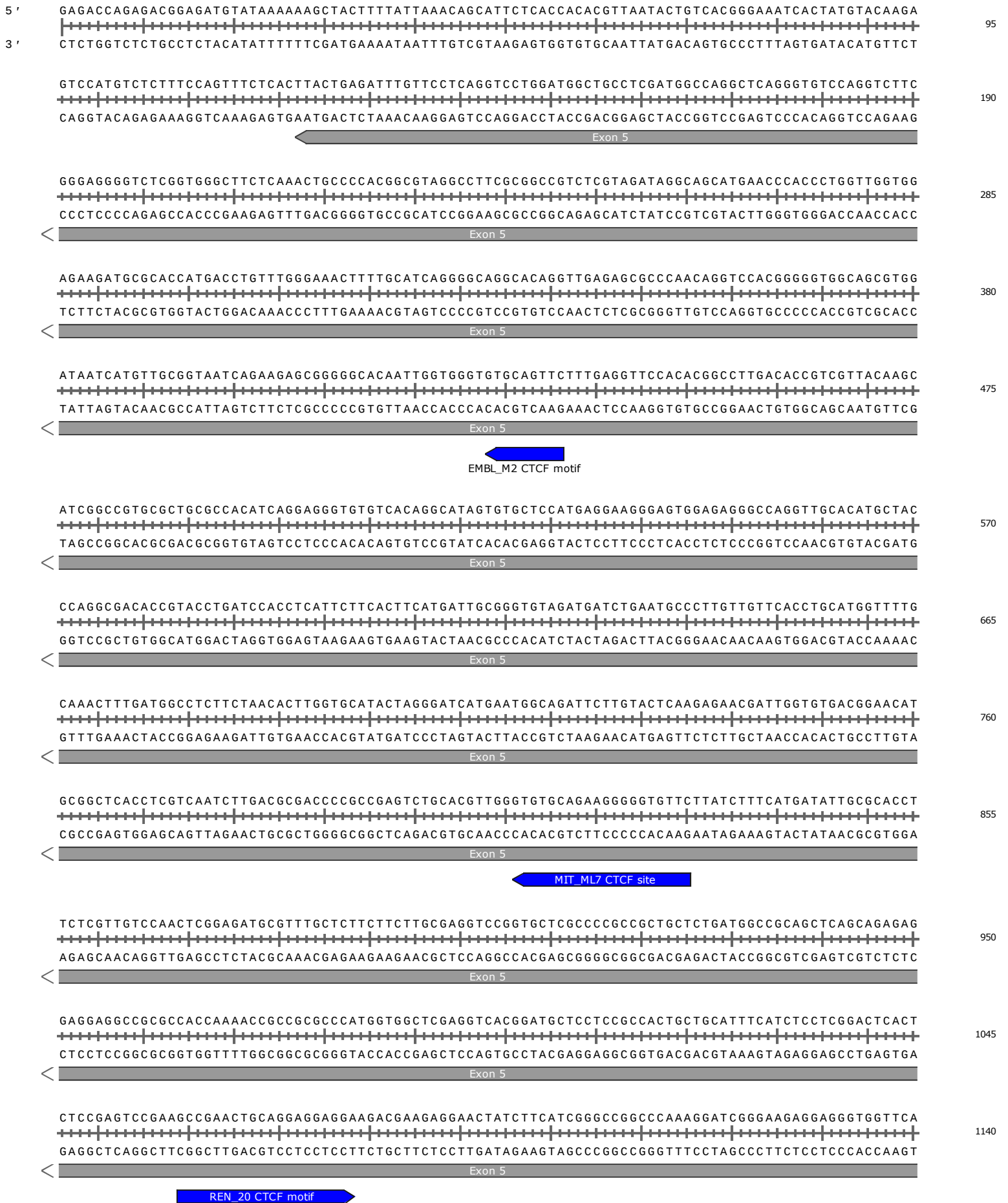


Supplementary Figure 1. The major immediate early locus (NC_006273.2 nucleotides 170,568-176,978) is displayed with annotated exons, the transcription start site (TSS), and the CTCFBSDB2.0-predicted CTCF binding sites, using the nomenclature of the database to describe each site.



TCTGGGAGAGCGGGTGGTGGGAGAGGTCACCTCGCGGGTGCCTGCGCGTGGGAAAGGGGAAGACGCGGTAGCACCGGGTCTCGACTTCTCA
AGACCCCTCTCGCCACGCACCCTCTCCAGTGAGCGCCGACGGCCACCTTCCCCTTCTGCGCCATCGTGCGGCCAGAGCTGAAGAAGT

1235



CCCTGTTCTTCTCGCTATCAGAGATCACGATACAGCCGGCGGTATCGATAATCTTGTGCGGTACTGGATGGTAAAGTCGGGCTCGGGCTTGT
GGGACAAGAAGGAGCGATAGTCTCTAGTGCTATGTCGGCCGCATAGCTATTAGAACAACGCCATGACCTACCATTTCAGCCCGAGCCGAACCTA

1330



GTCTTCTGTTTGTAGAGGGGACAGCATGATAGCGCGGGAGGCACGGCGGTTAATAATCACCTTAAAAGGACGCGTGGTTTTGCGCGGTTTCT
CAGAAGGACAACTACTCCCGTCTACTATCGCGCCCTCCGTGCCGCCAAATTATTAGTGAACCTTCTGCGCACAAAACGCGCCAAAGA

1425



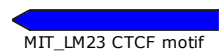
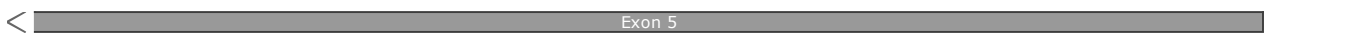
TACGCGGGCTGAGCTCGGGAGTAGCGGATGCCCGGGGAGAGGAGTGTAGTAACCGCGACGCTGGTGGGGTTCGGCTTGTAAAGAGGGGCGCTG
ATGCGCCCGACTCGAGCCCTCATCGCCTACGGGGCCCTCTCCTCACAATCATTGGCGCTGCGACACCCCGAGCCGAACAATTCTCCCGCGAC

1520



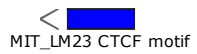
CTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGTACTGGAATCGATACCGGCATGATTGACAGCCTGGGCGAGGATGTCACCTGATGG
GATTGCGACGTTCTACCCAACAGTCGCACCCCGGCCACGATGACCTTAGCTATGGCCGTAATACTGTGCGACCCGCTCTACAGTGGACTACC

1615



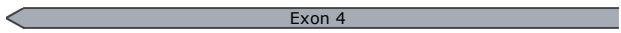
TGATAAGAAGACACGGGAGACTTAGTACGGTTTCACAGGCGTGACACGTTTATTGAGTAGGATTACAGAGTATAACATAGAGTATAATATAGAGT
ACTATTCTTCTGTGCCCTCTGAATCATGCCAAAGTGTCCGCACTGTGCAAACTCATCCTAATGTCTCATATTGTATCTCATATTATATCTCA

1710



ATACAATAGTGACGTGGGATCCATAACAGTAACTGATATATACAATAGTTTACTGGTCAGCCTTGCTTCTAGTCACCATAGGGTGGGTGCTCT
TATGTTATCACTGCACCCTAGGTATTGTCATTGACTATATATGTTATCAAAATGACCAGTCGGAACGAAGATCAGTGGTATCCACCCACGAGA

1805



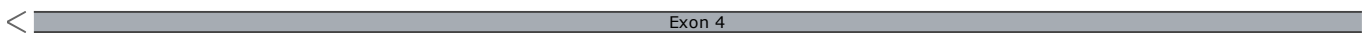
TGCCTCCAGAGGCGGTGGGTTCTCAGCACCATCCTCCTTCTCCTGAGGCAACTTCTCTATCTCAGACACTGGCTCAGACTTGACAGACACA
ACGGAGGTCTCCGCCACCCAAGGAGTCTGGTAGGAGGAGAAGGAGACTCCGTTGAAGGAGATAGAGTCTGTGACCGAGTCTGAACTGTCTGTGT

1900



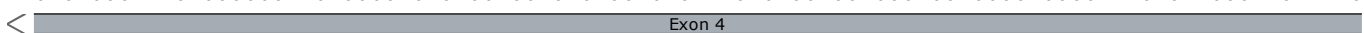
GTGTCCTCCCGCTCCTCCTGAGCACCTCCTCCTGTTCTCATCACTCTGTTCACTTTCTTCTGATCACTGTTCTCAGCCACAATTAAGTGGGA
CACAGGAGGGCGAGGAGGACTCGTGGGAGGAGGACAAGGAGTAGTGAGACAAGTGAAGAAGGACTAGTGACAAGAGTGGGTGTTAATGACTCCT

1995



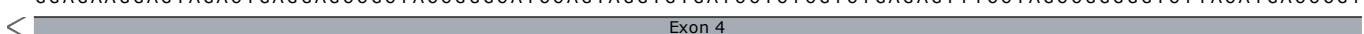
CAGAGGGATAGTCGCGGGTACAGGGGACTCTGGAGGTGACACCAGAGAATCAGAGGAGCTGGCACCGGGTGGCCAAAGTGTAGGCTACAATAG
GTCTCCCTATCAGCGCCCATGTCCCTGAGACCTCCTGTTGTTCTTAGTCTCCTCGACCGTGGTCCGCCACCGGTTTACATCCGATGTTATC

2090



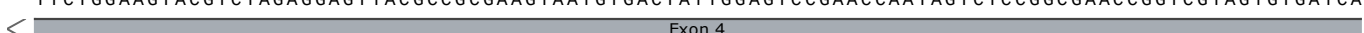
CCTTCTCCTCATCTGACTCCTCGCGATGGCCCGTAGGTCATCCACACTAGGAGAGCAGACTCTCAAAGGATCGGCCCCAGAATGTAAGTGGCA
GGAGAAAGGAGTAGACTGAGGAGCCGCTACCGGCATCCAGTAGGTGTGATCCTCTCGTCTGAGAGTTTCTAGCCGGGGTCTTACATGACCCGT

2185



AAGACCTTCATGCAGATCTCCTCAATGCGGGCTTCATTACACTGATAACCTCAGGCTTGGTTATCAGAGGCGGCTTGGCCAGCATCACACTAGT
TTCTGGAAGTACGTCTAGAGGAGTTACGCCGGAAGTAATGTGACTATTGGAGTCCGAACCAATAGTCTCCGGCGAACCGGTCGTAGTGTGATCA

2280



CTCCTTAAGACATAGCAGCACAGCACCCGACAGAACTCACTTAAGAGAGAGATGCCCCGTACATGGTCATCATAACAAGCGTCACTAGTGACCT
GAGGAGATTCTGTATCGTCGTGCTGTTGGCTGTCTTGAGTGAATTCTCTCTACGGGGGCATGTACCAGTAGTATGTTTCGAGTGATCACTGGA

2375



TGTA CT CATTACACATTGTTTCCACACATGTAGTGAGGATATCCATAAATATGTGATCAATGTGCGTGAGCACCTTGTCTCTCTCATCCAAA
ACATGAGTAATGTGTAACAAAGGTGTGTACATCACTCCTATAGGTATTTATACACTAGTTACACGCACTCGTGAACAGAGAGAGGAGTAGGTTT

2470



ATCTTAAAGATTTTCTGGGCATAAGACATAATCTCATCAGGAGAGCACTGAGGCAAGTTCTGCAATGCCGCCATGGCCTGACTGACGCCATTGGT
TAGAATTTCTAAAAGACCCGTATTCTGTATTAGAGTAGTCTCTCGTGACTCCGTTCAAGACGTTACGGCGGTACGGACTGACGTCGGTAACCA

2565



GGTCTTAGGGAAGGCTGAGTCTTGGTAAAGAACTCTATATTCCTGTAGCACATATACATCATCTTTCTCCTAAGTTCATCCTTTTTAGCACGGG
CCAGAATCCCTCCGACTCAAGAACCATTCTTGAGATATAAGGACATCGTGTATATGTAGTAGAAAGAGGATTCAAGTAGGAAAAATCGTGCC

2660



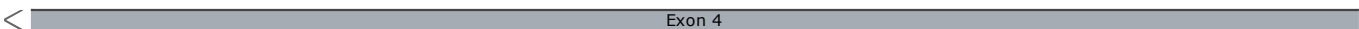
CCTTAGCCTGCAGTGCACCCCCAACTTGTAGCGGCGCCTTGCTCACATCATGCAGCTCCTTAATACAAGCCATCCACATCTCCGCTTATCC
GGAATCGGACGTACGTGGGGGGTGAACAATCGCCGCGGGAACGAGTGTAGTACGTGAGGAATTATGTTTCGGTAGGTGTAGAGGGCGAATAGG

2755



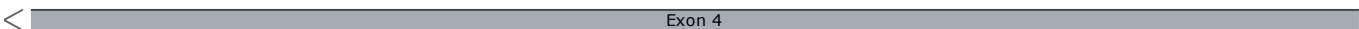
TCAGGTACAATGTAGTCTCATACATGCTCTGCATAGTTAGCCCAATACACTTCATGTCTCGAAAGGCTCATGAACCTTATCTAAGATATCTAA
AGTCCATGTTACATCAAGAGTATGTACGAGACGTATCAATCGGGTATGTGAAGTACAGGAGCTTCCGAGTACTTGAATAGATTCTATAGATT

2850



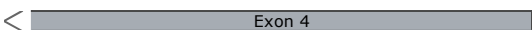
GGCATTCTGCAACATCCTCCCATCATATTAAGGCGCCAGTGAATTTTTCTTCGCTCGGGTATATTTTTTCAGCATGTGCTCCTTGATTCTAT
CCGTAAAGACGTTGTAGGAGGGTAGTATAATTTCCGCGGTCACTTAAAAAGAAGGCAGACCCATATAAAAAAGTCGTACACGAGGAAC TAAGATA

2945



GCCGCACCATGTCCACTCGAACCTTAATCTGTTTACTGTAGAGGAGGATAACAACACATATAAGTATCCGTCTCCTGACTCATTATCGCTAT
CGGCGTGGTACAGGTGAGCTTGAATTAGACAAACTGACATCTCCTCTATTGTTGTGTATATT CATAGGCAGGAGGACTGAGTAAATAGCGATA

3040



CTCGATGCCCGCTCACATGCAAGAGTTAATCTTCACTCTATCTGACATACACAAGTAAATCCACGTCCCATGCAGGTTAGTATATATCACATAC
GAGCTACGGGGCGAGTGTACGTTCTCAATTAGAAGTGAGATAGACTGTATGTGTT CATT TAGGTGCAGGGTACGTCCAATCATATATAGTGTATG

3135

ATGTCAACAGACTTACCGAGTCTGTGACAGGACATCTTTTTCGGGTTCTCGTTGCAATCCTCGGTCACTTGTTCAAAGGTTTTGAGGGATTCTTC
TACAGTTGCTGAATGGCTCAAGACAGTCTGTAGAAAAAGCCCCAAGAGCAACGTTAGGAGCCAGTGAACAAGTTTCCAAAACCTCCTAAGAAG

3230



GGCCAACTCTGGGAACAGCGGGTCTCCAGGCTCAGCTGACTGTTAACCTCCTTCTTAACATAGTCTGCAGGAACGTCGTGGCCTTGGTCACGG
CCGGTTAAGACCTTGTGCGCCAGAGGGTCCGAGTCACTGACAATTGGAGGAAGGAATTGTATCAGACGTCCTTGCAGCACCGGAACCAGTGCC

3325



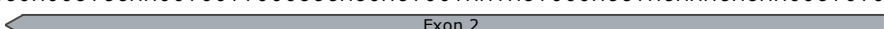
GTGTCTCGGGCCTAAACACATGATAAACAAGTCAAGCACATGGATCACATACAGGAAATATGTATATAACATTAAGATATAACTTTTTATT
CACAGAGCCCGGATTTGTTACTATTTGTTTCAGTATTCTGTGACTAGTGTATGTCCTTTATACATATATTGTAATTTCTATATTGAAAAATAA

3420



AAAAAAAGGGGAACAACAAGTCCCGACACGTACCGTGGCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAG
TTTTTTTCCCTTGTGTTACAGGCTGTGCATGGCACCGTGAACCTCCTTCCGGGAGCAGTCTTAATAGTCCAGGTAGAAAGAGAACCCTCTC

3515



GACTCCATCGTGTCAAGGACGGTGACTGCAGAAAGGATCCATGGAAAGGAACAGTCTGTTAGTCTGTCTGAGCTATTATGTCTGGTGGCGCGCGGG
CTGAGGTAGCACAGTTCTCTGCCACTGACGTCTTTCTAGGTACCTTTCTTGTCTGAGACAATCAGACAGTCTGATAATACAGACCACCGCGCGCGCC

3610

Exon 2

CAGCAACGAGTACTGCTCAGACTACACTGCCCTCCACCGTTAACAGCACCGCAACAGGAGTTACCTCTGACTCTCAACAGAACACAACCTCAGCTG
GTCGTTGCTCATGACGAGTCTGATGTGACGGGAGGTGGCAATTGTCGTGGCGTTGTCTCAATGGAGACTGAGAGTTGCTTGTGTTGAGTCGAC

3705

MIT_LM2 CTCF motif PUBLISHED

CCTGCTTCTTCTGCTGCTGCCTTAAGTCTTCCATCTGCGTCTGACGGTGAAGCCATTCCCGAGCTCATTTTCAGACACATACCCTACCGCCAC
GGACGAAGAAGACGACGACGGAATTCAGAAGGTAGACGCGAGTCGCCACGTTCCGGTAAGGGGCTCGAGTAAAAGTCTGTGTATGGGATGGCGGTG

3800

GGCCTTGTGCGGCACACTGGTGGTGGTGGGCATCGTCTGTGCCTAAGTCTGGCCTCCACTGTTAGGAGCAAGGAGCTGCCGAGCGACCATGAGC
CCGGAACACGCCGTGTGACCACCACCACCGTAGCACGACACGGATTAGACCGGAGGTGACAATCTCTGTTCTCGACGGCTCGCTGGTACTCG

3895

CGCTGGAGGCATGGGAGCAGGGCTCGGATGTGGAAGCTCCGCCGCTACCGGAGAAAGAGCCCATGTCCGGAACACGTACCCGAGATTCCGCTGGAG
GCGACCTCCGTACCCTCGTCCCGAGCCTACACCTTCGAGGCGGCGATGGCCTCTTCTCGGGTACAGGCCCTTGTGCATGGGCTCTAAGCGCACCTC

3990

ATCCACGCTATGTTTAAATAAAAACACTGCGGGTACTGGGGACGGTGTGTTGTATATGTGAATTTGTAATAATAAATGGGACCCCATCCTGTAAA
TAGGGTGCATACAAATTATTTTTGACGCCCATGACCCTGCCACAACAACATATACACTTAAACATTTATTATTTACCCTGGGGTAGGACATTT

4085

AATACAGAGTCCGTGTCAGTCTCTGAAGGACAGAGTATTGGCATATAGCCAATAAAGAGAGTTGTGGCAAAGAGCCATGTTATGGATTAGTAATG
TTATGTCTCAGGCACAGTCAGAGACTTCTGTCTCATAACCGTATATCGGTTATTTCTCTCAACACCGTTTCTCGGTACAATACCTAATCATTAC

4180

GAAAGTATCGCCACCAATAGAGGAGTGGTCAATAATGGTCAATAACCCACACCTATAGGCTAAGCTATACCATCACCTATAGCATAAGGAAGCGG
CTTTCATAGCGGTGGTTATCTCTCACCAGTTATTACCAGTTATTGGGTGTGGATATCCGATTGATATGGTAGTGGATATCGTATTCCTTCGCC

4275

GGGTGTATAGACCCCAAGCCAAAACAGTATAGCATGCATAAGAAGCCAAGGGGGTGGCCCTATAGAGTCTATAGGCGGTACTTACGCTCACTCTT
CCCACATATCTGGGGTTCGGTTTTTGTCTATCTGACGATATTCTTCGGTTCCCCACCCGGATATCTCAGATATCCGCCATGAATGCAGTGAGAA

4370

Exon 1

GGCACGGGAATCCGCGTTCCAATGCACCGTTCCCGGCCGCGGAGGCTGGATCGGTCCCGGTGCTTCTATGGAGGTCAAACAGCGTGGATGGC
CCGTGCCCTTAGGCGCAAGGTTACGTGGCAAGGGCCGGCGCTCCGACCTAGCCAGGGCCACAGAAGATACCTCCAGTTTTGTGCGACCTACCG

4465

Exon 1

GTCTCCAGGCGATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCCATAGTAAACGCCTACCGCCATTTGCGTCAATGGGGCGGAG
CAGAGGTCCGCTAGACTGCCAAGTGATTTGCTCGAGACGAATATATCTGGAGGGTATCATTGCGGATGGCGGGTAAACGCAGTTACCCCGCCTC

4560

Exon 1

transcription start site

TTATTACGACATTTTGGAAAGTCCCGTTGAATTTGGTGCCAAAACAACTCCCATGACGTCATGAGGTTGGAGACTTGGAAATCCCGTGAGTC
AATAATGCTGTAAACCTTTCAGGGCAACTTAAACCACGGTTTTGTTGAGGGTAACTGCAGTTACCCACCTCTGAACCTTATAGGGGCACTCAG

4655

AAACCGCTATCCACGCCATTGATGTACTGCCAAAACCGCATCACCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGT
TTTGGCGATAGGTGCGGGTAACTACATGACGGTTTTGGCGTAGTGGTACCATTATCGCTACTGATTATGCATCTACATGACGGTTCATCCTTCA

4750

CCCCTAAGGTCATGTACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCATAGGGGGCTACTTGGCATATGATACACTTGATGT
GGGCATTCCAGTACATGACCCTATTACGGTCCGCCGGTAAATGGCAGTAACTGCAGTTATCCCGCATGAACCGTATACTATGTGAACACTACA

4845

EMBL_M1 CTCF motif

