

## Instructions for setting up RStudioServer on Biowulf

### Running R Studio-Server on Biowulf

This follows the official Biowulf documentation <https://hpc.nih.gov/apps/rstudio-server.html> but adds a bit more detail for new folks

RStudio Server is used interactively via the browser and requires an SSH tunnel connection to Biowulf server.

Open a new terminal window (we'll call this window #1), login to Biowulf as normal:

```
ssh yourusername@biowulf.nih.gov
```

Start an interactive session in Biowulf:

```
sinteractive --mem=28g --gres=lscratch:10 --tunnel
```

(this gives you 10GB of temporary data in your scratch directory and 28GB of memory to do your work, and also gives you 2 CPUs)

It will output something like this and give you instructions for opening a tunnel to Biowulf:

```
salloc.exe: Pending job allocation 15323416
salloc.exe: job 15323416 queued and waiting for resources
salloc.exe: job 15323416 has been allocated resources
salloc.exe: Granted job allocation 15323416salloc.exe: Waiting for
resource configuration
salloc.exe: Nodes cn1640 are ready for job
```

```
Created 1 generic SSH tunnel(s) from this compute node to
biowulf for your use at port numbers defined
in the $PORTn ($PORT1, ...) environment variables.
```

```
Please create a SSH tunnel from your workstation to these ports on
biowulf.
```

```
On Linux/MacOS, open a terminal and run:
```

```
ssh -L 39689:localhost:39689 user@biowulf.nih.gov
```

```
For Windows instructions, see https://hpc.nih.gov/docs/tunneling
```

Following these instructions, open a new terminal window (#2) and copy and paste the ssh command from window#1 and press enter. In this case, it would be

```
ssh -L 39689:localhost:39689 user@biowulf.nih.gov
```

Don't use this exact command, use what the output from window#1 tells you to do. It will prompt you with the login page to Biowulf, so enter your password.

Then in the original terminal window #1, load the rstudio-server module

```
module load rstudio-server
```

It will then load the necessary modules like so:

```
[+] Loading gcc 11.3.0 ...
[+] Loading HDF5 1.12.2
[+] Loading netcdf 4.9.0
[-] Unloading gcc 11.3.0 ...
[+] Loading gcc 11.3.0 ...
[+] Loading openmpi/4.1.3/gcc-11.3.0 ...
[+] Loading pandoc 2.18 on cn4311
[+] Loading TeX 2024
[+] Loading pcre2 10.40
[+] Loading libtiff libraries v4.6.0 ...
[+] Loading R 4.4.1
[+] Loading rstudio-server 2024.04.1-748
```

Continuing in terminal window #1, run R Studio Server with the following command  
`rstudio-server`

It will output something like this:

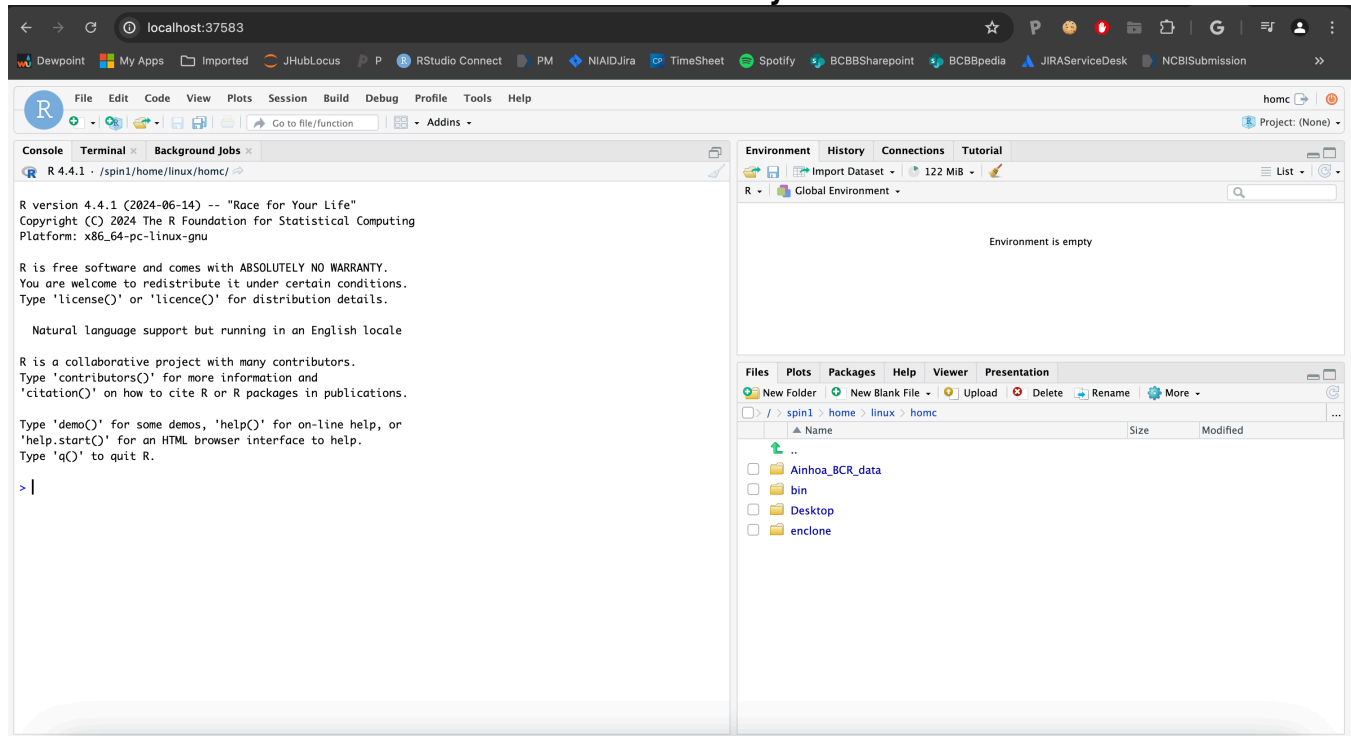
```
Please ensure you have set up the SSH port forwarding as described in
the sinteractive instructions.

Please connect to http://localhost:39689/auth-sign-
in?user=test2&password=nRmzfPWh_X8Z-03hbDjPz3bm
Use your username 'user' and the pasword 'nRmzfPWh_X8Z-03hbDjPz3bm'
to login
```

**Keep this terminal window #1 open until you are done using RStudio at the end of your analysis.**

Then, open up Google Chrome or other browser and go to the URL indicated and login with the autogenerated username and password. In the case of the above, you would open the URL: `http://localhost:39689/auth-sign-in?user=test2&password=nRmzfPWh_X8Z-03hbDjPz3bm`  
Don't use this URL, use what the output from window#1 tells you to do.

You should then be able to see RStudio loaded in your browser like so:



By default, your working directory will be set to `/home/username` but the amount of space you have in your home directory on Biowulf is very strictly limited, so you will want to work in your data directory. Within RStudio you will want to run

```
setwd("/data/username/someprojectfolder")
```

Modify the command (replace username with yours) above to point it to a project folder you have created within your `/data/username` directory on Biowulf (you can always request more space in your data directory as needed on the Biowulf HPC website <https://hpc.nih.gov/storage/>, you get 100GB to start with by default)

**At the end of your analysis, when you are done using RStudio, you can close your browser window/tab and then you can also close terminal window #1 (the one used to open the tunneling session) using **ctrl-C****

**Save your work! Remember to save any `.R` and `.RData` (or `.qmd` or `.rmd`) analysis and data files in your project folder within `/data/username/projectfolder` so that you can retrieve them again. Any data saved in the scratch directory is automatically wiped clean on Biowulf regularly.**