Gulbenkian Training Programme in Bioinformatics

CPANG22

Computational Pangenomics

May 23rd - 27th 2022

Course description

Reference genomes are central to most bioinformatics approaches. However, the use of a single reference genome to guide an analysis can result in reference bias: other genomes appear more similar to the reference than they actually are. We might miss, or misunderstand information about genome variability and relationship that cannot be expressed relative to a chosen reference genome. Moreover, new assembly methods are making it easier than ever to generate high-quality complete genome assemblies. To obtain a total understanding of variation between multiple whole genomes, we need to use a model that expresses many genomes and their mutual alignment. A general solution to these problems is to use a pangenome graph wherein genomes are described as labeled walks through an underlying s equence graph. In this course, we will work with methods to build such graphs from whole genome assemblies, and to use the built graphs in an array of downstream applications in comparative genomics, evolution, variation analysis, sequence alignment, and phenotype association.

Instructors

Erik Garrison
University of California, Santa Cruz, US

Andrea Guarracino
Human Technopole
Milano, IT

. . .

GTPB Organiser
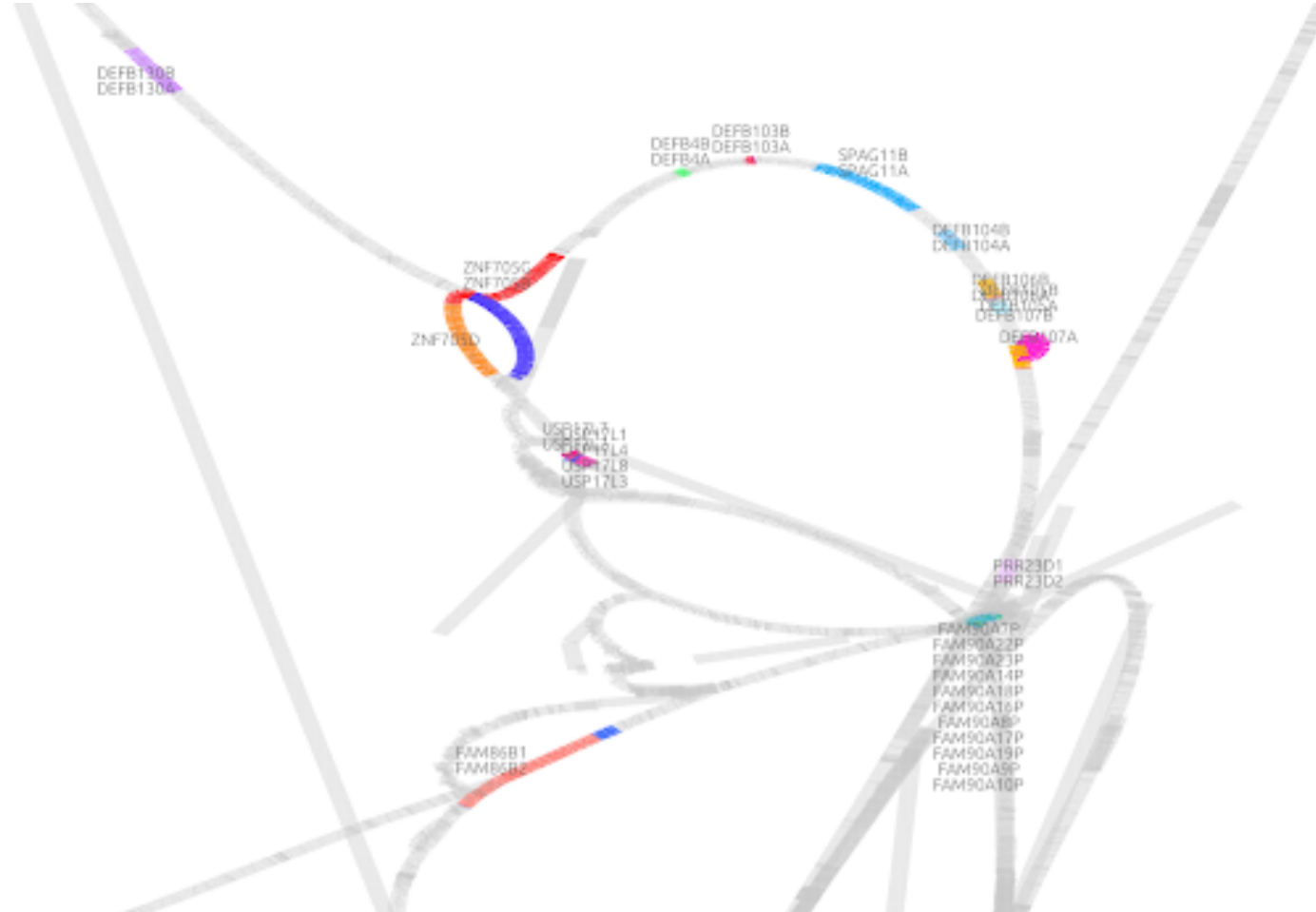
Pedro L. Fernandes
Instituto Gulbenkian de Ciência
Oeiras, PT
. . .

Sponsorship

OEIRAS VALLEY
PORTUGAL
MUNICÍPIO
OEIRAS

Course website
http://gtpb.igc.gulbenkian.pt/bicourses/2022/CPANG22