

Statističko i vjerojatnosno modeliranje DNA nizova s primjenama

Ivo Ugrina

Odjel za matematiku, Prirodoslovno-matematički fakultet,
Sveučilište u Splitu, Hrvatska

25. listopada 2018.

Sažetak

Palindromi su većini poznati kao nizovi slova koji se čitaju ili izgovaraju isto od početka ili kraja. U DNA nizovima palindrome definiramo na slučajan način kao i u prirodnim jezicima. Isto čitanje s oba kraja je nužno, ali uz prethodnu operaciju komplementiranja znakova. Recimo, za niz A C G T bi definiranjem komplementarnosti putem prirodnog uparivanja baza A T i C G definirali komplementaran niz sa T G C A te bi rekli da je početni niz palindrom jer se čita od početka isto kao i njemu komplementarni od kraja. Potreba za definiranjem palindroma u DNA nizovima, i proučavanje istih, može se činiti čudnom. Međutim, ako primijetimo da palindromi u DNA nizu opisuju spajanje između dva dijela DNA niza onda nam motivacija postaje bliža. Budući da su palindromi značajni za funkcioniranje genoma, korisno je imati nekakav oblik statističkog testa koji bi odavao ima li palindroma određene duljine značajno više ili manje od očekivanog broja unutar nekog npr. roda koji se proučava.

Ovaj uvodni primjer prikazati će važnost statističkog modeliranja DNA nizova i najjednostavnijim primjerima. Nakon uvodnog dijela predavanja prezentirat će se i teži problemi pri modeliranju DNA nizova poput procjene udjela bakterija u ljudskom probavnom traktu kroz proces od "raw" podataka do finalnih reportova koji bi se prikazivali u standardnom procesu u laboratorijima. Također, pri kraju predavanja kratko će se uvesti problem integracije različitih više-dimenzionalnih bio-kemijskih podataka.