

Introduction



Maxime U Garcia
SciLifeLab | National Genomics Infrastructure
nf-core

ANF - workflow et reproductibilité

2025-11-25..27



Maxime

- Community
- Memes
- Attention to details
- Chaotic
- Design
- Talk a lot
- Resourceful
- Resilient
- Recycle slides
- Multi-tasker
- No python
- Everywhere



The Challenge of Reproducibility

Hidden reproducibility issues are like an iceberg



First, we tried to re-run the analysis with the code and data provided by the authors.

Second, we reimplemented the whole method in a Python package...



Experimenting with reproducibility:
a case study of robustness in bioinformatics
<https://doi.org/10.1093/gigascience/giy077>

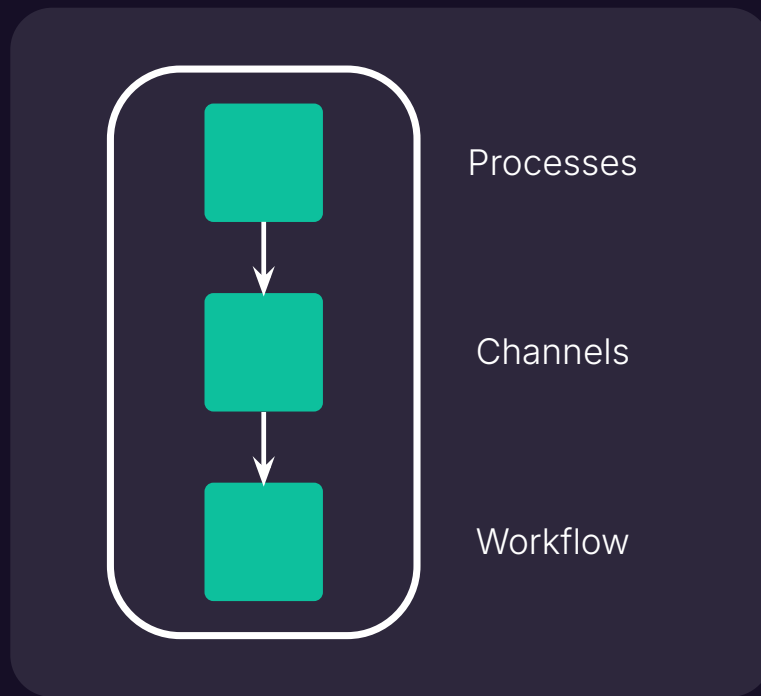
The five pillars of computational reproducibility:
bioinformatics and beyond
<https://doi.org/10.1093/bib/bbad375>





Introducing Nextflow

A reactive workflow framework and a programming DSL



Introducing Nextflow

Nextflow is a language, a runtime, and a community

 **nextflow**
pipeline



 **nextflow**
runtime

Write code
in any language



Orchestrate tasks with
dataflow programming



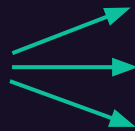
Define software
dependencies via containers



Built-in version control
with Git



Task orchestration
and execution



Introducing Nextflow

Nextflow is a language, a runtime, and a community



Reproducible

Integration with code management tools, with versioned releases.



Portable

Conda, Docker, Singularity/Apptainer works with most compute environments.



Scalable

5 samples on your laptop, 5k on an HPC or 5 million in the cloud.

nf-core



nf-core history

how nf-core came to be - told by Maxime



We had NGI pipelines for our internal usage.

People wanted to use them and contribute instead of re-inventing the wheel.

We joined effort and dropped the NGI branding and founded nf-core.

Key concept:

- Don't work alone, work TOGETHER



nf-core history

From modest starts towards more global usage

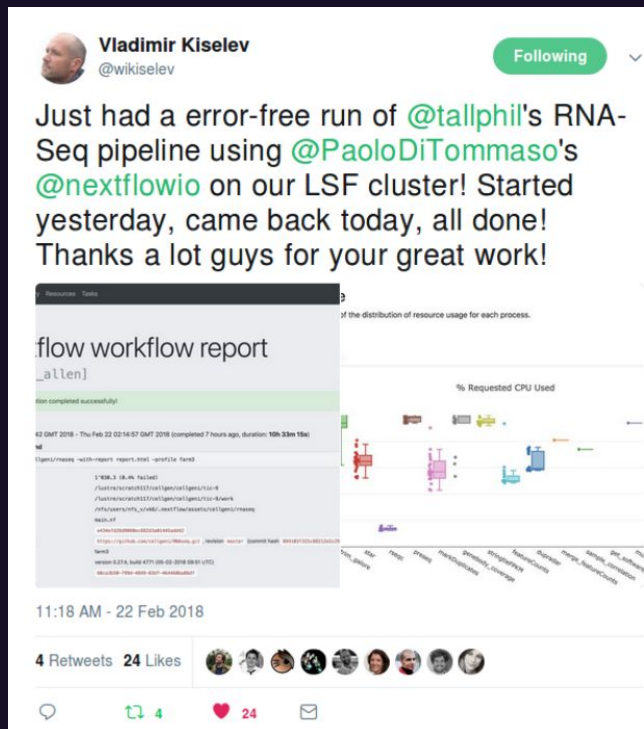
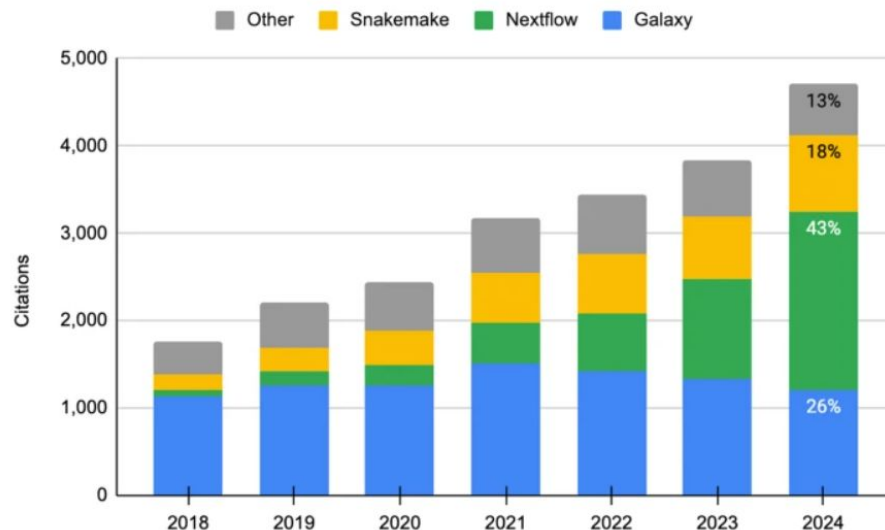


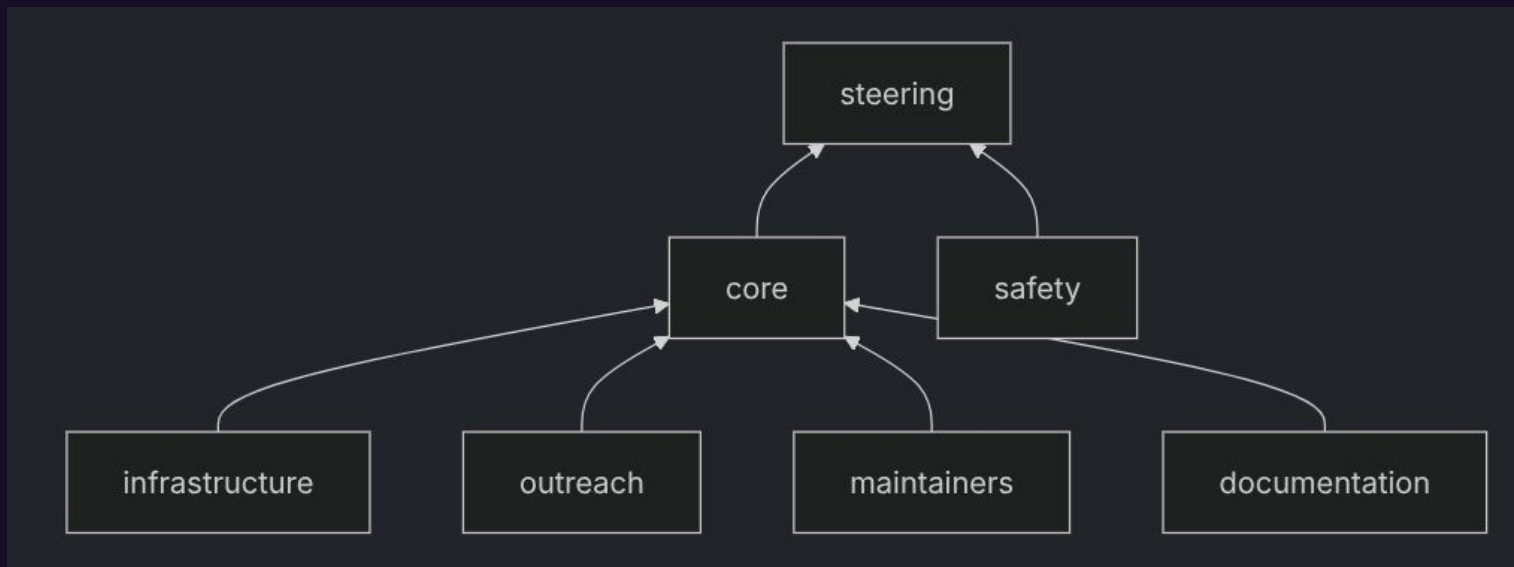
Fig. 1



Google Scholar citation counts for bioinformatics workflow management systems. Sum of citations of the major publications of Galaxy, Nextflow, and Snakemake between 2018 and 2024 (Data in Additional File 1: Supp. Table 1)



nf-core governance



nf-core community

A global community collaborating to build open-source Nextflow components and pipelines

12k+

Slack
users

1.5k+

GitHub
contributors

160+

GitHub
repositories

40k

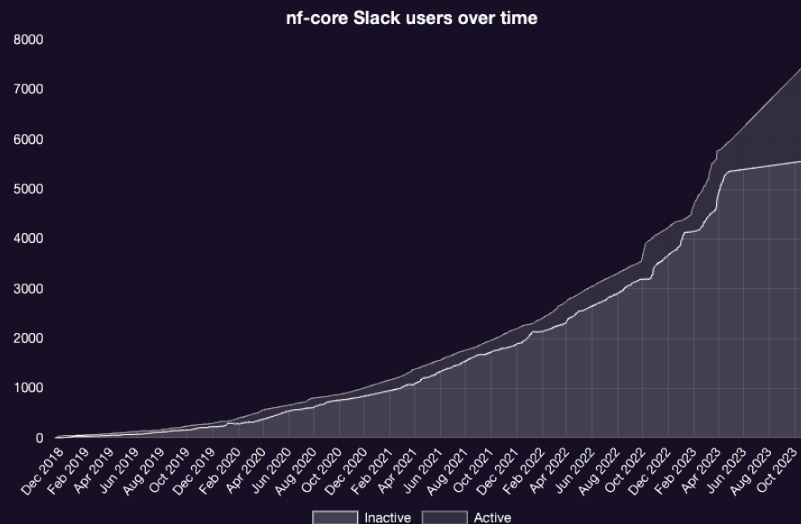
GitHub
commits

16k+

Pull
requests

7k+

GitHub
issues



nf-core principles

Principles that guide the nf-core community



Ecosystem

Nextflow based
pipelines
Dedicated tooling



Cooperation

Develop with the
communities



Stable tags

Release with extensive
CI/CD
Cloud ready



Open source

MIT license



Standards

Documentation
Use a common template
Use public containers
and/or virtual
environments



Best practices

1 pipeline per data type
and/or usage
No duplicates



nf-core components

Pick and choose which component you need



Pipelines

>129 pipelines



Modules

> 1672 modules
> 95 subworkflows



Test data

Minimal test data for CI
for all modules
+ tests with nf-test 🚀



Linting

Choose conventions to
test for consistency



Schema

Validation, channels and
user interface



Tooling

Development and
deployment
Pipeline template



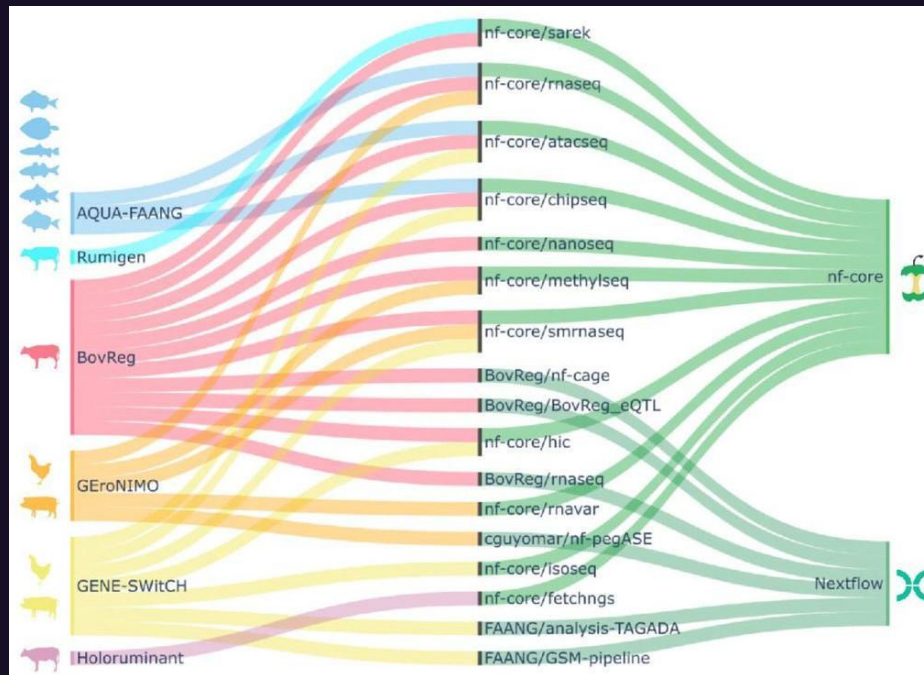
nf-core pipelines

Flagships

rnaseq ✓ ☆ 1143 RNA sequencing analysis pipeline using STAR, RSEM, HISAT2 or Salmon with gene/isoform counts and extensive quality control. rna rna-seq 🔍 3.21.0 released 2 months ago	sarek ✓ ☆ 502 Analysis pipeline to detect germline or somatic variants (pre-processing, variant calling and annotation) from WGS / targeted sequencing annotation cancer gatk4 genomics germline pre-processing somatic target-panels variant-calling whole-exome-sequencing whole-genome-sequencing 🔍 3.6.1 released 28 days ago	scrnaseq ✓ ☆ 301 Single-cell RNA-Seq pipeline for barcode-based protocols such as 10x, DropSeq or SmartSeq, offering a variety of aligners and empty-droplet detection 10x-genomics 10xgenomics alevin bustools cellranger kallisto rna-seq single-cell star-solo 🔍 4.1.0 released 22 days ago	mag ✓ ☆ 252 Assembly and binning of metagenomes annotation assembly binning long-read-sequencing metagenomes metagenomics nanopore nanopore-sequencing 🔍 5.2.0 released 14 days ago
chipseq ✓ ☆ 226 ChIP-seq peak-calling, QC and differential analysis pipeline. chip chip-seq chromatin-immunoprecipitation macs2 peak-calling 🔍 2.1.0 released about 1 year ago	ampliseq ✓ ☆ 225 Amplicon sequencing analysis workflow using DADA2 and QIIME2 16s 18s amplicon-sequencing edna illumina iontorrent its metabarcoding metagenomics metataxonomics microbiome pacbio qiime2 rna taxonomic-classification taxonomic-profiling 🔍 2.15.0 released about 2 months ago	atacseq ✓ ☆ 216 ATAC-seq peak-calling and QC analysis pipeline atac-seq chromatin-accessibility 🔍 2.1.2 released over 2 years ago	nanoseq ✓ ☆ 214 Nanopore demultiplexing, QC and alignment pipeline alignment demultiplexing nanopore qc 🔍 3.1.0 released over 2 years ago
eager ✓ ☆ 186 A fully reproducible and state-of-the-art ancient DNA analysis pipeline adna ancient-dna-analysis ancientdna genome metagenomics pathogen-genomics population-genetics 🔍 2.5.3 released 8 months ago	fetchngs ✓ ☆ 182 Pipeline to fetch metadata and raw FastQ files from public databases dbj download ena fastq geo sra synapse 🔍 1.12.0 released over 1 year ago	methyseq ✓ ☆ 181 Methylation (Bisulfite-Sequencing) analysis pipeline using Bismark or bwa-meth + MethylDackel bisulfite-sequencing dna-methylation em-seq epigenome epigenomics methyl-seq pbat rtbs 🔍 4.1.0 released 3 months ago	taxprofiler ✓ ☆ 170 Highly parallelised multi-taxonomic profiling of shotgun short- and long-read metagenomic data classification illumina long-reads metagenomics microbiome nanopore pathogen profiling shotgun taxonomic-classification taxonomic-profiling 🔍 1.2.4 released 3 months ago



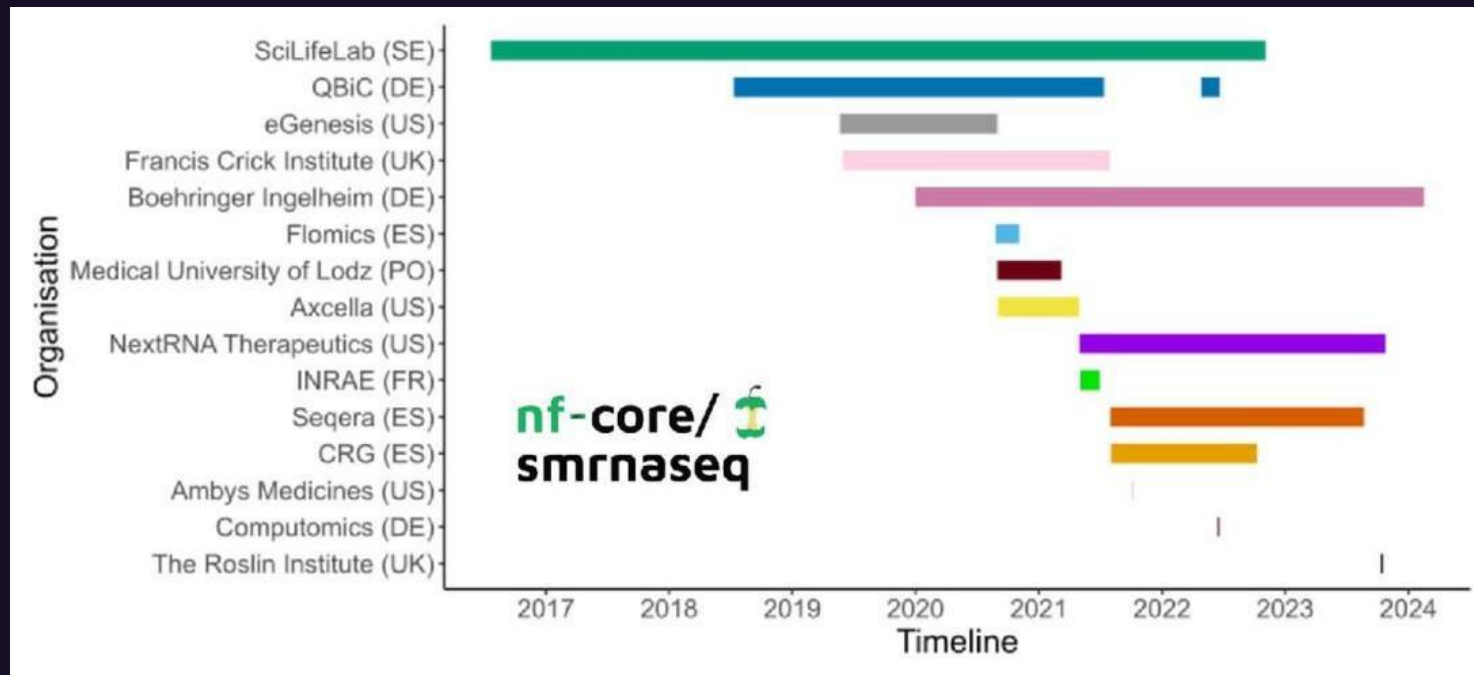
Working together with communities



Empowering bioinformatics communities with Nextflow and nf-core - <https://doi.org/10.1101/2024.05.10.592912>



Continuance in collaboration



Empowering bioinformatics communities with Nextflow and nf-core - <https://doi.org/10.1101/2024.05.10.592912>



Contribute and participate

Seminars, training, hackathons, and more

Join us on: <https://nf-co.re/join>

<https://community.seqera.io/>

- Bytesize seminars
- Training sessions
- Hackathons
- Blogs
- Community Forum and Slack
- Social media
- Documentation
- Mentorships
- Office hours (APAC - AMER||EMEA)
- Special Interest Groups







PLATFORM
STUDIO
COMPUTE
AI
COMMUNITY
EVENTS
TRAININGS
PODCASTS