Generation of a gene model for *Aedes aegypti* perilipin-1 gene and a probe to investigate number and size of the transcripts in mosquito tissues

John Adams

Biochemistry and Molecular Biology

Abstract

The mosquito *Aedes aegypti* is a known vector of disease for humans. Lipid metabolism is of key importance for survival and reproduction. Perilipin-1, PLIN1, is an important protein in the mobilization and storage of lipids in insects. Though the genome of *A. aegypti* has been sequenced, the gene structures must be confirmed by independent methods. Similarly, the nature of the transcripts from the gene must also be confirmed. From the available genomic and transcriptomic information, a cDNA clone was produced and used to obtain antibodies for western blot assays. Preliminary studies in which protein size was ascertained showed a protein much larger than was predicted by the reported sequences and the clones produced in the lab. Further studies which included cloning and sequencing led to the development of cDNA predicting a protein of the size found in the western blot. Using this information and the newly sequenced *Aedes albopictus* genome, a structural model of the PLIN1 gene has been constructed and a northern blot probe designed and produced to determine the number and size of the transcripts.

Introduction

Aedes aegypti is a vector associated with transmission of disease such as dengue, yellow fever, and chikungunya virus which affects millions of people annually[1]. There are many possible targets of insecticides which may someday include agents which inhibit lipid metabolism. However, a thorough understanding of lipid metabolism must be ascertained before this can occur. Of critical importance to understanding this biological process is characterizing the proteins associated with it. One of these proteins is PLIN1, which may also be referred to as lipid storage droplet protein 1 (LSD1). This protein is found in association with cytosolic lipid droplets, which in insects are made up of large accumulations of triglycerides[2,3]. Furthermore, triglycerides are the primary means by which fatty acids are stored. Fatty acids are the main source of energy for embryonic development and are a major component of eggs. Therefore, understanding the mechanisms by which fatty acids are stored and mobilized is of critical importance[3]. Previous studies conducted by the lab using the model insect *Manduca sexta* indicate that PLIN1 plays a central role in regulating fatty acid storage and mobilization. The genetic expression of PLIN1[4] and the modulation of lipase activity via phosphorylation[5] appear to be a critical factors in regulating the mobilization and storage of fatty acids. This protein has also been demonstrated to play a role in regulating fatty acid metabolism during wing development in *Drosophila melonagaster*[6]. The present study is part of effort directed to elucidate the function of PLIN1 in *Aedes aegypti*. To this end, an Aedes aegypti PLIN1 cDNA cloned in the lab was used to develop an RNA probe to determine the size and number of transcripts at various stages of the life cycle of *Aedes aegypti* by northern blot assays. We have

also generated a structural model of the PLIN1gene and compared it to the gene in the recently sequenced *Aedes albopictus* genome.

Materials and Methods

Preliminary studies. According to vectorbase.org, which hosts the genome information of Aedes aegypti, the Lsd1gene has 2 predicted transcripts. Lsd1 A (AAEL005951-RA) of length 714 base pairs and Lsd1 B (AAEL005951-RB) of length 1598 base pairs that encode proteins of 24kDa (211 aa) and 28kDa (245 aa), respectively (both transcripts are annotated as partial mRNA). At the time of these studies an additional transcript of 906 base pairs, AET-5405, encoding a protein of 34kDa (301 aa) protein was also reported in NCBI. This sequence was used to clone a cDNA of PLIN1 and express the recombinant protein in *E. coli*. The protein was then used to produce an antibody useful for western blotting. The immuno-blots of samples of *A. aegypti* fat body homogenates revealed a protein of 46kDa suggesting that AET-5405 was also a partial mRNA sequence. ESTs from NCBI were then used to search for missing regions of the previous sequences. This information was then used to design primers that would yield cDNA of the appropriate length. Cloning and sequencing with these primers yielded a cDNA product of 1548 base pairs with a predicted protein size of 46kDa. A multiple sequence alignment is shown in appendix.

PLIN1 Gene Model. The results of the cDNA sequencing were run in BLAST to find possible locations of exons using the Aag2 cell line genome assembly created by the Adino Lab at University of California, San Franscico and Pacific Biosciences (https://www.vectorbase.org/organisms/aedes-aegypti/aag2/aag2). *Aedes albopictus* is closely related to *A. aegypti* and shows a single transcript for Lsd1 AALF025165 of 1305 bp encoding a protein with a weight of 49 kDa. To compare the transcripts, the similarity of the proteins was ascertained using the Expasy local alignment tool. In order to further determine similarity, the cDNA sequence was queried against the *A. aegypti* genome which was followed by a query of the *A. albopictus* PLIN1 transcript against the *A. albopictus* genome. A possible genetic map was then made using the hits from the cDNA query in which the approximate size of the exons was postulated using the size of the hits and the size of the introns were calculated as the difference in one hit and the next.

Preparation of the RNA Probe. To determine regions of overlap between the 1548bp cDNA, AET-5405, Lsd1 A, and Lsd1 B, a multiple alignment was constructed (Fig 1) using the clustal omega tool found on the European Bioinformatics institute (http://www.ebi.ac.uk/). This information was then used to design the primers needed to make the antisense northern blot RNA-probe of approximately 700bp. The following primers were designed: forward primer AELsd1C-F (5'-TGCCAGC-AGCAATCGGATGACTCTCATAAAACG-CAC-3') and the reverse primer AELsd1C-R (5' TAATACGACTCACTATAGGGTCCATTGGCGCTCATCG). The reverse primer contains a T7 promoter (5'-GTCCATTGGCGCTCATCG) to be used in the

generation of the antisense RNA probe. The location of the primers in the cDNA sequence is shown in Figure 1.

Results

Gene Structure. To determine the location of introns and exons in the PLIN1 gene, the 1548bp cDNA cloned in the lab was used to run a BLAST search against the genomic sequence of *A. aegypti* reported in vector base. The search revealed 7 exons (Table 1). The first exon contains only 4 coding bases and the 5'-UTR. Comparison of the *A. aegypti* and *A. albopictus* predicted proteins yielded 363 overlapping amino acids with 97.5% similarity (Fig 1). Since the sequence of PLIN1 from *Aedes albopictus* is almost identical to that from *A. aegypti*, we also compared the predicted gene structure of *A. aegypti* with that reported for *A. albopictus* (Table 1). The comparison shows that both species of Aedes have similar gene structures (seven exons of similar size) and, as expected, some differences in the length of introns and untranslated regions (not shown).

A sketch of the generated gene model of A. aegypti PLIN1 is shown in figure 2 with a model of *A. albopictus* in figure 3.

Table 1. Summary of cDNA search against *A. aegypti* genome and *A. albopictus* transcript against genome.

cDNA against A. aegypti genome	Aligment Length	% Identity	Location (Aag2 cell line genome sequence)	A. albopictus PLIN1 transcript against A. albopictus genome	Alignment Length	% Identity	Location (GCA_001876365.2 assembly)
Hit 1 3937249 5' ATG	138bp	97.8% Identity	3937115-3937252	**Hit 1 5' ATG**	54bp	100%	1008612-1008665
Hit 2	127bp	100%	3937448-3937574	**Hit 2**	127bp	99.2%	1008898-1009024
Hit 3	196bp	95.96%	3937646-3937843	**Hit 3**	191bp	95.3%	1009098-1009288
Hit 4	215bp	97.67%	3939175-3939389	**Hit 4**	209bp	97.1%	1019564-1019722
Hit 5	439bp	97.9%	3952495-3952935	**Hit 5**	447bp	97.5%	1020266-1020712
Hit 6	114bp	98.2%	3953034-3953147	**Hit 6**	116bp	97.4%	1034832-1034947
Hit 7 3953426 TAG 3'	332bp	97.3%	3953293-3953628	**Hit 7 TGA 3'**	229bp	99.1%	1035083-1035311

Figure 1. Local alignment of protein sequence of *A. aegypti* and *A. albopictus*

```
 >>Albopictus 434 bp                                              (434 aa)

 Waterman-Eggert score: 2205;  590.7 bits; E(1) <  2.7e-173
94.5% identity (97.5% similar) in 363 aa overlap (1-363:1-361)

               10        20        30        40        50        60
Aegypt MVHQKLKRQNSGLPRMESISRVGSIPVVETGLKTANTVYQKIKESNGLFNWGLETAEAIT
       :::::::::::::::::::: :::::::::::::::.:.:::::::::::::::.::
Albopi MVHQKLKRQNSGLPRMESISRFGSIPVVETGLKTANTVYHKVKESNGLFNWGLETAETIT
               10        20        30        40        50        60

               70        80        90       100       110       120
Aegypt YAFVDSLRPAAKLIEGPLHRLDNFMCKSLDFVEQKVPSMYLPPEMMYWNTKEYMSDHLVK
       :::::::::::::::::::::::::::::::::::::::::::::::::::::::::.:::
Albopi YAFVDSLRPAAKLIEGPLHRLDNFMCKSLDFVEQKVPSMYLPPEMMYWNTKEYMSDRLVK
               70        80        90       100       110       120

              130       140       150       160       170       180
Aegypt PVLSRANSMKNLGHVVLESRVSNYAADRLDGALNVCDKYVDRYLPAEPEPAEDQTDSPNP
       ::::::::::::::::::::::::::::::::::::::::::::::  :: : ::: .:
Albopi PVLSRANSMKNLGHVVLESRVSNYAADRLDGALNVCDKYVDRYLPAE--PAGDCTDSLHP
              130       140       150       160         170

              190       200       210       220       230       240
Aegypt CQQQSDDSHKTHVIQTIHRGQLISRKLTRRLTFRTRQELTALKKQSTEAVHVVFYAAELI
        :::::::::::::::::::: ::::::::::::::::::::::::::::::::::::::
Albopi TQQQSDDSHKTHVIQTIHRGQQISRKLTRRLTFRTRQELTALKKQSTEAVHVVFYAAELI
      180       190       200       210       220       230

              250       260       270       280       290       300
Aegypt ATNPRLAMQKGVELWQYLSADEPENQARPQTLEQLVVLLTRESVRKVVHLINFTAGTVTK
       :::::::::::::::::::::::::::::::::::::::::::.:::.:::::.:::
Albopi ATNPRLAMQKGVELWQYLSADEPENQARPQTLEQLVVLLTRESVRKMVHLVNFTAGAVTK
      240       250       260       270       280       290

              310       320       330       340       350       360
Aegypt VPKTIRSQTRELLHHMMFATDRLIKAAHLENAKKATITEATGLMHRIQHTYEELQNQTNL
       ::.::: :::::::::::::::::::::::.:::::::::::::::::::::::::::::
Albopi VPRTIRVQTRELLHHMMFATDRLIKAAHLENAKNATITEATGLMHRIQHTYEELQNQTNL
      300       310       320       330       340       350


Aegypt ALE
       ::.
Albopi ALQ
      360
```

Figure 2. Gene model of PLIN1 in *Aedes aegypti*

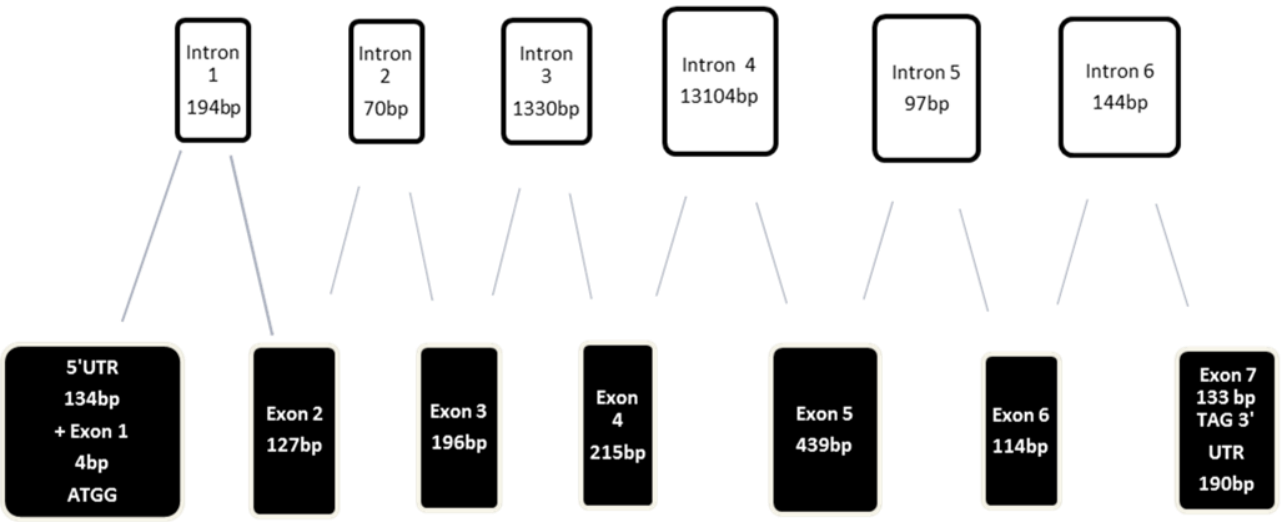
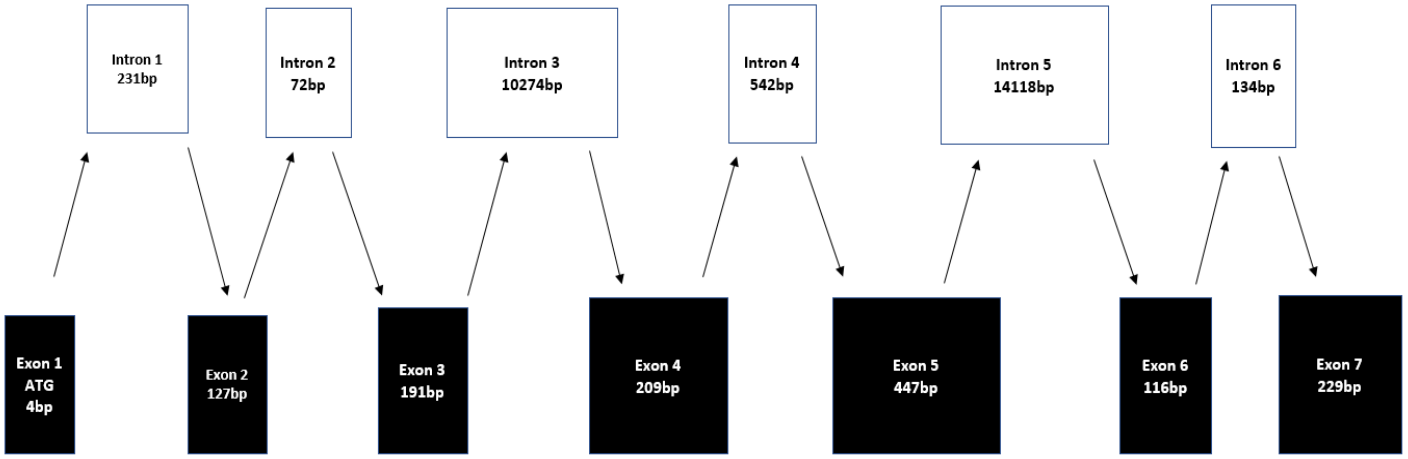


Figure 3. Gene model of PLIN1 in *Aedes albopictus*



Design and production of the anti-sense RNA probe. The sequence of the 1548 bp cDNA clone was used to obtain the RNA probe. The locations of the primers used to obtain the probe are highlighted in the figure 4. In order to generate sufficient quantities of cDNA (>20µg) to make the RNA probe, a total of 8 PCR reactions were conducted using the previously sequenced 1548bp cDNA as a template (Table 2). The PCR products were purified using an affinity column (Qiagen PCR Cleanup Kit) and the amount of purified PCR product obtained was estimated from the UV-absorption spectra. A picture of a representative agarose gel, used to confirm the presence and size of the PCR product is shown in figure 5. As expected we obtained a band of approximately 700 base pairs (Figure 5).

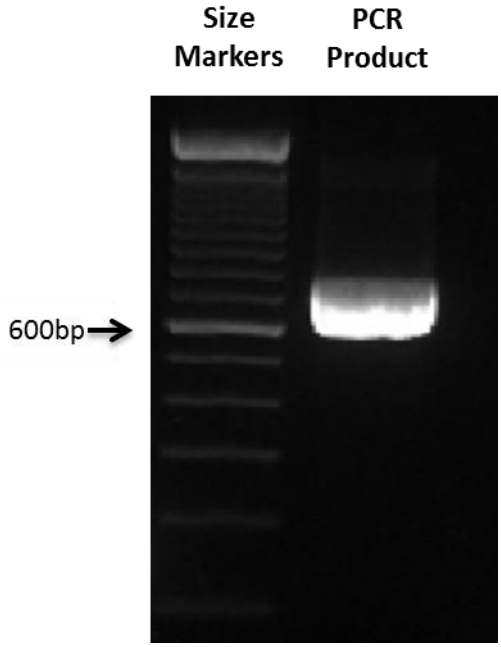
Figure 4. Aedes aegypti PLIN1 cDNA sequence. The locations of the forward and reverse primers used to obtain the RNA probe are highlighted in green and yellow.

GATCGTTTTCTTTTTTCCCGTGCGTTTGAGCACTCAAAGTTTCCAATCAAGAGTTGTGTTTTCTCGTGCGTGTTGTC
CCAAACCAGCTAAGGGTCATCGTCATTGTTCAAACAGCCAATACCTACGCTGCAACAATGGTGCATCAGAAGCTG
AAGCGCCAAAATTCGGGACTCCCGCGGATGGAATCCATATCGCGGGTCGGTAGCATTCCGGTCGTGGAAACCGG
ACTGAAAACCGCCAACACGGTTTATCAGAAGATTAAGGAAAGTAATGGATTGTTTAACTGGGGACTCGAGACGG
CTGAAGCAATCACATACGCCTTCGTGGATTCACTGCGTCCGGCTGCTAAGCTGATCGAAGGACCGCTGCATCGTTT
GGATAATTTTATGTGTAAAAGCTTGGACTTTGTCGAACAGAAGGTCCCTTCAATGTATCTTCCTCCGGAGATGATGT
ACTGGAACACCAAGGAATACATGTCGGATCACCTGGTCAAGCCGGTGCTGAGCCGTGCCAACTCGATGAAGAAC
CTCGGTCATGTAGTGCTAGAGTCCCGAGTGTCGAACTACGCAGCGGATCGTCTGGATGGGGCTTTAAATGTGTGC
GATAAATATGTGGATCGATACCTCCCTGCGGAGCCAGAACCCGCCGAAGATCAAACCGATTCACCTAATCCTTGCC
AGCAGCAATCGGATGACTCTCATAAAACGCACGTTATCCAAACCATTCACCGGGGACAACTAATATCTCGCAAACT
GACACGTCGATTAACGTTTCGCACCCGTCAGGAATTGACCGCCCTCAAGAAGCAGAGTACCGAAGCCGTACACGT
GGTGTTCTACGCAGCCGAACTGATTGCCACCAACCCACGGCTGGCCATGCAGAAGGGCGTCGAACTGTGGCAAT
ATCTTAGCGCTGACGAACCGGAAAACCAAGCCCGCCCCCAAACGCTAGAACAACTGGTGGTCCTACTGACGAGA
GAATCCGTCCGTAAGGTTGTACATCTGATCAACTTCACCGCTGGAACAGTAACGAAGGTCCCCAAGACGATTCGTT
CGCAAACGCGCGAATTGCTGCACCACATGATGTTTGCCACCGATCGATTGATAAAGGCTGCCCACTTGGAAAACG
CTAAGAAAGCAACGATTACCGAAGCGACTGGACTGATGCATCGAATTCAACACACCTATGAAGAACTGCAAAATC
AAACCAATTTGGCTTTGGAACGTTTAGCAGTGTTTCTTTCCGGCCGCCTGGAAGCGGAGAAGATCACCACCACCG
ACAATCCCCGGCGACGGATCCAAAACCGAGCGCACCATAATCCAATGCATACTAGCATCAATGGCGTTTATTAGCG
TACGAACGGCCGAGTTGCTGAACAGACGGATCCGCGAGGGCACCCCGCTCGAACCCGTACTCAACGTATTACGAT
GAGCGCCAATGGACCAATGGACGTGAATGAGATGTTTTTTTAAAATAAGTTCTGAAGCATAAAACATTTGTGGTGT
AATAATGTTACTACTTGTTGATAAATAATAAATTTCATATGATGTAC

Table 2. Estimates of PCR product yields

	ng/µl	µl	ng
PCR #1	62.6		
PCR #2	41.5	49	2035
PCR #3	25.6	49	1253
PCR #4	16.6	49	813
PCR #5	37.2	49	1821
PCR #6	37.3	49	1830
PCR #7	40.1	49	1963
PCR #8A	36.4	49	1786
PCR #8B	55.3	49	2709

Figure 5. Purity and size of PCR product. The PCR was analyzed by electrophoresis in an agarose gel, labelled with ethidium bromide and photographed under UV light. The left lane shows the used to make cDNA for the antisense RNA probe. 100 base pair ladder with dark band at 600 base pair



Discussion

The preliminary studies conducted involving immunoblotting indicate that the predicted transcripts from vectorbase and NCBI are incomplete. Comparison between the PLIN1 protein sequence predicted by the 1548 bp cDNA and that predicted by the *A. albopictus* genome were found to be highly similar. Further comparison between the predicted exons of the cDNA the *A. albopictus* exons supports the proposed genetic model. These findings also indicate that the genomic sequence for both *A. aegypti* and *A. albopictus* is correct.

The purified PCR product should be sufficient to produce enough anti-sense RNA probe for northern blot assays. However, this will remain uncertain until the northern blot has been conducted. These assays will be done in ovaries and fat body at various times of the *Aedes aegypti* life cycle to investigate the size and number of transcripts of PLIN1.

References

[1] Engdahl C, Knutsson S, Fredriksson S-Å, Linusson A, Bucht G, Ekström F. Acetylcholinesterases from the Disease Vectors Aedes aegypti and Anopheles gambiae: Functional Characterization and Comparisons with Vertebrate Orthologues. *Plos One*. 2015;10(10). doi:10.1371/journal.pone.0138598.

[2] Lin P, Chen X, Moktan H, et al. Membrane attachment and structure models of lipid storage droplet protein 1. *Biochimica et Biophysica Acta (BBA) - Biomembranes*. 2014;1838(3):874-881. doi:10.1016/j.bbamem.2013.12.003.

[3] Arrese EL, Rivera L, Hamada M, et al. Function and structure of lipid storage droplet protein 1 studied in lipoprotein complexes. *Archives of Biochemistry and Biophysics*. 2008;473(1):42-47. doi:10.1016/j.abb.2008.02.036.

[4] Arrese E, Mirza S, Rivera L, Howard A, Chetty P, Soulages J. Expression of lipid storage droplet protein-1 may define the role of AKH as a lipid mobilizing hormone in Manduca sexta. *Insect Biochemistry and Molecular Biology*. 2008;38(11):993-1000. doi:10.1016/j.ibmb.2008.08.004.

[5] Patel RT, Soulages JL, Hariharasundaram B, Arrese EL. Activation of the Lipid Droplet Controls the Rate of Lipolysis of Triglycerides in the Insect Fat Body. *Journal of Biological Chemistry*. 2005;280(24):22624-22631. doi:10.1074/jbc.m413128200.

[6] Men T, Binh T, Yamaguchi M, Huy N, Kamei K. Function of Lipid Storage Droplet 1 (Lsd1) in Wing Development of Drosophila melanogaster. *International Journal of Molecular Sciences*. 2016;17(5):648. doi:10.3390/ijms17050648.

Appendix

Multiple alignment of cDNA sequenced in the lab and the partial sequences, AET-5405, Lsd1A and Lsd1B. The primers for the cDNA are highlighted in green and yellow with the region covered by the probe highlighted in turquoise.

```
cDNA              GATCGTTTTCTTTTTTCCCGTGCGTTTGAGCACTCAAAGTTTCCAATCAAGAGTTGTGTT      60
DQ440406.1        ------------------------------------------------------------      0
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              TTCTCGTGCGTGTTGTCCCAAACCAGCTAAGGGTCATCGTCATTGTTCAAACAGCCAATA      120
DQ440406.1        ------------------------------------------------------------      0
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              CCTACGCTGCAACAATGGTGCATCAGAAGCTGAAGCGCCAAAATTCGGGACTCCCGCGGA      180
DQ440406.1        ------------------------------------------------------------      0
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              TGGAATCCATATCGCGGGTCGGTAGCATTCCGGTCGTGGAAACCGGACTGAAAACCGCCA      240
DQ440406.1        ------------------------------------------------------------      0
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              ACACGGTTTATCAGAAGATTAAGGAAAGTAATGGATTGTTTAACTGGGGACTCGAGACGG      300
DQ440406.1        ------------------------------------------------------------      0
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              CTGAAGCAATCACATACGCCTTCGTGGATTCACTGCGTCCGGCTGCTAAGCTGATCGAAG      360
DQ440406.1        ------------------------------------------------------------      0
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              GACCGCTGCATCGTTTGGATAATTTTATGTGTAAAAGCTTGGACTTTGTCGAACAGAAGG      420
DQ440406.1        ------------------------------------------------------------      0
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              TCCCTTCAATGTATCTTCCTCCGGAGATGATGTACTGGAACACCAAGGAATACATGTCGG      480
DQ440406.1        -------------------------ATGATGTGCTGGAACACCAAGGAATACATGTCGG      34
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              ATCACCTGGTCAAGCCGGTGCTGAGCCGTGCCAACTCGATGAAGAACCTCGGTCATGTAG      540
DQ440406.1        ATCACCTGGTCAAGCCGGTGCTGAGCCGTGCCAACTCGATGAAGAACCTCGGTCATGTAG      94
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0

cDNA              TGCTAGAGTCCCGAGTGTCGAACTACGCAGCGGATCGTCTGGATGGGGCTTTAAATGTGT      600
DQ440406.1        TGCTAGAGTCCCGAGTGTCGAACTACGCCGCGGATCGTCTGGATGGGGCTTTAAATGTGT      154
AAEL005951-RA     ------------------------------------------------------------      0
AAEL005951-RB     ------------------------------------------------------------      0
```

```
cDNA              GCGATAAATATGTGGATCGATACCTCCCTGCGGAGCCAGAACCCGCCGAAGATCAAACCG       660
DQ440406.1        GCGATAAATATGTGGATCGATACCTCCCTGCGGAGCCAGAACCCGCCGAAGATCAAACCG       214
AAEL005951-RA     --------------ATGGCACACGGATCAAAAGCCAAACAGCAGCAACTTCACAAAGGCT       46
AAEL005951-RB     --------------ATGGCACACGGATCAAAAGCCAAACAGCAGCAACTTCACAAAGGCT       46
                                * **    *    *    * * *         *  **  *

cDNA              ATTCACCTAATCCTTGCCAGCAGCAATCGGATGACTCTCATAAAACGCACGTTATCCAAA       720
DQ440406.1        ATTCACCTAATCCTTGCCAGCAGCAATCGGATGACTCTCATAAAACGCACGTTATCCAAA       274
AAEL005951-RA     CGGCACCTAATCCTTGCCAGCAGCAATCGGATGACTCTCATAAAACGCACGTTATCCAAA       106
AAEL005951-RB     CGGCACCTAATCCTTGCCAGCAGCAATCGGATGACTCTCATAAAACGCACGTTATCCAAA       106
                     ************************************************************

cDNA              CCATTCACCGGGGACAACTAATATCTCGCAAACTGACACGTCGATTAACGTTTCGCACCC       780
DQ440406.1        CCATTCACCGGGGACAACTAATATCTCGCAAACTGACACGTCGATTAACGTTTCGCACCC       334
AAEL005951-RA     CCATTCACCGGGGACAACTAATATCTCGCAAACTGACACGTCGATTAACGTTTCGCACCC       166
AAEL005951-RB     CCATTCACCGGGGACAACTAATATCTCGCAAACTGACACGTCGATTAACGTTTCGCACCC       166
                  ************************************************************

cDNA              GTCAGGAATTGACCGCCCTCAAGAAGCAGAGTACCGAAGCCGTACACGTGGTGTTCTACG       840
DQ440406.1        GTCAGGAATTGACCGCCCTCAAGAAGCAGAGTACCGAAGCCGTACACGTGGTGTTCTACG       394
AAEL005951-RA     GTCAGGAATTGACCGCCCTCAAGAAGCAGAGTACCGAAGCCGTACACGTGGTGTTCTACG       226
AAEL005951-RB     GTCAGGAATTGACCGCCCTCAAGAAGCAGAGTACCGAAGCCGTACACGTGGTGTTCTACG       226
                  ************************************************************

cDNA              CAGCCGAACTGATTGCCACCAACCCACGGCTGGCCATGCAGAAGGGCGTCGAACTGTGGC       900
DQ440406.1        CAGCCGAACTGATTGCCACCAACCCACGGCTGGCCATGCAGAAGGGCGTCGAACTGTGGC       454
AAEL005951-RA     CAGCCGAACTGATTGCCACCAACCCACGGCTGGCCATGCAGAAGGGCGTCGAACTGTGGC       286
AAEL005951-RB     CAGCCGAACTGATTGCCACCAACCCACGGCTGGCCATGCAGAAGGGCGTCGAACTGTGGC       286
                  ************************************************************

cDNA              AATATCTTAGCGCTGACGAACCGGAAAACCAAGCCCGCCCCCAAACGCTAGAACAACTGG       960
DQ440406.1        AATATCTTAGCGCTGACGAACCGGAAAACCAAGCCCGCCCCCAAACGCTAGAACAACTGG       514
AAEL005951-RA     AATATCTTAGCGCTGACGAACCGGAAAACCAAGCCCGCCCCCAAACGCTAGAACAACTGG       346
AAEL005951-RB     AATATCTTAGCGCTGACGAACCGGAAAACCAAGCCCGCCCCCAAACGCTAGAACAACTGG       346
                  ************************************************************

cDNA              TGGTCCTACTGACGAGAGAATCCGTCCGTAAGGTTGTACATCTGATCAACTTCACCGCTG      1020
DQ440406.1        TGGTCCTACTGACGAGAGAATCCGTCCGTAAGGTTGTACATCTGATCAACTTCACCGCTG       574
AAEL005951-RA     TGGTCCTACTGACGAGAGAATCCGTCCGTAAGGTTGTACATCTGATCAACTTCACCGCTG       406
AAEL005951-RB     TGGTCCTACTGACGAGAGAATCCGTCCGTAAGGTTGTACATCTGATCAACTTCACCGCTG       406
                  ************************************************************

cDNA              GAACAGTAACGAAGGTCCCCAAGACGATTCGTTCGCAAACGCGCGAATTGCTGCACCACA      1080
DQ440406.1        GAACAGTAACGAAGGTCCCCAAGACGATTCGTTCGCAAACGCGCGAATTGCTGCACCACA       634
AAEL005951-RA     GAACAGTAACGAAGGTCCCCAAGACGATTCGTTCGCAAACGCGCGAATTGCTGCACCACA       466
AAEL005951-RB     GAACAGTAACGAAGGTCCCCAAGACGATTCGTTCGCAAACGCGCGAATTGCTGCACCACA       466
                  ************************************************************

cDNA              TGATGTTTGCCACCGATCGATTGATAAAGGCTGCCCACTTGGAAAACGCTAAGAAAGCAA      1140
DQ440406.1        TGATGTTTGCCACCGATCGATTGATAAAGGCTGCCCACTTGGAAAACGCTAAGAAAGCAA       694
AAEL005951-RA     TGATGTTTGCCACCGATCGATTGATAAAGGCTGCCCACTTGGAAAACGCTAAGAAAGCAA       526
AAEL005951-RB     TGATGTTTGCCACCGATCGATTGATAAAGGCTGCCCACTTGGAAAACGCTAAGAAAGCAA       526
                  ************************************************************

cDNA                CGATTACCGAAGCGACTGGACTGATGCATCGAATTCAACACACCTATGAAGAACTGCAAA      1200
DQ440406.1        CGATTACCGAAGCGACTGGACTGATGCATCGAATTCAACACACCTATGAAGAACTGCAAA       754
AAEL005951-RA     CGATTACCGAAGCGACTGGACTGATGCATCGAATTCAACACACCTATGAAGAACTGCAAA       586
AAEL005951-RB     CGATTACCGAAGCGACTGGACTGATGCATCGAATTCAACACACCTATGAAGAACTGCAAA       586
                  ************************************************************

cDNA              ATCAAACCAATTTGGCTTTGGAACGTTTAGCAGTGTTTCTTTCCGGCCGCCTGGAAGCGG      1260
DQ440406.1        ATCAAACCAATTTGGCTTTGGAACGTTTAGCAGTGTTTCTTTCCGGCCGCCTGGAAGCGG       814
AAEL005951-RA     ATCAAACCAATTTGGCTTTGGTAAGGTTGCATAGCATCTTAGTTGATTGATGAAACTGTG       646
AAEL005951-RB     ATCAAACCAATTTGGCTTTGGAACGTTTAGCAGTGTTTCTTTCCGGCCGCCTGGAAGCGG       646
                  ********************* * * **        *  *    *   *     *    *

cDNA              AGAAGATCACCACCACCGACAATCC-CCGGCGACGGATCCAAAACCGAGCGCACCATAAT      1319
DQ440406.1        AGAAGATCACCACCACCGACAATCC-CCGGCGACGGATCCAAAACCGAGCGCACCATAAT       873
AAEL005951-RA     CAT--TAAACCATCATAAAAAATATAATTGCCAAAAGTGTCAATGCTAATGCATGTTAA-       703
```

```
AAEL005951-RB      AGAAGATCACCACCACCGACAATCC-CCGGCGACGGATCCAAAACCGAGCGCACCATAAT       705
                          **** **   * ***      ** *    *   **  * *  ***   ***

cDNA               CCAATGCATACTAGCATCAATGGCGTTTATTAGCGTACGAACGGCCGAGTTGCTGAACAG       1379
DQ440406.1         CCAATGCATACTAGCATCAATGGCGTTTATTAG----------------------------       906
AAEL005951-RA      ---------ATTAAAATTTA-----------------------------------------       714
AAEL005951-RB      CCAATGCATACTAGCATCAATGGCGTTTATTAGCGTACGAACGGCCGAGTTGCTGAACAG       765
                            * **  **  *

cDNA               ACGGATCCGCGAGGGCACCCCGCTCGAACCCGTACTCAACGTCGATGAGCGCCAATGGAC       1439
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      ACGGATCCGCGAGGGCACCCCGCTCGAACCCGTACTCAACGTCGATGAGCGCCAATGGAC       825

cDNA               CAATGGACGTGAATGAGAT----GTTTTTTTAAAATAAGTTCTGAAGCATAAAACATTTG       1495
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      CAATGGACGTGAATGAGATATATGTTTTTTTATAATAAGTTCTGAAGCATAAAACATTTG       885

cDNA               TGGTGTAATAATGTTACTACTTGTTGATAAATAATAAATTTCATATGATGTAC-------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      TGGTGTAATAATGTTACTACTTGTTGATAAATAATAAATTTCATATGATGTACAAAACTA       945

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      GTCTTAACTAATCTCAATTGATTGAGAAACTCTATCTCTTCTTCTTAATCATCGGTGTCC       1005

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      AAATGGGTGGTCCCCCCGAATCCGACGGTCGAGTAAGAAACACATTCGAAGCAAACTCCT       1065

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      CATCGATTTCGCTCGCATTTGCCAAATCGGAAATGTTCGTCACATCTTGGGATTCCGTTC       1125

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      CAGCAGAAGATTCCTCTGATTCCGCTTGCAAGGTCAGCTTCTCCAACGGAATTCCCACGG       1185

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      GTAGGTCTTCTTCGGCAATCTGGGGCATTTCTAAACCAAGATCGGCGTAAACCTCACTCA       1245

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      ACCGGGTGTATATTTTATCGGCAGTTTCATCGGCTTCCGCCAGCAGACGTTTCAGCTGTT       1305

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      CAGTGTCCTGGACGAACTCCGCGCTTTTTGCCTCGATGTCCTTCAAGGAGTTTTGGTTGT       1365

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
```

```
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      CGGCGATGAAGCGTCCGGCTTGATCAAGTTTGTTGGGGTTTTGCTGCACGAAGCCACACG       1425

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      CGAATTGGGCAGTTTTTTTGCGTAGTAAGCCCAGCTTGGTGTCTACTTCGTAGGCATGTT       1485

cDNA               ------------------------------------------------------------       1548
DQ440406.1         ------------------------------------------------------------       906
AAEL005951-RA      ------------------------------------------------------------       714
AAEL005951-RB      CCTTAGCGGTACGGAACATTTTCAATCCCAGGAACAGAACTTAAAACGTGCGATCTGGAT       1545

cDNA               ----------------------------------------------------------    1548
DQ440406.1         ----------------------------------------------------------    906
AAEL005951-RA      ----------------------------------------------------------    714
AAEL005951-RB      TTCACCGATTTCAACGTAAATGTTTTACTGGAACGATCAAATCCAAACTGCGC    1598
```