Seed modeling data, expression data, literature references verifying

SEED is used by many researchers for predicting genefunction and discovering new pathways. SEED also has a collection of proteinfamilies that are related functionally, and furthermore protein familiesderived from them. The core of RAST annotation system is the interconnection ofRAST and SEED. The SEED continuously combines different types of genomic datafrom a variety of sources. These include public genomes annotated by RAST, expert user annotation, metabolic modeling data, expression data, literaturereferences verifying annotations and links to data from popular resourcesincluding Swissport, Genbank, IMG. KEGG, CDD and so forth.

The SEED websiteworks like a google search engine for genome annotation and comparison. SEEDand RAST are a big tool to understand genome because these programs havemultiple genome analysis tools. RAST is an annotation system built on theframework provided by SEED. RAST can also identify protein coding regions, tRNA, rRNA, non-coding RNA, etc. Overall the article gives a detailed informationon RAST and SEED relationship and how future developments of RAST can make iteven more easier for researchers to annotate genome using wide variety oftools.

Mauve is a great tool when there is a need to constructmultiple genome alignments in the presence of rearrangement or inversion. Sincemutations will occur with the course of evolution of which some are large scalemutations (include gain or loss of large segments generated by unequalrecombination) or Local mutations (nucleotide substitution, insertion, deletion). With the help of Mauve, it is possible toexamine what has changed in a mutant genome sequence as compared to parent genomesequence. Progressive Mauve generates positional homology

multiple genomealignments. Mauve doesn’t require the user to use complex algorithms, insteadit is very user friendly.

The time frame in which an alignment is completed, dependson job load. Resfinder is a web based system that depends on BLAST foridentification of acquired antimicrobial resistant genes in whole genomic data. The article mentions that 1862 genbank files were tested.

The goal was toidentify and study the antimicrobial resistant genes. As researcher’s research, more and more genes on this platform, the database accumulates the informationand grows with every entry. To experiment, Resfinder was used on twenty-threeisolates of five different bacterial species and on WSG chromosomes andplasmids of 30 isolates.

A few of these isolates were annotated by the system, to have a microbial resistance. Overall resfinder is an easy and free resourcethat can help identify the antimicrobial resistant genes. Phylosift is a method to analyze metagenomics samples fromphylogenetic viewpoint of community structure among multiple related samples. Theanalysis can be broken down into four stages. In the first stage the query is searchedfor in the reference gene families. Then the second step is that inputsequences are added to a multiples alignment sequence with reference genes. Thethird step is that the input sequences is placed on a phylogeny of reference genes.

The fourth and the last step is the generation of taxonomic summaries. The standardphylosift database has 37 “ Elite” gene families identified as universal andpresent in a single copy. 16s and 18s ribosomal RNA genes,

mitochondrial genefamilies, viral gene families are also part of the database. In total 800 gene familiesexist of which most are viral.

Phylosift provides several advantages over OUT-basedor taxonomic analysis for metagenomics data. By correctly aligning unknownsequences within a known topology, with the help of evolutionary models; phylosiftreduces the risk of errors. Phylosift can also prevent errors in microbial forensics. With all these great functions, phylosift proves to be a great tool for researchin metagenomics.