

BovReg contribution to the adoption of nf-core as the EuroFAANG computational standard

Jose Espinosa-Carrasco (CRG)



BovReg Final Conference - Brussels

14th February 2024

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815668

Disclaimer: the sole responsibility of this presentation lies with the authors. The Research Executive Agency is not responsible for any use that may be made of the information contained therein.

Sustainable pipelines



How BovReg pipelines could become a resource for the community?

Article | [Open access](#) | [Published: 14 October 2022](#)

Introducing the FAIR Principles for research software

[Michelle Barker](#) , [Neil P. Chue Hong](#), [Daniel S. Katz](#), [Anna-Lena Lamprecht](#), [Carlos Martinez-Ortiz](#), [Fotis Psomopoulos](#), [Jennifer Harrow](#), [Leyla Jael Castro](#), [Morane Gruenpeter](#), [Paula Andrea Martinez](#) & [Tom Honeyman](#)

[Scientific Data](#) **9**, Article number: 622 (2022) | [Cite this article](#)

19k Accesses | **57** Citations | **232** Altmetric | [Metrics](#)

An aim shared by EuroFAANG

FAANG Shared Workshop: Foundation for the Future Agenda (Hinxton, February 2020)



 **EuroFAANG pipelines working group**



nextflow enables FAIR coding

Polyglot



Dependencies
sandboxed in containers



Uses dataflow
programming paradigm



Version control



Support of
multiple platforms



Other workflow manager systems could have been consider

Perspective | [Published: 23 September 2021](#)

Reproducible, scalable, and shareable analysis pipelines with bioinformatics workflow managers

[Laura Wratten](#), [Andreas Wilm](#) & [Jonathan Göke](#) 

[Nature Methods](#) **18**, 1161–1168 (2021) | [Cite this article](#)

Table 1 | Overview of workflow managers for bioinformatics (top, editable version; bottom, image version)

Tool	Class	Ease of use ^a	Expressiveness ^b	Portability ^c	Scalability ^d	Learning resources ^e	Pipeline initiatives ^f
Galaxy	Graphical	●●●	●○○	●●●	●●●	●●●	●●○
KNIME	Graphical	●●●	●○○	○○○	●●○	●●●	●●○
Nextflow	DSL	●●○	●●●	●●●	●●●	●●●	●●●
Snakemake	DSL	●●○	●●●	●●○	●●●	●●○	●●●
GenPipes	DSL	●●○	●●●	●●○	●●○	●●○	●●○
bPipe	DSL	●●○	●●●	●●○	●●○	●●○	●○○
Pachyderm	DSL	●●○	●●●	●○○	●●○	●●●	○○○
SciPipe	Library	●●○	●●●	○○○	○○○	●●○	○○○
Luigi	Library	●●○	●●●	●○○	●●○	●●○	○○○
Cromwell + WDL	Execution + workflow specification	●○○	●●○	●●●	●●○	●●○	●●○
cwltool + CWL	Execution + workflow specification	●○○	●●○	●●○	○○○	●●●	●●○
Toil + CWL/WDL/Python	Execution + workflow specification	●○○	●●●	●●○	●●●	●●○	●●○

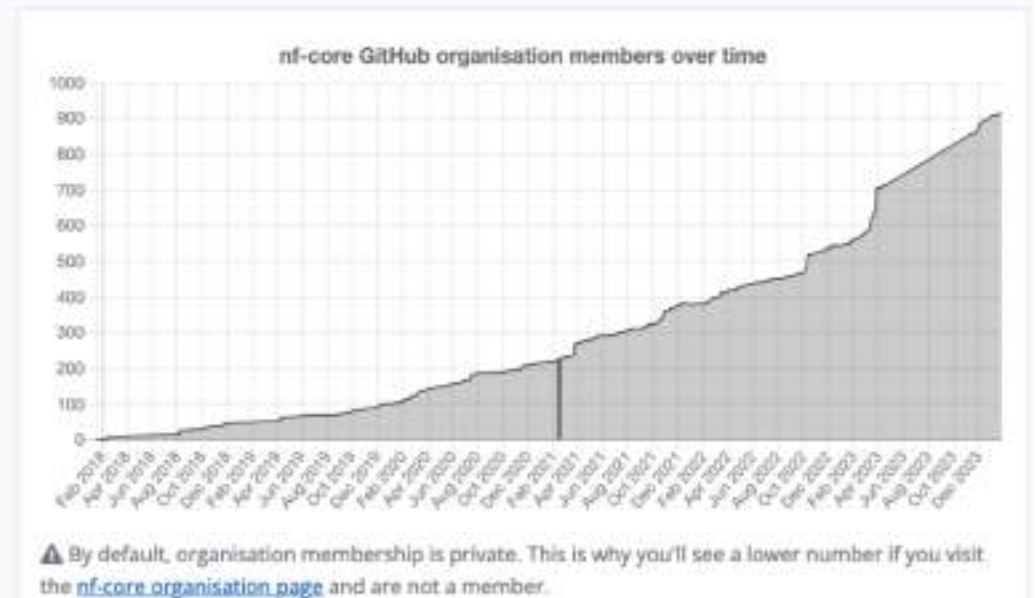
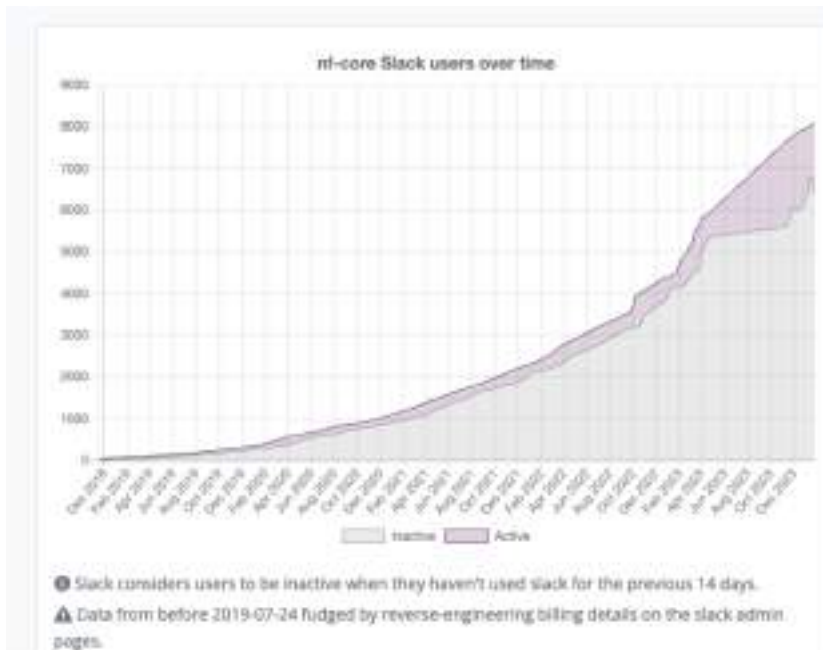
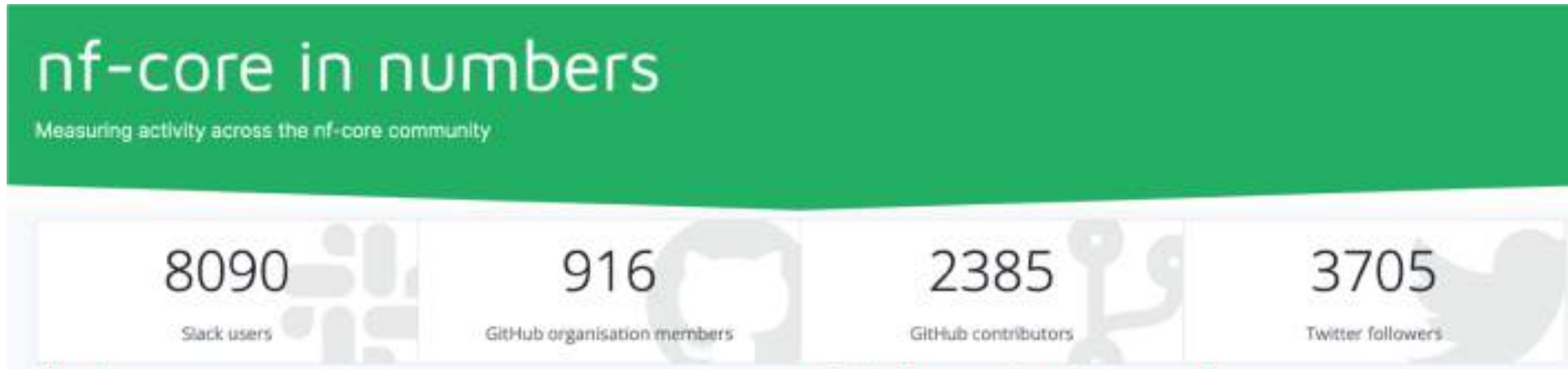
nf-core makes Nextflow a special workflow manager system



- A collection of Nextflow pipelines
- Coding guidelines (standard)
- Helper tools to use and develop pipelines
- A community



The community drives the evolution of the language and the standard



Nextflow modularization aligned perfectly with BovReg timing

- DSL2 was officially released on July 2020
- Modularization enables pipeline components to be shared and combined across projects
- nf-core pipelines started a process of reimplementing to follow DSL2
- Currently 1000 modules and 50 subworkflows available under: <https://nf-co.re/modules>



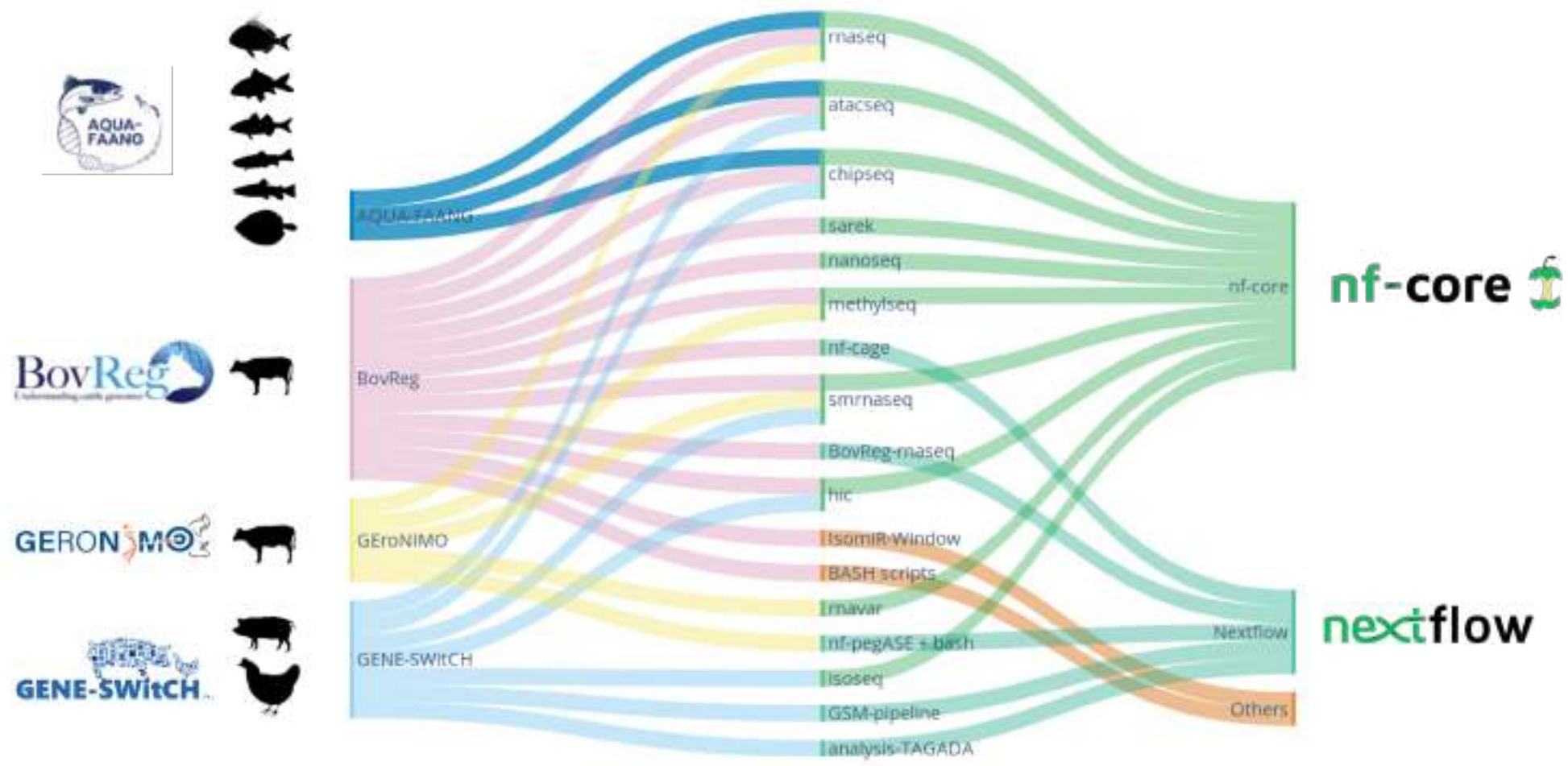
Examples of BovReg contributions to pipelines

**nf-core/
atacseq**

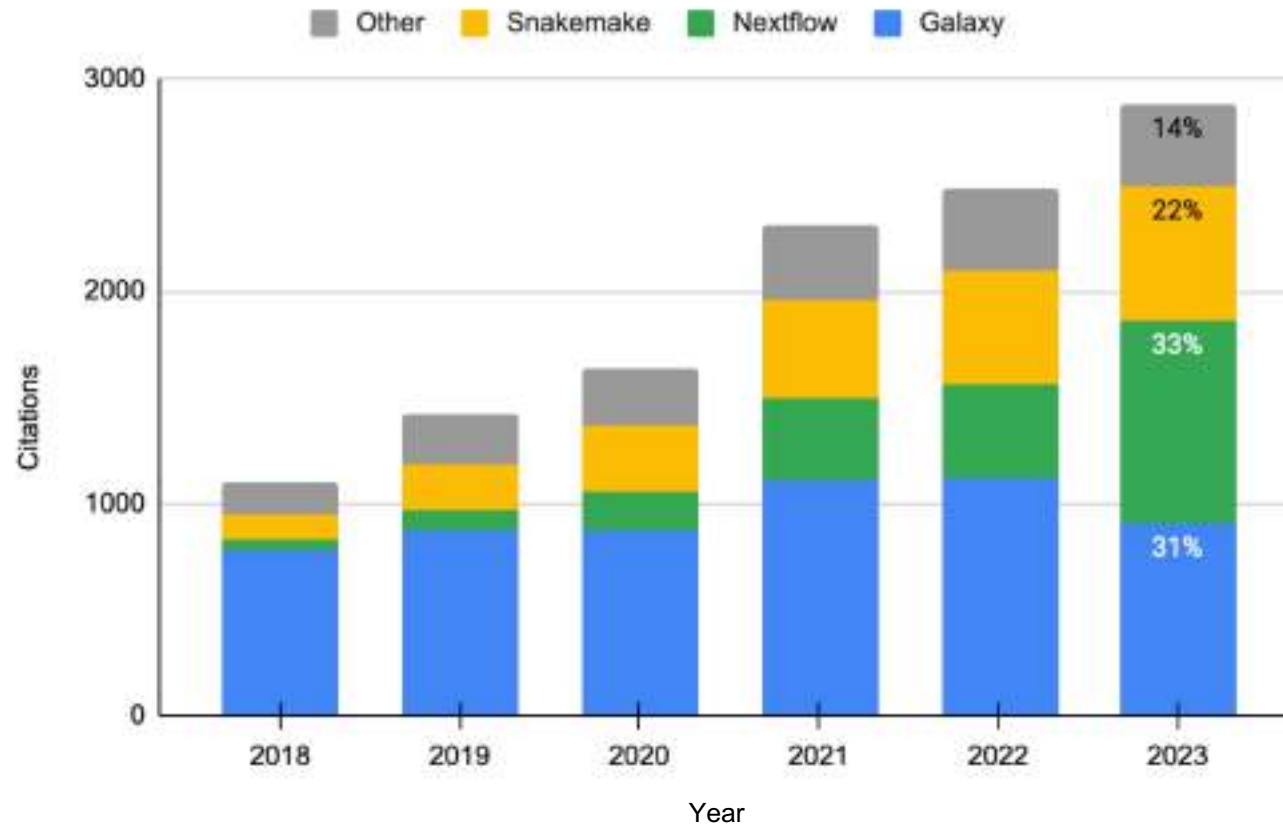


- Release v2.0.0 – Iron Iguana Major reimplementaion of the pipeline to follow DSL2
- Release v2.1.0 – Iron Cow Optional support of control data
- CRG became the main maintainer of the pipeline

Widespread Nextflow/nf-core adoption in EuroFAANG



Was a correct decision?



What about reproducibility?

Automatic Reproduction of Workflows in the Snakemake Workflow Catalog and nf-core Registries

Samuel Grayson

grayson5@illinois.edu

Department of Computer Science
University of Illinois Urbana-Champaign
Urbana, Illinois, USA

Daniel S. Katz

d.katz@ieee.org

NCSA & CS & ECE & iSchool
University of Illinois Urbana-Champaign
Urbana, Illinois, USA

Durko Marinov

marinov@illinois.edu

Department of Computer Science
University of Illinois Urbana-Champaign
Urbana, Illinois, USA

Reed Milewicz

rmilewi@sandia.gov

Department of Software Engineering and Research
Sandia National Laboratories
Albuquerque, New Mexico, US



Quantity	All	SWC	nf-core
# workflows	101	53	48
# revisions	584	333	251
% of revisions with no crash	28%	11%	51%
% of workflows with at least one non-crashing revision	53%	23%	88%

Table 2: Summary of data from automatic reproduction.

A paper is in the works

Empowering genomics communities with nf-core: The EuroFAANG success story



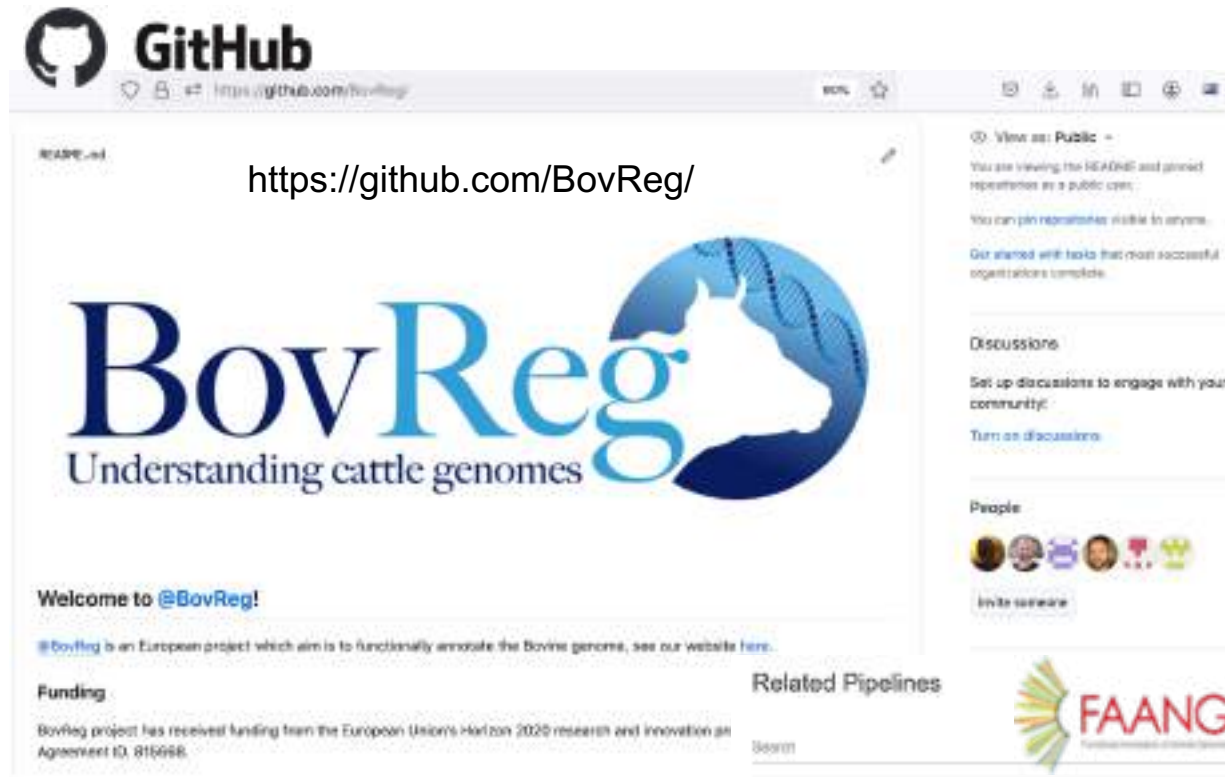
nf-core Special Interest Groups (SIG)

nf-core blog

Special interest groups

An exciting development is that nf-core wants to establish "*Special interest groups*" to encourage people within the nf-core community with similar interest to meet and work together. We spent a disproportionate amount of time arguing about the name for these: SIGs / COSIs / communities / working groups etc etc - naming things is hard! More information on this new initiative will be presented in the [bytesize talk on February 20th](#) and in a dedicated blog post / website section soon.

Where to find BovReg pipelines?



The image shows the GitHub repository page for BovReg. The URL is <https://github.com/BovReg/>. The repository name is `HEAD~1`. The main content area features the BovReg logo, which includes a blue silhouette of a cow's head inside a globe, with the text "BovReg" in large blue letters and "Understanding cattle genomes" below it. To the right of the logo, there is a "View as: Public" dropdown menu and a "Discussions" section. Below the logo, there is a "Welcome to @BovReg!" message and a "Funding" section. The "Related Pipelines" section is partially visible at the bottom right of the image.

<https://github.com/BovReg/>

Welcome to [@BovReg!](#)

[@BovReg](#) is an European project which aim is to functionally annotate the Bovine genome, see our website [here](#).

Funding

BovReg project has received funding from the European Union's Horizon 2020 research and innovation grant Agreement ID: 816668.

Related Pipelines



<https://data.faang.org/projects/BovReg>

Popular repositories

Sort

Pipeline name	Assay type	Pipeline link	Pipeline documentation	Filters
https://github.com/massq/bovReg	RNA-seq	https://github.com/BovReg/1-core-massq-bovReg	https://github.com/BovReg/1-core-massq-bovReg#readme	1f-core
https://github.com/massq/bovReg	RNA-seq	https://github.com/BovReg/1-core-massq	https://github.com/BovReg/1-core-massq#readme	1f-core
https://github.com/massq/bovReg	smRNA-seq	https://github.com/BovReg/1-core-massq	https://github.com/BovReg/1-core-massq#readme	1f-core
https://github.com/BovReg/1-core	CAGE-seq	https://github.com/BovReg/1-core	https://github.com/BovReg/1-core#readme	None
https://github.com/BovReg/1-core	ChIP-seq	https://github.com/BovReg/1-core	https://github.com/BovReg/1-core#readme	1f-core
https://github.com/BovReg/1-core	ATAC-seq	https://github.com/BovReg/1-core	https://github.com/BovReg/1-core#readme	1f-core
https://github.com/BovReg/1-core	Hi-C-seq	https://github.com/BovReg/1-core	https://github.com/BovReg/1-core#readme	1f-core
https://github.com/BovReg/1-core	Variant Calling	https://github.com/BovReg/1-core	https://github.com/BovReg/1-core#readme	1f-core
https://github.com/BovReg/1-core	WTL analysis	https://github.com/BovReg/1-core	https://github.com/BovReg/1-core#readme	None

Acknowledgements

Center for Genomic Regulation



CBCRG lab

Björn Langer
Ionas Erb
Cedric Notredame



members



FMV-ULisboa

Andreia Amaral
Luis da Gama



LUKE

Daniel Fischer
(HoloRuminant)

nf-core members



Friederike Hanssen
Harshil Patel
Gisela Gabernet
Sven Nahnsen
Phil Ewels
Core team



EuroFAANG pipelines working group



AQUA-FAANG

Daniel J. Macqueen
Delphine Lallias
Frederique Pitel
Gareth Gillard
Marie-Odile Baudement
Manu Kumar Gundappa



GENE-SWITCH

Cervin Guyomar
Jani de Vos
Sarah Djebali
Sylvain Foissac



GERONIMO

Klavdija Poklucar
Sandrine Lagarrigue
Sonia Eynard



INRAE

Mathieu Charles
Dominique Rocha



UEDIN

Mazdak Salavati
Emily Clark



GIGA

Gabriel Costa (Humigen)
Carole Charlier



FBN

Praveen Chitneedi
Christa Kühn



EMBL-EBI

Alexey Sokolov
Peter Harrison

Seqera members



Evan Floden
Paolo Di Tommaso

Communities:



BovReg *PARTNERS*



Thank you for your attention

www.bovregproject.eu



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 815668

Disclaimer: the sole responsibility of this presentation lies with the authors. The Research Executive Agency is not responsible for any use that may be made of the information contained therein.





The EuroFAANG pipelines working group

Who: EuroFAANG projects (BovReg, Gene-SWitCH, AQUA-FAANG, GEroNIMO, Rumigen, HoloRuminant) + nf-core core members

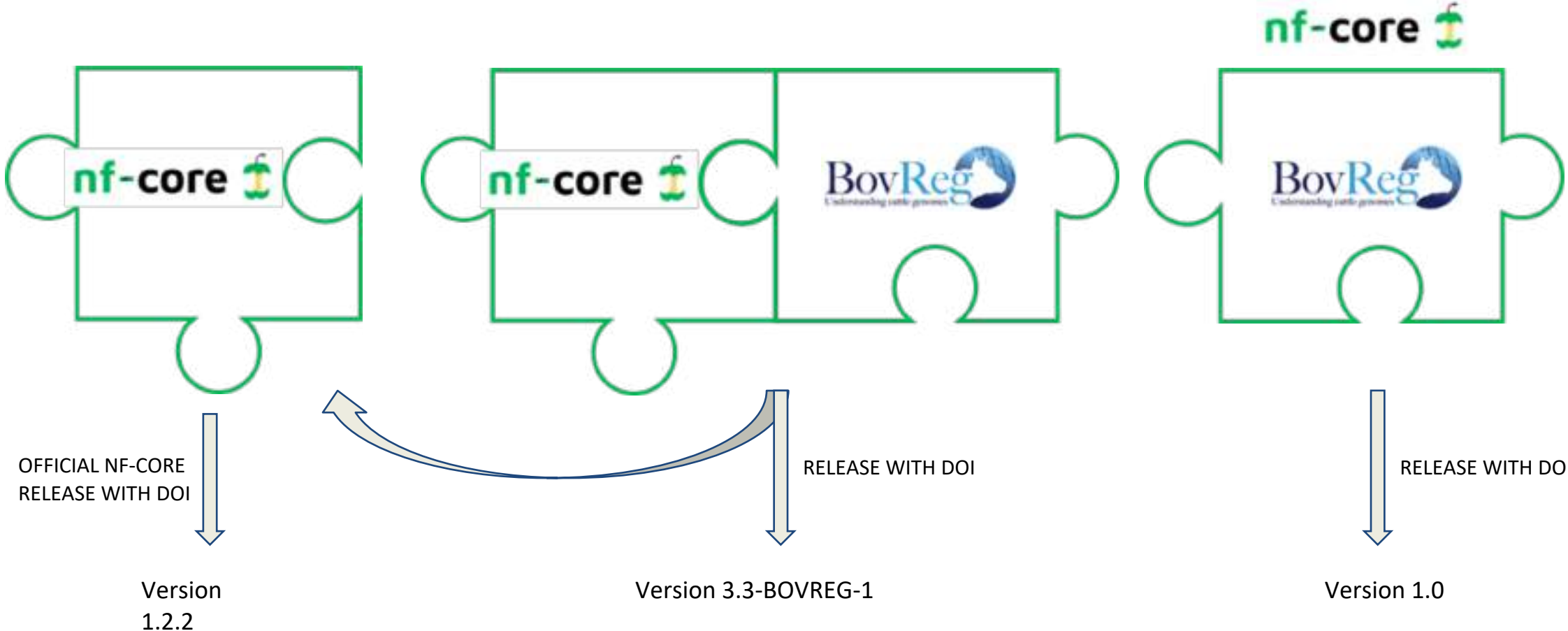
What: Collect existing resources, list good practices, identify gaps and overlaps.

When: Monthly meetings

Lead: EMBL-EBI and CRG

BovReg partners involved: Björn Langer, Jose Espinosa-Carrasco, Cedric Notredame (CRG), Alexey Solokov, Peter Harrison (EMBL-EBI), Praveen Chitneedi, Christa Kühn (FBN), Andreia Amaral (FMV), Mathieu Charles, Dominique Rocha (INRAe), Daniel Fischer (Luke), Mazdak Salavati, Emily Clark (UEDIN), Gabriel Costa (Uliege)

How does it work in practical terms



<https://github.com/BovReg/>

