

In the format provided by the authors and unedited.

Non-immunogenic utrophin gene therapy for the treatment of muscular dystrophy animal models

Yafeng Song^{1,8}, Leon Morales^{1,8}, Alock S. Malik^{1,8}, Andrew F. Mead¹, Christopher D. Greer¹, Marilyn A. Mitchell¹, Mihail T. Petrov¹, Leonard T. Su¹, Margaret E. Choi¹, Shira T. Rosenblum¹, Xiangping Lu¹, Daniel J. VanBelzen¹, Ranjith K. Krishnankutty¹, Frederick J. Balzer¹, Emanuele Loro², Robert French¹, Kathleen J. Propert³, Shangzhen Zhou⁴, Benjamin W. Kozyak^{1,5}, Peter P. Nghiem⁶, Tejvir S. Khurana², Joe N. Kornegay⁶ and Hansell H. Stedman^{1,7*}

¹Department of Surgery, Pennsylvania Muscle Institute, Penn Cardiovascular Institute, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ²Department of Physiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ³Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁴Department of Ophthalmology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁵Department of Anesthesiology and Critical Care Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. ⁶Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Texas A&M Institute for Neuroscience, Texas A&M University, College Station, TX, USA. ⁷Corporal Michael J. Crescenz Veterans Affairs Medical Center, Philadelphia, PA, USA.

⁸These authors contributed equally: Yafeng Song, Leon Morales, Alock S. Malik. *e-mail: hstedman@pennmedicine.upenn.edu

Species common name	Scientific name	abbreviation	genomic DNA accession number or filename	version	RNA transcript	polypeptide
Dystrophin homologs						
Purple sea urchin	Strongylocentrotus purpuratus	S.pur	NW_003577129 REGION: complement(169587..521351) GPS_000671889	NW_003577129.1 GI:346420255	NM_214496.1 GI:47550974	NP_999661.1 GI:47550975
Amphioxus, Florida lancelet	Branchiostoma floridae	B.flo	Brafl_scaffold_51 complement 1-302248	v1.0 JGI download	file "Bflo scf51 dys-u-like mRNA.docx"	ORF within file "Bflo scf51 dys-u-like mRNA.docx"
Amphioxus, Florida lancelet	Branchiostoma floridae	B.flo	Brafl_scaffold_194 complement 500000-805591	v1.0 JGI download	file "Bflo scf194 dys-u-like mRNA.docx"	ORF within file "Bflo scf194 dys-u-like mRNA.docx"
Sea lamprey	Petromyzon marinus	P.mar		GI:501301.1 GI:308126735 and	file "Pmar dys-u-like mRNA.docx"	ORF within file "Pmar dys-u-like mRNA.docx"
Elephant shark	Callorhynchus milii	C.mil	complement scaffold_GL501301 and complement scaffold_GL476499	GI:476499.1 GI:308151537		
			NW_006890171 REGION: complement(1..2330292) GPS_003798258	NW_006890171.1 GI:616174015	file "Cmil dystrophin mRNA.docx"	ORF within file "Cmil dystrophin.docx"
			and NW_006890194 REGION: complement(1443720..1732566)	and NW_006890194.1		
			GPS_003798281	GI:616168623		
Chinese alligator	Alligator sinensis	A.sin	NW_005841959 REGION: complement(1345083..3032920)	NW_005841959.1 GI:555311618	XM_006020580.1 GI:557274767	
			GPS_002785379	NC_014778.1 GI:315422583	XM_008107344.1 GI:637282639	XP_008105551.1 GI:637282640
Carolina green anole	Anolis carolinensis	A.car	NC_014778 REGION: 127257227..128903888 GPC_000000424	NC_014778.1 GI:315422583	XM_016943539.2	XP_016799028.1
Common chimpanzee	Pan troglodytes	P.tro	NC_006491 REGION: complement 31151290..33473903	NC_006491.3 GI:291061354		
			GPC_000000313	NC_000086.7 GI:372099090	NM_007868.5 GI:183396799	NP_031894.1 GI:6681203
Common house mouse	Mus musculus	M.mus	NC_000086 REGION: 82310769..85205050 GPC_000000793			
Domestic Dog	Canis familiaris	C.fam	NC_006621 REGION: complement(26290903..28444623)	NC_006621.3 GI:357579592	NM_001003343.1 GI:55742727	NP_001003343.1 GI:55742728
			GPC_000000698			
Human	Homo sapiens	H.sap	NC_000023 REGION: complement(30119222..34339609)	NC_000023.11 GI:568815575	NM_004006.2 GI:238018044	NP_003997.1 GI:5032283
			GPC_000001315			
Utrophin orthologs						
Elephant shark	Callorhynchus milii	C.mil	GPS_003798174	NW_006890087.1 GI:616205308	XM_007894455.1 GI:632953846	XP_007892646.1 GI:632953847
Chinese alligator	Alligator sinensis	A.sin	GPS_002785519	NW_005842099.1 GI:555311478	XM_006024237.1 GI:557282807	XP_006024299.1 GI:557282808
Carolina green anole	Anolis carolinensis	A.car	NC_014776 REGION: 206129426..206617010 GPC_000000422	NC_014776.1 GI:315422585	XM_008122821.1 GI:637255135	XP_008121028.1 GI:637255136
Duck-billed platypus	Ornithorhynchidae anatinus	O.ana	NC_009095 REGION: complement(16542292..17084893)	NC_009095.1 GI:149721824	XM_007656390.1 GI:620942232	XP_007654580.1 GI:620942233
Common house mouse	Mus musculus	M.mus	NC_000076 REGION: complement(12382187..12869360)			
			GPC_000000783	NC_000076.6 GI:372099100	NM_011682.4 GI:110431377	NP_035812.3 GI:110431378
Domestic Dog	Canis familiaris	C.fam	NC_006583 REGION: 35727522..36225533 GPC_000000660	NC_006583.3 GI:357579630	NM_001012395.1 GI:59958374	NP_001012395.1 GI:59958375
Human	Homo sapiens	H.sap	NC_000006 REGION: 144286354..144853034 GPC_000001298	GI:110611227	NM_007124.2 GI:110611227	NP_009055.2 GI:110611228

Bflo scf51 dys-u-like mRNA.docx

LOCUS Untitled 9645 bp ss-DNA linear SYN 02-Dec-2014

DEFINITION -

ACCESSION -

KEYWORDS -

SOURCE Bflo scaffold 51revcomp Hsap dystrophin hmm fgenes+ getfile

FEATURES Location/Qualifiers

 CDS 1..9645

BASE COUNT 2587 A 2524 C 2979 G 1554 T 0 OTHER

ORIGIN ?

```

1 ATGGCTGATC TACAGCAAAG CAACCTGGAG AAAACTCTCC TCGCCTGGTG CAGGCAAAC
61 ACTCAGGGGT ATAGAAGTGT AGACATTGTT AACTTCACCA GTAGCTGGAG GGATGGATAT
121 GCCTTTAATG CACTCATCCA CAAGTACAGG CCAGACCTGT TTGATTATAA TATCCTGTCT
181 AAGATGGGCA CAGAGAAGAG GCTGGAACAC GCCTTCAAGG TGGCCAACGA GCACCTCAAC
241 ATCGACAGAC TCCTCGACCC TGAAGATGTG AATGTGGAAT TCCCAGACAA GAAGTCCATC
301 ATCATGTACG TGACGCTGTT CTTCCAAGTC TTGCCGCACC AGAAGATAAC CCAGGAGGCC
361 ATGCAGGAGG CCCTGGAGGT GTCTGCTCCG GGCTACGCCG CCAAGCAGCA GACGCAGGGG
421 GCCACCCGTG CGAGCCCCGC TCGCCAGAGA AGCGGAGTGG ACCTGGGAGC TTACCAGAAT
481 GCTCTGGAGG AGGTCCTGGC CTGGCTCCTG GCCGCGGAGG AGAAACTGCA GCATCAGGGA
541 GACATCTCCG ACCTGGTGGG TGATGTCAAG GACCAGTTTC ATGAACACGA GAGCTTCATG
601 ATGGAGCTGA CATCCCACCA GAGCCGCGTG GGGAGCGTTC TCCAGGCCGG GAACCAGCTG
661 ATCATGGAGG GGAAGATCAC GGAGGACGAG GAAAACGAGA TCCGGGAGCA GATGAGCCTC
721 CTAACAGCC GCTGGGAGGT GCTCCGCGTC CAGTCCATGG ACAGGCAGTC CAAGTTGCAT
781 GAGGTGCTGA TGAACCTGCA GCAGAAGCAG CTGGACCAGC TCGACGCCTG GCTGTCCGGC
841 ACGGAACAGC GTATCAAGGT CATGGACACG CAGGGCATCG GGCCCGACCT TGACCAGGTC
901 CGCAGGCAGG TGGAAGAACA CAAGACGCTT CAGGAAGGCT TGGAGACTCA GCAACTGCAG
961 GTGAACTCCC TGACTCATAT GGTGGTGGTG GTGGATGAGA ACAGTCCTGA GAATGCCACC
1021 CAAGCCCTGG AGGCACAGCT GCAGAACCTG GGGGAGCGGT GGGCCTGGGT GTGTAAGTGG
1081 ACGGAGGACC GCTGGATGGT GCTACAGGAG GTGCTGGTGC GATGGCAGGT GTTCAACCTG
1141 GAGCAGATGG ACTTCCACGA CTGGCTGACC GACAAGGAAG GGGCTCTGCA GACCATGCAG
1201 CACGCCAACC TCAGCGACCC CAACGTCATT GTTGACACAG TCAAGAAACT AAAGGTACGG
1261 AGAGAAAGGA TTTTAGAACA TGAGATGGAG GCCAAGCAGC CGACCTTCCA GGTGTTGAGC
1321 GGAAGGCTG CGGAGGTTGT GCAGCATCTG GACCGCTCCT CCTCTGCCGT GGAGGGGATT
1381 CAGGGCAAGA TCGAGGAGTT CACGCAGCGC TGGGATGCCA TTGTGCAGAA GATGGAGAAC
1441 TACAGTAAAC TGCTCTCGAT CCCGCAGATG TCTGCCTCGG CTGCCGTCCT GAAGAAGACG
1501 ACCACCATCA CCGAGACGGT CACCAAGATG ACTCAGAAAC TGACCCAGAA AACACCGCCG
1561 CCACCACCTC CTAAGAAAGC AGTGAGCCAG CAGAGAAGGA GGTTTGACAG CGAGATCGCG
1621 ACACTTCTTG CCTGGATCGA CGACACAGAG GCCCTGCTGT TGTCGCCCCG CTACACCAAC
1681 CCAGCTACCT CAAACAGCCT CCAGGCTCAA ACCGACAAGT TCAACCAGCT GGACGAGAAC
1741 ATCATGTCCT ATCAGGGCCG CGTACAGACC ATCCAGGAGA CCGGCGAAAC CCTCATCGAG
1801 GAAGAGGATG ACAAGATTCA GCTGGAGACG CGCCTGAACA CGTTCATGGA GCGCTGGAAC
1861 GAGATCTGCG TGCTGATCAA CGACCGGAAG CGAAAGCTTC AGATGGCGCG ACTGTGTAAG
1921 CAGTTCTACA CGGAGCTGGA CGTCCTGAGG GACGTCACCG GGAGGAAGAT GACCTGGCTG
1981 CAGGGACAGG ACAAGCCCAG CGACGATGCA GTCAAGATTA AACAGCAGAT GGATGAACTA
2041 CAGGTTGAAG TTAAGTACAT GGCAGCACAC AGTGAGCGTG TAGAGCAGCT GAACAGCCTG
2101 GGCTCTCAGC TGGCGGACTC CTCAGACCCG CCAAGAACC TGCAGCAGGA CCTGACTCAG
2161 TTCCAGCTGC AGTGGGAGGA ACTGGTGGCC AAGATGGCCG CCAGGGAGAA GGACCTCACC
2221 AAAGCTCTTG AGCAGGCCCC TCCTAAACAG TACCTGGAGG CCATGAAGGC GCTGACACAG
2281 TGGATCGATA ACGTGGAGTC TCTCCTCACC ACGGAGAAAG TCTACATCGC TGACATCCCC
2341 ACCATGAAGG ATCAACAGCA GAAGTTTAAG GAGGTACAGC AGGCCCTGAA GGACCAGCAG
2401 AAGAGTCTGG ACTTCGTGAA CACCACGGGG AAGCAGGTGG TGGAGAGTTC CCCGCCGGAC
2461 CGGTCCCGCG GCCTGCAGAA GGAGCTGAAC GGGCTGAACG AGCGCTGGAG TAACGTCTGC
2521 GCCATGACAG AGGATCGGCT GGAACACCTG GACAAGGGGG TCGAGCTCCT CACGCAATTT
2581 CAGAAAGAAG GTGATGGCCT GCGTACCTGG ATGGATGAGG TGGATGTGTA CATCCACTCA
```

2641	GAAGAGGCAG	CAGTAGGTGA	CACAGAAACC	CTACAGGCTC	AGCTAGAACA	AAGCAAGTCC
2701	CTTCAGAAGG	ACATCGCCTC	CCTCCAGCAC	AACATGAACA	ACGTGAACGA	GGCGGGGAAG
2761	AGGCTGGTGC	CTGAGGCGCA	CCCCGACTTC	TCCAAGAAAC	TCAAGCAGGA	GCTGAAGGAC
2821	CTGAACGAAC	GCTGGGACGC	CATCACCAAG	AAGCAAAAGT	CTCAGAAGCA	GAGTCTTGAG
2881	GGAGCGATGA	AGAAGACGAG	CTCTTTACAG	CTGCAGCTGG	AGGCCTTCAG	CGAGTGGCTG
2941	ACTGACAAGG	AGAAGGAGGG	ACTGAAGGAA	GACGGGCCTC	TGGAGGACCC	TGATGACATC
3001	CAGCCCAGGA	TCAAGAAATA	TAAGGCCCTT	GAGAACCACG	TGCGGAACCA	CCCGTCAGAC
3061	AACAGGGACA	AGATCCAGGA	GATCTCCAG	CAGCTGGTGG	AGCACAAGGT	TCTGGTGGAG
3121	CCTGTACAGA	CGGAGGTCAC	GACCTTCACC	TCACGATGGG	ACACTCTGCA	GCAACAGGCT
3181	AAGCTGCGCC	AGACAGCCCT	TGAGCAGCGG	ATTGCCAGTG	CCCAGGACCT	GGAGAAGGAT
3241	GTGCTTGAGC	TGGAGCGCTG	GTTACACGG	ACAGACAAGC	TTCTGCACTC	GCGCATCACC
3301	TCGGGCGTCT	GCTCCTCCGA	TGTGCCTGAG	GAGTTTGAGC	AACTACAGGA	CGAGTTTGAT
3361	GCTCGAGACG	AAGACCTGGA	GGAGATCGGG	CGGAAATGTC	AGGCACTGAT	GACGGAGGGC
3421	AGCACAGAGG	CCACCACCAG	GATCAGGCAG	CAGTCTGAGA	TACTGCAGAA	GCGAATGGTT
3481	GAGATAAAGG	CAAAGTTCCG	CAAGTACCAG	AAGCCGGCAG	ACTTCGAGCC	GAAGATGGGT
3541	CGCGTTCGAG	ACATCCTGAA	CAACGTTGAG	AAGGAGCTTG	GAGTGCTGGA	GCTCAAGCAC
3601	CATGAGCCTG	AGATGGTGCA	GGGGCAACTC	GACATCTGTA	TGGGCCTGTA	CAAGATCCTG
3661	AGCGATGTGA	AACCTGAGGT	GACGGAAGGA	AAACAGGATC	TTGAGAAGGC	CATGAAGCTG
3721	TCTAAGAAGC	TAAAGAAGGA	GATGGTGTCC	CTTACGGAGT	GGCTGGGCAC	TACCGAGGCC
3781	GAGCTCAACA	AGAAGGAGGC	AGAGGGCGCC	ATCCCGGCTA	ATGTTGAAGA	GGAACCTAAG
3841	TGGTGCAAGA	CCCTTCAGCA	AGAGATGACC	AGGAAACAGT	CCGTGGTGCT	GATTATCAAT
3901	GAGACAGGGA	AGGAGATGGT	GGAACTCACC	CTTCAGCAAG	AGATGACCAG	GAAGCAGTCC
3961	GTGGTACTGA	TTATCAATGA	GACAGGGAAG	GAGATGGTGG	AACTCGCGGA	ACCAGGAACT
4021	CTCGGTCCGC	TGGAGGAAAA	ACTCTCCGCC	CTGAACCAGC	ACTGGTCGGA	CGTGTGCGCC
4081	CGGGCAGCTG	AGAGGAGGCT	CAGGCTTCAG	GAGTTTCATGT	CTGAGATGGA	GAAGTTCGGG
4141	GAGGGTCTTA	ACGAGGTGCA	CACCTGGATG	GAGAAGGTGG	AGATGTTCTT	TAAGATCTCC
4201	GACAGACTTA	GTCCCCAGGA	ACAGGCCAAG	GAGGTGGACA	CGTACAAGAA	GCTTCAGCTG
4261	GAAGCGAACC	GGTTGCGCGG	CACAGTGGAC	CAGGTGCGAG	ACCAGGCCAT	GGAGCTGATG
4321	CAGCACGGGC	CGGCTGTGCA	GGGCATGGTG	GAGCCAGAAC	TCATCAACCT	CAACAGCAGA
4381	TGGGAGAGTG	TGTGTGAGAA	GATCAGGGAG	AAGCAGCAGG	TCCAGCCATT	GCAGCAAGTC
4441	CAAGTTCAAC	AGGCCGTTGC	CGTGGCGTCA	GCCGCTCCC	CTCCACCAGT	CCCCGCACGC
4501	ACCATGAAGT	CCGAAACCAC	CACTGTCACC	ACGGTTACCA	CCGTCACCAC	GACGACGAAG
4561	ATCATCGAGT	TGGAGACGGA	CTTCCACAAG	GCCAGCTCTG	TCATCGCTCA	GTTTCGACGAC
4621	ACGTTACTGA	CGGACAGCGG	AGACTTGCGT	GAGGAAGACT	TCACCGATAA	TGTGGAGGAC
4681	AGGGTCAAGG	CAACCGAGGA	TGAAATCTCA	AGAATACAGG	CCGAGATCGA	CGACATCTTC
4741	GAGAAGGGAG	ACAAAGCCAT	CAAGGAGACC	GACGACTCGA	CCAAGAAGGA	GCAGATTTCGT
4801	CTTAAACTGA	ACGACCTTAA	AGGAAGATGG	AGCTCGGTGA	AGGATGATGC	TGAGCTGAAG
4861	AAGAAGACCC	TGCGGACGGT	GGTGCCACAG	TGGTACCAGT	TCCGTTCCCA	GACCGATGAT
4921	CTGAACATGT	ATCTCACGGA	GATCGAGATC	AGAATACGGG	AGGCCAAGGA	TGATCCCAGC
4981	AAGCTACAGG	CTATCCATGA	GGAGCTTCAG	CATAAGCAGG	AGGACCTGGA	CAGGCTGAAC
5041	AGGAGAGCCG	AGGAGCTGAA	ACAGCAGGGG	GCCAAGCCTG	TGGCCATAAA	GGAGCAGCTA
5101	GACACGTTGG	AGCCACAGGT	GCGGTCGGCC	CACGAGCGCA	AGGGGCCCGC	CATCTCTCAG
5161	GTGGCGCCCT	CCGAACGGAA	GAATATCGAG	AATGCCATGT	CTAATCTCAA	GGTCGAGTGG
5221	GACAAGGTCA	ACAAGGATTA	TGTGGACAGA	TACAGTCGCT	TTAAGAAGTC	ACTGGCGCAC
5281	TGGCAGCAGT	TCCACTGTGA	CATGCTCGCT	TTAAGAAGTC	ACTGGCGCAC	TGGCAGCAGT
5341	TCCACTAAAT	TGCAAGACGG	GATTGAGAAG	CACAGAAAGC	AAGTGGAGGG	ACTGAACGCA
5401	AACGGGCAGG	AGATCGTTGA	GCAGTGCACG	TCCATCGACG	CAACCCTTCT	GAGGGAAAAG
5461	CTAGACGGGC	TGAACAGAAAG	GTGGAGGGTG	CTGTGCGCAG	AGGTCAACGA	CAGGAAAAGAA
5521	AGGTTGGAGG	AGGCGAGCAT	GCAGACAGAG	GAGTTCCAGG	AAGAGATTGA	CGAGCTGCTG
5581	TTTTGGCTGG	AGGACACAGA	GGGCATCCTT	AACAAGCAGG	TGTGCCCAGC	TGACGAGGAT
5641	CAGCTAGAGG	AAGCTCTGGA	GAAAGTTAAG	GACAAGGAGA	ATGACCTTGG	ATCGAGACAG
5701	GACAACGTGA	ACACCATCAA	CCAGACGGGA	AACCGGATCA	TGACGACGCC	GGGCGCCTCA
5761	AAAGCTCTGG	TCAGCAACGT	CAAGAAGAAG	GTCGACAACC	TCAACACAAG	ATGGCTGACG
5821	GTGACGCCAG	AGATTCCCCA	GAAGAGGAGA	GGACTGCAGG	ACCAGCTGCA	CTTCTGTTCG

5881	ACCTTCCTGG	AGGAGCTGGA	GGAGCTTCTG	GTGTGGATGG	CTCACACGAT	GCAGTTGCTA
5941	GAGCAACAGA	AGGGACCCAT	CTCCTCCGCA	ACATCTTCTG	ATGAGGGGGA	TCAGGTCATC
6001	ATCGACCCAA	AGACGATGAA	AGATGCCCTG	ACGGCCCGTC	AATCAAACGT	CGATAACGTC
6061	AACCTCACGT	ATAACAGCCT	CGTAGCCGAG	AGCAAGGACC	TGGAAGTGGA	CGTCCCGGAC
6121	TACGTCAGAG	ACAAGATCAG	GAAGCTGAAC	GCCGACTGGG	CCAAGATCCG	CTACATGGCG
6181	GCCAACCTGA	GGCCGACCTC	GGATTCCGAC	GTGGAATCTC	TCGTTATGGA	GCAGAGAAGG
6241	TCTATGGAGG	AGATGGAAAC	CGCTAGGATA	GAAGTCGGAG	CCATGCATGC	GAGGGCGGAC
6301	ATGCCGTCGA	TGTATCCCGA	CTTCGACAAG	TCCGTGGCGG	AGCTCCGTGA	TTGGCTGATG
6361	CTCCTCGACC	ACATGCTGAA	GTCTCAGATT	GTCACCGTGG	GAGACGTCGA	GGAGATCGAA
6421	GAGATGATCA	TCAAGCAGAA	GACCGTCTTA	AACGACCTTG	ATGTCAAGAA	GCCCCAGCTG
6481	GATACGGTGG	TGGCTGCCGG	CAGACGGCTC	ATTCAAACCG	CCGACACAGA	CAGCGACAAG
6541	CAGCTCCTCA	AGGAAAAGGT	TGACAGAATA	CAGGAACAGT	GGGACGAAGC	GTCCACCCGA
6601	GTGGGTGCCA	GGAAACGGGA	GCTGGATAAC	ATGCTGACGG	ACTGCCACCA	CTGGGACGAG
6661	CTGAGGGAAG	ACGTGAGCCG	GTGGCTGGAG	CAGGCAGAGC	AGCAGGTCGA	GCACGACTCT
6721	GAAGTTGGAC	AGACGGTCGA	CGTCCTCGAC	AAGCAGATAG	CCCAGCACAA	GGAGTTCATG
6781	GACAACCTGA	GACAATGGCA	GCCGAGCATC	GAGGCAGTGA	ACGACCTGGC	CCAGAAGCTC
6841	ATCTCGGATT	ACAGTAACGA	CGACACCAGC	AAGATCAAAC	AGATGATCGA	TCGCATCAAC
6901	GGCAGGTGGC	AGCAGCTGAA	CAACAGGTGT	CGTGAGCGTC	AGCAGAACCT	GAAGAACGCG
6961	CTGGAGAAGC	TGCAGAAGTT	CAGCCTGCAG	CTGGAGGAGT	TCCTGGCCTG	GCTGTCCGAG
7021	GCTGAGGCGG	CCATGGAGGA	CCTCCAGGCG	GAGATCGACG	CGCACCAGGA	TGTCTGGTCG
7081	TCGCTCAACG	ACACCGGGCG	CCATATCCTG	CAGAGCATCG	AGCACTGTGA	GGACGCACAG
7141	ATGCTGCAGA	GGCGCCTCGA	GAACATGAAC	GACCGCTGGA	CCGCGCTCAG	GAGCAAGTCA
7201	CTGGACATCA	GACAACGACT	GGAGGCGAAC	ATTGAGCAGT	GGAACCAGCT	GCTGATCACC
7261	CTGCAGGACC	TCATCGAGTG	GTGCGGCATC	AAGGAGGACG	AGCTCACGCG	CATGCAGCCT
7321	GTGGGAGGGG	ACCTGAACAC	TGTGCAGCAG	CAGCAGGCAG	ACCACAAGCT	GTTCCGGCAG
7381	TCTGTGGAGG	ACAAGAGGCC	GCTGATCGAG	ACGAGTCTTG	AGACGGGTAA	GGAGTACCTG
7441	GTCGAGCAAG	CAACGAGCCT	TAGCGATGAG	CTGAGCACGG	AAGAAGACAA	GGTGGGCTTG
7501	ACACCTGAGG	AACAGGCTAA	GGAGATCGCC	AGGAACCTGA	AGAGGCAAGT	CAAGACGCTG
7561	ACGGACAAGT	GGGCACAGCT	CCAGGGCAGG	TCGGACATCT	GGCAGAAGAG	GATAGATGAC
7621	GTACAAAACA	AATTGCGCCA	GCTACACGAC	GCGATGAACG	ATTTGAGCGG	AAAACCTGCAC
7681	GAGGGAGAGA	CGCTGAAGGC	TTCGTGGCTG	CCAGTCGGAG	ACCTCATCAT	CGACAACCTC
7741	CAGGATCAGA	TAGATGACGT	CAAGGCATTC	CAGGCCCAGA	TCGCTCCCAT	CCAGAGCAAT
7801	GTGGACTACA	CCAACAACCT	GGCCAATAGC	TTCCCTCCTA	TGGACGTCCA	GCTGTCCAGC
7861	TCTAACCTTA	ACAGGTTAGA	GGACCTGAAC	ACCCGCTGGA	GGCTGCTGCA	AGTCTCGCCA
7921	CAGGACAGGC	TCAAACAGCT	GGAGGAGGCG	CTCAGGGACT	TCGGACCCAA	CTCACAACAC
7981	TTCCTGTCTA	CCTCTGTGGA	AGGGACCTGG	GAGAGGGCGG	TAGCTGCCAA	CAAGGTGCCA
8041	TATTATATCA	AATTCTCAGC	CTACCGCACA	GCCATGAAAC	TCAGGAGGTT	ACAGAAGGCT
8101	CTATGTTTGG	ACTTGCTAAG	CATGAACAAC	GCCATTGATG	CCTTTGACCA	GCACGGACTG
8161	CGGGGCCAGA	ACGATCGTCT	GATGGACGTC	ATCGAGATCA	TCAACTGCCT	GACGACCATC
8221	TACGACAATC	TCGAGCAGGA	CCACGGCAAC	CTGGTCAACG	TGCCGCTCTG	CGTAGACATG
8281	TGCCTCAACT	GGTTACTCAA	CGTGTATGAT	ACGGGTAGGA	GTGGGAAGAT	CAGGGTGCTA
8341	TCGTTCAAGG	TCGGGATCAT	CTCTCTCTGT	AGGGCGCATC	TCGAGGACAA	GTACAGATTC
8401	AACTTCAGGT	TGATAGCAGA	GGCTACAGGA	TTTGCCGACC	AAAGAAAGCT	GGGGCTACTG
8461	CTGCATGATC	TAATACAGGT	TCCTCGTCAG	CTGGGTGAGA	TCGCCTCGTT	CGGCGGCAGC
8521	AACATCGAGC	CCAGCGTGCG	GAGCTGCTTC	GAGCGGGCGG	GCGGGAAGCC	CGAGATCGAG
8581	GCAGCGCATT	TCCTGGACTG	GATGAAGCAG	GAGCCCCAGT	CCATGGTGTG	GCTGCCTGTC
8641	CTGCACAGGC	TCGCCGCCGC	CGAGACCGCC	AAACACCAGG	CCAAGTGCAA	TATCTGTAAA
8701	GAATACCCAA	TTGTAGGATT	TAGGTATCGC	TGTTTGAGGT	GCTTCAACTT	TGACATGTGC
8761	CAGAGCTGTT	TCCTGTCGGG	CAGGAAAGCC	AAGGGACACA	AGCTGTCCCA	CCCCATGCAG
8821	GAGTACTGCA	CAGCCACGAC	CTCAGGGGAG	GATGTGCGGG	ACTTTGCCAA	GGTCGTGCGC
8881	AACAAGTTCA	GGTCCAAGAA	GTCTTTGAAA	AAACACCCGC	GACTGGGGTA	CCTCCCGGTG
8941	CAGACGGTGT	TGGAGGGGGA	CAACCTTGAA	ACGCCGTCGT	CTTCGCCGCA	GCAGGCCGTT
9001	TCTCAGGACA	TGCATAGCAG	GCTGGAGCTG	TATGCCAGCA	GGGAGGACGA	ACACACACTG
9061	ATCCAGCAGT	ATTGTCAGAG	TCTGGGAGGA	GACTCCGGAG	TGCCAAAAAG	CCCGGCTCAG

9121 ATCGTGGTGG CCATCGACCA GGAGCAGAGG GCAGAGCTGG AGGCCATGAT CAAGGACCTG
9181 GAGGAGGAGA ATAGAAATTT ACAGATGGAG TACGAGCGGT TGAAGATCCA GAAGCAGCAG
9241 CAGCCATCGC CGCAGGCGGA GCCGAGGTCG GAGGTCAACC GCGACAACGA GCTGATCGCT
9301 GAGGCGAAGC TTCTCCGGCA GCACAAAGGT CGGCTGGAGG CGCGCATGCA GATCCTGGAG
9361 GACCACAACC GGCAGCTGGA GGCACAGCTG CAGAGGTTAC GGCAGCTTCT CGAGCAGCCG
9421 CAGAGTGACC GCAGCCCGAC CTCCCACACG TCGTCCAGCA GCCACACCCC CAACGAGGTC
9481 ACGCCCTCGT CCTCACAGGG CTCGCTGCCT GGCAACACCC CGCCCTCGCG GGCCCCGCGC
9541 TTCCCCACCA CGACCAACGG GCTTCCCAAC GGGGACGGGG CCGTCAATGG CGCCGCAGAC
9601 ATCAGAGGTG AGAACAGTCA CACACTCTCT CACTATCTGG TTTAG

//

Bflo scf194 dys-u-like mRNA.docx

LOCUS Blo_scaffold_194 9555 bp ss-DNA linear SYN 18-Aug-2014

DEFINITION -

ACCESSION -

KEYWORDS -

SOURCE -

FEATURES Location/Qualifiers

CDS 1..9555

BASE COUNT 2625 A 2429 C 2950 G 1551 T 0 OTHER

ORIGIN ?

```
1 GAACGCGAGG ACGTCCAGAA GAAAACCTTC GCAAAGTGGA TCAACAGCCA GCTTTCTAAG
61 GTTGAAGGCC AACAGGTTTC TGAGCTGTTT GAAGATCTGA AAGATGGAAC CAAACTTCTC
121 AGCCTGCTGG AGATCCTCTG TGGAACATAA TTGGCTCGCG AGAAGGGTCG TATGAGAGTA
181 CATTACCTGA ATAACGTCAA TAGAGCACTG CAAGTACTAG ACAAGCATAA CGTAAATGTG
241 AATGTGGAAA TCCCCGACAA GAAGTCCATC ATCATGTACG TGACGCTGTT CTTCCAAGTC
301 TTGCCTCACC AGAAGATTAC CCAGGAGGCC ATGCAGGAGG CCCTGGAGGT GTCTGCTCCG
361 GGCTATGCTG CCAAGCAGCT AAGGTTTCGGT GTTGTGTTTG TTGTTTCAGCA GACGCAGGGG
421 GCCACGCGAG CGAGCCCCGC CCGCCAGCGA AGCGGAGTCG ACCTGGGAGC GTACCAGAAC
481 GCTCTGGAGG AGGTCTTGGC CTGGCTGCTG GCCGCGGAGG AGAAACTGCA GCATCAGGGG
541 GACATCTCCG ATCTGGTGGA TGATGTCAAG GACCAGTTCC ATGAACATGA GGACCAGTTC
601 CATGAACACG AGGTGGGTAA AATAAACTG AACCAGTTTA AGCTCAAGGA ATATGCTACT
661 AATCTTCCAC TGGTCATGAA GAGAACTGC AGCATCAGGG AGACATCTCC GACCTGGTGG
721 ATGATGTCAA GGACCAGTTC CATGAACACG AGGTGGTCCA TGGACAGGCA GTCCAAGTTA
781 CATGAGGTGC TGATGAACCT GCAGCAGAAG CAGCTGGACC AGCTCGATGC CTGGCTGTCC
841 GGCACGGAAC AGCGCATCAA GGTCATGGAT ACGCAAGGCA TTGGGCCCCG CTTGACCAG
901 GTCCGCAGGC AGGTGGAAGA ACACAAGGTA GCCACCCAAG CCCTTGAGGC TCAGCTGCAG
961 AACCTGGGCG AGCGGTGGGC CTGGGTGTGT AAGTGGACGG AGGACCGCTG GATGGTGCTA
1021 CAGGAGGTGC TGGTGCGATG GCAGGTGTTT AACCTGGAGC AGATGGACTT CCACGACTGG
1081 CTGACCGACA AGGAGGGGGC TCTGCAGACC ATGCAGCACG CCAACCTCAG CGACCCCAAC
1141 GTCATTGTTG CACAGGTCAA GAACTAAAG GTACGGAGAG AAAGGATTTT AGAACATGAG
1201 ATGGAGGCCA AGCAGCCGAC CTTCCAGGTG TTGAGCGGGA AGGCTGCGGA GGTGGTGCAG
1261 CATCTGGACC GCTCCTCCTC AGCCGTGGAG GGGATTTCAG GCAAGATCGA GGAGTTCACG
1321 CAGCGCTGGG ACGCCATTGT GCAGAAGATG GAGAACTACA GTAAACTGGT AAAGATGGCG
1381 TTAGTGCTCG GACGGAAACT TGGCGAGCAT ATGTCTGCCT CGGCTGCCGT CCCCAGAAG
1441 ACGACCACCA TCACCGAGAC GGTCACCAAG ATGACGCAGA AGCTGACCCA GAAAAACCA
1501 CCGCCGCCGC CTCCTAAGAA AGCTGTCAGC CAACAGAGAA GGAGGTTTGA CAGCGAGATC
1561 GCGACACTTC TTGCCTGGAT CGATGACACA GAAGCCCTTC TGCTGTCGCC CGCCTACACC
1621 AACCCTGCTA CCTCAAACAG CTTCCAGGCT CAGACCGACA AGTTCAACCA ACTCGACGAG
1681 AACATCATGT CGTACCAGGG CCGTGTGCAG ACCATCCAGG AGACCGGCGA AACCCTCAT
1741 GAGGAAGAGG ATGACAAGAT TCAGCTGGAG ACGCGCCTGA ACACGTTTAC GGAGCGCTGG
1801 AACGAGATCT GCGTGCTGAT CAACGACCGG AAGCGGAAGC TGCAGATGGC GCGGCTGTGT
1861 AAGCAGTTCT ACACGGAGCT GGATGTCCTG AGGGACGTCA CCGGGAGGAA GTTGACCTGG
1921 CTCCAGGGGC AGGACAAGCC CAGCGACGAT GCAGTCAAGA TTAAACAGCA GATGGACGAA
1981 CTGCAGTTG AAGTTACGGA CATGGCAGCG CACAGCGAGC GCGTAGAGCA GCTGAACAGC
2041 CTGGGCTCGC AGCTGGCGGA CTCCTCAGAC CCGCCCAAGA ACCTGCAGCA GGACCTGACA
2101 CAGTTCCAGC TGCAGTGGGA GGAGCTGGTG GCCAAGATGG CTGCCAGGGA GAAGGACCTC
2161 ACCAAAGCTC TTGAGCAGGC CCCTCCTAAA CAGTACTTGG AGGCCATGAA GGCGCTGACC
2221 CAGTGGATCG ACAACGTCGA GTCTCTCCTC ACCACGGAGA AGTTCTACAT TGCAGACATC
2281 CCCACCATGA AGGACCAGCA GCAGAAGTTC AAGGTGAACA AGTTTGGAGG TACAGCAGGC
2341 CCTGAAGGAC CAGCAGAAGA GTCTGGACTT CGTGAACACC ACGGGGAAGC AAGTGGTGGA
2401 GAGTCCCCGC CGGAAAGGTC CCGCAGCCTG CAAAAGGATC TTAACGGGCT GAACGAGCGC
2461 TGGAGTAACG TCTGTGCTAT GACCGAGGAT CGGCTGGAAC ACCTCGACAA GGGGGTCGAG
2521 CTCCTCACGC AATATCAGAA AGAAGGTGAT GGCCTGCGTA CCTGGATGGA CGAGGTGGAT
2581 GTGTACATCC ACTCAGAGGA GGCAGCAGTG GGTGACACAG AAACCCTACA GGCTCAGCTA
```

2641	GAACAGAGCA	AGTCCCTTCA	GAAGGACATC	GCCTCCCTCC	AGCACAAACAT	GAACAACGTG
2701	AACGAGGCGG	GGAAGAGACT	GGTCCCCGAG	GCGCACCCCG	ACTTCTCCAA	GAAACTCAAG
2761	CAGGAGCTGA	AGGACCTGAA	CGAACGCTGG	GACGCCATCA	CCAAGAAGCA	AAAGTCTCAG
2821	AAGCAGAGTC	TTGAAGGAGC	GATGAAGAAG	ACGAGCTCCT	TACAGCTTCA	GCTGGAGGCC
2881	TTCAGCGAGT	GGCTGACAGA	CAAGGAGAAG	GAGGGACTGA	AGGAAGACGG	ACCTCTGGAG
2941	GACCCTGATG	ACATCCAGCC	CAGGATCAAG	AAATATAAGA	AACTGAAAGC	AGACCTAGAC
3001	AAGAAAGAGC	CTGATCTGAA	AGCCCTGAAC	ACCGTTGGTA	ACGAGATGGC	GACGAAGGCC
3061	GGTAGCCCGG	GTGGACTCGG	CAATGATCTT	CAGGTGGTCA	ACGAGAGGTG	GACTGTGCTG
3121	TACAACACTG	TCAACAAACG	CTACGAGGTT	TCACAGGAGG	TGTGGACCCT	GTGGCAGCAG
3181	TTCCAGGTGT	TACTAGAACA	GGAACAATCC	TGGCTGTCCA	CCCTGGAGAC	CAAGATCCAA
3241	CAGACCGAGG	CTGTGGGCGG	AGACGCAGAG	GAAATCTCTG	AAGCTCTGGA	TGCCCTTGAG
3301	AACCACGTGA	GGAACCACCC	GTCAGACAAC	AGGGACAAGA	TCCAGGAGAT	CTCCCAGCAG
3361	CTGGTGGAGC	ACAAGGTTCT	GGTGGAGCCG	GTACAGACGG	AGGTCACGAC	CTTCACCTCA
3421	CGATGGGACA	CTCTGCAGCA	ACAGGCTAAG	CTGCGCCAGA	CGGCCCTTGA	GCAGCGGATC
3481	GCCAGTGCAC	AGGACTTGGA	GAAGGACGTA	CTCGAGCTTG	AGCGCTGGTT	CACCCGGACA
3541	GACAAGCTTC	TGCACTCGCG	CATCACCTCT	GGTGTCTGCG	CCTCCGACGT	GCCTGAGGAG
3601	TTCGAGAAGC	GAATGGTTGA	GATAAAGGCG	AAGTTCCGCA	AGTACCAGAA	GCCGGCAGAC
3661	TTCGAGCCGA	AGATGGGCCG	TGTTTCGAGAC	ATCTTGAACA	ACGTTGAGAA	GGAGCTGGGT
3721	GTTCTTGAGC	TGAAACACCA	CGAGCCTGAG	ATGGTGCAGG	GGCAACTCGA	CATCTGTATG
3781	GGCCTGTACA	AGATCCTGAG	CGATGTGAAG	CCTGAGGTGG	AGTTTGTGAT	AAAGACTGGG
3841	CGACAGGTGG	TGGAGAAGAA	ACAAGTGGAC	AACACGGAGG	ACCTTAACGA	GCGGCTTAAC
3901	GAGCTGAAAC	TCCAGTATAA	CGACCTGGGT	GCTCAGGAGT	TCATGTCTGA	GATGGAGAAG
3961	TTCCGGGAGG	GTCTTAACGA	GGTGCACACC	TGGATGGAGA	AGGTGGAGAT	GTTCTTTAAG
4021	ATCTCCGACA	GACTTAGTCC	TGAGGAACAG	GCCAAGGAGG	TGGACACCTA	CAAGAAGCTT
4081	CAGCTGGAAG	CAAACCGGTT	GCGCGGCACG	GTGGACCAGG	TGCGAGACCA	GGCCATGGAG
4141	CTGATGCAGC	ACGGGCCCGG	TGTACAGGGC	ATGGTGGAGC	CGGAACTCAT	CAACCTCAAC
4201	AGCAGATGGG	AGAGTGTGTG	TGAGAAGATC	AGGGCAACCG	AGGATGAAAT	CTCAAGAATA
4261	CAGGCCGAGA	TCGACGACAT	CTTCGAGAAG	GGAGACAAAG	CCATCAAGGA	GACGGACGAC
4321	TCGACCAAGA	AGGAGCAGAT	ACGTCTTAAA	CTGAACGACC	TTAAAGGAAG	ATGGAGCTCG
4381	GTGAAGGATG	ATGCTGAGCT	GAAGAAGAAG	ACACTGCGGA	CGGTGGTGCC	ACAGTGGTAC
4441	CAGTTCCGTT	CCCAGACCGA	TGATCTCAAC	ATGTATCTCA	CGGAGATCGA	GATCAGAATA
4501	CGGGAGGCCA	AGGATGATCC	CAGCAAGCTA	CAGGCTATCC	ATGAGGAGCT	TCAGCATAAG
4561	CAGGAGGACC	TGGACAGGCT	GAACAGGAGA	GCCGAGGAGC	TGAAACAGCA	GGGAGCCAAA
4621	CCTGTGGTGG	AGCCGCAGCT	CATCCACATC	AACAGGAGAT	GGAAGGATGC	CATAAAGGAG
4681	CAGCTAGACA	CGTTGGAGCC	ACAAGTGAGG	TCGGCCCCACG	AGCGCAAGGG	GGCTGCCATC
4741	TCTCAGGTGG	CGCCCTCCGA	ACGGAAGAAC	ATCGAGAATG	CCATGTCTAA	TCTCAAGGTC
4801	GAGTGGGACA	AGGTCAACAA	GGATTATGTG	GACAGATACA	GTCGCTTTAA	GAAGTCACTG
4861	GCGCACTGGC	AGCAGTTCCA	CTGTGACATG	CGGGACCTGA	CGTCCTGGCT	GAACGAGGCA
4921	GAGCGGACGG	TGCAGGATAC	CAAGGCTCCA	GGGGGAGACT	TGGATCTGGA	GAAAGCACGG
4981	GCAGCGCAGA	AGAAATTGCA	AGACGGGATT	GAGAAGCACA	GGAAGCAAGT	GGAGGGACTG
5041	AACGCGAACG	GGCAGGAGAT	TGTTGAGCAG	TGCACGTCCA	TCGACGCCAC	CCTTCTGCGG
5101	GAAAAGCTTG	ACGGGCTGAA	CAGAAGGTGG	AGGGTGCTGT	GCGCAGAGGT	CAACGACAGG
5161	AAAGAAAGGT	CAGTGGAAAA	ATGGTTGGAG	GAGGCGAGCA	TGCAGACAGA	GGAGTTCAG
5221	GAGGAGATTG	ATGAGCTGCT	GTTTTGGCTG	GAGGACACAG	AGGGCATCCT	TAACAAGCAG
5281	GTGTGCCCAG	CTGACGAGGA	TCAGCTAGAG	GAAGCTCTGG	AGAAAGTTAA	GGACAAGGAG
5341	AATGACCTTG	GATCGAGACA	GGACAACGTG	AACACCATCA	ATCAGACCGG	AAACCGGATC
5401	ATGACGACGC	CGGGCGCCTC	CAAGGCACTG	GTCAGCAACG	TCAAGAAGAA	GGTCGACAAC
5461	CTCAACACAA	GATGGCTGAC	GGTGACGCCG	GAGATTCCCC	AGAAGAGGAG	AGGACTGCAG
5521	GACCAGCTGC	ACTTCTGTCT	CACCTTCCTG	GAGGAGCTGG	AGGAGCTTCT	GGTGTGGATG
5581	GCTCACACGA	TGCAGTTGCT	AGAGCAACAG	AAGGGACCCA	TCTCATCCGC	AACATCATCT
5641	GATGAAGGGG	ATCAGGTCAT	CATCGACCCT	AAGACGATGA	AGGATGCTCT	GACGGCCCGT
5701	CAATCAAACG	TCGATAACGT	CAACCTCACG	TACAACAGCC	TGGTGGCCGA	GAGCAAGGAC
5761	CTGGAAGTGG	ACGTCCCCTG	CTATGTCAGA	GACAAGATCA	GGAAGCTGAA	CGCCGACTGG
5821	GCCAAGATTC	GCTACATGGC	GGCCAACCTG	AGGCCGACCT	CGGATTCCGA	CGTGGAATCT

5881	CTCGTTATGG	AGCAGAGAAG	ATCTATGGAG	GAGATGGAAA	CCGCTAGGAT	AGAAGTCGGC
5941	GCCATGCAGG	CGAGGGCGGA	CATGCCGTCC	ATGTACCCCG	ACTTCGACAA	GTCTGTGGCG
6001	GAGCTCCGTG	ATTGGCTGAT	GCTCCTCGAC	CACATGCTGA	AGTCTCAGAT	CGTCACCGTG
6061	GGAGACGTCG	AGGAGATCGA	AGAGATGATC	ATCAAGCAGA	AGACTGTCTT	AAACGACCTT
6121	GACGTCAAGA	AGCCCCAGCT	GGATACGGTG	GTGGCTGCCG	GCAGACGGCT	CATTCAAACC
6181	GCCGACACGG	ACAGCGACAA	GCAGCTCCTC	AAGGAAAAGG	TCGACAGAAT	ACAAGAACAG
6241	TGGGACGAAG	CGTCCACGCG	AGTAGGTGCC	AGGAAACGAG	AGCTGGACAA	CATGCTGACG
6301	GACTGCCACC	ACTGGGACGA	GCTGAGGGAA	GACGTGAGCC	GGTGGCTGGA	GCAGGCAGAG
6361	CAGCAGGTCG	AGCACGACTC	TGAAGTTGGA	CAAACGGTCG	ACGTCCTTGA	CAAGCAGATC
6421	GCGCAGCACA	AGGAGTTCAT	GGACAACCTG	AGACAATGGC	AGCCGAGCAT	CGAGGCAGTG
6481	AATGACCTGG	CCCAGAAACT	CATCTCGGAT	TACAGTAACG	ACGACACCAG	CAAGATCAAA
6541	CAGATGATCG	ACCGCATCAA	TGGCAGGTGG	CAGCAGCTTA	ACAACAGGTG	TCGTGAGCGT
6601	CAGCAGAACC	TGAAGAACGC	GCTGGAGAAG	CTGCAGAAGT	TCAGCCTGCA	GCTGGAGGAG
6661	TTCTTGGCCT	GGCTGTGCGA	GGCTGAGGCG	GCCATGGAGG	TACTGGATGA	AGAGACAAGC
6721	AGAGAAGGTG	TGCTGGACGA	CCAAGAACAC	CTGAAAGTGT	GGAGGGACCA	GTACAGGGAC
6781	CTCCAGGCGG	AGATCGACGC	GCACCAGGAT	GTCTGGTCGT	CGCTCAACGA	CACCGGGCGT
6841	CACATCCTGC	AGAGCATCGA	GCACTGCGAG	GACGCGCAGA	TGCTGCAGAG	GCGCCTCGAG
6901	AACATGAACG	ATCGCTGGAC	CGCGCTCAGG	AGCAAGTCGC	TCGACATCAG	ACAACGACTG
6961	GAAGCAAACA	TTGAGCAGTG	GAACCAGCTG	CTGATCACCC	TGCAGGACCT	CATCGAGTGG
7021	TGCGGCATCA	AGGAGGACGA	GCTCACGCGC	ATGCAGCCTG	TGGGAGGGGA	CCTCAACACC
7081	GTGCAGCAAC	AGCAGGCGGA	CCACAAGCTG	TTCCGGCAGT	CAGTGGAGGA	CAAGAGGCCG
7141	CTGATCGAGA	CGAGTCTTGA	GACGGGTAAG	GAGTACCTGG	TGGAACAAGC	AACGAGCCTT
7201	AGCGATGAGC	TGAGCACGGA	AGAAGACAAG	GTGGGCCTGA	CGCCGGAGGA	ACAGGCTAAG
7261	GAGATTGCCA	GGAACCTGAA	GAGGCAAGTC	AAGACGCTGA	CGGACAAGTG	GGCACAGCTC
7321	CAGGGCAGGT	CGGACATCTG	GCAGAAGAGG	ATAGACGACG	TACAAAACAA	ATTACGCCAG
7381	CTACACGACG	CGATGAACGA	TTTGAGCGGA	AAACTGCACG	AGGGCGAGAC	GCTGAAGGCT
7441	TCGTGGCTGC	CAGTCGGAGA	CCTCATCATC	GACAACCTCC	AGGATCAGAT	AGATGACGTC
7501	AAGGCATTCC	AGGCCCGAGT	CGCCCCCATC	CAGAGTAATG	TGGACTACAC	CAACAACCTG
7561	GCCAACAGCT	TCCCTCCTAT	GGACGTCCAG	CTGTCCAGCT	CAAACCTAAA	CAGGTTAGAG
7621	GACCTGAACA	CGCGCTGGAG	GCTGCTGCAA	GTCTCCGCAC	AGGACAGGCT	CAAACAGCTG
7681	GAGGAGGCGC	TCAGGGACTT	CGGACCCAAC	TCACAACACT	TCCTGTCTAC	CTCTGTGGAA
7741	GGGACCTGGG	AGAGGGCAGT	AGCTGCCAAC	AAGGTGCCAT	ATTATATCAA	CCACACTACA
7801	GAAACAACCA	GCTGGGACCA	CCCCAAGATG	ACGGAACCTCT	TCCAATCATT	AGCTGACCTG
7861	AATGACGTGA	GATTCTCAGC	CTACCGCACA	GCCATGAAAC	TCAGGAGGTT	ACAGAAGGCT
7921	CTATGTTTGG	ACTTGCTAAG	CATGAACAAC	GCCATTGATG	CCTTTGACCA	GCACGGACTG
7981	CGGGGCCAGA	ACGATCGTCT	GATGGACGTG	ATCGAGATCA	TCAACTGCCT	GACGACCATC
8041	TACGACAATC	TGGAGCAGGA	CCACGGCAAC	CTGGTCAACG	TGCCGCTCTG	CGTAGACATG
8101	TGCCTCAACT	GGTTACTCAA	CGTGTATGAT	ACGGGTAGGA	GTGGGAAGAT	CAGGGTGCTA
8161	TCGTTCAAGG	TCGGGATCAT	CTCTCTCTGT	AGGGCGCATC	TCGAGGACAA	GTACAGATTG
8221	AACTTCAGGT	TGATAGCAGA	GGCTACAGGA	TTTGCCGACC	AAAGAAAGCT	GGGGCTACTG
8281	CTGCATGATC	TAATACAGGT	TCCTCGTCAG	CTGGGTGAGA	TCGCCTCGTT	CGGCGGGAGC
8341	AACATCGAGC	CCAGCGTGCG	GAGCTGCTTC	GAGCGGGCGG	GCGGGAAGCC	CGAGATCGAG
8401	GCCGCCCCT	TCCTGGACTG	GATGAAGCAG	GAGCCCCAGT	CCATGGTGTG	GCTGCCTGTC
8461	CTGCACAGGC	TCGCCGCAGC	CGAGACAGCC	AAACACCAGG	CCAAGTGCAA	TATCTGTAAA
8521	GAATACCCAA	TTGTAGGATT	TAGGTATCGC	TGTTTGAGGT	GCTTCAACTT	TGACATGTGC
8581	CAGAGCTGTT	TCCTGTGCGG	CAGGAAAGCC	AAGGGACACA	AACTGTCCCA	CCCCATGCAG
8641	GAGTACTGCA	CAGCCACGAC	CTCAGGAGAG	GATGTGCGGG	ACTTTGCCAA	GGTCGTGCGC
8701	AACAAGTTCA	GGTCCAAGAA	GTCTTTGAAA	AAACACCCGC	GACTTGGGTA	TCTCCAGTGT
8761	CAGACGGTGT	TGGAGGGAGA	CAACCTCGAA	ACGCCGTCGT	CTTCGCCACA	GCAGGCCGTT
8821	TCTCAGGACA	TGCATAGCAG	GCTGGAGCTG	TACGCCAGCA	GTAAGGAGCA	TGTGACTATC
8881	AGGCCATTCT	CACCTCAGCT	CCCGCGCACA	GAGCCACAGG	AATCTTCCGG	ACGTATTGTG
8941	CCAAAAGCC	CGGCTCAGAT	CGTGGTGGCG	ATCGACCAGG	AGCAGAGGGC	AGAGCTGGAA
9001	GCCATGATCA	AGGACCTGGA	GGAGGAAAAT	AGAAATTTAC	AGATGGAATA	CGAGCGGTTG
9061	AAGATTCAAA	AGCAGCAGCA	GCCGTCGCCG	CAGGCGGAGC	CGAGGTCGGA	GGTCAACCGC

9121 GACAATGAGC TGATCGCGGA GGCAAAGCTT CTCCGGCAGC ACAAGGGTCG GCTGGAGGCG
9181 CGCATGCAGA TCCTGGAGGA CCACAACCGG CAGCTGGAGG CACAGCTGCA GAGGTTACGG
9241 CAGCTACTCG AGCAGACTGT ACAAGGAGCG ACGCGCACCC CTCCCCCACT GCGAGACGGA
9301 CAGCGGCGCC AGCAGACTCC GGAAGAGCAG CTGGAGGAGG TCATGCGGGA GGTCAACAGC
9361 TCCTTCCCTC CAGACTCCAA TCACACACAA GGTAAGCTAC AGATGACCAA TCACACACAA
9421 GGGAAAAACA CGGTAGGAAG CCTGTTCAAC ATGGCGGGCG ACCTCAACAA GCGGGTGGGG
9481 AAGCTCGTTA CCGTCATGAC TAATGAGGAC GGGACGGAGA TGGGCGAGAA TGGTCACGTC
9541 AACAGCCTGT CGTAG

//

Pmar dys-u-like mRNA.docx

LOCUS Pmar_3-scaffold_ 9963 bp ss-DNA linear SYN 16-Sep-2019

DEFINITION -

ACCESSION -

KEYWORDS -

SOURCE -

FEATURES Location/Qualifiers

CDS 1..9963

/note="Pmar dys-u-like protein"

BASE COUNT 2142 A 2901 C 3253 G 1491 T 176 OTHER

ORIGIN ?

```
1 ATGACCGGAG GTCAGGTTTG GGGGTGGCCC TCCGAGGCAT TCAATCCTGG AGCAATTGCT
61 CGTATCGAGG CGATGGTTAT GGAAGATCGG CAAATTAGGG CGGCAGAGGA GCGGGAGTGC
121 TGCGCCGACA TCGAGGCTGG CAAGACGCCC ATAGAGGACC TGTTACCTC CCTGCGCGAT
181 GGCCGGAGGC TGCTCGAGTT GCTGGAAATC ATTTTGGGCC AGAAGTTGTC CAAGGAACGT
241 GGGACATCGC GGGTACATGC CCTGAATAAT GTCAACCGCG CTCTGCAGGT GCTGCAGAAA
301 AACAACGTGG AGCTCGTGAA CATTGGAGGC GCCGACATTG TGGACGGCAA CCACAAGCTG
361 ACGCTCGGCC TCATCTGGAG CATCATCATC CACTGGCAGG TAAAGGATGT GATGAAGGAT
421 GTGATGGCCG ACTTGCAGCA GACGAACAGC GAGAAGATCT TGCTGGGCTG GGTGCGCCAG
481 TCGACGCGGG ACTATCGCCC CGTCAACGTC GTCAACTTCA CCACCAGCTG GTCGGACGGG
541 CTGGCCTTCA ATGCCCTCAT ACACAGCTAC AGGCCTGACC TCTTTAACTG GAAGACCGTT
601 GCCAACCAAA TGACGGTCAC AGAGAGGCTC GACAACGCCT TCAATGTGCG CCAGGACCAT
661 CTGAACATCG CCCGCCTTCT GGATCCGGAA GACGTCGAGA CGGCCACCC GGACAAGAAG
721 TCAATTCTCA TGTACGTGAC GTCGCTCTTC CAAGTGCTCC CCAAGCAGGT GTCAATGGAG
781 GCCATCCGAG AGGTTGAGAC TCTCCACGG GCAATCACGG TTAGCCTCTC CCAGAGCAAT
841 GTCACGGCTT CCCCAAAGCC TCGCTACAAA ACCTTCCTCT ACGCGCAGCA CCCCAGTAGG
901 CTGTCCCCCG AGAGACAGAG GCGGGGCGTC CCGTCGCAGC TGGAGGGCGA GGTGGACCTC
961 GAGAGTTACC AGGGCGCCCT AGAGGAAGTG CTCACCTGGC TGCTGTCCGC CGAGGACACC
1021 CTGCAGGGCC AGCCCGACAT CTCGGCCGAC GTCGAGGAGG TCAAGGAACA GTTCCACGCG
1081 CACGAGGGCT TCATGCTGGA GCTGACGTCG CACCAGAGCA GCGTGGGTAA CGTGCTGCAC
1141 GCGGGCAACC AGCTGCTGGC CGAGGGCAAG CTGGGCGAGG ACGAGGAGAG CGAGGTGCGC
1201 GAGCAGATGA ACCTGCTGAA CTCGCGCTGG GAGAGCCTCC GCGTCGCCAG CATGGACCGC
1261 CAGTCGCGGC TGCACGAGGT GTTAATGGCG CTGCAGCAGA GGCAGCTGAC GGAGCTGGCG
1321 GACTGGCTGA GTCGCACGGA GGAGAGGATC CGCAAGCTCG ACACCCACCC CATTGGGCCC
1381 GAGCTCGACT CCGTGAAGGC GCAGATCGAA GAGCACAAAT CCTTGCAGTC GGACCTGGAG
1441 CAGGAGCAGG TGCGTGTGAA CTCTCTCACG CACATGGTGG TGGTGGTGGA CGAGAACTCG
1501 GGAGAGAACG CAACAGCTGC CTCGAGGCC CAACTGCACT CCCTGGGGGA GCGCTGGGCG
1561 GCCGTGTGCC GCTGGACAGA AGATCGCTGG CGCCTCCTGC AGGAGAGTCT CGTGCGGTGG
1621 CAGCAGTTCA CCGAGGAGCA GGTGCTGTTT GATGCCTGGC TGA CTGAGAA GGAGGATGCC
1681 CTCAGCAAAG TTCAGACTGG GAACTTCAAG GACCAGGCTG AGATGGTTGC GAATGTCCGA
1741 AAACTCGCCA TCTTGAAGGA GGACATGAGC ATGAAGCGGC CGGCGCTGAG CAAGCTGAGC
1801 ATACTAGGCC AGGCGCTGGT CACGGCCATG GGCAACGAGA CAACCGCCAG GGACATTGAG
1861 AAGAACTGG ACGAGTTTGC GCAGCGCTGG GACAACCTGG TTCAGAAGCT GGAGGGACAC
1921 TCCAAGCTGG TGTCGCACAC AGTGAGCCAG ACCCAGTCGT CGGTACCCA GTTGACGGTA
1981 CTGGAGACGG TGACGGTGAT CAGCAGCAGC AGCAGCAGTC GCGAGCAGGT GCTGGTGAAG
2041 CACGGGCGGG AGGAGCACGT GCTCGCCCCG CCCACCAAGA AGCCCTTCCT CGACCCCGAG
2101 CTGCACAAGA GCAGATTTGA CACCGAGGTG CAGGAGCTCA ACAGCTGGAT CAGCAGAGCG
2161 GAGACCACAC TCAACTCCCC CGAGTTCTCC ACCTTCAAGA AGTCTGGCGA CGTCTCGGAC
2221 CTCCAGGAGA AAGTTAACGG AGTGGAGAGG GAGAAGGCAG AGAGACAGAA GAAGCTCCAC
2281 GATATCCCGA ACACCGCGCA GACGCTGGCC GATCAGATGT CACAAGACGG TGTCCCCACG
2341 GAGGAGGTTT AGAGGGTGCT CAACGAGCTA TCTACCCGCT GGAAGCGCT CCTGCGGTTG
2401 CTAGGCGACC GCAAGGCCTG GCTAGTCTAC CAAGGCCTGG TGCTGGATTT CGAGCGCGAC
2461 GCGCGCACCG CCGAGCGGGG CCTCGCCGGG GCCGAGGCGC TGCTGCAGCA GGGCCCGGGC
2521 GACACGGGCG ACGCGGCACA CTGCGGCAC CGGATCGGCG AGCTGCAGTC GGAGCTGGAG
```

2581	AAAGTGCAG	GCATGAAGGG	CGAGGTGGAG	AGGCTCGGCG	TGCAGGCCAG	CTCCCTGAAG
2641	CTCCGCGACG	GGGCCCTCGC	TTTCCTGGAC	GCGGAGTTCG	TGCAGCTCAA	CAAGCGCTAC
2701	GCGCAGGTCA	TGTCTGAGCT	GGAGGCGAAG	AAGGGACAGC	TGCAGACAGC	CCTGGACAGC
2761	CTCCCTCCAC	CGCAGTACGC	TGAGACGCTC	GCAGCGCTCA	AGGCCTTCCT	GCAGGATGCG
2821	GGCACTCGCC	TCGGGCAGAT	CCCCGCGTCA	CTGGGCGACT	TGGAAGTCAT	GGACAAACGC
2881	CTGAAAGAGC	TAAAGGGGCT	GTCAGCCGCC	CAGCGCGAGA	GACAGAGCTC	CTTGATTCA
2941	GCGAGGGTCA	TGGCCGAGGA	GTTGAGCCGG	CGATCTCGCC	CGGACGCCGC	TCGCAAGCAC
3001	GCGGCCGAGC	TGGAGCCGAT	CGTGGAGCCG	TGGAGCAAGC	TGTCCACGCA	GCTCCCGGAG
3061	AGGATGCAGC	GAGCGGAGGA	CCTCTCTGCC	AAACTTGGA	AGTTCAAGGC	GGAGTGCCAG
3121	GCCCTTCGCA	AGTGGATGGA	TGAAGTGGAC	GTGTTTTTGC	GAGAGGAGCG	CCCGGCCCTC
3181	GGGGACACCG	ACACCCTGGA	GAAGCAGCTG	GACCAGTGCG	GCGCGCTCGT	GAGCGACATC
3241	AACACGCTGC	AGCCCAGCCT	GCGCAGCATC	CAGGACCTGG	GCCAGAGCCT	CAAGGGCCAG
3301	GCAGAGGCGC	CGCTCGCCAC	GTCCATTTCAG	GCGGAGCTCT	CGGCCATCTC	CTCGCGCTGG
3361	GACACGGTCT	GCAAGCAGGC	GAGGGCTCAG	AAGCAGGTGC	TGACGGAGGC	TCTGGCTCGC
3421	ACGCGGGAGC	TGCGCAAGGA	CCTGGCGGAG	ATGCAGGAGT	GGATCACGCA	GGCCGAGGAG
3481	GAGTACCTGG	AGAGGGACTT	CGACTGCAAG	ACCCCCGATG	ACCTCAAGTC	TGCGGTTCGCC
3541	GAGCTCAAGC	GTGCCAATGA	GGAGCTGTCA	CACAAGGAGG	TGAAGGTGAA	GATACTGAAG
3601	GATGCCGTGA	CTTCGTTTCGT	TGGCAAGAGC	TCAGGACCCG	GGCGTGACAA	GATTATGGGC
3661	GAGCTGTTCG	CTGTGTTGGC	AAACTACCAG	CGCCTTTGCA	GCCGCCTTCA	GGGCAAGCAC
3721	ACAATGCTGC	AGGAGGTGTG	GGCGTGCTGG	AGCGAGCTTC	TCTCCTACCT	GGAGGTGGAG
3781	AACAGGTGGC	TGGATGATTT	GGAGGGCGCA	ATCAAGCGTG	CGGACCTTTC	CCAGGGAGAC
3841	CCGGATGCGC	TCTCCGCCGC	TCTGAGTTCC	CTGGACGTGC	TGCTGGCCCA	CCCACGGAC
3901	AACAGGACCC	ATGTGCGTGA	GCTCTCGCAG	ACACTCATTG	ATGGCGGCAT	CCTGGACGAC
3961	ATCATCGGAG	AGAAGCTGGC	GGCCTTCAAT	CTCCGCTGGG	AGGAGCTCAC	CCAGCAGGAG
4021	ACGGACCGCG	CGCTCCGCAC	CGTCGCGGAC	GGCCTCTCTG	AGCTCGACCG	TCAGCTCGCC
4081	AAGCACCTGG	CCGACAAGGT	GGATGCGCTG	CAAGTCCCGC	AGGAGGCGCA	GAATAATTCAG
4141	ACGGCTCTCA	AGACGTACGA	ACTGAGCCTT	CAGCAGCTCG	GAGCAGAGGG	TCGCTCTCCA
4201	CGCCTCAAGG	GCCAGCTGGA	ATCTACGCAG	AAAAAGCTGG	GGCAGGTGGA	GACAAAATTC
4261	CGTCTCTTCC	AAAAGCCGTT	GGACTTTGAC	CGGCGCGTGA	CGGAGTGCCG	GCGCATCTTG
4321	ACGGACGTGA	ACTCGGAGGT	GCCGCTCGTG	AAAGTGTCGA	GTGCCGAGCC	GGAGGGAACC
4381	CAGGCCCAGC	TGGAGCACTG	CATGCAATTG	TACAAGACGC	TGAGCGAGGT	GAAGACCGAG
4441	GTGGAGGCTG	TGATCAAGAC	CGGGCGGCAG	ATCGTGACAG	AGCAGCAGAC	GGAAAGACCC
4501	CAGGAGTTGG	ACGAGCGGCT	CACCTCCCTC	AAGCTGCTCT	ACAATGATCT	GGGTGCCAG
4561	GTGACCGACA	GCAAGAGCGA	CCTGGAGAAG	GCGCTCAAGC	TGGCGCGCAA	GCTGCGGAAG
4621	GAGATCGCCG	CCCTCACCGA	TTGGCTGGCG	GCCACCGACT	CGGAGCTGAC	GCGCCGCTCC
4681	GCCGTGGACG	GAATGCCCGA	GAGCCTGGAG	ACGGAGCTCG	CCTGGAGTCA	GGCGACGCAA
4741	AAGGAGATCA	CCAAGCGGCA	GGGCTCGCTG	GCGGCCGCGG	CGGACACGGG	CGCAGCGCTG
4801	CAGTCTCTGC	TCAAGGGGCC	CGAGGGCTCG	GCCATCGACG	ACAAGCTCAG	CCTCCTCAAC
4861	AGCAACTGGA	TCGCGGTGAC	GGCACGCGCC	GAGGAGTGGC	TCAACCTGCT	GCAGGAATAC
4921	CAGGCCCAGA	TGAACAGCTT	CGACCACAAC	ATTGCGGATA	TAACCACGTG	GATATACCAC
4981	TCGCAAGTTA	TGATGGATGA	CGCGGAGAAG	CTGTGCGCCC	AACAAAGGGA	TGATGTGTTT
5041	AAGCGTCTGC	AGTCGGAGCT	GCCTGGTACG	AGCCATAAGG	TGGAGGTGGT	GCACGACCAA
5101	GCGGGCGAGC	TGATGGCAAA	CCGCGCGCAT	TACTGCCGCC	GACTCGTGGA	GCCCAAGGTG
5161	ACCGAGCTCC	ATCAGAGGTT	CAACTCCGTC	GCTCAGCGCA	TCAAGACCAG	GAAGGGAGGG
5221	GAGCGCATGG	CGTTTGGCGA	CGGCACGGAC	GGGCCCACGT	CGCCGGAGGC	CGCTGCTACC
5281	GCCGGCATCG	TCTTTGCGGC	GGCCTCTCGC	GAACAGGGCC	TGGAGCACGC	CCTGAGGCAG
5341	CTCGACGCGG	ACATTCAGCG	CGGGGAAGAT	TTGTACGAGG	AAGATTTTGA	AGAAGGAAAG
5401	GATGTTGTTG	AGGGTTTGGT	GGAAGAGGCT	TTGAAGACGG	GGGAGCAGCT	ACTGAAAGAA
5461	ACGAAGGATG	AAAGCGAGAG	GGAGAGAATC	GGTGCCAAAC	TCAAGGCGCT	GAGGAACAAG
5521	TATGCGACGA	TAAAGGAGGT	CAAGATTGTA	CGTCGACAGA	AGGTGATGGA	GATATGGCCA
5581	AGTTGGTACC	AGTACAGGCG	GGAGACGGAT	GATATGATGC	AGTGGTTGGA	TGACACGGAG
5641	CAGAGACTCA	CTGGAATGGA	GGGGACGCCC	AATGAGCTGG	AAGTTCAGGA	GCTGGTGAAG
5701	CAGCTGGAGG	AGCGCAGGGA	GGGCTACGAG	ACGCTGCAAA	GGCGGGCGCA	GCAGCTGGCA
5761	CGCGGTGGCG	CCGCCAAAGC	TCTGGAGCCC	AACATGCTGC	GCCTACCAA	GCGCTGGCGG

5821	GACCTGGAGA	GCCACTTCCT	GCAGTTCCGC	AGGGCGCAGC	ACCAGACGAA	CCTGGAGGAG
5881	CAGATAGAGC	AGCAGTGGCC	AGCAGTGGTA	GCCCTTGGCT	CAGCGGGCGA	GGAGATAAGA
5941	AGGCAGTGCT	CCCCGGCGGA	CGCGTGTGTT	CTTCGCGAGC	GCCTCGCTGA	GGTTCAGCGC
6001	CGCTGGCAGG	AAATCTGCTC	AGCGGTCCCT	CGCTGCCTCG	AACGCGACCG	CAAGGCTCGC
6061	ACCGAACTCA	AGGGGTTTGA	GGAGGGGACA	TCTCTGCAAG	CGTCTCTGCG	TGAGAGGATT
6121	GAAAGCCTCT	CCCCTGGGCT	GGACGAAGCC	CAGAGAATTG	TGGGTATTCA	GCCGGACCCA
6181	GACAACGAGA	TCCAAATCCA	GGAAACCCTG	GATAAGGTCA	AGGTGTCTGG	CACGCTTCCG
6241	GAGCGGGACA	GGCACCTGTC	GGATCTCCTG	CGCACTCTGA	CGGAGCTGCA	GCAGGAGGTG
6301	AGGCAGCTGG	GTGCCTGGAC	TGCCTCCACC	AAGGAGCAGC	TGGTGACCTT	CCAGCGGGCG
6361	GGCGTTCCCG	CCGACTCGTC	CCTGCAGGCC	GCTGAAGAGA	CGGTGCGAGC	CAAGAAGGCC
6421	GATGTAGAGG	AGACAATCTC	AAAAGCGGAG	CGGCTGGAGA	GGAAGGAGAC	TCTAAATCCA
6481	GCCGCAAAGG	AGCGGCTGGT	ACAGCTCAAA	TCCGACTGGG	CAGTTTTGAA	TGACCTGCTC
6541	AAAGACCTCA	AGCAGGCAAC	GCCTGAGGTG	AAGAAAGTGC	CCGGAGCCCA	CACAGCTGTA
6601	CAGACCCTCA	CCGTGGTGAC	GGTTACCAAG	GAGACGATCG	CCATGAAGCC	CGAGGCGGGC
6661	GCCACGGCGA	CGCCGCCGCT	GGCGGCCGAG	CTCACCGATT	GGCTCACGCT	GCAGGATCAC
6721	ATGATCCGCT	CGCACGTTGT	CACCGTCGGG	GACCTGGACG	AGATATCGGA	AGCAATACGG
6781	AAGCTGCAAA	GGGTGTTGGC	AGACCTGGAC	CAGAGGCGAC	CTCAGTTGGA	AGAGCTCATC
6841	ACGATATCTC	AAAACCTGAA	GAACAAAACG	GGTGACCCAG	AGACTCGAGC	TAACCTCACT
6901	GGAACCATCG	ACCGCGCGAT	GACGCTGTAC	GAGGACCTGC	AGCGACGCGT	CGCCGGGCGG
6961	CAGGCCCAGC	TGCAGGAGAT	GCTGCACGAC	TCGGCGCAGT	GGCACGAGGG	CCGCCAGGAG
7021	GCCGAGCGGC	TCGTGGCCGC	GGCCGAGGGA	CGCCTGGCGC	GCCTGGCGCA	GCCCCCTGCC
7081	AACTCGAGCG	AGCACATCGC	GAGGCAGATC	GGCGAGCTCA	AGAGTTTGGC	CAAGGAGCTC
7141	CGACAGTGGC	AGATAAGCGT	CGACGCGGCG	AACGACCTCG	CCCAAAAGCT	GCTTCGCGAC
7201	TACTCCGCCG	ATGACACCCA	CAAAGTCGAA	GAGATGACGG	GGAAACTCAA	CATGGCTTGG
7261	TCGGACATCA	TCCGACGCCCT	GAACGACCGC	GAGGCCGCCG	TGGAGCAGGC	CGGCAAGCGC
7321	GTTACAGCAGT	ACCACCTGGA	CCTGGAGCGC	TTCCTGGCGT	GGCTCTCCGA	GGCCGAGACC
7381	ACGGCCAACG	TGCTGGAGGA	CGCGAGCCAG	AAGGAAGGCG	TGCTGGAGGG	CTCGCGGGAC
7441	GCACGCAACG	TCCAGCGGCA	GTGGAAGGAA	CTGCAGGCAG	AGATCGACGC	CCATCAGGAG
7501	GTCCAGCGAG	GGCTGGACGA	TGGCGGCGCG	CGAGTGGTGG	CCTCCCTGCA	GGGCACGGAG
7561	CCAGCCGCCG	AGCTCCAGCG	CCGGCTGGAC	AACATGAACC	GCCGCTGGGC	CGACCTCAAA
7621	CGCAGGGCGC	TCACCATCAG	GGCTCGCCTG	GACCGGGATA	CCGAGCAGTG	GGGGCGACTC
7681	GGTCTCTCGC	TGGAGGAACT	CATCGCCTGG	GTGGACCTCA	AGGACTCGGA	GCTCGCGAGC
7741	GAGATGCCCA	TCGGTGGAGA	GGTGGCCATA	GTGCAGCAGC	AGCTGGACAA	CCACAGGGTC
7801	TTCCGCCGGG	ATCTGACCGG	GAATGAGCCG	ACGATCTTCA	ACGCGCTGGA	AGCTGCTCAC
7861	ATGTTCTCTG	CGGAGCAGCC	TATCGAGGCT	CCCACCGAGC	CTTTCCTCGA	GCAGAGAGAC
7921	ATTACGCAGG	AGGAGCCCGG	TCGCGCGTTC	GCGCGGTCCG	TCCGCCGCCA	GGCTCTGGAA
7981	CTGCAGCAGA	AGCAGGAGCG	ACTGTGGGCC	CTCTCCGACG	ACTGGCAGCG	GCGCCTCGAG
8041	CACGCGCTGG	AGCGGCTGGC	CGAGCTCCAG	GACGGTGCCG	ACGGCACAGT	GGTTCGCCCTG
8101	CGCCAGGCCG	AGATGCTGCG	CGACTCCTGG	CCGCCAGTCG	ACCAGCTGCT	GGTCACGTCC
8161	GTTCCCGAGC	AGGTGCGCCG	GGTGCAGGCG	TTCCGGGCGG	AGATCGCTCC	ACTGCAGGAG
8221	GAGGTGCGGC	AGGTGAACGA	CTTATCACGC	CGGTTTCCGC	CGCTCGACAT	CCAGCTCTCG
8281	CCCGCCAACC	TGCACCGCAT	CGAAGACGTC	AACATCCGGT	GGAAACGCCT	GCAGGCGAAA
8341	TCTGTCCAGA	CTCCCTGGGA	GAAGGCCGTT	TCCACCACCA	AAGTGCCCTA	CTACATCAAC
8401	CACCAGACTC	AGACAACCTT	TTGGGACCAT	CCCAGAATGA	CCGAACCTTA	CCAATCACTA
8461	TCTGACCTCA	ACAACGTGCG	ATTCTCGGCC	TACCGCACCG	CCATGAAGCT	ACGGCGACTC
8521	CAGAAGGCAC	TCTGCTTGGA	CCTCGTGAGC	CTGGATGCGC	TCACCAACGT	GTTCGAGCAG
8581	TCGGGCCTCA	AGCAGAACGA	CCGCCTCATG	GACGTGTCGG	AGGCCGTCGG	CTGCCTCACC
8641	ACCATCTACC	AGGGCCTGGA	GCGCAAGCAC	AGCGACCTGG	TGAACCTGCC	CCTCTGCGTC
8701	GACATGTGTC	TCAACTGGCT	GCTCAACGTT	TACGACACGG	GGCGTGGAGG	CAAGATACGA
8761	GAGCTGTCCT	TCAAGACAGG	GCTCGTTTCT	CTATCCAACG	CTCACCTCGA	GGACAAGTAC
8821	AGATTCTCTG	TCAAGCAGGT	GGCGAGCGTG	TCGGGCCTCT	GTGACCAGCG	GCGCCTGGGC
8881	CTCATGCTGC	ACGACGCCGT	GCAGATCCCG	CGGCAGCTCG	GCGAGGTGCG	CTCGTTCCGG
8941	GGCAGCAACA	TTGAGCCAG	CGTGCGCAGC	TGCTTCCAGT	TCGCTAACGG	GCGGCCCGAG
9001	ATCGAGGCGA	CCATGTTCTT	GGAGTGGATG	CGGATGGAGC	CACAGTCCAT	GGTGTGGCTA

9061 CCCGTACTGC ATCGGGTGGC CGCCTCGGAG ACCGCCAAGC ACCAGGCCAA GTGCAACATC
9121 TGCAAGGAGT GTCCCATCGT TGGCTTCCGG TACCGGAGCC TGAAGCATTT CAATTATGAC
9181 GTGTGCCAAA GCTGCTTCTT CTCCGGCCGG ACAGCCAAGG GCCACAAGCT CAGCTACCCG
9241 ATGGTGAAT ATTGCACGCC GNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9361 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9421 NNNNNNNNNN NNNNNNNGCC GTCTCCATCT CCGCAGCTCT CGAACGAAGA CACCCATGCC
9481 CGCATCGAGC ACCTGGCCAG CAGGTTGGCA GAGATGGAGT TCAAGAACGG ATCTTTCCTG
9541 AATGACAGCT CGCCCAATGA GAGCTTGGAG GACGAGCACT TGCTGATCCA GCACTATTGT
9601 CAGAGCCTGA GTGGAGAATC CCCAGGCAGC CAGCCGCAGA GCCCCGCACA GATCCTCATT
9661 TCTCTCGACA GCCAAGAGCG CAGCGAACTC GAGCGCATCC TGCGGGACTT GGAGGACGAA
9721 AACAGGAGCC TGCAGTCGGA GTACGATCGC CTGCGGCAGA AGCAGGAGCT GGGTGGCCTG
9781 GGC GCGCTGC TGCCGGCCCC TCCCACCATG CCACCTGAGA GCGCCCGTGA CTCGGAGCTC
9841 ATCGCAGAGG CCAAGCTGCT GCGGCAACAC AAGGGTCGGC TGGAAGCTCG CATGCAGGTG
9901 TTGGAAGATC ACAACCGTCA GCTCGAGTCC CAGCTGATTC GCTTGCGGCA GCTGCTGGAC
9961 CAG

//

Cmil dystrophin mRNA.docx

LOCUS Cmil_scaffold_11 10890 bp DNA linear SYN 17-Sep-2019

DEFINITION -

ACCESSION -

KEYWORDS -

SOURCE -

FEATURES

source

Location/Qualifiers

1..10890

/chromosome="Unknown"

/country="Australia: Tasmania"

/db_xref="taxon:7868"

/isolate="IMCB2004"

/mol_type=genomic DNA

/note="Hobart [Split]"

/organism="Callorhinchus milii"

/sex="male"

/tissue_type="testis"

CDS

1..10890

/note="Callorhinchus milii dystrophin mRNA"

mRNA

1..>831

/note="Transcription of Cmil dystrophin gene"

CDS

1..31

/codon_start=1

/db_xref="GI:632971674"

/db_xref="GeneID:103185537"

/gene="LOC103185537"

/note="Derived by automated computational analysis using gene prediction method: Gnomon."

/product="dystrophin-like"

/protein_id="XP_007902287.1"

/translation="MLWYEEPGEHYEREDVQKKTFTKWINSQFTKTRKPVIDDLFTDLC
DGRRLLELLEGLVGVDIVKEKGSTRVHALNNVNRALQILQKNNVDLVNIGGSDIVDGNH
KLTLGLIWSIILHWQVKDVMKGIMANLQQTNSEKILLSWVRQSSISNYTSVNVVNFTSSW
SNGLAFNALIHSRPLDFDWKNVAQQQSVTERLDHAFTVAKEHLGIDKLLDPEDMAKNA
DEVETGMYDKLPIHSVFPLFLKKFDFLTACVSIDVCTDNESSVLKGLYTFIDYVL"

mRNA

<1..31

/db_xref="GI:632971673"

/db_xref="GeneID:103185537"

/gene="LOC103185537"

/note="Derived by automated computational analysis using gene prediction method: Gnomon. Supporting evidence includes similarity to: 83% coverage of the annotated genomic feature by RNAseq alignments [Split]"

/product="dystrophin-like"

/transcript_id=XM_007904096.1

gene

<1..>31

/db_xref="GeneID:103185537"

/gene="LOC103185537"

/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"

CDS

32..93

mRNA

32..93

gene

<32..>93

/db_xref="GeneID:103185537"

	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	94..186
CDS	94..186
gene	<94..>186
	/db_xref="GeneID:103185537"
	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<187..>264
	/db_xref="GeneID:103185537"
	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	187..264
CDS	187..264
CDS	265..357
mRNA	265..357
gene	<265..>357
	/db_xref="GeneID:103185537"
	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	358..530
CDS	358..530
gene	<358..>530
	/db_xref="GeneID:103185537"
	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	531..649
CDS	531..649
gene	<531..>649
	/db_xref="GeneID:103185537"
	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<650..>655
	/db_xref="GeneID:103185537"
	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	650..655
CDS	650..655
mRNA	656..>831
gene	<656..>831
	/db_xref="GeneID:103185537"
	/gene="LOC103185537"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	656..>831
CDS	832..1002
	/codon_start=1

```

/db_xref="GI:632971672"
/db_xref="GeneID:103185536"
/gene="LOC103185536"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon."
/product="dystrophin-like"
/protein_id="XP_007902286.1"
/translation="MATEHPDKKSIIMYITSLFQVLPQKVITIEAIQEVELTPQQPKVTQ
EERREMQRFSQQQLSEVSEDKPAGSSLMETEVLDLSYQIELEEVLTWLLSAEDSLQAQDE
VSDDVEEVKEQFHIHEGFMELTAHQGSVGNILQAGNQLIVDGKLDDEEENEIQEQMNL
LNSRWENLRVASMERQSKLHEVLMDLQQQQLKHLAEWLTRTEVRIHQMEAQPLGPDLEA
LKQQVEEHKVLQEDLEQEQRVNSLTHMVVVDESNGESATVALEEELQMVHVNTALSS
QGALMSSQAYFKTGS LHLETVE"
mRNA      <832..1002
           /note=" [Split]"
gene      <832..>1002
           /db_xref="GeneID:103185536"
           /gene="LOC103185536"
           /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
mRNA      832..>1740
           /note="Transcription of Cmil dystrophin gene"
CDS       1003..1191
gene      <1003..>1191
           /db_xref="GeneID:103185536"
           /gene="LOC103185536"
           /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
mRNA      1003..1191
CDS       1192..1373
mRNA      1192..1373
gene      <1192..>1373
           /db_xref="GeneID:103185536"
           /gene="LOC103185536"
           /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
CDS       1374..1524
mRNA      1374..1524
gene      <1374..>1524
           /db_xref="GeneID:103185536"
           /gene="LOC103185536"
           /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
mRNA      1525..1644
gene      <1525..>1644
           /db_xref="GeneID:103185536"
           /gene="LOC103185536"
           /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
CDS       1525..1644
CDS       1645..>1740
mRNA      1645..>1740
gene      <1645..>1740
           /db_xref="GeneID:103185536"

```

```

/ gene="LOC103185536"
/ note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
CDS 1741..>6639
/ note="Cmil dystrophin"
gene <1741..>1827
/ db_xref="GeneID:103185535"
/ gene="LOC103185535"
/ note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
mRNA 1741..>6639
/ note="mRNA for Cmil ortholog to Hsap dystrophin"
mRNA 1741..>6639
/ note="Transcription of Cmil ortholog to Hsap dystrophin
gene"
mRNA <1741..1827
/ db_xref="GI:632971669"
/ db_xref="GeneID:103185535"
/ gene="LOC103185535"
/ note="Derived by automated computational analysis using
gene prediction method: Gnomon. Supporting evidence
includes similarity to: 2 Proteins, and 98% coverage of
the annotated genomic feature by RNAseq alignments,
including 2 samples with support for all annotated introns
[Split]"
/ product="dystrophin-like"
/ transcript_id=XM_007904094.1
CDS 1741..1827
/ codon_start=1
/ db_xref="GI:632971670"
/ db_xref="GeneID:103185535"
/ gene="LOC103185535"
/ note="Derived by automated computational analysis using
gene prediction method: Gnomon."
/ product="dystrophin-like"
/ protein_id="XP_007902285.1"
/ translation="MEKEKSVNKIQTTGFKDQNEVLANLRKLVILKGDLEMKRQVMDRL
CSLSQDLLTALQNKIVAQKLQARLEEFQQRWDNLVQRLNENSKMVSQFVTTTSQTSLTQT
TVMETVTMTVTKEQIQIKPKGEHSPPPPKRQVLLDSELKRKRFVDVDTTELHSMTRSEA
ILQSPEFAIYRKEASISDLQEKVNAIEKEKPEKYRKLQDANRSAQALVEQIVNEGLSAD
TIKHASQQLTNRWEEFCQLLVERTAWLQYQSKI IAVSFELQQLEQALISAETWLKAQQA
PATDSDTVKIQLDKCKDEIVRLTALQPQVDKLAEHGAELKQKEKVPKSLETDITAFNNH
FSQVLADLQARAKLLQSILSSLPTLRYRERVNTLALWQRSEAKLTVPCVAVTEYEIME
QRLRELKALQSSFQEQQSDNLNLT TTTTEEISRKAPAEISQKYRSEIEGLLGRWKKLSSS
LMEQSQKLEELMTKLQQFQNDTKTLKKWMAEVDVFLNEEWPALGDSEALEKQLEQCTAL
VNDIQMIQPSLNSINEVGEALKKQAEPAFANKLQTELAELNACWENICKQAHTKKTALK
GGLDKTM SLRKDLSEMQEWIAQAEEDYLERDFEYKTPEVLKKAVEELKRAKEDILQKEV
KVKLLSDGVNSFIAKAPPAHEALKSELDMLTGNYQRLCSRLNGKYNTFEEVWSCWCEL
QSYLEAENRWLDELEHKLGETENLQGGAAEISEALDSLENVMRHPEENRNQIRELAQTL
TDGEILDELINEKLEMFNTRWEELLQQAGRRQQVLVQSIQSAQETDKTLRLIQESLVLI
DKQLTSYIADRIDAAQVPQEAQKIQSELSSHEISLEEMKKRNQAINGAGQSKSAGRVLS
QIDVAQKKQQEVSTKFRLFQK PANFEQHLL ECKRLLDEVKAELDKLETSSVEPDFVQSQ
LDNFMKLYKSLSEVKSEVETVIKTGRQIVQKQQTENPKELDERLTALKLQYNELGAQVT
EGKQELEKCLKLSRKL RKEVNALTEWLAATDAELTKRSAVEGMPTNLDAEVAWSKAIQK

```

```

DIEKRRLLLNSITETGDSLRILMKKSNSAVDDKLSLLNCNWIAVTSRAEEWLNLLLEYQ
KQMFNFDQNVAQITAWMYRAEILLDESEKQKVPQKEEVLKRLKAELNDMRLKVDVRDQ
AVELMTNRGDHCREVVEPKLSELNQRFDDISHRIKTGKQLVLEHLSASPGIIVMESSEL
KPQKSFLAVQKKVQQQVSPKELEQFDTALQQMLGAVEAEIQQGENLNESDFKKETNEDD
ERRVNELTKRGDSLQQSITDERKQEEIRMKQQMLQAKFNILKDLRSIRRKKALEISPHW
YQYKKQTNNLLQWLDEIEQRLARLPEPKDDQKIKEMDAELEEKRGDLNELCRQAQSLSK
EGAAKALEPQLIQLNKRWREIENKFAQYRRLNYAKISVYEESTVTALTEGVMPETSYP
SAYLSEVYLLHQAVCEVEVRLNSLERQVKDYEDFSKQEDCLKTIKETLEKLSGRVEAVH
SKKPAVLQSASLTEAEQVRNAVAQLDAKWETVNKMYTNRQMKFDRSIEKWQGFHCDLNG
LDQWLTEAEQILAEVQFVEPAKSKQKYKVIVLQSLSHAALLESILLYFCHIILES"

mRNA 1828..2007
gene <1828..>2007
/db_xref="GeneID:103185535"
/gene="LOC103185535"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

CDS 1828..2007
CDS 2008..2177
gene <2008..>2177
/db_xref="GeneID:103185535"
/gene="LOC103185535"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

mRNA 2008..2177
gene <2178..>2301
/db_xref="GeneID:103185535"
/gene="LOC103185535"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

CDS 2178..2301
mRNA 2178..2301
mRNA 2302..2389
CDS 2302..2389
gene <2302..>2389
/db_xref="GeneID:103185535"
/gene="LOC103185535"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

CDS 2390..2631
mRNA 2390..2631
gene <2390..>2631
/db_xref="GeneID:103185535"
/gene="LOC103185535"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

mRNA 2632..2812
CDS 2632..2812
gene <2632..>2812
/db_xref="GeneID:103185535"
/gene="LOC103185535"
/note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

gene <2813..>2958
/db_xref="GeneID:103185535"

```

	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	2813..2958
CDS	2813..2958
CDS	2959..3171
mRNA	2959..3171
gene	<2959..>3171
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<3172..>3285
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	3172..3285
CDS	3172..3285
CDS	3286..3441
mRNA	3286..3441
gene	<3286..>3441
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<3442..>3612
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	3442..3612
CDS	3442..3612
CDS	3613..3795
mRNA	3613..3795
gene	<3613..>3795
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	3796..3930
mRNA	3796..3930
gene	<3796..>3930
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<3931..>4080
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	3931..4080
CDS	3931..4080

CDS	4081..4242
mRNA	4081..4242
gene	<4081..>4242
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<4243..>4371
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	4243..4371
CDS	4243..4371
mRNA	4372..4545
CDS	4372..4545
gene	<4372..>4545
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<4546..>4701
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	4546..4701
CDS	4546..4701
CDS	4702..4872
mRNA	4702..4872
gene	<4702..>4872
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<4873..>5052
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	4873..5052
CDS	4873..5052
CDS	5053..5181
mRNA	5053..5181
gene	<5053..>5181
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	5182..5352
mRNA	5182..5352
gene	<5182..>5352
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"

	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	5353..5463
mRNA	5353..5463
gene	<5353..>5463
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	5464..5580
gene	<5464..>5580
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	5464..5580
gene	<5581..>5718
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	5581..5718
CDS	5581..5718
CDS	5719..5871
mRNA	5719..5871
gene	<5719..>5871
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<5872..>6054
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	5872..6054
CDS	5872..6054
gene	<6055..>6249
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	6055..6249
CDS	6055..6249
CDS	6250..6422
mRNA	6250..6422
gene	<6250..>6422
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
gene	<6423..>6639
	/db_xref="GeneID:103185535"
	/gene="LOC103185535"

mRNA
CDS
CDS

```
/note="Derived by automated computational analysis using  
gene prediction method: Gnomon. [Split]"  
6423..>6639  
6423..>6639  
6640..6744  
/codon_start=1  
/db_xref="GI:632974305"  
/db_xref="GeneID:103186406"  
/gene="LOC103186406"  
/note="Derived by automated computational analysis using  
gene prediction method: Gnomon."  
/product="dystrophin-like isoform X2"  
/protein_id="XP_007903602.1"  
/translation="MNESGGVSLKSEALTPEQHFKLETELLEVNNRFTKVSHKLPEKQK  
EIEELLKDFVRYQQQLNHLALWAASTKDQLELYKQAPVPGAVDIKETEEGVLAKQADVE  
SILSKAHYFYKERPIAQPLKRKIEDLNSDWKTINKLLQDLNQRHQLGVGSTKQQTIVVTQ  
TLVTTETTSFTKLEMPSSLLMEVPALTDFNKAWVELTDWLSLLDRVIKSQIVTVGDVEEI  
NDMI IKQKATLQDLQKRPQLEELITAAQNLKNKTTNQEARAVITDRIEKVQTQWDEVQ  
NRLQARTLQLQEMLKDSNQWLEAKQEAEEHLERAKTKQGTGKVMSYTVDSLQNAELK  
QFFKELRQWQIHIDVANDLAGKLLRDYSTDDTRKVGMLADNINASWAAINKRVSAREAA  
LEAALRMLQQFYLDLDKFAWLTEAETTANVLQDATYKEGLDDAQLVRELTKQWQDLQ  
SEIEAHTDIYHNLDENGQKILKSLQGSSEDAILLQRRLLDMNSRWTELKKSLNIRACLE  
ASAEQWKRLQISLQELLVWLNKNDDELKSRMPVGGDVTAVQQQNDVHRSFKRELKSKEP  
VIMSALDTVRLFLADQPVEGPESLFPPLTKEKTPEENAQSARVIQKQSEDVKIEWEKL  
IRSVDWQKRLDEALERLLELQDAMDELNLKLRKAESIRDSWQPVGDLIDSLQDHIEKV  
KAFRAEIAPLKENVTHMNDLACQFTPTQIQLSPPYNLNRLLDNLNTRWKQLQVAIEEHLKQ  
LHEAHRDFGPTSQHFLSTSVQGAWEAISPKNVPYYINAKTAIGSKLSHQTTCTCWDHP  
KMTELYQSLADLNNVRFSAYRTAMKLRLQKALCLDLLSLPSACEAFDQRNLKQNDQLM  
DIMEIINCLTTIYDRLEQEHSLNVNPLCVDMLNWLNNVYDTGRTGKIRVLSFKTGIM  
SLCKAHLEDKYRYLQKQVASFPTGFCDQRRLLGLLLHDAIQVPRQLGEVASFGGSNIEPSV  
RSCFQFANNKPDIEAALFLDWMRLEPQSMVWMPVLHRVAAAETAKHQAKCNICKECPII  
GFRYRSLKHFNVDVCQSCFFSGRTAKGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFR  
TKRYFAKHPRMGYLPVQTVLEGDNLETPVTLINFWPVDYEPASSPQLSHDDTHSRIEY  
ASRLAEMENRNGSYLNDISPNESIDDEHLLIQHYCQSLNQESPLSQPRSPAQILISLE  
SEERGELELRLVSLDLEENRNLRQSEYERLQKQHDHKGSLPLSPPEMIPVSPQSPRDAEL  
IAEAKLLRQHKGRLEARMQILEDHNKQLESQHLRLRQLLEQPQTEVRVNGTTVSSPSSS  
SQRSDSSQPVLLHGVGSQTSEMLGEDELNPPQGSSTGLEDMVQNLNNSFPSSRSNNVG  
NLFHVVDLGRAMETLVTVMTDDEAAE"
```

CDS

```
6640..6744  
/codon_start=1  
/db_xref="GI:632974303"  
/db_xref="GeneID:103186406"  
/gene="LOC103186406"  
/note="Derived by automated computational analysis using  
gene prediction method: Gnomon."  
/product="dystrophin-like isoform X1"  
/protein_id="XP_007903601.1"  
/translation="MNESGGVSLKSEALTPEQHFKLETELLEVNNRFTKVSHKLPEKQK  
EIEELLKDFVRYQQQLNHLALWAASTKDQLELYKQAPVPGAVDIKETEEGVLAKQADVE  
SILSKAHYFYKERPIAQPLKRKIEDLNSDWKTINKLLQDLNQRHQLGVGSTKQQTIVVTQ  
TLVTTETTSFTKLEMPSSLLMEVPALTDFNKAWVELTDWLSLLDRVIKSQIVTVGDVEEI  
NDMI IKQKATLQDLQKRPQLEELITAAQNLKNKTTNQEARAVITDRIEKVQTQWDEVQ  
NRLQARTLQLQEMLKDSNQWLEAKQEAEEHLERAKTKQGTGKVMSYTVDSLQNAELK  
QFFKELRQWQIHIDVANDLAGKLLRDYSTDDTRKVGMLADNINASWAAINKRVSAREAA
```

```

LEAALRMLQQFYLDLDKFRAWLTEAETTANVLQDATYKEGLDDAQLVRELTKQWQDLQ
SEIEAHTDIYHNLDENGQKILKSLQGSEDAILLQRRLLDDMNSRWTELRRKKSINIRACLE
ASAEQWKRLQISLQELLVWLNLNKNDLKSMPVGGDVTAVQQQNDVHRSFKRELKSKEP
VIMSALDTVRLFLADQPVEGPESLFPLTKEKTPEENAQSARVVIQKQSEDEVKIEWEKLK
IRSVDWQKRLDEALERLLELQDAMDELNLKLRKAESIRDSWQPVGDLLIDSLQDHIEKV
KAFRAEIAPLKENVTHMNDLACQFTPTQIQLSPPYNNRLDDLNRWKQLQVAIEEHLKQ
LHEAHRDFGPTSQHFLSTSVQGAWERAISPKNVPYYINAKTAIGSKLSHQTTTCWDHP
KMTELYQSLADLNNVRFSAYRTAMKLRLQKALCLDLLSLPSACEAFDQRNLKQNDQLM
DIMEIINCLTTIYDRLEQEHSLNVNPLCVDMLNWLNNVYDTGRTGKIRVLSFKTGIM
SLCKAHLEDKYRYLFKQVASPTGFCQRRLLGLLLHDAIQVPRQLGEVASFGGSNIEPSV
RSCFQFANNKPDIEAALFLDWMRLPQSMVWMPVLHRVAAAETAKHQAKCNICKECPII
GFRYRSLKHFNVDVCQSCFFSGRTAKGHKMHYPMVEYCTPTTSGEDVRDFAKVLKNKFR
TKRYFAKHPRMGYLPVQTVLEGDNLETPASSPQLSHDDTHSRIEHYASRLAEMENRNGS
YLNDISPNEIDDEHLLIQHYCQSLNQESPLSQPRSPAQILISLESEERGELERVLSL
LEEENRNQLQSEYERLKQQHDHKGSLPLSPPEMIPVSPQSPRDAELIAEAKLLRQHKGK
LEARMQILEDHNKQLESQHLRLRQLLEQPQTEVRVNGTTVSSPSSSSQSRSDSSQPVLHH
GVGSQTSEMLGEDELMNPPQGSSTGLEDVMEQLNNSFPSSRSNNVGNLFHVVDLGRAM
ETLVTVMTDDEAAE"
mRNA <6640..6744
      /note=" [Split]"
mRNA <6640..6744
      /note=" [Split]"
gene <6640..>6744
      /db_xref="GeneID:103186406"
      /gene="LOC103186406"
      /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
CDS 6745..6909
CDS 6745..6909
gene <6745..>6909
      /db_xref="GeneID:103186406"
      /gene="LOC103186406"
      /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
mRNA 6745..6909
mRNA 6745..6909
gene <6910..>7011
      /db_xref="GeneID:103186406"
      /gene="LOC103186406"
      /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
mRNA 6910..7011
mRNA 6910..7011
CDS 6910..7011
CDS 6910..7011
gene <7012..>7096
      /db_xref="GeneID:103186406"
      /gene="LOC103186406"
      /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"
mRNA 7012..7096
mRNA 7012..7096
CDS 7012..7096
CDS 7012..7096

```

```

gene          <7097..>7329
               /db_xref="GeneID:103186406"
               /gene="LOC103186406"
               /note="Derived by automated computational analysis using
mRNA          7097..7329
mRNA          7097..7329
CDS           7097..7329
CDS           7097..7329
CDS           7330..7447
CDS           7330..7447
gene          <7330..>7447
               /db_xref="GeneID:103186406"
               /gene="LOC103186406"
               /note="Derived by automated computational analysis using
mRNA          7330..7447
mRNA          7330..7447
CDS           7448..7659
CDS           7448..7659
mRNA          7448..7659
mRNA          7448..7659
gene          <7448..>7659
               /db_xref="GeneID:103186406"
               /gene="LOC103186406"
               /note="Derived by automated computational analysis using
gene          <7660..>7814
               /db_xref="GeneID:103186406"
               /gene="LOC103186406"
               /note="Derived by automated computational analysis using
mRNA          7660..7814
mRNA          7660..7814
CDS           7660..7814
CDS           7660..7814
CDS           7815..8004
CDS           7815..8004
mRNA          7815..8004
mRNA          7815..8004
gene          <7815..>8004
               /db_xref="GeneID:103186406"
               /gene="LOC103186406"
               /note="Derived by automated computational analysis using
CDS           8005..8177
CDS           8005..8177
mRNA          8005..8177
mRNA          8005..8177
gene          <8005..>8177
               /db_xref="GeneID:103186406"
               /gene="LOC103186406"
               /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

```

CDS	8178..8334
CDS	8178..8334
mRNA	8178..8334
mRNA	8178..8334
gene	<8178..>8334
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
gene	<8335..>8455
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
mRNA	8335..8455
mRNA	8335..8455
CDS	8335..8455
CDS	8335..8455
mRNA	8456..8724
mRNA	8456..8724
CDS	8456..8724
CDS	8456..8724
gene	<8456..>8724
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
gene	<8725..>8871
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
mRNA	8725..8871
mRNA	8725..8871
CDS	8725..8871
CDS	8725..8871
gene	<8872..>8950
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
mRNA	8872..8950
mRNA	8872..8950
CDS	8872..8950
CDS	8872..8950
CDS	8951..9011
CDS	8951..9011
mRNA	8951..9011
mRNA	8951..9011
gene	<8951..>9011
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"

CDS	9012..9041
mRNA	9012..9041
mRNA	9012..9041
gene	<9012..>9041
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
CDS	9012..9041
gene	<9042..>9103
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
CDS	9042..9103
CDS	9042..9103
CDS	9042..9103
CDS	9042..9103
mRNA	9042..9103
mRNA	9042..9103
mRNA	9042..9103
mRNA	9042..9103
CDS	9042..9103
CDS	9042..9103
mRNA	9042..9103
mRNA	9042..9103
CDS	9104..9178
CDS	9104..9178
CDS	9104..9178
CDS	9104..9178
mRNA	9104..9178
mRNA	9104..9178
mRNA	9104..9178
mRNA	9104..9178
CDS	9104..9178
CDS	9104..9178
mRNA	9104..9178
mRNA	9104..9178
gene	<9104..>9178
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using
	gene prediction method: Gnomon. [Split]"
CDS	9179..9380
CDS	9179..9380
CDS	9179..9380
CDS	9179..9380
mRNA	9179..9380
mRNA	9179..9380
mRNA	9179..9380
mRNA	9179..9380
CDS	9179..9380
CDS	9179..9380
mRNA	9179..9380

mRNA	9179..9380
gene	<9179..>9380
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	9381..9466
CDS	9381..9466
mRNA	9381..9466
CDS	9381..9466
mRNA	9381..9466
mRNA	9381..9466
CDS	9381..9466
CDS	9381..9466
mRNA	9381..9466
mRNA	9381..9466
gene	<9381..>9466
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	9381..9466
CDS	9381..9466
CDS	9467..9624
CDS	9467..9624
CDS	9467..9624
CDS	9467..9624
mRNA	9467..9624
mRNA	9467..9624
mRNA	9467..9624
mRNA	9467..9624
CDS	9467..9624
CDS	9467..9624
mRNA	9467..9624
mRNA	9467..9624
gene	<9467..>9624
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	9625..9791
CDS	9625..9791
CDS	9625..9791
CDS	9625..9791
mRNA	9625..9791
mRNA	9625..9791
gene	<9625..>9791
	/db_xref="GeneID:103186406"
	/gene="LOC103186406"
	/note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	9625..9791
mRNA	9625..9791
CDS	9625..9791

```

CDS          9625..9791
mRNA         9625..9791
mRNA         9625..9791
CDS          9792..9903
CDS          9792..9903
CDS          9792..9903
CDS          9792..9903
mRNA         9792..9903
mRNA         9792..9903
mRNA         9792..9903
mRNA         9792..9903
CDS          9792..9903
CDS          9792..9903
mRNA         9792..9903
mRNA         9792..9903
gene         <9792..>9903
            /db_xref="GeneID:103186406"
            /gene="LOC103186406"
            /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

CDS          9904..10040
CDS          9904..10040
CDS          9904..10040
CDS          9904..10040
mRNA         9904..10040
mRNA         9904..10040
mRNA         9904..10040
CDS          9904..10040
mRNA         9904..10040
mRNA         9904..10040
gene         <9904..>10040
            /db_xref="GeneID:103186406"
            /gene="LOC103186406"
            /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

CDS          9904..10040
mRNA         9904..10040
CDS          10041..10106
CDS          10041..10106
CDS          10041..10106
CDS          10041..10106
mRNA         10041..10106
mRNA         10041..10106
mRNA         10041..10106
mRNA         10041..10106
CDS          10041..10106
CDS          10041..10106
mRNA         10041..10106
mRNA         10041..10106
gene         <10041..>10106
            /db_xref="GeneID:103186406"
            /gene="LOC103186406"
            /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

```

CDS	10107..10172
CDS	10107..10172
CDS	10107..10172
CDS	10107..10172
mRNA	10107..10172
mRNA	10107..10172
mRNA	10107..10172
mRNA	10107..10172
CDS	10107..10172
CDS	10107..10172
gene	<10107..>10172 /db_xref="GeneID:103186406" /gene="LOC103186406" /note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	10107..10172
mRNA	10107..10172
CDS	10173..10331
CDS	10173..10331
CDS	10173..10331
CDS	10173..10331
mRNA	10173..10331
mRNA	10173..10331
mRNA	10173..10331
mRNA	10173..10331
CDS	10173..10331
CDS	10173..10331
mRNA	10173..10331
mRNA	10173..10331
gene	<10173..>10331 /db_xref="GeneID:103186406" /gene="LOC103186406" /note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
CDS	10332..10575
CDS	10332..10575
CDS	10332..10575
CDS	10332..10575
mRNA	10332..10575
mRNA	10332..10575
mRNA	10332..10575
mRNA	10332..10575
CDS	10332..10575
CDS	10332..10575
mRNA	10332..10575
gene	<10332..>10575 /db_xref="GeneID:103186406" /gene="LOC103186406" /note="Derived by automated computational analysis using gene prediction method: Gnomon. [Split]"
mRNA	10332..10575
CDS	10576..10699
CDS	10576..10699
CDS	10576..10699

```

CDS      10576..10699
mRNA     10576..10699
mRNA     10576..10699
mRNA     10576..10699
mRNA     10576..10699
CDS      10576..10699
CDS      10576..10699
mRNA     10576..10699
mRNA     10576..10699
gene     <10576..>10699
        /db_xref="GeneID:103186406"
        /gene="LOC103186406"
        /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

CDS      10700..10792
CDS      10700..10792
CDS      10700..10792
CDS      10700..10792
mRNA     10700..10792
mRNA     10700..10792
mRNA     10700..10792
mRNA     10700..10792
CDS      10700..10792
CDS      10700..10792
mRNA     10700..10792
mRNA     10700..10792
gene     <10700..>10792
        /db_xref="GeneID:103186406"
        /gene="LOC103186406"
        /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

CDS      10793..10804
CDS      10793..10804
CDS      10793..10890
CDS      10793..10890
mRNA     10793..>10890
        /note=" [Split]"
mRNA     10793..>10890
        /note=" [Split]"
mRNA     10793..>10890
        /note=" [Split]"
mRNA     10793..>10890
        /note=" [Split]"
CDS      10793..10890
CDS      10793..10890
mRNA     10793..>10890
        /note=" [Split]"
mRNA     10793..>10890
        /note=" [Split]"
gene     <10793..>10890
        /db_xref="GeneID:103186406"
        /gene="LOC103186406"
        /note="Derived by automated computational analysis using
gene prediction method: Gnomon. [Split]"

```

BASE COUNT	3430	A	2289	C	2775	G	2396	T	0	OTHER
ORIGIN	?									
1	ATGCTGTGGT	ATGAAGAGCC	TGGAGAACAT	TATGAAAGAG	AAGATGTTCA	AAAGAAAACA				
61	TTCACAAAAT	GGATAAATTC	GCAGTTTACA	AAGACTAGAA	AACCTGTAAT	TGACGACCTG				
121	TTTACTGACC	TTTGTGATGG	AAGAAGGCTT	CTCGAGCTCC	TGGAGGGTCT	TGTTGGTTAT				
181	GATATAGTAA	AAGAAAAGGG	ATCCACGAGG	GTTTCATGCAC	TGAACAATGT	CAACAGAGCA				
241	CTGCAGATCT	TGCAGAAGAA	CAATGTTGAC	TTGGTAAATA	TTGGAGGAAG	TGATATAGTG				
301	GATGGAAATC	ACAAACTTAC	CCTGGGCTTG	ATCTGGAGCA	TCATCCTCCA	TTGGCAGGTT				
361	AAAGATGTTA	TGAAGGGAAT	CATGGCAAAC	CTGCAGCAGA	CTAACAGTGA	GAAGATTCTA				
421	CTGAGCTGGG	TCCGTCAGTC	TATCTCCAAC	TATACATCAG	TCAACGTTGT	CAATTTCACT				
481	TCCAGTTGGT	CGAATGGCTT	GGCTTTCAAT	GCACTCATCC	ACAGTCACAG	GCCTGACCTC				
541	TTTGA CTGGA	AGAATGTGGC	TCAGCAACAG	TCTGTAACCG	AACGACTGGA	TCATGCCTTC				
601	ACTGTAGCCA	AAGAACACCT	GGGTATTGAT	AAGCTGCTGG	ATCCTGAAGA	TATGGCAAAA				
661	AATGCTGATG	AGGTGGAAAC	AGGAATGTAC	GACAAGCTAC	CAATTCATTC	CGTCTTTCCT				
721	TTGTTCCCTGA	AGAAATTTGA	TTTTTCTAACA	GCTTGTGTTT	CCATTGACGT	TTGCACGGAC				
781	AATGAATCTT	CAGTTCTGAA	AGGGCTTTAT	ACCTTTATTG	ACTATGTGCT	GATGGCAACT				
841	GAACATCCAG	ACAAGAAGTC	CATCATTATG	TACATCACAT	CGCTTTTCCA	GGTTCTGCCT				
901	CAGAAAGTAA	CCATCGAAGC	CATCCAGGAG	GTTGAAACCC	TGCCACAACA	GCCAAAGGTC				
961	ACTCAGGAGG	AACGAAGAGA	GATGCAGCGA	TTCTCCCAGC	AGCTTTCAGA	AGTCTCAGAG				
1021	GATAAGCCAG	CTGGAAGCTC	ACTGATGGAA	ACAGAAGTGG	ACCTGGACAG	CTACCAGATT				
1081	GAAGTGGAGG	AGGTGTTGAC	CTGGCTGCTT	TCAGCTGAAG	ATTCACTACA	GGCTCAGGAT				
1141	GAGGTTTCAG	ATGATGTGGA	GGAAGTCAAG	GAGCAATTTT	ATATCCATGA	AGGCTTTATG				
1201	ATGGAGCTAA	CAGCACATCA	AGGCAGCGTG	GGTAACATTT	TGCAGGCTGG	AAACCAACTC				
1261	ATTGTAGATG	GCAAACTAGA	TGAAGATGAA	GAAAATGAAA	TTCAGGAGCA	GATGAACTTG				
1321	CTCAACTCCC	GATGGGAAAA	TCTCAGGGTA	GCCAGCATGG	AGAGACAAAG	CAAGTTACAT				
1381	GAAGTGCTAA	TGGACCTTCA	ACAGCAACAG	CTGAAGCATT	TAGCTGAGTG	GTTAACAAGA				
1441	ACGGAGGTCA	GAATTCACCA	AATGGAGGCA	CAGCCCCTCG	GACCAGACCT	GGAGGCATTA				
1501	AAACAACAAG	TGGAGGAGCA	TAAAGTTCTT	CAGGAAGACC	TGGAACAGGA	GCAGGTCGGG				
1561	GTCAACTCTC	TGACTCATAT	GGTGGTCGTG	GTGGATGAGA	GCAATGGAGA	GAGTGCTACA				
1621	GTTGCTCTGG	AGGAAGAGCT	TCAGATGGTT	CATGTCAACA	CAGCATTGTC	ATCACAAGGA				
1681	GCACTGATGA	GCAGCCAAGC	ATATTTTAAAG	ACAGGATCTC	TACATTTGGA	GA CTGTTGAG				
1741	ATGGAGAAGG	AGAAATCTGT	TAACAAAATT	CAGACAACTG	GCTTCAAGGA	TCAAAATGAG				
1801	GTGCTGGCTA	ATCTGCGGAA	GTTAGTGATT	CTGAAAGGAG	ACTTGGAGAT	GAAACGACAA				
1861	GTCATGGACA	GGCTGTGTTC	GCTCAGCCAA	GATCTGCTGA	CAGCATTGCA	AAATAAGATT				
1921	GTAGCTCAGA	AATTACAGGC	GAGGCTTGAG	GAATTTGCCC	AGCGCTGGGA	TAATCTTGTA				
1981	CAGAGACTCG	AGAACAACCTC	CAAGATGGTT	TCACAATTTG	TCACCACATC	TCAGACATCG				
2041	CTAACGCAGA	CCACAGTGAT	GGAAACAGTA	ACTATGGTGA	CCACCAAGGA	GCAGATTCAG				
2101	ATCAAACCGA	AAGGGGAGCA	TTCTCCACCC	CCTCCTAAGA	AGAGGCAGGT	GCTTTTGGAC				
2161	TCTGAGCTCA	GGAAAAGATT	TGATGTTGAT	ACAACAGAGC	TTCACAGCTG	GATGACACGA				
2221	TCAGAAGCCA	TCCTTCAAAG	CCCAGAGTTT	GCAATTTACC	GAAAAGAAGC	AAGCATTTCA				
2281	GACCTTCAAG	AAAAAGTTAA	TGCTATAGAG	AAGGAGAAAC	CTGAAAAATA	CAGAAAACCTG				
2341	CAAGATGCCA	ATAGATCAGC	TCAAGCCTTA	GTGGAACAGA	TTGTGAATGA	GGGTCTTTCT				
2401	GCCGATACTA	TCAAACACGC	CTCACAGCAG	CTGACGAACC	GCTGGGAGGA	GTTTTGTCAA				
2461	CTGCTGGTGG	AACGGACGGC	CTGGCTGCAA	TATCAGAGCA	AGATTATTGC	TGTTTCCCTT				
2521	GAGCTACAAC	AACTTGAGCA	AGCTCTGATA	TCAGCAGAGA	CCTGGCTCAA	AGCACAGCAA				
2581	GCACCAGCCA	CTGACTCAGA	TACAGTAAAG	ATCCAGCTGG	ACAAATGCAA	GGATGAAATT				
2641	GTACGATTGA	CAGCACTGCA	GCCTCAGGTT	GATAAGCTGG	CTGAGCATGG	GGCTGAGCTA				
2701	AAACAGAAGG	AGAAGGTGCC	GAAATCTTTG	GAGACTGACA	TCACTGCCTT	CAACAATCAC				
2761	TTTTCTCAAG	TTCTTGCAAG	TCTGCAGGCA	CGAGCGAAAC	TGCTGCAGTC	CATTCTGAGC				
2821	AGCTTGCCCA	CACTTCGATA	TAGGGAGAGA	GTGAACACAC	TAGCCCTGTG	GATCCAGCGA				
2881	TCAGAAGCCA	AACTAACTGT	CCCTTGTTGT	GCAGTCACAG	AGTATGAAAT	CATGGAGCAG				
2941	CGACTCAGGG	AACTGAAAGC	ATTGCAAAGT	TCCTTTCAGG	AGCAGCAGAG	TGACCTGAAT				
3001	TATCTCACCA	CAACTACAGA	AGAGATTTCC	AGGAAAGCCC	CGGCAGAAAT	CAGCCAGAAA				
3061	TATCGATCAG	AGATCGAGGG	GCTCCTTGGC	CGTTGGAAGA	AATTATCCTC	TTCTTTGATG				

3121	GAGCAGAGCC	AGAAACTTGA	GGAGCTGATG	ACAAAGCTTC	AGCAATTTCA	GAATGACACA
3181	AAAACGCTGA	AGAAATGGAT	GGCTGAAGTG	GATGTTTTTC	TGAATGAAGA	ATGGCCTGCT
3241	CTGGGTGACT	CAGAAGCACT	TGAAAAGCAG	CTGGAGCAAT	GCACTGCACT	GGTAAATGAC
3301	ATCCAAATGA	TTCAACCCAG	TTTAAACAGT	ATTAATGAGG	TTGGCGAAGC	ATTGAAAAAA
3361	CAGGCTGAGC	CTGCATTTGC	CAACAAATTA	CAAACTGAAC	TCGCCGAGCT	GAATGCTTGC
3421	TGGGAGAACA	TCTGTAAACA	GGCACATACC	AAAAAGACTG	CATTAAAGGG	TGGCCTGGAT
3481	AAGACCATGA	GTCTCAGGAA	AGACCTGTCT	GAAATGCAGG	AATGGATAGC	ACAAGCTGAG
3541	GAAGATTACC	TGGAAAGGGA	TTTTGAGTAC	AAAACACCAG	AAGTGCTCAA	GAAGGCTGTG
3601	GAAGAACTAA	AGAGAGCGAA	GGAGGACATT	TTGCAAAAGG	AGGTAAAGGT	GAAGCTCCTC
3661	AGTGATGGTG	TTAACAGCTT	CATAGCAAAG	GCACCACCTG	CAGCCCATGA	AGCTCTGAAG
3721	TCTGAACTGG	ACATGCTGAC	GGGCAATTAC	CAGCGACTCT	GCAGCAGATT	GAATGGAAAA
3781	TACAACACAT	TTGAGGAGGT	GTGGTCCTGC	TGGTGTGAAT	TGCAATCATA	TCTGGAAGCA
3841	GAGAATAGAT	GGCTTGATGA	ACTGGAACAT	AAACTCGGGG	AAACAGAGAA	TTTACAGGGA
3901	GGAGCAGAGG	AAATCTCAGA	GGCACTAGAT	TCCCTAGAGA	ATGTAATGCG	TCACCCAGAG
3961	GAAAATCGAA	ATCAAATCCG	TGAGTTAGCA	CAGACACTGA	CTGATGGCGA	GATCTTGGAT
4021	GAGCTGATCA	ATGAAAAGCT	GGAGATGTTT	AACACTCGTT	GGGAGGAGCT	CTTACAACAG
4081	GCTGGAAGAA	GACAGCAAGT	CCTTGTGCAG	AGTATTCAGT	CGGCTCAGGA	GAAGGCTGAG
4141	ACCTTGCGTT	TAATCCAAGA	ATCTCTCGTG	TTGATTGACA	AACAGCTGAC	GTCATACATT
4201	GCTGACAGAA	TTGATGCAGC	CCAGGTCCCT	CAGGAAGCTC	AGAAAATCCA	ATCTGAGTTG
4261	AGTAGTCATG	AGATAAGTTT	GGAAGAAATG	AAGAAAAGAA	ACCAGGCCAT	TAACGGTGCG
4321	GGGCAGAGTA	AATCTGCAGG	ACGAGTCCTT	TCACAGATTG	ACGTGGCACA	GAAGAAGCAG
4381	CAGGAAGTGT	CCACCAAGTT	CCGCTTGTTT	CAAAAACCAG	CAAACCTTTGA	GCAGCATCTT
4441	TTGGAGTGTA	AGAGGCTCCT	GGATGAGGTG	AAAGCAGAAC	TGGATAAATT	GGAGACCAGC
4501	AGTGTGGAGC	CTGACTTTGT	GCAATCACAA	CTTGATAATT	TTATGAAGTT	GTATAAAAGT
4561	CTGAGTGAAG	TGAAGTCTGA	AGTGGAACG	GTAATTAATA	CTGGACGCCA	GATTGTGCAG
4621	AAACAGCAAA	CAGAGAATCC	CAAAGAAGTG	GACGAGCGGC	TCACAGCCCT	GAAACTGCAA
4681	TATAATGAAC	TTGGAGCCCA	GGTGACTGAA	GGGAAGCAAG	AATTAGAAAA	GTGTCTGAAA
4741	TTGTCCCGCA	AGCTAAGGAA	GGAAGTCAAT	GCCTTGACGG	AGTGGCTGGC	AGCAACAGAT
4801	GCAGAGCTGA	CAAAGCGATC	AGCGGTTGAA	GGAATGCCAA	CTAATTTGGA	TGCTGAAGTT
4861	GCTTGGAGTA	AGGCTATCCA	GAAAGATATA	GAGAAGCGGC	GGCTGCTATT	AAACAGTATT
4921	ACCGAGACAG	GAGACTCTCT	GAGGATCTTG	ATGAAGAAGA	GTAACAGTGC	TGTGGATGAT
4981	AAACTCAGTC	TTTTGAACTG	TAATTGGATC	GCAGTCACAT	CACGGGCTGA	AGAGTGGCTC
5041	AATCTTCTAC	TGGAATATCA	GAAACAGATG	GAAAATTTTG	ATCAGAATGT	GGCTCAGATT
5101	ACTGCATGGA	TGTATCGTGC	TGAGATTTTA	TTGGATGAAT	CTGAAAAACA	GAAAGTGCCA
5161	CAGAAAGAGG	AAGTTTTGAA	GCGTCTGAAG	GCTGAACTCA	ATGATATGAG	ATTGAAGGTG
5221	GATGCGGTGC	GTGACCAGGC	TGTGGAGCTG	ATGACAAACC	GTGGAGATCA	CTGTAGGGAG
5281	GTCGTGGAAC	CCAAGCTGTC	AGAGCTTAAT	CAGAGGTTTG	ATGACATCTC	GCACAGAATA
5341	AAGACTGGCA	AGCAATTGGT	TCTGGAACAT	TTATCTGCTT	CACCAGGAAT	TATTGTAATG
5401	GAGTCCAGTG	AACTGAAACC	CCAGAAGAGT	TTTTTGGCAG	TGCAGAAAAA	AGTGCAACAG
5461	CAGGTATCTC	CAAAGGAGTT	AGAACAGTTT	GATACAGCTC	TTCAGCAAAT	GCTGGGTGCA
5521	GTGGAAGCAG	AAATACAGCA	GGGCGAAAAA	CTCAATGAAA	GCGACTTCAA	GAAAGAGACG
5581	AATGAGGATG	ATGAGCGGAG	AGTGAATGAA	TTGACAAAAC	GGGGAGACAG	TTTACAACAG
5641	AGCATCACGG	ATGAGAGAAA	GCAAGAAGAA	ATTAGGATGA	AACAGCAGAT	GCTGCAGGCT
5701	AAATTTAACA	TTCTCAAGGA	CCTGAGATCC	ATTAGAAGGA	AAAAGGCTCT	TGAGATATCA
5761	CCTCACTGGT	ACCAGTACAA	GAAACAGACC	AACAATCTAT	TGCAGTGGTT	GGATGAAATT
5821	GAACAAAGGT	TAGCCCGTCT	TCCAGAACCC	AAAGATGACC	AGAAGATTAA	GGAGATGGAT
5881	GCTGAACTGG	AAGAGAAGAG	AGGAGACCTG	AATGAATTGT	GCAGACAAGC	ACAGAGCCTG
5941	TCAAAAGAAG	GGGCTGCCAA	GGCACTGGAG	CCGCAGCTGA	TCCAGCTCAA	CAAACGCTGG
6001	CGTGAAATTG	AAAACAAATT	TGCTCAGTAT	CGAAGATTGA	ACTATGCAAA	AATTAGTGTT
6061	TATGAAGAGA	GTACCGTAAC	TGCGCTGACT	GAAGGGGTAA	TGCCTGAAAC	CTCTTACGTG
6121	CCTTCTGCCT	ACCTGTCAGA	GGTATATCTG	CTTCATCAAG	CTGTGTGTGA	AGTGGAAGTT
6181	CGGCTCAACT	CTCTGGAGCG	TCAAGTCAAG	GACTATGAAG	ATTTCTCCAA	ACAGGAGGAT
6241	TGTTTTAAAGA	CCATAAAAGA	GACCCTGGAG	AAACTCTCTG	GCCGAGTAGA	GGCCGTTTAC
6301	AGCAAGAAAC	CTGCAGTCCT	CCAGAGTGCC	TCGCTCACGG	AAGCAGAGCA	GGTACGGAAT

6361	GCCGTGGCTC	AGCTGGATGC	CAAATGGGAA	ACAGTAAACA	AGATGTACAC	TAATCGACAA
6421	ATGAAGTTTG	ACAGATCCAT	TGAGAAATGG	CAAGGTTTCC	ATTGCGATTT	GAATGGCTTG
6481	GACCAGTGGC	TGACAGAAGC	AGAGCAGATC	CTTGCTGAGG	TCCAGTTTGT	AGAACCAGCT
6541	AAATCGAAAC	AGAAGTACAA	GGTAATTGTG	CTTCAATCAC	TCAGTCATGC	TGCTTTAGAG
6601	TCAATATTGC	TATATTTTGT	CCATATCATC	TTGGAATCAA	TGAACGAGAG	TGGAGGAGTG
6661	TCACTGAAAA	GCGAGGCTCT	GACTCCAGAG	CAACATTTCA	AACTGGAAAC	TGAACTCCTG
6721	GAGGTTAACA	ACCGCTTCAC	AAAGGTTTCC	CATAAACTAC	CTGAGAAACA	GAAAGAGATC
6781	GAGGAGCTAC	TGAAGGATTT	TGTTGCTAT	CAGCAGCAGC	TAAACCACTT	GGCACTGTGG
6841	GCCGCTTCCA	CAAAAGATCA	ACTTGAATTG	TACAAGCAAG	CACCAGTACC	AGGGGCTGTT
6901	GACATCAAGG	AGACGGAGGA	AGGAGTGTTA	GCGAAGCAGG	CTGATGTGGA	AAGTATTTTG
6961	TCCAAGGCGC	ATTACTTCTA	TAAAGAAAGG	CCCATCGCTC	AGCCATTAAA	GAGAAAAATA
7021	GAAGATCTAA	ATTCTGACTG	GAAGACGATT	AATAAGTTGC	TTCAAGATTT	GAATCAAAGG
7081	CATCAACTTG	GTGTAGGTTT	AACCAAACAG	CAAACGTGTT	TGACACAGAC	ACTAGTTACC
7141	ACAGAAACCT	CCTTTACCAA	GCTAGAAATG	CCATCCTCTT	TACTCATGGA	AGTGCCAGCG
7201	CTGACTGACT	TCAACAAGGC	CTGGGTGGAA	TTGACAGACT	GGCTCTCTTT	GCTGGATCGT
7261	GTGATCAAAT	CTCAGATTGT	GACAGTAGGG	GATGTCGAAG	AAATCAATGA	CATGATCATA
7321	AAACAGAAGG	CTACTCTGCA	AGACTTGGAC	CAAAAGCGCC	CTCAGCTGGA	GGAGCTTATT
7381	ACTGCAGCCC	AGAACCTGAA	GAACAAGACA	ACTAACCAGG	AGGCTCGTGC	AGTCATTACA
7441	GATCGCATTG	AAAAGGTTCA	AACCCAGTGG	GATGAGGTGC	AGAACCGACT	ACAAGCCCGC
7501	ACTCTGCAGT	TGCAGGAGAT	GCTCAAGGAC	TCCAACCAGT	GGCTTGAAGC	TAAACAAGAA
7561	GCAGAAGAAC	ATCTGGAGAG	AGCTAAAACC	AAGCAGGGAA	CTGGGAAAGT	AATGTCCTAC
7621	ACCGTGGACT	CATTGATGAA	GCAGAATGCA	GAGCTTAAGC	AATTTTTTCAA	GGAGCTGCGA
7681	CAGTGGCAGA	TACACATAGA	TGTGGCTAAC	GACCTCGCTG	GCAAGCTTCT	ACGTGACTAT
7741	TCCACGGATG	ACACCAGGAA	GGTGGGACTG	ATGGCAGATA	ACATTAATGC	TTCATGGGCG
7801	GCCATTAATA	AAAGAGTCAG	TGAGAGGGAG	GCTGCGCTGG	AAGCAGCCCT	CAGAATGCTG
7861	CAGCAGTTCT	ATCTGGATTT	GGACAAATTC	CGGGCCTGGC	TTACAGAAGC	AGAGACCACT
7921	GCCAATGTCT	TGCAGGACGC	CACCTACAAA	GAAGGACTAT	TGGATGATGC	CCAGCTGGTT
7981	CGAGAGCTGA	CCAAACAGTG	GCAGGATCTG	CAGTCAGAAA	TTGAAGCTCA	CACTGACATT
8041	TACCATAACT	TGGATGAAAA	TGGACAGAAA	ATACTTAAGA	GTCTGCAGGG	CTCGGAAGAT
8101	GCAATTTTGC	TACAGAGGCG	CCTCGATGAT	ATGAATTCTC	GTTGGACTGA	ACTCAGGAAA
8161	AAATCCCTTA	ACATTAGAGC	CTGTTTGGAG	GCAAGTGCTG	AACAGTGGAA	GCGACTCCAG
8221	ATCTCCCTAC	AGGAGCTCCT	GGTGTGGCTC	AACTTAAAGA	ACGATGAGCT	GAAGAGCCGA
8281	ATGCCTGTTG	GTGGTGATGT	TACTGCAGTA	CAACAGCAGA	ACGATGTACA	CAGAAGCTTC
8341	AAGAGGGAGT	TGAAGTCGAA	GGAACCAGTG	ATCATGAGTG	CAC TTGACAC	AGTGCGTCTG
8401	TTCTTAGCAG	ACCAGCCCGT	AGAAGGACCA	GAAAGTCTCT	TTCCGCTAAC	AAAAGAAAAA
8461	ACACCAGAAG	AGAATGCCCC	GAGTGTGCT	AGAGTCATTC	AAAAGCAGTC	GGAGGACGTG
8521	AAAATAGAGT	GGGAGAAGCT	GAGCATTCGC	TCTGTTGATT	GGCAGAAGCG	TTTGGATGAA
8581	GCCCTGGAAA	GACTTCTGGA	GTTGCAGGAT	GCTATGGATG	AACTGAACCT	CAAACTGCGC
8641	AAGGCTGAAT	CTATCCGAGA	CTCGTGGCAA	CCTGTGGGAG	ACCTGCTGAT	TGACTCTCTG
8701	CAGGACCATA	TAGAAAAAGT	CAAGGCCTTC	AGAGCAGAAA	TTGCTCCTTT	GAAAGAGAAT
8761	GTGACGCACA	TGAATGATCT	CGCCTGCCAA	TTCACACCGA	CGCAAATCCA	ACTGTCTCCA
8821	TATAACTTGA	ACCGGTTGGA	CGACCTGAAT	ACGAGGTGGA	AACAATTACA	GGTGGCTATT
8881	GAGGAGCATC	TGAAGCAGCT	GCACGAGGCC	CATAGGGACT	TCGGCCCCAC	ATCTCAGCAT
8941	TTCTTGTTCA	CTTCTGTCCA	AGGTGCTTGG	GAGCGAGCAA	TATCGCCTAA	CAAAGTGCCC
9001	TACTACATTA	ATGCAAAGAC	TGCAATCGGA	AGTAAACTAA	GCCACCAGAC	CCAGACTACC
9061	TGCTGGGATC	ACCCCAAGAT	GACTGAGCTT	TATCAGTCGT	TAGCTGACCT	GAATAATGTC
9121	AGGTTTTTCTG	CGTACAGAAC	TGCCATGAAG	CTGCGAAGAC	TGCAGAAAGC	TCTCTGCTTG
9181	GACTTGCTCT	CTCTGCCAAG	TGCATGTGAG	GCCTTTGACC	AGCGCAACCT	GAAGCAAAAC
9241	GATCAACTAA	TGGATATTAT	GGAGATAATT	AACTGCCTGA	CTACAATCTA	CGATCGACTG
9301	GAGCAGGAAC	ACAGCAACCT	TGTCAATGTG	CCTCTCTGTG	TGGATATGTG	TCTTAACTGG
9361	CTCCTGAATG	TATATGACAC	TGGCCGAACA	GGGAAGATTC	GTGTCCTTTC	CTTCAAAACT
9421	GGCATCATGT	CTCTGTGCAA	AGCACACCTG	GAGGACAAGT	ACCGCTATCT	GTTCAAGCAA
9481	GTGGCAAGTC	CCACTGGATT	CTGTGATCAG	CGCAGGCTGG	GCCTTTTGTT	GCACGATGCT
9541	ATTCAGTTTC	CACGCCAACT	AGGAGAGGTG	GCATCTTTCG	GAGGAAGCAA	CATTGAACCT

9601	AGTGTTGCGAA	GTTGCTTCCA	GTTTGCAAAC	AACAAGCCAG	ATATCGAAGC	AGCCTTGTTT
9661	CTGGACTGGA	TGAGGTTGGA	ACCCCAGTCT	ATGGTCTGGA	TGCCAGTGCT	GCACAGAGTG
9721	GCAGCTGCTG	AGACTGCCAA	ACACCAGGCC	AAGTGCAATA	TCTGCAAGGA	ATGCCCTATC
9781	ATTGGATTCA	GGTACCGAAG	TTTAAAGCAC	TTTAACTATG	ATGTCTGCCA	AAGTTGCTTC
9841	TTTTCTGGTC	GAACAGCGAA	AGGACATAAA	ATGCATTATC	CAATGGTGGA	ATACTGCACA
9901	CCGACAACTT	CAGGAGAAGA	TGTCCGAGAT	TTTGCCAAGG	TATTAAAGAA	CAAGTTCAGA
9961	ACCAAGAGGT	ACTTTGCAAA	ACATCCTCGA	ATGGGCTATC	TCCCCGTTCA	GACGGTCCTG
10021	GAGGGAGACA	ACTTAGAAAC	ACCAGCCTCA	TCCCCTCAGT	TATCACATGA	TGATACTCAT
10081	TCACGTATTG	AACATTATGC	TAGCAGGCTA	GCAGAAATGG	AAAACAGAAA	TGGATCTTAC
10141	TTGAATGACA	GCATCTCGCC	CAATGAAAGC	ATAGATGATG	AACATTTGCT	GATCCAGCAC
10201	TACTGCCAGA	GTCTGAACCA	GGAGTCACCA	TTGAGCCAGC	CTCGCAGCCC	AGCTCAAATC
10261	CTCATTTTAC	TGGAAAGCGA	AGAGAGAGGG	GAATTGGAAA	GAGTGCTATC	TGATCTGGAA
10321	GAAGAAAACA	GGAACCTTCA	GTCTGAGTAC	GAGAGGCTGA	AGCAGCAGCA	TGATCACAAG
10381	GGCCTGTCCC	CCTTGCCATC	GCCTCCTGAG	ATGATTCCCTG	TTTCACCGCA	AAGTCCCAGG
10441	GACGCTGAGC	TCATCGCTGA	AGCTAAGCTG	CTTCGCCAGC	ACAAAGGTCG	CCTTGAGGCC
10501	CGGATGCAGA	TCCTCGAGGA	TCATAACAAG	CAGTTGGAAT	CCCAGCTGCA	CCGGCTCAGA
10561	CAGCTCCTGG	AGCAGCCTCA	GACGGAGGTC	AGGGTGAATG	GTACTACCGT	CTCCTCCCCT
10621	TCCAGCTCCT	CACAACGATC	GGATAGCAGT	CAACCAGTGC	TGCATCACGG	AGTTGGCAGT
10681	CAGACTTCGG	AAATGCTGGG	TGAGGATGAA	TTGATGAATC	CTCCCCAGGG	GTCAAGCACT
10741	GGACTIONAGG	ATGTGATGGA	ACAGCTTAAC	AATTCATTTT	CAAGCTCTCG	AAGCAACAAT
10801	GTAGGCAACC	TCTTCCATGT	GGTGGACGAT	TTAGGCAGAG	CGATGGAAAC	TTTAGTGACA
10861	GTCATGACTG	ATGACGAGGC	AGCAGAGTAA			

//