GEP Annotation Report

Student name: ___Kaitlin Leddy_________
Student email: ______kleddy@unm.edu___________
Faculty Advisor:_____Dr. Paul Szauter_______
College/University:__University of New Mexico

Project details

Project name:___*D. biarmipes* Contig 6
Project species:__*D. biarmipes*
Date of submission:__6-25-13__
Size of project in base pairs:__23,634__
Number Of genes in project:__1___

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

Gene report form

Gene name (i.e. *D. mojavensis eyeless*):_*D. biarmipes* Ankyrin
Gene symbol (i.e.dmoj_ey):_dbia_Ank
Approximate location in project (from 5' end to 3' end):__36965-13331___
Number of isoforms in *D. melanogaster*:_4___
Number of isoforms in this project:___4__

Complete the following table for all the isoforms in this project:

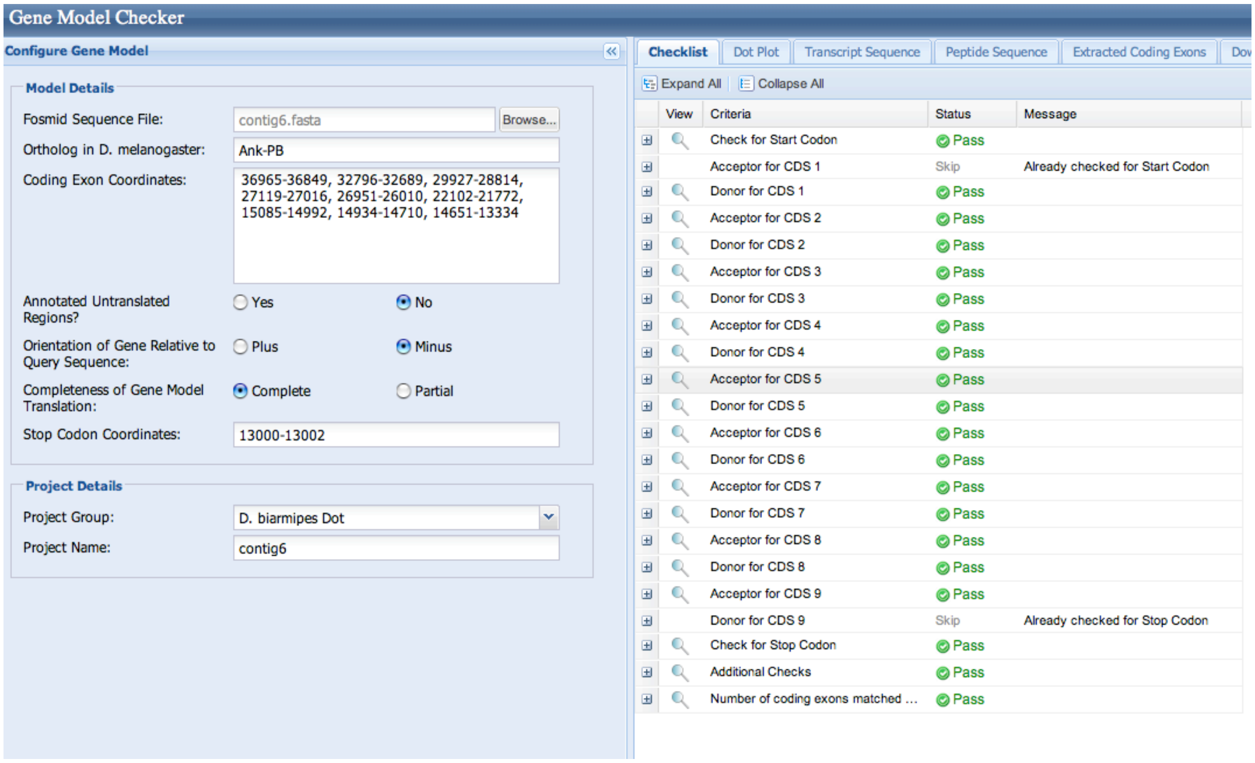
Name of unique isoform based on coding sequence	List of isoforms with identical coding sequences
Dbia_Ank-PB	Dbia_Ank-PC, Dbia_Ank-PD, Dbia_Ank-PE

Isoform report form

Gene--- isoform name (i.e. dmoj_ey--- PA):_dbia_Ank-PB
Names of the isoforms with identical coding sequences as this isoform
_____dbia_Ank-PC, dbia_Ank-PD, dbia_Ank-PE

Is the 5' end of this isoform missing from the end of project:__No_____
Is the 3' end of this isoform missing from the end of project:__No_____

1. Gene Model Checker checklist



Model Details

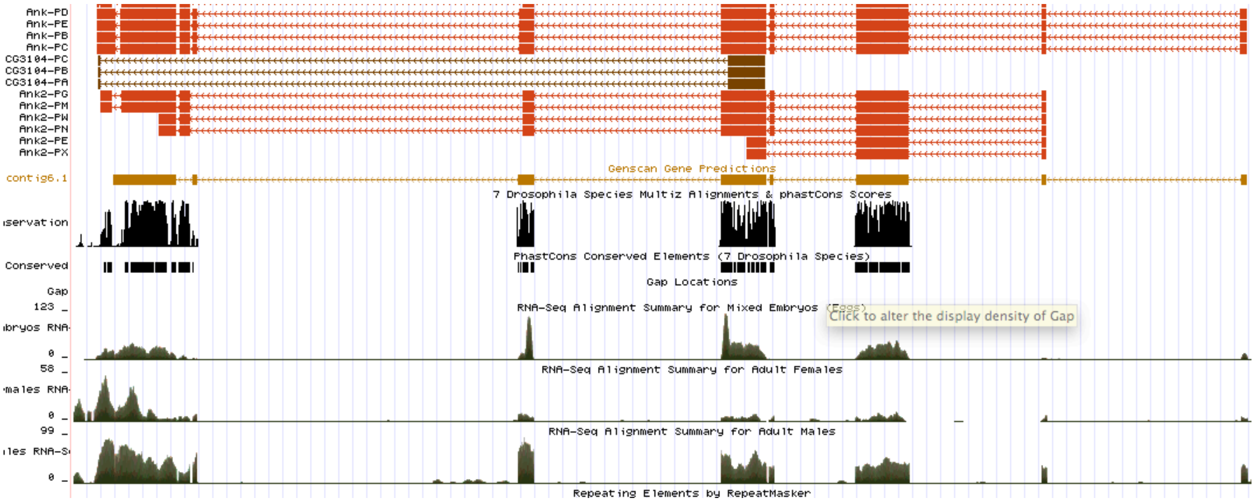
Fosmid Sequence File:	contig6.fasta
Ortholog in D. melanogaster:	Ank-PB
Coding Exon Coordinates:	36965-36849, 32796-32689, 29927-28814, 27119-27016, 26951-26010, 22102-21772, 15085-14992, 14934-14710, 14651-13334
Annotated Untranslated Regions?	○ Yes ● No
Orientation of Gene Relative to Query Sequence:	○ Plus ● Minus
Completeness of Gene Model Translation:	● Complete ○ Partial
Stop Codon Coordinates:	13000-13002

Project Details

Project Group:	D. biarmipes Dot
Project Name:	contig6

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 ...	Pass	

2. View the gene model on the Genome Browser



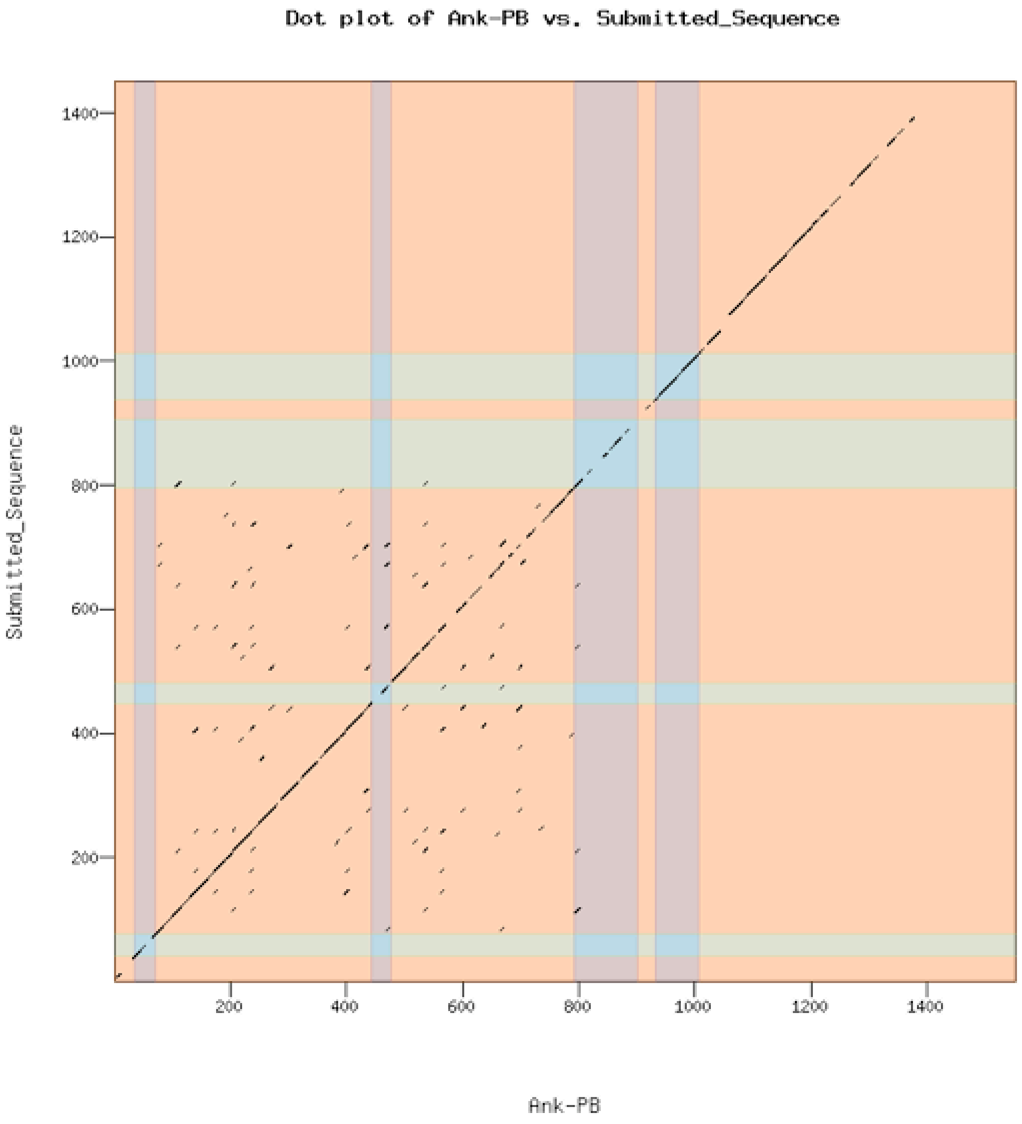
3. Alignment between the submitted model and the *D. melanogaster ortholog*

```
#=======================================
#
# Aligned_sequences: 2
# 1: Ank-PB
# 2: Submitted_Sequence
# Matrix: EBLOSUM62
# Gap_penalty: 12
# Extend_penalty: 2
#
# Length: 1566
# Identity:    1177/1566 (75.2%)
# Similarity:  1305/1566 (83.3%)
# Gaps:         132/1566 ( 8.4%)
# Score: 5823
#
#
#=======================================

Ank-PB             1 MTLGDVRVENKIQTRSETCAI-----NGMALDNKNGIIKQNDATISFLRA     45
                     ||||:|||||||.. :|.||.     |||..|..||.|||||||||||||
Submitted_Seq      1 MTLGEVRVENKIHN-TENCAQTMNINNGMTSDKFNGNIKQNDATISFLRA     49

Ank-PB            46 ARSGDIKKVMDFLDCGEISDINSCNANGLNALHLAAKDGYVDICCELLRR     95
                     ||||||||::|:|:.||||:|||||||||||||||||||:||||||||||
Submitted_Seq     50 ARSGDIKKLIDYLESGEISNINSCNANGLNALHLAAKDGFVDICCELLRR     99

Ank-PB            96 GIKIDNATKKGNTALHIASLAGQHDVINQLILYNANVNVQSLNGFTPLYM    145
                     ||.:|||||||||||||||||||.:||||||..|||||||||||||||||
Submitted_Seq    100 GIAVDNATKKGNTALHIASLAGQQEVINQLIKANANVNVQSLNGFTPLYM    149

Ank-PB           146 AAQENHDNCCRTLLANGANPSLSTEDGFTPLAVAMQQGHDKIVAVLLEND    195
                     |||||||||||.||||||||:||||||||||||||||||:||||||||.|
Submitted_Seq    150 AAQENHDNCCRILLANGANPTLSTEDGFTPLAVAMQQGHEKIVAVLLETD    199

Ank-PB           196 VRGKVRLPALHIAAKKNDVNAAKLLLQHDPNADIVSKSGFTPLHIAAHYG    245
                     |||||||||||||||||||||||||||||.||||||||||||||||||||
Submitted_Seq    200 VRGKVRLPALHIAAKKNDVNAAKLLLQHDQNADIVSKSGFTPLHIAAHYG    249

Ank-PB           246 NVDIATLLLNNKADVNYVAKHNITPLHVACKWGKLSLCTLLLCRGAKIDA    295
                     ||:||||||||.|||||||||||:||||||||||||:|:|||.|||||||
Submitted_Seq    250 NVEIATLLLNNDADVNYVAKHNISPLHVACKWGKLSVCSLLLSRGAKIDA    299

Ank-PB           296 ATRDGLTPLHCASRSGHVEVIKHLLQQNAPILTKTKNGLSALHMAAQGEH    345
                     |||||||||||||||||||||:|||:||||||||||||||||||:|||||
Submitted_Seq    300 ATRDGLTPLHCASRSGHVEVIEHLLRQNAPILTKTKNGLSALHMSAQGEH    349

Ank-PB           346 DEAAHLLLDNKAPVDEVTVDYLTALHVAAHCGHVKVAKLLLDYKANPNAR    395
                     ||||.|||:|||||||||:||||||||||||||||||||||||.||||||
Submitted_Seq    350 DEAARLLLNNKAPVDEVTIDYLTALHVAAHCGHVKVAKLLLDYNANPNAR    399

Ank-PB           396 ALNGFTPLHIACKKNRIKMVELLIKHGANIGATTESGLTPLHVASFMGCI    445
                     ||||||||||||||||||:||||||:|||||||||||||||||||||||:
Submitted_Seq    400 ALNGFTPLHIACKKNRIKVVELLIKYGANIGATTESGLTPLHVASFMGCM    449

Ank-PB           446 NIVIYLLQHEASADLPTIRGETPLHLAARANQADIIRILLRSAKVDAIAR    495
                     |||:||||.||..|:||||||||||||.||||.||||||||||:||||||
Submitted_Seq    450 NIVVYLLQQEAGPDIPTIRGETPLHLATRANQTDIIRILLRSARVDAIAR    499

Ank-PB           496 EGQTPLHVASRLGNINIIMLLLQHGAEINAQSNDKYSALHIAAKEGQENI    545
                     ||||||||||||||||||||||||||||||::||.|||||||||||||.:
Submitted_Seq    500 EGQTPLHVASRLGNINIIMLLLQHGAEINAKTNDMYSALHIAAKEGQEEV    549

Ank-PB           546 VQVLLENGAENNAVTKKGFTPLHLACKYGKQNVVQILLQNGASIDFQGKN    595
                     |||||:|||:.:|:|||||||||||.||||..:|.:|::.|.:::|||||
Submitted_Seq    550 VQVLLDNGAQLDAITKKGFTPLHLASKYGKAKIVTLLIERGCAVNFQGKN    599

Ank-PB           596 DVTPLHVATHYNNPSIVELLLKNGSSPNLCARNGQCAIHIACKKNYLEIA    645
                     |||.|||||||::..:||:|||||:.|||||||||.|:|||.|:|.||||
```

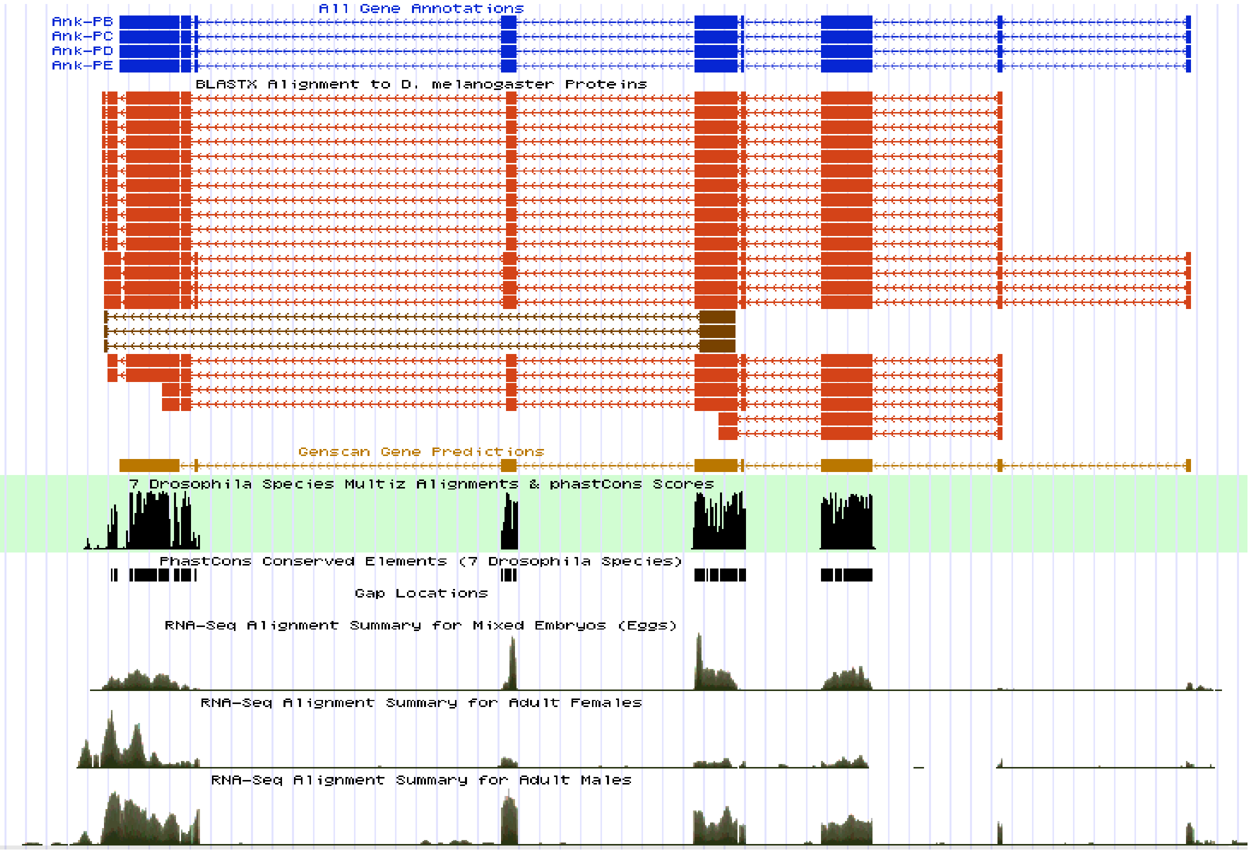
```
Submitted_Seq    600 DVTSLHVATHYDHHDVVEILLKNGALPNLCARNGQSALHIAAKQNNLEIA     649

Ank-PB           646 MQLLQHGADVNIISKSGFSPLHLAAQGGNVDMVQLLLEYGVISAAAKNGL     695
                     :|||||||||.:||||||||||||||.|:||||:||||||..|.|||:||
Submitted_Seq    650 LQLLQHGADVVVISKSGFSPLHLAAQEGHVDMVELLLEYGATSTAAKSGL     699

Ank-PB           696 TPLHVAAQEGHVLVSQILLEHGANISERTRNGYTPLHMAAHYGHLDLVKF     745
                     ||||:|||.|||.|.:|||||||.|||||:|||:|||:|||.||||||||
Submitted_Seq    700 TPLHLAAQGGHVHVCRILLEHGAKISERTKNGYSPLHIAAHNGHLDLVKF     749

Ank-PB           746 FIENDADIEMSSNIGYTPLHQAAQQGHIMIINLLLRHKANPNALTKDGNT     795
                     .:|||||||||:|:||||||||||||||||||||||:||||||||.:|||
Submitted_Seq    750 LLENDADIEMSTNVGYTPLHQAAQQGHIMIINLLLRNKANPNALTNNGNT     799

Ank-PB           796 ALHIASNLGYVTVMESLKIVTSTSVINSNIGAIEEKLKVMTPELMQETLL     845
                     ||||||||||:||||:||.|||||::::|.|.|:||.|||:||:|.||||
Submitted_Seq    800 ALHIASNLGYITVMETLKTVTSTSIVDTNSGVIQEKCKVMSPEVMHETLL     849

Ank-PB           846 SDSDDESCDDLLDHNHYKYMATDDLKANYGQDQKNFDTTNTDHDLTDVSV     895
                     ||.:||..|||||.|||||||||||||.|.|:.:||||||.:||.||||.
Submitted_Seq    850 SDCEDEVGDDLLDQNHYKYMATDDLKATYIQEHQNFDTTNPEHDQTDVSA     899

Ank-PB           896 --LNKKEILPNEMSCIELTEIGHKPDNVVIARSQVHLGFLVSFLVDARGG     943
                       :.:|||..|.:..|..:|:..|||||.|||.:|.||||||||||||||
Submitted_Seq    900 SPIVEKEIQHNGIKFIRTSEMELKPDNVPIARPRVQLGFLVSFLVDARGG     949

Ank-PB           944 SMRGYRHNGVRIIVPPKACAEPTRITCRYVKPQRVVNPPPLMEGEALVSR     993
                     |||||||:||||||||||||||||:|||||||.|||||||||||||||||
Submitted_Seq    950 SMRGYRHSGVRIIVPPKACAEPTRLTCRYVKPHRVVNPPPLMEGEALVSR     999

Ank-PB           994 ILEMSPVDGMFLSPITLEVPHYGTLRKNEREIIILRSDNGESWREHNLYK    1043
                     |||||||||:|||||||||||:||||:|||||||||||:||||||||||:
Submitted_Seq   1000 ILEMSPVDGIFLSPITLEVPHFGTLRENEREIIILRSDDGESWREHNLYE    1049

Ank-PB          1044 D-------IIGE---DINQTEEFHSDRIVRIVTQNVPHFFAVVSRVRQEV    1083
                     .       ..||   |.||.:|.|:||||||||||:||||||||||||||
Submitted_Seq   1050 SEGNINDFYKGEHVGDFNQMDELHTDRIVRIVTQNIPHFFAVVSRVRQEV    1099

Ank-PB          1084 HVIGPDGGTVFSTAVPQVKAIFPPHALTKKIRVGLQAQSVDLVECSKLLG    1133
                     |.||||||||||||||||:|||||||||||||||||||.||||.||||||
Submitted_Seq   1100 HAIGPDGGTVFSTAVPQVQAIFPPHALTKKIRVGLQAQPVDLVGCSKLLG    1149

Ank-PB          1134 QGVAVSPVVTVEPRRRKFHKAITLSIPAPKACTNSMVNACYGNGNSSSPT    1183
                     |||||||:|||||||||||||||||||||||||.||||||||||||:|||
Submitted_Seq   1150 QGVAVSPIVTVEPRRRKFHKAITLSIPAPKACTKSMVNACYGNGNSNSPT    1199

Ank-PB          1184 LRLLCSISGGQTRATWEDVTGSTPLSFVRDSVTFTTTVSARFWLIDCRNI    1233
                     |||||||:||||||||||||||||||||:||:||||||||||||:||||:
Submitted_Seq   1200 LRLLCSITGGQTRATWEDVTGSTPLSFVKDSITFTTTVSARFWLMDCRNV    1249

Ank-PB          1234 IDAGRMATELYSHLAKVPFYVKFVIFAKRISQTEAKFSVFCMTDDKEDKT    1283
                     :||||||:||||||.||||.||||:||||:|:.|||||||||||||||||
Submitted_Seq   1250 VDAGRMASELYSHLTKVPFLVKFVVFAKRLSKIEAKFSVFCMTDDKEDKT    1299

Ank-PB          1284 LEQQEYFKEVAKSRDIEVLQNQIVYLEFAGNIVPILKKGEQLYTKFQPFC    1333
                     |||||||.||||||||||||:|.:||||||:|||:||.||||.||||.||
Submitted_Seq   1300 LEQQEYFTEVAKSRDIEVLQDQNIYLEFAGSIVPVLKTGEQLSTKFQAFC    1349

Ank-PB          1334 ENRLSFSAHIKDQEFPHGRICFMTYPMVGPDEVPLKPLCTLNISVDFKTI    1383
                     ||||||.|||||.|||.|||||||.|..||||:|..||||||:|||.||:
Submitted_Seq   1350 ENRLSFIAHIKDPEFPQGRICFMTNPKAGPDELPQNPLCTLNVSVDLKTM    1399

Ank-PB          1384 TNHLERDNLHSLNDCINAHGKLNHNENIVFGVKEQQVKKIDITKACIMSS    1433
                     |.|||     ::|:.::..|.:|:|:.:.   .|::      |:......
Submitted_Seq   1400 TKHLE-----NINESVSLSGNMNYNDLLS---SEKR------TETAPSEK    1435

Ank-PB          1434 DIKLIHEADVILDDICSHLGSDWPLLANVLGVSQADIDLVKTEFLLNDSV    1483
                     |||                                  .|:..||:.|.:
Submitted_Seq   1436 DIK----------------------------------KLILPEFIRNSA-    1450

Ank-PB          1484 KQSMAMLQLWLEHGGILTGNVLAEALYKIGRSDIVEKSFKNAEFGTHQPE    1533
```

```
      Submitted_Seq   1451 --------------------------------------------------   1450

      Ank-PB          1534 KVLPTAGIEKDGDYEQ   1549
                                          :
      Submitted_Seq   1451 ---------------K   1451


      #---------------------------------------
#---------------------------------------
```

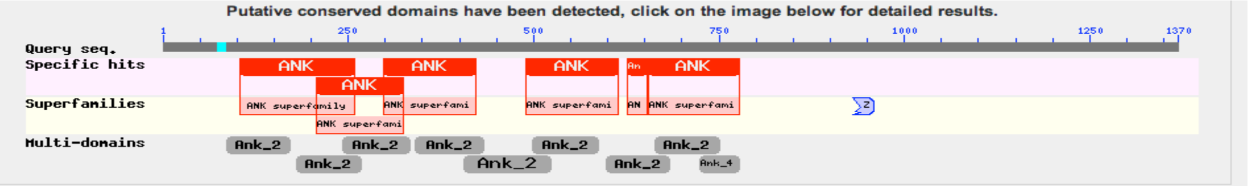
4. Dot plot between the submitted model and the *D. melanogaster ortholog*

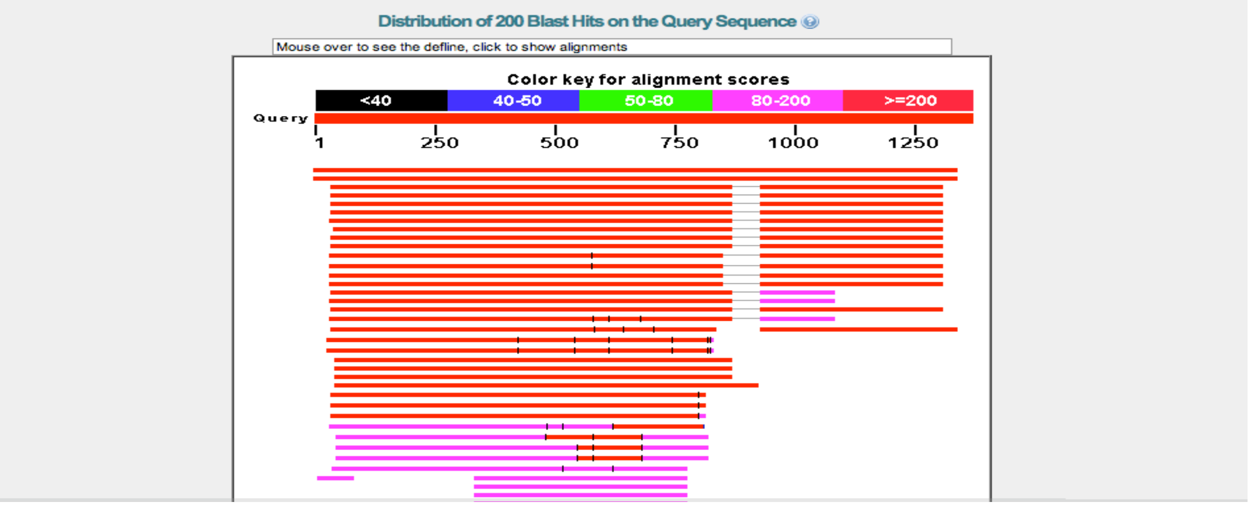


Preparing the project for submission

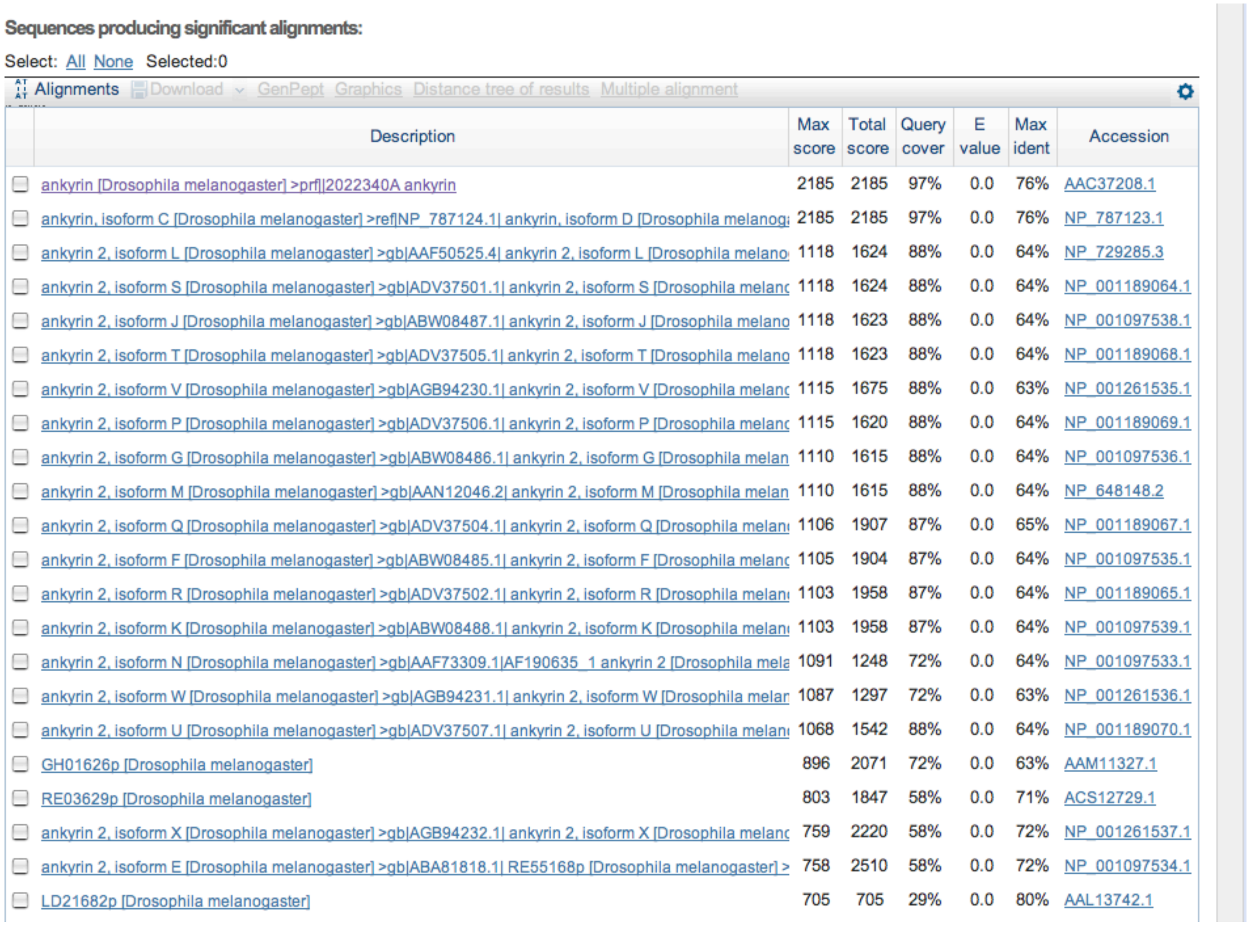


Have you annotated all the genes?


Putative conserved domains have been detected, click on the image below for detailed results.

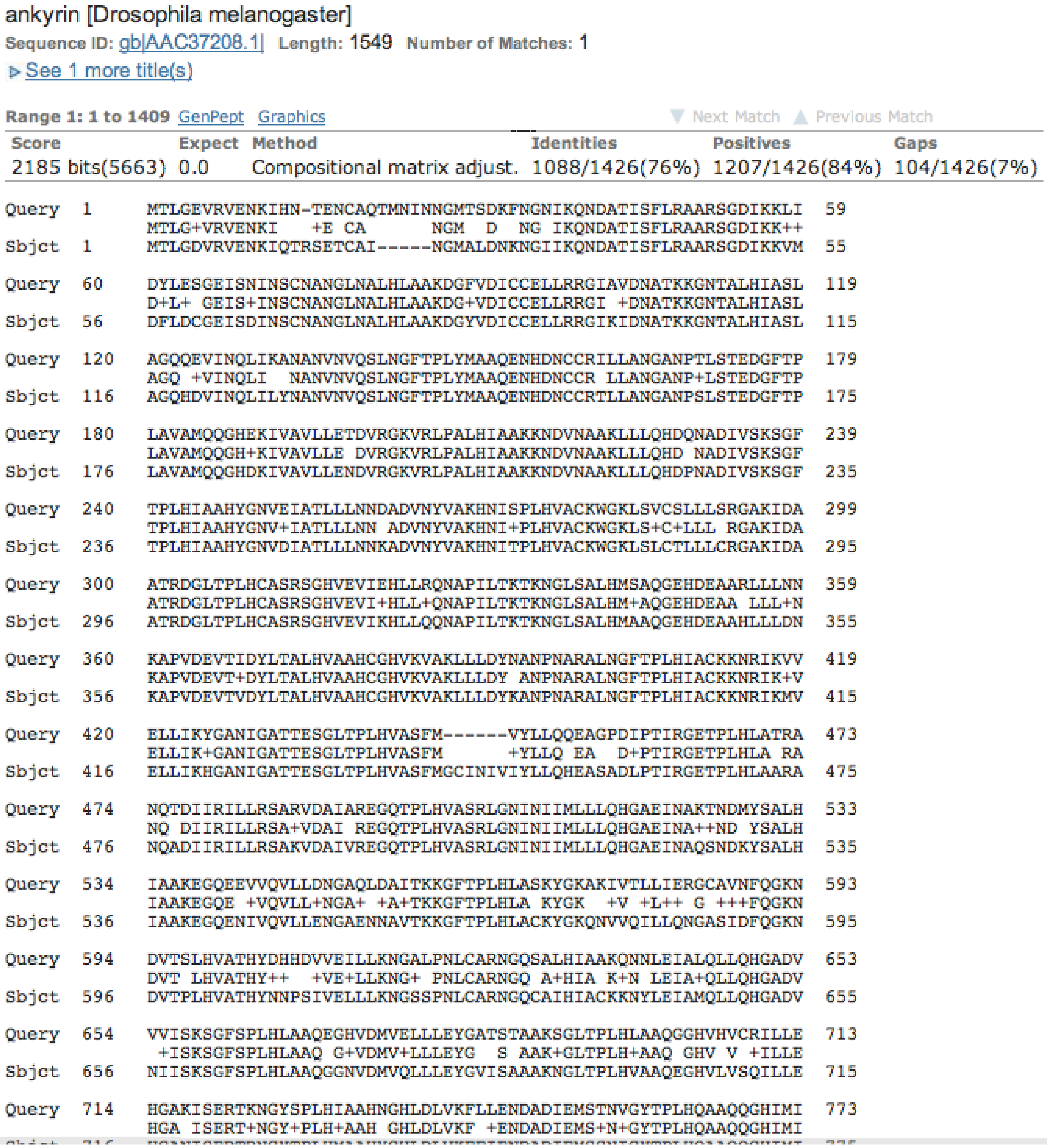

Distribution of 200 Blast Hits on the Query Sequence

Mouse over to see the defline, click to show alignments



	Description	Max score	Total score	Query cover	E value	Max ident	Accession
☐	ankyrin [Drosophila melanogaster] >prf||2022340A ankyrin	2185	2185	97%	0.0	76%	AAC37208.1
☐	ankyrin, isoform C [Drosophila melanogaster] >ref|NP_787124.1| ankyrin, isoform D [Drosophila melanog	2185	2185	97%	0.0	76%	NP_787123.1
☐	ankyrin 2, isoform L [Drosophila melanogaster] >gb|AAF50525.4| ankyrin 2, isoform L [Drosophila melano	1118	1624	88%	0.0	64%	NP_729285.3
☐	ankyrin 2, isoform S [Drosophila melanogaster] >gb|ADV37501.1| ankyrin 2, isoform S [Drosophila melan	1118	1624	88%	0.0	64%	NP_001189064.1
☐	ankyrin 2, isoform J [Drosophila melanogaster] >gb|ABW08487.1| ankyrin 2, isoform J [Drosophila melano	1118	1623	88%	0.0	64%	NP_001097538.1
☐	ankyrin 2, isoform T [Drosophila melanogaster] >gb|ADV37505.1| ankyrin 2, isoform T [Drosophila melano	1118	1623	88%	0.0	64%	NP_001189068.1
☐	ankyrin 2, isoform V [Drosophila melanogaster] >gb|AGB94230.1| ankyrin 2, isoform V [Drosophila melan	1115	1675	88%	0.0	63%	NP_001261535.1
☐	ankyrin 2, isoform P [Drosophila melanogaster] >gb|ADV37506.1| ankyrin 2, isoform P [Drosophila melan	1115	1620	88%	0.0	64%	NP_001189069.1
☐	ankyrin 2, isoform G [Drosophila melanogaster] >gb|ABW08486.1| ankyrin 2, isoform G [Drosophila melan	1110	1615	88%	0.0	64%	NP_001097536.1
☐	ankyrin 2, isoform M [Drosophila melanogaster] >gb|AAN12046.2| ankyrin 2, isoform M [Drosophila melan	1110	1615	88%	0.0	64%	NP_648148.2
☐	ankyrin 2, isoform Q [Drosophila melanogaster] >gb|ADV37504.1| ankyrin 2, isoform Q [Drosophila melan	1106	1907	87%	0.0	65%	NP_001189067.1
☐	ankyrin 2, isoform F [Drosophila melanogaster] >gb|ABW08485.1| ankyrin 2, isoform F [Drosophila melan	1105	1904	87%	0.0	64%	NP_001097535.1
☐	ankyrin 2, isoform R [Drosophila melanogaster] >gb|ADV37502.1| ankyrin 2, isoform R [Drosophila melan	1103	1958	87%	0.0	64%	NP_001189065.1
☐	ankyrin 2, isoform K [Drosophila melanogaster] >gb|ABW08488.1| ankyrin 2, isoform K [Drosophila melan	1103	1958	87%	0.0	64%	NP_001097539.1
☐	ankyrin 2, isoform N [Drosophila melanogaster] >gb|AAF73309.1|AF190635_1 ankyrin 2 [Drosophila mela	1091	1248	72%	0.0	64%	NP_001097533.1
☐	ankyrin 2, isoform W [Drosophila melanogaster] >gb|AGB94231.1| ankyrin 2, isoform W [Drosophila mela	1087	1297	72%	0.0	63%	NP_001261536.1
☐	ankyrin 2, isoform U [Drosophila melanogaster] >gb|ADV37507.1| ankyrin 2, isoform U [Drosophila melan	1068	1542	88%	0.0	64%	NP_001189070.1
☐	GH01626p [Drosophila melanogaster]	896	2071	72%	0.0	63%	AAM11327.1
☐	RE03629p [Drosophila melanogaster]	803	1847	58%	0.0	71%	ACS12729.1
☐	ankyrin 2, isoform X [Drosophila melanogaster] >gb|AGB94232.1| ankyrin 2, isoform X [Drosophila melan	759	2220	58%	0.0	72%	NP_001261537.1
☐	ankyrin 2, isoform E [Drosophila melanogaster] >gb|ABA81818.1| RE55168p [Drosophila melanogaster] >	758	2510	58%	0.0	72%	NP_001097534.1
☐	LD21682p [Drosophila melanogaster]	705	705	29%	0.0	80%	AAL13742.1

Though Ank and Ank2 are shown as matches for this BLASTp, the Ank2 results were disregarded because the Ank2 gene is on Chromosome 3 and Ank is actually found on the Dot Chromosome.



ankyrin [Drosophila melanogaster]

Sequence ID: gb|AAC37208.1| **Length:** 1549 **Number of Matches:** 1

▷ See 1 more title(s)

Range 1: 1 to 1409 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
2185 bits(5663)	0.0	Compositional matrix adjust.	1088/1426(76%)	1207/1426(84%)	104/1426(7%)

```
Query  1    MTLGEVRVENKIHN-TENCAQTMNINNGMTSDKFNGNIKQNDATISFLRAARSGDIKKLI  59
            MTLG+VRVENKI   +E CA      NGM  D  NG IKQNDATISFLRAARSGDIKK++
Sbjct  1    MTLGDVRVENKIQTRSETCAI-----NGMALDNKNGIIKQNDATISFLRAARSGDIKKVM  55

Query  60   DYLESGEISNINSCNANGLNALHLAAKDGFVDICCELLRRGIAVDNATKKGNTALHIASL  119
            D+L+ GEIS+INSCNANGLNALHLAAKDG+VDICCELLRRGI +DNATKKGNTALHIASL
Sbjct  56   DFLDCGEISDINSCNANGLNALHLAAKDGYVDICCELLRRGIKIDNATKKGNTALHIASL  115

Query  120  AGQQEVINQLIKANANVNVQSLNGFTPLYMAAQENHDNCCRILLANGANPTLSTEDGFTP  179
            AGQ +VINQLI  NANVNVQSLNGFTPLYMAAQENHDNCCR LLANGANP+LSTEDGFTP
Sbjct  116  AGQHDVINQLILYNANVNVQSLNGFTPLYMAAQENHDNCCRTLLANGANPSLSTEDGFTP  175

Query  180  LAVAMQQGHEKIVAVLLETDVRGKVRLPALHIAAKKNDVNAAKLLLQHDQNADIVSKSGF  239
            LAVAMQQGH+KIVAVLLE DVRGKVRLPALHIAAKKNDVNAAKLLLQHD NADIVSKSGF
Sbjct  176  LAVAMQQGHDKIVAVLLENDVRGKVRLPALHIAAKKNDVNAAKLLLQHDPNADIVSKSGF  235

Query  240  TPLHIAAHYGNVEIATLLLNNDADVNYVAKHNISPLHVACKWGKLSVCSLLLSRGAKIDA  299
            TPLHIAAHYGNV+IATLLLNN ADVNYVAKHNI+PLHVACKWGKLS+C+LLL RGAKIDA
Sbjct  236  TPLHIAAHYGNVDIATLLLNNKADVNYVAKHNITPLHVACKWGKLSLCTLLLCRGAKIDA  295

Query  300  ATRDGLTPLHCASRSGHVEVIEHLLRQNAPILTKTKNGLSALHMSAQGEHDEAARLLLNN  359
            ATRDGLTPLHCASRSGHVEVI+HLL+QNAPILTKTKNGLSALHM+AQGEHDEAA LLL+N
Sbjct  296  ATRDGLTPLHCASRSGHVEVIKHLLQQNAPILTKTKNGLSALHMAAQGEHDEAAHLLLDN  355

Query  360  KAPVDEVTIDYLTALHVAAHCGHVKVAKLLLDYNANPNARALNGFTPLHIACKKNRIKVV  419
            KAPVDEVT+DYLTALHVAAHCGHVKVAKLLLDY ANPNARALNGFTPLHIACKKNRIK+V
Sbjct  356  KAPVDEVTVDYLTALHVAAHCGHVKVAKLLLDYKANPNARALNGFTPLHIACKKNRIKMV  415

Query  420  ELLIKYGANIGATTESGLTPLHVASFM------VYLLQQEAGPDIPTIRGETPLHLATRA  473
            ELLIK+GANIGATTESGLTPLHVASFM      +YLLQ EA  D+PTIRGETPLHLA RA
Sbjct  416  ELLIKHGANIGATTESGLTPLHVASFMGCINIVIYLLQHEASADLPTIRGETPLHLAARA  475

Query  474  NQTDIIRILLRSARVDAIAREGQTPLHVASRLGNINIIMLLLQHGAEINAKTNDMYSALH  533
            NQ DIIRILLRSA+VDAI REGQTPLHVASRLGNINIIMLLLQHGAEINA++ND YSALH
Sbjct  476  NQADIIRILLRSAKVDAIVREGQTPLHVASRLGNINIIMLLLQHGAEINAQSNDKYSALH  535

Query  534  IAAKEGQEEVVQVLLDNGAQLDAITKKGFTPLHLASKYGKAKIVTLLIERGCAVNFQGKN  593
            IAAKEGQE +VQVLL+NGA+ +A+TKKGFTPLHLA KYGK  +V +L++ G +++FQGKN
Sbjct  536  IAAKEGQENIVQVLLENGAENNAVTKKGFTPLHLACKYGKQNVVQILLQNGASIDFQGKN  595

Query  594  DVTSLHVATHYDHHDVVEILLKNGALPNLCARNGQSALHIAAKQNNLEIALQLLQHGADV  653
            DVT LHVATHY++  +VE+LLKNG+ PNLCARNGQ A+HIA K+N LEIA+QLLQHGADV
Sbjct  596  DVTPLHVATHYNNPSIVELLLKNGSSPNLCARNGQCAIHIACKKNYLEIAMQLLQHGADV  655

Query  654  VVISKSGFSPLHLAAQEGHVDMVELLLEYGATSTAAKSGLTPLHLAAQGGHVHVCRILLE  713
             +ISKSGFSPLHLAAQ G+VDMV+LLLEYG  S AAK+GLTPLH+AAQ GHV V +ILLE
Sbjct  656  NIISKSGFSPLHLAAQGGNVDMVQLLLEYGVISAAAKNGLTPLHVAAQEGHVLVSQILLE  715

Query  714  HGAKISERTKNGYSPLHIAAHNGHLDLVKFLLENDADIEMSTNVGYTPLHQAAQQGHIMI  773
            HGA ISERT+NGY+PLH+AAH GHLDLVKF +ENDADIEMS+N+GYTPLHQAAQQGHIMI
```