GEP Annotation Report

Student name: Valerie Grieggs

Student email: valeriegmgrieggs@gmail.com

Faculty Advisor: Paul Pzauter

College/ University: University of New Mexico

Project Details

Project name: dbiarmipes_dot_Aug2012_contig52

Project Species: *Drosophila biarmipes*

Date of Submission: July 9, 2013

Size of project in base pairs: 40,000 bp

Number of genes in project: 2

Does this report cover all genes and all isoforms or is it a partial report? Full report

Gene Report Form

Gene name: metabotropic glutamate receptor

Gene Symbol: mGluRA

Approximate location in project: 6893-15719

Number of isoforms in *D. Melongaster:*3

Number of isoforms In this project: 3

Name of unique isoform based on coding sequence	List of isoforms with identical coding sequences
mGluRA-PB	mGluRA-PA
mGluRA-PC	

Isoform Report Form

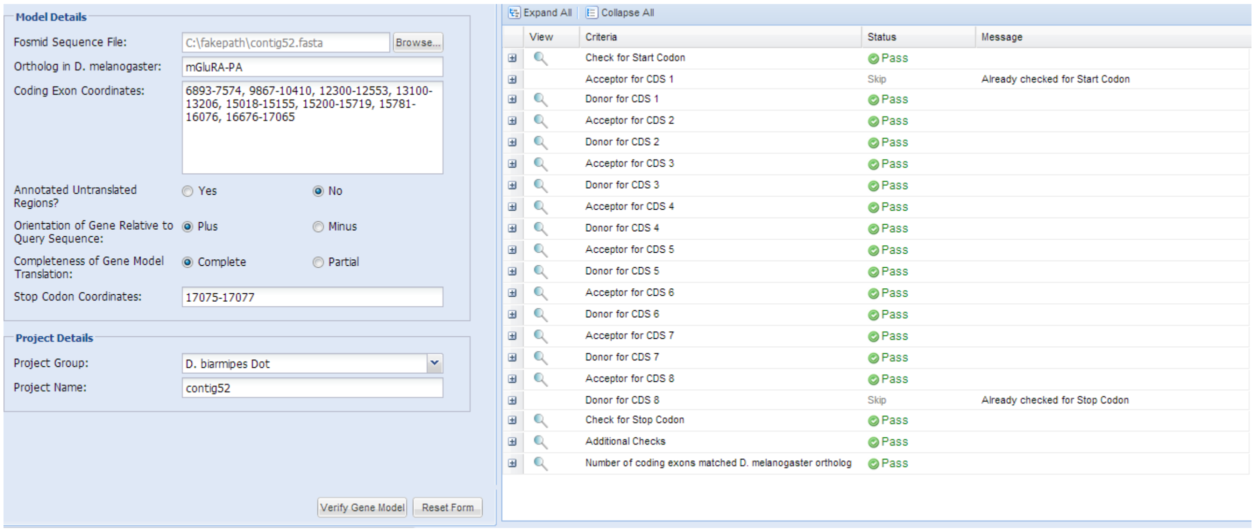
Gene-isoform name: mGluRA-PA

Names of the isoforms with identical coding sequences as this isoform: mGluRA-PB

Is the 5’ end of this isoform missing from the end of the project: No

Is the 3’ end of this isoform missing from the end of the project: No

Gene Model Checker Checklist



Model Details

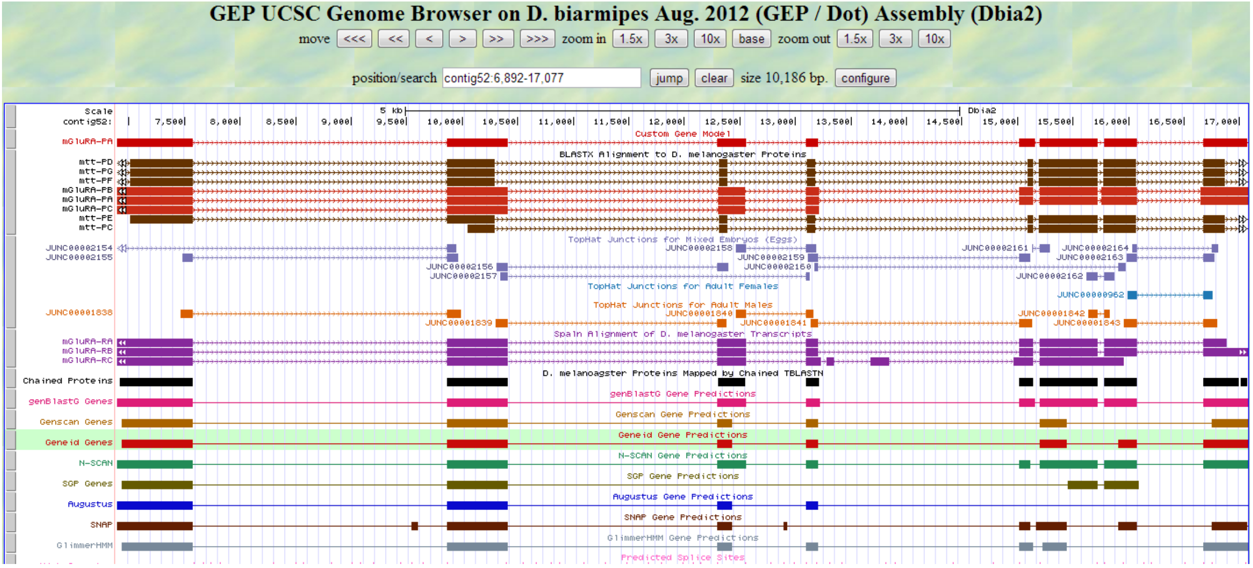
Fosmid Sequence File:	C:\fakepath\contig52.fasta	
Ortholog in D. melanogaster:	mGluRA-PA	
Coding Exon Coordinates:	6893-7574, 9867-10410, 12300-12553, 13100-13206, 15018-15155, 15200-15719, 15781-16076, 16676-17065	
Annotated Untranslated Regions?	○ Yes	● No
Orientation of Gene Relative to Query Sequence:	● Plus	○ Minus
Completeness of Gene Model Translation:	● Complete	○ Partial
Stop Codon Coordinates:	17075-17077	

Project Details

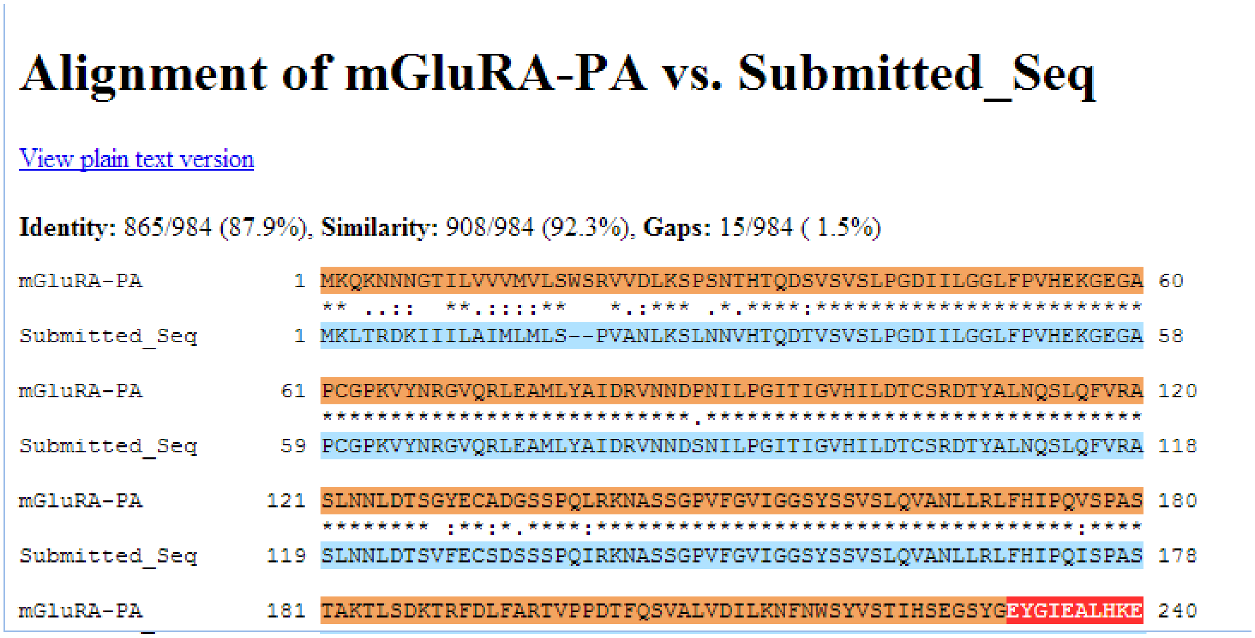
Project Group:	D. biarmipes Dot
Project Name:	contig52

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	✓ Pass	

Gene model on Genome Browser



Alignment between the submitted model and the *D. Melanogaster*



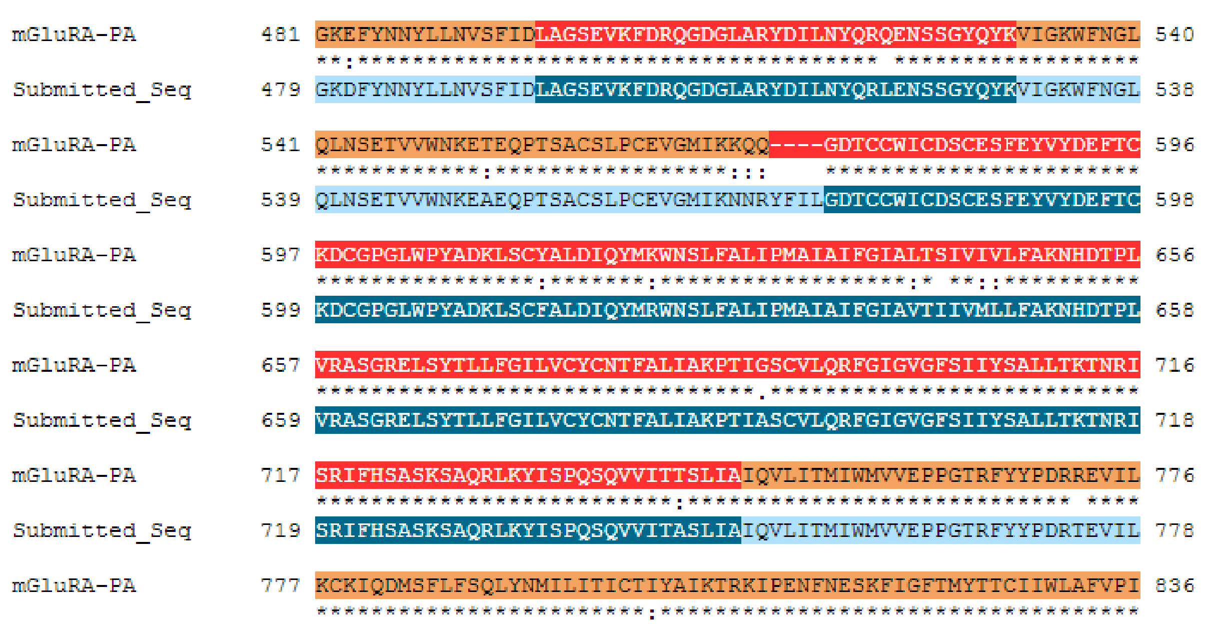
Alignment of mGluRA-PA vs. Submitted_Seq

View plain text version

Identity: 865/984 (87.9%), **Similarity:** 908/984 (92.3%), **Gaps:** 15/984 (1.5%)

```
mGluRA-PA           1 MKQKNNNGTILVVVMVLSWSRVVDLKSPSNTHTQDSVSVSLPGDIILGGLFPVHEKGEGA 60
                      ** ..::  **.::::**   *.:*** .*.****:************************
Submitted_Seq       1 MKLTRDKIIILAIMLMLS--PVANLKSLNNVHTQDTVSVSLPGDIILGGLFPVHEKGEGA 58

mGluRA-PA          61 PCGPKVYNRGVQRLEAMLYAIDRVNNDPNILPGITIGVHILDTCSRDTYALNQSLQFVRA 120
                      ***************************.********************************
Submitted_Seq      59 PCGPKVYNRGVQRLEAMLYAIDRVNNDSNILPGITIGVHILDTCSRDTYALNQSLQFVRA 118

mGluRA-PA         121 SLNNLDTSGYECADGSSPQLRKNASSGPVFGVIGGSYSSVSLQVANLLRLFHIPQVSPAS 180
                      ******** :**:*.****:***********************************:****
Submitted_Seq     119 SLNNLDTSVFECSDSSSPQIRKNASSGPVFGVIGGSYSSVSLQVANLLRLFHIPQISPAS 178

mGluRA-PA         181 TAKTLSDKTRFDLFARTVPPDTFQSVALVDILKNFNWSYVSTIHSEGSYGEYGIEALHKE 240
```

```
mGluRA-PA         181 TAKTLSDKTRFDLFARTVPPDTFQSVALVDILKNFNWSYVSTIHSEGSYGEYGIEALHKE 240
                      ********:***********************************************:***
Submitted_Seq     179 TAKTLSDKSRFDLFARTVPPDTFQSVALVDILKNFNWSYVSTIHSEGSYGEYGIEAFHKE 238

mGluRA-PA         241 ATERNVCIAVAEKVPSAADDKVFDSIISKLQKKPNARGVVLFTRAEDARRILQAAKRANL 300
                      ****:******************* **.********************************
Submitted_Seq     239 ATERHVCIAVAEKVPSAADDKVFDLIIGKLQKKPNARGVVLFTRAEDARRILQAAKRANL 298

mGluRA-PA         301 SQPFHWIASDGWGKQQKLLEGLEDIAEGAITVELQSEIIADFDRYMMQLTPETNQRNPWF 360
                      *************************************** ***:****************
Submitted_Seq     299 SQPFHWIASDGWGKQQKLLEGLEDIAEGAITVELQSEIIEDFDHYMMQLTPETNQRNPWF 358

mGluRA-PA         361 AEYWEDTFNCVLTSLSVKPDTSNSANSTDNKIGVKAKTECDDSYRLSEKVGYEQESKTQF 420
                      **********:* .:* :.::..* :**: **..*:** *:**:****************
Submitted_Seq     359 AEYWEDTFNCILEPVSDQTNSPTSIDSTEIKIATKSKTTCEDSFRLSEKVGYEQESKTQF 418

mGluRA-PA         421 VVDAVYAFAYALHNLHNDRCNTQSDQTTETRKHLQSESVWYRKISTDTKSQACPDMANYD 480
                      ***************************:* *** :********:****************
Submitted_Seq     419 VVDAVYAFAYALHNLHNDRCNTQSDQTSEQRKHHHSESVWYRKLSTDTKSQACPDMANYD 478

mGluRA-PA         481 GKEFYNNYLLNVSFIDLAGSEVKFDRQGDGLARYDILNYQRQENSSGYQYKVIGKWFNGL 540
                      **:************************************** ******************
```

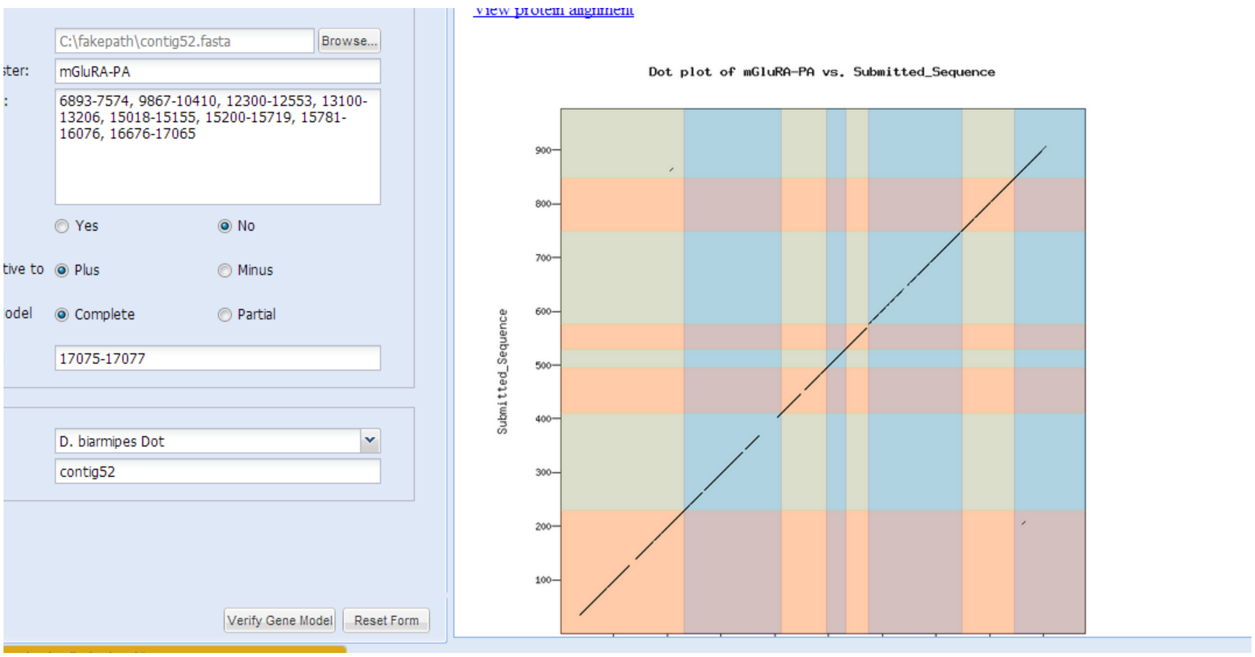


```
mGluRA-PA         481 GKEFYNNYLLNVSFIDLAGSEVKFDRQGDGLARYDILNYQRQENSSGYQYKVIGKWFNGL 540
                      **:************************************** ******************
Submitted_Seq     479 GKDFYNNYLLNVSFIDLAGSEVKFDRQGDGLARYDILNYQRLENSSGYQYKVIGKWFNGL 538

mGluRA-PA         541 QLNSETVVWNKETEQPTSACSLPCEVGMIKKQQ----GDTCCWICDSCESFEYVYDEFTC 596
                      ************:*****************:::    ***********************
Submitted_Seq     539 QLNSETVVWNKEAEQPTSACSLPCEVGMIKNNRYFILGDTCCWICDSCESFEYVYDEFTC 598

mGluRA-PA         597 KDCGPGLWPYADKLSCYALDIQYMKWNSLFALIPMAIAIFGIALTSIVIVLFAKNHDTPL 656
                      ****************:*******:******************:* **::**********
Submitted_Seq     599 KDCGPGLWPYADKLSCFALDIQYMRWNSLFALIPMAIAIFGIAVTIIVMLLFAKNHDTPL 658

mGluRA-PA         657 VRASGRELSYTLLFGILVCYCNTFALIAKPTIGSCVLQRFGIGVGFSIIYSALLTKTNRI 716
                      ********************************.***************************
Submitted_Seq     659 VRASGRELSYTLLFGILVCYCNTFALIAKPTIASCVLQRFGIGVGFSIIYSALLTKTNRI 718

mGluRA-PA         717 SRIFHSASKSAQRLKYISPQSQVVITTSLIAIQVLITMIWMVVEPPGTRFYYPDRREVIL 776
                      **************************:**************************** ****
Submitted_Seq     719 SRIFHSASKSAQRLKYISPQSQVVITASLIAIQVLITMIWMVVEPPGTRFYYPDRTEVIL 778

mGluRA-PA         777 KCKIQDMSFLFSQLYNMILITICTIYAIKTRKIPENFNESKFIGFTMYTTCIIWLAFVPI 836
                      ************************:***********************************
```

```
Submitted_Seq     659 VRASGRELSYTLLFGILVCYCNTFALIAKPTIASCVLQRFGIGVGFSIIYSALLTKTNRI 718

mGluRA-PA         717 SRIFHSASKSAQRLKYISPQSQVVITTSLIAIQVLITMIWMVVEPPGTRFYYPDRREVIL 776
                      **************************:**************************** ****
Submitted_Seq     719 SRIFHSASKSAQRLKYISPQSQVVITASLIAIQVLITMIWMVVEPPGTRFYYPDRTEVIL 778

mGluRA-PA         777 KCKIQDMSFLFSQLYNMILITICTIYAIKTRKIPENFNESKFIGFTMYTTCIIWLAFVPI 836
                      ************************:***********************************
```

```
Submitted_Seq     779 KCKIQDMSFLFSQLYNMILITICTVYAIKTRKIPENFNESKFIGFTMYTTCIIWLAFVPI 838

mGluRA-PA         837 YFGTGNSYEVQTTTLCISISLSASVALVCLYSPKVYILVFHPDKNVRKLTMNSTVYRRSA 896
                      ************************************************************
Submitted_Seq     839 YFGTGNSYEVQTTTLCISISLSASVALVCLYSPKVYILVFHPDKNVRKLTMNSTVYRRSA 898

mGluRA-PA         897 AAVAQGAPTSSGYSRTHAPGTSALTGGAVGTNASSSTLPTQNSPHLDEASAQTNVAHKT- 955
                      *: *****:** ****:* *.:. ****:**.***: * ****.:***.*.*: *.**:
Submitted_Seq     899 ATGAQGAPSSSVYSRTQA-GNTVPTGGALGTTASSA-LQTQNSSNLDEPSGQSAVVHKSS 956

mGluRA-PA         956 ---NGEFLPEVGERVEPICHIVNK 976
                         ****:**  * .*.:*   *:
Submitted_Seq     957 DYSNGEFMPEECECAEAFC---NR 977
```

Dot Plot between the submitted model and the *D. melanogaster* ortholog



C:\fakepath\contig52.fasta

ter: mGluRA-PA

: 6893-7574, 9867-10410, 12300-12553, 13100-13206, 15018-15155, 15200-15719, 15781-16076, 16676-17065

○ Yes ● No

tive to ● Plus ○ Minus

odel ● Complete ○ Partial

17075-17077

D. biarmipes Dot

contig52

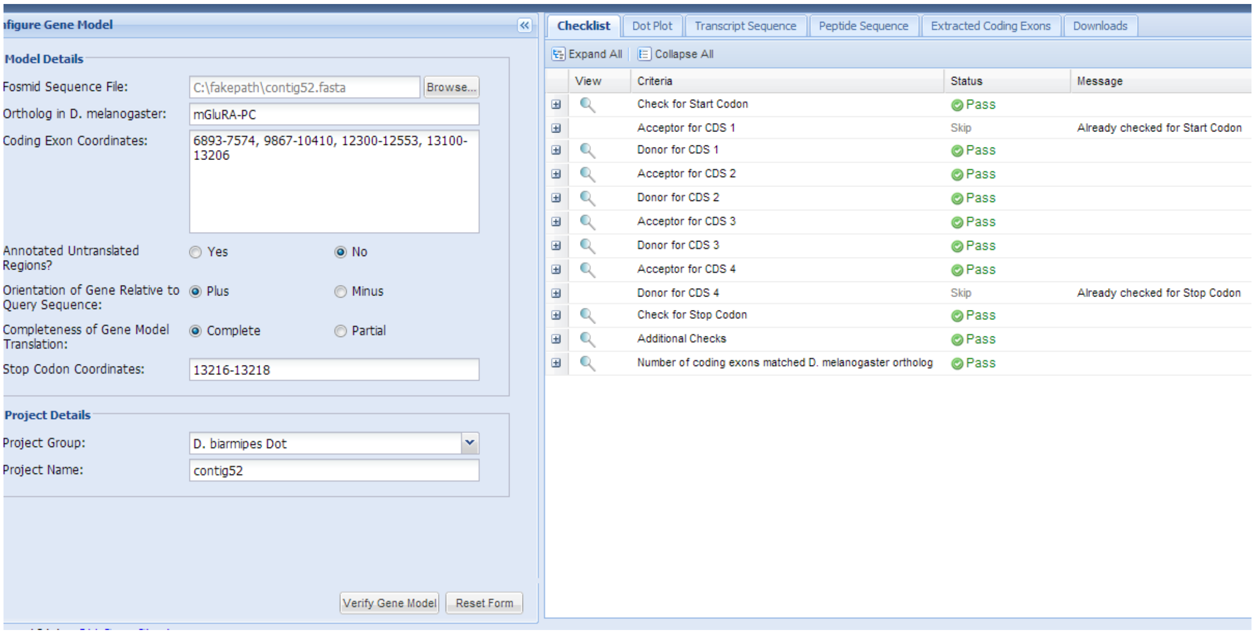
Gene-isoform name: MGluRA-RC

Names of the isoforms with identical coding sequences as this isoform: None

Is the 5’ end of this isoform missing from the end of the project: No

Is the 3’ end of this isoform missing from the end of the project: No

Gene Model Checker Checklist



Fosmid Sequence File: C:\fakepath\contig52.fasta

Ortholog in D. melanogaster: mGluRA-PC

Coding Exon Coordinates: 6893-7574, 9867-10410, 12300-12553, 13100-13206

Annotated Untranslated Regions? ○ Yes ● No

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

Completeness of Gene Model Translation: ● Complete ○ Partial

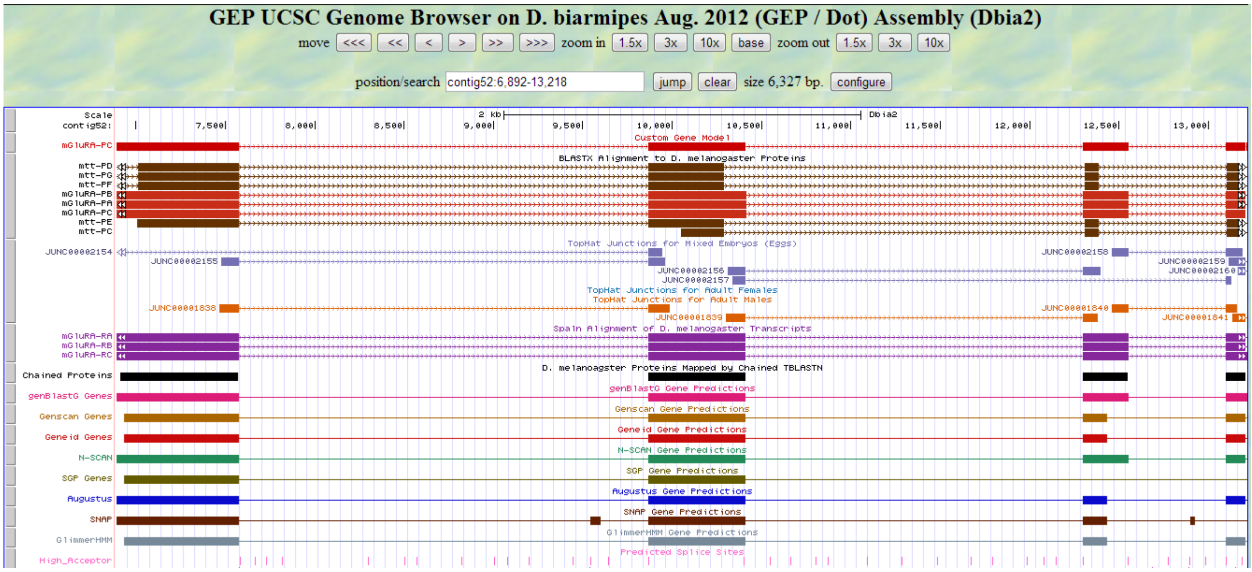
Stop Codon Coordinates: 13216-13218

Project Group: D. biarmipes Dot

Project Name: contig52

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	Skip	Already checked for Stop Codon
	Check for Stop Codon	Pass	
	Additional Checks	Pass	
	Number of coding exons matched D. melanogaster ortholog	Pass	

View of the gene model on the genome browser

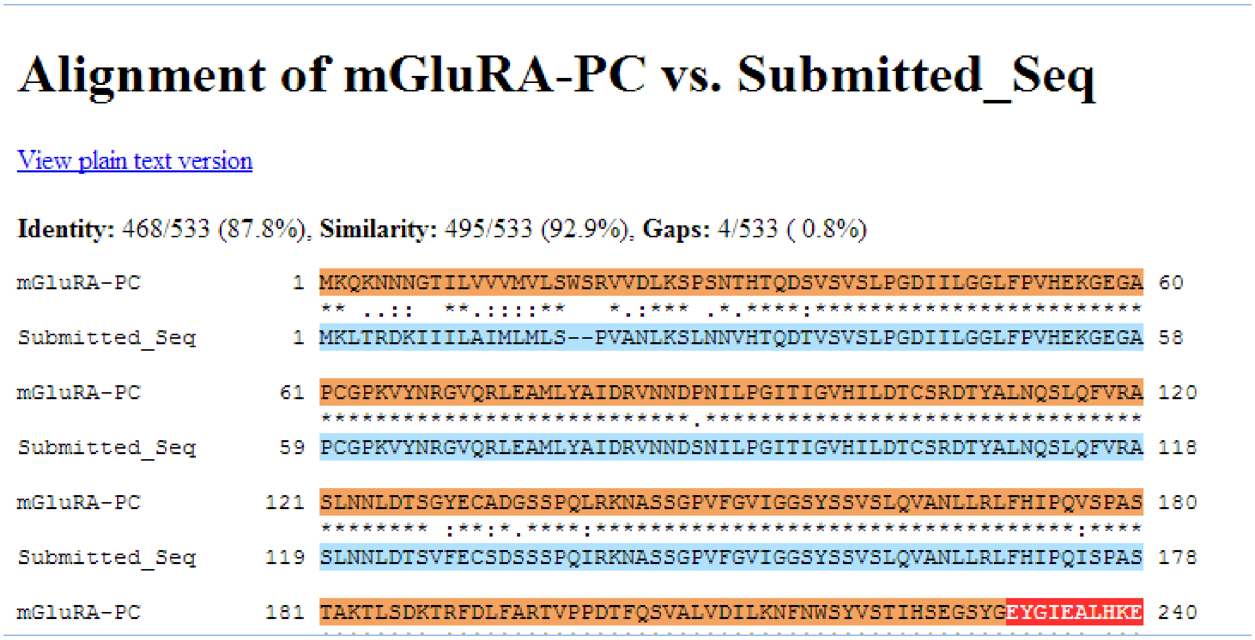


GEP UCSC Genome Browser on D. biarmipes Aug. 2012 (GEP / Dot) Assembly (Dbia2)

position/search contig52:6,892-13,218

size 6,327 bp.

Alignment between the submitted model and the *D. Melanogaster* ortholog

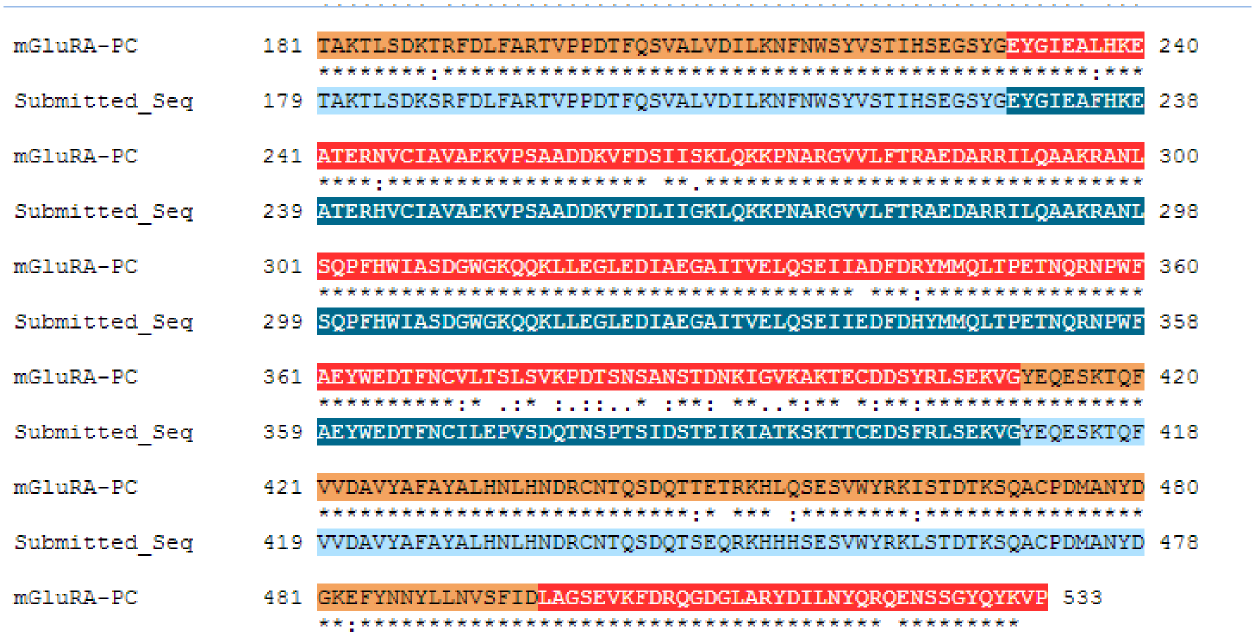


Alignment of mGluRA-PC vs. Submitted_Seq

View plain text version

Identity: 468/533 (87.8%), **Similarity:** 495/533 (92.9%), **Gaps:** 4/533 (0.8%)

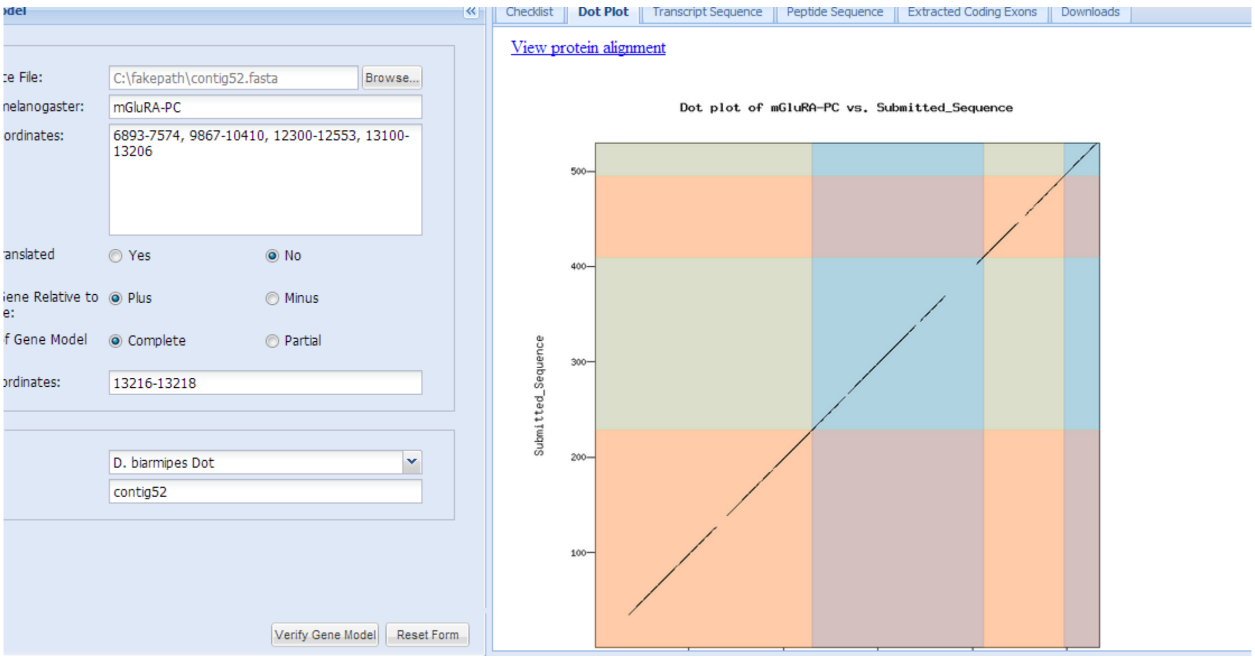
```
mGluRA-PC             1 MKQKNNNGTILVVVMVLSWSRVVDLKSPSNTHTQDSVSVSLPGDIILGGLFPVHEKGEGA 60
                        ** ..::  **.::::**   *.:*** .*.****:************************
Submitted_Seq         1 MKLTRDKIIILAIMLMLS--PVANLKSLNNVHTQDTVSVSLPGDIILGGLFPVHEKGEGA 58

mGluRA-PC            61 PCGPKVYNRGVQRLEAMLYAIDRVNNDPNILPGITIGVHILDTCSRDTYALNQSLQFVRA 120
                        ***************************.********************************
Submitted_Seq        59 PCGPKVYNRGVQRLEAMLYAIDRVNNDSNILPGITIGVHILDTCSRDTYALNQSLQFVRA 118

mGluRA-PC           121 SLNNLDTSGYECADGSSPQLRKNASSGPVFGVIGGSYSSVSLQVANLLRLFHIPQVSPAS 180
                        ******** :**:*.****:***********************************:****
Submitted_Seq       119 SLNNLDTSVFECSDSSSPQIRKNASSGPVFGVIGGSYSSVSLQVANLLRLFHIPQISPAS 178

mGluRA-PC           181 TAKTLSDKTRFDLFARTVPPDTFQSVALVDILKNFNWSYVSTIHSEGSYGEYGIEALHKE 240
```



```
mGluRA-PC           181 TAKTLSDKTRFDLFARTVPPDTFQSVALVDILKNFNWSYVSTIHSEGSYGEYGIEALHKE 240
                        ********:***********************************************:***
Submitted_Seq       179 TAKTLSDKSRFDLFARTVPPDTFQSVALVDILKNFNWSYVSTIHSEGSYGEYGIEAFHKE 238

mGluRA-PC           241 ATERNVCIAVAEKVPSAADDKVFDSIISKLQKKPNARGVVLFTRAEDARRILQAAKRANL 300
                        ****:******************* **.********************************
Submitted_Seq       239 ATERHVCIAVAEKVPSAADDKVFDLIIGKLQKKPNARGVVLFTRAEDARRILQAAKRANL 298

mGluRA-PC           301 SQPFHWIASDGWGKQQKLLEGLEDIAEGAITVELQSEIIADFDRYMMQLTPETNQRNPWF 360
                        *************************************** ***:****************
Submitted_Seq       299 SQPFHWIASDGWGKQQKLLEGLEDIAEGAITVELQSEIIEDFDHYMMQLTPETNQRNPWF 358

mGluRA-PC           361 AEYWEDTFNCVLTSLSVKPDTSNSANSTDNKIGVKAKTECDDSYRLSEKVGYEQESKTQF 420
                        **********:* .:* :.::..* :**: **..*:** *:**:****************
Submitted_Seq       359 AEYWEDTFNCILEPVSDQTNSPTSIDSTEIKIATKSKTTCEDSFRLSEKVGYEQESKTQF 418

mGluRA-PC           421 VVDAVYAFAYALHNLHNDRCNTQSDQTTETRKHLQSESVWYRKISTDTKSQACPDMANYD 480
                        ***************************:* *** :********:****************
Submitted_Seq       419 VVDAVYAFAYALHNLHNDRCNTQSDQTSEQRKHHHSESVWYRKLSTDTKSQACPDMANYD 478

mGluRA-PC           481 GKEFYNNYLLNVSFIDLAGSEVKFDRQGDGLARYDILNYQRQENSSGYQYKVP  533
                        **:************************************** *********
```

```
mGluRA-PC       241 ATERNVCIAVAEKVPSAADDKVFDSIISKLQKKPNARGVVLFTRAEDARRILQAAKRANL 300
                    ****:******************* **.********************************
Submitted_Seq   239 ATERHVCIAVAEKVPSAADDKVFDLIIGKLQKKPNARGVVLFTRAEDARRILQAAKRANL 298

mGluRA-PC       301 SQPFHWIASDGWGKQQKLLEGLEDIAEGAITVELQSEIIADFDRYMMQLTPETNQRNPWF 360
                    *************************************** ***:****************
Submitted_Seq   299 SQPFHWIASDGWGKQQKLLEGLEDIAEGAITVELQSEIIEDFDHYMMQLTPETNQRNPWF 358

mGluRA-PC       361 AEYWEDTFNCVLTSLSVKPDTSNSANSTDNKIGVKAKTECDDSYRLSEKVGYEQESKTQF 420
                    **********:* .:* :.::..* :**: **..*:** *:**:****************
Submitted_Seq   359 AEYWEDTFNCILEPVSDQTNSPTSIDSTEIKIATKSKTTCEDSFRLSEKVGYEQESKTQF 418

mGluRA-PC       421 VVDAVYAFAYALHNLHNDRCNTQSDQTTETRKHLQSESVWYRKISTDTKSQACPDMANYD 480
                    ***************************:* *** :********:****************
Submitted_Seq   419 VVDAVYAFAYALHNLHNDRCNTQSDQTSEQRKHHHSESVWYRKLSTDTKSQACPDMANYD 478

mGluRA-PC       481 GKEFYNNYLLNVSFIDLAGSEVKFDRQGDGLARYDILNYQRQENSSGYQYKVP 533
                    **:************************************** *********
Submitted_Seq   479 GKDFYNNYLLNVSFIDLAGSEVKFDRQGDGLARYDILNYQRLENSSGYQYK-- 529
```

Dot plot between the submitted model and the *D. melanogaster* ortholog



ze File: C:\fakepath\contig52.fasta Browse...

nelanogaster: mGluRA-PC

ordinates: 6893-7574, 9867-10410, 12300-12553, 13100-13206

anslated ○ Yes ● No

iene Relative to e: ● Plus ○ Minus

f Gene Model ● Complete ○ Partial

ordinates: 13216-13218

D. biarmipes Dot

contig52

Verify Gene Model Reset Form

Gene Report Form

Gene name: None

Gene symbol: CG32016

Approximate location in project: 22423-27672

Number of isoforms in D. Melanogaster: 5

Number of isoforms in project: 5

Name of unique isoform based on coding sequence	List of isoforms with identical coding sequences
CG32016-PC	CG32016-PB, CG32016-PG
CG32016-PF	
CG32016-PH	

Isoform Report Form

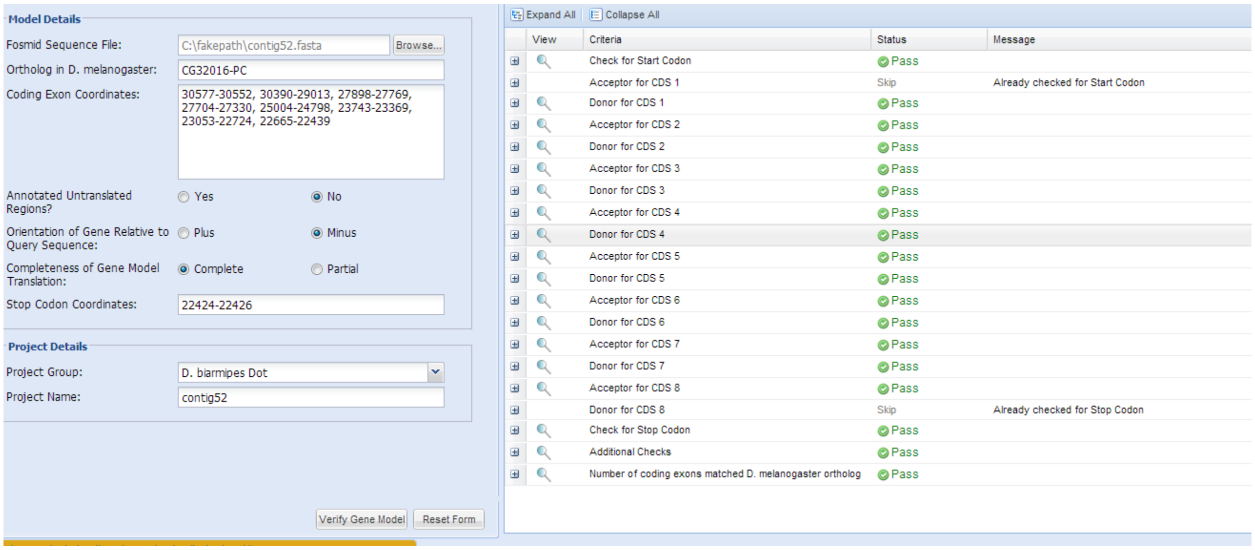
Gene-isoform name: CG32016-PC

Names of the isoforms with identical coding sequences: CG32016-PB, CG32016-PG

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

Is the 3' end of this isoform missing from the end of the project: No

Gene Model Checker Checklist



Model Details

Fosmid Sequence File: C:\fakepath\contig52.fasta

Ortholog in D. melanogaster: CG32016-PC

Coding Exon Coordinates: 30577-30552, 30390-29013, 27898-27769, 27704-27330, 25004-24798, 23743-23369, 23053-22724, 22665-22439

Annotated Untranslated Regions? ○ Yes ● No

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

Completeness of Gene Model Translation: ● Complete ○ Partial

Stop Codon Coordinates: 22424-22426

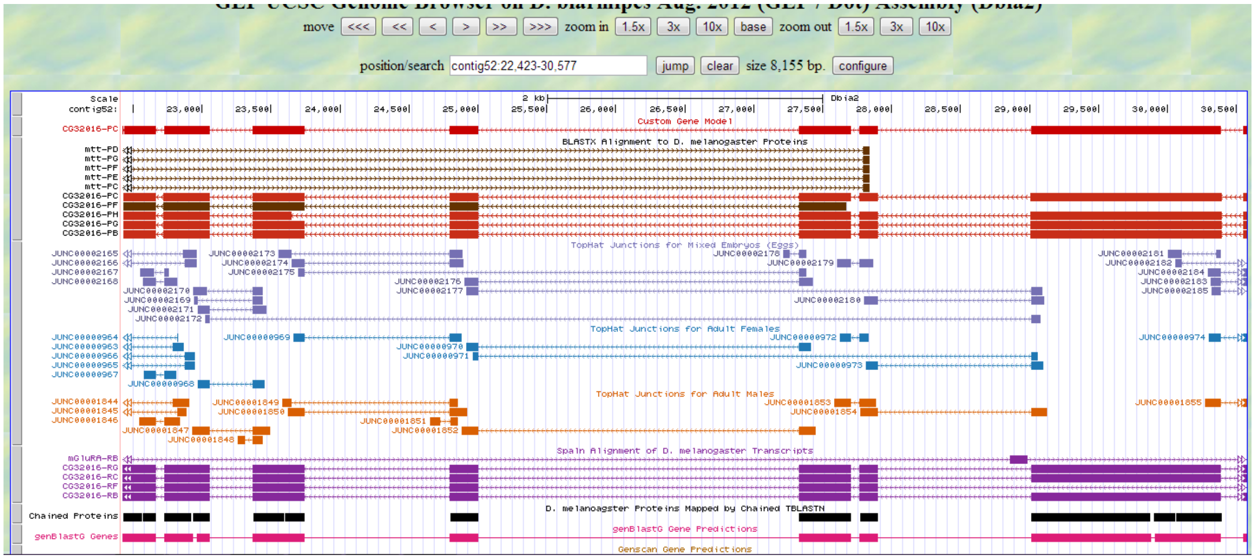
Project Details

Project Group: D. biarmipes Dot

Project Name: contig52

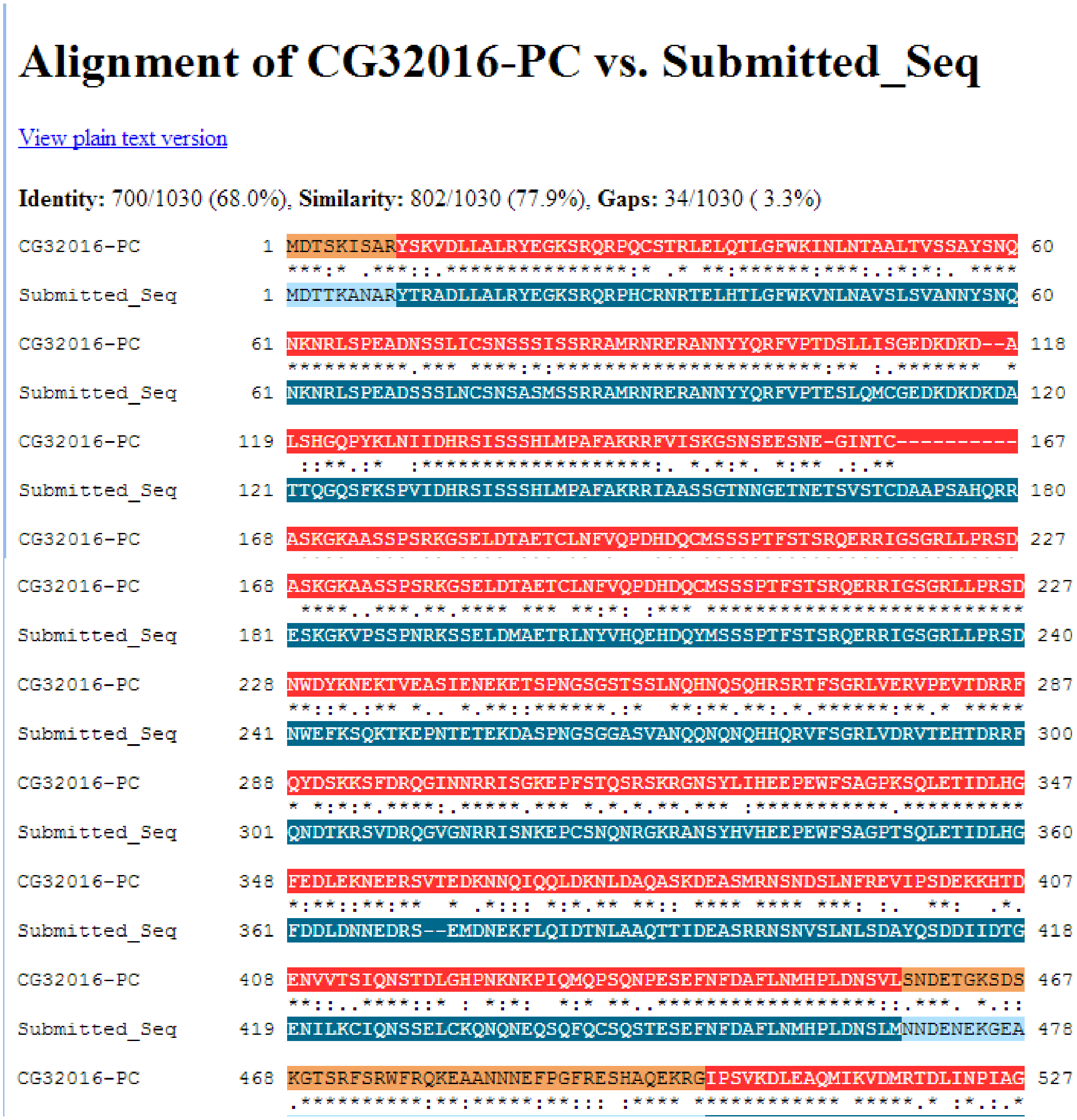
Criteria	Status	Message
Check for Start Codon	Pass	
Acceptor for CDS 1	Skip	Already checked for Start Codon
Donor for CDS 1	Pass	
Acceptor for CDS 2	Pass	
Donor for CDS 2	Pass	
Acceptor for CDS 3	Pass	
Donor for CDS 3	Pass	
Acceptor for CDS 4	Pass	
Donor for CDS 4	Pass	
Acceptor for CDS 5	Pass	
Donor 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	Pass	

View the gene model on the Genome Browser


position/search contig52:22,423-30,577

size 8,155 bp.

Alignment between the submitted model and the D. Melanogaster ortholog

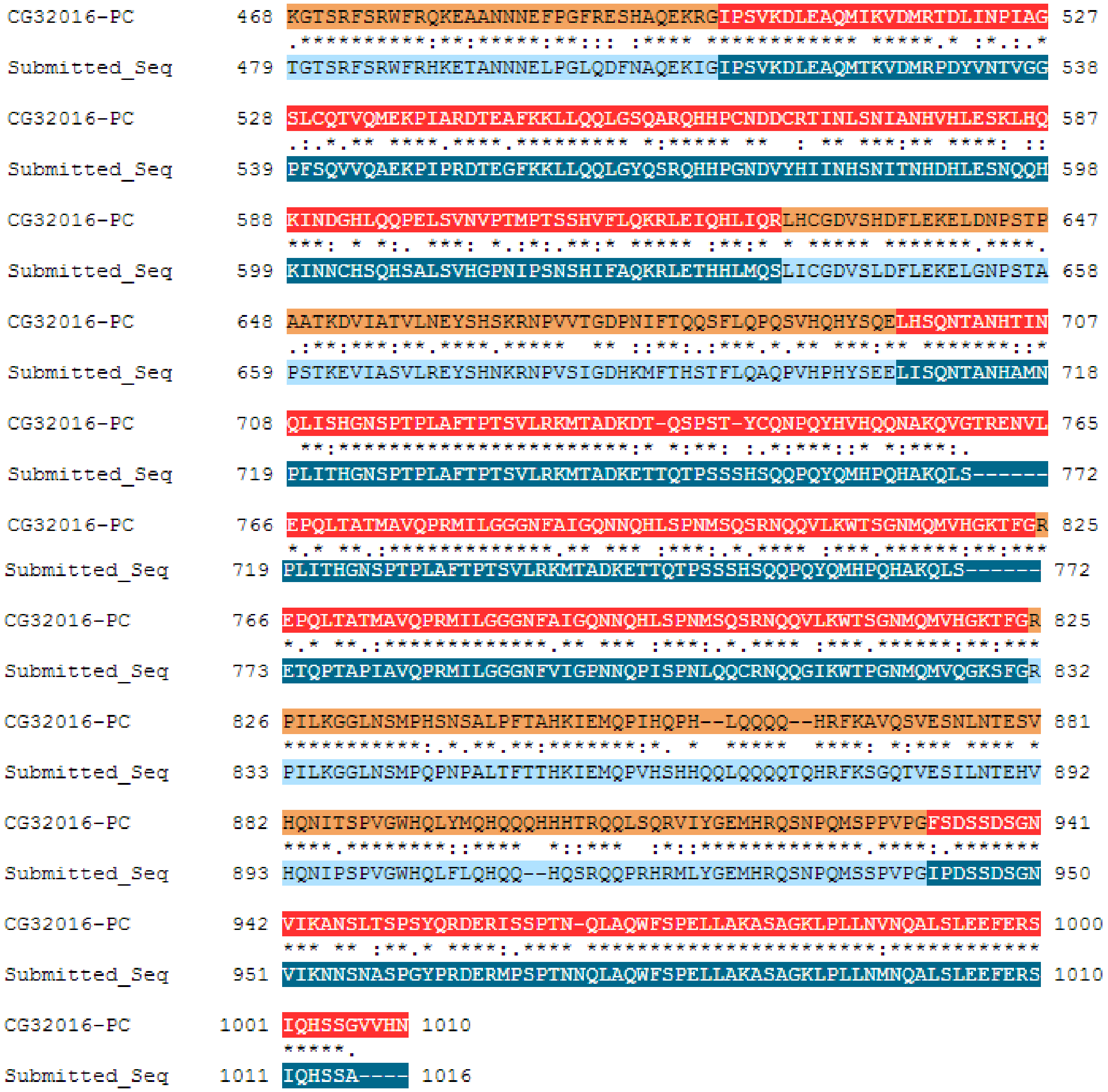


Alignment of CG32016-PC vs. Submitted_Seq

View plain text version

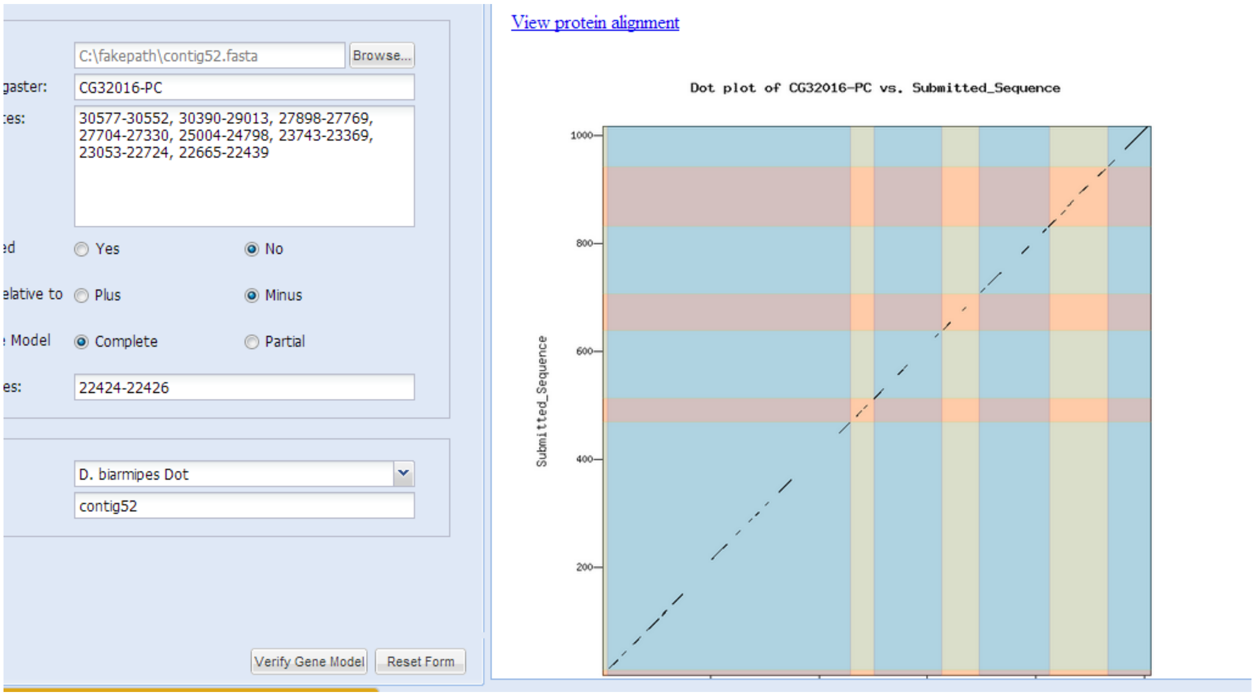
Identity: 700/1030 (68.0%), **Similarity:** 802/1030 (77.9%), **Gaps:** 34/1030 (3.3%)

```
CG32016-PC           1 MDTSKISARYSKVDLLALRYEGKSRQRPQCSTRLELQTLGFWKINLNTAALTVSSAYSNQ 60
                       ***:* .***::.***************:* .* **:******:***:.:*:*:. ****
Submitted_Seq        1 MDTTKANARYTRADLLALRYEGKSRQRPHCRNRTELHTLGFWKVNLNAVSLSVANNYSNQ 60

CG32016-PC          61 NKNRLSPEADNSSLICSNSSSISSRRAMRNRERANNYYQRFVPTDSLLISGEDKDKD--A 118
                       **********.*** ****:*:**********************:** :.*******  *
Submitted_Seq       61 NKNRLSPEADSSSLNCSNSASMSSRRAMRNRERANNYYQRFVPTESLQMCGEDKDKDKDA 120

CG32016-PC         119 LSHGQPYKLNIIDHRSISSSHLMPAFAKRRFVISKGSNSEESNE-GINTC---------- 167
                        ::**.:*  :*******************:. *.*:*. *:** .:.**
Submitted_Seq      121 TTQGQSFKSPVIDHRSISSSHLMPAFAKRRIAASSGTNNGETNETSVSTCDAAPSAHQRR 180

CG32016-PC         168 ASKGKAASSPSRKGSELDTAETCLNFVQPDHDQCMSSSPTFSTSRQERRIGSGRLLPRSD 227

CG32016-PC         168 ASKGKAASSPSRKGSELDTAETCLNFVQPDHDQCMSSSPTFSTSRQERRIGSGRLLPRSD 227
                        ****..***.**.**** *** **:*: :*** **************************
Submitted_Seq      181 ESKGKVPSSPNRKSSELDMAETRLNYVHQEHDQYMSSSPTFSTSRQERRIGSGRLLPRSD 240

CG32016-PC         228 NWDYKNEKTVEASIENEKETSPNGSGSTSSLNQHNQSQHRSRTFSGRLVERVPEVTDRRF 287
                       **::*.:** *.. *.**::******.:*  **:**.**:.*.******:**.* *****
Submitted_Seq      241 NWEFKSQKTKEPNTETEKDASPNGSGGASVANQQNQNQHHQRVFSGRLVDRVTEHTDRRF 300

CG32016-PC         288 QYDSKKSFDRQGINNRRISGKEPFSTQSRSKRGNSYLIHEEPEWFSAGPKSQLETIDLHG 347
                       * *:*:*.****:.*****.*** *.*.*.**.*** :***********.**********
Submitted_Seq      301 QNDTKRSVDRQGVGNRRISNKEPCSNQNRGKRANSYHVHEEPEWFSAGPTSQLETIDLHG 360

CG32016-PC         348 FEDLEKNEERSVTEDKNNQIQQLDKNLDAQASKDEASMRNSNDSLNFREVIPSDEKKHTD 407
                       *:**::**:**  * .*::: *:*.** **:: **** **** ***: :.  **:  .*.
Submitted_Seq      361 FDDLDNNEDRS--EMDNEKFLQIDTNLAAQTTIDEASRRNSNVSLNLSDAYQSDDIIDTG 418

CG32016-PC         408 ENVVTSIQNSTDLGHPNKNKPIQMQPSQNPESEFNFDAFLNMHPLDNSVLSNDETGKSDS 467
                       **::..****::* : *:*:  *:* **..******************::.***. *.::
Submitted_Seq      419 ENILKCIQNSSELCKQNQNEQSQFQCSQSTESEFNFDAFLNMHPLDNSLMNNDENEKGEA 478

CG32016-PC         468 KGTSRFSRWFRQKEAANNNEFPGFRESHAQEKRGIPSVKDLEAQMIKVDMRTDLINPIAG 527
                       .**********:**:*****:**:::  :**** ************ *****.* :*.:.*
```



```
CG32016-PC       468 KGTSRFSRWFRQKEAANNNEFPGFRESHAQEKRGIPSVKDLEAQMIKVDMRTDLINPIAG 527
                     .**********:**:*****:**::: :**** ************ *****.* :*.:.*
Submitted_Seq    479 TGTSRFSRWFRHKETANNNELPGLQDFNAQEKIGIPSVKDLEAQMTKVDMRPDYVNTVGG 538

CG32016-PC       528 SLCQTVQMEKPIARDTEAFKKLLQQLGSQARQHHPCNDDCRTINLSNIANHVHLESKLHQ 587
                     .:.*.** ****.****.********* *:***** **  : ** ***:** ****: ::
Submitted_Seq    539 PFSQVVQAEKPIPRDTEGFKKLLQQLGYQSRQHHPGNDVYHIINHSNITNHDHLESNQQH 598

CG32016-PC       588 KINDGHLQQPELSVNVPTMPTSSHVFLQKRLEIQHLIQRLHCGDVSHDFLEKELDNPSTP 647
                     ***: * *:. ***: *.:*:.**:* ***** :**:* * ***** *******.****.
Submitted_Seq    599 KINNCHSQHSALSVHGPNIPSNSHIFAQKRLETHHLMQSLICGDVSLDFLEKELGNPSTA 658

CG32016-PC       648 AATKDVIATVLNEYSHSKRNPVVTGDPNIFTQQSFLQPQSVHQHYSQELHSQNTANHTIN 707
                     .:**:***:**.****.*****  ** ::**:.:***.*.** ***:** *******::*
Submitted_Seq    659 PSTKEVIASVLREYSHNKRNPVSIGDHKMFTHSTFLQAQPVHPHYSEELISQNTANHAMN 718

CG32016-PC       708 QLISHGNSPTPLAFTPTSVLRKMTADKDT-QSPST-YCQNPQYHVHQQNAKQVGTRENVL 765
                      **:************************:* *:**: :.*:***::* *:***:.
Submitted_Seq    719 PLITHGNSPTPLAFTPTSVLRKMTADKETTQTPSSSHSQQPQYQMHPQHAKQLS------ 772

CG32016-PC       766 EPQLTATMAVQPRMILGGGNFAIGQNNQHLSPNMSQSRNQQVLKWTSGNMQMVHGKTFGR 825
                     *.* **.:*************.** *** :***:.*.**** :***.******:**:***
Submitted_Seq    719 PLITHGNSPTPLAFTPTSVLRKMTADKETTQTPSSSHSQQPQYQMHPQHAKQLS------ 772

CG32016-PC       766 EPQLTATMAVQPRMILGGGNFAIGQNNQHLSPNMSQSRNQQVLKWTSGNMQMVHGKTFGR 825
                     *.* **.:*************.** *** :***:.*.**** :***.******:**:***
Submitted_Seq    773 ETQPTAPIAVQPRMILGGGNFVIGPNNQPISPNLQQCRNQQGIKWTPGNMQMVQGKSFGR 832

CG32016-PC       826 PILKGGLNSMPHSNSALPFTAHKIEMQPIHQPH--LQQQQ--HRFKAVQSVESNLNTESV 881
                     ***********:.*.**.**:*******:*. * ***** ****: *:*** **** *
Submitted_Seq    833 PILKGGLNSMPQPNPALTFTTHKIEMQPVHSHHQQLQQQQTQHRFKSGQTVESILNTEHV 892

CG32016-PC       882 HQNITSPVGWHQLYMQHQQQHHHTRQQLSQRVIYGEMHRQSNPQMSPPVPGFSDSSDSGN 941
                     ****.********::****  *::***  :*::*************.****:.*******
Submitted_Seq    893 HQNIPSPVGWHQLFLQHQQ--HQSRQQPRHRMLYGEMHRQSNPQMSSPVPGIPDSSDSGN 950

CG32016-PC       942 VIKANSLTSPSYQRDERISSPTN-QLAQWFSPELLAKASAGKLPLLNVNQALSLEEFERS 1000
                     *** ** :**.* ****:.**** ************************:************
Submitted_Seq    951 VIKNNSNASPGYPRDERMPSPTNNQLAQWFSPELLAKASAGKLPLLNMNQALSLEEFERS 1010

CG32016-PC      1001 IQHSSGVVHN 1010
                     *****.
Submitted_Seq   1011 IQHSSA---- 1016
```

Dot Plot between the submitted model and the D. melanogaster ortholog



	C:\fakepath\contig52.fasta	
gaster:	CG32016-PC	
es:	30577-30552, 30390-29013, 27898-27769, 27704-27330, 25004-24798, 23743-23369, 23053-22724, 22665-22439	
ed	○ Yes	● No
elative to	○ Plus	● Minus
Model	● Complete	○ Partial
es:	22424-22426	
	D. biarmipes Dot	
	contig52	

Isoform Report Form

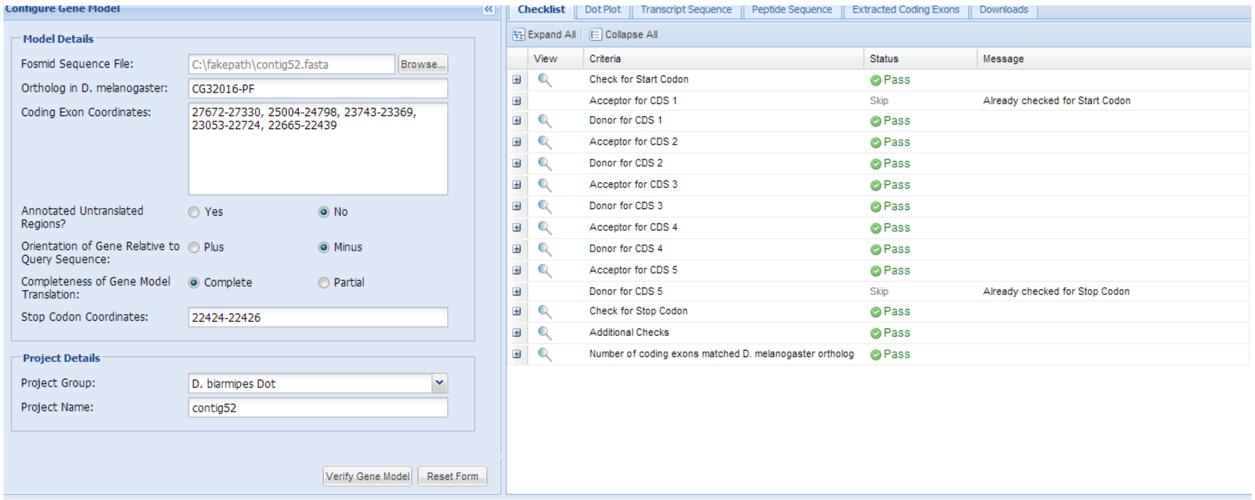
Gene-isoform name: CG32016-PF

Names of the isoforms with identical coding sequences: None

Is the 5’ end of this isoform missing from the end of the project: No

Is the 3’ end of this isoform missing from the end of the project: No

Gene Model Checker Checklist



Configure Gene Model

Model Details

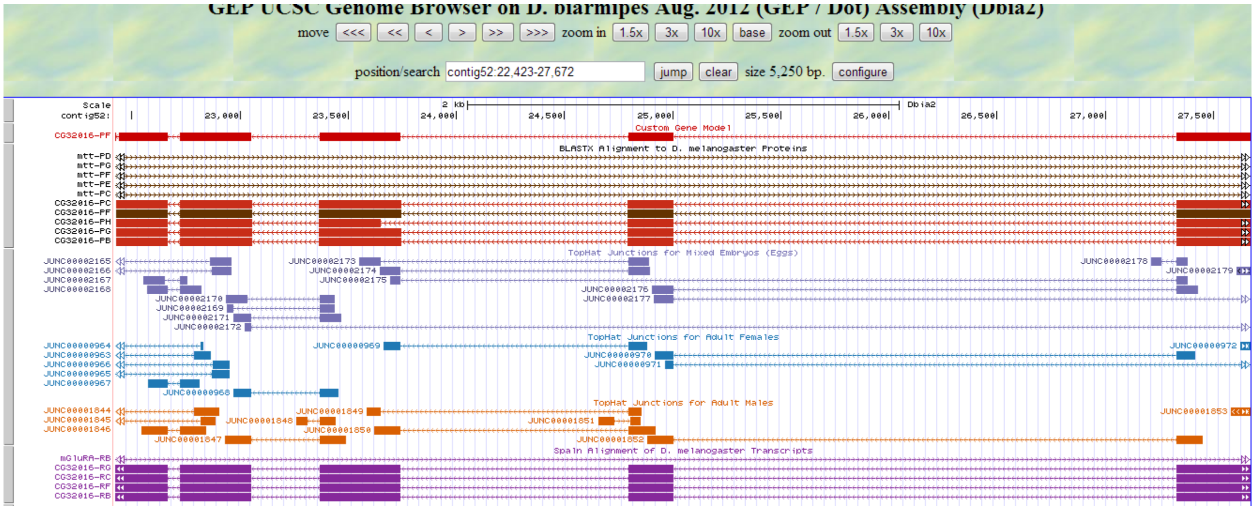
Fosmid Sequence File:	C:\fakepath\contig52.fasta	
Ortholog in D. melanogaster:	CG32016-PF	
Coding Exon Coordinates:	27672-27330, 25004-24798, 23743-23369, 23053-22724, 22665-22439	
Annotated Untranslated Regions?	○ Yes	● No
Orientation of Gene Relative to Query Sequence:	○ Plus	● Minus
Completeness of Gene Model Translation:	● Complete	○ Partial
Stop Codon Coordinates:	22424-22426	

Project Details

Project Group:	D. biarmipes Dot
Project Name:	contig52

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	Skip	Already checked for Stop Codon
	Check for Stop Codon	Pass	
	Additional Checks	Pass	
	Number of coding exons matched D. melanogaster ortholog	Pass	

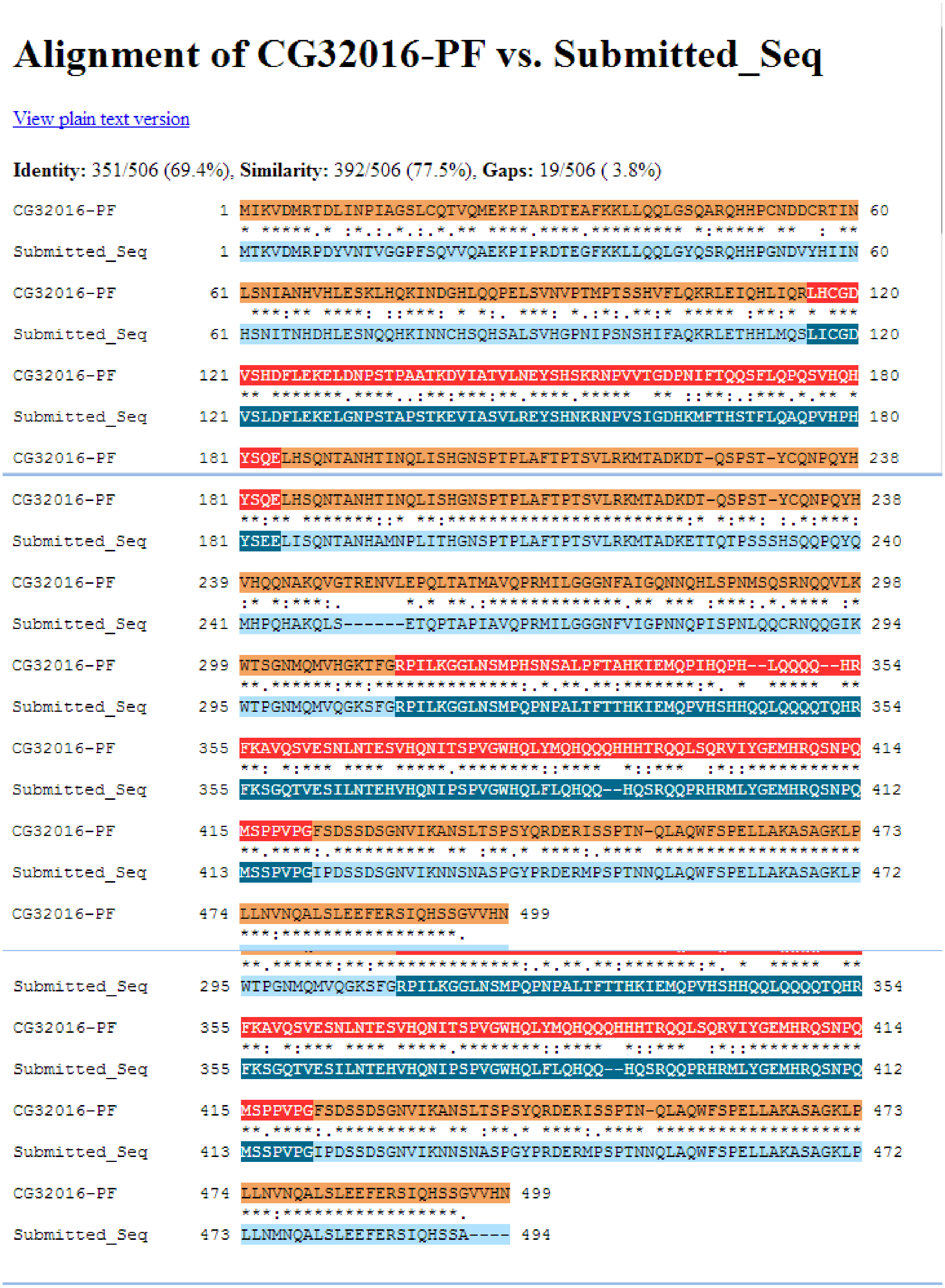
View the gene model on the Genome Browser



position/search contig52:22,423-27,672

size 5,250 bp.

Alignment between the submitted model and the D. Melanogaster ortholog



Alignment of CG32016-PF vs. Submitted_Seq

View plain text version

Identity: 351/506 (69.4%), **Similarity:** 392/506 (77.5%), **Gaps:** 19/506 (3.8%)

```
CG32016-PF          1 MIKVDMRTDLINPIAGSLCQTVQMEKPIARDTEAFKKLLQQLGSQARQHHPCNDDCRTIN 60
                      * *****.* :*.:.*.:.*.** ****.****.********* *:***** **  : **
Submitted_Seq       1 MTKVDMRPDYVNTVGGPFSQVVQAEKPIPRDTEGFKKLLQQLGYQSRQHHPGNDVYHIIN 60

CG32016-PF         61 LSNIANHVHLESKLHQKINDGHLQQPELSVNVPTMPTSSHVFLQKRLEIQHLIQRLHCGD 120
                       ***:** ****: ::***: * *:. ***: *.:*:.**:* ***** :**:* * ***
Submitted_Seq      61 HSNITNHDHLESNQQHKINNCHSQHSALSVHGPNIPSNSHIFAQKRLETHHLMQSLICGD 120

CG32016-PF        121 VSHDFLEKELDNPSTPAATKDVIATVLNEYSHSKRNPVVTGDPNIFTQQSFLQPQSVHQH 180
                      ** *******.****..:**:***:**.****.*****  ** ::**:.:***.*.** *
Submitted_Seq     121 VSLDFLEKELGNPSTAPSTKEVIASVLREYSHNKRNPVSIGDHKMFTHSTFLQAQPVHPH 180

CG32016-PF        181 YSQELHSQNTANHTINQLISHGNSPTPLAFTPTSVLRKMTADKDT-QSPST-YCQNPQYH 238
```

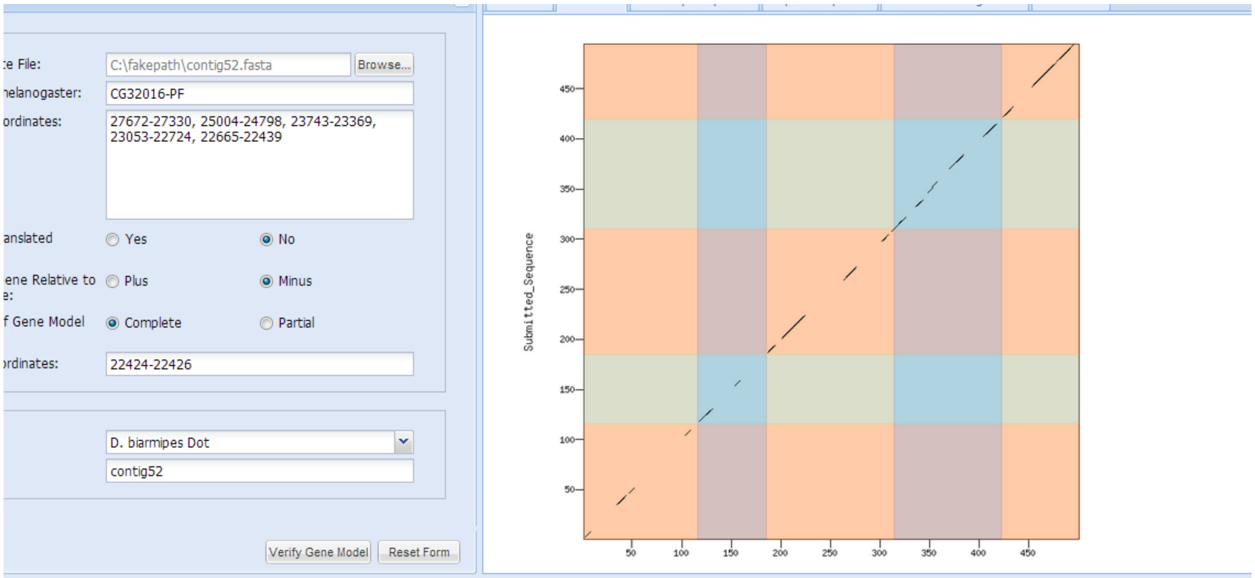
```
CG32016-PF        181 YSQELHSQNTANHTINQLISHGNSPTPLAFTPTSVLRKMTADKDT-QSPST-YCQNPQYH 238
```

```
                      **:** *******::* **:***********************:* *:**: :.*:***:
Submitted_Seq     181 YSEELISQNTANHAMNPLITHGNSPTPLAFTPTSVLRKMTADKETTQTPSSSHSQQPQYQ 240

CG32016-PF        239 VHQQNAKQVGTRENVLEPQLTATMAVQPRMILGGGNFAIGQNNQHLSPNMSQSRNQQVLK 298
                      :* *:***:.      *.* **.:*************.** *** :***:.*.**** :*
Submitted_Seq     241 MHPQHAKQLS------ETQPTAPIAVQPRMILGGGNFVIGPNNQPISPNLQQCRNQQGIK 294

CG32016-PF        299 WTSGNMQMVHGKTFGRPILKGGLNSMPHSNSALPFTAHKIEMQPIHQPH--LQQQQ--HR 354
                      **.******:**:**************:.*.**.**:*******:*. *  *****  **
Submitted_Seq     295 WTPGNMQMVQGKSFGRPILKGGLNSMPQPNPALTFTTHKIEMQPVHSHHQQLQQQQTQHR 354

CG32016-PF        355 FKAVQSVESNLNTESVHQNITSPVGWHQLYMQHQQQHHHTRQQLSQRVIYGEMHRQSNPQ 414
                      **: *:*** **** *****.********::****  *::***  :*::***********
Submitted_Seq     355 FKSGQTVESILNTEHVHQNIPSPVGWHQLFLQHQQ--HQSRQQPRHRMLYGEMHRQSNPQ 412

CG32016-PF        415 MSPPVPGFSDSSDSGNVIKANSLTSPSYQRDERISSPTN-QLAQWFSPELLAKASAGKLP 473
                      **.****:.********** ** :**.* ****:.**** ********************
Submitted_Seq     413 MSSPVPGIPDSSDSGNVIKNNSNASPGYPRDERMPSPTNNQLAQWFSPELLAKASAGKLP 472

CG32016-PF        474 LLNVNQALSLEEFERSIQHSSGVVHN 499
                      ***:*****************.
```

```
                      **.******:**:**************:.*.**.**:*******:*. *  *****  **
Submitted_Seq     295 WTPGNMQMVQGKSFGRPILKGGLNSMPQPNPALTFTTHKIEMQPVHSHHQQLQQQQTQHR 354

CG32016-PF        355 FKAVQSVESNLNTESVHQNITSPVGWHQLYMQHQQQHHHTRQQLSQRVIYGEMHRQSNPQ 414
                      **: *:*** **** *****.********::****  *::***  :*::***********
Submitted_Seq     355 FKSGQTVESILNTEHVHQNIPSPVGWHQLFLQHQQ--HQSRQQPRHRMLYGEMHRQSNPQ 412

CG32016-PF        415 MSPPVPGFSDSSDSGNVIKANSLTSPSYQRDERISSPTN-QLAQWFSPELLAKASAGKLP 473
                      **.****:.********** ** :**.* ****:.**** ********************
Submitted_Seq     413 MSSPVPGIPDSSDSGNVIKNNSNASPGYPRDERMPSPTNNQLAQWFSPELLAKASAGKLP 472

CG32016-PF        474 LLNVNQALSLEEFERSIQHSSGVVHN 499
                      ***:*****************.
```

```
Submitted_Seq     473 LLNMNQALSLEEFERSIQHSSA---- 494
```

Dot plot between the submitted model and the D. Melanogaster ortholog



e File:	C:\fakepath\contig52.fasta
nelanogaster:	CG32016-PF
ordinates:	27672-27330, 25004-24798, 23743-23369, 23053-22724, 22665-22439
anslated	○ Yes ● No
ene Relative to e:	○ Plus ● Minus
f Gene Model	● Complete ○ Partial
rdinates:	22424-22426
	D. biarmipes Dot
	contig52

Isoform Report Form

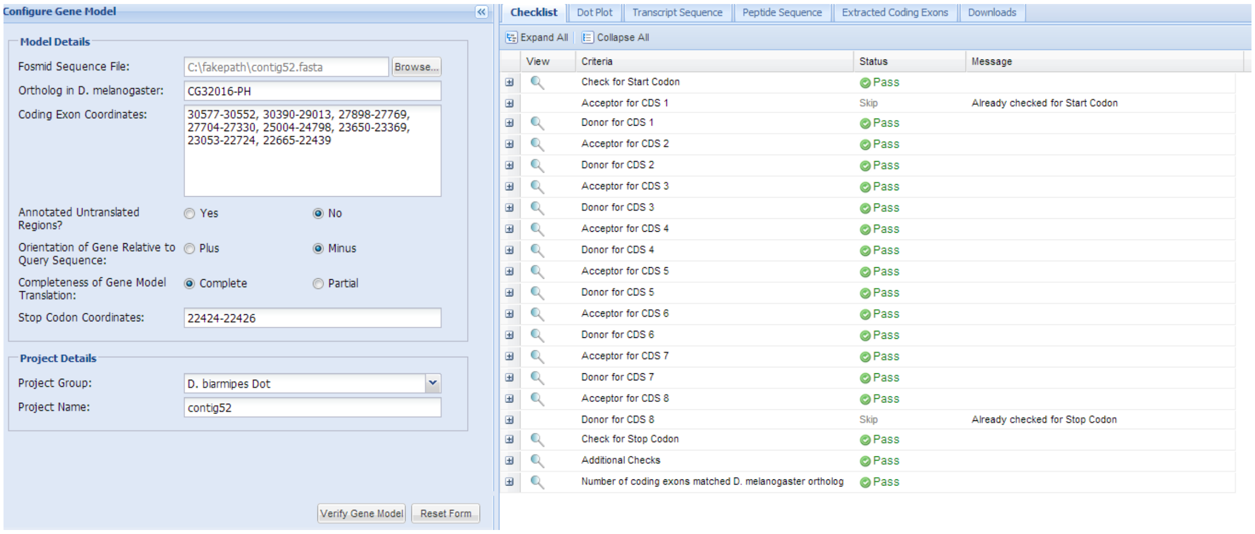
Gene-isoform name: CG32016-PH

Names of the isoforms with identical coding sequences: none

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

Is the 3' end of this isoform missing from the end of the project: No

Genome checker checklist



Model Details

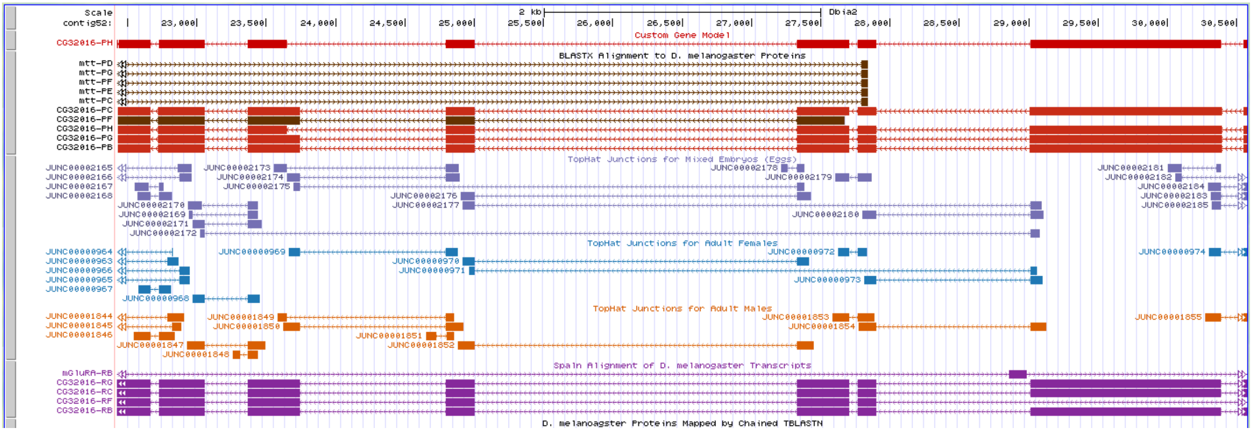
Fosmid Sequence File:	C:\fakepath\contig52.fasta
Ortholog in D. melanogaster:	CG32016-PH
Coding Exon Coordinates:	30577-30552, 30390-29013, 27898-27769, 27704-27330, 25004-24798, 23650-23369, 23053-22724, 22665-22439
Annotated Untranslated Regions?	○ Yes ● No
Orientation of Gene Relative to Query Sequence:	○ Plus ● Minus
Completeness of Gene Model Translation:	● Complete ○ Partial
Stop Codon Coordinates:	22424-22426

Project Details

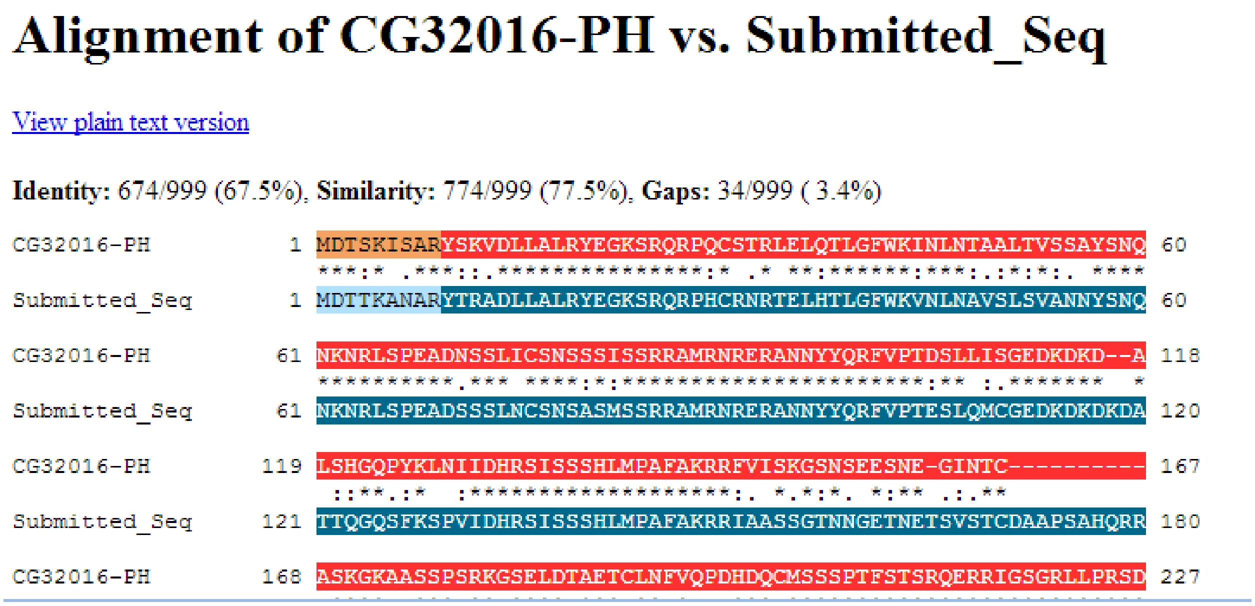
Project Group:	D. biarmipes Dot
Project Name:	contig52

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	✓ Pass	

View the gene model on the genome browser



Alignment between the submitted model and the D. melanogaster ortholog



Alignment of CG32016-PH vs. Submitted_Seq

View plain text version

Identity: 674/999 (67.5%), **Similarity:** 774/999 (77.5%), **Gaps:** 34/999 (3.4%)

```
CG32016-PH           1 MDTSKISARYSKVDLLALRYEGKSRQRPQCSTRLELQTLGFWKINLNTAALTVSSAYSNQ  60
                       ***:* .***::.***************:* .* **:******:***:.:*:*:. ****
Submitted_Seq        1 MDTTKANARYTRADLLALRYEGKSRQRPHCRNRTELHTLGFWKVNLNAVSLSVANNYSNQ  60

CG32016-PH          61 NKNRLSPEADNSSLICSNSSSISSRRAMRNRERANNYYQRFVPTDSLLISGEDKDKD--A 118
                       **********.*** ****:*:**********************:** :.*******  *
Submitted_Seq       61 NKNRLSPEADSSSLNCSNSASMSSRRAMRNRERANNYYQRFVPTESLQMCGEDKDKDKDA 120

CG32016-PH         119 LSHGQPYKLNIIDHRSISSSHLMPAFAKRRFVISKGSNSEESNE-GINTC---------- 167
                        ::**.:*  :*******************:. *.*:*. *:** .:.**
Submitted_Seq      121 TTQGQSFKSPVIDHRSISSSHLMPAFAKRRIAASSGTNNGETNETSVSTCDAAPSAHQRR 180

CG32016-PH         168 ASKGKAASSPSRKGSELDTAETCLNFVQPDHDQCMSSSPTFSTSRQERRIGSGRLLPRSD 227
```

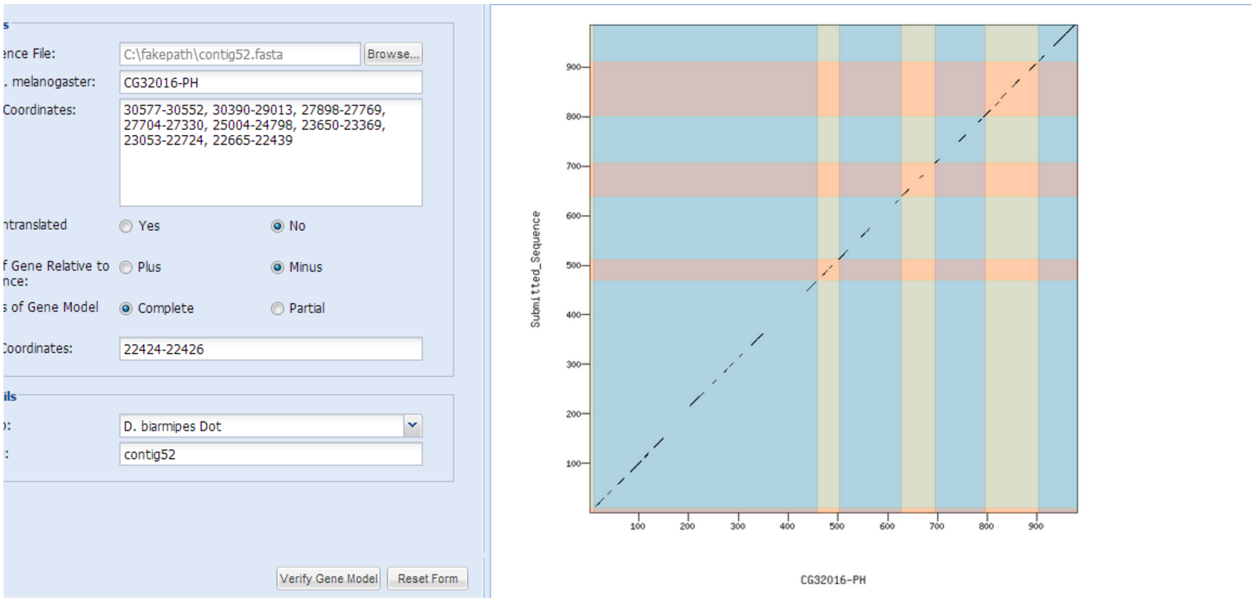
CG32016-PH   168 ASKGKAASSPSRKGSELDTAETCLNFVQPDHDQCMSSSPTFSTSRQERRIGSGRLLPRSD 227

```
                        ****..***.**.**** *** **:*: :*** **************************
Submitted_Seq      181 ESKGKVPSSPNRKSSELDMAETRLNYVHQEHDQYMSSSPTFSTSRQERRIGSGRLLPRSD 240

CG32016-PH         228 NWDYKNEKTVEASIENEKETSPNGSGSTSSLNQHNQSQHRSRTFSGRLVERVPEVTDRRF 287
                       **::*.:** *.. *.**::******.:*  **:**.**:.*.******:**.* *****
Submitted_Seq      241 NWEFKSQKTKEPNTETEKDASPNGSGGASVANQQNQNQHHQRVFSGRLVDRVTEHTDRRF 300

CG32016-PH         288 QYDSKKSFDRQGINNRRISGKEPFSTQSRSKRGNSYLIHEEPEWFSAGPKSQLETIDLHG 347
                       * *:*:*.****:.*****.*** *.*.*.**.*** :***********.**********
Submitted_Seq      301 QNDTKRSVDRQGVGNRRISNKEPCSNQNRGKRANSYHVHEEPEWFSAGPTSQLETIDLHG 360

CG32016-PH         348 FEDLEKNEERSVTEDKNNQIQQLDKNLDAQASKDEASMRNSNDSLNFREVIPSDEKKHTD 407
                       *:**::**:**  * .*::: *:*.** **:: **** **** ***: :.  **:  .*.
Submitted_Seq      361 FDDLDNNEDRS--EMDNEKFLQIDTNLAAQTTIDEASRRNSNVSLNLSDAYQSDDIIDTG 418

CG32016-PH         408 ENVVTSIQNSTDLGHPNKNKPIQMQPSQNPESEFNFDAFLNMHPLDNSVLSNDETGKSDS 467
                       **::..****::* : *:*:  *:* **..******************::.***. *.::
Submitted_Seq      419 ENILKCIQNSSELCKQNQNEQSQFQCSQSTESEFNFDAFLNMHPLDNSLMNNDENEKGEA 478

CG32016-PH         468 KGTSRFSRWFRQKEAANNNEFPGFRESHAQEKRGIPSVKDLEAQMIKVDMRTDLINPIAG 527
                       .**********:**:*****:**::: :**** ************ *****.* :*.:.*
```

```
CG32016-PH         468 KGTSRFSRWFRQKEAANNNEFPGFRESHAQEKRGIPSVKDLEAQMIKVDMRTDLINPIAG 527
                       .**********:**:*****:**::: :**** ************ *****.* :*.:.*
Submitted_Seq      479 TGTSRFSRWFRHKETANNNELPGLQDFNAQEKIGIPSVKDLEAQMTKVDMRPDYVNTVGG 538

CG32016-PH         528 SLCQTVQMEKPIARDTEAFKKLLQQLGSQARQHHPCNDDCRTINLSNIANHVHLESKLHQ 587
                       .:.*.** ****.****.********* *:***** **   : ** ***:** ****: ::
Submitted_Seq      539 PFSQVVQAEKPIPRDTEGFKKLLQQLGYQSRQHHPGNDVYHIINHSNITNHDHLESNQQH 598

CG32016-PH         588 KINDGHLQQPELSVNVPTMPTSSHVFLQKRLEIQHLIQRLHCGDVSHDFLEKELDNPSTP 647
                       ***: * *:. ***: *.:*:.**:* ***** :**:* * ***** *******.****.
Submitted_Seq      599 KINNCHSQHSALSVHGPNIPSNSHIFAQKRLETHHLMQSLICGDVSLDFLEKELGNPSTA 658

CG32016-PH         648 AATKDVIATVLNEYSHSKRNPVVTGDPNIFTQQSFLQPQSVHQHYSQVLRKMTADKDT-Q 706
                       .:**:***:**.****.*****  ** ::**:.:***.*.** ***:*********:* *
Submitted_Seq      659 PSTKEVIASVLREYSHNKRNPVSIGDHKMFTHSTFLQAQPVHPHYSEVLRKMTADKETTQ 718

CG32016-PH         707 SPST-YCQNPQYHVHQQNAKQVGTRENVLEPQLTATMAVQPRMILGGGNFAIGQNNQHLS 765
                       :**: :.*:***::* *:***:.      *.* **.:*************.** *** :*
Submitted_Seq      719 TPSSSHSQQPQYQMHPQHAKQLS------ETQPTAPIAVQPRMILGGGNFVIGPNNQPIS 772

CG32016-PH         766 PNMSQSRNQQVLKWTSGNMQMVHGKTFGRPILKGGLNSMPHSNSALPFTAHKIEMQPIHQ 825
                       **:.*.**** :***.******:**:**************:.*.**.**:*******:*.
```

```
Submitted_Seq      719 TPSSSHSQQPQYQMHPQHAKQLS------ETQPTAPIAVQPRMILGGGNFVIGPNNQPIS 772

CG32016-PH         766 PNMSQSRNQQVLKWTSGNMQMVHGKTFGRPILKGGLNSMPHSNSALPFTAHKIEMQPIHQ 825
                       **:.*.**** :***.******:**:**************:.*.**.**:*******:*.
Submitted_Seq      773 PNLQQCRNQQGIKWTPGNMQMVQGKSFGRPILKGGLNSMPQPNPALTFTTHKIEMQPVHS 832

CG32016-PH         826 PH--LQQQQ--HRFKAVQSVESNLNTESVHQNITSPVGWHQLYMQHQQQHHHTRQQLSQR 881
                        *  *****  ****: *:*** **** *****.********::****  *::***  :*
Submitted_Seq      833 HHQQLQQQQTQHRFKSGQTVESILNTEHVHQNIPSPVGWHQLFLQHQQ--HQSRQQPRHR 890

CG32016-PH         882 VIYGEMHRQSNPQMSPPVPGFSDSSDSGNVIKANSLTSPSYQRDERISSPTN-QLAQWFS 940
                       ::*************.****:.********** ** :**.* ****:.**** *******
Submitted_Seq      891 MLYGEMHRQSNPQMSSPVPGIPDSSDSGNVIKNNSNASPGYPRDERMPSPTNNQLAQWFS 950

CG32016-PH         941 PELLAKASAGKLPLLNVNQALSLEEFERSIQHSSGVVHN 979
                       ****************:*****************.
Submitted_Seq      951 PELLAKASAGKLPLLNMNQALSLEEFERSIQHSSA---- 985
```

Dot Plot between the submitted model and the d. melanogaster ortholog



ence File: C:\fakepath\contig52.fasta

. melanogaster: CG32016-PH

Coordinates: 30577-30552, 30390-29013, 27898-27769, 27704-27330, 25004-24798, 23650-23369, 23053-22724, 22665-22439

ntranslated ○ Yes ● No

f Gene Relative to nce: ○ Plus ● Minus

s of Gene Model ● Complete ○ Partial

Coordinates: 22424-22426

ils

): D. biarmipes Dot

: contig52

Merged GFF Files Custom Tracks