P1170

Java Merging Copy Number Variants (JM-CNV): A New Algorithm for Identifying Copy Number Variant Regions (CNVR)

Date: Monday, January 12, 2015
Room:

***Joaquim Manoel da Silva** , Universidade do Estado de Mato Grosso, Nova Xavantina, Brazil*
Alan Roberto Romaniuc , Embrapa Informática Agropecuária – Laboratório Multiusuário de Bioinformática, Campinas, Brazil
Alexandre R. Caetano , Embrapa Recursos Genéticos e Biotecnologia, Brasilia, Brazil
Poliana F Giachetto , Embrapa Informática Agropecuária, Campinas, SP, Brazil
Michel E Beleza Yamagishi , Embrapa Informática Agropecuária, Campinas, SAO PAULO, Brazil

CNVs (Copy Number Variations) are defined as copy number variations of DNA fragments typically larger than one kilobase (Kb), but less than five megabases (Mb). They represent a genomic disequilibrium that alters the ploidy in a specific locus within an individual, which may also be observed with varying frequencies within a population. An initial study estimated that 12% of the human genome shows this type of structural variation. Reliable tools have been developed to detect CNVs from molecular data produced with three main platforms: Comparative Genomic Hybridization (CGH) arrays, Single Nucleotide Polymorphism (SNP) genotyping arrays, and DNA Next-Generation Sequencing (NGS). However, processes for merging overlapping CNVs into a meaningful set of discrete Copy Number Variable Regions (CNVRs) need improvement, particularly when several CNV patterns co-exist within the same genomic locus. Available algorithms frequently merge noncontiguous CNVRs or fragment large CNVRs into multiple regions. A new web-based software (Java Merging Copy Number Variants: JM-CNV) was developed to address the aforementioned issues. The algorithm runs in three subsequent phases (parsing, sorting and merging). JM-CNV was evaluated using a dataset composed of 1,885 human individuals genotyped with 561,308 SNPs, using either Illumina 550K or Affymetrix 6.0 platforms, with a total of 30,466 CNVs detected using PennCNV. The major observed differences between results generated with JM-CNV and other tested software (CNVRuler and HD-CNV) were in the efficiency and accuracy JM-CNV dealt with complex genomic regions where several CNV patterns exist at the same locus.

Availability and implementation: JM-CNV is a web-based software freely available from http://www.lmb.cnptia.embrapa.br/tools/JMCNV

Back to: Bioinformatics: Algorithms - Even

<< Previous Poster | Next Poster >>

Home/Search

Browse by Type

Author Index

Poster Categories

Meeting Information

When:
January 10 - 14, 2015
Where:
San Diego, CA