

A MapReduce Algorithm for Detecting Genomic Structural Variation from Complete Paired-End Sequence Data Sets

Christopher Whelan Kemal Sönmez
Biomedical Engineering
Oregon Health and Science University
20000 NW Walker Road, Beaverton, OR 97006, USA

October 1, 2012

Abstract

This is the paper's abstract . . .

1 Introduction

Genomic structural variants are important..

Difficulty of detection.. Lack of concordance of current algorithms...

Typical algorithms use only unambiguous discordant read pair mappings...

Hadoop/MapReduce as framework for processing complete data sets

2 Methods

MapReduce algorithm description

Local deletion scores

3 Results

Simulated data

1000 Genomes Data
AML Data set

4 Discussion

5 Conclusions