

## 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\*\***

Gene Model Checker

Configure Gene Model

Model Details

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

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

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

Expand All

Collapse All

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

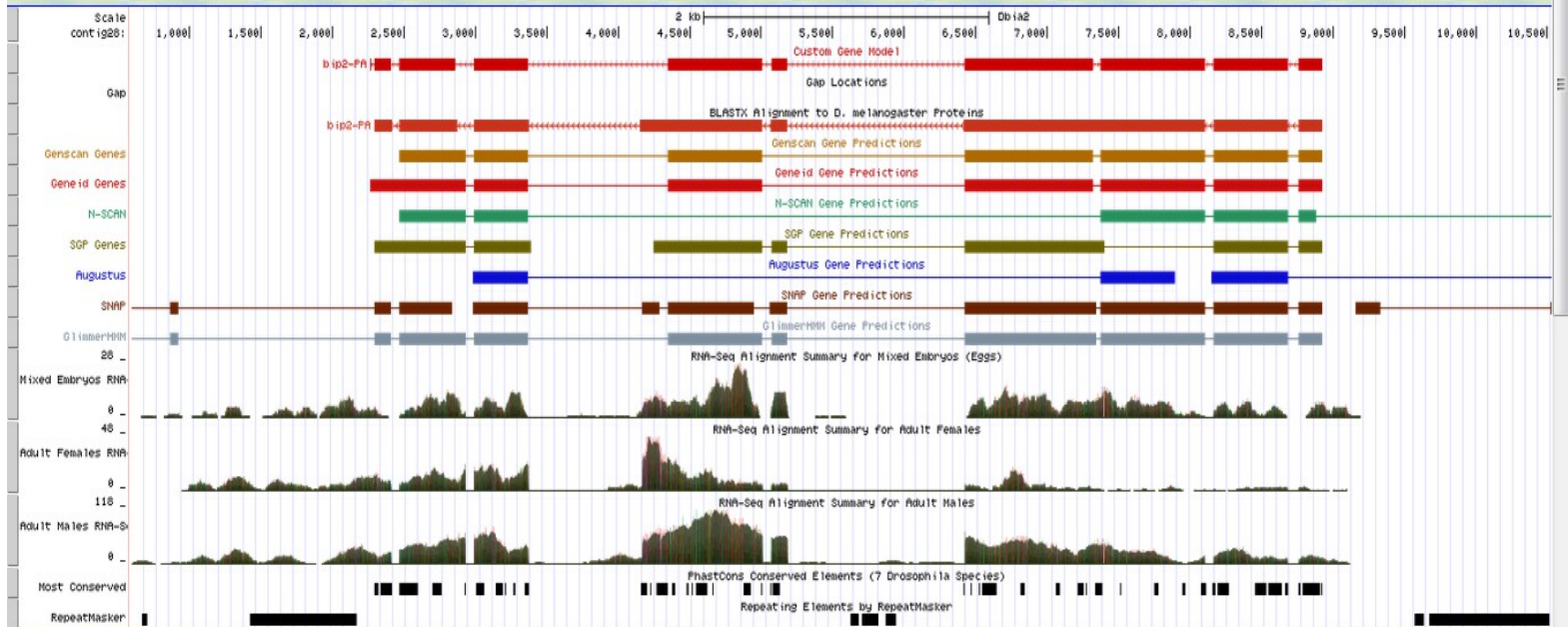
Verify Gene Model

Reset Form

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

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

position/search contig28:597-10,595 jump clear size 9,999 bp. configure



move start < 2.0 > 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. move end < 2.0 >

default tracks default order hide all manage custom tracks track hubs configure reverse resize refresh

collapse all

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

expand all

## Alignment of bip2-PA vs. Submitted\_Seq

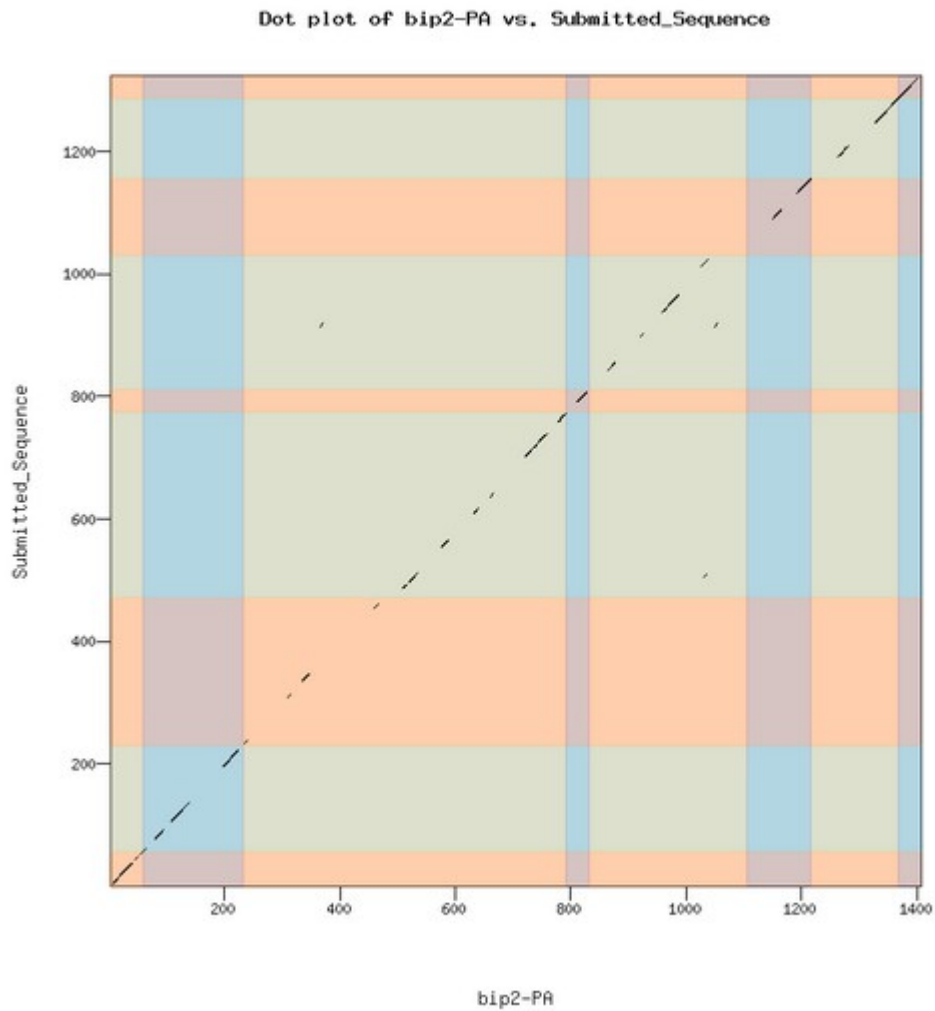
[View plain text version](#)

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

bip2-PA	1	MMADRYASDLALVVVAQITQTIGYSCITLSAPLELLQDILQKVFQEFARDMHRHMEHANRI	60
Submitted_Seq	1	M-ADRYASDLALVVIAQITQTIGYSCITLSAPLELLQDIMQKFTQEFARDLHCNMEHANRI	59
bip2-PA	61	EPNLKDARLSIKNLSINVQELLDYIGNVEPVGFIRDVPQFPFGKSVNMNFKPGSAETLT	120
Submitted_Seq	60	EPNLDAQLSMKLNINIQELLDYIGNVEPVGFTRDVPHPFIRKAINNMNFKPGSAETLT	119
bip2-PA	121	RPVYIFEYLPFMQDPELREIPADVQKEFSEKQEFCSKAQYSSSTNAADKLGAKHIDSISPN	180
Submitted_Seq	120	RPVYIFEYLPFMQDPEPRE3QSEAGREFPQKQELVSKAEFNVTSSEAKLSTNQVDS3SPN	179
bip2-PA	181	TVINFRSNAFELDVGS3SVREMSSVVMTTGGFISPAIEGKLPEDIIPDIVEKFLGLDAPFS	240
Submitted_Seq	180	AVINFSSNPLDS3DVGS3SVREMSSVVMTTGGFISPAIEGKLPPEPFIPDIIEKFLGLDAPFS	239
bip2-PA	241	PTIIVNSLQKSPQLALSDRDKTVNFRKIIPSETKIFKQNAALLTSGHSESSIIYASNNHT	300
Submitted_Seq	240	SLIAVSHLEHSEKELITSEKDAVNRTI-NSETKNFNHNAVLLISDSADASLLYSSNSLP	298
bip2-PA	301	M---LTVITPKKNRKQKHDICEPGQSELLTNPFKAQEK3QRKALKMYKQTSKNQRDSS	357
Submitted_Seq	299	MSSALTATISKKNRKPKPDIAIEHGQIVKSNISFGKSQEK3QRKALKMFQKLSK3QNDA3	358
bip2-PA	358	INQIQNMKKLKKKFNKGSF--DPNKKHLEKIFKKQSKLKQNDLQLDIDEKHFLQNSQTT3	415
Submitted_Seq	359	NSQILNMKKSKKRVNHN3SLDPSKVNIEKMFKKQNRHKQKSVQLE-----TQTLS	409
bip2-PA	416	GLAIKAN3N-ENALQADIP3QPPTVTSQIENNFRNSIYPVQPSVIQQTQVLLAEK3GSE	474
Submitted_Seq	410	DFPIEADMNTENFIAADIPCAEPVVLKQIEIG3Q3SVFPVQPGGIQQTQVSLQSKM--H	467
bip2-PA	475	PERSKLDIFKKISKPRIPRPDIGATLVPPGTGVSVFGSTMTPSALISLPSGTTITPTPSL	534
Submitted_Seq	468	QNRNE-----DGGATQVPPGIGVRVFGSTMPVSTLISLPSGTTITPTPPV	512
bip2-PA	535	GLNSENKNVPSMKINPCNIFDGTIPLTKAGVEMSIIDSPKPKRGRKPGGKNVIKQTNVV	594
Submitted_Seq	513	GQTPEDAMNPN3KINQYGVLGQKL3SPKDDIEMSTINPIKAKRGRKPGGKNLVKQTHPT	572
bip2-PA	595	SQPLINKVERKKSSQAITLPL3SSPMIITQSS3LNV3PPTTEPLNLCNTEQPSNNFL3NL3A	654

Submitted Seq	770	ALSNASIVPGIQTGMVFLPLQLQFFPRPGLIPTGPGLFPAVTVGLVGFNGHNRIPISPTI	769
bip2-FA	655	KEKKDRKKYK-----SLPENVMQLDKSCSPEKASTLHNSIRFTKQNNNDVQCSKRFVTISM	709
Submitted Seq	693	---KKERKKYKLFKFDITGLPQNVKDFNNVSETTSSTVTSIHLMKQKNEAPCPDSILMTAPN	690
bip2-FA	710	ALSNASIVPGIQTGMVFLPLQLQFFPRPGLIPTGPGLFPAVTVGLVGFNGHNRIPISPTI	769
Submitted Seq	691	TLNNTSLYPGNTGMVFLPLQLHFPFRPGLIPSGPLFPAVTVGLVGFNGHNRIPISPTI	750
bip2-FA	770	DYPGFEEESVADTVRCPPFKDSFVTDSDQFLSRSSSTQTDLQMDRNYCNVAPLVFDSMKFAP	829
Submitted Seq	751	AFPGTEGSSVADTVRCPPFKDSADTGNDQLLRSG-QMDLQIDRNYCNVAPLVFDSMKFSE	809
bip2-FA	830	CKSVSCASTILENPSAQATSKINIKPLLEASGNPDDPIEVSDSDSESMHNRQMVQRKTPI	889
Submitted Seq	810	CKPVSCGNTDMENTFT--PSKFKAKPSVQASGNLGDPIEVSDSDSEIQNRQIVQKKSPI	867
bip2-FA	890	SSPTVVKTSFAELNPSSFTSAAANGEEKSKMDLRNIV-SLSNNEPLKFKKLVKQSFDDV	948
Submitted Seq	868	TSPNHAKSSLVQSQHQTFATPSPLCEDKNQMDLRNVLPNTSPSESIKKLKKSVKLNLPDV	927
bip2-FA	949	KSVPHTPASHSSFPQFNLPNFMGGDKFSLAGGADLIPLSRVS-DSYSSKIVFSSSLGGT	1007
Submitted Seq	928	KNFIQIAPSQSSFPQFNLPNFIGGDKFSLAGGADLIPLARIDCGSAYSSHKVPSSSLTGG	987
bip2-FA	1008	IPNQI-----KISEEHNIFSTFSNYEDITITPTGLTSLPEPMKHHKHLKLVKQEGKNNK	1061
Submitted Seq	988	VASGVIPILPNHISEDQQFMPTFPNYEDISITPTGALSDEPKI-----LIQIPE-----	1036
bip2-FA	1062	KEKKDKSKKADQIGLPSFKSDRKIKANDKQKKEKKDKDRQILVHIPDDTEEFDKVPI	1121
Submitted Seq	1037	----VDEFSRAPFI-----NNLDYGTQSVFAST-----	1061
bip2-FA	1122	ANNDEPVLKSSSMTINPSLGAATSGI-----SPNQIPKLTILKLSGKSTLFSSEKEMTDA	1176
Submitted Seq	1062	----FVLKTSFMSVKPPLSATCAMSPRIQQSFSQIPKLTILKLSGKSTTCPEKERDITDA	1116
bip2-FA	1177	SKLKQTTILSSSENKKRERDMSPELARFSPVLTVGPPRNKQSETLHLGNSSTAVLPVPSFVA	1236
Submitted Seq	1117	VKVQPTMFVVENKERERDMSPELARFSPVLTVGPPRNKQ-----LIQEPFFMA	1164
bip2-FA	1237	VRAVQLFVSQTSNSAGWLSNPNNSNTASSTLSASSVLLPQQLMLAPHTIMNNFVPMCN	1296
Submitted Seq	1165	TRTSQISTSQNSSSSAGWMSNPNNSNVASSTLSASSVLLPQQLMLTSMNTMNSL-----	1219
bip2-FA	1297	STGTVSKSGLCSSPNTISEENANAMQIAESSRPSYVDAEGNRIWICPACGKVDDGSAMI	1356
Submitted Seq	1220	SSGG-PK3CSLSPANVPEENSH---IAETNRPSSYVDAEGNRIWICPACGKVDDGSAMI	1275
bip2-FA	1357	GCDGCDAWYHYICVGITFAPKDNDWFCRVCTKRIHSGEKKRRNRKKR	1406
Submitted Seq	1276	GCDGCDAWYHYICVGITFAPKDNDWFCRVCTKRVHVHSEKKRGEIRK-	1324

## Gene model dot plot

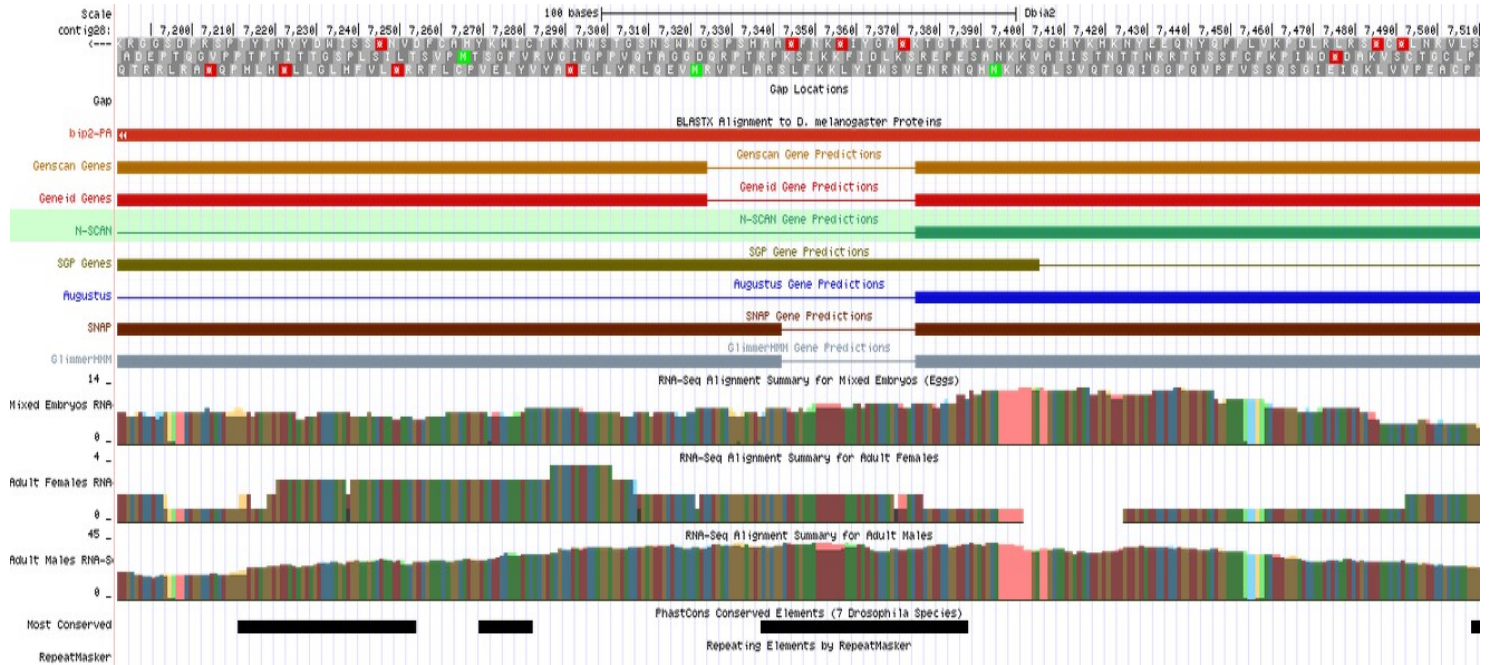


### **\*\*Gain of intron in *D. biarmipes*\*\***

The gene has an 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 toolkit (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 AAAAAATGCATCAGAACCGGAACGAAGTAAAGCTGGATATATTAAAAAAA 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 TTTCTAAGCCGCGCACTCCCCGTAGSATGGTGGAGCAACTCAGGTACCT 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 CCTGGAATAGCGTACGTGTATTGGGAAGTACCATGCCTGTTTCGACGTT 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 AGACTCCGGAGGACGCAATAATCCTAATTCGAAAATTAACCAATATGGT 1800
F1    93 T P E D A N P N S K I N Q Y G 108
F2     2 R L R D 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	caaacgcaagtgtattggcagaaaaaaaaagtgatcagaaccagaacg	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	aagtaattggacatttttaagaaaatttctaagccacgtactccacgtc	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	cagatattggggcgacttttagtacctcctggaaccgggtatctgtattt	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	ggaagcacaatgacccttcagcattaattagtcctccatcggggactac	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	tataactcctacaccttcacttgggtctaaattctgaaaacaaaaacgttc	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:

Gene Model Checker

Configure Gene Model

Model Details

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

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

Verify Gene Model Reset Form

Checklist

Expand All Collapse All

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	

Configure Gene Model

Model Details

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

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

Verify Gene Model Reset Form

Checklist

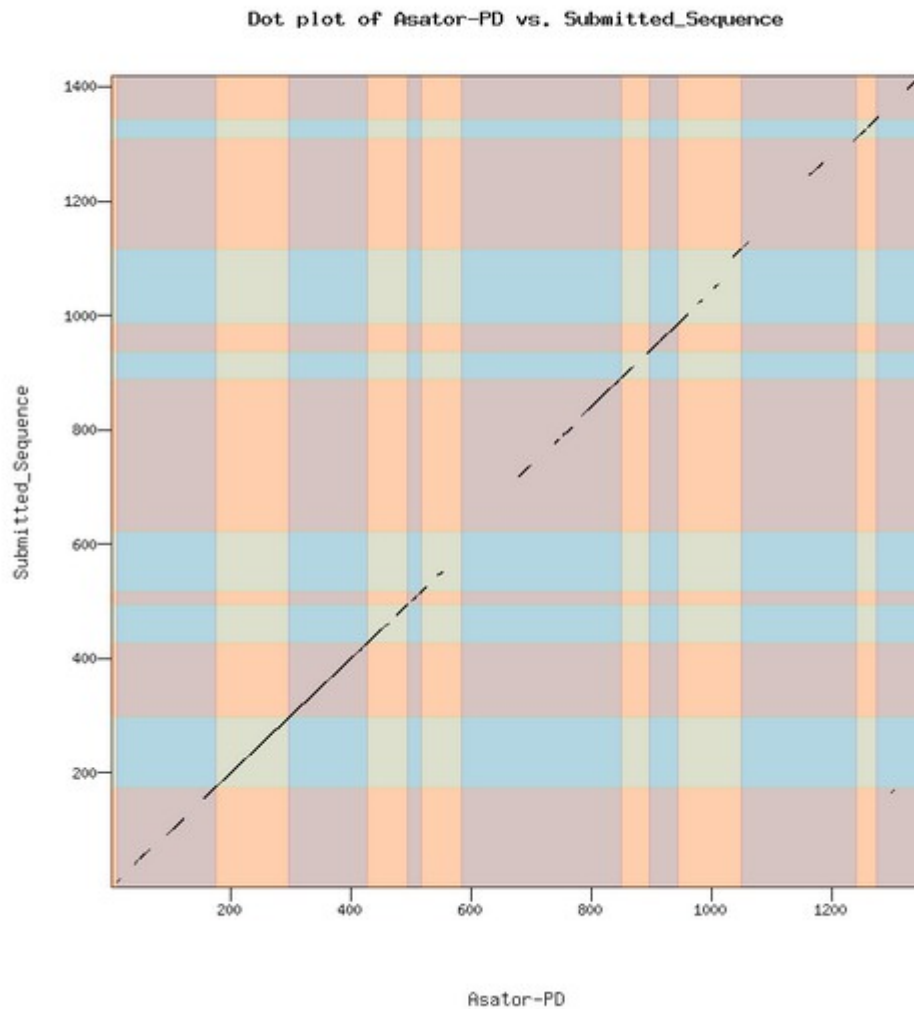
Expand All Collapse All

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-FD	1	MFHLLCVFNDENASAPDDGNQS--CQSSKQDQVLSFNRCQKNLLRLYPFPPSKPPPI	58
Submitted_Seq	1	MQF-LLFFRNDENASSPGDGNHNTCQPPCNQEQVISLNRDCQKNLFRHFPFPPSKPPPI	59
Asator-FD	59	VGAILQTRILLHQISPSAIADADADLNAVGELLYPNVLQRSATLPAKHNLGVRSRVTFKV	118
Submitted_Seq	60	AGAILQSRLLKQISSGANAE---NAEALAEHRYFNALQRSATLPAKHNLGVRSRVTFKV	116
Asator-FD	119	PSSNLPAQDSYSHQP--RNQVAAAKDGLVDVKAESVRMTSEDLLQPGHVVKERKVV	176
Submitted_Seq	117	PSSITPAVDPGSEPDGQNVAAERDGLVNFKERESCKMTSEDLLQPGHVVKERKVV	176
Asator-FD	177	RKIGGGGGEIYEGQDLITREQVALKVESARQPKQVLEMEVAVLKKLQKEHVCRFIGCG	236
Submitted_Seq	177	RKIGGGGGEIYEGQDLITREQVALKVESARQPKQVLEMEVAVLKKLQKEHVCRFIGCG	236
Asator-FD	237	RNDRFNYVVMQLQGENLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFHRDIKE	296
Submitted_Seq	237	RNDRFNYVVMQLQGENLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFHRDIKE	296
Asator-FD	297	SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNREMG	356
Submitted_Seq	297	SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNREMG	356
Asator-FD	357	RHDDLWSLFYMLVEFVNGQLFWRKIKDKQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS	416
Submitted_Seq	357	RHDDLWSLFYMLVEFVNGQLFWRKIKDKQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS	416
Asator-FD	417	LTYGDRPDYAMLIGLFCRCKRRGVKESDPYDWEKVDSTAGNISATGNFSIFIKSDYMH	476
Submitted_Seq	417	LTYADRPDYAMLIGLFCRCKRRGVKESDPYDWEKVDSTAGNISTSGNPPVPVKNNDYIH	476
Asator-FD	477	GNITQMTVAASNAGSTEYIRKRAPIETAHITATDPLNIKEKVDKNCNATSLAQPAKSGGE	536
Submitted_Seq	477	GNITQMTVAASNAGSTEYVRKRGDIETAHITATEPLHIKEVDKNCNATTLAQPKTSGE	536
Asator-FD	537	PMVQHGNAANNQMIITSGK-LQQQSTLT-NSQVAIANIQSAPSM-----IEREDVQ----	584
Submitted_Seq	537	ANVQLGCTANNQMITPKGMLQQQAALAINSQAAIPHMQSVPIKSPMVGMGSHDVQVHTKN	596
Asator-FD	585	-----YTKLEEGAPTKFITMKPNGECDN-VDIAARCIFEQ	618
Submitted_Seq	597	SQPQKGAASFSTNQNNSAFVYGNVQLQEDKAPNMFLETKPNAESESIVDVAPKSIPEE	656
Asator-FD	619	KHVPAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANQIQRTGTVTNDKTSEVNRSTEE	677
Submitted_Seq	657	KFVDSNDAECANKAFLSGEGQQQKSQVKKLNLPESANQQVQKLENVANKESSSEDFRASQE	716
Asator-FD	678	QRSTFGRLRVLTAPPMSVHDLPSGGGSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG	737
Submitted_Seq	717	PKSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLSEIKQDF--SSSNAGFAAGNSSSSKLA	773
Asator-FD	738	-SQHGQIFGLAAMPPIINRRSATSTNLRPSSSSAQ-----RINSGSTIGGAVGNGSNTARS	791
Submitted_Seq	774	INQHGQIFGITLMPQVNRSATSTNLRPSSSGGNTNPIHRINIGSA-GGGGGTGSNTARS	832
Asator-FD	792	SVAGDHSVTQFALIDDENVSALQQVTKGGALTILASQWKSQFDDSEDITDNEWNREHCLOE	851
Submitted_Seq	833	SVAGDHSVTQFALIDDENVSALQQVTKGGALTILASQWKSQFDDSEDITDNEWNREHCLOE	892
Asator-FD	852	NLEQLIKLDISLPLNEAKFPFQHGAVGTGKLNPPGEAKGRPKRYTLNITGIENYEALRI	911
Submitted_Seq	893	NLEQLIKLDIPLPLNEAKHFCQNVVTDITGILTKPPIEGNEKPKRYTLNITGIENYEALRI	952
Asator-FD	912	SIPNCWSEPAMGNVLRKGLEPPAVQQAAPDDTVYRMDIARNVCVRETYSEIHL--ARPS	969
Submitted_Seq	953	SIPHCWSEPAMGNVLRKLDLEPPAVQQAAPDDTVYRMDIARNVCVRETYSDITPLDKAKPA	1012
Asator-FD	970	TSSVLRNRLPSPFKKDSALQLNSTINDSLDKSRHRNSLPNVSVNDIFDDQLMKLNL----	1025
Submitted_Seq	1013	TSLVSRVVLPSPFKEDATFQLNTSNNSQAKLKRRSLPNVSVADLFDDQPIHNSDAMLA	1072
Asator-FD	1026	-----LGSAIQENNCISGRLEIRVIPKDTSHPDSDSVYYDAMGAV	1065
Submitted_Seq	1073	EDANKPKVKQTMACQBSWVALSSAVQENNGCISGRLEIRVIPKETSHLSDSVYYDALAI	1132
Asator-FD	1066	KNTPTANEGRHSDQAVNNCEMEATSAVIAFPNKSISKIMSPPGRDATEERTGASLCSE	1125
Submitted_Seq	1133	KNTTTANPDGISDKANIFCDEIEENAAVIALPSNTANKCKKQTYTDKTEACIEANPCSI	1192
Asator-FD	1126	YSAGVNLKRLNG-----NTAP-----RTQFKKGSTDGFGENESEFDFPLLNPSKI	1170
Submitted_Seq	1193	HATGVNWLKINGKSETCNALNDQPNYKSGDYKIPFTGGSTDVREITDGGCDLPLLNPSKI	1252
Asator-FD	1171	PVROSKCASWAGADF-ISASKPLESAEVPQEIYPHQPQSDTTYSVIDSIPVRKITYSIALE	1229
Submitted_Seq	1253	FIRQSKCASWAGADTTVYSSKILEPRDALFEIPFNPQTN-----TYSTALE	1298
Asator-FD	1230	CPFNISDLTPGLRRRRETEGKYVTDQTLPLKQRPASRTSSRTRGIPNAMLNSDDNG	1289
Submitted_Seq	1299	YFPNITDLTPGLRRRRESAEGKYVTDQTLQQLKQRPASRTSSRTRGISNTMLGNFDDNN	1358
Asator-FD	1290	TVMSAEKTHG--HVVKTEGVKNFTTSENSSEVNPKCTISPPPGDKPIENSARLRRYRHN	1348
Submitted_Seq	1359	TVSGEKQRLIGTQVQNEESNNVFTSIYSAELQDKCNISFPFGDKPIENSARLRRYRHN	1418
Asator-FD	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

Configure Gene Model

Model Details

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

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

Verify Gene Model

Reset Form

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

Expand All

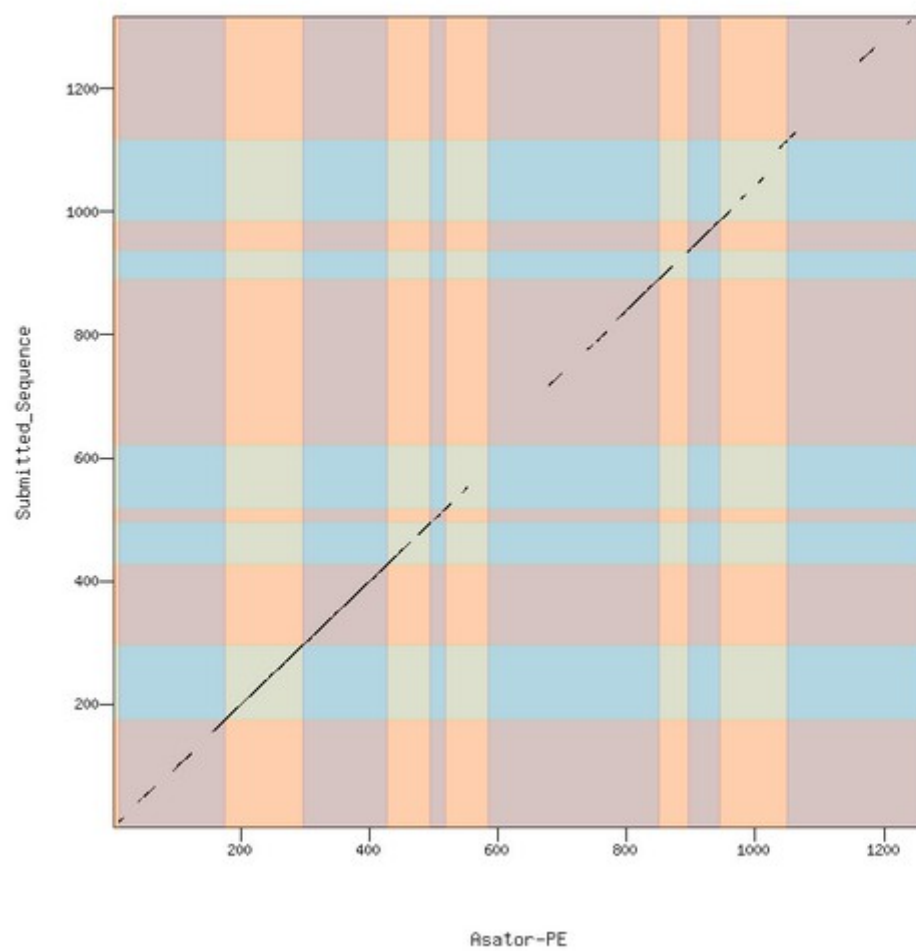
Collapse All

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	Gene model has 11 CDS's, ortholog has 12 CDS's



Asator-PE	1	MFVHLLCVFNDENASAPDDGNQS--CQPSKQDQYLSFNRNCQKMLLALYPFPPSKPPPL	58
Submitted_Seq	1	MQF-LLFFRNNDENASSPGDGNHNTCQFPNCQYISLNRDCQKMLRLHPPFPPSKPPPL	59
Asator-PE	59	VGAILQTRILLHQISPSAIDADADLNAVGEILLYPNVLQRSATLPKHNRLGVRSRVTFKV	118
Submitted_Seq	60	AGAILQSRLLKQISSGANAE---NAEALAEHRYPNALQRSATLPKHNRLGVRSRVTFKV	116
Asator-PE	119	PSSNLPAQDSYSHQF--RNQVAVAAKDGILVDVKAKESVVKMTSEDLLQPGHVVKERKRVV	176
Submitted_Seq	117	PSSITPAVDPGSEPDGQNVVAAERDGVILNPKERESCKMTSEDLLQPGHVVKERKRVV	176
Asator-PE	177	RKIGGGGFGIEYEGQDLITREQVALKVESARQPKQVLEMEVAVLKKLQKKEHVCRFICGG	236
Submitted_Seq	177	RKIGGGGFGIEYEGQDLITREQVALKVESARQPKQVLEMEVAVLKKLQKKEHVCRFICGG	236
Asator-PE	237	RNDRFNYVVMQLQGENLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGLHRDIKF	296
Submitted_Seq	237	RNDRFNYVVMQLQGENLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGLHRDIKF	296
Asator-PE	297	SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRVASINAHNRNRMG	356
Submitted_Seq	297	SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRVASINAHNRNRMG	356
Asator-PE	357	RHDDLWSLFYMLVEFVNGQLPWRKIKDKQVGLTKERYDHRILLKHLPSDLKQFLEHIQS	416
Submitted_Seq	357	RHDDLWSLFYMLVEFVNGQLPWRKIKDKQVGLTKERYDHRILLKHLPSDLKQFLEHIQS	416
Asator-PE	417	LYGDRPDYAMLIGLFCRCKRRGVKESDPYDWEKVDSTAIGNISATGNPSIPIKSDYMH	476
Submitted_Seq	417	LYADRPDYAMLIGLFCRCKRRGVKESDPYDWEKVDSTAIGNISTSGNPPVPVKNDYIH	476
Asator-PE	477	GNITQMTVAASNASGTEYIRKRAEIAETAHITATDPLNIKERVDKNCNATSLAQPAKSGGE	536
Submitted_Seq	477	GNITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKERVDKNCNATTLAQPKTSGE	536
Asator-PE	537	PMVQHGNAAANNQNTISKG-LQQQSILT-NSQVAIANIQSAPSM-----IEREDVQ-----	584
Submitted_Seq	537	ANVQLGCTANNQNTIPKGLQQQAALAINSQAAIPHMQSVPIKSPMVGMGSHDVQVHTKN	596
Asator-PE	585	-----YTKLEEGAPTKFITMKPNGECDN-VDIAAKCIFEQ	618
Submitted_Seq	597	SQPQKGAASFSTNQNNAPVYNGSYLQLEDKAPNMFLETKENAESESTVDVAPKSIREE	656
Asator-PE	619	KHVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQIQRTGTVTNDKTSVNRSTEE	677
Submitted_Seq	657	KEVDSNDAECANKAFLSGEGQQQKSQVKKLNLPEANQQVQKLENVANESSEDKRASQE	716
Asator-PE	678	QRSTFGRLRVLTAPPMSVHDLPSSGGGSHQVSDLSGKQDPYAATSNAAPIGINSSTKFG	737
Submitted_Seq	717	FKSTFGRLRVLTAPPMSVHDLTGG-HIQQTDLISIKQDP--SSSNAGFAAGNSSSSKLA	773
Asator-PE	738	-SQHGQIPGLAAMPPIINRSATSTNLPSSSASQ-----RINSGSTIGGAVGNGSNTARS	791
Submitted_Seq	774	INQHGQIPGITLMPQVNRASATSTNLPSSSGGNTNPIHRINIGSA-GGGGSGSNTARS	832
Asator-PE	792	SVAGDHSVTQFALIDDENVSALQQVTRGGALTASQWKSQFDDSEDITDNEWNREHCLOF	851
Submitted_Seq	833	SVAGDHSVTQFALIDDENVSALQQVTRGGALTASQWKSQFDDSEDITDNEWNREPCSQF	892
Asator-PE	852	NLEQLIKLDISLPLNEAKFPFQHGAVAGTGKLNPPGEAKGRPKRYTLNITGIEYEAALRI	911
Submitted_Seq	893	NLEQLIKLDIPLPLNEAKHFCQNVVDTGTILTKPIEGNEKPKRYTLNITGIEYEAALRI	952
Asator-PE	912	SIPNCWSEFAMGNVLRKLEPPAVQQAADFDTVYRMDIARNVCVRETYSEITHL--ARPS	969
Submitted_Seq	953	SIPHCWSEFAMGNVLRKDLFPFAVQQAADFDTVYRMDIARNVCVRETYSDITPLDKAKPA	1012
Asator-PE	970	TSSVLRNRLPSFFKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDLQMKLNLD----	1025
Submitted_Seq	1013	TSLVSRVLPSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHSNSDAMLA	1072
Asator-PE	1026	-----LGSATQENNCISGRLEIRVIPKDTSHPDSDSVYYDAMGAV	1065
Submitted_Seq	1073	EDANKFWKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKDTSHLDDSVYYDALAPI	1132
Asator-PE	1066	RNTPTANEGHDSQAVNNCDEMEATSAVIAFPNRSISKIMSPPGRDATEERTGASLCSL	1125
Submitted_Seq	1133	RNTTANPDGSDKANIFCDEIENAAVIALPNTANKCKKQTYTDKTEACIEANPCSI	1192
Asator-PE	1126	YSAGVNLKILNG-----NTAP-----RTQFRKGSTDGFGENESEDFPLLNPSKI	1170
Submitted_Seq	1193	HATGVNNLKINGKETCNALNDQPNYKSGDYKTFPTGGSTDSYRETDSGCDLPLLNPSKI	1252
Asator-PE	1171	PVRQSKCASWAGADF-ISASKPLESAEVPQEIYPHPQSDTTYSVIDSIPVRKTTYSIALE	1229
Submitted_Seq	1253	FIRQSKCASWAGADTTVYSKTLPRDALPEIPFPNQTN-----TYSTALE	1298
Asator-PE	1230	CPFNISDLTPGLSYFYCNIVVPLRLPSILLTES	1262
Submitted_Seq	1299	YFPNITDLTPGWSQLTCCI-----	1317

Dot plot of Asator-PE vs. Submitted\_Sequence



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

Configure Gene Model

Model Details

Fosmid Sequence File:

C:\fakepath\contig28.fasta

Browse...

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

Verify Gene Model

Reset Form

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

Expand All

Collapse All

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	
	Number of coding exons matched <i>D. melanogaster</i> ortholog	Pass	

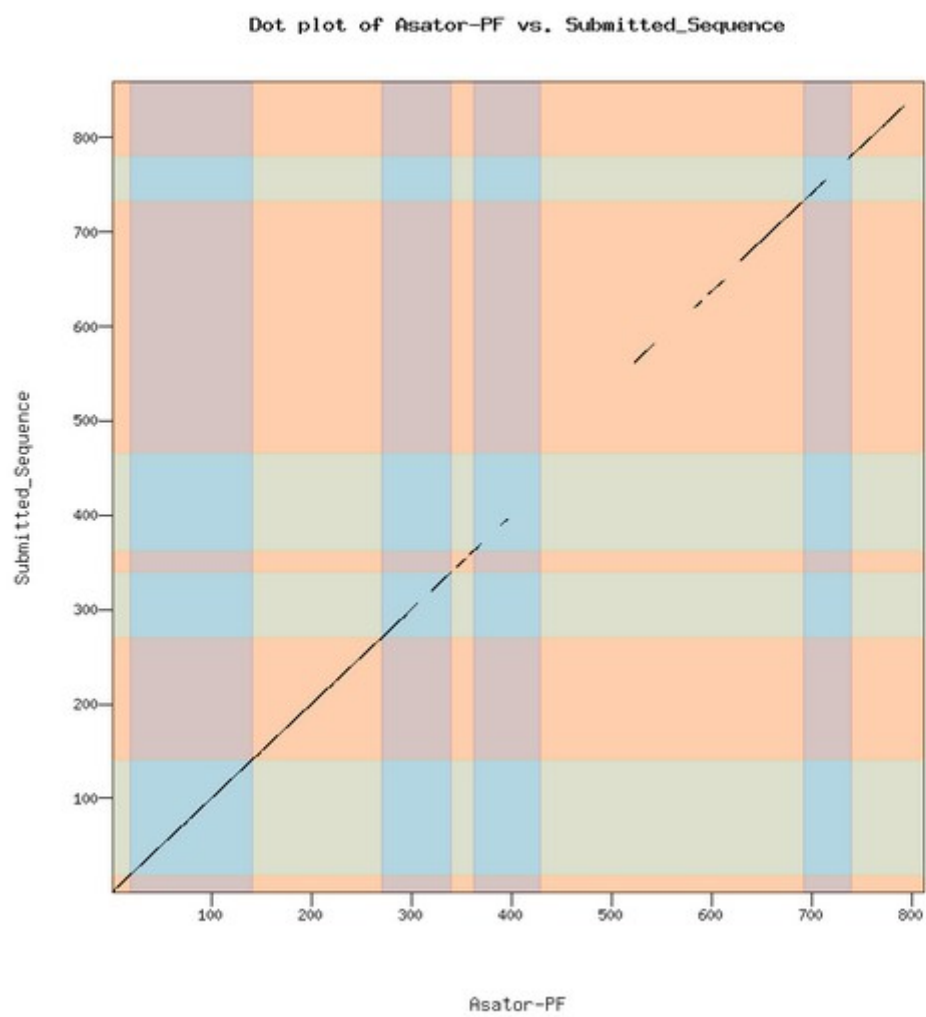


Identity: 651/863 (75.4%), Similarity: 701/863 (81.2%), Gaps: 56/863 ( 6.5%)

Asator-PF	1	MTSEDLLQPGHVVKERWVVRKIGGGSGEIEYEQDLITREQVALKVESARQPKQVLKMP	60
Submitted_Seq	1	MTSEDLLQPGHVVKERWVVRKIGGGSGEIEYEQDLITREQVALKVESARQPKQVLKMP	60
Asator-PF	61	VAVLKKLQSGKEHVCRFIGCGRNDRFNYYVVMQLQGNLAELRRAQPRGAFSLSTTLRLGLQ	120
Submitted_Seq	61	VAVLKKLQSGKEHVCRFIGCGRNDRFNYYVVMQLQGNLAELRRAQPRGAFSLSTTLRLGLQ	120
Asator-PF	121	ILKAIIESIHSVGFILHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA	180
Submitted_Seq	121	ILKAIIESIHSVGFILHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA	180
Asator-PF	181	GFRGTVRYASINAHNRNEMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH	240
Submitted_Seq	181	GFRGTVRYASINAHNRNEMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH	240
Asator-PF	241	RILLKHLPSDLKQFLEHIQSLTYGDRPDYAMLIGLPERCMKRRGVKESDPYDWEKVDSTA	300
Submitted_Seq	241	RILLKHLPSDLKQFLEHIQSLTYADRPDYAMLIGLPERCMKRRGVKESDPYDWEKVDSAT	300
Asator-PF	301	IGNISATGNPSIPIKSDYMEGNITQMTVAASNA3GTEYIRKRAEIETAHITATDPLNIKE	360
Submitted_Seq	301	IGNISTSGNPPVPVRNDYIHGNITQMTVAASNA3GTEYVRKRGDIETAHITATEPLHIKE	360
Asator-PF	361	RVDKNCNATSLAQPAKSGPEPMVOHGNAANNQNTSKG-LQQQSTLT-NSQVAIANIQSA	418
Submitted_Seq	361	RVDKNCNATTLAQPKTSGEANVOLGCTANNQNTIPKGMLOQQAAALAIN3QAAIPHMQSV	420
Asator-PF	419	PSM-----IEREDVQ-----YTKLEEGAPTKEFITM	443
Submitted_Seq	421	PIKSPMVGMGSHDVQVHTRNSQPQKGAASF3STNQNN3APVYGN3YLQLEDKAPNMFLPT	480
Asator-PF	444	KPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHYK3QIRKHN3SPEIANKQI	501
Submitted_Seq	481	KPNAE3ESTVDVAPK3IFEK3FVDSNDAECANKAF3SGEQ3QK3QVK3LN3LP3SANQ3QV	540
Asator-PF	502	QRTGIVTNDKISEVNRSTEEQKSTFGRLRVLTAPPMSVHDLPSGGGSH3QVSD3SGKQDE	561
Submitted_Seq	541	QKLENVANEK3SEDKRASQEPKSTFGRLRVLTAPPMSVHDLITGG-HIQQGTDL3IKQDE	599
Asator-PF	562	YAAT3NAAPIGIN33STKFG-3QHGGQIFGLAAMP3INRRSAT3TNLRP333ASQ-----R	615
Submitted_Seq	600	--333NAGPAAGN3333KLAINQHGGQIFGITLMPQVNR3SAT3TNLRP333GGNTNPIHR	657
Asator-PF	616	IN3GSTIGGAVNG3NTAR3SVAGDHSVTQFALIDDENVSALQQVTKGGALT3ASQWK3SQ	675
Submitted_Seq	658	INIG3A-GGGGGTGSNTAR3SVAGDHSVTQFALIDDENVSALQQVTKGGALT3ASQWK3SQ	716
Asator-PF	676	FDDSEDTTDNEWNREHQ3QPNLEQLIKLDISLPLNEAKPFPQHGVAGTGKLINPPGEARG	735
Submitted_Seq	717	FDDSEDTTDNEWNREHQ3QPNLEQLIKLDIPLPLNEAKHFCQNVVTDGTGILTKPPIEGNE	776
Asator-PF	736	RPKRYTILNITGIENYALRISIPNCWSE3PAMGNVLRKGLEPPAVQQA3FDDTV3SIL---N	792
Submitted_Seq	777	KPKRYTILNITGIENYALRISIPHCWSE3PAMGNVLRK3LEPPAVQQA3FDDTV3SILIKN	836
Asator-PF	793	K3LPV3IYIY----MHA3VTF3CF	811
Submitted_Seq	837	KESGRHFRCFN3IKSVNQ3LIVKNA	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

Gene Model Checker

Configure Gene Model

Model Details

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

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

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

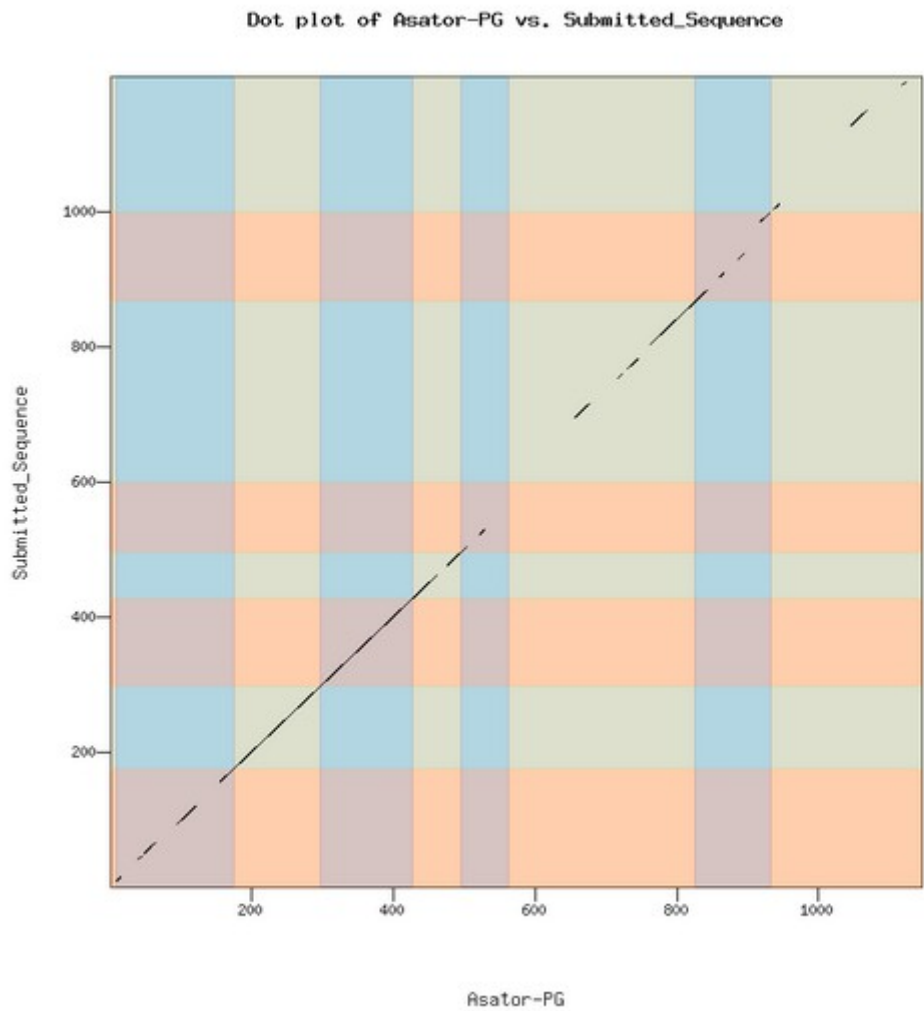
Expand All Collapse All

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

Identity: 823/1253 (65.7%), Similarity: 919/1253 (73.4%), Gaps: 12/1253 (1.0%)

Isator-PG	1	MFHLLCVFNDENASAPDDGNQS--CQSSKQDQYLSNRRNCQKNLLRLYPSPSKPPPI	58
Submitted_Seq	1	MQF-LLFFRNDENASSPGDGNHHTCQPPCNQEQYISLNRDCQKNLRLHPPSPSKPPPI	59
Isator-PG	59	VGAILQTRILLHQISPSAIAADADADLNAVCELLYPNVLQRSATLPKHNRLGVRSRVTFKV	118
Submitted_Seq	60	AGAILQSRLLKQISSGANAE---NAEALAEHRYFNALQRSATLPKHNRLGVRSRVTFKV	116
Isator-PG	119	PSSNLPAQDSYSHQF--RNQVAVAAKDGILVDVKAESVVKMTSEDLLQPGHVVKERWKVV	176
Submitted_Seq	117	PSSITPAVDPGSEFDPGQNVVAAERDGIIVLNPKERESCKMTSEDLLQPGHVVKERWKVV	176
Isator-PG	177	RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQKQKEHVCRFIGCG	236
Submitted_Seq	177	RKIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQKQKEHVCRFIGCG	236
Isator-PG	237	RNDRFNYYVMQLOGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFHHRDIKF	296
Submitted_Seq	237	RNDRFNYYVMQLOGKNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFHHRDIKF	296
Isator-PG	297	SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNRMG	356
Submitted_Seq	297	SNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNRMG	356
Isator-PG	357	RHDDLWSLFYMLVEFVNGQLPWRKIKDKQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS	416
Submitted_Seq	357	RHDDLWSLFYMLVEFVNGQLPWRKIKDKQVGLTKEKYDHRILLKHLPSDLKQFLEHIQS	416
Isator-PG	417	LYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAIGNISATGNPSIPIKSDYMH	476
Submitted_Seq	417	LYADRDPYAMLIGLFERCMKRRGVKESDPYDWEKVDSATIGNISTSGNPPVPVKNDYIH	476
Isator-PG	477	GNITQMTVAASNASGTEYVDKNCNATSLAQPAKSGEPMVQHGNAAANNQNTSKG-LQQQ	535
Submitted_Seq	477	GNITQMTVAASNASGTEYVDKNCNATTLAQPKTSGEANVQLGCTANNQNTPKGMLQQQ	536
Isator-PG	536	STLT-NSQVAIANIQSAPSM-----IEREDVQ-----	561
Submitted_Seq	537	AALAINSQAAIPHMQSVPKSPVMGMSHDVQVHTKNSQPKGAASFSTNQNSAPVYG	596
Isator-PG	562	--YTKLEEGAPTKFITMKPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHY	617
Submitted_Seq	597	NSYLQLEDKAPNMFLPTKPNASESTVDVAPKSIPEEKFVDSNDACANKAFLSGEQQQQ	656
Isator-PG	618	KSQIKKHNSPEIANKQIQRTGTVTNDKTSEVNRSTEEQKSTFGRLRVLTAPPMSVHDLPS	677
Submitted_Seq	657	KSQVKKLNLPESANQVQKLENVANEKSSDKRASQEPKSTFGRLRVLTAPPMSVHDLTT	716
Isator-PG	678	GGGHSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG-SQHGQIFGLAAMPFINRRSATS	736
Submitted_Seq	717	GG-HIQQGTDLISIKQDE---SSSNAGPAAGNSSSSKLAINQHGQIFGITLMPQVNRRSATS	773
Isator-PG	737	TNLRPSSSASQ-----RINSGSTIGGAVNGSNTARSSVAGDHSVTQFALIDDENVSALQ	791
Submitted_Seq	774	TNLRPSSSGGNTNP IHRINIGSA-GGGGGTGSNTARSSVAGDHSVTQFALIDDENVSALQ	832
Isator-PG	792	QVTKGGALTILASQWKSQFDDSEDITDNEWNREHQVYRMDIARNVCVRETYSEITHL--AR	849
Submitted_Seq	833	QVTKGGALTILASQWKSQFDDSEDITDNEWNREHQVYRMDIARNVCVRETYSDITFLDKAK	892
Isator-PG	850	PSTSSVLRLNRLSPFPKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDLQMKLNLD--	907
Submitted_Seq	893	PATSLVSRVVLPSFPKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHNSSDAM	952
Isator-PG	908	-----LGSATQENNCISGRLEIRVIPKDISHPDDSVYYDAMG	945
Submitted_Seq	953	LAEDANKFWKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKETSHLDDSVYYDALA	1012
Isator-PG	946	AVKNTPTANEGHDHSDQAVNNCDEMEATSAVIAFFPKSISKIMSPFGRDATEERTGASLC	1005
Submitted_Seq	1013	PIKNTITANPDGDISDKANIPDEIEENAAVIALPSNTANKCKKQTYTDKTEACIEANPC	1072
Isator-PG	1006	SLYSAGVNLKLNK-----NTAP-----RTQFKKGSTDGFGENESEFDFPLLNP	1050
Submitted_Seq	1073	SIHATGVNNLKINGKSETCNALNDQPNYKSGDYKTPFTGGSTDSYRETDGCDLPLLNP	1132
Isator-PG	1051	KIPVRQSKCASWAGADF-ISASKPLESAEVPQEIPIYHPQSDITYSVIDSIFVRKITYSIA	1109
Submitted_Seq	1133	KIPVRQSKCASWAGADTTVYSSKLTLEPRDALPEIPFPQTNT-----YSTA	1178
Isator-PG	1110	LECPNISDLTPGLSYFYCNIVVPLRLFSILLTES	1144
Submitted_Seq	1179	LEYPPNITDLTPGWSQLTCCI-----	1199

Dotplot



# 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

Gene Model Checker

Configure Gene Model

Model Details

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

Ortholog in D. melanogaster:  
Asator-PH

Coding Exon Coordinates:  
30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25394, 23001-21857, 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

Verify Gene Model

Reset Form

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

Expand All

Collapse All

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	

Configure Gene Model

Model Details

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

Ortholog in D. melanogaster:  
Asator-PH

Coding Exon Coordinates:  
30582-30529, 28535-28170, 28057-27668, 27605-27402, 26948-26880, 26822-26511, 26460-25657, 25433-25394, 23001-21857, 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

Verify Gene Model

Reset Form

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

Expand All

Collapse All

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	



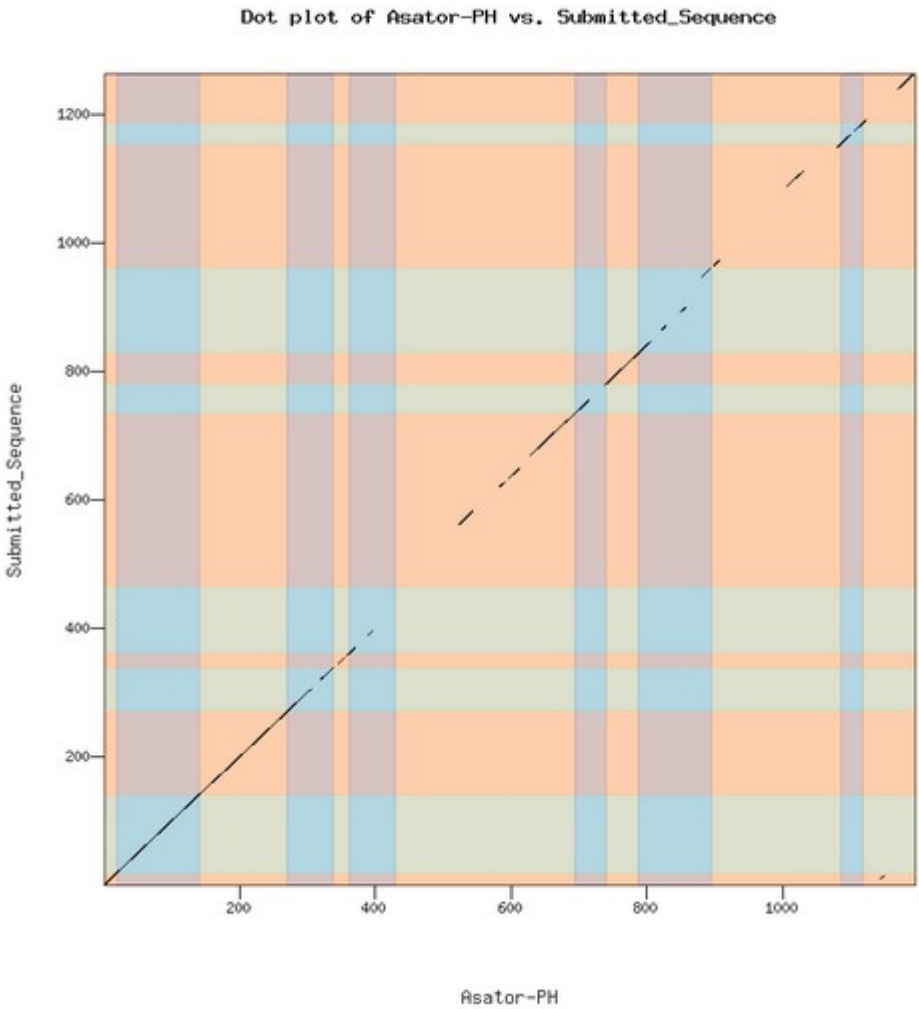
Identity: 894/1281 (69.8%), Similarity: 985/1281 (76.9%), Gaps: 106/1281 (8.3%)

Asator-PH	1	MTSEDLLQPGHVVKERWVVRKIGGGGFGIYEGQDLITREQVALKVESARQPKQVLKME	60
Submitted_Seq	1	MTSEDLLQPGHVVKERWVVRKIGGGGFGIYEGQDLITREQVALKVESARQPKQVLKME	60
Asator-PH	61	VAVLKKLQKKEHVCRFIGCGRNDRFNYVVMQLQSKNLAEPLRRAPRGAFSLSTTLRLGLQ	120
Submitted_Seq	61	VAVLKKLQKKEHVCRFIGCGRNDRFNYVVMQLQSKNLAEPLRRAPRGAFSLSTTLRLGLQ	120
Asator-PH	121	ILKAIESIHSVGLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCFRAAA	180
Submitted_Seq	121	ILKAIESIHSVGLHRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCFRAAA	180
Asator-PH	181	GFRGTVRYASINAHNRNEMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH	240
Submitted_Seq	181	GFRGTVRYASINAHNRNEMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKEKYDH	240
Asator-PH	241	RILLKHLPSDLKQFLEHIQSLTYGDRPDYA.MLIGLFCRMRKRGVKESDPYDWEKVDSTA	300
Submitted_Seq	241	RILLKHLPSDLKQFLEHIQSLTYGDRPDYA.MLIGLFCRMRKRGVKESDPYDWEKVDSTA	300
Asator-PH	301	IGNISATGNPSIPIKSDYMHGNIQTMTVAASNASGTEYIRKRAEIEHTAHITATDPLNIKE	360
Submitted_Seq	301	IGNISTSGNPPVPVKNYIHGNIQTMTVAASNASGTEYVRKRGDIETAHITATEPLHIKE	360
Asator-PH	361	FVDKNCNATSLAQPARSGEPMVQHGNAANNQNIISKGLQQQSTLT-NSQVAIANIQSA	418
Submitted_Seq	361	FVDKNCNATTLAQPKTSGEANTVQLGCTANNQNIIPKGM.LQQQAALAINQAAIPHMQSV	420
Asator-PH	419	PSM-----IEREDVQ-----YIKLEEGAPTKFITM	443
Submitted_Seq	421	PIKSPMVGMGSHDVQVHTKNSQPQKGAASFSTNQNSAPVYGNVYLQLEDKAPNMFLEPT	480
Asator-PH	444	KPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQI	501
Submitted_Seq	481	KPNAESESTVDVAFKSIFFEEKFVDSNDACANKAFLSGEQQQKSKQVKKLNLPESANQQV	540
Asator-PH	502	QRIGTVINDKISEVNRSTEEQKSTFGRLRVLTAPPMSVHDLPSGGGSHSHQVSDLSGKQDF	561
Submitted_Seq	541	QKLENVANERKSSSEDKRASQEPKSTFGRLRVLTAPPMSVHDLITGG-HIQQGTDLISIKQDF	599
Asator-PH	562	YAATSNAAPIGINSSSTKFG-SQHGQIFGLAAMPINRRSATSTNLRFSSSSASQ-----R	615
Submitted_Seq	600	--SSSNAGPAAGNSSSSKLAINQHQQIFGITLMPQVNRSATSTNLRFSSSGGNTNFIHR	657
Asator-PH	616	INSGSTIGGAVNGENTARSFVAGDHSVTQFALIDDENVSALQQVTRGGALTLASQWKSQ	675
Submitted_Seq	658	INIGSA-GGGGGTGSNTARSFVAGDHSVTQFALIDDENVSALQQVTRGGALTLASQWKSQ	716
Asator-PH	676	FDDSEDTTDEWNRHEQLQPNLEQLIKLDISPLNEAKPFPQHGVAAGTKLINPPGEARG	735
Submitted_Seq	717	FDDSEDTTDEWNRHEQLQPNLEQLIKLDIPLNEAKHFCQNVVTDGILTKPPIEGNE	776
Asator-PH	736	RPKRYTLNITGIENYEALRISIPNCWSEFAMGNVLRKLEPPAVQQAFFDDTVYRMDIAR	795
Submitted_Seq	777	KPKRYTLNITGIENYEALRISIPHCWSEFAMGNVLRKLEPPAVQQAFFDDTVYRMDIAR	836
Asator-PH	796	NVCVRETYSEITHL--ARPSTSSVLNRRLSPFKKDSALQLNSTNDSLDKSRHRNSLPNV	853
Submitted_Seq	837	NVCVRETYSDITPLDKAPATSLVSRVVLSPFKEDATFQLNTSNNSQAKLKHRRSLPNV	896
Asator-PH	854	SVNDIFDDLQMKLNL-----LGSALQENNNCCISGRLEIRV	889
Submitted_Seq	897	SVADLFDDQPIHNSDAMLAEADANKPWVQTMACQRSNVALSSAVQENNGCISGRLEIRV	956
Asator-PH	890	IPKDTSHPDSDSVYDAMGAVKNTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKI	949
Submitted_Seq	957	IPKETSHLDDSVYDALAPIKNTTIANPDGSDKANIFCDEIEENAAVIALPNTANKC	1016
Asator-PH	950	MSFPGRDATEERTGASLCSLYSAGVNLKLLNG-----NTAP-----RTQFKKGST	994
Submitted_Seq	1017	KKQTYTDKTEACIEANPCS IHATGVNNLRKINGKSETCNALNDQPNYKSGDYKTFPTGGST	1076
Asator-PH	995	DGFGENESEDFPLLNPISKIPVRQSKCASWAGADF-ISASKPLESAEVPQEIFYHPQSDT	1053
Submitted_Seq	1077	DSYRETDGCDLPLLNPISKIPVRQSKCASWAGADTIVYSSKILEPRDALPEIFPNPQNT	1136
Asator-PH	1054	TYSVIDSIPVRKITYSIALECFPNISDLTPGLRRRRETEGKYVTDQTLQPLKQRPISR	1113
Submitted_Seq	1137	-----YSTALEYFPNITDLTPGLRRRRESAEGKYVTDQTLQPLKQRPISR	1182
Asator-PH	1114	TSSRTRGIPNAMLENSDDNGIVMSAEKTHG-HVVKIEGVKNFTISENSSEVNPKCTISF	1172
Submitted_Seq	1183	TSSRTRGISNTMLGNFDDNNIVSGEKRLIGTQVQNEESNNVFTSIYSaelQRCNISF	1242
Asator-PH	1173	PPGDPKIENSARLRRYRHNID	1193
Submitted_Seq	1243	PPGDPKIENSARLRRYRHNLE	1263



Dotplot

[View protein alignment](#)



# 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

Gene Model Checker

Configure Gene Model

Model Details

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

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

Verify Gene Model Reset Form

Checklist

Dot Plot Transcript Sequence Peptide Sequence Extracted Coding Exons Downloads

Expand All Collapse All

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	

Gene Model Checker

Configure Gene Model

Model Details

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

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

Verify Gene Model Reset Form

Checklist

Dot Plot Transcript Sequence Peptide Sequence Extracted Coding Exons Downloads

Expand All Collapse All

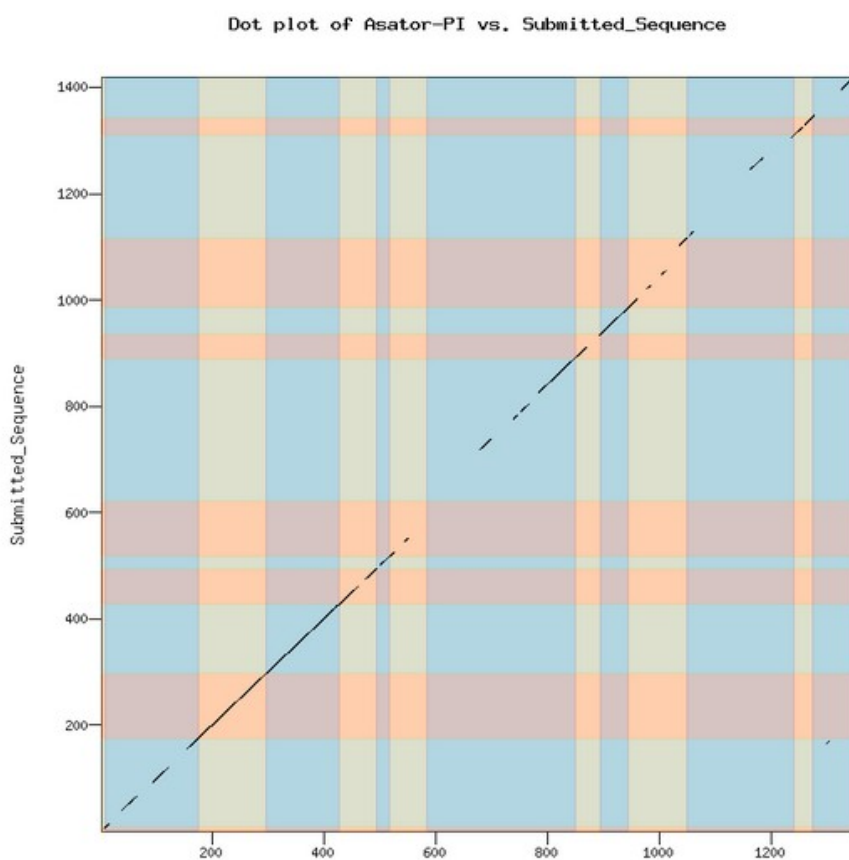
View	Criteria	Status	Message
	Acceptor for CDS 5	Pass	
	Donor for CDS 5	Pass	
	Acceptor for CDS 6	Pass	
	Donor for CDS 6	Pass	
	Acceptor for CDS 7	Pass	
	Donor for CDS 7	Pass	
	Acceptor for CDS 8	Pass	
	Donor for CDS 8	Pass	
	Acceptor for CDS 9	Pass	
	Donor for CDS 9	Pass	
	Acceptor for CDS 10	Pass	
	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 <i>D. melanogaster</i> ortholog	Pass	

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

Asator-FI	1	MHLQFYLRNDENASAPDDGNQS--CQPSKQDQYLSPNRCQKNLLRLYPSPSPKPPPLV	58
Submitted_Seq	1	MHLQDYLRNDENASAPDDGNHNTCQPPCQEQYISLNRDCQKNLFRLLHPPSPSPKPPPLA	60
Asator-FI	59	GAILQTRILLHQISPSAIDADADLNAGVCELLYFNVLRQSRATLPKHNRLGVRSRVTFKVF	118
Submitted_Seq	61	GAILQSRLLKQISSGANAE--NAEALCEHRYPNALQSRATLPKHNRLGVRSRVTFKVF	117
Asator-FI	119	SSNLPAQDSYSHQP--RNQVAAAKDGILVDVKAESVKMTSEDLLQPGHVVKERWVVR	176
Submitted_Seq	118	SSITPAVDGSGEPDQGNQVVAARDGIVLNPKERESCRMTSEDLLQPGHVVKERWVVR	177
Asator-FI	177	KIGGGGFGGEIYEGQDLITREQVALKVESARQPKQVLFMEVAVLKKLQKQKEHVCRFIGCGR	236
Submitted_Seq	178	KIGGGGFGGEIYEGQDLITREQVALKVESARQPKQVLFMEVAVLKKLQKQKEHVCRFIGCGR	237
Asator-FI	237	NDRFNYYVMQLQGRNLAEELRAQPRGAFSLSTTLRLGLQILKAIESIHSVGLFHRDIKFS	296
Submitted_Seq	238	NDRFNYYVMQLQGRNLAEELRAQPRGAFSLSTTLRLGLQILKAIESIHSVGLFHRDIKFS	297
Asator-FI	297	NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNRMGR	356
Submitted_Seq	298	NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNRMGR	357
Asator-FI	357	HDDLWSLFLYMLVEFVNGQLPWRKIKDKREQVGLTKEKYDHRILLKHLPSDLKQFLFHIQSTI	416
Submitted_Seq	358	HDDLWSLFLYMLVEFVNGQLPWRKIKDKREQVGLTKEKYDHRILLKHLPSDLKQFLFHIQSTI	417
Asator-FI	417	TYGDRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAGNISATGNPSIPIKSDVMHG	476
Submitted_Seq	418	TYADRPDYAMLIGLFERCMKRRGVKESDPYDWEKVDSTAGNISATGNPSIPIKSDVMHG	477
Asator-FI	477	NITQMTVAASNAGTIEYIRKRAPIETAHITATDPLNIKEKVDKNCNATSLAQPAKSGGEF	536
Submitted_Seq	478	NITQMTVAASNAGTIEYIRKRGDIETAHITATEPLHIKEKVDKNCNATSLAQPAKSGGEF	537
Asator-FI	537	MVQHGNAANNQNIITSGK-LQQOSTLT-NSQVAIANIQSAPSM-----IEREDVC-----	583
Submitted_Seq	538	NVQLGCTANNQNIITPKGMLQQQAALAINSQAALPHMQSVPIKSPMVGMGSHDVQVHTKNS	597
Asator-FI	584	-----YTKLEEGAPTKEITMKPNGECDN-VDIAAKCIFEQK	618
Submitted_Seq	598	QPKGAASFSTNQNNSAFVYGNSTYQLQEDKAPNMFPTKPNASESTVDVAPKSIPEEK	657
Asator-FI	619	HVEAND-DIVGRASLSGVEQHYKSQIKKHSPEIANKQIQRTGTVTNDKTSEVNRSTEEQ	677
Submitted_Seq	658	FVDSNDAECANKAFISGQQQQKQVKKLNLPEANQQVQKLENVANEKSSDEKRAQEF	717
Asator-FI	678	KSTFGRLRVLTAPPMSVHDLPSGGGSHSQVSDLSGKQDPYAATSNAAPIGINSSTKFGS	736
Submitted_Seq	718	KSTFGRLRVLTAPPMSVHDLTTGG-HIQQGTDLISIKQDP--SSNAGPAAGNSSSSKIAI	774
Asator-FI	737	SQHGQIFGLAAMPPIINRRSATSTNLRPSSASQ-----RINSGSTIGGAVGNGSNTARS	791
Submitted_Seq	775	NQHGQIFGITILMPQVNRSATSTNLRPSSSGGNTNP IHRINIGSA-GGGGTGNTARS	833
Asator-FI	792	VAGDHSVTQFALIDDENVSALQVTRGGALTLASQWKSQFDDSEDTTONEWNRHQ LQPN	851
Submitted_Seq	834	VAGDHSVTQFALIDDENVSALQVTRGGALTLASQWKSQFDDSEDTTONEWNRHQ LQPN	893
Asator-FI	852	LEQLIKLDISLPLNEAKFFQHGVAIGIKLINPPGEAKGRPKRYTLNITGIENYEALRIS	911
Submitted_Seq	894	LEQLIKLDIPLPLNEAKHFCQNVVTDIGILTKEFIEGNEKPKRYTLNITGIENYEALRIS	953
Asator-FI	912	IPNCWSEPAMGNVLRKGLPPAVQQAADFDTVYRMDIARNVCVRETYSEIHL--ARPST	969
Submitted_Seq	954	IPNCWSEPAMGNVLRKGLPPAVQQAADFDTVYRMDIARNVCVRETYSDITFLDKAKPAT	1013
Asator-FI	970	SSVLRNRLSPSPFKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDIFDDIQMKLNLD----	1024
Submitted_Seq	1014	SLVSRVVLSPSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQFIHNSNDAMLA	1073
Asator-FI	1025	-----LSSAIQENNCISGRLEIRVIPKDTSHPDSDSVYYDAMGAVR	1065
Submitted_Seq	1074	DANKPKVQTMACQRSNVALSSAVQENNGCISGRLEIRVIPKETSHLDSDSVYYDALAPIK	1133
Asator-FI	1066	NTPTANEHGHSDQAVNNCEMEATSAVIAFPNKSISKIMSPPGRDATEERTGASLCSLY	1125
Submitted_Seq	1134	NTTANPDPGISDKANIFCDEIENAAVIALPNTANKCKKQTYTDRTEACIEANPCSIH	1193
Asator-FI	1126	SAGVNLKILNG-----NTAP-----RTQFKRGSTDGFGENESEDFPLLNPSKIP	1170
Submitted_Seq	1194	ATGVNMLKINGKSETCNALNDQPNYKSGDYKTFPTGGSTDSYRETDSCDLPLLNPSKIF	1253
Asator-FI	1171	VRQSKCASWAGADF-ISASKPLESAEVQPIPYHPQSDITYSVIDSIPVRKTTYISIALEC	1229
Submitted_Seq	1254	IRQSKCASWAGADTTVYSSKTLPRDALPEIPFPNQTN-----TYSTALEY	1299
Asator-FI	1230	PPNISDLTPGLRRRRETEGKYVIDPQLQLKQFRFRSRTSSRTRGIPNAMLENSDDNNT	1289
Submitted_Seq	1300	PPNITDLTPGLRRRRESAEGKYVIDPQLQLKQFRFRSRTSSRTRGISNTMLGNFDDNNT	1359
Asator-FI	1290	VMSAETGEGH-HVVATEGVKNFTISENSSEVNPKCTISPPPGDPKIENSARLRRYRHNID	1348
Submitted_Seq	1360	VSGEKQRLIGTQVQNEESNNVPTSIYSAELQDKCNI SPFGDPKIENSARLRRYRHNLE	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

Gene Model Checker

Configure Gene Model

Model Details

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

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

Verify Gene Model

Reset Form

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

Expand All

Collapse All

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	

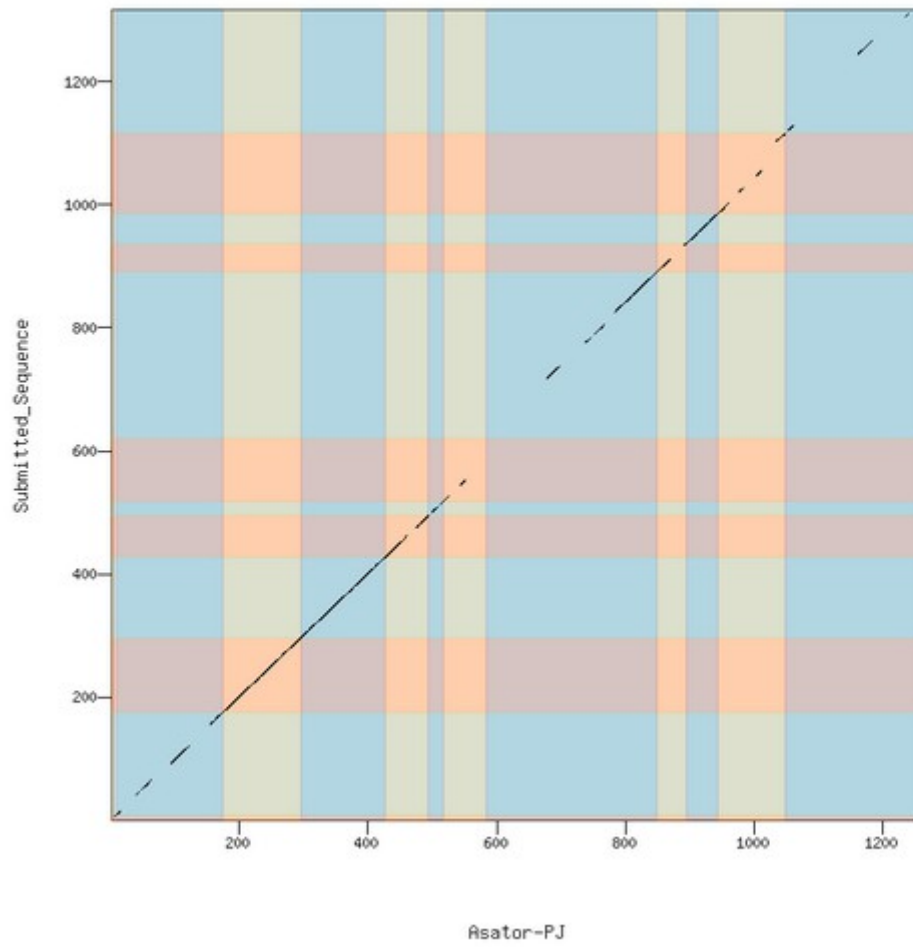
Version 1.1 (rev. 014) Hans Plaster



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

Asator-FJ	1	MHLQEYLRNDENASAPDDGNQS--CQPSKQDQYLSPNRNCQKNLLRLYPFPPSKPPPLV	58
Submitted_Seq	1	MHLQDYLRNDENASSPGDGNHMTCCPPCQEQYISLNRDCQKNLRLHPPPPSKPPPLA	60
Asator-FJ	59	GAILQTRLLHQISPSAIDADADLNAVGEILYPNVLQRSATLPKHNRLGVRSRVTFKVP	118
Submitted_Seq	61	GAILQSRLLKQISSGANAE--NAPALPEHRYPNALQRSATLPKHNRLGVRSRVTFKVP	117
Asator-FJ	119	SSNLPAQDSYSHQP--RMQVAAAKDGLVDVKAKESVKMTSEDLLQPGHVVKERKWKVVR	176
Submitted_Seq	118	SSITPAVDPGSEPDGPQNVVAAERDGLVLPKERESECKMTSEDLLQPGHVVKERKWKVVR	177
Asator-FJ	177	KIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQKEHVCRFIGCGR	236
Submitted_Seq	178	KIGGGGFGEIYEGQDLITREQVALKVESARQPKQVLKMEVAVLKKLQKEHVCRFIGCGR	237
Asator-FJ	237	NDRFNYVVMQLQGNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFHLRDIKFS	296
Submitted_Seq	238	NDRFNYVVMQLQGNLAELRRAQPRGAFSLSTTLRLGLQILKAIESIHSVGFHLRDIKFS	297
Asator-FJ	297	NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNREMGR	356
Submitted_Seq	298	NFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAAGFRGTVRYASINAHNRNREMGR	357
Asator-FJ	357	HDDLWSLFYMLVEFVNGQLPWRKIKDKQVGLTREKYDHRILLKHLPSDLKQFLEHIQSL	416
Submitted_Seq	358	HDDLWSLFYMLVEFVNGQLPWRKIKDKQVGLTREKYDHRILLKHLPSDLKQFLEHIQSL	417
Asator-FJ	417	TYGDRPDYAMLIGLFCRMRKRGVKSDFYDWEKVDSTAGNISATGNFSPIKSDYMHG	476
Submitted_Seq	418	TYADRPDYAMLIGLFCRMRKRGVKSDFYDWEKVDSTAGNISATGNFSPIKSDYMHG	477
Asator-FJ	477	NITQMTVAASNASGTEYIRKRAEIEHTAHITATDPLNIKEKVDKNCNATSLAQPAKGSSEF	536
Submitted_Seq	478	NITQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKEKVDKNCNATSLAQPAKGSSEF	537
Asator-FJ	537	MVQHGNAAANNQNTSKG-LQQQSTLT-NSQVAIANIQSAPSM-----IEREDVQ-----	583
Submitted_Seq	538	NVQLGCTANNQNTIPKMLQQQAALAINSQAAPHMQSVPIKSPMVGMSHDVQVHTKNS	597
Asator-FJ	584	-----YTKLEEGAPTKEFITMKPNGECDN-VDIAAKCIFEQK	618
Submitted_Seq	598	QFQKGAASFSTNQNSAPVYGNSTYQLEDKAPNMFPTKPNASESTVDVAPKSIPEEK	657
Asator-FJ	619	FVEAND-DIVGRASLSGVEQHYKSQIKKHNSPEIANKQIORTGIVINDKTSEVNRSTEEQ	677
Submitted_Seq	658	FVDSNDACANKAFLSGEGQQQKQSVKKNLNLPESANQQVQKLENVANEEKSSEDKRASQEF	717
Asator-FJ	678	KSTFGRLRLVTAPPMSVHDLPSGGGSHQVSDLSGKQDPYAATSNAAPIGINSSSTKFG-	736
Submitted_Seq	718	KSTFGRLRLVTAPPMSVHDLTIGG-HIQQGTDLISIKQDP--SSNAGPAAGNSSSSKLAI	774
Asator-FJ	737	SOHQGQIFGLAAMPFINRRSATSTNLRPSSSAQ-----RINSGSTIGGAVNGSNTARSS	791
Submitted_Seq	775	NOHQGQIFGITLMPQVNRSATSTNLRPSSSGGNTNPIHRINIGSA-GGGGGTGNTARSS	833
Asator-FJ	792	VAGDHSVTQFALIDDENVSALQQVTKGGALTASQWKSQFDDSEDTTDEWNRHQLOPN	851
Submitted_Seq	834	VAGDHSVTQFALIDDENVSALQQVTKGGALTASQWKSQFDDSEDTTDEWNRHQLOPN	893
Asator-FJ	852	LEQLIKLDISLFLNEAKFPFQHGVAAGTKLINPFGAERKPRKRYTLNITGIENYEALRIS	911
Submitted_Seq	894	LEQLIKLDIFLFLNEAKHFCQNVVIDTGILTKPFIEGNEKPRKRYTLNITGIENYEALRIS	953
Asator-FJ	912	IPNCWSEPMAGNVLRKGLEPPAVQQAADFDTVYRMDIARNVCVRETYSEITHL--ARPST	969
Submitted_Seq	954	IPHCWSEPMAGNVLRKLEPPAVQQAADFDTVYRMDIARNVCVRETYSDITPLDKAKPAT	1013
Asator-FJ	970	SSVLRNRLSPFKKDSALQLNSTNDSLDKSRHRNSLPNVSVNDFDDLQMKNLND-----	1024
Submitted_Seq	1014	SLVSRVVLSPFKEDATFQLNTSNNSQAKLKHRRSLPNVSVADLFDDQPIHNSNDAMLAE	1073
Asator-FJ	1025	-----LGSALQENNCISGRLEIRVIPKDTSHPDSDSVYYDAMGAVK	1065
Submitted_Seq	1074	DANKPFWQTMACQSRNVALSSAVQENNGCISGRLEIRVIPKDTSHLDSDSVYYDALAPIK	1133
Asator-FJ	1066	NTPTANEGHDSQAVNNCDEMEATSAVIAFPNKSISKIMSPPGRDATEREATGSLCSLY	1125
Submitted_Seq	1134	NTTTANPDGISDKANIPCEIEENAAVIALPSNTANKCKQTYYDRTEACIEANPCSIH	1193
Asator-FJ	1126	SAGVNRLKING-----NTAP-----RTQFKKSGTDGFGENESEFDFLLNPSKIP	1170
Submitted_Seq	1194	ATGVNRLKINGSETCNALNDQPNYKSGDYKTPFTGGSTDYRETDSCDLPNPSKIP	1253
Asator-FJ	1171	VRQSKCASWAGADF-ISASKPLESAEVPQEIPIYHPQSDTYSVIDSIPVRKTTYSIALEC	1229
Submitted_Seq	1254	IRQSKCASWAGADITVYSSKILEPRDALPEIPFNEQTN-----TYSTALEY	1299
Asator-FJ	1230	PPNISDLTPGLSYFYCNIVVPLRLFSILLTES	1261
Submitted_Seq	1300	PPNITDLTPGWSQLTCCI-----	1317

Dot plot of Asator-PJ vs. Submitted\_Sequence



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

Gene Model Checker

Configure Gene Model

Model Details

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

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

Verify Gene Model

Reset Form

Checklist

Dot Plot

Transcript Sequence

Peptide Sequence

Extracted Coding Exons

Downloads

Expand All

Collapse All

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](#)

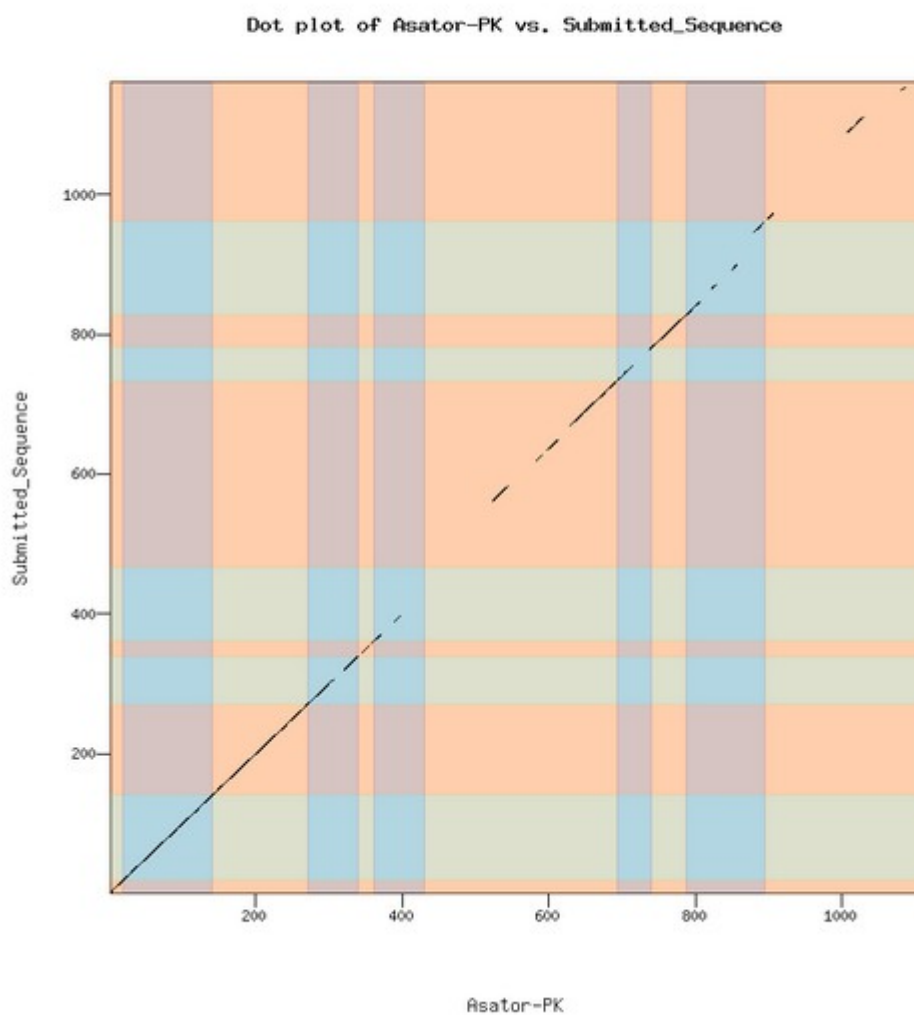
Identity: 821/1193 (68.8%), Similarity: 905/1193 (75.9%), Gaps: 119/1193 (10.0%)

Asator-PK	1	MTSEDLLQPGHVVKERWVVRKIGGGGPGIIEGQDLITREQVALKVESARQPKQVLKME	60
Submitted_Seq	1	MTSEDLLQPGHVVKERWVVRKIGGGGPGIIEGQDLITREQVALKVESARQPKQVLKME	60
Asator-PK	61	VAVLKKLQKKEHVCRFIGCGRNDRFNYVVMQLQKKNLAELARAQPRGAFSLSTTLRLGLQ	120
Submitted_Seq	61	VAVLKKLQKKEHVCRFIGCGRNDRFNYVVMQLQKKNLAELARAQPRGAFSLSTTLRLGLQ	120
Asator-PK	121	ILKAIESIHSVGFHLRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA	180
Submitted_Seq	121	ILKAIESIHSVGFHLRDIKPSNFSVGRLPYNCRRVYMLDFGLARQYTTGTGEVRCPRAAA	180
Asator-PK	181	GFRGTVRYASINAHNRNEMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKREKYDH	240
Submitted_Seq	181	GFRGTVRYASINAHNRNEMGRHDDLWSLFYMLVEFVNGQLPWRKIKDKEQVGLTKREKYDH	240
Asator-PK	241	RILLKHLPSDLKQFLEHIQSLTYGDRPDYAMLIGLPERCMKRRGVKESDPYDWEKVDSTA	300
Submitted_Seq	241	RILLKHLPSDLKQFLEHIQSLTYADRPDYAMLIGLPERCMKRRGVKESDPYDWEKVDSAT	300
Asator-PK	301	IGNISATGNPSIPKSDYMHGNTQMTVAASNASGTEYIRKRAEIEAHITATDPLNIKE	360
Submitted_Seq	301	IGNISTSGNPPVFPVKNDYIHGNTQMTVAASNASGTEYVRKRGDIETAHITATEPLHIKE	360
Asator-PK	361	KVDKNCNATSLAQPAKSGEPMVQHGNAAANNQMITSKG-LQQQSTLT-NSQVAIANIQSA	418
Submitted_Seq	361	KVDKNCNATTAPQPKTSGEANVQLGCTANNQMITPKGMLQQQAALAINSQAAIPHMQSV	420
Asator-PK	419	PSM-----IEREDVQ-----YTRKEEGAPTKFITM	443
Submitted_Seq	421	PIKSPMVGMGSHDVQVHTKNSQPQKGAASFSTNQNSAPVYGNQYLQLEDKAPNMFLEP	480
Asator-PK	444	KPNGECDN-VDIAAKCIFEQKHVEAND-DIVGRASLSGVEQHYK3QIKKHNSPEIANKQI	501
Submitted_Seq	481	KPNAESESTVDVAPKSIPEEKFVDSNDAECANKAFLSGEGQQQK3QVKKLNLPESANQQV	540
Asator-PK	502	QRIGTVINDKITEVNRSTEEQKSTFGRLRLVLTAPFMSVHDLPSGGGHSHQVSDLSGKQDF	561
Submitted_Seq	541	QKLENVANEKSSSEDKRASQEPKSTFGRLRLVLTAPFMSVHDLTTGG-HIQGGTDL3IKQDF	599
Asator-PK	562	YAATSNAAPIGINSSSIKFG-SQHGQIFGLAAMPINRRSATSTNLRPSSSASQ-----R	615
Submitted_Seq	600	--SSSNAGPAAGNSSSKLAINQHGGQIFGITLMPQVNRSATSTNLRPSSSGGNTNPIHR	657
Asator-PK	616	INSGSTIGGAVNGSNTARSSVAGDHSVITQFALIDDENVSALQQVTRGGALTLASQWK3Q	675
Submitted_Seq	658	INIGSA-GGGGGTGSNTARSSVAGDHSVITQFALIDDENVSALQQVTRGGALTLASQWK3Q	716
Asator-PK	676	FDDSEDTIDNEWNREHQIQPMLEQLIKLDISLPLNEAKPFPQHGVAAGTKLINPPGEAKG	735
Submitted_Seq	717	FDDSEDTIDNEWNREFQSQPMLEQLIKLDIPLPLNEAKHFCQNVVTDGILTKPPIEGNE	776
Asator-PK	736	RPKRVTLNITGIENYEALRISIPNCWSEFAMGNVLRKGLPEFAVQQAFFDDTVYRMDIAR	795
Submitted_Seq	777	RPKRVTLNITGIENYEALRISIPNCWSEFAMGNVLRKGLPEFAVQQAFFDDTVYRMDIAR	836
Asator-PK	796	MVCVRETYSEITHL--ARPSTSSVLNRRLSPFKKDSALQLNSTNDSLDKSRHRNSLPNV	853
Submitted_Seq	837	MVCVRETYSDITPLDKAKPATSLVSRVVLSPFKEDATFQLNTSNNSQAKLKHRRSLPNV	896
Asator-PK	854	SVNDFDQLQMKLND-----LGAIQENNCCTISGRLEIRV	889
Submitted_Seq	897	SVADLFDDQPIH3NSDAMLAEADANKPWVQTMACQ3NVALSSAVQENNGCTISGRLEIRV	956
Asator-PK	890	IPKDTSHPDSSVYYDAMGAVKNTPTANEGHDHSDQAVNNCDEMEATSAVIAFPNKSISKI	949
Submitted_Seq	957	IPKDTSHLDDSVYYDALAPIKNTTTANPDG3ISDKANIFCDEIEENAAVIALPSNTANKC	1016
Asator-PK	950	MSPPGRDATEERTIGASLCSLYSAGVNLKLN-----NTAP-----RTQFKKGST	994
Submitted_Seq	1017	KKQTYTDKTEACIEANPCSIHATGVNNLKINGKSETCNALNDQPNYKSGDYKTFPTGGST	1076
Asator-PK	995	DGFGENESEFDFPLLNPSKIPVRQSKCASWAGADF-ISASKPLESAEVPQEIPIYHPQSDT	1053
Submitted_Seq	1077	DSYRETD3GCDLPLLNPSKIPVIRQSKCASWAGADTTVY3SKTLEPRDALPEIFPNPQTNT	1136
Asator-PK	1054	TYSVIDSIPVRKITYSIALECPFNISDLTGLSYFYCNIVVPLRLFSILLTES	1106
Submitted_Seq	1137	-----YSTALEYFPNITDLTPGWSQLTCCI-----	1161



## Dotplot

[View protein alignment](#)

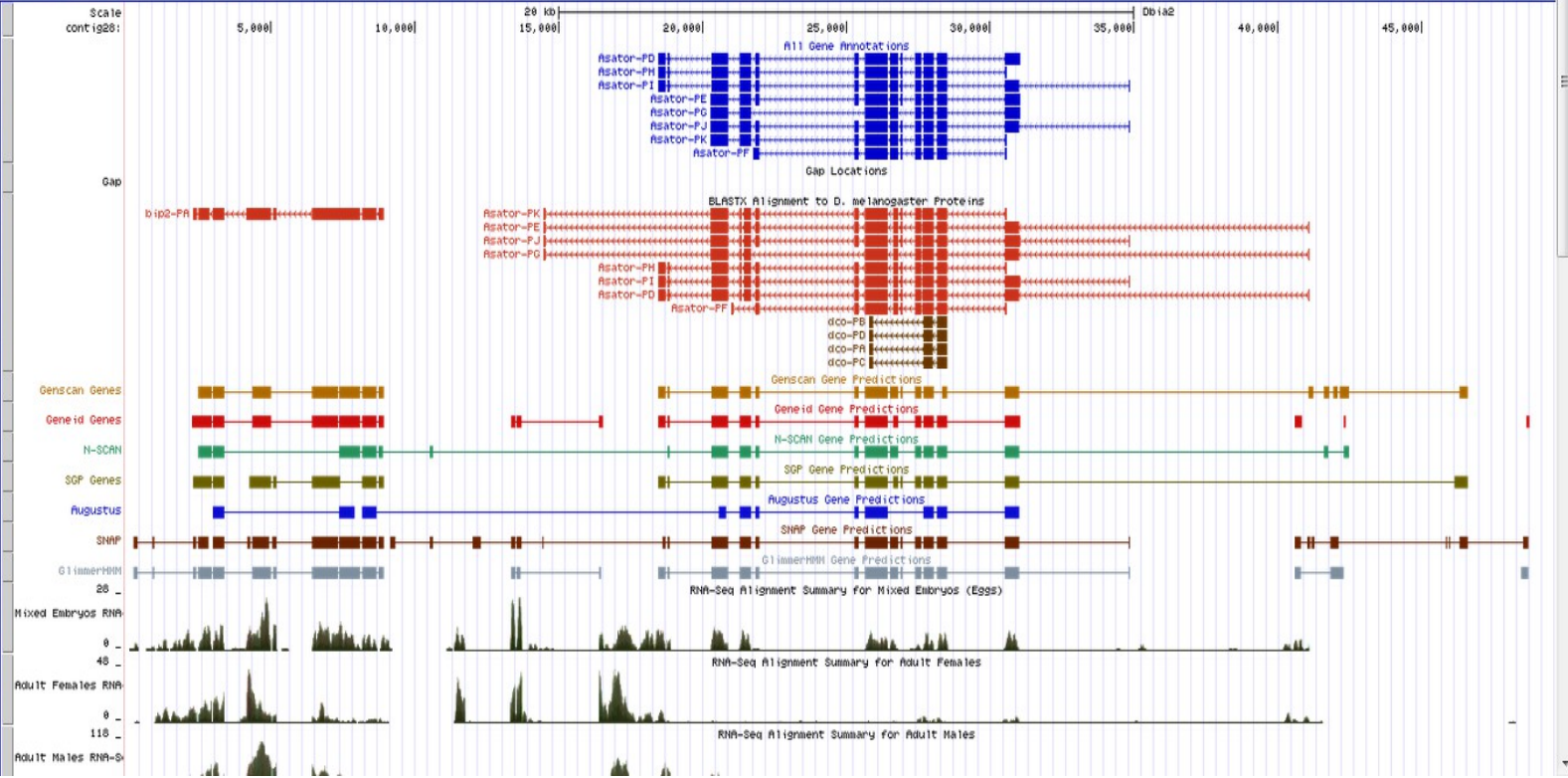


Custom track(s) viewed on the UCSC Genome Browser

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

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x

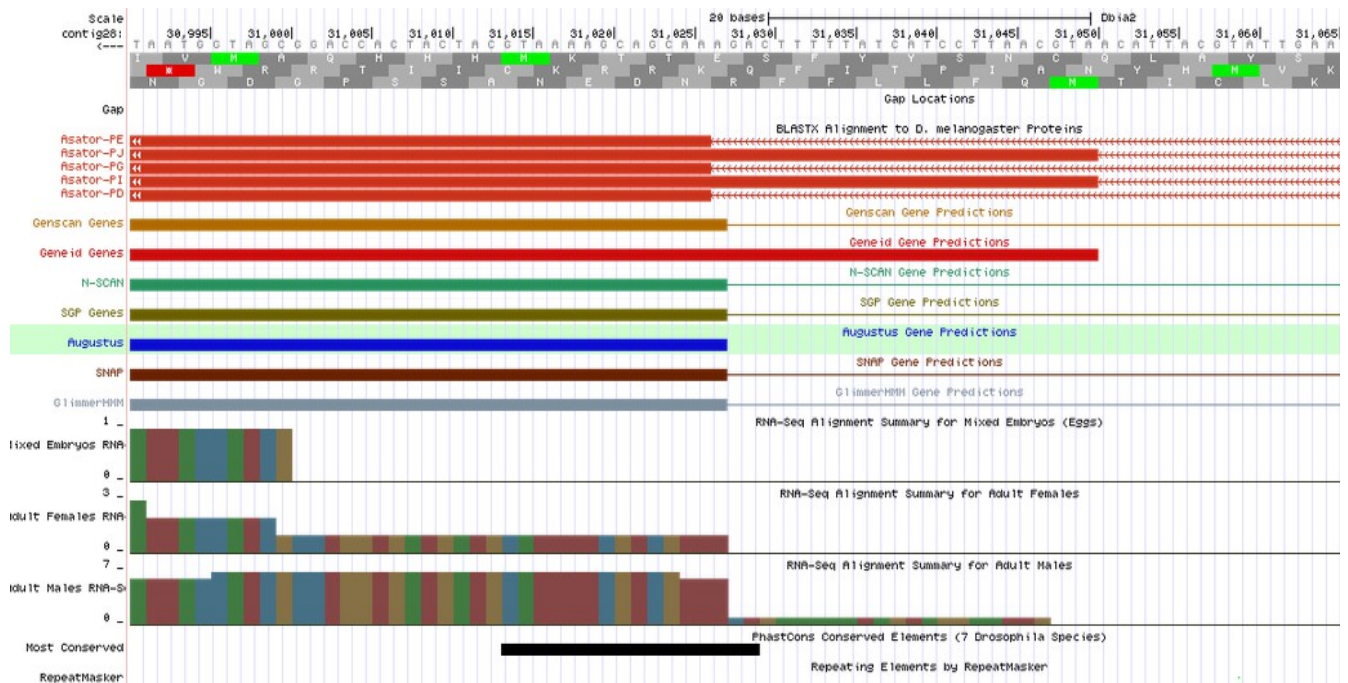
position/search contig28:1-50,000 jump clear size 50,000 bp. configure



**\*\*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 tcaaaatgatctatatatttttcgtattatagacaactgatgttttggcatt 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 tactatgcgttcgcgaagtgcgtctgaagatgcgttcaaatggacagaca 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 tggattttatttttttagaaacgacgaaaatgcatccgcgccccgacgat 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 taaccgcaattgccagaaaaacctgttgaggctttaccaccaccaccccat 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 ctaaaccaccacgcgtcgttggtgccatcctgcaaaccgctcttctgcat 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 aggggagcttctataccctaatgtcctgcagcggtcagcaactttgccag 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 CGAAACCACCGCGCTCGCGGGTGCTATCCTGCAATCCCGTCTTTTAAAG 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

```