

```

1
1'-929603- CGGGAAACTTATGAATCGTCAATTCCCGCGCAGGCGGGAATCTGGAAATTCAATGCCTCAAGAATTTATCGGAAAAAACCAAAACCCCTTCGTCATCATTCCCGCAAAAGCGGGAATCTAGAAATGAAAAAG
3'-1072218- AGCCTTTTTA-----
5-1591795- AACCAAAAACCGCAGGTTTCAA-----ATAC
4'-1427636- GAAGGCGGG-----GGATGGGCGTGTAGGCTGCTTTTTCGT-----AACG
Consensus .....a.Cc...a..c..c....tc.....a...

131
1'-929603- CAACAGGAATTTATCGGAATGACCGAACTGAACG GACTGGATTCCCTCTTTTGC GGGAATGACGGCGACAGGGTTGCTGTTATAGTGGATGAACAAAAACCGGTACGGCGTTGTCTCGCCTTAGTTTCG
3'-1072218- ----AAAATATTCCGCAAGCAATCCAATG----CCGCCTGAAATC-----
5-1591795- CTGCGGTTTTCCTTACACAATAAACCAACG---CTTCACATATC-----
4'-1427636- TTCTGCGTTTTCAGACCGTCGTTTACTGGGTTTACTGACTGAAATG-----AAACCAAAGTGCAGGCTGCCGGAAAGACTTTTCTGTCTTTTAGG
Consensus .....ttTt....ca....a..caa.g....C.g.CtgaaaTc.....

261
1'-929603- AAGAGAACGATTCTCTAAGGTGCTGAAGCACCAAGTGAATCGGTTCCGTACTATCTGTACTTTCTGCGGCTTCGTCGTCTTGTCTGATTTTTGTTAATTAATACTATATCGACATCGCCAAACGAACTTC
3'-1072218- -----TCATAATGTTC-----
5-1591795- -----CCACACTCCT-----
4'-1427636- CAGCCTGCATTCTTTTTCATTCT-----
Consensus .....tC....t..t.c.....

391
1'-929603- GTCATCGCCGTTTCGTCTTTGTCTAAAACCAAAACCGAAACCAACCCCAAGGTATCGCCATACTATCGAATACCTTAAAAAACACAAAGGTCGAGACCTTTGCAAAATTCGCCAAATCCCCTAAA
3'-1072218- -----AGGCGGAACCTTTGCAAAATTCGCCAAATCCCCTAAA
5-1591795- -----ATCCCGAGACCTTTGCAAAATTCGCCAAATCCCCTAAA
4'-1427636- -----ATCTAGAGACCTTTGCAAAATTCGCCAAATCCCCTAAA
Consensus .....A....GAgACCTTTGCAAAATTCGCCAAATCCCCTAAA

521
1'-929603- TTCCACCAAGACATTTAGGGGATTTCTCATGAGCACCTTCTTTCAACAAACCGCCCAAGCCATGATTGCCAAACACATCGACCGCTTCCCCTATTGAAGTTGGACAGGTGATTGATTGGCAGCCGAT
3'-1072218- TTCCACCAAGACATTTAGGGGATTTCTCATGAGCACCTTCTTCCGCAAAACCGCACAAAGCCATATTGCCAAACACATCGACCGTTCCCCTATTGAAGTTGGACCGGGTGATTGATTGGCAGCCGAT
5-1591795- TTCCACCAAGACATTTAGGGGATTTTCCATGAGCACCTTCTTTAGCAAAACCGCACAAAGCCATGATTGCCAAACACATCGACCGTTCCCCTATTGAAGTTGGATCAGGTAATTGATTGGCAACCGAT
4'-1427636- TTCCACCAAGACATTTAGGGGATTTCTCAAGAGCACCTTCTTCCAGCAAAACCGCCCAAGCCATGATTGCCAAACACATCGACCGCTTCCCCTATTGAAGTTGGACCGGGTGATTGATTGGCAGCTGAT
Consensus TTCCACCAAGACATTTAGGGGATTTctCatGAGCACCTTCTT.CagCAAAACCGC.CAAGCCATgATTGCCAAACACATCGACCG.TTCCCgCTATTGAAGTTGGAcC.GGTgATTGATTGGCAGcCGAT

651
1'-929603- CGAACATACCTGAACCGTCAAAAACCCGTTACCTCCGAGACCCACGCGGTCGTCCCGCTGTCCCTGTGTGTCATGTTCAAAGCCGTCCTGCTCGGACAATGGCACAACTCTCCGATCCCGAACTC
3'-1072218- CGAACCATACCTGAACCGTCAAA-AAACCCGTTACCTCAGAGATTACCGCGGCGGTCCCGCTATCCCTGTGTGTCATGTTCAAAGCCGTCCTGCTCGGACAATGGCACAGCTCTCCGATCCCGAACTC
5-1591795- CGAACAGTACCTGAACCGTCAAGAACCCGTTACCTTCGAGACCCACGCGGCGGTCCCGCTATCCCTGTGTGTCATGTTCAAAGCCGTCCTGCTCGGACAATGGCACAGCTCTCCGATCCCGAACTC

```

4'-1427636 CGAACAATACCTGAACCGTCAAAAACCCCGTTACCTTAGAGACCACCGCGGCCGTCCTGCCTATCCCTGCTGTCCATGTTCAAAGCCGTCCTGCTCGGACAATGGCACAGCCTCTCCGATCCCGAACTC  
Consensus CGAACaaTACCTGAACCGTCAAA.AACCCGTTACCT..GAGAccACCGCGGCCGTCCcGCCTaTCCCTG.TGTCCATGTTCAAAGCCGTCCTGCTCGGACAATGGCACAGcCCTCTCCGATCCCGAACTC

781 910

1'-929603- GAACACAGCCTCATCACCCGCATCGACTTCAACCTGTTTTGCCGTTTTGACGAACTGAGCATCCCCGATTACAGCACCTTATGCCGCTACCGCAACCGGCTGGCGCAAGACGACACCCTGTCCGAACtGT  
3'-1072218 GAACACAGCCTCATCACCCGCATCGACTTCAACCTGTTTTGCCGTTTTGACGAACTGAGCATCCCCGATTACAGCACCTTATGCCGCTACCGCAACTGGTTGGCGCAAGACAAcACCCTGTCCGAATTAT  
5-1591795- GAACACAGCCTCATCACCCGCATCGATTTCACCTGTTTTGCCGTTTTGACGAACTGAGCATCCCCGATTACAGCACCTTATGCCGCTACCGCAACTGGCTGGCGCAAGACGACACCCTGTCCGAACtGT  
4'-1427636 GAACACAGCCTCATTACCCGCATCGATTTCACCTGTTTTGCCGTTTTGACGAACTGAGCATCCCCGATTACAGCACCTTATGCCGCTACCGCAACCGGCTGGCGCAAGACAATACCCTGTCTGAACtGT  
Consensus GAACACAGCCTCATcACCCGCATCGA.TTCAACCTGTTTTGCCGTTTTGACGAACTGAGCATcCCCCGATTACAGCACCTTATGCCGCTACCGCAAC.GGcTGGCGCAAGAC.AcACCCTGTCCGAACtGT

911 1040

1'-929603- TGGAACTGATTAACCGCCAACTGACCGAAAAAGGCTTAAAAGTAGAGAAAGCATCCGCCGCCGCTCgTTGACGCCACCATTATTAGACCCGTCCGGCAGCAAAACAGCGCCAGGCCATAGAAGTCGATGAAGA  
3'-1072218 TGGAACTGATCAACCGCCAACTGACCGAAAAAGGCTTAAAAGTAGAGAAAGCATCCGCCGCCGCTCATTGATGCCACCATTATTAGACCCGCCGGCAGCACACAGCGTCAGGTCATAGAAGTCGATGAAGA  
5-1591795- TGGAACTGATTAACCGCCAACTGACCGAAAAAGGCTTAAAAGTAGAGAAAGCATCCGCCGCCGCTCgTTGATGCCACCATTATTAGACCCGTCCGGCAGCAAAACAGCGTCAGGCCATAGAAGTCGATGAAGA  
4'-1427636 TGGAACTGATTAACCGCCAACTGACCGAAAAAGGTTTAAAAATAGAGAAAGCATCCGCTGCCGTCgTTGACGCCACCATTATTAGACCCGCCGGCAGCAAAACAGCGTCAGGCCATAGAAGTTGACGAAGA  
Consensus TGGAACTGATtAACcGCCAACTGACCGAAAAAGGcTTAAAAgtTAGAGAAAGCATCCGCCcGCCGTCgTTGA.GCCACCATTATTAGACCCcCGCAGCAaACAGCGtCAGGCCATAGAAGTCgatGAAGA

1041 1170

1'-929603- GGGACAAGTCAGCGGTCAAACACACTGAGTAAGGACAAAAATGCCCGTTGGACAAGAAAAACAGCCTCTACAACTCGGTTACAAACAACATACACGTACCGATGCGGAAGGCTATATCGAGAACTG  
3'-1072218 AGGACAAATCAACGGCCAAACACACCGAGTAAGGACAGCGATGCCCGTTGGATCAAGAAAAACGGCCTCTACAACTCGGTTACAAACAACATACCCGTACCGATGCGGAAGGCTATATCGAGAACTG  
5-1591795- AGGACAAGTCAGCGGCCAAACACACCGAGTAAGGACAGCGATGCCCGTTGGATCAAGAAAAACGGCCTCTACAACTCGGTTACAAACAACATACCCGTACCGATGCGGAAGGCTATATCGAGAACTG  
4'-1427636 AGGACAAATCAGCGGTCAAACACACCGAGTAAGGACAGCGATGCCCGTTGGATAAAGAAAAACGGCCTCTACAACTCGGTTACAAACAACATACCCGTACCGATGCGGAAGGCTATATCGAGAACTG  
Consensus aGGACAA.TCagCGG.CAAACACACcGAGTAAGGACAgcgATGCCCGTTGGAt.AAGAAAAcGcCCTCTACAACTCGGTTACAAACAACATAcCGTACCGATGCGgGAAGGCTATATCGAGAACTG

1171 1300

1'-929603- CACATTACTCCCGCCAATGCCCATGAGTGCAAACACCTGTCGCCGTTGTTGGAAGGGTTACCCGAAGGTACGACCATCTATGCCGACAAAGGCTACGACAGTGCGGAAAACCGGCAACATCTGGAAGAAC  
3'-1072218 CACATTACCCCGCCAATGCCCATGAGTGCAAACACCTGCCGCCTTTGTGGAAGGACTGCCCAAGGTACGACCgTCTATGCCGACAAAGGCTATGACAGTGCGGAAAACCGGCAACATCTGGAAGAAC  
5-1591795- CACATTACCCCGCCAATGCCCATGAGTGCAAACACCTGTCGCCGTTGTTGGAAGGGTTACCCGAAGGTACGACCgTCTATGCCGACAAAGGCTATGACAGTGCGGAAAACCGGCAACATCTGGAAGAAC  
4'-1427636 CACATTACCCCGCCAATGCCCATGAGTGCAAACACCTGTCGCCGTTGTTGGAAGGACTGCCCAAGGTACGACCgTCTATGCCGACAAAGGCTACGACAGTGCGGAAAACCGGCAACATCTGGAAGAAC  
Consensus CACATTACcCCCGCCAATGCCCATGAGTGCAAACACCTGtCGCCgTTGTTGGAAGG..T.CCC.AAGGTACGACCgTCTATGCCGACAAAGGCTA.GACAGTGCGGAAAACCGGCAACATCTGGAAGAAC

1301 1430

1'-929603- ATCAGTTGCAGGACGGCATTATGCGCAAAGCCTCGCGCAACCGTCCGCTGACGGAAACGCAAACCAAACGCAACCGGTATTTATCTAAGACCCGTTAT-----  
3'-1072218 ATCAGTTGTTGGACGGCATTATGCGCAAAGCCTCGCGCAACCGTCCGCTGACGGAAACGCAAACCAAACGTAACCGATATTTATCGAAGACCCGTTATAGTGGATTAAATTTAAATCAGGACAAGGCGAC  
5-1591795- ATCAGTTGCAGGACGGCATTATGCGCAAAGCCTCGCGCAACCGCCCGCTGTCGGAAGTGCAAACCAAGCGTAACCGATATTTATCGAAGACCCGTTAT-----  
4'-1427636 ATCAGTTGCTGGACGGCATTATGCGCAAAGCCcACCGCAACCGTCCGCTGACGGAAAGTGCAAACCAAACGCAACCGATATTTGTCGAAGACCCGTTAT-----  
Consensus ATCAGTTGc.GGACGGCATTATGCGCAAAGCCtgCGCGAACCGTCCGCTGaCGGAA..GCAAACCAAaCG.AACCGaTATTTaTcGAAGACCCGTTAT.....

1431 1560

1'-929603- -----GTGGTcGAACAAGCTTCGGTACGCTGCACCGTAAATTCCGCTAcGCcCGGGcAGCC  
3'-1072218 GAAGCCGCAGACAGTACAAATAGTACGGCAAGGCGAGGCAACGCCGTACTGGTTTAAATTTAATCCACTATATGTGGTcGAACAGAGCTTCGGTACGCTGCACCGTAAATTCCGCTATcGCGGGGcAGCC  
5-1591795- -----GTGGTcGAACAAGCTTCGGTACGCTGCACCGTAAATTCCGCTAcGCcCGGGcAGCC  
4'-1427636 -----GTGGTGAACAGAGCTTCGGTACGCTGCACCGTAAATTCCGCTAcGCcCGGGcAGCC  
Consensus .....GTGGTcGAACA.AGCTTCGGTACGCTGCACCGTAAATTCCGCTAcGCcCGGGcAGCC

1561 1690

1'-929603- TATTTCCGGA CTGATT-----TGC GCCCGCTGCCGCCTAAAAGGCAGCCCGGATGCCTGATTATCGGG

3'-1072218- TATTTCCGGA CTGATTAAAGTGAGTGC GCAAAGCCACCTGAAGGCGATGTGTTTGAACCTGTTGAAAGCGGCCAACAGGTTAAGTGC GCCCGCTGCCGCCTAAAAGGCAGCCCGGATGCCTGATTATCGGG

5-1591795- TATTTCCGGA CTGATTAAAGTGAGTGTGCAAAGCCATCTGAAGGCGATGTGTTTGAACCTGTTGAAAGCGGCCAACAGGCTAAGTGC GCCCTGTTGCCGCCTAAAAGGCAGCACGGATGCCTGATTATCGGG

4'-1427636- TATTTCCGGA CTGTAAAGTGAGTGC GCAAAGCCACCTGAAGGCGATGTGTTTGAACCTGTTGAAAGCGGCCAACAGGCTAAGTGC GCCCGCTGCCGCCTAAAAGGCAGCACGGATGCCTGATTATCGGG

Consensus TATTTCCGGA CTGaTTaaagtgagtg.gcaaagcca.ctgaaggcgatgtgtttgaacctgttgaagc.gccaacagg.taagTGC GCCCGcTGCCGCCTAAAAGGCAGC.CGGATGCCTGATTATCGGG

1691 1820

1'-929603- TATCCGGGGAGGATTAAGGGGGTATTTGGGTAA AATTAGGAGGTATTTGGGGCGAAAACAGCTGAAAACCTGTGTTTGGGTTTCGGCTGTCGGGAGGGAAAGGAATTTTGCAAAGGTCTCAAC-----

3'-1072218- TATCCGGGGAGGATTAAGGGGGTATTTGGGTAA AATTAGGCGGTATTTGGGGCGAAAACAGCCGAAAACCTGTGTTTGGGTTTCGGCTGTCGGGAGGGAAAGGAATTTTGCAAAGGTCTCA-----

5-1591795- TATCCGGGGAGGATTAAGGGGGCGTTTGGGTAGAAATTAGGAGATATTTGGGGCGAAAACAGCCGAAAACCTGTGTTTGGGTTTCGGCTGTCGGGAGGGAAAGGAATTTTGCAAAGGTCTCATCCTGTTAT

4'-1427636- TATCCGGGGAGGATTAAGGGGATATTTGGGTAGAAATTAGGCGGTATTTGGGGCGAAAACAGCCGAAAACCTGTGTTTGGGTTTCGGCTGTCGGG--GAAAAGGAATTTTGCAAAGGTCTCTT-----

Consensus TATCCGGGGAGGATTAAGGGGGtaTTTGGGTAA.AATTAGG.GgTATTTGGGGCGAAAACAGCcGAAAACCTGTGTTtGGGTTTCGGCTGTCGGGagGgAAAGGAATTTTGCAAAGGTCTCa.....

1821 1950

1'-929603- -----TTGAA CAAAAA GAACCGCCCGAATCAGGC-----

3'-1072218- ----ACAAAAACAGAAACCTAAAGTCC-----CGTCATTCCCGCGCAGGCGGGAATCCAGACC-----CCCAACGCG

5-1591795- TTTCACAAAAACAGAAACCAAAACAGCAACCTGAAATTCGTCATTCCCCACGAAAGTGGGAATCCAGTGCCTTGAGTTTCAGCTATTAGAAATAAATTTTGAAACTCTAATCCGCTCATTCCACAGAA

4'-1427636- -----TTCGTCATTCCCGCGAAGCGGGGAATCTAGAA TCT-----CGGACTTT

Consensus .....ttcgtCAttccgcg.a.gcggGAATC.aG.....C.....

1951 2080

1'-929603- -----GGT TTTGCTTTGTGCGGGAACCGTGG-----

3'-1072218- GCAGGAATC-----TATCGGAAATAACCGAAAC-----CGGACGAACCTAGATTCCCGCTTTCGCGGGAATGACGGCAGGGTGGTTTCAGT----TGCTCCCGATAAAATGCC

5-1591795- GTGGGAATCCAGGACGCAAAATCTCAAGAAACCGTTTACCAGATAAGTTCCGCACCGACAACTCTAGATTCTCGCTTCGCGGGGAATGACGAATCCATCCATACGGAAACCTGCATCCCGTCATTCCC

4'-1427636- CAGATAATC-----TTTGAATATTGCTGTTGTT-----CTAAGGTCTAGATTCCCGCTTTCGCGGGGAATGACGGTTCAGTTGCTACGGT-----

Consensus .....aatc.....t..a.....c.g.....ctaGaTTc.CgctttcGCGGgAAtGacGg.....t..t.c.g.....

2081 2210

1'-929603- -----GATTCGACTAAATT-----

3'-1072218- GCCATCTCAAGTCTCGTCAATTCCTTAAACAGAAAACCGAA-----

5-1591795- ACGAACCTGCATCCGTCATTCCCACGAAAGTGGGAATCCAGTTTTTGAGTTTCAGTCATTCCCGATAAATTGCCTTAGCATTGAATGTCTAGATTCCCGCCTGCGCGGGAATGACGGGATTGAGATT

4'-1427636- -----TATGTCCAGGTTCGGTTATGTTGGAATTCG-----

Consensus .....t..c...aTTC...taaa...g....c.....

2211 2340

1'-929603- -----

3'-1072218- -----ATCAGAAACCTAAAA-----TCCCGTCATTCCCGCGCAGGCGGGAATCTAGGTTT-----

5-1591795- GCGGCATTATCAGGAGCAACAGAAGCCGCTCTGCCGTCATTCCCACGAAAGTGGGAATCCAGTTTTTGAGTTTCAGTCATTCCCGATAAATTGCCTTAGCATTGAATGTCTAGATTCCCGCCTGCGCG

4'-1427636- -----

Consensus .....c.....

2341 2470

1'-929603- -----

3'-1072218- -----GTCGGCACGGAACTT-----ATCGGGTAAAACGGTTTCTTTAGATT-----TTACGTT-----

5-1591795- GGAATGACGAATCCATCCATACGGAACTTGCACCACGTCATTCACGAACTACATTCCGTCATTCCCACGAAAGTGGGAATCCAGTTTTTGTAGTTTCAGTCATTCACGATAAATTGCCTTAGCAATT

4'-1427636- -----GGAACTTA-----

Consensus .....ggaaac.t.....

2471 2600

1'-929603- -----TTATTTCATTGATTTAAATATAT-----TTATT

3'-1072218- -----CTAGATTCCCGCCTGCGCGGGAATGACGA-----TGAAAAGATTGTTGTC

5-1591795- GAATGTCCTAGATTCCCGCCTGCGCGGGAATGACGAATCCATCCGTACGAAACCTGCATCCCGTCATTCCCACGAACTACATTCCGTCATTCCCACGAAAGTGGGAATCCAGTTTTTGTAGTTTCAGTC

4'-1427636- -----TGAAATTGAAACCTTTGCAAAAT-----AGTC

Consensus .....T..aTTC..ccT..gc..aAT.....gTC

2601 2730

1'-929603- TCTTTATGAAATTTAATTTACCATAAAA-----CAGCC-----

3'-1072218- GCTTCGGATAAATTTTGTTCGGTGGGTTCTAGATTCCCGCCTGCGCGGGAATGACGGCGGC-----GGGTTTCTGT

5-1591795- ATTTCCGATAAATTTGCCTTAGCATGAAATGTCCTAGATTCCCGCCTGCGCGGGAATGACGAATCCATCCGTACGAAACCTGCACCACGTCATTCCCACGAAAGTGGGAATCCAGTTGCTTGTAGTTTCAGT

4'-1427636- TGTTAACGAATTTG---ACGCATAAAA-----TGCGCCAAAAAT-----TTT

Consensus ..TT..ga..AA..TT....t.gCaT..aa.....ctGCgc...aa.....t

2731 2860

1'-929603- -----ATATACA

3'-1072218- -TTTTCGATAAATA-----CACACAAACTAAATTTTCGTCAATCCCATAAAAAC

5-1591795- CATTTCCGATAAATTTGCCTTAGCATTGAATGTCTAGATTCCCGCCTGCGCGGGAATGACGAATTCATCCGTACGAAACCTGCACCACGTCATTCCCACGAACCTACATTCCGTCATTCCCACGAAGT

4'-1427636- CAATTGCCTAAAA-----CCTTCCTAATATTGAGCA-----AAAAGTA

Consensus ...tt.c...aaa.....c...c.a...t...ca.....aaAa...

2861 2990

1'-929603- AAAATCTTGGAGTAACTATTGCATTACACTATTAGAAGATGCGTCTG-----AAGTGTTTTCGGACGGC-----

3'-1072218- AGAAACCAAGTGA-GAATAACAATTCGTTGTAAACAATACTATTGTAAATTT-----TATTAATATATGTAAAAATCCCCCCCCCCCCGAAGCT-----T

5-1591795- GGAAATCCAGTCGTTGAGTTTCACTCATTTCCAATAAATGCTTAGTATTGAATGTCTGGATTCCCGCCTGCGCGGGAATGACGAATTCATCCGTACGAAACCTGCATCCGTCATTCCCACGAAGT

4'-1427636- GGAAAAATCAGAAAGTTTGCATTTTGAAAATGAGATTGAGCATAAAATTT-----TAGTAACCTATGTTATTGCAAAGGTC---TCATGCT---A

Consensus .gAA.....g..a.g..Tt.CA.T....t...a..A....C....t.tt.....a.t.a...atgt....C...c..C....a..c.....

2991 3120

1'-929603- -----ATTTCCCAAGCTGCCGGAAGGCTAAATGCC-----

3'-1072218- AAGAAATATAATTGTAAGCGTAACGATTATTTACGTTATGTTACCATA-----

5-1591795- GGGAAATCCAGTTTTTGTAGTTTCAGTCATTCCCGATAAATTTGCCTTAGCATTGAATGTCTAGATTCCCGCCTGCGCGGGAATGACGGCGGAAATCTTGTATTATTTGAATCAAAAAAACCTGCACCTTA

4'-1427636- AATATCTAGCAAAAAAGTGGGTCAAAATTCACGCACCTTTTCGTT-----

Consensus ...a.....tt....g.G...c....Att.ac...A..Tt.Cc.t.....

```

3121                                     3182
1'-929603- -----GCAAACAGACAAGGAGCAGCGA
3'-1072218 -----TCCGACTACAAATCCAAATTTTAGAG-----GTTTAACTA
5-1591795- ATCAGTTGGCGGTTTAGTCGACTTTGGGGTGCAGATCAAGCTTTCAGACGGTATTCCTT
4'-1427636 -----CGAACAAATCGAAT
Consensus .....acA.tC.aa.....

```

REP sequences are in green background