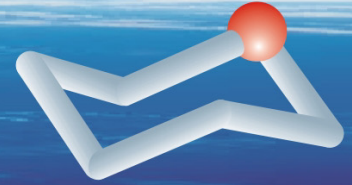


Bevezetés a bioinformatikába

Harangi János
DE, TEK, TTK Biokémiai
Tanszék



Bioinformatika



Interdiszciplináris tudomány, amely magába foglalja a biológiai adatok gyűjtésének, feldolgozásának, tárolásának, elemzésének, értelmezésének összes vonatkozását.

A számítástechnika alkalmazása biomolekulák vizsgálatából származó információk megszerezésére és értelmezésére.

Kialakulása:

**Kísérleti technikák fejlődése
(szekvenálás, szerkezet meghatározás)**

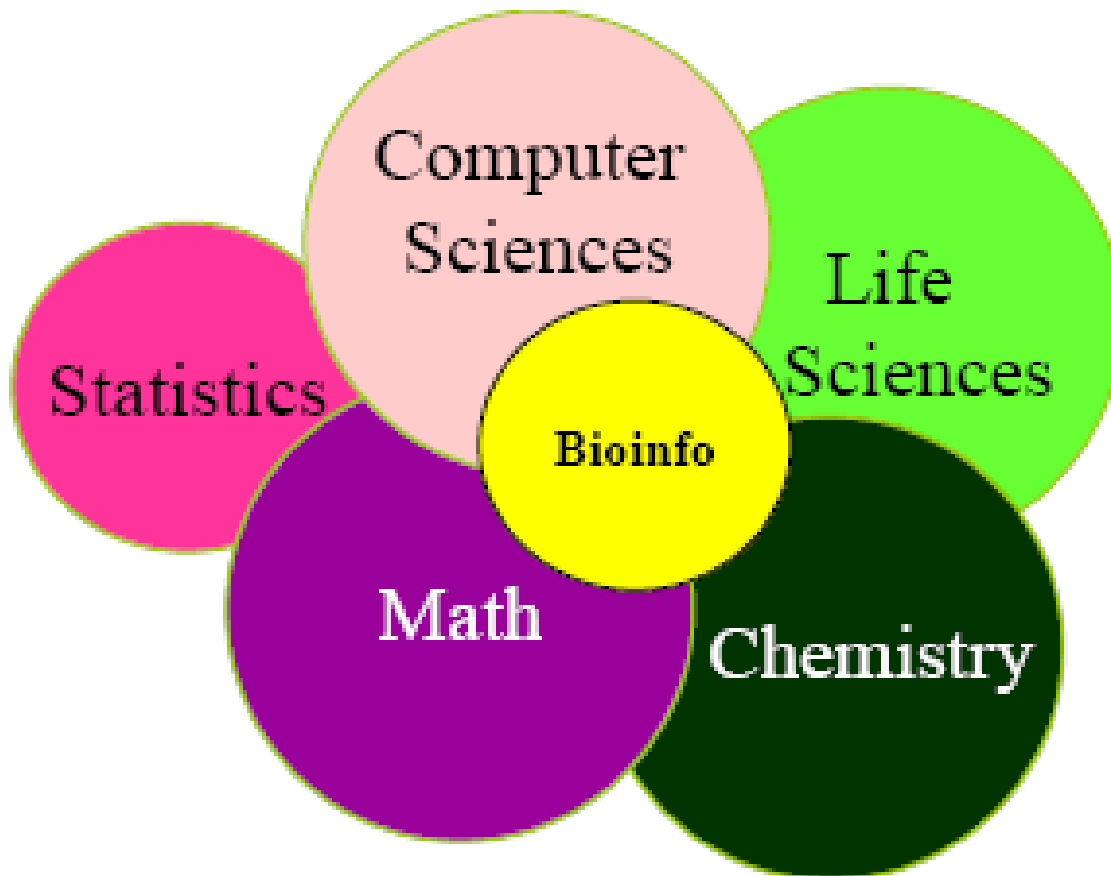
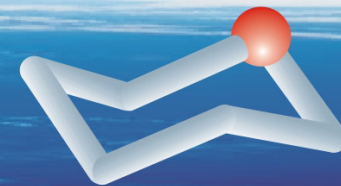


**„Nagy léptékű” biológiai kutatás
(genom projektek, DNS chippek)**

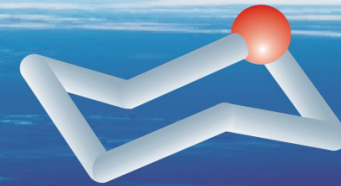


Nagy mennyiségű, komplex biológiai adat

A bioinformatika helye



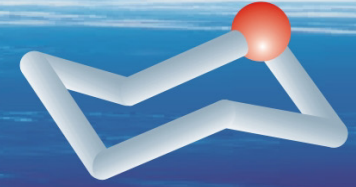
Bioinformatika, ahogy legtöbben használják



- PubMed keresés és eredményei
 - Szerző, publikáció helye, ...
- Információ egy gén szekvenciáról
- Információ egy fehérjéről



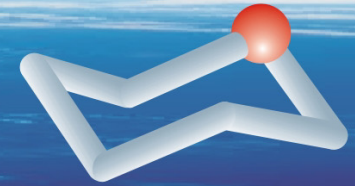
A bioinformatika fő területei



- DNS szekvenciaelemzés és adatbázisok – filogenetikus fa (genomika)
- Fehérje szekvenciaelemzés és adatbázisok (proteomika)
- Fehérjék 3D szerkezete
- Enzimműködés modellezése
- RNS szekvenciák és RNS működés
- Glikoproteinek és poliszacharidok szerkezete (glikomika)
- Annotáció
- Összefüggések keresése



Pubmed: human genome resources



NCBI Human Genome Resources - Microsoft Internet Explorer

Fájl Szerkesztés Nézet Kedvencek Eszközök Súgó

Vissza Keresés Kedvencek

Cím <http://www.ncbi.nlm.nih.gov/genome/guide/human/> Ugrás Hivatkozások

Google G Go Bookmarks 12 blocked Check AutoLink Settings

mywebsearch Search Screensavers Smiley Central Cursor Mania Fun Cards

NCBI Genomic Biology Homo sapiens

Search All Databases (Entrez) for Go Clear

Browse your Genome
Click on the Chromosome to show
Genes

1 2 3 4 5 6 7 8
9 10 11 12 13 14 15 16
17 18 19 20 21 22 X Y

Find A Gene
Search for

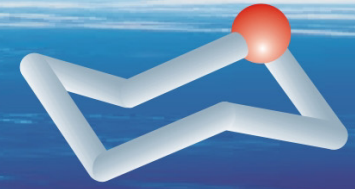
Human Genome Resources

A challenge facing researchers today is that of piecing together and analyzing the plethora of data currently being generated through the Human Genome Project and scores of smaller projects. NCBI's Web site serves an integrated, one-stop, genomic information infrastructure for biomedical researchers from around the world so that they may use these data in their research efforts. [More...](#)

Genes and Human Health

Internet

Bioinformatikai alkalmazások



- Adatgyűjtés és feldolgozás
- Szekvencia illesztés
- Molekuláris evolúció
- Genom térképezés és összehasonlítás

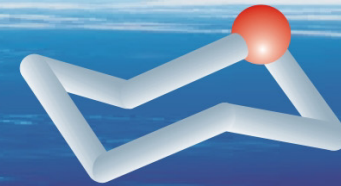
- **Szerkezet meghatározás**
- **Szerkezet összehasonlítás**
- **Molekuláris modellezés és szimuláció**
- **Szerkeztjóslás**
- **Gyógyszertervezés**



Másodlagos és harmadlagos szerkezet jóslása

- 3D szerkezeteket illesztő algoritmusok
- Fehérjegeometriai mérések
- Felszín, térfogat és alak számítása
- Intermolekuláris kölcsönhatások
- Molekulaszimulációk (molekuláris mozgások, dokkolás)

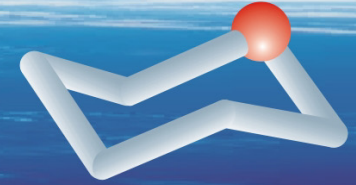
Fehérje szerkezet elemzés és azonosítás



- Fehérje enzimatis vagy kémiai hasítása
- Peptid mapping
 - Peptidek szekvenciájának elemzése
- Fragmensek összeillesztése informatikai eszközökkel
- Keresés adatbázisokban
- Adatbázisok információtartalma!!!



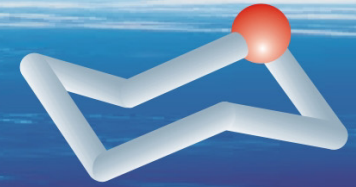
Enzim adatbázisok



- PDB: fehérje adatbázis
- BLAST: fehérje homológia keresés
- IUBMB: EC szám alapján
- Brenda: katalizált reakció, szerkezet, ligandok, forrás,
- ExPASy
- MACiE: aktív hely adatbázis
- CSA: Catalytic Site Atlas
- IntEnz: katalizált reakció, szerkezet
- EzCatDB: mechanizmusok, kevesebb kémiai részlet, publikációk.



Bioinformatika alkalmazási folyamat



Szerkezet meghatározás: NMR, röntgenkristallográfia

Szerkezet-jóslás

- A funkció vagy a térszerkezet megjóslása pl. hasonlóság alapján
- Ez alapján új szekvenciák funkciójának/szerkezetének azonosítására
- A már ismert, hasonló funkciójú/szerkezetű fehérjéket megvizsgálva
valamely, a funkcióra/szerkezetre jellemző, konzerválódott sajátosság keresése

Módosított enzimek tervezése

- Gyógyszertervezés** - szerkezet-hatás kapcsolat jóslása
- célmolekulák tervezése



Fehérje adatbázis



RCSB Protein Data Bank

RCSB **PDB**
PROTEIN DATA BANK

A MEMBER OF THE **PDB**
An Information Portal to Biological Macromolecular Structures
As of Tuesday Oct 09, 2007 there are 46557 Structures | PDB Statistics

CONTACT US | HELP | PRINT PAGE

PDB ID or keyword Author

Site Search

Advanced Search

Home Search

- Home
- Getting Started
- Download Files
- Deposit and Validate
- Structural Genomics
- Dictionaries & File Formats
- Software Tools
- General Education
- Site Tutorials
- BioSync
- General Information
- Acknowledgements
- Frequently Asked Questions
- Report Bugs/Comments

Quick Tips:

Visit mmcif.rcsb.org for detailed information about the macromolecular Crystallographic Information File (mmCIF) data dictionary.

Are you missing data updates? The PDB archive has moved to <ftp://ftp.wwpdb.org>. For more information click [here](#).

Welcome to the RCSB PDB

The RCSB PDB provides a variety of tools and resources for studying the structures of biological macromolecules and their relationships to sequence, function, and disease.

The RCSB is a member of the **wwPDB** whose mission is to ensure that the PDB archive remains an international resource with uniform data.

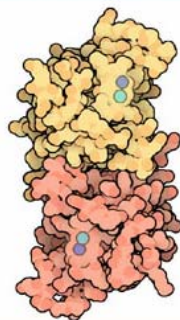
This site offers tools for browsing, searching, and reporting that utilize the data resulting from ongoing efforts to create a more consistent and comprehensive archive.

Information about compatible browsers can be found [here](#).

A **narrated tutorial** illustrates how to search, navigate, browse, generate reports and visualize structures using this new site. [This requires the [Macromedia Flash player download](#).]

Comments? info@rcsb.org

Molecule of the Month: Superoxide Dismutase



We can't live without oxygen. Our cells rely on oxygen as the final acceptor of electrons in respiration, allowing us to extract far more energy from food than would be possible without oxygen. But oxygen is also a dangerous compound. Reactive forms of oxygen, such as superoxide (oxygen with an extra electron), leak from the respiratory enzymes and wreak havoc on the cell. This superoxide can then cause mutations in DNA or attack enzymes that make amino acids and other essential molecules. This is a significant problem: one study showed that for every 10,000 electrons transferred down the respiratory pathway in *Escherichia coli* cells, about 3 electrons end up on superoxide instead of the proper place. To combat this potential danger, most cells make superoxide dismutase (SOD, shown here from PDB entry **2sod**), an enzyme that detoxifies superoxide.

- More ...
- Previous Features

The RCSB PDB is managed by two members of the RCSB: Rutgers, The State University of New Jersey and the San Diego Supercomputer Center and Skaggs School of Pharmacy and Pharmaceutical Sciences at the University of California, San Diego. It is supported by funds from the National Science Foundation, the Department of Energy, the National Institutes of Health, and the National Cancer Institute.

News

- Complete News
- Newsletter
- Discussion Forum

9-October-2007

PDB archive at <ftp://ftp.wwpdb.org>

As previously announced, the PDB archive has been moved to <ftp://ftp.wwpdb.org>. Updated weekly, this location serves the files from the **wwPDB Remediation Project** and all newly released files. In September, approximately 14.7 million files were downloaded from [ftp.wwpdb.org](ftp://ftp.wwpdb.org).

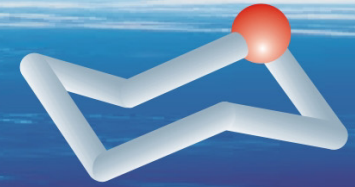
- Full article ...

2-October-2007

Structure Deposition Checklist

In citing the PDB please refer to: H.M. Berman, J. Westbrook, Z. Feng, G.

Basic Local Alignment Search Tool



BLAST: Basic Local Alignment and Search Tool - Microsoft Internet Explorer

Fájl Szerkesztés Nézet Kedvencek Eszközök Súgó

Vissza Keresés Kedvencek

Cím <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi> Ugrás Hivatkozások

Google Go Bookmarks 12 blocked Check AutoLink Settings

mywebsearch Search Screensavers Smiley Central Cursor Mania Fun Cards

BLAST Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help My NCBI [Sign In] [Register]

► NCBI/ BLAST Home

BLAST finds regions of similarity between biological sequences. [more...](#)

[Learn more](#) about how to use the new BLAST design

BLAST Assembled Genomes

Choose a species genome to search, or [list all genomic BLAST databases](#).

Human	Oryza sativa	Gallus gallus
Mouse	Bos taurus	Pan troglodytes
Rat	Danio rerio	Microbes
Arabidopsis thaliana	Drosophila melanogaster	Apis mellifera

Basic BLAST

Choose a BLAST program to run.

News

New Human and Mouse pre-indexed databases

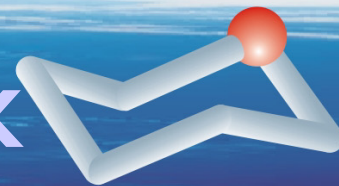
Human and mouse genomic + transcript megablast searches now use a faster, indexed algorithm that typically reduces run time by two thirds, as compared with standard megablast.

2007-09-04 10:55:00

[More BLAST news...](#)

Internet

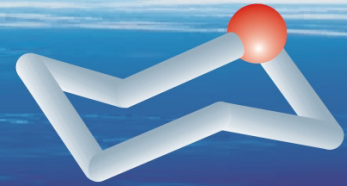
PDB: Protein Data Bank



- Biológiai makromolekulák szerkezetének információs portálja
- Keresés:
 - PDB ID (azonosító)
 - Kulcsszó
 - szerző
- Röntgenkristallográfiai mérések alapján
- Irodalmi hivatkozások
- Megjelenítés változatos grafikus formában:
 - Szalag
 - Vonal
 - Atomok jelölve...



Példa



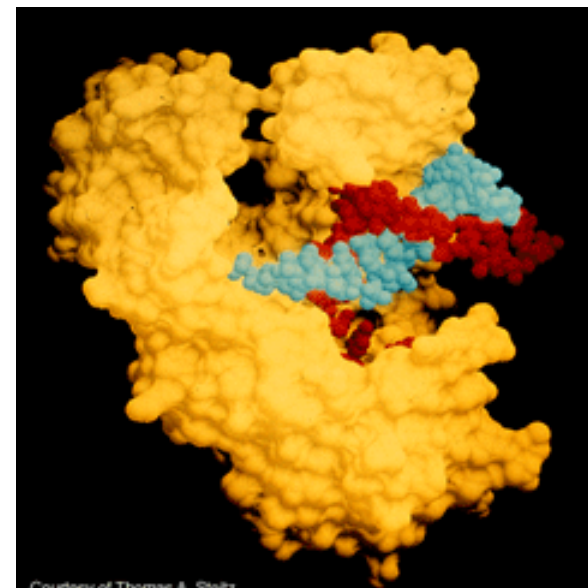
DNS polimeráz I enzim

Nukleotidot épít be
vagy hasít ki

Replikáció

Repair

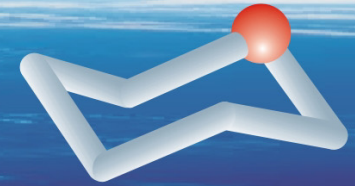
Hidroláz és transzferáz
aktivitás



Courtesy of Thomas A. Steitz.



Human Protein Atlas



Human Protein Atlas - Microsoft Internet Explorer

Eájl Szerkesztés Nézet Kedvencek Eszközök Súgó

Vissza Keresés Kedvencek

Cím http://www.proteinatlas.org/ Ugrás Hivatkozások

Google Go Bookmarks 12 blocked Check AutoLink Settings

mywebsearch Search Screensavers Smiley Central Cursor Mania Print Cards

Check spelling of text you type on any web form in English

hpa the project protein atlas dictionary disclaimer submission of antibodies

HUMAN PROTEIN ATLAS

The human protein atlas shows expression and localization of proteins in a large variety of normal human tissues, cancer cells and cell lines with the aid of immunohistochemistry (IHC) images.

Enter search: search

[Advanced search](#)

Or choose a chromosome:

1	5	9	13	17	21
2	6	10	14	18	22
3	7	11	15	19	X
4	8	12	16	20	Y

OTHER

2007-10-09
A new version of the Human Protein Atlas has been released. The new version contains more than 2.8 million images corresponding to more than 3000 antibodies. Protein Array-data and Western Blot-data have been added.

2007-10-09
A new feature with immunofluorescent (IF) images generated with confocal microscopy has been added. At present, the subcellular localization for 769 antibodies in three human cell lines are shown.

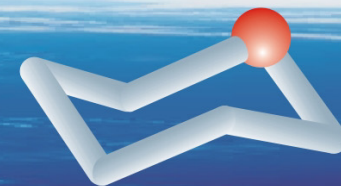
2007-10-09
A new feature has been added to allow the possibility to search for proteins with

Version: 3.0 Atlas updated: 2007-10-09 ([release history](#))
Atlas content: 3015 antibodies and 2,827,440 images.

Kész Internet



Y-kromoszóma



Human Protein Atlas - Microsoft Internet Explorer

http://www.proteinatlas.org/search_advanced.php?chr=y

Search Results

Search results for chromosome Y: **17 hits** (genes)

Choose, if available:

- an Antibody ID to view the annotation data
- a link button to open a new window with Ensembl/NCBI/RefSeq/Uniprot info

#	Gene name	Description	Chr	Links	Antibody ID	Validation
1	ASMTL	N-acetylserotonin O-methyltransferase-like protein (ASMTL).	X/Y:p22.33	U E	HPA003630	Medium
2	CD99	CD99 antigen precursor (T-cell surface glycoprotein E2) (E2 antigen) (Protein MIC2) (12E7).	X/Y:p22.33	U N E	CAB000020	N/A
3	CY014_HUMAN	Uncharacterized protein CYorf14.	Y:q11.222	U E	HPA003249	Very low
4	DDX3X	ATP-dependent RNA helicase DDX3Y (EC 3.6.1.-) (DEAD box protein 3, Y-chromosomal).	Y:q11.21	U N E	HPA001648 HPA005631	Medium Medium
5	EIF1AY	Eukaryotic translation initiation factor 1A, Y-chromosomal (eIF-1A Y isoform) (eIF-4C).	Y:q11.223	U N E	HPA002561	Low
6	FAM39DP	CXYorf1 protein (Fragment).	X/Y:q28	U N E	HPA002689	Very low
7	IL3RA	Interleukin-3 receptor alpha chain precursor (IL-3R-alpha) (CD123 antigen).	X/Y:p22.33	U N E	HPA003539	Low
8	IL9R	Interleukin-9 receptor precursor (IL-9R) (CD129 antigen).	X/Y:q28	U N E	CAB010497	N/A
9	P2RY8	P2Y purinoceptor 8 (P2Y8).	X/Y:p22.33	U N E	HPA003631	Low
		Protocadherin-11 Y-linked precursor				

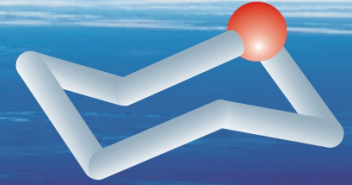
Show search results: 1 to 17

[help for this page](#)



ExpASY

Expert Protein Analysis System

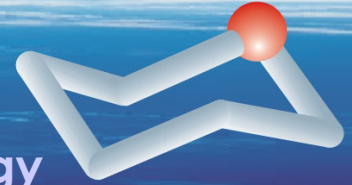


- www.expasy.org
- Fehérje szerkezetek és szekvenciák analízisére hozták létre
- 5 adatbázis, ebből egy az ENZYM
- Keresés:
 - EC szám
 - Enzim név
- Definiálja az enzim funkciót
- Megadja az ismert szerkezeteket
- Kapcsolat más adatbázisokkal



IUBMB:

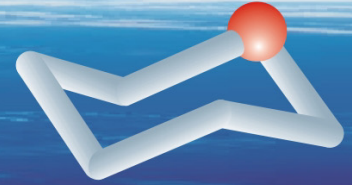
International Union of Biology and Molecular Biology



- Biológiai (szerves) nomenklatura és terminológia
- Keresés:
 - EC szám
 - Enzim név
 - kulcsszó
- Definiálja az enzimek funkcióját
- Enzimkinetikai terminológia is
- Nukleinsavak
- Peptidhormonok
- Transzportfehérjék
- Linkek egyéb adatbázisokhoz



BRENDA

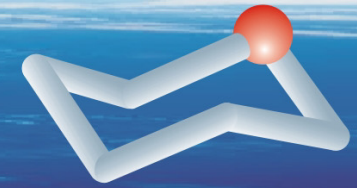


Egy átfogó enzim információs rendszer

- Nomenklatúra
- Preparálás
- Stabilitás
- Reakcióspecifitás
- Enzim szerkezet
- Funkcionális paraméterek
- Termelő szervezet
- Betegségek
- Alkalmazás és mérnökség



MACiE adatbázis



Mechanism, Annotation and Classification in Enzymes.

<http://www-mitchell.ch.cam.ac.uk/maciel/>

ID	42	Unique Identifier	0042.ov	Overall Reaction		Annotation	Overall ReactionType=(Hydrolysis); BondsMade=(O-H, O-P); AtomsInvolved=(O, P, O); Cofactors=(Mg2+); Product=(X00015, X00016); BondsBroken=(O-H, O-P); Substrate=(C00039, C00001); BondsInvolved=(O-H, O-P, O-P, O-H).
Name	Endonuclease						
EC	3.1.30.2						
PDB	1smn						

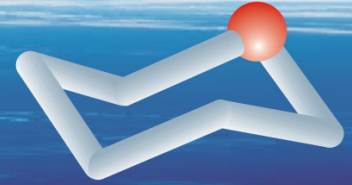
Step_No	Step_ID	Reaction	Comments	Annotation
1	0042.stg01		<p>The diagrammatic mechanism stated in the paper assumes that the mechanism is concerted.</p> <p>It is assumed that His89 is activating the water.</p>	<p>Substrates=(1smn-HIS89, C00001, C00039); BondFormed=(N-H, O-P, O-H); NucAttackOn=(C00001, C00039, 1smn-HIS89); Cofactors=(Mg2+); BondCleared=(O-H, O-P, N-H); RxnAtt=(Bond Cleavage, Bond Formation, Product Formed, Hydrolysis); ReactionCentres=(N, H, O, P, O); Reversible=(Yes); Glu127=(Side Chain, OnLooker, HBA); RDS=(Yes); Asn119=(HBA, HBD, Side Chain, OnLooker, MetalLigand); His89=(Side Chain, PA, Nuc, HBA, PD, Elec); ResiduesPresent=(His89, Glu127, Asn119, Arg57); RxnTypes=(NucleophilicAttack, Proton Transfer, Sn2); Products=(X00015, X00016); ProtonTransferFromTo=(C00001, 1smn-HIS89, 1smn-HIS89, C00039); Nucleophile=(1smn-HIS89, C00001, C00039); Arg57=(Side Chain, OnLooker, HBD, ChargeStabiliser).</p>

Ref_No	Ref	PubMed_ID
1	S. V. Shlyapnikov, V. V. Lunin, M. Perbandt, K. M. Polyakov, V. Yu Lunin, V. M. Levdivkov, C. Betzel, A. M. Mikhailov; Acta Cryst.: D66; 667-672; 2000	20235660

Subs_No	Subs_ID	Substrate_Structure	Substrate_Name	Product_No	Prod_ID	Product_Structure	Product_Name
1	C00039		DNA	1	X00015		DNA 5'-phosphate
2	C00001		Water	2	X00016		DNA 3'-hydroxyl



Kutatási kérdések



- Hogy történik a katalízis?
- Megjósolható-e az enzim működése a szerkezetből?
- Tervezhetünk-e új enzimeket?
- Összevethetők-e a különböző szervezetből származó azonos funkciójú enzimek?

Az aktív helyek ismerete szükséges.



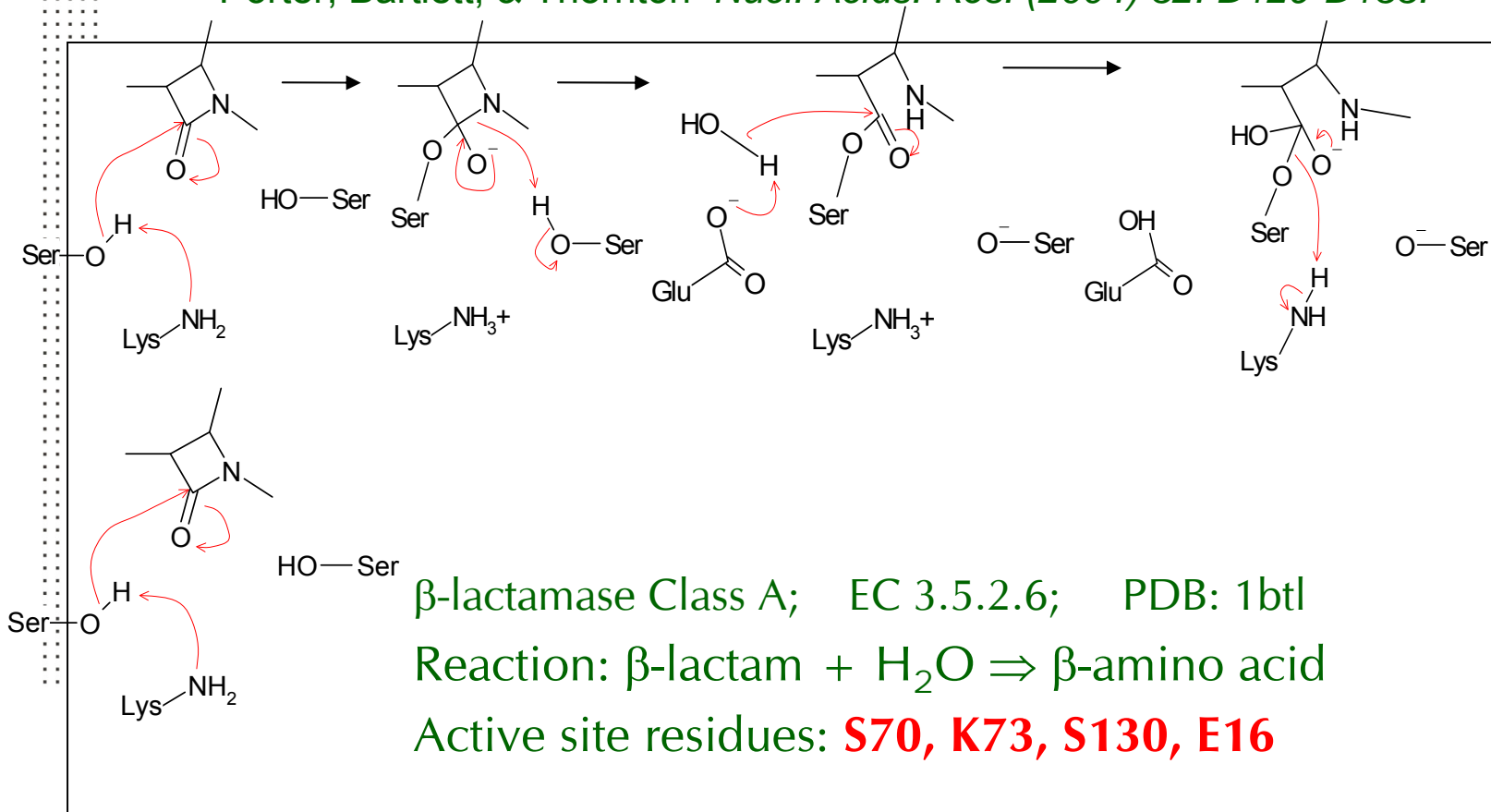
Catalytic Site Atlas



<http://www.ebi.ac.uk/thornton-srv/databases/CSA>

Catalytic Site Atlas: Az enzimekben szerkezeti adatok alapján azonosított katalitikus helyek és oldalláncok forrása

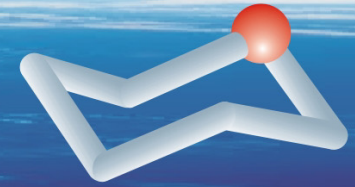
Porter, Bartlett, & Thornton *Nucl. Acids. Res.* (2004) 32: D129-D133.





IntEnz

Integrated Enzyme Database



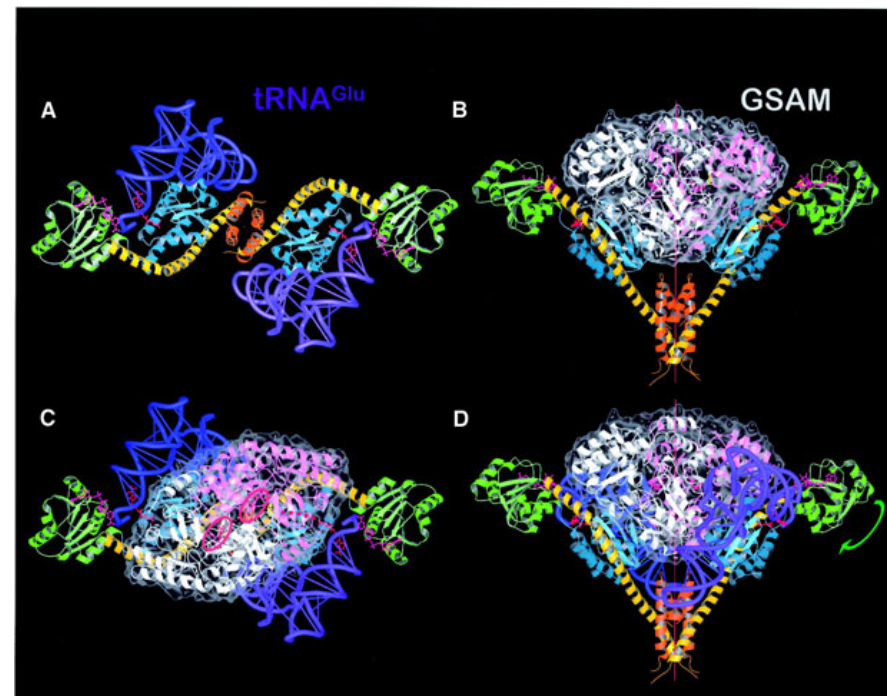
- Keresés: EC szám, név, katalizált reakció
- Abstract: mechanizmus, metabolikus kapcsolatok, hivatkozások
- Linkek: ExPASy
- Példa:

Glutamil-tRNS reduktáz

δ -amino levulinsav szintézis
glutaminsavból.

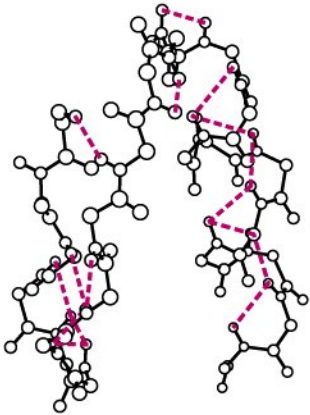
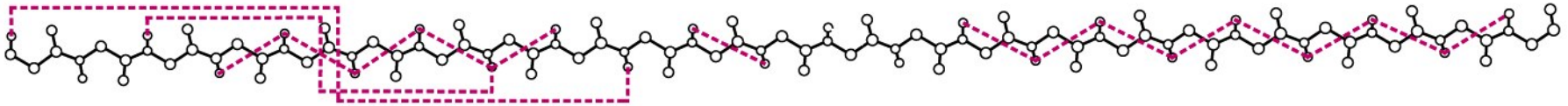
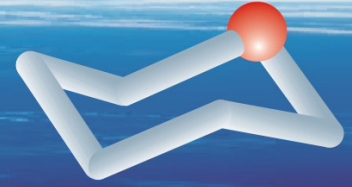
Dimer enzim, V alakú.

Általában hem tartalmú,
de ez arche-baktérium
eredetű, nincs prosztetikus
csoport.





Molekula modellezés



Calculated structure

Figure 3-52

Biochemistry, Sixth Edition

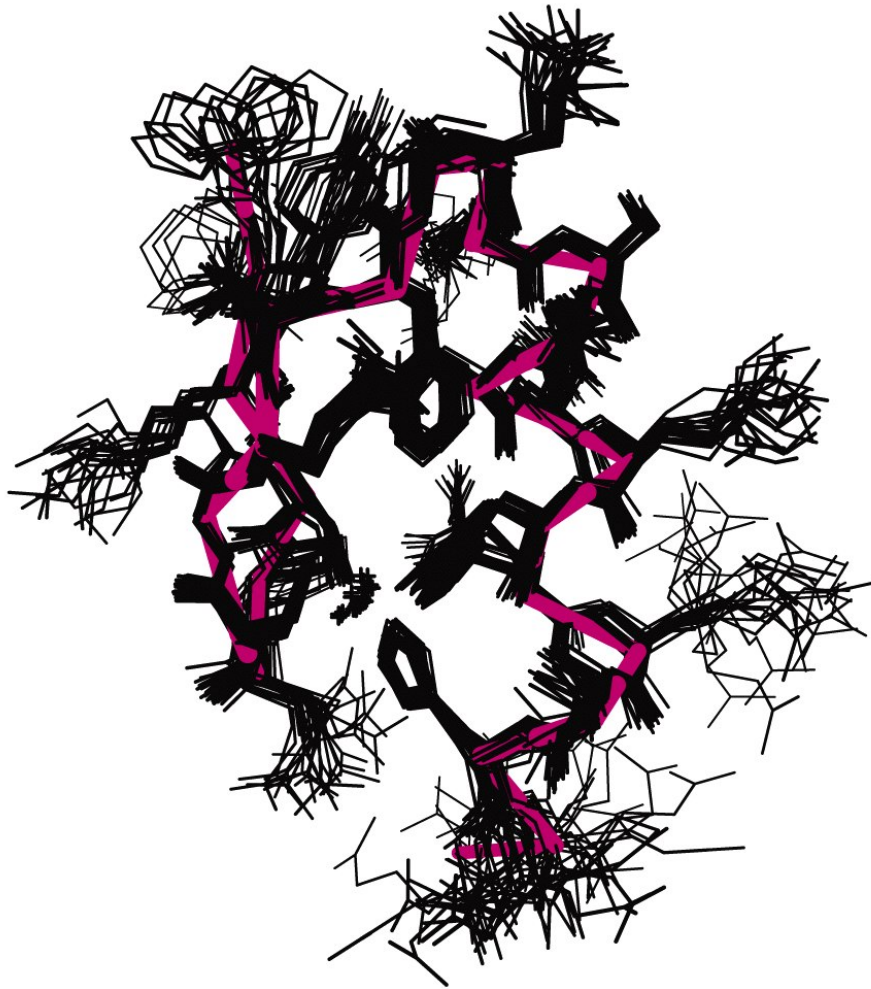
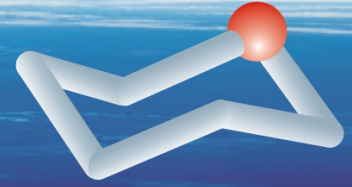
© 2007 W. H. Freeman and Company

- Számítógépes programmal szerkezet megadás.
- Felhasználja az ismert szekvencia-szerkezet összefüggéseket.
- Felhasználja a szerkezetvizsgáló módszerekkel kapott adatokat.





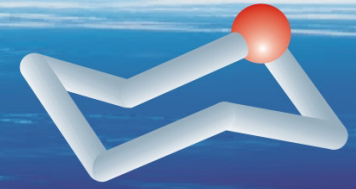
Dinamikus szerkezet



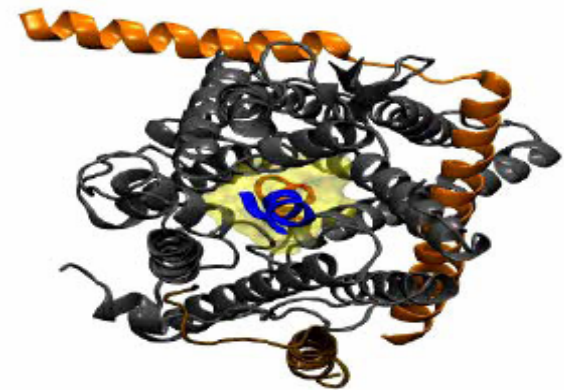
- Figyelembe veszi a molekula oldatbeli mozgásait.
- A szerkezetben vannak rögzített és mozgékony oldalláncok.

Figure 3-53
Biochemistry, Sixth Edition
© 2007 W. H. Freeman and Company

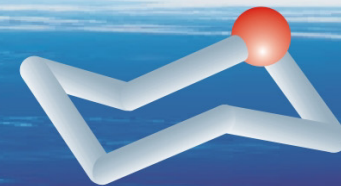
Modellezés folyamata



- A molekulamodellezés célja: térbeli szerkezet ábrázolása
- Kísérleti eredmények reprodukálása
- Molekulatervezés
 - Pénz és idő megtakarítás
 - Az enzimműködés megjósolható
 - Célmolekulák tervezése
 - Új termékek



Molekuláris modellező programok



TCE VMD - Visual Molecular Dynamics

THEORETICAL and COMPUTATIONAL BIOPHYSICS GROUP
NIH RESOURCE FOR MACROMOLECULAR MODELING AND BIOINFORMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Home
Overview
Publications
Research
Software
VMD Molecular Graphics Viewer
NAMM Molecular Dynamics Simulator


VMD
Visual Molecular Dynamics

VMD is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics and built-in scripting. VMD supports computers running MacOS-X, Unix, or Windows, is distributed free of charge, and includes source code.
[\(more details...\)](#)

ExpASY - SWISS-3DIMAGE

[ExpASY Home page](#) | [Site Map](#) | [Search ExpASY](#) | [Contact us](#)

Search for

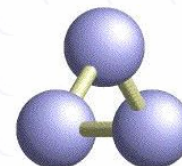
 **Database of annotated 3D images**

SWISS-3DIMAGE ([References](#)) is an image database which strives to provide high quality pictures of biological macromolecules with known three-dimensional structure. The database contains mostly images of experimentally elucidated structures, but also provides views of well accepted theoretical protein models. The images are provided in several useful formats; both mono and stereo pictures are generally available ([Disclaimer](#)).

Direct access to SWISS-3DIMAGE

- by description or identification (any word in the DE, OS, KW and ID lines, or PDB code)
- by SWISS-PROT accession number (AC line)

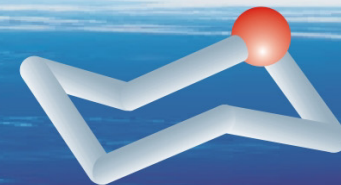
Home Page for RasMol and OpenRasMol



Molecular Graphics Visualisation Tool

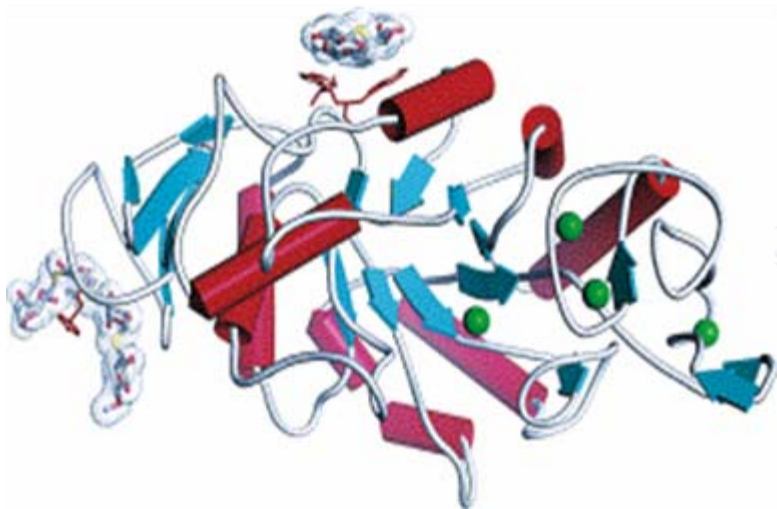
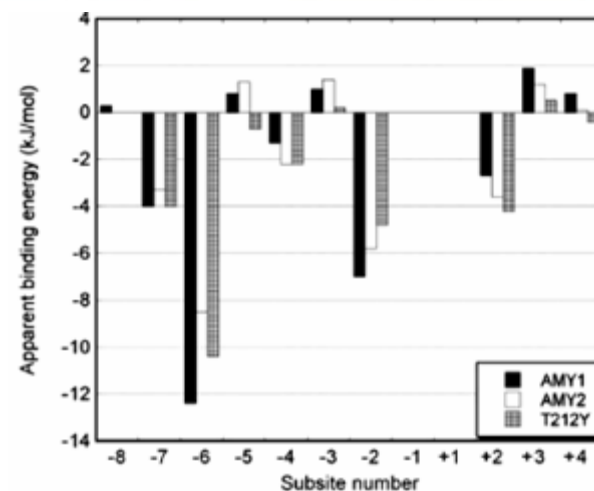


Molekula modellezés



1) működési sajátosságok meghatározása,
az enzim szubsztrát-kötésének vizsgálata,
kölsönhatási energiák meghatározása

2) molekulák keresése adatbázisból
(információk szekvenciáról, szerkezetről, működésről)



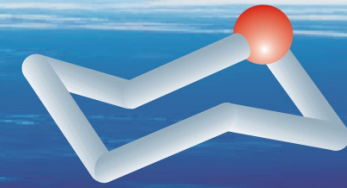
RCSB **PDB**
PROTEIN DATA BANK



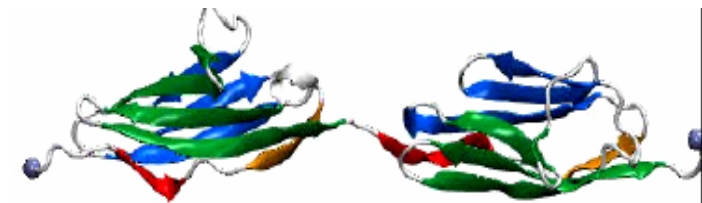
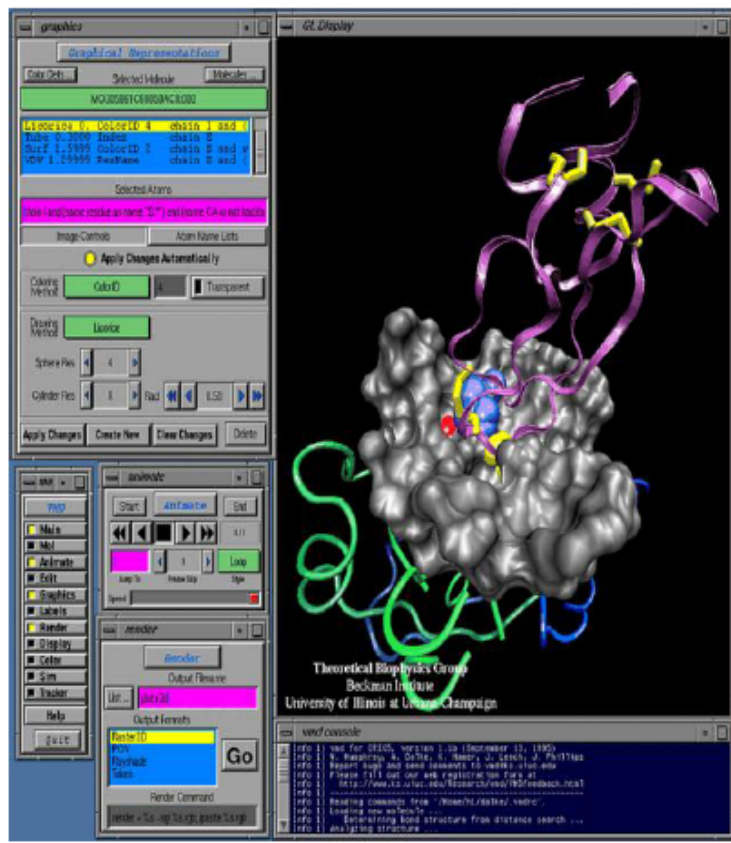
WORLDWIDE
wwPDB
PROTEIN DATA BANK



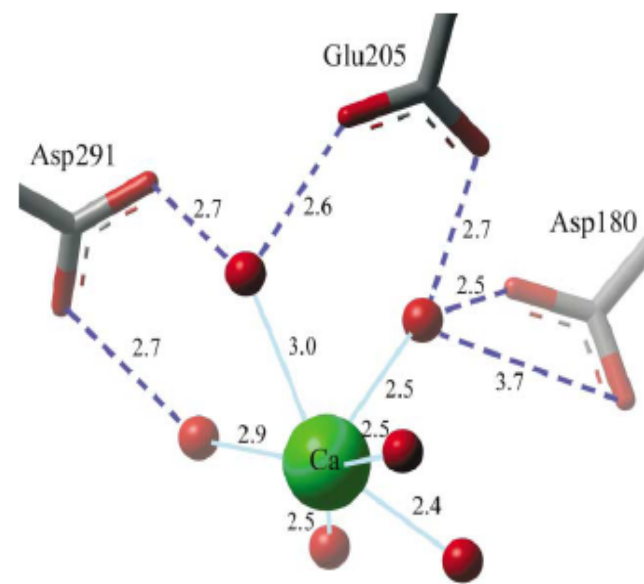
Molekula modellezés



3) molekulamodellezés számítógépes programok segítségével



Molekuladinamika: fehérjék szerkezetének időbeli változását is lehet modellezni
Az ábrán az izomműködés során a titin Z1Z2 fehérje kitekeredése látható



Kölcsönhatások az árpa amilázban

Home

[Functional Glycomics Update](#)
[Consortium for Functional Glycomics \(CFG\)](#)
[CFG Resources](#)
[CFG Data](#)
[CFG Databases](#)
[Links](#)
[About us](#)
[Suggestion](#)
[Sitemap](#)

Toolbox

E-alert sign up

Welcome to the home for Functional Glycomics research

Welcome to the Functional Glycomics Gateway, a comprehensive and regularly updated resource for functional glycomics research.

This gateway represents a unique collaboration between the Consortium for Functional Glycomics (CFG) and Nature Publishing Group (NPG). It is designed to provide up-to-date information and resources for research aimed at elucidating the roles of glycan-protein interactions in cell surface biology.

FUNCTIONAL GLYCOMICS UPDATE

Find out about the latest developments in the field. The Functional Glycomics Update is brought to you by Nature Publishing Group.

Last updated: 08 November 2007

This month's features:

[Carbohydrate recognition: A chemistry lab's lectin mimicry](#)

[Stem cell migration: Your brain on \(polysialic\) acid](#)

[read more](#)

E-alert sign up

CFG RESOURCES

[Materials and services available from CFG](#)



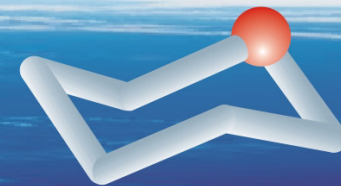
CFG DATA

[Phenotyping, profiling, array data, and more](#)



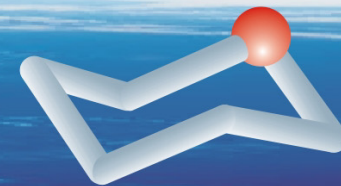
CFG DATABASES

Consortium for Functional Glycomics



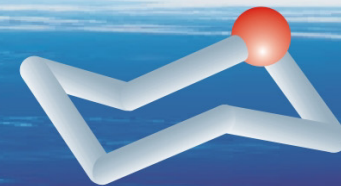
General	Reference	Genome	Proteome	Glycome	Biology
DC-SIGN (Human)					
CFG ID	cbp_hum_Ctlect_00121				
GBP Name	DC-SIGN				
Category	C-Type Lectin Family				
Sub-Family	2-Type 2 Receptor				
Other Names	CD209 antigen, Dendritic cell-specific ICAM-3-grabbing nonintegrin 1, DC-SIGN1				
Species	Human				
Summary					
Date Updated	05/24/2020				
Email	glycomics-web@mit.edu				

Consortium for Functional Glycomics



General	Reference	Genome	Proteome	Glycome	Biology
DC-SIGN (Human)					
Gene Name (Primary)	DC-SIGN				
Other Gene Names	CD209 antigen				
Chromosome	19				
Complete CDS (bp)	1215 bp				
Open Reading Frame (ORF)	<pre> atgagtgact ccaaggaacc aagactgcag cagctgggcc tcttggagga ggaacagctg agaggccttg gattccgaca gactcgagga tacaagagct tagcagggtg tcttggccat ggtcccctgg tgctgcaact cctctccttc acgetcttgg ctgggctcct tgtccaagtg tccaaggtec ccagctccat aagtcaggaa caatccaggc aagacgcgat ctaccagaac ctgaccagc ttaaagctgc agtgggtgag ctctcagaga aatccaagct gcaggagatc taccaggagc tgaccagct gaaggctgca gtgggtgagc ttccagagaa atctaagctg caggagatct accaggagct gacccggctg aaggctgcag tgggtgagct tccagagaaa tctaagctgc aggagatcta ccaggagctg acctggctga aggctgcagt ggggtgagctt </pre>				

Consortium for Functional Glycomics



General

Reference

Genome

Proteome

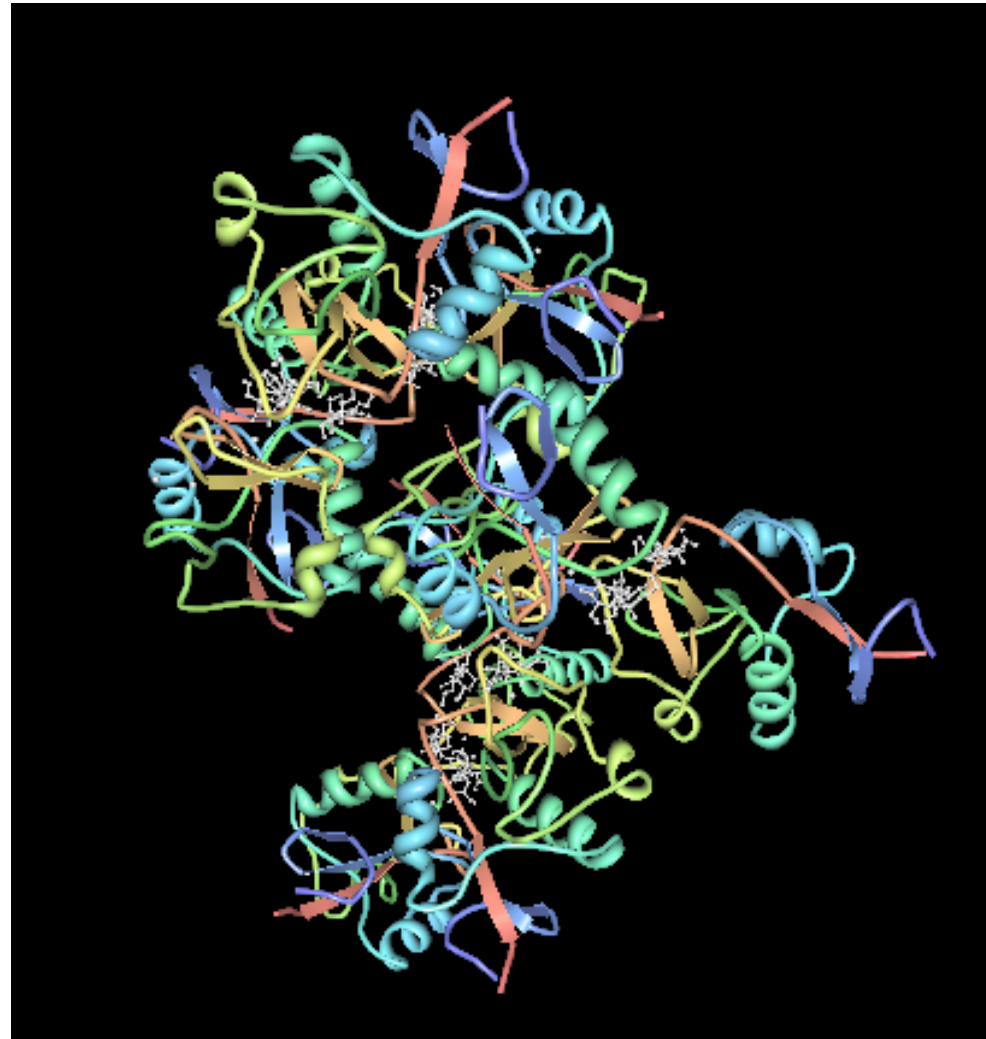
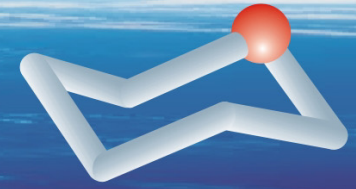
Glycome

Biology

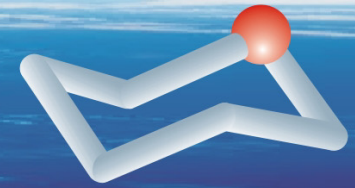
DC-SIGN (Human)

Protein Primary Sequence	MSDSKEPRLQ QLGLEEEQL RGLGFRQTRG YKSLAGCLGH GPLVLQLLSF TLLAGLLVQV SKVPSSISQE QSRQDAIYQN LTQLKAAVGE LSEKSKLQEI YQELTQLKAA VGELPEKSKL QEIIYQELTRL KAAVGELPEK SKLQEIYQEL TWLKAAVGEL PEKSKMQEIY QELTRLKAAV GELPEKSKQQ EIIYQELTRLK AAVGELPEKS KQQEIYQELT RLKAAVGELP EKSKQQEIYQ ELTQLKAAVE RLCHPCPWEW TFFQGNCYFM SNSQRNWHDS ITACKEVGAQ LVVIKSAEEQ NFLQLQSSRS NRFTWMGLSD LNQEGTWQWV DGSPLLPSFK QYWNRGEPNN VGEEDCAEFS GNGWNDKCN LAKFWICKKS AASCSRDEEQ FLSPAPATPN PPPAAPATPN
Protein Sequence Alignment	Blast DC-SIGN
Domain Organization	Carbohydrate Recognition Domain - Click on individual structures for more information
SwissProt	Q9NNX6
PDB	1K9I , 1SL4 , 1SL5
Isoelectric Point	5.0

Consortium for Functional Glycomics – links to PDBs



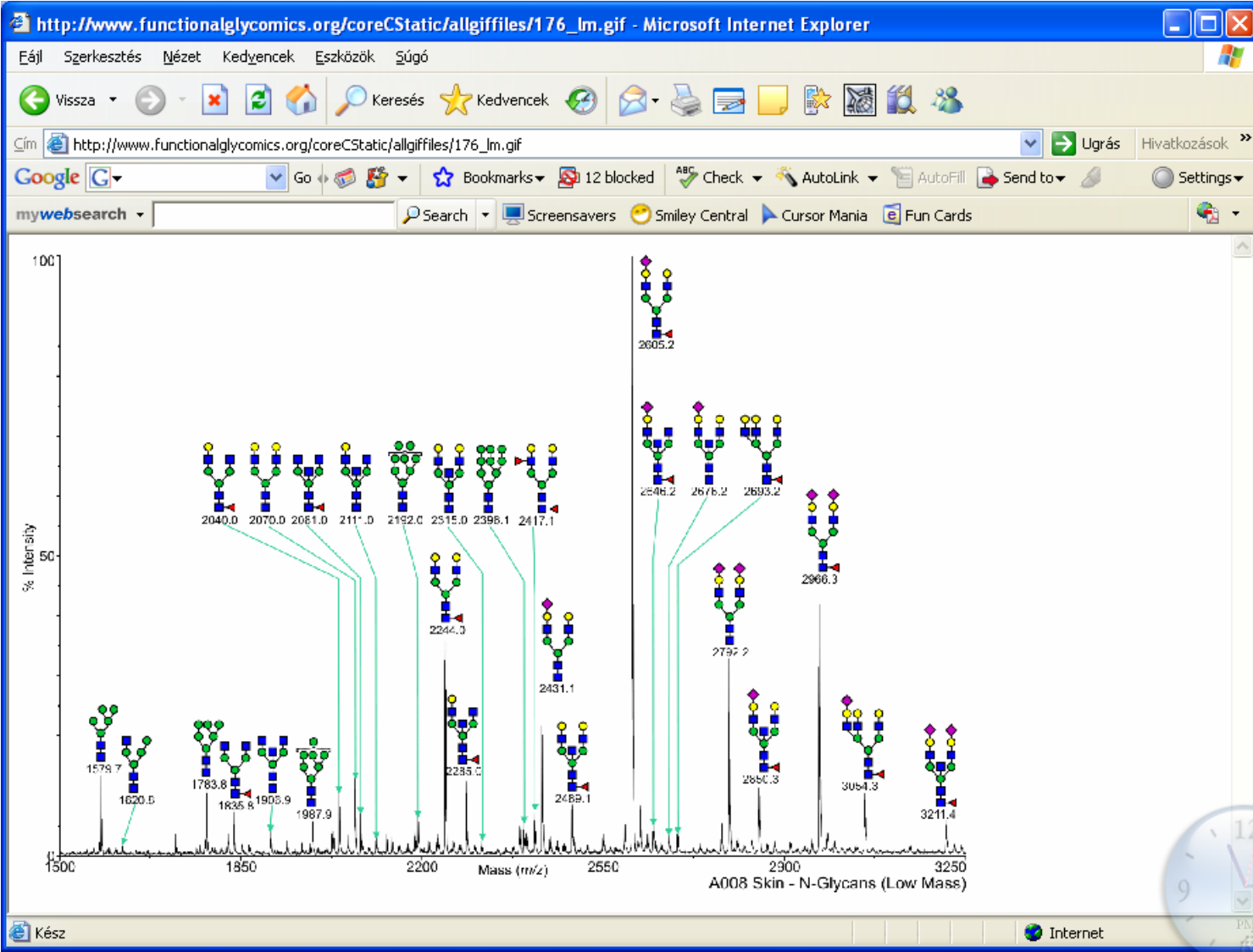
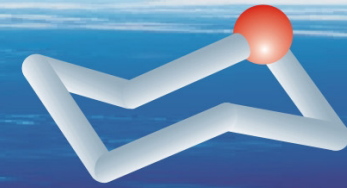
Consortium for Functional Glycomics – search



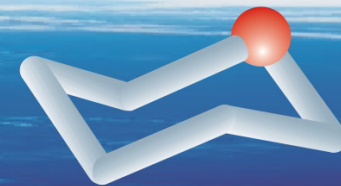
	GlcNAc;[NP]		
3683.34	<p>NeuAca2-3Galb1-4GlcNAcb1-2 (NeuAca2-3Galb1-4GlcNAcb1-6)</p> <p>Mana1-3 (NeuAca2-3Galb1-4GlcNAcb1-2 (NeuAca2-3Galb1-4GlcNAcb1-4) Mana1-6)</p> <p>Manb1-4GlcNAcb1-4 (Fuca1-6) GlcNAc</p>		<p>ex</p> <ul style="list-style-type: none"> • Homo sapiens(Human) • Bos taurus(Bovine)
1949.77	<p>Galb1-4GlcNAcb1-2Mana1-3 (Gala1-3Galb1-4GlcNAcb1-2M ana1-6)Manb1 4GlcNAcb1-4 (Fuca1-6) GlcNAc</p>		<p>ex</p> <ul style="list-style-type: none"> • Torpedo californica(pacific electric ray) • Friend murine leukemia virus(Friend murine leukemia virus) • Bos taurus(Bovine):Brain • Mus musculus (Mouse):Engelbreth-Holm-Swarm tumor, Cancer, Myeloma • Homo sapiens(Human)



Szénhidrát MS



Genography projekt

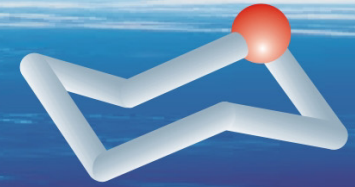


- Az emberiség eredete
- A népvándorlások bizonyítása
- A genetikai azonosság és különbözőség





Genography projekt



```
ATGGAGAAGGTGAAGGGTCTGCTGAAGCCGCGGGCCGACGCCGCAGCAGCA
GTGGCAGCGCAAGCTCCGCAACGAGTGCCGCGTCCTCGACCGCCAGATCCG
GAGGGAGGAGAAGAACGTGGAGAAGTCTATCAGGGAGGCTGCAAAGCGC
GATCCGCAAAGGCTCTTGCTAAGGAACTAGTGAGGTCAAGACGTGCTGTT
ATGAAAACAAGGCTCAACTAAATTCAGTATCAATGCATCTTGGAGAAATT
CAAGAACTGTCCGGTCATCTGTCAAAAAGTGCTGAAGTTATGAAAATCGTT
TGAAAGCTCCAGAATTGGCTGCCACCATGCAAGAATTTAGTGAAGAAATG
GGTGTGATGGAAGAGATGGTCAACGATGCAGTTGATTCAGCTTTGGACTC
GGAGGAGGAAATTGAAGAGGAGGTTGACAAGGTTCTTGCTTCAGTAGCTC
CCTCACAGCTACCCGATGCTGTGAGGAAACAGAGGATAAACCAAGCTTCA
GTCCCAGAAGAGCGACAAGCTGTTGCTGAGGGCGCTGATGATGATGAGGA
```

ORIGINAL SEQUENCE

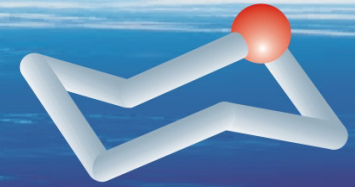
▲
CLICK TO COMPARE
▼

```
ATGGAGAAGGTGAAGGGTCTGCTGAAGCCGCGGGCCGACGCCGCAGCAGCA
GTGGCAGCGCAAGCTCCGCAACGAGTGCCGCGTCCTCGACCGCCAGATCCG
GAGGGAGGAGAAGAACGTGGAGAAGTCTATCAGGGAGGCTGCAAAGCGC
GATCCGCAAAGGCTCTTGCTAAGGAACTAGTGAGGTCAAGACGTGCTGTT
ATGAAAACAAGGCTCAACTAAATTCAGTATCAATGCATCTTGGAGAAATT
CAAGAACTGTCCGGTCATCTGTCAAAAAGTGCTGAAGTTATGAAAATCGTT
TGAAAGCTCCAGAATTGGCTGCCACCATGCAAGAATTTAGTAAAGAAATG
GGTGTGATGGAAGAGATGGTCAACGATGCAGTTGATTCAGCTTTGGACTC
GGAGGAGGAAATTGAAGAGGAGGTTGACAAGGTTCTTGCTTCAGTAGCTC
CCTCACAGCTACCCGATGCTGTGAGGAAACAGAGGATAAACCAAGCTTCA
GTCCCAGAAGAGCGACAAGCTGTTGCTGAGGGCGCTGATGATGATGAGGA
```

COPIED SEQUENCE



Genography projekt



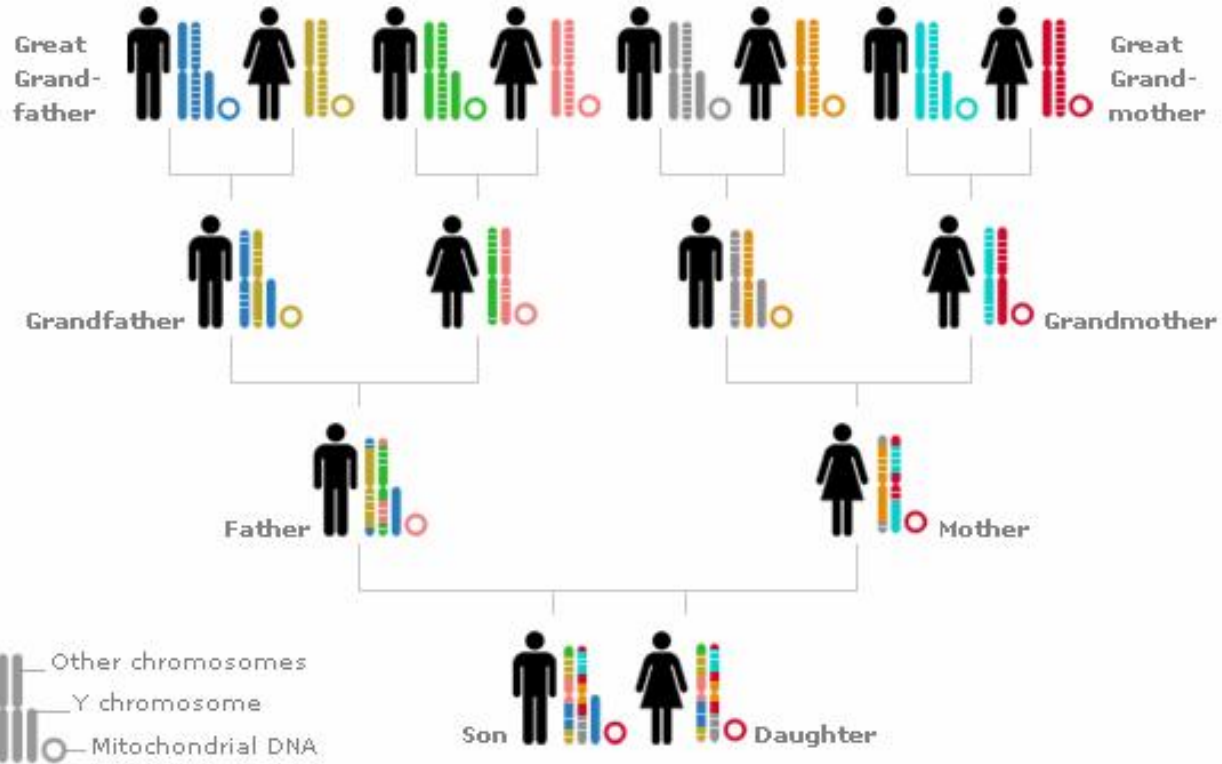
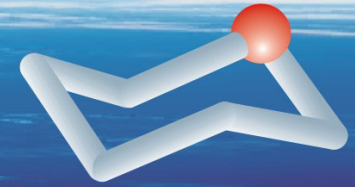
Congratulations! You found the mutation.
There are about 50 random single point changes that distinguish a child's DNA from his or her parents' DNA. **RESET**

AGTGAAG Original
↓
AGTAAAG Copy

```
ATGGAGAAGGTGAAGGGTCTGCTGAAGCCGCGGCCGACGCCGCAGCAGCA  
GTGGCAGCGCAAGCTCCGCAACGAGTGCCGCGTCCTCGACCCAGATCCG  
GAGGGAGGAGAAGAACGTGGAGAAGTCTATCAGGGAGGCTGCAAAGCGC  
GATCCGCAAAGGCTCTTGCTAAGGAACTAGTGAGGTCAAGACGTGCTGTT  
ATGAAAACAAGGCTCAACTAAATTCAGTATCAATGCATCTTGGAGAAATT  
CAAGAACTGTCGGTCATCTGTCAAAAAGTGCTGAAGTTATGAAAATCGTT  
TGAAAGCTCCAGAATTGGCTGCCACCATGCAAGAATTTAGTAAAGAAATG  
GGTGTGATGGAAGAGATGGTCAACGATGCAGTTGATTCAGCTTTGGACTC  
GGAGGAGGAAATTGAAGAGGAGGTTGACAAGGTTCTTGCTTCAGTAGCTC  
CCTCACAGCTACCCGATGCTGTGAGGAAACAGAGGATAAACCAAGCTTCA  
GTCCCAGAAGAGCGACAAGCTGTTGCTGAGGGCGCTGATGATGATGAGGA
```



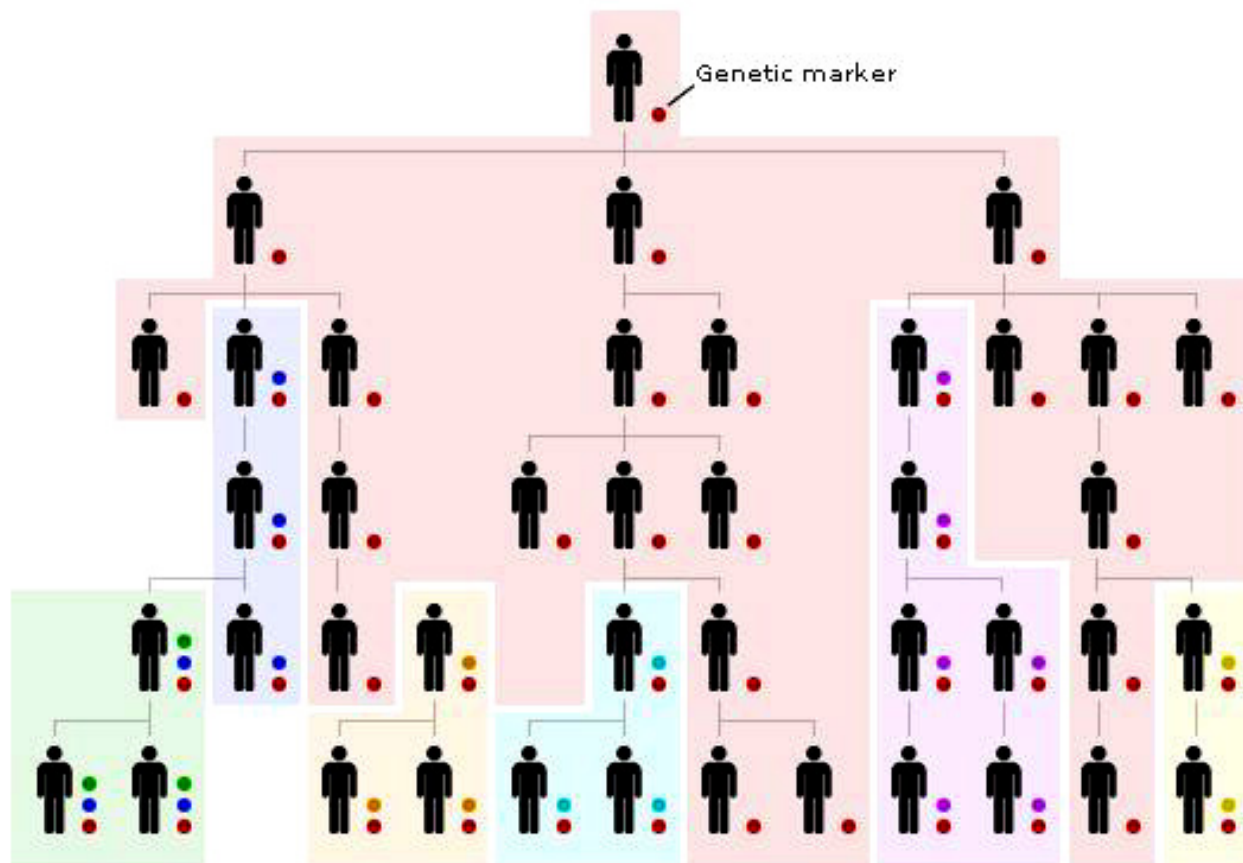
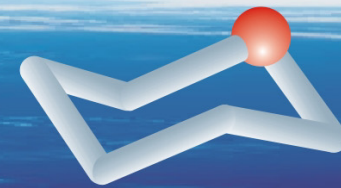
Genography projekt



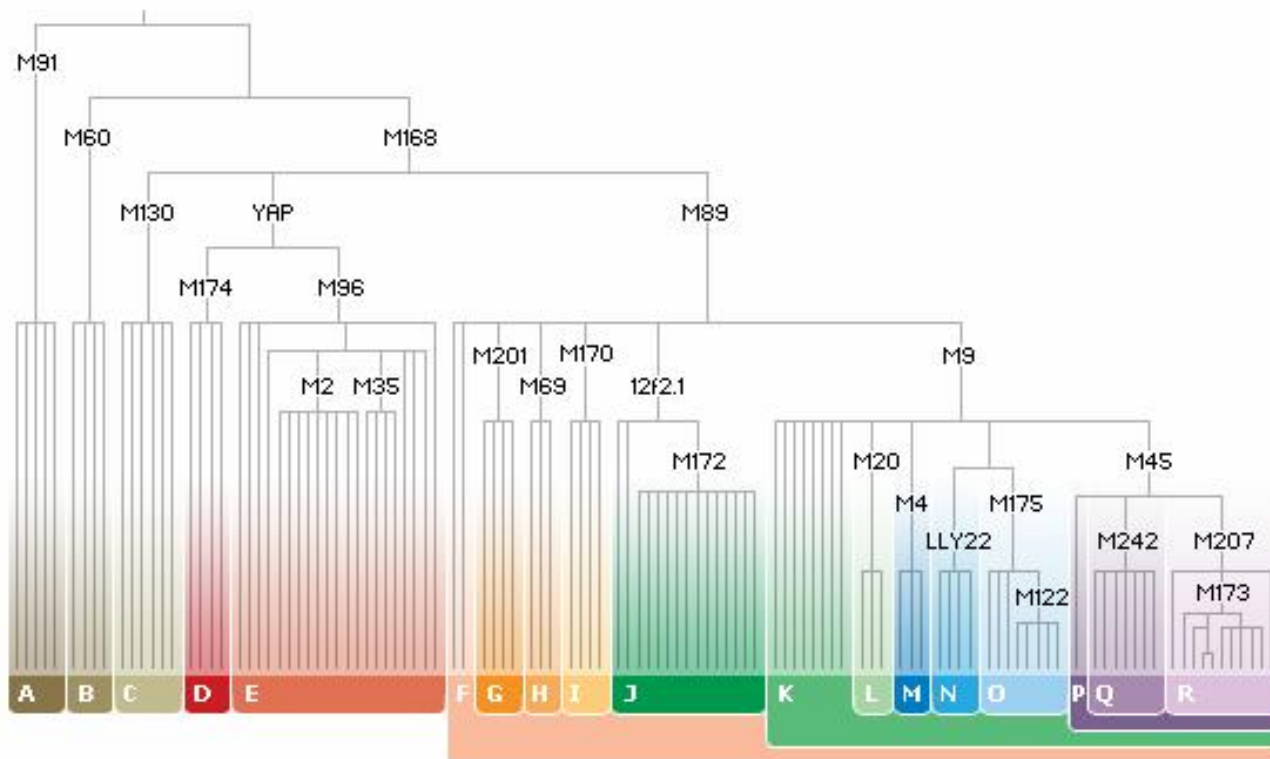
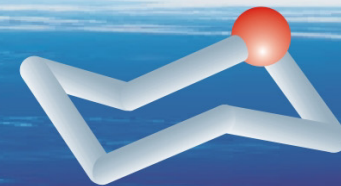
Show both types of DNA Show recombined DNA Show non-recombined DNA



Genography projekt

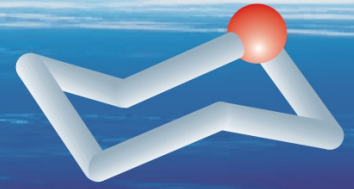


Genography projekt

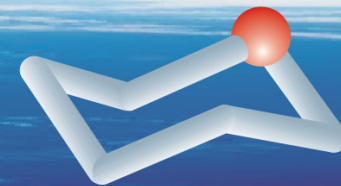




Genography projekt



A hasonlóság csapdája



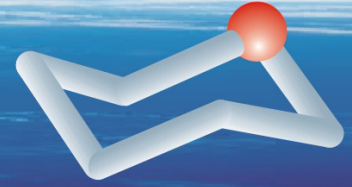
α -laktalbumin



lizozim

50% szekvencia azonosság, nagyon hasonló térszerkezet, teljesen különböző funkció

Ne higgyük el teljesen



- ami az adatbázisokban van (félrevezető vagy hibás lehet a szekvencia vagy a funkció hozzárendelés)
- amit a programok mondanak (a programban hiba lehet, az eredmény félrevezető lehet)
- amit a webszerverek mondanak (dettó)
- amit a szakirodalomban olvasunk (a tévedések nem ritkák)

