

nf-core



Getting started with nf-core

Nextflow Camp 2019 - Updated Feb 2020

Phil Ewels
phil.ewels@scilifelab.se

What is nf-core?

Deploy



Stable pipelines



Centralised
configs



List and update
pipelines



Download
for offline use

Participate



Documentation



Slack workspace



Twitter updates



Hackathons

Develop



Starter template



Code guidelines



CI code linting
and tests



Helper tools

Tutorial



https://nf-co.re/usage/nf_core_tutorial

- 1 Introduction
- 2 Listing
- 3 Running
- 4 Creating
- 5 Testing
- 6 Releasing

TLDR

TLDR pages *(too long, didn't read)*

Simplified and community-driven man pages

Super-fast and easy reference for the command line

<https://tldr.sh>

pip install tldr

brew install tldr

TLDR

- List currently running docker containers:
`docker ps`
- List all docker containers (running and stopped):
`docker ps -a`
- Start a container from an image, with a custom name:
`docker run --name container_name image`
- Start or stop an existing container:
`docker start|stop container_name`
- Pull an image from a docker registry:
`docker pull image`
- Open a shell inside of an already running container:
`docker exec -it container_name sh`
- Remove a stopped container:
`docker rm container_name`
- Fetch and follow the logs of a container:
`docker logs -f container_name`

Help!



Help!



slack

<https://nf-co.re/join>

Help!

nf-core ▾

- Phil Ewels 📝
- # general
- # github-invitations
- 🔒 hackathon-uk19
- # help**
- # igenomes
- # ismb-2019
- # linting
- # local-testing
- # mag
- 🔒 material
- # mhquant
- # modules
- # new-pipelines
- # pipelines
- # random
- # rnafusion
- # rnaseq
- # scrnaseq
- # singularity
- # tools
- # weblog
- # website

🔔

#help

☆ | 143 | 0 | Help me!

📞 ⓘ ⚙️ 🔍 Search @ ⭐️ ⋮

Today

hpatel 17:01

Should really check that the basic GitHub settings have been applied to all pipelines currently forked on [nf-core](#). Ill do that for the ones I can access [Settings](#)

hpatel 17:31

Untitled ▾

```
1 https://github.com/nf-core/scrnaseq
2 https://github.com/nf-core/bactmap
3 https://github.com/nf-core/nascent
4 https://github.com/nf-core/epitopeprediction
5 https://github.com/nf-core/hic
```

I cant seem to access the [Settings](#) tab for the pipelines above. Is this something that needs to be set upstream by the contributor of the fork?

Checked and amended the rest of the pipelines to have consistent settings

Alex Peltzer 🌴 17:32

Actually it shouldn't

We might have to add team members to these

hpatel 17:38

Ok. Im guessing you need [Owner](#) access to do that because I thought that was done via [Settings](#)? 😅

Phil Ewels 📝 21:22

Alex and I are owners so have full access to everything

We should add the admin team as admins to all repos though

Helper tools

nf-core/ **tools**



```
pip install nf-core
```



BIOCONDA[®]

```
conda install -c bioconda nf-core
```

Helper tools

nf-core --help

```
usage: nf-core [OPTIONS] COMMAND [ARGS]...

Options:
  --version      Show the version and exit.
  -v, --verbose   Verbose output (print debug statements)
  --help         Show this message and exit.

Commands:
  list          List nf-core pipelines with local info
  launch        Run pipeline, interactive parameter prompts
  download     Download a pipeline and singularity container
  licences      List software licences for a given workflow
  create        Create a new pipeline using the template
  lint          Check pipeline against nf-core guidelines
  bump-version  Update nf-core pipeline version number
```

Helper tools

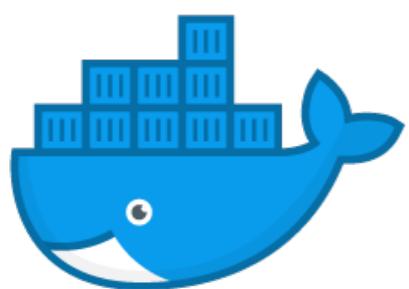
nf-core list

Name	Version	Released	Last Pulled	Have latest release?
nf-core/eager	2.0.7	2 weeks ago	-	-
nf-core/chipseq	1.0.0	2 weeks ago	-	-
nf-core/rnafusion	1.0.2	1 months ago	-	-
nf-core/hic	1.0.0	1 months ago	-	-
nf-core/bacass	1.0.0	1 months ago	-	-
nf-core/coproid	1.0	1 months ago	-	-
nf-core/bcellmagic	1.0.0	2 months ago	-	-
nf-core/nascent	1.0	2 months ago	-	-
nf-core/atacseq	1.0.0	2 months ago	-	-
nf-core/rnaseq	1.3	3 months ago	2 weeks ago	Yes
nf-core/hlatyping	1.1.4	3 months ago	9 months ago	No
nf-core/mhcquant	1.2.6	3 months ago	-	-
nf-core/methylseq	1.3	4 months ago	1 year ago	No
nf-core/ampliseq	1.0.0	6 months ago	-	-
nf-core/deepvariant	1.0	7 months ago	-	-
nf-core/bactmap	dev	-	-	-
nf-core/cageseq	dev	-	-	-
nf-core/clinvap	dev	-	-	-
nf-core/ddamsproteomics	dev	-	-	-
nf-core/epitopeprediction	dev	-	-	-
nf-core/exoseq	dev	-	-	-
nf-core/guideseq	dev	-	-	-
nf-core/lncpipe	dev	-	-	-

Running pipelines

Software required

nextflow



docker

OR



OR

CONDA

Software required

```
nextflow run nf-core/rnaseq
```

```
-r dev
```

```
nextflow pull nf-core/rnaseq
```

Documentation

Documentation

The nf-core/rnaseq pipeline comes with documentation about the pipeline, found in the `docs/` directory:

1. [Installation](#)
2. Pipeline configuration
 - [Local installation](#)
 - [Adding your own system](#)
 - [Reference genomes](#)
3. [Running the pipeline](#)
4. [Output and how to interpret the results](#)
5. [Troubleshooting](#)

<https://nf-co.re/PIPELINE-NAME/docs>

Documentation

```
nextflow run nf-core/rnaseq --help
```

```
N E X T F L O W ~ version 19.01.0
Launching `nf-core/rnaseq` [jovial_feynman] - revision: 37f260d360 [master]
-----
          ,--./,--.
          /,-.-.-~'
  ,{ } {
  \`-.,-`-
  .-,.,

nf-core/rnaseq v1.3
-----
```

Usage:

The typical command for running the pipeline is as follows:

```
nextflow run nf-core/rnaseq --reads '*_R{1,2}.fastq.gz' --genome GRCh37 -profile docker
```

Mandatory arguments:

--reads	Path to input data (must be surrounded with quotes)
-profile	Configuration profile to use. Can use multiple (comma separated) Available: conda, docker, singularity, awsbatch, test and more.

Options:

--genome	Name of iGenomes reference
--singleEnd	Specifies that the input is single end reads

Strandedness:

--forward_stranded	The library is forward stranded
--reverse_stranded	The library is reverse stranded
--unstranded	The default behaviour

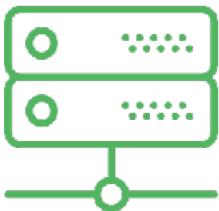
Config setup



Default 'base' config (always loaded)



Core profiles (eg. docker, conda, test)



Server profiles (nf-core/configs)



Your local config files (eg. the -c flag)

Config setup

<https://www.nextflow.io/docs/latest/config.html#config-profiles>

! Tip

Two or more configuration profiles can be specified by separating the profile names with a comma character, for example:

```
nextflow run <your script> -profile standard,cloud
```

Test data

```
nextflow run nf-core/rnaseq -profile test,docker
```

nf-core/ 
test-datasets

Test data

```
nextflow run nf-core/rnaseq -profile test,docker
```

```
Data Type : Single-End
Strandedness : None
Trimming : 5'R1: 0 / 5'R2: 0 / 3'R1: 0 / 3'R2: 0
Aligner : STAR
Fasta Ref : https://github.com/nf-core/test-datasets/raw/rnaseq/reference/genome.fa
GTF Annotation : https://github.com/nf-core/test-datasets/raw/rnaseq/reference/genes.gtf
Save prefs : Ref Genome: No / Trimmed FastQ: No / Alignment intermediates: No
Max Resources : 6 GB memory, 2 cpus, 2d time per job
Container : docker - nfcore/rnaseq:1.3
Output dir : ./results
Launch dir : /Users/philewels/test
Working dir : /Users/philewels/test/work
Script dir : /Users/philewels/.nextflow/assets/nf-core/rnaseq
User : philewels
Config Profile : test,docker
Config Description: Minimal test dataset to check pipeline function
[2m-----
[warm up] executor > local
executor > local (14)
[9b/6ab89a] process > get_software_versions [100%] 1 of 1 ✓
[c7/e53c44] process > output_documentation [100%] 1 of 1 ✓
[81/b17819] process > makeBED12 [100%] 1 of 1 ✓
[19/7c5e2e] process > makeSTARindex [100%] 1 of 1 ✓
[53/917cef] process > fastqc [100%] 4 of 4 ✓
[3d/15f1fc] process > trim_galore [100%] 4 of 4 ✓
[2d/9e59bb] process > star [  0%] 0 of 2
```

Running for real

If you've read the documentation..

```
nextflow run nf-core/rnaseq [...]
```

If not.. (command-line wizard with prompts)

```
nf-core launch rnaseq
```

If you're running offline..

```
nf-core download rnaseq
```

Exercises 1, 2, 3

Exercise 1 (installation)

- Install nf-core/tools
- Use the help flag to list the available commands

Exercise 2 (listing pipelines)

- Show the `nf-core list` command usage
- List all pipelines
- Sort pipelines alphabetically, then by popularity
- Fetch one of the pipelines using nextflow
- Use `nf-core list` to see if the pipeline you pulled is up to date
- Filter pipelines for any that work with RNA
- Save these pipeline details to a JSON file

Exercise 3 (using pipelines)

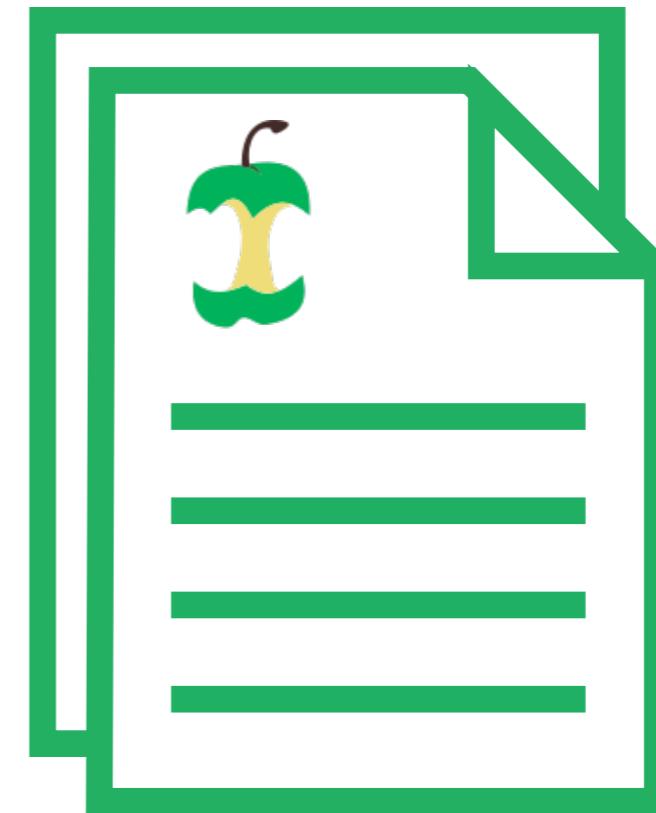
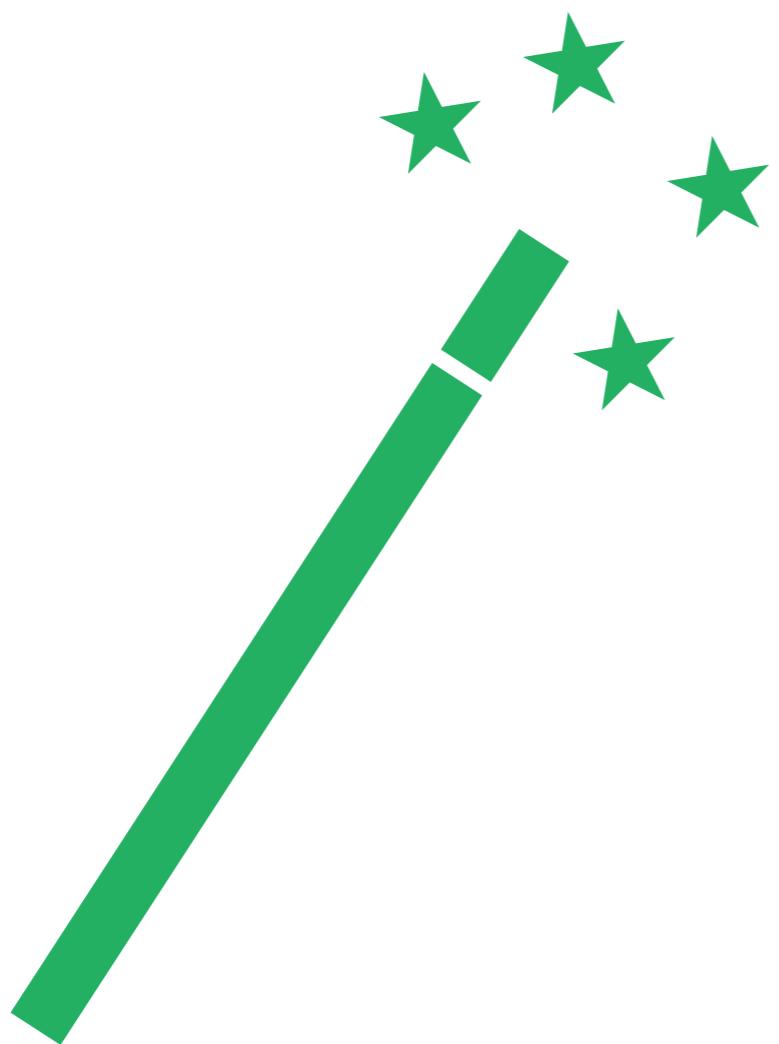
- Install software dependencies
- Print the command-line usage for the nf-core/rnaseq pipeline
- In a new directory, run the nf-core/rnaseq pipeline with the test profile
- Try launching the RNA pipeline using the `nf-core launch` command
- Download the nf-core/rnaseq pipeline for offline use using the `nf-core download` command

https://nf-co.re/usage/nf_core_tutorial

Creating pipelines



nf-core create



Template

nf-core create



Workflow Name: mypipeline

Description: This is a demo pipeline for the tutorial

Author: Phil Ewels

INFO: Creating new nf-core pipeline: nf-core/mypipeline

INFO: Initialising pipeline git repository

INFO: Done. Remember to add a remote and push to GitHub:

```
cd /Users/philewels/test/nf-core-mypipeline  
git remote add origin git@github.com:USERNAME/REPO_NAME.git  
git push --all origin
```

INFO: This will also push your newly created TEMPLATE branch for syncing.

Template

```
usage.md
49 NXF_OPTS='--Xms1g --Xmx4g'
50 ``
51 |
52 <!-- TODO nf-core: Document required command line parameters to run the pipeline-->
53 |
54 ## Running the pipeline
55 The typical command for running the pipeline is as follows:
56 |
57 ````bash
58 nextflow run nf-core/mypipeline --reads '*_R{1,2}.fastq.gz' -profile docker
59 ``
60 |
61 This will launch the pipeline with the `docker` configuration profile. See below for more
• information about profiles.
62 |
63 Note that the pipeline will create the following files in your working directory:
64 |
65 ````bash
66 work          # Directory containing the Nextflow working files
67 results        # Finished results (configurable, see below)
68 .nextflow_log # Log file from Nextflow
69 # Other Nextflow hidden files, e.g. history of pipeline runs and old logs.
70 ````
```

Template

```
usage.md
49 NXF_OPTS='--Xms1
50 ``
51 <-- TODO nf-co
52 ## Running the
53 The typical com
54 ``bash
55 nextflow run nf
56 ```
57 This will launc
58 • information abo
59 Note that the p
60 ``bash
61 work
62 results
63 .nextflow_log
64 # Other nextfl
65 ```
66 nf-core: Document re
67 nf-core: Update refer
68 nf-core: Describe ref
69 nf-core: Add required
70 nf-core: Add to this h
71 nf-core: Add any ref
72 nf-core: Report custo
```

are not needed

nf-core: Specify the paths to your test data on nf-core/test-datasets

nf-core: Give any req command line flags a

nf-core: Write this do workflow's output

nf-core: Document re run the pipeline

nf-core: Document re

nf-core: Update refer what is needed

nf-core: Describe ref

nf-core: Describe any

nf-core: Add required

nf-core: Add to this h parameters

nf-core: Add any ref

nf-core: Report custo

TODO

config/test

,--./,-.
/-._.-~\n } {
\`-.,-`-,
.~,.

Running pipeline tests [#####] 100% None

INFO: =====

LINTING RESULTS

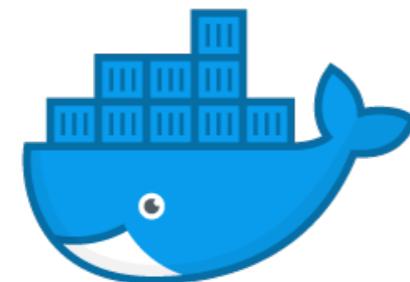
===== 60 tests passed 24 tests had warnings 0 tests failed

WARNING: Test Warnings:

http://nf-co.re/errors#10: TODO string found in 'environment.yml': Add require
http://nf-co.re/errors#10: TODO string found in 'README.md': Update the defa
http://nf-co.re/errors#10: TODO string found in 'README.md': Add a brief overv
http://nf-co.re/errors#10: TODO string found in 'README.md': Add citation fo
http://nf-co.re/errors#10: TODO string found in 'nextflow.config': Specify yo
http://nf-co.re/errors#10: TODO string found in 'main.nf': Add to this help
http://nf-co.re/errors#10: TODO string found in 'main.nf': Add any reference
http://nf-co.re/errors#10: TODO string found in 'main.nf': Report custom par
http://nf-co.re/errors#10: TODO string found in 'main.nf': Get all tools to
http://nf-co.re/errors#10: TODO string found in 'main.nf': Add in log files
http://nf-co.re/errors#10: TODO string found in 'main.nf': Specify which Mul
http://nf-co.re/errors#10: TODO string found in 'main.nf': If not using Mult

Software

CONDA



docker



environment.yml

Automated builds
hub.docker.com

Conversion done by
Nextflow at run time

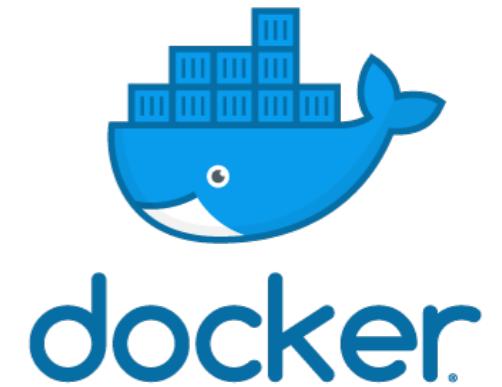
Software

```
1      This file to create a conda environment for this pipeline:  
2  #   conda env create -f environment.yml  
3  name: nf-core-methylseq-1.4dev  
4  channels:  
5  |  - bioconda  
6  |  - conda-forge  
7  |  - defaults  
8  dependencies:  
9  |  - fastqc=0.11.8  
10 |  # Default bismark pipeline  
11 |  - trim-galore=0.5.0  
12 |  - samtools=1.9  
13 |  - bowtie2=2.3.4.3  
14 |  - bismark=0.20.0  
15 |  - qualimap=2.2.2b  
16 |  - preseq=2.0.3  
17 |  - multiqc=1.7  
18 |  # bwa-meth pipeline  
19 |  - picard=2.18.21
```

environment.yml

Software

For local development with docker, build the image and tag



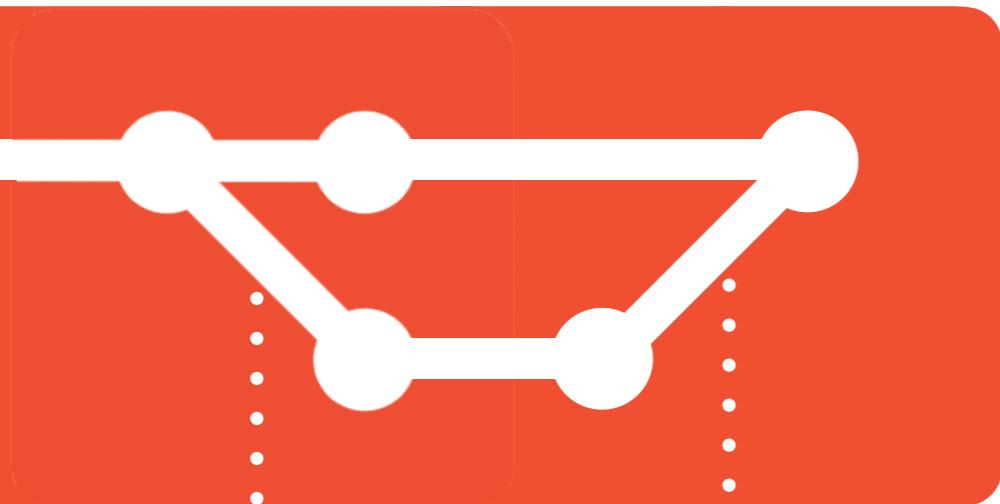
```
docker build -t nfcore/pipeline:dev .
```

Forks & branches



git

Forks & branches



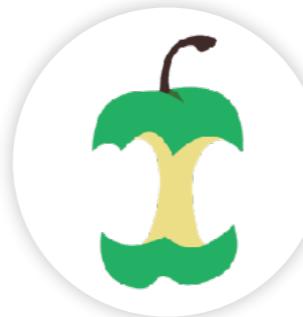
Branch

Merge

Forks & branches



🍴 Fork



@nf-core



@ewels

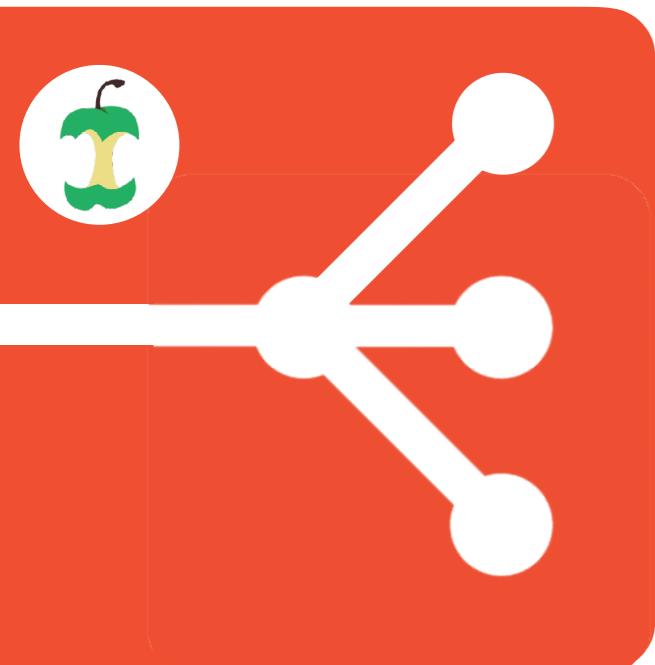
Forks & branches



Merge

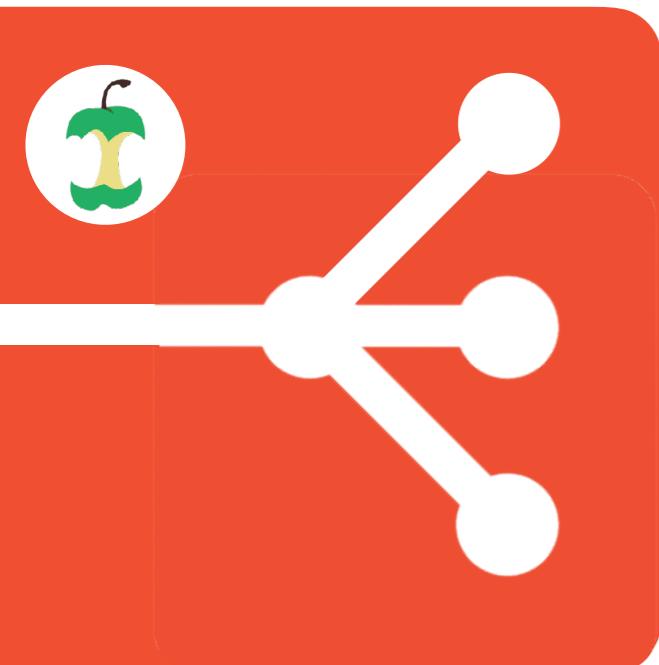
 Pull Request

Forks & branches



dev
master
TEMPLATE

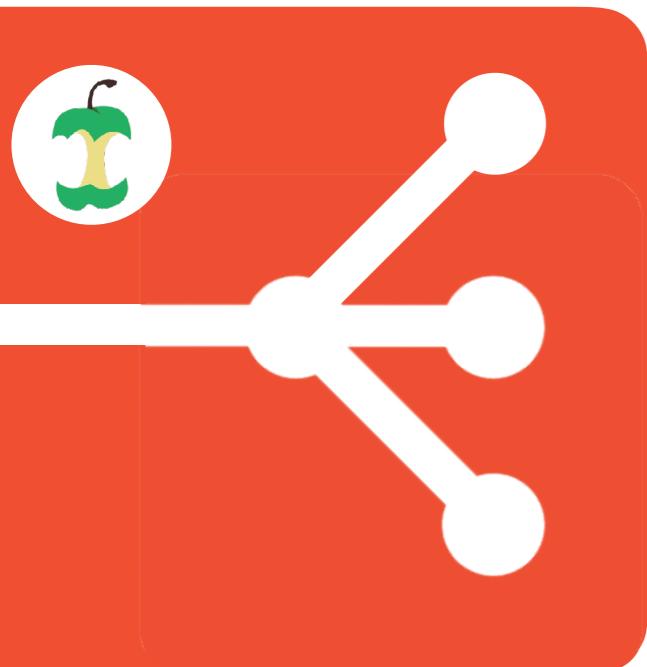
Forks & branches



dev
master
TEMPLATE

Stable releases only

Forks & branches



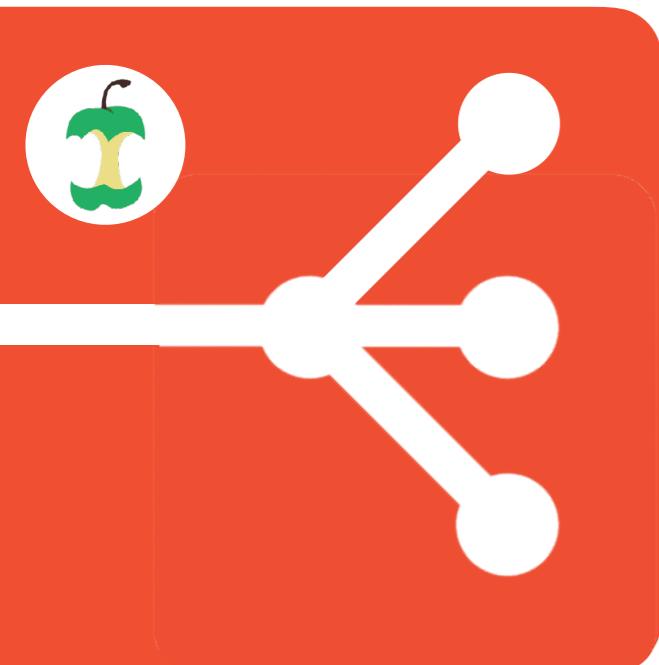
dev

master

TEMPLATE

Used by @nf-core-bot
to keep in sync

Forks & branches



dev
master
TEMPLATE

All development code
(and all pull requests)

Forks & branches

 REMEMBER!



PRs must be against the `dev` branch



Changelog must be updated



PRs need to be reviewed

Testing pipelines



Lint tests

nf-core lint .

```
Running pipeline tests [#####-----] 44% 'check_nextflow_config'  
ERROR: Found test failures in 'check_nextflow_config', halting lint run.
```

INFO: =====

LINTING RESULTS

=====

46 tests passed 0 tests had warnings 1 tests failed

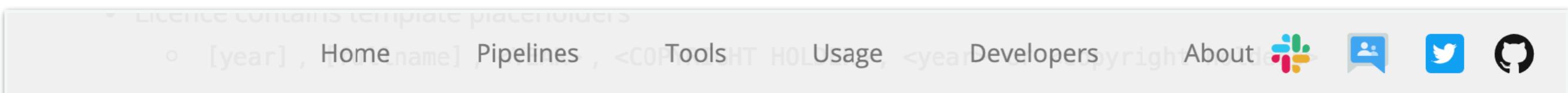
ERROR: Test Failures:

<http://nf-co.re/errors#4>: Config variable (incorrectly) found: params.container

ERROR: Sorry, some tests failed - exiting with a non-zero error code...

Lint tests

<https://nf-co.re/errors>



Error #4 - Nextflow config check failed

nf-core pipelines are required to be configured with a minimal set of variable names. This test fails or throws warnings if required variables are not set.

Note: These config variables must be set in `nextflow.config` or another config file imported from there.
Any variables set in nextflow script files (eg. `main.nf`) are not checked and will be assumed to be missing.

The following variables fail the test if missing:

- `params.outdir`
 - A directory in which all pipeline results should be saved
- `manifest.name`
 - The pipeline name. Should begin with `nf-core/`
- `manifest.description`
 - A description of the pipeline
- `manifest.version`
 - The version of this pipeline. This should correspond to a GitHub release.

Test data

```
nextflow run nf-core/rnaseq -profile test,docker
```

nf-core/ 
test-datasets

Test data

nf-core/ **test-datasets**



Make a new branch for your pipeline



Add a really (really!) tiny dataset



Set up the test config to use raw URLs

GitHub Actions



GitHub Actions

GitHub Actions

Screenshot of a GitHub repository page for `nf-core / sarek`.

The repository was forked from `MaxUlysse/sarek`. It has 6 watchers, 3 stars, and 82 forks.

The main navigation bar includes links for **Pull requests**, **Issues**, **Marketplace**, and **Explore**.

The repository has 1 pull request, 3 issues, and 16 checks.

A red circle highlights the **Checks** section.

The pull request #34, titled "update reference files", was merged by apeltzer yesterday. It has 16 checks and 4 files changed.

The CI status shows:

- Travis CI**: successful 2 days ago in 1m 7s
- sarek CI**: on: pull_request (status shown)
- test (19.04.0)**: successful 4s ago
- test**: successful 14s ago
- sarek linting**: on: pull_request (status shown)

A modal window displays the GitHub Actions log for the `nf-core` job, showing the following steps:

- ✓ Set up job
- ✓ Run actions/checkout@v1
- ✓ Install Nextflow
- ✓ Run actions/setup-python@v1
- ✓ Install pip

GitHub Actions

4s

✓ Install Nextflow

14s

✓ Run actions/setup-python@v1

0s

✓ Install pip

19s

✓ Install nf-core tools

18s

✓ Run nf-core lint

10s

```
1  ► Run nf-core lint ${GITHUB_WORKSPACE}
2
3
4
5
6
7
8
9
10
11
12
13 /opt/hostedtoolcache/Python/3.6.9/x64/lib/python3.6/site-packages/nf_core/lint.py:264: YAMLLoadWarning: calling yaml.load()
without Loader=... is deprecated, as the default Loader is unsafe. Please read https://msg.pyyaml.org/load for full details.
14     self.conda_config = yaml.load(fh)
15 /opt/hostedtoolcache/Python/3.6.9/x64/lib/python3.6/site-packages/nf_core/lint.py:449: YAMLLoadWarning: calling yaml.load()
without Loader=... is deprecated, as the default Loader is unsafe. Please read https://msg.pyyaml.org/load for full details.
16     ciconf = yaml.load(fh)
17
18 INFO: =====
19 LINTING RESULTS
20 =====
21 88 tests passed 10 tests had warnings 0 tests failed
22
23 WARNING: Test Warnings:
24 http://nf-co.re/errors#4: Config variable not found: params.reads
25 http://nf-co.re/errors#4: Config variable not found: params.singleEnd
26 http://nf-co.re/errors#8: Conda package is not latest available: control-freebayes=11.4, 11.5 available
27 http://nf-co.re/errors#8: Conda package is not latest available: ensembl-vep=95.2, 97.3 available
28 http://nf-co.re/errors#8: Conda package is not latest available: freebayes=1.2.0, 1.3.1 available
```

sarek linting
on: pull_request

✓ Markdown

✓ YAML

✓ nf-core

sarek extra CI
on: pull_request

Exercises 4, 5

Exercise 4 (creating pipelines)

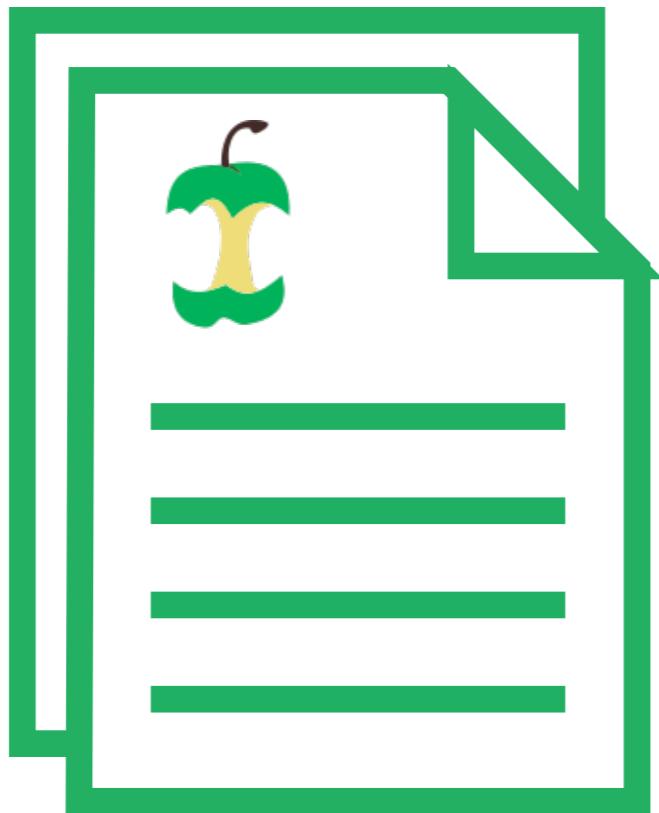
- Make a new pipeline using the template
- Update the readme file to fill in the TODO statements
- Add a new process to the pipeline in `main.nf`
- Add the new software dependencies from this process in to `environment.yaml`

Exercise 5 (testing pipelines)

- Run `nf-core lint` on your pipeline and make note of any test warnings / failures
- Read up on one or two of the linting rules on the nf-core website and see if you can fix some.
- Take a look at `conf/test.config` and switch the test data for another dataset on `nf-core/test_data`

https://nf-co.re/usage/nf_core_tutorial

Releasing pipelines



First release

https://nf-co.re/developers/adding_pipelines



Tell us about it: #new-pipelines



Join the @nf-core organisation



Fork your repo to nf-core



Set up Docker Hub & Zenodo

First release

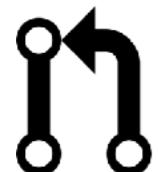
https://nf-co.re/developers/adding_pipelines



Make sure that all tests are passing



Community review: "Pseudo PR"



Open PR from dev to master



Create release on GitHub

Version numbers

Semantic versioning:

1.4.3

Major - results not backwards compatible

Minor - new feature(s)

Patch - bug fix

```
nf-core bump-version ./pipeline 1.4.3
```

```
nf-core bump-version ./pipeline --nextflow 19.07.0
```

Template updates

 nf-core / rnaseq
forked from ewels/nf-core-rnaseq

Unwatch 38 Star 127 Fork 136

Code Issues 32 Pull requests 6 Actions Security Insights Settings

Important pipeline nf-core update! (version 1.6) #194

Merged drpatelh merged 1 commit into `dev` from `TEMPLATE` on 12 Jun

Conversation 7 Commits 1 Checks 0 Files changed 22 +64 -375

 nf-core-bot commented on 9 Apr

Member + ...

Some important changes have been made in the nf-core pipelines templates.
Please make sure to merge this in ASAP and make a new minor release of your pipeline.
Follow the link [nf-core/tools](#)

 Update nf-core pipeline template. X 55d617e None yet

 ewels referenced this pull request on 11 Jun

Minor edit to `installation.md` #225 Merged

Reviewers Assignees Labels Projects Milestone

No reviews No one—assign yourself None yet None yet None yet

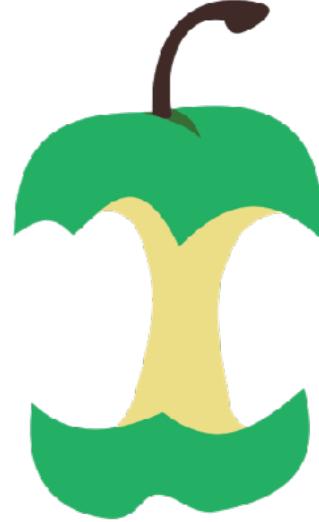
Exercise 6

Exercise 6 (releasing pipelines)

- Use `nf-core bump-version` to update the required version of Nextflow in your pipeline
- Bump your pipeline's version to 1.0, ready for its first release!
- Make sure that you're signed up to the nf-core slack (<https://nf-co.re/join>) and drop us a line about your latest and greatest pipeline plans!
- Ask to be a member of the nf-core GitHub organisation by commenting on this GitHub issue: <https://github.com/nf-core/nf-co.re/issues/3>
- If you're a twitter user, make sure to follow the @nf_core twitter account

https://nf-co.re/usage/nf_core_tutorial

nf-core



Phil Ewels

phil.ewels@scilifelab.se

 tallphil

 ewels

@nf-core/core



Alexander Peltzer

@apeltzer



Sven F.

@sven1103



Harshil Patel

@drpatelh



Maxime Garcia

@MaxUlysse



Johannes Alneberg

@alneberg

<https://nf-co.re>



nf_core