

CONFERENCIA

## **Indexación de colecciones altamente repetitivas**

**Gonzalo Navarro**

Universidad de Chile

Fecha: 19 de julio de 2018

Hora: 12:00h

Lugar: Seminario 3.1 - Facultad de Informática

Resumen

Las colecciones de genomas de individuos de una misma especie constituyen colecciones muy repetitivas, dado que todos ellos comparten una proporción muy importante del genoma. Por ello, para indexar dichas colecciones se puede usar esa propiedad de alta repetición para que la colección completa y el propio índice se puedan representar de modo comprimido.

Hay que entender que dicha indexación es crucial para facilitar la investigación en genómica en la que la búsqueda de subcadenas idénticas (o que se diferencien en pocos términos) a una dada como consulta es una tarea habitual que los investigadores necesitan.

La tecnología actual de alineamiento de genomas se basa precisamente en técnicas de indexación como el FM-index, sin embargo dichas técnicas de indexación no explotan adecuadamente la alta repetitividad de las colecciones que es lo que se trata de explotar ahora para crear toda una nueva generación de técnicas de indexación.