

Gene: ENSG00000109501 - Sequence: ENST00000226760  
Transcript: ENST00000226760 - Protein: ENSP00000226760  
Date : March 2, 2015

1<sup>st</sup> line: Base numbering. Full stops for intronic +/- 5, 10, 15...  
2<sup>nd</sup> line: Base sequence. lower case Introns, upper case Exons  
3<sup>rd</sup> line: Amino acid sequence. Printed on FIRST base of codon  
4<sup>th</sup> line: Amino acid numbering. Numbered on 1<sup>st</sup> and increments of 10

Exon 1 | Start: 502 | End: 666 | Length: 164

```

. . . . .
gaaggcgagcggcctcgctaagcaactggacgttccgcggggcgggcgggggcggggcccg

. . . . .
ggggcccgagtcgcgtcggaaactttcgctgtgggcgagccggacccgccttctggcccct

. . . . .
cggggcccaaccacgcaggggggagtgccggggccgcagctaggcgaggcgccacgtgattgg

. . . . .
cggagatgtggagtgattggcggctacaccggccactcagcaggccgagctggcgccgca

. . . . .
tccggggggccgcgtctggagtgaggaggccgaaggccccgccccctgccccgccccct

|-169      |-159      |-149      |-139      |-129      |-119
GTGCAGAAAGCCGCGCTAGCCGGCTCTTCAGCAGCGAGTGCAGATTGCTCCCCCGCGGCC

|-109      |-99       |-89       |-79       |-69       |-59
GCAGATCTCCCGTTTGCGCCGCGTTCAGCTGCTCCCGAACAACCTTTTCTGCCGGCCCAGA

|-49       |-39       |-29       |-19       |-9        .        .        .
GGCCCCAGGGCGTCGCAGCGCCGCGTGC GGCCCACTCACGGGCCGgtgagtacttcggcg

. . . . .
ctggggcagtgggcgcggtggctgtgggcagcgcgagggcgggcaagagccctgaggc

. . . . .
actgtcctcttcgggcctcagtttccccttccgagctgatgggtggctggccccaaagtc

. . . . .
ccgacaaggtccccgaagttggagggccgggggtcccgcccgctctgcaacgcgcaaggc

```

gacccctgttccgggcccgaacgggtcacccggggggcgcgccccggtccccgcgcgt

gtcgcctggagccccgccgcgggcgggacagcaggccccgagaggc

Exon 2 | Start: 8103 | End: 8339 | Length: 236

```
. . . . .
gaaggccagggttggtgtggctcatgcccagtgaggcagggactccaggcctcggcatgtg

. . . . .
caggtctgagattgtaagtgccatgccatctgtagagtcacgtgggtgagtgctcctcca

. . . . .
tggtttcctccctggaagcgggtgctggcccatggggactgtactgagtgtcagcgagatc

. . . . .
ctgtatggagtgctcggcagctccacctgcctccctctgcttttctgtctccagcagac

. . . . .
actaagtgccagagcgggctctgccggtgctggatgtgcctgaccttgacttttcttcca

      |1      |11      |21      |31      |41      |51
GCAGGATGGACTCCAACACTGCTCCGCTGGGCCCCCTCCTGCCACAGCCCCGCCAGCAC
      M D S N T A P L G P S C P Q P P P A P
      |1      |11
      |61      |71      |81      |91      |101      |111
CGCAGCCCCAGGCGCGTTCCCGACTCAATGCCACAGCCTCGTTGGAGCAGGAGAGGAGCG
      Q P Q A R S R L N A T A S L E Q E R S E
      |21      |31
      |121      |131      |141      |151      |161      |171
AAAGGCCCCGAGCACCCGACCCAGGCTGGCCCTGGCCCTGGTGTAGAGACGCAGCGG
      R P R A P G P Q A G P G P G V R D A A A
      |41      |51
      |181      |191      |201      |211      |221      |231
CCCCCGCTGAACCCAGGCCAGCATACCAGGAGCCGGGAAAGAGCAGACGGCACCgta
      P A E P Q A Q H T R S R E R A D G T G
      |61      |71

. . . . .
agggagcaggctgggaagcccaggctggggatgttcagggatagctgggtgggaacgggg

. . . . .
ttcagccaccctggaggggtcccccgcaggctcctctgcagttcagcattgtgcagctc
```

. . . . .  
ccatgctgtgcacaggcgtccatccagtggggctacccacctcctcagagccttgacct  
. . . . .  
gtcacctttgtggcaccactcgaggtggtgctggtgccccactcctctgcagtccttc  
. . . . .  
tcttcccagggcctctgcagcacttcacagcttccatttgcaacagcgtccaaacat

Exon 3 | Start: 17745 | End: 17827 | Length: 82

```
. . . . .
cttggctctgtgctcgctccttcccatgcctcccaccccgagcgctttgagttgctgtt

. . . . .
gaaccacctcaccgcgatagagtttgctcatcttgctctggcgctgcttgtgaccggaag

. . . . .
gcaaacagtggctttctgggcatcttccctgtctgtgtctgtgtctctctgtactcctgg

. . . . .
cctggatttgaaagtgacaagcagcagcagatctgaagaccctcatgccttgtcccctcc

. . . . .
atcctgacaagtgacaaagtctggctttgtgacatgtgtgtttgtttcttctgtgttaaa

      |241      |251      |261      |271      |281      |291
GGCCTACAAAGGGAGACATGGAAATCCCCTTTGAAGAAGTCCTGGAGAGGGCCAAGGCCG
  P  T  K  G  D  M  E  I  P  F  E  E  V  L  E  R  A  K  A  G
      |81                                |91

      |301      |311      . . . . .
GGGACCCCAAGGCACAGACTGAGgtgaggactgcggtgccggcagggacttcgggacgcg
  D  P  K  A  Q  T  E
      |101

. . . . .
gcccccggcacaacaggcctggccacgagctccacagcccacagagaagtgtcggcgcct

. . . . .
gagatcggggtcaggagccagcgtggtgcaccctacccacttgagccccatgttggtag

. . . . .
ggtgcccattgttactgtgccagttttcctcctggcactcctctggggagcagcgtcat

. . . . .
cccccttttgtccaactcacacctcatcttgggcatcacctcctccaggatgacctcctg

. . . . .
gcttcctgcagctgcctgctcag
```

Exon 4 | Start: 19639 | End: 19783 | Length: 144

```
. . . . .
tgagccttggcaggcaggagcaactcaaggaagagaacctgtaccagtaccagtcggagc

. . . . .
ccgtgtctccctcgccgtgtggatggggtggccacaccttcctcaccgtgttttgaggag

. . . . .
cgagtggccggaggctcagtagggcctagcctagtggacatgcctggtgtgacccattt

. . . . .
ctgccccttccttcctggcctgggtgacaaaggggaagtgggtgaaaggaggtgggctggc

. . . . .
agggagcatggggtgggagagggtcggagaatctggaggctgactggtgtctggcttgca

      |321      |331      |341      |351      |361      |371
GTGGGGAAGCACTACCTGCAGTTGGCCGGCGACACGGATGAAGAACTCAACAGCTGCACC
V  G  K  H  Y  L  Q  L  A  G  D  T  D  E  E  L  N  S  C  T
                |111                                |121

      |381      |391      |401      |411      |421      |431
GCTGTGGACTGGCTGGTCCTCGCCGCGAAGCAGGGCCGTCGCGAGGCTGTGAAGCTGCTT
A  V  D  W  L  V  L  A  A  K  Q  G  R  R  E  A  V  K  L  L
                |131                                |141

      |441      |451      . . . . .
CGCCGGTGCTTGGCGGACAGAAGAGgtgggtctgtgtgaggcttagaacagcctctggag
R  R  C  L  A  D  R  R  G
                |151

. . . . .
ggttgagcagcttgtaatgctgcttgctaactgaacaactaaaatcttacaaacctaac

. . . . .
gctggtgatgctgttgggaaatttcagtttctgttttgctggtggccttctcattttaga

. . . . .
cactgtttctggacttaacatgggatatttaacagaccaagccattttcattctctttgg

. . . . .
cttggtgttggtatctccaagtggtaactatacatcctgcttcctgctgggttctgatcc

. . . . .
aaactgagacatcgatcctgggttc
```

Exon 5 | Start: 21849 | End: 22019 | Length: 170

```
. . . . .
gcctccgggtctgcacctctgagagaggggaggaacaggaagaggctgccttcttcccgc

. . . . .
cagcatgtagggggcacttgggggcgtctggtgggagaccagtctggcctcccagctgga

. . . . .
gagtgggcgtggcgcatgtcctcttgagtcatgtccatgcatccttccctggtaacc

. . . . .
aagtccctgacaccttctatgagtctcgctcgaaagccttccaggcagagttggcagggtc

. . . . .
agagtggcaccgaaagcctaggcagggcacacaaggcctttgaccacatcctatccctca

|461      |471      |481      |491      |501      |511
GCATCACGTCCGAGAACGAACGGGAGGTGAGGCAGCTCTCCTCCGAGACCGACCTGGAGA
  I  T  S  E  N  E  R  E  V  R  Q  L  S  S  E  T  D  L  E  R
                        |161                        |171

|521      |531      |541      |551      |561      |571
GGGCCGTGCGCAAGGCAGCCCTGGTCATGTACTGGAAGCTCAACCCCAAGAAGAAGAAGC
  A  V  R  K  A  A  L  V  M  Y  W  K  L  N  P  K  K  K  K  Q
                        |181                        |191

|581      |591      |601      |611      |621      |631 .
AGGTGGCCGTGGCGGAGCTGCTGGAGAATGTCGCCAGGTCAACGAGCACGgtgcgagga
  V  A  V  A  E  L  L  E  N  V  G  Q  V  N  E  H  D
                        |201                        |211

. . . . .
ttcaccctgggcaccagccttccctgggcgccagccttcccacaggagccaggaccttcc

. . . . .
cataggggctgggaccttccctcaggggctgggtcttcccacaggagccgggaccttccc

. . . . .
tgtgaggacagggcccttccctgtggggaccaggggaccagaaccttccctgtagagaccg

. . . . .
tgccctagtgggtgaggtgtgtgggtggcattttgacagcatctgccctgggtcaagtgtc

. . . . .
cactcattgaataaaccagaggggtattctgcccagtgctctgtgaccacgt
```

Exon 6 | Start: 22569 | End: 22649 | Length: 80

```
. . . . .
tcactcattgaataaaccagaggggtattctgcccagtgctctgtgaccacgtctaccaat

. . . . .
gggacggactgtgtccatcaccaagtgggagcacgctacgtggtgctgagtccaccccag

. . . . .
ctactggaggtagacagaggtgtggcccctgctctgcctgccctgggggccctatgatcccc

. . . . .
agaacgtaggatgccctggaaactggcgtgccctaggaacagtgcgccagtttctggtgg

. . . . .
gctgcagggcacgaggagatagtcaacttgctgactgttaatccaccctgtcccctgca

      |641      |651      |661      |671      |681      |691
ATGGAGGGGCGCAGCCAGGCCCGTGCCCAAGTCCCTGCAGAAGCAGAGGCGCATGCTGG
  G  G  A  Q  P  G  P  V  P  K  S  L  Q  K  Q  R  R  M  L  E
                        |221                                |231

      |701      |711      . . . . .
AGCGCCTGGTCAGCAGCGAGTgtgagtgagccccctgccccgtctcacccatgcctccca
  R  L  V  S  S  E  S

. . . . .
gcctgcacctgcagggcgacctctccttcctgtgcgactccatcctggcctgccctatct

. . . . .
caccctgcctcccagcctgcgccctgcagggcgacctctccttcctgtgcgaccccatcc

. . . . .
tggccctgctaggatctcaggcgggccgtttggggctcagtgttctggacgctgggagta

. . . . .
gaccctgccacctggagcgcacgcactggaggggaaggcagaccaggacagaaaccatg

. . . . .
atgtgccagtcacctcttggac
```



Exon 7 | Start: 25693 | End: 25841 | Length: 148

. . . . .  
tgggtgagtggccccaggcataaggagctaggcagagagggacacttgggggtgggggaca

. . . . .  
gcacacccagggggccggggccaggagtggaggctggcacttggcaaacctgcccccttc

. . . . .  
ctcctcaccagcctggtcctcaaccctcaggccgcccaggggaagggtttcctccacctg

. . . . .  
aaccactcagctcctttcttagcttggccccacgccaccgtccccagcccattgctctg

. . . . .  
tgtgaggggtggcagtggggctgcagtgtggggcgcccatgctgttttctctcatgcttca

721	731	741	751	761	771
CCAAGAACTACATCGCGCTGGATGACTTTGTGGAGATCACTAAGAAGTACGCCAAGGGCG					
K N Y I A L D D F V E I T K K Y A K G V					
241			251		

781	791	801	811	821	831
TCATCCCCAGCAGCCTGTTCCTGCAGGACGACGAAGATGATGACGAGCTGGCGGGGAAGA					
I P S S L F L Q D D E D D D E L A G K S					
261			271		

841	851	861	.	.	.	.	.	.	.
GCCCTGAGGACCTGCCACTGCGTCTGAAGgtgagtgaccaagaccccggtcaggccggag									
P E D L P L R L K									
281									

. . . . .  
cctgcctcccaaggactcgcgcacctcaggcagggcaccttccaggaagctgcaggtggg

. . . . .  
gaggttcgcgctaacaagagtggtcttacagccgtgccgctggtacctttgggtcatca

. . . . .  
tctatcgtcataaggatgtgtcctcgggagagaggcctttcttttctgcgccgtcagggt

. . . . .  
cagaaaccaggggcggtgttgggcaggagtgctaggatggcaagcaagggggcccctgggt

. . . . .  
ctttctgtgcagtgtagggggcagtggg

Exon 8 | Start: 31309 | End: 33917 | Length: 2608

. . . . .  
ggaggtcttgcaaggagagaagcacacatgcatctagtcacgctggtagaagggtggggag  
. . . . .  
ccaggcacggggcagaggggggctccaggcccagaagaggagggtcacagggaaccgcg  
. . . . .  
agcatggggagggccacctggagaaggggggagggaggaccactaggatggggctggtga  
. . . . .  
tgggaaaacgcaagggtgcgggttccttttgccagaggcagggtggtcagagggaggcg  
. . . . .  
tgagatgggagcagtgggggtcctgtcccagcctcggtccacgtaccatctttcccca

          |871          |881          |891          |901          |911          |921  
GTGGTCAAGTACCCCCTGCACGCCATCATGGAGATCAAGGAGTACCTGATTGACATGGCC  
V V K Y P L H A I M E I K E Y L I D M A  
          |291                          |301

          |931          |941          |951          |961          |971          |981  
TCCAGGGCAGGCATGCACTGGCTGTCCACCATCATCCCCACGCACCACATCAACGCGCTC  
S R A G M H W L S T I I P T H H I N A L  
          |311                          |321

          |991          |1001          |1011          |1021          |1031          |1041  
ATCTTCTTCTTCATCGTCAGCAACCTCACCATCGACTTCTTCGCCTTCTTCATCCCGCTG  
I F F F I V S N L T I D F F A F F I P L  
          |331                          |341

          |1051          |1061          |1071          |1081          |1091          |1101  
GTCATCTTCTACCTGTCCTTCATCTCCATGGTGATCTGCACCCTCAAGGTGTTCCAGGAC  
V I F Y L S F I S M V I C T L K V F Q D  
          |351                          |361

          |1111          |1121          |1131          |1141          |1151          |1161  
AGCAAGGCCTGGGAGAACTTCCGCACCCTCACCACCTGCTGCTGCGCTTCGAGCCCAAC  
S K A W E N F R T L T D L L L R F E P N  
          |371                          |381

1171	1181	1191	1201	1211	1221
CTGGATGTGGAGCAGGCCGAGGTCAACTTCGGCTGGAACCACCTGGAGCCCTATGCCCAT					
L	D	V	E	Q	A
E	V	N	F	G	W
N	H	L	E	P	Y
A	H				
391		401			
1231	1241	1251	1261	1271	1281
TTCCTGCTCTCTGTCTTCTTCGTCATCTTCTCCTTCCCCATCGCCAGCAAGGACTGCATC					
F	L	L	S	V	F
F	V	I	F	S	F
P	I	A	S	K	D
C	I				
411		421			
1291	1301	1311	1321	1331	1341
CCCTGCTCGGAGCTGGCTGTCATCACCGGCTTCTTTACCGTGACCAGCTACCTGAGCCTG					
P	C	S	E	L	A
V	I	T	G	F	F
T	V	T	S	Y	L
S	L				
431		441			
1351	1361	1371	1381	1391	1401
AGCACCCATGCAGAGCCCTACACGCGCAGGGCCCTGGCCACCGAGGTCACCGCCGGCCTG					
S	T	H	A	E	P
Y	T	R	R	A	L
A	T	E	V	T	A
G	L				
451		461			
1411	1421	1431	1441	1451	1461
CTATCGCTGCTGCCCTCCATGCCCTTGAATTGGCCCTACCTGAAGGTCCTTGGCCAGACC					
L	S	L	L	P	S
M	P	L	N	W	P
Y	L	K	V	L	G
Q	T				
471		481			
1471	1481	1491	1501	1511	1521
TTCATCACCGTGCCCTGTCGCCACCTGGTCGTCCTCAACGTCAGCGTCCCGTGCCTGCTC					
F	I	T	V	P	V
G	H	L	V	V	L
N	V	S	V	P	C
L	L				
491		501			
1531	1541	1551	1561	1571	1581
TATGTCTACCTGCTCTATCTTCTTCTCCGCATGGCACAGCTGAGGAATTTCAAGGGCACC					
Y	V	Y	L	L	Y
L	F	F	R	M	A
Q	L	R	N	F	K
G	T				
511		521			
1591	1601	1611	1621	1631	1641
TACTGCTACCTTGTGCCCTACCTGGTGTGCTTCATGTGGTGTGAGCTCTCCGTGGTCATC					
Y	C	Y	L	V	P
Y	L	V	C	F	M
W	C	E	L	S	V
V	I				
531		541			
1651	1661	1671	1681	1691	1701
CTGCTGGAGTCCACCGGCCCTGGGGCTGCTCCGCGCCTCCATCGGCTACTTCCTCTTCCTC					
L	L	E	S	T	G
L	G	L	L	R	A
S	I	G	Y	F	L
F	L				
551		561			

	1711	1721	1731	1741	1751	1761
TTTGCCCTCCCCATCCTGGTGGCCGGCCTGGCCCTGGTGGGCGTGGCTGCAGTTCGCCCCGG						
F A L P I L V A G L A L V G V L Q F A R						
	571			581		
	1771	1781	1791	1801	1811	1821
TGGTTCACGTCTCTGGAGCTCACCAAGATCGCAGTCACCGTGGCGGTCTGTAGTGTGCCC						
W F T S L E L T K I A V T V A V C S V P						
	591			601		
	1831	1841	1851	1861	1871	1881
CTGCTGTTGCGCTGGTGGACCAAGGCCAGCTTCTCTGTGGTGGGGATGGTGAAGTCCCTG						
L L L R W W T K A S F S V V G M V K S L						
	611			621		
	1891	1901	1911	1921	1931	1941
ACGCGGAGCTCCATGGTCAAGCTCATCCTGGTGTGGCTCACGGCCATCGTGCTGTTCTGC						
T R S S M V K L I L V W L T A I V L F C						
	631			641		
	1951	1961	1971	1981	1991	2001
TGGTTCTATGTGTACCGCTCAGAGGGCATGAAGGTCTACAACCTCCACACTGACCTGGCAG						
W F Y V Y R S E G M K V Y N S T L T W Q						
	651			661		
	2011	2021	2031	2041	2051	2061
CAGTATGGTGCCTGTGCGGGCCACGCGCCTGGAAGGAGACCAACATGGCGCGCACCCAG						
Q Y G A L C G P R A W K E T N M A R T Q						
	671			681		
	2071	2081	2091	2101	2111	2121
ATCCTCTGCAGCCACCTGGAGGGCCACAGGTCACGTGGACCGCCGCTTCAAGTACGTC						
I L C S H L E G H R V T W T G R F K Y V						
	691			701		
	2131	2141	2151	2161	2171	2181
CGCGTGACTGACATCGACAACAGCGCCGAGTCTGCCATCAACATGCTCCCGTTCTTCATC						
R V T D I D N S A E S A I N M L P F F I						
	711			721		
	2191	2201	2211	2221	2231	2241
GGCGACTGGATGCGCTGCCTCTACGGCGAGGCCTACCCTGCCTGCAGCCCTGGCAACACC						
G D W M R C L Y G E A Y P A C S P G N T						
	731			741		



*111	*121	*131	*141	*151	*161
GACTGTGGCTGCAGAGACCTTGCGACCATGTGTAGATTGCGTGGACCCCGACAAAGGGAA					
*171	*181	*191	*201	*211	*221
GGCTGTGTGTAGCTCTGTCCACTCTGAATACCAAGTGTGTTGGGAATTGCATGCCATCT					
*231	*241	*251	*261	*271	*281
CCACCCTGAGCCTGACCTTTCTGAGTGACATGGGTGTGCCAGGCTAGACTAGGAGGTTCC					
*291	*301	*311	*321	*331	*341
GGTGTCTGAAAAGCACTTTACAGATGAGATTCCCTCTCCTCCCCACCTTCAAGCACCC					
*351	*361	*371	*381	*391	*401
TGTTCCCTCTTTCTTTCTTTGTGTTGGATTGTGTTAAAAACCAAATAAGCATCTGTGTA					
*411	*421	*431	*441	*451	*461
ACCTCCACAGTAGCATTTCTTATTTGTTGGTCACTGCTACACCTTAGCAGCTCTTCCCC					
*471	*481	*491	*501	*511	*521
TTTCCTGGGGGATGTGCACGGCAGCTTGAGCCTGTCACGTGGTCAAGGCCCGCCCCATC					
*531	*541	*551	*561	*571	*581
AGAGGCTGGGGGAGGCGGCACATTGGCAGTGTGTCACTGAGCTGGGCACCACAGGCTG					
*591	*601	*611	*621	*631	*641
CCTCATGACCCTCCTGTCCAGCAGGTAGTGGGTGAATGTGTGAAGGTCTTGCTGAATCC					
*651	*661	*671	*681	*691	*701
ATCAGGACTTGGGAAACAGAGAACCCTGTGGGGGCGGCTGTGGGGGAGGTCCCTGCCAGT					
*711	*721	*731	*741	*751	*761
GTTTAGAAGAGCCTGACTGTGTTCAGTGCCTTGAGCAGAAAGCCAGGGTCCTGAGTGGC					
*771	*781	*791	.	.	.
TGAAATAAAAAGCCTCTGGTGGAACCTGC	Agcgctttccttcctttctttaccgaaaagaa				
.	.	.	.	.	.
gtctttcttgtagcgtgcgtgagaatcagcagagcctgcactcctggtgaatgaaatgcaa					
.	.	.	.	.	.
gtgcaatttgagttataaaaagagcaaggttgatgtttcacagttgatggcttcctgccac					
.	.	.	.	.	.
agcgagaccctggcttcatctccagctggaggggcccctggggcatctgccgtaactgtg					

· · · · ·  
gggtggcctgggcatgggctgcctgtgcagagagacctgtgctgaagggtgaccatggagt  
· · · · ·  
gtcagcccagccatcttcagatcttactg

GBK Parser: Version: 1.1, Version Date: 11/02/2015  
Reader: Version: 1, Version Date: 11/02/2015  
Writer: Version: 1, Version Date: 11/02/2015  
Control: Version: 1, Version Date: 11/02/2015