

TUGAS 3 INDIVIDU

Jawab pertanyaan berikut. Untuk setiap pertanyaan, harap catat jawaban Anda, dan apa yang Anda lakukan / ketikkan untuk mendapatkan jawaban ini.

Q1.

Berapa panjang (jumlah total pasangan basa dalam) genom mitokondria *Schistosoma mansoni* (aksesi NCBI NC_002545), dan berapa banyak As, Cs, Gs, dan Ts yang dikandungnya?

Anda harus mencari urutan ini melalui situs web NCBI, karena tidak ada dalam database ACNUC. Catatan: *Schistosoma mansoni* adalah cacing parasit yang bertanggung jawab menyebabkan schistosomiasis, yang diklasifikasikan oleh WHO sebagai penyakit tropis terabaikan.

Q2.

Berapa panjang genom mitokondria *Brugia malayi* (aksesi NCBI NC_004298), dan berapa banyak As, Cs, Gs, dan Ts yang dikandungnya?

Anda harus mencari urutan ini melalui situs web NCBI, karena tidak ada dalam database ACNUC. Catatan: *Brugia malayi* adalah cacing parasit yang bertanggung jawab untuk menyebabkan filariasis limfatik, yang diklasifikasikan oleh WHO sebagai penyakit tropis terabaikan.

Q3.

Berapa probabilitas urutan genom mitokondria *Brugia malayi* (aksesi NCBI NC_004298), sesuai dengan model multinomial di mana probabilitas As, Cs, Gs dan Ts (p_A , p_C , p_G , dan p_T) disetel sama dengan fraksi As, Cs, Gs, dan Ts dalam genom mitokondria *Schistosoma mansoni*?

Q4.

Apa tiga kata 4-bp paling sering (4-mers) dalam genom bakteri *Chlamydia trachomatis* strain D / UW-3 / CX (aksesi NCBI NC_000117), dan berapa kali mereka muncul dalam urutannya?

Catatan: *Chlamydia trachomatis* adalah bakteri yang bertanggung jawab untuk trachoma, yang diklasifikasikan oleh WHO sebagai penyakit tropis terabaikan.

Q5.

Tulis fungsi R untuk menghasilkan urutan DNA acak yang panjangnya n huruf (yaitu, panjang n basis) menggunakan model multinomial di mana probabilitas p_A , p_C , p_G , dan p_T ditetapkan sama dengan fraksi A_s , C_s , G_s dan T_s dalam genom mitokondria *Schistosoma mansoni* (di sini p_A berarti probabilitas A_s , p_C adalah probabilitas C_s , dll.)
Petunjuk: lihat halaman bantuan untuk fungsi "sample ()" dalam R, karena mungkin berguna untuk digunakan dalam fungsi R Anda.

Q6.

Berikan contoh penggunaan fungsi Anda dari **Q5** untuk menghitung urutan acak yang panjangnya 20 huruf, menggunakan model multinomial dengan $p_A = 0,28$, $p_C = 0,21$, $p_G = 0,22$, dan $p_T = 0,29$.

Q7.

Berapa banyak urutan protein dari virus rabies yang ada di database NCBI Protein?
Anda harus mencari urutan ini melalui situs web NCBI, karena tidak mungkin melakukan pencarian ini menggunakan SeqinR. Catatan: virus rabies adalah virus yang bertanggung jawab atas rabies, yang diklasifikasikan oleh WHO sebagai penyakit tropis terabaikan.

Q8.

Apa akses NCBI untuk genom virus Mokola?

Catatan: Virus Mokola dan virus rabies adalah virus terkait erat yang keduanya termasuk dalam kelompok virus yang disebut Lyssavirus. Virus mokola menyebabkan infeksi seperti rabies pada mamalia termasuk manusia.