Supplementary Data

A Protocol for Seamless and Simultaneous Fusion of Multiple DNA Fragments via Homology-Annealing Extension Mediated Fusion PCR (HAEMF-PCR)

by

Xianyong Ma[1], Jie Tang[1], Yasin Oduk[1,2], Charles X Ma[3]*

[1]: EpigenoMax Therapeutics LLC, 101 College Street, New Haven, CT 06510.

[2]: Nanomedigene LLC, 101 College Street, New Haven, CT 06510.

[3]: Nuvance Health Medical Practice, 34 Maple St. Norwalk, CT 06856

*For correspondence: Xianyong Ma, Principal Investigator, EpigenoMax Therapeutics LLC, Suite 2#, BioLab, 101 College St, New Haven, CT 06510, USA. Email: xianyong.ma@epigenomax.com

1. Fragment sequences

Human TERT gene Promoter Sequence from (NCBI reference sequence NG_009265.1):

TTTCCTACTTGGCAGTACATCTACGTAACGGCATTCGTGGTGCCCGGAGCCCGACGCCCCGCGT
CCGGACCTGGAGGCAGCCCTGGGTCTCCGGATCAGGCCAGCGGCCAAAGGGTCGCCGCACGCAC
CTGTTCCCAGGGCCTCCACATCATGGCCCCTCCCTCGGGTTACCCCACAGCCTAGGCCGATTCG
ACCTCTCTCCGCTGGGGCCCTCGCTGGCGTCCCTGCACCCTGGGAGCGCGAGCGGCGCGCGGGC
GGGGAAGCGCGGCCCAGACCCCCGGGTCCGCCCGGAGCAGCTGCGCTGTCGGGGCCAGGCCGGG
CTCCCAGTGGATTCGCGGGCACAGACGCCCAGGACCGCGCTTCCCACGTGGCGGAGGGACTGGG
GACCCGGGCACCCGTCCTGCCCCTTCACCTTCCAGCTCCGCCTCCTCCGCGCGGACCCCGCCCC
GTCCCGAC

NSP4 (nonstructural protein 4, partial, synthetic sequence)

AGAGAGTTGCCCGTACTTGACAGCGCAGCGTTCAATGTTGAGTGTTTCAAGAAGTTCGCTTGTA
ATCAGGAATACTGGAAGACATTCAAGGAGAATCCCATTAGACTGACCACTGAGAACGTGACGAA
CTACATCACGAAGTTGAAAGGGCCTGAAGCAGCCGCATTCTTTGCG

SGP promoter+51 nt cis element

AGAGGACCTGTTATACACCTCTACGGCGGTCCTAGATTGGTGCGTTAATACACAGAATTCTGAT
TATAGCGCACTATTATAGCACC

miniCMV:

GGTAGGCGTGTACGGTGGGAGGCCTATATAAGCAGAGCT

Synthetic PLA2: UniProtKB/Swiss-Prot: Q9PVF1.1

ATGGCACGAACCCTGTGGATACTGGCCGTGCTGCTGGTGGGCGTAGAAGGATCTCTGGTGCAAT
TTGAGACCATGATCATGAAACTCGCCAAAAGGTCTGGCTTCTTCTGGTACTCTTTCTATGGTTG
TTATTGTGGTTGGGGCGGGCATGGACTGCCACAGGATCCTACAGATCGGTGTTGCTTCGTGCAC
GATTGTTGCTACGGCAAGGTCACCAACTGTAACCCTAAAACAGCTACATACTCCTACACAGAGG
AGAATGACGGGATTGTGTGTGGGGGTGATGATCCATGCAAAAAACAGGTCTGCGAATGTGATCG
AGTGGCAGCCATGTGTTTCCGAGACAACAAAGACACTTACGACGGCGATAAGTATTGGAAACTG
CCTCCTCAGAAGTGTCAGGAAGATCCTGAGCCATGTAGACCCCCCCTTCCTCCACCCAAT

Synthetic SVMP: UniProtKB/Swiss-Prot: U5PZ28.1

ATGATCCAGGTACTGCTGGTGACCATCTGCCTCGCCGCACTGCCATACCAGGGGAGCAGCATCA
TTCTGGAGAGCGGCAACGTGAACGACTACGAGGTGGTGTACCCCCGGAAGGTGACGGCTCTGCC
CAAGGGCGCCGGCCAGCCCAAGTACGAGGACGCCATGCAGTACGAGTTCAAAGTGAACGGCGAG
CCTGTGGTGCTGCACCTGGAGAAGAACAAGGGACTGTTTAGTAAGGACTACTCCGAAACTCACT
ACTCCAGCGACGGGCGGAAAATCACAACTAACCCCCCTGTGGAGGATCACTGCTACTACCACGG

GAGGATCGAGAATGACGCCGACTCTACTGGGTCCATCTCTGCCTGTAATGGCCTAAAGGGACAC
TTCAAGCTGCAGGGGGAGATGTATCTGATCGAGCCACTGAAACTGAGTGACAGCGAGGCCCATG
CCATCTACAAGTACGAGAACGTGGAAAAAGAAGACGAGGCCCCTAAAATGTGTGGCGTGACCGA
GACCAACTGGGAAAGCTACGAGCCTATTAAAAAGGCTAGCCAGAGCAACCTGACCCCCGAACAG
CAGAGGTTCAACCCCTTTAAGTACGTGGAGCTGGTGATCGTGGCCGATCACAGGATGTTCACTA
AATATAATGGGGATCTGGAAAAGATCAGGATTAAAATCTACGAGATCGTGAACATCCTGAACGA
GATGTTCAGGTATCTGTACATCAGGATCGCCCTGGTGGATCTCGAGATCTGGAGCAACCGCGAC
CTGATAAACGTGACCTCCGTGGCCGGCGACACCCTCGACTCATTCGGCAATTGGCGAGAAACCG
ACCTGCTCAAACGGAAAAGTCACGACAACGCACAGCTGCTGACAGGCATCGATTTCAACGGCAC
CACTATCGGAATTGCCTATATAGCCAGTATGTGCAATCCATACCTGTCCGTCGGCATCGTGCAG
GACCACTCCGAGATCAACTTTCTGATCGCTGTGACAATGGCCCACGAAATGGGCCATAACCTGG
GAATGCGCCACGACACCGACTACTGTACATGTGGGGGCTACAGCTGCATTATGTGCGCCGTGCT
GTCTGATCAGCCTTCTAAGTTTTTCTCCAACTGCTCCTACATTCAGTATGGCAAGTTCATCATG
AACCAGAACTCCCAGTGCATTCTGAACGAGCCCCTGGGCACCGATATCGTGTCTCCTCCCGTGT
GTGGGAACGAAATTCTGGAAGTGGGGGAAGAGTGCGATTGTGGCTGTCCTACCAATTGCCAGGA
TCCTTGCTGCAACGCCGCCACATGTAAGCAGTACTCCTGGGTGCAGTGCGAAAGCGGGGAGTGC
TGCGAGCAGTGTAGGTTCCGGGCCGCTGGCACAGTGTGCAGACGGGCCACTGACAACGACATGG
ACAACAGGTGTACAGGACAGAGCGCCGATTGCCCTTCAAATGGAAGACCCCCCCTTCCTCCACC
CAAT

Synthetic **SVSP:** UniProtKB/Swiss-Prot: [Q9PTU8.1]

ATGGTGCTGATTAGAGTGATTGCAAACCTGCTCATCCTGCAGCTGTCCAACGCCCAGAAGTCCT
CCGAGCTGGTGATCGGCGGCGATGAATGTAACATCACCGAACACAGATTTCTGGTGGAGATCTT
CAACTCCAGTGGGCTGTTTTGCGGCGGGACCCTGATCGACCAGGAGTGGGTGCTGAGCGCCGCC
CACTGCGACATGAGAAATATGAGAATCTACCTGGGCGTGCACAATGAGGGCGTGCAGCACGCAG
ACCAGCAGAGAAGGTTTGCCAGGGAGAAATTTTTCTGCCTGAGTAGCAGGAACTACACTAAATG
GGACAAGGACATTATGCTGATCAGACTGAATCGGCCAGTCAACAACAGCGAACACATCGCCCCA
CTGTCCCTGCCCAGCAACCCTCCCAGCGTTGGCTCCGTCTGTAGAATCATGGGCTGGGGCACAA
TCACAAGCCCCAACGCTACATTTCCAGATGTGCCCCACTGCGCCAACATCAACCTGTTTAATTA
CACCGTCTGCAGAGGCGCTCATGCAGGGCTGCCTGCCACCAGCAGGACCCTGTGTGCCGGCGTG
CTGCAAGGCGGGATCGACACTTGTGGCGGCGACTCTGGCGGGCCCCTGATCTGCAATGGGACCT
TCCAGGGCATCGTGTCTTGGGGGGGCCACCCCTGCGCCCAGCCTGGCGAGCCCGCTCTGTATAC
CAAAGTGTTTGATTACCTGCCTTGGATCCAGTCCATCATCGCTGGCAACACAACAGCCACCTGC
CCCCCCGGACCTCTTGGGCTGGCT

2. Primers used in the Paper

Table 1. Primers Used in the Paper

Name	Sequence (5'-3')	Overlap Length	Description
F01f	TTGGCAGTACATCTACGTAACGGCATTCGTGGTGCCCGGAGCCCG CA202		Forward primer for F1 fragment with Sna BI
F01r	**ACAGGGTTCGTGCCATGGTG**GCGGTGCTATAATAG	**20 bp**	Reverse primer for F1 fragment
F01r-25	**GCACGGCCAGTATCCACAGGGTTCGTGCCATGGTG****GCGGTGCTATAATAG**	**25 bp**	Reverse primer for F1 fragment (adjusted)
F02f	**CACCATGGCACGAACCCTGTGGATACTGGCCGTGC**	**25 bp**	Forward primer for F2 fragment
F02r-25	**AGGTCAGCAGGCTTCCCCTTCCTTC****ATTGGGTGGAGGAAGGGGGGGTCTAC**	**25 bp**	Reverse primer for F2 fragment
F02r	**CCCTTCCTTCATTGGGTGGA**GGAAGGGGGGGTCTACATGGCTCAGGATCTTCCTGACAC	**20 bp**	Reverse primer for F2 fragment
F03f-25	**GAAGGAAGGGGAAGCCTGCTGACCT**	**25 bp**	Forward primer for F3 fragment
F03f	**TCCACCCAATGAAGGAAGGG**GAAGCCTGCTGACCT	**20 bp**	Forward primer for F3 fragment
F03r-25	**GCTTGAGCAGGGAGAAGTTGGTGGC****ATTGGTGGA GG GGGGGTCTTCC**	**25 bp**	Reverse primer for F3 fragment
F04f-25	**GCCACCAACTTCTCCCTGCTCAAGC**	**25 bp**	Forword primer for fragment 4 amplification
F04r	TTAAGCCAGCCCAAGAGGTCCGGGG		Reverse Primer for For fragment 4 amplification
F04r-2	GAAAAAAGCTTGAATTGTTT**TT AAGCCAGCCCAAGAGGTCCGGGG**GGGCAGGTGGCTGTT GTGTT		Reverse Primer for For fragment 4 amplification
F1f-003-N	CTACGTAACGCGTCGGCATTCGTGGTGCCCGGAGCCCGACGCCCCGCGTCCGGACC		Forward Primer for CA401 fragment1 with MluI site
F1r-003	**CCATGGTGGCAGCTCTGCTT**ATATAGGCCTCCCAC	**20 bp**	Reverse Primer for CA401 F1 fragment
F2f(Tai)-003	**AAGCAGAGCTGCCACCATGG**CTTCTTTGCT	**20 bp**	Forward Primer for CA401 fragment2
F2r-004-3	**AGCAGGCTTCCCCTTCCTTCGCTGC**CCCCGCCGCCAGACCCGCCGCCGCC	25 bp	Reverse Primer for CA401 F2 fragment
F3f-004	**GCAGCGAAGGAAGGGGAAGCCTGCT**GACCTGC	25 bp	Forward Primer for CA401 F3 fragment
F3r-003	GAAAAAAGCTTGAATTGTTTAAACTTAAGCCAGCCCAAGAGG		Reverse Primer for CA401 plasmid with PmeI site

3. Efficacy Predication for HAEMF-PCR efficiency

Table1 Predicated fusion efficiency (Fn) based on HAEMF-PCR using the formula

Numer of fusion fragments	10 HAEMF-PCR cycles	15 HAEMF-PCR cycles
2	65.13%	79.50%
3	12.80%	18.40%
4	2%	3.00%
5	0.32%	0.48%
6	0.05%	0.08%

Table II Predicated fusion efficiency (Fn) based on HAEMF-PCR using the formula

Number of Cycles	2 fragments	3 fragments	4 fragments	5 fragments	6 fragments
5	41.00%	6.50%	1.02%	0.80%	0.03%
10	65.10%	12.80%	2.06%	1.60%	0.05%
15	79.50%	18.40%	3.11%	2.30%	0.08%

4. Plasmid Sequences used in the paper:

Ep-CA201

aaacaattcaagcttttttcaattctcgacctcgagacaaatggcagtattcatccacaattttaaaagaaaagggg
ggattgggggtacagtgcaggggaaagaatagtagacataatagcaacagacatacaaactaaagaattacaaaaa
caaattacaaaaattcaaaattttcgggtttattacagggacagcagagatccagtttggccgcggctcgaggggt
tggggttgcgccttttccaaggcagccctgggtttgcgcagggacgcggctgctctgggcgtggttccgggaaacgc
agcggcgccgaccctgggtctcgcacattcttcacgtccgttcgcagcgtcacccggatcttcgccgctacccttgt
gggccccccggcgacgcttcctgctccgcccctaagtcgggaaggttccttgcggttcgcggcgtgccggacgtgac
aaacggaagccgcacgtctcactagtaccctcgcagacggacagcgccagggagcaatggcagcgcgccgaccgcga
tgggctgtggccaatagcggctgctcagcagggcgcgccgagagcagcggccgggaaggggcggtgcgggaggcggg
gtgtggggcggtagtgtgggccctgttcctgcccgcgcggtgttccgcattctgcaagcctccggagcgcacgtcgg
cagtcggctccctcgttgaccgaatcaccgacctctctccccaggggatccaccggagcttaccatgggatcggcc
attgaacaagatggattgcacgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaaca
gacaatcggctgctctgatgccgccgtgttccggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacc
tgtccggtgccctgaatgaactgcaggacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgca
gctgtgctcgacgttgtcactgaagcgggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtc
atctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggcta

cctgcccattcgaccaccaagcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcag
gatgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacgg
cgatgatctcgtcgtgacccatggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattca
tcgactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctacccgtgatattgctgaagagctt
ggcggcgaatgggctgaccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcatcgccttctatcg
ccttcttgacgagttcttctgacgcccgccccacgacccgcagcgcccgaccgaaaggagcgcacgaccccatgcat
cgttaagagctcggtacctttaagaccaatgacttacaaggcagctgtagatcttagccacttttaaaagaaaagg
ggggactggaagggctaattcactcccaacgaagacaagatctgcttttgcttgtactgggtctctctggttagac
cagatctgagcctgggagctctctggctaactagggaacccactgcttaagcctcaataaagcttgccttgagtgct
tcaagtagtgtgtgcccgtctgttgtgtgactctggtaactagagatccctcagacccttttagtcagtgtggaaaa
tctctagcagtagtagttcatgtcatcttattattcagtatttataacttgcaaagaaatgaatatcagagagtgag
aggaacttgtttattgcagcttataatggttacaaataaagcaatagcatcacaaatttcacaaataaagcatttt
ttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgtctggctctagctatcccgcccct
aactccgcccatcccgcccctaactccgcccagttccgcccattctccgccccatggctgactaatttttttttatt
atgcagaggccgaggccgcctcggcctctgagctattccagaagtagtgaggaggctttttggaggcctagggacg
tacccaattcgccctatagtgagtcgtattacgcgcgctcactggccgtcgttttacaacgtcgtgactgggaaaac
cctggcgttacccaacttaatcgccttgcagcacatccccctttcgccagctggcgtaatagcgaagaggcccgcac
cgatcgcccttcccaacagttgcgcagcctgaatggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgg
gtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcc
tttctcgccacgttcgccggctttccccgtcaagctctaaatcgggggctccctttagggttccgatttagtgcttt
acggcacctcgaccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttc
gccctttgacgttggagtccacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcg
gtctattcttttgatttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatt
taacgcgaattttaacaaaatattaacgcttacaatttaggtggcacttttcggggaaatgtgcgcggaacccctat
ttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatatt
gaaaaaggaagagtatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtt
tttgctcacccagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaact
ggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatgagcacttttaaagttc
tgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaat
gacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgc
cataaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttt
tgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgag
cgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttc
ccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggct
ggtttattgctgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaag
ccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagat
aggtgcctcactgattaagcattggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttc
atttttaatttaaaaggatctaggtgaagatccttttttgataatctcatgaccaaaatcccttaacgtgagttttcg
ttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctg
cttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaagg
taactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaac
tctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtct
taccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagc
ccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaa
gggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaa
cgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcagggg
ggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatg
ttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccgcag
ccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgccctttgagtgagctgataccgctcgccg

cagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaaccgcctctccccg
cgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaaaa
tgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggctcgtatgttgtgtggaattgtga
gcggataacaatttcacacaggaaacagctatgaccatgattacgccaagcgcgcaattaaccctcactaaagggaa
caaaagctggagctgcaagcttaatgtagtcttatgcaatactcttgtagtcttgcaacatggtaacgatgagttag
caacatgccttacaaggagagaaaaagcaccgtgcatgccgattggtggaagtaaggtggtacgatcgtgccttatt
aggaaggcaacagacgggtctgacatggattggacgaaccactgaattgccgcattgcagagatattgtatttaagt
gcctagctcgatacataaacgggtctctctggttagaccagatctgagcctgggagctctctggctaactagggaac
ccactgcttaagcctcaataaagcttgccttgagtgcttcaagtagtgtgtgcccgtctgttgtgtgactctggtaa
ctagagatccctcagacccttttagtcagtgtggaaaatctctagcagtggcgcccgaacagggacttgaaagcgaa
agggaaaccagaggagctctctcgacgcaggactcggcttgctgaagcgcgcacggcaagaggcgaggggcggcgac
tggtgagtacgccaaaaattttgactagcggaggctagaaggagagagatgggtgcgagagcgtcagtattaagcgg
gggagaattagatcgcgatgggaaaaaattcggttaaggccaggggggaaagaaaaaatataaattaaaacatatagt
atgggcaagcagggagctagaacgattcgcagttaatcctggcctgttagaaacatcagaaggctgtagacaaatac
tgggacagctacaaccatcccttcagacaggatcagaagaacttagatcattatataatacagtagcaaccctctat
tgtgtgcatcaaaggatagagataaaagacaccaaggaagctttagacaagatagaggaagagcaaaacaaaagtaa
gaccaccgcacagcaagcggccggtgatcttcagacctggaggaggagatatgagggacaattggagaagtgaatta
tataaatataaagtagtaaaaattgaaccattaggagtagcacccaccaaggcaaagagaagagtggtgcagagaga
aaaaagagcagtgggaataggagctttgttccttgggttcttgggagcagcaggaagcactatgggcgcagcgtcaa
tgacgctgacggtacaggccagacaattattgtctggtatagtgcagcagcagaacaatttgctgagggctattgag
gcgcaacagcatctgttgcaactcacagtctggggcatcaagcagctccaggcaagaatcctggctgtggaaagata
cctaaaggatcaacagctcctggggatttggggttgctctggaaaactcatttgcaccactgctgtgccttggaatg
ctagttggagtaataaatctctggaacagatttggaatcacacgacctggatggagtgggacagagaaattaacaat
tacacaagcttaatacactccttaattgaagaatcgcaaaaccagcaagaaaagaatgaacaagaattattggaatt
agataaatgggcaagtttgtggaattggtttaacataacaaattggctgtggtatataaaattattcataatgatag
taggaggcttggtaggtttaagaatagtttttgctgtactttctatagtgaatagagttaggcagggatattcacca
ttatcgtttcagacccacctcccaaccccgaggggacccgacaggcccgaaggaatagaagaagaaggtggagagag
agacagagacagatccattcgattagtgaacggatctcgacggtatcgatactagtattatgcccagtacatgacct
tatgggactttcctacttggcagtacatctacgtaacggcattcgtggtgcccggagcccgacgccccgcgtccgga
cctggaggcagccctgggtctccggatcaggccagcggccaaagggtcgccgcacgcacctgttcccagggcctcca
catcatggcccctccctcgggttaccccacagcctaggccgattcgacctctctccgctggggccctcgctggcgtc
cctgcaccctgggagcgcgagcggcgcgcgggcggggaagcgcggcccagacccccgggtccgcccggagcagctgc
gctgtcggggccaggccgggctcccagtggattcgcgggcacagacgcccaggaccgcgcttcccacgtggcggagg
gactggggacccgggcacccgtcctgccccttcaccttccagctccgcctcctccgcgcggaccccgccccgtcccg
acggtaggcgtgtacggtgggaggcctatataagcagagctagagagttgcccgtacttgacagcgcagcgttcaat
gttgagtgtttcaagaagttcgcttgtaatcaggaatactggaagacattcaaggagaatcccattagactgaccac
tgagaacgtgacgaactacatcacgaagttgaaagggcctgaagcagccgcattctttgcgagaggacctgttatac
acctctacggcggtcctagattggtgcgttaatacacagaattctgattatagcgcactattatagcaccgccacca
tggcacgaaccctgtggatactggccgtgctgctggtgggcgtagaaggatctctggtgcaatttgagaccatgatc
atgaaactcgccaaaaggtctggcttcttctggtactctttctatggttgttattgtggttggggcgggcatggact
gccacaggatcctacagatcggtgttgcttcgtgcacgattgttgctacggcaaggtcaccaactgtaaccctaaaa
cagctacatactcctacacagaggagaatgacgggattgtgtgtgggggtgatgatccatgcaaaaaacaggtctgc
gaatgtgatcgagtggcagccatgtgtttccgagacaacaaagacacttacgacggcgataagtattggaaactgcc
tcctcagaagtgtcaggaagatcctgagccatgtagacccccccttcctccacccaatgaaggaaggggaagcctgc
tgacctgcggcgacgtggaggagaatcccggcccaatgatccaggtactgctggtgaccatctgcctcgccgcactg
ccataccaggggagcagcatcattctggagagcggcaacgtgaacgactacgaggtggtgtacccccggaaggtgac
ggctctgcccaagggcgccggccagcccaagtacgaggacgccatgcagtacgagttcaaagtgaacggcgagcctg
tggtgctgcacctggagaagaacaagggactgtttagtaaggactactccgaaactcactactccagcgacgggcgg
aaaatcacaactaaccccccctgtggaggatcactgctactaccacgggaggatcgagaatgacgccgactctactgg

gtccatctctgcctgtaatggcctaaagggacacttcaagctgcaggggagatgtatctgatcgagccactgaaac
tgagtgacagcgaggcccatgccatctacaagtacgagaacgtggaaaaagaagacgaggcccctaaaatgtgtggc
gtgaccgagaccaactgggaaagctacgagcctattaaaaaggctagccagagcaacctgacccccgaacagcagag
gttcaaccccttaagtacgtggagctggtgatcgtggccgatcacaggatgttcactaaatataatggggatctgg
aaaagatcaggattaaaatctacgagatcgtgaacatcctgaacgagatgttcaggtatctgtacatcaggatcgcc
ctggtggatctcgagatctggagcaaccgcgacctgataaacgtgacctccgtggccggcgacaccctcgactcatt
cggcaattggcgagaaaccgacctgctcaaacggaaaagtcacgacaacgcacagctgctgacaggcatcgatttca
acggcaccactatcggaattgcctatatagccagtatgtgcaatccatacctgtccgtcggcatcgtgcaggaccac
tccgagatcaactttctgatcgctgtgacaatggcccacgaaatgggccataacctgggaatgcgccacgacaccga
ctactgtacatgtgggggctacagctgcattatgtgcgccgtgctgtctgatcagccttctaagtttttctccaact
gctcctacattcagtatggcaagttcatcatgaaccagaactcccagtgcattctgaacgagcccctgggcaccgat
atcgtgtctcctcccgtgtgtgggaacgaaattctggaagtgggggaagagtgcgattgtggctgtcctaccaattg
ccaggatccttgctgcaacgccgccacatgtaagcagtactcctgggtgcagtgcgaaagcggggagtgctgcgagc
agtgtaggttccgggccgctggcacagtgtgcagacgggccactgacaacgacatggacaacaggtgtacaggacag
agcgccgattgcccttcaaatggaagacccccccttcctccacccaatgccaccaacttctccctgctcaagcaggc
cggagacgtggaagagaaccccggacccatggtgctgattagagtgattgcaaacctgctcatcctgcagctgtcca
acgcccagaagtcctccgagctggtgatcggcggcgatgaatgtaacatcaccgaacacagatttctggtggagatc
ttcaactccagtgggctgttttgcggcgggaccctgatcgaccaggagtgggtgctgagcgccgcccactgcgacat
gagaaatatgagaatctacctgggcgtgcacaatgagggcgtgcagcacgcagaccagcagagaaggtttgccaggg
agaaatttttctgcctgagtagcaggaactacactaaatgggacaaggacattatgctgatcagactgaatcggcca
gtcaacaacagcgaacacatcgccccactgtccctgcccagcaaccctcccagcgttggctccgtctgtagaatcat
gggctggggcacaatcacaagccccaacgctacatttccagatgtgccccactgcgccaacatcaacctgtttaatt
acaccgtctgcagaggcgctcatgcagggctgcctgccaccagcaggaccctgtgtgccggcgtgctgcaaggcggg
atcgacacttgtggcggcgactctggcgggcccctgatctgcaatgggaccttccagggcatcgtgtcttgggggggg
ccacccctgcgcccagcctggcgagcccgctctgtataccaaagtgtttgattacctgccttggatccagtccatca
tcgctggcaacacaacagccacctgccccccggacctcttgggctggcttaa

Ep-CA401

AAACAATTCAAGCTTTTTTCAATTCTCGACCTCGAGACAAATGGCAGTATTCATCCACAATTTTAAAAGAAAAGGGG
GGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAA
CAAATTACAAAAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCAGTTTGGCCGCGGCTCGAGGGGGT
TGGGGTTGCGCCTTTTCCAAGGCAGCCCTGGGTTTGCGCAGGGACGCGGCTGCTCTGGGCGTGGTTCCGGGAAACGC
AGCGGCGCCGACCCTGGGTCTCGCACATTCTTCACGTCCGTTCGCAGCGTCACCCGGATCTTCGCCGCTACCCTTGT
GGGCCCCCCGGCGACGCTTCCTGCTCCGCCCCTAAGTCGGGAAGGTTCCTTGCGGTTCGCGGCGTGCCGGACGTGAC
AAACGGAAGCCGCACGTCTCACTAGTACCCTCGCAGACGGACAGCGCCAGGGAGCAATGGCAGCGCGCCGACCGCGA
TGGGCTGTGGCCAATAGCGGCTGCTCAGCAGGGCGCGCCGAGAGCAGCGGCCGGGAAGGGGCGGTGCGGGAGGCGGG
GTGTGGGGCGGTAGTGTGGGCCCTGTTCCTGCCCGCGCGGTGTTCCGCATTCTGCAAGCCTCCGGAGCGCACGTCGG
CAGTCGGCTCCCTCGTTGACCGAATCACCGACCTCTCTCCCCAGGGGGATCCACCGGAGCTTACCATGGGATCGGCC
ATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACA
GACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACC
TGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCA
GCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTC
ATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTA
CCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAG
GATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGG
CGATGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCA
TCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTT

GGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCG
CCTTCTTGACGAGTTCTTCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGCAT
CGTTAAGAGCTCGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGG
GGGGACTGGAAGGGCTAATTCACTCCCAACGAAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGAC
CAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCT
TCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAA
TCTCTAGCAGTAGTAGTTCATGTCATCTTATTATTCAGTATTTATAACTTGCAAAGAAATGAATATCAGAGAGTGAG
AGGAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTT
TTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGGCTCTAGCTATCCCGCCCCT
AACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTT
ATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGGACG
TACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
CCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCAC
CGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGG
GTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCC
TTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTT
ACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTC
GCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCG
GTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATT
TAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTAT
TTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATT
GAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTT
TTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACT
GGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTC
TGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAAT
GACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGC
CATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTT
TGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAG
CGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTC
CCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCT
GGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAG
CCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGAT
AGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTC
ATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCG
TTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTG
CTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGG
TAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAAC
TCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
TACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGC
CCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAA
GGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAA
CGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG
GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATG
TTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAG
CCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCTTTGAGTGAGCTGATACCGCTCGCCG
CAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCG
CGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAAAA
TGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGA
GCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAA

CAAAAGCTGGAGCTGCAAGCTTAATGTAGTCTTATGCAATACTCTTGTAGTCTTGCAACATGGTAACGATGAGTTAG
CAACATGCCTTACAAGGAGAGAAAAAGCACCGTGCATGCCGATTGGTGGAAGTAAGGTGGTACGATCGTGCCTTATT
AGGAAGGCAACAGACGGGTCTGACATGGATTGGACGAACCACTGAATTGCCGCATTGCAGAGATATTGTATTTAAGT
GCCTAGCTCGATACATAAACGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAAC
CCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAA
CTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGCGAA
AGGGAAACCAGAGGAGCTCTCTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGAC
TGGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAAGCGG
GGGAGAATTAGATCGCGATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGT
ATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATAC
TGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTAT
TGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAA
GACCACCGCACAGCAAGCGGCCGGTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTA
TATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAGA
AAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCGTCAA
TGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAG
GCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATA
CCTAAAGGATCAACAGCTCCTGGGGATTTGGGGTTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATG
CTAGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAAT
TACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTGGAATT
AGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATAG
TAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATAGAGTTAGGCAGGGATATTCACCA
TTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAG
AGACAGAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGATACTAGTATTATGCCCAGTACATGACCT
TATGGGACTTTCCTACTTGGCAGTACATCTACGTAaCgcgtcggcattcgtggtgcccggagcccgacgccccgcgt
ccggacctggaggcagccctgggtctccggatcaggccagcggccaaagggtcgccgcacgcacctgttcccagggc
ctccacatcatggcccctccctcgggttaccccacagcctaggccgattcgacctctctccgctggggccctcgctg
gcgtccctgcaccctgggagcgcgagcggcgcgcgggcggggaagcgcggcccagacccccgggtccgcccggagca
gctgcgctgtcggggccaggccgggctcccagtggattcgcgggcacagacgcccaggaccgcgcttcccacgtggc
ggagggactggggacccgggcacccgtcctgccccttcaccttccagctccgcctcctccgcgcggaccccgccccg
tcccgacggtaggcgtgtacggtgggaggcctatataagcagagctGCCACCATGGCTTCTTTGCTCAATTTTGCTA
ACCTTATAGAGTGTGCTAATCATGGCACGCGCTCAGCACTCGCATACGCTGACTATGGATGTTATTGTGGTAAGGGA
GGAAGAGGAACCCCTCTCGACGACCTGGACCGCTGCTGCCATGTTCATGATGACTGCTATGGAGAGGCCGAGAAGCT
CCCCGCCTGTAATTATCTTATGTCTAGCCCTTACTTCAACTCCTATTCCTACAAGTGCAACGAAGGCAAAGTAACGT
GTACTGACGATAACGACGAGTGTAAGGCCTTTATATGTAACTGTGACCGAACGGCCGCGATATGTTTTGCTGGCGCG
ACTTACAACGATGAGAACTTTATGATTAGCAAGAAATATAACGACATCTGCCAGAGACCCCCCCTTCCTCCACCCCC
CAACGGCAGCGGGCCGCTGGGACTGTGGGCTGGCTCTGGCGGCGGCGGCGGGAGCGGCGGCGGCGGGTCTGGCGGCG
GGGGCAGCgaaggaaggggaagcctgctgacctgcggcgacgtggaggagaatcccggcccaatgatccaggtactg
ctggtgaccatctgcctcgccgcactgccataccaggggagcagcatcattctggagagcggcaacgtgaacgacta
cgaggtggtgtacccccggaaggtgacggctctgcccaagggcgccggccagcccaagtacgaggacgccatgcagt
acgagttcaaagtgaacggcgagcctgtggtgctgcacctggagaagaacaagggactgtttagtaaggactactcc
gaaactcactactccagcgacgggcggaaaatcacaactaaccccctgtggaggatcactgctactaccacgggag
gatcgagaatgacgccgactctactgggtccatctctgcctgtaatggcctaaagggacacttcaagctgcagggggg
agatgtatctgatcgagccactgaaactgagtgacagcgaggcccatgccatctacaagtacgagaacgtggaaaaa
gaagacgaggcccctaaaatgtgtggcgtgaccgagaccaactgggaaagctacgagcctattaaaaaggctagcca
gagcaacctgacccccgaacagcagaggttcaacccctttaagtacgtggagctggtgatcgtggccgatcacagga
tgttcactaaatataatggggatctggaaaagatcaggattaaaatctacgagatcgtgaacatcctgaacgagatg
ttcaggtatctgtacatcaggatcgccctggtggatctcgagatctggagcaaccgcgacctgataaacgtgacctc
cgtggccggcgacaccctcgactcattcggcaattggcgagaaaccgacctgctcaaacggaaaagtcacgacaacg

cacagctgctgacaggcatcgatttcaacggcaccactatcggaattgcctatatagccagtatgtgcaatccatac
ctgtccgtcggcatcgtgcaggaccactccgagatcaactttctgatcgctgtgacaatggcccacgaaatgggcca
taacctgggaatgcgccacgacaccgactactgtacatgtgggggctacagctgcattatgtgcgccgtgctgtctg
atcagccttctaagtttttctccaactgctcctacattcagtatggcaagttcatcatgaaccagaactcccagtgc
attctgaacgagcccctgggcaccgatatcgtgtctcctcccgtgtgtgggaacgaaattctggaagtgggggaaga
gtgcgattgtggctgtcctaccaattgccaggatccttgctgcaacgccgccacatgtaagcagtactcctgggtgc
agtgcgaaagcggggagtgctgcgagcagtgtaggttccgggccgctggcacagtgtgcagacgggccactgacaac
gacatggacaacaggtgtacaggacagagcgccgattgcccttcaaatggaggacctcttgggctggcttaaGTTTA
AACGGCCGC