

## BLAST üzrə tapşırıqlar

### Tapşırıq 1a:

Bir bakteriyaya aid 16S rRNA geni ardıcılığını əldə et və BLASTN vasitəsilə hansı növ və ya cinsə aid olduğunu müəyyən et. Cavabında bakteriyanın adını və identifikasiyanı göstər.

### Tapşırıq 1b:

Bir neçə **DNA ardıcılığını** FASTA formatında bir fayla yazın və bu faylı Biopython vasitəsilə oxuyub hər bir ardıcılığın başlığını (ID) çap edin.

### Tapşırıq 2a:

Bir heyvan zülalının (məsələn: insulin) ardıcılığını BLASTP ilə analiz et. Eyni zülalın digər növlərdə (məs: siçan, meymun, insanda) oxşarlıqlarını müqayisə et və təkamül əlaqələrini şərh et.

### Tapşırıq 2b:

**Amin turşusu ardıcılığının** ilə **genetik kodun** (genetic code) əsasında tərcümə edilməsini təmin edin. Tərcümə edilmiş zülal ardıcılığını faylda saxlayın.

### Tapşırıq 3a:

BLASTX istifadə edərək bilinməyən bir nukleotid ardıcılığının hansı zülalı kodladığını təyin et. Zülalın funksiyası və hansı orqanizmdən ola biləcəyi barədə təhlil ver.

### Tapşırıq 3b:

Verilmiş bir zülal ardıcılığını "**Pairwise Sequence Alignment**" ilə müqayisə edin və əldə olunan alignmentin əmsallarını çap edin.

*İstifadə:* Bio.pairwise2.align.globalxx() və ya Bio.pairwise2.align.localxx()

### Tapşırıq 4a:

İnsan genomundan bir genin mRNA ardıcılığını tap və tBLASTX istifadə edərək digər canlıların genomlarında bu genin konservasiyasını araşdır. Hansı canlılarda daha çox oxşarlıq var?

#### Tapşırıq 4b:

Bir zülal ardıcılığı ilə onun qarşılıq gələn **nukleotid ardıcılığını** (reverse complement) əldə edin. Hər iki ardıcılığın identik olub-olmaması məsələsini həll edin.

#### Tapşırıq 5a:

SARS-CoV-2 virusunun spike protein ardıcılığını əldə et və BLASTP ilə digər koronaviruslarla müqayisə et. Hansı növlərlə daha yaxındır? Mutasiyaları izah et.

#### Tapşırıq 5b:

Verilən zülal ardıcılığını **Prosite** database-i ilə təhlil edin və ardıcılıqda hansı protein domenlərinin olduğunu tapın.

*İstifadə:* Bio.ExPASy və Bio.SwissProt

#### Tapşırıq 6a:

İnsan mtDNA-dan bir hissə seç və BLASTN vasitəsilə fərqli populyasiyalar arasında bu ardıcılığın fərqlərini təhlil et. Ən çox fərqlilik hansı bölgəyə aiddir?

#### Tapşırıq 6b:

Bir **DNA ardıcılığından** amin turşusu ardıcılığına tərcümə etmək üçün Biopython vasitəsilə əlavə bir **genetik kodu** istifadə edərək tərcüməni icra edin.

#### Tapşırıq 7a:

Bir genom faylından (GenBank formatında) təkə CDS (coding sequence) hissələrini çıxarib FASTA formatına çevir. BLASTP ilə funksional zülalları tap və annotasiya et.

#### Tapşırıq 7b:

Biopython ilə **zülal ardıcılığının** hər bir amin turşusunun hansı xüsusiyyətlərə sahib olduğunu müəyyən etmək üçün amino acid properties kitabxanasından istifadə edərək bir statistik hesabat hazırlayın.

#### Tapşırıq 8a:

BLAST nəticələrini .xml formatında saxla və BioPython kitabxanası ilə (və ya onlayn vasitə ilə) bu fayldan nəticələri oxuyaraq "Top hit" olan gen adlarını çıxart.

#### Tapşırıq 8b:

**Pairwise alignment** üçün Biopython-un pairwise2 modulunu istifadə edərək verilmiş 2 nukleotid ardıcılığını **local alignment** üsulu ilə uyğunlaşdırın. Həmçinin alignmentin score və mismatches sayını çap edin.

*İstifadə:* Bio.pairwise2.align.localxx()

#### Tapşırıq 9a:

Nukleotid ardıcılığının daxilində yer alan bir mutasiyanın zülal səviyyəsinə təsirini BLASTX vasitəsilə təhlil et. Hər iki versiyanı müqayisə et.

#### Tapşırıq 9b:

Bir **nukleotid ardıcılığını** oxuyub Biopython ilə **complement** və **reverse complement** edin. Sonra reverse complement ardıcılığının uzunluğunu çap edin.

*İstifadə:* Seq.reverse\_complement()

#### Tapşırıq 10a:

Bir ardıcılığın BLAST nəticəsində əldə olunan *E-value*, *Query coverage* və *Max identity* anlayışlarını fərqli nəticələrdə müqayisə et və hansı nəticəyə daha çox güvənmək lazım olduğunu əsaslandır.

#### Tapşırıq 10b:

Verilmiş zülal ardıcılığını **FASTA formatında** saxlayın və Biopython ilə bu faylı oxuyaraq **zülalın funksiyasını** müəyyən etmək üçün FASTA başlığını təhlil edin. Başlıqda "**description**" hissəsində məlumatları çıxarın.

*İstifadə:* SeqRecord.description, SeqIO.parse()

#### Tapşırıq 11a:

Sənə təqdim olunan qısa bir nukleotid ardıcılığını BLASTN və BLASTX ilə analiz et. Aşağıdakı suallara cavab ver:

1. Bu ardıcılıq hansı orqanizmə aiddir?
2. Hansı gen və ya zülalə uyğun gəlir?

### Tapşırıq 11b:

Verilmiş bir **zülal ardıcılığı** üzərində Biopython istifadə edərək **motif** (konkret bir ardıcılıq) tapın və həmin motifin ardıcılıqda neçə dəfə təkrarlanmasını çap edin.

*İstifadə:* Seq.find(), Seq.count()

### Tapşırıq 12a:

**Zülalın kimə məxsus olduğunu tapın.**

Aşağıdakı **amin turşusu ardıcılığını** (protein) **BLASTP** istifadə edərək analiz et və suallara cavab ver:

>Unknown\_protein

MKWVTFISLLFLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQKHKI  
PVKYSVAV

**Cavablandırılacaq suallar:**

1. Bu zülal **hansı orqanizmə** aiddir?
2. Zülalın **adı və funksiyası** nədir?
3. BLAST nəticəsində çıxan **identity (%)**, **alignment uzunluğu** və **E-value** nədir?
4. Bu zülalın **bioloji əhəmiyyətini** 2-3 cümlə ilə yaz.

### Tapşırıq 12b:

Verilmiş bir **nukleotid ardıcılığını** **tBLASTN** ilə uyğunlaşdırın və ən yaxşı uyğunlaşmanın **genetik mənşəyi** və **oxşar genlər** haqqında məlumatları çıxarın.

*İstifadə:*

- **Bio.Blast.NCBIWWW.qblast()** - tBLASTN sorğusu göndərmək üçün
- **Bio.Blast.NCBIXML.parse()** - nəticələri təhlil etmək üçün