

Find-A-Gene Project Assignment

Name: Steven Gan
PID: A59020397
E-mail: digan@ucsd.edu

Quarter: Fall
Course: BGGN-213
Instructor: Dr. Barry Grant

Q1:

Name: CCCTC-binding factor (CTCF)

Accession: isoform 1 isoform 2 isoform 3
 NP_006556.1 NP_001177951.1 NP_001350845.1

Species: *Homo Sapiens*

Q2:

Isoform 1 will be used for downstream analysis, as it is the longest isoform.

Query: NP_006556.1

Method: TBLASTN (2.13.0+) search against with default parameters

Database: Expressed Sequence Tags (est)

Organism: Papilionoidea (taxid:37572)

TBLASTN Setting:

BLAST® » tblastn

Home Recent Results Saved Strategies Help

Translated BLAST: tblastn

blastn blastp blastx **tblastn** tblastx

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) Clear

Query subrange [?](#)

From

To

Or, upload file No file chosen [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Database ?

Organism ? ☐ exclude [Add organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

☐ Models (XM/XP) ☐ Uncultured/environmental sample sequences

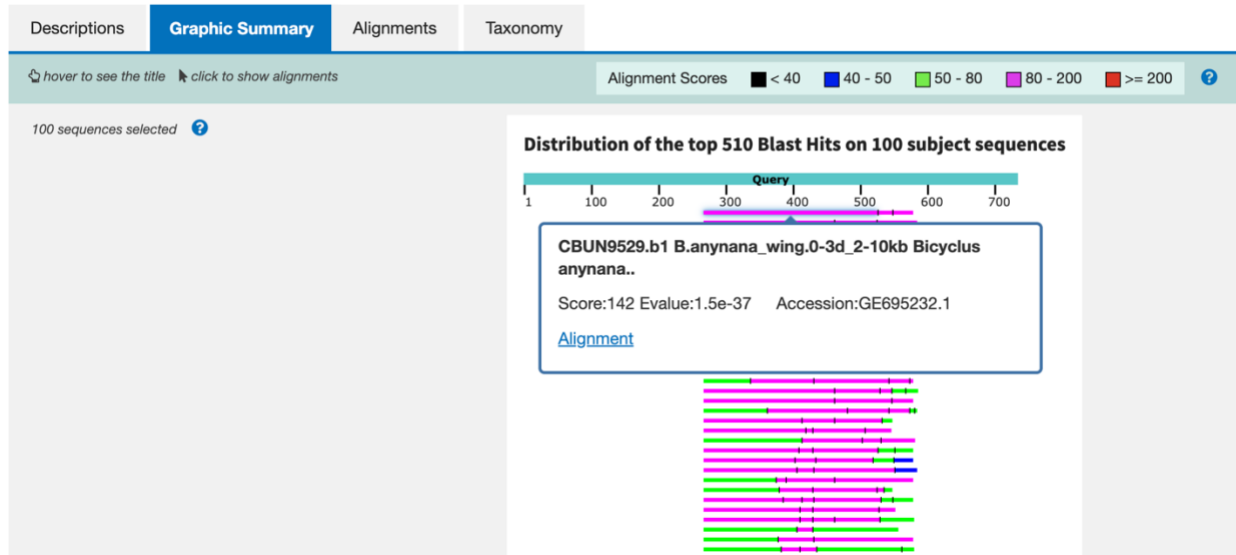
☐ Sequences from type material

Limit to [Create custom database](#)

Entrez Query [?](#)

FIND-A-GENE PROJECT ASSIGNMENT

Chosen Match: Accession GE695232.1, a 750 base pair clone from *Bicyclus anynana*. Alignment details see below.



Descriptions **Graphic Summary** **Alignments** Taxonomy

Alignment view Pairwise ? Restore defaults Download

100 sequences selected ?

Download GenBank Graphics Sort by: E value Next Previous Descriptions

CBUN9529.b1 B.anynana_wing.0-3d_2-10kb Bicyclus anynana cDNA clone CBUN9529, mRNA sequence
Sequence ID: [GE695232.1](#) Length: 750 Number of Matches: 3

Range 1: 3 to 734 GenBank Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
154 bits(389)	1e-41	Compositional matrix adjust.	95/284(33%)	144/284(50%)	47/284(16%)	+3
Query 267	QCELC	SYTC	PRRNLDRHMSHTDERPHKCHL----	CGRAFRTVTLLRNHLNTHGTGRPH		322
Sbjct 3	C++C	Y C +R	NL H++HT E+P+ C + C R R	LR+H+ THTG +P		170
Query 323	KCPDC	MAFVTS	GELVRHRRYKHTHEKPFKSCMCDYASVEVSKLRHRSHTGERPFQCS			382
Sbjct 171	C C+	LV H R	HT EKPF C +C+Y L H+++HTGE+PF C			347
Query 383	LCSYAS	RDITYKL	RHRMTHSGEKPYECYICHARFTQSGTMKMHILQKHTENVAKFHC			442
Sbjct 348	+C+Y +	L H+RTH	+GEKP+ C IC+ +F ICNKTGVKNSLVCHLR-THTGKPFCEICNYKF-----			449
Query 443	DTVIAR	KSDGLGVHL	RKQHSYIEQGK---CRYCDAVFHERYALIQHOKSHKNEKRFKCDQ			499
Sbjct 450	A K +L	H++ I G+K	C C+ + +L+ H ++H EK F C+ ----ALKHNLVNHMK-----			602
Query 500	CDYACR	QERHMM	HKRTHTEGKPYACSHCDKTFRQKQLDMHFK			543
Sbjct 603	C+Y ++R+++	H +THTG	KP++C C+ K L H + CNYKSARKRYLLNHMKTHTEGKPFSCDICNKTGIRNSLVRHMR			734

FIND-A-GENE PROJECT ASSIGNMENT

Range 2: 84 to 749 [GenBank](#) [Graphics](#) [▼ Next Match](#) [▲ Previous Match](#) [▲ First Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame	
142 bits(358)	1e-37	Compositional matrix adjust.	89/255(35%)	125/255(49%)	33/255(12%)	+3	
Query 266	FQCELC	SYTC	PPRRSNLDRHMKSH	DERPHK	LCGRA	FRTVTLLRNHLNTHGTRPHKCP 325	
Sbjct 84	YSCEMY	NYKC	ARKSRLRHHMTHTG	EKPF	SCGIC	NYKTGVKNSLVCHLRHTHTGEKPFCE 263	
Query 326	DCDMAF	VTSGEL	VRHRRYKHTHEK	PFKCS	MDYAS	VEVSKLKRHSHTGERPFQCSLCS 385	
Sbjct 264	ICNYK	FALKR	NLLNHMK-THTG	EKPF	SCGIC	NYKTGVKNSLVCHLRHTHTGEKPFCEICN 440	
Query 386	YASRD	TYKLKR	HMRTHSGEK	PYECY	ICHAR	FTQSGTMKMHILQKHTENNAKPHCDTV 445	
Sbjct 441	YKFA	LKHN	LVNHMKIHTG	EKPF	SCGIC	NYKTRVKNSLVSH-LRHTHTGE-KPFSCEICNYK 614	
Query 446	IARKS	DLGV	HLRKQHSYIE	QGKK	CRYC	DAVFHERYALIQHQS	HKNEKRFKCDQCDYACR 505
Sbjct 615	SARK	-----	RY-----	LLNH	MKTH	TGKPFSCDICNYKTG 704	
Query 506	QERH	MIMH	KRTH	TGTE	520		
Sbjct 705	IKNS	LV	RHMRI	HTGE	749		

Range 3: 6 to 728 [GenBank](#) [Graphics](#) [▼ Next Match](#) [▲ Previous Match](#) [▲ First Match](#)

Score	Expect	Method	Identities	Positives	Gaps	Frame
136 bits(343)	2e-35	Compositional matrix adjust.	87/250(35%)	126/250(50%)	9/250(3%)	+3
Query 324	CPDCMA	FVTSGELVRHRRYKHTHEKPFKCS	MDYASVEVSKLKRHSHTGERPFQCSL			383
Sbjct 6	C	DICHYKCAQGRGNLVCHIR-THTCEKPY	SCEMYNYKCARKSRLRHHMTHTTGKPFSCGI	S+L+ H+ +HTGE+PF C+		182
Query 384	CSYAS	RDYTKLKRHMRTSGEKPYECYICHARFTQ	SGSTMKMHILQKHTENNAKPHCD			443
Sbjct 183	CNYKT	GVKNSLVCHLRHTHTGEEKPFCCEICNYK	FALKRNLNLLNH-MKTHTGE-KPFSCEICN	+ H + + HT F C C +		356
Query 444	TVIAR	KSDLGVHLRKQHSYIEQGKKCRYCDA	VFHERYALIQHQS	HKNEKRFKCDQCDYA		503
Sbjct 357	YKTG	VKNSLVCHLRHTHTG---EKPFCEIC	NYKFKLNLLNHMKIHTGKPFSCICNYK	+K+ L+ H K+ F C C +		530
Query 504	CROER	HIMHKKRTHTEGKPYACSHCDKTRQKQL	DMHFKRYHDPNFVPAAFVCSKCGKT			563
Sbjct 531	TRV	KNSLVSHLRHTHTGEEKPFCCEICNYK	SARKRYLLNHMKTHTG----	EKPFSCDICNYK		698
Query 564	FTRR	NTMARH 573				
Sbjct 699	TGI	KNSLVHR 728				

FIND-A-GENE PROJECT ASSIGNMENT

Alignment Details:

CBUN9529.b1 B.anynana_wing.0-3d_2-10kb Bicyclus anynana cDNA clone CBUN9529, mRNA sequence

Sequence ID: [GE695232.1](#)

Length: 750

Number of Matches: 3

Range 1: 3 to 734

Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps	Frame
154 bits (389)	1e-41	Compositional matrix adjust.	95/284 (33%)	144/284 (50%)	47/284 (16%)	+3
Query 267	QCELCSYTCPRRSNLDHRMKSHTDERPHKCHL----CGRAFRTVTLLRNHLNTHTGTRPH					322
	C++C Y C +R NL H+++HT E+P+ C + C R R LR+H+ THTG +P					
Sbjct 3	SCDICHYKCAQRGNLVCHIRHTHTCEKPYSCEMYNYKCARKSR----LRHMTTHTGEKPF					170
Query 323	KCPDCDMAFVTSGELVRHRRYKHTHEKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCS					382
	C C+ LV H R HT EKPF C +C+Y L H+++HTGE+PF C					
Sbjct 171	SCGICNYKTGVKNSLVCHLR-THTGEKPFCEICNYKFALKRNLLNHMKHTHTGEKPFSCG					347
Query 383	LCSYASRDYKLRHMRTHSGEKPYECYICHTARFTQSGTMKMHILQKHTENVAKFHCPHC					442
	+C+Y + L H+RTH+GEKP+ C IC+ +F					
Sbjct 348	ICNYKTGVKNSLVCHLRHTHTGEKPFCEICNYKF-----					449
Query 443	DTVIARKSDLGVHLRKQHSYIEQGKK---CRYCDAVFHERYALIQHQKSHKNEKRFKCDQ					499
	A K +L H++ I G+K C C+ + +L+ H ++H EK F C+					
Sbjct 450	----ALKHNLVNHMK-----IHTGEKPFSCCEICNYKTRVKNSLVSHLRHTHTGEKPFSCCEI					602
Query 500	CDYACRQERHMIMHKRTHHTGEKPYACSHCDKTFRQKQLDMHFK					543
	C+Y ++R+++ H +THTGEKP++C C+ K L H +					
Sbjct 603	CNYKSARKRYLLNHMKHTHTGEKPFSCDICNYKTGIKNSLVRHMR					734

Range 2: 84 to 749

Alignment statistics for match #2

Score	Expect	Method	Identities	Positives	Gaps	Frame
142 bits (358)	1e-37	Compositional matrix adjust.	89/255 (35%)	125/255 (49%)	33/255 (12%)	+3
Query 266	FQCELCSYTCPRRSNLDHRMKSHTDERPHKCHLCGRAFRVTLLRNHLNTHTGTRPHKCP					325
	+ CE+ +Y C R+S L HM +HT E+P C +C L HL THTG +P C					
Sbjct 84	YSCEMYNYKCARKSRLRHMTTHTGEKPFSCGICNYKTGVKNSLVCHLRHTHTGEKPFCE					263
Query 326	DCDMAFVTSGELVRHRRYKHTHEKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSLCS					385
	C+ F L+ H + HT EKPF C +C+Y + + L H+R+HTGE+PF C +C+					
Sbjct 264	ICNYKFALKRNLLNHMK-THTGEKPFSCGICNYKTGVKNSLVCHLRHTHTGEKPFCEICN					440
Query 386	YASRDYKLRHMRTHSGEKPYECYICHTARFTQSGTMKMHILQKHTENVAKFHCPHCDTV					445

FIND-A-GENE PROJECT ASSIGNMENT

Sbjct	441	Y + L HM+ H+GEKP+ C IC+ + ++ H L+ HT F C C+	614
Query	446	IARKSDLGVHLRKQHSYIEQGKKCRYCDAVFHERYALIQHQKSHKNEKRFKCDQCDYACR	505
		ARK RY L+ H K+H EK F CD C+Y	
Sbjct	615	SARK-----RY-----LLNHMKHTHTGEKPFSCDICNYKTG	704
Query	506	QERHMIMHKRTHHTGE	520
		+ ++ H R HTGE	
Sbjct	705	IKNSLVRHMRHTHTGE	749

Range 3: 6 to 728

Alignment statistics for match #3

Score	Expect	Method	Identities	Positives	Gaps	Frame
136 bits (343)	2e-35	Compositional matrix adjust.	87/250 (35%)	126/250 (50%)	9/250 (3%)	+3
Query	324	CPDCDMAFVTSGELVRHRRYKHTHEKPFKCSMCDYASVEVSKLKRHIRSHTGERPFQCSL				383
		C C G LV H R HT EKP+ C M +Y S+L+ H+ +HTGE+PF C +				
Sbjct	6	CDICHYKCAQRGNLVCHIR-THTCEKPYSCEMYNYKCARKSRLRHMTTHTGEKPFSCGI				182
Query	384	CSYASRDYKLRHMRTHSGEKPYEYICHARFTQSGTMKMHILQKHTENVAKFHCPHCD				443
		C+Y + L H+RTH+GEKP+ C IC+ +F + H ++ HT F C C+				
Sbjct	183	CNYKTGVKNSLVCHLRTHHTGEKPFCEICNYKFALKRNLNLN-MKTHHTGE-KPFSCGICN				356
Query	444	TVIARKSDLGVHLRKQHSYIEQGKKCRYCDAVFHERYALIQHQKSHKNEKRFKCDQCDYA				503
		K+ L HLR E+ C C+ F ++ L+ H K H EK F C+ C+Y				
Sbjct	357	YKTGVKNSLVCHLRTHHTG--EKPFCEICNYKFALKHNLVNHMKIHTGEKPFSCDICNYK				530
Query	504	CRQERHMIMHKRTHHTGEKPYACSHCDKTFRQKQLLDMHFKRYHDPNFVPAAFVCSKCGKT				563
		R + ++ H RTHHTGEKP++C C+ +K+ L H K + F C C				
Sbjct	531	TRVKNSLVSHLRTHHTGEKPFSCDICNYKSARKRYLLNHMKTHHTG---EKPFSCDICNYK				698
Query	564	FTRRNTMARH				573
		+N++ RH				
Sbjct	699	TGIKNSLVRH				728

FIND-A-GENE PROJECT ASSIGNMENT

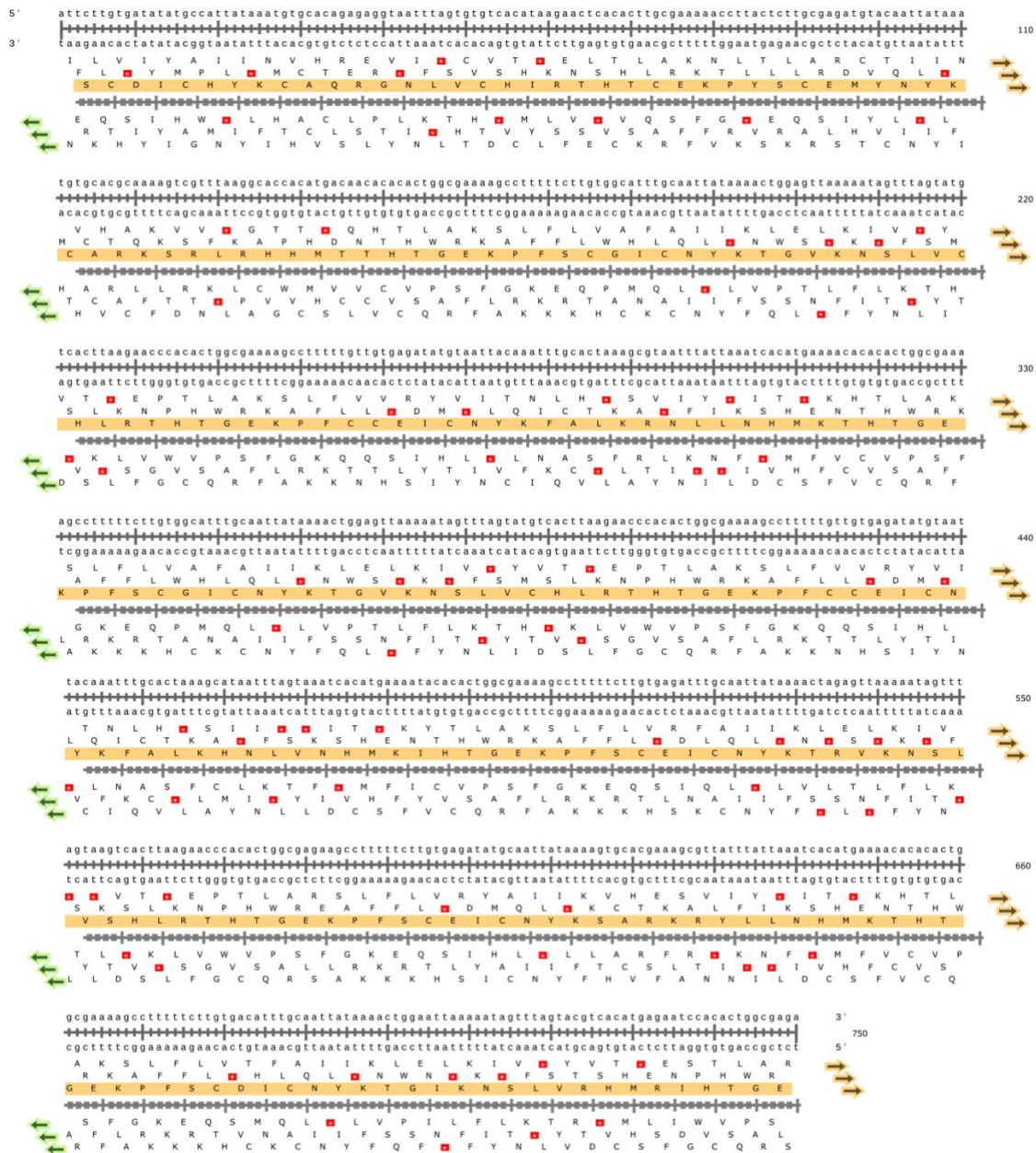
Q3:

Chosen sequence:

>B. anynana protein (from BLAST results)

SCDICHYKCAQRGNLVCHIRHTHTCEKPYSCEMYNYKCARKSRLRHHMTTHTGEEKPFSCGICNYKTGVKNSLVCHLRT
HTGEEKPFCEICNYKFALKRNLNLMKTHHTGEEKPFSCGICNYKTGVKNSLVCHLRTHTGEEKPFCEICNYKFALKHN
LVNHMKIHTGEEKPFCEICNYKTRVKNLSLVSHLRTHTGEEKPFCEICNYKSARKRYLLNLMKTHHTGEEKPFSCDICNY
KTGIKNSLVRHMRHTG

All six reading frame:



FIND-A-GENE PROJECT ASSIGNMENT

Name: *Bicyclus* CTCF-like protein

Species: *Bicyclus anynana*

Taxonomy: Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Papilionoidea; Nymphalidae; Satyrinae; Satyrini;
Mycalesina; *Bicyclus*.

Q4:

BLASTP search on non-redundant protein sequences (nr) hits top on zinc finger protein 84-like proteins on *Bicyclus anynana*, with identity percentage of 95.16%, suggesting a possible novel protein. See details below.

BLASTP setting:

The screenshot shows the NCBI BLASTP search interface. At the top, there are tabs for 'blastn', 'blastp' (selected), 'tblastn', and 'tblastx'. Below the tabs, the 'Enter Query Sequence' section contains a text box with the sequence: '>B. anynana protein (from BLAST results) SCDCIHYKCAQRGNLVCHIRHTTCEKPYSCEMYNYKCARKSLRHMMTTHTGE KPFSCGICNYKTGVKNSLVCHLRHTTGEKPFCEICNYKFALKRLLNMMKTHT GEKPFSCGICNYKTGVKNSLVCHLRHTTGEKPFCEICNYKFALKHNLVNHMM'. The 'Job Title' is 'B. anynana protein (from BLAST results)'. The 'Choose Search Set' section shows 'Standard databases (nr etc.):' selected, with a 'Try experimental clustered nr database' button. The 'Database' is set to 'Non-redundant protein sequences (nr)'. The 'Program Selection' section shows 'blastp (protein-protein BLAST)' selected.

blastn **blastp** tblastn tblastx

BLASTP programs search protein databases using a protein query. more... [Reset page](#) [Bookmark](#)

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#) [Query subrange](#) [?](#)

>B. anynana protein (from BLAST results)
SCDCIHYKCAQRGNLVCHIRHTTCEKPYSCEMYNYKCARKSLRHMMTTHTGE
KPFSCGICNYKTGVKNSLVCHLRHTTGEKPFCEICNYKFALKRLLNMMKTHT
GEKPFSCGICNYKTGVKNSLVCHLRHTTGEKPFCEICNYKFALKHNLVNHMM

From
To

Or, upload file No file chosen [?](#)

Job Title
Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Databases ☒ Standard databases (nr etc.): [New](#) ☐ Experimental databases [Try experimental clustered nr database](#) [?](#)
For more info see [What is clustered nr?](#)

Compare ☐ Select to compare standard and experimental database [?](#)

Standard

Database [?](#)

Organism ☐ exclude [Add organism](#)
Optional
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown. [?](#)

Exclude ☐ Models (XM/XP) ☐ Non-redundant RefSeq proteins (WP) ☐ Uncultured/environmental sample sequences
Optional

Program Selection

Algorithm ☐ Quick BLASTP (Accelerated protein-protein BLAST)
☒ blastp (protein-protein BLAST)
☐ PSI-BLAST (Position-Specific Iterated BLAST)
☐ PHI-BLAST (Pattern Hit Initiated BLAST)
☐ DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)
Choose a BLAST algorithm [?](#)

FIND-A-GENE PROJECT ASSIGNMENT

BLASTP results:

Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Select columns

Show

100

☒

select all

100 sequences selected

GenPept

Graphics

Distance tree of results

Multiple alignment

MSA Viewer

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> zinc finger protein 84-like [Bicyclus anynana]	Bicyclus anynana	488	1550	100%	1e-165	95.16%	810	XP_023937889.1
<input checked="" type="checkbox"/> histone-lysine N-methyltransferase PRDM9-like isoform X1 [Maniola jurtina]	Maniola jurtina	337	1414	100%	4e-113	62.50%	301	XP_045778716.1
<input checked="" type="checkbox"/> zinc finger protein 260-like isoform X1 [Bicyclus anynana]	Bicyclus anynana	334	1907	100%	9e-107	65.06%	704	XP_023953410.1
<input checked="" type="checkbox"/> zinc finger protein 260-like isoform X2 [Bicyclus anynana]	Bicyclus anynana	334	1906	100%	1e-106	65.06%	701	XP_023953411.1
<input checked="" type="checkbox"/> gastrula zinc finger protein XICGF57.1-like [Bicyclus anynana]	Bicyclus anynana	318	1852	100%	3e-104	62.75%	411	XP_023952909.1
<input checked="" type="checkbox"/> gastrula zinc finger protein XICGF8.2DB-like isoform X1 [Nilaparvata lugens]	Nilaparvata lugens	316	887	100%	1e-103	55.69%	381	XP_039298152.1
<input checked="" type="checkbox"/> zinc finger protein 182-like isoform X4 [Bicyclus anynana]	Bicyclus anynana	325	611	100%	2e-103	63.82%	684	XP_023952407.1
<input checked="" type="checkbox"/> zinc finger protein 182-like isoform X3 [Bicyclus anynana]	Bicyclus anynana	325	612	100%	4e-103	63.82%	729	XP_023952406.1
<input checked="" type="checkbox"/> zinc finger protein 182-like isoform X1 [Bicyclus anynana]	Bicyclus anynana	325	612	100%	4e-103	63.82%	732	XP_023952404.1
<input checked="" type="checkbox"/> zinc finger protein 260-like isoform X2 [Bicyclus anynana]	Bicyclus anynana	325	612	100%	4e-103	63.82%	732	XP_023952405.1
<input checked="" type="checkbox"/> gastrula zinc finger protein XICGF57.1-like isoform X2 [Maniola jurtina]	Maniola jurtina	307	756	100%	2e-102	62.61%	245	XP_045778717.1
<input checked="" type="checkbox"/> gastrula zinc finger protein XICGF57.1-like [Nilaparvata lugens]	Nilaparvata lugens	319	1782	100%	4e-102	56.68%	619	XP_022189878.2
<input checked="" type="checkbox"/> gastrula zinc finger protein XICGF17.1-like [Maniola hyperantus]	Maniola hyperantus	309	1708	100%	8e-102	59.27%	329	XP_034839155.1
<input checked="" type="checkbox"/> gastrula zinc finger protein XICGF8.2DB-like isoform X1 [Bicyclus anynana]	Bicyclus anynana	310	610	100%	9e-102	62.55%	364	XP_023950901.1

Alignment details:

Descriptions	Graphic Summary	Alignments	Taxonomy
Alignment view Pairwise Restore defaults Download			
100 sequences selected			
Download GenPept Graphics Sort by: E value			
zinc finger protein 84-like [Bicyclus anynana]			
Sequence ID: XP_023937889.1 Length: 810 Number of Matches: 4			
Range 1: 476 to 723 GenPept Graphics			
Score	Expect	Method	Identities Positives Gaps
488 bits(1257)	1e-165	Compositional matrix adjust.	236/248(95%) 238/248(95%) 0/248(0%)
Query 1	SCDICHYKCAQGNLVCHIRHTCEKPYSCEMYNYKCARKSRLRHHMTTHTGKPFSCGI	60	
Sbjct 476	SCDICHYKCAQGNLVCHIRHT EKPYSCEM NYKCA KSLRHHMTTHTGKPFSCGI	535	
Query 61	CNYKTGVKNSLVCHLRHTHTGKPFCEICNYKFALKRNLNHNKTHHTGKPFSCGICNYK	120	
Sbjct 536	CNYKTGVKNSLVCHLRHTHTGKPFCEICNYKFALKRNLNHNKTHHTGKPFSCGICNYK	595	
Query 121	TGVKNSLVCHLRHTHTGKPFCEICNYKFALKHNLVNHMKIHTGKPFSCICNYKTRVK	180	
Sbjct 596	TGVKNSLVCHLRHTHTGKPFCEICNYKFALKHNLVNHMKIHTGKPFSCICNYKTRVK	655	
Query 181	NSLVSHLRHTHTGKPFSCICNYKSARKRYLLNHNKTHHTGKPFSCDICNYKTGIKNSLV	240	
Sbjct 656	NSLVHLRHTHTGKPFCEICNYK A K L+NHMKTHHTGKPFSCDICNYKTGIKN+LV	715	
Query 241	RHMRHTG 248		
Sbjct 716	RHMRHTG 723		
Related Information			
Gene - associated gene details			
Genome Data Viewer - aligned genomic context			

FIND-A-GENE PROJECT ASSIGNMENT

Range 2: 532 to 779 [GenPept](#) [Graphics](#)[▼ Next Match](#) [▲ Previous Match](#) [▲ First Match](#)

Score	Expect	Method	Identities	Positives	Gaps
409 bits(1051)	1e-134	Compositional matrix adjust.	200/248(81%)	216/248(87%)	0/248(0%)
Query 1	SCDICHYKCAQRGNLVCHIRHTCEKPYSCEMYNYKCARKSRLRHHMTTHTGEKPFSCGI	60			
Sbjct 532	SC IC+YK + +LVCH+RTHT EKP+ CE+ NYK A K L +HM THTGEKPFSCGI				
Query 61	CNYKTGVKNSLVCHLRHTHTGEKPFCEICNYKFALKRNLNLMKTHHTGEKPFSCGICNYK	120			
Sbjct 592	CNYKTGVKNSLVCHLRHTHTGEKPFCEICNYKFALKRNLNLMKTHHTGEKPFSCGICNYK				
Query 121	TGVKNSLVCHLRHTHTGEKPFCEICNYKFALKRNLNLMKTHHTGEKPFSCGICNYKTRVK	180			
Sbjct 652	T VKNLSVCHLRHTHTGEKPFCEICNYKFALKRNLNLMKTHHTGEKPFSCGICNYKT +K				
Query 181	NSLVSHLRHTHTGEKPFCEICNYKSARKRYLLNLMKTHHTGEKPFSCDICNYKTGIKNSLV	240			
Sbjct 712	N+LV H+RTHTGEKPFCEICN+KSA K LL+HMKTHHTGEKPFSC ICNYK K L+				
Query 241	RHMRHTG 248				
Sbjct 772	HM+ HTG				
	GHMKTHTG 779				

Range 3: 588 to 809 [GenPept](#) [Graphics](#)[▼ Next Match](#) [▲ Previous Match](#) [▲ First Match](#)

Score	Expect	Method	Identities	Positives	Gaps
334 bits(857)	8e-106	Compositional matrix adjust.	165/222(74%)	187/222(84%)	0/222(0%)
Query 1	SCDICHYKCAQRGNLVCHIRHTCEKPYSCEMYNYKCARKSRLRHHMTTHTGEKPFSCGI	60			
Sbjct 588	SC IC+YK + +LVCH+RTHT EKP+ CE+ NYK A K L +HM HTGEKPFSC I				
Query 61	CNYKTGVKNSLVCHLRHTHTGEKPFCEICNYKFALKRNLNLMKTHHTGEKPFSCGICNYK	120			
Sbjct 648	CNYKT VKNLSVCHLRHTHTGEKPFCEICNYKFALK NL+NHMKTHHTGEKPFSC ICNYK				
Query 121	TGVKNSLVCHLRHTHTGEKPFCEICNYKFALKRNLNLMKTHHTGEKPFSCGICNYKTRVK	180			
Sbjct 708	TG+KN+LV H+RTHTGEKPF CEICN+K ALKH+L+HMK HTGEKPFSC+ICNYK K				
Query 181	NSLVSHLRHTHTGEKPFCEICNYKSARKRYLLNLMKTHHTGEK 222				
Sbjct 768	L+ H++THTGEK F C++CNYK ARKR L +HMKTHTG K				
	QHLGHMKTHTGEKSFCCCLCNYKCARKRDLESHMKTHTGGK 809				

Range 4: 447 to 667 [GenPept](#) [Graphics](#)[▼ Next Match](#) [▲ Previous Match](#) [▲ First Match](#)

Score	Expect	Method	Identities	Positives	Gaps
317 bits(813)	3e-99	Compositional matrix adjust.	161/221(73%)	182/221(82%)	3/221(1%)
Query 31	EMYN-YKCARKSRLRHH--MTTHTGEKPFSCGICNYKTGVKNSLVCHLRHTHTGEKPFCE	87			
Sbjct 447	++Y+ +K +K+ L + + T T E +SC IC+YK + +LVCH+RTHTHTGEKPF CE				
Query 88	ICNYKFALKRNLNLMKTHHTGEKPFSCGICNYKTGVKNSLVCHLRHTHTGEKPFCEICNY	147			
Sbjct 507	+CNYK A K L +HM THTGEKPFSCGICNYKTGVKNSLVCHLRHTHTGEKPFCEICNY				
Query 148	KFALKRNLNLMKTHHTGEKPFCEICNYKTRVKNLSVSHLRHTHTGEKPFCEICNYKSAR	207			
Sbjct 567	KFALK NL+NHMK HTGEKPFSC ICNYKT VKNLSV HLRHTHTGEKPF CEICNYK A				
Query 208	KRYLLNHMKTHHTGEKPFSCDICNYKTGIKNSLVRRHRIHTG 248				
Sbjct 627	K L+NHMK HTGEKPFSC+ICNYKT +KNSLV H+R HTG				
	KHNLNLMKTHHTGEKPFCEICNYKTRVKNLSVCHLRHTHTG 667				