Processed sequence:

51         TCTCAGCTCACCACAACCTCCGCCTCCCAGTTTCAAGCGATTCTCCTC           101         TTAGCCTCCCAAGAGCCTGCAACTACAGGGCCCCCCCCCC	ICCTGCC CTAATTA FGGTTTC ACTGGGA FTTTTTC 3GGTTTC 3GCAAGC GGAAACG CGTGATC CCACAGC CCTGGTG ACTGAAA 3TAAAAA 3TAGAGC
101       TTAGCCTCCCAAGTAGCTGAACTACAGGTGCCACCATGCCCGGCTAAT         151       TTATATTTTGGTAGAGATGGGGTTCACCATGCTCTCCCAGGCTGGTGATTAATTA	CTAATTA FGGTTTC ACTGGGA FTTTTTC 3GGTTTC 3GCAAGC 3GCAAGC CGTGATC CCACAGC 2CTGGTG 4CTGAAA 3TAAAAA 3TAGAGC
151       TTGTATTTTGGTAGAGATGGGGTTTCACCATGCTGTCCAGGCTGGT         201       AAACTCCTGACTCTCGATCCTCCGGCTCCGAAAGTACTGG         201       TTATATGGCGAGATCACTGCGCCGCGCCGCGCGCAAAGTAATTAAT	IGGTTTC ACTGGGA ITTTTTC 3GGTTTC 3GCAAGC 3GCAAGC 2GTGATC 2CACAGC 2CTGGTG 4CTGAAA 3TAAAAA 3TAGAGC
AAACTCCTGACCTTCGATCCTGCCGCCTCCGGCTCCCAAAGTACTG TTATATGCGTGAGTCACTGCGCTGGCCTGGCCTGCGCTCACGGGTT AGCAATTGCCTGCCTCGGCTACTTTTGTGTTTTTTTTTT	ACTGGGA ITTTTTC GGTTTC GGCAAGC GGCAAGC CGTGATC CCACAGC CCTGGTG ACTGAAA FTAAAAA FTAGAGC
1         TTATATGCGTGAGTCACTGCGCTGCCGCGGTGATATAATTATTTTT           251         TTATATGCGTGAGTCACTGCACTGCCGGTGATATAATTATTTTT           351         AAGCAATTCCCGGCCTCGCCGCGCGGATATAATTATTTTTT           351         AAGCAATTCCCGCCCCAGCCCGCCGCGGATATAATTACAGGCAA           451         GGGTTTCACCATGTTGGTCAGGCTAGACTCGAACTCCTAACGCGCAC           551         ACCGGCTGAACTCTTTTTACATAATTAACTAATTGGCCCCCCC           551         ACCTGGCTGACTCTTCCTCCCAACCCCTTTACCAATAACTGAC           651         TACTCTGGATGCCGCGCAATATTCTTACATATTATTATTTTTTTT	TTTTTTC GGGTTTC GGCAAGC GGGAGACG CGTGATC CCACAGC CCTGGTG ACTGAAA FTAAAAA FTAGAGC
<ul> <li>TARGTGANTAGCANTCRETCRECCEGORGALMALTIANTIANTIANTIANTIANTIANTIANTIANTIANTIAN</li></ul>	GGGTTTC GGCAAGC AGAGACG CGTGATC CCACAGC CCTGGTG ACTGAAA FTAAAAA FTAGAGC
<ul> <li>Sofi Chao IGAMIAGCARICIC IGACIGUACIC TEACIGUA TECCAGOR TIACAGGAN</li> <li>GCAACCATGCCTGGCTAATTTTTGTATTTTTTTTTTTTT</li></ul>	GGGTTTC GGCAAGC AGAGACG CGTGATC CCACAGC CCTGGTG ACTGAAA FTAAAAA FTAGAGC
AAGCAATTCTCTCTCCCTAAGCTCCTGAGTAGCTCGGATTACAGGAA         GGGTTTCACCATGTTGGTCAGGCTAGACTCGAACTCCTAACCTCGTGA         GGGTTTCACCATGTTGGTCAGGCTAGACTCGAACTCCTAACCTCGTGA         CACCCACCTCCACCTCTCCCCAAAGTGCTGGAGTACAGGCAGG	AGAGACG CGTGATC CCACAGC CCTGGTG ACTGAAA FTAAAAA FTAGAGC
401       GCCACCATGCCTGGCTAAGTTTTGTATTTTTTTTTTTTT	AGAGACG CGTGATC CCACAGC CCTGGTG ACTGAAA FTAAAAA FTAGAGC
451 GGGTTICACCATGGTTGGTCAGGCTAGACTCGAACTCCTAACCTCGTGACA 501 CACCCACCTCAGCCTCCCAAGGTGCGGGGATTACAGGCAGG	CGTGATC CCACAGC CCTGGTG ACTGAAA FTAAAAA FTAGAGC
501       CACCECACCTCACCTCCCAAAGTGCTGGGGATTACAGGCAGGAGCCACP         551       ACCTGGCTGAACTCTTTCTTTTTTTTTTTTTTTTTTTTT	CCACAGC CCTGGTG ACTGAAA 3TAAAAA 3TAAAAA
551       ACCTGGCTGAACTCTTTCTTTTTCCTTTTAAATTAACTTATTTGCCCTGG         601       GCAGGGACCTCTTGCCCCCTCTCCTCACCACCCCTTTACAATAACTGA         601       GCTTTTAATACTACTGCCCCCCCCTTTCCTAACTATTTTTTTT	CCTGGTG ACTGAAA GTAAAAA GTAGAGC
601         GCAGGGACCTCTTGCCCCCTCTCCAACCCCTTTACAATAACTG2           651         TACTCTGATGTCCTGGCATATTTTTACTATTTATTTTTTTT	ACTGAAA GTAAAAA GTAGAGC
651         TACTCTGATGTCCTGCAATATTCTTACTATTTATTTTTTTGTAAP           701         GCTTTTAATACTACTCAACTGCAGTAGTGAGGGGGGGGGG	GTAAAAA GTAGAGC
701       GCTTTTAATACTACTCAACTGCAGTAGTGAGGGGGGGGGG	GTAGAGC
751AAGGAGTATGATCTGTAACTGATCCTGAACAATCAATTGGGGATAACTCG801TGCCTAGATAATTCACTATTTAGATGCTGTATTATTAGGGCCAACAGCG801CTTTTCTCTCCTCCAACATCAAGGTAAGGAGAATTGTCATAGGTGCCCC901CTTTTCTCTCCTCCTCACACTCACCTCCCTGAACTCCGAAGGCCCCA901CTCTTCTGGCCAATAATATTATACTATTTCAGAGAAATTGAGAAGGGA1001CTCTTCTGGCCAATAATATTATACTATTCCACCTCCCCCGGGGCTAGAC101ACTAGGGGTGAGAGGCATTCTTAGATATTCCTCCTCCTCCTGGGGCAACAACATAGT1051AAACTTTGGAAGGCAGGAGGAAGTCATAATACCCTGGGCAAAAGGTCCCC1151AGGTAAGGTCAGGACCATTTTTTTACATTGGGCCAAGGAAGACCAAAGGTCCCC1201TGGTCAAGGTCAGGACCATTTTTTTGCCCTCTATTTTTCCAGGACACAAAGGTC1201TGGTCAAGGACGACTTTAGGTAAGATCGGCAACAAAGGTCC1301AACTCTCCAGTAGACCTTAGGACATGGGTATACATAGTAATCCAAGGCCT1301AACTCTCCAGTAGACCTTAGGTAGGTAAGATCCAAAGCTCC1401CCGTGAGTCAGAAATTTTTTCTCAGCACCATAGGAAAAAAACCATAA1451AAAAAATTAAGAGAAGAAATGTGTTGTTTGTTTTTGGGTTATTTTTT	
801TGCCTAGATAATTCACTATTTAGATGCTGTATTATTAGGGTCAACAGC851ACTGGAGCAACAGCTAAAGTCAAGATAAGAGAATTGTCATAGGGTGCTG901CTTTTCTCTCCTCAACATCATCACACTTCACCTTCTGAACTAGGACGGAGG951GTTCTTCTCCTCCTCTTCTAACACCTACTGAGAAATTGAGAAGGGCGA1001CTCTTCTGGCCAATAATATTATACTATTTCAGAGAAAGGTCCTCATATC1051AAACTTTGGAAGGCATTCTTAGATATTCCTCCTCCTGCGGGCTAGAC1061ACTAGGGGTGGGAAGGAGGAGCACTATAATACCTGGGCAACAATAGT1201TGGTCAAGGTCAGGACCATTTTTTTTTCCTCCTATTTTCAGGAACCA1201TGGTCAGGACCATTTTTGACAATCGGGTTGAGGTTAGTCTCA1201AAATGGTCAGGACCATTTTTGACAATCGGCTTGACAAATTCAAGGCC1301AACTCTCCAGTAGACCTTAGACATAGGATAACATAGTAAATTCAAGGCCT1301AACTCTCCAGTAGACCTTAGTTAGGTATACATAGTAAATTCAAGGCCG1401CCGTGAGTCAGAAATTTTTTTTCTCCAGGACCATAGGAAAAAAATCATAA1451AAAAAATTAAAGAGAAGAAATGTGTGTGTTTTGTTTTG	ACTCAC
851ACTGGAGCAACAGCTAAAGTCAAGATAAGAGAATTGTCATAGGTGCTCG901CTTTTCTCCTCCTCAACATCACCACCTACTGAGAAATTGAGAAGGGCAG951GTTCTTCTGCCCCTCTCTCTAACACCTACTGAGAAATTGAGAAGGGCAGGGCA1001CTCTTCTGGCCAATAATATTATACTATTTCAGAGCAAAGTTAGGAAGGCAGGGCTAAAGG101ACTAGGGGTTAAAAAGAAATCCCTTTTACATTGGGTCAAAGGTCCCC101ACTAGGGGTGAGGAGGAGGAGGAGTCATAATTCGGCTCGCAACAATAGT1201TGGTCAAGGTCAGGACCATTTTTTTTCCTCCCTGCTGCCGCAACAATAGT1201TGGTCAAGGTCAGGACCATTTTTTTTCCTCCCCTGCTGGATGGTTTAGTCTC1301AACTCTCCAGTAGACCTTAGAACTGGGTATACATAGTAATTCAAGTCTTC1301AACTCTCCCAGTAGACCTTAGGATAAGAACTGACAAATTCAAAGCCT1401CCGTGAGTCAGAAATTTTTTTCCAGCACCATAGGAAAAAAACATAA1451AAAAATTAAGAGAAGAAATGTGTTGTTTGTTTTGGGTATACTCAAGCGTGATC1551GGCTCACCGCAACCTCTGGCCTCCTGGGTCAAGCGAATGCCGTGACC1551GGCTCACCGCAACCTCTGGCTCCTGGGTCCACCACCATCCTTGGCTAA1551GGCTCACCGGAACCTCTGGCTCCCGGGCCTCCCAAAGTGGCTGAT1701CGAACTCCTGACCTGGACCGCCCCCCCCGGCCCCCCCAAAGTGGCTG1701CGAACTCCTGACCTGGACCGCCCCCCCCGGCCCCCCACAAGTGGCTG1701CGAACTCCTGACCTCCCCAGAGGAAGGAGGACTGTTGTAAATTCA1801ACAGGCAAAATAAGTGAAAGCACGAATTTCTGTATATCCAATGGTA1801ACAGGCAAAATAAGTGAACCCCCCTCTCCTCACCACTGGAGCCACTGAGGC2001AGAGCATTGTTGTTTCCAAGAGATATGCCCATCCTTCACACAATAAGGAT2101AACCCTAGTTCTCTATAGCCCATCCTTCACACAATAAAGAATAAGC2011AACCCTAGTGTGTTTCAAGGAAATGACTCCTGGCCCAACCACCACACACA	ACAGCTA
901CTTTTCTCTCCTCAACATCATCACCTTCACCTTCCTGAACTCCGCCCAC951GTTCTTCTTCTCCTCCTTCTCTAACACCTACTGAGAAATTGAGAAGGCGA1001CTCTTCTGGCCAATAATATTATACTATTTCAGAGCAAGGTTCTATATC1051AAACTTTGGAAGGCATTCTTAGATATTCCTCCTCCTCGCGGGCTAGAC1011ACTAGGGGTAAAAAGAAATCCCTTTTACATTGGGTCAAAGGTCCC1201TGGTCAAGGCCAGGACCATTTTTTTTTCCTCCTATTTTCAGGACCA1201TGGTCAGGACCATTTTTGGACAATTCGGCTTGATGGTTAGTCTC1301AACTCTCCAGTAGACCTTAGAACTTGGTATACATAGTAATTCAGTCCT1301AACTCTCCAGTAGACCTTAGTAGAGTAAGATCTGACAAATCCAAGGCG1401CCGTGAGTCAGAACTTTTTTTTCCCAGCACATAGGAAAAAAAA	IGCTGGA
951GTTCTTCTTCTCCTCCTTCTCAACACCTACTGAGAAATTGAGAAGGCGA1001CTCTTCTGGCCAATAATATTATACTATTTCAGAGCAAGGTTCTATATC1051AAACTTTGGAAGGCATTCTTAGATATTCCTCCTCCTGCGGGCCAGAG1011ACTAGGGGTTAAAAAGAAATCCCTTTTTACATTGGGTCAAAGGTCCCC1151AGGGTAGGACGACCATTTTTTTCCCCCTATTTTTCAGGACACAATAGT1201TGGTCAAGGTCAGGACCATTTTTTTTCCCCCTATTTTTCAGGACCAA1201TGGTCAAGGACCATATTTTGACAATTCGGCTTGATGGTTAGACCTC1201AAATGGTCAGGACCATTTTTGACAATTCGGCTGATGGTTTAGATCATC1201AAATGGTCAGGACCATTTTGACAATTCGGCTGATGGTTTAGATCATAGTAGTCTC1301AAATGTCAGAAATTTTTTTTCCAGCACCATAGGAAAAAAATCATAAC1401CCGTGAGTCAGAAAATTTTTTTTCCCAGCACCATAGGAAAAAAATCATAAA1451AAAAAATTAAGAGAAAGAAATGGTGTTGTTTTGGTTTTGGGTTATTTTTT	CCACAT
1001CTCTTCTGGCCAATAATATTATACTATTCAGAGCAAGGTTCTATATG1051AAACTTTGGAAGGCATTCTTAGATATTCCTCCTCCCTGCGGGCTAGAG1011ACTAGGGGTTAAAAAGAAATCCCTTTTTACATTGGGTCAAAGGTCCCC1151AGGTAGGGTGGGAAGGAGGAGGAAGTCATAAATACCCTGGCAACAATAGT1201TGGTCAAGGTCAGGACTATTTTTGACAATACCTGGTGATTAGTCTTC1301AAATGTCACGAGCTATATTTTGACAATTCGGCTTGATGGTTTAGTCTC1301AACTCTCCAGTAGACCTTAGAACTTGGTATACATAGTAATCAAGCCT1401CCGTGAGTCAGAAATTTTTTTTCTCAGCACCATAGGAAAAAAAA	GCGATT
1051AAACTTTGGAAGGCATTCTTAGATATTCCTCCTCCTGCGGGGCTAGAC1101ACTAGGGGTTAAAAAGAAATCCTTTAGATATTCCTCCTACTTGGGAAGAGGAGGGAG	АТАТССТ
101ACTAGGGGTTAAAAAGAAATCCCTTTTTACATTGGGTCAAAGGTCCCC1101ACGGTAGGGTGGGAAGGAGAGAGTCATAAATACCCTGGCAACAATAGT1201TGGTCAAGGTCAGGACCATTTTTTTTTTTCCCCCTATTTTTCAGGAACCA1201TGGTCAAGGCCAGGACCATTTTTTTTTTCCCCCCTATTTTTCAGGAACCA1201AACTCTCCCAGTAGACCTTAGAACTTGGCATACATAGTAATTCAGTCTC1301AACTCTCCCAGTAGACCTTAGAACTTGGTATACATAGTAATTCAGTCCT1301AACTCTCCCAGTAGACCTTAGTAGAGAAGATCTGACAAATTCAAGTCCT1301AACTCTCCCAGTAGACCTTAGTAGGACAAGATCTGACAAATTCAAGGCCT1401CCGTGAGTCAGAAATTTTTTTTTCCCAGCACCATAGGAAAAAAATCATAA1451AAAAATTAAGAGAAGAAATGTGTTGTTTTGTTTTGGTTTTGGGTTATTTTTT	РАСАСТТ
1151AGGGTAGGGTGGGAAGGAGGAGGAGGAGTATAAATACCCTGGCAACAATAGT1201TGGTCAAGGTCAGGACCATTTTTTTTTCCCCCTATTTTTCAGGAACCA1251AAATGGTCAGGACTATTTTTTGACAATTCGGCTTGATGGTTTAGTCTCC1301AACTCTCCAGTAGACCTTAGAACTTGGTATACATAGTAATTCAAGCCT1301AACTCTCCAGTAGACCTTAGAAGTAGAGTAGAAATTCAAAGCCT1401CCGTGAGTCAGAAATTTTTTTTTCCAGCACCATAGGAAAAAAAA	ГССССТС
1101NGGGTAGGGTGGGAGGAGGAGGAGGAGTATAAATAAAAAAA	
1251AAATGGTCAAGGACCAATTTTTGACAATTCGGCTTGATGTTAGTCATGGTCAGGACCATTTTGACAATGCGCTTAGTTAG	
1251AAATGGTCAGGACTATITTTTGACAATTCGGCTGTGATGGTTTAGTGTTTGGT1301AACTCTCCAGTAGACCTTAGAACTTGGTATACATAGTAATTCAGTCCT1351AATATTAAAAGCCTTTAGTTAGAGTAAGATCTGACAAATTCAAAGCTG1401CCGTGAGTCAGAAATTTTTTTTCTCAGGCACCATAGGAAAAAAATCATAA1451AAAAAATTAAGAGAAAGAAATGTGTTGTTTGTTTTGGGTTATTTTTT	
1301AACTCTCTCAGTAGACCTTAGAACTTGGTATACATAGTAATTCAAAGTCCT1351AATATTAAAAGCCTTTAGTTAGAGTAAGATCTGACAAAATTCAAAGCTG1401CCGTGAGTCAGAAATTTTTTTTCTCAGCACCATAGGAAAAAAATCATAA1451AAAAAATTAAGAGAAGAAATGTGTTGTTTTGGGTGGGGGCAATGGCGTGATT1501GACGGAATTTCGCTCTTGTGTCCCAGGCTGGAGGGCAATGGCGGGGCTGCT1551GGCTCACCGCAACCTCTGCCTCGTGGGTTCAAGCGATTCTCCTGGCTCA1651TTTGTATTTTTAGTAGAGACAGGATTTACAAACACATGCCACCATCGTGGGCTGGT1701CGAACTCCTGACCTCGTGATCCGCCCGCCCCGGCCTCCCAAAGTGCTG1751ATTACAGGCGTGAGCCACTGCACCTAGCTGCTGTGTTTTTATAGGTA1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAAATTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGAAATTAA1951GCTCATTGTTGTGTTTTCAAGAGTTATTCGTAAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCGTAATTCCAATTGATGTG2101AACCCTAGTTCTCATTTTATAGCCAATTCCTTCACACAAATAAAGACT2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCAGAGGAGAGTAGTTCTTAAACAATAGC2201TCAGAAAATCACCTGGTTCTCTCAAGAGTAGCTCCTACTCTAACAATAGACT251AAATCACAGGGAGAATTTTAAAACAGGAAGGCTGCCAACCAA	
1351AATATTTAAAAGCCTTTAGTTAGAGTAAGATCTGACAAATTCAAAGCTG1401CCGTGAGTCAGAAATTTTATTTTCTCAGCACCATAGGAAAAAAATCATAP1451AAAAAATTAAGAGAAGAAATGTGTTGTTTTGTTTTGGGTTATTTTTT	TCCTTG
1401CCGTGAGTCAGAAATTTTTTTTTTTCTCAGCACCATAGGAAAAAAATCATAA1451AAAAAATTAAGAGAAGAAATGTGTTGTTTTGTTTTGGGTTATTTTTT	AGCTGAG
1451AAAAATTAAGAGAAGAAATGTGTTGTTTTGTTTTGGGTTATTTTTT	CATAATA
1501GACGGAATTTCGCTCTTGTTGCCCAGGCTGGAGGGCAATGGCGTGATC1551GGCTCACCGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTCCCTGCCTC1601CCTCCTGAGTAGCTGGAATTACAAACACATGCCACCATCCTTGGCTAA1651TTTGTATTTTTAGTAGAGAGACAGGATTTTACCATGTTGGTCAGGCTGGT1701CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTG1751ATTACAGGCGTGAGCCACTGCACCTAGCTGGTGTTTTTATAGGTA1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAAATCT1851AAACTAGCAGTTTCCGTAGCTGTGGTGTTCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2011CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAAC2201TCAGAAAAATCACCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAAC2201TCAGAAAAATCACCTGGTGAAGCCCATCATCAGCATCTCTAACAAAAAA2251AAATCACAGGGAGAATTTTAAAACACAGATTCCTGGTCCCAACCACCAC2301GATTCCTATTCAGTGGTGAAGCCCATCAATCAGCATCTCTAACAAGCT2351CAAGTGGTGCTGAGGATGCTGGTGAAATAAAAAAAAAAA	ſTTTTGA
1551GGCTCACCGCAACCTCTGCCTCGCGGTTCAAGCGATTCTCCTGCCTC1601CCTCCTGAGTAGCTGGAATTACAAACACATGCCACCATCCTTGGCTAA1651TTTGTATTTTTAGTAGAGACAGGATTTTACCATGTTGGTCAGGCTGGT1701CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTG1751ATTACAGGCGTGAGCCACTGCACCTAGCTGGTGTTTTTATAGGTA1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAATCT1851AAACTAGCAGTTTCCGTAGCTGTGGGTGTTCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGAGCTGTTGTAAATTAA1951GCTCATTGTTGTTTTCAAGACTTATTCTGTATATCACATTGATGTGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCATCAACAAATAAGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	ſGATCTT
1601CCTCCTGAGTAGCTGGAATTACAAACACATGCCACCATCCTTGGCTAA1651TTTGTATTTTTAGTAGAGACAGGATTTTACCATGTTGGTCAGGCTGGT1701CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTG1751ATTACAGGCGTGAGCCACTGCACCTAGCTGCTGTGTTTTTATAGGTA1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAATCT1851AAACTAGCAGTTTCCGTAGCTGTGGTGTTCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTGTTTTCAAGCCTCATTGAATTGTCAGTGTAGGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTCG2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAAATCACCTGGTTCTCTAGAAGAGTAGTTCTTAAACAAAAAA2251AAATCACAGGGAGAATTTTAAAACACAGGTGGTGCCCAACCACCA2301GATTCCTATTCAGTGGTGAAGCCCATCAATCAGCATCTCTAACAAGCT2351CAAGTGGTGCTGAGGATGCTGGTGAATGATCCATTCATACTCTTCAACAAGCT2401CTGTAGATAATAAATTAAACAGGCAAAAATAAAAAAAAAA	JCCTCAG
1651TTTGTATTTTTAGTAGAGACAGGATTTTACCATGTTGGTCAGGCTGGT1701CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTG1751ATTACAGGCGTGAGCCACTGCACCTAGCTGCTGTGTTTTTATAGGTA1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAATCT1851AAACTAGCAGTTTCCGTAGCTGTGGTGTTCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCATATTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCAACAAAAAAAA	JCTAATT
1701CGAACTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTG1751ATTACAGGCGTGAGCCACTGCACCTAGCTGCTGTGTTTTTATAGGTA1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAATCT1851AAACTAGCAGTTTCCGTAGCTGTGGTGTTCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCATTATGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCAACAAAAAAAA	CTGGTCT
1751ATTACAGGCGTGAGCCACTGCACCTAGCTGCTTGTGTTTTTATAGGTA1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAATCT1851AAACTAGCAGTTTCCGTAGCTGTGGTGTTCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	ſGCTGGG
1801ACAGGCAAAATAAGTGAAGAATATGCTGTATAATTCCATCTGTAATCT1851AAACTAGCAGTTTCCGTAGCTGTGGTGTGTATCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	AGGTATA
1851AAACTAGCAGTTTCCGTAGCTGTGGTGTTCATCTTCCCATCATTGGTA1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	ATCTTC
1901TGGAGAGTAACTCTACCTCCCAGAGGAAGGAGGACTGTTGTAAATTAA1951GCTCATTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	FGGTAAT
1951GCTCATTGTTCAAGCCTCATTGAATTGTCAGTGTAGGGGACCTGAGGC2001AGAGCATTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGC2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAAATCACCTGGTTTCCAATGGTAGGCTGTCATGAAAAAAAA	ATTAAAG
2001AGAGCATTGTTGTTGTTTTCAAGAGTTATTCTGTATATCACATTGATGTGG2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	GAGGCAA
2051TTCAATATTTAAATGTAAGGAAATGACTCCTACTCTCTATTTGTTTAC2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	FGTGCAG
2101AACCCTAGTTCTCATTTATGCCCATTCCTTCATCAACAAATAATGATT2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTTAAACAATAGC2201TCAGAAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	ГТТАСАА
2151CTACTTTTTGCCTGGTTCTCTCTAGAAGAGTAGTTCTTAAACAATAGC2201TCAGAAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	rgate
2201TCAGAAAAATCACCTGGTTTCAATGGTAGGCTGTCATGAAAAAAAA	
2251       ICAGAMAMATCACCIGGTTTCAATGGTAGGCTGTCATGAMAMAMAAAAAAAAAAA	11/10/2/11
2251       AAATCACAGGGAGAATTTTTAAAACACAGATTCCTGGTCCCAACCACCACCACCACCACCACCACCACCACCACCAC	
2301       GATTCCTATTCAGTGGTGAAGCCCATCAATCAGCATCTCTAACAAGCT         2351       CAAGTGGTGCTGAGGATGCTGGTGAATGATCCATTCATACTATCTTCTAG         2401       CTGTAGATAATAAAATTAAACAGGCAAAAATAAAAATAGGCTGGCGCGCG         2451       GCTTACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTG	
2351       CAAGTGGTGCTGAGGATGCTGGTGAATGATCCATTCATACTCTTCTAG         2401       CTGTAGATAATAAATTAAACAGGCAAAAATAAAAATAGGCTGGCGCGG         2451       GCTTACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTG	AGCITC
2401 CTGTAGATAATAAATTAAACAGGCAAAAATAAAATAGGCTGGCGCGC 2451 GCTTACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGTGGGTG	reraggg
2451 GCTTACACCTGTAATCCCCAGCACTTTGGGAGGCCAAGGTGGGTG	CGCGGTG
2501 CTGAGGTCAGGAGTTCTAGACCAGCCTGGCCAACATGGTGAACCCCGT	JGACCAC
	CCCGTCT
2551 CTACTAAAAATACAAAAAATTAGCTAGGCATGGTGGTGTGTGCCTGT	CTGTAA
2601 TCCCAGCTACTCTGGAGGCTGAGGCAGAAGAATCGCTTGAACCCAGAA	CAGAAGG
2651 CGGAGGTTGCGGTGAGCCGAGATTGCGCCATTACACTGTAGCCTGGGC	ſGGGCAA
2701 CAAGAGGGAAACTCCATCTCAAAAAAAAAAAAAAAAAAA	CTCACA
2751 GAGCTTATATTCTAGTGGTAGTAGGGGGGATAAATCAGATAGGTAGCTA	AGCTAGC
2801 TAGCTAGCTAGATAAATAGATAGATAGACAGATAGTATATTACCTGTG	CTGTGTC
	CAAGCTG

10/07/13 15.45

2901	CACTTGACCTTTCCTAATCATCTCCAACACTACCTATTGCCCTATCCAAA
2951	TCCTATCACTTGAATTATTCAAAAAGTCTCCCATCTTCATTTTGGTGGGA
3001	TTTGGTCAGCTCCTTTACTGCAACCTGTTTTATCAGGAAGGTCTTTATGA
3051	CCTGTATTTGTGCTGACCTCCTATCTCATCCTATGACTTAGAACACCCTT
3101	AACCATCTGGGAATGCAGCCCAGTAGGTTTCAGCCTCATTTACCCAGCT
3151	ССТАТТСТССССТСССССССССССССССССССССССССС
3201	СПСАСТААСТСТИТОСТСОСТСТСТСТСТСТСТСТСТСТСТСТСТСТС
3251	
3201	
3351	
3401	
2401	
2501	
2551	
2601	
2651	
2701	
3701	
3/51	
3801	
3851	
3901	
3951	GTAGTAAGTCGGCCTTCCAGCCACCAGCCCCTTCCCTTTGGTCTTTCACT
4001	CCGGAGGCTCTTACCCTAGACACAATGGGACAGGGAGCGGGGGATGGGGGG
4051	AATTCAGCTCAGGCTTTTATGCAAAGACCCCCTTCTGCAAAGAACAAAGC
4101	TTCTGGTACCTGCCCTTTGGAGAGCTGCGGGCAAGCTCAGCCTCGGTGAG
4151	TCTTGGTGGCCTTGACAGCCCCCACTTAACAAACTGTGCTGATTAAGAGA
4201	GACAGGAGGGCAAGTTTTTCCCTTCTTTTAAAGAAATCATCCTATTTCCT
4251	ACGAGACATAGACTATCTGCCTGAAGCATGATGTACTAGCCCCACTCACC
4301	GGCTCCCTGATGCCCCTATGCTTAATCTTCTCCGGAATGGTAGTCTGAGA
4351	AGAAAAAAGATTACGCCCAATTTCATTTCCTTGTTTCACATCAAGCAATA
4401	CTTTTCGAGTCTTTGCATTGTGAACAAAAGTCAGCTTGTGTGGGAGCAAA
4451	GCCAGCTGCTCTGGGTGCAGACCCAGGAGCAGAGTGCAGAGGAGAATGAG
4501	TCAAAGAGTTTTGTCTTCAAAAATTACATAATCGGGATTTGCTAAGAGTT
4551	TACTTTTCGGTATGGAAGACTGGAAAAGAGAAAGAAATCTTAGGTTTCTT
4601	GAATGTTGGGTTTGGGAATAGGAAGGAAAATCGAAAACTGTAGACTTTGT
4651	CCATAAATGTTAGTGCTGGAACCCCACTCTAAAAACTTTGTTCCTTTGGA
4701	AAACACCTCCCTTCCCCCAGAAACACACACACCACACGAGATGGGCACG
4751	GAGTAGTCTTGAAAGACATGACAAATCACCAGACCTGGGAAGAAGCTAAA
4801	GAGCCAGAGGGAAAAAGCCAGAAGTCGACTACCTGGGAGGAGGGATAGAC
4851	AAGAAACCAAACTAAAGGAAACTAAGGTAGGTGCTGAAAACAAGTACCAT
4901	TTTCAACATTAACTGATGCCTTGGCTTCATGCTATAATGCCATGTTGTGT
4951	TTCACTATAACCTCAGAGTGAATGAAAGAGGAAAATGGAGCTAGTTGAAA
5001	TTTCTGCCTAAACTAGCCAGATTTTGAGACACTAAGTTATCTCAAATCAA
5051	GAAATCACCCTAATGAGAATTTCAATAACCTCAGGAATTTAAGGTGCATG
5101	CATCCCCCACCCCCCTTTTTTTTTTGAGACGTAGTCCCGCTCTGTTGC
5151	CCAGGCTGGAGTACAGTGGCGCGATATCGGCTCACCACAACCTCTGCCTC
5201	CCAGGTTCAAGGGATTCTCCCGCCTCAGCTTCCAGAGTAGCTGGGACTAC
5251	AGACACCCACCACCATGCGTGGCTAATTTTTGTATTTTTAGTAGAGAGGG
5301	GGTTTCGCCATGTTGGCCAGGCTGGTTTCAAACTCCTGACTTCAGGTGAT
5351	CCGCCTGCCACGGCCTCCCAATTTACTGGGATTACAGGGGTGGGCCACCG
5401	CGCCCGGCCTTTTTCTTAAATTTTTAAAATATTTAAAGTTTTATCCCATTC
5451	CTGTTGAACCATATTCCTGATTTAAAAGTTGGAAACGTGGTGAACCTAGA
5501	AGTATTTGTTGCTGGGTTTGTCTTCAGGTTCTGTTGCTCGGTTTTCTAGT
5551	TCCCCACCTAGTCTGGGTTACTCTGCAGCTACTTTTGCATTACAATGGCC
5601	TTGGTGAGACTGGTAGACGGGATTAACTGAGAATTCACAAGGGTGGGT
5651	GTAGGGGGTGTGCCCGCCAGGAGGGGGTGGGTCTAAGGTGATAGAGCCTTC
5701	ATTATAAATCTAGAGACTCCAGGATTTTAACGTTCTGCTGGACTGAGCTG
5751	GTTGCCTCATGTTATTATGCAGGCAACTCACTTTATCCCAATTTCTTGAT
5801	ACTTTTCCTTCTGGAGGTCCTATTTCTCTAACATCTTCCAGAAAAGTCTT
5851	AAAGCTGCCTTAACCTTTTTTCCAGTCCACCTCTTAAATTTTTTCCTCCT
5901	CTTCCTCTATACTAACATGAGTGTGGATCCAGCTTGTCCCCAAAGCTTGC

Proscan: Version 1.7 Processed Sequence: 6001 Base Pairs

Promoter region predict	ed on fo	orward s	trand in 9	947 to 1197
Promoter Score: 56.27 (	Promote	r Cutoff	= 53.0000	000)
TATA found at 1174, Est	.TSS = 2	1204		
Significant Signals:				
Name	TFD #	Strand	Location	Weight
Sp1	<u>S00064</u>	-	947	5.934000
Sp1	<u>S01542</u>	-	947	3.608000
junB-US2	<u>S01738</u>	+	1008	2.717000
JCV_repeated_sequenc	<u>S01193</u>	-	1086	1.658000
PuF	<u>S02016</u>	+	1157	1.082000
JCV repeated sequenc	S01193	+	1157	1,427000

Promoter region predicted on forward strand in 5474 to 5724 Promoter Score: 54.96 (Promoter Cutoff = 53.000000) TATA found at 5702, Est.TSS = 5732 Significant Signals: Name TFD # Strand Location Weight SIF <u>s02021</u> - 5621 1.161000 1.082000 PuF <u>S02016</u> + 5641 <u>S00801</u> + 5663 2.755000 Sp1 5668 2.772000 Sp1 S00781 -

SDR_RS	<u>S01561</u>	+	5671	1.554000
AP-2	<u> 501936</u>	-	5680	1.091000
TFIID	<u>\$00087</u>	+	5703	2.618000
TFIID	<u>S01540</u>	+	5703	1.971000

Promoter region predicted on reverse strand in 1186 to 936 Promoter Score: 57.47 (Promoter Cutoff = 53.000000)

Significant Signals:

Name	Strand	Location	Weight
JCV_repeated_sequenc	+	1157	1.658000
PuF	+	1157	1.391000
JCV_repeated_sequenc	-	1086	1.427000
junB-US2	+	1008	1.510000
Sp1	-	947	6.661000
Sp1	-	947	10.681000
Sp1	-	947	17.211000
JCV_repeated_sequenc	-	946	1.427000
Sp1	-	946	7.086000
Sp1	-	946	3.013000
Sp1	+	941	3.061000
EARLY-SEQ1	+	939	5.795000