# Seminar 8

# Evidentierea legaturii dintre gena si secventa proteica corespunzatoare

(Identificarea genei GFP si determinarea mutațiilor importante)

# Pentru a demonstra realizarea activitatilor urmatoare, creati un fisier Word (nume\_seminar8.doc) in care sa salvati ("print screen" si apoi "paste") imaginile create!

1) Determinarea secventei ADN a proteinei "Green Fluorescent Protein" (in baza de date UniProtKB)

2) Translatia secventei AND

3) Compararea secventei de aminoacizi obtinuta cu secvente similare din baza de date PDB4) Vizualizarea mutatiilor genetice

# 1) Determinarea secventei ADN a proteinei "Green Fluorescent Protein" (in baza de date UniProtKB)

Baza de date **UniProtKB** (www.uniprot.org/

www.expasy.org/):

- organizeaza si noteaza secventele proteice.

- contine informatii importante pentru studierea relatiei dintre secventa proteica si functia proteinei.

- foloseste 2 tipuri de notatii:

-- una alocata manual (de cercetatori) bazata pe articolele publicate

-- alta alocata automat de un software sofisticat.

Pentru aceasta activitate practica vom folosi notatii alocate manual!

Daca ati folosit adresa "www.expasy.org/" se selecteaza "UniProtKB" in lista bazelor de date si se scrie in campul de cautare "Green Fluorescent Protein".

→ C  www.expasy.or	9		☆ <mark></mark>		
	FASY		Home About Contact		
	UniProti/B	search low			
	Sill for D	active protonity and active protonity			
visual Guidance	ExPASy is the SIB Bioinformatics Resource Portal w	hich provides access to scientific databases and	Popular resources		
ategories proteomics genomics	software tools (i.e., resources) in different areas of life sci systems biology, population genetics, transcriptomics etc. you find resources from many different SIB groups as well	ences including proteomics, genomics, phylogeny, (see Categories in the left menu). On this portal as external institutions.	○ UniProtKB Ø SWISS-MODEL STRING		
structural bioinformatics	Featuring today	- PROSITE			
systems biology					
phylogenylevolution	World-2DPAGE Constellar	Latest News			
oopulation genetics	2DPAGE, Make2D-DB II, MiapeGeIDB,	ss.			
ranscriptomics	(details)	(details)			
biophysics			When we come into this world, we		
maging		$\sim$	usually emerge whole. Not everyone is as fortunate though. There are a variety		
T infrastructure			of rare afflictions that push the		
drug design					
esources AZ	How to use this portal?		UniProt Knowledgebase release 2015 11 - 2015-11-11		
inks/Documentation	Features and updates     New to ExPASy     Experienced ExPASy users: what is different		Release notes 549,832 UniProtKB/Swiss-Prot entries (More.) 54,540,801 UniProtKB/TrEMBL entries (More.)		
			(More news) (SIB news)		

Daca ati folosit adresa "www.uniprot.org/" se selecteaza sectiunea "Swiss-Prot"

	t.org			E 🔊 😒
UniProt	UniProtikB - green fluores Database: Pr	cent protein oteinModelPortal		× Advanced - Q Search
BLAST Align Retrieve/ID	mapping	2 VILPT		Help Contact
The mission of UniProt is to p	rovide the scientific communit	y with a comprehensive, high-	quality and freely accessible i	resource of protein sequence and functional information.
UniProtKB	UniRef	UniParc	Proteomes	News 🔤 💟 🛃 🔊
UniProt Knowledgebase	Sequence clusters	Sequence archive	**	Forthcoming changes ^
Swiss-Prot (549,832)	( <u>,</u>	)))		UniProt release 2015_11
Manually annotated and reviewed.		Supporting data		The sense of a motion   Change of the cross- references to eggNOG
TrEMBL (54,540,801)	Literature citations	Taxonomy	Subcellular locations	UniProt release 2015_10 The small of the sea in UniProtKB 1 Cross-
Automatically annotated and not	Cross-ref. databases	Diseases	Keywords	references to WBParaSite   Removal of the cross- references to CYGD   UniParc cross-reference t
reviewed.	0 r.	XXX	ЦA	News archive
Getting started	You Tube	UniProt data		Protein spotlight
Q Text search Our basic text search allows resources available	you to search all the	± Download latest release Get the UniProt data		The Makings Of A Face November 2015
♣ BLAST Find regions of similarity bety	reen vour sequences	d Statistics View Swiss-Prot and TrEMB	L statistics	When we come into this world, we usually emerge whole. Not everyone is

si se scrie in campul de cautare "Green Fluorescent Protein".

UniProtKB ×	1							_
C 👔 🗋 www.unipro	ot.org/							숬 🔤
Scopus preview - Scop	e Blogilat	es — Fitness,	F 🙁 Google 🚳	Analize	Medicale - Col 🧿 Live Streaming Vir	deo / 👳 Veria   Heal	thy Recipe 📴 2048 😁 Culo de ywsfol 🔬 Com	ori Delicioase: Cui
			ø	/			A AFT AN ANALY AND	
LloiProt			- 4		otKB 🗸		Ac Ac	lvanced 🚽 🔍 🧃
						and the second second		11.25 No. 1
DLACT Align Detrigue	/ID Max	mine						Hales Contracts
BLAST Aligit Retrieve,	по мар	ping					Show b	alp for UniProtVP
-							510% 11	A Packat -
Results								B DOSKOU V
Filter by	R	Columns	SBLAST 🖷	Align	± Download	ket	┥ 1 to 25 of 84,501,513 🕨	Show 25 🔻
Peviewed								
(546,439)	-		Entry name 🗘		Protein names 🖨 🛛 🕨	l Gene names 🖨	Organism 🗢	Length 🗘 🗶
Sviss-Prot		Q6GZX3	002L_FRG3G	<b>⊷</b>	Uncharacterized protein	FV3-002L	Frog virus 3 (isolate Goorha) (FV-3)	320
(83,955,074)		06GZY1	004P EPG3G		UU2L Uncharacterized protein	EV3-004P	Eron virus 3 (isolate Goorba) (EV=3)	60
TrEMBL		QUALITY	0011(_11030	☆	004R	1.10.0041	(i og viras 5 (isolate doorna) (i v 5)	00
Popular organisms		Q6GZX0	005R_FRG3G	÷	Uncharacterized protein	FV3-005R	Frog virus 3 (isolate Goorha) (FV-3)	204
Human (138,517)		060772	0020 50020		UU5K	EV2-002P	Erron virus 2 (isolate Goorba) (EV-2)	439
Rice (99,936)		QUULTE	0001011000	<b>\$</b>	3R	110 0001		100
Mouse (74,058)		Q6GZW8	007R_FRG3G	٨	Uncharacterized protein	FV3-007R	Frog virus 3 (isolate Goorha) (FV-3)	128
Zebrafish (56,277)		0607002	0101 50000		UU/R	EV2-019	Free virus 2 (isolate Coerba) (EV-2)	207
A. trialiaria (52,986) Other organisms		0002003	012L_1 KG3G	Υ.	012L	1 45 0121	(rv-3)	297
Go		Q197F3	007R_IIV3	<b>A</b>	Uncharacterized protein	IIV3-007R	Invertebrate iridescent virus 3 (IIV-3)	447
		0407044	0110 50636		007R	510.0110	(Mosquito iridescent virus)	70
View by		Q6G2W4	UIIK_FRG3G	ţ,	011R	FV3-011R	Frug virus a (isulate Gooma) (FV-3)	70
Taxonomy		Q197F5	005L_IIV3		Uncharacterized protein	IIV3-005L	Invertebrate iridescent virus 3 (IIV-3)	217
Keywords				Π.	005L		(Mosquito iridescent virus)	

Rafinati cautarea facand click pe "reviewed"

← → C 🗋 www.unipro	t.org/	/uniprot/?qu	ery=Green+Fluores	scent+F	rotein&sort=score			\$
UniProt	( [	IniProtikB 👻 🕻	green fluorescent prote	ein			Advance	d 🗸 🔍 Sear
BLAST Align Retrieve/ID	mappir	ng		ΠÍ		A PLANT	201 70	Help Cont
UniProtKB r	esi	ults					② About UniProt	KB 🏦 Baske
Filter by <sup>i</sup>	*	BLAST 🗮 A	lign 土 Download f	add t	o basket 🖉 Columns	>	◀ 1 to 25 of 474	Show 25
Reviewed (70)	Qu	uote terms:	"green fluorescent p	rotein"				
Upperviewed (404)		Entry 🔷	Entry name 🗘		Protein names 🗘 🗵	Gene names 🖨	Organism 🖨	Length 🗘
TrembL		P42212	GFP_AEQVI	<b>k</b>	Green fluorescent protein	GFP	Aequorea victoria (Jellyfish)	238
Popular organisms A. thaliana (12)		PODM59	UNAG_ANGJA		Bilirubin-inducible fluorescent pro		Anguilla japonica (Japanese eel)	139
Zebrafish (8)		F4IVL6	GRV2_ARATH	<b>.</b>	DnaJ homolog subfamily C GRV2	GRV2 GFS2,KAM2,At2q26890,F12C20.7	Arabidopsis thaliana (Mouse-ear cress)	2,554
S. cerevisiae (7) Human (6)		F4JY12	GFS12_ARATH		Protein GFS12	GF812 BCHD,At5g18525,T28N17.10	Arabidopsis thaliana (Mouse-ear cress)	1,639
Mouse (6)		Q9U6Y4	GFPL2_ZOASP		GFP-like fluorescent chromoprotein		Zoanthus sp. (Green polyp)	231
Other organisms		Q9GZ28	NFCP_ANESU		GFP-like non- fluorescent chromoprot		Anemonia sulcata (Mediterranean snakelocks sea anemone)	232
Subcellular		D7PM05	D7PM05_CLYGR		Green-fluorescent protein 10	GFP2 GFP10,GFP12,GFP25,GFP4,GFP45	Clytia gregaria (Phialidium gregarium)	235
location Multi-pass membrane		Q9U6Y8	RFP_DISSP		Red fluorescent protein drFP583		Discosoma sp. (Sea anemone)	225
protein (474) Perinheral membrane		Q8ISF8	RFP_ENTQU		Red fluorescent	A Et Condal	Entacmaea quadricolor (Bubble-tip anemone)	231

# Selectati "protein name" in sectiunea Filter "green" as

нари 🕒 экороз растом - Экор 🤹	ciogna		and a stope w	1110120			conditioned and the second sec	an boundador cur
Filter by	×	Columns	SBLAST E	lign	🕹 Download 🗎 🔒 Ac	ld to basket	◀ 1 to 25 of 66 ►	Show 25 •
Reviewed (66) *					Protein names 🖨		•	
Popular	-		Entry name 🔻		»	Gene names 🛡	Organism 🛡	Length <b>=</b> ,
organisms		P42212	GFP_AEQVI	<mark>ک</mark>	Green fluorescent protein	GFP	Aequorea victoria (Jellyfish)	238
S. cerevisiae (7)		P0DM59	UNAG_ANGJA	÷	Bilirubin-inducible fluorescent pro		Anguilla japonica (Japanese eel)	139
Human (6) Mouse (6) Zebrafish (2)		F4IVL6	GRV2_ARATH	<mark>.</mark> ∕₽	DnaJ homolog subfamily C GRV2	GRV2, GFS2, KAM2, At2g26890, F12C20.7	Arabidopsis thaliana (Mouse-ear cress)	2,554
Other organisms		Q9U6Y4	GFPL2_ZOASP	A	GFP-like fluorescent chromoprotein		Zoanthus sp. (Green polyp)	231
Search terms Filter "fluorescent" as: protein name (11)		Q9GZ28	NFCP_ANESU	<mark>.</mark> ∕₽	GFP-like non- fluorescent chromoprot		Anemonia sulcata (Mediterranean snakelocks sea anemone) (Anemonia viridis)	232
Filter "protein" as: gene ontology (35)		Q9U6Y8	RFP_DISSP	÷	Red fluorescent protein drFP583		Discosoma sp. (Sea anemone)	225
keyword (27)		Q8ISF8	RFP_ENTQU	<b>☆</b> `	Red fluorescent protein eqFP611		Entacmaea quadricolor (Bubble-tip anemone) (Parasicyonis actinostoloides)	231
protein family (7) protein name (38)		Q9U6Y6	GFPL_ANEMA	<b>☆</b>	GFP-like fluorescent chromoprotein		Anemonia manjano (Sea anemone)	229
Filter "green" as: author (13)		Q9U6Y3	GFPL_CLASP	<b>∳</b>	GFP-like fluorescent chromoprotain		Clavularia sp. (Brown star polyp)	266
organism (2)					chromoprotein			

#### Se obtine:

Child Call     Call Call     Call Call     <	(Answerdt) magnen                  (Answerdt)                 (Answerdt)                (Answerdt)                 (Answerdt)                 (Answerdt)                 (Answerdt)                 (Answerdt)                 (Answerdt)                 (Answerdt)	1 1272	-	_					NUMBER OF THE OWNER	ALCON STR
Lati Algan Restructed Demonsport In Process Results The advect of algan Restruction State Stat	n determend to many and the case of the second to be set of the second to be s	niProt	Un	ProtKB -	name:green fluoresc	cent pr	rotein AND reviewed yes		Adianos	s ← Q Search
Imported results	Constraint         Constra	AST Align Retrieve/ID mappir	ng					Arrive Mark		Help Cont
Bit by:     Istorical discretional discretiona discretional discretional discretion	6 (1) ***********************************	niProtKB result	s						About UniProt	KB 🏦 Bask
Image: Second	City         Entry         Entry <the< th=""><th>er bv</th><th><b>%</b> BI</th><th>LAST I</th><th>Align 🛓 Downlos</th><th>ad 10</th><th>Add to basket 🖉 Columns 🝃</th><th></th><th>1 to 4 of 4</th><th>Shov 25</th></the<>	er bv	<b>%</b> BI	LAST I	Align 🛓 Downlos	ad 10	Add to basket 🖉 Columns 🝃		1 to 4 of 4	Shov 25
text <ul> <li>P 4222</li> <li>P P MCP</li> <li>P MCP</li></ul>	Magnetic Section         P 4222         OP _ M (VI)         Section         Gene Management on Value         Gene Management on Value         Accurace without (Multiplum)         C230           Control         P 42022         OP _ M (VI)         Section         Magnetic Malana (Moures ear own)         230           P 42022         P F312_0         P 512_0         Magnetic Malana (Moures ear own)         2,554           P 10 cellion         P 0000         P 0000         P 0000         P 0000         Academin on Magnetic Malana (Moures ear own)         1,393           P 10 cellion         P 0000         P 0000         P 00000         P 000000         P 000000         P 000000         P 000000         P 000000         P 0000000         P 00000000         P 000000000000000000000000000000000000	Reviewed (4)		ntry 🗘	Entry name 🌲		Protein names 🗘 🛛 🕅	Gene names 🌲	Organism \$	Length 🌲
Vill of Organisms     P 45VL     RV2_ARTH     Point Demoking and family C GRV2     GRV2 052,XAV2_AR322690,F12C2C.7     Anabidepiis thalians (dourse-arr over)       VII (s)     P 50/F2     GR32_ARATH     Points GF312     GF312 64D0.Af5218255,728H17.10     Anabidepiis thalians (dourse-arr over)       VII (s)     P 00M59     UM0_MAMA     Points GF312     GF312 64D0.Af5218255,728H17.10     Anabidepiis thalians (dourse-arr over)       VII (s)     P 00M59     UM0_MAMA     Points GF312     GF312 64D0.Af5218255,728H17.10     Anabidepiis thalians (dourse-arr over)       Urg as mathrane patient entername patient	g2n1sms         r         rUL         eV2_ARATH         p         bnankogsubfandigt C GRV2         GRV2 GF2_XAM2_A22_22690_F12C2.07         Arabidopsit thaliana (Moure-ear cost)         2,054           r         FV2_L2         GF31_ARATH         p         Potana CF512         GF31_Z BC10_A52_91252_T72BN17.10         Arabidopsit thaliana (Moure-ear cost)         1,459           r         D00M59         UM4_ABABA         p         Bindoin inducible floorescont pro         Arabidopsit thaliana (Moure-ear cost)         319           r         Do1ma         Bindoin inducible floorescont pro         I to 4 of 4         bin 20           r <td>r-Prot</td> <td>•</td> <td>42212</td> <td>GFP_AEQVI</td> <td>e.</td> <td>Green fluorescent protein</td> <td>GFP</td> <td>Aequorea victoria (Jellyfish)</td> <td>238</td>	r-Prot	•	42212	GFP_AEQVI	e.	Green fluorescent protein	GFP	Aequorea victoria (Jellyfish)	238
vg (g) P5/T2 0/S12_AAATH N Poken G512 GF512 640-A/S26252,728171.710 Asabidepsit shallana (Moure-ear oess) 36 (g) P60M59 UNAQ_AMDA N M Minishinduchle fluorescent pro Angulla japonic (japonese ear) 1949 Ambana poken an itorage vacalle an itorage vacalle an itorage vacalle 1970 Ambana Minishinduchle fluorescent pro P00M59 UNAQ_M10A M MINISHING (Japonese ear) P00	Image: Preving and Preving and Preving CFS12         CFS12 BCH0_A55g1525_7.72817.30         Anabidopsis thaliana (Mouse-ear-ossis)         1.439           Including approximation of the preving and preving approximation of the pre	ular organisms naliana (2)	O F	4TVL6	GRV2_ARATH		DnaJ homolog subfamily C GRV2	GRV2 GFS2,KAM2,At2g26890,F12C20.7	Arabidopsis thaliana (Mouse-ear cress)	2,554
A(1)       P00M9       UMAG_ANGAN       Minubin-inducible fluorescent pro       Anguillo japonica (Japanesa eel)         cellular location       pages membranes protein       1 to 4 of 4         sharal membranes protein       1 to 4 of 4         in norace service       1 to 4 of 4         repart membranes protein       1 to 4 of 4         in norace service       1 to 4 of 4         in norace service       1 to 4 of 4	Image: Probability in probability i	(1) IV	BF	4JY12	GFS12_ARATH	5	Protein GFS12	GFS12 BCHD,At5g18525,T28N17.10	Arabidopsis thaliana (Mouse-ear cress)	1,639
ellular incentional enter international enternational entern enternational enternational enternatio	ribotion     interview       nembrane protein     it to 4 of 4       seque vacuula     membrane       membrane     it to 4 of 4       y' as:     g       op (3)     by (2)	A (1)		ODM59	UNAG_ANGJA	154	Bilirubin-inducible fluorescent pro		Anguilla japonica (Japanese eel)	139
rpadar Amerikanan porkan Amerikanan Amerikana ain storage vacuale die pasar amerikana die maser metrik terms geneer aas an amerikana	herebrane protein age vestode membrane MB art art art art art art art art art art	cellular location								
ain storage vacuale life pass marbrane off mich terms "gener" at an nome	aga vacuda membrane TS *** * ** yr (3)	nheral membrane protein							1 to 4 of 4	Shov 25
ain storage vacuale ain critegraam manhormane ain critegraam aan graam aan	age vacuale membrane S age v* age og (3) by (2)									
lar pas i montolano and i terms forean as a	memorane TS (* ατ σr (3) by (2)	tein storage vacuole								
srch terms *green* #s: ahn name # *protein* as:	ms **** **** \$7 (3)	gle-pass membrane tein								
"green" s: In nome ★ "protein" s:	**** ** *** 97 (3)	irch terms								
"protein" as:	97 (3) 19 (2)	"green" as:								
ain (1)	97 (3) hy (2)	"protein" as:								
ontology (3)	η (2)	ontology (3)								
vord (3)	ily (2)	vord (3)								
sin family (2)		ein family (2)								
<u>/ by</u>										

Se selecteaza codul **P42212** (corespunde proteinei GFP extrasa din Aequorea victoria (Jellyfish))

(Codul P42212 se afla si in pagina dedicata proteinei 1EMA in baza de date PDB!). Verificati unde se afla acest cod!!!

Studiati informatiile cuprinse in pagina de sumar a proteinei GFP

Sequence	features	View only features	(star, domains, PTHs)		
	Status	Residued - Ar	notation score: ***** - Experimental exidence at protein law	¢	
isplay	Pare	S BLACT	E Format Windd to bashat @matury	🕈 Faedbath 🛛 Halp video	Other tutorials and videor
funding .		Function			
Artist & Tarahamy		Energy-transfer a receiving energy (	cospton. Its role is to transduce the blue chemiluminescence of $\theta$ rom the Ca <sup>be</sup> -activated photoprotein acquorin.	e protein aequivin into green fluorescent light by energy transfer.	Pluoresces in vivo upon
attorings & Bristock TM / Proceeding		Absorption <sup>4</sup> Abs(max)=295 m Exhibits a smalle	n absorbance peak at 470 nm. The fluorescence emission spectru	m peaks at 509 nm with a shoulder at 540 nm.	
		GO - Biological pro	cess" In # Same contribution - Description contributions and energy # lower laborations		
anny & Don sinc		<ul> <li>protein-dware</li> </ul>	ophare trik age # Scare, Instantion)		
		Complete GO and Reywords - Molec Photoprotein	ular function <sup>5</sup>		
		Keywords - Bielog Luminescence	feat process <sup>1</sup>		
		Keywords - Ligan Chranophore	·		
Somifiar professor		Names & Ta	xonomy		
		Protein names*	Recommended neme: Green Recement protein		
		Oerie names <sup>4</sup>	Nama: GFP		
		Crganism <sup>4</sup>	Aeguorea victoria (Jallyfish)		
		Texonomic Identifier <sup>4</sup>	eroo [scst]		
		Taxonomic Inreage <sup>4</sup>	Eukaryota + Metazoa + Gnidaria + Hedrozoa + Hydroidolina + Lap	othecata + Aeguoreidae + Aeguorea 🕅	
		Pathology &	Biotech		

Folosind codul "P42212" se poate ajunge direct la aceasta pagina (fara pasii anteriori)!

#### ▶ Pentru a studia gena proteinei GFP selectati "*Cross-references*" (meniul din stanga)

⊢ → C 🗋 www.uni	iprot.org/Uniprot/P-	42212				Q 😭
Sequence feature	es View only features	(sites, domains, PTMs	)			
State	us 🛛 🐴 Reviewed - Ar	notation score:	<ul> <li>Experimental evidence at protein la</li> </ul>	veř		
Display and	SBLAST = Alig	n 🗟 Format 📾 Add t	o basket OHistory		🏶 Feedback 💷 Help video	Other tutorials and side
C Function	Cross-refere	inces				
Names & Taxonomy	Web resources <sup>1</sup>					
	Protein Section					
Pathology & Biotech	The greenest of	us all - Issue 11 of Jun	. 2001			
PTM / Processing	Protein Spotlight					
Expression	Paint my thought	ts - Issue 100 of Augus	4 2009			
Indepartment	Green fluorescen	t protein entry				
Structure	Protein Spotlight					
amily & Domaina	Paint my thought	ts - Issue 100 of Nover	nber 2007			
lequence	Sequence databas	ies				
Trospiteferences	Select the link	M62654 mRNA, Transl	ation: AAA27722.1.			
Publications	destinations:	M62653 mRNA. Transl	ation: AAA27721.1.			
Entry information	* EMEL*	L29345 mRNA, Transli X26418 mRNA, Transli	ation: AAA59246.1. ation: CAA65279.1.			
Miscellaneous	0 CCBJ <sup>4</sup>	U73901 Genomic DNA	Translation: AAB18957.1.			
Similar profeito	PRI	J80692. JQ1514.				
op	3D structure datab	10100				
	Select the link					
	destinations:	Entry Met	nod Resolution (A)	Chain	Positions	PD8sum
	# PCDe <sup>4</sup>	189C X-ra	y 2.40	A/B/C/D	1-238	[*]
	© POB/	1048	2.10	2	1-220	[1]
		1047 X-78	2.50	2	1.220	(*)
		1EMA X-ra	v 1.90	A	1-230	[2]
		1EMB X-ra	2.19	A	1-200	[*]
		1EMC X-ra	y 2.30	A/8/C/D	2-237	[*]
		3EME X-ra	y 2.50	A	2-237	[*]
		1EMF X-ra	2.40		2-238	[2]

La capitolul "Sequence databases" se selecteaza GenBank, apoi se activeaza linkul "M62654":

GenBank ≑	M62654 mRNA. Translation: AAA27722.1.
	M62653 mRNA. Translation: AAA27721.1.
	L29345 mRNA. Translation: AAA58246.1.
	X96418 mRNA, Translation; CAA65278.1.
	U73901 Genomic DNA. Translation: AAB18957.1.

Pagina "http://www.ncbi.nlm.nih.gov/nuccore/M62654" afiseaza structura pre-ARNm pentru molecula GFP (ARN mesager precursor - inainte de splicing si translatie)!

Display Sett	ings: ⊡ GenBank	Send: 🕑		_
			Change region shown	
Aequoi	rea victoria green-fluorescent protein mRNA, comple	te		
ds			Customize view	
SenBank: N	462654.1		Customize view	
ASTA Gra	aphics		Basic Features     Default features     Gene, RNA, and CDS features only	
<u>io to:</u> 🕑			Features added by NCBI	
ocus	AEVOFPB 5170 bp n#WA linear INV 26-APR-1993		I conserved domain	
FINITION	Acquorea victoria green-fluorescent protein mPNA, complete cds.		Display options	
CCESSION	<b>M62654</b>		Show sequence	
ERSION	262654.1 GI:155662		U Show reverse complement	
EYWORDS	acquorin-associated energy-transfer acceptor; bioluminescence;		Update View	W
OFFICE	green-fluorescent protein; photocyte.			
ORGANTEN	leguorea victoria			_
	Eukaryota; Netazoa; Cnidaria; Nydrozoa; Nydroida; Leptomedusae;		Analyze this sequence	
	Acquoreidae; Acquorea.		Run BLAST	
EFERENCE	1 (bases 1 to 5170)		Pick Primers	
AUTHORS	Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G. and			
TITLE	Primary structure of the leguorea wictoria green-fluorescent		Highlight Sequence Features	
	protein		Find in this Sequence	
JOURNAL	Gene 111 (2), 229-233 (1992)			
PUBMED	1347277			
OMMENT	Original source text: Aequorea victoria (library: lambda gt10) cDNA		Related information	
	TO WERE		Delated Company	

Secventa pre-ARNm obtinuta contine atat zone purtatoare de informatie genetica (exoni) cat si zone nepurtatoare de informatie genetica (introni). exon:



intron:

	/product="green-fluorescent protein"	
	/protein id=" <u>A&amp;A27722.1</u> "	
	/db_xref="GI:155663"	
	/translation="MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL	
	KFICTTGKLPVPUPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFYKD	
	DGNYKSRAEVKFEGDTLVNRIELEGIDFKEDGNILGHKMEYNYNSHWVYINADKQKNG	
	IKVNFKIRNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDH#	
d and a second	ILLEVITAGOILGONELTK"	
Incron	141910 //ana#/020#	
	/ generation of P	
exon	946.1240	
	/gene="GFP"	
	/number=3	
intron	12412307	
	/gene="GFp"	
	/number=3	
exon	2308>2744	
	/gene="GFP"	
TOTAL	/number=4	
1 angettenne		
f1 atgaatgatg	stangstage telecanaty manyabasa anglytagit tanganatatit	
121 ttggaaaatt	ageocompa cancearanto geogramanto a consectat tatanaccom	
181 ttottatatt	ttacaggat accaagatg agtaaaggag aagaacttt cactggagt	
241 gtcccaattc	ttgttgaatt agatggtgat gttaatggge acaaattete tgteagtgga	
301 gagggtgaag	gtgatgcaac atacggaaaa cttaccctta aatttatttg cactactgga	
361 aagctacctg	ttccatggcc aacastigts astastist stiatggigt isa <mark>gtaagig</mark>	
421 catttatac	tettttaata teagtgitaa gaaaateaag igtetigeta tittttegat	
481 tattggtgca	attetagtea aattattgeg tttttttace caaaatgtta atgtaaaact	
541 gaaatttggc	acacttgogo aaatatatac agggtatttt gaaaaaatta aacaggatga	
601 taaaagttgc	acagaaactt atctcaagat ttacccgcag aaagatgett naaaaattga	
661 tatttgacag	ageaaaactt gagatteaeg tettitagit gittgaettg aaattitggt	
721 gacaggtagg	tatetgaaa aacaaacaaa acgtaaaaat atcacgtgat taaagtgtat	
941 totgoottto	agamacagte testamete contatte attergenat atmendatog	
901 cascattere	tigagitati igaagitata degacetate ateagaatti eaagetata	
961 cccadatcat	Automatical algorithms and algorithms an	
1021 ggaaagaagt	atattitaa aagatgacgi gaactacaaa toocgigcig aactoaacti //mber=2	
1081 tgaaggtgat	accotcgtta atagaattga gttaaaaggt attgattta aagaagatgg	

Se poate ajunge direct la informatia anterioara prin acesarea paginii http://www.ncbi.nlm.nih.gov/

se selecteaza baza de date "*Nucleotide*"
se introduce codul "M62654"

3 NCBI Resources ⊡ How	To 🖸	Sign in to NCBI
SNCBI National Center for Biotechnology	Nucleotide • M62654	Search
NCBI Home	Welcome to NCBI	Popular Resources
Resource List (A-Z)	The National Center for Rietechnology Information advances science and health by providing	PubMed
All Resources	access to biomedical and genomic information.	Bookshelf
Chemicals & Bioassays		PubMod Contral
Data & Software	About the NCBI   Mission   Organization   Research   NCBI News	Fubmed Central
DNA & RNA		PubMed Health
Domains & Structures		BLAST
Genes & Expression	Get Started	Nucleotide
Genetics & Medicine	Tools: Analyze data using NCBI software	Genome
Genomes & Maps	<ul> <li><u>Downloads</u>: Get NCBI data or software</li> </ul>	SNP
Homology	How Tos: Learn how to accomplish specific tasks at NCBI	Gene
Literature	<ul> <li><u>Submissions</u>: Submit data to GenBank or other NCBI databases</li> </ul>	Gene
Proteins		Protein
Sequence Analysis		PubChem
Taxonomy	Genotypes and Phenotypes	
Training & Tutorials	Data from Genome Wide Association	NCBI Appoundements
Variation	tiset tudy variables, protocols, and analysis. 1 1 2 3 4 5 8 7 8	Next NCBI webinar on November 5th Oct 23, 2014

Pentru a evidentia secventa ADN corespunzatoare moleculei ARNm ("matrita" ce serveste ca model pentru translatie) se activeaza "mRNA".

FEATURES	Location/Qualifiers
source	15170
	/organism="Aequorea victoria"
	/mol_type="mRNA"
	/db_xref="taxon: <u>6100</u> "
	/tissue_lib="lambda gt10"
gene	1>5170
	/gene="GFP"
intron	<1197
	/gene="GFP"
	/number=1
<u>misc_feature</u>	193201
	/gene="GFP"
	<pre>/experiment="experimental evidence, no additional details</pre>
	recorded"
	/note="fluorescent chromophore"
mRNA	join(198413,9461240,2308>2744)
	/gene="GFP"
exon	198413
	/gene="GFP"
	/number=2
CDS	join(208413,9461240,23082523)
	/gene="GFP"

Obs:

- secventele ADN obtinute prin activarea "*mRNA*" difera de secventele ADN obtinute prin activarea "*CDS*" (CoDing Sequence), deoarece secventele CDS contin doar regiuni definite de codonii start si stop.

- secventele "CDS" incep cu gruparea de nucleotide "ATG" si se termina cu una din gruparile "TGA", "TAA" sau "TAG".

Pentru a vedea secventa completa a ADN trebuie vizualizata intreaga pagina. Copiati cele 3 zone "mRNA" intr-un fisier word pentru a obtine secventa ADN ceruta!

Se obtine un sir de caractere care pe langa literele ce indica bazele, contine si cifre, care nu fac parte din secventa ADN corespunzatoare moleculei ARNm:

## Deci nu asa se obtine ceea ce dorim!!!!! Daca nu asa atunci cum?

2) **Translatia secventei ADN** (*obtinerea secventei proteice pornind de la ADN*)

Pentru a obtine secventa ADN corespunzatoare proteinei GFP se selecteaza "FASTA" (dreapta jos sau stanga sus)

Intr-o fereastra noua (new "Tab") deschideti pagina: http://web.expasy.org/translate/

Selectati secventa ADN (nu linia de comanda!) din pagina obtinuta prin activarea comenzii FASTA si copiati-o: "Copy" sau "Ctrl+c"



Puneti secventa ADN ("Paste" sau "Ctrl+v") in pagina: http://web.expasy.org/translate/

In sectiunea "Output format" selectati "Includes Nucleotide Sequences"

Botati      Songia preview - Scopi.     Plante Madda - Planes, F.     Red cools     Control Delaborate: Cal.     Control Delaborate: Cal.     Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence.     Please enter a DNA or RNA sequence in the box below (numbers and blanks are ignored).     Cantact Mathematication and the control of th	← → C fi 🗋 web.expasy.org/translate/					\$	2045 🔊	Ξ
Image:	🗄 Aplicații   🔘 Scopus preview - Scop 🔹 Blogilates — Fitness, F 🛚 🔀 Google	🛞 Analize Medicale - Col	🧿 Live Streaming Video /	📎 Veria   Healthy Recipe	🔤 2048  😁 Cub de viespi	👷 Comori Delicioase: Cui		*
Translate tool Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence. Please enter a DNA or RNA sequence in the box below (numbers and blanks are ignored). GATALACAMAGRAMAGATAMOTTACTOGATOTOCCANTCHTOTOTALTAGATOT CANTERTINATION CANTERTINATION CONTENTS AND	SIB CONTRACTOR Resource Portal		Translate			Home	Contact	
Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence. Please enter a DNA or RNA sequence in the box below (numbers and blanks are ignored). GATAGAMASTAACTAHAGAMACTITICACTOGOATTOTOCOATTOTTOTTOTATATAGATOT	Translate tool							
Please enter a DNA or RNA sequence in the box below (numbers and blanks are ignored). GATAACAAWATAGATAMAGAAMAACTITICACTGOATTGETCCAATTCHTUTGAATTAGATGAT GATATAATGGAAMAGATCAATTGAATGAGAGAGATGATGAATTAGAATGAATTGAATTAGATGA	Translate is a tool which allows the translation	of a nucleotide (D	NA/RNA) sequen	ce to a protein se	equence.			
CATACAMAGINAL TRANSMERSIONAL AND THE CARE TO THE CARE TO THE CARE AND	Please enter a DNA or RNA sequence in the bo	x below (numbers	and blanks are ig	nored).				
TTALATTERTITATTECACTACTOGALAGCTACCTOTECATOGCCALACCTOTECACTACTTECTITATOG TETCAATECTITICAGATACCAACTACTATAGAACAGACACTOTECACTACTTECTITATOG GOTTATECTACAGOALAGAACTATATAGATACAGATAGATACAACTACAGTACTOALAGCTACA TETAAGTOTATACTATACATTATAATCACAGTACTATAGATTAGAT	GATAACAAAGATGAGTAAAGGAGAAGAACTTTTTCACTGGAGTTGTCCCCAAT GATGTTAATGGGCACAAATTCTCTGTCAGTGGAGAGGGGTGAAGGTGATGCA	TCTTGTTGAATTAGATGGT ACATACGGAAAACTTACCC						
боттатьскае обламовают атактитите самызатов собовают така самото ставаято самото само	TTAAATTTATTTGCACTACTGGAAAGCTACCTGTTCCATGGCCAACACTTG TGTTCAATGCTTTTCAAGATACCCAGATCATATGAAACAGCATGACTTTT	TCACTACTTTCTCTTATGG CAAGAGTGCCATGCCCGAA						
Показото на Local стала на коакт на коакт на коакт на коакта на к	GGTTATGTACAGGAAAGAACTATATTTTACAAAGATGACGGGAACTACAAA	TCACGTGCTGAAGTCAAGT						
атсамаюттаюсттсамаяттаюсьсяхсяттоямовтодаютсяхствосаюссаттатсаюс малиатсастаяттовосаятовостотостистаковая салоститестоское салоствоест ттссамаялекосаятовостотоститаская салоститестоское салоствоест словозатовсямовся таксаятальнаятотсовает салостаетовое салоста словозатовсямовся таксаятальнаятотсовое словатовое салоствое салоста такамается патаготальнаятос салостите салагтая стальают такамается патаговоетальнаятовоета салостаетоваета салоствое салостаетов такамается патаговоета салостаетоваета салостаетовое салостаетоваета салостаетовоета такамается патаговаета салостая салостаетоваета салостаетоваета салостаетоваета салостаетоваета такамается патаговоета салостаетоваета салостаетоваета салостаетоваета салостаетоваета такамается патаговаета салостаетоваета салостаетоваетоваета салостаетоваетов такамается патаговаета салостаетоваета салостаетоваетоваета салостаетоваетоваетоваетоваетоваетоваетоваето	TIGARGOTGATACCCCCCCCTTARTAGERTIGAGTTARAGGTRTIGATTTIA	AGACAAACAAAAGAATGGA						
алал на стора и подока подока со подока са стала со подокала со подок ПОДИ ПОДИ ПОДИ ПОДИ СО ПОД	atcaaagttaacttcaaaattagacacaacattgaagatggaagcgttcaa	CTAGCAGACCATTATCAAC						
Ситовсятоватованствия: Анальятотоскаяствосаясты Аналотосованскаятаяс такаятостовоятовитовоговой ситаятовоставой ситаяствой такаятов такаятов такаятов такаятов такаята спасовой соответся и соответся соответся Соответся соответся соответс Соответся соответся соо	AAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACC TTCCAAAGATCCCAACGAAAAGAGAGATCACATGATCCTTCTTGAGTTTGT	TGTCCACACAATCTGCCCT AACAGCTGCTGGGATTACA						
TAMATECEAGGETECTGOSTTANTTCAGGETGAGATATTATTATATATTATAATTCATTAAAATT TATGATATATTATTATTATTATATATATAGGGETATTTCAGATGAGTCATTCC TAATTCATATTATTCAATTTGATTGATTGAGTCATTCAATAGGGECATCCGCATCC Output format: [includes nucleotide sequence	CATGGCATGGATGAACTATACAAATAAATGTCCAGACTTCCAATTGACACT	AAAGTGTCCGAACAATTAC						
ALIGATIATTIATTATTATTATTATTOATTIGATTIGATTIG	TAAAATCTCAGGGTTCCTGGTTAAATTCAGGCTGAGATATTATTTAT	TTATAGATTCATTAAAATT						
Output format: Includes nucleotide sequence	TAATTCTATATTAATTACAATTTGATTTGACTTGCTCA	5501201652616021100						
Output format: [Includes nucleotide sequence •			6					
	Output format: Includes nucleotide sequence	•						
Genetic code: Standard	Genetic code: Standard		•					

Click "Translate Sequences"

Se obtin *6 secvente diferite* ce provin din <u>schimbarea cadrului de citire</u> (incepe cu prima, a doua sau a treia nucleotida) si a <u>directiei de citire</u> (5'3' sau 3'5').

#### Numai una din aceste secvente este cea buna!

De obicei secventa buna (directia si cadru de citire bun) este cea care produce **cea mai lunga translatie**. (secventa neintrerupta de aminoacizi, fara codoni "Stop" in interiorul secventei proteice)

#### $\rightarrow$ Studiati cele 6 secvente si incercati sa gasit secventa potrivita!

Directia 5'3' cadrul 2: Codonul Start (ATG) este destul de aproape de inceput, iar primul codon Stop (TAG) se afla aproape de sfarsit!

				의 있는	 =
🗄 Apicații 🕘 Scopus preview - Scop 👒 Bioglates — Fitness, F 🚺 Google 🚫 Analze Medicale - Col	Elve Streaming Video /	💫 Veria   Healthy Recipe	2018 😁 Cub de viespi	👷 Comori Delicioase: Cul	
5'3' Frame 2					
patasonasgatgagtasggasgtasggasgttttooctggagttgtoocastbettgtgagtl L $\odot$ K H $\otimes$ K $\otimes$ B $\otimes$ L $P$ T $\otimes$ V $V$ P L $S$ V $\otimes$ P L $S$ V $V$ P L $S$ V V $V$ P L $S$ V $V$ P V V V V V V V V V V V V V V V V V V					

Click pe link-ul "5'3' Frame 2 " In pagina obtinuta, este evidentiata metionina ( $M \rightarrow codon ATG$ ):

ITK MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL KFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEG YVQERTIFYKDDGNYKSRAEVKFEGDTLVNRIELKGIDFKEDGNILG HKMEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQ NTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMILLEFVTAAGITH GMDELYKStopMSRLPIDTKVSEQLLKSQGSWLNSGStopDIIYIFIDSLK FYEStopFIDVINRGYFLIKStopATGVHSStopFYINYNLIStopLA

Click pe primul "M": Se obtine secventa de aminoacizi din care este compusa proteina "Green Fluorescence Protein"



Protein" (1EMA) din baza de date **pdb.org** si activati "*Sequence*". Ar trebui sa obtineti secventa de aminoacizi urmatoare:

SCOP	Green f	uorescent pro	tein, GFP (d1er	naa.)		_	Ł		
DSSP	~~ <u> </u>	^		<u> </u>			1		
PDB MS KG	EELFTGVVPIL	VELDGDVNG	<u>HKFSVSGEG</u>	EGDATYGKLT	LKFICTTGK	LPVPWPTL	÷		
100 2	10	20	30	40	50	60			
SCOP	Green fl	Jorescent prot	ein, GFP (d1em	aa_)		_	¢.		
DSSP -		~~~~	$\sim$	<u> </u>		^			
PDBVTTF	XVQCFSRYPDH	MKRHDFFKS	AMPEGYVQE	RTIFFKDDGN	YKTRAEVKF	EGDTLVNR	ž.		
PDB 61	70	80	90	100	110	120			
SCOP	Green fl	iorescent prot	ein, GFP (d1em	aa_)		_	ŧ		
DSSP		<u> </u>		$\sim$	<u> </u>	-			
PDB IELK	GIDFKEDGNIL	GHKLEYNYN	SHNVYIMAD	KQKNGIKVNF	K I RHN I EDG	SVQLADHY	ŕ		
PDB 123	130	140	150	160	170	150	F		
SCOP C	reen fluorescent	protein, GFP (d	1emaa)						
DSSP		_		_	<u> </u>				
RDR OON T									
PDB 183	190	200	210	220	AAUT INUMD	LIK			
DSSP Legend									
T: tum									

# Comparati cele 2 secvente de aminoacizi?

## 3) Compararea secventei de aminoacizi obtinuta cu secvente similare din baza de date PDB

In pagina rezultata dupa translatarea ADN in secventa de aminoacizi se selecteaza "*Fasta Format*" (scris cu albastru!). Se obtine:

Aplicații (	Scopus preview - Scop 🔹 Bioglates — Pitness, P	Google 🔞 Analize Medicale - Col	() Live Streaming Video J	📀 Veria   Healthy Recipe	🔤 2048 🛛 😁 Cuib de viespi	Comori Delicioase: Cui		
<b>*</b> • 🔏	EXPASY storicmatol Pesource Pontal		Translate			Home	Contact	
ID VI AC VI DB Tr DB on CC -i DR SM SQ SB M3 VT NR NR HY // Sequen	RE24501 Unreviewed; RE24501; analation of nucleotide sequence g (27-oct-2014 by 108.24.22.252; - This virtual protain sequence wi from the server after a few days machine server after a few days network of the server after a few days from the server after a few days network of the server after a sequence wi from the server after a few days network of the sequence with a sequence of the sequence of the sequence of the server after a sequence with a sequence of the sequence of the sequence of the sequence of the sequence of the sequence of the ce in FASTA format	238 AA. enerated on ExpAsy ill automatically be d CRC64. BROATY INCOMPETATION IN ADDONG INVERTION IN ADDONG INVERTION IN	eleted LPVPMPTL KRKODIV DGRVQLAD RDKLYK					
>VIRT24	501							
MSKGEEL VTTFSYG NRIELKG HYQQNTP	PTGVUPILVELDGDVNGHKPSVBGEGEGEGEATYGK VQCPSRVDDHHKCHDPPKEAAKDBGVVQERTIFYR VQCPSRVDDHHKCHDPPKEAAKDBGVVQERTIFYR	LTLKFICTTGKLPVDWPTL DDGNYKSRAEVKFEGDTLV KVNFKIRHNIEDGSVQLAD EFVTAAGITHGMDELYK						
BLAST B	LAST submission on ExPASy/SIB	Sequence a	analysis tools: ProtPa	aram, ProtScale, Co	mpute pl/Mw, Pepti	deMass, PeptideC	Cutter,	
prosite s	canProsite	Direct Subn	nission to SWISS-MC	DDEL				

Comparati aceasta secventa de aminoacizi cu secvente proteice din PDB. Daca gasiti secvente similare, inseamna ca un cercetator a rezolvat structura 3D a acestei proteine! Copiati secventa proteica obtinuta in format Fasta.

Deschideti pagina http://www.pdb.org/ si click "Advanced Search" Din meniul "Chose a Query Type" selectati "Sequence (Blast/Fasta)" In fereastra "Sequence" puneti secventa copiata. Click "Submit Querry"

Scopus preview -	Scop 👒 Bloglates — Fitness, F 🚦 Google 🔞 Analize Medicale - Col 💿 Live Streaming Video / 🔗 Veria   Healthy Recipe 🔤 2048 😁 Cub d	le viespi 🛛 🔛 Comori Delicioase: Cul
Search Advanced Browse	Everything Author Macromolecule Sequence Ligand @	
anced Search Inte	dace	
Sequence (BLAST)	(ASTA/PSFBLAST) • 0	
quence search (BL/	ST or FASTA)	
Structure Id		Result Count
Chain Id		
Sequence	MISINGECTUS/VII/LICE/LICE/UNIV/VII/LICE/LICE/LICE/LICE/LICE/LICE/LICE/LI	*
Search Tool	BLAST V	
Mask Low Complexity	Yes •	
E-Value Cutoff	10.0	
Sequence Identity Cutoff (%)	0	
		Add Search Criteria 🖸

Pagina rezultata contine o lista de proteine a caror secventa de aminoacizi este apropiata de secventa data.

Similaritatile se pot observa activand "Display Full Alignment" in sectiunea "Alignment"

	opus preview - Scop	🔹 Bloglates — Fitness, F. 🔣 Google 🔞 Analize Medicale - Col 🐵 Live Streaming Video / 👶 Veria   Healthy Recipe 🚞 2048 📼 Cub d	le viespi 🔬 Comori Deli	cioase: Cui
2.04	Chain(s):	A.B.C.D		
	Authors:	Van There, J.J., & Generatiny, G.Y., & Terrerine, M. P., Sanne, J.T., P		
100	Release:	2005-07-20		
然論	Experiment:	X-RAY DIFFRACTION with resolution of 1.85 Å	Residue Count	944
	Compound: 1	olymer [ Display Full Polymur Dutails   Display for All Results ]		
	Alignment: Le	ngth: 238 E-value: 1.12503E-138 Score: 489.574bits (1259) Identities: 233/238 (98%) Positives: 238/238 (100%) Gaps: 0/2	238 (0%)	
	77 (A Citation: U Valecule of G Nelecule of G	10 10 10 10 10 10 10 10 10 10 10 10 10 1	IIO IZO II III IKSARVKEBOTUMBI IKSTARVKEBOTUMBI IKSTARVKEBOTUMBI IIO IZO 007 - Insertion	LING 1 1 100 1 1 LING DYPEONILAR ELECTOPYPEONILAR ELECTOPYPEONILAR 1 1 1 130 1
e 4ANJ	MYOSIN VI (	MDinsert2-GFP fusion) PRE-POWERSTROKE STATE (MG.ADP.AlF4)		
≥ 4ANJ ≛ ≋ ∎	MYOSIN VI ( Chain(s):	MDInsert2-GFP fusion) PRE-POWERSTROKE STATE (MG.ADP.AlF4) A		
≥ 4ANJ 2 8 4	MYOSIN VI ( Chain(s): Authors:	MDInsert2-GFP fusion) PRE-POWERTIROKE STATE (MG.ADP.AIF4) A Manatry, J.A. Lubet, T.S. Tapar, Y.S. Maharja, N.S. Pilyanin, O.S. Lu, X.S. Para, J.S. Valatta, P.S. Summy, RLS. Hudans, AJ.S		
<ul> <li>4ANJ</li> <li>≜ ≜</li> </ul>	MYOSIN VI (I Chain(s): Authors: Release:	MDinsert2-GFP fusion) PRE-POWERSTROKE STATE (MG.ADP.AlF4) A Heastray, J.S., Stalet, T.S., Rapen, Y.S., Habbrin, H.S., Pelyanda, G.S., K.K., Penn, J.S., Varbetta, P.S., Steenary, H.J., S., Headana, A.H.S. 2021210-17		
AANJ	MYOSIN VI ( Chain(s): Authors: Release: Experiment:	NEInsent2-OFP fusion) PRE-POWERSTROKE STATE (MC.ADP.ABF4) Ansenty J.S. Bahet, T.D. Rapan, Y.J. Yadaraja, K.J. Pajyeeke, G.J. Liu, K.J. Pare, J.J. Varlette, P.J. Sommer, H.J. S. Huedure, A.S.S. 2021210-27 X-RAY DIPRACTION with resolution of 2:40 Å	Residue Count	1201
AANJ	MYOSIN VI ( Chain(s): Authors: Release: Experiment: Compound:	NDINSORT2-OEP Fusion) PRE-POWERETROKE STATE (MG.ADP.ABF4) A mannary J.S. Isabet, T.P. Rapen, Y.S. Hadherjan, H.S. Pejapenko, G.S. Lin, K.F. Paren, J.F. Vanhette, P.J. Stewaney, H.L.P. Headane, A.M.P. 2012;10:17 - KART DIPPROCENTION with resolution of 2:40 Å 2 Polyment Simpler full Engineer baselin (Simpler for All Basedon J 4 Updath (Simpler Vall Engineer baselin).	Residue Count	1201
	MYOSIN VI ( Chain(s): Authors: Release: Experiment: Compound: Alignment:	MDInsert2- GPF fusion) PRE-POWERETROKE BTATE (MG.ADP.ABF4) A Masslvey, J.G., Tudet, T.D., Rupur, Y.D. Yakhariya, N.D. Polyansho, O.D. Liu, K.D., Parez, J.D., Yakatta, P.D., Summary, H.L.D., Haudana, A.H.D. 2022-10-12 X-RAY DIPRACTION with resolution of 2.60 Å 2 Polymeter ( <i>Display full Dyado Basisis</i> ). ( <i>Display for All Naculus</i> J 4 Ugotta'. ( <i>Display full Dyado Basisis</i> ). ( <i>Display for All Naculus</i> J 4 Ugotta'. ( <i>Display full Dyado Basisis</i> ). ( <i>Display for All Naculus</i> J 4 Ugotta'. ( <i>Display full Dyado Basisis</i> ). ( <i>Display for All Naculus</i> J	Residue Count	1201
± anj ± a ∎	MYOSIN VI (i Chain(s): Authors: Release: Experiment: Compound: Alignment: Citation:	NDInsert2-OFP fusion) PRE-POWERSTROKE STATE (MC.ADP.AIF-4) Ansates Li, Bubet Li), Rupan, Yu), Haharjan, KU, Pajyanaha, G, Lik, KU, Pana, Li), Yahasha, PU, Simeney, NLU, Haharia, A.K.P 2023/0-07 X-RAY DIPRACTION with resultation of 2-00 Å 2 Polymetry [Display Foll System Statish [Display for All Ansates ] 4 Opends [Engine] Foll System Statish [Display for All Ansates ] Procession Statish in the Riverse Direction Result: Display for All Ansates ] Procession Statish in the Riverse Direction Result: Display for All Ansates ] Procession Statish in the Riverse Direction Result [Display for All Ansate]	Residue Count	1201

Se obtin proteine similare cu GFP si forme mutante!

# → Studiati proteina "1HCJ"

- numai 5 aminoacizi diferiti
- modificarile sunt conservative: aminoacizii din cele 2 forme sunt similari
- ex: schimbare de la izoleucina la valina, ambele hidrofobice

₫ 1HCJ ≧ 🖻 🐨 😡	PHOTOPRODUCT	OF THE WILL	D-TYPE AD	QUOREA V	ICTORIA	GREEN FL	UORESCE	INT PROTE	EN				
	Chains	A, B, C, D											
	E Value	7.6E-136											
	Characteristics	Resolution:	1.80 Å	1001 Exp.	Method: 3	CHONY DUPP	RACTION						
	Classification	Luminescen	t Protein										
	Compound	Molecule: 4	GREEN FLUX	DRESCENT	PROTEIN								
		Polymer: 1 Chains: 7 Mutation: 1	1 A, B, C, D /ES	Түр	s: bojābet	ticke(L)							
ଭୁଭ	Authors Alignment State	Van Thor, 3. Length: 238	J.P. Gen Score: 47	ach, T. A. 9.559bits (	Hollingwa 1233) E-s	orf, K.J.P. alus: 7.63	Johnson 06E-136	n, L.P Edentifies:	233/238-(9	0%) Positi	ven: 238/2	238 (100%)	Gape:
775 - 201		51	60	10	80	90	1.0.0	110	12.0	1.00	140	150	140
						· ·				I -			
		CTURE POPPER	CAN'T TRUE DURING	CREATING NE	DESITIVE 2 ADER	COLUMN T 1	C REDUCTION OF	BARYPERSOT.	AND DESCRIPTION OF	PROCEED LOPE	CALC & TRUE &	NET DIALINGRO	STRVEPOLT
	Alignment	CTURLEVEWER	LATTE DOTO	CETEMPORNE	NUMBER OF TAXABLE	CONTRACTOR	TESSONES	NAMES DOT:	ATTRACTOR	PRESS NUMBER OF COMPANY	C LINING	NTINALIQUE	III/II/IIIII
	Pinginiteite												
		1.1	80	10	80	90	1.00	110	120	1.30	140	1.00	1.60
					_								

## 4) Vizualizarea mutatiilor genetice

Cautati in pagina de pe site-ul PDB la care ati ajuns in partea precedenta codul **1BFP**. (o proteina mutanta creata de cercetatori, care are emisia de fluorescenta albastru in loc de verde)

→ Determinati diferenta dintre structurile primareale celor doua proteine.

In tabelul urmator sunt indicate mutatiile care duc la modificarea emisiei de fluorescenta a proteinei GFP

Green Fluorescent	No mutation
Yellow Fluorescent	S65G, S72A, T203F
Cyan Fluorescent	Y66W
Blue Fluorescent	Y66H, Y145F

"S65G" inseamna ca in pozitia 65 serina a fost inlocuita cu glicina → Identificati mutatiile Y66H si Y145F (in structural alignment)

Pentru vizualizarea mutatiilor ce apar in 1BFP se va folosi "Protein workshop" selectati "Visibility"

selectati "Atoms and Bonds" selectati aminoacidul "145" selectati aminoacidul "66"

Mutatii

- tirozina (Y) este inlocuita de histidina (H) in pozitia 66,
- tirozina (Y) este inlocuita de fenilalanina (F) in pozitia 145

In pozitia 66 era situat cromoforul

proteinei "Green fluorescent"

 $\rightarrow$  Comparati cu proteina 1 EMA!



#### Tema 8:

- 1. Evidentiati mutatiile ce apar la proteinele: 1GFL, 1EMG
- 2. Indicati primele 12 nucleotide ale secventei codante ARNm pentru GFP.
- 3. Indicati o referinta pentru articolul in care s-a raportat secventa codanta a GFP.
- 4. Explorati moleculele din pagina interactiva Molecular Machinery poster (http://mm.rcsb.org/).

Click pe molecula care de interes activeaza vizualizarea interactive.