

Additional File 8: Alignment of the gene sequences isolated from BACs. The aligned nucleotide and amino acid sequences for the five unique genes isolated from two BACs are shown. Clones 091-01, 091-02 and 091-07 were isolated from BAC 091D11. Clones 190-03 and 190-04 were isolated from BAC 190G12.

910 920 930 940 950 960 970 980 990 1000 1010 1020 1030 1040 1050
091-02 GGACCTGGATTTGGTGGCTCCAGACCAGTTGGTGGTGGAGGAGACCTTTCTTCGGCCAAGGAGGCAGGCGTGGTGGTGGAGAAGAAGAACTGATGCTGCCAACAAATTTGGTGGTCTAGGAGGGCGCGGTTCAGTTTGATGGT 423
G P G F G G S R P V G A G G R P F F G Q G G R R G D G E E E T D A A Q Q I G D G L G G R G Q F D G
091-07 GGACCTCGATTTGGTGGCTCCAGACCAGATGGTGGTGGAGGAGACCTTTCTTCGGCCAAGGAGGCAGGCGTGGTGGTGGAGAAGAAGAACTGATGCTGCCAACAAATTTGGTGGTCTAGGAGGGCGCGGTTCAGTTCAATGGT 334
G P R F G G S R P D G A G G R P F F G Q G G R R G D G E E E T D A A Q Q I G D G L G G R G Q F N G
091-01 GGACCTGGATTTGGTGGCTCCAGACCAGATGGTGGTGGAGGAAGACCTTTCTTCGGCCAAGGAGGAAGGCGTGGTGGTGGAGAAGAAGAACTGATGCTGCCAACAAATTTGGTGGTCTAGGAGGGCGCGGTTCAGTTTGATGGT 836
G P G F G G S R P D G A G G R P F F G Q G G R R G D G E E E T D A A Q Q I G D G L G G R G Q F D G
190-03 GGACCTGGATTTGGTGGCTCCAGACCAGATGGTGGTGGAGGAAGACCTTTCTTCGGCCAAGGAGGCAGGCGTGGTGGTGGAGAAGAAGAACTGATGCTGCCAACAAATTTGGTGGTCTAGGAGGGCGCGGTTCAGTTTGATGGT 909
G P G F G G S R P D G A G G R P F F G Q G G R R G D G E E E T D A A Q Q I G D G L G G P G Q F D G P
190-04 GGACCTGGATTTGGTGGCTCCAGACCAGATGGTGGTGGAGGAAGACCTTTCTTCGGCCAAGGAGGCAGGCGTGGTGGTGGAGAAGAAGAACTGATGCTGCCAACAAATTTGGTGGTCTAGGAGGGCGCGGTTCAGTTTGATGGT 773
G P G F G G S R P D G A G G R P F F G Q G G R R G D G E E E T D A A Q Q I G D G L G G P G Q F D G P

1060 1070 1080 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
091-02 CATGGACGTGGACATCATGGTCAACCGTCAAGGTCCTCCCTCAGGACCGACAGAGGAACAAACCGTTCCGGTCAGCGCAACGAAAGCAGCGATGAGGGTGGCCGCTCCTCACCCCTGCCACCATGGCCGCCACCAC 555
H G R G H H G H R Q G P P Q D R P E E Q P F G Q R N E S S D E G G R P H P R H H G R H H
091-07 CATGGACGTGGACATCATGGTCAACCGTCAAGGTCCTCCCTCAGGACCGACAGAGGAACAAACCGTTCCGGTCAGCGCAACGAAAGCAGCGATGAGGGTGGCCGCTCCTCACCCCTGCCACCATGGCCGCCACCAC 466
H G R G H H G H R Q G P P H D R P E E Q P F G Q R N E S S D E D G R P H P R H H G R H H
091-01 CATGGACGTGGACATCATGGTCAACCGTCAAGGTCCTCCCTCAGGACCGACAGAGGAACAAACCGTTCCGGTCAGCGCAACGAAAGCAGCGATGAGGGTGGCCGCTCCTCACCCCTGCCACCATGGCCGCCACCAC 962
H G R R H H G H R Q G P P Q D R P E E Q P F G Q R N E R N E E D G R P H P H H H G H
190-04 GGACGTGGACATCATGGTCAACCGTCAAGGTCCTCCCTCAGGACCGACAGAGGAACAAACCGTTCCGGTCAGCGCAACGAAAGCAGCGAGGAGGATGGCCGCTCCTCACCCCTGCCACCATGGCCGCCACCAC 1032
G R R H H G H R Q G H P Q D Q A E E Q P F G Q R N E S S E E D G R P H P H H H R H
190-03 GGACGTGGACATCATGGTCAACCGTCAAGGTCCTCCCTCAGGACCGACAGAGGAACAAACCGTTCCGGTCAGCGCAACGAAAGCAGCGAGGAGGATGGCCGCTCCTCACCCCTGCCACCATGGCCGCCACCAC 791
G R R H H G H R Q G H P Q D Q A E E Q P F G Q R N E S S E E D G R P H P H H H R H

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350
091-02 CAGCATCATCAT Q H H H 567
091-07 CAGCATCATCAT Q H H H 478
091-01 CATGGCCACCATGGCCACCACCCTCATCATCAACCAGACAGAGGGTCAACCAAGGACATATGAGACGGGAGATCAAGATCAGGACAAACCAATGATACGAGGCCCTTCCGGTTCACCCCTTCGGC 1091
H G H H G H H H R H H N Q T E G H Q G H N E T G D Q D Q D K P N D T R P F R F N H F G
190-03 CATCATCAACCACACAGAAGGTCACCAAGGTCACCAAGGTCATCAATGAGACGGGAGATCAAGATCAGGACAAACCAATGATACGAGGCCCTTCCGGTACAAACCCTTCGGTATGAAAGCCTTTCGGAGAC 1170
H G R H H R H H N H T E G H Q G H N E T G D Q D Q D K L H D T R P F R Y N H F G R K P F G D
190-04 AGAAAGCCTTTCGGAGAC 809
R K P F G D

1360 1370 1380 1390 1400 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
091-02 CGCAACCACACAGAAGGTCACCAAGGTCACCAATGAGACAGGAGATCACCCCCACCCTCATCACAAACAGACCAGGAGATGGAGATCAGGACAGACCAATGTTTGGATGAGGCCCTTCGGCTTCAACCCCTTCGGT 702
R N H T E G H Q G H N E T G D H P H R H H N K T G D G D Q D R P M F F E M R P F R F N P F G
091-07 CGCAACCACACAGAAGGTCACCAAGGTCACCAATGAGACAGGAGATCACCCCCACCCTCATCACAAACAGACCAGGAGATGGAGATCAGGACAGACCAATGTTTGGATGAGGCCCTTCCGGTTCACCCCTTCGGT 613
R N H T E G H Q G H N E T G D H P H R H H N K T G D G D Q D R P M F F E M R P F R F N P F G
091-01 AGACGCCAACCCACAGAAGGTCACCAAGGTCATAATGAGACGGGAGATCACCCCCCTCATCACAAACAGACCAGGAGATGGAGATCAGGACAGACCAATGTTTGGATGAGG 1205
R N H T E G H Q G H N E T G D H P H R H H N K T G D G D Q D R P M F E M R
190-03 CGTCCCTTCGGCAGACCAACCATACAGAAGGTCACCAAGGTCATAATGAGACGGGAGACCCCCCATCGTATCACAAACAGACCAGGAGATGGAGATCAGGACAGACCAATGTTTGGATGAGGCCCTTCCGGTTCACCCCTTCGGT 1320
R P F G R R N H T E G H Q G H N E T G D H P H R H H N K T R D G D Q D R P M F E M R P F R F N P F G
190-04 CGTCCCTTCGGCAGACCAACCATACAGAAGGTCACCAAGGTCATAATGAGACGGGAGACCCCCCATCGTATCACAAACAGACCAGGAGATGGAGATCAGGACAGACCAATGTTTGGATGAGGCCCTTCCGGTTCACCCCTTCGGT 959
R P F G R R N H T E G H Q G H N E T G D H P H R H H N K T R D G D Q D R P M F E M R P F R F N P F G

1510 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620 1630 1640 1650
091-02 AGAAAGCCTTTCGGAGACCGTCCCTTCGGCAGACGC AACGGAAACCGAAGGATCTCCAGGCGTATGGCCAACGTCGGCCCTATGGT 792
R K P F G D R P F G R R N G T E E G S P R R D G Q R R P Y G
091-07 AGAAAGCCTTTCGGAGACCGTCCCTTCGGCAGACGC AACGGAAACCGAAGGATCTCCAGGCGTATGGCCAACGTCGGCCCTATGGT 703
R K P F G D R P F G R R N G T E E G S P R R D G Q R R P Y G
091-01 CCCCTTCGGTCAACCCCTTCGGTATGAAAGCCTTTCGGAGACCGTCCCTTCGACAGACGC AACGGAAACCGAAGGATCTCCAGGCGTATGGCCACCCTCATCCCTATGGT 1319
P F W V N P F G R K P F G D R P F D R R N G T E E G S P R R D G H R H P Y G
190-03 AGAAAGCCTTTCGGAGCCGTCCCTTCGACAGACGC AACGGAAACCGAAGGATCTCCAGGCGTATGGCCAACGTCGGCCCTATGGT 1410
R K P F G G R P F D R R N G T E E G S P R R D G H R R P Y G
190-04 AGAAAGCCTTTCGGAGACCGTCCCTTCGGCAGACGC AACGGAAACCGAAGGATCTCCAGGCGTATGGCCAACGTCGGCCCTATGGT 1049
R K P F G D R P F G R R N G T E E G S P R R D G Q R R P H G

```

      1660      1670      1680      1690      1700      1710      1720      1730      1740      1750      1760      1770      1780
091-02  ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
      AACCGAGGACGTTGGGGTGAGAAATGAAAGTGAGGAGAAGGAGCATCCAACGACGGAAGCGTAACGACATCTTCACCACCTTAAGTGGTCGAGATCGCAGTCAATGAAGAAGACATCTATGTGGTCGCCGAGGTGTAG 930
      N R G R W G E N E S E E K E H P T T E S V T T S S P P * V V E I A V N E E D I Y V V A E V *
091-07  AACCGAGGACGTTGGGGTGAGAAATGAAAGTGAGGAGAAGGAGCATCCAACGACGGAAGCGTAACGACATCTTCACCACCTTAAGTGGTCGAGATCGCAGTCAATGAAGAAGACATCAATGTGGTCGCCGAGGTGTAG 841
      N R G R W G E N E S E E K E H P T T E S V T T S S P P * V V E I A V N E E D I N V V A E V *
091-01  AACCGAGGACGTTGGGGTGAGAAATGAAAGTGAGGAGAAGGAGCATCCAACGACGGAAGCGTAACGACATCTTCACCACCTGAAGTGGTTGAGATCGCAGTCAATGAAGAAGACGTCCTCAAGTGGTCGCCGAGGTGTAG 1457
      N R G R W G E N E S E E K E H P T T E S I T T S S P P E V V E I A V N E E D V Q V V A E V *
190-03  AACCGAGGACGTTGGGGTGAGAAATGAAAGTGAGGAGAAGGAGCATCCAACGACGGAAGCGTAACGACATCTTCACCACCG                                     1491
      N R G R W G E N E S E E K E H P T T E S V T T S S P P
190-04  AACCGAGGACGTTGGGGTGAGAAATGAAAGTGAGGAGAAGGAGCATCCAACGACGGAAGCGTAACGACATCTTCACCACCTGACGTTGGTCGAGATCGCGATCAAT-----GATGTCGCCGAGGTGTAG 1172
      N R G R W G E N E S E E K E H P T T E S V T T S S P P D V V E I A I N - - - - - D V A E V *

```