

What is claimed is:

1. An isolated HPV16 E7₁₁₋₁₉ (YMLDLQPET) peptide-MHC (HLA-A*02) (pMHC) complex binding protein comprising:

a) a T cell receptor (TCR) alpha chain variable domain comprising CDR1 having the amino acid sequence of TTLSN, CDR2 having the amino acid sequence of LVKSGEV, and CDR3 having the amino acid sequence of CAGIGSSNTGKLIF; and

b) a TCR beta chain variable domain comprising CDR1 having the amino acid sequence of SNHLYCDR2 having the amino acid sequence of FYNNEI, and CDR3 having the amino acid sequence of TGRVSYEQYF.

2. The binding protein of claim 1 comprising:

a) a TCR alpha chain variable (V_{α}) domain comprising a sequence with at least about 80% identity to the amino acid sequence of
MLLITSMLVLWMQLSQVNGQQVMQIPQYQHVEGEDFTTYCNSSTTLSNIQWYKQRPG
GHPVFLIQLVKSGEVKKQKRLTFQFGEAKKNSSLHITATQTDDVGTYFCAGIGSSNTGKL
IFGQGTTLQVKPD; and

b) a TCR beta chain variable (V_{β}) domain comprising a sequence with at least about 80% identity to the amino acid sequence of
MDTWLVCWAIFSLLKAGLTEPEVTQTPSHQVTQMGQEVILRCVPISNHLFYFWYRQILG
QKVEFLVSFYNNEISEKSEIFDDQFSVERPDGSNFTLKIRSTKLEDSAMYFCAITGRVSYE
QYFGPGTRLTVTE.

3. The binding protein of claim 2 comprising:

a) a TCR alpha chain variable (V_{α}) domain comprising the amino acid sequence of
MLLITSMLVLWMQLSQVNGQQVMQIPQYQHVEGEDFTTYCNSSTTLSNIQWYKQRPG
GHPVFLIQLVKSGEVKKQKRLTFQFGEAKKNSSLHITATQTDDVGTYFCAGIGSSNTGKL
IFGQGTTLQVKPD; and

b) a TCR beta chain variable (V_{β}) domain comprising the amino acid sequence of
MDTWLVCWAIFSLLKAGLTEPEVTQTPSHQVTQMGQEVILRCVPISNHLFYFWYRQILG
QKVEFLVSFYNNEISEKSEIFDDQFSVERPDGSNFTLKIRSTKLEDSAMYFCAITGRVSYE
QYFGPGTRLTVTE.

4. The binding protein of any one of claims 1-3 comprising:

1) a) a TCR alpha chain comprising a sequence with at least about 80% identity to the amino acid sequence of

MLLITSMLVLWMQLSQVNGQQVMQIPQYQHVEGEDFTTYCNSSTTLSNIQWYKQRPG
GHPVFLIQLVKSSEVKKQKRLTFQFGEAKKNSSLHITATQTTDVGTYFCAGIGSSNTGKL
IFGQGTTLQVKPDiqnppdavyqlrdsksdksvclftdfdsqtnvsqskdsdvyyitdktvldmrsmdfksnsavawsnks
dfacanafnnsiipedtffpsesscdvklveksfetdtnlnfqnlsvigfrilllkvagfnllmtlrlwss; and

b) a TCR beta chain comprising a sequence with at least about 80% identity to the amino acid sequence of

MDTWLVCWAIFSLLKAGLTEPEVTQTPSHQVTQMGQEVLRCVPISNHL YFYWYRQILG
QKVEFLVSFYNNEISEKSEIFDDQFSVERPDGSNFTLKIRSTKLEDSAMYFCAITGRVSYE
QYFGPGTRLTVTEdlnkvfppevavfepseaeishtqkatlvclatgffpdhvelswwwngkevhsqsgvstdpqplkeqpall
dsryclssrlrvsatfwqnpnrnhfrcqvfyglsendewtqdrakpvtqivsaeawgradcgftsvsyqqgvlstilyeillgkatlyavl
vsalvlmamvkrkdf, or

2) a) a TCR alpha chain comprising a sequence with at least about 80% identity to the amino acid sequence of

MLLITSMLVLWMQLSQVNGQQVMQIPQYQHVEGEDFTTYCNSSTTLSNIQWYKQRPG
GHPVFLIQLVKSSEVKKQKRLTFQFGEAKKNSSLHITATQTTDVGTYFCAGIGSSNTGKL
IFGQGTTLQVKPDiqnppdavyqlrdsksdksvclftdfdsqtnvsqskdsdvyyitdktvldmrsmdfksnsavawsnks
dfacanafnnsiipedtffpsdvpdcvklveksfetdtnlnfqnlvivrilllkvagfnllmtlrlws; and

b) a TCR alpha chain comprising a sequence with at least about 80% identity to the amino acid sequence of

MDTWLVCWAIFSLLKAGLTEPEVTQTPSHQVTQMGQEVLRCVPISNHL YFYWYRQILG
QKVEFLVSFYNNEISEKSEIFDDQFSVERPDGSNFTLKIRSTKLEDSAMYFCAITGRVSYE
QYFGPGTRLTVTEdlnkvfppevavfepskaeiahtqkatlvclatgffpdhvelswwwngkevhsqsgvstdpqplkeqpall
ndsryclssrlrvsatfwqnpnrnhfrcqvfyglsendewtqdrakpvtqivsaeawgradcgitsasyhqqgvlstilyeillgkatlya
vlvsalvlmamvkrkdfggrakrsgsg.

5. The binding protein of claim 4 comprising:

1) a) a TCR alpha chain comprising the amino acid sequence of
MLLITSMVLVWMQLSQVNGQQVMQIPQYQHVEGEDFTTYCNSSTLSNIQWYKQRPG
GHPVFLIQLVKSGEVKKQKRLTFQFGEAKKNSSLHITATQTDDVGTYFCAGIGSSNTGKL
IFGQGTTTLQVKPDiqnppdavyqlrdsksdksvclftdfdsqtnvsqksdsvyitdktvldmrsmdfksnsavawsnks
dfacanafnnsiipetffpsesscdvklveksfetdtnlnfnqnsvigfrillllkvagfnllmtlrlwss; and

b) a TCR beta chain comprising the amino acid sequence of
MDTWLVCWAIFSLLKAGLTEPEVTQTPSHQVTQMGQEVILRCVPISNHLIFYWYRQILG
QKVEFLVSFYNNSEISEKSEIFDDQFSVERPDGSNFTLKIRSTKLEDSAMYFCAITGRVSYE
QYFGPGTRLTVTEdlnkvfppevavfepseaeishtqkatlvclatgffpdhvelswwwngkevhsqgvstdpqpplkeqpal
dsryclssrlrvsatfwqnpnrnhfrcqvqfyglsendewtqdrakpvtqivsaeawgradcgftsvsyqqgvlstilyeillgkatlyavl
vsalvlmamvkrkdf, or

2) a) a TCR alpha chain comprising the amino acid sequence of
MLLITSMVLVWMQLSQVNGQQVMQIPQYQHVEGEDFTTYCNSSTLSNIQWYKQRPG
GHPVFLIQLVKSGEVKKQKRLTFQFGEAKKNSSLHITATQTDDVGTYFCAGIGSSNTGKL
IFGQGTTTLQVKPDiqnppdavyqlrdsksdksvclftdfdsqtnvsqksdsvyitdktvldmrsmdfksnsavawsnks
dfacanafnnsiipetffpsdvpdcvklveksfetdtnlnfnqnlvivrillllkvagfnllmtlrlws; and

b) a TCR alpha chain comprising the amino acid sequence of
MDTWLVCWAIFSLLKAGLTEPEVTQTPSHQVTQMGQEVILRCVPISNHLIFYWYRQILG
QKVEFLVSFYNNSEISEKSEIFDDQFSVERPDGSNFTLKIRSTKLEDSAMYFCAITGRVSYE
QYFGPGTRLTVTEdlnkvfppevavfepskaeiahtqkatlvclatgffpdhvelswwwngkevhsqgvstdpqpplkeqpal
ndsryclssrlrvsatfwqnpnrnhfrcqvqfyglsendewtqdrakpvtqivsaeawgradcgitsasyhqgvlstilyeillgkatlya
vlvsalvlmamvkrkdfgsggrakrsgsg.

6. The binding protein of any one of claims 1-5, wherein 1) the TCR V α domain or TCR alpha chain is encoded by a TRAV25 and TRAJ37 gene, and 2) the TCR V β domain or TCR beta chain is encoded by a TRBV2 and TRBJ2-7 gene.

7. The binding protein of any one of claims 1-6, wherein the binding protein is chimeric, humanized, or human.

8. The binding protein of any one of claims 1-7, wherein the binding protein is a TCR, an antigen-binding fragment of a TCR, a single chain TCR (scTCR), a chimeric antigen receptor (CAR), or a fusion protein comprising a TCR and an effector domain, optionally wherein the binding domain comprises a transmembrane domain and an effector domain that is intracellular.
9. The binding protein of any one of claims 1-8, wherein the TCR alpha chain and the TCR beta chain are covalently linked, optionally wherein the TCR alpha chain and the TCR beta chain are covalently linked through a linker peptide.
10. The binding protein of any one of claims 1-9, wherein the TCR alpha chain and/or the TCR beta chain are covalently linked to a moiety, optionally wherein the covalently linked moiety comprises an affinity tag or a label.
11. The binding protein of claim 10, wherein the affinity tag is selected from the group consisting of CD34 enrichment tag, Glutathione-S-Transferase (GST), calmodulin binding protein (CBP), protein C tag, Myc tag, HaloTag, HA tag, Flag tag, His tag, biotin tag, and V5 tag, and/or wherein the label is a fluorescent protein.
12. The binding protein of either one of claims 10 or 11, wherein the covalently linked moiety is selected from the group consisting of an inflammatory agent, cytokine, toxin, cytotoxic molecule, radioactive isotope, or antibody or antigen-binding fragment thereof.
13. The binding protein of any one of claims 1-12, wherein the binding protein binds to the pMHC complex on a cell surface.
14. The binding protein of any one of claims 1-13, wherein the MHC is a MHC multimer, optionally wherein the MHC multimer is a tetramer.
15. The binding protein of any one of claims 1-14, wherein the MHC is a MHC class I molecule.

16. The binding protein of any one of claims 1-15, wherein the MHC comprises an MHC alpha chain that is an HLA serotype HLA-A*02.
17. The binding protein of any one of claims 1-16, wherein the HLA allele is selected from the group consisting of HLA-A*0201, HLA-A*0202, HLA-A*0203, HLA-A*0205, HLA-A*0206, and HLA-A*0207 allele.
18. The binding protein of any one of claims 1-17, wherein binding of the binding protein to a HPV16 E7₁₁₋₁₉ peptide-MHC (pMHC) complex elicits an immune response, optionally wherein the immune response is a T cell response.
19. The binding protein of any one of claims 1-18, wherein the T cell response is selected from the group consisting of T cell expansion, cytokine release, and/or cytotoxic killing.
20. The binding protein of any one of claims 1-19, the binding protein is capable of specifically and/or selectively binding to a HPV16 E7₁₁₋₁₉ immunogenic peptide-MHC (pMHC) complex with a K_d less than or equal to 1×10^{-4} M.
21. The binding protein of claim 20, wherein the binding protein does not bind to a peptide-MHC (pMHC) complex, wherein the peptide is derived from SPTA1, MPL, HERC1, CPAMD8, INTS4, or NUTM1.
22. An isolated nucleic acid molecule that encodes a binding protein of any one of claims 1-21, wherein the nucleic acid is codon optimized for expression in a host cell.
23. A vector comprising the isolated nucleic acid of claim 22.
24. The vector of claim 23, wherein the vector comprises the nucleotide sequence of
- 1)
- tggaagggctaattcactcccaaagaagacaagatatccttgatctgtggatctaccacacacaaggctacttcctgattagcagaactaca
caccagggccaggggtcagatatccactgaccttggatgggtctacaagctagtagccagataaggtagaagaggccaata
aaggagagaacaccagctgttacacctgtgagcctgcatgggatggatgacccggagagagaagtgttagagtggaggtttgacagcc

gcctagcatttcatcacgtggcccgagagctgcatccggagtacttcaagaactgctgatatcgagcttgctacaagggactttccgctggg
gactttccagggaggcggtggcctgggaggactggggagtgaggagccctcagatcctgcatataagcagctgcttttgcctgtactggg
tctctctggtagaccagatctgagcctgggagctctctggctaactagggaaaccactgcttaagcctcaataaagcttgcttgagtcttc
aagtagtgtgtgcccgtctgttgtgactctggtaactagagatccctcagacccttttagtcagtgtggaaaatctctagcagtgggcgccc
aacagggacttgaaagcgaaggggaaccagaggagctctctcgacgcaggactcggcttgctgaagcgcgcacggcaagaggcgag
gggaggcgactggtgagtacgcaaaaatttgactagcggaggctagaaggagagagatgggtgagagcgtcagtattagcggg
ggagaattagatcgcatgggaaaaattcggtaaggccagggggaaagaaaaataataaaacatatagtagggcaagcaggg
agctagaacgattcgagttaatcctggcctgttagaacaatcagaaggctgtagacaaatactgggacagctacaacatcccttcagaca
ggatcagaagaacttagatcattatataacagtagcaaccctctattgtgtgcatcaaaggatagagataaaagacaccaaggaagcttta
gacaagatagaggaagagcaaaaacaaagtaagaccaccgcacagcaagcggccggccgctgatcttcagacctggaggaggagata
tgagggacaattggagaagtgaattatataaaatataaagtagtaaaaattgaaccattaggagtagcaccaccaaggcaagagaagagt
ggtgcagagagaaaaagagcagtgaggataggagcttggcttctgggttctgggagcagcaggaagcactatgggagcagcgctcaat
gacgtgacggtacaggccagacaattattgtctgggtatagtcagcagcagaacaattgtgagggctattgagggcgaacagcatctgt
tgcaactcacagctctggggcatcaagcagctccaggcaagaatcctggctgtggaaagatacctaaaggatcaacagctcctggggatttg
gggtgtcttgaaaactcattgcaccactgctgtgccttggaatgctagtggagtaataaatctctggaacagatttggaatcacacgacct
ggatggagtgggacagagaaattaacaattacacaagcttaatacactccttaattgaagaatcgaaaaccagcaagaaaagaatgaaca
agaattattggaattagataaatgggcaagtttggaattgggttaacatacaaaattggctgtggtatataaaattattcataatgatagtagga
ggcttgtaggttaagaatagttttgctgtactttctatagtgaaatagagttaggcagggatattcaccattatcgtttcagaccacctcccaa
ccccgaggggacccgacaggccgaaggaatagaagaagaaggtggagagagagacagagacagatccattcgattagtgaacggat
ctcgacgggtatcgccgaattaattcacaatggcagttatccacaattttaaagaaaaggggggattgggggggtacagtgaggggga
aagaatagtagacataatagcaacagacatacaaaactaaagaattacaaaaacaaattacaaaaattcaaaattttcgggtttattacaggCG
cGCcagagatccagtttgacCTgcAGGTGAAAGACCCACCTGTAGGTTTGGCAAGtTAGCTTA
AGTAACGCCATTTTGC AAGGCATGGAAAATACATAACTGAGAATAGAGAAGTTCAG
ATCAAGGTTAGGAACAGAGAGACAGCAGAATATGGGCCAAACAGGATATCTGTGGT
AAGCAGTTCCTGCCCGGCTCAGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCC
GCCCTCAGCAGTTTCTAGCGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGA
AATGACCCTGTGCCTTATTTGAACTAACCAATCAGTTtGCTTCTtGCTTCTGTTtGtGtGC
TTCTGCTCCCTGAGCTCAATAAAAGAGCCCAACCCCTCACTtGGtGgGCCAGTCCTC
tGATAGACTGtGTCcCCtGGaTACCCGTAcgggtaccgctagcgccaccATGGATACCTGGCTCGTG
TGTTGGGCCATCTTTAGCCTGCTGAAGGCCGGAAGTACCGAGCCTGAAGTGACCCAG
ACTCCAAGCCATCAAGTGACTCAGATGGGGCAAGAAGTCATTCTGCGTTGCGTGCCC
ATCAGCAACCACCTGTACTTTTATTGGTATCGCCAGATCCTGGGCCAGAAAGTGGA
TTCCTGGTGTCTTCTACAACAATGAGATCTCCGAGAAGTCCGAGATCTTCGACGAC
CAGTTCTCCGTGGAAAGACCCGACGGCAGCAACTTCACACTGAAGATCCGGTCTAC
CAAACCTGAGGACTCCGCTATGTATTTTTGTGCAATCACAGGTCGCGTTTCATATGA
GCAATATTTCCGGCCGGGCACCAGGCTCACGGTCACAGAAGATCTGAACAAGGTGT
TCCCTCCAGAGGTGGCCGTGTTTCGAGCCTTCTAAGGCCGAGATCGCCACACACAAA
AAGCCACCCTCGTGTGCCTGGCCACCGGCTTTTTCCCCGACCACGTGGAAGTGTCTT
GGTGGGTCAACGGCAAAGAGGTGCACTCCGGCGTGTCAACGGATCCCCAGCCTCTG
AAAGAACAGCCTGCCCTGAACGACAGCCGGTACTGCCTGAGCTCCAGACTGAGAGT
GTCCGCCACCTTCTGGCAGAACCCCCGGAACCACTTCAGATGCCAGGTGCAGTTTAA
CGGCCTGAGCGAGAACGACGAGTGACCCAGGACAGAGCCAAGCCCGTGACACAA
ATCGTGTCTGCCGAAGCCTGGGGAAGAGCCGATTGCGGCATCACCAGCGCCTCCTAT
CACCAGGGCGTGCTGAGCGCCACAATCCTGTACGAAATCCTGCTGGGCAAGGCCAC
CCTGTACGCCGTGCTGGTGTCTGCTCTGGTGTCTGATGGCCATGGTCAAGCGGAAGGA
CTTTGGCAGCGGCAGAGCCAAAAGGTCCGGGAGCGGTGCGACAAACTTTAGCCTGT

TGAAACAAGCCGGCGACGTTGAAGAGAACCCCGGACCTATGCTGCTGATCACCTCC
 ATGCTGGTGTGTGGATGCAGCTGAGCCAAGTGAACGGCCAGCAAGTGATGCAGAT
 CCCTCAGTACCAGCACGTGCAAGAAGGCGAGGACTTCACCACCTACTGCAACAGCA
 GCACCACACTGAGCAACATCCAGTGGTACAAGCAGCGGCCTGGCGGACACCCTGTG
 TTTCTGATCCAGCTGGTCAAGTCCGGCGAAGTGAAGAAGCAGAAGCGGCTGACCTT
 CCAGTTCGGCGAGGCCAAGAAGAACAGCAGCCTGCACATCACCGCCACACAGACCA
 CAGATGTGGGCACCTACTTCTGCGCTGGCATCGGTAGCAGCAACACCGGTAAGCTC
 ATCTTTGGGCAAGGGACAACCTTTACAAGTAAAACCAGACATCCAGAACCCCGACCC
 CGCCGTGTACCAGCTGAGGGACTCCAAGTCCAGCGACAAGAGCGTGTGTCTGTTTAC
 GGACTTCGACAGCCAGACCAACGTGAGTCAAAGCAAGGACAGCGACGTCTACATAA
 CGGATAAGACCGTGCTGGACATGCGGAGCATGGACTTCAAGAGCAACAGCGCCGTG
 GCCTGGTCCAACAAGAGCGACTTCGCCTGCGCCAACGCCTTCAACAACAGCATCATC
 CCCGAGGACACCTTCTTCCCCAGCAGCGACGTGCCCTGCGACGTGAAACTGGTGGA
 GAAGTCCTTCGAGACAGACACCAATCTGAACTTTTCAAGAACCTGCTGGTGATCGTGCT
 GCGGATTCTGCTGCTGAAAGTGGCCGGCTTCAATCTGCTGATGACCCTGCGGCTGTG
 GAGCAGCAGGGCTAAGAGGTCCGGCAGCGGAGCCACCAATTTTTCCCTGCTGAAAC
 AGGCTGGTGACGTGGAAGAAAACCTGGCCCCATGGCGCTGCCCGTCACCGCGCTG
 CTGCTGCCCCCTGGCGCTGCTGTTACACGCCGCTCGGCCAGAGCTTCCCACCCAGGGC
 ACATTCTCCAACGTGTCCACCAATGTGTCTGGGAGGCGGCGGATCGTCCCAGTTCAGA
 GTGTCCCCTCTGGACCGCACCTGGAACCTGGGCGAGACCGTGAGCTGAAATGTCA
 GGTCCTGCTGAGCAACCCGACCTCCGGGTGCAGTTGGCTGTTCCAGCCGCGTGGTG
 TGCCGCAAGCCCTACGTTCTGCTTTACCTGAGCCAGAACAAGCCCAAGGCGGCCG
 AGGGCCTGGACACCCAGAGATTCTCCGGCAAGCGCCTGGGGGACACATTCGTGCTT
 ACTTTGAGCGATTTCCGCAGAGAGAACGAGGGCTACTATTTCTGTTCCGGCGCTGAGC
 AATTCATCATGTATTTACGCCACTTTGTGCCAGTGTTCTGCTGCCAAGCCTACCA
 CAACACCAGCTCCCCGTCCCCCGACTCCGGCGCCTACCATCGCGAGTCAACCGTTGA
 GCCTGAGGCCTGAGGCTTGTTCGGCCCGCTGCGGGGGGTGCCGTCCACACCAGGGGC
 CTCGACTTTGCGTGCGACATCTATATTTGGGCGCCTCTGGCGGGTACCTGCGGGGTG
 CTGCTGCTGTCATTGGTGATTACCCTGTACTGCAATCACCGCAACCGCCGGCGGGTC
 TGTAAGTGCCACGGCCTGTGGTCAAGTCCGGTGACAAACCGTTCGCTCTCGGCTCGC
 TACGTGCGCGCTAAGCGCAGCGGTTCCGGGGCCACCAACTTTTCATTGCTGAAGCAG
 GCCGGTGATGTGGAGGAGAATCCAGGGCCCATGCGCCCCAGGCTTTGGCTCCTTCTT
 GCTGCTCAGCTCACTGTCTTGCATGGCAACTCCGTTCTGCAGCAGACTCCCGCCTAC
 ATCAAGGTGCAGACGAACAAGATGGTGATGCTGTCATGCGAGGCCAAGATCTCTCT
 TTCAAATATGAGAATTTATTGGCTACGACAGCGCCAGGCCCCCTCCAGCGACAGCCA
 CCACGAGTTCCTGGCGCTTTGGGATTCTGCTAAAGGCACCATCCATGGAGAGGAGGT
 GGAACAGGAGAAGATAGCTGTCTTCCGCGACGCATCCCGCTTCATCCTGAACCTGAC
 CAGCGTGAAGCCGGAGGACAGCGGCATCTACTTCTGTATGATCGTTGGCTCCCCGA
 GCTGACCTTCGGCAAAGGCACCCAGCTGTCCGTGGTGGACTTCCTGCCACCCACAGC
 CCAGCCAACCAAGAAATCCACCCTCAAGAAGCGCGTGTGCCGACTGCCCCGCCCTG
 AAACCCAGAAGGGCCCTCTGTGCTCCCCCATCACCTTGGACTGCTGGTGGCGGGAG
 TCCTGGTGCTGCTCGTATCTCTGGGTGTCGCCATCCACCTGTGCTGCCGCCGCCGCCG
 CGCCCCGCTGAGGTTTATGAAACAGTTTTACAAGTGATAAatcgatagatcctaataacctctggatt
 acaaaatttgtgaagattgactggtattcttaactatgttgctccttttacgctatgtggatacgtgctttaatgcctttgtatcatgtattgcttc
 ccgtagtgctttcattttctccttgataaactcctgggtgctgtctctttatgaggagttgtggcccggtgtcaggcaacgtggcggtggtgtgc
 actgtgttgctgacgcaacccccactggttggggcattgccaccacgtgtcagctcctttccgggactttcgctttccccctccctattgccac

gcctagcatttcatcacgtggcccgagagctgcatccggagtacttcaagaactgctgatatcgagcttgctacaagggactttccgctggg
gactttccagggaggcggtggcctgggaggactggggagtgaggagccctcagatcctgcatataagcagctgcttttgcctgtactggg
tctcttggttagaccagatctgagcctgggagctctctggctaactagggaaaccactgcttaagcctcaataaagcttgcttgagtcttc
aagtagtgtgtgcccgtctgtgtgactctggttaactagagatccctcagacccttttagtcagtgtggaaaatctctagcagtgggcgccc
aacagggacttgaaagcgaaggggaaccagaggagctctctgcagcgaggactcggcttgctgaagcgcgacggcaagaggcgag
gggaggcgactggtgagtacgcaaaaattttagtagcggaggctagaaggagagagatgggtgagagcgctcagtattagcgagg
ggagaattagatcgcatgggaaaaattcggtaaggccagggggaaagaaaaataaaatataatagtagggcaagcagggg
agctagaacgattcgagtttaactcctggcctgtagaacaatcagaaggctgtagacaaatactgggacagctacaacatcccttcagaca
ggatcagaagaacttagatcattatataacagtagcaaccctctattgtgtgcatcaaaggatagagataaaagacaccaaggaagcttta
gacaagatagaggaagagcaaaaagtaagaccaccgcacagcaagcggccggccgctgatcttcagacctggaggaggagata
tgagggacaattggagaagtgaattatataaaatataaagtagtaaaaattgaaccattaggagtagcaccaccaaggcaagagaagagt
ggtgcagagagaaaaagagcagtgagggaataggagcttggcttgggttctgggagcagcaggaagcactatgggagcagcgctaat
gacgctgacggtacaggccagacaattattgtctggtatagtcagcagcagaacaattgctgagggctattgaggcgcaacagcatctgt
tgcaactcacagctctggggcatcaagcagctccaggcaagaatcctggctgtggaaagatacctaaaggatcaacagctcctggggatttg
gggttgccttggaactcattgcaccactgctgtgccttggaatgctagtggagtaataatctctggaacagatttggaatcacacgacct
ggatggagtgggacagagaaattaacaattacaaagcttaatacactccttaattgaagaatcgaaaaccagcaagaaaagaatgaaca
agaattattggaattagataaatgggcaagtttggaattgggttaacatacaaaattggctgtggtatataaaattattcataatgatagtagga
ggcttgtaggttaagaatagttttgctgtactttctatagtgaaatagagttaggcagggatattcaccattatcgtttcagaccacctcccaa
ccccgaggggacccgacaggccgaaggaatagaagaagaaggtggagagagagacagagacagatccattcgattagtgaacggat
ctgcagcgtatcgccgaattaattcacaatggcagttatccacaattttaaagaaaaggggggattgggggggtacagtgaggggga
aagaatagtagacataatagcaacagacatacaaaactaaagaattacaaaaaattacaaaaattcaaaattttcgggtttattacaggCG
cGCcagagatccagtttggacCTgcAGGTGAAAGACCCACCTGTAGGTTTGGCAAGtTAGCTTA
AGTAACGCCATTTTGCAAGGCATGGAAAATACATAACTGAGAATAGAGAAGTTCAG
ATCAAGGTTAGGAACAGAGAGACAGCAGAATATGGGCCAAACAGGATATCTGTGGT
AAGCAGTTCCTGCCCGGCTCAGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCC
GCCCTCAGCAGTTTCTAGCGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGA
AATGACCCTGTGCCTTATTTGAACTAACCAATCAGTTtGCTTCTtGCTTCTGTtGtGtGC
TTCTGCTCCCTGAGCTCAATAAAAGAGCCCACAACCCCTCACTtGGtGgGCCAGTCCTC
tGATAGACTGtGTCcCCtGGaTACCCGTAcgggtaccgctagcgccaccatggatacctggctcgtgtgttgggccatc
tttagcctgctgaaggccggactgaccgagcctgaagtgaccagactccaagccatcaagtgactcagatggggcaagaagtcattctgc
gttgctgctcccatcagcaaccacgtgacttttattggtatcgccagatcctgggccagaaagtgggaattcctggtgtccttctacaacaatga
gatctccgagaagtccgagatcttcgacgaccagttctccgtggaaagacccgacggcagcaacttcacactgaagatccggtctacaaa
cttgaggactccgctatgtattttgtgcAATCACAGGTCGCGTTTCATATGAGCAATATTTTCGGGCCG
GGCACCAGGCTCACGGTCACAGAAGATCTGAACAAGGTGTTCCCTCCAGAGGTGGC
CGTGTTCGAGCCTTCTAAGGCCGAGATCGCCACACACAAAAAGCCACCCTCGTGTG
CCTGGCCACCGGCTTTTTCCCCGACCACGTGGAACCTGTCTTGGTGGGTCAACGGCAA
AGAGGTGCACTCCGGCGTGTCAACGGATCCCCAGCCTCTGAAAGAACAGCCTGCCC
TGAACGACAGCCGGTACTGCCTGAGCTCCAGACTGAGAGTGTCCGCCACCTTCTGGC
AGAACCCCCGGAACCACTTCAGATGCCAGGTGCAGTTTTACGGCCTGAGCGAGAAC
GACGAGTGGACCCAGGACAGAGCCAAGCCCGTGACACAAATCGTGTCTGCCGAAGC
CTGGGGAAGAGCCGATTGCGGCATCACCAGCGCCTCCTATCACCAGGGCGTGCTGA
GCGCCACAATCCTGTACGAAATCCTGCTGGGCAAGGCCACCCTGTACGCCGTGCTGG
TGTCTGCTCTGGTGCTGATGGCCATGGTCAAGCGGAAGGACTTTGGCAGCGGCAGA
GCCAAAAGGTCCGGGAGCGGTGCGACAACTTTAGCCTGTTGAAaagccggCGACGT
TGAAGAGAACCCCGGACCTatgctgctgatcacctccatgctgggtgctgtggatgcagctgagccaagtgaacggccag
caagtgatgcagatccctcagtagcagcagtgcaagaaggcgaggacttcaccacctactgcaacagcagcaccacactgagcaacat

ccagtgggtacaagcagcggcctggcgacacctgtgtttctgatccagctgggtcaagtccggcggaagtgaagaagcagaagcggctga
ccttcagttcggcgaggccaagaagaacagcagcctgcacatcaccgccacagaccagatgtgggcacctacttcTGCGCT
GGCATCGGTAGCAGCAACACCGGTAAGCTCATCTTTGGGCAAGGGACAACTTTACA
AGTAAAACCAAGacatccagaacccccgaccccgctgtaccagctgagggactccaagtccagcgacaagagcgtgtgtctgt
ttacggacttcgacagccagaccaacgtgagtc aaagcaaggacagcgacgtctacataacggataagaccgtgtgtggacatgcggagc
atggacttcaagagcaacagcgccgtggcctggccaacaagagcgacttcgcctgcgccaacgccttcaacaacagcatcatccccgag
gacaccttctccccagcagcgacgtgccctgcgacgtgaaactgggtggagaagtccttcgagacagacaccaatctgaacttcagaacc
tgctggtgatcgtgtgcggattctgtgCTGAAAGTGGCCGGCTTCAATCTGCTGATGACCCTGCGG
CTGTGGAGCAGCAGGGCTAAGAGGTCCGGCAGCGGAGCCACCAATTTTTCCCTGCT
GAAACAGGCTGGTGACGTGGAAGAAAACCCTGGCCCCATGGGTGCGGGGGCTGCTCA
GGGGCCTGTGGCCGCTGCACATCGTCCTGTGGACGCGTATCGCCAGCACGATCCCAC
CGCACGTTCAGAAGTCGGTTAATAACGACATGATAGTCACTGACAACAACGGTGCA
GTCAAGTTTCCACAACCTGTGTAAATTTTGTGATGTGAGATTTTCCACCTGTGACAACC
AGAAATCCTGCATGAGCAACTGCAGCATCACCTCCATCTGTGAGAAGCCACAGGAA
GTCTGTGTGGCTGTATGGAGAAAGAATGACGAGAACATAACACTAGAGACAGTTTG
CCATGACCCCAAGCTCCCCTACCATGACTTTATTCTGGAAGATGCTGCTTCTCCAAA
GTGCATTATGAAGGAAAAAAAAAAGCCTGGTGAGACTTTCTTCATGTGTTCTGTAG
CTCTGATGAGTGCAATGACAACATCATCTTCTCAGAAGAATATAACACCAGCAATCC
TGACTTGTGCTAGTCATATTTCAAGTGACAGGCATCAGCCTCCTGCCACCCTGGG
AGTTGCCATATCTGTCATCATCATCTTCTACTGCTACCGCGTTaaccggcagcagaagTAGTG
ATAAatc gatagatcctaataacacctgtgattacaaaattgtgaaagattgactggattcttaactatgtgtcctttacgctatgtggat
acgctgcttaatgcctttgatcatgctattgctcccgtatggctttcattttctcctcctgtataaatcctgggtgtgtctctttatgaggagttgt
ggccccgtgtcaggcaacgtggcgtgggtgtgactgtgtttgtgacgaacccccactgggtggggcattgccaccacctgtcagctcctt
ccgggactttcgctttccccctccctattgccacggcggaactcatgccgctgtccttgcctgtgctggacaggggctcggctgttgggc
actgacaattccgtgtgtgtcggggaaatcatcgctcttcttggtgtctgcctgtgttggcacctggattctgcgcgggacgtccttctg
ctacgtcccttcggccctcaatccagcggaccttcttcccgggcctgtgcggctctgcggccttctccgcgtcttcgcttcgctcctca
gacgagtcggatctccctttggcgccctccccgcctgagatccttaagaccaatgacttacaaggcagctgtagatcttagccacttttta
aagaaaaggggggactggaagggttaattcactcccaacgaagacaagatctgcttttgcctgtactgggtctctctggttagaccagatct
gagcctgggagctctctggttaactaggggaaccactgcttaagcctcaataaagcttgcttgagtgttcaagtagtgtgtgcccgtctgtt
gtgtgactctggttaactagagatccctcagacccttttagtcagtgtggaaaatctctagcagtagtagttcatgtcatcttattattcagtatttat
aacttgcaagaaatgaatatcagagagtgagaggccgggttaattaaggaaagggctagatcattctgaagacgaaagggcctcgtga
tacgcctattttataggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctattgttta
tttttctaaatacattcaaatatgatccgctcatgagacaataaccctgataaatgctcaataatattgaaaaaggaagagtatgagtattcaac
atttccgtgtcgccttattcccttttttgcggcattttgccttctgttttgcctcaccagaaacgctgggtgaaagtaaaagatgctgaagatcag
ttgggtgcacgagtggtgttacatcgaactggatctcaacagcggtaagatccttgagagtttcgccccgaagaacgtttccaatgatgagc
acttttaaagtctgctatgtggcgcggtattatcccgtgttgacggcggaagagcaactcggctgcgcgcatacatttctcagaatgactt
gggtgagtactaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgtgcataaccatgagtataaca
ctgcggccaacttacttctgacaacgatcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcatgtaactgccttgatc
gttgggaaccggagctgaatgaagccatacacaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgttgcgcaaacat
taactggcgaactacttacttagcttccccggcaacaattaatagactggatggaggcggataaagttgcaggaccacttctgcgtcggcc
cttccggctgggtgtttattgtgataaatctggagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagatggtaagcc
ctcccgtatcgtattctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcactgatta
agcattggtaactgtcagaccaagttactcatatatacttttagattgatttaaaacttcatttttaattaaaaggatctaggtgaagatccttttga
taatctcatgacaaaaatccctaactgagtttctgctccactgagcgtcagaccccgtagaaaagatcaaggatcttctgagatccttttt
ctgcgcgtaactgtgtgcttgaacaaaaaaaaccaccgctaccagcgggtgtgttgcggatcaagagctaccaactcttttccgaag
gtaactggcttcagcagagcgcagatacacaataactgttcttagttagccgtagtaggaccacttcaagaactctgtagcaccgcct

acatacctcgctctgctaactctgttaccagtggtgctgctgccagtggtcgataagtcgtgtcttaccgggttgactcaagacgatagttaccgg
ataaggcgacgcggctcgggctgaacgggggggttcgtgcacacagcccagcttggagcgaacgacctacacgaactgagatacctaca
gcgtgagctatgagaaagcgccacgcttcccgaagggagaaagcgggacaggtatccggtaagcggcaggggtcggaacaggagagc
gcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggttcgccacctctgacttgagcgtcgattttgtgatgctc
gtcagggggggcggagcctatggaaaaacgccagcaacgcggccttttacgggtcctggccttttgctggcctttgctcacatgttcttcct
gcgttatccCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATAACCGCTC
GCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCG
CCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCAAGCT
CATGGCTGACTAATTTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAG
CTATTCCAGAAAGTAGTGAGGAGGCTTTTTTGGAGGCCCTAGGCTTTTGCAAAAAGCTC
CCCGTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTA
ATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCG
TATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACAT
GATTACGAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCC
AAACTCATCAATGTATCTTATCATGTCTGGATCAACTGGATAACTCAAGCTAACCAA
AATCATCCCAAACCTTCCCACCCCATACCCTATTACCACTGCCAATTACCTGTGGTTTC
ATTTACTCTAAACCTGTGATTCTCTGAATTATTTTCATTTTAAAGAAATTGTATTTG
TTAAATATGTACTACAAACTtagtagt;

3)

GAATTCGTCGACGCTAGCTGGCTTGTTGTCCACAACCATTAACCTTAAAAGCTTTA
AAAGCCTTATATATTCTTTTTTTTCTTATAAACTTAAAACCTTAGAGGCTATTTAAG
TTGCTGATTTATATTAATTTTATTGTTCAAACATGAGAGCTTAGTACGTGAAACATGA
GAGCTTAGTACATTAGCCATGAGAGCTTAGTACATTAGCCATGAGGGTTTAGTTCAT
TAAACATGAGAGCTTAGTACATTAAACATGAGAGCTTAGTACATACTATCAACAGGT
TGAAGTGTCTGATCTGTACAGTAGAATTGGTAAAGAGAGTTGTGTAAAATATTGAGTT
CGCACATCTTGTTGTCTGATTATTGATTTTTTGGCGAAACCATTTGATCATATGACAAG
ATGTGTATCTACCTTAACTTAATGATTTTTGATAAAAATCATTAGGTACCAATTACATT
GCTTGCAATTAACCTTTAACGGTTATAAGGATCTAGATGAGATAGAAAGATTTGGT
TTTCGGATTTGTGTTACATAAGATGCCTAAAATAAAAATTGAGATTCAATTTTTTTTA
AACTTTTTTTTAATTGGTGGTAAGAATATCCCTCTACCTGTTTGAGAGTAATGAAAT
TGTAGTATGATTTTTCAACAACTAAAAAACAACATAAATCTCACATAATAACTTT
ATTTCAATCACACAATTGAATACCAATAGGTTGACAGTACTTACCAGCCTGCAGGTG
AAAGACCCACCTGTAGGTTTGGCAAGTTAGCTTAAGTAACGCCATTTTGCAAGGCA
TGGAATAACATAACTGAGAATAGAGAAGTTCAGATCAAGGTTAGGAACAGAGAG
ACAGCAGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTC
AGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCCGCCCTCAGCAGTTTCTAGCGA
ACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTG
AACTAACCAATCAGTTTGCTTCTTGCTTCTGTTTGTGTGCTTCTGCTCCCTGAGCTCA
ATAAAAGAGCCCACAACCCCTCACTTGGTGGGCCAGTCCTCTGATAGACTGTGTCCC
CTGGATACCCGTACGGTACCGCTAGCGCCACCATGGATACCTGGCTCGTGTGTTGGG
CCATCTTTAGCCTGCTGAAGGCCGACTGACCGAGCCTGAAGTGACCCAGACTCCA
AGCCATCAAGTGACTCAGATGGGGCAAGAAGTCATTCTGCGTTGCGTGCCCATCAGC
AACCACCTGTACTTTTATTGGTATCGCCAGATCCTGGGCCAGAAAGTGGAATTCCTG
GTGTCCTTCTACAACAATGAGATCTCCGAGAAGTCCGAGATCTTCGACGACCAGTTC
TCCGTGGAAAGACCCGACGGCAGCAACTTCACACTGAAGATCCGGTCTACCAAACCT

TGAGGACTCCGCTATGTATTTTTGTGCAATCACAGGTCGCGTTTCATATGAGCAATA
TTTCGGGGCCGGGCACCAGGCTCACGGTCACAGAAGATCTGAACAAGGTGTTCCCTCC
AGAGGTGGCCGTGTTTCGAGCCTTCTAAGGCCGAGATCGCCACACACAAAAAGCCA
CCCTCGTGTGCCTGGCCACCGGCTTTTTCCCCGACCACGTGGAAGTGTCTTGGTGGG
TCAACGGCAAAGAGGTGCACTCCGGCGTGTCAACGGATCCCCAGCCTCTGAAAGAA
CAGCCTGCCCTGAACGACAGCCGGTACTGCCTGAGCTCCAGACTGAGAGTGTCCGC
CACCTTCTGGCAGAACCCCCGGAACCACTTCAGATGCCAGGTGCAGTTTTACGGCCT
GAGCGAGAACGACGAGTGGACCCAGGACAGAGCCAAGCCCGTGACACAAATCGTG
TCTGCCGAAGCCTGGGGAAGAGCCGATTGCGGCATCACCAGCGCCTCCTATCACCA
GGGCGTGCTGAGCGCCACAATCCTGTACGAAATCCTGCTGGGCAAGGCCACCCTGT
ACGCCGTGCTGGTGTCTGCTCTGGTGCTGATGGCCATGGTCAAGCGGAAGGACTTTG
GCAGCGGCAGAGCCAAAAGGTCCGGGAGCGGTGCGACAAACTTTAGCCTGTTGAAA
CAAGCCGGCGACGTTGAAGAGAACCCCGGACCTATGCTGCTGATCACCTCCATGCT
GGTGCTGTGGATGCAGCTGAGCCAAGTGAACGGCCAGCAAGTGATGCAGATCCCTC
AGTACCAGCACGTGCAAGAAGGCGAGGACTTCACCACCTACTGCAACAGCAGCACC
ACACTGAGCAACATCCAGTGGTACAAGCAGCGGCCTGGCGGACACCCTGTGTTTCT
GATCCAGCTGGTCAAGTCCGGCGAAGTGAAGAAGCAGAAGCGGCTGACCTTCCAGT
TCGGCGAGGCCAAGAAGAACAGCAGCCTGCACATCACCGCCACACAGACCACAGAT
GTGGGCACCTACTTCTGCGCTGGCATCGGTAGCAGCAACACCGGTAAGCTCATCTTT
GGGCAAGGGACAACCTTTACAAGTAAAACCAGACATCCAGAACCCCGACCCCGCCGT
GTACCAGCTGAGGGACTCCAAGTCCAGCGACAAGAGCGTGTGTCTGTTTACGGACTT
CGACAGCCAGACCAACGTGAGTCAAAGCAAGGACAGCGACGTCTACATAACGGATA
AGACCGTGCTGGACATGCGGAGCATGGACTTCAAGAGCAACAGCGCCGTGGCCTGG
TCCAACAAGAGCGACTTCGCCTGCGCCAACGCCTTCAACAACAGCATCATCCCCGA
GGACACCTTCTTCCCCAGCAGCGACGTGCCCTGCGACGTGAAACTGGTGGAGAAGT
CCTTCGAGACAGACACCAATCTGAACTTTCAGAACCTGCTGGTGATCGTGCTGCGGA
TTCTGCTGCTGAAAGTGGCCGGCTTCAATCTGCTGATGACCCTGCGGCTGTGGAGCA
GCAGGGCTAAGAGGTCCGGCAGCGGAGCCACCAATTTTTCCCTGCTGAAACAGGCT
GGTGACGTGGAAGAAAACCCTGGCCCCATGGCGCTGCCCGTCACCGCGCTGCTGCT
GCCCCCTGGCGCTGCTGTTACACGCCGCTCGGCCAGAGCTTCCCACCCAGGGCACATT
CTCCAACGTGTCCACCAATGTGTGCGGAGGCGGCGGATCGTCCCAGTTCAGAGTGTC
CCCTCTGGACCGCACCTGGAACCTGGGCGAGACCGTGAGCTGAAATGTCAGGTCC
TGCTGAGCAACCCGACCTCCGGGTGCAGTTGGCTGTTCCAGCCGCGTGGTGCTGCCG
CAAGCCCTACGTTTCTGCTTTACCTGAGCCAGAACAAGCCCAAGGCGGCCGAGGGC
CTGGACACCCAGAGATTCTCCGGCAAGCGCCTGGGGGACACATTCGTGCTTACTTTG
AGCGATTTCCGCAGAGAGAAACGAGGGCTACTATTTCTGTTCCGGCGCTGAGCAATTCC
ATCATGTATTTACGCCACTTTGTGCCAGTGTTCTGCTGCCAAGCCTACCACAACA
CCAGCTCCCCGTCCCCGACTCCGGCGCCTACCATCGCGAGTCAACCGTTGAGCCTG
AGGCCTGAGGCTTGTCGGCCCCGCTGCGGGGGGTGCCGTCCACACCAGGGGCCTCGA
CTTTGCGTGCGACATCTATATTTGGGCGCCTCTGGCGGGTACCTGCGGGGTGCTGCT
GCTGTCAATTGGTGATTACCCTGTACTGCAATCACCGCAACCGCCGGCGGGTCTGTAA
GTGCCCACGGCCTGTGGTCAAGTCCGGTGACAAACCGTCGCTCTCGGCTCGCTACGT
GCGCGCTAAGCGCAGCGGTTCCGGGGCCACCAACTTTTCATTGCTGAAGCAGGCCG
GTGATGTGGAGGAGAATCCAGGGCCCATGCGCCCCAGGCTTTGGCTCCTTCTTGCTG
CTCAGCTCACTGTCTTGTCATGGCAACTCCGTTCTGCAGCAGACTCCCGCCTACATCA
AGGTGCAGACGAACAAGATGGTGATGCTGTCATGCGAGGCCAAGATCTCTCTTTCA

AATATGAGAATTTATTGGCTACGACAGCGCCAGGCCCCCTCCAGCGACAGCCACCA
CGAGTTCCTGGCGCTTTGGGATTCTGCTAAAGGCACCATCCATGGAGAGGAGGTGG
AACAGGAGAAGATAGCTGTCTTCCGCGACGCATCCCGCTTCATCCTGAACCTGACCA
GCGTGAAGCCGGAGGACAGCGGCATCTACTTCTGTATGATCGTTGGCTCCCCCGAGC
TGACCTTCGGCAAAGGCACCCAGCTGTCCGTGGTGGACTTCCTGCCCACCACAGCCC
AGCCAACCAAGAAATCCACCCTCAAGAAGCGCGTGTGCCGACTGCCCCGCCCTGAA
ACCCAGAAGGGCCCTCTGTGCTCCCCCATCACCTTGGACTGCTGGTGGCGGGAGTC
CTGGTGCTGCTCGTATCTCTGGGTGTGCGCCATCCACCTGTGCTGCCGCCGCCGCCGC
GCCCCCCTGAGGTTTATGAAACAGTTTTACAAGTGATAAATCGATGGAAGGGTGGC
ATCCCTGTGACCCCTCCCCAGTGCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGC
CCACCAGCCTTGTCTAATAAAATTAAGTTGCATCATTTTTGTCTGACTAGGTGTCCTT
CTATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGGGCAAGTTGGGAAGA
CAACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAA
TCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTC
CCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTTG
GTAGAGACGGGGTTTCACCATATTGGCCAGGCTGGTCTCCAACCTCCTAATCTCAGGT
GATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCC
CTTCCCTGTCTTCTGATTACTAGTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCAC
ATCGCCACAGTCCCCGAGAAGTTGTGGGGAGGGGTCCGGCAATTGAACCGGTGCCT
AGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTT
TTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTT
TTTCGCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGGG
CCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTGGCTGC
AGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTTCGAGGC
CTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCT
GGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTTCGATA
AGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGA
TAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTTGGGGCCGC
GGGCGGCGACGGGGCCCCGTGCGTCCCAGCGCACATGTTTCGGCGAGGCGGGGCCTGC
GAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTG
GTGCTTGGCCTCGCGCCGCGGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGG
TCGGCACCAAGTTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAG
CTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAA
AGGAAAAGGGCCTTTCCGTCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCGG
GCGCCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGTCTTTAGGTT
GGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAA
GTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAAATTTGCCCTTTTTGAGTTTG
GATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTTCTTCCATTTCA
GGTGTGCTGAACTAGTCCAGTGTGGTGGAAATCTGCAGATATCACGGCTAGCGCCAC
CATGGGTGCGGGGGCTGCTCAGGGGCCTGTGGCCGCTGCACATCGTCCTGTGGACGC
GTATCGCCAGCACGATCCCACCGCACGTTTCAAGTTCGGTGAATAACGACATGATA
GTCACTGACAACAACGGTGCAGTCAAGTTTCCACAACCTGTGTAAATTTTGTGATGTG
AGATTTTCCACCTGTGACAACCAGAAATCCTGCATGAGCAACTGCAGCATCACCTCC
ATCTGTGAGAAGCCACAGGAAGTCTGTGTGGCTGTATGGAGAAAGAATGACGAGAA
CATAACACTAGAGACAGTTTGCCATGACCCCAAGCTCCCCTACCATGACTTTATTCT
GGAAGATGCTGCTTCTCCAAAGTGCATTATGAAGGAGAAGAAAAAGCCTGGTGAGA

CTTTCTTCATGTGTTCTCTGTAGCTCTGATGAGTGCAATGACAACATCATCTTCTCAGA
AGAATATAACACCAGCAATCCTGACTTGTTGCTAGTCATATTTCAAGTGACAGGCAT
CAGCCTCCTGCCACCACTGGGAGTTGCCATATCTGTCATCATCATCTTCTACTGCTAC
CGCGTGAACCGGCAGCAGAAGGCTAGTGGTTCAGGCGCAACGAATTTCTCTTTGCTG
AAGCAGGCTGGGGATGTCTGAAGAAAATCCGGGTCCAATGGTGGGCTCGCTCAACTG
CATCGTAGCAGTCTCCCAGAATATGGGCATCGGGAAGAACGGTGATTTCCCGTGGCC
CCCACTTCGCAACGAGAGCCGTTATTTCCAAAGAATGACTACAACCTCCTCCGTGGA
GGGTAAGCAGAACCTGGTCATCATGGGGAAGAAGACCTGGTTCTCTATCCCTGAAA
AAAACCGCCCCCTGAAGGGCCGCATCAACCTGGTGCTGAGCAGGGAACCTCAAGGAG
CCTCCTCAGGGGCGCGCATTTTCTGAGCCGCTCATTGGATGACGCTCTCAAACCTGACC
GAACAGCCGGAGCTAGCCAACAAGGTGGACATGGTGTGGATCGTCGGAGGCTCCTC
CGTGTACAAGGAGGCCATGAATCACCCCGGCCACTTGAAGCTGTTTCGTCACCCGGAT
CATGCAGGACTTCGAGTCGGACACGTTCTTTCCAGAGATTGACCTGGAGAAGTACAA
GCTGCTGCCCCGAGTACCCGGGAGTTCTTAGTGATGTGCAGGAGGAGAAAGGCATCA
AGTACAAATTTGAGGTGTACGAGAAGAACGACTAACGGTCCGTCCTGACCAATGCT
GGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
CAATAGCATCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGT
TTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACAGGTTACCTCAGTCTCCT
AGGTACGTCTTATATCTATGAAAAAACATTCAAAGCACAACATCTAGAAGAACTT
ACCTTTTTTTCACCACTCTATTGCAAAGATATGTACCGATTTCTCTCGAAGTACAAAAA
ACCGCTAGTTTTCAAATTCACCTCAAGACTTTGAAAAAAAATTGAATCTGTCAATGT
CAAATAAAATCAGAAACAAATGTCATAATGTTACGTTAATGTTGTCAGGTGCAAAA
ATAAAATTGCAAATAGAAATTTTGTTCCTTTTTTATTGGTTTTTTATTGGTGGGAAAAA
TATTCCTCTAACTGCAAAAGGGTTAATTATGTTAGAGGTAGAGTCGACAAGCTT;

or

4)

GAATTCGTCGACGCTAGCTGGCTTGTTGTCCACAACCATTAAACCTTAAAAGCTTTA
AAAGCCTTATATATTCTTTTTTTTCTTATAAACTTAAACCTTAGAGGCTATTTAAG
TTGCTGATTTATATTAATTTTATTGTTCAAACATGAGAGCTTAGTACGTGAAACATGA
GAGCTTAGTACATTAGCCATGAGAGCTTAGTACATTAGCCATGAGGGTTTAGTTCAT
TAAACATGAGAGCTTAGTACATTAAACATGAGAGCTTAGTACATACTATCAACAGGT
TGAAGTGTCTGATCTGTACAGTAGAATTGGTAAAGAGAGTTGTGTAAAATATTGAGTT
CGCACATCTTGTTGTCTGATTATTGATTTTTTGCCGAAACCATTTGATCATATGACAAG
ATGTGTATCTACCTTAACTTAATGATTTTGATAAAAATCATTAGGTACCAATTACATT
GCTTGCAATTAACCCTTTAACGGTTATAAGGATCTAGATGAGATAGAAAGATTTGGT
TTTCGGATTTGTGTTACATAAGATGCCTAAAAATAAAAATTGAGATTCAATTTTTTTTA
AACTTTTTTTTAATTGGTGGTAAGAATATTCCTCTACCTGTTTGAGAGTAATGAAAT
TGTAGTATGATTTTTCAACAACTAAAAAACAACATAAATCTCACATAATAACTTT
ATTTCAATCACACAATTGAATACCAATAGGTTGACAGTACTTACCAGCCTGCAGGTG
AAAGACCCACCTGTAGGTTTGCCAAGTTAGCTTAAGTAACGCCATTTTGCAAGGCA
TGAAAAATACATAACTGAGAATAGAGAAGTTCAGATCAAGGTTAGGAACAGAGAG
ACAGCAGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTC
AGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCCGCCCTCAGCAGTTTCTAGCGA
ACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTG

AACTAACCAATCAGTTTGCTTCTTGCTTCTGTTTGTGTGCTTCTGCTCCCTGAGCTCA
ATAAAAGAGCCCACAACCCCTCACTTGGTGGGCCAGTCCTCTGATAGACTGTGTCCC
CTGGATACCCGTACGGTACCGCTAGCGCCACCATGGATACCTGGCTCGTGTGTTGGG
CCATCTTTAGCCTGCTGAAGGCCGACTGACCGAGCCTGAAGTGACCCAGACTCCA
AGCCATCAAGTGACTCAGATGGGGCAAGAAGTCATTCTGCGTTGCGTGCCCATCAGC
AACCACCTGTACTTTTTATTGGTATCGCCAGATCCTGGGCCAGAAAGTGGAATTCCTG
GTGTCCTTCTACAACAATGAGATCTCCGAGAAGTCCGAGATCTTCGACGACCAGTTC
TCCGTGGAAAGACCCGACGGCAGCAACTTCACACTGAAGATCCGGTCTACCAAACCT
TGAGGACTCCGCTATGTATTTTTGTGCAATCACAGGTCGCGTTTCATATGAGCAATA
TTTCGGGCCGGGCACCAGGCTCACGGTCACAGAAGATCTCAATAAAGTGTTCCCCC
TGAGGTTGCGGTGTTTGAGCCGTCCAAAGCGGAGATTGCCACACACAGAAAGCGA
CTTTGGTTTGTTTGCGGACAGGCTTTTTCCCTGACCACGTAGAGCTGTCTTGGTGGGT
CAACGGCAAGGAGGTTACAGCGGTGTGTCAACGGATCCCCAGCCTCTGAAAGAAC
AGCCTGCCCTGAACGACAGCCGGTACTGCCTGAGCTCCAGACTGAGAGTGTCCGCC
ACCTTCTGGCAGAACCCCGGAACCACTTCAGATGCCAGGTGCAGTTTTACGGCCTG
AGCGAGAACGACGAGTGGACCCAGGACAGAGCCAAGCCCGTGACACAAATCGTGT
CTGCCGAAGCCTGGGGAAGAGCCGATTGCGGCATCACCAGCGCCTCCTATCACCAG
GGCGTGCTGAGCGCCACAATCCTGTACGAAATCCTGCTGGGCAAGGCCACCCTGTA
CGCCGTGCTGGTGTCTGCTCTGGTGCTGATGGCCATGGTCAAGCGGAAGGACTTTGG
CAGCGGCAGAGCCAAAAGGTCCGGGAGCGGTGCGACAACTTTAGCCTGTTGAAAC
AAGCCGGCGACGTTGAAGAGAACCCCGGACCTATGCTGCTGATCACCTCCATGCTG
GTGCTGTGGATGCAGCTGAGCCAAGTGAACGGCCAGCAAGTGATGCAGATCCCTCA
GTACCAGCACGTGCAAGAAGGCGAGGACTTCACCACCTACTGCAACAGCAGCACCA
CACTGAGCAACATCCAGTGGTACAAGCAGCGGCCTGGCGGACACCCTGTGTTTCTG
ATCCAGCTGGTCAAGTCCGGCGAAGTGAAGAAGCAGAAGCGGCTGACCTTCCAGTT
CGGCGAGGCCAAGAAGAACAGCAGCCTGCACATCACCGCCACACAGACCACAGAT
GTGGGCACCTACTTCTGCGCTGGCATCGGTAGCAGCAACACCGGTAAAGCTCATCTTT
GGGCAAGGGACAACCTTTACAAGTAAAACCAGACATCCAGAACCCCGACCCCGCCGT
GTACCAGCTGAGGGACTCCAAGTCCAGCGACAAGAGCGTGTGTCTGTTTACGGACTT
CGACAGCCAGACCAACGTGAGTCAAAGCAAGGACAGCGACGTCTACATAACGGATA
AGACCGTGCTGGACATGCGGAGCATGGACTTCAAGAGCAACAGCGCCGTGGCCTGG
TCCAACAAGAGCGACTTCGCCTGCGCCAACGCCTTCAACAACAGCATCATCCCCGA
GGACACCTTCTTCCCCAGCAGCGACGTGCCCTGCGACGTGAAACTGGTGGAGAAGT
CCTTCGAGACAGACACCAATCTGAACCTTTCAGAACCTGCTGGTGATCGTGCTGCGGA
TTCTGCTGCTGAAAGTGGCCGGCTTCAATCTGCTGATGACCCTGCGGCTGTGGAGCA
GCAGGGCTAAGAGGTCCGGCAGCGGAGCCACCAATTTTTCCCTGCTGAAACAGGCT
GGTGACGTGGAAGAAAACCCTGGCCCCATGGCGCTGCCCGTCACCGCGCTGCTGCT
GCCCCTGCGCTGCTGTTACACGCCGCTCGGCCAGAGCTTCCCACCCAGGGCACATT
CTCCAACGTGTCCACCAATGTGTGCGGAGGCGGCGGATCGTCCCAGTTCAGAGTGTC
CCCTCTGGACCGCACCTGGAACCTGGGCGAGACCGTGAGCTGAAATGTCAGGTCC
TGCTGAGCAACCCGACCTCCGGGTGCAGTTGGCTGTTCCAGCCGCGTGGTGCTGCCG
CAAGCCCTACGTTCCCTGCTTTACCTGAGCCAGAACAAAGCCCAAGGCGGCCGAGGGC
CTGGACACCCAGAGATTCTCCGGCAAGCGCCTGGGGGACACATTCGTGCTTACTTTG
AGCGATTTCCGCAGAGAGAACGAGGGCTACTATTTCTGTTCGGCGCTGAGCAATTCC
ATCATGTATTTACGCCACTTTGTGCCAGTGTTCCCTGCCTGCCAAGCCTACCACAACA
CCAGCTCCCCGTCCCCGACTCCGGCGCCTACCATCGCGAGTCAACCGTTGAGCCTG

AGGCCTGAGGCTTGTCTGGCCCGCTGCGGGGGGTGCCGTCCACACCAGGGGGCCTCGA
CTTTGCGTGCACATCTATATTTGGGCGCCTCTGGCGGGTACCTGCGGGGTGCTGCT
GCTGTCATTGGTGATTACCCTGTACTGCAATCACCGCAACCGCCGGCGGGTCTGTAA
GTGCCCACGGCCTGTGGTCAAGTCCGGTGACAAACCGTCGCTCTCGGCTCGCTACGT
GCGCGCTAAGCGCAGCGGTTCCGGGGGCCACCAACTTTTCATTGCTGAAGCAGGCCG
GTGATGTGGAGGAGAATCCAGGGGCCATGCGCCCCAGGCTTTGGCTCCTTCTTGCTG
CTCAGCTCACTGTCTTGCATGGCAACTCCGTTCTGCAGCAGACTCCCGCCTACATCA
AGGTGCAGACGAACAAGATGGTGATGCTGTCATGCGAGGCCAAGATCTCTCTTTCA
AATATGAGAATTTATTGGCTACGACAGCGCCAGGCCCCCTCCAGCGACAGCCACCA
CGAGTTCCTGGCGCTTTGGGATTCTGCTAAAGGCACCATCCATGGAGAGGAGGTGG
AACAGGAGAAGATAGCTGTCTTCCGCGACGCATCCCGCTTCATCCTGAACCTGACCA
GCGTGAAGCCGGAGGACAGCGGCATCTACTTCTGTATGATCGTTGGCTCCCCCGAGC
TGACCTTCGGCAAAGGCACCCAGCTGTCCGTGGTGGACTTCCTGCCCACCACAGCCC
AGCCAACCAAGAAATCCACCCTCAAGAAGCGCGTGTGCCGACTGCCCCGCCCTGAA
ACCCAGAAGGGCCCTCTGTGCTCCCCCATCACCCCTTGGACTGCTGGTGGCGGGAGTC
CTGGTGCTGCTCGTATCTCTGGGTGTGCCCATCCACCTGTGCTGCCGCCGCCGCCGC
GCCCCGCTGAGGTTTATGAAACAGTTTACAAGTGATAAATCGATGGAAGGGTGGC
ATCCCTGTGACCCCTCCCCAGTGCCCTCTCCTGGCCCTGGAAGTTGCCACTCCAGTGC
CCACCAGCCTTGTCTAATAAAATTAAGTTGCATCATTTTTGTCTGACTAGGTGTCCTT
CTATAATATTATGGGGTGGAGGGGGGTGGTATGGAGCAAGGGGGCAAGTTGGGAAGA
CAACCTGTAGGGCCTGCGGGGTCTATTGGGAACCAAGCTGGAGTGCAGTGGCACAA
TCTTGGCTCACTGCAATCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTC
CCGAGTTGTTGGGATTCCAGGCATGCATGACCAGGCTCAGCTAATTTTTGTTTTTTTG
GTAGAGACGGGGTTTACCATATTGGCCAGGCTGGTCTCCAACCTCCTAATCTCAGGT
GATCTACCCACCTTGGCCTCCCAAATTGCTGGGATTACAGGCGTGAACCACTGCTCC
CTTCCCTGTCCTTCTGATTACTAGTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCAC
ATCGCCACAGTCCCCGAGAAGTTGTGGGGAGGGGTCCGGCAATTGAACCGGTGCCT
AGAGAAGGTGGCGCGGGGTAAACTGGGAAAGTGATGTCGTGTAAGTGGCTCCGCCTT
TTTCCCGAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTT
TTTCGCAACGGGTTTGGCGCCAGAACACAGGTAAAGTGCCGTGTGTGGTTCCCGCGGG
CCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCCTGAATTACTTCCACCTGGCTGC
AGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGAGTTCGAGGC
CTTGCGCTTAAGGAGCCCCCTTCGCCTCGTGCTTGAGTTGAGGCCTGGCCTGGGCGCT
GGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCTGTCTCGCTGCTTTTCGATA
AGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGCGACGCTTTTTTTCTGGCAAGA
TAGTCTTGTAATGCGGGCCAAGATCTGCACACTGGTATTTTCGGTTTTTGGGGCCGC
GGGCGGCGACGGGGCCCGTGCGTCCCAGCGCACATGTTCCGGCAGGCGGGGCCTGC
GAGCGCGGCCACCGAGAATCGGACGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTG
GTGCCTGGCCTCGCGCCCGCCGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCCG
TCGGCACCAAGTTGCGTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAG
CTCAAAATGGAGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAA
AGGAAAAGGGCCTTTCCGTCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCGG
GCGCCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGTCTTTAGGTT
GGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGACTGAA
GTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTTTGAGTTTG
GATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTCTTCCATTTCA

GGTGTCTGTAAGTCCAGTGTGGTGAATTCTGCAGATATCACGGCTAGCGCCAC
CATGGGTCGGGGGCTGCTCAGGGGCCTGTGGCCGCTGCACATCGTCCTGTGGACGC
GTATCGCCAGCACGATCCCACCGCACGTTTCCAGAAAGTCGGTGAATAACGACATGATA
GTCAGTACAAACGAGTCAAGTTTCCACAAGTGTGTAAATTTTGTGATGTG
AGATTTTCCACCTGTGACAACCAGAAATCCTGCATGAGCAACTGCAGCATCACCTCC
ATCTGTGAGAAGCCACAGGAAGTCTGTGTGGCTGTATGGAGAAAGAATGACGAGAA
CATAACACTAGAGACAGTTTGCCATGACCCCAAGCTCCCCTACCATGACTTTATTCT
GGAAGATGCTGCTTCTCCAAAGTGCATTATGAAGGAGAAGAAAAAGCCTGGTGAGA
CTTTCTTCATGTGTTCTGTAGCTCTGATGAGTGCAATGACAACATCATCTTCTCAGA
AGAATATAACACCAGCAATCCTGACTTGTTGCTAGTCATATTTCAAGTGACAGGCAT
CAGCCTCCTGCCACCACTGGGAGTTGCCATATCTGTCATCATCATCTTCTACTGCTAC
CGCGTGAACCGGCAGCAGAAGGCTAGTGGTTCAGGCGCAACGAATTTCTCTTTGCTG
AAGCAGGCTGGGGATGTCGAAGAAAATCCGGGTCCAATGGTGGGCTCGCTCAACTG
CATCGTAGCAGTCTCCCAGAATATGGGCATCGGGAAGAACGGTGATTTCCCGTGGCC
CCCACTTCGCAACGAGAGCCGTTATTTCCAAAGAATGACTACAACCTCCTCCGTGGA
GGGTAAGCAGAACCTGGTCATCATGGGGAAGAAGACCTGGTTCTCTATCCCTGAAA
AAAACCGCCCCCTGAAGGGCCGCATCAACCTGGTGCTGAGCAGGGAAGTCAAGGAG
CCTCCTCAGGGCGCGCATTTTCTGAGCCGCTCATTGGATGACGCTCTCAAAGTACC
GAACAGCCGGAGCTAGCCAACAAGGTGGACATGGTGTGGATCGTCGGAGGCTCCTC
CGTGTACAAGGAGGCCATGAATCACCCCGGCCACTTGAAGCTGTTTCGTCACCCGGAT
CATGCAGGACTTCGAGTCGGACACGTTCTTTCCAGAGATTGACCTGGAGAAGTACAA
GCTGCTGCCCAGTACCCGGGAGTTCTTAGTGATGTGCAGGAGGAGAAAGGCATCA
AGTACAAATTTGAGGTGTACGAGAAGAACGACTAACGGTCCGTCCTGACCAATGCT
GGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
CAATAGCATCACAAATTTACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGT
TTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACAGGTTACCTCAGTCTCCT
AGGTACGTCTTATATCTATGAAAAAACATTCAAAGCACAACATCTAGAAGAACTT
ACCTTTTTTACCACCTCTATTGCAAAGATATGTACCGATTCTCTCGAAGTACAAAAA
ACCGCTAGTTTTCAAATTCACCTCAAGACTTTGAAAAAAAATTGAATCTGTCAATGT
CAAATAAAATCAGAAACAAATGTCATAATGTTACGTTAATGTTGTCAGGTCGAAAA
ATAAAATTGCAAATAGAAATTTTGTTCCTTTTTTATTGGTTTTTATTGGTGGGAAAAA
TATCCCTCTAACTGCAAAAGGGTTAATTATGTTAGAGGTAGAGTCGACAAGCTT.

25. The vector of claim 23 or 24, wherein the vector is a cloning vector, expression vector, or viral vector.

26. The vector of any one of claims 23-25, wherein the vector further comprises a nucleic acid sequence encoding CD8 α , CD8 β , a dominant negative TGF β receptor II (DN-TGF β RII), selectable protein marker, optionally wherein the selectable protein marker is dihydrofolate reductase (DHFR).

27. The vector of any one of claims 23-26, wherein the nucleic acid sequence encoding CD8 α , CD8 β , the DN-TGF β RII, and/or the selectable protein marker is operably linked to a nucleic acid encoding a tag.
28. The vector of any one of claims 23-27, wherein the nucleic acid encoding a tag is at the 5' upstream of the nucleic acid sequence encoding CD8 α , CD8 β , the DN-TGF β RII, and/or the selectable protein marker such that the tag is fused to the N-terminal of CD8 α , CD8 β , the DN-TGF β RII, and/or the selectable protein marker.
29. The vector of any one of claims 23-28, wherein the tag is a CD34 enrichment tag.
30. The vector of any one of claims 23-29, wherein the isolated nucleic acid encoding the binding protein, and the nucleic acid sequence encoding TCR8 α , TCR8 β , CD8 α , CD8 β , the DN-TGF β RII, and/or the selectable protein marker are interconnected with an internal ribosome entry site or a nucleic acid sequence encoding a self-cleaving peptide.
31. The vector of claim 30, wherein the self-cleaving peptide is P2A, E2A, F2A or T2A.
32. A host cell which comprises the vector of any one of claims 23-31, and/or expresses the binding protein of any one of claims 1-21, optionally wherein the cell is genetically engineered.
33. The host cell of claim 32, wherein the host cell comprises a chromosomal gene knockout of a TCR gene, an HLA gene, or both.
34. The host cell of claim 32 or 33, wherein the host cell comprises a knockout of an HLA gene selected from an α 1 macroglobulin gene, α 2 macroglobulin gene, α 3 macroglobulin gene, β 1 microglobulin gene, β 2 microglobulin gene, and combinations thereof.
35. The host cell of any one of claims 32-34, wherein the host cell comprises a knockout of a TCR gene selected from a TCR α variable region gene, TCR β variable region gene, TCR constant region gene, and combinations thereof.

36. The host cell of any one claims 32-35, wherein the host cell expresses CD8 α , CD8 β , a DN-TGF β RII, and/or a selectable protein marker, optionally wherein the selectable protein marker is DHFR.
37. The host cell of claims 36, wherein the CD8 α , CD8 β , the DN-TGF β RII, and/or the selectable protein marker is fused to a CD34 enrichment tag.
38. The host cell of any one of claims 32-37, wherein the host cell is an immune cell.
39. The host cell of any one of claims 32-38, wherein the immune cell is a cytotoxic lymphocyte, cytotoxic lymphocyte precursor cell, cytotoxic lymphocyte progenitor cell, cytotoxic lymphocyte stem cell, CD4⁺ T cell, CD8⁺ T cell, CD4/CD8 double negative T cell, gamma delta ($\gamma\delta$) T cell, natural killer (NK) cell, NK-T cell, dendritic cell, or combination thereof.
40. The host cell of claim 39, wherein the T cell is a naive T cell, central memory T cell, effector memory T cell, or a combination thereof.
41. The host cell of claim 39 or 40, wherein the T cell is a primary T cell or a cell of a T cell line.
42. The host cell of any one of claims 39-41, wherein the T cell does not express or has a lower surface expression of an endogenous TCR.
43. The host cell of any one of claims 39-42, wherein the host cell is capable of producing a cytokine or a cytotoxic molecule when contacted with a target cell that comprises a peptide-MHC (pMHC) complex comprising the HPV16 E7₁₁₋₁₉ peptide epitope in the context of an MHC molecule.
44. The host cell of claim 43, wherein the cytokine is TNF- α , IL-2, and/or IFN- γ .

45. The host cell of any one of claims 43-44, wherein the cytotoxic molecule is perforins and/or granzymes, optionally wherein the cytotoxic molecule is granzyme B.
46. The host cell of any one of claims 43-45, wherein the host cell is capable of producing a higher level of cytokine or a cytotoxic molecule when contacted with a target cell expressing HPV16 E7₁₁₋₁₉ peptide epitope.
47. The host cell of claim 46, wherein the host cell is capable of producing an at least 1.05-fold higher level of cytokine or a cytotoxic molecule.
48. The host cell of any one of claims 43-47, wherein the host cell is capable of killing a target cell that comprises a peptide-MHC (pMHC) complex comprising the HPV16 E7₁₁₋₁₉ peptide epitope in the context of an MHC molecule.
49. The host cell of claim 48, wherein the killing is determined by a killing assay.
50. The host cell of claim 48 or 49, wherein the ratio of the host cell and the target cell in the killing assay is from 20:1 to 0.625:1.
51. The host cell of any one of claims 48-50, wherein a target cell is a T2 cell pulsed with 1 µg/mL to 50 pg/mL of HPV16 E7₁₁₋₁₉ peptide.
52. The host cell of any one of claims 48-50, wherein the host cell is capable of killing a higher number of target cells when contacted with target cells expressing HPV16 E7₁₁₋₁₉ peptide epitope.
53. The host cell of claim 52, wherein the host cell is capable of killing an at least 1.05-fold higher number of target cells.

54. The host cell of any one of claims 48, 49, 52 and 53, wherein the target cell is CaSki, SCC152, or SCC090 cell line.
55. The host cell of any one of claims 43-54, wherein the HPV16 E7₁₁₋₁₉ immunogenic peptide comprises the amino acid sequence YMLDLQPET.
56. The host cell of any one of claims 43-55, wherein the MHC molecule is an MHC class I molecule.
57. The host cell of any one of claims 43-56, wherein the MHC molecule comprises an MHC alpha chain that is an HLA serotype HLA-A*02.
58. The host cell of any one of claims 43-57, wherein the HLA allele is selected from the group consisting of HLA-A*02:01, HLA-A*02:02, HLA-A*02:03, HLA-A*02:05, HLA-A*02:06, and HLA-A*02:07 allele.
59. The host cell of any one of claims 43-58, wherein the target cell is a cell line selected from the group consisting of CaSki, SCC152 and SCC090 cell lines, is a cancer cell expressing HPV16 E7₁₁₋₁₉ immunogenic peptide, or is not SiHa cell line and/or is not NCI-H1792 cell line.
60. The host cell of claim 59, wherein the cancer cell is selected from the group consisting of head & neck cancer cell, an oropharyngeal cancer cell, a cervical cancer cell, an anal cancer cell, a vaginal cancer cell, a vulval cancer cell, and a penile cancer cell.
61. The host cell of any one of claims 43-60, wherein a) the host cell does not induce T cell expansion, cytokine release, or cytotoxic killing when contacted with a target cell that comprises a peptide-MHC (pMHC) complex comprising a SPTA1, MPL, HERC1, CPAMD8, INTS4, or NUTM1 peptide epitope in the context of an MHC molecule and/or b) the host cell does not express HPV16 E7₁₁₋₁₉ antigen, is not recognized by a binding protein of any one of claims 1-21, is not of serotype HLA-A*02, and/or does not express an HLA-A*02 allele, optionally wherein the HLA-A*02 allele is HLA-A*02:01 and/or HLA-A*02:06.

62. A population of host cells of any one of claims 32-61.
63. A composition comprising: a) a binding protein according to any one of claims 1-21, b) an isolated nucleic acid according to claim 22, c) a vector according to any one of claims 23-31, d) a host cell according to any one of claims 32-61, and/or e) a population of host cells according to claim 62, and a carrier.
64. A device or kit comprising: a) a binding protein according to any one of claims 1-21, b) an isolated nucleic acid according to claim 22, c) a vector according to any one of claims 23-31, d) a host cell according to any one of claims 32-61, and/or e) a population of host cells according to claim 62, said device or kit optionally comprising a reagent to detect binding of a), d) and/or e) to a pMHC complex.
65. A method of producing a binding protein according to any one of claims 1-21, wherein the method comprises the steps of: (i) culturing a transformed host cell which has been transformed by a nucleic acid comprising a sequence encoding a binding protein according to any one of claims 1-21 under conditions suitable to allow expression of said binding protein; and (ii) recovering the expressed binding protein.
66. A method of producing a host cell expressing a binding protein according to any one of claims 1-21, wherein the method comprises the steps of: (i) introducing a nucleic acid comprising a sequence encoding a binding protein according to any one of claims 1-21 into the host cell; (ii) culturing the transformed host cell under conditions suitable to allow expression of said binding protein.