Find-a-Gene Project

[Q1]

Protein name: Tumor protein p53

Species: Homo sapiens

Accession number: NP_000537.3

Function known: Encodes tumor suppressor protein; regulates expression of target genes by

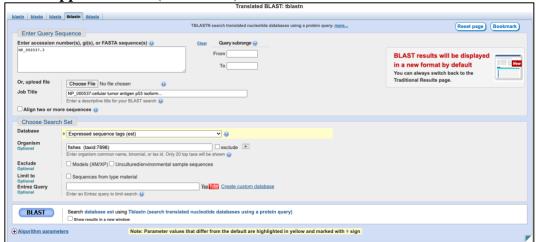
inducing apoptosis, DNA repair, or changes in metabolism

[Q2]

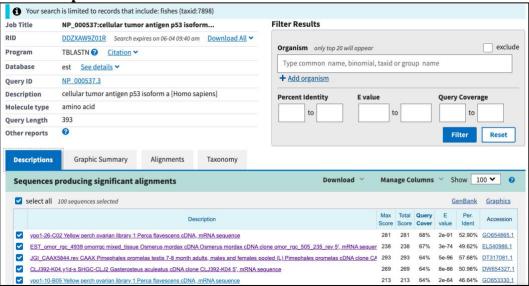
BLAST method: TBLASTN

Database searched: Expressed sequence tags (est)

Limits applied: Fishes (taxid: 7898)



Search output list:



Alignment of choice: a cDNA clone from *Osmerus mordax* (rainbow smelt), with accession number EL540986.1.

E-value and other alignment stats: E-value = 3e-74, score = 238, query coverage = 67%, percent identity = 49.62%.

```
EST_omor_rgc_4939 omorrgc mixed_tissue Osmerus mordax cDNA Osmerus mordax cDNA clone omor_rgc_505_235_rev 5', mRNA sequence
 Sequence ID: EL540986.1 Length: 907 Number of Matches: 1
 Range 1: 184 to 906 GenBank Graphics
                                                                                                                                                                                                                                                                                         ▼ Next Match ▲ Previous Match

        Score
        Expect
        Method
        Identities
        Positives
        Gaps
        Fram

        238 bits(607)
        3e-74
        Compositional matrix adjust.
        134/266(50%)
        171/266(64%)
        26/266(9%)
        +1

 Query 1 MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP 60
M+EP + S+ SQE+F LW N+++ P++ + LM +P E+ E+P
Sbjct 184 MDEPVGNESIT--FSQESFHYLW-----NLMAGQPTETENCLMEAPLFDEEDLFEEP-- 333
 Query 61 DEaprmpeaappvapapaaptpaapapapSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK 120
AP + L+5+VPS Y GS F L F S TAK
Sbjct 334 -----TIPAPQAPCHSSDSGSSLTSTVPSTTDYPGSLNFSLKFQQSSTAK 468
 Query 121 SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE 180
SVTCTYSP LNK+FCQLAKTCPVQH VH PPPG +RA+A+YK+S+ + +VVRRCPHHE 548
SDJCt 469 SVTCTYSPELNKLFCQLAKTCPVQMVVEQAPPPGALLRAVAVYKKSKDVADVVRRCPHHE 648
Query 181 RCS-DSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTT1HYNYMCN 239
RT ++++G AP HL+RVECN + YL D +T RHSV+VPYEPPVGSD TT+ HNYMCN 500 CT + RTENNEOPPAGALDRVGENDAMTCHSVLVPYEPPQVCSDITTVLFNAWD 500 CT - RTENNEOPPAGALDRVGENDAMTCHSVLVPYEPPQUCSDITTVLFNAWD 500 CT - RTENNEOPPAGALDRV
 Query 240 SSCMGGMNRRPILTIITLEDSSGNLL 265
 SSCMGGMN RPILTIITLE G LL
Sbjct 829 SSCMGGMNPRPILTIITLETQEGQLL 906
```

EsT_omor_rgc_4939 omorrgc mixed_tissue Osmerus mordax cDNA Osmerus mordax cDNA clone omor_rgc_505_235_rev 5', mRNA sequence
Sequence ID: EL540986.1|Length: 907|Number of Matches: 1|Range 1: 184 to 906

Alignment statistics for match #1

Score		Expec t	Method			Identities	Positives	Gaps	Fram e
238 bits(6	07)	3e-74	Composition adjust.	onal matri	Х	134/266(50%)	171/266(64%)	26/266(9%)	+1
Query	1	MEEPQSDI	PSVEPPLSQET	'FSDLWKLLP	ENNVLSP	LPSQAMDDLMLS	PDDIEQWFTEDP	GP 60	
		M+EP +	S+ SQE+	F LW	N+++	P++ + LM +	P E+ E+P		
Sbjct	184	MDEPVGNI	ESITFSQES	SFHYLW	NLMAG	QPTETENCLMEA	PLFDEEDLFEEP	333	
Query	61	DEaprmpe	eaappvapapa	aptpaapap	apSWPLS	SSVPSQKTYQGS	YGFRLGFLHSGT	AK 120	
				AP +	L+	S+VPS Y GS	FLF ST	AK	
Sbjct	334		TIE	PAPQAPCHSS	DSGSSLT	STVPSTTDYPGS	LNFSLKFQQSST	AK 468	
0	101								
Query	121		_	_		_	QHMTEVVRRCPHI		
Sbjct	469						+ + +VVRRCPHI KDVADVVRRCPHI		
30)00	409	50101151	г пликль Сбль	MCICE VQMVV.	LQAFFFG.	ALLINAVAVINNS	NDVADV VNNCEIII	HE 040	
Query	181	RCS-DSD0	GLAPPOHLIRV	EGNLRVEYL	DDRNTFR	HSVVVPYEPPEV	GSDCTTIHYNYM	CN 239	
~ 1		R + +++					GSD TT+ +NYM		
Sbjct	649	RTTENNE	GPAPAGHLVRV	EGNNQARYL	ADNHTGR	HSVLVPYEPPQV	GSDITTVLFNYM	CN 828	
Query	240	SSCMGGM	NRRPILTIITI	EDSSGNLL	265				
		SSCMGGM	N RPILTIITI	E G LL					
Sbjct	829	SSCMGGM	NPRPILTIITI	ETQEGQLL	906				

Protein sequence of choice matches subject above:

> Osmerus mordax protein

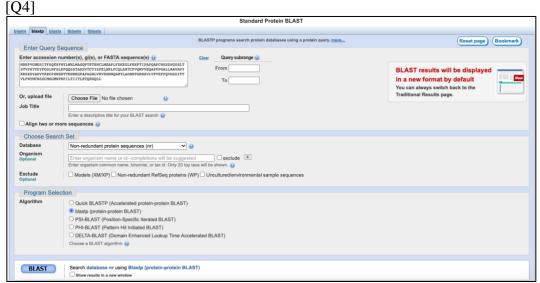
 ${\tt MDEPVGNESITFSQESFHYLWNLMAGQPTETENCLMEAPLFDEEDLFEEPTIPAPQAPCHSSDSGSSLTSTVPSTTDYPGSLNFSLKFQQSSTAKSVTCTYSPELNKLFCQLAKTCPVQMVVEQAPPPGALLRAVAVYKKSKDVADVVRRCPHHERTTENNEGPAPAGHLVRVEGNNQARYLADNHTGRHSVLVPYEPPQVGSDITTVLFNYMCNSSCMGGMNPRPILTIITLETQEGQLL}$

Name: Osmerus mordax protein

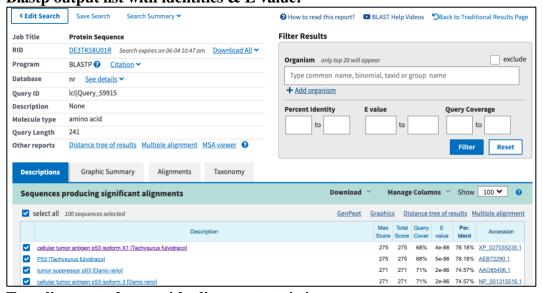
Species: Osmerus mordax

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii;

Neopterygii; Teleostei; Stomiatii; Osmeriformes; Osmeridae; Osmerus.



Blastp output list with identities & E value:



Top alignment shown with alignment statistics:

Results indicates a "novel" gene found: The percent identity of the top hit is not 100%, so my protein should be "novel."

[Q5]

Trimmed and labeled sequences for alignment:

> Human p53

 ${\tt TYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHERCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLL$

> Osmerus mordax protein

DYPGSLNFSLKFQQSSTAKSVTCTYSPELNKLFCQLAKTCPVQMVVEQAPPPGALLRAVAVYKKSKDVADVVRRCPHHERTTENNE GPAPAGHLVRVEGNNQARYLADNHTGRHSVLVPYEPPQVGSDITTVLFNYMCNSSCMGGMNPRPILTIITLETQEGQLL

>Yellow_catfish XP_027035235.1 cellular tumor antigen p53 isoform X1 [Tachysurus fulvidraco]

DYPGLHNFTLHFQKSSTAKSVTCTYSPELNKLFCQLAKTCPVLMAVSFSPPHGSVLRATAVYKRSEHVADVVRRCPHHERSNDNNE GPAPPGHLLRVEGNSRAVYHEDLNTQRHSVVVPYEPPQVGSECTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGQLL

>Zebrafish AAO85406.1 tumor suppressor p53 [Danio rerio] DYPGDHGFRLRFPQSGTAKSVTCTYSPDLNKLFCQLAKTCPVQMVVDVAPPQGSVVRATAIYKKSEHVAEVVRRCPHHERTPDGDN LAPAGHLIRVEGNQRANYREDNITLRHSVFVPYEAPQLGAEWTTVLLNYMCNSSCMGGMNRRPILTIITLETQEGQLL

>Brown_bullhead QGQ62196.1:80-244 tumor protein 53 [Ameiurus nebulosus]
DYPGLHDFTLHFQESSTAKSVTCTYSPGLNKLFCQLAKTCPVLMAVSSSPPPGSVLRATAVYKRSEHVAEVVRRCPHHERSNDSSD
GPAPPGHLLRVEGNSRAVYQEDGNTQRHSVVVPYEPPQVGSQCTTVLYNYMCNSSCMGGMNRRPILTIITLETQNGHLL

>Wuchang_bream QBM00765.1:59-230 tumor antigen p53 [Megalobrama amblycephala] DYPGEHGFKLQFPQSGTAKSVTCTYSPDLNKLFCQLAKTCPVQMVVDIAPPQGSLLRATAIYKKSEHVAEVVRRCPHHERTPDTDG LAPAAHLIRVEGNLRATYKEDDSTSRHSVVVPYEAPQLGAGFTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGQML

>Channel_catfish NP_001187005.1:77-241 cellular tumor antigen p53 [Ictalurus punctatus]

 ${\tt DYPGLLNFTLHFQESSGTKSVTCTYSPDLNKLFCQLAKTCPVLMAVSSSPPPGSVLRATAVYKRSEHVAEVVRRCPHHERSNDSSDGPAPPGHLLRVEGNSRAVYQEDGNTQAHSVVVPYEPPQVGSQSTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGHLL}$

>Grass_carp ANP93608.1:59-230 tumor antigen p53 [Ctenopharyngodon idella] DYPGEHGFKLQFPQSGTAKSVTCTYSPDLNKLFCQLAKTCPVQMVVDVAPPQGSVLRATAIYKKSEHVAEVVRRCPHHERTPDTDG LAPAAHLIRVEGNLRATYKEDDVTSRHSVVVPYEAPQLGAGFTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGQML

>West_African_lungfish QC084573.1:180-355 tumor protein suppressor 53 [Protopterus annectens]

 $\verb|DYPGECSFQLVFQESGTAKSVTYTYSPTLNKLYCQLAKTCPVQIKVESTPEVGAIIRATAVYKKSEHVADVVKRCPHHERSTEHKD|\\ GLVPPSHLIRVEGNSQAQYLEDSNTRRQSVIVPYEKPQVGSDCTTVLYNYMCNSSCMGGMNRRPILTIITLESKDGKLL|\\$

> Rainbow trout NP 001118164.1:7-254 cellular tumor antigen p53 [Oncorhynchus mykiss]

DYPGALGFQLRFLQSSTAKSVTCTYSPDLNKLFCQLAKTCPVQIVVDHPPPPGAVVRALAIYKKLSDVADVVRRCPHHQSTSENNE GPAPRGHLVRVEGNQRSEYMEDGNTLRHSVLVPYEPPQVGSECTTVLYNFMCNSSCMGGMNRRPILTIITLETQEGQLL

CLUSTAL multiple sequence alignment by MUSCLE (3.8):

Human p53 West African lungfish Zebrafish Wuchang bream Grass carp Yellow catfish Brown bullhead Channel catfish Osmerus mordax Rainbow trout

TYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMA DYPGECSFOLVFOESGTAKSVTYTYSPTLNKLYCOLAKTCPVOIKVESTPEVGAIIRATA DYPGDHGFRLRFPQSGTAKSVTCTYSPDLNKLFCQLAKTCPVQMVVDVAPPQGSVVRATA DYPGEHGFKLQFPQSGTAKSVTCTYSPDLNKLFCQLAKTCPVQMVVDIAPPQGSLLRATA DYPGEHGFKLQFPQSGTAKSVTCTYSPDLNKLFCQLAKTCPVQMVVDVAPPQGSVLRATA DYPGLHNFTLHFQKSSTAKSVTCTYSPELNKLFCQLAKTCPVLMAVSFSPPHGSVLRATA DYPGLHDFTLHFQESSTAKSVTCTYSPGLNKLFCQLAKTCPVLMAVSSSPPPGSVLRATA DYPGLLNFTLHFQESSGTKSVTCTYSPDLNKLFCQLAKTCPVLMAVSSSPPPGSVLRATA DYPGSLNFSLKFQQSSTAKSVTCTYSPELNKLFCQLAKTCPVQMVVEQAPPPGALLRAVA DYPGALGFOLRFLOSSTAKSVTCTYSPDLNKLFCOLAKTCPVOIVVDHPPPPGAVVRALA * * .* * * * . :**** *** ******** : *. .* *: :** *

Human p53 Wuchang bream Grass_carp
Yellow_catfish Grass carp Brown bullhead Channel_catfish Osmerus mordax Rainbow trout

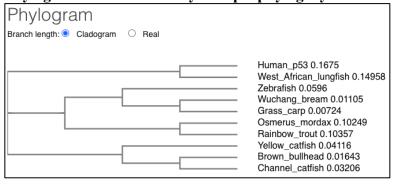
IYKQSQHMTEVVRRCPHHERCSD-SDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPY IYKKSEHVAEVVRRCPHHERTPD-TDGLAPAAHLIRVEGNLRATYKEDDSTSRHSVVVPY IYKKSEHVAEVVRRCPHHERTPD-TDGLAPAAHLIRVEGNLRATYKEDDVTSRHSVVVPY VYKRSEHVADVVRRCPHHERSNDNNEGPAPPGHLLRVEGNSRAVYHEDLNTQRHSVVVPY VYKRSEHVAEVVRRCPHHERSNDSSDGPAPPGHLLRVEGNSRAVYQEDGNTQRHSVVVPY VYKRSEHVAEVVRRCPHHERSNDSSDGPAPPGHLLRVEGNSRAVYQEDGNTQAHSVVVPY VYKKSKDVADVVRRCPHHERTTENNEGPAPAGHLVRVEGNNQARYLADNHTGRHSVLVPY IYKKLSDVADVVRRCPHHQSTSENNEGPAPRGHLVRVEGNQRSEYMEDGNTLRHSVLVPY

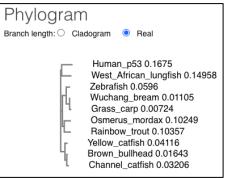
Human p53 Human_pus
West_African_lungfish Zebrafish Wuchang bream Grass carp Yellow catfish Brown bullhead Channel catfish Osmerus mordax Rainbow trout

EPPEVGSDCTTIHYNYMCNSSCMGGMNRRPILTIITLEDSSGNLL EKPQVGSDCTTVLYNYMCNSSCMGGMNRRPILTIITLESKDGKLL EAPQLGAEWTTVLLNYMCNSSCMGGMNRRPILTIITLETQEGQLL EAPQLGAGFTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGQML EAPQLGAGFTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGQML EPPQVGSECTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGQLL EPPQVGSQCTTVLYNYMCNSSCMGGMNRRPILTIITLETQNGHLL EPPQVGSQSTTVLYNYMCNSSCMGGMNRRPILTIITLETQDGHLL EPPQVGSDITTVLFNYMCNSSCMGGMNPRPILTIITLETQEGQLL EPPQVGSECTTVLYNFMCNSSCMGGMNRRPILTIITLETQEGQLL

[Q6]

Phylogenetic tree created by "simple phylogeny" from EBI:

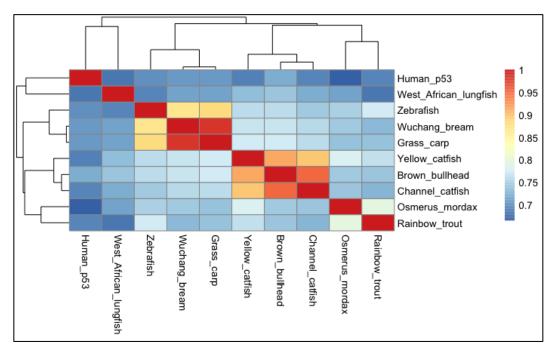




[Q7]

Sequence identity based heatmap:



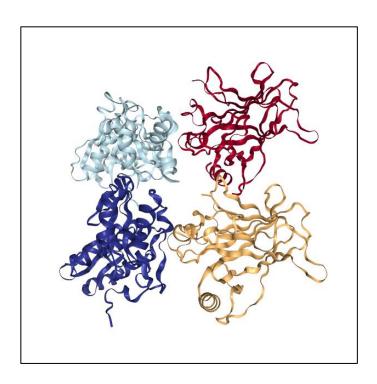


[Q8] **PDB** identifiers from multiple species:

Structure ID	Technique	Resolution	Source	E value	Identity
1KZY_A	x-ray	2.5	Ното	3e-99	71.43%
	diffraction		sapiens		
1HU8_A	x-ray	2.7	Mus	3e-90	70.81%
	diffraction		musculus		
4D1L_A	x-ray	1.97	Danio rerio	2e-15	66.67%
	diffraction				

[Q9]

The structure of Chain A, which is colored by red in the figure, can be similar to the *Osmerus mordax* protein studied in this project, given there is a high percent of identity.



[Q10]

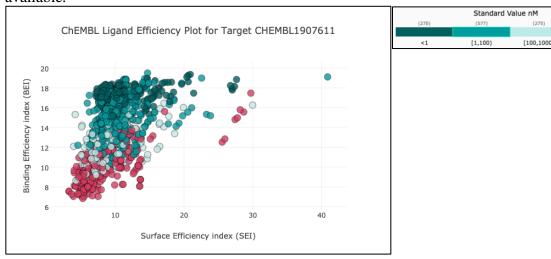
CHEMBL search:

The top hit (CHEMBL3885544) showed a binding assay and no functional assay. There is no ligand efficiency data found.

The binding assay (CHEMBL3778681) tested peptidyl-prolyl cis-trans isomerase D and observed an "inhibition of Flag-tagged p53/GST-tagged PPID (unknown origin) expressed in Escherichia coli," suggesting it inhibition effect on p53.

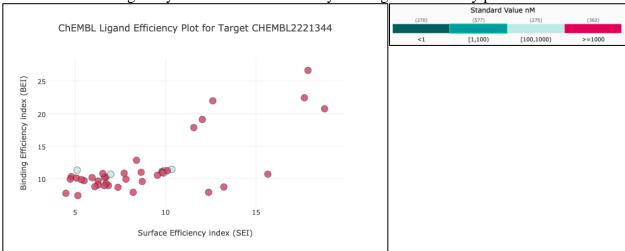
https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL3885544/ https://www.ebi.ac.uk/chembl/assay_report_card/CHEMBL3778681/

The second hit (CHEMBL1907611) is a protein-protein interaction between p53 and oncoprotein MDM2. There are 147 binding assays and 1 functional assay. The ligand efficiency plot is also available.



https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL1907611/

The third hit (CHEMBL2221344) is a protein-protein interaction between p53 and MDM4. There are 10 binding assays and no functional assay. The ligand efficiency plot is available.



https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL2221344/