

PLAZA-Cloud: A framework for integrative genomics



Sander Thierens^{1,2}, Michiel Van Bel^{1,2}, and Klaas Vandepoele^{1,2}

¹Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium
²Center for Plant Systems Biology, VIB, 9052 Ghent, Belgium

Problem

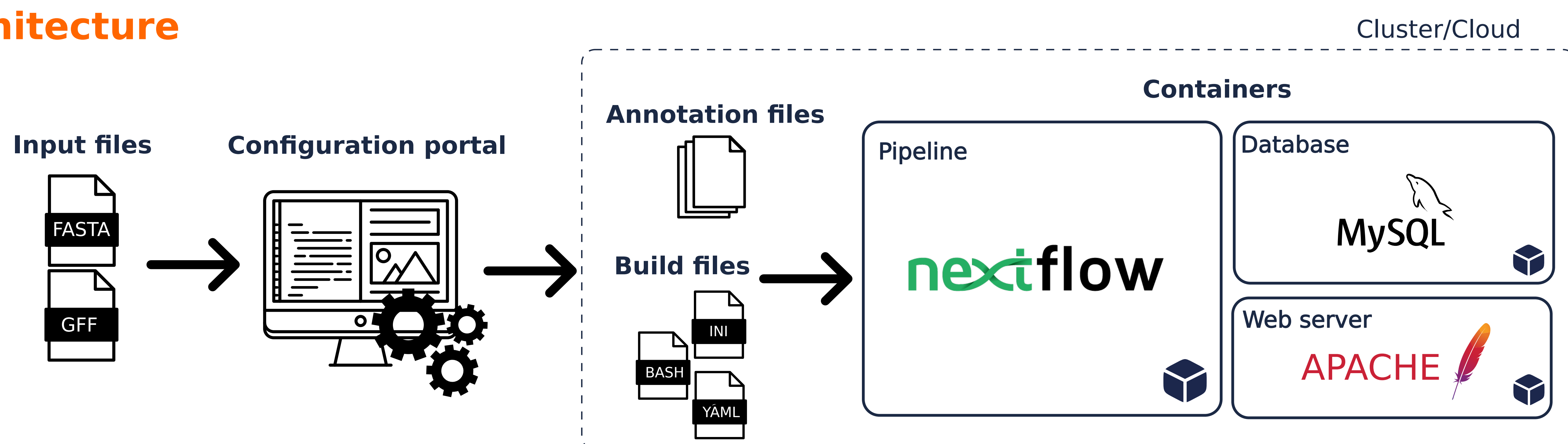
The ever-expanding volume of available genomic data has fueled the need for efficient and accessible **platforms for comparative and functional genomics**. In 2009, **PLAZA** emerged as a specialized platform dedicated to plant genomes. However, the rapid influx of novel plant genomes has left the community with a constant **demand for frequent additions and updates**.

Solution

PLAZA-Cloud

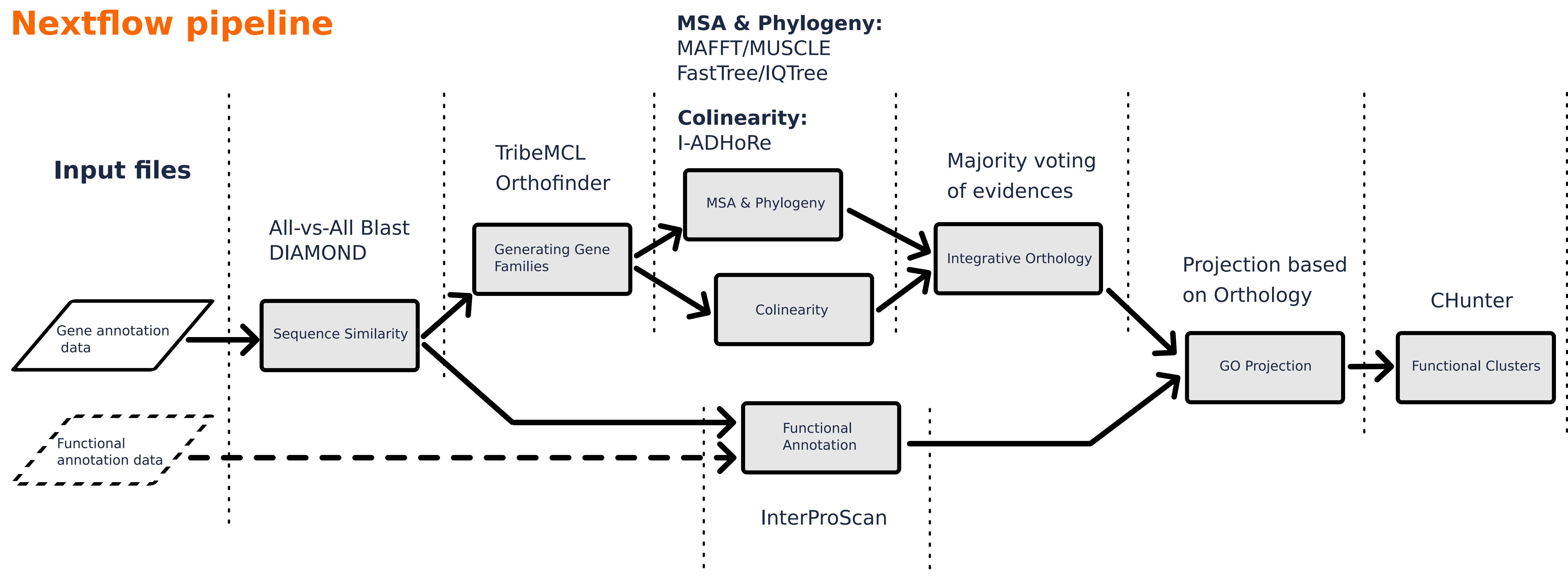
- You can not only select species, you can also upload your genome
- Configure the tools and pipeline based on your needs
- Build and deploy on your own resources/ cloud

Architecture



Starting from **coding sequences** and **GFF files**, users can import novel species data in the configuration web application and combine it with **reference data from PLAZA or Ensembl**. The application will validate new data and offers several **configuration options per process**. After this configuration, the user can deploy the build on his own resources.

Nextflow pipeline



Conclusion

We developed a **framework** which enables users to build their own **comparative, evolutionary, and functional genomics platform** based on novel and existing genomes. Using Nextflow and its features as the backbone, the framework minimizes both the dependency tree and the runtime.

Nextflow Techniques

- Container support
- Custom HTTP endpoint for logging
- Cluster support

