

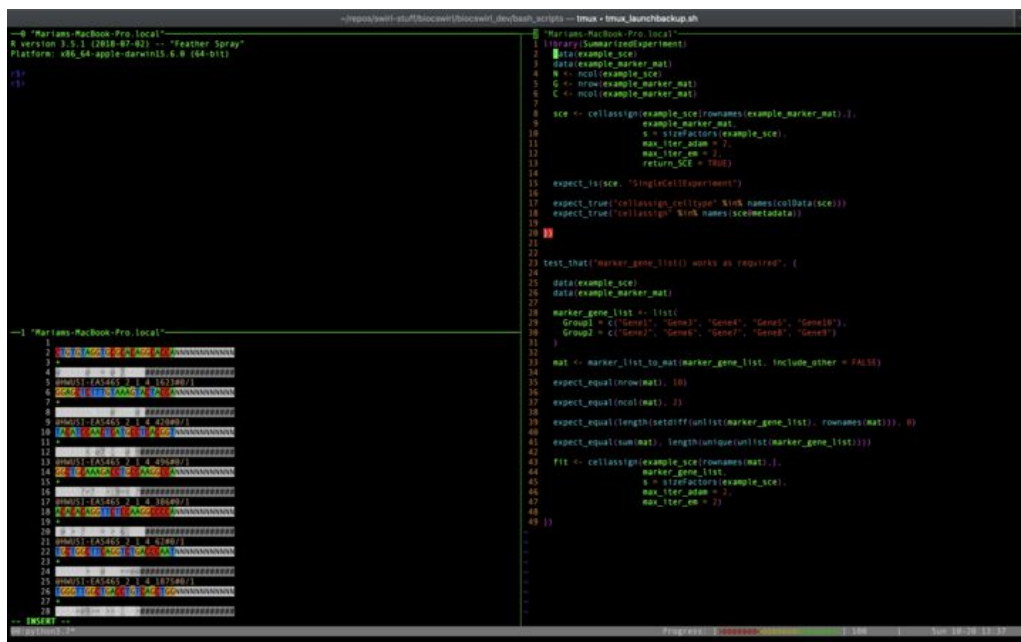
# BIOINFORMATICS TUTORIALS & TOOLS

The interface (BiocTerm) and R package (BiocSwirl) can be used independently of each other but are best used together for people who are interested in making the most of their learning experience.

## BiocTerm

BiocTerm is a standalone terminal application that acts as the ideal interface for those conducting bioinformatics as it integrates our custom editor, biovim (a vim installation bundled with bio plugins), radian (an r console), and tmux (our customized terminal wrapper) in a contained instance. Our interface supports R, Bash, and vim. It is ideal for multi language workflows.

- + Automatically saves session info + restores and generates a log
- + Multiple panes for native file viewing, has robust syntax highlighting that is ideal for bioinformaticians such as BioSyntax or Radian
- + Tmux.conf file and plugin capabilities for vim and tmux allow for high configurability depending on what your needs are
- + Comes with our editor biovim, a powerful editor that makes reading gene files easy



```
--@ "Marius-MacBook-Pro.local"
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Platform: x86_64-apple-darwin13.6.0 (64-bit)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

library(SummarizedExperiment)
data(example_sce)
data(example_marker_mat)
n <- ncol(example_sce)
C <- rownames(example_marker_mat)
C <- mcol(example_marker_mat)

sce <- cellassign(example_sce[rownames(example_marker_mat).],
  example_marker_mat,
  n = sizeFactors(example_sce),
  max_iter_adam = 7,
  max_iter_fm = 2,
  return_SCE = TRUE)

expect_is(sce, "SingleCellExperiment")
expect_true("cellassign_celltype" %in% names(colData(sce)))
expect_true("cellassign" %in% names(rowData(sce)))

test_that("marker_gene_list() works as required", {
  data(example_sce)
  data(example_marker_mat)
  marker_gene_list <- list(
    Group1 = c("Gene1", "Gene2", "Gene3", "Gene4", "Gene5"),
    Group2 = c("Gene6", "Gene7", "Gene8", "Gene9", "Gene10")
  )
  mat <- marker_list_to_mat(marker_gene_list, include_other = FALSE)
  expect_equal(nrow(mat), 10)
  expect_equal(ncol(mat), 3)
  expect_equal(length(unique(unlist(marker_gene_list, rownames(mat))), 0)
  expect_equal(sum(mat[, length(unique(unlist(marker_gene_list)))]
  fit <- cellassign(example_sce[rownames(mat).],
    marker_gene_list,
    n = sizeFactors(example_sce),
    max_iter_adam = 7,
    max_iter_fm = 2)
})
```

# BiocSwirl

Our `swirlify()` R package and course installation client, used to deliver our interactive courses and can be run within R console. Biocswirl aims to make learning bioinformatics concepts hands on through the development of course material that takes you through the common bioinformatics workflows.

- + Is easily configurable and breaks down complex bioinformatics workflows into simple, easy to chew steps. Includes standardized datasets to work with and checks your work
- + Highly emphasizes good coding practices, open science, reproducibility, and
- + YAML file format and templates make it easy to update a workflow to match current best practices and create/communicate your own workflows for within lab

```
Console Terminal x
C:/Users/Kate96/Desktop/BiocSwirl/ ↗

|=====| 20%
| There are many open source and commercial alignment tools with varying
| sensitivities and speeds. which aligner you choose to use depends on available
| computing power and the acceptable trade-off between accuracy and speed. Commonly
| used aligners include BWA, Bowtie2, HISAT2, Bfast, and Stampy.
...

|=====| 40%
| In our sample data, 100bp single-end reads were aligned using RSEM to the mm10
| mouse genome build with the RefSeq annotation downloaded on 11 June 2013. Raw
| fastq files are available at Gene Expression Omnibus, accession ID GSE71585.
...

|=====| 60%
| The standard output of alignment packages is a BAM file. Before downstream
| processing of aligned reads (.bam) in R, we must assess the quality of the
| alignments.
...

|=====| 80%
| What tool can we use to quality check the aligned .bam files?

1: STAR
2: samtools
3: DESeq2
4: FastQC

selection: 4

| You are doing so well!

|=====| 100%
| Go to https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ and follow instructions to perform QC on .bam
...|
```

