PSI Blast for

gi|112497046|gb|ABI20163.1| 14.5 kDa salivary...

a member of the 15-17 kDa insect protein family

- No new sequences were found above the 1e-05 threshold

Reference
Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Reference - composition-based statistics starting in round 2
Alejandro A. Schäffer, L. Aravind, Thomas L. Madden, Sergei Shavirin, John L. Spouge, Yuri I. Wolf, Eugene V. Koonin, and Stephen F. Altschul (2001), "Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements", Nucleic Acids Res. 29:2994-3005.

Search Parameters

Search parameter name	Search parameter value
Program	blastp
Word size	3
Expect value	10
Hitlist size	10000
Gapcosts	11,1
Matrix	BLOSUM62
Filter string	F
Genetic Code	1
Window Size	40
Threshold	11
Composition-based stats	1

Database

Database parameter name	Database parameter value
Posted date	Feb 16, 2010 5:41 PM
Number of letters	3,560,334,713
Number of sequences	10,435,126
Entrez query	none

Karlin-Altschul statistics

Params	Ungapped	Gapped
Lambda	0.315147	0.267
K	0.133848	0.0449541
H	0.378164	0.14

Results Statistics

Results Statistics parameter name	Results Statistics parameter value

Length adjustment	101
Effective length of query	35
Effective length of database	2506386987
Effective search space	87723544545
Effective search space used	87723544545

```
Sequences with    pattern at position  and  E-value BETTER than threshold

                                                                   Score    E
Sequences producing significant alignments:                       (Bits)  Value

● ☑  gb|ABI20163.1|  14.5 kDa salivary protein [Phlebotomus duboscqi]   128    3e-28
● ☑  gb|ACF72876.1|  unknown [Ochlerotatus taeniorhynchus]              126    6e-28
● ☑  ref|XP_001850045.1|  conserved hypothetical protein [Culex qui...  126    8e-28 UG
● ☑  ref|XP_001650838.1|  hypothetical protein AaeL_AAEL005428 [Aed...  126    9e-28 UG
● ☑  ref|XP_002101617.1|  GE15521 [Drosophila yakuba] >gb|EDX02725....  125    1e-27 G
● ☑  ref|XP_002039288.1|  GM22809 [Drosophila sechellia] >ref|XP_00...  125    2e-27 G
● ☑  ref|NP_573299.1|  CG15044 [Drosophila melanogaster] >gb|AAL488...  125    2e-27 UG
● ☑  gb|ABV44740.1|  13.6 kDa midgut protein [Phlebotomus papatasi]     122    1e-26
● ☑  gb|ABV60331.1|  11.6 kDa midgut protein [Lutzomyia longipalpis]    121    2e-26
● ☑  ref|NP_001033959.1|  CG33998 [Drosophila melanogaster] >tpg|DA...  121    2e-26 UG
● ☑  ref|XP_001961218.1|  GF11112 [Drosophila ananassae] >gb|EDV380...  121    3e-26 G
● ☑  ref|XP_002026810.1|  GL27028 [Drosophila persimilis] >gb|EDW33...  121    3e-26 G
● ☑  ref|XP_002059070.1|  GJ15191 [Drosophila virilis] >gb|EDW58139...  120    3e-26 G
● ☑  ref|XP_002015909.1|  GL10767 [Drosophila persimilis] >gb|EDW31...  120    4e-26 G
● ☑  ref|XP_001974726.1|  GG21920 [Drosophila erecta] >gb|EDV55126....  120    4e-26 G
● ☑  gb|ACI30042.1|  hypothetical secreted protein [Anopheles darli...  120    4e-26
● ☑  ref|XP_002071202.1|  GK25257 [Drosophila willistoni] >gb|EDW82...  120    5e-26 G
● ☑  ref|XP_002048773.1|  GJ21142 [Drosophila virilis] >gb|EDW59966...  119    8e-26 G
● ☑  ref|XP_001354364.2|  GA13455 [Drosophila pseudoobscura pseudoo...  119    9e-26 G
● ☑  ref|XP_001230950.2|  AGAP003713-PA [Anopheles gambiae str. PES...  118    2e-25 UG
● ☑  ref|XP_001977671.1|  GG18116 [Drosophila erecta] >gb|EDV46598....  118    2e-25 G
● ☑  ref|XP_001354363.1|  GA13456 [Drosophila pseudoobscura pseudoo...  118    2e-25 G
● ☑  ref|XP_002138284.1|  GA24478 [Drosophila pseudoobscura pseudoo...  117    4e-25 G
● ☑  gb|ACD81831.1|  IP21621p [Drosophila melanogaster]                 116    6e-25
● ☑  ref|XP_002071201.1|  GK25258 [Drosophila willistoni] >gb|EDW82...  116    6e-25 G
● ☑  ref|XP_001987728.1|  GH22082 [Drosophila grimshawi] >gb|EDW025...  116    8e-25 G
● ☑  ref|XP_002009841.1|  GI15587 [Drosophila mojavensis] >gb|EDW07...  116    9e-25 G
● ☑  ref|XP_002082142.1|  GD11406 [Drosophila simulans] >gb|EDX0772...  116    1e-24 G
● ☑  ref|XP_002034496.1|  GM21912 [Drosophila sechellia] >gb|EDW485...  115    2e-24 G
● ☑  ref|XP_001990970.1|  GH12337 [Drosophila grimshawi] >gb|EDV995...  115    2e-24 G
● ☑  gb|ABV60298.1|  14.3 kDa midgut protein [Lutzomyia longipalpis]    115    2e-24
● ☑  ref|XP_002009842.1|  GI15588 [Drosophila mojavensis] >gb|EDW07...  114    4e-24 G
```

●	☑	ref|XP_001990969.1|	GH12339 [Drosophila grimshawi] >gb|EDV995...	113	7e-24	G
●	☑	ref|XP_002061455.1|	GK20920 [Drosophila willistoni] >gb|EDW72...	112	9e-24	G
●	☑	ref|XP_002101618.1|	GE15520 [Drosophila yakuba] >gb|EDX02726....	112	1e-23	G
●	☑	ref|XP_002050287.1|	GJ20290 [Drosophila virilis] >gb|EDW61480...	112	1e-23	G
●	☑	ref|XP_002004523.1|	GI19567 [Drosophila mojavensis] >gb|EDW08...	112	1e-23	G
●	☑	ref|XP_001986843.1|	GH21597 [Drosophila grimshawi] >gb|EDW017...	111	2e-23	G
●	☑	gb|AAQ09813.1|	CG15043 [Drosophila yakuba]	111	2e-23	
●	☑	ref|XP_001965516.1|	GF22411 [Drosophila ananassae] >gb|EDV350...	111	2e-23	G
●	☑	ref|XP_001965515.1|	GF22412 [Drosophila ananassae] >gb|EDV350...	111	2e-23	G
●	☑	ref|XP_001238394.2|	AGAP011630-PA [Anopheles gambiae str. PES...	111	2e-23	UG
●	☑	ref|XP_001360670.1|	GA12205 [Drosophila pseudoobscura pseudoo...	111	2e-23	G
●	☑	gb|ABR27888.1|	putative salivary secreted protein [Triatoma i...	111	3e-23	
●	☑	ref|XP_002016256.1|	GL10591 [Drosophila persimilis] >gb|EDW32...	110	3e-23	G
●	☑	ref|XP_002062967.1|	GK21639 [Drosophila willistoni] >gb|EDW73...	110	4e-23	G
●	☑	gb|ABR27890.1|	putative salivary secreted peptide [Triatoma i...	110	7e-23	
●	☑	ref|XP_002059069.1|	GJ15192 [Drosophila virilis] >gb|EDW58138...	109	1e-22	G
●	☑	ref|XP_002059818.1|	GJ15031 [Drosophila virilis] >gb|EDW57155...	109	1e-22	G
●	☑	ref|XP_001997725.1|	GH17985 [Drosophila grimshawi] >gb|EDW049...	108	2e-22	G
●	☑	ref|XP_001977672.1|	GG18115 [Drosophila erecta] >gb|EDV46599....	108	2e-22	G
●	☑	gb|AAV84229.1|	unknown [Culicoides sonorensis]	108	2e-22	
●	☑	ref|XP_002005185.1|	GI19218 [Drosophila mojavensis] >gb|EDW09...	108	2e-22	G
●	☑	ref|XP_001660534.1|	hypothetical protein AaeL_AAEL009985 [Aed...	108	3e-22	UG
●	☑	ref|XP_002050017.1|	GJ21904 [Drosophila virilis] >gb|EDW61210...	107	3e-22	G
●	☑	ref|XP_001986842.1|	GH21596 [Drosophila grimshawi] >gb|EDW017...	107	4e-22	G
●	☑	ref|XP_002004350.1|	GI19667 [Drosophila mojavensis] >gb|EDW08...	107	5e-22	G
●	☑	ref|NP_573300.1|	CG15043 [Drosophila melanogaster] >gb|AAF488...	107	5e-22	UG
●	☑	gb|AAV84228.1|	unknown [Culicoides sonorensis]	106	6e-22	
●	☑	ref|XP_001868963.1|	14.5 kDa salivary protein [Culex quinquef...	106	6e-22	UG
●	☑	ref|XP_002063501.1|	GK21373 [Drosophila willistoni] >gb|EDW74...	106	7e-22	G
●	☑	gb|ACT79296.1|	putative fatbody protein 3Rev-G1 [Bombyx mori]	106	7e-22	
●	☑	ref|XP_001975823.1|	GG22531 [Drosophila erecta] >gb|EDV56223....	106	7e-22	G
●	☑	ref|XP_002005815.1|	GI18871 [Drosophila mojavensis] >gb|EDW09...	106	7e-22	G
●	☑	gb|ABV44744.1|	14.4 kDa midgut protein [Phlebotomus papatasi]	106	7e-22	
●	☑	ref|XP_002081284.1|	GD25794 [Drosophila simulans] >gb|EDX0686...	106	7e-22	G
●	☑	ref|XP_001664091.1|	hypothetical protein AaeL_AAEL013885 [Aed...	106	9e-22	UG
●	☑	ref|XP_001987508.1|	GH21960 [Drosophila grimshawi] >gb|EDW023...	106	9e-22	G
●	☑	gb|ACN69173.1|	hypothetical conserved secreted protein [Stomo...	106	1e-21	
●	☑	ref|XP_002107315.1|	GD15631 [Drosophila simulans] >gb|EDX1829...	104	3e-21	G
●	☑	ref|ZP_05005660.1|	conserved hypothetical protein [Streptomyc...	104	3e-21	
●	☑	ref|XP_002101745.1|	GE17797 [Drosophila yakuba] >gb|EDX02853....	103	4e-21	G
●	☑	ref|XP_002028226.1|	GL16235 [Drosophila persimilis] >gb|EDW38...	103	4e-21	G

●	☑	ref|XP_002039289.1|	GM22808 [Drosophila sechellia] >gb|EDW560...	103	7e-21	G
●	☑	ref|XP_001959559.1|	GF11989 [Drosophila ananassae] >gb|EDV363...	103	8e-21	G
●	☑	ref|XP_002076927.1|	GD24776 [Drosophila simulans] >gb|EDX1652...	103	8e-21	G
●	☑	ref|NP_001155164.1|	venom protein R [Nasonia vitripennis]	103	9e-21	UG
●	☑	ref|NP_001033839.1|	CG34026 [Drosophila melanogaster] >gb|ABC...	102	1e-20	UG
●	☑	ref|ZP_05006207.1|	conserved hypothetical protein [Streptomyc...	102	1e-20	
●	☑	ref|XP_001868961.1|	14.5 kDa salivary protein [Culex quinquef...	102	1e-20	UG
●	☑	ref|XP_001868962.1|	conserved hypothetical protein [Culex qui...	102	1e-20	UG
●	☑	ref|XP_002063499.1|	GK21944 [Drosophila willistoni] >gb|EDW74...	102	1e-20	G
●	☑	ref|XP_001865641.1|	conserved hypothetical protein [Culex qui...	101	2e-20	G
●	☑	ref|XP_001975824.1|	GG22532 [Drosophila erecta] >gb|EDV56224....	101	2e-20	G
●	☑	ref|XP_001987095.1|	GH21729 [Drosophila grimshawi] >gb|EDW019...	101	2e-20	G
●	☑	ref|XP_001865369.1|	conserved hypothetical protein [Culex qui...	101	3e-20	UG
●	☑	ref|XP_002041778.1|	GM11373 [Drosophila sechellia] >gb|EDW456...	101	3e-20	G
●	☑	ref|XP_001664100.1|	hypothetical protein AaeL_AAEL013895 [Aed...	101	4e-20	UG
●	☑	gb|ABH10141.1|	HMG176 [Helicoverpa armigera] >gb|ABK29472.1| ...	100	5e-20	
●	☑	gb|ABV44709.1|	18 kDa midgut protein [Phlebotomus papatasi]	100	6e-20	
●	☑	ref|XP_001604583.1|	PREDICTED: hypothetical protein [Nasonia ...	100	6e-20	UG
●	☑	ref|XP_002016257.1|	GL10590 [Drosophila persimilis] >gb|EDW32...	100	7e-20	G
●	☑	ref|XP_001965768.1|	GF20010 [Drosophila ananassae] >gb|EDV414...	100	7e-20	G
●	☑	ref|XP_001649836.1|	hypothetical protein AaeL_AAEL004809 [Aed...	99.6	8e-20	UG
●	☑	ref|XP_001604165.1|	PREDICTED: hypothetical protein [Nasonia ...	99.6	9e-20	UG
●	☑	ref|XP_001599861.1|	PREDICTED: similar to CG34026-PA [Nasonia...	99.2	1e-19	UG
●	☑	ref|XP_002028229.1|	GL16227 [Drosophila persimilis] >gb|EDW38...	99.2	1e-19	G
●	☑	ref|XP_002099025.1|	Acp225 [Drosophila yakuba] >gb|EDW98737.1...	99.2	1e-19	G
●	☑	gb|ABW82136.2|	IP21714p [Drosophila melanogaster]	98.8	1e-19	
●	☑	gb|ACD99549.1|	IP21514p [Drosophila melanogaster] >gb|ACD9955...	98.8	2e-19	
●	☑	ref|XP_001865640.1|	hypothetical protein CpipJ_CPIJ015171 [Cu...	98.5	2e-19	UG
●	☑	gb|ACO72873.1|	MIP04578p [Drosophila melanogaster]	98.5	2e-19	
●	☑	ref|XP_001599834.1|	PREDICTED: similar to putative salivary s...	98.5	2e-19	UG
●	☑	gb|ABD97958.1|	ACP225 [Drosophila yakuba]	98.1	2e-19	
●	☑	gb|ABD97961.1|	ACP225 [Drosophila yakuba]	98.1	3e-19	
●	☑	gb|ABD97965.1|	ACP225 [Drosophila yakuba]	97.7	3e-19	
●	☑	gb|ABD97960.1|	ACP225 [Drosophila yakuba]	97.7	3e-19	
●	☑	ref|NP_726308.1|	CG30413 [Drosophila melanogaster] >gb|AAM682...	97.7	4e-19	UG
●	☑	ref|XP_001603272.1|	PREDICTED: similar to ENSANGP00000031402 ...	97.7	4e-19	G
●	☑	gb|ABD97959.1|	ACP225 [Drosophila yakuba]	97.7	4e-19	
●	☑	ref|XP_001965767.1|	GF20011 [Drosophila ananassae] >gb|EDV414...	97.3	5e-19	G
●	☑	gb|ABQ18257.1|	unknown [Lygus lineolaris]	97.3	5e-19	
●	☑	ref|XP_002050018.1|	GJ20423 [Drosophila virilis] >gb|EDW61211...	97.3	5e-19	G
●	☑	gb|ABD97966.1|	ACP225 [Drosophila yakuba]	97.3	5e-19	

●	☑	gb|ABD97964.1|	ACP225 [Drosophila yakuba]	96.5	7e-19	
●	☑	ref|XP_002033647.1|	GM20317 [Drosophila sechellia] >gb|EDW476...	96.5	8e-19	G
●	☑	ref|NP_610832.1|	CG13323 [Drosophila melanogaster] >gb|AAF584...	96.5	9e-19	UG
●	☑	ref|XP_002081283.1|	GD25795 [Drosophila simulans] >gb|EDX0686...	96.2	1e-18	G
●	☑	ref|NP_610833.1|	CG13324 [Drosophila melanogaster] >gb|AAF584...	96.2	1e-18	UG
●	☑	ref|XP_001602278.1|	PREDICTED: similar to conserved hypotheti...	95.4	2e-18	G
●	☑	ref|XP_002071834.1|	GK10198 [Drosophila willistoni] >gb|EDW82...	95.0	2e-18	G
●	☑	gb|ABI20159.1|	16 kDa salivary protein [Phlebotomus duboscqi]	94.6	3e-18	
●	☑	gb|ABD97962.1|	ACP225 [Drosophila yakuba]	94.6	3e-18	
●	☑	gb|ABD97963.1|	ACP225 [Drosophila yakuba]	92.7	1e-17	
●	☑	ref|XP_001604205.1|	PREDICTED: hypothetical protein [Nasonia ...	92.3	1e-17	G
●	☑	ref|XP_001661180.1|	hypothetical protein AaeL_AAEL002263 [Aed...	92.3	1e-17	UG
●	☑	ref|XP_002028227.1|	GL16228 [Drosophila persimilis] >gb|EDW38...	91.9	2e-17	G
●	☑	ref|XP_002091873.1|	GE11995 [Drosophila yakuba] >gb|EDW91585....	91.9	2e-17	G
●	☑	ref|XP_001661179.1|	hypothetical protein AaeL_AAEL002263 [Aed...	91.5	2e-17	UG
●	☑	gb|ABD97967.1|	ACP225 [Drosophila teissieri]	90.8	4e-17	
●	☑	gb|ACY69959.1|	putative putative salivary secreted protein [C...	90.8	4e-17	
●	☑	gb|AAY34548.1|	putative 11.4 kDa secreted salivary gland prot...	90.4	5e-17	
●	☑	ref|XP_976138.1|	PREDICTED: hypothetical protein [Tribolium c...	90.0	8e-17	UG
●	☑	gb|ACI30202.1|	hypothetical conserved mosquito protein [Anoph...	90.0	8e-17	
●	☑	ref|XP_002090966.1|	GE13403 [Drosophila yakuba] >gb|AAQ09833....	89.6	9e-17	G
●	☑	ref|XP_002040057.1|	GM15559 [Drosophila sechellia] >gb|EDW569...	89.6	1e-16	G
●	☑	ref|XP_002082759.1|	GD25061 [Drosophila simulans] >gb|EDX0834...	89.6	1e-16	G
●	☑	ref|XP_002086118.1|	GE14474 [Drosophila yakuba] >ref|XP_00209...	89.2	1e-16	G
●	☑	gb|ABR27885.1|	salivary secreted protein [Triatoma infestans]	89.2	1e-16	
●	☑	ref|XP_316478.4|	AGAP006442-PA [Anopheles gambiae str. PEST] ...	88.8	2e-16	UG
●	☑	ref|XP_001976402.1|	GG22852 [Drosophila erecta] >gb|EDV56802....	88.4	2e-16	G
●	☑	ref|XP_001987096.1|	GH20161 [Drosophila grimshawi] >gb|EDW019...	88.4	2e-16	G
●	☑	ref|XP_002033648.1|	GM20316 [Drosophila sechellia] >gb|EDW476...	88.1	3e-16	G
●	☑	gb|ABN54491.1|	conserved secreted protein [Oncopeltus fasciatus]	87.7	4e-16	
●	☑	ref|XP_001976379.1|	GG20045 [Drosophila erecta] >gb|EDV56779....	87.7	4e-16	G
●	☑	ref|XP_002090965.1|	GE13402 [Drosophila yakuba] >gb|EDW90677....	87.3	4e-16	G
●	☑	ref|NP_724118.1|	CG31789 [Drosophila melanogaster] >gb|AAN110...	87.3	4e-16	UG
●	☑	ref|XP_002038990.1|	GM17281 [Drosophila sechellia] >gb|EDW556...	87.3	5e-16	G
●	☑	ref|NP_726325.1|	CG30411 [Drosophila melanogaster] >gb|AAL485...	86.9	5e-16	UG
●	☑	ref|XP_002082783.1|	GD11761 [Drosophila simulans] >gb|EDX0836...	86.9	6e-16	G
●	☑	ref|XP_001604117.1|	PREDICTED: similar to HDC07203 [Nasonia v...	86.5	8e-16	UG
●	☑	ref|XP_002005814.1|	GI20673 [Drosophila mojavensis] >gb|EDW09...	86.1	1e-15	G
●	☑	ref|XP_002079800.1|	GD24143 [Drosophila simulans] >gb|EDX0538...	85.7	1e-15	G
●	☑	229002332 dbj|BAH57948.1|	hypothetical protein [Bombyx mori]	85.7	2e-15	

```
● ☑ 156538599  ref|XP_001607547.1|  PREDICTED: similar to HDC07203 [Nasonia v...  85.4  2e-15 UG
● ☑ 195489141  ref|XP_002092612.1|  GE14288 [Drosophila yakuba] >gb|EDW92324....  85.0  2e-15 G
● ☑ 195347078  ref|XP_002040081.1|  GM16010 [Drosophila sechellia] >gb|EDW569...  84.6  3e-15 G
● ☑ 226342894  ref|NP_001139709.1|  Bm122 protein [Bombyx mori] >gb|ACN81326....  84.2  4e-15 UG
● ☑ 158288659  ref|XP_001237263.2|  AGAP000570-PA [Anopheles gambiae str. PES...  82.7  1e-14 UG
● ☑ 156540630  ref|XP_001599581.1|  PREDICTED: hypothetical protein [Nasonia ...  82.7  1e-14 UG
● ☑ 195489090  ref|XP_002092590.1|  GE11581 [Drosophila yakuba] >gb|EDW92302....  82.7  1e-14 G
● ☑ 189233613  ref|XP_001811691.1|  PREDICTED: similar to conserved hypotheti...  82.3  2e-14 G
● ☑ 189236015  ref|XP_001807284.1|  PREDICTED: similar to 11.6 kDa midgut pro...  81.1  3e-14 UG
● ☑ 194755056  ref|XP_001959808.1|  GF13054 [Drosophila ananassae] >gb|EDV366...  80.4  5e-14 G
● ☑ 157111220  ref|XP_001651441.1|  hypothetical protein AaeL_AAEL005789 [Aed...  78.8  2e-13 UG
● ☑ 195440326  ref|XP_002067993.1|  GK10923 [Drosophila willistoni] >gb|EDW78...  78.8  2e-13 G
● ☑ 218505759  ref|NP_001136225.1|  hypothetical protein LOC100216500 [Bombyx...  76.5  8e-13 UG
● ☑ 194880143  ref|XP_001974374.1|  GG21120 [Drosophila erecta] >gb|EDV54774....  76.1  1e-12 G
● ☑ 270004645  gb|EFA01093.1|  hypothetical protein TcasGA2_TC004016 [Triboli...  75.7  2e-12
● ☑ 195455264  ref|XP_002074639.1|  GK23181 [Drosophila willistoni] >gb|EDW85...  75.3  2e-12 G
● ☑ 195384233  ref|XP_002050822.1|  GJ22362 [Drosophila virilis] >gb|EDW62015...  75.0  2e-12 G
● ☑ 194890484  ref|XP_001977319.1|  GG18315 [Drosophila erecta] >gb|EDV46246....  75.0  3e-12 G
```

```
>☐ gb|ABI20163.1|  14.5 kDa salivary protein [Phlebotomus duboscqi]
Length=153

 Score =  128 bits (321),  Expect = 3e-28, Method: Composition-based stats.
 Identities = 136/136 (100%), Positives = 136/136 (100%), Gaps = 0/136 (0%)

Query  1    TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQ  60
            TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQ
Sbjct  18   TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQ  77

Query  61   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  120
            DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE
Sbjct  78   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  137

Query  121  ARNTYTFSYIAEIYGL  136
            ARNTYTFSYIAEIYGL
Sbjct  138  ARNTYTFSYIAEIYGL  153
```

```
>☐ gb|ACF72876.1|  unknown [Ochlerotatus taeniorhynchus]
Length=117

 Score =  126 bits (317),  Expect = 6e-28, Method: Composition-based stats.
 Identities = 28/115 (24%), Positives = 47/115 (40%), Gaps = 1/115 (0%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                    L     +  +++ +   G +   D L+ R     S   LQ   + + YP  G
Sbjct  2    KSLVLAAVLVSFGALVLSQSHNYFWGYKGPYDVLLNRTFAIKSSSILQVKTMDLIYPLKG  61

Query  82   GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G  +S ++I     +  G    +L GGIG     I ++++    FS+I EIYG
Sbjct  62   QLGRNISAINITDQYINGKGGYASLLAGGIGYNHTKIHLKSQRGNGFSFIVEIYG  116
```

```
>☐ ref|XP_001850045.1| UG conserved hypothetical protein [Culex quinquefasciatus]
```

```
 gb|EDS31353.1| G conserved hypothetical protein [Culex quinquefasciatus]
Length=117

 GENE ID: 6040690 CpipJ_CPIJ008016 | hypothetical protein
[Culex quinquefasciatus]

 Score =  126 bits (317),  Expect = 8e-28, Method: Composition-based stats.
 Identities = 29/115 (25%), Positives = 49/115 (42%), Gaps = 1/115 (0%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                  V L  L  +   ++ H   G R   D L+ R     S   LQ   + + YP  G
Sbjct  2    KTLALTVVLFSLGALVFTQSHHYQWGFRGPYDVLLNRTIAIKSSSILQVKTMDLYYPLKG  61

Query  82   GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G  ++++++     +  G    +L GG+G     I ++++    FS+I EIYG
Sbjct  62   QLGRNITFMNVTDQYINGKGGYASLLAGGVGFNNTKIHLKSQRGNGFSFIVEIYG  116


>☐ ref|XP_001650838.1| UG hypothetical protein AaeL_AAEL005428 [Aedes aegypti]
 gb|EAT43100.1| G conserved hypothetical protein [Aedes aegypti]
Length=117

 GENE ID: 5566457 AaeL_AAEL005428 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score =  126 bits (316),  Expect = 9e-28, Method: Composition-based stats.
 Identities = 28/115 (24%), Positives = 48/115 (41%), Gaps = 1/115 (0%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                  + L     +  +++ +   G +   D L+ R     S   LQ   + + YP  G
Sbjct  2    KTIVFALILVSFGALVISQSHNYFWGYKGPYDVLLNRTFAIKSSSILQVKTMDLYYPLKG  61

Query  82   GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G  +S ++I     +  G    +L GGIG     I ++++    FS+I EIYG
Sbjct  62   QLGRNISAINITDQFINGKGGYASLLAGGIGYNHTTIHLKSQRGNGFSFIVEIYG  116


>☐ ref|XP_002101617.1|  GE15521 [Drosophila yakuba]
 gb|EDX02725.1| G GE15521 [Drosophila yakuba]
Length=157

 GENE ID: 6525796 Dyak\GE15521 | GE15521 gene product from transcript GE15521-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score =  125 bits (315),  Expect = 1e-27, Method: Composition-based stats.
 Identities = 41/128 (32%), Positives = 70/128 (54%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWGS  67
               IT V D   +    NP   LT+++ V+ ++     + G R SGDRL+  + D  S+ S
Sbjct  29   YTDITPVTDDTRYTSLNPDTQLTEMNKVKHSKGNIVFTAGDRASGDRLIVNHYDGESFPS  88

Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
             ++V++ ++YPA    G  ++ + + VD +++    GY+  GGIGQ  + IL+ +  T +F
Sbjct  89   AKDVEVLMSYPAGASTGVTLTSIEVYVDMSADDAGGYLTKGGIGQTNVEILLTSNTTRSF  148

Query  128  SYIAEIYG  135
             Y   IYG
Sbjct  149  VYETFIYG  156


>☐ ref|XP_002039288.1| G GM22809 [Drosophila sechellia]
 ref|XP_002107314.1| G GD15632 [Drosophila simulans]
 gb|EDW56030.1| G GM22809 [Drosophila sechellia]
 gb|EDX18296.1| G GD15632 [Drosophila simulans]
Length=156

 GENE ID: 6614870 Dsec\GM22809 | GM22809 gene product from transcript GM22809-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score =  125 bits (314),  Expect = 2e-27, Method: Composition-based stats.
 Identities = 40/128 (31%), Positives = 70/128 (54%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWGS  67
               IT V D   +    NP   LT+++ V+ ++     + G+R SGDRL+  + D  ++ S
Sbjct  28   YTDITPVTDETRYTTLNPDTQLTEMNKVKHSKGNIVFTAGERASGDRLIVNHYDDENFPS  87
```

```
Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
             ++V++ ++YPA    G  ++ + + VD  ++   GY+  GGIGQ  + IL+ +  T +F
Sbjct  88   AKDVEVLMSYPAGASTGVTLTSIEVYVDMAADDAGGYLTKGGIGQTNVEILLTSNQTRSF  147

Query  128  SYIAEIYG  135
             Y   IYG
Sbjct  148  VYETFIYG  155


>☐ ref|NP_573299.1| UG CG15044 [Drosophila melanogaster]
 gb|AAL48888.1| G RE30049p [Drosophila melanogaster]
 gb|AAF48849.2| G CG15044 [Drosophila melanogaster]
 gb|ACL86723.1|  CG15044-PA [synthetic construct]
 gb|ACL91333.1|  CG15044-PA [synthetic construct]
Length=160

 GENE ID: 32834 CG15044 | CG15044 gene product from transcript CG15044-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score =  125 bits (313),  Expect = 2e-27, Method: Composition-based stats.
 Identities = 41/128 (32%), Positives = 70/128 (54%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIH-NSLGQRVSGDRLVARNQDSASWGS  67
               IT V D   +   NP   L +++ V+ ++     + G+R SGDRL+  + D  S+ S
Sbjct  32   YTDITPVTDETRYTTLNPDAQLNEMNKVKHSKGNVVFTAGKRASGDRLIVNHYDDDSFPS  91

Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
             ++V++ ++YPA    G  ++ + + VD T++   GY+  GGIGQ  + IL+ +  T +F
Sbjct  92   AKDVEVLMSYPAGASTGVTLTSIEVYVDMTADDAGGYLTKGGIGQTNVEILLTSNQTRSF  151

Query  128  SYIAEIYG  135
             Y   IYG
Sbjct  152  VYETFIYG  159


>☐ gb|ABV44740.1|  13.6 kDa midgut protein [Phlebotomus papatasi]
Length=120

 Score =  122 bits (306),  Expect = 1e-26, Method: Composition-based stats.
 Identities = 25/120 (20%), Positives = 45/120 (37%), Gaps = 6/120 (5%)

Query  20   EFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPA  79
            +F  A   V L     V  A++ +   G     D+L+ R     S   +Q      ++P
Sbjct  2    KFAVAF--VVLVAFCGVLYAQSHNMQWGNTTHYDQLLHRAFVKKSSSFMQTKTTEYSWPN  59

Query  80   NGG---YGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                      ++ +  L    +  G    +  GG+G     I  +++  + F +I EIYG
Sbjct  60   ATNSQPRNRTITCIRALDQYVNGKGGYATLRAGGVGFNHTTIRFKSQRGHGFDFILEIYG  119


>☐ gb|ABV60331.1|  11.6 kDa midgut protein [Lutzomyia longipalpis]
Length=121

 Score =  121 bits (305),  Expect = 2e-26, Method: Composition-based stats.
 Identities = 25/114 (21%), Positives = 46/114 (40%), Gaps = 5/114 (4%)

Query  27   GVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG----  82
             + L     V  A++ +   G +   DRL+ R     S   +Q V   VN+P
Sbjct  7    LIVLAVFCGVLYAQSHNIQWGNQTHYDRLLHREFVHKSSKWMQVVTHEVNWPNRTQAAQA  66

Query  83   YGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              A ++Y+  +    +  G    +  GG+G     +  +++ +    +I EIYG
Sbjct  67   RNATITYIRAMDQYVNGKGGYATLRAGGVGFNHTEVRFKSQRSQGIDFILEIYG  120


>☐ ref|NP_001033959.1| UG CG33998 [Drosophila melanogaster]
 tpg|DAA02612.1| G TPA_inf: HDC07203 [Drosophila melanogaster]
 gb|ABC66050.1| G CG33998 [Drosophila melanogaster]
Length=119

 GENE ID: 3885594 CG33998 | CG33998 gene product from transcript CG33998-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score =  121 bits (305),  Expect = 2e-26, Method: Composition-based stats.
```

```
 Identities = 20/120 (16%), Positives = 39/120 (32%), Gaps = 2/120 (1%)

Query  18   VDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY  77
            +  +V                 R+   + G R   D  + R   +     L+ V     +
Sbjct  1    MKIYVILLVFALTAFAAASGRGRSHSITWGARTYRDMHLHREIITEKSKFLRVVTREFVF  60

Query  78   PANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
                     ++ + I        G   Y+  GG       I ++++    FS+I +IYG+
Sbjct  61   DQK-KLARTITQIVITDQIRDGNGGYAYLTAGGPQTTYAKIHLKSQRNQGFSFIIDIYGV  119


>☐ ref|XP_001961218.1| G GF11112 [Drosophila ananassae]
 gb|EDV38040.1| G GF11112 [Drosophila ananassae]
Length=120

 GENE ID: 6493977 Dana\GF11112 | GF11112 gene product from transcript GF11112-RA
[Drosophila ananassae] (10 or fewer PubMed links)

 Score =  121 bits (304),  Expect = 3e-26, Method: Composition-based stats.
 Identities = 20/120 (16%), Positives = 41/120 (34%), Gaps = 1/120 (0%)

Query  18   VDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY  77
            +  ++  +       +       +   + G R   D  V R   +     L+ V     +
Sbjct  1    MKFWIVLSVVSLAAFVKAEGRGPSHSVTWGARAFRDMHVRREIITEKSKFLRVVTRDFVF  60

Query  78   PANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
              N      ++ + I       LG   Y+  GG     + I +++     FS+I +IY +
Sbjct  61   TQNPKLARTITQIVITDQVREGLGGYAYLTGGGPQSDFVKIHFKSQRNKGFSFIVDIYAI  120


>☐ ref|XP_002026810.1| G GL27028 [Drosophila persimilis]
 gb|EDW33792.1| G GL27028 [Drosophila persimilis]
Length=166

 GENE ID: 6601775 Dper\GL27028 | GL27028 gene product from transcript GL27028-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score =  121 bits (303),  Expect = 3e-26, Method: Composition-based stats.
 Identities = 50/138 (36%), Positives = 80/138 (57%), Gaps = 6/138 (4%)

Query  3    PDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA-----RAIHNSLGQRVSGDRLVA  57
            P   +   +IT+V++V  F   +PG+ + +L            ++  +LG RV+ D LVA
Sbjct  28   PLDSSTLREITVVKNVRLFAAQHPGLKIQRLAKQPATARGGVESVRYTLGARVTSDSLVA  87

Query  58   RNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGI  117
            +  D  S+ + Q+V + + YP  G  GA+++YV ++  Q ++ GN YV+ GGIGQR I I
Sbjct  88   QGADVFSYPASQDVSVQLTYPEEGA-GAIITYVELICTQDNSEGNAYVVAGGIGQRFISI  146

Query  118  LIEARNTYTFSYIAEIYG  135
            ++EA  T  FSY A+ YG
Sbjct  147  ILEASQTRNFSYQAQYYG  164


>☐ ref|XP_002059070.1| G GJ15191 [Drosophila virilis]
 gb|EDW58139.1| G GJ15191 [Drosophila virilis]
Length=162

 GENE ID: 6635434 Dvir\GJ15191 | GJ15191 gene product from transcript GJ15191-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  120 bits (302),  Expect = 3e-26, Method: Composition-based stats.
 Identities = 48/135 (35%), Positives = 85/135 (62%), Gaps = 5/135 (3%)

Query  5    PETVSAKITLVEDVDEFVRANPGVTLTKLDNV----RTARAIHNSLGQRVSGDRLVARNQ  60
               +   I +V++V +F   +PG+ + +L+       + +++  +LG RVSGDRLVA+
Sbjct  27   DADIHRDIVVVKNVRDFAAQHPGLKMQRLEKQPAVRASGQSVRYTLGARVSGDRLVAQGA  86

Query  61   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  120
            D  S+ + ++V L + YP +G  G++V++V ++  Q ++ GN YV+ GGIGQR I I++E
Sbjct  87   DVYSYPARKDVSLQLTYPESGA-GSIVTFVELICTQDNSEGNAYVVAGGIGQRFISIVVE  145

Query  121  ARNTYTFSYIAEIYG  135
            A+ T  F+Y A+ +G
Sbjct  146  AKQTENFAYQAQYFG  160
```

```
>☐ ref|XP_002015909.1| G GL10767 [Drosophila persimilis]
 gb|EDW31799.1| G GL10767 [Drosophila persimilis]
Length=120

 GENE ID: 6591185 Dper\GL10767 | GL10767 gene product from transcript GL10767-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score =  120 bits (302),  Expect = 4e-26, Method: Composition-based stats.
 Identities = 20/120 (16%), Positives = 40/120 (33%), Gaps = 1/120 (0%)

Query  18   VDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY  77
            +  ++                  +   + G R   D  + R+  +     L+ V   + +
Sbjct  1    MRVYIFLVIISVTALCQAAGRGPSHSVTWGSRSFRDMHLRRDIITEKSSFLRVVTRDLIF  60

Query  78   PANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
              N      ++ + I        G   Y+  GG       I  +++    FS+I +IYGL
Sbjct  61   KQNPKLARTITQIVITDQVRDGQGGYAYLQAGGPQTTYAKIHFKSQRNKGFSFIVDIYGL  120


>☐ ref|XP_001974726.1| G GG21920 [Drosophila erecta]
 gb|EDV55126.1| G GG21920 [Drosophila erecta]
Length=119

 GENE ID: 6547300 Dere\GG21920 | GG21920 gene product from transcript GG21920-RA
[Drosophila erecta] (10 or fewer PubMed links)

 Score =  120 bits (302),  Expect = 4e-26, Method: Composition-based stats.
 Identities = 24/119 (20%), Positives = 45/119 (37%), Gaps = 3/119 (2%)

Query  20   EFVRANPGVTLTKL-DNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
            +F      V+L+ L +     R+   + G R   D L+ R   +     L+ V     +
Sbjct  2    KFSVILLVVSLSALSEASGRGRSHSITWGARSYRDMLLRREIITEKSKFLRVVTKDYEFD  61

Query  79   ANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
                    ++ + I        G   Y+  GG       I ++++    FS+I +IYG+
Sbjct  62   QK-KLARTITQIVITDQIRDGNGGYAYLTAGGPQTTYAKIHLKSQRNQGFSFIIDIYGI  119


>☐ gb|ACI30042.1|  hypothetical secreted protein [Anopheles darlingi]
Length=119

 Score =  120 bits (302),  Expect = 4e-26, Method: Composition-based stats.
 Identities = 23/117 (19%), Positives = 46/117 (39%), Gaps = 1/117 (0%)

Query  20   EFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPA  79
            +       + +     +  A+  +   G R+  D L  +     S   L+   L ++YP+
Sbjct  2    KITICFTVLAVLGFSCLAVAQTHNYFFGARIPYDSLANQTTVIQSGSFLRVKSLNIDYPS  61

Query  80   NGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             G  G  ++ + +        G    +  GGIGQ    I ++++     ++  EIYG
Sbjct  62   KGQRGRNITAIYVYDRLGGGRGGYASITRGGIGQNFTRINLKSQRGNGMNFQVEIYG  118


>☐ ref|XP_002071202.1| G GK25257 [Drosophila willistoni]
 gb|EDW82188.1| G GK25257 [Drosophila willistoni]
Length=169

 GENE ID: 6648809 Dwil\GK25257 | GK25257 gene product from transcript GK25257-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score =  120 bits (301),  Expect = 5e-26, Method: Composition-based stats.
 Identities = 57/132 (43%), Positives = 82/132 (62%), Gaps = 5/132 (3%)

Query  8    VSAKITLVEDVDEFVRANPGVTLTKLDNVRTA----RAIHNSLGQRVSGDRLVARNQDSA  63
                IT+V+DV  F   +PG+ + +L     A    +++  SLG R+SGDRLVA+  D
Sbjct  37   SLRDITVVKDVRAFAAKHPGLKMQRLTKHPEARSSTQSVRYSLGARISGDRLVAQGADVF  96

Query  64   SWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARN  123
            S+ S+++V L + YP +G  GAVV+YV ++  Q +  GN YV+ GGIGQR I I++EA
Sbjct  97   SYPSVKDVSLQLTYPESG-TGAVVTYVELICSQDNTEGNAYVVAGGIGQRFISIVLEANQ  155

Query  124  TYTFSYIAEIYG  135
            T   FSY A+ YG
Sbjct  156  TKNFSYQAQYYG  167
```

```
>☐ ref|XP_002048773.1| G GJ21142 [Drosophila virilis]
 gb|EDW59966.1| G GJ21142 [Drosophila virilis]
Length=116

 GENE ID: 6626260 Dvir\GJ21142 | GJ21142 gene product from transcript GJ21142-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  119 bits (299),  Expect = 8e-26, Method: Composition-based stats.
 Identities = 20/115 (17%), Positives = 49/115 (42%), Gaps = 2/115 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L     ++A   + G+R S D L++R  +  +        + V++P +G
Sbjct  1    MKFLYSCVLLLALCAASQAYSATWGKRNSTDYLLSRQTEVRAPLKNSYWNINVDFPRSGT  60

Query  83   YG-AVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 ++ ++++ +  ++ G    +  GG G R   I + ++ +   +   EI+G
Sbjct  61   SNLYTITAINVIDNFRNSSGATPSLWSGGPGYRFAQINLRSQVSKGLNSTVEIWG  115


>☐ ref|XP_001354364.2| G GA13455 [Drosophila pseudoobscura pseudoobscura]
 gb|EAL31417.2| GA13455 [Drosophila pseudoobscura pseudoobscura]
Length=166

 GENE ID: 4814282 Dpse\GA13455 | GA13455 gene product from transcript GA13455-RA
[Drosophila pseudoobscura pseudoobscura] (10 or fewer PubMed links)

 Score =  119 bits (299),  Expect = 9e-26, Method: Composition-based stats.
 Identities = 51/138 (36%), Positives = 80/138 (57%), Gaps = 6/138 (4%)

Query  3    PDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA-----RAIHNSLGQRVSGDRLVA  57
            P   +   +IT+V++V  F   +PG+ + +L            ++  +LG RV+ D LVA
Sbjct  28   PLDSSTLREITVVKNVRLFAAQHPGLKIQRLAKQPATARGGVESVRYTLGARVTSDSLVA  87

Query  58   RNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGI  117
            +  D  S+ + Q+V + + YP  G  GA+V+YV ++  Q ++ GN YV+ GGIGQR I I
Sbjct  88   QGADVFSYPASQDVSVQLTYPEEGA-GAIVTYVELICTQDNSEGNAYVVAGGIGQRFISI  146

Query  118  LIEARNTYTFSYIAEIYG  135
            ++EA  T  FSY A+ YG
Sbjct  147  ILEASQTRNFSYQAQYYG  164


>☐ ref|XP_001230950.2| UG AGAP003713-PA [Anopheles gambiae str. PEST]
 gb|EAU76936.2| G AGAP003713-PA [Anopheles gambiae str. PEST]
Length=119

 GENE ID: 4577451 AgaP_AGAP003713 | AGAP003713-PA [Anopheles gambiae str. PEST]
(10 or fewer PubMed links)

 Score =  118 bits (297),  Expect = 2e-25, Method: Composition-based stats.
 Identities = 25/113 (22%), Positives = 49/113 (43%), Gaps = 1/113 (0%)

Query  24   ANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGY  83
                  +     +  +++ +   G R   D L+ R     S   LQ   + + YP  G
Sbjct  6    LAVVAAVLLFACLVASQSNNYFWGVRDPRDVLLNRTIAVRSGTILQVKSIDLVYPLKGQV  65

Query  84   GAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            G  +S +S++   T+  G    +  GGIG     + ++++  + +++I EIYG
Sbjct  66   GRNISAISVVDQYTNGKGGYASLYAGGIGYNYTTVHLKSQRGHGYNFIVEIYG  118


>☐ ref|XP_001977671.1| G GG18116 [Drosophila erecta]
 gb|EDV46598.1| G GG18116 [Drosophila erecta]
Length=156

 GENE ID: 6550161 Dere\GG18116 | GG18116 gene product from transcript GG18116-RA
[Drosophila erecta] (10 or fewer PubMed links)

 Score =  118 bits (297),  Expect = 2e-25, Method: Composition-based stats.
 Identities = 39/128 (30%), Positives = 69/128 (53%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWGS  67
               IT V D   +   NP + L +++ V+ ++     + G+R SGDR++  + D  S  S
Sbjct  28   YTDITPVTDETRYTTLNPDIQLNEMNKVKHSKGNIVFTAGERASGDRMIVNHFDGDSLTS  87

Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
```

```
             ++V++ ++YPA    G   ++ + + VD  ++   GY+  GGIGQ  + IL+ +  T +F
Sbjct  88   AKDVEVLMSYPAGASTGVTLTSIEVYVDMMADDAGGYLTKGGIGQTHVEILLTSNLTRSF  147

Query  128  SYIAEIYG  135
             Y   IYG
Sbjct  148  VYETFIYG  155


>☐ ref|XP_001354363.1| G GA13456 [Drosophila pseudoobscura pseudoobscura]
 ref|XP_002026811.1| G GL27029 [Drosophila persimilis]
 gb|EAL31416.1| G GA13456 [Drosophila pseudoobscura pseudoobscura]
 gb|EDW33793.1| G GL27029 [Drosophila persimilis]
Length=157

 GENE ID: 4814310 Dpse\GA13456 | GA13456 gene product from transcript GA13456-RA
[Drosophila pseudoobscura pseudoobscura] (10 or fewer PubMed links)

 Score =  118 bits (295),  Expect = 2e-25, Method: Composition-based stats.
 Identities = 40/128 (31%), Positives = 70/128 (54%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWGS  67
             + I+ V D   F    P V L++++ V+ ++     + G+R+SGDRLV  + D  S+ +
Sbjct  29   YSDISPVTDDKRFTHLYPEVELSEMNKVKHSKGNIVFTSGERISGDRLVVNHYDDESFAT  88

Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
              ++++ + YPA    G   ++ + I VD T++   GY+  GGIG+    IL+ +  T +F
Sbjct  89   AVDIEVLMRYPAESSTGVTLTCIEIYVDTTASDAGGYLTRGGIGETIAEILLTSNGTRSF  148

Query  128  SYIAEIYG  135
             Y   IYG
Sbjct  149  VYETFIYG  156


>☐ ref|XP_002138284.1| G GA24478 [Drosophila pseudoobscura pseudoobscura]
 gb|EDY68842.1| G GA24478 [Drosophila pseudoobscura pseudoobscura]
Length=133

 GENE ID: 6898218 Dpse\GA24478 | GA24478 gene product from transcript GA24478-RA
[Drosophila pseudoobscura pseudoobscura] (10 or fewer PubMed links)

 Score =  117 bits (293),  Expect = 4e-25, Method: Composition-based stats.
 Identities = 21/133 (15%), Positives = 41/133 (30%), Gaps = 14/133 (10%)

Query  18   VDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY  77
            +  ++                  +   + G R   D  + R+  +     L+ V   + +
Sbjct  1    MRVYIFLVLISVTALCQAAGRGPSHSVTWGSRSFRDMHLRRDIITEKSSFLRVVTRDLIF  60

Query  78   PANGG-------------YGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARN  123
              NG                  ++ + I        G   Y+  GG       I  +++
Sbjct  61   KQNGSAQRLIATCCQKSKLARTITQIVITDQVRDGQGGYAYLQAGGPQTTYAKIHFKSQR  120

Query  124  TYTFSYIAEIYGL  136
               FS+I +IYGL
Sbjct  121  NKGFSFIVDIYGL  133


>☐ gb|ACD81831.1|  IP21621p [Drosophila melanogaster]
Length=112

 Score =  116 bits (292),  Expect = 6e-25, Method: Composition-based stats.
 Identities = 19/113 (16%), Positives = 36/113 (31%), Gaps = 2/113 (1%)

Query  25   NPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYG  84
                           R+   + G R   D  + R   +     L+ V     +
Sbjct  1    LVFALTAFAAASGRGRSHSITWGARTYRDMHLREIITEKSKFLRVVTREFVFDQK-KLA  59

Query  85   AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
              ++ + I        G   Y+  GG       I  ++++   FS+I +IYG+
Sbjct  60   RTITQIVITDQIRDGNGGYAYLTAGGPQTTYAKIHLKSQRNQGFSFIIDIYGV  112


>☐ ref|XP_002071201.1| G GK25258 [Drosophila willistoni]
 gb|EDW82187.1| G GK25258 [Drosophila willistoni]
Length=156
```

```
 GENE ID: 6648808 Dwil\GK25258 | GK25258 gene product from transcript GK25258-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score =  116 bits (292),  Expect = 6e-25, Method: Composition-based stats.
 Identities = 35/128 (27%), Positives = 67/128 (52%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWGS  67
               IT V D   +   NP   L + + V+ ++     + G+R++GDRL+  + D  S+ +
Sbjct  28   YTAITPVSDETRYTTLNPNAELKEFNKVKHSKGNIVFTSGERITGDRLIVNHYDDESFTT  87

Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
              ++++ + YP++   GA ++ + I VD +++    Y++ GGIG+    IL+    T T+
Sbjct  88   TADIEVGMKYPSSSSTGATLTCIEIYVDMSADDAGAYMINGGIGKANAEILLTCNQTRTY  147

Query  128  SYIAEIYG  135
             Y    YG
Sbjct  148  VYETYFYG  155


>☐ ref|XP_001987728.1| GH22082 [Drosophila grimshawi]
 gb|EDW02595.1| GH22082 [Drosophila grimshawi]
Length=121

 GENE ID: 6560811 Dgri\GH22082 | GH22082 gene product from transcript GH22082-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  116 bits (291),  Expect = 8e-25, Method: Composition-based stats.
 Identities = 19/117 (16%), Positives = 44/117 (37%), Gaps = 2/117 (1%)

Query  21   FVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPAN  80
            F             +  T ++   + G R+  D  + R   +     L+ V    ++
Sbjct  6    FTTLLLVSLAAVSHSAGTNQSHSVAWGGRMFRDTHLERVIITEKSRFLRVVTRDYHFQQI  65

Query  81   GGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
            G     ++ + +     +  G   +++ GG     + I  +++    FS+I +IYG+
Sbjct  66   G-LQRRITQIVVTDQVRNGGGGYAHLVAGGPNSHFVKIHFKSQRNKGFSFIVDIYGI  121


>☐ ref|XP_002009841.1| GI15587 [Drosophila mojavensis]
 gb|EDW07158.1| GI15587 [Drosophila mojavensis]
Length=164

 GENE ID: 6584185 Dmoj\GI15587 | GI15587 gene product from transcript GI15587-RA
[Drosophila mojavensis] (10 or fewer PubMed links)

 Score =  116 bits (290),  Expect = 9e-25, Method: Composition-based stats.
 Identities = 50/137 (36%), Positives = 84/137 (61%), Gaps = 7/137 (5%)

Query  5    PETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTAR------AIHNSLGQRVSGDRLVAR  58
               V   I +V++V +FV  +PG+ L ++     AR      ++  +LG R++GDRLVA+
Sbjct  27   DADVHRDIIVVKNVRDFVAQHPGLKLQRMVKQPAARATTRSLSVRYTLGARINGDRLVAQ  86

Query  59   NQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGIL  118
              D   + + ++V L + YP +G  G++V++V ++  Q +N GN YV+ GGIGQR I I+
Sbjct  87   GADVFHYQARKDVSLQITYPESGA-GSIVTFVELICTQDNNEGNAYVVAGGIGQRFISIV  145

Query  119  IEARNTYTFSYIAEIYG  135
            +EA+ T  F+Y A+ +G
Sbjct  146  LEAKQTDDFAYQAQYFG  162


>☐ ref|XP_002082142.1| GD11406 [Drosophila simulans]
 gb|EDX07727.1| GD11406 [Drosophila simulans]
Length=119

 GENE ID: 6735167 Dsim\GD11406 | GD11406 gene product from transcript GD11406-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score =  116 bits (290),  Expect = 1e-24, Method: Composition-based stats.
 Identities = 19/120 (15%), Positives = 39/120 (32%), Gaps = 2/120 (1%)

Query  18   VDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY  77
            +   V          +      R+   + G R   D  + R   +     L+ V     +
Sbjct  1    MKISVILLVISLSAFVAASGRGRSHSITWGARSYRDMHLRREIITEKSKFLRVVTRDYVF  60

Query  78   PANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
```

```
                     ++ + I        G    ++ GG       I ++++    FS+I +IYG+
Sbjct  61   DQK-KLARTITQIVITDQIRDGNGGYAFLTAGGPQTTYAKIHLKSQRNQGFSFIIDIYGI  119


>☐  ref|XP_002034496.1|  GM21912 [Drosophila sechellia]
 gb|EDW48509.1|  GM21912 [Drosophila sechellia]
Length=119

 GENE ID: 6609844 Dsec\GM21912 | GM21912 gene product from transcript GM21912-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score =  115 bits (288),  Expect = 2e-24, Method: Composition-based stats.
 Identities = 19/120 (15%), Positives = 38/120 (31%), Gaps = 2/120 (1%)

Query  18   VDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY  77
            +   V                 R+   + G R   D  + R   +     L+ V     +
Sbjct  1    MKISVILLVISLSAFAAASGRGRSHSITWGARSYRDMHLRREIITEKSKFLRVVTRDYVF  60

Query  78   PANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
                     ++ + I        G    ++ GG       I ++++    FS+I +IYG+
Sbjct  61   DQK-KLARTITQIVITDQIRDGNGGYAFLTAGGPQTTYAKIHLKSQRNQGFSFIIDIYGI  119


>☐  ref|XP_001990970.1|  GH12337 [Drosophila grimshawi]
 gb|EDV99594.1|  GH12337 [Drosophila grimshawi]
Length=162

 GENE ID: 6564732 Dgri\GH12337 | GH12337 gene product from transcript GH12337-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  115 bits (288),  Expect = 2e-24, Method: Composition-based stats.
 Identities = 49/133 (36%), Positives = 84/133 (63%), Gaps = 5/133 (3%)

Query  7    TVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA----RAIHNSLGQRVSGDRLVARNQDS  62
             V  +I +V++V +F   +PG+ L +L          +++  +LG RVSGDRLVA+  D
Sbjct  29   DVQRQIVVVKNVRDFAAEHPGLKLQRLQKQPAGRMSGQSVRYTLGARVSGDRLVAQGADV  88

Query  63   ASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEAR  122
             S+ + ++V + + YP +G  G++V++V ++  Q ++ GN YV+ GGIGQR I I++EA+
Sbjct  89   FSYPAQKDVSVQLTYPESGA-GSIVTFVELICTQDNSEGNAYVVAGGIGQRFISIVVEAK  147

Query  123  NTYTFSYIAEIYG  135
             T +F Y A+ +G
Sbjct  148  QTNSFGYQAQYFG  160


>☐  gb|ABV60298.1|  14.3 kDa midgut protein [Lutzomyia longipalpis]
Length=154

 Score =  115 bits (288),  Expect = 2e-24, Method: Composition-based stats.
 Identities = 82/136 (60%), Positives = 98/136 (72%), Gaps = 2/136 (1%)

Query  3    PDPETVSAKITLVEDVDEFVRANPGVTLTKLDN--VRTARAIHNSLGQRVSGDRLVARNQ  60
             D E VSAKI +V+ + +F  ANP V L KLD+  VR  + IH SLG RV+GDRLV  ++
Sbjct  19   SDAEDVSAKIRIVDSISDFKLANPDVQLMKLDSQSVRAGKTIHYSLGGRVAGDRLVGTSR  78

Query  61   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  120
             S SW   QN+QL++NYP NGG G VV+YVSI V+QTSN G GYVL GGIGQR I I+IE
Sbjct  79   GSHSWAFPQNIQLSLNYPTNGGSGGVVTYVSIHVNQTSNQGRGYVLNGGIGQRFISIIIE  138

Query  121  ARNTYTFSYIAEIYGL  136
            A NT  FSY AEIYG+
Sbjct  139  ANNTLNFSYNAEIYGV  154


>☐  ref|XP_002009842.1|  GI15588 [Drosophila mojavensis]
 gb|EDW07159.1|  GI15588 [Drosophila mojavensis]
Length=170

 GENE ID: 6584186 Dmoj\GI15588 | GI15588 gene product from transcript GI15588-RA
[Drosophila mojavensis] (10 or fewer PubMed links)

 Score =  114 bits (285),  Expect = 4e-24, Method: Composition-based stats.
 Identities = 47/142 (33%), Positives = 75/142 (52%), Gaps = 7/142 (4%)

Query  1    TAPDPET------VSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGD  53
```

```
            + PDP++      V   I  V D + +   NP V LT L+ V+ A+     S G+R+SGD
Sbjct  28   SPPDPQSQRATDLVYTPIKPVTDENRYTTLNPDVKLTALNKVKHAKGSMVFSSGERISGD  87

Query  54   RLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQR  113
            +L+  N D  S+ +  +V++ ++Y +  G    ++ + I VD T++  + Y   GGIG
Sbjct  88   KLLVNNYDDESFTTAVDVEVVMSYSSKTGSATTLTSIEIYVDTTADDASAYFTAGGIGSS  147

Query  114  QIGILIEARNTYTFSYIAEIYG  135
            +  IL+    T TFSY A  YG
Sbjct  148  EASILLACNQTRTFSYEAVFYG  169


>☐ ref|XP_001990969.1| G GH12339 [Drosophila grimshawi]
 gb|EDV99593.1| G GH12339 [Drosophila grimshawi]
Length=166

 GENE ID: 6564731 Dgri\GH12339 | GH12339 gene product from transcript GH12339-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  113 bits (282),  Expect = 7e-24, Method: Composition-based stats.
 Identities = 42/128 (32%), Positives = 70/128 (54%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWGS  67
             + I  V +   F   NP V L  L  V+ A+     S G+R++GD+L+  N D  S+ +
Sbjct  38   YSPIQTVTNETRFTTLNPSVQLVPLKKVKHAKGNIVFSAGERITGDKLLVNNYDDESFTT  97

Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
              +V++ + YP++ G    ++ + I VD +++  N Y + GGIG+ +  IL++   T TF
Sbjct  98   AVDVEVQMRYPSSSGISITLTSIEIYVDASADDANAYFIDGGIGKHKASILLDCNLTRTF  157

Query  128  SYIAEIYG  135
            SY A  YG
Sbjct  158  SYEAFFYG  165


>☐ ref|XP_002061455.1| G GK20920 [Drosophila willistoni]
 gb|EDW72441.1| G GK20920 [Drosophila willistoni]
Length=121

 GENE ID: 6638174 Dwil\GK20920 | GK20920 gene product from transcript GK20920-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score =  112 bits (281),  Expect = 9e-24, Method: Composition-based stats.
 Identities = 19/115 (16%), Positives = 41/115 (35%), Gaps = 3/115 (2%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
            +        T  + V    +   + G R   D L++R   +     L+ V     +
Sbjct  8    LVLAIVSVATFCEGVGRGGSNSVTWGARKYRDTLLSREIITVKSRFLRVVTHDYVFRQK-  66

Query  82   GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             +   ++ + +     +  G    ++ GG       I  +++    FS+I +IYG
Sbjct  67   -WPRTITQIVVTDQIRNGEGGYATLVGGGPMATYAKIHFKSQRNNGFSFIVDIYG  120


>☐ ref|XP_002101618.1| G GE15520 [Drosophila yakuba]
 gb|EDX02726.1| G GE15520 [Drosophila yakuba]
Length=164

 GENE ID: 6525797 Dyak\GE15520 | GE15520 gene product from transcript GE15520-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score =  112 bits (281),  Expect = 1e-23, Method: Composition-based stats.
 Identities = 50/135 (37%), Positives = 76/135 (56%), Gaps = 7/135 (5%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRT-------ARAIHNSLGQRVSGDRLVARNQD  61
            S  I +V D+  F   +PG+ +  ++           ++ +  ++G R+ GD LVA+  D
Sbjct  29   STGIRVVNDIHVFAAQHPGLQIQPMEKEIVPGKARVGSQTVRYNMGARIPGDALVAQTAD  88

Query  62   SASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEA  121
            +  +   Q+V L + YP N    A VSY+ +L  Q S+ G  YV+ GGIGQR I I++EA
Sbjct  89   TFEYQRAQDVSLQLTYPENSSDAAAVSYLQLLCTQDSSEGTAYVVAGGIGQRLISIVLEA  148

Query  122  RNTYTFSYIAEIYGL  136
            +NT  FSY AE YG+
Sbjct  149  KNTKYFSYQAEYYGV  163
```

```
>☐ ref|XP_002050287.1| G GJ20290 [Drosophila virilis]
 gb|EDW61480.1| G GJ20290 [Drosophila virilis]
Length=121

 GENE ID: 6626912 Dvir\GJ20290 | GJ20290 gene product from transcript GJ20290-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  112 bits (280),  Expect = 1e-23, Method: Composition-based stats.
 Identities = 22/118 (18%), Positives = 47/118 (39%), Gaps = 3/118 (2%)

Query  21   FVRANPGVTLTKLD-NVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPA  79
            F+     ++L  L  +    ++   + G R+  D  + R   +     L+ V    N+
Sbjct  5    FLAILGLLSLVALAQSEGQNQSHSVTWGARIFRDMHLERAIITEKSKFLRVVSRYYNFTQ  64

Query  80   NGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
             G Y   ++ + +     +  G    +  GG       I  +++ +  FS+I +IYG+
Sbjct  65   IG-YQRRITQIVVTDQIRNGQGGFARLEAGGPQATFAKIYFKSQRSKGFSFIVDIYGI  121


>☐ ref|XP_002004523.1| G GI19567 [Drosophila mojavensis]
 gb|EDW08458.1| G GI19567 [Drosophila mojavensis]
Length=118

 GENE ID: 6578615 Dmoj\GI19567 | GI19567 gene product from transcript GI19567-RA
[Drosophila mojavensis] (10 or fewer PubMed links)

 Score =  112 bits (280),  Expect = 1e-23, Method: Composition-based stats.
 Identities = 22/115 (19%), Positives = 48/115 (41%), Gaps = 2/115 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG-  81
                   L  L     ++A  +S G+R + D L++R  +  +        + ++YP  G
Sbjct  1    MKFLYSCLLLLAVFAASQAFSSSWGKRNNTDILLSRQNERRAPLKNNYWSVNIDYPRAGT  60

Query  82   GYGAVVSYVSILVDQTS-NLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 +S+++++ +  + +     +  GG G R   I + ++ +       EIYG
Sbjct  61   SNNYTISFINVIDNFRNTSGATPSLWSGGPGYRYAQINLRSQVSRGIDSTVEIYG  115


>☐ ref|XP_001986843.1| G GH21597 [Drosophila grimshawi]
 gb|EDW01710.1| G GH21597 [Drosophila grimshawi]
Length=116

 GENE ID: 6560462 Dgri\GH21597 | GH21597 gene product from transcript GH21597-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  111 bits (279),  Expect = 2e-23, Method: Composition-based stats.
 Identities = 22/115 (19%), Positives = 48/115 (41%), Gaps = 2/115 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L     ++A   S G+R S D L++R  +           + V++P +G
Sbjct  1    MKFQYSCVLLLALFAASQAYSASWGKRNSNDYLLSRQTEVRYPIKNNYWNVNVDFPRSGT  60

Query  83   YG-AVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 +++V+++ +  ++ G    +  GG G R + I + +      +   EI+G
Sbjct  61   TNMYSITFVNVIDNFRNSSGATPSLWSGGPGYRFVQINLRSLVGKGLNSTVEIWG  115


>☐ gb|AAQ09813.1|  CG15043 [Drosophila yakuba]
Length=164

 Score =  111 bits (279),  Expect = 2e-23, Method: Composition-based stats.
 Identities = 49/135 (36%), Positives = 76/135 (56%), Gaps = 7/135 (5%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRT-------ARAIHNSLGQRVSGDRLVARNQD  61
            S  I +V D+  F   +PG+ +  ++           ++ +  ++G R+ GD LVA+  D
Sbjct  29   STGIRVVNDIHVFAAQHPGLQIQPMEKEIVPGKARVGSQTVRYNMGARIPGDALVAQTAD  88

Query  62   SASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEA  121
            +  +   Q+V L + YP N    A VSY+ ++  Q S+ G  YV+ GGIGQR I I++EA
Sbjct  89   TFEYQRAQDVSLQLTYPENSSDAAAVSYLQLVCTQDSSEGTAYVVAGGIGQRLISIVLEA  148

Query  122  RNTYTFSYIAEIYGL  136
            +NT  FSY AE YG+
Sbjct  149  KNTKYFSYQAEYYGV  163
```

```
>☐ ref|XP_001965516.1| G GF22411 [Drosophila ananassae]
 gb|EDV35031.1| G GF22411 [Drosophila ananassae]
Length=157

 GENE ID: 6505072 Dana\GF22411 | GF22411 gene product from transcript GF22411-RA
[Drosophila ananassae] (10 or fewer PubMed links)

 Score =  111 bits (279),  Expect = 2e-23, Method: Composition-based stats.
 Identities = 59/135 (43%), Positives = 84/135 (62%), Gaps = 9/135 (6%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRT--------ARAIHNSLGQRVSGDRLVARNQ  60
            SA+I +V+D+ EFV  +PGV L  ++            +  +  +LG R+ GDRLVA+
Sbjct  23   SARILVVKDIKEFVVQHPGVHLQAMEKEVVPSKAREAGSLTVRYNLGARIGGDRLVAQGA  82

Query  61   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  120
            D+ ++  LQ+V L + YP +G  GA+VSYV IL  Q  + G+ YV+ GGIGQR I I++E
Sbjct  83   DTFNYSQLQDVSLQLTYPESGA-GAIVSYVQILCTQDDSDGSAYVIAGGIGQRLISIVLE  141

Query  121  ARNTYTFSYIAEIYG  135
            A+NT  FSY  E YG
Sbjct  142  AKNTRNFSYHVEYYG  156


>☐ ref|XP_001965515.1| G GF22412 [Drosophila ananassae]
 gb|EDV35030.1| G GF22412 [Drosophila ananassae]
Length=162

 GENE ID: 6505073 Dana\GF22412 | GF22412 gene product from transcript GF22412-RA
[Drosophila ananassae] (10 or fewer PubMed links)

 Score =  111 bits (279),  Expect = 2e-23, Method: Composition-based stats.
 Identities = 40/128 (31%), Positives = 71/128 (55%), Gaps = 1/128 (0%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIH-NSLGQRVSGDRLVARNQDSASWGS  67
              +IT V D   +   +P V LT++D V+ ++     + G+R  GDR++  + DS S+ S
Sbjct  34   YTEITPVTDDTRYTTIHPEVVLTEMDKVKHSKGNVVFTSGERTKGDRVIVNHYDSESFAS  93

Query  68   LQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTF  127
              + ++ +NYP+    G  ++ + I V+ ++N   GY+  GGIG   + +L+ +  T TF
Sbjct  94   AVDTEVLMNYPSGSSTGITLTCIEIYVEMSANDTGGYITSGGIGNFNVEVLLTSNLTRTF  153

Query  128  SYIAEIYG  135
            +Y   IYG
Sbjct  154  AYETFIYG  161


>☐ ref|XP_001238394.2| UG AGAP011630-PA [Anopheles gambiae str. PEST]
 gb|EAU75892.2| G AGAP011630-PA [Anopheles gambiae str. PEST]
Length=120

 GENE ID: 4577506 AgaP_AGAP011630 | AGAP011630-PA [Anopheles gambiae str. PEST]
(10 or fewer PubMed links)

 Score =  111 bits (278),  Expect = 2e-23, Method: Composition-based stats.
 Identities = 24/103 (23%), Positives = 43/103 (41%), Gaps = 1/103 (0%)

Query  34   DNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSIL  93
              + TA+  +   G R+  D LV +     S   L+     ++YP  G  G  ++ + +
Sbjct  17   ACMVTAQTHNYFFGSRIPYDSLVNQTTVIQSSSFLRVKSAYLDYPLKGQRGRNITAIYVY  76

Query  94   VDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                S  G    +  GGIG+    I ++ +     ++  EIYG
Sbjct  77   DRLGSGRGGYASITSGGIGKNYTRINLKTQRGNGMNFQVEIYG  119


>☐ ref|XP_001360670.1| G GA12205 [Drosophila pseudoobscura pseudoobscura]
 gb|EAL25245.1| G GA12205 [Drosophila pseudoobscura pseudoobscura]
Length=116

 GENE ID: 4804045 Dpse\GA12205 | GA12205 gene product from transcript GA12205-RA
[Drosophila pseudoobscura pseudoobscura] (10 or fewer PubMed links)

 Score =  111 bits (278),  Expect = 2e-23, Method: Composition-based stats.
 Identities = 21/115 (18%), Positives = 41/115 (35%), Gaps = 2/115 (1%)
```

```
Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                      +  +        S G+R S D L++R  +  S        + V++P  G
Sbjct  1    MKFLLSLTILVAAIFVVYGSSASWGRRNSNDYLLSRQVEVRSPIKNNYWNVNVDFPRAGT  60

Query  83   YG-AVVSYVSILVDQTS-NLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 +S + +  +  + +     +  GG G R   I + ++     +   EIYG
Sbjct  61   VNYYNISAIFVYDNFKNTSGATPALWSGGPGFRFAQINLRSQVNRGLNSTIEIYG  115


>☐ gb|ABR27888.1|  putative salivary secreted protein [Triatoma infestans]
Length=136

 Score =  111 bits (277),  Expect = 3e-23, Method: Composition-based stats.
 Identities = 29/122 (23%), Positives = 52/122 (42%), Gaps = 6/122 (4%)

Query  21   FVRANPGVTLTKL----DNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVN  76
            F        +        N    ++ + + G R  GD+L+ + +  + W  L  VQ  V
Sbjct  12   FTVGIVIALIIPSPAWAANKCEGKSHNATFGMRKPGDKLIHKERIKSVWKLLSFVQKDVT  71

Query  77   YPANGGYG-AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
            YPA       +++Y+ I    T   G    ++ GG+G   + I  +++ T    +I EIY
Sbjct  72   YPAKDKKRKYIITYIKITDRYTDGHGGCASIVKGGVGYDHVKIHTKSQFTRGLDFIIEIY  131

Query  135  GL  136
            G+
Sbjct  132  GI  133


>☐ ref|XP_002016256.1| G GL10591 [Drosophila persimilis]
 gb|EDW32146.1| G GL10591 [Drosophila persimilis]
Length=116

 GENE ID: 6591102 Dper\GL10591 | GL10591 gene product from transcript GL10591-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score =  110 bits (276),  Expect = 3e-23, Method: Composition-based stats.
 Identities = 22/115 (19%), Positives = 41/115 (35%), Gaps = 2/115 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                      +  +        S G+R S D L++R  +  S        + V++P  G
Sbjct  1    MKFLLSLTILVAAIFVVYGSSASWGRRNSNDYLLSRQVEVRSPIKNNYWNVNVDFPRAGT  60

Query  83   YG-AVVSYVSILVDQTSNLGNG-YVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 +S + +  +  +  G    +  GG G R   I + ++     +   EIYG
Sbjct  61   VNYYNISAIFVYDNFKNTSGAAPALWSGGPGFRFAQINLRSQVNRGLNSTIEIYG  115


>☐ ref|XP_002062967.1| G GK21639 [Drosophila willistoni]
 gb|EDW73953.1| G GK21639 [Drosophila willistoni]
Length=116

 GENE ID: 6640510 Dwil\GK21639 | GK21639 gene product from transcript GK21639-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score =  110 bits (276),  Expect = 4e-23, Method: Composition-based stats.
 Identities = 20/115 (17%), Positives = 46/115 (40%), Gaps = 2/115 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                  ++  L  +  +     S G+R + D L++R  +  +        + V++P  G
Sbjct  1    MKFLLSSVFLLALLGASYGYSASWGKRANNDYLLSRQIEVRNPIKNNYWSVNVDFPRAGT  60

Query  83   YG-AVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 +S V +  +  ++ G +  +  GG G R   + + +  +   +   EI+G
Sbjct  61   VNYYNISAVFVYDNFKNSSGASPSLWSGGPGYRFATVNLRSLVSKGINSTVEIWG  115


>☐ gb|ABR27890.1|  putative salivary secreted peptide [Triatoma infestans]
Length=136

 Score =  110 bits (274),  Expect = 7e-23, Method: Composition-based stats.
 Identities = 29/122 (23%), Positives = 52/122 (42%), Gaps = 6/122 (4%)

Query  21   FVRANPGVTLTKL----DNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVN  76
            F        +        N    ++ + + G R  GD+L+ + +  + W  L  VQ  V
Sbjct  12   FTVGIVIALIIPSPAWAANKCEGKSHNVTFGMRTPGDKLIHKERIKSIWKLLSFVQKDVT  71
```

```
Query  77   YPANGGYG-AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
            YPA       +++Y+ I    T   G    ++ GG+G   + I  +++ T    +I EIY
Sbjct  72   YPAKDKKRKYIITYIKITDRYTDGHGGCASIVKGGVGYDHVKIHTKSQFTRGLDFIIEIY  131

Query  135  GL  136
            G+
Sbjct  132  GI  133


>☐ ref|XP_002059069.1| G GJ15192 [Drosophila virilis]
 gb|EDW58138.1| G GJ15192 [Drosophila virilis]
Length=140

 GENE ID: 6635433 Dvir\GJ15192 | GJ15192 gene product from transcript GJ15192-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  109 bits (272),  Expect = 1e-22, Method: Composition-based stats.
 Identities = 42/129 (32%), Positives = 72/129 (55%), Gaps = 1/129 (0%)

Query  8    VSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWG  66
            V   IT V D   +   NPGV L  L+ V+ A+     S G+R++GD+L+  + D  S+
Sbjct  11   VYTPITSVTDATRYTSKNPGVELLALNKVKHAKGNIVFSAGERITGDKLLVNSYDDESFT  70

Query  67   SLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYT  126
            +  ++++ + Y ++ G G  ++ + I VD +++  + Y + GGIG+ +  IL+    T T
Sbjct  71   TTVDIEVEMTYSSSSGTGVTLTSIEIYVDTSADDADAYFINGGIGKSKASILLACNQTRT  130

Query  127  FSYIAEIYG  135
            FSY A  YG
Sbjct  131  FSYEAFFYG  139


>☐ ref|XP_002059818.1| G GJ15031 [Drosophila virilis]
 gb|EDW57155.1| G GJ15031 [Drosophila virilis]
Length=118

 GENE ID: 6636459 Dvir\GJ15031 | GJ15031 gene product from transcript GJ15031-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  109 bits (272),  Expect = 1e-22, Method: Composition-based stats.
 Identities = 17/114 (14%), Positives = 38/114 (33%), Gaps = 1/114 (0%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                G+ L  L             G+    D  +A++  S ++           +  +
Sbjct  4    VTLVGLVLCCLFVTSAYAGNDYLWGEIGPNDYQLAKDTVSKAFFLGLVQTKKYVFKQSNN  63

Query  83   YGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              A  ++ + +   + +N     +  GG G +   I   ++  Y    + EI+G
Sbjct  64   LDALTITAIRVTDKKKNNGATAVLTAGGPGSKGATIQFTSQRGYGIKDLVEIWG  117


>☐ ref|XP_001997725.1| G GH17985 [Drosophila grimshawi]
 gb|EDW04958.1| G GH17985 [Drosophila grimshawi]
Length=116

 GENE ID: 6571537 Dgri\GH17985 | GH17985 gene product from transcript GH17985-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  108 bits (271),  Expect = 2e-22, Method: Composition-based stats.
 Identities = 21/108 (19%), Positives = 49/108 (45%), Gaps = 2/108 (1%)

Query  30   LTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYG-AVVS  88
            +  L     ++A   S G++ S D L++R  +  +        + V++P +G      ++
Sbjct  8    VLLLAFFAASQAYSASWGKKNSNDYLLSRQTEVRNPIKNNYWNVNVDFPRSGTTNMYNIT  67

Query  89   YVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            +V+++ +  ++ G    +  GG G R + I + +  +       EI+G
Sbjct  68   FVNVIDNFRNSSGATPSLWSGGPGYRFVQINLRSLVSQGMDSTVEIWG  115


>☐ ref|XP_001977672.1| G GG18115 [Drosophila erecta]
 gb|EDV46599.1| G GG18115 [Drosophila erecta]
Length=163

 GENE ID: 6550162 Dere\GG18115 | GG18115 gene product from transcript GG18115-RA
```

```
[Drosophila erecta] (10 or fewer PubMed links)

 Score =  108 bits (271),  Expect = 2e-22, Method: Composition-based stats.
 Identities = 53/135 (39%), Positives = 78/135 (57%), Gaps = 8/135 (5%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRT---ARAIHNSL----GQRVSGDRLVARNQD  61
            S  + +V D+ EF   +PG+ + +++       ARA   ++    G R+ GD LVA+  D
Sbjct  29   STGVRVVSDIHEFAAQHPGLQIQQMEKELVPGKARAGSLTVRYNMGARIPGDALVAQTAD  88

Query  62   SASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEA  121
            +  +   Q+V L + YP  G  GA VSYV ++  Q SN G  YV+ GGIGQ  I I++EA
Sbjct  89   TYEYPRAQDVSLQLTYPEKGK-GAAVSYVELVCTQDSNEGTAYVVAGGIGQTLISIVVEA  147

Query  122  RNTYTFSYIAEIYGL  136
            +NT  FSY A+ YG+
Sbjct  148  KNTKNFSYQAQYYGV  162


>☐ gb|AAV84229.1|  unknown [Culicoides sonorensis]
Length=168

 Score =  108 bits (271),  Expect = 2e-22, Method: Composition-based stats.
 Identities = 65/136 (47%), Positives = 88/136 (64%), Gaps = 3/136 (2%)

Query  2    APDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIH--NSLGQRVSGDRLVARN  59
              +P  + AKI LV+D+++F  A+P + LT+L             +LG RV GD LVA
Sbjct  33   EDNPVPLDAKIVLVDDINKFKAAHPDLQLTELVRSEKQPRNSIRYTLGARVVGDTLVATK  92

Query  60   QDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILI  119
             D AS+G+ QNV+LT+ YP +G  GAVV++V + V Q+SN+G GYV  GGIGQR I I+I
Sbjct  93   SDFASYGTAQNVKLTLTYPTSG-TGAVVTHVLVDVQQSSNVGKGYVTGGGIGQRYIQIVI  151

Query  120  EARNTYTFSYIAEIYG  135
            EA++T  F Y A IYG
Sbjct  152  EAQSTTYFGYSAYIYG  167


>☐ ref|XP_002005185.1| G GI19218 [Drosophila mojavensis]
 gb|EDW09120.1| G GI19218 [Drosophila mojavensis]
Length=121

 GENE ID: 6579289 Dmoj\GI19218 | GI19218 gene product from transcript GI19218-RA
[Drosophila mojavensis] (10 or fewer PubMed links)

 Score =  108 bits (270),  Expect = 2e-22, Method: Composition-based stats.
 Identities = 20/119 (16%), Positives = 42/119 (35%), Gaps = 3/119 (2%)

Query  20   EFVRANPGVTLTKLD-NVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
             ++      +L  +        +   + G R   D  + R   S     L+ V  T  +
Sbjct  4    VYLALLVSFSLLAVGRAAGKNESHSVTWGSRSFRDMHLERVIISEKSKFLRVVTRTYTFK  63

Query  79   ANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYGL  136
              G    +++ + +     +  G    +  GG       I  +++    FS+I +IYG+
Sbjct  64   QTGVQ-RLITQIVVTDQIKNGNGGYARLEAGGPNTNFAKIYFKSQRNKEFSFIVDIYGI  121


>☐ ref|XP_001660534.1| UG hypothetical protein AaeL_AAEL009985 [Aedes aegypti]
 gb|EAT38090.1| G conserved hypothetical protein [Aedes aegypti]
Length=158

 GENE ID: 5572717 AaeL_AAEL009985 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score =  108 bits (269),  Expect = 3e-22, Method: Composition-based stats.
 Identities = 64/136 (47%), Positives = 90/136 (66%), Gaps = 2/136 (1%)

Query  1    TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA-RAIHNSLGQRVSGDRLVARN  59
               +P  + AKI +V+D+  F  A+P + +  L+  R+  + I  +LG R+SGDRLV  +
Sbjct  23   AEEEPLQIDAKIEVVQDIAAFKAAHPELDVVPLEVARSTRQQIVYTLGNRLSGDRLVGTS  82

Query  60   QDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILI  119
            QD  SW +LQ+V L + YP  G  GAVVSYV ++V+Q+SN G GYV+ GGIGQR I ++I
Sbjct  83   QDGTSWSNLQDVTLNLRYPQAGS-GAVVSYVQVVVNQSSNQGRGYVVSGGIGQRYIQLVI  141

Query  120  EARNTYTFSYIAEIYG  135
            EA +T  F Y A+IYG
Sbjct  142  EAYSTSYFQYNAQIYG  157
```

```
>☐ ref|XP_002050017.1| GJ21904 [Drosophila virilis]
 gb|EDW61210.1| GJ21904 [Drosophila virilis]
Length=118

 GENE ID: 6625994 Dvir\GJ21904 | GJ21904 gene product from transcript GJ21904-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  107 bits (268),  Expect = 3e-22, Method: Composition-based stats.
 Identities = 27/117 (23%), Positives = 48/117 (41%), Gaps = 3/117 (2%)

Query  21   FVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPAN  80
            F    P   L  L  V  A ++ ++ G   +  +L+            + V   + YP N
Sbjct  2    FFSVLPSFLLVFLCVVP-AYSVSSTWGNISNSAQLLHAENVVYGSSPGKYVDHEIKYPTN  60

Query  81   G-GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            G G G +++ +      T+N G    +  GG G   + I ++++  Y   +  EIYG
Sbjct  61   GIGNGRIITGIRAFDQVTNNTGGHATIYSGGPGFNFVNIKLQSQYNYGLIFRVEIYG  117


>☐ ref|XP_001986842.1| GH21596 [Drosophila grimshawi]
 gb|EDW01709.1| GH21596 [Drosophila grimshawi]
Length=116

 GENE ID: 6560029 Dgri\GH21596 | GH21596 gene product from transcript GH21596-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  107 bits (267),  Expect = 4e-22, Method: Composition-based stats.
 Identities = 21/108 (19%), Positives = 49/108 (45%), Gaps = 2/108 (1%)

Query  30   LTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYG-AVVS  88
            +  L     ++A  +S G++ S D L++R  +           + V++P +G      ++
Sbjct  8    VLLLAFFAASQAYSSSWGKKNSNDYLLSRQTEVRYPIKNNYWNVNVDFPRSGTTNMYNIT  67

Query  89   YVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            +V+++ +  ++ G    +  GG G R + I + +  +       EI+G
Sbjct  68   FVNVIDNFRNSSGATPSLWSGGPGYRFVQINLRSLVSQGMDSTVEIWG  115


>☐ ref|XP_002004350.1| GI19667 [Drosophila mojavensis]
 gb|EDW08285.1| GI19667 [Drosophila mojavensis]
Length=120

 GENE ID: 6578439 Dmoj\GI19667 | GI19667 gene product from transcript GI19667-RA
[Drosophila mojavensis] (10 or fewer PubMed links)

 Score =  107 bits (267),  Expect = 5e-22, Method: Composition-based stats.
 Identities = 16/119 (13%), Positives = 38/119 (31%), Gaps = 1/119 (0%)

Query  18   VDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY  77
            +        G+ L  +             G+    D  +A++  S ++           +
Sbjct  1    MGFHSITLVGLLLCAVGLATVHAGNDYLWGEIGPNDYHLAKDTVSKAFFVGLVQTKKYVF  60

Query  78   PANGGYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                   A  ++ + +   + S+     ++ GG G +   I   +   Y    + EI+G
Sbjct  61   KQKDNLDALTITAIKVTDKKKSHGATAVLVSGGPGSKGATIQFTSERGYGIKDVVEIWG  119


>☐ ref|NP_573300.1| UG CG15043 [Drosophila melanogaster]
 gb|AAF48850.1| CG15043 [Drosophila melanogaster]
 gb|AAM29471.1| RE41194p [Drosophila melanogaster]
 gb|ACL86770.1| CG15043-PA [synthetic construct]
 gb|ACL91442.1| CG15043-PA [synthetic construct]
Length=163

 GENE ID: 32835 CG15043 | CG15043 gene product from transcript CG15043-RA
[Drosophila melanogaster] (Over 10 PubMed links)

 Score =  107 bits (267),  Expect = 5e-22, Method: Composition-based stats.
 Identities = 48/134 (35%), Positives = 76/134 (56%), Gaps = 8/134 (5%)

Query  10   AKITLVEDVDEFVRANPGVTLTKLDNVRT-------ARAIHNSLGQRVSGDRLVARNQDS  62
              I +V+++  F    PG+ + +++           ++ +  ++G R+ GD LVA+  ++
```

```
Sbjct  30   TGIQVVDNIHVFAAQYPGLQIQQMEKEIVPAKARVGSQTVRYNMGARIPGDELVAQTANT  89

Query  63   ASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEAR  122
              +   Q+V L + YP NG  GA VSYV +L  Q +N G  YV+ GGIGQ  I I++EA+
Sbjct  90   YEFPRAQDVSLQLTYPENGK-GATVSYVELLCTQDTNEGTAYVVAGGIGQSLISIVLEAK  148

Query  123  NTYTFSYIAEIYGL  136
            NT  FSY A  YG+
Sbjct  149  NTKNFSYQALYYGV  162


>☐ gb|AAV84228.1|  unknown [Culicoides sonorensis]
Length=168

 Score =  106 bits (266),  Expect = 6e-22, Method: Composition-based stats.
 Identities = 64/136 (47%), Positives = 87/136 (63%), Gaps = 3/136 (2%)

Query  2    APDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIH--NSLGQRVSGDRLVARN  59
              +P  + AKI LV+D+++F  A+P +  T+L             +LG RV GD LVA
Sbjct  33   EDNPVPLDAKIVLVDDINKFRAAHPDLQFTELVRSEKQPRNSIRYTLGARVVGDTLVATK  92

Query  60   QDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILI  119
             D AS+G+ QNV+LT+ YP +G  GAVV++V + V Q+SN+G GYV  GGIGQR I I+I
Sbjct  93   SDFASYGTAQNVKLTLTYPTSG-TGAVVTHVLVDVQQSSNVGKGYVTAGGIGQRYIQIVI  151

Query  120  EARNTYTFSYIAEIYG  135
            EA++T  F Y A IYG
Sbjct  152  EAQSTTYFGYSAYIYG  167


>☐ ref|XP_001868963.1| UG 14.5 kDa salivary protein [Culex quinquefasciatus]
 gb|EDS28050.1| G 14.5 kDa salivary protein [Culex quinquefasciatus]
Length=166

 GENE ID: 6052739 CpipJ_CPIJ018795 | 14.5 kDa salivary protein
[Culex quinquefasciatus]

 Score =  106 bits (266),  Expect = 6e-22, Method: Composition-based stats.
 Identities = 57/137 (41%), Positives = 93/137 (67%), Gaps = 4/137 (2%)

Query  2    APDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA---RAIHNSLGQRVSGDRLVAR  58
              D  + + +I L++D++ F  AN  + +T L+ +++    + I  + GQR++GD LV+
Sbjct  29   DSDQTSTAVRIDLIDDIEAFRAANANLKVTPLERIKSNDVRQQIIYAQGQRIAGDSLVST  88

Query  59   NQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGIL  118
            N +S  W  LQ+V + ++YPA+G  GAV++YV ++V+Q+SN G GYV+ GGIGQR I ++
Sbjct  89   NSESKQWPRLQDVTMNLSYPASG-VGAVITYVQVIVNQSSNSGKGYVIAGGIGQRFIHLV  147

Query  119  IEARNTYTFSYIAEIYG  135
            IEA +T  FSY ++IYG
Sbjct  148  IEAHSTSYFSYNSQIYG  164


>☐ ref|XP_002063501.1| G GK21373 [Drosophila willistoni]
 gb|EDW74487.1| G GK21373 [Drosophila willistoni]
Length=124

 GENE ID: 6640062 Dwil\GK21373 | GK21373 gene product from transcript GK21373-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score =  106 bits (265),  Expect = 7e-22, Method: Composition-based stats.
 Identities = 16/110 (14%), Positives = 37/110 (33%), Gaps = 1/110 (0%)

Query  27   GVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGA-  85
             V          +       G+    D  +A++    ++        T  +       A
Sbjct  14   LVVCIAYLVPGASAGNDYLWGEIGPNDYKLAKDTVKKAFFVGLVQTSTYTFTQKDNLNAL  73

Query  86   VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             ++ + +   + SN     ++ GG G +   I  +++  Y    + EI+G
Sbjct  74   TITAIQVTDKKKSNGATAQLISGGPGSKGATIKFKSKRGYGIKDVVEIWG  123


>☐ gb|ACT79296.1|  putative fatbody protein 3Rev-G1 [Bombyx mori]
Length=118

 Score =  106 bits (265),  Expect = 7e-22, Method: Composition-based stats.
```

```
 Identities = 23/114 (20%), Positives = 48/114 (42%), Gaps = 2/114 (1%)

Query  24   ANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGY  83
            A   + +  L +  +A++    LGQ    D ++ +  +             + YP  G
Sbjct  3    AFTVLLIACLLSACSAQSHDLGLGQLTYRDIVLYKINEYKYGFPFIVRTSEIEYPEPGQQ  62

Query  84   G-AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              A +S + +    T   G    V  GG+GQ+ + + ++++  + F++   IYG
Sbjct  63   NFAYISAIYVKDHYTDGNGGYPIVKSGGVGQKFVKLKLKSQRGHGFNFTVTIYG  116


>☐ ref|XP_001975823.1| G GG22531 [Drosophila erecta]
 gb|EDV56223.1| G GG22531 [Drosophila erecta]
Length=112

 GENE ID: 6549589 Dere\GG22531 | GG22531 gene product from transcript GG22531-RA
[Drosophila erecta] (10 or fewer PubMed links)

 Score =  106 bits (265),  Expect = 7e-22, Method: Composition-based stats.
 Identities = 24/114 (21%), Positives = 45/114 (39%), Gaps = 4/114 (3%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 +    L  +    A   + G+RV+ D L++R ++  +        L VNYPA
Sbjct  1    MRLLLILSVVLAVILGCHAYSATWGRRVNNDFLLSRTREVRNPIKNNYWNLNVNYPA---  57

Query  83   YGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                +S V +  +  ++ G +  +  GG G R   + +  +         EI+G
Sbjct  58   GFYNISAVIVYDNFKNSSGASPSLYSGGPGYRFATVNLRGQVNRGIDSTVEIWG  111


>☐ ref|XP_002005815.1| G GI18871 [Drosophila mojavensis]
 gb|EDW09750.1| G GI18871 [Drosophila mojavensis]
Length=119

 GENE ID: 6579942 Dmoj\GI18871 | GI18871 gene product from transcript GI18871-RA
[Drosophila mojavensis] (10 or fewer PubMed links)

 Score =  106 bits (265),  Expect = 7e-22, Method: Composition-based stats.
 Identities = 26/117 (22%), Positives = 49/117 (41%), Gaps = 2/117 (1%)

Query  21   FVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPAN  80
            FV       L  L  V    ++ ++ G   +  +L+       +    + V+  + YP N
Sbjct  2    FVSVVVPCFLLVLYCVAPVLSVSSTWGNISNSAQLLHAENVVYTSSPEKYVEHEIKYPTN  61

Query  81   G-GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            G G G +++ +      T+N G    +  GG G   + I ++++  Y   +  EIYG
Sbjct  62   GIGNGRIITGIRAFDQVTNNTGGHATIYSGGPGFNFVNIKLQSQYNYGLIFRVEIYG  118


>☐ gb|ABV44744.1|  14.4 kDa midgut protein [Phlebotomus papatasi]
Length=153

 Score =  106 bits (265),  Expect = 7e-22, Method: Composition-based stats.
 Identities = 125/136 (91%), Positives = 131/136 (96%), Gaps = 0/136 (0%)

Query  1    TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQ  60
            TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDN+RTA AIH SLGQRVSGDRLVAR Q
Sbjct  18   TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNLRTAGAIHYSLGQRVSGDRLVARGQ  77

Query  61   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  120
            DSASWGS Q++QLT+NYPANGGYGAVVSYVS+LVDQTSNLG GYVLYGGIGQRQIGILIE
Sbjct  78   DSASWGSPQDLQLTLNYPANGGYGAVVSYVSVLVDQTSNLGGGYVLYGGIGQRQIGILIE  137

Query  121  ARNTYTFSYIAEIYGL  136
            ARNTYTFSYIAE+YGL
Sbjct  138  ARNTYTFSYIAEVYGL  153


>☐ ref|XP_002081284.1| G GD25794 [Drosophila simulans]
 gb|EDX06869.1| G GD25794 [Drosophila simulans]
Length=112

 GENE ID: 6734255 Dsim\GD25794 | GD25794 gene product from transcript GD25794-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score =  106 bits (265),  Expect = 7e-22, Method: Composition-based stats.
```

```
 Identities = 23/114 (20%), Positives = 44/114 (38%), Gaps = 4/114 (3%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 +    L  +    A   + G+RV+ D L++R ++  +        L VNYP
Sbjct  1    MRLLLILSVVLAVILGCHAYSATWGRRVNNDFLLSRTREVRNPIKNNYWNLNVNYP---N  57

Query  83   YGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                +S V +  +  ++ G +  +  GG G R   + +  +         EI+G
Sbjct  58   GFYNISAVIVYDNFKNSSGASPSLYSGGPGYRFATVNLRGQVNRGIDSTVEIWG  111


>☐ ref|XP_001664091.1| UG hypothetical protein AaeL_AAEL013885 [Aedes aegypti]
 gb|EAT33840.1| G conserved hypothetical protein [Aedes aegypti]
Length=158

 GENE ID: 5578882 AaeL_AAEL013885 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score =  106 bits (264),  Expect = 9e-22, Method: Composition-based stats.
 Identities = 65/136 (47%), Positives = 90/136 (66%), Gaps = 2/136 (1%)

Query  1    TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA-RAIHNSLGQRVSGDRLVARN  59
            T  +P  + AKI +V+D+  F  A+P + +  L+  R+  + I  +LG R+SGDRLV  +
Sbjct  23   TEEEPLQIDAKIEVVQDIAAFKAAHPELDVVPLEVARSTRQQIVYTLGNRLSGDRLVGTS  82

Query  60   QDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILI  119
            QD  SW SLQ+V L + YP  G  GAVVSYV ++V+Q+SN G GYV+ GGIGQR I ++I
Sbjct  83   QDGTSWSSLQDVTLNLRYPQAGS-GAVVSYVQVVVNQSSNQGRGYVVSGGIGQRYIQLVI  141

Query  120  EARNTYTFSYIAEIYG  135
            E  +T  F Y A+IYG
Sbjct  142  EVYSTSYFQYNAQIYG  157


>☐ ref|XP_001987508.1| G GH21960 [Drosophila grimshawi]
 gb|EDW02375.1| G GH21960 [Drosophila grimshawi]
Length=122

 GENE ID: 6561150 Dgri\GH21960 | GH21960 gene product from transcript GH21960-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  106 bits (264),  Expect = 9e-22, Method: Composition-based stats.
 Identities = 15/115 (13%), Positives = 35/115 (30%), Gaps = 1/115 (0%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                 G+ +                G+    D  +A++  S ++           +
Sbjct  7    TVTLLGLFICCALVTSVYGGNDYLWGEIGPNDYHLAKDTVSKAFFVGLVQTKKYVFKQKN  66

Query  82   GYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
               A  ++ + +   + SN     +  GG G +   I+  +        + EI+G
Sbjct  67   DLNALTITAIKVTDKKKSNGATAVLTSGGPGSKGATIVFTSERGSGIKDLVEIWG  121


>☐ gb|ACN69173.1|  hypothetical conserved secreted protein [Stomoxys calcitrans]
Length=116

 Score =  106 bits (264),  Expect = 1e-21, Method: Composition-based stats.
 Identities = 19/114 (16%), Positives = 39/114 (34%), Gaps = 2/114 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                      +       A   S G+R S D L+ R     +    +N  + V++P  G
Sbjct  1    MKFLTSFCLIVSLATLGWAASASWGRRNSSDYLMLRENVVRTPIRNRNWSVNVDFPKPGQ  60

Query  83   YG-AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                 ++ + +    T+  G    +  GG       + + ++ +   +   EIY
Sbjct  61   INRRTITAIFVYDRFTNTSGAMPSLWSGGPYLTFANVNLRSQTSRGINSTVEIY  114


>☐ ref|XP_002107315.1| G GD15631 [Drosophila simulans]
 gb|EDX18297.1| G GD15631 [Drosophila simulans]
Length=163

 GENE ID: 6726355 Dsim\GD15631 | GD15631 gene product from transcript GD15631-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score =  104 bits (260),  Expect = 3e-21, Method: Composition-based stats.
```

```
 Identities = 48/134 (35%), Positives = 75/134 (55%), Gaps = 8/134 (5%)

Query  10   AKITLVEDVDEFVRANPGVTLTKLDNVRT-------ARAIHNSLGQRVSGDRLVARNQDS  62
              I +V+D+  F    PG+ + +++           ++ +  ++G R+  D LVA+  ++
Sbjct  30   TGIRVVDDIHVFAAQYPGLQIQQMEKEIVPGKARAGSQTVRYNMGARIPSDELVAQTANT  89

Query  63   ASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEAR  122
              +   Q+V L + YP NG  GA+VSYV +L  Q +  G  YV+ GGIGQ  I I++EAR
Sbjct  90   YEFPRAQDVSLQLTYPENGK-GAIVSYVELLCTQDTQEGTAYVVAGGIGQSLISIVLEAR  148

Query  123  NTYTFSYIAEIYGL  136
            NT  FSY A  YG+
Sbjct  149  NTKNFSYQALYYGI  162


>☐ ref|ZP_05005660.1|  conserved hypothetical protein [Streptomyces clavuligerus ATCC
27064]
 gb|EDY49959.1|  conserved hypothetical protein [Streptomyces clavuligerus ATCC
27064]
Length=125

 Score =  104 bits (259),  Expect = 3e-21, Method: Composition-based stats.
 Identities = 22/107 (20%), Positives = 43/107 (40%), Gaps = 3/107 (2%)

Query  30   LTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSY  89
                     A++ +  +G R+ GDRLV       S   ++ V  T  + A   +   ++
Sbjct  19   TVPNAQRAAAQSHNLIVGSRMPGDRLVLSQAVVKSSSWMRIVTETKTFQAP--HNERITL  76

Query  90   VSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            V  L  +T   G    +L GG G   + I  +++  +   ++  +Y
Sbjct  77   VEALDQRTDGTGAYASILNGGPGHNSVTIRFKSQRGHGIDFVVNLYA  123


>☐ ref|XP_002101745.1| G GE17797 [Drosophila yakuba]
 gb|EDX02853.1| G GE17797 [Drosophila yakuba]
Length=117

 GENE ID: 6525925 Dyak\GE17797 | GE17797 gene product from transcript GE17797-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score =  103 bits (258),  Expect = 4e-21, Method: Composition-based stats.
 Identities = 19/116 (16%), Positives = 40/116 (34%), Gaps = 3/116 (2%)

Query  23   RANPGVTLTKLDN--VRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPAN  80
               P + L  L    + T  A     G+    D  +A++  S ++           +  +
Sbjct  1    MKYPTIILFALAAFVLPTFAANDYLWGEVGENDYKLAKDTVSKAFFVGLVQTKKYVFKQS  60

Query  81   GGYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                A  ++ + +   + S+     ++ GG G +   I   +   Y      EI+G
Sbjct  61   DNLNALTITAIKVTDKKKSHGATAVLVSGGPGSKGATIKFTSERGYGIKDTVEIWG  116


>☐ ref|XP_002028226.1| G GL16235 [Drosophila persimilis]
 gb|EDW38775.1| G GL16235 [Drosophila persimilis]
Length=120

 GENE ID: 6603249 Dper\GL16235 | GL16235 gene product from transcript GL16235-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score =  103 bits (258),  Expect = 4e-21, Method: Composition-based stats.
 Identities = 26/109 (23%), Positives = 48/109 (44%), Gaps = 2/109 (1%)

Query  29   TLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG-GYGAVV  87
             L     V TA A+ ++ G   +  +L++     + +    + V   + YP NG G G ++
Sbjct  11   LLVVTSCVATAYAVSSTWGNISNSAQLLHAENVANASSPGRYVNHEIKYPKNGIGNGRII  70

Query  88   SYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            + +      T+N G    +  GG G   + I ++++  Y   +  EIYG
Sbjct  71   TGIRAFDQVTNNTGGHATIYSGGPGFNFVHIKLQSQYNYGLIFRVEIYG  119


>☐ ref|XP_002039289.1| G GM22808 [Drosophila sechellia]
 gb|EDW56031.1| G GM22808 [Drosophila sechellia]
Length=163

 GENE ID: 6614871 Dsec\GM22808 | GM22808 gene product from transcript GM22808-RA
```

```
[Drosophila sechellia] (10 or fewer PubMed links)

 Score =  103 bits (257),  Expect = 7e-21, Method: Composition-based stats.
 Identities = 49/134 (36%), Positives = 76/134 (56%), Gaps = 8/134 (5%)

Query  10   AKITLVEDVDEFVRANPGVTLTKLDNVRT-------ARAIHNSLGQRVSGDRLVARNQDS  62
              I +V+D+  F    PG+ + +++           ++ +  ++G R+  D LVA+  ++
Sbjct  30   TGIQVVDDIHVFAAQYPGLQIQQMEKEIVPGKARAGSQTVRYNMGARIPSDELVAQTANT  89

Query  63   ASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEAR  122
              +   Q+V L + YP NG  GA+VSYV +L  Q S+ G  YV+ GGIGQ  I I++EAR
Sbjct  90   YEFPRAQDVSLQLTYPENGK-GAIVSYVDLLCTQDSSEGTAYVVAGGIGQSLISIVLEAR  148

Query  123  NTYTFSYIAEIYGL  136
            NT  FSY A  YG+
Sbjct  149  NTKNFSYQALYYGI  162


>☐ ref|XP_001959559.1| G GF11989 [Drosophila ananassae]
 gb|EDV36381.1| G GF11989 [Drosophila ananassae]
Length=115

 GENE ID: 6494848 Dana\GF11989 | GF11989 gene product from transcript GF11989-RA
[Drosophila ananassae] (10 or fewer PubMed links)

 Score =  103 bits (256),  Expect = 8e-21, Method: Composition-based stats.
 Identities = 20/114 (17%), Positives = 44/114 (38%), Gaps = 1/114 (0%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
              +  +    L  +  A A   + G+  + D  +    +           + +NYPA+G
Sbjct  1    MRSVLILSVVLAVILGAHAYSATWGKAKNTDYQLLHQVEIRYPIKNNYWNVNINYPASGV  60

Query  83   YGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                +S + +  +  ++ G +  +  GG G R   I ++++     +    IYG
Sbjct  61   GYYNISAIYVYDNFKNSSGASPTLYSGGPGYRFATINLKSQVNRGINSTVVIYG  114


>☐ ref|XP_002076927.1| G GD24776 [Drosophila simulans]
 gb|EDX16521.1| G GD24776 [Drosophila simulans]
Length=117

 GENE ID: 6740043 Dsim\GD24776 | GD24776 gene product from transcript GD24776-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score =  103 bits (256),  Expect = 8e-21, Method: Composition-based stats.
 Identities = 17/117 (14%), Positives = 40/117 (34%), Gaps = 3/117 (2%)

Query  20   EFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPA  79
            ++           +  + T  A     G+  + D  +A++  S ++           +
Sbjct  2    KYTTIILFALAAFV--LPTFAANDYLWGEVGANDYKLAKDTVSKAFFVGLVQTKKYVFKQ  59

Query  80   NGGYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            +    A  ++ + I   + S+     ++ GG G +   I   +   Y      EI+G
Sbjct  60   SDNLNALTITAIKITDKKKSHGATAVLVSGGPGSKGATIKFTSERGYGIKDTVEIWG  116


>☐ ref|NP_001155164.1| UG venom protein R [Nasonia vitripennis]
Length=136

 GENE ID: 100120101 LOC100120101 | venom protein R [Nasonia vitripennis]

 Score =  103 bits (256),  Expect = 9e-21, Method: Composition-based stats.
 Identities = 17/102 (16%), Positives = 48/102 (47%), Gaps = 4/102 (3%)

Query  35   NVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILV  94
                 ++ +  +G R +GDRLV +         LQ +++  ++   G    +++ +  +
Sbjct  33   AAAPNKSHNLIIGNRQAGDRLVYQENIVKPSKWLQVIEVKKSFNITG---YLITQIRAMD  89

Query  95   DQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             +T+  G     + GG+G   + +  +++ ++  +++ +IY
Sbjct  90   QKTNGNGAIASRVDGGVGYSNVTLKFKSQRSHGINFVVQIYA  131


>☐ ref|NP_001033839.1| UG CG34026 [Drosophila melanogaster]
 gb|ABC67180.1| G CG34026 [Drosophila melanogaster]
```

```
 gb|ACD81771.1| G IP20615p [Drosophila melanogaster]
Length=117

 GENE ID: 3885569 CG34026 | CG34026 gene product from transcript CG34026-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score =  102 bits (255),  Expect = 1e-20, Method: Composition-based stats.
 Identities = 21/116 (18%), Positives = 42/116 (36%), Gaps = 3/116 (2%)

Query  23   RANPGVTLTKLDN--VRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPAN  80
               P + L  L    + T  A     G+  + D  +A++  S ++           +  +
Sbjct  1    MKYPTIILFALAAFILPTFAANDYLWGEVGADDYQLAKDTVSKAFFVGLVQTKKYVFKQS  60

Query  81   GGYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                A  ++ + I   + S+     ++ GG G +   I   +   Y    I EI+G
Sbjct  61   DNLNALTITAIKITDKKKSHGATAVLVSGGPGSKGATIKFTSERGYGIKDIVEIWG  116


>☐ ref|ZP_05006207.1|  conserved hypothetical protein [Streptomyces clavuligerus ATCC
27064]
 gb|EDY50506.1|  conserved hypothetical protein [Streptomyces clavuligerus ATCC
27064]
Length=125

 Score =  102 bits (255),  Expect = 1e-20, Method: Composition-based stats.
 Identities = 21/115 (18%), Positives = 44/115 (38%), Gaps = 3/115 (2%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                             A++ +  +G R+ GDRLV       +   ++ V  T  + A
Sbjct  11   ASTATVADTVPNAQRAAAQSHNLIVGSRMPGDRLVLSQAVVKNSSWMRIVTETKTFQAP-  69

Query  82   GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             +   ++ V  L  +T   G    +L GG G   + I  +++ ++   ++  +Y
Sbjct  70   -HNERITLVEALDQRTDGTGAYASILNGGPGHNSVTIRFKSQRSHGIDFVVNLYA  123


>☐ ref|XP_001868961.1| UG 14.5 kDa salivary protein [Culex quinquefasciatus]
 gb|EDS28048.1| G 14.5 kDa salivary protein [Culex quinquefasciatus]
Length=153

 GENE ID: 6052742 CpipJ_CPIJ018793 | 14.5 kDa salivary protein
[Culex quinquefasciatus]

 Score =  102 bits (254),  Expect = 1e-20, Method: Composition-based stats.
 Identities = 63/136 (46%), Positives = 93/136 (68%), Gaps = 2/136 (1%)

Query  1    TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA-RAIHNSLGQRVSGDRLVARN  59
            +   P  + A + +V+++  F  ANP + L  LD  R+  + I  S+G R+SGDRLVA +
Sbjct  18   SDEQPAAIEAHVEIVQNIGAFKAANPDLELVPLDITRSTRQQIVYSVGNRISGDRLVAAD  77

Query  60   QDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILI  119
             +  SW SLQ+V LT+NYP +G  GAVV+YV ++V+Q+SN G GY+  GGIGQR+I ++I
Sbjct  78   SNGQSWSSLQDVALTLNYPRSG-TGAVVTYVQVVVNQSSNQGRGYITAGGIGQRRITLVI  136

Query  120  EARNTYTFSYIAEIYG  135
            EA +T  F+Y A+I+G
Sbjct  137  EAYSTNYFNYNAQIFG  152


>☐ ref|XP_001868962.1| UG conserved hypothetical protein [Culex quinquefasciatus]
 gb|EDS28049.1| G conserved hypothetical protein [Culex quinquefasciatus]
Length=153

 GENE ID: 6052738 CpipJ_CPIJ018794 | hypothetical protein
[Culex quinquefasciatus]

 Score =  102 bits (254),  Expect = 1e-20, Method: Composition-based stats.
 Identities = 65/136 (47%), Positives = 94/136 (69%), Gaps = 2/136 (1%)

Query  1    TAPDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTA-RAIHNSLGQRVSGDRLVARN  59
            +   P  + A + +V+++  F  ANP V L  LD  R+  + I  S+G R+SGDRLVA +
Sbjct  18   SDEQPAAIEAYVEIVQNIGAFKAANPDVELVPLDITRSTRQQIVYSVGNRISGDRLVAAD  77

Query  60   QDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILI  119
             ++ SW SLQ+V LT+NYP +G  GAVV+YV ++V+Q+SN G GY+  GGIGQR+I ++I
Sbjct  78   SNAQSWSSLQDVALTLNYPRSG-TGAVVTYVQVVVNQSSNQGRGYITAGGIGQRRITLVI  136
```

```
Query  120  EARNTYTFSYIAEIYG  135
            EA +T  F+Y A+IYG
Sbjct  137  EAYSTNYFNYNAQIYG  152


>☐ ref|XP_002063499.1| G GK21944 [Drosophila willistoni]
 gb|EDW74485.1| G GK21944 [Drosophila willistoni]
Length=112

 GENE ID: 6640573 Dwil\GK21944 | GK21944 gene product from transcript GK21944-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score =  102 bits (254),  Expect = 1e-20, Method: Composition-based stats.
 Identities = 25/109 (22%), Positives = 46/109 (42%), Gaps = 2/109 (1%)

Query  29   TLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG-GYGAVV  87
             L     V TA A+ ++ G   +  +L+       +    +     + YP NG G G ++
Sbjct  3    VLLVASCVATAYAVSSTWGNISNSAQLLHAENVVNASSPGKYAHHEIKYPKNGIGNGRII  62

Query  88   SYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            + +      T+N G    +  GG G   + I ++++   Y   +  EIYG
Sbjct  63   TGIRAFDQVTNNTGGHATIYSGGPGFNFVHIKLQSQYNYGLIFRVEIYG  111


>☐ ref|XP_001865641.1| G conserved hypothetical protein [Culex quinquefasciatus]
 gb|EDS42031.1| G conserved hypothetical protein [Culex quinquefasciatus]
Length=102

 GENE ID: 6048772 CpipJ_CPIJ015172 | hypothetical protein
[Culex quinquefasciatus]

 Score =  101 bits (252),  Expect = 2e-20, Method: Composition-based stats.
 Identities = 19/99 (19%), Positives = 40/99 (40%), Gaps = 3/99 (3%)

Query  40   RAIHNSLGQRVSGDRLVARNQDSASW--GSLQNVQLTVNYPANGGYGAVVSYVSILVDQT  97
            +  + S G R   D L++R            +   + + YP  G  G  ++ ++++
Sbjct  3    KQHNFSWGVRGPYDVLLSRTTILEPPARSPYRTRTVNLYYPEKGQPGRTIAAINVIDGFG  62

Query  98   SNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
               G    +  GG+GQ    + + ++     ++  EIYG
Sbjct  63   DGTGGTARLWSGGLGQNSTTVRLRSQRGRGLNFTVEIYG  101


>☐ ref|XP_001975824.1| G GG22532 [Drosophila erecta]
 gb|EDV56224.1| G GG22532 [Drosophila erecta]
Length=112

 GENE ID: 6549590 Dere\GG22532 | GG22532 gene product from transcript GG22532-RA
[Drosophila erecta] (10 or fewer PubMed links)

 Score =  101 bits (252),  Expect = 2e-20, Method: Composition-based stats.
 Identities = 24/98 (24%), Positives = 42/98 (42%), Gaps = 4/98 (4%)

Query  39   ARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTS  98
              A   + G+RV+ D L++R  +  +        L VNYPA       +S V +  +  +
Sbjct  17   CHAYSATWGRRVNNDFLLSRTNEVRNPIKNNYWNLNVNYPA---GFYNISAVIVYDNFKN  73

Query  99   NLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            N G +  +  GG G R   + +  +     +   EI+G
Sbjct  74   NSGASPSLYSGGPGYRFATVNLRGQVNRGINSTVEIWG  111


>☐ ref|XP_001987095.1| G GH21729 [Drosophila grimshawi]
 gb|EDW01962.1| G GH21729 [Drosophila grimshawi]
Length=120

 GENE ID: 6560516 Dgri\GH21729 | GH21729 gene product from transcript GH21729-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  101 bits (252),  Expect = 2e-20, Method: Composition-based stats.
 Identities = 23/116 (19%), Positives = 48/116 (41%), Gaps = 2/116 (1%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                    +  +  V  A ++ ++ G   +  +L+     +      + V+  + YP NG
Sbjct  4    TVVAIVPLVFFIWFVAPALSVSSTWGNISNSAQLLHAENVAQGSSPEKYVEDEIKYPTNG  63
```

```
Query  82   -GYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             G G +++ +      T+N G    +  GG G   + I ++++  Y   +   EIYG
Sbjct  64   IGNGRIITGIRAFDQVTNNTGGHATIYSGGPGFNFVHIKLQSQYNYGLIFRVEIYG  119


>☐ ref|XP_001865369.1| UG conserved hypothetical protein [Culex quinquefasciatus]
 gb|EDS41618.1| G conserved hypothetical protein [Culex quinquefasciatus]
Length=118

 GENE ID: 6048450 CpipJ_CPIJ014861 | hypothetical protein
[Culex quinquefasciatus]

 Score =  101 bits (252),  Expect = 3e-20, Method: Composition-based stats.
 Identities = 29/119 (24%), Positives = 51/119 (42%), Gaps = 6/119 (5%)

Query  20   EFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY-P  78
            +F+ A+  + L     +  A +     G+RV GDRL+ R     +   +      V Y P
Sbjct  2    KFLLASATL-LAIFAVLILADSN--QWGRRVPGDRLLNRTTVVNNSLPIPIKTAVVQYLP  58

Query  79   ANGGYGAVVSYVSILVDQTS--NLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     V+Y+    +  +    G   ++ GGI Q    + + +R +  F++  EIYG
Sbjct  59   PPGVRPPNVTYIVARDNVKNPGQGGLAALISGGINQTHATVKLSSRPSLGFNFTVEIYG  117


>☐ ref|XP_002041778.1| G GM11373 [Drosophila sechellia]
 gb|EDW45626.1| G GM11373 [Drosophila sechellia]
Length=117

 GENE ID: 6617461 Dsec\GM11373 | GM11373 gene product from transcript GM11373-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score =  101 bits (251),  Expect = 3e-20, Method: Composition-based stats.
 Identities = 17/117 (14%), Positives = 39/117 (33%), Gaps = 3/117 (2%)

Query  20   EFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPA  79
            ++           +  + T  A     G+    D  +A++  S ++           +
Sbjct  2    KYTTIILFALAAFV--LPTFAANDYLWGEVGVNDYKLAKDTVSKAFFVGLVQTKKYVFKQ  59

Query  80   NGGYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            +    A  ++ + I   + S+     ++ GG G +   I   +   Y      EI+G
Sbjct  60   SDNLNALTITAIKITDKRKSHGATAVLVSGGPGSKGATIKFTSERGYGIKDTVEIWG  116


>☐ ref|XP_001664100.1| UG hypothetical protein AaeL_AAEL013895 [Aedes aegypti]
 gb|EAT33832.1| G conserved hypothetical protein [Aedes aegypti]
Length=155

 GENE ID: 5578892 AaeL_AAEL013895 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score =  101 bits (251),  Expect = 4e-20, Method: Composition-based stats.
 Identities = 58/135 (42%), Positives = 83/135 (61%), Gaps = 5/135 (3%)

Query  2    APDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQD  61
            A   + +SA I   ++V  F   N         + +  + +  +LGQRV GDRLV+   D
Sbjct  26   ADVSDQISATIEAEKNVGVFTVGNQL----PGPSGKVGQQLVYTLGQRVVGDRLVSTKSD  81

Query  62   SASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEA  121
            S  W +LQ+V L++ YP +G  GAVV+YV ++V Q+SNLG GYV+ GGIGQR I + IEA
Sbjct  82   SKQWPTLQDVTLSLAYPQSG-VGAVVTYVQVIVKQSSNLGKGYVVSGGIGQRFIHLNIEA  140

Query  122  RNTYTFSYIAEIYGL  136
             +T  F+Y  +I+GL
Sbjct  141  YSTNYFTYTVQIFGL  155


>☐ gb|ABH10141.1|  HMG176 [Helicoverpa armigera]
 gb|ABK29472.1|  HMG176 [Helicoverpa armigera]
Length=176

 Score =  100 bits (249),  Expect = 5e-20, Method: Composition-based stats.
 Identities = 27/127 (21%), Positives = 49/127 (38%), Gaps = 7/127 (5%)

Query  9    SAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSL  68
             A    + ++  F+     V L   +    A     SLG   + DRL++ +  S +  +
```

```
Sbjct  44   DADEVPMSEMAVFITK---VLLADEEMPFVAPRSGMSLGNIGASDRLLSASTHSRNPIAN  100

Query  69   QNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFS  128
            Q   + V Y   G    ++  V      +       V+ G +G+  I I +++     F
Sbjct  101  QVQTVNVRYT--GSSSIIILAVRAY--GSGQGATARVVEGYLGRNSITIQLQSARGRGFH  156

Query  129  YIAEIYG  135
            Y  EI+G
Sbjct  157  YRIEIWG  163


>☐ gb|ABV44709.1|  18 kDa midgut protein [Phlebotomus papatasi]
Length=161

 Score =  100 bits (249),  Expect = 6e-20, Method: Composition-based stats.
 Identities = 50/136 (36%), Positives = 70/136 (51%), Gaps = 2/136 (1%)

Query  2    APDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARA-IHNSLGQRVSGDRLVARNQ  60
              D    +     V ++ +FV  NP   L  +      R  I+  LG R  GDRLV
Sbjct  27   DDDAPDSNVNFLSVNNLADFVAENPEAELEMMHMYSDYRGQINYVLGARQRGDRLVGTAS  86

Query  61   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  120
            +   W   QNV + + YP  G  GA+V+YVS+ + Q++NLG  YV  GGI QR I I++E
Sbjct  87   NGQRWPRPQNVDIEIKYPTKG-RGALVTYVSLNITQSTNLGRAYVTKGGIHQRNITIVVE  145

Query  121  ARNTYTFSYIAEIYGL  136
            A +T    YI  I+G+
Sbjct  146  AYSTTFVHYICNIFGM  161


>☐ ref|XP_001604583.1| UG PREDICTED: hypothetical protein [Nasonia vitripennis]
Length=129

 GENE ID: 100120995 LOC100120995 | hypothetical protein LOC100120995
[Nasonia vitripennis]

 Score =  100 bits (248),  Expect = 6e-20, Method: Composition-based stats.
 Identities = 22/104 (21%), Positives = 51/104 (49%), Gaps = 3/104 (2%)

Query  33   LDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSI  92
            +++   +++ H  +G RV GDRL+ R      +   + V+ T  +P  G   + ++ +
Sbjct  26   VNSTNNSKSHHLIIGSRVWGDRLLQRVHVQKDYSWFRVVKETRTFP--GDGISKITEIQA  83

Query  93   LVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            L  ++   G    ++ GG  Q  + +  +++ + +  ++ EIYG
Sbjct  84   LDQRSDRSGADVALINGGPDQTFVTLQFKSQRSQSIDFVVEIYG  127


>☐ ref|XP_002016257.1| G GL10590 [Drosophila persimilis]
 gb|EDW32147.1| G GL10590 [Drosophila persimilis]
Length=116

 GENE ID: 6591103 Dper\GL10590 | GL10590 gene product from transcript GL10590-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score =  100 bits (248),  Expect = 7e-20, Method: Composition-based stats.
 Identities = 19/115 (16%), Positives = 40/115 (34%), Gaps = 2/115 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                      +  +        S G+R + D L++R  +  S        + V++P
Sbjct  1    MKLLLCLTILVAAIFVVHGSSVSWGRRKNNDFLLSRQVEVRSAIKNNFWNIDVDFPRADT  60

Query  83   YG-AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 +S + +  +  ++ G    +  GG G R   + + +      +   EIYG
Sbjct  61   ENYYKISAIFVYDNFKNSTGAVPGLWAGGPGYRFAQVNLRSFVNVGLNSTIEIYG  115


>☐ ref|XP_001965768.1| G GF20010 [Drosophila ananassae]
 gb|EDV41416.1| G GF20010 [Drosophila ananassae]
Length=117

 GENE ID: 6502742 Dana\GF20010 | GF20010 gene product from transcript GF20010-RA
[Drosophila ananassae] (10 or fewer PubMed links)

 Score =  100 bits (248),  Expect = 7e-20, Method: Composition-based stats.
 Identities = 18/115 (15%), Positives = 42/115 (36%), Gaps = 1/115 (0%)
```

```
Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                  V       + +    +   G+    D ++A+++ S ++           +
Sbjct  2    KSRLIFVLAITAYIIPSFAFNNYLWGEISGNDYILAKDKVSKAFFVGLVQTKKYVFKQKD  61

Query  82   GYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
               A  ++ + I   +  N     +L GG G +   IL +++  Y  + + EI+G
Sbjct  62   NLNALTITAIKITDKKEKNGATAELLSGGPGSKGATILFQSKRGYGINDVVEIWG  116


>☐ ref|XP_001649836.1| UG hypothetical protein AaeL_AAEL004809 [Aedes aegypti]
 gb|ABF18045.1|  putative secreted salivary protein [Aedes aegypti]
 gb|EAT43758.1| G conserved hypothetical protein [Aedes aegypti]
Length=155

 GENE ID: 5565507 AaeL_AAEL004809 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score = 99.6 bits (247),  Expect = 8e-20, Method: Composition-based stats.
 Identities = 58/135 (42%), Positives = 83/135 (61%), Gaps = 5/135 (3%)

Query  2    APDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQD  61
            A   + +SA I   ++V  F   N         + +  + +  +LGQRV GDRLV+   D
Sbjct  26   ADVSDQISATIEAEKNVGVFTVGNQL----PGPSGKVGQQLVYTLGQRVVGDRLVSTKSD  81

Query  62   SASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEA  121
            S  W +LQ+V L++ YP +G  GAVV+YV ++V Q+SNLG GYV+ GGIGQR I + IEA
Sbjct  82   SKQWPTLQDVTLSLAYPQSG-VGAVVTYVQVIVKQSSNLGKGYVVSGGIGQRFIHLNIEA  140

Query  122  RNTYTFSYIAEIYGL  136
             +T  F+Y  +I+GL
Sbjct  141  YSTNYFTYTVQIFGL  155


>☐ ref|XP_001604165.1| UG PREDICTED: hypothetical protein [Nasonia vitripennis]
Length=119

 GENE ID: 100120529 LOC100120529 | hypothetical protein LOC100120529
[Nasonia vitripennis]

 Score = 99.6 bits (247),  Expect = 9e-20, Method: Composition-based stats.
 Identities = 21/96 (21%), Positives = 42/96 (43%), Gaps = 3/96 (3%)

Query  41   AIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNL  100
                 +G R  GD L+ R      +  L++V+LT  +   G   + ++ V +L
Sbjct  25   PNSLIVGDREPGDHLMQREFVMKGYKWLRSVKLTKTFT--GSVHSKITQVQLLDQNEKGN  82

Query  101  G-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            G    +L GG G   + +  ++   +  +Y+ ++YG
Sbjct  83   GATAAILSGGPGDNFVTVQFKSVRGHGVNYVVKLYG  118


>☐ ref|XP_001599861.1| UG PREDICTED: similar to CG34026-PA [Nasonia vitripennis]
Length=116

 GENE ID: 100117346 LOC100117346 | similar to CG34026-PA [Nasonia vitripennis]

 Score = 99.2 bits (246),  Expect = 1e-19, Method: Composition-based stats.
 Identities = 20/115 (17%), Positives = 34/115 (29%), Gaps = 2/115 (1%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                  +         T       +G  V+GDRL+                  +     G
Sbjct  2    TSQYVFILAVFATLAVTFANNDLIIGDHVAGDRLLQLEHIEKDAAWWGEKGS-ITRTFEG  60

Query  82   GYGAVVSYVSILVDQTSN-LGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
               A ++ V  L    +       ++ GG+G   + I   +       +I EIYG
Sbjct  61   DTFAKITMVRALDKHDNGHGATAEIIAGGVGHSYVTIKFVSERLRGIDFIVEIYG  115


>☐ ref|XP_002028229.1| G GL16227 [Drosophila persimilis]
 gb|EDW38778.1| G GL16227 [Drosophila persimilis]
Length=123

 GENE ID: 6603243 Dper\GL16227 | GL16227 gene product from transcript GL16227-RA
[Drosophila persimilis] (10 or fewer PubMed links)
```

```
 Score = 99.2 bits (246),  Expect = 1e-19, Method: Composition-based stats.
 Identities = 17/117 (14%), Positives = 40/117 (34%), Gaps = 2/117 (1%)

Query  20   EFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPA  79
             FV     +              +   G+  + D  +A++    ++           +
Sbjct  7    AFVVTVAFILFVAFH-EPVCATNNYLWGEIGANDYKLAKDTVKKAFFVGLVQTKKYVFKQ  65

Query  80   NGGYGA-VVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            +    A  ++ + I   + +N     ++ GG G +   IL  +   Y  + + EI+G
Sbjct  66   SDNLDALTITAIQITDKKKNNGATAELISGGPGSKGATILFNSVRGYGINDVVEIWG  122


>☐ ref|XP_002099025.1| G Acp225 [Drosophila yakuba]
 gb|EDW98737.1| G Acp225 [Drosophila yakuba]
Length=120

 GENE ID: 6538503 Dyak\Acp225 | GE10689 gene product from transcript GE10689-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score = 99.2 bits (246),  Expect = 1e-19, Method: Composition-based stats.
 Identities = 19/112 (16%), Positives = 38/112 (33%), Gaps = 3/112 (2%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+          + G    GD+L+ + +   +   L+ V   V++
Sbjct  8    IFVAVAIVAVLNVAFAGHDGSTTFGVLKPGDKLIHKEEIVEAKKFLRVVTRKVHFKPQ--  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   +  ++  EIY
Sbjct  66   -PYKLSAIVITDHSESKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  116


>☐ gb|ABW82136.2|  IP21714p [Drosophila melanogaster]
Length=131

 Score = 98.8 bits (245),  Expect = 1e-19, Method: Composition-based stats.
 Identities = 19/117 (16%), Positives = 34/117 (29%), Gaps = 1/117 (0%)

Query  19   DEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
              F+            N  +      + G + + D L+A    + S   L     T
Sbjct  14   GVFLLLVGTHVCFIDANFGSGEGNDYTYGTQATTDTLIASETITKSKSLLGITTKTYTLT  73

Query  79   ANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     ++Y+ I   +        +  GG+G   + I   +           IYG
Sbjct  74   QAG-TAKTITYIKITDLKKMRGATAEITSGGVGSTTVTIKFTSARGAGIKSQVVIYG  129


>☐ gb|ACD99549.1|  IP21514p [Drosophila melanogaster]
 gb|ACD99554.1|  IP21614p [Drosophila melanogaster]
Length=129

 Score = 98.8 bits (245),  Expect = 2e-19, Method: Composition-based stats.
 Identities = 19/117 (16%), Positives = 34/117 (29%), Gaps = 1/117 (0%)

Query  19   DEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
              F+            N  +      + G + + D L+A    + S   L     T
Sbjct  12   GVFLLLVGTHVCFIDANFGSGEGNDYTYGTQATTDTLIASETITKSKSLLGITTKTYTLT  71

Query  79   ANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     ++Y+ I   +        +  GG+G   + I   +           IYG
Sbjct  72   QAG-TAKTITYIKITDLKKMRGATAEITSGGVGSTTVTIKFTSARGAGIKSQVVIYG  127


>☐ ref|XP_001865640.1| UG hypothetical protein CpipJ_CPIJ015171 [Culex quinquefasciatus]
 gb|EDS42030.1| G hypothetical protein CpipJ_CPIJ015171 [Culex quinquefasciatus]
Length=128

 GENE ID: 6048771 CpipJ_CPIJ015171 | hypothetical protein
[Culex quinquefasciatus]

 Score = 98.5 bits (244),  Expect = 2e-19, Method: Composition-based stats.
 Identities = 28/120 (23%), Positives = 45/120 (37%), Gaps = 5/120 (4%)

Query  20   EFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSA----SWGSLQNVQLTV  75
            +F+     +       +  +R      G R   D L+           SW S + V  T+
Sbjct  2    KFLVFAIVLISFSALALTQSRQNTIRWGNRKQSDVLIGHYVAIKKAKSSWFSSKAVNQTL  61
```

```
Query  76   NYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             YPA G  G  +S +           N  +L GGIG+    IL+ ++     +   E+YG
Sbjct  62   QYPAMGQLGRNISAIYFYE-APKGRANPKLLSGGIGRNYTTILLTSQKGQGLNVTVELYG  120


>☐ gb|ACO72873.1|  MIP04578p [Drosophila melanogaster]
Length=118

 Score = 98.5 bits (244),  Expect = 2e-19, Method: Composition-based stats.
 Identities = 24/116 (20%), Positives = 45/116 (38%), Gaps = 4/116 (3%)

Query  21   FVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPAN  80
            F      +    L  +    A   + G+RV+ D L++R ++  +        + VNYP
Sbjct  5    FNMRLLLILSVVLAVILGCHAYSATWGRRVNNDFLLSRTREVRNPIKNNYWNVNVNYP--  62

Query  81   GGYGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                  +S V +  +  +N G +  +  GG G R   + +  +         EI+G
Sbjct  63   -NGFYNISAVIVYDNFKNNSGASPSLYSGGPGYRFATVNLRGQVNRGIDSTVEIWG  117


>☐ ref|XP_001599834.1| UG PREDICTED: similar to putative salivary secreted peptide [Nasonia
vitripennis]
Length=125

 GENE ID: 100118223 LOC100118223 | similar to putative salivary secreted peptide
[Nasonia vitripennis]

 Score = 98.5 bits (244),  Expect = 2e-19, Method: Composition-based stats.
 Identities = 19/101 (18%), Positives = 40/101 (39%), Gaps = 2/101 (1%)

Query  36   VRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVD  95
            +  +       G R+ GDRL+ R +       +    +T+     G   + +S +  L
Sbjct  25   LAASLPNDVIAGSRIPGDRLIQREEVKKDANWVGMK-VTITKTFEGDKYSKISQLRALDQ  83

Query  96   QTSN-LGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             ++       ++ GG+G   + I  E+       ++ E+YG
Sbjct  84   HSNGHGATAKIIAGGVGFSYVTIKFESERFRGIDFVVELYG  124


>☐ gb|ABD97958.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 98.1 bits (243),  Expect = 2e-19, Method: Composition-based stats.
 Identities = 18/112 (16%), Positives = 37/112 (33%), Gaps = 2/112 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+          + G    G +L+ + +   +   L+ V   V++
Sbjct  8    IFVAIAIVAVLNVAFAGHDGSTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   + ++   EIY
Sbjct  66   QPYKLSAIVITDHSESKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ gb|ABD97961.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 98.1 bits (243),  Expect = 3e-19, Method: Composition-based stats.
 Identities = 18/112 (16%), Positives = 37/112 (33%), Gaps = 2/112 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+          + G    G +L+ + +   +   L+ V   V++
Sbjct  8    IFVAIAIVAVLNVAFAGHDGSTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   + ++   EIY
Sbjct  66   QPYKLSAIVITDHSESKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ gb|ABD97965.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 97.7 bits (242),  Expect = 3e-19, Method: Composition-based stats.
 Identities = 19/112 (16%), Positives = 38/112 (33%), Gaps = 2/112 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
```

```
                   +  L+ V        + G    G +L+ + +   +   L+ V   V++
Sbjct  8    IFVAIAIVAVLNVVFAGHDGTTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   +  ++  EIY
Sbjct  66   QPYKLSAIVITDHSESKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ gb|ABD97960.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 97.7 bits (242),  Expect = 3e-19, Method: Composition-based stats.
 Identities = 19/112 (16%), Positives = 37/112 (33%), Gaps = 2/112 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+          + G    G +L+ + +   +   L+ V   V+Y
Sbjct  8    IFVAIAIVAVLNVAFAGHDGTTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHY--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   +  ++  EIY
Sbjct  66   QPYKLSAIVITDHSESKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ ref|NP_726308.1| UG CG30413 [Drosophila melanogaster]
 gb|AAM68241.1| G CG30413 [Drosophila melanogaster]
Length=122

 GENE ID: 246601 CG30413 | CG30413 gene product from transcript CG30413-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score = 97.7 bits (242),  Expect = 4e-19, Method: Composition-based stats.
 Identities = 19/117 (16%), Positives = 34/117 (29%), Gaps = 1/117 (0%)

Query  19   DEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
              F+            N  +      + G + + D L+A    + S   L     T
Sbjct  5    GVFLLLVGTHVCFIDANFGSGEGNDYTYGTQATTDTLIASETITKSKSLLGITTKTYTLT  64

Query  79   ANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     ++Y+ I   +        +  GG+G   + I   +           IYG
Sbjct  65   QAG-TAKTITYIKITDLKKMRGATAEITSGGVGSTTVTIKFTSARGAGIKSQVVIYG  120


>☐ ref|XP_001603272.1| G PREDICTED: similar to ENSANGP00000031402 [Nasonia vitripennis]
Length=118

 GENE ID: 100120502 LOC100120502 | similar to ENSANGP00000031402
[Nasonia vitripennis]

 Score = 97.7 bits (242),  Expect = 4e-19, Method: Composition-based stats.
 Identities = 28/115 (24%), Positives = 49/115 (42%), Gaps = 5/115 (4%)

Query  23   RANPGVTLTKLDNVRTARAI-HNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
                 +  T +    +A++  H  +G R  GD L  R         LQ V+L   +   G
Sbjct  6    LLLFVLLATSVAVALSAKSAKHVIIGVRQPGDHLARREFVKKDSSWLQVVKLQQTF---G  62

Query  82   GYGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 +++V +L       G    VL GG+G I +  +++ +++ +YI EIY
Sbjct  63   VDYGKITHVKLLDQNLKGKGATANVLAGGLGYSYITVHFKSKRSHSINYIVEIYA  117


>☐ gb|ABD97959.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 97.7 bits (242),  Expect = 4e-19, Method: Composition-based stats.
 Identities = 18/112 (16%), Positives = 37/112 (33%), Gaps = 2/112 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+          + G    G +L+ + +   +   L+ V   V++
Sbjct  8    IFVAVAIVAVLNVAFAGHDGTTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   +  ++  EIY
Sbjct  66   QPYKLSAIVITDHSESKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ ref|XP_001965767.1| G GF20011 [Drosophila ananassae]
```

```
 gb|EDV41415.1|  GF20011 [Drosophila ananassae]
Length=122

 GENE ID: 6502743 Dana\GF20011 | GF20011 gene product from transcript GF20011-RA
[Drosophila ananassae] (10 or fewer PubMed links)

 Score = 97.3 bits (241),  Expect = 5e-19, Method: Composition-based stats.
 Identities = 20/97 (20%), Positives = 35/97 (36%), Gaps = 1/97 (1%)

Query  39   ARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTS  98
                  + G + + D L+AR   + S   LQ V  T      G     ++Y+ I   +
Sbjct  25   GEGNDYTYGSKATTDTLIARETITKSKMMLQTVTKTYTLTQAG-TAKTITYIQITDLKRM  83

Query  99   NLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
               +  +  GG+G   + I   +          EI+G
Sbjct  84   RGASAEITSGGVGSTTVTIKFTSARGAGIKSQVEIWG  120


>☐ gb|ABQ18257.1|  unknown [Lygus lineolaris]
Length=133

 Score = 97.3 bits (241),  Expect = 5e-19, Method: Composition-based stats.
 Identities = 20/101 (19%), Positives = 44/101 (43%), Gaps = 4/101 (3%)

Query  36   VRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVD  95
                ++ +  +G R   D+L+  +        L+ V   V   A G    +++Y+ +L
Sbjct  33   CGRNKSHNAFIGSRHWDDKLLFMDHPKKESKWLRTVSADV---ATGNLRGIINYIEVLDQ  89

Query  96   QTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            +T+  G   Y+  GG G   +   I+++ ++   +I   +G
Sbjct  90   KTNGNGGCAYLTSGGAGTNNVAFHIKSQRSHGLDFIIRAWG  130


>☐ ref|XP_002050018.1|  GJ20423 [Drosophila virilis]
 gb|EDW61211.1|  GJ20423 [Drosophila virilis]
Length=122

 GENE ID: 6625995 Dvir\GJ20423 | GJ20423 gene product from transcript GJ20423-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score = 97.3 bits (241),  Expect = 5e-19, Method: Composition-based stats.
 Identities = 19/122 (15%), Positives = 38/122 (31%), Gaps = 2/122 (1%)

Query  15   VEDVDEFVRANPGVTLTKLDNVRT-ARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQL  73
            +  +  F+                       + G +++   L+AR   S S   LQ
Sbjct  1    MNSILVFILLMCTCVCVTHGLGGNNGEGYDYTYGAKLTTSTLIARETISKSKTLLQTTSK  60

Query  74   TVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEI  133
            T      G     + Y++I   +        ++ GG+G   + +   +          EI
Sbjct  61   TYTLTQAG-TAKNIDYINITNLKKMRGATAEIMSGGVGSTTVTVKFTSARGSGIKSQIEI  119

Query  134  YG  135
            +G
Sbjct  120  WG  121


>☐ gb|ABD97966.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 97.3 bits (241),  Expect = 5e-19, Method: Composition-based stats.
 Identities = 18/112 (16%), Positives = 37/112 (33%), Gaps = 2/112 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+          + G    G +L+ + +   +   L+ V   V++
Sbjct  8    IFVAIAIVAVLNVAFAGHDGTTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   +  ++  EIY
Sbjct  66   QPYKLSAIVITDHSESKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ gb|ABD97964.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 96.5 bits (239),  Expect = 7e-19, Method: Composition-based stats.
 Identities = 18/112 (16%), Positives = 37/112 (33%), Gaps = 2/112 (1%)
```

```
Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+ V        + G    G +L+ + +   +   L+ V   V++
Sbjct  8    IFVAIAIVAVLNVVFAGHDGTTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I        G   +L GG   +   I   +   +  ++  EIY
Sbjct  66   QPYKLSAIVITDHSEXKGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ ref|XP_002033647.1| G GM20317 [Drosophila sechellia]
 gb|EDW47660.1| G GM20317 [Drosophila sechellia]
Length=112

 GENE ID: 6608933 Dsec\GM20317 | GM20317 gene product from transcript GM20317-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score = 96.5 bits (239),  Expect = 8e-19, Method: Composition-based stats.
 Identities = 24/114 (21%), Positives = 45/114 (39%), Gaps = 4/114 (3%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 +    L  +    A   + G+R S D L++R  ++ +        + VNYPA
Sbjct  1    MRLLLILSVVLAVILGCHAYGATWGRRNSNDYLLSRTNEARNPIKNNYWNVNVNYPA---  57

Query  83   YGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                +S V +  +  +N G +  +  GG G R   + +  +     +   EI+G
Sbjct  58   GFYNISAVIVYDNFKNNSGASPSLYSGGPGYRFATVNLRGQVNRGINSTVEIWG  111


>☐ ref|NP_610832.1| UG CG13323 [Drosophila melanogaster]
 gb|AAF58432.1| G CG13323 [Drosophila melanogaster]
 gb|AAY55445.1| G IP04046p [Drosophila melanogaster]
 gb|ACL88185.1|  CG13323-PA [synthetic construct]
Length=112

 GENE ID: 36433 CG13323 | CG13323 gene product from transcript CG13323-RA
[Drosophila melanogaster] (Over 10 PubMed links)

 Score = 96.5 bits (239),  Expect = 9e-19, Method: Composition-based stats.
 Identities = 23/114 (20%), Positives = 44/114 (38%), Gaps = 4/114 (3%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 +    L  +    A   + G+R + D L++R  +  +        + VNYPA
Sbjct  1    MRLLLILSVVLAVILGCHAYGATWGRRNNNDYLLSRTTEVRNPIKNNYWNVNVNYPA---  57

Query  83   YGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                +S V +  +  +N G +  +  GG G R   + +  +     +   EI+G
Sbjct  58   GFYNISAVIVYDNFKNNSGASPSLYSGGPGYRFATVNLRGQVNRGINSTVEIWG  111


>☐ ref|XP_002081283.1| G GD25795 [Drosophila simulans]
 gb|EDX06868.1| G GD25795 [Drosophila simulans]
Length=112

 GENE ID: 6734254 Dsim\GD25795 | GD25795 gene product from transcript GD25795-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score = 96.2 bits (238),  Expect = 1e-18, Method: Composition-based stats.
 Identities = 23/114 (20%), Positives = 45/114 (39%), Gaps = 4/114 (3%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 +    L  +    A   + G+R + D L++R  ++ +        + VNYPA
Sbjct  1    MRLLLILSVVLAVILGCHAYGATWGRRNNNDYLLSRTNEARNPIKNNYWNVNVNYPA---  57

Query  83   YGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                +S V +  +  +N G +  +  GG G R   + +  +     +   EI+G
Sbjct  58   GFYNISAVIVYDNFKNNSGASPSLYSGGPGYRFATVNLRGQVNRGINSTVEIWG  111


>☐ ref|NP_610833.1| UG CG13324 [Drosophila melanogaster]
 gb|AAF58431.1| G CG13324 [Drosophila melanogaster]
Length=112

 GENE ID: 36434 CG13324 | CG13324 gene product from transcript CG13324-RA
```

```
[Drosophila melanogaster] (Over 10 PubMed links)

 Score = 96.2 bits (238),  Expect = 1e-18, Method: Composition-based stats.
 Identities = 23/114 (20%), Positives = 44/114 (38%), Gaps = 4/114 (3%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 +    L  +    A   + G+RV+ D L++R ++  +        + VNYP
Sbjct  1    MRLLLILSVVLAVILGCHAYSATWGRRVNNDFLLSRTREVRNPIKNNYWNVNVNYP---N  57

Query  83   YGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                +S V +  +  +N G +  +  GG G R   + +  +         EI+G
Sbjct  58   GFYNISAVIVYDNFKNNSGASPSLYSGGPGYRFATVNLRGQVNRGIDSTVEIWG  111


>☐ ref|XP_001602278.1| G PREDICTED: similar to conserved hypothetical protein [Nasonia
vitripennis]
Length=126

 GENE ID: 100118260 LOC100118260 | similar to conserved hypothetical protein
[Nasonia vitripennis]

 Score = 95.4 bits (236),  Expect = 2e-18, Method: Composition-based stats.
 Identities = 23/115 (20%), Positives = 45/115 (39%), Gaps = 2/115 (1%)

Query  21   FVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPAN  80
            F      + ++       AR     +G+RV GD L+ R+        +  V +T
Sbjct  8    FFTIFATLAVSAQSESTVARQNDLVIGRRVEGDYLLNRSFIKKDASWIGKV-VTFTKTFT  66

Query  81   GGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            G   + ++ V  L    +   +  V+ GG+    + + +E+       ++ EIYG
Sbjct  67   GDGVSTITQVKALEQHENG-ASAKVIAGGVDYLYVTLRLESERYRHIDFVVEIYG  120


>☐ ref|XP_002071834.1| G GK10198 [Drosophila willistoni]
 gb|EDW82820.1| G GK10198 [Drosophila willistoni]
Length=115

 GENE ID: 6648731 Dwil\GK10198 | GK10198 gene product from transcript GK10198-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score = 95.0 bits (235),  Expect = 2e-18, Method: Composition-based stats.
 Identities = 17/115 (14%), Positives = 45/115 (39%), Gaps = 2/115 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   L     +    A + + G ++S   L+     +    +      T+++P +G
Sbjct  1    MKLLQFVLVFSGILALGYAANATWGSKLSNSVLLNTQNVTLYKKNNDYQSSTISFPTSGS  60

Query  83   YG-AVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                 ++Y+++    T++ G    +  GG+G     I ++++     +  A+ +G
Sbjct  61   VNTKTITYITVTDRFTNSSGPTSTLWTGGVGFTFATINVKSQYAQGINVTAQFWG  115


>☐ gb|ABI20159.1|  16 kDa salivary protein [Phlebotomus duboscqi]
Length=161

 Score = 94.6 bits (234),  Expect = 3e-18, Method: Composition-based stats.
 Identities = 49/136 (36%), Positives = 68/136 (50%), Gaps = 2/136 (1%)

Query  2    APDPETVSAKITLVEDVDEFVRANPGVTLTKLDNVRTARA-IHNSLGQRVSGDRLVARNQ  60
              D          V  + +FV  NP   L  +      R  I+  LG R  GDRLV
Sbjct  27   DDDAPDRQVNFLSVNSLADFVAENPEAELEMMHMYSDYRGQINYVLGARQRGDRLVGTAS  86

Query  61   DSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIE  120
            +   W   QNV + + YP  G  GA+V++VS+ + Q++NLG  YV  GGI QR I I++E
Sbjct  87   NGHRWPRPQNVDIEIKYPTKG-RGALVTFVSLNITQSTNLGRAYVTKGGIHQRNITIVVE  145

Query  121  ARNTYTFSYIAEIYGL  136
            A +T    YI  I+G+
Sbjct  146  AYSTTFVQYICNIFGM  161


>☐ gb|ABD97962.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 94.6 bits (234),  Expect = 3e-18, Method: Composition-based stats.
 Identities = 18/112 (16%), Positives = 37/112 (33%), Gaps = 2/112 (1%)
```

```
Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   +  L+          + G    G +L+ + +   +   L+ V   V++
Sbjct  8    IFVAIAIVAVLNVAFAGHDGTTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQ  65

Query  83   YGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                +S + I     S  G   +L GG   +   I   +   +  ++  EIY
Sbjct  66   QPYKLSAIVITDHSESKGGTASLLEGGPPGKCAVIGFRSDRNHGLNFTLEIY  117


>☐ gb|ABD97963.1|  ACP225 [Drosophila yakuba]
Length=121

 Score = 92.7 bits (229),  Expect = 1e-17, Method: Composition-based stats.
 Identities = 17/96 (17%), Positives = 34/96 (35%), Gaps = 2/96 (2%)

Query  39   ARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTS  98
                  + G    G +L+ + +   +   L+ V   V++         +S + I     S
Sbjct  24   GHDGTTTFGVLKPGAKLIHKEEIVEAKKFLRVVTRKVHF--KPQQPYKLSAIVITDHSES  81

Query  99   NLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
              G   +L GG   +   I   +   +  ++  EIY
Sbjct  82   KGGTASLLEGGPPGKFAVIGFRSDRNHGLNFTLEIY  117


>☐ ref|XP_001604205.1| G PREDICTED: hypothetical protein [Nasonia vitripennis]
Length=137

 GENE ID: 100120578 LOC100120578 | hypothetical protein LOC100120578
[Nasonia vitripennis]

 Score = 92.3 bits (228),  Expect = 1e-17, Method: Composition-based stats.
 Identities = 23/110 (20%), Positives = 43/110 (39%), Gaps = 4/110 (3%)

Query  25   NPGVTLTKLDNVRTARA-IHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGY  83
               +  T L  V  A+   H  +  + S D L+           LQ V+LT  +P  G
Sbjct  7    VLVLLATSLTIVLAAKPPNHLIIDAKKSSDHLLQIESVGKPSKWLQKVKLTKTFP--GNN  64

Query  84   GAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAE  132
             + ++ V +L       G    +L GG G   + +  ++   ++  Y+ +
Sbjct  65   HSNITQVQLLDQDQKGNGTTAAILTGGPGNNFVTVQFKSTRGHSVDYVVK  114


>☐ ref|XP_001661180.1| UG hypothetical protein AaeL_AAEL002263 [Aedes aegypti]
 gb|EAT46579.1| G conserved hypothetical protein [Aedes aegypti]
Length=118

 GENE ID: 5574109 AaeL_AAEL002263 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score = 92.3 bits (228),  Expect = 1e-17, Method: Composition-based stats.
 Identities = 25/117 (21%), Positives = 45/117 (38%), Gaps = 5/117 (4%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY--PA  79
               +  V L  +  +  A +     G+RV GDRL+       +   +        Y  P
Sbjct  3    TILSTLVVLLAVTALIYADSN--QWGRRVVGDRLLQSATIVNNSLPVPIKTAVFQYLPPQ  60

Query  80   NGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                   V+Y+    +  + LG    ++ GGI Q    + + +R    F++  E+YG
Sbjct  61   GSARAPNVTYILARDNVRNGLGGFAALISGGINQTHATVKLSSRPNLGFNFTIEVYG  117


>☐ ref|XP_002028227.1| G GL16228 [Drosophila persimilis]
 gb|EDW38776.1| G GL16228 [Drosophila persimilis]
Length=122

 GENE ID: 6603250 Dper\GL16228 | GL16228 gene product from transcript GL16228-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score = 91.9 bits (227),  Expect = 2e-17, Method: Composition-based stats.
 Identities = 21/120 (17%), Positives = 41/120 (34%), Gaps = 6/120 (5%)

Query  21   FVRANPGVTLTKLDNV-----RTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTV  75
            F      +  + +  +      ++     + G +++ D L+AR   + +   LQ    T
Sbjct  3    FSSVFLLLLGSYVCLINGLGGSSSEGNDYTWGTKLTTDTLIAREVITKTKMLLQTTTKTY  62

Query  76   NYPANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
```

```
                 G     +SY+ I   +        +  GG+G   + I   +          EIYG
Sbjct  63   TLTQAG-TAKTISYIHITDLKRMRGATAEITSGGVGDTTVTIKFTSSRGSGIKSQVEIYG  121


>☐ ref|XP_002091873.1| G GE11995 [Drosophila yakuba]
 gb|EDW91585.1| G GE11995 [Drosophila yakuba]
Length=83

 GENE ID: 6531037 Dyak\GE11995 | GE11995 gene product from transcript GE11995-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score = 91.9 bits (227),  Expect = 2e-17, Method: Composition-based stats.
 Identities = 15/84 (17%), Positives = 30/84 (35%), Gaps = 2/84 (2%)

Query  54   RLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGN-GYVLYGGIGQ  112
              + R   +     L+ V     +         ++ + I         G   Y+  GG
Sbjct  1    MHLRREIITEKSKFLRVVTRDYVFEQK-KLARTITQIVITDQVRDGNGGHAYLTEGGPQT  59

Query  113  RQIGILIEARNTYTFSYIAEIYGL  136
                I ++++    FS+I +IYG+
Sbjct  60   TYAKIHLKSQRNQGFSFIIDIYGI  83


>☐ ref|XP_001661179.1| UG hypothetical protein AaeL_AAEL002263 [Aedes aegypti]
 gb|EAT46578.1| G conserved hypothetical protein [Aedes aegypti]
Length=117

 GENE ID: 5574109 AaeL_AAEL002263 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score = 91.5 bits (226),  Expect = 2e-17, Method: Composition-based stats.
 Identities = 26/116 (22%), Positives = 46/116 (39%), Gaps = 4/116 (3%)

Query  22   VRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANG  81
               +  V L  +  +  A +     G+RV GDRL+       +   +        Y   G
Sbjct  3    TILSTLVVLLAVTALIYADSN--QWGRRVVGDRLLQSATIVNNSLPVPIKTAVFQYLPPG  60

Query  82   GYGA-VVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
               A  V+Y+    +  + LG    ++ GGI Q    + + +R    F++  E+YG
Sbjct  61   SARAPNVTYILARDNVRNGLGGFAALISGGINQTHATVKLSSRPNLGFNFTIEVYG  116


>☐ gb|ABD97967.1|  ACP225 [Drosophila teissieri]
Length=120

 Score = 90.8 bits (224),  Expect = 4e-17, Method: Composition-based stats.
 Identities = 20/90 (22%), Positives = 36/90 (40%), Gaps = 3/90 (3%)

Query  45   SLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLGNGY  104
            S G    G +L+ + +   S   L+ V   V++         +S + I     SN G
Sbjct  30   SFGVLKPGAKLIHKEEIVESKKFLRVVTRKVHFKPQ---PFKLSAIVITDHSESNGGTAS  86

Query  105  VLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
            +L GG   +   I   ++  +  ++  EIY
Sbjct  87   LLEGGPPGKFAVIGFRSKRNHGLNFTLEIY  116


>☐ gb|ACY69959.1|  putative putative salivary secreted protein [Cimex lectularius]
Length=136

 Score = 90.8 bits (224),  Expect = 4e-17, Method: Composition-based stats.
 Identities = 16/123 (13%), Positives = 40/123 (32%), Gaps = 8/123 (6%)

Query  21   FVRANPGVTLTKLDN----VRTARAIHNSLGQRVSGD---RLVARNQDSASWGSLQNVQL  73
            F+     ++  +  +     R     +   G +   D   +L+            ++
Sbjct  8    FLAILLNISAQQSASTTRRCRKNDQHNAVFGTKQPDDNTQQLLRSENIKVGPRQKKSTTY  67

Query  74   TVNYPANGGYGAVVSYVSILVDQTSNLGNG-YVLYGGIGQRQIGILIEARNTYTFSYIAE  132
               YP  G     ++Y+          G   +++ GG+G   + +     + +    ++
Sbjct  68   PFVYPEQGDIRGNITYIEASDRTADGKGGCVFLIEGGVGYDNVTLHFRTQKSLGMEFVVS  127

Query  133  IYG  135
            IYG
Sbjct  128  IYG  130
```

```
>☐ gb|AAY34548.1|  putative 11.4 kDa secreted salivary gland protein [Stomoxys calcitrans]
Length=123

 Score = 90.4 bits (223),  Expect = 5e-17, Method: Composition-based stats.
 Identities = 18/97 (18%), Positives = 36/97 (37%), Gaps = 2/97 (2%)

Query  41   AIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVV-SYVSILVD-QTS  98
              + + G     DR++AR++           + T+ YP  G     + + + +
Sbjct  23   GNNATWGHIGYFDRVIARDRVEHPANWFTKHKETILYPPKGHENFNILTAIRVTDYGGPK  82

Query  99   NLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            N G   +L GG G     I + ++     +   E +G
Sbjct  83   NFGQATLLQGGPGHFNATIEVRSQTRRPLNMTIEYFG  119


>☐ ref|XP_976138.1| UG PREDICTED: hypothetical protein [Tribolium castaneum]
 gb|EFA01092.1| G hypothetical protein TcasGA2_TC004015 [Tribolium castaneum]
Length=169

 GENE ID: 655848 TcasGA2_TC004015 | hypothetical protein LOC655848
[Tribolium castaneum] (10 or fewer PubMed links)

 Score = 90.0 bits (222),  Expect = 8e-17, Method: Composition-based stats.
 Identities = 20/99 (20%), Positives = 38/99 (38%), Gaps = 4/99 (4%)

Query  41   AIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPA--NGGYGA-VVSYVSILVDQT  97
            +    +GQ    DRL+          S    +  + YP    G +    ++ + I
Sbjct  70   SGDVIIGQIGYLDRLLFNQIYYKPRRSWSGREKIIEYPKDMTGKHNYETINSIRIYNKFY  129

Query  98   SN-LGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
                    +L GG+G R + I + ++    F Y+ +I+G
Sbjct  130  DGLDSKAKILSGGVGSRFVKIKLSSKRNKGFKYLVQIFG  168


>☐ gb|ACI30202.1|  hypothetical conserved mosquito protein [Anopheles darlingi]
Length=119

 Score = 90.0 bits (222),  Expect = 8e-17, Method: Composition-based stats.
 Identities = 22/111 (19%), Positives = 43/111 (38%), Gaps = 3/111 (2%)

Query  28   VTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY-PANGG-YGA  85
            + L    +V          G+RV+GDR +       +   +        Y PA G
Sbjct  8    LVLFLSLSVVLIYGESRQWGRRVAGDRQLDYTVLVNNSLPVPQKSSIYRYLPAPGAYRPP  67

Query  86   VVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
             ++ +    +  +  G    ++ GGI Q  + + + ++    F++  EIYG
Sbjct  68   NITAIYARDNVPAGRGGYAAIVSGGINQANVTVKLTSQPYQRFNFTIEIYG  118


>☐ ref|XP_002090966.1| G GE13403 [Drosophila yakuba]
 gb|AAQ09833.1| G CG13324 [Drosophila yakuba]
 gb|EDW90678.1| G GE13403 [Drosophila yakuba]
Length=112

 GENE ID: 6530020 Dyak\GE13403 | GE13403 gene product from transcript GE13403-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score = 89.6 bits (221),  Expect = 9e-17, Method: Composition-based stats.
 Identities = 23/98 (23%), Positives = 43/98 (43%), Gaps = 4/98 (4%)

Query  39   ARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTS  98
              A   + G+RV+ D L++R ++  +        + VNYPA       +S V +  +  +
Sbjct  17   CHAYSATWGRRVNNDFLLSRTREVRNPIKNNYWNVNVNYPA---GFYNISAVIVYDNFKN  73

Query  99   NLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            N G +  +  GG G R   + +  +     +   EI+G
Sbjct  74   NSGASPSLYSGGPGYRFATVNLRGQVNRGINSTVEIWG  111


>☐ ref|XP_002040057.1| G GM15559 [Drosophila sechellia]
 gb|EDW56922.1| G GM15559 [Drosophila sechellia]
Length=121

 GENE ID: 6615686 Dsec\GM15559 | GM15559 gene product from transcript GM15559-RA
```

```
[Drosophila sechellia] (10 or fewer PubMed links)

 Score = 89.6 bits (221),  Expect = 1e-16, Method: Composition-based stats.
 Identities = 19/117 (16%), Positives = 34/117 (29%), Gaps = 1/117 (0%)

Query  19   DEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
              F+            N  +      + G + + D L+A    + S   L     T
Sbjct  5    AVFLLLVGTHICFIDANFGSGEGNDYTYGTQATTDTLIASETVTKSKALLGTTTKTYTLT  64

Query  79   ANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     ++Y+ I   +        +  GG+G   + I   +           IYG
Sbjct  65   QAG-TAKTITYIKITDLKRMRGATAEITSGGVGSTTVTIKFTSARGAGIKSQVVIYG  120


>☐ ref|XP_002082759.1| G GD25061 [Drosophila simulans]
 gb|EDX08344.1| G GD25061 [Drosophila simulans]
Length=121

 GENE ID: 6735849 Dsim\GD25061 | GD25061 gene product from transcript GD25061-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score = 89.6 bits (221),  Expect = 1e-16, Method: Composition-based stats.
 Identities = 19/117 (16%), Positives = 33/117 (28%), Gaps = 1/117 (0%)

Query  19   DEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
              F+            N  +      + G   + D L+A    + S   L     T
Sbjct  5    GVFLLLVGTHICFIDANFGSGEGNDYTYGTHATTDTLIASETVTKSKALLGTTTKTYTLT  64

Query  79   ANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     ++Y+ I   +        +  GG+G   + I   +           IYG
Sbjct  65   QAG-TAKTITYIKITDLKRMRGATAEITSGGVGSTTVTIKFTSARGAGIKSQVVIYG  120


>☐ ref|XP_002086118.1| G GE14474 [Drosophila yakuba]
 ref|XP_002090583.1| G GE13194 [Drosophila yakuba]
 gb|EDW90295.1| G GE13194 [Drosophila yakuba]
 gb|EDW99588.1| G GE14474 [Drosophila yakuba]
Length=117

 GENE ID: 6539219 Dyak\GE14474 | GE14474 gene product from transcript GE14474-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score = 89.2 bits (220),  Expect = 1e-16, Method: Composition-based stats.
 Identities = 22/114 (19%), Positives = 47/114 (41%), Gaps = 2/114 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 + L     V  A A + S G ++S  +LV+    +    + Q V   +++P  G
Sbjct  1    MKVLQIALVFSAIVAIAFAANASWGAKLSSSKLVSTQNVTLHKKANQYVSSLISFPLTGQ  60

Query  83   YG-AVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                 + Y+SI    T++ G +  +  GG G     + + ++ +   +  A+ +
Sbjct  61   TNTKTIRYISITDRFTNSSGPSSTLWSGGPGFISAKVNVTSQLSQGINVTAQFW  114


>☐ gb|ABR27885.1|  salivary secreted protein [Triatoma infestans]
Length=129

 Score = 89.2 bits (220),  Expect = 1e-16, Method: Composition-based stats.
 Identities = 19/121 (15%), Positives = 41/121 (33%), Gaps = 5/121 (4%)

Query  20   EFVRANPGVTLTKLDNVRTARA-IHNSLGQRVSGDRLVARNQDSA-SWGSLQNVQLTVNY  77
                                       +G++   D ++ +   S  +W     V   V Y
Sbjct  6    VITLLLVTCLYVTYGWAPHQYGNHSLIIGRKGYNDVILFQKTVSKKNWNPFGKVSEDVTY  65

Query  78   PANGGYGAV--VSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
            P     G    ++ +  +   +++ G   YVL GGI +  + I  +++    + +   I+
Sbjct  66   PVRPIPGRRPLITEIDAIDLDSNDKGGYAYVLKGGINRNNVTIHFKSQKGRGYKFNLTIF  125

Query  135  G  135
            G
Sbjct  126  G  126


>☐ ref|XP_316478.4| UG AGAP006442-PA [Anopheles gambiae str. PEST]
```

```
 gb|EAA11094.5|  AGAP006442-PA [Anopheles gambiae str. PEST]
Length=119

 GENE ID: 1277050 AgaP_AGAP006442 | AGAP006442-PA [Anopheles gambiae str. PEST]
(10 or fewer PubMed links)

 Score = 88.8 bits (219),  Expect = 2e-16, Method: Composition-based stats.
 Identities = 21/99 (21%), Positives = 39/99 (39%), Gaps = 3/99 (3%)

Query  40   RAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY-PANGG-YGAVVSYVSILVDQT  97
                   G+RVSGDRL+       +   +        Y PA G      ++ +    +
Sbjct  20   YGDSRQWGRRVSGDRLLDYAVVVNNSLPVPEKSSIYRYLPAAGAYRPPNITAIFTRDNVP  79

Query  98   SNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
               G    ++ GG+ Q  + + + ++    F++  EIYG
Sbjct  80   GGKGGFAGIVSGGVNQSNVTLKLTSKPYQRFNFTIEIYG  118


>☐ ref|XP_001976402.1|  GG22852 [Drosophila erecta]
 gb|EDV56802.1|  GG22852 [Drosophila erecta]
Length=133

 GENE ID: 6547809 Dere\GG22852 | GG22852 gene product from transcript GG22852-RA
[Drosophila erecta] (10 or fewer PubMed links)

 Score = 88.4 bits (218),  Expect = 2e-16, Method: Composition-based stats.
 Identities = 17/107 (15%), Positives = 29/107 (27%), Gaps = 10/107 (9%)

Query  36   VRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNY-PANGGYGAVVSYVSILV  94
            V        + G       ++         GS       V Y   NG     ++ V +
Sbjct  29   VLLGNGASVTWGHVHKYASVLHTEDVYLPPGSPG---RDVAYGSQNGSDNYRITGVHVED  85

Query  95   DQTSNLGNGYVLYGGIGQRQIGILIEARNT-----YTFSYIAEIYGL  136
               +      +  GGIGQ+ + +                    IYG+
Sbjct  86   LG-AQGAEAVLTAGGIGQQFVVLHCRRLRGAEEQPQAAHLEVSIYGV  131


>☐ ref|XP_001987096.1|  GH20161 [Drosophila grimshawi]
 gb|EDW01963.1|  GH20161 [Drosophila grimshawi]
Length=122

 GENE ID: 6560517 Dgri\GH20161 | GH20161 gene product from transcript GH20161-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score = 88.4 bits (218),  Expect = 2e-16, Method: Composition-based stats.
 Identities = 17/117 (14%), Positives = 37/117 (31%), Gaps = 1/117 (0%)

Query  19   DEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
              F+          ++ +  A     + G + +   L+AR     S   L     T
Sbjct  6    AIFILTVCASMALGINALGNAAEGDYTYGAKQATSELIARETIITSKTLLGTTTRTYTLV  65

Query  79   ANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     + Y++I   +        +  GG+    + +   ++ +       EI+G
Sbjct  66   QAG-TPKTIHYINIRNLKRRRGATATIESGGVTATTLTVRFTSQRSSGIKSEIEIWG  121


>☐ ref|XP_002033648.1|  GM20316 [Drosophila sechellia]
 gb|EDW47661.1|  GM20316 [Drosophila sechellia]
Length=112

 GENE ID: 6608934 Dsec\GM20316 | GM20316 gene product from transcript GM20316-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score = 88.1 bits (217),  Expect = 3e-16, Method: Composition-based stats.
 Identities = 21/98 (21%), Positives = 41/98 (41%), Gaps = 4/98 (4%)

Query  39   ARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTS  98
              A   + G+RV+ D L++R ++  +        + VNYP        +S V +  +  +
Sbjct  17   CHAYSATWGRRVNNDFLLSRTREVRNPIKNNYWNVNVNYP---NGFYNISAVIVYDNFKN  73

Query  99   NLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            + G +  +  GG G R   + +  +         EI+G
Sbjct  74   SSGASPSLYSGGPGYRFATVNLRGQVNRGIDSTVEIWG  111
```

```
>☐ gb|ABN54491.1|  conserved secreted protein [Oncopeltus fasciatus]
Length=116

 Score = 87.7 bits (216),  Expect = 4e-16, Method: Composition-based stats.
 Identities = 18/99 (18%), Positives = 38/99 (38%), Gaps = 1/99 (1%)

Query  27   GVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAV  86
             VT           + +  LG R  GD+L+    +      L+     V++P +G    V
Sbjct  15   LVTCYPKTACGDGPSHNLLLGNRTYGDKLLYSGSEHIDSSLLRVKTKDVHWPLHGVSPEV  74

Query  87   VSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNT  124
            ++ + ++       G   ++  GG G R   + ++ +
Sbjct  75   ITRLEVVDKAKDGSGGCAFLSGGGPGSRVAKLHLKTQRG  113


>☐ ref|XP_001976379.1| G GG20045 [Drosophila erecta]
 gb|EDV56779.1| G GG20045 [Drosophila erecta]
Length=121

 GENE ID: 6547443 Dere\GG20045 | GG20045 gene product from transcript GG20045-RA
[Drosophila erecta] (10 or fewer PubMed links)

 Score = 87.7 bits (216),  Expect = 4e-16, Method: Composition-based stats.
 Identities = 20/117 (17%), Positives = 34/117 (29%), Gaps = 1/117 (0%)

Query  19   DEFVRANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP  78
              F+            N  +      + G + + D L+A    + S   L     T
Sbjct  5    GVFLLLVGTHVCFIDANFGSGEGNDYTYGTQATTDTLIASETVTKSKSLLGTTTKTYTLT  64

Query  79   ANGGYGAVVSYVSILVDQTSNLGNGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
              G     +SY+ I   +        +  GG+G   + I   +           IYG
Sbjct  65   QAG-TAKTISYIKITDLKRMRGATAEITSGGVGSTTVTIKFTSARGGGIKSQVVIYG  120


>☐ ref|XP_002090965.1| G GE13402 [Drosophila yakuba]
 gb|EDW90677.1| G GE13402 [Drosophila yakuba]
Length=112

 GENE ID: 6530019 Dyak\GE13402 | GE13402 gene product from transcript GE13402-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score = 87.3 bits (215),  Expect = 4e-16, Method: Composition-based stats.
 Identities = 22/98 (22%), Positives = 42/98 (42%), Gaps = 4/98 (4%)

Query  39   ARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGGYGAVVSYVSILVDQTS  98
              A   + G+RV+ D L++R ++  +        + VNYPA       +S V +  +  +
Sbjct  17   CHAYSATWGRRVNNDFLLSRTREVRNPIKNNYWNVNVNYPA---GFYNISAVIVYDNFKN  73

Query  99   NLG-NGYVLYGGIGQRQIGILIEARNTYTFSYIAEIYG  135
            N G +  +  GG G R   + +  +         E++G
Sbjct  74   NSGASPSLYSGGPGYRFATVNLRGQVNRGIDSTVEVWG  111


>☐ ref|NP_724118.1| UG CG31789 [Drosophila melanogaster]
 gb|AAN11008.1| G CG31789 [Drosophila melanogaster]
 gb|ACL68720.1| G LP09219p [Drosophila melanogaster]
Length=117

 GENE ID: 318943 CG31789 | CG31789 gene product from transcript CG31789-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score = 87.3 bits (215),  Expect = 4e-16, Method: Composition-based stats.
 Identities = 21/114 (18%), Positives = 45/114 (39%), Gaps = 2/114 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                   L     +  A A + S G ++S  +LV+    +    + Q V   +++P  G
Sbjct  1    MKALQFALVFSAILAIAFAANASWGAKLSSSKLVSTQNVTIYKKANQYVSSLISFPLAGQ  60

Query  83   YG-AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                 + Y+SI    T++ G    +  GG G     + + ++ +   +  A+ +
Sbjct  61   SNTKTIRYISITDRFTNSSGPYSTLWSGGPGFINAQVNVTSQFSQGINVTAQFW  114


>☐ ref|XP_002038990.1| G GM17281 [Drosophila sechellia]
```

```
 gb|EDW55636.1|  GM17281 [Drosophila sechellia]
Length=117

 GENE ID: 6614562 Dsec\GM17281 | GM17281 gene product from transcript GM17281-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score = 87.3 bits (215),  Expect = 5e-16, Method: Composition-based stats.
 Identities = 22/114 (19%), Positives = 46/114 (40%), Gaps = 2/114 (1%)

Query  23   RANPGVTLTKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYPANGG  82
                 + L     V  A A + S G ++S  +LV+    +    + Q V   +++P  G
Sbjct  1    MKALQIALVFSAIVAIAFAANASWGAKLSSSKLVSTQNVTIYKKANQYVSSIISFPLAGQ  60

Query  83   YG-AVVSYVSILVDQTSNLGN-GYVLYGGIGQRQIGILIEARNTYTFSYIAEIY  134
                 + Y+SI    T++ G    +  GG G     + + ++ +   +  A+ +
Sbjct  61   SNTKTIRYISITDRFTNSSGPYSTLWSGGPGFISAKVNVTSQFSQGINVTAQFW  114


>☐  ref|NP_726325.1|  CG30411 [Drosophila melanogaster]
 gb|AAL48593.1|  RE06996p [Drosophila melanogaster]
 gb|AAM68251.1|  CG30411 [Drosophila melanogaster]
 gb|ACL86455.1|  CG30411-PA [synthetic construct]
 gb|ACL91066.1|  CG30411-PA [synthetic construct]
Length=141

 GENE ID: 246600 CG30411 | CG30411 gene product from transcript CG30411-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score = 86.9 bits (214),  Expect = 5e-16, Method: Composition-based stats.
 Identities = 16/108 (14%), Positives = 29/108 (26%), Gaps = 11/108 (10%)

Query  36   VRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP-ANGGYGAVVSYVSILV  94
            V    A   + G       ++       + GS       V Y          ++ V +
Sbjct  36   VLLGNAASVTWGHVNKYASVLHTEDVYLTPGSPG---RDVAYGLQKASDNYRITGVHVED  92

Query  95   DQTSNLGNGYVLYGGIGQRQIGILIE------ARNTYTFSYIAEIYGL  136
                      +  GGIGQ+ + +          +          IYG+
Sbjct  93   LGAD-EAEAVLTSGGIGQQFVVLRCRRLRDSEGQPERAAHLEVSIYGV  139


>☐  ref|XP_002082783.1|  GD11761 [Drosophila simulans]
 gb|EDX08368.1|  GD11761 [Drosophila simulans]
Length=138

 GENE ID: 6735873 Dsim\GD11761 | GD11761 gene product from transcript GD11761-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score = 86.9 bits (214),  Expect = 6e-16, Method: Composition-based stats.
 Identities = 16/105 (15%), Positives = 28/105 (26%), Gaps = 8/105 (7%)

Query  36   VRTARAIHNSLGQRVSGDRLVARNQDSASWGSLQNVQLTVNYP-ANGGYGAVVSYVSILV  94
            V    A   + G       ++          G+       V Y   N      ++ V +
Sbjct  36   VLLGNAASVTWGHVNKYASVLHTEDVYLPPGTPG---RDVAYGLENASDNYRITGVHVED  92

Query  95   DQTSNLGNGYVLYGGIGQRQIGILIEARNT---YTFSYIAEIYGL  136
                      +  GGIGQ+ + +                  IYG+
Sbjct  93   LGADG-AEALLTAGGIGQQFVVLHCRRLPNSEEQAAHLEVSIYGV  136


>☐  ref|XP_001604117.1|  PREDICTED: similar to HDC07203 [Nasonia vitripennis]
Length=133

 GENE ID: 100120478 LOC100120478 | similar to HDC07203 [Nasonia vitripennis]

 Score = 86.5 bits (213),  Expect = 8e-16, Method: Composition-based stats.
 Identities = 18/130 (13%), Positives = 40/130 (30%), Gaps = 13/130 (10%)

Query  18   VDEFVRANPGVTL-----------TKLDNVRTARAIHNSLGQRVSGDRLVARNQDSASWG  66
            +  F   +    +             L+ V    +    +G R + D LV     +  +
Sbjct  1    MTSFKTLSLVAIVLVVAATVSSSPVPLEKVEVFTSNDLVVGSRQARDSLVYTESITEKFL  60

Query  67   SLQNVQLTVNYPANGGYGAVVSYVSILVDQTSNLG-NGYVLYGGIGQRQIGILIEARNTY  125
             +   +L +        G V++ V          G    V+ GG     I +  +++
Sbjct  61   -VSGKKLIITRNIQVPNGYVITQVRATDKAADGTGAEATVIGGGPDMEFISLRFKSQRWR  119

Query  126  TFSYIAEIYG  135
```

```
                 +  ++Y
Sbjct  120  GIYFDIQVYA  129
```