PSI-Blast for AD-114 Anopheles darlingi salivary protein member of the 41.9 kDa family

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

Search Parameters

Program	blastp
Query range	0-405
Word size	3
Expect value	10
Hitlist size	500
Gapcosts	11,1
Matrix	BLOSUM62
Threshold	11
Composition-based stats	2
Window Size	40

Database

Posted date	Aug 19, 2008 12:36 PM
Number of letters	2,393,238,591
Number of sequences	6,929,940
Entrez query	none

Karlin-Altschul statistics

Params	Gapped	Ungapped
Lambda	0.317415	0.267
K	0.13269	0.0353583
H	0.386003	0.14

Results Statistics

Length adjustment	137
Effective length of query	269
Effective length of database	1443836811
Effective search space	388392102159
Effective search space used	388392102159

Distribution of 102 Blast Hits on the Query Sequence [?]

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

Sequences producing significant alignments:	Score (Bits)	E Value	
ref|XP_001648978.1| hypothetical protein AaeL_AAEL004382 [Aed...	346	2e-93	UG
gb|ABF18273.1| putative 41 kDa salivary secreted protein [Aed...	344	9e-93	
gb|ABF18024.1| putative 41 kDa salivary secreted protein [Aed...	341	5e-92	
ref|XP_001850512.1| 41.9 kDa basic salivary protein [Culex qu...	269	2e-70	UG
gb|AAR18425.1| putative 41.9 kDa basic salivary protein [Cule...	259	2e-67	
ref|YP_732608.1| hypothetical protein Shewmr4_0471 [Shewanell...	257	1e-66	G
ref|NP_800765.1| hypothetical protein VPA1255 [Vibrio parahae...	250	2e-64	G
ref|YP_001554416.1| hypothetical protein Sbal195_1985 [Shewan...	250	2e-64	G
gb|AAV90688.1| putative 41 kDa salivary secreted basic protei...	249	4e-64	
ref|XP_001648977.1| hypothetical protein AaeL_AAEL004353 [Aed...	235	5e-60	UG
gb|AAX54873.1| putative 30.3 kDa secreted protein [Aedes aegy...	221	9e-56	

gb|ABA43063.1|	46 kDa salivary protein SP19 [Phlebotomus pern...	207	1e-51
ref|XP_001850511.1|	salivary mucin [Culex quinquefasciatus] >...	202	4e-50
gb|AAX55661.1|	39 kDa salivary protein [Phlebotomus ariasi]	194	1e-47
gb|AAS16908.1|	37.2 kDa salivary protein [Lutzomyia longipalpis]	181	1e-43
gb|ABF18165.1|	putative salivary mucin [Aedes aegypti]	170	1e-40
ref|XP_310399.1|	AGAP003841-PA [Anopheles gambiae str. PEST] ...	167	2e-39
gb|AAU06485.1|	unknown salivary protein [Culicoides sonorensis]	165	7e-39
gb|AAR18443.1|	putative threonine/serine-rich mucin [Culex pi...	163	2e-38
ref|XP_001864883.1|	threonine/serine-rich mucin [Culex quinqu...	158	7e-37
ref|XP_001599762.1|	PREDICTED: similar to GA18227-PA [Nasonia...	148	6e-34
gb|ABI83794.1|	gSG10 salivary mucin [Anopheles funestus]	144	1e-32
ref|XP_001661479.1|	hypothetical protein AaeL_AAEL011183 [Aed...	131	1e-28
gb|ABJ97695.1|	putative threonine/serine-rich mucin [Toxorhyn...	108	1e-21
emb|CAC35527.1|	gSG9 protein [Anopheles gambiae]	76.5	4e-12

> ref|XP_001648978.1| hypothetical protein AaeL_AAEL004382 [Aedes aegypti]
gb|EAT44234.1| conserved hypothetical protein [Aedes aegypti]
Length=396

GENE ID: 5564658 AaeL_AAEL004382 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

Score = 346 bits (888), Expect = 2e-93, Method: Composition-based stats.
Identities = 104/421 (24%), Positives = 176/421 (41%), Gaps = 57/421 (13%)

```
Query  4    SMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLA  63
             +SV+   LI ++   PW+        PW  +   +I                  VT++
Sbjct  10   LLSVMGLFLIPSLEASPWF-------WPWQSRASPSI------------------VTIS  43

Query  64   LGNRINTDTALTDYGTNVNGLGINVNRRR---------------RRRRDIWNIESNPVLA  108
            +GNR+N++T +TD GT++  L  N++R +               R    + NI  +
Sbjct  44   IGNRVNSNT-VTDEGTSIRDLTFNLSRWKGNLTFEKNGTEILITRGNDGVTNIHIDFGSN  102

Query  109  GLAANN---PVNGSYIAIGNQVFHLPNNLSNVSLNIASGSSGGAG-QTVPSNFSTFLPMI  164
              +  N   P N +      ++F LP+ L  + +   +G+    G    P       P
Sbjct  103  SKSQQNETVPSNVTQSRKKREIFGLPS-LPAIPIPTLNGTQAAIGENFFPLPNEIANPAS  161

Query  165  -PDNYTQAFQKLVTQ-FQTLPLDAMDSIVQTLSDAYSNTFKGLFTAAGLNGLNETLEHGR  222
             N+ Q +Q L T   + LP+  +D+ V  LS  YS+T K +    G+N L ET++
Sbjct  162  NASNFLQPWQFLQTLITRRLPMQFLDTAVSYLSQLYSSTAKSMLDT-GVNTLAETMDTLS  220

Query  223  SVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNRK  282
            S++   +  I DQ A G   V  SF +     Q CVG    P    + VV K T CV  +
Sbjct  221  SIVNSTLDYIEDQTAAGIRGVAISFRELPQARQSCVGQ--VPDDAAQPVVTKATGCVRDR  278

Query  283  WRQLSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVR  342
            W +L  I  +    +  ++     +L  L  CN  NF   D +       +A +R CYV+
Sbjct  279  WNELDGIVNQFLVAVSDSEEAYGGWLRALDGCNARNFAGVDEAE-----LDAAQRSCYVQ  333

Query  343  AIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCASEMGLEIGLVTAQVSSKIAFCQ  402
             I    + ++ +P+       +  + +   Q  V LC + +G ++  V+A +  +I  CQ
Sbjct  334  NISKSVEKMVDIPIQWANLAARASSIVMQFQPQVGLCVAGVGAKVAAVSANIGVQIMLCQ  393

Query  403  L  403

Sbjct  394  A  394
```

```
>    gb|ABF18273.1|  putative 41 kDa salivary secreted protein [Aedes aegypti]
Length=396

 Score =  344 bits (881),  Expect = 9e-93, Method: Composition-based stats.
 Identities = 102/421 (24%), Positives = 174/421 (41%), Gaps = 57/421 (13%)

Query  4    SMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLA  63
             +SV+   LI ++   PW+        PW  +   +I                   VT++
Sbjct  10   LLSVMGLFLIPSLEASPWF-------WPWQSRASPSI------------------VTIS  43

Query  64   LGNRINTDTALTDYGTNVNGLGINVNRRR---------------RRRRDIWNIESNPVLA  108
            +GNR N++ ++TD GT++  L  N++R +               R    + NI  +
Sbjct  44   IGNRFNSN-SVTDEGTSIRDLTFNLSRWKGNLTFEKNGTEILITRGNDGVTNIHIDFGSN  102

Query  109  GLAANN---PVNGSYIAIGNQVFHLPNNLSNVSLNIASGSSGGAG-QTVPSNFSTFLPMI  164
              +  N   P N +      ++F LP+ L  + +   +G+    G            P
Sbjct  103  SKSQQNETVPSNATQSRKKREIFGLPS-LPAIPIPTINGTQAAIGENFFQLPNEIANPAS  161

Query  165  -PDNYTQAFQKLVTQ-FQTLPLDAMDSIVQTLSDAYSNTFKGLFTAAGLNGLNETLEHGR  222
               N+ Q +Q L T   + LP+  +D+ V  LS  YS+T K +    G+N L ET++
Sbjct  162  NASNFLQPWQFLQTLITRRLPMQFLDTAVSYLSQLYSSTAKNMLDT-GVNTLAETMDTLS  220

Query  223  SVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNRK  282
            S++   +  I DQ A G   V  SF +     Q CVG    P    + VV K T CV  +
Sbjct  221  SIVNSTLDYIEDQTAAGIRGVAISFRELPQARQSCVGQ--VPDDAAQPVVTKATGCVRDR  278

Query  283  WRQLSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVR  342
            W +L  I  +    +  ++     +L  L  CN  NF   D +       +A +R CYV+
Sbjct  279  WNELDGIVNQFLVAVSDSEEAYGGWLRALDGCNARNFAGVDEAE-----LDAAQRSCYVQ  333

Query  343  AIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCASEMGLEIGLVTAQVSSKIAFCQ  402
             I    + ++ +P+       +  + +   Q  V LC + +G ++  V+A +  +I  CQ
Sbjct  334  NISKSVEKMVDIPIQWANLAARASSIVMQFQPQVGLCVAGVGAKVAAVSANIGVQIMLCQ  393

Query  403  L  403

Sbjct  394  A  394


>    gb|ABF18024.1|  putative 41 kDa salivary secreted protein [Aedes aegypti]
Length=395

 Score =  341 bits (875),  Expect = 5e-92, Method: Composition-based stats.
 Identities = 104/421 (24%), Positives = 175/421 (41%), Gaps = 58/421 (13%)

Query  4    SMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLA  63
             +SV+   LI ++   PW+        PW  +   +I                   VT++
Sbjct  10   LLSVMGLFLIPSLEASPWF-------WPWQSRASPSI------------------VTIS  43

Query  64   LGNRINTDTALTDYGTNVNGLGINVNRRR---------------RRRRDIWNIESNPVLA  108
            +GNR+N++T +TD GT++  L  N++R +               R    + NI  +
Sbjct  44   IGNRVNSNT-VTDEGTSIRDLTFNLSRWKGNLTFEKNGTEILITRGNDGVTNIHIDFGSN  102

Query  109  GLAANN---PVNGSYIAIGNQVFHLPNNLSNVSLNIASGSSGGAG-QTVPSNFSTFLPMI  164
              +  N   P N +      ++F LP+ L  + +   +G+    G    P       P
Sbjct  103  SKSQQNETVPSNVTQSRKKREIFGLPS-LPAIPIPTLNGTQAAIGENFFPLPNEIANPAS  161

Query  165  -PDNYTQAFQKLVTQ-FQTLPLDAMDSIVQTLSDAYSNTFKGLFTAAGLNGLNETLEHGR  222
               N+ Q +Q L T   + LP+  +D+ V  LS  YS+T K +    G+N L ET++
Sbjct  162  NASNFLQPWQFLQTLITRRLPMQFLDTAVSYLSQLYSSTAKSMLDT-GVNTLAETMDTLS  220

Query  223  SVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNRK  282
            S++   +  I DQ A G   V  SF +     Q CVG    P    + VV K T CV  +
Sbjct  221  SIVNSTLDYIEDQTAAGIRGVAISFRELPQARQSCVGQ--VPDDAAQPVVTKATGCVRDR  278

Query  283  WRQLSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVR  342
            W +L  I  +    +  ++     +L  L  CN  NF   D +       +A +R CYV
Sbjct  279  WNELDGIVNQFLVAVSDSEEAYGGWLRALDGCNARNFAGVDEAE-----LDAAQRSCYV-  332

Query  343  AIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCASEMGLEIGLVTAQVSSKIAFCQ  402
             I    + ++ +P+       +  + +   Q  V LC + +G ++  V+A +  +I  CQ
Sbjct  333  NISKSVEKMVDIPIQWANLAARASSIVMQFQPQVGLCVAGVGAKVAAVSANIGVQIMLCQ  392

Query  403  L  403

Sbjct  393  A  393
```

> ref|XP_001850512.1| 41.9 kDa basic salivary protein [Culex quinquefasciatus]
gb|EDS32123.1| 41.9 kDa basic salivary protein [Culex quinquefasciatus]
Length=338

GENE ID: 6041220 CpipJ_CPIJ008900 | 41.9 kDa basic salivary protein [Culex quinquefasciatus]

Score = 269 bits (688), Expect = 2e-70, Method: Composition-based stats.
Identities = 82/322 (25%), Positives = 149/322 (46%), Gaps = 29/322 (9%)

```
Query  89   NRRRRRRR--DIWNIESNPVLAGLAANNPV-----NGSYIAIGNQVFHLPNNLSNVSLNI  141
            +RR+R      I  + + P +  + A   +     N +YI+IG ++  LP +++ ++ N
Sbjct  35   SRRKRETGFLPILPVTATPFVPSVPAAPAIPIPTMNSNYISIGGKLVGLPESITKLTANT  94

Query  142  ASGSSGGAGQTVPSNFSTFLPMIPDNYTQAFQKLVTQFQTLPLDAMDSIVQTLSDAYSNT  201
                S G G  V SN             Q +Q L T  +  P   +D +V  L+ +Y +T
Sbjct  95   ----SIGTGSAVQSN----------TMQQPWQFLKTLMERRPSKILDPVVDFLATSYLST  140

Query  202  FKGLFTAAGLNGLNETLEHGRSVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPD  261
             K +   A L  LN+T     ++    +    DQ + G + V  SFND +   + CVG
Sbjct  141  GKVMVDGA-LTALNKTEAFFYNIGNSSLNYATDQVSLGLQGVTRSFNDLTQMKRTCVG--  197

Query  262  LSPTSIGRRVVGKGTDCVNRKWRQLSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNA  321
             +P + GRRV+ K T CV  +  ++  I  +   ++ A +   S +L  +  CN  NF
Sbjct  198  GAPENAGRRVMEKATGCVRERLDEVIGIVDQFRTNVEATEGVFSGWLIEMQKCNAKNFPT  257

Query  322  ADASSGGTGTRNALRRQCYVRAIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCAS  381
               +       +A +R+CY  A+ +    ++ +P+  T    ++  A+++ +  + LCA
Sbjct  258  LADTE-----LDAAQRKCYAEALVSPTGKVVDIPLRWTTLASRISTAVTTFKAQLGLCAG  312

Query  382  EMGLEIGLVTAQVSSKIAFCQL  403
            ++ LE+  +  ++ SKI  C L
Sbjct  313  KVILEVAAIANELGSKITSCAL  334
```

> gb|AAR18425.1| putative 41.9 kDa basic salivary protein [Culex pipiens quinquefasciatus]
Length=412

Score = 259 bits (662), Expect = 2e-67, Method: Composition-based stats.
Identities = 82/321 (25%), Positives = 147/321 (45%), Gaps = 29/321 (9%)

```
Query  89   NRRRRRRR--DIWNIESNPVLAGLAANNPV-----NGSYIAIGNQVFHLPNNLSNVSLNI  141
            +RR+R      I  +   P +  + A   +     N +YI+IG ++  LP +++ ++ N
Sbjct  114  SRRKRETGFLPILPVTVTPFVPSVPAAPAIPIPTMNSNYISIGGKLVGLPESITKLTANT  173

Query  142  ASGSSGGAGQTVPSNFSTFLPMIPDNYTQAFQKLVTQFQTLPLDAMDSIVQTLSDAYSNT  201
                S G G  V SN             Q +Q L T  +  P   +D +V  L+ +Y +T
Sbjct  174  ----SIGTGSAVQSN----------TMQQPWQFLKTLMERRPSKILDPVVDFLATSYLST  219

Query  202  FKGLFTAAGLNGLNETLEHGRSVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPD  261
             K +   A L  LN+T     ++    +    DQ + G + V  SFND +   + CVG
Sbjct  220  GKVMVDGA-LTALNKTEAFFYNIGNSSLNYATDQVSLGLQGVTRSFNDLTQMKRTCVG--  276

Query  262  LSPTSIGRRVVGKGTDCVNRKWRQLSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNA  321
             +P + GRRV+ K T CV  +  ++  I  +   ++ A +   S +L  L  CN  NF
Sbjct  277  GAPENAGRRVMEKATGCVRERLDEVIGIVDQFRTNVEATEGVFSGWLTELQKCNAKNFPT  336

Query  322  ADASSGGTGTRNALRRQCYVRAIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCAS  381
               +       +A +R+CY  A+ +    ++ +P+  T    ++  A+++ +  + LCA
Sbjct  337  LADAE-----LDAAQRKCYAEALVSPTGKVVDIPLRWTTLASRISTAVTTFKAQLGLCAG  391

Query  382  EMGLEIGLVTAQVSSKIAFCQ  402
            ++ LE+  +  ++ SKI  C
Sbjct  392  KVILEVAAIANELGSKITSCA  412
```

> ref|YP_732608.1| hypothetical protein Shewmr4_0471 [Shewanella sp. MR-4]
gb|ABI37551.1| conserved hypothetical protein [Shewanella sp. MR-4]
Length=771

GENE ID: 4250806 Shewmr4_0471 | hypothetical protein [Shewanella sp. MR-4]

Score = 257 bits (656), Expect = 1e-66, Method: Composition-based stats.

```
 Identities = 51/349 (14%), Positives = 108/349 (30%), Gaps = 48/349 (13%)

Query  75   TDYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNGSYIAIGNQVFHLPNNL  134
             D  T +N         +   R   NI+++  L G   +  +N +  ++ N +    +
Sbjct  150  DDLDTEIN---------KFLTRTAKNIKTSLKLIGTEDSKYLNYNKTSLEN-LIRTSSAE  199

Query  135  SNVSLNIASGSSGGAGQTVPSNFSTFLPMIPDNYTQ---AFQKLVTQFQTLPLDAMDSIV  191
                 ++ +           S           N               + L   A++ ++
Sbjct  200  IVDPASVLTDEQITLHTKSASPVEKDKISFAANTISQHYFESSFNELTKLLSKSAVNEVI  259

Query  192  QTLSD------------AYSNTFKG--------LFTAAGLNGLNE--TLEH--GRSVIED  227
            + L D            +  +  K             + L  LN   + E    +  I +
Sbjct  260  EFLRDNPDIQTWVSSGMSVHDEHKSEKCHFCGNKIDESRLTSLNNHFSNEFKILKQEITE  319

Query  228  GIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNRKWRQLS  287
              +   D        V + F++     +  V P +   +  + +V K   C+  K
Sbjct  320  ATKYCPDLPELTLPSVESFFDEFQTEYKDAVAPLVGVVTEIKSIVDKWQVCIENKSNDPF  379

Query  288  DIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVRAIGNF  347
            DI  +I+   +      +  + ++ TC   + + ++     T +        Y
Sbjct  380  DISQKISSVTVDLIDQYNKVILDIATCVKKHNDKSENFQSVTTSHKKALELHYAA-----  434

Query  348  PQSLLFLPVSLTIEGGK-LYAALSSLQTDVALCASEMGLEIGLVTAQVS  395
               ++    S  IE  K  YA +  L  D+    S +  EI  + A++S
Sbjct  435  -DEVIDFEYSKKIEDRKTAYANVQKLSDDI----SAIMKEIQRLEAELS  478
```

> ref|NP_800765.1| hypothetical protein VPA1255 [Vibrio parahaemolyticus RIMD 2210633]
dbj|BAC62598.1| hypothetical protein [Vibrio parahaemolyticus RIMD 2210633]
Length=771

GENE ID: 1191951 VPA1255 | hypothetical protein
[Vibrio parahaemolyticus RIMD 2210633] (10 or fewer PubMed links)

```
 Score =  250 bits (637),  Expect = 2e-64, Method: Composition-based stats.
 Identities = 44/348 (12%), Positives = 108/348 (31%), Gaps = 48/348 (13%)

Query  76   DYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNGSYIAIGNQVFHLPNNLS  135
            D  T +N         +   R   +I+++  + G   +  +N +  ++ N +    +
Sbjct  151  DLDTEIN---------KFLSRTAKSIKTSLKVIGTEDSKYLNYNKTSLEN-LIRTSSAEV  200

Query  136  NVSLNIASGSSGGAGQTVPSNFSTFLPMIPDNYTQ---AFQKLVTQFQTLPLDAMDSIVQ  192
                ++ +           S         P +  +            + L   A++ +++
Sbjct  201  VDPASVLTDEQIILHTKSASPVEKDKISFPASTIRKDYFESSFNALTKLLGKSAVNEVIE  260

Query  193  TLSD------------AYSNTFKG--------LFTAAGLNGLNE--TLEH--GRSVIEDG  228
             L D            A  +  K             + L  LN   + E    +  I +
Sbjct  261  FLRDNPDVQTWVSSGMAVHDKHKSEECHFCGNKIDESRLTSLNNHFSNEFKLLKQEITEA  320

Query  229  IQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNRKWRQLSD  288
             +   D        V + F++     ++ V P     +    +V +  +C+++K     D
Sbjct  321  TKYCCDLPELTLPSVESFFDEFQTEYKKVVAPLEGVVTEINSIVEQWRECIDKKSNDPFD  380

Query  289  IGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVRAIGNFP  348
            I  +I+   +      ++ + ++  C   + + +      T          Y
Sbjct  381  ISLKISSVTVDLIDQYNNVILDIAACVKKHNDKSGNFQAVTTAHKKALELHYAA------  434

Query  349  QSLLFLPVSLTIEGGK-LYAALSSLQTDVALCASEMGLEIGLVTAQVS  395
              ++    S  +E  K     +  L  ++    S +  EI  + A++S
Sbjct  435  DEVIDFDYSKKVEDRKTANGNVLKLSDEI----SAIKKEIQRLEAELS  478
```

> ref|YP_001554416.1| hypothetical protein Sbal195_1985 [Shewanella baltica OS195]
gb|ABX49156.1| conserved hypothetical protein [Shewanella baltica OS195]
Length=771

GENE ID: 5753729 Sbal195_1985 | hypothetical protein [Shewanella baltica OS195]

```
 Score =  250 bits (637),  Expect = 2e-64, Method: Composition-based stats.
 Identities = 48/349 (13%), Positives = 110/349 (31%), Gaps = 48/349 (13%)

Query  75   TDYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNGSYIAIGNQVFHLPNNL  134
             D  T +N         +   R   +I++N  + G   +  +N +  ++ N +    +
Sbjct  150  DDLDTEIN---------KFLSRTAKSIKTNLKVIGTEDSKYLNYNKTSLEN-LIRSSSAE  199

Query  135  SNVSLNIASGSSGGAGQTVPSNFSTFLPMIPDNYTQ---AFQKLVTQFQTLPLDAMDSIV  191
```

```
                ++ +           S           N  +        V   Q L   A++ ++
Sbjct  200  IVDPASVLTDEQILLHTKSASPVEKDKISFASNTIKQDYFESGFVAVTQLLNKSAVNEVI  259

Query  192  QTLSD------------AYSNTFKG--------LFTAAGLNGLNE--TLEH--GRSVIED  227
            + L D            +  +  K         +   + L  LN   + E    +  I +
Sbjct  260  EFLKDNPDVQTWVFSGMSVHDKHKSEECHFCGNMIDESRLTSLNNHFSNEFKQLKQEITE  319

Query  228  GIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNRKWRQLS  287
             I+   D        V   F++   + ++ V P    T+    +V +   C+++K +
Sbjct  320  AIKYCRDLPELSLPVVELFFDEFRPKYKEAVAPLERVTTEINSIVEQWRKCIDKKSKDPF  379

Query  288  DIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVRAIGNF  347
            DI  +I    +      +  + ++  C   + + +      T T        Y
Sbjct  380  DISQKIQSVNIELIEQYNKVIIDIAACIQKHNDKSGDFQTVTTTHKNALELHYAA-----  434

Query  348  PQSLLFLPVSLTIEGGK-LYAALSSLQTDVALCASEMGLEIGLVTAQVS  395
               ++    +  +E      + +  L  +    ASE+  EI  +  ++S
Sbjct  435  -DEVIDFDYNEKLENRNTANSNVQKLNDE----ASEIKKEIHRLETELS  478
```

> gb|AAV90688.1| putative 41 kDa salivary secreted basic protein [Aedes albopictus]
Length=332

Score = 249 bits (635), Expect = 4e-64, Method: Composition-based stats.
Identities = 81/343 (23%), Positives = 135/343 (39%), Gaps = 59/343 (17%)

```
Query  14   AAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLALGNRINTDTA  73
             ++   PW+        PW  +   +I                   + +++GNR+NT+T+
Sbjct  20   PSLEASPWF-------WPWQSRSAPSI-------------------INISIGNRVNTNTS  53

Query  74   LTDYGTNVNGLGINVNRRR---------------RRRRDIWNIESNPVLAG---------  109
            + D GT++  L  N++R +               R    + NI  +
Sbjct  54   V-DEGTSIRDLTFNLSRWKGNLTFEKNGTEIFITRGNGGVTNIHIDFSSNSGNQRSSANR  112

Query  110  -LAANNPVNGSYIAIGNQVFHLPNNLSNVSLNIASGSSGG-AGQTVPSNFSTFLPMI-PD  166
                +N    S       +F LP+ L ++ +   +G+     G ++P   +  +P I
Sbjct  113  TTPLSNATQPSTSRRKRAIFGLPS-LPSIPIPTLNGTQAFLGGNSIPLPTAITIPAINAS  171

Query  167  NYTQAFQKLVTQ-FQTLPLDAMDSIVQTLSDAYSNTFKGLFTAAGLNGLNETLEHGRSVI  225
               Q +Q L T   + LP+  +D+ V  L+  YS+T K +    G+N L ET++   SV+
Sbjct  172  TLLQPWQFLQTLITRRLPMQLLDTAVSFLAQLYSSTAKSMLDT-GVNTLGETMDTLSSVV  230

Query  226  EDGIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNRKWRQ  285
               + +I DQ A G   V  SF   +   Q CVG    P    R VV K T CV  +W +
Sbjct  231  NSTLDTIEDQVASGIAGVAQSFKQLTRAGQSCVGQ--VPDDSARPVVTKATGCVRDRWNE  288

Query  286  LSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGG  328
            L  I  +    +   +     +L  L  CN  NF     +  G
Sbjct  289  LDGIVNQFLVAVSDTEDAYGGWLRALDACNARNFAGLSDTGAG  331
```

> ref|XP_001648977.1| hypothetical protein AaeL_AAEL004353 [Aedes aegypti]
gb|EAT44233.1| hypothetical protein AaeL_AAEL004353 [Aedes aegypti]
Length=308

GENE ID: 5564644 AaeL_AAEL004353 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

Score = 235 bits (599), Expect = 5e-60, Method: Composition-based stats.
Identities = 38/269 (14%), Positives = 89/269 (33%), Gaps = 19/269 (7%)

```
Query  141  IASGSSGGAGQTVPSNFSTFLPMIPDNYTQAFQKL--VTQFQTLPLDAMDSIVQTLSDAY  198
            + S  +   G    +N S     +  N      +            + ++ ++ T+ + Y
Sbjct  49   VPSMYNISIGNRENTNASIDYGTMITNLQINLARFNGSALLGDATENRLNEVIYTVLEGY  108

Query  199  SNTFKGLFTAAGLNGLNETLEHGRSVIEDGIQSIVDQGAQGFEQ----VIASFNDSSARV  254
             N  + +  +     L    E    +++   Q   D+  +  +Q    V  S    +  V
Sbjct  109  RNRSEAMVDSR----LAWIDEVLLDIVQYARQLRQDENNRWLKQFPEHVRTSVAQLNETV  164

Query  255  QQCVGPDLSPTSIGRRVVGKGT-DCVNRKWRQLSDIGGRIAGDILAADRGASDFLGNLTT  313
            + C+  ++    +   V  +    C+  + ++L ++      ++      A D    L T
Sbjct  165  RSCLEREVVVEELIHAVENRSRSGCLEGRLQKLFELREVAKNNLTEFLASAGDLEDRLET  224

Query  314  CNGANFNAADASSGGTGTRNALRRQCYVRAIGNFPQSLLFLPVSLTIEGGKLYAALSSLQ  373
            C   +F+ A          N  +  C    +       L L  +++   G +  AL   +
Sbjct  225  CVNPDFDDALD--------NVYQEACISSILLEVEMESLKLGYTVSNLTGAMEPALGLAK  276
```

```
Query  374  TDVALCASEMGLEIGLVTAQVSSKIAFCQ  402
              +  CA+++       +  +   I  C
Sbjct  277  AVLLECAADLAWYAFDASMSLRQWINSCS  305
```

Score = 68.0 bits (164), Expect = 1e-09, Method: Composition-based stats.
Identities = 17/63 (26%), Positives = 30/63 (47%), Gaps = 1/63 (1%)

```
Query  28  FTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLALGNRINTDTALTDYGTNVNGLGIN  87
           F   W      T       + ++ ++    S   +++GNR NT+ ++ DYGT +  L IN
Sbjct  21  FGWSWFFPRQTTTTTASSGTTSATSSNAVPSMYNISIGNRENTNASI-DYGTMITNLQIN  79

Query  88  VNR  90
           + R
Sbjct  80  LAR  82
```

> gb|AAX54873.1| putative 30.3 kDa secreted protein [Aedes aegypti]
Length=295

Score = 221 bits (562), Expect = 9e-56, Method: Composition-based stats.
Identities = 66/307 (21%), Positives = 110/307 (35%), Gaps = 54/307 (17%)

```
Query  4    SMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLA  63
             +SV+   LI ++   PW+        PW  +   +I                   VT++
Sbjct  10   LLSVMGLFLIPSLEASPWF-------WPWQSRASPSI------------------VTIS  43

Query  64   LGNRINTDTALTDYGTNVNGLGINVNRRR---------------RRRRDIWNIESNPVLA  108
            +GNR N++ ++TD GT++  L  N++R +                R    + NI  +
Sbjct  44   IGNRFNSN-SVTDEGTSIRDLTFNLSRWKGNLTFEKNGTEILITRGNDGVTNIHIDFGSN  102

Query  109  GLAANN---PVNGSYIAIGNQVFHLPNNLSNVSLNIASGSSGGAG-QTVPSNFSTFLPMI  164
              +  N   P N +      ++F LP+ L  + +   +G+    G            P
Sbjct  103  SKSQQNETVPSNATQSRKKREIFGLPS-LPAIPIPTINGTQAAIGENFFQLPNEIANPAS  161

Query  165  -PDNYTQAFQKLVTQ-FQTLPLDAMDSIVQTLSDAYSNTFKGLFTAAGLNGLNETLEHGR  222
              N+ Q +Q L T   + LP+  +D+ V  LS  YS+T K +     L  L ET+ H
Sbjct  162  NASNFLQPWQFLQTLITRRLPMQFLDTAVSYLSQLYSSTAKNMLGHRELTPLAETMGHPS  221

Query  223  SVIEDGI-QSIVDQ-GAQGFEQVIASFND-SSARVQQCVGPDLSPTSIGRRVVGKG--TD  277
            V    +  +  DQ   +        F        Q CV P          +V K
Sbjct  222  PVSSTPLWTTSRDQTVGRNPRGSHIPFRGTFLKLAQSCVWPRCPDECCFSPIVTKAHRDG  281

Query  278  CVNRKWR  284
                +W
Sbjct  282  VTGIRWE  288
```

> gb|ABA43063.1| 46 kDa salivary protein SP19 [Phlebotomus perniciosus]
Length=431

Score = 207 bits (527), Expect = 1e-51, Method: Composition-based stats.
Identities = 39/301 (12%), Positives = 95/301 (31%), Gaps = 34/301 (11%)

```
Query  92   RRRRRDIWNIESNPVLAGLAANNPVNGSYIAIGNQVFHLPNNLSNVSLNIASGSSGGAGQ  151
            ++ +  +      P      ++N  + +       +         + ++    +    G+
Sbjct  143  KKLKGMLKPKPQTPSGNQTDSSNTTSETQSRKKRALTDF------IPMDSLKDAISKTGE  196

Query  152  TVPSNFSTFLPMIPDNYTQAFQKLVTQFQTLPLDAMDSIVQTLSDAYSNTFKGLFT-AAG  210
             +  + ++      D       KL      L  ++M  I + L+ +       ++   +
Sbjct  197  VLIPSSASANSSPLD----FMSKLSDIANDLIQNSMKEISENLASSV-----AMYQVNSQ  247

Query  211  LNGLNETLEHGRSVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQ--CVGPDLSPTSIG  268
            L+ + ++++  +  I+   Q I     +   Q   +      +++   C    ++P  +
Sbjct  248  LDAIKQSMDIIKQEIDKT-QKIQKYVKEALNQAKNATKSLGEKLKSSNCFAQFINPFKLF  306

Query  269  RRVVGKGTDCVNRKWRQLSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGG  328
                 KG  CV  K       I      ++  A    SD    ++ C          +
Sbjct  307  ----EKGITCVKNKIDNGLKIAKDTFKNLQQAMSVPSDIQSEVSKC--------SQNQQL  354

Query  329  TGTRNALRRQCYVRAIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCASEMGLEIG  388
                   +  CY+R         L LP   T    ++    ++++ D+  C  E    IG
Sbjct  355  NPI---AKLLCYLRTPLQLDDEKLLLPFEFTRRIREITNYFATMRMDLIRCGIETIQSIG  411

Query  389  L  389
```

```
Sbjct  412  D  412
```

> ref|XP_001850511.1| salivary mucin [Culex quinquefasciatus]
gb|EDS32122.1| salivary mucin [Culex quinquefasciatus]
Length=317

GENE ID: 6041218 CpipJ_CPIJ008899 | salivary mucin [Culex quinquefasciatus]

Score = 202 bits (513), Expect = 4e-50, Method: Composition-based stats.
Identities = 30/275 (10%), Positives = 72/275 (26%), Gaps = 19/275 (6%)

```
Query  140  NIASGSSGGAGQTVPSNFS-TFLPMIPDNYTQA---FQKLVTQFQTLPLDAMDSIVQTLS  195
            N  +      G    +  + T    +  N       F              +D  V +++
Sbjct  53   NTPNNYQLSIGNRQNTFANLTDYGTMISNVNINLARFNGTAASGNDAGATLLDDAVLSVA  112

Query  196  DAYSNTFKGLFTAAGLNGLNETLEHGRSVIEDGI----QSIVDQGAQGFEQVIASFNDSS  251
            +AY N    +        L    +               Q I D          A     S
Sbjct  113  EAYRNESTAVIAGR----LAWFNDQIAQANRSASTVHHQHISDLVEDFARDTRAGTLGLS  168

Query  252  ARVQQCVGPDLSPTSIGRRVVGKGTD-CVNRKWRQLSDIGGRIAGDILAADRGASDFLGN  310
               + C+   +    + R V  +    C+  K +++ ++    + ++        +
Sbjct  169  NATRDCLARSVQVEDVIRSVENRSEGGCLRAKIQRMLELRAEASSNLTEFLESQQNVEDR  228

Query  311  LTTCNGANFNAADASSGGTGTRNALRRQCYVRAIGNFPQSLLFLPVSLTIEGGKLYAALS  370
            L  C                  +  +  C    + +       L ++      +    +
Sbjct  229  LEICVD------LQDDFDDDMSDFYKVACVSSILFDVQMETAKLELTAEELTAEAGVTVR  282

Query  371  SLQTDVALCASEMGLEIGLVTAQVSSKIAFCQLAQ  405
             L+  +  C  ++       + ++   I  C  A+
Sbjct  283  QLRAGLLECVVDVANYAFDASMKLRHWINVCSAAR  317
```

> gb|AAX55661.1| 39 kDa salivary protein [Phlebotomus ariasi]
Length=375

Score = 194 bits (492), Expect = 1e-47, Method: Composition-based stats.
Identities = 33/295 (11%), Positives = 88/295 (29%), Gaps = 32/295 (10%)

```
Query  100  NIESNPVLAGLAANNPVN-GSYIAIGNQVFHLPNNLSNVSLNIASGSSG-GAGQTVPSNF  157
             +++ P      ++N  +  +       +         + +N    +        + +
Sbjct  89   FLKTKPSGNQTESSNSTSTKTQSRKRRGLTDF------IPVNSLKDAISQATSGAMKAFK  142

Query  158  STFLPMIPDNYTQAFQKLVTQFQTLPLDAMDSIVQTLSDAYSNTFKGLFT-AAGLNGLNE  216
             +       N       L    + L  +++  +   L  +       L+   + L+ + +
Sbjct  143  PSSENKTSSNPLDFLASLSDISRDLVQNSIKEVSGNLVSSV-----ALYQVNSKLDAIKQ  197

Query  217  TLEHGRSVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQ--CVGPDLSPTSIGRRVVGK  274
            ++      I+   + +        +Q     N    +++   C    ++P  +   V+
Sbjct  198  SIGIINQEIDRT-KKVQQYVMNALQQASNITNSIGEQLKSNNCFAQFINPFKLFEEVIT-  255

Query  275  GTDCVNRKWRQLSDIGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNA  334
               CV  K      I      ++  A    SD +  ++ C          +
Sbjct  256  ---CVKNKIENGLKIAEETFKNLNQALSVPSDIVSEVSKC--------SQNQNLNPLT--  302

Query  335  LRRQCYVRAIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCASEMGLEIGL  389
             +  CY+R      +  L LP+       ++    ++++ D+  C       IG
Sbjct  303  -KLLCYLRVPLQLDEEKLLLPIEFARRIREITNYFATMRMDLIQCGIATIQSIGD  356
```

> gb|AAS16908.1| 37.2 kDa salivary protein [Lutzomyia longipalpis]
Length=350

Score = 181 bits (458), Expect = 1e-43, Method: Composition-based stats.
Identities = 38/265 (14%), Positives = 87/265 (32%), Gaps = 30/265 (11%)

```
Query  130  LPNNLSNVSLNIA--SGSSGGAGQTVPSNFST-FLPMIPDNYTQAFQKLVTQFQTLPL--  184
            LP+  S  S+     S +       +    S   L  +P+      ++    +  +
Sbjct  108  LPSMPSMPSIPSLWSSQTQAAPNTALALPESDYSLLDMPNIVKNFLKETRDLYNDVGAFL  167

Query  185  DAMDSIVQTLSDAYSNTFKGLFTAAGLNGLNETLEHGRSVIEDGIQSIVDQGAQGFEQVI  244
             A+   +   S +       + +       N+T E  R+ I+   + + +   +  +++
Sbjct  168  KAITEALTNRSSSSQLLSSPMVST------NKTKEFIRNEIQKVRK-VRNFVQETLQKIR  220

Query  245  ASFNDSSARVQQ--CVGPDLSPTSIGRRVVGKGTDCVNRKWRQLSDIGGRIAGDILAADR  302
```

```
                 + +V+   C+          + +V  G +C+  K+     I  ++  ++L   +
Sbjct  221  DISAAIAKKVKSSECLSNLTD----IKGLVSDGINCLKEKFNDGKRIILQLYNNLLKGLK  276

Query  303  GASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVRAIGNFPQSLLFLPVSLTIEG  362
             +D +  L  C+    N               R  CY        +  + LPV
Sbjct  277  IPNDLMVELKKCDTNQNNTLG------------RIICYFLTPLQLEKEQILLPVEFIKRI  324

Query  363  GKLYAALSSLQTDVALCASEMGLEI  387
             +L    S+++ D+  C       I
Sbjct  325  LELTHYFSTMKEDLINCGITTIASI  349
```

```
>    gb|ABF18165.1|  putative salivary mucin [Aedes aegypti]
Length=248

 Score =  170 bits (431),  Expect = 1e-40, Method: Composition-based stats.
 Identities = 29/210 (13%), Positives = 68/210 (32%), Gaps = 19/210 (9%)

Query  141  IASGSSGGAGQTVPSNFSTFLPMIPDNYTQAFQKL--VTQFQTLPLDAMDSIVQTLSDAY  198
            + S  +   G    +N S     +  N      +            + ++ ++ T+ + Y
Sbjct  49   VPSMYNISIGNRENTNASIDYGTMITNLQINLARFNGSALLGDATENRLNEVIYTVLEGY  108

Query  199  SNTFKGLFTAAGLNGLNETLEHGRSVIEDGIQSIVDQGAQGFEQ----VIASFNDSSARV  254
             N  + +  +     L    E    +++   Q   D+  +  +Q    V  S    +  V
Sbjct  109  RNRSEAMVDSR----LAWIDEVLLDIVQYARQLRQDENNRWLKQFPEHVRTSVAQLNETV  164

Query  255  QQCVGPDLSPTSIGRRVVGKGT-DCVNRKWRQLSDIGGRIAGDILAADRGASDFLGNLTT  313
            + C+  ++    +   V  +    C+  + ++L ++      ++      A D    L T
Sbjct  165  RSCLEREVVVEELIHAVENRSRSGCLEGRLQKLFELREVAKNNLTEFLASAGDLEDRLET  224

Query  314  CNGANFNAADASSGGTGTRNALRRQCYVRA  343
            C   +F+ A          N  +  C
Sbjct  225  CVNPDFDDALD--------NVYQEACISSV  246


 Score = 67.2 bits (162),  Expect = 2e-09, Method: Composition-based stats.
 Identities = 17/63 (26%), Positives = 30/63 (47%), Gaps = 1/63 (1%)

Query  28  FTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLALGNRINTDTALTDYGTNVNGLGIN  87
           F   W      T       + ++ ++    S   +++GNR NT+ ++ DYGT +  L IN
Sbjct  21  FGWSWFFPRQTTTTTASSGTTSATSSNAVPSMYNISIGNRENTNASI-DYGTMITNLQIN  79

Query  88  VNR  90
           + R
Sbjct  80  LAR  82
```

```
>    ref|XP_310399.1|  AGAP003841-PA [Anopheles gambiae str. PEST]
 emb|CAC35525.1|  hypothetical protein [Anopheles gambiae]
 gb|EAA05987.2|  AGAP003841-PA [Anopheles gambiae str. PEST]
Length=195

 GENE ID: 1271577 AgaP_AGAP003841 | AGAP003841-PA [Anopheles gambiae str. PEST]
(10 or fewer PubMed links)

 Score =  167 bits (422),  Expect = 2e-39, Method: Composition-based stats.
 Identities = 50/195 (25%), Positives = 80/195 (41%), Gaps = 9/195 (4%)

Query  1    MNRSMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQ--PPRTINQLPVTSATSATTPTQQS  58
            M +   + + LL AA+ + P       +  PW+     P T      +   +++  +  S
Sbjct  1    MLQMRIISVPLLSAAVLLQPLLAAPIFWFLPWSFPALSPTTTTLATTSGTAASSGASNSS  60

Query  59   NVTLALGNRINTDTALTDYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNG  118
            NV++A+GNR+NT T L DYGTN+  + INVNR RR   D    +    L   A     NG
Sbjct  61   NVSVAIGNRVNTSTGLDDYGTNITNVQINVNRMRRAVTDADLAKLERKLRQAADEGSTNG  120

Query  119  SYIAIGNQVFHLPNNLSNVSLNI--ASGSS-----GGAGQTVPSNFSTFLPMIPDNYTQA  171
            + I IG     LP+N+SN+++N    +G            +   N +     +
Sbjct  121  TVITIGEHTIRLPHNISNLTVNTYLINGIPENYMMLADTGSATINANETYTPLLSAIRLF  180

Query  172  FQKLVTQFQTLPLDA  186
                 + F
Sbjct  181  KNFFSSLFSKQGSQT  195
```

```
>    gb|AAU06485.1|  unknown salivary protein [Culicoides sonorensis]
Length=243

 Score =  165 bits (417),  Expect = 7e-39, Method: Composition-based stats.
 Identities = 45/232 (19%), Positives = 77/232 (33%), Gaps = 14/232 (6%)

Query  175  LVTQFQTLPLDAMDSIVQTLSDAYSNTFKGLFTA--AGLNGLNETLEHGRSVIEDGIQSI  232
            +     +L    ++S +  L    ++  K       + L+GL        + I    +S
Sbjct  18   VSNLGNSLSNGTVESALDNLQSKVNDAAKRANDTTNSALSGLQSVGNTLSNEINKVSKSA  77

Query  233  VDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIG-RRVVGKGTDCVNRKWRQLSDIGG  291
              Q     + + +         + CV  + +P S   +        CV  K  +   I
Sbjct  78   EAQVNSTMKGLSSDI-------RDCVNQNGNPGSAALKTAKTSAQKCVQDKIDEAKSIIN  130

Query  292  RIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCYVRAIGNFPQSL  351
                DI  A  G     GNL+ C+         +     T NA +  C   A+ +
Sbjct  131  GTRNDIQQAKLGVDYLRGNLSNCHLN----VSMNLASLPTFNASQPACLASALLSVQSET  186

Query  352  LFLPVSLTIEGGKLYAALSSLQTDVALCASEMGLEIGLVTAQVSSKIAFCQL  403
            L +P+SL   G +    +  LQ     C + +  EI   T   S  +A C
Sbjct  187  LLIPLSLASHGAQAVGLVQGLQGYAMKCMANLMEEIAKATLSTSLSVAKCYA  238
```

```
>    gb|AAR18443.1|  putative threonine/serine-rich mucin [Culex pipiens quinquefasciatus]
Length=173

 Score =  163 bits (412),  Expect = 2e-38, Method: Composition-based stats.
 Identities = 48/183 (26%), Positives = 85/183 (46%), Gaps = 13/183 (7%)

Query  1    MNRSMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQ--QS  58
            M R  + ++ ++++A+           +  PW+     +      +  +S +T T    S
Sbjct  1    MTRVQTYLIVVIVSAL-FSTTLGSPIFWLLPWSSWLNPSTTSSTSSGTSSTSTATARTPS  59

Query  59   NVTLALGNRINTDTALTDYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNG  118
            NV++A+G+R NT   L D GT ++GL IN  RRRR      + +  P +  +    P N
Sbjct  60   NVSIAIGDRTNTTAGLDDMGTVIDGLTINAGRRRR------SAQQLPEVDFIV---PSNE  110

Query  119  SYIAIGNQVFHLPN-NLSNVSLNIASGSSGGAGQTVPSNFSTFLPMIPDNYTQAFQKLVT  177
            SY+ IGN V  LP+ N+SN+++N+  G        +P+      P + D+    F  +
Sbjct  111  SYVQIGNHVIRLPSRNISNITINVVDGVPYVGAGGLPAPPQDVPPNLFDHIKNFFDGVFN  170

Query  178  QFQ  180
            +
Sbjct  171  RTN  173
```

```
>    ref|XP_001864883.1|  threonine/serine-rich mucin [Culex quinquefasciatus]
 gb|EDS40846.1|  threonine/serine-rich mucin [Culex quinquefasciatus]
Length=173

 GENE ID: 6047887 CpipJ_CPIJ014439 | threonine/serine-rich mucin
[Culex quinquefasciatus]

 Score =  158 bits (399),  Expect = 7e-37, Method: Composition-based stats.
 Identities = 48/183 (26%), Positives = 84/183 (45%), Gaps = 13/183 (7%)

Query  1    MNRSMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATT--PTQQS  58
            M R  + ++ ++++A+           +  PW+     +      +  +S +T      S
Sbjct  1    MTRVQTYLIVVIVSAL-FSTTLGSPIFWLLPWSSWLSPSTTSSTSSGTSSTSTATARTPS  59

Query  59   NVTLALGNRINTDTALTDYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNG  118
            NV++A+G+R NT   L D GT ++GL IN  RRRR      + +  P +  +    P N
Sbjct  60   NVSIAIGDRTNTTAGLDDMGTVIDGLTINAGRRRR------SAQQPPEVDFIV---PSNE  110

Query  119  SYIAIGNQVFHLPN-NLSNVSLNIASGSSGGAGQTVPSNFSTFLPMIPDNYTQAFQKLVT  177
            SYI IGN V  LP+ N+SN+++N+  G        +P+      P + D+    F  +
Sbjct  111  SYIQIGNHVIRLPSRNISNITINVVDGVPYVGAGGLPAPPQDVPPNLFDHIKNFFDGVFN  170

Query  178  QFQ  180
            +
Sbjct  171  RTN  173
```

```
>    ref|XP_001599762.1|  PREDICTED: similar to GA18227-PA [Nasonia vitripennis]
```

Length=2301

GENE ID: 100114886 LOC100114886 | similar to GA18227-PA [Nasonia vitripennis]

Score = 148 bits (374), Expect = 6e-34, Method: Composition-based stats.
Identities = 28/226 (12%), Positives = 57/226 (25%), Gaps = 20/226 (8%)

```
Query  162   PMIPDNYTQAFQKLVTQFQTLPLDAMDSIVQTLSDAYSNTFKGLFTAAGLNGLNETLEHG  221
                     Q     + +     ++ +D   Q +SD    T K +        ++ET +
Sbjct  1968  TNEISKVQQIESSFINKTNEAVVNTIDETKQAVSDHMDETHKIVSDQ-----IDETNKTI  2022

Query  222   RSVIEDGIQSIVDQGAQGFEQVIASFNDSSARVQQCVGPDLSPTSIGRRVVGKGTDCVNR  281
                I D  +   DQ  +    V    +++   V+                + +    V
Sbjct  2023  SEEINDTKKVASDQIDKTKRSVTDRIDETKKNVKD--------------TIDETKKMVAE  2068

Query  282   KWRQLSD-IGGRIAGDILAADRGASDFLGNLTTCNGANFNAADASSGGTGTRNALRRQCY  340
             K  +    I  +I   I  +D                     + +S       +
Sbjct  2069  KIDETKKSISDKIQRSISKSDEDGQTIDAEKQVSAEKKEGTSISSVDRILETSKYLATNL  2128

Query  341   VRAIGNFPQSLLFLPVSLTIEGGKLYAALSSLQTDVALCASEMGLE  386
             V    +     +        E     +    L++       EM  E
Sbjct  2129  VDTTVDASNKAVLAAQEKAEEVENALSNTDLLKSVSTSLPDEMKEE  2174
```

> gb|ABI83794.1| gSG10 salivary mucin [Anopheles funestus]
Length=190

Score = 144 bits (363), Expect = 1e-32, Method: Composition-based stats.
Identities = 49/190 (25%), Positives = 76/190 (40%), Gaps = 14/190 (7%)

```
Query  3    RSMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPP----RTINQLPVTSATSATTPTQQS  58
            +   ++  LL  A+   P       +  PW+         T +     +A+S +  +  S
Sbjct  1    QIRIIIAPLLSVAVLTQPLLAAPIFWFLPWSFPNTFTTSSTTSSSTSGTASSTSGASNSS  60

Query  59   NVTLALGNRINTDTALTDYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNG  118
            NV++A+GNR+NT T L DYGTN+  + INVNR RR   +    +     A   A    N
Sbjct  61   NVSVAIGNRVNTSTGLDDYGTNITNVQINVNRMRRAITENDLAKLERRYARQTA----NE  116

Query  119  SYIAIGNQVFHLPNNLSNVSLNI--ASGSSGGAGQ----TVPSNFSTFLPMIPDNYTQAF  172
            S I IG     LP N+SN+++N    +G            +  N       +
Sbjct  117  STITIGEHTIRLPANISNLTVNTYLINGIPDSTMNLPDTGMTINTDEASTPLLTAVNIFK  176

Query  173  QKLVTQFQTL  182
              L + F
Sbjct  177  NFLSSFFGKQ  186
```

> ref|XP_001661479.1| hypothetical protein AaeL_AAEL011183 [Aedes aegypti]
gb|ABF18096.1| putative mucin [Aedes aegypti]
gb|EAT36759.1| conserved hypothetical protein [Aedes aegypti]
Length=180

GENE ID: 5574465 AaeL_AAEL011183 | hypothetical protein [Aedes aegypti]
(10 or fewer PubMed links)

Score = 131 bits (328), Expect = 1e-28, Method: Composition-based stats.
Identities = 50/181 (27%), Positives = 85/181 (46%), Gaps = 17/181 (9%)

```
Query  5    MSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQ--QSNVTL  62
             S+V+AL + A+   P  +   +F  PW+   P T +     +A+S+++ T        +
Sbjct  6    FSIVVALTLCAVL--PSTLGVPIFWFPWSYLLPTTSSSSASGTASSSSSSTTLRPIMANI  63

Query  63   ALGNRINTDTALTDYGTNVNGLGINVNRRRRRRRDIWNIESNPVLAGLAANNPVNGSYIA  122
            ++GNR NT   + D GT +NGL I+     R RRD+  +E+      +      + S+I
Sbjct  64   SIGNRRNTSAMV-DSGTVINGLTIS-----RGRRDV-PLETLRNDPSV-----TSESFIQ  111

Query  123  IGNQVFHL-PNNLSNVSLNIASGSSGGAGQTVPSNFSTFLPMIPDNYTQAFQKLVTQFQT  181
            +GNQV  L P N++N+++NI  G     G T     S       +  +  F ++   F
Sbjct  112  VGNQVIRLPPGNVANLTINIIDGVPYVFGNTESDVSSGSATGSSEGMSGLFSRVRNFFGD  171

Query  182  L  182
            L
Sbjct  172  L  172
```

```
>    gb|ABJ97695.1|  putative threonine/serine-rich mucin [Toxorhynchites amboinensis]
Length=172

 Score =  108 bits (268),  Expect = 1e-21, Method: Composition-based stats.
 Identities = 42/158 (26%), Positives = 66/158 (41%), Gaps = 18/158 (11%)

Query  34   PQPPRTINQLPVTSATSATTPTQQSNVTLALGNRINTDTALTDYGTNVNGLGINVNRRRR  93
               P   +    TS+TS++T     N T++LGNR NT   + D GTN++GL I  +R +R
Sbjct  17   FGMPTYTSGTSSTSSTSSSTARSPVNATISLGNRENTSA-VEDLGTNIDGLTI--SRSKR  73

Query  94   -------RRRDIWNIESNPVL----AGLAANNP--VNGSYIAIGNQVFHL-PNNLSNVSL  139
                          ++ P +       A + P   N S I +GNQ+  L P  +SN+++
Sbjct  74   FVPLLIPNVTGFSFFQNLPGIFTRQKRAAESIPPLSNESVIRVGNQIIKLPPRGISNLTI  133

Query  140  NIASGSSGGAG-QTVPSNFSTFLPMIPDNYTQAFQKLV  176
            +I  G S         +  +       D     F  L
Sbjct  134  HIIDGGSYSEKFTPFQTPANGTWAPFVDQIKNFFDGLF  171


>    emb|CAC35527.1|  gSG9 protein [Anopheles gambiae]
Length=76

 Score = 76.5 bits (186),  Expect = 4e-12, Method: Composition-based stats.
 Identities = 31/78 (39%), Positives = 44/78 (56%), Gaps = 10/78 (12%)

Query  4   SMSVVLALLIAAIGIIPWYVDAFLFTTPWNPQPPRTINQLPVTSATSATTPTQQSNVTLA  63
           S  +V+ ++       P++   F +  P+  QP   +       A++A    Q+SNVT+A
Sbjct  9   STLLVVVVIAPRHQCSPFF---FQYNRPYLSQPSSQL-------ASTAANVVQRSNVTVA  58

Query  64  LGNRINTDTALTDYGTNV  81
           LGNRINTDTAL DYGT V
Sbjct  59  LGNRINTDTALDDYGTRV  76
```