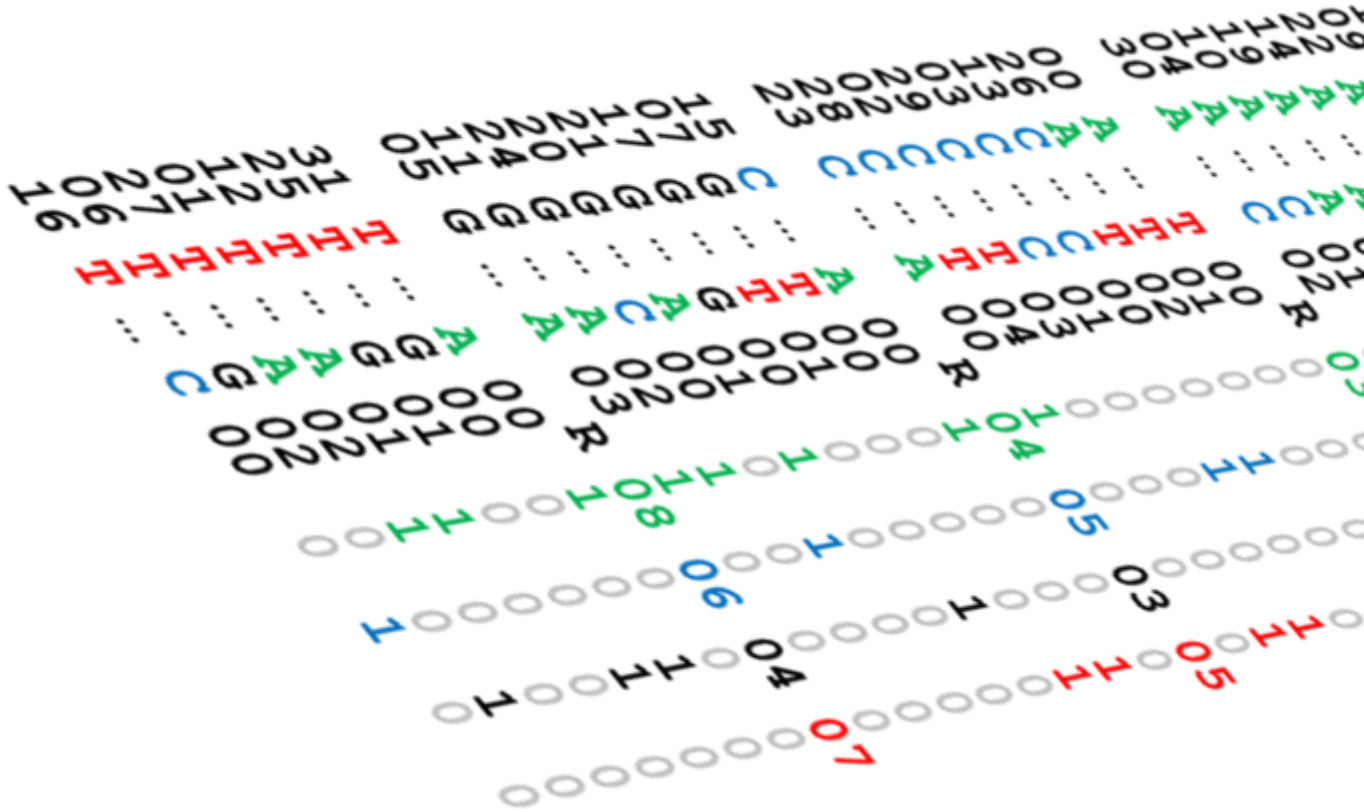


Algorithms for genomics data compression



Data transmission is one of the major bottlenecks in data management. We seek to contribute to the development of new algorithms of genomics data compression with good compression efficiency, but also to allow genomics data search and analysis to be performed in the compressed form.

Summary

Data transmission is one of the major bottlenecks in data management, especially in large genomics projects. Our group es involved in the data management if a series of large genomics projects. Although several data compression algorithms are available in the genomics field, all of them requires a full decompression before any analysis is to be performed, since requiring a significant processing time and extra storage. We seek to contribute to the development of new algorithms of genomics data compression, and to evaluate existing ones, but besides of compression efficiency, also we are studying strategies to allow search and analysis to be performed directly on the compressed files, and eventually could be performed using streaming protocols.

Objectives

- Develop and benchmark genomics compression algorithms
- Study new data models to compress genomics data compatible with analysis without decompression.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 10 Mar 2025 - 13:57): <https://www.bsc.es/ca/research-development/research-areas/genomics/algorithms-genomics-data-compression>