

GenBank және геномдық дерекқорлар

СНЕ9131 ДНҚ технологиялары · 6B05105 Биотехнология

9-апта · Жаниязов Ж.А., аға оқытушы

Дәріс жоспары

- GenBank туралы жалпы мәлімет
- Жазбаның құрылымы
- Іздеу және деректерді жүктеу
- BLAST және FASTA форматы
- Басқа биоинформатикалық қорлар

9 / 15

апта • GenBank

GenBank дегеніміз не?

1982

GenBank құрылған
жыл

NCBI

АҚШ ұлттық
биотехнология
ақпарат орталығы

3

INSDC қоры: GenBank,
ENA, DDBJ

Ашық

тегін қолжетімді

Үш қор деректерімен күнделікті алмасады, сондықтан мазмұны бірдей.

GenBank жазбасының құрылымы

Өріс	Мазмұны
LOCUS	Атауы, ұзындығы, молекула түрі
DEFINITION	Қысқаша сипаттама
ACCESSION	Бірегей нөмір
ORGANISM	Ағзаның атауы мен жіктелуі
FEATURES	Гендер, CDS, экзондар, реттеуші аймақтар
ORIGIN	Нуклеотид тізбегі

NCBI қорларының түрлері

Nucleotide

ДНҚ және РНҚ тізбектері

Protein

Аударылған ақуыз тізбектері

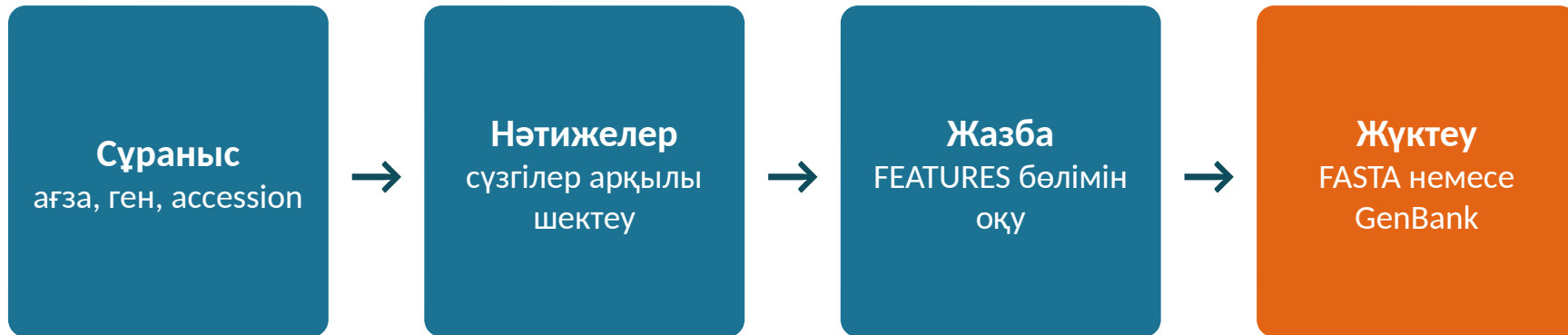
Gene

Ген туралы біріктірілген мәлімет

SRA

NGS-тің шикі оқулары

Деректерді іздеу



Нақты іздеу үшін accession нөмірін немесе гендің атын қолданыңыз.

BLAST: ұқсастықты іздеу

- Берілген тізбекті қордағы тізбектермен салыстырады
- blastn — нуклеотидтер, blastp — белоктар, blastx — аударылған тізбек
- Identity — сәйкес келетін позициялар үлесі
- E-value неғұрлым кіші болса, сәйкестік соғұрлым сенімді

E-value

кездейсоқтық бағасы

FASTA форматы

- Бірінші жол «>» белгісінен басталады: идентификатор мен сипаттама
- Келесі жолдарда тізбек жазылады (A, T, G, C)
- Көптеген бағдарламалар осы форматты қабылдайды
- Бір файлда бірнеше тізбек болуы мүмкін



FASTA тақырыбы

Басқа дерекқорлар

Ensembl, UCSC

Геном браузерлері: гендер, нұсқалар, аннотация

UniProt

Ақуыз тізбектері мен қызметтері

PDB

Ақуыздар мен нуклеин қышқылдарының 3D құрылымы

ENA, DDBJ

GenBank-пен ұқсас халықаралық нуклеотид қорлары

Қорытынды және бақылау сұрақтары

- GenBank жазбасы қандай бөлімдерден тұрады?
- Accession нөмірі не үшін керек?
- BLAST-та E-value нені көрсетеді?
- Тапсырма: ПР №9, СӨЖ: GenBank туралы презентация

141–159

бет • [1] Brown