



ӘЛ-ФАРАБИ АТЫНДАҒЫ ҚАЗАҚ ҰЛТТЫҚ УНИВЕРСИТЕТІ
БИОЛОГИЯ ЖӘНЕ БИОТЕХНОЛОГИЯ ФАКУЛЬТЕТІ
МОЛЕКУЛАЛЫҚ БИОЛОГИЯ ЖӘНЕ ГЕНЕТИКА
КАФЕДРАСЫ



Дәріс 11.

**Ұлттық Биотехнологиялық Ақпарат Орталығы
(National Center for Biotechnology Information – NCBI)**

Лектор: к.б.н., Алтыбаева Н.А.

Дәрістің мақсаты:

NCBI ресурстарымен жұмыс істеуді, геномдық және биоинформатикалық деректер қорларын қолдануды үйрету.

Қарастырылатын сұрақтар:

1. NCBI тарихы
2. NCBI құрылымы мен қызметтері
3. GenBank, PubMed, BLAST, GEO және басқа деректер базалары
4. Ресурстарды қолдану мысалдары

Ұлттық биотехнологиялық ақпарат орталығы (ағылш. National Center for Biotechnology Information –NCBI) — биомедициналық және генетикалық ақпараттарды сақтау, өңдеу және талдау үшін құрылған халықаралық ғылыми орталық. Ол 1988 жылы АҚШ-тың Ұлттық медицина кітапханасы (National Library of Medicine, NLM) мен Ұлттық денсаулық сақтау институтының (NIH) қолдауымен ұйымдастырылған.

Бүгінгі таңда NCBI – әлемдегі ең ірі биоинформатикалық деректер орталығы болып табылады. Онда миллиондаған ДНҚ және РНҚ тізбектері, ақуыз құрылымдары, ғылыми мақалалар мен генетикалық зерттеу нәтижелері сақталады.

Орталық зерттеушілерге GenBank, PubMed, BLAST, Gene, Protein сияқты танымал деректер базалары арқылы ғылыми ақпаратпен жұмыс істеуге жағдай жасайды.



Құрылу Тарихы

NCBI - Мэриленд штатының Бетезда қаласында орналасқан және 1988 жылы АҚШ конгресспені Клод Пеппердің демеушілігімен заң негізінде құрылған .NCBI 1988 жылы АҚШ-та биоинформатика және молекулалық биология салаларының қарқынды дамуына байланысты құрылған. Сол кезде ғалымдар жинақталған генетикалық деректердің көлемі жыл сайын еселеп өсіп жатқанын байқап, оларды орталықтандырылған жүйеде сақтау қажеттілігі туындады. Ұлттық медицина кітапханасы мен Ұлттық денсаулық сақтау институты осы мәселені шешу мақсатында NCBI орталығын ұйымдастырды. Бастапқыда оның басты міндеті — GenBank дерекқорын дамыту және биологиялық деректермен жұмыс істеудің бірыңғай жүйесін құру болды.

Орталық биомедициналық деректерді сақтау ғана емес, оларды талдау, салыстыру және визуализациялау құралдарын да ұсынады.

Уақыт өте келе NCBI қызметі кеңейіп, ол бүгінде биоинформатикалық зерттеулердің негізгі ақпараттықхабы болып саналады.

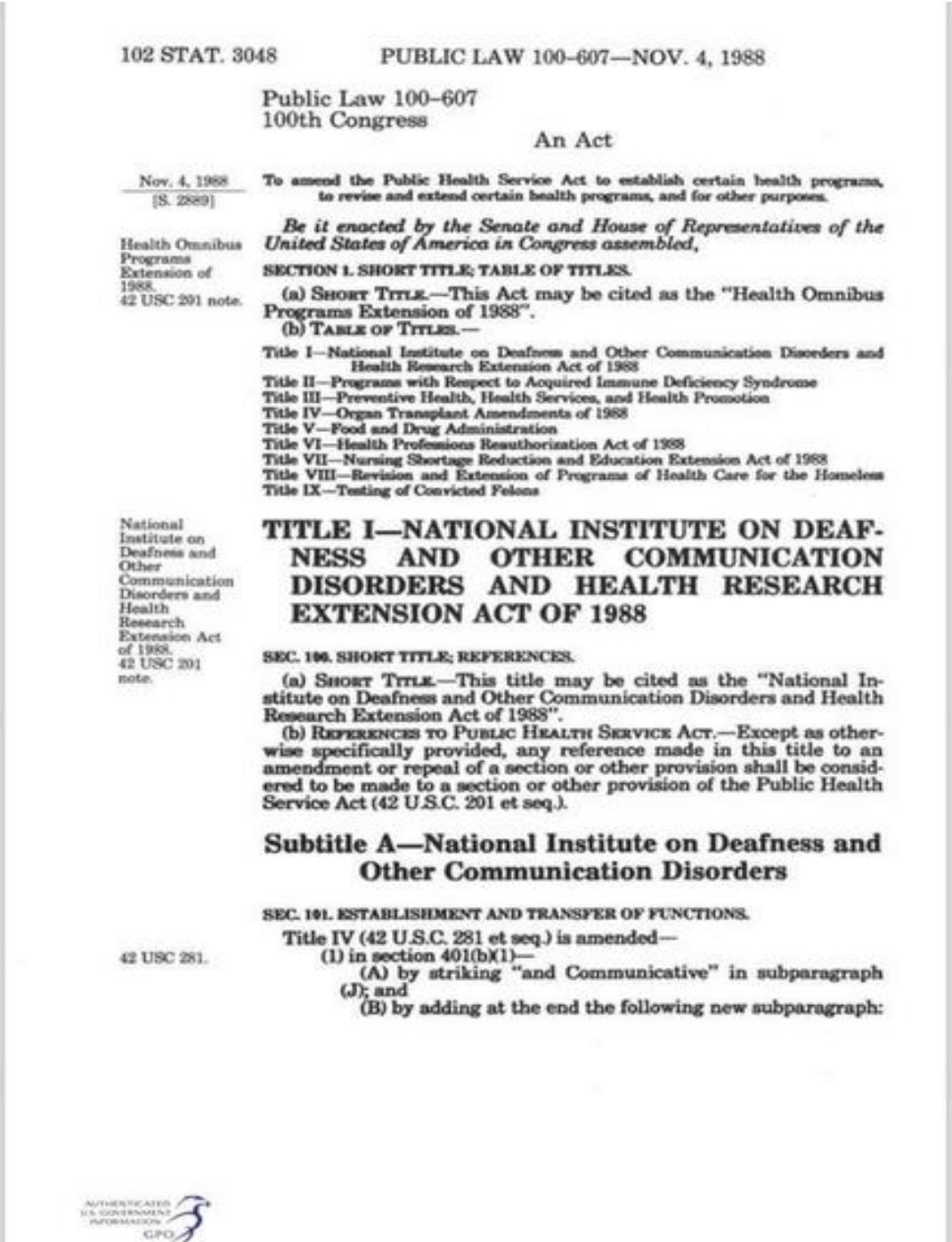


NCBI, Бетезда, Штат Мэриленд, США



Заңнамалық қолдау және бастамалар:

NCBI-ге заңнамалық қолдау және бастамалар негізінен АҚШ Конгресінде қабылданған заңдар мен актілер арқылы жүзеге асырылды, әсіресе 1980- жылдардың соңында және 1990-жылдардың басында. Негізгі заңнамалық негізі 1988 жылы 4 қарашада күшіне енген Health Omnibus Extension Act (P.L.100-607) болды. Бұл заң жобасы H.R.393 деп аталды және ол ұлттық биотехнология мен геномика саласындағы ақпараттық ресурстарды құру және дамытуға бағытталды. 1988 жылы 4 қарашада АҚШ Конгресі қабылдаған Health Omnibus Extension Act (P.L.100-607) заңымен NCBI-ге заңнамалық қолдау көрсетілді. Бұл заң нөмірі S. 2889 және H.R.393 деп аталған, ол ұлттық биотехнология мен геномика саласындағы ақпараттық ресурстарды құру мен дамытуға бағытталған. Заң биомедициналық зерттеулерді қолдауға арналған, оның ішінде геномика және ақпараттық технологиялар дамуына құқықтық негіз болды.



АҚШ Конгресінің Health Omnibus Extension Act (P.L.100-607) құжаты

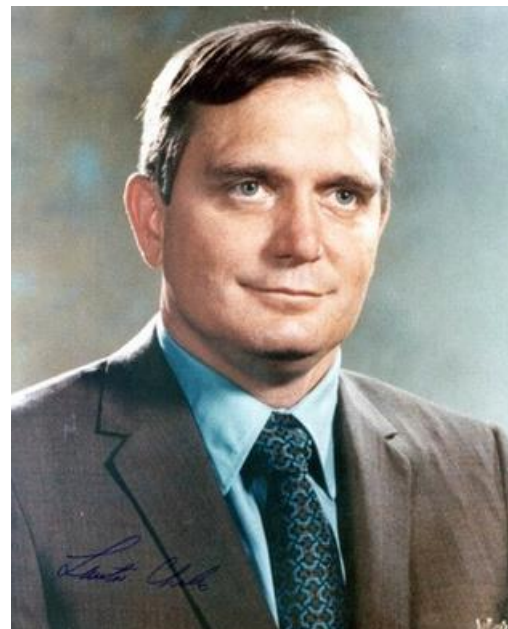


1987 жылы Сенатор Клод Пеппер төрағалық еткен Конгресстің Ардагерлер істері бойынша комитеті Ұлттық биотехнологиялық ақпарат орталығын (NCBI) құруға арналған заң жобасын (H.R.393) енгізді. Бұл бастама бірнеше жылдар бойы биомедицина саласындағы ғалымдардың және ғылыми қауымдастықтардың Конгресстегі қолдауына негізделді, оның ішінде Конгрессмен Генри Уэксман мен сенаторлар Эдвард Кеннеди және Лотон Чайлз белсенді қатысты. Заң жобасы биотехнология мен геномика саласындағы ақпараттық жүйелерді құруға бағытталып, 1988 жылы Health Omnibus Extension Act (P.L.100-607) ретінде қабылданып, Президент Рональд Рейганның қол қоюымен күшіне енді. Заң қабылдау рәсімінде сенатор Пеппер жаңа орталықтың адам өмірін ұзартуға және денсаулығын жақсартуға ықпал ететініне деген сенімін білдірді: *«Мен шын жүректен үміттенемін, бұл орталық біз ұзақ жылдар бойы армандаған мақсаттарды жүзеге асырады және адам өмірін жақсартып, олардың бақытына үлес қосады»*. Бұл тарихи заң биотехнологиялық зерттеулердің дамуына зор серпін беріп, NCBI-дің ақпараттық инфрақұрылымын құруға құқықтық негіз болды.



Henry A. Waxman

Конгрессмен, денсаулық сақтау саласында маңызды заңнамаларға қатысқан саясаткер. (4-сурет)



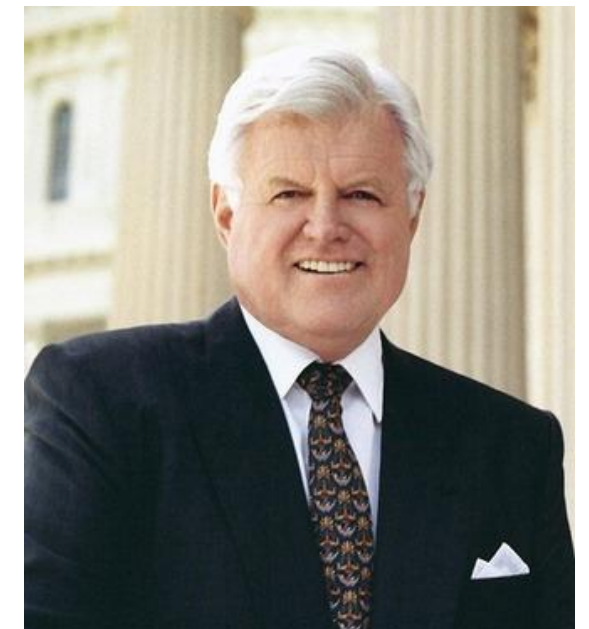
Lawton Chiles

Флоридаданшыққан сенатор және губернатор. (5-сурет)



Edward M. Kenned

«Тед» Кеннеди ретінде танымал сенатор, денсаулық сақтау мен әлеуметтік салаларда белсенді болған. (6-сурет)



Дэниел Масис (Daniel Masys) және **Денис Бенсон** (Dennis Benson) NCBI-дің ерте кезеңдегі заңнамалық қолдау және бастамаларында маңызды рөл атқарған тұлғалар ретінде белгілі. Олар АҚШ Ұлттық Медицина Кітапханасының құрамындағы Lister Hill Biomedical Communications Center-де жұмыс істеді және NLM-ның биотехнология ақпараттық қызметінің алғашқы жұмыстарын бастады. 1987-1988 жылдары Сенатор Клод Пеппердің қолдауымен ұлттық биотехнологиялық ақпараттық орталық құру туралы заң жобалары кезінде Дэниел Масис пен Денис Бенсон NLM-да бұл бастамаларды іске асыруға қатысып, NCBI-дың қалыптасуында жетекші қызметтер атқарды. Олар жаңа орталықтың ақпараттық ресурстар мен компьютерлік талдау құралдарын дамытуға бағытталған іргетасын қалады. Бұл заң жобалары 1988 жылы қабылданып, NCBI құрылды. Сонымен бірге, олар ғылыми кадрларды тарту және жобаның жүзеге асуына ықпал етті, осылайша NCBI ұлттық биотехнология саласында маңызды мекеме ретінде тез дами бастады



Daniel R. Masys

*ListerHill NationalCenter for
Biomedical Communications орталығында
жұмыс істеген ғылыми қызметкер.*



Dennis A. Benson

*Биотехнологиялық ақпарат жүйесін дамытуда алғашқы кезеңде белсенді болған ғалым. Суреттерде
олкомандаішінде, зертхана жағдайында*



Ғылыми Бағыт

NCBI-де компьютерлік ғалымдардан, молекулалық биологтардан, математиктерден, биохимиктерден, зерттеуші дәрігерлерден және есептеу молекулалық биологиясындағы іргелі және қолданбалы зерттеулерге шоғырланатын құрылымдық биологтардан тұратын көп салалы зерттеу тобы бар. Бұл зерттеушілер іргелі ғылымға маңызды үлес қосып қана қоймай, қолданбалы зерттеу қызметінің жаңа әдістерінің бастауы ретінде де қызмет етеді. Олар бірге математикалық және есептеу әдістерін қолдана отырып, молекулалық деңгейде іргелі биомедициналық мәселелерді зерттейді. Бұл мәселелерге гендерді ұйымдастыру, дәйектілікті талдау және құрылымды болжау кіреді.

Басты ғылыми бағыттары: Гендердің ұйымдасуын анықтау және талдау, қайталанатын реттілік үлгілерін, ақуыз домендерін және құрылымдық элементтерді, адам геномының гендік картасын құру, АИТВ-инфекциясының кинетикасын математикалық модельдеу, дерекқорды іздеу үшін секвенирлеу қателерінің әсерін талдау, дерекқорды іздеудің жаңа алгоритмдерін және ретсіздігінің реттілігін құру, дерекқорды теңестірудің жаңа алгоритмдерін әзірлеу, дерекқордың ретсіздігінің реттілігін құру.

Осы ғылыми бағыттардың барлығы NCBI-ді қазіргі заманғы биотехнология мен генетика саласындағы ең маңызды ақпараттық және зерттеу орталықтарының біріне айналдырды.



NCBI-Ұлттық биотехнологиялық ақпарат орталығы

Негізгі мақсаты:

- Биомедициналық және генетикалық ақпаратты жинау, сақтау және тарату
- Ғылыми зерттеулерді қолдау

Негізгі ғылыми бағыттары :

Геномика

Геномдық деректерді жинақтау (GenBank, RetSea) Адам және басқа организмдердің геномын талдау

Биоинформатика

Ақпараттық технологиялар арқылы биологиялық деректерді өңдеу. Алгоритімдер мен дерекқорларды жасау.

Молекулалық биология

Ген, құрылымы қызметін зерттеу. Транскрипция трансляция процестерін талдау.

Биомедицина және дәрілік зерттеулер

Аурулармен байланысты гендерді анықтау. Дәрілік заттарды жобалау.

Деректер базаларын дамыту

Генетикалық және медициналық ақпарат базаларын біріктіру

NCBI-дің негізгі мақсаты мен міндеттері

NCBI-дің мақсаты — биотехнология, молекулалық биология және биомедицина саласындағы ақпараттық ресурстарды дамыту және қамтамасыз ету арқылы ғылыми-зерттеу қызметіне қолдау көрсету. Ол ғалымдар мен медициналық қызметкерлерге геномдық, генетикалық және биомедициналық деректердің ашық, сенімді және кең қолжетімді базасын ұсынады. Бұл орталық жаңа биомедициналық ақпаратты жинау, сақтау, өңдеу мен таратуға бағытталған заманауи биоинформатикалық құралдар мен бағдарламаларды дамытады.

NCBI мәртебесі молекулалық зерттеулердің дамуына, медициналық диагностика мен емдеудің жетілуіне ықпал етеді, ғылыми білімді таратуда маңызды рөл атқарады және ғылыми қоғамдастықта байланыс орнатады.

NCBI міндеті — биомедициналық және генетикалық ақпаратты жинақтау, сақтау, талдау және ғалымдар мен қоғамға қолжетімді ету арқылы биотехнология мен медицинаның дамуына үлес қосу.



NCBI-дің Мақсаты

1. Биологиялық ақпаратты жинақтау және сақтау

– Әлемдегі барлық генетикалық және биомедициналық деректерді жүйелі түрде сақтау (мысалы, GenBank, PubMed).

2. Ғалымдарға ашық қолжетімділік беру

– Ғылыми қауымдастыққа сенімді және тегін ақпарат көзін ұсыну.

3. Биоинформатикалық құралдар мен ресурстар жасау

– Деректерді талдау, салыстыру және түсіндіру үшін бағдарламалар (мысалы, BLAST, Entrez) құрастыру.

4. Ғылыми зерттеулерді қолдау

– Геномика, молекулалық биология, медицина сияқты салаларда зерттеу жұмыстарын дамыту.

5. Биомедициналық білім мен ақпаратты тарату

– Зерттеушілермен студенттерге арналған оқу және ғылыми материалдар ұсыну.

6. Деректер базаларының өзара байланысын қамтамасыз ету

– Әртүрлі ақпараттық жүйелермен базалардың үйлесімді жұмыс істеуін ұйымдастыру.

Баршаға қолжетімді деректер
базасын әзірлеу



Бағдарламалық құралдарды
әзірлеу



Биомедициналық ақпаратты тарату



Міндеттері

1. Деректерді жинау және сақтау



Әлем бойынша жүргізілетін генетикалық және молекулалық зерттеулердің нәтижелерін қабылдау және GenBank, Gene, Protein сияқты базаларда сақтау.

3. Зерттеушілерді қолдау



Ғалымдарға деректермен жұмыс істеуге арналған бағдарламалық интерфейстер, іздеу және визуализация құралдарын ұсыну.

2. Ақпараттық талдау жүйесін дамыту



BLAST сияқты талдау құралдары арқылы генетикалық тізбектердің ұқсастығын анықтау және биологиялық мәліметтерді салыстыруға мүмкіндік беру.



NCBI өзінің міндеттерін орындау үшін мына қызметтерді атқарады:

- Молекулалық биология, биомедицина және генетика бойынша деректерді сақтау, өңдеу және талдау үшін автоматтандырылған жүйелерді құру.
- Биологиялық белсенді молекулалардың және заттардың құрылымы мен функциясы туралы зерттеу нәтижелерін компьютерлік түрде өңдеу.
- Биотехнология саласындағы зерттеушілер мен медициналық қызметкерлерге арналған мәліметтер базасы мен бағдарламалық қамтамасыз етуді дамыту және кеңінен қолдануды қолдау.
- Әлем бойынша биотехнологиялық ақпаратты жинақтау және үйлестіру жұмыстарын жүргізу.
- Жыл сайын молекулалық биология, генетика және биомедицинаға арналған семинарлар мен ғылыми конференциялар өткізу және білім беру қызметін, оның ішінде онлайн оқытуды ұйымдастыру.



NCBI құрамындағы негізгі деректер базалары

Ұлттық биотехнологиялық ақпарат орталығы (NCBI) көптеген биоинформатикалық деректер базаларын біріктіріп, ғалымдарға молекулалық биология мен генетика саласында кең көлемде ақпарат алуға мүмкіндік береді. Бұл деректер базалары өзара байланысты және бірыңғай жүйе арқылы жұмыс істейді.

NCBI базаларының көмегімен зерттеушілер генетикалық тізбектерді, ақуыз құрылымдарын, ғылыми мақалаларды және биомедициналық зерттеу нәтижелерін іздеп, талдай алады. NCBI-де көптеген базалар бар, бірақ ең маңызды және жиі қолданылатындары 5 негізгі база.

Олар:

- **GenBank** — ДНК және РНК тізбектері
- **PubMed** — ғылыми мақалалар мен әдебиеттер
- **BLAST** — тізбектерді салыстыру құралы
- **Gene және Protein** — гендер мен ақуыздар деректер туралы
- **Structure (MMDB)** — молекулалық модельдер



GenBank — ДНҚ және РНҚ нуклеотидтік тізбектерінің ұлттық мәліметтер қоры. Бұл база АҚШ Ұлттық Биотехнологиялық Ақпарат Орталығы ұсынады және онда әртүрлі организмдерден алынған генетикалық мәліметтер сақталады. GenBank ғалымдарға геномдық зерттеулер жүргізуге мүмкіндік береді әрі Еуропалық және Жапон әріптестерімен мәлімет алмасады.

PubMed — медицина мен биология саласындағы ғылыми мақалалардың, журналдардың мәліметтер қоры, NCBI арқылы қолжетімді. Бұл база зерттеушілерге зерттеу материалдары мен ғылыми жаңалықтарға қол жеткізуге көмектеседі.

BLAST (Basic Local Alignment Search Tool) — генетикалық тізбектерді салыстыруға арналған құрал. Бұл бағдарлама сұрау тізбегі бойынша GenBank және басқа дерекқорлардан ұқсас тізбектерді іздейді, бұл гендердің функциясын анықтауға және эволюциялық байланысты зерттеуге көмектеседі.

Gene — NCBI-дің гендер туралы толық мәліметтерді қамтитын дерекқоры, әрбір геннің орналасқан жері, құрылымы, функциясы және организмдегі рөлі туралы ақпарат береді. **Protein** — ақуыз құрылысына қатысты ақпараттық база, онда аминқышқыл тізбегі, құрылымы және ақуыздың биологиялық қызметі көрсетіледі.

Structure — биомолекулалардың, әсіресе ақуыздардың және нуклеин қышқылдарының үшөлшемді құрылымын сипаттайтын мәліметтер базасы, ол молекуланың формасын және өзара әрекеттесуін зерттеуге мүмкіндік береді.



GenBank — ДНҚ және РНҚ тізбектері сақталатын ең ірі халықаралық деректер қоры. Онда әр түрлі организмдердің геномдық ақпараттары жинақталған. GenBank-та сақталған деректер әлемнің барлық елдеріндегі зерттеушілерден қабылданып, жаңартылып отырады. Мысалы, өсімдік гендерінің, вирустар мен бактериялардың толық геномдық тізбектері осы базадан табуға болады.

PubMed — медициналық және биологиялық бағыттағы ғылыми мақалалар мен журналдарды қамтитын деректер базасы. Бұл жүйеде миллиондаған ғылыми еңбектер сақталады және оларды автор, тақырып немесе кілт сөз бойынша іздеуге болады.

BLAST — нуклеотидтік және ақуыздық тізбектердің ұқсастығын анықтауға арналған биоинформатикалық бағдарлама. Зерттеуші өзінде бар ДНҚ немесе ақуыз тізбегін енгізіп, оны әлемдік деректер базасындағы тізбектермен салыстыра алады. Мысалы, жаңа анықталған геннің қандай организмге ұқсас екенін немесе оның мүмкін функциясын BLAST арқылы анықтауға болады.

Gene — NCBI-дің Gene дерекқоры генге қатысты барлық ақпаратты біріктіріп, әрбір генге жеке жазба түрінде ұсынады.

Protein — NCBI-дің Protein дерекқоры ақуыздардың аминқышқылдық тізбегін, оның құрылымы мен функциясы туралы ақпаратты қамтиды.

MMDB — ақуыздар мен басқа да биомолекулалардың үшөлшемді құрылымдарын сақтайтын база. Ол Protein Data Bank (PDB) деректерін пайдаланып, молекулалардың кеңістіктік моделін визуалдауға мүмкіндік береді. Бұл база дәрілік препараттарды жобалау және молекулалық өзара әрекеттесуді зерттеу салаларында ерекше маңызға ие.



NCBI толық деректер базалары

1	Nucleotide (GenBank)	ДНҚ және РНҚ нуклеотидтік тізбектері сақталатын база
2	Genome	Әртүрлі организмдердің толық геномдық мәліметтері
3	Taxonomy	Организмдердің жіктелуі мен таксономиялық құрылымы
4	Structure (MMDB)	Ақуыздардың үшөлшемді (3D) құрылымдық модельдері
5	Domains (CDD)	Консервацияланған ақуыз домендері туралы мәліметтер
6	3D Domains	Ақуыздардың ықшам 3D домендері
7	OMIM	Адамның тұқым қуалайтын аурулары мен генетикалық бұзылыстар дерекқоры
8	SNP	Генетикалық полиморфизм мен мутациялар дерекқоры
9	UniSTS	Геномдағы маркерлер мен белгіленген аймақтар туралы мәлімет
10	GEO	Ген экспрессиясы мен микрочиптік талдау нәтижелері

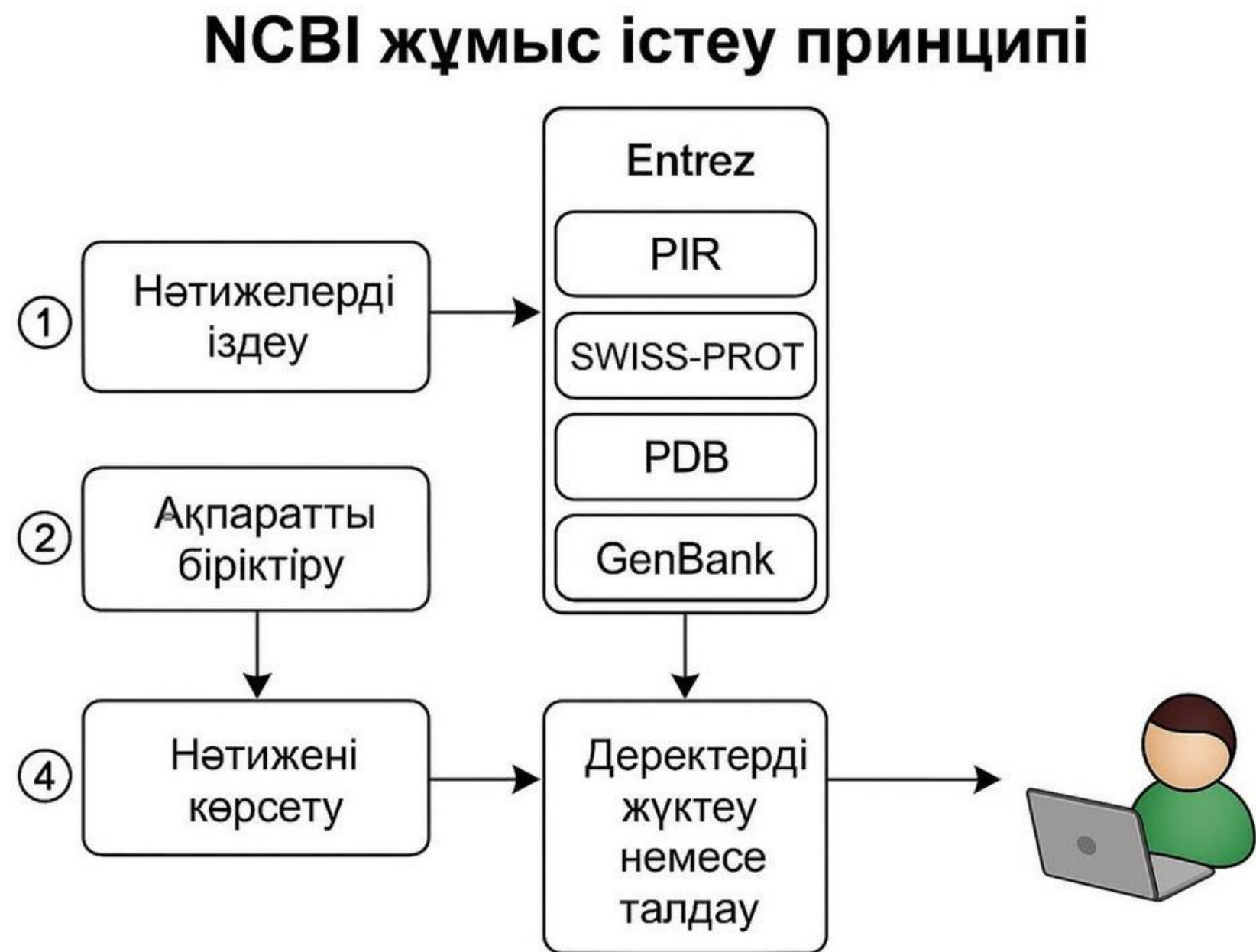
11	PopSet	Популяциялық зерттеулердің генетикалық деректері
12	UniGene	Ген экспрессиясына негізделген тізбектер топтамасы
13	HomoloGene	Эукариоттар арасындағы гомологты гендер топтамасы
14	Cancer Chromosomes	Ісіктік (қатерлі) аурулардағы хромосомалық өзгерістер дерекқоры
15	GENSAT	Тышқанның орталық жүйке жүйесіндегі ген экспрессиясы
16	Protein	Ақуыз тізбектері мен құрылымдары жинақталған база
17	PubMed	Биомедициналық ғылыми мақалалар мен әдебиеттер қоры
18	PubMed Central (PMC)	Толық мәтінді ашық ғылыми мақалалар базасы
19	Books	Онлайн оқулықтар мен ғылыми кітаптар жинағы
20	Entrez	Барлық деректер базаларын біріктіретін іздеу жүйесі

NCBI жұмыс істеу принципі

NCBI жұмыс істеу принципі – биологиялық ақпаратты жинау, сақтау және пайдаланушыларға Entrez жүйесі арқылы біріктірілген дерекқорлардан (мысалы, GenBank, PubMed, Protein, Genome және т.б.) іздеу мен талдау жасау мүмкіндігін беру.

NCBI қалай жұмыс істейді?

Ғалымдар өз мәліметтерін NCBI базасына жүктейді, ал Entrez жүйесі оларды бір жүйеге біріктіреді. Пайдаланушы іздеу арқылы қажет ақпаратты табады, нәтижелерді қарап, талдау немесе BLAST салыстыруын жасай алады. Деректер тұрақты түрде жаңартылып отырады.



(9-
сурет)

5. Биотехнология мен генетикада қолданылуы



Ғылыми деректер алмасуда

Әлем ғалымдары зерттеу нәтижелерін NCBI базаларына енгізіп, бір-бірімен ашық түрде бөліседі.



Биоинформатикада

BLAST сияқты құралдар арқылы ғалымдар геномдар мен ақуыз тізбектерін салыстырып, олардың ұқсастығын анықтайды.



Биотехнологиялық өнімдер жасауда

NCBI мәліметтері жан-дәрілер, вакциналар және гендік терапия түрлері әзірлеуде қолданылады.



Генетикалық зерттеулерде

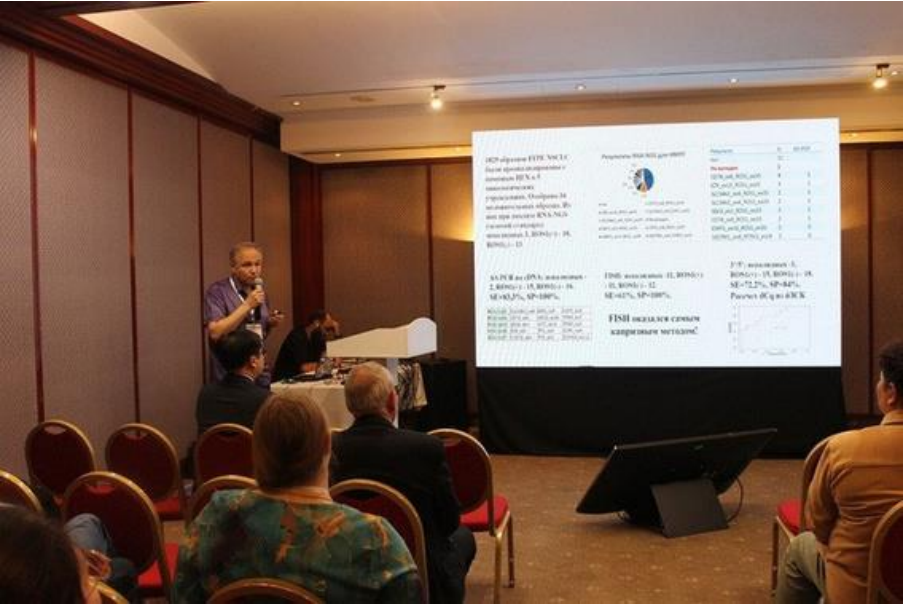
NCBI деректерінен тізбегін талдауға, мутацияларды анықтауға және тұқым қуалайтын ауруларды зерттеуге пайдаланылады.



NCBI кездесулерге, семинарларға және лекциялар топтамаларына демеушілік жасау арқылы молекулалық биология мен генетикаға қолданылатын компьютерлер саласындағы ғылыми байланысты дамытады. Сырттан оқитын ғалымдармен ынтымақтастықты дамыту үшін «Ғылыми келушілер» бағдарламасы құрылды. Докторантурадан кейінгі стипендиаттар NIH Intramural Research бағдарламасының бөлігі ретінде қол жетімді.



Суретте Ұлттық биотехнологиялық ақпарат ұйымдастырған BioEd Summit орталығы қатысушыларының топтық фотосы көрсетілген.



Суретте,, медициналық немесе ғылыми конференцияда өткен презентация көрсетілген. Слайдта FFPE NSCLC (өкпенің ұсақ жасушалы емес қатерлі ісігі) үлгілерін талдау нәтижелері берілген.



Бақылау сұрақтары:

1. NCBI қандай мақсатта құрылған?
2. BLAST қызметі не үшін қолданылады?
3. GenBank базасында қандай деректер сақталады?
4. NCBI-дің негізгі ақпараттық модульдерін атаңыз.
5. NCBI деректерін қолданудың ғылыми маңызы неде?

Қолданылған әдебиеттер

1. NCBI Resources — *National Center for Biotechnology Information, U.S. National Library of Medicine* (<https://www.ncbi.nlm.nih.gov>).
2. Lesk A.M. *Introduction to Bioinformatics* — Oxford University Press, 2023.
3. Конничев А.С. *Биоинформатика: основы анализа геномных данных* — Москва: Академия, 2021.
4. Әбілаев С.А. *Молекулалық биология және генетика* — Шымкент: Асқаралы, 2008.
5. Zhang, J. et al. *Data Mining in Genomics and Proteomics* — Springer, 2020.