

**O‘ZBEKISTON RESPUBLIKASI OLIY VA  
O‘RTA MAXSUS TA‘LIM VAZIRLIGI**

**GULISTON DAVLAT UNIVERSITETI**

**BIOLOGIYA KAFEDRASI**



**“BIOINFORMATIKA”**

**Fanidan**

**O‘QUV - USLUBIY MAJMUA**

Bilim sohasi: 100 000 – Gumanitar soha

Ta‘lim sohasi: 140000 – Tabiiy fanlar

Ta‘lim yo‘nalishi: 5140100 – Biologiya (Turlar bo‘yicha)

**Guliston – 2023**

Fanning ishchi o'quv dasturi Guliston davlat universitetining 2023 yil "\_\_\_ avgustdagi "1"-sonli bayonnomasi bilan tasdiqlangan "Bioinformatika" fani dasturi asosida tayyorlangan.

**Tuzuvchilar:**

H.Ro'ziboyev,  
Tursunov M.M.

Biologiya falsafasi Doktori PhD  
Boiligiya kafedrası o'qituvchisi

**Taqrizchilar:**

**Ergashev M.M.**

Biologiya kafedrası dotsenti.

## MUNDARIJA

Soʻz boshi .....	4
Fanni oʻqitishdan maqsad va vazifalar.. .....	4
Nazariy materiallar (maʼruzalar kursi)... .....	6
Amaliy ishlarni bajarish boʻyicha uslubiy koʻrsatmalar.....	167
Mustaqil taʼlim boʻyicha materiallar .....	270
Golossariy .....	278

## ILOVALAR

### Fan dasturi

Ishchi fan dasturi.....	281
Test savollari.....	292
Adabiyotlar.....	295

## SO‘Z BOSHI

**Bioinformatika** – fani biologik tizimlarda axborotni saqlash va uzatishning nazariy masalalari bilan shug‘ullanadigan, informatikaning jadal rivojlanayotgan sohasi hisoblanadi. Bioinformatikaning asosiy tarmoqlari DNK, RNK nukleotidlar ketma- ketligida saqlanadigan genetik “matnlarni” ochish muammosini hal qiladigan kompyuter genomikasi va hujayra metabolizmining tashkil etilishi, uning genom tomonidan boshqarilishini o‘rganadi. Genomika va metabolomika fan sohalarining rivojlanishi uchun zarur bo‘lgan eksperimental ma’lumotlar bilan yetarli miqdorda va qulay shaklda ta’minlaydigan kompyuter ma’lumotlar bazalarini yaratish bioinformatikaning asosiy bo‘limlarini rivojlantirish uchun katta ahamiyatga ega.

### I. Fanni o‘qitishdan maqsad va vazifalar:

**1.1.Fanning maqsadi:** Bioinformatika biologik tizimlar asosida olingan bilimlarni to‘plash, saqlash va ulardan foydalanishni ta’minlaydigan axborot xizmatlarini o‘z ichiga oladi. Shuning uchun bioinformatikaning bosh maqsadi biologik bilimlarni samarali foydalanishni ta’minlaydigan shaklda to‘plash, biologik tizimlar elementlarini matematik modellarini qurish va tahlil qilishdan iborat. Organizmning faoliyatini ta’minlaydigan moddiy elementlarning tuzilishi haqidagi ma’lumotlar uning genomini tashkil etuvchi DNK yoki RNK nukleotidlar ketma-ketligida saqlanadi.

Organizmlar genomlari DNKsining nukleotidlar ketma-ketligini yaratish (sekvenirlash) 21-asrning boshlarida rivojlangan va ancha tejamli va samarali texnologiyaga aylandi. Aniqlangan genom ketma-ketliklari soni tez sur‘atlarda o‘sib bormoqda. Bioinformatikada genomika deb nomlangan maxsus bo‘lim mavjud bo‘lib, uning mavzusi DNK va RNK ketma-ketligida kodlangan biologik tizimlarning asosiy moddiy elementlarining tuzilishi haqidagi ma’lumotlarni saqlash usullarini modellashtirish va o‘rganishdir. Genomika sohasida ishlaydigan butun dunyo olimlarining asosiy sa’yi-harakatlari hozirda hujayra genomining nukleotidlari ketma-ketligi bo‘lgan genetik “matnlarni” kompyuterda tahlil qilishning samarali usullarini ishlab chiqishga qaratilgan. Genetik matnni tahlil qilish, DNK ketma-ketligining funksiyalarini aniqlashni anglatadi ya’ni, ularning boshqaruvchi va oqsil kodlovchi mintaqalari va genlar faoliyatini tartibga solish va muvofiqlashtirishni ta’minlovchi mintaqalar o‘rganiladi. Hozirda eng dolzarb masala - bu inson genlarini hosil qiluvchi nukleotidlar ketma-ketligini, Xalqaro inson genomi loyihasiga muvofiq, oxirga yillarda inson genomining to‘liq ketma- ketligini aniqlash, izohlash hisoblanadi. Aytish joizki, DNKning nukleotidlar ketma-ketligi bo‘yicha aniq funksiyalari bo‘lgan mintaqalarni ajratish oson ish emas, chunki ular tabiatan ancha xilma-xildir. Hozirgi vaqtda nukleotidlar ketma- ketligi bo‘yicha genlarni kompyuter usullari bilan aniqlash ehtimolligi 70% dan oshmaydi. Ammo bugungi kunda bioinformatik tadqiqot ishlari



hujayradagi metabolism jarayonlarini o'rganishga ko'plab bag'ishlangan. Hujayradagi metabolism jarayonlarini modellashtirish va biologik tizimning moddiy elementlarining birgalikdagi faoliyatini o'rganishdan iborat bo'lgan bioinformatikaning tegishli bo'limini metabonomika deb atash mumkin.

**1.2.Fanning vazifasi:** Bioinformatika fanining asosiy vazifasi bu biologik molekulalar, eng avvalo nuklein kislotalar va oqsillarning strukturalari, funksiyalari bo'yicha ma'lumotlarni tahlil qilish va tizimlashtirish uchun hisoblash algoritmlarini ishlab chiqishdir. DNK nukleotid ketma-ketliklarini sekvenirlashning jadal usuli ishlab chiqilgandan so'ng ma'lumotlar bazasida to'planayotgan genetik axborotlar hajmi yuqori tezlik bilan orta boshladi. Informatika, lingvistika va informatsiya nazariyasi yutuqlari genetik matnlarni tahlil qilish imkoniyatlarini ochib berdi. Bioinformatikaning boshqa fan sohalari bilan o'zaro bog'liq holdagi rivojlanishi organizm va hujayrada yuz berayotgan biologik jarayonlarni tushunishning yangi darajasi shakllantirishga imkon beradi.

Amaliy ma'noda bioinformatika - bu biologlar manfaatlari uchun xizmat qiladigan amaliy fandır. Ma'lumotlarni birlamchi tahlil qilish texnik bioinformatika sohasiga tegishlidir. Olingan ma'lumotlarni qayerdadir saqlash va ulardan foydalanish imkoniyatlarini ta'minlash zarurati mavjud, jumladan B geni qaysidir jarayonda qatnashadi va hokazo, bu esa bioinformatika fanining amaliy ahamiyatidan dalolat beradi.

**1.3.Talabalar bilimiga qo'yiladigan talablar:** Talabalar quyidagi bilimlar bilan qurollangan bo'lishlari lozim:

-bioinformatika fanlarini chuqur o'zlashtirgan bo'lishlari;

-tirik organizmlar va ularning ichki tuzilishining murakkab ko'rsatkichlarini dasturlash to'g'risida bilimlarga ega bo'lishlari lozim:

-bioinformatik dasturlardan yaxshi foydalanishi va ularni tadqiqotlarda qo'llay olishi lozim:

**1.4. Fanni o'rganishda:** talabalar bir nechta biologiya sohasidagi fanlar-bioximya, biofizika, molekulyar genetikaning bir qancha yo'nalishlar, shuningdek matematika, kompyuter dasturlari bo'yicha amaliy ko'nikmalarga ega bo'lishlari lozim: Bioinformatika fani eng rivojlangan davlatlar AQSH, yevropa va boshqa davlatlarda yuqori darajada rivojlangan. Shuni hisobga olib, ushbu fanni o'zlashtirishda chet tili (ayniqsa ingliz tili) fanlarini bilish muhim ahamiyatga ega.

## Nazariy materiallar (ma'ruza kursi)

## 1-mavzu: Kirish.

1. Bioinformatikaning fan sifatida shakllanish tarixi. Uning predmeti, vazifalari va ob'ektlari.
2. Zamonaviy biologik tadqiqotlarda bioinformatikaning ahamiyati.
3. Bioinformatika rivojlanish bo'chqichlari va yutuqlari.

**Mavzuga oid tayanch tushunchalar va iboralar:** *bioinformatika, sekvenirlash, genomika, proteomika, DNK va oqsil ketma-ketliklari. Ma'lumotlar bazalari*

**Dars maqsadi:** Talabalarga bioinformatika fanining kelib chiqishi, rivojlanish bosqichlari, biologiya sohasidagi ro'li haqida umumiy tushuncha berish. Bundan tashqari bu fanni o'qitishdan maqsadi va vazifalarini aniq belgilab berish hamda kelajakda biologiya sohasida ham raqamli texnologiyalardan foydalanishni o'rganish uchun fundamental hamda amaliy bilimlarga ega bo'lishini taminlash.

### **Birinchi savolning asosiy bayoni.**

Informatika fanining XX asming ikkinchi yarmida paydo bo'lgan davrdan boshlab fizika-matematika, texnika, gumanitar va boshqa fanlarga ham tadbiiq qilinishi hamda ular bilan hamkorlikda ishlashi tobora kengayib bormoqda. Hozirgi kunda informatika fani usullarini chetlab o'tadigan biron-bir fan sohasini topish mushkul. Tabiiy fanlar ham bundan mustasno emas. O'tgan asming 60- yillar oxiri 70-yillar boshlarida biologiyada EHM -elektron hisoblash mashinalari faol qo'llanila boshlandi: Shu bilan birgalikda ularning xotiralari, operatsion tezliklari oshdi, o'lchamlari kuchraytirildi hamda biologiya sohasida information tahlillarni talab etuvchi katta miqdordagi eksperimental ma'lumotlar to'planib qoldi. Bunga misol qilib bir qancha davlat olimlari hamkorligida 2003 yildayoq odam genomining sekvenirlanishini bo'yicha olingan ma'lumotlarni keltirish mumkin. 90-yillarda genomika fani paydo bo'la boshladi. Hozirgi kunga kelib bir qancha organizmlar, jumladan odam, sichqon, tovuq, qurbaqa, bir qancha baliq turlari, chuvalchanglar, yuzlab viruslar va bakteriyalar hamda yuzlab o'simlik turlarining genom ketma-ketliklari aniqlangan. Bakteriya genomining o'qilishi biologlar uchun Mendeleevning ximiklar uchun yaratilgan davriylik qonunini ochish bilan tenglashtiriladi. Shu boisdan ham bunday hajmdagi biologik ma'lumotlarni tahlil qilishda kompyuter texnologiyasidan foydalanila boshlandi. Gen ketma-ketliklarini tenglashtirish bo'yicha birinchi algoritm 1970 yilda yaratildi. Kompyuterlar axborotlarni virtual ma'lumotlar bazasida saqlash va ular ustida yuqori tezlikda operatsiyalar o'tkazish imkonini berdi.

Bioinformatika ham boshqa zamonaviy fanlar singari bir qancha fanlar, ya'ni molekulyar biologiya, genetika, matematika va kompyuter texnologiyalari fanlari birlashuvi asosida vujudga keldi. Uning asosiy vazifasi bu biologik molekulalar, eng avvalo nuklein kislotalar va oqsillarning strukturalari, funksiyalari bo'yicha

ma'lumotlarni tahlil qilish va tizimlashtirish uchun hisoblash algoritmlarini ishlab chiqishdir. DNK nukleotid ketma-ketliklarini sekvenirlashning jadal usuli ishlab chiqilgandan so'ng ma'lumotlar bazasida to'planayotgan genetik axborotlar hajmi yuqori tezlik bilan orta boshladi. Informatika, lingvistika va informatsiya nazariyasi yutuqlari genetik matnlarni tahlil qilish imkoniyatlarini ochib berdi. Bioinformatikaning boshqa fan sohalari bilan o'zaro bog'liq holdagi rivojlanishi organizm va hujayrada yuz berayotgan biologik jarayonlarni tushunishning yangi darajasi shakllantirishga imkon beradi.

Bioinformatikaning yaralish tarixi XIII asrlarga borib taqaladi. Matematika tarixiga Fibonachchi (Fibonacci) nomi bilan kirib kelgan yosh italyan Pizalik Leonardo Leonardo of Pisa (1-rasm) biologik jarayonning birinchi matematik modelini tuzgan holda quyonlarnig ko'payishi to'g'risidagi masalani tavsiflab bergan.

XX asrning 20 yillariga kelib esa yana bir italyan olimi Vito Volterra -Vito Volterra "yirtqich-o'lja" ko'rinishidagi ikki biologik turning o'zaro harakati modelini yaratdi. 40 yillar oxirida biologiyaga fizik va matematiklar kirib kela boshladi. Biologiyaning zamonaviy tarixi 1953 yildan, Amerika olimlari Jeyms Uotson hamda Frensis Krik tomonidan DNK ning qo'sh spiralligi kashf qilingan davrdan boshlandi (2-rasm.). 1950-yillardan shu kunga qadar bazalar ma'lumotida ko'p miqdordagi turli xil organizmlarning nukleotidlari qatori ma'lumotlari yig'ilgan bo'lib, ularni oddiy hisoblash usullarida taqqoslash va tahlil qilish juda murakkab bo'lgan.

Bioinformatika fanining mustaqil fan sifatida shaklanishining asosiy xronologik sanalari sifatida qo'yidialarni ko'rsatish mumkin.

1866 yil - Gregor Mendel tomonidan no'xat o'simligida irsiy omillarnig irsilanishi to'g'risidagi tajribalarini nasrda e'lon qilinishi.

1928 yil - Ervin Shyodinger bunday omillar 1000 angstromga teng bo'lishi to'g'risidagi fikrlarini aytgan.

1933 yil - A.Tizelius eritmadagi oqsil aralashmalarini elektroforetik ajratish usulini taklif etgan.

1951 yil - L.Ploing va R.Kyuri a-spiral va P-barg strukturasi hosil qiluvchi



*1-rasm. Pizalik Leonardo  
(1170-1250)*

polipeptid zanjirlarining struktur modelini taklif qilishgan.

1952 yil - R. Franklin va M. Wilkins rentgenostruktur tahlil yordamida DNK reguliyalar strukturasi xarakterini aniqlashgan.

1953 yil - J. Watson va F. Crick DNK ikkilamchi strukturasi taklif qilishgan.

1954 yil - M. Perutz va u boshqargan ilmiy guruh xodimlari oqsillar kristallografiyasida fazalar muammosini, og'ir atomlarning o'rin almashinishi usuli yordamida aniqlashgan.

1955 yil - F. Sanger qoramol insulin oqsilining ketma-ketligini aniqlaydi.

1957 yil - Artur Kornberg ilk bor sintetik DNK molekulasini yaratgan.

1965 yil - Margaret Deyhoff ilk bor "Biotibbiyot tadqiqotlar milliy fondi" (BTMF), Vashington, xodimlari bilan hamkorlikda birinchi marotaba oqsil ketma-ketliklarini ma'lumotlar bazasiga birlashtirgan.

1968 yil - Verner Arber, Hamilton Smit va Daniel Nathans restriktazalar ishlash prinsipini aniqlashgan.

1969 yil - Los Angelesda "Stanford universiteti" va "Kaliforniya universiteti" tomonidan kompyuterlarni birlashtirish tizimining yaratilishi "ARPAnet" tarmog'ining yaratilishiga olib elgan.

1970 yil - Nidman - Vunshning ketma-ketliklarni taqqoslash maqsadida algoritim tavsifi beriladi.

A. D. J. Gibbs va G. A. Makintayer aminokislotalar va nukleotid qatorlari uchun nuqtali matrisa yordamida aniqlash uslubini yoritib berishgan.

1972 yil - Paul Berg, ligaza yordamida birinchi sun'iy rekombinant molekulasini konstruksiyasini yaratgan.

Stanley Koen, Enni Chan, Gerbert Boyer rekombinant RNKli birinchi organizmni yaratishadi.

1973 yil - Djozef Sembruk o'zining ishchi guruhi bilan birgalikda DNK elektroforezi uslubini, agaroz gelini qo'llab mukammallashtirishadi.

Stanley Koen DNK klonlashni amalga oshiradi.

"Brukheyven oqsil bazalar ma'lumoti" yaratiladi.

Robert Metcalf o'zining doktorlik dissertatsiya ishida "Ethernet" tarmog'iga ta'rif beradi.

1974 yil - Vint Karf va Robert Kan kompyuter tarmoqlarini "Internet" global tarmog'iga birlashtirish konsepsiyasini rivojlantirishdi va ma'lumotni uzatishning protokollarini ishlab chiqishdi - Transmission Control Protocol, TCP.

1975 yil - P. H. O'Farrell poliakrilamid gelida dodesil sulfat ishlatgan holda ikki o'lchamli elektroforez uslubini yaratdi.

Edvard Saucer-Saucer blot tahlilini yoritib bergan.

Bill Geyts va Pol Allen "Maykrosoft" - Microsoft Corporation, korporatsiyasiga asos solishadi.

1977 yil - Frederik Senger, Allen Maksam va Uolter Gilbert DNK sekvenlash uslubini ishlab chiqishdi.

1979 yil - “Los Alamos milliy laboratoriyasida” Nyu Meksiko shtatida, Uolter Goud xamkasblari bilan birgalikda, ilk bor “GenBank” ma’lumotlar bazasining prototipi asosida DNK ma’lumotlar bazasini birashtirishdi.

1980 yil - Mark Skolnik, Rey Uayt, David Botshteyn va Ronald Deyvis inson genomining PDR-marker (Pestriktazalar uzunligi polimorfizmi) xaritasini yaratdilar.

“FX-174” organism genlarining to‘liq ketma-ketligi aniqlandi.

Ko‘p o‘lchamli yadroli magnitli rezonans usuida Vutix oqsil struktirasini aniqlash ishlarini amalga oshirdi

Kaliforniyada “IntelliGenetics Inc.” kompaniyasi tashkil etildi va uning birinchi mahsuloti “IntellGenetics Suit” oqsillarni aniqlash va DNK ketma- ketliklarini aniqlaydigan va tahlil qiladigan dasturlar paketi ishlab chiqildi.

Smit-Uotermanning ketma-ketliklarni tahrirlash algoritmi yaratildi.

AQSH Oliy sudi suniy modifisirlangan bakteriya patentini tan oldi.

1981 yil - IBM korporasiyasi bozorga personal kompyuterlarni chiqardi.

Inson mitoxondrial genomi sekvens qilindi.

D.Benson va LD.Lipmen tomonidan “GENINFO” - menyu orqali ketma-ketliklar ma’lumotlar bazasini boshqaradigan dasturni ishlab chiqishdi.

Mayzel va Lenk rangli tasvir va sxemalar filtrasi dasturini ishlab chiqishdi va nuqtali matrisalar uslubini ishlash prinsipi soddalashtirilar.

1982 yil - rekombinant DNK asosida yaratilgan dorivor vosita ilk bor bozorga chiqarildi.

Viskonsin biotexnologiya markazida “genetics Computer Group, GCG”, axborot bo‘limi tashkil etildi.

1983 yil - Sotuvda CD kompakt diskar chiqarildi.

1984 yil - Djon Postelning “Internet” tarmog‘iga SID nomli domen tizimlari joylashtirildi.

1985 yil - Tomas Roderik genlarni xaritalash, sekvenlash va tahlil qilish muammolari bilan shug‘ullanadigan aloxida ilmiy fanni “Genomika” deb atalishini taklif etdi.

Jeneva universiteti tibbiy biokimy bo‘limi va Evropa molekulyar biologiya laboratoriyasi bilan xamkorlikda SwissPROT ma’lumotlar bazasi yaratildi.

Leroy Hud va Lloyd Smit DNK ni sekvenlash usulini avtomatlashtirildi.

1986 yil - AQSH lari tabiatni muhafaza qilish vazirligi “Inson genomi” loyihasini rasmiy ishga tushurganligini e’lon qildi.

1987 yil - Devid T.Byork hammualliflar bilan birgalikda, zamburug‘lar sun’iy xromosomasini ishlatish uslubini yoritdilar.

Pirson va Lipmen FASTA algoritmini e’lon qilishdi.

AQSHlari “Rak milliy institute”da “Biotexnologik ma’lumotlar milliy markazi” tashkil etildi.

1988 yil-AQSH lari Sog‘lomlashtirish milliy institute - NIH “Inson genomini tadqiqot qilish milliy markazini” tashkil etdi.

Oxford Molecular Group Ltd., OMG, “Anakonda”, “ASP”, “Cameleon” dasturlarini ishlab chiqishdi va oqsil strukturalarini ishlab konstruksiyalaydigan, dorivor vosiyalarni molekulyar modellashtiraigan va ushalb chiqadigan dasturlarni yaratdilar.

1989 yil - Altshtul bir qancha dasturchilar bilan birgalikda, BLAST algoritmini ishlab chishdi.

1990 yil - “CERN” Jeneva, tomonidan “Xalqaro o‘rgimchak to‘ri” ning boshlang‘ich ishlariga asos solishganin e‘lon qilishdi.

Kreyg Venter ekspressiyalanadigan ketma-ketliklar yarliklari yordamida genlarni aniqlash texnologiyasini yaratdilar.

Kaliforniyada “Insyte Pharmaceuticals” farmasevtik genomika bilan shug‘ullanadigan kompaniya tashkil etildi.

AQSH larining Yuta shtatida Myriad genetics Inc. kompaniyasi tashkil etildi va uning asosiy faoliyati asosiy kasalliklar genini aniqlash va ularning irsiylanishini o‘rganishdan iborat.

1993 yil-Frensis Kollins “Inson genomi” loyihasini boshqardi. Buyuk Brutaniyada “Senger markazi” tashkil etildi, loyihaga boshqa davlatlar ham qo‘shildi. Loyiha faoliyatini 2005 yilda tugatish rejalashtirildi.

1994 yil - Ettvud va Bek PRINTS oqsillar motivi ma’lumotlar bazasini yaratishdi.

Merilend shtatida Gene Logic kompaniyasi tashkil etildi.

1995 yil - “TIGR” kompaniyasi olimlari erkin yashaydigan *Haemophilus influenzae* organism genom ketma-ketliklarini aniqlashdi.

1996 yil - Achitqi zamburug‘i *Saccharomyces cerevisiae* genomi sekvenlandi.

PROSITE ma’lumotlar bazasi ishlab chiqildi.

Affimetrix kompaniyasi birinchi komersiya DNK chiplarini ishlab chiqishdi.

1998 yil - *Caenorhabditis elegans* chuvalchangi va *Saccharomyces cerevisiae* genomi taxrirlandi.

AQSH lari NIH - National Institute of Health - instituti tomonidan “PON” - aloxida oligan nukleotidlar polimorfozmi loyihasiga asos solindi, ushbu loyiha inson genomida o‘tayotgan o‘zgarishlarni tadqiqot etishga qaratilgan.

1999 yil-Inson birinchi xromosomasining ketma-ketligi e‘lon qilindi.

2000 yil - *Pseudomonas aeruginosa*, *Arabidopsis thaliana* va *Drosophila melanogaster* genomlari sekvenlandi.

2001 yil-fevralda “Science” va “Nature” jumallarida inson genomi va uning

tahlillari annotasiyalari e'lon qilindi.

2002 yil - boshqa organizmlar genom ketma-ketliklari e'lon qilindi.

Bioinformatika shuningdek genetika, molekulyar biologiya, biofizika, ekologiya, va qator tabiiy fanlar sohalarida foydalaniladi. Bioinformatika o'z ichiga qo'yidagilarni oladi:

S qiyosiy genomikada kompyuter tahlilining matematik usullari (genom bioinformatikasi);

S oqsil strukturalarini bashorat qilish uchun algoritm va dasturlarni ishlab chiqish (strukturaviy bioinformatika);

S muvofiq hisoblash uslubiyatlari strategiyasi tadqiqoti hamda informatsion murakkablikning biologik tizimlar tomonidan umumiy boshqarilishi.

Amaliy ma'noda bioinformatika - bu biologlar manfaatlari uchun xizmat qiladigan amaliy fandır. Ma'lumotlarni birlamchi tahlil qilish texnik bioinformatika sohasiga tegishlidir. Olingan ma'lumotlarni qayerdadir saqlash va ulardan foydalanish imkoniyatlarini ta'minlash zarurati mavjud, jumladan B geni qaysidir jarayonda qatnashadi va hokazo, bu esa bioinformatika fanining amaliy ahamiyatidan dalolat beradi.

Bioinformatika biologiya sohasining qo'yidagi yo'nalishlarida qo'llaniladi:

- genomika, transkriptomika va proteomika;
- rivojlanish biologiyasida kompyuter modellashtirish;
- gen tarmoqlarining kompyuter tahlili;
- populyatsion genetikada modellashtirish.

Bioinformatika dori preparatlarini loyihalashtirish muddatini 5-6 yildan bir necha oylarga qisqartirish imkoniyatini yaratib farmakologiya sohasiga ham osongina kirib bordi. Shuningdek, bu fan ko'plab boshqa tibbiyotga va biologiyaga oid fanlar bilan integratsiyalandi. Bugungi kunda bioinformatikaning qo'yidagi bo'limlari mavjud:

- umumiy bioinformatika;
- klinik bioinformatika;
- strukturaviy genomika;
- funktsional genomika;
- farmakogenomika;
- klinik proteomika;
- funktsional proteomika;
- strukturaviy proteomika.

Bioinformatika usullari yordamida katta hajmdagi biologik ma'lumotlarni shunchaki tahlil qilish emas, balki har doim ham oddiy tajribalarda aniqlab bo'lmaydigan qonuniyatlarni isbotlash, genlar va ular kodlaydigan oqsillar funksiyalarini bashorat qilish, hujayradagi genlarning o'zaro ta'siri modelini qurish,

dori preparatlarini yaratish mumkin. Bu ma'lumotlar oqsil ketma- ketliklarini va regulyator uchastkalarini aniqlash uchun foydalaniladi. Ma'lumotlar miqdorining ko'payishi bilan endi ketma-ketliklarni qo'lda tahlil qilish mumkin bo'lmay qoldi. Hozirgi kunda milliardlab juft nukleotidlardan tashkil topgan minglab organizmlar genomlari bo'yicha qidiruvlar olib borish uchun kompyuter dasturlaridan foydalaniladi.

Genomika kontekstida annotatsiya - bu DNK ketma-ketligida genlarni va boshqa ob'ektlarni markirovkalash nishonlash jarayonidir. Genomlar annotatsiyasi birinchi dasturiy tizimi Owen White tomonidan 1955 yildayoq yaratilgan edi. Bu orqali evolyutsion biologiya turlarning kelib chiqish va paydo bo'lishini, ularning davrlar bo'yicha rivojlanishini o'rganadi. Informatika evolyutsiyani o'rganuvchi biologlarga bir necha jihatlarida yordam berdi:

- 1) barcha DNK dagi o'zgarishlarni o'rgangan holda ko'p sonli organizmlar evolyutsiyalarini tadqiq qilishda;
- 2) yanada kompleks evolyutsion hodisalarni o'rganish imkonini beruvchi genomlarni bir-biriga taqqoslashda;
- 3) populyatsiyalar kompyuter modellarini qurishda;
- 4) ko'p miqdordagi turlar haqida ma'lumotni o'z ichiga oluvchi nashrlarni sifatida aniqlanishi mumkin. Ixtisoslashtirilgan dasturiy ta'minot mahsulotlari qidirish, vizualizatsiya qilish, axborotni tahlil qilish va eng muhimi, natijalarni boshqa tadqiqotchilar bilan bo'lishda foydalaniladi.

Bundan tashqari matematik biologiya ham mavjud bo'lib, u ham bioinformatika singari biologik muammolarni yechishda ishlatiladi, biroq unda qo'llaniladigan usullar natijasi son bilan ifodalanmaydi va ularni amalga oshirishda dasturiy va jihoz ta'minoti talab etilmaydi. Oqsillar fazoviy tuzilmalarini bashorat qilishda ishlatiladigan algoritm va dasturlar ishlab chiqish bilan shug'ullanuvchi srukturaviy bioinformatika boshqalaridan ajralib turadi.

### **Ikkinchi savol bayoni.**

#### **Zamonaviy biologik tadqiqotlarda bioinformatikaning ahamiyati.**

Zamonaviy biologik tadqiqotlarda bioinformatikaning ahamiyati. Bioinformatika biologiyaning ilmiy tajribalari asosida olingan natijalarni tahlil qiladi. Olingan ma'lumotlarni tadqiqotchi ma'lumotlar bazasida mavjud bo'lgan barcha to'plamlar bilan solishtiradi. Bordini, u o'zi aniqlagan ketma- ketlikni ma'lumotlar bazasidan topa olmasa bunda u bu ma'lumotni shu joyga kiritib qo'yadi va bu bilan bazani yanada boyitadi. Ma'lumotlar bazasi funksiyalariga saqlash, tizimlashtirish, axborotlarni yangilab turish unga kirish huquqi bilan ta'minlashlar kiradi. Bu operatsiyalar esa katta qudratlardagi kompyuterlarni talab qiladi. Shuningdek biologik mavzular majmuidagi ilmiy nashriyotlar bazalari ham mavjud. Biologiya bo'yicha istalgan ilmiy jurnalning barcha sonlarida chiqadigan har bir maqola ma'lumotlar bazasiga joylashtiriladi



izlanuvchi uni internet tarmog'i orqali oson topib olishi uchun qisqa ta'rif berib qo'yiladi.

Integral ma'lumotlar bazasi va ensiklopediyalar konkret gen, oqsil, funksiyalarni amalga oshishi haqida ma'lumot beradi. Ular katta miqdordagi boshqa ma'lumotlar bazalari axborotlarini umumlashtiradi va uni hamisha yangilab turadi. Har qanday yangidan o'qilgan genom harflaming turli xil kombinatsiyalarida takrorlanuvchi ulkan ketma-ketliklar ko'rinishida namoyon bo'ladi. Bioinformatika bunday xilma-xillikdagi matndan genlarni ajratib olish imkoniyatini beradi.

Genomdan genni ajratib olish kabi bunday operatsiya genomni belgilash deb ataladi. Barcha genlar funksiyalarini tajribalar asosida aniqlash yetarli darajada murakkablikni yuzaga keltiradi. Bu holatda bioinformatika funksiyalari allaqachon aniqlangan genlar bilan solishtirib ko'rishga tayangan holda ularni bashorat qilishda ko'maklashadi. Oqsil molekulasida biologik vazifalarning har xil turlariga javob beruvchi uchastkalar mavjud. Bioinformatika usullari yordamida ushbu uchastkalarni aniqlash konkret bir oqsilning barcha spektr funksiyasini ochib beradi. Oqsil strukturalarini tajribalar asosida, ya'ni masalan oqsil molekulalaridan tashkil topgan mikroskopik kristalni rentgen nurlari bilan nurlantirish orqali aniqlash mumkin. Bu esa yetarli darajada uzoq va qimmatli jarayon hisoblanadi. Ayrim oqsillar kristall tuzilmalarga ega bo'lmaganligi sababli ularni tahlil qilishning umuman iloji yo'q. Bioinformatika kompyuter modellashtirish yordamida hech bo'lmaganda oqsil strukturasini uzoqroq o'xshash ketma-ketligi ma'lum bo'lgan holatlarda oqsilning fazoviy modelini yasashda yordam beradi. Bioinformatika metodlari asosida olingan molekulaning fazoviy strukturasini bilgan holda uning qanday ishlashini va uning ishlashiga qanday ta'sir eta olishni bashorat qilish mumkin. Dori preparatlarini fazoda har xil ximiyoviy bog'lanishlar bilan oqsil-nishonlarning o'zaro ta'sirini modellashtirish asosida tayyorlash mumkin. Bunda katta miqdori bog'lanishlarni saralash va eng maqbullarini tanlab olish kerak bo'ladi.

Bioinformatika texnologiyalaridan foydalanib qilingan biologiya sohasidagi yangi kashfiyotlar tez suratda tibbiyot, farmakologiya, kosmetologiya, biotexnologiya, qishloq xo'jaligi, ekologiya va boshqa sohalarda jalb qilinadi. Bioinformatika mustaqil ravishda amaliy ahamiyatga ega bo'lgan natijalar beradi va shuningdek biologiyaning turli sohalarida ishlash uchun sharoit bilan ta'minlaydi. Bioinformatika bo'yicha ishning katta qismi biologik axborotni saqlash va uni tahlil qilish uchun ma'lumotlar bazasidan foydalanishtexnologiyalari atrofiga jamlangan. Bunday ma'lumotlar bazasi ommabop yoki shaxsiy bo'lishi mumkin. Ularga ochiq standartlar orqali ommaviy kirish huquqini olish esa muhim ahamiyat kasb etadi. Garchi ma'lumotlar bazasidan foydalanishga nisbatan bu usullar anchagina keng tarqalgan bo'lsada biologik axborotlarni tahlil qilish uchun ontologiya va mantiqiy usullardan foydalanish rivojlanib bormoqda.

### **Uchinchi savol bayoni.**

**Bioinformatikaning rivojlanishi bosqichlari va yutuqlari.** Bir qancha xorijiy davlatlarda 20-21 asrlarda bioinformatika jadal suratda rivojlanayotgan dunyo biotibbiyot fanlari sohasiga aylanib bordi. Bioinformatcion texnologiyalar iste'molchilari tadqiqotchilar, fundamental ishlanmalar mualliflari bilan bir qatorda tibbiyot, farmakologiya, biotexnologiya hamda o'quv muassasalari hisoblanadi. Fanning bu sohasi AQSHda va shuningdek boshqa rivojlangan davlatlarda muhim yo'nalish sifatida qaraladi.

Yevropa, Osiyo, AQSH hamda Avstraliya davlatlarida bioinformatika markazlari soni yildan-yilga ko'payib bormoqda. Bioinformatika bo'yicha yuqoridagi davlatlarda, akademiya hamda ta'lim markazlari bilan bir qatorda so'nggi yillarda sohada olingan tadqiqot natijalardan tijorat maqsadida foydalanishga yo'naltirilgan sezilarli darajadagi tashkilot va loyihalar yuzaga keldi. Bu eng avvalo genomlarning, shuningdek odam genomining strukturaviy, funktsional hamda qiyosiy tahlili bo'yicha faoliyat yurituvchi tashkilotlardir. Bioinformatika sohasi bo'yicha yaratilgan usullarni qo'llash bilan birga amaliy muammolarni yechish yo'lida, xususan farmakologiyada texnik hamda dasturiy bazalar jadal suratda rivojlanib bormoqda. Bunday muammolarni bartaraf etishda dasturiy ta'minot sanoati ham takomillashib bormoqda.

Mamlakatimizda genomika va bioinformatika fanlarining rivojlanishiga qaratilayotgan alohida e'tibor tufayli dunyo fanida o'z o'rniga ega nufuzli ilmiy maktab va muhit shakllantirildi, zamonaviy laboratoriyalar tashkil etilib, keng miqyosda xalqaro ilmiy aloqalar yo'lga qo'yildi. Xususan, O'zbekiston Respublikasi Fanlar akademiyasi "Genomika va bioinformatika" markazida sohada anchagina muvaffaqiyatli dasturlar amalga oshirildi. Markazda yetakchi xorijiy ilmiy markaz tajribalariga ega, bioinformatcion texnologiyalar bo'yicha bilim va ko'nikmalarni puxta egallagan ilmiy xodimlarning faoliyat olib borishi va shular hisobga olingan holda markazda bioinformatika laboratoriyasining tashkil etilganligi bunga yaqqol misol bo'la oladi. Markaz ilmiy jamoasi hanuzgacha noaniq bo'lgan g'o'za genomidagi rekombinatsion bloklar (ya'ni, avloddan avlodga ko'chib o'tadigan gen allellari to'plami) o'lchamlarini topib, topib, zamonaviy tezkor "assotsiativ kartalashtirish" usulini kashf etdi. Natijada g'o'za genomidagi genlardan foydalanishning yangi imkoniyatlari ochilib, g'o'zada zamonaviy markerlarga asoslangan seleksiya usullari ishlab chiqildi. Gen-nokaut yoki RNK interferentsiyasi molekulyar genetika va bioinformatika usullari mahsuli bo'lib, organizmning belgilangan genlari faolligini to'xtatish imkonini beradi. Shu tufayli genlari "o'chirilgan" (nokaut qilingan) organizm vujudga keladi. Bu nukleotid ketma-ketligi ma'lum bo'lgan genlarning funksiyasini aniqlashga yordam beradi. Nokaut qilingan va normal organizm namunalari orasidagi farqlar, o'rganilayotgan gen funksiyasini ko'rsatib beradi. Qishloq xo'jaligi ekinlarining biologik ko'rsatkichlari - hosildorlik,

ertapisharlik, zararkunanda va hasharotlarga chidamlilikning namoyon bo'lishida ishtirok etuvchi genning tarkibi va funksiyasi aniqlangandan so'ng maqsadga muvofiq ravishda ushbu gen faoliyatini kuchaytirish yoki aksincha uni to'xtatish mumkin. Markaz olimlari erishgan eng so'nggi yutuqlardan biri - bu ular tomonidan g'o'za uchun yaratilgan dunyodagi ilk -nokaut texnologiyasidir.

Bioinformatika, shuning uchun bio bo'limlariga kiritilgan - anatomiya, botanika, virusologiya, mikrobiologiya, tsitologiya, paleontologiya, fiziologiya va boshqalar bilan birga. Bioinformatika eksperimental biologiyada olingan dalillarni tahlil qiladi. Eksperimentator olingan ma'lumotlarni ma'lumotlar bankida mavjud bo'lgan butun to'plam bilan solishtiradi. Ketma-ketliklar aniqlanganch, ma'lumotlar bazasiga kiritadi. Ma'lumotlar banklarning vazifalariga axborotlarni saqlash, tizimlashtirish, yangilash va ulardan foydalanishni ta'minlash kiradi. Biologik mavzulardagi ilmiy nashrlar banklari ham mavjud.

*Gen funksiyasini bashorat qilish.* Barcha genlarning funksiyalarini eksperimental ravishda aniqlash juda ko'p vaqt talab qiladi. Bu holda, bioinformatika ularni asoslangan xolda bashorat qilishda yordam beradi va vazifalari allaqachon aniqlangan genlar bilan taqqoslash imkonini beradi. Biron xil ketma-ketlikning alohida olingan bo'limlarining rolini baholash maqsadida masalan oqsil molekulasining biologik vazifalariga ma'sul bo'lgan joylar mavjud bo'ladi. Ushbu bo'limlarni quyidagi usullar yordamida aniqlashdan oldin bioinformatika ma'lum bir oqsilning barcha funksiyalarini bashorat qiladi. Oqsillarning ketma-ketliklarining molekulyar modellarini qurish asosida, Protein tuzilmalari eksperimental tarzda aniqlanadi, masalan, oqsil molekulalaridan iborat mikroskopik kristalni rentgen nurlari bilan nurlantirish orqali. Bu uzoq va qimmat jarayon. Ba'zi oqsillarni kristall strukturaga ega bo'lmaganligi uchun umuman bunday tahlil qilib bo'lmaydi. Kompyuter modellashtirish yordamida bioinformatika oqsilning fazoviy modelini tuzishga yordam beradi. Makromodellarning faoliyat mexanizmi tadqiqoti o'tkazilgandan so'ng, ularning modellari asosida olingan molekulaning fazoviy tuzilishini asosida bioinformatika usullaridan foydalanib, uning qanday ishlashini taxmin qilish mumkin ish.

*Kompyuter dasturlariga asoslangan dori dizaynlarini yaratish.* Dorilar turli kimyoviy birikmalarning fazodagi maqsadli oqsil bilan o'zaro ta'sirini simulyatsiya qilish yo'li bilan loyihalanishi mumkin. SHu bilan birga bunday ta'sirlarning optimal bo'lgan shaklini tanlash kerak. Axborotlarni tahlil qilish biologiya, kimyo, fizika, matematika va informatika fanlari uyg'unligi biologik tizimni ko'p jihatdan tavsiflashga imkon beradi. Foydalanish kompyuter resurslari tahlil jarayonini tezlashtirishga yordam beradi vanatijalarni olish aniqligi va tezligini oshiradi.

*Informatika yutuqlaridan foydalanish.* Bioinformatika texnologiyalari yordamida biologiyada yangi kashfiyotlar tibbiyot, farmakologiya, kosmetologiya,

biotexnologiya, qishloq xo‘jaligi, ekologiya va boshqa sohalarda keng qo‘llanilmoqda. Bioinformatika mustaqil ravishda amaliy qiymatga ega bo‘lgan natijalarni beradi va turli sohalarda ishlash uchun sharoit yaratadi.



Bioinformatika servis markazlari va resurslari

*Bioinformatikaning asosiy metodlari.* Bioinformatikaga oid ishlarning aksariyati biologik axborotni saqlash va keyin uni qayta ishlash uchun ma’lumotlar bazalaridan foydalanish texnologiyasiga qaratilgan. Bu ma’lumotlar bazalari davlatniki va xususiy bo‘lishi mumkin. Bunday ma’lumotlar bazalariga ochiq standartlar orqali ommaviy kirishni ta’minlash juda muhimdir. Biologik axborotni qayta ishlash uchun ontologiyalar va mantiqiy usullardan foydalanish rivojlanib bormoqda.

*Aniq dasturiy paketlarga misollar.* Bioinformatika usullari tizimlar biologiyasida muhim rol o‘ynaydi. Ular kompyuterlar yordamida keng imkonlar yaratadi.

- tajriba ma’lumotlarini ma’lumotlar banklariga to‘plash va integratsiya qilish;
- kompyuter tahlilini amalga oshirish uchun;
- turmush tizimlarini strukturaviy va funksional tashkil etishni matematik modellashtirishni o‘tkazish;
- tirik tizimlarning yangi xususiyatlarini bashorat qilish;
- shu asosda eksperimental tadqiqotlarning yangi bosqichlarini rejalashtirish. Qo‘yida bugungi kunda keng tarqalgan asosiy ma’lumotlar bazalari va

ma'lumotlar banklari keltirilgan:

- GenBank - <http://www.ncbi.nlm.nih.gov/GenBank>
  - Nukleotid ketma-ketliklarining ma'lumotlar banki (3400000000juft asoslari 461000 ketma-ketliklarda).
  - SWISS-PROT - <http://www.expasy.ch/sprot/sprot-top.html>
  - Oqsillarning ketma-ketligi asosida annotatsiyalangan ma'lumotlar banki.
  - PIR - <http://www.nbrf.georgetown.edu/pir/searchdb.html>
  - gomologiya va taksonomiyaga ko'ra, oqsillarning ketma-ketligiga ko'ra annotatsiyalangan bank ma'lumotlari.
  - PDB - <http://www.rcsb.org/pdb/>
  - 3D strukturaga ko'ra, biologik makromolekulalarning ma'lumotlar banki. Bank dambix po strukture biologicheskix makromolekul.
  - NDB - <http://ndbserver.rutgers.edu>
  - Nuklein kislotalar bo'yicha bank ma'lumotlari. DNK va RNK ning uchlamchi fazoviy tasvirlari kiritilgan.
  - ProDom - <http://protein.toulouse.inra.fr/prodom.html>
  - Oqsillar domeniga ko'ra ma'lumotlar banki.
  - ACT - (Artemis Comparison Tool) - genom tahlili;
  - Arlequin - populyatsion-genetik ma'lumotlar tahlili.
  - Bio Edit - aminokislota va nuklein kislotalar ketma-ketligini taxrirlanishi redaktori;
  - Bio Numerics - bioinformatika bo'yicha tijorat universal paketlar dasturi;
  - BLAST - aminokislota va nuklein kislotalar ketma-ketligining bazalarida bir-biriga yaqin ketma-ketliklarni izlash dasturi;
  - ClustalW - keng ravishdagi aminokislota va nuklein kislotalar ketma-ketligini taxrirlash
  - FASTA - aminokislota va nuklein kislotalar ketma-ketligining uxshash algortmlari to'plami
  - Mesquite - Java tilidagi qiyosiy biologiya uchun dastur
  - Muscle - ClustalW ga nisbatan tez va aniq dastur. Aminokislota va nuklein kislotalar ketma-ketligining solishtirish dasturi.
  - PopGene - populyatsiyalarning genetik xilma xilligini tahlili dasturi.
- Populations - populyatsion-genetik tahlil.

Keng qo'llaniladigan saytlar:

1. <http://www.bioinformatics.ru>
2. <http://www.rusbiotech.ru>
3. <http://elementy.ru/lib/430895>
4. <http://www.sibai.ru>
5. <http://lake.baikal.ru>

6. <http://www.jcbi.ru>
7. [http://www.rtcbl.iitp.ru/index\\_r.html](http://www.rtcbl.iitp.ru/index_r.html)
8. <http://medi.ru/pbmc/8800101.htm>

### **Muhokama uchun savollar:**

1. *Bioinformatika fani qaysi davlatlarda muhim fan sifatida qaraladi?*
2. *Bioinformatikaning qanday bo'limlari mavjud?*
3. *Genomika kontekstida annotasiya nimani anglatadi?*
4. *Gen funksiyasini bashorat qilish nimaga asoslanadi?*
5. *Zamonaviy biologik tadqiqotlarda bioinformatikaning ahamiyati qanday?*
6. *PubMed nima?*
7. *Bioinformatika qaysi sohalarda keng qo'llaniladi?*
8. *Bioinformatika fani nimani o'rganadi?*
9. *Bioinformatika fani qaysi fanlar bilan uzviy bog'langan?*

## **2-Mavzu: Zamonaviy bioinformatson ma'lumotlar bazalari**

### **Asosiy savollar**

1. Axborot va bioaxborot tushunchasi.
2. Oqsil, DNK va RNK nukleotidlar ketma-ketliklari ma'lumot bazalari.
3. Molekulalarni modellashtirish bo'yicha ma'lumotlar bazalari (MMDB, PDB, NCBI).

**Mavzuga oid tayanch tushuncha va iboralar:** *NCBI, WormBase, EMBL, DDBJ, ENA, GenBank, FlyBase, UniProt, TREMBL, UniProtKB, Swiss-Prot, MMDB*

**Dars maqsadi:** Talabalarga axborot va bioaxborot haqida tushuncha berish hamda biomolekulalar ketma-ketliklari haqida ma'lumotlarga ega bo'lish va ularning ma'lumotlar bazalari bilan tanishish. Talabalarga EMBL, DDBJ, GenBank, INSDC ma'lumotlar bazalari haqida tushuntirish. Talabalarga *PDB, SWISS-PROT, PIR, UniProt* ma'lumotlar bazalari haqida tushuntirish.

### **1-savol bayoni.**

#### **Axborot va bioaxborot tushunchasi.**

NIH genetik ketma-ketligining ma'lumotlar bazasi, barcha DNK ketma-ketliklarining sharhlangan to'plami GenBank Yaponiyaning DNK ma'lumot bazasi (DDBJ), Evropadagi nukleotidlar arxivi (ENA) va NCBI-da GenBank ni o'z ichiga olgan Xalqaro nukleotidlar ketma-ketligi ma'lumotlar bazasi bilan hamkorlikning bir qismidir. Ushbu uchta tashkilot har kuni ma'lumot almashadilar. GenBank har ikki



oyda bir marta chiqariladi va FTP saytidan olinadi. GenBankning hozirgi versiyasi uchun nashr eslatmalarida batafsil ma'lumotlar va GenBankda bo'lajak o'zgarishlar to'g'risida bildirishlar mavjud. GenBankning oldingi versiyalari uchun chiqarilgan eslatmalar ham mavjud. GenBankning an'anaviy GenBank va WGS bo'linmalari bo'yicha o'sish statistikasi har nashrda mavjud.

Sequin - GenBank ketma-ketlik bazasiga yozuvlarni yuborish va yangilash uchun NCBI tomonidan ishlab chiqilgan mustaqil dasturiy vositadir. Bitta qisqa mRNK ketma-ketligini va uzoq ketma-ketliklarni, ko'p izohlarni, jagged ketma-ketliklarni yoki filogenetik va populyatsion tadqiqotlarni o'z ichiga olgan oddiy tasvirlarni qayta ishlashga qodir. Bitta Sequin faylida maksimal ishlash uchun 10000 dan kam ketma-ketliklar bo'lishi kerak. Katta takliflar tablitsasidan foydalangan holda amalga oshirilishi kerak.

Bioinformatika sohasining shiddat bilan rivojlanishi biotexnologiya va seleksiya yo'nalishlari taraqqiyotiga turtki berib, gen muxandisligining amaliy sohasini yuzaga keltirdi. Biroq an'anaviy gen muxandisligi usullari bir qator kamchilik ega bo'lib, bulardan bittasi - odam va hayvonlarning katta genomlarini manipulyatsiya qilish o'ta murakkabligidir.

Odam genomi Milliy tadqiqot instituti 2003 yilda yangi halqaro loyiha ENCODE (ingl. Encyclopedia of DNA Elements, o'zb. DNK elementlari entsiklopediyasi) ustida ish boshladi. Loyihadan maqsad – olimlarning intilish va izlanishlarini birlashtirgan holda RNK va oqsillar darajasida faol bo'lgan elementlar, fundamental genetik jarayonlarni (transkripsiya, translatsiya va replikatsiya) nazorat qiluvchi regulyator elementlar va odam genomi funktsional elementlarining to'liq ro'yxatini olish edi. Bu kabi funktsional o'zaro aloqadorliklarni aniqlash uchun quyidagi ikki strategiya qo'llaniladi: genni o'chirish (nakout yoki nokdaun) hamda gen faoliyatini yoki uning ektopik ekspressiyasini kuchaytirish. An'anaviy usullar - gomologik rekombinatsiyalar qo'llangan transgenез sichqonlarda, bundan tashqari virusli va lentivirusli vektorlarning qo'llanilishi nafaqat qimmat, balki juda katta mehnat talab etadi, ular o'ta qat'iy belgilangan genom lokusida aniq o'zgarishlar kiritish imkonini bermaydi.

Hozirgi kunda olimlar ixtiyorida bir necha texnologiyalar paydo bo'ldi, bular orqali o'simliklar, hayvonlar va odam genomlarini o'ta yuqori aniqlikda tahrirlash imkonini beradi.

### **Ikkinchi savolning bayoni.**

Zamonaviy genom tadqiqotlarining o'ziga xos belgilaridan biri bu juda katta miqdordagi ketma-ketlik ma'lumotlarini yaratishdir. Genom ma'lumotlar hajmi o'sib borishi bilan, ma'lumotlarni boshqarish uchun murakkab hisoblash metodologiyasi talab qilinadi. Shunday qilib, genomika va bioinformatika oldida turgan birinchi vazifa - bu kompyuter ma'lumotlar bazalarini yaratish va ulardan foydalanish orqali keng

hajmdagi ma'lumotlarni saqlash va boshqarish tashkil qiladi. Biologik ma'lumotlar bazasi - bu tizimda saqlanadigan ma'lumotlarning tarkibiy qismlarini yangilash, so'rash va yuklab olish uchun mo'ljallangan, kompyuterlashtirilgan dasturiy ta'minot bilan bog'langan jarayon hisoblanadi. Oddiy ma'lumotlar bazasi har xil ma'lumot to'plamini o'z ichiga olgan bitta fayl bo'lishi mumkin. Ma'lumotlar bazasini rivojlantirishning asosiy maqsadi ma'lumotlarni oson qidirib topishga imkon beradigan tuzimli to'plamida ma'lumotlarni tashkil qilishdir. Misol uchun bir nechta mashhur ma'lumotlar bazalari:

NCBI - Genotika Milliy Biotexnologiya Markazi, Shveytsariya Bioinformatika Institutidan-SwissProt va Protein Axborot Resursidan-PIR ma'lumotlar bazalarini keltirish mumkin.

*Biologik ma'lumotlar bazalarining turlari.* Ularning tarkibiga qarab biologik ma'lumotlar bazalarini taxminan ikki toifaga bo'lish mumkin:

1. Birlamchi ma'lumotlar bazalari: Birlamchi ma'lumotlar bazalari arxivlashgan ma'lumotlar bazasi deb ham ataladi. Ular nukleotidlar ketma-ketligi, oqsillar ketma-ketligi yoki makromolekulyar tuzilish kabi eksperimental ravishda olingan ma'lumotlar kabilar bilan to'ldirilgan. Eksperimental tadqiqot natijalari tadqiqotchilar tomonidan to'g'ridan-to'g'ri ma'lumotlar bazasiga topshiriladi va bu ma'lumotlar arxivlanib boradi. Ma'lumotlar quyidagi bazalarga taqdim etiladi:

- EMBL, GenBank NCBI va DDBJ-nukleotidlar ketma-ketligi,
- Array Express Archive va GEO-funksional genomik ma'lumotlar,
- Protein ma'lumotlari banki -PDB; uch o'lchovli makromolekulyar tuzilmalar koordinatalari.

2. Ikkilamchi ma'lumotlar bazalari: Ikkilamchi ma'lumotlar bazalari dastlabki ma'lumotlarni tahlil qilish natijasida olingan ma'lumotlarni o'z ichiga oladi. Ikkilamchi ma'lumotlar bazalari ko'pincha ko'plab manbalardan olingan ma'lumotlarga, shu jumladan boshqa ma'lumotlar bazalariga boshlang'ich va ikkinchi darajali ma'lumotlar bazalariga asoslanadi.

Ushbu bazalarga quyidagilar misol bo'ladi:

- InterPro- oqsil oilalari, motivlari va domenlari,
- UniProt ma'lumotlar bazasi- oqsillar to'g'risidagi ketma-ketlik va funksional ma'lumotlar
- Ensembl-genlar o'zgaruvchanligi, funktsiyasi, boshqarish funksiyalari to'g'risidagi ma'lumotlar

Bundan tashqari ko'pgina ma'lumotlar manbalari ham birlamchi, ham ikkilamchi xususiyatlarga ega. Masalan, UniProt peptidlarni sekvenirlash tajribalarida olingan birlamchi ketma-ketlikni qabul qiladi. Shu bilan birga, UniProt peptidlar ketma-ketligini genom ma'lumotidan TrEMBL va SwissProt ma'lumotlar bazasidan oladi va tahlil qiladi.



## **Zamonaviy nukleotid ma'lumotlar bazasi.**

*Nukleotidlar ketma-ketligi bazalari-turlari va ahamiyati.* Biologiya sohasi tobora ma'lumotlarga boy fanga aylanib borgan sari, katta ma'lumotlar to'plamlarini saqlash va ular bilan aloqa qilish zarurati tug'iladi. Aniq misollarga - bu nukleotidlarning ketma-ketligi, oqsillar ketma-ketligi va rentgen kristallografiyasi, makromolekulyar NMR markazi tomonidan ishlab chiqarilgan oqsillarning 3D tarkibiy va tuzilish ma'lumotlarini o'z ichiga oladi. Nukleotidlar ketma-ketligidan tashkil topgan bunday ma'lumotlar bazalari nuklein kislotalari ketma-ketligining bazalari deb nomlanadi.

Nuklein kislotalari ketma-ketligini saqlaydigan va ommaga taqdim etadigan uchta asosiy ma'lumotlar bazasi mavjud: GenBank, NCBI, EMBL, DDBJ. Ular nukleotidlarning ketma-ketlik bazalari deb nomlanadi, chunki ular barcha nuklein kislotalari ketma-ketligining omboridir. Nukleotid ma'lumotlar bazasi GenBank, RefSeq, TPA va PDB kabi bir nechta bazalar to'plamidir. Genom, gen va transkripsiya ketma-ketligi ma'lumotlari biotibbiy tadqiqotlar va kashfiyot uchun asos bo'lib xizmat qiladi.

*a. GenBank* AQShda joylashgan bo'lib, NCBI portali orqali malaka oshiruvchilar foydalanishi mumkin. EMBL -Yevropa molekulyar biologiya laboratoriyasi Buyuk Britaniyada va DDJB -Yaponiyaning DNK ma'lumotlar bazasi, Yaponiyada joylashgan. Uchalasi ham nukleotidlar ketma-ketliklarini qabul qilishadi, so'ngra ular orasidagi optimal sinxronizatsiyaga erishish uchun har kuni yangi va yangilangan ma'lumotlarni almashadilar. Ushbu uchta ma'lumotlar bazasi asl ketma-ketlik ma'lumotiga ega bo'lganligi uchun birlamchi ma'lumotlar bazasiga birikadi - INSDC va ular Sequence Read Archive -SRA bilan hamkorlik qiladi, u yuqori o'tkazuvchanlikdagi ketma-ketliklaridan olingan ma'lumotlarni arxivlaydi. GenBankning ketma-ket ma'lumotlar bazasi ochiq foydalanish, barcha ommaga ma'lum bo'lgan nukleotidlarning izohli to'plami va ularning proteinli tarjimalaridan iborat. Ushbu ma'lumotlar bazasi nukleotidlarning ketma-ketligi bo'yicha xalqaro ma'lumotlar bazasi -INSDC bilan hamkorlik qilish doirasida Milliy biotexnologiya ma'lumotlari markazi -NCBI tomonidan ishlab chiqiladi va saqlanadi. Dunyo bo'ylab laboratoriyalarda 100000 dan ortiq alohida organizmlardan ishlab chiqariladigan ketma-ketliklarni jamlaydi. GenBank biologik sohalarda tadqiqotlar olib borish uchun muhim ma'lumotlar bazasiga aylandi va so'nggi 18 yilda har ikki oyda ikki baravar ko'payib, geometrik progressiv o'sdi.

*b. EMBL (Yevropa molekulyar biologiya laboratoriyasi).* Yevropa molekulyar biologiya laboratoriyasi-EMBL. Nukleotidlarning ketma-ketligi to'g'risidagi ma'lumotlar bazasi - Yevropa bioinformatika institutida -EBI saqlanadigan birlamchi nukleotidlarning ketma-ket to'plamini saqlaydi. Ma'lumotlarni genomlarni sekvenirlash markazlaridan, alohida olimlardan va patent idoralaridan oladi.

c. *DDBJ-Yaponiya DNK ma'lumotlar banki*. U Yaponiyaning Shizuoka prefekturasidagi Milliy Genetika Institutida -NIH joylashgan. Bu Osiyodagi yagona nukleotidlar ketma-ketligi ma'lumotlar banki hisoblanadi. Garchi DDBJ o'z ma'lumotlarini asosan yapon tadqiqotchilaridan qabul qilsa-da, u har qanday boshqa mamlakatlarning tadqiqotchilaridan ma'lumotlarni qabul qilishi va taqdim etishi mumkin.

### **DDBJ- ma'lumotlar bazasi Yaponiya**

*DNA Data Bank Japan*- Yaponiya DNK ma'lumotlar bazasi bo'lib, turli genlar va organizmlar bilan bog'liq nukleotid ketma-ketliklar haqida ma'lumotga ega bo'lgan elektron resurslar bazasidir (29-rasm). DDBJ markazi INSDC a'zosi sifatida nukleotid natija ma'lumotlarini to'playdi.

*DDBJ ma'lumotlar bazasining tashkillashtirilishi:*

- 1980 EMBL ma'lumotlar kutubxonasi tashkil etildi va Yaponiya davlatidan nukleotid
    - ma'lumotlar banki uchun xalqaro hamkorlikni so'radi.
  - 1982 EMBL va GenBank xalqaro hamkorlikni boshladi, ma'lumotlar bankida ishtirok etish uchun Yaponiya davlatiga xamkorik taklif qilishdi.
  - 1983 xalqaro ma'lumotlar banki uchun hissa qo'shish maqsadida nukleotid ma'lumotlarini to'plash, baholash va tajriba ma'lumotlarini yuklash boshlandi.
  - 1984 NIG, Genetika Milliy instituti universitetlararo ilmiy-tadqiqot instituti sifatida qayta tashkil etildi. DDBJ NIH tarkibida ishlay boshladi.
- 29-rasm. DDBJ ma'lumotlar bazasi oynasining umumiy ko'rinishi.*
- 1986 DNKning ma'lumotlar bazasi maslahat qo'mitasi tashkil etildi.
  - 1987 DDBJ mustaqil bo'ldi va DDBJ ma'lumotlar bazasi operatsiyasining rasmiy boshlanishi deb hisoblanadi.
  - 1995 DDBJning yanada samarali faoliyati uchun, CIB axborot Biologiya markazi NIG tashkil etildi.
  - 2001 CIB, CIB-DDBJ sifatida qayta tashkil etildi, Yaponiya axborot Biologiya va DNK ma'lumotlar banki markaziga aylandi.
  - 2005 DDBJ, EMBL, GenBank o'zaro hamkorlikni kuchaytirish maqsadida kelishib yagona Xalqaro nukleotid ketma-ketlik ma'lumotlar bazasi-INSDC ni tashkil qilishdi.
  - 2009 yilda DDBJ fakulteti xodimlari xizmati bilan DBCLS va DDBJ hamkorlik yanada kuchaygan.

INSD o'zlarining ma'lumotlar bazalarida mavjud bo'lgan barcha gen annotasiyalariga bepul va cheklovsiz kirishning yagona siyosatiga ega. Dunyo miqyosidagi olimlar

ushbu annotasiyalardan tajribalarni rejalashtirishda yoki tahlillarni nashr qilish uchun foydalanishlari mumkin. Ilmiy nashrlarni chop etishda ilmiy adabiyotlardan foydalangan holda, ilmiy izlanishlar asl nusxasini keltirish lozim bo'ladi.

INSID Yaponiya, Koreya, Yevropa va AQShda Patent idoralari tomonidan to'plangan patent- mualliflik ixtirolari ilovalar bilan bog'liq nukleotid natija ma'lumotlarni o'z ichiga oladi. DDBJ markazi, shuningdek, Yaponiya va Koreyada patent idoralari tomonidan to'plangan patent ilovalar bilan bog'liq aminokislotalar natija ma'lumotlarni beradi.

### **EMBL-tadqiqot markazi va ma'lumotlar bazasi haqida umumiy ma'lumot.**

EMBL -1974 yilda tashkil topgan, tirik tabiat haqidagi fanlar bo'yicha Yevropaning yetakchi laboratoriyalari-molekulyar biologiya spektrlarini qamrab oluvchi 80 dan ortiq mustaqil tadqiqot guruhlari bo'lgan hukumatlararo tashkilot(31-rasm) hisoblanadi. EMBL-tadqiqot markazi va ma'lumotlar bazasi oltita saytlar hamjihatligida ishlaydi: Heidelberg, Barselona, Gamburg, Grenobl, Rim va EMBL-EBI Xinxton.

Yevropa molekulyar biologiya laboratoriyasi -EMBL dunyodagi yetakchi ilmiy-tadqiqot institutlari va Yevropaning tirik tabiat haqidagi ilm-fan laboratoriyalari jamlanmalaridan biridir. EMBL Yevropa bo'ylab oltita markaz va uning saytlardan foydalanib ishlaydi:

1. Heidelberg, Germaniya\* - asosiy laboratoriya;
2. Xinxton, Buyuk Britaniya\* - Yevropa bioinformatika instituti (EMBL-EBI);
3. Grenobl, Fransiya\* - struktur biologiya bo'yicha tadqiqotlar va xizmatlar;
4. Gamburg, Germaniya\* - struktur biologiya bo'yicha tadqiqotlar va xizmatlar;
5. Rim, Italiya\* - epigenetika va neyrobiologiya;
6. Barselona, Ispaniya\* - to'qima biologiyasi va kasalliklarni modellashtirish.

*EMBL boshqaruvi.* EMBL - tirik tabiat bilan bog'liq bolgan fanlar bo'yicha fundamental tadqiqotlarga ixtisoslashgan hukumatlararo tashkilot bo'lib, 20 dan ortiq a'zo davlatlarning, shu jumladan, Yevropa va Isroilning, shuningdek, ikkita assotsiatsiyalashgan a'zo bo'lgan Argentina va Avstraliyaning jamoat tadqiqot markazlari tomonidan moliyalashtiriladi. EMBLni Bosh direktor, hozirda EMBL Kengashi tomonidan boshqaruvchi organ tomonidan tayinlangan professor Edit Xard boshqaradi. Kengash tarkibiga barcha a'zo va unga a'zo bo'lgan davlatlar vakillari kiradi. EMBL missiyasining asosiy funksiyalari quyidagilardan iborat:

- molekulyar biologiyada asosiy tadqiqotlarni o'tkazish;

- o'rgatish, barcha darajadagi olimlar, talabalar va yangi a'zolarga, tirik tabiat haqidagi bilimlarni o'rgatish
- a'zo davlatlar olimlariga xizmatlarni taklif qilish, tirik tabiat bilan bog'liq bo'lgan hamma ma'lumotlar yuzasidan xizmat ko'rsatish
- yangi vositalar va usullarni ishlab chiqish;
- texnologiyalar transferida faol ishtirok etish;
- Yevropa ilmiy tadqiqotlarini birlashtirish, zamonaviy biologiya soxasidagi hamma ilmiy tadqiqot ishlarini birlashtirish

*EMBL-dagi mavjud tadqiqotlar.* EMBL-da olib borilgan tadqiqotlar biologik tashkilotlarning ko'p darajalarida, molekuladan organizmgacha, shuningdek, hisoblash biologiyasi, bioinformatika va tizimlar biologiyasida eksperimental tahliliga urg'u beradi. Tadqiqotlar molekulyar biologiya spektrini qamrab oluvchi 80 dan ortiq mustaqil guruhlar tomonidan olib borilmoqda. EMBL xalqaro, innovatsion va fanlararo o'zaro hamkorlikni tashkil qilgan. Uning ko'pgina mamlakatlardan kelgan 1700 dan ortiq xodimlari bioinformatika, genomika, biologiya, fizika, kimyo va informatika fanlarini o'z ichiga oladi.

*Ilmiy xizmatlar haqida.* EMBL tomonidan taqdim etiladigan xizmatlar quyidagilardan iborat:

- biomolekulyar ma'lumotlar bazalari va bioinformatika vositalari, xususan EMBL-EBI;
- Gamburg va Grenoblda struktura biologiyasi uchun texnologik jihozlar va yuqori o'tkazuvchanlik texnologiyalari bilan ta'minlash;
- o'rnatish yoki saqlash uchun qimmatga tushadigan yoki katta xarajatlarni talab qiladigan usullar va texnologiyalardan tejamkor va samarali usullardan foydalanishni ta'minlash.

### **EMBL-EBI Foydalanish shartlari.**

1. EMBL-EBI vakolati bilan shaffof ravishda ilm-fanni tadqiqotlarini targ'ib qiladi, bu biologiya sohasidagi ilmiy tajribalardan olinadigan eng keng jamoatchilikka taqdim etiladigan ma'lumotlar bilan bog'liq bo'lgan bepul onlayn xizmatlar, ma'lumotlar bazasi va dasturiy ta'minotni taqdim qilishni oz ichiga oladi. Olimlar tomonidan yaratilgan ilmiy ma'lumotlar taqdim etilganda, taqdim etilgan ma'lumotlardan foydalanishga hech qanday cheklovlar qo'yilmaydi.

2. EMBL-EBI ilg'or ilmiy amaliyotga muvofiq har qanday onlayn xizmatlar, ma'lumotlar bazalari yoki dasturiy ta'minotga kiritiladi va nashrlarda yoki mahsulotlarda aks ettirilishini talab qiladi. Kutilayotgan natijalar tegishli veb-sahifada ko'rsatiladi.

3. EMBL-EBI-ga uning onlayn-xizmatlarida taqdim etilgan har qanday ma'lumot maxfiy hisoblanmaydi.

4. Barcha ilmiy ma'lumotlar ma'lum vaqt oraliq'ida saqlanadi va

ma'lumotlar turiga mos ravishda taqdim etiladi, masalan kirish ma'lumotlari kirish qo'mitasi tomonidan ko'rib chiqilishi kerak bo'lgan olimlar ma'lumotlari, ma'lum vaqtgacha nashr etishdan oldin taqiqlangan.

5. EMBL-EBI tomonidan saqlanadigan shaxsiy ma'lumotlar qonun yoki sud tartibga soluvchi buyruq talab qilgan hollardagina alohida holatlarda oshkor etiladi. EMBL-EBI ma'lum bir dasturiy ta'minot yoki ma'lumotlar bazalaridan biron-bir shaxsning foydalanishi uchun jamoat tashkilotlariga taqdim qilishi mumkin.

6. OpenScience-dagi majburiyati saqlab qolgan holda, ushbu foydalanish shartlarini istalgan vaqtda yangilash huquqini saqlab qoladi. O'zgarishlar muqarrar bo'lgan taqdirda, veb-saytlarga xabar yuborish orqali har qanday o'zgarishlar to'g'risida oqilona xabar berishga harakat qiladi, va ushbu o'zgarishlarni veb- saytdan foydalanganda tekshirib ko'rish mumkin bo'ladi. Eng yangi qayta o'zgarishlar, 'EMBL-EBI foydalanish shartlari' sahifasida ko'rinadi.

7. Ushbu foydalanish shartlariga tegishli har qanday savol yoki izohga quyidagi manzilga murojaat qilish mumkin: Administrator, EMBL-EBI, Wellcome Trust Genome Campus, Hinxton CB10 1SD



### **NCBI - ma'lumotlar bazasi -GenBank-AQSH**

[\(https://www.ncbi.nlm.nih.gov/genbank/\)](https://www.ncbi.nlm.nih.gov/genbank/)

NCBI-“Biotexnologik Ma'lumotlar Milliy Markazi” *National Center Biotechnology Information*, 1988 yilda AQSH da tashkil etilgan, Milliy tibbiyot kutubxonasi filiali sifatida- *National Library of Medicine* asosida yuzaga kelgan (32-rasm).

U AQShning Vifezda, Merilend shtatidagi Sog'liqni Saqlash Milliy Institutida joylashgan. NCBIning asosiy maqsadi sog'lom va kasal organizmda o'tadigan molekulyar va genetik jarayonlarni keng tushunish uchun yangi informatsion texnologiyalarni ishlab chiqarishdan iborat. Maxsus maqsadlariga: biologik ma'lumotlarni saqlash va tahlil qilish uchun avtomatlashtirilgan tizimlarni ishlab chiqish, ma'lumotlarni mashinali qayta ishlash yuqori texnologiyalarini rivojlantirish, foydalanuvchilarga bazalar ma'lumotiga va dasturiy ta'minotga kirishini osonlashtirish, butun jahon bo'yicha biotexnologik ma'lumotlarni yig'ish ishlarini olib borishdan iborat.

Shuning uchun NCBI Genbank (<https://www.ncbi.nlm.nih.gov/genbank/>) tarmog'iga xizmat ko'rsatadi. GenBank - bu nuklein kislotalar genetik ketma-ketligining ma'lumotlar bazasi, asosan NCBI tomonidan boshqariladi, barcha ommaga ma'lum bo'lgan DNK ketma- ketliklarining sharhlangan to'plamidir (Nuklein kislotalar tadqiqotlari, 2013 yil yanvar; 41 (D1): D3642). GenBank Yaponiyaning DNK ma'lumotbazasi (DDBJ), Yevropadagi nukleotidlar arxivi (ENA) va NCBI-da

GenBankni o'z ichiga olgan Xalqaro nukleotidlar ketma-ketligi ma'lumotlar bazasi bilan hamkorlikning bir qismidir.

NCBI bazasi quyidagi xizmat ko'rsatish resurslarini taqdim etadi:

*NCBI ma'lumotlar bazasining asosiy oynasi.*

- Mahalliy ketma-ketliklarni qidiruv vositasi (BLAST)
- BioProject (ilgari inson Genomi loyihasi)
- BioSistemalar
- BLAST (yakka tartibda)
- Konservativ domenlar ma'lumotlar bazasi (CDD)
- Bazada Saqlanadigan domenlarni qidirish xizmati (CD qidirish)
- Genomik tarkibiy o'zgarishlarning ma'lumotlar bazasi (dbVar)
- Genotiplar va fenotiplar to'g'risidagi ma'lumotlar bazasi (dbGaP)
- Single nukleotid Polimorphism Ma'lumotlar Bazasi (dbSNP)
- FTP: FASTA BLAST ma'lumotlar bazasi
- FTP: Genom xaritasi ma'lumotlari
- GenBank
- Genom BLAST
- Primer-BLAST
- PubMed
- Taksonomiya umumiy daraxti

- Vektorni tahrirlash vositasi (VAST)



## **INSDC - nukleotidlarning ketma-ketligi bo'yicha xalqaro ma'lumotlar bazasi**

(<http://www.insdc.org>)

*Nukleotidlarning ketma-ketligi bo'yicha xalqaro ma'lumotlar bazasi -INSDC.* DNK, RNK va aminokislotalar ketma-ketligini 30 yillar mobaynida o'z ichiga olgan ma'lumotlar bazalarini yig'ish va tarqatish bo'yicha kompleks sa'yl-harakatlarini boshqarishdan iborat (33-rasm). Biologik ma'lumotlar arxivlarining uzoq vaqtdan beri davom etayotgan global izlanishlardan biri, ommaviy nukleotidlarning ketma-ketligi haqidagi ma'lumotlarni to'playdi, saqlaydi va kerakli ma'lumotlar bilan ta'minlaydi.

- Buyuk Britaniyaning Xinxton shahridagi Yevropa Molekulyar Biologiya Laboratoriyasining Yevropa Bioinformatika Instituti- EMBL-EBI; <http://www.ebi.ac.uk/ena>.

Vaqt o'tishi bilan an'anaviy INSDC ketma-ketlik arxiviga kiritilgan nukleotidlar ketma-ketliklar sonining korrelyativ o'sishi kuzatildi. Jumladan, ommaviy ketma-ketlik ma'lumotlari-WGS hamda ommaviy bo'lmagan ya'ni Expressed Sequence Tag (EST), Genome Survey Sequence (GSS), Patent va Transkriptome Shotgun Assambleyasi (TSA) ketma-ketlik ma'lumotlari o'rin oladi. Keyinchalik (2013 yil) barcha ma'lumotlar BioSample dasturi yordamida optimallashtirilib umumlashtirildi. Ma'lumot o'rnida shuni aytish mumkinki, BioSample dasturi NCBI, DDBJ va EMBL ma'lumotlar bazasidan o'rin olgan.

Nukleotidlar ketma-ketligi bo'yicha xalqaro ma'lumotlar bazasi bilan ishlash siyosati.

1. INSDC o'zlarining ma'lumotlar bazalarida mavjud bo'lgan barcha yozuvlarga bepul va cheklovsiz kirishning yagona siyosatiga ega. Dunyo miqyosidagi olimlar ushbu yozuvlarga eksperimentlarni rejalashtirish yoki har qanday tahlil yoki tanqidlarni nashr etish uchun kirishlari mumkin. Tegishli maqola, chop etilgan ilmiy adabiyotlardan foydalangan holda, olimlarning asl nusxalariga tayanib, asl taqdimnomaga tayanib beriladi.

2. INSDC ma'lumotlariga kirishni cheklaydigan, ushbu yozuvlardagi ma'lumotlardan foydalanishni cheklaydigan yoki ushbu yozuvlar asosida nashrlarning ayrim turlarini taqiqlaydigan yozuvlarga bayonnomalarni biriktirmaydi. Xususan, foydalanish bo'yicha cheklovlar yoki litsenziyalash talablari biron-bir ketma-ketlik ma'lumotlari yozuvlariga kiritilmaydi va hech kim tomonidan ma'lumotlar bazasini qayta taqsimlash yoki undan foydalanishda hech qanday cheklovlar yoki litsenzion to'lovlar bo'lmaydi.

3. INSDC-ma'lumotlar bazasiga taqdim etilgan barcha annotasiyalar ilmiy

nashrlar sifatida doimiy ravishda ochiq bo'lib qoladi. Xatolarni tuzatish va mualliflar tomonidan yozuvlarning yangilanishi qabul qilinadi va noto'g'ri annotasiyalar bazaning keyingi nashrlaridan olib tashlanishi mumkin, ammo barchasi kirish raqami orqali doimiy ravishda saqlanib qoladi.

4. Taqdim etuvchilarga INSDC tomonidan qo'llab-quvvatlanadigan veb-saytlarda aks ettirilgan ma'lumotlar keng jamoatchilikka oshkor qilinishi tavsiya etiladi. Ma'lumotni topshirish huquqiga ega ekanliklarini aniqlash uchun yuboruvchilar javobgar.

5. Tahririyat nazorati va ba'zi ichki tekshiruvlaridan tashqari (masalan, INSD formatlaridan to'g'ri foydalanish va CDS yozuvlarida ko'rsatilgan kodlash hududlarining tarjimai tekshirilganda), ma'lumotning sifati va aniqligi ma'lumotlar bazasi emas, balki taqdim etuvchi muallifning zimmasidadir. Ma'lumotlar bazalari iloji boricha sifatli manbaga erishish uchun ma'lumotlar bazasiga ma'lumot taqdim etuvchilari va foydalanuvchilari bilan doimiy aloqada bo'ladi.

### **Uchunchi savol bayoni**

**Protein bazalari turlari va ahamiyati.** Proteinlar bazasi oqsillar haqida bir yoki bir nechta ma'lumotlar to'plamidir. Ular tarkibiga oqsilning aminokislotalar ketma-ketligi, tuzilishi, shuningdek funktsiyalari kabi ma'lumotlar kiradi.

Protein ma'lumotlar bazalari turli xil gen bazalaridan DNK ketma-ketligini translyatsiya qilish orqali tuziladi va protein tarkibiy ma'lumotlarini o'z ichiga oladi. Ular muhim manba hisoblanadi, chunki oqsillar ko'p biologik funktsiyalarni bajaradi.

*Protein ma'lumotlar bazalarining ahamiyati:* Protein tuzilmalari, funktsiyalari va ayniqsa ketma-ketligi to'g'risida juda ko'p ma'lumotlar olinmoqda. Ma'lumotlar bazalarini qidirish ko'pincha yangi oqsilni o'rganishda katta ahamiyatga egadir.

Protein ma'lumotlar bazalari quyidagi maqsadlarda foydalanadi:

1. Proteinlar va oqsillar oilalari o'rtasidagi taqqoslash genom tarkibidagi yoki turli xil turlardagi oqsillar o'rtasidagi munosabatlar haqida ma'lumot beradi va shu sababli faqat ajratib olingan oqsilni o'rganish orqali olinadigan ko'proq ma'lumotni taqdim etadi.

2. Eksperimental bazalardan olingan ikkilamchi ma'lumotlar bazalari ham keng tarqalgan. Ushbu ma'lumotlar bazalari ma'lumotlarni qayta tashkil qiladi va izohlaydi yoki bashorat qiladi.

3. Ko'pgina ma'lumotlar bazalaridan foydalanish ko'pincha tadqiqotchilarga oqsilning tuzilishi va funktsiyasini tushunishga yordam beradi.

*1. Protein ma'lumotlar birlamchi bazalari: a. PRIMARY* - ma'lumotlar bazalarida eksperimental tadqiqotlar natijasida aniqlangan oqsillar ketma-ketligi mavjud, ular nukleotidlar ketma-ketligining kontseptual translyasiyasidan hosil



qilingan. Konseptual translyasiya, albatta, eksperimental ravishda olingan ma'lumotlarga kirmaydi, shuning uchun nukleotidlar ketma-ketligi ma'lumotlarini talqin qilinishi natijasida paydo bo'lgan noto'g'ri ma'lumotlarga ega bo'lishi mumkin. Protein ketma-ketligi bo'yicha bir qator boshlang'ich ma'lumotlar bazalari mavjud va ularning har biri alohida e'tibor talab qiladi.

b. *Protein ma'lumot manbasi (PIR) - oqsillar ketma-ketligi bo'yicha ma'lumotlar bazasi (PIR-PSD):* PIR-PSD PIR, MIPS (Protein Sequences uchun Myunxen axborot markazi, Germaniya) va JIPID (Yaponiya xalqaro oqsil ma'lumotlari bazasi, Yaponiya) o'rtasidagi hamkorlikdagi ma'lumot bazasidir. PIR-PSD ning o'ziga xos xususiyati shundaki, u super-oila tushunchasi asosida oqsillar ketma-ketligini tasniflaydi. PIR-PSD-da ketma-ketlik gomologiya domeni va ketma-ketliklar bo'yicha ham tasniflanadi. Gomologiya domenlari evolyutsion shajara bloklariga to'g'ri kelishi mumkin, ketma-ketliklar funktsional joylarni yoki saqlanib qolgan hududlarni anglatadi. Bunday tasniflash yondoshuvi funktsional- struktura munosabatlari ketma-ketligini to'liqroq tushunishga imkon beradi.

c. *SWISS-PROT:* Taniqli va keng ishlatiladigan proteinlar bazasi bu SWISS-PROT hisoblanadi. PIR-PSD singari, bu saralangan oqsillar ketma-ketligi bazasi hamda annotasiyalarning yuqori darajasini ta'minlaydi. Har bir tahrirdagi ma'lumotlarni alohida ma'lumotlar va annotasiyalar sifatida ko'rib chiqish mumkin. Asosiy ma'lumotlar umumiy bitta harfli aminokislotalar kodiga kiritilgan ketma-ketliklar va tegishli ma'lumot va bibliografiyadan iborat. Tahrirda oqsilning tuzilishi yoki funktsiyalari, fosforillanish, asetilatsiya va boshqalar kabi post- translatsional o'zgarishlar, tarkibiy sohalar, saytlar, kaltsiyni bog'laydigan hududlar, ATP-ulanish joylari, ruh barmoqlari va boshqalar to'g'risida ma'lumotlar mavjud.

d. *TrEMBL (tarjima qilingan EMBL uchun) -* bu SWISS-PROTga qo'shimcha sifatida chiqarilgan dasturga kiritilgan oqsillar ketma-ketligi ma'lumotlar bazasi hisoblanadi. Unda EMBL nukleotid ma'lumotlar bazasida to'liq tahrirlanmagan kodlash ketma-ketliklarining tarjimasi mavjud. Shunday qilib, ushbu ma'lumot bazasida hech qachon tahrirlanmagan va organizmda aniqlanmagan oqsillar ketma-ketligini o'z ichiga olishi mumkin.

e. *Protein ma'lumotlar banki (PDB):* PDB-oqsil tuzilishi to'g'risidagi asosiy ma'lumotlar bazasi hisoblanadi. Ma'lumot bazasida yuqori molekulyar oqsillar biologik molekulalarning uch o'lchamli tuzilishi to'g'risida kristallografik ma'lumotlar bazasi. PDB nafaqat oqsillarning, balki nuklein kislotalar bo'laklari, RNK molekulalari, antibiotik gramitsidin kabi yirik peptidlar, oqsil va nuklein kislotalar komplekslarining uch o'lchamli tuzilishini ham arxivlaydi. Ma'lumotlar bazasida asosan uchta manbadan olingan ma'lumotlar mavjud: rentgen kristallografiyasi, tadqiqotchilar tajribalari va molekulyar modellashtirish.

2. *Protein ma'lumotlar ikkilamchi bazalari*: ularda birlamchi ma'lumotlar bazalarida joylashgan ketma-ketlikni chuqur tahlil qilish natijalarini ham o'z ichiga oladi. Ko'pgina oqsillarning ikkinchi darajali ma'lumotlar bazalari turli xil oqsillarga tegishli xususiyatlarni izlash natijasidir. Ba'zi bir ishlatiladigan ketma-ketlik va tuzilishning ikkinchi darajali ma'lumotlar bazalari quyidagilar kiradi:

a. *MHC Pep:MHC Pep* - bu 13000 dan ortiq peptid ketma-ketligidan tashkil topgan ma'lumotlar bazasi, bu immunitet tizimining asosiy gistomansublik kompleksi to'g'risidagi ma'lumotlarni saqlaydi.

**PDB - protein ma'lumotlari banki** Protein ma'lumotlari banki (PDB) 1971 yil oktyabr oyida PROTEIN DATA BANE Nature New Biologiyada

### **PIR -protein axborot resursi**

**PIR** - *protein axborot resursi* bazasi, Jorjtaun universiteti tibbiyot markazi (*GUMC*) tomonidan bioinformatikani qo'llab-quvvatlash maqsadida genom va proteom tadqiqot natijalari va oqsil ketma-ketliklarini taqdim etadi. *PIR -Protein axborot resursi markazining asosiy oynasi*.



Tadqiqotlar Fondi (*NBRF*) tomonidan tadqiqotchilar va mijozlarga oqsillar ketma-ketligi ma'lumotlarini aniqlash va izohlashda yordam beradigan manba sifatida tashkil etilgan. Bungacha NBRF 1964-1974 yillarda Margaret Dayhoff tahriri ostida nashr etilgan "Proteinlar ketma-ketligi va tuzilishi atlas" da birinchi makromolekulyar ketma-ketlik to'plamini tuzgan. Dayxof va uning tadqiqot guruhi oqsillar ketma-ketligini taqqoslash, ketma-ketliklar orasidagi masofa va ketma-ketlikni aniqlash hamda oqsillar ketma-ketligi evolyutsion tarixini aniqlash uchun kompyuter usullarini ishlab chiqishda kashf etdilar.

PIR boshqa ma'lumotlar bazalari jumladan, EBI (Yevropa Bioinformatika Instituti), oqsillar ketma-ketligi va funksiyalarining markaziy manbai bo'lgan UniProt (Birlashgan Protein ma'lumotlar bazasi) va SIB (Shveytsariya Bioinformatika Instituti) bilan birgalikda ish yuritishadi.

PIR o'zida oqsillar ketma-ketligi, oilalari, fermentlar va fermentativ yo'llar, molekulyar tuzilmalar va strukturalar ularning sinflari, genlar va genomlar, adabiyot va taksonomiyalarning ko'p sonli molekulyar ma'lumotlar bazalariga keng ko'lamli murojaatlarni va havolalarni taqdim etadi (1-rasm). Ushbu sayt ma'lumotlarini UniProt (Birlashgan Protein ma'lumotlar bazasi) resurslari bilan dastlabki tahlilini o'tkazgan holda taqdim etib, tahlil qilishda ham UniProt dasturlaridan foydalanadi.

### **UniProt ma'lumotlar bazasi**

**UniProt** - bu oqsillar ketma-ketligi va funksional ma'lumotlarning erkin kirishi mumkin bo'lgan ma'lumotlar bazasi bo'lib, ularning ko'plari genomni tartiblash

loyihalaridan olinadi (38-rasm). Unda tadqiqotlar natijasida olingan oqsillarning biologik funksiyasi haqida juda ko'p ma'lumotlar mavjud. Uni bir nechta Yevropa bioinformatika tashkilotlari va Vashington, Kolumbiya okrugi (AQSh) dan tashkil topgan UniProt konsorsiumi qo'llab- quvvatlaydi. UniProt keng doirada boshqa bazalar bilan hamkorlikda ish olib boradi. Jumladan:

- ✓ INSDC
- ✓ EMBL
- ✓ GenBank (NCBI)
- ✓ DDBJ
- ✓ Ansambl
- ✓ Yevropa patent idorasi (EPO)
- ✓ FlyBase: Drosophilidae hasharotlar oilasi uchun genetik va molekulyar ma'lumotlarning asosiy ombori.

- ✓ H-Invitasion ma'lumotlar bazasi (H-Inv)
- ✓ Proteinlarning xalqaro indeksi (IPI)
- ✓ Yaponiya Patent idorasi (JPO)
- ✓ Protein haqida ma'lumot manbasi (PIR-PSD)
- ✓ Protein ma'lumotlari banki (PDB)
- ✓ Oqsillarni tadqiqot qilish fondi (PRF) <sup>[18]</sup>
- ✓ Saccharomyces Genome Database (SGD)
- ✓ Arabidopsis haqida ma'lumot manbai (TAIR)
- ✓ AQSh Patent idorasi (USPTO)
- ✓ Shveysariya-Prot (SWISS-Prot)
- ✓ TrEMBL
- ✓ Umurtqali va Genomli Annotatsiya Ma'lumotlar Bazasi (VEGA)

Saytning yaratilishi Shveysariya-PROT tomonidan 1986 yilga borib taqaladi. Uning UniProt tarzida to'liq ishga tushirilishi 2003 yil dekabr oyida sodir bo'lgan. UniProt Inson Genomini o'rganish Milliy Instituti, Milliy Sog'liqni Saqlash Institutlari (NIH), Yevropa Komissiyasi, Shveysariya Federal Hukumati Ta'lim va Ilmiy Federal Xizmati, NCI-caBIG va AQSh Mudofaa Vazirligi tomonidan moliyalashtiriladi. UniProt to'rtta asosiy ma'lumotlar bazasini taqdim etadi:

1. UniProtKB (Swiss-Prot qismlari bilan)
2. UniProtKB (TrEMBL-qismlari bilan)
3. UniParc
4. UniRef.

UniProt bilimlar bazasi (UniProtKB) - bu proteinlar bazasi bo'lib, ikki qismdan iborat: UniProtKB / Swiss-Prot (ko'rib chiqilgan, qo'lda izohlangan ma'lumotlarni o'z

ichiga olgan) va UniProtKB / TrEMBL (ko'rib chiqilmagan, avtomatik izohlangan ma'lumotlarni o'z ichiga olgan). 2014 yil 19 mart holatiga ko'ra UniProtKB / Swiss-Prot da 193,019,802 aminokislotalar va UniProtKB / TrEMBL da 17,207,833 aminokislotalar ketma-ketligi ma'lumoti mavjud.

*UniProtKB / Swiss-Prot* - bu qo'lda izohlangan, ortiqcha bo'lmagan proteinlar ketma-ketligining ma'lumotlar bazasi. Bu ilmiy adabiyotlar va olingan ma'lumotlarni birlashtiradi va tahlil qiladi. UniProtKB / Swiss-Protning maqsadi ma'lum bir protein haqida barcha ma'lum ma'lumotlarni taqdim etish va oqsillar ketma-ketligi va ilmiy adabiyotlarning batafsil tahlilini o'z ichiga oladi. *UniProtKB / TrEMBL* - avtomatik izoh bilan boyitilgan yuqori sifatli hisoblash, tahlil ma'lumotlarini o'z ichiga oladi.

*UniProt Archive (UniParc)* - bu oqsillar ketma-ketligining asosiy, hamma uchun ochiq bo'lgan to'liq ma'lumotlar bazasi. Bazada proteinlarning bir necha xil turini topish mumkin. Oqsilning har bir ketma-ketligiga barqaror va noyob identifikator (UPI) va proteinni turli xil manbalar bazalarida aniqlash imkonini beradi. Bundan tashqari ma'lumotlar bazasidagi protein ketma-ketliklari o'zgariganda, bu o'zgarishlar UniParc tomonidan kuzatiladi va barcha o'zgarishlar tarixi arxivlanadi.

*UniProt Referents Klasterlari (UniRef)* - yuqoridagi UniProtKB (Swiss-Prot qismlari bilan), UniProtKB (TrEMBL-qismlari bilan) va UniParc da mavjud xususiy va umumiy oqsillar ketma-ketligi yagona UniRef yozuviga birlashtiradi.



## **SWISS-PROT-proteinlar ketma-ketligi ma'lumotlar bazasi**

Genom loyihalaridan ketma-ket ma'lumotlar bazalariga suis^jrot ma'lumotlar oqimi oshib borishi sababli ma'lumotlar bazasini izohlashda bir qator qiyinchiliklarga duch kelinmoqda. SWISS-PROT-da yuqori sifatli ketma-ketlik va izohlarning saqlanishi bu kabi muammolarning yechimini topish imkonini beradi (39-rasm)

SWISS-PROT - 1987 yilda Jeneva Universitetining Tibbiy biokimyo bo'limida yaratilgan. U Yevropa molekulyar biologiya laboratoriyasi (EMBL) va Shveysariya Bioinformatika Instituti (SIB) o'rtasida teng huquqli sheriklik asosida birgalikdagi sa'y-harakatlari bilan ish yuritadi. Hozirgi vaqtda SWISS-PROT 31 ta turli xil ma'lumotlar bazalari bilan bog'langan va bio-molekulyar ma'lumotlar almashinib turadi. Hozirgi vaqtda (1999 yil oktyabr holatiga ko'ra) SWISS-PROT ~ 81 000 ketma-ket yozuvlarni o'z ichiga oladi, bunda ~ 65 000 ta ma'lumotlardan olingan 30 million aminokislotalar haqida ma'lumotlar mavjud.

SWISS-PROT ma'lumotlar bazasidagi oqsillar ketma-ketligi boshqa ma'lumotlar bazasidan uchta aniq mezon bilan ajratib turadi: (I) aniq izohlar, (II) minimal ma'lumotlar va (III) boshqa ma'lumotlar bazalari bilan to'laqonli



Qabul qiling Protein ketma-ketligi uchun

EBI Home About EBI Research Services Toolbox **Databases** Downloads Submissions

[SWISS-PROT MA'LUMOTLARI](#)



▪ [SWISS-PROT Bosh sahifa](#)

▪ [Ma'lumot](#)

▪ [Kirish](#)

▪ [Topshiriqlar](#)

▪ [Asboblari](#)

▪ [FTP](#)

▪ [Guruh haqida ma'lumot](#)

▪ [Hujjatlar](#)

▪ [Aloqa](#)

SWISS-PROT Ma'lumotlar bazasiga kirish

SWISS-PROT ma'lumotlar bazasiga quyidagi qidiruv tizimi orqali kirish mumkin

[Qidirmoq](#) Proteinlar uchun :

Qidiruv mexanizmi [SRS](#) tomonidan quvvatlanadi, ushbu izlash vositasi yanada murakkab va / yoki bir nechta ma'lumotlar bazasi so'rovlari uchun ishlatilishi mumkin.

Kirishning boshqa usullari SWISS-PROT ma'lumotlar bazasiga quyidagilar yordamida kirish mumkin.

### SWISS-PROT ma'lumotlar asosiy oynasi

PROT-da ma'lumotlarning ikki sinfini ajratish mumkin: asosiy ma'lumotlar va izohlar. Har bir ketma-ketlik uchun asosiy ma'lumotlar ketma-ketlik ma'lumotlaridan iborat; iqtibos ma'lumotlari (bibliografik havolalar) va taksonomik ma'lumotlar (oqsilning biologik manbasi tavsifi), izoh esa quyidagi elementlarning tavsifidan iborat:

- Oqsilning funktsiyalari

Post-translational modifikatsiyalar, masalan, uglevodlar, fosforillanish, atsetilatsiya, GPI-langar va boshqalar.

- Domenlar va saytlar. Masalan, kaltsiyni bog'laydigan mintaqalar, ATP-ulanish joylari, sink barmoqlari, gomeobokslar, CH<sub>2</sub> va CH<sub>3</sub> domenlari va boshqalar.

- Ikkilamchi tuzilish. Masalan, alfa spiral, beta zanjir va boshqalar.

To'rtlamchi tuzilish. Masalan, homodimer, heterotrimer va boshqalar.

- Boshqa oqsillarga o'xshashliklar
- Oqsil yetishmovchiligi bilan bog'liq kasalliklar
- Tarkibiy nizolar, variantlar va boshqalar.

*Odam proteomikasi bo'yicha loyiha (HPI)*-Bir necha oy ichida bir qator sekvens markazlari va kompaniyalarining birgalikdagi sa'y-harakatlari inson genomlari ketma-ketligining birinchi loyhasini ishlab chiqishdi. Bunday urinish insonning biologik jarayonlarini tushunishda juda muhim qadamdir. Bu loyihani amalga oshirishda ko'pgina muammolar mavjud. Jumladan, barcha kodlash hududlarini genom ketma-ketligi bo'yicha aniqlash murakkabdir. Hozirgi algoritmlar juda kuchli bo'lishiga qaramay, barcha ekzonlarni aniq aniqlashga qodir emas, turli xil variantlarni ajratish uchun yaxshi jihozlanmagan va mayda oqsillarni aniqlay olmaydilar va bunday oqsillar ko'pgina biologik jarayonlar uchun hal qiluvchi rol o'ynaydi. Bir marta ribosomalarda sintez qilingan oqsillar ko'p sonli o'zgartirish bosqichlariga duchor bo'ladi. Ushbu barcha modifikatsiyalar tufayli xilma-xillik yuqori daraja bilan murakkablashmoqda.

Shunday qilib, inson genomini ifoda etadigan turli xil protein molekulalarining soni millionga yaqinroqdir. Hisobga olish kerak bo'lgan murakkablikning yana bir omili - bu proteinlar ketma-ketligi darajasidagi polimorfizm miqdorining mavjudligi hisoblanadi. Ushbu polimorfizmlarning ba'zilar kasallik holatlari bilan bog'liq, shuning uchun SWISS-PROT sifat standartlariga muvofiq barcha ma'lum bo'lgan odamlarning ketma-ketligini izohlash uchun katta loyihani amalga oshirmoqda. Bu har bir ma'lum protein uchun uning funktsiyasi, uning domen tuzilishi, hujayra ichidagi joylashishi, translyasidan keyingi o'zgarishlar, variantlar, boshqa oqsillarga o'xshashlik va hokazolarni o'z ichiga olgan juda ko'p ma'lumotni taqdim etishni anglatadi. Hozirda 5400 dan oshiq izoh mavjud.

HPI loyihasi quyida qisqacha tavsiflangan bir qator kichik tarkibiy qismlarni o'z ichiga oladi:

- Barcha ma'lum bo'lgan proteinlarga izoh berish. Keyingi 9 oy davomida (2000 yil aprelgacha) SWISS-PROT-da oldin mavjud bo'lmagan protein oqsillari to'liq izohlandi.
- Sut emizuvchilarining inson oqsillari ortologlarini izohlash. Har qanday inson oqsillari uchun, sutemizuvchi boshqa turlardagi mavjud ortologlarga izoh beriladi.
- Proteinlar ketma-ketligi darajasida barcha ma'lum bo'lgan polimorfizmlarga izoh berish. Yuqorida ta'kidlab o'tilganidek, SWISS-PROT bunday polimorfizmlarning katta miqdori haqida ma'lumotga ega va u protein darajasida barcha «mayda» o'zgarishlarni saqlash va izohlash harakatini sezilarli darajada kengaytiradi.
- Odam oqsillaridagi barcha ma'lum post-translatsion o'zgarishlarni izohlash. Kelgusi 9 oy davomida SWISS-PROT-da taqdim etilgan oqsillardagi ma'lum translyasidan keyingi modifikatsiyalarning batafsil tavsifini to'ldirishga katta kuch sarflanadi.

- Strukturaviy ma'lumotlarga qattiq bog'lanish. SWISS-PROT PDB / RCSB 3D-tuzilmasi ma'lumotlar bazasi bilan chambarchas bog'langan biologlar uchun foydali bo'lgan ko'plab xususiyatlarni o'z ichiga oladi. Bunday yaqin aloqalar, ilmiy yondashuvga mos keladigan, insonning barcha oqsillari uchun homologydan olingan modellarni taqdim etish orqali yanada kengaytiriladi. HPI loyihasi va uning hozirgi holati haqida batafsil ma'lumot <http://www.expasy.ch/sprot/hpi/> da berilgan.

### **Muhokama uchun savollar:**

1. *PDB nima va u kim tomonidan qachon tashkil etilgan?*
2. *SWISS-PROT nima va u kim tomonidan qachon tashkil etilgan?*
3. *PIR nima va u kim tomonidan qachon tashkil etilgan?*
4. *UniProt nima va u kim tomonidan qachon tashkil etilgan?*
5. *EMBL nima va qachon tashkil etilgan?*
6. *DDBJ nima va qachon tashkil etilgan?*
7. *GenBank nima va qachon tashkil etilgan?*
8. *INSDC nima va qachon tashkil etilgan?*

### **3- mavzu: Biologik ketma-ketliklarni taqqoslash.**

#### **Asosiy savollar**

1. Biologik ketma-ketliklarda makromolekulalarning polimerlanishi.
2. Blast dasturi. Blast oilasining asosiy gruppalari.
3. Blast dasturining ishlash printsi

**Mavzuga oid tayanch tushuncha va iboralar:** *DNK, RNK va oqsil strukturalari, megablast, Dmegablast, blastn, blastp, Cdar, rpsblast, psi-blast, phi-blast, blastx, Tblastn, tblastx, magicblast, bl2seq, VecScreen,*

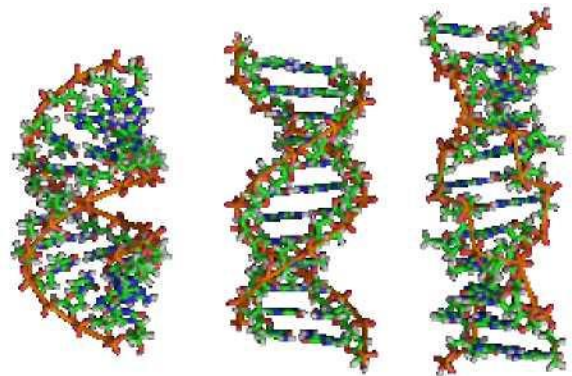
**Dars maqsadi:** DNK, RNK va oqsil strukturalarining molekulyar va kimyoviy tuzilishi haqida tushunchalarga ega bo'lish hamda ularning ketma-ketliklarini taqqoslashda foydalaniladigan blast dasturining turlari bilan tanishish.

#### **Birinchi savolning bayoni.**



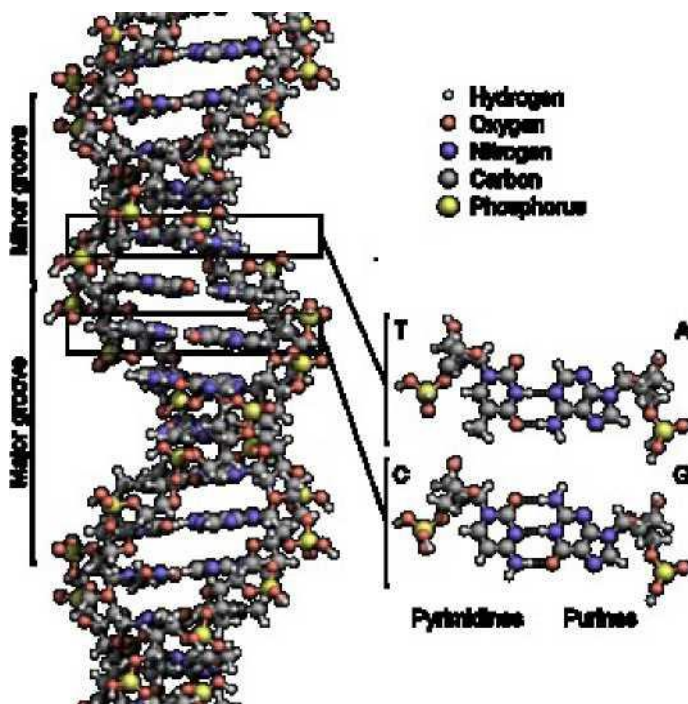
Dezoksiribonuklein kislota (DNK) - bu tirik organizmlarning rivojlanishi va faoliyati uchun genetik dasturni saqlash, avloddan avlodga yetkazish va amalga oshirishni ta'minlaydigan makromolekuladir. DNK molekulasi biologik ma'lumotlarni nukleotidlar ketma-ketligidan iborat genetik kod shaklida saqlaydi. DNKda har xil turdagi RNK va oqsillarning tuzilishi haqida ma'lumotlar mavjud. Eukariot hujayralarda - hayvonlar, o'simliklar va zamburug'larda DNK xromosomalarning bir qismi sifatida hujayra yadrosida, shuningdek ba'zi hujayra organoidlarida - mitoxondriya va plastidlarda joylashgan. Prokariot organizmlar - bakteriyalar va arxeylar hujayralarida aylana yoki chiziqli DNK molekulasi, ya'ni nukleoid deb ataladigan hujayra membranasiga biriktirilgan. Quyi taraqqiy etgan eukariotlar masalan, achitqi zamburug'ida plazmidlar deb nomlangan kichik, avtonom, asosan aylana shaklidagi DNK molekulalari mavjud. Bundan tashqari, bitta yoki ikki zanjirli DNK molekulalari DNK - tutuvchi viruslarda viruslar genomini hosil qilishi mumkin.

Aksariyat hollarda bir qatorli DNKni o'z ichiga olgan ba'zi viruslar bundan mustasno. DNK makromolekulasidagi azotli asoslar bir-biriga komplementar ikkita zanjir asosida tashkil topgan. Ushbu ikki zanjirli molekula spiral chiziqqa o'ralgan. Umuman olganda, DNK molekulasining tuzilishi an'anaviy "qo'sh spiral" shaklida tan olingan, lekin, ushbu ko'rinish noto'g'ri nomlangan bo'lib, aslida "ikkita vint" shaklidagi nomlanish to'g'ri hisoblanadi. Spiral o'ng tarafga og'gan bo'lsa, DNKning A- va B shakllari, chap tarafga og'gan bo'lsa, DNKning Z-shakli deb ataladi (4-rasm).



4-rasm. Ionlarning konsentratsiyasi va molekulaning nukleotid tarkibiga qarab, tirik organizmlarda DNKning ikki karra spirali turli shakllarda mavjud. Rasmda A, B va Z shakllari ko'rsatilgan (chapdan o'ngga)





5-rasm. DNK tuzilishi (juft spiral). Tuzilishdagi turli xil atomlar turli xil ranglarda ko'rsatilgan; ikkita asosiy juftlikning batafsil tuzilishi quyida ko'rsatilgan

Bundan tashqari, eukariotlarning genomida ko'pincha "genetik parazitlar" ga tegishli mintaqalar, masalan, transpozonlar mavjud bo'ladi.

DNK tuzilishini aniqlash 1953 yilda, biologiya tarixidagi burilish nuqtalaridan biri bo'ldi. Francis Krik, Jeyms Uotson, va Maurice Wilkins DNK struktursini aniqlashda qo'shgan buyuk hissasi uchun 1962 yilda fiziologiya va tibbiyot soxasi bo'yicha Nobel mukofoti bilan taqdirlandi. Ushbu mukofotga hissa qo'shgan Rozalinda Franklin rentgenografiya usulida DNK tuzilishi to'g'risida ma'lumotlar olgan, ushbu strukturani esa Uotson va Krik ochib bergan.

DNK juft spiraling tuzilishi Frensis Krik va Jeyms Uotson tomonidan 1953 yilda Mauris Uilkins va Rozalind Franklin tomonidan olingan rentgen ma'lumotlari va Chargaff qoidalari asosida taklif qilingan.



[Francis Krik](#)



[Jeyms Uotson](#)



[Moris Uilkins](#)

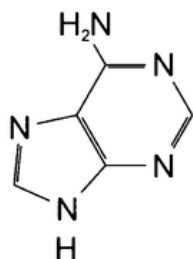


[Rosalind Franklin](#)

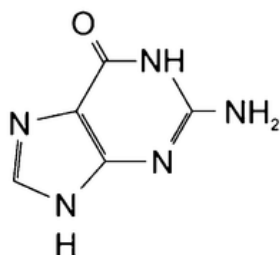
1957 yilda amerikaliklar Aleksandr Rich, Gari Felsenfeld va Devid Devis uchta ipdan tashkil topgan nuklein kislotani tasvirlashdi. 1985-1986 yillarda Moskvada Maksim Davidovich Frank-Kamenetskiy ikkita emas, balki uchta DNK zanjiridan tashkil topgan H-shaklini isbotlashgan va qanday hosil bo'lishini ko'rsatishdi.

Molekulalarning tuzilishi asosida nukleotidlarni tashkil etuvchi asoslar ikki guruhga bo'linadi: purinlar - adenin (A) va guanin (G) bir-biriga bog'langan besh va olti a'zoli geterosikllar orqali hosil bo'ladi va pirimidinlar - sitozin (S) va timin (T) olti a'zodan iborat geterosikl orqali hosil bo'ladi.

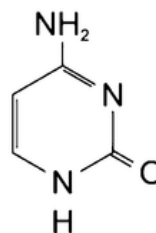
Shuni ta'kidlash kerakki, timin (T) va uratsil (U) ilgari mavjud bo'lgan fikrlarga ko'ra, mos ravishda DNK va RNK da mavjud bo'lishi chegaralanmagan, ba'zi RNK molekulalarining sintezidan so'ng, ushbu molekulardagi uratsillarning katta qismi maxsus fermentlar yordamida timinga aylanib metillanadi. Bunday jarayonlar transport va ribosomal RNKlarda uchraydi. DNK polimeri juda murakkab tuzilishga ega. Nukleotidlar uzun polinukleotid zanjirlarida kovalent ravishda bog'langan.



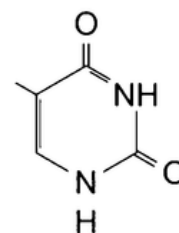
Adenine



Guanine



Cytosine



Thymine

*DNK molekulasida tarkibidagi asosiy azotli asoslari.*

Bu zanjirlar aksariyat hollarda ayrim bir zanjirli DNK-genomiga ega bo'lgan viruslardan tashqari, o'zaro vodorod bog'lari bilan bog'langan bo'ladi va ikkilamchi strukturalarni hosil qilib, ikkilamchi spiral nomini oladi. Har bir zanjir asosi ketma-ket keladigan fosfatlar va shakardan tashkil topgan. Bir zanjir ichki qismidagi qo'shni nukleotidlar fosfodiefir bog'lar bilan bog'langan, ular o'z navbatida bir nukleotidning dezoksiriboza molekulasining (3'—OH) 3'-gidroksil guruhi va boshqanukleotidning (5'—PO<sub>3</sub>) 5'- fosfat guruhi o'rtasida vujudga keladi. DNKning assimetrik zanjirlar uchi 3' - va 5' deb ataladi. DNK sintezida zanjir qutblanishi muhim ahamiyatga egadir, zanjir uzayishi faqatgina yangi nukleotidlarning erkin 3'-uchiga qo'shilishi natijasida yuz beradi.

Ikki karra spiralning kengligi 22 dan 24 A ga tengdir, yoki 2,2-2,4 nm bo'ladi, har bir nukleotid uzunligi 3,3 A yoki 0,33 nm g teng. Shu bilan birga, spiraldagi kichkina 12 A li va katta 22 A li qatorlarni farqlash mumkin. Oqsillar, masalan transkripsiya omillari, ikki zanjirli DNK ning ma'lum ketma-ketliklariga bog'lanadi, asosan katta qatorlarda asoslar uchlari bilan o'zaro ta'sirlashadi.

**Asoslar orasidagi bog'lar.**

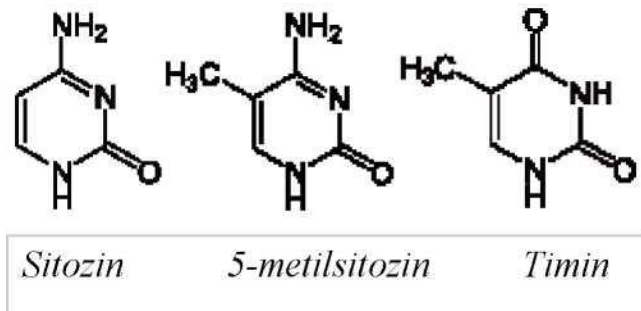
DNK zanjiridgi iplarning biridagi har bir tayanch asos ikkinchi ipning bitta aniq asosiga bog‘lanadi. Ushbu bunday o‘ziga xos majburiy bog‘lanish komplementarlik deb ataladi. Purinlar pirimidinlar bilan bir - birini to‘ldiradi ya’ni ular bilan vodorod bog‘larini hosil qila oladi: adenin faqat timin bilan, sitoziin esa guanin bilan bog‘lanishni hosil qiladi. Ikki karra spiralning bir-birini to‘ldirishi shuni anglatadiki, bir qatorda joylashgan ma’lumotlar ikkinchi qatorda ham mavjud. Bir-birini to‘ldiruvchi asos juftlari o‘rtasidagi o‘zaro ta’sirlarning qaytaruvchanligi va o‘ziga xosligi DNKning replikatsiyasi va tirik organizmlardagi boshqa DNK funktsiyalari uchun muhimdir.

Vodorod bog‘lari kovalent bo‘lmaganligi sababli, ular osonlikcha buziladi va tiklanadi. Ikki karra spiralning zanjirlari fermentlar- helikaza ta’sirida yoki yuqori harorat ta’sirida ajralib ketishi mumkin. Turli xil asosli juftliklar har xil miqdordagi vodorod bog‘lanishlarini hosil qiladi. A-T ikkita, G-C esa uchta vodorod bog‘i bilan bog‘lanadi, shuning uchun G-C ni uzish uchun ko‘proq energiya talab qilinadi. G-C juftlarining ulushi va DNK molekulasining uzunligi zanjirning dissotsiyalanishi uchun zarur bo‘lgan energiya miqdorini aniqlaydi: tarkibida G-C miqdori yuqori bo‘lgan uzun DNK molekulalari ko‘proq chidamli bo‘ladi. DNK molekulalarining ushbu funktsiyalari tufayli osonlik bilan ajralishi kerak bo‘lgan qismlar, masalan, bakterial promotorlarda TATA ketma-ketligi, odatda ko‘p miqdordagi A va T ni o‘z ichiga oladi.

### **Azotli asoslarning kimyoviy modifikatsiyasi.**

DNK tarkibidagi azotli asoslar kovalent ravishda o‘zgartirilishi mumkin, bu gen ekspressiyasini boshqarishda qo‘llaniladi. Masalan, umurtqali hayvonlar hujayralarida sitoziin metillanishi natijasida 5-metiltsitoziin hosil bo‘lishi somatik hujayralar tomonidan gen ekspressiyasini qiz hujayralariga yetkazish uchun ishlatiladi. Sitoziin metillanishi DNK juft spiralidagi asoslarning juftlanishiga ta’sir qilmaydi. Umurtqali hayvonlarda somatik hujayralaridagi DNK metillanishi CG ketma-ketligidagi sitoziin metillanishi bilan cheklanadi. O‘rtacha metillanish darajasi turli xil organizmlarda farq qiladi, masalan, *Caenorhabditis elegans* nematodasida sitoziin metillanishi kuzatilmaydi va yuqori metillanish darajasi umurtqali hayvonlarda uchraydi- 1% gacha. Asoslar boshqa modifikatsiyalari bakteriyalarda adenin metillanishi jarayoni va kinetoplastlarda “J-asoslaming” glikozillirlanishi uchraydi.

Genning promotor qismida 5-metilsitozin hosil bo'lishi bilan sitoziinni metillashtirish uning faol bo'lmagan holati bilan o'zaro bog'liq. Sitozin metillanishi sut emizuvchilarda X-xromosomaning inaktivasiyasi uchun muximdir. DNK metillanishi genom imprintingida ishlatiladi.

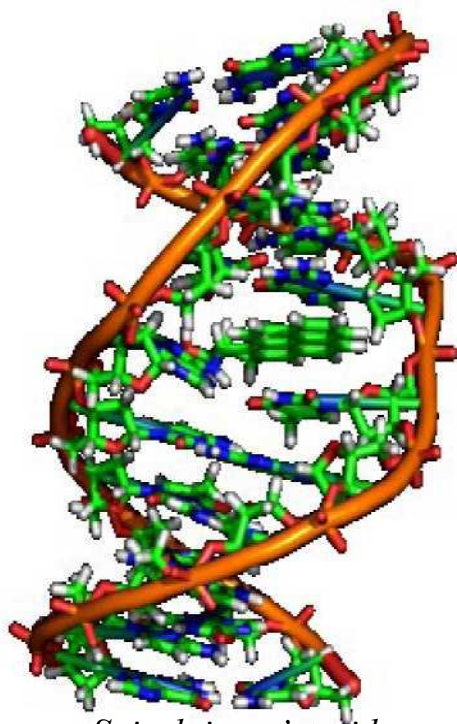


7-rasm. Sitozin, 5-metilsitozin va timinning tuzilishi. Timin 5-metilsitozinni dezaminirlanishi natijasida paydo bo'lishi mumkin

Kanserogenez jarayonida DNK metillanish profili buziladi. Biologik roliga qaramasdan, 5-metilsitozin spontan ravishda amin guruhini yo'qotish va timinga aylanish xususiyatiga ega, shuning uchun metillangan sitozinlar yuqori sondagi mutatsiyalar hosil bo'lishi manbai hisoblanadi (7-rasm).

### DNK ning shikastlanishi.

DNKga turli xil mutagenlar zarar yetkazishi mumkin, ular tarkibiga oksidlovchi va alkillovchi moddalar, shuningdek yuqori energiyali elektromagnit nurlanish - ultrabinafsha va rentgen nurlari kiradi. DNK zararlanish turi mutagen turiga bog'liq. Masalan, ultrabinafsha nurlari DNKga zarar yetkazadi, ular tarkibidagi timin dimerlarini hosil qiladi, ular qo'shni asoslar o'rtasida kovalent bog'lanish hosil bo'lishidan paydo bo'ladi.



*Spiralning o'rtasida joylashgan interkallatsiyalangan kimyoviy birikma benzopiren, tamaki tutunidagi asosiy mutagen*

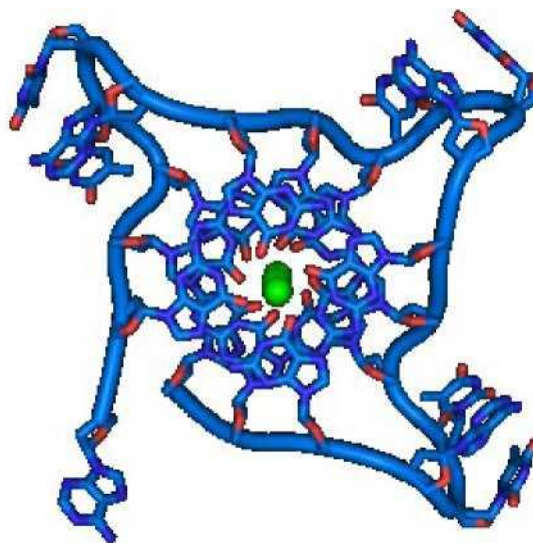
Erkin radikallar yoki vodorod peroksid kabi oksidlovchilar DNKning bir necha turdagi zararlanishiga olib keladi, shu jumladan asoslarda modifikatsiyalar, xususan, guanozin va ikki zanjirli DNK sinishiga sababchi bo'lishi mumkin. Ba'zi hisob-kitoblarga ko'ra, har bir inson hujayrasi kuniga 500 ga yaqin asos oksidlovchi birikmalar tomonidan zarar ko'radi (8-rasm). Zararlarning har xil turlari orasida eng xavflisi qo'sh zanjirda hosil bo'ladigan uzilishlardir, chunki ularni tiklash qiyin va xromosoma mintaqalarining yo'qolishiga- o'chirilishiga va translokatsiyaga olib kelishi mumkin.

*Superspirallashish.* Superspirallashish asosida DNK zanjirining uchlarining yo‘nalishlarida burilishi uning qisqarishi, “super spirallar” hosil bo‘lishi tushuniladi. Oddiy holatlarda DNK zanjiri har 10, 4 juft asosi uchun bitta burilishni amalga oshiradi, ammo o‘ta o‘ralgan holatda ham bo‘lishi mumkin. Super- burilishning ikki turi mavjud: ijobiy - asoslar bir-biriga yaqinroq joylashgan normal burilish yo‘nalishi bo‘yicha va qarama-qarshi yo‘nalishda buralgan ham bo‘lishi mumkin-salbiy superburalish deb ataladi. Tabiatda DNK molekullari odatda salbiy superburalishi uchraydi, bu jarayonlar asosan fermentlar- topoizomerazalar tomonidan amalga oshiriladi. Bu fermentlar DNKda transkripsiya va replikasiya jarayonida hosil bo‘lgan qo‘shimcha supersprallrni olib tashlaydi. Xromosomalarning uchlaridagi strukturalar. Chiziqli xromosomalarning uchida DNKning telomerlari deb nomlangan ixtisoslashgan tuzilmalari joylashgan. Ushbu mintaqalarning asosiy vazifasi xromosomalar uchlari yaxlitligini saqlashdir. Telomerlar, shuningdek DNK uchlarini endonukleazalar degradasiyasidan saqlaydi va reparasiya tizimining faolligini olidini oladi. DNK-polimeraza fermentlari xromosomalar 3’-uchini replikasiyasini amalga oshira olmagan uchun, bu jarayonni maxsus ferment-telomeraza amalga oshiradi (10-rasm). Inson hujayralarida telomerlar ko‘pincha bitta zanjirli DNK bilan ifodalanadi va TTAGGG ketma-ketligining bir necha ming takrorlanishlaridan tashkil topgan. Bu ketma-ketliklarda yuqori miqdorda guanin asoslari uchraydi, ular xromosoma uchlarini stabillashtiradi, va g‘ayrioddiy strukturalar – G-kvadruplekslar hosil qiladi, ular o‘zaro ta’sirda bo‘lgan ikkita emas, balki to‘rtta asoslardan tashkil topgan bo‘ladi. To‘rtta guanin asoslarining atomlari, bitta tekislikda joylashadi, asoslar o‘rtasidagi stabillashgan vodorod bog‘lari yordamida markazda metal ioni-ko‘pgina hollarda kaliy bilan xelatlangan plastinka hosil qiladi. Ushbu plastinkalar bir-birining ustiga joylashtirilgan bo‘ladi. Xromosomalarning uchida boshqa tuzilmalar ham paydo bo‘lishi mumkin: asoslar bir zanjirda yoki turli parallel zanjirlarda joylashgan bo‘lishi mumkin. Bunday strukturalardan tashqari telomerlar, T-halqalar yoki telomer halqalar deb ataladigan halqasimon tugunli strukturalar hosil qilish xususiyatiga egadir. Bularda bir zanjirli DNK stabillashgan telomer oqsillari yordamida keng halqa shaklida joylashadi. T- tugun oxirgi uchida bir zanjirli telomer DNK ikki zanjirli DNK ga qo‘shiladi, va ushbu molekulada zanjirlar qo‘shilishini buzadi, va bog‘ar faqat bir



zanjir bilan bog‘langan holda qoladi. Bunday uch zanjirli tuzilma D-tugun den ataladi, ingliz tilidan displacement loop ma’nosini anglatadi.

Nukleotidlar ketma-ketligi turli xil RNK turlari to‘g‘risidagi ma’lumotlarni kodlaydi: informatsion yoki axborot-mRNA, ribosomal-rRNK va transport-tRNK. Ushbu turdagi RNKlarning barchasi transkripsiya jarayonida DNKdan sintezlanadi. Ularning oqsil biosintezidagi roli turlicha, m-RNK oqsil tarkibidagi aminokislotalar ketma-ketligi haqidagi ma’lumotlarni o‘z ichiga oladi, ribosomal RNK ribosomalar uchun asos bo‘lib xizmat qiladi, murakkab nukleoproteinlar bo‘lib, ularning vazifasi m-RNK asosida aloxida olingan aminokislotalardan oqsillarni shakllantiradi, t-RNK aminokislotalarni oqsil shakllanadigan manzilga- ribosomalar faol markaziga yetkazib berishda ishtirok etadi.



10-rasm Telomerlarning tuzilishi. Markazda xelatlangan metall ionni yashil rangda ko‘rsatilgan.



STAT3 transkripsiya omilining DNK bilan o‘zaro ta’siri (ko‘k spiral shaklida ko‘rsatilgan)

*Oqsillar bilan o‘zaro ta’sir.* DNKning barcha funksiyalari uning oqsillar bilan o‘zaro ta’siriga bog‘liq. O‘zaro ta’sirlar nospesifik bo‘lishi mumkin, har qanday DNK molekulasiga oqsil qo‘shilganda yoki ma’lum bir ketma-ketlikning mavjudligiga bog‘liq (13-rasm). Shuningdek, fermentlar DNK bilan o‘zaro ta’sirlashishi mumkin, ulardan eng muhimi RNK-polimerazalar bo‘lib, ular DNK asoslarining ketma-ketligini RNK molekulasini transkripsiya jarayonida, yoki yangi DNK zanjirini – replikatsiya jarayonida sintez qiladi.

*Strukturaviy va boshqaruvchi oqsillar.* DNK ning nukleotidlar ketma- ketligiga bog‘liq bo‘lmagan, oqsillarning DNK bilan o‘zaro ta’siri, ularning struktur oqsillar bilan o‘zaro ta’siri hisoblanadi. Hujayrada DNK ushbu oqsillar bilan bog‘langan bo‘ladi va kompakt strukturani hosil qiladi, bunga xromatin deb ataladi. Eukariotlarda xromatin DNKga katta bo‘lmagan ishqoriy xususiyatga ega bo‘lgan oqsil-gistonlarning bog‘lanishi natijasida hosil bo‘ladi, prokariotlarda esa xromatin biroz tartibsiz bo‘lib, giston-o‘xshash oqsillar bilan bog‘langan. Gistonlar disksimon oqsil strukturalarni shakllantiradi ularga nukleosomalar deb ataladi, gistonlar xar biriga DNK ikki barobar spiral shaklida o‘raladi. Gistonlar va DNK o‘rtasidagi nospesifik

bog‘lar gistonlar tarkibidagi ishqorli xususiyatga ega bo‘lgan aminokislotalarning va DNK shakar fosfat karkasining kislotali qoldiqlarining ion bog‘lari hisobiga hosil bo‘ladi. Bu aminokislotalar kimyoviy modifikatsiyalari metillanish, fosfolirlanish va asetillanish hisobiga ro‘y beradi. Bu kimyoviy modifikatsiyalar DNK va gistonlar o‘rtasidagi o‘zaro ta’sir kuchini o‘zgartiradi, transkripsiya omillari uchun xizmat qiladigan maxsus ketma-ketliklarning transkripsiya jarayoniga moyilligini oshiradi va transkripsiya tezligini o‘zgartiradi. Xromatin tarkibidagi nospesifik ketma-ketliklarga bog‘langan boshqa oqsillar gel elektroforezda yuqori harakatchan bo‘lib, buralgan DNK ning katta miqdorini tashkil qiladi. Bu oqsillar xromatin tarkibida yuqori tartibdagi strukturalarni tashkil qilishda katta ahamiyatga ega. DNK ga birikadigan boshqa guruh oqsillariga- bir zanjirli DNK bilan bog‘lanadigan oqsillar kiradi. Insonda bu guruhga kiruvchi oqsillardan yahshi o‘rganilgan oqsillarga- replikasining A oqsili hisoblanadi, ushbu oqsil ikki zanjir ishtirok etadigan zanjir ochilishi bilan bog‘liq bo‘ladigan jarayonlarda shu jumladan, rekombinasiya va reparasiya jarayonlarida ishtirok etadi. Bu guruh oqsillari bir zanjirli DNKni halqa tugunlar hosil bo‘lishi va nukleazalar ta’sirida degradasiya bo‘lishidan saqlaydi. Shuningdek, boshqa oqsillar spetsifik ketma-ketliklarga bog‘lanadi va tanishda ishtirok etadi. Bunday yahshi o‘rganilgan guruh oqsillariga-turli xil transkripsiya omillari kiradi, ya’ni transkripsiya jarayonini boshqaradigan oqsillar. Har bir oqsillar ma’lum ketma- ketlikni taniydi, asosan promotor qismlarda va genlar transkripsiyasini yoki bostiradi, yoki faollashtiradi. Bunday holatlar asosan transkripsiya omillarining RNK-polimeraza bilan to‘g‘ridan-to‘g‘ri assotsiasiyasi yoki, yordamchi oraliq- vositachi oqsillar yordamida yuz beradi. Polimeraza oqsillar bilan avvaliga assotsiyalashadi, keyin transkripsiya boshlanadi. Boshqa xollarda transkripsiya omillari promotor qismlardagi joylashgan gistonlarni modifikaslaydigan fermentlar bilan bog‘lanadi va DNK ni polimerazaga ochib beradi. Spetsifik ketma- ketliklar genom ko‘p qismida uchragani uchun bitta transkripsiya omili tipining faolligining o‘zgarishi minglab genlar faolligining o‘zgarishiga olib kelishi mumkin. Shunga ko‘ra, bu oqsillar atrof-muhit o‘zgarishiga organism javob reaksiyasida, organism rivojlanishida va hujayralar differensirovkasida tez-tez boshqarilib turiladi. Transkripsiya omillarining DNK bilan o‘zaro ta’sirining maxsusligi, aminokislotalar va DNK asoslari o‘rtasidagi munosabatlar orqali ta’minlanadi, bu esa o‘z navbatida DNK ketma-ketliklarini o‘qish imkonini yaratadi. Asoslar bilan o‘zaro ta’sirlar asosiy halqada o‘tadi, bu yerda asoslar transkripsiya jarayonining o‘tishi uchun ochiq bo‘ladi.

**DNK ni modifikatsiyalovchi fermentlar.** Hujayrada DNK kompakt, yani superspirallashgan holatda joylashadi, boshqa holatda hujayra yadrosiga sig‘mas edi. Asosiy hayotiy jarayonlari o‘tishi uchun DNK ochilishi kerak bo‘ladi, ushbu jarayon ikki hil oqsil guruhlari- topoizomerazalar va xelikazalar tomonidan amalga oshiriladi.

*Topoizomerazalar*- nukleazalar va ligazalar faolligiga ega bo'lgan fermentlardir. Ular DNK ning superburalish darajasini o'zgartirish xususiyatiga ega. Bu fermentlarning ayrimlari DNK spiralini kesadi va hosil bo'lgan bitta zanjirning harakatlanishiga imkon yaratadi, superburalish darajasini kamaytiradi, keyinchalik yana kesilgan joyni birlashtirib qo'yadi. Boshqa fermentlar ikki zanjirdan birini kesadi, hosil bo'lgan bitta zanjir kesilgan joydan o'tish imkoniga ega bo'ladi, keyinchalik birinchi zanjirdagi kesilgan joyni yana ligirlaydi. Topoizomerazalar DNK bilan bog'liq bo'lgan - transkripsiya va replikasiya jarayonlarida ishtirok etadi.

*Xelikazalar*- molekulyar motor oqsillari hisoblanadi. Ular nukleotiduchfosfatlarning va ko'pgina hollarda ATF kimyoviy energiyasidan foydalanib, asoslar o'rtasidagi vodorod bog'larini uzadi, ikki zanjirni alohida olingan zanjirlarga ajratish xususiyatiga ega.

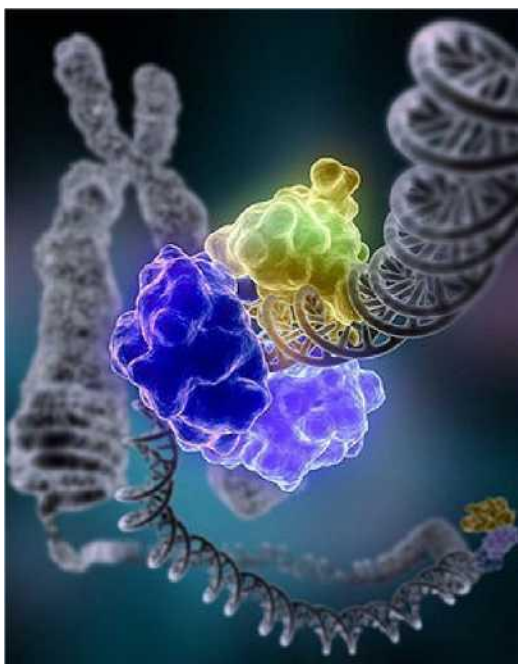
*Polimerazalar*. DNK metabolizmi uchun kerakli bo'lgan fermentlar guruhi mavjud bo'lib, ular nukleozidtrifosfatlardan polinukleotid zanjirini sintez qiladi, bularga DNK-polimerazalar deb ataladi. Ular DNK zanjiridagi oldingi nukleotidning 3'-gidroksil guruhiga yangi nukleotidlarni qo'shish xususiyatiga ega. hamma polimerazalar 5'--> 3' yonalishda faoliyat ko'rsatadi. Ferment faol markazida substrat-nukleozidtrifosfat-komplementar asos bilan bir zanjirli polinukleotid zanjirdagi-matrisa bilan juftlashadi. DNK replikasiyasi jarayonida DNK bog'liq DNK polimeraza boshlang'ich DNK ketma-ketligining nusxasini sintezlaydi. Bu jarayonda polimerizasiya aniqligi juda katta ahamiyatga ega bo'lib, xatoliklar mutatsiyalarga olib keladi, shuning uchun ko'pgina polimerazalar mustaqil ravishda "tahrirlash" xususiyatiga egadir, xatoliklarni bartaraf eta oladi.

Polimeraza sintez davrida xatoliklarni noto'g'ri nukleotidlar o'rtasidagi juftlashishlarning bo'lmasligi orqali taniydi. Juftlashish yo'qligi aniqlangandan so'ng, fermentning 3'--> 5' yo'nalishdagi ekzonukleazali faolligi oshadi va noto'g'ri juftlashish joylarini kesib tashlanadi. Ko'pgina organizmlarda DNK-polimerazalar katta majmua sifatida faoliyat ko'rsatadi, ularga replisomalar deb ataladi va ular tarkibida ko'plab qo'shimcha subbirliklar bo'lib, xelikazalar ham ular tarkibida faoliyat ko'rsatishi mumkin.

RNK bog'liq DNK polimerazalar- polimerazalarning maxsuslashgan turi bo'lib, RNK matrisasida DNK sintezini amalga oshiradi. Bu polimerazalar turiga teskari transkriptazalar kiradi, ushbu fermentlar retroviruslarda uchraydi va ular hujayrani zararlantirganda ferment faollashadi, xuddi shuningdek telomeraza, telomerlar replikasiyasi uchun foydalaniladi. Telomeraza- noan'anaviy ferment turlariga kiradi, chunki xususiy m-RNKga ega bo'ladi. Transkripsiya DNK-bog'liq RNK-polimeraza tomonidan amalga oshiriladi, yani DNK matrisasi asosida bir zanjirli m-RNK nusxa tushiradi. Gen transkripsiyasi boshlanishida RNK- polimeraza genning boshlang'ich



promotor qismiga bog‘lanadi, va DNK zanjirini ochadi. Keyinchalik gen ketma-ketligini m-RNKga terminator qismiga yetguncha



*Zararlangan DNK zanjirini bog‘laydigan DNK ligaza I (turli xil ranglarda ko‘rsatilgan bir nechta bir xil oqsil molekulalaridan tashkil topgan halqa shaklidagi tuzilish)*

nusxa ko‘chiradi, terminator qismida to‘xtaydi va DNK dan ajraladi. DNK- bog‘liq DNK polimeaza singari, RNK- polimeraza II inson genomida genlarning ko‘p qismini transkripsiyasini amalga oshiradi, ko‘p miqdordagi oqsil majmuasi- boshqaruvchilar va qo‘shimcha birliklar tarkibida faoliyat ko‘rsatadi.

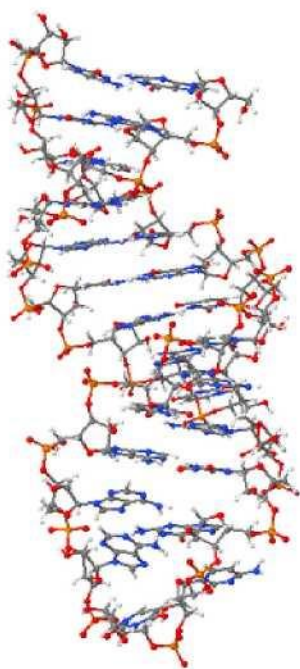
*Nukleazalar va ligazalar.* Hujayrada sodir bo‘ladigan turli xil jarayonlar, masalan, rekombinatsiya va reparasiya jarayonlarida, DNK zanjirlarini kesishi va yaxlitligini tiklashi (ligaza, 14-rasm) mumkin bo‘lgan fermentlarni o‘z ichiga oladi. DNKni kesuvchi fermentlarga nukleazalar deyiladi. DNK molekulasining uchlarida nukleotidlarni gidrolizlaydigan nukleazalar ekzonukleazalar deb ataladi va endonukleazalar DNKni zanjir ichida kesib tashlaydi. Molekulyar biologiya va genetik muhandislikda eng ko‘p ishlatiladigan nukleazalar DNKni ma‘lum ketma- ketliklar

qismida kesuvchi restriksion endonukleazalar ishlatiladi. Masalan, EcoRV fermenti - E. colidagi ferment olti nukleotidli 5’-GAT | ATC-3’ ketma-ketligini taniydi va DNKni vertikal holda ikki zanjir joyida kesadi. Tabiatda bu fermentlar bakteriyalarni bakteriofaglar bilan zararlanishdan, bakterial hujayraga kiradigan fagni kesish orqali himoya qiladi. Bu holatlarda nukleazalar-restriksiya- modifikasiyalar tizimining asosiy qismi sifatida ishtirok etadi. DNK-ligazalar DNK fragmentlarining uchlarini tikadi, ATF energiyasidan foydalangan holda, fosfodiefir bog‘larini katalizlaydi. Restriksion nukleazalar va ligazalar klonlashda va finjerprintingda qo‘llaniladi.

**Ribonuklein kislota.**Ribonuklein kislota- uchta asosiy makromolekulalardan biri bo‘lib, hamma tirik organizmlar hujayralarida uchraydi va kodlashda, irsiy materialni o‘qishda, genlarni boshqarishda va tasvirlashda ishtirok etadi.

DNK singari, RNK uzun zanjirdan iborat, undagi har bir zveno nukleotid deb ataladi. Har bir nukleotid azot asosidan, shakar va fosfar guruhidan iborat. RNK ketma-ketligi irsiy ma‘lumotni kodlash imkonini beradi. Hamma tirik organizmlar RNKni oqsil sintezini dasturlashda ishlatiladi. Hujayra RNKlari transkripsiya jarayonida, ya‘ni maxsus fermentlar RNK-polimerazalar asosida DNK matritsada asosida RNK hosil

bo'ladi. Keyinchalik matrisali RNKlar translyasiya jarayonida ishtirok etadilar. Translyasiya- bu m-RNK matrisasida oqsil sintezlanishi hisoblanadi. Boshqa RNKlar transkripsiyadan so'ng, kimyoviy modifikasiyalarga uchraydi, undan so'ng hosil bo'lgan ikkilamchi va uchlamchi strukturalar RNK tipiga qarab turli xil vazifalarni bajaradi. Bir zanjirli RNKlar uchun turli xil fazoviy strukturalar xos, ularda bitta zanjirning nukleotidlar bir qismi o'zaro juftlashgan bo'ladi. Ayrim yuqori sturturali RNKlar oqsil sintezida ishtirok etishadi, masalan t-RNKlar kodonlarni tanishda va aminokislotalarni oqsil sintezi jarayonida manziliga yetkazishda, ribosomal RNKlar esa ribosomalar strukturaviy va katalitik vazifalarni ishtirok etishadi. Shu bilan birga, zamonviy tasavvurlarga ko'ra, RNKlar vazifalari faqatgina translayasiya jarayonlari bilan chegaralanib qolmaydi, yadroviy kichik RNKlar eukariot matrisali RNKlar splayingsida va boshqa jarayonlarda ishtirok etadi. Bundan tashqari, RNKlar ayrim fermentlar tarkibiga kiradi masalan telomerazalar, va ayrim RNKlarning fermentativ faolligi aniqlangan, ya'ni boshqa RNKlarni zanjirini kesish xususiyatiga ega bo'lib, va aksincha RNK fragmentlarini tikish xususiyatiga ham ega bo'ladi. Bunday RNKlar ribozimlar deb ataladi. Ko'pgina viruslar genomi RNK dan tashkil topgan, ya'ni DNK rolini bajaradi. Hujayradagi RNKlarning turli tuman funksiyasiga qarab, hujayra biologiyasida RNK biologik tizimlar paydo bo'lgunga qadar birinchi irsiy ma'lumotni tashuvchi molekulalar degan gipoteza mavjuddir.



Pro [mRNK halqasi](#).  
Asoslardagi [azot](#)  
atomlari ko'k  
rangda, fosfat  
asosidagi [kislород](#)  
qizil

Oqsillar sintezida RNK roli 1939 yilda Torbbern Oskar Kaspersson, Jan Brashe va Djek Shuls ishlarida taklif qilingan. Jerard Mairbaks birinchi marta quyon gemoglobinini kodlaydigan, matrisali RNK ni ajratib olingan, va ushbu RNK oositlarga o'tkazilganda, xuddi shu oqsil kodlanishi aniqlangan. 1956-1957 yillarda A.Belozerskiy, A.Spirin, E.Volkin, L.Astraxan tomonidan hujayralardagi RNK tarkibini va hujayradagi RNK lar asosiy massasini ribosomal RNK tashkil qilishini aniqlaganlar. Severo Ochoa 1959 yilda tibbiyot bo'yicha RNK sinteining mexanizmini ochgani uchun Nobel mukofotiga sazovor bo'lgan. 1990-yillarning boshlarida o'simlik genomiga begona genlarning kiritilishi analog genlarining ekspresiyasini bostirishga olib kelishi aniqlandi. Shu bilan birga 22 nukleotid qatoriga ega bo'lgan RNK lar mikro RNK lar bo'lib, S. elegans nematodalar ontogeneizida boshqaruvchi rolini bajarishi isbotlangan (16-rasm).

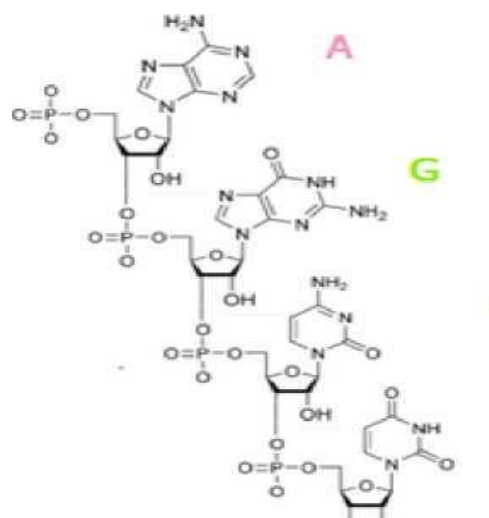
*Monomerlarning kimyoviy tarkibi va modifikatsiyalari.* RNK nukleotidlari tarkibida shakar –

riboza bo‘lib, unga 1’ holatida asoslar- adenin, guanin, sitozin yoki uratsillardan biri birikkan bo‘ladi. Fosfat guruhi ribozani zanjirga bog‘lab, bir ribozaning 3 ‘uglerod atomi bilan, ikkinchisining 5’ holatida bog‘lanish hosil qiladi. Fosfat guruhlari fiziologik pH ko‘rsatkichida salbiy zaryadlangan, shuning uchun RNK polianion xususiyatga ega. RNK to‘rt asosli - adenin (A), guanin (G), uratsil (U) va sitozinli (S) polimer sifatida transkripsiyalanadi. Ammo "yetuk" RNKda ko‘plab o‘zgartirilgan asoslar va shakar mavjud. Hammasi bo‘lib, RNK da 100 ga yaqin modifikasiyalangan turli xil nukleotidlar mavjud bo‘lib, ulardan 2’- metilriboza shakarning ko‘p uchraydigan modifikasiya variantlari hisoblanadi, psevdouridin esa -asoslarning modifikasiyalari hisoblanadi. Psevdouridin, riboza bilan C-N emas, balkim C-C bog‘lari orqali bog‘lanadi va RNK molekulasida turli xil xolatlarda uchraydi. Shuningdek, psevdouridin, t-RNK ning funksiyasi uchun muhim hisoblanadi. Ikkinchi modifikasiya turlariga gipoksantin kiradi, adeninning dezaminirlanishi natijasida hosil bo‘ladi, va nukleozid inozin nomi bilan ataladi. Inozin genetik kodnig turg‘unligini ta‘minlashda katta ahamiyatga egadir. Boshqa ko‘plab uchraydigan modifikasiyalarning roli oxirigacha o‘rganilmagan. Masalan, peptid bog‘larni hosil qiluvchi ribonukleotidlar roli shular jumlasiga kiradi.

*Tuzilishi.* RNKdagi azotli asoslar sitozin va guanin, adenin va uratsil, shuningdek, guanin va uratsil o‘rtasida vodorod bog‘larini hosil qilishi mumkin. Lekin boshqa ta‘sirler ham mavjud bo‘lib, bularga bir nechta adenin nukleotidlarining o‘zaro tugunlar hosil qilishi, yoki to‘rtta adenin o‘zaro tugunlar hosil qilib o‘zaro ta‘sirler adenin-guanin o‘rtasida bo‘lishini misol keltirish mumkin (17-rasm).

RNKning uni DNKdan ajratib turadigan muhim tarkibiy xususiyati ribozaning 2’ holatida gidroksil guruhining mavjudligidir, bu RNK molekulasining B-konformatsiyasida emas, balki A konformasiyada ega bo‘lishiga imkon beradi, bu holat ko‘pincha DNKda kuzatiladi. DNK ning A- shaklida chuqur- tor qator va chuqur bo‘lmagan, keng kichik qatorli struktura uchraydi.

RNK tarkibidagi 2’ gidroksil guruhining mavjudligi, konformatsion plastik, ya’ni spiral hosil bo‘lishida ishtirok etmaslik imkonini beradi, RNK molekulasining hududlari boshqa fosfat bog‘lanishlariga kimyoviy ta‘sir ko‘rsatishi va ularni uzishida ishtirok etadi. Bir zanjirli RNK ning “ishchi” shakli, oqsillar singari, uchlamchi strukturaga ega bo‘ladi. Uchlamchi struktura ikkilamchi struktura elementlari asosida



OH  
RNK polinukleotidning kimyoviy tuzilishi

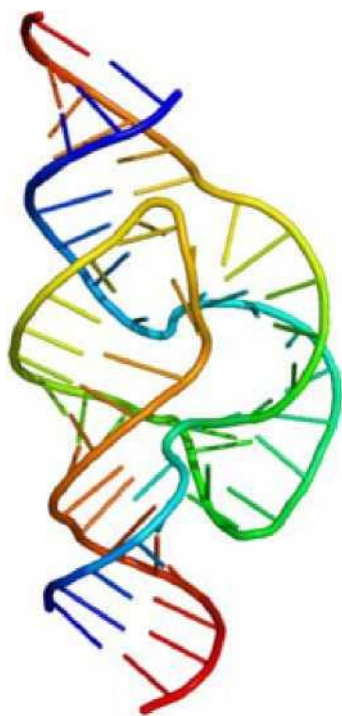




sintezning tugashini belgilaydi. Ko'pgina RNK molekulalari boshlang'ich molekulalar sifatida sintezlanadi, ular "tahrirlash" dan o'tadi - RNK- oqsil komplekslari yordamida keraksiz qismlari olib tashlanadi.

Masalan, *Escherichia coli*- da rRNK genlari bitta operonda joylashgan- lokusda joylashish tartibi quyidagicha: 16S - tRNK Glu<sub>2</sub> - 23S-5S, bitta uzun molekula sifatida o'qiladi, so'ngra avval rRNK hosil bo'lishi bilan bir nechta mintaqalarga bo'linadi, va keyin yetuk rRNK molekulalari hosil bo'ladi. Sintezdan so'ng RNKning nukleotidlar ketma-ketligini o'zgartirish jarayoni RNKni qayta ishlash yoki tahrirlash deb ataladi.

Transkripsiya tugaganidan so'ng, RNK molekulalari bajaradigan funksiyalariga qarab modifikasiyaga uchraydi. Eukariotlarda RNK ning yetilish jarayoni ya'ni oqsil sinteziga tayyorlanishi, splaysing jarayonini o'z ichiga oladi: oqsilni kodlamaydigan ketma-ketliklarning intronlarning splaysasoma ribonukleotidi orqali olib tashlanishi ro'y beradi. Keyinchalik eukariotlarda boshlang'ich i-RNK ning 5' uchiga maxsus modifikasiyalangan nukleotid cap qo'shiladi, 3' uchiga esa bir nechta adenin qo'shiladi, ularga poli A- uchi deb ataladi.



*19-rasm. RNKni ajratib turadigan bolg'a ribozimasining tuzilishi*

Matrisali RNK- bu RNK turi DNK dagi kodlangan ma'lumotni ribosomalarga tirik organizmlarning oqsil sintezlovchi molekulyar mashinalarga o'tkazishda vositachi vazifasini bajaradi. i-RNK kodlaydigan ketma-ketliklari oqsil tarkibidagi aminokislotalar ketma-ketliklarini belgilaydi. Lekin ko'p RNK lar oqsilni kodlamaydi. Bu kodlamaydigan RNKlar alohida olingan genlar tomonidan transkripsiyalanadi, masalan, ribosomal RNK lar yoki intronlar hosilalari bo'lishi mumkin. Klassik, yahshi o'rganilgan kodlamaydigan RNK turlari - bular transport RNK va ribosomal RNK lar bo'lib, translatsiya jarayonida ishtirok etadi. RNK ning i-RNK prosesingida va boshqa vazifalarni bajaruvchi genlarni boshqarishda ishtirok etadigan sinflari ham mavjud (19-rasm). Bundan tashqari, RNK ning kodlamaydigan molekulalari mavjud bo'lib, RNK molekulalarini ligirlash va kesish kabi kimyoviy reaksiyalarni katalizlash xususiyatiga ega. Oqsillar kabi, enzimlar bilan kimyoviy reaksiyalarni

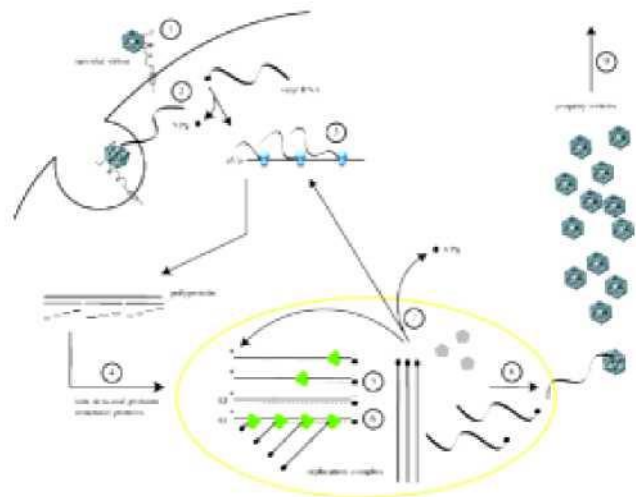
katalizlaydi, va RNK ning katalitik molekulalari ribozimlar deb ataladi.

Kichik interferensiyalaydigan RNKlar o'hshash mikro-RNK mexanizmlarining RNK-interferensiyasi orqali harakat qilishadi. Hayvonlarda RNK ning yangi turi

topilgan bo‘lib, ular pi-RNK 29-30 nukleotiddan iborat bo‘lib, Piwi bilan o‘zaro ta’sirlasha oladi, jinsiy hujayralarda transpozonlar nusxalarining sonining ko‘payishiga yo‘l qo‘ymaydi va gametalarni hosil bo‘lishida asosiy rol o‘ynaydi. Bundan tashqari, pi-RNK epigenetik tarzda ona tomondan irsiylanadi, va avlodlarga transpozonlarni ekspressiyasini ingibirleydigan xususiyatlarning irsiylanishini ta’minlaydi. Ma’nosiz RNKlar bakteriyalarda keng tarqalgan bo‘lib, i-RNK ga birikadi va genlarni namoyon etuvchi xususiyatlarini bostiradi, lekin bir hillari ekpressiyasini faollashtiradi. Bunday RNKlar i-RNK ga birikkanda, RNK ning ikki zanjirli molekulasi hosil bo‘ladi va fermentlar tomonidan degradasiya qilinadi. Eukariotlarda yuqori molekulyar - RNK ga o‘hshash RNK molekulalari topilgan, ular oqsil molekulasini kodlamaydi. Ular ham genlarning namoyon bo‘lishida ishtirok etadi. Misol tariqasida sut emizuvchilarda Xist genini keltirish mumkin, X-xromosomalarning biriga birikadi va faolsizlantiradi.

Genlarni boshqarishda individual molekulalarning rovidan tashqari, regulyator elementlar mRNKning translyasiya qilinmagan 5’ va 3’ mintaqalarida hosil bo‘ladi. Ushbu elementlar translyasiyaning boshlanishiga to‘sqinlik qilib, mustaqil ravishda harakat qilishi yoki oqsillarni biriktirishi mumkin. *RNK processing*. Ko‘pgina RNKlar boshqa RNKlar modifikatsiyasida ishtirok etadi. Intronlar boshlang‘ich iRNKdan splaysasoma yordamida kesiladi, splaysasomalarda tarkibida oqsillardan tashqari bir qancha kichik yadro RNKlari ham bo‘ladi. Bundan tashqari, intronlar o‘zlarining xususiy kesilishini katalizlaydi.

Transkripsiyani natijasida sintezlangan RNK kimyoviy jihatdan modifikatsiya qilinishi mumkin. Eukariotlarda RNK nukleotidlarining kimyoviy modifikatsiyasi, masalan, ularning metillanishi, kichik yadro RNKlari - 60-300 nukleotidlardan tashkil topgan ketma-ketliklar tomonidan amalga oshiriladi. RNKning bu turi yadrochada va Kahal tanachalarida joylashgan. Kichik yadro RNKlari fermentlar bilan assotsiatsiyasi yuz bergandan so‘ng, ular RNK-nishon bilan bog‘lanadi ikki molekula asoslari orasida juftliklar hosil qiladi, fermentlar esa RNK-nishonlarining nukleotidlarini modifikatsiya qiladi. Ribosomal va transport RNKlar ko‘pgina



22-rasm. Poliovirus misolida RNK genomiga ega bo‘lgan virusning hayot aylanishi : 1 – asl virionni retseptorga biriktirish; 2 – virion hujayraga kiradi; 3 – virus RNKlaridan virus oqsillarini polipeptid hosil bo‘lishi bilan tarjima qilish; 4 – virusli polimerazalar uning RNKini ko‘paytiradi

modifikatsiyalarga ega bo'lib, evolyutsiya davomida saqlanib qoladi.

Oqsil makromolekulalar to'rt xil darajadagi tuzilishga ega - **birlamchi, ikkilamchi, uchlamchi va to'rtlamchi** tuzilishlardan iborat.

### **Birlamchi tuzilma**

Hujayralar tomonidan oqsil qurilishi uchun ishlatiladigan 20 xil standart aminokislotalar mavjud. Aminokislotalar, ularning nomidan ko'rinib turibdiki, asosiy aminokislotalar va kislotali karboksil guruhini o'z ichiga oladi. Bu disfunktsionallik individual aminokislotalarga peptid aloqalarini hosil qilish orqali uzluksiz zanjirlarga qo'shilish imkonini beradi. Bunda bir aminokislota amino -NH<sub>2</sub> gruppasi orasidagi **amid** bog'lari va boshqasining –karboksil-COOH gruppasi bilan bog'langan. Odatda 50 tadan kam aminokislotalarga ega ketma-ketliklar peptidlar deb ataladi.

Yon zanjir o'rnini bosuvchiga qarab, aminokislotalarni kislotali, asosli yoki neytral deb tasniflash mumkin. Odamlarda mavjud bo'lgan turli xil oqsillarni sintez qilish uchun 20 ta aminokislota kerak bo'lsa-da, biz faqat o'ntasini sintez qila olamiz. Qolgan 10 tasi muhim aminokislotalar deb ataladi va ular dietada olinishi kerak. Oqsilning aminokislotalar ketma-ketligi DNKda kodlangan bo'ladi. Oqsillar transkripsiya (information RNK zanjirini yaratish uchun DNK zanjiridan foydalanish - mRNK) va nusxa (mRNK ketma-ketligi aminokislotalar zanjiri sintezini boshqarish uchun shablon sifatida ishlatiladi) deb ataladigan bir qator bosqichlar orqali sintezlanadi. proteinni oshiradi). Ko'pincha, oqsilning biologik funksiyasi uchun zarur bo'lgan glikozillanish yoki fosforillanish kabi post-translatsion modifikatsiyalar sodir bo'ladi. Aminokislotalar ketma-ketligi oqsilning asosiy tuzilishini tashkil etsa-da, oqsilning kimyoviy va biologik xususiyatlari uch o'lchovli yoki uchlamchi darajali tuzilishga juda bog'liq.

**Ikkilamchi tuzilma.** Oqsil yoki peptidlarning cho'zilishi yoki iplarning vodorod bog'lanishiga bog'liq bo'lgan aniq, xarakterli lokal strukturaviy konformatsiyalariga ega yoki ikkilamchi tuzilishli bo'ladi. Ikkilamchi strukturaning ikkita asosiy turi -  $\alpha$ -spiral va  $\beta$ -spiral zanjirdan iborat.

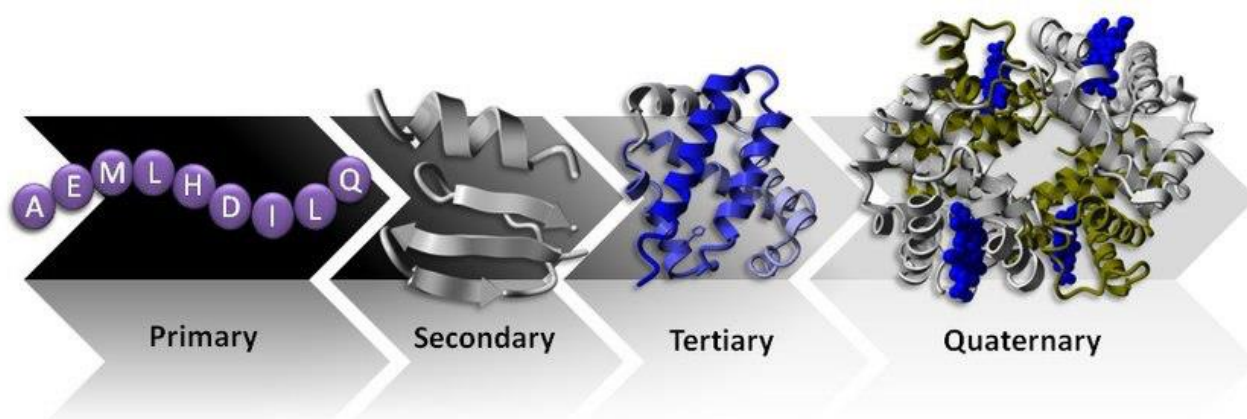
$\alpha$ -spiral o'ng qo'li coiled ipli hisoblanib,  $\alpha$ -spiraldagi aminokislotalar guruhlarining yon zanjirli o'rinbosarlari tashqi tomonga cho'ziladi. Vodorod aloqalari ipdagi har bir C=O bog'ining kislorodi va spiraldagi o'zidan pastda joylashgan to'rtta aminokislotalarning har biri NH<sub>2</sub> guruhining vodorodi o'rtasida hosil bo'ladi. Vodorod aloqalari bu strukturani ayniqsa barqaror qiladi. Aminokislotalarning yon zanjirli o'rinbosarlari NH guruhlarini yonida joylashgan.

$\beta$ -zanjir vodorod bog'lanishi iplar ichida (intra-strand) emas, balki iplar o'rtasida (iplararo) bo'ladi. Cho'yshab ko'rinishdagi konformatsiyasi yonma-yon

joylashgan juft iplardan iborat. Bir zanjirdagi karboksil kislorodlari orqali qo'shni zanjirning aminokislotalari bilan bog'lanadi. Ip yo'nalishlari (N-terminusdan C-terminusgacha) bir xil yoki qarama-qarshi bo'lishiga qarab, ikkita ip parallel yoki antiparallel bo'lishi mumkin. Anti-parallel  $\beta$ -zanjiri yanada yaxshi tafovutlashgan vodorod bog'lari tufayli yanada barqaror holatga kelishidir.

**Uchinchi darajali tuzilma.** Protein molekulasining umumiy uch o'lchamli shakli uchinchi darajali tuzilishdir. Oqsil molekulasi maksimal barqarorlikka yoki eng past energiya holatiga erishadigan tarzda egilib, buriladi. Proteinning uch o'lchovli shakli tartibsiz va tasodifiy ko'rinishi mumkin bo'lsa-da, u aminokislotalarning yon zanjir guruhleri o'rtasidagi bog'lanish o'zaro ta'siri tufayli ko'plab barqarorlashtiruvchi kuchlar tomonidan yaratilgan.

Fiziologik sharoitda fenilalanin yoki izoleytsin kabi neytral, qutbsiz aminokislotalarning gidrofob yon zanjirlari oqsil molekulasining ichki qismiga birikadi, shu bilan birga ularni suvli muhitdan himoya qiladi. Alanin, valin, leytsin va izoleytsinning alkil guruhleri ko'pincha bir-biri bilan gidrofob o'zaro ta'sirlarni hosil qiladi, ko'pincha fenilalanin va tirozin kabi aromatik guruhlar esa birga yig'iladi. Kislotali yoki asosiy aminokislotalarning yon zanjirlari odatda gidrofill bo'lganligi sababli oqsil yuzasiga ta'sir qiladi. Sisteindagi sulfidgidrid guruhlarining oksidlanishi natijasida disulfid bog'larining shakllanishi oqsil zanjirining turli qismlarini kovalent tarzda birga ushlab turishga imkon beruvchi uchinchi darajali oqsil tuzilishini barqarorlashtirishning muhim jihati hisoblanadi. Bundan tashqari, turli yon zanjir guruhleri o'rtasida vodorod aloqalari paydo bo'lishi mumkin.



*Oqsil strukturalarining tuzilishi*

*Disulfid bog'larida* bo'lgani kabi, bu vodorod bog'lari ketma-ketlik jihatidan kichik masofada joylashgan zanjirning ikki qismini birlashtirishi mumkin. Tuz ko'prigi, aminokislotalar yon zanjirlaridagi musbat va manfiy zaryadlangan joylar orasidagi ionli o'zaro bog'lanishlarni, shuningdek, oqsilning uchinchi darajali tuzilishini barqarorlashtirishga yordam beradi.



## To'rtlamchi tuzilish

Ko'pgina oqsillar bir nechta polipeptid zanjirlaridan iborat bo'lib, ular ko'pincha oqsil subbirliklari deb ataladi. Bu subbirliklar gomodimerdagi kabi bir xil yoki geterodimerdagi kabi har xil bo'lishi mumkin. To'rtlamchi struktura bu oqsil bo'linmalarining bir-biri bilan o'zaro ta'sirini va kattaroq agregat oqsil kompleksini hosil qilish uchun o'zini qanday tartibga solishini anglatadi. Oqsil kompleksining yakuniy shakli vodorod-bog'lanish, disulfid-bog'lar va asos bog'larini o'z ichiga olgan turli o'zaro ta'sirlar bilan yana bir bor barqarorlashadi. Oqsil tuzilishining to'rtta darajasi rasmda ko'rsatilgan.

Protein Structure, Stability and Folding, Methods in Molecular Biology, 168, Edited by Kenneth P. Murphy  
Protein Stability and Folding, Theory and Practice, Methods in Molecular Biology, Vol. 40, Edited by Bret Shirley

### Uchinchi savolning bayoni.

BLAST (Basic Local Alignment Search Tool) - asosiy tuzilish (ketma-ketlik) yoki uning bo'lagi ma'lum bo'lgan oqsillar yoki nuklein kislotalarning gomologlarini qidirishda foydalaniladigan kompyuter dasturlari oilasi. BLAST-dan foydalanib, tadqiqotchi o'z ketma-ketligini ma'lumotlar bazasidagi ketma-ketliklar bilan taqqoslashi va taxmin qilinayotgan gomologlarning ketma-ketligini topishi mumkin. Bu molekulyar biologlar, bioinformatika va taksonomlar uchun eng muhim vositadir. BLAST dasturi olimlar Stiven Altschul, Uorren Gish, Uebb Miller, Evgeniy Myers va Devid J. Lipman tomonidan AQShning Milliy Sog'liqni saqlash institutida ishlab chiqilgan va 1990 yilda Molekulyar Biologiya jurnalida nashr etilgan

#### **Blast dasturi oilasi asosan 5 ta gruppasi farqlanadi:**

**Nukleotidli blast** o'rganilgan nukleotidlar ketma-ketligini va ketma-ket joylashgan nuklein kislotalarni -va ularning bo'limlari bazasi bilan taqqoslash uchun mo'ljallangan. Ular o'z navbatida:

- *megablast* - juda o'xshash ketma-ketlikni topish uchun tez taqqoslash,
- *dmegablast* - unchalik o'xshash bo'lmagan ketma-ketlikni qidirish uchun tez taqqoslash,
- *blastn* - o'xshash ketma-ketlikni qidirish uchun sekin taqqoslash va hk.

**Proteinli Blast.** O'rganilgan oqsilning aminokislotalar ketma-ketligini mavjud bo'lgan oqsillar va ularning bo'limlari bazasi bilan taqqoslash uchun mo'ljallangan. Ular o'z navbatida:

*blastp* - shunga o'xshash barcha ketma-ketlikni topish uchun sekin taqqoslash  
*Cdart* - domen tuzilishi bo'yicha gomologik oqsillarni qidirish uchun taqqoslash  
*rpsblast* - konservativ domenlarning ma'lumotlar bazasidagilari bilan taqqoslash

*psi-blast* - bir-biriga o'xshash bo'lmagan ketma-ketlikni qidirish uchun taqqoslash,

*phi-blast* - foydalanuvchi tomonidan belgilangan namunani o'z ichiga olgan oqsillarni qidirish va hk.

**Translyantli Blast.** Nukleotidlar ketma-ketligini aminokislotaga aylantirishi mumkin. Ular o'z navbatida:

*blastx* - o'rganilgan nukleotidlar ketma-ketligini kodlangan aminokislotalarga tarjima qiladi va keyin mavjud aminokislotalar oqsillari bazasi bilan taqqoslaydi;

*Tblastn* - o'rganilgan aminokislotalar ketma-ketligi nuklein kislotalar bazasining tarjima qilingan ketma-ketligi bilan taqqoslanadi,

*tblastx* - o'rganilgan nukleotidlar ketma-ketligini aminokislotalar ketma-ketligiga o'zgartiradi va keyin uni ketma-ket nuklein kislotalar bazasining tarjima qilingan ketma-ketligi bilan taqqoslaydi.

**Genomli Blast** [genomni tahrirlash ] o'rganilayotgan nukleotidlar ketma-ketligini organizmning (odam, sichqon va boshqalar) genom ma'lumotlari bazasi bilan taqqoslash uchun mo'ljallangan. Ular o'z navbatida:

- *magicblast* - xaritalarni to'liq genom yoki transkriptlarni o'qiydi.

**Maxsus Blast** [tahrirlash ] BLAST-dan foydalanuvchi dasturlar:

*bl2seq* - Lokal ketma-ketlik bo'yicha ikkita ketma-ketlikni moslashtirish.

*VecScreen* - nuklein kislotaning vektorli bo'lishi mumkin bo'lgan nukleotidlar ketma-ketligini aniqlash va hk.

**3.3. Blast dasturining ishlash printsi.** Barcha moslashtirish odatda global (ketma-ketliklar to'liq taqqoslanadi) va lokalga (faqat ma'lum bo'limlar taqqoslanadi) bo'linadi. BLAST seriyali dasturlari turli xil oqsillarda o'xshash domenlar va buralishlarning mavjudligi bilan bog'liq bo'lgan lokal moslashuvlarni ishlab chiqaradi. Bundan tashqari, lokal moslashtirish mRNKni genom DNK si bilan taqqoslashga imkon beradi. Global taqqoslash holatida ketma-ketlikning kamroq o'xshashligi, ayniqsa ularning **domenlari** va **buralishlari** aniqlanadi.

O'rganilgan nukleotid yoki aminokislotalar ketma-ketligi (so'rov) BLAST veb-sahifalaridan biriga yuborilgandan so'ng, u boshqa kirish ma'lumotlari (ma'lumotlar bazasi, "so'z" (so'rov) hajmi, E qiymati va boshqalar) bilan birgalikda serverga yuboriladi. BLAST barcha "so'zlar" jadvalini tuzadi (oqsilda bu uchta aminokislotadan iborat nuklein kislotalar va 11 ta nuklein kislotadan iborat ketma-ketliklar bo'limidir.

Keyin ularni ma'lumotlar bazasida qidirishadi. Gugurt topilganda, "so'z" ning hajmini (4 yoki undan ko'p aminokislotalar va 12 yoki undan ko'p nukleotidlarga qadar), avval bo'shliqlarsiz (bo'shliqlarsiz), so'ngra ulardan foydalanishga urinish amalga oshiriladi. O'rganilayotgan ketma-ketlikning barcha "so'zlari" o'lchamlarini maksimal kengaytirgandan so'ng, moslashtirish har bir so'rov - juftliklar uchun ma'lumotlar bazasining ketma-ketligi bo'yicha maksimal natijalar bilan belgilanadi va olingan ma'lumotlar SeqAlign tuzilmasida qayd etiladi. BLAST serverida joylashgan formatlash vositasi SeqAlign ma'lumotlaridan foydalanadi va uni turli usullar bilan (an'anaviy, grafik, jadval ko'rinishida) taqdim etadi.

BLAST dasturlari tomonidan ma'lumotlar bazasida aniqlangan har bir ketma-ketlik uchun, o'rganilayotgan ketma-ketlikka (so'rovga) qanchalik o'xshashligini va bu o'xshashlik ahamiyatli ekanligini aniqlash kerak. Buning uchun BLAST har bir ketma-ketlik uchun bit sonini va E qiymatini (kutilgan qiymat, E-qiymat) hisoblab chiqadi.

O'xshashlikni aniqlashda asosiy element o'rin almashtirish matritsasi hisoblanadi, chunki u har qanday mumkin bo'lgan nukleotidlar yoki aminokislotalar juftligi uchun o'xshashlik indekslarini aniqlaydi. BLAST seriyasining ko'plab dasturlarida

- LOSUM62 matritsasiidan foydalanadi (Bloklnrni almashtirish matritsasi 62% identifikatsiya, 62% identifikatsiyali blokni almashtirish matritsasi). Istisno holatlar blastn va megablast (nukleotidni ishlatadigan dasturlar -nukleotidlarni taqqoslash va aminokislotalarni almashtirish matritsalarini ishlatmaslik).

- O'zgartirilgan Smith-Waterman yoki Sellers algoritmlaridan foydalanib, barcha segmentlar (kengaytirilgan "so'zlar") aniqlanmagan, chunki ular o'xshashlik ko'rsatkichlarining pasayishiga olib keladi.

Bunday kengaytirilgan "so'zlar" juftlari yuqori ko'rsatkich segmentlari (HSP) deb nomlanadi. O'rganilgan ketma-ketliklar (m) va ma'lumotlar bazasining ketma-ketligi (n) etarlicha katta bo'lsa, HSP o'xshashlik ko'rsatkichlari ikkita parametr K (qidiruv maydonining o'lchami) va P bilan tavsiflanadi.

Ushbu ko'rsatkichlar o'rganilayotgan ketma-ketlik va ma'lumotlar bazasi ketma-ketligining o'xshashlik ko'rsatkichlarini (S) keltirishda ko'rsatilishi kerak.

Amaldagi matritsadan qat'i nazar, turli xil taqqoslanishlarning o'xshashlik ko'rsatkichlarini taqqoslash uchun ular o'zgartirilishi kerak. O'zgartirilgan o'xshashlik indeksini olish uchun (bitlar soni, B) foydalaning

$$\text{formula: } B = (P \cdot S - \ln \{K\}) / \ln \{2\}$$

B qiymati ketma-ketliklar bir-biriga qanchalik o'xshashligini ko'rsatadi (bit soni qancha ko'p bo'lsa, o'xshashlik shuncha katta bo'ladi). B hisoblash formulasi K va P ko'rsatkichlarini o'z ichiga olganligi sababli, B qiymatlarini o'zgartirganda ularni ko'rsatishga hojat yo'q,

BLAST dasturlari asosan E ni emas, balki P ning qiymatini aniqlaydilar (indekslari S dan katta yoki unga teng bo'lgan kamida bitta HESga ega bo'lish ehtimoli). Ammo  $E < 0.01$  uchun P va E qiymatlari deyarli bir xil.

E ning qiymati faqat ikkita aminokislota yoki nukleotidning ketma-ketligini taqqoslashda formula bilan aniqlanadi. Uzunligi m o'rganilgan ketma-ketlikni ko'plab ma'lumotlar bazalari ketma-ketligi bilan taqqoslash ikki nuqtaga asoslanishi mumkin. Birinchi nuqta, ma'lumotlar bazasining barcha ketma-ketliklari o'rganilayotganga o'xshashdir. Bu ma'lumotlar bazasida mavjud bo'lgan qisqa ketma-ketlik bilan hizalanish uchun E qiymatini uzun ketma-ketlik bilan tekislash uchun E qiymatiga tenglashtirish kerakligini anglatadi. Ma'lumotlar bazasidan E qiymatini hisoblash uchun olingan E qiymatini undagi ketma-ketliklar soniga juft-juft taqqoslash orqali ko'paytirish kerak. Ikkinchi nuqta, o'rganilgan ketma-ketlik uzoq ketma-ketliklarga qaraganda qisqaroqroqqa o'xshaydi, chunki ikkinchisi ko'pincha turli qismlardan iborat (ko'p oqsillar domenlardan iborat). Agar o'xshashlik ehtimolligi ketma-ketlik uzunligiga mutanosib deb hisoblasak, n uzunlikdagi ma'lumotlar bazasi uchun E ning juft qiymatini  $N / n$  ga ko'paytirish kerak, bu erda N - bazadagi aminokislotalar yoki nukleotidlarning umumiy uzunligi. BLAST dasturlari asosan ushbu yondashuvdan ma'lumotlar bazasidagi E qiymatlarini hisoblashda foydalanadilar.

Nazariy jihatdan, mahalliy hizalanish har qanday nukleotid yoki aminokislotalarning hizalanadigan ketma-ketligidan boshlanishi mumkin. Biroq, GES, qoida tariqasida, chekkaga yaqin boshlamaydi (boshida yoki oxirida)

ketma-ketliklar. Bunday chekka effektini tuzatish uchun ketma-ketlikning samarali uzunligini hisoblash kerak. 200 dan ortiq qoldiqlar ketma-ketligi holatida zararsizlantirish amalga oshiriladi

BLAST (Basic Basic Alignment Search Tool) - asosiy tuzilish (ketma-ketlik) yoki uning bo'lagi ma'lum bo'lgan oqsillar yoki nuklein kislotalarning homologlarini qidirishda ishlatiladigan kompyuter dasturlari oilasi. BLAST-dan foydalanib, tadqiqotchi o'z ketma-ketligini ma'lumotlar bazasidagi ketma-ketliklar bilan taqqoslashi va taxmin qilinayotgan homologlarning ketma-ketligini topishi mumkin. Bu molekulyar biologlar, bioinformatika, sistematikalar uchun eng muhim vositadir. BLAST dasturi olimlar Stiven Altschul, Uorren Gish, Uebb Miller, Evgeniy Myers va Devid J. Lipman tomonidan AQShning Milliy Sog'liqni saqlash institutida ishlab chiqilgan va 1990 yilda Molekulyar Biologiya jurnalida nashr etilgan [1].

O'rganilgan nukleotid yoki aminokislotalar ketma-ketligi (so'rov) BLAST veb-sahifalaridan biriga kiritilgandan so'ng, u boshqa kirish ma'lumotlari (ma'lumotlar

bazasi, "so'z" (fitna) hajmi, E qiymati va boshqalar) bilan birgalikda serverga yuboriladi. BLAST barcha "so'zlar" jadvalini (oqsilda, bu uchta aminokislotadan tashkil topgan va 11 ta nukleotidning nuklein kislotalaridan iborat ketma-ketliklar bo'limi) va shunga o'xshash "so'zlar" ning jadvalini tuzadi.

BLAST dasturlari tomonidan ma'lumotlar bazasida aniqlangan har bir ketma-ketlik uchun, o'rganilayotgan ketma-ketlikka (so'rovga) qanchalik o'xshashligini va bu o'xshashlik ahamiyatli ekanligini aniqlash kerak. Buning uchun BLAST har bir ketma-ketlik uchun bit sonini va E qiymatini (kutilgan qiymat, E-qiymat) hisoblab chiqadi.

O'xshashlikni aniqlashda asosiy element o'rin almashtirish matritsasi hisoblanadi, chunki u har qanday mumkin bo'lgan nukleotidlar yoki aminokislotalar juftligi uchun o'xshashlik indekslarini aniqlaydi. BLAST seriyali dasturlarning aksariyati BLOSUM62 matritsasiidan foydalanadi (Blokklarni almashtirish matritsasi 62% identifikatsiya, 62% identifikatsiyali blokni almashtirish matritsasi). Istisnolardan blastn va megablast (nukleotid-nukleotidni taqqoslashni amalga oshiradigan va aminokislotalar matritsasini ishlatmaydigan dasturlar).

### **Muhokama uchun savollar.**

1. Nukleotidli blast nima?
2. Proteinli Blast nima?
3. Translyantli Blast nima?
4. Genomli Blast nima?
5. Maxsus Blast nima?
6. Blast dasturining ishlash printspi nima?
7. DNKni modifikasilovchi fermentlar nima?
8. Genom tuzilishini tusuntiring.
9. Azotli asoslarning kimyoviy modifikatsiyalari nima?
10. Chargaff qoidasi nima?
11. Genetik rekombinatsiya nima?
12. RNK genamlari nima?
13. Oqsil necha xil darajadagi tuzilishga ega va ularning strukturaviy farqi nimada?
14. Birinchi marta qaysi organism genomi to'liq ochib berilgan?

#### **4- mavzu: Eukariot organizmlar gen strukturalarini bashorat qilish.**

##### **Asosiy savollar**

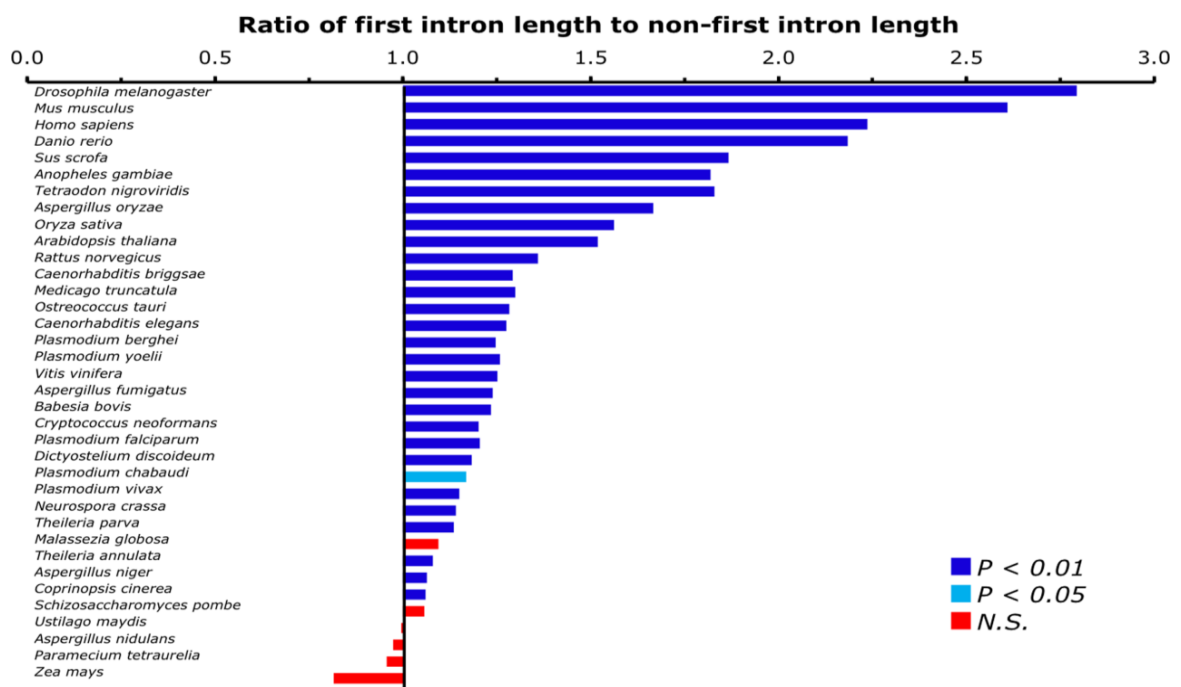
1. Gen strukturaviy tarkibi.
2. Culstal dasturi va uning turlari
3. Eukariot organizmlar gen strukturalarini bashorat qilish

**Mavzuga oid tayanch tushuncha va iboralar:** *Gen annotatsiyasi, ,Intron, Ezkzon, Culstal, Promotr, Strukturaviy genlar, gen regulyator.*

**Dars maqsadi:** Talabalar gen strukturasi hamda uning turlari haqida tushuncha berish va ularni taqqoslash uchun ishlatiladigan Culstal hamda boshqa dasturlar bilan tanishish.

**Gen strukturaviy tarkibi.** Eukaryotik gen tuzilishining ko'pgina xususiyatlari yaxshi tavsiflangan bo'lsa-da, transkriptning turli pozitsiyalarida paydo bo'ladigan intronlarning shakli va funktsiyasidagi farqlar unchalik yaxshi tushunilmagan. Xususan, intron pozitsiyasi intron uzunligi o'zgarishi dinamikasiga nisbatan kam e'tibor berilgan. Ushbu tadqiqot GenBankdagi intron uzunligi bo'yicha barcha mavjud ma'lumotlarni tahlil qiladi va turlarning keng doirasi bo'ylab birinchi intronlarda uzunlikning sezilarli darajada oshishi tendentsiyasini topadi. Ushbu tendentsiya uchta model organizmlar (*Arabidopsis thaliana*, *Caenorhabditis elegans* va *Drosophila melanogaster*) uchun yuqori ishonchli gen annotatsiyasi ma'lumotlaridan foydalanganda yanada kuchliroq ekanligi aniqlandi) tadqiq qilish shuni ko'rsatadiki, 5' UTRdagi birinchi intron - o'rtacha - gen ichidagi barcha quyi oqim intronlaridan sezilarli darajada uzunroq. *A. thaliana*dagi birinchi intron uzunligining oshishiga qisman tushuntirish birinchi intronlarda mavjud bo'lgan ba'zi motivlarning ko'payishi

bilan izohlanadi. Demak, barcha eukaryotlarda birinchi intron uzunligining ortishi umumiylik haligacha noaniq. Genning turli pozitsiyalarida paydo bo'ladigan intronlarni farqlashning sabablaridan biri shundaki, genning 5' proksimal mintaqasidagi intronlar ("erta" intronlar) ko'pincha gen ekspressiyasi bilan bog'liq bo'lgan muhim funktsional xususiyatlarga ega ekanligi ko'rsatilgan. Intronning gen ekspressiyasini kuchaytirish qobiliyati intron vositachiligini kuchaytirish (IME) deb ataladi, ammo barcha intronlar IME effektini keltirib chiqarmaydi va har qanday intronning ketma-ketligi uning pozitsiyasidan kamroq ahamiyatga ega bo'lishi mumkin. Yuqori ifoda uchun zarur bo'lgan yoki normal ifoda uchun zarur bo'lgan erta intronlarga misollar odamlarda, sichqonlarda, *A. thaliana*, guruchda, topilgan. va *C. elegans*. IME orqali ifodani kuchaytirishga hissa qo'shishi mumkin bo'lgan ketma-ketlik motivi *A. thaliana* va guruchda bashorat qilingan. Erta intronlar keyinchalik genda paydo bo'ladiganlardan funktsional jihatdan farq qilishi mumkinligini ko'rsatadigan boshqa sabablar ham mavjud. Odamlardagi birinchi intronlar (keyingi intronlarga nisbatan) kamroq SNP va transkripsiya faktorini bog'lash joylari ni o'z ichiga oladi va SINE elementlarini kiritish chastotasini kamaytiradi. Turli xil umurtqali hayvonlar turlarida ortologik intronlar o'rtasidagi taqqoslashlar shuni ko'rsatdiki, birinchi intron (va ayniqsa birinchi intronning 5' eng ko'p mintaqasi) barcha intronlarning eng saqlanib qolgani bo'ladi.

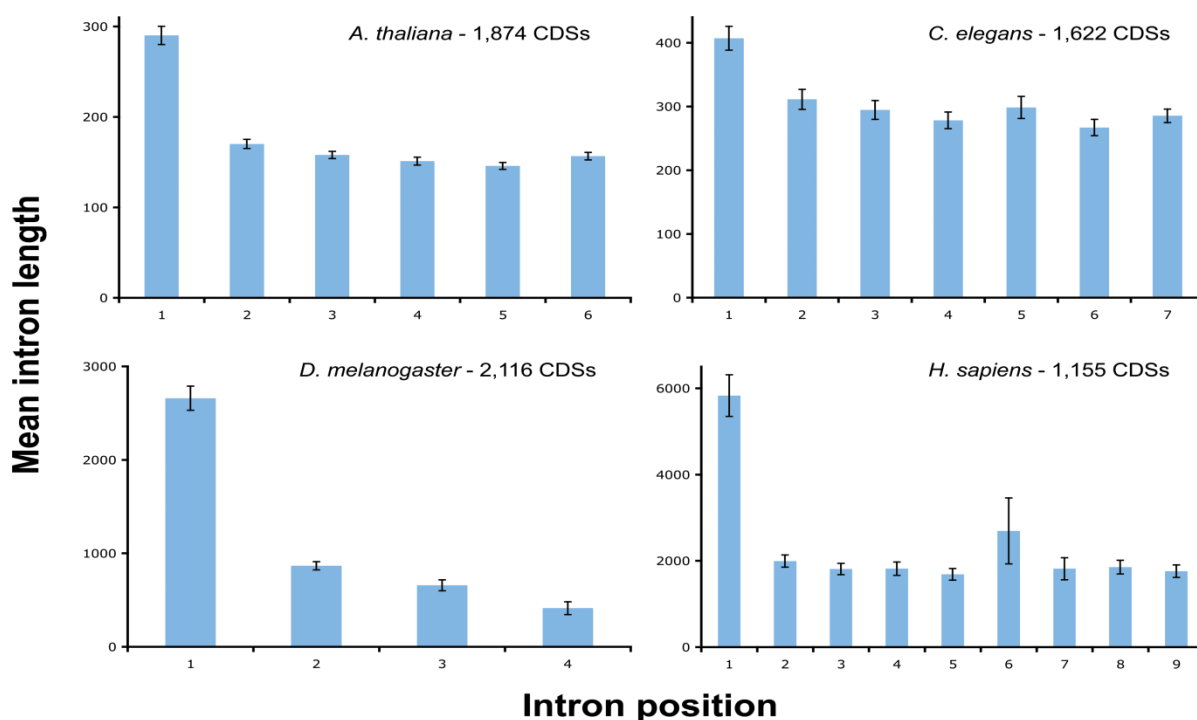


#### Birinchi intronlar ko'pchilik turlardagi eng uzun intronlardir.

GenBankdagi barcha turlar uchun ko'rsatilgan natijalar 164-sonli relizlar, ularda bir nechta intronlarni aniqlaydigan kamida 500 CDS mavjud. Muhimlikni aniqlash uchun Z-testlari ishlatilgan va rang muhimlik darajasini bildiradi (qarang, afsona, NS = ahamiyatsiz). <https://doi.org/10.1371/journal.pone.0003093.g001>

Uzunroq birinchi intronlar tandentsiyasi "birinchi bo'lmagan" intronlar bo'yicha ma'lumotlarni alohida sinflarga bo'lish yo'li bilan batafsil ko'rib chiqildi, ya'ni

intronlarning soni aynan bir xil bo'lgan CDSlarda 1 va 2-intron uzunliklari solishtirilganda. Hammasi bo'lib, ma'lum miqdordagi intronlarga ega bo'lgan turlar uchun 118 ta taqqoslash amalga oshirildi (bitta intronli CDSlar bundan mustasno). Ulardan ko'pchiligi (78%) 1 va 2-intron uzunligi o'rtasida sezilarli farqni ko'rsatdi (Z testi:  $P < 0,05$ , tanlangan misollar 2-rasmda ko'rsatilgan.). Kam sonli hollarda 2-intron ham keyingi intronlarga qaraganda ancha uzunroq edi, lekin odatda u boshqa intronlardan ajralib turishi mumkin bo'lgan birinchi intron edi (ma'lumotlar ko'rsatilmagan). 1 va 2-intron uzunligi o'rtasidagi sezilarli farqlar 2 introndan kam bo'lgan (masalan, *Aspergillus niger*da 1 va 2-intronlarning o'rtacha uzunligi : 99 nt va 91 nt,  $P < 0,01$ ) yoki 20 ta intron (masalan,  $P < 0,01$ ) bo'lgan CDSlarda sodir bo'lgan. *Oryza sativa* : 554 nt va 344 nt,  $P < 0,01$ ).



#### Turli xil intronlar soniga ega tanlangan turlar uchun intron o'lchamining o'zgarishi.

Intron uzunligi 4, 6, 7 yoki 9 intronni o'z ichiga olgan CDS'li turlar uchun ko'rsatilgan (mos ravishda *D. melanogaster*, *A. thaliana*, *C. elegans* va *H. sapiens*). Grafikdagi chiziqlar o'rtachaning standart xatosini ko'rsatadi. Har bir tur uchun ishlatiladigan CDS raqamlari ko'rsatilgan. <https://doi.org/10.1371/journal.pone.0003093.g002>

**Gen tuzilishi** - bu gen ichidagi maxsus ketma-ketlik elementlarini tashkil qilish. Genlar tirik hujayralar omon qolishi va ko'payishi uchun zarur bo'lgan ma'lumotlarni o'z ichiga oladi. Aksariyat organizmlarda genlar DNK dan iborat bo'lib, bu yerda DNKning ma'lum ketma-ketligi gen funksiyasini belgilaydi. Gen DNK dan RNK ga transkripsiya qilinadi (nusxalanadi), u to'g'ridan-to'g'ri funktsiyaga ega kodlanmagan (ncRNK) yoki oraliq xabarchi (mRNK) bo'lishi mumkin, keyinchalik u oqsilga aylanadi.. Ushbu bosqichlarning har biri gen ichidagi ma'lum ketma-ketlik elementlari



yoki hududlar tomonidan boshqariladi. Shuning uchun har bir gen funksional bo'lishi uchun bir nechta ketma-ketlik elementlarini talab qiladi. <sup>[2]</sup> Bu funksional oqsil yoki ncRNK ni kodlaydigan ketma-ketlikni, shuningdek, bir nechta tartibga soluvchi ketma-ketlik mintaqalarini o'z ichiga oladi. Bu hududlar bir necha tayanch juftligacha qisqa, minglab tayanch juftlarga uzun bo'lishi mumkin.

Regulyatsiya ketma-ketligi genlarning ekstremalarida joylashgan. Ushbu ketma-ketlik hududlari yoki transkripsiya qilingan hudud ( promotor ) yonida bo'lishi mumkin yoki ko'plab kilobazalar bilan ajralib turishi mumkin ( kuchaytirgichlar va susturucular ). <sup>[8]</sup> Promotor genning 5' uchida joylashgan bo'lib, asosiy promotor ketma-ketligi va proksimal promotor ketma-ketligidan iborat. Asosiy promoter DNKni RNKga nusxalash uchun zarur bo'lgan RNK polimeraza va boshqa oqsillarni bog'lash orqali transkripsiyaning boshlang'ich joyini belgilaydi. Proksimal promotor hududi RNK polimeraza uchun yadro promouterining yaqinligini o'zgartiruvchi transkripsiya omillarini bog'laydi. Genlar faollashtiruvchi yoki repressor oqsillarini bog'lash orqali promotorlarning faolligini yanada o'zgartiradigan bir nechta kuchaytiruvchi va susturucu ketma-ketliklari bilan tartibga solinishi mumkin. Kuchaytirgichlar va susturucular gendan bir necha ming tayanch juftlikdan uzoqda joylashgan bo'lishi mumkin. Turli xil transkripsiya omillarining bog'lanishi, shuning uchun turli vaqtlarda va turli hujayralardagi transkripsiyaning boshlanishi tezligini tartibga soladi. <sup>[13]</sup>

Garchi barcha organizmlar transkripsiya faollashtiruvchi va repressorlardan foydalansa ham, eukaryotik genlar "standart o'chirilgan", prokaryotik genlar esa "sukut bo'yicha yoqilgan" deyiladi. Eukaryotik genlarning asosiy promotori, odatda, ekspressiya paydo bo'lishi uchun promotor elementlari tomonidan qo'shimcha faollashtirishni talab qiladi. Prokaryotik genlarning asosiy promoteri, aksincha, kuchli ifodalash uchun etarli va repressorlar tomonidan tartibga solinadi.

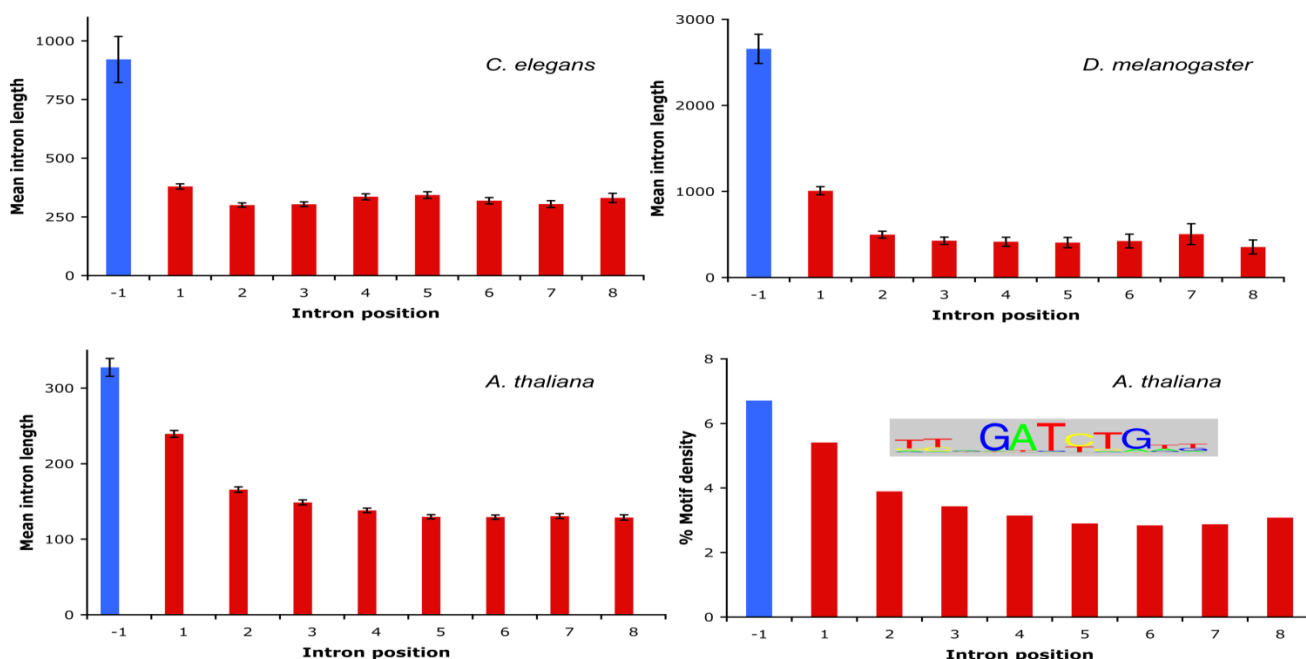
**Eukariotlarning gen strukturasi.** Eukaryotik genlarning tuzilishi prokariotlarda uchramaydigan xususiyatlarni o'z ichiga oladi. Ulardan eng taalluqli modifikatsiya-transkripsiyonel post ning oldindan mRNA'lardan ishlab chiqarish mRNA etuk oqsil tarjima qilish uchun tayyor. Eukaryotik genlar, odatda, prokaryotlarga nisbatan gen ekspressiyasini nazorat qilish uchun ko'proq tartibga soluvchi elementlarga ega. <sup>[5]</sup> Bu, ayniqsa, ko'p hujayrali eukariotlarda, masalan, odamlarda, gen ekspressiyasi turli to'qimalarda keng farq qiladigan hollarda to'g'ri keladi.

Eukaryotik genlar tuzilishining asosiy xususiyati shundaki, ularning transkriptlari odatda ekson va intron mintaqalariga bo'linadi. Exon viloyatlar final saqlanadi etuk mRNA hududlar Intro esa, molekulada amalga spliced (eksize) post-transkripsiyonel qayta ishlash jarayonida. <sup>[22]</sup> Darhaqiqat, genning intron hududlari

ekson hududlariga qaraganda ancha uzun bo'lishi mumkin. Bir-biriga bog'langandan so'ng, ekzonlar yagona uzluksiz oqsil kodlash hududlarini hosil qiladi va qo'shilish chegaralari aniqlanmaydi. Eukaryotik transkripsiyadan keyingi ishlov berish, shuningdek, mRNKning boshlanishiga 5' qopqoq va poli-adenozin dumini qo'shadi.mRNKning oxirigacha. Bu qo'shimchalar mRNA barqarorlashtirish va uning bevosita transport dan yadro uchun sitoplazma bu xususiyatlar ham bevosita gen tarkibida kodlangan bo'lsa-da.

### **Yuqori ishonchli gen izohlari yordamida intronlarni tahlil qilish.**

GenBankning CDS ma'lumotlaridan foydalanishning asosiy cheklovi shundaki, bu ma'lumotlar genning 5' UTR qismida joylashgan bo'lishi mumkin bo'lgan har qanday intronlarni o'z ichiga olmaydi. Shuning uchun GenBank ma'lumotlaridagi ko'plab "birinchi" intronlar aslida gendagi ikkinchi yoki uchinchi intron bo'lishi mumkin. Yana bir muammo shundaki, GenBank zaxiralarni o'z ichiga oladi; bir xil intron ketma-ketliklari bir nechta yozuvlarda muqobil qo'shilish variantlari va/yoki bir xil turdagi turli shtammlardan olingan ketma-ketliklar sifatida mavjud bo'lishi mumkin. Ushbu muammolarni bartaraf etish uchun biz uchta model organizm turidagi yuqori ishonchli gen annotatsiyalari to'plamidan olingan intron ma'lumotlaridan foydalandik (har bir lokus uchun faqat bitta izoform). Ushbu izohlarning barchasi cDNK dalillari bilan tasdiqlangan va ko'plab kodlash transkriptlarining 5' UTR da intronlar aniqlangan (5' UTR intronli genlarning ulushi *D. melanogaster* uchun 23%, 15% va 7% edi *A. thaliana* va *C. elegans* ). Ushbu intronlarning tahlili uzoqroq birinchi intronlarning oldindan kuzatilganligini tasdiqladi va shuningdek, kodlash mintaqasining birinchi introni quyi oqim kodlash intronlariga qaraganda ancha uzoqroq bo'lsa-da, bu 5' UTR ning birinchi introni, agar mavjud bo'lsa, bu eng uzun ekanligini ko'rsatdi. transkript ichidagi intron. Aksariyat genlarda 5' UTR intronlari mavjud emasligi sababli, biz yuqori oqim 5' UTR introniga ega yoki ega bo'lmagan genlardagi CDS ning birinchi intronlari o'rtasidagi farqlarni qidirdik. CDS ning birinchi introni UTRda oldingi introna ega bo'lmagan genlarda mavjud bo'lganlarga qaraganda uzunroq ekanligi aniqlandi. Masalan, *A. thaliana*da barcha CDSlardan birinchi intronning o'rtacha uzunligi 239 nt edi. Yuqori oqimdagi UTR introni bo'lmagan CDS larning 77% da uzunlik biroz oshib, 248 nt ga etadi va 5' UTR introni bo'lgan CDS larning 23 foizida uzunlik 183 nt ga tushadi.

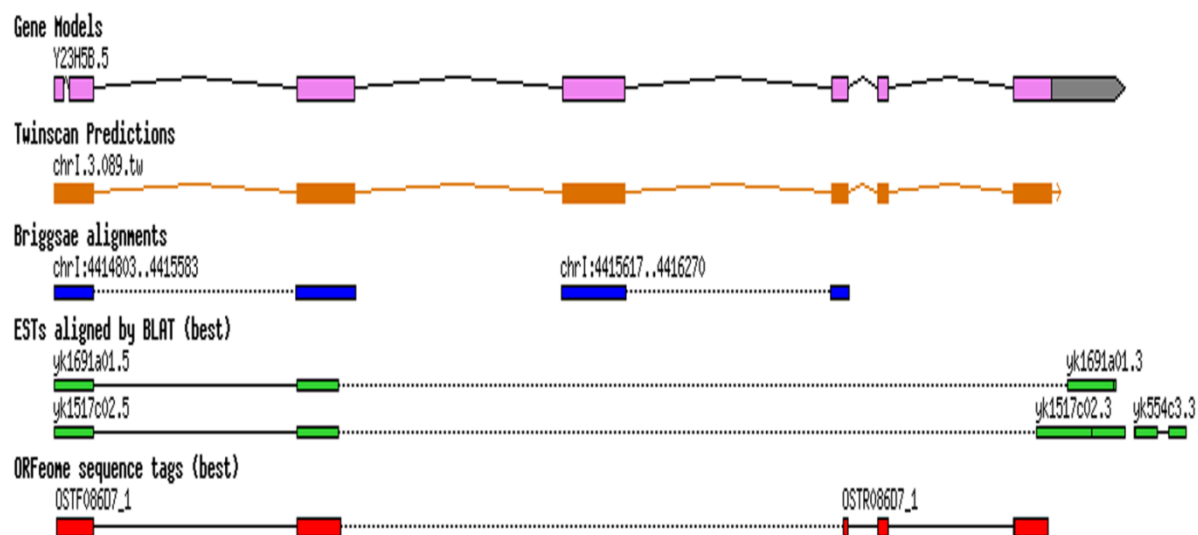


### Uchta model organizmda intron uzunligining o'zgarishi.

O'rtacha intron uzunligi 5' UTRdagi birinchi intron uchun (pozitsiya -1, ko'k rangda) va uchta nomlangan tur uchun kodlash ketma-ketligining birinchi sakkiz introni uchun (qizil rangda) hisoblanadi. Xato satrlari o'rtachaning standart xatosini ko'rsatadi. Pastki o'ng panelda *A. thaliana* intronlarida potentsial IME motivining (rasmda) paydo bo'lishi ko'rsatilgan. %Motif zichligi har bir toifadagi barcha intronlarni birlashtirib, keyin motiv umumiy ketma-ketlikning qaysi qismini egallashini hisoblash orqali hisoblanadi. <https://doi.org/10.1371/journal.pone.0003093.g003>

*C. elegans* genomining g'ayrioddiy xususiyati shundaki, ko'p genlar (~15%) operonlarda joylashgan. Shuning uchun operondagi quyi oqim genining birinchi introni transkriptdagi birinchi introni bo'lmaydi. Operonlardagi intronlarni tekshirish uchun biz 1-intronlarni operonning birinchi genidan, operonlarda quyi oqim genlaridan va operonlarda bo'lmagan genlardan bo'lganlarga ajratdik. 1-intronlarning ushbu uchta guruhining o'rtacha uzunligi mos ravishda 230 nt, 318 nt va 458 nt. Bu shuni ko'rsatadiki, operonlar ichidagi intronlar o'rtacha *C. elegans* intronidan (344 nt) qisqaroq bo'ladi, operon transkripsiyasidagi birinchi intron esa o'rtachadan sezilarli darajada qisqaroq.

**Gen annotatsiyasidagi xatolarni aniqlash uchun intron uzunligi ma'lumotlaridan foydalanish.** Biz turli pozitsiyalardagi intronlar turli uzunliklarga ega ekanligi haqidagi bilimlar ommaviy ravishda mavjud bo'lgan gen izohlarida xatolarni topishda foydali bo'lishi mumkinligini tekshirdik. Potentsial nomzod annotatsiya xatolari *C. elegans*da o'rtachadan qisqaroq birinchi intronga, so'ngra o'rtachadan ancha uzunroq ikkinchi intronga ega genlarni izlash orqali aniqlandi. Ushbu misollarning 20 tasi tekshirildi va ulardan to'rttasi (20%) annotatsiya xatosi ekanligi aniqlandi. Ba'zi hollarda izoh bilan bog'liq muammo aniq edi, chunki birinchi intronni izohdan o'chirish kerakligi haqida aniq tasdiqlovchi dalillar mavjud edi. Ushbu kuratsiya xatolar WormBase ma'lumotlar bazasiga xabar qilindi va izohlar keyinchalik tuzatildi.



**Shakl 4. Intron uzunligini tekshirish orqali aniqlangan noto‘g‘ri *C. elegans* gen annotatsiyasi.**

Ushbu gen bashorati birinchi eksonda noto‘g‘ri intron ketma-ketligini o‘z ichiga olgan. Transkript dalillari, *C. briggsae* dan homologiya dalillari va muqobil gen bashorati (Twinscan) birinchi intron annotatsiya xatosi ekanligini ko‘rsatdi. Tasvir WormBase WS180 ( <http://ws180.wormbase.org> ) versiyasining Genom brauzeri ekranidan olingan. <https://doi.org/10.1371/journal.pone.0003093.g004>

### **Intron uzunligi ma'lumotlaridan foydalanish.**

Genom annotatsiyasi turli xil eksperimental va hisoblash dalillarini birlashtirish uchun murakkab kompyuter vositalarini talab qiladi. Hisoblash gen topuvchilari haqiqiy genlarning ko‘plab ma'lum xususiyatlarini birlashtiradi (kodlash tarafdoshligi, ulanish joyi konsensus ketma-ketligi va boshqalar) va intron uzunligining turli sinflarini aks ettiruvchi modellarni o‘z ichiga olishi mumkin. Biroq, biz genning turli pozitsiyalarida nisbiy intron uzunligi bo‘yicha genlarni bashorat qiladigan biron bir dastur haqida bilmaymiz. *C. elegans* uchun dastlabki natijalar shuni ko‘rsatadiki, bu izohlarda xatolarni topish uchun asosli usuldir. Ma'lumki, uzun intronlar genlarni aniqlash dasturlari uchun muammoli bo‘lishi mumkin [39], bu birinchi intronlarning ko‘paygan uzunligini hisobga olish uchun genlarni topish algoritmlari uchun foydali bo‘lishi mumkin.

**GenBank ma'lumotlari.** GenBank ma'lumotlar bazasining 164-versiyasidan ma'lumotlarni olish uchun Perl skripti yozilgan. Faqat INV, MAM, PLN, PRI, ROD, VRT va HTG bo'linmalaridan DNK molekula turiga ega bo'lgan ketma-ketliklar va Butun Genom Shotgun (WGS) katalogidagi ketma-ketliklardan foydalanilgan. Intron uzunligi gen ichidagi oqsillarni kodlovchi hududlarning joylashishini aniqlaydigan kodlash ketma-ketligi (CDS) xususiyatlaridan kelib chiqqan holda aniqlandi. Qisman CDS (tarjimani boshlash va/yoki to'xtash joylari noma'lum) bir nechta GenBank yozuvlarini qamrab olgan CDS funksiyalari kabi chiqarib tashlandi. Uzunligi 20 nukleotiddan kam bo'lgan intronlarni o'z ichiga olgan har qanday CDS ham e'tiborga olinmadi. Ushbu filtrlash deyarli 3000 turdan ikki millionga yaqin CDS to'plamini ishlab chiqardi.

Har bir tur uchun CDS ichidagi turli pozitsiyalardagi intronlarning o'rtacha uzunligi ikki xil usulda hisoblab chiqilgan. Birinchidan, CDS turli sinflarga bo'lingan - ular qancha intron borligiga qarab - va har bir sinf ichidagi har bir intron pozitsiyasida o'rtacha intron uzunligi hisoblangan (har bir pozitsiyada kamida 100 intron talabi bilan). Ikkinchidan, barcha intronlar "birinchi intron" yoki "birinchi bo'lmagan" deb tasniflangan. Qo'shimcha filtrlash ushbu ma'lumotlarni GenBankda kamida 500 CDSga ega bo'lgan turlarga cheklab qo'ydi va har qanday bitta intronli CDS'larni chiqarib tashladi. O'rtacha intron uzunliklari Z-testi yordamida taqqoslandi.

### **Yuqori ishonchli genom annotatsiya ma'lumotlari.**

*A. thaliana*, *C. elegans* va *D. melanogaster*ning genom izohlari tafsilotlari TAIR, WormBase va FlyBase tegishli model organizm ma'lumotlar bazalaridan olingan. Ushbu ma'lumotlar bazalaridan yuklab olingan GFF va DNK fayllari transkriptning barcha pozitsiyalarida joylashgan intronlardan uzunlik va ketma-ketlik ma'lumotlarini olish uchun Perl skriptlari yordamida tahlil qilindi. Uchta ma'lumotlar bazasi tarkibi o'rtasidagi farqlar transkript ma'lumotlari (cDNA'lar, ESTlar) tomonidan eng yaxshi qo'llab-quvvatlanadigan genlardan faqat intron tafsilotlarini olish uchun har bir turga boshqacha munosabatda bo'lishini anglatadi. Bir nechta transkript izoformlari bo'lgan genlar uchun faqat birinchi nomlangan yoki raqamlangan transkript saqlanib qoldi.

### **Ikkinchi savolning bayoni.**

**Culstal dasturi va uning turlari.** Clustal – bu ketma-ketlikni tenglashtirish uchun bioinformatikada ishlatiladigan keng tarqalgan kompyuter dasturlari, algoritmi ishlab chiqish bo'yicha Clustalning ko'plab versiyalari mavjud. Har bir vositaning tahlili va uning algoritmi ham tegishli toifalarida batafsil bayon qilingan. Clustal vositalarining kombinatsiyasi mavjud bolib har bir joriy versiyasida qo'llab-quvvatlanmasligi mumkin.

**Clustal Omega** barcha Clustal vositalaridan eng keng doiradagi operatsion tizimlarga ega. Clustal dasturining turli xil variantlari mavjud bo'lib, ularning hammasi quyida keltirilgan:

**Clustal2** -1988 yilda Des Xiggins tomonidan yaratilgan bir nechta ketma-ketliklarni moslashtirish uchun original dastur filogenetik daraxtlarni aminokislotalar yoki nukleotidlarning juft ketma-ketliklaridan olishga asoslangan. **ClustalV**: Clustal dasturining ikkinchi avlodi 1992 yilda chiqarilgan va original Clustal to'plamini qayta yozgan. U oxirgi taqqoslash bo'yicha filogenetik daraxtlarni qayta qurish, mavjud moslashtirishdan burmalar yaratish qobiliyati va qo'shnilar qo'shilishi deb nomlangan usul yordamida hizalamalardan daraxt yaratish variantini kiritdi.

**ClustalW**: 1994 yilda chiqarilgan uchinchi avlod avvalgi versiyalarga qaraganda ancha yaxshilandi. Bu progressiv algoritm bo'yicha turli yo'llar bilan yaxshilandi, shu jumladan qisman hizalanishda o'xshashlik yoki tafovutlarga ko'ra individual ketma-ketliklarni pastga yoki yuqoriga ko'tarish imkonini berdi. Shuningdek, dasturni buyruq satridan ommaviy rejimda ishlatish imkoniyati ham mavjud edi. **ClustalX**: 1997 yilda chiqarilgan ushbu versiya birinchi bo'lib foydalanuvchi grafik interfeysiga ega edi. **Clustal $\Omega$**  (Omega): hozirgi standart versiya hisoblanadi.

**Clustal2: ClustalW va ClustalX**-ning yangilangan versiyalari yuqori aniqlik va samaradorlik bilan ishlaydi. **Clustal dasturiy** ta'minotni tavsiflovchi hujjatlar juda yuqori darajada keltirilgan, ulardan ikkitasi barcha davrlarning eng iqtibos qilingan gazetalari qatoriga kiritilgan. Windows, Mac OS va Unix / Linux uchun mavjud bo'lgan dasturiy ta'minotning eng so'nggi versiyasi va bundan tashqari, u veb-interfeys orqali o'z uy sahifasida yoki Evropa Bioinformatika Institutida joylashgan.

Clustal Harajatlar bo'yicha eng ko'p ishlatiladigan dasturlar **ClustalW** va **ClustalX**. Clustal algoritmda ketma-ketliklar juftlik bilan ketma-ket joylashtirilgan bo'lib, ular ketma-ketlikning qo'pol boshlang'ich daraxtini yaratish uchun ishlatilishi mumkin; Ushbu qo'pol daraxt, qo'llanma daraxtidagi dallanma tartibiga ko'ra, ketma-ketlikni ko'paytirish uchun ishlatiladi.

### **Uchinchi savolning bayoni.**

**Eukariot organizmlar gen strukturalarini bashorat qilish.** Molekular filogenetika Umurtqali hayvonlar (ko'proq sut emizuvchilar) genlarini xaritalash orqali genlar evolyutsion kelib chiqishiga ko'ra bir hil ekanligi va xar hil turlarda bir hil gomologik genlarning bajaradigan vazifasi ham bir hil ekanligi, o'zaro aloqaga kirishishi tomonlama bir hil genlar bilan bog'lanishi ma'lum bo'ldi. Yuqoridagilarga asoslangan xolda biror genning xromasomada joylashgan o'rni, uning boshqa turlarda

qaysi genlar bilan bog‘langanligiga qarab taxmin qilinadi, ya‘ni “solishtirma xaritalash” amalga oshiriladi. Albatta bu kabi axborotlar fizik (DNK ketma – ketligining joylashgan o‘rni) va genetik xaritalash bilan ham tasdiqlanishi kerak.

Genom ma‘lumot bazalari aniq bir turdagi lokalizatsiyalashgan genetik markerlar xaqida axborot va boshqa turlardagi gomologik genlarni izlash uchun yaratiladi. Bundan tashqari genom ma‘lumot bazalari genomning umumiy ko‘rinishi, genetik haritalar uchun zarur omil hisoblangan fenotip mutatsiyalari haqida ma‘lumotga ega bo‘lishga yordam beradi.

Hozirda internet tarmog‘ida juda ko‘plab genom ma‘lumot bazalarini uchratish mumkin, lekin ularning hammasini ham ishonchli deb bo‘lmaydi. Bunda ma‘lumotlar bazasida joylashtirilgan axborotlarning doimiy yangilab turilishi muxim omil sanaladi. Odatda bunday ma‘lumotlar bosh saxifaga (home page) joylashtirilgan bo‘ladi. Ba‘zi genom ma‘lumot bazalari yagona tur uchun tuzilgan, boshqalari esa o‘z strukturasiiga ko‘ra qolganlaridan keskin farqlanadi. Xozirgi kunda genom ma‘lumot bazalarini yartuvchilari uchun barchaga foydalanishi qulay bo‘lgan yagona interfeysni yaratishga harakatqilishmoqda.

**Biologik ma‘lumotlarni qayta ishlash vositalari.** Biologik ma‘lumotlarni qayta ishlovchi vositalar jamlanmasi yetarlicha. Bu yerga nukleotid va aminokislota ketma –ketliklarini solishtiruvchi, juft va ko‘plab tenglashtiruvchi dasturlar, oqsillarning ikkilamchi va uchlamchi strukturalarini taxmin qiluvchi va yana boshqa ko‘plab dasturlarni kiritish mumkin.

Komersion ma‘lumot bazalaridan eng sifatliari foydalanuvchilarga ma‘lumotlarni to‘liq qayta ishlash imkonini beruvchi vositalar tarmog‘ini (GCG, DNASIS va boshqalar.) taklif etib, yirik biologik markazlar va laboratoriyalarda qaysidir darajada qo‘llanib kelinadi, bunday dasturiy vositalarni narxi tahminan bir necha ming dollarni tashkil etadi.

Bioinformatikada ko‘plab masalalarni hal etishda bepul dasturiy vositalar ham mavjud bo‘lib, ularni internet tarmog‘i olib yuklab olib, lokal kompyuterga joylashtirish mumkin. Biologik ma‘lumotlarni intensiv va uzluksiz ravishda qayta ishlaganda esa tarmoqqa bog‘liq bo‘lmagan dasturlarni o‘rnatish maqsadga muvofiq xisoblanadi. Biologlar uchun olish mumkin bo‘lgan dasturiy vositalar haqidagi axborotni BioCatalog (EBI) orqali topish mumkin.

Genomi to‘liq tahlil etilgan turlar. Bu mavzuda tirik organizmlar genom tadqiqotlarini birlashtiruvchi internet resurslari xaqida so‘z yuritiladi. Genomi tahlil etilgan birinchi eukariot organizmlar hamirturush achitqisi *Sacharomyces sereviciae* (1996) va yumaloq chuvalchang *Cayenorhabditis elegans* (1998) bo‘lgan. 2000 yilda *Drosophila melanogaster* meva pashshasi to‘liq tahlil etilgan va FlyBase (<http://flybase.bio.indiana.edu/>) ma‘lumotlar bazasi shu pashsha genomining struktur –funktional axborotlari va fenotipiga bag‘ishlangan. Bu drozofila pashshasining

birinchi murakkab tuzilgan ko'p hujayrali organism va genetikaning klassik ob'ekti bo'lganligi bilan aloqador. Shu bilan bir qatorda aynan drozofilaning alohida genlari, gen guruhlari, xromasoma tuzilishi va vazifalarni amalga oshirishi bo'yicha katta miqdorda ma'lumotlar to'plangan. Yuqoridagilarga asoslangan holda aynan FlyBase genom ma'lumotlar tizim shakllanishida namuna etib olingan. Biror genom haqida so'z yuritilganda uning to'liq nomini atash joiz ekanligini ham unutmaslik kerak. Misol uchun drozofilaning "melanogaster" turidan tashqari *Drosophila pseudoobscura* turning ham genom ketma –ketligi to'liq sekvenirlangan.

Bundan tashqari o'sha varaqalarga mos keluvchi aminokislota ketma ketliklari ma'lumotlar bazasiga, pashsha rivojlanish bosqichlari, tashqi va ichki tuzilishiga doir rasmlarga, genlarning xromasomada joylashgan o'rni to'g'risida tushuncha xosil qilish imkonini beruvchi grafik ilovalarga va yana bir qator boshqa kataloglarga ham yo'naltirgichlar bor.

*Drosophila melanogaster* genomining sekvenirlanishi va annotatsiyalanishiga bag'ishlangan loyixa saytini bilan to'liqroq tanishib chiqamiz. Yaqin yillargacha ushbu loyixa genom ketma ketligini taxlil etuvchi va ushbu taxlil natijalarini vizuallashtiruvchi dasturlarni ishlab chiqish bo'yicha etakchi hisoblanar edi. 1999 yilda ushbu loyixaning tashabbusi bilan ketma –ketligi o'rganilmagan genomlarni taxlil etish, ya'ni ularning annotatsiyasi, aniqlangan oqsil ketma –ketliklari bo'yicha o'hshash tomonlarini namoyon etuvchi usullarini ishlab chiqarish bilan shug'ullanuvchi 12 dan ortiq tadqiqot guruxlari o'rtasida GASP (Genome Annotation Assessment in *Drosophila melanogaster*) bellashuvi o'tkaziladi. Ushbu solishtirma tahlil natijalariga asoslangan holda genlar qidiruvi va istalgan organism genomini struktur –funktional annotatsiyasi uchun mezonlar bo'yicha nisbatan sezgirliigi yuqori va spetsifik bo'lgan dasturlar ishlab chiqarilgan.

Gap shundaki, nematodalarning rivojlanishi aniq determinatsiyalangan, ya'ni birinchi kundan boshlab xar bir hujayraning taqdiri aniq belgilab qo'yilgan va bo'linish miqdori hamda xar bir organdagi hujayralarning miqdori chegaralangan -1000 tadan sal ortiq. Nematodaning bunday tuzilishi uni molekular mexanizmlari, ayniqsa hujayralararo ta'sirlanishlar va signal yo'llarini oson aniqlash imkonini beradi.

WormBase ma'lumotlar bazasida maxsus savollarni o'rganish imkonini beruvchi ma'lumotlarning probazalari yoki yo'ldosh bazalariga yo'naltirgichlarni ham ko'rishimiz mumkin bo'ladi. Barcha bazalar bo'yicha sodda va murakkab qidiruvni amalga oshiruvchi vositlar bor.

*Cayenorhabditis elegans* genomi struktur komponentlari xosil qiluvchilarni ko'rib chiqamiz. Qandaydir bir geni va qo'shni genlarning tuzilishi yoki 3'- gen soxalarini ajratib olish mumkin. Barcha rasmlar interaktiv hisoblanganligi uchun



ta'minlangan yo'naltirgichlar bo'yicha tanlangan ob'ektni tavsiflovchi saxifaga o'tishi mumkin bo'ladi.

AnoBase-ma'lumotlar bazasi Anopheles gamibiaega doir genom va biologik axborotlarni saqlaydi. Bu ma'lumotlar bazasida ham genetic harita, birlamchi ketma – ketliklar, gen ko'rsatkichi kabilarga yo'naltirgichlar joylashgan.

O'simlik genomiga bag'ishlangan ma'lumot bazalari bilan ham tanishib chiqaylik. Arabidopsis thaliana krestguldoshlar oilasiga kiruvchi bu o'simlik hech qanday qishloq ho'jalik axamiyatiga ega emas. Lekin o'simlik organizmlari rivojlanishining genetik mexanizmlari va to'qima va organlarning bajaradigan vazifasi, ayniqsa gullash tog'risidagi dastlabki tasavvurlar aynan shu osimlik yordamida paydo bo'lgan.

Ushbu o'simlik genomining sekvenirlanishi 2000 yilda Arabidopsis thaliana genomining sekvenirlanishi va annotatsiyasi bo'yicha xalqaro loyixa - Arabidopsis Genome Initiative qoshida amalga oshirilgan. Ma'lumotlar bazasi yoki TAIR (The Arabidopsis Information Resource) gen yo'naltirgichi yordamida ushbu o'simlik haqida to'liq ma'lumotga ega bo'lish mumkin bo'ladi..

Qishloq ho'jaligi axamiyatiga ega bo'lgan ko'plab o'simlik organizmlari yirik agrotexnik firmalarning kuch va vositalari yordamida genom ketma –ketligi tahliliga jalb etilmoqda. Bu kabi o'simliklar qatoriga guruch – Oryza sativa, beda – Medicago truncatula, bug'doy – Triticum aestivum, arpa – Hordeum vulgare, soya – Glycine max, makkajohori- Zea mayslarni kiritish mumkin.

Bu organizm ko'p vaqtdan beri molekular va umumiy genetika, tibbiyot genetikasi va ayniqsa inson bilan bog'liq bo'lgan xodisalarni tadqiq etishda model organizm sifatida qo'llanib keladi. Bundan tashqari sichqon oziq –ovqat, parfyumeriya, farmakologik vositalarini sinab ko'richda ham standart ob'ekt sanaladi.

Uy sichqonining genetika, genomika va biologiyasi haqida axborotga Mouse Genome Informatics (MGI) orqali ega bo'lish mumkin.

Endi yuksak tuzilgan organizmlardan sut emizuvchilar sinfiga kiruvchi uy sichqoni – Mus Musculus genom resurslari bilan tanishib chiqaylik. Uy sichqonining genom tahlilini qoralama variant 2002 yilda tayyor bo'lgan edi.

Bu organizm ko'p vaqtdan beri molekular va umumiy genetika, tibbiyot genetikasi va ayniqsa inson bilan bog'liq bo'lgan xodisalarni tadqiq etishda model organizm sifatida qo'llanib keladi. Bundan tashqari sichqon oziq –ovqat, parfyumeriya, farmakologik vositalarini sinab ko'richda ham standart ob'ekt sanaladi.

Uy sichqonining genetika, genomika va biologiyasi haqida axborotga Mouse Genome Informatics (MGI) orqali ega bo'lish mumkin.

HAP genlari oilasi tomonidan kodlangan NF-Y transkripsiya omillari uchta subbirlikdan (HAP2/NF-YA, HAP3/NF-YB va HAP5/NF-YC) iborat bo'lib, maqsadli

genlarni yuqori o'ziga xoslikka ega bo'lgan transkripsiyaviy CCAAT o'z ichiga olgan promotorlar ketma-ketligini tartibga solishga qodir.

### **Muhokama uchun savollar.**

1. Genomi tahlil etilgan birinchi eukariot organizmlar qaysilar?
2. FlyBase nima?
3. Qaysi sinf vakillarining genlarini xaritalash orqali genlarning evolyutsion kelib chiqishiga ko'ra bir hil ekanligi va xar hil turlarda bir hil gomologik genlarning bajaradigan vazifasi ham bir hil ekanligi, o'zaro aloqaga kirishishi tomonlama bir hil genlar bilan bog'lanishi ma'lum bo'ldi.
4. Intron uzunligi ma'lumotlaridan foydalanish deganda nimani tushunasiz?
5. Nima uchun birinchi intronlar uzunroq?
6. Intron ma'lumotlari eukaryotik gen tuzilishining asosiy strukturasi ochib beradimi?

### **5- mavzu: Molekulyar filogenetika.**

#### **Asosiy savollar**

1. Molekulyar filogenetika. Filogenetikaning asosiy tushunchalari.
2. Filogenetik tahlil uchun bioinformatika vositalari
3. Clustaw2-phylogeny dasturi yordamida filogenetik daraxtni tuzish.

**Mavzuga oid tayanch tushuncha va iboralar:** *Filogeniya*, Evolyutsion daraxt, *ClustalW*, *MSA*, *MAFFT* va *T-Coffee*, *UPGMA*, Algoritm, *MEGA 7*

**Dars maqsadi:** Talabalarga Molekulyar filogenetika haqida asosiy tushunchalarni berish hamda DNK, RNK va oqsillarning tuzilishini o'rganish asosida tirik organizmlar o'rtasidagi munosabatlarni komputerdasturlarida tuzishni o'rgatish.

#### **Birinchi savol bayoni.**

**Filogenetikaning asosiy tushunchalari.** DNK yoki oqsillar ketma-ketligi ko'rinishidagi molekulyar ma'lumotlar mavjud organizmlarning juda foydali

evolyutsion istiqbollari ham ta'minlaydi, chunki organizmlar rivojlanib borishi bilan genetik materiallar vaqt o'tgani sari mutatsiyalarni fenotipik o'zgarishlarga olib keladi. Genlar to'plangan mutatsiyalarni qayd qilish uchun vosita bo'lganligi sababli, ular molekulyar qazilmalar bo'lib xizmat qilishi mumkin. Bir qator o'zaro bog'liq organizmlarning molekulyar qoldiqlarini qiyosiy tahlil qilish orqali genlarning va hatto organizmlarning evolyutsion tarixini aniqlanish imkonini beradi. Filogenetik tadqiqotlarni amalga oshirish juda mashaqqatli ishdir, chunki ma'lumotlar va tadqiqotlar shunchalik darajada ko'p, ammo hisoblash va qator bioinformatika vositalarini ishlab chiqish va ulardan foydalanish bilan amaliy hisoblash vaqtlarida katta ma'lumotlar to'plamini tahlil qilish va yuqori ehtimollik bilan eng maqbul yechimlarni topish imkoniyati mavjud. Ushbu tendentsiyaga javoban, filoinformatika (ya'ni, hisoblash filogenetikasi) bo'yicha olib borilayotgan izlanishlarning aksariyati samarali zamonaviy yondashuvlarni ishlab chiqishga qaratilgan.

Filogeniya - biologiyaning bir qismi bo'lib, organizmlarni bir-biridan kelib chiqish muammolarini o'rganadi. Filogeniya odatda sistematik nomlar yoki «evolyutsion shajara» ko'rinishida tasvirlanadi.

Molekulyar filogenetika-polimer makromolekulalar- DNK, RNK va oqsillarning tuzilishini o'rganish asosida tirik organizmlar o'rtasidagi munosabatlarni o'rnatish usulidir. Molekulyar filogenetik tahlil natijasi tirik organizmlarning filogenetik daraxtini qurishdir.

Molekulyar filogenetika tirik organizmlarning ilmiy tasnifiga kuchli ta'sir ko'rsatdi. Makromolekulalar bilan ishlash usullari turli ixtisosdagi biologlar uchun mavjud bo'lib, bu tirik organizmlar haqida yangi ma'lumotlar yig'ilishiga olib keldi. Bu ma'lumotlar asosida tirik organizmlar evolyutsiyasi haqidagi eski taxminlar qayta ko'rib chiqilmoqda. Ular yangi guruhlarni, shu jumladan faqat molekulyar filogenetik ma'lumotlar asosida aniqlangan guruhlarni ta'riflaydilar. Tirik organizmlar tuzilishiga qarab o'xshash va farqlarga ega bo'lgan guruhlarga bo'linadi. Agar ikki xil organizm bir biri bilan juda bog'liq bo'lsa, ularning ajdodlari bitta deb qabul qilinadi. Filogenetik tahlil bu evolyutsion munosabatlarni baholashga kiradi. Evolyutsion tarix esa filogenetik tahlil asosida shoxlangan, daraxtsimon diagrammalar asosida shakllantiriladi, unda taqribiy avlodlar o'rtasidagi nasliy munosabatlar keltiriladi. Bunday munosabatlar molekulalar, organizmlar o'rtasida keltirilishi mumkin. Turli xil organizmlar o'rtasidagi filogoniyalar ular o'rtasidagi gomologiyalarni tahmin qiladi va klassifikatsiyaga bog'liq bo'ladi. Filogeniya bir necha belgilar to'plamiga ko'ra, yoki klassifikatsiya asosida munosabatlar topologiyasini o'rnatadi, yoki bu munosabatlarda evolyutsion jarayonlar modelini tuzadi.

*Evolyutsion daraxt.* Organizmlar, populyatsiyalar, turlar va genlar o'rtasidagi munosabatlar tom ma'noda ularning qarindoshligiga yoki genealogiyasiga qarab, ya'ni bitta ajdoddan kelib chiqishiga qarab quriladi. Natijalar ko'pgina holatlarda daraxt

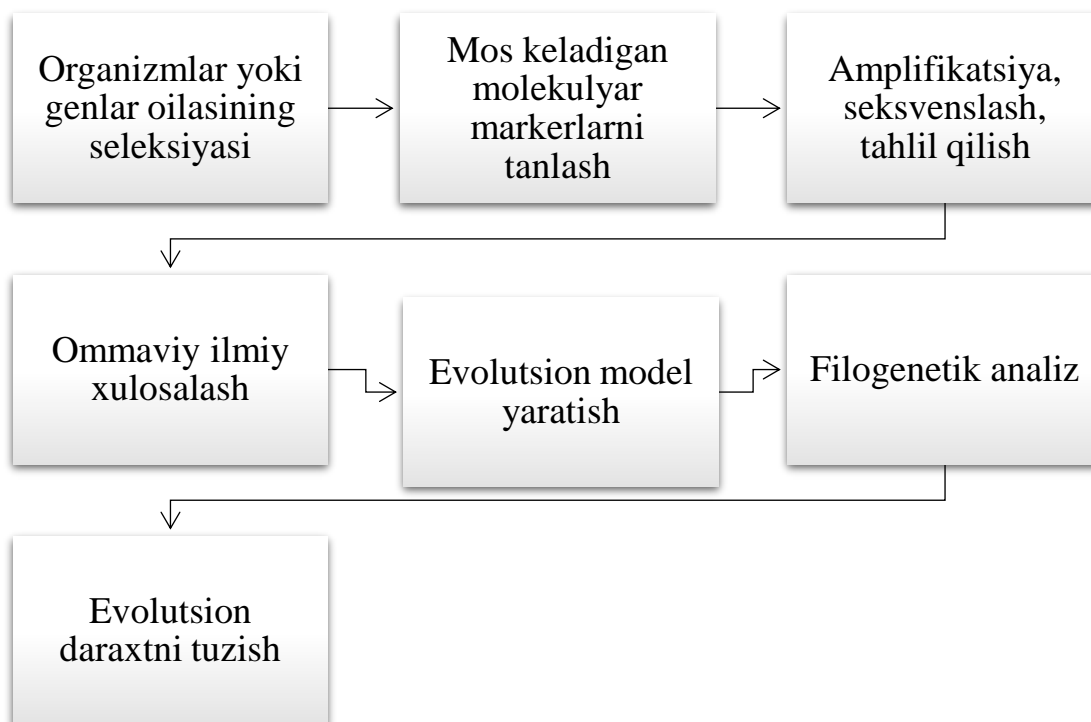
diagrammasi shaklida beriladi. Agar oxirida keltirilgan avlodlar hammasi bitta ajdoddan rivojlangan bo'lsa, bunga ildizli asos deb ataladi. Evolyutsion daraxtlar genetik ma'lumotlar asosida ham shakllantiriladi. Bir-biriga qarindosh bo'lgan nukleotil ketma-ketligi yoki oqsillar ketma-ketligining filogenetik tahlili evolyutsiya davomida oilalar rivojlanishini yo'nalishlarini aniqlashda katta ahamiyatga egadir. Ketma-ketliklarning evolyutsion munosabatlarini shakllantirishda tashqi shoxlanish diagrammasidan foydalaniladi. Bunda daraxt asosidagi shoxlangan bog'lanishlar ketma-ketliklar o'rtasidagi munosabatlar shakllanishini aks ettiradi. Filogenetik tahlilning asosiy maqsadi, daraxtdagi shoxlangan bog'lanishlarni aniqlash va shoxlarning uzunligini aniqlashga qaratiladi. Nuklein kislotalar va oqsillarning ketma-ketliklarning tahlilida ko'pincha qo'shni shoxlarda joylashgan bir biriga yaqin bo'lgan pozitsiyalar belgilanadi. Agar organizmlarda yoki organizmlar guruhida genlar guruhi aniqlansa, bu oila genlarining o'rtasidagi filogenetik munosabatlar ularning ekvivalent funksiyalari to'g'risida ma'lumot berishi mumkin. Agar ikki turli xil organizmlarda topilgan ikki xil nuklein kislotalari yoki oqsillar bir-biriga o'xshasa, ularning bir ajdoddan kelib chiqqanligi to'g'risida ma'lumot olish mumkin. Bunday ketma-ketliklarni taxrirlash natijasida ketma-ketliklarda o'zgarmay qolgan pozitsiyalarni aniqlash imkonini beradi va qaysilari o'zgarganligini bilish mumkin. Bu ikki pozitsiya ketma-ketliklari evolyutsion qarindosh bo'ladi va ularni gomologik deb hisoblash mumkin. Filogenetikada qo'yidagi sistematik yondashuvlar mavjud: guruhli-fenetik, vaqtinchalik - kladistik, va evolyutsion.

Uzoq ajdodlarida bo'lmagan belgilariga qarab, umumiy xarakterli belgilariga ko'ra guruhlanadi. Bunday orttirilgan belgilariga vizual jihatdan tasvirlanadigan turli xil belgilar kiritiladi. Kladistik tahlil fenotipik belgilar yoki aminokislotalar, nuklein kislotalar asoslarining ketma-ketligiga qarab olib boriladi.

Kladistikada uch xil mumkin bo'lgan holatlar mavjud:

1. Organizmlar o'zaro umumiy ajdoddan kelib chiqishiga qarab bog'liq bo'ladi.
2. Evolyutsion chiziqlar vaqt oralig'ida shoxlanadi.

3. Vaqt o'tishi bilan avlodlarda xarakteristikalarda o'zgarishlar sodir bo'ladi



***Filogenetik tahlil bosqichlari:***

***Ma'lumotlar bazasini yig'ish va tekislash***

- Birinchi qadam protein yoki DNKning ketma-ketligini aniqlash va boshqa tegishli ketma-ketlikdan iborat ma'lumotlar to'plamini yig'ishdir.
- DNK qiziqishlarining ketma-ketligini NCBI- BLAST yoki shunga o'xshash qidirish vositalaridan foydalanib olish mumkin.
- Agar ketma-ketliklar tanlangan va olingan bo'lsa, bir nechta ketma-ketliklar yaratiladi.
- Bu olingan ketma-ketliklarda gomologiyani aniqlash uchun ketma-ketliklar to'plamini tashkil qilishni o'z ichiga oladi.
- ClustalW, MSA, MAFFT va T-Coffee kabi ko'plab veb-saytlar va dasturlar mavjud bo'lib, ular ma'lum bir molekulyar ma'lumot to'plamida bir nechta ketma-ketlikni bajarish uchun mo'ljallangan.

**Ikkinchi savolning bayoni.**

**Filogenetik tahlil uchun bioinformatika vositalari.** Filogenetik tahlil uchun ishlatilishi mumkin bo'lgan bir nechta bioinformatika vositalari va ma'lumotlar bazalari mavjud. Bularga PANTHER, P- Pod, PFam, TreeFam, UPGMA, PhyloFacts, MEGA-7 va Clustaw2\_phylogeny dasturi tarkibiy filogenomik entsiklopediyasi kiradi. Ushbu ma'lumotlar bazalarining har biri turli xil algoritmlardan foydalanadi va ketma-ketlik ma'lumotlarini olish uchun turli xil manbalardan foydalanadi va shuning uchun

PANTHER tomonidan hisoblangan daraxtlar, masalan, P-Pod yoki PFam tomonidan yaratilgan daraxtlardan sezilarli farq qilishi mumkin. Ushbu turdagi barcha bioinformatika vositalarida bo'lgani kabi turli xil usullarni sinab ko'rish, natijalarni taqqoslash, keyin har xil ma'lumotlar to'plamini jalb qilgan tadqiqotlar uchun qaysi ma'lumotlar bazasi yaxshi ishlashini aniqlash (konsensus natijalariga ko'ra) muhimdir.

Clustal Omega dasturi yordamida nukleotidlar ketma - ketligini ko'p marotaba to'g'irlanadi. Keyingi bosqichda filogenetik daraxtni tuzishda MEGA-7 programmasida ko'proq haqiqatga o'xshash usuli (*maximal likelihood*), maksimal iqtisod (*maximal parsimony*), chamalab ko'rilgan o'rtacha juftlik (UPGMA) va yaqin qo'shnilar (Neighbor-joining)dasturlari orqali tekshiriladi va olingan nukleotidlar ketma-ketligini halqaro Genbank (NCBI) joylashtiriladi.

#### UPGMA juftliklar guruhi usuli

Pairwise intra-group unweighted (arifmetik o'rtacha, UPGMA juftliklar guruhi usuli) usuli eng oddiylaridan hisoblanadi. Uning hozirgi shaklida, usul 1973 Singh va Sokal tomonidan taqdim qilingan. Dastlab uslubni filogenetikada qo'llash morfologik xususiyatlar asosida fenogrammlar qurish bilan bog'liq. Metoddan foydalanishning zaruriy sharti o'rganilayotgan nukleotid ketma-ketliklarining doimiy evolyutsiya tezligi hisoblanadi. Agar ketma-ketlik evolyutsiyasi tezligi notekis bo'lsa (molekulyar soat modeli mos kelmasa), UPGMA usuli daraxt topologiyasida xatoliklarga olib kelishi mumkin.

*Algoritm usuli.* Birinchi bosqichda masofa matritsasida eng kichik masofa qiymatiga ega bo'lgan ikkita takson topiladi. Bu ikki takson bitta klasterga (yoki kompozitsion taksonga) birlashtiriladi. Bu usul molekulyar evolyutsiyaning yagona tezligini nazarda tutgani uchun shoxlanish (divergensiya) nuqtasi ikki takson orasidagi genetik masofaning yarmini tashkil etadi. Kelajakda bu ikki taksonning klasteri bir butun hisoblanadi. Masofa matritsasi kompozitsion takson bilan boshqa takson orasidagi masofa ga teng deb faraz qilib, qayta hisoblanadi:

$$d_{uk} = (d_{u1k} + d_{u2k})/2$$

buyerdagi  $d$  - genetik masofa bo'lsa,  $u$  kompozit ketma-ketlik.

*Qo'shnilarni qo'shish usuli.* Genetik masofalar xaritasi 2002 yilda tuzilgan bo'lib, qo'shnilarni qo'shish orqali 18 ta insonlar guruhi baxolangan, 23 xil genetik ma'lumotlar baxolangan. Xarita Yaponiya Milliy Genetika instituti professori Naruya Saytou tomonidan yaratilgan.

PANTHER-(Protein Analysis The Evolutionary Relationships) tasniflash tizimi oqsillar uchun mo'ljallangan. Proteinlar quyidagicha tasniflangan:

- Oilalar evolyutsiyaga bog'liq bo'lgan oqsillar guruhi.
- Molekulyar funksiyasiga bog'liq bo'lgan oqsillar guruhi: oqsilning o'z-o'zidan yoki bevosita biokimyoviy darajada ta'sir qiladigan oqsillarning funktsiyasi, masalan, protein kinaza.

- Biologik jarayon: hujayra yoki organizm darajasida, masalan mitozda, jarayonni bajarish uchun o‘zaro ta’sir qiladigan oqsillar guruhi.

### **Uchunchi savol bayoni**

**Clustaw2-phylogeny dasturi yordamida filogenetik daraxtni tuzish.** Tahlil uchun olingan ketma-ketliklarni tuzish. Alohida matnli faylga (Microsoft Word) filogenetik daraxt tuzish uchun xizmat qiluvchi organizmlarning (FASTA formatida), ketma-ketligini kiritiladi.

- Ketma-ketlikni raqamlanadi. Boshqa matnli faylga organizm nomlariga mos keluvchi raqamlar ketma-ketligini yozib chiqiladi.
- Clustal Omega dasturi yordamida nukleotid kislotalar va oqsillar ketma - ketligini aniqlab olinadi, so‘ngra ularni ko‘p marotabali to‘g‘irlashga mo‘ljallanadi. Clustal Omega guruhli satr yoki onlayn tarzda ishlaydi.
- Ko‘p marotabali to‘g‘irlash uchun Clustal Omega sahifasiga kiriladi: <http://www.ebi.ac.uk/Tools/msa/clustalo/index.html>.
- Clustal Omega bosh sahifasida to‘rt ilovaly menyu bor:
  - Step 1 (Qadam 1) - ilova ketma-ketlikda FASTA formatida Microsoft Word dokumentiga tahlil qilinayotgan nukleotidlar ketma-ketligini kirituvchi oynani o‘zida saqlaydi. Grafada Enter or paste da DNA tanlanadi.
  - Step 2 (Qadam 2) - ilova (Pairwise Alignment Options) juft to‘g‘irlash variantlarini o‘zida saqlaydi: sekinroq (Slow) yoki tezroq (Fast). Parametrlarni o‘zgartiriladi (Slow);
  - Step 3 (Qadam 3) - ilova ko‘plab to‘g‘irlash variantlarini o‘zida saqlaydi (Multiple Sequence Alignment Options): kiritish formatini aniqlanadi PHYLIP;
  - Step 4 (Qadam 4) - natijalarni elektcron manzil orqali yuborish uchun oyna (o‘zingizning elektron manzilngizni grafada EMAIL ko‘rsating).
  - To‘g‘irlashni ishga solish uchun SUBMIT tugmasini bosiladi. To‘g‘irlash natijalarini bir necha daqiqadan so‘ng elektron manzilga yuboriladi.
  - Step 5 (Qadam 5) - Alignments oynasinig o‘ng tarafida filogeniya tuzish uchun ilova saqlanadi. (Send to ClustaW2\_Phylogeny)
  - Send to ClustaW2\_Phylogeny tugmasini bosing, boshqa oynada filogeniya tuzish uchun ketma - ketlik ochiladi.
  - Clustal W - Phylogeny dasturi bilan filogenetik malumotlarni tayyorlash
  - Boshqa oynada Submit tugmasini bosing, bir necha daqiqa mobaynida fayllar va filogrammalar birdan filogenetik daraxt ochiladi.

### **Muhokama uchun savollar**

1. Filogenetik daraxt qanday tuziladi?



2. Ma'lumotlar bazasini yig'ish va tekislash qanday jarayonlarni o'z ichiga oladi?
3. MEGA 7 dasturi yordamida filogenetik daraxtni tuzish qanday amalga oshiriladi
4. Filogenetik daraxtning qanday terminlari bor?
5. Algoritm usuli nima?
6. Filogenetik tahlil uchun ishlatilishi mumkin bo'lgan qanday bioinformatika vositalari va ma'lumotlar bazalari mavjud.
7. Qoshnilarni qo'shish usuli nima?
8. Filogeniya nima?
9. Evolyutsion daraxt deganda nimani tushunasiz?
10. Fenetik va kladistik yondashuvlar qaysi sohada qo'llaniladi?
11. Hisoblash usullari va stoxastik modellar yordamida ketma-ketliklardan filogenetik daraxtlarni qurishqanday amalga oshiriladi?

## **6- mavzu. Biologik makromolekulalarni vizualizatsiyalashtirishning zamonaviy usullari**

### **Asosiy savollar**

1. Malekulalarning fazoviy strukturalari va ularning yaratilish tarixi.
2. Biologik makromolekulalarning strukturasi asosida ularni vizualizatsiyalashtirish.
3. PyMol va I-TASSER dasturlarida vizual tasvirlar.

**Mavzuga oid tayanch tushuncha va iboralar:** *Molekulyar simulyatsiya, 3D, NMR, Modellar 1, vdW, SES, GPU, MSS*

**Dars maqsadi:** Talabalarga biomolekulalarning vizual tarkibi hamda ularning yaratilishining bosqichma bosqich rivojlanishi haqida tushuncha berish. Vizual tasvirlarni turli xil dasturlar yordamida yaratish haqida asosiy tushunchalarni berish.

### **Birinchi savol bayoni**

Molekulalarning strukturaviy xossalari ko'p sohalarida birinchi o'rinda turadi. Ushbu hisobot molekulyar grafika va vizualizatsiya sohalarida ishlab chiqilgan dasturlar haqida to'liq ma'lumot beradi. Bunda asosiy e'tibor strukturaviy biologiyada qo'llaniladigan sohalar ko'p jihatdan uch o'lchovli, murakkab, katta va vaqt ta'sirida o'zgaruvchan molekulyar tuzilmalarning kompyuterlashtirilgan geometrik va vizual tasvirlariga tayanadi. Hisobotda molekulyar vizualizatsiyaning qaysi sohalarida allaqachon keng qamrovli o'rganilganligi va sohaning rivojlanayotganligi ko'rsatadigan taksonomiya taqdim etilgan. U molekulyar tuzilmalar uchun vizualizatsiyani, tasvir sifati va kadr tezligini samarali namoyish qila olish strategiyalarini muhokama qiladi hamda tafsilotlar darajasining turli tomonlarini

qamrab oladi va molekulyar simulyatsiya ma'lumotlarining dinamik jihatlarini aks ettiruvchi vizualizatsiyalarni ko'rib chiqadi. Interaktiv molekulyar vizualizatsiya, ma'lumotlar bazalarining vizualizatsiyasi ichida eng qadimgi tarmoqlaridan biri bo'lib, kompyuterdan oldingi davrda chuqur o'rganilgan edi. Ushbu maqola so'nggi yigirma yil ichida eng ko'p rivojlangan va yaqinda paydo bo'lgan soha biomolekulyar tuzilmalarning interaktiv vizualizatsiyasini ko'rib chiqadi. Oddiy materiya atomlar va molekulalardan iborat bo'lib, ular o'z navbatida protonlar, neytronlar va elektronlarni o'z ichiga oladi. Protonlar va neytronlar yadroviy kuchlar bilan bog'lanib, atomlarning yadrolarini hosil qiladi. Musbat zaryadlangan yadrolar manfiy zaryadlangan elektronlarni tortadi; kvant mexanik ta'siri tufayli zarralar to'qnashmaydi, lekin elektronlar barqaror va elektr neytral *atomlarni* o'z ichiga olgan ma'lum masofalarda yadrolarni o'rab oladi. Bu kimyoviy elementning eng kichik birliklaridir. Atomdagi elektronlar orbitallarda, ya'ni fazoning mintaqalarida tashkil etilgan bo'lib, ularda elektronlar katta ehtimollik bilan qoladi. *Molekulyar vizualizatsiyaning* asosiy maqsadi molekulyar tuzilmalar, ularning xususiyatlari va o'zaro ta'sirini tushunarli qilish orqali boy, murakkab moddiy dunyo haqidagi tushunchamizni qo'llab-quvvatlashdir. Bundan tashqari, u farmatsevtik faol birikmalar yoki o'ziga xos xususiyatlarga ega moslashtirilgan moddalar kabi yangi molekulalarning "ratsional" dizaynini qo'llab-quvvatlashga qaratilgan.

Keyingi tadqiqotlar shuni ko'rsatdiki, agar kovalent bog'lanishlar yangitdan hosil bo'lmasa yoki uzilmasa hamda bir vaqtning o'zida tizimning nozik xatti-harakatiga sozlangan energiya qiymatlariga sezgir bo'lmasa, molekulyar tizimlarni klassik tarzda taxminan yaxshi tasvirlash mumkin. Molekulyar *dinamikada (MD) simulyatsiyalarda*, molekulyar orbitallar hisoblanmaydi. Buning o'rniga atomlar kvant effektlarini taqlid qiluvchi sun'iy ko'p tanali kuchlar ("kuch maydonlari") ta'siri ostida harakatlanadigan klassik ob'ektlar sifatida qaraladi. Neytral atomlar va molekulalar orasidagi kuchli itarilish tufayli atomlarni taxminan "qattiq" sharlar deb hisoblash mumkin. Bu shuni anglatadiki, atomlar to'liq massasi, radiusi va boshqa atomlarga ta'sir qiladigan ko'p tanali kuchlari bilan tavsiflanadi, "ichki" elektron erkinlik darajalari e'tiborga olinmaydi. MD simulyatsiyalarining ko'pchiligi, xususan, biomolekulyar tuzilmalar ushbu "klassik" yaqinlashuv yordamida amalga oshiriladi. Shunday qilib, van der Vaals sferalarining tasviri Li va Richards (1971) ishidan boshlangan zamonaviy molekulyar kompyuter grafikasining boshlang'ich nuqtalaridan biri edi. Bu ish 40 yildan ortiq davom etdi va molekulalarning fazoviy mavjudligini ifodalovchi molekulyar sirtlarning keyingi turlari ixtiro qilindi. Biroq, biologik tizimlarning ayrim turlari batafsil tushunish uchun kvant mexanik mulohazalarini talab qiladi. Rivojlanayotgan "kvant biologiyasi"ning ommabop ilmiy taqdimotlari uchun biologik va tibbiy ahamiyatga ega bo'lgan misollar fermentativ reaksiyalar yoki fotosintezdir. Bu molekulyar vizualizatsiya bo'yicha yangi tadqiqot sohasini ochadi,

biz bu haqda juda qisqacha ma'lumot beramiz. James Watson va Francis Crick tomonidan 1953 yilda DNKning fazoviy tuzilishini kashf qilish va "Nature" jurnalida "Dezoksiribonuklein kislotaning tuzilishi", nomli maqola chop etilgandan beri, zamonaviy tibbiyot va biologiya inson tabiatini tushunish va kasalliklarga tashxis qo'yish va davolashning yangi usullarini ishlab chiqish borasida katta yutuqlarga erishdi. Oxirgi yillarda molekulyar biologiya tadqiqotlarida YaMR, immunofluorescence, mikroskopiya, PZR (polimeraza zanjir reaksiyasi), FRET (Fluorescence Resonance Energy Transfer) kabi usullarning qo'llanilishi, sitometriyaning turli usullarida to'qima kulturalar, qon hujayralari va asos hujayralari ajratish va saqlash usullari organizm uchun funktsional muhim oqsillarning sintezi, signalizatsiya va metabolik ta'sirlar, tirik hujayralarning strukturalarini o'rganish uchun asos bo'lib xizmat qildi. SHu jumladan elektrofiziologiya sohasida keng qo'llaniladigan "patch clamp" asosida hujayralar kichik guruhning oqimlarini ro'yxatga olish imkonini beruvchi turli modifikatsiyadagi usullar, tutash ion kanallari yoki bitta kanallarning fiziologik xolatlari tadqiqot qilinmoqda. Yuqorida keltirilgan usullar aniq usullar qatoriga mansub bo'lib, ular asosida organizmda kechadigan jarayonlar to'g'risida ma'lumot olish mumkin emas, lekin xali bu ma'lumotlar to'liq shakllangan emas. Oxirgi yillarda olimlar asosiy e'tibori irsiy axborot makromolekulalarining tuzilishi (DNK va RNK), ularning alohida qismlarining vazifasi va birlamchi, ikkilamchi, uchlamchi va to'rtlamchi strukturalar oqsil makromolekulalari, ularning o'zgaruvchanligi, faolligi, fizik va biologik xususiyatlarini tadqiqot qilishga qaratilgan.

Biologik ma'lumotlarning vizualizatsiyasi intuitiv tasavvurda tushirish hisoblash biologiyasining xal qiluvchi vazifalaridan biri hisoblanadi. Ma'lumki, inson 90% ma'lumotni ko'rish organlari orqali oladi, shuning uchun tadqiqotchilar makromolekulalarni va ular komplekslarini vizualizatsiya qilish oqsil strukturasining fazoviy tuzilishi ularning guruhlarining, oqsil-DNK komplekslari, Oqsil-ligand komplekslari bilan samarali ish olib borishda katta ahamiyatga egadir va maxsus dastur ta'minotlari mavjud bo'lib, ma'lumotlar bankidan kerakli strukturalarni tanlab olishda va tadqiqot qilingan molekulalarni tasvirlab berish yoki guruhlarini tushirish va tahlil qilish uchun tasvirlash xususiyatiga egadir. Xuddi shuningdek, bu paketlar murakkab bo'lmagan animatsiyalarni tuzishda, bir yoki bir nechta uslub asosida molekulalar potentsial energiyasini minimalizatsiyalashda, de novo sharoitida tadqiqot qilingan qismlarni modellashtirishda ya'ni in silico da mutatsiyalar hosil bo'lishi uchun sharoit yaratib berishi, tugunlar hosil bo'lishi, gomologlar rekonstruksiyasida aminokislotalar radikallarining rotamerlarining o'zgarishi, dokingni amalga oshirishda (retseptor va ligand bog'lanishining bashorat qilinishida) vositalarga ega bo'lib, yangi dori vositalarini loyihalashtirishda katta ahamiyatga egadir. Bugungi kunda bunday dasturiy ta'minotlarga oqsillarni modellashtirish va vizualizatsiya qilishda Deep View

Swiss PDB Viewer (the Swiss Institute of Bioinformatics ishlab chiqqan) Dastur ma'lumotlar ustida bazali operatsiyalar o'tkazish xususiyatiga ega. Dastur GROMACS96 uslubi yordamida molekular potentsial energiyasini imkoniyatini hisoblash va minimalizatsiyalash xususiyatiga ega bo'lib, gomologlarni modellashtirishni amalga oshiradi, aminokislotalar ketma-ketligini taxrirlaydi, va oqsil molekularining strukturasi taxrirlaydi. Foydalanuvchi polipeptid zanjiri bo'yicha bazali operatsiyalar quya oladi-tugunlar tuzishi, mutatsiyalar hosil qilishi, torzion burchaklar diagrammasi nazoratida zanjir konfiguratsiyasini o'zgartira oladi.

### **Ikkinchi savol bayoni**

**Biologik makromolekulalarning birlamchi strukturasi asosida ularni vizualizatsiyalashtirish.** *In vitro* tajribalari quyidagi uchta texnikaga asoslangan molekulyar strukturaviy ma'lumotlar uchun asosiy manbani ta'minlaydi; kristallarni olish mumkin bo'lgan eng yuqori aniqlikdagi ma'lumotlarga olib keladigan *rentgen kristallografiyasi*; *yadro magnit-rezonans spektroskopiyasi* (NMR) yagona tuzilmani emas, balki strukturaviy ansamblarni aniqlash; *krio-elektron mikroskop* (kriyo-EM) katta tuzilmalarni aniqlash imkonini beradi, lekin cheklangan tasvirga asoslangan rekonstruksiyaning talab qiladi. Vizualizatsiya tuzilmani aniqlash jarayoniga tasvirni qayta ishlash va tasniflash algoritmlariga qo'shimcha sifatida yordam berishi mumkin.

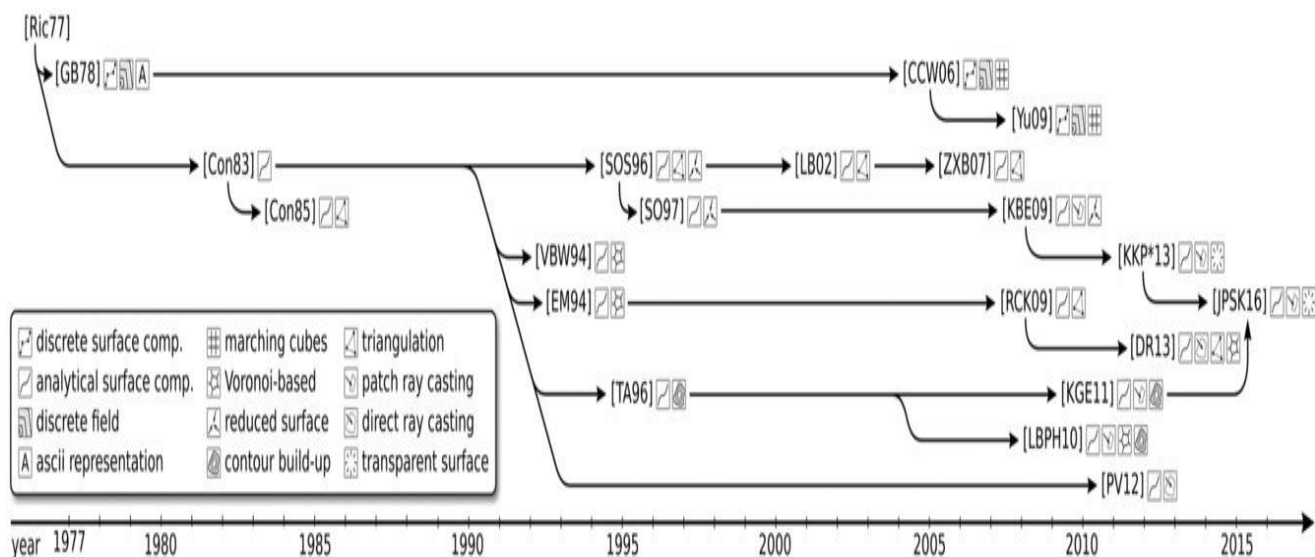
*Molekulyar simulyatsiya* oldindan aniqlangan molekulyar tuzilmalarning dinamik harakatlarini o'rganishning foydali usuli hisoblanadi. Bu olimlarga turli xil atrof-muhit parametrlarining ta'sirini va boshqa molekular bilan o'zaro aloqasini o'rganish imkonini beradi. Zamonaviy GPU tezlashtirilgan kvant mexanikasi simulyatsiyalari hali ham faqat kichik oqsillarni simulyatsiya qilishi mumkin. Shunday qilib, kattaroq tizimlar uchun eng ko'p ishlatiladigan usullar Monte-Karlo (MC) namunalari va molekulyar dinamika (MD) simulyatsiyalari hisoblanadi.

Vertikal o'q vizuallashtirilgan asosiy ma'lumotlarning shkalasiga mos keladi. Bu o'q uzluksiz bo'lsada, molekulyar vizualizatsiya bo'yicha ikkita asosiy sohaga bo'linishi mumkin. *Intramolekulyar shkala* atom shkalasi bo'yicha atomistik ma'lumotlardan tortib dag'al taneli molekulyar modellarigacha. Molekulyar *shkala* mezoskopik darajagacha bo'lgan dag'al modellarni qamrab oladi. Bu yerda butun molekular yagona ob'ekt sifatida qaraladi. Ma'lumotlarning haqiqiy miqyosi asosan ma'lumotlarni yig'ish bilan bog'liq, masalan, NMR tomonidan olingan molekulyar tuzilmalar yoki mezoskopik hujayra ichidagi simulyatsiya natijalari. Molekulalararo shkaladagi strukturaviy ma'lumotlarning atomistik shkaladagi detallar, ya'ni alohida atomlar bilan almashtirilishi bunday ko'paytirishga misol bo'la oladi. Bundan tashqari, filogenetik daraxtlar va boshqa biomolekulyar ma'lumotlar kabi qo'shimcha bioinformatika ma'lumotlarini ham kiritish mumkin.

Yuqoridagi rasmdagi rangli joylar keyingi bo'limlarda muhokama qilingan turli tushunchalarga mos keladi. Ularning joylashuvi odatda tegishli usullar va algoritmlar

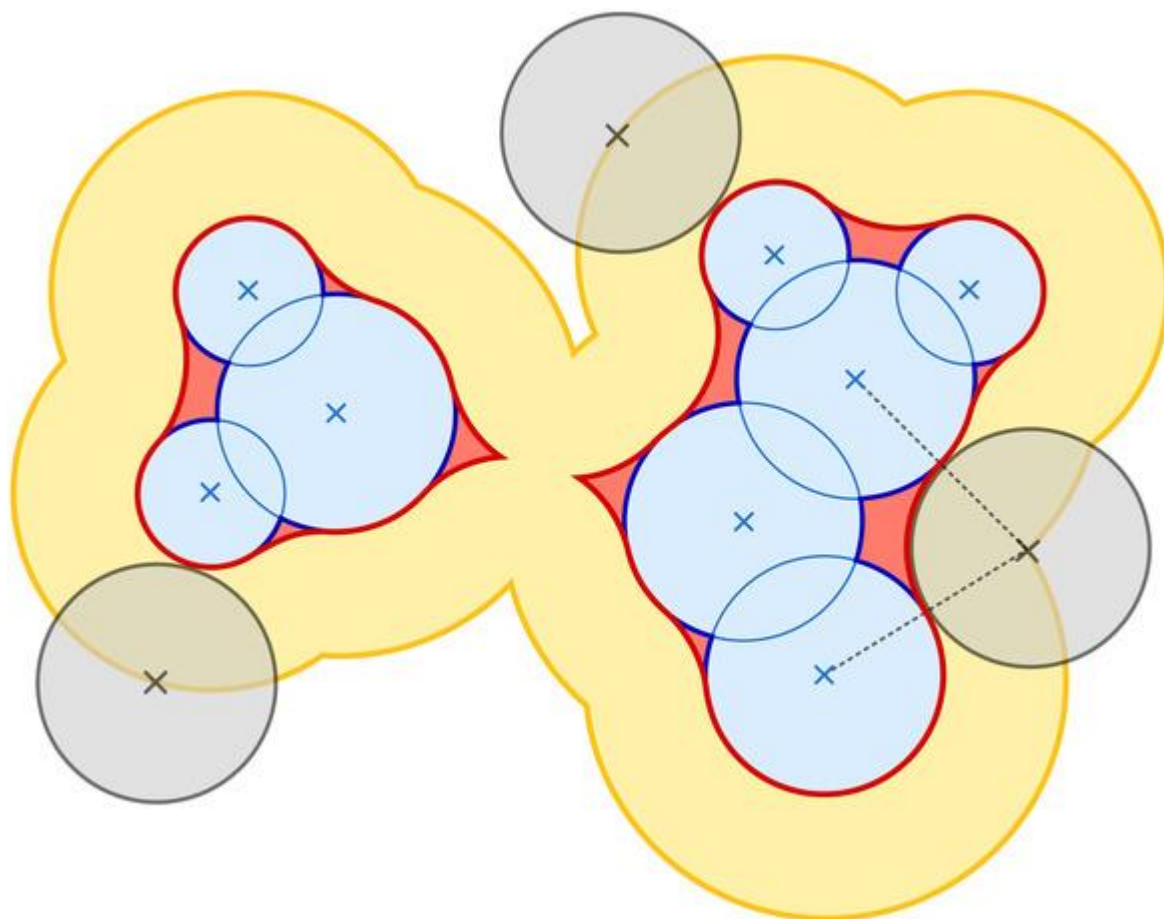
qo'llanilishi mumkin bo'lgan vizualizatsiya va ma'lumotlar shkalasi turiga to'g'ri keladi. Ushbu tasvirlarni atomistik modellarga tasviriy va mavhum va strukturaviy tafsilotlar darajasiga ( 4.3-bo'lim) bo'lish mumkin.). Ular molekulyar miqyosda statik va dinamik atributlarni ko'rish uchun qo'llanilishi mumkin. Istisnolardan biri yuqorida tavsiflangan boyitishdan foydalangan holda molekulalararo miqyosda atomistik tafsilotning tasviridir ( 4.3 -bo'limga qarang ). Qolgan maydonlarni molekulyar dinamikani vizualizatsiya qilish muddati ostida umumlashtirish mumkin ( 6 -bo'lim ). Bunga moslashuvchanlikni vizuallashtirish (qizil, 6.1 -bo'lim ), hajmli tasvirlar va agregatsiya (sariq, 6.2 -bo'lim ), interfaol va boshqariladigan simulyatsiyalar (to'q sariq, 6.3 -bo'lim ), molekulyar reaksiyalarning vizualizatsiyasi (binafsha, 6.4 -bo'lim ), va kvant effektlarini vizualizatsiya qilish (ko'k, 6.5-bo'lim.). 5 -bo'limda tasvirlangan molekulyar ko'rsatish usullari taksonomiyaga kiritilmagan, chunki ular odatda molekulyar vizualizatsiyalarning ko'pchiligiga tegishli. Kimyoda tasvirlangan molekulaning turli atributlarini ko'rsatadigan ko'plab uch o'lchovli molekulyar modellar ishlab chiqilgan. Ma'lumotlarni vizualizatsiya qilish uchun ishlatiladigan molekulyar modelni tanlash mo'ljallangan tahlil vazifasiga bog'liq. **Modellar** yuqoridagi rasmda tasvirlangan taksonomiyada ko'rsatilganidek, atomistik va mavhumba bo'linishi mumkin. Katta molekulyar tizimlar ko'pincha batafsil vizualizatsiya darajasidan foydalangan holda tasvirlanadi **ular Goodsell** tomonidan aniqlanganidek, atom detallarini soddalashtiradigan doimiy tasvirlarni o'z ichiga oladi.

1977 yilda Richards SAS g'oyasi asosida birinchi silliq molekulyar sirtni aniqladi. Atomlar ustida aylanadigan probning markazini olish o'rniga, u sferik probning chegarasidan foydalanishni taklif qildi. Bu ikkala oldingi yuzalarning afzalliklarini, vdW sirtini yaxshiroq o'lchamli ko'rsatishni va SAS ning mavjud vizualizatsiyasini birlashtiradi. Greer va Bush muqobil ta'rifni berdi va bu Richards ta'rifiga mutunosib edi. Ular sirtni molekulaning hech qanday atomiga kirmaydigan barcha mumkin bo'lgan zond sohalari birlashuvining topologik chegarasi sifatida aniqladilar. *Ularning ishlari erituvchidan tashqari sirt (SES)* atamasini ishlab chiqdi. Rasmda SES vizualizatsiyasiga oid barcha nashrlarning umumiy ko'rinishini beradi. Matematik jihatdan SES uch turdagi yamoqlardan iborat: Qavariq sferik yamalar prob aynan bitta atomga tegib turgan joyda paydo bo'ladi; toroidal yamoqlar - bu zond aynan ikkita atomga tegadigan izlar; konkav sferik yamoqlar prob aniq uchta atomga tegib, sobit holatda yotgan joyda paydo bo'ladi. Ikki yoki undan ortiq yamoqlar bir-biriga mos keladigan yamoq chegaralarida sirt  $C^1$  - uzluksiz, ya'ni SES silliq bo'ladi. Biroq, sirt o'z-o'zidan kesishmalarni o'z ichiga olishi mumkin, ular "singularliklar" deb ham ataladi. Bu erda sirt o'tkir qirralarga ega va faqat  $C^0_{dir}$ -davomiydir.



Erituvchisiz yuzalarni vizualashtirishga qaratilgan nashrlarning xronologiyasi va o‘zaro bog‘liqligini ko‘rsatadigan grafik.

SESni hisoblash algoritmlari ikki toifaga bo‘linadi. Birinchisi, bo‘shliqni diskretlash orqali sirtni hisoblaydigan barcha usullarni o‘z ichiga oladi  $\mathbb{R}^3$ . Ushbu yondashuvlar odatda izo-sirt olinadigan diskret skalyar maydonni hisoblab chiqadi, yo Marching Cubes orqali triangulyatsiya yoki to‘g‘ridan-to‘g‘ri izosurfa nurlarining yurishi. Ushbu tadqiqot sohasidagi eng tezkor ikkita yondashuv Can va Yu boshqalar tomonidan taqdim etilgan. Ushbu algoritmlarni amalga oshirish odatda oson bo‘lsa-da, hisoblash vaqti va xotira talablari tarmoq o‘lchamlari bilan kubik ravishda ortadi. Ikkinchi toifa barcha yamoqlarning yashirin sirt tenglamalarini aniqlash orqali sirtning analitik tasvirini hisoblaydigan barcha usullarni o‘z ichiga oladi. 1983 yilda Connolly SES ni analitik tarzda hisoblash uchun tenglamalarni va shunga asoslangan birinchi algoritmni taqdim etdi.



vdW sirtining 2D sxemasi (ko'k), SAS (sariq) va SES (qizil). SAS va SES vdW yuzasida aylanadigan sferik zond (kulrang) bilan belgilanadi.

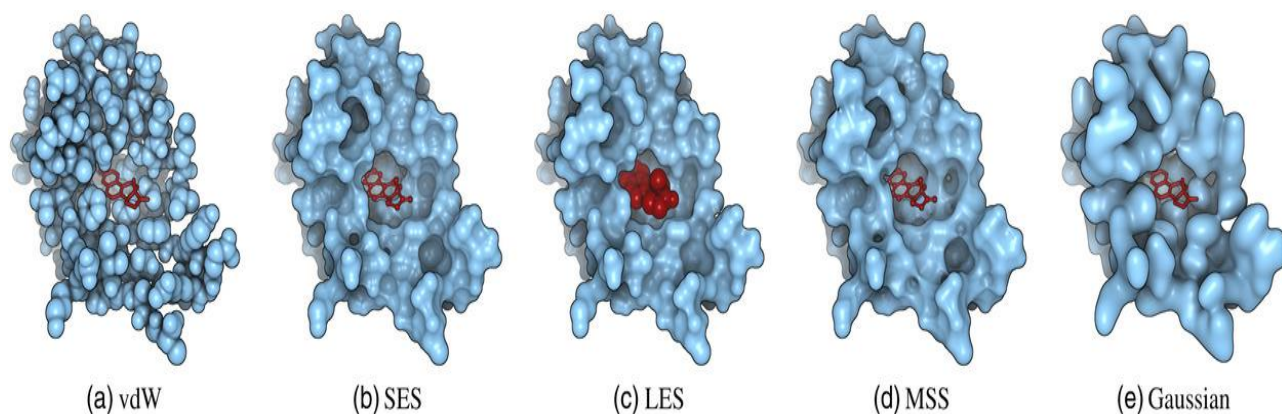
Krone va *boshqalar* GPU lar uchun CB algoritmini parallellashtirdi hamda, bu SES hisoblashni yanada tezlashtirdi va  $10^5$  atomgacha bo'lgan dinamik molekulalarni interaktiv vizualizatsiya qilish imkonini berdi. Bu ikki usul hozirda SESni hisoblashning eng tezkor analitik usullari hisoblanadi. Ko'rsatish uchun SES an'anaviy tarzda mozaiklangan. Juda aniq mozaikalarga misollar Laug va Borouchaki, Sanner va *boshqalar* tomonidan berilgan va Keyinchalik, Zhao va *boshqalar* triangulyatsiyani soddalashtirish uchun yamoqlarni spline sirtlari bo'yicha yaqinlashtiruvchi triangulyatsiyani taklif qildi. Keyinchalik *eng tezkor usullardan biri* bo'linish sirtlari yordamida Ryu va *boshqalar* tomonidan taklif qilingan. Biroq, ularning yondashuvi barcha mumkin bo'lgan o'ziga xosliklarni bartaraf eta olmaydi. SESni triangulyatsiya qilish hisoblash qimmat va odatda o'rta o'lchamdagi oqsillar uchun soniyalarni oladi. Shunday qilib 2009 yilda Krone va *boshqalar* uch turdagi sirt yamoqlarini ko'rsatish uchun GPU-ga asoslangan nurli to'qimalardan foydalangan. Yuqorida aytib o'tilganidek, u nafaqat piksellik tasvir sifatini beradi, balki kvartal tenglamalarni echish kerak bo'lsa ham, ancha tezroq hisoblanadi. Krone va *boshqalar*. shuningdek, nurli quyish yordamida SES yamoqlarining o'z-o'zidan kesishishlarini boshqargan. Lindow va *boshqalar* ibtidoiylarning rasterlanishini optimallashtirish uchun geometriya shaderidan foydalanadigan, taxminan 30% tezroq bo'lgan biroz takomillashtirilgan



nurli to‘qimalarni taqdim etdi. Nurlarni quyish samaradorligini optimallashtirish uchun SES ichida joylashgan konveks sharsimon yamoqlarning qismlari oldingi usullarda kesilmagan. Shunday qilib, sirtni faqat shaffof bo‘lmagan holda yoki old yuzni oddiy aralashtirish bilan ko‘rish mumkin edi. Yarim shaffof yoki qirqilgan vizualizatsiya, ammo bu yamoqlarni to‘liq kesishni talab qiladi. Buning uchun yechim Kauker va *boshqalar* tomonidan tasvirlangan. Yaqinda Jurkík va *boshqalar* Krone va *boshqalarning* tezkor GPU tezlashtirilgan SES hisobiga asoslangan SESning yaxshilangan shaffof renderini taqdim etdi. Tafsilotlarni STAR-da Patane va Spagnuolo tomonidan yaratilgan molekulyar yuzalar uchun geometrik va yashirin modellashtirish bo‘yicha topish mumkin.

Decherchi va Rocchia triangulyatsiya va ray quyish kombinatsiyasini taqdim etdi. Ularning algoritmi SES ning analitik tavsifini hisoblab chiqadi va 3D panjara bo‘ylab nurlarni quyishni amalga oshiradi, undan sirt Marching Cubes yordamida triangulyatsiya qilinadi. Garchi ular SES triangulyatsiyasini tezlashtirishga muvaffaq bo‘lishsa-da, umumiy tezlik va vizual sifat to‘g‘ridan-to‘g‘ri nurlanish bilan raqobatlasha olmaydi.

*Edelsbrunner po‘st yuzasi* deb ataladigan cheklangan kirish sharlari to‘plami uchun yangi silliq sirtni taqdim etdi. Uning shakli bitta parametrغا bog‘liq  $s \in (0, 1]$ , qisqarish omili. Molekulyar *teri yuzasi* (MSS) teri sirtini atomlarning vdW sohalariga qo‘llashdir. MSS ning SESga nisbatan asosiy afzalligi shundaki, sirt butunlay  $C^1$  - uzluksiz Bundan tashqari, u kvadriklarning yamoqlariga ajralishi mumkin. Biroq, MSS biofizik ma'lumotga ega emas. Kruithof va Vegter MSS uchun mozaik yondashuvni taqdim etdilar. Cheng va Shi yuqori sifatga erishadigan, lekin juda ko‘p vaqt talab qiluvchi triangulyatsiya algoritmini ishlab chiqdi. Decherchi va Rocchia tomonidan SES yondashuvi bilan bir xil strategiyadan so‘ng juda tez triangulyatsiya taqdim etildi. Biroq, u to‘liq sirt topologiyasini saqlab qolishi shart emas. Tez, yuqori sifatli vizualizatsiyaga erishish uchun Chavent va *boshqalar* MSS-ning birinchi GPU-ga asoslangan nurli to‘qimasini taqdim etdi. MSSni qurish uchun ularni amalga oshirishning uzoq muddatlari dinamik molekulyar ma'lumotlardan foydalanishga to‘sqinlik qildi. 2010 yilda Lindow va *boshqalar* Varshney va *boshqalar* bilan bir xil fikrdan foydalangan holda tezlashtirilgan hisoblashni taqdim etdi. SESni hisoblash uchun qo‘llaniladi. Ular, shuningdek, MSS ning nurlanishini optimallashtirishdi. Ikkala takomillashtirish natijasida bir necha ming atomli dinamik molekulalarning MSS interaktiv vizualizatsiyasi mumkin bo‘ldi.



Protein izomerazasining turli molekulyar sirtlari o'rtasidagi taqqoslash (PDB ID: 1OGZ). (a) vdW yuzasi, (b) zond radiusi 1,4 Å bo'lgan SES, (c) ekilenin uchun LES, (d) 0,35 qisqarish faktorli MSS va (e) atom radiusiga teng bo'lgan standart og'ish bilan Gauss konvolyutsiya yuzasi. Ligand ekilenin (qizil) (a), (e) tayog, (b), (d) shar-and-stick yoki (c) vdW yuzasi sifatida ko'rsatilgan. Barcha misollarda chuqurlikni aniqlash, ekran-kosmosdagi muhit okklyuziyasi va siluetlar qo'llanilgan.

### ***Ligandan tashqari sirtlar***

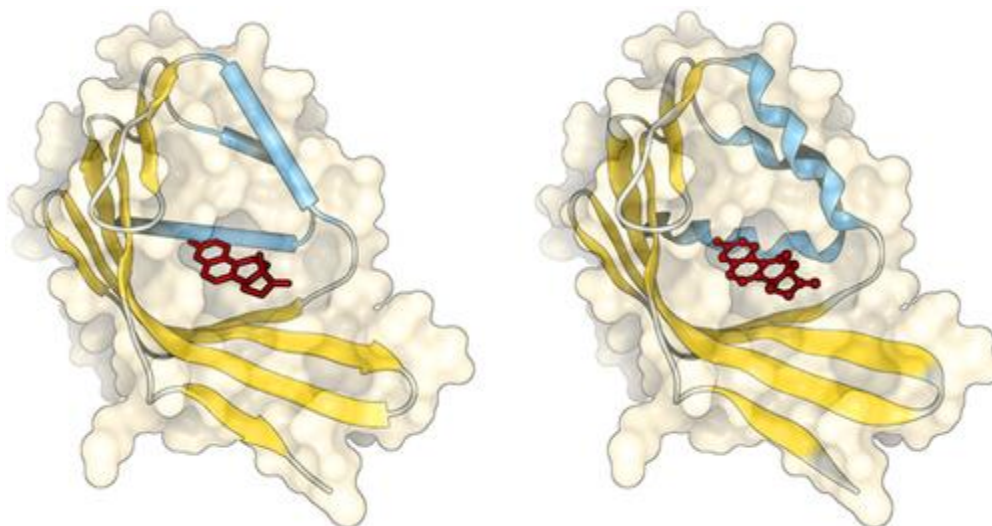
Yaqinda ligandan tashqari sirt va SES ning umumlashtirilishi Lindow va boshqalar tomonidan taklif qilingan. Bu usul SESdan farqli o'laroq, LES ligandni shar bilan yaqinlashtirmaydi, lekin ligandning vdW sirtlari bilan aniqlangan to'liq va potentsial dinamik geometriyadan foydalanadi. Shunday qilib, LES molekulaga yaqinlashganda ma'lum bir ligand kirishi mumkin bo'lgan geometrik sirtni ko'rsatadi. LESni analitik hisoblash qiyin. Lindow va boshqalar shuning uchun mumkin bo'lgan ligand o'rinlari, yo'nalishlari va dinamikasini diskretlash orqali sirtni hisoblash algoritmini taklif qildi. LES ma'lum bir ligand uchun eng aniq kirishni ta'minlasa-da, o'rta o'lchamdagi oqsillar va o'rtacha sirt sifati uchun uni hisoblash bir necha daqiqa davom etadi. Shunday qilib, agar interaktivlik zarur bo'lsa, SES qulaydir. Agar statik molekulaning batafsilroq ko'rinishi kerak bo'lsa, LESni afzal ko'rish kerak.

Molekula atomlarini bevosita tasvirlaydigan molekulyar modellardan tashqari bir qancha mavhum modellar ham yaratilgan. Mavhum model molekulaning atomistik modelda aniq va oson sezilmaydigan yoki hech bo'lmaganda aniq bo'lmagan xususiyatini ko'rsatishi mumkin. Ushbu modellar, shuningdek, siyrak tasvirlarga olib kelishi mumkin, ularni tushunish osonroq yoki okklyuzionni kamaytirish mumkin. Mavhum tasvirlash juda katta molekulyar komplekslar uchun ham foydali bo'lishi mumkin, ular uchun ko'pincha alohida atomlar emas, balki umumiy shakl qiziqish uyg'otadi.

Juda erta, murakkab makromolekulyar birikmalarning kontseptualizatsiyasi olimlarni ushbu ob'ektlarni ifodalovchi kompyuter grafik tasvirlarini soddalashtirishga undadi. Molekulyar arxitekturaning vizual abstraksiyasi ko'pincha muhim strukturaviy xususiyatlarni to'liq tafsilotli atomistik tasvirga qaraganda aniqroq ko'rsatadi. Masalan, molekulyar subbirlik tuzilmalari uchun abstraksiyalardan

foydalanish va boshqalar. Goddard va Ferrin muqobil ravishda molekulyar birikmalarning asosiy tuzilmaviy ierarxiyasiga mos keladigan ko'p darajali tafsilotlar kabi abstraksiyalarga murojaat qiladilar. Biologik tuzilmalar haqidagi tushunchamiz rivojlanib borayotganligi sababli, nuklein kislota polimerlari va yaqinda uglevodlar asoslarini ifodalash uchun yangi abstraksiyalarga ehtiyoj paydo bo'lishi mumkin.

Vehlow va boshqalar oqsil tarkibidagi aminokislotalarning kontakt xaritalarini 3D tasviri bilan birga ko'rsatadigan vositani taqdim etdi. Foydalanuvchilar oqsil tuzilishini tahlil qilishlari va oqsilning turli qatlamlarining aminokislotalar kontaktlarini solishtirishlari mumkin. Vizualizatsiya **Ramachandran** chizmalaridan ilhomlantirilgan bo'lib, ular oqsilning orqa miya burilish burchaklarini ko'rsatadi. Bu chizmalar oqsillarning ikkilamchi struktura elementlarini (masalan, spirallar yoki varaqlar) aniqlash uchun hamda eksperimental ravishda olingan tuzilmalar sifatini ko'rsatuvchi ko'rsatkich sifatida ishlatiladi.



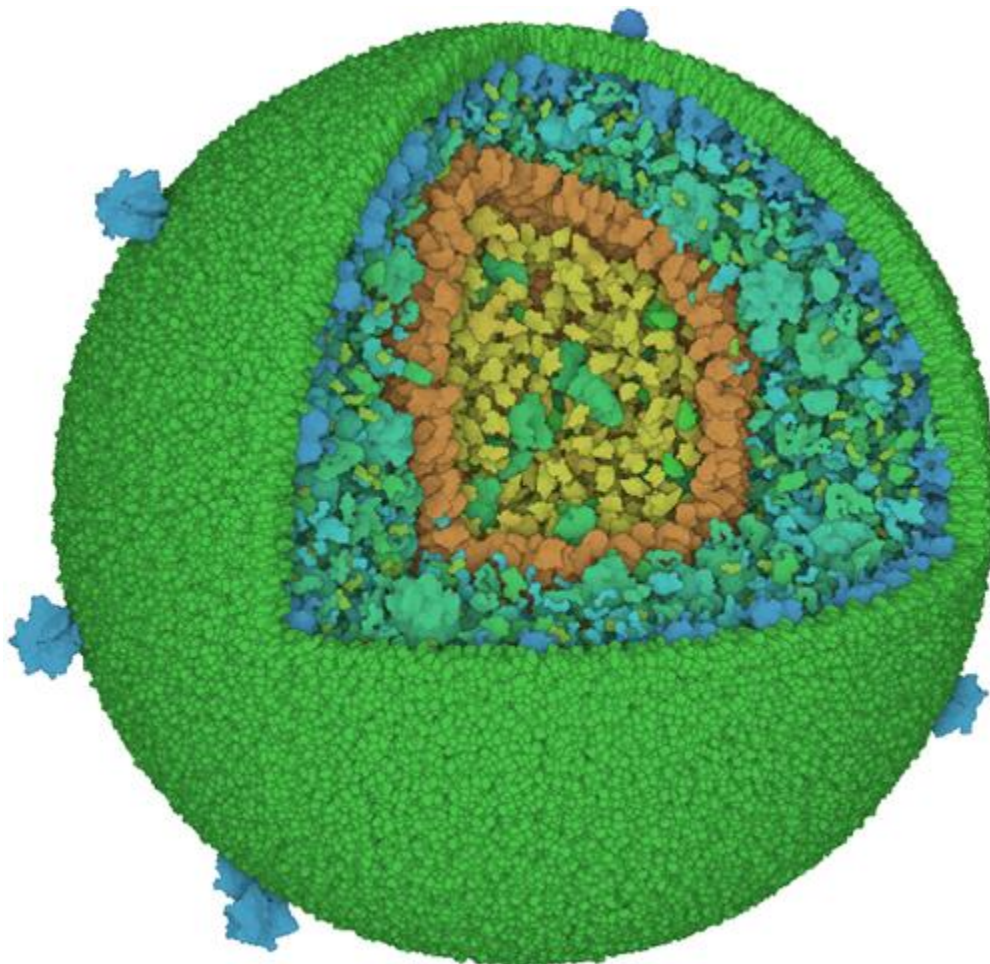
Xuddi shu oqsilning ikkita mumkin bo'lgan multfilm renderlari (PDB ID: 1OGZ). Chapda: lenta shaklidagi o'qlar b-varaqlar uchun aminokislotalar zanjirining yo'nalishini ko'rsatadi, a-spirallar silindr shaklida stilize qilingan. O'ngda: dumaloq lentalar varaqlar va spirallarni tasvirlash uchun ishlatiladi. Malumot uchun yarim shaffof SES ko'rsatilgan. Ligand ekilenin (qizil) shar va tayoq tasvirida ingl.

Mavhum tasvirlar DNK va RNK uchun ham qo'llaniladi. DNK odatda zinapoyaga o'xshash qo'sh spiral bilan tasvirlangan, bu fosfat-shakar magistralini lenta yoki naycha orqali va nukleotid asoslarini tayoqchalar yoki ellipsoidlar bilan ifodalaydi. Ko'pgina asboblarda bunday tasvirlar mavjud, masalan, *VMD*, *PyMOL* yoki *Chimera*. Ellipsoidlar, shuningdek, molekularning turli sinflaridagi turli strukturaviy elementlar uchun umumiy abstraksiya shakli sifatida ham qo'llaniladi. Bernier va boshqalar tomonidan yaratilgan RiboVision ribosomalardagi RNK tuzilishini ko'rish uchun maxsus vositadir. U bog'langan ko'rinishlardan foydalangan



holda 1D chizmalar, 2D ketma-ketlik diagrammalari va 3D vizualizatsiya kombinatsiyasidan foydalanadi. Bu foydalanuvchilarga RNK molekulalarining tuzilishini har tomonlama tahlil qilish imkonini beradi.

Molekulyar sirt to‘rini (odatda uchburchakli SES) sferik koordinatalar tizimiga xaritalashning bir qancha usullari taklif qilingan. Rahi va Sharp [RS07] molekulyar yuzaning uchburchaklarini sharga xaritalash uchun sferik koordinatalarga asoslangan parametrlashdan foydalanadigan usulni ishlab chiqdilar. Postarnakevich va Singx texnikasi SESga mos kelguncha chegaralovchi sharni deformatsiya qilish uchun kuchga yo‘naltirilgan yondashuvdan foydalanadi va shu bilan SES va sfera o‘rtasida xaritani yaratadi. Ushbu xaritalashdan foydalanib, sharni molekulaning fizik-kimyoviy xususiyatlariga ko‘ra yoki asl SES shaklini ta’kidlash uchun sfera deformatsiyasining yo‘l uzunligiga qarab rang berish mumkin. Xass va Kohl molekula qanchalik sharsimon ekanligini o‘lchash uchun molekulyar sirt va chegaralovchi sfera o‘rtasidagi konformal xaritadan foydalaning. Shuningdek, ular molekulalarni solishtirish uchun o‘zlarining sharsimon tasviridan foydalanishni taklif qilishadi.



cellVIEW [LMAPV15] tomonidan yaratilgan virusning (OIV) tasviriy vizualizatsiyasi.

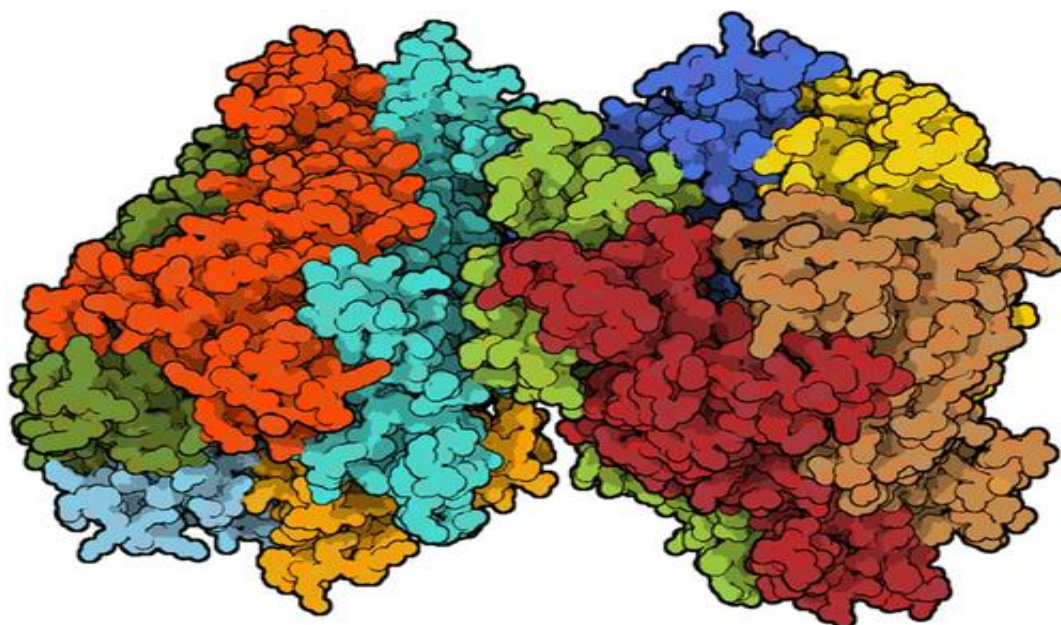
Ushbu yondashuv milliardlab atomlar bilan uzunlik shkalasi bo‘yicha beshta buyurtmani bog‘laydigan atom detallaridagi biologik sahnalarni interaktiv ravishda

vizualizatsiya qilish uchun ishlatilishi mumkin. Ko'p o'tmay, Falk va boshqalar ierarxik LOD yordamida texnikani tezlashtirdi: agar to'r katakchasining proyeksiyasi pikseldan kichik bo'lsa, bu katakdagi sharlar uchun nurlanishni amalga oshirish kerak emas. U faqat hujayra bo'sh yoki yo'qligi tekshiriladi. Xuddi shu narsa butun panjara bir pikseldan kichikroq bo'lganda ham amal qiladi. Shuningdek, ular sahnani bir nechta renderlash paslariga bo'lishdi. Har bir o'tishda oldingi o'tishning chuqurlik buferi keraksiz nurlarni quyish operatsiyalarini oldini olish uchun chuqurlik sinovi uchun ishlatiladi. Ular, shuningdek, foydalanuvchiga molekulyar yuzalar kabi murakkab modellarni tasavvur qilish imkonini beruvchi uchburchakli ob'ektlar misollari uchun yondashuvning umumlashtirilishini taqdim etdilar.

Parulek va boshqalar molekulyar sirtlarni tez ko'rsatish uchun LOD usulini joriy qildi. Ularning usuli chiziqli interpolyatsiya yordamida uchta molekulyar sirt tasvirini - SES, Gauss konvolyutsiyasi yuzasi va vdW sirtini birlashtiradi. Tegishli modelni tanlash, kameradan masofaga qarab sahnani uchta maydonga ajratadigan muhimlik funksiyasidan kelib chiqadi. Ierarxik abstraksiya LODni yanada ta'kidlaydigan moslashtirilgan soyani o'z ichiga oladi. Ishlashni yaxshilash uchun A-bufer texnikasi qo'llaniladi.

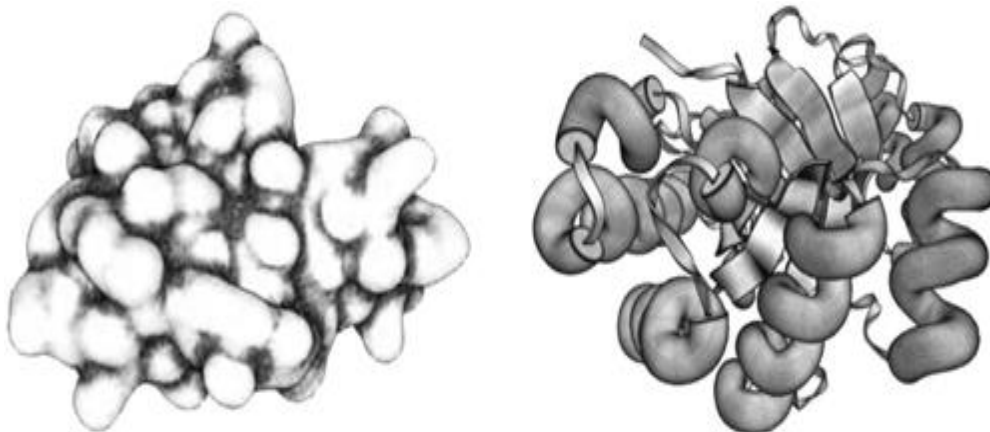
**Molekulyar renderlash.** Molekulyar dinamika ma'lumotlarining vizualizatsiyasi ko'pincha murakkab va yuqori chuqurlikdagi murakkablikdan tashqari yuqori vizual murakkablikka ega. Haqiqiy vaqtda renderlash va soylashning ilg'or usullari nafaqat tasvir sifatini yaxshilaydi, balki sahnadagi geometrik shakllar va chuqurlik murakkabligini idrok etishni ham yaxshilaydi. Molekulyar vizualizatsiya bilan bog'liq asosiy jihatlar soyalar va turli xil chuqurlik belgilari, shu jumladan atrof-muhit okklyuziyasi effektlari. Ushbu kontekstda eng ko'p qo'llaniladigan texnikalar quyida muhokama qilinadi. Quyida sanab o'tilgan barcha usullarning umumiy jihati shundaki, ularni real vaqtda dinamik ma'lumotlar uchun hisoblash mumkin.

Devid Gudsell tomonidan bajarilgan hujayralarning molekulyar ichki qismining qo'lda chizilgan rasmlaridan ilhomlanib, komiksga o'xshash ko'rinishga ega badiiy yoki fotorealistik bo'lmagan renderlarni yaratish uchun ko'pincha *toon soyasi* qo'llaniladi. Bu turdagi soyalar B-Raf oqsiliga qo'llaniladi.



Goodsell tomonidan "Oyning molekulasi" uchun ishlatgan uslubga o'xshash ikkita oqsilning fotoreal bo'lmagan renderi (PDB ID: 4A97). MegaMol [GKM\*15] yordamida yaratilgan tasvir.

Xususiyat chiziqlari va lyuklardan iborat *chizmalardan* foydalangan holda tasviriy tasvirlar molekulyar renderlashda uzoq an'anaga ega. Chiziqli chizmalar haqida umumiy ma'lumot uchun [RCDF08] ga qarang. Xususan, kontur chiziqlari molekulyar vizualizatsiyada keng qo'llaniladi (masalan, [TCM06, LVRH07, KBE09]). Goodsell va Olson molekulyar sirtlarni tasvirlash uchun bir necha turdagi lyuklardan foydalanadilar. Atomistik model va oqsilning multfilm modeli [vdZLBI11] o'rtasida uzluksiz abstraktsiyani olish uchun kontur chiziqlari va lyukirovka ham qo'llanilgan. Weber tomonidan ProteinShader vositasi oqsillarning multfilm tasvirlari uchun chiziqli real vaqt rejimida tasviriy renderlashni taklif etdi. Lawonn va *boshqalar* molekulyar yuzalardagi muhim xususiyatlarni ta'kidlash uchun xususiyat chiziqlari va lyuklarni birlashtirgan. Bu usul yorug'lik gradientining vektor maydonidagi chiziqli integral konvolyutsiyaga (LIC) asoslangan bo'lib, u sezilarli sirt hududlarini ta'kidlaydi. Yuqoridagi oqsillarni tasviriy vizualizatsiya qilish uchun misollar ko'rsatilgan.





Ikki molekulaning tasviriy chiziqli renderlari: sirt tuzilishi (chapda, tasvir manbai: [LKEP14] ; PDB ID: 1OGZ) va multfilm tasviri (o'ngda, ProteinShader [Web09] ; PDB ID: 1RWE bilan yaratilgan).

Atrof-muhit okklyuziyasi (AO) - *Miller va Jukov va boshqalarning* ishlariga asoslangan usulda ko'rsatishicha ob'ektlar orasidagi tarqoq yorug'likning o'tkazilishini taqlid qiladi, bu chuqurlikni idrok etishni oshirishi mumkin bo'lgan burmalarda mahalliy soyaga olib keladi. AO zich zarracha ma'lumotlar to'plamlari uchun eng yaxshi ishlaydi, bu esa uni ko'pgina molekulyar ma'lumotlar vizualizatsiyasi uchun mos qiladi. Yuqoridagi rasmda mahalliy yoritish va OSAO o'rtasidagi farqlar ko'rsatilgan. AO hisoblash qiymat bo'lgani uchun interaktiv vizualizatsiya uchun bir nechta tezlashtirilgan yondashuvlar ishlab chiqilgan. Screen-Space AO (SSAO) - bu qayta ishlashdan keyingi bosqichda AO effektlarini yaqinlashtiruvchi tasvir-makon texnikasi, masalan [Kaj09]. Molekulyar ma'lumotlar to'plamlari uchun Object-Space AO (OSAO) texnikasi yanada ishonchli natijalar berishi mumkin. OSAO faqat ko'rinadigan mahallani hisobga oladigan SSAO yondashuvlaridan farqli o'laroq, butun mahalliy mahallani ko'rib chiqadi. *Grottel va boshqalar* juda katta, dinamik zarrachalar ma'lumotlar to'plamlari uchun ham interaktiv kadrlar tezligiga yetadigan OSAO usulini ishlab chiqdi. Usul atrof-muhitni okklyuziv omillarni saqlash uchun mahalliy mahallaning volumetrik yaqinlashuvidan foydalanadi. *Yaqinda bu yondashuv Staib va boshqalar* tomonidan kengaytirildi. [SGG15] to'liq rangli AO xaritasidan namuna olishni yaxshilaydigan ierarxik voksel-konusni kuzatish usulidan foydalanish. Ularning usuli shaffof zarralar uchun ham ishlaydi. *Eichelbaum va boshqalar* global va mahalliy tizimli ma'lumotlarni saqlashga qaratilgan zarrachalarni ko'rsatish uchun SSAO usuli bo'lgan PointAO taqdim etdi. *Wahle and Wriggers* biomolekulalarning strukturaviy xususiyatlarini ta'kidlash uchun mo'ljallangan ko'p miqyosli SSAO usulini ishlab chiqdi. *Hermosilla va boshqalar* halo va AO effektlarini yaratish uchun interaktiv usulni taqdim etdi. 4 -rasmda molekulyar sirtlar uchun chuqurlik, siluet va SSAO kombinatsiyasi tasvirlangan. Yuqorida aytib o'tilgan interaktiv AO yondashuvlari molekulyar vizualizatsiya uchun eng ko'p qo'llaniladigan usullardir, chunki AO usullarining to'liq ro'yxati ushbu hisobot doirasidan tashqarida bo'ladi.

Aniq ob'ekt chegaralari oqsillar yoki simulyatsiya natijalari kabi ko'plab ob'ektlarga ega sahnalar uchun foydali *chuqurlik belgisidir*. Chuqurlikka bog'liq *siluettlarni* [ST90] tasvir maydonida chuqurlikdagi uzilishlarni aniqlash va shunga mos ravishda chiziq kengligini sozlash orqali qayta ishlashdan keyingi bosqichda hisoblash mumkin. *Xuddi shunday ta'sir Tarini va boshqalar* tomonidan taklif qilinganidek, ob'ekt chegaralaridan cho'zilgan *halolarni* qo'llash orqali olinadi. [TCM06]. Ob'ektning chegarasida halo ob'ekt bilan bir xil chuqurlikka ega. Ob'ektdan masofa ortishi bilan halo chuqurligi ham ortadi. Shunga o'xshash texnika, *Luft va boshqalar tomonidan chuqur qorayish* yondashuvi. [LCD06], uzoqdagi bir-birining ustiga tushadigan



ob'ektlarni vizual ravishda ajratib turadi va tasvir maydonida chuqurlikka bog'liq haloslarni yaratadi. Qo'shimcha chuqurlik belgilari sifatida oddiy tumanlash yoki chuqurlikka bog'liq desaturatsiyadan foydalanish mumkin.

Odatda, ko'rish nuqtasi va kamera parametrlari molekulyar sahnalarni ko'rsatish va o'rganishda foydalanuvchi tomonidan tanlanadi. Muayyan molekula uchun eng yaxshi ko'rinishni avtomatik tanlash 3D strukturani ekranga joylashtirish uchun tizimli ma'lumotlardan tashqari qo'shimcha ma'lumotni talab qiladi. *Vazqéz va boshqalar* nuqtai nazaridan entropiya tushunchasidan foydalanadi va uni orfografik molekulyar ko'rinishlarga kengaytiradi. Proteinga qo'shimcha semantik ma'lumotni kiritish optimal kamera sozlamalarini tanlashni yaxshilashi mumkin.

Ma'lumotlarning shakli va chuqurligi murakkabligini ta'kidlaydigan renderlash usullaridan tashqari, stereoskopik ko'rsatish molekulyar grafikada keng qo'llaniladi (masalan, [GF07]). Stereoskopik renderlash boshga o'rnatilgan displeylar (HMD), 3D ko'zoynaklar yoki avtostereoskopik ekranlar kabi maxsus jihozlarni talab qilsa-da, ko'rsatish qismi odatda nisbatan sodda: har bir ko'z uchun har birida tegishli kamera sozlamalari bilan alohida tasvir beriladi. Shubhasiz, renderlash ham ikki baravar hisoblash quvvatini talab qiladi. Yaqinda Stone *va boshqalar* real vaqt rejimida nurlarni kuzatishdan foydalanadigan Oculus Rift HMD uchun masofaviy renderlash tizimini taqdim etdi. Shunisi e'tiborga loyiqki, shaklni ajratib ko'rsatish uchun yuqorida ko'rsatilgan renderlash usullaridan foydalanish stereoskopik ko'rsatish uchun ehtiyotkorlik bilan ko'rib chiqilishi kerak, chunki bu usullar monoskopik ko'rsatish uchun mo'ljallangan va idrok etish muammolariga olib kelishi mumkin.

Molekulalar o'z-o'zidan moslashuvchan mavjudotlardir, ammo vizualizatsiyalarning aksariyati statik strukturaviy suratni ifodalaydi. Pozitsiyaviy noaniqlikni hisobga olish uchun aniq belgilangan atom pozitsiyalari o'zgaruvchan molekulyar konformatsiyalarni [RJ99] tasvirlash uchun ehtimollik taqsimoti bilan almashtirilishi mumkin. Dinamik molekulyar konformatsiyalar uchun taqdimotlar Shmidt-Ehrenberg *va boshqalar* tomonidan qo'shimcha tekshirildi. [SEBH02]. Ular atom yoki qoldiq xususiyatlarni tasvirlash uchun rangni o'z ichiga olgan panjaraga shar-and-stick va vdW tasvirlarini namuna olish usulini ishlab chiqdilar. Shunday qilib hisoblangan konformatsion loyqalik so'ngra izosurface yoki to'g'ridan-to'g'ri hajmni ko'rsatish yordamida ko'rsatiladi. *MolMol* [KBW96] va bir nechta boshqa dasturlar ushbu usulga o'xshash "kolbasa" ko'rinishini taqdim etadi, bu erda protein magistral naychasi kabi mavhum tasvirlar oldindan hisoblangan moslashuvchanlik parametriga muvofiq modulyatsiya qilinadi. Olingan trubaning kengligi moslashuvchanlikni ta'kidlaydi. Li va Varshney ko'p qatlamli yarim shaffof yuzalar orqali atomlarning termal tebranishlarini tasvirlagan. Proteinlardagi halqalar yoki domenlar kabi tanlangan moslashuvchan elementlar voksel xaritalari bilan ifodalanishi mumkin. Bryden *va boshqalar* oddiy rejim tahlilidan hisoblangan molekulyar

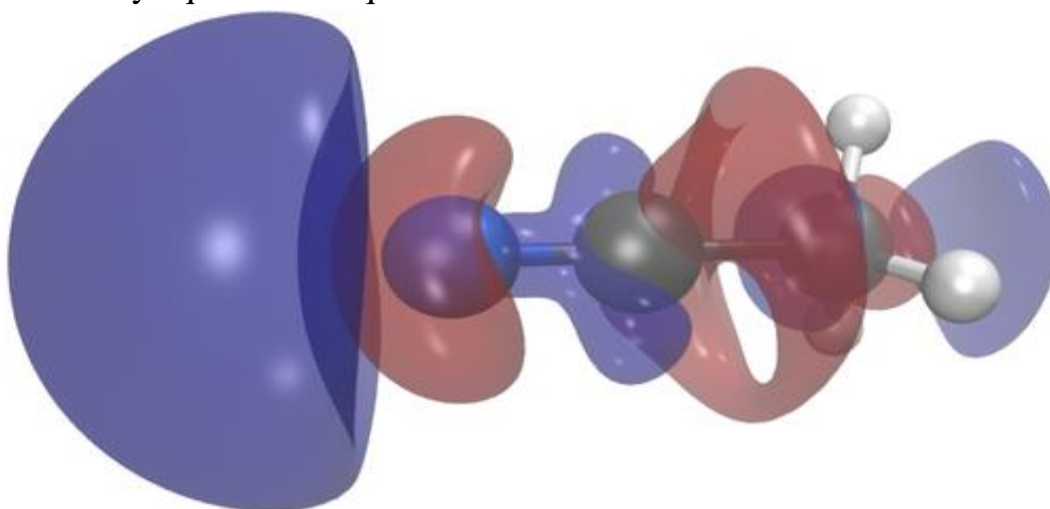
moslashuvchanlikni ko'rsatish uchun gliflardan foydalangan. Ularning yondashuvi sinxronlashtirilgan aylanish harakatini ko'rsatadigan atomlar guruhlarini birlashtiradi. Klasterlar ta'kidlangan va aylanishni tasvirlaydigan mos keladigan dumaloq yo'lar bilan jihozlangan. Ushbu yo'ylarning tepasidagi o'qlar aylanish yo'nalishini va tezlik, xatolik yoki qattiq bo'lmagan energiya kabi boshqa qiymatlarni ko'rsatadi. Fioravante va boshqalar oqsillardagi harakat korrelyatsiyalarini tahlil qilish uchun asosiy komponentlar tahlili va kovariatsiya klasteridan foydalanadigan vizualizatsiya usullarini taqdim etdi. Ushbu tahlillar natijalari oqsil tuzilishining 3D vizualizatsiyasini boyitish uchun ishlatiladi, masalan, rang yoki konus gliflari yordamida.

Yaqinda Dabdoub va boshqalar molekulyar strukturaning tanlangan atomlari, masalan, magistral atomlarning yo'llarini ko'rsatish orqali molekula dinamikasini vizualizatsiya qiluvchi *MoFlow asbobini* taqdim etdi. Vaqt bosqichlari orasidagi atom pozitsiyalari splinelar yordamida interpolatsiya qilinadi. Olingan egri chiziqlar vaqt shkalasi rang xaritasiga ko'ra ranglanadi, bu esa atomlarning vaqt o'tishi bilan harakatlarini oson tushunish imkonini beradi. Bog'larning harakatini aks ettiruvchi yarim shaffof lentalar orqali ko'proq vizual belgilar qo'shiladi. *MoFlow* traektoriyaning qisqa qismlarini oson tushunishga imkon bersa-da, vizual tasvir juda murakkab harakatlar uchun tezda chalkashishi mumkin.

Vizualizatsiya interaktiv simulyatsiyalarning muhim elementidir. Foydalanuvchi simulyatsiyani to'g'ri boshqara olishi uchun vizualizatsiya interaktiv bo'lishi kerakligi sababli, simulyatsiya ishlashi odatda asosiy cheklovchi omil hisoblanadi. Interaktivlik 1960-yillardan beri molekulyar grafiklar uchun maqsad bo'lib kelgan. O'sha paytda o'zaro ta'sir asosan kamera harakatini boshqarishni anglatardi. Faol manipulyatsiya elementi keyinroq qo'shildi, birinchi navbatda 20 dan 80 gacha qoldiqlar tizimidan boshlab energiyani minimallashtirishning ixtisoslashgan yondashuvi va oxir-oqibat elektrostatik o'zaro ta'sirlarni **qoldiruvchi**. *MDScope* interaktiv vizualizatsiya va MD simulyatsiyalari uchun bir necha yuz qoldiqgacha to'liq elektrostatika bilan boshqarishni amalga oshirdi va vaqt chegaralari masalasini ko'tardi. Ayniqsa, boshqariladigan simulyatsiyalar kontekstida ixtisoslashtirilgan o'zaro ta'sir qurilmalari yordamida haptik aloqa qiziqarli bo'ladi, chunki u kuchlarni uzatish uchun ishlatilishi mumkin. Molekulyar ma'lumotlarni to'g'ridan-to'g'ri manipulyatsiya qilish usullaridan foydalanadigan boshqa dastur sohasi - bu interaktiv molekulyar modellashtirish. Maxsus apparat yordamida intuitiv haptik tadqiqot amalga oshirildi va 4000 atomli membrana kanaliga tatbiq etildi. Haptik renderlash uchun ishlash talabi grafiklarni ko'rsatishdan ko'ra qattiqroqdir, chunki u taxminan 1000 Gts yangilanish tezligini talab qiladi. 4 -bo'limda batafsil tavsiflangan zamonaviy molekulyar vizualizatsiya usullari real vaqtda dinamik ma'lumotlarni qayta ishlashga qodir. Shunday qilib, ular interaktiv simulyatsiyalarni vizualizatsiya qilish uchun ishlatilishi mumkin. Oddiy o'rnatish yuqoridagi rasmlarda ko'rsatilgan. Hozirgi vaqtda

hatto kimyoviy reaktivlikni ham QM simulyatsiyalari yordamida interaktiv tarzda o'rganish mumkin. Bunday interaktiv tajribalar Kreylos va boshqalar tomonidan muhokama qilingan vizual manipulyatsiya bo'yicha qo'llanmalar tomonidan osonlashtiriladi. Ular molekula ichidagi cheklovlar uchun teskari kinematikadan foydalanadigan *ProteinShop* molekulyar modellashtirish vositasini ishlab chiqdilar hamda, bu foydalanuvchilarga optimal oqsil qatlamlarini aniqlashga yordam berdi. Hozirgi vaqtda arzonroq uskuna, yaxshi grafik kartalar va tezroq kompyuterlar bilan haptik boshqaruv juda jozibador bo'lib qoldi va hatto bir milliondan ortiq atomdan iborat tizimlarga ham qo'llanilishi mumkin.

Vizualizatsiya tomonida maydonlar va ko'p maydonlarni tasvirlash talab qilinadi. Bunday maydonlarni vizuallashtirish boy va hayratlanarli hodisalarni (eng oddiy molekulyar tizimlarda ham sodir bo'ladigan) ochishga yordam berishi ko'rsatildi [BHI\*09, ABB\*11, HKM\*11]. Elektron zichlik maydonlarining topologik vizual tahlili alohida atomlarga tegishli bo'lgan fazoviy domenlar haqida ma'lumot beradi [Bad90]. Kovalent va kovalent bo'lmagan aloqalarni [GBCG \* 14], zaif o'zaro ta'sirlarni [JKMS \* 10, CGJK \* 11] va molekulyar orbitallarni [SSH \* 09] vizual tahlil qilish uchun bir qator usullar va vositalar mavjud ( 16 -rasmga qarang .) hamda tegishli elektron zichliklari [HG08]. Bundan tashqari, hajm ko'rsatish yordamida hosil bo'lgan loyqa molekulyar sirtlarni vizualizatsiya qilish taklif qilindi



Asetonitril molekulasida uchun molekulyar orbitallarning tasviri (VMD [HDS96] bilan yaratilgan).

So'nggi o'n yilliklarda molekulyar vizualizatsiya uchun ko'plab vositalar va tizimlar paydo bo'ldi. Ulardan ba'zilar ma'lum bir maqsad uchun ishlab chiqilgan va ularning rivojlanishi to'xtagan. Boshqa tomondan, domen mutaxassisleri o'z tadqiqotlarida vizual tahlil qilish va natijalarni tarqatish uchun keng qo'llaniladigan bir nechta juda muvaffaqiyatli va mustahkam tizimlar mavjud. Biz mavjud tizimlarni to'rt guruhga bo'lishga qaror qildik: ba'zi zamonaviy usullarni birlashtirgan erkin mavjud bo'lgan funktsional jihatdan boy tizimlar, samarali algoritmlar va kengaytirilishiga qaratilgan ochiq manba prototip vositalari, tijorat tizimlari va veb-ga asoslangan echimlar. Ushbu bo'lim ushbu toifaga qarab tuzilgan.

Birinchi turkumda *VMD* [HDS96], *PyMOL* [DeL02], *Chimera* [PGH\*04], *YASARA View* [KV14] yoki *CAVER Analyst* [KSS\*14] kabi mustahkam va ommabop vositalar mavjud. Ushbu tizimlar notijorat maqsadlarda bepul mavjud va shuning uchun ilmiy hamjamiyat tomonidan keng qo'llaniladi. Ushbu tizimlarning ba'zilari, shuningdek, o'z plaginlarini qo'shish orqali hissa qo'shadigan foydalanuvchilar hamjamiyatidan olinadi. Aksariyat tizimlar 4 -bo'limda muhokama qilingan molekulyar modellarning barcha asosiy ko'rinishlarini qo'llab-quvvatlaydi.. Ko'pgina vositalar qo'shimcha ravishda an'anaviy molekulyar modellarni molekulyar tizimdagi turli fizik-kimyoviy xususiyatlar va munosabatlar (masalan, atom zichligi, molekulyar orbitallar, qutblanish yoki elektrostatik potentsial va maydonlar) haqida qo'shimcha ma'lumot bilan jihozlash uchun vositalarni taqdim etadi. Ularning to'g'ri vizual tasviri bog'lanish va boshqa munosabatlar haqida muhim tushuncha berishi mumkin. Molekulyar orbitallar ( 16 -rasmga qarang) GPU tezashtirilgan algoritmlar [SHLK11] yordamida dinamik ma'lumotlar uchun hisoblanishi va ko'rsatilishi mumkin.. VMD, ximera va PyMOL kabi vositalar, shuningdek, foydalanuvchilarga oddiy tarmoqlarda saqlangan dala ma'lumotlarini yuklash imkonini beradi, keyin ularni to'rni ajratib olish, izo-konturlar yoki hajmni ko'rsatish orqali ko'rish mumkin. Ular, shuningdek, elektrostatik ma'lumotlar uchun foydali bo'lishi mumkin bo'lgan maydon chizig'ini vizualizatsiya qilishni taklif qiladilar. *Molden* paketi [SN00], *Molekel* [PL00], *Gabedit* [All11 ], *GaussView* [ DKM09], *Chemcraft* [And15 ] va kabi fizik-kimyoviy xususiyatlarni molekulyar vizualizatsiya qilish uchun turli xil ixtisoslashgan mustaqil vositalar ham mavjud. *Avogadro* [HCL\*12]. Ushbu vositalarning barchasi, shuningdek, VMD va Chimera ham kub fayllardan o'qiladigan molekulyar orbitallarni (ular *Gauss* [FTS\*09] yoki *GAMESS-US* [SBB\*93] kabi asboblardan tomonidan chiqariladi ) yoki ko'rish imkoniyatiga ega. Vizualizatsiya vositalari orqali bevosita hisoblab chiqiladi.

Oxirgi toifa molekulyar ko'rsatish uchun veb-ga asoslangan echimlarni o'z ichiga oladi. Garchi bunday vositalar, odatda, texnik cheklovlar tufayli ushbu hisobotda ko'rsatilgan eng so'nggi zamonaviy texnikalarni birlashtira olmasa ham, ularni eslatib o'tish joiz, chunki ular bugungi kunda yirik molekulyar komplekslarni interaktiv vizualizatsiya qilish imkoniyatiga ega. Ular tizimli ma'lumotlar bazalaridagi yozuvlarning ixtisoslashtirilgan vizualizatsiyasini yoki tuzilish bilan bog'liq hisob-kitoblar natijalarini taqdim etish uchun veb-saytlarga joylashtirilishi mumkin. Hali ham eng ko'p foydalaniladigan veb-asosli vositalardan biri Java-applet *Jmol* [jmo09] dir. U ko'plab fayl formatlarini yuklashni, molekulyar sirtlarni, orbitallarni, sxematik multfilmlarni va boshqa xususiyatlarni ko'rsatishni qo'llab-quvvatlaydi. *OpenAstexViewer* [Har02] tuzilishga asoslangan dori dizaynida yordam berishga qaratilgan yana bir Java-ga asoslangan dastur. U applet sifatida ham, mustaqil dastur sifatida ham foydalanish mumkin. Boshqa funktsiyalar bilan bir qatorda, shaffoflik va mulkiy xaritalash bilan soyali molekulyar sirtlarni taqdim etadi. *JSmol*, Java o'rniga

faqat HTML5 va JavaScript ishlatadigan Jmol kengaytmasi hozirda ishlab chiqilmoqda. *NGL viewer* [RH15] va *iview* [LLNW14] brauzerda apparat tezashtirilgan renderlash imkonini beruvchi WebGL-dan foydalanadigan zamonaviy, funksiyalarga boy veb-asoslangan vositalarga misoldir. Molekulyar vizualizatsiya kutubxonasi *3Dmol.js* [RK15] JavaScript va WebGL dan ham foydalanadi. Shuningdek, u molekulaning ko'pgina standart tasvirlarini, jumladan yarim shaffof molekulyar sirtlarni va orbitallarni vizualizatsiya qilishni qo'llab-quvvatlaydi. Yuqorida aytib o'tilganidek, veb-ga asoslangan vizualizatsiya vositalari veb-ga asoslangan grafiklarning cheklovlari tufayli uchburchakka asoslangan ko'rsatishga tayanadi. Yaqinda Mwalongo va boshqalar. [MKK\*14, MKB\*15] WebGL brauzerda GPU-ga asoslangan nurlarni uzatish imkonini berishini ko'rsatdi. Bunday texnologik yutuqlar, ehtimol, yaqin kelajakda yanada rivojlangan veb-ga asoslangan molekulyar vizualizatsiyaga olib keladi.

Birinchi mavzuga kelsak, molekulyar dinamikani shunchaki fenomenologik darajada tasvirlash o'rniga, molekulyar hodisalarning fizik va kimyoviy *sabablarini* ham klassik, ham kvant mexanik darajada tasvirlash kerak. Hisoblash klasterlari va simulyatsiya usullari takomillashgani sayin, hal qilinishi mumkin bo'lgan erkinlikning kvant mexanik darajalari soni ortadi. Shuning uchun dinamik molekulyar tizimlarda kvant hodisalarini tasvirlash uchun yangi vizualizatsiya usullari kerak bo'ladi.

Murakkab molekulyar tuzilmalar uchun chuqurlikni idrok etishni yaxshilash usullari allaqachon keng ko'lamda o'rganilgan ( 5 -bo'limga qarang ). Biroq, joriy vizual tasvirlarni qo'shimcha ishoralar bilan oshirish imkoniyatlari hali ham mavjud (masalan, [SVGR15] ). Vizual tartibsizlikni maxsus kesmalar, ochish yoki portlash kabi yangi tasviriy vizualizatsiya usullarini ishlab chiqish orqali hal qilish mumkin.

Hozirgi vaqtda ushbu darajalarni mazmunli ravishda etkazadigan hech qanday abstraktsiya mexanizmi mavjud emas, masalan, multfilm tasvirlari ikkinchi darajali tuzilmalar uchun. Ehtimol, bunday yirik molekulyar komplekslar uchun biz tez orada quinar tuzilmani va hatto undan tashqarida ham mazmunli ta'rif va vizual tasvirni tekshirishga guvoh bo'lamiz.

### Uchinchi savol bayoni

**PyMol dasturida ishlash.** PyMOL ochiq manbali, foydalanuvchi homiyligidagi molekulyar grafik tizim bo'lib, o'rnatilgan Python tarjimoniga ega hamda, real vaqt rejimida vizuallashtirish va yuqori sifatli molekulyar grafik tasvirlar va animatsiyalarni tez yaratish uchun mo'ljallangan. PyMOL skript bo'lib, Python tilidan foydalanib kengaytirilishi mumkin. PyMOL faol loyiha bo'lib, unga tez-tez yangi xususiyatlar qo'shiladi. Dunyo bo'ylab o'n minglab olimlar o'zlarining molekulyar ma'lumotlarini ko'rish, almashish va tahlil qilish uchun PyMOLni tanlaydilar. PyMOL biologik



maqsadlarning 3D tuzilmalarini ko'rsatish uchun kuchli vosita bo'lib, 12 tagacha turli stereo vizualizatsiya rejimlarini taklif etadi. Foydalanuvchilar maqsadlardagi turli muhim tarkibiy xususiyatlarni, xususan, dori molekulalari uchun mos bog'lanish joylarini samarali tarzda ajratib ko'rsatishlari va farqlashlari mumkin. Yaqinda o'tkazilgan sharhda PyMOL-da vizualizatsiya va tahlilni yaxshilash, oqsil-ligand modellashtirish, molekulyar simulyatsiya va dori skriningi uchun mo'ljallangan turli xil molekulyar modellashtirish modullari muhokama qilinadi. Yangi dasturiy ta'minot va ilovalar kundalik hayotimizga, shuningdek, giyohvand moddalarni topish bo'yicha tadqiqotlarimizga qulaylik keltiradi. Aksariyat kasalliklar tanadagi oqsillar yoki DNKning noto'g'ri ishlashi tufayli yuzaga kelganligi sababli, olimlar fiziologik javoblarni tuzatish uchun ushbu biologik maqsadlarga bog'laydigan va ular bilan o'zaro ta'sir qiladigan to'g'ri molekulani topishga intilishadi. Birinchidan, olimlar maqsadda dori molekulasi bog'lanishi mumkin bo'lgan joylarni aniqlashlari kerak. Ikkinchidan, ko'p miqdordagi potentsial dori molekulalari ularning maqsad bilan o'zaro ta'siri uchun tekshiriladi. Va nihoyat, olimlar molekulaning nishonga ta'sirini tushunishlari kerak. Birgalikda bu vazifalar to'liq va qiyin. Kasallik bilan bog'liq maqsadlarning uch o'lchovli (3D) tuzilmalarini vizualizatsiya qilish orqali olimlar asosiy muammolarni yaxshiroq tushunishlari va tegishli muammolarni hal qilish uchun yangi molekulyar konstruktsiyalarni loyihalashlari mumkin. PyMOL biologik maqsadlarning 3D tuzilmalarini ko'rsatish uchun kuchli vosita bo'lib, 12 tagacha turli stereo vizualizatsiya rejimlarini taklif etadi. Foydalanuvchilar maqsadlardagi turli muhim tarkibiy xususiyatlarni, xususan, dori molekulalari uchun mos bog'lanish joylarini samarali tarzda ajratib ko'rsatishlari va farqlashlari mumkin. Python-ga asoslangan dasturiy ta'minot bo'lgan PyMOL dori-darmonlarni aniqlash hisob-kitoblari uchun funksiyalarini yaxshilash uchun ko'plab foydali Python-ga asoslangan plaginlarni o'rnatish afzalligiga ega. Pythonning o'zi "oson skript" dasturlash tilidir. Keng dasturlash tajribasiga ega bo'lmagan olimlar ham plaginlarni juda oson ishlab chiqishlari mumkin. Onlaynda mavjud bo'lgan ko'plab plaginlar kompyuter yordamida dori vositalarini loyihalashning asosiy bosqichlarini qamrab oladi, masalan, in situ molekulalarni tahrirlash kabi oddiy vazifalardan dori-maqsad o'zaro ta'sirini tahlil qilish kabi murakkabroq vazifalargacha. Samarali, PyMOL dori dizayni uchun universal platformaga aylandi. *WIREs Computational Molecular Science* kompaniyasining so'nggi sharhida, Shuguang Yuan, HC Stiven Chan va Zhenquan Xu PyMOL-da 3D tuzilmalarni vizuallashtirish va hisoblash dori dizaynini taqdim etish uchun ajoyib kirishni taqdim etadi.

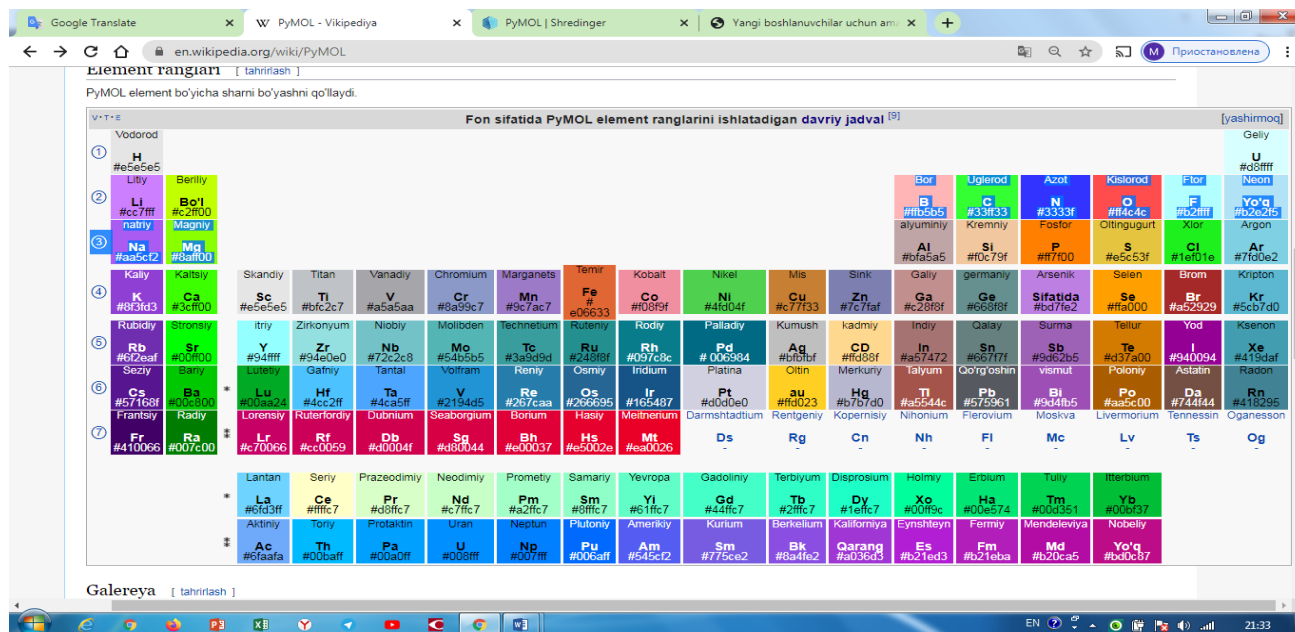
PyMolning dastlabki versiyalari Python litsenziyasi ostida chiqarilgan. 2006 yil 1 avgustda DeLano Scientific kompaniya tomonidan tarqatilgan oldindan kompilyatsiya qilingan PyMOL tuzilmalari (shu jumladan beta versiyalari) uchun boshqariladigan kirish yuklab olish tizimini qabul qildi. Ushbu bajariladigan fayllarga kirish endi

mijozlarga pul to'layotgan ro'yxatdan o'tgan foydalanuvchilar uchun cheklangan; o'quv inshootlari talabalar va o'qituvchilarga bepul taqdim etiladi. Biroq, hozirgi manba kodining aksariyati , oldingi kompilyatsiya qilingan tuzilmalar kabi bepul mavjud bo'lib qolmoqda. Boshqa platformalar uchun qurish tizimlari ochiq bo'lsa-da, Windows API (WinAPI, Win32) qurish tizimi mavjud emas, garchi norasmiy Windows ikkilik fayllari onlaynda mavjud. Har kim kompilyatsiya qilishi mumkinPython-litsenziyalangan manba kodidan bajariladigan fayl yoki oldindan kompilyatsiya qilingan bajariladigan fayllarga kirish uchun qo'llab-quvvatlash xizmatlariga obuna uchun to'lov.

2010 yil 8 yanvarda Schrödinger, Inc. PyMOLni sotib olish bo'yicha kelishuvga erishdi. Firma PyMOLni ishlab chiqish, texnik xizmat ko'rsatish, qo'llab-quvvatlash va sotishni o'z zimmasiga oldi, shu jumladan o'sha paytdagi barcha obunalar. Ular, shuningdek, PyMOL ochiq manba hamjamiyatini faol qo'llab-quvvatlashda davom etmoqdalar. 2017 yilda Shredinger Qt ostida foydalanuvchi interfeysini va Anaconda ostida paketlarni boshqarishni birlashtirish uchun tarqatish tizimini yangiladi va uni PyMol v2 sifatida chiqardi. Ushbu versiya ba'zi yangi funksiyalarni cheklaydi va agar 30 kunlik sinov muddatidan keyin litsenziyasiz foydalanilsa, vizualizatsiyaga moybo'yoq qo'shadi; umumiy litsenziya siyosati DeLano tizimiga o'xshaydi. Manba kodi asosan mavjud bo'lib qoladi, bu safar BSD-ga o'xshash litsenziya ostida. Oldingi tarqatishda bo'lgani kabi, g'ildirak formatidagi norasmiy Windows ikkilik fayllari mavjud va haqiqatan ham Linux distributivlari ochiq kodli kodning o'z tuzilmalarini taqdim etishda davom etmoqda. PyMolning dastlabki versiyalari Python litsenziyasi ostida chiqarilgan. 2006 yil 1 avgustda DeLano Scientific kompaniya tomonidan tarqatilgan oldindan kompilyatsiya qilingan PyMOL tuzilmalari (shu jumladan beta versiyalari) uchun boshqariladigan kirish yuklab olish tizimini qabul qildi. Ushbu bajariladigan fayllarga kirish endi mijozlarga pul to'layotgan ro'yxatdan o'tgan foydalanuvchilar uchun cheklangan; o'quv inshootlari talabalar va o'qituvchilarga bepul taqdim etiladi. Biroq, hozirgi manba kodining aksariyati , oldingi kompilyatsiya qilingan tuzilmalar kabi bepul mavjud bo'lib, qolmoqda. Boshqa platformalar uchun qurish tizimlari ochiq bo'lsa-da, Windows API (WinAPI, Win32) qurish tizimi mavjud emas, garchi norasmiy Windows ikkilik fayllari onlaynda mavjud.<sup>[7]</sup> Har kim kompilyatsiya qilishi mumkinPython-litsenziyalangan manba kodidan bajariladigan fayl yoki oldindan kompilyatsiya qilingan bajariladigan fayllarga kirish uchun qo'llab-quvvatlash xizmatlariga obuna uchun to'lov. 2010 yil 8 yanvarda Schrödinger, Inc. PyMOLni sotib olish bo'yicha kelishuvga erishdi. Firma PyMOLni ishlab chiqish, texnik xizmat ko'rsatish, qo'llab-quvvatlash va sotishni o'z zimmasiga oldi, shu jumladan o'sha paytdagi barcha obunalar. Ushbu versiya ba'zi yangi funksiyalarni cheklaydi va agar 30 kunlik sinov muddatidan keyin litsenziyasiz foydalanilsa, vizualizatsiyaga



moybo‘yoq qo‘shadi; umumiy litsenziya siyosati DeLano tizimiga o'xshaydi. Manba kodi asosan mavjud bo'lib qoladi, bu safar BSD-ga o'xshash litsenziya ostida. Oldingi tarqatishda bo'lgani kabi, g'ildirak formatidagi norasmiy Windows ikkilik fayllari mavjud va haqiqatan ham Linux distributivlari ochiq kodli kodning o'z tuzilmalarini taqdim etishda davom etmoqda.



PyMol dasturining elementlar rangli jadvali

## Muhokama uchun savollar

1. Molekulaning visual tasviri nima?
2. Molekulyar vizualizatsiyaning asosiy maqsadi ayting!
3. 3D struktura nima?
4. Solventsiz sirt nima?
5. Po‘stning molekulyar yuzasi deganda nimani tushunasiz?
6. Molekulyar modellarning interaktiv vizualizatsiyasi va manipulyatsiyasi nima?
7. Molekulyar reaksiyalarning vizualizatsiyasi nima?
8. Kvant effektlarini vizuallashtirish deganda nimani tushunasiz?
9. PyMol nima?

## 7-mavzu: Oqsillarning strukturasi va xususiyatlarini in silico sharoitida o‘rganish.

### Asosiy savollar:

1. In silico yaralish tarixi.
2. In silicoda qo'llaniladigan usullar.
3. In siliconing tibbiyotdagi ahamiyati.

**Mavzuga oid tayanch tushuncha va iboralar:** *In silici, farmakologiya, farmakokinetik, PD, Deskriptor, PK, QSAR, ligand, I-TASSER, M4T, RaptorX, MODELLER 9.14, IntFOLD, Phyre2, Errat, Ramachandran, Protparam Expasy, AutoDock, UCSF Chimera, LIGPLOT, Chem Draw ultra, GRAMM-X, ClusPro, PyMOL, KCNE.*

**Dars maqsadi:** Talabalarga In siliconing yaralish tarixi va qo'llaniladigan usullari hamda tibbiyotda va boshqa sohalardagi ahamiyati haqida tushuncha berish.

### **Birinchi asosiy savol bayoni.**

"*In silico*" atamasi zamonaviy so'z bo'lib, odatda kompyuterda o'tkaziladigan eksperimentni bildirish uchun ishlatiladi va *in vivo* va *in vitro*da ko'proq ma'lum bo'lgan biologik atamalar bilan bog'liq. "*In silico*" atamasining tarixi kam ta'riflangan, bir qancha tadqiqotchilar uning paydo bo'lishidagi rolini da'vo qilmoqdalar. Biroq, so'zning eng qadimgi nashr etilgan misollaridan ba'zilari Sieburg (1990) va Danchin va boshqalar tomonidan qo'llanilishini o'z ichiga oladi. (1991). Yaqin kunlarda nashr etilgan kitobda, Danchin (2002) kimyo, biologiya va farmakologiyada hisoblash vositalarining imkoniyatlarini qisqa va ishonchli tasvirlashni taklif qiluvchi iqtibosni taqdim etadi hamda informatika biologik funksiyalarni tahlil qilishda kashfiyotga haqiqiy yordamchi hisoblanadi. *Shuni hisobga olib in vivo va in vitro* eksperimentlarga qo'shimcha sifatida uning ahamiyatini ta'kidlash uchun *silico* deb atagan hisoblash yondashuvining potentsialini tushunish kerak.

Asosan, bioaktiv birikmalar va biologik tizimlar o'zaro ta'sirlashganda ikkita natija mavjud (1-rasm) (Testa va Kramer, 2006). E'tibor bering, "biologik tizim" bu erda juda keng ta'riflangan va funktsional oqsillarni (masalan, retseptorlar), ko'p hujayrali organizmlardan ajratilgan bir hujayrali organizmlar va hujayralarni, ajratilgan to'qimalar va organlarni, ko'p hujayrali organizmlarni va hatto bir yoki ko'p hujayrali shaxslar populyatsiyasini o'z ichiga oladi.. Dori (yoki har qanday ksenobiotik) va biologik tizim o'rtasidagi o'zaro ta'sirga kelsak, ular "birikma biotizimga nima qiladi" va "biotizim birikmaga nima qiladi" deb soddalashtirilgan bo'lishi mumkin. Biologik tizimga ta'sir qiluvchi dori farmakologik va/yoki toksik javobni, boshqacha aytganda, farmakodinamik (PD) hodisani keltirib chiqarishi mumkin. Simmetrik tarzda, biologik tizim ksenobiotikni so'rib olish, tarqatish, metabolizm va chiqarish orqali unga ta'sir qiladi. Bular farmakokinetik (FK) hodisalari. Ammo shuni tushunish kerakki, ksenobiotiklar harakatining bu ikki jihati bir-biri bilan chambarchas bog'liq. So'rilish, tarqatish va yo'q qilish PD ta'sirining intensivligi va davomiyligiga hal qiluvchi ta'sir ko'rsatadi, biotransformatsiya esa

o'ziga xos PD ta'siriga ega bo'lishi mumkin bo'lgan metabolitlarni hosil qiladi. Aksincha, o'zining PD ta'siri bilan birikma organizmning holatiga ta'sir qilishi mumkin (masalan, gemodinamik o'zgarishlar va ferment faolligi) va shuning uchun uning ksenobiotiklarni boshqarish qobiliyati. Faqat PK/PD modellashtirishda va klinik farmakologiyada qo'llaniladigan tizimli yondashuv ushbu o'zaro bog'liqlikning global tabiatini baholashga qodir. Ushbu munozarani aniqlashtirish uchun, deb belgilash foydali bo'lishi mumkin. Ushbu munozarani aniqlashtirish uchun, deb belgilash foydali bo'lishi mumkin birikma organizmning holatiga ta'sir qilishi mumkin (masalan, gemodinamik o'zgarishlar va ferment faolligi) va shuning uchun uning ksenobiotiklarni boshqarish qobiliyati. Faqat PK/PD modellashtirishda va klinik farmakologiyada qo'llaniladigan tizimli yondashuv ushbu o'zaro bog'liqlikning global tabiatini baholashga qodir. Ushbu munozarani aniqlashtirish uchun, deb belgilash foydali bo'lishi mumkindori yoki boshqa ksenobiotik bilan o'zaro ta'siridan keyin PD hodisasini keltirib chiqaradigan turli xil biologik komponentlarga qaratilgan. Bunday maqsadlarga retseptorlar, ion kanallari, nuklein kislotalar, anabolik va katabolik fermentlar va boshqalar kiradi. Xuddi shunday, biologik komponentlar (ksenobiotik-metabolizuvchi fermentlar, tashuvchilar, aylanma oqsillar, membranalar va boshqalar) uchun vositalarga murojaat qilish mumkin, ular dori-darmonlarni metabolizatsiya qilish, tashish, tarqatish yoki chiqarish orqali ta'sir qiladi.

### **In silico yondashuvlarining tarixi va evolyutsiyasi**

Giyohvand moddalarni ishlab chiqish va giyohvand moddalarni kashf qilishda tegishli fanlar informatikaning paydo bo'lishini va fan sifatida rivojlanishini kutmadi. Albert (1971, 1985) tomonidan mohirlik bilan umumlashtirilganidek, tuzilma-faoliyat munosabatlaridagi eng dastlabki sezgi va tushunchalar XIX asrda kuzatilishi mumkin. Faoliyat va fizik-kimyoviy xususiyat o'rtasidagi bog'liqlik Meyer (1899) va Overton (1901) tomonidan mustahkam o'rnatilgan., "hujayra depressiyasining lipoid nazariyasini" taklif qilgan, shunday qilib, lipid erituvchi va suv o'rtasidagi bo'linish koeffitsienti qanchalik yuqori bo'lsa, depressant ta'siri shunchalik yuqori bo'ladi. Bunday hujjatlar lipofillik va elektron xususiyatlarni PD va PK reaksiyalarining asosiy determinantlari sifatida tan olish uchun yo'l ochdi, bu eng yaxshi Korvin Xanshning davr yaratuvchi va davom etayotgan ishi (Hansch va Fujita, 1964; Hansch, 1972) tomonidan tasvirlangan. dori dizaynining asoschisi.

Molekulyar tuzilish kontseptsiyasini tushunishning ortib borishiga parallel ravishda, XIX asr oxiri va XX asr boshlarida bir nechta ko'rgan tadqiqotchilar

(masalan, Jon Lengli, Pol Erlich va Alfred Klark; ko'rib chiqilgan ( Ariens, 1979 ; Parascandola, 1980 )) retseptorlar tushunchasi, ya'ni dori ta'sirining maqsadlari. Retseptorlar va fermentlar o'rtasidagi o'xshashlik Albert (1971) tomonidan belgilab berilgan.

Kimyo va biologiya sohasidagi bir-biriga yaqinlashib kelayotgan taraqqiyot yo'nalishlari ma'lumotlar va bilimlar to'liqligini keltirib chiqardi, ular " *in serebro* " ning odatdagi imkoniyatlaridan tashqariga chiqdi.' ma'lumotlar bilan ishlash va kompyuter fanlarining paydo bo'lishi va rivojlanishida harakatlantiruvchi kuch bo'lgan. Xansh 1950-yillarda birinchilardan bo'lib tuzilma (aslida parametrlar va tavsiflovchilar) va faoliyat o'rtasidagi miqdoriy munosabatlarga erishish uchun kalkulyator va statistikadan foydalangan. 1980 va 1990-yillarda kompyuter grafikasi va molekulyar modellashtirish orqali 1980 va 1990 yillarda miqdoriy tuzilma-faoliyat munosabatlari (QSAR) paydo bo'ldi. Biroq, kompyuter fanlari tezda dori-darmonlarni kashf qilish va farmakologiyada oddiy vosita bo'lishni to'xtatdi va taraqqiyotga katta hissa qo'shdi. Kimyo-biologiya-informatika triadasi endi o'z hayotiga aylandi va farmakologiyani yangi cho'qqilarga olib chiqmoqda, chunki bu sharh qisqacha tasvirlashga harakat qiladi.

### **Miqdoriy tuzilma-faoliyati munosabatlari**

*In silico* farmakologiyaning dastlabki davri 1960-yillarning boshlarida, kimyoviy tuzilish va biologik tizimlardagi PD va PK ta'siri o'rtasidagi miqdoriy bog'lanishlar hisoblash usullari bilan ochila boshlaganda o'rnatilishi mumkin. O'shandan beri QSARni tahlil qilish va tan olish zamonaviy dorivor kimyo va farmakologiyaning muhim tarkibiy qismiga aylandi. Dastlabki e'tibor molekularning bioaktivligi uchun hisob-kitoblarni taqdim etishga qaratilgan edi ( Hansch va Fujita, 1964).). Shunga ko'ra, nomenklaturadan aniq uzilishda kimyoviy tuzilish va biologik ta'sir (ya'ni faollik, toksiklik, so'rilish, tarqatish, metabolizm, ekskretsiya va toksiklik (ADME/Tox) yoki fizik-kimyoviy xususiyatlar) o'rtasidagi bog'liqlikni o'rnatishga qaratilgan har qanday urinish. ) aniqlik uchun ushbu sharhda odatda QSAR deb nomlanadi. Shu sababli, keng ma'noda QSARlar statistik usullar yordamida molekulyar strukturani kimyoviy xususiyat yoki biologik ta'sir bilan bog'laydigan matematik modelni qurishdan iborat. Bu, bir tomondan, turli xil molekularning turli mexanizmlar bilan harakat qilish yoki retseptor bilan turli xil bog'lanish rejimlarida o'zaro ta'sir qilish imkoniyatini ko'rib chiqishda oson ish emas, bu esa har qanday QSAR modeliga mos kelmaydigan chegaralar mavjudligiga olib keladi (Verma va Hansch, 2005 ), shuningdek, boshqa tomondan, QSAR modelini qurishda ishtirok etadigan asl ma'lumotlar va aniq uslubiy jihatlar bilan bog'liq bo'lgan ichki shovqin ( Polanski va boshq., 2006 ). Oxir oqibat, agar ishonchli biologik ma'lumotlar mavjud bo'lgan o'quv molekularlari to'plami uchun muhim korrelyatsiyaga erishilsa, model boshqa molekular uchun biologik ta'sirni bashorat qilish uchun ishlatilishi mumkin, garchi qo'shimcha sharhda tavsiflanganidek, ba'zi bir bo'lishi mumkin. e'tiborga olinishi kerak bo'lgan modelni qo'llash uchun cheklovlar ( Ekins va boshq., 2007).). Oxirgi 40 yil ichida bu harakatlar minglab QSAR modellarini yaratdi, ularning

ko'plari C-QSAR ma'lumotlar bazasida to'plangan va saqlangan ( Hansch va boshqalar , 2002 ; Kurup, 2003 ).

Natijalar shuni ko'rsatdiki, mavjud bilimga asoslangan potentsiallar hali ham universal qo'llanilishidan uzoqdir, hozirgi samaradorlik tahlil qilinadigan ligand-oqsil komplekslarining turiga kuchli bog'liqdir. Ma'lumotlardan samarali foydalanish va uzatishni maksimal darajada oshirish imkonini beradigan tezlik va aniqlikni muvozanatlash uchun protein-ligand potentsiallarini yaxshilash uchun qanday choralar ko'rish mumkinligi haqida ba'zi munozaralar bo'ldi ( Shimada, 2006 ).

**Virtual ligand skriningi.** Katta kimyoviy kutubxonalarda molekullarni ma'lum bir maqsadga yaqinlik ehtimoli bo'yicha baholash va tartiblash jarayoni odatda virtual skrining ( Oprea va Matter, 2004 ) deb ataladi. Shu nuqtai nazardan, virtual skrining QSAR kontsepsiyasini dastlab konjenerik birikmalarning kichik to'plamlariga yo'naltirilgan, mavjud sintezlangan molekullar va shuningdek, ishonchli sintez qilinadigan molekullar tomonidan aniqlangan kimyoviy o'lchov bo'ylab kengaytirishga urinish sifatida qaralishi mumkin. Bu atamaning o'zi 1990-yillarning oxirida, kompyuterga asoslangan usullar, avvaliga kutilganidan umidsizlikka tushadigan darajada yomon ishlash va yuqori xarajatlarga ega bo'lgan eksperimental yuqori o'tkazuvchanlik skrining (HTS) usullariga alternativa taklif qilish uchun yetarli etuklikka erishganida yaratilgan ( Lahana, 1999).). Yillar davomida farmatsevtika sanoati virtual skrining usullari haqiqatan ham HTSga samarali qo'shimcha bo'lishi mumkinligini qabul qilishni o'rgandi ( Stahura va Bajorath, 2004 ), ular shubhasiz bugungi kunda qo'rg'oshin ishlab chiqarish jarayonining ajralmas qismiga aylandi ( Bajorath, 2002). ; Bleicher va boshqalar , 2003 ).

Texnologiyaga asoslangan HTS dan farqli o'laroq, virtual skrining bilimga asoslangan yondashuv bo'lib, u qiziqish maqsadi uchun bioaktiv ligandlar (ligandga asoslangan virtual skrining) yoki maqsadning o'zi (maqsadga asoslangan virtual skrining) haqida tarkibiy ma'lumotlarni talab qiladi. Qiyosiy tadqiqotlar shuni ko'rsatdiki, ma'lum bioaktiv ligandlardan olingan maqsad haqidagi ma'lumot virtual skrining orqali yangi bioaktiv iskalalarni aniqlash uchun maqsadli tuzilmalar haqidagi bilimlar kabi qimmatlidir ( Evers va boshq., 2005a ; Zhang va Muegge, 2006 ). Shu sababli, foydalanish usuli bo'yicha yakuniy tanlov oxir-oqibatda ishlashga katta ta'sir ko'rsatmasdan mavjud *bo'lgan ma'lumotlarning turi va miqdoriga bog'liq bo'ladi.*

**Virtual yaqinlik profilini yaratish.** Agar virtual ligand skriningi QSARni kimyoviy o'lchov bo'ylab kengaytirsam, virtual yaqinlik profilidagi so'nggi tendentsiyalar unga qo'shimcha biologik o'lchov qo'shmoqda. Yaqinda bir nechta maqsadlarda molekullarning farmakologik profilini baholashga qodir bo'lgan yangi usullar to'liqini haqida xabar berildi. Ular optimallashtirish jarayonida birikmalarning maqsaddan tashqari yaqinliklari tufayli yuzaga kelishi mumkin bo'lgan nojo'ya ta'sirlarini aniqlash vositasi sifatida dori-darmonlarni topishga kuchli ta'sir ko'rsatishga va'da beradi ( O'Konnor va Roth, 2005 ; Paolini va boshqalar , 2006).). Shunga qaramay, shuni e'tirof etish kerakki, ushbu usullarning hozirgi gullab-yashnashi, asosan, ligandlar uchun farmakologik ma'lumotlarni olishda, shuningdek, tarkibiy ma'lumotlarni yig'ishda ma'lumotlarni to'plash, tasniflash va saqlashga bag'ishlangan

ba'zi muvofiqlashtirilgan tashabbuslar tomonidan erishilgan muhim yutuqlarning natijasidir. oqsillar uchun ( Mestres, 2004 ).

**Ligandga asoslangan usullar.** Ligandga asoslangan yaqinlik profilini aniqlash usullarini ishlab chiqish an'anaviy kimyoviy omborlarga adabiyotga asoslangan farmakologik ma'lumotlarni o'z ichiga olgan izohli kimyoviy kutubxonalarini qurishdan katta foyda keltirdi ( Savchuk va *boshqalar.* , 2004.). Ular orasida WOMBAT ma'lumotlar bazasi (Sunset Molecular Discovery LLC, Santa Fe, NM, AQSH) so'nggi 30 yil ichida tibbiy kimyo jurnallarida e'lon qilingan 120 400 molekula uchun biologik ma'lumotlarni taqdim etadi. MDL Drug Data Report yoki MDDR (MDL Information Systems, San Ramon, CA, AQSH) patent adabiyotlari, jurnallar va kongresslardan to'plangan 132 000 dan ortiq birikmalar uchun terapevtik ta'sir va biologik faollik haqidagi ma'lumotlarni o'z ichiga oladi. AurSCOPE ma'lumotlar bazalari (Aureus Pharma, Parij, Frantsiya) 38 000 dan ortiq nashrlarda yoritilgan, terapevtik ahamiyatga ega bo'lgan protein oilalari a'zolari bilan bog'liq 1 300 000 ga yaqin biologik faolliklarga izohlangan 320 000 dan ortiq molekulalarni o'z ichiga olgan kimyoviy kutubxonalar to'plamini taklif qiladi. Va nihoyat, MedChem va Target Inhibitor ma'lumotlar bazalari (GVK Biosciences, Haydarobod, Hindiston) 20 000 dan ortiq nashrlardan olingan terapevtik ahamiyatga ega protein oilalari uchun biologik faollik, toksiklik va farmakologik ma'lumotlarga ega 2 000 000 ga yaqin molekulalarni tuzadi. Bu katta izohli tashabbuslar oxir-oqibatda oxirgi foydalanuvchiga ilmiy nashrlarda chop etilgan ma'lumotlar asosida maqsadli oqsillarga kichik molekulalarni ulash imkonini beradi. Keyinchalik bu virtual yaqinlik profilini yaratish uchun ishlatilishi mumkin bo'lgan ligandga asoslangan protein modellarini yaratishda foydalanish mumkin (Shuffenhauer va Yakobi, 2004 ).

**Ma'lumotlarni vizualizatsiya qilish.** Hisoblash usullari har bir molekula tuzilishi uchun ko'plab turli xil farmakologik va fizik-kimyoviy xususiyatlar uchun bashoratlarni yaratish potentsialiga ega, bunday ma'lumotlarni tahlil qilish ma'lumotlarni qazib olish uchun ko'p o'lchovli usullar va, ehtimol, murakkab vizualizatsiya vositalariga ehtiyoj borligini ko'rsatadi ( Cheng va *boshq.*, 2002 ); Ekins va *boshqalar* , 2002 , 2006 ). Diva va Spotfire ( Ahlberg, 1999 ) kabi tijoratda mavjud vositalar ADME va fizik-kimyoviy xususiyat ma'lumotlarini tahlil qilish uchun keng qo'llanilgan ( Ekins va *boshq.*, 2002 , 2006 ; Stoner va *boshq.* , 2004.) yoki xususiy qarorlarni qo'llab-quvvatlash tizimlariga integratsiyalangan ( Rojnuckarin va *boshq.*, 2005 ), shu bilan birga, shunga o'xshash 3D grafik va filtrlash imkoniyatlari bilan yangi usullar ham mavjud ( Oellien va *boshq.*, 2005 ). 2D strukturaviy o'xshashlikka asoslangan aglomerativ ierarxik klasterlash, rekursiv bo'linish, Sammon xaritalari, o'z-o'zini tashkil etuvchi xaritalar va generativ topografik xaritalash kabi boshqa usullardan hisoblash bashoratlari bilan foydalanish mumkin ( Balakin va *boshqalar* , 2005 ; Kibbey va Calvet, 2005 ; Maniyar va *boshqalar...*, 2006 ; Yamashita va *boshqalar* , 2006 ).





*In silico* oqsili tuzilishini bashorat qilish va bioinformatika orqali molekulyar tomondan o'rganish, Protein tuzilishini bashorat qilish va bioinformatika orqali molekulyar modellarini yaratish tahlili va oqsil tuzilishini bashorat qilish tahlili bo'yicha Insilco yondashuvlarini ko'rib chiqamiz.

**Universal Protein Resurs (UniProt)** oqsil ma'lumotlari uchun markaziy markaz vazifasini bajaradi. Uniprotning maqsadi qo'shimcha ma'lumotlarga ega bo'lgan bir nechta ma'lumotlar bazalarida protein ma'lumotlarini qidirishni osonlashtirish uchun samarali kirish strategiyalari bilan barqaror va yuqori sifatli protein ma'lumotlar bazalarini yaratish va shakllantirishdir (Cathy va boshqalar, 2006).

**BLAST (lokal moslashtirish qidirishining asosiy vositasi).** BLAST evristik algoritmda ishlaydi. Bu tezkor qidiruvni amalga oshirish uchun ba'zi yorliqlarga bog'liq. BLAST mahalliy moslashtirishni amalga oshiradi. Turli xil ketma-ketliklarni taqqoslash uchun foydalanuvchi uchun BLAST tahlilining juda ko'p turlari mavjud, masalan, DNK ma'lumotlar bazasidan DNK so'rovi, oqsil ma'lumotlar bazasidan oqsil so'rovi va oqsil ketma-ketligi ma'lumotlar bazasidan barcha olti o'qish ramkalarida tarjima qilingan DNK so'rovi. (Tomas Madden, 2013).

**MODELLER 9.14 Dasturi.** MODELLER - bu homologik modellashtirish yondashuviga asoslangan kompyuter dasturi bo'lib, uchinchi darajali oqsil va to'rtlamchi tuzilmalarning gomologik modellarini yaratadi. Kirish shablonlar ketma-ketligi, shablonlarning atomik koordinatalari va oddiy skript fayli bilan modellashtiriladigan ketma-ketlikni tekislash faylini o'z ichiga oladi. MODELLER keyin avtomatik ravishda barcha vodorod atomlari bo'lmagan modelni hisoblab chiqadi. MODELLER skript tilidan foydalanadi va grafik interfeysga ega emas. U standart FORTRAN 90 da yozilgan va UNIX, Windows yoki Mac operatsion tizimlarida ishlashi mumkin (Narayanan va boshq., 2008).

**I-TASSER** veb-server avtomatlashtirilgan to'liq uzunlikdagi protein tuzilmalarini ishlab chiqaradi. I-TASSER Suite - bu yuqori aniqlikdagi oqsil tuzilishini bashorat qilish, takomillashtirish va strukturaga asoslangan funktsiya izohlari uchun ishlab chiqilgan mustaqil kompyuter dasturlari to'plami. Serverning chiqishi beshta modeldan iborat bo'lib, ishonchlilik reytingi, taxminiy TM-ballari va RMSD va taxminlarning standart og'ishi (Yang Zhang., 2008).

**M4T** (Bir nechta shablonli bir nechta xaritalash usuli) qiyosiy va to'liq avtomatlashtirilgan oqsil tuzilishini modellashtirish serveridir. U bir nechta shablonlarning kombinatsiyasidan foydalangan holda qiyosiy modellashtirish va muqobil tekislashlarni iterativ optimallashtirish imkonini beradi (Narcis va boshqalar, 2007).

**RaptorX** oqsil tuzilishini bashorat qilish serveri bo'lib, uchta komponentdan iborat (bitta shablonli ipni ajratish, tekislash sifatini bashorat qilish va bir nechta shablonli tishlash. RaptorX jarayonining kiritish ketma-ketligi uning ikkilamchi va uchinchi tuzilmalarini, shuningdek, erituvchiga kirish imkoniyati va tartibsiz hududlarni bashorat qilish orqali. ReptorX quyidagilardan iborat: bir nechta tishli komponent va hizalanish sifatini bashorat qilish uchun yangi modulga ega (Jian Peng va Jinbo Xu., 2011).

**IntFOLD** serveri berilgan ketma-ketlikdan strukturani bashorat qilish va funktsiya uchun turli zamonaviy usullarni birlashtiradi. Serverdan olingan natijalar

chizmalar va izohli 3D modellar orqali umumiy natijalarni grafik tarzda umumlashtiradigan oddiy jadval sifatida taqdim etiladi (Daniel va boshq., 2011).

**Phyre2** (Protein gomologiyasining analogiyasini aniqlash mexanizmi) oqsil tuzilishi va funksiyasini bashorat qilish va tahlil qilish uchun vositalar to'plamidir. Phyre2 ning maqsadi biologlarga oddiy va intuitiv interfeys bilan hissa qo'shishdir (Kelley va boshq., 2015).

**Errat** - kristallografiya yordamida aniqlangan oqsil tuzilmalarini tekshirish uchun bilimlar bazasi dasturi. Errat haqiqiy oqsillardan statistik ma'lumotlardan foydalanadi. Errat bog'lanmagan atom-atom (azot, uglerod va kislorod) hisobot tuzilmasidagi o'zaro ta'sir statistikasini sharhlaydi (Björn va Arne, 2006).

**Ramachandran** syujeti oqsil tuzilishidagi Ramachandran burchaklari deb ham ataladigan phi (P) va psi (P) burchaklarini ko'rsatadi. Ramachandran uchastkalari oqsil uch o'lchovli tuzilmalarining stereokimyoviy sifatini tahlil qilish uchun ishlatiladi. Syujet hududlari phi va psi qiymatlarining kombinatsiyasi (Oliviero va Kristina, 2013).

**Protparam Expasy** vositasi ma'lum bir protein uchun turli xil fizik va kimyoviy parametrlarni hisoblash imkonini beradi. Hisoblangan parametrlarga molekulyar og'irlik, nazariy Iso-elektr nuqtasi (pI), aminokislotalar tarkibi, atom tarkibi, so'nish koeffitsienti, taxminiy yarim yemirilish davri, beqarorlik indeksi, alifatik indeks va gidropatiklikning umumiy o'rtacha qiymati kiradi (John va boshqalar, 2005)..

**AutoDock** - bu avtomatik o'rnatish vositalari to'plami. AutoDock4 ham, AutoDock Vina ham kichikroq molekulalarni bir nechta aylanadigan bog'lanishlarga o'rnatishga qodir. AutoDock Vina tezroq ishlaydi va keng molekulalarni aniqroq tartiblash qobiliyatiga ega. Virtual skrining maqsadi uchun tadqiqotchilar birinchi navbatda unga qarashlari kerak (Maks va boshq., 2010).

**UCSF Chimera** - bu erkin mavjud bo'lgan va molekulyar tuzilmalarni interaktiv vizualizatsiya qilish va tahlil qilish uchun ishlatiladigan keng funktsiyalarga ega molekulyar grafik to'plami (Tomas va boshq., 2005). Chimera shuningdek, Internetga asoslangan xizmatlardan foydalangan holda ma'lum tuzilmalarni vizuallashtirishni qo'llab-quvvatlaydi (Erik va boshq., 2004).

**LIGPLOT** - bu ish stoliga asoslangan dastur bo'lib, u protein-ligand komplekslarining sxematik 2-D ko'rinishini avtomatik ravishda ishlab chiqadi. Chiqish - bu molekulalararo o'zaro ta'sirlarning informatsion tasvirini o'z ichiga olgan postscript fayli. Bundan tashqari, oqsil va nuklein kislotalarning o'zaro ta'sirini LIGPLOTdan ham ko'rish mumkin (Endryu va boshq., 1995).

**Chem Draw ultra** ligandlarni chizish va ligand molekulalarining energiyasini minimallashtirish uchun ishlatiladi. Ushbu dasturiy ta'minot qayta ishlash, saralash va tahrirlash bo'yicha kimyoviy tuzilish ma'lumotlariga ixtisoslashgan (Zhenjiang va boshq., 2004).

**GRAMM-X** - oqsil-oqsil o'rnatuvchi veb-server. U foydalanuvchilarga oddiy interfeysni taqdim etadi. GRAMM-X C++ va python tillarida amalga oshirilgan va u bepul mavjud (Andrey va Ilya, 2006).

**ClusPro** - protein-oqsillarni hisoblash uchun to'liq avtomatlashtirilgan va veb-ga asoslangan dastur. Foydalanuvchi kirish sifatida tegishli strukturaning PDB kodi bilan ikkita protein strukturasi PDB faylini yuklashi mumkin. Chiqarish - bu

klasterlash xususiyatlariga ko'ra tartiblangan komplekslar ro'yxati (Stephen va boshq., 2004).

**PyMOL** - bu ko'pgina grafik paketlarni taqdim etadigan molekulyar grafik tizim. PyMOL makromolekulyar tuzilmalar uchun keng tarqalgan ko'rinishlarning ko'pini qo'llab-quvvatlaydi, shu jumladan to'p va tayoq, nuqta sirlari, qattiq yuzalar, lentalar va multfilm lentalar (Uorren DeLano., 2002).

### **Uchunchi savolning bayoni.**

**In siliconing tibbiyotdagi axamiyati.** Ommaviy ma'lumotlar bazalarida mavjud bo'lgan nukleotidlar ketma-ketligi ma'lumotlarining tez kengayishi biotibbiyot tadqiqotlarida inqilob qildi. Nukleotidlar ketma-ketligi ma'lumotlar bazalarining o'sishi gen ekspressiyasining "virtual" yoki elektron profilini yaratishga olib keldi. Ushbu ko'rib chiqish uchun hisoblash usullarini qo'llash misollari asosan saraton tadqiqotlari loyihalarida aniqlanadi.

mo'ljallangan, ko'pincha kutilgan zarba darajasi 1% yoki undan kam bo'lib, keyingi sinovlardan so'ng yana kamroq haqiqiy natijalar kutilmoqda (dori kashfiyoti bo'limiga qarang ).

Misol tariqasida, ushbu texnika COVID-19 (SARS-CoV-2) uchun potentsial davolash usullarini izlash uchun dori vositalarini qayta ishlashni o'rganish uchun ishlatilgan.

### **Muhokama uchun savollar**

1. In silico yondashuvlarining tarixi va evolyutsiyasi nima?
2. Miqdoriy tuzilma-faoliyati munosabatlarini tushuntiring!
3. Deskriptorga asoslangan usullar nima?
4. Virtual ligand skriningi deganda nimani tushunasiz?
5. Ligandga asoslangan usullarni tushuntiring!
6. Virtual yaqinlik profilini yaratish nima?
7. in silicoda qo'llaniladigan qanday dasturlarni bilasiz?
8. To'qimalarga xos gen ifodasi deganda nimani tushunasiz?

## **8-mavzu: Neyron to'rlari**

### **Asosiy savollar**

1. Bioinformatikada Neyron to'rlari va rivojlanish tarixi.
2. Suniy intellekt.
3. Bioinformatikada Neyron to'rlarining ro'li.

**Mavzuga oid tayanch tushuncha va iboralar:** ANN, *GENESIS*, *NEXUS*, *MATLAB*, *SNN*, *Neyrologiya*, *LSTM*, *Sun'iy neyron*, *SEAForth40*, *parallel algoritm*

**Dars maqsadi:** Talabalarga Neyron to'rlarining (NT) yaralish tarixi va qo'llaniladigan usullari hamda bioinformatika va boshqa sohalardagi ahamiyati haqida tushuncha berish. Suniy intellektni haqida ma'lumot berish.

### **Birinchi savol bayoni**

Ma'lumki, organizmda juda ko'plab molekulyar biologik jarayonlar o'tadi. Proteoliz — peptidilglutamil fermentlari yoki proteazalar bilan katalizlanadigan oqsillarning gidroliz jarayoni. Proteoliz organizmdagi quyidagi jarayonlarda katta rol o'ynaydi: oshqozon va ingichka ichakda ovqat hazm qilish fermentlarining ta'siri tufayli oziq-ovqat oqsillarini aminokislotalarga parchalash; moddalar almashinuvi jarayonida organizmda oqsillarining sintezini amalga oshirish; fermentlarni, gormonlarni va biologik faol peptidlarni ularning faol bo'lmagan progormonlardan shakllantirish; o'simliklarda proteoliz unib chiqish vaqtida zaxira urug' oqsillarini safarbar qilishda ishtirok etadi. Proteolitik fermentlarning ta'sirini ikki toifaga bo'lish mumkin: cheklangan proteoliz, unda proteaza fermeti oqsil tarkibidagi bir yoki bir nechta peptid bog'larini parchalaydi, bu odatda funktsional holatning o'zgarishiga olib keladi: masalan, fermentlar faollashadi va progormonlar gormonlarga aylanadi; cheksiz, yoki umumiy proteoliz, unda oqsillar alohida aminokislotalarga bo'linadi. Substrat molekulariga ta'sir etishiga ko'ra proteolitik fermentlar endopeptidazalar va ekzopeptidazalarga bo'linadi: endopeptidazalar yoki proteinazalar peptid zanjiri ichidagi peptid bog'larini parchalaydi. Ular organizm va substratlar tarkibidagi peptid tarkibidagi aminokislotalarning ketma-ketligining qoldiqlari o'rtasidagi bog'larni gidroliz qilishda ishtirok etadi. Ekzopeptidazalar zanjirning uchidagi aminopeptidazalar-N- uchidan, karboksipeptidazalar-C-uchidan aminokislotalarni gidrolizlaydi. Proteazalar kataliz mexanizmi turiga ko'ra ham tasniflanadi. Xalqaro biokimyo va molekulyar biologiya ittifoqi bir necha proteaza sinflariga tasnif berishgan, shu jumladan:

- Serine proteases - serin proteazalar
- Aspartic proteases - aspartate proteazalar

- Cysteine proteases - sistein proteazalar
- Metalloproteaza-metallproteazalar

Chegaralangan proteoliz-oqsil molekulasidagi bir yoki bir necha peptid bog'larning proteaza fermenti yordamida ajralishi jarayoni tushsuniladi. Cheklangan proteoliz tartibga soluvchi posttranslational modifikatsiyalardan biridir. Cheklangan proteoliz fermentativ faoliyat bo'lib, boshqa oqsillarni bog'lash qobiliyati va hujayra ichidagi joylashgan oqsil xususiyatlarini o'zgartirishi mumkin.

*Cheklangan proteolizga misollar.* Cheklangan proteoliz turli maqsadlar uchun hujayra tomonidan ishlatilishi mumkin: hujayra ichidagi oqsillarni tashish paytida N- va C-terminal signal ketma-ketliklarini ajratish uchun; uchlamchi tuzilishining (proinsulin C-peptid) hosil bo'lishida yordam beradi, yani polipeptid zanjirning yordamchi qismini olib tashlash uchun; ferment faol markazlarini faollashtirish uchun- ovqat hazm qilish fermentlari, qon ivishi proteazalari; oqsil joylarining o'zgarishi- ba'zi sytoplasma oqsillari cheklangan proteolizni o'tgandan so'ng, yadroga o'tishi: SREBP, [NF-kB](#), YB-1, progormonlardan fiziologik faol oligopeptidlarning hosil bo'lishi- endorfin, adrenokortikotrop gormon,  $\alpha$ - va  $\gamma$ - melanotsit gormonlar va boshqa fiziologik faol peptidlar hosil qilish uchun, proopiomelanokortinning ajralishi; domenlar odatda bir-biri difteriyatoksini, hujayra proliferatsiya omili HCF-1 shakllantirish omillari bilan aloqada bo'ladi, oqsil domenlar o'rtasida bog'lar hosil bo'lishi; bir nechta oqsil izoformalarini olish uchun- STAT5 va STAT6 oqsillarining cheklangan proteolizi transkripsiyani faollashtiradigan domenlardan mahrum bo'lgan izoformalarining shakllanishiga olib keladi. Ushbu jarayonlarni kompyuter modellashtirishda SPLEN-K, COSMOS/M, ANSYS dasturlaridan foydalaniladi.

Sinaptik uzatish - neurotransmissiya deb ham ataladi — nerv impulslarining tarqalishi oqibatida sinapslarda elektr harakatlarini yuzaga keltiradi. Har bir nerv hujayrasi presinaptik neyron dan yoki postsinaptik neyron dan yoki ikkilamchi neyronning dendridlaridan neurotransmitterni oladi va jarayonni takrorlaydigan bir necha neyronlarga yuboradi, shuning uchun impuls to'liqini ma'lum bir organ yoki muayyan neyronlar guruhiga yetib borguncha tarqaladi.

Signallarning tarqalishi uchun nerv impulslari zarur. Ushbu signallar markaziy asab tizimidan efferent va afferent neyronlar orqali silliq, skelet va yurak mushaklari, bez sekretsiyasi va sut emizuvchilar kabi ko'p hujayrali umurtqali hayvonlarning yashashi uchun muhim organ vazifasini muvofiqlashtirish uchun yuboriladi.

Neyronlar nerv impulslari o'tadigan neyron tarmoqlarini hosil qiladi. Har bir neyron boshqa neyronlar bilan kamida 15,000 tarmoqlar hosil qiladi. Neyronlar bir-biriga tegmaydi; ular sinapslar deb ataladigan aloqa nuqtalarini hosil qiladi. Neyronlar axborotni nerv impulsi yordamida uzatadi. Neyronning impulsi sinapsga yetganda, u boshqa hujayralarga ta'sir qiluvchi mediatorlarni sintez qilishga va ingibitsiyasiga yoki qo'zg'alishga olib keladi. Keyingi neyron boshqa ko'plab neyronlar bilan bog'lanishi mumkin va agar hayajonlantiruvchi jarayonlar tushkunlikka tushib qolsa, akson bazasida harakat potentsiali ishlab chiqiladi va shu tariqa axborotni keyingi neyronga uzatib, xotira yoki harakatga olib keladi.

Sun'iy neyron (macculloch-Pitts matematik neyron, formal neyron) – tabiiy neyronning soddalashtirilgan modeli bo'lgan sun'iy neyron tarmog'ining tuguni hisoblanadi. Matematik jihatdan sun'iy neyron odatda bitta argumentning chiziqli bo'lmagan funksiyasi — barcha kirish signallarining chiziqli kombinatsiyasi sifatida namoyon bo'ladi. Bu funksiyaga aktivatsiya funksiyasi yoki aktuatsiya uzatish funksiyasi deyiladi.

Sun'iy neyronlar tarmoqlarda ulanadi-ba'zi neyronlarning chiqishlarini boshqalarning kirishlari bilan bog'laydi. Sun'iy neyronlar va tarmoqlar ideal neyrokompyuterning asosiy elementlari hisoblanadi.

*Biologik prototip.* Biologik neyron diametri 3 dan 100 mikrongacha bo'lgan, o'z ichiga yadro (yadro teshiklari ko'p bo'lgan) va boshqa organellalar, jumladan, faol ribosomalar, Golji apparatlari bilan yuqori darajada rivojlangan dagal endoplazmatik retikulum organoidlarini va molukulyar jarayonlarni o'z ichiga olgan organizmdan iborat. Jarayonlarning ikki turi mavjud. Akson bir tarmoqli bo'lib, odatda neyron tanasida qo'zg'alishni o'tkazish uchun moslashgan uzun nerv uchidir. Dendritlar odatda qisqa va yuqori shoxlangan bo'lib, ular neyronga ta'sir etuvchi qo'zg'aluvchan va ingibitor sinapslar hosil bo'lishining asosiy joyi bo'lib xizmat qiladi- turli neyronlar

akson va dendritlar uzunligining turlicha nisbatiga ega boladi. Neyron bir nechta dendritlarga va odatda faqat bitta aksonga ega bo'lishi mumkin. Bitta neyron 20,000 boshqa neyronlarga ulanishi mumkin. Inson kontekstida 80 milliard neyron mavjud.

*Rivojlanish tarixi.* Sun'iy neyronning birinchi matematik modelini W. McCulloch va W. Pitts tomonidan tarmoq modeli sifatida taklif etishdi. Mualliflar tarmoqning bunday elementlar bo'yicha sonli va mantiqiy operatsiyalarni bajara olishini ko'rsatishgan. Amalda, tarmoq Frank Rosenblatt tomonidan kompyuter dasturi yordamida 1958 - yilda oshirgan. Dastlab neyron faqat mantiqiy nol va mantiqiy birlik signallari bilan ishlashi mumkin edi, chunki u biologik prototip asosida qurilgan bo'lib, faqat ikki holatda — hayajonlangan yoki hayajonlanmagan holatlarni aks ettirgan. Neyron tarmoqlarining rivojlanishi shuni ko'rsatdiki, ularni qo'llash doirasini kengaytirish uchun neyron nafaqat binar, balki uzluksiz (analog) signallar bilan ham ishlay olishi zarur. Neyron modelining bu umumlashmasi Widrow va Hoff tomonidan tuzilgan bo'lib, u logistik egri chiziqdan neyron aktivlanish funksiyasi sifatida foydalanishni taklif qilgan.

### Sun'iy neyronlar orasidagi bog'lanishlar

---

Ba'zi neyronlarning chiqish signallari boshqalaming kirishiga keladigan bog'lanishlar ko'pincha biologik neyronlar orasidagi bog'lanishlarga o'xshash sinapslar deb ataladi. Har bir bog'lanish o'z og'irligi bilan ajralib turadi. Musbat og'irlikka ega bo'lgan bog'lanishlar qo'zg'aluvchan, manfiy og'irlikka ega bo'lgan bog'lanishlar esa ingibitor deyiladi. Neyron bitta chiqishi bor, ko'pincha biologik prototipga o'xshatib akson deb ataladi. Neyronning bitta chiqishidan boshqa neyronlarning har qanday kirishiga signal yuborish mumkin.

*Molekulyar jarayonlarni modellashtirishda AutoDock dasturining ishlatilishi.* Doking- dorivor vositalarni kompyuterda modellashtirish jarayonining asosiy va muhim bosqichlaridan biri bo'lib, asosiy vazifasi ligand molekulasini kompleksini biologik faol moddan va reseptor molekulasini bionishonning strukturasi modelini qurish hisoblanadi. Reseptor molekulasini oqsil molekulasini bo'lib, ligand molekulasini kichik hajmga ega bo'ladi. Kam hollarda oqsil-oqsil doking munosabatlari uchrab



turadi. Molekulyar doking bir yoki bir necha molekulalardan tuzilgan molekulalar aro kompleks strukturasi bashorat qiladi. Doking vaqtida reseptor va ligand molekulalar orasidagi konformasion harakatchanlikni hisobga olish zarur bo'ladi. AutoDock dasturi qo'yidagidalarda qo'llaniladi:

1. Rentgen kristallografiyasida
2. Dorivor vositalarga nomzod strukturalar dizaynida
3. Optimazasiyani o'tkazishda
4. Virtual skriningda (HTS)
5. Kombinator kutubxonalar dizaynida
6. Oqsil-oqsil o'zarota'sirlari
7. Kimyoviy tadqiqotlar mehanizmlari.

**Neyron tarmoq** - bu biologik neyronlar tarmog'i yoki sxemasi yoki zamonaviy ma'noda sun'iy neyronlar yoki tugunlardan tashkil topgan sun'iy neyron tarmoq. Shunday qilib, neyron tarmoq bu biologik neyronlardan tashkil topgan biologik neyron tarmoq yoki sun'iy intellektni hal qilish uchun sun'iy neyron tarmoqdir. Biologik neyronning ulanishlari tugunlar orasidagi og'irliklar sifatida sun'iy neyron tarmoqlarda modellashtirilgan. Ijobiy og'irlik qo'zg'atuvchi aloqani aks ettiradi, salbiy qiymatlar esa inhibitiv aloqalarni bildiradi. Barcha kirishlar og'irlik bilan o'zgartiriladi va yig'iladi. Bu faoliyat chiziqli birikma deb ataladi. Nihoyat, faollashtirish funksiyasi chiqish amplitudasini boshqaradi. Masalan, qabul qilinadigan chiqish diapazoni odatda 0 dan 1 gacha yoki -1 va 1 bo'lishi mumkin. Ushbu sun'iy tarmoqlar prognozli modellashtirish, moslashuvchan boshqaruv va ma'lumotlar to'plami orqali o'qitilishi mumkin bo'lgan ilovalar uchun ishlatilishi mumkin. Tajriba natijasida kelib chiqadigan o'z-o'zini o'rganish murakkab va bir-biriga bog'liq bo'lmagan ma'lumotlar to'plamidan xulosa chiqarishi mumkin bo'lgan tarmoqlar ichida sodir bo'lishi mumkin. Biologik neyron tarmoq kimyoviy bog'langan yoki funksional jihatdan bog'langan neyronlar guruhidan iborat. Bitta neyron ko'plab boshqa neyronlarga ulanishi mumkin va tarmoqdagi neyronlar va ulanishlarning umumiy soni keng bo'lishi mumkin. Sinapslar deb ataladigan ulanishlar odatda aksonlardan dendritlarga hosil bo'ladi, ammo dendrodendritik sinapslar va boshqa ulanishlar mumkin. Elektr signalizatsiyasidan tashqari, neurotransmitter diffuziyasidan kelib chiqadigan signalizatsiyaning boshqa shakllari ham mavjud.

Neyron tarmoqlari yoki ko'pincha oddiy neyron tarmoqlari deb atalib, rivojlanishining dastlabki bosqichida inson miyasining ishlashidan ilhomlangan

chiziqli bo'lmagan hisoblash usullari oilasini o'z ichiga oladi. Darhaqiqat, birinchi NT lari inson markaziy asab tizimida nerv stimullari va signallarining uzatilishini tushunish uchun yaratilgan integral mikrosxemalar sifatida yaratilgan. Keyinchalik, hisoblash NT lari asta-sekin an'anaviy matematik va statistik usullar uchun qiyin yoki ba'zi hollarda imkonsiz deb hisoblangan vazifalarni bajarish yoki muammolarni hal qilishga qodir bo'lgan vositalar sifatida paydo bo'ldi. SHuning uchun so'nggi bir necha yil ichida NT ikkita asosiy yo'nalish bo'ylab rivojlandi. Biz ko'proq neyrofiziologik yo'naltirilardan biri deb atalgan rivojlanishga qaratilgan *in silicoda rivojlanishning* barcha xatti-harakatlar mexanizmlarini chuqurroq tushunish uchun imkon qadar aniq bo'lgan inson miyasining modellari tushunishimiz lozim. Ikkinchisi esa NT ni murakkab muammolarni hal qilish uchun hisoblash vositasi sifatida ishlab chiqilgan va ular odatda murakkab chiziqli bo'lmagan. Ushbu ikkinchi doiradagi qiziqish inson miyasining eng yaxshi modelini olishga emas, balki insonning ba'zi vazifalarni har qanday kompyuterga qaraganda yaxshiroq va juda qisqa vaqt ichida keng sxematik algoritmik shaklda bajarishiga yordam beradigan biologik xususiyatlarni aniqlash va qo'lga kiritish va ularni amalga oshirishga qaratilgan. Ushbu hisoblash NT lari biologlar uchun sezilarli darajada uzoq, ammo ular ekonometrika, ob-havo prognozi, signalni filtrlash va naqshni aniqlash kabi keng sohalarda aniq modellar va bashoratlarni taqdim etishga qodir. Haqiqatdan ham hozirgi kunda bashorat qilish, tasniflash yoki nazorat qilish muammolari mavjud bo'lgan har qanday vaziyatda NN lari joriy etilmoqda. Bu katta muvaffaqiyatlar bir necha asosiy omillar bilan bog'liq bo'lishi mumkin. Birinchidan, NT - bu juda murakkab funktsiyalarni modellashtirishga qodir bo'lgan juda murakkab chiziqli bo'lmagan hisoblash vositalari. Xususan, kirishlar to'plami va mos keladigan chiqishlar to'plami o'rtasidagi har qanday berilgan funktsional bog'liqlik tasodifiy tanlangan NT arxitekturasi bilan ifodalanishi mumkin. Shu doirada, NT ham nazorat qiladigan kirishlar to'plami va mos keladigan chiqishlar to'plami o'rtasidagi har qanday berilgan funktsional aloqani o'z vaqtida tanlangan NT arxitekturasi bilan ifodalash mumkin. Shu doirada, NT ham ko'p sonli o'zgaruvchilar bilan chiziqli bo'lmagan funktsiyalarni modellashtirishga harakat qiladigan o'lchovlilik muammosini nazorat qiladi. Bundan tashqari, NN larini misollar orqali o'rganadilar: ma'lumotlar strukturasi mos ravishda ishlab chiqilgan o'quv algoritmlari yordamida vakillik ma'lumotlaridan avtomatik ravishda o'rganiladi. Garchi foydalanuvchi ma'lumotlarni tanlash va tayyorlash, tegishli NNni qanday tanlash va natijalarni qanday talqin qilish bo'yicha ba'zi evristik ma'lumotlarga ega bo'lishi kerak bo'lsa-da, NNni muvaffaqiyatli qo'llash uchun foydalanuvchi bilimlari darajasi avvalgisidan ancha past. masalan, an'anaviy nochiziqli statistik usullardan foydalanish.

NEXUS (Network Experiments Use Simulation) Pol Sajda va Leif Finkel tomonidan ishlab chiqilgan bo'lib, o'zaro bog'liq bo'lgan hisoblash xaritalarini yaratish

va simulyatsiya qilish uchun mo'ljallangan. NEXUS neyron birliklari xaritalari orasidagi murakkab ulanish naqshlarini tavsiflash uchun tilni taqdim etadi. Neyron birliklari murakkablashtirish uchun dasturlashtirilishi mumkin bo'lsa-da, birliklar odatda integratsiya va olov yoki sigmasimon faollashtiruvchi neyronlar sifatida mavhumlanadi. Yagona neyronlar va kichik tarmoqlarning biofizik xususiyatlarini modellashtirishga qiziqqan tadqiqotchilar GENESISdan foydalanishni ko'rib chiqishlari kerak. NEXUS yordamida ishlab chiqilgan ko'plab modellar vizual korteksqa qaratilgan, chunki u retinotopik ulanishni yaratish uchun yaxshi mo'ljallangan.

Neyron tarmog'i asboblari to'plami turli xil oldindan belgilangan neyron tarmoq arxitekturalarini qo'llab-quvvatlaydi, jumladan MLP, o'z-o'zini tashkil etuvchi tarmoqlar va takroriy tarmoqlar. Transfer funktsiyalarining ko'p sinflari qo'llab-quvvatlanadi va turli xil o'rganish qoidalari amalga oshirilishi mumkin.

Sun'iy intellekt, kognitiv modellashtirish va neyron tarmoqlar biologik neyron tizimlari ma'lumotlarni qayta ishlash usulidan ilhomlangan axborotni qayta ishlash paradigmatidir. Sun'iy intellekt va kognitiv modellashtirish biologik neyron tarmoqlarning ba'zi xususiyatlarini taqlid qilishga harakat qiladi. Sun'iy intellekt sohasida dasturiy ta'minot agentlari (kompyuter va video o'yinlarda) yoki avtonom robotlarni yaratish uchun sun'iy neyron tarmoqlar nutqni aniqlash, tasvirni tahlil qilish va moslashuvchan boshqarish uchun muvaffaqiyatli qo'llanildi.

Tarixiy jihatdan raqamli kompyuterlar fon Neyman modelidan kelib chiqqan va bir qator protsessorlar tomonidan xotiraga kirish orqali aniq ko'rsatmalarni bajarish orqali ishlaydi. Boshqa tomondan, neyron tarmoqlarning kelib chiqishi biologik tizimlarda axborotni qayta ishlashni modellashtirish harakatlariga asoslanadi. Fon Neyman modelidan farqli o'laroq, neyron tarmoqni hisoblash xotira va ishlov berishni ajratmaydi.

Zamonaviy neyron tarmoqlar uchun dastlabki nazariy asos mustaqil ravishda Aleksandr Beyn (1873) va Uilyam Jeyms (1890) tomonidan taklif qilingan. Ularning ishlarida fikrlar va tana faoliyati miya ichidagi neyronlarning o'zaro ta'siri natijasida yuzaga kelgan. Neyron tarmoqlarning tarixi ko'pchilik o'ylagandan uzoqroq. "O'yaydigan mashina" g'oyasi qadimgi yunonlarda kuzatilishi mumkin bo'lsa-da, biz neyron tarmoqlar atrofidagi fikrlash evolyutsiyasiga olib kelgan asosiy voqealarga to'xtalib o'tamiz, bu yillar davomida mashhurligi pasaygan va oqib kelgan:

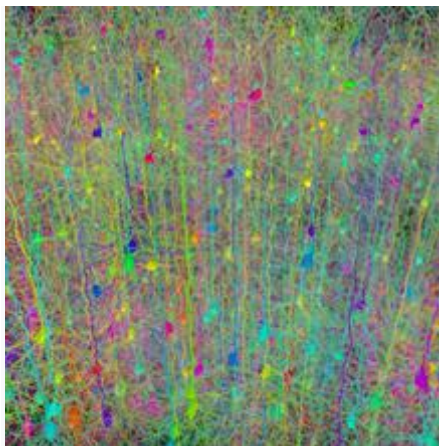
**1943 yil:** Uorren S. MakKullokk va Uolter Pitts "Asab faoliyatida immanent bo'lgan g'oyalarning mantiqiy hisobi (PDF, 1 MB) (havola IBM'dan tashqarida joylashgan)" nomli tadqiqot inson miyasi bog'langan miya orqali qanday qilib murakkab naqshlarni ishlab chiqishi mumkinligini tushunishga harakat qildi. hujayralar yoki neyronlar. Ushbu ishdan chiqqan asosiy g'oyalardan biri ikkilik

chegaraga ega neyronlarni Boolean mantiqqa (ya'ni, 0/1 yoki to'g'ri/noto'g'ri bayonotlar) solishtirish edi.

**1958 yil: Frenk Rozenblatt o'zining " Perceptron: Axborotni saqlash va miyada tashkil etishning ehtimollik modelini tadqiqotlarida hujjatlashtirilgan perseptronni ishlab chiqishda ishtirok etgan. U tenglamaga og'irliklarni kiritish orqali Makkaloch va Pittning ishini bir qadam oldinga olib boradi. IBM 704 dan foydalangan holda, Rosenblatt chapda belgilangan kartalarni o'ngda belgilangan kartalarni qanday ajratishni o'rganish uchun kompyuter olishga muvaffaq bo'ldi.**

**1974 yil: Ko'plab tadqiqotchilar orqaga tarqalish g'oyasiga hissa qo'shgan bo'lsa-da, Pol Verbos AQShda uning doktorlik dissertatsiyasida neyron tarmoqlarda qo'llanilishini qayd etgan birinchi odam bo'ldi (PDF, 8.1 MB) (havola IBM'dan tashqarida joylashgan).**

**1989 yil: Yann LeKun maqola chop etdi, unda orqaga tarqalishda cheklovlardan foydalanish va uning neyron tarmog'i arxitekturasiga integratsiyalashuvi algoritmlarni o'rgatish uchun qanday ishlatilishi mumkin. Ushbu tadqiqot AQSh pochta xizmati tomonidan taqdim etilgan qo'lda yozilgan pochta indeksi raqamlarini tanib olish uchun neyron tarmoqdan muvaffaqiyatli foydalandi.**



Piramidal neyronlarning dendritlari dallanma arxitekturasini kompyuter simulyatsiyasi.

Jeyms nazariyasi Beynnikiga o'xshardi, ammo u xotiralar va harakatlar miyadagi neyronlar orasidan oqib o'tadigan elektr toklari natijasida paydo bo'lishini taklif qildi. Uning modeli, elektr toklari oqimiga e'tibor qaratib, har bir xotira yoki harakat uchun individual neyron ulanishlarni talab qilmadi.

CS Sherrington (1898) Jeyms nazariyasini sinab ko'rish uchun tajribalar o'tkazdi. U elektr tokini kalamushlarning umurtqa pog'onasiga o'tkazdi. Biroq, Jeyms tomonidan prognoz qilingan elektr tokining o'sishini namoyish qilish o'rniga, Sherrington vaqt o'tishi bilan sinov davom etar ekan, elektr tokining kuchi pasayganligini aniqladi. Muhimi, bu ish odatlanish kontseptsiyasining ochilishiga olib keldi.

McCulloch va Pitts (1943) matematika va algoritmlarga asoslangan neyron tarmoqlar uchun hisoblash modelini yaratdilar. Ular bu modelni chegara mantiqi deb atashdi. Model neyron tarmog'ini tadqiq qilishning ikkita alohida yondashuvga bo'linishiga yo'l ochdi. Bir yondashuv miyadagi biologik jarayonlarga, ikkinchisi esa neyron tarmoqlarni sun'iy intellektga qo'llashga qaratilgan.

1940-yillarning oxirida psixolog Donald Xebb neyron plastisiya mexanizmiga asoslangan ta'lim gipotezasini yaratdi, u hozirda Hebbian o'rganish deb nomlanadi. Hebbian o'rganish "odatiy" nazoratsiz o'rganish qoidasi hisoblanadi va uning keyingi variantlari uzoq muddatli potentsialning dastlabki modellari edi. Ushbu g'oyalar 1948 yilda Turingning B tipidagi mashinalari bilan hisoblash modellarida qo'llanila boshlandi.

Farli va Klark (1954) MITda Hebbian tarmog'ini simulyatsiya qilish uchun dastlab hisoblash mashinalaridan foydalanganlar, keyinchalik esa u kalkulyatorlar deb ataldi. Boshqa neyron tarmoq hisoblash mashinalari Rochester, Holland, Habit va Duda (1956) tomonidan yaratilgan.

Rosenblatt (1958) oddiy qo'shish va ayirish yordamida ikki qavatli o'rganish kompyuter tarmog'iga asoslangan naqshni aniqlash algoritmi bo'lgan perceptronni yaratdi. Matematik belgilar bilan Rosenblatt asosiy perseptronda bo'lmagan sxemani ham tasvirlab berdi, masalan, eksklyuziv yoki sxema, Verbos [13] (1975) tomonidan teskari tarqalish algoritmi yaratilgunga qadar matematik hisoblashni qayta ishlash mumkin bo'lmagan sxema.

Dartmut loyihasidan keyingi yillarda Jon fon Neyman telegraf rölesi yoki vakuum naychalari yordamida oddiy neyron funktsiyalariga taqlid qilishni taklif qildi. Shuningdek, Kornelning neyrobiologi Frenk Rozenblatt Perceptron ustida ish boshladi. U pashsha ko'zining ishlashi bilan qiziqdi. Pashshaga qochishni bildiruvchi ko'p ishlov berish uning ko'zida amalga oshiriladi. Ushbu tadqiqot natijasida paydo bo'lgan Perceptron apparatda qurilgan va bugungi kunda ham foydalaniladigan eng qadimgi neyron tarmoq hisoblanadi. Bir qatlamli perseptron doimiy qiymatli kirishlar to'plamini ikkita sinfdan biriga tasniflashda foydali ekanligi aniqlandi. Perseptron kiritilgan ma'lumotlarning vaznli yig'indisini hisoblab chiqadi, chegarani olib tashlaydi va natijada ikkita mumkin bo'lgan qiymatdan birini chiqaradi. Afsuski, idrok etish qobiliyati cheklangan va u "ko'ngil aynish yillari" davomida isbotlangan. *Perseptronlar*.

1982 yilda bir nechta voqealar yangi qiziqish uyg'otdi. Kaltekdan Jon Xopfield milliy Fanlar akademiyasiga maqola taqdim etdi. Xopfieldning yondashuvi shunchaki miyalarni modellashtirish emas, balki foydali qurilmalar yaratish edi. Aniqlik va matematik tahlil bilan u bunday tarmoqlar qanday ishlashi va ular nima qilishi mumkinligini ko'rsatdi. Shunga qaramay, Xopfieldning eng katta boyligi uning xarizmasi edi. U so'zli, yoqimli va harakatsiz texnologiyaning chempioni edi.

Shu bilan birga, yana bir voqea sodir bo'ldi. Yaponiyaning Kioto shahrida konferentsiya bo'lib o'tdi. Ushbu konferentsiya hamkorlikdagi/raqobatbardosh neyron tarmoqlar bo'yicha AQSh-Yaponiya qo'shma konferentsiyasi edi. Yaponiya keyinchalik Beshinchi avlod harakatlarini e'lon qildi. AQSh davriy nashrlari bu voqeani ko'rib chiqdi va AQSh ortda qolishi mumkinligi haqida xavotir uyg'otdi. Ko'p o'tmay, yana moliyalash boshlandi.

1989 yilga kelib Mudofaa uchun neyron tarmoqlari yig'ilishida Bernard Vidrou o'z tinglovchilariga ular IV Jahon urushida qatnashganliklarini aytdi, "Uchinchi Jahon urushi hech qachon sodir bo'lmagan", bu yerda urush maydonlari jahon savdosi va ishlab chiqarishdir. 1990 yildagi AQSh Mudofaa vazirligining Kichik biznes innovatsion tadqiqot dasturi neyron tarmoqlarga qaratilgan 16 ta mavzuni, qo'shimcha 13 tasi esa neyron tarmoqlardan foydalanish mumkinligini eslatib o'tgan.

Bugungi kunda neyron tarmoqlar bo'yicha munozaralar hamma joyda sodir bo'lmoqda. Ularning va'dalari juda yorqin ko'rinadi, chunki tabiatning o'zi bu turdagi narsalarning ishlashini isbotlaydi. Shunga qaramay, uning kelajagi, haqiqatan ham butun texnologiyaning kaliti apparatni ishlab chiqishda yotadi. Hozirgi vaqtda ko'pchilik neyron tarmoqlarni ishlab chiqish printsiplal ishlayotganini isbotlamoqda. Ushbu tadqiqot neyron tarmoqlarni ishlab chiqmoqda, ularni qayta ishlash cheklovlari tufayli o'rganish haftalar davom etadi. Ushbu prototiplarni laboratoriyadan olib chiqib, foydalanishga topshirish uchun maxsus chiplar kerak bo'ladi. Kompaniyalar uchta turdagi neyrochiplar ustida ishlamoqda - raqamli, analog va optik. Ba'zi kompaniyalar neyron tarmoq Application Specific Integrated Circuit (ASIC) yaratish uchun "kremniy kompilyator" yaratish ustida ishlamoqda. Ushbu ASIC va neyronga o'xshash raqamli chiplar yaqin kelajak to'lqini bo'lib ko'rinadi. Oxir oqibat, optik chiplar juda istiqbolli ko'rinadi. Shunga qaramay, optik chiplar tijorat ilovalarida kun yorug'ligini ko'rishi uchun yillar o'tishi mumkin.

### **Ikkinchi savol bayoni**

**Suniy intellekt.** Sun'iy intellektda qo'llaniladigan neyron tarmoqlar an'anaviy ravishda miyada neyronni qayta ishlashning soddalashtirilgan modellari sifatida ko'rib chiqiladi, garchi bu model va miya biologik arxitekturasi o'rtasidagi bog'liqlik muhokama qilinsa ham, sun'iy neyron tarmoqlar funktsiyasi miyani qay darajada aks ettirishi aniq emas.

Neyron *tarmog'i* (NT), sun'iy neyronlar uchun *sun'iy neyron tarmog'i* (ANN) yoki *simulyatsiya qilingan neyron tarmog'i* (SNN) deb ataladigan bo'lsa, u ma'lumotlarni qayta ishlash uchun matematik yoki hisoblash modelidan foydalanadigan tabiiy yoki sun'iy neyronlarning o'zaro bog'langan guruhidir. hisoblashda konneksionistik yondashuv. Aksariyat hollarda NT adaptiv

tizim bo'lib, tarmoq orqali o'tadigan tashqi yoki ichki ma'lumotlarga asoslangan holda o'z tuzilishini o'zgartiradi.

Neyron tarmoqlarda o'rganish, ayniqsa, ma'lumotlar yoki vazifaning murakkabligi bunday funktsiyalarni qo'lda loyihalashni amaliy bo'lmagan ilovalarda foydalidir. Neyron tarmoqlar turli sohalarida qo'llanilishi mumkin. Sun'iy neyron tarmoqlari qo'llaniladigan vazifalar odatda quyidagi keng toifalarga kiradi:

- Funktsiyani yaqinlashtirish yoki regressiya tahlili , shu jumladan vaqt seriyalarini bashorat qilish va modellashtirish.
- Tasniflash , shu jumladan naqsh va ketma-ketlikni aniqlash, yangilikni aniqlash va ketma-ket qaror qabul qilish.
- Ma'lumotlarni qayta ishlash , jumladan, filtrlash, klasterlash, ko'r signallarni ajratish va siqish.

NTni qo'llash sohalariga chiziqli bo'lmagan tizimni identifikatsiyalash va boshqarish (avtomobilni boshqarish, jarayonni boshqarish), o'yin o'ynash va qaror qabul qilish (nard, shaxmat, poyga), naqshni aniqlash (radar tizimlari, yuzni identifikatsiyalash , ob'ektni tanib olish), ketma-ketlikni aniqlash kiradi. (imo-ishora, nutq, qo'lda yozilgan matnni aniqlash ), tibbiy diagnostika, moliyaviy ilovalar, ma'lumotlarni qidirish (yoki ma'lumotlar bazalarida bilimlarni topish, "KDD"), vizualizatsiya va elektron pochta spamlarini filtrlash. Masalan, ob'ektni tanib olishga o'rgatilgan rasmlardan foydalanuvchi qiziqishlarining semantik profilini yaratish mumkin. **Neyrologiya.** Nazariy va hisoblash nevrologiyasi biologik asab tizimlarini tahlil qilish va hisoblash modellashtirish bilan shug'ullanadigan sohadir. Neyron tizimlar kognitiv jarayonlar va xatti-harakatlar bilan chambarchas bog'liq bo'lganligi sababli, bu soha kognitiv va xatti-harakatni modellashtirish bilan chambarchas bog'liq. Ushbu sohaning maqsadi biologik tizimlar qanday ishlashini tushunish uchun biologik asab tizimlarining modellarini yaratishdir. Ushbu tushunchaga ega bo'lish uchun nevrologlar kuzatilgan biologik jarayonlar (ma'lumotlar), neyronni qayta ishlash va o'rganish uchun biologik asosli mexanizmlar ( biologik neyron tarmoq modellari) va nazariya (statistik o'rganish nazariyasi va axborot nazariyasi ) o'rtasida aloqa o'rnatishga intiladi. Modellar turlari.Ko'pgina modellar qo'llaniladi; mavhumlikning turli darajalarida aniqlangan va asab tizimining turli tomonlarini modellashtirish. Ular individual neyronlarning qisqa muddatli xatti-harakatlari modellaridan tortib, individual neyronlarning o'zaro ta'siridan kelib chiqadigan neyron zanjiri dinamikasi modellari orqali to'liq quyi tizimlarni ifodalovchi mavhum neyron modullaridan kelib chiqadigan xatti-harakatlar modellarigacha. Bularga nerv sistemalarining uzoq muddatli va qisqa muddatli plastisitivligi hamda uning individual neyrondan tizim darajasigacha bo'lgan o'rganish va xotiraga aloqadorligi modellari kiradi.**Ulanish.** 2020-yil avgust oyida olimlar ikki yo'nalishli ulanishlar yoki qo'shilgan tegishli teskari aloqa aloqalari miya yarim



korteksining modulli neyron tarmoqlari orasidagi va ulardagi aloqani tezlashtirishi va yaxshilashi hamda ularning muvaffaqiyatli aloqa chegarasini pasaytirishi haqida xabar berishdi. Ular rezonans juftligi o'rtasida qayta aloqa ulanishlarini qo'shish butun tarmoq bo'ylab bitta impuls paketining muvaffaqiyatli tarqalishini qo'llab-quvvatlashi mumkinligini ko'rsatdi. Neyron tarmoqlarning, xususan, robototexnika sohasida keng tarqalgan tanqidi shundaki, ular haqiqiy hayotda ishlash uchun turli xil o'quv namunalarini talab qiladi. Buning ajablanarli joyi yo'q, chunki har qanday o'quv mashinasi yangi holatlarni umumlashtirishga imkon beradigan asosiy tuzilmani qo'lga kiritish uchun etarli vakillik misollariga muhtoj. Din Pomerleau "Avtonom robotlarni boshqarish uchun sun'iy neyron tarmoqlarni bilimga asoslangan o'qitish" maqolasida taqdim etilgan tadqiqotida robotlashtirilgan transport vositasini bir nechta turdagi yo'llarda (bir qatorli, ko'p polosali, tuproqli) haydashga o'rgatish uchun neyron tarmoqdan foydalanadi va boshqalar. Uning tadqiqotining katta qismi (1) bitta mashg'ulot tajribasidan bir nechta o'quv stsenariylarini ekstrapolyatsiya qilish va (2) tizimni haddan tashqari oshirib yubormaslik uchun oldingi mashg'ulotlarning xilma-xilligini saqlashga bag'ishlangan (masalan, agar u o'ngga bir qator burilishlar bilan taqdim etiladi - u doimo o'ngga burilishni o'rganmasligi kerak). Bu muammolar neyron tarmoqlarda keng tarqalgan bo'lib, ular turli xil javoblar orasidan qaror qabul qilishi kerak, lekin ularni bir necha usullar bilan hal qilish mumkin, masalan, o'quv misollarini tasodifiy aralashtirish, raqamli optimallashtirish algoritmidan foydalangan holda, juda katta qadamlar qo'ymaydi. misol bo'yicha tarmoq ulanishlarini o'zgartirish yoki misollarni mini-to'plamlarda guruhlash orqali. Neyron tarmoqlar ko'pincha *samarali* dasturlarni keltirib chiqaradigan bo'lsa-da, ular ko'pincha *samaradorlik* xarajati bilan amalga oshiradilar (ular ko'p vaqt va pul sarflashga moyil). Dyudni pozitsiyasiga qarshi argumentlar shundan iboratki, neyron tarmoqlar ko'plab murakkab va xilma-xil vazifalarni, masalan, avtonom uchadigan samolyotlarni hal qilish uchun muvaffaqiyatli ishlatilgan. Texnologiya yozuvchisi Rojer Bridgman Dyudnining neyron tarmoqlar haqidagi so'zlarini quyidagicha izohladi: Masalan, neyron tarmoqlar nafaqat baland osmonga ko'tarilganligi uchun (nima yo'q?), balki siz uning qanday ishlashini tushunmasdan muvaffaqiyatli tarmoq yaratishingiz mumkinligi uchun ham dokda: uni ushlaydigan raqamlar to'plami. xatti-harakati, ehtimol, "shaffof, o'qib bo'lmaydigan jadval... ilmiy manba sifatida qadsiz" bo'ladi. O'zining ilm-fan texnologiya emasligini qat'iy ta'kidlaganiga qaramay, Dyudni bu erda neyron tarmoqlarni yomon ilm deb baholaganga o'xshaydi, chunki ularni yaratuvchilarning aksariyati shunchaki yaxshi muhandis bo'lishga harakat qilishadi. Foydali mashina o'qishi mumkin bo'lgan o'qilmaydigan jadvalga ega bo'lishga arziydi. Sun'iy neyron tarmog'i tomonidan o'rganilgan narsalarni tahlil qilish qiyin ekanligi haqiqat bo'lsa-da, biologik neyron tarmoq tomonidan o'rganilgan narsalarni tahlil qilishdan ko'ra buni qilish ancha

oson. Bundan tashqari, so'nggi paytlarda AIning tushuntirilishiga urg'u berilgan neyron tarmoqlarni vizualizatsiya qilish va tushuntirish uchun usullarni, xususan diqqat mexanizmlariga asoslangan usullarni ishlab chiqishga yordam berdi. Bundan tashqari, neyron tarmoqlar uchun o'rganish algoritmlarini o'rganish bilan shug'ullanadigan tadqiqotchilar asta-sekin o'rganish mashinasining muvaffaqiyatli bo'lishiga imkon beruvchi umumiy tamoyillarni ochib berishadi. Masalan, Bengio va LeCun (2007) mahalliy va mahalliy bo'lmagan o'rganish, shuningdek, sayoz va chuqur arxitektura haqida maqola yozgan. Ba'zi boshqa tanqidlar gibrid modellarga (neyron tarmoqlari va ramziy yondashuvlarni birlashtirgan) ishonuvchilardan keldi. Ular ushbu ikki yondashuvning aralashuvini yoqlaydilar va gibrid modellar inson ongining mexanizmlarini yaxshiroq qamrab olishiga ishonishadi (Sun va Bookman, 1990).

**Bioinformatikada Neyron to'rlarining ro'li.** Hozirgi vaqtda sun'iy neyron to'rlari va masalalarni parallel ishlash ustida nazariy izlanishlar va amaliy qo'llanishlar keskin rivojlanmoqda. Neyron to'rlar analitik tavsifi bo'lmagan va faqatgina eksperimental ma'lumotlar bilan berilgan katta ko'lamdagi amaliy masalalarni yechish imkonini beradi.

Neyron to'rlarini sintez qilishda algoritmlarning nozik tomoni bu qaror qabul qilishni tushuntirish bo'lib hisoblanadi. Bu muammoni yechish bilan ko'pchilik tadqiqotchilar shug'ullanmoqdalar. Bu maqsadda ishlatadigan usullar evristik bo'lganligi uchun ular asosida korrekt qaror qabul qilish foydalanuvchining subektiv mulohazasiga bog'liq bo'ladi.

Ko'p o'lchovli chiziqsiz optimizasiyaning an'anaviy iterativ gradiyent algoritmlari bilan o'rganadigan neyron to'rlari modellarining eng ko'p tarqalgani - bu ko'p qatlamli sun'iy neyron to'rlari sinfidir. Ma'lumki, ko'p qatlamli sun'iy neyron to'rlari o'rganishda iterativ algoritmlar yaqinlashuvi, o'rganiladigan berilganlarning (tanlovning) hajmiga, vaznlarning boshlang'ich qiymatiga, shuningdek, o'rganishdagi maksimal xatolarga (o'rganishning sifat mezonlariga), o'rganishdagi takrorlanishlar soniga (o'rganish vaqtining uzayishi mezonlariga) bog'liq.

Shuning uchun, qo'yilgan masalani yechish uchun optimal modellarni tanlashda ularni solishtirish va qaror qabul qilishda neyroto'rlarning xususiyatlarini yetarli darajada baholashga imkon beruvchi xususiy va umumiy mezonlar majmuasini ishlab chiqish zarur.

Hozirda keng tarqalgan xatolarning teskari tarqalish algoritmlarida va Xopfield neyron to'rlarida qaror qabul qilish jarayonini tushuntirishga harakatlar qilindi. Bu modellardagi algoritmlarning evristik xarakterda ekanligi qaror qabul qilishda neyron

to'rlarining shaffoflik muammosini yechishni yetarli darajada matematik formallashtirishga imkon bermaydi. Natijada, tasvirlarni ajratib olish neyron to'rlari bo'yicha mutaxassisga bog'liq va asosan tavsifiya xususiyatiga ega bo'ladi.

### **Bir qatlamli sun'iy neyron to'ri**

**Sun'iy neyron.** Sun'iy neyron birinchi yaqinlashishda biologik neyron xossalarini imitatsiya qiladi. Har bir sun'iy neyronga boshqa neyronlar chiqishi bo'lgan qandaydir signallar to'plami kiradi. Har bir kiruvchi signal sinaptik kuchga mos vaznga ko'paytiriladi va ularning yig'indisi neyronning aktivlik darajasini aniqlaydi. Har bir signal o'ziga mos keluvchi  $w_1, w_2, \dots, w_n$  vaznlarga ko'paytiriladi va  $\Sigma$  bilan belgilangan yig'uvchi blokka kelib tushadi. Har bir vazn bitta biologik sinapsis «kuchiga» mos keladi. Vaznlar to'plami  $W$  vektori orqali belgilanadi. Biologik yelemnt tanasiga mos keluvchi yig'uvchi blok, mos vaznlariga ko'paytirilgan kiruvchi qiymatlarni algebraik tarzda yig'adi va neyron chiqishini shakllantiradi.

Shu sababli teskari bog'lanishli to'rlar inson miyasining qisqa muddatli xotirasi xossalariga o'xshash xossalarga yega bo'ladi. To'r chiqishlari qisman oldingi kirishlarga bog'liq bo'ladi.

### **Suniy neyron to'rlarini o'rgatishda parallel algorimlardan foydalanish**

Ko'plab hisoblash ishlari katta miqdordagi amallarni bajarishni talab etadi va zamonaviy mashinalar ham bu masalalarni hal etishda ko'p vaqt sarflaydi. Bir qancha masalalarni hal etishda qat'iy belgilangan vaqt talab etiladi, aks holda masala o'z dolzarbligini yo'qotadi. Ko'plab masalalarni hal etishda neyron tarmoqlari mexanizmi qo'llaniladi, ko'rsatilgan parametrlarni tanlash me'zoni sifatida qo'yidagini formulani ko'rsatish mumkin[19,20,21,22]:

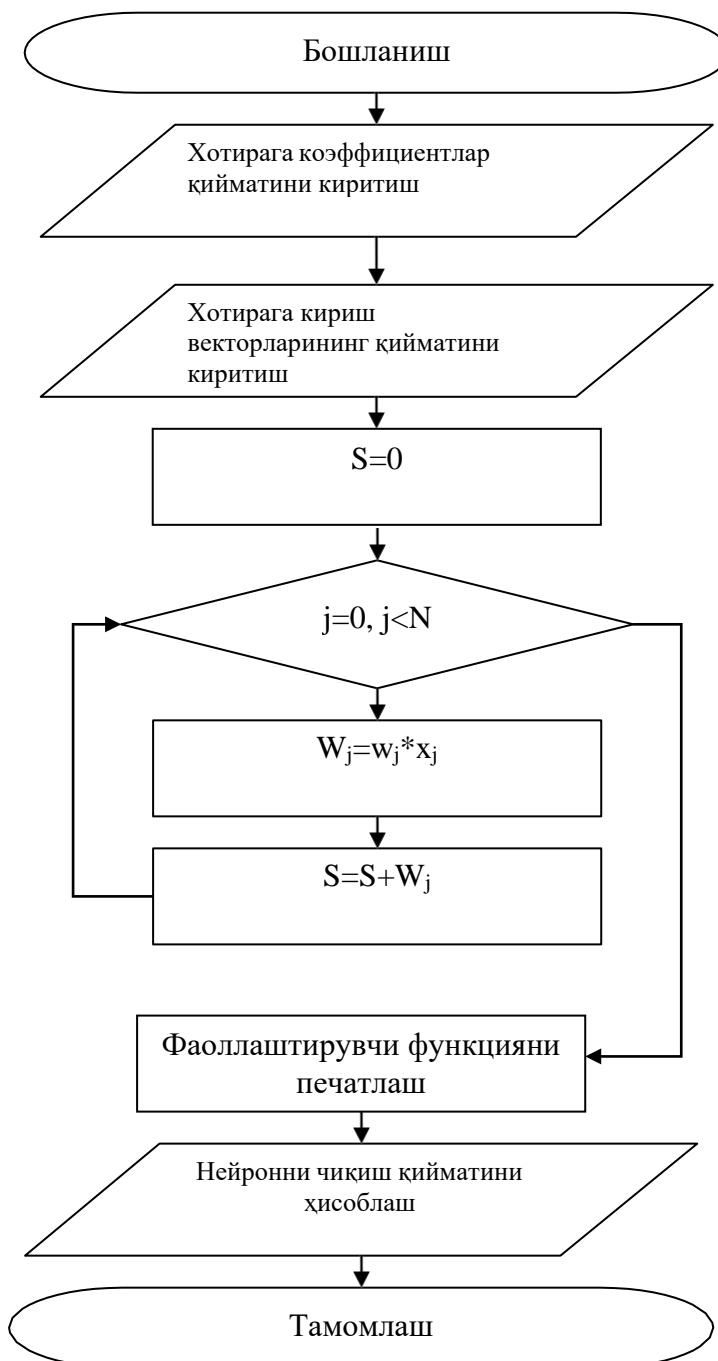
$$\varepsilon = \|d - y\| = \sqrt{\sum_{i=1}^N \sum_{j=1}^p (d_{i,j} - y_{i,j})^2} = \sqrt{\sum_{i=1}^N \sum_{j=1}^p (d_{i,j} - F(\bar{x}_i, \bar{w})_j)^2} \quad (1)$$

bu yerda  $d_{i,j}, y_{i,j}$  chiquvchi qiymatlar.

Ushbu ishda har xil mashinalarda dasturni ishlatish orqali parallel algoritmlarni ishlash vaqtining samaradorligi tekshirildi. Masalan Intel Core i7 920 prosessori har biri ikkita mantiqiy prosessoridan tashkil topgan 4 fizik yadroli kompyuterda ko'rsatkich  $2 \cdot 4 = 8$ .

Tajriba natijalari qo'yidagi jadvalda ko'rsatilgan.

Qo'yida neyron to'rlarini o'rgatishda ko'pyadroli muhitdan foydalanish usulini ko'rsatib utamiz.



Нейроннинг умумлашган чиқиш қийматини ҳисоблашнинг блок sxemasi

Векторлар қийматини икки ҳил yo'l yordamida кiritиш mumkin: Barcha vektorning қиймати оператив yadroga кiritилadi yoki ketma ket ravishda birorta portdan кiritилadi.

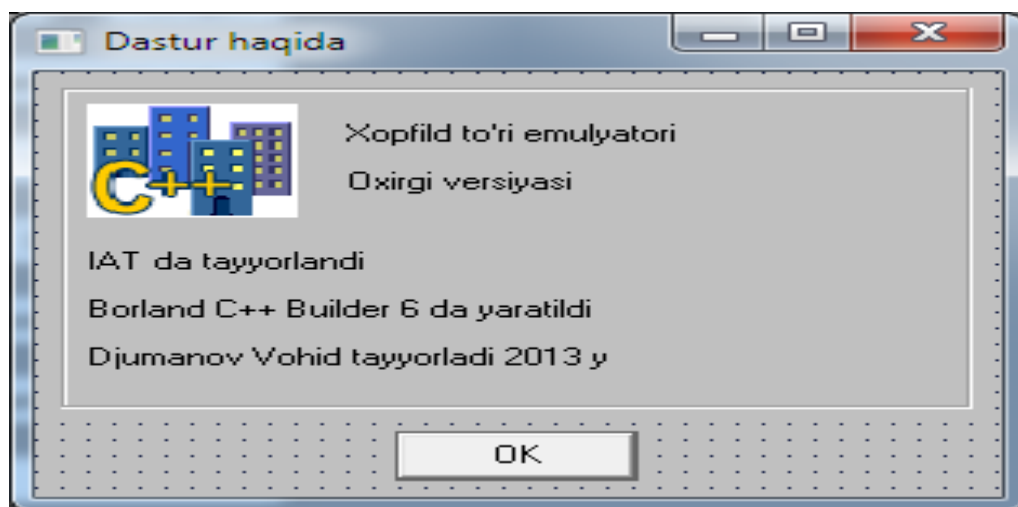
**Tasvirlarni tanishda neyron tarmoqlarining dasturiy vositasi.**

### Dasturiy vositasini ishlab chiqish

Dasturiy vositani S++ dasturlash tilida yaratamiz, buning uchun Borland C++ Bulder dasturini ishga tushiramiz va shaklga zarur elementlarni o'rnatamiz (4.1-rasm) shakl nomini "Xopfild neyron to'ri" deb nomlaymiz va unga StringGrid1, StringGrid2, StringGrid3 larni o'rnatamiz. Shundan sung, zarur tugmachalarni shaklga qo'yib chiqamiz. Ularga mos datur kodlarini ilovada keltiramiz.

Eng oldin paydo bo'ladigan "Matrisaning o'lchamlarini kiritish" shaklni yaratamiz, bu shaklda dastur matrisadagi qatorlar soni, ustunlar soni, bosh shakldagi kataklar o'lchami, qo'shish dialog oynasidagi kataklar o'lchami, ko'rishdagi kataklar o'lchami va ularni ishga to'shiruvchi tugmalardan shaklni hosil qilish, bekor qilish, chiqish tugmachalarini o'rnatamiz va ularga mos dastur kodlarini kiritamiz. Dastur kodlarini ilovada keltiramiz.

Dastur haqida ma'lumotlarni beruvchi shaklni yaratamiz, buning uchun unga oddiy Label elementlaridan foydalanib, dasstur nomi, dastur versiyasi, tayerlangan joyi, qaysi tilda yaratilganligi va kim tomonidan yaratilganligi haqida ma'lumotlarni kiritamiz va u elementlarni mos ravishda nomlaymiz.



Dastur haqida.

Ushbu shakl dasturni yaratilganligi to'g'risidagi barcha ma'lumotlarni chiqarish uchun ishlatiladi.

Keyingi shakl, yangi shablonlarni yaratish uchun yordam beruvchi dasturiy kodlarni shakllantiradi. Bu shaklni yaratish uchun RadioGroup, StringGrid va Button elementlaridan foydalanamiz, bu elementlar mos ravishda shaklni yarattishda, unga ishlov berish usullari va unga yuklatilgan vazifalarni bajarish uchun foydalaniladi. Tugmachalarga mos kodlarni yaratamiz, bu tugmachalar buyash, tozalash, invertlashva ishni bajarish uchun foydalaniladi.

1.Tanlov obektlarini sinflarga ajralganlik darajasining umumlashgan bahosi hisoblanadi.

2. Algoritmlar tuzish va ularni dastur ko'rishida amalga oshirildi.

3. Olingan natijalarni tahlil qilindi.

4. Xopfeld tarmog'i orqali tasvirlarni tanishning dasturiy vositalari yaratildi.

Shunday qilib, o'qitilgan NT noma'lum qiziqish holatining sinfini bashorat qilish uchun mustaqil bajariladigan tizim sifatida ishlatilishi mumkin va shuning uchun diagnostikada potentsial qo'llanilishi mumkin.

### **Muhokama uchun savollar**

1. Neyron to'rlari nima?
2. Piramidal neyronlar nima?
3. Suniy intellekt nima?
4. Bioinformatikada Neyron to'rlarining ro'li qanday ahamiyatga ega?
5. Bir qatlamli sun'iy neyron to'ri
6. Ko'p qatlamli neyron to'rlari nima?
7. Teskari bog'lanishli to'rlar nima?
8. Suniy neyron to'rlarini o'rgatishda parallel algorimlardan foydalanish
9. Tasvirlarni tanishda neyron tarmoqlarining dasturiy vositasini tushuntiring!

## 9-Amaliy mashg'ulot

**Mavzu:** Dori vositalarini ishlab chiqishda bioinformatsion yondashuvlarning qo'llanilishi.

### Asosiy savollar

1. Bioinformatikada dori vositalarini bashorat qilish.
2. Epigenetika, genomlar arxitekturasi va dorilarni topishda sistromalar
3. Dori vositalarini ishlab chiqish jarayonida bioinformatika va farmakogenomikaning roli.

**Kalit so'zlar:** *Dori maqsadi, Dori-darmon nomzodi, Dori skriningi, Genomika, Epigenetika, Transkriptomika, Proteomika, Farmokogenomika*

**Dars maqsadi:** Talabalarga Dori vositalarini ishlab chiqishda bioinformatsion yondashuvlarning ro'li va qo'llaniladigan usullari hamda bioinformatika va boshqa sohalardagi ahamiyati haqida tushuncha berish.

### Birinchi savol bayoni

Bioinformatik tahlil nafaqat dori maqsadini aniqlash va dori nomzodini skrining va takomillashtirishni tezlashtirishi, balki yon ta'sirlarning tavsifini osonlashtirishi va dori qarshiligini bashorat qilishi mumkin. Genomik, epigenetik, genom arxitekturasi, tsistromik, transkriptomik, proteomik va ribosoma profili ma'lumotlari kabi yuqori o'tkazuvchanlik ma'lumotlarining barchasi mexanizmga asoslangan dori kashfiyoti va dori vositalarini qayta ishlashga muhim hissa qo'shdi. Protein va RNK tuzilmalarining



to'planishi, shuningdek, gomologik modellashtirish va oqsil strukturasi simulyatsiyasining rivojlanishi, kichik molekulalar va metabolitlarning katta strukturaviy ma'lumotlar bazalari bilan birgalikda, yanada realroq protein-ligand o'rnatish tajribalari va ko'proq ma'lumot beruvchi virtual skrining uchun yo'l ochdi. Giyohvand moddalarni kashf qilish hayot sifatini pasaytiradigan yaxshi xarakterli alomatlarga ega bo'lgan kasallikning tashxisi bilan boshlanadi. An'anaviy ravishda kerakli dori kimyoviy (oddiy kimyoviy yoki murakkab oqsil bo'lishi mumkin) yoki bemorda jiddiy yon ta'sirga olib kelmasdan simptomlarni kamaytiradigan kimyoviy moddalar birikmasidir. Kerakli dorining boshqa xususiyatlariga dori ishlab chiqaruvchi kompaniyalar uchun arzonligi va foydasi, doriga chidamlilik ehtimolining pastligi dorining tijorat qiymatining keskin pasayishiga olib kelishi, atrof-muhitga past zararli ta'siri, *masalan*, qayta yo'qligi kiradi. -inson foydalanishidan keyin bakterial turlarning faollashishi. Shunday qilib, kerakli dori - bu nafaqat kam yon ta'siri bilan samarali, balki bemorga, jamiyatga va atrof-muhitga minimal uzoq muddatli salbiy ta'sir ko'rsatadigan doridir. Ushbu sharh bioinformatikaning bunday kerakli dori-darmonlarni topishga qanday yordam berishiga qaratilgan. Bioinformatika - bu genomika, transkriptomika, proteomika, populyatsiya genetikasi va molekulyar filogenetikani o'z ichiga olgan fanlararo fan. Dori-darmonlarni kashf qilishda bioinformatikachilar yuqori samarali molekulyar ma'lumotlardan foydalanadilar *simptom tashuvchilar (bemorlar, hayvonlarning kasallik modellari, saraton hujayralari liniyalari va boshqalar)* va oddiy nazorat o'rtasidagi taqqoslashda. Bunday taqqoslashning asosiy maqsadlari 1) kasallik belgilarini genetik mutatsiyalar, epigenetik modifikatsiyalar va gen ekspressiyasini modulyatsiya qiluvchi boshqa atrof-muhit omillari bilan bog'lash, 2) hujayra funksiyasini tiklaydigan yoki noto'g'ri ishlaydigan hujayralarni, *masalan*, saraton hujayralarini yo'q qila oladigan dori maqsadlarini aniqlashdir.) mo'ljallangan terapevtik natijaga erishish va nojo'ya ta'sirlarni kamaytirish uchun dori maqsadiga ta'sir qilishi mumkin bo'lgan dori nomzodlarini bashorat qilish yoki aniqlashtirish va 4) atrof-muhit salomatligiga ta'sirini va dori qarshiligining potentsialini baholash.

Bioinformatikaning dori maqsadini aniqlashga qo'shgan dastlabki hissalaridan biri simian sarkoma virusi onc geni, *v-sis* va trombositlardan kelib chiqadigan o'sish omili (PDGF) o'rtasidagi ketma-ketlik homologiyaning oddiy qatorlarni moslashtirish orqali aniqlashdir. Ushbu topilma nafaqat PDGFning saraton dori maqsadi sifatida ishlatilishiga olib keldi, balki ikkita yangi fikrlash yo'nalishiga ham olib keldi. Birinchidan, virusni o'zgartiruvchi omil oddiygina o'sish omilining vaqtinchalik ifodasini konstitutsiyaviy ifodaga o'zgartirish orqali ishlashi mumkin, bu o'sish omillarini saratonga qarshi dori ishlab chiqarish maqsadlari sifatida taklif qiladi. Ikkinchidan, gen ekspresyon modellarini modulyatsiya qiluvchi har qanday omillar saraton rivojlanishiga hissa qo'shishi mumkin. Saraton biologiyasining ushbu

yangi kontseptual asosi keyingi yillarda saratonga qarshi dori ishlab chiqish mexanizmiga asoslangan taraqqiyotga hissa qo'shdi.

**Genetik kasalliklar.** Irsiy kasalliklarga chalingan bemorlarning genomik va butun ekzom ketma-ketligi genetik kasalliklar bilan bog'liq bo'lgan ko'plab somatik mutatsiyalarni tikladi va potentsial dori maqsadlari bo'lishi mumkin. Somatik mutatsiyalar bo'yicha bioinformatik tadqiqotlar bilan bog'liq asosiy qiyinchilik, mos keladigan bemor va normal nazorat o'rtasidagi ko'plab kuzatilgan genetik farqlar orasida kasallikni keltirib chiqaruvchi mutatsiyalarni aniqlashda yotadi. Saraton kabi ba'zi kasalliklar, hatto bitta o'simta ichidagi hujayralar orasida ham yuqori genetik geterogenlik namoyon bo'ladi. Ushbu somatik mutatsiyalarning ko'pchiligi hujayraning noto'g'ri ishlashining sababi emas, balki oqibati bo'lishi mumkin.

Somatik mutatsiyalarning uch turini ajratishga harakat qilish kerak: 1) kasallikni keltirib chiqaradigan va dori maqsadi bo'lishi mumkin bo'lganlar, 2) kasallik geni bilan chambarchas bog'liq bo'lgan va shuning uchun kasallik bilan bog'liq bo'lganlar va 3) bo'lmaganlar. Kasallik bilan bog'liq, lekin nazorat guruhida emas, balki bemor guruhida sodir bo'ladi. Ikkinchi turdagi mutatsiyalar kasallik tashxisi uchun ishlatilishi mumkin, ammo dori maqsadi sifatida emas. Uchinchi turni ikki yo'l bilan chiqarib tashlash mumkin. Birinchisi, namuna hajmini oshirish. Agar minglab ko'krak saratonni bir xil somatik mutatsiyaga ega bo'lsa, unda ko'krak bezi saratoniga mutatsiyaning aloqadorligi faqat bitta ko'krak saratonida uchraydigan somatik mutatsiyaga nisbatan yuqori bo'ladi. Ikkinchisi, ko'plab kasalliklar kasallikning namoyon bo'lishidan ancha oldin genetik determinantga ega bo'lishi mumkinligini tan olgan holda, uzunlamasına ma'lumotlarni to'plashdir Aytaylik, X mutatsiyasi odamni Altsgeymer kasalligiga (AD) moyil qiladi. Agar biz AD bilan og'rigan bemorlarning bir guruhini AD bo'lmagan nazorat guruhi bilan solishtirsak va agar nazorat guruhida allaqachon X mutatsiyasiga ega bo'lgan, ammo hali AD rivojlanmagan odamlar bo'lsa, biz X mutatsiyasining ahamiyatini tushuna olmasligimiz mumkin. AD guruhida noyob. Vaqt o'tishi bilan bemorlarga yoki tegishli hayvonlar modellariga rioya qilsak, X mutatsiyasiga ega bo'lgan odam oxir-oqibat ADni rivojlantiradi degan xulosaga kelishimiz mumkin.

Oxirgi yondashuvga ma'lum tartibga soluvchi motivlarning keng qamrovli tuzilgan va tanlangan ma'lumotlar bazalari mavjudligi yordam beradi. Bioinformatika vositalari ko'pincha genomlarni tartibga solish motivlarini skanerlash uchun ishlatiladi. Bunday vositalarga ma'lum motivning genomik joylashuvini topish uchun pozitsiyaning og'irligi matritsasi (PWM), *de novo* motif kashfiyoti uchun Gibbs namuna oluvchisi va ikki guruh o'rtasidagi farqlarni chiqarish uchun ishlatilishi mumkin bo'lgan qo'llab-quvvatlovchi vektor mashinalari kiradi. ketma-ketliklar (*masalan*, motif-hozir va motiv-yo'q) va olingan ma'lumotlardan genomlardagi motivlarni aniqlash/skanerlash uchun foydalanish. Tartibga solish motivlari yadro retseptorlarining javob elementlari bo'lishi mumkin, ularning identifikatsiyasi

ko'pincha dori maqsadlarini takomillashtirishga olib keladi. Bunday tadqiqotlar DAMBE kabi dasturiy ta'minot yordamida amalga oshiriladi] izohli genomik ketma-ketlik berilganda, sichqonchani bir necha marta bosish bilan kodlash ketma-ketliklari, rRNKlar, tRNKlar, intronlar, ekzonlar, 5' va 3' qo'shilish joylari, gen xususiyatlarining yuqori yoki quyi oqimlari ketma-ketliklari va *hokazolarni ajratib olish mumkin*. PWM, Gibbs namuna oluvchisi va minimal katlama energiyasini baholash funksiyalariga qo'shimcha ravishda, DAMBE oqsil izoelektrik nuqtasi va oqsillarni tarjima qilish samaradorligi indekslarini ham hisoblashi mumkin.

Patogenga qarshi dori-darmonlarni ishlab chiqishning ikkinchi bosqichi - bu muhim genlarning xostda gomologlari bor-yo'qligini tekshirish. Agar ular shunday bo'lsa, patogenda bunday muhim genlarni inhibe qilish xost gomologining funksiyasiga salbiy ta'sir ko'rsatishi mumkin va shuning uchun patogen va xost gomologlari o'rtasida ketma-ketlik va tizimli taqqoslashlarni amalga oshirishimiz kerak, bu patogen gomologning noyob qismini aniqlashga yordam beradi. patogenga xos dori-darmonlarni loyihalashda.

Uchinchidan, patogenning dori qarshiligini rivojlantirish ehtimolini minimallashtirish uchun preparat patogen bo'lmagan uning filogenetik qarindoshlariga emas, balki o'ziga xos patogenga qaratilgan bo'lishi muhimdir. Shu sababli, patogen bakteriyalarda noyob bo'lgan, ammo patogen bo'lmagan qarindoshlarida bo'lmagan patogenlik orollari tobora ko'proq dori maqsadlari uchun afzal qilingan manbaga aylandi.

*Bioinformatik tahlil patogen Clostridium perfringens* tarkibida mavjud bo'lgan, ammo sutemizuvchilar va qushlarda yo'q bo'lgan glutamat tashish tizimini aniqladi. Bunday transport tizimiga qarshi ishlab chiqilgan dorilar nafaqat odamlarni, balki xonaki sutemizuvchilar va parrandalarni ham himoya qiladi. Inson paraziti *Giardia intestinalis*da fosfoinositid-3 kinaz (PI3K) signalizatsiya yo'llari muhim va dori maqsadi sifatida xizmat qilishi mumkin. Biroq, PI3K yo'li ko'plab eukariotlarda ham muhim ahamiyatga ega, shuning uchun sutemizuvchilarga nisbatan *G. intestinalis*dagi PI3K gomologlarida (*Gipi3k1* va *Gipi3k2*) nima noyob ekanligini aniqlash muhimdir. Ketma-ket taqqoslashlar faqat parazitda patogenga xos dori maqsadi bo'lib xizmat qilishi mumkin bo'lgan noyob qo'shimchani aniqladi. *Xuddi shu yondashuv Pseudomonas aeruginosa* ni nishonga olishda qo'llaniladi. Xuddi shunday, OIV-1ga qarshi dori-darmonlarni ishlab chiqishda uning tarjima qilingan poliproteinining teskari transkripsiyasi va proteaz hazm bo'lishida ishtirok etuvchi genlarni nishonga olish mumkin, chunki bu jarayonlar nafaqat virusning omon qolishi va tarqalishi uchun zarur, balki insonda yaqin gomologlarga ega emas, shuning uchun ularning inhibisyonu. insonga minimal yon ta'sir ko'rsatishi kerak.

Genomik tahlil mavjud dorilarni boshqa patogenlarga qarshi qayta ishlashga yordam beradi. Galaktofuranoza ( $Gal_f$ ) turli xil bakterial patogenlarning hujayra

yuzasida muhim tarkibiy qismdir] va uning sintezi UDP-galaktopiranoz mutazasini (UGM) talab qiladi. Gal<sub>f</sub> insonda yo'qligi sababli, UGM kerakli dori maqsadi sifatida ishlatilgan *GLF* geni bilan kodlangan UGM keyinchalik bir nechta eukaryotik bir hujayrali patogenlarda shuningdek nematodlarda topilgan. Eukaryotik bir hujayrali patogenlarga qarshi kurashish uchun bakterial patogenlarga qarshi ishlab chiqilgan dorilarni qayta ishlata olamizmi?]? Giyohvand moddalarni qayta ishlash dori vositalarini ishlab chiqishda tejamkor. Genomik tahlil shuni ko'rsatadiki, eukaryotik UGMlar bir-biriga o'xshash bo'lsa-da, prokaryotik UGMlardan sezilarli darajada farq qiladi, bu esa preparatni bakterial patogendan eukaryotik patogenlarga qayta tiklashda qiyinchiliklarni ko'rsatadi. Biroq, agar bitta eukaryotik UGMga qarshi samarali dori ishlab chiqilsa, preparat boshqa eukaryotik patogenga qayta qo'llanilishi uchun juda yaxshi imkoniyatga ega bo'ladi.

Genomika, shuningdek, giyohvand moddalar ta'sirini tushunishga hissa qo'shdi. *Psalmopoeus cambridgei* o'rgimchakning PcFK1 zaharli oqsili *Plasmodium falciparum* o'sishini inhibe qila oldi, ammo mexanizmi noma'lum edi. Ketma-ket tahlil PcFK1 va *P. falciparum* fermenti PfSUB1 oqsil substrati o'rtasidagi ketma-ketlik homologiyasini aniqladi, bu PcFK1 PfSUB1 antagonisti degan gipotezaga olib keldi. Keyinchalik docking bashorati va *in vitro* tajribalari bu farazni tasdiqlaydi, PfSUB1 ga dori maqsadi sifatida ishora qiladi.

Bioinformatika, o'ziga xos evolyutsion nuqtai nazari va molekulyar filogenetikaning integratsiyasi bilan ko'pincha molekulyar mexanizmlar bo'yicha bahslarni hal qilishga hissa qo'shishi mumkin. Bunday misollardan biri o'z-o'zidan dezaminatsiya orqali sodir bo'lgan keyingi C→T mutatsiyasi orqali CpG etishmovchiligini keltirib chiqaradigan CpG metilatsiyasining sababiy talqinini o'z ichiga oladi. *Mycoplasma genitalium* va *M. pneumoniae* genomlarida DNK CpG metiltransferaza yo'qligi aniqlanganda, *M. genitalium* genomida *M. pneumoniae* genomiga qaraganda ancha kuchli CpG tanqisligi namoyon bo'lganida munozara yuzaga keldi, bu ikki tur o'rtasidagi farq CpG etishmovchiligi degan xulosaga keldi. CpG metilatsiyasiga aloqador emas. Evolyutsion istiqbolsiz genomik tadqiqotlardan olingan bunday xulosa ko'pincha noto'g'ri. DAMBE dasturiy ta'minotidan foydalangan holda keng qamrovli filogenetik tadqiqot shuni ko'rsatdiki, ikki turning ajdodlari bir nechta CpG metiltransferazalariga ega bo'lishi kerak, chunki *M. pulmonis* va *M. genitalium* va *M. pneumoniae* dan oldinroq bo'lgan boshqa qarindoshlar bir nechta CpG metiltransferazalariga ega. *M. genitalium* va *M. pneumoniae* ajdodida CpG metiltransferazalarini yo'qotgandan so'ng, ikkala tur ham CpG chastotasiga ega bo'la boshladi, lekin *M. pneumoniae* ancha tez rivojlandi (juda uzunroq filial bilan) va *M. genitalium* dan ancha tezroq CpG ni qaytarib oldi. Ushbu topilmalar CpG-ga xos DNK metilatsiyasi va CpG etishmovchiligi o'rtasidagi sabab-oqibat munosabatlarining haqiqiylikini tikladi va biologik jarayonlarni tushunishda

evolyutsion nuqtai nazarga ega bo'lish muhimligini ko'rsatadi. Ko'pgina bunday tadqiqotlar yuqori darajada ajralib turadigan bakterial yoki virusli turlarni o'z ichiga olganligi sababli va juda ko'p divergent ketma-ketliklar bilan ishonchli bir nechta ketma-ketlik moslashuvini olish ko'pincha qiyin bo'lganligi sababli, yaqinda juftlik ketma-ketligini moslashtirishga asoslangan yangi filogenetik usul ishlab chiqildi juda xilma-xil turlar.

### **Ikkinchi savol bayoni**

**Epigenetika, genomlar arxitekturasi va dorilarni topishda sistromalar.** Yuqorida aytib o'tilgan ALD mutatsiyasi kabi bir xil zararli mutatsiyalarga ega bo'lgan monozigotik egizaklar ko'pincha fenotipda juda farq qiladi. Bunday kuzatishlar epigenetik modifikatsiyalar va inson kasalliklari o'rtasidagi munosabatlarni ta'kidlash uchun xizmat qiladi. Epigenetik modifikatsiya ikkita o'zaro bog'liq hodisa, DNK metilatsiyasi va giston modifikatsiyasini o'z ichiga oladi. Sutmizuvchilarda DNK metilatsiyasi naqshini saqlash sutemizuvchilar DNK metiltransferaza 1 (DNMT1) tomonidan amalga oshiriladi, uning CatD domeni yarim metillangan CpG saytlarini taniydi shunday qilib, DNK metilatsiyasi sxemasi ota-onadan qiz hujayralarigacha saqlanib qolishi mumkin. Sutmizuvchilarda metillangan CpG MBD1, MBD2, MBD3 va MeCP2 kabi metil-CpG bog'lovchi domeniga ega bo'lgan oqsillarni o'z ichiga oladi, so'ngra ular asetil guruhini olib tashlash va lizin qoldiqlarining (yoki giston N-terminalining) musbat zaryadini tiklash uchun giston deasetilazani jalb qiladi. giston, shuning uchun DNKning manfiy zaryadlangan magistralini genni o'chirish uchun musbat zaryadlangan giston atrofiga mahkam o'ralishi mumkin. Ovozsiz gen ko'p jihatdan funktsiyani yo'qotuvchi mutatsiyaga teng. Ba'zi saratonlar DNK metilatsiyasi va giston deasetilatsiyasi orqali apoptoz yo'lida ishtirok etadigan genlarning doimiy ravishda o'chirilishi natijasida paydo bo'lgan ko'rinadi histon deasetilaz apoptoz yo'lini qayta faollashtirishga qaratilgan inhibitorlari bilan dori maqsadi sifatida ishlatilgan. Ushbu yondashuvdagi asosiy muammo o'ziga xoslikdir, chunki deasetilaz inhibitorlari ko'pincha boshqa ko'plab genlarning regulyatsiyasiga chuqur ta'sir ko'rsatadi, bu nima uchun bunday dorilar ko'pincha klinik sinovlarga kirmasligini tushuntirishi mumkin. Hozirgi vaqtda epigenomni aniq tahrirlash, DNK bilan bog'lanish uchun komponentlar va o'ziga xos ketma-ketlikni aniqlash va o'zgartirish usullari ishlab chiqilmoqda.

DNK metilatsiyasi va giston deasetilatsiyasi asosan genlarni doimiy ravishda susaytirish maqsadiga xizmat qiladi, degan an'anaviy qarash endi epigenetik modifikatsiya va genlarni tartibga solishning umumiy kontseptual asosi bilan almashtirildi. Ushbu kontseptual siljish bir necha turdagi yuqori o'tkazuvchanlik ma'lumotlarini kompleks tahlil qilishni talab qiladi: bisulfit ketma-ketligidan metilatsiya sxemasi, ChIP-on-chip va ChIP-Seq dan DNK va oqsillarni bog'lash ma'lumotlari (**sistrom**) va genom Hi-C yoki uning hosilalari arxitektura

ma'lumotlari. DNK metilatsiyasi DNK/oqsil bog'lanishini o'zgartiradi, bu esa o'z navbatida genom arxitekturasini o'zgartiradi, ya'ni chiziqli DNK bo'ylab bir-biridan uzoqda joylashgan ikkita DNK segmentini birlashtirish mumkin. Genom arxitekturasi ma'lumotlari bir-biridan bir million bazagacha bo'lishi mumkin bo'lgan kuchaytirgichlar va promoterlar o'rtasidagi fazoviy o'zaro ta'sirni o'rganish uchun yo'l ochadi. Genning genomdagi joylashuviga bog'liqligi 1930 yildan beri translokatsiyani o'rganish orqali ma'lum, ammo empirik dalillar ko'p vaqt o'tishi bilan oqsil DNK o'zaro ta'siri nukleosomalarning qayta konfiguratsiyasi va kuchaytirgich va promotor o'rtasidagi o'zaro ta'sirga olib kelganligini ko'rsatish uchun to'plangan. Bu genlarni tartibga solishning kuchaytiruvchi uyasi modelining formulasini yaratdi. Ya'ni, markaz bir yoki bir nechta kuchaytirgichni o'z ichiga oladi va uning promotori markazga yaqin bo'lgan gen ifodalanadi; bunday markazni yo'q qilish ularning ifodalanadigan markazga jismoniy yaqinligiga bog'liq bo'lgan barcha genlarning ifodasini o'chiradi.

Bioinformatika nuqtai nazaridan asosiy savol DNKdagi metillanish signali nima va epigenetik modifikatsiyani o'zgartirish uchun bunday signalni modulyatsiya qilish mumkinmi yoki yo'qmi. Men ba'zi b-talassemiya bilan og'riqan bemorlarda b-globin genining to'g'ri versiyasi borligi haqida aytib o'tgan edim, ammo gen o'zidan uzoqda sodir bo'lgan mutatsiyalar tufayli ifodalanmaydi. Ikkita farazni shakllantirish mumkin. Birinchidan, b-globin genining ifodasini nazorat qiluvchi kuchaytirgich bemorda mutatsiyaga uchragan yoki o'chiriladi. Ikkinchidan, normal genom arxitekturasida b-globin geni promotoriga yaqinlashtirilgan kuchaytirgich g'ayritabiiy epigenetik modifikatsiyalar va oqsil/DNK bilan bog'lanish tufayli boshqa joyga ko'chiriladi. Genom arxitekturasi, metilatsiya naqshlari va tsistromlarning (barcha oqsil/DNK bog'lanish joylari to'plami) yuqori o'tkazuvchanlik ma'lumotlari mavjudligi bilangina mumkin bo'lgan ushbu gipotezalarni sinab ko'rish biz kuchaytirgich va b ni qanday o'zgartirishimiz mumkinligiga oydinlik kiritadi. -globin promotori, shuning uchun gen ifodalanadi. Xuddi shunday, agar b-globin geni DNK metilatsiyasi orqali o'chirilsa, metilatsiya sxemasini o'zgartirish uchun signalni qanday modulyatsiya qilish haqida bilim bizni o'chirilgan b-globin genini qayta faollashtirishga yaqinlashtiradi. Xuddi shu fikrga ko'ra, agar xomilalik globin genlari metilatsiya bilan o'chirilgan bo'lsa va bu homila globin genlarining qayta faollashishi kattalar globin genlaridagi mutatsiyalar tufayli yuzaga keladigan muammoni yengillashtirsa, u holda saytga xos demetilatsiyani bilish juda ma'qul bo'lar edi.

Agar noto'g'ri DNK metilatsiyasi namunasi shakllangan bo'lsa, u holda ideal dori (yoki epigenomni tiklaydigan nano-mashina) noto'g'ri naqshni aniq aniqlashi va uni tuzatishi kerak. Bunday dori yoki nano-mashinani ishlab chiqish uchun biz birinchi navbatda to'g'ri metilatsiya naqshini bilishimiz yoki ideal tarzda shunday to'g'ri naqshni kodlaydigan molekulalar to'plamini topishimiz kerak. Eksperimental natijalar

RNKning epigenetik modifikatsiyadagi rolini qo'llab-quvvatlash uchun to'plangan. Zigotadagi DNK pluripotentsiyani qayta tiklash uchun demetilatsiyaga uchraganligini hisobga olsak, epigenomik kod, ehtimol DNKda emas. Proteinlar kod yozishda yaxshi emasdek tuyuladi va ko'pchilik asosiy gistonlar erkak jinsiy hujayralarida protamin bilan almashtiriladi, epigenetik kodlar, ayniqsa de novo DNK metilatsiyasini aniqlaydiganlar ] va Kcnq1ot1 [Kcnq1ot1] kabi lncRNKning bog'lanishini aniqladi. DNK metilatsiyasini oqsillarda topish dargumon. Biroq, bunday kodlar oosit va sperma bosqichida mavjud bo'lishi mumkin bo'lgan yuqori darajada saqlanib qolgan va tizimli barqaror RNK molekulalarida mavjud bo'lishi mumkin. Uzoq kodlanmagan RNKlar (lncRNA) epigenetik modifikatsiyada ishtirok etishi va xromatin holatini boshqarishi mumkin. ChiRP-seq usuli bilan DNK va oqsilga bog'langan lncRNAlarning tavsifi DNKda ko'plab ketma-ketliklarga xos bog'lanish joylarini va HOTAIR kabi lncRNKning bog'lanishini aniqladi.] bunday saytlarga giston H3 lizin-27 trimetilatsiyasini vositachilik qilish uchun Polycomb Repressiv Kompleks 2 (PRC2) ni jalb qilishni osonlashtiradi. Qisqa RNKlar epigenetik o'zgarishlarni ham modulyatsiya qilishi mumkin. Yetuk spermatozoidlar bir qancha kichik RNK turlarini o'z ichiga oladi va bu kichik RNKlar nasl fenotipiga ta'sir qiladi. Bundan tashqari, nasldagi bu kichik RNKlar epigenetik modifikatsiyaga hissa qo'shadi. ENCODE pilot loyihasi shuni ko'rsatadiki, "genom keng tarqalgan bo'lib transkripsiyalanadi, shuning uchun uning asoslarining aksariyati birlamchi transkriptlarda topilishi mumkin. Ushbu kodlanmagan transkriptlar bioinformatikachilar uchun epigenomni o'zgartiruvchi RNKlarni dori maqsadlari sifatida topish uchun xazina bo'lishi mumkin.

### **Uchinchi savol bayoni**

Giyohvand moddalarni topish va ishlab chiqish murakkab, yuqori xavfli, vaqt talab qiluvchi va potentsial yuqori darajada foydali jarayondir. Farmatsevtika kompaniyalari uni bozorga olib chiqish uchun har bir dori uchun millionlab dollarlarni yoqib yuborishadi. Yangi dori vositasini yaratish texnologik tajriba, inson resurslari va katta kapital qo'yilmalarni talab qiladi. Bundan tashqari, yangi dori bozorga chiqmasdan oldin sinov va ishlab chiqarish standartlari bo'yicha qoidalarga qat'iy rioya qilishni talab qiladi va aholining umumiy qatlamida qo'llanilishi mumkin, aslida u bozorga chiqmaydi. Bu omillarning barchasi yangi kimyoviy ob'ektni tadqiq qilish va ishlab chiqish xarajatlarini oshiradi. Dori-darmonlarni loyihalash jarayoniga ijobiy ta'sir ko'rsatgan va umumiy xarajatlar va xavfni kamaytiradigan ikkita tarmoq - Bioinformatika va Farmakogenomika. Ularning dori-darmonlarni loyihalash jarayonidagi amaliyoti umumiy jarayonga ijobiy ta'sir ko'rsatdi va ular dori-darmonlarni loyihalashning turli bosqichlarini tezlashtirishi, narxini pasaytirishi va har



doim ham mumkin. Joriy eslatma bioinformatika va farmakogenomikaning dori vositalarini kashf qilish va ishlab chiqish jarayonidagi roliga qaratilgan.

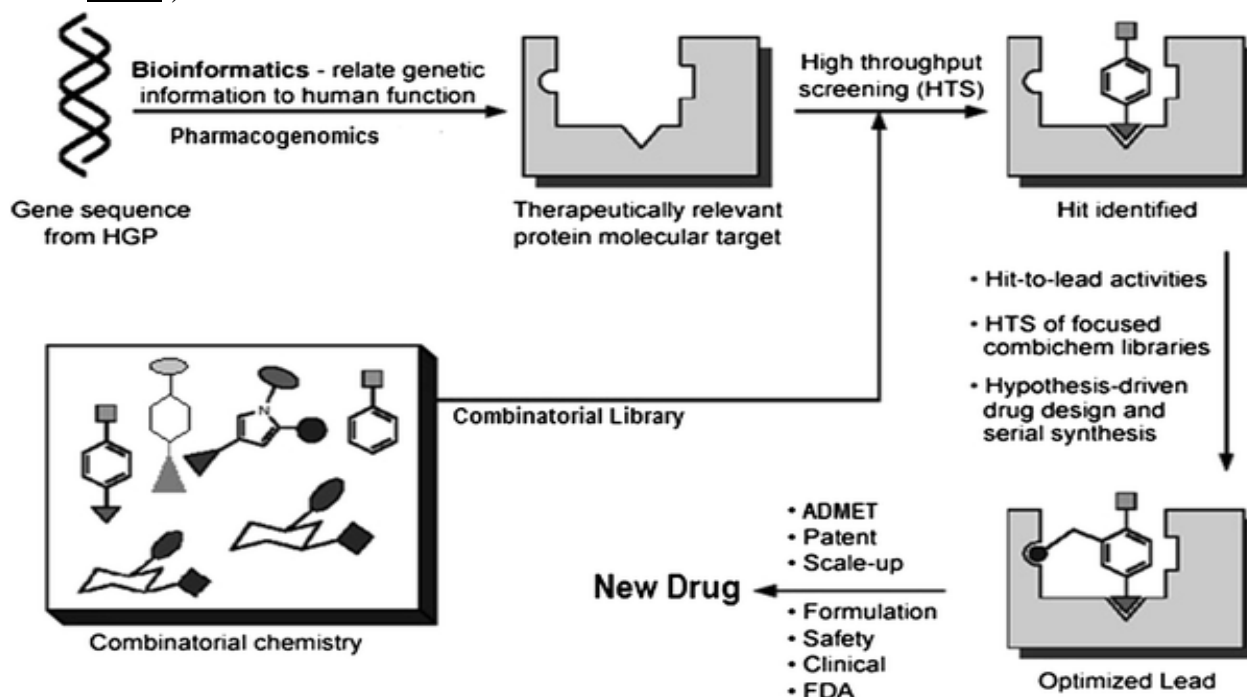
**Bioinformatika va dori kashfiyoti.** Giyohvand moddalarni aniqlash - bu yangi nomzod dorilarni kashf qilishning bosqichma-bosqich jarayoni. An'anaga ko'ra, farmatsevtika kompaniyalari yaxshi tasdiqlangan farmakologiya va kimyoga asoslangan dori-darmonlarni kashf qilish yondashuvlariga amal qiladi va yangi dori-darmonlarni topishda turli qiyinchiliklarga duch keladi (Iskar va boshq. 2012 ). Yuqori raqobatbardosh farmatsevtika sanoatida "g'olib hamma narsani oladi" bo'lsa, yangi kimyoviy ob'ektni (NCE, ya'ni, yangi dori nomzodi) muayyan davolash uchun patentlagan birinchi kompaniya barcha o'ljalarni oladi va boshqa raqobatchilar asosan patentning amal qilish muddati tugashini kutishga majbur bo'ladi. yiriklik (Iskar va boshq. 2012 ). Shunday qilib, bugungi kunda farmatsevtika kompaniyalari dori vositalarini ishlab chiqish jarayonining istalgan bosqichini tezlashtirish potentsialini ko'rsatadigan barcha yondashuvlarga katta miqdorda sarmoya kiritadilar (Whittaker 2003 ). Qisqa vaqt ichida past xavf bilan ko'proq va ko'proq dori ishlab chiqarish uchun ortib borayotgan bosim bioinformatikaga (Ortega va boshq. 2012 ) ajoyib qiziqish uyg'otdi . Darhaqiqat, hozirda kompyuter yordamida dori vositalarini loyihalash (CADD) deb nomlanuvchi yangi, alohida soha mavjud (Song et al. 2009 ; Speck-Planche and Cordeiro 2013 ).

**Dori ta'sirini aniqlash.** Hozirgi bioinformatika yondashuvlarining asosiy yo'nalishlaridan biri biologik faol nomzodlarni bashorat qilish va aniqlash (Whittaker 2003 ) va tegishli ma'lumotlarni qazib olish va saqlashdir. Giyohvand moddalar odatda faqat ushbu dorilarning ta'siri uchun aniq dori maqsadi aniqlangan va o'rganilganda ishlab chiqariladi. Dori-darmonlarni kashf qilish jarayonining potentsial maqsadlari soni eksponent ravishda ortib bormoqda. Bioinformatikadan foydalangan holda inson genomi ketma-ketligini qazib olish va saqlash yangi dorilar uchun ko'proq potentsial taklif qiluvchi yangi maqsadlarni aniqlashdan tashqari, maqsadli oqsillarni kodlash uchun mas'ul bo'lgan ushbu genlarning nukleotid tarkibini aniqlash va tasniflashga yordam berdi (Chen va Chen). 2008 ; Katara va boshqalar 2011). Bu inson genomi ma'lumotlari asosiy rol o'ynashi kutilayotgan sohadir (Yamanishi va boshq. 2010 ). Giyohvand moddalarni ishlab chiquvchilarga odatiy bo'lmagan hashamatli tanlov taqdim etiladi, chunki ko'proq genlar aniqlanadi va giyohvand moddalarni kashf qilish sikli ko'proq ma'lumotga ega bo'ladi (Loh and Soong 2011 ). Bioinformatika tobora ko'proq biologik dori maqsadlarini aniqlash va tahlil qilish imkonini beradi; Shunday qilib, farmatsevtika kompaniyalari quvurlarida potentsial dori vositalarining nafasini sezilarli darajada oshirishi kutilmoqda (Whittaker 2003 ; Ortega et al. 2012 ).

**Dori vositalarining maqsadini tekshirish.** Bioinformatika, shuningdek, yangi dori maqsadlarini bashorat qilish va mavjud dori maqsadli ma'lumotlarini saqlash va

boshqarish uchun strategiyalar va algoritmlarni taqdim etadi. "Potensial" dori maqsadlari topilgandan so'ng, taxminiy maqsad va qiziqish kasalligi o'rtasida kuchli bog'lanishni o'rnatish uchun beqiyos ehtiyoj bor (Yamanishi va boshq. 2010 ). Bunday asosiy birlashmaning tashkil etilishi dori vositalarini ishlab chiqish jarayonini asoslab beradi. Maqsadni tekshirish deb nomlanuvchi bu jarayon bioinformatika muhim rol o'ynaydigan sohadir ( 1- rasm ). Dori-darmonlarni maqsadli tekshirish klinik sinov va tasdiqlash bosqichlarida muvaffaqiyatsizlik potentsialini mo'tadil qilishga yordam beradi (Ratti va Trist 2001 ; Gilbert va boshq. 2003 ; Whittaker 2003 ).

Dori-darmonlarni topish va ishlab chiqishning hozirgi yuqori narxi farmatsevtika kompaniyalari orasida tashvishlanishning asosiy sababidir (Dickson and Gagnon 2004a , b ; Adams and Brantner 2010 ). Hosildorlikni oshirish bilan bir qatorda, farmatsevtika kompaniyalari ham dori-darmonlarni kashf qilish jarayonida yuqori muvaffaqiyatsizlik darajasini pasaytirishni maqsad qilgan, shuning uchun bozorga chiqishga qodir bo'lgan dorilar soni ko'paygan (Tsaionun et al. 2009 ). Klinik sinovlarning turli bosqichlarining yuqori narxi farmatsevtika kompaniyalari tomonidan ishlab chiqilishi mumkin bo'lgan dori vositalari sonini cheklovchi omil bo'lib xizmat qiladi va shuning uchun tasdiqlash uchun eng yaxshi imkoniyatga ega birikmalarni tanlash juda muhim (Klein va Tabarrok 2003 ; Wierenga and Eaton 2004 ); Tamimi va Ellis 2009 ).



Dori vositalarini topish jarayonining turli bosqichlarida bioinformatikaning roli

Bioinformatika maqsadli kashfiyotlar va tekshirish yondashuvlarini yanada samaraliroq ta'minlaydi, shuning uchun tasdiqlash jarayonida ko'proq dori nomzodlari muvaffaqiyatli bo'lishini ta'minlashga yordam beradi va uni yanada tejamkor qiladi (Ortega va boshq. 2012 ).

**Yangi dori vositalarini ishlab chiqishni rag'batlantirish.** Farmatsevtika sanoatini bezovta qiladigan ba'zi garov xarajatlari mavjud (Collier 2009). Ushbu xarajatlar tijoratlashtirish xarajatlarini o'z ichiga oladi (Gilbert va boshq. 2003; Mullin 2003; Liang and Mackey 2011), sud jarayoni va giyohvand moddalarni chaqirib olish xarajatlari (Klein and Tabarrok 2003) va jamiyatga umumiy xarajatlar (Lazarou va boshq. 1998; Rawlins 2004). Yangi dori-darmonlarni tijoratlashtirish xarajatlari har bir tasdiqlangan dori uchun taxminan 250 million dollarni tashkil qiladi (Gilbert va boshq. 2003; Dickson va Gagnon 2004a, b.), yuqori, asosan, tasdiqlangan "yangi" dorilarning aksariyati allaqachon mavjud bo'lgan dori vositalarining funksional nusxalari bo'lganligi sababli. Ko'pchilik dori vositalari allaqachon mavjud bo'lgan kasallikni davolash uchun tijoratlashtirilmoqda; Shunday qilib, shunga o'xshash dori-darmonlarga kirish imkoniga ega bo'lgan shifokorlar va bemorlarning e'tiborini tortadigan interfeysga ehtiyoj bor (Meyers and Baker 2001). Bioinformatika to'g'ri interfeys sifatida harakat qilishi mumkin va farmatsevtika kompaniyalariga potentsial dori maqsadlarini samarali aniqlash va yangi dori vositalarini ishlab chiqish uchun yangi yondashuvlar va imkoniyatlarni taqdim etadi (Whittaker 2003). Agar dori vositalari allaqachon mavjud ekvivalentlari bilan raqobatda tijoratlashtirilmasa, ularni tijoratlashtirish xarajatlari sezilarli darajada pasayishi kutilmoqda (Meyers va Beyker 2001);; Simoens 2011; Liang va Makki 2011).

**Dori-darmonlarni kashf qilish va ishlab chiqishda farmakogenomika.** Farmakogenomika genetik polimorfizm va genomik variantlarning dori ta'siriga ta'sirini bildiradi, uning bilimlari optimal dori, doza va davolash jarayonini tanlashda yordam beradi va salbiy dori reaksiyalaridan qochishga yordam beradi (Amstutz and Carleton 2011). Kasallik bilan bog'liq bo'lgan genetik biomarkerlar haqidagi bilimlar farmatsevtika kompaniyalariga farmakogenomikaga asoslangan, aniqroq va individuallashtirilgan dori va dozalarni ishlab chiqishga yordam beradi (Liou va boshq. 2012). Bu dori vositalari yoki uning qo'shimcha mahsulotlari bilan o'zaro ta'sirini ko'rsatadigan va qandaydir tarzda farmakodinamikaning dorilar farmakokinetikasi bilan bog'liq bo'lgan genetik naqshlarni va o'sha genetik elementlarning polimorfizmini kuzatish orqali amalga oshiriladi (Bernard 2003; Whittaker 2003; Liou va boshq. 2012).

Farmakogenomik yondashuv farmatsevtika kompaniyalariga umumiy aholining alohida genetik kichik guruhlariga talablariga javob beradigan dori-darmonlarni ishlab chiqish imkonini beradi (Nelson va boshq. 2009). Farmakogenomikaning asosiy qiziqishi dori samaradorligini bashorat qilish mumkin bo'lgan bemorlarni aniqlash va dori vositalarining salbiy ta'siri xavfini kamaytirishdir (Nelson va boshq. 2009). Bemorlarning genetik profili asosida dori-darmonlarni buyurishning farmakogenomik yondashuvlari va'dasi "Shaxsiylashtirilgan tibbiyot" deb ataladi. Bu dori retseptlaridagi spekulyatsiyani kamaytiradi va shu tariqa shifokor va bemorning

ishonchini oshiradi va dori-darmonlarni topish va ishlab chiqish, diagnostika, davolash va kasalliklarning oldini olish strategiyalariga ustunlik qiladigan yondashuvlarni o'zgartiradi (Prows and Prows 2004 ).; Zemlo 2004 ).

Dori-darmonlarni loyihalash juda murakkab, qimmat va vaqt talab qiluvchi jarayondir. Bioinformatika ham, farmakogenomika ham xarajat va vaqt kontekstini turli yo'llar bilan yengish uchun katta yordam beradi. Bioinformatika dori vositalarini loyihalash va ishlab chiqish jarayoni bilan bog'liq bo'lgan turli maqsadlarda foydalanish mumkin bo'lgan dori vositalariga oid ma'lumotlar bazalari va dasturiy ta'minotlarining keng doirasini taqdim etadi. Xuddi shunday, farmakogenomika o'zgaruvchan dori reaksiyasi haqida genom darajasida ma'lumot beradi, bu farmatsevtika kompaniyalari uchun etim dori bilan bir qatorda yangi dori-darmonlarni ishlab chiqish va mavjud dori-darmonlarni saqlash uchun juda muhimdir. Bioinformatika va farmakogenomika hali boshlang'ich bosqichida bo'lsa-da va hozirda ba'zi to'siqlarga duch kelgan bo'lsa-da, ular yaqin kelajakda dori vositalarini ishlab chiqish jarayoniga yordam berish uchun etarli salohiyatga ega.

<https://link.springer.com/article/10.1007/s13721-013-0039-5>

### **Muhokama uchun savollar**

1. Dori vositalarini ishlab chiqarishda bioinformatik tahlil nima?
2. Patogenlar keltirib chiqaradigan inson kasalliklarini ayting?
3. Dori vositalarini ishlab chiqarishda genomikani ro'li qanday?
4. Dori vositalarini ishlab chiqarishda sistromalar nima?
5. DrugBank nima?
6. Dorilarni loyihalash jarayonida bioinformatika taraqqiyotidagi to'siqlarni ayting!

### **Amaliy mashg'ulotlarni bajarish yuzasidan ko'rsatmalar**

#### **1-amaliy mashg'ulot**

**Mavzu:** Zamonaviy bioinformatsion ma'lumotlar bazalari haqidagi ma'lumotlar bilan tanishish. Bibliografik ma'lumotlar bazalari. Matnli ma'lumotlarni olish instrumentlari.

**Dars maqsadi:** Talabalarga Zamonaviy bioinformatsion ma'lumotlar bazalari bilan foydalanishni o'rganish uchun fundamental hamda amaliy bilimlarga ega bo'lishini taminlash.

**Ishning mazmuni:** Zamonaviy biologik tadqiqotlarda bioinformatikaning ahamiyati. Bioinformatika biologiyaning ilmiy tajribalari asosida olingan natijalami tahlil qiladi. Olingan ma'lumotlarni tadqiqotchi ma'lumotlar bazasida mavjud bo'lgan

barcha to'plamlar bilan solishtiradi. Bordini, u o'zi aniqlagan ketma- ketlikni ma'lumotlar bazasidan topa olmasa bunda u bu ma'lumotni shu joyga kiritib qo'yadi va bu bilan bazani yanada boyitadi. Ma'lumotlar bazasi funksiyalariga saqlash, tizimlashtirish, axborotlarni yangilab turish unga kirish huquqi bilan ta'minlashlar kiradi. Bu operatsiyalar esa katta qudratlardagi kompyuterlarni talab qiladi. Shuningdek biologik mavzular majmuidagi ilmiy nashriyotlar bazalari ham mavjud. Biologiya bo'yicha istalgan ilmiy jurnalning barcha sonlarida chiqadigan har bir maqola ma'lumotlar bazasiga joylashtiriladi izlanuvchi uni internet tarmog'i orqali oson topib olishi uchun qisqa ta'rif berib qo'yiladi.

*NCBI ma'lumotlar bazasining asosiy oynasi.*

- Mahalliy ketma-ketliklarni qidiruv vositasi (BLAST)
- BioProject (ilgari inson Genomi loyihasi)
- BioSistemalar
- BLAST (yakka tartibda)
- Konservativ domenlar ma'lumotlar bazasi (CDD)
- Bazada Saqlanadigan domenlarni qidirish xizmati (CD qidirish)
- Genomik tarkibiy o'zgarishlarning ma'lumotlar bazasi (dbVar)
- Genotiplar va fenotiplar to'g'risidagi ma'lumotlar bazasi (dbGaP)
- Single nukleotid Polimorphism Ma'lumotlar Bazasi (dbSNP)
- FTP: FASTA BLAST ma'lumotlar bazasi
- FTP: Genom xaritasi ma'lumotlari
- GenBank

- Genom BLAST
- Primer-BLAST
- PubMed
- Taksonomiya umumiy daraxti
- Vektorni tahrirlash vositasi (VAST)

База информационных ресурсов для <i>Arabidopsis thaliana</i> (TAIR)	Кодирующие последовательности для <i>Arabidopsis thaliana</i>
База данных геномных и других биологических характеристик <i>Caenorhabditis elegans</i> (WormBase)	Кодирующие последовательности для нематоды <i>Caenorhabditis elegans</i>
<a href="#">Японская база данных ДНК (DDBJ)</a> <a href="#">Европейский архив нуклеотидов (ENA)</a> <a href="#">База данных ДНК и РНК (GenBank)</a>	Кодирующие последовательности
Основное хранилище генетических и молекулярных данных для насекомых семейства <i>Drosophilidae</i> (FlyBase)	Кодирующая последовательность для видов из семейства <i>Drosophilidae</i>
UniProtKB/TrEMBL	Автоматически курируемые последовательности белка, полученные из кодирующих последовательностей в базах данных нуклеотидных последовательностей
<a href="#">Японская база данных ДНК (DDBJ)</a> <a href="#">Европейский архив нуклеотидов (ENA)</a> <a href="#">База данных ДНК и РНК (GenBank)</a>	Кодирующие последовательности
<a href="#">Patent Offices in Europe, US and Japan (USPTO)</a>	Кодирующие последовательности, связанные с патентами из патентных ведомств
<a href="#">TROME</a>	Прогнозируемые аминокислотные последовательности
UniProtKB/Swiss-Prot	Обработанные вручную белковые последовательности, главным образом производные от TrEMBL
UniProtKB/TrEMBL	Автоматически курируемые последовательности белка, полученные из кодирующих последовательностей в базах данных нуклеотидных последовательностей
База информационных ресурсов для <i>Arabidopsis thaliana</i> (TAIR)	Кодирующие последовательности для <i>Arabidopsis thaliana</i>

База данных геномных и других биологических характеристик *Caenorhabditis elegans* (WormBase)

Кодирующие последовательности для нематоды *Caenorhabditis elegans*



## INSDC - nukleotidlarning ketma-ketligi bo'yicha xalqaro ma'lumotlar bazasi

(<http://www.insdc.org>)

*Nukleotidlarning ketma-ketligi bo'yicha xalqaro ma'lumotlar bazasi* -INSDC. DNK, RNK va aminokislotalar ketma-ketligini 30 yillar mobaynida o'z ichiga olgan ma'lumotlar bazalarini yig'ish va tarqatish bo'yicha kompleks sa'yl-harakatlarini boshqarishdan iborat (33-rasm). Biologik ma'lumotlar arxivlarining uzoq vaqtdan beri davom etayotgan global izlanishlardan biri, ommaviy nukleotidlarning ketma-ketligi haqidagi ma'lumotlarni to'playdi, saqlaydi va kerakli ma'lumotlar bilan ta'minlaydi.

- Mishima shahridagi Milliy Genetika Institutida Yaponiya DNK ma'lumotlar banki (DDBJ; <http://www.ddbj.nig.ac.jp/>);

- Buyuk Britaniyaning Xinxton shahridagi Yevropa Molekulyar Biologiya Laboratoriyasining Yevropa Bioinformatika Instituti- EMBL-EBI; <http://www.ebi.ac.uk/ena>.

- AQShning Biotexnologiya Axborot Milliy Markazi- NCBI; <http://www.ncbi.nlm.nih.gov/genbank/>, Bethesda shtatida, MD, AQSh.

INSDC ma'lumotlariga bepul va cheklanmagan kirishning birgalikdagi yagona siyosatiga ega. Ushbu siyosatga binoan, INSDC har kuni nukleotidlar ketma-ketligi va ular bilan bog'liq ma'lumotlarni to'playdi, saqlaydi, ta'minlaydi va almashadi.

INSDC xalqaro ma'lumotlar bazasida quyidagi yo'naltirilgan platformalari mavjud:

- Bioinformatika
- Biologik ma'lumotlar bazasi
- Yaponiya DNK ma'lumotlari banki
- Yevropa bioinformatika instituti
- Biologik ma'lumotlar bazalari ro'yxati
- Milliy Biotexnologiya Axborot markazi
- Sequence ma'lumotlar bazasi

INSDC xalqaro ma'lumotlar bazasida quyidagi bog'laniladigan URL platformalari mavjud:

- ›Rasmiy veb-sayt
- ›EMBL INSDC sayti



- › EMBL nukleotid bazasi
  - › Yaponiya DNK ma'lumotlari banki
  - › GenBank nukleotidini qidirish
- INSDC xalqaro ma'lumotlar bazasida quyidagi hujjatlar platformalari mavjud:
- › INSDC tomonidan boshqariladigan so'z birikmalari
  - › Genetik kod jadvallari\*
  - › INSDC holati to'g'risidagi hujjat
  - › Genom to'plamini topshirish uchun INSDC standartlari
  - › INSDSeq XML v1.5 dtd
  - › INSDC Advisor jurnalining muharrirlariga ochiq xat
  - › TPA topshirish bo'yicha ko'rsatmalar
  - › INSDSeq XML holati
  - › / inference qualifier lug'at bo'yicha tavsiyanoma
  - › / eksperiment kvalifikatori lug'at bo'yicha tavsiyanoma
  - › / Submitter\_seqid kvalifikatori tavsiyalari to'g'risidagi hujjat
  - › INSDC kelishilgan uslubiy kalit so'zlari
- \*Genetik kod jadvallari platformasidan ma'lumot (2016 yil tuzilgan):

*Achitqi zamburug'i mitoxondriyasidan (transl\_table = 3)*

Aminokislotalar FLLSSSSYY \*\*

CCWWTTTTPPPPHHQRRRRRIIMMTTTTNNKKSSRRVVVVAAAADDE

EGGGG

Boshlash kodlari ----- \*\* ----- MM -----

Baza1TTTTTTTTTTTTTTTTTTTTCCCCCCCCCCCCCAAAAAAAAAAAAAAAAAA

GGGGGGGGGGGGGGGGGGGG

Baza2TTTTCCCCAAAAGGGGTTTTCCCCAAAAGGGGTTTTCCCCAAA

AGGGGTTTTCCCCAAAAGGGG

Baza3TCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCA

GTCAGTCAGTCAGTCAGTCAG

Vaqt o'tishi bilan an'anaviy INSDC ketma-ketlik arxiviga kiritilgan nukleotidlar ketma-ketliklar sonining korrelyativ o'sishi kuzatildi. Jumladan, ommaviy ketma-ketlik ma'lumotlari-WGS hamda ommaviy bo'lmagan ya'ni Expressed Sequence Tag (EST), Genome Survey Sequence (GSS), Patent va Transkriptome Shotgun Assambleyasi (TSA) ketma-ketlik ma'lumotlari o'rin oladi. Keyinchalik (2013 yil) barcha ma'lumotlar BioSample dasturi yordamida optimallashtirilib umumlashtirildi. Ma'lumot o'rnida shuni aytish mumkinki, BioSample dasturi NCBI, DDBJ va EMBL ma'lumotlar bazasidan o'rin olgan.

Nukleotidlar ketma-ketligi bo'yicha xalqaro ma'lumotlar bazasi bilan ishlash siyosati.

6. INSDC o'zlarining ma'lumotlar bazalarida mavjud bo'lgan barcha

yozuvlarga bepul va cheklovsiz kirishning yagona siyosatiga ega. Dunyo miqyosidagi olimlar ushbu yozuvlarga eksperimentlarni rejalashtirish yoki har qanday tahlil yoki tanqidlarni nashr etish uchun kirishlari mumkin. Tegishli maqola, chop etilgan ilmiy adabiyotlardan foydalangan holda, olimlarning asl nusxalariga tayanib, asl taqdimnomaga tayanib beriladi.

7. INSDC ma'lumotlariga kirishni cheklaydigan, ushbu yozuvlardagi ma'lumotlardan foydalanishni cheklaydigan yoki ushbu yozuvlar asosida nashrlarning ayrim turlarini taqiqlaydigan yozuvlarga bayonnomalarni biriktirmaydi. Xususan, foydalanish bo'yicha cheklovlar yoki litsenziyalash talablari biron-bir ketma-ketlik ma'lumotlari yozuvlariga kiritilmaydi va hech kim tomonidan ma'lumotlar bazasini qayta taqsimlash yoki undan foydalanishda hech qanday cheklovlar yoki litsenzion to'lovlar bo'lmaydi.

8. INSDC-ma'lumotlar bazasiga taqdim etilgan barcha annotasiyalar ilmiy nashrlar sifatida doimiy ravishda ochiq bo'lib qoladi. Xatolarni tuzatish va mualliflar tomonidan yozuvlarning yangilanishi qabul qilinadi va noto'g'ri annotasiyalar bazaning keyingi nashrlaridan olib tashlanishi mumkin, ammo barchasi kirish raqami orqali doimiy ravishda saqlanib qoladi.

9. Taqdim etuvchilarga INSDC tomonidan qo'llab-quvvatlanadigan veb-saytlarda aks ettirilgan ma'lumotlar keng jamoatchilikka oshkor qilinishi tavsiya etiladi. Ma'lumotni topshirish huquqiga ega ekanliklarini aniqlash uchun yuboruvchilar javobgar.

10. Tahririyat nazorati va ba'zi ichki tekshiruvlaridan tashqari (masalan, INSD formatlaridan to'g'ri foydalanish va CDS yozuvlarida ko'rsatilgan kodlash hududlarining tarjimasini tekshirilganda), ma'lumotning sifati va aniqligi ma'lumotlar bazasi emas, balki taqdim etuvchi muallifning zimmasidadir. Ma'lumotlar bazalari iloji boricha sifatli manbaga erishish uchun ma'lumotlar bazasiga ma'lumot taqdim etuvchilari va foydalanuvchilari bilan doimiy aloqada bo'ladi.

INSDC xalqaro ma'lumotlar bazasining maslahatchi qo'mitalari.

*DDBJ maslahatchilari:*

- Sumio Sugano, Tibbiyot fanlari instituti, Tokio universiteti
- Ken Kurokawa, Yer-hayot ilmiy instituti, Tokio texnologiya instituti
- Kaoru Fukami-Kobayashi, Bioresource ma'lumot bo'limi, RIKEN

BioResource markazi

*EMBL maslahatchilari:*

- Antuan Danchin, AMAbiotics SAS, Evri, Frantsiya va CEA ilmiy maslahatchisi, Frantsiya
- Babis Savakis, Krit universiteti va IMBB-FORTH, Irakilion
- Jan Vaysenbax, Genoskop, Evri

- Mark Blaxter, GenePool Genomics fondi va Edinburg universiteti, Evolyutsiya biologiyasi instituti

GenBank (NCBI) maslahatchilari:

- Stiven Salzberg, Djons Xopkins nomidagi tibbiyot maktabi, Baltimor, MD, AQSh

- Rich Roberts, Yangi Angliya Biolabs, Beverli, MA, AQSh

- Valeri de Kresi-Lagard, Florida universiteti, Geynvill, FL, AQSh

**Protein bazalari turlari va ahamiyati.** Proteinlar bazasi oqsillar haqida bir yoki bir nechta ma'lumotlar to'plamidir. Ular tarkibiga oqsilning aminokislotalar ketma-ketligi, tuzilishi, shuningdek funktsiyalari kabi ma'lumotlar kiradi.

Protein ma'lumotlar bazalari turli xil gen bazalaridan DNK ketma-ketligini translyatsiya qilish orqali tuziladi va protein tarkibiy ma'lumotlarini o'z ichiga oladi. Ular muhim manba hisoblanadi, chunki oqsillar ko'p biologik funktsiyalarni bajaradi.

*Protein ma'lumotlar bazalarining ahamiyati:* Protein tuzilmalari, funktsiyalari va ayniqsa ketma-ketligi to'g'risida juda ko'p ma'lumotlar olinmoqda. Ma'lumotlar bazalarini qidirish ko'pincha yangi oqsilni o'rganishda katta ahamiyatga egadir.

### **PDB - protein ma'lumotlari banki**

Protein ma'lumotlari banki (PDB) 1971 yil oktyabr oyida PROTEIN DATA BANK Nature New Biologiyada Kembrij Kristallografik Ma'lumotlar Markazi, Buyuk Britaniya va Bruksaven Milliy Laboratoriyasi, AQShdagi qo'shma korxonada e'lon qilindi. Ushbu veb saytga odatda rentgen kristallografiyasi, NMR pektroskopiyasi va krioelektron mikroskopiya usullari orqali olingan, dunyoning turli mamlakatlari biolog va kimyogarlar tomonidan taqdim etilgan ma'lumotlar kirishi mumkin (PDBe, PDBj, RCSB, va BMRB). PDB butun dunyo bo'ylab oqsil ma'lumotlari banki deb nomlangan tashkilot tomonidan nazorat qilinadi. PDB global arxivini yagonalashtirish maqsadida bir harakat PDBe (Buyuk Britaniya), PDBj (Yaponiya), va BMRB (AQSh) bazalari hamkorlikda ish olib borishadi (35-rasm).

PDB arxivi har hafta yangilanadi va ma'lumotlarni foydalanuvchilarga bepul taqdim etiladi. Taqdim etiladigan xizmat turlari ko'p jumladan:

- Kristallografik ma'lumotlar bazasi



35-rasm. PDB – protein ma'lumot bazasining boshqa bazalar bilan a'loqasi.

- Protein tuzilishi
- Protein tarkibini bashorat qilish
- Protein tuzilishi ma'lumotlar bazasi
- PDBREPORT PDB tuzilishidagi barcha anomaliyalarni ko'rsatadi
- PDBsum - boshqa ma'lumotlar bazalaridan foydalanadi
- Proteopedia - bu molekulalarning 3D ensiklopediya tipida ko'rsatadi.

WORLDWIDE PDB PROTEIN DATA BANK

TEKSHIRISH ▾ CHO'KISH ▾ LUG'ATLAR ▾ HUJJATLAR ▾ VAZIFA KUCHLARI ▾ FTP ▾ STATISTIKA ▾ HAQIDA ▾ wwPDB Foundation

1971 yildan boshlab Protein Ma'lumotlar Bazasi (PDB) oqsillar, nuklein kislotalar va murakkab yig'ilishlarning 3D tuzilmalari haqidagi ma'lumotlarning yagona saqlanadigan ombori bo'lib xizmat qildi.

Butunjahon PDB tashkiloti PDB arxivini boshqaradi va PDB dunyo hamjamiyatiga bemaol va ochiq bo'lishini ta'minlaydi.

PDB tarixi va kelajak haqida ko'proq ma'lumot oling .



**Tekshiruv tarkibi**  
yoki tekshirish hisobotlarini ko'rish

**Omonat tarkibi**  
Barcha cho'kma manbalari

**Arxivni yuklab oling**  
Ko'rsatmalar

**Ko'rish va vazifa**

**Vizyon**  
Biologik makromolekulalar uchun ma'lumot va metadata ma'lumotlarini erkin kirishga imkon beradigan, birgalikda ishlaydigan Asosiy arxivlarini doimiy fanlar bo'yicha fundamental va amaliy tadqiqotlar va ta'limni ilgari surish uchun saqlab turing.

**Missiya**

- FAIR printsiplariga muvofiq wwPDB Core Archives-ni jamoat foydasi sifatida boshqaring.
- Dunyo bo'yicha Ma'lumot Omonatchilariga bepul to'lovlarni amalga oshirish, tekshirish, biokuratsiya va tiklash xizmatlari bilan ta'minlash.
- Foydalanish bo'yicha cheklolovsiz jamoat mulki bo'lgan biologik ma'lumotlarga umumiy kirishni ta'minlash.
- Global strukturaviy biologiya ma'lumotlarini arxivlash va almashish uchun tasdiqlangan ma'lumotlar standartlarini ishlab chiqish va ilgari surish.

**wwPDB a'zolari**

**Biologik magnit-rezonans ma'lumotlari banki**  BMRB  
Har qanday tajribadan NMR ma'lumotlarini to'playdi va turli xil makromolekulalar uchun belgilangan kimyoviy siljishlar, ulash doimiylik va cho'qqilar ro'yxatini oladi; vodorod almashinuv kurslari, pKa qiymatlari va gevseme parametrlari kabi olingan izohlarni o'z ichiga oladi.

**Evropadagi proteinlar banki**  EPPDB  
PDB-ning barcha yozuvlari, ko'plab qidirish va ko'rib chiqish vositalari, ilg'or xizmatlar, shu jumladan PDBaPISA, PDBeFold va PDBeMotif, NMR va EM tuzilmalarini vizualizatsiyalash va tekshirish, bioinformatistlar uchun vositalar to'g'risida boy ma'lumotlar.

**Protein ma'lumotlari banki**  PDBj  
Yaponiya  
Ko'plab tillarda, masalan, yapon, xitoy va koreys tillarida ko'rib chiqishni qo'llab-quvvatlaydi; SeSAW funksional yoki evolyutsion saqlanib qolgan motivlarni aniqlash va izohlash orqali ketma-ketlik va tuzilish o'xshashligini, bioinformatistlar uchun vositalarni va boshqalarni aniqlaydi.

**Strukturaviy bioinformatika oqsillari ma'lumotlar banki bo'yicha ilmiy-tadqiqot laboratoriyasi**  PDB  
Makromolekulalar va ligandlar, jadval jadvallari, vizualizatsiyaning ixtisoslashtirilgan vositalari, ketma-ketlik tuzilishini taqqoslash, RCSB PDB Mobile, O'zining molekulasi va PDB-101-dagi boshqa ta'lim manbalarini oddiy va chuqurroq qidirish.

**wwPDB manbalari**

**Ma'lumot lug'atlari**

- Makromolekulyar lug'at (PDBx / mmCIF)
- Kichik molekula lug'ati (CCD)
- Peptidga o'xshash antibiotik va ingibitor molekullari (BIRD)

**Biokuratsiya**

- Protседuralar va qoidalar
- Muvofiqlik va aniqlikni yaxshilash

**Jamiyatni jalb qilish: Ishchi kuchlari va ishchi guruhlari**

- Tekshiruv ish kuchlari (rentgen, NMR, 3DEM)
- Kichik burchakka sochish vazifalari guruhi
- PDBx / mmCIF ishchi guruhi
- Gibrid / integratsiyalashgan usullar
- Ligandni tekshirish bo'yicha seminar

**PDB ma'lumotlarining o'sishi va foydalanish statistikasi**

- Depozitlar: ma'lumot markazi, yil bo'yicha va omonatchining joylashuvi bo'yicha
- Yuklanishlar: barcha yozuvlar uchun yiliga

**Seminarlar va simpozialar**

- O'tgan uchrashuv va tadbirlarning qisqacha mazmuni va taqdimotlari

**Jurnallar uchun ma'lumot**

- Siyosatlar, protseduralar, noshirlar bilan muvofiqlashtirish va muallifarga ko'rsatmalar

**WwPDB-ga ishora qiling:**  
*Tabiatning strukturaviy biologiyasi* 10 , 980 (2003)  
doi: 10.1038 / nsb1203-980

**PDB arxivini keltiring:**  
Nuklein kislotalarini tadqiq qilish (2019), doi: 10.1093 / nar / gky949  
Boshqa nashrlar

**Yangiliklar va e'lonlar**

**04.09.2010**

- Yangilangan PDB-Dev veb-sayti

Endi yangi va takomillashtirilgan PDB-Dev veb - sayti mavjud. PDB-Dev prototip tizimi bo'lib, u wwPDB sheriklariga integral tuzilmalarni arxivlashda talablarni tushunishga yordam beradi.

**Ko'proq o'qing**

**04.07.07**

- EMDB xaritasini kiritish jarayonini soddalashtirish

15-apreldan boshlab, Elektron Mikroskopiya Ma'lumotlar Banki (EMDB) asosiy xarita va tegishli fayllarni chiqarishdan oldin xaritaga kirish metadata (XML sarlavhali fayllar deb ataladi) ni chiqarish amaliyotini tugatadi.

**Ko'proq o'qing**

**02/25/2020**

- 2020 yilgacha PDBda uglevodlarni yaxshilash



Ushbu tuzilmalarni va ularning murakkab kimyosini qidirish qobiliyatini yaxshilash uchun (masalan, stereo-izomerlar, anomerik konfiguratsiyalar, tarmoqli zanjirlar) wwPDB yangi qayta tiklash loyihasiga kirishdi.

**Ko'proq o'qing**

**Barcha yangiliklar**

Arxivni yuklab oling  
RCSB PDB ftp | PDBe ftp | PDBj ftp ko'rsatmalar

WwPDB-ga ishora qiling:  
*Tabiatning strukturaviy biologiyasi* 10 , 980 (2003)  
doi: 10.1038 / nsb1203-980  
Boshqa nashrlar

36-rasm. PDB-Protein ma'lumotlari bankining asosiy oynasi.

## PIR -protein axborot resursi

PIR - protein axborot resursi bazasi, Jorjtaun universiteti tibbiyot markazi (GUMC) tomonidan bioinformatikani qo'llab-quvvatlash maqsadida genom va



proteom tadqiqot natijalari va oqsil ketma-ketliklarini taqdim etadi. *PIR -Protein axborot resursi markazining asosiy oynasi.*

PIR 1984 yilda Uinona Barker va Robert Ledlilar boshchiligida Milliy

**PIR** A UniProt CONSORTIUM MEMBER  
Protein Information Resource

About PIR | Resources | Search/Analysis | Download | Support

**INTEGRATED PROTEIN INFORMATICS RESOURCE FOR GENOMIC, PROTEOMIC AND SYSTEMS BIOLOGY RESEARCH**

**UniProt** Universal Protein Resurslari (UniProt) ilmiy jamoalarga oqsillar ketma-ketligi va funksional ma'lumotlar uchun yagona, markazlashtirilgan, yakolatli manba beradi.  
UniProtKB | UniRef | UniParc | Joriy nashr: 2020\_02

**PRO** Protein Ontology

- Protein ob'ektlarini tavsiflash va o'zaro bog'liqlik
- [PRO-ni ko'rib chiqish](#)
- [RACE-PRO bilan izohlang](#)

\* [PRO hisobotining namunasi](#) \*

**iPTMnet** Integrated Protein PTM Resource

- PTM fermenti-substrat-sayt aloqalari
- PTM tarmoqlarini vizual tahlil qilish, o'zaro so'zlashuvlar va saqlash
- [Partiya olish](#)

\* [IPTMnet hisobotining namunasi](#) \*

**iProLINK** Literature Information & Knowledge

- Matnni yaratish vositalariga va izohlangan korporatsiyalarga kirish
- [RLIMS-P](#) kinaz, substrat va saytni qazib olish
- [miRtex](#) miRNA / maqsadli ma'lumotni ajratib olish

\* [RLIMS-P hisobotining namunasi](#) \*

**BOSHQA RESURS**

- [Malumot Proteomlari](#)
- [i ProClass](#)
- [i ProXpress](#)
- [QONUN](#)

**XALQNING qidirish** MA'LUMOT: UniProtKB

Use single letter amino acid code

**TEXT qidirish** MA'LUMOT: i ProClass

Bioinformatika va kompyuter biologiyasi bo'yicha magistrlik dasturlari:

- [Jorjtaun universitetida](#)
- [magistrlik dissertatsiyasi, magistrlik, PSM va Delever universitetida magistrlik dasturlari.](#)

Uy | PIR haqida | Ma'lumotlar bazalari | Qidiruv / tahlil | Yuklab olish | Qo'llab-quvvatlash | SITE Xaritasi | FOYDALANISH S

© 2018 Protein ma'lumotlari manbai

Biotibbiy Tadqiqotlar Fondi (*NBRF*) tomonidan tadqiqotchilar va mijozlarga oqsillar ketma-ketligi ma'lumotlarini aniqlash va izohlashda yordam beradigan manba sifatida tashkil etilgan. Bungacha *NBRF* 1964-1974 yillarda Margaret Dayhoff tahriri ostida nashr etilgan "Proteinlar ketma-ketligi va tuzilishi atlas" da birinchi makromolekulyar ketma-ketlik to'plamini tuzgan. Dayxof va uning tadqiqot guruhi

oqsillar ketma-ketligini taqqoslash, ketma-ketliklar orasidagi masofa va ketma-ketlikni aniqlash hamda oqsillar ketma-ketligi evolyutsion tarixini aniqlash uchun kompyuter usullarini ishlab chiqishda kashf etdilar.

PIR boshqa ma'lumotlar bazalari jumladan, EBI (Yevropa Bioinformatika Instituti), oqsillar ketma-ketligi va funksiyalarining markaziy manbai bo'lgan UniProt (Birlashgan Protein ma'lumotlar bazasi) va SIB (Shveysariya Bioinformatika Instituti) bilan birgalikda ish yuritishadi.

PIR o'zida oqsillar ketma-ketligi, oilalari, fermentlar va fermentativ yo'llar, molekulyar tuzilmalar va strukturalar ularning sinflari, genlar va genomlar, adabiyot va taksonomiyalarning ko'p sonli molekulyar ma'lumotlar bazalariga keng ko'lamli murojaatlarni va havolalarni taqdim etadi (1-rasm). Ushbu sayt ma'lumotlarini UniProt (Birlashgan Protein ma'lumotlar bazasi) resurslari bilan dastlabki tahlilini o'tkazgan holda taqdim etib, tahlil qilishda ham UniProt dasturlaridan foydalanadi.

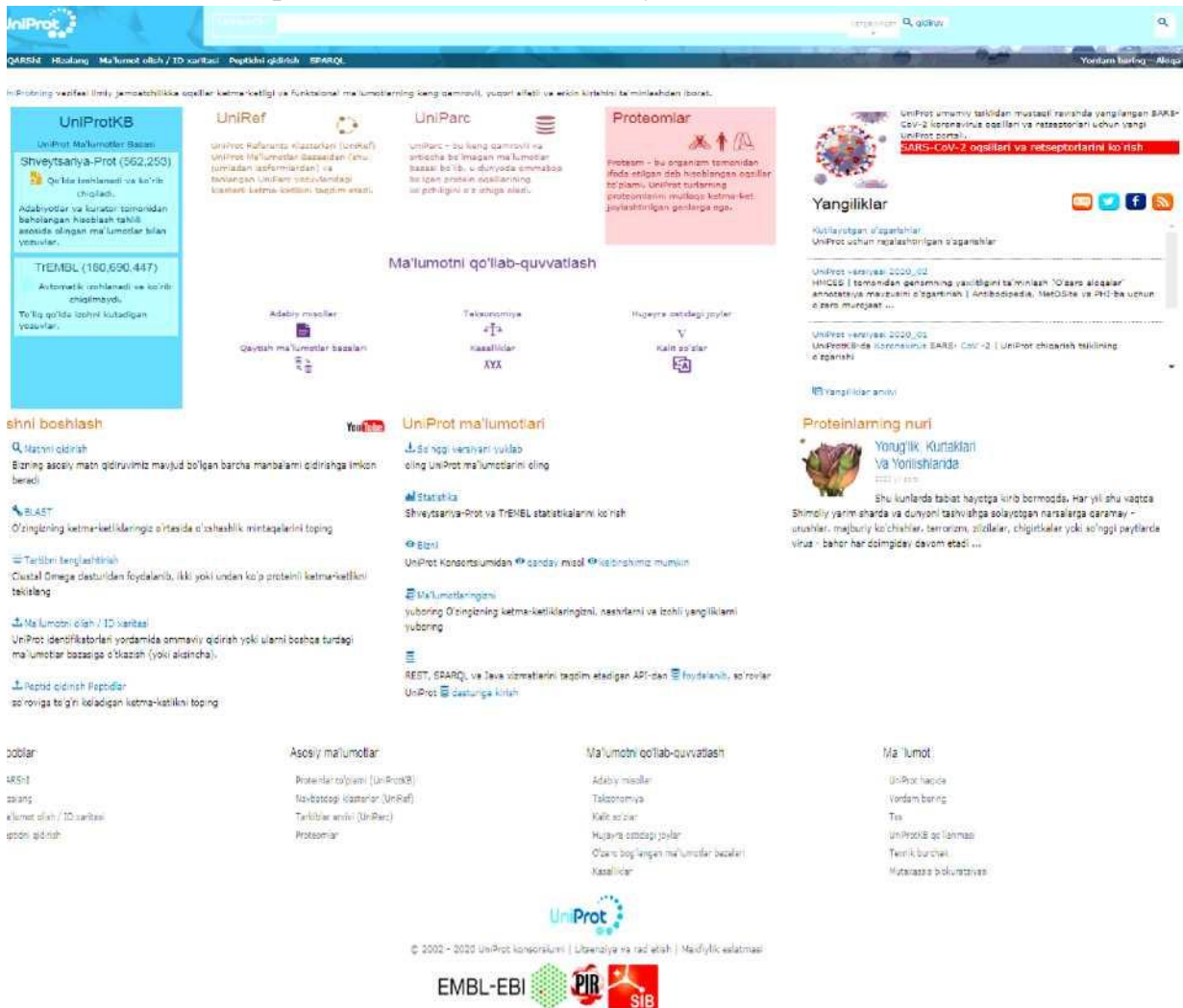
### **UniProt ma'lumotlar bazasi**

**UniProt** - bu oqsillar ketma-ketligi va funksional ma'lumotlarning erkin kirishi mumkin bo'lgan ma'lumotlar bazasi bo'lib, ularning ko'plari genomni tartiblash loyihalaridan olinadi (38-rasm). Unda tadqiqotlar natijasida olingan oqsillarning biologik funksiyasi haqida juda ko'p ma'lumotlar mavjud. Uni bir nechta Yevropa bioinformatika tashkilotlari va Vashington, Kolumbiya okrugi (AQSh) dan tashkil topgan UniProt konsorsiumi qo'llab- quvvatlaydi. UniProt keng doirada boshqa bazalar bilan hamkorlikda ish olib boradi. Jumladan:

- ✓ INSDC
- ✓ EMBL
- ✓ GenBank (NCBI)
- ✓ DDBJ
- ✓ Ansambl
- ✓ Yevropa patent idorasi (EPO)
- ✓ FlyBase: Drosophilidae hasharotlar oilasi uchun genetik va molekulyar ma'lumotlarning asosiy ombori.
- ✓ H-Invitation ma'lumotlar bazasi (H-Inv)
- ✓ Proteinlarning xalqaro indeksi (IPI)
- ✓ Yaponiya Patent idorasi (JPO)
- ✓ Protein haqida ma'lumot manbasi (PIR-PSD)
- ✓ Protein ma'lumotlari banki (PDB)
- ✓ Oqsillarni tadqiqot qilish fondi (PRF) <sup>[18]</sup>
- ✓ Saccharomyces Genome Database (SGD)



- ✓ Arabidopsis haqida ma'lumot manbai (TAIR)
- ✓ AQSh Patent idorasi (USPTO)
- ✓ Shveysariya-Prot (SWISS-Prot)
- ✓ TrEMBL
- ✓ Umurtqali va Genomli Annotatsiya Ma'lumotlar Bazasi (VEGA)



### UniProt ma'lumotlar bazasining asosiy oynasi.

U genom loyihalari natijasida olingan ma'lumotlar oqimining ko'payishiga javoban joriy qilingan, chunki UniProtKB / Shveysariya-Protning vaqt va mehnat sarf qiladigan qo'l yozuvlaridan ko'ra EMBL-Bank / GenBank / DDBJ nukleotidlar ketma-ketlik ma'lumotlar bazasidagi kodlangan ketma-ket tarjimalarini avtomatik ravishda qayta ishlaladi.

*UniProt Archive (UniParc)* - bu oqsillar ketma-ketligining asosiy, hamma uchun ochiq bo'lgan to'liq ma'lumotlar bazasi. Bazada proteinlarning bir necha xil turini topish mumkin. Oqsilning har bir ketma-ketligiga barqaror va noyob identifikator (UPI) va proteinni turli xil manbalar bazalarida aniqlash imkonini beradi. Bundan tashqari ma'lumotlar bazasidagi protein ketma-ketliklari o'zgaranda, bu o'zgarishlar UniParc tomonidan kuzatiladi va barcha o'zgarishlar tarixi arxivlanadi.

*UniProt Referents Klasterlari (UniRef)* - yuqoridagi UniProtKB (Swiss-Prot qismlari bilan), UniProtKB (TrEMBL-qismlari bilan) va UniParcda mavjud xususiy va umumiy oqsillar ketma-ketligi yagona UniRef yozuviga birlashtiradi.



## SWISS-PROT-proteinlar ketma-ketligi ma'lumotlar bazasi

Genom loyihalaridan ketma-ket ma'lumotlar bazalariga suvli ma'lumotlar oqimi oshib borishi sababli ma'lumotlar bazasini izohlashda bir qator qiyinchiliklarga duch kelinmoqda. SWISS-PROT-da yuqori sifatli ketma-ketlik va izohlarning saqlanishi bu kabi muammolarning yechimini topish imkonini beradi (39-rasm)

SWISS-PROT ma'lumotlar bazasidagi oqsillar ketma-ketligi boshqa ma'lumotlar bazasidan uchta aniq mezon bilan ajratib turadi: (I) aniq izohlar, (II) minimal ma'lumotlar va (III) boshqa ma'lumotlar bazalari bilan to'laqonli integratsiya.



- 
- [SWISS-PROT Bosh sahifa](#)
  - [Ma'lumot](#)
  - [Kirish](#)
  - [Topshiriqlar](#)
  - [Asboblari](#)
  - [FTP](#)
  - [Guruh haqida ma'lumot](#)
  - [Hujjatlar](#)
  - [Aloqa](#)

SWISS-PROT Ma'lumotlar bazasiga kirish

SWISS-PROT ma'lumotlar bazasiga quyidagi qidiruv tizimi orqali kirish mumkin

[Qidirmoq](#) Proteinlar uchun :

Qidiruv mexanizmi [SRS](#) tomonidan quvvatlanadi, ushbu izlash vositasi yanada murakkab va / yoki bir nechta ma'lumotlar bazasi so'rovlari uchun ishlatilishi mumkin.

Kirishning boshqa usullari SWISS-PROT ma'lumotlar bazasiga quyidagilar yordamida kirish mumkin.

### SWISS-PROT ma'lumotlar asosiy oynasi

SWISS-PROT-da ma'lumotlarning ikki sinfini ajratish mumkin: asosiy ma'lumotlar va izohlar. Har bir ketma-ketlik uchun asosiy ma'lumotlar ketma-ketlik ma'lumotlaridan iborat; iqtibos ma'lumotlari (bibliografik havolalar) va taksonomik ma'lumotlar (oqsilning biologik manbasi tavsifi), izoh esa quyidagi elementlarning tavsifidan iborat:

- Oqsilning funksiyalari

Post-translational modifikatsiyalar, masalan, uglevodlar, fosforillanish, atsetilatsiya, GPI-langar va boshqalar.

- Domenlar va saytlar. Masalan, kaltsiyni bog'laydigan mintaqalar, ATP-ulanish joylari, sink barmoqlari, gomeobokslar, CH<sub>2</sub> va CH<sub>3</sub> domenlari va boshqalar.
- Ikkilamchi tuzilish. Masalan, alfa spiral, beta zanjir va boshqalar.
- To'rtlamchi tuzilish. Masalan, homodimer, heterotrimer va boshqalar.
- Boshqa oqsillarga o'xshashliklar
- Oqsil yetishmovchiligi bilan bog'liq kasalliklar
- Tarkibiy nizolar, variantlar va boshqalar.

### **Nazorat savollari**

1. NCBI nima?
2. PDB nima?
3. PIR nima?
4. Qaysi Bibliografik ma'lumotlar bazalarini bilasiz.
5. Biologik ketma-ketliklarning matnli ma'lumotlari qanday yaratiladi?
6. Protein ketma-ketligi ma'lumotlar bazalarini aytin.
- 7.

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
2. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

1. (<http://www.insdc.org>)
2. (<https://www.ncbi.nlm.nih.gov/genbank/>)

## 2-amaliy mashg'ulot

**Mavzu:** Biologik makromolekulalarning birlamchi strukturasi (matn ko'rinishidagi) haqidagi ma'lumotlarni saqlovchi ma'lumotlar bazalari bilan ishlash.

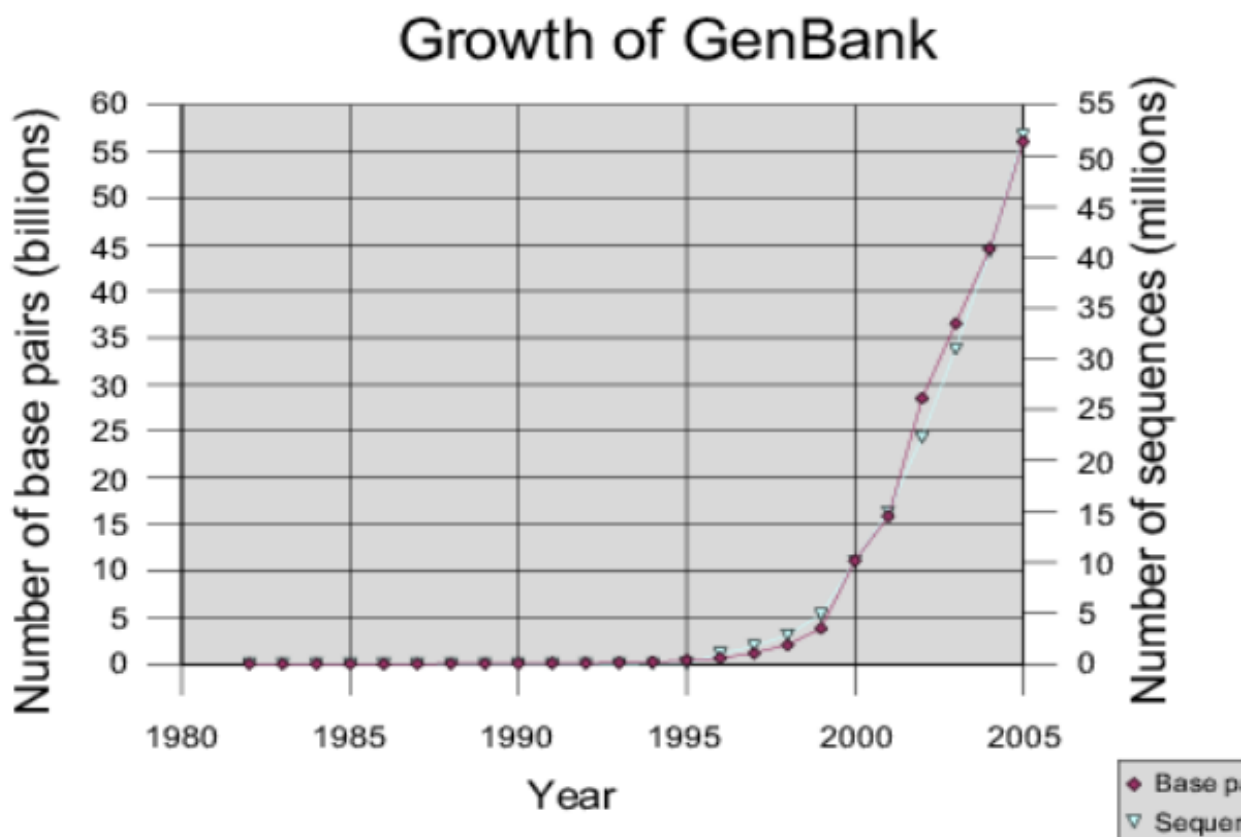
**Dars maqsadi:** Talabalarga biologik makromolekulalarning birlamchi strukturasi (matn ko'rinishidagi) haqidagi ma'lumotlarni saqlovchi GenBank, NCBI, MMDB kabi ma'lumotlar bazalari bilan tanishishtirish.

**Kerakli jihoz va preparatlar:** Kampyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** GenBankning ketma-ketlik bazasi- bu ochiq kirish, barcha ommaga taqdim etiladigan nukleotidlar ketma-ketligi va ularning proteinli tarjimalarining izohli to'plami. Ushbu ma'lumotlar bazasi nukleotidlar ketma-ketligi bo'yicha Xalqaro ma'lumotlar bazasi hamkorligi (INSDC)tarkibida Milliy biotexnologiya ma'lumotlari markazi (NCBI; AQShdagi Sog'liqni Saqlash milliy institutlarining bir qismi) tomonidan ishlab chiqiladi va saqlanadi.

GenBank va uning hamkorlari butun dunyo bo'ylab laboratoriyalarda 100000 dan ortiq alohida organizmlardan ishlab chiqariladigan ketma-ketlikni oladilar. Ma'lumotlar bazasi 1982-yilda Valter Gad va Los Alamos milliy laboratoriyasi tomonidan tashkil etilgan. GenBank biologik sohalarda tadqiqotlar uchun muhim ma'lumotlar bazasiga aylandi va so'nggi yillarda har 18 oyda ikki baravarga ko'paydi.

2013-yil fevral oyida ishlab chiqarilgan 194 versiyasida 162 mlndan ortiq ketma-ketlikda 150 mlrddan ortiq nukleotid asoslari mavjud edi.



GenBankdagi ketma-ketliklarning yillar hisobida ko'payishi

GenBankga faqat asl tartiblarni topshirish mumkin. Tartibot taqdim etilgandan so'ng, GenBank xodimlari ma'lumotlarning aslligini tekshiradi va ketma-ketlik uchun

kirish raqamini beradi va sifatni tekshirishni amalga oshiradi Keyinchalik taqdimnomalar ommaviy ma'lumotlar bazasiga chiqariladi, u yerda yozuvlarni Entrez tomonidan yuklab olinishi mumkin.

NCBI Resources How To Sign in to NCBI

GenBank Nucleotide Search

GenBank Submit Genomes WGS Metagenomes TPA TSA INSDC Other

Data regarding the SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2, 2019-nCoV) outbreak sequences can be found in [GenBank/SRA](#), the [NCBI Virus](#) resource, and a specialized [BLAST](#) page that searches Betacoronavirus sequences.

### GenBank Overview

#### What is GenBank?

GenBank<sup>®</sup> is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research, 2013 Jan; 41\(D1\):D36-42](#)). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth statistics for both the traditional GenBank divisions and the WGS division are available from each release. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An [annotated sample GenBank record](#) for a *Saccharomyces cerevisiae* gene demonstrates many of the features of the GenBank flat file format.

#### Access to GenBank

There are several ways to search and retrieve data from GenBank.

- Search GenBank for sequence identifiers and annotations with [Entrez Nucleotide](#).
- Search and align GenBank sequences to a query sequence using [BLAST](#) (Basic Local Alignment Search Tool). BLAST searches CoreNucleotide, dbEST, and dbGSS independently; see [BLAST info](#) for more information about the numerous BLAST databases.
- Search, link, and download sequences programatically using [NCBI e-utilities](#).
- The ASN.1 and flatfile formats are available at NCBI's anonymous FTP server: [ftp://ftp.ncbi.nlm.nih.gov/ncbi-asn1](#) and [ftp://ftp.ncbi.nlm.nih.gov/genbank](#).

#### GenBank Data Usage

### GenBank Resources

- [GenBank Home](#)
- [Submission Types](#)
- [Submission Tools](#)
- [Search GenBank](#)
- [Update GenBank Records](#)

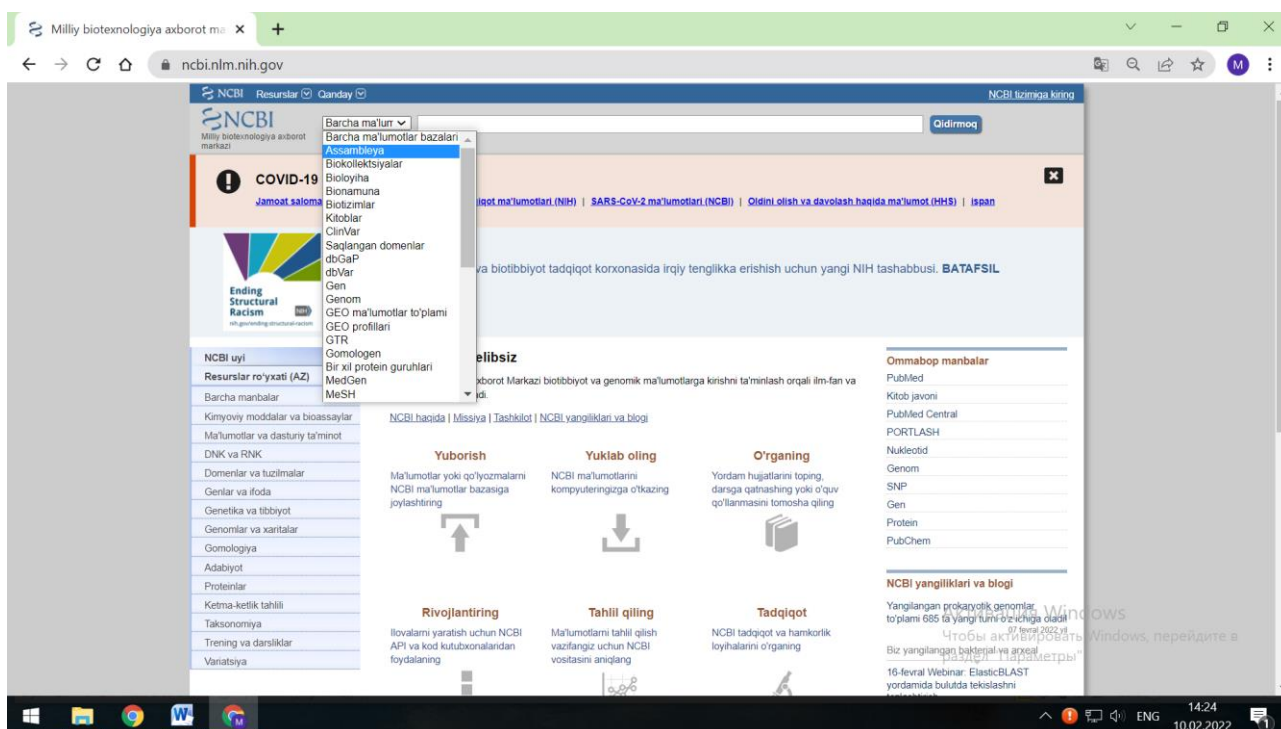
GenBankning eng katta afzalligi ochiq kirish formati bo'lib, u yagona formatda markazlashtirilgan omborga ega bo'lishga imkon beradi. Laboratoriyalar tomonidan yaratilgan juda katta miqdordagi ma'lumotlar (mikroarraylar va mikroRNK massivlari kabi) tadqiqot maqolasida nashr etilmaydi. Biroq, GenBankga joylashtirilgan va yuklangan ma'lumotlar jurnallarning veb-saytlariga ulanishi mumkin va havolalar maqolalarning bosma versiyalarida ham taqdim etilishi mumkin.

## NCBI

Biotexnologiya axborotlashtirish Milliy markazi ( NCBI ) hisoblanib Amerika Qo'shma Shtatlari Tibbiyot Milliy kutubxonasi (NLM), ning ma'lum bir filiali Milliy Sog'liqni saqlash institutlari (NIH) ning bir qismidir. Amerika Qo'shma Shtatlari hukumati tomonidan tasdiqlangan va moliyalashtirilgan. NCBI Merilend shtatining Bethesda shahrida joylashgan bo'lib , 1988 yilda AQSh Kongressi a'zosi Klod Pepper homiyligidagi qonunchilik asosida tashkil etilgan.



NCBI biotexnologiya va biomeditsinaga tegishli bir qator ma'lumotlar bazalarini o'z ichiga oladi va bioinformatika vositalari va xizmatlari uchun muhim manba hisoblanadi. Asosiy ma'lumotlar bazalariga DNK ketma-ketligi uchun GenBank va biotibbiy adabiyotlar uchun bibliografik ma'lumotlar bazasi PubMed kiradi . Boshqa ma'lumotlar bazalariga NCBI Epigenomics ma'lumotlar bazasi kiradi. Ushbu ma'lumotlar bazalarining barchasini **Entrez** qidiruv tizimi orqali onlayn kuzatish mumkin .



NCBI oynasi

## NCBI biologik ketma-ketliklarni Tahrirlash

NCBI **Bookshelf** - bu tanlangan biotibbiyot kitoblarining erkin foydalanish mumkin bo'lgan, yuklab olinadigan, onlayn versiyalari to'plami mavjud. Kitob javoni molekulyar biologiya , biokimyo , hujayra biologiyasi , genetika , mikrobiologiya , molekulyar va hujayra nuqtai nazaridan kasallik holatlari, tadqiqot usullari va virusologiya kabi keng mavzularni qamrab oladi . Ba'zi kitoblar avval nashr etilgan kitoblarning onlayn versiyalari, boshqalari, masalan, **COFFE BREAK** dasturi NCBI xodimlari tomonidan yozilgan va tahrirlangan. Kitob

javoni - bu Entrez PubMed repositoriyasining ko'rib chiqilgan nashri uchun to'ldiruvchidir. Bunda Bookshelf mazmuni rivojlanayotgan ta'lim yo'nalishlari bo'yicha belgilangan istiqbollarni va hisobot qilingan tadqiqotning bir-biridan farq qiladigan alohida qismlarini tashkil qilish mumkin bo'lgan kontekstni taqdim etadi.

## MMDB

Molekulyar modellashtirish ma'lumotlar bazasi (MMDB) eksperimental ravishda aniqlangan uch o'lchovli biomolekulyar tuzilmalar ma'lumotlar bazasi bo'lib, "Entrez Structure" ma'lumotlar bazasi deb ham yuritiladi. Bu nazariy modellardan tashqari RCSB Oqsil Ma'lumotlar Bankidan (PDB) olingan uch o'lchovli tuzilmalar to'plami hisoblanadi.

## MMDB oynasi

Shunday qilib biologik makromolekulalarning birlamchi strukturasi (matn ko'rinishidagi) haqidagi ma'lumotlarni saqlovchi ma'lumotlar bazalari bioinfatikada muhim hisoblanib, ularni strukturaviy va fazoviy ko'rinishlarini qayta ishlashda muhim ko'rsatkich hisoblanadi.

## Nazorat savollari

1. GenBank nima?
2. MMDB nima?



3. GenBankda nima saqlanadi?
4. MMDB qanday moddalar bilan ishlaydi?
5. NCBI nima?

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
2. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

1. [https://www.ncbi.nlm.nih.gov/Structure/MMDB/docs/mmdb\\_help.html](https://www.ncbi.nlm.nih.gov/Structure/MMDB/docs/mmdb_help.html)
2. (<https://www.ncbi.nlm.nih.gov/genbank/>)

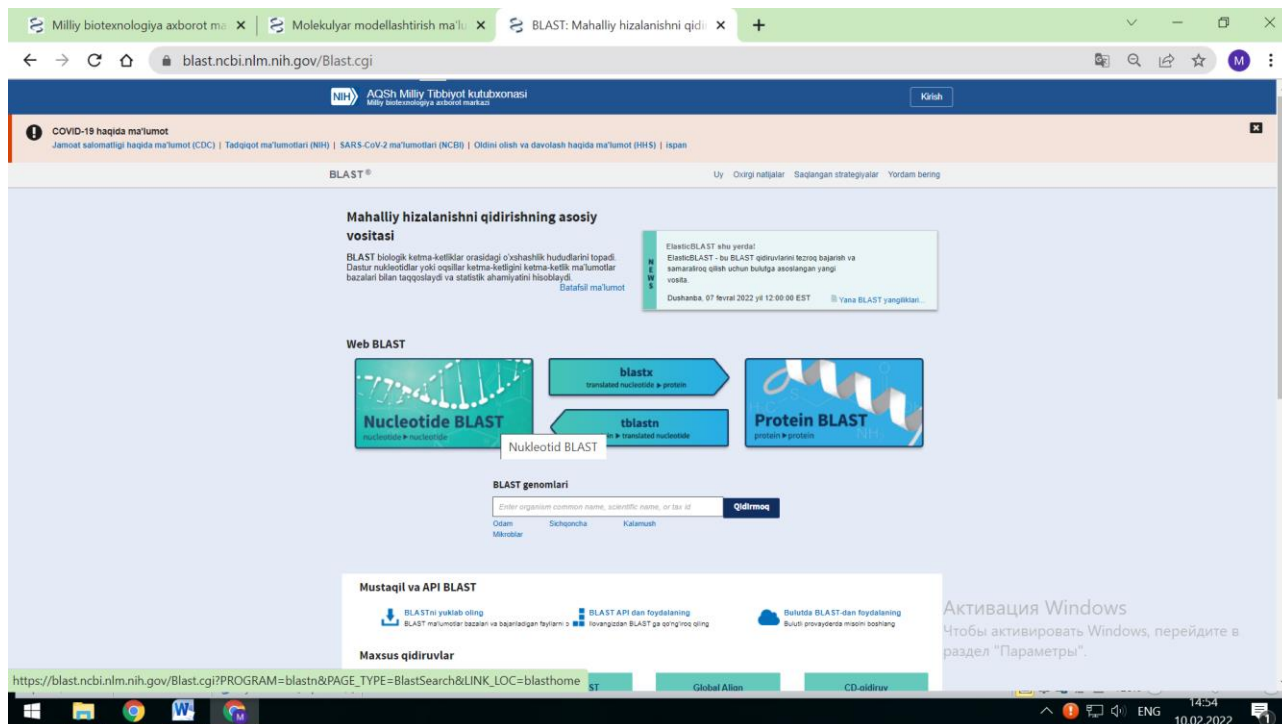
### **3- Amaliy mashg'ulot:**

**Mavzu:** BLAST dasturi yordamida nukleotid ketma-ketliklarini solishtirish.

**Dars maqsadi:** Talabalarga BLAST dasturi yordamida nukleotid ketma-ketliklarini solishtirish ni o'rgatish.

**Kerakli jihoz va preparatlar:** Kampyuter va uning dasturlari, internet tarmog'i, videoprektor.

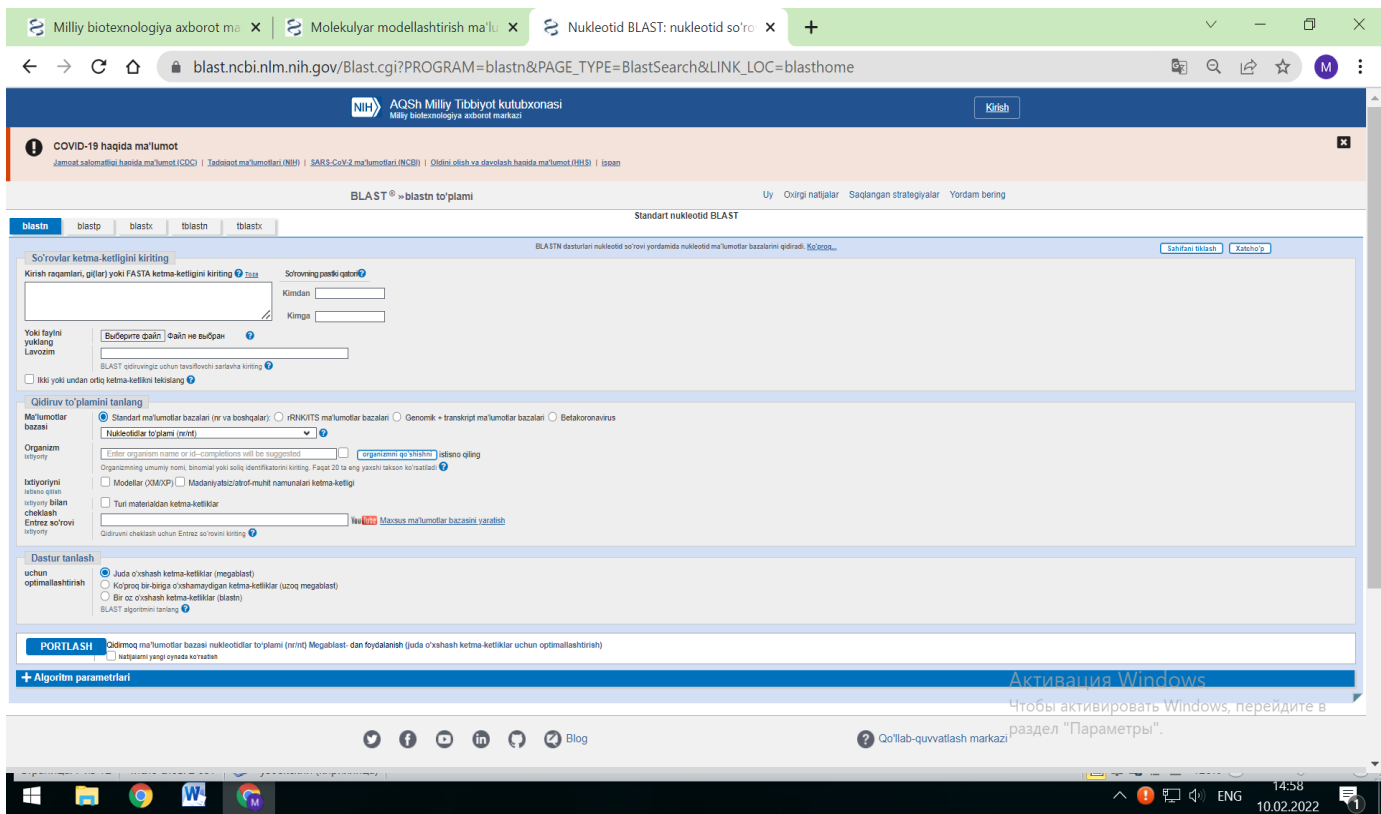
**Ishning mazmuni:** BLAST (Berkeley Lazy Abstraction Software Verification Tool) - bu C tilini namunaviy tekshirish dasturi. BLAST (Basic Basic Alignment Search Tool) - asosiy tuzilish (ketma-ketlik) yoki uning bo'lagi ma'lum bo'lgan oqsillar yoki nuklein kislotalarning homologlarini qidirishda foydalaniladigan kompyuter dasturlari oilasi. BLAST-dan foydalanib, tadqiqotchi o'z ketma-ketligini ma'lumotlar bazasidagi ketma-ketliklar bilan taqqoslashi va taxmin qilinayotgan homologlarning ketma-ketligini topishi mumkin. Bu molekulyar biologlar, bioinformatika, sistematikalar uchun eng muhim vositadir.



## Blast oynasi

**BLAST nukleotidlarining ketma-ketligi dasturi.** Belgilanmaganga o'xshash oqsillarni kodlovchi genlarning aniqlanmagan genomlarini qidiring. *E. coli* GlmS ga o'xshash oqsillarni kodlovchi genlarni qidirish amalga oshirildi. Qidiruv quyidagi organizmlarning aniqlanmagan genamlari bo'yicha olib borildi: *Pseudomonas aeruginosa*, *Pasteurella multocida* va *Vibrio* vabo. NCBI Entrez taksonomiyasining ma'lumotlariga ko'ra, bu organizmlarning barchasi *E. coli*ning juda yaqin qarindoshlari bo'lib, ular bir xil Gammaproteobakteriyalar sinfiga kiradi.

Format dasturi yordamida qidiruvni amalga oshirishdan oldin har bir genom uchun indeksli fayllar, shuningdek uchta genom uchun indeksli fayllar alohida yaratildi. Ikkala holatda ham dastur nhr, nin va nsq kengaytmali uchta faylni yaratdi. Nsq kengaytmali fayl (nukleotidlar ketma-ketligi) 4 marta siqilgan nukleotidlar ketma-ketligi to'g'risidagi ma'lumotlarni o'z ichiga oladi. Ushbu fayl boshqa ikkalasiga qaraganda ancha katta. Nhr va ninali fayllar (nukleotid indekslari) yordamchi bo'lib, mos ravishda to'liq genomni tashkil etuvchi yozuvlar sarlavhalarini va indekslarning ro'yxatini o'z ichiga oladi (formatdb dasturi uchun hujjatlarni ko'ring). Qidiruvni amalga oshirish uchun BLAST to'plamidagi TBLASTN dasturidan foydalanildi, u oltita freymdagi tarjima mahsuloti kirish oqsillari ketma-ketligiga o'xshash nukleotidlar ketma-ketliklarini aniqlash uchun ishlab chiqilgan.



## Nukleotid Blast oynasi

Elektron qiymatning chegaraviy qiymati 0,01 ga teng tanlangan. Qidiruv natijalari jadvalda keltirilgan.

Nukleotidning ketma-ketligini taqqoslaganda (2) formula bilan aniqlanadi. Uzunligi  $m$  o'rganilgan ketma-ketlikni ko'plab ma'lumotlar bazalari ketma-ketligi bilan taqqoslash ikki nuqtaga asoslanishi mumkin. Birinchi nuqta, ma'lumotlar bazasining barcha ketma-ketliklari o'rganilayotganga o'xshashdir. Bu ma'lumotlar bazasida mavjud bo'lgan qisqa ketma-ketlik bilan hizalanish uchun  $E$  qiymatini uzun ketma-ketlik bilan tekislash uchun  $E$  qiymatiga tenglashtirish kerakligini anglatadi. Ma'lumotlar bazasidan  $E$  qiymatini hisoblash uchun olingan  $E$  qiymatini undagi ketma-ketliklar soniga juft-juft taqqoslash orqali ko'paytirish kerak. Ikkinchi nuqta, o'rganilgan ketma-ketlik uzoq ketma-ketliklarga qaraganda qisqaroqroqqa o'xshaydi, chunki ikkinchisi ko'pincha turli qismlardan iborat (ko'p oqsillar domenlardan iborat). Agar o'xshashlik ehtimolligi ketma-ketlik uzunligiga mutanosib deb hisoblasak,  $n$  uzunlikdagi ma'lumotlar bazasi uchun  $E$  ning juft qiymatini  $N / n$  ga ko'paytirish kerak, bu erda  $N$  - bazadagi aminokislotalar yoki nukleotidlarning umumiy uzunligi. BLAST dasturlari asosan ushbu yondashuvdan ma'lumotlar bazasidan  $E$  qiymatlarini hisoblashda foydalanadilar.

Nazariy jihatdan, mahalliy moslashtirish har qanday nukleotid yoki aminokislotalarning hizalanadigan ketma-ketligidan boshlanishi mumkin. Biroq, GES, qoida tariqasida, ketma-ketlikning chetiga (boshiga yoki oxiriga) yaqin boshlamaydi.

Bunday chekka effektni tuzatish uchun ketma-ketlikning samarali uzunligini hisoblash kerak. 200 dan ortiq qoldiq bo'lsa, chekka effekti neytrallanadi.

BLASTN AC-ni chiqarishda bu yozuv AE004967. AE004967 yozuvi 2006 yil 12 iyulda AE004091 bilan almashtirildi. Shubhasiz, *P. aeruginosa* genom yozuvlari EMBL-ning oldingi versiyalaridan olingan. EMBL-ning SRS-qidiruvi EMBL-ning 87-versiyasida AE004967-ni topdi. CDS koordinata qiymatlari topilgan yozuvdan olinadi.

Query: 1460 tgaagagatctttacattcacgctgaagcctacgctgctggcgaactgaaacacggtc 1519

||||| ||||| ||||| || || ||||| || || || || |

Sbjct: 4535 tgaagagatctttacatccacgcagaagcctatgcagcgggcgagctaaagcatggcc 4594

Query: 1520 cgctggcgctaattgatgccgatatgccggttattgttggcaccgaacaacgaattgc 1579

| ||||| ||||| ||||| || ||||| ||||| | || || ||

Sbjct: 4595 cattggcgtaattgatgcggatatgccagtggttggtgcaccaagcaatgaactgt 4654

Query: 1580 tggaaaaactgaaatccaacattgaagaagttcgcgcgcg 1619

| ||||| || ||||| ||||| || |||||

Sbjct: 4655 tagaaaagcttaaatccaatattgaagaagtgcgtgcgcg 4694

Query: 889 cagatcctgcctgtggtacttctataactccggtatggttcccgctactggttgaa 948

||||| ||||| ||||| ||||| | || || | || |||||

Sbjct: 3961 cagatcctgcctgcgtggtacttctataatgcaggatgacggcacgttactggttgaa 4020

Query: 949 tcgctagcaggtattccgtgcgacgtcgaaatcgctctgaattccgctatcgcaaat 1006

||| ||||| || | || || ||||| ||||| |||||

Sbjct: 4021 tcgtagcgggtgtgagctgtgatgcgaaatcgctctgaattccgctatcgcaaat 4078

Query: 1754 cttaccatgtcgcgctgatcaaaggcaccgacgttgaccagccgctaacctggcaaaat 1813

||||||| || | ||||| ||||| ||||| ||||| || |||||

Sbjct: 4829 cttaccatgtggctttaatcaaaggtagcgttgaccagcctcgtaaccttgctaaag 4888

Query: 1814 cggttacggttgagtaa 1830

||||| || || |||||

Sbjct: 4889 cggtaacgtcgagtaa 4905

Query: 1123 tgtaacgttccgggttcttctctggtgcgcaatc 1157

||||||| ||||||||| ||||| |||||

Sbjct: 4195 tgtaacgttgcgggttcttctctctggtgcgtaatc 4229

Query: 285 ggtggtgcataacggcatcatcgaaaacctgaaccgctgcgt 327

||||| || ||||||| ||||||||| |||||||

Sbjct: 3357 ggtggtacacaacggcattatcgaaaacctgaaatgctgcgt 3399

Query: 1 atgtgtggaattgttggcgcatcgcgcaacgtgatgtagcagaaatccttcttgaaggt 60

||||||||||| ||| | || ||||| ||||| || ||||| | | ||||

Sbjct: 3073 atgtgtggaattgttggtgcggttgacacaacgcgatgttgctgaaatttagtacaaggc 3132

Query: 61 ttacgtcgtctggaataccgcggatgactctgc 95

|||| ||||| ||||||||| |||||||||

Sbjct: 3133 ctacgccgtcttgaataccgcggctatgactctgc 3167

**BLASTN** dasturining ixtisosligi *E. coli* glmS geniga etkazish xususiyatlarining natijalari

*Pseudomonas aeruginosa*, *Pasteurella multocida* va *Vibrio cholerae* genomlari bo'yicha. hakl 1. *E. coli* glmS genlar ketma-ketligi va *E. coli* GlnS ortologini kodlovchi *V. cholerae* genlari ketma-ketligi bo'yicha mintaqalarni tekislash. Tizimlar BLASTN dasturi tomonidan qurilgan.

Genlarning to'liq bo'linishini bir necha qismlarga bo'lishining sababi bu genetik kodning nasli. Ko'pgina aminokislotalar qoldiqlari uchun kodonning uchinchi pozitsiyasi ahamiyatga ega emas (ya'ni har xil uchinchi pozitsiyaga ega kodonlar odatda bir xil aminokislotalarni kodlaydi). Shunday qilib, genning deyarli har uchinchi nukleotidi seleksiya bosimi bilan topilmaydi. Bu, uchinchi bir-biriga yaqin bo'lgan organizmlar (*E. coli* va *V. cholerae*) genlarining ayrim qismlarining har uchinchi qoldiq uchun farq qilishi mumkinligiga olib keladi, bu ham rasmda ko'rsatilgan izalanishlar bilan tasdiqlanadi.

Agar kamida har uchinchi kodonning uchinchi nukleotidi homolog genlar ketma-ketligining etarlicha uzun qismida farq qiladigan bo'lsa, unda BLASTN dasturi ushbu bo'limlarning hizalanishini qura olmaydi (mutlaqo bir xil bo'lgan ketma-ketliklarning minimal uzunligi - armatura uzunligi - kamida 11 np) bo'lishi kerak. Shuning uchun BLASTN ba'zi genlar ketma-ketliklari o'rtasidagi o'xshashlikni aniqlay

olmaydi va to'liq hizalanish qisqa qismlarga bo'linadi. Bundan tashqari, juda qisqa bo'laklarning hizalanishining elektron qiymati chegaraviy qiymatdan oshib ketishi mumkin, bu dasturning chiqishida bunday hizalanish yo'qligiga olib keladi.

### **Nazorat savollari**

1. Blast nima?
2. Nukleotid Blast nima?
3. Nukleotid blast vazifasi nima?
4. Biologik ketma-ketliklarni tenglashtirishdan maqsad nima?
5. Blast qanday dastur?

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
2. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

1. <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
2. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
3. (<https://www.ncbi.nlm.nih.gov/genbank/>)

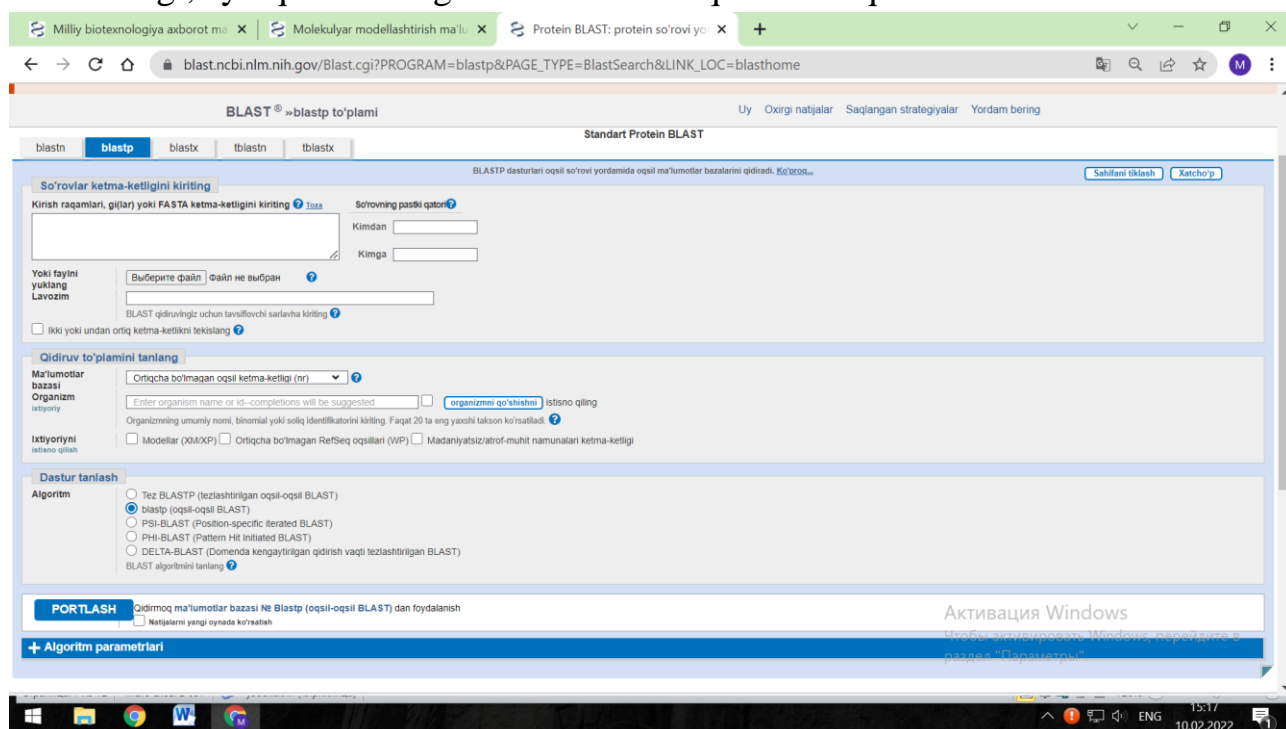
## **4- Amaliy mashg'ulot.**

**Mavzu:** BLAST dasturi vositasida aminokislotalar ketma-ketliklarini taqqoslash va tranlyantlarning olinishi.

**Dars maqsadi:** Talabalarga BLAST dasturi vositasida aminokislotalar ketma-ketliklarini taqqoslash va tranlyantlarning olinishi o'rgatish.

**Kerakli jihoz va preparatlar:** Kampyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni: BLAST qanday ishlaydi.** Barcha tekisliklar odatda global (ketma-ketliklar to'liq taqqoslanadi) va mahalliy bo'linadi (faqat ma'lum bo'limlar taqqoslanadi). BLAST seriyali dasturlari turli xil oqsillarda o'xshash domenlar va naqshlarning mavjudligi bilan bog'liq bo'lgan mahalliy moslashuvlarni ishlab chiqaradi. Bundan tashqari, mahalliy hizalanish mRNKni genom DNK bilan taqqoslashga imkon beradi. Global hizalanish holatida ketma-ketlikning kamroq o'xshashligi, ayniqsa ularning domenlari va naqshlari aniqlanadi.



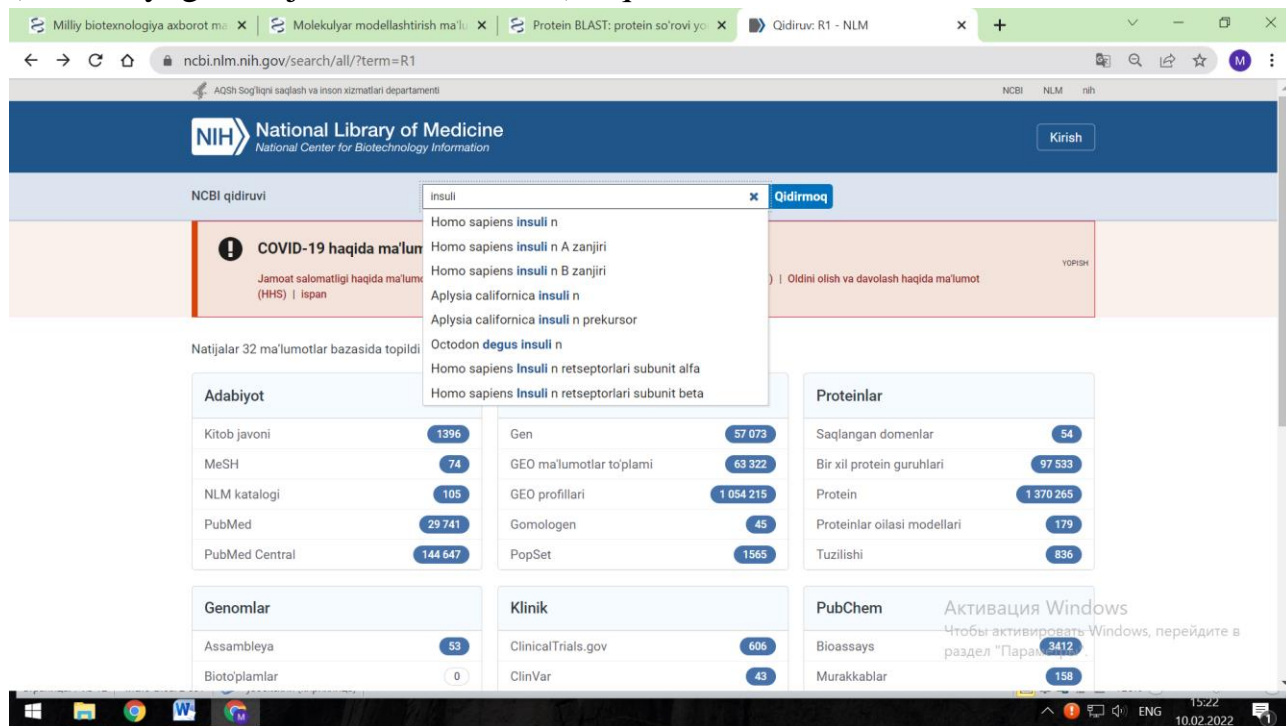
Protein Blast oynasi (blastp)

O'rganilgan aminokislotalar ketma-ketligi (so'rov) BLAST veb-sahifalaridan biriga yuborilgandan so'ng, u boshqa kirish ma'lumotlari (ma'lumotlar bazasi, "so'z" (fitna) hajmi, E qiymati va boshqalar) bilan birgalikda serverga yuboriladi. BLAST barcha "so'zlar" jadvalini tuzadi (oqsilda bu uchta aminokislotalardan iborat nukleotidlar va 11 ta nukleotid kislotalardan iborat ketma-ketliklar bo'limidir).

Keyin ularni ma'lumotlar bazasida qidirishadi. Gugurt topilganda, "so'z" ning hajmini (4 yoki undan ko'p aminokislotalar va 12 yoki undan ko'p nukleotidlarga qadar), avval bo'shliqlarsiz (bo'shliqlarsiz), so'ngra ulardan foydalanishga urinish amalga oshiriladi. O'rganilayotgan ketma-ketlikning barcha "so'zlari" o'lchamlarini maksimal kengaytirgandan so'ng, moslashtirish har bir so'rov - juftliklar uchun ma'lumotlar bazasining ketma-ketligi bo'yicha maksimal natijalar bilan belgilanadi va

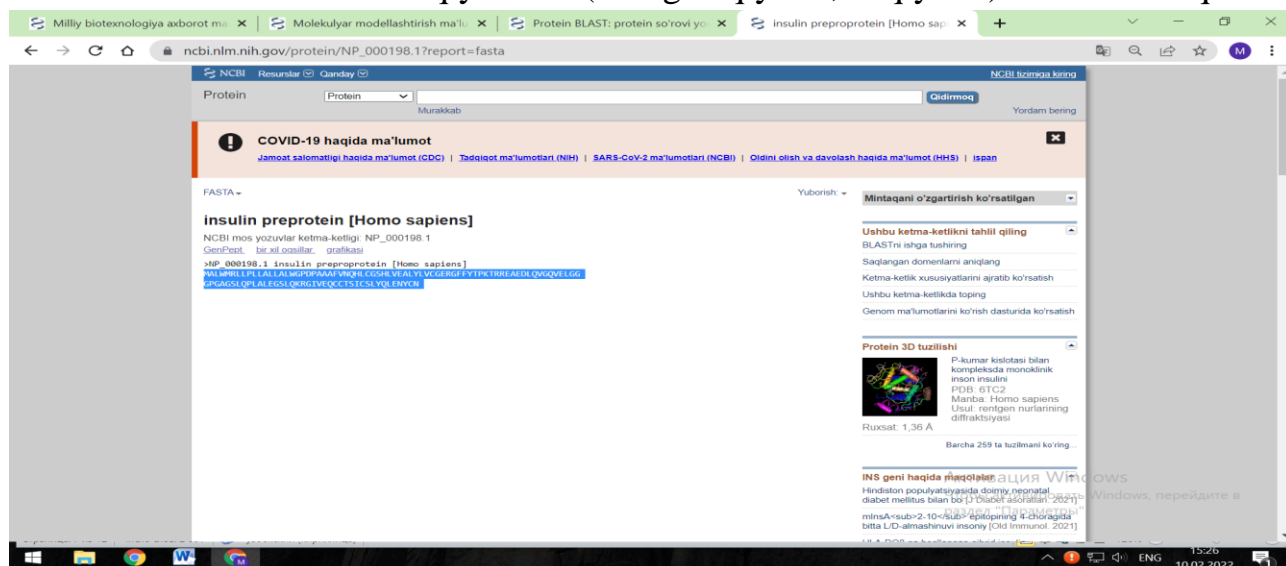


olingan ma'lumotlar SeqAlign tuzilmasida qayd etiladi. BLAST serverida joylashgan formatlash vositasi SeqAlign ma'lumotlaridan foydalanadi va uni turli usullar bilan (an'anaviy, grafik, jadval ko'rinishida) taqdim etadi.

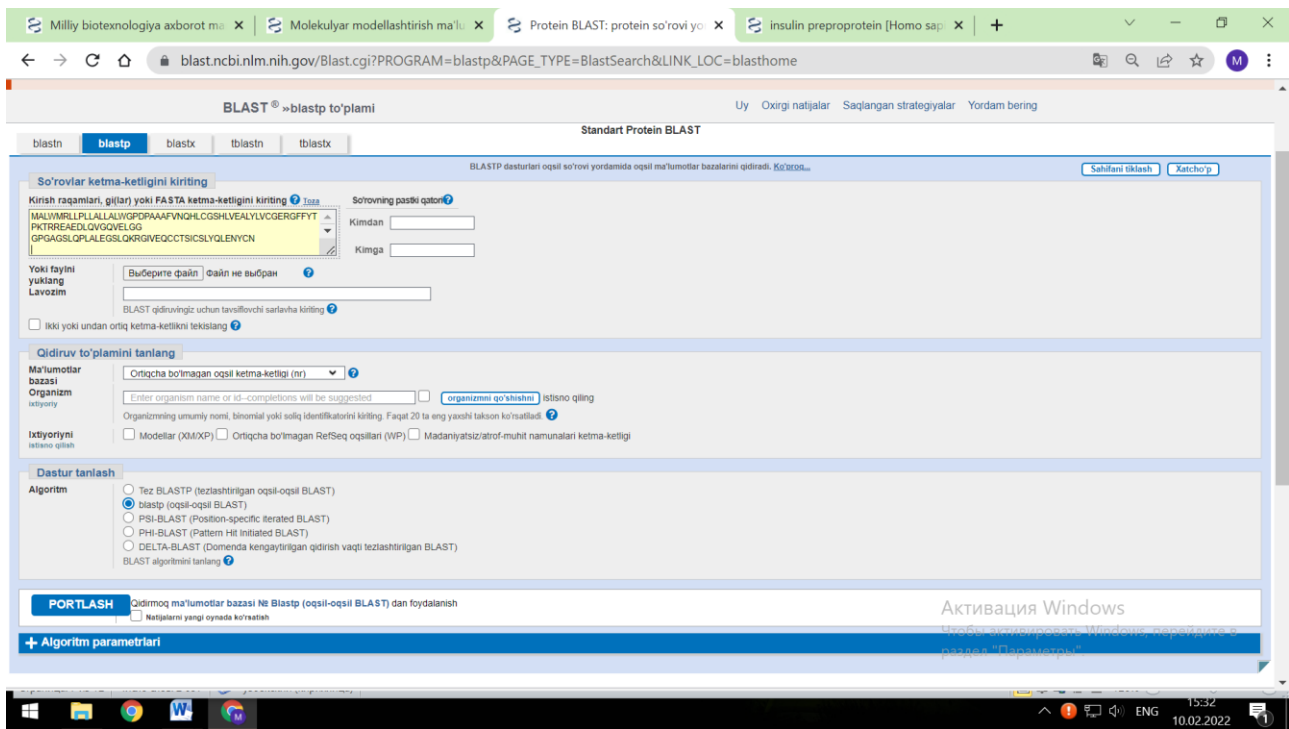


NCBI ma'lumotlar bazasidan istalgan ketma-ketlikni izlash (bu yerda Homo sapiens insulin oqsili tanlangan).

BLAST dasturlari tomonidan ma'lumotlar bazasida aniqlangan har bir ketma-ketlik uchun, o'rganilayotgan ketma-ketlikka (so'rovga) qanchalik o'xshashligini va bu o'xshashlik ahamiyatli ekanligini aniqlash kerak. Buning uchun BLAST har bir ketma-ketlik uchun bit sonini va E qiymatini (kutilgan qiymat, E-qiymat) hisoblab chiqadi.



NCBI fasta formatidan oqsil aminokislotalaridan ketma-ketligidan nusxa olish



### Yuklangan aminokislotalar ketma-ketligini blast ramkasiga joylashtirish

O'xshashlikni aniqlashda asosiy element o'rin almashtirish matritsasi hisoblanadi, chunki u har qanday mumkin bo'lgan nukleotidlar yoki aminokislotalar juftligi uchun o'xshashlik indekslarini aniqlaydi. BLAST seriyasining ko'plab dasturlarida

- LOSUM62 matritsasiidan foydalanadi (Blokarni almashtirish matritsasi 62% identifikatsiya, 62% identifikatsiyali blokni almashtirish matritsasi). Istisno holatlar blastn va megablast (nukleotidni ishlatadigan dasturlar -nukleotidlarni taqqoslash va aminokislotalarni almashtirish matritsalarini ishlatmaslik).
- O'zgartirilgan Smith-Waterman yoki Sellers algoritmlaridan foydalanib, barcha segmentlar (kengaytirilgan "so'zlar") aniqlanmagan, chunki ular o'xshashlik ko'rsatkichlarining pasayishiga olib keladi.

Bunday kengaytirilgan "so'zlar" juftlari yuqori ko'rsatkich segmentlari (HSP) deb nomlanadi. O'rganilgan ketma-ketliklar (m) va ma'lumotlar bazasining ketma-ketligi (n) etarlicha katta bo'lsa, HSP o'xshashlik ko'rsatkichlari ikkita parametr K (qidiruv maydonining o'lchami) va P bilan tavsiflanadi.

Ushbu ko'rsatkichlar o'rganilayotgan ketma-ketlik va ma'lumotlar bazasi ketma-ketligining o'xshashlik ko'rsatkichlarini (S) keltirishda ko'rsatilishi kerak.

BLAST (Basic Basic Alignment Search Tool) - asosiy tuzilish (ketma-ketlik) yoki uning bo'lagi ma'lum bo'lgan oqsillar yoki nuklein kislotalarning homologlarini qidirishda ishlatiladigan kompyuter dasturlari oilasi. BLAST-dan foydalanib,

tadqiqotchi o'z ketma-ketligini ma'lumotlar bazasidagi ketma-ketliklar bilan taqqoslashi va taxmin qilinayotgan homologlarning ketma-ketligini topishi mumkin. Bu molekulyar biologlar, bioinformatika, sistematikalar uchun eng muhim vositadir. BLAST dasturi olimlar Stiven Altschul, Uorren Gish, Uebb Miller, Evgeniy Myers va Devid J. Lipman tomonidan AQShning Milliy Sog'liqni saqlash institutida ishlab chiqilgan va 1990 yilda Molekulyar Biologiya jurnalida nashr etilgan [1].

O'rganilgan nukleotid yoki aminokislotalar ketma-ketligi (so'rov) BLAST veb-sahifalaridan biriga kiritilgandan so'ng, u boshqa kirish ma'lumotlari (ma'lumotlar bazasi, "so'z" (fitna) hajmi, E qiymati va boshqalar) bilan birgalikda serverga yuboriladi. BLAST barcha "so'zlar" jadvalini (oqsilda, bu uchta aminokislotalardan tashkil topgan va 11 ta nukleotidning nuklein kislotalaridan iborat ketma-ketliklar bo'limi) va shunga o'xshash "so'zlar" ning jadvalini tuzadi.

### **Nazorat savollari**

1. BLAST seriyali dasturlarning aksariyati qanday matritsadan foydalanadi?
2. Yuklangan aminokislotalar ketma-ketligini blast ramkasiga joylashtirish
3. Proteinli Blast nima?
4. Proteinli blast vazifasi nima?
5. Biologik ketma-ketliklarni tenglashtirishdan maqsad nima?

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
2. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

3. <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
4. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
5. (<https://www.ncbi.nlm.nih.gov/genbank/>)

## 5- Amaliy mashg'ulot.

**Mavzu:** Oqsil molekulalarida gomologik domenlarni izlash.

**Dars maqsadi:** Talabalarga Oqsil molekulalarida gomologik domenlarni izlash bo'yicha amaliy ko'rsatmalar va tushunchalar berish.

**Kerakli jihoz va preparatlar:** Kamyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** Protein domeni - bu oqsil polipeptid zanjirining o'z-o'zidan turg'un bo'lgan va qolgan qismidan mustaqil ravishda katlanadigan qismi. Har bir domen ixcham katlanmish uch o'lchamli strukturani hosil qiladi. Ko'pgina oqsillar bir nechta domenlardan iborat. Bitta domen turli xil oqsillarda paydo bo'lishi mumkin. Molekulyar evolyutsiya domenlarni qurilish bloklari sifatida ishlatadi va ular turli funktsiyalarga ega bo'lgan oqsillarni yaratish uchun turli xil tuzilmalarda qayta birlashtirilishi mumkin. Umuman olganda, domenlar uzunligi taxminan 50 ta aminokislotadan 250 ta aminokislotagacha o'zgarib turadi. Eng qisqa domenlar, masalan, sink barmoqlari, metall ionlari yoki disulfid ko'prigi bilan barqarorlashadi.

**Protein tuzilishi birliklari.** Proteinning birlamchi tuzilishi (aminokislotalar qatori) oxir-oqibat uning noyob katlami uch o'lchovli (3D) konformatsiyasini kodlaydi. Proteinning 3D tuzilishiga katlanishini boshqaradigan eng muhim omil qutbli va qutbsiz yon zanjirlarning taqsimlanishidir. Buklanish suvli muhit bilan aloqa qilmaslik uchun gidrofobik yon zanjirlarning molekulaning ichki qismiga ko'milishiga olib keladi. Odatda oqsillar gidrofilik qoldiqlar qobig'i bilan o'ralgan hidrofobik qoldiqlardan iborat yadroga ega. Peptid aloqalarining o'zi qutbli bo'lgani uchun ular hidrofobik muhitda bir-biri bilan vodorod bog'lanishi orqali neytrallanadi. Bu ikkilamchi struktura deb ataladigan muntazam 3D strukturaviy naqshlarni hosil qiluvchi polipeptid hududlarini keltirib chiqaradi. Ikkilamchi strukturaning ikkita asosiy turi mavjud: a-spiral va b-varaq. Ikkilamchi tuzilish elementlarining ba'zi oddiy birikmalari oqsil tarkibida tez-tez uchraydi va ular supersekondar tuzilma yoki motivlar deb ataladi. Misol uchun, b-sochli naqsh kichik halqa bilan birlashtirilgan ikkita qo'shni antiparallel b-iplardan iborat. Ko'pgina antiparallel b tuzilmalarda ham izolyatsiya qilingan lenta sifatida, ham murakkab b-varaqlarning bir qismi sifatida mavjud. Yana bir keng tarqalgan super-ikkinchi darajali tuzilma b-a-b motivi bo'lib, u tez-tez ikkita parallel b-torni ulash uchun ishlatiladi. Markaziy a-spiral birinchi ipning C-uchini ikkinchi ipning N-uchlari bilan bog'lab, uning yon zanjirlarini b-varaqqaga o'raladi va shuning uchun b-torchalarning hidrofobik qoldiqlarini sirtdan himoya qiladi. Ikki domenning kovalent assotsiatsiyasi funktsional va tizimli afzalliklarni anglatadi, chunki kovalent bog'liq bo'lmagan bir xil tuzilmalar bilan solishtirganda barqarorlik oshadi. Boshqa afzalliklar - suvli muhitda beqaror bo'lishi mumkin bo'lgan domenlararo fermentativ yoriqlar ichidagi oraliq mahsulotlarni himoya qilish va ketma-ket reaksiyalar to'plami uchun zarur bo'lgan fermentativ faollikning sobit

stexiometrik nisbati. Strukturaviy moslashuv domenlarni aniqlash uchun muhim vositadir.

**Uchinchi darajali tuzilish.** Domenlar deb ataladigan ixcham, mahalliy, yarim mustaqil birliklarni hosil qilish uchun bir nechta motivlar birlashadi. Polipeptid zanjirining umumiy 3D tuzilishi oqsilning uchinchi darajali tuzilishi deb ataladi. Domenlar uchinchi darajali tuzilmaning asosiy birliklari bo'lib, har bir domen halqa hududlari bilan bog'langan ikkilamchi strukturaviy birliklardan qurilgan individual hidrofobik yadroni o'z ichiga oladi. Polipeptidning qadoqlanishi odatda qattiq yadro va suyuqlikka o'xshash sirt hosil qiluvchi domenning tashqi qismiga qaraganda ichki qismda ancha qattiqroqdir. Yadro qoldiqlari ko'pincha oqsillar oilasida saqlanadi, halqalardagi qoldiqlar esa, oqsil funksiyasida ishtirok etmasa, kamroq saqlanadi. Proteinning uchinchi darajali tuzilishini domenning ikkilamchi tarkibiy tarkibiga qarab to'rtta asosiy sinfga bo'lish mumkin. Barcha-a domenlari faqat a-spirallardan qurilgan domen yadrosiga ega. Bu sinfda kichik burmalar ustunlik qiladi, ularning ko'pchiligi yuqoriga va pastga harakatlanadigan spirallar bilan oddiy to'plamni tashkil qiladi. Barcha b domenlari antiparallel b-varaqlardan tashkil topgan yadroga ega, odatda bir-biriga o'ralgan ikkita varaq. Iplarni joylashtirishda turli naqshlarni aniqlash mumkin, bu ko'pincha takrorlanuvchi naqshlarni, masalan, yunoncha kalit naqshlarni aniqlashga olib keladi. a+b domenlari all-a va all-b motivlarining aralashmasidir. Proteinlarni ushbu sinfga ajratish qiyin, chunki boshqa uchta sinf bilan bir-biriga mos keladi va shuning uchun CATH domeni ma'lumotlar bazasida ishlatilmaydi.

a/b domenlari asosan amfipatik a-spirallar bilan o'ralgan parallel b-varaqli tashkil etuvchi b-a-b motivlarining kombinatsiyasidan yaratilgan. Ikkilamchi tuzilmalar qatlamlar yoki bochkalarda joylashgan.

**Hajmi bo'yicha cheklovlar.** Domenlar hajmi bo'yicha cheklovlarga ega. Alohida strukturaviy domenlarning o'lchami E-selektindagi 36 qoldiqdan lipoksigenaza-1dagi 692 qoldiqgacha o'zgarib turadi, lekin ko'pchilik, 90%, o'rtacha 100 ta qoldiq bilan 200 dan kam qoldiqga ega. Juda qisqa domenlar, 40 dan kam qoldiqlar ko'pincha metall ionlari yoki disulfid bog'lari bilan barqarorlashadi. Kattaroq domenlar, 300 dan ortiq qoldiqlar, bir nechta hidrofobik yadrolardan iborat bo'lishi mumkin.

**To'rtlamchi tuzilish.** Ko'pgina oqsillar to'rtlamchi tuzilishga ega bo'lib, ular oligomerik molekulaga birikadigan bir nechta polipeptid zanjirlaridan iborat. Bunday oqsildagi har bir polipeptid zanjiri subbirlik deb ataladi. Masalan, gemoglobin ikkita a va ikkita b bo'linmalaridan iborat. To'rt zanjirning har birida gem cho'ntagasi bilan to'liq a globin burmasi mavjud.

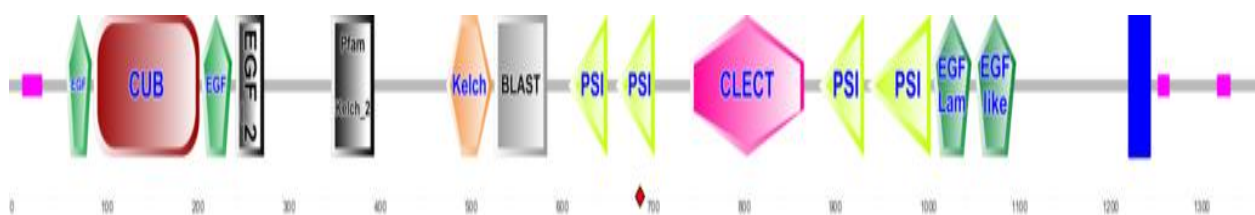
**Domenni almashtirish.** Domenni almashtirish - bu oligomerik yig'ilishlarni shakllantirish mexanizmi. Domenni almashtirishda monomerik oqsilning ikkilamchi yoki uchinchi elementi boshqa oqsilning bir xil elementi bilan almashtiriladi. Domenni

almashtirish ikkilamchi struktura elementlaridan butun strukturaviy domenlargacha bo'lishi mumkin. Shuningdek, u oligomerizatsiya orqali funktsional moslashish uchun evolyutsiya modelini ifodalaydi, masalan. oligomerik fermentlar, ularning faol joyi subunit interfeyslarida.

Molekulyar evolyutsiya o'xshash ketma-ketlik va tuzilishga ega bo'lgan oqsillar oilalarini keltirib chiqaradi. Biroq, bir xil tuzilishga ega bo'lgan oqsillar orasida ketma-ketlik o'xshashligi juda past bo'lishi mumkin. Protein tuzilmalari o'xshash bo'lishi mumkin, chunki oqsillar umumiy ajdoddan ajralib chiqqan. Shu bilan bir qatorda, ba'zi burmalar boshqalardan ko'ra ko'proq ma'qul bo'lishi mumkin, chunki ular ikkilamchi tuzilmalarning barqaror joylashuvini ifodalaydi va ba'zi oqsillar evolyutsiya jarayonida bu burmalar tomon birlashishi mumkin.

**Ko'p domenli oqsillar.** Oqsillarning ko'p qismi, bir hujayrali organizmlarda uchdan ikki qismi va metazoalarda 80% dan ortig'i ko'p domenli oqsillardir. Biroq, boshqa tadqiqotlar shuni ko'rsatdiki, prokaryotik oqsillarning 40% bir nechta domenlardan iborat, eukaryotlarda esa taxminan 65% ko'p domenli oqsillar mavjud.

Eukaryotik multidomenli oqsillardagi ko'plab domenlar prokaryotlarda mustaqil oqsillar sifatida topilishi mumkin, bu ko'p domenli oqsillardagi domenlar bir vaqtlar mustaqil oqsillar sifatida mavjud bo'lganligini ko'rsatadi. Masalan, umurtqali hayvonlarda GAR sintetaza, AIR sintetaza va GAR transformilaza domenlarini o'z ichiga olgan ko'p fermentli polipeptid mavjud (GARs-AIRs-GARt; GAR: glitsinamid ribonukleotid sintetaza/transferaza; AIR: aminoimidazol ribonukleotid sintetaza). Hasharotlarda polipeptid GARs-(AIRs)<sub>2</sub>-GARt ko'rinishida ko'rinadi, xamirturushda GARs-AIRs GARt dan alohida kodlangan, bakteriyalarda esa har bir domen alohida kodlangan.



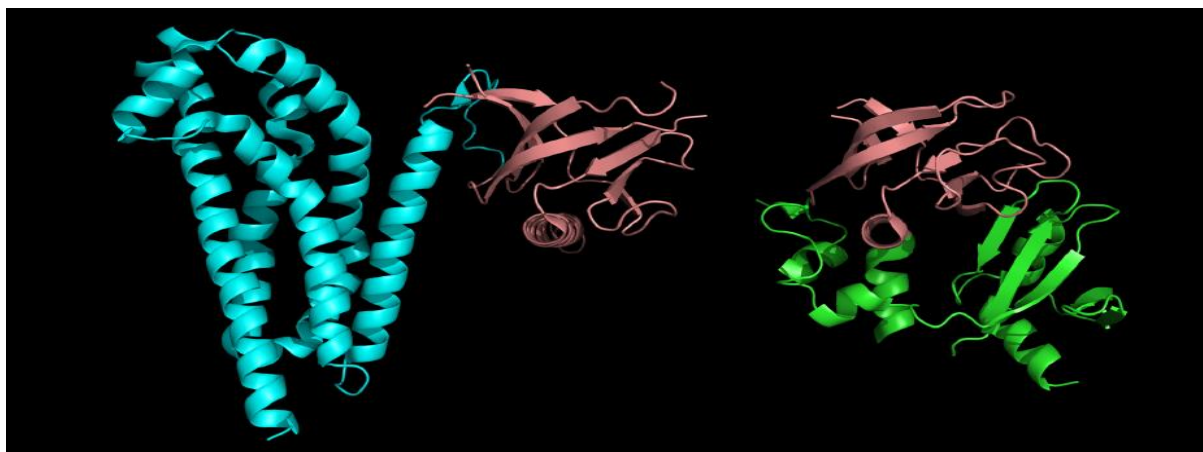
Attractinga o'xshash protein 1 (ATRNL1) hayvonlarda, jumladan, odamlarda topilgan ko'p domenli oqsildir. Har bir birlik bitta domendir, masalan. EGF yoki Kelch domenlari.

[https://en.wikipedia.org/wiki/Protein\\_domain#/media/File:ATRNL1\\_Bitmap.png](https://en.wikipedia.org/wiki/Protein_domain#/media/File:ATRNL1_Bitmap.png)

**Tashkilot turlari.** Ikki xil oqsilga o'xshash PH domen modullarini (maroon) kiritish. Proteinlarda ko'rilgan eng oddiy ko'p domenli tashkilot tandemda takrorlanadigan bitta domendir. Domenlar bir-biri bilan o'zaro ta'sir qilishi mumkin (domen-domen o'zaro ta'siri) yoki ipdagi boncuklar kabi izolyatsiya qilingan holda qolishi mumkin. Gigant 30 000 qoldiq mushak oqsili titin 120 ga yaqin fibronektin-III

va Ig tipidagi domenlarni o'z ichiga oladi. Serin proteazalarida genlarning duplikatsiyasi hodisasi ikkita b-barrelli domen fermentining shakllanishiga olib keldi. Takrorlashlar shu qadar keng tarqalganki, ular orasida aniq ketma-ketlik o'xshashligi yo'q. Faol sayt ikkita b-barrel domenlari orasidagi bo'shliqda joylashgan bo'lib, unda har bir domendan funktsional muhim qoldiqlar qo'shiladi. Ximotripsin serin protezasining genetik jihatdan ishlab chiqilgan mutantlari, ularning faol joy qoldiqlari bekor qilingan bo'lsa ham, bir oz proteinaza faolligiga ega ekanligi ko'rsatildi va shuning uchun takrorlanish hodisasi ferment faolligini kuchaytirdi, deb taxmin qilingan. Modullar ko'pincha kinesinlar va ABC tashuvchilar tomonidan tasvirlanganidek, turli xil ulanish munosabatlarini namoyish etadi. Kinesin motor domeni polipeptid zanjirining har ikki uchida bo'lishi mumkin, bu o'ralgan-o'ralgan mintaqa va yuk domenini o'z ichiga oladi. ABC transporterlari bir-biriga bog'liq bo'lmagan ikkita moduldan, ATP-bog'lovchi kassetada va turli xil kombinatsiyalarda joylashgan integral membrana modulidan iborat to'rttagacha domen bilan qurilgan. Domenlar nafaqat qayta birlashadi, balki boshqa domenga kiritilgan domenning ko'plab misollari mavjud. Boshqa domenlar bilan ketma-ketlik yoki strukturaviy o'xshashliklar kiritilgan va asosiy domenlarning gomologlari mustaqil ravishda mavjud bo'lishi mumkinligini ko'rsatadi. Bunga misol qilib, Pol I oilasining polimerazalari ichida "xurmo" domenga kiritilgan "barmoqlar"ni keltirish mumkin. Domen boshqa domenga kiritilishi mumkinligi sababli, ko'p domenli oqsilda har doim kamida bitta doimiy domen bo'lishi kerak. Bu strukturaviy domenlar va evolyutsion/funktsional domenlarning ta'riflari o'rtasidagi asosiy farq. Evolyutsion domen domenlar orasidagi bir yoki ikkita ulanish bilan chegaralanadi, strukturaviy domenlar esa umumiy yadro mavjudligining berilgan mezoni doirasida cheksiz ulanishlarga ega bo'lishi mumkin. Evolyutsion domenga bir nechta strukturaviy domenlar tayinlanishi mumkin. Superdomen nominal mustaqil kelib chiqishi bo'lgan ikki yoki undan ortiq saqlangan domenlardan iborat bo'ladi, lekin keyinchalik bitta strukturaviy/funktsional birlik sifatida meros bo'lib o'tadi. Ushbu birlashtirilgan superdomen faqat genlarning ko'payishi bilan bog'liq bo'lmagan turli xil oqsillarda paydo bo'lishi mumkin. Superdomenga misol sifatida PTEN, tensin, auksilin va TPTE2 membrana oqsilidagi tirozin fosfataza-C2 domen juftligi kiradi. Ushbu superdomen hayvonlar, o'simliklar va zamburug'lardagi oqsillarda mavjud. PTP-C2 superdomenining asosiy xususiyati domen interfeysida aminokislota qoldiqlarini saqlashdir.





Ikki xil oqsilga o'xshash PH domen modullarini (maroon) kiritish.

Shu bilan birga, domenlararo o'zaro ta'sirlarning oqsillarni katlamada va mahalliy tuzilmani barqarorlashtirish energetikasida roli, ehtimol, har bir oqsil uchun farq qiladi. T4 lizozimida bir domenning ikkinchisiga ta'siri shunchalik kuchliki, butun molekula proteolitik parchalanishga chidamli. Bunday holda, katlama ketma-ket jarayon bo'lib, C-terminal domeni dastlabki bosqichda mustaqil ravishda katlanishi kerak, boshqa domen esa katlama va barqarorlashtirish uchun katlanmish C-terminal domenining mavjudligini talab qiladi.

Izolyatsiya qilingan domenning katlanishi integratsiyalashgan domennikiga qaraganda bir xil tezlikda yoki ba'zan tezroq sodir bo'lishi mumkinligi aniqlandi, buklanish paytida oqsilning qolgan qismi bilan noqulay o'zaro ta'sirlar paydo bo'lishi mumkinligini ko'rsatadi. Bir nechta argumentlar shuni ko'rsatadiki, katta oqsillarni katlamaning eng sekin bosqichi buklangan domenlarni juftlashtirishdir. Buning sababi yoki domenlar to'liq to'g'ri yig'ilmaganligi yoki ularning o'zaro ta'siri uchun zarur bo'lgan kichik sozlashlar energiya jihatidan noqulay bo'lganligi, domen interfeysidan suvni olib tashlash kabi.

**Strukturaviy koordinatalardan domen ta'rifi.** Domenlarning strukturaviy qurilish bloklari va evolyutsiya elementlari sifatidagi ahamiyati ma'lum strukturadagi oqsillarda ularni aniqlash va tasniflashning ko'plab avtomatlashtirilgan usullarini keltirib chiqardi. Ishonchli domenni tayinlash uchun avtomatik protseduralar domen ma'lumotlar bazalarini yaratish uchun juda muhim, ayniqsa ma'lum protein tuzilmalari soni ortib bormoqda. Domen chegaralarini vizual tekshirish orqali aniqlash mumkin bo'lsa-da, avtomatlashtirilgan usulni qurish oson emas. Muammolar uzluksiz yoki yuqori darajada bog'langan domenlarga duch kelganda paydo bo'ladi.

RigidFinder - bu ikki xil konformatsiyadan oqsil qattiq bloklarini (domenlar va halqalar) aniqlashning yangi usuli. Qattiq bloklar barcha qoldiqlar orasidagi masofalar konformatsiyalar bo'yicha saqlanadigan bloklar sifatida aniqlanadi.

Pandurangan va Topf tomonidan ishlab chiqilgan RIBFIND usuli oqsillardagi ikkilamchi strukturaviy elementlarning fazoviy klasterlanishini amalga

o'shinish orqali oqsil tuzilmalarida qattiq jismlarni aniqlaydi . RIBFIND qattiq jismlari oqsil tuzilmalarini kriyoelektron mikroskop zichlik xaritalariga moslashuvchan tarzda moslashtirish uchun ishlatilgan.

### **Nazorat savollari**

1. Gomologik oqsil nima?
2. Domen nima?
3. Oqsil domen nima?
4. Domenlar va oqsil moslashuvchanligi deganda nima tushuniladi?
5. Domenning avtonom katlama birliklari deganda nimani tushunasiz?

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

4. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
5. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

6. [https://en.wikipedia.org/wiki/Protein\\_domain](https://en.wikipedia.org/wiki/Protein_domain)
7. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
8. (<https://www.ncbi.nlm.nih.gov/genbank/>)

## 6 - Amaliy mashg'ulot.

**Mavzu:** Hisoblashning ochiq ramkasini izlash. ORF Finder dasturida ishlash.

**Dars maqsadi:** Talabalarga hisoblashning ochiq ramkasini izlash va ORF Finder dasturida ishlash haqida umumiy ma'lumot berish.

**Kerakli jihoz va preparatlar:** Komyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** DNK (deoksiribonuklein kislotasi) - bu tirik organizmlardagi barcha genetik ma'lumotlarni o'z ichiga olgan genetik material. Ma'lumot adenin (A), guanin (G), sitozin (C) va timin (T) yordamida genetik kodlar sifatida saqlanadi. Transkripsiya jarayonida DNK mRNKga ko'chiriladi. Ushbu tayanch juftlarining har biri shakar va fosfat molekulasi bilan bog'lanib, nukleotid hosil qiladi. Tarjima paytida ma'lum bir aminokislotalarni kodlaydigan uchta nukleotid kodon deb ataladi. Nukleotidning boshlang'ich kodondan boshlanib, to'xtash kodon bilan tugaydigan hududi **Ochiq o'qish ramkasi (ORF) deb ataladi..** Proteinlar ORF dan hosil bo'ladi. ORFn tahlil qilish orqali biz tarjima paytida hosil bo'lishi mumkin bo'lgan aminokislotalarni taxmin qilishimiz mumkin. ORF topuvchisi NCBI veb-saytida mavjud bo'lgan dasturdir. U olti xil o'qish ramkasidan barcha ORF yoki mumkin bo'lgan protein kodlash mintaqasini aniqlaydi.

Google Tarjimon x Bosh sahifa - ORFfinder - NCBI x Yulduz Usmonova-U Senga MP3 x +

ncbi.nlm.nih.gov/orffinder/

ORFfinder PubMed Qidirmoq

COVID-19 haqida ma'lumot  
Jamoa salomatligi haqida ma'lumot (CDC) | Tadqiqot ma'lumotlari (NIH) | SARS-CoV-2 ma'lumotlari (NCBI) | Oldini olish va davolash haqida ma'lumot (HHS) | ispan

O'qish uchun ramka qidiruvchini oching

ORF topuvchisi siz kiritgan DNK ketma-ketligidagi ochiq o'qish ramkalarini (ORF) qidiradi. Dastur har bir ORF diapazonini oqsil tarjimasini bilan birga qaytaradi. Potensial oqsil kodlash segmentlari uchun yangi ketma-ket DNKni qidirish uchun ORF qidiruvchisidan foydalaning, yangi ishlab chiqilgan SMART BLAST yoki oddiy BLAST yordamida prognoz qilingan protinni tekshiring.

ORF qidiruvchisining ushbu veb-versiyasi uzunligi 50 kb gacha bo'lgan so'rovlar ketma-ketligining pastki diapazoni bilan cheklangan. So'rovlar ketma-ketligi uzunligi chekloviga ega bo'lmagan mustaqil versiya Linux x64 uchun mavjud.

**Misolalar** (qimmatlarni o'rnatish uchun bosing, keyin Yuborish tugmasini bosing).

- NC\_011604 Salmonella enterica plazmid pWES-1; genetik kod: 11; "ATG" va muqobil boshlash kodonlari; minimal ORF uzunligi: 300 nt
- NM\_000059; genetik kod: 1; boshlang'ich kodoni: "taqat ATG"; minimal ORF uzunligi: 150 nt

So'rovlar ketma-ketligini kiriting

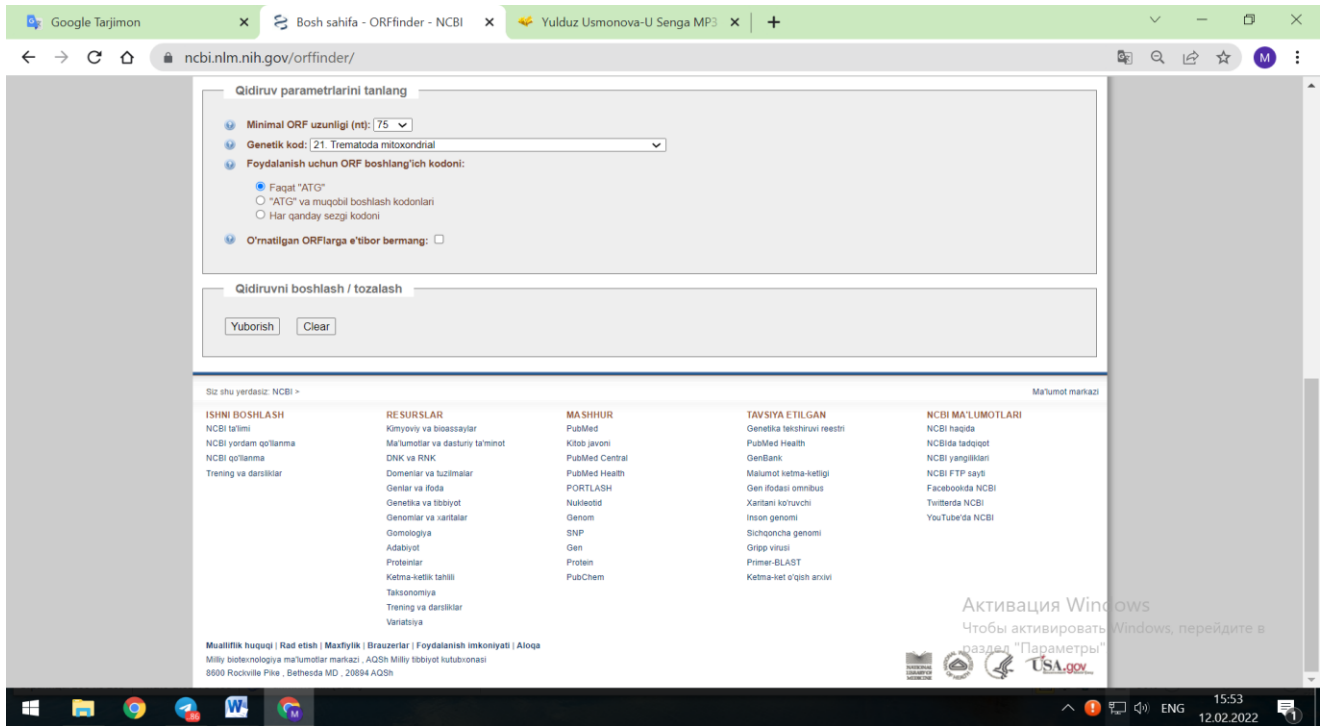
FASTA formatida kirish raqami, gi yoki nukleotidlar ketma-ketligini kiriting:

Kimdan: Kimga:

Qidiruv parametrlarini tanlang

Активация Windows  
Чтобы активировать Windows, перейдите в раздел "Параметры"

15:50  
12.02.2022



## ORF Finder dasturining oynasi

### ORFni qanday topish mumkin

1. Gipotetik ketma-ketlikni ko'rib chiqing:

**CGCTACGTCCTTACGCTGGAGCTCTCATGGATCGGTTCCGGTAGGGCTC  
GATCCACATCGCTAGCCAT**

2. Ketma-ketlikni 6 xil o'qish ramkalariga bo'ling (+1, +2, +3, -1, -2 va -3). Birinchi o'qish ramkasi 3 so'zlardagi ketma-ketlikni hisobga olgan holda olinadi.

**FRAME +1: CGC TAC GTC TTA CGC TGG AGC TCT CAT GGA TCG  
GTT CGG TAG GGC TCG ATC ACA TCG CTA GCC AT**

Ikkinchi o'qish ramkasi birinchi nukleotidni qoldirib, keyin ketma-ketlikni 3 ta nukleotid so'zlariga guruhlashdan keyin hosil bo'ladi.

**FRAME +2: C GCT ACG TCT TAC GCT GGA GCT CTC ATG GAT CGG  
TTC GGT AGG GCT CGA TCA CAT CGC TAG CCA T**

Uchinchi o'qish ramkasi dastlabki 2 nukleotidni qoldirib, keyin ketma-ketlikni 3 ta nukleotid so'zlariga guruhlagandan so'ng hosil bo'ladi.

**FRAME +3: CG CTA CGT CTT ACG CTG GAG CTC TCA TGG ATC  
GGT TCG GTA GGG CTC GAT CAC ATC GCT AGC CAT**

Qolgan 3 ta o'qish ramkasini faqat teskari to'ldiruvchi topilgandan keyingina topish mumkin.

To'ldiruvchi

: **GCGATGCAGAATGCGACCTCGAGAGTACCTAGCCAAGCCATC  
CCGAGCTAGTGTAGCGATCGGTA**

Teskari

to'ldiruvchi: **ATGGCTAGCGATGTGATCGAGCCCTACCGAACCGAT  
CCATGAGAGCTCCAGCGTAAGACGTAGCG**

Endi +1, +2 va +3 iplardagi kabi jarayon teskari komplementlar ketma-ketligi bilan -1, -2 va -3 iplar uchun takrorlanadi.

**FRAME -1: ATG GCT AGC GAT GTG ATC GAG CCC TAC CGA ACC  
GAT CCA TGA GAG CTC CAG CGT AAG ACG TAG CG**

**FRAME -2: A TGG CTA GCG ATG TGA TCG AGC CCT ACC GAA CCG  
ATC CAT GAG AGC TCC AGC GTA AGA CGT AGC G**

**FRAME -3: AT GGC TAG CGA TGT GAT CGA GCC CTA CCG AAC  
CGA TCC ATG AGA GCT CCA GCG TAA GAC GTA GCG**

3. Endi o'qish ramkalarida boshlang'ich va to'xtash kodonlarini belgilang

**FRAME +1: CGC TAC GTC TTA CGC TGG AGC TCT CAT GGA TCG  
GTT CGG TAG GGC TCG ATC ACA TCG CTA GCC AT**

**FRAME +2: C GCT ACG TCT TAC GCT GGA GCT CTC ATG GAT  
CGG TTC GGT AGG GCT CGA TCA CAT CGC TAG CCA T**

**FRAME +3: CG CTA CGT CTT ACG CTG GAG CTC TCA TGG ATC  
GGT TCG GTA GGG CTC GAT CAC ATC GCT AGC CAT**

**FRAME -1: ATG GCT AGC GAT GTG ATC GAG CCC TAC CGA ACC  
GAT CCA TGA GAG CTC CAG CGT AAG ACG TAG CG**

**FRAME -2: A TGG CTA GCG ATG TGA TCG AGC CCT ACC GAA  
CCG ATC CAT GAG AGC TCC AGC GTA AGA CGT AGC G**

**FRAME -3: AT GGC TAG CGA TGT GAT CGA GCC CTA CCG AAC  
CGA TCC ATG AGA GCT CCA GCG TAA GAC GTA GCG**

4. Ochiq o'qish ramkasini (ORF) aniqlang - boshlang'ich kodoni bilan boshlanib, to'xtash kodonida tugaydigan ketma-ketlik cho'zilishi.

**FRAME +2: ATG GAT CGG TTC GGT AGG GCT CGA TCA CAT  
CGC TAG**

**FRAME -1: ATG GCT AGC GAT GTG ATC GAG CCC TAC CGA ACC  
GAT CCA TGA**

**FRAME -3: ATG AGA GCT CCA GCG TAA**

5. Aminokislotalar jadvali asosida peptidlar ketma-ketligi topiladi

### Aminokislotalar jadvali

FRAME +2: **ATG** GAT CGG TTC GGT AGG GCT CGA TCA CAT CGC **TAG**  
**met asp arg phe gly arg ala arg ser uning arg stop**

FRAME -1: **ATG** GCT AGC GAT GTG ATC GAG CCC TAC CGA ACC GAT  
CCA **TGA**

**met ala ser asp val ile glu pro tyr arg thr asp pro stop**

FRAME -3: **ATG** AGA GCT CCA GCG **TAA**

**met arg ala pro ala stop**

ORFni tahlil qilish orqali biz tarjima jarayonida ishlab chiqarilishi mumkin bo'lgan aminokislotalarni taxmin qilishimiz mumkin. Yangi tartiblangan gendan to'g'ri ORFni bashorat qilish muhim qadamdir. ORFni topish PCR, ketma-ketlik va boshqalar kabi tajribalar uchun zarur bo'lgan primerlarni loyihalashga yordam beradi.

### Nazorat savollari

1. ORF nima?
2. ORFni qanday topish mumkin?
3. ORF ni qaysi bazadan toppish mumkin?

### Informatsion-uslubiy ta'minot

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### QO'SHIMCHA ADABIYOTLAR

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
2. А. Н. Огурцов Введение в Бионформатику, 2011г.

### Internet ma'lumotlari

1. <https://www.ncbi.nlm.nih.gov/orffinder/>
2. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
3. (<https://www.ncbi.nlm.nih.gov/genbank/>)



## 7- Amaliy mashg'ulot.

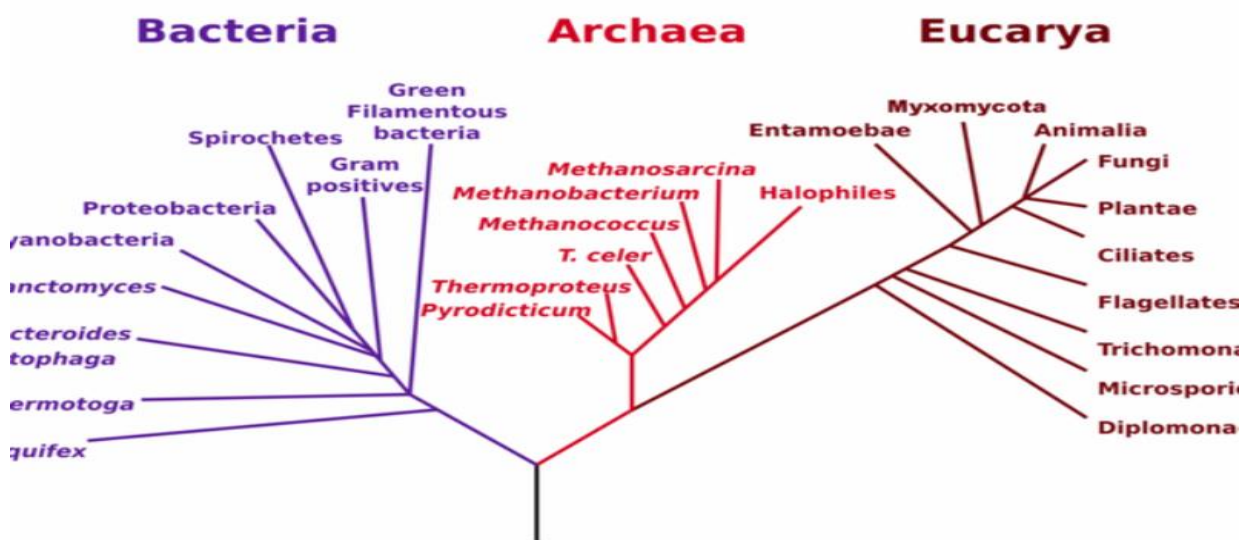
**Mavzu:** Ko'plik taqqoslanishlar va filogenetik daraxtni yaratish. Clustal W i T-Coffee dasturlari. Ma'lumotlar tahlili.

**Dars maqsadi:** Talabalarga Ko'plik taqqoslanishlar va filogenetik daraxtni yaratish bo'yicha ma'lumot berish hamda Clustal W va T-Coffee dasturlarida ma'lumotlar tahlilini yaratishni o'rgatish.

**Kerakli jihoz va preparatlar:** Kampyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** Evolyutsion aloqalarni tushunish zamonaviy biologiyaning asosiy yo'nalishi bo'lib, filogenetik daraxt ushbu birlashmalarni tasvirlash uchun asosiy vosita hisoblanadi. Biroq, turli xil biologik jarayonlarning o'xshashligi va miqdorini aniqlash maqsadida daraxtlarni taqqoslash muhim muammo bo'lib qolmoqda.

Filogenetik daraxtlarni taqqoslashdan oldin filogenetik daraxt qurish bilan tanishib chiqish lozim.

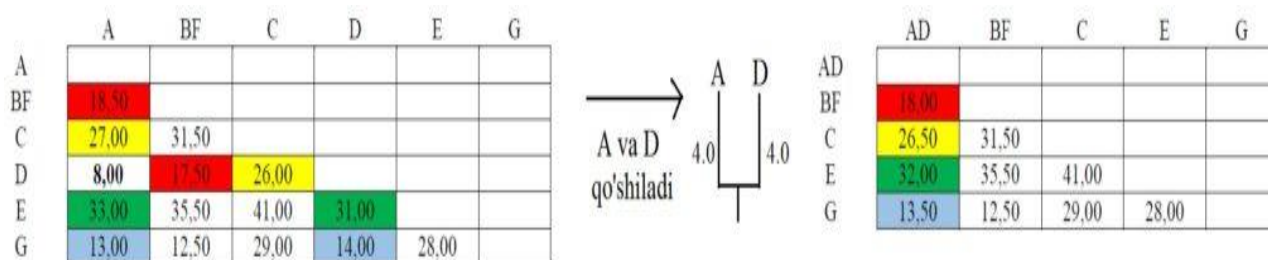


Tirik organizmlarning molekulyar ko'rsatkichlaridan foydalangan holda ularni filogenetik daraxtdagi o'rni hisoblashning juda ko'plab turlari (UPGMA, WPGMA, Neighbor-joining, Maximum-likelihood, Bayesian inference va boshqalar) mavjud. **UPGMA (Unweighted Pair Group Method with Arithmetic Mean)** – guruhdagi juftliklarni o'rtacha taqqoslanmagan hisoblash usuli bo'lib, uni ilk marotaba Sneath va Sokalning (1973) ishlarida ko'rish mumkin. UPGMA usuli asosida filogenetik daraxt qurish uchun bizga berilgan organizmlarning molekulyar ko'rsatkichlari zarur bo'ladi. Quyidagi jadvalda toshbaqa, odam, tunets balig'i, tovuq, tunlam, maymun va itning sitoxrom C oqsilidagi aminokislotalar orasidagi farq ko'rsatilgan.

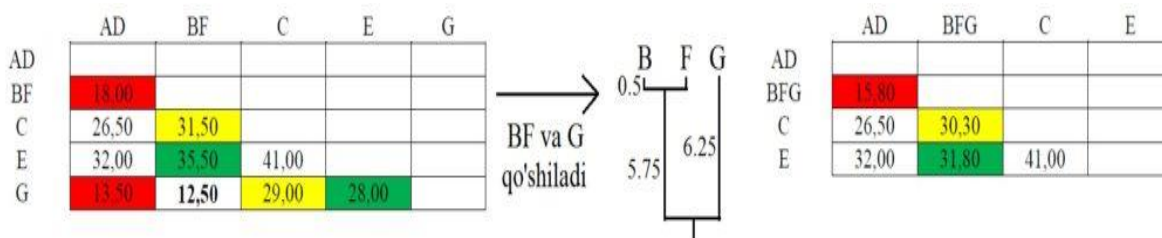




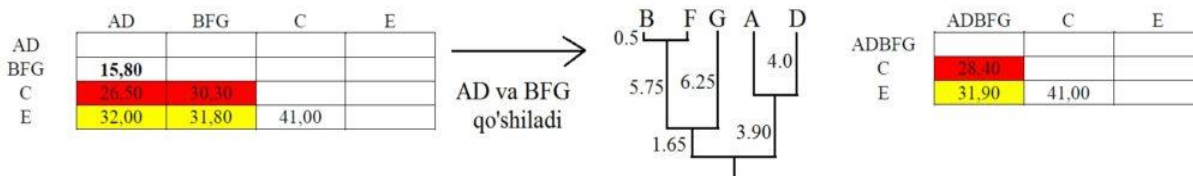
o'zaro 4 ga teng. Bu ko'rsatkich yuqoridagi ular orasidagi farqni ikkiga bo'lish orqali topiladi:  $8/2=4$ .



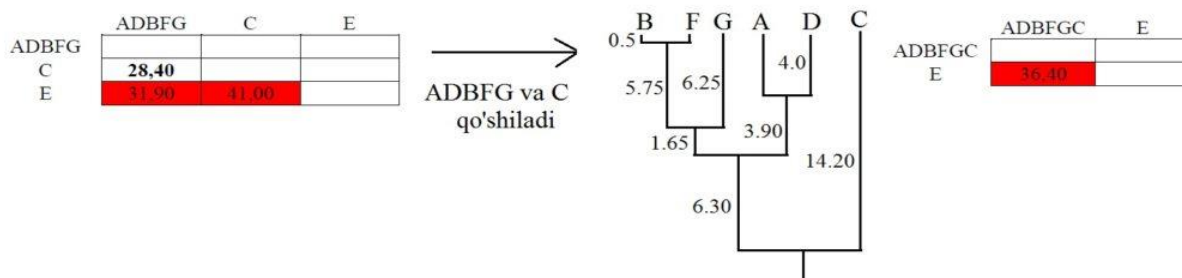
Hosil bo'lgan ushbu yangi jadvaldagi eng kichik sonni izlaymiz, bu yerda u 12,50 bo'lib, BF va G orasida turibdi. Bu degani G filogenetik daraxtda BF ga qo'shni shoxda joylashadi. Hosil qilinadigan jadvaldagi raqamlar ham xuddi yuqoridagi kabi hisoblab boriladi, ya'ni AD va BFG orasidagi raqamni topish uchun avvalgi jadvaldan AD va BF orasidagi raqam AD va G orasidagi raqamga qo'shib, so'ng ikkiga bo'linadi (qizil rangli katakchaga qarang). BF va G filogenetik daraxtda yonma-yon shoxlarda joylashadi deb aytdik. Ular joylashgan shoxning uzunligi ular orasidagi farqli sonni ikkiga bo'lish orqali topiladi:  $12,50/2=6,25$ . Lekin chap tomondagi shoxda BF avvaldan bor edi, ularning shoxlari 0,5 ga teng ekanligini aytgan edik. Demak ushbu holatda 6,25 dan 0,5 ni ayrib, ana undan so'ng hosil bo'lgan sonni ko'rsatamiz. Bu yerda BF uchun umumiy hisoblangan shox G bilan umumiy hosil qilinayotgan shoxgacha 5,75 ga farq qiladi, ya'ni shuncha uzunlikda bo'ladi.



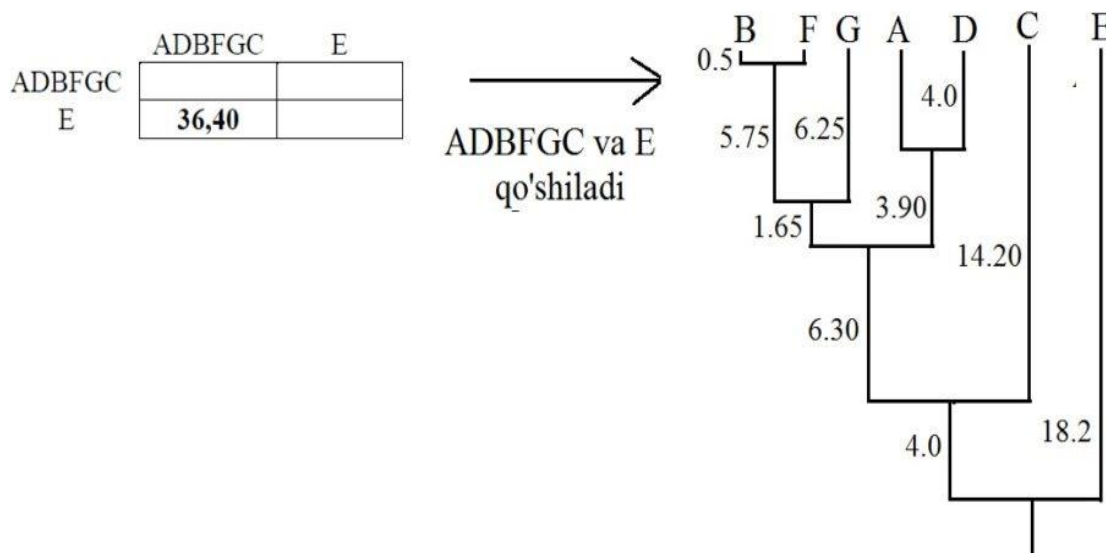
Hosil bo'lgan ushbu jadvaldan ko'rinib turibdiki eng kichik ko'rsatkich AD va BFG orasida turibdi. Demak ularni birlashtiramiz va ko'rsatkichlarni xuddi yuqoridagi kabi hisoblab topamiz. Qaysi katakchalar o'zaro qo'shilayotganligini ularning ranglariga qarab topib, tushunishingiz mumkin. Filogenetik jadvalda BFG va AD shoxlari umumiy shoxda joylashadi. Umumiy shoxgacha bo'lgan masofani ularning o'zaro farqini ikkiga bo'lish orqali topiladi:  $15,80/2=7,90$ . Lekin bir narsaga e'tibor berish lozimki, AD shoxlarning uzunligi 4,0 ekanligini yuqorida aytgan edik, yoki BFG shoxlarining umumiy uzunligi 6,25 ekanligini ham aytgan edik. Shundan kelib chiqqan holda AD va BFG umumiy shoxga mos ravishda 3,90 ( $7,90-4,00=3,90$ ) va 1,65 ( $7,90-6,25=1,65$ ) uzunlikdagi shox orqali birlashadi.

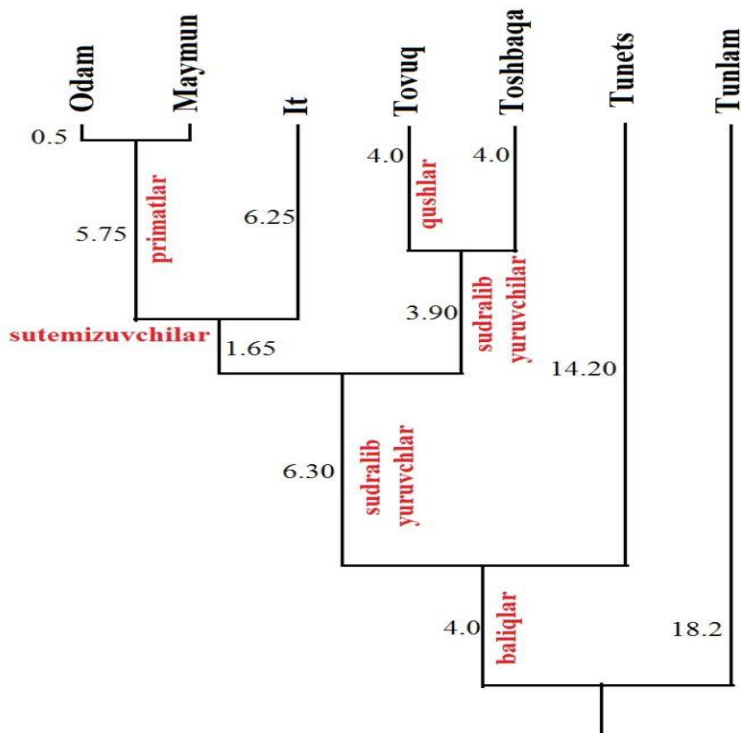


Yuqorida hosil bo'lgan jadvaldagi eng kichik son ADBFG va C orasida joylashgan. Demak ushbu ikki kataklarni xuddi yuqoridagi usulda birlashtiramiz. Filogenetik daraxt shoxini ham yuqorida ko'rsatilganidek quramiz.



E'tibor bersangiz jadvalda faqatgina bir dona raqam qoldi, bu ADBFGC va E orasidagi raqam, demak E ushbu filogenetik daraxtda oxirgi bo'lib ADBFGC shoxlariga qo'shiladi. Ularning orasidagi raqamni ikkiga bo'lish orqali shoxlarning uzunligini topiladi. Shox uzunligi ADBFGC shoxlari uchun tadbiq qilishda yuqorida ko'rsatilgan usul qo'llaniladi.





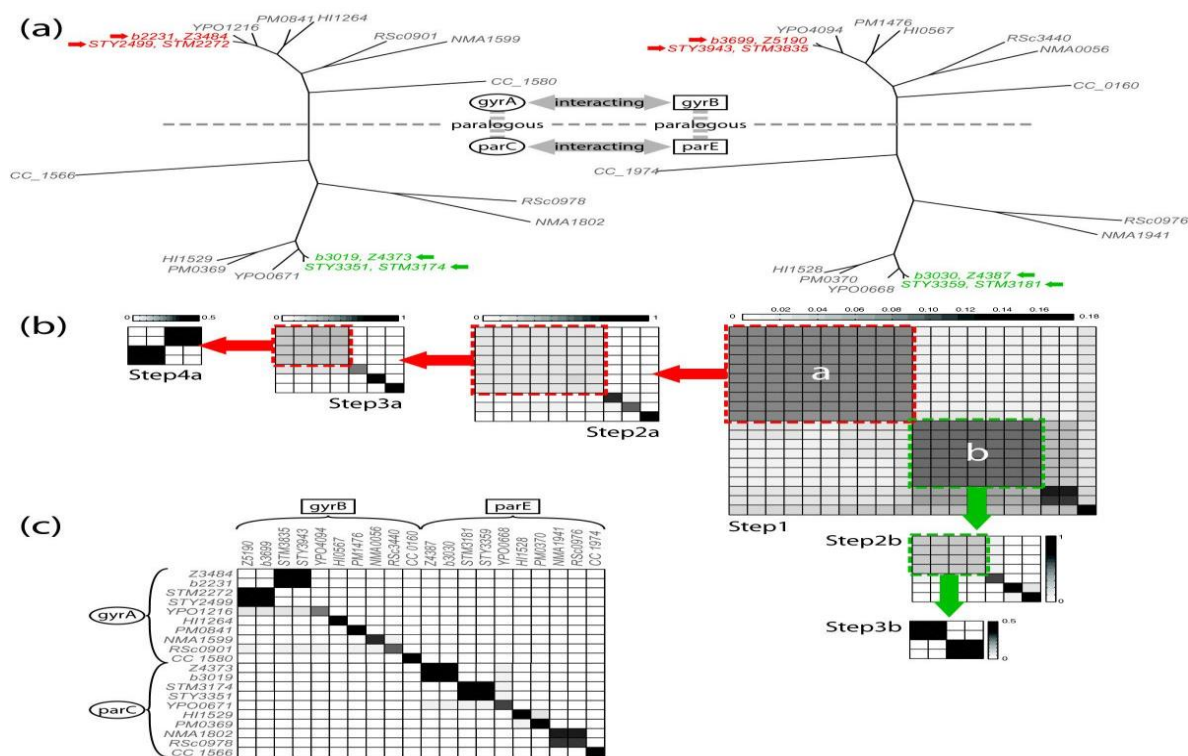
Filogenetik daraxtga e'tibor bersangiz odam va maymun bir shoxda turibdi, undan keyin ularga eng yaqin holatda it qo'shiladi. Odam, maymun, it sutemizuvchilar sinfiga kirgani bois filogenetik daraxtda bir klada (shox) da joylashganini ko'rish mumkin. Uning oldidagi kladada tovuq va toshbaqa joylashgan, demak ular o'zaro qo'shni, ya'ni qushlar va sudralib yuruvchilar sinfi o'zaro umumiy ajdodga ega ekanligi ko'rinib turadi. Sutemizuvchilar joylashgan klada hamda qushlar va sudralib yuruvchilar sinfi joylashgan klada o'zaro birlashib baliqlar sinfi vakillari bilan umumiy ajdodga ega ekanligini namoyon qilmoqda. Demak umumiy ajdod ikkiga ajraladi, bir tarmoqdan baliqlar, ikkinchi tarmoqdan esa sutemizuvchilar, qushlar va sudralib yuruvchilarning umumiy ajdodi rivojlanadi, keyinchalik esa ikkinchi kladaning umumiy ajdodi divergensiya qiladi, ya'ni ikkiga ajraladi.

Masalan, gen daraxtlarini organizm daraxtlari bilan taqqoslash, gorizontaal genlarni o'tkazish kabi nostandart hodisalarni aniqlashga imkon beradi. Turli daraxtlarning taqqoslashidan ko'rinib turganidek, mezbon-parazit simbiozining rasmini berish uchun foydalanish mumkin, masalan, chumolilar, ularning qo'ziqorin turlari va Escovopsis paraziti misolida olamiz. Protein-oqsillarning o'zaro ta'sirini bashorat qilish yana bir misol, chunki o'zaro ta'sir qiluvchi oqsillar ko'pincha bir-biri bilan qo'shib ketishi ko'rinadi. Koevolyutsiyaning bunday holatlari asosan ularning o'zaro ta'sirini (va shu bilan ularning keng funktsionalligini) saqlab qolish uchun bitta gen protein tarkibidagi o'zgarishlar boshqasining o'zgarishi bilan muvofiqlashtiriladi degan fikrga asoslanadi va bu kevolyutsiya jarayoni yoki korrelyatsiya qilingan

evolyutsiya bo'lishi mumkin. ularning filogenetik daraxtlarining o'xshashligi orqali kuzatiladi.

Daraxtlarni taqqoslash uchun turli xil usullar mavjud bo'lsa-da, ikkita umumiy toifani aniq ajratib ko'rsatish mumkin.

- **Birinchi darajali** yondoshish daraxtlarni topologik xususiyatlar bo'yicha taqqoslashga qaratilgan, masalan, bir juft daraxtlar orasidagi o'zaro taqsimlangan va taqsimlanmagan substruktsiyalar sonini (masalan, to'rt bargli tugunlarning pastki qismlarini) aniqlash yoki minimal operatsiyalar sonini aniqlash. (masalan, yaqin qo'shnilarning almashinuvi) bitta daraxtni boshqa daraxtga aylantirish uchun
- **Ikkinchi yondashuvlar** masofa yoki yo'l uzunligi to'g'risidagi ma'lumotni to'g'ridan-to'g'ri taqqoslaydi.



Rasmlarda Filogenetik daraxtlarni taqqoslash uchun yangi yondashuvni va uni qo'llashning bir nechta namunalarida, Birinchidan, oqsillarning o'zaro ta'sirini bashorat qilishda boshqa mavjud yondashuvlar bilan taqqoslaganda rCEED vCEED yondashuvidan foydalangan holda bashorat qilish aniqligi yaxshilanganligi, bu ikkala usulda yuqori o'xshashlik a'zolar o'rtasida jismoniy o'zaro ta'sirni talab qilmasligi, lekin bu faqat mumkinligini ko'rsatadi. Xuddi shunday, ikkita oqsil o'rtasidagi jismoniy ta'sir kevolyutsiyani talab qilmaydi. Shunday qilib, bu yerda keltirilgan koevolyutsion yondashuvlar ma'lum bir tur doirasidagi to'liq interaktomning faqat bir qismini aniqlay oladi. Proteinlarning o'zaro ta'sirini kuchaytirish uchun bashorat qilish uchun, rCEED

/ vCEED kabi yondoshuvlar boshqa hisoblash usullari bilan birgalikda eng yuqori samaradorligini ko'rsatishi mumkin.

### **Clustal dasturi va uning turlari.**

**Clustal** - bu ketma-ketlikni tenglashtirish uchun bioinformatikada ishlatiladigan keng tarqalgan kompyuter dasturlari, algoritmi ishlab chiqish bo'yicha Clustalning ko'plab versiyalari mavjud. Har bir vositaning tahlili va uning algoritmi ham tegishli toifalarida batafsil bayon qilingan. Clustal vositalarining kombinatsiyasi mavjud bolib har bir joriy versiyasida qo'llab-quvvatlanmasligi mumkin.

**Clustal Omega** barcha Clustal vositalaridan eng keng doiradagi operatsion tizimlarga ega. Clustal dasturining turli xil variantlari mavjud bo'lib, ularning hammasi quyida keltirilgan:

**Clustal2** -1988 yilda Des Xiggins tomonidan yaratilgan bir nechta ketma-ketliklarni moslashtirish uchun original dastur filogenetik daraxtlarni aminokislotalar yoki nukleotidlarning juft ketma-ketliklaridan olishga asoslangan. **ClustalV**: Clustal dasturining ikkinchi avlodi 1992 yilda chiqarilgan va original Clustal to'plamini qayta yozgan. U oxirgi hizalama bo'yicha filogenetik daraxtlarni qayta qurish, mavjud hizalama-dan burmalar yaratish qobiliyati va qo'shnilar qo'shilishi deb nomlangan usul yordamida hizalamalardan daraxt yaratish variantini kiritdi.

**ClustalW**: 1994 yilda chiqarilgan uchinchi avlod avvalgi versiyalarga qaraganda ancha yaxshilandi. Bu progressiv algoritm bo'yicha turli yo'llar bilan yaxshilandi, shu jumladan qisman hizalanishda o'xshashlik yoki tafovutlarga ko'ra individual ketma-ketliklarni pastga yoki yuqoriga ko'tarish imkonini berdi. Shuningdek, dasturni buyruq satridan ommaviy rejimda ishlatish imkoniyati ham mavjud edi. **ClustalX**: 1997 yilda chiqarilgan ushbu versiya birinchi bo'lib foydalanuvchi grafik interfeysiga ega edi. **ClustalΩ (Omega)**: hozirgi standart versiya hisoblanadi.

**Clustal2: ClustalW va ClustalX**-ning yangilangan versiyalari yuqori aniqlik va samaradorlik bilan ishlaydi.**Clustal dasturiy** ta'minotni tavsiflovchi hujjatlar juda yuqori darajada keltirilgan, ulardan ikkitasi barcha davrlarning eng iqtibos qilingan gazetalari qatoriga kiritilgan. Windows, Mac OS va Unix / Linux uchun mavjud bo'lgan dasturiy ta'minotning eng so'nggi versiyasi va bundan tashqari, u veb-interfeys orqali o'z uy sahifasida yoki Evropa Bioinformatika Institutida joylashgan.

### **T-Coffe dasturi imkoniyatlari**

**T-Coffee** (moslashtirishni baholash uchun daraxtga asoslangan muvofiqlik ob'ektiv funksiyasi) - bu progressiv yondashuvdan foydalangan holda ketma-ketlik bo'yicha tasvirlanadigan dasturi. Bir nechta ketma-ketlik bo'yicha izlanishni boshqarish uchun juft yo'nalishli tasvirlanish kutubxonasi yaratadi. Bundan tashqari, u ilgari olingan va ketma-ket keladigan ketma-ketlikdagi izlanishlarni birlashtirishi



mumkin va PDB fayllaridan tarkibiy ma'lumotlardan (3D-Coffee) foydalanishi mumkin.

U hizalaishlar sifatini va motiflarning paydo bo'lishini aniqlash uchun bir qator imkoniyatlarni baholash uchun ilg'or xususiyatlarga ega (Mokka). Sukut bo'yicha aln formatida (Clustal) hizalama ishlab chiqaradi, lekin PIR, MSF va FASTA formatlarini ham ishlab chiqarishi mumkin. Eng keng tarqalgan kirish formatlari qo'llab-quvvatlanadi (FASTA, PIR).

Kutubxonada kiritilgan ketma-ketlik o'rtasidagi o'zaro moslashuv moslamalari ma'lumotlarini ko'rib chiqqani uchun, algoritm **ClustalW-ga** qaraganda aniqroq ekanligi xabar qilingan. Shunday qilib, u ko'proq turli xil ketma-ketlikni tasvirlanish uchun uchun foydali deb hisoblanadi.

### **Nazorat savollari**

1. Bioinformatikada ko'plik taqqoslanish nima?
2. Filogenetik daraxt nima?
3. Filogenetik daraxt qanday tuziladi?
4. Clustal qanday dastur?
5. Clustal dasturining qanday versiyalari mavjud?
6. T-COFFEE qanday dasur?

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
2. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

1. <https://www.ncbi.nlm.nih.gov/orffinder/>
2. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
3. (<https://www.ncbi.nlm.nih.gov/genbank/>)

## **8-Amaliy mashg'ulot.**



**Mavzu:** Biologik makromolekulalarni vizualizatsiya qilish dasturlari. Aminokislota ketma-ketligi bo'yicha fazoviy strukturani yaratish.

**Dars maqsadi:** Talabalarga Biologik makromolekulalarni vizualizatsiya qilish dasturlari haqida ma'lumot berish hamda aminokislota ketma-ketligining fazoviy strukturalarini yaratish bo'yicha tushuncha berish.

**Kerakli jihoz va preparatlar:** Kamyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** Molekulalarning strukturaviy xossalari ko'p sohalarda birinchi o'rinda turadi. Ushbu hisobot molekulyar grafika va vizualizatsiya sohalarida ishlab chiqilgan dasturlar haqida to'liq ma'lumot beradi. Bunda asosiy e'tibor strukturaviy biologiyada qo'llaniladigan sohalar ko'p jihatdan uch o'lchovli, murakkab, katta va vaqt ta'sirida o'zgaruvchan molekulyar tuzilmalarning kompyuterlashtirilgan geometrik va vizual tasvirlariga tayanadi. Hisobotda molekulyar vizualizatsiyaning qaysi sohalarida allaqachon keng qamrovli o'rganilganligi va sohaning rivojlanayotganligi ko'rsatadigan taksonomiya taqdim etilgan. U molekulyar tuzilmalar uchun vizualizatsiyani, tasvir sifati va kadr tezligini samarali namoyish qila olish strategiyalarini muhokama qiladi hamda tafsilotlar darajasining turli tomonlarini qamrab oladi va molekulyar simulyatsiya ma'lumotlarining dinamik jihatlarini aks ettiruvchi vizualizatsiyalarni ko'rib chiqadi. Interaktiv molekulyar vizualizatsiya, ma'lumotlar bazalarining vizualizatsiyasi ichida eng qadimgi tarmoqlaridan biri bo'lib, kompyuterdan oldingi davrda chuqur o'rganilgan edi. Ushbu maqola so'nggi yigirma yil ichida eng ko'p rivojlangan va yaqinda paydo bo'lgan soha biomolekulyar tuzilmalarning interaktiv vizualizatsiyasini ko'rib chiqadi.

*Molekulyar vizualizatsiyaning* asosiy maqsadi molekulyar tuzilmalar, ularning xususiyatlari va o'zaro ta'sirini tushunarli qilish orqali boy, murakkab moddiy dunyo haqidagi tushunchamizni qo'llab-quvvatlashdir. Bundan tashqari, u farmatsevtik faol birikmalar yoki o'ziga xos xususiyatlarga ega moslashtirilgan moddalar kabi yangi molekulalarning "ratsional" dizaynini qo'llab-quvvatlashga qaratilgan.

Molekulalarning uch o'lchamli tuzilmalarini vizualizatsiya qilish uchun dasturiy ta'minot, *molekulyar vizualizatsiya dasturini* molekulyar modellashtirish dasturidan ajratish mumkin. To'g'risini aytganda, *Vizualizatsiya* dasturi oldindan mavjud bo'lgan molekulyar modelni o'zgartirmasdan ko'rsatadi, *modellashtirish* dasturi esa modelni yaratishi yoki qismlarni qo'shish yoki olib tashlash, kovalent aloqalarni, bog'lanish burchaklarini, konformatsiyani yoki kovalent bo'lmagan o'zaro ta'sirlarni o'zgartirish orqali uni o'zgartirishi mumkin.

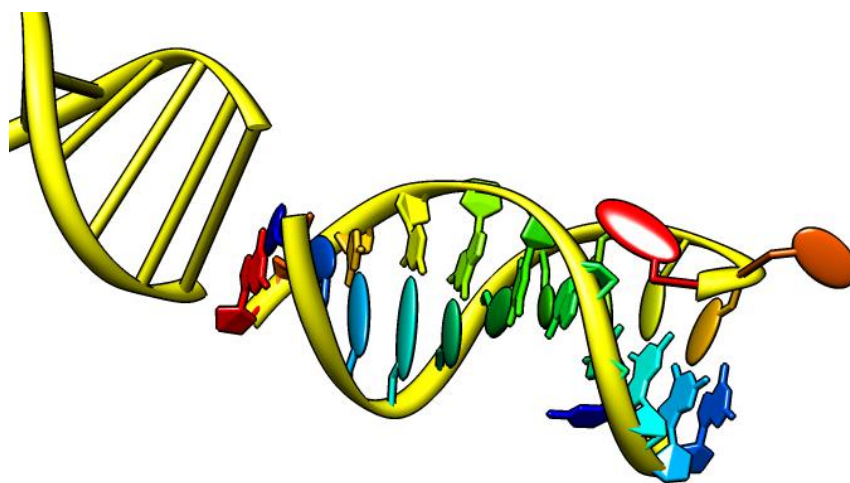
**Bepul molekulyar vizualizatsiya dasturlari.** Mashhur bepul molekulyar vizualizatsiya dasturiy paketlariga misollar:

- **Jmol**, java-ga asoslangan ochiq manbali dastur mustaqil yoki applet shakllarida mavjud. Applet Proteopedia-da va **Molecular Workbench** bepul o'quv dasturida qo'llaniladi. Jmoldan samarali foydalanish buyruq skript tilini o'rganishni talab qiladi.

- [FirstGlance in Jmol](#) , Jmol uchun ochiq manbali foydalanuvchi interfeysi [Nature](#) jurnalidagi maqolalarda yangi makromolekulyar tuzilmalar haqida xabar beruvchi *3D View havolalarida foydalaniladi*. Proteopedia-ning avtomatik ravishda ekilgan sahifalarida ([PDB kodi bilan sarlavhali sahifalar](#)) Jmol in First Glance-da strukturani yanada o'rganish uchun havolalar mavjud . Jmoldagi *FirstGlance* hech qanday holatda Jmoldagi barcha quvvatdan foydalanmasa ham, u makromolekulaning asosiy strukturaviy xususiyatlarini **hech qanday buyruq skripti tilini o'rganmasdan** ko'rsatadi . [Ko'proq ...](#)
- [Chimera](#) (faqat notijorat maqsadlarda foydalanish uchun bepul).
- [Kinemages, Mage va KiNG](#) (KiNG = Kinemage, Keyingi avlod) molekulyar tuzilishga muallifning nuqtai nazarini taqdim etish uchun mo'ljallangan. Ular ixtisoslashgan sohalarda ustunlik qiladi, lekin umumiy maqsadli vizualizatsiya uchun maqbul emas. KiNG Proteopedia-da ishlatilishi mumkin; misolini [Gemoglobin #Gemoglobin\\_subbirligi\\_bog'lash\\_O<sub>2</sub>](#) da ko'rish mumkin .
- [RasMol](#) , 1993 yilda chiqarilgan ochiq manbali mustaqil dastur va hali ham mashhur. RasMol-dan samarali foydalanish buyruq skript tilini o'rganishni talab qiladi.
- [Protein Explorer](#) , [Chime](#) uchun keng va kuchli ochiq manbali foydalanuvchi interfeysi . 2009-yilda Protein Explorer-ning quvvati va foydalanish qulayligini birlashtirgan boshqa hech narsa mavjud bo'lmasa-da, [Chime](#) brauzer plaginini o'rnatish va ishlatish bilan bog'liq qiyinchiliklar tufayli u ishlatilmaydi. Jmol hozirda ancha kuchliroq va ishga kirishish osonroq bo'lgani uchun, ayniqsa OS X da, Protein Explorer-ga texnik xizmat ko'rsatish 2007 yilda to'xtatildi. [Protein Explorer-ni Jmol bilan amalga oshirish boshlandi](#) , ammo bu loyiha boshlang'ich bosqichida. (Ko'ngillilar kerak!)
- [Chime](#) , 1996-yilda chiqarilgan bepul brauzer plagini, endi uning o'rniga [Jmol](#) . Ochiq manba emas.
- Molsoft tomonidan ishlab chiqarilgan [ICM-brauzer](#) va bepul [ActiveICM](#) plagini endi [Molecular and Cellular Proteomics Journal](#) va [PLoS \(Public Library of Science\)](#) jurnallarida takomillashtirilgan tarkibni ko'rish uchun ishlatiladi , masalan, [kengaytirilgan versiyada ko'rish mumkin](#). brauzer va alohida [Datapack bilan ushbu misol maqola](#) . Yuklab olish va/yoki plaginni talab qiladi. Ochiq manba emas.
- [HOLLOW](#) mustaqil molekulyar vizualizatsiya dasturi emas; ammo oqsillarning sirt tasvirlarini ishlab chiqarishni osonlashtiradi. Siz [PDB](#) faylini kiritasiz va oqsil bo'shliqlari va kanallarining "quymasini" tashkil etadigan qo'g'irchoq suv atomlarining [PDB](#) faylini olasiz.

- **RINalyzer** Protein ma'lumotlar bankidagi PDB fayllarida saqlanadigan oqsilning uch o'lchovli tuzilishidan tuzilgan qoldiq o'zaro ta'sir tarmoqlarini (RIN) tahlil qilish va vizualizatsiya qilish imkonini beradi. RINalyzer bir vaqtning o'zida, interaktiv 2D vizualizatsiya va RINni **Cytoscape'da** , **UCSF Chimera** vizualizatsiya/modellashtirish tizimida ko'rsatilgan molekulyar 3D strukturasi bilan birgalikda o'rganish imkonini beradi. **Cytoscape** - molekulyar o'zaro ta'sir tarmoqlarini tahlil qilish va vizualizatsiya qilish uchun bepul, ochiq kodli dasturiy platforma. Har qanday PDB yozuvi uchun RIN ma'lumotlarini **PDB kodidan** foydalanib yuklab olish **mumkin** , muqobil ravishda **RING veb-serveridan** RIN yaratish uchun foydalanish mumkin.

**Jmol First Glance dasturi.** Proteopediadagi har bir molekula molekula ostidagi havoladan foydalangan holda (1d66 kabi PDB kodi bilan sarlavhalangan sahifalarda ) Jmol'da First Glance - da o'rganilishi mumkin .Jmol-dagi FirstGlance-ga to'g'ridan-to'g'ri kirganingizda , istalgan PDB kodini kiritishingiz yoki kompyuteringizdan PDB faylini yuklashingiz mumkin, masalan, AlphaFold tomonidan bashorat yoki homologiya modeli .FirstGlance in Jmol ( [firstglance.jmol.org](http://firstglance.jmol.org) ) bu veb-brauzerda on-layn rejimda ishlaydigan ochiq, ochiq manbali, makromolekulyar vizualizatsiya dasturiy ta'minot to'plami.



- narvon zinapoyalari
- to'ldirilgan halqali atom tasvirlari
- "lolipoplari" bo'lib, unda asoslar ellipsoidlar va shakarlar naychalar sifatida ko'rsatilgan.

Bazalar, shuningdek, yo'nalishni ko'rsatish uchun bo'rtmalari bo'lgan yoki bo'lmasdan qutilar yoki elliptik naychalar sifatida ko'rsatilishi mumkin. Maxsus tasvirlarning ranglari mos keladigan atomlarga mos kelishi uchun avtomatik ravishda yangilanadi.

## Nazorat savollari

1. **Molekulyar vizualizatsiya deganda nimani tushunasiz?**
2. Vizual tasvirlarni yaratadigan qanday dasturlarni bilasiz?
3. Jmol qanday dastur?
4. Jmolning qanday shakllari bor?
5. CHimera qanday dastur?
6. CHimera qanday versiyalari mavjud

## Informatsion-uslubiy ta'minot

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

## QO'SHIMCHA ADABIYOTLAR

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
3. А. Н. Огурцов Введение в Бионформатику, 2011г.

## Internet ma'lumotlari

1. <https://www.ncbi.nlm.nih.gov/orffinder/>
2. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
3. <https://www.cgl.ucsf.edu/chimera/>
4. <https://proteopedia.org/wiki/index.php/Chimera>
5. [https://proteopedia.org/wiki/index.php/Molecular\\_modeling\\_and\\_visualization\\_software](https://proteopedia.org/wiki/index.php/Molecular_modeling_and_visualization_software)
6. <https://proteopedia.org/wiki/index.php/Jmol>

## 9-Amaliy mashg'ulot.

## **Mavzu:** I – TASSER va Modeller dasturlarining ishlash prinsipi

**Dars maqsadi:** Talabalarga I – TASSER va Modeller dasturlarining ishlash prinsipi haqida ma'lumot berish.

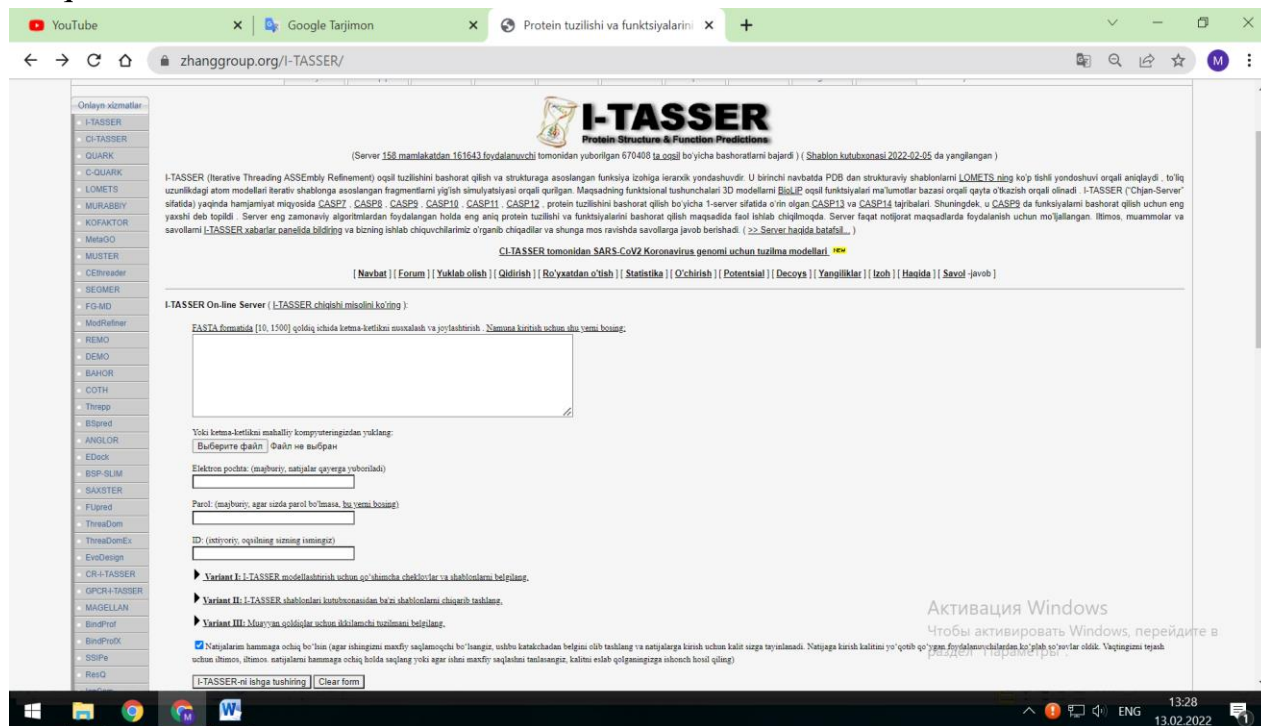
**Kerakli jihoz va preparatlar:** Kampyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** I-TAISSER yangi rivojlangan oqsil tuzilishi va funksiyasini bashorat kiluvchi server. I-TAISSER serveri avtomatlashtirilgan oqsil strukturasi bashorat qilish va tuzilishini kursatishga asoslangan onlayn resurs hisoblanadi. I-TAISSERda strukturaviy shablonlar birinchi navbatda PDB dan bir nechta threading alignment yondashuvlari yordamida tan olinadi. To'liq uzunligi, tuzilishi modellari, keyin interaktiv qism montaj stimulyatorlari tomonidan qurilgan. Funktsional tushunchalar, nihoyat, bashorat qilingan tuzilish modellarini funktsiya bazalaridagi ma'lum oqsillar bilan moslash orqali olinadi. Server turli biologik va biomedikal tekshiruvlar uchun keng qo'llanilgan bo'lsada, foydalanuvchi jamoasidan ko'plab sharhlar va takliflar bildirilgan. Ushbu maqolada I-TAISSERserveridagi so'nggi ishlanmalarni umumlashtiramiz, ular foydalanuvchilar jamoasidan qo'yilgan talablarni hal qilish va modellashtirish bashoratlari aniqligini oshirish uchun mo'ljallangan edi.

**I-TAISSERga kirish.** Oqsil taraqqiyoti tuzilishini bashorat qilish bilan, molekulyar va sitologik tadqiqotchilar eksperimental tadqiqotlar o'tkazishdan oldin ularning oqsillar uchun avtomatlashtirilgan server bashorat izlab muntazam izlanishadi. Jamiyat ko'p CAPS tajribalar bunday I-TAISSER, Rosetta va HHpred kabi ko'plab avtomatlashtirilgan serverlar, endi eng yaxshi inson ekspert modellashtirish bilan solishtirish, aniqlik bilan tarkibiy modellari mumkin 'kanligini ko'rsatdi. Boshlab computational tuzilishi modellashtirish, bir qator usullari taklif etildi. Shuning uchun annotate biologik funktsiyasi oqsil molekulalar, misollar bilan, shu jumladan, COFACTOR, MURABBIY, ConCavity, FINDSITE, Firestar va 3D ligandsite, o'rtasida va boshqalar. Jonli dastgoh asosidagi platforma yaqinda ligand-majburiy sayt bashoratini baholashga kirishdi, bu esa avtomatlashtirilgan oqsil funktsiyasi izohlarining aniqligi va foydaliligini namoyish etdi.

Bu foydalanishga qaramay, biz tizimni yaxshilash yo'llari haqida, foydalanuvchi hamjamiyat ko'plab izoh va takliflar oldik. Ushbu maqolada I-TAISSERserveriga qilingan so'nggi ishlanmalar haqida xabar berib o'tamiz, ular I-TAISSERmodellashtirishning sifati va server tizimining funksionalligini keskin yaxshilagan. Asosiy yangi o'zgarishlar o'z ichiga oladi: (1) omil biolog foydalanuvchilar tomonidan funktsional tadqiqotlar takdim qilish uchun muhim bo'lgan tarkibiy modellari qoldiq darajali mahalliy sifatini baholash, yangi yondashuv; (2) omil bashorat uchun algoritmi; (3) atom tuzilishli darajalarini tozalash

usullari I-TASSERmodellari vodorod ulash tarmoqlari va jismoniy realizmni yaxshilash; (4) konsensusga asoslangan ligand majburiy sayt tuzilishi va taqqoslash profili (5) proteinning funktsional izohlari; va (6) foydalanuvchi hamjamiyati bilan muhokama va muloqotni engillashtirish uchun yangi xabar kengashi tizimini ishlab chiqish.



I-TASSER serverining ish stoli

## I-TASSER matriallari va usullari

### I-TASSER quvurining umumiy ko'rinishi

I-TASSER serveri I-TASSER asosida qurilgan bo'lib, oqsil tuzilishi va funksiyasini bashorati uchun ierarxik Andoza asosidagi uchta usulli umumiy qadamdan iborat.

1). Berilgan so'rovlar ketma-ketligi uchun I-TASSERavval PDB kutubxonasi LOMETs yordamida tizimli shablonlar yoki super ikkilamchi struktura motivlarini, bir nechta uchli algoritmlardan iborat meta-uchli dasturni aniqlaydi. To'liq metrajli modellari topologiyasi keyin andozalari eksizi doimiy qism tuzilmalarni qayta qurilgan tomonidan, qaerda mos bo'lmagan viloyatlar tuzilmalari nusxa almashish o'rnatilgan Karlo stimulyatorlari asosida AB initio tomonidan noldan qurilgan. Tuzilishi traektoriyalar kam erkin energiya holatlarini aniqlash uchun SPICKER tomonidan kiritilgan. SPICKER klasterlaridan boshlab strukturaviy modellarni takomillashtirish uchun ikkinchi tur struktura reassemblyasi o'tkaziladi. Kam erkin energiyali konformatsiyalar FG-MD va Modrefiner yordamida to'la atomli simulyatsiyalar orqali yanada tuldirdildi.



Nihoyat, proteinning funktsional tushunchalari strukturaviy modelni BioLip funktsiyasi kutubxonasiidagi oqsillar bilan tuzilishi va ketma-ketlik Profil taqqoslashlari orqali moslab olinadi.

I-TAISSERquvuri CAPS tajribalarida Zhang-Server tomonidan ishlatiladigan yondashuv bilan bir xil bo'ladi. CAPS9 ammo, yangi ammo initiostruc ture bashorat yondashuv, QUARKdan va qattiq bepul modellashtirish (FM) maqsadlar uchun andozalar tartiblashtirish uchun Zhang-Server quvuri joriy etildi. QUARK dasturi hali I-TAISSERserveriga qo'shilmagan, bu ab initio katlama maqsadlari uchun mustaqil onlayn server sifatida mavjud .

I-TAISSER tomonidan qoldiq maxsus sifat baholash I-TAISSERtomonidan hosil bulgan ularning tuzilishi modellari bilan 635 oqsillarni ortiqcha bo'lmagan majmui bo'yicha baholandi, qaerda I-TAISSERmodellari taxmin va kuzatilgan masofa xatolar o'rtasidagi o'rtacha farq 1.4 Å hisoblanadi 506 C-hisobida bilan oqsillar uchun shuningdek, CAPS9 va CAPS10 tajribalarida turli xil prediktorlar tomonidan yaratilgan tuzilish modellari bilan sinovdan o'tkazildi. Natijalar Resq uchun teng ekanligini ko'rsatdi, yoki eng ortda, model sifat baholash dasturlari (Mqaps) mahalliy tuzilishi sifat baholash uchun (qo'shimcha jadvallar S1–S4 qarang). Resq dasturidan ham I-TAISSER, ham boshqa strukturani bashorat qilish usullari bilan hosil qilingan struktura modellarining aniqligini baholash uchun foydalanish mumkin. I-TAISSER Suite orqali mustaqil paketi bir ko'chirib sifatida onlayn server orqali yoki kirish mumkin.

B-omil bashorati. Yana bir juda tegishli, lekin ko'pincha-o'tkazib yuborilgan mahalliy xususiyati tuzilishini bashorat quvurlari oqsillarni qoldiqlari tabiiy harakati issiqlik hisoblanadi. Mutlaq nol haroratda oqsildagi atomlar eng past energiyaning muvozanat holatida qolishi taxmin qilinadi; ammo harorat ortishi bilan atrof-muhit issiqlik energiyasi atomlarning muvozanat holati atrofida tebranishiga olib keladi, ularning miqdori ko'pincha 3D tuzilishi va ligand va hal qiluvchi atomlar bilan o'zaro ta'siriga qarab o'zgaradi. Atom harakati eksperimental rentgen kristallografiyasida b-faktor (yoki harorat omili) sifatida o'lchanishi mumkin, bu struktura omili tenglamasiga o'zgartirish omili sifatida kiritilgan, chunki rentgen nurlarining sochilish ta'siri qolgan atomlarga nisbatan tebranuvchi atomlarga kamayadi. Oqsil kristallarida issiqlik harakat omillarining taqsimlanishiga eksperimental rezolyutsiya, kristalli kontaktlar va tozalash protseduralari kabi sistematik xatolar ta'sir etishi mumkin, chunki b-omil qiymatlari odatda turli eksperimental tuzilmalar o'rtasida taqqoslanmaydi. Shuning uchun ta'sirni kamaytirish uchun 3 balli transformatsiyaga ega bo'lgan normalangan b-omilni hisoblaymiz.

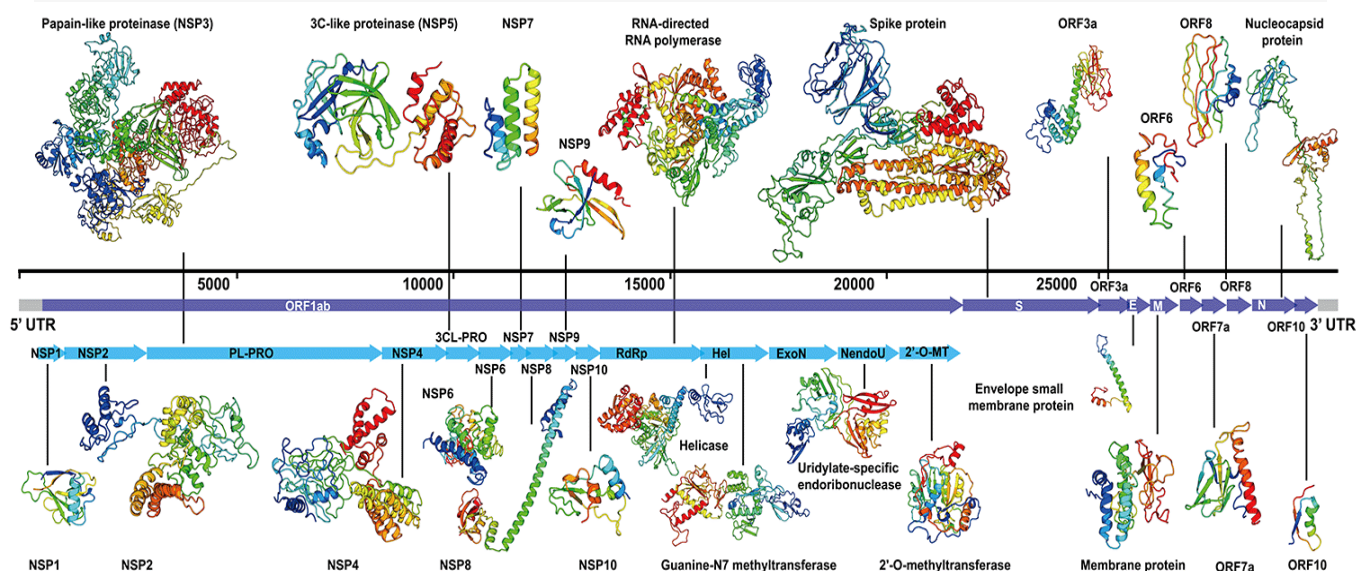
Normalangan b-omil natija Profil ish bilan ta'minlangan Andoza asoslangan topshiriq va mashina ta'lim asoslangan bashorat kombinatsiyasi yordamida b-omil bashorati va tarkibiy xususiyatlarini bashorat qilindi (Yang et al., taqdim etilgan).



Yuqorida aytib o'tilgan 635 oqsillar bo'yicha testlar b-omil tomonidan taxminan b-omil x-ray kristallografiya ma'lumotlariga nisbatan 0.79 egri (AUC) ostida o'rtacha maydoni bor, deb ko'rsatadi.

Atom darajasi tuzilishi aniqlash. Oldingi I-TASSERmodellashtirish versiyaga kelib tushadigant shikoyatlardan biri tarkibiy modellari mahalliy tuzilishiga kelardi. Masalan; misol uchun, ba'zi modellar xayoliy o'rta tuzilishi xususiyatga ega yoki sterik to'qnashuvlar bor edi, global topologiya odatda to'g'ri bo'lsa-da,. Bu tez-tez tarkibiy montaj stimulyatorlarga mos kelmaydigan threading alignments tufayli farqsiz decoys moyil nodavlat homolog protein maqsadlar uchun sodir bo'ladi. Shuning uchun, diverged decoy tuzilmalardan birlashgan tuzilishi modellari buzib mahalliy tuzilishi uzoqdan olinardi.

### SARS-CoV-2 virusining genom bo'yicha tuzilishi va funksiyasini TASSER da modellashtirish



Ushbu sahifada SARS-CoV-2, shuningdek, 2019-nCoV nomi bilan ham tanilgan, COVID-19 pandemiyasini keltirib chiqargan yangi koronavirus genomi tomonidan kodlangan barcha oqsillar uchun 3D strukturaviy modellar va funktsiya izohlari mavjud. Struktura modellari DI-TASSER/ CI-TASSER quvur liniyasi tomonidan yaratilgan bo'lib, u I-TASSER fragmentlarini yig'ish simulyatsiyasini boshqarish uchun chuqur konvolyutsion neyron tarmoqqa asoslangan masofa xaritasi/kontakt xaritasi bashoratlardan foydalanadi. Benchmark va ko'r CASP testlari shuni ko'rsatdiki, DI-TASSER/CI-TASSER I-TASSERga qaraganda ancha yuqori aniqlikdagi modellarni yaratadi, ayniqsa gomologik shablonlarga ega bo'lmagan oqsil maqsadlari uchun. Ko'p domenli maqsadlar uchun alohida domenlarning DI-TASSER/CI-TASSER tuzilmasi DEMO tomonidan yig'iladi.to'liq uzunlikdagi tuzilishga.

Ikki usul turli maqsadlar modellarini takomillashtirish uchun I-TASSERquvuriga kiritilgan. Umuman, FG-MD gomolog andozalari bilan oson maqsadlar uchun ishlatiladi va Modrefiner FG-MD tomonidan ta'qib qilinadi, to'liq atom modellari ishlab chiqarish uchun gomolog andozalari holda qattiq maqsadlar uchun ishlatiladi. Benchmark va CASP testlaridan olingan ma'lumotlar shuni ko'rsatdiki, I-TASSERmodellarining HB-score (vodorod-bog'lanish tarmog'ining sifatini o'lchash) va molprobitiy score (steric clash va Ramachandran muntazamligini o'lchash) rafinatsiya simulyatsiyalari bilan sezilarli darajada yaxshilanishi mumkin.

### **Nazorat savollari**

1. I-TAISSEER nima?
2. Modeller nima?
3. I-TAISSEER ga qanday kiriladi?
4. I-TAISSEER Atom darajasi tuzilishi aniqlash qanday amalga oshiriladi?

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
4. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

1. <https://zhanggroup.org/I-TASSER/>
2. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
3. <https://www.cgl.ucsf.edu/chimera/>
4. <https://proteopedia.org/wiki/index.php/Chimera>
5. [https://proteopedia.org/wiki/index.php/Molecular\\_modeling\\_and\\_visualization\\_software](https://proteopedia.org/wiki/index.php/Molecular_modeling_and_visualization_software)

### **10-Amaliy mashg'ulot.**

**Mavzu:** RasMol dasturida oqsil strukturalarining qo'yilishi.

**Dars maqsadi:** Talabalarga RasMol dasturida oqsil strukturalarining qo'yilishi haqida amaliy ko'nikma berish.

**Kerakli jihoz va preparatlar:** Kamyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** Rasmol molekulyar grafik tasvirlar uchun mo'ljallangan kompyuter dasturi va asosan biologik makromolekula tuzilmalarni tasvirlash uchun ishlatiladi. U dastlab 1990-yillar boshida Rojer Sayle tomonidan ishlab chiqilgan.

Rasmol dastlab Rojer Sayle tomonidan ishlab chiqilgan molekulyar grafik ko'rgazma uchun dastur hisoblangan. Ushbu sayt Rasmol foydalanuvchilar va Rasmol ochiq manba versiyalari ishlab chiquvchilar uchun qulaylik uchun taqdim etiladi. Rasmolni ta'minlash, rivojlantirish ko'plab hamjamiyatlar tomonidan o'zgartirishlar integratsiya amalga oshiriladi.

Source Forge tomonidan yaratilgan sayt orkali jamiyat a'zolari bizni saytga a'zo bo'lishlari mumkin.

Rasmol ustida ishlash rasman Aqsh energetika departamenti, Aqsh Milliy Fanlar jamg'armasi va Aqsh (NIH) Milliy tibbiyot fanlari instituti grantlari tomonidan qo'llab quvvatlandi. Ushbu materialda bildirilgan har qanday fikr, topilmalar va xulosalar yoki tavsiyalar muallif(lar)dir va moliyalashtirish idoralarining fikrlarini aks ettirmaydi.

Manba kodi va hujjatlar mavjudligini ta'minlash maqsadida ushbu sayt dasturlari va hujjatlar mualliflik huquqiga tortiladi. Bu Rasmolning ochiq kodli versiyalaridan foydalanishda, nusxa va o'zgartirishlar kiritishga to'sqinlik qilmaydi, balki ochiq kodli versiyalardan "yopiq manba" versiyalarini yaratishga to'sqinlik qiladi. Tegishli mualliflik huquqi va litsenziyalar tegishli manbalar va hujjatlar bilan paydo buladi.

**Rasmol va Open RasMol web saytiga xush kelibsiz.** Ushbu sayt 2000 yilning sentyabr oyida tashkil etilgan bulib, Rasmol ochiq manba versiyalari ishlab chiqarila boshladi va 2002 yilning may oyidan boshlab Rasmol foydalanuvchilari uchun bosh sahifaga aylandi

Rasmol 1992-yilda Rojer Sayle tomonidan yaratilgan molekularni vizuallashtirishning muhim ilmiy vositasi hisoblanadi. Rasmol makromolekularni ko'rish va nashr sifatli tasvirlarni tayyorlash uchun yuz minglab foydalanuvchilar tomonidan jahon miqyosida qo'llaniladi.

Biz foydalanadigan vositalar ushbu vositalarni ishlatadiganlar va ushbu vositalar bilan olingan natijalardan foydalanadiganlar tomonidan to'liq tushunishga xizmat qiladi. Ilmiy vosita dasturiy ta'minot sifatida mavjud bo'lsa, manba kodiga kirish ushbu vositani to'liq tushunishda muhim element hisoblanadi. Bizning sohamiz

rivojlanib, dasturiy ta'minotning yangi versiyalari talab qilinar ekan, manbaga kirish bizning dasturimiz tez va samarali moslashtirishga imkon beradi.

Rasmol rivojlanishining asosiy liniyasi manbaiga har doim bepul kirish imkoni mavjud. Rasmol 2.7 seriyasining 1999 yildada boshlangan nashrlarini yaratish bilan Rasmol rasmiy ravishda ochiq manba dasturiga aylandi. "Ochiq manba" iborasining ma'nosi haqida ba'zi chalkashliklar mavjud. Dasturiy ta'minotni ishlab chiqishning dastlabki kunlarida ko'pchilik ilmiy dasturiy ta'minot manba kodi erkin almashildi.

Shu kunlarda, u diqqat bilan chizilgan huquqiy hujjatlar ilmiy dasturiy ta'minot va manba kodi erkin foydalanish uchun zarur deb chikildi. Biz mualliflik huquqi va cheklovchi litsenziyalar ijodiy birikmasi bizga dasturlarni ishlatish uchun, albatta, tekin erkinlik berishi mumkin. Torichard Stallman ularning manba kodini o'qish va yangi versiyalarini ishlab chiqishni ko'rsatdi.

**Oqsillar strukturasini ko'rish uchun Rasmol dasturiga qo'llanma.** Rasmol-PDB fayllarini ko'ruvchi dastur. PDB bunday fayllar Brookhaven protein ma'lumotlar bank internetdagi topish yoki boshqa dasturlar yordamida yaratilgan oqsil yoki DNK kabi katta molekulalar uchun ishlatiladigan fayl formati xisoblanadi. Rasmol dasturidan oqsillar molekulasi tiklash uchun foydalanish mumkin. Muayyan qoldiqlarni, shuningdek, turli tuzilmalarni aniqlash mumkin. Rasmolda faqat bir vaqtning o'zida bir nechta PDB fayllarni ochish mumkin. Rasmol tasvirlar internetdagi yoki hujjat ichida ko'rsatish uchun turli rasm formatlarida wordda saqlanishi mumkin. Rasmol "freeware" deb, faylni qaysi joyda o'rnatish mumkin, degan ma'noni anglatadi.

---

### **Rasmol ishchi stoli**

**Rasmol ish stoli.** Biriktirilgan ko'rsatmalarga qarang, har qanday kompyuterdan Rasmoldan qanday foydalanish yoki kompyuteringizda Rasmol dasturi bulsa, u xolda quyidagicha ketma-ketlikka qarang.

#### **Strukturani sozlash**

- Netscape ochib shundan sung  
<http://www.pdb.bnl.gov>ga o'tib
- 3D brauzer tanlang (sahifaga yarim yo'l)

## Welcome to the Protein Data

The Protein Data Bank (PDB) is an archive of experimental macromolecules, serving a global community of researchers. For more information on subscribing to the [PDB List server](#).



Send your suggestions and complaints to the

- ▶ Oct. 1997 PDB Newsletter
- ▶ PDB WWW Bulletin
- ▶ PDB Data Throughput Plot
- ▶ AutoDep at EBI
- NEW** ▶ Nature Structural Biology's Survey on the IUC

**NEW** En  
7081 Coordinate Entries / 6548

- [3DB Browser](#) and other Searching an
- [Submitting Data to the PDB](#)
- [PDB FTP Server](#)
- [PDB Documentation](#)
- [PDB Mirror Sites](#)
- [PDB Users Group](#)
- [mmCIF - Macromolecular Crystallog](#)
- [Molecular Biology Servers/Database](#)
- [Software and Related Information](#)

- Kalit so'z ostiga meoglobin suzini qidiruv shaklida yozib qismlari to'ldiring.
- Qidiruv tugmasini bosing.

## Protein Data Bank 3DB Browser

closerSite® From United States you have these 2 alternative mirrors [ [BNL](#) ] [ [UGA](#) ]

Simple searches	
<input type="button" value="Clear form"/>	
<b>PDB ID code</b>	<input type="text"/> <input type="button" value="Search"/> Four-character accession code
<b>Keyword</b>	<input type="text" value="myoglobin"/> Molecule name, class or family, or related term [HEADER, TITLE, KEYWDS and COMPND fields]
<b>Author</b>	<input type="text"/> Family name of depositor or author of associated publication [AUTHOR and JRNL fields]
<b>Text query</b>	<input type="text"/> Any word in the complete PDB text
<input type="button" value="Search"/>	<a href="#">FASTA search</a> <a href="#">HELP</a>

- Keyin Kepp Working bosing.

## Protein Data Bank 3DB Browser

Site updated on Wed Feb 4 0:22:26 EST 1998 [total entries: 7157]

Query results: 107 released, 16 being processed.		
Field	Hits	Your query
Keyword	107	myoglobin

[Display] the list of 107 Hits  
or click to [Keep Working] with all of them now.

- Oyna o'z-o'zini yangilash va joy mezonlar ostida sizning qidiruv natijalari bo'ladi.
- Oliy nur qarash istayman aralashma. (Bu ish uchun bor.)
- "Ma'lumotlarni olish" qutisini bosing.

Query results: 107 released, 16 being processed

Field	Hits	Yo
Keyword	107	myoglobin

Please select the entry or entries that you want to focus on, and click on the 'Retrieve data'

PDBid	Short description of the entry	Author	Res
0GHS	Myoglobin (Met)		1.50
1AJS	Photolysed Carbonmonoxy-Myoglobin At 20 K	Schlichting	1.50
1AJG	Carbonmonoxy Myoglobin At 40 K	Teng	1.70
1AJH	Photoproduct Of Carbonmonoxy Myoglobin At ...	Teng	1.70
1BJE	H64t Variant Of Myoglobin (Horse Heart) Re...	Hauras	1.80
1EHY	Myoglobin with Cyanide	Bisig	1.70
1FCS	Myoglobin mutant His 64 Val And Thr 67 Arg ...	Fizzi	1.60
1HJT	Spere Whale Myoglobin (Ferrous, Nitric Oxide)	Brucker	1.70
1KH1	Myoglobin mutant His 93 Tyr (H93Y)	Burk	1.70
1R5Y	Myoglobin mutant His 64 Thr (H64t)	Hauras	1.90

You may also select multiple lines on the list or select  ALL 107 to retrieve them as

[\[Download\]](#) this list of 107 ID codes for further refer

- Muallif, jurnal manbai va murojaatlarni ko'rsatadigan matn paydo bo'ladi. Aminokislotalar ketma-ketligi ham kursatib beradi.
- PDBni yuklab olish uchun Rasmol "assimetrik birlik" havolasini bosing.

This is 1HJT [\[HELP\]](#)

```

HEADER      OXYGEN TRANSPORT
TITLE       SPERE WHALE MYOGLOBIN (FERROUS, NI
COMPND      MOL_ID: 1;
COMPND      2 MOLECULE: MYOGLOBIN;
COMPND      3 CHAIN: NULL
AUTHOR      E. A. BRUCKER, G. N. PHILLIPS JUNIOR
    
```

---

**Data retrieval:**  
 Asymmetric unit, PDB entry: [\[header on Sequence: \[1hjt\\_1\]\]](#)  
[Ligand-Protein Contacts](#) for 1HJT  
[Complete Macromolecule\(s\) file](#) for 1HJT  
[Structure Factors](#) in CIF format (r1hjt.cif)  
[Retrieve 1HJT in mmCIF format](#)

**Molecule visualization:** [\[HELP\]](#) [How to view a VRML](#) | [Look here for Virtual Reality M](#)  
 Rasmol [Asymmetric unit](#)

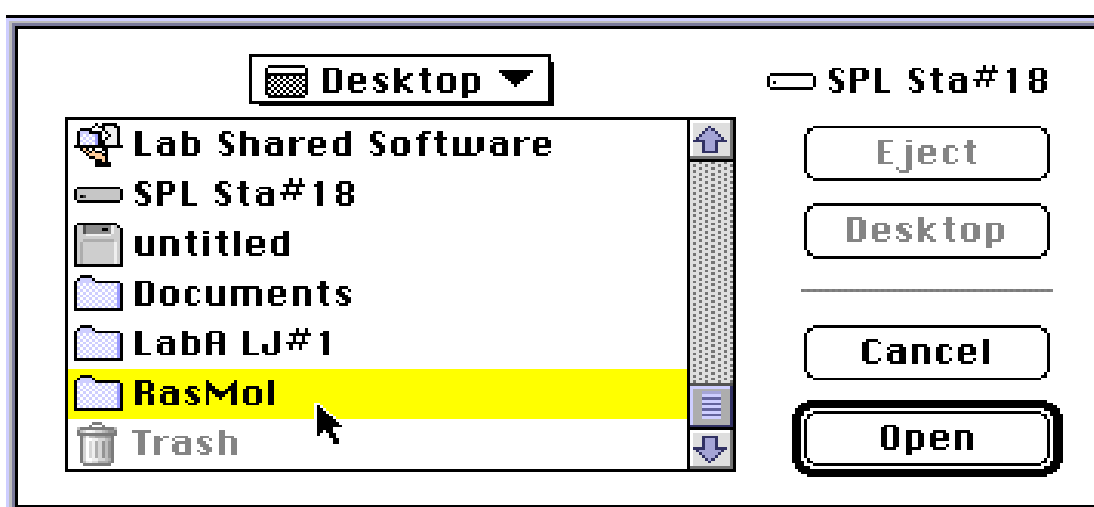
**Other resources with information on 1HJT**  
[scop](#) (Structural Classification of Protein)  
[PDBREPORT](#) (protein verification by W)  
[MMDB](#) (Entrez's Structure Database)  
 Fold representative is [1mbd](#) from [Dali/Fs](#)  
[Domain Definition for 1hjt](#) from [3Dec](#) ( T)  
[PDBSUM to CATH](#) (PDB Summary files: Classification)

- [Pick appni bosing.](#)



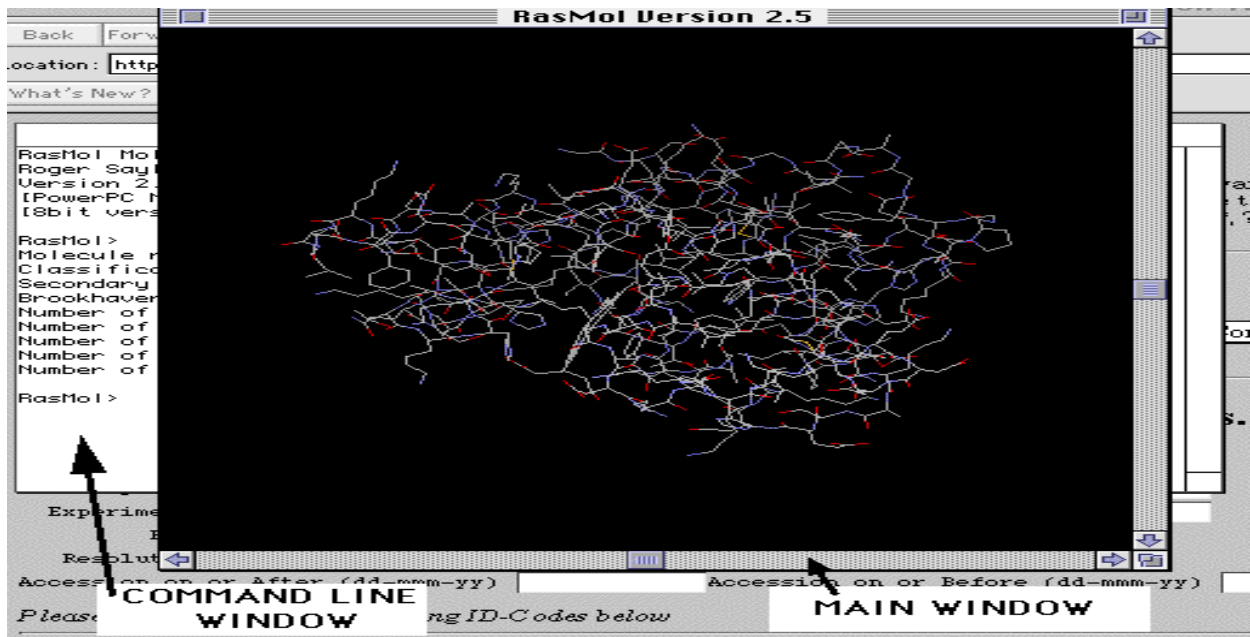
Stol ustidagi Rasmol papkasini oching

- Rasmolni tanlan va bosing
- Ochish uchun birinchi marta ariza tanlash kerak bo'ladi.



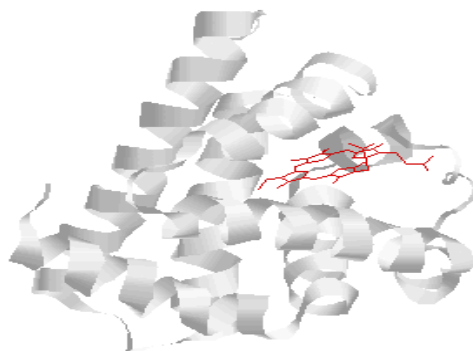
**Rasmol ish stoli.** Rasmol oqsilni turli xil konfiguratsiyalar va ranglarda ko'rsatishga qodir. Rasmolning ikkita oynasi bor: biri tasvirni aks ettiruvchi asosiy oyna va ikkinchisi buyruqlar satri oynasi. Buyruqlar satri oynasi foydalanuvchiga buyruqlar kiritish imkonini beradi. Buyruqlar ko'p, shuningdek, ekranning yuqori qismida menyu imkoniyatlari yordamida bir foydalanuvchi bilan do'st tarzda kirish mumkin. Biroq, faqat buyruq yozib amalga oshirilishi mumkin buyruqlar bor. Bu keyingi bo'lim sizga buyruqlar har bir turi bir necha ko'rsatadi. Rasmol siz tanlagan proteinni ochib beradi wireframe ikki xil derazani farqlang old oyna asosiy oyna molekulasi ko'rsatiladi asosiy oynaning orqasida boshqa oyna buyruq satr oynasi





*Rasmol ishchi oynasi*

- Display menyusida Ribbon ni tanlang. Bu tasmada faqat oqsil asosini ko'rsatadi.
  - Buyruqlar satri oynasini bosing.
  - Rasmol xohish turda: heme ni tanlang.
  - Keyingi xohishi turida: qizil rang.
  - Keyingi tezkor turda: wireframeni
  - Bundan tashqari, **View** menyusiga o'ting va hemeni ko'rishni turli yo'llarini tanlash mumkin. Iplarga o'xshab faqat o'zagini ko'rsatadiganlari xammasini ko'rsatmaydi.



### **Ko'rinishi va rangni tanlash**

- **View** menyusidan tanlagan buyruqlar faqat bitta tanlangan elementni ishlaydi,

agar siz butun oqsilni o'zgartirmoqchi bo'lsangiz, tahrirlash menyusi ostida va kerakli ko'rish rejimini tanlang. **View** menyusida ostida joy va tasodifiy variantlari tanlab turli ko'rish va rang uslublari bilan atrofida o'ynash mumkin.

- Rangli Menyu ostida bir xil narsani bajaring. Fon rangi Ba'zan boshqa fon rangini chop etishda kerak.

### **Buyruqlar satri oynasida**

- Turli fon oq yoki hech qanday rangsiz rasmlarni saqlash uchun bo'ladi.

### **Aminokislotalar en zanjirlarni tanlash**

- Buyruqlar satri oynasida (tasma ko'rinishida va monoxrom rangda boshlang.)
- Turi "uning tanlang" xohishida
- "color yellow" turi sizga asosiy oynadagi bittasiga bosadigan proteindagi barcha gistidinlarni ko'rsatadi, gistidin soni buyruqlar satri oynasida chop etiladi.

**Eslatma:** ba'zan gistidinni bosish qiyin va tasodifan boshqa qoldiqda emas.

## **Nazorat savollari**

1. RasMol nima?
2. Rasmoldan tasvirni chop etish qanday amalga oshadi?
3. Rasmolda Ko'rinishi va rangni tanlash qanday amalga oshadi?
4. Rasmolda strukturani sozlash qanday amalga oshadi?

## **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008

## **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
1. А. Н. Огурцов Введение в Бионформатику, 2011г.

## **Internet ma'lumotlari**

1. <https://www.ncbi.nlm.nih.gov/orffinder/>
2. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
3. <https://www.cgl.ucsf.edu/chimera/>

4. <https://proteopedia.org/wiki/index.php/Chimera>

## 11-Amaliy mashg'ulot.

**Mavzu:** Ligand, retseptor ta'sirlashuvlarini oldindan aytish. AutoDock dasturida ishlash.

**Dars maqsadi:** Talabalarga Ligand, retseptor ta'sirlashuvlarini oldindan aytish bo'yicha ma'lumot berish va AutoDock dasturida ishlash bo'yicha amaliy ko'nikma uyg'otish.

**Kerakli jihoz va preparatlar:** Kamyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** Protein-ligandlarning o'zaro ta'sirini o'rganish va dori-darmonlarni kashf qilish va ishlab chiqish uchun hisoblash uchun docking keng qo'llaniladi. Odatda jarayon ma'lum strukturaning maqsadi bilan boshlanadi, masalan, dorivor ahamiyatga ega bo'lgan fermentning kristallografik tuzilishi. Docking keyinchalik bog'langan konformatsiyani va kichik molekulalarning nishonga bog'langan erkin energiyasini bashorat qilish uchun ishlatiladi. Yagona o'rnatish tajribalari maqsad funksiyasini o'rganish uchun foydalidir va virtual skrining, bu erda katta birikmalar kutubxonasi joylashtirilgan va tartiblangan, dori ishlab chiqarish uchun yangi inhibitorlarni aniqlash uchun ishlatilishi mumkin.

AutoDock - bu kichik molekulalarni makromolekulyar retseptorlarga hisoblash va virtual skrining qilish uchun ochiq kodli bepul dasturiy ta'minot to'plami. Hozirgi vaqtda to'plam bir nechta qo'shimcha vositalarni o'z ichiga oladi:

- AutoDock Vina : oddiy ball funksiyasi va tezkor gradientni optimallashtirish konformatsion qidiruviga asoslangan kalit topshirqli hisoblash docking dasturi. [1](#)
- AutoDock : empirik erkin energiya kuchi maydoniga va tezkor Lamark genetik algoritmini qidirish usuli [2](#) , [3](#) -ga asoslangan hisoblash docking dasturi .
- Raccoon2 : virtual skrining va tahlil qilish uchun interaktiv grafik vosita [4](#) .
- AutoDockTools : koordinatalarni tayyorlash, joylashtirish va tahlil qilish uchun interaktiv grafik vosita. [5](#)
- AutoLigand : retseptorlari bilan bog'lanishning optimal joylarini bashorat qilish uchun dastur.

AutoDock to'plami, shu jumladan manba, bepul mavjud va tadqiqot va dori-darmonlarni topishda keng qo'llanilgan.

Hozirgi vaqtda ligandlarni hisoblash uchun turli xil akademik va tijorat usullari mavjud ( joriy usullarni keng ko'lamli ko'rib chiqish uchun Ref [1 ga qarang](#)). Ushbu usullarning aksariyati hisob-kitoblarni qulay qilish uchun muammoni ikki usulda soddalashtiradi. Birinchidan, konformatsion bo'shliq tizimga qattiq retseptor va liganddagi qattiq bog'lanish burchaklari va uzunliklari kabi cheklovlar qo'yish orqali kamayadi. Ikkinchidan, konformatsiyani qidirishning har bir bosqichida pozalarni tez baholash uchun ko'pincha bog'lanishning empirik erkin energiyalariga asoslangan soddalashtirilgan ball funksiyasi qo'llaniladi.

Bularning ikkalasi ham jiddiy cheklovlardir va agar yanada aniqroq konformatsion qidiruv yoki energiyani bashorat qilish zarur bo'lsa, foydalanuvchilar molekulyar dinamika yoki erkin energiya buzilishi kabi vositalardan foydalanishlari kerak. Ushbu vositalar hisoblash o'rnatish usullarini to'ldiradi, chunki o'rnatish usullari odatda kengroq konformatsion makonni qidiradi, ammo yanada ilg'or usullar konformatsion landshaftning mahalliy hududida konformatsiya va energiyani aniqroq bashorat qilishi mumkin.

Energiyani baholash uchun tezkor usulni talab qilish cheklovlari juda cheklangan bo'lgan hollarda natijalarni yaxshilash uchun ilg'or o'rnatish usullaridan foydalanish mumkin. Misol uchun, ko'plab joylashtirish usullari retseptor uchun qattiq modelni qo'llaydi, bu ko'pincha bog'langanda sezilarli induktsiyalangan oqsillar uchun noto'g'ri natijalarga olib keladi. AutoDock retseptorlaridagi cheklangan konformatsion o'zgarishlarni hisobga olish uchun retseptorlarning yon zanjirlarini aniq davolash usulini o'z ichiga oladi. Bundan tashqari, buyurtma qilingan suv molekullari ko'pincha ligandlar va retseptorlar o'rtasidagi o'zaro ta'sirga vositachilik qiladi va tanlangan suvlarni aniq tozalashning ilg'or usullari AutoDock-da amalga oshirildi. Ushbu ilg'or usullarning ikkalasi ham ushbu protokolda ko'rsatilgan.

Ko'pgina hisobotlarda AutoDock (yaqinda Sousa va boshq. [7](#) tomonidan ko'rib chiqilgan ) kabi mashhur docking usullarining ishlashi taqqoslangan. Turli usullar aniq maqsadlarga qarab turli muvaffaqiyat stavkalariga erishishi mumkin, lekin umuman olganda, ularning barchasi bir qator turli xil protein-ligand komplekslarida sinovdan o'tganda bir xil ishlaydi: ularning barchasi dori o'lchamidagi molekullar uchun bog'langan komplekslarni bashorat qilishda yaxshi ishlaydi, taxminlar bilan. Agar retseptorda sezilarli harakat talab qilinmasa, taxminan 2-3 kkal/mol xatolik bilan bog'lanishning erkin energiyalari. Muayyan tizim uchun o'rnatish usulini sozlash yoki tizimning yanada murakkab va hisoblash intensiv parametrlariga o'tish orqali yaxshi natijalarga erishish mumkin.

**AutoDock va AutoDock vina dasturlari.** Ikki xil ehtiyojga javob berish uchun ikkita o'rnatish usuli parallel ravishda ishlab chiqilgan. Ishlab chiqish

AutoDock [2](#), [3](#), [5](#), [21](#), [22](#) dan boshlandi va u o'rnatish usullari bo'yicha tajribalar uchun platforma bo'lib qolmoqda. AutoDock Vina foydalanuvchilardan keng ekspert bilimlarini talab qilmaydigan kalit taslim o'rnatish usuliga bo'lgan ehtiyojni qondirish uchun yaqinda ishlab chiqilgan [1](#). Bu yaxshi sinovdan o'tgan standart usullardan foydalangan holda o'rnatish tajribalarini bajarish uchun yuqori darajada optimallashtirilgan. Hozirda ikkala usul ham bepul. AutoDock Vina ko'pgina tizimlar uchun tez va samarali, AutoDock esa qo'shimcha uslubiy yaxshilanishlarni talab qiladigan tizimlar uchun mavjud.

Ikkala usul ham umumiy hisoblash moslamalari, retseptorlar va ligandlar uchun koordinata fayllarini qabul qilish va optimal o'rnatilgan konformatsiyalarni bashorat qilish uchun mo'ljallangan. Odatda, foydalanuvchilar kristallografiya yoki NMR spektroskopiyasidan olingan retseptorlarning koordinatalaridan va SMILES satrlari yoki boshqa usullardan olingan ligand koordinatalaridan boshlanadi.

Qidiruv usullari stokastik bo'lganligi sababli, optimal o'rnatilgan konformatsiyalar to'plami bashorat qilinadi, so'ngra natijalarning izchilligini tahlil qilish uchun odatda fazoviy ravishda klasterlanadi. Yuqori klasterlangan natijalar konformatsion qidiruv protsedurasi qulay konformatsion makonni qamrab olishni ta'minlash uchun etarlicha to'liq ekanligidan dalolat beradi. Qidiruvning stokastik tabiati tufayli usul global minimal topilganligini ta'minlay olmaydi. Shu sababli, foydalanilayotgan docking protokolini baholash uchun o'xshash konformatsion murakkablikdagi ma'lum komplekslar bilan qayta o'rnatish tajribalarini qo'llash muhimdir.

AutoDock va AutoDockVina hozirda olingan natijalarga ta'sir qiluvchi bir nechta soddalashtirishlardan foydalanadi. Eng muhim soddalashtirish - bu qattiq retseptordan foydalanish. Ushbu yaqinlashish konformatsion fazoning hajmini qisqartiradi, uni ishonchli izlashga imkon beradi va har bir sinov konformatsiyasini baholash uchun hisoblash harakatini kamaytiradi. Ushbu o'rnatish usullarini ma'lum bir retseptorga qo'llashda ushbu cheklovning mumkin bo'lgan ta'sirini hisobga olish kerak va agar tizim retseptorlarning sezilarli harakatini o'z ichiga olsa, bir qator usullardan foydalanish mumkin, jumladan:

- Retseptor-ligand komplekslaridan olingan retseptor tuzilmalaridan foydalanish, bu erda retseptor tegishli konformatsiyada bo'lishini kutish.
- Retseptorda kutilgan moslashuvchanlik diapazonini qamrab oluvchi turli xil retseptor tuzilmalari to'plamiga ulanish. Ularni bir nechta strukturaviy aniqlashlar yoki simulyatsiyalar yordamida olish mumkin.
- Tegishli yon zanjirlar haqida ma'lumot mavjud bo'lsa (protokolda tavsiflangan) o'rnatish vaqtida aniq retseptorlarning yon zanjiri moslashuvchanligidan foydalanish.

Baholash usullari, shuningdek, natijalarga ta'sir qiladigan turli xil soddalashtirishlarni qo'llaydi. AutoDock Vina skoring funksiyasi juda taxminiy bo'lib, sferik simmetrik vodorod bog'lanish potentsiallari, yashirin vodorodlar va elektrostatik hissasi yo'q. Odatda biologik o'lcham va tarkibga ega bo'lgan ligandlar bilan yaxshi ishlashi isbotlangan. AutoDock kuch maydoni fizik jihatdan asoslangan hissalarini, jumladan, aniq qutbli vodorodlar bilan yo'naltirilgan vodorod bog'lanish atamasini va elektrostatikani o'z ichiga oladi. Agar ushbu hissalar ma'lum bir tizimda muhim bo'lsa, AutoDock mos vosita bo'ladi. Bundan tashqari, agar kerak bo'lsa, ma'lum tizimlarni sozlash imkonini beruvchi AutoDock skoring funksiyasining parametrlari foydalanuvchi uchun mavjud.

**AutoDock instrumentlari bilan tayyorgarlikni muvofiqlashtirish.** Muvaffaqiyatli joylashtirish va virtual skrining qilish retseptorlar va ligandlar uchun ishlatiladigan koordinatalar sifatiga diqqat bilan e'tibor berishni talab qiladi. AutoDock ham, AutoDock Vina ham molekularning soddalashtirilgan tasviridan foydalanadi, ular PDBQT deb ataladigan o'zgartirilgan Protein ma'lumotlar banki (PDB) fayl formatiga kiritilgan:

- Birlashgan atom tasviri: Ikkala usul ham qutbli vodorod atomlarini o'z ichiga olgan koordinatalar to'plamini talab qiladi. AutoDock o'rnatish vaqtida qutbli vodorod koordinatalaridan foydalanadi; AutoDock Vina ulardan heteroatomlarning vodorod bog'lanish holatini belgilash uchun foydalanadi, lekin o'rnatish vaqtida aniq vodorodlardan foydalanmaydi.
- Atom tiplash: Ikkala usul ham atomlarni soddalashtirilgan tiplashni talab qiladi, jumladan aromatik va alifatik uglerod atomlarini aniqlash va geteroatomlarning vodorod bog'lanish holatini aniqlash.
- Atom zaryadlari: AutoDock elektrostatik o'zaro ta'sirlar va desolvatsiya energiyalarini hisoblash uchun Gasteiger-Marsili atom zaryadlaridan foydalanadi. AutoDock Vina atom zaryadlaridan foydalanmaydi.
- Ligandning moslashuvchanligi: Ikkala usul ham foydalanuvchidan moslashuvchan bo'lishi kerak bo'lgan ligand molekularida va har qanday retseptor yon zanjirlarida burilish erkinlik darajalarini belgilashni talab qiladi.
- Qidiruv maydoni: Ikkala usul ham foydalanuvchidan qidiriladigan retseptor atrofidagi bo'sh joyni qoplaydigan o'rnatish qutisini ko'rsatishni talab qiladi.

AutoDockTools grafik foydalanuvchi interfeysi turli xil umumiy formatlardagi fayllardan mos koordinatali fayllarni yaratish usullarini taqdim etadi. Avtomatlashtirilgan skriptlar birikmalar kutubxonalarini ommaviy qayta ishlash uchun ham mavjud.

AutoDock va AutoDock Vina-da qo'llaniladigan standart usullar odatdagi dori-darmonlarga o'xshash ligandlar uchun juda samarali bo'lib, virtual skrining 25 kabi ilovalar uchun keng qo'llaniladi. Biroq, standart usullar bilan qiyinchiliklar tug'diradigan muammolarni hal qilish uchun bir nechta takomillashtirish ishlab chiqilgan.

Ushbu dizayn parametrlaridan chetga chiqqan tizimlar o'zgaruvchan natijalar beradi va ehtiyotkorlik bilan yondashish kerak. Foydalanuvchilar ko'pincha ikkita muhim cheklovlarga duch kelishadi. Birinchidan, foydalanuvchilar ko'pincha dekapeptid kabi juda katta ligandlarni joylashtirishni xohlashadi. Ushbu ligandlar juda ko'p erkinlik darajasiga ega va joylashtirish usullari mavjud konformatsion makonni qidirishga qodir emas. Ko'pincha, bu muammoni kichik qismlarga ajratish orqali hal qilinadi. Ikkinchidan, oqsil maqsadlari ko'pincha AutoDock to'plamida modellashtirilmagan muhim konformatsion moslashuvchanlikni ko'rsatadi, bu Protokolni o'z ichiga olgan tanlangan yon zanjir harakatidan tashqari. Bu muammoni hal qilish, odatda, o'rnatishdan oldin boshqa usullar (masalan, molekulyar dinamika) bilan oqsilning konformatsiyasini yaratish orqali amalga oshiriladi. Ushbu protokol odatiy dori-darmonlarni kashf qilish loyihasi bilan bog'liq ko'plab bosqichlarni tavsiflaydi. Protokol uchun maqsad proto-onkogen tirozin protein kinaz c-Ablning kinaz domenidir. Protein saraton kimyoterapiyasi, xususan, surunkali miyelogen leykemiya (CML) ni davolash uchun muhim maqsaddir. Ushbu misol tasdiqlangan Gleevec (Imatinib) preparati bilan muhim saraton nishonida protokoldan foydalanishni ko'rsatishi va maqsadning dori-darmonlarga chidamli mutatsiyalarining ta'sirini ko'rsatishi mumkinligi sababli tanlangan. Protokol quyidagilarni o'z ichiga oladi:

- AutoDock Vina va AutoDock-dan foydalangan holda, saratonga qarshi imatinib (PDB kirish [1iep](#) 27 ) bilan c-Abl strukturasiidan foydalangan holda tajribalarni qayta tiklash.
- C-Abl ga qarshi birikmalar kutubxonasi Raccoon2 bilan virtual skrining, PDB kirish [1iep](#) dan oqsil koordinatalari yordamida.
- PDB kirish [1fpu](#) [28](#) dan c-Abl koordinatalari bilan imatinibni o'zaro bog'lash, preparat bilan o'zaro ta'sir qiluvchi treonindagi moslashuvchanlikni modellashtirish.
- AutoLigand yordamida c-Abl uchun optimal ligandlarni bashorat qilish.
- Fragmentga asoslangan dori dizayni natijalarini yaxshilashi mumkin bo'lgan aniq suv molekullari bilan ulash misoli.

## **Boshlanish ma'lumotlari**



- Retseptor uchun koordinata fayli (pdb, mol2, cif va sdf kabi turli formatlarda)
- Coordinate file for ligand (in a variety of formats, including pdb, mol2, cif & sdf)
- Har bir protokol uchun qo'llanma sifatida foydalanish uchun Qo'shimcha ma'lumotlarda bir nechta fayllar mavjud: 1iep\_receptorH.pdb (PDB 1iep yozuvidagi c-Abl kinaz domenining koordinatalari, AutoDockTools'ga vodorod atomlari qo'shilgan), 1iep\_ligandH.pdb (imatinib koordinatalari). PDB yozuvi 1iep, vodorod atomlari AutoDockTools ilovasiga qo'shilgan), 1fpu\_receptorH.pdbqt (PDB kirish 1fpu dan c-Abl kinaz domenining koordinatalari, PDBQT formatida), imatinib.pdbqt (PDBQT formatidagi imatinib koordinatalari (PDBQTIIa formatlari\_ub), NCIdiset PDBQT fayllari sifatida formatlangan ZINK kutubxonasi dan 499 ta birikmani o'z ichiga oladi)

### Uskuna va dasturiy ta'minot

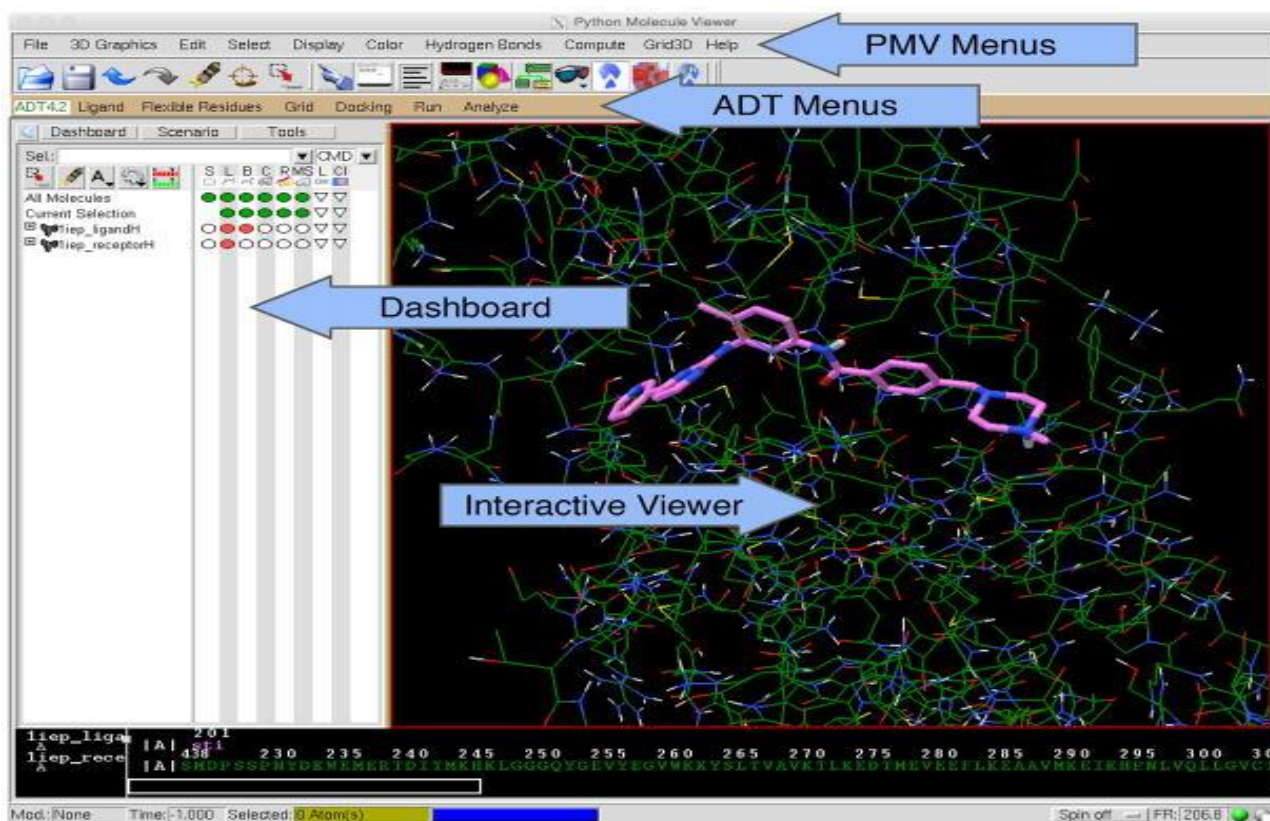
- Kompyuter: Linux, Macintosh yoki Windows PC; Internetga ulanish
- Raccoon bilan virtual skrining uchun, PBS yoki SGE rejalashtiruvchisi bo'lgan Linux klasteri/HPC
- Matn muharriri
- AutoDock: <http://autodock.scripps.edu> (o'rnatish haqida ma'lumotni quyidagi manzilda olishingiz mumkin: <http://autodock.scripps.edu/downloads/autodock-registration/autodock-4-2-download-page/> )
- AutoDock Vina: <http://vina.scripps.edu> (o'rnatish haqida ma'lumotni <http://vina.scripps.edu/manual.html#faq> saytida olishingiz mumkin )
- AutoDockTools (MGLTools qismi): <http://mgltools.scripps.edu> (information on installation is available at: <http://mgltools.scripps.edu/downloads>)
- AutoLigand AutoDockTools-ning bir qismidir
- [Rakun http://autodock.scripps.edu/resources/raccoon](http://autodock.scripps.edu/resources/raccoon) saytida mavjud.

**AutoDock instrumentlari yordamida tayyorgarlikni muvofiqlashtirish - vaqt 10 min.** Ligand koordinata faylini yarating. Vodorod atomlarini o'z ichiga olgan koordinatalar to'plami talab qilinadi. Buni turli yo'llar bilan olish mumkin, jumladan Protein ma'lumotlar banki ( [www.pdb.org](http://www.pdb.org) ) yoki Kembrij kristallografik ma'lumotlar bazasi (ccdc.cam.ac.uk) dan eksperimental koordinatalar yoki CACTUS serveri (kaktus) kabi strukturani yaratish usullari ([nci.nih.gov/translate](http://nci.nih.gov/translate)). 1iep\_ligandH.pdb fayli ushbu protokol bo'yicha o'quv qo'llanmasi sifatida foydalanish uchun taqdim

etilgan (barcha misol fayllar [Qo'shimcha ma'lumotlarda](#) keltirilgan ), u PDB liep yozuvidan olingan, barcha vodorod atomlari qo'shilgan va ma'lum protonatsiyaga qo'lda sozlangan ligand koordinatalarini o'z ichiga oladi. davlat. AutoDockTools (ADT) dasturini ishga tushiring ([1-rasm](#)) va "Fayl->Preferences->Set"-ni bosish orqali ishchi katalogni o'rnatish. "Ishga tushirish katalogi" maydoniga ishchi katalogingiz yo'li nomini kiriting va "O'rnatish" tugmasini bosish. Oynaning pastki qismidagi "O'chirish" tugmasini bosish.

Atom koordinatalarini o'qing. Buni amalga oshirish uchun avval "Ligand->Input->Open" -ni tanlang va "PDB fayllari" ni tanlash uchun "Fayllar turi" menyusidan foydalaning. Koordinata faylingizni, bu holda liep\_ligandH.pdb ustiga bosish va "Ochish" tugmasini bosish. ADT koordinatalarni o'qiydi, agar kerak bo'lsa zaryad qo'shadi, qutbsiz vodorodlarni birlashtiradi va tegishli atom turlarini tayinlaydi. Shu nuqtada, ligand yashil rangda aromatik uglerodlar bilan tomoshabin oynasida ko'rsatiladi. Davom etish uchun qalqib chiquvchi oynada "OK" tugmasini bosish.

Agar koordinatalar to'plamida vodorod pozitsiyalari bo'lmasa, "Fayl->Molekulani o'qish" tugmasini bosish va koordinata faylingizni tanlang. Keyin, "Tahrirlash->Vodorodlar->Qo'shish" tugmasini bosish orqali siz barcha vodorodlarni sukut bo'yicha qo'shasiz. Ligand molekulasini tanlash uchun "Ligand->Kirish->Tanlash"-ni tanlang.



AutoDock instrumentlari

ADT Python Molecule Viewer ichida qurilgan. Koordinatalarni tayyorlash, joylashtirish va tahlil qilish uchun ADT buyruqlari pastki asboblar panelidagi menyular orqali mavjud. Yuqori darajadagi vizualizatsiya uchun PMV buyruqlari yuqori asboblar panelida mavjud. Boshqaruv paneli interaktiv ko‘rish panelida ko‘rsatiladigan molekulyar obyektlarning ko‘rinishini, ranglarini va yorliqlarini boshqaradi. PMV va ADT imkoniyatlari haqida qo‘shimcha ma‘lumot olish uchun <http://mglttools.scripps.edu/documentation> ga qarang .

**Docking simulyatsiyasi usullari.** Retseptor va ligand koordinatalari formatlangandan so‘ng, AutoDock to‘plami o‘rnatish simulyatsiyasi uchun bir qator usullarni taqdim etadi. Ushbu protokol quyidagi jadvalda tavsiflanganidek, oddiy o‘rnatishdan tortib, ilg‘or usullargacha bo‘lgan oltita usulni o‘z ichiga oladi.

Variant	Usul	Tavsif
A	AutoDock Vina bilan bitta o‘rnatish tajribasi	Bitta retseptorli bitta ligandni o‘rganish uchun asosiy o‘rnatish usuli
B	AutoDock bilan bitta o‘rnatish tajribasi	Yagona retseptorli bitta ligandni o‘rganish uchun asosiy ulanish usuli, yaqinlik xaritalarini aniq hisoblash.
C	Raccoon2 va AutoDock Vina bilan virtual skrining	Bitta retseptorli ligandlar kutubxonasining virtual ekrani, ko‘pincha dori-darmonlarni topish uchun ishlatiladi
D	Moslashuvchan yon zanjirlar bilan AutoDock Vina	Cheklangan retseptorlarning moslashuvchanligini o‘z ichiga olgan bitta retseptorli bitta ligand uchun joylashtirish usuli
E	AutoLigand yordamida saytni faol bashorat qilish	Retseptorlarning bog‘lanish joylarini tahlil qilish, dorivor joylarni bashorat qilish usuli
F	Aniq suvlar bilan tutash	Aniq ko‘priqli suv molekularini o‘z ichiga olgan bitta retseptorli bitta ligand uchun ilg‘or biriktirish usuli

*(A) AutoDock Vina bilan bitta o‘rnatish tajribasi - vaqt 30 min*

AutoDock Vina uchun ligand va retseptor uchun PDBQT fayllarini belgilaydigan va o‘rnatish parametrlarini belgilaydigan konfiguratsiya faylini ([1 - bo‘lim](#)) yarating. Buning uchun ADT-ni qayta ishga tushiring va standart ishchi katalogni o‘rnating (1-bosqichga qarang). “Grid->Makromolekula->Choose/Open” ni tanlang. “Tanlash” koordinatalar ADTda allaqachon o‘qilgan bo‘lsa (yuqoridagi koordinata sozlamalarida bo‘lgani kabi) va “Ochish” fayldan koordinatalarni o‘qish uchun ishlatiladi. Ushbu protokol uchun biz ADT ni qayta ishga tushirdik va

koordinata fayllaridan o'qishimiz kerak deb hisoblaymiz. PDBQT retseptor faylini “oching”, fayldagi mavjud to'lovlarni saqlab qolish uchun “Ha” tugmasini bosing va qabul qilish uchun “OK” tugmasini bosing. To'lovlarda kichik qoidabuzarliklar bo'lsa, ogohlantirish oynasi ham bo'lishi mumkin. Agar paydo bo'lsa, "OK" tugmasini bosing.

## 1 blok

### Vina konfiguratsiya fayli

retseptor = 1iep\_receptorH.pdbqt

ligand = 1iep\_ligandH.pdbqt

center\_x = 15.19

center\_y = 53.903

center\_z = 16.917

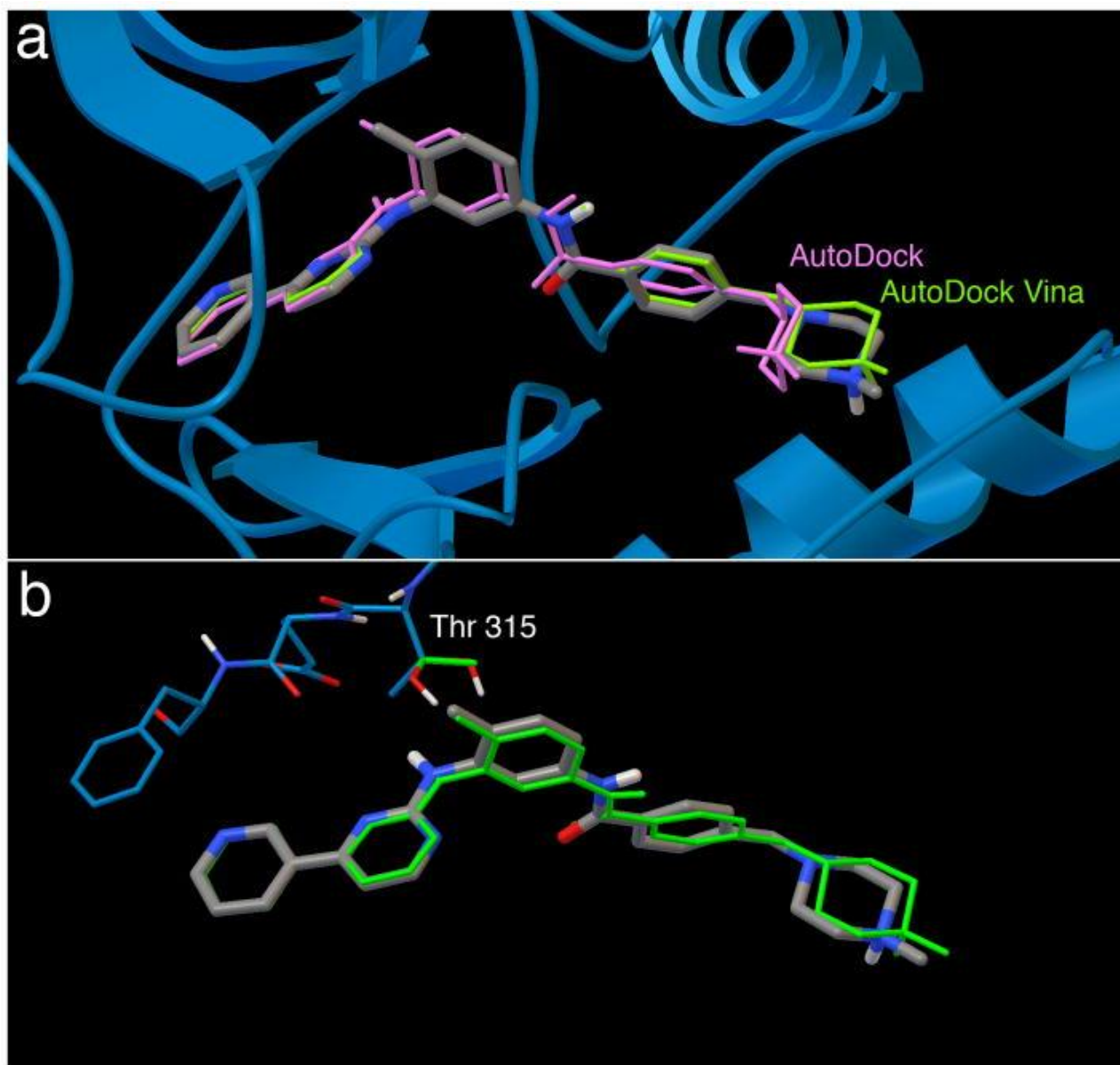
size\_x = 15.0

size\_y = 17.25

size\_z =

15\_bt.dp out.

“Grid->SetMapTypes->OpenLigand” ni tanlang, PDBQT ligand faylini tanlang va “Ochish” tugmasini bosing. “Grid->GridBox” qidiruv maydonining markazi va hajmini aniqlash uchun oynani ochadi. "Center->CenteronLigand" ligand asosida qutini belgilaydi. Boshqa variantlar mavjud yoki qiymatlar barmoq g'ildiraklari yordamida qo'lda o'zgartirilishi mumkin. Muhim: tugallangach, “Docking->Output->VinaConfig” oynasida “File->CloseSavingCurrent” bandini tanlang, “Saqlash” tugmasini bosish orqali konfiguratsiya faylini (standart nomi: config.txt) yozing.



Bog'langan va apo konformasyonlarda imatinibni retseptoriga ulash natijalari

a . Moslashuvchan imatinibni AutoDock (binafsha rang) va AutoDock Vina (yashil) yordamida qattiq Abl (PDB yozuvi 1iep) ga qayta joylashtirish. Rentgen kristallografik ligand holati kumush rangda.

b . Moslashuvchan imatinibni AutoDock Vina (yashil) yordamida bitta moslashuvchan qoldiq yon zanjiri bilan Abl (PDB kirish 1fpu) ga o‘zaro o‘rnatish. E’tibor bering, Vina o‘rnatish vaqtida vodorod atomi pozitsiyalarini saqlamaydi, shuning uchun treonin gidroksil vodorodi o‘rnatilgan koordinatalar to‘plamida tasodifiy holatda joylashtiriladi.

Tahlil qilish-> Dockings->ShowInteractions” ligand va retseptor o‘rtasidagi o‘zaro ta’sirni ko‘rsatadigan vizualizatsiyani ishga tushiradi.

AutoDock uchun retseptor uchun PDBQT fayllarini va atomik yaqinlik xaritalarini yaratish parametrlarini belgilaydigan grid parametr faylini yarating. ADT-



ni ishga tushiring va standart ishchi katalogni o'rnating (1-bosqichga qarang). "Ligand->Input->Open" ni tanlang va ligand PDBQT faylini tanlang (1iep\_ligandH.pdbqt). "Grid->Macromolecule->Open" ni tanlang va PDBQT retseptor faylini (1iep\_receptorH.pdbqt) tanlang. "Grid->SetMapTypes->ChooseLigand" va ligand PDBQT faylini (1iep\_ligandH.pdbqt) tanlash orqali xarita turlarini o'rnating. "Grid->GridBox" ni tanlab, qidiruv maydonini aniqlang, bu qidiruv maydonining markazi va hajmini belgilash uchun oynani ochadi. Minimal qutini aniqlash uchun "Center->CenteronLigand" dan foydalaning. Tugatgandan so'ng, "Fayl->CloseSavingCurrent" ni tanlang. Nihoyat, grid parametr faylini saqlash uchun "Output-> SaveGPF" ni tanlang,

AutoGrid-ni ADT-dan yoki buyruq satridan ishga tushiring. ADT dan "Run->RunAutoGrid" ni tanlang va ochilgan oynada "Browse" tugmalaridan foydalanib, AutoGrid bajariladigan faylga yo'lni va grid parametr fayliga yo'lni belgilang. Grid parametr faylini ko'rsatganingizdan so'ng, ADT jurnal fayli uchun nom taklif qiladi. AutoGrid-ni ishga tushirish uchun "Ishga tushirish" tugmasini bosing.

Shu bilan bir qatorda, buyruq satrida terminal oynasini oching va koordinata fayllari va grid parametr faylini o'z ichiga olgan katalogga o'ting. Keyin buyruqni bering:

```
./autogrid4 -p 1iep.gpf -l 1iep.glg
```

Bu buyruq AutoGrid bajariladigan *autogrid4* ham xuddi shu katalogda joylashganligini nazarda tutadi.

**Muammolarni bartaraf qilish va nosozliklarni tuzatish.** Majburiy bo'lmagan ADTda AutoGrid xaritalarini ko'ring (quyida E bo'limiga qarang). ADT-ni ishga tushiring va standart ishchi katalogni o'rnating (1-bosqichga qarang). Yuqori asboblar panelida "Grid3D->O'qish" ni tanlang va xarita faylini tanlang (masalan, uglerodga yaqinlik xaritasi uchun 1iep\_receptorH.C.map). Qalqib chiquvchi boshqaruv panelida "Tormoq nomi" panelidagi faylni bosing, "Izokontur" tugmasini bosing va ADT xarita qiymatlarining gistogrammasini ko'rsatadi. "Maksimum" qiymatini 0,0 ga o'zgartiring va qaytish tugmasini bosing (yoki "enter"). Izokonturni hisoblash uchun gistogrammada Shift tugmasini bosing; keyin kontur darajasini o'zgartirish uchun gistogrammadagi satrni o'ngga va chapga sudrab olishingiz mumkin. -0,5 dan -0,1 gacha bo'lgan qiymatlar samarali hisoblanadi. Retseptorni ko'rsatish uchun "File->ReadMolecule" menyu opsiyasidan foydalanish mumkin.

Ligand uchun PDBQT faylini va o'rnatish simulyatsiyasi parametrlarini belgilaydigan docking parametr faylini yarating. "Docking->Macromolecule->SetRigidFilename" ni tanlang va PDBQT retseptor faylini (1iep\_receptorH.pdbqt) tanlang. "Docking->Ligand->Choose" ni tanlang, ligand PDBQT faylini (1iep\_ligandH.pdbqt) tanlang va standart ligand o'rnatish parametrlarini qabul qiling,

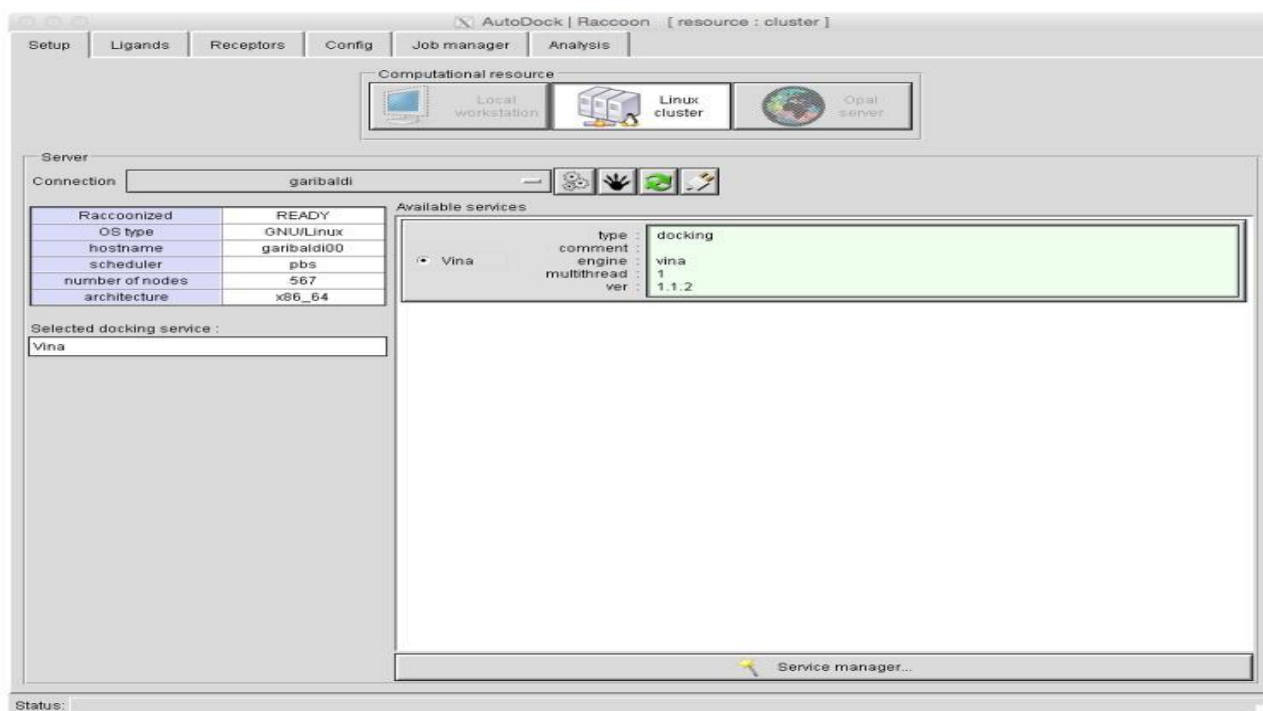
ular o'rnatishdan oldin ligand pozasini tasodifiylashtiradi. "Docking->SearchParameters->GeneticAlgorithmParameters" bandini tanlang va ko'pgina dori o'lchamli ligandlar uchun standart parametrlardan foydalaning, bundan "GA yugurishlar soni"ni 50 ga o'rning. Qo'shimcha o'rnatish parametrlari "Docking->DockingParameters;" bilan belgilanadi; ko'pchilik dori o'lchamli ligandlar uchun standart parametrlardan foydalaning. Nihoyat, "Docking->Output->LamarckianGA,

AutoDock-ni ADT-dan yoki buyruq satridan ishga tushiring. ADT-dan ishga tushirish uchun "Run->RunAutoDock" ni tanlang va ochilgan oynada "Browse" tugmalaridan foydalanib, AutoDock4 bajariladigan faylga yo'lni va docking parametr fayliga yo'lni belgilang. Docking parametr faylini ko'rsatganingizdan so'ng, ADT jurnal fayli uchun nom taklif qiladi. Dasturni ishga tushirish uchun "Ishga tushirish" tugmasini bosing.

Shu bilan bir qatorda, buyruq satrida terminal oynasini oching va koordinata fayllari va o'rnatish parametrlari faylini o'z ichiga olgan katalogga o'ting. Keyin buyruqni bering:

```
./autodock4 -p liep.dpf -l liep.dlg
```

Bu buyruq AutoDock bajariladigan autodock4 ham xuddi shu katalogda joylashganligini nazarda tutadi.



RasMol ish stoli

Muammolarni bartaraf qilish; nosozliklarni tuzatish. Raccoon2-ni ishga tushiring va serverni sozlang. Raccoon Linux klasteri kabi katta hisoblash resursida virtual skrinini ishga tushirish uchun mo'ljallangan. Yangi server qo'shsangiz, ulanishni sozlashingiz va bir yoki bir nechta o'rnatish xizmatlarini o'rnatishingiz



kerak. Raccoon2 ni ishga tushiring ([3-rasm](#)) va ulanish menejerini ochish uchun uchta tishli belgini bosish orqali yangi server ulanishini yarating. Bir nechta variantlarni o'rnatish kerak: server nomi (resursni aniqlash uchun nom); manzil (xost nomi yoki serverning IP manzili); va foydalanuvchi nomi va parol. "Saqlash" tugmasini bosning va ulanish menejerini yoping.

Ulanish menyusida serverni tanlash orqali docking xizmatini o'rnatish. Serverni tayyorlash ("rakunlashtirish") uchun rakun panjasi belgisini bosning. "Xizmat menejeri" belgisini (tayoqcha) bosning va ortiqcha belgisini bosish orqali yangi xizmatni qo'shing va unga o'zingiz tanlagan tavsiflovchi nom bering ("Vina docking xizmati"). "Serverga AutoDock Vina-ni o'rnatish" tugmasini bosning va bir nechta standart tanlovlarni qabul qiling. Saqlash tugmasini bosning va Xizmat menejerini yoping. Nihoyat, "Mavjud xizmatlar" panelidan "Vina docking xizmati" ni tanlang.

Muammolarni bartaraf qilish; nosozliklarni tuzatish. Ligand kutubxonasini sozlang. Ushbu protokol ma'lum inhibitor imatinibni o'z ichiga olgan 500 ta birikmaning kichik kutubxonasi bilan C-Abl virtual skriningini tavsiflaydi. Ligand kutubxonalarini yaratish va qayta ishlash vositalari <http://autodock.scripps.edu/resources/databases> saytida mavjud. Ligand kutubxonalarini serverda saqlanadi va ortiqcha ishlarni kamaytirish va tajribalarni kuzatish va takrorlash imkonini berish uchun docking ishlari uchun mavjud bo'ladi.

LIGANDS yorlig'ida "AddLigand->ScanDirectory" ni tanlang va ligandlar kutubxonasi joylashgan katalogni ko'rib chiqing. Hisobot oynasini yoping va "Yuklash" ni tanlang, o'zingiz tanlagan tavsiflovchi kutubxona nomini kiriting va "Ishga tushirish" tugmasini bosning. Yuklash tugallangach, kutubxona menejerini yoping. Kerakli ligand kutubxonasini tanlash uchun o'ng tugmasini bosning.

Qabul qiluvchilarning koordinatalarini o'rnatish. RECEPTORLAR ko'rinishida "Qo'shish->ScanDirectory" ni tanlang, PDBQT retseptor faylini ko'rib chiqing (4-bosqichda yaratilgan) va "OK" tugmasini bosning. Hisobot oynasini yoping.

AutoDock Vina docking parametrlarini sozlang. CONFIG yorlig'ida: Retseptorlar ro'yxatidagi fayllarni tanlash orqali 3D ko'ruvchiga retseptor strukturasi yuklang. Qidiruv maydonini aniqlash uchun "Markaz" va "O'lcham" tugmalaridan foydalaning yoki sichqoncha kursorni qutilar ustiga qo'ying va raqamli qiymatlarni kiriting. Agar xohlasangiz, o'rnatish parametrlarini o'zgartiring yoki 5(A)i da yaratilgan config.txt faylini yuklang.

Perform the virtual screening calculation. In the JOB MANAGER tab, ensure that all "Cluster submission requirements" are green, then click "Submit." Several options will need to be specified: set "Project" to "new" and type a descriptive name of your

choice; set “Experiment” to “new” and type a descriptive name of your choice; and click “OK” to start the calculation. To follow the status of the docking, right click on the experiment name and select “update status.” When one or more jobs in the experiment are finished, download them individually by right-clicking on the job name, or together by right-clicking on the experiment and select “Download results.”

Natijalarni filtrlang va tahlil qiling. Filtrlash va tanlashda yordam berish uchun Raccoon o'zaro ta'sir, ball va ligand samaradorligi kabi joylashtirish pozalarini hisoblab chiqadi. TAHLIL->MA'LUMOT MANBAI yorlig'ida "Natijalarni qo'shish->Yuklab olingan natijalarni tanlash" tugmasini bosib va natijalar katalogini ko'rib chiqing. Natijalarni qayta ishlash uchun ishlatiladigan retseptor faylini tanlang va tavsiflovchi jurnal fayli nomini kiriting yoki 5(A)i-bosqichda yaratilgan config.txt faylini tanlang.

"Ma'lumotlar manbai" oynasida (yuqorida) virtual ekran natijalarini filtrlash uchun bir nechta variant mavjud. Bu erda ikkita shovqin filtri qo'llaniladi. "Visualizatsiya" oynasi (pastki) filtrlangan natijalarni vizuallashtirishga imkon beradi va eksport uchun kerakli birikmalarni tanlash uchun katakchalardan foydalanish mumkin.

Natijalarni eksport qilish. Ligandlar to'plami filtrlangandan so'ng, Natijalar panelidagi tugmani bosish orqali qiziqarlilarini tanlang. Tanlangan nomzodlarni eksport qilish uchun EKSPORT yorlig'idan foydalaning.

(F) Aniq suv bilan ulash - vaqt 60 min

Bu usul hidratsiyaning barcha mumkin bo'lgan joylariga mos keladigan ligandga qo'g'irchoq atomlarni qo'shadi. O'rnatish vaqtida o'zgartirilgan AutoGrid xaritasi qo'llaniladi, bu suv yaxshi joylashtirilganda qulay ball beradi va agar u retseptor bilan bir-biriga to'g'ri kelsa, suvni o'tkazib yuboradi. Yakuniy skript o'rnatilgan natijalarni tahlil qiladi, faqat o'sha suvlarni tegishli pozitsiyalarda saqlaydi.

Ushbu protokol ligand va retseptor standart AutoDock o'rnatish uchun tayyorlangan deb taxmin qiladi. Yuqoridagi misoldagi 1–4-bosqichlarda tasvirlanganlar bilan bir xil protokol yordamida yaratilgan ikkita koordinatali fayl, ligand.pdbqt va protein.pdbqt, ushbu protokol uchun kuch maydoni uchun o'zgartirilgan parametr fayli bilan birga taqdim etilgan. o'zgartirilgan docking parametr fayli va har bir qadamni bajarish uchun bir nechta python skriptlari.

Ligandga suv pozitsiyalarini qo'shish uchun buyruq satriga yozing:

```
$ python wet.py -i ligand.pdbqt
```

Ushbu skript ligand PDBQT fayliga "W" atomlarini qo'shib, uni "ligand\_HYDRO.pdbqt" fayliga saqlaydi.

Standart atom panjara xaritalarini hisoblang. ADT-ni ishga tushiring va standart ishchi katalogni o'rnatib (1-bosqichga qarang). "Ligand->Input->Open" ni tanlang va ligand PDBQT faylini (ligand.pdbqt) tanlang va "Grid->Macromolecule->Open" ni tanlang va PDBQT retseptor faylini (receptor.pdbqt) tanlang. "Grid->SetMapTypes->Directly..." da xarita turlarini o'rnatib va "HD" va "OA" turlari ro'yxatda ekanligiga ishonch hosil qiling. Qidiruv maydonining markazi va hajmini belgilash uchun "Grid->GridBox" dan foydalaning. Qalqib chiquvchi oynada minimal qutini aniqlash uchun "Center->CenteronLigand" dan foydalaning. Tugatgandan so'ng, "Fayl->CloseSavingCurrent" yordamida parametrlarni saqlang. Grid parametr faylini "Output->SaveGPF" yordamida "protein.gpf" nomi bilan yozing. AutoGrid-ni ishga tushiring (qo'shimcha ma'lumot uchun 5(B)ii-bosqichga qarang) "Run->RunAutoGrid" yoki buyruq satrida:

```
/autogrid4 -p protein.gpf -l protein.glg."
```

"W" xaritasini yarating. Agar xaritalar uchun standart fayl nomlari ishlatilsa, suv energiyasini baholash xaritasini yaratuvchi skript uchun faqat retseptor nomi ko'rsatilishi kerak:

Standart docking parametr fayli ikki usulda qo'lda o'zgartirilishi kerak: "parameter\_file AD4\_water\_forcefield.dat" yuqoriga yaqin qo'shilishi kerak va "protein.W.map" xaritalar ro'yxatiga qo'shilishi kerak. Ushbu o'zgarishlarning ikkalasi [3](#) -bandda ta'kidlangan .

AutoDock-ni ishga tushiring (qo'shimcha ma'lumot olish uchun 5(B)v-bosqichga qarang) ADT-da "Run->RunAutoDock" yoki buyruq satrida:

```
./autodock4 -p ligand_HYDRO_protein.dpf -l
```

```
ligand_HYDRO_protein.dlg.
```

Extract and score the results at the command line with:

```
$ python dry.py -c -r protein.pdbqt -m protein.W.map -i
```

```
ligand_HYDRO_protein.dlg
```

Bu skript ajraladigan suvni aniqlash uchun retseptor yordamida o'rnatish natijalarini filtrlaydi va saqlanib qolganlarni KUCHLI yoki ZAFIYAT deb tasniflash uchun W xaritasi. Bu hisoblangan energiya bilan ligand\_LELC\_DRY\_SCORED.pdbqt nomli faylni yozadi.

## Nazorat savollari

1. Ligant nima?
2. Ligant retseptor ta'sirlashuv nima?
3. Ligant retseptor ta'sirlashuvni oldindan aytish qanday amalga oshadi?
4. AutoDock nima?
5. AutoDock dasturini qanday ishga tushirish mumkin?

### **Informatsion-uslubiy ta'minot**

5. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
6. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
7. Каменская Г.И. Биоинформатика. Москва, 2008

### **QO'SHIMCHA ADABIYOTLAR**

1. Yi-Ping Phoebe Chen (Ed.) Bioinformatics Technologies, 1998.
2. А. Н. Огурцов Введение в Бионформатику, 2011г.

### **Internet ma'lumotlari**

4. <https://www.ncbi.nlm.nih.gov/orffinder/>
5. [https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE\\_TYPE=BlastSearch&LINK\\_LOC=blasthome](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome)
6. <https://www.cgl.ucsf.edu/chimera/>
8. <https://proteopedia.org/wiki/index.php/Chimera>
9. [https://proteopedia.org/wiki/index.php/Molecular\\_modeling\\_and\\_visualization\\_software](https://proteopedia.org/wiki/index.php/Molecular_modeling_and_visualization_software)
10. <https://proteopedia.org/wiki/index.php/Jmol>

## **12-Amaliy mashg'ulot.**

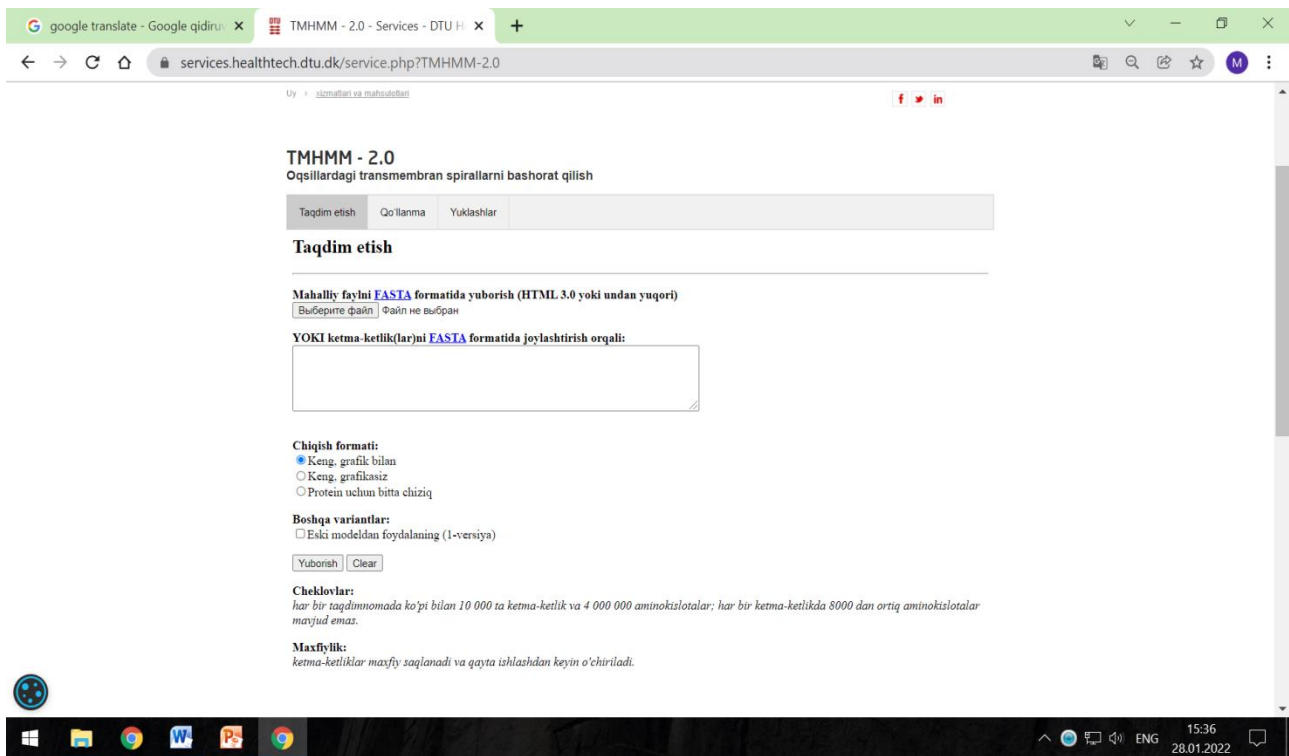
**Mavzu:** TMHMM dasturi yordamida transmembran spirallarini oldinday aytish.

**Dars maqsadi:** Talabalarga TMHMM dasturi yordamida transmembran spirallarini oldinday aytish bo'yicha amaliy ko'nikma uyg'otish.

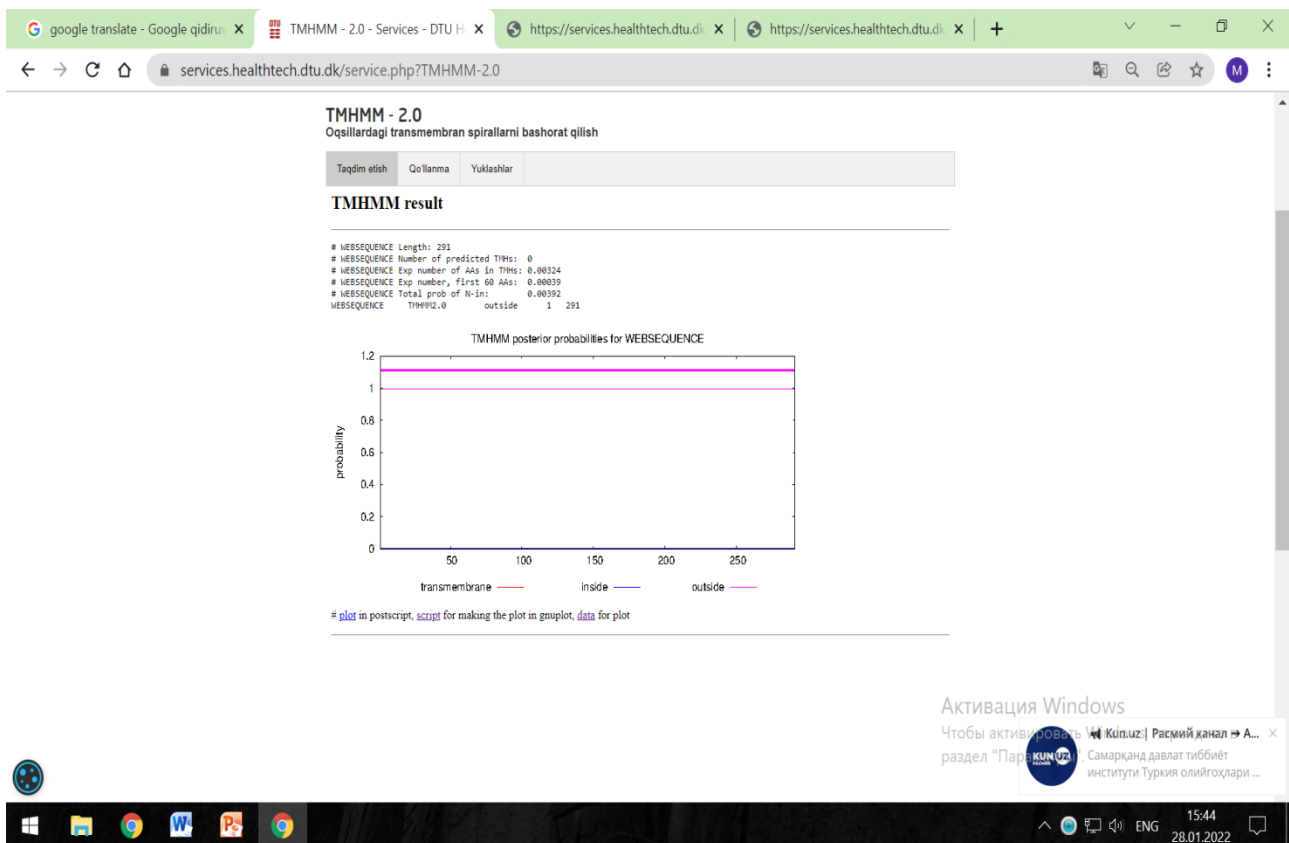
**Kerakli jihoz va preparatlar:** Kampyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** Transmembran spiralning topologiyasini bilish bog'lanish joylarini aniqlashga va membrana oqsillarining funktsiyalarini aniqlashga yordam beradi. Biroq, membrana oqsillarini eritish va tozalash qiyin bo'lganligi sababli, faqat juda oz miqdordagi membrana oqsillari eksperimental tarzda aniqlangan tuzilishga va topologiyaga ega. Bu membrana oqsillari topologiyasini bashorat qilish uchun turli xil hisoblash usullarini talab qiladi. Markov modeliga asoslangan membrana oqsillari topologiyasini bashorat qilishning yangi usuli TMHMM ni tasvirlaymiz va sinab ko'ramiz. Biz TMHMM ning batafsil ishlash tahlilini taqdim etamiz va u transmembran bobinlarining 97-98% ni to'g'ri taxmin qilishini ko'rsatamiz. Bundan tashqari, TMHMM eruvchan va membrana oqsillarini o'ziga xosligi va sezgirligi 99% dan yuqori bo'lgan holda ajrata oladi, garchi signal peptidlari mavjudligida aniqlik kamayadi. Bu yuqori darajadagi aniqlik bizga genomlarning katta to'plamidagi integral membrana oqsillarini ishonchli bashorat qilish imkonini berdi. Ushbu bashoratlarga asoslanib, biz ko'pchilik genomlardagi barcha genlarning 20-30 foizini membrana oqsillarini kodlashini taxmin qilamiz, bu avvalgi taxminlarga mos keladi. Bundan tashqari, biz N(in)-C(in) topologiyasiga ega bo'lgan oqsillar barcha o'rganilgan organizmlar uchun ko'proq afzal ekanligini aniqladik. *Caenorhabditis elegans* bundan mustasno, bu erda ko'p sonli 7TM retseptorlari N (tashqarida) -C (in) topologiyalari sonini oshiradi. Biz membrana oqsillarini yig'ish mexanizmlarini tushunish uchun ushbu kashfiyotning mumkin bo'lgan ahamiyatini muhokama qilamiz.

**Uni qanday ishga tushirish kerak.** Oynaning yuqori yarmida oqsillar mavjud bo'lgan mahalliy fayl nomini bering yoki ketma-ketlikni (lar) oynaning pastki qismiga joylashtiring. Keyin "Yuborish" tugmasini bosing. Unda siz mahalliy faylni yoki FASTA formatdan olgan ketma-ketliklarni joylashtirishingiz mumkin.



Dastur oqsillarni **FASTA formatida oladi**. U 20 ta aminokislota va B, Z va X ni taniydi, ularning barchasi noma'lum deb hisoblanadi. Boshqa har qanday belgi X ga o'zgartiriladi, shuning uchun ketma-ketliklarning sezgir oqsillar ekanligiga ishonch hosil qiling.



Transmembran oqsillarini identifikatsiya qilish va topologiyani bashorat qilish uchun yaxshilangan yashirin Markov modeli TMMODni taqdim etamiz. Bizning modelimiz prototip sifatida TMHMM dan foydalanadi, lekin membrananing har ikki tomonidagi halqalar uchun submodellar arxitekturasida, shuningdek, modelni o'qitish tartibida TMHMM dan farq qiladi. Ma'lum topologiyaga ega bo'lgan 83 ta transmembran oqsillari to'plamidan foydalangan holda o'zaro tekshirish tajribalarida TMMOD topologiya va joylashuv bo'yicha 89% aniqlik bilan TMHMM va boshqa mavjud usullardan ustun keldi. 160 ta transmembran oqsilining bitta to'plamidan foydalangan holda boshqa tajribada TMMOD 84% topologiyaga va 89% joylashuvga ega edi. Transmembran oqsillarini transmembran bo'lmagan oqsillardan, ayniqsa signal peptidlaridan aniqlash uchun foydalanilganda, TMMOD doimiy ravishda TMHMM ga qaraganda kamroq noto'g'ri musbat hosil qiladi. An'anaviy transmembran topologiyasi va TMHMM va SignalP kabi signal peptidlari bashorat qiluvchilaridan foydalanganda, bu ikki turdagi bashoratlar sezilarli darajada mos keladi. Ushbu usullarni beshta to'liq proteomaga qo'llash orqali biz taxmin qilingan barcha signal peptidlarining 30-65 foizi va barcha prognoz qilingan transmembran topologiyalarining 25-35 foizi bir-biriga mos kelishini aniqladik. Bu proteomaning 5-10% prognozlarini yomonlashtiradi, shuning uchun bu protein annotatsiyasida muhim masala. Ushbu muammoni hal qilish uchun biz ilgari transmembran topologiyasi va signal peptidlari bashoratlarini birlashtirgan Phobius Hidden Markov Modelini ishlab chiqdik. Usul transmembran segmentlari va signal peptidlari o'rtasida optimal tanlovni amalga oshiradi, shuningdek, cheklovlar va homologiya bilan bashorat qilish imkonini beradi.

### **Nazorat savollari**

1. TMHMM nima?
2. TMHMM qanday ishga tushirish kerak?
3. Transmembran spirallarini oldinday aytish deganda nimani tushunasiz?
4. Markov modeli?
5. TMHMM dastur oqsillarni qayerdan olishi mumkin?

### **Informatsion-uslubiy ta'minot**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Горбань А.Н. Нейроинформатика. Новосибирск: Наука 1998.
3. Каменская Г.И. Биоинформатика. Москва, 2008



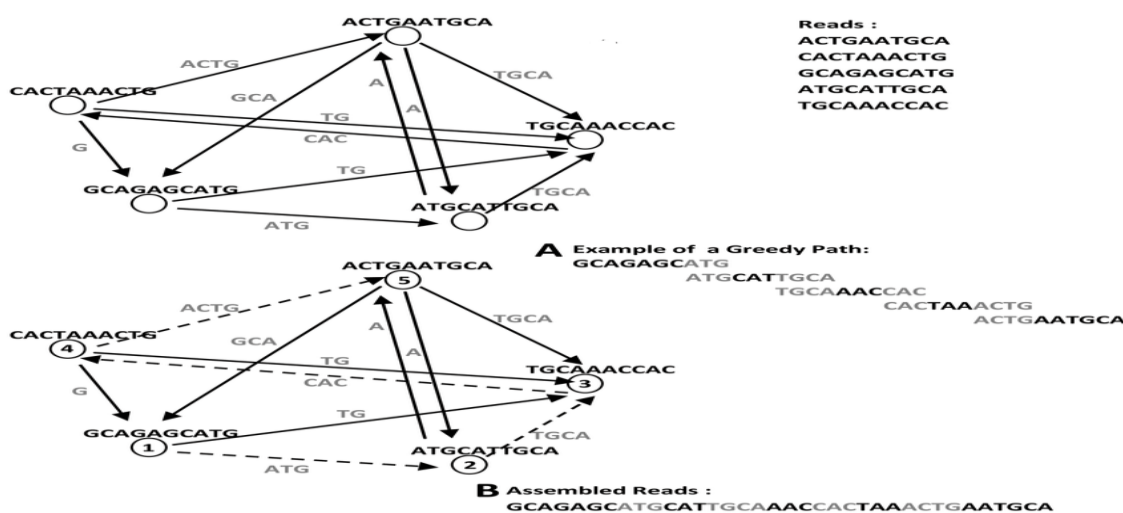
### 13-Amaliy mashg'ulot.

**Mavzu:** GREED texnologiya. Biologiyada superkompyuterlarning qo'llanilishi.

**Dars maqsadi:** Talabalarga GREED texnologiya. Biologiyada superkompyuterlarning qo'llanilishi bo'yicha amaliy ko'nikma uyg'otish.

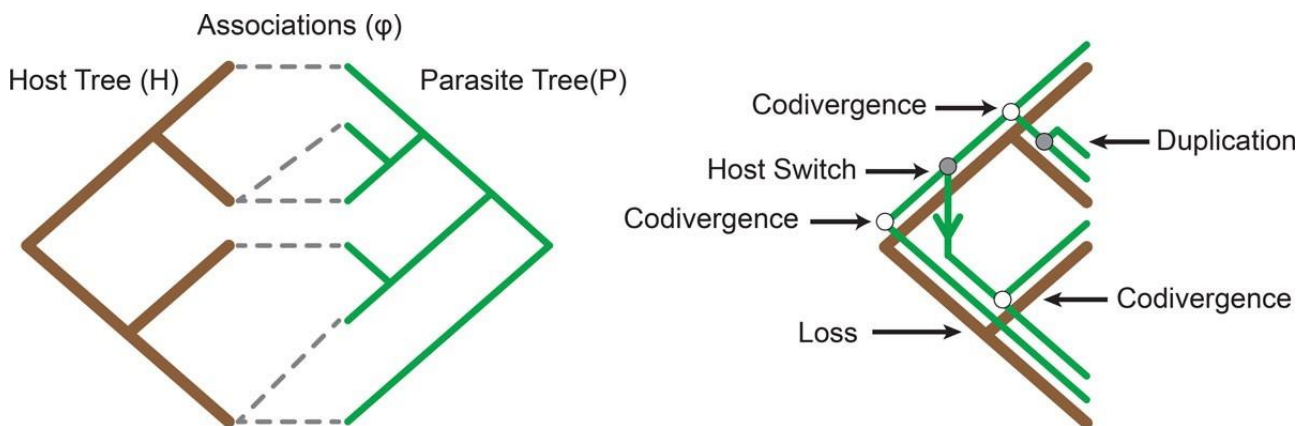
**Kerakli jihoz va preparatlar:** Kampyuter va uning dasturlari, internet tarmog'i, videoprektor.

**Ishning mazmuni:** Samarali va mijozlarga qulay komponentlarni (ya'ni superkompyuterlarni) jalb qilish tufayli ma'lumotlarni tezkor tahlil qilish amalga oshirilmoqda. Bioinformatikada u genomika, proteomika, metabolomika va boshqalar kabi ko'plab tadqiqot yo'nalishlarini kengaytirmoqda. Strukturaga asoslangan dori dizayni inson kasalliklarini davolash uchun tadqiqotning asosiy yo'nalishlaridan biridir. Ushbu bob ketma-ketlik tahlili uchun foydalaniladigan dasturiy ta'minot va veb-ga asoslangan dasturlarni jamlagan batafsil jadval bilan ketma-ketlik tahlilida super-kompyuterlarni muhokama qilishni boshlaydi. DOCK, ZINC, EDULISS va boshqalar kabi ma'lumotlar bazalari bilan tanishtirilgan virtual skrinindagi superkompyuterlar haqida qisqacha ma'ruza qilinadi. Bo'lim keyingi bosqichga o'tayotganda, Fragmentga asoslangan 2D QSAR, Multiple-Field 3D QSAR, va Aminokislotalarga asoslangan peptidlarni bashorat qilish abstraktsiya tushunchasiga o'xshash tarzda keltirilgan. Protein-ligand docking, Protein-protein docking va Multi-protein docking uchun docking dasturiy ta'minoti ta'minlangan joyda docking tadqiqotlaridagi superkompyuterlar ta'kidlanadi. Bo'lim keng qo'llaniladigan mikroarray ma'lumotlar tahlilida superkompyuterlarni qo'llash bilan yakunlanadi.



Grafik qurish uchun GREED texnologiyasiga asoslangan yondashuv. (A) Maksimal bir-biriga yopishish uzunligi bo'yicha tugunlarga tashrif buyuradigan GREED yo'lining misoli (nuqtali o'qlar) (eslatma: boshlang'ich tugun tasodifiy tanlanadi; har bir tugunda GREED algoritmi keyingi tashrifchini o'rtasidagi maksimal qoplama uzunligidan kelib chiqqan holda tanlaydi. bu tugun va uning ulangan qo'shnilari). (B) ochko'z yo'lda o'tgan tugunlarga mos keladigan yig'ilgan o'qishlar.

Kofilogeniya xaritasi ikki yoki undan ortiq filogenetik tarixlar orasidagi chuqur koevolyutsion assotsiatsiyalarni makro koevolyutsion miqyosda ochish uchun ishlatiladi. Kofilogeniya xaritasi NP-Hard bo'lganligi sababli, bu usul eng ahamiyatsiz holatlardan tashqari hamma narsani hal qilish uchun ko'p jihatdan evristikaga tayanadi. E'tiborga molik yondashuvlardan biri ko'rib chiqilayotgan filogenetik tarixlar uchun mumkin bo'lgan sobit tugun tartiblarining eksponensial sonining faqat bir qismini qidirish uchun metaevristikdan foydalanadi. Bu alohida qiziqish uyg'otadi, chunki u biologik mumkin bo'lgan echimlarni kafolatlaydigan yagona ma'lum evristik hisoblanadi. Bu anjir va ularning changlatuvchi ari o'rtasidagi koevolyutsion assotsiatsiyalar, shu jumladan 200 dan ortiq taksonlar kabi yirik koevolyutsion tizimlarga e'tibor qaratish imkonini berdi. Ko'p sonli biologik stsenariylarni o'z ichiga olgan bo'lsa-da, har bir koevolyutsiya tizimi mustaqil evolyutsiya tarixidan iborat bo'lib, ko'pincha *mezbon* filogeniya deb ataladi va unga bog'liq bo'lgan evolyutsiya tarixi deb nomlanadi. *parazitlarning* filogeniyasi.



**Tanglegramma va kofilogeniya xaritasi.** Oddiy tanglegramma, unda optimal xaritalash barcha to'rtta qayta tiklanadigan hodisani o'z ichiga oladi.

*Kortejni* ( $H, P, \phi$ ) baholashning keng tarqalgan usuli - bu assotsiatsiyalar ( $\phi$ ) saqlanib qolishi uchun parazit daraxti xost daraxtiga tushirilgan *kofilogeniya xaritasi*. Kofilogeniya xaritasi ( $P(\phi)$ ) *kodivergensiya*, *duplikatsiya*, *xostni almashtirish* va *yo'qotish* kabi to'rtta tiklanadigan koevolyutsion hodisalardan foydalangan holda, uning xostiga nisbatan parazitning evolyutsion tarixini baholaydi. Kodivergensiya hodisasi ham xost, ham parazit turlarining bir vaqtda ajralishidir. Duplikatsiya *hodisasi*, aksincha, parazit turlarining mustaqil divergensiyasidir. Ushbu ajralish hodisasidan so'ng ikkala yangi parazit turlari

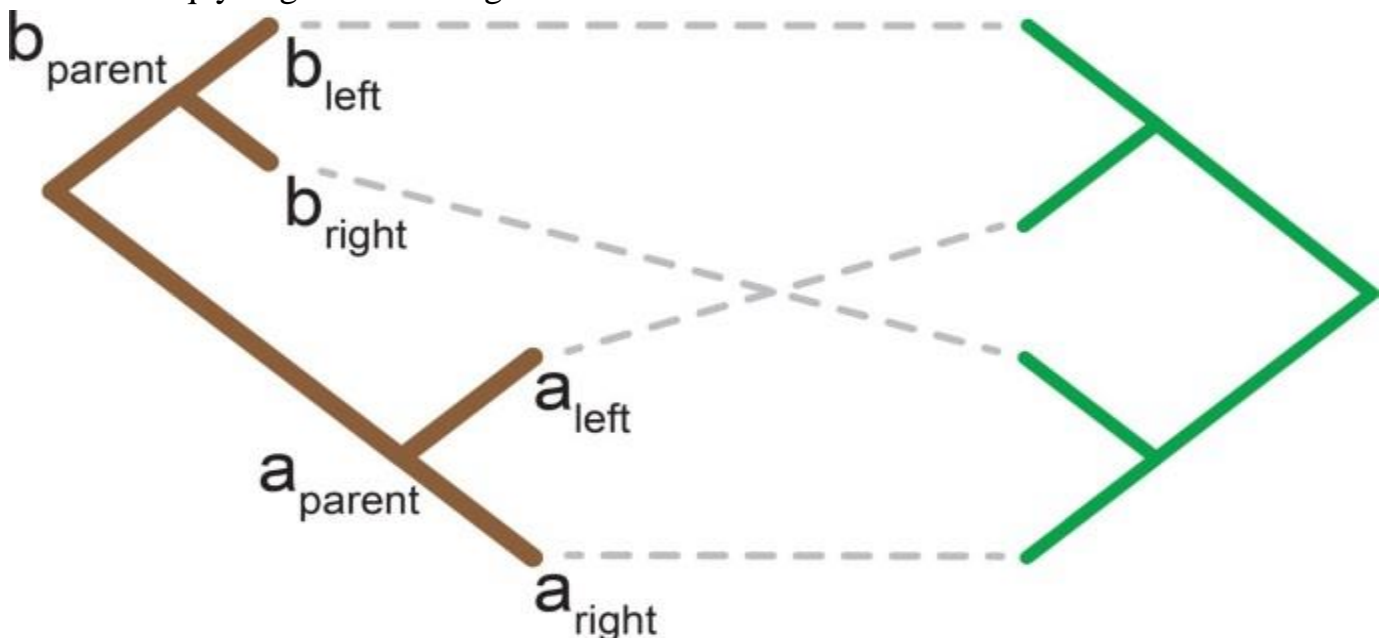
o'zlarining dastlabki xostlarini kuzatishda davom etadilar. Ko'paytirish hodisasiga o'xshash xost o'tkazgichi parazit turlarining mustaqil divergentsiyasidir. Ko'paytirish hodidasidan farqli o'laroq, yangi parazit turlaridan biri yangi xostga o'tadi, qolgan parazit turlari esa dastlabki xostda yashashda davom etadi. *Yo'qotish* hodisa uchta imkoniyatdan birini ifodalaydi: mezbon tomonidan divergentsiya hodidasidan keyin parazit turlari tomonidan ajralmasligi, parazit turining yo'q bo'lib ketish hodisasi yoki parazit filogeniyasining filogenetik rekonstruksiyasida namuna olish xatosi. Ushbu to'rtta hodisa birgalikda qo'llanilganda, har bir parazit faqat bitta xostda yashashi mumkin bo'lgan xaritalarning to'liq to'plamini tiklashga qodir. Bu erda ko'rib chiqiladigan kofilogeniyani qayta qurish muammosining o'zgarishi. Oddiy tanglegrammani yechish uchun qo'llaniladigan barcha to'rtta qayta tiklanadigan hodisalarning misolini yuqoridagi rasmda ko'rish mumkin. Kofilogeniya xaritasini tiklashda maqsad, natijada xarajat minimallashtirilgan yechimni olishdir. Minimal xarajatga ega bo'lgan barcha xaritalashlar to'plami Pareto to'plamida joylashgan. Kofilogeniyani rekonstruksiya qilish muammosi NP-Hard sifatida tashkil etilgan xaritalash narxi minimallashtirilgan Pareto to'plamidan xaritani qayta qurish muammosi sifatida belgilanadi, bu esa o'z navbatida bir qator evristik yondashuvlarni keltirib chiqardi. . Hozirgi vaqtda evristik usullarning ikkita klassi, naqsh va hodisaga asoslangan usullar mavjud. Voqealarga asoslangan algoritmlar tiklangan yechimlarning maqbulligini kafolatlaydigan yagona yondashuv ekanligi ko'rsatilgan. Ronquist, ammo naqshga asoslangan algoritmlar yechim maydoniga cheklovlar qo'yadigan voqealarga asoslangan usullar kabi yaxshi bo'lishi mumkin bo'lgan ishonchli taxminlarni ishlab chiqishi mumkinligini ta'kidladi. Ushbu muammoni hisoblashning qiyinligi tufayli cheklovlarni qo'yish talab qilinadi va shuning uchun naqshga asoslangan usullar hodisalarga asoslangan usullar bilan taqqoslanadigan aniqlikni taklif qilish potentsialiga ega. Mavjud naqshga asoslangan yondashuv - bu sahifaning muvofiqlashtirilgan daraxt tahlili, bu muammoni asosiy kalitlarga ruxsat berilmagan holda optimal tarzda hal qiladi. Bu algoritmi Pareto optimal yechimini  $O(n)$ . Ushbu yondashuv, optimal yechimni tiklash bilan birga, xostpatogen tizimlar yoki yaqindan bog'liq parazitlar uzoqdan bog'liq bo'lgan xostlarni yuqtirgan holda, xostni almashtirish hodisalari keng tarqalgan bo'lgan koevolutsion o'zaro ta'sirlarni aniq ifodalovchi echimlarni tiklay olmaydi. Yana bir taniqli naqshga asoslangan usul - Brooks Parsimony Analysis (BPA). Bu usul parazit filogeniyasining ikkilik kodlash tasviridan foydalanadi va parazitni xostga solishtirish uchun bu ikkilik kodlashdan foydalanadi. Bu usul yaxshi yechimlar ishlab chiqarish uchun ko'rsatilgan, lekin xo'jayin filogeniyasida hodisalarni tanlash va tugunlar tartib nomuvofiqliklari tufayli hali amalga oshirilmagan.

Ushbu ikkala yondashuv duch keladigan cheklovlar tufayli, naqshga asoslangan algoritmlar bu muammoga tez-tez qo'llanilmaydi, so'nggi kofilogenetik tahlillarning aksariyati hodisaga asoslangan usullardan foydalangan holda ushbu muammoni hodisaga asoslangan usullardan foydalangan holda yaqinlashish global ball minimallashtirilgan kofilogeniya xaritalarini tiklashga qaratilgan. Ushbu yondashuv

Ronquistning kofilogeniyani qayta qurish muammosining umumlashtirilgan ta'rifiga asoslanadi, unda barcha hodisalarning umumiy xarajatlarini minimallashtirish orqali algoritm Paretoning optimal rekonstruksiyasini tiklashi mumkin. Ushbu umumlashtirilgan ta'rifdan foydalangan holda so'nggi ikkita algoritmik yondashuv kofilogeniyani qayta qurish muammosining o'zgartirilgan holatlariga dinamik dasturlashni qo'llaydi. Birinchi yondashuv asosiy daraxtning ichki tugunlarining nisbiy tartibini e'tiborsiz qoldiradi va ikkinchi yondashuv asosiy daraxtning ichki tugunlarining nisbiy tartibini aniqlaydi.

Xost daraxtning ichki tugunlarining nisbiy tartibini e'tiborsiz qoldirib, kofilogeniyani qayta qurish masalasini polinom vaqtda hal qilish mumkin. Ushbu yondashuv, shuningdek, NP-Hard bo'lgan Feedback Arcs Set muammosini taxmin qilish uchun ishlatiladigan evristikdan olingan. Tarzan va CoRe-PA polinom vaqtda kofilogeniya xaritalarini tiklash uchun ushbu texnikani qo'llaydi. Ushbu yondashuv hozirda  $O(n^2)$  da ishlaydigan eng tez taklif qilingan algoritm bilan to'rtta qayta tiklanadigan hodisalarga ruxsat berilgan yechimlarni tiklash uchun eng tez ma'lum bo'lgan yondashuvdir.

TreeCollapse birinchi marta Ronquist tomonidan taklif qilingan naqshni aniqlash metodologiyasining kengaytmasidir. Ronquistning naqshni aniqlash tizimidagi sezilarli o'zgarish Ronquistning tarix tasniflash texnikasidan ko'ra mahalliy naqshni aniqlashdan foydalanishdir. Taklif etilayotgan naqshni aniqlash tizimi har bir qadamda faqat doimiy miqdordagi gilosni hisobga olgan holda har bir naqshning qidiruv maydonini doimiy o'lcham bilan chegaralaydi. Gilos ( $C$ ) McKenzie va Steel tomonidan quyidagicha ta'riflangan:



**Asosiy naqshlar.** TreeCollapse tomonidan aniqlanishi mumkin bo'lgan asosiy to'rtta naqsh va ularning natijaviy xaritalari.

Naqshlar chuqurlikdagi birinchi qidiruvni amalga oshirish orqali xost daraxtidagi har bir gilos uchun aniqlanadi. Bu algoritm tanglegram misolini  $(H, P, ph)$  aylanib o'tish uchun mo'ljallangan bo'lib, u gilosning barglaridan boshlanib, asosiy daraxtdagi bargga qaytish yo'li topilgandan keyin tugaydi. Ushbu algoritm  $O(1)$  da ishlashini ta'minlash uchun birinchi qidiruv algoritmi parazit daraxti orqali o'tishi mumkin bo'lgan darajalar soni doimiy koeffitsientga o'rnatiladi. Algoritm tugallangandan so'ng, chuqurlikdagi birinchi qidiruv algoritmi tomonidan topilgan natija yo'li ruxsat etilgan naqshlar to'plami bilan taqqoslanadi. *Ushbu qidiruv algoritmi  $O$  ga bog'langan(1)* chunki izlanadigan darajalarning doimiy soni bor va ikkiga bo'lingan daraxtning shoxlanish tezligi ham doimiydir. Shuni ta'kidlash kerakki, bu naqshni aniqlash algoritmini doimiy vaqtda amalga oshirishning faqat bitta usuli.

Mahalliy naqshni aniqlash algoritmi tomonidan tiklangan har bir naqsh xost yoki parazit daraxtlaridan bir yoki bir nechta gilosning qulashiga olib keladi, shuning uchun algoritm TreeCollapse nomini oladi. Bu operatsiya pastdan yuqoriga iteratsiya texnikasiga asoslangan bo'lib, bu erda mezbon daraxtdagi gilos ildizdan masofaga qarab ketma-ket qayta ishlanadi. Natijada, bu algoritm ikkala daraxt ham faqat bitta tugundan iborat bo'lgunga qadar davom etadi (ham xost, ham parazit daraxtining ildizi).

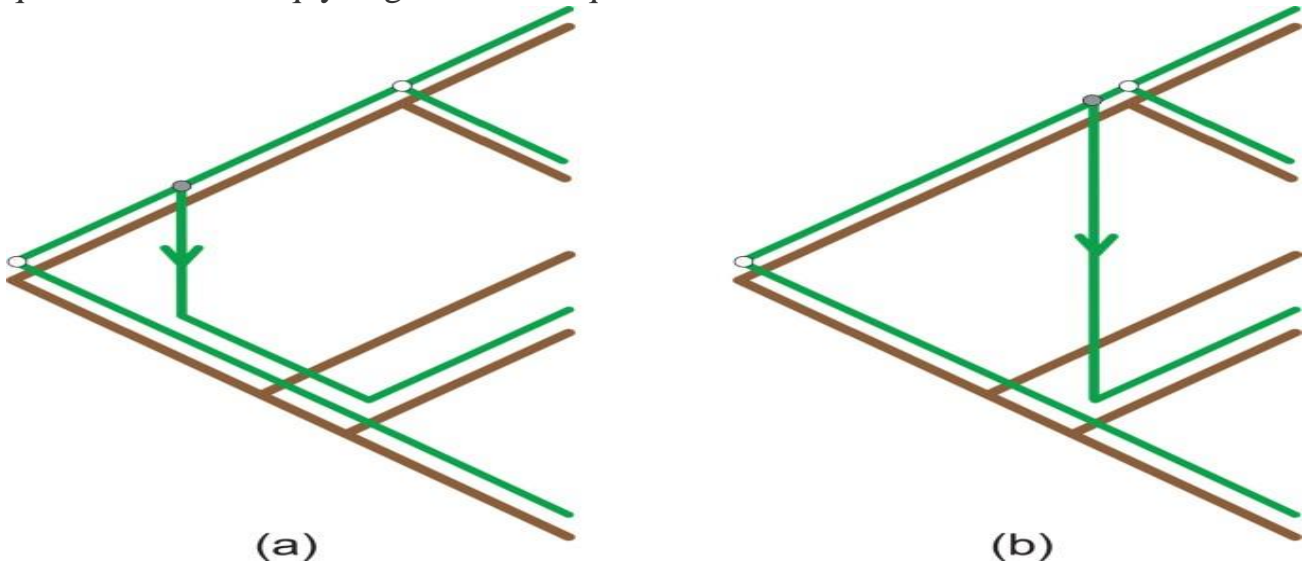
TreeCollapse mahalliy naqshni aniqlash tizimi uch xil siqilish funksiyasini bajarishga imkon beradi. Birinchi holat xo'jayin ham, parazit olchasi ham qulab tushadi, bu kodivergensiya hodisasini ko'rsatadi. Ikkinchi holat - parazit daraxtidan gilos qulab tushishi, bu dublikatsiya yoki xostni almashtirish hodisasini ko'rsatadi. Nihoyat, agar parazit daraxti bilan naqsh hosil bo'lmasa, u holda yo'qotish hodisasi aniqlanadi va uy egasi gilos qulab tushadi.

Ushbu jarayon tugaguniga qadar amalga oshirilganda uchta yakuniy holat mavjud. Birinchidan, mezbon daraxt va parazit daraxti ikkalasi ham tegishli ildiz tugunlarigacha yiqiladi. Bunday holda, xost daraxtining ildizida kodivergensiya hodisasi mavjud. Ikkinchi holat, parazit daraxti to'liq qulab tushganidan keyin ham mezbon daraxtda bir nechta tugunlar mavjud. Bunday holda, biz xost daraxtidagi barcha qolgan tugunlarni e'tiborsiz qoldiramiz. Nihoyat, xo'jayin daraxti parazit daraxti paydo bo'lgunga qadar to'liq qulab tushsa, parazit daraxtidagi qolgan tugunlar ko'payish hodisalarini sifatida xost daraxtining ildizi oldiga qo'shiladi. TreeCollapse barcha uchta stsenariyni ko'rib chiqadi, u kofilogeniyaning qayta qurish muammosining barcha holatlari uchun echimlarni tiklashga qodir.

TreeCollapse vaqtga mos keladigan yechimlarni tiklashini kafolatlang

TreeCollapse xost daraxtidagi har bir gilos tugun chuqurligiga qarab kamayish tartibida qayta ishlanishini talab qiladi. Chuqurlik bu kontekstda ildizgacha bo'lgan masofa filial uzunliklarining yig'indisi sifatida aniqlanadi.  $A$  tugunining chuqurligiga  $d(a)$  deb murojaat qilamiz. Vaqtning nomuvofiqligini oldini olish

uchun TreeCollapse xost daraxtidagi chekka og'irliklari har bir ichki tugun chuqurligi noyob bo'ladigan tarzda o'rnatilishini va 0 va  $(n - 2)$  oraliq'ida chegaralanganligini talab qiladi, bu erda ichki va barg tugunlarining umumiy soni daraxt  $(2n - 1)$  bo'ladi. Ushbu yondashuv ilgari tiklangan yechimlarning biologik jihatdan amalga oshirilishini ta'minlash uchun dinamik dasturlash yondashuvlari tomonidan qo'llanilgan. Keyin  $n$  barg tugunlari umumiy chuqurlikka  $(n - 1)$  tayinlanadi. Bunday qurilish ostida biz quyidagilarni da'vo qilamiz:



**O'ngga surish** . O'ngga surishdan oldin va keyin kofilogeniya xaritasiga misol qo'llaniladi.

**Ta'rif.** Faraz qilaylik,  $p \in PH(P)$  ning  $p$  chap va  $p$  o'ngga,  $PH(p_{chap}) = s$  va  $PH(p)$  bo'lgan bolalari bor.  $(e_i, s) = t$  bo'lgan bolalari bor. Keyin optimal xostni almashtirish joyi  $e_i$  va  $e_j \in E(H)$  ( $e_i, s$ ) va  $(e_j, t)$  orasidagi yo'qotish hodisalarining umumiy soni minimal bo'ladi va bu erda  $e_i$  va  $e_j$  bir xil vaqt oraliq'ida yotadi .

Joriy kofilogeniyaning xaritalashda xost kalitini optimal holatga o'tkazish orqali o'ngga surish algoritmi joriy kodivergeniya, takrorlash va xostni almashtirish hodisalarini tartibini saqlab, yo'qotish hodisalarini sonini kamaytiradi, bu esa o'z navbatida joriy xaritalash saqlanib qolishini ta'minlaydi. vaqtga mos. Right Push-ning TreeCollapse ramkasining aniqligiga ta'sirini tahlil qilish Natijalarning bir qismi sifatida batafsil muhokama qilinadi.

TreeCollapse algoritmidagi birinchi qadam asosiy daraxtdagi joriy va kelajakdagi gilolar ro'yxatini yaratishni talab qiladi, ular ildizdan masofaga qarab kamayish tartibida tartiblanadi. Odatda bunday tartibsiz ro'yxatni saralash uchun  $O(m \log m)$  ish vaqti talab qilinadi. Biroq, bu holda, biz quyidagilarni ko'rsatish uchun sobit tugun tartiblarining ba'zi xususiyatlaridan foydalanishimiz mumkin:

Ikkinchi bosqich har bir gilolar  $C$  uchun barcha naqshlarni ochishni talab qiladi;  $i$  uchun barcha naqshlarni ochishni talab qiladi, ulardan  $O(k)$  gacha, chunki parazit daraxtida har bir tugun uchun naqsh mavjud. Ushbu naqshlarni qayta ishlash uchun ikkita mumkin bo'lgan yondashuvni ko'rib chiqamiz. Birinchisi, qo'pol kuch yondashuvi, qayta ishlangan har bir hodisa uchun barcha mumkin bo'lgan naqshlarni ochishdir. Bu usul gilolar  $C_i$  uchun barcha hodisalarini qayta ishlash uchun  $O(k^2)$



vaqtni. Muqobil yondashuv shundan iboratki, har bir gilos  $C_i$  uchun tiklangan barcha naqshlar qayta ishlanadigan naqshlar ro'yxati sifatida saqlanadi. Agar  $C_i$  bu naqshlar ro'yxati qayta ishlanmasidan oldin yiqilib, keyin bu ro'yxat keyingi qayta ishlash uchun ota-onasiga ajratiladi. Bu yondashuv, oldingi qo'pol kuch usulidan farqli o'laroq, har bir naqsh har bir tugun uchun faqat bir marta ochilishini va  $O(k)$  naqshlar ro'yxatini ketma-ket qayta ishlanishini talab qiladi. Agar ikkinchi yondashuv qo'llanilsa, biz quyidagilarni isbotlashimiz mumkin:

Lemma 2 Xost daraxtining har bir tugunida hosil bo'lgan naqshlar soni  $O(km)$

**Super kompyuterlarni qo'llash.** Turli xil hisoblash muammolarini hal qilish uchun zamonaviy superkompyuterlardan foydalanishga misollar ilmiy va kosmik tadqiqotlar, dizaynni ishlab chiqish, yangi dori-darmonlarni izlash va meteorologiyada.

O'tgan yilning noyabr oyida Livermor milliy laboratoriyasida IBM Blue Gene/L superkompyuteri o'rnatildi. AQSh Energetika vazirligi xodimi Lourens dunyodagi eng kuchli superkompyuterlar TOP500 ([www.top500.org](http://www.top500.org)) rasmiy reytingida to'rt yillik yetakchiligini tasdiqladi. Bu dunyodagi eng tezkor kompyuter bo'lib chiqdi - bu reytingga kiritilgan barcha boshqa tizimlardan deyarli uch barobar tezroq. O'tgan yozda Blue Gene/L superkompyuterining konfiguratsiyasi 478 TFLOPS nominal ishlashga erishish uchun kengaytirildi.

Dunyodagi 2-raqamli va Yevropadagi eng tezkor superkompyuter Jülich tadqiqot markazida (Shimoliy Reyn-Vestfaliya, Germaniya) o'rnatilgan IBM Blue Gene/P superkompyuter liniyasining yangi modelidir. JUGENE hisoblash tizimida 65 000 dan ortiq PowerPC 450 protsessorlari mavjud va ushbu Blue Gene/P konfiguratsiyasining ishlashi 167 TFLOPS ni tashkil qiladi. JUGENE, JUMP va JUBL tizimlarining hisoblash quvvati taxminan ikki yuz tadqiqot guruhi uchun mavjud. Superkompyuterlarning resurslari, jumladan, yangi materiallarni ishlab chiqish, yangi avlod dori vositalarini izlash bilan bog'liq loyihalarni hisoblashda, shuningdek, iqlim o'zgarishini modellashtirishda, elementar zarrachalar harakati, murakkab kimyoviy reaksiyalar va boshqalarda qo'llaniladi. Mustaqil ekspertlar guruhi loyihalar o'rtasida hisoblash quvvatini taqsimlash uchun mas'uldir.

Superkompyuter sohasida sanoat yetakchisi sifatida IBM ([www.ibm.com](http://www.ibm.com)) sekundiga 1000 trillion (1 kvadrillion) hisob-kitoblarni amalga oshirishga teng bo'lgan petaflops (PFLOPS) bilan o'lchanadigan keyingi samaradorlik to'sig'iga yaqinlashmoqda. Petaflops tizimlari yuqori aniqlikdagi bashorat va simulyatsiya imkoniyatlari bilan fan va texnologiyani inqilob qilishi kutilmoqda. Masalan, zilzilalarni modellashtirishda Kaliforniyadagi San-Andreas yorig'i deb ataladigan hududlarda yer qobig'ining tom ma'noda binodan binoga siljishini kuzatish mumkin, bu esa zilzilalarga chidamli bino va inshootlarning loyihasini takomillashtirish imkonini beradi.



IBM petaflops tizimlari davriga yo'l ochish uchun mo'ljallangan bir nechta superkompyuter platformalarini ishlab chiqmoqda. Kelajakdagi petaflop va yuqori samaradorlik uchun mo'ljallangan Blue Gene/P dastlab ilmiy-tadqiqot ishlariga mo'ljallangan, ammo uning kengaytirilgan xotira resurslari va simmetrik ko'p ishlov berish (SMP) klaster tugunlari uni boshqa ilovalarning keng doirasi uchun juda jozibador qiladi. Bundan tashqari, bu yilgi POWER protsessorlarining so'nggi avlodi asosida qurilgan IBM superkompyuterlari HPC bozorida ob-havo prognozi, iqlim o'zgarishini modellashtirish va yangi energiya tadqiqotlari kabi tijorat va texnik ilovalar uchun yetakchi bo'lishi kutilmoqda. , avtomobilsozlik va aerokosmikda dizayn. tarmoqlar.

IBMning petaflops apparat tashabbuslari ilovalar va ishlab chiqish vositalarini qo'llab-quvvatlash uchun tegishli dasturiy investitsiyalar bilan mos keladi. Ushbu investitsiyalar tizimning unumdorligini, foydalanish qulayligini va loyihaning tijorat samaradorligini oshirishga qaratilgan. Masalan, bu yil IBM AQSh Energetika vazirligining Argonna milliy laboratoriyasi, Blue Gene/P joylashtirilishi rejalashtirilgan ilmiy markaz bilan birgalikda amalga oshiriladigan yangi ochiq kodli hamjamiyat ishlab chiquvchi dasturi orqali ilovalarni qo'llab-quvvatlashni kengaytiradi.

Blue Gene superkompyuterining operatsion tizimi ochiq kodli Linux operatsion tizimiga asoslangan. Bu muhit standartga asoslangan MPI (Message Passing Interface) aloqa protokollaridan foydalanadigan Fortran, C va C++ kabi umumiy dasturlash tillarida yozilgan ilovalarni ishga tushiradi. Blue Gene/P hozirda fizika, kimyo, biologiya, astrofizika, genetika, kosmologiya, seysmologiya kabi fanning ko'plab sohalarida tadqiqotlar uchun Blue Gene/L tizimlarida qo'llaniladigan turli xil ilovalarni qo'llab-quvvatlaydi.

### **Nazorat savollari**

1. GREED nima?
2. TMHMM qanday ishga tushirish kerak?
3. Transmembran spirallarini oldinday aytish deganda nimani tushunasiz?
4. Markov modeli?
5. TMHMM dastur oqsillarni qayerdan olishi mumkin?
6. Biologiyada super kompyuterlarni qo'llash deganda nimaani tushunasiz?
7. IBM Blue Gene/P oilasi haqida nimalarni bilasiz?

### **Mustaqil ta'lim bo'yicha materiallar**

## Mustaqil ish №1

Amaliy mashg'ulotlarga nazariy tayyorlanish.

**Maqsad:** Talaba amaliy mashg'ulotlarga o'zining maqsad va vazifalaridan kelib chiqib, yuqoridagi amaliy mashg'ulotlardan birontasini tanlab olib, (quyidagi tartibda) mustaqil bajarishi kerak.

**Mavzu:** Zamonaviy bioinformatsion ma'lumotlar bazalari haqidagi ma'lumotlar bilan tanishish. Bibliografik ma'lumotlar bazalari. Matnli ma'lumotlarni olish instrumentlari.

### Mavzu rejasi:

- 1) NCBI ma'lumotlar bazalari.
- 2) GenBank
- 3) EBI ma'lumotlar bazalari
- 4) EMBL ma'lumotlar bazalari
- 5) DDBJ ma'lumotlar bazalari
- 6) Meta bazalar
- 7) Malumotlar bazalarining amaliy ahamiyati.

### Foydalaniladigan adabiyotlar:

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерный исследований. 2002.
3. Каменская Г.И. Биоинформатика. Москва, 2008.
4. Ro'ziboev X.S., Radjabova G.G', Atamuratova N.R. Bioinformatika fanidan O'quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству insilico Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.
9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.

10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.

11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

## **Mustaqil ish №2**

**Mavzu:** Nukleotidli ketma – ketliklar bo‘yicha ma’lumotlar bazalari; polimorfizmlarning, mutatsiyalarning, nasliy kasalliklar; genetik xaritalar, to‘liq genomlar.

**Maqsad:** : Nukleotidli ketma – ketliklar bo‘yicha ma’lumotlar bazalari haqida ma’lumotga ega bo‘lish, mutatsiyalar va ularning kelib chiqish sabablarini tahlil qilish, irsiy kasalliklarni bashorat qilish, genetik xaritalash bo‘yicha dasturiy taminotlarni tahlil qilish, genomi to‘liq o‘qilgan organizmlarni tahlil qilish.

### **Mavzu rejasi:**

- 1) Nukleotidlar ketm-ketligi ma’lumotlar bazasi.
- 2) Mutatsiyalar va irsiy kasalliklar hamda ularning kelib chiqish sabablarini bashorat qilish.
- 3) Genomni o‘qish hamda genetik xaritalovchi dasturiy taminotlar. Genomi to‘liq o‘qilgan organizmlar.

### **Foydalaniladigan adabiyotlar:**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерный исследований. 2002.
3. Каменская Г.И. Биоинформатика. Москва, 2008.
4. Ro‘ziboev X.S. , Radjabova G.G‘., Atamuratova N.R. Bioinformatika fanidan O‘quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству insilico Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.

9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.
11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

### **Mustaqil ish №3**

**Mavzu:** Filogenetik daraxtlarni tuzish algoritmlari.

**Maqsad:** Filogenetik daraxtlarni tuzish algoritmlari haqida tushunchalarning shakllanishi, DNK va RNK hamda oqsil filogenetik daraxtlarini tuzishni o'rganish.

**Mavzu rejasi:**

- 1) Filogenetik daraxtlarni tuzish algoritmlarining modellari.
- 2) DNK va RNK filogeniyasi.
- 3) Oqsil aminokislotalar joylashish o'rni va gomologiyasini o'xshashlik va farqlarini tuzish algoritmlari.
- 4) Filogenetik daraxt qurishning UPGMA usuli.

### **Foydalaniladigan adabiyotlar:**

5. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
6. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерный исследований. 2002.
7. Каменская Г.И. Биоинформатика. Москва, 2008.
8. Ro'ziboev X.S., Radjabova G.G', Atamuratova N.R. Bioinformatika fanidan O'quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству insilico Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.

9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.
11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

#### **Mustaqil ish №4**

**Mavzu:** Eukariot organizmlarning gen strukturasi bashorat qilish.

**Maqsad:** Gen strukturalarini bashorat qilish metodlari. Gomologik ketma-ketliklar va ularni taqqoslash bo'yicha ko'nikmalarga ega bo'lish.

**Mavzu rejasi:**

- 1) Genetik xilma-xillik. Bir nukleotidli polimorfizmlar. Genom evolyutsiyasi.
- 2) Gen strukturalarini bashorat qilish metodlari.
- 3) Ekzon va intronlar hamda ularni taqqoslash.

#### **Foydalaniladigan adabiyotlar:**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерных исследований. 2002.
3. Каменская Г.И. Биоинформатика. Москва, 2008.
4. Ro'ziboev X.S., Radjabova G.G., Atamuratova N.R. Bioinformatika fanidan O'quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству insilico Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.
9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.

11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

### **Mustaqil ish №5**

**Mavzu:** Zamonaviy dunyoda bioaxborot xavfsizligi tushunchasi.

**Maqsad:** Zamonaviy dunyoda bioaxborot xavfsizligi, ishonchliligi, bioinformatik ma'lumotlarning amaliyotga qanchalik mosliligini hamda ularning ahamiyatini tahlil qilish.

#### **Mavzu rejasi:**

- 1) Zamonaviy ma'lumotlar bazalari va uni xavfsizligini ta'minlash bo'yicha say harakatlar.
- 2) Bioinformatik ma'lumotlar bazalari va ularning ishonchliligi.
- 3) Ma'lumotlar bazalari va ularning amaliy tadqiqotlarda qo'llash.

#### **Foydalaniladigan adabiyotlar:**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
2. Астахов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерных исследований. 2002.
3. Каменская Г.И. Биоинформатика. Москва, 2008.
4. Ro'ziboev X.S., Radjabova G.G., Atamuratova N.R. Bioinformatika fanidan O'quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству insilico Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.
9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.
11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

## Mustaqil ish №6

**Mavzu:** Dunyo bioinformatson va bibliografik ma'lumotlar bazasi va kooperatsiyasi.

**Maqsad:** Bioinformatik ma'lumotlar bazalari, hamda bibliografik hamkorliklar (cooperation), vazifa va maqsadlari haqida umumiy tanishish.

### Mavzu rejasi:

- 1) Bioinformatik ma'lumotlar bazalari va ularning rivojlanish tarixi.
- 2) Bibliografik ma'lumotlar bazalari hamda ularning eng asosiy turlari.
- 3) Bibliografik ma'lumotlar bazasida saqlanadigan axborotlar hamda ularning bir tizimda ekanligi.

### Foydalaniladigan adabiyotlar:

5. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
6. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерный исследований. 2002.
7. Каменская Г.И. Биоинформатика. Москва, 2008.
8. Ro'ziboev X.S., Radjabova G.G', Atamuratova N.R. Bioinformatika fanidan O'quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству *insilico* Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.
9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.
11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.



## Mustaqil ish №7

**Mavzu:** Evolyutsion biologiya, biologik xilma-xillikni baxolash.

**Maqsad:** Biologiyaning rivojlanish bosqichlari hamda hozir zamon evolyutsiyasi. Bioinformatik ma'lumotlar bazalarida yangi turlarni aniqlash bo'yicha tushunchalar.

### Mavzuning rejasi:

- 1) Biologiyaning ilk sivilizatsiyasidan toki hozirgi zamonaviy biologiya sohalarigacha bo'lgan davrda rivojlanish bosqichlar.
- 2) Biologik ketma-ketliklarni sekvinirlash usullari va ularning rivojlanishi
- 3) Bioinformatik ma'lumotlar bazalarida yangi turlarni aniqlash.

### Foydalaniladigan adabiyotlar:

9. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015
10. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерных исследований. 2002.
11. Каменская Г.И. Биоинформатика. Москва, 2008.
12. Ro'ziboev X.S., Radjabova G.G', Atamuratova N.R. Bioinformatika fanidan O'quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству *insilico* Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.
9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.
11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

## GLOSSARIY

<b>Bioinformatika</b>	biologik ma'lumotlarni qayta ishlash uchun kompyuter texnikasidan foydalanishdir.
<b>Arxivli ma'lumotlar bazalari</b>	(PDB, GenBank) har bir kiritilgan ma'lumot uchun muallif-tadqiqotchi javob beradi
<b>Nazorat qilinuvchi ma'lumotlar bazalari</b>	ma'lumotlar uchun maxsus xodimlar – kuratorlar javob beradilar
<b>Swiss- Prot</b>	Oqsillarning aminokislota ketma-ketliklarini saqlovchi baza
<b>KEGG</b>	metabolizm haqidagi ma'lumotlar bazasi
<b>FlyBase</b>	Drosophila haqidagi ma'lumotlar bazasi
<b>COG</b>	ortologik genlar haqidagi ma'lumotlar.
<b>COP</b>	Oqsillarning strukturaviy klassifikatsiyasi
<b>PFAM</b>	Oqsil oilalari bo'yicha ma'lumotlar banki
<b>GO (Gene Ontology)</b>	Genlar klassifikatsiyasi
<b>ProDom</b>	Oqsil domenlari
<b>AsMamDB</b>	Sutemizuvchilardagi alternativ splaysing bo'yicha ma'lumotlar bazasi
<b>NCBI Entrez</b>	Nukleotid yoki aminokislotalar ketma-ketliklari haqidagi ma'lumotlarga kirish uchun meta baza
<b>Ecocyc</b>	barchasi E. coli haqida – genlar, oqsillar, metabolizm va boshqalar.
<b>PubMed</b>	biologik tibbiyotga doir ma'lumotlar, adabiyotlar. Tarkibiga Medline –jurnallardan olingan maqolalar (tibbiyotga doir b'lgan) va taqrizlar kiradi.
<b>PMC</b>	Pub Med Central. bepul, jurnallardagi to'liq matnli maqolalar biologik tibbiyotga doir malumotlar, adabiyotlar.

<b>OMIM</b>	Insondagi Mendel qonunlari asosida irsiylanuvchi belgilar, genlar, mutatsiyalar
<b>OMIA</b>	Hayvonlardagi Mendel qonunlari asosida irsiylanuvchi belgilar, genlar, mutatsiyalar
<b>Nucleotide</b>	nukleotid ketma –ketliklarga doir ma’lumtlar bazasi (GenBank)
<b>Taxonomy</b>	GenBankdagi taksonlashtirilgan organizmlar
<b>SNP</b>	Single Nucleotide polymorphisms. yakka nukleotidning polimorfizmi
<b>HomoloGene</b>	eukariot organizmlar uchun xos bo’lgan genlar
<b>PubChem</b>	kichik molekulaning kimyoviy strukturalarini saqlovchi ma’lumotlar bazasi
<b>UniGene</b>	gen va gen guruxlari ketma –ketligiga asoslangan holda stenogrammani rasshifrovka qilish
<b>CDD</b>	oqsil domenlariga doir ma’lumotlar bazasi
<b>BLAST</b>	— nukleotidlar va aminokislotalar ketma-ketliklari bo’yicha ma’lumotlar bazalaridan qarindosh ketma –ketliklarni izlash, solishtirish
<b>Clustal W</b>	Nukleotidlar va aminokislotalar ketma-ketliklarining ko‘plik taqqoslanishlari
<b>DNASP</b>	DNK ketma – ketliklaridagi polimorfizmlarni tahlili
<b>UGENE</b>	Nukleotidlar va aminokislotalar ketma-ketliklarining ko‘plik taqqoslanishlari, filogenetik tahlil, annotirlash, ma’lumotlar bazalari bilan ishlash uchun erkin holatdagi rus tilidagi dastur.
<b>T-Coffee</b>	Nukleotidlar va aminokislotalar ketma-ketliklarining ko‘plik taqqoslanishlari dasturi. ClustalW va ClustalX ga qaraganda ancha sezgirroq
<b>PSI-Protein Classifier</b>	PSI-BLAST dasturi vositasida olingan natijalarning jamlanishi
<b>PHYLIP</b>	filogenetik dasturlar paketi

<b>JalView</b>	nukleotidlar va aminokislotalar ketma-ketliklarining ko‘plik taqqoslanishlari dastur muharriri
<b>FigTree</b>	filogenetik daraxtlarning muharriri
<b>FASTA</b>	Fast Alignment. Solishtiruvlar uchun eng qulay bo‘lgan matn formati.
<b>PyMol</b>	Makromolekularani vizualizatsiyalash uchun dastur
<b>EMBL</b>	Yevropa Molekulyar Biologiya Laboratoriyasi. Eng yirik ma’lumotlar bazalaridan biri.
<b>PDB</b>	Oqsillarning birlamchi, ikkilamchi, uchlamchi va to‘rtlamchi strukturasi haqida ma’lumotlar bazasi
<b>CRISPR</b>	Clustered Regularly Interspaced Short Palindromic Repeats, o‘zbek tilida - muntazam guruhlarda joylashgan qisqa palindromik takrorlar. Genomlarni tahrirlash texnologiyasi

**O‘ZBEKISTON RESPUBLIKASI**  
**OLIY VA O‘RTA MAXSUS TA‘LIM VAZIRLIGI**  
**GULISTON DAVLAT UNIVERSITETI**

**«TASDIQLAYMAN»**

O‘quv ishlari bo‘yicha prorektor

\_\_\_\_\_ J.H.Qarshiboyev

«\_\_» \_\_\_\_\_ 2022 yil

№ \_\_\_\_\_

**“BIOINFORMATIKA”**  
**FANINING ISHCHI O‘QUV DASTURI**

**Bilim sohasi: 100 000 – Gumanitar soha**

**Ta‘lim sohasi: 140000 – Tabiiy fanlar**

**Ta‘lim yo‘nalishi: 5140100 – Biologiya (Turlar bo‘yicha)**

Umumiy o‘quv soati: – 150 soat

Shu jumladan:

Ma‘ruza – 20 (5-semestr- 32 soat)

Amaliy mashg‘ulot – 40 (5-semestr- 52 soat)

Mustaqil ta‘lim – 90 (5-semestr-82 soat)

## GULISTON – 2022

Mazkur fanning ishchi o'quv dasturi O'zbekiston Respublikasi Oliy va o'rta maxsus ta'lim vazirligining 2020 yil 29 avgustdagi 452-sonli buyrug'i bilan tasdiqlangan "Bioinformatika" fani dasturi asosida tayyorlangan.

Mazkur ishchi fan dasturi "Biologiya" kafedrasining 2022 yil "\_\_\_" \_\_\_ dagi \_\_\_-sonli bayonnoma bilan yig'ilishida muhokama qilinib, tasdiqlash uchun tavsiya etilgan.

Mazkur ishchi fan dasturi "Tabiiy fanlar" fakulteti Kengashining 2022 yil "\_\_\_" \_\_\_ dagi \_\_\_-sonli bayonnoma bilan yig'ilishida muhokama qilinib, tasdiqlash uchun tavsiya etilgan.

Mazkur ishchi fan dasturi Guliston davlat universiteti Kengashining 2022 yil "\_\_\_" \_\_\_ dagi \_\_\_-sonli bayonnoma bilan tasdiqlangan

### **Tuzuvchi:**

H.Ro'ziboev,

Biologiya falsafa Doktori (PhD)

M.Tursunov

Biologiya kafedra o'qituvchisi

### **Taqrizchilar:**

M.Ergashev

GulDU Biologiya kafedra dotsenti.

GulDU Tabiiy fanlar

fakulteti dekani:

2022 yil "\_\_\_" "\_\_\_\_\_" \_\_\_\_\_ M.Ergashev

GulDU "Biologiya"

kafedra mudiri:

2022 yil "\_\_\_" "\_\_\_\_\_" \_\_\_\_\_ Z.Abdikulov

GulDU O'quv-uslubiy bo'lim boshlig'i :

2022 yil "\_\_\_" "\_\_\_\_\_" \_\_\_\_\_ I.Xudoyberdiev

## **I. Fanning mazmuni**

Fanni o‘qitishdan maqsad – Kompyuter texnologiyalari, bioinformatika usullari va genetik ma’lumotlarni intellektual tahlil qilish ko‘nikmalariga ega, shuningdek zamonaviy biomedikal uskunalardan uchun yangi bioinformatika texnologiyalardan foydalana oladigan mutaxassislarni tayyorlashdan iborat.

Fanning vazifasi – fanni chuqur o‘zlashtirishda nazariy bilimlar bilan amaliy mashg‘ulotlar uyg‘unlashtirilgan holda amalga oshirish, bioinformatikaning nazariy asoslari, asosiy tushunchalari, uning amaliyotda tadbqiq etish, tirik hujayrani molekulyar darajada o‘rganish, organizm genomining strukturaviy asoslarini tahlil qilish, umumiy biologik muammolarni makromolekulalar va hujayra asosida mantiqan yechimini tafakkur qila olish orqali insonning hayotdagi o‘rni va ahamiyatini ochib berishda

### **II. Asosiy nazariy qism (ma’ruza mashg‘ulotlari)**

#### **II.I. Fan tarkibiga quyidagi mavzular kiradi:**

#### **1-mavzu. “Bioinformatika” faniga kirish. Asosiy atamalar va tushunchalar.**

Bioinformatika faniga kirish. Bioinformatika fanining rivojlanish tarixi va bioinformatika fani istiqbollari. Asosiy atamalar va tushunchalar.

#### **2- mavzu. Zamonaviy bioinformatika ma’lumotlar bazalari**

“Axborot” va “Bioaxborot” tushunchasi. Axborot nazariyasi. Axborot xususiyatlari. Zamonaviy bioinformatika ma’lumot bazalari turlari. DNK va RNK nukleotidlar ketma,ketliklari ma’lumot bazalari (GenBank, EMBL, DDBJ). Metabazalar. Genom bazalari. Oqsil ketma,ketliklari bazalari (PIR, SWISS,PROT, UniProt, TrEMBL). Oqsil strukturalari bazalari. Metabolik yo‘llar bazalari. Molekulalarni modellashtirish bo‘yicha ma’lumotlar bazalari (MMDB, PDB, NCBI). PSR (Polimeraza zanjir reaksiyasi) bazalari.

#### **3- mavzu. Biologik ketma-ketliklarni taqqoslash**

Biologik ketma, ketliklarni taqqoslash asoslari. Gomologik ketma-ketliklar. Biologik ketma-ketliklarning yakka va ko‘plik taqqoslanishi. BLAST algoritmi. BLAST turlari. NCBI da BLAST. Biologik ketma,ketliklarni juft va ko‘plik taqqoslashlarni solishtirish. Mark yashirin modellari. Genlarni taqqoslash asosida turlarning filogenetik yaqinligini aniqlash.



#### **4- mavzu. Genom tahlili va ukariot organizmlar gen strukturalarini bashorat qilish.**

Genetik axborotning uzatilishi. Genlarning genomdagi lokalizatsiyasi. Pro, va eukariotlarning yaxshi o'rganilgan genamlari. Bir nukleotidli polimorfizmlar. Genetik xilma,xillik. Genom evolyusiyasi. Gen strukturalarini bashorat qilish metodlari. Gomologik ketma-ketliklar. Ekzon va intronlar. Hisoblashning ochiq ramkasini izlash. ORF Finder dasturi.

#### **5- mavzu. Molekulyar filogenetika**

Filogenetikaning asosiy tushunchalari. Filogenetik daraxtlarning tiplari. Zamonaviy bioinformatsion dasturlar (Clustal W2, T,Coffee). Genlarni solishtirish asosida filogenetik yaqinlikni aniqlash. Filogenetik daraxtlar klassifikatsiyasi. Filogenetik bog'lanishlarni aniqlash va filogenetik qarindoshlikni o'rnatish.

#### **6- mavzu. Biologik makromolekulalarni vizualizatsiyalashtirishning zamonaviy usullari**

Fazoviy strukturani vizualizatsiyalashtirishning asosiy prinsiplari. RasMol dasturi va unda ishlash tartibi. Biologik makromolekulalarning birlamchi strukturasi asosida ularni vizualizatsiyalashtirish. PyMol va I , TASSER dasturlarida ishlash. Yaratilgan strukturalarni PDB, MMDB ma'lumotlar bazalariga joylashtirish.

#### **7- mavzu. Oqsillarning strukturasi va xususiyatlarini in silico sharoitida o'rganish**

Oqsil strukturasi oldindan aytish va o'rganish bo'yicha zamonaviy yondashuvlar. Ramachandra xaritalari. Barqarorlik va oqsillar foldingi. Hidrofoblik profilining tahlili. Strukturaviy tekislanishlar. Oqsil strukturalarini modellashtirish va oldindan aytish. Genomlarda oqsil strukturalarini aniqlash. Evolyusiyada oqsil funksiyasining divergensiyasi.

#### **8- mavzu. Neyron to'rlari.**

Neyronlar, signal uzatilishi prinsipi. Sun'iy neyron to'rlari tushunchasi. Neyron to'rlarining mantig'i. Bir qavatli va ko'p qavatli perseptron. Ko'p qavatli perseptronni o'rgatish. Neyron to'ri tuzish. Neyron to'rlarining qo'llanilishi.

#### **9- mavzu. Dori vositalarini ishlab chiqishda bioinformatsion yondashuvlarning qo'llanilishi**

Farmakologik nishonlar. Nishonni aniqlash va tasdiqlash. Struktura –faollik munosabati. Yangi dori birikmalarining kompyuterli konstruksiyalanishi. Zamonaviy drug-designinstrumentlariva usullari. Dori vositalarining ratsional dizayni va personallashtirilgan tibbiyot. Oqsil,ligand bog‘lanishlarining modellashtirilishi. Zamonaviy drug-design maʼlumotlar bazalari va dasturlari. Kompyuterli toksikologiya va immunoinformatika.

<b>№</b>	<b>Ma’ruza mavzulari</b>	<b>Dars soatlari hajmi</b>
1.	Bioinformatika faniga kirish. Asosiy atamalar va tushunchalar.	2
2.	Zamonaviy bioinformatsion ma’lumotlar bazalari.	2
3.	Biologik ketma-ketliklarni taqqoslash.	2
4.	Genom tahlili va eukariot organizmlar gen strukturalarini bashorat qilish.	2
5.	Molekulyar filogenetika.	2
6.	Biologik makromolekulalarni vizualizatsiyalashtirishning zamonaviy usullari.	2
7.	Oqsillarning strukturasi va xususiyatlarini in silico sharoitida o‘rganish.	2
8.	Neyron to‘rlari.	2
9.	Dori vositalarini ishlab chiqishda bioinformatsion yondashuvlarning qo‘llanilishi.	2
<b>Jami:</b>		<b>20</b>

### **III. Amaliy mashg‘ulotlar bo‘yicha tavsiya va ko‘rsatmalar**

Amaliy mashg‘ulotlar multimedia qurilmalari bilan jihozlangan auditoriyada bir akademik guruhga bir professor-o‘qituvchi tomonidan o‘tkazilishi zarur. Mashg‘ulotlar faol va interfaktiv usullar yordamida o‘tilishi, mos ravishda munosib pedagogik va axborot texnologiyalar qo‘llanilishi maqsadga muvofiq. Mustaqil o‘zlashtiriladigan mavzular bo‘yicha talabalar tomonidan referatlar tayyorlash va uni taqdimot qilish tavsiya etiladi.

#### **Amaliy mashg‘ulotlar uchun tavsiya etiladigan ishlar ro‘yxati:**

<b>№</b>	<b>Amaliy mashg‘ulot mavzulari</b>	<b>Dars soatlari hajmi</b>
----------	------------------------------------	----------------------------

1.	Zamonaviy bioinformatson ma'lumotlar bazalari haqidagi ma'lumotlar bilan tanishish. Bibliografik ma'lumotlar bazalari. Matnli ma'lumotlarni olish instrumentlari.	2
2.	BLAST dasturi yordamida nukleotid ketma-ketliklarini solishtirish.	2
3.	BLAST dasturi vositasida aminokislotalar ketma-ketliklarini taqqoslash va tranlyantlarning olinishi.	4
4.	Oqsil molekularida gomologik domenlarni izlash.	2
5.	Hisoblashning ochiq ramkasini izlash. ORF Finder dasturida ishlash.	2
6.	Ko'plik taqqoslanishlar va filogenetik daraxtni yaratish. Clustal W va T-Coffee dasturlari. Ma'lumotlar tahlili.	6
7.	Biologik makromolekulalarni vizualizatsiya qilish dasturlari.	4
8.	Aminokislota ketma-ketligi bo'yicha fazoviy strukturani yaratish.	4
9.	I – TASSER va Modeller dasturlarining ishlash prinsipi	4
10.	RasMol dasturida oqsil strukturalarining qo'yilishi.	4
11.	Ligand, retseptor ta'sirlashuvlarini oldindan aytish.	4
12.	AutoDock dasturida ishlash.	2
<b>Jami:</b>		<b>40</b>

#### IV. Mustaqil ishni tashkil etishning shakli va mazmuni

Mustaqil ish uchun berilgan mavzularni talabalar mustaqil ravishda ko'rsatilgan adabiyotlar yordamida o'zlashtirib, oraliq nazorat shaklida yoki darslardan tashqari vaqtda referat, prezentatsiya yoki muloqot tarzida topshiradilar.

№	Mustaqil ta'lim mavzulari	Hajmi (soat)
1.	Amaliy mashg'ulotlarga nazariy tayyorlanish.	30
2.	Nukleotidli ketma – ketliklar bo'yicha ma'lumotlar bazalari; polimorfizmlarning, mutatsiyalarning, nasliy kasalliklar; genetik xaritalar, to'liq genomlar.	10
3.	Filogenetik daraxtlarni tuzish algoritmlari.	10
4.	Eukariot organizmlarning gen strukturasi bashorat qilish.	10
5.	Zamonaviy dunyoda bioaxborot xavfsizligi tushunchasi	10

6.	Dunyo bioinformatsion va bibliografik ma'lumotlar bazasi va kooperatsiyasi.	10
7.	Evolyutsion biologiya, biologik xilma-xillikni baxolash	10
	<b>Jami</b>	<b>90</b>

## V. Fan o'qitilishining natijalari (shakllanadigan kompetentsiyalar)

Fanni o'zlashtirish natijasida talaba:

- Zamonaviy bioinformatsion ma'lumot bazalari turlari, biologik ketma-ketliklarni taqqoslash asoslari, gen strukturalarini bashorat qilish va genom tahlili metodlari, filogenetik bog'lanishlarni aniqlash va filogenetik qarindoshlikni o'rnatish nazariyalari haqida tasavvur va bilimga ega bo'lishi;
- Fazoviy strukturani vizualizatsiyalashning asosiy printsiplari, oqsil strukturalarini modellashtirish va oldindan aytish, neyron tarmoqlari haqida asosiy bilimlar, zamonaviy drug-design instrumentlari, usullari hamda ulardan foydalanish ko'nikmalariga ega bo'lishi;
- statistik ma'lumotlarni tahlil qilish uchun zamonaviy kompyuter dasturlaridan foydalanish, ma'lumot olish uchun algoritmlardan foydalanish, ma'lumotlar bazasini boshqarish usullaridan foydalanish, biologik ketma-ketlikni taqqoslashda dasturlardan foydalanish, grafik modellashtirish qobiliyati, filogenetik tahlili, genomik ma'lumotlarni tahlil qilish dasturidan foydalani malakasiga ega bo'lishi kerak.

## VI. Ta'lim texnologiyalari va metodlari:

- ma'ruzalar;
- seminarlar ( mantiqiy fiklash, tezkor savol-javoblar);
- guruhlarda ishlash;
- taqdimotlar qilish;
- individual loyihalar qilish;
- jamoa bo'lib ishlash va himoya qilish uchun loyihalar;

## VII. Kreditni olish uchun talabalar:

Fanga oid nazariy va uslubiy tushunchalarni to'la o'zlashtirish, tahlil natijalarini to'g'ri aks ettira olish, o'rganilayotgan jarayonlar haqida mustaqil mushohada yuritish va joriy, oraliq nazorat shakllarida berilgan vazifa va topshiriqlarni bajarish, yakuniy nazorat bo'yicha yozma ishni topshirish.

## VIII. Asosiy va qo'shimcha o'quv adabiyotlar hamda axborot manbalari

### ASOSIY ADABIYOTLAR

5. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015

6. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерных исследований. 2002.
7. Каменская Г.И. Биоинформатика. Москва, 2008

### **Qo‘shimcha adabiyotlar**

1. Ro‘ziboev X.S., Radjabova G.G., Atamuratova N.R. Bioinformatika fanidan O‘quv uslubiy majmua, 2019 y.
2. Иванов А.С. Биоинформатика: путь от генома к лекарству insilico Вест. РГМУ. 2003. №4.
3. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
4. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
5. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.
6. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
7. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.
8. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

### **Axborot manbaalari**

1. [www.ziyonet.uz](http://www.ziyonet.uz)
2. [www.catuzmu](http://www.catuzmu)
3. [www.natl.uz](http://www.natl.uz)
4. [www.nature.uz](http://www.nature.uz)
5. [www.pedagog.uz](http://www.pedagog.uz)
6. [www.maik.ru](http://www.maik.ru)
7. [www.blast.nih.gov](http://www.blast.nih.gov)
8. [www.floranimal.ru](http://www.floranimal.ru)
9. [www.biology.ru](http://www.biology.ru)
10. [www.cellbio.com](http://www.cellbio.com)
11. [www.wikipedia.org](http://www.wikipedia.org)
12. [www.ncbi.gov.org](http://www.ncbi.gov.org)
13. [www.ornl.gov/hgmis/project/info.html](http://www.ornl.gov/hgmis/project/info.html)
14. [www.phylogeny.arizona.edu/tree](http://www.phylogeny.arizona.edu/tree)
15. [www.ebi.ac.uk/biocat/](http://www.ebi.ac.uk/biocat/)
16. [www.ebi.ac.uk/tools/index.html](http://www.ebi.ac.uk/tools/index.html),
17. [www.uniprot.org](http://www.uniprot.org),

### Talaba bilimini baholash

Talaba bilimini baholash kredit-modul tizimiga muvofiq ishlab chiqilgan Nizom asosida amalga oshiriladi.

Nazorat turi	1-OB	YaB
O'tkazilish vaqti	6- hafta	12-hafta
Nazorat shakli	Yozma*	Yozma*

**Oraliq baholash:** fanning ma'ruza qismiga tegishli teng yarmi o'tib bo'lingandan so'ng so'ng OB olinadi. Bunda o'tilgan mavzularga doir 3 ta nazariy yozma savollari varianti tarqatiladi. Oldindan tuzilgan 3 ta yozma variantlarini to'la echgan talabaga xar bir to'g'ri javob uchun maksimal 5 baho beriladi.

**Yakuniy baholash** o'tilgan barcha mavzular bo'yicha tuzilgan variantlari asosida o'tkaziladi. Bunda xar bir talabaga semestr davomida o'tilgan mavzular bo'yicha 3 ta nazariy va 1 tadan laboratoriya ishi bo'yicha og'zaki savol variantlari tarqatiladi. Talaba og'zaki javobning xar biridan maksimal 5 baho to'plash imkoniyatiga ega. Umumiy baxo o'rtacha arifmetika asosida chiqariladi.

*\*Izoh. Nazoratlardagi har bir savol va topshiriqlar quyidagi baholash mezonlari bo'yicha baholanadi.*

#### Talabalar bilimini baholash mezonlari

a) "5" (a'lo) baho uchun talabaning bilim darajasi quyidagilarga javob berishi lozim:

- Hulosa va qaror qabul qilish;
- Ijodiy fikrlay olish;
- Mustaqil mushohada yurita olish;
- Olgan bilimlarini amalda qo'llay olish;
- Mohiyatini tushunish;
- Bilish, aytib berish;
- Tasavvurga ega bo'lish;

b) "4" (yaxshi) baho uchun talabaning bilim darajasi quyidagilarga javob berishi lozim:

- Mustaqil mushohada yurita olish;
- Olgan bilimlarini amalda qo'llay olish;
- Mohiyatini tushunish;
- Bilish, aytib berish;
- Tasavvurga ega bo'lish;

v) **“3” (qoniqarli)** baho uchun talabning bilim darajasi quyidagilarga javob berishi lozim:

- Mohiyatini tushunish;
  - Bilish, aytib berish;
  - Tasavvurga ega bo'lish;
- g) talabning bilim darajasi **“2” (qoniqarsiz)** deb quyidagi hollarda baholanadi:
- Aniq tasavvurga ega bo'lmaslik;
  - Javoblarda xatoliklarga yo'l qo'yilganlik;
  - Bilmaslik.



**Nazorat uchun test savollari.**

**Bioinformatikada qo'llaniladigan asosiy dasturlarni ko'rsating?**

- A. ACT, Arlequin, Bioconductor,
- B. BioEdit, BioNumerics,
- C. A va B javoblar
- D. LAN, RNAi, cDNA

**Bioinformatika biologiya sohasining qaysi yo'nalishlarida qo'llaniladi?**

- A. Genomika, transkriptomika va proteomika;
- B. Gen tarmoqlarining kompyuter tahlili;
- C. Populyatsion genetikada modellashtirish.
- D. Barchasida.

**Parsimoniya usulidan foydalanib amalga oshiriladigan filogenetik tahlil dasturi**

- A. PAUP
- B. Tcoffee
- C. UGENE
- D. ZENBU

**Quyida keltirilgan dasturiy ta'minotlar dan qaysi biri birikkanlik kartalarini tuzishda ishlatiladi?**

- A. MapQTL;
- B. JoinMap;
- C. BLAST;
- D. Ugene.

**Odam genomi nechta nukleotiddan tashkil topgan?**

- A. 1 milliondan ortiq;
- B. 3,2 milliard;
- C. 10 million;
- D. 30 - 40 ming oralig'ida.

**Miqdoriy belgilar lokuslarini kartalashtirishda foydalaniladigan dasturiy ta'minot?**

- A. BLAST;
- B. Ugene;
- C. PubMed;
- D. MapQTL.

**Genomlar annotatsii birinchi dasturiy tizimi kim tomonidan qachon yaratilgan?**

- A. Vito Volterra tomonidan 1920 yilda;
- B. Owen White tomonidan 1955 yilda;
- C. Paul Naim Berg tomonidan 1972 yilda;
- D. Feng Zhang tomonidan 2015 yilda.

**Biologik tajribalarni kompyuter modellashtirish (simulyatsiya) qanday ataladi?**

- A. In vivo;
- B. In silico;
- C. CRISPR;
- D. Zinc Finger.

**Quyidagi javoblardan qaysi biri nukleotid ketma-ketliklar ma'lumotlar bazasiga to'g'ri keladi.**

- A. NCBI;
- B. PubMed;
- C. MMDB;
- D. ENTREZ;

**"Odam genomi" loyihasi qachon bajarilgan?**

- A. 1990-2005 yillar;
- B. 2000-2015 yillar;
- C. 1980-2000 yillar;
- D. Hali tugatilmagan.

**Odam genomidagi genlar soni nechta?**

- A. 1 milliondan ortiq;
- B. 30 - 40 ming oralig'ida;
- C. 2,5 milliard;
- D. 100 mingdan ortiq;

**Bioinformatikaning qanday bo'limlari bor?**

- A. Umumiy bioinformatika va klinik bioinformatika;
- B. Strukturaviy genomika va funktsional genomika;
- C. Klinik proteomika, funktsional proteomika strukturaviy proteomika;
- D. Barchasi.

**Populyatsion-genetik tahlil dasturi?**

- A. ACT
- B. PHYLIP
- C. Populations
- D. Velvet

**Filogenetik shajarani tahrirlashda qo'llaniladigan dastur,**

- A. Populations
- B. PSI Protein Classifier
- C. FigTree
- D. Sequin

**Bakterial genomlarni umumlashtirishda foydalaniladigan dastur?**

- A. JalView
- B. MacClade
- C. MEGA
- D. SPAdes

**Solishtirma biologiya uchun Java tizimida ishlaydigan dastur?**

- A. MEGA
- B. Mesquite

C.MacClade

D.JalView

**Populyatsion-genetik tahlil dasturi?**

A. ACT

B. PHYLIP

C.Genepop

D.Velvet

**Genetik xilma-xillik populyatsiyalarigi tahlil qilishda qo'llaniladigan dastur?**

A. BLAST,

B. Clustal,

C.DnaSP,

D.PopGene

**DNK ketma-ketligi polimorfizmi tahlili qaysi dastur yordamida amalga oshiriladi?**

A. DnaSP

B. Muscle

C.UGENE

D.MacClade

**Genomlarni umumlashtirishda foydalaniladigan dastur?**

A. SPAdes

B. JalView

C.Velvet

D.MEGA

**Filogenetik tahlilda qo'llaniladigan dastur?**

A. Seaview

B. FigTree

C.Populations

D.Sequin

**Natijalarni umumlashtiruvchi dastur?**

A. FigTree

B. MEGA

C.Populations

D.ZENBU

**Molekulyar-evolyutsion, genetik tahlil qilishda foydalaniladigan dastur?**

A. Populations

B. MEGA

C.UGENE

D.Muscle

## **ADABIYOTLAR**

1. Леск А. Введение в Биоинформатику. М., БИНОМ, 2015

2. Астахонов Т.В. Сравнительный анализ информационных биополимеров. Компьютеры и суперкомпьютеры в биологии. М. Ижевск: Институт компьютерный исследований. 2002.
3. Каменская Г.И. Биоинформатика. Москва, 2008
4. Ro'ziboev X.S., Radjabova G.G., Atamuratova N.R. Bioinformatika fanidan O'quv uslubiy majmua, 2019 y.
5. Иванов А.С. Биоинформатика: путь от генома к лекарству insilico Вест. РГМУ. 2003. №4.
6. Бауэр Ф.Л., Гооз Г. Информатика. Вводный курс. В 2 ч. М. Мир, 1990. Todaro, Michael P.
7. М. Бородовский, С. Екишева Задачи и решения по анализу биологических последовательностей М.-Ижевск : РХД, 2008.
8. Дромашко С.Е. Очерки биоинформатики. Минск, Беларуская наука, 2009.
9. Handbook of Molecular Microbial Ecology (F.J. de Bruijn, Ed.), Wiley-Blackwell, 2011.
10. Metagenomics: Theory, Methods and Applications. (D. Marco, Ed.), Caister Academic Press, 2010.
11. Metagenomics Methods and Protocols. (S. Wolfgang, D. Rolf, Eds.), Springer, 2010.

#### **Axborot manbaalari**

19. [www.ziyonet.uz](http://www.ziyonet.uz)
20. [www.catuzmu](http://www.catuzmu)
21. [www.natl.uz](http://www.natl.uz)
22. [www.nature.uz](http://www.nature.uz)
23. [www.pedagog.uz](http://www.pedagog.uz)
24. [www.maik.ru](http://www.maik.ru)
25. [www.blast.nih.gov](http://www.blast.nih.gov)
26. [www.floranimal.ru](http://www.floranimal.ru)
27. [www.biology.ru](http://www.biology.ru)
28. [www.cellbio.com](http://www.cellbio.com)
29. [www.wikipedia.org](http://www.wikipedia.org)
30. [www.ncbi.gov.org](http://www.ncbi.gov.org)
31. [www.ornl.gov/hgmis/project/info.html](http://www.ornl.gov/hgmis/project/info.html)
32. [www.phylogeny.arizona.edu/tree](http://www.phylogeny.arizona.edu/tree)
33. [www.ebi.ac.uk/biocat/](http://www.ebi.ac.uk/biocat/)
34. [www.ebi.ac.uk/tools/index.html](http://www.ebi.ac.uk/tools/index.html),
35. [www.uniprot.org](http://www.uniprot.org),
36. [www.expasy.org](http://www.expasy.org), w