PSI blast for Simulium/Culicoides insect-conserved proteins

gi|221768755|gb|ACM40903.1| secreted salivary...

Results for: lcl|50231 gi|221768755|gb|ACM40903.1| secreted salivary protein [Culicoides nubeculosus](398aa) ▾

Skip to the first new sequence

Your BLAST job specified more than one input sequence. This box lets you choose which input sequence to show BLAST results for.

Query ID
    gi|221768755|gb|ACM40903.1| secreted salivary protein [Culicoides nubeculosus]
Molecule type
    amino acid
Query Length
    398
Database Name
    nr
Description
    All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects See details
Program
    BLASTP 2.2.22+ Citation

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

                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

  ● ☑  ref|XP_002101210.1|  GE17494 [Drosophila yakuba] >gb|EDX02318....   524    9e-147 G
  ● ☑  ref|XP_001978593.1|  GG17594 [Drosophila erecta] >gb|EDV47520....   522    3e-146 G
  ● ☑  ref|XP_001963941.1|  GF20993 [Drosophila ananassae] >gb|EDV343...   517    9e-145 G
  ● ☑  ref|XP_002067370.1|  GK16218 [Drosophila willistoni] >gb|EDW78...   516    1e-144 G
  ● ☑  ref|XP_002057094.1|  GJ16894 [Drosophila virilis] >gb|EDW62580...   514    5e-144 G
  ● ☑  ref|XP_002025529.1|  GL15246 [Drosophila persimilis] >gb|EDW31...   513    2e-143 G
  ● ☑  ref|XP_001355344.2|  GA15301 [Drosophila pseudoobscura pseudoo...   512    2e-143 G
  ● ☑  ref|XP_002076900.1|  GD24578 [Drosophila simulans] >gb|EDX1649...   511    4e-143 G
  ● ☑  ref|NP_572427.1|  lethal (1) G0193, isoform A [Drosophila mela...   508    5e-142 UG
  ● ☑  ref|XP_002010644.1|  GI21656 [Drosophila mojavensis] >gb|EDW06...   506    2e-141 G
  ● ☑  ref|NP_996366.1|  lethal (1) G0193, isoform B [Drosophila mela...   501    4e-140 UG
  ● ☑  ref|XP_001992028.1|  GH24541 [Drosophila grimshawi] >gb|EDV917...   492    3e-137 G
  ● ☑  ref|XP_001851110.1|  conserved hypothetical protein [Culex qui...   490    9e-137 UG
  ● ☑  ref|XP_969643.2|  PREDICTED: similar to GA15301-PA [Tribolium ...   484    9e-135 UG
  ● ☑  gb|EFA08554.1|  hypothetical protein TcasGA2_TC006209 [Triboli...   478    4e-133
  ● ☑  gb|AAY55307.1|  IP12501p [Drosophila melanogaster]                  478    5e-133
  ● ☑  ref|XP_001851111.1|  conserved hypothetical protein [Culex qui...   475    4e-132 G
  ● ☑  ref|XP_002077791.1|  GD22890 [Drosophila simulans] >gb|EDX0337...   474    9e-132 G
  ● ☑  ref|XP_001960321.1|  GF11569 [Drosophila ananassae] >gb|EDV371...   474    9e-132 G
  ● ☑  ref|XP_002077790.1|  GD22891 [Drosophila simulans] >gb|EDX0337...   474    1e-131 G
  ● ☑  gb|AAY55318.1|  IP12601p [Drosophila melanogaster]                  473    2e-131
```

●	☑	gb|ACH95281.1|	FI07603p [Drosophila melanogaster]	473	2e-131	
●	☑	ref|NP_722735.1|	CG31664 [Drosophila melanogaster] >gb|AAF513...	471	5e-131	UG
●	☑	ref|XP_001648416.1|	hypothetical protein AaeL_AAEL014309 [Aed...	470	1e-130	UG
●	☑	ref|XP_002087581.1|	GE15356 [Drosophila yakuba] >gb|EDW87293....	469	2e-130	G
●	☑	ref|XP_001968326.1|	GG24571 [Drosophila erecta] >gb|EDV57385....	464	8e-129	G
●	☑	ref|XP_002044470.1|	GM11959 [Drosophila sechellia] >gb|EDW536...	463	1e-128	G
●	☑	ref|XP_001851109.1|	conserved hypothetical protein [Culex qui...	462	3e-128	UG
●	☑	ref|XP_001648417.1|	hypothetical protein AaeL_AAEL014309 [Aed...	462	3e-128	UG
●	☑	ref|XP_002015853.1|	GL10792 [Drosophila persimilis] >gb|EDW31...	462	4e-128	G
●	☑	gb|ACM40903.1|	secreted salivary protein [Culicoides nubeculo...	459	4e-127	
●	☑	ref|XP_001360382.2|	GA15824 [Drosophila pseudoobscura pseudoo...	456	2e-126	G
●	☑	ref|XP_002092594.1|	GE14278 [Drosophila yakuba] >gb|EDW92306....	451	7e-125	G
●	☑	ref|XP_002427759.1|	conserved hypothetical protein [Pediculus...	447	1e-123	G
●	☑	ref|XP_001946173.1|	PREDICTED: similar to GA15301-PA [Acyrtho...	443	2e-122	UG
●	☑	ref|XP_001121007.1|	PREDICTED: similar to lethal (1) G0193 CG...	439	2e-121	UG
●	☑	ref|XP_002049066.1|	GJ21381 [Drosophila virilis] >gb|EDW60259...	437	9e-121	G
●	☑	ref|XP_002062683.1|	GK19560 [Drosophila willistoni] >gb|EDW73...	429	2e-118	G
●	☑	ref|XP_001603409.1|	PREDICTED: similar to CG2206-PA [Nasonia ...	412	3e-113	UG
●	☑	ref|XP_001603438.1|	PREDICTED: similar to GA15301-PA [Nasonia...	405	4e-111	UG
●	☑	gb|EFA08551.1|	hypothetical protein TcasGA2_TC006206 [Triboli...	397	2e-108	
●	☑	ref|XP_969565.1|	PREDICTED: similar to lethal (1) G0193 CG220...	392	4e-107	UG
●	☑	ref|XP_310721.4|	AGAP000385-PA [Anopheles gambiae str. PEST] ...	389	2e-106	UG
●	☑	ref|XP_001603377.1|	PREDICTED: similar to CG2206-PA [Nasonia ...	385	3e-105	G
●	☑	ref|XP_001606530.1|	PREDICTED: similar to CG2206-PA [Nasonia ...	378	7e-103	G
●	☑	gb|EFA08552.1|	hypothetical protein TcasGA2_TC006207 [Triboli...	372	6e-101	
●	☑	gb|EFA12872.1|	hypothetical protein TcasGA2_TC010724 [Triboli...	366	3e-99	
●	☑	gb|ACM40897.1|	putative cysteine endopeptidase [Culicoides nu...	365	6e-99	
●	☑	ref|XP_001851115.1|	conserved hypothetical protein [Culex qui...	360	2e-97	UG
●	☑	ref|XP_001606517.1|	PREDICTED: similar to CG2206-PA [Nasonia ...	352	6e-95	UG
●	☑	gb|ACZ28336.1|	hypothetical conserved protein [Simulium nigri...	350	2e-94	
●	☑	ref|XP_001601809.1|	PREDICTED: similar to GA15301-PA [Nasonia...	347	1e-93	G
●	☑	ref|XP_001600223.1|	PREDICTED: similar to GA15301-PA [Nasonia...	343	2e-92	G
●	☑	ref|XP_001602724.1|	PREDICTED: similar to CG2206-PA [Nasonia ...	340	2e-91	G
●	☑	ref|XP_001810077.1|	PREDICTED: similar to GA15301-PA [Triboli...	337	1e-90	UG
●	☑	gb|ACE75210.1|	conserved hypothetical protein [Glyptapanteles...	330	2e-88	
●	☑	ref|XP_001601835.1|	PREDICTED: similar to ef hand protein [Na...	322	7e-86	UG
●	☑	gb|ABY19389.1|	venom protein 2 [Microctonus hyperodae]	314	1e-83	
●	☑	gb|ACZ28338.1|	hypothetical conserved protein [Simulium nigri...	309	5e-82	
●	☑	ref|XP_001603579.1|	PREDICTED: similar to GA15301-PA [Nasonia...	308	5e-82	UG
●	☑	gb|ACH56860.1|	hypothetical protein [Simulium vittatum]	307	2e-81	
●	☑	ref|XP_002040062.1|	GM15999 [Drosophila sechellia] >gb|EDW569...	306	3e-81	G

	●	☑	gb|ACE75299.1|	conserved hypothetical protein [Glyptapanteles...	306	4e-81	
	●	☑	ref|XP_001605051.1|	PREDICTED: similar to GA15301-PA [Nasonia...	296	2e-78	UG
	●	☑	ref|XP_967785.2|	PREDICTED: similar to conserved hypothetical...	275	8e-72	G
	●	☑	ref|XP_001599112.1|	PREDICTED: similar to CG31664-PA [Nasonia...	241	1e-61	G
	●	☑	ref|XP_001605439.1|	PREDICTED: similar to ENSANGP00000010019 ...	229	4e-58	G
	●	☑	gb|ACZ28354.1|	hypothetical protein [Simulium nigrimanum]	228	1e-57	
	●	☑	gb|ACH56924.1|	hypothetical protein [Simulium vittatum]	224	2e-56	
	●	☑	gb|ACZ28153.1|	hypothetical protein [Simulium nigrimanum]	215	7e-54	
	●	☑	gb|ACH56909.1|	hypothetical protein [Simulium vittatum]	214	1e-53	
	●	☑	gb|ACH56925.1|	hypothetical protein [Simulium vittatum]	209	3e-52	
	●	☑	gb|ACZ28152.1|	hypothetical conserved protein [Simulium nigri...	204	1e-50	
	●	☑	gb|AAU06548.1|	unknown salivary protein [Culicoides sonorensis]	202	5e-50	
	●	☑	ref|XP_002082764.1|	GD11752 [Drosophila simulans] >gb|EDX0834...	200	2e-49	G
	●	☑	gb|ACZ28340.1|	hypothetical conserved protein [Simulium nigri...	196	4e-48	
	●	☑	gb|AAU06547.1|	unknown salivary protein [Culicoides sonorensis]	195	6e-48	
	●	☑	gb|EFA12874.1|	hypothetical protein TcasGA2_TC010726 [Triboli...	193	3e-47	
	●	☑	ref|XP_002041740.1|	GM16585 [Drosophila sechellia] >gb|EDW455...	187	3e-45	G
	●	☑	gb|AAU06499.1|	unknown salivary protein [Culicoides sonorensis]	173	3e-41	
	●	☑	gb|AAU06503.1|	unknown salivary protein [Culicoides sonorensis]	171	1e-40	
	●	☑	gb|AAS87304.1|	CG2206-like protein [Drosophila miranda]	163	3e-38	
	●	☑	gb|ACZ28339.1|	hypothetical protein [Simulium nigrimanum]	145	5e-33	
	●	☑	gb|EFA11732.1|	hypothetical protein TcasGA2_TC002380 [Triboli...	125	1e-26	
	●	☑	gb|AAR20336.1|	CG2206-like protein [Drosophila pseudoobscura]...	124	2e-26	
	●	☑	gb|AAU06508.1|	unknown salivary protein [Culicoides sonorensis]	119	7e-25	
	●	☑	gb|AAU06510.1|	unknown salivary protein [Culicoides sonorensis]	107	2e-21	
	●	☑	gb|AAU06490.1|	unknown salivary protein [Culicoides sonorensis]	107	3e-21	
	●	☑	ref|NP_611782.2|	CG30409 [Drosophila melanogaster] >gb|AAF469...	70.5	4e-10	UG
	●	☑	ref|XP_001976383.1|	GG22842 [Drosophila erecta] >gb|EDV56783....	67.4	3e-09	G
NEW	●	☑	gb|ACZ28349.1|	hypothetical conserved protein [Simulium nigri...	65.5	1e-08	
NEW	●	☑	gb|ACZ28230.1|	hypothetical secreted protein [Simulium nigrim...	65.5	1e-08	

Run PSI-Blast iteration 5 with max 10000

Go

Sequences with E-value WORSE than threshold

☐	gb|ACZ28321.1|	hypothetical protein [Simulium nigrimanum]	52.0	1e-04	
☐	ref|XP_001430683.1|	hypothetical protein [Paramecium tetraure...	39.7	0.64	G
☐	gb|ACO11777.1|	Serine protease 33 precursor [Lepeophtheirus s...	39.3	0.89	
☐	ref|ZP_04196096.1|	hypothetical protein bcere0026_8120 [Bacil...	39.3	0.91	
☐	ref|XP_001944083.1|	PREDICTED: similar to AGAP003506-PA [Acyr...	39.3	0.92	UG
☐	emb|CAI43276.1|	Isp3b protein [Bacillus thuringiensis]	38.9	1.2	

☐	sp|P56765.2|ACCD_ARATH	RecName: Full=Acetyl-coenzyme A carbox...	38.9	1.2	
☐	ref|NP_051068.1|	acetyl-CoA carboxylase beta subunit [Arabido...	38.5	1.3	UG
☐	ref|YP_001123559.1|	carboxytransferase beta subunit [Draba ne...	38.5	1.4	G
☐	ref|YP_001123382.1|	acetyl-CoA carboxylase beta subunit [Caps...	38.5	1.5	G
☐	ref|YP_001271557.1|	1-deoxy-D-xylulose-5-phosphate synthase [...	38.5	1.5	G
☐	ref|XP_001606911.1|	PREDICTED: hypothetical protein [Nasonia ...	38.5	1.5	UG
☐	ref|YP_001843821.1|	1-deoxy-D-xylulose-5-phosphate synthase [...	38.5	1.6	G
☐	ref|NP_001035549.1|	shroom family member 4 [Mus musculus] >sp...	38.1	1.6	UG
☐	dbj|BAD32407.1|	mKIAA1202 protein [Mus musculus]	38.1	1.9	G
☐	ref|YP_001123471.1|	acetyl-CoA carboxylase beta subunit [Cruc...	38.1	1.9	G
☐	emb|CAM20542.1|	shroom family member 4 [Mus musculus]	38.1	2.0	G
☐	gb|EDL33881.1|	mCG49121 [Mus musculus]	37.7	2.1	
☐	gb|AAY88247.1|	Vip3D [Bacillus thuringiensis]	37.7	2.4	
☐	ref|YP_001123124.1|	acetyl-CoA carboxylase beta subunit [Olim...	37.7	2.5	G
☐	gb|ABH10614.1|	VipS [Bacillus thuringiensis]	37.7	2.5	
☐	ref|ZP_01046841.1|	TonB-dependent receptor [Nitrobacter sp. N...	37.4	2.9	
☐	ref|XP_001426273.1|	hypothetical protein [Paramecium tetraure...	37.4	3.0	G
☐	gb|AAY41427.1|	vegetative insecticidal protein [Bacillus thur...	37.4	3.0	
☐	gb|AAH93584.1|	LOC733205 protein [Xenopus laevis]	37.4	3.4	G
☐	ref|NP_001074557.1|	low density lipoprotein receptor-related ...	37.4	3.6	UG
☐	ref|YP_001123646.1|	acetyl-CoA carboxylase beta subunit [Lepi...	37.0	4.2	G
☐	gb|AAL69542.1|	Vip3A [Bacillus thuringiensis] >gb|AAO32350.1|...	37.0	4.4	
☐	ref|YP_001123208.1|	carboxytransferase beta subunit [Arabis h...	37.0	4.5	G
☐	ref|XP_606509.3|	PREDICTED: similar to Proprotein convertase ...	37.0	4.6	UG
☐	gb|EDL27035.1|	mCG129621 [Mus musculus]	37.0	4.7	
☐	ref|YP_001123823.1|	acetyl-CoA carboxylase beta subunit [Nast...	37.0	4.7	G
☐	ref|XP_001104098.1|	PREDICTED: low density lipoprotein-relate...	36.6	4.7	UG
☐	gb|AAX88965.1|	unknown [Homo sapiens]	36.6	4.7	G
☐	ref|NP_004516.2|	low density lipoprotein-related protein 2 pr...	36.6	5.0	UG
☐	gb|EAX11280.1|	low density lipoprotein-related protein 2 [Hom...	36.6	5.0	G
☐	gb|AAP88586.1|	glycoprotein receptor gp330/megalin precursor ...	36.6	5.0	G
☐	gb|AAP88585.1|	glycoprotein receptor gp330/megalin precursor ...	36.6	5.0	G
☐	gb|AAB41649.1|	gp330 precursor [Homo sapiens]	36.6	5.0	G
☐	ref|XP_001104179.1|	PREDICTED: low density lipoprotein-relate...	36.6	5.1	UG
☐	ref|XP_001023022.1|	Giardia variant-specific surface protein ...	36.6	5.4	G
☐	ref|XP_786902.2|	PREDICTED: similar to G protein-coupled rece...	36.6	5.5	G
☐	ref|YP_001123295.1|	carboxytransferase beta subunit [Barbarea...	36.6	6.0	G
☐	ref|XP_001156482.1|	PREDICTED: low density lipoprotein-relate...	36.2	6.1	UG
☐	ref|XP_001459453.1|	hypothetical protein [Paramecium tetraure...	36.2	6.4	G
☐	ref|XP_515882.2|	PREDICTED: low density lipoprotein-related p...	36.2	6.8	UG
☐	ref|XP_001580470.1|	hypothetical protein [Trichomonas vaginal...	36.2	7.0	G

☐ gb|ABI18031.1| acetyl co-A carboxylase [Ongokea gore] 35.8 8.5

☐ ref|XP_001426265.1| hypothetical protein [Paramecium tetraure... 35.8 9.5 G

☐ ref|XP_397194.3| PREDICTED: similar to distracted CG5634-PA [... 35.8 9.6 UG

Run PSI-Blast iteration 5 with max 10000
Go

Alignments ☐ Select All Get selected sequences Distance tree of results Multiple alignment NEW

>☐ ref|XP_002101210.1| G GE17494 [Drosophila yakuba]
gb|EDX02318.1| G GE17494 [Drosophila yakuba]
Length=844

GENE ID: 6525374 Dyak\GE17494 | GE17494 gene product from transcript GE17494-RA
[Drosophila yakuba] (10 or fewer PubMed links)

```
 Score =  524 bits (1349),  Expect = 9e-147, Method: Composition-based stats.
 Identities = 111/405 (27%), Positives = 180/405 (44%), Gaps = 45/405 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            L   R +   DP +H++G TY+E+TR  Q Y    E +     +C   C  Y + R
Sbjct  455  LRADRIVYRCDPPKHVQGVTYDEVTRLLQGYIENEVDLNNEETCRETCGFYQSTRSEGCF  514

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
             + Y   Q +CTG +YNC          V       P    S +RY ++E
Sbjct  515  KDLYCSRQPKCTGRLYNCQ--------FVDSDMWVCPSPQNSTRRYEFIE----------  556

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                     Y    V G +      T   D+W R     C  C CLCDE +  +S R F+
Sbjct  557  ---------YENGRVLGQRGKCTRGTTKVDSWWRYLLWHCSYCFCLCDE-EGLKSDRFFN  606

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS----  237
            L    +D K N++V G++   +++IF+ +I++  LLP G  +Q  L WK   +
Sbjct  607  LRDTVADVKRNRIVTGLRFVKQNRIFHLQIQEGELLPRGTVNQSTLEWKPVEKYNIFDRH  666

Query  238  -SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
                 D  KL   KR+I +    +  + FV+TGV+F  +  +L L    ++FD  +GKL+
Sbjct  667  VKNGIDYHKLSYEKRTIDLDDVDTEDNSFVITGVRFRVVGTHLNLEAYYSEFDFRTGKLI  726

Query  297  ND--NNQWIYNSLSK-----QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +      ++  +   A+   +   +          ++F  + L  D
Sbjct  727  QPEANSYWKSNDNTDVSGARREKLSLNNADVSTRTVAHSIPLSRHNQYIDFTNTGLDKDA  786

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPFID++ +VS     L G+G+ ++ R+   GGF+ P ++T+
Sbjct  787  AQSTVPFIDIQDVVSNPPVPLAGIGIYYKGRNGYGGFLGPKIITY  831
```

>☐ ref|XP_001978593.1| G GG17594 [Drosophila erecta]
gb|EDV47520.1| G GG17594 [Drosophila erecta]
Length=844

GENE ID: 6551613 Dere\GG17594 | GG17594 gene product from transcript GG17594-RA
[Drosophila erecta] (10 or fewer PubMed links)

```
 Score =  522 bits (1344),  Expect = 3e-146, Method: Composition-based stats.
 Identities = 111/405 (27%), Positives = 176/405 (43%), Gaps = 45/405 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            L   R +   DP +H++G TY+E+TR  Q Y    E +     +C   C  Y + R
Sbjct  455  LRADRIVYRCDPTKHVQGVTYDEVTRLLQGYIENEVDLNNEETCRETCSFYQSTRSEGCF  514
```

```
Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
               Y   Q +C+G +YNC          V       P    S +RY Y+E
Sbjct  515  KELYCSRQPKCSGRLYNCQ--------FVDSDMWVCPSPQNSTRRYEYIE----------  556

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                     Y    V G +      T   D+W R     C  C CLCDE +  +S R F+
Sbjct  557  ---------YENGRVLGKRGKCTRGTTKVDSWWRYLLWHCSYCFCLCDE-EGLKSDRFFN  606

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL-----  236
            L    +D K N+VV G++   +++IF+ +I++  LLP G  +Q  L WK  +
Sbjct  607  LRDTIADVKRNRVVTGLRFVKQNRIFHLQIQEGELLPRGTVNQSTLEWKPVEKYNVFDRH  666

Query  237  SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
                 D  KL    R+I +    +  + FVVTGV+F  I  +L L    ++FD  +G+L+
Sbjct  667  VKNGVDYHKLSYENRTIDLDDVDTEDNSFVVTGVRFRVIGTHLNLEAYYSEFDFRTGQLI  726

Query  297  ND--NNQWIYNSLSKQKLN-----NGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +          N   A+   +              ++F  + +  D
Sbjct  727  QPEGNSYWKSNDNTDVSGTRRQKLNLNNADVSTRTIAPSIPLSRHNQYIDFTNTGMDKDA  786

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPFID++ ++S     L G+G+ ++ R++ GGF+ P ++T+
Sbjct  787  AQSTVPFIDIQDVISNPPVPLAGIGIYYKGRNSYGGFLGPKIITY  831
```

>☐ ref|XP_001963941.1| GF20993 [Drosophila ananassae]
 gb|EDV34390.1| GF20993 [Drosophila ananassae]
Length=851

 GENE ID: 6503682 Dana\GF20993 | GF20993 gene product from transcript GF20993-RA
[Drosophila ananassae] (10 or fewer PubMed links)

```
 Score =  517 bits (1331),  Expect = 9e-145, Method: Composition-based stats.
 Identities = 111/406 (27%), Positives = 178/406 (43%), Gaps = 46/406 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            L   R +   DP +HI   TY+E+TR  Q Y   E +     +C   C+ Y + R
Sbjct  461  LRADRIVYRCDPPKHIPKVTYDEVTRLLQGYVENEVDLNNEETCRETCDFYQSTRAEGCF  520

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
              +Y   Q++C+G IYNC          +       P    S +RY ++E
Sbjct  521  KEKYCARQQRCSGRIYNCQ--------FIDSDMWVCPSPQNSTRRYEFIE----------  562

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                     Y    V G +      T   D+W R     C  C CLCDE +  +S R F+
Sbjct  563  ---------YESGRVLGQRTKCHRGTTKVDSWWRYLLWHCSYCFCLCDE-QGLKSDRFFN  612

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS----  237
            L    +D K N+VV G++L   ++IF+ +I++  L+P G  +Q  L W+  +
Sbjct  613  LRDSIADVKKNRVVTGLRLVKSNRIFHLQIQEGELMPRGVINQSTLEWRPVEKYNIFDRD  672

Query  238  -SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
                 D   L    RSI +    ++ + FVVTGV+F  +  +L L    T+FD  +GKL+
Sbjct  673  VKKGVDYHTLSYESRSIDLDDVNTDDNSFVVTGVRFRLVGAHLNLEAYFTEFDFKTGKLI  732

Query  297  ND--NNQWIYNSLSK------QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDD  348
                N+ W  N  +       ++       + P +       +      ++F  + L  D
Sbjct  733  QPEINSYWKSNDNTDVGGGARREKLRLANTDLPTRTIATSLVQSRHNQYIDFTNTGLDKD  792

Query  349  VGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
              QSTVPFID++ +VS     L G+G+ ++ R+  GGF+ P ++T+
Sbjct  793  AAQSTVPFIDIQDVVSNPPVPLSGIGIYYKGRNGYGGFVGPKIITY  838
```

>☐ ref|XP_002067370.1| GK16218 [Drosophila willistoni]
 gb|EDW78356.1| GK16218 [Drosophila willistoni]
Length=711

 GENE ID: 6644823 Dwil\GK16218 | GK16218 gene product from transcript GK16218-RA
[Drosophila willistoni] (10 or fewer PubMed links)

```
 Score =  516 bits (1330),  Expect = 1e-144, Method: Composition-based stats.
 Identities = 109/405 (26%), Positives = 175/405 (43%), Gaps = 46/405 (11%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYAAPRHNFKED  63
               R +   DP +H    TY+EITR  Q Y   E +  +   C   C  Y A R
Sbjct  321  RADRIVWRCDPPKHKAAVTYDEITRLLQGYIENEVDLNTDETCRETCSFYQATRSEGCFK  380

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
              +   Q +C+G +++C          V       P +  S +RY YVE
Sbjct  381  ELFCARQPKCSGRLFDCQ--------FVDSDMWVCPSAQNSTRRYEYVE-----------  421

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSL  182
                    Y    V G +      T   D+W R  F  C  C C+CDE +  +S R F+L
Sbjct  422  --------YENGRVLGQRGQCVRGTTKVDSWWRYLFWHCSYCFCMCDE-QGLKSDRFFNL  472

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS-----  237
                SD K N+VV G++    ++IF+ +I+++ LLP G  ++  L W+  D
Sbjct  473  RESLSDVKRNRVVTGLRFVKLNRIFHLQIQEAELLPRGAVNESTLEWRKIDSYNIFDRDV  532

Query  238  SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMN  297
                D   L    RSI +    ++ + +VVTGV+F  +  +L L  +L++ D  +G+L+
Sbjct  533  KKGVDYHTLSYESRSIDLDDIYTDDNSYVVTGVRFRVVGAHLNLEARLSELDFRTGQLIQ  592

Query  298  D--NNQWIYNSLSKQKLNN------GGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
               N+ W  N  +    ++         A+ P +              ++F  + L  D
Sbjct  593  PEVNSFWKSNDNTDVSGDDRRKKIPLSNADLPTRTITTSIPISRHNQYIDFTYTGLDKDA  652

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPFID++ + S     L G+G+ ++ R   GGF+AP ++T+
Sbjct  653  AQSTVPFIDIQDVTSKPPVPLSGIGIYYKGRQGYGGFLAPKIITY  697


>☐ ref|XP_002057094.1| G GJ16894 [Drosophila virilis]
 gb|EDW62580.1| G GJ16894 [Drosophila virilis]
Length=824

 GENE ID: 6633805 Dvir\GJ16894 | GJ16894 gene product from transcript GJ16894-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  514 bits (1325),  Expect = 5e-144, Method: Composition-based stats.
 Identities = 116/405 (28%), Positives = 177/405 (43%), Gaps = 46/405 (11%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYAAPRHNFKED  63
               R +   DP +H+   TY+E+TR  Q Y   E +  S   C   C  Y + R
Sbjct  435  RADRIVWRCDPDKHVPQVTYDEVTRLLQGYIENEVDLNSDETCRETCAFYQSTRSEGCFK  494

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
            N Y   Q +CTG +Y C          V       P    S +RY Y+E
Sbjct  495  NLYCSRQPKCTGKLYGCQ--------FVDSDMWVCPAPLSSTRRYEYIE-----------  535

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFSL  182
                    Y    V G +++    T   D+W R     C  C CLCDE +  +S R F+L
Sbjct  536  --------YENGRVLGQRKNCVRGTTKVDSWWRYLLWHCSYCFCLCDE-QGLKSDRFFNL  586

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LS  237
                SD K NKV+ G++   K++IF+ +I++  LLP G  ++  LSWK  D
Sbjct  587  RESVSDVKQNKVITGLRFLKKNRIFHLQIQEGQLLPRGAINESTLSWKPVDEYKIFDRNV  646

Query  238  SASKDIFKLDVNKRSIHIGSAT-SNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
            +   D   L    RSI +     ++ + FVVTGV+F  +  +L L  +L++FD   G+L+
Sbjct  647  AKGIDYHTLSYESRSIDLDDIMKTDDNSFVVTGVRFRVLGAHLNLEARLSEFDFKKGELI  706

Query  297  ND--NNQWIYNSLSKQKLNNGG-----VANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +              A+ P +              ++F  S +  D
Sbjct  707  QPEVNSVWQSNDNTDVSGERRQKLLLYNADVPTRTVVTSLPLSKHNQYIDFVNSGMDKDA  766

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPF+D++ +VS     L G+G+ ++ R   GGF+AP ++T+
Sbjct  767  AQSTVPFLDIQDVVSQPPVPLAGIGIYYKGRPGYGGFLAPKIITY  811
```

```
>☐ ref|XP_002025529.1| G GL15246 [Drosophila persimilis]
 gb|EDW31051.1| G GL15246 [Drosophila persimilis]
Length=806

 GENE ID: 6600424 Dper\GL15246 | GL15246 gene product from transcript GL15246-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score =  513 bits (1321),  Expect = 2e-143, Method: Composition-based stats.
 Identities = 110/407 (27%), Positives = 178/407 (43%), Gaps = 48/407 (11%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKED  63
               R +   DPK+H  G TY+E+TR  Q Y   E +     +C   C  Y + R
Sbjct  415  RADRIVWRCDPKKHEAGVTYDEVTRLLQGYIENEVDLNNEETCRETCSFYQSTRSEGCYK  474

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
            + Y   Q +C+G +YNC          V       P    S +RY Y+E
Sbjct  475  DLYCARQPRCSGRLYNCQ--------FVDSDMWVCPSPKNSTRRYEYIE-----------  515

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSL  182
                    Y      G + +    T   D+W R  F  C  C C+CDE +  +S R F+L
Sbjct  516  --------YENGRTLGQRANCVRGTTKVDSWWRYLFWHCSYCFCMCDE-QGLKSDRFFNL  566

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS-----  237
                SD K N+VV G++   K++IF+ +I++  LLP G  +   LSWK  +
Sbjct  567  REAVSDVKRNRVVTGLRFVKKNRIFHLQIQEGELLPRGGINHTTLSWKPVENYNIFDRDV  626

Query  238  SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMN  297
                D   L    RS+ +     ++ + FVVTGV+F  +  +L L    ++FD  +G+L+
Sbjct  627  KKGTDYHTLSYESRSVDLDDVNTDDNSFVVTGVRFRVVGAHLNLEACYSEFDFKTGQLLQ  686

Query  298  D--NNQWIYNSLSK--------QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKD  347
               N+ W  N  +         ++  +   A+   +       +      +EF  + L
Sbjct  687  PEANSYWKSNDNTDVSGDSKTRREKVSLSNADVSTRTITYSIPQSRHNQYIEFTNTGLDK  746

Query  348  DVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            D  QSTVPF+D++ ++S     L G+G+ ++ R+  GGF+ P ++T+
Sbjct  747  DAAQSTVPFMDVQDVISNPPVPLSGIGIYYKGRNGYGGFLGPKIITY  793


>☐ ref|XP_001355344.2| G GA15301 [Drosophila pseudoobscura pseudoobscura]
 gb|EAL32401.2|  GA15301 [Drosophila pseudoobscura pseudoobscura]
Length=806

 GENE ID: 4816053 Dpse\GA15301 | GA15301 gene product from transcript GA15301-RA
[Drosophila pseudoobscura pseudoobscura] (10 or fewer PubMed links)

 Score =  512 bits (1319),  Expect = 2e-143, Method: Composition-based stats.
 Identities = 110/407 (27%), Positives = 178/407 (43%), Gaps = 48/407 (11%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKED  63
               R +   DPK+H  G TY+E+TR  Q Y   E +     +C   C  Y + R
Sbjct  415  RADRIVWRCDPKKHEAGVTYDEVTRLLQGYIENEVDLNNEETCRETCSFYQSTRSEGCYK  474

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
            + Y   Q +C+G +YNC          V       P    S +RY Y+E
Sbjct  475  DLYCARQPRCSGRLYNCQ--------FVDSDMWVCPSPKNSTRRYEYIE-----------  515

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSL  182
                    Y      G + +    T   D+W R  F  C  C C+CDE +  +S R F+L
Sbjct  516  --------YENGRTLGQRANCVRGTTKVDSWWRYLFWHCSYCFCMCDE-QGLKSDRFFNL  566

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS-----  237
                SD K N+VV G++   K++IF+ +I++  LLP G  +   LSWK  +
Sbjct  567  REAVSDVKRNRVVTGLRFVKKNRIFHLQIQEGELLPRGGINHTTLSWKPVENYNIFDRDV  626

Query  238  SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMN  297
                D   L    RS+ +     ++ + FVVTGV+F  +  +L L    ++FD  +G+L+
Sbjct  627  KKGTDYHTLSYESRSVDLDDVNTDDNSFVVTGVRFRVVGAHLNLEACYSEFDFKTGQLLQ  686

Query  298  D--NNQWIYNSLSK--------QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKD  347
               N+ W  N  +         ++  +   A+   +       +      +EF  + L
```

```
Sbjct  687  PEANSYWKSNDNTDVSGDSKTRREKVSLSNADVSTRTITYSIPQSRHNQYIEFTNTGLDK  746

Query  348  DVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            D  QSTVPF+D++ ++S     L G+G+ ++ R+  GGF+ P ++T+
Sbjct  747  DAAQSTVPFMDVQDVISNPPVPLSGIGIYYKGRNGYGGFLGPKIITY  793


>☐ ref|XP_002076900.1| G GD24578 [Drosophila simulans]
 gb|EDX16494.1| G GD24578 [Drosophila simulans]
Length=646

 GENE ID: 6740015 Dsim\GD24578 | GD24578 gene product from transcript GD24578-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score =  511 bits (1317),  Expect = 4e-143, Method: Composition-based stats.
 Identities = 110/405 (27%), Positives = 179/405 (44%), Gaps = 45/405 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            L   R +   DP +H++G TY+E+TR  Q Y   E +     +C   C  Y + R
Sbjct  257  LRADRIVYRCDPTKHVQGVTYDEVTRLLQGYIENEVDLNNEETCRETCGFYQSTRSEGCF  316

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
             + Y   Q +CTG +YNC          V       P    S +RY ++E
Sbjct  317  KDLYCSRQPKCTGRLYNCQ--------FVDSDMWVCPSPQNSTRRYEFIE----------  358

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                     Y    V G +      T   D+W R     C  C CLCDE +  +S R F+
Sbjct  359  ---------YENGRVLGQRGKCTRGTTKVDSWWRYLLWHCSYCFCLCDE-EGLKSDRFFN  408

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL-----  236
            L    +D K N++V G++   +++IF+ +I++  LLP G  +Q  L WK  +
Sbjct  409  LRDTIADIKRNRIVTGLRFVKQNRIFHLQIQEGELLPRGTVNQSTLEWKPVEKYNVFDRH  468

Query  237  SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
                 D  KL   KR+I +    +  + FVVTGV+F  +  +L L    ++FD ++G+L+
Sbjct  469  VKNGVDYHKLSYEKRTIDLDDVDTEDNSFVVTGVRFRVVGTHLNLEAYYSEFDFLTGQLI  528

Query  297  ND--NNQWIYNSLSK-----QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +      ++      A+   +              ++F  + L  D
Sbjct  529  QPEYNSYWKSNDNTDVSGARREKLRLSNADVSTRTIAPSLPLSRHNQFIDFTNTGLDKDA  588

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPF+D++ +VS     L G+G+ ++ R+  GGF+ P ++T+
Sbjct  589  AQSTVPFMDIQDVVSNPPVPLAGIGIYYKGRNGYGGFLGPKIITY  633


>☐ ref|NP_572427.1| UG lethal (1) G0193, isoform A [Drosophila melanogaster]
 gb|AAF46301.1| G lethal (1) G0193, isoform A [Drosophila melanogaster]
 gb|AAL13770.1| G LD24308p [Drosophila melanogaster]
 gb|ACL84971.1|  l(1)G0193-PA [synthetic construct]
Length=664

 GENE ID: 31712 l(1)G0193 | lethal (1) G0193 [Drosophila melanogaster]
(Over 10 PubMed links)

 Score =  508 bits (1308),  Expect = 5e-142, Method: Composition-based stats.
 Identities = 110/405 (27%), Positives = 181/405 (44%), Gaps = 45/405 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            L   R +   DP +H++G TY+E+TR  Q Y   E +     +C   C+ Y + R
Sbjct  275  LRADRIVYRCDPTKHVRGVTYDEVTRLLQGYIENEVDLNKEETCRETCDFYQSTRSEGCF  334

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
             + Y   Q +C+G +YNC          V       P    S +RY ++E
Sbjct  335  KDLYCSRQPKCSGRLYNCR--------FVDSDMWVCPSPQNSTRRYEFIE----------  376

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                     Y    V G +      T   D+W R     C  C CLCDE +  +S R F+
Sbjct  377  ---------YENGRVLGQRGKCTRGTTKVDSWWRYLLWHCSYCFCLCDE-EGLKSDRFFN  426
```

```
Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL-----  236
            L    +D K N+VV G++   +++IF+ +I++  LLP G  +Q  L WK  +
Sbjct  427  LRDTIADVKRNRVVTGLRFVKQNRIFHLQIQEGELLPRGIVNQSTLEWKPVEKYNVFDRH  486

Query  237  SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
                 D  KL   KR+I +    ++ + FVVTGV+F  +  +L L    ++FD  +G+L+
Sbjct  487  VKNGVDYHKLSYEKRTIDLDDVDTDDNSFVVTGVRFRVVGTHLNLEAYYSEFDFRTGQLI  546

Query  297  ND--NNQWIYNSLSK-----QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +      ++      A+   +   +          ++F  + L  D
Sbjct  547  RPEYNSYWKSNDNTDVSGARREKLRLSNADVSTRTIAHSIPLSRHNQYIDFTNTGLDKDA  606

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPF+D++ +VS      L G+G+ ++ R+  GGF+ P ++T+
Sbjct  607  AQSTVPFMDIQDVVSNPPVPLAGIGIYYKGRNGYGGFLGPKIITY  651


>☐ ref|XP_002010644.1| G GI21656 [Drosophila mojavensis]
 gb|EDW06299.1| G GI21656 [Drosophila mojavensis]
Length=727

 GENE ID: 6585004 Dmoj\GI21656 | GI21656 gene product from transcript GI21656-RA
[Drosophila mojavensis] (10 or fewer PubMed links)

 Score =  506 bits (1303),  Expect = 2e-141, Method: Composition-based stats.
 Identities = 113/405 (27%), Positives = 176/405 (43%), Gaps = 46/405 (11%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKED  63
               R +   DP+  +  +TY+E+TR  Q Y   E +     +C   C  Y + R
Sbjct  338  RSDRIVWRCDPEHFVLNKTYDEVTRLLQGYIENEVDLNADETCRETCSFYQSTRSEGCFK  397

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
              +   Q +CTG +YNC          V       P +  S +RY Y+E
Sbjct  398  EYFCARQPKCTGRLYNCQ--------FVDSDMWVCPAARNSTRRYEYIE-----------  438

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSL  182
                    Y    V G ++     T   D+W R  F  C  C CLCDE +  +S R F+L
Sbjct  439  --------YENGRVLGRRQRCVRGTTKVDSWWRYLFWHCSYCFCLCDE-QGLKSDRFFNL  489

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LS  237
                +D K N VV G++   ++++F+ +I++  LLP G  ++  LSWK  D
Sbjct  490  RESVADVKKNNVVTGLRFVKRNRVFHLQIQEGQLLPRGAINESTLSWKPVDEYKIYDRNV  549

Query  238  SASKDIFKLDVNKRSIHIGSAT-SNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
            +   D   L    RSI +     ++   FVVTG++F  +  +L L  +LT+FD   G+L+
Sbjct  550  AKGIDYHTLSYESRSIDLDDIMKTDDYSFVVTGLRFRVLGAHLNLEARLTEFDFTKGELI  609

Query  297  ND--NNQWIYNSLSKQKLNNGG-----VANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +              A+ P +              +EF  S +  D
Sbjct  610  QPEVNSVWQSNDNTDVSGERRQKINIYNADVPTRSVVTSIPMSKHNQYIEFVNSGMDKDA  669

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPF+D++ +VS      L G+GL ++ R   GGF+AP ++T+
Sbjct  670  AQSTVPFLDIQDVVSQPPVPLAGIGLYYKGRPGYGGFLAPKIITY  714


>☐ ref|NP_996366.1| UG lethal (1) G0193, isoform B [Drosophila melanogaster]
 gb|AAS65275.1| G lethal (1) G0193, isoform B [Drosophila melanogaster]
Length=646

 GENE ID: 31712 l(1)G0193 | lethal (1) G0193 [Drosophila melanogaster]
(Over 10 PubMed links)

 Score =  501 bits (1291),  Expect = 4e-140, Method: Composition-based stats.
 Identities = 110/405 (27%), Positives = 181/405 (44%), Gaps = 45/405 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            L    R +   DP +H++G TY+E+TR  Q Y   E +     +C   C+ Y + R
Sbjct  257  LRADRIVYRCDPTKHVRGVTYDEVTRLLQGYIENEVDLNKEETCRETCDFYQSTRSEGCF  316

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
```

```
             + Y   Q +C+G +YNC          V       P    S +RY ++E
Sbjct  317  KDLYCSRQPKCSGRLYNCR--------FVDSDMWVCPSPQNSTRRYEFIE----------  358

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                     Y    V G +      T   D+W R     C  C CLCDE +  +S R F+
Sbjct  359  ---------YENGRVLGQRGKCTRGTTKVDSWWRYLLWHCSYCFCLCDE-EGLKSDRFFN  408

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL-----  236
            L    +D K N+VV G++   +++IF+ +I++  LLP G  +Q  L WK  +
Sbjct  409  LRDTIADVKRNRVVTGLRFVKQNRIFHLQIQEGELLPRGIVNQSTLEWKPVEKYNVFDRH  468

Query  237  SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
                 D  KL   KR+I +    ++ + FVVTGV+F  +  +L L    ++FD  +G+L+
Sbjct  469  VKNGVDYHKLSYEKRTIDLDDVDTDDNSFVVTGVRFRVVGTHLNLEAYYSEFDFRTGQLI  528

Query  297  ND--NNQWIYNSLSK-----QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +      ++      A+   +   +          ++F  + L  D
Sbjct  529  RPEYNSYWKSNDNTDVSGARREKLRLSNADVSTRTIAHSIPLSRHNQYIDFTNTGLDKDA  588

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             QSTVPF+D++ +VS     L G+G+ ++ R+  GGF+ P ++T+
Sbjct  589  AQSTVPFMDIQDVVSNPPVPLAGIGIYYKGRNGYGGFLGPKIITY  633


>☐ ref|XP_001992028.1| G GH24541 [Drosophila grimshawi]
 gb|EDV91735.1| G GH24541 [Drosophila grimshawi]
Length=827

 GENE ID: 6566309 Dgri\GH24541 | GH24541 gene product from transcript GH24541-RA
[Drosophila grimshawi] (10 or fewer PubMed links)

 Score =  492 bits (1267),  Expect = 3e-137, Method: Composition-based stats.
 Identities = 114/411 (27%), Positives = 181/411 (44%), Gaps = 48/411 (11%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYAAPRHNFKED  63
               R +   DP++HIK ETY+E+TR  Q Y   E +  +   C   C+ Y + R +
Sbjct  438  RSDRIVYRCDPQKHIKNETYDEVTRLLQGYIVNEVDLNTDDTCRETCDFYQSTRSDSCFK  497

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
              Y   Q +CTG IYNC          V       P      +RY Y+E
Sbjct  498  ELYCARQPKCTGKIYNCQ--------FVDSDMWVCPSPLNGSRRYEYIE-----------  538

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSL  182
                    Y      G ++     T   D+W R  F  C  C CLCDE +  +S R F+L
Sbjct  539  --------YENGRTLGQRKTCSRGTTKVDSWWRYLFWHCSYCFCLCDE-QGIKSDRFFNL  589

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LS  237
                +D K NKVV G++   ++++F+ +I++  LLP G  +   L+WK  D
Sbjct  590  RESVADVKQNKVVTGLRFVKRNRVFHLQIQEGQLLPRGSINASTLAWKPVDDYKIFDRNV  649

Query  238  SASKDIFKLDVNKRSIHIGSAT-SNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
            +  +D   +    RSI +      + + FVVTGV+F  +  +L L  +L++FD   G+L+
Sbjct  650  AKGRDYHTMSYESRSIDLDDIMKMDDNSFVVTGVRFRVLGAHLNLEARLSEFDFKRGELV  709

Query  297  ND--NNQWIYNSLSKQKLNNGG-----VANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                N+ W  N  +               + P +   +          +EF  S +  D
Sbjct  710  QPEVNSVWQSNDNTDVSGERRQKLKIHNVDVPTRTVVSSLVASKHNEYIEFVNSGMDKDA  769

Query  350  GQSTVPFIDLRSIV-STGTALIGVGLQFRTRDTSGGFIAPVVVTH--PTYT  397
             QSTVP++D++ +V S    L G+G+ ++TR   GGF+AP ++T+    +
Sbjct  770  AQSTVPYLDIQDVVSSPPVPLAGIGIYYKTRPGFGGFVAPKIITYDFSPHV  820


>☐ ref|XP_001851110.1| UG conserved hypothetical protein [Culex quinquefasciatus]
 gb|EDS33063.1| G conserved hypothetical protein [Culex quinquefasciatus]
Length=642

 GENE ID: 6041920 CpipJ_CPIJ009483 | hypothetical protein
[Culex quinquefasciatus]

 Score =  490 bits (1262),  Expect = 9e-137, Method: Composition-based stats.
```

```
 Identities = 115/405 (28%), Positives = 182/405 (44%), Gaps = 49/405 (12%)

Query  8    FSRTLDNVDPK--EHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYA-APRHNFK  61
             SR     DP+  +H +GE++ +ITR  Q Y   E N     +C   C  Y+   +
Sbjct  259  ASREYWRCDPERSQHKEGESFVQITRLLQGYVENEVNLNRNDTCRKDCAYYSSDVKQEQC  318

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
            + + Y   Q  C+G IY+C        E +       P +  SD+RY Y+E
Sbjct  319  KKDMYCTKQLICSGKIYDC--------EFIDSAMWICPAANNSDRRYEYIE---------  361

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSF  180
                      Y    V+G K+     T    +W R  F  C  C CLCD++K   S R
Sbjct  362  ----------YENGMVFGQKKTCSRGTTKVKSWWRWIFWHCSNCFCLCDDAKH--SDRYI  409

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----  235
            ++    SD K+NKVV G++   K++I +  ++Q  LLP G+ D   L W
Sbjct  410  NMRESVSDLKNNKVVTGMRFVKKNRIIHLIVQQGRLLPRGQIDNGTLEWVEPHNYNLLST  469

Query  236  LSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
                 +D   L+ N RS+ +         +V+TGV+F  +D +L L +++T+ D  +G+L
Sbjct  470  NVRVGRDYHTLNRNNRSLDLDDLYV-PQGYVLTGVRFRLLDNHLNLEIRMTEIDFATGQL  528

Query  296  MN-DNNQWIYNSLSKQKLN-----NGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
            ++ D + WI N  +++ +N          + P+        +      ++FR SD    D
Sbjct  529  IDPDKSIWIGNDNTERTVNKRTEFELDRPHVPLLSLVKSLPDSASNQFIQFRASDRNMDA  588

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             Q+TVPF D +S+V      L G GL  + R   GGF+AP V+T+
Sbjct  589  AQTTVPFFDAQSVVPVKPVPLSGAGLFHKGRPGFGGFVAPKVMTY  633


>☐ ref|XP_969643.2| UG PREDICTED: similar to GA15301-PA [Tribolium castaneum]
Length=660

 GENE ID: 658141 LOC658141 | similar to CG2206-PB, isoform B
[Tribolium castaneum]

 Score =  484 bits (1245),  Expect = 9e-135, Method: Composition-based stats.
 Identities = 108/403 (26%), Positives = 171/403 (42%), Gaps = 48/403 (11%)

Query  9    SRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKEDNR  65
            SR L   DPK H+ GETY E+T+  Q Y   E +     +C   C +Y   + +    N
Sbjct  270  SRDLWKCDPKRHVSGETYVEVTQLLQGYIQNEVDLNPEGTCRENCAEYTYTKSHGCFQNL  329

Query  66   YKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYA  125
            +   QR+C G + NC          V       P      +RY YVE
Sbjct  330  WCRKQRRCHGKVINCK--------YVDSDMWVCPAD-NQSRRYEYVE-------------  367

Query  126  IIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSLNA  184
                  Y    V G K+     T   D+W R  F  C  C CLCDE +   S R  +L A
Sbjct  368  ------YENGRVLGRKQGCRRGTSKVDSWWRWLFWHCSYCFCLCDE-QGVHSDRYINLRA  420

Query  185  VESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LSSA  239
              +D K+N+VV G++    ++I + +I++  LLP G  D   + W   +         +
Sbjct  421  SIADIKNNRVVTGLRFVKHNRIIHLQIQEGKLLPRGNIDVTTVHWVPVEEYKITDSNVAN  480

Query  240  SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDN  299
            ++D   L   KR++ +    +    +VVTGV+F  I  +L   +  T FD  +G+L++
Sbjct  481  AQDYHTLTWEKRAVDLDDLVA-DEGYVVTGVRFKVIGSHLNFEIYTTPFDFETGQLIDPE  539

Query  300  NQWIY--NSLSK------QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQ  351
             + +Y  N  +       +        +   +       +      +EF  SD+  D  Q
Sbjct  540  TKSMYKDNPNTDSSLYKPRTRVRLTNPDISTRSPSPSLPDSKTDQYIEFTNSDMDRDAAQ  599

Query  352  STVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            +TVPF+D + + S     L G G+  + +   GGFIAP ++T+
Sbjct  600  TTVPFLDAQPVDSLQPVPLSGAGIFHKGQKYFGGFIAPKLITY  642


>☐ gb|EFA08554.1|  hypothetical protein TcasGA2_TC006209 [Tribolium castaneum]
Length=640
```

```
 Score =  478 bits (1231),  Expect = 4e-133, Method: Composition-based stats.
 Identities = 108/403 (26%), Positives = 171/403 (42%), Gaps = 48/403 (11%)

Query  9    SRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKEDNR  65
            SR L   DPK H+ GETY E+T+  Q Y   E +     +C   C +Y   + +    N
Sbjct  250  SRDLWKCDPKRHVSGETYVEVTQLLQGYIQNEVDLNPEGTCRENCAEYTYTKSHGCFQNL  309

Query  66   YKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYA  125
            +   QR+C G + NC          V       P      +RY YVE
Sbjct  310  WCRKQRRCHGKVINCK--------YVDSDMWVCPAD-NQSRRYEYVE-------------  347

Query  126  IIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSLNA  184
                  Y    V G K+     T   D+W R  F  C  C CLCDE +   S R  +L A
Sbjct  348  ------YENGRVLGRKQGCRRGTSKVDSWWRWLFWHCSYCFCLCDE-QGVHSDRYINLRA  400

Query  185  VESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LSSA  239
              +D K+N+VV G++    ++I + +I++  LLP G  D   + W   +         +
Sbjct  401  SIADIKNNRVVTGLRFVKHNRIIHLQIQEGKLLPRGNIDVTTVHWVPVEEYKITDSNVAN  460

Query  240  SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDN  299
            ++D   L   KR++ +    +    +VVTGV+F  I  +L   +  T FD  +G+L++
Sbjct  461  AQDYHTLTWEKRAVDLDDLVA-DEGYVVTGVRFKVIGSHLNFEIYTTPFDFETGQLIDPE  519

Query  300  NQWIY--NSLSK------QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQ  351
             + +Y  N  +       +        +   +       +      +EF  SD+  D  Q
Sbjct  520  TKSMYKDNPNTDSSLYKPRTRVRLTNPDISTRSPSPSLPDSKTDQYIEFTNSDMDRDAAQ  579

Query  352  STVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            +TVPF+D + + S     L G G+  + +   GGFIAP ++T+
Sbjct  580  TTVPFLDAQPVDSLQPVPLSGAGIFHKGQKYFGGFIAPKLITY  622


>☐ gb|AAY55307.1|  IP12501p [Drosophila melanogaster]
Length=579

 Score =  478 bits (1230),  Expect = 5e-133, Method: Composition-based stats.
 Identities = 113/409 (27%), Positives = 187/409 (45%), Gaps = 47/409 (11%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA  54
            LP +        R++   DP+EH  G T+EE+TR  Q Y   E +  S   CN +C  Y
Sbjct  188  LPKIKVLMSQADRSVWRCDPQEHQSGVTHEEVTRLLQGYVENEVDLNSDGACNQECGYYK  247

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
              ++    DN++   Q +CTG +++C          V            S +RY +++
Sbjct  248  LAKNEGCFDNKFCSEQPKCTGGVHDCR--------FVESNMKICQAEKNSSRRYEFIK--  297

Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR-CDICRCLCDESKD  173
                             Y    V+GN++       +  +W R     C  C CLCD+ +
Sbjct  298  -----------------YESGLVHGNEKPCATWLNSAKSWNRWLFMECSYCLCLCDD-QS  339

Query  174  SESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD  233
             +S R F+L  V +D K+N+VV+G++   +++IF+ ++++  LLPLG  +   L WK  D
Sbjct  340  PQSDRYFNLRPVIADVKNNRVVVGLRFVKRNRIFHLQVQEGELLPLGDINGTTLRWKPVD  399

Query  234  PSLSSA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKF  288
                S      + D   L   KRSI +G   + ++  VVTG++F     +L L  + T
Sbjct  400  SYSISDKDVSENVDYHTLTYEKRSIDLGDLNAKINS-VVTGLRFRVFGAHLQLEAQFTAV  458

Query  289  DVMSGKLMNDN--NQWIYNSLSKQK-LNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +  SGKL++    + WI     + +       A    K +        D   +EF  +
Sbjct  459  EFESGKLIDPKKTSHWISMGTDQPRERLPLMDAQVSTKSSVLSIPYPSDNQFIEFTNTGF  518

Query  346  KDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
              D  Q TVPFID++ +VS     L G+G+ ++ +D  GGF+AP ++T+
Sbjct  519  VQDAAQRTVPFIDIQDVVSDPPVPLSGIGIYYKGQDGYGGFLAPRIITY  567


>☐ ref|XP_001851111.1| conserved hypothetical protein [Culex quinquefasciatus]
 gb|EDS33064.1| conserved hypothetical protein [Culex quinquefasciatus]
Length=721
```

GENE ID: 6041924 CpipJ_CPIJ009484 | hypothetical protein
[Culex quinquefasciatus]

```
 Score =  475 bits (1222),  Expect = 4e-132, Method: Composition-based stats.
 Identities = 106/405 (26%), Positives = 176/405 (43%), Gaps = 48/405 (11%)

Query  8    FSRTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA-APRHNFK  61
             S      DP  ++H +GET+ ++TR  Q Y   E    +   C   C  Y+   +
Sbjct  322  ASPAYWRCDPVQRQHKEGETFVQLTRLLQGYVENEVALNTDNTCKETCGYYSWGVQQERC  381

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
              + Y   Q +C+G IY+C        + +  +    P +  +++RY YVE
Sbjct  382  YKDTYCSKQSKCSGKIYDC--------DFIDSHMWICPAAKNTNRRYEYVE---------  424

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSF  180
                      Y    V G+++     T   ++W R    RC  C CLCD++    S R
Sbjct  425  ----------YENGIVLGHRKQCASGTTKVESWWRWLIWRCSYCFCLCDDA-GKNSDRYI  473

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----  235
            ++    SD  +NKVV G++   K +I +  ++Q  LLP G+ D   L W
Sbjct  474  NMRESVSDVMNNKVVTGMRFIKKSQIIHLIVQQGRLLPRGQIDNGTLEWVQPHNYTVMSQ  533

Query  236  LSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
                 +D   L+ N RS+ +         +VVTGV+F  +  +L L +++T+ D  +G+L
Sbjct  534  NVRNGRDYHTLNRNNRSLDLDDLYV-PQGYVVTGVRFRLLGNHLNLEIRMTEKDFATGQL  592

Query  296  MND-NNQWIYNSLSKQKLN-----NGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
            ++   + WI N  +++ +N          N P+          +    + FR+SD + D
Sbjct  593  IDPEKSIWIGNVNTERSVNKRTEIKLDQPNVPLLSPLKSEPSFVSNRFIRFRESDRRTDA  652

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             Q+TVPF D + +       L G GL  + R   GGF+AP V+T+
Sbjct  653  AQTTVPFFDAQPVAPVKPVPLSGAGLFHKGRPKFGGFVAPKVMTY  697
```

>☐ ref|XP_002077791.1| G GD22890 [Drosophila simulans]
gb|EDX03376.1| G GD22890 [Drosophila simulans]
Length=490

GENE ID: 6730612 Dsim\GD22890 | GD22890 gene product from transcript GD22890-RA
[Drosophila simulans] (10 or fewer PubMed links)

```
 Score =  474 bits (1219),  Expect = 9e-132, Method: Composition-based stats.
 Identities = 114/409 (27%), Positives = 188/409 (45%), Gaps = 47/409 (11%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA  54
            LP +        R++   DP+EH  G TYEE+TR  Q Y   E +  S   CN  C  Y
Sbjct  91   LPKIKVLMSQAERSVWRCDPQEHQSGVTYEEVTRLLQGYVENEVDLNSDGACNHDCGYYN  150

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
            + ++    DN++   Q +CTG +++C          V            S +RY +++
Sbjct  151  SAKNEGCFDNKFCSEQPKCTGGVHDCR--------FVESSMQICQAEKNSSRRYEFIK--  200

Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKD  173
                             Y    V+GN++       +  +W R   + C  C CLCD+ +
Sbjct  201  -----------------YESGLVHGNEKPCASWLTSAKSWNRWLIMECSYCLCLCDD-QS  242

Query  174  SESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD  233
             +S R F+L  V +D  +NKVV+G++   +++IF+ +I++  LLPLG  ++  L WK  D
Sbjct  243  PQSDRYFNLRPVIADVDNNKVVVGLRFVKRNRIFHLQIQEGELLPLGAINETTLLWKPVD  302

Query  234  PSLSSA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKF  288
                S      + D   L   KRSI +G   + ++  VVTG++F     +L L  + T
Sbjct  303  SYSISDKDVSENVDYHTLTYQKRSIDLGDLNAKINS-VVTGLRFRVFGAHLQLEAQFTAV  361

Query  289  DVMSGKLMNDN--NQWIYNSLSKQKLNN-GGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +  SGKL++    + WI     + +       A    K +        D   +EF   +
Sbjct  362  EFESGKLVDPKKTSHWISMGTDQPREKLPLMDAQVSTKSSVLSIPYPSDNQFIEFTNTGF  421

Query  346  KDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
              D  Q TVPFID++ +VS     L G+G+ ++ ++  GGF+AP ++T+
Sbjct  422  IQDAAQRTVPFIDIQDVVSEPPVPLSGIGIYYKGQNGYGGFLAPRLITY  470
```

```
>☐ ref|XP_001960321.1| G GF11569 [Drosophila ananassae]
 gb|EDV37143.1| G GF11569 [Drosophila ananassae]
Length=662

 GENE ID: 6494433 Dana\GF11569 | GF11569 gene product from transcript GF11569-RA
[Drosophila ananassae] (10 or fewer PubMed links)

 Score =  474 bits (1219),  Expect = 9e-132, Method: Composition-based stats.
 Identities = 109/404 (26%), Positives = 164/404 (40%), Gaps = 47/404 (11%)

Query  10   RTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYAAPRHNFKE--  62
            R     DP    H++GET+  +TR  Q Y   E N  S   C + C+DY     N
Sbjct  275  RRYWRCDPAKDRHVEGETFARVTRLLQGYVENEINLNSDQSCWNSCDDYHDTLSNGCYQP  334

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
            +      Q  C G +YNC          +       P S    +RY Y+     T
Sbjct  335  EEDLCGQQPPCNGRLYNC--------AFIDSALTICPASENGTRRYEYIHYGDGTSLG--  384

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                            G             +W R   +RC  C CLCDE    +S R F+
Sbjct  385  ----------------GKSAGCKSKEKEAASWSRWVVMRCHYCFCLCDEP-GPKSDRYFN  427

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL-----  236
            L A  SD   N++V GV+     ++F+ +++Q+ L+P G  +   L W   D
Sbjct  428  LRASMSDFMRNRIVTGVRFVKSRRVFHLQLQQAQLVPWGAINTSTLEWLPLDDYNVSDAD  487

Query  237  SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLM  296
                 D   L  + RS+ +    +N +D VVTGV+F  I+ +L L V+ T F+  +G L+
Sbjct  488  IRNDYDYHTLGPDSRSLDLDEIVANATDQVVTGVRFRIINKHLNLEVRFTPFNFSTGLLV  547

Query  297  NDN--NQWIYNSLSKQKLNN----GGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVG  350
                 + W+ N+ S             A+ P                ++F  S  + DV
Sbjct  548  EPQTMSFWLGNNNSHVDGKRQKLVLKEADLPTASEMPSLPLSQTNQYMDFVSSSRQKDVA  607

Query  351  QSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTH  393
            Q+T+PFID++ +V      L G+G+  + R   GGF AP VVT+
Sbjct  608  QNTLPFIDIQEVVPHPAVPLSGLGIFHKGRSGYGGFFAPKVVTY  651


>☐ ref|XP_002077790.1| G GD22891 [Drosophila simulans]
 gb|EDX03375.1| G GD22891 [Drosophila simulans]
Length=490

 GENE ID: 6730611 Dsim\GD22891 | GD22891 gene product from transcript GD22891-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score =  474 bits (1219),  Expect = 1e-131, Method: Composition-based stats.
 Identities = 113/409 (27%), Positives = 188/409 (45%), Gaps = 47/409 (11%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA  54
            LP +        R++   DP+EH  G TYEE+TR  Q Y   E +  S   CN  C  Y
Sbjct  91   LPKIKVLMSQAERSVWRCDPQEHQSGVTYEEVTRLLQGYVENEVDLNSDGACNHDCGYYN  150

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
            + ++    DN++   Q +CTG +++C          V            S +RY +++
Sbjct  151  SAKNEGCFDNKFCSEQPKCTGGVHDCR--------FVESSMQICQAEKNSSRRYEFIK--  200

Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKD  173
                             Y    V+GN++       +  +W R   + C  C CLCD+ +
Sbjct  201  -----------------YESGLVHGNEKPCASWLTSAKSWNRWLIMECSYCLCLCDD-QS  242

Query  174  SESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD  233
             +S R F+L  V +D  +N+VV+G++   +++IF+ +I++  LLPLG  ++  L WK  D
Sbjct  243  PQSDRYFNLRPVIADVDNNRVVVGLRFVKRNRIFHLQIQEGELLPLGAINETTLLWKPVD  302

Query  234  PSLSSA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKF  288
                S      + D   L   KRSI +G   + ++  VVTG++F     +L L  + T
Sbjct  303  SYSISDKDVSENVDYHTLTYQKRSIDLGDLNAKINS-VVTGLRFRVFGAHLQLEAQFTAV  361
```

```
Query  289  DVMSGKLMNDN--NQWIYNSLSKQKLNN-GGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +  SGKL++    + WI     + +       A    K +        D   +EF  +
Sbjct  362  EFESGKLVDPKKTSHWISMGTDQPREKLPLMDAQVSTKSSVLSIPYPSDNQFIEFTNTGF  421

Query  346  KDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
              D  Q TVPFID++ +VS     L G+G+ ++ ++  GGF+AP ++T+
Sbjct  422  IQDAAQRTVPFIDIQDVVSEPPVPLSGIGIYYKGQNGYGGFLAPRLITY  470


>☐ gb|AAY55318.1|  IP12601p [Drosophila melanogaster]
Length=571

 Score =  473 bits (1217),  Expect = 2e-131, Method: Composition-based stats.
 Identities = 113/409 (27%), Positives = 187/409 (45%), Gaps = 47/409 (11%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA  54
            LP +        R++   DP+EH  G T+EE+TR  Q Y   E +  S   CN +C  Y
Sbjct  180  LPKIKVLMSQADRSVWRCDPQEHQSGVTHEEVTRLLQGYVENEVDLNSDGACNQECGYYK  239

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
              ++    DN++   Q +CTG +++C          V            S +RY +++
Sbjct  240  LAKNEGCFDNKFCSEQPKCTGGVHDCR--------FVESNMKICQAEKNSSRRYEFIK--  289

Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR-CDICRCLCDESKD  173
                             Y    V+GN++       +  +W R     C  C CLCD+ +
Sbjct  290  -----------------YESGLVHGNEKPCATWLNSAKSWNRWLFMECSYCLCLCDD-QS  331

Query  174  SESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD  233
             +S R F+L  V +D K+N+VV+G++   +++IF+ ++++  LLPLG  +   L WK  D
Sbjct  332  PQSDRYFNLRPVIADVKNNRVVVGLRFVKRNRIFHLQVQEGELLPLGDINGTTLRWKPVD  391

Query  234  PSLSSA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKF  288
                S      + D   L   KRSI +G   + ++  VVTG++F     +L L  + T
Sbjct  392  SYSISDKDVSENVDYHTLTYEKRSIDLGDLNAKINS-VVTGLRFRVFGAHLQLEAQFTAV  450

Query  289  DVMSGKLMNDN--NQWIYNSLSKQK-LNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +  SGKL++    + WI     + +       A    K +        D   +EF  +
Sbjct  451  EFESGKLIDPKKTSHWISMGTDQPRERLPLMDAQISTKSSVLSIPYPSDNQFIEFTNTGF  510

Query  346  KDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
              D  Q TVPFID++ +VS     L G+G+ ++ +D  GGF+AP ++T+
Sbjct  511  VQDAAQRTVPFIDIQDVVSDPPVPLSGIGIYYKGQDGYGGFLAPRIITY  559


>☐ gb|ACH95281.1|  FI07603p [Drosophila melanogaster]
Length=571

 Score =  473 bits (1217),  Expect = 2e-131, Method: Composition-based stats.
 Identities = 113/409 (27%), Positives = 187/409 (45%), Gaps = 47/409 (11%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA  54
            LP +        R++   DP+EH  G T+EE+TR  Q Y   E +  S   CN +C  Y
Sbjct  180  LPKIKVLMSQADRSVWRCDPQEHQSGVTHEEVTRLLQGYVENEVDLNSDGACNQECGYYK  239

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
              ++    DN++   Q +CTG +++C          V            S +RY +++
Sbjct  240  LAKNEGCFDNKFCSEQPKCTGGVHDCR--------FVESNMKICQAEKNSSRRYEFIK--  289

Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR-CDICRCLCDESKD  173
                             Y    V+GN++       +  +W R     C  C CLCD+ +
Sbjct  290  -----------------YESGLVHGNEKPCATWLNSAKSWNRWLFMECSYCLCLCDD-QS  331

Query  174  SESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD  233
             +S R F+L  V +D K+N+VV+G++   +++IF+ ++++  LLPLG  +   L WK  D
Sbjct  332  PQSDRYFNLRPVIADVKNNRVVVGLRFVKRNRIFHLQVQEGELLPLGDINGTTLRWKPVD  391

Query  234  PSLSSA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKF  288
                S      + D   L   KRSI +G   + ++  VVTG++F     +L L  + T
Sbjct  392  SYSISDKDVSENVDYHTLTYEKRSIDLGDLNAKINS-VVTGLRFRVFGAHLQLEAQFTAV  450

Query  289  DVMSGKLMNDN--NQWIYNSLSKQK-LNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +  SGKL++    + WI     + +       A    K +        D   +EF  +
```

```
Sbjct  451  EFESGKLIDPKKTSHWISMGTDQPRERLPLMDAQVSTKSSVLSIPYPSDNQFIEFTNTGF  510

Query  346  KDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
              D  Q TVPFID++ +VS     L G+G+ ++ +D  GGF+AP ++T+
Sbjct  511  VQDAAQRTVPFIDIQDVVSDPPVPLSGIGIYYKGQDGYGGFLAPRIITY  559


>☐ ref|NP_722735.1| UG CG31664 [Drosophila melanogaster]
 gb|AAF51336.2| G CG31664 [Drosophila melanogaster]
Length=482

 GENE ID: 33359 CG31664 | CG31664 gene product from transcript CG31664-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score =  471 bits (1213),  Expect = 5e-131, Method: Composition-based stats.
 Identities = 113/409 (27%), Positives = 187/409 (45%), Gaps = 47/409 (11%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA  54
            LP +        R++   DP+EH  G T+EE+TR  Q Y   E +  S   CN +C  Y
Sbjct  91   LPKIKVLMSQADRSVWRCDPQEHQSGVTHEEVTRLLQGYVENEVDLNSDGACNQECGYYK  150

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
             ++     DN++   Q +CTG +++C          V            S +RY +++
Sbjct  151  LAKNEGCFDNKFCSEQPKCTGGVHDCR--------FVESNMKICQAEKNSSRRYEFIK--  200

Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR-CDICRCLCDESKD  173
                             Y    V+GN++       +  +W R     C  C CLCD+ +
Sbjct  201  -----------------YESGLVHGNEKPCATWLNSAKSWNRWLFMECSYCLCLCDD-QS  242

Query  174  SESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD  233
             +S R F+L  V +D K+N+VV+G++   +++IF+ ++++  LLPLG  +   L WK  D
Sbjct  243  PQSDRYFNLRPVIADVKNNRVVVGLRFVKRNRIFHLQVQEGELLPLGDINGTTLRWKPVD  302

Query  234  PSLSSA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKF  288
                S      + D   L   KRSI +G   + ++  VVTG++F     +L L  + T
Sbjct  303  SYSISDKDVSENVDYHTLTYEKRSIDLGDLNAKINS-VVTGLRFRVFGAHLQLEAQFTAV  361

Query  289  DVMSGKLMNDN--NQWIYNSLSKQK-LNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +  SGKL++    + WI     + +       A    K +        D   +EF  +
Sbjct  362  EFESGKLIDPKKTSHWISMGTDQPRERLPLMDAQVSTKSSVLSIPYPSDNQFIEFTNTGF  421

Query  346  KDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
              D  Q TVPFID++ +VS     L G+G+ ++ +D  GGF+AP ++T+
Sbjct  422  VQDAAQRTVPFIDIQDVVSDPPVPLSGIGIYYKGQDGYGGFLAPRIITY  470


>☐ ref|XP_001648416.1| UG hypothetical protein AaeL_AAEL014309 [Aedes aegypti]
 gb|EAT33416.1| G conserved hypothetical protein [Aedes aegypti]
Length=663

 GENE ID: 5564091 AaeL_AAEL014309 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

 Score =  470 bits (1209),  Expect = 1e-130, Method: Composition-based stats.
 Identities = 112/405 (27%), Positives = 177/405 (43%), Gaps = 48/405 (11%)

Query  8    FSRTLDNVDPK--EHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA-APRHNFK  61
             SR     DP   +H +GETY ++TR  Q Y   E +  +   C   C  Y    R
Sbjct  264  ASREYWRCDPDHSKHKEGETYVQLTRLLQGYVENEVDMNTDNTCKENCAYYNWGVRQEQC  323

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
              + Y   Q +C G +Y+C        E V       P +  S++RY ++E
Sbjct  324  YKDLYCSKQPKCAGKMYSC--------EYVDSDMWICPAASSSNRRYEFIE---------  366

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSF  180
                      Y    V G K+H    T   D+W R  F  C  C CLCD+S   +S R
Sbjct  367  ----------YENGMVMGQKKHCTRGTTKVDSWWRWLFWHCSYCFCLCDDS-GPKSDRYI  415

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL----  236
            ++    SD  +NKVV G++   K++I +  +++  LLP G+ D   L W   +
Sbjct  416  NMRESVSDVMNNKVVTGLRFIKKNRIIFLIVQEGQLLPRGQIDNTTLQWVEPEAYNILSP  475
```

```
Query  237  -SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
               A  D   ++    R++ +         +VVTGV+F  +  ++ L V++T+ D  SGKL
Sbjct  476  YIRAKVDYNVMNYENRAMDLDDIILQP-PYVVTGVRFRMLGTHMNLEVRMTEMDFGSGKL  534

Query  296  MN-DNNQWIYNSLSKQKLN-----NGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
            ++ D + W+ +  ++   +          + PI        +      ++FR+SD   D
Sbjct  535  IDPDKSIWVGSDKTEHSEDKRTEMKLDKPHIPILSPTKSVPDSEPNQFIKFRESDRGMDA  594

Query  350  GQSTVPFIDLRSIVST-GTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             Q+TVPF D + +V +  T L G GL  + R   GGF+AP V T+
Sbjct  595  AQTTVPFFDAQPVVPSKQTPLGGAGLFHKGRPKFGGFMAPRVFTY  639


>☐ ref|XP_002087581.1| GE15356 [Drosophila yakuba]
 gb|EDW87293.1| GE15356 [Drosophila yakuba]
Length=489

 GENE ID: 6526462 Dyak\GE15356 | GE15356 gene product from transcript GE15356-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score =  469 bits (1208),  Expect = 2e-130, Method: Composition-based stats.
 Identities = 113/400 (28%), Positives = 188/400 (47%), Gaps = 41/400 (10%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYAAPRHNFKED  63
               R++   DP EH  G TYEE+TR  Q Y   E +  S   C+  C  Y + R+    D
Sbjct  98   RADRSVWRCDPWEHKSGVTYEEVTRLLQGYVENEVDLNSDGVCSRDCGYYKSARNEGCFD  157

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
            +++   Q +C+G +++C          V            S +RY ++E
Sbjct  158  SKFCSEQPKCSGGVHDCR--------FVESSMKICQAEKTSSRRYEFIE-----------  198

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR-CDICRCLCDESKDSESIRSFSL  182
                    Y    V+G +EH D  + +  +W R     C  C CLCD+ + ++S R F+L
Sbjct  199  --------YKSGLVHGTQEHCDTWSNSAKSWNRWLFMECSYCLCLCDD-QSAQSDRYFNL  249

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LS  237
              V +D K+N+VV G++   K++IF+ ++++  LLPLG  ++  L WK  D
Sbjct  250  RPVVADVKNNRVVTGLRFVKKNRIFHLQVQEGELLPLGVINKPTLQWKPVDSYSIFDKNV  309

Query  238  SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMN  297
            + + D  KL    RS+ +G   +  +  VVTG++F  + G+L L V+ +  +  SGKL++
Sbjct  310  TDTVDYHKLSYESRSLDLGDVKAENNSAVVTGLRFRVMGGHLQLQVQFSAVEFESGKLID  369

Query  298  D--NNQWIYNSLSKQKLNN-GGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTV  354
                + W      +++       A    K +        D   +EF  +    D  QSTV
Sbjct  370  PTKTSHWTSLGTDQKREKLPLRDAQVSSKSSVPSIPYPSDNQYIEFTNTGFVQDAAQSTV  429

Query  355  PFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTH  393
            PFID++ +V+     L G+G+ ++ RD  GGF+AP + T+
Sbjct  430  PFIDIQDVVTEPAVPLSGIGIYYKGRDGYGGFLAPRLTTY  469


>☐ ref|XP_001968326.1| GG24571 [Drosophila erecta]
 gb|EDV57385.1| GG24571 [Drosophila erecta]
Length=488

 GENE ID: 6543280 Dere\GG24571 | GG24571 gene product from transcript GG24571-RA
[Drosophila erecta] (10 or fewer PubMed links)

 Score =  464 bits (1194),  Expect = 8e-129, Method: Composition-based stats.
 Identities = 111/415 (26%), Positives = 186/415 (44%), Gaps = 50/415 (12%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYA  54
            LP +        R++   DP +H  G TYEE+TR  Q Y   E +     +C   C  Y
Sbjct  90   LPKIKELMSLADRSVWRCDPWKHEAGVTYEEVTRLLQGYVENEVDLNDDGTCRRDCGYYK  149

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
            + R+    D+ +   Q +C+G +++C          V            S +RY ++E
Sbjct  150  SARNEGCFDSEFCSEQPKCSGGVHDCR--------FVESNMQICQAEKNSSRRYEFIE--  199
```

```
Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKD  173
                             Y    V+GN++    T  +  +W R  FV C  C CLCD+ +
Sbjct  200  -----------------YQSGLVHGNQKPC-GTRNSAKSWNRWLFVECSYCLCLCDD-QS  240

Query  174  SESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD  233
            + S R F+L  V +D ++N+VV G++     +IF+ ++++  LLPLG  ++  L WK  D
Sbjct  241  ARSDRYFNLRPVIADVENNRVVTGLRFVKSRRIFHLQVQEGQLLPLGVINETTLQWKPLD  300

Query  234  PSLSSASK-----DIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKF  288
                S        D   L   KR++ +G   +  +  VVTG++F  +  +L L  + +
Sbjct  301  AYSISDKDVTEHADYHTLSYEKRALDLGDLKA-DNSAVVTGLRFRVVGAHLQLEAQFSAL  359

Query  289  DVMSGKLMNDN--NQWIYNSLSKQKLNN-GGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +  SGKL++    +QW     S+ +       A    K  +       D   +EF  +
Sbjct  360  EFESGKLIDPKRTSQWTSLGASQHREKLPLKDAQVSSKSPELSIPYPSDNQYIEFTNTGF  419

Query  346  KDDVGQSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTHP--TYT  397
             +D  Q+TVPFID++ +VS     L G+G+ ++  +  GGF+AP + T+    Y
Sbjct  420  VEDAAQTTVPFIDIQDVVSEPAVPLSGIGIYYKGLEGYGGFLAPRITTYDFTPYV  474


>☐ ref|XP_002044470.1| G GM11959 [Drosophila sechellia]
 gb|EDW53677.1| G GM11959 [Drosophila sechellia]
Length=799

 GENE ID: 6620262 Dsec\GM11959 | GM11959 gene product from transcript GM11959-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score =  463 bits (1192),  Expect = 1e-128, Method: Composition-based stats.
 Identities = 101/400 (25%), Positives = 166/400 (41%), Gaps = 50/400 (12%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            L   R +   DP +H++G TY+E+TR  Q Y   E +     +C   C  Y + R
Sbjct  425  LRADRIVYRCDPTKHVQGVTYDEVTRLLQGYIENEVDLNNEETCRETCGFYQSTRSEGCF  484

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
             + Y   Q +CTG +YNC          V       P    S +RY ++E
Sbjct  485  KDLYCSRQPKCTGRLYNCQ--------FVDSDMWVCPSPQNSTRRYEFIE----------  526

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                     Y    V G +       T   D+W R     C  C CLCDE +  +S R F+
Sbjct  527  ---------YENGRVLGQRGKCTRGTTKVDSWWRYLLWHCSYCFCLCDE-ESLKSDRFFN  576

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASK  241
            +    +D K N++V G + +           QS    +G   ++   +   D
Sbjct  577  MRDTIADIKRNRIVTGGRTSAAWHC------QSKHSGMGNPVEK---YNVFDRH-VKNGV  626

Query  242  DIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMND--N  299
            D  KL    R+I +    +  + FVVTGV+F  +  +L L    ++FD ++G+L+    N
Sbjct  627  DYHKLSYEYRTIDLDDVDTEDNSFVVTGVRFRVVGTHLNLEAYYSEFDFLTGQLIQPEYN  686

Query  300  NQWIYNSLSK-----QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTV  354
            + W  N  +      ++      A+   +              ++F  + L  D  QSTV
Sbjct  687  SYWKSNDNTDVSGARREKLRLSNADVSTRTIAPSLPLSRHNQFIDFTNTGLDKDAAQSTV  746

Query  355  PFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            PF+D++ +VS     L G+G+ ++ R+  GGF+ P ++T+
Sbjct  747  PFMDIQDVVSNPPVPLAGIGIYYKGRNGYGGFLGPKIITY  786


>☐ ref|XP_001851109.1| UG conserved hypothetical protein [Culex quinquefasciatus]
 gb|EDS33062.1| G conserved hypothetical protein [Culex quinquefasciatus]
Length=582

 GENE ID: 6041957 CpipJ_CPIJ009482 | hypothetical protein
[Culex quinquefasciatus]

 Score =  462 bits (1190),  Expect = 3e-128, Method: Composition-based stats.
 Identities = 104/405 (25%), Positives = 171/405 (42%), Gaps = 48/405 (11%)

Query  8    FSRTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA-APRHNFK  61
```

```
             SR     DP   +H +G+T+ +ITR  Q Y   E N  +   C  +C  Y+   +
Sbjct  198  ASREYWRCDPERNQHKEGKTFVQITRLLQGYVELELNMNTDDTCMEECSFYSWGVQQEGC  257

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
              ++Y   Q +C+G IY+C        E +       P +  S +RY+++E
Sbjct  258  YKDQYCTMQPKCSGKIYDC--------EFIDSDMWICPAAKTSHRRYHFIE---------  300

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR-CDICRCLCDESKDSESIRSF  180
                      Y    V G K++         +W R F   C  C C+CD++    S R
Sbjct  301  ----------YDNGKVLGQKKYCAQGRTKVMSWTRYFFYRCSYCFCMCDDA-GKNSDRYI  349

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----  235
            ++    SD  DNKVV G++   K++I +  ++Q  LLP G  D   L W
Sbjct  350  NMRESVSDVMDNKVVTGMRFIKKNQIIHLIVQQGRLLPRGHIDNGTLEWVEPHNYTVMSQ  409

Query  236  LSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
                 ++   L+   RS+ +         +V+TGV+F  +  +L L ++ T+ +  +G+L
Sbjct  410  NVRTGREYHTLNRINRSLDLDDLYV-PQGYVMTGVRFRLLGNHLNLEIRKTEMNFATGQL  468

Query  296  MNDN-NQWIYNSLS-----KQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
            ++ N N WI N  +     K+        + P+        +      ++FR SD   D
Sbjct  469  IDSNKNIWIGNDNTERSVDKRTELKLDRPHVPLLSPSKSIPDSASNQFIQFRASDRDMDA  528

Query  350  GQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             Q+TVPF D + +       L G GL  + R   GGF+A  V+T+
Sbjct  529  AQTTVPFFDTQPVAPVRPVPLSGAGLFHKGRPKFGGFVASKVMTY  573
```

>☐ ref|XP_001648417.1| UG hypothetical protein AaeL_AAEL014309 [Aedes aegypti]
 gb|EAT33417.1| G conserved hypothetical protein [Aedes aegypti]
Length=684

 GENE ID: 5564091 AaeL_AAEL014309 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

```
 Score =  462 bits (1189),  Expect = 3e-128, Method: Composition-based stats.
 Identities = 112/426 (26%), Positives = 177/426 (41%), Gaps = 69/426 (16%)

Query  8    FSRTLDNVDPK--EHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA-APRHNFK  61
             SR     DP   +H +GETY ++TR  Q Y   E +  +   C   C  Y    R
Sbjct  264  ASREYWRCDPDHSKHKEGETYVQLTRLLQGYVENEVDMNTDNTCKENCAYYNWGVRQEQC  323

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
              + Y   Q +C G +Y+C        E V       P +  S++RY ++E
Sbjct  324  YKDLYCSKQPKCAGKMYSC--------EYVDSDMWICPAASSSNRRYEFIE---------  366

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSF  180
                      Y    V G K+H    T   D+W R  F  C  C CLCD+S   +S R
Sbjct  367  ----------YENGMVMGQKKHCTRGTTKVDSWWRWLFWHCSYCFCLCDDS-GPKSDRYI  415

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL----  236
            ++    SD  +NKVV G++   K++I +  +++  LLP G+ D   L W   +
Sbjct  416  NMRESVSDVMNNKVVTGLRFIKKNRIIFLIVQEGQLLPRGQIDNTTLQWVEPEAYNILSP  475

Query  237  -SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
               A  D   ++   R++ +         +VVTGV+F  +  ++ L V++T+ D  SGKL
Sbjct  476  YIRAKVDYNVMNYENRAMDLDDIILQP-PYVVTGVRFRMLGTHMNLEVRMTEMDFGSGKL  534

Query  296  MN-DNNQWIYNSLSKQKLN--------------------------NGGVANRPIKFTDNH  328
            ++ D + W+ +  ++   +                               + PI
Sbjct  535  IDPDKSIWVGSDKTEHSEDKRLVPEQLPSEVPFFNSSSISSRTEMKLDKPHIPILSPTKS  594

Query  329  ASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVST-GTALIGVGLQFRTRDTSGGFIA  387
              +      ++FR+SD   D  Q+TVPF D + +V +  T L G GL  + R   GGF+A
Sbjct  595  VPDSEPNQFIKFRESDRGMDAAQTTVPFFDAQPVVPSKQTPLGGAGLFHKGRPKFGGFMA  654

Query  388  PVVVTH  393
            P V T+
Sbjct  655  PRVFTY  660
```

```
>☐ ref|XP_002015853.1|  GL10792 [Drosophila persimilis]
 gb|EDW31743.1|  GL10792 [Drosophila persimilis]
Length=670

 GENE ID: 6590462 Dper\GL10792 | GL10792 gene product from transcript GL10792-RA
[Drosophila persimilis] (10 or fewer PubMed links)

 Score =  462 bits (1188),  Expect = 4e-128, Method: Composition-based stats.
 Identities = 112/415 (26%), Positives = 168/415 (40%), Gaps = 49/415 (11%)

Query  6    LGFSRTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNF  60
                R     DP   +H+ GETY+  TR  Q Y   EAN G    C   C DY   R N
Sbjct  266  ARSGRVYWRCDPGKDKHVSGETYDRATRLLQGYVENEANLGENQGCWGTCGDYPDTRSNG  325

Query  61   KED--NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTD  118
                       Q  C G +YNC       +          P    S +RY Y++
Sbjct  326  CYKGDKELCGKQTPCNGRLYNCDDMDSDVE--------ICPSEGNSSRRYEYIQSDDGAT  377

Query  119  KKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDT-WIRGFVRCDICRCLCDESKDSESI  177
              +                  +    +  T    + W     RC  C CLCDE +   S
Sbjct  378  LGRKS----------------SNGRCESATKTASSGWYYMLWRCSYCFCLCDE-QGPLSD  420

Query  178  RSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS--  235
            R F+L    SD  +NKVV GV+    +++F+ +++Q  LLP G  +   L W   D
Sbjct  421  RYFNLRDSLSDFMENKVVTGVRFVKSNRVFHLQLQQGELLPRGAINVSSLEWLPLDSYHL  480

Query  236  ---LSSASKDIFKLDVNKRSIHIGSATSNLS-DFVVTGVKFAEIDGNLALGVKLTKFDVM  291
                     D   L V+ R++ +   TS  +   VVTG +F     +L L V+ + FD
Sbjct  481  NDVDVRNDYDYHTLTVDSRAMDLDEITSGSNRSLVVTGARFRVWHKHLNLEVRFSYFDFP  540

Query  292  SGKLMND--NNQWIYNSLSKQKLNNG------GVANRPIKFTDNHASEKLDLGLVEFRQS  343
            +G+L+    N+ W+ NS +++ +           A+ P                + F  S
Sbjct  541  TGRLVEPKRNSVWLNNSNTEKTIEKPRQELIVKEADVPTAADQPSLPLSTSNQYMNFFSS  600

Query  344  DLKDDVGQSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTHPTYT  397
                D  Q+T+PFID++ +V      L G+G+ +++R  SGGF AP VVT+   Y+
Sbjct  601  SKVKDASQNTLPFIDVQEVVPHPAVPLAGLGIFYKSRPGSGGFFAPKVVTY-DYS  654


>☐ gb|ACM40903.1|  secreted salivary protein [Culicoides nubeculosus]
Length=417

 Score =  459 bits (1180),  Expect = 4e-127, Method: Composition-based stats.
 Identities = 398/398 (100%), Positives = 398/398 (100%), Gaps = 0/398 (0%)

Query  1    NKLPSLGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAAPRHNF  60
            NKLPSLGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAAPRHNF
Sbjct  20   NKLPSLGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAAPRHNF  79

Query  61   KEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK  120
            KEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK
Sbjct  80   KEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK  139

Query  121  KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSF  180
            KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSF
Sbjct  140  KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSF  199

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSAS  240
            SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSAS
Sbjct  200  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSAS  259

Query  241  KDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNN  300
            KDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNN
Sbjct  260  KDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNN  319

Query  301  QWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLR  360
            QWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLR
Sbjct  320  QWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLR  379

Query  361  SIVSTGTALIGVGLQFRTRDTSGGFIAPVVVTHPTYTG  398
            SIVSTGTALIGVGLQFRTRDTSGGFIAPVVVTHPTYTG
```

```
Sbjct  380  SIVSTGTALIGVGLQFRTRDTSGGFIAPVVVTHPTYTG  417


>☐ ref|XP_001360382.2|  GA15824 [Drosophila pseudoobscura pseudoobscura]
 gb|EAL24957.2|  GA15824 [Drosophila pseudoobscura pseudoobscura]
Length=670

 GENE ID: 4803708 Dpse\GA15824 | GA15824 gene product from transcript GA15824-RA
[Drosophila pseudoobscura pseudoobscura] (10 or fewer PubMed links)

 Score =  456 bits (1173),  Expect = 2e-126, Method: Composition-based stats.
 Identities = 111/415 (26%), Positives = 167/415 (40%), Gaps = 49/415 (11%)

Query  6    LGFSRTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNF  60
                R     DP   +H+ GETY+  TR  Q Y   EAN G    C   C DY   R N
Sbjct  266  ARSGRVYWRCDPGKDKHVSGETYDRATRLLQGYVENEANLGENQGCWGTCGDYPDTRSNG  325

Query  61   KED--NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTD  118
                       Q  C G +YNC       +          P    S +RY Y++
Sbjct  326  CYKGDKELCGKQTPCNGRLYNCGDMDSDVE--------ICPSEGNSSRRYEYIQSDDGAT  377

Query  119  KKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESI  177
              +                  +    +  T    +       RC  C CLCDE +   S
Sbjct  378  LGRKS----------------SNGRCESATKTASSGCYYMLWRCSYCFCLCDE-QGPLSD  420

Query  178  RSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS--  235
            R F+L    SD  +NKVV GV+    +++F+ +++Q  LLP G  +   L W   D
Sbjct  421  RYFNLRDSLSDFMENKVVTGVRFVKSNRVFHLQLQQGELLPRGAINVSSLEWLPLDSYHL  480

Query  236  ---LSSASKDIFKLDVNKRSIHIGSATSNLS-DFVVTGVKFAEIDGNLALGVKLTKFDVM  291
                     D   L V+ R++ +   TS  +   VVTG +F     +L L V+ + FD
Sbjct  481  NDVDVRNDYDYHTLTVDSRAMDLDEITSGSNRSLVVTGARFRVWHKHLNLEVRFSYFDFP  540

Query  292  SGKLMND--NNQWIYNSLSKQKLNNG------GVANRPIKFTDNHASEKLDLGLVEFRQS  343
            +G+L+    N+ W+ NS +++ +           A+ P                + F  S
Sbjct  541  TGRLVEPKRNSVWLNNSNTEKTIEKPRQELIVKEADVPTAADQPSLPLSTSNQYMNFFSS  600

Query  344  DLKDDVGQSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTHPTYT  397
                D  Q+T+PFID++ +V      L G+G+ +++R  SGGF AP VVT+  Y+
Sbjct  601  SKVKDASQNTLPFIDVQEVVPHPAVPLAGLGIFYKSRPGSGGFFAPKVVTY-DYS  654


>☐ ref|XP_002092594.1|  GE14278 [Drosophila yakuba]
 gb|EDW92306.1|  GE14278 [Drosophila yakuba]
Length=678

 GENE ID: 6531807 Dyak\GE14278 | GE14278 gene product from transcript GE14278-RA
[Drosophila yakuba] (10 or fewer PubMed links)

 Score =  451 bits (1160),  Expect = 7e-125, Method: Composition-based stats.
 Identities = 107/406 (26%), Positives = 165/406 (40%), Gaps = 50/406 (12%)

Query  8    FSRTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
             SR     DP   EH++G+TY+ +TR  Q +   E N     SC S C DY   R +
Sbjct  272  SSRLYWRCDPGKNEHVEGQTYDRVTRLLQGFVENEINLNPNQSCRSSCPDYHDTRSSGCY  331

Query  63   --DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK  120
              +  +   Q  C G +YNC   +           +    +  S +RY Y+     +
Sbjct  332  QPEEEFCGQQPACKGRLYNCNTIESDIS-------VCLSPNATSARRYEYITFHEDETSP  384

Query  121  KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRS  179
                                      T     +W    F RC  C CLCDE     S R
Sbjct  385  V----------------------CKSTVRKPSSWSSWIFWRCHYCFCLCDEP-GPNSDRF  421

Query  180  FSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL---  236
            F+L    SD K NKVV GV+     ++F+ +++Q  LLP G      L W   +
Sbjct  422  FNLRDSMSDFKQNKVVTGVRFVKIKRVFHLQLQQGELLPGGVIKNSSLEWLPVEAYNVSD  481

Query  237  --SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGK  294
                   D   L V+ R++ +   + N +D VVTG +F     +L L V+ T F+  +G+
Sbjct  482  VDIRDGYDYHTLSVDSRAMDLDEISVNTTDQVVTGFRFRVFKKHLNLEVRFTSFNFSTGQ  541
```

```
Query  295  LMND--NNQWIYNSLS----KQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDD  348
            L+     + W+ N  S    ++K      ++ P            D   ++F  S    D
Sbjct  542  LIEPQTKSFWLGNQNSHLEGERKRLILKESDLPTASELPSLPLSQDNQYMDFVSSSQLKD  601

Query  349  VGQSTVPFIDLRSIVSTGT-ALIGVGLQFRTRDTSGGFIAPVVVTH  393
              Q+T P++D++ +V      L G+G+ ++ R   GGF AP VVT+
Sbjct  602  AAQNTFPYVDVQEVVPQQAGPLAGLGIYYKGRPGYGGFFAPKVVTY  647


>☐ ref|XP_002427759.1| G conserved hypothetical protein [Pediculus humanus corporis]
 gb|EEB15021.1| G conserved hypothetical protein [Pediculus humanus corporis]
Length=635

 GENE ID: 8231845 Phum_PHUM343650 | hypothetical protein
[Pediculus humanus corporis]

 Score =  447 bits (1149),  Expect = 1e-123, Method: Composition-based stats.
 Identities = 115/405 (28%), Positives = 183/405 (45%), Gaps = 49/405 (12%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            +   R+L   DPK+H++G TY ++T+  Q Y   E N     SC+  C  Y    +N
Sbjct  254  MNADRSLYKCDPKQHVEGVTYVKLTKLLQGYIENEVNLNGGGSCSENCGYYEYTEYNGCY  313

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
             N++   Q  C G + NC          V         S  S++RY ++E
Sbjct  314  HNQFCSHQPPCRGKVLNCK--------FVDSDMTICRSS-NSERRYDWIE----------  354

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTV-NFDTWIRG-FVRCDICRCLCDESKDSESIRSF  180
                     Y      GN    ++  +   D+W R  F  C  C CLCD+ +D  S R F
Sbjct  355  ---------YENGNNLGNPGVCENRNLFKVDSWWRYLFWHCSYCLCLCDDPED--SDRYF  403

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----  235
            +L  V SD + NKVV G++L   ++I + +I+Q  L P G  +   LSWK  D
Sbjct  404  NLRDVTSDVEKNKVVTGIRLVKHNRIIHIQIQQGSLNPRGIINNGTLSWKPVDNYTIDDL  463

Query  236  LSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
                + D   +   +R+I +           VTGV+F ++  +L L +++T F+  +G++
Sbjct  464  NVENNVDYHTMTWAERTIDMDDVFGPSE-HAVTGVRFRKLGSHLNLEIRVTPFNFTTGQV  522

Query  296  -MNDNNQWIYNSLSK------QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDD  348
               D + W+ N  +       +        + PI+    H  +      ++F+ SD   D
Sbjct  523  SSTDKSYWLANDNTDGNLKNPRTEIRLNNPDIPIR-KGPHTLDSKHDQYLQFKNSDFNTD  581

Query  349  VGQSTVPFIDLRSIVSTGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             GQ+T+P+ID +++ S G  L G+GL  R R  SGGFIAP +VT+
Sbjct  582  AGQNTIPYIDSQALTSHGFPLSGLGLIHRGRKNSGGFIAPKLVTY  626


>☐ ref|XP_001946173.1| UG PREDICTED: similar to GA15301-PA [Acyrthosiphon pisum]
Length=677

 GENE ID: 100162921 LOC100162921 | similar to GA15301-PA [Acyrthosiphon pisum]

 Score =  443 bits (1140),  Expect = 2e-122, Method: Composition-based stats.
 Identities = 104/409 (25%), Positives = 166/409 (40%), Gaps = 49/409 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
            +  S      DP +H+K  TY E+T+  Q Y   EA+     SC   C  Y    H
Sbjct  283  ISASNIYWRCDPAKHVKDTTYSEVTQLLQGYIENEADMNREQSCMKDCSYYEVAEHKSCF  342

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
             ++    QR C G + NC          +            S +RY Y+
Sbjct  343  RDQLCAKQRVCNGRLLNCQ--------YIDSDMRICQSKFGSSRRYDYI-----------  383

Query  123  QYAIIEGSNYTYRYVYGNKE-HSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSF  180
                     Y    V G+       TT   D+W       C  C C+CD S DS S R F
Sbjct  384  --------AYEKGKVLGDSSVECKGTTSMVDSWWTFLVWHCSYCFCICD-SNDSGSDRYF  434

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSD-------D  233
            SL    S+ K+N V+ GV+     ++F+ + ++  L P G   +    W
```

```
Sbjct  435  SLKPSVSNIKNNMVITGVRFIKVKQMFHLQAQEGELGPRGGISESSRQWLEVPGSEFSYK  494

Query  234  PSLSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSG  293
                    D   L    R I I +  +   + VVTG+KF +I  +L + V++T FD M+G
Sbjct  495  NKTLKDGVDYHTLTYESRIIDIDNVIA-PENTVVTGLKFRKIGSHLNIEVRITPFDFMTG  553

Query  294  KLMND--NNQWIYNSLSK-----QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLK  346
            +L+     + WI N  ++     +   +   ++ P K +     +      ++F  S +
Sbjct  554  ELIEPELKSYWIGNENTELSINPRTEIDINNSDLPTKSSTASEPDMSSNKFLKFTHSSID  613

Query  347  DDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTHP  394
             DV Q+T+PFID++++    G  L G+G+ ++   + GGFI   V T+
Sbjct  614  RDVAQTTLPFIDMQTVAPFPGVPLSGIGIYYKGTKSFGGFIGANVFTYN  662


>☐ ref|XP_001121007.1| UG PREDICTED: similar to lethal (1) G0193 CG2206-PB, isoform B [Apis
mellifera]
Length=827

 GENE ID: 725122 LOC725122 | similar to lethal (1) G0193 CG2206-PB, isoform B
[Apis mellifera] (10 or fewer PubMed links)

 Score =  439 bits (1130),  Expect = 2e-121, Method: Composition-based stats.
 Identities = 107/398 (26%), Positives = 169/398 (42%), Gaps = 41/398 (10%)

Query  8    FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKEDN  64
              R +   D + H   ETY ++T+  Q Y   E +     SC   C  Y   +      N
Sbjct  432  APRQIWKCDARIHKLDETYTKLTQLFQGYIVNEVDLNKDSSCRENCAYYEYSKVYGCYKN  491

Query  65   RYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQY  124
            ++   QR+C G I  C        E +       P S  +D+RY Y+E
Sbjct  492  QFCAQQRKCNGRILKC--------EYIDSDMWICPSSEHNDRRYEYIE------------  531

Query  125  AIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSFSLN  183
                   Y    V+G K   +  T   D+W R  F  C  C C CD+  +  S R FSL
Sbjct  532  -------YENGQVFGEKSTCNKGTTKVDSWWRWLFWHCSYCFCYCDD-NNMNSDRYFSLR  583

Query  184  AVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LSS  238
             V SD  +NKVV GV+    ++I + +I+Q  L+P G  D+  + WKS +
Sbjct  584  KVVSDVANNKVVTGVRFKKVNQIIHIQIQQGELMPRGNIDKASVHWKSIEEFNVLDSNIK  643

Query  239  ASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMND  298
               D   L   KR +          ++++TG KF  +   L L V++T F+  +GKL+
Sbjct  644  NGIDYHTLSWEKRGLD-LDDLILDKNYLLTGFKFRMVGSRLNLEVRMTPFNFTTGKLIEP  702

Query  299  --NNQWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPF  356
              N+ W  +  + +        N P +       +      + F  SD K D  Q+T+PF
Sbjct  703  LNNSFWFSHDRTDRTELKLENPNIPTRLPLLALPDSKTEQYLNFAPSDRKTDAAQNTIPF  762

Query  357  IDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            +D++S+ S     + G G+  + R  SGG++A  ++T+
Sbjct  763  LDVQSVESNPPIPIAGAGIFHKGRLGSGGYVALRLITY  800


>☐ ref|XP_002049066.1| G GJ21381 [Drosophila virilis]
 gb|EDW60259.1| G GJ21381 [Drosophila virilis]
Length=647

 GENE ID: 6625370 Dvir\GJ21381 | GJ21381 gene product from transcript GJ21381-RA
[Drosophila virilis] (10 or fewer PubMed links)

 Score =  437 bits (1125),  Expect = 9e-121, Method: Composition-based stats.
 Identities = 113/400 (28%), Positives = 172/400 (43%), Gaps = 37/400 (9%)

Query  10   RTLDNVDPK--EHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKEDN  64
            R     DP+  EH+ G+TY+EITR  Q Y   E N     SC   CEDY   R     + N
Sbjct  260  RIYWRCDPRPREHVLGKTYDEITRLMQGYVENEINLNTESSCYQSCEDYQDVRVTGCQKN  319

Query  65   RYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQY  124
             +   Q +C G +++C          +            S +RY ++    +      +
Sbjct  320  EFCNDQPKCAGRVHSCR--------FIDSNMDVCQSYGNSSRRYEFIHYDNDRLLNIDED  371
```

```
Query  125  AIIEGSNYTYRYVYGNKEHSDHTTVNFDTWI-RGFVRCDICRCLCDESKDSESIRSFSLN  183
             I                     +    +W    F RC  C CLCDE    +S R FSL
Sbjct  372  GIQNLV-------------CSRGSNRVKSWSFYLFWRCHYCFCLCDEP-GPQSDRFFSLR  417

Query  184  AVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LSS  238
               SD K NK+V G +     +IF+ +++Q  LLP G  +Q  L W   +
Sbjct  418  DTLSDFKQNKIVTGARFIKDQRIFHLQLQQGELLPGGLINQSSLDWIPLESFDVNHVDIR  477

Query  239  ASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMND  298
               D  +L    RS+ +   T+     VVTG +F  ID +L L V+ + F+  +G+L+N
Sbjct  478  NGFDYHELTFESRSLGLDEITAKNDSLVVTGARFRVIDKHLNLAVRFSLFNFTTGRLLNP  537

Query  299  --NNQWIYNSLSKQKLN-NGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVP  355
              N+ W+ N+  K +        N+P K T        +   +EF  S +  D  QSTVP
Sbjct  538  EVNSLWLGNNADKPRQKLLLKNPNKPTKTTVKSVPLSKNNQFMEFISSSMDMDAAQSTVP  597

Query  356  FIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTHP  394
            FID++ +V+    ++ G+G+  + +   GGF AP +VT
Sbjct  598  FIDVQDVVTNPAMSISGLGIYHKGKKGFGGFFAPKIVTFN  637


>☐ ref|XP_002062683.1| G GK19560 [Drosophila willistoni]
 gb|EDW73669.1| G GK19560 [Drosophila willistoni]
Length=672

 GENE ID: 6639766 Dwil\GK19560 | GK19560 gene product from transcript GK19560-RA
[Drosophila willistoni] (10 or fewer PubMed links)

 Score =  429 bits (1104),  Expect = 2e-118, Method: Composition-based stats.
 Identities = 104/409 (25%), Positives = 167/409 (40%), Gaps = 43/409 (10%)

Query  8    FSRTLDNVDPKE-HIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE-  62
              R     DP   +I G+TY+EITR  Q Y   E N      C S C +Y   +
Sbjct  273  SQRVYWRCDPDGGYILGKTYDEITRLLQGYLENEVNLNENQQCWSTCNNYYETQQFGCYQ  332

Query  63   -DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
             +  Y   Q++C G +YNC          +       P +  S  RY ++E    T   K
Sbjct  333  PEEEYCGKQKKCNGKLYNCR--------YIDSDMWVCPSAENSTHRYKHIEFESGTILGK  384

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSF  180
                        +     N E         ++W R     C  C CLCDE K  +S R F
Sbjct  385  I-----------FPQPPPNPEQCQTEMEKVESWWRYLLWHCSYCFCLCDE-KGPQSDRFF  432

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----  235
            +L    +D ++ +V  G++     ++F+ +++Q  LLP G  +   L W   D
Sbjct  433  NLRDTITDYQNGRVATGIRFVKSQRVFHLQLQQGQLLPNGAINGSSLEWLPIDSYDPDNV  492

Query  236  LSSASKDIFKLDVNKRSIHIGSATSNLS--DFVVTGVKFAEIDGNLALGVKLTKFDVMSG  293
                  D   L  + R++ +    S     D VVTGV+   +  +L L V+ ++++  +G
Sbjct  493  DVRNGIDYHILSFDSRTLDLDEVVSESHSSDHVVTGVRLRMLGKHLNLEVRFSEYNFTTG  552

Query  294  KLMN--DNNQWIYNSLSKQKLNN------GGVANRPIKFTDNHASEKLDLGLVEFRQSDL  345
            +L++  + + W+ N  +                + PI               ++F  S
Sbjct  553  RLLDAKNKSFWLSNDNTDVSGTKRRQRLWLNNPDLPINSKLKSLPLSSTNQFMDFFASSP  612

Query  346  KDDVGQSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTH  393
            + D  QSTVPFID++ +V      L G+GL  + R+  GGF AP V+T+
Sbjct  613  ESDASQSTVPFIDIQDVVPNPVVPLTGLGLYHKGREGFGGFFAPKVITY  661


>☐ ref|XP_001603409.1| UG PREDICTED: similar to CG2206-PA [Nasonia vitripennis]
Length=675

 GENE ID: 100119678 LOC100119678 | similar to CG2206-PA [Nasonia vitripennis]

 Score =  412 bits (1059),  Expect = 3e-113, Method: Composition-based stats.
 Identities = 102/407 (25%), Positives = 179/407 (43%), Gaps = 45/407 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCN------SKCEDYAAPRHN  59
            L   R   + DP +H++G TY ++    Q Y T E + G         +KC D  + R +
```

```
Sbjct  283  LAAPRDYWSCDPDDHVEGLTYAKLDPVFQAYVTSEKDVGISKYFEPFTNKCADSCSDRSD  342

Query  60   FKEDNR--YKFP--QRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPV  115
            F +  R  Y F   Q++C G ++NC     + +               S +R+ ++EL
Sbjct  343  FHDICRSAYCFKHNQKRCYGKLHNCKSSGKSLRM--------CLSDKNSTRRFEFIELTG  394

Query  116  NTDKKKYQYAIIEGSNYTYRYVYGNKEHS-DHTTVNFDTWIRGFVRCDICRCLCDESKDS  174
               +                   G+K+    +T V           C+ C C+CD+  ++
Sbjct  395  KKTE-----------------FLGSKKSCGSNTQVKVSQTTLFLETCEYCFCICDD-TNN  436

Query  175  ESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDP  234
            E  R FSL  V SD   N VV GV+ A ++ IF+ +I+Q  L+  G      + WK  +
Sbjct  437  EMDRYFSLRNVNSDVDKNMVVTGVRFAKENHIFHIQIQQGKLVENGNIKASTVEWKPVEN  496

Query  235  SLSSA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFD  289
                      ++D   +    R +++ +  S L   V+TG++F   +G+L L +++T F+
Sbjct  497  FTIHDDSVVLNRDYHNITWANRLVYLDTFDSPLH-HVLTGLRFQLQNGDLKLEIQVTPFN  555

Query  290  VMSGKLMNDNNQWIYNSLSKQKLNNGGVANRPIKFTDNH-ASEKLDLGLVEFRQSDLKDD  348
              + +L  + + W+ NS+         + N  +    +    +  +   ++F  S +K D
Sbjct  556  FYTAELKTNESTWMSNSIGAYGRGELDLENLDLGTNGHSNLLDSSENQYIKFTTSSMKTD  615

Query  349  VGQSTVPFIDLRSIV-STGTALIGVGLQFRTRDTSGGFIAPVVVTHP  394
            VGQ+ VPF+DL+ +  S    L G GL  R    S GF++  V+T+
Sbjct  616  VGQTVVPFMDLQDVTSSPPVPLSGAGLVHRGDKASAGFLSFKVLTYN  662


>☐ ref|XP_001603438.1| UG PREDICTED: similar to GA15301-PA [Nasonia vitripennis]
Length=669

 GENE ID: 100119712 LOC100119712 | venom protein E [Nasonia vitripennis]

 Score =  405 bits (1042),  Expect = 4e-111, Method: Composition-based stats.
 Identities = 111/404 (27%), Positives = 176/404 (43%), Gaps = 44/404 (10%)

Query  8    FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEAN------FGSCNSKCEDYAAPRHNFK  61
              R + + DP EH + ETY ++    Q Y  +E +           +KC D    R +
Sbjct  285  APRDVWSCDPPEHKEDETYTKLAPIFQTYIMKENDAGISSYLNKYTNKCSDTCTERTDTY  344

Query  62   EDNRYKFPQ---RQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTD  118
            +  R K+ Q   R+C+G +++C       K               S KRY  + L
Sbjct  345  DLCRTKYCQQHGRKCSGRVHDCKSSSAELK--------ICEADKNSTKRYNSITLTGKQA  396

Query  119  KKKYQYAIIEGSNYTYRYVYGNKEHSDHTT-VNFDTWIRGFVRCDICRCLCDESKDSESI  177
            +                 V G+       T +N    I    RC+ C C+CD+S ++E
Sbjct  397  E-----------------VLGSIVPCARNTAINLREKIHYIYRCEYCFCICDDS-NTEID  438

Query  178  RSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS  237
            R FSL  V SD   N VV G++ A K+ I + +I Q  LL  G  ++    WK+ D
Sbjct  439  RYFSLRNVTSDANKNYVVTGIRFAKKNNIIHIQIEQGQLLKNGNVNESTAHWKAVDNFTI  498

Query  238  SA-----SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMS  292
                    +D   +  ++RS+ + +  S L++ V+TG++F  +DG+L L +++T  D
Sbjct  499  YDKMILIDRDYHNITWSQRSVDLDTVDSPLNN-VITGLRFQNVDGHLKLEIQVTPIDFYR  557

Query  293  GKLMNDNNQWIYNSLSKQKLNNGGVANRPIKFTDNH-ASEKLDLGLVEFRQSDLKDDVGQ  351
            G+L  D + WI NS S        V N  +  +      +      ++F  S +K DVGQ
Sbjct  558  GELKIDESTWISNSKSVHGRIELNVENLDLGTSTYVNLPDSSTNQFLKFTTSSMKTDVGQ  617

Query  352  STVPFIDLRSIV-STGTALIGVGLQFRTRDTSGGFIAPVVVTHP  394
            + VPFIDL+ +  S    L G GL  R    S GF++   +T+
Sbjct  618  TVVPFIDLQDVTSSPPVLLSGAGLIHRGDKQSAGFLSFKAITYN  661


>☐ gb|EFA08551.1|  hypothetical protein TcasGA2_TC006206 [Tribolium castaneum]
Length=447

 Score =  397 bits (1019),  Expect = 2e-108, Method: Composition-based stats.
 Identities = 103/405 (25%), Positives = 168/405 (41%), Gaps = 48/405 (11%)

Query  8    FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANF---GSCNSKCEDYAAP-RHNFKED  63
             SR +   DP++HI+ ET+ EIT F   +   +AN    G C  +C +Y    +    E
```

```
Sbjct  66   ASREVWRCDPRKHIRNETFVEITNFLHGFILNKANLDSRGQCRDECSEYKMVLKPKNCEK  125

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
            N +   Q  C   I +C    +     V    +      + ++RY YV+     +
Sbjct  126  NYWCKKQTTCP-KIIDCQ--YVDSSMTVCSNIM-----AKHNRRYDYVQFADGKN-----  172

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFSL  182
                          +G K+  +  +   +T+       C  C CLCDE     S R  +L
Sbjct  173  --------------FGAKKSCNGDSDYLETYKFWLIWSCSYCFCLCDEGPKYFSDRYINL  218

Query  183  NAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LS  237
                S+  DNKVV G++    ++I + ++++  LL  G  D   + W   +
Sbjct  219  RLAMSNISDNKVVTGLRFVKHNQIIHLQVQEGKLLEHGTIDPDSVQWVPPENYKRTDRKI  278

Query  238  SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMN  297
               +D   +   KRSI +    +  +  VVTGV+F +I   L L + +T F+  +GKL N
Sbjct  279  HEGQDYQVITWEKRSIELTEIMA-RNGSVVTGVRFKKIGSRLELEIMITPFNFTTGKLRN  337

Query  298  --DNNQWIYNSLSKQKLNN------GGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
              D +     + +    +           + P + T    +       +EF  +D+  D
Sbjct  338  YSDQSSIFVEAPNIHLRHRFGKKITLSSPDVPTRSTAPS-THYTTGKYIEFTTTDVYKDA  396

Query  350  GQSTVPFIDLRSI-VSTGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             Q+TVPF D + +       L GVGL  R+R  SGGFIAP + T+
Sbjct  397  AQTTVPFFDSQQVYNKNPVPLSGVGLFHRSRRGSGGFIAPKIFTY  441


>☐ ref|XP_969565.1| UG PREDICTED: similar to lethal (1) G0193 CG2206-PA [Tribolium castaneum]
Length=396

 GENE ID: 658059 LOC658059 | similar to CG2206-PA, isoform A
[Tribolium castaneum]

 Score =  392 bits (1007),  Expect = 4e-107, Method: Composition-based stats.
 Identities = 106/401 (26%), Positives = 171/401 (42%), Gaps = 43/401 (10%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANF---GSCNSKCEDYAAPR-HNFKE  62
              SR L   DP +H++ +TY  I  F       +A+    G C ++C+ Y   +
Sbjct  3    RASRELWKCDPPKHVENQTYLRIGNFLHACIFNKAHLDPLGKCQNECDAYKMTKPQENCR  62

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
            +N++   Q      I NC    +     V   +     + ++++RY YVEL
Sbjct  63   ENKWCNKQLP-ASRIINCQ--YLDSSMTVCSNK----NTRKNNRRYQYVELG--------  107

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFS  181
                          V G K+       +  T+I     RC  C CLCDE +   +    +
Sbjct  108  -----------NGKVLGKKQSCRGYKTDLSTYIYWIVWRCSYCFCLCDE-QSRGTDVYIN  155

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSA--  239
            L    ++  DN+VV G++    ++IF+ +I++  LL  GK D   + W   D
Sbjct  156  LRLSLANMNDNRVVTGLRFVKHNEIFHLQIQEGKLLKRGKIDPDSVRWVPVDNYKHFDTI  215

Query  240  -SKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMND  298
              KD  K+   KRS  + +  +  +  VVTGV+F  I  +L L V  T F+  +GKL+
Sbjct  216  EGKDFHKITWEKRSFELSTIMAQ-NGSVVTGVRFRLIRNSLDLEVLTTPFNFTTGKLLPS  274

Query  299  NNQWIYNSL-----SKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQST  353
             + +            ++       + P   T+ H +       +EFR +DL  D+ Q+T
Sbjct  275  LSTFKGTRNLNLRHKFKQQVKLSKPDVPTGSTEPH-TLITTGQYIEFRNTDLYKDLAQTT  333

Query  354  VPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            VPF D + + S     L GV L  + R+ SGGFIAP + T+
Sbjct  334  VPFFDSQPVNSKNPVPLSGVSLIHKGREGSGGFIAPKIFTY  374


>☐ ref|XP_310721.4| UG AGAP000385-PA [Anopheles gambiae str. PEST]
 gb|EAA06714.4| G AGAP000385-PA [Anopheles gambiae str. PEST]
Length=667

 GENE ID: 1271866 AgaP_AGAP000385 | AGAP000385-PA [Anopheles gambiae str. PEST]
(10 or fewer PubMed links)
```

```
 Score =  389 bits (1000),  Expect = 2e-106, Method: Composition-based stats.
 Identities = 85/421 (20%), Positives = 150/421 (35%), Gaps = 74/421 (17%)

Query  9    SRTLDNVDPKE--HIKGETYEEITRFNQVYFTQEANF---GSCNSKCEDYAAP-RHNFKE  62
            S      DP    H +  TY + TR  Q Y   E +     +C   C  Y      +
Sbjct  254  SGEYWRCDPDRGAHRENVTYVQFTRLVQGYVENEVDMTTDNTCRETCSHYKYGHEQHQCY  313

Query  63   DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
               Y   Q +C G IY+C        E +       P +P   +RY Y+E
Sbjct  314  KELYCSQQPKCAGRIYDC--------EYIDSDMWICPAAPGGTRRYEYIE----------  355

Query  123  QYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDIC--RCLCDESKDSESIRSF  180
                     Y    + G K+     T   D+       C  C  R + ++  +      +
Sbjct  356  ---------YENGEMRGRKQSCPRGTTKVDS-------CVFCVFRAVFEQLTNGT----Y  395

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSL----  236
            + + + S      VV G++   + ++ +  ++Q  LLP G+     L W    P
Sbjct  396  TKSHICSSASLRSVVTGLRFVKRRQMVHLIVQQGLLLPGGEIANSTLEWVVPKPFTHTDK  455

Query  237  -SSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
                  D   L  ++R+I +         +VVTGV+F  +   L L +++T+ +  +GKL
Sbjct  456  GVRDGHDYHTLSYDRRAIDLDDVHVPP-GYVVTGVQFRVLGARLNLLLRMTEMNFTAGKL  514

Query  296  MN-DNNQWIYNSLSK--------------------QKLNNGGVANRPIKFTDNHASEKLD  334
            +  D + W  +  ++                    +        + P
Sbjct  515  VGLDASIWSGSLETERAHLCVKQFITFHALSLSPSRTEIPLKYMDVPTLAPAKSVPMSTQ  574

Query  335  LGLVEFRQSDLKDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
               + F  +  + D  Q+T+PF+D + +     T L G GL ++     GGF+AP ++T+
Sbjct  575  NQYLRFGPTGGRADAAQTTIPFLDAQPVSPVRPTPLAGAGLMYKGVPKYGGFVAPKIMTY  634

Query  394  P  394

Sbjct  635  N  635


>☐ ref|XP_001603377.1| PREDICTED: similar to CG2206-PA [Nasonia vitripennis]
Length=671

 GENE ID: 100119643 LOC100119643 | similar to CG2206-PA [Nasonia vitripennis]

 Score =  385 bits (990),  Expect = 3e-105, Method: Composition-based stats.
 Identities = 103/408 (25%), Positives = 183/408 (44%), Gaps = 46/408 (11%)

Query  6    LGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANF---------GSCNSKCEDYAAP  56
            +   R     DP EH +G TY +++   Q Y   E +            C   C+D+
Sbjct  284  IKAPRDFWYCDPDEHAEGSTYTKLSPIFQAYIMNEKDSGLTFHHKKRNKCKGSCKDHNET  343

Query  57   RHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVN  116
             ++       +  +++C+G + +C                    + ES +RY ++E+  +
Sbjct  344  -NDLCRSKHCRNYEKKCSGLVRDCTSSAEELN--------ICYANEESSRRYEFIEILSS  394

Query  117  TDKKKYQYAIIEGSNYTYRYVYGNKEHS--DHTTVNFDTWIRGFVRCDICRCLCDESKDS  174
                              R  YG  +    ++T    + + + F  C+ C C C++  ++
Sbjct  395  G----------------NRKFYGKPKSCSGNNTLQVKNRFNKLFYECEYCVCTCEDI-NT  437

Query  175  ESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDP  234
            +  R FSL  V SD++ + VV GV+ A K+ IF+ +I++  LL  G  D+  + WK  D
Sbjct  438  KINRYFSLKNVTSDSQYSYVVTGVRFAKKNNIFHIQIQRGKLLVNGNIDKSSVDWKPVDN  497

Query  235  SLSSAS-----KDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFD  289
               S       +D +K+  N+R++++ +  + L++ V+ G++F     NL   +++T F+
Sbjct  498  FTVSDKHFVDHRDYYKISWNQRTVYLDTLDAPLNN-VLIGLRFQVDGSNLKFEIQMTPFN  556

Query  290  VMSGKLMNDN-NQWIYNSLSKQKLNNGGVANRPIKFTDNH-ASEKLDLGLVEFRQSDLKD  347
              SG+L   N  +WI NS S++        +     T N    E     L +F +S  +
Sbjct  557  FSSGELKKKNETRWISNSSSERTEIYLKNRDLGTSTTQNSYHYEAGSNQLFKFTRSSFES  616

Query  348  DVGQSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTHP  394
            D GQ+ VPFIDL+ I ST    L G GL  R    S GF++  ++T+
Sbjct  617  DAGQNVVPFIDLQEIESTPFVPLSGAGLILRNTKESAGFLSFKLLTYN  664
```

```
>☐ ref|XP_001606530.1| G PREDICTED: similar to CG2206-PA [Nasonia vitripennis]
Length=880

 GENE ID: 100122920 LOC100122920 | similar to CG2206-PA [Nasonia vitripennis]

 Score =  378 bits (970),  Expect = 7e-103, Method: Composition-based stats.
 Identities = 85/404 (21%), Positives = 165/404 (40%), Gaps = 51/404 (12%)

Query  9    SRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANF---GSCNSKCEDYAAPRHNFKEDNR  65
            SR L   +P +HI+  TY E+    Q Y   E +      C ++C DY  P
Sbjct  280  SRELWRNNPVQHIEDVTYTELKHVFQKYIINEYDISPGDQCKNQCSDYRRPESKAPHRCA  339

Query  66   YK---FPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKY  122
            Y+        C G +++C  +       +       P   +S +RY Y++
Sbjct  340  YQGEICEITHCDGYLFDCQYK-------IDTEMSICPSGRDSSRRYDYID----------  382

Query  123  QYAIIEGSNYTYRYVYGN-KEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFS  181
                     Y     YG+ K       +     + G   CD C C+C ++      R F+
Sbjct  383  ---------YANGLRYGDSKHKCSMNPLTTTRKLSGLWWCDPCMCVCADTSKKSESR-FN  432

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASK  241
            L+ V +++ +N V+ G +L  ++++ + +I++  LL  G+ +   L W   + +  S
Sbjct  433  LDPVMANSAENYVLTGARLVKRNRMIHIQIQEGQLLQNGRINASTLRWSKINENPIS---  489

Query  242  DIFKLDVN-KRSIHIGSATSNLSDFVVTGVKFAEI------DGNLALGVKLTKFDVMSGK  294
                L    KR IH+    +   + V+TG+ F  I         L L          SG+
Sbjct  490  -YHTLQWEQKREIHLDDLMA-PKNHVLTGLSFTTIQVDSAYTKRLKLQAYAAPLYFHSGE  547

Query  295  LMNDNNQWI----YNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVG  350
            +++ +   +    ++  +          + P + +    ++      ++F  + L  DVG
Sbjct  548  VIDKHQMPLYPKDFSPTNATNEIKLDNLDVPTRTSVQTVTKSKQDQFLKFGPTGLDKDVG  607

Query  351  QSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTH  393
            QST+PFID++ + ++    + G G+  ++R  +GGF+A  + ++
Sbjct  608  QSTIPFIDIQEVTTSPAIPMSGAGIYHKSRSGNGGFLALRLFSY  651


 Score = 52.8 bits (125),  Expect = 7e-05, Method: Composition-based stats.
 Identities = 9/48 (18%), Positives = 18/48 (37%), Gaps = 1/48 (2%)

Query  8    FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAA  55
             S      DP  H++ +T+ ++    QV      +      +C +Y
Sbjct  815  SSTEFYRGDPNNHVEDKTFTQLKNVMQVIAMNMIDINRA-DRCYNYCD  861


>☐ gb|EFA08552.1|  hypothetical protein TcasGA2_TC006207 [Tribolium castaneum]
Length=607

 Score =  372 bits (954),  Expect = 6e-101, Method: Composition-based stats.
 Identities = 100/387 (25%), Positives = 163/387 (42%), Gaps = 43/387 (11%)

Query  21   IKGETYEEITRFNQVYFTQEANF---GSCNSKCEDYAAPR-HNFKEDNRYKFPQRQCTGT  76
             + +TY  I  F       +A+    G C ++C+ Y   +      +N++   Q
Sbjct  228  KQDQTYLRIGNFLHACIFNKAHLDPLGKCQNECDAYKMTKPQENCRENKWCNKQLP-ASR  286

Query  77   IYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRY  136
            I NC    +     V   +     + ++++RY YVEL
Sbjct  287  IINCQ--YLDSSMTVCSNK----NTRKNNRRYQYVELG-------------------NGK  321

Query  137  VYGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVV  195
            V G K+       +  T+I     RC  C CLCDE +   +    +L    ++  DN+VV
Sbjct  322  VLGKKQSCRGYKTDLSTYIYWIVWRCSYCFCLCDE-QSRGTDVYINLRLSLANMNDNRVV  380

Query  196  IGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSA---SKDIFKLDVNKRS  252
             G++    ++IF+ +I++  LL  GK D   + W   D          KD  K+   KRS
Sbjct  381  TGLRFVKHNEIFHLQIQEGKLLKRGKIDPDSVRWVPVDNYKHFDTIEGKDFHKITWEKRS  440

Query  253  IHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNNQWIYNSL-----  307
              + +  +  +  VVTGV+F  I  +L L V  T F+  +GKL+   + +
Sbjct  441  FELSTIMAQ-NGSVVTGVRFRLIRNSLDLEVLTTPFNFTTGKLLPSLSTFKGTRNLNLRH  499
```

```
Query  308  SKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVS-TG  366
              ++       + P   T+ H +       +EFR +DL  D+ Q+TVPF D + + S
Sbjct  500  KFKQQVKLSKPDVPTGSTEPH-TLITTGQYIEFRNTDLYKDLAQTTVPFFDSQPVNSKNP  558

Query  367  TALIGVGLQFRTRDTSGGFIAPVVVTH  393
              L GV L   + R+ SGGFIAP + T+
Sbjct  559  VPLSGVSLIHKGREGSGGFIAPKIFTY  585


>☐ gb|EFA12872.1|  hypothetical protein TcasGA2_TC010724 [Tribolium castaneum]
Length=1530

 Score =  366 bits (939),  Expect = 3e-99, Method: Composition-based stats.
 Identities = 97/396 (24%), Positives = 163/396 (41%), Gaps = 49/396 (12%)

Query  22   KGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFK-EDNRYKFPQRQCT---  74
            +G+TY EIT F   Y   + N      C ++C+ +          +  +   +  C+
Sbjct  41   EGKTYLEITNFLHGYIINKKNLNPYQRCENECDSFKMTTTQNCVYEKEWCKEEIPCSKII  100

Query  75   ------GTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIE  128
                   T++ C    I     +     K    P+ ++RY++++L   T
Sbjct  101  SCTSLHSTMHVCTAVSIRPLLGITTVSFKNL--PKDNRRYHHIKLSNGT-----------  147

Query  129  GSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNAVESD  188
                     +G K           T+      C  C CLC+E + S S R  +L   +++
Sbjct  148  --------FFGPKRQCTGNI-KIHTYSYRNSYCPYCFCLCEEGQVSYSERYVNLRITKAN  198

Query  189  TKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASK-----DI  243
              DN+VV G++    ++I +  I++  LL  G  D + + W   D             D
Sbjct  199  ISDNRVVTGLRFIKHNQIVHLHIQEGKLLEKGYVDPKTVQWVPPDNYKIDDRNIYEEEDY  258

Query  244  FKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNNQW-  302
              L   +RS+ +    +      VV GV+F +  G + L V  T FD  +GKL  DN  W
Sbjct  259  HTLSWEERSLELTEIVAE-EGSVVIGVRFKKTLGGVKLEVLTTTFDFDTGKL-EDNTTWN  316

Query  303  -IYNSLSKQKLN---NGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFID  358
               N+++   +N        + PI    +  + +  +   +  F  +D   D GQ+TVPF D
Sbjct  317  IFKNTITPNLINQEIKLDSPDIPI-HKNLASIDFSENTYINFTNTDYDKDAGQTTVPFFD  375

Query  359  LRSIVST-GTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             + + ST  T L G+GL  + R  SGGFI P ++T+
Sbjct  376  AQPVYSTAPTPLSGIGLFHKGRKGSGGFITPKILTY  411


>☐ gb|ACM40897.1|  putative cysteine endopeptidase [Culicoides nubeculosus]
Length=395

 Score =  365 bits (937),  Expect = 6e-99, Method: Composition-based stats.
 Identities = 197/391 (50%), Positives = 266/391 (68%), Gaps = 26/391 (6%)

Query  9    SRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAAPRHNFKEDNRYKF  68
            S  L N+DPK+H+KGETY E+TRFNQVYFTQEANFGSCNSKC DY+ P+HNF+E + Y+F
Sbjct  24   SSELYNLDPKQHVKGETYGEVTRFNQVYFTQEANFGSCNSKCSDYSKPQHNFRESDSYQF  83

Query  69   PQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIE  128
             QR CTGTI++C  E+  +   V  Y+++ P+S + D RYYYV                 +
Sbjct  84   DQRTCTGTIHSCNNEEREDNSHVDNYKVRVPQSSKVDARYYYV-------------FKQ  129

Query  129  GSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNAVESD  188
             +   ++YVYGN + S++ TV F++W+RGFV C ICRC+CDE  D  S+RSFSLN++ESD
Sbjct  130  FAPKMFQYVYGNIDASNYRTVTFNSWVRGFVMCAICRCICDEPNDDRSVRSFSLNSLESD  189

Query  189  TKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASKDIFKLDV  248
            T  NKV+ GVQ+  ++ +F  KI QS L   GK ++   SWKSD+P+      + F L
Sbjct  190  TLQNKVITGVQIVQQNNVFCLKIHQSCLYANGKINKNGGSWKSDNPNN-----NTFNLAF  244

Query  249  NKRSIHIGSATSNLSD------FVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNNQW  302
            NK +IH+GS+  + +D      +VVTGVKFA+I+GNLALGVK T F++ +GKL+    +W
Sbjct  245  NKHNIHMGSSILSTADPKKTREYVVTGVKFADIEGNLALGVKFTLFEISTGKLL-PEQKW  303

Query  303  IYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSI  362
```

```
            ++NS SK  LN     +RPI++TDN  + +  LGL++F  S L DDVGQSTVPF+DL+ I
Sbjct  304  VFNSFSKNLLNYHKHGHRPIEYTDNIPTAQTKLGLIQFEHSHLHDDVGQSTVPFVDLQPI  363

Query  363  VSTGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
             S  TAL GVGL  R R +SGGFIAP + T+
Sbjct  364  NSNETALSGVGLHLRGRKSSGGFIAPTLTTY  394
```

>☐ ref|XP_001851115.1| UG conserved hypothetical protein [Culex quinquefasciatus]
gb|EDS33068.1| G conserved hypothetical protein [Culex quinquefasciatus]
Length=548

GENE ID: 6041925 CpipJ_CPIJ009488 | hypothetical protein
[Culex quinquefasciatus]

Score = 360 bits (924), Expect = 2e-97, Method: Composition-based stats.
Identities = 86/321 (26%), Positives = 138/321 (42%), Gaps = 42/321 (13%)

```
Query  8    FSRTLDNVDPK--EHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA-APRHNFK  61
             SR     DP+  +  +GET+ +ITR  Q Y   E +  +   C   C  Y+   R
Sbjct  257  ASRDYWRCDPERSQQKEGETFVQITRLLQGYVENEVDMNTDNTCKENCAYYSWGARQEQC  316

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
              + Y   Q +C+G IY+C        E +       P +  S++RY Y+E
Sbjct  317  FKDLYCSKQPKCSGKIYDC--------EYLDSDMWICPAATSSNRRYEYIE---------  359

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRSF  180
                      Y    V G K+H    +   D+W R  F  C  C CLCD++    S R
Sbjct  360  ----------YENGMVLGQKKHCARGSTKVDSWWRWLFWHCSYCFCLCDDA-GKNSDRYI  408

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----  235
            ++    SD  DNKVV G++   K++I +  ++Q  LLP G+ D   L W   +
Sbjct  409  NMRESVSDVMDNKVVTGMRFIKKNRIIHLIVQQGRLLPRGQIDNATLEWVQPEDYNVLSP  468

Query  236  LSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
                 +D   L+ + R I        L  +V+TGV+F  +  +L L +++T+ D  +G L
Sbjct  469  NVRNGRDYHTLNRDNR-ILDLDDLDVLQGYVLTGVRFRLLGSHLNLEIRMTEMDFATGHL  527

Query  296  MN-DNNQWIYNSLSKQKLNNG  315
            M+ D + WI +  ++   +
Sbjct  528  MDPDKSIWIGSDKTEHSDDKR  548
```

>☐ ref|XP_001606517.1| UG PREDICTED: similar to CG2206-PA [Nasonia vitripennis]
Length=662

GENE ID: 100122910 LOC100122910 | similar to CG2206-PA [Nasonia vitripennis]

Score = 352 bits (902), Expect = 6e-95, Method: Composition-based stats.
Identities = 88/400 (22%), Positives = 156/400 (39%), Gaps = 41/400 (10%)

```
Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANF---GSCNSKCEDYAAPRHNFKED  63
            G SR L   +P EH +  TY E+ R  Q +   E +      C   C +Y  P
Sbjct  277  GASRELMRFNPSEHKEDVTYTELKRVFQEFIVNEYDVKPSDGCKKTCSNYRIPDGKALNG  336

Query  64   NRYK---FPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK  120
              Y       + C G +Y+C  +  T     I           SD+ Y +V+
Sbjct  337  CFYADEVCSVKTCDGYLYDCRKKVDTTMNVCISPMF-------SDRNYEFVDFANGAG--  387

Query  121  KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSF  180
               Y    G  +              T     + +     CD C C+C +   S+S R F
Sbjct  388  --HYGHATGKCWAD------------TIKTAASHVSIPWSCDSCMCVCAD-TSSQSQRFF  432

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSAS  240
            +LN   +   +N VV G +   K+++ + +I++  LLP G  +     W   + + ++
Sbjct  433  NLNPSMATIAENYVVTGARFVKKNRVVHIQIQEGQLLPNGMINSSTTRWAPINENPTT--  490

Query  241  KDIFKLDVNK-RSIHIGSATSNLSDFVVTGVKFAEIDGN-----LALGVKLTKFDVMSGK  294
             D   L     R +++    +     V+TG+ F+    +     L L +        +G+
Sbjct  491  LDYHTLAWENGREVNLDDLHAPKKS-VLTGLAFSTESTSNFQRRLHLKIFSAPVSYETGR  549
```

```
Query  295  LMNDNNQWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTV  354
            L  ++ +   N    +        + P + +   A        V+F  + L+ DVGQST+
Sbjct  550  LWINDQKEKVNEP-AKDEITLKDVDIPTRTSVVTADRSKPDSFVKFGPTGLEQDVGQSTI  608

Query  355  PFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTH  393
            PFIDL+ + S+    + G G+  + R  +GGF+   + ++
Sbjct  609  PFIDLQEVSSSPQIPMSGAGIYLKRRSNNGGFLTLKLFSY  648


>☐ gb|ACZ28336.1|  hypothetical conserved protein [Simulium nigrimanum]
Length=408

 Score =  350 bits (897),  Expect = 2e-94, Method: Composition-based stats.
 Identities = 97/404 (24%), Positives = 161/404 (39%), Gaps = 43/404 (10%)

Query  3    LPSLGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANF-GSCNSKCEDYAAPRHNFK  61
            LP L   R+    D     +  T+ E+ +        EA+  G+C++ C D+  PR
Sbjct  27   LPRL--ERSYWKCDADNWQRDVTFYELNKLMYGLVEYEADVTGACSTGCYDFQKPRG---  81

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
                 ++P  +C G I  C         P   ++    E+  + + Y Y+   +
Sbjct  82   ITTLTQWPDYKCPGYIEACEAS------PKETFEFCFNENAMAPETYSYISSDL------  129

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTV-NFDTWIRGFVRCDICRCLCDESKDSESIRSF  180
                             G KE      V    ++  G  +C+ C CLC  S+ +E+ R
Sbjct  130  --------------MRRGTKEQCRDEKVHKVSSYTSGVFKCEYCACLC-HSEGAEAHRYL  174

Query  181  SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD-----PS  235
            SL    S+  +N+V++GV+    +++ Y +  ++PLLP    D  +  W   D     P
Sbjct  175  SLLPQVSNIDENEVIVGVKFVKHNRVLYLQTMKAPLLPFASVDTDQAVWNELDELGLVPP  234

Query  236  LSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
                   ++   D   R I +       +  V+TGVKF   DG   L ++ T  D  +G+L
Sbjct  235  DYHNFSGLYTFDHKHREILMKELEV-PAGTVLTGVKFIVKDGVPDLSIRYTPVDWTTGEL  293

Query  296  MNDNNQWIYNSLSKQKLNNGGVAN-RPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTV  354
                + WI  +      N+    N  P +       +        F  S ++ D GQ  +
Sbjct  294  NPAASLWITEAHDAYDTNDMLYQNRLPDQCKTPSFIDTASGQKFRFSTSAMELDGGQHVL  353

Query  355  PFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVT-HPTY  396
            P+ D + +V  +   L GVG+  +  D  GGFI+PVV T H  Y
Sbjct  354  PYFDAQPVVPVSMVPLSGVGITHKGDDRCGGFISPVVFTVHEDY  397


>☐ ref|XP_001601809.1| G PREDICTED: similar to GA15301-PA [Nasonia vitripennis]
Length=666

 GENE ID: 100117628 LOC100117628 | similar to GA15301-PA [Nasonia vitripennis]

 Score =  347 bits (890),  Expect = 1e-93, Method: Composition-based stats.
 Identities = 96/430 (22%), Positives = 175/430 (40%), Gaps = 78/430 (18%)

Query  8    FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQE-----ANFGSCNSKCEDYAAPRHNFKE  62
             SR +   D ++H K  TY ++    Q +F  E     +N+ +   +C      R   K
Sbjct  262  ASREVWRCDTEKHEKDVTYSKLQEVFQGFFVYEPDTTRSNWHAAVFRCSALKDDRLRCKS  321

Query  63   D-NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
               + +  +R+CTG + +C          +          P SD+RY +++    T
Sbjct  322  APCKAQPEERRCTGHLRDCQS--------IAHNMQLCHSPPGSDRRYDWMKTGAGT----  369

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFS  181
                           +YG      +T   +D         + C C C ES+  ++ R FS
Sbjct  370  ---------------LYGRDGPCANTQGIWDIPPYLI---ENCLCNC-ESEQRDADRYFS  410

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSD---------  232
            L  V SD   NKVV GV+L   + +F+ +I+Q  L   G      L WK
Sbjct  411  LLDVTSDVDQNKVVTGVRLRKHNHVFHIQIQQGTLEANGTI-SGTLEWKEIKEISSSDLL  469

Query  233  --------------------------DPSLSSASKDIFKLDVNKRSIHIGSATSNLSDFV  266
                                      D +    +  ++ +  ++ ++++    S   + +
Sbjct  470  SDKVSKHFYKVIANPHENFDKEFKYDDWTEFYKNDLVYAVTYDQNTVYLDDLDSGP-NAL  528
```

```
Query  267  VTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNNQWIYNSLSKQKLNN--GGVANRPIKF  324
            +TG++F +++ +L L ++ T FD  +GKL  D + W  N+ ++   +       + P +
Sbjct  529  LTGLRFQQVNDDLRLEIQSTPFDYTTGKLKTDQSSWTANTNTRPTRHEVFLTDPDDPTRT  588

Query  325  TDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRS-IVSTGTALIGVGLQFRTRDTSG  383
             ++      +   V FR +DL+ D+GQ TVPF+DL+  + +    + G GL   R R  SG
Sbjct  589  KESDNF-SKNKQFVRFRMTDLRHDLGQYTVPFLDLQDAVTTPALPISGAGLFHRGRLHSG  647

Query  384  GFIAPVVVTH  393
            GF+   ++ +
Sbjct  648  GFLGFKLIHY  657


>☐ ref|XP_001600223.1| G PREDICTED: similar to GA15301-PA [Nasonia vitripennis]
Length=497

 GENE ID: 100115520 LOC100115520 | similar to GA15301-PA [Nasonia vitripennis]

 Score =  343 bits (880),  Expect = 2e-92, Method: Composition-based stats.
 Identities = 92/426 (21%), Positives = 158/426 (37%), Gaps = 54/426 (12%)

Query  8    FSRTLDNVDPKEHIKGETYEEITRFNQVYFT---QEANFGSCNSKCEDYAAPRHNFK-ED  63
             SR +   DPK HI+  TY E+    Q Y         N  +C   C ++
Sbjct  91   ASRDVWRCDPKSHIEDVTYTELKEVFQKYVVNDEDMNNENNCKKTCGEFKYATVTNCANK  150

Query  64   NRY------------KFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYV  111
              Y               QR+C G +++C         E  I      P +  S +RY ++
Sbjct  151  QEYYDTEVQYYEYSICDQQRKCNGALFDCEDYFPDYNELWI-----CPSTLNSSRRYEFM  205

Query  112  ELPVNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVN-----FDTWIRGFVRCDICRC  166
            E                G         G+K+  D            ++   +  C++C C
Sbjct  206  EHGSVDPVIFKHDVRQYG-------PKGDKKKCDRHITYDYREFGQSFKDAWRYCEVCLC  258

Query  167  LCDESKDSESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRK  226
             C + K   S +  ++  V +D  ++ VV G +L  K+ I + +I Q  +L  G+ D
Sbjct  259  TCHDFKSPTSEQQINVRPVVAD--NDYVVTGARLVKKNHILHVQIEQRKILADGRIDNST  316

Query  227  LSWKSDDPSLSSASKDIFKLDVN-KRSIHIGSATSNLSDFVVTGVKFAEI----DGNLAL  281
            L WK+          D+     N  R+I++   T    + V+ G+K   +       ++
Sbjct  317  LHWKTIAEP--RNVYDLRSFSWNGFRTIYL-EVTEVPKNHVLVGLKLTNLKNGDGQMMSF  373

Query  282  GVKLTKFDVMSGKLMNDNNQWIYNSLSKQKLNNGG------VANRPIKFTDNHASEKLDL  335
                  +D  +G+L  D +   Y+   K  L +          + P         +
Sbjct  374  KTLSVPYDYSTGELTYDGSY--YDEPEKLDLRDRQAEVSVENCDLPQNTHFPSLPQPDSK  431

Query  336  GLVEFRQSDLKDDVGQSTVPFIDLRSI-VSTGTALIGVGLQFRTRDTSGGFIAPVVVT--  392
              ++F   D K D+GQS +PFID++   +     L G+G   + R+  GGF+   +VT
Sbjct  432  KTIKFTTLDFKKDLGQSMIPFIDIQEFSIRPAAPLAGIGFYHKGREGYGGFVGLKLVTFD  491

Query  393  HPTYTG  398
            H  + G
Sbjct  492  HSPHIG  497


>☐ ref|XP_001602724.1| G PREDICTED: similar to CG2206-PA [Nasonia vitripennis]
Length=894

 GENE ID: 100118844 LOC100118844 | venom protein W [Nasonia vitripennis]

 Score =  340 bits (871),  Expect = 2e-91, Method: Composition-based stats.
 Identities = 90/399 (22%), Positives = 170/399 (42%), Gaps = 41/399 (10%)

Query  7    GFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANF---GSCNSKCEDYAAP--RHNFK  61
               R L       H+   T+ E+    Q Y   EA+F   GSC   C+D+       +
Sbjct  509  RAQRNLGEQSAPTHV---TFAELKNVFQSYVINEADFKKEGSCGHTCKDFKLALAGLDCY  565

Query  62   EDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKK  121
             D ++    ++C+GT+++C+ +   E            E   S +RY Y+++
Sbjct  566  YDKQFCKKHQRCSGTLHDCVHQFDEE-------MFVCLEEKNSTRRYQYIDVG-------  611

Query  122  YQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFS  181
            YQ   +E  +   + +  ++ +S H              CD C CLCD+    ++ R FS
```

```
Sbjct  612  YQRYGLEDKSCPSKAIELDRHYSAHN----------MAYCDSCMCLCDDDAHPKTERYFS  661

Query  182  LNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASK  241
            L +  +  K+N V+ G +L  +++  + +I+Q  LLP G  ++  + W      + +  +
Sbjct  662  LISSIAAVKENYVLTGARLVKRERTIHVQIQQGQLLPYGHVNESTVEWVPIKADVLAKDE  721

Query  242  DIFKLDV-NKRSIHIGSATSNLSDFVVTGVKFAEIDGN----LALGVKLTKFDVMSGKLM  296
             ++K+     R++ +       ++ VV G+ F   + N    L L       +  +G L
Sbjct  722  -VYKIKWLGSRTVFLDDIEVPRNN-VVVGLGFKAPEQNNIKVLRLVAYSAPMNYSAGVLT  779

Query  297  NDNNQWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPF  356
            ++  +  ++S            + P + +     +  +   V+FR + L  DVGQST+PF
Sbjct  780  SETYE-SHSSEESSDEVVLEDLDIPTRRSSQIVQKSANHQYVKFRTTSLLTDVGQSTIPF  838

Query  357  IDLRSIVSTGT-ALIGVGLQFRTRDTSGGFIAPVVVTHP  394
            ID +S+ +     + G G+  R     GGF+A  + T
Sbjct  839  IDSQSVTTNPASPMSGAGIYHRRSSGDGGFLALKLFTFN  877


>☐ ref|XP_001810077.1| PREDICTED: similar to GA15301-PA [Tribolium castaneum]
Length=426

 GENE ID: 100142514 LOC100142514 | similar to GA15301-PA [Tribolium castaneum]

 Score =  337 bits (865),  Expect = 1e-90, Method: Composition-based stats.
 Identities = 89/391 (22%), Positives = 153/391 (39%), Gaps = 79/391 (20%)

Query  24   ETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKEDNR---YKFPQRQCTGTI  77
            +TY EIT F   Y   + N      C ++C+ +        +D+    + F +  C+  I
Sbjct  82   KTYLEITNFLHGYIINKKNLDPKKQCVNECDSFKMTAIQNCDDDIDEPWCFKESPCS-RI  140

Query  78   YNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRYV  137
             +C          +
Sbjct  141  ISCT--------YIDSSMDVCLA-------------------------------------  155

Query  138  YGNKEHSDHTTVNFDTWIRGF-VRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVI  196
                       +N  T+      +C  C CLCDE + S S R  +L    S+  DN+VV
Sbjct  156  -----------INLATYSESLIFKCSYCFCLCDEGQGSYSDRYINLRISMSNISDNRVVT  204

Query  197  GVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LSSASKDIFKLDVNKR  251
            G++    ++I + ++++  LL  G  +   + W   +            +D   +   +R
Sbjct  205  GLRFIKHNRIVHLQVQEGKLLEKGNINPTTVQWVPPEDYKITDRDIYKGQDYHTISWEER  264

Query  252  SIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL-MNDNNQW-------I  303
            +I +    +  +  VVTGV+F    G+L L V  T FD  +GKL  N ++ W
Sbjct  265  TIELTKVIAK-AGSVVTGVRFKRTFGSLNLQVLTTAFDFNTGKLEKNSDSDWGIFESAPY  323

Query  304  YNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIV  363
            +N   K +L      +     T+   ++  +   +EF  +D + D GQ+T+PF D +++
Sbjct  324  FNIRHKFRLKVNLNPSDVPVRTNLDFTDFSENMYIEFANTDYRKDAGQTTIPFFDAQAVN  383

Query  364  S-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            S T T L G+GL  + R   GGFIAP ++T+
Sbjct  384  STTPTPLSGIGLFHKGRKGYGGFIAPKIITY  414


>☐ gb|ACE75210.1|  conserved hypothetical protein [Glyptapanteles flavicoxis]
Length=645

 Score =  330 bits (846),  Expect = 2e-88, Method: Composition-based stats.
 Identities = 90/427 (21%), Positives = 163/427 (38%), Gaps = 77/427 (18%)

Query  12   LDNVDPKEHIKGETYEEITRFNQVYFTQE-----ANFGSCNSKCEDYAAPRHNFKE----  62
            L   DP + I+GET+ E     Q     E      +   C+++C+     R
Sbjct  217  LQRCDPPKDIRGETFSEFQGLFQGIIINELQTDSIDTTHCHNQCDAVDYSRLVRCYSFDA  276

Query  63   --DNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK  120
                     ++ C G +  C           I +         S +R+ +     +   +
Sbjct  277  GNRTITHCHKKPCNGLLSACQ---------WIGFIKACEMYENSSRRFLW----SDDVNE  323

Query  121  KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESK-DSESIRS  179
               Y    G   T+ YV                W     RC +CRC+C E + +S ++R+
```

```
Sbjct  324  DSSYRKCPGRLVTFDYV----------------WQNDVSRCSVCRCICVEQEGNSTAMRT  367

Query  180  FSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDP-----  234
             SL    +D   N V+ GV+ A KD+IF+ +I QS L P G+     + WK
Sbjct  368  ISLIPQLADIHRNMVMTGVRFAQKDRIFHLQIEQSELGPFGEIVPGTVEWKELGDFQYDS  427

Query  235  -----SLSSAS---------KDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAE--IDGN  278
                                 D   + +++R+I++    +N    V+ GV+F    +D
Sbjct  428  TGEGSFSIKKGDEFVKLTEFVDFSFVALDQRTINLDELFAND-GQVLIGVRFIYNGMDDA  486

Query  279  LALGVKLTKFDVMSGKLMNDNN---QWIYNSLSKQKLNNGGVA---------NRPIKFTD  326
              L VK    D+ +GKL + N    +WI +  ++++  N             + P+
Sbjct  487  FELQVKSWPVDIRTGKLADGNPSDVEWIGSENAQKRSENYDGPRSTFEIGETDEPLSTNK  546

Query  327  NHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVSTG--TALIGVGLQFRTRDTSGG  384
             +    +    +  R ++ +DD+ Q TVP++DL+ +  +     L G+ + +R  +  GG
Sbjct  547  LNIPLSVPNQKLMIRATNFEDDLHQHTVPYLDLQPVTYSTAKIPLSGLEIIYRGVEGFGG  606

Query  385  FIAPVVV  391
            F+   +
Sbjct  607  FLGFRIF  613


>☐ ref|XP_001601835.1| UG PREDICTED: similar to ef hand protein [Nasonia vitripennis]
Length=796

 GENE ID: 100117668 LOC100117668 | similar to ef hand protein
[Nasonia vitripennis]

 Score =  322 bits (824),  Expect = 7e-86, Method: Composition-based stats.
 Identities = 91/392 (23%), Positives = 164/392 (41%), Gaps = 54/392 (13%)

Query  9    SRTLDNVDPKEHIKGETYEEITRFNQVYFTQEAN-----FGSCNSKCEDYAAPRHNFKED  63
             R L   DP  H K  TY  +    Q  F  E +     + +   +CE   + R    +
Sbjct  141  DRKLLRADPDSHQKDITYSRLYPVLQGVFIYEPDTTAGGWYTFWHQCESLKSTRLKVPDH  200

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
                   + C G + +C          +       P +   D+ Y +
Sbjct  201  -------KSCAGVLRDCR--------YIDSSIDICPST-TDDRIYSW-------------  231

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLN  183
                  +      VYG+K +S+      +T+  G +  D+C C C ES   +  R F L
Sbjct  232  ------AVTKSNDVYGHK-NSECPVKLQNTY--GVLTYDVCLCTC-ESAYVQKDRYFCLR  281

Query  184  AVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDD---PSLSSAS  240
             V SD+ +NKV+ G +L  K  + Y +I+Q  L   G        W+      P     +
Sbjct  282  NVVSDSDNNKVITGARLVKKGNVLYIQIQQGTLGANGSIV-DVPQWQPVKEESPRNLHNN  340

Query  241  KDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNN  300
            K++++++ ++ ++ +    S  + + +TG++F + + +L L  ++T F+  +G L   N+
Sbjct  341  KNVYQVNYDQNTVFLDVLKSQPA-YALTGLRFQKAENSLRLEAQITNFNYETGNLRLTNS  399

Query  301  QWIYNSL--SKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFID  358
             W+ N      ++       + PI   +++     +   V FR ++L+ D+GQ+ VPFID
Sbjct  400  -WLSNPNGLETRERLKLVNPDVPIIMPEHNKI-THNNRYVSFRMTNLRKDLGQTIVPFID  457

Query  359  LRSIVS-TGTALIGVGLQFRTRDTSGGFIAPV  389
             + IV+     L G GL  + R  SGGF+
Sbjct  458  TQEIVTNPPMPLAGAGLYHKGRKHSGGFLGQK  489


>☐ gb|ABY19389.1|  venom protein 2 [Microctonus hyperodae]
Length=679

 Score =  314 bits (805),  Expect = 1e-83, Method: Composition-based stats.
 Identities = 106/422 (25%), Positives = 164/422 (38%), Gaps = 81/422 (19%)

Query  8    FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFG--SCNSKCEDYA------APRHN  59
             S TLD  DPK   +GE + E+ +  Q++   E +    SC+  C +          R+
Sbjct  268  ASLTLDYCDPKNFTRGENFVELEKSIQMFVGNEVDINGNSCSDVCSNIDSSGIRSDARNC  327

Query  60   FKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDK  119
```

```
             K +     P + C G IY+CI                      S +RY +VE
Sbjct  328  DKNNGNSWCPLQSCYGRIYDCISASTAS---------VCELGKSSLRRYDWVE-------  371

Query  120  KKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLC-DESKDSESIR  178
                       ++   + YGN    +    N   +  G   CD+C C C DE  +S + R
Sbjct  372  -----------DHYNNHNYGNTNSCNGRIRNPSYYRSGLTSCDVCICTCADEIGESPANR  420

Query  179  SFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS-  237
            + +L    S+ + N V+ GV+    DKI + +I++  L+  G+       W   +
Sbjct  421  TINLRPQLSNIEANMVITGVKFISSDKIIHIQIKEGELMSGGRIKPETEHWIEPENFEYD  480

Query  238  ------------------------SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFA  273
                                    D   +   +R I++    +  S +VVTGVKF
Sbjct  481  PKGILKSRKVGSFFLSTGNKMTLMKHGMDFTFVTFFQREINLDDVKA-PSGYVVTGVKFV  539

Query  274  EIDG--------NLALGVKLTKFDVMSGKLMNDNNQWIYNSLSKQKLNNGGV--------  317
                         L L + +TKFD ++G LM D++ WI       +  N G+
Sbjct  540  REGNDTDGLSRRRLQLAIYVTKFDFLNGVLMKDSSYWISQENMPNRPENYGLRRNEVELS  599

Query  318  -ANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVST-GTALIGVGLQ  375
              + P K  D H         V+F  SD  +D GQ+T+PF D +S+ ST  TAL G+GL
Sbjct  600  APDDPTKGKD-HKQISKPGQYVKFTTSDRVNDAGQATIPFFDGQSVTSTISTALSGLGLF  658

Query  376  FR  377
             R
Sbjct  659  HR  660


>☐ gb|ACZ28338.1|  hypothetical conserved protein [Simulium nigrimanum]
Length=522

 Score =  309 bits (791),  Expect = 5e-82, Method: Composition-based stats.
 Identities = 94/401 (23%), Positives = 157/401 (39%), Gaps = 50/401 (12%)

Query  7    GFSRTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANF-GSCNSKCEDYAAPRHNFKED  63
              SR     +P    EH    T+ ++ +F   Y   E++  G+C   C      R
Sbjct  146  KVSRDYWVCNPSVNEHT---TF-QMNKFIYGYMDIESDLTGTCTGTCN----LRLQRGIV  197

Query  64   NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQ  123
                  QR+C G + +C             ++    ++ E+   Y Y++  V
Sbjct  198  AMTNEVQRKCVGLVSHCRTTSDQ------TFEYCFAKNEEASYMYEYIKNDVIK------  245

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNF---DTWIRGFVRCDICRCLCDESKDSESIRSF  180
                           G+KE              +  GFV+CD C C C ES+D E+ R
Sbjct  246  --------------RGSKEMCSSDINPLQHKHAYTAGFVKCDYCVCEC-ESEDDETHRYL  290

Query  181  SLNAVESDTKDNKVVIGVQLAL-KDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSA  239
            SL    S+  +NKV++G++    + K+   +++++PL P G  D+ K  W+  DP L
Sbjct  291  SLLPQMSNIVENKVIVGMKFTKYETKLISLQVKEAPLQPYGYVDETKAEWRDLDPMLLPN  350

Query  240  SK----DIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL  295
            +         +    R+I +      L   V+TG+KF  ++G   + ++ T F+  SG +
Sbjct  351  TYWDYPGYHIMSGKLRAIRMKKLEGPLH-TVLTGIKFIVVNGVPDISIRHTYFNFTSGVM  409

Query  296  MNDNNQWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVP  355
              + + WI  +    K  +      P         +      + F  + L  D G   VP
Sbjct  410  QPEKSYWITETHDDPKPIDYNGLPEPGACATL-FPDTKSGQKILFGHNGL-GDAGSKVVP  467

Query  356  FIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTHPT  395
            F D + +V     AL G G+  +     GG +APVV T P
Sbjct  468  FFDSQPVVPDKPLALSGAGILHKGAQNCGGLVAPVVTTLPE  508


>☐ ref|XP_001603579.1| UG PREDICTED: similar to GA15301-PA [Nasonia vitripennis]
Length=612

 GENE ID: 100119874 LOC100119874 | similar to GA15301-PA [Nasonia vitripennis]

 Score =  308 bits (790),  Expect = 5e-82, Method: Composition-based stats.
 Identities = 88/415 (21%), Positives = 166/415 (40%), Gaps = 74/415 (17%)

Query  23   GETYEEITRFNQVYFTQEAN-----FGSCNSKCEDYAAPRHNFKED-NRYKFPQRQCTGT  76
```

```
            G+T+ ++    Q +FT +++     +      CE     R   +    + +  +R+C G
Sbjct  221  GQTFLKLNPVFQTFFTYQSDIKYTAWYGAVHTCESMVDKRLWCRSHPCKVQPTERRCYGK  280

Query  77   IYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRY  136
            + +C        E V ++       P+SD+RY ++++    +
Sbjct  281  LRDC--------EYVAKWFQICHAPPKSDRRYDWIKIYKGGEYTDIHGV-----------  321

Query  137  VYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVI  196
                 E  D+T     TW         C C C ES+  E  R FSL  V+SD  DNKV+
Sbjct  322  ---QNERCDNTQGVGSTWP-----FHTCACNC-ESESREEDRFFSLLDVKSDIGDNKVIT  372

Query  197  GVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS---------------------  235
            G++L  K  +++ +I+Q  L            WK+ D
Sbjct  373  GLRLVKKGHVYHIEIQQGTLGINASIIDAS-QWKNTDELFNQTTLSSDEIATYHYQQISS  431

Query  236  ---------------LSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLA  280
                               +  +F +  ++ ++++ +      D V+TGV+F +    L
Sbjct  432  PQQNDYRITMFNEWVDWHKNNLVFSVTYDQSAVNLNTRELGPED-VLTGVRFHQSGSALD  490

Query  281  LGVKLTKFDVMSGKLMNDNNQWIY-NSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVE  339
            + V  T FD  +G+L   N +WI+ N     K       + P+    +    K +   +
Sbjct  491  IQVHGTPFDYATGRLDTKNGRWIFINRARNVKEIVLKDPDIPVNSKKSDRFSKNNNEFIR  550

Query  340  FRQSDLKDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            FR+++ + DVGQ+ +P++D++ IV+     L G G+  + ++  GGF+   +V +
Sbjct  551  FRRTNFRKDVGQNVIPYLDIQDIVTNPPIPLSGAGIYHKYQEGYGGFLGLKIVHY  605


>☐ gb|ACH56860.1|  hypothetical protein [Simulium vittatum]
Length=351

 Score =  307 bits (786),  Expect = 2e-81, Method: Composition-based stats.
 Identities = 82/367 (22%), Positives = 137/367 (37%), Gaps = 39/367 (10%)

Query  37   FTQEANFGSCNSKCEDYAAPRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQI  96
            F  + + G C + C +   PR             R C G I  C             ++
Sbjct  3    FESDLSLGHCKTACYELQQPRQMNLVKRM---SNRSCNGYIGQCGASSK------GTFEY  53

Query  97   KKPESPESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVN-FDTWI  155
                +  + + Y YV   V                       G+KE      V    T++
Sbjct  54   CYSNNEFASEMYSYVSSDV--------------------MRKGSKELCQGEEVKSVSTYL  93

Query  156  RGFVRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSP  215
             G   C+ C C+C  S+D  + R  SL    S+  +N++++GV+   ++++ Y +I Q+P
Sbjct  94   SGVFHCEYCACIC-HSEDKTAHRFLSLLPQFSNIDENQIIVGVRFVKRNRVLYLQIMQAP  152

Query  216  LLPLGKADQRKLSWKSDD-----PSLSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGV  270
            LLP    D     W+  D     P   +        D   R +++          ++TGV
Sbjct  153  LLPYAAVDMDSAEWRDIDELGMLPPEHTNLSGYHAFDHKNREMYMKELEV-PGGSLLTGV  211

Query  271  KFAEIDGNLALGVKLTKFDVMSGKLMNDNNQWIYNSLSKQKLNNGGVAN-RPIKFTDNHA  329
            KF   DG   L ++ T  +  +G+L   ++ W+         N+    N  P +
Sbjct  212  KFIIKDGVPDLSIRYTFVNWTTGELNQSSSLWMTEKHDTYDTNDMLYQNQLPDQCKTPSF  271

Query  330  SEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVSTGT-ALIGVGLQFRTRDTSGGFIAP  388
             +        F  S L  D  Q  +P+ D + +V      L G+G+  +  D  GGFI+P
Sbjct  272  IDTASGQRFRFGTSALDQDGAQHVLPYFDAQPVVPMMMVPLSGIGVTHKGDDRCGGFISP  331

Query  389  VVVTHPT  395
            VV T
Sbjct  332  VVFTVNE  338


>☐ ref|XP_002040062.1|  GM15999 [Drosophila sechellia]
 gb|EDW56927.1|  GM15999 [Drosophila sechellia]
Length=549

 GENE ID: 6615691 Dsec\GM15999 | GM15999 gene product from transcript GM15999-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score =  306 bits (784),  Expect = 3e-81, Method: Composition-based stats.
```

```
 Identities = 76/322 (23%), Positives = 115/322 (35%), Gaps = 64/322 (19%)

Query  8    FSRTLDNVDP--KEHIKGETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKE  62
             SR     DP   +H++G+TY  +TR  Q +   E N     SC + C DY   R +
Sbjct  272  SSRLYWRCDPGKNKHVEGQTYGRVTRLLQGFVENEVNLNTEQSCMNSCGDYHDTRSSGCY  331

Query  63   D--NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK  120
                 +   Q  C G +YNC   +           +    +  S +RY Y+    +
Sbjct  332  KPEEEFCGQQPACKGRLYNCNTIESDIS-------VCPSSNASSARRYEYISFKDD----  380

Query  121  KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRG-FVRCDICRCLCDESKDSESIRS  179
                              GN    +       +W    F RC  C CLCDE   + S R
Sbjct  381  ------------------GNSPDCESQVRKPSSWSSWIFWRCHYCFCLCDEP-STNSDRY  421

Query  180  FSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLS--  237
            F+L    SD K NKV                     LLP G      L W   +
Sbjct  422  FNLRDSLSDFKQNKV-------------------GELLPKGAIRNSSLEWLPVEDYNVTD  462

Query  238  ---SASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGK  294
                   D   L ++ R++ +   + N +D VVTG +F     +L L V  + F+  +G+
Sbjct  463  VDVRDGYDYHTLSMDSRAMDVDEISVNSTDQVVTGFRFRVFKQHLNLEVLFSSFNFSTGR  522

Query  295  LMND--NNQWIYNSLSKQKLNN  314
            L+     + W+ N  S  + +
Sbjct  523  LIEPQTKSFWLGNQNSHLEGHR  544


>☐ gb|ACE75299.1|  conserved hypothetical protein [Glyptapanteles indiensis]
Length=643

 Score =  306 bits (783),  Expect = 4e-81, Method: Composition-based stats.
 Identities = 85/415 (20%), Positives = 147/415 (35%), Gaps = 79/415 (19%)

Query  11   TLDNVDPKEHIKGETYEEITRFNQVYFTQE-----ANFGSCNSKCEDYAAPRHNFKED--  63
             L N DP  +I+GET+ E     Q     E      +   C  +C+     R     D
Sbjct  216  NLLNCDPPNYIRGETFSEFLGLFQGVIINELQTDSIDASQCQGECDAVDYSRLVRCYDFN  275

Query  64   ----NRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDK  119
                       + C G +  C           I           S +R+ +
Sbjct  276  AKNRTITHCHTKPCNGLLTACDE---------IGAIKTCEMYENSSRRFLWAHHIDEDSS  326

Query  120  KKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFD-TWIRGFVRCDICRCLCDESK-DSESI  177
                                         V FD  W     RC +C C+C E + +S ++
Sbjct  327  HL---------------------ECPGRVVKFDNAWENDISRCSVCVCICVEQEGNSTAM  365

Query  178  RSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDP---  234
            R+ SL    +D   N V+ GV+ A KD++F+ +I Q  L   G+       WK
Sbjct  366  RTISLIPQLADIHRNMVMTGVRFAEKDRMFHLQIEQGELGSFGEIVPGTAEWKELGDFQY  425

Query  235  -------SLSSAS---------KDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAE--ID  276
                                   D   + +++R+I++    +N    V+ GV+F    +D
Sbjct  426  DSAGEGSFSIKKGEKFVKLTEFVDFSFVALDQRTINLDELFAND-GQVLIGVRFIYNGMD  484

Query  277  GNLALGVKLTKFDVMSGKLMNDNN---QWIYNSLSKQKLNNGGVA---------NRPIKF  324
                L VK    D+ +G+L N N    +WI +  ++++  N             N P+K
Sbjct  485  DAFELQVKSWPVDIRTGELANGNPSDVEWIGSENAQKRSENYDGPRRMFEIGETNEPLKT  544

Query  325  TDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVSTG--TALIGVGLQFR  377
             + +    +    +  R ++ +DD+ Q TVP++DL+ +  +     L G+ +  R
Sbjct  545  REWNHPLSVPNQRLMIRATNYEDDLHQHTVPYLDLQPVTYSTAKIPLSGLEIIHR  599


>☐ ref|XP_001605051.1| UG PREDICTED: similar to GA15301-PA [Nasonia vitripennis]
Length=592

 GENE ID: 100121438 LOC100121438 | similar to GA15301-PA [Nasonia vitripennis]

 Score =  296 bits (759),  Expect = 2e-78, Method: Composition-based stats.
 Identities = 92/379 (24%), Positives = 149/379 (39%), Gaps = 58/379 (15%)

Query  31   RFNQVYFTQEANFGS---CNSKCEDYAA-PRHNFKEDNRYKFPQRQCTGTIYNCIGEKIT  86
            +  Q Y   E +   S   C+  C  ++   +     DN+    QR C G ++NC  E
```

```
Sbjct  243  QVFQNYIVNEVDLSSEKSCSKDCPHFSGGSQVQGCFDNKICAKQRTCEGKLWNCYEESK-  301

Query  87   EKEPVIRYQIKKPES-PESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRYVVYGNKEHSD  145
                     I  P S   ++ RY Y+++                     RY     ++
Sbjct  302  --------IIVCPASKENNNIRYDYIDI------------------KKKRYGEEKSDNCK  335

Query  146  HTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVIGVQLALKDK  205
            +  V   TW+R   +C  C C+C E   ++  R  SL  V SD   N VV G++  LK+
Sbjct  336  NEVVEAKTWLRLPYKCSNCMCIC-EVSSPKTDRYISLRPVISDAHLNYVVTGMKFELKNG  394

Query  206  IFYFKIRQSPLLPLGKADQ------RKLSWKSDDPSLSSA-----SKDIFKLDVNKRSIH  254
            I + +I+QS LL  G  +       +   WK  D    +       KD + L    R++
Sbjct  395  IIHIRIKQSKLLANGLIENAVSELPKSEQWKKIDDYTINDKDIFEGKDYYMLSWQNRTMC  454

Query  255  IGSATSNLSDFVVTGVKFAEIDGN----LALGVKLTKFDVMSGKLMNDNNQWIYNSLSKQ  310
            +    S     V+TG++F     N    L L V+LT ++  +G+L + N+ W  N L +
Sbjct  455  LDELDSLNKSDVLTGLRFKLESVNDQTYLKLEVRLTPYNFTTGRLSDKNSYWYSNGLQQN  514

Query  311  KLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSI-VSTGTAL  369
                        K   +  +        +F  ++L+ D GQSTVP ++   +       L
Sbjct  515  DRTEL-------KLDSSVTANPQLDSYFKFT-TNLEIDAGQSTVPHVEHVELETKVSVPL  566

Query  370  IGVGLQFRTRDTSGGFIAP  388
             G GL    +D S  + A
Sbjct  567  SGAGLF-INKDLSFKWKAF  584


>☐ ref|XP_967785.2| G PREDICTED: similar to conserved hypothetical protein [Tribolium
castaneum]
Length=519

 GENE ID: 656143 LOC656143 | similar to CG2206-PA, isoform A
[Tribolium castaneum]

 Score =  275 bits (703),  Expect = 8e-72, Method: Composition-based stats.
 Identities = 73/249 (29%), Positives = 116/249 (46%), Gaps = 14/249 (5%)

Query  156  RGFVRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSP  215
                 C  C CLC+E + S S R  +L   +++  DN+VV G++    ++I +  I++
Sbjct  90   YRNSYCPYCFCLCEEGQVSYSERYVNLRITKANISDNRVVTGLRFIKHNQIVHLHIQEGK  149

Query  216  LLPLGKADQRKLSWKSDDPSLSSASK-----DIFKLDVNKRSIHIGSATSNLSDFVVTGV  270
            LL  G  D + + W   D             D   L   +RS+ +    +     VV GV
Sbjct  150  LLEKGYVDPKTVQWVPPDNYKIDDRNIYEEEDYHTLSWEERSLELTEIVAE-EGSVVIGV  208

Query  271  KFAEIDGNLALGVKLTKFDVMSGKLMNDNNQW--IYNSLSKQKLN---NGGVANRPIKFT  325
            +F +  G + L V  T FD  +GKL  DN  W    N+++   +N        + PI
Sbjct  209  RFKKTLGGVKLEVLTTTFDFDTGKL-EDNTTWNIFKNTITPNLINQEIKLDSPDIPI-HK  266

Query  326  DNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVST-GTALIGVGLQFRTRDTSGG  384
            +  + +  +   + F  +D   D GQ+TVPF D + + ST  T L G+GL  + R  SGG
Sbjct  267  NLASIDFSENTYINFTNTDYDKDAGQTTVPFFDAQPVYSTAPTPLSGIGLFHKGRKGSGG  326

Query  385  FIAPVVVTH  393
            FI P ++T+
Sbjct  327  FITPKILTY  335


>☐ ref|XP_001599112.1| G PREDICTED: similar to CG31664-PA [Nasonia vitripennis]
Length=530

 GENE ID: 100113795 LOC100113795 | similar to CG31664-PA [Nasonia vitripennis]

 Score =  241 bits (615),  Expect = 1e-61, Method: Composition-based stats.
 Identities = 69/330 (20%), Positives = 140/330 (42%), Gaps = 59/330 (17%)

Query  86   TEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRYVVYGNKEHSD  145
              K+PV  +     E+   + R++Y+ +                  + Y  +Y ++  ++
Sbjct  231  ETKDPVRGWVQYCNENKNKNSRHFYIHIF-----------------WIYLRLYLHRHETN  273

Query  146  HTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVIGVQLALKDK  205
              +V                C C++  D +  R F+L  V SD  +NK+V GV++
```

```
Sbjct  274  LPSV-------------FEYCTCEDPAD-DIDRHFNLRPVTSDVANNKIVTGVKIVKDQN  319

Query  206  IFYFKIRQSPLLPLGKADQRKLSWKSDDPSL-----SSASKDIFKLDVNKRSIHIGSAT-  259
            +F+ +I+Q  L   G  +   ++WK  +             D + ++   +++ +
Sbjct  320  LFHIQIKQGEL-SNGFVNTSSIAWKPLEKYSPLDIGYERGVDFYTVESRSKTVCLDDLDA  378

Query  260  ---SNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNNQWIYNSLSKQKLNNGG  316
               +N S++++TG++FA     L L ++LT +D  +G L+ +N+ WI N     K +NG
Sbjct  379  NKSANKSNYLLTGLRFAV-GPYLKLEIQLTPYDPETGYLIAENSTWISNINKPAKYSNGP  437

Query  317  VANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVSTGT-ALIGVGLQ  375
                  ++ + + +       +   +S+   D GQ+TVP ID++ I +     L G GL
Sbjct  438  FDWMDAEYDNYYIT----NRCIRISRSEDSLDFGQATVPLIDIQEITTDPVIPLSGAGLY  493

Query  376  FRTRD------------TSGGFIAPVVVTH  393
            ++++               GG+    + TH
Sbjct  494  YKSKQLYSHNDVKYELLRYGGYFGLRLFTH  523


>☐ ref|XP_001605439.1| G PREDICTED: similar to ENSANGP00000010019 [Nasonia vitripennis]
Length=585

 GENE ID: 100121830 LOC100121830 | similar to ENSANGP00000010019
[Nasonia vitripennis]

 Score =  229 bits (584),  Expect = 4e-58, Method: Composition-based stats.
 Identities = 74/278 (26%), Positives = 127/278 (45%), Gaps = 38/278 (13%)

Query  118  DKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESI  177
             +KK +Y  + G+   Y +  G++  +    +   ++   F   ++C C C+E   +
Sbjct  296  SEKKIEYIELLGTKKIYGF-KGSECPTKVEQLRGSSFAAQFRAVELCFCTCEEDPAASDD  354

Query  178  RSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSW---KSDDP  234
            R FSL  V S+   NKV+ GV++ L ++   F+++Q  L+  G   QR ++W   +  DP
Sbjct  355  RYFSLKPVTSNVVKNKVLTGVRIVLMNRTVQFQVQQGRLVENGNKAQRSVNWGRLQPLDP  414

Query  235  SLSS--ASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMS  292
               S  A KD  ++  + R+++I S  +  +D  +TG+KF +   +L L ++ T  D  +
Sbjct  415  LDRSFVAGKDYHRITYDARTVNIDSLEA-PADCALTGLKFEKEGSSLKLAMQYTPIDFAN  473

Query  293  GKLMNDNNQWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQS  352
            GKL+   + WI + +   +                               S L+ D+G S
Sbjct  474  GKLIASGSTWISSPVGSSR-----------------------------DSSLESDLGGS  503

Query  353  TVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPV  389
             VPFID + I ST    L G GL ++    S GFI
Sbjct  504  VVPFIDSQPITSTPAVLLSGAGLVYKGGRKSSGFITMK  541


>☐ gb|ACZ28354.1|  hypothetical protein [Simulium nigrimanum]
Length=291

 Score =  228 bits (581),  Expect = 1e-57, Method: Composition-based stats.
 Identities = 65/279 (23%), Positives = 116/279 (41%), Gaps = 16/279 (5%)

Query  126  IIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNAV  185
            +    N   ++V           +  +TW+     CD C C+CD+  +        L
Sbjct  1    LCIQQNPQKKFVKYQVPDEKCHQLVLNTWLYDLRNCDFCSCICDKDDEEN---YVYLGPS  57

Query  186  ESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASKDIFK  245
            +  T  ++VV GV+L L++ + Y++++       G  D   L W+            +FK
Sbjct  58   KV-TDPDRVVTGVRLVLENGVLYWQLQTGKPRTFGFIDSETLEWQPLPKVSDKKDVRLFK  116

Query  246  LDVNKRSIHIGSATSNLSDFVVTGVKFAEIDG---NLALGVKLTKFDVMSGKLMNDNNQW  302
              +N++   I         +V+TG++F +      +  L +   + ++M G+L ND N +
Sbjct  117  -KINRKQSVILRNLEVPVPYVLTGIQFVKSSNPPDSFDLNLFGRQIELMEGRLFNDTNPF  175

Query  303  -----IYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFI  357
                 +++    +  +     N+  K T            V F  S++K D  Q  VPF
Sbjct  176  GPEKALFDRYRAENRSALINVNKEEKITVTAKDNHTV--YVVFGHSNMKTDFSQYLVPFF  233

Query  358  DLRSIVST-GTALIGVGLQFRTRDTSGGFIAPVVVTHPT  395
```

```
            D+R++  +    L GVGL  RT D  GG IAP ++
Sbjct  234  DIRNVTPSIPMPLKGVGLYHRTNDDYGGLIAPRLIAINP  272


>☐ gb|ACH56924.1|  hypothetical protein [Simulium vittatum]
Length=389

 Score =  224 bits (570),  Expect = 2e-56, Method: Composition-based stats.
 Identities = 65/242 (26%), Positives = 106/242 (43%), Gaps = 19/242 (7%)

Query  167  LCDESKDSESIRSF-------SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPL  219
            +CD+ +  E  + F       SL   ESD KD KV++G+QL  KD + YFKI+Q+PLLP
Sbjct  140  VCDQGRSGEHYKKFKNSLPMVSLKPQESDIKDGKVIVGLQLTTKDDVLYFKIKQAPLLPH  199

Query  220  GKADQRKLSWKSDDPSLSS-ASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGN  278
               ++   SW   +  LS    K++  ++     + +   T    + V+TG+ F     N
Sbjct  200  FHVNESASSWMDIEYMLSRPDDKNLTTVEPYHWKLMMKELTV-PENTVLTGIGFGYDSKN  258

Query  279  -LALGVKLTKF-DVMSGKLMNDNNQWIYNSLSKQKLNNGGVAN-RPIKFTDNHASEKLDL  335
             L + +K T   +  SG+L    + W+       +                +   + ++
Sbjct  259  QLDIQLKYTPVLNASSGELDVLASGWMAERHDAHRTKEFDNKVKVSTSCELDSFPDMMNG  318

Query  336  GLVEFRQSDLKDDVGQSTVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTHP  394
              +      L   V    +PFID + +V     AL G G+  +  D  GGF+APV +T
Sbjct  319  QCL------LMKKVSNDIIPFIDTQELVPKPMMALSGAGITHKGHDNCGGFLAPVALTLS  372

Query  395  TY  396
             Y
Sbjct  373  DY  374


>☐ gb|ACZ28153.1|  hypothetical protein [Simulium nigrimanum]
Length=296

 Score =  215 bits (547),  Expect = 7e-54, Method: Composition-based stats.
 Identities = 63/281 (22%), Positives = 115/281 (40%), Gaps = 16/281 (5%)

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLN  183
              +    N   ++V           +  +TW+     CD C C+CD+  +        L
Sbjct  22   NTLCIQQNPQKKFVKYQVPDEKCHQLVLNTWLYDLRNCDFCSCICDKDDEEN---YVYLG  78

Query  184  AVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASKDI  243
              +  T  ++VV GV+L L++ + Y++++       G  D   L W+            +
Sbjct  79   PSKV-TDPDRVVAGVRLVLENGVLYWQLQTGKPRTFGFIDSETLEWQPLPKVTDKKDARL  137

Query  244  FKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDG---NLALGVKLTKFDVMSGKLMNDNN  300
            FK  +N++   I         +V+TG++F +      +  L +   + ++M G+L ND N
Sbjct  138  FK-KINRKQSVILRNLEVPVPYVLTGIQFVKSSNPPDSFDLNLFGRQIELMEGRLFNDTN  196

Query  301  QW-----IYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVP  355
            ++     +++    +  +     N+  K T            V F  S++K D  Q  VP
Sbjct  197  RFGPEKALFDRYRAENRSALVNVNKEEKITVTAKDNHTV--YVVFGHSNMKTDFSQYLVP  254

Query  356  FIDLRSIVST-GTALIGVGLQFRTRDTSGGFIAPVVVTHPT  395
            F D+R++  +    L GVGL  RT D  GG I   ++
Sbjct  255  FFDIRNVTPSIPMPLKGVGLYHRTNDDYGGLIRHGLIGINP  295


>☐ gb|ACH56909.1|  hypothetical protein [Simulium vittatum]
Length=313

 Score =  214 bits (546),  Expect = 1e-53, Method: Composition-based stats.
 Identities = 61/282 (21%), Positives = 110/282 (39%), Gaps = 17/282 (6%)

Query  124  YAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLN  183
             ++   ++   ++V   K       +    W+  F  C+ C C+CD+           L
Sbjct  25   NSLCMENDLKEKFVKYTKPKDKCHQLPLTNWLYDFRNCEYCSCICDKDDQEN---FVYLG  81

Query  184  AVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASKDI  243
            A++     + VV GV+L + + + +++++       G  D   L+W+   P +
Sbjct  82   AIQVANLGH-VVTGVRLVMANNVLFWQLQTGKPRSFGFIDADTLAWQPL-PKIDRKKDAK  139
```

```
Query  244  FKLDVNKRSIHIGSATSNLSDFVVTGVKF--AEIDGN--LALGVKLTKFDVMSGKLMNDN  299
                VNK+   I       + +V+TGV+F    ++ N    + +     D+M GKL N++
Sbjct  140  LLHRVNKKQSFIMRNLEVPAPYVLTGVQFFARSLNQNAAYDMNLYGRPIDLMEGKLYNES  199

Query  300  NQWI-----YNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTV  354
             Q+      ++             N+    T            +    S +  D  Q  V
Sbjct  200  RQFTPNKEQFDPYQTNTRLALVNENKEEDTTIPRKDNYTVNVFL--GHSSITGDFSQYLV  257

Query  355  PFIDLRSIVST-GTALIGVGLQFRTRDTSGGFIAPVVVTHPT  395
            PF D R++ S+    L GVGL  RT D   G IAP ++
Sbjct  258  PFFDTRNVTSSVPMPLKGVGLFHRTNDFYAGLIAPRLIALNP  299
```

>☐ gb|ACH56925.1| hypothetical protein [Simulium vittatum]
Length=221

```
 Score =  209 bits (533),  Expect = 3e-52, Method: Composition-based stats.
 Identities = 52/214 (24%), Positives = 94/214 (43%), Gaps = 12/214 (5%)

Query  187  SDTKDNKVVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASKDIFKL  246
            SD  + KV++G+QL  KD +  FKI ++PLLP    +  +  WK  +  + S+  +   +
Sbjct  1    SDIANGKVIVGLQLLAKDGVLTFKILEAPLLPHFHVNASEKKWKELE-YIRSSPDNKTVV  59

Query  247  DVNKRSIHIGSATSNLSDFVVTGVKFAEIDGN-LALGVKLTKF-DVMSGKLMNDNNQWIY  304
            + +   + +   T    + V+TG+ F     N L + +K T   +  +G+L    + W+
Sbjct  60   EPHHWKLMMKELTVP-ENTVLTGIGFRYDGKNQLDIQLKYTPVLNASTGELDVLASGWMT  118

Query  305  NSLSKQKLNNGG-VANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIV  363
                 Q+          P     N   + ++   +      L   V    +PFI+ + +V
Sbjct  119  ERHDAQRTKEFDKNVQIPTSCEVNSFPDMMNGQCL------LMKKVSNDIIPFIETQEVV  172

Query  364  S-TGTALIGVGLQFRTRDTSGGFIAPVVVTHPTY  396
                 AL G G+  +  D  GG++APV +T   Y
Sbjct  173  PEPMMALSGAGITHKGHDNCGGYLAPVALTLSDY  206
```

>☐ gb|ACZ28152.1| hypothetical conserved protein [Simulium nigrimanum]
Length=293

```
 Score =  204 bits (520),  Expect = 1e-50, Method: Composition-based stats.
 Identities = 58/249 (23%), Positives = 104/249 (41%), Gaps = 16/249 (6%)

Query  156  RGFVRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSP  215
                   +  C+CD+  +        L   +  T  ++VV GV+L L++ + Y++++
Sbjct  45   YHLAYYHLDSCICDKDDEEN---YVYLGPSKV-TDPDRVVTGVRLVLENGVLYWQLQTGK  100

Query  216  LLPLGKADQRKLSWKSDDPSLSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEI  275
                G  D   L W+            +FK  +N++   I         +V+TG++F +
Sbjct  101  PRTFGFIDSETLEWQPLPKVTDKKDARLFK-KINRKQSVILRNLEVPVPYVLTGIQFVKS  159

Query  276  DG---NLALGVKLTKFDVMSGKLMNDNNQW-----IYNSLSKQKLNNGGVANRPIKFTDN  327
                 +  L +   + ++M G+L ND N++     +++    +        N+  K T
Sbjct  160  SNPPDSFDLNLFGRQIELMEGRLFNDTNRFGPEKALFDRYRAENRLALVNVNKEEKITVT  219

Query  328  HASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVST-GTALIGVGLQFRTRDTSGGFI  386
                      V F  S++K D  Q  VPF D+R++  +    L GVGL  RT D  GG I
Sbjct  220  AKDNHTV--YVVFGHSNMKTDFSQYLVPFFDIRNVTPSIPMPLKGVGLYHRTNDDYGGLI  277

Query  387  APVVVTHPT  395
            AP ++
Sbjct  278  APRLIAINP  286
```

>☐ gb|AAU06548.1| unknown salivary protein [Culicoides sonorensis]
Length=225

```
 Score =  202 bits (514),  Expect = 5e-50, Method: Composition-based stats.
 Identities = 115/210 (54%), Positives = 140/210 (66%), Gaps = 20/210 (9%)

Query  5    SLGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAAPRHNFKEDN  64
```

```
            S G SR L N DPK H++G+TYEEITRFNQVVYFTQEA FGSC+S CEDYAAPRHNFKED+
Sbjct  35   SAGESRELYNTDPKAHVRGDTYEEITRFNQVVYFTQEATFGSCDSVCEDYAAPRHNFKEDD  94

Query  65   RYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQY  124
            +YKFPQRQCTGTI++C  E   E      Y +++P+S  SD RYYY       + +
Sbjct  95   QYKFPQRQCTGTIHSCTSEYSPE------YHVRQPKSETSDARYYYFSKQKGFNGEL---  145

Query  125  AIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNA  184
                       Y+YGNK+  ++    F  W R FVRCDIC+CLCDE +D +S+RSFSLN
Sbjct  146  -----------YIYGNKKEPEYRFEIFKGWERVFVRCDICKCLCDEPEDKKSVRSFSLNT  194

Query  185  VESDTKDNKVVIGVQLALKDKIFYFKIRQS  214
            + SD  DN VV G  LA KDK+F +KI  S
Sbjct  195  MTSDKSDNMVVTGSPLAWKDKVFSWKIHIS  224


>☐ ref|XP_002082764.1| GD11752 [Drosophila simulans]
 gb|EDX08349.1| GD11752 [Drosophila simulans]
Length=163

 GENE ID: 6735854 Dsim\GD11752 | GD11752 gene product from transcript GD11752-RA
[Drosophila simulans] (10 or fewer PubMed links)

 Score =  200 bits (509),  Expect = 2e-49, Method: Composition-based stats.
 Identities = 41/153 (26%), Positives = 73/153 (47%), Gaps = 7/153 (4%)

Query  248  VNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMND--NNQWIYN  305
            ++ R++ +   +   +D VVTG +F     +L L V+ + F+  +G+L+     + W+ N
Sbjct  1    MDSRALDLDEISVKSTDQVVTGFRFRVFKQHLNLEVRFSSFNFSTGRLIEPQTKSFWLGN  60

Query  306  SLSK----QKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRS  361
              S     +K      ++ P            +   +EF  S   D   Q+TVPFID++
Sbjct  61   QNSHLEGHRKRLILKESDLPTASELPSLPLSQNNQFLEFGSSSQLKDAAQNTVPFIDVQE  120

Query  362  IVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTH  393
            +V      L G+G+ ++ R   GGF AP V+T+
Sbjct  121  VVPRPAMPLAGLGIYYKGRPGYGGFFAPKVMTY  153


>☐ gb|ACZ28340.1|  hypothetical conserved protein [Simulium nigrimanum]
Length=346

 Score =  196 bits (498),  Expect = 4e-48, Method: Composition-based stats.
 Identities = 51/246 (20%), Positives = 105/246 (42%), Gaps = 24/246 (9%)

Query  167  LCD-ESKDSESIRSF-------SLNAVESDTKDNKVVIGVQLALKDKIFYFKIRQSPLLP  218
            +C+ +S++ ++ + F       SL   ESD K+ KV++G+QL +   +  F+I Q+PL+
Sbjct  94   VCNPQSRNPDTFKKFEKKTPLVSLTPQESDIKNGKVIVGLQLTIVFDVLTFQIMQAPLIS  153

Query  219  LGKADQRKLSWKSDDPSLSSASKDIFKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGN  278
                +  +  W     +  S + +   ++     + +   T    + V+TG+ F     N
Sbjct  154  NFYVNATQSEWIKT-YNQHSLANNNTVVEPYHWKLMMKELTV-PENTVLTGIGFRYDSKN  211

Query  279  -LALGVKLTKF-DVMSGKLMNDNNQWIYNSLSKQKLNNGGVANR-PIKFTDNHASEKLDL  335
             L + +K T   +  +G+L    + W+      Q         +       +   +  +
Sbjct  212  QLDIQLKYTPVLNATTGELDVLASGWMAERHDAQMTKEFSTKVKASASCEVDPFPDMPNG  271

Query  336  GLVEFRQSDLKDDVGQ----STVPFIDLRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVV  390
              +      L  ++G     + +P+++ + +V     AL G G+  +  D  GGF+APV
Sbjct  272  QCL------LMKNLGNGHQANWIPYVETQQVVPQPMMALSGAGITHKGHDDCGGFLAPVT  325

Query  391  VTHPTY  396
            +T   Y
Sbjct  326  LTLSDY  331


>☐ gb|AAU06547.1|  unknown salivary protein [Culicoides sonorensis]
Length=218

 Score =  195 bits (496),  Expect = 6e-48, Method: Composition-based stats.
```

```
 Identities = 111/203 (54%), Positives = 134/203 (66%), Gaps = 20/203 (9%)

Query  5    SLGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAAPRHNFKEDN  64
            S G SR L N DPK H++G+TYEEITRFNQVYFTQEA FGSC+S CEDYAAPRHNFKED+
Sbjct  35   SAGESRELYNTDPKAHVRGDTYEEITRFNQVYFTQEATFGSCDSVCEDYAAPRHNFKEDD  94

Query  65   RYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQY  124
            +YK PQRQCTGTI++C  E   E      Y +++P+S  SD RYYY       + +
Sbjct  95   QYKLPQRQCTGTIHSCTSEYSPE------YHVRQPKSETSDARYYYFSKQKGFNGEL---  145

Query  125  AIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNA  184
                       Y+YGNK+  ++    F  W R FVRCDIC+CLCDE +D +S+RSFSLN
Sbjct  146  -----------YIYGNKKEPEYRFEIFKGWERVFVRCDICKCLCDEPEDKKSVRSFSLNT  194

Query  185  VESDTKDNKVVIGVQLALKDKIF  207
            + SD  DN VV GV LA K  IF
Sbjct  195  MTSDKSDNMVVTGVHLAWKISIF  217


>☐ gb|EFA12874.1|  hypothetical protein TcasGA2_TC010726 [Tribolium castaneum]
Length=304

 Score =  193 bits (491),  Expect = 3e-47, Method: Composition-based stats.
 Identities = 47/191 (24%), Positives = 86/191 (45%), Gaps = 15/191 (7%)

Query  194  VVIGVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPS-----LSSASKDIFKLDV  248
            VV G++    ++I + ++++  LL  G  +   + W   +            +D   +
Sbjct  44   VVTGLRFIKHNRIVHLQVQEGKLLEKGNINPTTVQWVPPEDYKITDRDIYKGQDYHTISW  103

Query  249  NKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKL-MNDNNQW-----  302
             +R+I +    +  +  VVTGV+F    G+L L V  T FD  +GKL  N ++ W
Sbjct  104  EERTIELTKVIAK-AGSVVTGVRFKRTFGSLNLQVLTTAFDFNTGKLEKNSDSDWGIFES  162

Query  303  --IYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLR  360
               +N   K +L      +     T+   ++  +   +EF  +D + D GQ+T+PF D +
Sbjct  163  APYFNIRHKFRLKVNLNPSDVPVRTNLDFTDFSENMYIEFANTDYRKDAGQTTIPFFDAQ  222

Query  361  SIVS-TGTALI  370
            ++ S T T L
Sbjct  223  AVNSTTPTPLS  233


>☐ ref|XP_002041740.1| G GM16585 [Drosophila sechellia]
 gb|EDW45556.1| G GM16585 [Drosophila sechellia]
Length=248

 GENE ID: 6617415 Dsec\GM16585 | GM16585 gene product from transcript GM16585-RA
[Drosophila sechellia] (10 or fewer PubMed links)

 Score =  187 bits (474),  Expect = 3e-45, Method: Composition-based stats.
 Identities = 44/184 (23%), Positives = 71/184 (38%), Gaps = 36/184 (19%)

Query  3    LPSLG-----FSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGS---CNSKCEDYA  54
            LP +        R++   DP++H  G TYEE+ R  Q Y   E +  S   CN  C  Y
Sbjct  91   LPKIKVLMSQAERSVWRCDPQKHQSGITYEEVNRLLQGYVENEVDLNSDGACNHDCGYYN  150

Query  55   APRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELP  114
            + ++    DN++   Q +CTG +++C          V            S +RY +++
Sbjct  151  SAKNEGCFDNKFCSEQPKCTGGVHDCR--------FVESSMQICQAEKNSSRRYEFIK--  200

Query  115  VNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR-CDICRCLCDESKD  173
                             Y    V+GN++       +  +W R     C  C CLCD+
Sbjct  201  -----------------YESGLVHGNEKPCASWLTSAKSWNRWLFMECSYCLCLCDDQSP  243

Query  174  SESI  177
            S  +
Sbjct  244  STCV  247


>☐ gb|AAU06499.1|  unknown salivary protein [Culicoides sonorensis]
```

```
Length=157

 Score =  173 bits (439),  Expect = 3e-41, Method: Composition-based stats.
 Identities = 124/155 (80%), Positives = 140/155 (90%), Gaps = 0/155 (0%)

Query  244  FKLDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGNLALGVKLTKFDVMSGKLMNDNNQWI  303
            FKLDV KR++HIGSA S L+DFVVTGVKFAE+DGNLALGVK+TKFDVMSGKL+ND NQWI
Sbjct  1    FKLDVKKRTVHIGSAKSKLTDFVVTGVKFAEVDGNLALGVKVTKFDVMSGKLINDINQWI  60

Query  304  YNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIV  363
            +NSL  QKLNN G A RPIKFTD+ A+E  ++GLV+F+QSDLKDDVGQSTVPFID+R IV
Sbjct  61   FNSLXTQKLNNKGGAYRPIKFTDHRATEDNNVGLVQFKQSDLKDDVGQSTVPFIDVRPIV  120

Query  364  STGTALIGVGLQFRTRDTSGGFIAPVVVTHPTYTG  398
            +TGTALIGVGLQ RTR+TSGGFI PVVVTHPTYTG
Sbjct  121  ATGTALIGVGLQLRTRETSGGFIQPVVVTHPTYTG  155


>☐ gb|AAU06503.1|  unknown salivary protein [Culicoides sonorensis]
Length=183

 Score =  171 bits (434),  Expect = 1e-40, Method: Composition-based stats.
 Identities = 130/160 (81%), Positives = 147/160 (91%), Gaps = 0/160 (0%)

Query  1    NKLPSLGFSRTLDNVDPKEHIKGETYEEITRFNQVYFTQEANFGSCNSKCEDYAAPRHNF  60
            NKLPSLGFSR LDNVDPKEHIKG TYEEITRF+QVYFTQE+NFGSCNSKCEDY+APRHNF
Sbjct  23   NKLPSLGFSRPLDNVDPKEHIKGVTYEEITRFDQVYFTQESNFGSCNSKCEDYSAPRHNF  82

Query  61   KEDNRYKFPQRQCTGTIYNCIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKK  120
            KED+RYKFPQRQCTGTIY C+GEKITE+EPVIRY++KKP+S +S++RYY + +PVN +K
Sbjct  83   KEDDRYKFPQRQCTGTIYGCVGEKITEREPVIRYRVKKPKSSDSNERYYDISIPVNNNKT  142

Query  121  KYQYAIIEGSNYTYRYVYGNKEHSDHTTVNFDTWIRGFVR  160
             YQYAII+G NYTYRYVYGN EHSD+  VNFDTWIRGFVR
Sbjct  143  IYQYAIIDGLNYTYRYVYGNNEHSDYEIVNFDTWIRGFVR  182


>☐ gb|AAS87304.1|  CG2206-like protein [Drosophila miranda]
Length=132

 Score =  163 bits (413),  Expect = 3e-38, Method: Composition-based stats.
 Identities = 38/153 (24%), Positives = 54/153 (35%), Gaps = 31/153 (20%)

Query  23   GETYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKEDNRYKFPQRQCTGTIYN  79
            G TY+E+TR  Q Y   E +     +C   C  Y + R      + Y   Q +C+G +Y+
Sbjct  3    GVTYDEVTRLLQGYIENEVDLNNEETCRETCSFYQSTRSEGCYKDLYCARQPRCSGRLYD  62

Query  80   CIGEKITEKEPVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRYVYG  139
            C          V       P    S +RY Y+E                   Y      G
Sbjct  63   CQ--------FVDSDMWVCPSPKNSTRRYEYIE-------------------YENGRTLG  95

Query  140  NKEHSDHTTVNFDTWIRG-FVRCDICRCLCDES  171
             + +    T   D+W R     C  C C+CDE
Sbjct  96   RRANCVRGTTKVDSWWRYILWHCSYCFCMCDEQ  128


>☐ gb|ACZ28339.1|  hypothetical protein [Simulium nigrimanum]
Length=167

 Score =  145 bits (367),  Expect = 5e-33, Method: Composition-based stats.
 Identities = 31/159 (19%), Positives = 63/159 (39%), Gaps = 15/159 (9%)

Query  246  LDVNKRSIHIGSATSNLSDFVVTGVKFAEIDGN-LALGVKLTKF-DVMSGKLMNDNNQWI  303
            ++     + +   T    + V+TG+ F     N L + +K T   +  +G+L    + W+
Sbjct  1    VEPYHWKLMMKELTVP-ENTVLTGIGFRYDSKNQLDIQLKYTPVLNATTGELDVLASGWM  59

Query  304  YNSLSKQKLNNGGVANRPIK-FTDNHASEKLDLGLVEFRQSDLKDDVGQ----STVPFID  358
                  Q         +P      +   +  +   +      L  ++G     + +P+++
Sbjct  60   AERHDAQMTKEFSTKVKPSASCEVDPFPDMPNGQCL------LMKNLGNGHQANWIPYVE  113

Query  359  LRSIVS-TGTALIGVGLQFRTRDTSGGFIAPVVVTHPTY  396
```

```
             + +V     AL G G+  +  D  GGF+APV +T   Y
Sbjct  114  TQQVVPQPMMALSGAGITHKGHDDCGGFLAPVTLTLSDY  152


>☐ gb|EFA11732.1|  hypothetical protein TcasGA2_TC002380 [Tribolium castaneum]
Length=440

 Score =  125 bits (314),  Expect = 1e-26, Method: Composition-based stats.
 Identities = 29/105 (27%), Positives = 54/105 (51%), Gaps = 3/105 (2%)

Query  290  VMSGKLMNDNNQWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDV  349
                +L++D  + + +    ++       + PI+  +  +++      ++F  +D + D
Sbjct  158  FEETQLISDLKR-LDSPPKFKENVKLESPDIPIR-KNLASTDFSKNKCIKFTNTDYQKDA  215

Query  350  GQSTVPFIDLRSIVST-GTALIGVGLQFRTRDTSGGFIAPVVVTH  393
            GQ+T+PF D + + ST  T L GVGL  +    SGGFIAP + T+
Sbjct  216  GQTTIPFFDAQPVNSTIPTPLSGVGLFHKGHKGSGGFIAPKIFTY  260


>☐ gb|AAR20336.1|  CG2206-like protein [Drosophila pseudoobscura]
 gb|AAR20337.1|  CG2206-like protein [Drosophila pseudoobscura]
 gb|AAR20338.1|  CG2206-like protein [Drosophila pseudoobscura]
 gb|AAR20339.1|  CG2206-like protein [Drosophila pseudoobscura]
Length=104

 Score =  124 bits (311),  Expect = 2e-26, Method: Composition-based stats.
 Identities = 27/131 (20%), Positives = 39/131 (29%), Gaps = 30/131 (22%)

Query  33   NQVYFTQEANFG---SCNSKCEDYAAPRHNFKEDNRYKFPQRQCTGTIYNCIGEKITEKE  89
             Q Y   E +     +C    C  Y + R      + Y   Q +C+G +YNC
Sbjct  1    LQGYIENEVDLNNEETCRETCSFYQSTRSEGCYKDLYCARQPRCSGRLYNCQ--------  52

Query  90   PVIRYQIKKPESPESDKRYYYVELPVNTDKKKYQYAIIEGSNYTYRYVYGNKEHSDHTTV  149
             V       P    S +RY Y+E                   Y      G + +    T
Sbjct  53   FVDSDMWVCPSPKNSTRRYEYIE-------------------YENGRTLGQRANCVRGTT  93

Query  150  NFDTWIRGFVR  160
              D+W R
Sbjct  94   KVDSWWRYLFW  104


>☐ gb|AAU06508.1|  unknown salivary protein [Culicoides sonorensis]
Length=135

 Score =  119 bits (298),  Expect = 7e-25, Method: Composition-based stats.
 Identities = 70/127 (55%), Positives = 85/127 (66%), Gaps = 5/127 (3%)

Query  137  VYGNKEHSDHTTVNFDTWIRGFVRCDICRCLCDESKDSESIRSFSLNAVESDTKDNKVVI  196
            VYG+   S   TV F  WIRGFV C ICRC+ DE KD+ S+RSFSLN+++SDTK NKVV
Sbjct  1    VYGHLGGSKFKTVTFKAWIRGFVMCAICRCVRDEPKDNLSVRSFSLNSLKSDTKKNKVVT  60

Query  197  GVQLALKDKIFYFKIRQSPLLPLGKADQRKLSWKSDDPSLSSASKDIFKLDVNKRSIHIG  256
            GVQ+   + +F  KI QSPLL  GK + +K SWKSD P     S + FKL  NKRSIHIG
Sbjct  61   GVQIIQINNVFCLKIHQSPLLRKGKVNFKKGSWKSDSP-----SNNTFKLAFNKRSIHIG  115

Query  257  SATSNLS  263
            ++  N
Sbjct  116  TSCVNPK  122


>☐ gb|AAU06510.1|  unknown salivary protein [Culicoides sonorensis]
Length=163

 Score =  107 bits (268),  Expect = 2e-21, Method: Composition-based stats.
 Identities = 26/118 (22%), Positives = 52/118 (44%), Gaps = 3/118 (2%)

Query  280  ALGVKLTKFDVMSGKLMNDN--NQWIYNSLSKQKLNNGGVANRPIKFTDNHASEKLDLGL  337
             L V   + ++++G          W       + ++   +         N  + + ++ L
Sbjct  15   NLRVFGRQINILNGNFTPTPIVTHWNEIPKKAKVIDTENLIGGVDAKPQNVTTGQDEVKL  74
```

```
Query  338  VEFRQSDLKDDVGQSTVPFIDLRSIVSTG-TALIGVGLQFRTRDTSGGFIAPVVVTHP  394
            V+F  S L D+ GQ++VP+ D + + S       GVGL +RT +   G++AP ++
Sbjct  75   VKFSHSPLTDNFGQNSVPYFDTQQVTSRPKMVQAGVGLYYRTTEKYAGYLAPRLIAFN  132


>☐ gb|AAU06490.1|  unknown salivary protein [Culicoides sonorensis]
Length=94

 Score =  107 bits (266),  Expect = 3e-21, Method: Composition-based stats.
 Identities = 48/86 (55%), Positives = 66/86 (76%), Gaps = 0/86 (0%)

Query  312  LNNGGVANRPIKFTDNHASEKLDLGLVEFRQSDLKDDVGQSTVPFIDLRSIVSTGTALIG  371
            + +G   +RP  +TDN A+EK  +GL++  QS + DDVGQSTVPFIDL+ I ST +ALIG
Sbjct  2    ITDGKSGHRPTAYTDNVATEKTKIGLIQLSQSHMHDDVGQSTVPFIDLQPITSTESALIG  61

Query  372  VGLQFRTRDTSGGFIAPVVVTHPTYT  397
            +GL +R R+ SGGFIAP ++THPT++
Sbjct  62   LGLLYRGREFSGGFIAPSLITHPTFS  87


>☐ ref|NP_611782.2| UG CG30409 [Drosophila melanogaster]
 gb|AAF46984.2| G CG30409 [Drosophila melanogaster]
Length=98

 GENE ID: 37696 CG30409 | CG30409 gene product from transcript CG30409-RA
[Drosophila melanogaster] (10 or fewer PubMed links)

 Score = 70.5 bits (171),  Expect = 4e-10, Method: Composition-based stats.
 Identities = 15/68 (22%), Positives = 22/68 (32%), Gaps = 5/68 (7%)

Query  25  TYEEITRFNQVYFTQEANFG---SCNSKCEDYAAPRHNFKED--NRYKFPQRQCTGTIYN  79
           TY  +    QV+   E N     SC + C      R +        +   Q  C G +Y
Sbjct  30  TYGHVACLLQVFVENEINLSPDQSCMNCCGGCHDTRSSGCYKAEEEFCGQQAACKGRLYK  89

Query  80  CIGEKITE  87
           C   +
Sbjct  90  CNTIESES  97
```