An Efficient Means for Calling Moderate to Long Structural Variations Through Pairwise Genomic Comparison And Kmers.

W. KORANI*; J. CLEVENGER. Hudson Alpha Institute for Biotechnology, 601 Genome Way Huntsville, AL 35806

Although structure variations are less abundant than SNPs, they control many economical traits in different species. Calling valid structural variations is a computationally expensive and time-consuming process. We developed a computational approach, designated Cyclops, to compare two genomes and extract moderate to long insertions and deletions utilizing high multiple threading, which allows running each chromosomal pair in a thread in pairwise comparison of all chromosomes from the two genomes. A second round of runs is implemented to extract high quality structure variances by using Kmers extracted from Deletions/Insertions and aligned to the opposite genome, then re-stacked based on the hit results. In a case study using a novel assembled NC94022 genome against the Tifrunner genome, Cyclops recognized 309,186 deletions and 17,009 insertions, varied from 40 bp to 39 kbp in less than three hours.