download epigenomic data
- 4 Benchmark chromswitch performance
- 5 Robustness experiments
- 6 Genome-wide validation
- 7 Session Info

## chromswitch: A flexible method to detect chromatin state switches - *Analysis*

Code ▾

*Selin Jessa and Claudia L. Kleinman*

*11 December, 2017*

### 1 Introduction

This document executes all the analysis presented in *chromswitch: A flexible method for detecting chromatin state switches*, from downloading the data, to running experiments, to generating the tables and figures included in the paper. The dropdown in the top right corner of this HTML document controls whether code is shown or hidden. The document can be navigated using the menu on the left.

<https://github.com/sjessa/chromswitch-analysis>

<https://cdn.rawgit.com/sjessa/chromswitch-analysis/85b605c105df1e34eabe1603a92aea14c122c9ce/analysis.html>



# Take home messages

- Writing an R package: not as daunting as it seems
- Packing your code saves your future self time, encourages reproducibility and robustness
- Use Bioconductor requirements as a guide during the package development process

# Acknowledgements

Kleinman Lab

**Claudia Kleinman**

Nicolas De Jay

Steven Hébert

Matt D'Iorio

Yang Yang

Nicholas Anthony

Karine Choquet

Maud Marques

Alexis Blanchet-Cohen



# Links



[Home](#)

[Install](#)

[Help](#)

[Home](#) » [Bioconductor 3.6](#) » [Software Packages](#) » chromswitch

## chromswitch

platforms **all** downloads **available** posts **0** in Bioc **< 6 months**  
build **ok**

DOI: [10.18129/B9.bioc.chromswitch](https://doi.org/10.18129/B9.bioc.chromswitch) [f](#) [t](#)

An R package to detect chromatin state switches from epigenomic data

[bioconductor.org/packages/chromswitch](https://bioconductor.org/packages/chromswitch)

[bit.ly/selin-monbug](https://bit.ly/selin-monbug)