

# 生物信息学本科课程R语言部分教学材料

---

理论课部分：

## 第五章 R 语言与Bioconductor 在生物信息学中的应用

- 重点：R 语言安装和运行、变量、语法和函数；Bioconductor 与R 程序包。
- 难点：R 语言变量、语法和函数。

教学内容：

2019-11-04 3-5节

1. 掌握R 语言安装和运行。
2. 掌握Bioconductor 与R 程序包。
3. 掌握R 语言变量、语法和函数。
  1. 系统变量、
  2. 用户自定义变量、
  3. 数据类型（布尔型、数值型、字符型、列表、向量、矩阵、数据框、复杂类型等）、
  4. 函数定义和使用（包括系统函数和用户自定义的函数）

教学资源

- [R语言环境的准备](#)
- [（视频）R语言环境的准备](#)
- [R的初级知识](#)
- [矩阵，因子，数据框，流程控制，for循环，apply](#)
- [（视频）R语言基础：向量，矩阵，数据框和列表](#)

## 第六章 ggplot2 在生物信息绘图中的应用

2019-11-07 3-5节

- 重点：使用ggplot2 绘制简单图形与复杂图形。
- 难点：使用低级绘图函数进行图形精确控制。

教学内容：

1. 掌握使用ggplot2 绘制简单图形。
  1. 学习安装并使用ggplot2。了解以下概念：
    1. 数据（Data）和映射（Mapping）
    2. 几何对象（Geometric）
    3. 统计变换（Statistics）
    4. 标度（Scale）
    5. 坐标系（Coordinate）
    6. 分面（Facet）和图层（Layer）

2. 学习用ggplot2 绘制条形图、柱状图、散点图、箱线图、折线图、饼图等。
2. 了解使用ggplot2 绘制复杂图形。
  1. 学习使用ggplot2 将多个单图进行组合, 学习绘制热图、GSEA 富集图、网络图等复杂图形。

## 教学资源

### ggplot课程代码

## 第七章 二代测序数据处理一般过程

2019-11-11 3-5

- 重点：二代测序原理、常用软件和一般过程。
- 难点：使用HISAT、StringTie、Ballgown 等软件进行数据处理。

### 教学内容：

1. 熟悉二代测序原理和常用软件。
  1. RNA-seq 技术
  2. 了解RNA-seq 常用的分析软件FastQC、HISAT、SAMTools、StrignTie、Ballgown 等。
2. 掌握二代测序数据处理一般过程。
  1. 下载比对参考基因组文件, 然后构建索引。
  2. 利用索引将reads 拼接到参考基因组上。
  3. 下载基因注释文件, 对拼接到基因组上的reads 进行基因注释
  4. 合并转录本, 达到表达矩阵。

---

### 实验课部分

教材：高山, R 语言与Bioconductor 生物信息学应用、ISBN : 9787543333604.天津：天津科技翻译出版公司, 2014 年, 第1 版。

### 注意：

1. 我负责的实验课部分, 需要书写纸质版实验报告, 同时也需要在问卷星中按照要求提交电子文档；
2. 请不要抄袭其他人的作业；

## 实验二 生物信息分析环境的搭建

11-28, 1-3 学时：3

### 实验说明

## 实验三 R 语言变量和基本语法

12-5, 1-3 学时：3

### 实验说明

## 实验四 R 语言函数和程序包

12-6, 1-3, 学时：3

[实验说明](#)

## 实验五 生物信息简单统计分析

12-12, 1-3, 学时 : 3

[实验说明](#)

## 实验六 生物信息基本图形绘制

12-13, 1-3, 学时 : 3

[实验说明](#)

## 实验七 差异表达分析

12-19, 1-3, 学时 : 3

[实验说明](#)