

# *Perspectives and challenges of Computational Pan-Genomics*

Eric Rivals - Methods and Algorithms for Bioinformatic  
LIRMM, CNRS, Univ. Montpellier

July 23 2019

LifeTime Unconference, Barcelona, Spain, July 22-23 2019

French title: *Enjeux et défis de la pan-génomique computationnelle*

## **Abstract**

The acquisition of a collection of individual genome sequences taken from a population of the same species aims at building a representative catalog of variant genomes. However, bioinformatic analyses are mostly performed against a single reference genome that does not account for the genetic variability in the population. Such a paradigm is inadequate to leverage the variability acquired in population genomic studies. An alternative approach is needed. Such a set of variant genomes can be represented as a graph in which paths represent sequences, and where nodes can record whether the piece of sequence they represent is shared among individuals or specific to some of them. This kind of representations, often termed "pan-genome graph" or "variation graph", can effectively represent multiple reference genomes. New questions arise: how can we build and use variation graphs? Which algorithms can address variation graphs to perform procedures commonly used with a single reference? This is the domain of computational pan-genomics. In this presentation, I will give an overview of data structures capable of storing such graphs, and of the perspectives and challenges of analysing collectively (at once) a set of individual genomes. It is foreseeable that this will be the common situation of sequence analysis in a very near future.

## **Pointers and references**

1. Computational pan-genomics: status, promises and challenges

The Computational Pan-Genomics Consortium

Briefings in Bioinformatics, Vol. 19(1), p. 118–135, 2018.

<https://doi.org/10.1093/bib/bbw089>

2. Rebooting the Human Genome

MIT tech review 2015

<https://www.technologyreview.com/s/537916/rebooting-the-human-genome/>

