



HAL
open science

Biocluster: an NGS-dedicated HPC cluster in IMBBC, HCMR Upcoming Upgrade Usage/Partners Training

Jacques Lagnel, Tereza Manousaki, Georgios Kotoulas, Antonios Magoulas

► **To cite this version:**

Jacques Lagnel, Tereza Manousaki, Georgios Kotoulas, Antonios Magoulas. Biocluster: an NGS-dedicated HPC cluster in IMBBC, HCMR Upcoming Upgrade Usage/Partners Training. 9. Hellenic Bioinformatics 2016, Nov 2016, Thessalonica, Greece. 2016. hal-02799037

HAL Id: hal-02799037

<https://hal.inrae.fr/hal-02799037>

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.


See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/319932900>

Biocluster: an NGS-dedicated HPC cluster in IMBBC, HCMR Upcoming Upgrade Usage/Partners Training


Conference Paper · November 2016

CITATIONS
0

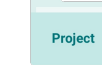
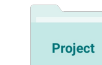
4 authors, including:

 **Jacques Lagne**
INRA French National Institute for Agricultural Research
49 PUBLICATIONS 1,574 CITATIONS
[SEE PROFILE](#)

READS
20

 **Tereza Manousaki**
30 PUBLICATIONS 743 CITATIONS
[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:

-  Regional cooperation for the transnational ecosystem sustainable development (RECONNECT) [View project](#)
-  RobustBass - Advanced selective breeding for robustness, disease and stress resistance in European sea bass (*Dicentrarchus labrax*) through the use of NGS techniques for genetic improvement [View project](#)

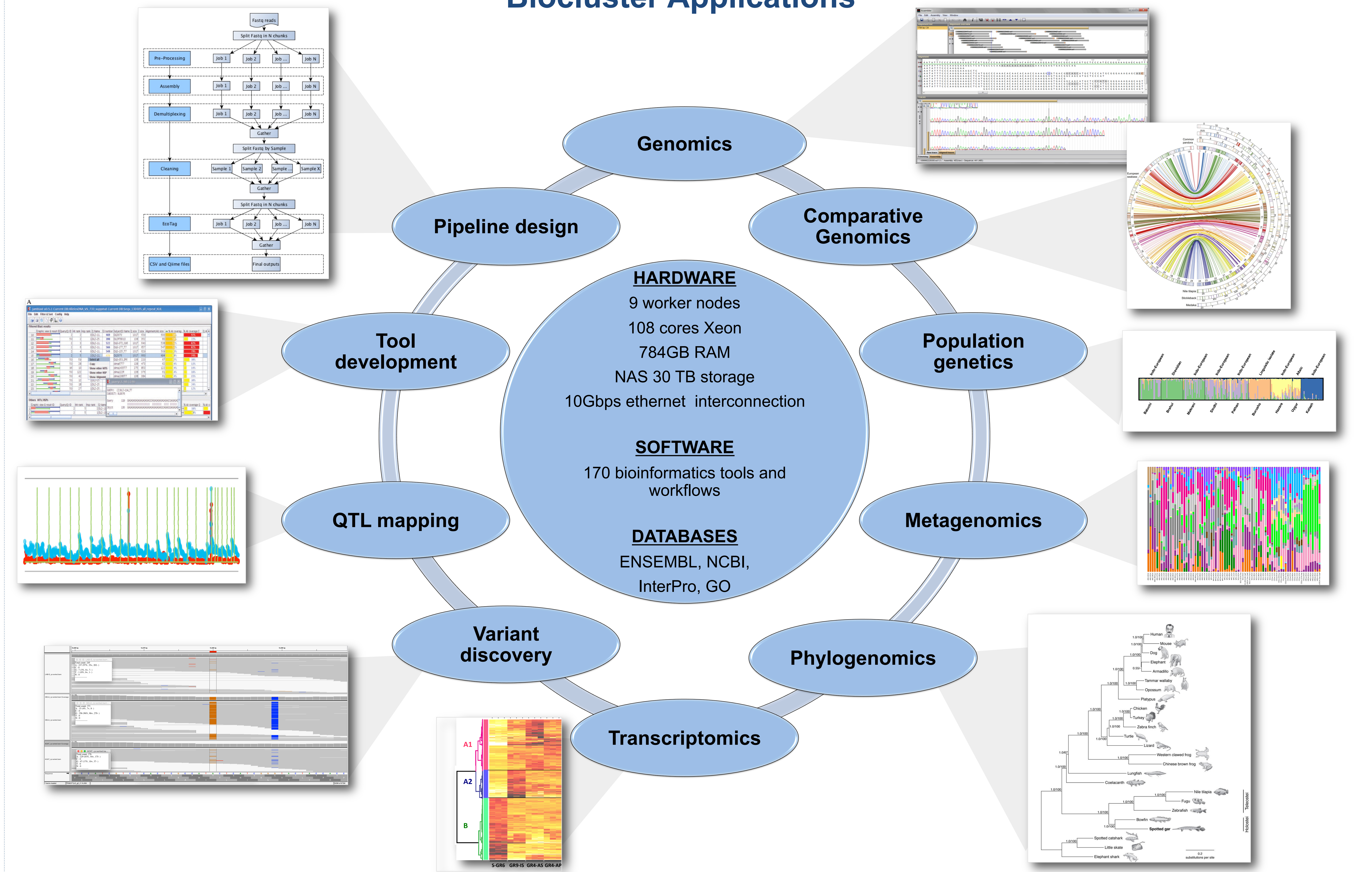
Biocluster: an NGS-dedicated HPC cluster in IMBBC, HCMR

Lagnel J, Manousaki T, Kotoulas G., Magoulas A.
Institute of Marine Biology, Biotechnology and Aquaculture (IMBBC)
Hellenic Centre for Marine Research (H.C.M.R.) Crete



Modern biology is largely shaped by the development of next generation sequencing (NGS) technology. Researchers sequence massively model and non-model organisms from single cell transcriptomes, to whole genomes with applications that range from medicine up to ecology and conservation. However, the great challenge following the scaling up of sequencing throughput is data analysis. The amounts of data produced by NGS experiments require high computational power and cannot be analyzed in desktop computers. This fact underlines the need for HPC platforms that allow the analysis of high throughput data in reasonable timeframes. In IMBBC (HCMR), an HPC cluster named Biocluster has been built, dedicated to bioinformatics applications and NGS data analysis. Since 2010, when the cluster was first launched, we have been configuring it to accommodate more than 200 state-of-the-art pieces of software. On top of that, we have been developing parallelized pipelines for NGS data analysis with special focus on the challenges rising from sequencing non-model species. The parallelized pipelines include raw reads pre-preprocessing, gene annotation, variant discovery, population genetic analyses, metabarcoding analyses and others. Biocluster can accommodate all possible OMICS data schemes in an efficient and optimized way, allowing analyses for various experimental designs in a speed comparable to that of modern sequencing data production. The unprecedented collection of sequence analysis software and the availability of parallelized pipelines turn this platform to a unique bioinformatics tool.

Biocluster Applications

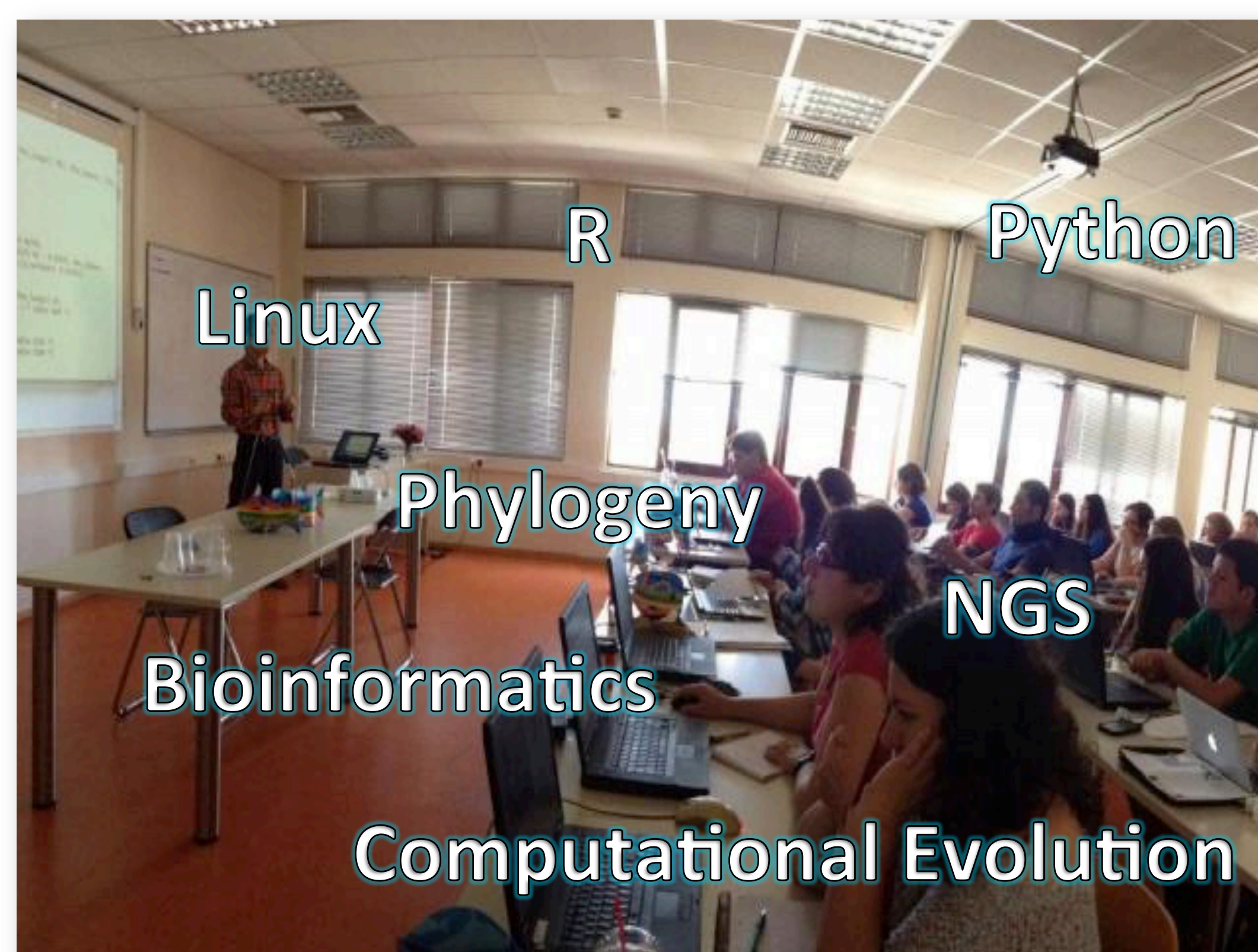


Usage/Partners

- 43 users
- 11 Research institutes
- 5 Countries: Greece, Italy, France, Norway, Cyprus
- > 30 peer-reviewed papers



Training



Upcoming Upgrade

- 12 nodes
- 232 cores Xeon
- 2 TB RAM
- 120 TB storage
- Infiniband network