GEP Annotation Report

Student name: ___Jill Ackerman___
Student email: __jillackerman@unm.edu__
Faculty Advisor: ___Dr. Paul Szauter
College/University: __University of New Mexico_

Project Details

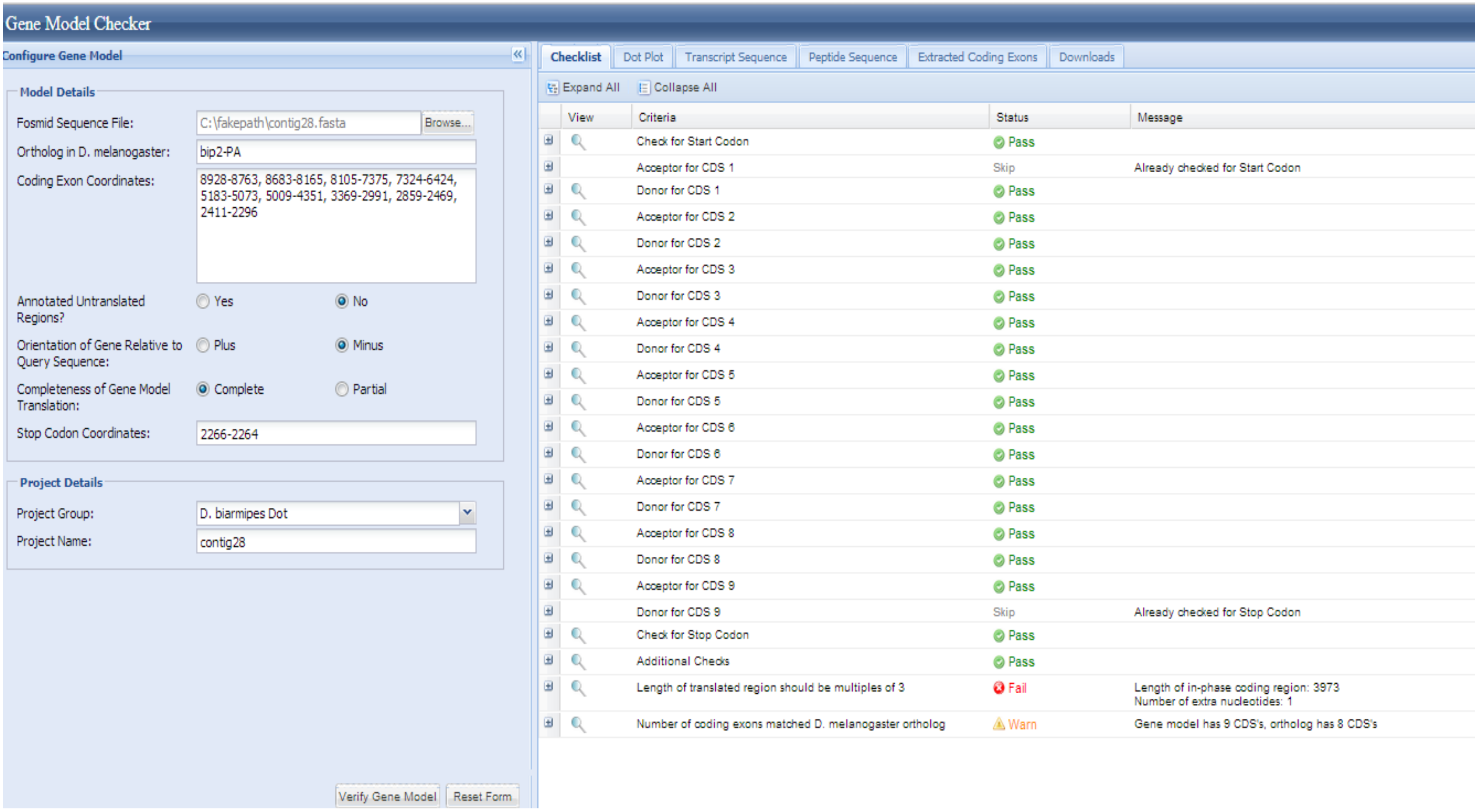
Project name: __*D. biarmipes* Contig 28__
Project species: _*Drosophila biarmipes*_
Date of submission: _________7/25/2013___________
Size of project in base pairs: 50,000
Number of genes in project: _2_

Does this report cover all genes and all isoforms or is it a partial report? _All genes/isoforms_

Gene Report Form

Gene name: *bip2*
Gene Symbol: *bip2*
Approximate location in project: 2296-8928
Number of isoforms in D. Melongaster: 1
Number of isoforms In this project: 1

Gene Model Checker Checklist
Regarding the warning/fail, see the "gain of intron" note at the end of the *bip2* report



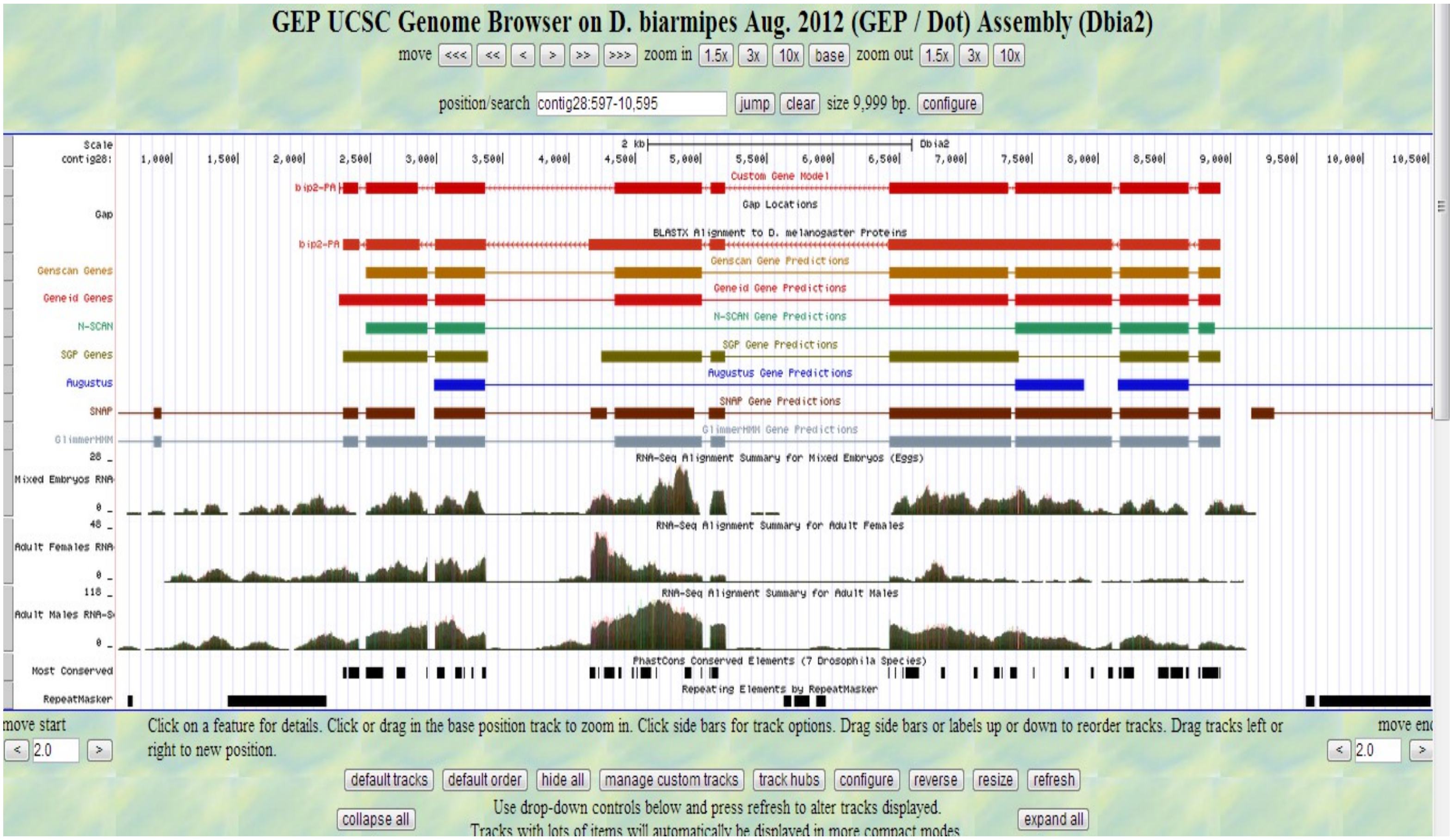
Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta
Ortholog in D. melanogaster: bip2-PA
Coding Exon Coordinates: 8928-8763, 8683-8165, 8105-7375, 7324-6424, 5183-5073, 5009-4351, 3369-2991, 2859-2469, 2411-2296
Annotated Untranslated Regions? ○ Yes ● No
Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus
Completeness of Gene Model Translation: ● Complete ○ Partial
Stop Codon Coordinates: 2266-2264

Project Details

Project Group: D. biarmipes Dot
Project Name: contig28

View	Criteria	Status	Message
🔍	Check for Start Codon	✓ Pass	
	Acceptor for CDS 1	Skip	Already checked for Start Codon
🔍	Donor for CDS 1	✓ Pass	
🔍	Acceptor for CDS 2	✓ Pass	
🔍	Donor for CDS 2	✓ Pass	
🔍	Acceptor for CDS 3	✓ Pass	
🔍	Donor for CDS 3	✓ Pass	
🔍	Acceptor for CDS 4	✓ Pass	
🔍	Donor for CDS 4	✓ Pass	
🔍	Acceptor for CDS 5	✓ Pass	
🔍	Donor for CDS 5	✓ Pass	
🔍	Acceptor for CDS 6	✓ Pass	
🔍	Donor for CDS 6	✓ Pass	
🔍	Acceptor for CDS 7	✓ Pass	
🔍	Donor for CDS 7	✓ Pass	
🔍	Acceptor for CDS 8	✓ Pass	
🔍	Donor for CDS 8	✓ Pass	
🔍	Acceptor for CDS 9	✓ Pass	
	Donor for CDS 9	Skip	Already checked for Stop Codon
🔍	Check for Stop Codon	✓ Pass	
🔍	Additional Checks	✓ Pass	
🔍	Length of translated region should be multiples of 3	✗ Fail	Length of in-phase coding region: 3973 Number of extra nucleotides: 1
🔍	Number of coding exons matched D. melanogaster ortholog	⚠ Warn	Gene model has 9 CDS's, ortholog has 8 CDS's

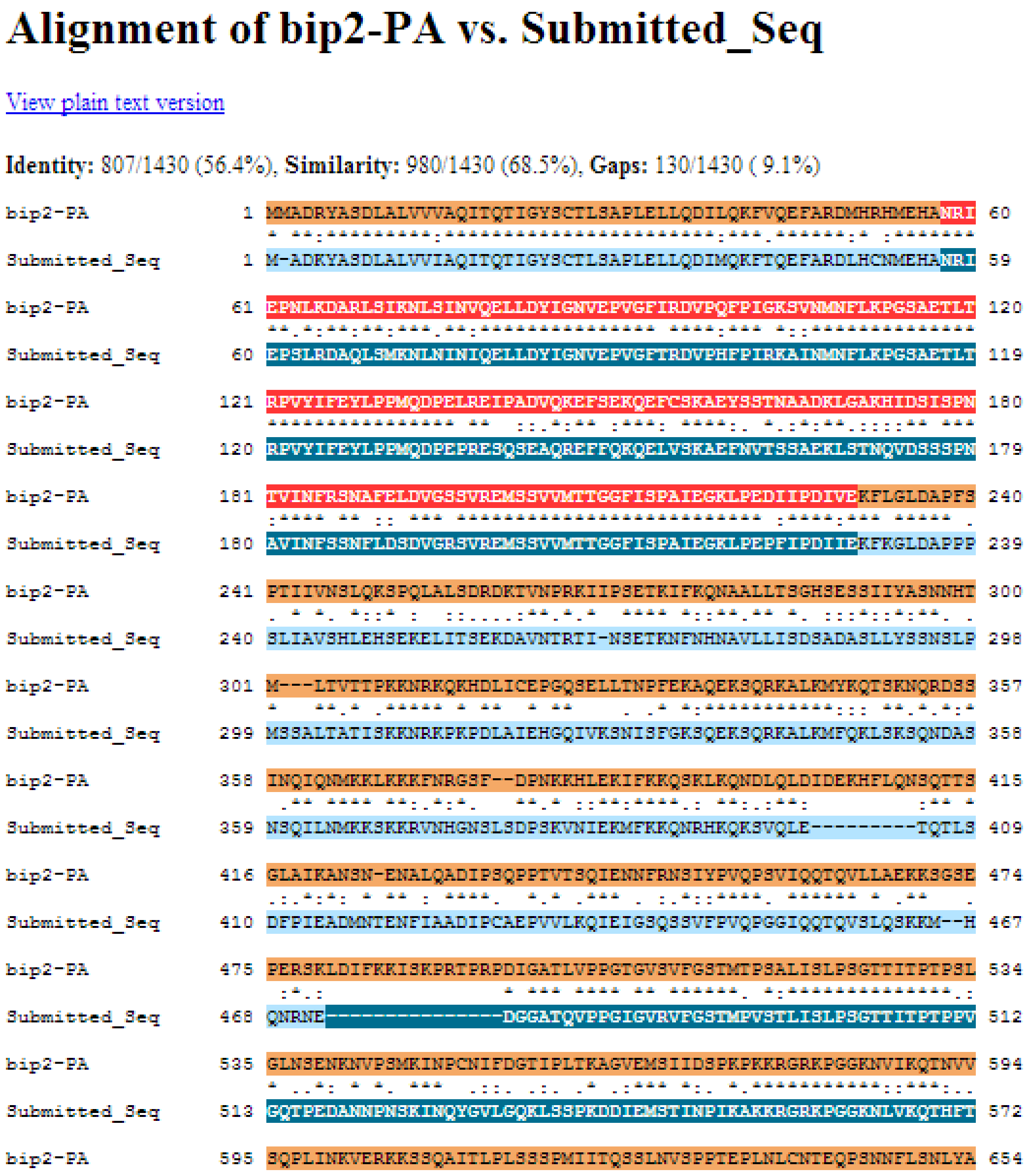


GEP UCSC Genome Browser on D. biarmipes Aug. 2012 (GEP / Dot) Assembly (Dbia2)

Click on a feature for details. Click or drag in the base position track to zoom in. Click side bars for track options. Drag side bars or labels up or down to reorder tracks. Drag tracks left or right to new position.

Use drop-down controls below and press refresh to alter tracks displayed.
Tracks with lots of items will automatically be displayed in more compact modes

Alignment between model and *D. melanogaster*

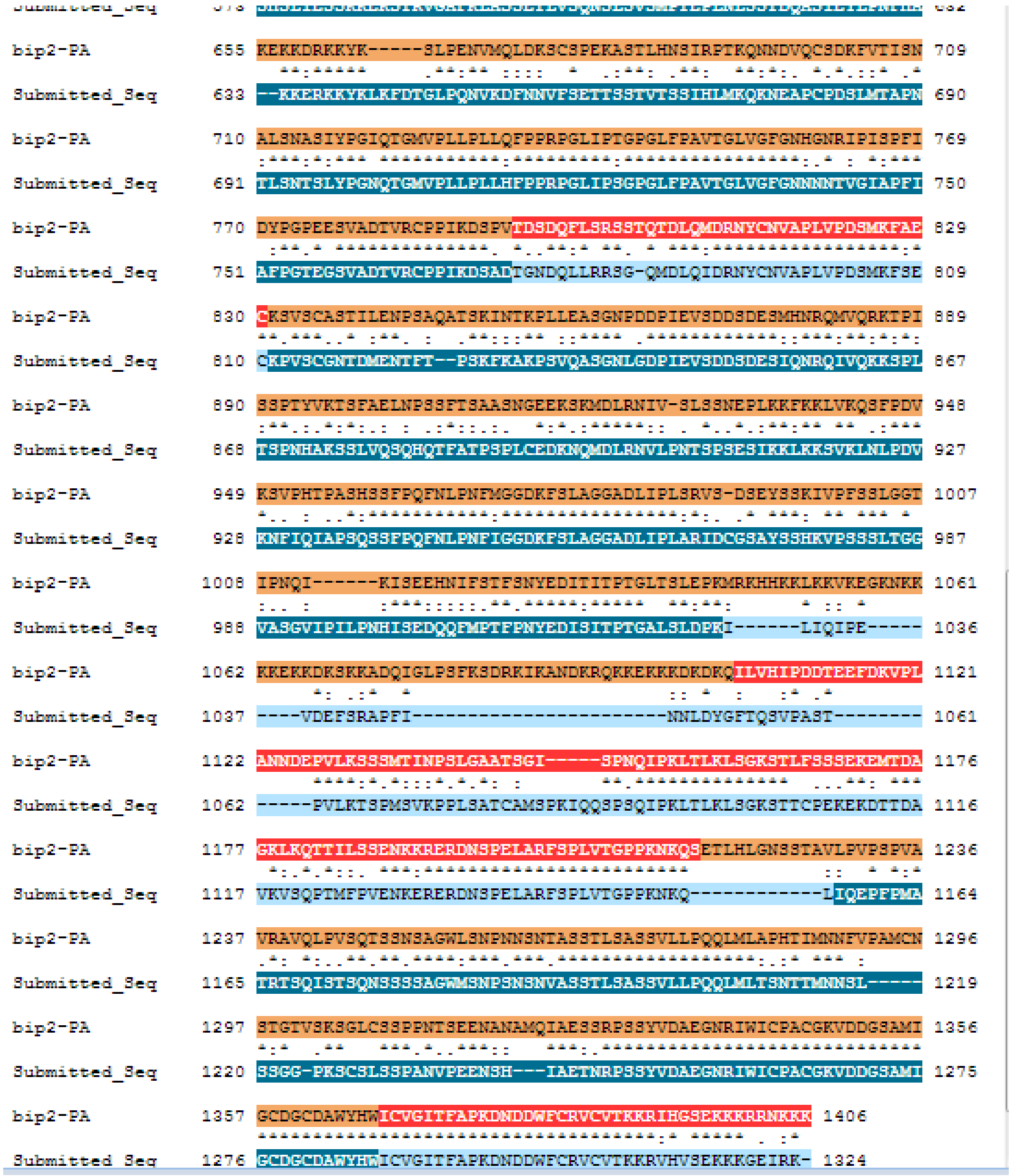


Alignment of bip2-PA vs. Submitted_Seq

View plain text version

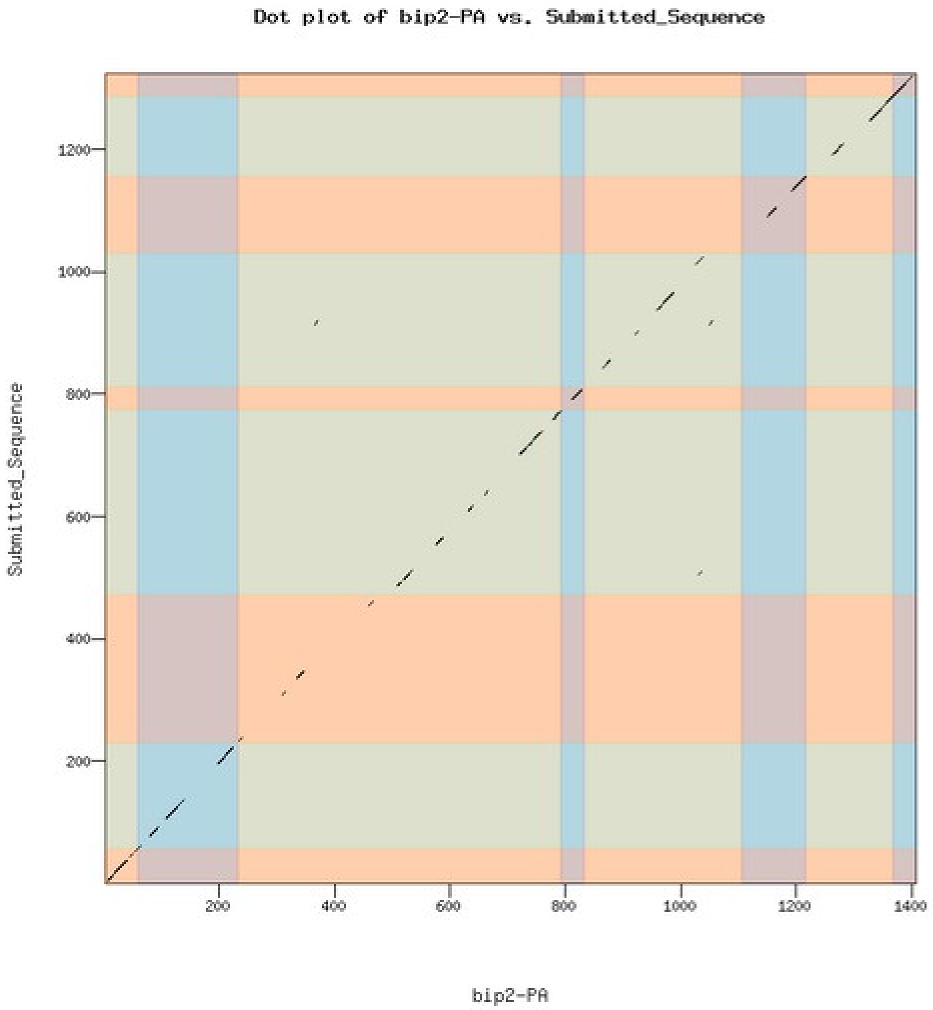
Identity: 807/1430 (56.4%), **Similarity:** 980/1430 (68.5%), **Gaps:** 130/1430 (9.1%)

```
bip2-PA            1 MMADRYASDLALVVVAQITQTIGYSCTLSAPLELLQDILQKFVQEFARDMHRHMEHANRI  60
                     * **:*********:***********************:***.******:* :*******
Submitted_Seq      1 M-ADKYASDLALVVIAQITQTIGYSCTLSAPLELLQDIMQKFTQEFARDLHCNMEHANRI  59

bip2-PA           61 EPNLKDARLSIKNLSINVQELLDYIGNVEPVGFIRDVPQFPIGKSVNMNFLKPGSAETLT 120
                     **.*:**:**:***.**:*************** ****:*** *::**************
Submitted_Seq     60 EPSLRDAQLSMKNLNINIQELLDYIGNVEPVGFTRDVPHFPIRKAINMNFLKPGSAETLT 119

bip2-PA          121 RPVYIFEYLPPMQDPELREIPADVQKEFSEKQEFCSKAEYSSTNAADKLGAKHIDSISPN 180
                     **************** **  ::.*:** :***: ****:. *.:*:**.::::** ***
Submitted_Seq    120 RPVYIFEYLPPMQDPEPRESQSEAQREFFQKQELVSKAEFNVTSSAEKLSTNQVDSSSPN 179

bip2-PA          181 TVINFRSNAFELDVGSSVREMSSVVMTTGGFISPAIEGKLPEDIIPDIVEKFLGLDAPFS 240
                     :**** ** :: *** ************************** :****:*** ***** .
Submitted_Seq    180 AVINFSSNFLDSDVGRSVREMSSVVMTTGGFISPAIEGKLPEPFIPDIIEKFKGLDAPPP 239

bip2-PA          241 PTIIVNSLQKSPQLALSDRDKTVNPRKIIPSETKIFKQNAALLTSGHSESSIIYASNNHT 300
                     . * *. *::* :  ::....:**.*.*  **** *::**.** *. :::*::*:**. .
Submitted_Seq    240 SLIAVSHLEHSEKELITSEKDAVNTRTI-NSETKNFNHNAVLLISDSADASLLYSSNSLP 298

bip2-PA          301 M---LTVTTPKKNRKQKHDLICEPGQSELLTNPFEKAQEKSQRKALKMYKQTSKNQRDSS 357
                     *   **.* .***** * **  * **    . .* *:***********::: **.*.*:*
Submitted_Seq    299 MSSALTATISKKNRKPKPDLAIEHGQIVKSNISFGKSQEKSQRKALKMFQKLSKSQNDAS 358

bip2-PA          358 INQIQNMKKLKKKFNRGSF--DPNKKHLEKIFKKQSKLKQNDLQLDIDEKHFLQNSQTTS 415
                      .** **** **:.*:*.   **.* ::**:****.: **:.:**:         :** *
Submitted_Seq    359 NSQILNMKKSKKRVNHGNSLSDPSKVNIEKMFKKQNRHKQKSVQLE---------TQTLS 409

bip2-PA          416 GLAIKANSN-ENALQADIPSQPPTVTSQIENNFRNSIYPVQPSVIQQTQVLLAEKKSGSE 474
                     .:.*:*: * ** : ****.  *.* .*** . :.*::****. ****** * .**   .
Submitted_Seq    410 DFPIEADMNTENFIAADIPCAEPVVLKQIEIGSQSSVFPVQPGGIQQTQVSLQSKKM--H 467

bip2-PA          475 PERSKLDIFKKISKPRTPRPDIGATLVPPGTGVSVFGSTMTPSALISLPSGTTITPTPSL 534
                      :*.:               * *** **** ** ******. *:**************.:
Submitted_Seq    468 QNRNE---------------DGGATQVPPGIGVRVFGSTMPVSTLISLPSGTTITPTPPV 512

bip2-PA          535 GLNSENKNVPSMKINPCNIFDGTIPLTKAGVEMSIIDSPKPKKRGRKPGGKNVIKQTNVV 594
                     * ..*: * *. ***  .::. .:. .* .:*** *:. *.***********::***:..
Submitted_Seq    513 GQTPEDANNPNSKINQYGVLGQKLSSPKDDIEMSTINPIKAKKRGRKPGGKNLVKQTHFT 572

bip2-PA          595 SQPLINKVERKKSSQAITLPLSSSPMIITQSSLNVSPPTEPLNLCNTEQPSNNFLSNLYA 654
```



```
Submitted_Seq     573 [illegible] 632

bip2-PA           655 KEKKDRKKYK-----SLPENVMQLDKSCSPEKASTLHNSIRPTKQNNDVQCSDKFVTISN 709
                        **:*****     .**:** ::::  *  .:**: .**:  **:*:. *.*.::* .*
Submitted_Seq     633 --KKERKKYKLKFDTGLPQNVKDFNNVFSETTSSTVTSSIHLMKQKNEAPCPDSLMTAPN 690

bip2-PA           710 ALSNASIYPGIQTGMVPLLPLLQFPPRPGLIPTGPGLFPAVTGLVGFGNHGNRIPISPFI 769
                      :***:*:*** ***********:*********:****************:.* : *:***
Submitted_Seq     691 TLSNTSLYPGNQTGMVPLLPLLHFPPRPGLIPSGPGLFPAVTGLVGFGNNNNTVGIAPFI 750

bip2-PA           770 DYPGPEESVADTVRCPPIKDSPVTDSDQFLSRSSTQTDLQMDRNYCNVAPLVPDSMKFAE 829
                       :**.* **************. *..**:* **. * ***:*****************:*
Submitted_Seq     751 AFPGTEGSVADTVRCPPIKDSADTGNDQLLRRSG-QMDLQIDRNYCNVAPLVPDSMKFSE 809

bip2-PA           830 CKSVSCASTILENPSAQATSKINTKPLLEASGNPDDPIEVSDDSDESMHNRQMVQRKTPI 889
                      **.***..* :**. :  .**:::** ::**** .************::***:**:*:*:
Submitted_Seq     810 CKPVSCGNTDMENTFT--PSKFKAKPSVQASGNLGDPIEVSDDSDESIQNRQIVQKKSPL 867

bip2-PA           890 SSPTYVKTSFAELNPSSFTSAASNGEEKSKMDLRNIV-SLSSNEPLKKFKKLVKQSFPDV 948
                      :**.:.*:*:.: : .:*::.:.  *:*.:*****:: . *..*.:**:** ** .:***
Submitted_Seq     868 TSPNHAKSSLVQSQHQTFATPSPLCEDKNQMDLRNVLPNTSPSESIKKLKKSVKLNLPDV 927

bip2-PA           949 KSVPHTPASHSSFPQFNLPNFMGGDKFSLAGGADLIPLSRVS-DSEYSSKIVPFSSLGGT 1007
                      *.. : ..*:***********:****************.*:. .* ***: ** *** *
Submitted_Seq     928 KNFIQIAPSQSSFPQFNLPNFIGGDKFSLAGGADLIPLARIDCGSAYSSHKVPSSSLTGG 987

bip2-PA          1008 IPNQI------KISEEHNIFSTFSNYEDITITPTGLTSLEPKMRKHHKKLKKVKEGKNKK 1061
                      :.. :      :***:::::.**.*****:*****  **:**:      * :: *
Submitted_Seq     988 VASGVIPILPNHISEDQQFMPTFPNYEDISITPTGALSLDPKI------LIQIPE----- 1036

bip2-PA          1062 KKEKKDKSKKADQIGLPSFKSDRKIKANDKRQKKEKKKDKDKQILVHIPDDTEEFDKVPL 1121
                           *: .:*  *                       :: *  :   :* .*
Submitted_Seq    1037 ----VDEFSRAPFI-----------------------NNLDYGFTQSVPAST-------- 1061

bip2-PA          1122 ANNDEPVLKSSSMTINPSLGAATSGI-----SPNQIPKLTLKLSGKSTLFSSSEKEMTDA 1176
                           ****:*.*:::*.*.*: :       **.**************  ...**: ***
Submitted_Seq    1062 -----PVLKTSPMSVKPPLSATCAMSPKIQQSPSQIPKLTLKLSGKSTTCPEKEKDTTDA 1116

bip2-PA          1177 GKLKQTTILSSENKKRERDNSPELARFSPLVTGPPKNKQSETLHLGNSSTAVLPVPSPVA 1236
                       *:.*.*::. ***:************************            ::  * *:*
Submitted_Seq    1117 VKVSQPTMFPVENKERERDNSPELARFSPLVTGPPKNKQ------------LIQEPFPMA 1164

bip2-PA          1237 VRAVQLPVSQTSSNSAGWLSNPNNSNTASSTLSASSVLLPQQLMLAPHTIMNNFVPAMCN 1296
                      .*: *:..**.**.****:***.***.******************:.:* *** :
Submitted_Seq    1165 TRTSQISTSQNSSSSAGWMSNPSNSNVASSTLSASSVLLPQQLMLTSNTTMNNSL----- 1219

bip2-PA          1297 STGTVSKSGLCSSPPNTSEENANAMQIAESSRPSSYVDAEGNRIWICPACGKVDDGSAMI 1356
                      *:*  .**   ***.*..***::   ***:.*****************************
Submitted_Seq    1220 SSGG-PKSCSLSSPANVPEENSH---IAETNRPSSYVDAEGNRIWICPACGKVDDGSAMI 1275

bip2-PA          1357 GCDGCDAWYHWICVGITFAPKDNDDWFCRVCVTKKRIHGSEKKKRRNKKK 1406
                      ************************************:* ***** . :*
Submitted_Seq    1276 GCDGCDAWYHWICVGITFAPKDNDDWFCRVCVTKKRVHVSEKKKGEIRK- 1324
```

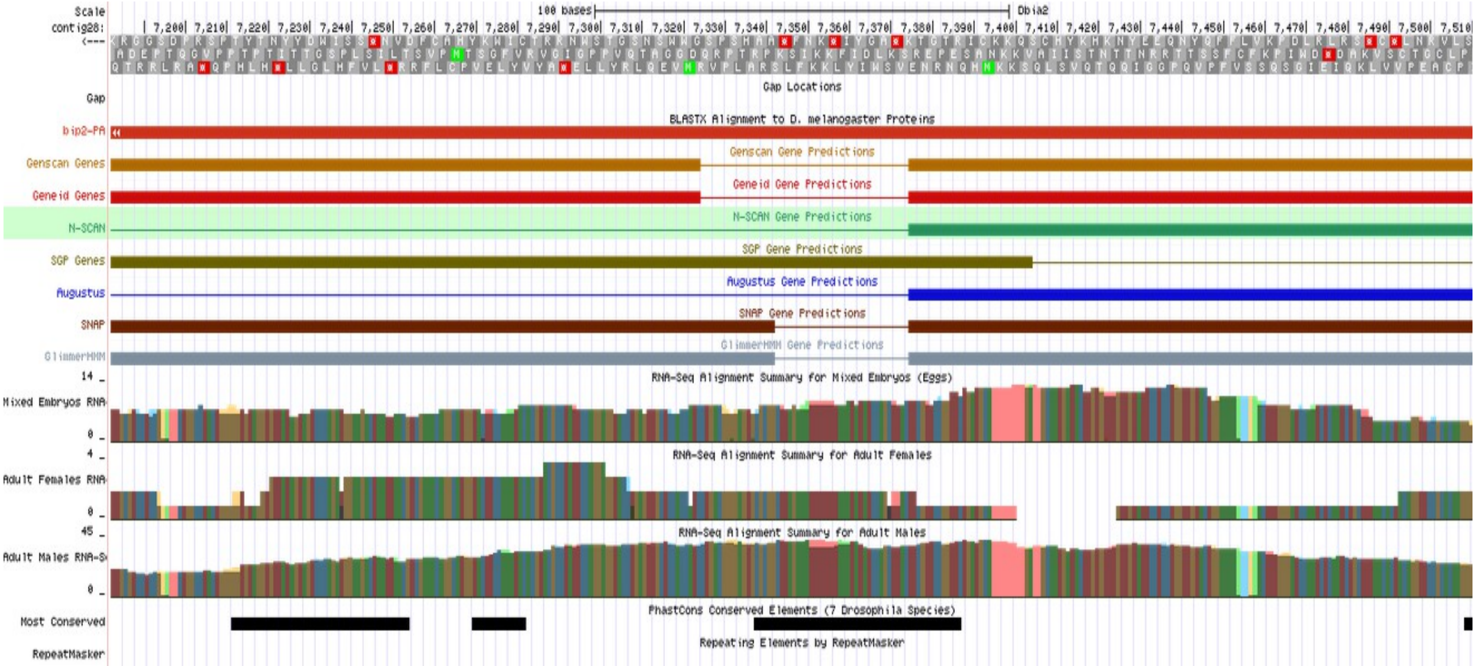
Gene model dot plot



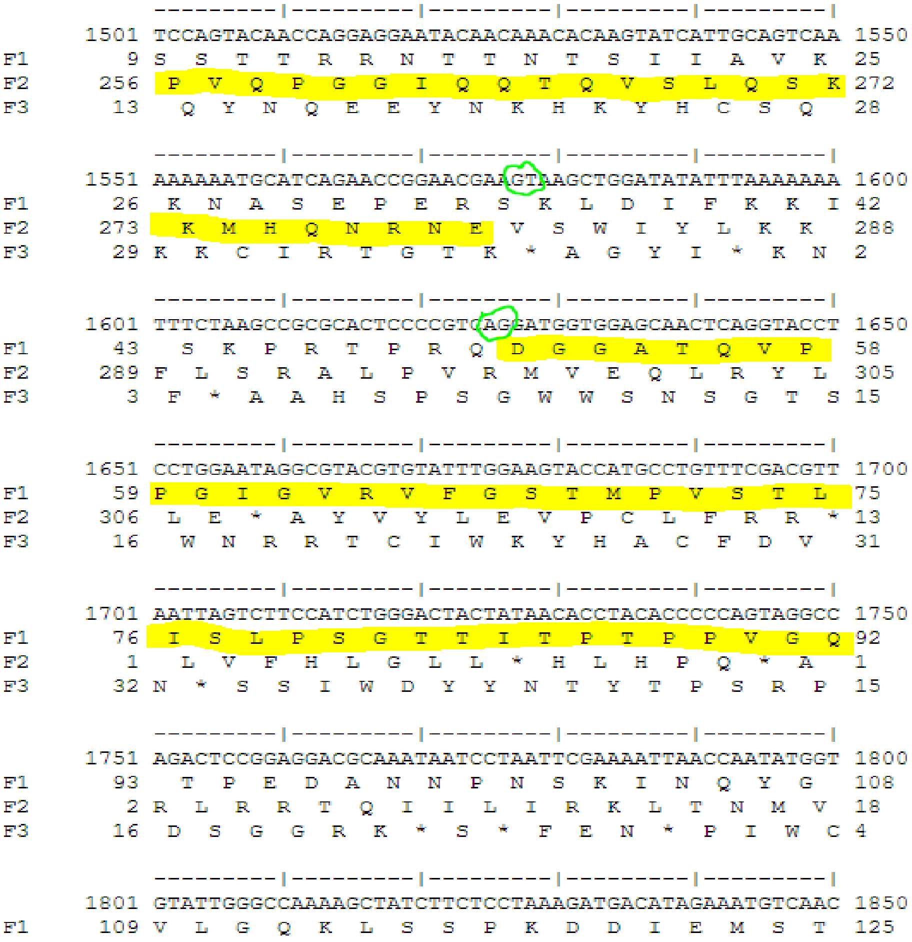
Gain of intron in *D. biarmipes*

The gene has in intron (coordinates 7375 and 7324 on contig28) that is not present in *D. melanogaster.* There is various evidence to support this claim: for CDS 6 is in two different reading frames, which is of course not possible. There is also an appropriate splice donor and splice acceptor site at the 5' and 3' ends of the (two) exons. The intron is above the minimum 40 nucleotides.

The exon spans two reading frames (-3 and -2, respectively), as shown below:

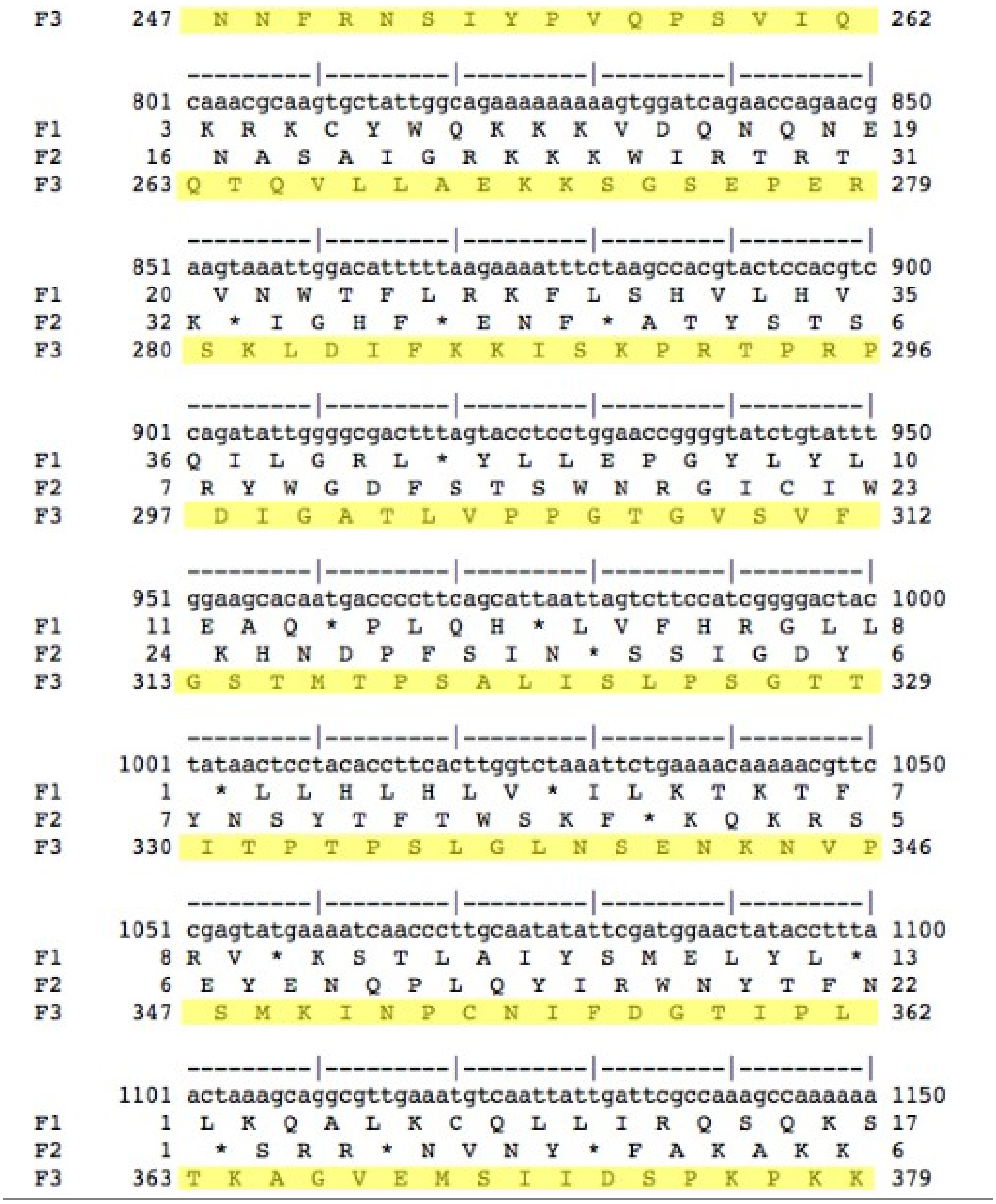


Using the EMBOSS tookit (showorf), I highlighted the peptide sequence in *D. biarmipes* to show the intron and splice sites:



```
           ---------|---------|---------|---------|---------|
      1501 TCCAGTACAACCAGGAGGAATACAACAAACACAAGTATCATTGCAGTCAA 1550
F1       9 S  S  T  T  R  R  N  T  T  N  T  S  I  I  A  V  K  25
F2     256  P  V  Q  P  G  G  I  Q  Q  T  Q  V  S  L  Q  S  K 272
F3      13   Q  Y  N  Q  E  E  Y  N  K  H  K  Y  H  C  S  Q   28

           ---------|---------|---------|---------|---------|
      1551 AAAAAATGCATCAGAACCGGAACGAAGTAAGCTGGATATATTTAAAAAAA 1600
F1      26  K  N  A  S  E  P  E  R  S  K  L  D  I  F  K  K  I 42
F2     273   K  M  H  Q  N  R  N  E  V  S  W  I  Y  L  K  K   288
F3      29 K  K  C  I  R  T  G  T  K  *  A  G  Y  I  *  K  N  2

           ---------|---------|---------|---------|---------|
      1601 TTTCTAAGCCGCGCACTCCCCGTCAGGATGGTGGAGCAACTCAGGTACCT 1650
F1      43   S  K  P  R  T  P  R  Q  D  G  G  A  T  Q  V  P   58
F2     289 F  L  S  R  A  L  P  V  R  M  V  E  Q  L  R  Y  L  305
F3       3  F  *  A  A  H  S  P  S  G  W  W  S  N  S  G  T  S 15

           ---------|---------|---------|---------|---------|
      1651 CCTGGAATAGGCGTACGTGTATTTGGAAGTACCATGCCTGTTTCGACGTT 1700
F1      59 P  G  I  G  V  R  V  F  G  S  T  M  P  V  S  T  L  75
F2     306  L  E  *  A  Y  V  Y  L  E  V  P  C  L  F  R  R  * 13
F3      16   W  N  R  R  T  C  I  W  K  Y  H  A  C  F  D  V   31

           ---------|---------|---------|---------|---------|
      1701 AATTAGTCTTCCATCTGGGACTACTATAACACCTACACCCCCAGTAGGCC 1750
F1      76  I  S  L  P  S  G  T  T  I  T  P  T  P  P  V  G  Q 92
F2       1   L  V  F  H  L  G  L  L  *  H  L  H  P  Q  *  A   1
F3      32 N  *  S  S  I  W  D  Y  Y  N  T  Y  T  P  S  R  P  15

           ---------|---------|---------|---------|---------|
      1751 AGACTCCGGAGGACGCAAATAATCCTAATTCGAAAATTAACCAATATGGT 1800
F1      93   T  P  E  D  A  N  N  P  N  S  K  I  N  Q  Y  G   108
F2       2 R  L  R  R  T  Q  I  I  L  I  R  K  L  T  N  M  V  18
F3      16  D  S  G  G  R  K  *  S  *  F  E  N  *  P  I  W  C 4

           ---------|---------|---------|---------|---------|
      1801 GTATTGGGCCAAAAGCTATCTTCTCCTAAAGATGACATAGAAATGTCAAC 1850
F1     109 V  L  G  Q  K  L  S  S  P  K  D  D  I  E  M  S  T  125
```

Here is the same span of sequence in *D. biarmipes*- note that there is no intron



```
F3      247   N  N  F  R  N  S  I  Y  P  V  Q  P  S  V  I  Q   262

            ---------|---------|---------|---------|---------|
        801 caaacgcaagtgctattggcagaaaaaaaaagtggatcagaaccagaacg 850
F1        3  K  R  K  C  Y  W  Q  K  K  K  V  D  Q  N  Q  N  E 19
F2       16   N  A  S  A  I  G  R  K  K  K  W  I  R  T  R  T   31
F3      263 Q  T  Q  V  L  L  A  E  K  K  S  G  S  E  P  E  R  279

            ---------|---------|---------|---------|---------|
        851 aagtaaattggacatttttaagaaaatttctaagccacgtactccacgtc 900
F1       20   V  N  W  T  F  L  R  K  F  L  S  H  V  L  H  V   35
F2       32 K  *  I  G  H  F  *  E  N  F  *  A  T  Y  S  T  S  6
F3      280  S  K  L  D  I  F  K  K  I  S  K  P  R  T  P  R  P 296

            ---------|---------|---------|---------|---------|
        901 cagatattggggcgactttagtacctcctggaaccggggtatctgtattt 950
F1       36 Q  I  L  G  R  L  *  Y  L  L  E  P  G  Y  L  Y  L  10
F2        7  R  Y  W  G  D  F  S  T  S  W  N  R  G  I  C  I  W 23
F3      297   D  I  G  A  T  L  V  P  P  G  T  G  V  S  V  F   312

            ---------|---------|---------|---------|---------|
        951 ggaagcacaatgaccccttcagcattaattagtcttccatcggggactac 1000
F1       11  E  A  Q  *  P  L  Q  H  *  L  V  F  H  R  G  L  L 8
F2       24   K  H  N  D  P  F  S  I  N  *  S  S  I  G  D  Y   6
F3      313 G  S  T  M  T  P  S  A  L  I  S  L  P  S  G  T  T  329

            ---------|---------|---------|---------|---------|
       1001 tataactcctacaccttcacttggtctaaattctgaaaacaaaaacgttc 1050
F1        1   *  L  L  H  L  H  L  V  *  I  L  K  T  K  T  F   7
F2        7 Y  N  S  Y  T  F  T  W  S  K  F  *  K  Q  K  R  S  5
F3      330  I  T  P  T  P  S  L  G  L  N  S  E  N  K  N  V  P 346

            ---------|---------|---------|---------|---------|
       1051 cgagtatgaaaatcaacccttgcaatatattcgatggaactataccttta 1100
F1        8 R  V  *  K  S  T  L  A  I  Y  S  M  E  L  Y  L  *  13
F2        6  E  Y  E  N  Q  P  L  Q  Y  I  R  W  N  Y  T  F  N 22
F3      347   S  M  K  I  N  P  C  N  I  F  D  G  T  I  P  L   362

            ---------|---------|---------|---------|---------|
       1101 actaaagcaggcgttgaaatgtcaattattgattcgccaaagccaaaaaa 1150
F1        1  L  K  Q  A  L  K  C  Q  L  L  I  R  Q  S  Q  K  S 17
F2        1   *  S  R  R  *  N  V  N  Y  *  F  A  K  A  K  K   6
F3      363 T  K  A  G  V  E  M  S  I  I  D  S  P  K  P  K  K  379
```

**Alignment – DIGATL (*dmel*) = DGGATQ (*biarm*)

Gene Report Form
Gene name: *Asator*
Gene Symbol: *Asator*
Approximate location in project: 2296-8928
Number of isoforms in D. Melongaster: 8
Number of isoforms In this project: 8

Regarding the warnings in the gene model checklists, see the "loss of intron" note at the end

Name of unique isoform based on coding sequence	List of isoforms with identical coding sequences
Asator-PD	
Asator-PE	
Asator-PF	
Asator-PG	
Asator-PH	
Asator-PI	
Asator-PJ	
Asator-PK	

**In case the gene model checklist screenshots are too small, I have included the individual exon coordinates and the isoform coordinates below:

19) 34873-34851
17) 31050-30529
16_1) 31027-30529
16_0) 30582-30529
15) 28535-28170
14) 28057-27668
13) 27605-27402
12) 26948-26880
11) 26822-26511
10) 26460-25657
9) 25433-25294
8) 22001-21857
7) 22001-21764
6) 21680-21287
5) 20894-20316
4) 20894-20290
3) 18880-18781
1) 18717-18486
STOP (exon 1): 18485-18483
STOP (exon 7): 21763-21761
STOP (exon 4): 20289-20287

I:
34873-34851, 31027-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

H:
30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

F:
30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21764

D:
31050-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

J:
34873-34851, 31027-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20290

K:
30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20290

G:
31050-30529, 28535-28170, 28057-27668, 27605-27402, 26822-26511, 26460-25657, 21680-21287, 20894-20290

E:
31050-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20290

Isoform Report Form

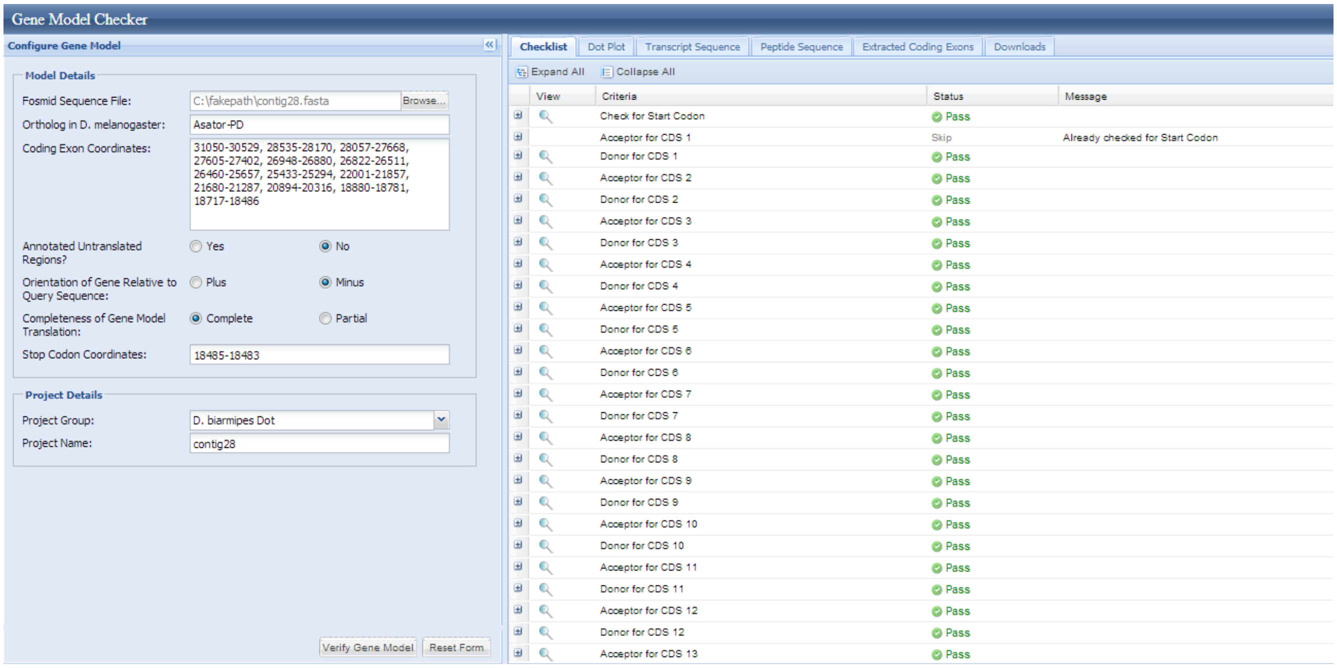
Gene--- isoform name: *Asator-PD*

Names of isoforms with identical coding sequences: none

Is the 5' end of this isoform missing from the end of the project: no

Is the 5' end of this isoform missing from the end of the project: no

Gene Model Checker Checklist:



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PD

Coding Exon Coordinates: 31050-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

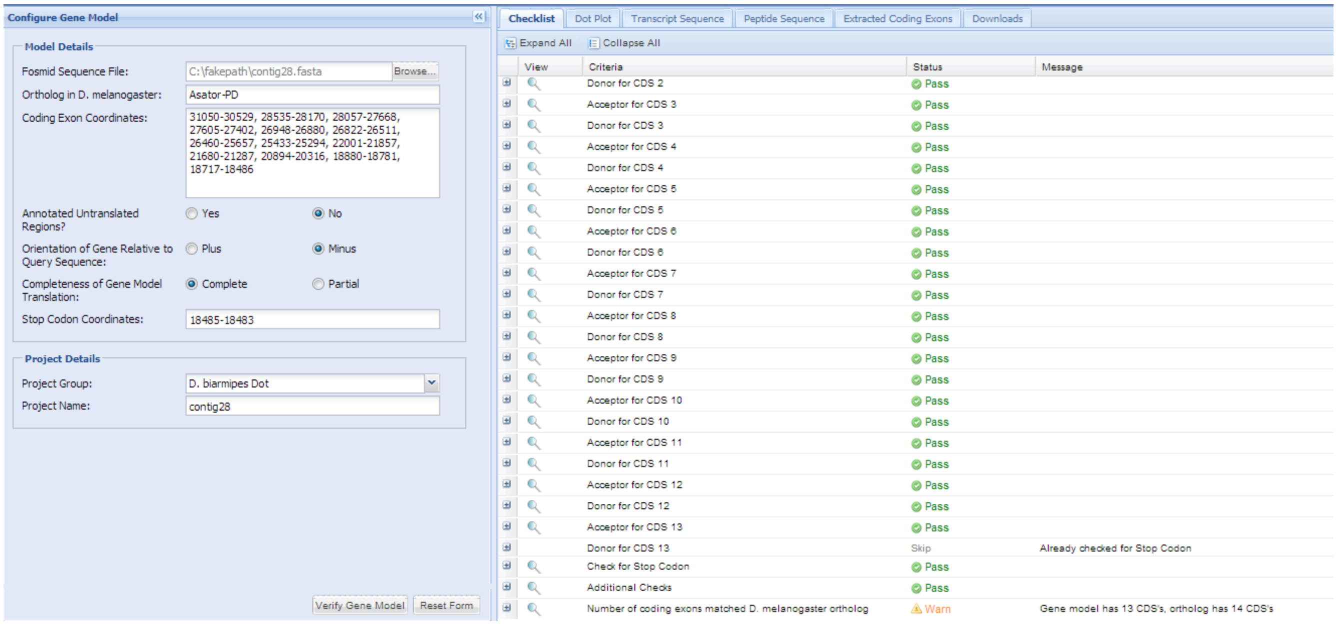
Stop Codon Coordinates: 18485-18483

Project Details

Project Group: D. biarmipes Dot

Project Name: contig28

View	Criteria	Status	Message
	Check for Start Codon	Pass	
	Acceptor for CDS 1	Skip	Already checked for Start Codon
	Donor for CDS 1	Pass	
	Acceptor for CDS 2	Pass	
	Donor for CDS 2	Pass	
	Acceptor for CDS 3	Pass	
	Donor for CDS 3	Pass	
	Acceptor for CDS 4	Pass	
	Donor for CDS 4	Pass	
	Acceptor for CDS 5	Pass	
	Donor for CDS 5	Pass	
	Acceptor for CDS 6	Pass	
	Donor for CDS 6	Pass	
	Acceptor for CDS 7	Pass	
	Donor for CDS 7	Pass	
	Acceptor for CDS 8	Pass	
	Donor for CDS 8	Pass	
	Acceptor for CDS 9	Pass	
	Donor for CDS 9	Pass	
	Acceptor for CDS 10	Pass	
	Donor for CDS 10	Pass	
	Acceptor for CDS 11	Pass	
	Donor for CDS 11	Pass	
	Acceptor for CDS 12	Pass	
	Donor for CDS 12	Pass	
	Acceptor for CDS 13	Pass	



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PD

Coding Exon Coordinates: 31050-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

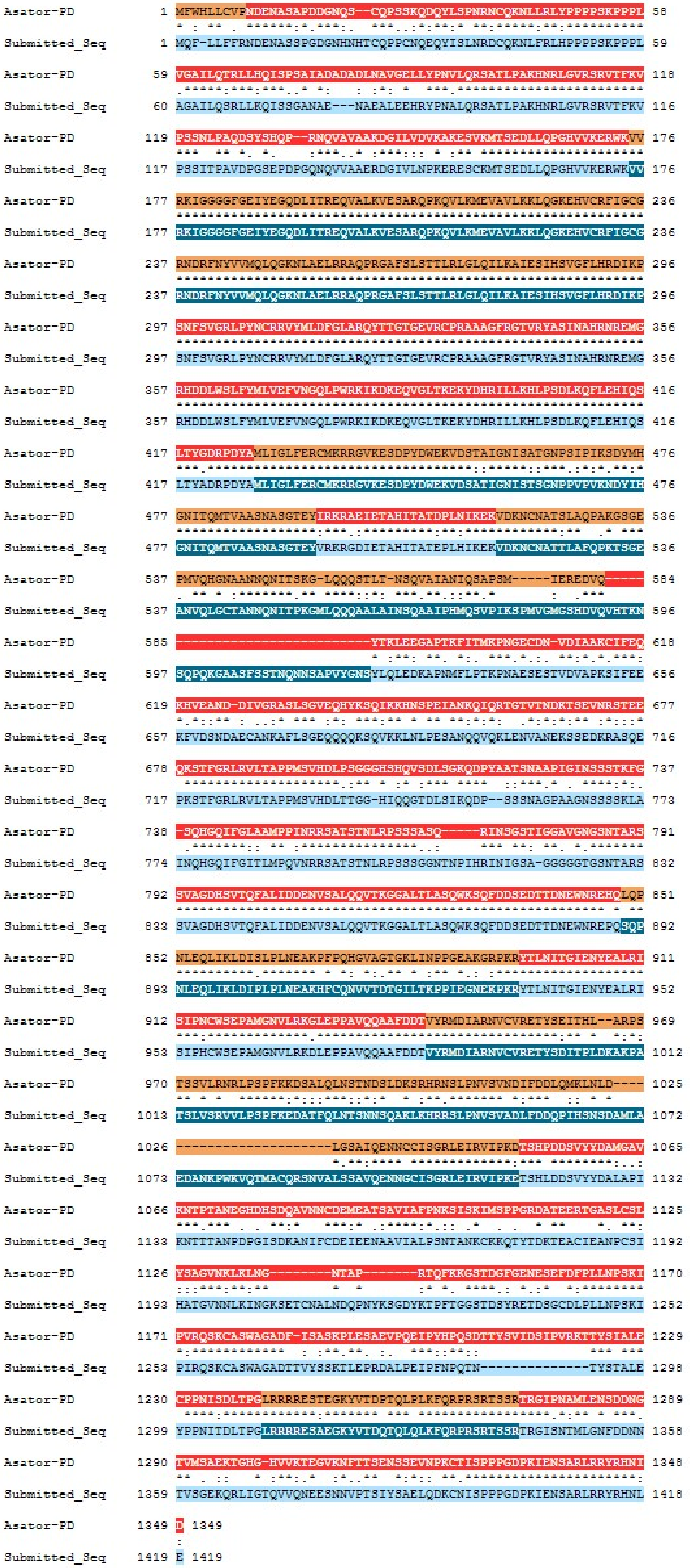
Stop Codon Coordinates: 18485-18483

Project Details

Project Group: D. biarmipes Dot

Project Name: contig28

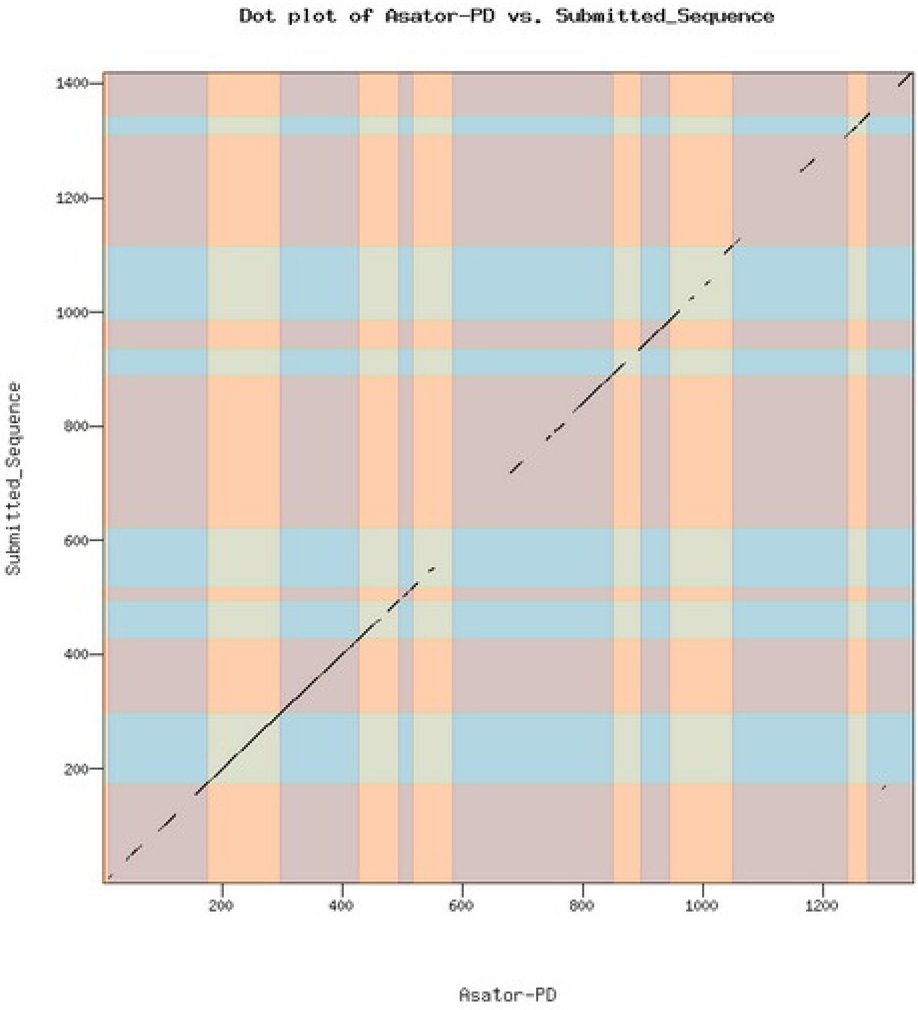
View	Criteria	Status	Message
	Donor for CDS 2	Pass	
	Acceptor for CDS 3	Pass	
	Donor for CDS 3	Pass	
	Acceptor for CDS 4	Pass	
	Donor for CDS 4	Pass	
	Acceptor for CDS 5	Pass	
	Donor for CDS 5	Pass	
	Acceptor for CDS 6	Pass	
	Donor for CDS 6	Pass	
	Acceptor for CDS 7	Pass	
	Donor for CDS 7	Pass	
	Acceptor for CDS 8	Pass	
	Donor for CDS 8	Pass	
	Acceptor for CDS 9	Pass	
	Donor for CDS 9	Pass	
	Acceptor for CDS 10	Pass	
	Donor for CDS 10	Pass	
	Acceptor for CDS 11	Pass	
	Donor for CDS 11	Pass	
	Acceptor for CDS 12	Pass	
	Donor for CDS 12	Pass	
	Acceptor for CDS 13	Pass	
	Donor for CDS 13	Skip	Already checked for Stop Codon
	Check for Stop Codon	Pass	
	Additional Checks	Pass	
	Number of coding exons matched D. melanogaster ortholog	⚠ Warn	Gene model has 13 CDS's, ortholog has 14 CDS's



```
Asator-PD         1 MFWHLLCVPNDENASAPDDGNQS--CQPSSKQDQYLSPNRNCQKNLLRLYPPPPSKPPPL 58
                    * : ** . ******:*.***:.  ***..:*:**:* **:*****:**:**********
Submitted_Seq     1 MQF-LLFFRNDENASSPGDGNHNHTCQPPCNQEQYISLNRDCQKNLFRLHPPPPSKPPPL 59

Asator-PD        59 VGAILQTRLLHQISPSAIADADADLNAVGELLYPNVLQRSATLPAKHNRLGVRSRVTFKV 118
                    .*****:***:***..* *:   : :*: *  ***.************************
Submitted_Seq    60 AGAILQSRLLKQISSGANAE---NAEALEEHRYPNALQRSATLPAKHNRLGVRSRVTFKV 116

Asator-PD       119 PSSNLPAQDSYSHQP--RNQVAVAAKDGILVDVKAKESVKMTSEDLLQPGHVVKERWKVV 176
                    ***  ** *. *.    :***..* :***::: * :** *********************
Submitted_Seq   117 PSSITPAVDPGSEPDPGQNQVVAAERDGIVLNPKERESCKMTSEDLLQPGHVVKERWKVV 176

Asator-PD       177 RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCG 236
                    ************************************************************
Submitted_Seq   177 RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCG 236

Asator-PD       237 RNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKP 296
                    ************************************************************
Submitted_Seq   237 RNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKP 296

Asator-PD       297 SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMG 356
                    ************************************************************
Submitted_Seq   297 SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMG 356

Asator-PD       357 RHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS 416
                    ************************************************************
Submitted_Seq   357 RHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS 416

Asator-PD       417 LTYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAIGNISATGNPSIPIKSDYMH 476
                    ***.**********************************::*****::***.:*:*.**:*
Submitted_Seq   417 LTYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSATIGNISTSGNPPVPVKNDYIH 476

Asator-PD       477 GNITQMTVAASNASGTEYIRKRAEIETAHITATDPLNIKEKVDKNCNATSLAQPAKGSGE 536
                    ******************:***.:*********:**:************:**  .* ***
Submitted_Seq   477 GNITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKEKVDKNCNATTLAFQPKTSGE 536

Asator-PD       537 PMVQHGNAANNQNITSKG-LQQQSTLT-NSQVAIANIQSAPSM-----IEREDVQ----- 584
                    . ** * :*******.** ****::*: ***.**.::**.*       :  .***
Submitted_Seq   537 ANVQLGCTANNQNITPKGMLQQQAALAINSQAAIPHMQSVPIKSPMVGMGSHDVQVHTKN 596

Asator-PD       585 -------------------------YTKLEEGAPTKFITMKPNGECDN-VDIAAKCIFEQ 618
                                             * :**: **. *:. ***.*.:. **:*.*.***:
Submitted_Seq   597 SQPQKGAASFSSTNQNNSAPVYGNSYLQLEDKAPNMFLPTKPNAESESTVDVAPKSIFEE 656

Asator-PD       619 KHVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQIQRTGTVTNDKTSEVNRSTEE 677
                    *.*::** : ..:* *** :*: ***:** * ** **:*:*:  .*:*:*:** :*:::*
Submitted_Seq   657 KFVDSNDAECANKAFLSGEQQQQKSQVKKLNLPESANQQVQKLENVANEKSSEDKRASQE 716

Asator-PD       678 QKSTFGRLRVLTAPPMSVHDLPSGGGHSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG 737
                     ********************.:** * :* :*** ****  ::***.* . ****:*:.
Submitted_Seq   717 PKSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSIKQDP--SSSNAGPAAGNSSSSKLA 773

Asator-PD       738 -SQHGQIFGLAAMPPINRRSATSTNLRPSSSASQ-----RINSGSTIGGAVGNGSNTARS 791
                     .*******:: ** :***************..:     *** **: **. *.*******
Submitted_Seq   774 INQHGQIFGITLMPQVNRRSATSTNLRPSSSGGNTNPIHRINIGSA-GGGGGTGSNTARS 832

Asator-PD       792 SVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREHQLQP 851
                    ******************************************************* * **
Submitted_Seq   833 SVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREPQSQP 892

Asator-PD       852 NLEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEAKGRPKRYTLNITGIENYEALRI 911
                    **********.******* * *: *:.** * :** *.: :*******************
Submitted_Seq   893 NLEQLIKLDIPLPLNEAKHFCQNVVTDTGILTKPPIEGNEKPKRYTLNITGIENYEALRI 952

Asator-PD       912 SIPNCWSEPAMGNVLRKGLEPPAVQQAAFDDTVYRMDIARNVCVRETYSEITHL--ARPS 969
                    ***:*************.*******************************:** *  *:*:
Submitted_Seq   953 SIPHCWSEPAMGNVLRKDLEPPAVQQAAFDDTVYRMDIARNVCVRETYSDITPLDKAKPA 1012

Asator-PD       970 TSSVLRNRLPSPFKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDLQMKLNLD---- 1025
                    ** * *  ******:*:::***::*:*  * :**.******* *:***  :: * *
Submitted_Seq  1013 TSLVSRVVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHSNSDAMLA 1072

Asator-PD      1026 --------------------LGSAIQENNCCISGRLEIRVIPKDTSHPDDSVYYDAMGAV 1065
                                        *.**:**** *************:*** ********:..:
Submitted_Seq  1073 EDANKPWKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKETSHLDDSVYYDALAPI 1132

Asator-PD      1066 KNTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKIMSPPGRDATEERTGASLCSL 1125
                    ***.*** . . **:*   ***:* .:****:*.:: .*  . .  * **    *. **:
Submitted_Seq  1133 KNTTTANPDPGISDKANIFCDEIEENAAVIALPSNTANKCKKQTYTDKTEACIEANPCSI 1192

Asator-PD      1126 YSAGVNKLKLNG--------NTAP-------RTQFKKGSTDGFGENESEFDFPLLNPSKI 1170
                    :::***:**:**        *  *       :* *. ****.: *.:*  *:********
Submitted_Seq  1193 HATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGSTDSYRETDSGCDLPLLNPSKI 1252

Asator-PD      1171 PVRQSKCASWAGADF-ISASKPLESAEVPQEIPYHPQSDTTYSVIDSIPVRKTTYSIALE 1229
                    *:************  : :**.**. :.  ***::**::              *** ***
Submitted_Seq  1253 PIRQSKCASWAGADTTVYSSKTLEPRDALPEIPFNPQTN--------------TYSTALE 1298

Asator-PD      1230 CPPNISDLTPGLRRRRESTEGKYVTDPTQLPLKFQRPRSRTSSRTRGIPNAMLENSDDNG 1289
                     ****:************:******* *** *****************.*:** * ***.
Submitted_Seq  1299 YPPNITDLTPGLRRRRESAEGKYVTDQTQLQLKFQRPRSRTSSRTRGISNTMLGNFDDNN 1358

Asator-PD      1290 TVMSAEKTGHG-HVVKTEGVKNFTTSENSSEVNPKCTISPPPGDPKIENSARLRRYRHNI 1348
                    ** . ::   * :**:.*  :*..**  *:*:: **.**********************:
Submitted_Seq  1359 TVSGEKQRLIGTQVVQNEESNNVPTSIYSAELQDKCNISPPPGDPKIENSARLRRYRHNL 1418

Asator-PD      1349 D 1349
                    :
Submitted_Seq  1419 E 1419
```

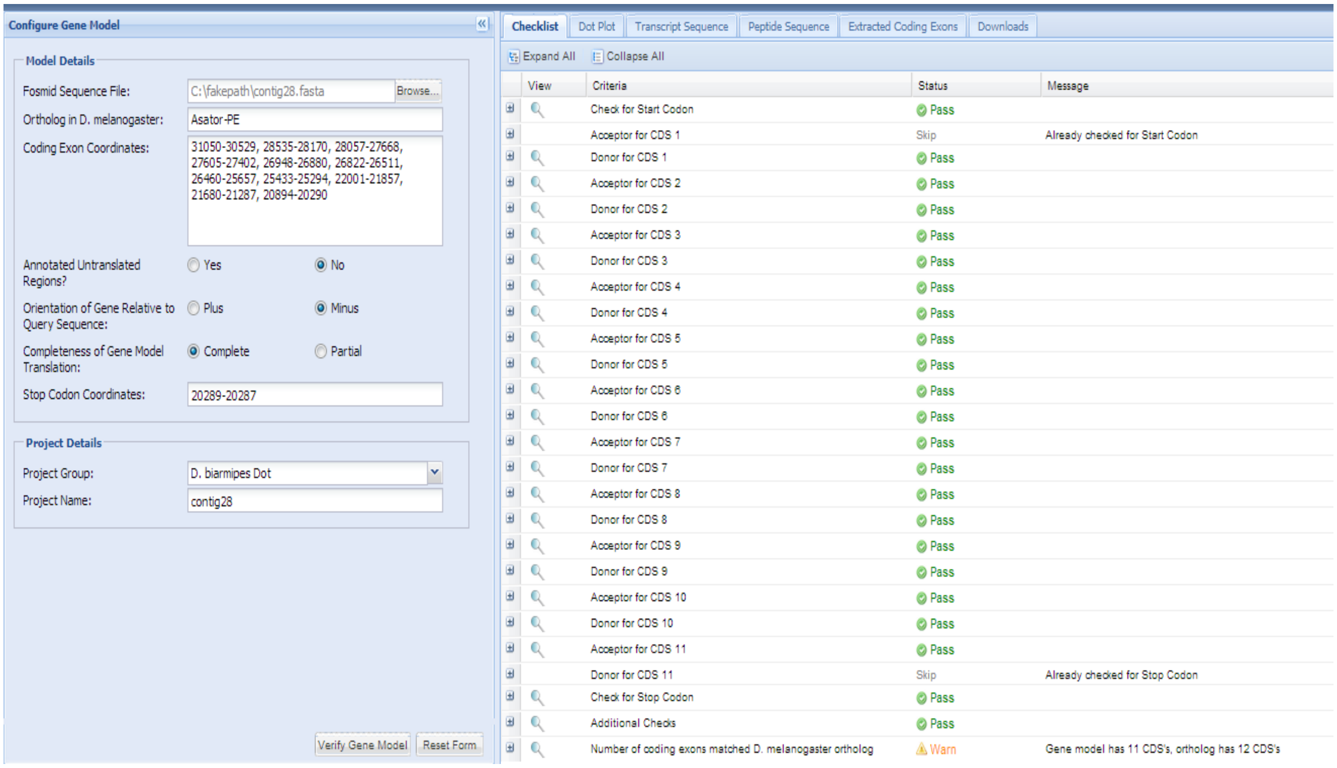
Dotplot

View protein alignment



Isoform Report Form
Gene--- isoform name: *Asator-PE*
Names of isoforms with identical coding sequences: none
Is the 5' end of this isoform missing from the end of the project: no
Is the 5' end of this isoform missing from the end of the project: no

Checklist



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PE

Coding Exon Coordinates: 31050-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20290

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

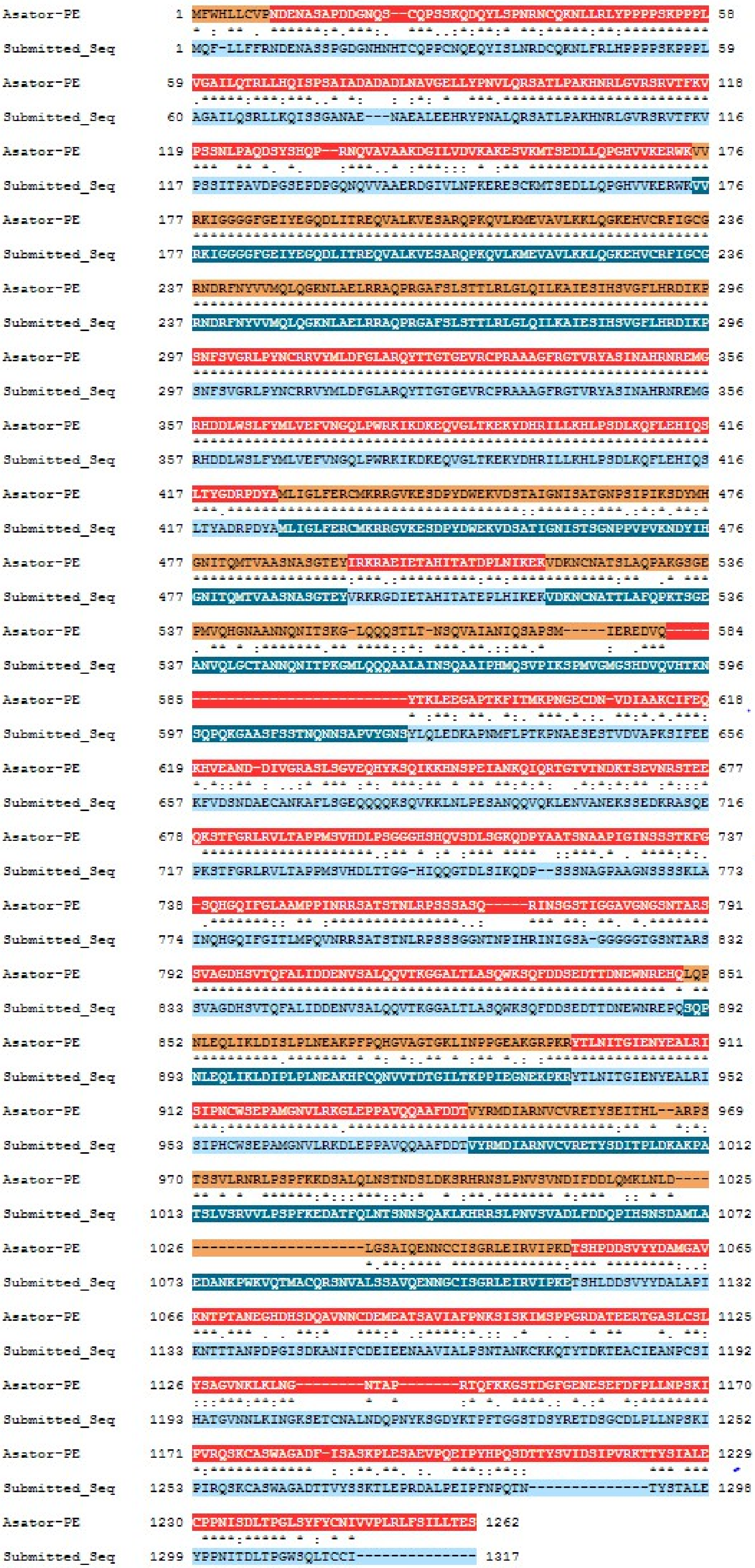
Stop Codon Coordinates: 20289-20287

Project Details

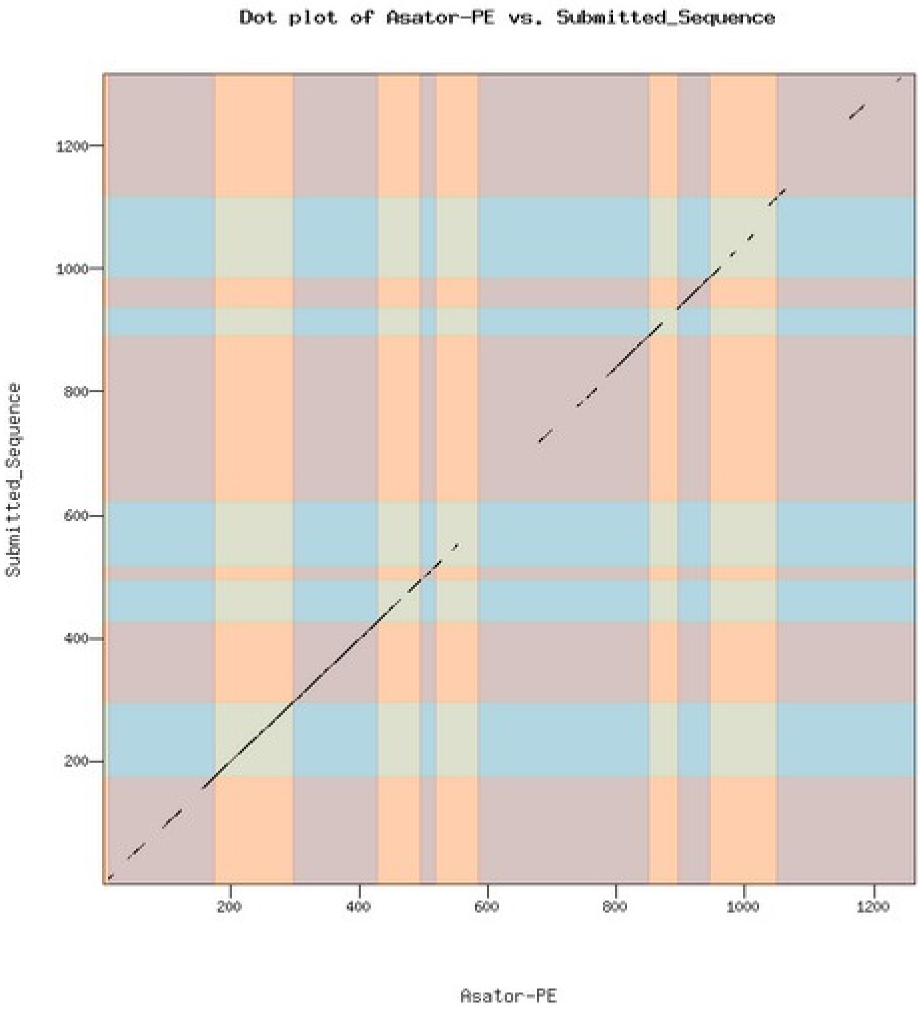
Project Group: D. biarmipes Dot

Project Name: contig28

Criteria	Status	Message
Check for Start Codon	Pass	
Acceptor for CDS 1	Skip	Already checked for Start Codon
Donor for CDS 1	Pass	
Acceptor for CDS 2	Pass	
Donor for CDS 2	Pass	
Acceptor for CDS 3	Pass	
Donor for CDS 3	Pass	
Acceptor for CDS 4	Pass	
Donor for CDS 4	Pass	
Acceptor for CDS 5	Pass	
Donor for CDS 5	Pass	
Acceptor for CDS 6	Pass	
Donor for CDS 6	Pass	
Acceptor for CDS 7	Pass	
Donor for CDS 7	Pass	
Acceptor for CDS 8	Pass	
Donor for CDS 8	Pass	
Acceptor for CDS 9	Pass	
Donor for CDS 9	Pass	
Acceptor for CDS 10	Pass	
Donor for CDS 10	Pass	
Acceptor for CDS 11	Pass	
Donor for CDS 11	Skip	Already checked for Stop Codon
Check for Stop Codon	Pass	
Additional Checks	Pass	
Number of coding exons matched D. melanogaster ortholog	Warn	Gene model has 11 CDS's, ortholog has 12 CDS's

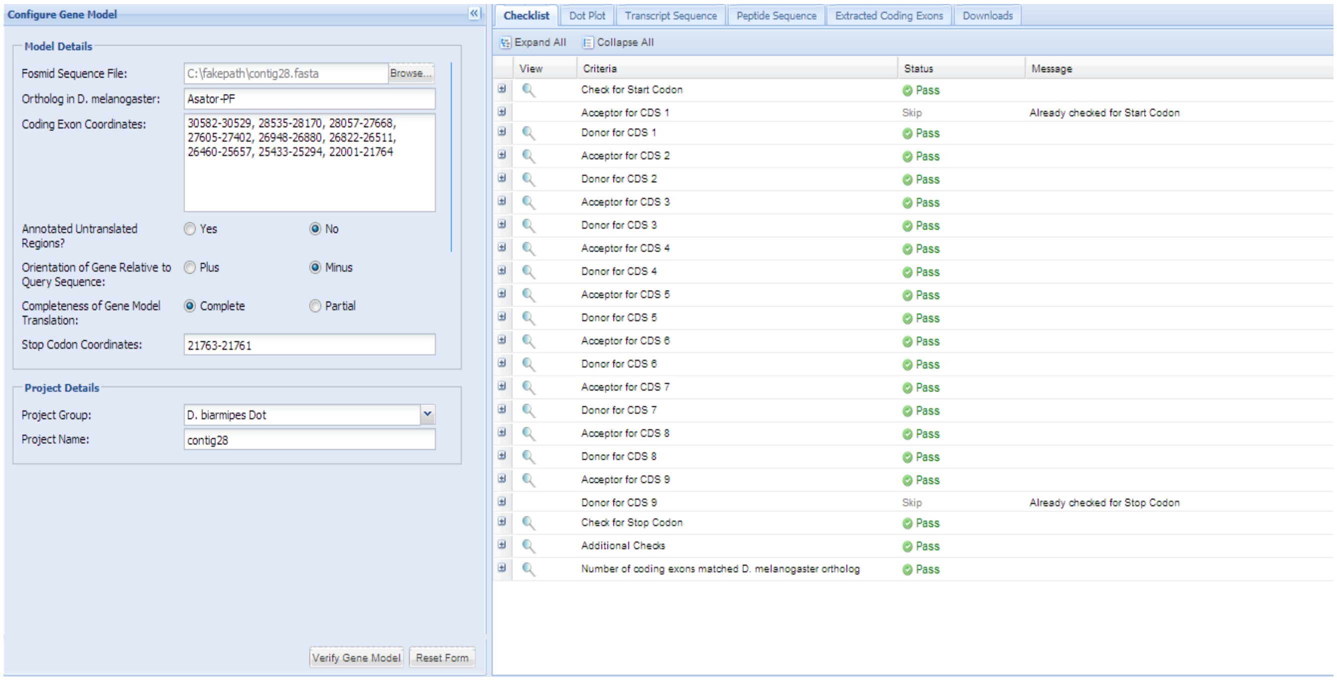


```
Asator-PE         1 MFWHLLCVPNDENASAPDDGNQS--CQPSSKQDQYLSPNRNCQKNLLRLYPPPPSKPPPL 58
                    * : ** . ******:*.***:.  ***..:*:**:* **:*****:**:**********
Submitted_Seq     1 MQF-LLFFRNDENASSPGDGNHNHTCQPPCNQEQYISLNRDCQKNLFRLHPPPPSKPPPL 59

Asator-PE        59 VGAILQTRLLHQISPSAIADADADLNAVGELLYPNVLQRSATLPAKHNRLGVRSRVTFKV 118
                    .*****:***:***..* *:   : :*: *  ***.************************
Submitted_Seq    60 AGAILQSRLLKQISSGANAE---NAEALEEHRYPNALQRSATLPAKHNRLGVRSRVTFKV 116

Asator-PE       119 PSSNLPAQDSYSHQP--RNQVAVAAKDGILVDVKAKESVKMTSEDLLQPGHVVKERWKVV 176
                    ***  ** *. *.    :***..* :***::: * :** *********************
Submitted_Seq   117 PSSITPAVDPGSEPDPGQNQVVAAERDGIVLNPKERESCKMTSEDLLQPGHVVKERWKVV 176

Asator-PE       177 RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCG 236
                    ************************************************************
Submitted_Seq   177 RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCG 236

Asator-PE       237 RNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKP 296
                    ************************************************************
Submitted_Seq   237 RNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKP 296

Asator-PE       297 SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMG 356
                    ************************************************************
Submitted_Seq   297 SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMG 356

Asator-PE       357 RHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS 416
                    ************************************************************
Submitted_Seq   357 RHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS 416

Asator-PE       417 LTYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAIGNISATGNPSIPIKSDYMH 476
                    ***.**********************************::*****::***.:*:*.**:*
Submitted_Seq   417 LTYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSATIGNISTSGNPPVPVKNDYIH 476

Asator-PE       477 GNITQMTVAASNASGTEYIRKRAEIETAHITATDPLNIKERVDKNCNATSLAQPAKGSGE 536
                    ******************:***.:*********:**:***:********:**  .* ***
Submitted_Seq   477 GNITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKEKVDKNCNATTLAFQPKTSGE 536

Asator-PE       537 PMVQHGNAANNQNITSKG-LQQQSTLT-NSQVAIANIQSAPSM-----IEREDVQ----- 584
                    . ** * :*******.** ****::*: ***.**.::**.*       :  .***
Submitted_Seq   537 ANVQLGCTANNQNITPKGMLQQQAALAINSQAAIPHMQSVPIKSPMVGMGSHDVQVHTKN 596

Asator-PE       585 -------------------------YTKLEEGAPTKFITMKPNGECDN-VDIAAKCIFEQ 618
                                             * :**: **. *:. ***.*.:. **:*.*.***:
Submitted_Seq   597 SQPQKGAASFSSTNQNNSAPVYGNSYLQLEDKAPNMFLPTKPNAESESTVDVAPKSIFEE 656

Asator-PE       619 KHVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQIQRTGTVTNDKTSEVNRSTEE 677
                    *.*::** : ..:* *** :*: ***:** * ** **:*:*:  .*:*:*:** :*:::*
Submitted_Seq   657 KFVDSNDAECANKAFLSGEQQQQKSQVKKLNLPESANQQVQKLENVANEKSSEDKRASQE 716

Asator-PE       678 QKSTFGRLRVLTAPPMSVHDLPSGGGHSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG 737
                     ********************.:** * :* :*** ****  ::***.* . ****:*:.
Submitted_Seq   717 PKSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSIKQDP--SSSNAGPAAGNSSSSKLA 773

Asator-PE       738 -SQHGQIFGLAAMPPINRRSATSTNLRPSSSASQ-----RINSGSTIGGAVGNGSNTARS 791
                     .*******:: ** :***************..:     *** **: **. *.*******
Submitted_Seq   774 INQHGQIFGITLMPQVNRRSATSTNLRPSSSGGNTNPIHRINIGSA-GGGGGTGSNTARS 832

Asator-PE       792 SVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREHQLQP 851
                    ******************************************************* * **
Submitted_Seq   833 SVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREPQSQP 892

Asator-PE       852 NLEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEAKGRPKRYTLNITGIENYEALRI 911
                    **********.******* * *: *:.** * :** *.: :*******************
Submitted_Seq   893 NLEQLIKLDIPLPLNEAKHFCQNVVTDTGILTKPPIEGNEKPKRYTLNITGIENYEALRI 952

Asator-PE       912 SIPNCWSEPAMGNVLRKGLEPPAVQQAAFDDTVYRMDIARNVCVRETYSEITHL--ARPS 969
                    ***:*************.*******************************:** *  *:*:
Submitted_Seq   953 SIPHCWSEPAMGNVLRKDLEPPAVQQAAFDDTVYRMDIARNVCVRETYSDITPLDKAKPA 1012

Asator-PE       970 TSSVLRNRLPSPFKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDLQMKLNLD---- 1025
                    ** * *  ******:*:::***::*:*  * :**.******* *:***  :: * *
Submitted_Seq  1013 TSLVSRVVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHSNSDAMLA 1072

Asator-PE      1026 --------------------LGSAIQENNCCISGRLEIRVIPKDTSHPDDSVYYDAMGAV 1065
                                        *.**:**** *************:*** ********:..:
Submitted_Seq  1073 EDANKPWKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKETSHLDDSVYYDALAPI 1132

Asator-PE      1066 KNTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKIMSPPGRDATEERTGASLCSL 1125
                    ***.*** . . **:*   ***:* .:****:*.:: .*  . .  * **    *. **:
Submitted_Seq  1133 KNTTTANPDPGISDKANIFCDEIEENAAVIALPSNTANKCKKQTYTDKTEACIEANPCSI 1192

Asator-PE      1126 YSAGVNKLKLNG--------NTAP-------RTQFKKGSTDGFGENESEFDFPLLNPSKI 1170
                    :::***:**:**        *  *       :* *. ****.: *.:*  *:********
Submitted_Seq  1193 HATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGSTDSYRETDSGCDLPLLNPSKI 1252

Asator-PE      1171 PVRQSKCASWAGADF-ISASKPLESAEVPQEIPYHPQSDTTYSVIDSIPVRKTTYSIALE 1229
                    *:************  : :**.**. :.  ***::**::              *** ***
Submitted_Seq  1253 PIRQSKCASWAGADTTVYSSKTLEPRDALPEIPFNPQTN--------------TYSTALE 1298

Asator-PE      1230 CPPNISDLTPGLSYFYCNIVVPLRLFSILLTES 1262
                     ****:***** * : * *
Submitted_Seq  1299 YPPNITDLTPGWSQLTCCI-------------- 1317
```


Dot plot of Asator-PE vs. Submitted_Sequence
Submitted_Sequence
1200
1000
800
600
400
200
200
400
600
800
1000
1200
Asator-PE

Isoform Report Form
Gene--- isoform name: *Asator-PF*
Names of isoforms with identical coding sequences: none
Is the 5' end of this isoform missing from the end of the project: no
Is the 5' end of this isoform missing from the end of the project: no

Checklist



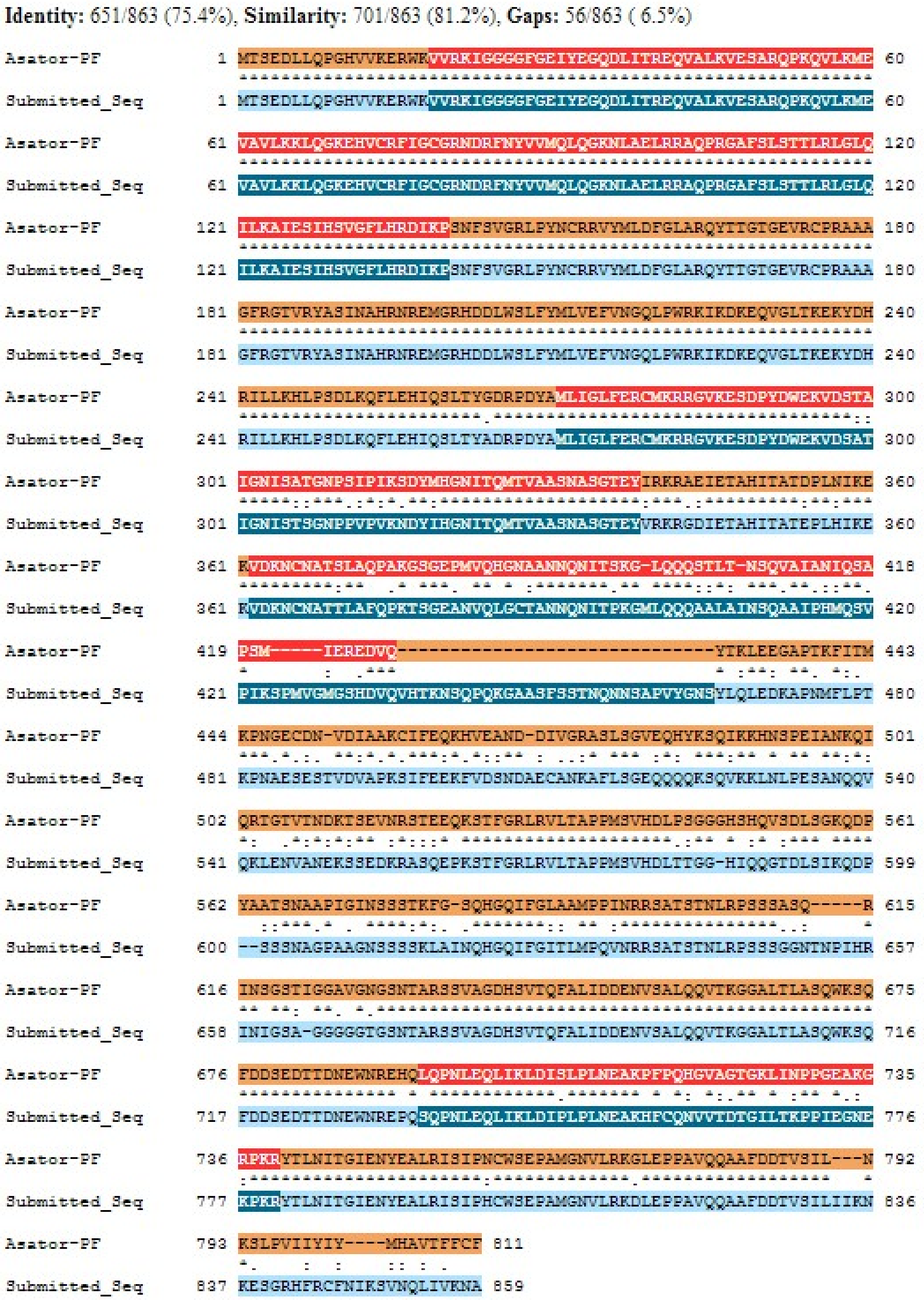
Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta
Ortholog in D. melanogaster: Asator-PF
Coding Exon Coordinates: 30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21764
Annotated Untranslated Regions? ○ Yes ● No
Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus
Completeness of Gene Model Translation: ● Complete ○ Partial
Stop Codon Coordinates: 21763-21761

Project Details

Project Group: D. biarmipes Dot
Project Name: contig28

Criteria	Status	Message
Check for Start Codon	Pass	
Acceptor for CDS 1	Skip	Already checked for Start Codon
Donor for CDS 1	Pass	
Acceptor for CDS 2	Pass	
Donor for CDS 2	Pass	
Acceptor for CDS 3	Pass	
Donor for CDS 3	Pass	
Acceptor for CDS 4	Pass	
Donor for CDS 4	Pass	
Acceptor for CDS 5	Pass	
Donor for CDS 5	Pass	
Acceptor for CDS 6	Pass	
Donor for CDS 6	Pass	
Acceptor for CDS 7	Pass	
Donor for CDS 7	Pass	
Acceptor for CDS 8	Pass	
Donor for CDS 8	Pass	
Acceptor for CDS 9	Pass	
Donor for CDS 9	Skip	Already checked for Stop Codon
Check for Stop Codon	Pass	
Additional Checks	Pass	
Number of coding exons matched D. melanogaster ortholog	Pass	

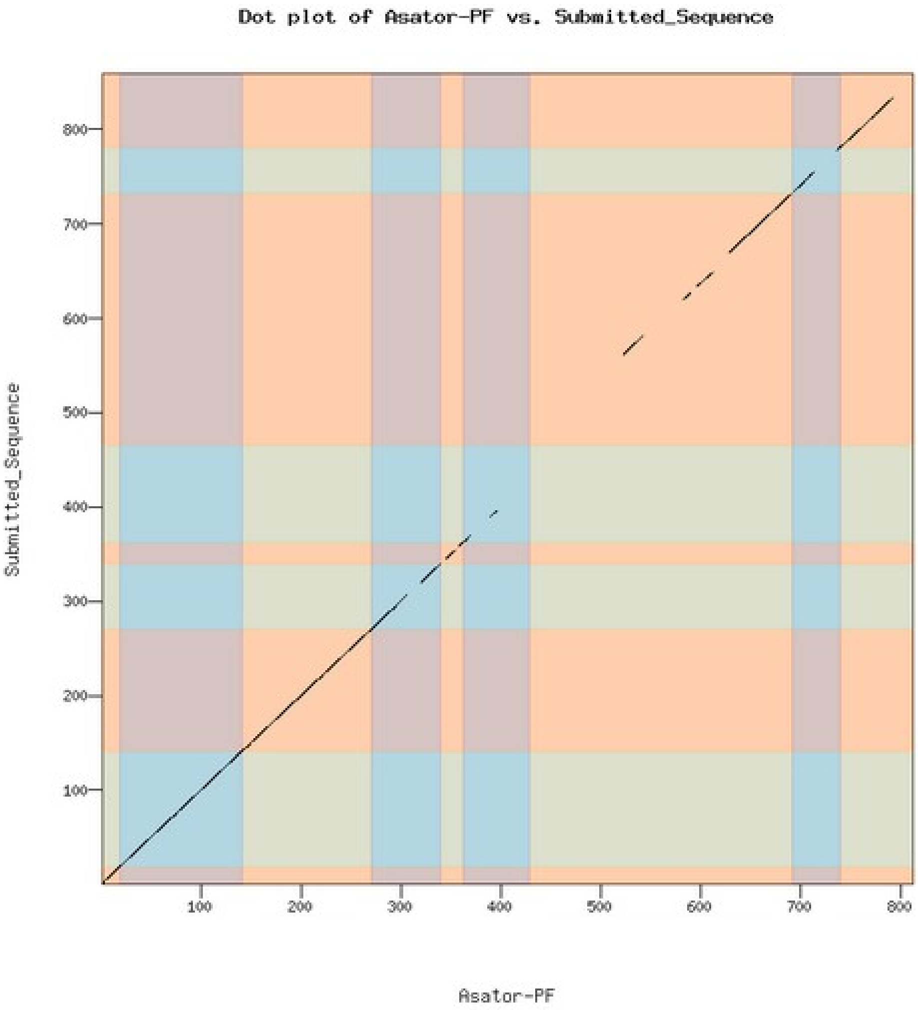


Identity: 651/863 (75.4%), **Similarity:** 701/863 (81.2%), **Gaps:** 56/863 (6.5%)

```
Asator-PF          1 MTSEDLLQPGHVVKERWKVVRKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKME 60
                     ************************************************************
Submitted_Seq      1 MTSEDLLQPGHVVKERWKVVRKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKME 60

Asator-PF         61 VAVLKKLQGKEHVCRFIGCGRNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQ 120
                     ************************************************************
Submitted_Seq     61 VAVLKKLQGKEHVCRFIGCGRNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQ 120

Asator-PF        121 ILKAIESIHSVGFLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA 180
                     ************************************************************
Submitted_Seq    121 ILKAIESIHSVGFLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA 180

Asator-PF        181 GFRGTVRYASINAHRNREMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH 240
                     ************************************************************
Submitted_Seq    181 GFRGTVRYASINAHRNREMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH 240

Asator-PF        241 RILLKHLPSDLKQFLEHIQSLTYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTA 300
                     **********************.***********************************::
Submitted_Seq    241 RILLKHLPSDLKQFLEHIQSLTYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSAT 300

Asator-PF        301 IGNISATGNPSIPIKSDYMHGNITQMTVAASNASGTEYIRKRAEIETAHITATDPLNIKE 360
                     *****::***.:*:*.**:*******************:***.:*********:**:***
Submitted_Seq    301 IGNISTSGNPPVPVKNDYIHGNITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKE 360

Asator-PF        361 KVDKNCNATSLAQPAKGSGEPMVQHGNAANNQNITSKG-LQQQSTLT-NSQVAIANIQSA 418
                     *********:**  .* ***. ** * :*******.** ****::*: ***.**.::**.
Submitted_Seq    361 KVDKNCNATTLAFQPKTSGEANVQLGCTANNQNITPKGMLQQQAALAINSQAAIPHMQSV 420

Asator-PF        419 PSM-----IEREDVQ------------------------------YTKLEEGAPTKFITM 443
                     *       :  .***                              * :**: **. *:.
Submitted_Seq    421 PIKSPMVGMGSHDVQVHTKNSQPQKGAASFSSTNQNNSAPVYGNSYLQLEDKAPNMFLPT 480

Asator-PF        444 KPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQI 501
                     ***.*.:. **:*.*.***:*.*::** : ..:* *** :*: ***:** * ** **:*:
Submitted_Seq    481 KPNAESESTVDVAPKSIFEEKFVDSNDAECANKAFLSGEQQQQKSQVKKLNLPESANQQV 540

Asator-PF        502 QRTGTVTNDKTSEVNRSTEEQKSTFGRLRVLTAPPMSVHDLPSGGGHSHQVSDLSGKQDP 561
                     *:  .*:*:*:** :*:::* ********************.:** * :* :*** ****
Submitted_Seq    541 QKLENVANEKSSEDKRASQEPKSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSIKQDP 599

Asator-PF        562 YAATSNAAPIGINSSSTKFG-SQHGQIFGLAAMPPINRRSATSTNLRPSSSASQ-----R 615
                       ::***.* . ****:*:. .*******:: ** :***************..:     *
Submitted_Seq    600 --SSSNAGPAAGNSSSSKLAINQHGQIFGITLMPQVNRRSATSTNLRPSSSGGNTNPIHR 657

Asator-PF        616 INSGSTIGGAVGNGSNTARSSVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQ 675
                     ** **: **. *.***********************************************
Submitted_Seq    658 INIGSA-GGGGGTGSNTARSSVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQ 716

Asator-PF        676 FDDSEDTTDNEWNREHQLQPNLEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEAKG 735
                     *************** * ************.******* * *: *:.** * :** *.: 
Submitted_Seq    717 FDDSEDTTDNEWNREPQSQPNLEQLIKLDIPLPLNEAKHFCQNVVTDTGILTKPPIEGNE 776

Asator-PF        736 RPKRYTLNITGIENYEALRISIPNCWSEPAMGNVLRKGLEPPAVQQAAFDDTVSIL---N 792
                     :**********************:*************.******************   *
Submitted_Seq    777 KPKRYTLNITGIENYEALRISIPHCWSEPAMGNVLRKDLEPPAVQQAAFDDTVSILIIKN 836

Asator-PF        793 KSLPVIIYIY----MHAVTFFCF 811
                     *.    :  :    :: : .
Submitted_Seq    837 KESGRHFRCFNIKSVNQLIVKNA 859
```

Dotplot

View protein alignment



Isoform Report Form

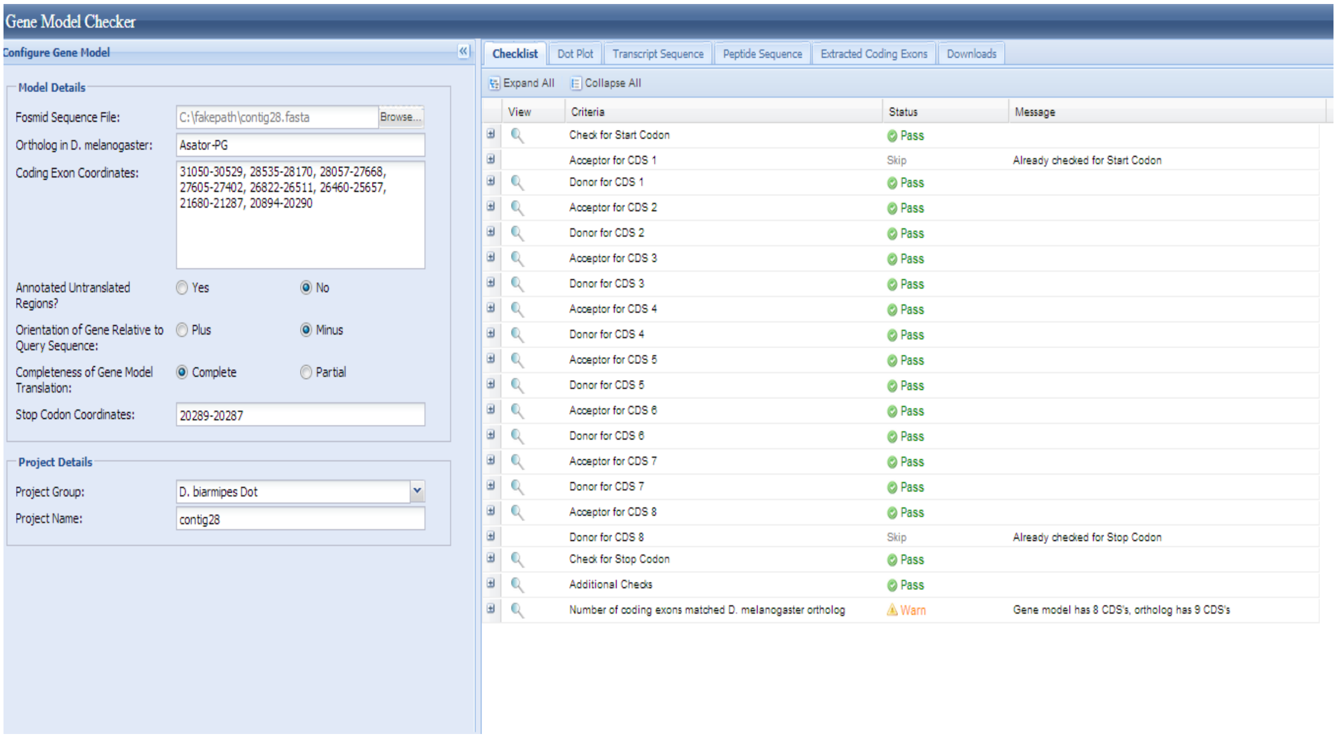
Gene--- isoform name: *Asator-PG*

Names of isoforms with identical coding sequences: none

Is the 5' end of this isoform missing from the end of the project: no

Is the 5' end of this isoform missing from the end of the project: no

Checklist



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PG

Coding Exon Coordinates: 31050-30529, 28535-28170, 28057-27668, 27605-27402, 26822-26511, 26460-25657, 21680-21287, 20894-20290

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

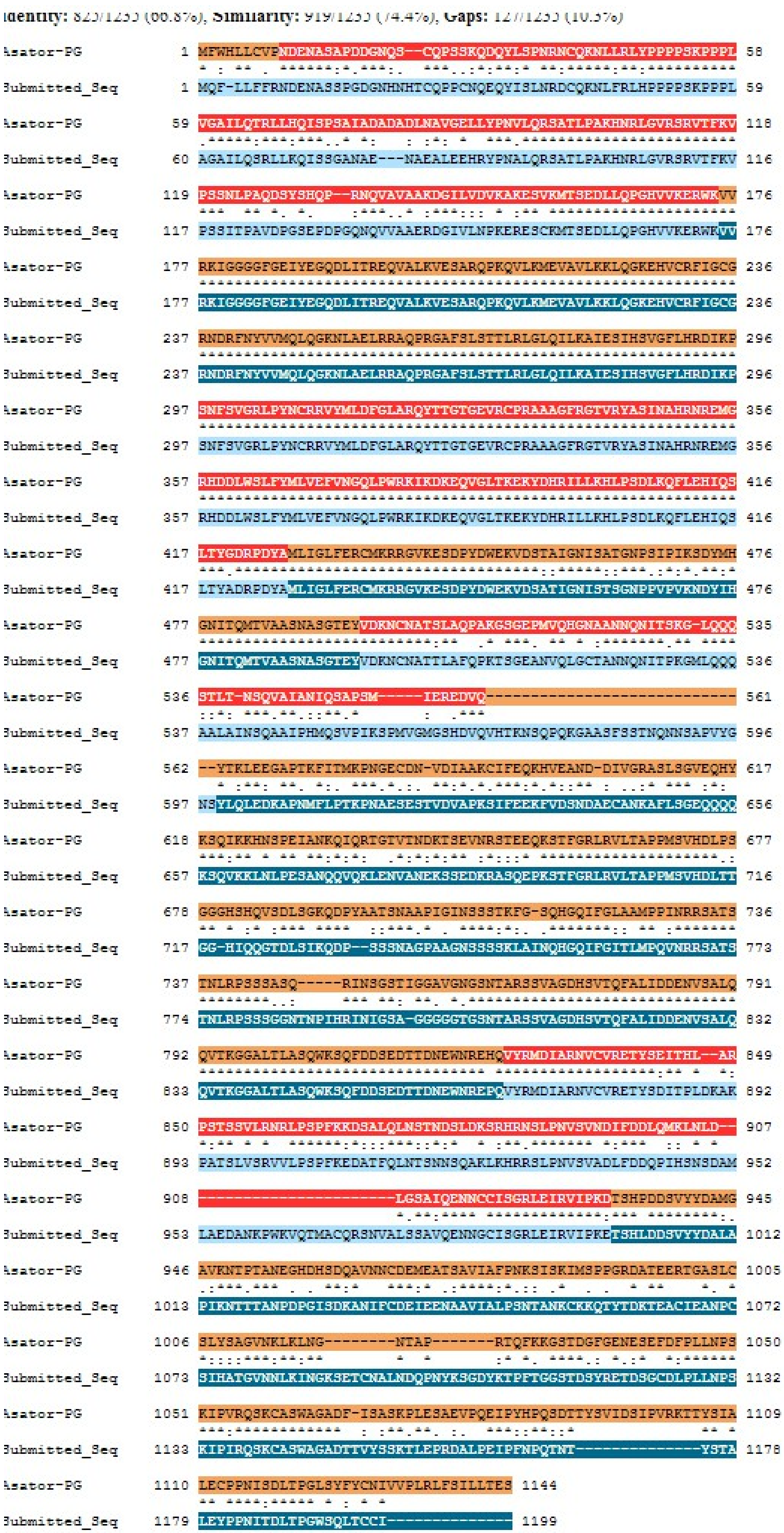
Stop Codon Coordinates: 20289-20287

Project Details

Project Group: D. biarmipes Dot

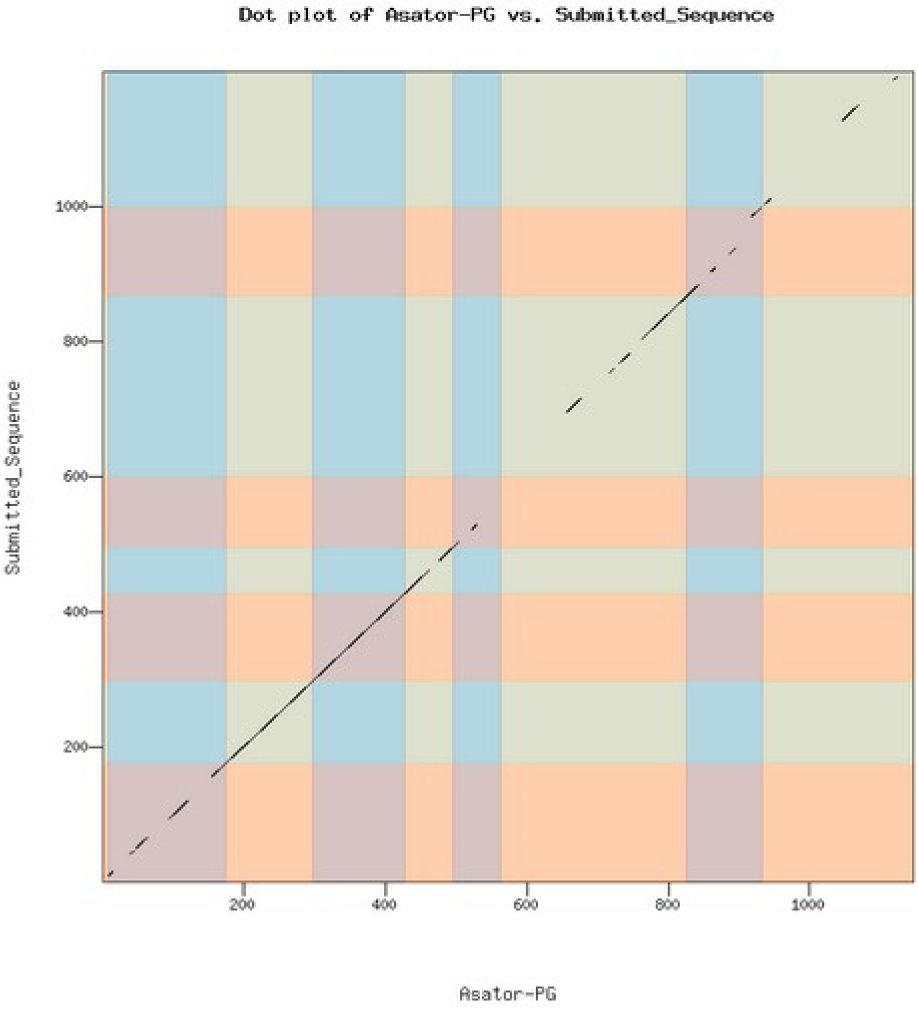
Project Name: contig28

	View	Criteria	Status	Message
⊞	🔍	Check for Start Codon	Pass	
⊞		Acceptor for CDS 1	Skip	Already checked for Start Codon
⊞	🔍	Donor for CDS 1	Pass	
⊞	🔍	Acceptor for CDS 2	Pass	
⊞	🔍	Donor for CDS 2	Pass	
⊞	🔍	Acceptor for CDS 3	Pass	
⊞	🔍	Donor for CDS 3	Pass	
⊞	🔍	Acceptor for CDS 4	Pass	
⊞	🔍	Donor for CDS 4	Pass	
⊞	🔍	Acceptor for CDS 5	Pass	
⊞	🔍	Donor for CDS 5	Pass	
⊞	🔍	Acceptor for CDS 6	Pass	
⊞	🔍	Donor for CDS 6	Pass	
⊞	🔍	Acceptor for CDS 7	Pass	
⊞	🔍	Donor for CDS 7	Pass	
⊞	🔍	Acceptor for CDS 8	Pass	
⊞		Donor for CDS 8	Skip	Already checked for Stop Codon
⊞	🔍	Check for Stop Codon	Pass	
⊞	🔍	Additional Checks	Pass	
⊞	🔍	Number of coding exons matched D. melanogaster ortholog	⚠ Warn	Gene model has 8 CDS's, ortholog has 9 CDS's



```
dentity: 825/1235 (66.8%), Similarity: 919/1235 (74.4%), Gaps: 127/1235 (10.3%)

Asator-PG         1 MFWHLLCVPNDENASAPDDGNQS--CQPSSKQDQYLSPNRNCQKNLLRLYPPPPSKPPPL 58
                    * : ** . ******:*.***:.  ***..:*:**:* **:*****:**:**********
Submitted_Seq     1 MQF-LLFFRNDENASSPGDGNHNHTCQPPCNQEQYISLNRDCQKNLFRLHPPPPSKPPPL 59

Asator-PG        59 VGAILQTRLLHQISPSAIADADADLNAVGELLYPNVLQRSATLPAKHNRLGVRSRVTFKV 118
                    .*****:***:***..* *:   : :*: *  ***.************************
Submitted_Seq    60 AGAILQSRLLKQISSGANAE---NAEALEEHRYPNALQRSATLPAKHNRLGVRSRVTFKV 116

Asator-PG       119 PSSNLPAQDSYSHQP--RNQVAVAAKDGILVDVKAKESVKMTSEDLLQPGHVVKERWKVV 176
                    ***  ** *. *.    :***..* :***::: * :** *********************
Submitted_Seq   117 PSSITPAVDPGSEPDPGQNQVVAAERDGIVLNPKERESCKMTSEDLLQPGHVVKERWKVV 176

Asator-PG       177 RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCG 236
                    ************************************************************
Submitted_Seq   177 RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCG 236

Asator-PG       237 RNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKP 296
                    ************************************************************
Submitted_Seq   237 RNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKP 296

Asator-PG       297 SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMG 356
                    ************************************************************
Submitted_Seq   297 SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMG 356

Asator-PG       357 RHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS 416
                    ************************************************************
Submitted_Seq   357 RHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS 416

Asator-PG       417 LTYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAIGNISATGNPSIPIKSDYMH 476
                    ***.**********************************::*****::***.:*:*.**:*
Submitted_Seq   417 LTYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSATIGNISTSGNPPVPVKNDYIH 476

Asator-PG       477 GNITQMTVAASNASGTEYVDKNCNATSLAQPAKGSGEPMVQHGNAANNQNITSKG-LQQQ 535
                    **************************:**  .* ***. ** * :*******.** ****
Submitted_Seq   477 GNITQMTVAASNASGTEYVDKNCNATTLAFQPKTSGEANVQLGCTANNQNITPKGMLQQQ 536

Asator-PG       536 STLT-NSQVAIANIQSAPSM-----IEREDVQ---------------------------- 561
                    ::*: ***.**.::**.*       :  .***
Submitted_Seq   537 AALAINSQAAIPHMQSVPIKSPMVGMGSHDVQVHTKNSQPQKGAASFSSTNQNNSAPVYG 596

Asator-PG       562 --YTKLEEGAPTKFITMKPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHY 617
                      * :**: **. *:. ***.*.:. **:*.*.***:*.*::** : ..:* *** :*:
Submitted_Seq   597 NSYLQLEDKAPNMFLPTKPNAESESTVDVAPKSIFEEKFVDSNDAECANKAFLSGEQQQQ 656

Asator-PG       618 KSQIKKHNSPEIANKQIQRTGTVTNDKTSEVNRSTEEQKSTFGRLRVLTAPPMSVHDLPS 677
                    ***:** * ** **:*:*:  .*:*:*:** :*:::* ********************.:
Submitted_Seq   657 KSQVKKLNLPESANQQVQKLENVANEKSSEDKRASQEPKSTFGRLRVLTAPPMSVHDLTT 716

Asator-PG       678 GGGHSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG-SQHGQIFGLAAMPPINRRSATS 736
                    ** * :* :*** ****  ::***.* . ****:*:. .*******:: ** :*******
Submitted_Seq   717 GG-HIQQGTDLSIKQDP--SSSNAGPAAGNSSSSKLAINQHGQIFGITLMPQVNRRSATS 773

Asator-PG       737 TNLRPSSSASQ-----RINSGSTIGGAVGNGSNTARSSVAGDHSVTQFALIDDENVSALQ 791
                    ********..:     *** **: **. *.******************************
Submitted_Seq   774 TNLRPSSSGGNTNPIHRINIGSA-GGGGGTGSNTARSSVAGDHSVTQFALIDDENVSALQ 832

Asator-PG       792 QVTKGGALTLASQWKSQFDDSEDTTDNEWNREHQVYRMDIARNVCVRETYSEITHL--AR 849
                    ******************************** ******************:** *  *:
Submitted_Seq   833 QVTKGGALTLASQWKSQFDDSEDTTDNEWNREPQVYRMDIARNVCVRETYSDITPLDKAK 892

Asator-PG       850 PSTSSVLRNRLPSPFKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDLQMKLNLD-- 907
                    *:** * *  ******:*:::***::*:*  * :**.******* *:***  :: * *
Submitted_Seq   893 PATSLVSRVVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHSNSDAM 952

Asator-PG       908 ----------------------LGSAIQENNCCISGRLEIRVIPKDTSHPDDSVYYDAMG 945
                                          *.**:**** *************:*** ********:.
Submitted_Seq   953 LAEDANKPWKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKETSHLDDSVYYDALA 1012

Asator-PG       946 AVKNTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKIMSPPGRDATEERTGASLC 1005
                    .:***.*** . . **:*   ***:* .:****:*.:: .* .  .  * **    *. *
Submitted_Seq  1013 PIKNTTTANPDPGISDKANIFCDEIEENAAVIALPSNTANKCKKQTYTDKTEACIEANPC 1072

Asator-PG      1006 SLYSAGVNKLKLNG--------NTAP-------RTQFKKGSTDGFGENESEFDFPLLNPS 1050
                    *::::***:**:**        *  *       :* *. ****.: *.:*  *:******
Submitted_Seq  1073 SIHATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGSTDSYRETDSGCDLPLLNPS 1132

Asator-PG      1051 KIPVRQSKCASWAGADF-ISASKPLESAEVPQEIPYHPQSDTTYSVIDSIPVRKTTYSIA 1109
                    ***:************  : :**.**. :.  ***::**::*              ** *
Submitted_Seq  1133 KIPIRQSKCASWAGADTTVYSSKTLEPRDALPEIPFNPQTNT--------------YSTA 1178

Asator-PG      1110 LECPPNISDLTPGLSYFYCNIVVPLRLFSILLTES 1144
                    ** ****:***** * : * *
Submitted_Seq  1179 LEYPPNITDLTPGWSQLTCCI-------------- 1199
```

Dotplot


Dot plot of Asator-PG vs. Submitted_Sequence
Submitted_Sequence
1000
800
600
400
200
200
400
600
800
1000
Asator-PG

Isoform Report Form

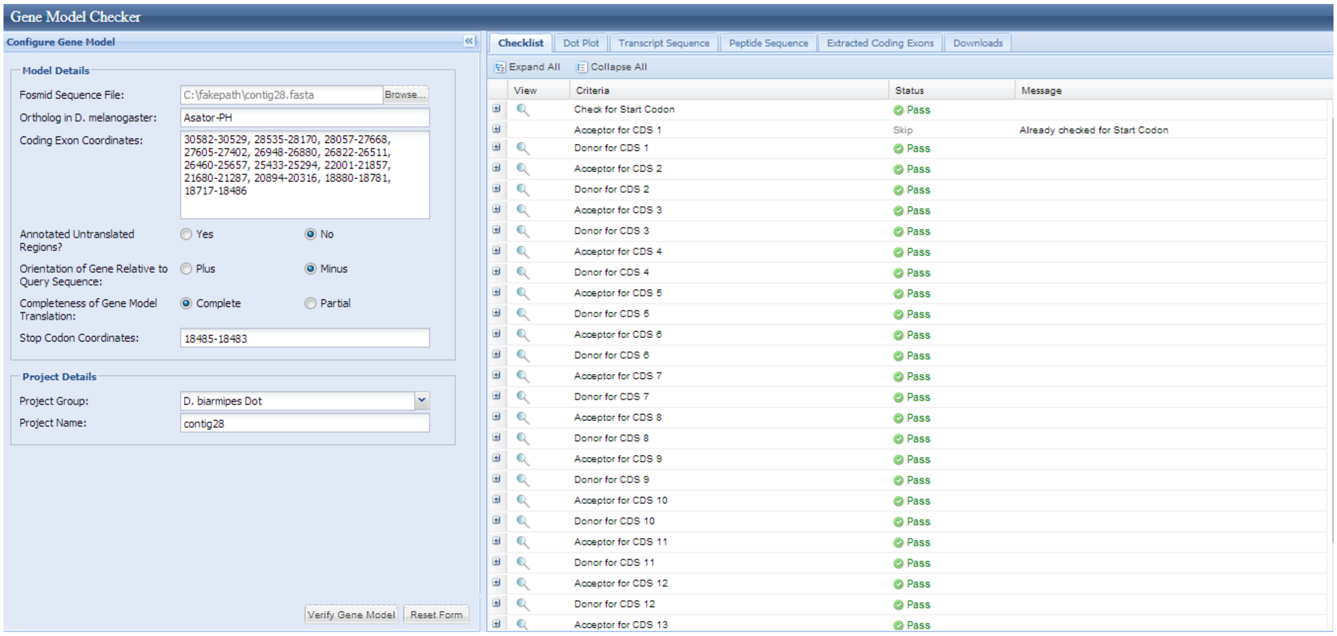
Gene--- isoform name: *Asator-PH*

Names of isoforms with identical coding sequences: none

Is the 5' end of this isoform missing from the end of the project: no

Is the 5' end of this isoform missing from the end of the project: no

Checklist



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PH

Coding Exon Coordinates: 30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

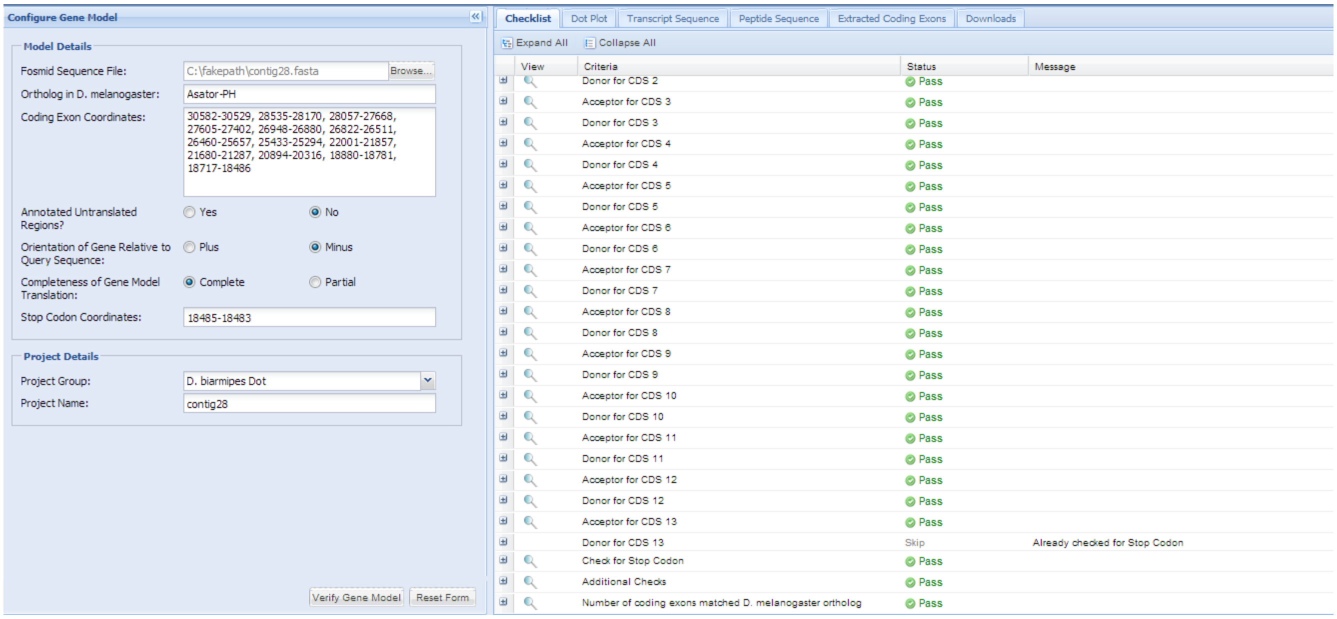
Stop Codon Coordinates: 18485-18483

Project Details

Project Group: D. biarmipes Dot

Project Name: contig28

View	Criteria	Status	Message
🔍	Check for Start Codon	✓ Pass	
	Acceptor for CDS 1	Skip	Already checked for Start Codon
🔍	Donor for CDS 1	✓ Pass	
🔍	Acceptor for CDS 2	✓ Pass	
🔍	Donor for CDS 2	✓ Pass	
🔍	Acceptor for CDS 3	✓ Pass	
🔍	Donor for CDS 3	✓ Pass	
🔍	Acceptor for CDS 4	✓ Pass	
🔍	Donor for CDS 4	✓ Pass	
🔍	Acceptor for CDS 5	✓ Pass	
🔍	Donor for CDS 5	✓ Pass	
🔍	Acceptor for CDS 6	✓ Pass	
🔍	Donor for CDS 6	✓ Pass	
🔍	Acceptor for CDS 7	✓ Pass	
🔍	Donor for CDS 7	✓ Pass	
🔍	Acceptor for CDS 8	✓ Pass	
🔍	Donor for CDS 8	✓ Pass	
🔍	Acceptor for CDS 9	✓ Pass	
🔍	Donor for CDS 9	✓ Pass	
🔍	Acceptor for CDS 10	✓ Pass	
🔍	Donor for CDS 10	✓ Pass	
🔍	Acceptor for CDS 11	✓ Pass	
🔍	Donor for CDS 11	✓ Pass	
🔍	Acceptor for CDS 12	✓ Pass	
🔍	Donor for CDS 12	✓ Pass	
🔍	Acceptor for CDS 13	✓ Pass	



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PH

Coding Exon Coordinates: 30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

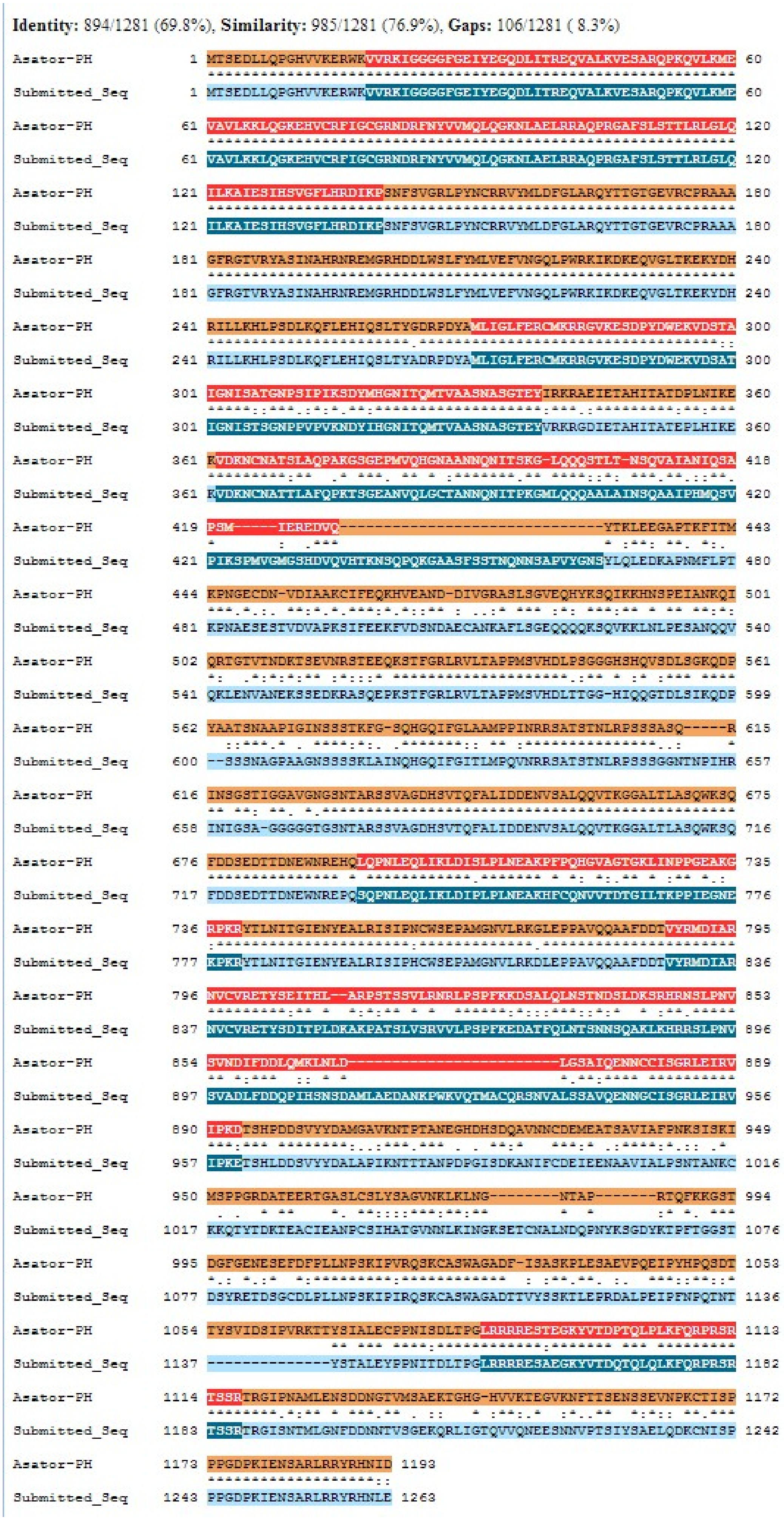
Stop Codon Coordinates: 18485-18483

Project Details

Project Group: D. biarmipes Dot

Project Name: contig28

View	Criteria	Status	Message
🔍	Donor for CDS 2	✓ Pass	
🔍	Acceptor for CDS 3	✓ Pass	
🔍	Donor for CDS 3	✓ Pass	
🔍	Acceptor for CDS 4	✓ Pass	
🔍	Donor for CDS 4	✓ Pass	
🔍	Acceptor for CDS 5	✓ Pass	
🔍	Donor for CDS 5	✓ Pass	
🔍	Acceptor for CDS 6	✓ Pass	
🔍	Donor for CDS 6	✓ Pass	
🔍	Acceptor for CDS 7	✓ Pass	
🔍	Donor for CDS 7	✓ Pass	
🔍	Acceptor for CDS 8	✓ Pass	
🔍	Donor for CDS 8	✓ Pass	
🔍	Acceptor for CDS 9	✓ Pass	
🔍	Donor for CDS 9	✓ Pass	
🔍	Acceptor for CDS 10	✓ Pass	
🔍	Donor for CDS 10	✓ Pass	
🔍	Acceptor for CDS 11	✓ Pass	
🔍	Donor for CDS 11	✓ Pass	
🔍	Acceptor for CDS 12	✓ Pass	
🔍	Donor for CDS 12	✓ Pass	
🔍	Acceptor for CDS 13	✓ Pass	
	Donor for CDS 13	Skip	Already checked for Stop Codon
🔍	Check for Stop Codon	✓ Pass	
🔍	Additional Checks	✓ Pass	
🔍	Number of coding exons matched D. melanogaster ortholog	✓ Pass	



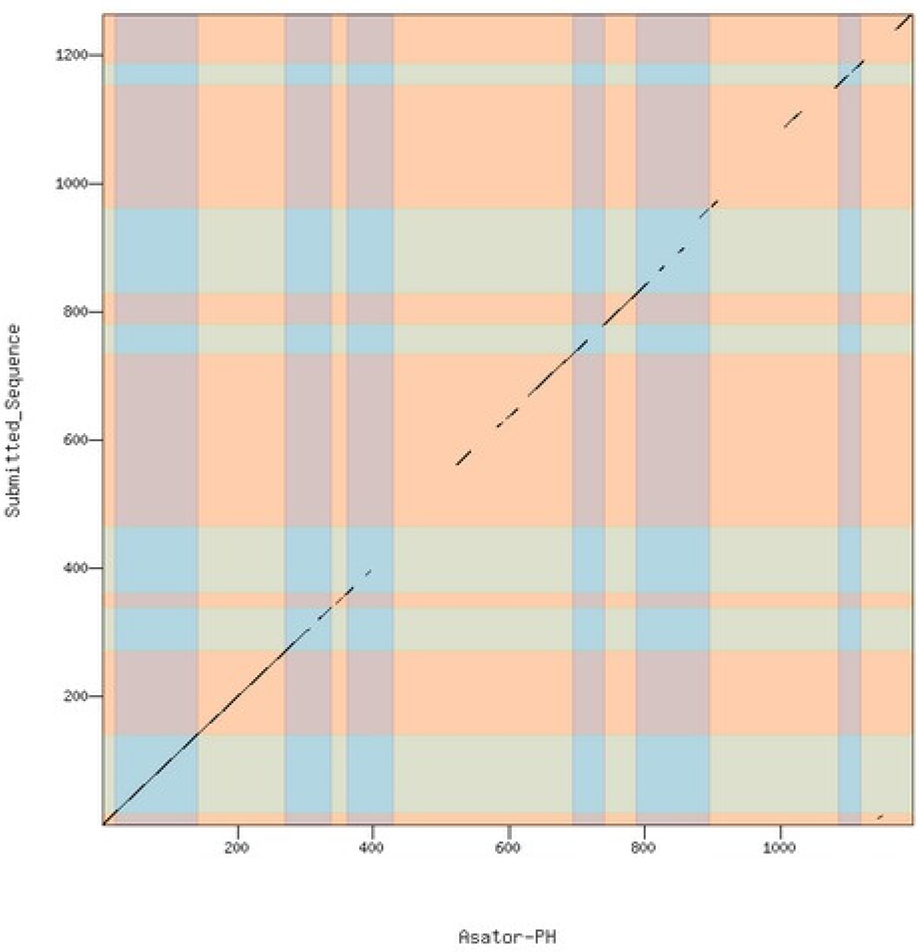
Identity: 894/1281 (69.8%), **Similarity:** 985/1281 (76.9%), **Gaps:** 106/1281 (8.3%)

```
Asator-PH               1 MTSEDLLQPGHVVKERWKVVRKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKME 60
                          ************************************************************
Submitted_Seq           1 MTSEDLLQPGHVVKERWKVVRKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKME 60

Asator-PH              61 VAVLKKLQGKEHVCRFIGCGRNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQ 120
                          ************************************************************
Submitted_Seq          61 VAVLKKLQGKEHVCRFIGCGRNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQ 120

Asator-PH             121 ILKAIESIHSVGFLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA 180
                          ************************************************************
Submitted_Seq         121 ILKAIESIHSVGFLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA 180

Asator-PH             181 GFRGTVRYASINAHRNREMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH 240
                          ************************************************************
Submitted_Seq         181 GFRGTVRYASINAHRNREMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH 240

Asator-PH             241 RILLKHLPSDLKQFLEHIQSLTYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTA 300
                          ***********************.**********************************::
Submitted_Seq         241 RILLKHLPSDLKQFLEHIQSLTYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSAT 300

Asator-PH             301 IGNISATGNPSIPIKSDYMHGNITQMTVAASNASGTEYIRKRAEIETAHITATDPLNIKE 360
                          *****::***.:*:*.**:*******************:***.:*********:**:***
Submitted_Seq         301 IGNISTSGNPPVPVKNDYIHGNITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKE 360

Asator-PH             361 KVDKNCNATSLAQPAKGSGEPMVQHGNAANNQNITSKG-LQQQSTLT-NSQVAIANIQSA 418
                          *********:**  .* ***. ** * :*******.** ****::*: ***.**.::**.
Submitted_Seq         361 KVDKNCNATTLAFQPKTSGEANVQLGCTANNQNITPKGMLQQQAALAINSQAAIPHMQSV 420

Asator-PH             419 PSM-----IEREDVQ------------------------------YTKLEEGAPTKFITM 443
                          *       :  .***                              * :**: **. *:.
Submitted_Seq         421 PIKSPMVGMGSHDVQVHTKNSQPQKGAASFSSTNQNNSAPVYGNSYLQLEDKAPNMFLPT 480

Asator-PH             444 KPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQI 501
                          ***.*.:. **:*.*.***:*.*::** : ..:* *** :*: ***:** * ** **:*:
Submitted_Seq         481 KPNAESESTVDVAPKSIFEEKFVDSNDAECANKAFLSGEQQQQKSQVKKLNLPESANQQV 540

Asator-PH             502 QRTGTVTNDKTSEVNRSTEEQKSTFGRLRVLTAPPMSVHDLPSGGGHSHQVSDLSGKQDP 561
                          *:  .*:*:*:** :*:::* ********************.:** * :* :*** ****
Submitted_Seq         541 QKLENVANEKSSEDKRASQEPKSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSIKQDP 599

Asator-PH             562 YAATSNAAPIGINSSSTKFG-SQHGQIFGLAAMPPINRRSATSTNLRPSSSASQ-----R 615
                            ::***.* . ****:*:. .*******:: ** :***************..:     *
Submitted_Seq         600 --SSSNAGPAAGNSSSSKLAINQHGQIFGITLMPQVNRRSATSTNLRPSSSGGNTNPIHR 657

Asator-PH             616 INSGSTIGGAVGNGSNTARSSVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQ 675
                          ** **: **. *.***********************************************
Submitted_Seq         658 INIGSA-GGGGGTGSNTARSSVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQ 716

Asator-PH             676 FDDSEDTTDNEWNREHQLQPNLEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEAKG 735
                          *************** * ************.******* * *: *:.** * :** *.:
Submitted_Seq         717 FDDSEDTTDNEWNREPQSQPNLEQLIKLDIPLPLNEAKHFCQNVVTDTGILTKPPIEGNE 776

Asator-PH             736 RPKRYTLNITGIENYEALRISIPNCWSEPAMGNVLRKGLEPPAVQQAAFDDTVYRMDIAR 795
                          :**********************:*************.**********************
Submitted_Seq         777 KPKRYTLNITGIENYEALRISIPHCWSEPAMGNVLRKDLEPPAVQQAAFDDTVYRMDIAR 836

Asator-PH             796 NVCVRETYSEITHL--ARPSTSSVLRNRLPSPFKKDSALQLNSTNDSLDKSRHRNSLPNV 853
                          *********:** *  *:*:** * *  ******:*:::***::*:*  * :**.*****
Submitted_Seq         837 NVCVRETYSDITPLDKAKPATSLVSRVVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNV 896

Asator-PH             854 SVNDIFDDLQMKLNLD------------------------LGSAIQENNCCISGRLEIRV 889
                          ** *:***  :: * *                        *.**:**** **********
Submitted_Seq         897 SVADLFDDQPIHSNSDAMLAEDANKPWKVQTMACQRSNVALSSAVQENNGCISGRLEIRV 956

Asator-PH             890 IPKDTSHPDDSVYYDAMGAVKNTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKI 949
                          ***:*** ********:..:***.*** . . **:*   ***:* .:****:*.:: .*
Submitted_Seq         957 IPKETSHLDDSVYYDALAPIKNTTTANPDPGISDKANIFCDEIEENAAVIALPSNTANKC 1016

Asator-PH             950 MSPPGRDATEERTGASLCSLYSAGVNKLKLNG--------NTAP-------RTQFKKGST 994
                           . .  * **    *. **::::***:**:**        *  *       :* *. ***
Submitted_Seq        1017 KKQTYTDKTEACIEANPCSIHATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGST 1076

Asator-PH             995 DGFGENESEFDFPLLNPSKIPVRQSKCASWAGADF-ISASKPLESAEVPQEIPYHPQSDT 1053
                          *.: *.:*  *:*********:************  : :**.**. :.  ***::**::*
Submitted_Seq        1077 DSYRETDSGCDLPLLNPSKIPIRQSKCASWAGADTTVYSSKTLEPRDALPEIPFNPQTNT 1136

Asator-PH            1054 TYSVIDSIPVRKTTYSIALECPPNISDLTPGLRRRRESTEGKYVTDPTQLPLKFQRPRSR 1113
                                        ** *** ****:************:******* *** *********
Submitted_Seq        1137 --------------YSTALEYPPNITDLTPGLRRRRESAEGKYVTDQTQLQLKFQRPRSR 1182

Asator-PH            1114 TSSRTRGIPNAMLENSDDNGTVMSAEKTGHG-HVVKTEGVKNFTTSENSSEVNPKCTISP 1172
                          ********.*:** * ***.** . ::   * :**:.*  :*..**  *:*:: **.***
Submitted_Seq        1183 TSSRTRGISNTMLGNFDDNNTVSGEKQRLIGTQVVQNEESNNVPTSIYSAELQDKCNISP 1242

Asator-PH            1173 PPGDPKIENSARLRRYRHNID 1193
                          *******************::
Submitted_Seq        1243 PPGDPKIENSARLRRYRHNLE 1263
```

Dotplot

View protein alignment

Dot plot of Asator-PH vs. Submitted_Sequence



Isoform Report Form

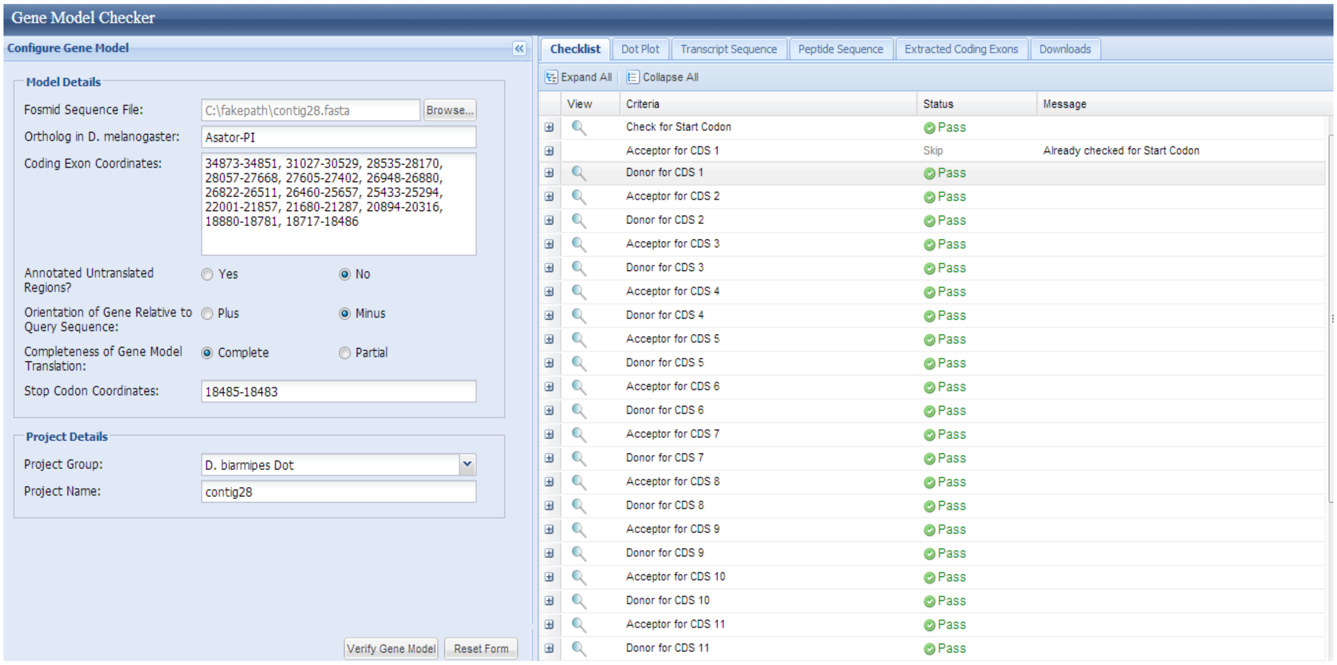
Gene--- isoform name: *Asator-PI*

Names of isoforms with identical coding sequences: none

Is the 5' end of this isoform missing from the end of the project: no

Is the 5' end of this isoform missing from the end of the project: no

Checklist



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PI

Coding Exon Coordinates: 34873-34851, 31027-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

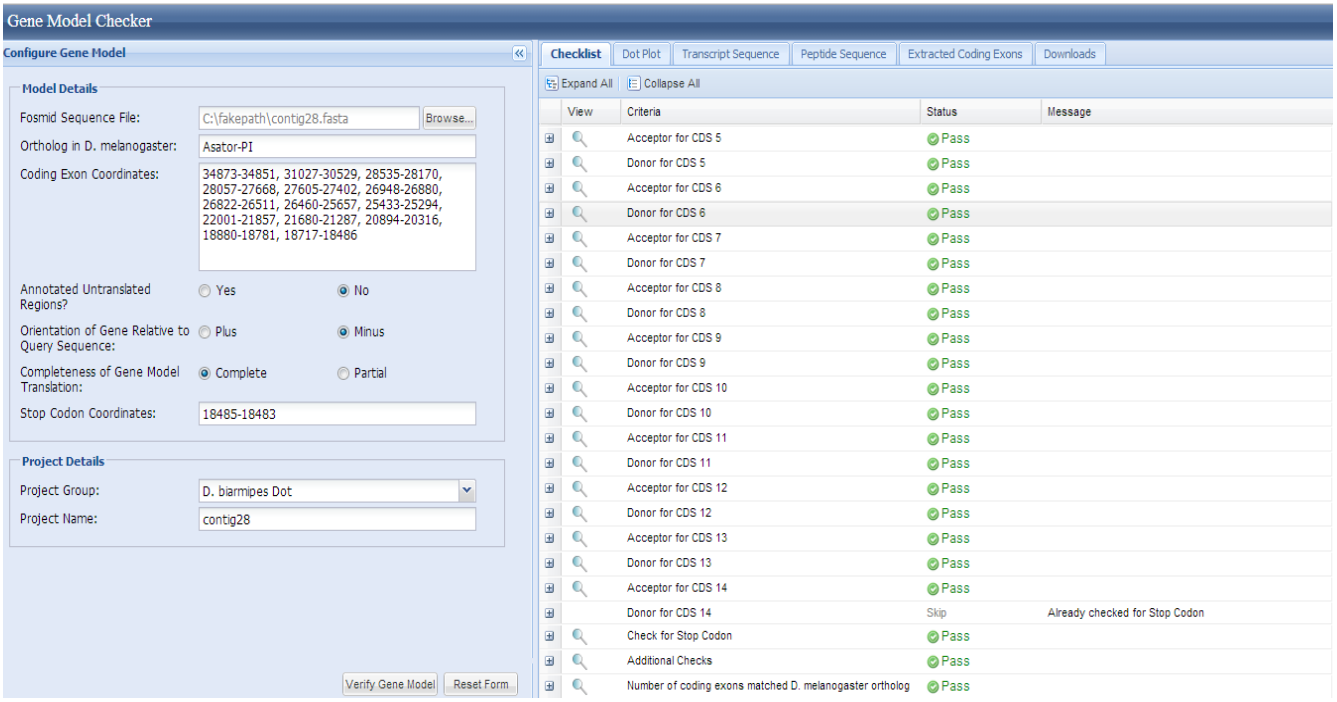
Stop Codon Coordinates: 18485-18483

Project Details

Project Group: D. biarmipes Dot

Project Name: contig28

View	Criteria	Status	Message
	Check for Start Codon	Pass	
	Acceptor for CDS 1	Skip	Already checked for Start Codon
	Donor for CDS 1	Pass	
	Acceptor for CDS 2	Pass	
	Donor for CDS 2	Pass	
	Acceptor for CDS 3	Pass	
	Donor for CDS 3	Pass	
	Acceptor for CDS 4	Pass	
	Donor for CDS 4	Pass	
	Acceptor for CDS 5	Pass	
	Donor for CDS 5	Pass	
	Acceptor for CDS 6	Pass	
	Donor for CDS 6	Pass	
	Acceptor for CDS 7	Pass	
	Donor for CDS 7	Pass	
	Acceptor for CDS 8	Pass	
	Donor for CDS 8	Pass	
	Acceptor for CDS 9	Pass	
	Donor for CDS 9	Pass	
	Acceptor for CDS 10	Pass	
	Donor for CDS 10	Pass	
	Acceptor for CDS 11	Pass	
	Donor for CDS 11	Pass	



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PI

Coding Exon Coordinates: 34873-34851, 31027-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20316, 18880-18781, 18717-18486

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

Stop Codon Coordinates: 18485-18483

Project Details

Project Group: D. biarmipes Dot

Project Name: contig28

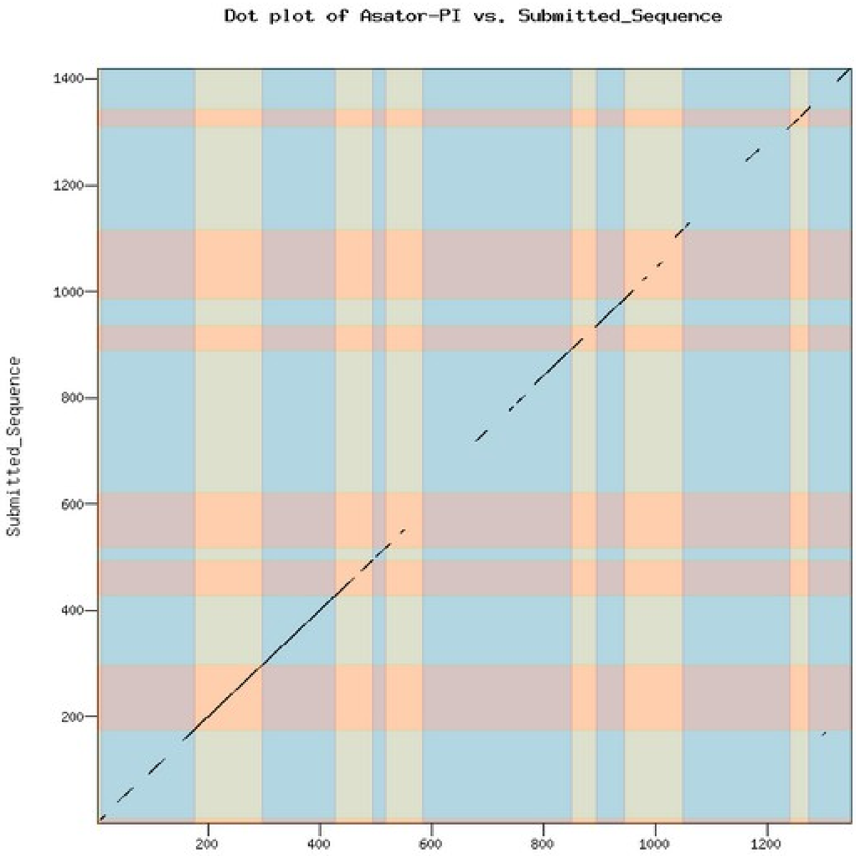
View	Criteria	Status	Message
	Acceptor for CDS 5	Pass	
	Donor for CDS 5	Pass	
	Acceptor for CDS 6	Pass	
	Donor for CDS 6	Pass	
	Acceptor for CDS 7	Pass	
	Donor for CDS 7	Pass	
	Acceptor for CDS 8	Pass	
	Donor for CDS 8	Pass	
	Acceptor for CDS 9	Pass	
	Donor for CDS 9	Pass	
	Acceptor for CDS 10	Pass	
	Donor for CDS 10	Pass	
	Acceptor for CDS 11	Pass	
	Donor for CDS 11	Pass	
	Acceptor for CDS 12	Pass	
	Donor for CDS 12	Pass	
	Acceptor for CDS 13	Pass	
	Donor for CDS 13	Pass	
	Acceptor for CDS 14	Pass	
	Donor for CDS 14	Skip	Already checked for Stop Codon
	Check for Stop Codon	Pass	
	Additional Checks	Pass	
	Number of coding exons matched D. melanogaster ortholog	Pass	

Identity: 997/1440 (69.2%), **Similarity:** 1105/1440 (76.7%), **Gaps:** 113/1440 (7.8%)

```
Asator-PI           1 MHLQEYLRNDENASAPDDGNQS--CQPSSKQDQYLSPNRNCQKNLLRLYPPPPSKPPPLV 58
                      ****:*********:*.***:.  ***..:*:**:* **:*****:**:**********.
Submitted_Seq       1 MHLQDYLRNDENASSPGDGNHNHTCQPPCNQEQYISLNRDCQKNLFRLHPPPPSKPPPLA 60

Asator-PI          59 GAILQTRLLHQISPSAIADADADLNAVGELLYPNVLQRSATLPAKHNRLGVRSRVTFKVP 118
                      *****:***:***..* *:   : :*: *  ***.*************************
Submitted_Seq      61 GAILQSRLLKQISSGANAE---NAEALEEHRYPNALQRSATLPAKHNRLGVRSRVTFKVP 117

Asator-PI         119 SSNLPAQDSYSHQP--RNQVAVAAKDGILVDVKAKESVKMTSEDLLQPGHVVKERWKVVR 176
                      **  ** *. *.    :***..* :***::: * :** :*********************
Submitted_Seq     118 SSITPAVDPGSEPDPGQNQVVAAERDGIVLNPKERESCRMTSEDLLQPGHVVKERWKVVR 177

Asator-PI         177 KIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCGR 236
                      ************************************************************
Submitted_Seq     178 KIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCGR 237

Asator-PI         237 NDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKPS 296
                      ************************************************************
Submitted_Seq     238 NDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKPS 297

Asator-PI         297 NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMGR 356
                      ************************************************************
Submitted_Seq     298 NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMGR 357

Asator-PI         357 HDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQSL 416
                      ************************************************************
Submitted_Seq     358 HDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQSL 417

Asator-PI         417 TYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAIGNISATGNPSIPIKSDYMHG 476
                      **.**********************************::*****::***.:*:*.**:**
Submitted_Seq     418 TYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSATIGNISTSGNPPVPVKNDYIHG 477

Asator-PI         477 NITQMTVAASNASGTEYIRKRAEIETAHITATDPLNIKEKVDKNCNATSLAQPAKGSGEP 536
                      *****************:***.:*********:**:************:**  .* ***.
Submitted_Seq     478 NITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKEKVDKNCNATTLAFQPKTSGEA 537

Asator-PI         537 MVQHGNAANNQNITSKG-LQQQSTLT-NSQVAIANIQSAPSM-----IEREDVQ------ 583
                       ** * :*******.** ****::*: ***.**.::**.*       :  .***
Submitted_Seq     538 NVQLGCTANNQNITPKGMLQQQAALAINSQAAIPHMQSVPIKSPMVGMGSHDVQVHTKNS 597

Asator-PI         584 ------------------------YTKLEEGAPTKFITMKPNGECDN-VDIAAKCIFEQK 618
                                              * :**: **. *:. ***.*.:. **:*.*.***:*
Submitted_Seq     598 QPQKGAASFSSTNQNNSAPVYGNSYLQLEDKAPNMFLPTKPNAESESTVDVAPKSIFEEK 657

Asator-PI         619 HVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQIQRTGTVTNDKTSEVNRSTEEQ 677
                      .*::** : ..:* *** :*: ***:** * ** **:*:*:  .*:*:*:** :*:::*
Submitted_Seq     658 FVDSNDAECANKAFLSGEQQQQKSQVKKLNLPESANQQVQKLENVANEKSSEDKRASQEP 717

Asator-PI         678 KSTFGRLRVLTAPPMSVHDLPSGGGHSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG- 736
                      ********************.:** * :* :*** ****  ::***.* . ****:*:.
Submitted_Seq     718 KSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSIKQDP--SSSNAGPAAGNSSSSKLAI 774

Asator-PI         737 SQHGQIFGLAAMPPINRRSATSTNLRPSSSASQ-----RINSGSTIGGAVGNGSNTARSS 791
                      .*******:: ** :***************..:     *** **: **. *.********
Submitted_Seq     775 NQHGQIFGITLMPQVNRRSATSTNLRPSSSGGNTNPIHRINIGSA-GGGGGTGSNTARSS 833

Asator-PI         792 VAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREHQLQPN 851
                      ****************************************************** * ***
Submitted_Seq     834 VAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREPQSQPN 893

Asator-PI         852 LEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEAKGRPKRYTLNITGIENYEALRIS 911
                      *********.******* * *: *..** * :** *.: :********************
Submitted_Seq     894 LEQLIKLDIPLPLNEAKHFCQNVVTDTGILTKPPIEGNEKPKRYTLNITGIENYEALRIS 953

Asator-PI         912 IPNCWSEPAMGNVLRKGLEPPAVQQAAFDDTVYRMDIARNVCVRETYSEITHL--ARPST 969
                      **:*************.*******************************:** *  *:.:*
Submitted_Seq     954 IPHCWSEPAMGNVLRKDLEPPAVQQAAFDDTVYRMDIARNVCVRETYSDITPLDKAKPAT 1013

Asator-PI         970 SSVLRNRLPSPFKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDLQMKLNLD----- 1024
                      * * *  ******:*:::***::*:*  * :**.******* *:***  :: * *
Submitted_Seq    1014 SLVSRVVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHSNSDAMLAE 1073

Asator-PI        1025 -------------------LGSAIQENNCCISGRLEIRVIPKDTSHPDDSVYYDAMGAVK 1065
                                         *.**:**** *************:*** ********:..:*
Submitted_Seq    1074 DANKPWKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKETSHLDDSVYYDALAPIK 1133

Asator-PI        1066 NTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKIMSPPGRDATEERTGASLCSLY 1125
                      **.*** . . **:*   ***:* .:****:*.:: .*  . .  * **    *. **::
Submitted_Seq    1134 NTTTANPDPGISDKANIFCDEIEENAAVIALPSNTANKCKKQTYTDKTEACIEANPCSIH 1193

Asator-PI        1126 SAGVNKLKLNG--------NTAP-------RTQFKKGSTDGFGENESEFDFPLLNPSKIP 1170
                      ::***:**:**        *  *       :* *. ****.: *.:*  *:*********
Submitted_Seq    1194 ATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGSTDSYRETDSGCDLPLLNPSKIP 1253

Asator-PI        1171 VRQSKCASWAGADF-ISASKPLESAEVPQEIPYHPQSDTTYSVIDSIPVRKTTYSIALEC 1229
                      :************  : :**.**. :.  ***::**::              *** ***
Submitted_Seq    1254 IRQSKCASWAGADTTVYSSKTLEPRDALPEIPFNPQTN--------------TYSTALEY 1299

Asator-PI        1230 PPNISDLTPGLRRRRESTEGKYVTDPTQLPLKFQRPRSRTSSRTRGIPNAMLENSDDNGT 1289
                      ****:************:******* *** *****************.*:** * ***.*
Submitted_Seq    1300 PPNITDLTPGLRRRRESAEGKYVTDQTQLQLKFQRPRSRTSSRTRGISNTMLGNFDDNNT 1359

Asator-PI        1290 VMSAEKTGHG-HVVKTEGVKNFTTSENSSEVNPKCTISPPPGDPKIENSARLRRYRHNID 1348
                      * . ::   * :**:.*  :*..**  *:*:: **.**********************::
Submitted_Seq    1360 VSGEKQRLIGTQVVQNEESNNVPTSIYSAELQDKCNISPPPGDPKIENSARLRRYRHNLE 1419
```

Dotplot

View protein alignment



Isoform Report Form

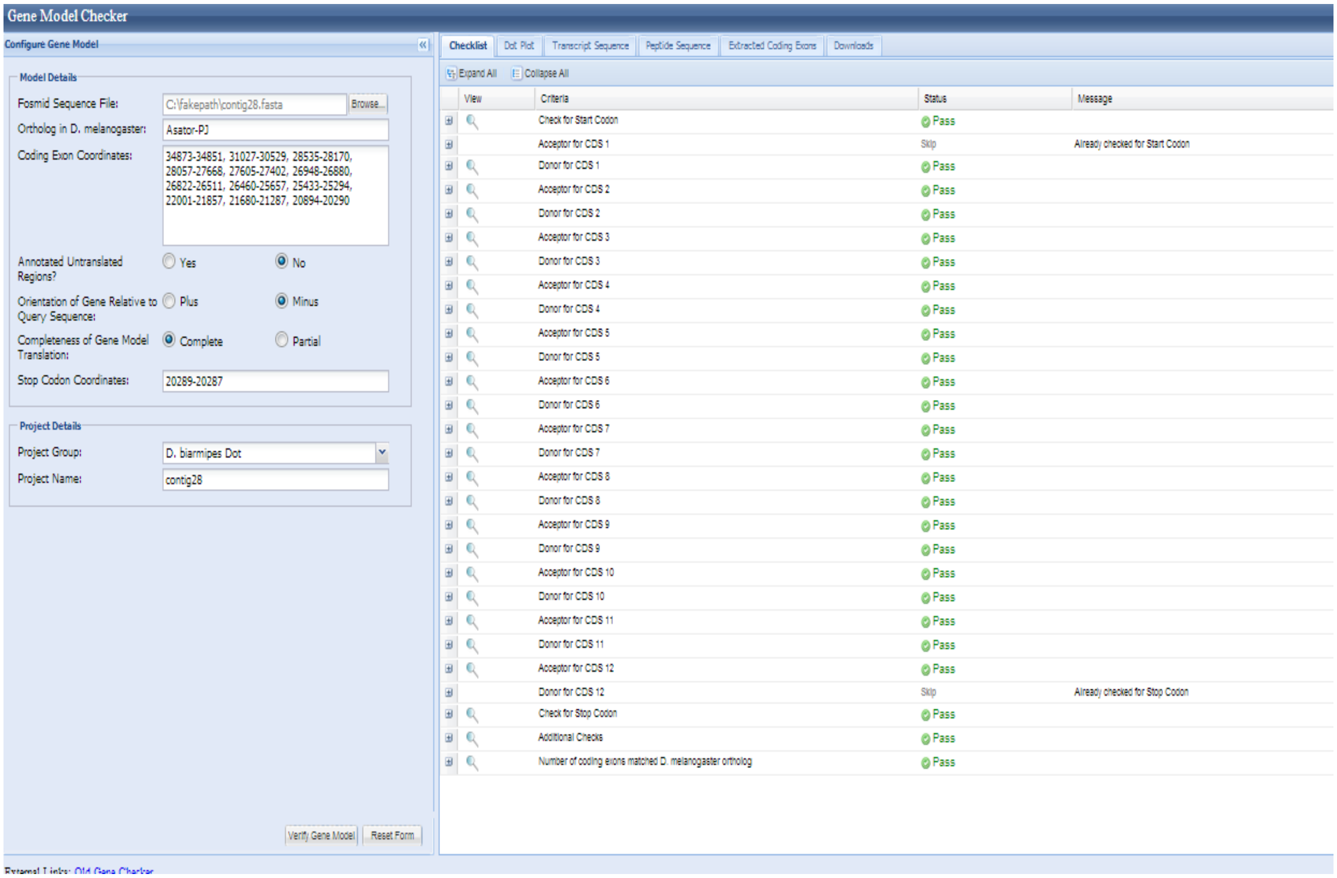
Gene--- isoform name: *Asator-PJ*

Names of isoforms with identical coding sequences: none

Is the 5' end of this isoform missing from the end of the project: no

Is the 5' end of this isoform missing from the end of the project: no

Checklist



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-PJ

Coding Exon Coordinates: 34873-34851, 31027-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20290

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

Stop Codon Coordinates: 20289-20287

Project Details

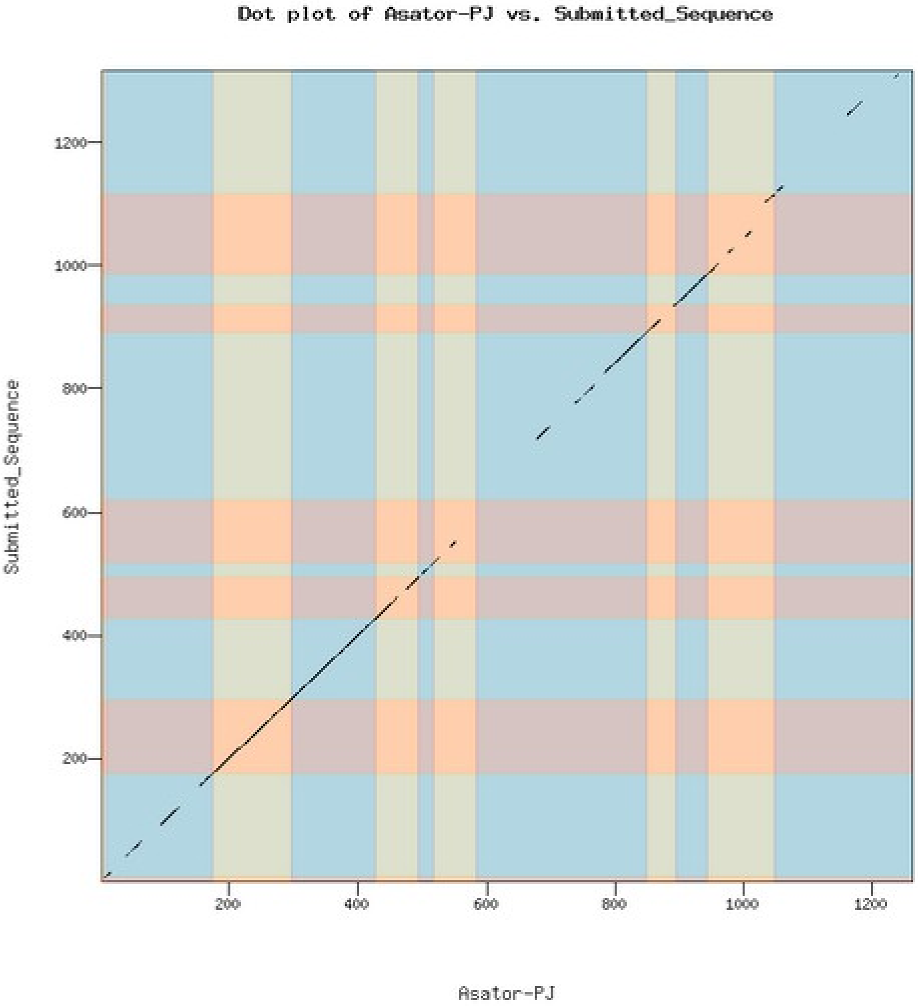
Project Group: D. biarmipes Dot

Project Name: contig28

View	Criteria	Status	Message
🔍	Check for Start Codon	✓ Pass	
	Acceptor for CDS 1	Skip	Already checked for Start Codon
🔍	Donor for CDS 1	✓ Pass	
🔍	Acceptor for CDS 2	✓ Pass	
🔍	Donor for CDS 2	✓ Pass	
🔍	Acceptor for CDS 3	✓ Pass	
🔍	Donor for CDS 3	✓ Pass	
🔍	Acceptor for CDS 4	✓ Pass	
🔍	Donor for CDS 4	✓ Pass	
🔍	Acceptor for CDS 5	✓ Pass	
🔍	Donor for CDS 5	✓ Pass	
🔍	Acceptor for CDS 6	✓ Pass	
🔍	Donor for CDS 6	✓ Pass	
🔍	Acceptor for CDS 7	✓ Pass	
🔍	Donor for CDS 7	✓ Pass	
🔍	Acceptor for CDS 8	✓ Pass	
🔍	Donor for CDS 8	✓ Pass	
🔍	Acceptor for CDS 9	✓ Pass	
🔍	Donor for CDS 9	✓ Pass	
🔍	Acceptor for CDS 10	✓ Pass	
🔍	Donor for CDS 10	✓ Pass	
🔍	Acceptor for CDS 11	✓ Pass	
🔍	Donor for CDS 11	✓ Pass	
🔍	Acceptor for CDS 12	✓ Pass	
	Donor for CDS 12	Skip	Already checked for Stop Codon
🔍	Check for Stop Codon	✓ Pass	
🔍	Additional Checks	✓ Pass	
🔍	Number of coding exons matched D. melanogaster ortholog	✓ Pass	

Identity: 924/1352 (68.3%), **Similarity:** 1025/1352 (75.8%), **Gaps:** 126/1352 (9.3%)

```
Asator-PJ          1 MHLQEYLRNDENASAPDDGNQS--CQPSSKQDQYLSPNRNCQKNLLRLYPPPPSKPPPLV 58
                     ****:*********:*.***:.  ***..:*:**:* **:*****:**:**********.
Submitted_Seq      1 MHLQDYLRNDENASSPGDGNHNHTCQPPCNQEQYISLNRDCQKNLFRLHPPPPSKPPPLA 60

Asator-PJ         59 GAILQTRLLHQISPSAIADADADLNAVGELLYPNVLQRSATLPAKHNRLGVRSRVTFKVP 118
                     *****:***:***..* *:   : :*: *  ***.*************************
Submitted_Seq     61 GAILQSRLLKQISSGANAE---NAEALEEHRYPNALQRSATLPAKHNRLGVRSRVTFKVP 117

Asator-PJ        119 SSNLPAQDSYSHQP--RNQVAVAAKDGILVDVKAKESVKMTSEDLLQPGHVVKERWKVVR 176
                     **  ** *. *.    :***..* :***::: * :** **********************
Submitted_Seq    118 SSITPAVDPGSEPDPGQNQVVAAERDGIVLNPKERESCKMTSEDLLQPGHVVKERWKVVR 177

Asator-PJ        177 KIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCGR 236
                     ************************************************************
Submitted_Seq    178 KIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQGKEHVCRFIGCGR 237

Asator-PJ        237 NDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKPS 296
                     ************************************************************
Submitted_Seq    238 NDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFLHRDIKPS 297

Asator-PJ        297 NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMGR 356
                     ************************************************************
Submitted_Seq    298 NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHRNREMGR 357

Asator-PJ        357 HDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQSL 416
                     ************************************************************
Submitted_Seq    358 HDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDHRILLKHLPSDLKQFLEHIQSL 417

Asator-PJ        417 TYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAIGNISATGNPSIPIKSDYMHG 476
                     **.**********************************::*****::***.:*:*.**:**
Submitted_Seq    418 TYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSATIGNISTSGNPPVPVKNDYIHG 477

Asator-PJ        477 NITQMTVAASNASGTEYIRKRAEIETAHITATDPLNIKEKVDKNCNATSLAQPAKGSGEP 536
                     *****************:***.:*********:**:************:**  .* ***.
Submitted_Seq    478 NITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKEKVDKNCNATTLAFQPKTSGEA 537

Asator-PJ        537 MVQHGNAANNQNITSKG-LQQQSTLT-NSQVAIANIQSAPSM-----IEREDVQ------ 583
                      ** * :*******.** ****::*: ***.**.::**.*       :  .***
Submitted_Seq    538 NVQLGCTANNQNITPKGMLQQQAALAINSQAAIPHMQSVPIKSPMVGMGSHDVQVHTKNS 597

Asator-PJ        584 ------------------------YTKLEEGAPTKFITMKPNGECDN-VDIAAKCIFEQK 618
                                             * :**: **. *:. ***.*.:. **:*.*.***:*
Submitted_Seq    598 QPQKGAASFSSTNQNNSAPVYGNSYLQLEDKAPNMFLPTKPNAESESTVDVAPKSIFEEK 657

Asator-PJ        619 HVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQIQRTGTVTNDKTSEVNRSTEEQ 677
                     .*::** : ..:* *** :*: ***:** * ** **:*:*:  .*:*:*:** :*:::*
Submitted_Seq    658 FVDSNDAECANKAFLSGEQQQQKSQVKKLNLPESANQQVQKLENVANEKSSEDKRASQEP 717

Asator-PJ        678 KSTFGRLRVLTAPPMSVHDLPSGGGHSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG- 736
                     ********************.:** * :* :*** ****  ::***.* . ****:*:.
Submitted_Seq    718 KSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSIKQDP--SSSNAGPAAGNSSSSKLAI 774

Asator-PJ        737 SQHGQIFGLAAMPPINRRSATSTNLRPSSSASQ-----RINSGSTIGGAVGNGSNTARSS 791
                     .*******:: ** :***************..:     *** **: **. *.********
Submitted_Seq    775 NQHGQIFGITLMPQVNRRSATSTNLRPSSSGGNTNPIHRINIGSA-GGGGGTGSNTARSS 833

Asator-PJ        792 VAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREHQLQPN 851
                     ****************************************************** * ***
Submitted_Seq    834 VAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQFDDSEDTTDNEWNREPQSQPN 893

Asator-PJ        852 LEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEAKGRPKRYTLNITGIENYEALRIS 911
                     *********.******* * *: *:.** * :** *.: :********************
Submitted_Seq    894 LEQLIKLDIPLPLNEAKHFCQNVVTDTGILTKPPIEGNEKPKRYTLNITGIENYEALRIS 953

Asator-PJ        912 IPNCWSEPAMGNVLRKGLEPPAVQQAAFDDTVYRMDIARNVCVRETYSEITHL--ARPST 969
                     **:*************.*******************************:** *  *:*:*
Submitted_Seq    954 IPHCWSEPAMGNVLRKDLEPPAVQQAAFDDTVYRMDIARNVCVRETYSDITPLDKAKPAT 1013

Asator-PJ        970 SSVLRNRLPSPFKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDLQMKLNLD----- 1024
                     * * *  ******:*:::***::*:*  * :**.******* *:***  :: * *
Submitted_Seq   1014 SLVSRVVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHSNSDAMLAE 1073

Asator-PJ       1025 -------------------LGSAIQENNCCISGRLEIRVIPKDTSHPDDSVYYDAMGAVK 1065
                                        *.**:**** *************:*** ********:..:*
Submitted_Seq   1074 DANKPWKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKETSHLDDSVYYDALAPIK 1133

Asator-PJ       1066 NTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKIMSPPGRDATEERTGASLCSLY 1125
                     **.*** . . **:*   ***:* .:****:*.:: .*       * **    *. **::
Submitted_Seq   1134 NTTTANPDPGISDKANIFCDEIEENAAVIALPSNTANKCKKQTYTDKTEACIEANPCSIH 1193

Asator-PJ       1126 SAGVNKLKLNG--------NTAP-------RTQFKKGSTDGFGENESEFDFPLLNPSKIP 1170
                     ::***:**:**        *  *       :* *. ****.: *.:*  *:*********
Submitted_Seq   1194 ATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGSTDSYRETDSGCDLPLLNPSKIP 1253

Asator-PJ       1171 VRQSKCASWAGADF-ISASKPLESAEVPQEIPYHPQSDTTYSVIDSIPVRKTTYSIALEC 1229
                     :************  : :**.**. :.  ***::**::              *** ***
Submitted_Seq   1254 IRQSKCASWAGADTTVYSSKTLEPRDALPEIPFNPQTN--------------TYSTALEY 1299

Asator-PJ       1230 PPNISDLTPGLSYFYCNIVVPLRLFSILLTES 1261
                     ****:***** * : * *
Submitted_Seq   1300 PPNITDLTPGWSQLTCCI-------------- 1317
```


Dot plot of Asator-PJ vs. Submitted_Sequence
Submitted_Sequence
1200
1000
800
600
400
200
200
400
600
800
1000
1200
Asator-PJ

Isoform Report Form

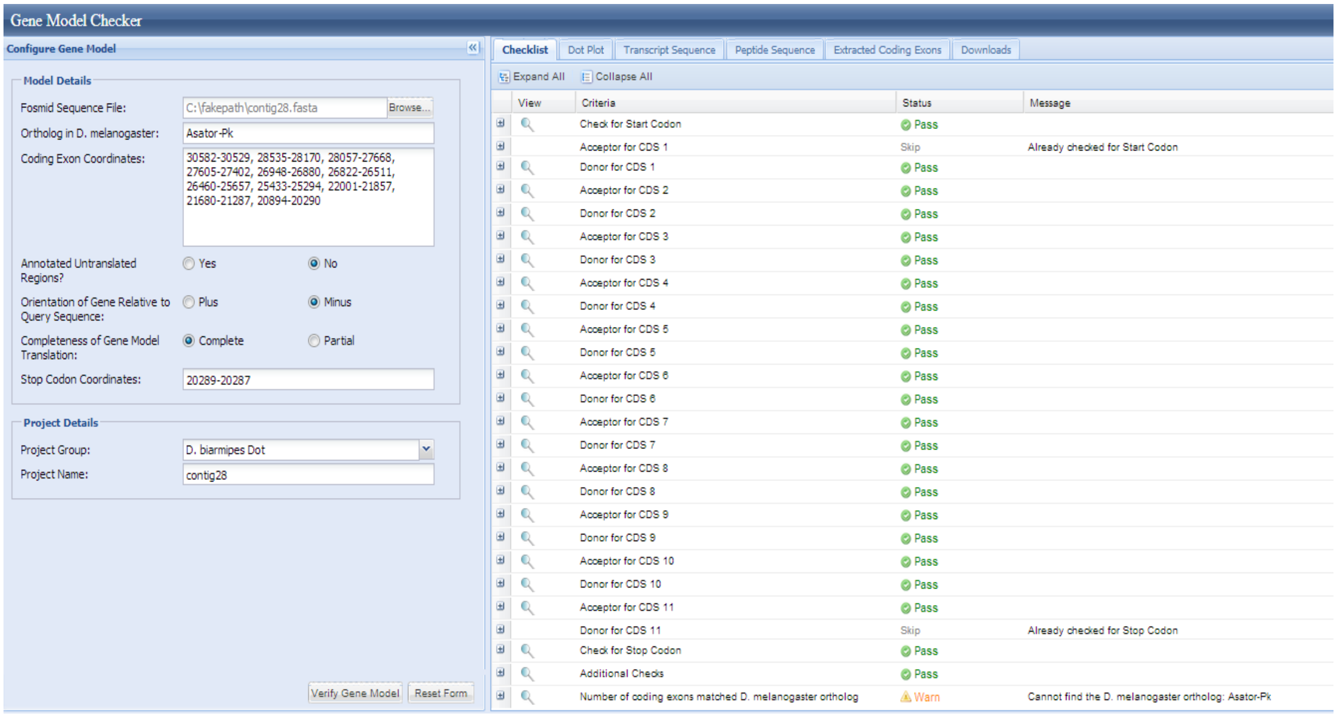
Gene--- isoform name: *Asator-PK*

Names of isoforms with identical coding sequences: none

Is the 5' end of this isoform missing from the end of the project: no

Is the 5' end of this isoform missing from the end of the project: no

Checklist



Model Details

Fosmid Sequence File: C:\fakepath\contig28.fasta

Ortholog in D. melanogaster: Asator-Pk

Coding Exon Coordinates: 30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25294, 22001-21857, 21680-21287, 20894-20290

Annotated Untranslated Regions? ○ Yes ● No

Orientation of Gene Relative to Query Sequence: ○ Plus ● Minus

Completeness of Gene Model Translation: ● Complete ○ Partial

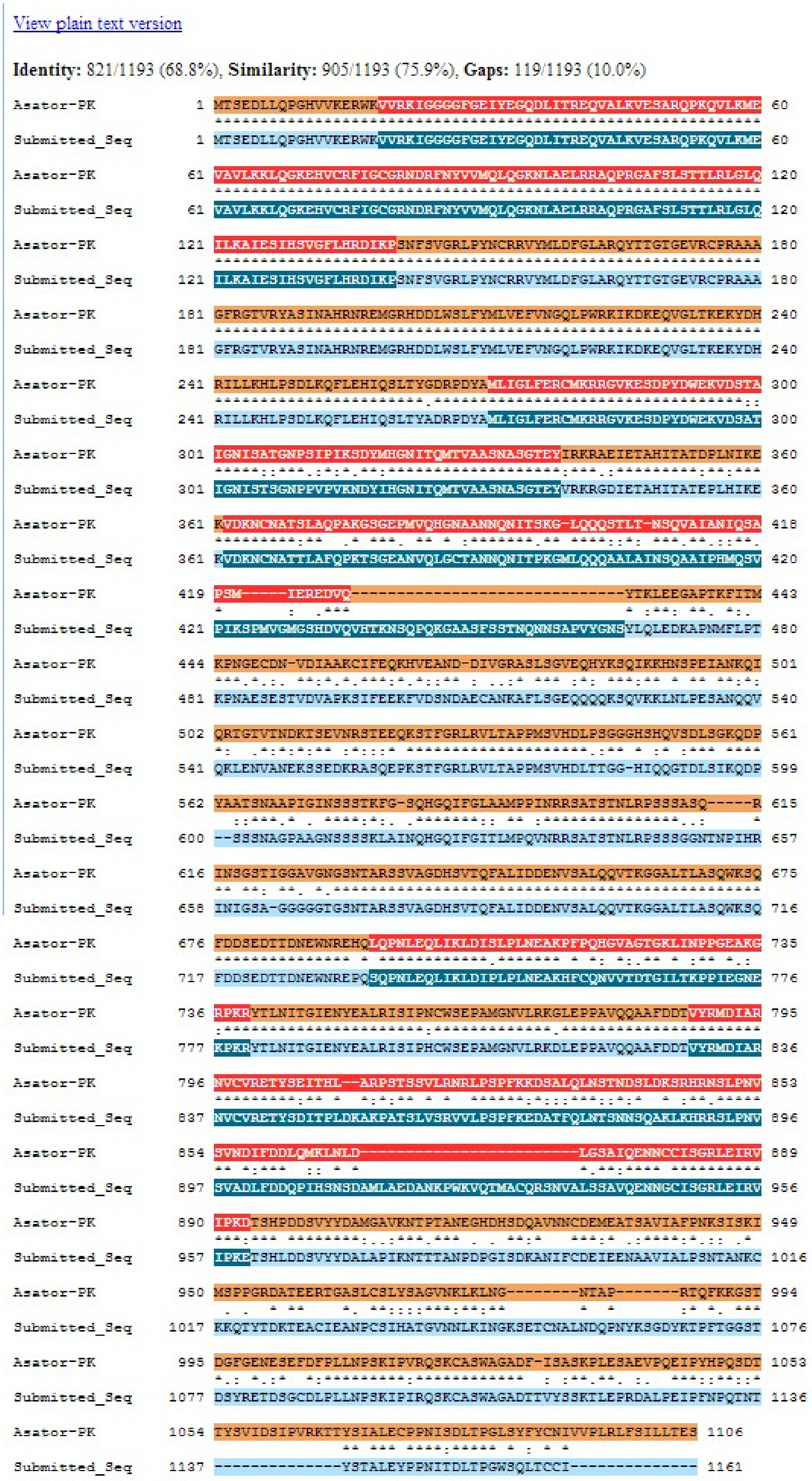
Stop Codon Coordinates: 20289-20287

Project Details

Project Group: D. biarmipes Dot

Project Name: contig28

View	Criteria	Status	Message
🔍	Check for Start Codon	✓ Pass	
	Acceptor for CDS 1	Skip	Already checked for Start Codon
🔍	Donor for CDS 1	✓ Pass	
🔍	Acceptor for CDS 2	✓ Pass	
🔍	Donor for CDS 2	✓ Pass	
🔍	Acceptor for CDS 3	✓ Pass	
🔍	Donor for CDS 3	✓ Pass	
🔍	Acceptor for CDS 4	✓ Pass	
🔍	Donor for CDS 4	✓ Pass	
🔍	Acceptor for CDS 5	✓ Pass	
🔍	Donor for CDS 5	✓ Pass	
🔍	Acceptor for CDS 6	✓ Pass	
🔍	Donor for CDS 6	✓ Pass	
🔍	Acceptor for CDS 7	✓ Pass	
🔍	Donor for CDS 7	✓ Pass	
🔍	Acceptor for CDS 8	✓ Pass	
🔍	Donor for CDS 8	✓ Pass	
🔍	Acceptor for CDS 9	✓ Pass	
🔍	Donor for CDS 9	✓ Pass	
🔍	Acceptor for CDS 10	✓ Pass	
🔍	Donor for CDS 10	✓ Pass	
🔍	Acceptor for CDS 11	✓ Pass	
	Donor for CDS 11	Skip	Already checked for Stop Codon
🔍	Check for Stop Codon	✓ Pass	
🔍	Additional Checks	✓ Pass	
🔍	Number of coding exons matched D. melanogaster ortholog	⚠ Warn	Cannot find the D. melanogaster ortholog: Asator-Pk

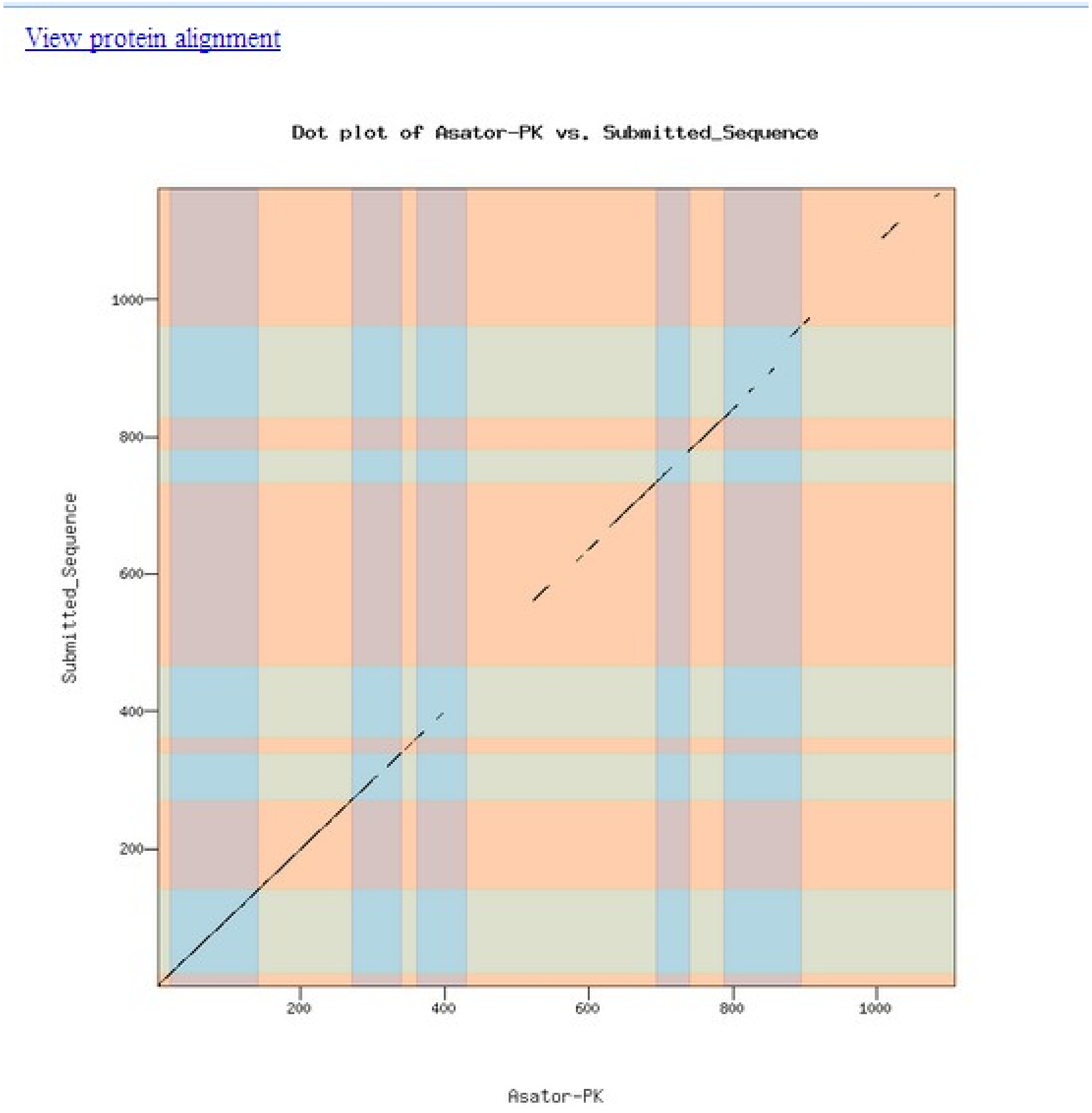


View plain text version

Identity: 821/1193 (68.8%), **Similarity:** 905/1193 (75.9%), **Gaps:** 119/1193 (10.0%)

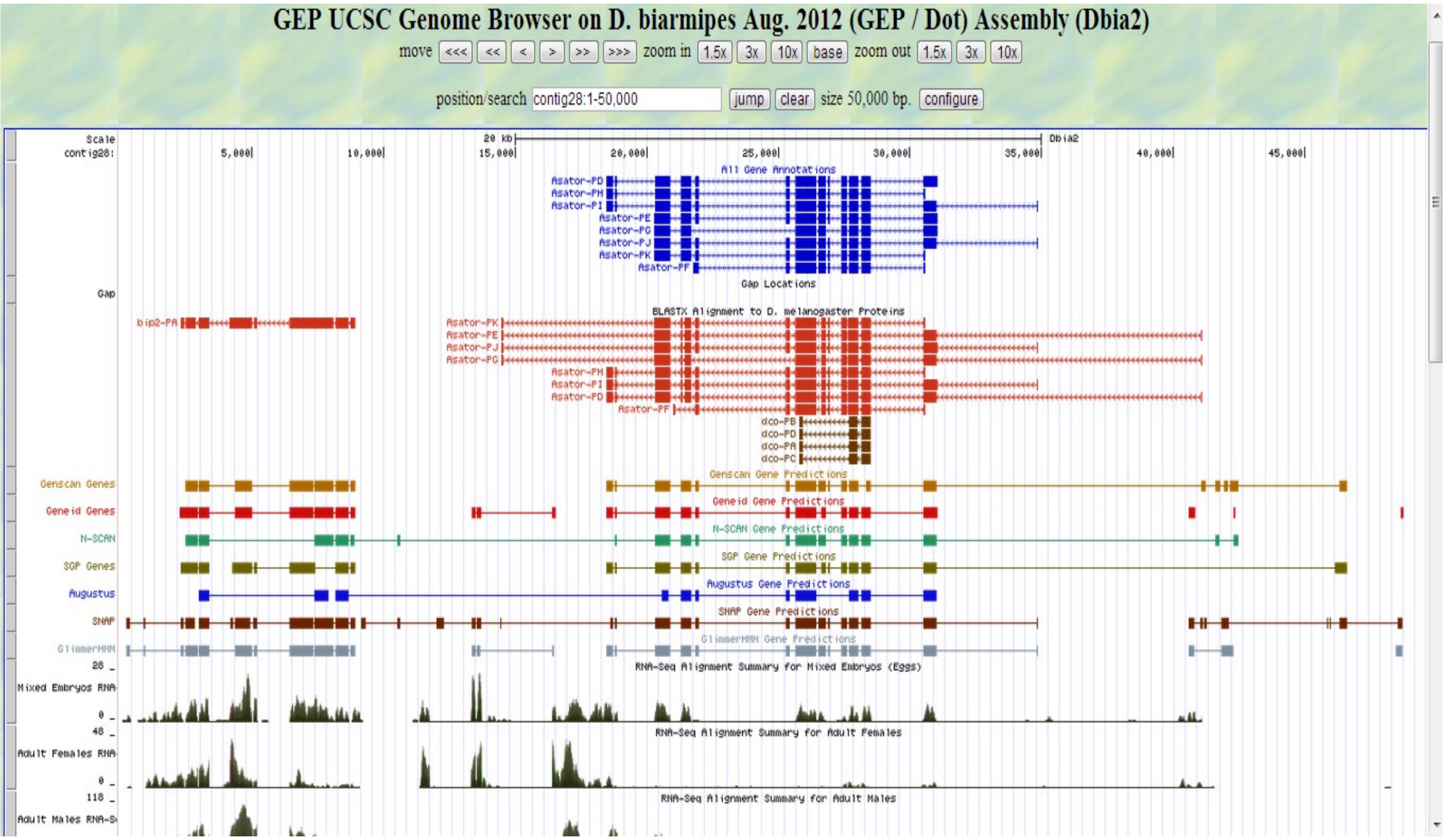
```
Asator-PK           1 MTSEDLLQPGHVVKERWKVVRKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKME 60
                      ************************************************************
Submitted_Seq       1 MTSEDLLQPGHVVKERWKVVRKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKME 60

Asator-PK          61 VAVLKKLQGKEHVCRFIGCGRNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQ 120
                      ************************************************************
Submitted_Seq      61 VAVLKKLQGKEHVCRFIGCGRNDRFNYVVMQLQGKNLAELRRAQPRGAFSLSTTLRLGLQ 120

Asator-PK         121 ILKAIESIHSVGFLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA 180
                      ************************************************************
Submitted_Seq     121 ILKAIESIHSVGFLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA 180

Asator-PK         181 GFRGTVRYASINAHRNREMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH 240
                      ************************************************************
Submitted_Seq     181 GFRGTVRYASINAHRNREMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH 240

Asator-PK         241 RILLKHLPSDLKQFLEHIQSLTYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTA 300
                      ***********************.**********************************::
Submitted_Seq     241 RILLKHLPSDLKQFLEHIQSLTYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSAT 300

Asator-PK         301 IGNISATGNPSIPIKSDYMHGNITQMTVAASNASGTEYIRKRAEIETAHITATDPLNIKE 360
                      *****::***.:*:*.**:*******************:***.:*********:**:***
Submitted_Seq     301 IGNISTSGNPPVPVKNDYIHGNITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKE 360

Asator-PK         361 KVDKNCNATSLAQPAKGSGEPMVQHGNAANNQNITSKG-LQQQSTLT-NSQVAIANIQSA 418
                      *********:**  .* ***. ** * :*******.** ****::*: ***.**.::**.
Submitted_Seq     361 KVDKNCNATTLAFQPKTSGEANVQLGCTANNQNITPKGMLQQQAALAINSQAAIPHMQSV 420

Asator-PK         419 PSM-----IEREDVQ------------------------------YTKLEEGAPTKFITM 443
                      *       :  .***                              * :**: **. *:.
Submitted_Seq     421 PIKSPMVGMGSHDVQVHTKNSQPQKGAASFSSTNQNNSAPVYGNSYLQLEDKAPNMFLPT 480

Asator-PK         444 KPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQI 501
                      ***.*.:. **:*.*.***:*.*::** : ..:* *** :*: ***:** * ** **:*:
Submitted_Seq     481 KPNAESESTVDVAPKSIFEEKFVDSNDAECANKAFLSGEQQQQKSQVKKLNLPESANQQV 540

Asator-PK         502 QRTGTVTNDKTSEVNRSTEEQKSTFGRLRVLTAPPMSVHDLPSGGGHSHQVSDLSGKQDP 561
                      *:  .*:*:*:** :*:::* ********************.:** * :* :*** ****
Submitted_Seq     541 QKLENVANEKSSEDKRASQEPKSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSIKQDP 599

Asator-PK         562 YAATSNAAPIGINSSSTKFG-SQHGQIFGLAAMPPINRRSATSTNLRPSSSASQ-----R 615
                        ::***.* . ****:*:. .*******:: ** :***************..:     *
Submitted_Seq     600 --SSSNAGPAAGNSSSSKLAINQHGQIFGITLMPQVNRRSATSTNLRPSSSGGNTNPIHR 657

Asator-PK         616 INSGSTIGGAVGNGSNTARSSVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQ 675
                      ** **: **. *.***********************************************
Submitted_Seq     658 INIGSA-GGGGGTGSNTARSSVAGDHSVTQFALIDDENVSALQQVTKGGALTLASQWKSQ 716

Asator-PK         676 FDDSEDTTDNEWNREHQLQPNLEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEAKG 735
                      *************** * ************.******* * *: *:.** * :** *.:
Submitted_Seq     717 FDDSEDTTDNEWNREPQSQPNLEQLIKLDIPLPLNEAKHFCQNVVTDTGILTKPPIEGNE 776

Asator-PK         736 RPKRYTLNITGIENYEALRISIPNCWSEPAMGNVLRKGLEPPAVQQAAFDDTVYRMDIAR 795
                      :**********************:**************.*********************
Submitted_Seq     777 KPKRYTLNITGIENYEALRISIPHCWSEPAMGNVLRKDLEPPAVQQAAFDDTVYRMDIAR 836

Asator-PK         796 NVCVRETYSEITHL--ARPSTSSVLRNRLPSPFKKDSALQLNSTNDSLDKSRHRNSLPNV 853
                      *********:** *  *:*:** * *  ******:*:::***::*:*  * :**.*****
Submitted_Seq     837 NVCVRETYSDITPLDKAKPATSLVSRVVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNV 896

Asator-PK         854 SVNDIFDDLQMKLNLD------------------------LGSAIQENNCCISGRLEIRV 889
                      ** *:***  :: * *                        *.**:**** **********
Submitted_Seq     897 SVADLFDDQPIHSNSDAMLAEDANKPWKVQTMACQRSNVALSSAVQENNGCISGRLEIRV 956

Asator-PK         890 IPKDTSHPDDSVYYDAMGAVKNTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKI 949
                      ***:*** ********:..:***.*** . . **:*   ***:* .:****:*.:: .*
Submitted_Seq     957 IPKETSHLDDSVYYDALAPIKNTTTANPDPGISDKANIFCDEIEENAAVIALPSNTANKC 1016

Asator-PK         950 MSPPGRDATEERTGASLCSLYSAGVNKLKLNG--------NTAP-------RTQFKKGST 994
                       . .  * **    *. **::::***:**:**        *  *       :* *. ***
Submitted_Seq    1017 KKQTYTDKTEACIEANPCSIHATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGST 1076

Asator-PK         995 DGFGENESEFDFPLLNPSKIPVRQSKCASWAGADF-ISASKPLESAEVPQEIPYHPQSDT 1053
                      *.: *.:*  *:*********:************  : :**.**. :.  ***::**::*
Submitted_Seq    1077 DSYRETDSGCDLPLLNPSKIPIRQSKCASWAGADTTVYSSKTLEPRDALPEIPFNPQTNT 1136

Asator-PK        1054 TYSVIDSIPVRKTTYSIALECPPNISDLTPGLSYFYCNIVVPLRLFSILLTES 1106
                                    ** *** ****:***** * : * *
Submitted_Seq    1137 --------------YSTALEYPPNITDLTPGWSQLTCCI-------------- 1161
```

Dotplot



View protein alignment

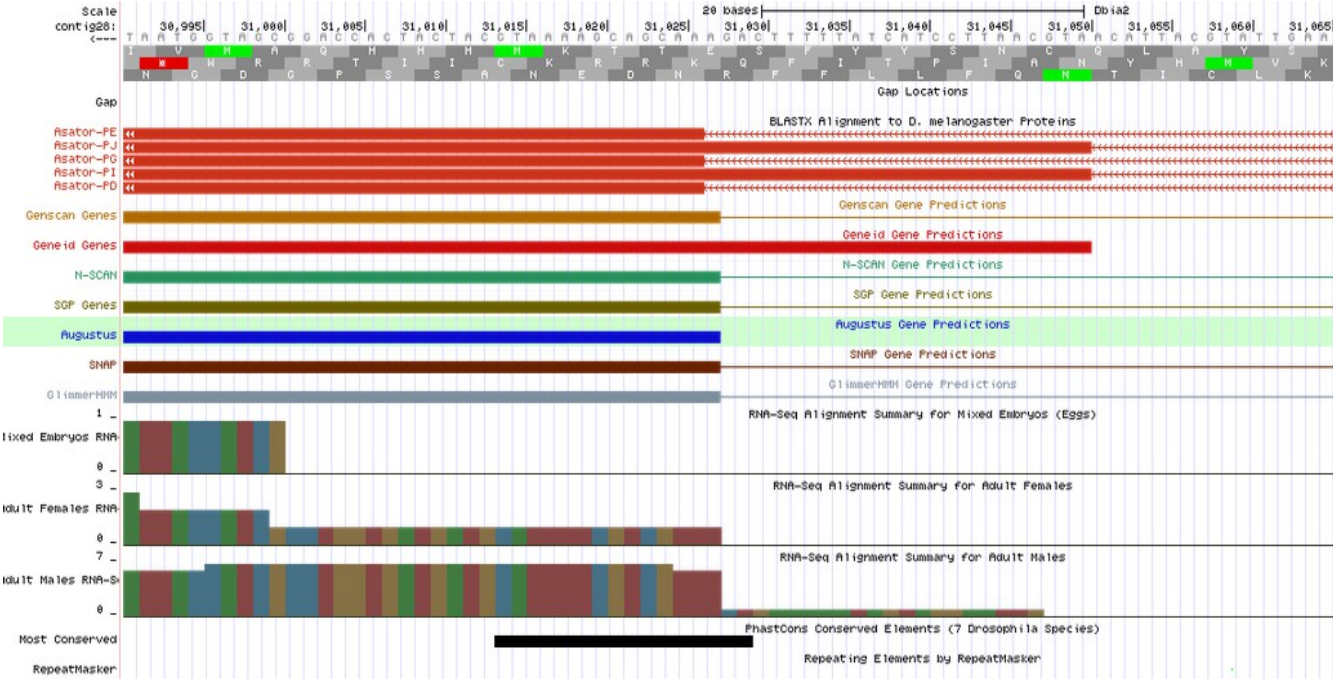
Custom track(s) viewed on the UCSC Genome Browser



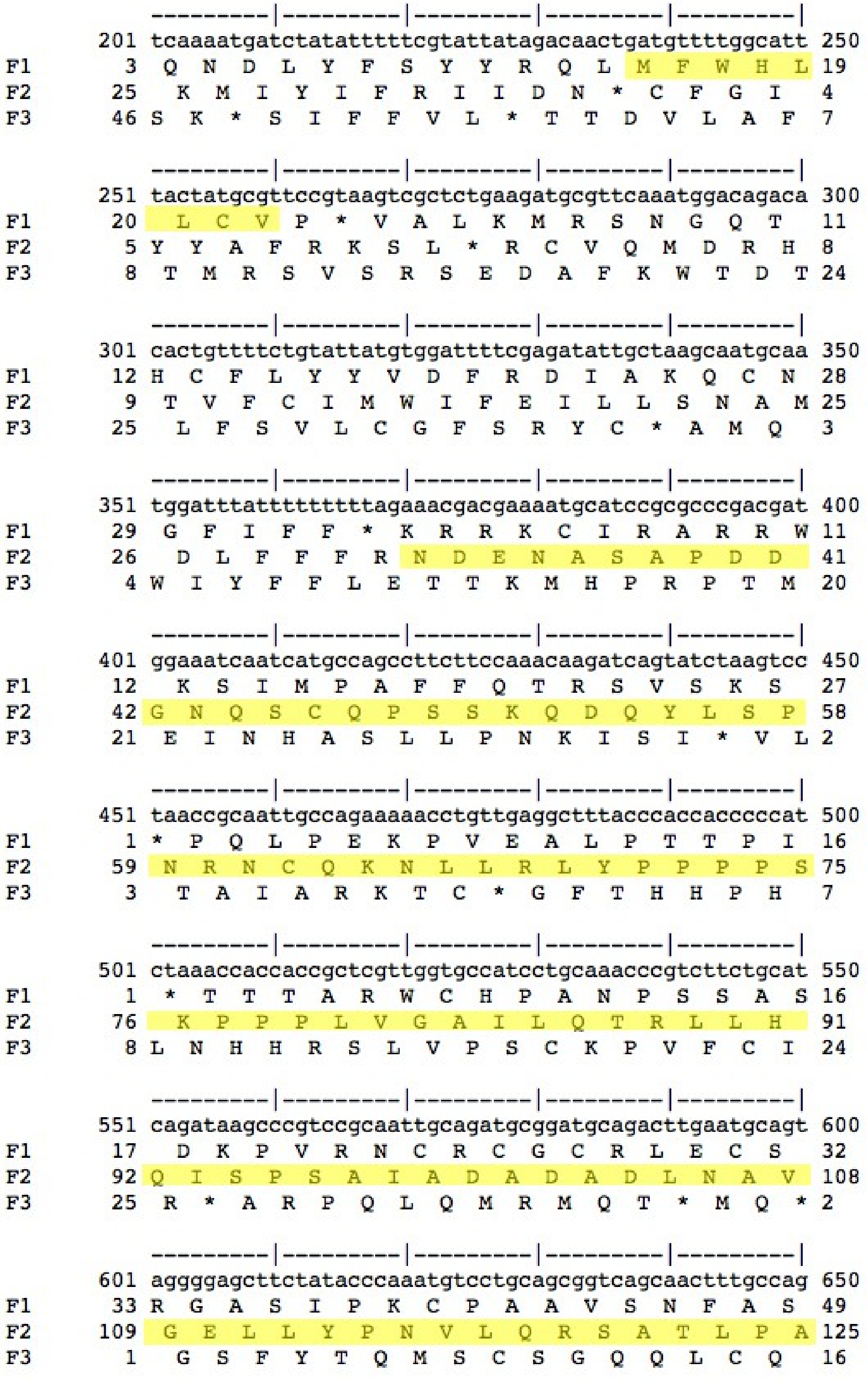
****Loss of intron**: The *Asator* gene in *D. biarmipes* is missing an intron that is present in *D. melanogaster* between the exons 17 and 16_1. Blasting was fairly unhelpful, as exon 17 is too short to give significant hits. My initial suspicion was due to the fact that when exon 17 is present, it is always followed by 16_1, and this pair is mutually exclusive to the utilization of exon 16_0. So, using the genome browser and the EMBOSS toolkit (showorf), I compared the peptide sequences of the relevant area for *D. melanogaster* and *D. biarmipes* and highlighted the surrounding sequence to display the lost intron.

**This is also why, in the isoform models that have exon 17/16_1, there is a warning in the gene model checklists about a different number of CDSs than in the *D. melanogaster* model.

Exon 17/16_1 as shown in the UCSC Genome Browser:

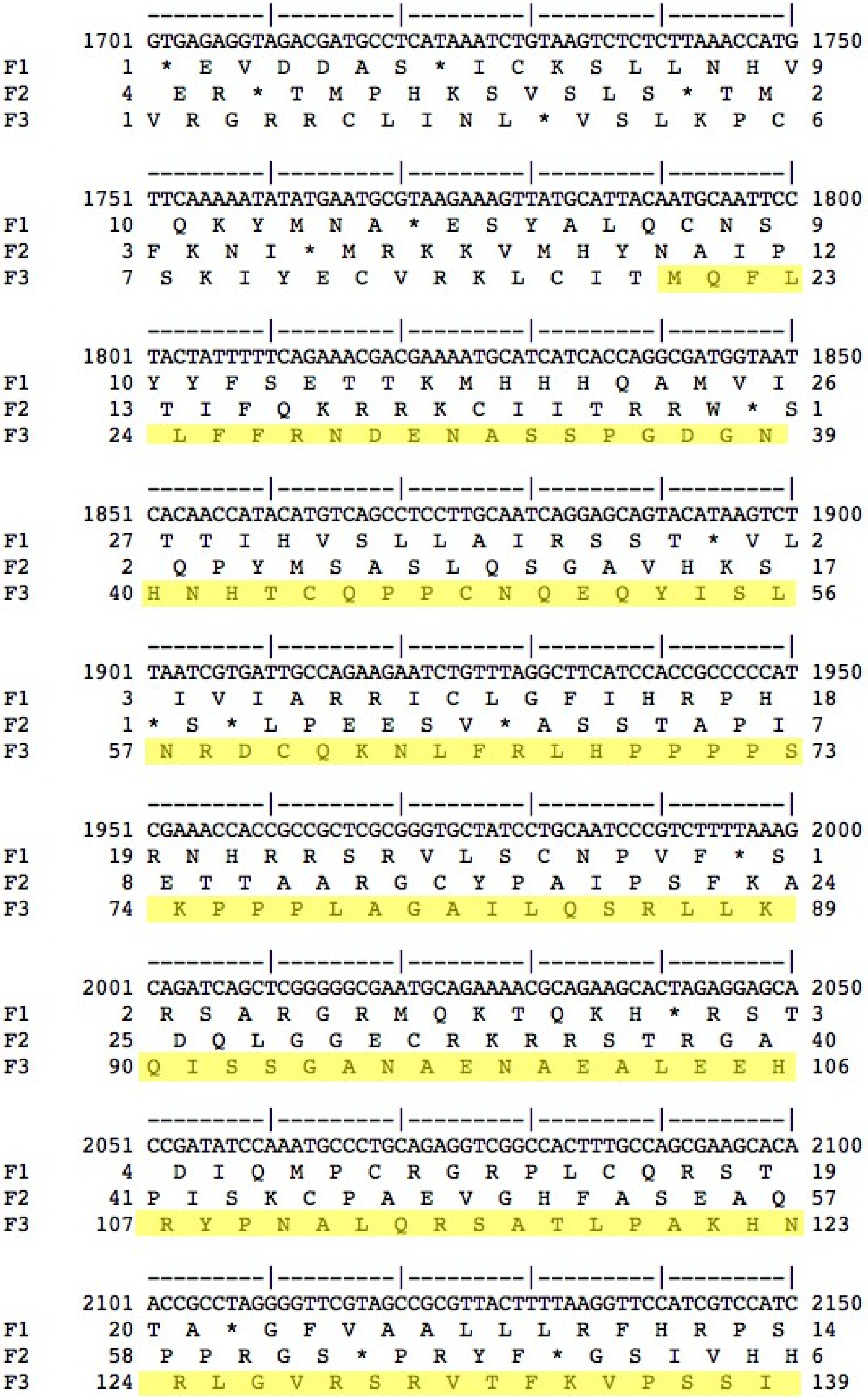


Peptide sequence (with intron) for exons 17 and 16_1 in D. melanogaster:



```
          ---------|---------|---------|---------|---------|
      201 tcaaaatgatctatatttttcgtattatagacaactgatgttttggcatt 250
F1      3  Q  N  D  L  Y  F  S  Y  Y  R  Q  L  M  F  W  H  L 19
F2     25   K  M  I  Y  I  F  R  I  I  D  N  *  C  F  G  I   4
F3     46 S  K  *  S  I  F  F  V  L  *  T  T  D  V  L  A  F  7

          ---------|---------|---------|---------|---------|
      251 tactatgcgttccgtaagtcgctctgaagatgcgttcaaatggacagaca 300
F1     20   L  C  V  P  *  V  A  L  K  M  R  S  N  G  Q  T   11
F2      5 Y  Y  A  F  R  K  S  L  *  R  C  V  Q  M  D  R  H  8
F3      8  T  M  R  S  V  S  R  S  E  D  A  F  K  W  T  D  T 24

          ---------|---------|---------|---------|---------|
      301 cactgttttctgtattatgtggattttcgagatattgctaagcaatgcaa 350
F1     12 H  C  F  L  Y  Y  V  D  F  R  D  I  A  K  Q  C  N  28
F2      9  T  V  F  C  I  M  W  I  F  E  I  L  L  S  N  A  M 25
F3     25   L  F  S  V  L  C  G  F  S  R  Y  C  *  A  M  Q   3

          ---------|---------|---------|---------|---------|
      351 tggatttatttttttttagaaacgacgaaaatgcatccgcgcccgacgat 400
F1     29  G  F  I  F  F  *  K  R  R  K  C  I  R  A  R  R  W 11
F2     26   D  L  F  F  F  R  N  D  E  N  A  S  A  P  D  D   41
F3      4 W  I  Y  F  F  L  E  T  T  K  M  H  P  R  P  T  M  20

          ---------|---------|---------|---------|---------|
      401 ggaaatcaatcatgccagccttcttccaaacaagatcagtatctaagtcc 450
F1     12   K  S  I  M  P  A  F  F  Q  T  R  S  V  S  K  S   27
F2     42 G  N  Q  S  C  Q  P  S  S  K  Q  D  Q  Y  L  S  P  58
F3     21  E  I  N  H  A  S  L  L  P  N  K  I  S  I  *  V  L 2

          ---------|---------|---------|---------|---------|
      451 taaccgcaattgccagaaaaacctgttgaggctttacccaccacccccat 500
F1      1 *  P  Q  L  P  E  K  P  V  E  A  L  P  T  T  P  I  16
F2     59  N  R  N  C  Q  K  N  L  L  R  L  Y  P  P  P  P  S 75
F3      3   T  A  I  A  R  K  T  C  *  G  F  T  H  H  P  H   7

          ---------|---------|---------|---------|---------|
      501 ctaaaccaccaccgctcgttggtgccatcctgcaaacccgtcttctgcat 550
F1      1  *  T  T  T  A  R  W  C  H  P  A  N  P  S  S  A  S 16
F2     76   K  P  P  P  L  V  G  A  I  L  Q  T  R  L  L  H   91
F3      8 L  N  H  H  R  S  L  V  P  S  C  K  P  V  F  C  I  24

          ---------|---------|---------|---------|---------|
      551 cagataagcccgtccgcaattgcagatgcggatgcagacttgaatgcagt 600
F1     17   D  K  P  V  R  N  C  R  C  G  C  R  L  E  C  S   32
F2     92 Q  I  S  P  S  A  I  A  D  A  D  A  D  L  N  A  V  108
F3     25  R  *  A  R  P  Q  L  Q  M  R  M  Q  T  *  M  Q  * 2

          ---------|---------|---------|---------|---------|
      601 aggggagcttctatacccaaatgtcctgcagcggtcagcaactttgccag 650
F1     33 R  G  A  S  I  P  K  C  P  A  A  V  S  N  F  A  S  49
F2    109  G  E  L  L  Y  P  N  V  L  Q  R  S  A  T  L  P  A 125
F3      1   G  S  F  Y  T  Q  M  S  C  S  G  Q  Q  L  C  Q   16
```

Peptide sequence (displaying the missing intron between exons 17 and 16_1) in *D. biarmipes:*



```
          ---------|---------|---------|---------|---------|
     1701 GTGAGAGGTAGACGATGCCTCATAAATCTGTAAGTCTCTCTTAAACCATG 1750
F1      1  *  E  V  D  D  A  S  *  I  C  K  S  L  L  N  H  V 9
F2      4   E  R  *  T  M  P  H  K  S  V  S  L  S  *  T  M   2
F3      1 V  R  G  R  R  C  L  I  N  L  *  V  S  L  K  P  C  6

          ---------|---------|---------|---------|---------|
     1751 TTCAAAAATATATGAATGCGTAAGAAAGTTATGCATTACAATGCAATTCC 1800
F1     10   Q  K  Y  M  N  A  *  E  S  Y  A  L  Q  C  N  S   9
F2      3 F  K  N  I  *  M  R  K  K  V  M  H  Y  N  A  I  P  12
F3      7  S  K  I  Y  E  C  V  R  K  L  C  I  T  M  Q  F  L 23

          ---------|---------|---------|---------|---------|
     1801 TACTATTTTTCAGAAACGACGAAAATGCATCATCACCAGGCGATGGTAAT 1850
F1     10 Y  Y  F  S  E  T  T  K  M  H  H  H  Q  A  M  V  I  26
F2     13  T  I  F  Q  K  R  R  K  C  I  I  T  R  R  W  *  S 1
F3     24   L  F  F  R  N  D  E  N  A  S  S  P  G  D  G  N   39

          ---------|---------|---------|---------|---------|
     1851 CACAACCATACATGTCAGCCTCCTTGCAATCAGGAGCAGTACATAAGTCT 1900
F1     27  T  T  I  H  V  S  L  L  A  I  R  S  S  T  *  V  L 2
F2      2   Q  P  Y  M  S  A  S  L  Q  S  G  A  V  H  K  S   17
F3     40 H  N  H  T  C  Q  P  P  C  N  Q  E  Q  Y  I  S  L  56

          ---------|---------|---------|---------|---------|
     1901 TAATCGTGATTGCCAGAAGAATCTGTTTAGGCTTCATCCACCGCCCCCAT 1950
F1      3   I  V  I  A  R  R  I  C  L  G  F  I  H  R  P  H   18
F2      1 *  S  *  L  P  E  E  S  V  *  A  S  S  T  A  P  I  7
F3     57  N  R  D  C  Q  K  N  L  F  R  L  H  P  P  P  P  S 73

          ---------|---------|---------|---------|---------|
     1951 CGAAACCACCGCCGCTCGCGGGTGCTATCCTGCAATCCCGTCTTTTAAAG 2000
F1     19 R  N  H  R  R  S  R  V  L  S  C  N  P  V  F  *  S  1
F2      8  E  T  T  A  A  R  G  C  Y  P  A  I  P  S  F  K  A 24
F3     74   K  P  P  P  L  A  G  A  I  L  Q  S  R  L  L  K   89

          ---------|---------|---------|---------|---------|
     2001 CAGATCAGCTCGGGGGCGAATGCAGAAAACGCAGAAGCACTAGAGGAGCA 2050
F1      2  R  S  A  R  G  R  M  Q  K  T  Q  K  H  *  R  S  T 3
F2     25   D  Q  L  G  G  E  C  R  K  R  R  S  T  R  G  A   40
F3     90 Q  I  S  S  G  A  N  A  E  N  A  E  A  L  E  E  H  106

          ---------|---------|---------|---------|---------|
     2051 CCGATATCCAAATGCCCTGCAGAGGTCGGCCACTTTGCCAGCGAAGCACA 2100
F1      4   D  I  Q  M  P  C  R  G  R  P  L  C  Q  R  S  T   19
F2     41 P  I  S  K  C  P  A  E  V  G  H  F  A  S  E  A  Q  57
F3    107  R  Y  P  N  A  L  Q  R  S  A  T  L  P  A  K  H  N 123

          ---------|---------|---------|---------|---------|
     2101 ACCGCCTAGGGGTTCGTAGCCGCGTTACTTTTAAGGTTCCATCGTCCATC 2150
F1     20 T  A  *  G  F  V  A  A  L  L  L  R  F  H  R  P  S  14
F2     58  P  P  R  G  S  *  P  R  Y  F  *  G  S  I  V  H  H 6
F3    124   R  L  G  V  R  S  R  V  T  F  K  V  P  S  S  I   139
```