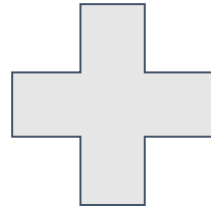




ԿԵՆՏԱՔԻՆՖՈՐՄԱՏԻԿԱ

Ավանեսյան Գևորգ
Կուրդիկյան Միեր
Հարությունյան Անի
Սոհրաբյան Նաիրի

Տվյալների հավաքագրման սկիզբը



Ինչ է Կենսաաինֆորմատիկան

- Բիոինֆորմատիկան ըստ Նիկոլաս Լասկըմբի՝

Կենսաբանություն + Ինֆորմատիկա = Կենսաինֆորմատիկա

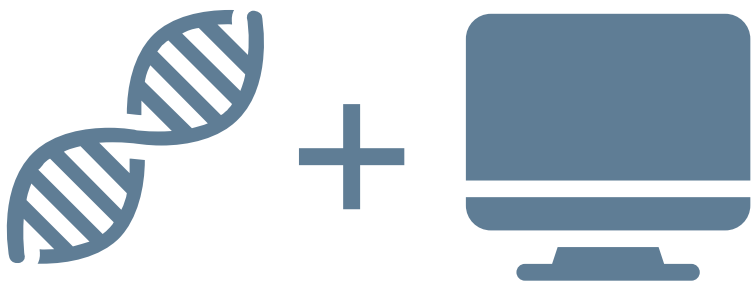


Սակայն...

- Համակարգչային կենսաբանություն՝ ավելի լայն հասկացություն է



- Կենսաինֆորմատիկա (համակարգչային մոլեկուլային կենսաբանություն) սահմանափակվում է գեների, գենոմների և սպիտակուցների հաջորդականության, կառուցվածքի և գործունեության վերլուծությամբ:



ԴՆԹ-ի հաջորդականացում (սեկվենավորում)

ԴՆԹ-ի սեկվենավորումը ԴՆԹ-ի նմուշում նուկլեոտիդային հաջորդականության որոշումն է:

Կիրառություններ

1. Մոլեկուլային կենսաբանություն
2. Բժշկություն
3. Դատաբժշկական փորձաքննություն
4. Վիրուսաբանություն
5. Էվոլյուցիոն կենսաբանություն



1-ին սերնդի (Սամգերի) սեկվենավորում

Դիդեօքսինուկլեոտիդներ՝ դդՆԵՖ-ներ,
ԴևԹ Պոլիմերազ, դՆԵՖ-ներ

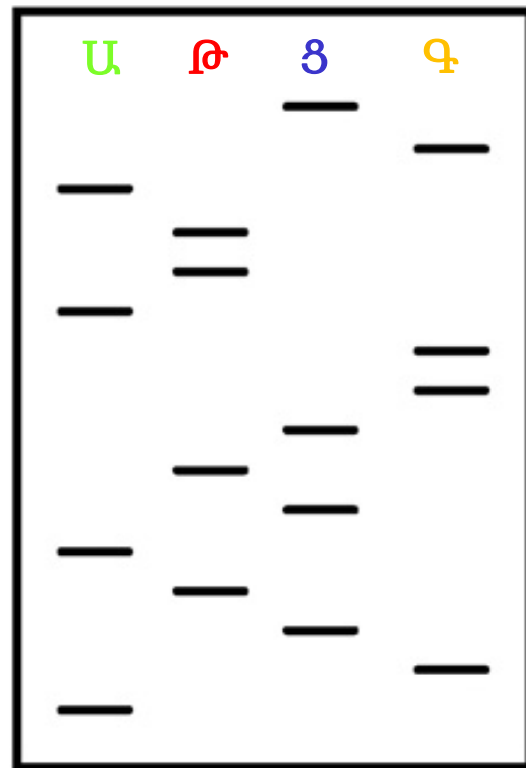
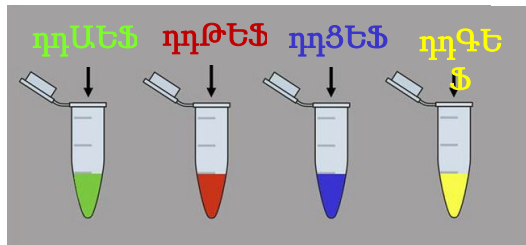
a)

ATGCAGCGTTACCATG...

b)

U

8

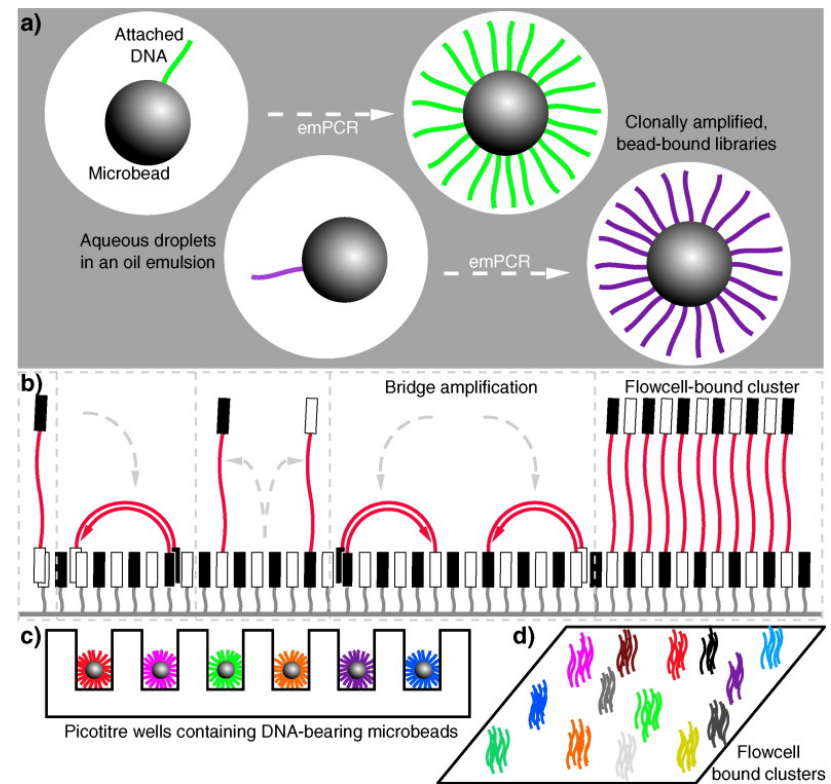
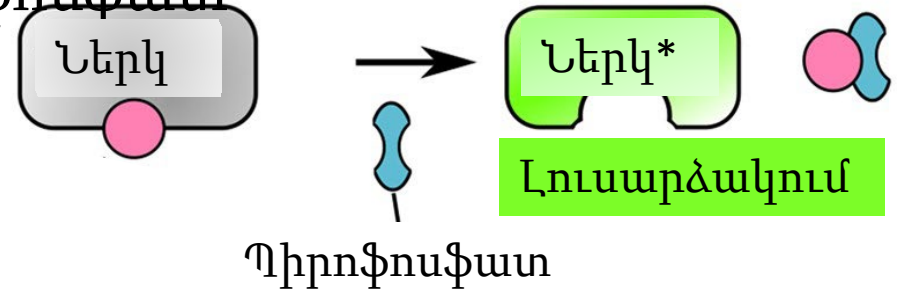
 Q_f 

ԷԼԵԿՏՐԱՖՈՐԵԳԻ
ԱՐԴՅՈՒՆՔՆԵՐԸ

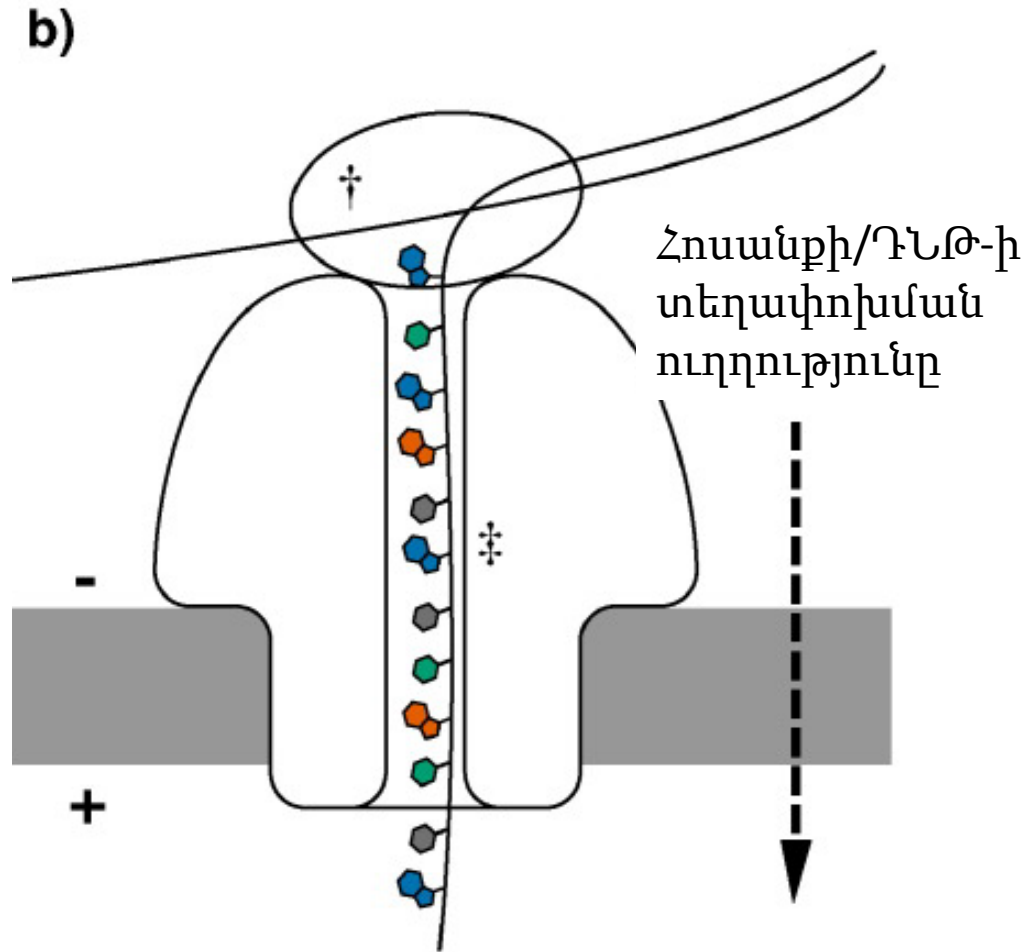
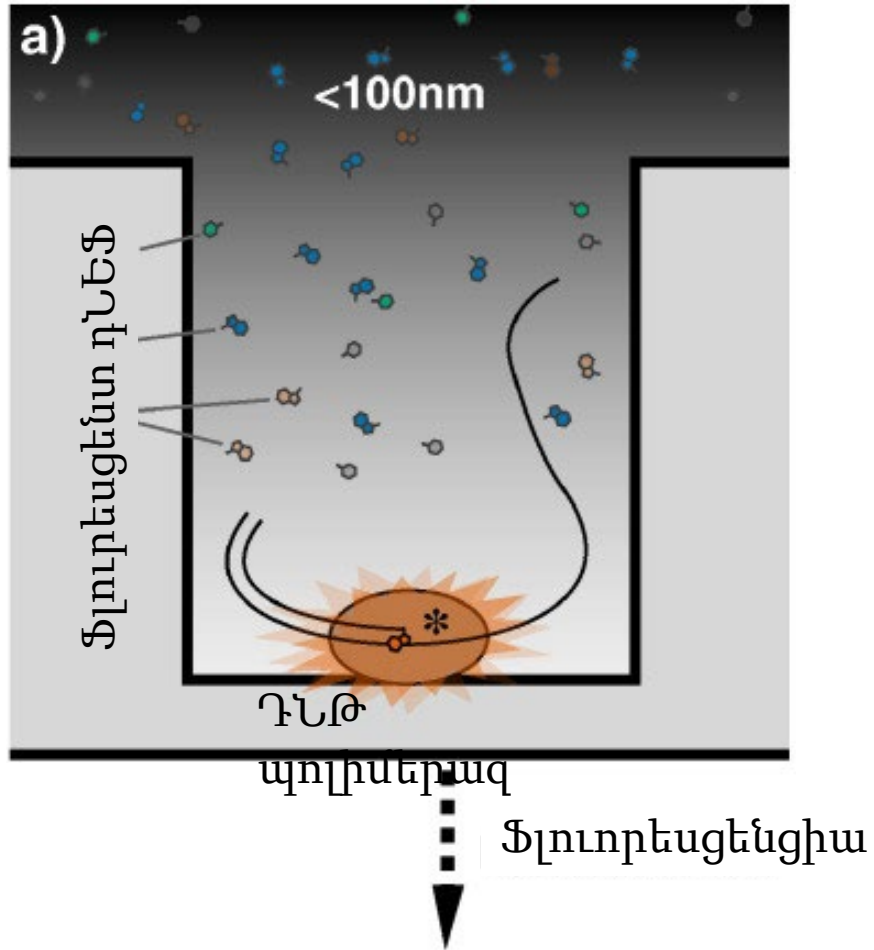
2-րդ մերմդի սեկվենավորում



Պիրոֆոսֆատ



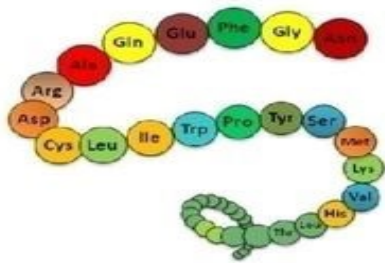
Տ-իմ սպինային սեկվենավորում



- Մեկ մոլեկուլի սեկվենավորում է
- Հիմնված չէ ԴՆԹ պոլիսերագի աշխատանքի վրա (b-ն)
- Գործում է ավելի մեծ հաջորդականությունների համար

Կենսաաինֆորմատիկական վերլուծություններ Կառուցվածքի վերլուծություն

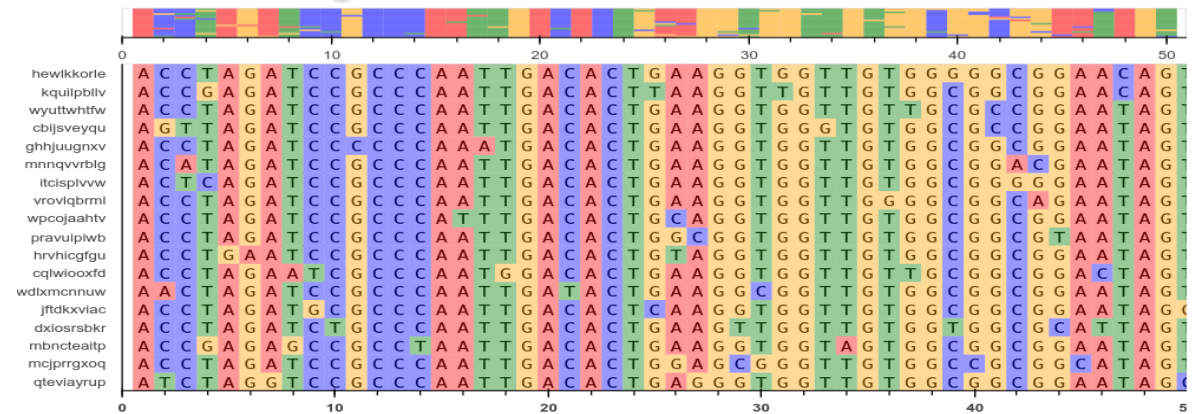
- Նուկլեինաթթվի կառուցվածքի կանխատեսում
(Nucleic acid structure prediction)



- Սպիտակուցի կառուցվածքի կանխատեսում
(Protein structure prediction)
- Սպիտակուցի կառուցվածքի դասակարգում
(Protein structure classification)
- Սպիտակուցի կառուցվածքի համեմատում
(Protein structure comparison)

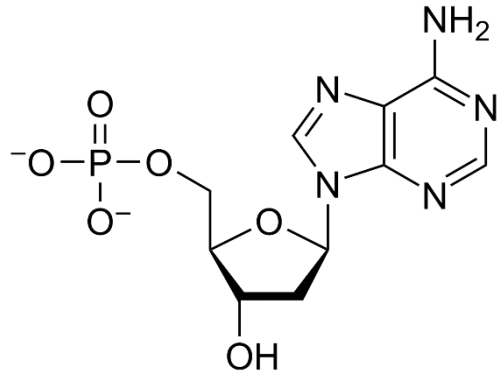
Կենսաախինֆորմատիկական վերլուծություններ Հաջորդականության վերլուծություն

- Գենոմի համեմատում
(Genome comparison)
- Մոտիվի հայտնաբերում
(Motif discovery)
- Հաջորդականության հավասարեցում
(Sequence alignment)
- Հաջորդականության տվյալների բազայի որոնում
(Sequence database searching)
- Գենի և պրոմոտերի կանխատեսում
(Gene and promoter prediction)



- Ֆիլոգենիա
(Phylogeny)

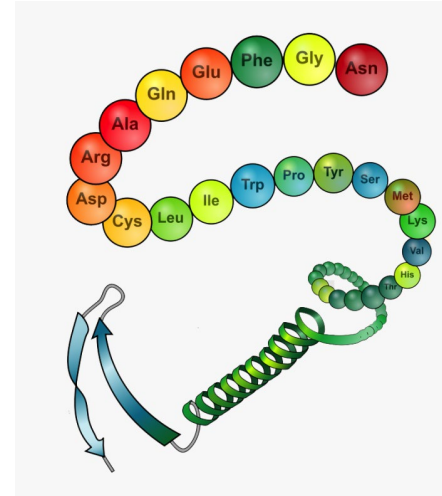
Հաջորդականությունների նմանություն և նույնություն



նմանություն = նույնություն

Նմանությունը վերաբերում է միևնույն դիրքում եղած այն ամինաթթուներին, որոնք ունեն ֆիզիկաքիմիական նմանատիպ հատկություններ և սպիտակուցի կազմում կարող են փոխարինվել մեկը մյուսով:

Նույնությունը վերաբերում է տրված հաջորդականությունների միջև միևնույն ամինաթթուների համընկնմանը:



նմանություն !=
նույնություն

Հաջորդականությունների Նմանությունը%

Միմյանց նման մոնոմերներ

$$S = [(L_s \times 2) / (L_a + L_b)] \times 100$$

Ուսումնասիրվող հաջորդականություններ

$$I = [(L_i \times 2) / (L_a + L_b)] \times 100$$

Հաջորդականությունների Նույնություն%

Միևնույն մոնոմերներ



Հարաբերվող տվյալների բազաներ

- **Տվյալների բազա**-տվյալների հավաքագրման և կարգավորման համար նախատեսված համակարգչային արխիվ

Աղյուսակ Ա		
Ուսանող #	Անուն	Ծննդավայր
1.	Անի	Հրազդան
2.	Գևորգ	Մարտունի
3.	Մհեր	Աբովյան
4.	Նաիրի	Թեհրան

Աղյուսակ Բ	
Ուսանող #	Տարիքը
1.	19
2.	17
3.	17
4.	17

Աղյուսակ Գ	
Ծննդավայր	Մարզ (Շրջան)
Հրազդան	Կոտայք
Մարտունի	Մարտունի
Աբովյան	Կոտայք
Թեհրան	Թեհրան

Հավասարեցում

Տեսակները

- Գլոբալ
- Տեղային

Ձևերը

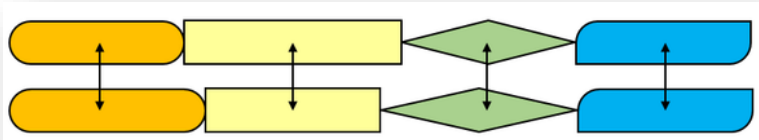
- Զույգերով
- Բազմակի

Հիստոն 1 (120-180 դիրքեր)

Մարդ
Շիմպանզե
Մուկ
Առնետ
Կով

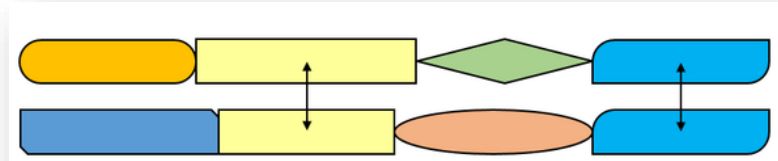
KKASKPKKAASKAPT	KKPKATPVKKAKKKLAATPKKAKKPKTVKAKPVKASKPKKAKPVK
KKASKPKKAASKAPT	KKPKATPVKKAKKKLAATPKKAKKPKTVKAKPVKASKPKKAKPVK
KKAAKPKKAASKAPS	KKPKATPVKKAKKKRAATPKKAKKPKVVKVPVKASKPKKAKTVK
KKAAKPKKAASKAPS	KKPKATPVKKAKKKRAATPKKAKKPKIVKVPVKASKPKKAKPVK
KKAAKPKKAASKAPS	KKPKATPVKKAKKKRAATPKKTKKPKTVKAKPVKASKPKKTKPVK
***	*****:***** ** .*****:* **

Գլոբալ հավասարեցում



- Համեմտվող հաջորդականությունների երկարությունների նմանություն:
- Դասավորումը իրականացվում է երկու հաջորդականությունների սկզբի ծայրից մինչև վերջ:
- Օգտագործվում է կարգաբանորեն միմյանց մոտիկ խմբերի առանձնյակների միջև:
- Օգտագործվում է միմյանցից քիչ տարբերվող հաջորդականությունների համեմատման մեջ:
- Օգտագործվում են պոլիմորֆիզմների հայտնաբերման մեջ:

Տեղային հավասարեցում



- Համեմատվող հաջորդականությունները կարող են ունենալ միմյանցից զգալիորեն տարբերվող երկարություններ:
- Համեմատում է հաջորդականությունները ոչ ամբողջությամբ, այլ որոշակի հատվածներով:
- Օգտագործվում է կարգաբանորեն միմյանցից հեռու եղած խմբերի առանձնյակների միջև:
- Լայնորեն օգտագործվում է առանձնյակների իՌՆԹ-ների, ինտրոնների, էքսոնների համեմատման մեջ:

5' ACTACTAGATTACTTACGGATCAGGTACTTTAGAGGCTTGCAACCA 3'

||||| ||||| ||||| |||||

5' ACTACTAGATT---ACGGATC--GTACTTTAGAGGCTAGCAACCA 3'

5' ACTACTAGATTACTTACGGATCAGGTACTTTAGAGGCTTGCAACCA 3'

|||| ||||| ||||| |||||

5' TACTCACGGATGAGGTACTTTAGAGGC 3'

Հաջորդականությունների զույգերով հավասարեցում

- Երկու հաջորդականությունների հավասարեցում, որը հաջորդականությունների նմանությունները գտնելու և բազմակի հավասարեցման հիմքն է:

EEELTKPRLWALYFNMRDALSSG

VEKPRILYALYFNMRDSSDE



EEELTKPRILWALYFNMRDALSSG—
---VEKPRILYALYFNMRD---SSDE

Հաջորդականությունների զույգերով հավասարեցում

Տեսակները

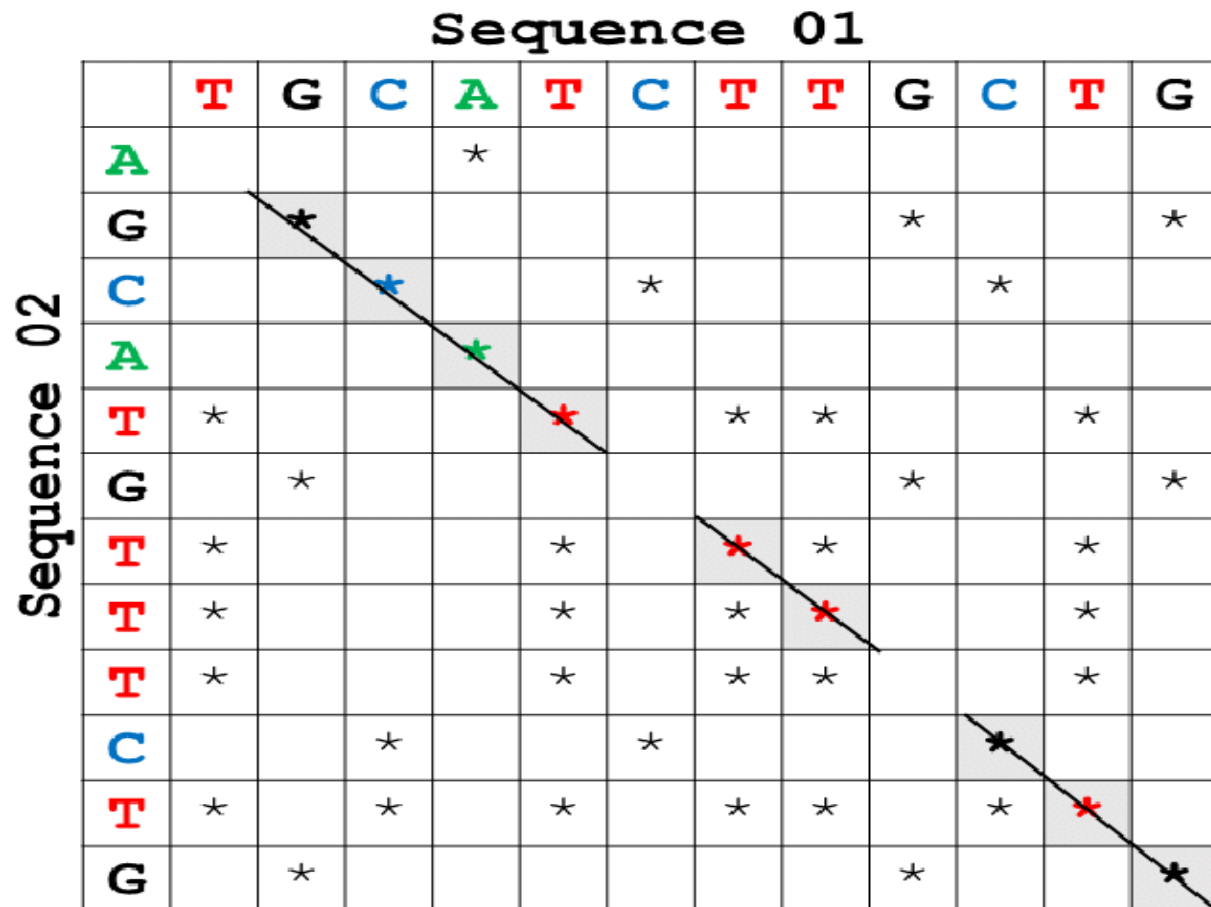
- Կետային մատրիցայի (գրաֆիկի) մեթոդ
- Դինամիկ ծրագրավորման մեթոդ
- Բառային մեթոդ

Q	K	G	S	Y	P	V	R	S	T	C
Q	K	G	S	G	P	V	R	S	T	C
1	2	3	4	5	6	7	8	9	10	11

Կետային մատրիցայի (գրաֆիկի) մեթոդ

(Dot matrix (plot) method)

- 2 ամինաթթվային կամ նուկլեոտիդային հաջորդականություններ համեմատում



- 1970, Գիբս և Մակինտայր (Gibbs and McIntyre)

1 Հաջորդականություններից յուրաքանչ-յուրը գրվում է մատրիցայի առանցքներին (հորիզոնական և ուղղահիգ)

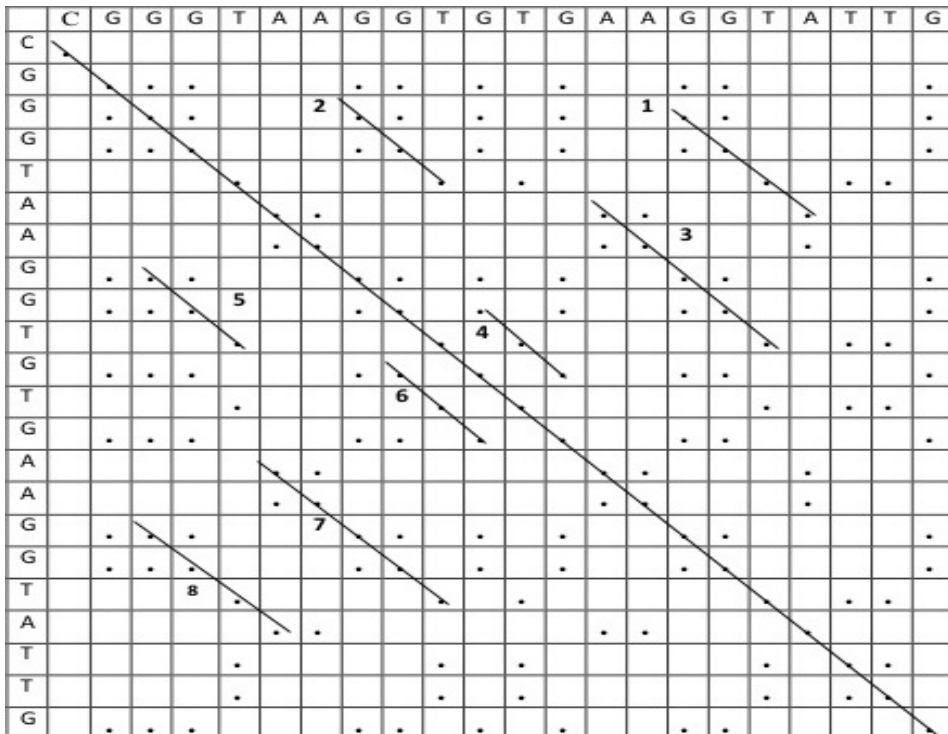
2 Տառերի համընկման դեպքում կետ է դրվում

3 Հաջորդական կետերը միացնում են

Կետային մատրիցայի (գրաֆիկի) մեթոդ

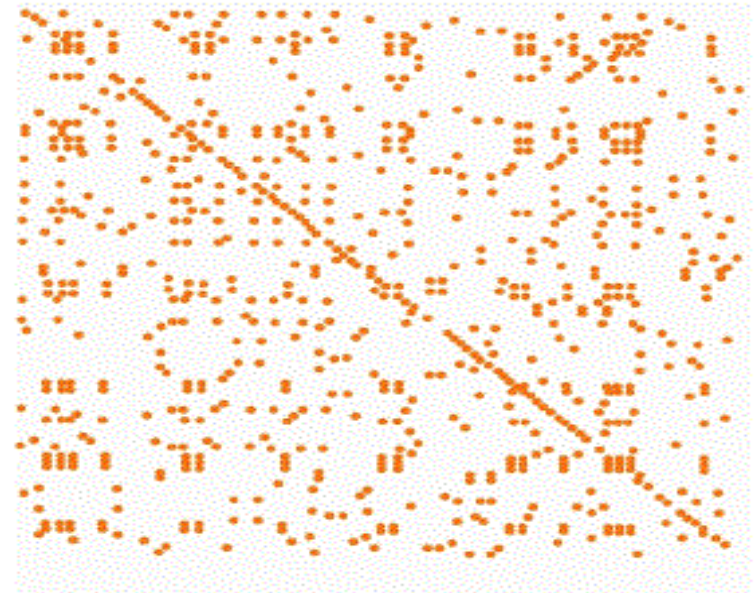
(Dot matrix (plot) method)

- Հաջորդական կետերի միացման անկյունագիծը ցույց է տալիս այդ հաջորդականությունների մնացորների նմանությունը և ենթադրում է հոմոլոգություն դրանց միջև, բացատները ցույց են տալիս ինսերցիաներ և դելեցիաներ, իսկ անկյունագծին զուգահեռ գծերը ցույց են տալիս կրկնվող հատվածները



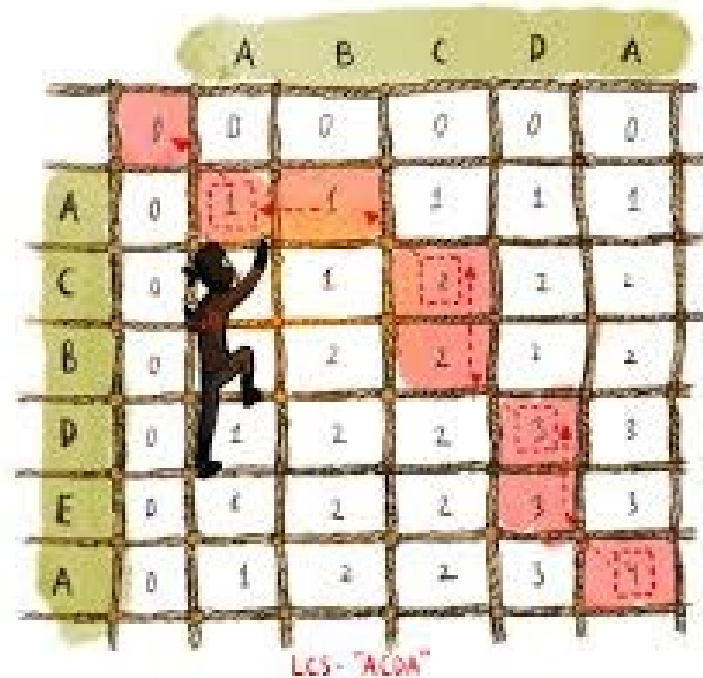
- Օգտագործում են զտող «պատուհան»
(Filtering window)

- Ինքնահամեմատում
(Self-comparison)



Դինամիկ ծրագրավորման մեթոդ

- Նման է կետային մատրիցի մեթոդին
 - Հավասարեցման երկչափ մեթոդ
- Որակական ճշգրիտ վերլուծություն
 - Միավորներ կետերի փոխարեն
- Աշխատում է դանդաղ



Բառային մեթոդ (word method)

- Էվրիստիկական մեթոդ է
- Օգտագործում է հաջորդականությունների կարճ հատվածներ՝ «բառեր» հաջորդականությունները համեմատելու համար:
- Ծրագրերից են՝
 1. **BLAST** (**B**asic **L**ocal **A**lignment **S**earch **T**ool՝ Տեղային հավասարեցման որոնման հիմնական գործիք)
 2. **FASTA** (FAST ALL՝ բոլորը արագ)

BLAST

- Սա հաջորդականությունների վերլուծության ամենահայտնի ծրագրերից է, ստեղծվել է 1990թ. Ստեֆեն Ալտշուլի կողմից:
- Նուկլեոտիդային «բառերը» 11 տառանի են, ամինաթթվայինները՝ 3
- Վիճակագրական նշանակալիությունը՝
 - E-արժեք
 - բիթ միավոր S' (bit score)
- Տարբերակներ՝
BLASTN, BLASTP, BLASTX, TBLASTN, TBLASTX.

Հարցադրված հաջորդականությունը: MRD**PYN**KLIS

1

Հարցադրում	PYN	PYN	PYN	PYN	...
Շտեմարան	PYN	PFN	PFQ	PFE	...

2

Հարցադրում	PYN	PYN	PYN	PYN	...
Շտեմարան	PYN	PFN	PFQ	PFE	...
Միավորներ ի գումարը	20	16	10	10	...

3

Հարցադրում	M R D	PYN	K L I S
Շտեմարան	M H E	PYN	D V P W

← →
 Դեպի ձախ Դեպի աջ

4

Հարցադրում	M R D	PYN	K L I S	
Շտեմարան	M H E	PYN	D V P W	
	5 0 2	20	-1 1 -3	-3

Ընդհանուր միավորը՝
24

Բարձր միավորով հատվածի գույգ (high scoring segment pair՝ **HSP**)

BLAST Արտաքին տեսքը

Հարցադրվող հաջորդականությունը ներմուծելու տեղը

14:43blast.ncbi.nlm.nih.gov

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s)ClearQuery subrangeFromTo

ATTTAGTCAAATAAAATTTGAGAGAATCTTCTACATAAGTAGCTCACAGACCCAACCAAAGGACC
ATTGAATACCATATATATAGATACACAGACATATAAACACACAAATATTCGTGTTTTTCAAACTGT
GAGAGAAAAAGAGAGAGAAAGAGATGGGAGAGATTGGGTTTACAGAGAAGCAAGAAGC

Or, upload fileВыбрать файлфайл не выбран

Job TitleNC_003074.8:c3277930-3276163 Arabidopsis thaliana...
Enter a descriptive title for your BLAST search

☐ Align two or more sequences

BLAST results will be displayed in a new format by default

You can always switch back to the Traditional Results page.

Choose Search Set

DatabaseStandard databases (nr etc.):
Nucleotide collection (nr/nt)

OrganismOptional
Enter organism name or id—completions will be suggested
ExcludeOptional
Models (XM/XP)
Limit toOptional
Sequences from type material
Entrez QueryOptional
Enter an Entrez query to limit search

Program Selection

Optimize for
Highly similar sequences (megablast)
More dissimilar sequences (discontiguous megablast)
Somewhat similar sequences (blastn)
Choose a BLAST algorithm

BLAST

Search database Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences)
☐ Show results in a new window

Հարցադրման արդյունքները

14:44blast.ncbi.nlm.nih.gov

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignmentsDownloadManage ColumnsShow100

☒ select all52 sequences selected

GenBankGraphicsDistance tree of results

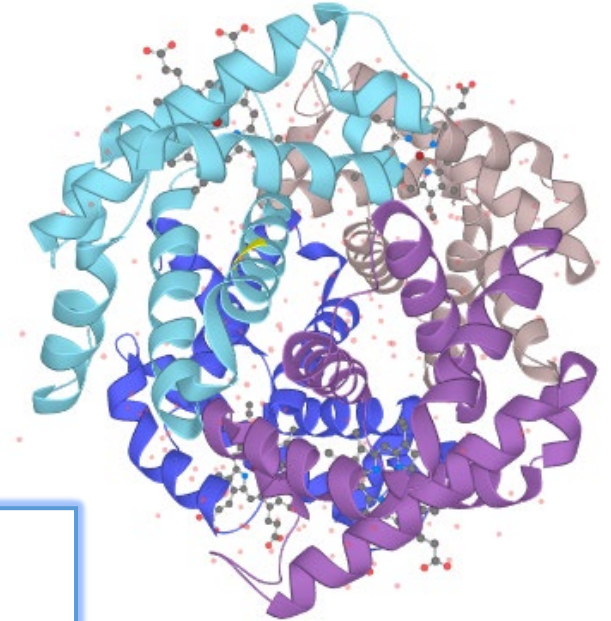
	Description	Max Score	Total Score	Query Cove	E value	Per. Ident	Accession
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3265	3265	100%	0.0	100.00%	LR782544.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3265	3265	100%	0.0	100.00%	LR699772.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3265	3265	100%	0.0	100.00%	LR215054.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3265	3265	100%	0.0	100.00%	LR797809.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3265	3265	100%	0.0	100.00%	LR797794.1
☒	Arabidopsis thaliana chromosome 3 sequence	3265	3265	100%	0.0	100.00%	CP002686.1
☒	Arabidopsis thaliana chromosome III BAC F18K10 genomic sequence, complete sequence	3265	3265	100%	0.0	100.00%	AC013428.6
☒	Arabidopsis thaliana chromosome 3 BAC F13M14 genomic sequence, complete sequence	3265	3265	100%	0.0	100.00%	AC011560.6
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3260	3260	100%	0.0	99.94%	LR699757.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3238	3238	100%	0.0	99.72%	LR797799.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3232	3232	100%	0.0	99.66%	LR699767.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3225	3225	100%	0.0	99.55%	LR699747.2
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3223	3223	100%	0.0	99.55%	LR699752.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	3206	3206	100%	0.0	99.38%	LR797804.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	2878	2958	94%	0.0	98.35%	LR699762.1
☒	Arabidopsis thaliana genome assembly, chromosome: 3	2878	2958	94%	0.0	98.35%	LR797789.1
☒	Arabidopsis thaliana class 2 non-symbiotic hemoglobin (AHB2) gene, complete cds	2464	2464	76%	0.0	99.70%	U94999.1
☒	Arabidopsis thaliana hemoglobin 2 (HB2), mRNA	917	1852	56%	0.0	100.00%	

Feedback

Հաջորդականությունների ներկայացումը FASTA ֆորմատով

- FASTA-ն տեքստի վրա հիմնված ձևաչափ է:
- FASTA-ի հիմքն է ծառայում կետային մատրիքսը:
- Օգտագործվում է նուկլեոտիդային հաջորդականությունները և ամինաթթվային հաջորդականությունները ներկայացնելու համար:
- Նուկլեոտիդները կամ ամինաթթուները ներկայացված են մեկ տառանի կոդերի միջոցով:

Հեմոգլոբինի ալֆա շղթա



ԴՆԹ –ն ներկայացնող «բառերը» 6 նուկլեոտիդ երկարություն ունեն
Սպիտակուց ներկայացնող «բառերը» 2 ամինաթթու երկարություն ունեն



Իսկ ինչպե՞ս է կարելի մեզ հետաքրքրող հաջորդականությունները համեմատել՝ օգտագործելով մասնագիտացված ծրագրերից

Name/Gene ID	Description	Location	Aliases	MIM
☐ HBA1 ID: 3039	hemoglobin subunit alpha 1 [<i>Homo sapiens</i> (human)]	Chromosome 16, NC_000016.10 (176680..177522)	ECYT7, HBA-T3, HBH, METHBA 	141800

[Hba-a1 – hemoglobin alpha, adult chain 1](#)

[Mus musculus \(house mouse\)](#)

Also known as: Hba, Hba1, Hbat1



[hbaa1 – hemoglobin, alpha adult 1](#)

[Danio rerio \(zebrafish\)](#)

Also known as: AA1, dZ118J2.2, hbaa1.1



NC_000016.10:17..	1	-----ACTCTTCTGGT----CCCCACA-GACTCAGAGAGAAC-CCACCATGGTGCTGTCT	49
NC_000077.6:322..	1	----GACACTTCTGAT----TCTGACA-GACTCAGGAAGAA----ACCATGGTGCTCTCT	47
NC_007114.7:551..	1	AACAGCTTTTCTGAAGAAACCTCACTTGAGAAAAAGAAAACGCAGCGATGAGTCTCTCT	60
		. : *****: * **: ** :*.***.*** . * **.	
NC_000016.10:17..	50	CCTGCCGACAAGACCAACGTCAAGGCCGCTGGGGTAAGGTCGGCGCGCACGCTGGCGAG	109
NC_000077.6:322..	48	GGGGAAGACAAAAGCAACATCAAGGCTGCCTGGGGGAAGATTGGTGCCATGGTGCTGAA	107
NC_007114.7:551..	61	GATACGGACAAGGCTGTTGTTAAGGCCATCTGGGCTAAGATCAGCCCCAAGGCCGATGAA	120
		.. *****. :. * ***** . ***** ***. * . * * * * **.	
NC_000016.10:17..	110	TATGGTGCGGAGGCCCTGGAGAGGTGAGGCTCCCTCCCTGCT-----CCGA-----	156
NC_000077.6:322..	108	TATGGAGCTGAAGCCCTGGAAAGGTGAGAACAGG-ACCTTGAT-----CTGT-----	153
NC_007114.7:551..	121	ATTGGTGCTGAAGCCCTCGCCAGGTAAAGCTCACC--TATGTTTTTTTTTCTATTTGAA	178
		: : ***: ** **.****** * . *****.*... . ** * * * . :	
NC_000016.10:17..	157	-C-----CCGGGCTCCTCGCCCGCCCGGACC-----CAC-----AGGCC	189
NC_000077.6:322..	154	-AAGGATCACAGGATCCAATATGGACCTGGCACTCGCTCAG-----TGGGC	198
NC_007114.7:551..	179	TTAGGATTACTG-----ATTGACAGAACTTTATAGAAGTTTATTATTATTATTAAGA	230
		. * * . * . * . . * . * .	
NC_000016.10:17..	190	ACCCTCAA-CCGTCCTG-GCCCCGG-----ACCCAAACCCCAACCCCTCACTCTGCTTCTC	242
NC_000077.6:322..	199	AGCTTCTAACTATGCTT-TTCTGTG-----ACCTCAAC-----TTCTCTTCTCTCCTT	245
NC_007114.7:551..	231	ATAATCTATGATAACTTATAATGTGTAAAAACATAAA--TAACCTATGT--ATTTTTTC	285
		* . **:* : ** . * **.*** * : * *	
NC_000016.10:17..	243	CCCGCAGGATGTTCTGTCCTTCCCCACCACCAAGACCTACTTCCCGCACTTCG---ACC	299
NC_000077.6:322..	246	CTCCCAGGATGTTTGCTAGCTTCCCCACCACCAAGACCTACTTCCCTCACTTTG---ATG	302
NC_007114.7:551..	286	CTCCCAGAATGCTGACCGTCTACCCTCAAACCAAGACCTACTTTTCTCACTGGGCTGACT	345
		* * **.**** * **:**** ..***** * **** * *	



Ծառերի կառուցում

NC_007114.7:551001...
NC_000016.10:17668...
NC_000077.6:322836...

Danio rerio
Homo sapiens
Mus musculus

(Ժամանակ)	Running time	15.9 seconds
(Միասնական դիրքեր)	Identical positions	345
(Նույնություն)	Identity	36.938%
(Նման դիրքեր)	Similar positions	139

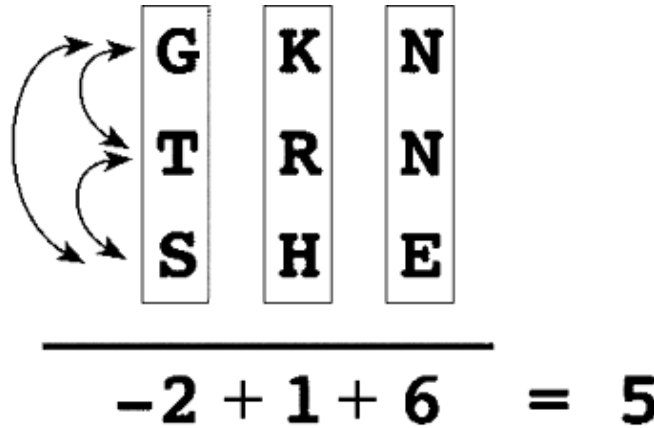
Հաջորդականությունների Բազմակի Հավասարեցում (Multiple Sequence Alignment)

Հաջորդականություն 1

Հաջորդականություն 2

Հաջորդականություն 3

Ձույգերի միավորների
գումարը



- Ճշգրիտ (exhaustive) ալգորիթմներ
- Էվրիստիկական ալգորիթմներ
 - Պրոգրեսիվ հավասարեցման մեթոդ
 - Կրկնվող հավասարեցում
 - Բլոկերի վրա հիմնված հավասարեցում

```

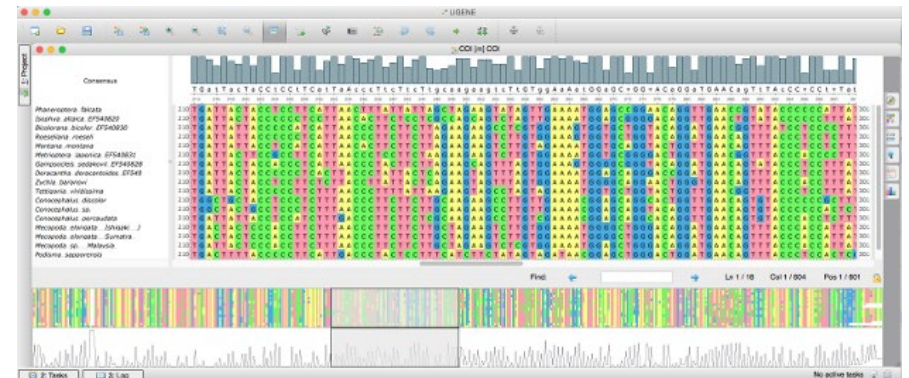
Target      VTNSP-VVVA LDYHNRDDAL AFVDKI-DPR DCRLKVGKEM FTLFGPQFVR
3LDV A      AMNDPKVIVA LDYDNLADAL AFVDKI-DPS TCRLKVGKEM FTLFGPDEFVR
3TR2 A      ---DPKVIVA IDAGTVEQAR AQINPL-TPE LCHLKIGSIL FTRYGPAFVE
3TFX A      ---DRPVIVA LDLDNEEQLN KILSKLGDPH DVFVKVGXEL FYNAGIDVIK

Target      ELQQRGFDFI LDLKFHDIPN TAAHAAAAA DLGVWMNVNH ASGGARMMTA
3LDV A      ELHKRGFSVF LDLKFHDIPN TCSKAVKAAA ELGVWMNVNH ASGGERMMAA
3TR2 A      ELXQKGYRIF LDLKFYDIPQ TVAGACRAVA ELGVWXXNIH ISGGRTXXET
3TFX A      KLTQQGYKIF LDLKXHDIPN TVYNGAKALA KLGITFTTVH ALGGSQXIKS

Target      AREALVPFG-- -KDAPLLIAV TVLTSMEASD LVD-LGMTLS PADYAERLA
3LDV A      SREILEPYG-- -KERPLLIGV TVLTSMESAD LQG-IGILSA PQDHVLRRLA
3TR2 A      VVNAIQSITL- -KEKPLLIGV TILTSLDGSD LKT-LGIQEK VPDIVCRXA
3TFX A      AKDGLIAGTPA GHSVPKLLAV TELTSSIDDV LRNEQNCRLP XAEQVLSLA

Target      ALTQKCGLDG VVCSAQEAVERF KQVFGQEFKL VTPGIRPQGS EAGDQRRIM
3LDV A      TLTKNAGLDG VVCSAQEASLL KQHLGREFKL VTPGIRPAGS EQGDQRRIM
3TR2 A      TLAKSAGLDG VVCSAQEAALL RKQFDRNFLL VTPGIR-----RVX
3TFX A      KXAKHSGADG VICSPLEVKKL HENIGDDFLY VTPGIRP-----A

Target      TPEQALSAGV DYMVIGRPVITQ SVDPAQTLKA INASLQ-----
3LDV A      TPAQAIASGS DYLVIGRPITQ AAHPEVVLEE INSSL-----
3TR2 A      TPRAAIQAGS DYLVIGRPITQ STDPLKALEA IDKDI-----
3TFX A      TPKXAKEWGS SAIVVGRPITL ASDPKAAAEA IKKEFNAENLYFQS
    
```



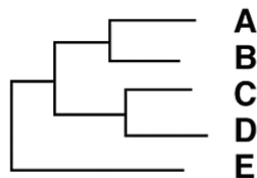
Պրոգրեսիվ հավասարեցման մեթոդ

A _____
B _____
C _____
D _____
E _____

Առանձին զույգերով
հավասարեցում և
հեռավորությունների մատրիքի
կառուցում

	A	B	C	D	E
A	—				
B	11	—			
C	20	30	—		
D	27	36	9	—	
E	30	33	20	27	—

Ուղորդով ծառի կառուցում



C/D և A/B զույգերի առանձին
հավասարեցում

C _____
D _____

C/D և A/B զույգերի առանձին
հավասարեցում

A _____
B _____

C/D և A/B 2 համաձայնեցված
(consensus) հաջորդականությունների
առաջացում և դրանց հավասարեցում

C/D ~~~~~
A/B ~~~~~

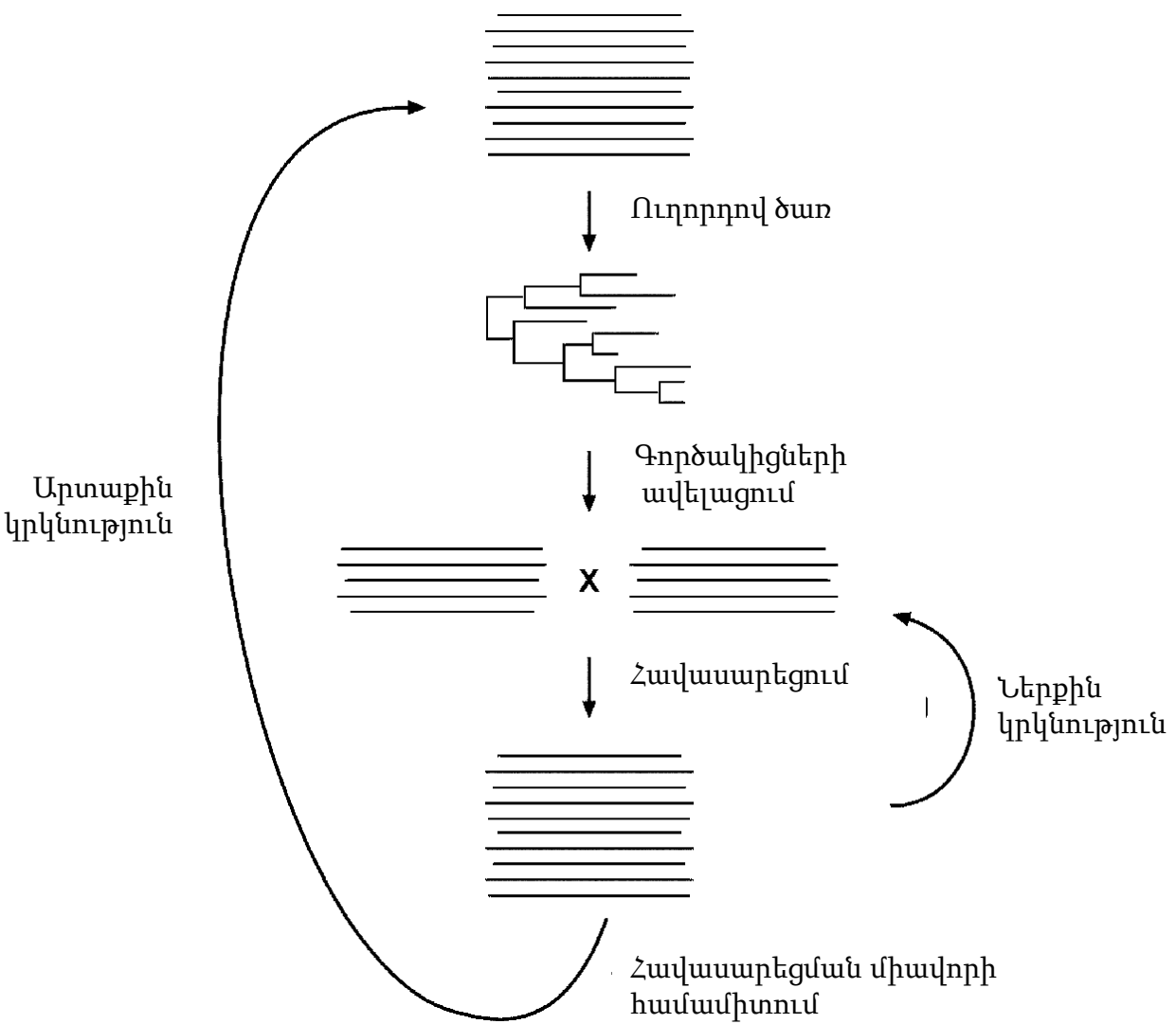
C/D/A/B համաձայնեցված (consensus)
հաջորդականության ստեղծում և դրա
հավասարեցում E հաջորդականության
հետ

A/B/C/D ~~~~~
E _____

Հավասարեցման ավարտ

A _____
B _____
C _____
D _____
E _____

Կրկնվող հավասարեցում



Բլոկերի վրա հիմնված հավասարեցում

Հիմնված է հաջորդականությունների նման հատվածների (բլոկների) հավասարեցման վրա

(A)

WA	XA1	XA2	YA	ZA
WB	ZB	YB	XB	
ZC1	WC	ZC2	XC	

Species A
Species B
Species C

(B)

W	X	Y	Z
A	A1	A	A
B	A2	B	B
C	B		C1
	C		C2

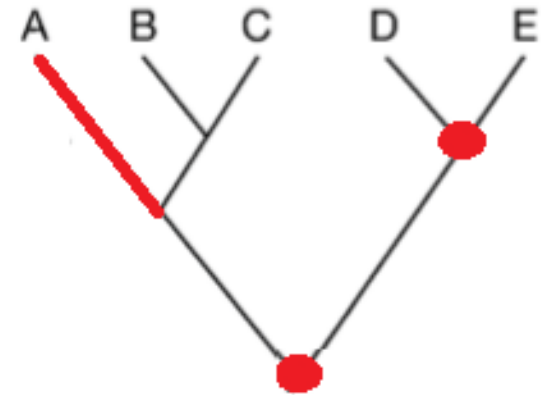
(C)

A	TT-CTAAGTG
B	CTACTAAG-G
C1	CTACT--GTG
C2	CTACC--GTG

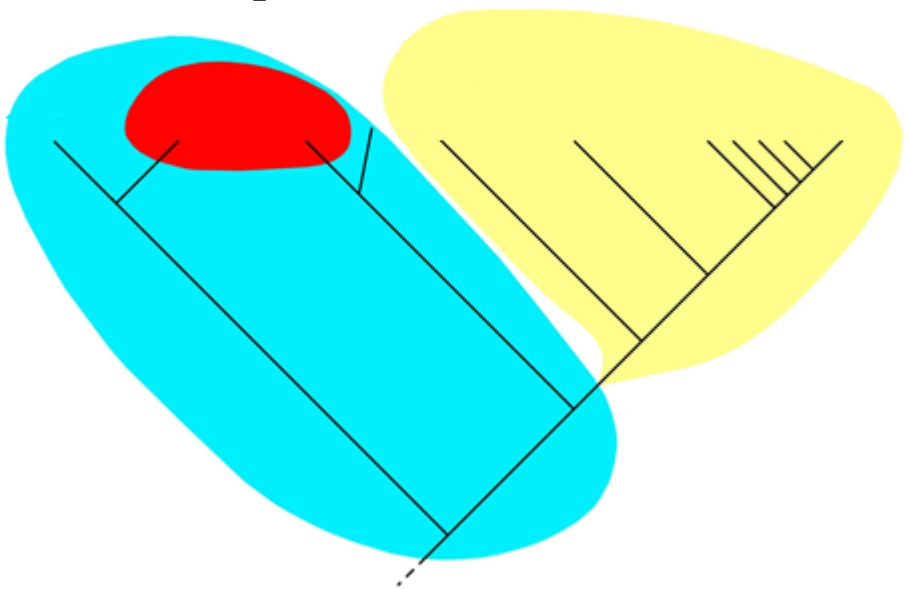
Ֆիլոգենետիկ ծառեր

Ֆիլոգենիան թույլ է տալիս տարբեր կարգաբանական խմբերի ուսումնասիրությանը և փոփոխված հատկանիշների պատկերավոր դիտարկմանը:

Ֆիլոգենետիկ ծառը ներկայացնում է օրգանիզմների միջև էվոլյուցիոն հարաբերություններ:

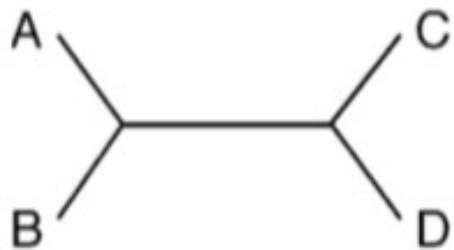


- մոնոֆիլետիկ
- պարաֆիլետի
- քլոիֆիլետիկ



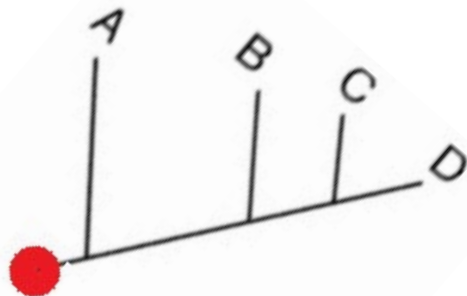
Ֆիլոգենետիկ ծառի կայուն հիմք են հանիսանում **մուլեկուլային տվյալները**

ԴՆԹ, ՌՆԹ, սպիտակուցներ



Անարմատ ծառ

Ցույց չի տալիս
ընդհանուր նախնուն



Արմատավոր ծառ

Ցույց է տալիս
ընդհանուր նախնուն



$$N_R = (2n - 3)! / 2^{n-2} (n - 2)!$$

$$= (2 * 4 - 3)! / 2^{4-2} (4 - 2)! =$$

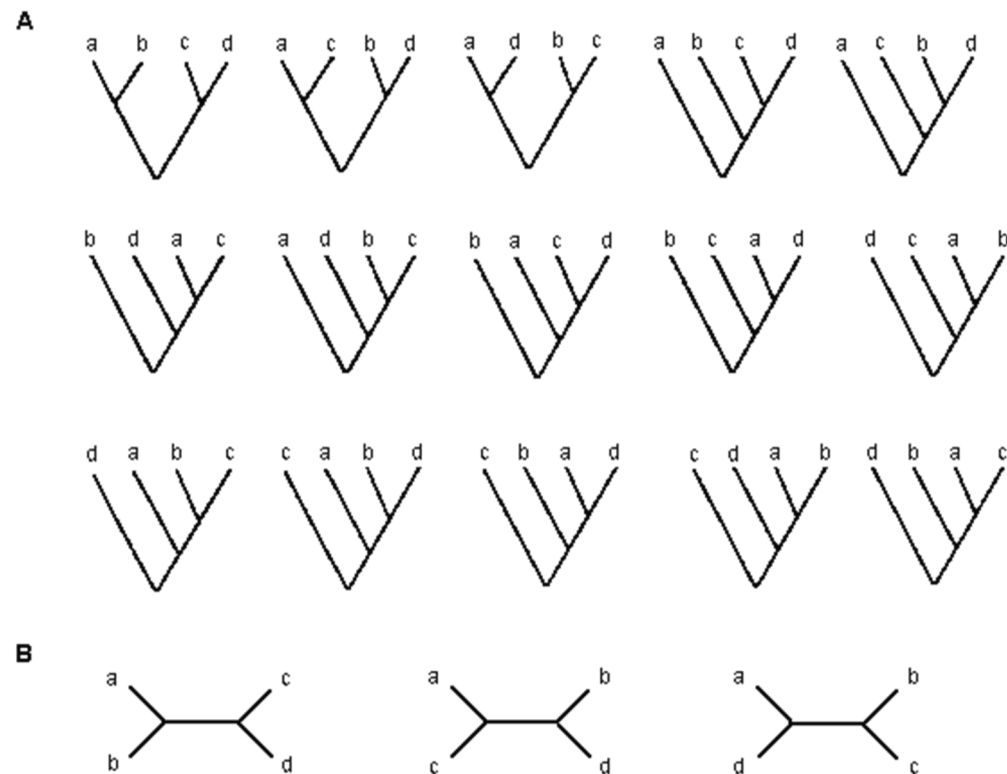
$$= 120 / (4 * 2) = 15$$

$$N_U = (2n - 5)! / 2^{n-3} (n - 3)!$$

$$= (2 * 4 - 5)! / 2^{4-3} (4 - 3)! =$$

$$= 6 / 2 = 3$$

Էվոլուցիոն ծառեր



Ֆիլոգենետիկ անալիզի մեթոդներ

A ATCGTGGTACTG
B CCGGAGAACTAG
C AACGTGCTACTG
D ATGGTGAAAGTG
E CCGGAAAACCTG
F TGGCCCTGTATC

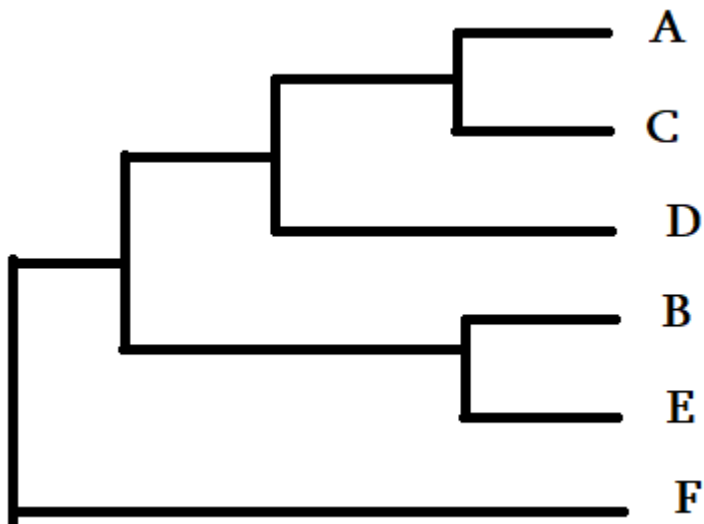
մոլեկուլային
հաջորդականությունները
գարգանում են որոշակի
արագությամբ, այնպես որ
կուտակված մուտացիաների
քանակը համընկնում է
էվոլյուցիոն ժամանակին:

1

	A	B	C	D	E	F
A		9	2	4	2	10
B			9	6	2	10
C				5	9	10
D					6	10
E						10
F						

2

	A/C	B	D	E	F
A/C		9	4.5	9	10
B			6	2	10
D				6	10
E					10
F					



1 UPGMA Method
2 Neighbor Joining Method
(էվոլյուցիոն արագության ճշտությամբ)

3

	A/C	B/E	D	F
A/C		9	4.5	10
B/E			6	10
D				10

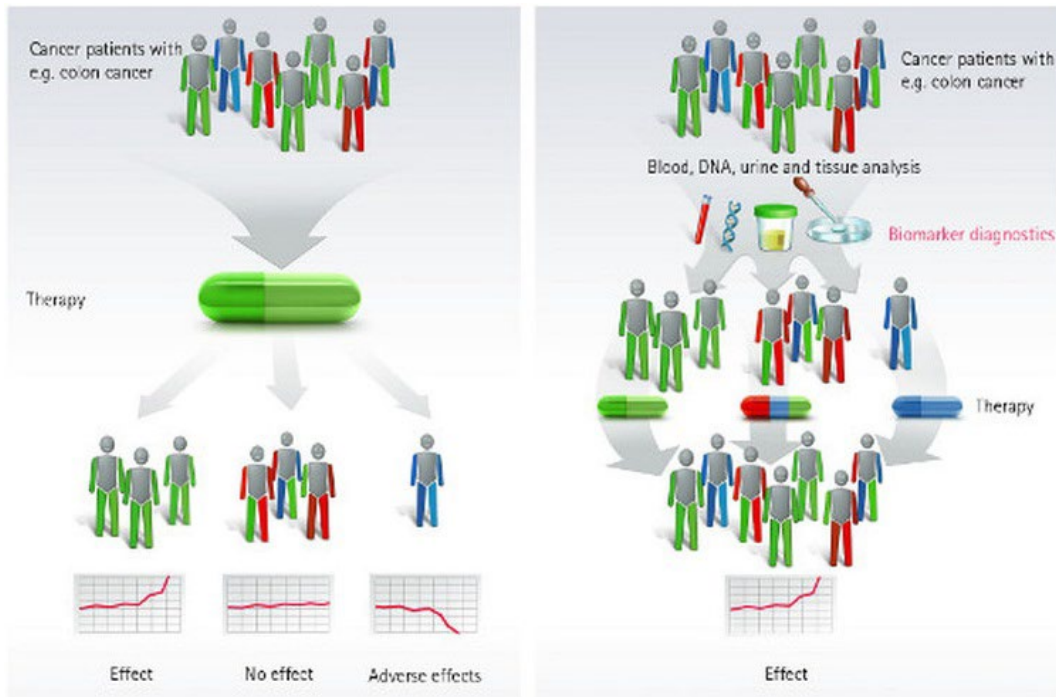
4

	A/C/D	B/E	F
A/C/D		7.5	10
B/E			10

Կենսաախինֆորմատիկայի կիրառությունները

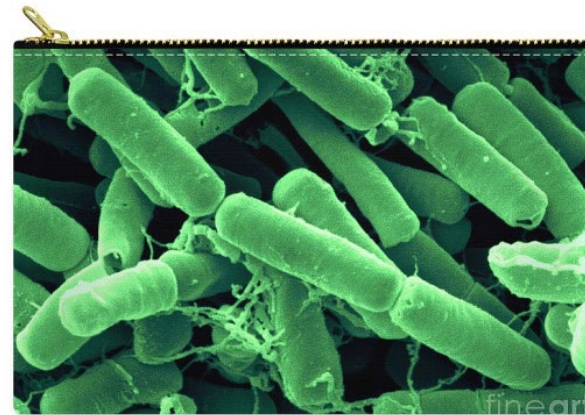
Բժշկություն

1. Դեղամիջոցների հայտնաբերում
2. Գենային թերապիա
3. Անհատական բժշկություն



Գյուղատնտեսություն

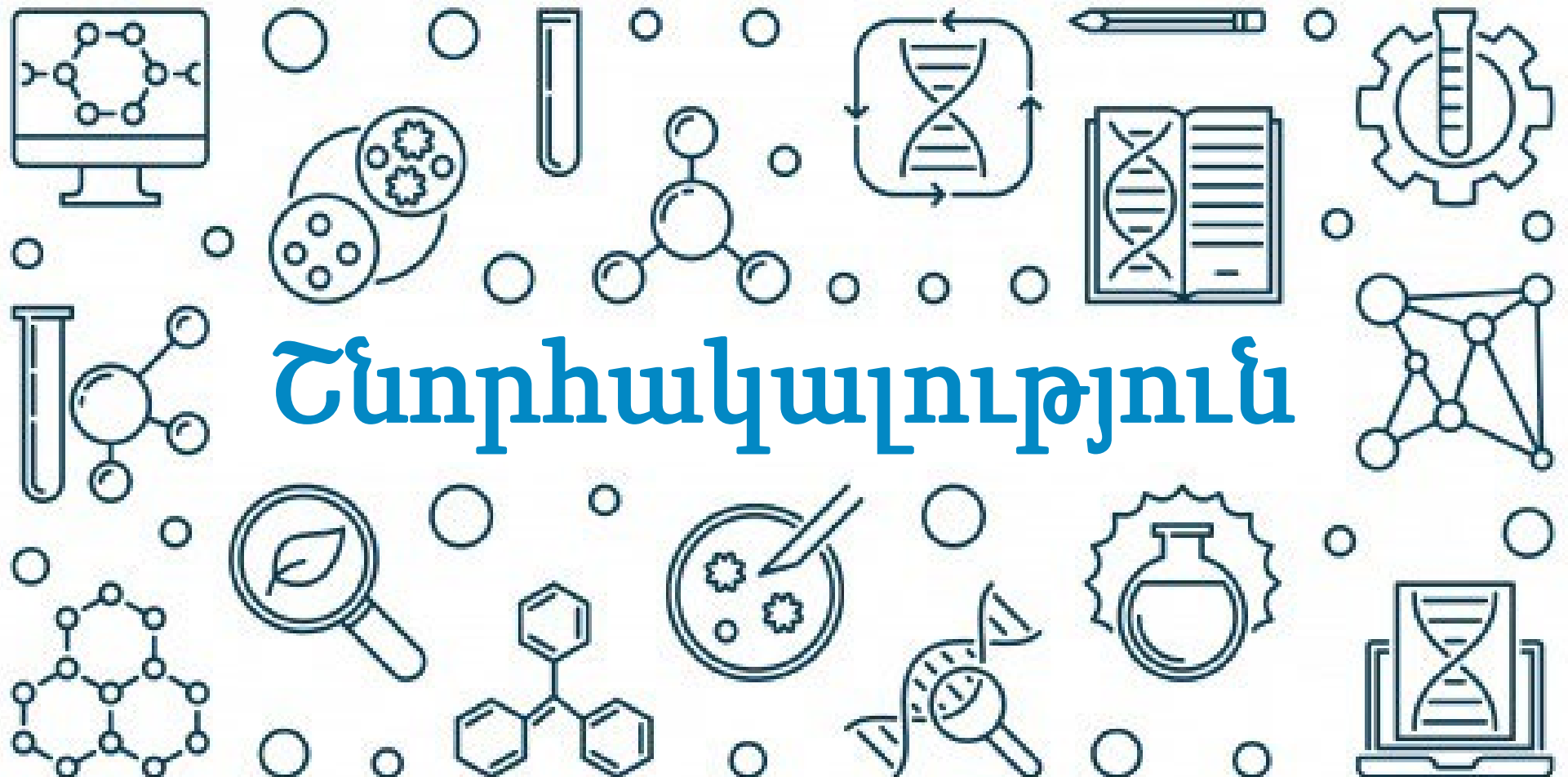
1. Բերքի բարելավում
2. Միջատների հանդեպ դիմադրողականություն
3. Սննդամթերքի որակի բարելավում



Bacillus thuringiensis



Oryza sativa



Հնորհակալություն