

1. Are you ready to start?:)

1. Yes

2. How many hits did you have in which databases?

2. Results found in 6 databases – MeSH, PubMed, PubMedCentral, Gene, Protein and Nucleotide

3. In which organism is the cheapdate expressed?

3. Drosophila

4. What was the date when the cheapdate was first described?

4. July 7, 1998

5. What is the cheapdate causing in an organism that is having this allele?

5. Ethanol intoxication

6. What is the other name of the gene? What are the functions of this gene?

6. Gene name is amnesiac and its function as mentioned in the research paper is amnesiac has been postulated to encode a neuropeptide that activates the cAMP pathway.

7. Cite the publication that first described cheapdate allele.

7. 1: Moore MS, DeZazzo J, Luk AY, Tully T, Singh CM, Heberlein U. Ethanol intoxication in Drosophila: Genetic and pharmacological evidence for regulation by the cAMP signaling pathway. Cell. 1998 Jun 12;93(6):997-1007. PubMed PMID:9635429.

8. Of how many nucleotides does the cheapdate gene consists?

8. 3953 nucleotides with combinations of C,A,T and G are present in the cheapdate

9. What is characteristic about FASTA format?

9. It is a single long string containing the sequence of C,A,T and G and is 3953 characters long.

10. From which database was the information about cheapdate gene downloaded? What is specific about this database?

10. It is inside the Gene database of NCBI website. The unique thing is that it contains a lot of information about the gene including -Genomic context, Genomic regions, transcripts, and products, bibliography etc. It is also dynamic in nature

11. On which chromosome is the gene located? On which position is the gene located? Provide the assembly of reference genome that you have chosen to give the positions of gene.

11. Below is the information I found of the assembly reference in the NCBI-

```
Gene: amn
Title: amnesiac
RNA title: mRNA-amnesiac
Protein title: amnesiac
Merged features: NM_078692.2 and NP_523416.2
Location: complement(19,886,108..19,890,060)
[Length]
Span on NC_004354.4: 3,953 nt
Protein length: 180 aa
[NM_078692.2]
Exon: 1 of 1
mRNA position: 1,396
mRNA sequence: AAAAAATAATATAAA[G]GATCTCTTAGTGA
```

Genomic

1. NC_004354.4 Reference assembly	Range
19886108..19890060 complement	
	Download
GenBank, FASTA, Sequence Viewer (Graphics)	

mRNA and Protein(s)

1. NM_078692.2 → NP_523416.2 amnesiac [Drosophila melanogaster]
See identical proteins and their annotated locations for NP_523416.2
Status: REVIEWED
UniProtKB/Swiss-Prot
Q24049
UniProtKB/TrEMBL
C9QPI0
Related
FBpp0074571, FBtr0074802

12. What are the other genes, that cheapdate interacts with? Hint: go to interactions section.

12. The genes it interacts with can be seen from the table below which are mentioned in Other Gene column in the below table-

Products	Interactant	Other Gene	Complex	Source	Pubs	Description
BioGRID:59293	BioGRID:62436	Arf51F		BioGRID	PubMed	Phenotypic Enhancement
BioGRID:59293	BioGRID:67421	H		BioGRID	PubMed	Phenotypic Suppression
BioGRID:59293	BioGRID:57834	dnc		BioGRID	PubMed	Phenotypic Suppression
BioGRID:59293	BioGRID:64423	orb2		BioGRID	PubMed	Protein-RNA
BioGRID:59293	BioGRID:64306	pbl		BioGRID	PubMed	Phenotypic Enhancement
BioGRID:59293	BioGRID:69326	pnr		BioGRID	PubMed	Phenotypic Suppression
BioGRID:59293	BioGRID:58765	rut		BioGRID	PubMed	Phenotypic Suppression

13. By clicking on the links, try to find IDs for cheapdate gene for a) RefSeq and b) Ensembl. Next, on the NCBI Gene website find the Entrez ID for the gene. Hint: Entrez ID is also named geneID.

14. What is the subcellular location of the protein?

14. Extracellular region or secreted

15. What is the fasta sequence of the amino-acids in the protein?

15. >sp|Q24049|AMN_DROME Amnesiac neuropeptides OS=Drosophila melanogaster OX=7227 GN=amn PE=1 SV=3

MRSFCCCFYPAVALHCVLLFYTFLLFRASALRRRVVSGSKGSAALALCRQFEQLSASR
RERAEECRTTQLRYHYHRNGAQSRSLCAAVLCCKRSYIPRPNFSCFSLVFPVGQRFAAAR
TRFGPTLVASWPLCNDSETKVLTKWPSCSLIGRRSVPRGQPKFSRENPRALSPSLLGEMR

16. Are there any similar proteins

16. RT03217p 100% similarity and UPI0007E70A8D as 50% similarity

17. What are the functions of the protein? Do the biological processes correspond to the function of the protein?

17. Required for associative learning and memory in adults. Expression pattern suggests a modulatory role in memory formation. Controls neurotransmitter-mediated signalling pathways associated with the structure of the larval peripheral nerve. Yes the biological process correspond to function of protein

18. Are the interactions in the STRING-db similar to the interactions found in Gene Database?

18.No

19. Using Legend option, decode the relations between the proteins. What kind of relations does the amn protein has?

19. Amnesiac neuropeptides; Required for associative learning and memory in adults. Expression pattern suggests a modulatory role in memory formation. Controls neurotransmitter-mediated signaling pathways associated with the structure of the larval peripheral nerve (180 aa)