

Git versioning: Keep track of your changes

Physalia course 2023

Instructor: Jacques Serizay

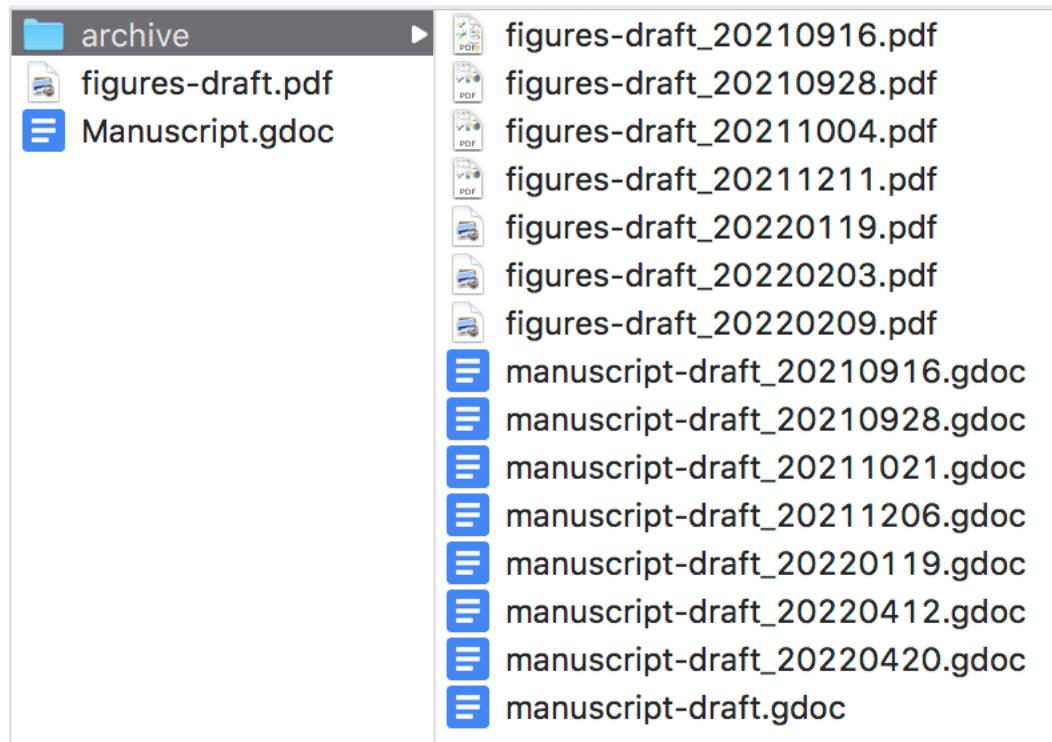
Keeping track of changes with git

<https://happygitwithr.com/>

- Git is a version control system.
- Git manages the evolution of a set of files – called a **repository** – in a sane, highly structured way.
- Its original purpose was to help groups of developers work collaboratively on big software projects.
- Essentially, this is “Track Changes” features from Microsoft Word on steroids.

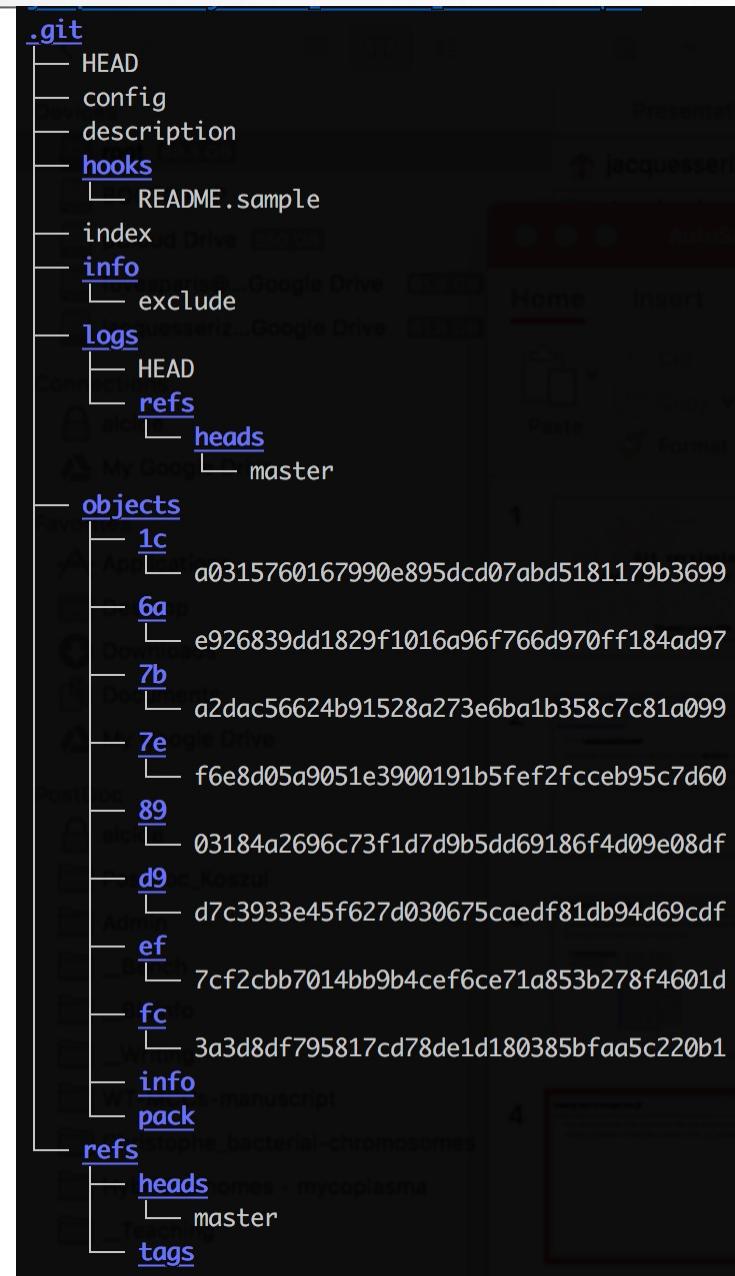
Keeping track of changes with git

- Ever time-stamped a file? Git does the same, for one or several files, and allows you to revert back to previous files or compare changes (for text files).



Keeping track of changes with git

- Your working directory only contains the 2 files (pdf and gdoc) you are working on, and changes of these files are stored in the ` `.git` subfolder, in a compressed way.



Keeping track of changes with git

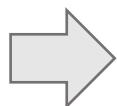
Git's main downside is... git is not user-friendly!

- It's painful to visualize changes, to browse history, to compare files, ...
- No native graphical interface

Git vs github

- Git is a software, locally installed on one's computer

```
$> git --help
```

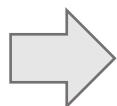


“Track Changes” features

Git vs github

- Git is a software, locally installed on one's computer

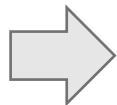
```
$> git
```



“Track Changes” features

- Github is an Internet hosting service, storing your local `git` repository (files AND changes) on-line (“remotely”):

```
https://github.com/
```



Dropbox features

Git vs github

- You are working on your manuscript draft in Word: you may rewrite sentences, resolve previous comments, reply to others still open, ... As long as you are working on your own, this works.
- Now you want to share it with others → you decide to put the document online as a Google Doc, allow access to others, and enable synchronization with your local Google Drive.
- Your colleagues will eventually rewrite/resolve/reply together with you (sometimes with conflicting changes!), and everyone will see who is doing what, etc.

The Google Doc is your “remote repository” and your local Google Drive folder is your “local repository”

Typically, both are continuously synced, whereas Git and Github are not sync-ed “live”, you do it manually and whenever you want.

Git vs github

Github, gitlab, bitbucket: all different internet hosting services doing the same thing:

1. Storing your local git repositories online, to make available to others.
2. Allow specific online-based features: issues and pull requests
3. Easy “time-travel” (i.e. browsing files back in time) with permanent links
4. Search function!

Git vs github

Github, gitlab, bitbucket: all different internet hosting services doing the same thing:

1. Storing your local git repositories online, to make available to others.

The screenshot shows a GitHub repository page for 'js2264 / HiContacts'. The top navigation bar includes 'Pin', 'Unwatch 2', 'Fork 1', and 'Star 0'. Below the bar are links for 'Code', 'Issues', 'Pull requests', 'Actions', 'Projects', 'Wiki', 'Security', 'Insights', and 'Settings'. The main content area shows a file tree on the left and a 'Clone' section on the right. The 'Clone' section offers three methods: HTTPS (selected), SSH, and GitHub CLI. The SSH URL 'git@github.com:js2264/HiContacts.git' is highlighted with a red box. Below it are buttons for 'Open with GitHub Desktop' and 'Download ZIP', also highlighted with red boxes. To the right of the clone section is an 'About' summary.

js2264 / HiContacts Public

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master 4 branches 0 tags Go to file Add file Code

js2264 bump 0.99.11

.github add localDistance

R replace data.table

data replace data files

inst/extdata add biocthis GA

man Cleanup methods

tests Cleanup methods

vignettes contacts class/function renamed to Contacts 25 days ago

.Rbuildignore add biocthis GA last month

.gitignore remove unnecessary functions last month

DESCRIPTION bump 0.99.11 4 days ago

LICENSE Initial commit 16 months ago

LICENSE.md Initial commit 16 months ago

Clone

HTTPS SSH GitHub CLI

git@github.com:js2264/HiContacts.git

Use a password-protected SSH key.

Open with GitHub Desktop

Download ZIP

About

HiContacts: R interface to cool files and analysis of Hi-C contacts

js2264.com/HiContacts/

Readme

Unknown, MIT licenses found

0 stars

2 watching

1 fork

Releases

No releases published

Create a new release

Packages

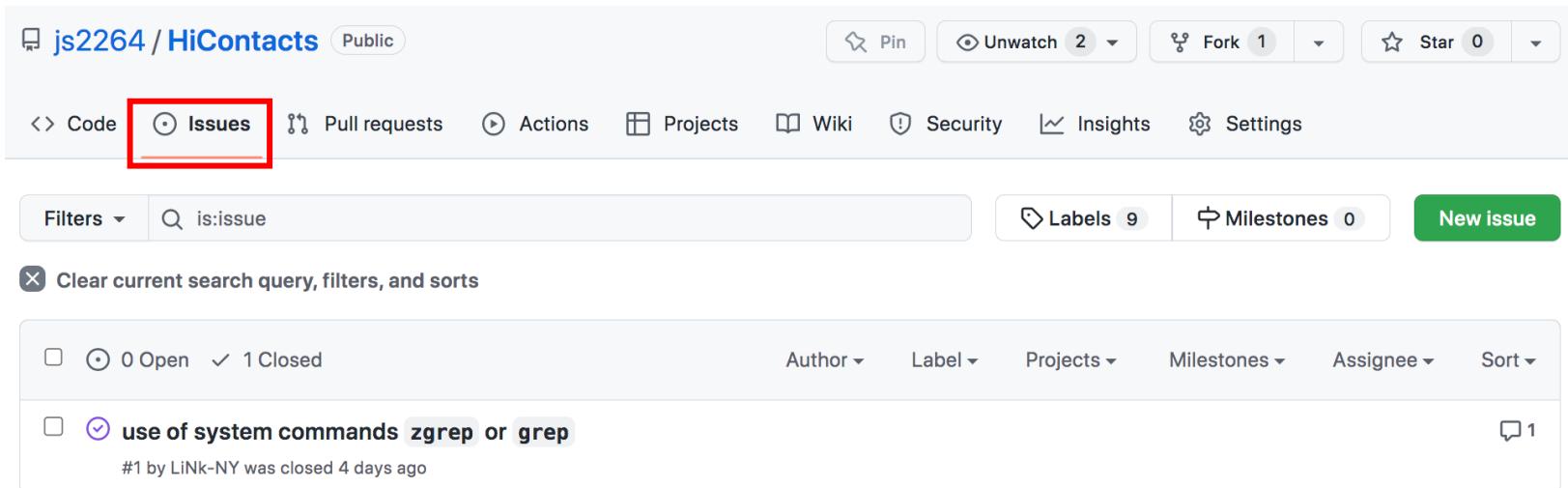
No packages published

Publish your first package

Git vs github

Github, gitlab, bitbucket: all different internet hosting services doing the same thing:

2. Allow specific online-based features: issues and pull requests



The screenshot shows a GitHub repository page for 'js2264 / HiContacts'. The 'Issues' tab is highlighted with a red box. The page includes standard GitHub navigation links like Code, Pull requests, Actions, Projects, Wiki, Security, Insights, and Settings. Below the navigation is a search bar with 'is:issue' and buttons for Filters, Labels (9), Milestones (0), and New issue. A clear search query button is also present. At the bottom, there are filters for Open (radio button) and Closed (checkbox) issues, along with dropdowns for Author, Label, Projects, Milestones, Assignee, and Sort. A specific issue is listed: '#1 by LiNk-NY was closed 4 days ago' with a note about using system commands zgrep or grep.

Git vs github

Github, gitlab, bitbucket: all different internet hosting services doing the same thing:

2. Allow specific online-based features: issues and pull requests

The screenshot shows a GitHub repository page for 'koszullab / hicstuff'. The 'Pull requests' tab is highlighted with a red box. The interface includes standard GitHub navigation like Code, Issues (3), Actions, Projects, Wiki, Security, Insights, and Settings. Below the tabs are filters ('is:pr'), labels (9), milestones (0), and a 'New pull request' button. A search bar at the top says 'Clear current search query, filters, and sorts'. The main area displays a list of pull requests:

checkbox	title	author	status	last update	comment count
<input type="checkbox"/>	add prefix to fig_dir	js2264	✓ 11 Closed	Jan 27	1
<input type="checkbox"/>	Replace travis with github workflow	cmdoret	✗	Jan 21	1
<input type="checkbox"/>	Digest map	ABignaud	✓	Apr 19, 2021	Approved
<input type="checkbox"/>	Move to BSD 3-Clause License	cmdoret	✗	Oct 15, 2020	
<input type="checkbox"/>	Throw instead of returning input in normalization	scovit	✗	Apr 30, 2020	2

Git vs github

Github, gitlab, bitbucket: all different internet hosting services doing the same thing:

3. Easy "time-travel" (i.e. browsing files back in time) with permanent links

The screenshot shows a GitHub repository page for `koszullab/hicstuff`. The top navigation bar includes options like Edit Pins, Watch (8), Fork (8), and Star (28). Below the header, there are tabs for Code, Issues (3), Pull requests, Actions, Projects, Wiki, Security, Insights, and Settings. The `Code` tab is selected. The main content area displays the history of the `README.md` file. It shows four commits from different dates, each with a detailed view of the commit message, author (cmdoret), date, and a green checkmark indicating it's a verified commit. The commits are as follows:

- Commits on Oct 10, 2022:
 - Add instructions for conda and docker installation (cmdoret committed 7 days ago) - Verified, SHA: 89891c3
- Commits on Jan 20, 2022:
 - Replace travis status badge with gh-workflow (cmdoret committed on Jan 20) - Verified, SHA: 5f606be
- Commits on Aug 24, 2021:
 - update travis badge in readme (cmdoret committed on Aug 24, 2021) - Verified, SHA: ca735b7
- Commits on May 19, 2021:
 - fix pairs specification in readme (cmdoret committed on May 19, 2021) - SHA: 17b5a21

Git vs github

Github, gitlab, bitbucket: all different internet hosting services doing the same thing:

3. Easy "time-travel" (i.e. browsing files back in time) with permanent links

The screenshot shows a GitHub repository page for 'koszullab/hicstuff'. The 'Code' tab is selected, and the URL 'hicstuff / README.md' is visible. The page displays the history of the README.md file, which has 100644 commits. The first commit is an 'Initial commit' from 7 years ago. The last commit is 'Add instructions for conda and docker installation' from 7 days ago. The commit history includes various updates like 'Update README.md', 'Packaging, merging hicstuff and yahcp', and 'Add instructions for conda and docker installation'. The right side of the screen shows the actual content of the README.md file, which is a detailed documentation for the 'hicstuff' library.

```
100644 | 390 lines (287 sloc) | 18.1 KB
Initial commit 7 years ago
Update README.md 7 years ago
Packaging, merging hicstuff and yahcp 4 years ago
Add instructions for conda and docker installation 7 days ago
Replace travis status badge with gh-workflow 9 months ago
Add zenodo DOI badge to README 3 years ago
Coverage badge 4 years ago
Packaging, merging hicstuff and yahcp 4 years ago
Binder link 4 years ago
Packaging, merging hicstuff and yahcp 4 years ago
Update README.md 7 years ago
update readme with new pipeline commands 4 years ago
Add instructions for conda and docker installation 7 days ago
```

hicstuff

[! [PyPI version](https://badge.fury.io/py/hicstuff.svg)](https://pypi.org/project/hicstuff/)

[! [PyPI - Python Version](https://img.shields.io/pypi/pyversion/hicstuff))

[! [Install with bioconda](https://img.shields.io/badge/install%20with-bioconda-brightgreen)](https://bioconda.github.io/recipes/hicstuff/)

[! [Build Status](https://github.com/koszullab/hicstuff/actions))

[! [DOI](https://zenodo.org/badge/DOI/10.5281/zenodo.2620601))

[! [codecov](https://codecov.io/gh/koszullab/hicstuff/branch/main))

[! [Read the docs](https://readthedocs.org/projects/hicstuff/))

[! [Binder](https://mybinder.org/badge_logo.svg)](https://mybinder.org/v2/gh/koszullab/hicstuff/main)

[! [Code style: black](https://img.shields.io/badge/code-style-black)](https://img.shields.io/badge/code-style-black)

Table of contents

- * [Installation](#installation)
- * [pip](#pip)
- * [conda](#conda)
- * [docker](#docker)
- * [usage](#usage)
- * [full pipeline](#full-pipeline)
- * [Individual components](#individual-components)
- * [library](#library)
- * [Connecting the modules](#connecting-the-modules)
- * [File formats](#file-formats)
- * [Contributing](#contributing)

Github and organizations

- A public repository is readable by the world. The owner can grant higher levels of permission to others, such as the ability to push commits.
- A private repository is invisible to the world. The owner can grant read, write (push), or admin access to others.
- There is also a notion of an organization, which can be useful for managing repository permissions for entire teams of people.

Github and organizations

- A public repository is readable by the world. The owner can grant higher levels of permission to others, such as the ability to push commits.
- A private repository is invisible to the world. The owner can grant read, write (push), or admin access to others.
- There is also a notion of an organization, which can be useful for managing repository permissions for entire teams of people.

→ Git has been re-purposed by the data science community.

- We now also use it, IN COLLABORATION WITH A CLOUD-BASED GIT HOST, to manage the diverse collection of files that make up typical data analytical projects, which often consist of data, figures, reports, and, yes, source code.

Github and organizations

R infrastructure

Overview Repositories 136 Projects Packages People 15

Pinned

- devtools** Public Tools to make an R developer's life easier
- roxygen2** Public Generate R package documentation from inline R comments
- testthat** Public An R to make testing 😊
- pkgdown** Public Generate static html documentation for an R package
- useRis** Public Set up commonly used components
- http** Public http: a friendly http package for R

Bioconductor

Software for the analysis and comprehension of high-throughput genomic data
49 followers <https://bioconductor.org>

Overview Repositories 260 Projects 2 Packages People 2

Pinned

- Contributions** Public Contribute Packages to Bioconductor
- BioStickers** Public Stickers for some Bioconductor packages - feel free to contribute and/or modify.
- bioconductor_docker** Public Docker Containers for Bioconductor - NEW!
- bioconductor.org** Public Source code for the Bioconductor website
- BBS** Public The Bioconductor Build System

tidyverse

The tidyverse is a collection of R packages that share common principles and are designed to work together seamlessly
361 followers <http://tidyverse.org> Verified

Overview Repositories 37 Projects Packages People 27

Pinned

- dplyr** Public dplyr: A grammar of data manipulation
- ggplot2** Public An implementation of the Grammar of Graphics in R
- readr** Public Read flat files (csv, tsv, fwf) into R
- purrr** Public A functional programming toolkit for R
- tidyr** Public Tidy Messy Data
- rvest** Public Simple web scraping for R

rOpenSci

190 followers Berkeley, CA <https://ropensci.org/> @ropensci info@ropensci.org Verified

Overview Repositories 328 Projects 1 Packages People 79

README.md

Welcome to the rOpenSci Project on GitHub. Our repositories are currently distributed across a few different organizations.

You can explore our [resources page](#) to find what you're after, it includes a list of all our GitHub organizations.

Also refer to our [code of conduct](#).

Pinned

- software-review** Public rOpenSci Software Peer Review
- software-review-meta** Public For organizing projects related to rOpenSci Software Peer Review
- dev_guide** Public rOpenSci Packages: Development, Maintenance, and Peer Review
- rtweet** Public R client for interacting with Twitter's [stream and REST] APIs

How to set things up

1. Register for a GitHub account
2. **MAKE SURE YOU ARE RUNNING R/RStudio/BiocManager WITH THE APPROPRIATE VERSION!**
3. Install Git: <https://git-scm.com/downloads>, or:
 - Linux: `sudo apt-get install git`
 - Mac: `xcode-select –install` (needs Xcode)
 - Windows: `choco install git.install` (needs Chocolatey)

How to set things up

1. Register for a GitHub account
2. MAKE SURE YOU ARE RUNNING R/RStudio/BiocManager WITH THE APPROPRIATE VERSION!
3. Install Git: <https://git-scm.com/downloads>, or:
 - Linux: `sudo apt-get install git`
 - Mac: `xcode-select –install` (needs Xcode)
 - Windows: `choco install git.install` (needs Chocolatey)

```
usethis::git_sitrep()
gert::git_config_global_set("user.name", "Jacques")
gert ::git_config_global_set("user.email", "jacquessserizay@gmail.com")
usethis ::git_vaccinate()
usethis ::use_git()
usethis ::use_git_ignore("*.Rproj")

# Don't forget to make an initial commit!
```

4. Configure git from R (easy)

How to set things up

5. Configure github (harder):

- Git can communicate with a remote server using one of two protocols, HTTPS or SSH, and the different protocols use different credentials.
- We must use SSH, as this means of communications is required later on, for Bioconductor submission.
- For that, we need to generate a pair of private/public SSH keys, connect to GitHub from R and push the public key to GitHub.

How to set things up

5. Configure github (harder):

- Git can communicate with a remote server using one of two protocols, HTTPS or SSH, and the different protocols use different credentials.
- We must use SSH, as this means of communications is required later on, for Bioconductor submission.
- For that, we need to generate a pair of private/public SSH keys, connect to GitHub from R and push the public key to GitHub.

```
gh::gh_whoami()
usethis::gh_token_help()

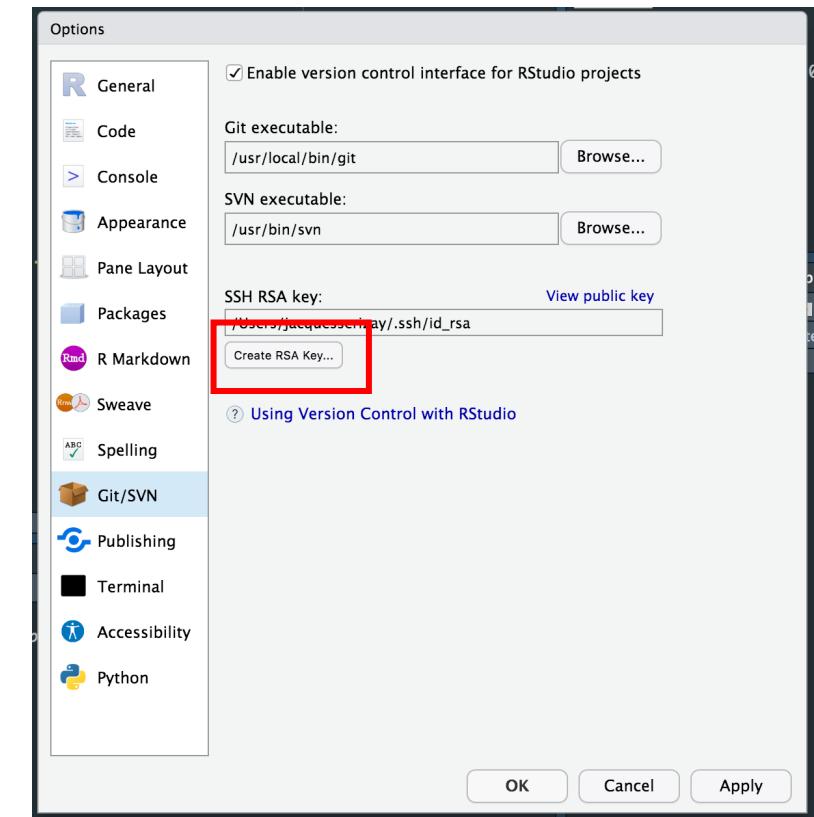
#Create a GH PAT
usethis ::create_github_token()
gitcreds ::gitcreds_set()
#Paste your GitHub token, not your password!

#Add an SSH key to GitHub
usethis ::use_github()
credentials::ssh_setup_github()
```

How to set things up

5. Configure github (harder):

- Fallback approach: create a key in RStudio



How to set things up

5. Configure github (harder):

- Then, we need to upload the public key to GitHub

Jacques Serizay
Your personal account [Switch to another account ▾](#)

Public profile Account Appearance Accessibility Notifications Access Billing and plans Emails **Password and authentication** **SSH and GPG keys** Organizations Moderation

Code, planning, and automation Repositories Packages GitHub Copilot Pages Saved replies

SSH keys

This is a list of SSH keys associated with your account. Remove any keys that you do not recognize.

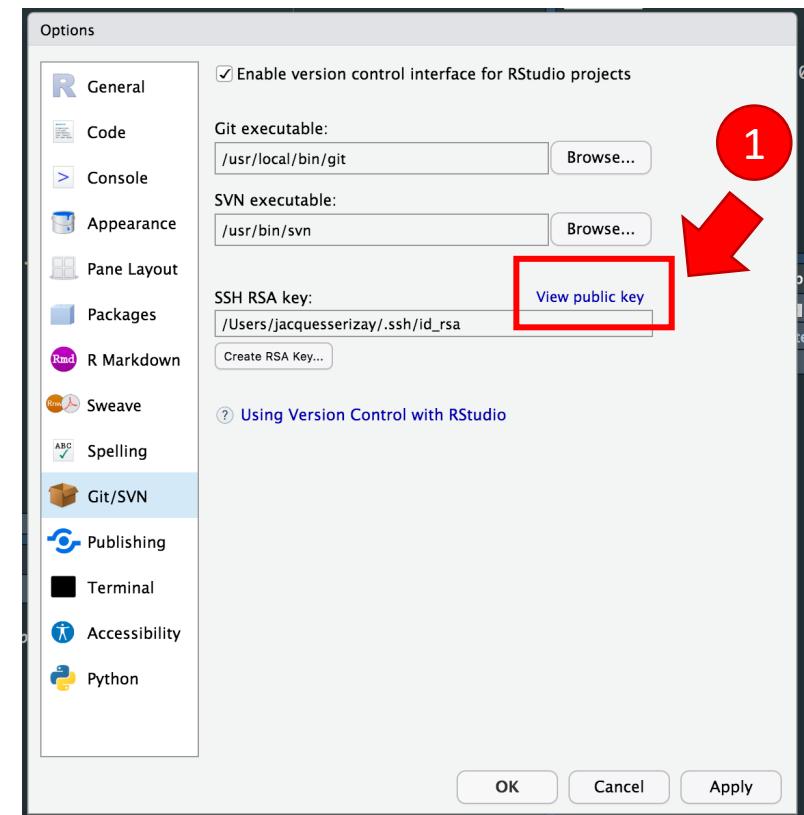
Authentication Keys

Key	Name	Added	Last used	Type	Action
SSH	rsg_koszul	Added on Dec 18, 2020	Last used within the last 2 weeks — Read/write	SSH	Delete
SSH	aws	Added on Sep 16, 2022	Never used — Read/write	SSH	Delete
SSH	aws	Added on Sep 16, 2022	Never used — Read/write	SSH	Delete

[New SSH key](#)

GPG keys

New GPG key



How to set things up

5. Configure github (harder):

- Test it with `ssh -T git@github.com`

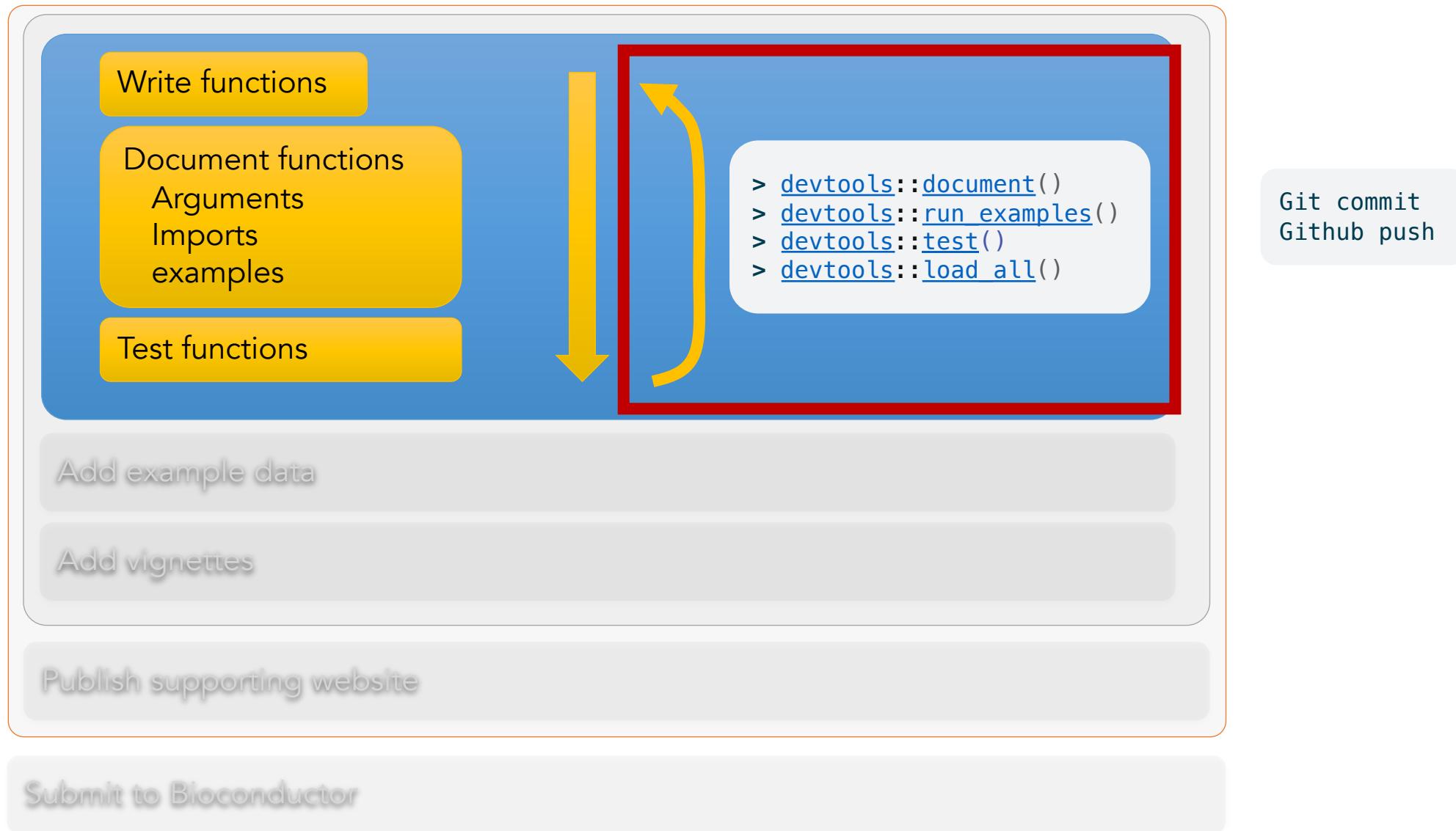
```
jacquesserizay@LOCAL[13:54:27]:~ $ ssh -T git@github.com
▶Hi js2264! You've successfully authenticated, but GitHub
does not provide shell access.
```

How to set things up

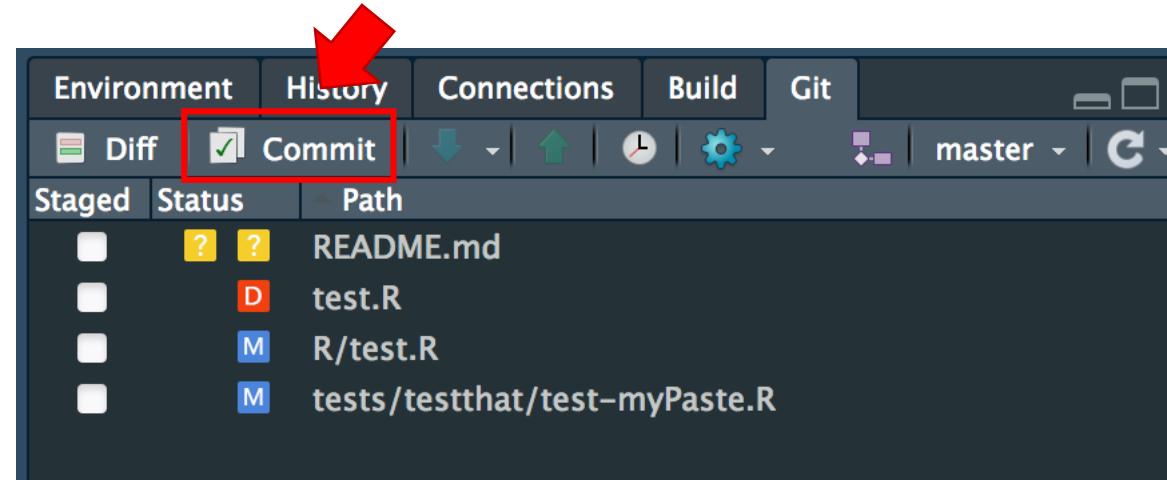
If you are still struggling...

<https://happygitwithr.com/usage-intro.html#usage-intro>

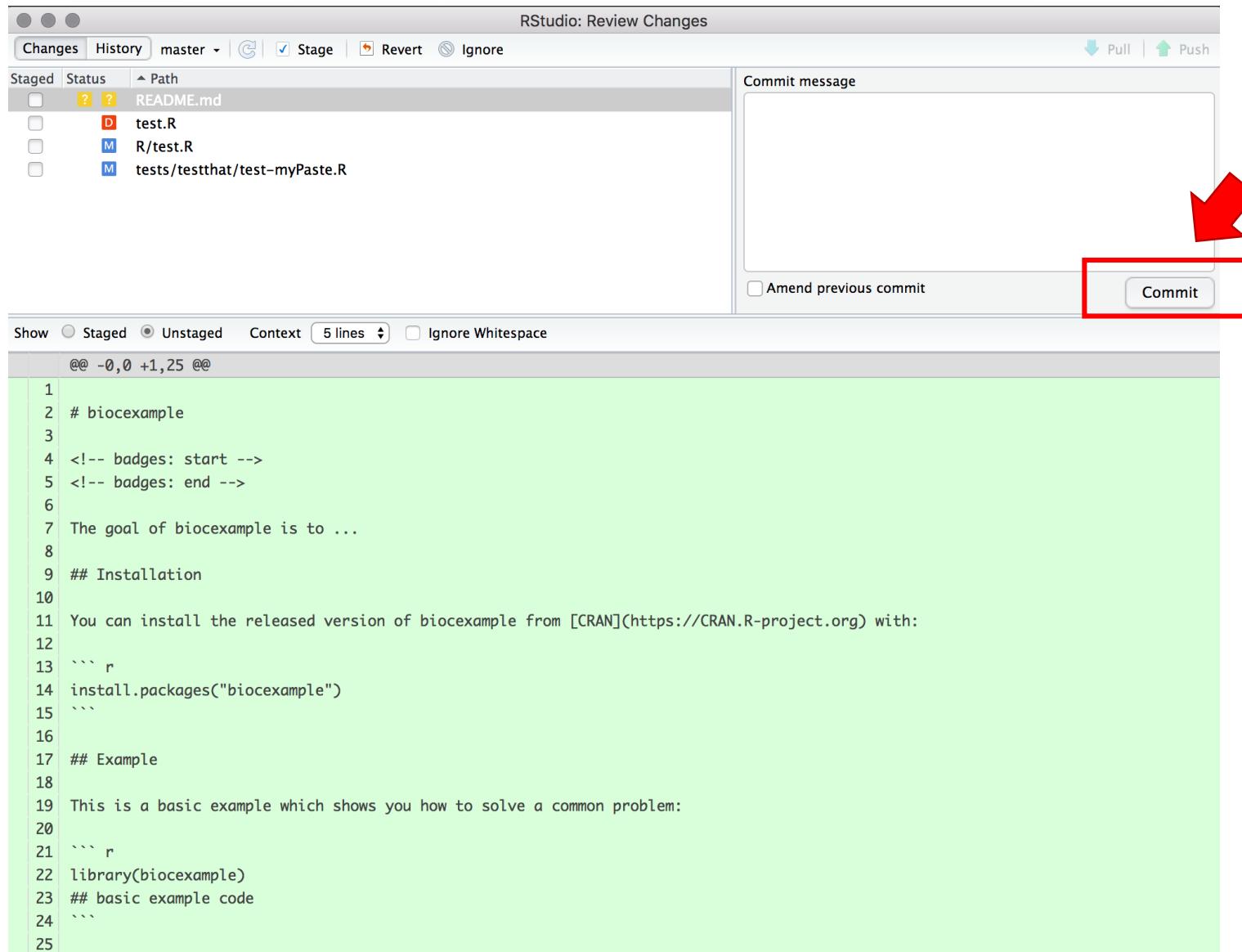
Lather-rinse-repeat



Lather-rinse-repeat



Lather-rinse-repeat



Lather-rinse-repeat

