Abstract ID:- 195

Abstract Topic:- Evolutionary and population genetics

Abstract Title:- Identification and determination of parent of origin for de novo mutations

Presenting author name :- Thota Kiran

Presenting author institute:- CSIR-CCMB

Co-authors name:- Avvaru Akshay Kumar, Divya Tej Sowpati

Co-authors institute:-CSIR-CCMB

Aims:-de novo mutations are genetic changes that occur for the first time in an individual and are not inherited from their parents, potentially contributing to genetic disorders or diversity. Profiling de novo mutations is vital in clinical contexts, understanding population dynamics and also sheds light on developmental biology, DNA replication and maintenance. Delineating the parent-of-origin (PoO) of de novo variants additionally reveals parental specific effects on germline mutation rates. Inclusion of trio datasets allows us to identify de novo mutations that raises the need for a computational tool to identify and phase them confidently. DeMinTR (pronounced dementor) is a tool to identify and phase de novo mutations from WGS datasets for STRs, SNVs and Indels.

Methods:- DeMinTR is designed to identify de novo SNV, indels and STR variants and phase them. The approach is similar for all three genomic variations. Addressing STR variations requires additional methods owing to their unique length polymorphisms. DeMinTR starts with an input joint vcf file of the trio samples. For STRs we currently support outputs of both GangSTR and HipSTR tools. The first step of de novo variant identification picks an allele in the offspring which is not possible to be inherited from parents. We filter out allele drop out cases by checking for reads with the de novo allele in parent samples. Correspondingly calculates de novo likelihoods using genotype probabilities in parental vs offspring samples. Phasing of de novo variants is based on two methods, allele sharing and read tracing. In allele sharing we infer the PoO of the de novo allele where the inherited allele has only a single possibility of origin. In read tracing, PoO is determined by building a haplotype with proximal phased heterozygous SNV using read/read-pairs.

Results:- We phased de novo variants identified from 602 Trio samples of the 1000 genome project. DeMinTR phased results are overlapping more than 98% with existing tool MonSTR for STRs and adding up the Read tracing method we are able to phase 20% additional de novo variants. DeMinTR is capable of phasing SNV and INDELS given the de novo sites. We validated with a number of high quality short and long reads and phased with multiple proximal heterozygous SNVs.

Conclusions:- DeMinTR determines the PoO using two methods which uncovers the most cases where existing tools fail to phase. MonSTR limits to only allele sharing methods which leads to drop in many cases. Unfazed, a tool overlooks nearby heterozygous SNVs while phasing. Both methods we used overcome the limitations of existing tools and increase the sensitivity of results. We are further focusing on phasing SVs.

Keywords:- de novo, Germline, Parent of origin