

# Reproducible bioinformatics workflows with Nextflow and nf-core

Phil Ewels

[phil.ewels@scilifelab.se](mailto:phil.ewels@scilifelab.se)



NATIONAL  
GENOMICS  
INFRASTRUCTURE



SciLifeLab

<https://scilifelab.se>  
<https://ngisweden.scilifelab.se>



NATIONAL  
GENOMICS  
INFRASTRUCTURE



The background of the image is a photograph of a vast, misty mountain range with a dense forest of evergreen trees in the foreground. The sky is a clear, pale blue.

# nextflow

<https://www.nextflow.io/>

# nextflow

```
#!/usr/bin/env nextflow
input = Channel.fromFilePairs(params.reads)

process fastqc {
    input:
        file reads from input

    output:
        file "*_fastqc.{zip,html}" into results

    script:
    """
        fastqc -q $reads
    """

}
```



SGE



Google Cloud



LSF

PBS

<https://www.nextflow.io/>



A community effort to collect a  
curated set of analysis pipelines  
built using Nextflow.

<https://nf-co.re>



Guidelines



Tools



Pipelines

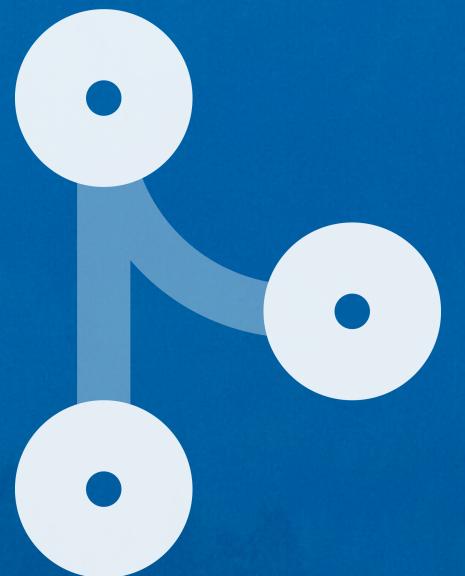
# nf-core

The logo consists of the word "nf-core" in a bold, sans-serif font. The "n" is green, while "f", "core", and the "x" in "core" are white. A green apple icon with a yellow core and a brown stem is positioned to the right of the text.

Develop with  
the community



Start from the  
template



Collaborate,  
don't duplicate

# nf-core



16  
UNDER DEVELOPMENT



5  
ARCHIVED

# Available Pipelines

Can you think of another pipeline that would fit in well? [Let us know!](#)

Search keywords

Filter:

Released 25

Under development 14

Archived 3

Sort:

Last Release

Alphabetical

Stars

Display:



## nf-core/viralrecon ✓

☆ 18

amplicon assembly covid-19 covid19 illumina metagenomics

sars-cov-2 variant-calling viral virus

Assembly and intrahost/low-frequency variant calling for viral samples

Version 1.1.0

Published 7 days ago

## nf-core/imcyto ✓

☆ 5

cytometry image-analysis image-processing image-segmentation

Image Mass Cytometry analysis pipeline

Version 1.0.0

Published 1 month ago

## nf-core/coproid ✓

☆ 2

adna ancient-dna coprolite microbiome

Sample-level identification pipeline

## nf-core/sarek ✓

☆ 49

cancer germline somatic variant-calling

Analysis pipeline to detect germline or somatic variants from WGS / targeted sequencing

Version 2.6.1

Published 1 week ago

## nf-core/slamseq ✓

☆ 3

differential-expression quantseq slamseq transcriptomics

SLAMSeq processing and analysis pipeline

Version 1.0.0

Published 2 months ago

## nf-core/mhcquant ✓

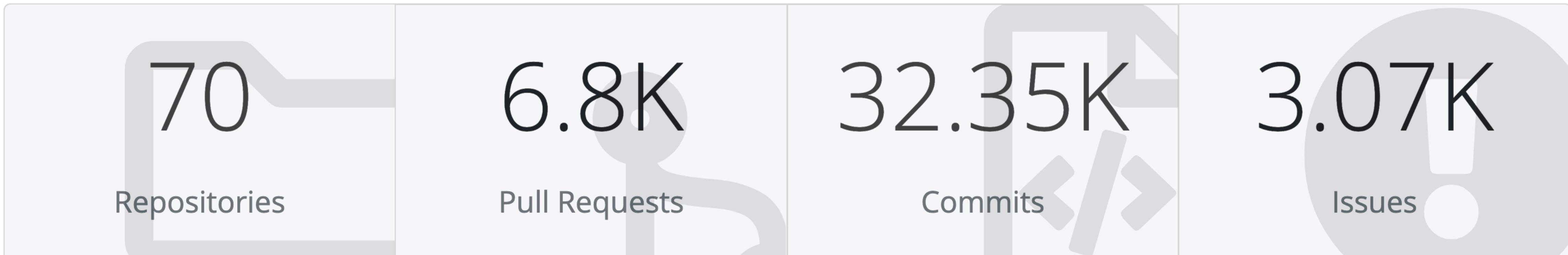
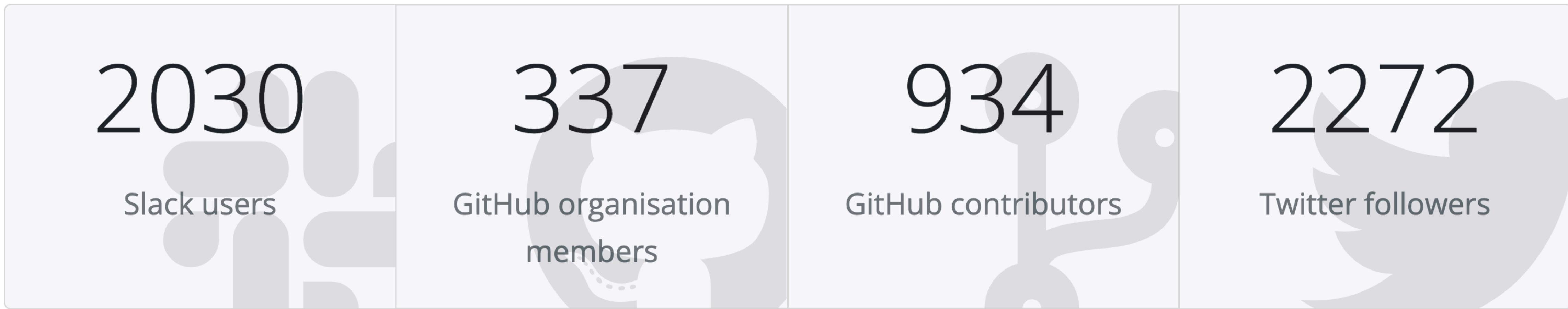
☆ 12

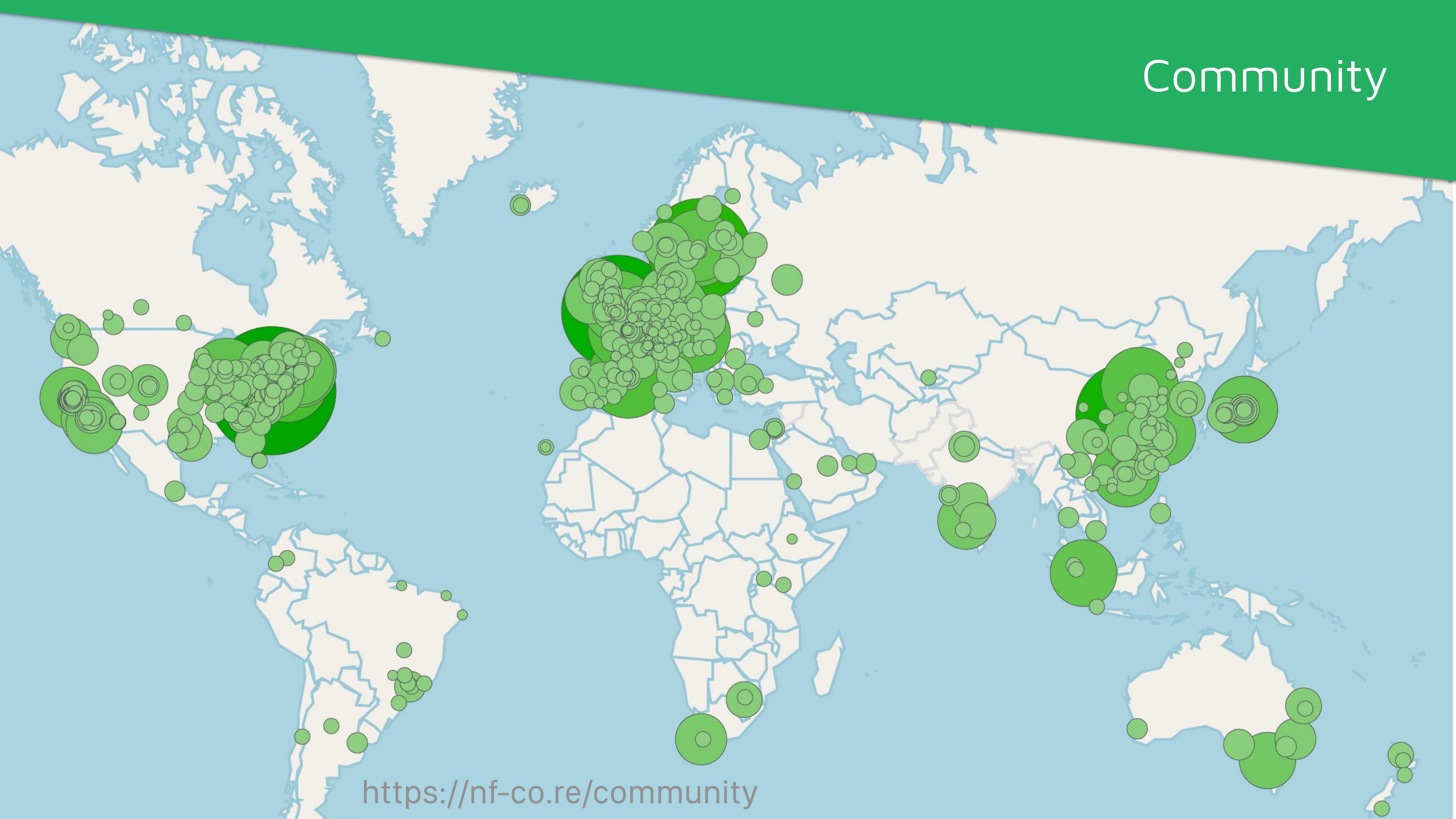
mass-spectrometry mhc peptides

Identify and quantify MHC peptides from mass-spectrometry data

# Community

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





# Community

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

## Community



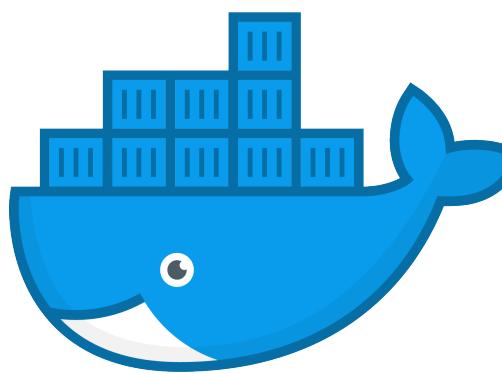
<https://nf-co.re/join>

# Requirements

1

# nextflow

2



docker®

or



or

# CONDA

3

# nf-core/tools



Requirements

(Optional)

# nf-core/tools



BIOCONDA®

pip install nf-core

conda install nf-core

# Running a pipeline

Let Nextflow fetch the pipeline for you!

```
nextflow run nf-core/<pipeline> -r <version>
```

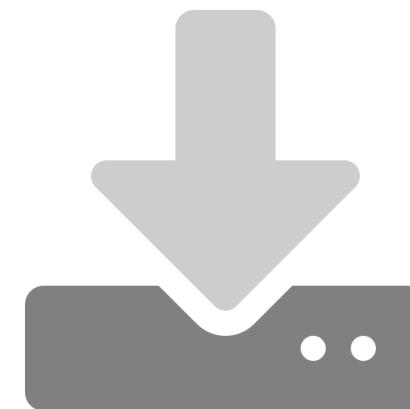
GitHub pipeline name  
Clones to `~/.nextflow`

Pipeline release tag  
Repository branch

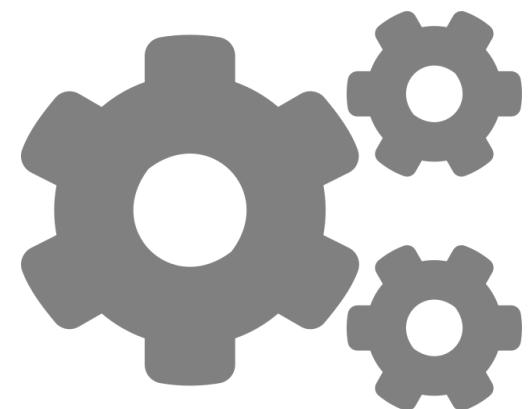
# Offline usage

Offline? Fetch using nf-core helper tool

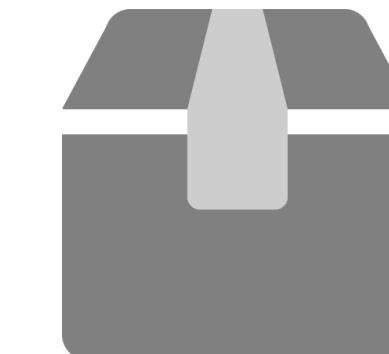
```
nf-core download nf-core/<pipeline> -r <version>
```



Downloads pipeline



Downloads configs



Fetches Singularity containers

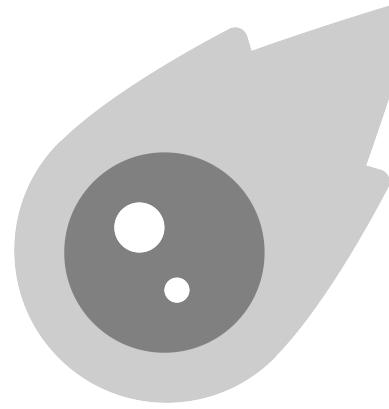


Creates archive

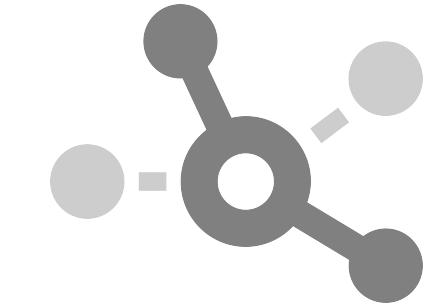
## Central location of all nf-core pipelines

```
module load nf-core-pipelines
```

```
nextflow run $NF_CORE_PIPELINES/<pipeline>/<version>/workflow
```



All pipelines, all versions  
available out of the box



Uses central store  
of Singularity containers



Available on Bianca 😊



Very new!  
Please report problems

Launch

# nf-core/atacseq

ATAC-seq peak-calling, QC and differential analysis pipeline

atac-seq chromatin-accessibility

 Launch version 1.2.0

 <https://github.com/nf-core/atacseq>

Readme

Documentation

Statistics

Releases

## Introduction

**nfcore/atacseq** is a bioinformatics analysis pipeline used for ATAC-seq data.

The pipeline is built using [Nextflow](#), a workflow tool to run tasks across multiple compute infrastructures in a very portable manner. It comes with docker containers making installation trivial and results highly reproducible.

## Pipeline summary

>\_ command

```
» nextflow run nf-core/atacseq -r 1.2.0 -profile tes
```

↓ clones in last 12 months

1633



stars

37

watchers

50

last release

last updated

Launch

# Launch pipeline

Configure workflow parameters for a pipeline run.

Pipeline: `atacseq ( dev )`

Launch ID: `1593526475_51e7bd935c0f` (expires 07/14/2020)

Go through the pipeline inputs below, setting them to the values that you would like. When you're done, click *Launch* and your parameters will be saved. The `nf-core launch` command running in the background should detect your changes and give you further instructions.

 Jump to section ▾

 Show hidden params

Nextflow command-line flags

 Launch

## Nextflow command-line flags

```
nf-core launch nf-core/<pipeline>
```



Works with any Nextflow pipeline (local or remote)  
Online web-form or offline command-line wizard



Search...

### nf-core/chipseq

big\_euler

### nf-core/rnaseq

thirsty\_minsky  
Duration: -  
Started: 2020-10-06 22:27:09

### nf-core/rnaseq

boring\_laplace  
Duration: 16 m 4 s  
Started: 2020-07-01 00:15:38

### nf-core/rnaseq

spontaneous\_pesquet  
Duration: 14 m 23 s  
Started: 2020-06-18 11:44:27

### nf-core/rnaseq

dreamy\_fermat  
Duration: 5 m 49 s  
Started: 2019-12-09 14:45:41

### nf-core/rnaseq

desperate\_rosalind  
Duration: 3 m 53 s  
Started: 2019-09-21 08:58:40

### nextflow-io/hello

elated\_morse  
Duration: -  
Started: 2019-06-22 08:55:57

### nextflow-io/hello

stoic\_leavitt  
Duration: 2 s  
Started: 2019-09-19 17:41:45



# nextflow tower

Command line Parameters Configuration Execution log

Download log files

```
1 NEXTFLOW ~ version 20.07.1
2 Pulling nf-core/chipseq ...
3 downloaded from https://github.com/nf-core/chipseq.git
4 Launching `nf-core/chipseq` [big_euler] - revision: 0f487ed76d [master]
5 -----
6
7
8   N E X T F L O W   ~
9   \ \ / / \ \ / /   } {
10    \| / \| / \|   \_/
11      nf-core/chipseq v1.2.1
12 -----
13 Run Name      : big_euler
14 Data Type     : Paired-End
15 Design File   : https://raw.githubusercontent.com/nf-core/test-datasets/chipseq/design.csv
16 Genome        : Not supplied
17 Fasta File    : https://raw.githubusercontent.com/nf-core/test-datasets/atacseq/reference/genome.fa
18 GTF File      : https://raw.githubusercontent.com/nf-core/test-datasets/atacseq/reference/genes.gtf
19 MACS2 Genome Size : 1.2E+7
20 Min Consensus Reps : 1
21 MACS2 Narrow Peaks : No
22 MACS2 Broad Cutoff : 0.1
```

General

Status

id F0z5aPZLjfLyJ

big\_euler

2020-10-06 22:27:09

Started: 2019-06-22 08:55:57

0f487ed76dc9779fb456270d1305202a5 (master)

b19410b8-52f3-471e-9a34-f4a4935db345

phil-ewells

s3://nf-core-awsmegatests/scratch/phil-test

https://tower.nf



NATIONAL  
GENOMICS  
INFRASTRUCTURE



nextflow tower



nextflow  
nf-core   
aws 

nextflow  
nf-core   
  
slurm  
workload manager

# Phil Ewels

<https://phil.ewels.co.uk>

 phil.ewels@scilifelab.se

 @ewels

 @tallphil



<https://nf-co.re>

 <https://www.nextflow.io>

 **nextflow** tower <https://tower.nf>

<https://ngisweden.scilifelab.se>

<https://scilifelab.se>



Ewels, P.A., Peltzer, A., *et al.*  
*Nat Biotechnol* **38**, 276–278 (2020).  
<https://doi.org/10.1038/s41587-020-0439-x>



<https://multiqc.info>

