

Biđin, S<sup>1</sup>., I. Vujaklija<sup>1\*</sup>, T. Paradžik, A. Bielen, D. Vujaklija\*, *Leitmotif: protein motif scanning 2.0*, Bioinformatics, 2020; 36(11): 3566-2020.

## Doprinos predloženog rada području istraživanja

U predloženom radu razvijena je slobodno dostupna web aplikacija/server pod nazivom **Leitmotif** (<https://leitmotif.irb.hr>). Mogućnosti ovog bioinformatičkog alata su jedinstvene i omogućavaju korisniku brzu analizu proteinskih motiva ili njihovih varijacija u velikim bazama podataka. Rezultati testiranja potvrdili su vrijednost ove aplikacije i omogućili su nam objavu u prestižnom časopisu Bioinformatics, koji spada u prvih 5 % najboljih časopisa u području (IF 5,61).

Potreba za boljim bioinformatičkim alatima javlja se paralelno s brzim napretkom tehnologija za sekvenciranje DNA. Razvoji sekvencatora novih generacija značajno smanjuju cijenu sekvenciranja ali i rezultiraju prikupljanjem ogromnih količina podataka s pohranjenim sekvencama. Usporedno dolazi do nedostatka eksperimentalnih testova visoke propusnosti za potvrdu funkcija anotiranih proteina i zato se u analizama sve više koriste računalne metode. Međutim, glavni nedostatak računalnog pristupa je u sve većem broju pogrešaka u anotaciji proteina. Te pogreške negativno utječu na sve kasnije analize. Na primjer, pseudoenzimi su proteini s očuvanom sekvencom koja im određuje pripadnost određenoj porodici ali kojima u motivima nedostaju katalitičke aminokiseline ključne za enzimsku reakciju. Sve je jasnije kako pseudoenzimi pripadaju posebnoj klasi proteina koji preuzimaju nove i važne funkcije u stanicama i zato su izuzetno interesantni za istraživanja. Upravo zbog kompleksnih funkcija koje proteini obavljaju u stanicama njihova klasifikacija je izuzetno važna u razjašnjavanju njihovih bioloških uloga i molekularnih mehanizama u kojima proteini sudjeluju.

U radu smo pokazali kako je motif-Hidden Markov Model, do sada rijetko korišten statistički model, posebno prikladan za detekciju kratkih aminokiselinskih nizova – motiva koji karakteriziraju evolucijski srodne proteine i koji su ključni su za njihove funkcije. Leitmotif učinkovito traži varijacije ili mutacije aminokiselina u motivima što omogućuje distinkciju proteinskih podskupina ili pseudoenzima. Brojne mogućnosti i pregledno sučelje daju jedinstvenu prednost aplikaciji Leitmotif u odnosu na standardno korištene metode pretraživanja proteina (npr. HMMER, PSI-BLAST). U traženim motivima korisnik može 'fiksirati' odnosno odabrati jednu ili više aminoikselina kao nepromijenjivu te odabrati očekivanu udaljenost između proteinskih motiva, kao i udaljenost motiva od kraja proteina te razinu penalizacije za odstupanje od očekivanih vrijednosti. Leitmotif je detaljno testiran na dvije proteinske porodice, a rezultati prikazani u radu su pokazali kako aplikacija pruža korisniku jedinstvene prednosti pri traženju očuvanih tipova/podtipova proteinskih motiva i kako se može detektirati nova podporodica i/ili pseudoenzimi unutar neke proteinske porodice.

Ubrzo nakon objave Leitmotif su citirali Ping *et al* u „Global quantitative analysis of the human brain proteome and phosphoproteome in Alzheimer’s disease“ (doi: 10.1038/s41597-020-00650-8), što pokazuje njegovu vrijednost u istraživanjima strukture/funkcije proteina.