

NBEAL1 Conceptually Annotated Transcript

1 gctcgtctgctggtgctgctgggggtgacacggggcttgccttgggaaggggtcgaggaag 60
61 cagttagacggctgctggggggctgcccggcgacacaaattttatttaattggcc 120
121 aactaccactgatgaagatatattggagtgactgctgaaattgcctttttgttttaacc 180
181 agaggacagtcatttgttcaactcttttgccttctttactgctatgagctttactga 240
241 acggctgaaaaaacttggaaaaataaaatggacatgctgtagcttgaacataatttttta 300
M A S R E R L F E 9
301 aggaaaaactaaagtccagagtgaagccagaatggcatccagagagagctcttttgaa 360
L W M L Y C T K K D P D Y L K L W L D T 29
361 ctttgatgctttattgtacaaaagaagatccagattactgagctgtggctggacact 420
F V S S Y E Q F L D V D F E K L P T R V 49
421 tttgtttctagctatgaacaatttttagacgttgacttggaaaagctgcctaccagggtta 480
D D M P P G I S L L P D N I L Q V L R I 69
481 gatgatagcctccaggaatatctctgcttctgataatattctgcaggttctgaggatc 540
Q L L Q C V Q K M A D G L E E Q Q Q A L 89
541 cagcttctacagtggttcagaaaatggcagatgggttagaggacaacacgaagccttg 600
(S) I L L V K F F I I L C R N L S N V E E 109
601 tcaattttgtcaggttcttcttatttcttggagaatctcaaatgtggaagaa 660
I G T C S Y I N Y V I T M T T L Y I Q Q 129
661 attggactgtcgtacattaattatgtcatcaccatgacaacactctatattcagcaa 720
L K S K K K E K E M A D Q T C I E E F V 149
721 ttaaaaagcaaaaaaagagaaggaaatggcagatcagacatgtattgaagaattttgtg 780
I H A L L A F C E S L Y D P Y R N W R H R 169
781 atccagcattggcattttgtgaagcttatatgcatatcggaattggagacataga 840
I S G R I L S T V E K S R Q K Y K P A (S) 189
841 atttcagagcaatccttagtactgtggaaaagagcagacagaaaataaacagcttct 900
L T V E F V P F F Y Q C F Q E S E H L K 209
901 ctcacagtgaattcgtccctttcttttatcaatgttttcaggaaggtgaacctctcaag 960
E S L K C C L L H L F G A I V A G G Q R 229
961 gaaagtcttaaatgttgcttattgcctctttggagccattgagccggtgggcaagg 1020
N A L Q A I S P A T M E V L M R V L A D 249
1021 aatgctttgcaagcaatttctccagccactatggaagttcttatgagcagattggcagat 1080
C D S W E D G D P E E V G R K A E L T L 269
1081 tgtgattcctggaggatggagatcctgaagaagtggttaggaagccagaactaactctg 1140
K C L T E V V H I L L S S N S D Q R Q V 289
1141 aagtgccttacagaagtggtacatatccttctcagtagcaactctgatcagcgtcaagtg 1200
E T S T I L E N Y F K L L N S D H S A L 309
1201 gaaaccagtaactattctggagaactattttaaattgctaattcagatcattcagcttta 1260
P N Q R R (S) R Q W E N R F I A L Q I K M 329
1261 cctaatacaagggatccagacagtgggaaaacgatttattgctctacagatcaaaaatg 1320
L N T I T A M L D C T D R P V L Q A I F 349
1321 ctgaaataccatcacagccatggttagattgtacagatagacctgttctcagccattttt 1380
L N S N C F E H L I R L L Q N C K V F Q 369
1381 cttaacagcaattgcttgaacatctcatacagctgctacagaactgcaagggtttcag 1440
G Q L D C L A I S T I Q A L T A V M N K 389
1441 ggacaattggattggttggccatatcaaccattcaggctttgaccgagtaatagaacaaa 1500
S P A A K E V F K E R I G Y T H M L E V 409
1501 tctcagctgctaaaggatatttaagaaagaattggttatcacacatagccttgaagta 1560
L K S L G Q P P L E L L K E L M N M A V 429
1561 ttaaaatccctgggtgacccaccactggaattacttaagaacttatgaatatggctgta 1620
E G D H T S V G I L G I S N V Q P L L L 449
1621 gagggtgaccacacttcagttgggattttggcattagtaatgtccaacctctctgtt 1680
L I Q W L P E L Q S H D L Q I F I S D W 469
1681 cttatocagtgctccagaaactacaatcccacatgacctgcaaatctctcctgattgg 1740
L K R I C C I A N R Q S R T T C V N A N M 489
1741 ctgaaagaatttgggtatttaatagacagagtcgaactacttgtgtcaatgcaaacatg 1800

Comment [ER1]: Ex1/2

Comment [ER2]: N-terminal acetylation

Comment [ER3]: Upstream in-frame stop codon

Comment [ER4]: G → C and R → T

Comment [ER5]: E2/ex3

Comment [ER6]: T → G, Leu → Trp

Comment [ER7]: Ex3/4

Comment [ER8]: G → A, Val → Ile

Comment [ER9]: C → T, Ser → Phe

Comment [ER10]: C → T, Gln → nothing

Comment [ER11]: Ex4/5

Comment [ER12]: Frame shift, goes to Gln (Q)

Comment [ER13]: Ex5/6

Comment [ER14]: Insertion of A, causes frameshift to Arg (R)

Comment [ER15]: Ex6/7

Comment [ER16]: Ex7/8

Comment [ER17]: Ex8/9

Comment [ER18]: A → G, Asn → Ser

Comment [ER19]: C → T, Arg → no recognized codon

Comment [ER20]: Ex 9/10

Comment [ER21]: Ex 11/12

Comment [ER22]: Ex 12/13

Comment [ER23]: Deletion of 13 bp, frameshift

Comment [ER24]: Ex 13/14

Comment [ER25]: A → G, Lys → Glu

G I R I I E T L D L H S S L H Q T C A E 509
 1801 gggattagaatcattgaaacccttgacttgattcttccctccatcaaaccttgctgag 1860
 N L I A I H G S L G S Q S V S S E E I R 529
 1861 aacttgattgcaatccatgggtccctggggagtcagtcagtgagccagagaatccgt 1920
 R L L R L L R V D E S E S V H P Y V T P 549
 1921 cgactactgagattgctgagagtgatgaatctgagctgttacccttatgtactccc 1980
 V T R A I L T M A R K L S L E S A L Q Y 569
 1981 gtgactcgagcaatcctgacaatggccgaaaactaagtctagagagtgccctccagtat 2040
 F N L S H S M A G I S V P P I Q K W P G 589
 2041 ttcaattgtcacatagatggcaggaattctgtgctcccatacagaaatggccaggg 2100
 S A F S F S A W F C L D Q D Q L T L G I 609
 2101 totgcttttctttcagtgctgtgtttgcttagaccaggatcagttgactcttgccatt 2160
 A N K G G K R K Q L Y S F F T G S G M G 629
 2161 gctaacaaaggagggaaaaggaaacaattgtacagtttttttacaggaatggcatgggt 2220
 F E A F I T H S G M L V V A V C T K R E 649
 2221 tttgaagcctttattaccattcaggtatgttgctgtgagtgccacaaaagagaa 2280
 Y A T V M L P D H S F C D S L W H N I T 669
 2281 tatgcaacggttatgcttctgaccacagtttctgtgattccctctgaccacataact 2340
 V V H M P G K R P F G Q S F V Y I Y D N 689
 2341 gttgtccacatgcttgaaaaggccttttgctcagagcttctctatctatgacaat 2400
 G Q Q K V S A P L R F P A M N E P F T S 709
 2401 ggacacagaaggtttctgcccctctcagatttctgctgcatatgacccctttacttcc 2460
 C C I A G S A G O R T T T P P P S Q I P D 729
 2461 tgttgcaattggttcagctgggcaagaaccaccactcctccaccatcccaatcccagat 2520
 P P F S S P C I T P H R T S F G G I L S S 749
 2521 ccactttctcttccattaccctcatcggacatcatttggggaattctgtcatca 2580
 A S W G G T I E K S K L I T K L I S A G 769
 2581 gcctcctggggaggaacaattgaaaaatcaaatgattaccaaatgatcagctgga 2640
 T Q D S E W G C P T S L E G Q L G S V I 789
 2641 acccaagacagtgaaatgggggtgtccacatctctggaggtcagctaggatctgtatc 2700
 I F Y E P L Q P P Q V K A L Y L A G P N 809
 2701 atctttatgaaccactacaacctcctcagtggaagcattatattagcaggtccaat 2760
 C L S P W K C Q E S D M A D L P G N I L 829
 2761 tgtttaagcccttggaagtgtcaagagctgacatggccagcctgctgtatcattcctt 2820
 L Y Y T A K A C K N S I C L D L S T N C 849
 2821 cttactacacagcaaaaggcctgcaaaaattcaatctgcttatttactactaattgt 2880
 L H G R L T G N K V V N W D I K D I I N 869
 2881 ttgcatggaagattaacaggaaacaaagttagtgaactgggacattaagatatcataaac 2940
 C I G G L N V L F P L L E Q I S H F S E 889
 2941 tgcataagtggttaaatgtactcttctttatttggaacaaatcagccactttagttaa 3000
 G Q I P E E K N E S T V P E S V T P V E 909
 3001 ggacagattcctgaaagaaagaatgaaagcacagttcctgaaatcagtaaacacctgtttaa 3060
 G D W L V W T S T K A S E S R L E R N L 929
 3061 ggagattgctcgtatggacttccacaaggcctcagagtcagactagagagaaacctaa 3120
 V A T F I L I V K H F I Q R H P I N Q G 949
 3121 gttgcaacatttatcttaattgtgaaacattttattcagagacatcctatcaaccagggc 3180
 N L I H S H G V A T L G A L L Q K V P S 969
 3181 aatcttattcactcccaggttgcaactcttggtgctttactcagaaagtgccaagc 3240
 T L M D V N V L M A V Q L L I E Q V S L 989
 3241 accttgatggatgtaattgtgtgatggcagttcagttactaattgaaacagatcatta 3300
 E K N M Q L L Q Q M Y Q Y L L F D F R I 1009
 3301 gagaaaaatcagcctcctgcaacaaatgatcaatatttactctttgactttcgtatt 3360
 W N R G D F P F R I G H I Q Y L S T I I 1029
 3361 tggaaacctggagatttcccttccgaatcggcacatacagtatctttcaaccatcatt 3420
 K D S R R V F R K K Y G V Q F L L D T L 1049
 3421 aaagacagcaggagatttccgaaagaagtatggtgtcagtttctcctagatcactt 3480
 R I Y Y G N G C K Y N E L S L D D I R T 1069
 3481 aggatttatttgggaattggttgaatataatgaactatctctagatgatcgaaca 3540

Comment [ER26]: G→T, Asp→Tyr

Comment [ER27]: Insertion of second T, frameshift to Leucine

Comment [ER28]: A→G, Gln→Arg

Comment [ER29]: Ex14/15

Comment [ER30]: Ex15/16

Comment [ER31]: Ex16/17

Comment [ER32]: Insertion of additional TA, frameshift to Serine

Comment [ER33]: Deletion, causes frameshift to Pro (P)

Comment [ER34]: Ex17/18

Comment [ER35]: A → C, Asn→Thr

Comment [ER36]: Ex18/19

Comment [ER37]: Ex19/20

Comment [ER38]: Ex20/21

Comment [ER39]: Ex 21/22

Comment [ER40]: Ex 22/23

Comment [ER41]: Ex 23/24

I R T **S** L Y G L I K Y F L C K G G **S** H E 1089
3541 ataaggacttctttgtatggactaattaataatcttctgtgcaaaggatcctcatgaa 3600
E I Q S I **M** G Y I A A **T** N E E E Q L F G 1109
3601 gagatacaaaagtattatggggtacatagctgctactaatgaagagaaca**ctctttt**gga 3660
I L D V L F **S** L L R **T** **S** P **T** R G Q L F L 1129
3661 attttggagctgctctcagctcctacgtaccagccaaccagaggtcagcttttctta 3720
L L F E P G N A D I L Y A L L L N Q K Y 1149
3721 ctgctttttgaaccaggaaatgctgacatactgtacgcattgctcttaaatcagaagtac 3780
S D R L R E I I F K I **M** E Q **M** L K C T N 1169
3781 tctgacagactaagagaaatcatttttaag**at**tatggaaacaaatggtgaaatgcacgaac 3840
V **Y** E R S K Q H I R L R E V G **Y** **S** G L G 1189
3841 gtttatgagcgtagtaacaacatattcgactcagagaagttggctactcgggactggga 3900
L L L N E A A L V N T **S** L I K N L T H Q I 1209
3901 ctcttcttaataagcacttgttaataacttctcttattaaaacctcaccatcaaatc 3960
I N T D P V I N F K D L L **S** V V Y I **S** H 1229
3961 ataaatacag**at**cctcttatttaatttcaaatgactactatctgtgtatataatctcac 4020
R A H I N V R V A I C R K V L Q I L Q F 1249
4021 agagcacatataaatgttagagtgccatctgcagaa**gg**ttttgcaattttgcagttc 4080
Q P D A A A H Q I **S** Q Q V G W Q D **T** L V R 1269
4081 cagccagatcagcacaatcacagcaagtggtgggcaaacaccttagttagg 4140
L F L K A K F E N G N T L H K H **S** R A V 1289
4141 cttttttaaaagcaaaatgaaacggaataactcttcataagcagatagagctgtt 4200
L **M** K D N D K N **M** (**S** | **T**) E D **T** K K N **S** D E 1309
4201 ttaatgaaagacaatgataaaaatgtcaactgaagataccaagaagaactctgatgaa 4260
K **T** D E E K I T S F A S A N V S S D Q W 1329
4261 aaaacagatgaggaaaatcactcttttgctcagctaatgtgtcttcggatcagtg 4320
S L E D R H **S** L D **S** N **T** P L F P E D **S** **S** 1349
4321 agtttggagatagacactcttttagactcaaacacaccattatttccagaagatagctc 4380
V G E L **S** F K **S** E N Q E E F W H **S** N P S 1369
4381 vtgggaaatgtcttcaaatcagagaatcaagaggaattctggcatagtaacccttca 4440
H L S L D L **S** G I D **S** C E **M** **S** D **S** G **S** Q 1389
4441 catttgagtttagacctcagtggaattgactcatgtgaaatgagtgatagtggaagtcaa 4500
V P D **S** L P **S** T P (**S**) P V E **S** T K **S** F **S** V 1409
4501 gtgccagacagctgctagcacaccatccccagtagagctactaaatcgtttctgtg 4560
H **S** D R E **S** (**S**) I T N D **M** G F (**S**) D D F **S** L 1429
4561 cacttgacagagaaagcagcacaatgatatggccttttagtgatgacttctcttta 4620
L E **S** Q E R C E E E L L Q L L T H I L N 1449
4621 ctgaaagccaagag**gat**gtgaggaggagcttcttcaattactgacacatattttgaa 4680
Y V **M** C K G L E K **S** D D D **T** W I E R G Q 1469
4681 tatgtaattgtgaaggactagaaaagtctgatgatgatacttgatgaaacaggacaa 4740
V F **S** A L **S** K P G I **S** S E L L R P **S** D E 1489
4741 gtgttttcagcactaagtaaacaggaaatccagtgaaactacttccacatcagatgaa 4800
I K L **T** L L Q K **M** L E W A I **S** E N R E A 1509
4801 ataaaactaa**ct**ttgtacaaaagatgttagaatgggcaatctcagaaaacagagaagca 4860
K T N P V T A E N A F R L V L I I Q D F 1529
4861 aaaactaatccagtaactgtgaaaacgccttccgactagtgtgatcatacaggacttt 4920
L Q S E G L V N S N **M** W **T** E K L L E D **M** 1549
4921 cttcagtcagaggactagtttaattcaacatgtggaccgagaag**ct**tttagagatag 4980
M L L F D C L **S** V C Y S E **S** P V W V K L 1569
4981 atgctgctctttgactgtctgtcagctctgctattotgaaagtccagtatgggtaaaactc 5040
S Q I Q I Q L L L G F I G R G N L Q V C 1589
5041 tctcaaatcagatccagttgctcttagattcattggaaggggtaatttgcag**gt**ttgt 5100
A **M** A **S** A K L N **T** L L Q T K V I E N Q D 1609
5101 gcaatggcatcagctaagctaaatcccttcttcagaccaaagtgtgaaatcaggat 5160
E A C Y I L G K L E H V L **S** Q **S** I K E Q 1629
5161 gaagcatgttacattttaggaagctggaacatgttctaagtcaatcaatcaaggaacag 5220
T E I Y **S** F L I P L V R **T** L V **S** K I Y E 1649
5221 actgaaatcactcatttctgatcccttcttctgacccctgttccaaaatttatgag 5280

Comment [ER42]: A→G, Tyr→Cys
Comment [ER43]: Ex 24/25

Comment [ER44]: Ex 25/26

Comment [ER45]: Ex 26/27

Comment [ER46]: Ex 27/28

Comment [ER47]: Deletion, frameshift to Arg (R)

Comment [ER48]: Ex 28/29

Comment [ER49]: C→T, Arg→nothing

Comment [ER50]: Ex 29/30

Comment [ER51]: Ex 30/31

Comment [ER52]: Ex 31/32

L L F M N L H L P S L P F T N G S S S F 1669
 5281 cttctcttcacatgaacctgcacctacctctttaccttttaccatggtagctcctcattt 5340
 F E D F Q E Y C N S N E W Q V Y I E K Y 1689
 5341 tttagaattttcaagaattgttaattcaaatgaatggcaagtttaccattgaaaaata 5400
 I V P Y M K Q Y E A H T F Y D G H E N M 1709
 5401 attgtaccttatatgaagcagatgaagctcacaatctttacgatggcatgagaacatg 5460
 A L Y W K D C Y E A L M V N M H K R D R 1729
 5461 gcactttattggaagattgtttatgaaagcctttaaaggtaaatatgcataaacgagaccgg 5520
 E G G E S K L K F Q E L F V E P F N R K 1749
 5521 gaaggaggggaaagcaagctcaaatctcagagctgtttgtggagccatttaacgaaaa 5580
 A R Q E N L R Y N N M L K Q L S S Q Q L 1769
 5581 gcacgcaagagAACctgaggtataataatgcttaaaacaacttagcagtcaacagtta 5640
 A T L R R W K A I Q L Y L T C E R G P W 1789
 5641 gccactcttagacgctggaagcaatacagctctctacatgtgaaagggaccttgg 5700
 A K R K Q N P I H W K L A N V E N Y S R 1809
 5701 gctaaaaaggaaacagaatccaatcactggaagctagcaatgtagagaattattcccgc 5760
 M R L K L V P N Y N F K T H E E A S A L 1829
 5761 atgagacttaagctggtaccgaattataattcaaaaccatgaggaagctagtgccttg 5820
 R D N L G G I Q H S Q P S S D T L L L E V 1849
 5821 agagataatctggatccaacacagccttccagtgatcacattgcttttggaaagta 5880
 V K Q V K V S D M V E D K L D L P E E D 1869
 5881 gtgaaacaagtaaaagtttagtgatggtgaggataaattagaccttccctgaagaggat 5940
 I T A R V N V D E K E E Q D Q K E K L V 1889
 5941 ataacagctagagtaaatgtttgatgagaagaagaacaggatcaaaaagaaaaattggta 6000
 L M E D C E L I T I I D V I P G R L E I 1909
 6001 ttgatggagactgtgaaactattacaataattgtaattcctggcagattagaaatc 6060
 T T Q H I Y F Y D G S I E K E D G V G F 1929
 6061 actactcaacacacttactctatgatggcagcattgaaaaagaagatggagtaggcttt 6120
 D F K W P H S Q I R E I H L R R Y N L R 1949
 6121 gatttcaagtgccctcattctcaaatcagagagattcatctccggcgttacaatttaaga 6180
 R S A L E I F H V D Q S N Y F L N F K K 1969
 6181 agatcagcccttgagatttttcatggtgaccaatccaactactttctcaatttcaaaaaa 6240
 E V R N K I Y S R L L S L H S P N S Y Y 1989
 6241 gaggttagaacaaaatataatagccgactgtgtcacttccatccccaaatagttattat 6300
 G S R S P Q E L F K A S G L T Q K W V N 2009
 6301 ggaagcagatcaccacagagttattcaaaagcatcaggattgacacagaaatgggtaaac 6360
 R E I S N F D Y L I Q I N T M A G R T Y 2029
 6361 agagagatatcaaattttgactacctcattcaataataatcaatggcaggacgaacctat 6420
 N D L A Q Y P V P W I L Q D Y T S E E 2049
 6421 aatgaccttgacagatctctgttttccctggattttacaagattatacttcggaagag 6480
 L D L N N P A V F R D L S K P I G V V N 2069
 6481 ttggaccttaaacctgctgtatttcgagatctttccaaccaattggggtagttaat 6540
 E K N A K A M R E K Y E N F E D P M G T 2089
 6541 gaaaaaacgccaagctatgagagaaaatatgaaattttgaggatcctatgggaact 6600
 I D K F H Y G T H Y S N S A G V M H Y L 2109
 6601 attgataagtttactatggtactcactattcaaatctcgccgggcatgactatctc 6660
 I R V E P F T T L H I Q L Q S G R F D C 2129
 6661 attcgtgtagaaccttaccacctccacatccaactcagagtggaaggttgactgt 6720
 A D R Q F H S I P A T W Q A L M D N P Y 2149
 6721 gcagatcgacagttcattctattctcctgctacctggcaagctcttatggataatccatat 6780
 D V K E L I P E F F Y F P E F L E N Q N 2169
 6781 gatgttaagaacttattctgaattcttatttccagagtttttggaaaatcaaat 6840
 Q F N L G R L Q I S K E L V N D V I L P 2189
 6841 caatttaacttgggtcgtctacagatttccaagaattagtaaatgatgtcattctccg 6900
 K W A K S A E D F I Y K H R K A L E S E 2209
 6901 aaatgggctaaatcagctgaagatttcatctataaacataggaagctttggagtctgaa 6960
 Y V S A H L H E W I D L I F G Y K Q R G 2229
 6961 tatgtttcagctcatctcatgaatggatagatctgactttggctataaacagagggga 7020

Comment [ER53]: Ex 32/33

Comment [ER54]: Insertion of A, frameshift

Comment [ER55]: G→A, Arg → Gln

Comment [ER56]: Insertion of A, causes frameshift

Comment [ER57]: Ex 33/34

Comment [ER58]: Ex 34/35

Comment [ER59]: Ex 35/36

Comment [ER60]: Ex 36/37

Comment [ER61]: Ex 37/38

Comment [ER62]: Ex 38/39

Comment [ER63]: Ex 39/40

Comment [ER64]: Ex 40/41

Comment [ER65]: Ex 41/42

Comment [ER66]: Ex 42/43

Comment [ER67]: ILPK – vacuolar-targeting peptide motif

Comment [ER68]: Ex 43/44

Comment [ER69]: Ex 44/45

P A A V E A L N V F Y Y C S Y E G A V D 2249
 7021 ccagctgcagtagaggcaactcaacgttttctattattgtagttatgaa**gg**agctgtggat 7080
 L D A L T D E K E R K A L E G M I N N F 2269
 7081 ctggatgccttaacagatgagaaagaagaagccttagaaggatgattaataat 7140
 G Q T P C Q L L K E P H P P R L S A E E 2289
 7141 gggcaaacacctgtcaattataaa**gg**aaccacacctccaagattatcagcagaagaa 7200
 A V Q K P T K I D T S T L N L F Q H L P 2309
 7201 gcagtgcaagaagccaacaaaatagacacttcaacacctaaacctgtttcaacaccttct 7260
 E L K S F F I E G I S D G I P L L K A T 2329
 7261 gaactcaagtcatttttataga**gg**gattagtgatggtattccactattaaaggccacc 7320
 I P K N Q Y R S F M S Q G S P E L L I T 2349
 7321 atccccaaaatcagatcgttcttttattgtctcaaggcagccttagtact**gat**taaca 7380
 I S M N Y V I G T H G W L P Y D R N I (S) 2369
 7381 ataagcatgaattatgttttggaacctggatggttgcccttatgacagaacatttct 7440
 N Y F T F I K D Q T V T N P K T Q R S I 2389
 7441 aattactttacattcatcaaggatcaaacctgtga**ca**aatccaaaactcagcgcagatata 7500
 N G S F A P G L E I T S K L F V V S H D 2409
 7501 aatggttcttttctccgggctagagatcacttctaagctattttagtatcacatgat 7560
 A K L L T F S A G Y W D N S I Q V M S L T 2429
 7561 gcaaaagtgtctctcagtgctggataactggataatagcattcaagtgatgctactaca 7620
 K G K I I S H I I R H M D I V T C L A T 2449
 7621 aaaggcaaaattctcacacatctccggcatatgg**at**attgtgactgtcttagctaca 7680
 D Y C G I H L I S G S R D T T C M I W Q 2469
 7681 gattactgtggaatacatttgatttctggttccagagatactacatgatgatggcaa 7740
 I T Q Q Q G G V P V G L A S K P F Q I L Y 2489
 7741 ataacaca**gg**gagvtctcctggtgcttagcattcaaccctttcagattctttat 7800
 G H T N E V L S V G I S T E L D M A V S 2509
 7801 ggacacaccaagcaggtactgagtgctggcactcagcactgagctagacatggcagtgca 7860
 G S R D G T V I H T I Q K G Q Y M R T 2529
 7861 ggatcaag**gg**atggaaacggtgattatacataccattcagaaaaggtcagtcacataggact 7920
 L R P P C E S S L F L T I P N L A I S W 2549
 7921 ttaccaccacttgtgagattctctgttctcagcattctcaatttggctatatcttgg 7980
 E G H I V V Y S T E E K T T L K D K N 2569
 7981 gaaggacatattgtgtctactccagcactgaagaaagaccacctca**gg**ataagaat 8040
 A L H L F S I N G K Y L G S Q I L K E Q 2589
 8041 gcattacatctgtttctataaaatggcaagtatctagggtctcaaatcctgaaggaacaa 8100
 V S D I C I I G E H I V T G S I Q G F L 2609
 8101 gtatcagatataatgtataatcgggagaacacattgtcacaggcagcatacaaggattcctg 8160
 S I R D L H S L N L S I N P L A M R L P 2629
 8161 tctataagagatctccaca**gct**tgaatctcagcatcaaccattagccatgcgactgct 8220
 I H C V C V T K E Y S H I L V G L E D G 2649
 8221 atccattgtgttgtgtcaccaaagaatacagccatattctttaggtttagaagatggc 8280
 K L I V V G V G K P A E M R S G Q L S R 2669
 8281 aaattgattgtagtggtgtggcaagcctgctgag**at**gcgttcaggtcagctttctcga 8340
 K F W G S S K R L S Q I S A G E T E Y N 2689
 8341 aaatttggggatcgagcaagcggctcagccagatttcagctggagaactgaatataat 8400
 (T) Q D S K * 2694
 8401 actcaagattccaag**tg**attgttatttccattttctgttatgattactgaaacctgattt 8460
 at**tgctttt**gcactttaaaccacatctctcaact**ctctg**caatgttgaaggcttttatcc 8520
 ctgaaaat**cat**ttacagataaccacaatttctgtggtatataaaactaaattcttggctca 8580
 factaagatgatttgagaaaatacatttgatttgattttgtggccattctcaaggctc 8640
 attgtatccatttttaaaacaaaactaaaatgagaacattaggttcaattttcttat**tatt** 8700
 ccaaatgataaaaattaaagatttttctaataaaaagatcagataatgggacagttgaga 8760
 gagatggctttaaatacattttaa**gaatcat**tttctatttactgaccactgtaatga 8820
 aaatatatcaatttattt**gg**aactcctgattgggataaatattttaaaggatctgtt 8880
 gcacacttggattttcaaaactcgggaaagttacaagtttgcattgtaagaataaaaata 8940
 agaatattgaaactggtacattagct**aa**tctattactacttagcgtgtttctaatgaga 9000
 9001 agttact**g**aaatctat**factg**ctt**taataaaa**atttgagtagaaaaaagtggaactag 9058

Comment [ER70]: Ex 45/46

Comment [ER71]: Ex 46/47

Comment [ER72]: Ex 47/48

Comment [ER73]: Ex 48/49

Comment [ER74]: Ex 49/50

Comment [ER75]: Ex 50/51

Comment [ER76]: Ex 51/52

Comment [ER77]: Ex 52/53

Comment [ER78]: Ex 53/54

Comment [ER79]: Ex 54/55

Comment [ER80]: Ex 55/56

Comment [ER81]: Predicted Sulfated Tyrosine

Comment [ER82]: Polyadenylation signal

Comment [ER83]: 3' UTR

Key to colors and analyzing:

Red *: Stop codon

Letterletter: Intron sites, first letter is ending of one exon, second is beginning of another exon.

Blue M: Methionine

Bolded Letters: the different domains within the gene

Alpha helix

Beta sheet

Bold red letters: DUF4704

Bold italicized letters: DUF4804

Bold black letters: PH_BEACH domain

Bold italicized letters: BEACH domain

Bold underlined letters: WD40

Bold orange letters – Phosphorylated Tyrosine, Serine or threonine (NetPhos)

(XX) - Phosphorylated residue per GPS