

Biological context in C-Siberia

Ekaterina Starostina

Advisors: Ilya Minkin, Son Pham, Nikolay Vyahhi

30.03.2013

Tasks

Goal of project: develop variant annotation and effect prediction module for C-Sibelia.

- Add VCF output
Output variants and rearrangements found by C-Sibelia in VCF format
- Add annotation of SNPs and indels
Develop an application that is capable to annotate found variants using external databases.
- Add annotation of mobile elements
Add capabilities of recognition of mobile elements

SnpEff

SnpEff is a variant annotation and effect prediction tool. It annotates and predicts the effects of genetic variants (such as amino acid changes).

Work with: indels, SNPs, MNPs

Effects: START_GAINED, SPLICE_SITE_ACCEPTOR, SPLICE_SITE_DONOR, START_LOST, SYNONYMOUS_START, NON_SYNONYMOUS_START, EXON_DELETED, NON_SYNONYMOUS_CODING, SYNONYMOUS_CODING, FRAME_SHIFT, CODON_CHANGE, CODON_INSERTION, CODON_CHANGE_PLUS_CODON_INSERTION, etc

Databases: SnpEff supports all bacterial genomes available at NCBI

Plans

- Already done:
 - VCF output for SNPs found by C-Sibelia
 - First version of script for variant annotation
- To do:
 - Improve VCF output for indels and CNVs
 - Add annotation of CNVs and mobile elements

Questions

Thank you for your attention!

