

FTB 55 (2) 251-257.

(FTB-4749)

MEGGASENSE – pretraživač (meta)genomski anotiranih sekvencija pomoću govornog jezika – platforma za izradu bioloških skladišta podataka

Sažetak

Platforma MEGGASENSE služi za izradu relacijskih baza podataka koje sadržavaju nukleotidne ili proteinske sekvencije. Osnovna funkcionalna analiza zasniva se na primjeni 14 106 profila skrivenih Markovljevih modela (HMM), temeljenih na sekvencijama dostupnim u bazi podataka KEGG. Pomoću tražilice Solr mogu se zadati napredni upiti u sprezi s implementiranom pretragom BLAST. Osnovne funkcionalnosti platforme omogućile su izradu baze podataka SCATT, temeljene na predviđenom proteomu bakterije *Streptomyces cattleya*. U radu je opisana implementacija specijalizirane metagenomske baze podataka (AMYLOMICS) za „bioprospecting“ enzima koji modificiraju ugljikohidrate. Uz standardno slaganje očitanih kratkih sljedova DNA, razvijen je funkcionalni postupak pretraživanja HMM profila u očitanim slijedovima DNA prije slaganja. Baza podataka AMYLOMICS sadržava i dodatne HMM profile enzima za modifikaciju ugljikohidrata. U radu je prikazano kako se kombinacijom analiza HMM i BLAST mogu identificirati ciljani geni. Platforma MEGGASENSE upotrijebljena je za izradu raznih proteomskih i metagenomskih baza podataka.

Ključne riječi: „bioprospecting“, enzimi za modifikaciju ugljikohidrata, slaganje DNA