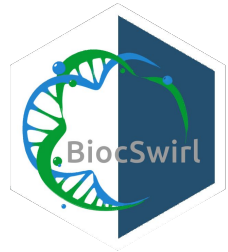


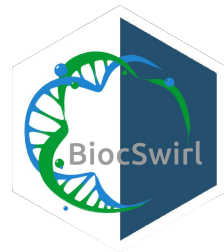
Introduction to R and RNAseq

with BiocSwirl



Zoom etiquette

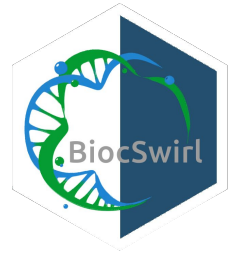
- Put your preferred pronouns into your zoom username if you'd like
- Keep camera on if possible - that will allow us to “read the room”
- Keep yourself muted to reduce background noise, but feel free to unmute and ask questions anytime
- We encourage questions on the etherpad over questions in the chat!
- https://etherpad.wikimedia.org/p/BiocSwirl_RNA
- Closed captions available!
- Let us or Jenelys know of any accessibility needs
- This workshop is recorded, recording will be made available at the end



Agenda

- Who we are
- What this workshop is about (and what it isn't about)
- Part I - Introduction to R
- Break!
- Part II - Introduction to RNAseq and RNAseq data analysis
- What is Bioconductor & BiocSwirl?
- Part III - RNAseq data analysis exercise using the BiocSwirl platform & live demos
- Open Science
- Post workshop assessment

Meet Team BiocSwirl - @bioctools



Lisa N. Cao

Simon Fraser
University

@lisancao



Matt A. Moss

Zucker School of
Medicine

@MattMoss5595



Almas Khan

University of
British Columbia

@VancityAlmas



Julia Philipp

UC Santa Cruz

@its_JPhilipp



Jasdeep Singh

Punjab Agricultural
University

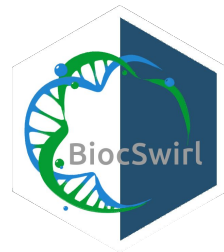
@Jas_deepSingh



Sourav Singh

Vishwakarma
Institute of
Information
Technology

@MrSouravSign

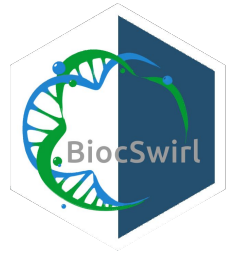


What this workshop is and isn't

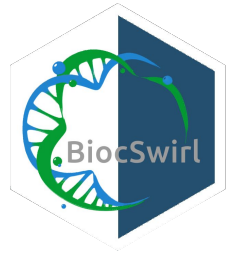
- A brief introduction to coding in R
 - A brief introduction to RNAseq
 - A brief introduction to differential expression analysis in R
 - An introduction to learning tools
-
- Not a complete guide to planning and analyzing RNAseq experiments

Pre-workshop assessment

Include link here



Part I - Introduction to R

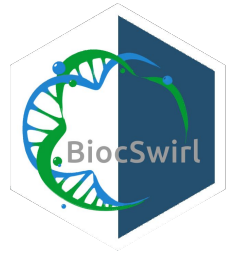


Update R packages



```
update.packages(ask = FALSE)
```


Break time!



Team BiocSwirl - @bioctools



Jasdeep Singh

@Jas_deepSingh

Matt A. Moss

@MattMoss5595

Lisa N. Cao

@lisancao

Sourav Singh

@MrSouravSign

Twitter: @DIAS_org1

Mailing list:

DIASeboard@cshl.edu

See you at 1:25pm!!

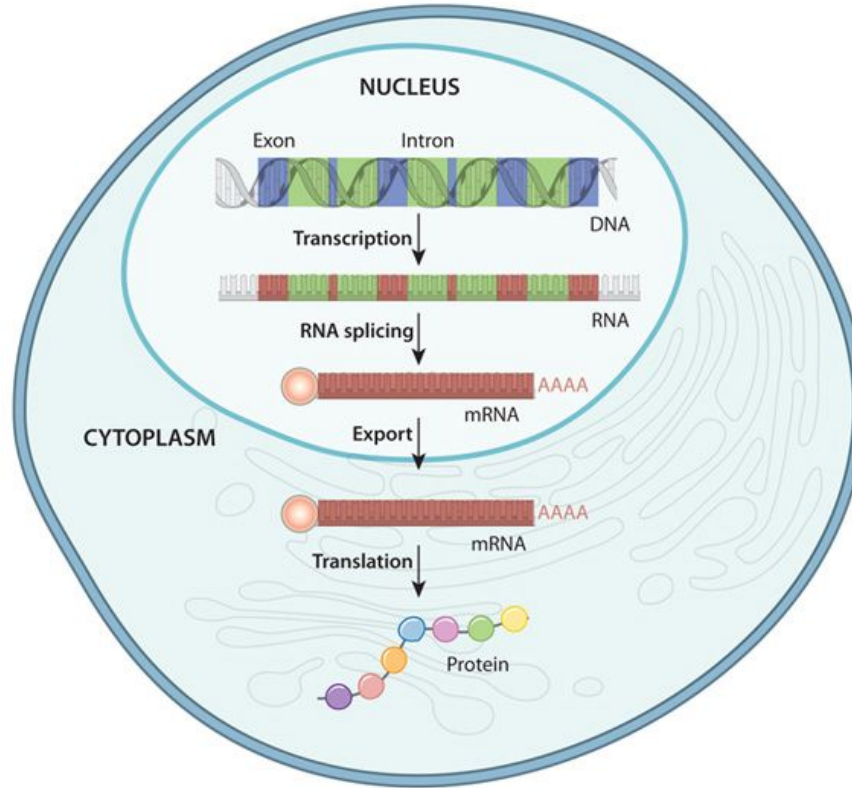
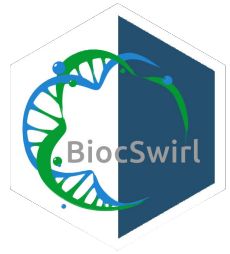
Almas Khan

@VancityAlmas

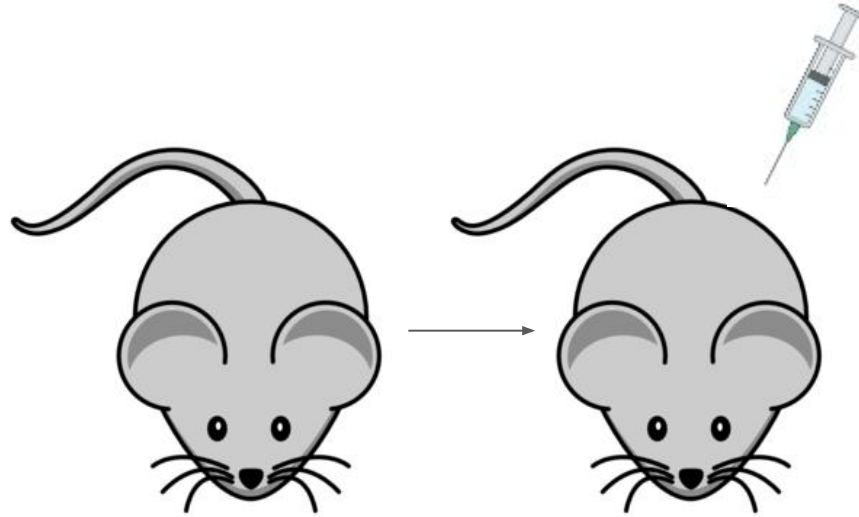
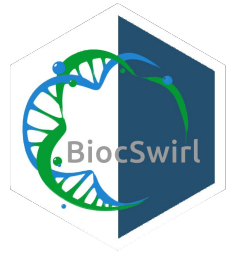
Julia Philipp

@its_JPhilipp

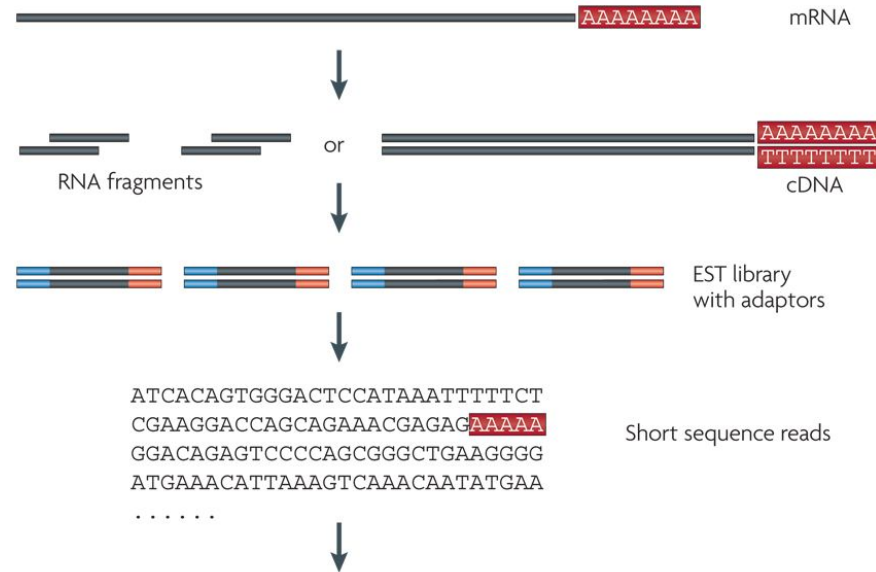
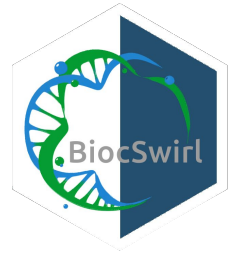
Part II - RNAseq // Gene expression



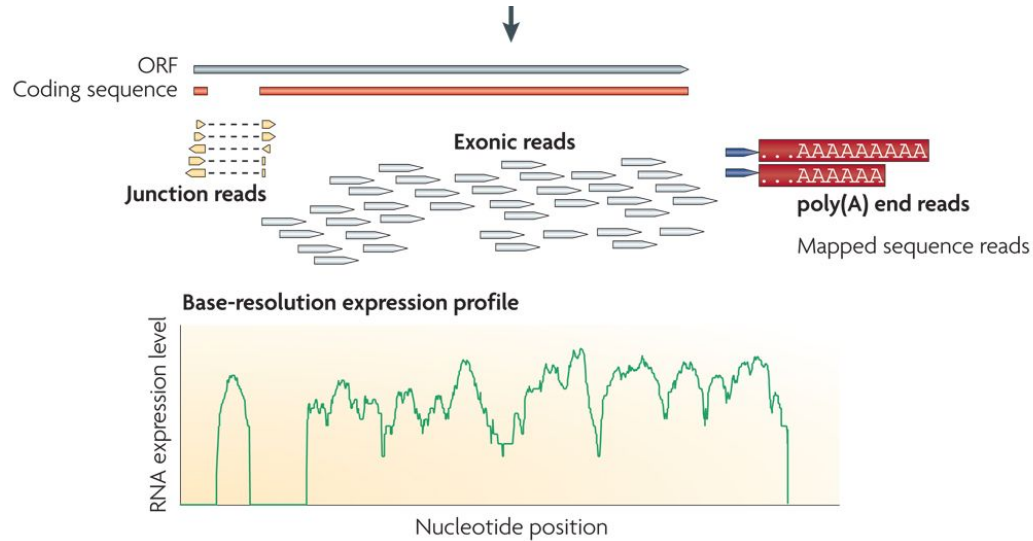
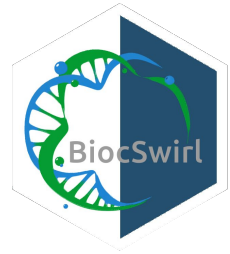
Differential Gene Expression



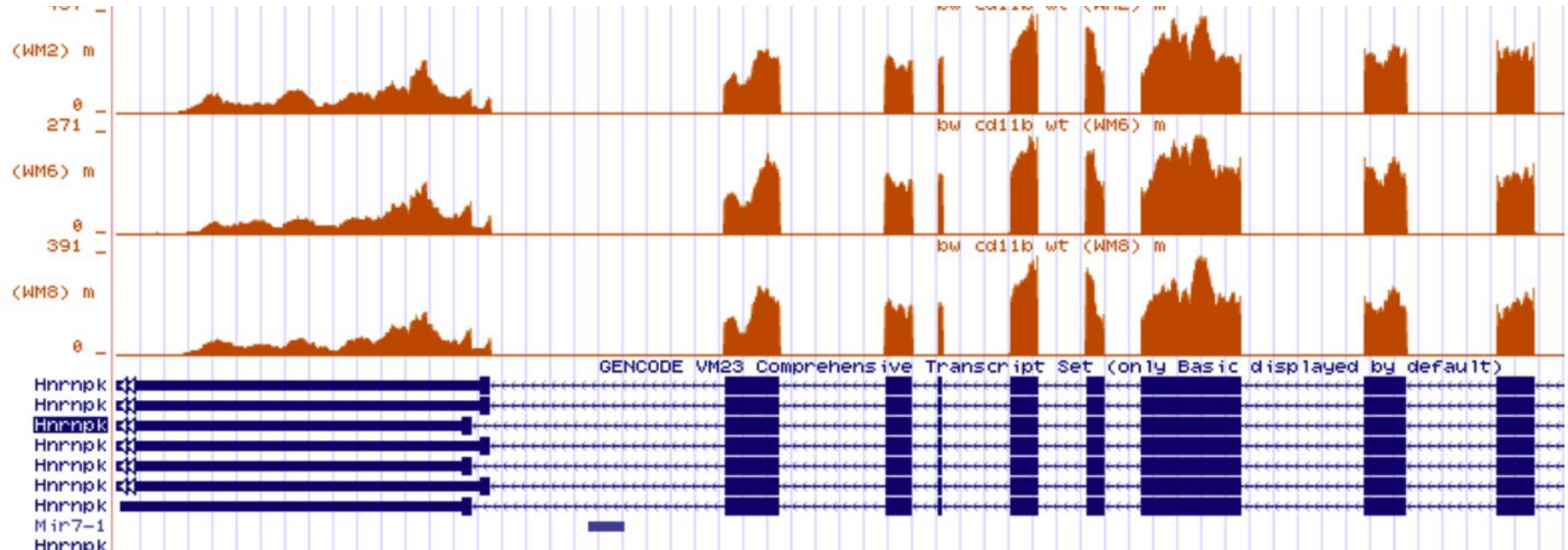
Introduction to RNAseq



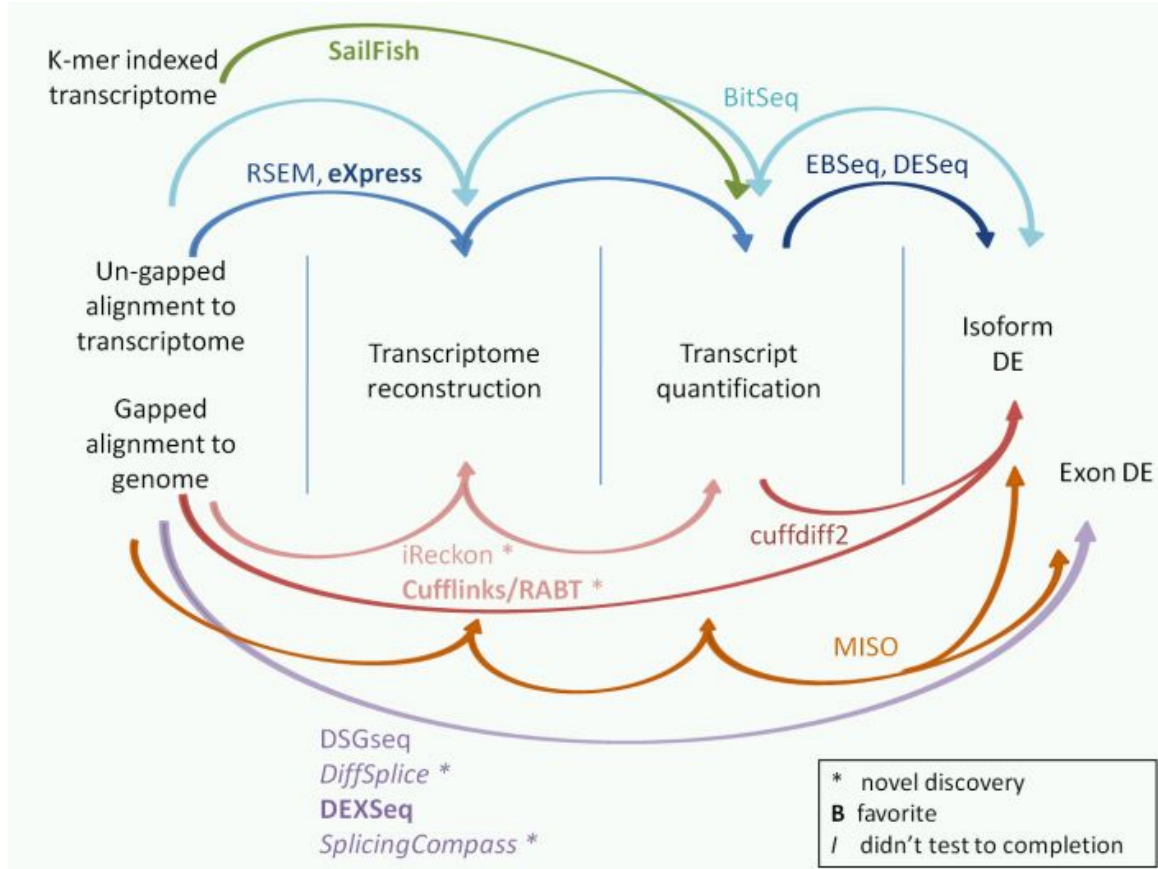
Introduction to RNAseq

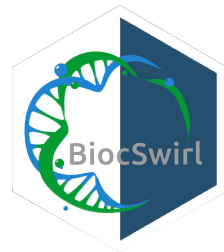


Example RNA-seq data in the genome browser



Analysis workflows and tools





RNAseq analysis workflow for today

RNA-seq workflow: gene-level exploratory analysis and differential expression

Michael I. Love^{1,2}, Simon Anders³, Vladislav Kim⁴ and Wolfgang Huber⁴

¹Department of Biostatistics, UNC-Chapel Hill, Chapel Hill, NC, US

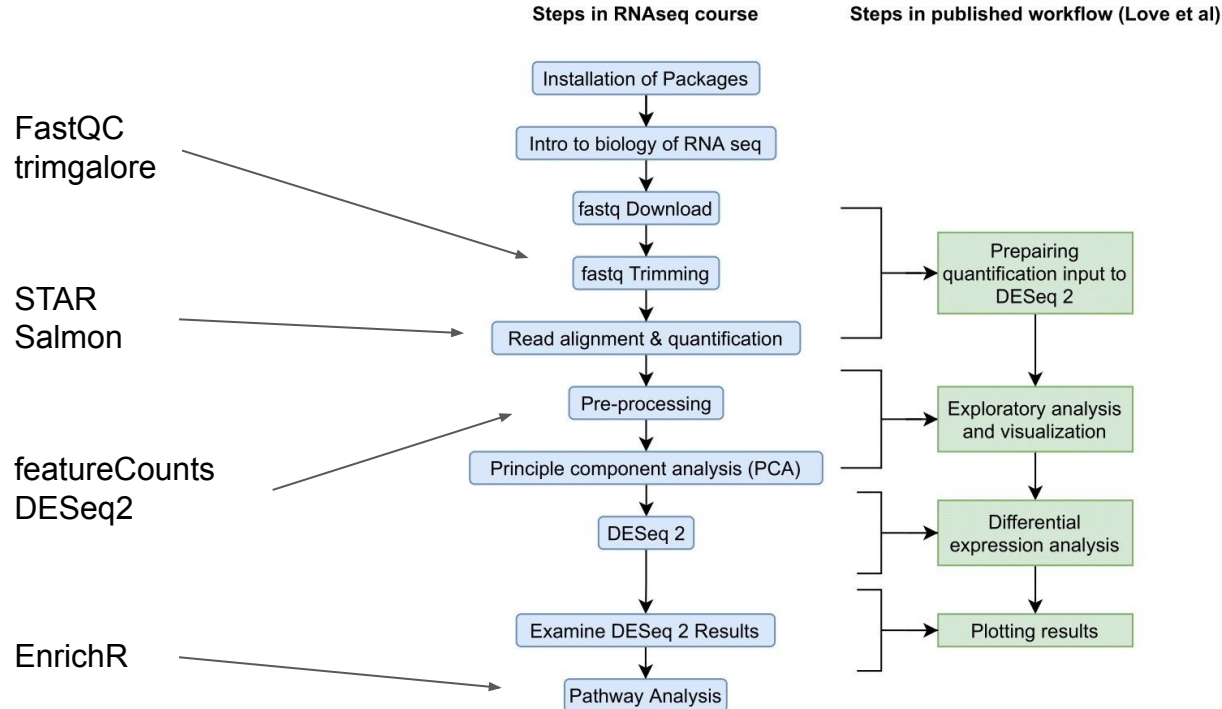
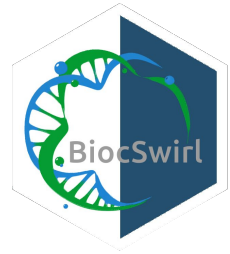
²Department of Genetics, UNC-Chapel Hill, Chapel Hill, NC, US

³Zentrum für Molekulare Biologie der Universität Heidelberg, Heidelberg, Germany

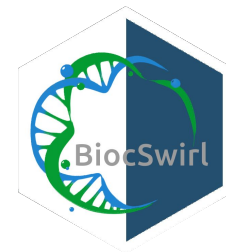
⁴European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

16 October, 2019

RNAseq workflow for today



What is Bioconductor?

[Home](#)[Install](#)[Help](#)[Developers](#)[About](#)Search:

About *Bioconductor*

Bioconductor provides tools for the analysis and comprehension of high-throughput genomic data.

Bioconductor uses the R statistical programming language, and is open source and open development. It has two releases each year, and an active user community. *Bioconductor* is also available as an [AMI](#) (Amazon Machine Image) and [Docker](#) images.

News

- *Bioconductor* [3.12](#) is available.
- [BioC Europe](#) virtual conference registration and abstract submission open December 14-18, 2020.
- See our [google calendar](#) for events, conferences, meetings, forums, etc. Add your event with email to events at [bioconductor.org](#).
- Core team **job opportunities** available.

Install »

- Discover [1974 software packages](#) available in *Bioconductor* release 3.12.

Get started with *Bioconductor*

- [Install *Bioconductor*](#)
- [Get support](#)
- [Latest newsletter](#)
- [Follow us on twitter](#)
- [Install R](#)

Learn »

Master *Bioconductor* tools

- [Courses](#)
- [Support site](#)
- [Package vignettes](#)
- [Literature citations](#)
- [Common work flows](#)
- [FAQ](#)
- [Community resources](#)
- [Videos](#)

Use »

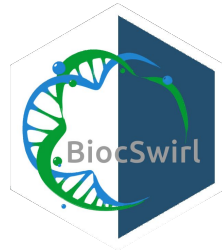
Create bioinformatic solutions with *Bioconductor*

- [Software](#), [Annotation](#), and [Experiment](#) packages
- [Docker](#) and [Amazon](#) machine images

Develop »

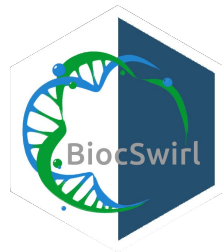
Contribute to *Bioconductor*

- [Developer resources](#)
- [Use Bioc 'devel'](#)
- ['Devel' packages](#)
- [Package guidelines](#)



What is BiocSwirl?

- Interactive courses teaching Bioconductor analysis workflows
- Can be taken within your R Studio
- Easy to install
- Example input and output
- Feedback on your syntax / coding input
- Self-paced, progress can be saved



BiocSwirl & course installation

```
install.packages("devtools")

library(devtools)

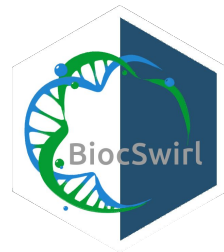
devtools::install_github("biocswirl-dev-team/BiocSwirl", build_vignettes
= TRUE)

library(BiocSwirl)
library(swirl)

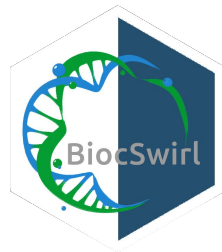
list_courses()

load_course('RNAseq')

start_course()
```



Break time?



Part III - RNAseq analysis exercise

```
> start_course()
```

```
| Welcome to swirl! Please sign in. If you've been here before, use the same name as you did  
| then. If you are new, call yourself something unique.
```

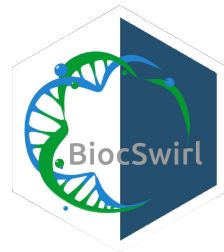
```
What shall I call you?
```

```
bye() #exits the course environment
```

```
play() #suspends feedback, allows you to 'play' with code
```

```
nxt() #continue feedback
```

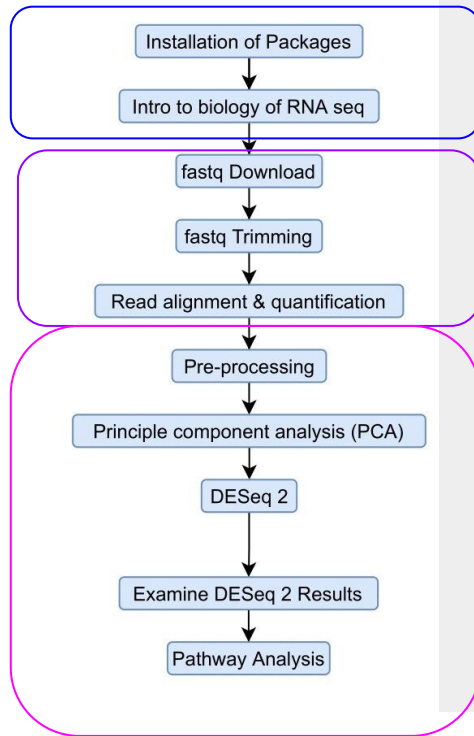
```
skip() #skip question, not recommended
```



RNAseq workflow for today

We'll do these together!

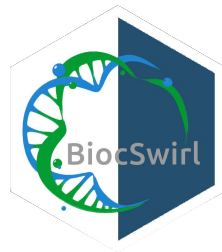
Steps in RNAseq course



I'll live demo some non-R stuff!

Self paced :)

At any time: Ask questions & check etherpad!



Part III - RNAseq analysis exercise

```
> start_course()
```

| Welcome to swirl! Please sign in. If you've been here before, use the same name as you did
| then. If you are new, call yourself something unique.

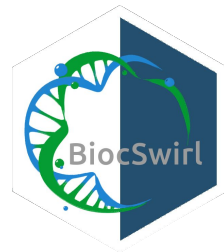
```
What shall I call you?
```

```
swirl::bye() #exits the course environment
```

```
swirl::play() #suspends feedback, allows you to 'play' with  
code
```

```
swirl::nxt() #continue feedback
```

```
swirl::skip() #skip question, not recommended
```

Other analysis goals for RNAseq data

- differential gene expression
- variants detection and allele-specific expression
- small RNA profiling
- characterization of alternative splicing patterns
- Transcript identification / annotation
- RNA editing
- Single cell RNAseq instead of bulk RNAseq

Differences Between Bulk and Single Cell RNA-Seq

Bulk

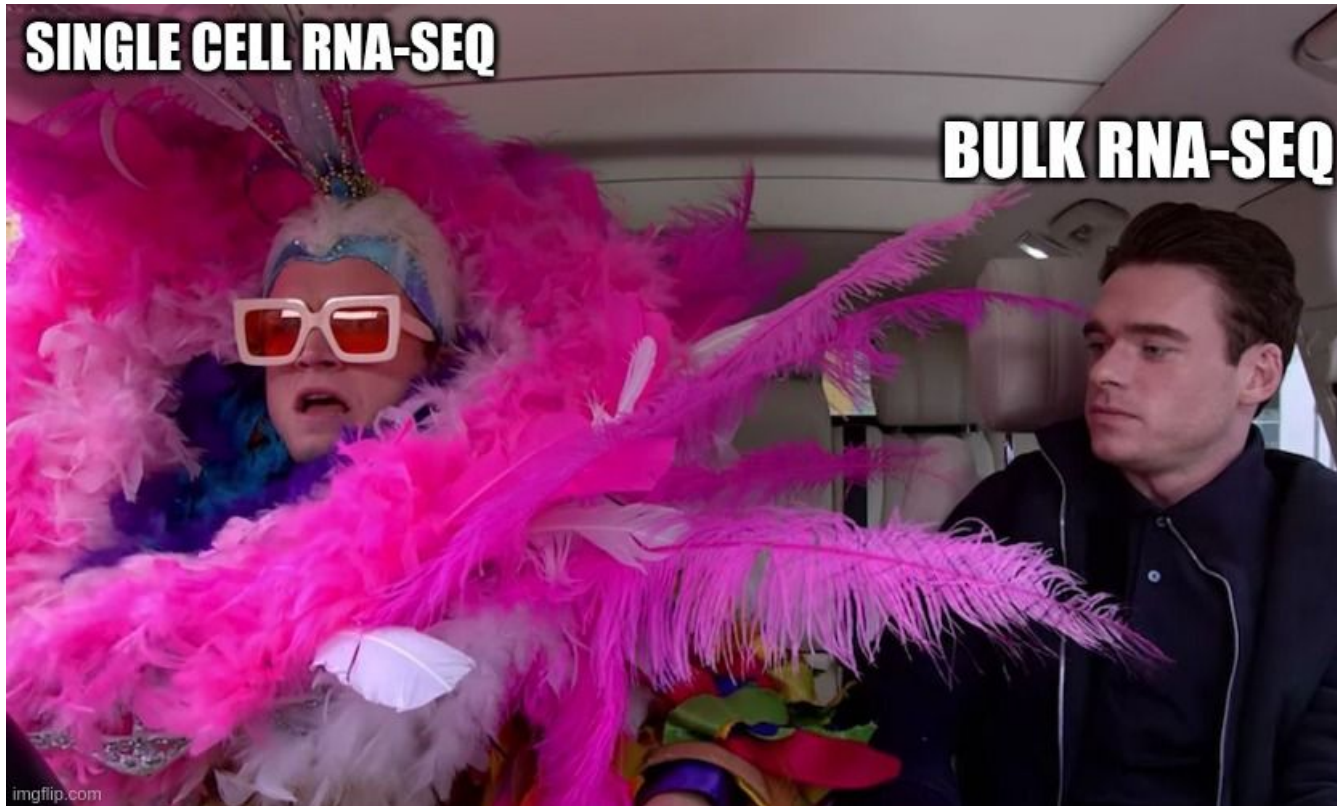
- Higher sequencing depth allows exploration of more genes
- Unable to distinguish gene expression changes due to tissue composition changes
 - Tissues sequenced should be as homogenous as possible
 - For heterogeneous tissues, may be beneficial to FACS sort and sequence the different cell types separately
- Lower cost and ease of preparation allow for more samples to be analyzed
 - More conditions can be studied
 - More replicates can be used to ensure significance of results

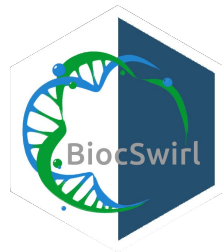
Single Cell

- Cell specific tagging allows for exploration of heterogeneity in your tissue of interest
 - Allows for discovery of rare cell types
- Due to high cost, removing batch effects can be a substantial problem
 - Software solutions to this problem are rapidly being developed

SINGLE CELL RNA-SEQ

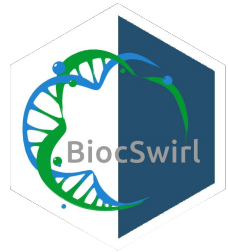
BULK RNA-SEQ





Open source science

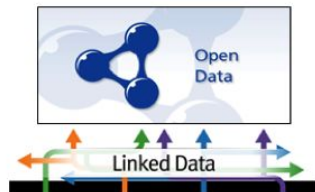
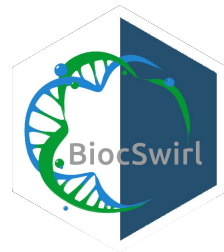
- Publicly available code
 - Can be reviewed, reused, re'mixed'
 - Can be contributed to by community
 - Forums
 - Good documentation
 - Shareable
- Free distribution
- Less emphasis on individual, more on community



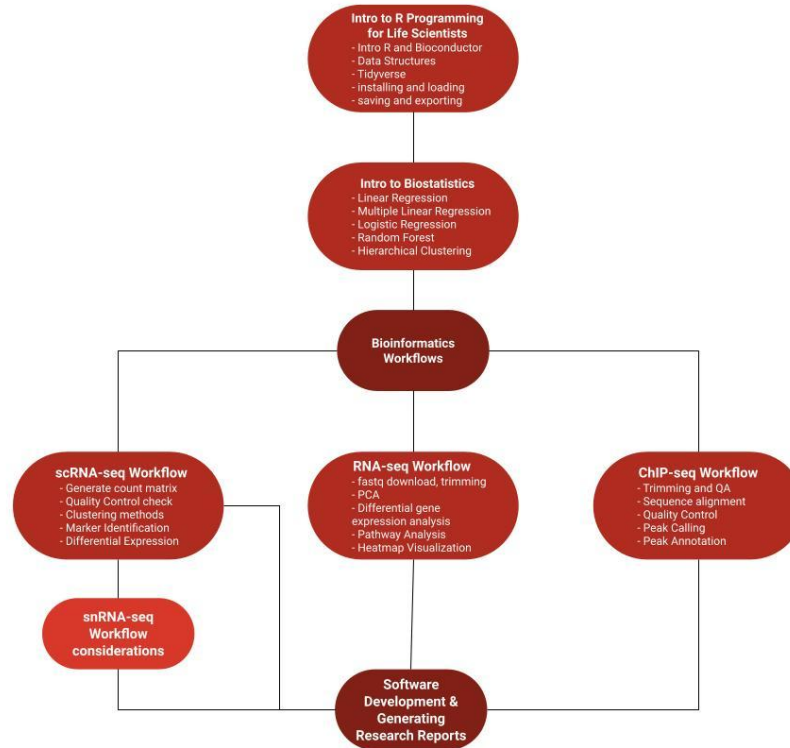
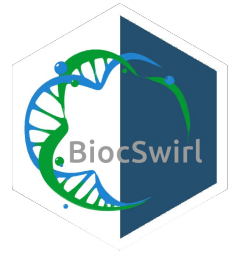
The need for and benefit of open science

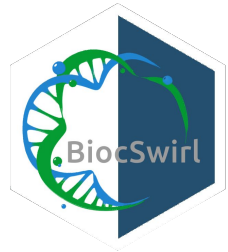
- Code written by “not software devs” / independent researchers
 - High redundancy
 - Low reproducibility / buggy code
 - Not maintained
 - Not well documented
- With open science:
 - Good documentation
 - Maintenance
 - Reproducibility
 - User-friendly code

Open Science Resources



Other BiocSwirl courses you could try!





Other BiocSwirl courses you could try!

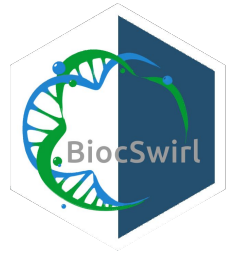
```
> library(BiocSwirl)
```

```
> list_courses()
```

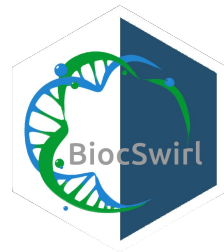
	course_name	description
1	Intro_to_Data_Science	Intro to R and data analysis
2	Intro_to_Stats	Intro to Statistics
3	RNAseq	bulk RNAseq analysis
4	scRNAseq	single-cell RNAseq analysis
5	ChIPseq	ChIPseq analysis

```
> |
```


Post-workshop assessment



Stay in touch!



Follow us on twitter [@bioctools](https://twitter.com/bioctools) for updates on new course releases & workshops

Check out our github github.com/biocswirl-dev-team/BiocSwirl for more courses, help, etc...

Our website: <https://bioinformaticstutorials.github.io/> is still under construction :]

Thank you!

