

Syllabus Title

Bioconductor for Genomic Data Science

Instructor

Kasper Daniel Hansen, Ph.D

Description

This class will provide an introduction to the Bioconductor project (<http://bioconductor.org>). The class will cover how to install and use Bioconductor software. We will discuss common data structures, including ExpressionSets, SummarizedExperiment and GRanges used across several types of analyses. Several popular packages for analysis and annotation of your data will be covered, including packages such as GenomicRanges, biomaRt, rtracklayer, AnnotationHub and others.

Course Content

Week 1

Introduction and pre-requisites

What is Bioconductor

Installing Bioconductor

The Bioconductor Website

Useful Online Resources

R Base Types

Core Bioconductor functionality: GRanges and AnnotationHub

IRanges

GRanges from GenomicRanges (Definition, Usage, seqinfo)

AnnotationHub

Usecase - AnnotationHub and GRanges

Week 2

Biological strings in Bioconductor

Biostrings

BSgenome

Intermediate usage of GenomicRanges

Lists

Rle

Transcription annotation in Bioconductor

GenomicFeatures

Importing data using rtracklayer

Week 3

Data containers and annotations

ExpressionSet

SummarizedExperiment

Getting data online

GEOquery

biomaRt

R programming

S4 Classes

S4 Methods

Week 4

Handling raw and aligned short reads.

Basic usage of statistical models for differential expression.


Preprocessing microarrays.

Pre-requisites




We assume some familiarity with the R language.

A great introduction to R, is the class R programming taught by Roger Peng.

Resources

For installation of software, see [Software Installation](#) 

Other sources of help

[Rstudio](#) , [R project](#) , and [Bioconductor](#) 

[Rstudio's cheatsheets](#) 

[PH525x Statistics for Genomics Book](#) 

Weekly quizzes

There are four weekly quizzes. You may begin submitting them as soon as the course opens. Quiz 1 is due at the end of the first week, Quiz 2 is due at the end of the second week, Quiz 3 is due at the end of the third week, and Quiz 4 is due at the end of the fourth week.

Quiz Scoring

You may attempt each quiz up to 3 times in 8 hours. The score from your most successful attempt will count toward your grade.

Grading policy

You must receive a final grade of 70% or better on each assignment (quizzes and project) to pass the course.

Your final grade will be calculated as follows:

Quiz 1 = 25%

Quiz 2 = 25%

Quiz 3 = 25%

Quiz 4 = 25%

Differences of opinion







Keep in mind that currently data analysis is as much art as it is science - so we may have a difference of opinion - and that is ok! Please refrain from angry, sarcastic, or abusive comments on the message boards. Our goal is to create a supportive community that helps the learning of all students, from the most advanced to those who are just seeing this material for the first time.


Plagiarism

Johns Hopkins University defines plagiarism as "...taking for one's own use the words, ideas, concepts or data of another without proper attribution. Plagiarism includes both direct use or paraphrasing of the words, thoughts, or concepts of another without proper attribution." We take plagiarism very seriously, as does Johns Hopkins University.

We recognize that many students may not have a clear understanding of what plagiarism is or why it is wrong.

Please see the following guide for more information on plagiarism:

<http://www.jhsph.edu/academics/degree-programs/master-of-public-health/current-students/JHSPH-ReferencingHandbook.pdf> 

It is critically important that you give people/sources credit when you use their words or ideas. If you do not give proper credit -- particularly when quoting directly from a source -- you violate the trust of your fellow students.

The Coursera Honor code includes an explicit statement about plagiarism:

I will register for only one account. My answers to homework, quizzes and exams will be my own work (except for assignments that explicitly permit collaboration). I will not make solutions to homework, quizzes or exams available to anyone else. This includes both solutions written by me, as well as any official solutions provided by the course staff. I will not engage in any other activities that will dishonestly improve my results or dishonestly improve/hurt the results of others.