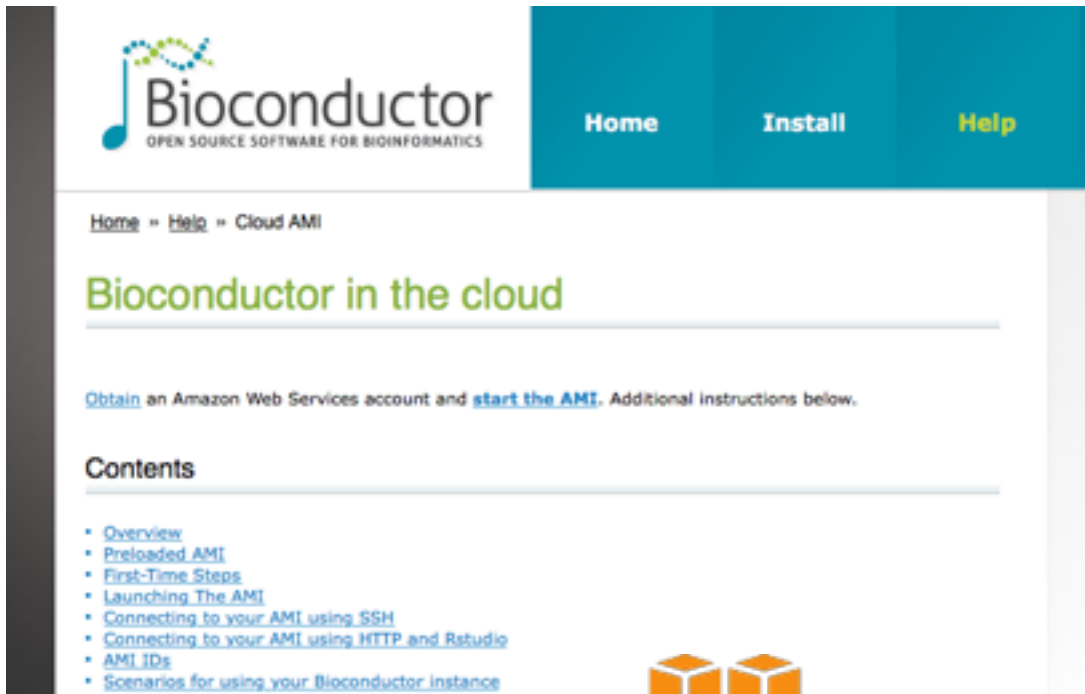



Using Cloud Computing for Workshops

- **Problem:** Setting up the computing environment on each participant's computer takes time away from workshop content; versioning can cause problems
- **Solution:** Provide each participant with a cloud server with a preconfigured environment to access and use during the workshop



Bioconductor supports preconfigured AMIs



 **Bioconductor**
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

[Home](#) [Install](#) [Help](#)

[Home](#) » [Help](#) » [Cloud AMI](#)

Bioconductor in the cloud

[Obtain](#) an Amazon Web Services account and [start the AMI](#). Additional instructions below.

Contents

- [Overview](#)
- [Preloaded AMI](#)
- [First-Time Steps](#)
- [Launching The AMI](#)
- [Connecting to your AMI using SSH](#)
- [Connecting to your AMI using HTTP and Rstudio](#)
- [AMI IDs](#)
- [Scenarios for using your Bioconductor instance](#)

Pros:

- Pre-loaded with latest R version, top 80 BioC software pkgs plus many annotation packages
- Supported & updated

Cons:

- Limited customizability of AMI properties
- Pricing structure can be confusing (e.g. bidding on Spot Instances)




Digital Ocean

 [Droplets](#) [Images](#) [Networking](#) [API](#) [Support](#)

Create Droplets

Choose an image 

[Distributions](#) [One-click apps](#) [Snapshots](#)

 Ubuntu 16.04.1 x64 	 FreeBSD Select version 	 Fedora Select version 	 Debian Select version 
--	--	---	---



Pros:

- Fully customizable droplet properties
- Simple pricing structure

Cons:

- All setup of R/Rstudio Server, packages, etc done manually

Steps to Create Rstudio Server Instances

1. Launch empty droplet (instance)
 - Choose OS version, RAM, # cores, disk size, location, etc
2. Configure environment
 - Install R, Rstudio Server, packages, etc
3. Take a snapshot image
4. Launch many droplets based on snapshot





1. Launch empty droplet

 [Droplets](#) [Images](#) [Networking](#) [API](#) [Support](#)

Create Droplets

Choose an image ?

[Distributions](#) [One-click apps](#) [Snapshots](#)

 Ubuntu	 FreeBSD	 Fedora	 Debian
16.04.1 x64 ▼	Select version ▼	Select version ▼	Select version ▼

1. Launch empty droplet

Choose a size

Standard High memory

<p>\$5/mo \$0.007/hour</p> <p>512 MB / 1 CPU 20 GB SSD disk 1000 GB transfer</p>	<p>\$10/mo \$0.015/hour</p> <p>1 GB / 1 CPU 30 GB SSD disk 2 TB transfer</p>	<p>\$20/mo \$0.030/hour</p> <p>2 GB / 2 CPUs 40 GB SSD disk 3 TB transfer</p>	<p>\$40/mo \$0.060/hour</p> <p>4 GB / 2 CPUs 60 GB SSD disk 4 TB transfer</p>	<p>\$80/mo \$0.119/hour</p> <p>8 GB / 4 CPUs 80 GB SSD disk 5 TB transfer</p>	<p>\$160/mo \$0.238/hour</p> <p>16 GB / 8 CPUs 160 GB SSD disk 6 TB transfer</p>
<p>\$320/mo \$0.476/hour</p> <p>32 GB / 12 CPUs 320 GB SSD disk 7 TB transfer</p>	<p>\$480/mo \$0.714/hour</p> <p>48 GB / 16 CPUs 480 GB SSD disk 8 TB transfer</p>	<p>\$640/mo \$0.952/hour</p> <p>64 GB / 20 CPUs 640 GB SSD disk 9 TB transfer</p>			

1. Launch empty droplet

Choose a size

Standard High memory

<p>\$5/mo \$0.007/hour</p> <p>512 MB / 1 CPU 20 GB SSD disk 1000 GB transfer</p>	<p>\$10/mo \$0.015/hour</p> <p>1 GB / 1 CPU 30 GB SSD disk 2 TB transfer</p>	<p>\$20/mo \$0.030/hour</p> <p>2 GB / 2 CPUs 40 GB SSD disk 3 TB transfer</p>	<p>\$40/mo \$0.060/hour</p> <p>4 GB / 2 CPUs 60 GB SSD disk 4 TB transfer</p>	<p>\$80/mo \$0.119/hour</p> <p>8 GB / 4 CPUs 80 GB SSD disk 5 TB transfer</p>	<p>\$160/mo \$0.238/hour</p> <p>16 GB / 8 CPUs 160 GB SSD disk 6 TB transfer</p>
<p>\$320/mo \$0.476/hour</p> <p>32 GB / 12 CPUs 320 GB SSD disk 7 TB transfer</p>	<p>\$480/mo \$0.714/hour</p> <p>48 GB / 16 CPUs 480 GB SSD disk 8 TB transfer</p>	<p>\$640/mo \$0.952/hour</p> <p>64 GB / 20 CPUs 640 GB SSD disk 9 TB transfer</p>			

...







1. Launch empty droplet

Choose a size

Standard High memory

\$5/mo \$0.007/hour	\$10/mo \$0.015/hour	\$20/mo \$0.030/hour	\$40/mo \$0.060/hour	\$80/mo \$0.119/hour	\$160/mo \$0.238/hour
512 MB / 1 CPU 20 GB SSD disk 1000 GB transfer	1 GB / 1 CPU 30 GB SSD disk 2 TB transfer	2 GB / 2 CPUs 40 GB SSD disk 3 TB transfer	4 GB / 2 CPUs 60 GB SSD disk 4 TB transfer	8 GB / 4 CPUs 80 GB SSD disk 5 TB transfer	16 GB / 8 CPUs 160 GB SSD disk 6 TB transfer
\$320/mo \$0.476/hour	\$480/mo \$0.714/hour	\$640/mo \$0.952/hour			
32 GB / 12 CPUs 320 GB SSD disk 7 TB transfer	48 GB / 16 CPUs 480 GB SSD disk 8 TB transfer	64 GB / 20 CPUs 640 GB SSD disk 9 TB transfer			

Choose a datacenter region

 New York 1 2 3	 San Francisco 1 2	 Amsterdam 2 3	 Singapore 1	 London 1	 Frankfurt 1
 Toronto 1	 Bangalore 1				

...

1. Launch empty droplet

Choose a size

Standard High memory

<p>\$5/mo \$0.007/hour</p> <p>512 MB / 1 CPU 20 GB SSD disk 1000 GB transfer</p>	<p>\$10/mo \$0.015/hour</p> <p>1 GB / 1 CPU 30 GB SSD disk 2 TB transfer</p>	<p>\$20/mo \$0.030/hour</p> <p>2 GB / 2 CPUs 40 GB SSD disk 3 TB transfer</p>	<p>\$40/mo \$0.060/hour</p> <p>4 GB / 2 CPUs 60 GB SSD disk 4 TB transfer</p>	<p>\$80/mo \$0.119/hour</p> <p>8 GB / 4 CPUs 80 GB SSD disk 5 TB transfer</p>	<p>\$160/mo \$0.238/hour</p> <p>16 GB / 8 CPUs 160 GB SSD disk 6 TB transfer</p>
<p>\$320/mo \$0.476/hour</p> <p>32 GB / 12 CPUs 320 GB SSD disk 7 TB transfer</p>	<p>\$480/mo \$0.714/hour</p> <p>48 GB / 16 CPUs 480 GB SSD disk 8 TB transfer</p>	<p>\$640/mo \$0.952/hour</p> <p>64 GB / 20 CPUs 640 GB SSD disk 9 TB transfer</p>			

Finalize and create

How many Droplets?

Deploy multiple Droplets with the same [configuration](#).

— 1 Droplet +

Choose a hostname

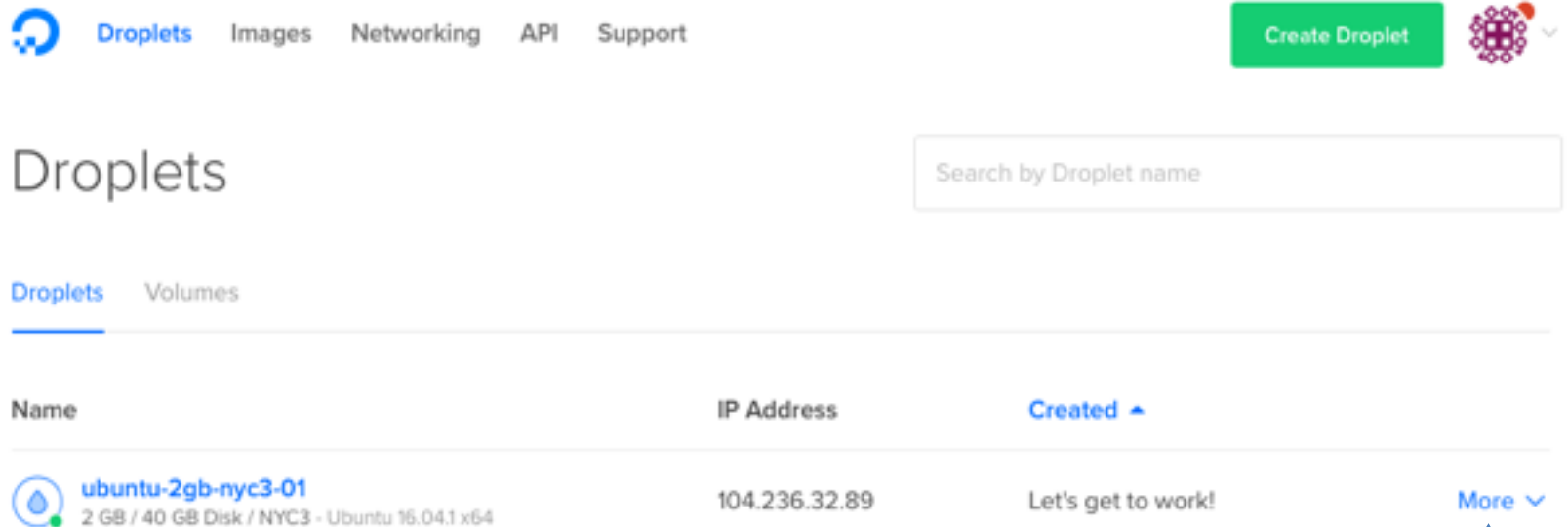
Give your Droplets an identifying name you will remember them by. Your Droplet name can only contain alphanumeric characters, dashes, and periods.

ubuntu-2gb-nyc3-01


...

Create

2. Configure Environment



The screenshot shows the DigitalOcean management console. At the top, there are navigation links for 'Droplets', 'Images', 'Networking', 'API', and 'Support'. A green 'Create Droplet' button is visible on the right, along with a user profile icon. Below the navigation is a search bar labeled 'Search by Droplet name'. Underneath, there are tabs for 'Droplets' and 'Volumes'. A table lists the droplets with columns for 'Name', 'IP Address', and 'Created'. One droplet is listed: 'ubuntu-2gb-nyc3-01' with IP '104.236.32.89' and status 'Let's get to work!'. A 'More' dropdown menu is shown next to the droplet name, with a blue arrow pointing to it from the text below.

Name	IP Address	Created
 ubuntu-2gb-nyc3-01 2 GB / 40 GB Disk / NYC3 - Ubuntu 16.04.1 x64	104.236.32.89	Let's get to work!

Options here to delete,
resize, etc

2. Configure Environment


Droplets Images Networking API Support

Create Droplet

Droplets

Search by Droplet name


Droplets Volumes


Name	IP Address	Created
 ubuntu-2gb-nyc3-01 2 GB / 40 GB Disk / NYC3 - Ubuntu 16.04.1 x64	104.236.32.89	Let's get to work! More

SSH into this IP address using username and password emailed to you

Options here to delete, resize, etc

2. Configure Environment

GitHub Gist Search... All gists GitHub New gist 

 [kdkorthauer](#) / [RstudioServerSetup.sh](#) Created 2 hours ago Edit Delete ★ Star 0 ⓘ

↔ Code ↔ Revisions 1 Embed <script src="https://gist. ..." 📄 📄 Download ZIP

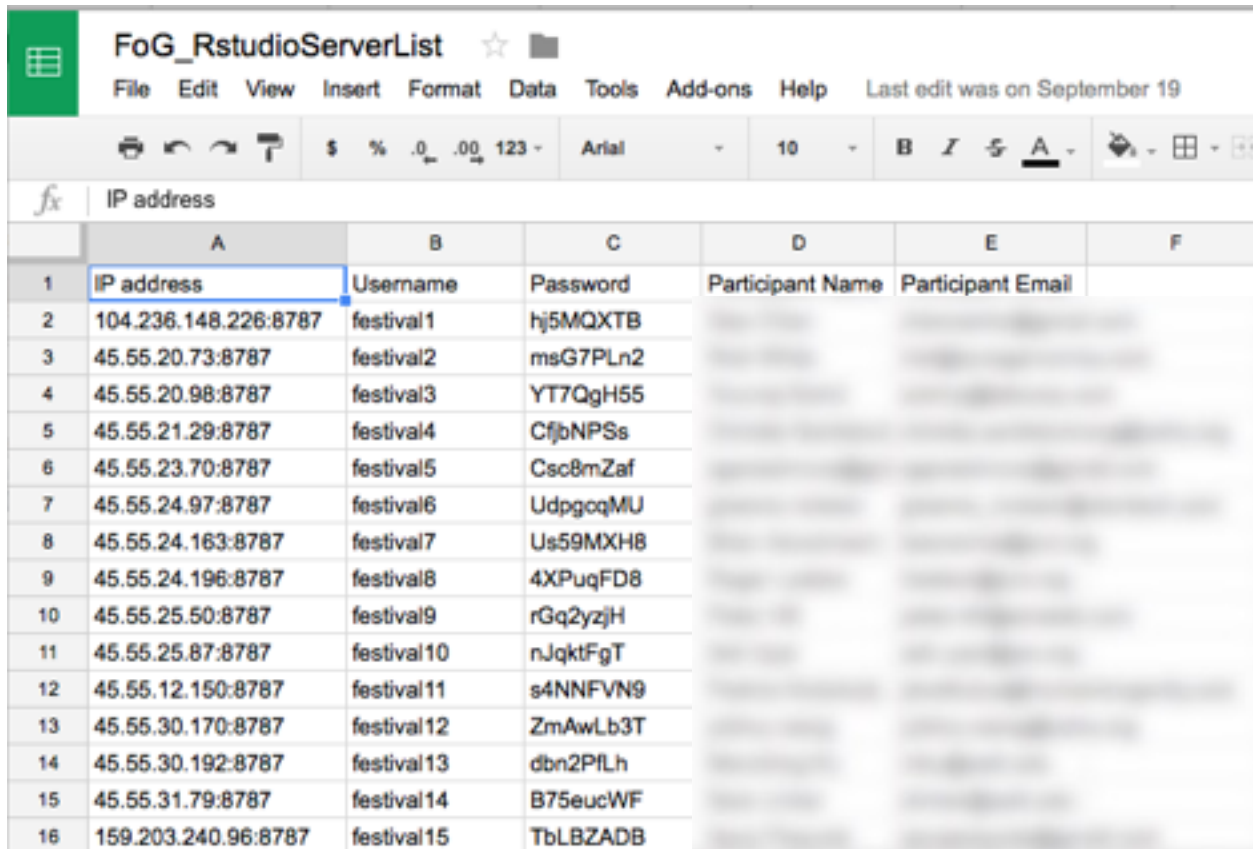
Bash script to set up R, install a few R packages, and get Rstudio Server running on ubuntu.

```
RstudioServerSetup.sh Raw
1  sudo sh -c 'echo "deb http://cran.rstudio.com/bin/linux/ubuntu trusty/" >> /etc/apt/sources.list'
2  gpg --keyserver keyserver.ubuntu.com --recv-key E884DAB9
3  gpg -a --export E884DAB9 | sudo apt-key add -
4  sudo apt-get update
5  sudo apt-get -y install r-base libapparmor1 libcurl4-gnutls-dev libxml2-dev libssl-dev gdebi-core
6  sudo apt-get install libcairo2-dev
7  sudo apt-get install libxt-dev
8  sudo apt-get install git-core
9
10 sudo /bin/dd if=/dev/zero of=/var/swap.1 bs=1M count=1024
11 sudo /sbin/mkswap /var/swap.1
12 sudo /sbin/swapon /var/swap.1
13 sudo sh -c 'echo "/var/swap.1 swap swap defaults 0 0 " >> /etc/fstab'
14
15 sudo su - -c "R -e \"install.packages('devtools', repos='http://cran.rstudio.com/')\"
16 sudo su - -c "R -e \"install.packages('Rcpp', repos='http://cran.rstudio.com/')\"
17 sudo su - -c "R -e \"install.packages('RcppEigen', repos='http://cran.rstudio.com/')\"
18 sudo su - -c "R -e \"install.packages('ggplot2', repos='http://cran.rstudio.com/')\"
19 sudo su - -c "R -e \"install.packages('Cairo', repos='http://cran.rstudio.com/')\""
```

GitHub Gist: [/kdkorthauer/RstudioServerSetup.sh](#)

2. Configure Environment

- Create a non-root account for yourself with sudo privileges, e.g. 'adduser keegan sudo'
- Set up an account for each user with a password and homedir using the 'adduser' command



The screenshot shows a spreadsheet application window titled "FoG_RstudioServerList". The window has a menu bar with "File", "Edit", "View", "Insert", "Format", "Data", "Tools", "Add-ons", and "Help". Below the menu bar is a toolbar with various icons for printing, undo, redo, and text formatting. The spreadsheet itself has a header row with columns labeled "A" through "F". The data rows contain IP addresses, usernames (festival1 through festival15), and passwords. The "Participant Name" and "Participant Email" columns are present but contain blurred text.

	A	B	C	D	E	F
1	IP address	Username	Password	Participant Name	Participant Email	
2	104.236.148.226:8787	festival1	hj5MQXTB			
3	45.55.20.73:8787	festival2	msG7PLn2			
4	45.55.20.98:8787	festival3	YT7QgH55			
5	45.55.21.29:8787	festival4	CfjbNPSs			
6	45.55.23.70:8787	festival5	Csc8mZaf			
7	45.55.24.97:8787	festival6	UdpgcqMU			
8	45.55.24.163:8787	festival7	Us59MXH8			
9	45.55.24.196:8787	festival8	4XPuqFD8			
10	45.55.25.50:8787	festival9	rGq2yzjH			
11	45.55.25.87:8787	festival10	nJqktFgT			
12	45.55.12.150:8787	festival11	s4NNFVN9			
13	45.55.30.170:8787	festival12	ZmAwLb3T			
14	45.55.30.192:8787	festival13	dbn2PILh			
15	45.55.31.79:8787	festival14	B75eucWF			
16	159.203.240.96:8787	festival15	TbLBZADB			


3. Take a snapshot image

Images

Take a Snapshot

Power-down Droplets before taking a snapshot to ensure data consistency. Snapshots cost is based on space used and charged at a rate of \$0.05/GB/mo.

 **rstudio-fog-09192016-2gb-sfo1-01** 
2 GB / 40 GB / SFO1

Enter image name
rstudio-fog-09192016-2gb-sfo1-01-1475860738C 

[Take Snapshot](#)





4. Launch many droplets based on snapshot

 [Droplets](#) [Images](#) [Networking](#) [API](#) [Support](#)

Create Droplets

Choose an image 

[Distributions](#) [One-click apps](#) [Snapshots](#)

 Ubuntu rstudio-FoG-09192016	 Ubuntu rstudio-FoG-09192016_n..
 Ubuntu rstudio-1-1472566373570	 Ubuntu rstudio-1-1472159987879

4. Launch many droplets based on snapshot

 [Droplets](#) [Images](#) [Networking](#) [API](#) [Support](#)

Create Droplets

Choose an image 

Same process as before, but select a Snapshot instead of a Distribution (empty)

[Distributions](#) [One-click apps](#) [Snapshots](#)



 Ubuntu rstudio-FoG-09192016

 Ubuntu rstudio-FoG-09192016_n...

 Ubuntu rstudio-1-1472566373570

 Ubuntu rstudio-1-1472159987879

4. Launch many droplets based on snapshot

Create Droplets

Choose an image 

Same process as before, but select a Snapshot instead of a Distribution (empty)

[Distributions](#) [One-click apps](#) [Snapshots](#)




 Ubuntu rstudio-FoG-09192016

 Ubuntu rstudi

Finalize and create

 Ubuntu rstudio-1-1472566373570

 Ubuntu rstudi

How many Droplets?

Deploy multiple Droplets with the same [configuration](#).



5 Droplets



Choose a hostname

Give your Droplets an identifying name you will remember them by. Your Droplet name can only contain alphanumeric characters, dashes, and periods.

rstudio-fog-09192016-2gb-sfo1-01

rstudio-fog-09192016-2gb-sfo1-02

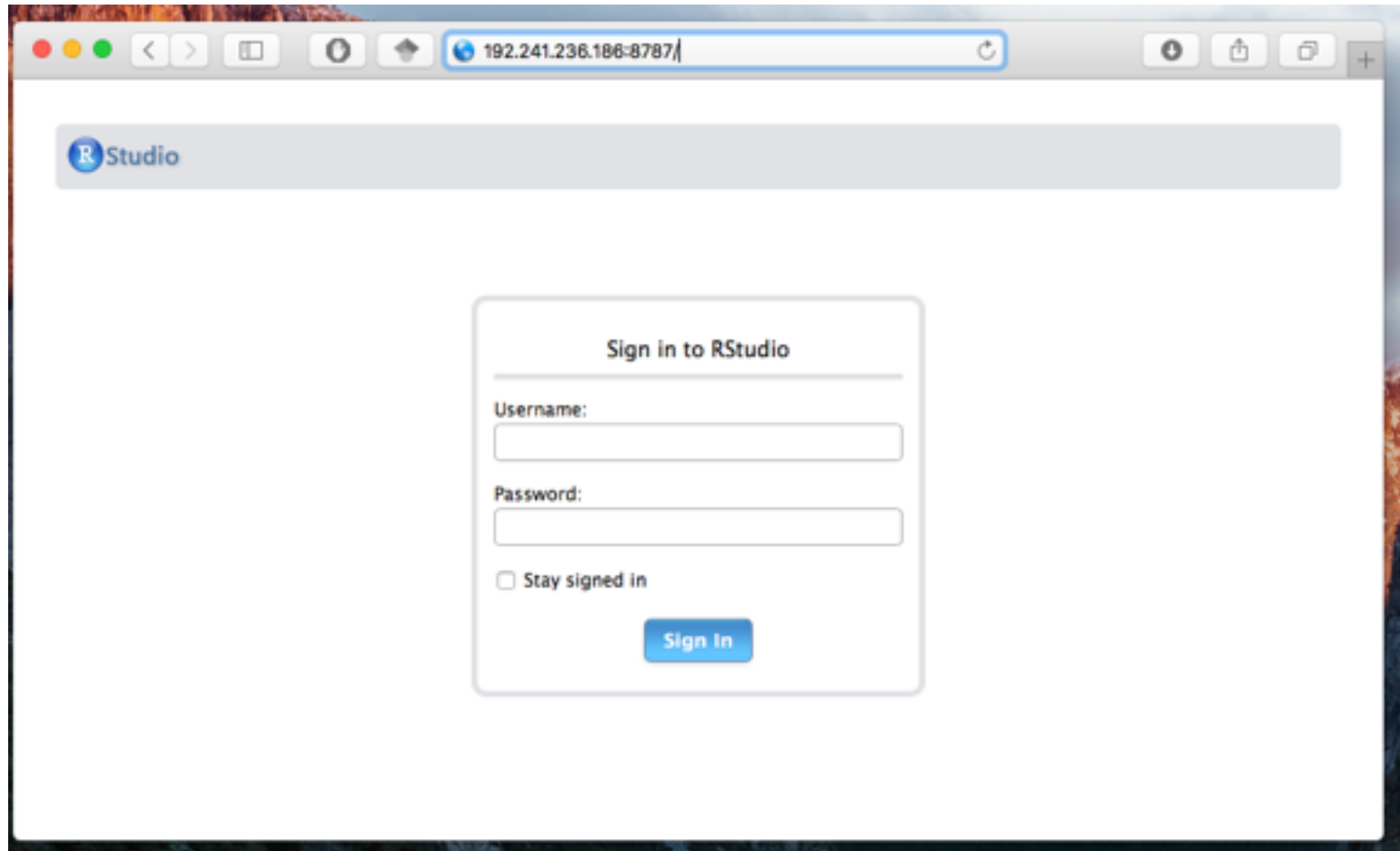
rstudio-fog-09192016-2gb-sfo1-03

rstudio-fog-09192016-2gb-sfo1-04

rstudio-fog-09192016-2gb-sfo1-05

Create

End Result



End Result

The screenshot displays the RStudio environment. The main editor window shows a script named 'SingleCellAnalyses.Rmd' with the following R code:

```
25 - ```{r Check for data files, eval=TRUE, echo=TRUE}
26 setwd("~/FestivalWorkshopSC/BrainAtlas")
27 file.exists("cell_metadata.csv")
28 file.exists("genes_counts.csv")
29 file.exists("genes_rpkm.csv")
30 file.exists("ercc_counts.csv")
31 file.exists("README.txt")
32 - ```
33
34 If any of the preceding lines return 'FALSE', double check
```

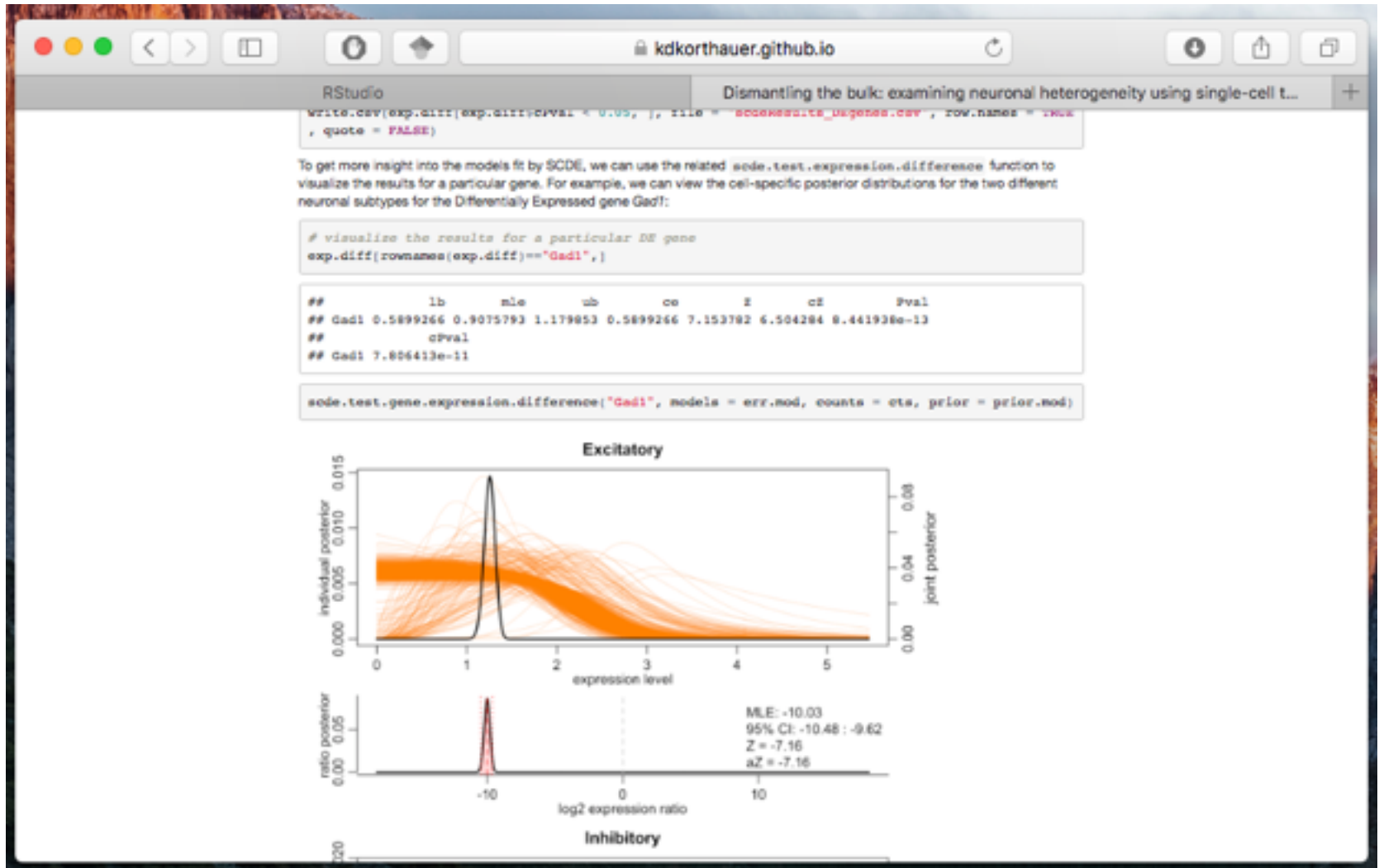
The console window at the bottom left shows the R prompt and the following text:

```
1:1 [1] Dismantling the bulk: examining neuronal heterogeneity using single-cell te
> |
```

The file explorer on the right shows the directory structure: /home > FestivalWorkshopSC > BrainAtlas. The files listed are:

Name	Size	Modified
..		
cell_classification.csv	197.8 KB	Aug 26, 2016, 2:13 PM
cell_data.csv	2.1 MB	Aug 26, 2016, 2:13 PM
cell_metadata.csv	182.2 KB	Aug 26, 2016, 2:13 PM
cluster_metadata.csv	7.1 KB	Aug 26, 2016, 2:13 PM
ercc_counts.csv	741.4 KB	Aug 26, 2016, 2:13 PM

End Result



<https://kdkorthauer.github.io/FestivalWorkshopVignettes/>