Bioinformatics in the last decade: between quantitative biology and sequence analysis

Roderic Guigó Serra[1]

[1] Center for Genomic Regulation, Universitat Pompeu Fabra

Bioinformatics oscillates between sequence analyses, which is largely discrete, and quantitative biology. The initial success of bioinformatics built on the relevant biological information encoded in the sequence of genomes and proteins. Linguistic-related methods, mostly based on string comparisons, become powerful tools to mine this information. As high throughput, highly automated methods to monitor genome activity quantitatively (i.e transcription) became increasingly sophisticated, biostatistical methods became dominant within the field. With the recent advent of Artificial Intelligence methods based on Large Language Models, string based methods are gaining in popularity again, emphasizing both the similarities and the differences between the human and the genome languages.

Keywords: bioinformatics, biostatistics, sequence analysis, quantiative biology.