

# Unlocking Genomic Insights: Structural and Functional Annotation Services at CNRGV

Charlotte Cravero, Stéphane Cauet, William Marande, S. Vautrin, A. Theron, N. Rodde, I. Dufau, C. Callot, N. Gautier, R. Schwartz, N. Arnal, D. Pujol, L. Lancon, A. Bellec  
Centre National de Ressources Génomiques Végétales (CNRGV), INRAE

The massive and exponential production of genomic data using high-throughput sequencing technologies has revolutionized the field of genomics. However, these raw data remain unusable for biological applications in the absence of annotation. The annotation of a genome therefore represents a key step, essential for extracting and interpreting the information encoded in the DNA of an organism.

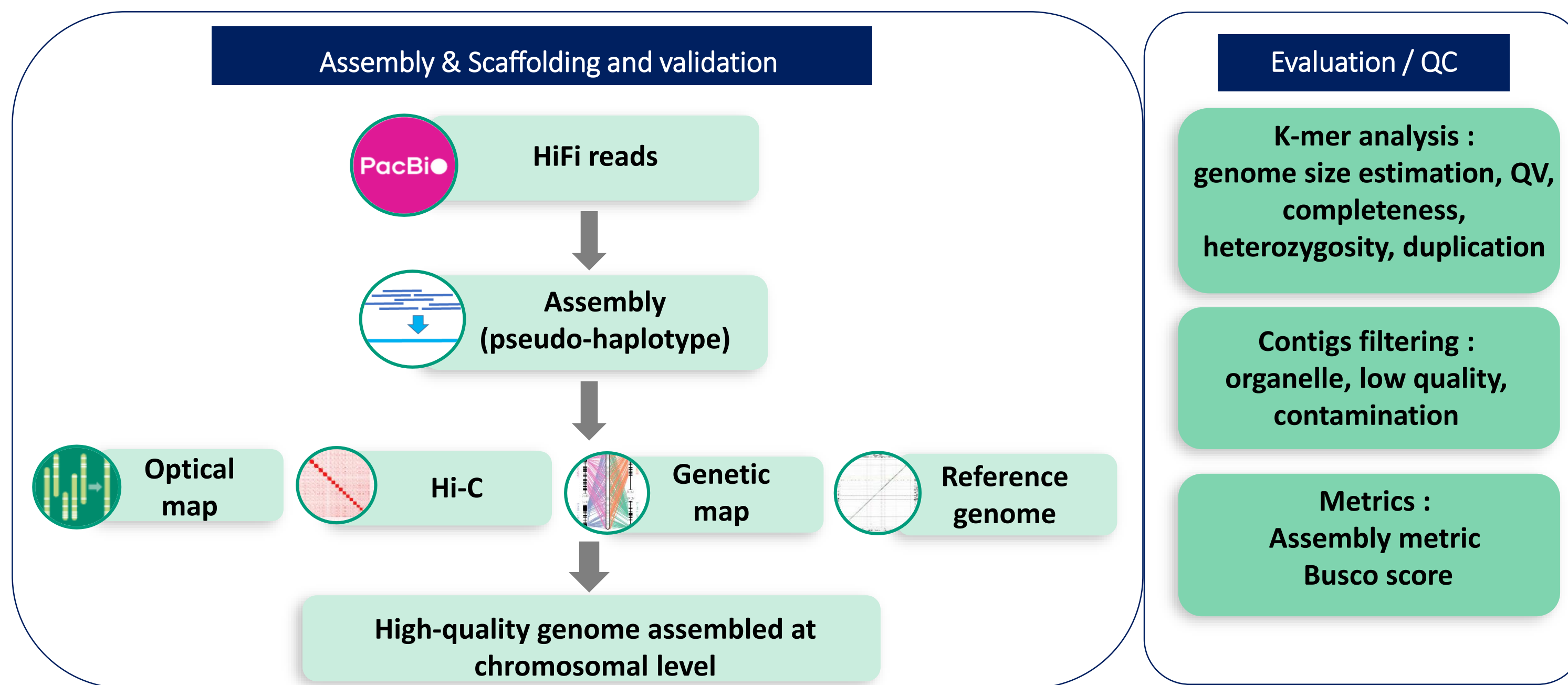
The annotation process is broken down into **two major stages**:

- **Structural annotation** which makes it possible to **identify the structural elements of the genome** (protein-coding genes, regulatory regions and non-coding sequences) and to **map its genetic organization**.
- **Functional annotation** which consists of **characterizing the functional elements** of a genome by attributing to them a **biological function**.

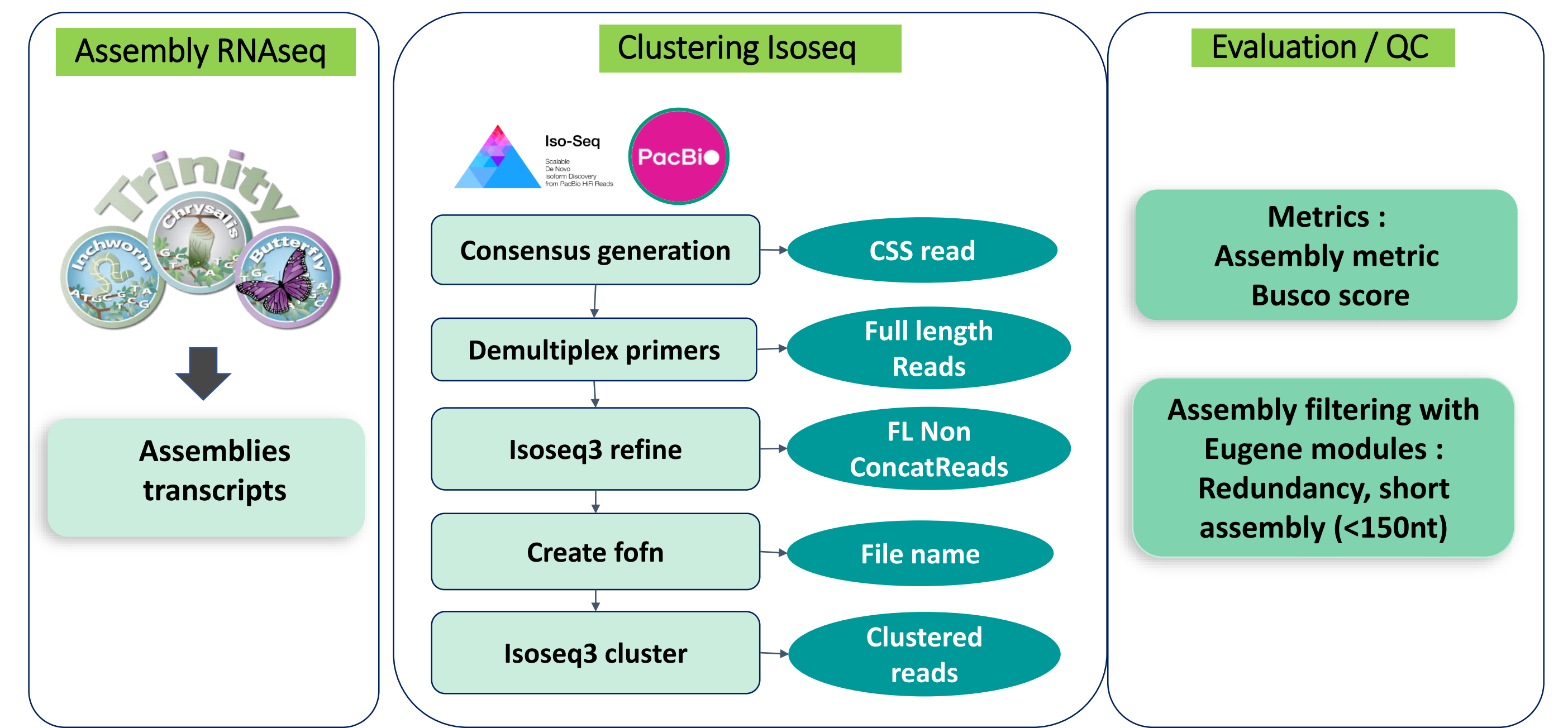
Together, these two annotations provide an exhaustive and interpretable view of the genome, essential for understanding biological mechanisms, discovering new therapeutic targets, or even optimizing organisms for industrial or agricultural applications.

In this context, the development of advanced services dedicated to genome annotation has become crucial to fully exploit the potential of genomic data. At **CNRGV**, we offer a comprehensive **"seed to annotated genome" service**, capitalizing on recognized technical expertise in the **production of high quality plant material and high molecular weight DNA**, as well as data analysis across **assembly, scaffolding and annotation** of genomes.

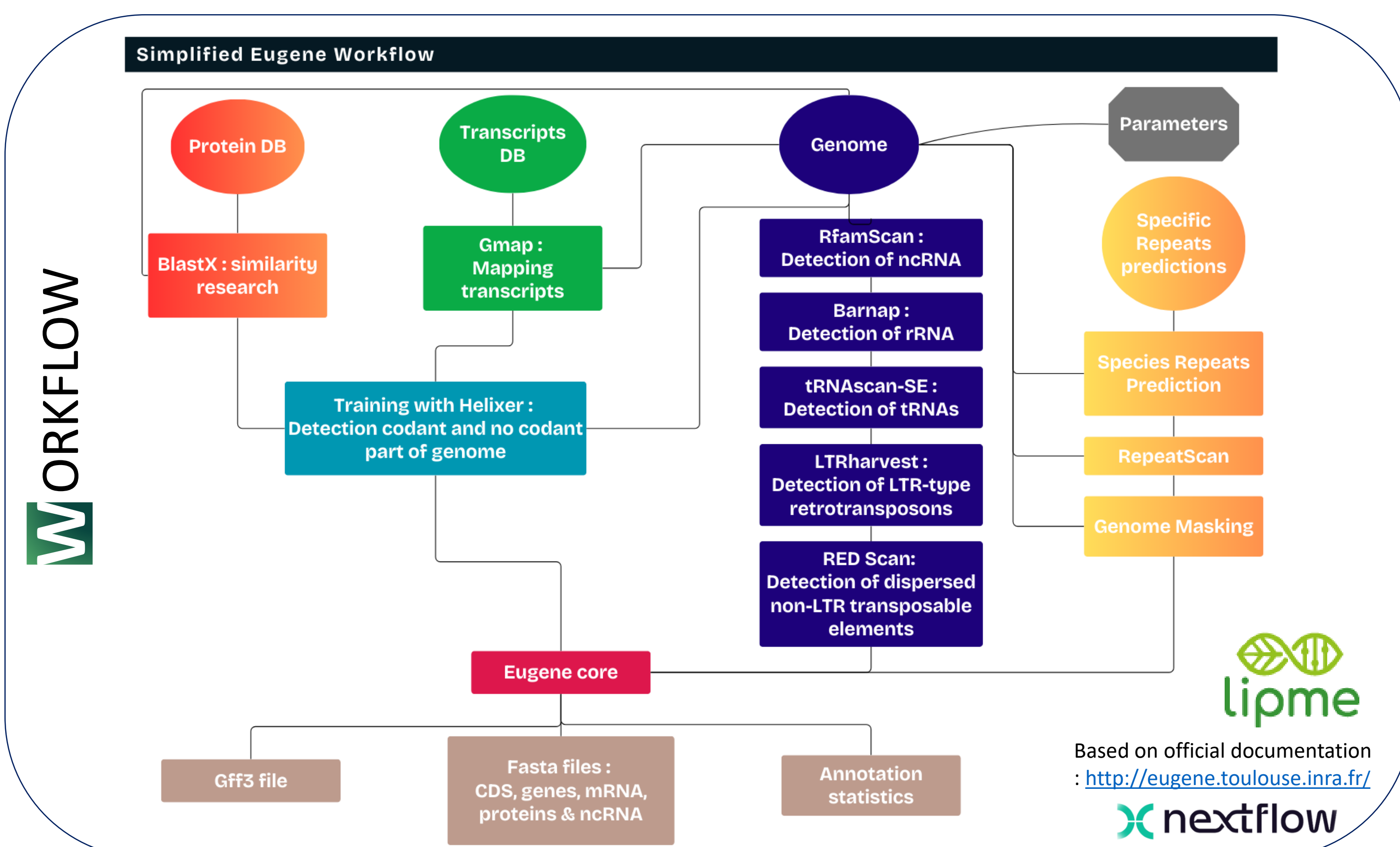
## Genome Assembly & Scaffolding



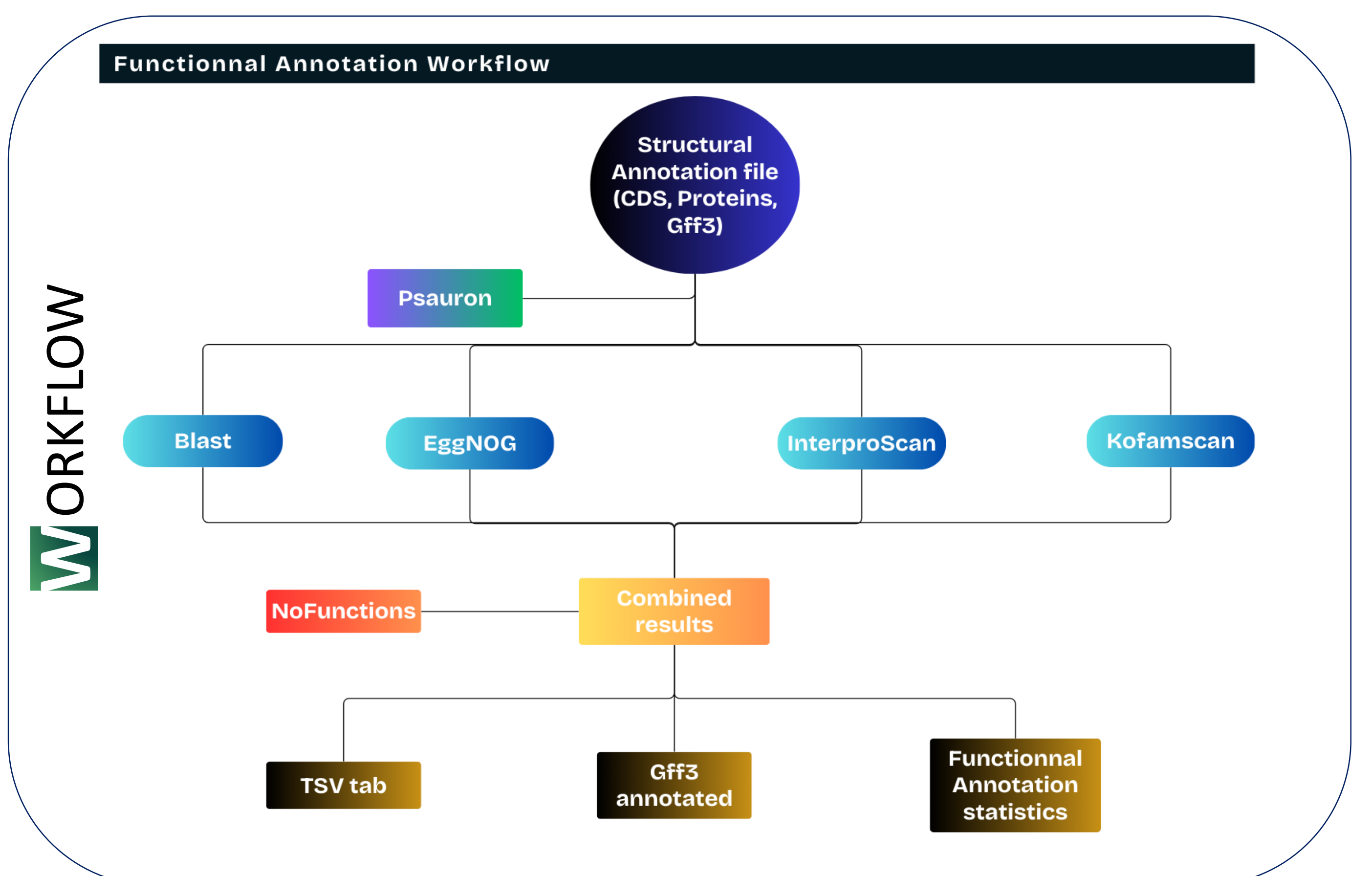
## Transcripts Assembly or Clustering



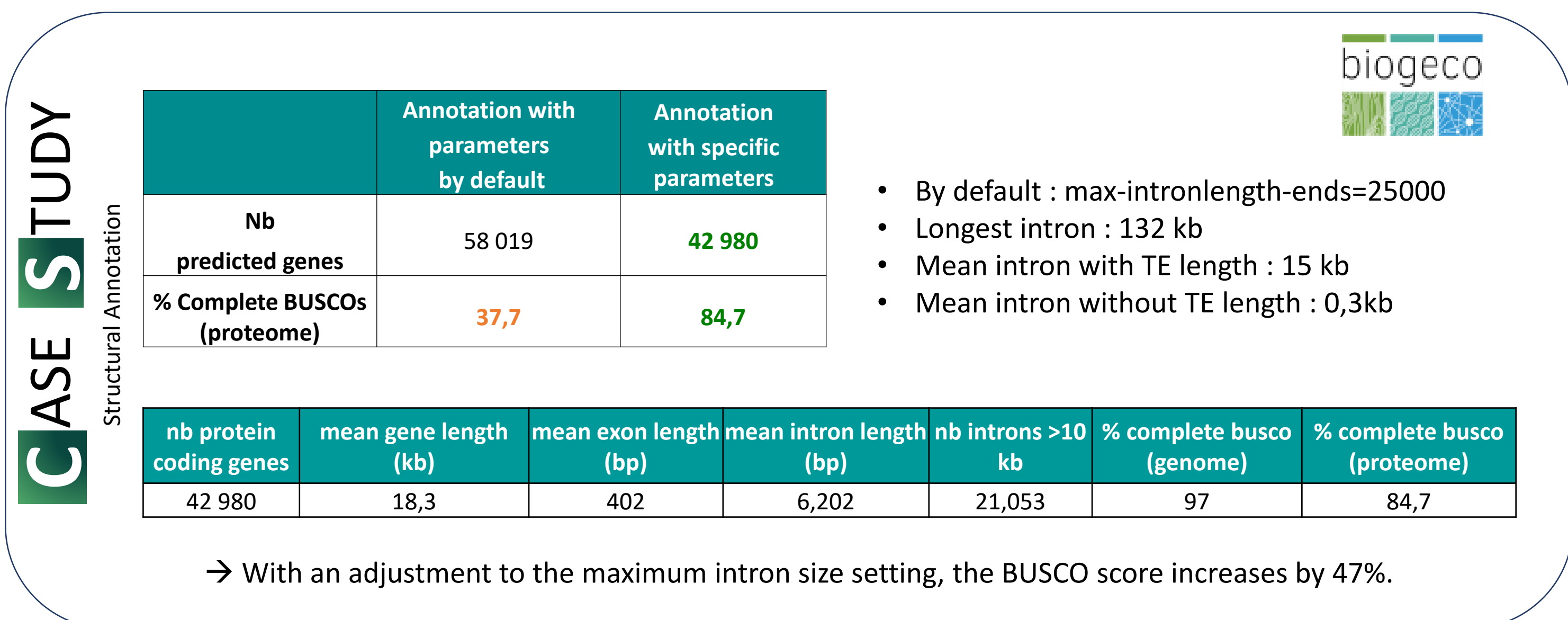
## Pipeline of structural annotation



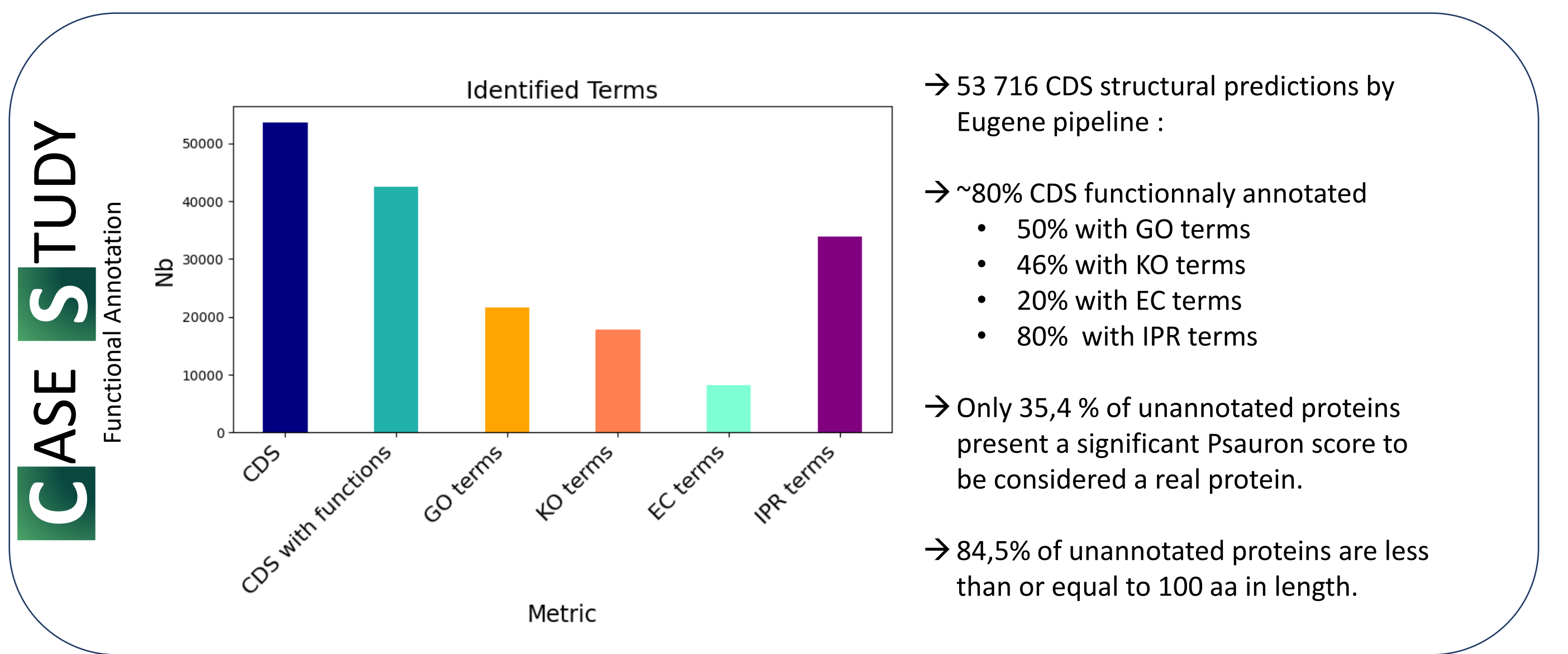
## Pipeline of functional annotation



## Cupressus sempervirens genome



## Rosaceae genome



## Highlights

### Structural annotation :

- Powerful software, less than one day per annotation
- Use transcripts data
- Nextflow pipeline
- Updated and improvements with new release Eugene v2.2.1 (december 2024)
- Increase calculation performance

### Functional annotation :

- Developed by CNRGV and adaptive for our collaborators
- Help for structural annotation validation

**Acknowledgements :** The authors would like to thank everyone who helped generate and analyze data for these poster :

- ❖ LIPME, INRAE-CNRS, Castanet tolosan
- ❖ UMR BIOGECO, INRAE & Cupressus project collaborators.
- ❖ Genotoul bioinformatics platform Toulouse (<https://doi.org/10.15454/1.5572369328961167E12>)



## CONCLUSION

The structural and functional annotation of a genome is not limited to a technical step: it constitutes a strategic lever for exploiting the full potential of genomic data. By translating raw sequences into actionable biological information, it opens the way for major advances in fundamental research, health, agriculture, and the biotechnology industry.

The service we offer aims to simplify and accelerate this complex process, by offering tools and expertise adapted to the needs of researchers and companies. Thanks to robust methodologies and cutting-edge technologies, we enable our users to better understand the genomes they study and to achieve their scientific or industrial objectives with precision and efficiency.

Thus, genomic annotation is not only a key step, but a gateway to innovation and discovery, meeting the challenges of a constantly evolving scientific and technological world.

**Ask a genome analysis service?**



<https://cnrgv.toulouse.inrae.fr/Services/Whole-genome-characterization2>



24 Chemin de Borde Rouge  
Auzeville - CS 52627  
31326 Castanet tolosan cedex FRANCE  
Tél : 05.61.28.52.53  
Email : infocnrgv-toulouse@inrae.fr