

ADENOVIRUS POLYNUCLEOTIDES AND POLYPEPTIDES

FIELD OF THE INVENTION

5

The present invention relates to isolated polynucleotide and polypeptide sequences derived from novel chimp adenovirus ChAd155, as well as to recombinant polynucleotides, vectors, adenoviruses, cells and compositions comprising said polynucleotide and polypeptide sequences.

10

BACKGROUND OF THE INVENTION

Adenovirus has been widely used for gene transfer applications due to its ability to achieve highly efficient gene transfer in a variety of target tissues and large transgene capacity.

15

Conventionally, E1 genes of adenovirus are deleted and replaced with a transgene cassette consisting of the promoter of choice, cDNA sequence of the gene of interest and a poly A signal, resulting in a replication defective recombinant virus.

Recombinant adenoviruses are useful in gene therapy and as vaccines. Viral vectors based on chimpanzee adenovirus represent an alternative to the use of human derived Ad vectors for the development of genetic vaccines. Adenoviruses isolated from chimpanzees are closely related to adenoviruses isolated from humans as demonstrated by their efficient propagation in cells of human origin. However, since human and chimp adenoviruses are close relatives, serologic cross reactivity between the two virus species is possible.

20

There is a demand for vectors which effectively deliver molecules to a target and minimize the effect of pre-existing immunity to selected adenovirus serotypes in the population. One aspect of pre-existing immunity that is observed in humans is humoral immunity, which can result in the production and persistence of antibodies that are specific for adenoviral proteins. The humoral response elicited by adenovirus is mainly directed against the three major structural capsid proteins: fiber, penton and hexon.

25

Vectors, compositions and methods of the present invention may have one or more following improved characteristics over the prior art, including but not limited to higher productivity, improved immunogenicity and increased transgene expression.

30

SUMMARY OF THE INVENTION

35

40

There is provided an isolated polynucleotide, wherein the polynucleotide encodes a polypeptide selected from the group consisting of:

- (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
- 5 (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1, and
- (c) a polypeptide having the amino acid sequence according to SEQ ID NO: 3.

10 Also provided is a recombinant polynucleotide comprising a polynucleotide selected from the group consisting of:

- (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
- (b) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1, and
- 15 (c) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3.

20

Also provided is a recombinant vector comprising a polynucleotide selected from the group consisting of:

- (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
- 25 (b) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1, and
- (c) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3.

30

Also provided is a recombinant adenovirus comprising at least one polynucleotide or polypeptide selected from the group consisting of:

- (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
- 35 (b) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1,
- 40 (c) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3,

- (d) a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
(e) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1, and
5 (f) a polypeptide having the amino acid sequence according to SEQ ID NO: 3.

Also provided is a composition comprising at least one of the following:

- (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
10 (b) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1,
(c) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3,
15 (d) a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
(e) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1,
20 (f) a polypeptide having the amino acid sequence according to SEQ ID NO: 3,
(g) a vector comprising a polynucleotide as described in (a), (b) or (c) above, and
(h) a recombinant adenovirus comprising a polynucleotide as described in (a), (b) or (c) above,
and a pharmaceutically acceptable excipient.

25

Also provided is a cell comprising at least one of the following:

- (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
(b) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1,
30 (c) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3,
(d) a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
35 (e) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1,
(f) a polypeptide having the amino acid sequence according to SEQ ID NO: 3,
40 (g) a vector comprising a polynucleotide as described in (a), (b) or (c) above, and

(h) a recombinant adenovirus comprising a polynucleotide as described in (a), (b) or (c) above,
and a pharmaceutically acceptable excipient.

- 5 Also provided is an isolated adenoviral polypeptide selected from the group consisting of:
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1,
(b) a functional derivative of a polypeptide having the amino acid sequence according to
SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at
least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1, and
10 (c) a polypeptide having the amino acid sequence according to SEQ ID NO: 3.

Also provided is an isolated polynucleotide, vector, recombinant adenovirus, composition
or cell comprising or consisting of the sequence according to SEQ ID NO: 6.

15 DESCRIPTION OF THE FIGURES

Figure 1A-C - Alignment of fiber protein sequences from the indicated simian
adenoviruses.

- 20 ChAd3 (SEQ ID NO:27)
PanAd3 (SEQ ID NO:28)
ChAd17 (SEQ ID NO:29)
ChAd19 (SEQ ID NO:30)
ChAd24 (SEQ ID NO:31)
ChAd155 (SEQ ID NO:1)
25 ChAd11 (SEQ ID NO:32)
ChAd20 (SEQ ID NO:33)
ChAd31 (SEQ ID NO:34)
PanAd1 (SEQ ID NO:35)
PanAd2 (SEQ ID NO:36)

- 30 Figure 2 – Flow diagram for production of specific ChAd155 BAC and plasmid vectors
Figure 3 – Species C BAC Shuttle #1365 schematic
Figure 4 – pArsChAd155 Ad5E4orf6-2 (#1490) schematic
Figure 5 – pChAd155/RSV schematic
Figure 6 – BAC ChAd155/RSV schematic
35 Figure 7 – Productivity of ChAd3 and ChAd155 vectors expressing an HIV Gag
transgene (Experiment 1)
Figure 8 – Productivity of ChAd3 and ChAd155 vectors expressing an HIV Gag
transgene (Experiment 2)
Figure 9 – Productivity of PanAd3 and ChAd155 vectors expressing RSV transgene
40 Figure 10 – Expression levels of ChAd3 and ChAd155 vectors expressing an HIV Gag
transgene

- Figure 11 – Expression levels of PanAd3 and ChAd155 vectors expressing an HIV Gag transgene – Western Blot
- Figure 12 – Immunogenicity of ChAd3 and ChAd155 vectors expressing an HIV Gag transgene – IFN-gamma ELISpot
- 5 Figure 13 – Immunogenicity of PanAd3 and ChAd155 vectors expressing an HIV Gag transgene – IFN-gamma ELISpot

DESCRIPTION OF THE SEQUENCES

- 10 SEQ ID NO: 1 - Polypeptide sequence of ChAd155 fiber
 SEQ ID NO: 2 - Polynucleotide sequence encoding ChAd155 fiber
 SEQ ID NO: 3 - Polypeptide sequence of ChAd155 penton
 SEQ ID NO: 4 - Polynucleotide sequence encoding ChAd155 penton
 SEQ ID NO: 5 - Polypeptide sequence of ChAd155 hexon
 15 SEQ ID NO: 6 - Polynucleotide sequence encoding ChAd155 hexon
- SEQ ID NO: 7 - Polynucleotide sequence encoding ChAd155#1434
 SEQ ID NO: 8 - Polynucleotide sequence encoding ChAd155#1390
 SEQ ID NO: 9 - Polynucleotide sequence encoding ChAd155#1375
 20 SEQ ID NO: 10 - Polynucleotide sequence encoding wild type ChAd155
 SEQ ID NO: 11 - Polynucleotide sequence encoding ChAd155/RSV
 SEQ ID NO: 12 - Polynucleotide sequence encoding the CASI promoter
 SEQ ID NO: 13 - Ad5orf6 primer 1 polynucleotide sequence
 SEQ ID NO: 14 - Ad5orf6 primer 2 polynucleotide sequence
 25 SEQ ID NO: 15 - BAC/CHAd155 Δ E1_TetO hCMV RpsL-Kana primer 1
 polynucleotide sequence
 SEQ ID NO: 16 - BAC/CHAd155 Δ E1_TetO hCMV RpsL-Kana (#1375) primer 2
 polynucleotide sequence
- SEQ ID NO: 17 - 1021-FW E4 Del Step1 primer polynucleotide sequence
 30 SEQ ID NO: 18 - 1022-RW E4 Del Step1 primer polynucleotide sequence
 SEQ ID NO: 19 - 1025-FW E4 Del Step2 primer polynucleotide sequence
 SEQ ID NO: 20 - 1026-RW E4 Del Step2 primer polynucleotide sequence
 SEQ ID NO: 21 - 91-SubMonte FW primer polynucleotide sequence
 SEQ ID NO: 22 - 890-BghPolyA RW primer polynucleotide sequence
 35 SEQ ID NO: 23 - CMVfor primer polynucleotide sequence
 SEQ ID NO: 24 - CMVrev primer polynucleotide sequence
 SEQ ID NO: 25 - CMVFAM-TAMRA qPCR probe polynucleotide sequence
 SEQ ID NO: 26 - Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element
 (WPRE) polynucleotide sequence
- 40 SEQ ID NO: 27 - Amino acid sequence for the fiber protein of ChAd3
 SEQ ID NO: 28 - Amino acid sequence for the fiber protein of PanAd3

- SEQ ID NO: 29 - Amino acid sequence for the fiber protein of ChAd17
 SEQ ID NO: 30 - Amino acid sequence for the fiber protein of ChAd19
 SEQ ID NO: 31 - Amino acid sequence for the fiber protein of ChAd24
 SEQ ID NO: 32 - Amino acid sequence for the fiber protein of ChAd11
 5 SEQ ID NO: 33 - Amino acid sequence for the fiber protein of ChAd20
 SEQ ID NO: 34 - Amino acid sequence for the fiber protein of ChAd31
 SEQ ID NO: 35 - Amino acid sequence for the fiber protein of PanAd1
 SEQ ID NO: 36 - Amino acid sequence for the fiber protein of PanAd2
 SEQ ID NO: 37 - RSV F Δ TM amino acid sequence
 10 SEQ ID NO: 38 - HIV Gag polynucleotide sequence

DETAILED DESCRIPTION OF THE INVENTION

Adenovirus

15

Adenoviruses have a characteristic morphology with an icosahedral capsid comprising three major proteins, hexon (II), penton base (III) and a knobbed fiber (IV), along with a number of other minor proteins, VI, VIII, IX, IIIa and IVa2. The virus genome is a linear, double-stranded DNA. The virus DNA is intimately associated with the highly basic protein

20

VII and a small peptide pX (formerly termed mu). Another protein, V, is packaged with this DNA-protein complex and provides a structural link to the capsid via protein VI. The virus also contains a virus-encoded protease, which is necessary for processing of some of the structural proteins to produce mature infectious virus.

25

The adenoviral genome is well characterized. There is general conservation in the overall organization of the adenoviral genome with respect to specific open reading frames being similarly positioned, e.g. the location of the E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5 genes of each virus. Each extremity of the adenoviral genome comprises a sequence known as an inverted terminal repeat (ITR), which is necessary for viral replication. The virus also comprises a virus-encoded protease, which is necessary for processing some of the structural proteins required to produce infectious virions. The structure of the adenoviral genome is described on the basis of the order in which the viral genes are expressed following host cell transduction. More specifically, the viral genes are referred to as early (E) or late (L) genes according to whether transcription occurs prior to or after onset of DNA replication. In the early phase of transduction, the E1A, E1B, E2A, E2B, E3 and E4 genes of adenovirus are expressed to prepare the host cell for viral replication. During the late phase of infection, expression of the late genes L1-L5, which encode the structural components of the virus particles, is activated.

30

35

40

Adenoviruses are species-specific and different serotypes, i.e., types of viruses that are not cross-neutralized by antibodies, have been isolated from a variety of mammalian

species. For example, more than 50 serotypes have been isolated from humans which are divided into six subgroups (A–F; B is subdivided into B1 and B2) based on sequence homology and on their ability to agglutinate red blood cells (Tatsis and Ertl *Molecular Therapy* (2004) 10:616–629). Numerous adenoviruses have been isolated from nonhuman
5 simians such as chimpanzees, bonobos, rhesus macaques and gorillas, and they are classified into the same human groups based on phylogenetic relationships based on hexon or fiber sequences (Colloca *et al.* (2012) *Science Translational Medicine* 4:1-9; Roy *et al.* (2004) *Virology* 324: 361-372; Roy *et al.* (2010) *Journal of Gene Medicine* 13:17-25).

10 **Adenovirus Capsid Proteins Including the Fiber Protein and Polynucleotides Encoding These Proteins**

As outlined above, the adenoviral capsid comprises three major proteins, hexon, penton and fiber. The hexon accounts for the majority of the structural components of the capsid,
15 which consists of 240 trimeric hexon capsomeres and 12 penton bases. The hexon has three conserved double barrels, while the top has three towers, each tower containing a loop from each subunit that forms most of the capsid. The base of hexon is highly conserved between adenoviral serotypes, while the surface loops are variable (Tatsis and Ertl *Molecular Therapy* (2004) 10:616–629).

20 Penton is another adenoviral capsid protein that forms a pentameric base to which fiber attaches. The trimeric fiber protein protrudes from the penton base at each of the 12 vertices of the capsid and is a knobbed rod-like structure. A remarkable difference in the surface of adenovirus capsids compared to that of most other icosahedral viruses is the
25 presence of the long, thin fiber protein. The primary role of the fiber protein is the tethering of the viral capsid to the cell surface via its interaction with a cellular receptor.

The fiber proteins of many adenovirus serotypes share a common architecture: an N-terminal tail, a central shaft made of repeating sequences, and a C-terminal globular knob
30 domain (or “head”). The central shaft domain consists of a variable number of beta-repeats. The beta-repeats connect to form an elongated structure of three intertwined spiralling strands that is highly rigid and stable. The shaft connects the N-terminal tail with the globular knob structure, which is responsible for interaction with the target cellular
35 receptor. The globular nature of the adenovirus knob domain presents large surfaces for binding the receptor laterally and apically. The effect of this architecture is to project the receptor-binding site far from the virus capsid, thus freeing the virus from steric constraints presented by the relatively flat capsid surface.

40 Although fibers of many adenovirus serotypes have the same overall architecture, they have variable amino acid sequences that influence their function as well as structure. For example, a number of exposed regions on the surface of the fiber knob present an easily

adaptable receptor binding site. The globular shape of the fiber knob allows receptors to bind at the sides of the knob or on top of the fiber knob. These binding sites typically lie on surface-exposed loops connecting beta-strands that are poorly conserved among human adenoviruses. The exposed side chains on these loops give the knob a variety of surface features while preserving the tertiary and quaternary structure. For example, the electrostatic potential and charge distributions at the knob surfaces can vary due to the wide range of isoelectric points in the fiber knob sequences, from pI approximately 9 for Ad 8, Ad 19, and Ad 37 to approximately 5 for subgroup B adenoviruses. As a structurally complex virus ligand, the fiber protein allows the presentation of a variety of binding surfaces (knob) in a number of orientations and distances (shaft) from the viral capsid.

One of the most obvious variations between some serotypes is fiber length. Studies have shown that the length of the fiber shaft strongly influences the interaction of the knob and the virus with its target receptors. Further, fiber proteins between serotypes can also vary in their ability to bend. Although beta-repeats in the shaft form a highly stable and regular structure, electron microscopy (EM) studies have shown distinct hinges in the fiber. Analysis of the protein sequence from several adenovirus serotype fibers pinpoints a disruption in the repeating sequences of the shaft at the third beta-repeat from the N-terminal tail, which correlates strongly with one of the hinges in the shaft, as seen by EM. The hinges in the fiber allow the knob to adopt a variety of orientations relative to the virus capsid, which may circumvent steric hindrances to receptor engagement requiring the correct presentation of the receptor binding site on the knob. For example, the rigid fibers of subgroup D Ads thus require a flexible receptor or one prepositioned for virus attachment, as they are unable to bend themselves. (Nicklin et al *Molecular Therapy* 2005 12:384–393)

The identification of specific cell receptors for different Ad serotypes and the knowledge of how they contribute to tissue tropism have been achieved through the use of fiber pseudotyping technology. Although Ads of some subgroups use CAR as a primary receptor, it is becoming clear that many Ads use alternate primary receptors, leading to vastly different tropism in vitro and in vivo. The fibers of these serotypes show clear differences in their primary and tertiary structures, such as fiber shaft rigidity, the length of the fiber shaft, and the lack of a CAR binding site and/or the putative HSPG binding motif, together with the differences in net charge within the fiber knob. Pseudotyping Ad 5 particles with an alternate fiber shaft and knob therefore provides an opportunity to remove important cell binding domains and, in addition, may allow more efficient (and potentially more cell-selective) transgene delivery to defined cell types compared to that achieved with Ad 5. Neutralization of fiber-pseudotyped Ad particles may also be reduced if the fibers used are from Ads with lower seroprevalence in humans or experimental models, a situation that favours successful administration of the vector (Nicklin et al *Molecular Therapy* (2005) 12:384–393). Furthermore, full length fiber as well as isolated fiber knob

regions, but not hexon or penton alone, are capable of inducing dendritic cell maturation and are associated with induction of a potent CD8+ T cell response (Molinier-Frenkel et al. *J. Biol. Chem.* (2003) 278:37175-37182). Taken together, adenoviral fiber plays an important role in at least receptor-binding and immunogenicity of adenoviral vectors.

5

Illustrating the differences between the fiber proteins of Group C simian adenoviruses is the alignment provided in Figure 1. A striking feature is that the fiber sequences of these adenoviruses can be broadly grouped into having a long fiber, such as ChAd155, or a short fiber, such as ChAd3. This length differential is due to a 36 amino acid deletion at approximately position 321 in the short fiber relative to the long fiber. In addition, there are a number of amino acid substitutions that differ between the short versus long fiber subgroup yet are consistent within each subgroup. While the exact function of these differences have not yet been elucidated, given the function and immunogenicity of fiber, they are likely to be significant. It has been shown that one of the determinants of viral tropism is the length of the fiber shaft. It has been demonstrated that an Ad5 vector with a shorter shaft has a lower efficiency of binding to CAR receptor and a lower infectivity (Ambriović-Ristov A. et al.: *Virology*. (2003) 312(2):425-33): It has been speculated that this impairment is the results of an increased rigidity of the shorter fiber leading to a less efficient attachment to the cell receptor (Wu, E et al.: *J Virol.* (2003) 77(13): 7225–7235). These studies may explain the improved properties of ChAd155 carrying a longer and more flexible fiber in comparison with the previously described ChAd3 and PanAd3 carrying a fiber with a shorter shaft.

In one aspect of the invention there is provided isolated fiber, penton and hexon capsid polypeptides of chimp adenovirus ChAd155 and isolated polynucleotides encoding the fiber, penton and hexon capsid polypeptides of chimp adenovirus ChAd155.

All three capsid proteins are expected to contribute to low seroprevalence and can, thus, be used independently from each other or in combination to suppress the affinity of an adenovirus to preexisting neutralizing antibodies, e.g. to manufacture a recombinant adenovirus with a reduced seroprevalence. Such a recombinant adenovirus may be a chimeric adenovirus with capsid proteins from different serotypes with at least a fiber protein from ChAd155.

The ChAd155 fiber polypeptide sequence is provided in SEQ ID NO: 1.
The ChAd155 penton polypeptide sequence is provided in SEQ ID NO: 3.
The ChAd155 hexon polypeptide sequence is provided in SEQ ID NO: 5.

Polypeptides, Recombinant Adenoviruses, Compositions or Cells Comprising Polypeptide Sequences of ChAd155 Fiber or a Functional Derivative Thereof

40

Suitably the isolated polypeptide, recombinant adenovirus, composition or cell of the invention comprises a polypeptide having the amino acid sequence according to SEQ ID NO: 1.

5 Suitably the polypeptide, recombinant adenovirus, composition or cell of the invention comprises a polypeptide which is a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Suitably the functional derivative of a polypeptide having the
10 amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 80% identical, such as at least 85.0% identical, , such as at least 90% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0% identical, such as at least 99.2% identical, such as at least 99.4% identical,
15 such as 99.5% identical, such as at least 99.6% identical, such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more
20 suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 1.

25 Suitably the polypeptide, recombinant adenovirus, composition or cell according to the invention further comprises:

(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to
30 SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3, and/or
(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to
35 SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3 has an amino acid sequence which is at least 60.0%, such as
40 at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0%

- identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 3.
- 5 Alternatively the functional derivative has no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 3.
- 10
- 15 Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.
- 20 Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 450, more suitably no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.
- 25
- 30

35 Polypeptides, Recombinant Adenoviruses, Compositions or Cells Comprising Polypeptide Sequences of ChAd155 Penton

- Suitably the polypeptide, recombinant adenovirus, composition or cell of the invention comprises a polypeptide having the amino acid sequence according to SEQ ID NO: 3.
- 40 Suitably the polypeptide, recombinant adenovirus, composition or cell of the invention further comprises:

(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1

5 and/or

(a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 60% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

10

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 60.0% identical, such as at least 70.0% identical, such as at least 80.0% identical, such as at least 85.0% identical, such as at least 87.0% identical, such as at least 89.0% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0% identical, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 1.

25

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 95.0%, such as at least 97.0%, such as at least 99.0%, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO:5. Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 450, more suitably no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more

35

40

than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.

5 Isolated Polynucleotides, Vectors, Recombinant Adenoviruses, Compositions or Cells comprising Polynucleotides Encoding ChAd155 Fiber or a Functional Derivative Thereof

Suitably the isolated polynucleotide, vector, recombinant adenovirus, composition or cell of the invention comprises a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1. Suitably the polynucleotide has a sequence according to SEQ ID NO: 2.

Alternatively, the polynucleotide, vector, recombinant adenovirus, composition or cell of the invention comprises a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 80% identical, such as at least 85.0% identical, such as at least 90% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0% identical, such as at least 99% identical, such as at least 99.4% identical, such as at least 99.6% identical, such as at least 99.8% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 1.

Suitably the polynucleotide, vector, recombinant adenovirus, composition or cell of the invention further comprises a polynucleotide encoding:

- (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; or
- (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 3, wherein the functional derivative has an amino acid sequence which is at least 50.0% identical over its entire length to the amino acid sequence of SEQ ID NO: 3,
- and/or
- (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or

(b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

5 Suitably the functional derivative of the polypeptide having the amino acid sequence according to SEQ ID NO: 3 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99%, such as at least 99.4%, such as at least 99.6%, such as at least 99.8% identical over its entire length to the amino acid sequence of SEQ ID NO: 3. Alternatively the functional derivative has no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 3.

Suitably the functional derivative of the polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 95.0%, such as at least 97.0%, such as at least 98.0%, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 5. Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 350, more suitably no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.

40 Isolated Polynucleotides, Vectors, Recombinant Adenoviruses, Compositions or Cells comprising Polynucleotides Encoding ChAd155 Penton

Suitably the isolated polynucleotide, vector, recombinant adenovirus, composition or cell of the invention comprises a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3. Suitably the polynucleotide has a sequence according to SEQ ID NO: 4.

Suitably the polynucleotide, vector, recombinant adenovirus, composition or cell of the invention further comprises a polynucleotide encoding:

- (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; or
- (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 1 and/or
- (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; or
- (b) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5, wherein the functional derivative has an amino acid sequence which is at least 50% identical over its entire length to the amino acid sequence of SEQ ID NO: 5.

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1 has an amino acid sequence which is at least 60.0% identical, such as at least 70.0% identical, such as at least 80.0% identical, such as at least 85.0% identical, such as at least 87.0% identical, such as at least 89.0% identical, such as at least 91.0% identical, such as at least 93.0% identical, such as at least 95.0% identical, such as at least 97.0% identical, such as at least 98.0% identical, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 1. Alternatively the functional derivative has no more than 130, more suitably no more than 120, more suitably no more than 110, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 1.

Suitably the functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 5 has an amino acid sequence which is at least 60.0%, such as at least 70.0%, such as at least 80.0%, such as at least 85.0%, such as at least 90.0%, such as at least 95.0%, such as at least 97.0%, such as at least 98.0%, such as at least 99.0%, such as at least 99.2%, such as at least 99.4%, such as 99.5% identical, such as

at least 99.6%, such as 99.7% identical such as at least 99.8% identical, such as 99.9% identical over its entire length to the amino acid sequence of SEQ ID NO: 5. Alternatively the functional derivative has no more than 500, more suitably no more than 400, more suitably no more than 450, more suitably no more than 300, more suitably no more than 250, more suitably no more than 200, more suitably no more than 150, more suitably no more than 125, more suitably no more than 100, more suitably no more than 90, more suitably no more than 80, more suitably no more than 70, more suitably no more than 60, more suitably no more than 50, more suitably no more than 40, more suitably no more than 30, more suitably no more than 20, more suitably no more than 10, more suitably no more than 5, more suitably no more than 4, more suitably no more than 3, more suitably no more than 2, more suitably no more than 1 addition(s), deletion(s) or substitutions(s) compared to SEQ ID NO: 5.

15 ChAd155 Backbones

The invention provides isolated polynucleotide sequences of chimp adenovirus ChAd155, including that of wild type, unmodified ChAd155 (SEQ ID NO: 10) and modified backbone constructs of ChAd155. These modified backbone constructs include ChAd155#1434 (SEQ ID NO: 7), ChAd155#1390 (SEQ ID NO: 8) and ChAd155#1375 (SEQ ID NO: 9). ChAd155 backbones may be used in the construction of recombinant replication-competent or replication-incompetent adenoviruses for example for the delivery of transgenes.

25 Annotation of the ChAd155 wild type sequence (SEQ ID NO: 10) sequence is provided below.

```

LOCUS       ChAd155                      37830 bp    DNA        linear        10-JUN-2015
DEFINITION  Chimp adenovirus 155, complete genome.
30 COMMENT  Annotation according to alignment of ChAd155 against the human
            Adenovirus 2 reference strain NC_001405
            Two putative ORFs in the E3 region added manually
FEATURES             Location/Qualifiers
    source           1..37830
35                 /organism="Chimpanzee adenovirus 155"
                    /mol_type="genomic DNA"
                    /acronym="ChAd155"
    repeat_region    1..101
40                 /standard_name="ITR"
                    /rpt_type=inverted
    gene             466..1622
                    /gene="E1A"
    TATA_signal      466..471
                    /gene="E1A"
45    prim_transcript 497..1622
                    /gene="E1A"
    CDS              join(577..1117,1231..1532)
                    /gene="E1A"
                    /product="E1A_280R"

```

	CDS	join(577..979,1231..1532)
		/gene="E1A"
		/product="E1A_243R"
5	polyA_signal	1600..1605
		/gene="E1A"
	gene	1662..4131
		/gene="E1B"
	TATA_signal	1662..1667
		/gene="E1B"
10	prim_transcript	1692..4131
		/gene="E1B"
	CDS	1704..2267
		/gene="E1B"
		/product="E1B_19K"
15	CDS	2009..3532
		/gene="E1B"
		/product="E1B_55K"
	gene	3571..4131
		/gene="IX"
20	TATA_signal	3571..3576
		/gene="IX"
	prim_transcript	3601..4131
		/gene="IX"
25	CDS	3628..4092
		/gene="IX"
		/product="IX"
	polyA_signal	4097..4102
		/note="E1B, IX"
30	gene	complement(4117..27523)
		/gene="E2B"
	prim_transcript	complement(4117..27494)
		/gene="E2B"
	gene	complement(4117..5896)
		/gene="IVa2"
35	prim_transcript	complement(4117..5896)
		/gene="IVa2"
	CDS	complement(join(4151..5487,5766..5778))
		/gene="IVa2"
		/product="E2B_IVa2"
40	polyA_signal	complement(4150..4155)
		/note="IVa2, E2B"
	CDS	complement(join(5257..8838,14209..14217))
		/gene="E2B"
		/product="E2B_polymerase"
45	gene	6078..34605
		/gene="L5"
	gene	6078..28612
		/gene="L4"
50	gene	6078..22658
		/gene="L3"
	gene	6078..18164
		/gene="L2"
	gene	6078..14216
		/gene="L1"
55	TATA_signal	6078..6083
		/note="L"
	prim_transcript	6109..34605
		/gene="L5"
60	prim_transcript	6109..28612
		/gene="L4"
	prim_transcript	6109..22658
		/gene="L3"
	prim_transcript	6109..18164
		/gene="L2"
65	prim_transcript	6109..14216

		/gene="L1"
	CDS	join(8038..8457,9722..9742)
		/gene="L1"
5	CDS	/product="L1_13.6K"
		complement(join(8637..10640,14209..14217))
		/gene="E2B"
		/product="E2B_pTP"
	gene	10671..10832
		/gene="VAI"
10	misc_RNA	10671..10832
		/gene="VAI"
		/product="VAI"
	gene	10902..11072
		/gene="VAII"
15	misc_RNA	10902..11072
		/gene="VAII"
		/product="VAII"
	CDS	11093..12352
		/gene="L1"
20		/product="L1_52K"
	CDS	12376..14157
		/gene="L1"
		/product="L1_pIIIIa"
25	polyA_signal	14197..14202
		/gene="L1"
	CDS	14254..16035
		/gene="L2"
		/product="L2_penton"
30	CDS	16050..16646
		/gene="L2"
		/product="L2_pVII"
	CDS	16719..17834
		/gene="L2"
		/product="L2_V"
35	CDS	17859..18104
		/gene="L2"
		/product="L2_pX"
	polyA_signal	18143..18148
		/gene="L2"
40	CDS	18196..18951
		/gene="L3"
		/product="L3_pVI"
	CDS	19063..21945
		/gene="L3"
45		/product="L3_hexon"
	CDS	21975..22604
		/gene="L3"
		/product="L3_protease"
50	polyA_signal	22630..22635
		/gene="L3"
	gene	complement(22632..27523)
		/gene="E2A"
	prim_transcript	complement(22632..27494)
		/gene="E2A"
55	gene	complement(22632..26357)
		/gene="E2A-L"
	prim_transcript	complement(22632..26328)
		/gene="E2A-L"
	polyA_signal	complement(22649..22654)
60		/note="E2A, E2A-L"
	CDS	complement(22715..24367)
		/gene="E2A"
		/note="DBP; genus-common; DBP family"
		/codon_start=1
65		/product="E2A"

	CDS	24405..26915
		/gene="L4"
		/product="L4_100k"
5	TATA_signal	complement(26352..26357)
		/gene="E2A-L"
	CDS	join(26602..26941,27147..27529)
		/gene="L4"
		/product="L4_33K"
10	CDS	26602..27207
		/gene="L4"
		/product="L4_22K"
	TATA_signal	complement(27518..27523)
		/note="E2A, E2B; nominal"
15	CDS	27604..28287
		/gene="L4"
		/product="L4_pVIII"
	gene	27969..32686
		/gene="E3B"
20	gene	27969..31611
		/gene="E3A"
	TATA_signal	27969..27974
		/note="E3A, E3B"
	prim_transcript	27998..32686
		/gene="E3B"
25	prim_transcript	27998..31611
		/gene="E3A"
	CDS	28288..28605
		/gene="E3A"
		/product="E3 ORF1"
30	polyA_signal	28594..28599
		/gene="L4"
	CDS	29103..29303
		/gene="E3A"
		/product="E3 ORF2"
35	CDS	29300..29797
		/gene="E3A"
		/product="E3 ORF3"
	CDS	29826..30731
		/gene="E3A"
40		/product="E3 ORF4"
	CDS	30728..31579
		/gene="E3A"
		/product="E3 ORF5"
45	CDS	31283..31579
		/gene="E3A"
		/product="E3 ORF6"
	polyA_signal	31578..31584
		/gene="E3A"
50	CDS	31591..31863
		/gene="E3B"
		/product="E3 ORF7"
	CDS	31866..32264
		/gene="E3B"
		/product="E3 ORF8"
55	CDS	32257..32643
		/gene="E3B"
		/product="E3 ORF9"
	polyA_signal	32659..32664
		/gene="E3B"
60	gene	complement(<32678..32838)
		/gene="U"
	CDS	complement(<32678..32838)
		/gene="U"
65		/note="exon encoding C terminus unidentified; genus-common"

```

5      CDS      /product="protein U"
          32849..34585
          /gene="L5"
          /product="L5_fiber"
polyA_signal 34581..34586
          /gene="L5"
gene         complement(34611..37520)
          /gene="E4"
10      prim_transcript complement(34611..37490)
          /gene="E4"
polyA_signal complement(34625..34630)
          /gene="E4"
CDS         complement(join(34794..35069,35781..35954))
          /gene="E4"
15      CDS         /product="E4 ORF7"
          complement(35070..35954)
          /gene="E4"
          /product="E4 ORF6"
20      CDS         complement(35875..36219)
          /gene="E4"
          /product="E4 ORF4"
CDS         complement(36235..36582)
          /gene="E4"
          /product="E4 ORF3"
25      CDS         complement(36579..36971)
          /gene="E4"
          /product="E4 ORF2"
CDS         complement(37029..37415)
          /gene="E4"
30      CDS         /product="E4 ORF1"
          complement(37515..37520)
          /gene="E4"
TATA_signal 37740..37830
repeat_region
35      /standard_name="ITR"
          /rpt_type=inverted

```

In one embodiment, fragments of the sequences of SEQ ID NO: 7, 8, 9, 10 and their complementary strands, cDNA and RNA complementary thereto are provided. Suitably, fragments are at least 15 nucleotides in length, more suitably 30 nucleotides in length, more suitably 60 nucleotides in length, more suitably 120 nucleotides in length, more suitably 240, more suitably 480 nucleotides in length and encompass functional fragments, i.e., fragments which are of biological interest. For example, a functional fragment can express a desired adenoviral product or may be useful in production of recombinant viral vectors. Such fragments include the gene sequences listed above.

Gene products of the ChAd155 adenovirus, such as proteins, enzymes, and fragments thereof, which are encoded by the adenoviral nucleic acids described herein are provided. Such proteins include those encoded by the open reading frames identified above and the proteins encoded by the polynucleotides provided in the Sequence Listing.

Further ChAd155 Polynucleotides and Polypeptides

In some embodiments the polynucleotide of the invention comprises a polynucleotide encoding a fiber polypeptide; a penton polypeptide; a hexon polypeptide and penton polypeptide; a hexon polypeptide and fiber polypeptide; penton polypeptide and fiber polypeptide; or hexon polypeptide, penton polypeptide and fiber polypeptide of the invention; and may further comprise additional adenoviral polynucleotides, suitably ChAd155 polynucleotides. Thus, suitably the polynucleotide according to the invention comprises one or more of the following, the sequence coordinates relative to SEQ ID NO:10 provided in the previous annotation:

- (a) an adenoviral 5'-inverted terminal repeat (ITR);
- 10 (b) an adenoviral E1A region, or a fragment thereof selected from among the E1A_280R and E1A_243R regions;
- (c) an adenoviral E1B or IX region, or a fragment thereof selected from among the group consisting of the E1B_19K, E1B_55K and IX regions;
- 15 (d) an adenoviral E2B region; or a fragment thereof selected from among the group consisting of the E2B_pTP, E2B_polymerase and E2B_IVa2 regions;
- (e) an adenoviral L1 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L1_13.6K, L1_52K and L1_pIIIa protein;
- 20 (f) an adenoviral L2 region or a L2 region comprising a polynucleotide encoding the penton protein of the invention, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L2_penton protein, the L2_pVII protein, the L2_V protein and the L2_pX protein;
- 25 (g) an adenoviral L3 region or a L3 region comprising a polynucleotide encoding the hexon protein of the invention, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L3_pVI protein, the L3_hexon protein and the L3_protease protein;
- (h) an adenoviral E2A region;
- 30 (i) an adenoviral L4 region, or a fragment thereof said fragment encoding an adenoviral protein selected from the group consisting of the L4_100k protein, the L4_33K protein, the L4_22K protein and protein L4_VIII;
- (j) an adenoviral E3 region, or a fragment thereof selected from the group consisting of E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, and E3 ORF9;
- 35 (k) an adenoviral L5 region or a L5 region comprising a polynucleotide encoding the L5_fiber fiber polypeptide of the invention
- (l) an adenoviral (such as Ad5) E4 region, or a fragment thereof selected from the group consisting of E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, and E4 ORF1; in particular ORF6 of said E4 region;
- 40 (m) an adenoviral 3'-ITR; and/or

(n) an adenoviral VAI or VAI_{II} RNA region, preferably an adenoviral VAI or VAI_{II} RNA region from an adenovirus other than ChAd155, more preferably from Ad5.

Definitions

5

Suitably the polynucleotides or polypeptides of the invention are isolated. An "isolated" polynucleotide is one that is removed from its original environment. For example, a naturally-occurring polynucleotide is isolated if it is separated from some or all of the coexisting materials in the natural system. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of its natural environment or if it is comprised within cDNA.

10

Suitably the polynucleotides of the invention are recombinant. Recombinant means that the polynucleotide is the product of at least one of cloning, restriction or ligation steps, or other procedures that result in a polynucleotide that is distinct from a polynucleotide found in nature. A recombinant adenovirus is an adenovirus comprising a recombinant polynucleotide. A recombinant vector is a vector comprising a recombinant polynucleotide. 'A recombinant virus' includes progeny of the original recombinant virus. 'A recombinant vector' includes replicates of the original recombinant vector. 'A recombinant polynucleotide' includes replicates of the original recombinant polynucleotide.

20

Suitably, the polypeptide sequence of the present invention contains at least one alteration with respect to a native sequence. Suitably, the polynucleotide sequences of the present invention contain at least one alteration with respect to a native sequence. For example, a polynucleotide introduced by genetic engineering techniques into a plasmid or vector derived from a different species (and often a different genus, subfamily or family) is a heterologous polynucleotide. A promoter removed from its native coding sequence and operatively linked to a coding sequence with which it is not naturally found linked is a heterologous promoter. A specific recombination site that has been cloned into a genome of a virus or viral vector, wherein the genome of the virus does not naturally contain it, is a heterologous recombination site. A heterologous nucleic acid sequence also includes a sequence naturally found in an adenoviral genome, but located at a non-native position within the adenoviral vector.

25

30

Typically, "heterologous" means derived from a genotypically distinct entity from that of the rest of the entity to which it is being compared. A heterologous nucleic acid sequence refers to any nucleic acid sequence that is not isolated from, derived from, or based upon a naturally occurring nucleic acid sequence of the adenoviral vector. "Naturally occurring" means a sequence found in nature and not synthetically prepared or modified. A sequence is "derived" from a source when it is isolated from a source but modified (e.g., by

40

deletion, substitution (mutation), insertion, or other modification), suitably so as not to disrupt the normal function of the source gene.

A "functional derivative" of a polypeptide suitably refers to a modified version of a polypeptide, e.g. wherein one or more amino acids of the polypeptide may be deleted, inserted, modified and/or substituted. A derivative of an unmodified adenoviral capsid protein is considered functional if, for example:

- (a) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a lower seroprevalence compared to an adenovirus comprising the unmodified capsid protein and/or
- (b) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a higher host cell infectivity compared to an adenovirus comprising the unmodified capsid protein and/or
- (c) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a higher immunogenicity compared to an adenovirus comprising the unmodified capsid protein and or
- (d) an adenovirus comprising the derivative capsid protein within its capsid retains substantially the same or a higher level of transgene productivity compared to an adenovirus comprising the unmodified capsid protein.

Properties (a)-(d) above may suitably be measured using the methods described in the Examples section below.

Suitably, the polypeptide, vector or recombinant adenovirus has a low seroprevalence in a human population. "Low seroprevalence" may mean having a reduced pre-existing neutralizing antibody level as compared to human adenovirus 5 (Ad5). Similarly or alternatively, "low seroprevalence" may mean less than about 20% seroprevalence, less than about 15% seroprevalence, less than about 10% seroprevalence, less than about 5% seroprevalence, less than about 4% seroprevalence, less than about 3% seroprevalence, less than about 2% seroprevalence, less than about 1% seroprevalence or no detectable seroprevalence. Seroprevalence can be measured as the percentage of individuals having a clinically relevant neutralizing titre (defined as a 50% neutralisation titer >200) using methods as described in Aste-Amézaga *et al.*, Hum. Gene Ther. (2004) 15(3):293-304.

The terms polypeptide, peptide and protein are used interchangeably herein.

The term "simian" is typically meant to encompass nonhuman primates, for example Old World monkeys, New World monkeys, apes and gibbons. In particular, simian may refer to nonhuman apes such as chimpanzees (*Pan troglodyte*), bonobos (*Pan paniscus*) and gorillas (genus *Gorilla*). Non-ape simians may include rhesus macaques (*Macaca mulatta*)

Sequence Comparison

For the purposes of comparing two closely-related polynucleotide or polypeptide sequences, the "% identity" between a first sequence and a second sequence may be calculated using an alignment program, such as BLAST® (available at blast.ncbi.nlm.nih.gov, last accessed 09 March 2015) using standard settings. The % identity is the number of identical residues divided by the number of residues in the reference sequence, multiplied by 100. The % identity figures referred to above and in the claims are percentages calculated by this methodology. An alternative definition of % identity is the number of identical residues divided by the number of aligned residues, multiplied by 100. Alternative methods include using a gapped method in which gaps in the alignment, for example deletions in one sequence relative to the other sequence, are accounted for in a gap score or a gap cost in the scoring parameter. For more information, see the BLAST® fact sheet available at ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_BLASTGuide.pdf, last accessed on 09 March 2015.

Sequences that preserve the functionality of the polynucleotide or a polypeptide encoded thereby are likely to be more closely identical. Polypeptide or polynucleotide sequences are said to be the same as or identical to other polypeptide or polynucleotide sequences, if they share 100% sequence identity over their entire length.

A "difference" between sequences refers to an insertion, deletion or substitution of a single amino acid residue in a position of the second sequence, compared to the first sequence. Two polypeptide sequences can contain one, two or more such amino acid differences. Insertions, deletions or substitutions in a second sequence which is otherwise identical (100% sequence identity) to a first sequence result in reduced percent sequence identity. For example, if the identical sequences are 9 amino acid residues long, one substitution in the second sequence results in a sequence identity of 88.9%. If the identical sequences are 17 amino acid residues long, two substitutions in the second sequence results in a sequence identity of 88.2%. If the identical sequences are 7 amino acid residues long, three substitutions in the second sequence results in a sequence identity of 57.1%. If first and second polypeptide sequences are 9 amino acid residues long and share 6 identical residues, the first and second polypeptide sequences share greater than 66% identity (the first and second polypeptide sequences share 66.7% identity). If first and second polypeptide sequences are 17 amino acid residues long and share 16 identical residues, the first and second polypeptide sequences share greater than 94% identity (the first and second polypeptide sequences share 94.1% identity). If first and second polypeptide sequences are 7 amino acid residues long and share 3 identical residues, the first and

second polypeptide sequences share greater than 42% identity (the first and second polypeptide sequences share 42.9% identity).

Alternatively, for the purposes of comparing a first, reference polypeptide sequence to a second, comparison polypeptide sequence, the number of additions, substitutions and/or deletions made to the first sequence to produce the second sequence may be ascertained. An addition is the addition of one amino acid residue into the sequence of the first polypeptide (including addition at either terminus of the first polypeptide). A substitution is the substitution of one amino acid residue in the sequence of the first polypeptide with one different amino acid residue. A deletion is the deletion of one amino acid residue from the sequence of the first polypeptide (including deletion at either terminus of the first polypeptide).

For the purposes of comparing a first, reference polynucleotide sequence to a second, comparison polynucleotide sequence, the number of additions, substitutions and/or deletions made to the first sequence to produce the second sequence may be ascertained. An addition is the addition of one nucleotide residue into the sequence of the first polynucleotide (including addition at either terminus of the first polynucleotide). A substitution is the substitution of one nucleotide residue in the sequence of the first polynucleotide with one different nucleotide residue. A deletion is the deletion of one nucleotide residue from the sequence of the first polynucleotide (including deletion at either terminus of the first polynucleotide).

Suitably substitutions in the sequences of the present invention may be conservative substitutions. A conservative substitution comprises the substitution of an amino acid with another amino acid having a chemical property similar to the amino acid that is substituted (see, for example, Stryer et al, *Biochemistry*, 5th Edition 2002, pages 44-49). Preferably, the conservative substitution is a substitution selected from the group consisting of: (i) a substitution of a basic amino acid with another, different basic amino acid; (ii) a substitution of an acidic amino acid with another, different acidic amino acid; (iii) a substitution of an aromatic amino acid with another, different aromatic amino acid; (iv) a substitution of a non-polar, aliphatic amino acid with another, different non-polar, aliphatic amino acid; and (v) a substitution of a polar, uncharged amino acid with another, different polar, uncharged amino acid. A basic amino acid is preferably selected from the group consisting of arginine, histidine, and lysine. An acidic amino acid is preferably aspartate or glutamate. An aromatic amino acid is preferably selected from the group consisting of phenylalanine, tyrosine and tryptophane. A non-polar, aliphatic amino acid is preferably selected from the group consisting of glycine, alanine, valine, leucine, methionine and isoleucine. A polar, uncharged amino acid is preferably selected from the group consisting of serine, threonine, cysteine, proline, asparagine and glutamine. In contrast to a conservative amino acid substitution, a non-conservative amino acid substitution is the

exchange of one amino acid with any amino acid that does not fall under the above-outlined conservative substitutions (i) through (v).

Vectors and Recombinant Adenovirus

5

The ChAd155 sequences of the invention are useful as therapeutic agents and in construction of a variety of vector systems, recombinant adenovirus and host cells. Suitably the term "vector" refers to a nucleic acid that has been substantially altered (e.g., a gene or functional region that has been deleted and/or inactivated) relative to a wild type sequence and/or incorporates a heterologous sequence, *i.e.*, nucleic acid obtained from a different source (also called an "insert"), and replicating and/or expressing the inserted polynucleotide sequence, when introduced into a cell (e.g., a host cell). For example, the insert may be all or part of the ChAd155 sequences described herein. In addition or alternatively, a ChAd155 vector may be a ChAd155 adenovirus comprising one or more deletions or inactivations of viral genes, such as E1 or other viral gene or functional region described herein. Such a ChAd155, which may or may not comprise a heterologous sequence, is often called a "backbone" and may be used as is or as a starting point for additional modifications to the vector.

20 A vector may be any suitable nucleic acid molecule including naked DNA, a plasmid, a virus, a cosmid, phage vector such as lambda vector, an artificial chromosome such as a BAC (bacterial artificial chromosome), or an episome. Alternatively, a vector may be a transcription and/or expression unit for cell-free *in vitro* transcription or expression, such as a T7-compatible system. The vectors may be used alone or in combination with other adenoviral sequences or fragments, or in combination with elements from non-adenoviral sequences. The ChAd155 sequences are also useful in antisense delivery vectors, gene therapy vectors, or vaccine vectors. Thus, further provided are gene delivery vectors, and host cells which contain the ChAd155 sequences.

30 The term "replication-competent" adenovirus refers to an adenovirus which can replicate in a host cell in the absence of any recombinant helper proteins comprised in the cell. Suitably, a "replication-competent" adenovirus comprises the following intact or functional essential early genes: E1A, E1B, E2A, E2B, E3 and E4. Wild type adenoviruses isolated from a particular animal will be replication competent in that animal.

35

The term "replication-incompetent" or "replication-defective" adenovirus refers to an adenovirus which is incapable of replication because it has been engineered to comprise at least a functional deletion (or "loss-of-function" mutation), *i.e.* a deletion or mutation which impairs the function of a gene without removing it entirely, e.g. introduction of artificial stop codons, deletion or mutation of active sites or interaction domains, mutation or deletion of a regulatory sequence of a gene etc, or a complete removal of a gene

40

encoding a gene product that is essential for viral replication, such as one or more of the adenoviral genes selected from E1A, E1B, E2A, E2B, E3 and E4 (such as E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, E3 ORF9, E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 and/or E4 ORF1). Particularly suitably
5 E1 and optionally E3 and/or E4 are deleted. If deleted, the aforementioned deleted gene region will suitably not be considered in the alignment when determining % identity with respect to another sequence.

The present invention provides vectors such as recombinant adenovirus that deliver a
10 protein, suitably a heterologous protein, to cells, either for therapeutic or vaccine purposes. A vector may include any genetic element including naked DNA, a phage, transposon, cosmid, episome, plasmid, or a virus. Such vectors contain DNA of ChAd155 as disclosed herein and a minigene. By "minigene" (or "expression cassette") is meant the combination
15 of a selected heterologous gene (transgene) and the other regulatory elements necessary to drive translation, transcription and/or expression of the gene product in a host cell.

Typically, a ChAd155-derived adenoviral vector is designed such that the minigene is located in a nucleic acid molecule which contains other adenoviral sequences in the region native to a selected adenoviral gene. The minigene may be inserted into an existing gene
20 region to disrupt the function of that region, if desired. Alternatively, the minigene may be inserted into the site of a partially or fully deleted adenoviral gene. For example, the minigene may be located in the site of a mutation, insertion or deletion which renders non-functional at least one gene of a genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and E4. The term "renders non-functional" means that a
25 sufficient amount of the gene region is removed or otherwise disrupted, so that the gene region is no longer capable of producing functional products of gene expression. If desired, the entire gene region may be removed (and suitably replaced with the minigene).

For example, for a production vector useful for generation of a recombinant virus, the
30 vector may contain the minigene and either the 5' end of the adenoviral genome or the 3' end of the adenoviral genome, or both the 5' and 3' ends of the adenoviral genome. The 5' end of the adenoviral genome contains the 5' cis-elements necessary for packaging and replication; i.e., the 5' ITR sequences (which function as origins of replication) and the native 5' packaging enhancer domains (that contain sequences necessary for packaging
35 linear Ad genomes and enhancer elements for the E1 promoter). The 3' end of the adenoviral genome includes the 3' cis-elements (including the ITRs) necessary for packaging and encapsidation. Suitably, a recombinant adenovirus contains both 5' and 3' adenoviral cis-elements and the minigene (suitably containing a transgene) is located between the 5' and 3' adenoviral sequences. A ChAd155-based adenoviral vector may
40 also contain additional adenoviral sequences.

Suitably, ChAd155-based vectors contain one or more adenoviral elements derived from the adenoviral ChAd155 genome of the invention. In one embodiment, the vectors contain adenoviral ITRs from ChAd155 and additional adenoviral sequences from the same adenoviral serotype. In another embodiment, the vectors contain adenoviral sequences
5 that are derived from a different adenoviral serotype than that which provides the ITRs.

As defined herein, a pseudotyped adenovirus refers to an adenovirus in which the capsid proteins of the adenovirus are from a different adenovirus than the adenovirus which provides the ITRs.
10

Further, chimeric or hybrid adenoviruses may be constructed using the adenoviruses described herein using techniques known to those of skill in the art (e.g., US 7,291,498).

ITRs and any other adenoviral sequences present in the vector of the present invention may be obtained from many sources. A variety of adenovirus strains are available from the American Type Culture Collection, Manassas, Virginia, or available by request from a variety of commercial and institutional sources. Further, the sequences of many such strains are available from a variety of databases including, e.g., PubMed and GenBank. Homologous adenovirus vectors prepared from other chimp or from human adenoviruses
15 are described in the published literature (for example, US 5,240,846). The DNA sequences of a number of adenovirus types are available from GenBank, including type Ad5 (GenBank Accession Number M73370). The adenovirus sequences may be obtained from any known adenovirus serotype, such as serotypes 2, 3, 4, 7, 12 and 40, and further including any of the presently identified human types. Similarly adenoviruses known to
20 infect nonhuman animals (e.g., simians) may also be employed in the vector constructs of this invention (e.g., US 6,083,716). The viral sequences, helper viruses (if needed), and recombinant viral particles, and other vector components and sequences employed in the construction of the vectors described herein may be obtained as described below.

30 **Sequence, Vector and Adenovirus Production**

The sequences of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Suitable production techniques are well known to those of skill in the art. Alternatively, peptides
35 can also be synthesized by well known solid phase peptide synthesis methods.

The adenoviral plasmids (or other vectors) may be used to produce adenoviral vectors. In one embodiment, the adenoviral vectors are adenoviral particles which are replication-incompetent. In one embodiment, the adenoviral particles are rendered replication-incompetent by deletions in the E1A and/or E1B genes. Alternatively, the adenoviruses
40 are rendered replication-incompetent by another means, optionally while retaining the E1A

and/or E1B genes. Similarly, in some embodiments, reduction of an immune response to the vector may be accomplished by deletions in the E2B and/or DNA polymerase genes. The adenoviral vectors can also contain other mutations to the adenoviral genome, e.g., temperature-sensitive mutations or deletions in other genes. In other embodiments, it is desirable to retain an intact E1A and/or E1B region in the adenoviral vectors. Such an intact E1 region may be located in its native location in the adenoviral genome or placed in the site of a deletion in the native adenoviral genome (e.g., in the E3 region).

In the construction of adenovirus vectors for delivery of a gene to a mammalian (such as human) cell, a range of modified adenovirus nucleic acid sequences can be employed in the vectors. For example, all or a portion of the adenovirus delayed early gene E3 may be eliminated from the adenovirus sequence which forms a part of the recombinant virus. The function of E3 is believed to be irrelevant to the function and production of the recombinant virus particle. Adenovirus vectors may also be constructed having a deletion of at least the ORF6 region of the E4 gene, and more desirably because of the redundancy in the function of this region, the entire E4 region. Still another vector of the invention contains a deletion in the delayed early gene E2A. Deletions may also be made in any of the late genes L1 to L5 of the adenovirus genome. Similarly, deletions in the intermediate genes IX and IVa2 may be useful for some purposes. Other deletions may be made in the other structural or non-structural adenovirus genes. The above discussed deletions may be used individually, i.e., an adenovirus sequence for use as described herein may contain deletions in only a single region. Alternatively, deletions of entire genes or portions thereof effective to destroy their biological activity may be used in any combination. For example, in one exemplary vector, the adenovirus sequence may have deletions of the E1 genes and the E4 gene, or of the E1, E2A and E3 genes, or of the E1 and E3 genes, or of E1, E2A and E4 genes, with or without deletion of E3, and so on. Any one or more of the E genes may suitably be replaced with an E gene (or one or more E gene open reading frames) sourced from a different strain of adenovirus. Particularly suitably the ChAd155 E1 and E3 genes are deleted and the ChAd155E4 gene is replaced with E4Ad5orf6. As discussed above, such deletions and/or substitutions may be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

An adenoviral vector lacking one or more essential adenoviral sequences (e.g., E1A, E1B, E2A, E2B, E4 ORF6, L1, L2, L3, L4 and L5) may be cultured in the presence of the missing adenoviral gene products which are required for viral infectivity and propagation of an adenoviral particle. These helper functions may be provided by culturing the adenoviral vector in the presence of one or more helper constructs (e.g., a plasmid or virus) or a packaging host cell.

Complementation of Replication-Incompetent Vectors

To generate recombinant adenoviruses deleted in any of the genes described above, the function of the deleted gene region, if essential to the replication and infectivity of the virus, must be supplied to the recombinant virus by a helper virus or cell line, i.e., a
5 complementation or packaging cell line.

Helper Viruses

Depending upon the adenovirus gene content of the viral vectors employed to carry the minigene, a helper adenovirus or non-replicating virus fragment may be used to provide
10 sufficient adenovirus gene sequences necessary to produce an infective recombinant viral particle containing the minigene. Useful helper viruses contain selected adenovirus gene sequences not present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. In one embodiment, the helper virus
15 is replication-defective and contains adenovirus genes in addition, suitably, to one or more of the sequences described herein. Such a helper virus is suitably used in combination with an E1 expressing (and optionally additionally E3 expressing) cell line.

A helper virus may optionally contain a reporter gene. A number of such reporter genes
20 are known to the art as well as described herein. The presence of a reporter gene on the helper virus which is different from the transgene on the adenovirus vector allows both the adenoviral vector and the helper virus to be independently monitored. This reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

25

Complementation Cell Lines

In many circumstances, a cell line expressing the one or more missing genes which are essential to the replication and infectivity of the virus, such as human E1, can be used to
30 transcomplement a chimp adenoviral vector. This is particularly advantageous because, due to the diversity between the chimp adenovirus sequences of the invention and the human adenovirus sequences found in currently available packaging cells, the use of the current human E1-containing cells prevents the generation of replication-competent adenoviruses during the replication and production process.

35

Alternatively, if desired, one may utilize the sequences provided herein to generate a packaging cell or cell line that expresses, at a minimum, the E1 gene from ChAd155 under the transcriptional control of a promoter for expression in a selected parent cell line. Inducible or constitutive promoters may be employed for this purpose. Examples of such
40 promoters are described in detail elsewhere in this document. A parent cell is selected for the generation of a novel cell line expressing any desired ChAd155 gene. Without

limitation, such a parent cell line may be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], HEK 293, KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, among others. These cell lines are all available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209.

5 Such E1-expressing cell lines are useful in the generation of recombinant adenovirus E1 deleted vectors. Additionally, or alternatively, cell lines that express one or more adenoviral gene products, e.g., E1A, E1B, E2A, E3 and/or E4, can be constructed using essentially the same procedures as used in the generation of recombinant viral vectors.

10 Such cell lines can be utilised to transcomplement adenovirus vectors deleted in the essential genes that encode those products, or to provide helper functions necessary for packaging of a helper-dependent virus (e.g., adeno-associated virus). The preparation of a host cell involves techniques such as assembly of selected DNA sequences.

15 In another alternative, the essential adenoviral gene products are provided in trans by the adenoviral vector and/or helper virus. In such an instance, a suitable host cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells.

20 Host cells may be selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, HEK 293 cells or Per.C6 (both of which express functional adenoviral E1) [Fallaux, FJ et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and

25 hamster.

A particularly suitable complementation cell line is the Procell92 cell line. The Procell92 cell line is based on HEK 293 cells which express adenoviral E1 genes, transfected with the Tet repressor under control of the human phosphoglycerate kinase-1 (PGK) promoter, and the G418-resistance gene (Vitelli et al. *PLOS One* (2013) 8(e55435):1-9). Procell92.S

30 is adapted for growth in suspension conditions and is useful for producing adenoviral vectors expressing toxic proteins (www.okairos.com/e/innners.php?m=00084, last accessed 13 April 2015).

35 Assembly of a Viral Particle and Transfection of a Cell Line

Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 µg to about 100 µg DNA, and preferably about 10 to about 50 µg DNA to about 1×10^4 cells to about 1×10^{13} cells, and preferably about 10^5

40 cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking

into consideration such factors as the selected vector, the delivery method and the host cells selected.

Introduction into the host cell of the vector may be achieved by any means known in the art, including transfection, and infection. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently.

Introduction of vectors into the host cell may also be accomplished using techniques known to the skilled person. Suitably, standard transfection techniques are used, e.g., CaPC transfection or electroporation.

Assembly of the selected DNA sequences of the adenovirus (as well as the transgene and other vector elements) into various intermediate plasmids, and the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques. Such techniques include conventional cloning techniques of cDNA, use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence. Standard transfection and co-transfection techniques are employed, e.g., CaPC precipitation techniques. Other conventional methods employed include homologous recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

For example, following the construction and assembly of the desired minigene-containing viral vector, the vector is transfected in vitro in the presence of a helper virus into the packaging cell line. Homologous recombination occurs between the helper and the vector sequences, which permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the recombinant viral vector particles. The resulting recombinant adenoviruses are useful in transferring a selected transgene to a selected cell. In in vivo experiments with the recombinant virus grown in the packaging cell lines, the E1-deleted recombinant adenoviral vectors of the invention demonstrate utility in transferring a transgene to a non-simian mammal, preferably a human, cell.

Transgenes

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a protein of interest. The nucleic acid coding sequence is

operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

5 The composition of the transgene sequence will depend upon the use to which the resulting vector will be put. For example, the transgene may be a therapeutic transgene or an immunogenic transgene. Alternatively, a transgene sequence may include a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, 10 CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc. These coding sequences, when associated with regulatory elements which drive their expression, 15 provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry.

20 In one embodiment, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as a therapeutic transgene or an immunogenic transgene such as proteins, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, catalytic RNAs, and antisense RNAs. One 25 example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

The transgene may be used for treatment, e.g., of genetic deficiencies, as a cancer therapeutic or vaccine, for induction of an immune response, and/or for prophylactic 30 vaccine purposes. As used herein, induction of an immune response refers to the ability of a protein to induce a T cell and/or a humoral immune response to the protein.

Regulatory Elements

35 In addition to the transgene the vector also includes conventional control elements which are operably linked to the transgene in a manner that permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression 40 control sequences that act in trans or at a distance to control the gene of interest.

Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (poly A) signals including rabbit beta-globin polyA; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (e.g., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. Among other sequences, chimeric introns may be used.

In some embodiments, the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE) (Zuffrey et al. (1999) J Virol; 73(4):2886-9) may be operably linked to the transgene. An exemplary WPRE is provided in SEQ ID NO: 26.

A "promoter" is a nucleotide sequence that permits binding of RNA polymerase and directs the transcription of a gene. Typically, a promoter is located in the 5' non-coding region of a gene, proximal to the transcriptional start site of the gene. Sequence elements within promoters that function in the initiation of transcription are often characterized by consensus nucleotide sequences. Examples of promoters include, but are not limited to, promoters from bacteria, yeast, plants, viruses, and mammals (including humans). A great number of expression control sequences, including promoters which are internal, native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

Examples of constitutive promoters include, without limitation, the TBG promoter, the retroviral Rous sarcoma virus LTR promoter (optionally with the enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer, see, e.g., Boshart et al, Cell, 41:521-530 (1985)), the CASI promoter, the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1a promoter (Invitrogen).

In some embodiments, the promoter is a CASI promoter (see, for example, WO2012/115980). The CASI promoter is a synthetic promoter which contains a portion of the CMV enhancer, a portion of the chicken beta-actin promoter, and a portion of the UBC enhancer. In some embodiments, the CASI promoter can include a nucleic acid sequence having at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or more, sequence identity to SEQ ID NO: 12. In some embodiments, the promoter comprises or consists of a nucleic acid sequence of SEQ ID NO: 12.

Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are

available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. For example, inducible promoters include the zinc-inducible sheep metallothioneine (MT) promoter and the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter. Other inducible systems include the T7 polymerase promoter system (WO 98/10088); the ecdysone insect promoter (No et al, Proc. Natl. Acad. Sci. USA, 93:3346-3351 (1996)), the tetracycline-repressible system (Gossen et al, Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)), the tetracycline-inducible system (Gossen et al, Science, 378:1766-1769 (1995), see also Harvey et al, Curr. Opin. Chem. Biol, 2:512-518 (1998)). Other systems include the FK506 dimer, VP16 or p65 using castradiol, diphenol murislerone, the RU486-inducible system (Wang et al, Nat. Biotech., 15:239-243 (1997) and Wang et al, Gene Ther., 4:432-441 (1997)) and the rapamycin-inducible system (Magari et al, J. Clin. Invest., 100:2865-2872 (1997)). The effectiveness of some inducible promoters increases over time. In such cases one can enhance the effectiveness of such systems by inserting multiple repressors in tandem, e.g., TetR linked to a TetR by an IRES.

In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

The transgene may be operably linked to a tissue-specific promoter. For instance, if expression in skeletal muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally occurring promoters (see Li et al, Nat. Biotech., 17:241-245 (1999)). Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake et al, J. Virol, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig et al, Gene Ther., 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbutnot et al., Hum. Gene Ther., 7: 1503-14 (1996)), bone osteocalcin (Stein et al, Mol. Biol. Rep., 24:185-96 (1997)); bone sialoprotein (Chen et al., J. Bone Miner. Res., 11:654-64 (1996)), lymphocytes (CD2, Hansal et al, J. Immunol, 161:1063-8 (1998); immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen et al, Cell. Mol. Neurobiol, 13:503-15 (1993)), neurofilament light-chain gene (Piccioli et al, Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)), and the neuron-specific vgf gene (Piccioli et al, Neuron, 15:373-84 (1995)), among others.

Optionally, vectors carrying transgenes encoding therapeutically useful or immunogenic products may also include selectable markers or reporter genes which may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be packaged into a viral particle) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the vector may include an origin of replication.

These vectors are generated using the techniques and sequences provided herein, in conjunction with techniques known to those of skill in the art. Such techniques include conventional cloning techniques of cDNA such as those described in texts, use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence.

15 **Therapeutics and Prophylaxis**

The recombinant ChAd155-based vectors are useful for gene transfer to a human or non-simian mammal in vitro, ex vivo, and in vivo.

The recombinant adenovirus vectors described herein can be used as expression vectors for the production of the products encoded by the heterologous transgenes in vitro. For example, the recombinant replication-incompetent adenovirus containing a transgene may be transfected into a complementation cell line as described above.

A ChAd155-derived recombinant adenoviral vector provides an efficient gene transfer vehicle that can deliver a selected transgene to a selected host cell in vivo or ex vivo even where the organism has neutralizing antibodies to one or more adenovirus serotypes. In one embodiment, the vector and the cells are mixed ex vivo; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. These techniques are particularly well suited to gene delivery for therapeutic purposes and for immunisation, including inducing protective immunity.

Immunogenic Transgenes

The recombinant ChAd155 vectors may also be administered in immunogenic compositions. An immunogenic composition as described herein is a composition comprising one or more recombinant ChAd155 vector capable of inducing an immune response, for example a humoral (e.g., antibody) and/or cell-mediated (e.g., a cytotoxic T cell) response, against a transgene product delivered by the vector following delivery to a mammal, suitably a human. A recombinant adenovirus may comprise (suitably in any of

its gene deletions) a gene encoding a desired immunogen and may therefore be used in a vaccine. The recombinant adenoviruses can be used as prophylactic or therapeutic vaccines against any pathogen for which the antigen(s) crucial for induction of an immune response and able to limit the spread of the pathogen has been identified and for which the cDNA is available.

Such vaccine or other immunogenic compositions may be formulated in a suitable delivery vehicle. Generally, doses for the immunogenic compositions are in the range defined below under 'Delivery Methods and Dosage'. The levels of immunity of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in the serum, optional booster immunizations may be desired.

Optionally, a vaccine or immunogenic composition of the invention may be formulated to contain other components, including, e.g., adjuvants, stabilizers, pH adjusters, preservatives and the like. Examples of suitable adjuvants are provided below under 'Adjuvants'. Such an adjuvant can be administered with a priming DNA vaccine encoding an antigen to enhance the antigen-specific immune response compared with the immune response generated upon priming with a DNA vaccine encoding the antigen only. Alternatively, such an adjuvant can be administered with a polypeptide antigen which is administered in an administration regimen involving the ChAd155 vectors of the invention (as described below under 'Administration Regimens').

The recombinant adenoviruses are administered in an immunogenic amount, that is, an amount of recombinant adenovirus that is effective in a route of administration to transfect the desired target cells and provide sufficient levels of expression of the selected gene to induce an immune response. Where protective immunity is provided, the recombinant adenoviruses are considered to be vaccine compositions useful in preventing infection and/or recurrent disease.

The recombinant vectors described herein are expected to be highly efficacious at inducing cytolytic T cells and antibodies directed to the inserted heterologous antigenic protein expressed by the vector.

Immunogens expressed by the inventive vectors which are useful to immunize a human or non-human animal against other pathogens include, e.g., bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell. For example, immunogens may be selected from a variety of viral families. Examples of viral families against which an immune response would be desirable include respiratory viruses such as respiratory syncytial virus (RSV) and other paramyxoviruses such as human metapneumovirus, hMPV and parainfluenza viruses (PIV).

Suitable antigens of RSV which are useful as immunogens to immunize a human or non-human animal can be selected from: the fusion protein (F), the attachment protein (G), the matrix protein (M2) and the nucleoprotein (N). The term "F protein" or "fusion protein" or "F protein polypeptide" or "fusion protein polypeptide" refers to a polypeptide or protein having all or part of an amino acid sequence of an RSV Fusion protein polypeptide. Similarly, the term "G protein" or "G protein polypeptide" refers to a polypeptide or protein having all or part of an amino acid sequence of an RSV Attachment protein polypeptide. The term "M protein" or "matrix protein" or "M protein polypeptide" refers to a polypeptide or protein having all or part of an amino acid sequence of an RSV Matrix protein and may include either or both of the M2-1 (which may be written herein as M2.1) and M2-2 gene products. Likewise, the term "N protein" or "Nucleocapsid protein" or "N protein polypeptide" refers to a polypeptide or protein having all or part of an amino acid sequence of an RSV Nucleoprotein.

Two groups of human RSV strains have been described, the A and B groups, based mainly on differences in the antigenicity of the G glycoprotein. Numerous strains of RSV have been isolated to date, any of which are suitable in the context of the antigens of the immunogenic combinations disclosed herein. Exemplary strains indicated by GenBank and/or EMBL Accession number can be found in US published application number 2010/0203071 (WO2008114149), which is incorporated herein by reference for the purpose of disclosing the nucleic acid and polypeptide sequences of RSV F and G proteins suitable for use in present invention. In an embodiment, the RSV F protein can be an ectodomain of an RSV F Protein (F Δ TM).

Exemplary M and N protein nucleic acids and protein sequences can be found, e.g., in US published application number 2014/0141042 (WO2012/089833), which are incorporated herein for purpose of disclosing the nucleic acid and polypeptide sequences of RSV M and N proteins suitable for use in present invention.

Suitably, for use with in present invention, a nucleic acid encodes an RSV F antigen and RSV, M and N antigens. More specifically, the nucleic acid encodes an RSV F Δ TM antigen and RSV M2-1 and N antigens, wherein a self-cleavage site is included between the RSV F Δ TM antigen and the RSV M2-1 and a flexible linker is included between the RSV M2-1 and N antigens. In one embodiment a suitable nucleic acid encodes the polypeptide represented by SEQ ID NO:37

In one embodiment, the immunogen may be from a retrovirus, for example a lentivirus such as the Human Immunodeficiency Virus (HIV). In such an embodiment, immunogens may be derived from HIV-1 or HIV-2.

The HIV genome encodes a number of different proteins, each of which can be immunogenic in its entirety or as a fragment when expressed by vectors of the present invention. Envelope proteins include gp120, gp41 and Env precursor gp160, for example. Non-envelope proteins of HIV include for example internal structural proteins such as the products of the gag and pol genes and other non-structural proteins such as Rev, Nef, Vif and Tat. In an embodiment the vector of the invention encodes one or more polypeptides comprising HIV Gag.

The Gag gene is translated as a precursor polyprotein that is cleaved by protease to yield products that include the matrix protein (p17), the capsid (p24), the nucleocapsid (p9), p6 and two space peptides, p2 and p1, all of which are examples of fragments of Gag.

The Gag gene gives rise to the 55-kilodalton (kD) Gag precursor protein, also called p55, which is expressed from the unspliced viral mRNA. During translation, the N terminus of p55 is myristoylated, triggering its association with the cytoplasmic aspect of cell membranes. The membrane-associated Gag polyprotein recruits two copies of the viral genomic RNA along with other viral and cellular proteins that triggers the budding of the viral particle from the surface of an infected cell. After budding, p55 is cleaved by the virally encoded protease (a product of the pol gene) during the process of viral maturation into four smaller proteins designated MA (matrix [p17]), CA (capsid [p24]), NC (nucleocapsid [p9]), and p6, all of which are examples of fragments of Gag. In one embodiment, the vectors of the present invention comprise a Gag polypeptide of SEQ ID NO: 38.

Adjuvants

An "adjuvant" as used herein refers to a composition that enhances the immune response to an immunogen. Examples of such adjuvants include but are not limited to inorganic adjuvants (e.g. inorganic metal salts such as aluminium phosphate or aluminium hydroxide), organic adjuvants (e.g. saponins, such as QS21, or squalene), oil-based adjuvants (e.g. Freund's complete adjuvant and Freund's incomplete adjuvant), cytokines (e.g. IL-1 β , IL-2, IL-7, IL-12, IL-18, GM-CSF, and INF- γ) particulate adjuvants (e.g. immuno-stimulatory complexes (ISCOMS), liposomes, or biodegradable microspheres), virosomes, bacterial adjuvants (e.g. monophosphoryl lipid A, such as 3-de-O-acylated monophosphoryl lipid A (3D-MPL), or muramyl peptides), synthetic adjuvants (e.g. non-ionic block copolymers, muramyl peptide analogues, or synthetic lipid A), synthetic polynucleotides adjuvants (e.g. polyarginine or polylysine) and immunostimulatory oligonucleotides containing unmethylated CpG dinucleotides ("CpG").

One suitable adjuvant is monophosphoryl lipid A (MPL), in particular 3-de-O-acylated monophosphoryl lipid A (3D-MPL). Chemically it is often supplied as a mixture of 3-de-O-

acylated monophosphoryl lipid A with either 4, 5, or 6 acylated chains. It can be purified and prepared by the methods taught in GB 2122204B, which reference also discloses the preparation of diphosphoryl lipid A, and 3-O-deacylated variants thereof. Other purified and synthetic lipopolysaccharides have been described (U.S. Pat. No. 6,005,099 and EP 0 729 473 B1; Hilgers et al., 1986, *Int.Arch.Allergy.Immunol.*, 79(4):392-6; Hilgers et al., 1987, *Immunology*, 60(1):141-6; and EP 0 549 074 B1).

Saponins are also suitable adjuvants (see Lacaille-Dubois, M and Wagner H, A review of the biological and pharmacological activities of saponins. *Phytomedicine* vol 2 pp 363-386 (1996)). For example, the saponin Quil A (derived from the bark of the South American tree *Quillaja Saponaria Molina*), and fractions thereof, are described in U.S. Pat. No. 5,057,540 and Kensil, *Crit. Rev. Ther. Drug Carrier Syst.*, 1996, 12:1-55; and EP 0 362 279 B1. Purified fractions of Quil A are also known as immunostimulants, such as QS21 and QS17; methods of their production is disclosed in U.S. Pat. No. 5,057,540 and EP 0 362 279 B1. Also described in these references is QS7 (a non-haemolytic fraction of Quil-A). Use of QS21 is further described in Kensil et al. (1991, *J. Immunology*, 146: 431-437). Combinations of QS21 and polysorbate or cyclodextrin are also known (WO 99/10008). Particulate adjuvant systems comprising fractions of QuilA, such as QS21 and QS7 are described in WO 96/33739 and WO 96/11711.

Another adjuvant is an immunostimulatory oligonucleotide containing unmethylated CpG dinucleotides ("CpG") (Krieg, *Nature* 374:546 (1995)). CpG is an abbreviation for cytosine-guanosine dinucleotide motifs present in DNA. CpG is known as an adjuvant when administered by both systemic and mucosal routes (WO 96/02555, EP 468520, Davis et al, *J.Immunol*, 1998, 160:870-876; McCluskie and Davis, *J.Immunol.*, 1998, 161:4463-6). CpG, when formulated into vaccines, may be administered in free solution together with free antigen (WO 96/02555) or covalently conjugated to an antigen (WO 98/16247), or formulated with a carrier such as aluminium hydroxide (Brazolot-Millan et al., *Proc. Natl. Acad. Sci., USA*, 1998, 95:15553-8).

Adjuvants such as those described above may be formulated together with carriers, such as liposomes, oil in water emulsions, and/or metallic salts (including aluminum salts such as aluminum hydroxide). For example, 3D-MPL may be formulated with aluminum hydroxide (EP 0 689 454) or oil in water emulsions (WO 95/17210); QS21 may be formulated with cholesterol containing liposomes (WO 96/33739), oil in water emulsions (WO 95/17210) or alum (WO 98/15287); CpG may be formulated with alum (Brazolot-Millan, *supra*) or with other cationic carriers.

Combinations of adjuvants may be utilized in the present invention, in particular a combination of a monophosphoryl lipid A and a saponin derivative (see, e.g., WO 94/00153; WO 95/17210; WO 96/33739; WO 98/56414; WO 99/12565; WO 99/11241),

more particularly the combination of QS21 and 3D-MPL as disclosed in WO 94/00153, or a composition where the QS21 is quenched in cholesterol-containing liposomes (DQ) as disclosed in WO 96/33739. Alternatively, a combination of CpG plus a saponin such as QS21 is an adjuvant suitable for use in the present invention. A potent adjuvant
5 formulation involving QS21, 3D-MPL & tocopherol in an oil in water emulsion is described in WO 95/17210 and is another formulation for use in the present invention. Saponin adjuvants may be formulated in a liposome and combined with an immunostimulatory oligonucleotide. Thus, suitable adjuvant systems include, for example, a combination of monophosphoryl lipid A, preferably 3D-MPL, together with an aluminium salt (e.g. as
10 described in WO00/23105). A further exemplary adjuvant comprises QS21 and/or MPL and/or CpG. QS21 may be quenched in cholesterol-containing liposomes as disclosed in WO 96/33739.

Other suitable adjuvants include alkyl Glucosaminide phosphates (AGPs) such as those
15 disclosed in WO9850399 or U.S. Pat. No. 6,303,347 (processes for preparation of AGPs are also disclosed), or pharmaceutically acceptable salts of AGPs as disclosed in U.S. Pat. No. 6,764,840. Some AGPs are TLR4 agonists, and some are TLR4 antagonists. Both are thought to be useful as adjuvants.

It has been found (WO 2007/062656, which published as US 2011/0293704 and is incorporated by reference for the purpose of disclosing invariant chain sequences) that the fusion of the invariant chain to an antigen which is comprised by an expression system used for vaccination increases the immune response against said antigen, if it is administered with an adenovirus. Accordingly, in one embodiment of the invention, the
20 immunogenic transgene may be co-expressed with invariant chain in a recombinant ChAd155 viral vector.
25

In another embodiment, the invention provides the use of the capsid of ChAd155 (optionally an intact or recombinant viral particle or an empty capsid is used) to induce an immunomodulatory effect response, or to enhance or adjuvant a cytotoxic T cell response to another active agent by delivering a ChAd155 capsid to a subject. The ChAd155 capsid can be delivered alone or in a combination regimen with an active agent to enhance the immune response thereto. Advantageously, the desired effect can be accomplished without infecting the host with an adenovirus.
30

35

Administration Regimens

Commonly, the ChAd155 recombinant adenoviral vectors will be utilized for delivery of therapeutic or immunogenic molecules (such as proteins). It will be readily understood for
40 both applications, that the recombinant adenoviral vectors of the invention are particularly well suited for use in regimens involving repeat delivery of recombinant adenoviral vectors.

Such regimens typically involve delivery of a series of viral vectors in which the viral capsids are alternated. The viral capsids may be changed for each subsequent administration, or after a pre-selected number of administrations of a particular serotype capsid (e.g. one, two, three, four or more). Thus, a regimen may involve delivery of a recombinant adenovirus with a first capsid, delivery with a recombinant adenovirus with a second capsid, and delivery with a recombinant adenovirus with a third capsid. A variety of other regimens which use the adenovirus capsids of the invention alone, in combination with one another, or in combination with other adenoviruses (which are preferably immunologically non-crossreactive) will be apparent to those of skill in the art. Optionally, such a regimen may involve administration of recombinant adenovirus with capsids of other non-human primate adenoviruses, human adenoviruses, or artificial sequences such as are described herein.

The adenoviral vectors of the invention are particularly well suited for therapeutic regimens in which multiple adenoviral-mediated deliveries of transgenes are desired, e.g., in regimens involving redelivery of the same transgene or in combination regimens involving delivery of other transgenes. Such regimens may involve administration of a ChAd155 adenoviral vector, followed by re-administration with a vector from the same serotype adenovirus. Particularly desirable regimens involve administration of a ChAd155 adenoviral vector, in which the source of the adenoviral capsid sequences of the vector delivered in the first administration differs from the source of adenoviral capsid sequences of the viral vector utilized in one or more of the subsequent administrations. For example, a therapeutic regimen involves administration of a ChAd155 vector and repeat administration with one or more adenoviral vectors of the same or different serotypes.

In another example, a therapeutic regimen involves administration of an adenoviral vector followed by repeat administration with a ChAd155 vector which has a capsid which differs from the source of the capsid in the first delivered adenoviral vector, and optionally further administration with another vector which is the same or, preferably, differs from the source of the adenoviral capsid of the vector in the prior administration steps. These regimens are not limited to delivery of adenoviral vectors constructed using the ChAd155 sequences. Rather, these regimens can readily utilize other adenoviral sequences, including, without limitation, other adenoviral sequences including other non-human primate adenoviral sequences, or human adenoviral sequences, in combination with the ChAd155 vectors.

In a further example, a therapeutic regimen may involve either simultaneous (such as co-administration) or sequential (such as a prime-boost) delivery of (i) one or more ChAd155 adenoviral vectors and (ii) a further component such as non-adenoviral vectors, non-viral vectors, and/or a variety of other therapeutically useful compounds or molecules such as antigenic proteins optionally simultaneously administered with adjuvant. Examples of co-

administration include homo-lateral co-administration and contra-lateral co-administration (further described below under 'Delivery Methods and Dosage').

5 Suitable non-adenoviral vectors for use in simultaneous or particularly in sequential delivery (such as prime-boost) with one or more ChAd155 adenoviral vectors include one or more poxviral vectors. Suitably, the poxviral vector belongs to the subfamily chordopoxvirinae, more suitably to a genus in said subfamily selected from the group consisting of orthopox, parapox, yatapox, avipox (suitably canarypox (ALVAC) or fowlpox (FPV)) and molluscipox. Even more suitably, the poxviral vector belongs to the orthopox
10 and is selected from the group consisting of vaccinia virus, NYVAC (derived from the Copenhagen strain of vaccinia), Modified Vaccinia Ankara (MVA), cowpoxvirus and monkeypox virus. Most suitably, the poxviral vector is MVA.

"Simultaneous" administration suitably refers to the same ongoing immune response.
15 Preferably both components are administered at the same time (such as simultaneous administration of both DNA and protein), however, one component could be administered within a few minutes (for example, at the same medical appointment or doctor's visit), within a few hours. Such administration is also referred to as co-administration. In some embodiments, co-administration may refer to the administration of an adenoviral vector, an
20 adjuvant and a protein component. In other embodiments, co-administration refers to the administration of an adenoviral vector and another viral vector, for example a second adenoviral vector or a poxvirus such as MVA. In other embodiments, co-administration refers to the administration of an adenoviral vector and a protein component, which is optionally adjuvanted.

25 A prime-boost regimen may be used. Prime-boost refers to two separate immune responses: (i) an initial priming of the immune system followed by (ii) a secondary or boosting of the immune system many weeks or months after the primary immune response has been established.

30 Such a regimen may involve the administration of a recombinant ChAd155 vector to prime the immune system to second, booster, administration with a traditional antigen, such as a protein (optionally co-administered with adjuvant), or a recombinant virus carrying the sequences encoding such an antigen (e.g., WO 00/11140). Alternatively, an immunization
35 regimen may involve the administration of a recombinant ChAd155 vector to boost the immune response to a vector (either viral or DNA-based) encoding an antigen. In another alternative, an immunization regimen involves administration of a protein followed by booster with a recombinant ChAd155 vector encoding the antigen. In one example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or
40 other organism from which the antigen is derived. In another embodiment, the prime-boost regimen provides a therapeutic effect that can be measured using conventional

assays for detection of the presence of the condition for which therapy is being administered.

5 Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming composition to the subject. The administration of the boosting composition is accomplished using an effective amount of a boosting composition containing or capable of delivering the same antigen or a different antigen as administered by the priming vaccine. The boosting composition may be composed of a recombinant viral vector derived from the same viral source or from another source. Alternatively, the
10 boosting composition can be a composition containing the same antigen as encoded in the priming vaccine, but in the form of a protein, which composition induces an immune response in the host. The primary requirements of the boosting composition are that the antigen of the composition is the same antigen, or a cross-reactive antigen, as that encoded by the priming composition.

15

Delivery Methods and Dosage

The vector may be prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline; isotonic
20 salts solution or other formulations that will be apparent to those skilled in the art. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The compositions described herein may be administered to a mammal in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes.

25

In some embodiments, the recombinant adenovirus of the invention is administered to a subject by intramuscular injection, intravaginal injection, intravenous injection, intraperitoneal injection, subcutaneous injection, epicutaneous administration, intradermal administration, nasal administration or oral administration.

30

If the therapeutic regimen involves co-administration of one or more ChAd155 adenoviral vectors and a further component, each formulated in different compositions, they are favourably administered co-locally at or near the same site. For example, the components can be administered (e.g. via an administration route selected from
35 intramuscular, transdermal, intradermal, sub-cutaneous) to the same side or extremity ("co-lateral" administration) or to opposite sides or extremities ("contra-lateral" administration).

Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For
40 example, a therapeutically effective adult human or veterinary dosage of the viral vector

generally contains 1×10^5 to 1×10^{15} viral particles, such as from 1×10^8 to 1×10^{12} (e.g., 1×10^8 , 2.5×10^8 , 5×10^8 , 1×10^9 , 1.5×10^9 , 2.5×10^9 , 5×10^9 , 1×10^{10} , 1.5×10^{10} , 2.5×10^{10} , 5×10^{10} , 1×10^{11} , 1.5×10^{11} , 2.5×10^{11} , 5×10^{11} , 1×10^{12} particles). Alternatively, a viral vector can be administered at a dose that is typically from 1×10^5 to 1×10^{10} plaque forming units (PFU), such as 1×10^5 PFU, 2.5×10^5 PFU, 5×10^5 PFU, 1×10^6 PFU, 2.5×10^6 PFU, 5×10^6 PFU, 1×10^7 PFU, 2.5×10^7 PFU, 5×10^7 PFU, 1×10^8 PFU, 2.5×10^8 PFU, 5×10^8 PFU, 1×10^9 PFU, 2.5×10^9 PFU, 5×10^9 PFU, or 1×10^{10} PFU. Dosages will vary depending upon the size of the animal and the route of administration. For example, a suitable human or veterinary dosage (for about an 80 kg animal) for intramuscular injection is in the range of about 1×10^9 to about 5×10^{12} particles per mL, for a single site. Optionally, multiple sites of administration may be used. In another example, a suitable human or veterinary dosage may be in the range of about 1×10^{11} to about 1×10^{15} particles for an oral formulation.

The viral vector can be quantified by Quantitative PCR Analysis (Q-PCR), for example with primers and probe designed on CMV promoter region using as standard curve serial dilution of plasmid DNA containing the vector genome with expression cassette including HCMV promoter. The copy number in the test sample is determined by the parallel line analysis method. Alternative methods for vector particle quantification can be analytical HPLC or spectrophotometric method based on A_{260} nm.

An immunologically effective amount of a nucleic acid may suitably be between 1 ng and 100 mg. For example, a suitable amount can be from 1 μ g to 100 mg. An appropriate amount of the particular nucleic acid (e.g., vector) can readily be determined by those of skill in the art. Exemplary effective amounts of a nucleic acid component can be between 1 ng and 100 μ g, such as between 1 ng and 1 μ g (e.g., 100 ng-1 μ g), or between 1 μ g and 100 μ g, such as 10 ng, 50 ng, 100 ng, 150 ng, 200 ng, 250 ng, 500 ng, 750 ng, or 1 μ g. Effective amounts of a nucleic acid can also include from 1 μ g to 500 μ g, such as between 1 μ g and 200 μ g, such as between 10 and 100 μ g, for example 1 μ g, 2 μ g, 5 μ g, 10 μ g, 20 μ g, 50 μ g, 75 μ g, 100 μ g, 150 μ g, or 200 μ g. Alternatively, an exemplary effective amount of a nucleic acid can be between 100 μ g and 1 mg, such as from 100 μ g to 500 μ g, for example, 100 μ g, 150 μ g, 200 μ g, 250 μ g, 300 μ g, 400 μ g, 500 μ g, 600 μ g, 700 μ g, 800 μ g, 900 μ g or 1 mg.

Generally a human dose will be in a volume of between 0.1ml and 2 ml. Thus the composition described herein can be formulated in a volume of, for example 0.1, 0.15, 0.2, 0.5, 1.0, 1.5 or 2.0 ml human dose per individual or combined immunogenic components.

One of skill in the art may adjust these doses, depending on the route of administration and the therapeutic or vaccine application for which the recombinant vector is employed. The levels of expression of the transgene, or for an adjuvant, the level of circulating antibody, can be monitored to determine the frequency of dosage administration.

If one or more priming and/or boosting steps are used, this step may include a single dose that is administered hourly, daily, weekly or monthly, or yearly. As an example, mammals may receive one or two doses containing between about 10 µg to about 50 µg of plasmid in carrier. The amount or site of delivery is desirably selected based upon the identity and condition of the mammal.

The therapeutic levels of, or level of immune response against, the protein encoded by the selected transgene can be monitored to determine the need, if any, for boosters. Following an assessment of CD8+ T cell response, or optionally, antibody titers, in the serum, optional booster immunizations may be desired. Optionally, the recombinant ChAd155 vectors may be delivered in a single administration or in various combination regimens, e.g., in combination with a regimen or course of treatment involving other active ingredients or in a prime-boost regimen.

The present invention will now be further described by means of the following non-limiting examples.

EXAMPLES

20

Example 1: Isolation of ChAd155

Wild type chimpanzee adenovirus type 155 (ChAd155) was isolated from a healthy young chimpanzee housed at the New Iberia Research Center facility (New Iberia Research Center; The University of Louisiana at Lafayette) using standard procedures as described in Colloca *et al.* (2012) and WO 2010086189, which is hereby incorporated by reference for the purpose of describing adenoviral isolation and characterization techniques

Example 2: ChAd155 Vector Construction

30

The ChAd155 viral genome was then cloned in a plasmid or in a BAC vector and subsequently modified (Figure 2) to carry the following modifications in different regions of the ChAd155 viral genome:

- 35
- a) deletion of the E1 region (from bp 449 to bp 3529) of the viral genome;
 - b) deletion of the E4 region (from bp 34731 to bp 37449) of the viral genome;
 - c) insertion of the E4orf6 derived from human Ad5.

40

2.1: Deletion of E1 region: Construction of BAC/ChAd155 ΔE1 TetO hCMV RpsL-Kana#1375

The ChAd155 viral genome was cloned into a BAC vector by homologous recombination in *E. coli* strain BJ5183 electroporation competent cells (Stratagene catalog no. 2000154) co-transformed with ChAd155 viral DNA and Subgroup C BAC Shuttle (#1365). As shown in the schematic of Figure 3, the Subgroup C Shuttle is a BAC vector derived from pBeloBAC11 (GenBank U51113, NEB) and which is dedicated to the cloning of ChAd belonging to species C and therefore contains the pIX gene and DNA fragments derived from right and left ends (including right and left ITRs) of species C ChAd viruses.

The Species C BAC Shuttle also contains a RpsL-Kana cassette inserted between left end and the pIX gene. In addition, an Amp-LacZ-SacB selection cassette, flanked by IScel restriction sites, is present between the pIX gene and right end of the viral genome. In particular, the BAC Shuttle comprised the following features: Left ITR: bp 27 to 139, hCMV(tetO) RpsL-Kana cassette: bp 493 to 3396, pIX gene: bp 3508 to 3972, IScel restriction sites: bp 3990 and 7481, Amp-LacZ-SacB selection cassette: bp 4000 to 7471, Right ITR: bp 7805 to 7917.

BJ5183 cells were co-transformed by electroporation with ChAd155 purified viral DNA and Subgroup C BAC Shuttle vector digested with IScel restriction enzyme and then purified from gel. Homologous recombination occurring between pIX gene and right ITR sequences (present at the ends of Species C BAC Shuttle linearized DNA) and homologous sequences present in ChAd155 viral DNA lead to the insertion of ChAd155 viral genomic DNA in the BAC shuttle vector. At the same time, the viral E1 region was deleted and substituted by the RpsL-Kana cassette, generating BAC/ChAd155 Δ E1/ TetO hCMV RpsL-Kana#1375.

2.2: Plasmid construction by homologous recombination in *E.coli* BJ5183

2.2.1: Deletion of E4 region - Construction of pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434)

To improve propagation of the vector, a deletion of the E4 region spanning from nucleotide 34731-37449 (ChAd155 wild type sequence) was introduced in the vector backbone by replacing the native E4 region with Ad5 E4orf6 coding sequence using a strategy involving several steps of cloning and homologous recombination in *E.coli*. The E4 coding region was completely deleted while the E4 native promoter and polyadenylation signal were conserved. To this end, a shuttle vector was constructed to allow the insertion of Ad5orf6 by replacing the ChAd155 native E4 region by homologous recombination in *E.coli* BJ5183 as detailed below.

Construction of pARS SpeciesC Ad5E4orf6-1

A DNA fragment containing Ad5orf6 was obtained by PCR using Ad5 DNA as template, with the oligonucleotides 5'-ATACGGACTAGTGGAGAAGTACTCGCCTACATG-3' (SEQ ID NO: 13) and 5'-ATACGGAAGATCTAAGACTTCAGGAAATATGACTAC-3' (SEQ ID NO: 14). The PCR fragment was digested with BglII and SpeI and cloned into Species C RLD-EGFP shuttle digested with BglII and SpeI, generating the plasmid pARS Species C Ad5orf6-1. Details regarding the shuttle can be found in Colloca *et al*, Sci. Transl. Med. (2012) 4:115ra.

Construction of pARS SpeciesC Ad5E4orf6-2

To delete the E4 region, a 177 bp DNA fragment spanning bp 34586 to bp 34730 of the ChAd155 wt sequence (SEQ ID NO: 10) was amplified by PCR using the plasmid BAC/ChAd155 ΔE1_TetO hCMV RpsL-Kana (#1375) as a template with the following oligonucleotides:

5'-ATTCAAGTGTACAGGCGCGCCAAAGCATGACGCTGTTGATTTGATTC-3' (SEQ ID NO: 15) and 5'-ACTAGGACTAGTTATAAGCTAGAATGGGGCTTTGC-3' (SEQ ID NO: 16). The PCR fragment was digested with BsrGI and SpeI and cloned into pARS SubGroupC Ad5orf6-1 digested with BsrGI and SpeI, generating the plasmid pARS SpeciesC Ad5orf6-2 (#1490). A schematic diagram of this shuttle plasmid is provided in Figure 4. In particular, the shuttle plasmid comprised the following features: Left ITR: bp 1 to 113, Species C first 460bp: bp 1 to 460, ChAd155 wt (bp 34587 to bp 34724 of SEQ ID NO:10) : bp 516 to 650, Ad5 orf6: bp 680 and 1561, Species C last 393 bp: bp 1567 to 1969, Right ITR: bp 1857 to 1969.

Construction of pChAd155 ΔE1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434)

The resulting plasmid pARS SubGroupC Ad5orf6-2 was then used to replace the E4 region within the ChAd155 backbone with Ad5orf6. To this end the plasmid BAC/ChAd155 ΔE1_TetO hCMV RpsL-Kana (#1375) was digested with PacI/PmeI and co-transformed into BJ5183 cells with the digested plasmid pARS SubGroupC Ad5orf6-2 BsrGI/Ascl, to obtain the pChAd155 ΔE1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434) pre-adenovirus plasmid.

2.2.2: Insertion of RSV Expression Cassette - Construction of pChAd155 ΔE1, E4_Ad5E4orf6/TetO hCMV RSV

An RSV cassette was cloned into a linearised pre-adenovirus acceptor vector via homologous recombination in *E.coli* by exploiting the homology existing between HCMV promoter and BGH polyA sequences. The plasmid pTetOhCMV-bghpolyA_RSV was cleaved with SfiI and SpeI to excise the 4,65 Kb fragment containing the HCMV promoter with tetO, RSV and BGHpolyA sequence. The resulting RSV 4,65 Kb fragment was cloned by

homologous recombination into the pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1434) acceptor vector carrying the RpsL-Kana selection cassette under control of HCMV and BGHpA. The acceptor pre-adenoviral plasmid was linearized with the restriction endonuclease SnaBI. The resulting construct was the pChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RSV vector (Figure 5).

2.3: BAC vector construction by recombineering

2.3.1: Deletion of E4 region - Construction of BAC/ChAd155 Δ E1, E4_Ad5E4orf6/TetO hCMV RpsL-Kana#1390

A deletion of the E4 region spanning from nucleotide 34731-37449 of the ChAd155 wt sequence was introduced in the vector backbone by replacing this native E4 region with the Ad5 E4orf6 coding sequence using a strategy involving two different steps of recombineering in *E. Coli* SW102 competent cells.

The first step resulted in insertion of a selection cassette including the suicide gene SacB, ampicillin-R gene and lacZ (Amp-LacZ-SacB selection cassette) in the E4 region of ChAd155, for the purpose of positive/negative selection of recombinants.

First step – Substitution of ChAd155 native E4 region with Amp-LacZ-SacB selection cassette

The Amp-LacZ-SacB selection cassette was amplified by PCR using the oligonucleotides provided below containing E4 flanking sequences to allow homologous recombination:

1021-FW E4 Del Step1 (5'-
TTAATAGACACAGTAGCTTAATAGACCCAGTAGTGCAAAGCCCCATTCTAGCTTATAA
CCCCTATTTGTTTATTTTCT-3') (SEQ ID NO: 17) and 1022-RW E4 Del Step1
(5'-
ATATATACTCTCTCGGCACTTGGCCTTTTACACTGCGAAGTGTTGGTGCTGGTGCTGC
GTTGAGAGATCTTTATTTGTAACTGTTAATTGTC-3') (SEQ ID NO: 18).

The PCR product was used to transform *E. Coli* SW102 competent cells containing the pAdeno plasmid BAC/ChAd155 (Δ E1) tetO hCMV - RpsLKana#1375. The transformation of SW102 cells allowed the insertion of the selection cassette in the E4 region of ChAd155 via lambda (λ) Red-mediated homologous recombination, thus obtaining BAC/ChAd155 (Δ E1) TetOhCMV - RpsL Kana #1379 (including Amp-LacZ-SacB cassette by substituting ChAd155 native E4 region).

Second step – Substitution of Amp-lacZ-SacB selection cassette with Ad5E4orf6 region

The resulting plasmid BAC/ChAd155 (DE1) TetOhCMV - RpsL Kana #1379 (with Amp-LacZ-SacB cassette in place of ChAd155 E4 region) was then manipulated to replace the Amp-lacZ-SacB selection cassette with Ad5orf6 within the ChAd155 backbone. To this end, a DNA fragment containing the Ad5orf6 region was obtained by PCR, using the

5 oligonucleotides 1025-FW E4 Del Step2 (5'-TTAATAGACACAGTAGCTTAATA-3') (SEQ ID NO: 19) and 1026-RW E4 Del Step2 (5'-GGAAGGGAGTGTCTAGTGTT-3') (SEQ ID NO: 20). The resulting DNA fragment was introduced into *E. coli* SW102 competent cells containing the pAdeno plasmid BAC/ChAd155 (DE1) TetOhCMV - RpsL Kana)#1379, resulting in a final plasmid BAC/ChAd155 (Δ E1, E4 Ad5E4orf6) TetOhCMV - RpsL

10 Kana#1390 containing Ad5orf6 substituting the native ChAd155 E4 region.

2.3.2: Insertion of RSV Expression Cassette: Construction of BAC/ChAd155 Δ E1, E4_Ad5E4orf6 /TetOhCMV RSV#1393

15 An RSV transgene was cloned into the BAC/ChAd155 Δ E1, E4_Ad5E4orf6 /TetOhCMV RSV#1393 vector by substituting the RpsL-Kana selection cassette. The construction strategy was based on two different steps of recombineering in *E. Coli* SW102 competent cells.

20 *First step – Substitution of RpsL-Kana cassette with Amp-LacZ-SacB selection cassette*

The Amp-LacZ-SacB selection cassette was obtained from plasmid BAC/ChAd155 (DE1) TetO hCMV Amp-LacZ-SacB#1342 by PCR using the oligonucleotides 91-SubMonte FW (5'-CAATGGGCGTGGATAGCGGTTTGAC-3') (SEQ ID NO: 21) and 890-BghPolyA RW

25 (5'-CAGCATGCCTGCTATTGTC-3') (SEQ ID NO: 22). The product was transformed into *E. Coli* SW102 competent cells containing the pAdeno plasmid BAC/ChAd155 (DE1, E4 Ad5E4orf6) TetOhCMV - RpsL Kana#1390, resulting in BAC/ChAd155 (DE1, E4 Ad5E4orf6) TetOhCMV – Amp-LacZ-SacB#1386.

30 *Second step – Substitution of Amp-lacZ-SacB selection cassette with RSV transgene*

The RSV transgene was inserted in plasmid BAC/ChAd155 (DE1, E4 Ad5E4orf6) TetOhCMV – Amp-LacZ-SacB#1386 by replacing the Amp-lacZ-SacB selection cassette by homologous recombination. To this end, the plasmid pvjTetOhCMV-

35 bghpolyA_RSV#1080 (containing an RSV expression cassette) was cleaved with SpeI and SfiI to excise the 4.4 Kb fragment including the HCMV promoter, RSV and BGHpolyA. The resulting RSV 4.4 Kb fragment was transformed into *E. Coli* SW102 competent cells containing the pAdeno plasmid BAC/ChAd155 (DE1, E4 Ad5E4orf6) TetOhCMV – Amp-LacZ-SacB#1386, resulting in the final plasmid BAC/ChAd155 Δ E1, E4_Ad5E4orf6 / TetO hCMV Kana#1390. The structure of the BAC carrying ChAd155/RSV (SEQ ID NO: 11) is

40 illustrated in Figure 6. In particular, ChAd155/RSV comprised the following features:

Species C Left ITR: bp 1 to 113, hCMV(tetO) bp 467 to 1311, RSV gene: bp 1348 to 4785, bghpolyA: bp 4815 to 5032, Ad5E4orf6: bp 36270 to 37151, Species C Right ITR: bp 37447 to 37559.

5 Example 3: Vector production

The productivity of ChAd155 was evaluated in comparison to ChAd3 and PanAd3 in the Procell 92 cell line.

10 3.1: Production of vectors comprising an HIV Gag transgene

Vectors expressing the HIV Gag protein were prepared as described above (ChAd155/GAG) or previously (ChAd3/GAG Colloca *et al*, Sci. Transl. Med. (2012) 4:115ra). ChAd3/GAG and ChAd155/GAG were rescued and amplified in Procell 92 until
 15 passages 3 (P3); P3 lysates were used to infect 2 T75 flasks of Procell 92 cultivated in monolayer with each vector. A multiplicity of infection (MOI) of 100 vp/cell was used for both infection experiments. The infected cells were harvested when full CPE was evident (72 hours post-infection) and pooled; the viruses were released from the infected cells by 3 cycles of freeze/thaw (-70°/37°C) then the lysate was clarified by centrifugation. The
 20 clarified lysates were quantified by Quantitative PCR Analysis with primers and probe complementary to the CMV promoter region. The oligonucleotide sequences are the following: CMVfor 5'-CATCTACGTATTAGTCATCGCTATTACCA-3' (SEQ ID NO: 23), CMVrev 5'-GACTTGGAAATCCCCGTGAGT-3' (SEQ ID NO: 24), CMVFAM-TAMRA probe 5'-ACATCAATGGGCGTGGATAGCGGTT-3' (SEQ ID NO: 25) (QPCRs were run on ABI
 25 Prism 7900 Sequence detector – Applied Biosystem). The resulting volumetric titers (vp/ml) measured on clarified lysates and the specific productivity expressed in virus particles per cell (vp/cell) are provided in Table 1 below and illustrated in Figure 7.

Table 1. Vector productivity from P3 lysates.

30

Vector	vp/ml	Total vp (20 ml conc.)	vp/cell
ChAd3/GAG	9,82E+09	1,96E+11	6,61E+03
ChAd155/GAG	1,11E+10	2,22E+11	7,46E+03

To confirm the higher productivity of the ChAd155 vector expressing HIV Gag transgene, a second experiment was performed by using purified viruses as inoculum. To this end, Procell 92 cells were seeded in a T25 Flask and infected with ChAd3/GAG and
 35 ChAd155/GAG when the confluence of the cells was about 80%, using a MOI=100 vp/cell

of infection. The infected cells were harvested when full CPE was evident; the viruses were released from the infected cells by freeze/thaw and clarified by centrifugation. The clarified lysates were quantified by Quantitative PCR Analysis by using following primers and probe: CMVfor 5'-CATCTACGTATTAGTCATCGCTATTACCA-3' (SEQ ID NO: 23),
 5 CMV rev GACTTGGAAATCCCCGTGAGT (SEQ ID NO: 24), CMV FAM-TAMRA probe 5'-ACATCAATGGGCGTGGATAGCGGTT-3' (SEQ ID NO: 25) complementary to the CMV promoter region (samples were analysed on an ABI Prism 7900 Sequence detector-Applied Biosystems). The resulting volumetric titers (vp/ml) measured on clarified lysates and the specific productivity expressed in virus particles per cell (vp/cell) are provided in
 10 Table 2 below and illustrated in Figure 8.

Table 2. Vector productivity from purified viruses.

Vector	vp/ml	Total vp/T25 flask (5ml of lysate)	vp/cell
ChAd3/GAG	1,00E+10	5,00E+10	1,67E+04
ChAd155/GAG	1,21E+10	6,05E+10	2,02E+04

15 3.2: Production of vectors comprising an RSV transgene

A different set of experiments were performed to evaluate the productivity of RSV vaccine vectors in Procell 92.S cultivated in suspension. The experiment compared PanAd3/RSV (described in WO2012/089833) and Chad155/RSV in parallel by infecting Procell 92.S at a
 20 cell density of 5×10^5 cells/ml. The infected cells were harvested 3 days post infection; the virus was released from the infected cells by 3 cycles of freeze/thaw and the lysate was clarified by centrifugation. The clarified lysates were then quantified by Quantitative PCR Analysis as reported above. The volumetric productivity and the cell specific productivity are provided in Table 3 below and illustrated in Figure 9.

25

Table 3

Virus	Volumetric productivity (Vp/ml)	Total vp	Cell specific productivity (vp/cell)
PanAd3/RSV	5,82E+09	2,91E+11	1,16E+4
ChAd155/RSV	3,16E+10	1,58E+12	6,31E+04

Example 4: Transgene Expression Levels

4.1: Expression level of HIV Gag transgene

Expression levels were compared in parallel experiments by infecting HeLa cells with ChAd3 and ChAd155 vectors comprising an HIV Gag transgene. HeLa cells were seeded in 24 well plates and infected in duplicate with ChAd3/GAG and ChAd155/GAG purified viruses using a MOI=250 vp/cell. The supernatants of HeLa infected cells were harvested 48 hours post-infection, and the production of secreted HIV GAG protein was quantified by using a commercial ELISA Kit (HIV-1 p24 ELISA Kit, PerkinElmer Life Science). The quantification was performed according to the manufacturer's instruction by using an HIV-1 p24 antigen standard curve. The results, expressed in pg/ml of GAG protein, are illustrated in Figure 10.

4.1: Expression level of RSV F transgene

Expression levels were compared in parallel experiments by infecting HeLa cells with the above-described PanAd3 and ChAd155 vectors comprising an RSV F transgene. To this end, HeLa cells were seeded in 6 well plates and infected in duplicate with PanAd3/RSV and ChAd155/RSV purified viruses using a MOI=500 vp/cell. The supernatants were harvested 48 hours post-infection, and the production of secreted RSV F protein was quantified by ELISA. Five different dilutions of the supernatants were transferred to microplate wells which are coated with a commercial mouse anti-RSV F monoclonal antibody. The captured antigen was revealed using a secondary anti-RSV F rabbit antiserum followed by Biotin-conjugated anti-rabbit IgG, then by adding Streptavidin-AP conjugate (BD Pharmingen cat. 554065). The quantification was performed by using an RSV F protein (Sino Biological cat. 11049-V08B) standard curve. The results obtained, expressed as ug/ml of RSV F protein, are provided in Table 4 below.

Table 4

Sample	µg/ml RSV F protein
ChAd155/RSV	5,9
PanAd3/RSV	4

A western blot analysis was also performed to confirm the higher level of transgene expression provided by the ChAd155 RSV vector relative to the PanAd3 RSV vector. HeLa cells plated in 6 well plates were infected with PanAd3/RSV and ChAd155/RSV purified viruses using MOI=250 and 500 vp/cell. The supernatants of HeLa infected cells were harvested and the production of secreted RSV F protein were analysed by non-reducing SDS gel followed by Western Blot analysis. Equivalent quantities of supernatants were loaded on non-reducing SDS gel; after electrophoresis separation, the proteins were

transferred to a nitrocellulose membrane to be probed with an anti-RSV F mouse monoclonal antibody (clone RSV-F-3 catalog no: ABIN308230 available at antibodies-online.com (last accessed 13 April 2015). After the incubation with primary antibody, the membrane was washed and then incubated with anti-mouse HRP conjugate secondary antibody. Finally the assay was developed by electrochemiluminescence using standard techniques (ECL detection reagents Pierce catalog no W3252282). The Western Blot results are shown in Figure 11. A band of about 170 kD indicated by the arrow was revealed by monoclonal antibody mAb 13 raised against the F protein, which corresponds to the expected weight of trimeric F protein. It can be seen that the ChAd155 RSV vector produced a darker band at both MOI=250 and 500vp/cell.

Example 5: Evaluation of immunological potency by mouse immunization experiments

5.1: Immunogenicity of vectors comprising the HIV Gag transgene

The immunogenicity of the ChAd155/GAG vector was evaluated in parallel with the ChAd3/GAG vector in BALB/c mice (5 per group). The experiment was performed by injecting 10^6 viral particles intramuscularly. T-cell response was measured 3 weeks after the immunization by ex vivo IFN-gamma enzyme-linked immunospot (ELISpot) using a GAG CD8+ T cell epitope mapped in BALB/c mice. The results are shown in Figure 12, expressed as IFN-gamma Spot Forming Cells (SFC) per million of splenocytes. Each dot represents the response in a single mouse, and the line corresponds to the mean for each dose group. Injected dose in number of virus particles and frequency of positive mice to the CD8 immunodominant peptide are shown on the x axis.

5.2 Immunogenicity of vectors comprising the RSV transgene

The immunological potency of the PanAd3/RSV and ChAd155/RSV vectors was evaluated in BALB/c mice. Both vectors were injected intramuscularly at doses of 10^8 , 10^7 and 3×10^6 vp. Three weeks after vaccination the splenocytes of immunized mice were isolated and analyzed by IFN-gamma-ELISpot using as antigens immunodominant peptide F and M epitopes mapped in BALB/c mice. The levels of immune-responses were reduced in line with decreasing dosage (as expected) but immune responses were clearly higher in the groups of mice immunized with ChAd155/RSV vector compared to the equivalent groups of mice immunized with PanAd3/RSV vaccine (Figure 13). In Figure 13, symbols show individual mouse data, expressed as IFN-gamma Spot Forming Cells (SFC)/million splenocytes, calculated as the sum of responses to the three immunodominant epitopes (F_{51-66} , F_{85-93} and $M2-1_{282-290}$) and corrected for background. Horizontal lines represent the mean number of IFN-gamma SFC/million splenocytes for each dose group.

Conclusion

Taken together the results reported above demonstrated that ChAd155 is an improved adenoviral vector in comparison to ChAd3 and PanAd3 vectors. ChAd155 was shown to be more productive therefore facilitating the manufacture process, able to express higher level of transgene in vitro and also in vivo providing a stronger T-cell response against the antigens expressed in animal models.

DESCRIPTION OF THE SEQUENCES**10 SEQ ID NO: 1 - Polypeptide sequence of ChAd155 fiber**

MKRTKTSDESNPVYPYDTESGPPSPVFLTPPFVSPDGFQESPPGVLSLNLAELVTSHGMLALKMGS
GLSLDDAGNLTSQDITTASPLKTKTNLSLETSSPLTVSTSGALTVAAPLAVAGTSLTMQSEAPL
TVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSLTVSATPPLSTSNGLGIDMQAPIYTTNGKL
GLNFGAPLHVVDLSNALTVVVGQGLTINGTALQTRVSGALNYDTSGNLELRAAGGMRVDANGQLIL
DVAYPFDAQNNLSRLGQGPLFVNSAHNLDVNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAI
NAGDGLQFDSGSDTNPLKTKLGLGLDYDSSRAIIAKLGTGLSFDNTGAITVGNKNDDKLTWTPDP
SPNCRIYSEKDAKFTLVLTCKGSQVLASVSVLSVKGSLAPISGTVTSAQIVLRFDENGVLSSNSLDPQ
YWNRYRKGDLTEGTAYTNAVGFMPNLTAYPKTSQTAKSNIQSQVYLNGDKSKPMTLTITLNGTNET
GDATVSTYSMSFSWNWNGSNYINETFQTSFTFSYIAQE

20

SEQ ID NO: 2 - Polynucleotide sequence encoding ChAd155 fiber

ATGAAGCGCACCAAAACGTCTGACGAGAGCTTCAACCCCGTGTACCCCTATGACACGGAAAGCG
GCCCTCCCTCCGTCCCTTTCCTCACCCCTCCCTTCGTGTCTCCCGATGGATTCCAAGAAAGTCCCC
CCGGGGTCTGTCTCTGAACCTGGCCGAGCCCCCTGGTCACTTCCCACGGCATGCTCGCCCTGAAA
ATGGGAAGTGGCCTCTCCCTGGACGACGCTGGCAACCTCACCTCTCAAGATATCACCACCGCTA
GCCCTCCCCTCAAAAAAACCAAGACCAACCTCAGCCTAGAAACCTCATCCCCCTAACTGTGAG
CACCTCAGGCGCCCTCACCGTAGCAGCCGCCGCTCCCCCTGGCGGTGGCCGGCACCTCCCTCACC
ATGCAATCAGAGGCCCCCTGACAGTACAGGATGCAAAACTCACCTGGCCACCAAGGCCCCC
TGACCGTGTCTGAAGGCAAACTGGCCTTGCAAACATCGGCCCCGCTGACGGCCGCTGACAGCAG
CACCTCAGAGTCAGTGCCACACCACCCCTTAGCACAAGCAATGGCAGCTTGGGTATTGACATG
CAAGCCCCCATTTACACCACCAATGGAAAAGTAACTTTGGCGCTCCCTGCATGTGGT
AGACAGCCTAAATGCACTGACTGTAGTTACTGGCCAAGGTCTTACGATAAACGGAACAGCCCTA
CAAACTAGAGTCTCAGGTGCCCTCAACTATGACACATCAGGAAACCTAGAATTGAGAGCTGCAG
GGGGTATGCGAGTTGATGCAAATGGTCAACTTATCCTTGATGTAGCTTACCCATTTGATGCACAA
AACAATCTCAGCCTTAGGCTTGACAGGGACCCCTGTTTGTTAACTCTGCCCACAACCTGGATGT
TAACTACAACAGAGGCCTCTACCTGTTACATCTGGAAATACCAAAAAGCTAGAAGTTAATATC
AAAACAGCCAAGGGTCTCATTTATGATGACACTGCTATAGCAATCAATGCGGGTGATGGGCTAC
AGTTTGACTCAGGCTCAGATACAAATCCATTAAAACTAACTTGGATTAGGACTGGATTATGA
CTCCAGCAGAGCCATAATTGCTAACTGGGAAGTGGCCTAAGCTTTGACAACACAGGTGCCATC
ACAGTAGGCAACAAAAATGATGACAAGCTTACCTTGTGGACCACACCAGACCCATCCCCTAACT
GTAGAATCTATTCAGAGAAAGATGCTAAATTCACACTTGTTTTGACTAAATGCGGCAGTCAGGT

35

40

GTTGGCCAGCGTTTCTGTTTATCTGTAAAAGGTAGCCTTGCGCCCATCAGTGGCACAGTAACTA
 GTGCTCAGATTGTCCTCAGATTTGATGAAAATGGAGTTCTACTAAGCAATTCTCCCTTGACCCT
 CAATACTGGAACACAGAAAAGGTGACCTTACAGAGGGCACTGCATATACCAACGCAGTGGGA
 TTTATGCCCAACCTCACAGCATACCCAAAAACACAGAGCCAACTGCTAAAAGCAACATTGTAA
 5 GTCAGGTTTACTTGAATGGGGACAAATCCAAACCCATGACCCTCACCATTACCCTCAATGGAAC
 TAATGAAACAGGAGATGCCACAGTAAGCACTTACTCCATGTCATTCTCATGGAACGGAATGGA
 AGTAATTACATTAATGAAACGTTCCAAACCAACTCCTTCACCTTCTCCTACATCGCCCAAGAA

SEQ ID NO: 3 - Polypeptide sequence of ChAd155 penton

10 MRRAAMYQEGPPPSYESVVGAAAAAPSSPFASQLLEPPYVPPRYLRPTGGRNSIRYSELAPLFDITTRV
 YLVDNKSADVASLNYQNDHSNFLTTVIQNNDYSPSEASTQTINLDDRSHWGGDLKILHTNMPNVN
 EFMFTNKFKARVMVSRSHTKEDRVELKYEWEFELPEGNYSETMTIDL MNNAIVEHYLKVGRQNG
 VLES DIGVKFDTRNFRGLDPVTGLVMPGVYTNEAFHPDIILLPGCGVDFTYSRLSNLLGIRKRQPFQE
 GFRITYEDLEGGNIPALLDVEAYQDSLKENEAGQEDTAPAASAAAEQGEDAADTAAADGAEDPA
 15 MVVEAPEQEEDMNDSAVRGDTFVTRGEEKQAEAEAAAEKQLAAAAAAAAALAAEAESEGTKPA
 KEPIKPLTEDSKKRSYNLLKDSTNTAYRSWYLA NYGDPSTGVRSWTLLCTPDVTCGSEQVYWSL
 PDMMQDPVTRSTRQVS NFVVGAE LLPVHSKSFYNDQAVYSQLIRQFTSLTHVFNRFENQILARPP
 APTITTVSENVPALTDHGTLP LRSIGGVQRVTVTDARRRTPYVYKALGIVSPRVLSRTE

SEQ ID NO: 4 - Polynucleotide sequence encoding ChAd155 penton

20 ATGCGGCGCGCGCGGATGTACCAGGAGGGACCTCCTCCCTCTTACGAGAGCGTGGTGGGCGCGG
 CGGCGGCGGCGCCCTCTTCTCCCTTTGCGTCGCAGCTGCTGGAGCCGCCGTACGTGCCTCCGCGC
 TACCTGCGGCCTACGGGGGGGAGAAACAGCATCCGTTACTCGGAGCTGGCGCCCCTGTTGACA
 CCACCCGGGTGTACCTGGTGGACAACAAGTCGGCGGACGTGGCCTCCCTGAACTACCAGAACGA
 25 CCACAGCAATTTTTTGACCACGGTCATCCAGAACAATGACTACAGCCCGAGCGAGGCCAGCACC
 CAGACCATCAATCTGGATGACCGGTCGCACTGGGGCGGCGACCTGAAAACCATCCTGCACACCA
 ACATGCCCAACGTGAACGAGTTCATGTTACCAATAAGTTCAAGGCGCGGGTGATGGTGTGCGG
 CTCGCACACCAAGGAAGACCGGGTGGAGCTGAAGTACGAGTGGGTGGAGTTCGAGCTGCCAGA
 GGGCAACTACTCCGAGACCATGACCATTGACCTGATGAACAACGCGATCGTGGAGCACTATCTG
 30 AAAGTGGGCAGGCAGAACGGGGTCCTGGAGAGCGACATCGGGGTCAAGTTCGACACCAGGAAC
 TTCCGCCTGGGGCTGGACCCCGTGACCGGGCTGGTTATGCCCCGGGTGTACACCAACGAGGCCT
 TCCATCCCGACATCATCCTGCTGCCCCGGCTGCGGGGTGGACTTCACTTACAGCCGCTGAGCAAC
 CTCCTGGGCATCCGCAAGCGGCAGCCCTTCCAGGAGGGCTTCAGGATCACCTACGAGGACCTGG
 AGGGGGGCAACATCCCCGCGCTCCTCGATGTGGAGGCCTACCAGGATAGCTTGAAGGAAAATG
 35 AGGCGGGACAGGAGGATACCGCCCCCGCCGCTCCGCGCCGCGGAGCAGGGCGAGGATGCTG
 CTGACACCGCGGCCGCGGACGGGGCAGAGGCCGACCCGCTATGGTGGTGGAGGCTCCCGAGC
 AGGAGGAGGACATGAATGACAGTGCAGTGCAGGAGACACCTTCGTACCCGGGGGGAGGAAA
 AGCAAGCGGAGGCCGAGGCCGCGGCCGAGGAAAAGCAACTGGCGGCAGCAGCGCGCGGGCG
 GCGTTGGCCGCGGCGGAGGCTGAGTCTGAGGGGACCAAGCCCGCCAAGGAGCCCGTGATTAAG
 40 CCCCTGACCGAAGATAGCAAGAAGCGCAGTTACAACCTGCTCAAGGACAGCACCAACACCGCG
 TACCGCAGCTGGTACCTGGCCTACAACACGGCGACCCGTCGACGGGGGTGCGCTCCTGGACCC

TGCTGTGCACGCCGGACGTGACCTGCGGCTCGGAGCAGGTGTACTGGTCGCTGCCCCGACATGAT
 GCAAGACCCCGTGACCTTCCGCTCCACGCGGCAGGTCAGCAACTTCCCGGTGGTGGGCGCCGAG
 CTGCTGCCCCGTGCACTCCAAGAGCTTCTACAACGACCAGGCCGTCTACTCCCAGCTCATCCGCCA
 GTTCACCTCTCTGACCCACGTGTTCAATCGCTTTCCTGAGAACCAGATTCTGGCGCGCCCCGCCG
 5 CCCCCACCATCACCACCGTCAGTGAAAACGTTCTGCTCTCACAGATCACGGGACGCTACCGCT
 GCGCAACAGCATCGGAGGAGTCCAGCGAGTGACCGTTACTGACGCCAGACGCCGCACCTGCCCC
 TACGTTTACAAGGCCTTGGGCATAGTCTCGCCGCGCGTCCTTTCCAGCCGCACTTTT

SEQ ID NO: 5 - Polypeptide sequence of ChAd155 hexon

10 MATPSMMPQWSYMHISGQDASEYLSPLVQFARATDSYFSLSNKFRNPTVAPTHDVTDRSQRLTL
 RFIPVDREDTAYSYKARFTLAVGDNRVLDMASTYFDIRGVLDRGPTFKPYSGTAYNSLAPKGAPNSC
 EWEQEETQTAEAAQDEEEDAEAEEMPQEEQAPVKKTHVYAQAPLSGEKITKDGLQIGTDATE
 QKPIYADPTFQPEPQIGESQWNEADASVAGGRVLKKTTPMKPCYGSYARPTNANGGQGVLEKDGG
 KMESQVDMQFFSTSENARNEANNIQPKLVLYSEVHMETPDTHISYKPAKSDDNSKVMLGQQSMPN
 15 RPNYIGFRDNFIGLMYYNSTGNMGVLAGQASQLNAVVDLQDRNTELSYQLLLDSMGDRTRYFSMW
 NQAVDSYDPDVRIENHGTDELPNYCFPLGGIGVTDITYQAIKTNGNGNGGGNTTWTKDETFADRN
 EIGVGNNFAMEINLSANLWRNFLYSNVALYLPDKLKYNPSNVEISDNPNITYDYMNRVAPGLVDC
 YINLGARWSLDYMDNVNPFNHHRNAGLRYSMLLGNRGYVPFHIQVPQKFFAIKNLLLLPGSYTYE
 WNFRKDVNMVLQSSLGNDLRVDGASIKFESICLYATFFPMAHNTASTLEAMLRNDTNDQSFNDYLS
 20 AANMLYIPANATNPISIPSRNWAAFRGWAFTRLKTKETPSLGSFGDPYYTYSGSIPYLDGTFYLNH
 TFKKVSVTFDSSVSWPGNDRLLTPNEFEIKRSVDGEGYNVAQCNMTKDWFLIQMLANYNIGYQGFY
 IPESYKDRMYSFFRNFPMSRQVVDETKYKDYQQVGIIHQHNNSGFVGYLAPTMREGQAYPANFPY
 PLIGKTAVDSTQKKFLCDRTLWRIPSSNFMSMGALTDLGQNLLYANSAHALDMTFEVDPMDEPT
 LLYVLFEVFDVVRVHQPHRGVIETVYLRTPFSAGNATT

25

SEQ ID NO: 6 - Polynucleotide sequence encoding ChAd155 hexon

ATGGCGACCCCATCGATGATGCCGCAGTGGTCGTACATGCACATCTCGGGCCAGGACGCCTCGG
 AGTACCTGAGCCCCGGGCTGGTGCACTTCGCCCCGCGCCACCGAGAGCTACTTCAGCCTGAGTAA
 CAAGTTTAGGAACCCACGGTGGCGCCACGCACGATGTGACCACCGACCGGTCTCAGCGCCTG
 30 ACGCTGCGGTTTATCCCGTGGACCGCGAGGACACCGCGTACTCGTACAAGGCGCGGTTACCC
 TGGCCGTGGGCGACAACCGCGTGCTGGACATGGCCTCCACCTACTTTGACATCCGCGGGGTGCT
 GGACCGGGGTCCCACCTTTCAAGCCCTACTCTGGCACCGCCTACAACCTCCCTGGCCCCCAAGGGC
 GCTCCCAACTCCTGCGAGTGGGAGCAAGAGGAACTCAGGCAGTTGAAGAAGCAGCAGAAGAG
 GAAGAAGAAGATGCTGACGGTCAAGCTGAGGAAGAGCAAGCAGCTACCAAAAAGACTCATGTA
 35 TATGCTCAGGCTCCCTTTCTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAACGGACG
 CTACAGCTACAGAACAAAAACCTATTTATGCAGACCTACATTCCAGCCCGAACCCCAAATCGG
 GGAGTCCCAGTGGAATGAGGCAGATGCTACAGTCGCCGCGGTAGAGTGCTAAAGAAATCTACT
 CCCATGAAACCATGCTATGGTTCCTATGCAAGACCCACAAATGCTAATGGAGGTCAGGGTGTAC
 TAACGGCAATGCCAGGGACAGCTAGAATCTCAGGTTGAAATGCAATTCTTTCAACTTCTGA
 40 AAACGCCCGTAACGAGGCTAACAACATTAGCCCAAATTTGGTGCTGTATAGTGAGGATGTGCAC
 ATGGAGACCCCGGATACGCACCTTTCTTACAAGCCCGCAAAAAGCGATGACAATTCAAAAATCA

TGCTGGGTCAGCAGTCCATGCCAACAGACCTAATTACATCGGCTTCAGAGACAACTTTATCGG
 CCTCATGTATTACAATAGCACTGGCAACATGGGAGTGCTTGCAGGTCAGGCCTCTCAGTTGAAT
 GCAGTGGTGGACTTGCAAGACAGAAACACAGAACTGTCTACCAGCTCTTGCTTGATTCCATGG
 GTGACAGAACCAGATACTTTTCCATGTGGAATCAGGCAGTGGACAGTTATGACCCAGATGTTAG
 5 AATTATTGAAAATCATGGAAGTGAAGACGAGCTCCCCAACTATTGTTTCCCTCTGGGTGGCATAG
 GGGTAACTGACACTTACCAGGCTGTTAAAACCAACAATGGCAATAACGGGGGCCAGGTGACTTG
 GACAAAAGATGAACTTTTGCAGATCGCAATGAAATAGGGGTGGGAAACAATTCGCTATGGA
 GATCAACCTCAGTGCCAACCTGTGGAGAACTTCCTGTACTCCAACGTGGCGCTGTACCTACCA
 GACAAGCTTAAGTACAACCCCTCCAATGTGGACATCTCTGACAACCCCAACACCTACGATTACA
 10 TGAACAAGCGAGTGGTGGCCCCGGGGCTGGTGGACTGCTACATCAACCTGGGCGCGCGCTGGTC
 GCTGGACTACATGGACAACGTCAACCCCTTCAACCACCACCGCAATGCGGGCCTGCGCTACCGC
 TCCATGCTCCTGGGCAACGGGCGCTACGTGCCCTTCCACATCCAGGTGCCCCAGAAGTTCTTTGC
 CATCAAGAACCTCCTCCTGCGGGCTCCTACACCTACGAGTGGAACCTTCAGGAAGGATGTC
 AACATGGTCTCCTCAGAGCTCTCTGGGTAACGATCTCAGGGTGGACGGGGCCAGCATCAAGTTCG
 15 AGAGCATCTGCCTCTACGCCACCTTCTTCCCCATGGCCACAACACGGCCTCCACGCTCGAGGCC
 ATGCTCAGGAACGACACCAACGACCAGTCTTCAATGACTACCTCTCCGCCGCCAACATGCTCT
 ACCCCATACCCGCCAACGCCACCAACGTCCCCATCTCCATCCCCTCGCGCAACTGGGCGGCCTTC
 CGCGGTGGGCCTTACCCGCTCAAGACCAAGGAGACCCCTCCTGGGCTCGGGATTTCGACC
 CCTACTACACCTACTCGGGCTCCATTCCCTACCTGGACGGCACCTTCTACCTCAACCACACTTTC
 20 AAGAAGGTCTCGGTACCTTCGACTCCTCGGTGAGTGGCCGGGCAACGACCGTCTGCTACCC
 CCAACGAGTTCGAGATCAAGCGCTCGGTGACGGGGAGGGCTACAACGTGGCCCAGTGCAACA
 TGACCAAGGACTGGTTCCTGGTCCAGATGCTGGCCAACTACAACATCGGCTACCAGGGCTTCTA
 CATCCCAGAGAGCTACAAGGACAGGATGTACTCCTTCTTCAGGAACCTCCAGCCCATGAGCCGG
 CAGGTGGTGGACCAGACCAAGTACAAGGACTACCAGGAGGTGGGCATCATCCACCAGCACAAC
 25 AACTCGGGCTTCGTGGGCTACCTCGCCCCACCATGCGCGAGGGACAGGCCTACCCCGCCAACT
 TCCCCTATCCGCTCATAGGCAAGACCGCGGTGACAGCATCACCCAGAAAAAGTTCCTCTGCGA
 CCGCACCTCTGGCGCATCCCCTTCTCCAGCAACTTCATGTCCATGGGTGCGCTCTCGGACCTGG
 GCCAGAACTTGCTCTACGCCAACTCCGCCACGCCCTCGACATGACCTTCGAGGTGACCCCATG
 GACGAGCCACCTTCTCTATGTTCTGTTGGAAGTCTTTGACGTGGTCCGGGTCCACCAGCCGCA
 30 CCGCGGCGTCATCGAGACCGTGTACCTGCGTACGCCCTTCTCGGCCGGCAACGCCACCACC

SEQ ID NO: 7 - Polynucleotide sequence encoding ChAd155#1434 backbone construct

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGATGGGCGGCGGGG
 GCGGGGCGCGGGGCGGGAGGCGGGTTTGGGGGCGGGCCGGCGGGCGGGGCGGTGTGGCGGAA
 35 GTGGACTTTGTAAGTGTGGCGGATGTGACTTGCTAGTGCCGGGCGCGGTAAAAGTGACGTTTTT
 CGTGCGGACAACGCCCCCGGAAGTGACATTTTTCCCGCGGTTTTTACCGGATGTTGTAGTGAA
 TTTGGGCGTAACCAAGTAAGATTTGGCCATTTTCGCGGGAAAACTGAAACGGGGAAGTGAAATC
 TGATTAATTTTTCGTTAGTCATACCGCGTAATATTTGTCTAGGGCCGAGGGACTTTGGCCGATTA
 CGTGAGGACTCGCCAGGTGTTTTTGGAGTGAATTTCCGCGTTCCGGGTCAAAGTCTGCGTTT
 40 TATTATTATAGGATATCCCATTCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCT
 CATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACG

GGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCCGCC
TGGCTGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACCTGGCAGT
ACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCC
5 TGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
CATCGCTATTACCATGGTGTATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCGAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTGTTTGGCACCAAAAATCAA
CGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAG
10 AGATCGTCGACGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTG
ACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGC
GGATTCCCCGTGCCAAGAGTGAGATCTTCCGTTTATCTAGGTACCGGGCCCCCCTCGAGGTCTGA
CGGTATCGATAAGCTTCACGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGG
AACACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCGGATGAATGTCAGCTACTGGGCTA
15 TCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAAGCAGGTAGCTTGCAGTGGGCTTACATGGC
GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTC
TGTAAGGTTGGGAAGCCCTGCAAAGTAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGG
CGCAGGGGATCAAGATCTAACCAGGAGCTATTTAATGGCAACAGTTAACCAGCTGGTACGCAAA
CCACGTGCTCGCAAAGTTGCGAAAAGCAACGTGCCTGCGCTGGAAGCATGCCCCGAAAAACGTG
20 GCGTATGTACTCGTGTATATACTACCACTCTAAAAAACCGAACTCCGCGCTGCGTAAAGTATGC
CGTGTTCTGCTGACTAACGGTTTCGAAGTGACTTCCTACATCGGTGGTGAAGGTCACAACCTGCA
GGAGCACTCCGTGATCCTGATCCGTGGCGGTGCTGTTAAAGACCTCCCGGGTGTTCTGTTACCACA
CCGTACGTGGTGCCTTGACTGCTCCGGCGTTAAAGACCGTAAGCAGGCTCGTTCCAAGTATGG
CGTGAAGCGTCCCTAAGGCTTAATGGTAGATCTGATCAAGAGACAGGATGACGGTCGTTTCGCAT
25 GCTTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGCTAT
GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGC
GCCCCGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGC
GCGGCTATCGTGGCTGGCCACGACGGGCGTTTCCTTGCGCAGCTGTGCTCGACGTTGTCAGTAA
GCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTATCTACCTTG
30 CTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGC
TACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTAATCGGATGGAAGCC
GGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAAGTTCG
CCAGGCTCAAGGCGCGCATGCCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTT
GCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTG
35 GCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAAT
GGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGACGCGCATCGCCTTCTAT
CGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCC
CAACCTGCCATCACGAGATTTGATTCCACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCG
TTTTCCGGGACGCGGCTGGATGATCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCAC
40 CCCGGGCTCGATCCCCCTCGGGGGGAATCAGAATTCAGTCGACAGCGGCCGCGATCTGCTGTGCC
TTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCTTCCTTGACCCTGGAAGGTGCCAC

TCCCACTGTCCTTTCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTA
TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
CTGGGGATGCGGTGGGCTCTATGGCCGATCAGCGATCGCTGAGGTGGGTGAGTGGGCGTGGCCT
GGGGTGGTCATGAAAATATATAAGTTGGGGGTCTTAGGGTCTCTTTATTTGTGTTGCAGAGACCG
5 CCGGAGCCATGAGCGGGAGCAGCAGCAGCAGCAGTAGCAGCAGCGCCTTGGATGGCAGCATCG
TGAGCCCTTATTTGACGACGCGGATGCCCCACTGGGCCGGGGTGCCTCAGAATGTGATGGGCTC
CAGCATCGACGGCCGACCCGTCTTGCCCGCAAATTCGCCACGCTGACCTATGCGACCGTCGCG
GGGACGCCGTTGGACGCCACCGCCGCCGCCGCCACCGCAGCCGCCTCGGCCGTGCGCAGCC
TGGCCACGGACTTTGCATTCTTGGGACCACTGGCGACAGGGGCTACTTCTCGGGCCGCTGCTGCC
10 GCCGTTTCGCGATGACAAGCTGACCGCCCTGCTGGCGCAGTTGGATGCGCTTACTCGGGAAGTGG
GTGACCTTTCTCAGCAGGTCATGGCCCTGCGCCAGCAGGTCTCCTCCCTGCAAGCTGGCGGGAAT
GCTTCTCCCAAAAATGCCGTTTAAGATAAAATAAAACCAGACTCTGTTTGGATTAAAGAAAAGTA
GCAAGTGCATTGCTCTCTTTATTTTCATAATTTTCCGCGCGGATAGGCCCTAGACCAGCGTTCTC
GGTCGTTGAGGGTGGGTGTATCTTCTCCAGGACGTGGTAGAGGTGGCTCTGGACGTTGAGATA
15 CATGGGCATGAGCCCGTCCCGGGGGTGGAGGTAGCACCCTGCAGAGCTTCATGCTCCGGGGTG
GTGTTGTAGATGATCCAGTCGTAGCAGGAGCGCTGGGCATGGTGCCTAAAAATGTCCTCAGCA
GCAGGCCGATGGCCAGGGGGAGGCCCTTGGTGTAAGTGTTTACAAAACGGTTAAGTTGGGAAG
GGTGCATTGCGGGAGAGATGATGTGCATCTTGGACTGTATTTTATAGATTGGCGATGTTTCCGCCC
AGATCCCTTCTGGGATTCATGTTGTGCAGGACCACCAGTACAGTGTATCCGGTGCCTTGGGGA
20 ATTTGTATGCAGCTTAGAGGGAAAAGCGTGGAAGAACTTGGAGACGCCTTTGTGGCCTCCCAG
ATTTTCCATGCATTCGTCCATGATGATGGCAATGGGCCCGCGGGAGGCAGCTTGGGCAAAGATA
TTTCTGGGGTCGCTGACGTCGTAGTTGTGTTCCAGGGTGAGGTGCTCATAGGCCATTTTTACAAA
GCGCGGGCGGAGGGTGCCCGACTGGGGGATGATGGTCCCCTCTGGCCCTGGGGCGTAGTTGCC
TCGCAGATCTGCATTTCCCAGGCCTTAATCTCGGAGGGGGGAATCATATCCACCTGCGGGGCGA
25 TGAAGAAAACGGTTTCCGGAGCCGGGGAGATTAACTGGGATGAGAGCAGGTTTCTAAGCAGCT
GTGATTTTCCACAACCGGTGGGCCCATAAATAACACCTATAACCGGTTGCAGCTGGTAGTTTAG
AGAGCTGCAGCTGCCGTCGTCCCGAGGAGGGGGGCCACCTCGTTGAGCATGTCCCTGACGCGC
ATGTTCTCCCCGACCAGATCCGCCAGAAGGCGCTCGCCGCCAGGGACAGCAGCTCTTGCAAGG
AAGCAAAGTTTTTCAGCGGCTTGAGGCCGTCCGCCGTGGGCATGTTTTTCAGGGTCTGGCTCAGC
30 AGCTCCAGGCGGTCCAGAGCTCGGTGACGTGCTCTACGGCATCTCTATCCAGCATATCTCCTCG
TTTCGCGGGTTGGGGCGACTTTTCGCTGTAGGGCACCAAGCGGTGGTTCGTCCAGCGGGGCCAGAG
TCATGTCTTCCATGGGCGCAGGGTCTCTCGTCAGGGTGGTCTGGGTACCGGTGAAGGGGTGCGC
TCCGGGCTGAGCGCTTGCCAAGGTGCGCTTGAGGCTGGTTCTGCTGGTGCTGAAGCGCTGCCGG
TCTTCGCCCTGCGCGTCGGCCAGGTAGCATTTGACCATGGTGTGCATAGTCCAGCCCCCTCCGCGGC
35 GTGTCCCTTGGCGCGCAGCTTGCCCTTGAGAGGTGGCGCCGCACGAGGGGCAGAGCAGGCTCTTG
AGCGCGTAGAGCTTGGGGGCGAGGAAGACCGATTGCGGGGAGTAGGCGTCCGCGCCGCAGACC
CCGCACACGGTCTCGCACTCCACCAGCCAGGTGAGCTCGGGGCGCGCCGGGTCAAAAACAGGT
TTCCCCCATGCTTTTTGATGCGTTTCTTACCTCGGGTCTCCATGAGGTGGTGTCCCCGCTCGGTGA
CGAAGAGGCTGTCCGTGTCTCCGTAGACCGACTTGAGGGGTCTTTTCTCCAGGGGGGTCCCTCGG
40 TCTTCTCGTAGAGGAACTCGGACCACTCTGAGACGAAGGCCCGCGTCCAGGCCAGGACGAAGG
AGGCTATGTGGGAGGGGTAGCGGTGCTTGTCCACTAGGGGGTCCACCTTCTCCAAGGTGTGAAG

ACACATGTCGCCTTCCTCGGCGTCCAGGAAGGTGATTGGCTTGTAGGTGTAGGCCACGTGACCG
GGGGTTCTGACGGGGGGGTATAAAAGGGGGTGGGGGCGCGCTCGTCGTCACTCTCTTCCGCAT
CGCTGTCTGCGAGGGCCAGCTGCTGGGGTGAGTATTCCCTCTCGAAGGCGGGCATGACCTCCGC
GCTGAGGTTGTCAGTTTCCAAAAACGAGGAGGATTTGATGTTACCTGTCCCGAGGTGATACCTT
5 TGAGGGTACCCGCGTCCATCTGGTCAGAAAACACGATCTTTTATTGTCCAGCTTGGTGCGAAC
GACCCGTAGAGGGCGTTGGAGAGCAGCTTGGCGATGGAGCGCAGGGTCTGGTTCTTGTCCCTGT
CGGCGCGCTCCTTGGCCGCGATGTTGAGCTGCACGTA CTGCGCGCGACGCAGCGCCACTCGGG
GAAGACGGTGGTGCGCTCGTCGGGCACCAGGCGCACGCGCCAGCCGCGTTGTGCAGGGTGAC
CAGGTCCACGCTGGTGGCGACCTCGCCGCGCAGGCGCTCGTTGGTCCAGCAGAGACGGCCGCC
10 TTGCGCGAGCAGAAGGGGGGCGAGGGGGTTCGAGCTGGGTCTCGTCCGGGGGGTCCGCGTCCACG
GTGAAAACCCCGGGGCGCAGGCGCGCGTCGAAGTAGTCTATCTTGCAACCTTGCATGTCCAGCG
CCTGCTGCCAGTCGCGGGGCGGCGAGCGCGCGCTCGTAGGGGTTGAGCGGCGGGCCCCAGGGCAT
GGGGTGGGTGAGTGCGGAGGCGTACATGCCGCAGATGTCATAGACGTAGAGGGGCTCCCGCAG
GACCCCGATGTAGGTGGGGTAGCAGCGGCCCGCGGATGCTGGCGCGCACGTAGTCATACAGC
15 TCGTGCGAGGGGGCGAGGAGGTCGGGGCCCAGGTTGGTGCGGGCGGGGCGCTCCGCGCGGAAG
ACGATCTGCCTGAAGATGGCATGCGAGTTGGAAGAGATGGTGGGGCGCTGGAAGACGTTGAAG
CTGGCGTCTGCAGGCCGACGGCGTCGCGCACGAAGGAGGCGTAGGAGTCGCGCAGCTTGTGTA
CCAGCTCGGCGGTGACCTGCACGTCGAGCGCGCAGTAGTCGAGGGTCTCGCGGATGATGTCATA
TTAGCCTGCCCCCTTCTTTTCCACAGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTCCAGTA
20 CTCTTGATCGGGAACCGTCCGTTCCGAACGGTAAGAGCCTAGCATGTAGA ACTGGTTGACG
GCCTGGTAGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGTAGGCCTGCGCGGCCCTTGCGGAGCG
AGGTGTGGGTCAGGGCGAAGGTGTCCCTGACCATGACTTTGAGGTACTGGTGCTTGAAGTCGGA
GTCGTGCGAGCCGCCCCGCTCCCAGAGCGAGAAGTCGGTGCGCTTCTTGAGCGGGGGTTGGGC
AGAGCGAAGGTGACATCGTTGAAGAGGATTTTGGCCGCGCGGGGCATGAAGTTGCGGGTGATGC
25 GGAAGGGCCCCGGCACTTCAGAGCGGTTGTTGATGACCTGGGCGGCGAGCACGATCTCGTCGAA
GCCGTTGATGTTGTGGCCACGATGTAGAGTTCCAGGAAGCGGGGCGGCCCTTACGGTGGGC
AGCTTCTTTAGCTCTTCGTAGGTGAGCTCCTCGGGCGAGGCGAGGCCGTGCTCGGCCAGGGCCC
AGTCCGCGAGGTGCGGGTTGTCTCTGAGGAAGGACTTCCAGAGGTGCGGGGCCAGGAGGGTCTG
CAGGCGGTCTCTGAAGGTCCTGAACTGGCGGCCACGGCCATTTTTTCGGGGGTGATGCAGTAG
30 AAGGTGAGGGGGTCTTGCTGCCAGCGGTCCAGTCGAGCTGCAGGGCGAGGTGCGCGCGGGCG
GTGACCAGGCGCTCGTCGCCCCGAATTCATGACCAGCATGAAGGGCACGAGCTGCTTCCGA
AGGCCCCCATCCAAGTGTAGGTCTCTACATCGTAGGTGACAAAGAGGCGCTCCGTGCGAGGATG
CGAGCCGATCGGGAAGAACTGGATCTCCCGCCACCAGTTGGAGGAGTGGCTGTTGATGTGGTGG
AAGTAGAAGTCCCGTCGCGGGGCCGAACACTCGTGCTGGCTTTTGTAAGCGAGCGCAGTACT
35 GGCAGCGCTGCACGGGCTGTACCTCATGCACGAGATGCACCTTTCGCCCCGCGCACGAGGAAGCC
GAGGGGAAATCTGAGCCCCCGCCTGGCTCGCGGCATGGCTGGTTCTTCTACTTTGGATGCGT
GTCCGTCTCCGTCTGGCTCCTCGAGGGGTGTTACGGTGGAGCGGACCACCACGCCGCGCAGCC
GCAGGTCCAGATATCGGCGCGCGGCGGTGCGAGTTTGATGACGACATCGCGCAGCTGGGAGCTG
TCCATGGTCTGGAGCTCCCGCGGCGGCGGCAGGTGACCCGGGAGTTCTTGACGGTTCACCTCGC
40 AGAGTCGGGCCAGGGCGCGGGGCAGGTCTAGGTGGTACCTGATCTCTAGGGGCGTGTTGGTGGC
GGCGTCGATGGCTTGACAGGAGCCCGCAGCCCCGGGGGCGACGACGGTGCCCCGCGGGGTGGT

GGTGGTGGTGGCGGTGCAGCTCAGAAGCGGTGCCGCGGGCGGGCCCCCGAGGTAGGGGGGGC
TCCGGTCCCGCGGGCAGGGGCGGCAGCGGCACGTCCGGCGTGGAGCGCGGGCAGGAGTTGGTGC
TGTGCCCCGAGGTTGCTGGCGAAGGCGACGACGCGCGGTTGATCTCCTGGATCTGGCGCCTCT
GCGTGAAGACGACGGGCCCCGGTGAGCTTGAACCTGAAAGAGAGTTTCGACAGAATCAATCTCGG
5 TGTCATTGACCGCGGCCTGGCGCAGGATCTCCTGCACGTCTCCCGAGTTGTCTTGGTAGGCGATC
TCGGCCATGAACTGCTCGATCTCTTCTCCTGGAGGTCTCCGCGTCCGGCGCGTTCCACGGTGGC
CGCCAGGTCGTTGGAGATGCGCCCCATGAGCTGCGAGAAGGCGTTGAGTCCGCCCTCGTTCCAG
ACTCGGCTGTAGACCACGCCCCCTGGTCATCGCGGGCGCGCATGACCACCTGCGCGAGGTTGA
GCTCCACGTGCCGCGCAAGACGGCGTAGTTGCGCAGACGCTGGAAGAGGTAGTTGAGGGTGG
10 TGGCGGTGTGCTCGGCCACGAAGAAGTTCATGACCCAGCGGCGCAACGTGGATTCTGTTGATGTC
CCCCAAGGCCTCCAGCCGTTCCATGGCCTCGTAGAAGTCCACGGCGAAGTTGAAAACTGGGAG
TTGCGCGCCGACACGGTCAACTCCTCCTCCAGAAGACGGATGAGCTCGGCGACGGTGTGCGCA
CCTCGCGCTCGAAGGCTATGGGGATCTCTTCTCCGCTAGCATCACACCTCCTCCTCTTCTCCT
CTTCTGGCACTTCCATGATGGCTTCTCCTCTTCGGGGGGTGGCGGCGGCGGCGGTGGGGGAGG
15 GGGCGCTCTGCGCCGGCGGCGGCGCACCGGGAGGCGGTCCACGAAGCGCGCGATCATCTCCCCG
CGGCGGCGGCGCATGGTCTCGGTGACGGCGCGGCCGTTCTCCCGGGGGCGCAGTTGGAAGACGC
CGCCGACATCTGGTGTGGGGCGGTGGCCGTGAGGCAGCGAGACGGCGCTGACGATGCATCT
CAACAATTGCTGCGTAGGTACGCCGCCGAGGGACCTGAGGGAGTCCATATCCACCGGATCCGAA
AACCTTTCGAGGAAGGCGTCTAACCAGTCGCAGTCGCAAGGTAGGCTGAGCACCGTGGCGGGCG
20 GCGGGGGGTGGGGGAGTGTCTGGCGGAGGTGCTGCTGATGATGTAATTGAAGTAGGCGGACTT
GACACGGCGGATGGTCGACAGGAGACCATGTCCTTGGGTCCGGCCTGCTGGATGCGGAGGCGG
TCGGCTATGCCCCAGGCTTCGTTCTGGCATCGGCGCAGGTCTTGATGATGCTTGCATGAGCCT
TTCCACCGGCACCTCTTCTCCTTCTTCTGCTTCTTCCATGTCTGCTTCGGCCCTGGGGCGGCG
CCGCGCCCCCTGCCCCCATGCGCGTGACCCGAACCCCTGAGCGGTTGGAGCAGGGCCAGG
25 TCGGCGACGACGCGCTCGGCCAGGATGGCCTGCTGCACCTGCGTGAGGGTGGTTTGAAGTCAT
CCAAGTCCACGAAGCGGTGGTAGGCGCCCGTGTGATGGTGTAGGTGCAGTTGGCCATGACGGA
CCAGTTGACGGTCTGGTGGCCCGGTTGCGACATCTCGGTGTACCTGAGTCGCGAGTAGGCGCGG
GAGTCGAAGACGTAGTCGTTGCAAGTCCGCACCAAGTACTGGTAGCCACCAGGAAGTGC GGCG
GCGGCTGGCGGTAGAGGGGCCAGCGCAGGGTGGCGGGGGCTCCGGGGGCCAGGTCTTCCAGCA
30 TGAGGCGGTGGTAGGCGTAGATGTACCTGGACATCCAGGTGATAACCGCGGCGGTGGTGGAGGC
GCGCGGGAAGTCGCGCACCCGTTCCAGATGTTGCGCAGGGGCAGAAAGTGCTCCATGGTAGGC
GTGCTCTGTCCAGTCAGACGCGCGCAGTCGTTGATACTCTAGACCAGGGAAAACGAAAGCCGGT
CAGCGGGCACTCTTCCGTGGTCTGGTGAATAGATCGCAAGGGTATCATGGCGGAGGGCCTCGGT
TCGAGCCCCGGGTCCGGGCCGACGGTCCGCCATGATCCACGCGGTTACCGCCCGCGTGTGAA
35 CCCAGGTGTGCGACGTCAGACAACGGTGGAGTGTTCTTTTGGCGTTTTTCTGGCCGGGCGCCGG
CGCCGCGTAAGAGACTAAGCCGCGAAAGCGAAAGCAGTAAGTGGCTCGCTCCCCGTAGCCGGA
GGGATCCTTGCTAAGGGTTGCGTTGCGGCGAACCCCGGTTGCAATCCCGTACTCGGGCCGGCCG
GACCCGCGGCTAAGGTGTTGGATTGGCCTCCCCCTCGTATAAAGACCCCGTTGCGGATTGACTC
CGGACACGGGGACGAGCCCCTTTTATTTTGGCTTTCCCCAGATGCATCCGGTGTGCGGCAGATG
40 CGCCCCCGCCCCAGCAGCAGCAACAACACAGCAAGAGCGGCAGCAACAGCAGCGGGAGTCA
TGCAGGGCCCCCTACCCACCCTCGGCGGGCCGGCCACCTCGGCGTCCGCGGCCGTGTCTGGCG

CCTGCGGCGGCGGGGGGCCGGCTGACGACCCCGAGGAGCCCCCGGGCGCAGGGCCAGAC
ACTACCTGGACCTGGAGGAGGGCGAGGGCTGGCGCGGCTGGGGGCGCCGTCTCCCGAGCGCC
ACCCGCGGGTGCAGCTGAAGCGCGACTCGCGCGAGGCGTACGTGCCTCGGCAGAACCTGTTCAG
GGACCGCGGGGCGAGGAGCCCGAGGAGATGCGGGACAGGAGGTTACGCGCAGGGCGGGAGC
5 TGCGGCAGGGGCTGAACCGCGAGCGGCTGCTGCGCGAGGAGGACTTTGAGCCCGACGCGCGGA
CGGGGATCAGCCCCGCGCGCGCACGTGGCGGCCCGACCTGGTGACGGCGTACGAGCAGA
CGGTGAACCAGGAGATCAACTTCCAAAAGAGTTTCAACAACCACGTGCGCACGCTGGTGGCGCG
CGAGGAGGTGACCATCGGGCTGATGCACCTGTGGGACTTTGTAAGCGCGCTGGTGCAGAACCCC
AACAGCAAGCCTCTGACGGCGCAGCTGTTCTGATAGTGACGACAGCAGGGACAACGAGGCG
10 TTTAGGGACGCGCTGCTGAACATCACCGAGCCCGAGGGTCGGTGGCTGCTGGACCTGATTAACA
TCCTGCAGAGCATAGTGGTGCAGGAGCGCAGCCTGAGCCTGGCCGACAAGGTGGCGGCCATCA
ACTACTCGATGCTGAGCCTGGGCAAGTTTTACGCGCGCAAGATCTACCAGACGCCGTACGTGCC
CATAGACAAGGAGGTGAAGATCGACGGTTTTTACATGCGCATGGCGCTGAAGGTGCTCACCTG
AGCGACGACCTGGGCGTGTACCGCAACGAGCGCATCCACAAGGCCGTGAGCGTGAGCCGGCGG
15 CGCGAGCTGAGCGACCGCGAGCTGATGCACAGCCTGCAGCGGGCGCTGGCGGGCGCCGGCAGC
GGCGACAGGGAGGCGGAGTTCCTACTTCGATGCGGGGGCGGACCTGCGCTGGGCGCCAGCCGG
CGGGCCCTGGAGGCCGCGGGGGTCCGCGAGGACTATGACGAGGACGGCGAGGAGGATGAGGAG
TACGAGCTAGAGGAGGGCGAGTACCTGGACTAAACCGCGGGTGGTGTTCGCGTAGATGCAAG
ACCCGAACGTGGTGGACCCGCGCTGCGGGCGGCTCTGCAGAGCCAGCCGTCCGGCCTTAATC
20 CTCAGACGACTGGCGACAGGTCATGGACCGCATCATGTGCTGACGGCGCGTAACCCGGACGCG
TTCCGGCAGCAGCCGAGGCCAACAGGCTCTCCGCCATCCTGGAGGCGGTGGTGCCTGCGCGCT
CGAACCCACGCACGAGAAGGTGCTGGCCATAGTGAACGCGCTGGCCGAGAACAGGGCCATCC
GCCCCGACGAGGCCGGGCTGGTGTACGACGCGCTGCTGCAGCGCGTGGCCCGCTACAACAGCG
GCAACGTGCAGACCAACCTGGACCGGCTGGTGGGGGACGTGCGCGAGGCGGTGGCGCAGCGCG
25 AGCGCGCGGATCGGCAGGGCAACCTGGGCTCCATGGTGGCGCTGAATGCCTTCCTGAGCACGCA
GCCGGCCAACGTGCCGCGGGGGCAGGAAGACTACACCAACTTTGTGAGCGCGCTGCGGCTGATG
GTGACCGAGACCCCCAGAGCGAGGTGTACCACTCGGGCCCCGACTACTTCTTCCAGACCAGCA
GACAGGGCCTGCAGACGGTGAACCTGAGCCAGGCTTTCAAGAACCTGCGGGGGCTGTGGGGCG
TGAAGGCGCCACCGGCGACCGGGCGACGGTGTCCAGCCTGCTGACGCCCAACTCGCGCCTGCT
30 GCTGCTGCTGATCGCGCCGTTACGGACAGCGGCAGCGTGTCCCGGGACACCTACCTGGGGCAC
CTGCTGACCCTGTACCGCGAGGCCATCGGGCAGGCGCAGGTGGACGAGCACACCTTCCAGGAGA
TCACCAGCGTGAGCCGCGCGCTGGGGCAGGAGGACACGAGCAGCCTGGAGGCGACTCTGAAC
ACCTGCTGACCAACCGGCGGCAGAAGATTCCCTCGCTGCACAGCCTGACCTCCGAGGAGGAGCG
CATCTTGCGCTACGTGCAGCAGAGCGTGAGCCTGAACCTGATGCGCGACGGGGTGACGCCACG
35 GTGGCGCTGGACATGACCGCGCGCAACATGGAACCGGGCATGTACGCCGCGCACCGGCCCTTACA
TCAACCGCCTGATGGACTACCTGCATCGCGCGGCGGCCGTGAACCCCGAGTACTTTACCAACGC
CATCCTGAACCCGCACTGGCTCCCGCCGCCCGGGTTCTACAGCGGGGGCTTCGAGGTCCCGGAG
ACCAACGATGGCTTCCTGTGGGACGACATGGACGACAGCGTGTTCCTCCCGCGGCCGACGGCGC
TGGCGGAAGCGTCCCTGCTGCGTCCCAAGAAGGAGGAGGAGGAGGAGGCGAGTCGCCGCCGCG
40 GCAGCAGCGCGCTGGCTTCTCTGTCCGAGCTGGGGGCGGCAGCCGCCGCGCGCCCCGGGTCCCT
GGGCGGCAGCCCCTTTCCGAGCCTGGTGGGGTCTCTGCACAGCGAGCGCACCAACCCGCCCTCGG

CTGCTGGGCGAGGACGAGTACCTGAATAACTCCCTGCTGCAGCCGGTGCGGGAGAAAAACCTGC
CTCCCGCCTTCCCCAACACGGGATAGAGAGCCTGGTGGACAAGATGAGCAGATGGAAGACCT
ATGCGCAGGAGCACAGGGACGCGCCTGCGCTCCGGCCGCCACGCGGCGCCAGCGCCACGACC
GGCAGCGGGGGCTGGTGTGGGATGACGAGGACTCCGCGGACGATAGCAGCGTGCTGGACCTGG
5 GAGGGAGCGGCAACCCGTTTCGCGCACCTGCGCCCCCGCCTGGGGAGGATGTTTTAAAAA
AAAAAAGCAAGAAGCATGATGCAAAAATTAAATAAACTACCAAGGCCATGGCGACCGAG
CGTTGGTTTCTTGTGTTCCCTTCAGTATGCGGCGCGCGGCGATGTACCAGGAGGGACCTCCTCCC
TCTTACGAGAGCGTGTTGGGCGCGGCGGCGGCGGCCCTCTTCTCCCTTTGCGTCGCAGCTGCT
GGAGCCCGCGTACGTGCCTCCGCGCTACCTGCGGCCTACGGGGGGGAGAAACAGCATCCGTTAC
10 TCGGAGCTGGCGCCCCTGTTTCGACACCACCCGGGTGTACCTGGTGGACAACAAGTCGGCGGACG
TGGCCTCCCTGAACTACCAGAACGACCACAGCAATTTTTTGACCACGGTCATCCAGAACAATGA
CTACAGCCCGAGCGAGGCCAGCACCCAGACCATCAATCTGGATGACCGGTCGCACTGGGGCGGG
GACCTGAAAACCATCCTGCACACCAACATGCCAACGTGAACGAGTTCATGTTACCAATAAGT
TCAAGGCGCGGGTGATGGTGTGCGGCTCGCACACCAAGGAAGACCGGGTGGAGCTGAAGTACG
15 AGTGGGTGGAGTTCGAGCTGCCAGAGGGCAACTACTCCGAGACCATGACCATTGACCTGATGAA
CAACGCGATCGTGGAGCACTATCTGAAAGTGGGCAGGCAGAACGGGGTCTGGAGAGCGACAT
CGGGGTCAAGTTCGACACCAGGAATTCCGCTGGGGCTGGACCCCGTGACCGGGCTGGTTATG
CCCGGGGTGTACACCAACGAGGCCTTCCATCCCGACATCATCCTGCTGCCCGGCTGCGGGGTGG
ACTTCACTTACAGCCGCCTGAGCAACCTCCTGGGCATCCGCAAGCGGCAGCCCTTCCAGGAGGG
20 CTTACAGGATCACCTACGAGGACCTGGAGGGGGCAACATCCCCGCGCTCCTCGATGTGGAGGCC
TACCAGGATAGCTTGAAGGAAAATGAGGCGGGACAGGAGGATACCGCCCCCGCCGCTCCGCC
GCCGCCGAGCAGGGCGAGGATGCTGCTGACACCGCGGCCGCGGACGGGGCAGAGGCCGACCCC
GCTATGGTGGTGGAGGCTCCCGAGCAGGAGGAGGACATGAATGACAGTGCGGTGCGCGGAGAC
ACCTTCGTACCCGGGGGGAGGAAAAGCAAGCGGAGGCCGAGGCCGCGGCCGAGGAAAAGCA
25 ACTGGCGGCAGCAGCGGCGGCGGCGGCGTGGCCGCGGCGGAGGCTGAGTCTGAGGGGACCAA
GCCCCCAAGGAGCCCGTGATTAAGCCCCTGACCGAAGATAGCAAGAAGCGCAGTTACAACCT
GCTCAAGGACAGCACCAACACCGCGTACCGCAGCTGGTACCTGGCCTACAACTACGGCGACCCG
TCGACGGGGGTGCGCTCCTGGACCTGCTGTGCACGCCGACGTGACCTGCGGCTCGGAGCAGG
TGTA CTGGTGCCTGCCGACATGATGCAAGACCCCGTGACCTTCCGCTCCACGCGGCAGGTCAG
30 CAACTTCCCGGTGGTGGGCGCCGAGCTGCTGCCCGTGCACTCCAAGAGCTTCTACAACGACCAG
GCCGTCTACTCCAGCTCATCCGCCAGTTCACCTCTCTGACCCACGTGTTCAATCGCTTTCCTGAG
AACCAGATTCTGGCGCGCCCGCCCCCCCCACCATCACCACCGTCAGTGAACGTTTCCTGCTCT
CACAGATCACGGGACGCTACCGCTGCGCAACAGCATCGGAGGAGTCCAGCGAGTGACCGTTACT
GACGCCAGACGCCGCACCTGCCCCCTACGTTTACAAGGCCTTGGGCATAGTCTCGCCGCGCGTCT
35 TTCCAGCCGCACTTTTTGAGCAACACCACCATCATGTCCATCCTGATCTACCCAGCAATAACTC
CGGCTGGGGACTGCTGCGCGCGCCCAGCAAGATGTTTCGGAGGGGCGAGGAAGCGTTCCGAGCA
GCACCCCGTGCGCGTGC GCGGGCACTTCCGCGCCCCCTGGGGAGCGCACAAACGCGGCCGCGCG
GGGCGCACCAACCGTGACGACGCCATCGACTCGGTGGTGGAGCAGGCGCGCAACTACAGGCC
GCGGTCTCTACCGTGACGCGGCATCCAGACCGTGGTGGGGGCGCGCGGCGGTACGCCAAGC
40 TGAAGAGCCGCCGGAAGCGCGTGGCCCCGCCGCCACCGCCGCCGACCCGGGGCCGCCGCCAAC
GCGCCGCCGCGGCCCTGCTTCGCCGGGCCAAGCGCACGGGCGCCGCGCGGCCCATGAGGGCCGC

GCGCCGCTTGCCGCCCGGCATCACCGCCGCCACCATGGCCCCCGTACCCGAAGACGCGCGGCC
GCCGCCGCCGCCGCCGCCATCAGTGACATGGCCAGCAGGCGCCGGGGCAACGTGTACTGGGTGC
GCGACTCGGTGACCGGCACGCGCGTGCCCGTGCGCTTCCGCCCCCGCGGACTTGAGATGATGT
GAAAAACAACACTGAGTCTCCTGCTGTTGTGTGTATCCCAGCGGCGGCGCGCGCAGCGTC
5 ATGTCCAAGCGCAAAATCAAAGAAGAGATGCTCCAGGTGCTCGCGCCGAGATCTATGGGCCCC
CGAAGAAGGAAGAGCAGGATTCAAGCCCCGCAAGATAAAGCGGGTCAAAAAGAAAAAGAAA
GATGATGACGATGCCGATGGGGAGGTGGAGTTCCTGCGCGCCACGGCGCCAGGCGCCCGGTGC
AGTGGAAGGGCCGCGCGTAAAGCGCGTCTTGCGCCCCGGCACCGCGGTGGTCTTACGCCCCG
CGAGCGCTCCACCCGACTTTCAAGCGCGTCTATGACGAGGTGTACGGCGACGAAGACCTGCTG
10 GAGCAGGCCAACGAGCGCTTCGGAGAGTTTGCTTACGGGAAGCGTCAGCGGGCGCTGGGGAAG
GAGGACCTGCTGGCGCTGCCGCTGGACCAGGGCAACCCACCCCACTCTGAAGCCCGTGACCC
TGCAGCAGGTGCTGCCGAGCAGCGCACCTCCGAGGCGAAGCGGGGTCTGAAGCGCGAGGGCG
GCGACCTGGCGCCACCGTGCAGCTCATGGTGCCCAAGCGGCAGAGGCTGGAGGATGTGCTGGA
GAAAATGAAAGTAGACCCCGGTCTGCAGCCGGACATCAGGGTCCGCCCCATCAAGCAGGTGGC
15 GCCGGGCCCTCGGCGTGACAGCCGTGGACGTGGTCATCCCCACCGGCAACTCCCCCGCCGCCGCC
ACCACTACCGCTGCCTCCACGGACATGGAGACACAGACCGATCCCGCCGACCCGCAGCCGCAG
CCGCCGCCCGACCTCCTCGGCGGAGGTGCAGACGGACCCCTGGCTGCCGCCGGCGATGTCAGC
TCCCCGCGCGCTCGCGGGCGCAGGAAGTACGGCGCCGCCAACCGCGTCTTGCCGAGTACGCC
TTGCATCCTTCCATCGCGCCACCCCGGCTACCGAGGCTATACCTACCGCCCGCGAAGAGCCA
20 AGGGTTCCACCCGCCGTCCCCGCCGACGCGCCGCCACCACCCGCCGCCGCCGCCGAGACG
CCAGCCCGCACTGGCTCCAGTCTCCGTGAGGAAAGTGGCGCGCGACGGACACACCCTGGTGCTG
CCCAGGGCGCGCTACCACCCAGCATCGTTTAAAAGCCTGTTGTGGTTCTTGAGATATGGCCCT
CACTTGCCGCCTCCGTTTCCCGGTGCCGGGATACCGAGGAGGAAGATCGCGCCGACGAGGGGT
CTGGCCGGCCGCGGCCTGAGCGGAGGCAGCCGCCGCGCGCACCGGCGGGCGACGCGCCACCAGC
25 CGACGCATGCGCGGCGGGGTGCTGCCCCGTGTAATCCCCCTGATCGCCGCGGCGATCGGCGCCG
TGCCCGGGATCGCCTCCGTGGCCTTGCAAGCGTCCCAGAGGCATTGACAGACTTGCAAACTTGC
AAATATGAAAAAAAACCCCAATAAAAAAGTCTAGACTCTCACGCTCGCTTGGTCCTGTGACT
ATTTGTAGAATGGAAGACATCAACTTTGCGTCGCTGGCCCCGCGTCACGGCTCGCGCCCGTTCC
TGGGACACTGGAACGATATCGGCACCAGCAACATGAGCGGTGGCGCCTTCAAGTTGGGGCTCTCT
30 GTGGAGCGGCATTAAAGTATCGGGTCTGCCGTTAAAAATTACGGCTCCCGGGCCTGGAACAGC
AGCACGGGCCAGATGTTGAGAGACAAGTTGAAAGAGCAGAACTTCCAGCAGAAGGTGGTGGAG
GGCCTGGCCTCCGGCATCAACGGGGTGGTGGACCTGGCCAACCAGGCCGTGCAGAATAAGATCA
ACAGCAGACTGGACCCCCGGCCGCCGGTGGAGGAGGTGCCGCCGGCGCTGGAGACGGTGTCCC
CCGATGGGCGTGCGGAGAAGCGCCCGCGCCCGATAGGGAAGAGACCACTCTGGTCACGCAGA
35 CCGATGAGCCGCCCCGTATGAGGAGGCCCTGAAGCAAGGTCTGCCCACCACGCGGCCCATCGC
GCCCATGGCCACCGGGGTGGTGGGCCGCCACACCCCGCCACGCTGGACTTGCTCCGCCCGCC
GATGTGCCGACGAGCAGAAGGCGGCACAGCCGGGCCCGCCGCGACCGCCTCCCGTTCTCCG
CCGGTCTCTGCGCCGCGCGGCCAGCGGCCCCCGCGGGGGGGTTCGCGAGGCACGGCAACTGGC
AGAGCACGCTGAACAGCATCGTGGGTCTGGGGGTGCGGTCCGTGAAGCGCCGCCGATGCTACTG
40 AATAGCTTAGCTAACGTGTTGTATGTGTGTATGCGCCCTATGTGCGCCGACAGAGGAGCTGCTGAG
TCGCCGCCGTTTCGCGCGCCACCACCACCGCCACTCCGCCCTCAAGATGGCGACCCCATCGAT

GATGCCGCAGTGGTCGTACATGCACATCTCGGGCCAGGACGCCTCGGAGTACCTGAGCCCCGGG
CTGGTGCAAGTTCGCCCCGCCACCGAGAGCTACTTCAGCCTGAGTAACAAGTTTAGGAACCCCA
CGGTGGCGCCACGCACGATGTGACCACCGACCGGTCTCAGCGCCTGACGCTGCGGTTTATTCC
CGTGGACCGCGAGGACACCGCGTACTCGTACAAGGCGCGGTTACCCCTGGCCGTGGGCGACAAC
5 CGCGTGCTGGACATGGCCTCCACCTACTTTGACATCCGCGGGGTGCTGGACCGGGGTCCCACCTT
CAAGCCCTACTCTGGCACCGCCTACAACCTCCCTGGCCCCCAAGGGCGCTCCCAACTCCTGCGAGT
GGGAGCAAGAGGAAACTCAGGCAGTTGAAGAAGCAGCAGAAGAGGAAGAAGAAGATGCTGAC
GGTCAAGCTGAGGAAGAGCAAGCAGCTACCAAAAAGACTCATGTATATGCTCAGGCTCCCCCTT
CTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAACGGACGCTACAGCTACAGAACAAA
10 AACCTATTTATGCAGACCCTACATTCCAGCCCGAACCCCAAAATCGGGGAGTCCCAGTGGAATGA
GGCAGATGCTACAGTCGCCGGCGGTAGAGTGCTAAAGAAATCTACTCCCATGAAACCATGCTAT
GGTTCCTATGCAAGACCCACAAATGCTAATGGAGGTCAGGGTGCTACTAACGGCAAATGCCCAGG
GACAGCTAGAATCTCAGGTTGAAATGCAATTCTTTTCAACTTCTGAAAACGCCCGTAACGAGGC
TAACAACATTCAGCCCAAATTGGTGCTGTATAGTGAGGATGTGCACATGGAGACCCCGGATACG
15 CACCTTTCTTACAAGCCCGCAAAAAGCGATGACAATTCAAAAATCATGCTGGGTCAGCAGTCCA
TGCCCAACAGACCTAATTACATCGGCTTCAGAGACAACTTTATCGGCCTCATGTATTACAATAGC
ACTGGCAACATGGGAGTGCTTGCAGGTCAGGCCTCTCAGTTGAATGCAGTGGTGGACTTGCAAG
ACAGAAACACAGAACTGTCTTACCAGCTCTTGCTTGATTCCATGGGTGACAGAACCAGATACTT
TTCCATGTGGAATCAGGCAGTGGACAGTTATGACCCAGATGTTAGAATTATTGAAAATCATGGA
20 ACTGAAGACGAGCTCCCCAACTATTGTTTCCCTCTGGGTGGCATAGGGGTAAGTACACTTACCA
GGCTGTTAAAACCAACAATGGCAATAACGGGGGCCAGGTGACTTGGACAAAAGATGAAACTTTT
GCAGATCGCAATGAAATAGGGGTGGGAAACAATTCGCTATGGAGATCAACCTCAGTGCCAACC
TGTGGAGAACTTCCTGTACTCCAACGTGGCGCTGTACCTACCAGACAAGCTTAAGTACAACCC
CTCCAATGTGGACATCTCTGACAACCCCAACACCTACGATTACATGAACAAGCGAGTGGTGGCC
25 CCGGGGCTGGTGGACTGCTACATCAACCTGGGCGCGCGCTGGTCGCTGGACTACATGGACAACG
TCAACCCCTTCAACCACCACCGCAATGCGGGCCTGCGCTACCGCTCCATGCTCCTGGGCAACGG
GCGCTACGTGCCCTTCCACATCCAGGTGCCCCAGAAGTTCTTTGCCATCAAGAACCTCCTCCTCC
TGCCGGGCTCCTACACCTACGAGTGGAACCTCAGGAAGGATGTCAACATGGTCCTCCAGAGCTC
TCTGGGTAAACGATCTCAGGGTGGACGGGGCCAGCATCAAGTTCGAGAGCATCTGCCTCTACGCC
30 ACCTTCTTCCCCATGGCCCCACAACACGGCCTCCACGCTCGAGGCCATGCTCAGGAACGACACCA
ACGACCAGTCCTTCAATGACTACCTCTCCGCCGCCAACATGCTCTACCCCATACCCGCCAACGCC
ACCAACGTCCCCATCTCCATCCCCTCGCGCAACTGGGCGGCCCTTCCGCGGCTGGGCCTTACCCCG
CCTCAAGACCAAGGAGACCCCTCCCTGGGCTCGGGATTGACCCCTACTACACCTACTCGGGC
TCCATTCCCTACCTGGACGGCACCTTCTACCTCAACCACACTTCAAGAAGGTCTCGGTACCTT
35 CGACTCCTCGGTCAGCTGGCCGGGCAACGACCGTCTGCTCACCCCCAACGAGTTCGAGATCAAG
CGCTCGGTGACGGGGAGGGCTACAACGTGGCCCAGTGCAACATGACCAAGGACTGGTTCCTGG
TCCAGATGCTGGCCAACTACAACATCGGCTACCAGGGCTTCTACATCCCAGAGAGCTACAAGGA
CAGGATGTACTCCTTCTCAGGAACCTCCAGCCCATGAGCCGGCAGGTGGTGGACCAGACCAAG
TACAAGGACTACCAGGAGGTGGGCATCATCCACCAGCACAACTCGGGCTTCGTGGGCTACC
40 TCGCCCCCACCATGCGCGAGGGACAGGCCTACCCCGCCAACTTCCCCTATCCGCTCATAGGCAA
GACCGCGGTGACAGCATCACCCAGAAAAAGTTCTCTGCGACCGCACCCCTCTGGCGCATCCCC

TTCTCCAGCAACTTCATGTCCATGGGTGCGCTCTCGGACCTGGGCCAGAACTTGCTCTACGCCAA
CTCCGCCCACGCCCTCGACATGACCTTCGAGGTCGACCCCATGGACGAGCCACCCCTTCTCTATG
TTCTGTTTCAAGTCTTTGACGTGGTCCGGGTCCACCAGCCGCACCGCGGCGTCATCGAGACCGTG
TACCTGCGTACGCCCTTCTCGGCCGGCAACGCCACCACCTAAAGAAGCAAGCCGCAGTCATCGC
5 CGCCTGCATGCCGTCGGGTTCACCGAGCAAGAGCTCAGGGCCATCGTCAGAGACCTGGGATGC
GGGCCCTATTTTTTGGGCACCTTCGACAAGCGCTTCCCTGGCTTTGTCTCCCCACACAAGCTGGC
CTGCGCCATCGTCAACACGGCCGGCCGCGAGACCGGGGGCGTGCACTGGCTGGCCTTCGCCTGG
AACCCGCGCTCCAAAACATGCTTCCTCTTTGACCCCTTCGGCTTTTCGGACCAGCGGCTCAAGCA
AATCTACGAGTTCGAGTACGAGGGCTTGCTGCGTCGCAGCGCCATCGCCTCCTCGCCCAGCCGT
10 GCGTCACCCCTCGAAAAGTCCACCCAGACCGTGAGGGGCCCGACTCGGCCGCCTGCGGTCTCTT
CTGCTGCATGTTTCTGCACGCCTTTGTGCACTGGCCTCAGAGTCCCATGGACCGCAACCCCCACCA
TGAACCTTGCTGACGGGGGTGCCAACTCCATGCTCCAGAGCCCCCAGGTCGAGCCCACCCTGCG
CCGCAACCAGGAGCAGCTCTACAGCTTCCTGGAGCGCCACTCGCCTTACTTCCGCCGCCACAGC
GCACAGATCAGGAGGGCCACCTCCTTCTGCCACTTGCAAGAGATGCAAGAAGGGTAATAACGAT
15 GTACACACTTTTTTCTCAATAAATGGCATCTTTTTATTTATAACAAGCTCTCTGGGGTATTCAATT
CCCACCACCACCCGCCGTTGTGCGCATCTGGCTCTATTTAGAAATCGAAAGGGTCTGCCGGGAG
TCGCCGTGCGCCACGGGCAGGGACACGTTGCGATACTGGTAGCGGGTGCCCCACTTGAACTCGG
GCACCACCAGGCGAGGCAGCTCGGGGAAGTTTTCGCTCCACAGGCTGCGGGTCAGCACCAGCGC
GTTTCATCAGGTCGGGCGCCGAGATCTTGAAGTCGCAGTTGGGGCCGCCGCCCTGCGCGCGCGAG
20 TTGCGGTACACCGGGTTGCAGCACTGGAACACCAACAGCGCCGGGTGCTTACGCTGGCCAGCA
CGCTGCGGTGCGAGATCAGCTCGGCGTCCAGGTCTCCGCGTTGCTCAGCGCGAACGGGGTCAT
CTTGGGCACTTGCCGCCCCAGGAAGGGCGCGTGCCCCGGTTTCGAGTTGCAGTCGCAGCGCAGC
GGGATCAGCAGGTGCCCCGTGCCCGGACTCGGCGTTGGGGTACAGCGCGCGCATGAAGGCCTGCA
TCTGGCGGAAGGCCATCTGGGCCCTTGCGGCCCTCCGAGAAGAACATGCCGCAGGACTTGCCCGA
25 GAACTGGTTTTCGGGGCAGCTGGCGTCGTGCAGGCAGCAGCGCGCGTCGGTGTGGCGATCTGC
ACCACGTTGCGCCCCCACCAGTTCTTACGATCTTGGCCTTGACGATTGCTCCTTCAGCGCGCG
CTGCCCCGTTCTCGCTGGTCACATCCATCTCGATCACATGTTCCCTTGTTACCATGCTGCTGCCGTG
CAGACACTTCAGCTCGCCCTCCGTCTCGGTGCAGCGGTGCTGCCACAGCGCGCAGCCCCGTGGGC
TCGAAAGACTTGTAGGTACCTCCGCGAAGGACTGCAGGTACCCCTGCAAAAAGCGGCCCATCA
30 TGGTCACGAAGGTCTTGTTGCTGCTGAAGGTCAGCTGCAGCCCGCGGTGCTCCTCGTTCAGCCAG
GTCTTGACACGCGCCGAGCGCCTCCACCTGGTCGGGCAGCATCTTGAAGTTCACCTTCAGCTC
ATTCTCCACGTGGTACTTGTCATCAGCGTGCGCGCCGCTCCATGCCCTTCTCCCAGGCCGACA
CCAGCGGCAGGCTCACGGGGTTCTTACCATCACCGTGCCCGCCGCTCCGCCGCGCTTTCGCTT
TCCGCCCCGCTGTTCTCTTCTCTCTCTCTCTCTCTCTCGCCGCCGCCACTCGCAGCCCCCGCACC
35 ACGGGGTCGTCTTCTGCAAGGCGCTGCACCTTGCGCTTGCCGTTGCGCCCCTGCTTGATGCGCAC
GGGCGGGTTGCTGAAGCCCACCATCACAGCGCGGCCTCTTCTTGCTCGTCCTCGCTGTCCAGAA
TGACCTCCGGGAGGGGGGGTGGTTCATCTCAGTACCGAGGCACGCTTCTTTTTCTTCTGCGG
GCGTTGCGCAGCTCCGCGGCTGCGGCCGCTGCCGAGGTCGAAGGCCGAGGGCTGGGCGTGCGCG
GCACCAGCGCGTCTGCGAGCCGTCTCGTCTCTCGGACTCGAGACGGAGGGCGGGCCCGCTT
40 CTTGCGGGGCGCGCGGGGCGCGGAGGCGCGGCGGCGACGAGACGGGGACGAGACATCGTC
CAGGGTGGGTGGACGGCGGGCCGCGCCGCTCCGCGCTCGGGGGTGGTCTCGCGCTGGTCCTCT

TCCCGACTGGCCATCTCCCACTGCTCCTTCTCTATAGGCAGAAAGAGATCATGGAGTCTCTCAT
GCGAGTCGAGAAGGAGGAGGACAGCCTAACC GCCCCTCTGAGCCCTCCACCACCGCCGCCACC
ACCGCCAATGCCGCCGCGGACGACGCGCCACCGAGACCACCGCCAGTACCACCCTCCCCAGCG
ACGCACCCCCGCTCGAGAATGAAGTGCTGATCGAGCAGGACCCGGGTTTTGTGAGCGGAGAGG
5 AGGATGAGGTGGATGAGAAGGAGAAGGAGGAGGTCGCCGCCTCAGTGCCAAAAGAGGATAAA
AAGCAAGACCAGGACGACGCAGATAAGGATGAGACAGCAGTCGGGCGGGGAACGGAAGCCA
TGATGCTGATGACGGCTACCTAGACGTGGGAGACGACGTGCTGCTTAAGCACCTGCACCGCCAG
TGCGTCATCGTCTGCGACGCGCTGCAGGAGCGCTGCGAAGTGCCCCCTGGACGTGGCGGAGGTCA
GCCGCGCCTACGAGCGGCACCTCTTCGCGCCGCACGTGCCCCCAAGCGCCGGGAGAACGGCAC
10 CTGCGAGCCCAACCCGCGTCTCAACTTCTACCCGGTCTTCGCGGTACCCGAGGTGCTGGCCACCT
ACCACATCTTTTTCCAAAAGTCAAGATCCCCCTCTCCTGCCGCGCCAACCGCACCCGCGCCGAC
AAAACCTTGACCCTGCGGCAGGGCGCCACATACCTGATATCGCCTCTCTGGAGGAAGTGCCCA
AGATCTTCGAGGGTCTCGGTGCGACGAGAAACGGGCGGCGAACGCTCTGCACGGAGACAGCG
AAAACGAGAGTCACTCGGGGGTGTGGTGGAGCTCGAGGGCGACAACGCGCGCCTGGCCGTAC
15 TCAAGCGCAGCATAGAGGTCACCCACTTTGCCTACCCGGCGCTCAACCTGCCCCCAAGGTCAT
GAGTGTGGTCATGGGCGAGCTCATCATGCGCCGCGCCAGCCCCCTGGCCGCGGATGCAAACTTG
CAAGAGTCCTCCGAGGAAGGCCTGCCCCGCGTCAGCGACGAGCAGCTGGCGCGCTGGCTGGAG
ACCCGCGACCCCGCGCAGCTGGAGGAGCGCGCAAGCTCATGATGGCCGCGGTGCTGGTCACCG
TGGAGCTCGAGTGTCTGCAGCGCTTCTTCGCGGACCCCGAGATGCAGCGCAAGCTCGAGGAGAC
20 CCTGCACTACACCTTCCGCCAGGGCTACGTGCGCCAGGCCTGCAAGATCTCCAACGTGGAGCTC
TGCAACCTGGTCTCCTACCTGGGCATCCTGCACGAGAACCGCCTCGGGCAGAACGTCCTGCACT
CCACCCTCAAAGGGGAGGCGCGCCGCGACTACATCCGCGACTGCGCCTACCTCTTCCTCTGCTAC
ACCTGGCAGACGGCCATGGGGGTCTGGCAGCAGTGCCTGGAGGAGCGCAACCTCAAGGAGCTG
GAAAAGCTCCTCAAGCGCACCTCAGGGACCTCTGGACGGGCTTCAACGAGCGCTCGGTGGCCG
25 CCGCGCTGGCGGACATCATCTTTCCCGAGCGCCTGCTCAAGACCCTGCAGCAGGGCCTGCCCCA
CTTACCAGCCAGAGCATGCTGCAGAACTTCAGGACTTTTCATCCTGGAGCGCTCGGGCATCCTGC
CGGCCACTTGCTGCGCGCTGCCAGCGACTTCGTGCCATCAAGTACAGGGAGTGCCCGCCGCC
GCTCTGGGGCCACTGCTACCTCTTCCAGCTGGCCAACTACCTCGCCTACCACTCGGACCTCATGG
AAGACGTGAGCGGCGAGGGCCTGCTCGAGTGCCACTGCCGCTGCAACCTCTGCACGCCCCACCG
30 CTCTCTAGTCTGCAACCCGCGAGCTGCTCAGCGAGAGTCAGATTATCGGTACCTTCGAGCTGCAGG
GTCCCTCGCCTGACGAGAAGTCCGCGGCTCCAGGGCTGAACTCACTCCGGGGCTGTGGACTTC
CGCCTACCTACGCAAATTTGTACCTGAGGACTACCACGCCCACGAGATCAGGTTCTACGAAGAC
CAATCCCGCCCCGCCAAGGCGGAGCTACCGCCTGCGTCATCACCCAGGGGCACATCCTGGGCC
AATTGCAAGCCATCAACAAAGCCCGCCGAGAGTTCTTGCTGAAAAAGGGTCGGGGGGTGTACCT
35 GGACCCCCAGTCCGGCGAGGAGCTAAACCCGCTACCCCGCCGCCGCCAGCAGCGGGACCTT
GCTTCCAGGATGGCACCCAGAAAGAAGCAGCAGCCGCCGCCGCCGCCGAGCCATACATGCTT
CTGGAGGAAGAGGAGGAGGACTGGGACAGTCAGGCAGAGGAGGTTTCGGACGAGGAGCAGGA
GGAGATGATGGAAGACTGGGAGGAGGACAGCAGCCTAGACGAGGAAGCTTCAGAGGCCGAAG
AGGTGGCAGACGCAACACCATCGCCCTCGGTGCGAGCCCCCTCGCCGGGGCCCCCTGAAATCCTC
40 CGAACCCAGCACCAGCGCTATAACCTCCGCTCCTCCGGCGCCGGCGCCACCCGCCCGCAGACCC
AACCGTAGATGGGACACCACAGGAACCGGGGTGCGTAAGTCCAAGTGCCCCGCCGCCGCCACCG

CAGCAGCAGCAGCAGCAGCGCCAGGGCTACCGCTCGTGGCGCGGGCACAAGAACGCCATAGTC
GCCTGCTTGCAAGACTGCGGGGGCAACATCTCTTTGCCCCGCCGCTTCCTGCTATTCCACCACGG
GGTCGCCTTTCCCCGCAATGTCCTGCATTACTACCGTCATCTCTACAGCCCCCTACTGCAGCGGGC
ACCCAGAGGCGGCAGCGGCAGCCACAGCGGCGACCACCACCTAGGAAGATATCCTCCGCGGGC
5 AAGACAGCGGCAGCAGCGGCCAGGAGACCCGCGGCAGCAGCGGCGGGAGCGGTGGGCGCACT
GCGCCTCTCGCCAACGAACCCCTCTCGACCCGGGAGCTCAGACACAGGATCTTCCCCACTTTGT
ATGCCATCTTCCAACAGAGCAGAGGCCAGGAGCAGGAGCTGAAAATAAAAAACAGATCTCTGC
GCTCCCTCACCCGACGCTGTCTGTATCAGAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAGGA
CGCGGAGGCACTCTTCAGCAAATACTGCGCGCTCACTCTTAAAGACTAGCTCCGCGCCCTTCTCG
10 AATTTAGGCGGGAGAAAACCTACGTCATCGCCGGCCGCCGCCAGCCCCGCCAGCCGAGATGAGC
AAAGAGATTCCACGCCATACATGTGGAGCTACCAGCCGCAGATGGGACTCGCGGCGGGAGCG
GCCCAGGACTACTCCACCCGCATGAACTACATGAGCGCGGGACCCACATGATCTCACAGGTCA
ACGGGATCCGCGCCCAGCGAAACCAAATACTGCTGGAACAGGCGGCCATCACCGCCACGCCCC
GCCATAATCTCAACCCCCGAAATTGGCCCCGCCGCCCTCGTGTACCAGGAAACCCCTCCGCCAC
15 CACCGTACTACTTCCGCGTGACGCCCAGGCCGAAGTCCAGATGACTAACTCAGGGGCGCAGCTC
GCGGGCGGCTTTTCGTACGGGGCGCGGCCGCTCCGACCAGGTATAAGACACCTGATGATCAGAG
GCCGAGGTATCCAGCTCAACGACGAGTCGGTGAGCTCTTCGCTCGGTCTCCGTCCGACGGAAC
TTTCAGCTCGCCGGATCCGGCCGCTCTTCGTTACGCCCCGCCAGGCGTACCTGACTCTGCAGA
CCTCGTCTCGGAGCCCCGCTCCGGCGGCATCGGAACCCCTCCAGTTCGTGGAGGAGTTCGTGCCC
20 TCGGTCTACTTCAACCCCTTCTCGGGACCTCCCGGACGCTACCCCGACCAGTTCATTCCGAACCT
TGACGCGGTGAAGGACTCGGCGGACGGCTACGACTGAATGTCAGGTGTCGAGGCAGAGCAGCT
TCGCCTGAGACACCTCGAGCACTGCCGCCGCCACAAGTGCTTCGCCCCGCGTTCTGGTGAGTTCT
GCTACTTTTACGCTACCCGAGGAGCATACCGAGGGGCCGGCGCACGGCGTCCGCCTGACCACCCA
GGGCGAGGTTACCTGTTCCCTCATCCGGGAGTTTACCCTCCGTCCCCCTGCTAGTGGAGCGGGAGC
25 GGGGTCCCTGTGTCTAACTATCGCCTGCAACTGCCCTAACCTGGATTACATCAAGATCTTTGC
TGTCATCTCTGTGCTGAGTTAATAAACGCTGAGATCAGAATCTACTGGGGCTCCTGTGCCATC
CTGTGAACGCCACCGTCTTCACCCACCCCGACCAGGCCAGGCGAACCTCACCTGCGGTCTGCA
TCGGAGGGCCAAGAAGTACCTCACCTGGTACTTCAACGGCACCCCTTTGTGGTTTACAACAGCT
TCGACGGGGACGGAGTCTCCCTGAAAGACCAGCTCTCCGGTCTCAGCTACTCCATCCACAAGAA
30 CACCACCTCCAACCTTCCCTCCCTACCTGCCGGGAACCTACGAGTGCGTCACCGGCCGCTGCA
CCCACCTACCCGCCTGATCGTAAACCAGAGCTTTCCGGGAACAGATAACTCCCTCTTCCCCAGA
ACAGGAGGTGAGCTCAGGAAACTCCCCGGGGACCAGGGCGGAGACGTACCTTCGACCCTTGTG
GGGTTAGGATTTTTTATTACCGGGTTGCTGGCTCTTTTAATCAAAGTTTCCTTGAGATTTGTCTT
TCCTTCTACGTGTATGAACACCTCAACCTCCAATAACTCTACCCTTTCTTCGGAATCAGGTGACTT
35 CTCTGAAATCGGGCTTGGTGTGCTGCTTACTCTGTTGATTTTTTTCCTTATCATACTCAGCCTTCT
GTGCCTCAGGCTCGCCGCTGCTGCGCACACATCTATATCTACTGCTGGTTGCTCAAGTGCAGGG
GTCGCCACCCAAGATGAACAGGTACATGGTCTATCGATCCTAGGCCTGCTGGCCCTGGCGGCC
TGCAGCGCCGCCAAAAAAGAGATTACCTTTGAGGAGCCCGCTTGCAATGTAACCTTCAAGCCCG
AGGGTGACCAATGCACCACCTCGTCAAATGCGTTACCAATCATGAGAGGCTGCGCATCGACTA
40 CAAAAACAAAACCTGGCCAGTTTGGGTCTATAGTGTGTTTACGCCCCGAGACCCCTCTAACTACT
CTGTACCGTCTTCCAGGGCGGACAGTCTAAGATATTCAATTACACTTTCCTTTTTATGAGTTAT

[illegible]

CTGACGCCCTAGATCTAGAAATGGACGGCATCAGTACCGAGCAGCGTCTCCTAGAGAGGCGCAG
GCAGGCGGCTGAGCAAGAGCGCCTCAATCAGGAGCTCCGAGATCTCGTTAACCTGCACCAGTGC
AAAAGAGGCATCTTTTGTCTGGTAAAGCAGGCCAAAGTCACCTACGAGAAGACCGGCAACAGC
CACCGCCTCAGTTACAAATTGCCACCCAGCGCCAGAAGCTGGTGCTCATGGTGGGTGAGAATC
5 CCATCACCGTCACCCAGCACTCGGTAGAGACCGAGGGGTGTCTGCACTCCCCCTGTCGGGGTCC
AGAAGACCTCTGCACCCTGGTAAAGACCCTGTGCGGTCTCAGAGATTTAGTCCCCTTTAACTAAT
CAAACACTGGAATCAATAAAAAGAATCACTTACTTAAAATCAGACAGCAGGTCTCTGTCCAGTT
TATTCAGCAGCACCTCCTTCCCCCTCCTCCCAACTCTGGTACTCCAAACGCCTTCTGGCGGCAAAC
TTCCTCCACACCCTGAAGGGAATGTCAGATTCTTGCTCCTGTCCCTCCGCACCCACTATCTTCATG
10 TTGTTGCAGATGAAGCGCACCAAAACGTCTGACGAGAGCTTCAACCCCGTGTACCCCTATGACA
CGGAAAGCGGCCCTCCCTCCGTCCCTTTCCTACCCCTCCCTTCGTGTCTCCCGATGGATTCCAA
GAAAGTCCCCCGGGTCTGTCTCTGAACCTGGCCGAGCCCCCTGGTCACTTCCCACGGCATGCT
CGCCCTGAAAATGGGAAGTGGCCTCTCCCTGGACGACGCTGGCAACCTCACCTCTCAAGATATC
ACCACCGCTAGCCCTCCCTCAAAAAAACCAAGACCAACCTCAGCCTAGAAACCTCATCCCCC
15 TAACTGTGAGCACCTCAGGCGCCCTACCGTAGCAGCCGCCGCTCCCCTGGCGGTGGCCGGCAC
CTCCCTCACCATGCAATCAGAGGCCCCCTGACAGTACAGGATGCAAAACTCACCTTGGCCACC
AAAGGCCCCCTGACCGTGTCTGAAGGCAAACCTGGCCTTGCAAACATCGGCCCCGCTGACGGCCG
CTGACAGCAGCACCTCACAGTCAGTGCCACACCACCCCTTAGCACAAGCAATGGCAGCTTGGG
TATTGACATGCAAGCCCCCATTTACACCACCAATGGAAAACCTAGGACTTAACTTTGGCGCTCCCC
20 TGCATGTGGTAGACAGCCTAAATGCACTGACTGTAGTTACTGGCCAAGGTCTTACGATAAACGG
AACAGCCCTACAACTAGAGTCTCAGGTGCCCTCAACTATGACACATCAGGAAACCTAGAATTG
AGAGCTGCAGGGGGTATGCGAGTTGATGCAAAATGGTCAACTTATCCTTGATGTAGCTTACCCATT
TGATGCACAAAACAATCTCAGCCTTAGGCTTGGACAGGGACCCCTGTTTGTTAACTCTGCCACA
ACTTGGATGTAACTACAACAGAGGCCTCTACCTGTTACATCTGGAAATACCAAAAAGCTAGA
25 AGTTAATATCAAAACAGCCAAGGGTCTCATTTATGATGACACTGCTATAGCAATCAATGCGGGT
GATGGGCTACAGTTTGACTCAGGCTCAGATACAAATCCATTAAAAACTAAACTTGGATTAGGAC
TGGATTATGACTCCAGCAGAGCCATAATTGCTAAACTGGGAACTGGCCTAAGCTTTGACAACAC
AGGTGCCATCACAGTAGGCAACAAAAATGATGACAAGCTTACCTTGTGGACCACACCAGACCCA
TCCCCTAACTGTAGAATCTATTAGAGAAAGATGCTAAATTCACACTTGTTTTGACTAAATGCGG
30 CAGTCAGGTGTTGGCCAGCGTTTCTGTTTTATCTGTAAAAGGTAGCCTTGCGCCCATCAGTGCCA
CAGTAACTAGTGCTCAGATTGTCTCAGATTGATGAAAATGGAGTTCTACTAAGCAATTCTTCC
CTTGACCCTCAATACTGGAACTACAGAAAAGGTGACCTTACAGAGGGCACTGCATATACCAACG
CAGTGGGATTTATGCCCAACCTCACAGCATACCCAAAAACACAGAGCCAAACTGCTAAAAGCAA
CATTGTAAGTCAGGTTTACTTGAATGGGGACAAATCCAAACCCATGACCCTCACCATACCCTCA
35 ATGGAATAATGAAACAGGAGATGCCACAGTAAGCACTTACTCCATGTCATTCTCATGGAACTG
GAATGGAAGTAATTACATTAATGAAACGTTCCAAACCAACTCCTTCACCTTCTCCTACATCGCCC
AAGAATAAAAAGCATGACGCTGTTGATTGATTCAATGTGTTTCTGTTTTATTTCAAGCACAAC
AAAATCATTCAAGTCATTCTTCCATCTTAGCTTAATAGACACAGTAGCTTAATAGACCCAGTAGT
GCAAAGCCCCATTCTAGCTTATAACTAGTGGAGAAGTACTCGCCTACATGGGGGTAGAGTCATA
40 ATCGTGCATCAGGATAGGGCGGTGGTGTCTGACGAGCGCGGAATAAACTGCTGCCGCCGCCG
TCCGTCTCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCGATGATTTCGACCGCCCGCAGCA

TAAGGCGCCTTGTCTCCGGGCACAGCAGCGCACCCCTGATCTCACTTAAATCAGCACAGTAACT
GCAGCACAGCACCACAATATTGTTCAAAATCCACAGTGCAAGGCGCTGTATCCAAAGCTCATG
GCGGGGACCACAGAACCCACGTGGCCATCATACCACAAGCGCAGGTAGATTAAAGTGGCGACCC
CTCATAAACACGCTGGACATAAACATTACCTCTTTTGGCATGTTGTAATTCACCACCTCCCGGTA
5 CCATATAAACCTCTGATTAAACATGGCGCCATCCACCACCATCCTAAACCAGCTGGCCAAAACC
TGCCCGCCGGCTATACACTGCAGGGAACCGGGACTGGAACAATGACAGTGGAGAGCCCAGGAC
TCGTAACCATGGATCATCATGCTCGTCATGATATCAATGTTGGCACAACACAGGCACACGTGCA
TACACTTCCTCAGGATTACAAGCTCCTCCCGGTTAGAACCATATCCCAGGGAACAACCCATTCC
TGAATCAGCGTAAATCCACACTGCAGGGAAGACCTCGCACGTAACCTCACGTTGTGCATTGTCA
10 AAGTGTTACATTGCGGCAGCAGCGGATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTCTCAAA
AGGAGGTAGACGATCCCTACTGTACGGAGTGCGCCGAGACAACCGAGATCGTGTTGGTCGTAGT
GTCATGCCAAATGGAACGCCGGACGTAGTCATATTTCTGAAGTCTTAGATCTCTCAACGCAGC
ACCAGCACCAACACTTCGCAGTGTAAGGCCAAGTGCCGAGAGAGTATATATAGGAATAAAA
AGTGACGTAAACGGGCAAGTCCAAAAAACGCCAGAAAAACCGCACGCGAACCTACGCCCCG
15 AAACGAAAGCCAAAAAACACTAGACACTCCCTTCCGGCGTCAACTTCCGCTTCCCACGCTACG
TCACTTGCCCCAGTCAAACAACTACATATCCCGAAGTTCGAAGTCGCCACGCCAAAAACCCG
CCTACACCTCCCCGCCGCCGGCCCCGCCCAACCCGCTCCCGCCCCGCGCCCCGCCCGCGC
CGCCCATCTCATTATCATATTGGCTTCAATCCAAAAATAAGGTATATTATTGATGATGGTTTAAAC
GGATCCAATTCTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGAT
20 AATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTT
TATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAA
TAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC
GGCATTTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATC
AGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTT
25 TCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTAT
CCCGTGTGACGCCGGGCAAGAGCAACTCGGTGCGGCATACACTATTCTCAGAATGACTTGGT
TGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGT
GCTGCCATAACCATGAGTGATAAACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGA
AGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAAGTTCGCTTGATCGTTGGGAACC
30 GGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAAC
AACGTTGCGCAAATTAATACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACT
GGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATT
GCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATG
GTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA
35 TAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAAGCATTGGTAACTGTCAGACCAAGTTTAC
TCATATATACTTTAGATTGATTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACC
AAAATCCCTTAACGTGAGTTTTCTGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT
CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAACAAAAAAACCACCGCTACCA
GCGGTGGTTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAG
40 AGCGCAGATACCAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTG
TAGCACCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAG

TCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCTGGGCTGAA
 CGGGGGGTTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
 GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAG
 CGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTA
 5 TAGTCCTGTCTGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGC
 GGAGCCTATGGA AAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTGA
 AGCTGTCCCTGATGGTCGTCATCTACCTGCCTGGACAGCATGGCCTGCAACGCGGGCATCCCCGAT
 GCCGCCGGAAGCGAGAAGAATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCAGATCCGAA
 TTCGTTTAAAC
 10

SEQ ID NO: 8 - Polynucleotide sequence encoding ChAd155#1390 backbone construct

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGATGGGCGGCGCGGG
 15 GCGGGGCGCGGGGCGGGAGGCGGGTTTGGGGGCGGGCCGGCGGGCGGGGCGGTGTGGCGGAA
 GTGGACTTTGTAAGTGTGGCGGATGTGACTTGCTAGTGCCGGGCGCGGTAAAGTGACGTTTTTC
 CGTGCGCGACAACGCCCCCGGGAAGTGACATTTTTCCCGCGTTTTTACCGGATGTTGTAGTGAA
 TTTGGGCGTAACCAAGTAAGATTTGGCCATTTTCGCGGGAAAAC TGAAACGGGGGAAGTGAAATC
 TGATTAATTTTTCGTTAGTCATACCGCGTAATATTTGTCTAGGGCCGAGGGACTTTGGCCGATTA
 20 CGTGAGGACTCGCCAGGTGTTTTTGGAGTGAATTTCCGCGTTCCGGGTCAAAGTCTGCGTTT
 TATTATTATAGGATATCCCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCT
 CATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACG
 GGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCC
 TGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 25 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACCTGGCAGT
 ACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCC
 TGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
 CATCGCTATTACCATGGTGTATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
 CACGGGGATTTC AAGTCTCCACCCCATGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAA
 30 CGGGACTTTCCAAAATGTCGTAACAAC TCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
 GGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAG
 AGATCGTCGACGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTG
 ACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGC
 GGATTCCCCGTGCCAAGAGTGAGATCTTCGTTTATCTAGGTACCGGGCCCCCCCCCTCGAGGTCTGA
 35 CGGTATCGATAAGCTTACGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGG
 AACACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGCTA
 TCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGC
 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACC GGAATTGCCAGCTGGGGCGCCCTC
 TGGTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGG
 40 CGCAGGGGATCAAGATCTAACCAGGAGCTATTTAATGGCAACAGTTAACCAGCTGGTACGCAAA
 CCACGTGCTCGCAAAGTTGCGAAAAGCAACGTGCCTGCGCTGGAAGCATGCCCCGAAAAACGTG

GCGTATGTACTCGTGTATATACTACCACTCCTAAAAAACCGAACTCCGCGCTGCGTAAAGTATGC
CGTGTTCTGTCTGACTAACGGTTTCGAAGTGACTTCCTACATCGGTGGTGAAGGTCACAACCTGCA
GGAGCACTCCGTGATCCTGATCCGTGGCGGTCGTGTTAAAGACCTCCCGGGTGTTCGTTACCACA
CCGTACGTGGTGCCTTACTGCTCCGGCGTTAAAGACCGTAAGCAGGCTCGTTCCAAGTATGG
5 CGTGAAGCGTCCTAAGGCTTAATGGTAGATCTGATCAAGAGACAGGATGACGGTCGTTTCGCAT
GCTTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGCTAT
GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGC
GCCCCGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGC
GCGGCTATCGTGGCTGGCCACGACGGCGTTCTTTCGCGCAGCTGTGCTCGACGTTGTCACTGAA
10 GCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTCATCTCACCTTG
CTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGC
TACCTGCCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCC
GGTCTTGTCTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAAGTTCG
CCAGGCTCAAGGCGCGCATGCCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTT
15 GCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTG
GCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAAT
GGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGACGCGCATCGCCTTCTAT
CGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCC
CAACCTGCCATCACGAGATTTGATTCCACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCG
20 TTTCCGGGACGCGCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCAC
CCCGGGCTCGATCCCCTCGGGGGGAATCAGAATTCAGTCGACAGCGGCCGCGATCTGCTGTGCC
TTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCTTCCTTGACCCCTGGAAGGTGCCAC
TCCCACTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTA
TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
25 CTGGGGATGCGGTGGGCTCTATGGCCGATCAGCGATCGCTGAGGTGGGTGAGTGGGCGTGGCCT
GGGGTGGTCATGAAAATATATAAGTTGGGGGTCTTAGGGTCTCTTTATTTGTGTTGCAGAGACCG
CCGGAGCCATGAGCGGGAGCAGCAGCAGCAGCAGTAGCAGCAGCGCCTTGATGGCAGCATCG
TGAGCCCTTATTTGACGACGCGGATGCCCCACTGGGCCGGGGTGCCTCAGAATGTGATGGGCTC
CAGCATCGACGGCCGACCCGTCTGCCCCGAAATTCCGCCACGCTGACCTATGCGACCGTCGCG
30 GGGACGCCGTTGGACGCCACCGCCGCCGCCGCCGCCACCGCAGCCGCTCGGCCGTGCGCAGCC
TGGCCACGGACTTTGCATTCTTGGGACCACTGGCGACAGGGGCTACTTCTCGGGCCGCTGCTGCC
GCCGTTTCGCGATGACAAGCTGACCGCCCTGCTGGCGCAGTTGGATGCGCTTACTCGGGAAGTGG
GTGACCTTTCTCAGCAGGTCTATGGCCCTGCGCCAGCAGGTCTCCTCCCTGCAAGCTGGCGGGAAT
GCTTCTCCACAAATGCCGTTTAAGATAAATAAAAACCACTCTGTTTGGATTAAAGAAAAGTA
35 GCAAGTGCAATTGCTCTCTTTATTTTCATAATTTTCCGCGCGCGATAGGCCCTAGACCAGCGTTCTC
GGTCGTTGAGGGTGGCGGTGTATCTTCTCCAGGACGTGGTAGAGGTGGCTCTGGACGTTGAGATA
CATGGGCATGAGCCCGTCCCGGGGGTGGAGGTAGCACTGACAGGCTTCATGCTCCGGGGTG
GTGTTGTAGATGATCCAGTCGTAGCAGGAGCGCTGGGCATGGTGCCTAAAAATGTCCTTCAGCA
GCAGGCCGATGGCCAGGGGGAGGCCCTTGGTGTAAGTGTTTACAAAACGGTTAAGTTGGGAAG
40 GGTGCATTCCGGGAGAGATGATGTGCATCTTGGACTGTATTTTAGATTGGCGATGTTTCCGCCC
AGATCCCTTCTGGGATTCATGTTGTGCAGGACCACAGTACAGTGTATCCGGTGCCTTGGGGA

ATTTGTCATGCAGCTTAGAGGGAAAAGCGTGGAAGAACTTGGAGACGCCTTTGTGGCCTCCCAG
ATTTTCCATGCATTTCGTCCATGATGATGGCAATGGGCCCCGCGGAGGCAGCTTGGGCAAAGATA
TTTCTGGGGTCGCTGACGTCGTAGTTGTGTTCCAGGGTGAGGTCGTCATAGGCCATTTTTACAAA
GCGCGGGCGGAGGGTGCCCCACTGGGGGATGATGGTCCCCTCTGGCCCTGGGGCGTAGTTGCC
5 TCGCAGATCTGCATTTCCCAGGCCTTAATCTCGGAGGGGGGAATCATATCCACCTGCGGGGCGA
TGAAGAAAACGGTTTCCGGAGCCGGGGAGATTAACTGGGATGAGAGCAGGTTTCTAAGCAGCT
GTGATTTTCCACAACCGGTGGGCCCATAAATAACACCTATAACCGGTTGCAGCTGGTAGTTTAG
AGAGCTGCAGCTGCCGTCGTCCCGAGGAGGGGGGCCACCTCGTTGAGCATGTCCCTGACGCGC
ATGTTCTCCCCGACCAGATCCGCCAGAAGGCGCTCGCCGCCAGGGACAGCAGCTCTTGCAAGG
10 AAGCAAAGTTTTTCAGCGGCTTGAGGCCGTCCGCCGTGGGCATGTTTTTCAGGGTCTGGCTCAGC
AGCTCCAGGCGGTCCCAGAGCTCGGTGACGTGCTCTACGGCATCTCTATCCAGCATATCTCCTCG
TTTCGCGGGTTGGGGCGACTTTCGCTGTAGGGCACCAAGCGGTGGTTCGTCCAGCGGGGCCAGAG
TCATGTCCTTCCATGGGCGCAGGGTCTCTCGTCAGGGTGGTCTGGGTACGGTGAAGGGGTGCGC
TCCGGGCTGAGCGCTTGCCAAGGTGCGCTTGAGGCTGGTTCTGCTGGTGCTGAAGCGCTGCCGG
15 TCTTCGCCCTGCGCGTCGGCCAGGTAGCATTGACCATGGTGTTCATAGTCCAGCCCCTCCGCGGC
GTGTCCCTTGGCGCGCAGCTTGCCCTTGGAGGTGGCGCCGCACGAGGGGCAGAGCAGGCTCTTG
AGCGCGTAGAGCTTGGGGGCGAGGAAGACCGATTTCGGGGGAGTAGGCGTCCGCGCCGCAGACC
CCGCACACGGTCTCGCACTCCACCAGCCAGGTGAGCTCGGGGCGCGCCGGGTCAAAAACAGGT
TTCCCCCATGCTTTTTGATGCGTTTCTTACCTCGGGTCTCCATGAGGTGGTGTCCCCGCTCGGTGA
20 CGAAGAGGCTGTCCGTGTCTCCGTAGACCGACTTGAGGGGTCTTTTCTCCAGGGGGGTCCCTCG
TCTTCCTCGTAGAGGAACTCGGACCACTCTGAGACGAAGGCCCGCTCCAGGCCAGGACGAAGG
AGGCTATGTGGGAGGGGTAGCGGTCTGTTGTCCACTAGGGGGTCCACCTTCTCCAAGGTGTGAAG
ACACATGTCGCCTTCCTCGGCGTCCAGGAAGGTGATTGGCTTGTAGGTGTAGGCCACGTGACCG
GGGGTTCCTGACGGGGGGGTATAAAAGGGGGTGGGGGCGCGCTCGTCGTCCTCTTCCGCAT
25 CGCTGTCTGCGAGGGCCAGCTGCTGGGGTGAGTATTCCCTCTCGAAGGCGGGCATGACCTCCGC
GCTGAGGTTGTCAGTTTCCAAAAACGAGGAGGATTTGATGTTACCTGTCCCGAGGTGATACCTT
TGAGGGTACCCGCGTCCATCTGGTCAGAAAACACGATCTTTTATTGTCCAGCTTGGTGGCGAAC
GACCCGTAGAGGGCGTTGGAGAGCAGCTTGGCGATGGAGCGCAGGGTCTGGTTCTTGTCCCTGT
CGGCGCGCTCCTTGGCCGCGATGTTGAGCTGCACGTACTCGCGCGCGACGCAGCGCCACTCGGG
30 GAAGACGGTGGTTCGCTCGTCGGGCACCAGGCGCACGCGCCAGCCGCGGTTGTGCAGGGTGAC
CAGGTCCACGCTGGTGGCGACCTCGCCGCGCAGGCGCTCGTTGGTCCAGCAGAGACGGCCGCC
TTGCGCGAGCAGAAGGGGGGCGAGGGGGTTCGAGCTGGGTCTCGTCCGGGGGGTCCGCGTCCACG
GTGAAAACCCCGGGGCGCAGGCGCGCGTCAAGTAGTCTATCTTGCAACCTTGCATGTCCAGCG
CCTGCTGCCAGTCGCGGGCGGCGAGCGCGCGCTCGTAGGGGTTGAGCGGCGGGCCCCAGGGCAT
35 GGGGTGGGTGAGTGCGGAGGCGTACATGCCGAGATGTCATAGACGTAGAGGGGCTCCCGCAG
GACCCCGATGTAGGTGGGGTAGCAGCGGCCGCCGCGGATGCTGGCGCGCACGTAGTCATACAGC
TCGTGCGAGGGGGCGAGGAGGTCGGGGCCCAGGTTGGTGCGGGCGGGGCGCTCCGCGCGGAAG
ACGATCTGCCTGAAGATGGCATGCGAGTTGGAAGAGATGGTGGGGCGCTGGAAGACGTTGAAG
CTGGCGTCTGCAGGCCGACGGCGTCGCGCACGAAGGAGGCGTAGGAGTCGCGCAGCTTGTGTA
40 CCAGCTCGGCGGTGACCTGCACGTCGAGCGCGCAGTAGTCGAGGGTCTCGCGGATGATGTCATA
TTTAGCCTGCCCTTCTTTTTCCACAGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTCCAGTA

CTCTTGGATCGGGAAACCGTCCGGTTCCGAACGGTAAGAGCCTAGCATGTAGAACTGGTTGACG
GCCTGGTAGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGTAGGCCTGCGCGGCCCTTGCGGAGCG
AGGTGTGGGTCAGGGCGAAGGTGTCCCTGACCATGACTTTGAGGTACTGGTGCTTGAAGTCGGA
GTCGTGCGAGCCGCCCCGCTCCCAGAGCGAGAAGTCGGTGCGCTTCTTGAGCGGGGGTTGGGC
5 AGAGCGAAGGTGACATCGTTGAAGAGGATTTTGTCCGCGCGGGGCATGAAGTTGCGGGTGATGC
GGAAGGGCCCCGGCACTTCAGAGCGGTTGTTGATGACCTGGGCGGCGAGCACGATCTCGTCGAA
GCCGTTGATGTTGTGGCCACGATGTAGAGTTCCAGGAAGCGGGGCCGGCCCTTTACGGTGGGC
AGCTTCTTTAGCTCTTCGTAGGTGAGCTCCTCGGGCGAGGCGAGGCCGTGCTCGGCCAGGGCCC
AGTCCGCGAGGTGCGGGTTGTCTCTGAGGAAGGACTTCCAGAGGTCGCGGGCCAGGAGGGTCTG
10 CAGGCGGTCTCTGAAGGTCCTGAAC TGGCGGCCACGGCCATTTTTTCGGGGGTGATGCAGTAG
AAGGTGAGGGGGTCTTGCTGCCAGCGGTCCCAGTCGAGCTGCAGGGCGAGGTCGCGCGCGGCG
GTGACCAGGCGCTCGTCGCCCCGAATTTTCATGACCAGCATGAAGGGCACGAGCTGCTTTCCGA
AGGCCCCCATCCAAGTGTAGGTCTCTACATCGTAGGTGACAAAGAGGCGCTCCGTGCGAGGATG
CGAGCCGATCGGGAAGAACTGGATCTCCCGCCACCAGTTGGAGGAGTGGCTGTTGATGTGGTGG
15 AAGTAGAAGTCCCGTCGCCGGGCCGAACACTCGTGCTGGCTTTTGTAAGCGAGCGCAGTACT
GGCAGCGCTGCACGGGTGTACCTCATGCACGAGATGCACCTTTTCGCCCCGCGCACGAGGAAGCC
GAGGGGAAATCTGAGCCCCCGCCTGGCTCGCGGCATGGCTGGTTCTTCTACTTTGGATGCGT
GTCCGTCTCCGTCTGGCTCCTCGAGGGGTGTTACGGTGGAGCGGACCACCACGCCGCGCGAGCC
GCAGGTCCAGATATCGGCGCGCGGCGGTCCGAGTTTGATGACGACATCGCGCAGCTGGGAGCTG
20 TCCATGGTCTGGAGCTCCCGCGGCGGCGGCAGGTCAGCCGGGAGTTCTTGACAGGTTACCTCGC
AGAGTCGGGCCAGGGCGCGGGGCAGGTCTAGGTGGTACCTGATCTCTAGGGGCGTGTTGGTGGC
GGCGTCGATGGCTTGACAGGAGCCCGCAGCCCCGGGGGGCGACGACGGTGCCCCGCGGGGTGGT
GGTGGTGGTGGCGGTGCAGCTCAGAAGCGGTGCCGCGGGCGGGCCCCCGAGGTAGGGGGGGC
TCCGGTCCCGCGGGCAGGGGCGGCAGCGGCACGTCCGGCGTGAGCGCGGGCAGGAGTTGGTGC
25 TGTGCCCCGAGGTTGCTGGCGAAGGCGACGACGCGGCGGTTGATCTCCTGGATCTGGCGCCTCT
GCGTGAAAGACGACGGGCCCGGTGAGCTTGAACCTGAAAGAGAGTTGACAGAATCAATCTCGG
TGTCATTGACCGCGCCTGGCGCAGGATCTCCTGCACGTCTCCCGAGTTGTCTTGGTAGGCGATC
TCGGCCATGAACTGCTCGATCTCTTCTCCTGGAGGTCTCCGCGTCCGGCGCGTTCCACGGTGGC
CGCCAGGTGCTTGAGATGCGCCCCATGAGCTGCGAGAAGGCGTTGAGTCCGCCCTCGTTCCAG
30 ACTCGGCTGTAGACCACGCCCCCTGGTCATCGCGGGCGCGCATGACCACCTGCGCGAGGTTGA
GCTCCACGTGCCGCGGAAGACGGCGTAGTTGCGCAGACGCTGGAAGAGGTAGTTGAGGGTGG
TGGCGGTGTGCTCGGCCACGAAGAAGTTCATGACCCAGCGGCGCAACGTGGATTGCTTGATGTC
CCCCAAGGCCTCCAGCCGTTCCATGGCCTCGTAGAAGTCCACGGCGAAGTTGAAAACTGGGAG
TTGCGCGCCGACACGGTCAACTCCTCCTCCAGAAGACGGATGAGCTCGGCGACGGTGTGCGCA
35 CCTCGCGCTCGAAGGCTATGGGGATCTCTTCTCCTCCGCTAGCATCACACCTCCTCCTCTCCTCCT
CTTCTGGCACTTCCATGATGGCTTCTCCTCTTCGGGGGGTGGCGGCGGCGGCGGTGGGGGAGG
GGGCGCTCTGCGCCGGCGGCGGCGCACCGGGAGGCGGTCCACGAAGCGCGCGATCATCTCCCCG
CGGCGGCGGCGCATGGTCTCGGTGACGGCGCGGCCGTTCTCCCGGGGGCGCAGTTGGAAGACGC
CGCCGACATCTGGTGTGTTGGGCGGGTGGCCGTGAGGCAGCGAGACGGCGCTGACGATGCATCT
40 CAACAATTGCTGCGTAGGTACGCCGCCGAGGGACCTGAGGGAGTCCATATCCACCGGATCCGAA
AACCTTTCAGGAAGGCGTCTAACCAAGTCGCAGTCGCAAGGTAGGCTGAGCACCGTGGCGGGCG

GCGGGGGGTGGGGGGAGTGTCTGGCGGAGGTGCTGCTGATGATGTAATTGAAGTAGGCGGACTT
GACACGGCGGATGGTCGACAGGAGACCATGTCCTTGGGTCCGGCCTGCTGGATGCGGAGGCGG
TCGGCTATGCCCCAGGCTTCGTTCTGGCATCGGCGCAGGTCTTGTAGTAGTCTTGCATGAGCCT
TTCCACCGGCACCTCTTCTCCTTCCTCTTCTGCTTCTTCCATGTCTGCTTCGGCCCTGGGGCGGCG
5 CCGCGCCCCCTGCCCCCATGCGCGTGACCCGAACCCCTGAGCGGTTGGAGCAGGGCCAGG
TCGGCGACGACGCGCTCGGCCAGGATGGCCTGCTGCACCTGCGTGAGGGTGGTTTGGAAAGTCAT
CCAAGTCCACGAAGCGGTGGTAGGCGCCCGTGTGATGGTGTAGGTGCAGTTGGCCATGACGGA
CCAGTTGACGGTCTGGTGGCCCGGTTGCGACATCTCGGTGTACCTGAGTCGCGAGTAGGCGCGG
GAGTCGAAGACGTAGTCGTTGCAAGTCCGCACCAAGTACTGGTAGCCACCAGGAAGTGCGGCG
10 GCGGCTGGCGGTAGAGGGGCCAGCGCAGGGTGGCGGGGGCTCCGGGGGCCAGGTCTTCCAGCA
TGAGGCGGTGGTAGGCGTAGATGTACCTGGACATCCAGGTGATACCCGCGGCGGTGGTGGAGGC
GCGCGGGAAGTCGCGCACCCGGTTCCAGATGTTGCGCAGGGGCAGAAAGTGCTCCATGGTAGGC
GTGCTCTGTCCAGTCAGACGCGCGCAGTCGTTGATACTCTAGACCAGGGAAAACGAAAGCCGGT
CAGCGGGCACTCTTCCGTGGTCTGGTGAATAGATCGCAAGGGTATCATGGCGGAGGGCCTCGGT
15 TCGAGCCCCGGTCCGGGCCGGACGGTCCGCCATGATCCACGCGGTTACCGCCCGCGTGTGAA
CCCAGGTGTGCGACGTCAGACAACGGTGGAGTGTTCCTTTTGGCGTTTTTCTGGCCGGGCGCCGG
CGCCCGCTAAGAGACTAAGCCGCGAAAGCGAAAGCAGTAAGTGGCTCGCTCCCCGTAGCCGGA
GGGATCCTTGCTAAGGGTTGCGTTGCGGCGAACCCCGGTTCAATCCCGTACTCGGGCCGGCCG
GACCCGCGGCTAAGGTGTTGGATTGGCCTCCCCCTCGTATAAAGACCCCGCTTGCGGATTGACTC
20 CGGACACGGGGACGAGCCCCCTTTATTTTGGCTTTCCCCAGATGCATCCGGTGCTGCGGCAGATG
CGCCCCCGCCCCAGCAGCAGCAACAACACCAGCAAGAGCGGCAGCAACAGCAGCGGGAGTCA
TGCAGGGCCCCCTCACCCACCCTCGGCGGGCCGGCCACCTCGGCGTCCGCGGCCGTGTCTGGCG
CCTGCGGCGGCGGCGGGGGGCCGGCTGACGACCCCGAGGAGCCCCGCGGCGCAGGGCCAGAC
ACTACCTGGACCTGGAGGAGGGCGAGGGCTGGCGCGGCTGGGGGCGCCGTCTCCCGAGCGCC
25 ACCCGCGGGTGCAGCTGAAGCGCGACTCGCGCAGGGCGTACGTGCCTCGGCAGAACCTGTTTCAG
GGACCGCGCGGGCGAGGAGCCCCGAGGAGATGCGGGACAGGAGGTTTCAGCGCAGGGCGGGAGC
TGCGGCAGGGGCTGAACCGCGAGCGGCTGCTGCGCGAGGAGGACTTTGAGCCCGACGCGCGGA
CGGGGATCAGCCCCGCGCGCGCGCACGTGGCGGCCGCCGACCTGGTGACGGCGTACGAGCAGA
CGGTGAACCAGGAGATCAACTTCCAAAAGAGTTTCAACAACCACGTGCGCACGCTGGTGGCGCG
30 CGAGGAGGTGACCATCGGGCTGATGCACCTGTGGGACTTTGTAAGCGCGCTGGTGCAGAACCCC
AACAGCAAGCCTCTGACGGCGCAGCTGTTCTGATAGTGCAGCACAGCAGGGACAACGAGGCG
TTTAGGGACGCGCTGCTGAACATCACCGAGCCCGAGGGTCCGGTGGCTGCTGGACCTGATTAACA
TCCTGCAGAGCATAGTGGTGCAGGAGCGCAGCCTGAGCCTGGCCGACAAGGTGGCGGCCATCA
ACTACTCGATGCTGAGCCTGGGCAAGTTTTACGCGCGCAAGATCTACCAGACGCCGTACGTGCC
35 CATAGACAAGGAGGTGAAGATCGACGGTTTTTACATGCGCATGGCGCTGAAGGTGCTCACCTG
AGCGACGACCTGGGCGTGTACCGCAACGAGCGCATCCACAAGGCCGTGAGCGTGAGCCGGCGG
CGCGAGCTGAGCGACCGCGAGCTGATGCACAGCCTGCAGCGGGCGCTGGCGGGCGCCGGCAGC
GGCGACAGGGAGGCGGAGTCCTACTTCGATGCGGGGGCGGACCTGCGCTGGGCGCCCAGCCGG
CGGGCCCTGGAGGCCGCGGGGGTCCGCGAGGACTATGACGAGGACGGCGAGGAGGATGAGGAG
40 TACGAGCTAGAGGAGGGCGAGTACCTGGACTAAACCGCGGGTGGTGTTCGGGTAGATGCAAG
ACCCGAACGTGGTGGACCCGCGCTGCGGGCGGCTCTGCAGAGCCAGCCGTCCGGCCTTAATC

CTCAGACGACTGGCGACAGGTCATGGACCGCATCATGTCGCTGACGGCGCGTAACCCGGACGCG
TTCCGGCAGCAGCCGAGGCCAACAGGCTCTCCGCCATCCTGGAGGCGGTGGTGCCTGCGCGCT
CGAACCCACGCACGAGAAGGTGCTGGCCATAGTGAACGCGCTGGCCGAGAACAGGGCCATCC
GCCCCGACGAGGCCGGGCTGGTGTACGACGCGCTGCTGCAGCGCGTGGCCCGCTACAACAGCG
5 GCAACGTGCAGACCAACCTGGACCGGCTGGTGGGGGACGTGCGCGAGGCGGTGGCGCAGCGCG
AGCGCGCGGATCGGCAGGGCAACCTGGGCTCCATGGTGGCGCTGAATGCCTTCCTGAGCACGCA
GCCGGCCAACGTGCCGCGGGGGCAGGAAGACTACACCAACTTTGTGAGCGCGCTGCGGCTGATG
GTGACCGAGACCCCCAGAGCGAGGTGTACCAGTCGGGCCCCGACTACTTCTTCCAGACCAGCA
GACAGGGCCTGCAGACGGTGAACCTGAGCCAGGCTTTCAAGAACCTGCGGGGGCTGTGGGGCG
10 TGAAGGCGCCACCGGCGACCGGGCGACGGTGTCCAGCCTGCTGACGCCCAACTCGCGCCTGCT
GCTGCTGCTGATCGCGCCGTTACGGACAGCGGCAGCGTGTCCCGGGACACCTACCTGGGGCAC
CTGCTGACCCTGTACCGCGAGGCCATCGGGCAGGCGCAGGTGGACGAGCACACCTTCCAGGAGA
TCACCAGCGTGAGCCGCGCGCTGGGGCAGGAGGACACGAGCAGCCTGGAGGCGACTCTGAACT
ACCTGCTGACCAACCGGCGGCAGAAGATTCCCTCGCTGCACAGCCTGACCTCCGAGGAGGAGCG
15 CATCTTGCGCTACGTGCAGCAGAGCGTGAGCCTGAACCTGATGCGCGACGGGGTGACGCCCAGC
GTGGCGCTGGACATGACCGCGCGCAACATGGAACCGGGCATGTACGCCGCGCACCGGCCCTTACA
TCAACCGCCTGATGGACTACCTGCATCGCGCGGCGGCCGTGAACCCCGAGTACTTTACCAACGC
CATCCTGAACCCGCACTGGCTCCCGCCCGCCGGGTTCTACAGCGGGGGCTTCGAGGTCCCGGAG
ACCAACGATGGCTTCCTGTGGGACGACATGGACGACAGCGTGTTCCTCCCGCGGCCGCGAGGCGC
20 TGGCGGAAGCGTCCCTGCTGCGTCCCAAGAAGGAGGAGGAGGAGGAGGCGAGTCGCCGCCGCG
GCAGCAGCGCGTGGCTTCTGTCCGAGCTGGGGGCGGCAGCCGCCGCGCGCCCCGGGTCCCT
GGGCGGCAGCCCCCTTCCGAGCCTGGTGGGGTCTCTGCACAGCGAGCGCACCAACCCGCCCTCGG
CTGCTGGGCGAGGACGAGTACCTGAATAACTCCCTGCTGCAGCCGGTGCGGGAGAAAAACCTGC
CTCCCGCCTTCCCCAACAAACGGGATAGAGAGCCTGGTGGACAAGATGAGCAGATGGAAGACCT
25 ATGCGCAGGAGCACAGGGACGCGCCTGCGCTCCGGCCGCCACGCGGCGCCAGCGCCACGACC
GGCAGCGGGGGCTGGTGTGGGATGACGAGGACTCCGCGGACGATAGCAGCGTGTGGACCTGG
GAGGGAGCGGCAACCCGTTTCGCGCACCTGCGCCCCCGCCTGGGGAGGATGTTTTAAAAA
AAAAAAGCAAGAAGCATGATGCAAAAATTAAATAAACTACCAAGGCCATGGCGACCGAG
CGTTGGTTTCTTGTGTTCCCTTCAGTATGCGGCGCGCGCGCATGTACCAGGAGGGACCTCCTCCC
30 TCTTACGAGAGCGTGGTGGGCGCGGCGGCGGCGGCCCTCTTCTCCCTTTGCGTCGCAGCTGCT
GGAGCCGCCGTACGTGCCTCCGCGCTACCTGCGGCCTACGGGGGGGAGAAACAGCATCCGTTAC
TCGGAGCTGGCGCCCCCTGTTTCGACACCACCCGGGTGTACCTGGTGGACAACAAGTCGGCGGACG
TGGCCTCCCTGAACTACCAGAACGACCACAGCAATTTTTTGACCACGGTCATCCAGAACAATGA
CTACAGCCCGAGCGAGGCCAGCACCCAGACCATCAATCTGGATGACCGGTCGCACTGGGGCGGC
35 GACCTGAAAACCATCCTGCACACCAACATGCCAACGTGAACGAGTTCATGTTACCAATAAGT
TCAAGGCGCGGGTGTATGGTGTGCGCGCTCGCACACCAAGGAAGACCGGTGGAGCTGAAGTACG
AGTGGGTGGAGTTCGAGCTGCCAGAGGGCAACTACTCCGAGACCATGACCATTGACCTGATGAA
CAACGCGATCGTGGAGCACTATCTGAAAGTGGGCAGGCAGAACGGGGTCTGGAGAGCGACAT
CGGGGTCAAGTTCGACACCAGGAACCTCCGCTGGGGCTGGACCCCGTGACCGGGCTGGTTATG
40 CCCGGGGTGTACACCAACGAGGCCTTCCATCCCGACATCATCCTGCTGCCCGGCTGCGGGGTGG
ACTTCACTTACAGCCGCCTGAGCAACCTCCTGGGCATCCGCAAGCGGCAGCCCTTCCAGGAGGG

CTTCAGGATCACCTACGAGGACCTGGAGGGGGGCAACATCCCCGCGCTCCTCGATGTGGAGGCC
TACCAGGATAGCTTGAAGGAAAATGAGGCGGGACAGGAGGATACCGCCCCCGCCGCTCCGCC
GCCGCCGAGCAGGGCGAGGATGCTGCTGACACCGCGGCCGCGGACGGGGCAGAGGCCGACCCC
GCTATGGTGGTGGAGGCTCCCGAGCAGGAGGAGGACATGAATGACAGTGCGGTGCGCGGAGAC
5 ACCTTCGTACCCGGGGGGAGGAAAAGCAAGCGGAGGCCGAGGCCGCGGCCGAGGAAAAGCA
ACTGGCGGCAGCAGCGGCGGCGGCGGCGTTGGCCGCGGCGGAGGCTGAGTCTGAGGGGACCAA
GCCCCCAAGGAGCCCGTGATTAAGCCCCTGACCGAAGATAGCAAGAAGCGCAGTTACAACCT
GCTCAAGGACAGCACCAACACCGCGTACCGCAGCTGGTACCTGGCCTACAACCTACGGCGACCCG
TCGACGGGGGTGCGCTCCTGGACCTGCTGTGCACGCCGACGTGACCTGCGGCTCGGAGCAGG
10 TGTACTGGTCGCTGCCCGACATGATGCAAGACCCCGTGACCTTCCGCTCCACGCGGCAGGTCAG
CAACTTCCCGGTGGTGGGCGCCGAGCTGCTGCCCCTGCACTCCAAGAGCTTCTACAACGACCAG
GCCGTCTACTCCAGCTCATCCGCCAGTTCACCTCTCTGACCCACGTGTTCAATCGCTTTCTCTGAG
AACCAGATTCTGGCGCGCCCGCCCCCCCCACCATCACCAACCGTCAGTGAAAACGTTTCTGCTCT
CACAGATCACGGGACGCTACCGCTGCGCAACAGCATCGGAGGAGTCCAGCGAGTGACCGTTACT
15 GACGCCAGACGCCGCACCTGCCCCCTACGTTTACAAGGCCTTGGGCATAGTCTCGCCGCGCGTCT
TTCCAGCCGCACTTTTTGAGCAACACCACCATCATGTCCATCCTGATCTCACCCAGCAATAACTC
CGGCTGGGGACTGCTGCGCGCGCCAGCAAGATGTTTCGGAGGGGCGAGGAAGCGTTCCGAGCA
GCACCCCGTGCGCGTGCGCGGGCACTTCCGCGCCCCCTGGGGAGCGCACAAACGCGGCCGCGCG
GGGCGCACCAACCGTGGACGACGCCATCGACTCGGTGGTGGAGCAGGCGCGCAACTACAGGCCC
20 GCGGTCTCTACCGTGACGCGGCCATCCAGACCGTGGTGCGGGGCGCGCGGCGGTACGCCAAGC
TGAAGAGCCGCCGGAAGCGCGTGGCCCCGCCGCCACCGCCGCCGACCCGGGGCCGCCGCCAAAC
GCGCCGCCGCGGCCCTGCTTCGCCGGGCCAAGCGCACGGGGCCGCCGCGCGGCCATGAGGGCCGC
GCGCCGCTTGCCCGCCGGCATCACCGCCGCCACCATGGCCCCCGTACCCGAAGACGCGCGGCC
GCCGCCGCCGCCGCCGCCATCAGTGACATGGCCAGCAGGCGCCGGGGCAACGTGTACTGGGTGC
25 GCGACTCGGTGACCGGCACGCGCGTGCCCGTGCGCTTCCGCCCCCGCGGACTTGAGATGATGT
GAAAAAACAACACTGAGTCTCCTGCTGTTGTGTGTATCCAGCGGCGGCGCGCGCAGCGTC
ATGTCCAAGCGCAAAATCAAAGAAGAGATGCTCCAGGTGCTCGCGCCGGAGATCTATGGGCCCC
CGAAGAAGGAAGAGCAGGATTCAAGCCCCGCAAGATAAAGCGGGTCAAAAAGAAAAAGAAA
GATGATGACGATGCCGATGGGGAGGTGGAGTTCCTGCGCGCCACGGCGCCAGGCGCCCGGTGC
30 AGTGGAAGGGCCGCGCGTAAAGCGCGTCCTGCGCCCCGGCACCGCGGTGGTCTTACGCCCCG
CGAGCGCTCCACCCGACTTTCAAGCGCGTCTATGACGAGGTGTACGGCGACGAAGACCTGCTG
GAGCAGGCCAACGAGCGCTTCGAGAGTTTGCTTACGGGAAGCGTCAGCGGGCGCTGGGGAAAG
GAGGACCTGCTGGCGCTGCCGCTGGACCAGGGCAACCCCAACCCCACTCTGAAGCCCGTGACCC
TGCAGCAGGTGCTGCCGAGCAGCGCACCTCCGAGGCGAAGCGGGGTCTGAAGCGCGAGGGCG
35 GCGACCTGGCGCCCACCGTGACGCTCATGGTGCCCAAGCGGCAGAGGCTGGAGGATGTGCTGGA
GAAAATGAAAGTAGACCCCGGTCTGCAGCCGGACATCAGGGTCCGCCCCATCAAGCAGGTGGC
GCCGGGCCCTCGGCGTGACACCGTGGACGTGGTCATCCCCACCGGCAACTCCCCCGCCGCCGCC
ACCACTACCGCTGCCTCCACGGACATGGAGACACAGACCGATCCCCCGCAGCCGCAGCCGCAG
CCGCGCGCCGACCTCCTCGGCGGAGGTGCAGACGGACCCCTGGCTGCCGCCGGCGATGTCAGC
40 TCCCCGCGCGCGTCGCGGGCGCAGGAAGTACGGCGCCGCCAACGCGTCTCTGCCCGAGTACGCC
TTGCATCTTCCATCGCGCCACCCCGGCTACCGAGGCTATACCTACCGCCCCGCAAGAGCCA

AGGGTTCCACCCGCCGTCCCCGCCGACGCGCCGCCGCCACCACCCGCCGCCGCCGCCGACG
CCAGCCCGCACTGGCTCCAGTCTCCGTGAGGAAAGTGCGCGCGACGGACACACCCTGGTGCTG
CCCAGGGCGCGCTACCACCCCAGCATCGTTTAAAAGCCTGTTGTGGTTCTTGCAGATATGGCCCT
CACTTGCCGCCTCCGTTTCCCGGTGCCGGGATACCGAGGAGGAAGATCGCGCCGCAGGAGGGGT
5 CTGGCCGGCCGCGGCCTGAGCGGAGGCAGCCGCCGCGCGCACCCGGCGGGCAGCGGCCACCAGC
CGACGCATGCGCGGCGGGGTGCTGCCCCTGTTAATCCCCCTGATCGCCGCGGCGATCGGCGCCG
TGCCCGGGATCGCCTCCGTGGCCTTGCAAGCGTCCCAGAGGCATTGACAGACTTGCAAATTGC
AAATATGGAAAAAAACCCCAATAAAAAAGTCTAGACTCTCACGCTCGCTTGGTCCTGTGACT
ATTTTGTAGAATGGAAGACATCAACTTTGCGTCGCTGGCCCCGCGTCACGGCTCGCGCCCGTTCC
10 TGGGACACTGGAACGATATCGGCACCAGCAACATGAGCGGTGGCGCCTTCAAGTTGGGGCTCTCT
GTGGAGCGGCATTAAAAGTATCGGGTCTGCCGTTAAAAATTACGGCTCCCGGGCCTGGAACAGC
AGCACGGGCCAGATGTTGAGAGACAAGTTGAAAGAGCAGAACTTCCAGCAGAAGGTGGTGGAG
GGCCTGGCCTCCGGCATCAACGGGGTGGTGGACCTGGCCAACCAGGCCGTGCAGAATAAGATCA
ACAGCAGACTGGACCCCCGCCGCCGGTGGAGGAGGTGCCGCCGGCGCTGGAGACGGTGTCCC
15 CCGATGGGCGTGGCGAGAAGCGCCCGCGCCCGATAGGGAAGAGACCACTCTGGTCACGCAGA
CCGATGAGCCGCCCCCGTATGAGGAGGCCCTGAAGCAAGGTCTGCCCACCACGCGGCCCATCGC
GCCCATGGCCACCGGGGTGGTGGGCCGCCACACCCCCGCCACGCTGGACTTGCCCTCCGCCCGCC
GATGTGCCGAGCAGCAGAGAAGGCGGCACAGCCGGGCCCGCCGCGACCGCCTCCCGTTCCCTCCG
CCGGTCCTCTGCGCCGCGCGGCCAGCGGCCCCCGCGGGGGGGTTCGCGAGGCACGGCAACTGGC
20 AGAGCACGCTGAACAGCATCGTGGGTCTGGGGGTGCGGTCCGTGAAGCGCCGCCGATGCTACTG
AATAGCTTAGCTAACGTGTTGTATGTGTGTATGCGCCCTATGTGCGCCGACAGGAGCTGCTGAG
TCGCCGCCGTTTCGCGCGCCACCACCACCGCCACTCCGCCCTCAAGATGGCGACCCCATCGAT
GATGCCGAGTGGTCGTACATGCACATCTCGGGCCAGGACGCCTCGGAGTACCTGAGCCCCGGG
CTGGTGCAGTTTCGCCGCGCCACCGAGAGCTACTTCAGCCTGAGTAACAAGTTTAGGAACCCCA
25 CGGTGGCGCCACGCACGATGTGACCACCGACCGGTCTCAGCGCCTGACGCTGCGGTTTCATTCC
CGTGGAACGCGAGGACACCGCGTACTCGTACAAGGCGCGTTTACCCTGGCCGTGGGCGACAAC
CGCGTGCTGGACATGGCCTCCACCTACTTTGACATCCGCGGGGTGCTGGACCGGGGTCCCACTTT
CAAGCCCTACTCTGGCACCGCCTACAACTCCCTGGCCCCCAAGGGCGCTCCCAACTCCTGCGAGT
GGGAGCAAGAGGAAACTCAGGCAGTTGAAGAAGCAGCAGAAGAGGAAGAAGAAGATGCTGAC
30 GGTCAAGCTGAGGAAGAGCAAGCAGCTACCAAAAAGACTCATGTATATGCTCAGGCTCCCCTTT
CTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAACGACGCTACAGCTACAGAACAAA
AACCTATTTATGCAGACCCTACATTCCAGCCCGAACCCCAATCGGGGAGTCCCACTGGAATGA
GGCAGATGCTACAGTCGCCGCGGTAGAGTGCTAAAGAAATCTACTCCCATGAAACCATGCTAT
GGTTCCTATGCAAGACCCACAAATGCTAATGGAGGTCAGGGTGTACTAACGGCAAATGCCAGG
35 GACAGCTAGAATCTCAGGTTGAAATGCAATTCTTTTCAACTTCTGAAAACGCCCGTAACGAGGC
TAACAACATTCAGCCCAAATTGGTGCTGTATAGTGAGGATGTGCACATGGAGACCCCGGATACG
CACCTTTCTTACAAGCCCGCAAAAAGCGATGACAATTCAAAAATCATGCTGGGTGAGCAGTCCA
TGCCCAACAGACCTAATTACATCGGCTTCAGAGACAACTTTATCGGCCTCATGTATTACAATAGC
ACTGGCAACATGGGAGTGCTTGACGGTCAGGCCTCTCAGTTGAATGCAGTGGTGGACTTGCAAG
40 ACAGAAACACAGAACTGTCTTACCAGCTCTTGCTTGATTCCATGGGTGACAGAACCAGATACTT
TTCCATGTGGAATCAGGCAGTGGACAGTTATGACCCAGATGTTAGAATTATTGAAAATCATGGA

ACTGAAGACGAGCTCCCCAACTATTGTTTCCCTCTGGGTGGCATAGGGGTAAGTACACTTACCA
GGCTGTTAAAACCAACAATGGCAATAACGGGGGCCAGGTGACTTGGACAAAAGATGAAACTTTT
GCAGATCGCAATGAAATAGGGGTGGGAAACAATTCGCTATGGAGATCAACCTCAGTGCCAACC
TGTGGAGAACTTCCTGTACTCCAACGTGGCGCTGTACCTACCAGACAAGCTTAAGTACAACCC
5 CTCCAATGTGGACATCTCTGACAACCCCAACACCTACGATTACATGAACAAGCGAGTGGTGGCC
CCGGGGCTGGTGGACTGCTACATCAACCTGGGCGCGCGCTGGTCGCTGGACTACATGGACAACG
TCAACCCCTTCAACCACCACCGCAATGCGGGCCTGCGCTACCGCTCCATGCTCCTGGGCAACGG
GCGCTACGTGCCCTTCCACATCCAGGTGCCCCAGAAGTTCTTTGCCATCAAGAACCTCCTCCTCC
TGCCGGGCTCCTACACCTACGAGTGGAACCTCAGGAAGGATGTCAACATGGTCCTCCAGAGCTC
10 TCTGGGTAAACGATCTCAGGGTGGACGGGGCCAGCATCAAGTTCGAGAGCATCTGCCTCTACGCC
ACCTTCTTCCCCATGGCCCACAACACGGCCTCCACGCTCGAGGCCATGCTCAGGAACGACACCA
ACGACCAGTCCTTCAATGACTACCTCTCCGCCGCCAACATGCTCTACCCCATACCCGCCAACGCC
ACCAACGTCCCCATCTCCATCCCCTCGCGCAACTGGGCGGCCCTTCCGCGGCTGGGCCTTCAACCG
CCTCAAGACCAAGGAGACCCCCCTCCCTGGGCTCGGGATTGACCCCTACTACACCTACTCGGGC
15 TCCATTCCCTACCTGGACGGCACCTTCTACCTCAACCACACTTTCAAGAAGGTCTCGGTACCTT
CGACTCCTCGGTACGTGGCCGGGCAACGACCGTCTGCTCACCCCCAACGAGTTCGAGATCAAG
CGCTCGGTGACGGGGAGGGCTACAACGTGGCCCAGTGCAACATGACCAAGGACTGGTTCCTGG
TCCAGATGCTGGCCAACTACAACATCGGCTACCAGGGCTTCTACATCCCAGAGAGCTACAAGGA
CAGGATGTACTCCTTCTCAGGAACTTCCAGCCCATGAGCCGGCAGGTGGTGGACCAGACCAAG
20 TACAAGGACTACCAGGAGGTGGGCATCATCCACCAGCACAACTCGGGCTTCGTGGGCTACC
TCGCCCCCACCATGCGCGAGGGACAGGCCTACCCGCCAACTTCCCCTATCCGCTCATAGGCAA
GACCGCGGTGACAGCATCACCCAGAAAAAGTTCTCTGCGACCGCACCCCTCTGGCGCATCCCC
TTCTCCAGCAACTTCATGTCCATGGGTGCGCTCTCGGACCTGGGCCAGAACTTGCTCTACGCCAA
CTCCGCCACGCCCTCGACATGACCTTCGAGGTGACCCCATGGACGAGCCACCCCTTCTCTATG
25 TTCTGTTTCAAGTCTTTGACGTGGTCCGGGTCCACCAGCCGCACCGCGGCGTCATCGAGACCGTG
TACCTGCGTACGCCCTTCTCGGCCGGCAACGCCACCACCTAAAGAAGCAAGCCGCAGTCATCGC
CGCCTGCATGCCGTGGGTTCACCGAGCAAGAGCTCAGGGCCATCGTCAGAGACCTGGGATGC
GGGCCCTATTTTTTGGGCACCTTCGACAAGCGCTTCCCTGGCTTTGTCTCCCCACACAAGCTGGC
CTGCGCCATCGTCAACACGGCCGGCCGCGAGACCGGGGGCGTGCCTGGCTGGCCTTCGCTGG
30 AACCCGCGCTCCAAAACATGCTTCCTCTTTGACCCCTTCGGCTTTTCGGACCAGCGGCTCAAGCA
AATCTACGAGTTCGAGTACGAGGGCTTGCTGCGTCGACGCGCCATCGCCTCCTCGCCGACCGCT
GCGTCACCCCTCGAAAAGTCCACCCAGACCGTGCAGGGGCCCCGACTCGGCCGCCTGCGGTCTCTT
CTGCTGCATGTTTCTGCACGCCTTTGTGCACTGGCCTCAGAGTCCCATGGACCGCAACCCACCA
TGAACCTTGCTGACGGGGGTGCCAACTCCATGCTCCAGAGCCCCCAGGTGAGCCACCCCTGCG
35 CCGCAACCAGGAGCAGCTCTACAGCTTCCTGGAGCGCCACTCGCCTTACTTCCGCCGCCACAGC
GCACAGATCAGGAGGGCCACCTCCTTCTGCCACTTGCAAGAGATGCAAGAAGGGTAATAACGAT
GTACACACTTTTTTCTCAATAAATGGCATCTTTTTATTTATACAAGCTCTCTGGGGTATTCAATTT
CCCACCACCACCCGCCGTTGTCGCCATCTGGCTCTATTTAGAAATCGAAAGGGTCTGCCGGGAG
TCGCCGTGCGCCACGGGCAGGGACACGTTGCGATACTGGTAGCGGGTGCCCCACTTGAACCTCGG
40 GCACCACCAGGCGAGGCAGCTCGGGGAAGTTTTGCTCCACAGGCTGCGGGTCAGCACCAGCGC
GTTTCATCAGGTGCGGCGCCGAGATCTTGAAGTCGAGTTGGGGCCGCCGCCCTGCGCGCGCGAG

TTGCGGTACACCGGGTTGCAGCACTGGAACACCAACAGCGCCGGGTGCTTCACGCTGGCCAGCA
CGCTGCGGTTCGGAGATCAGCTCGGCGTCCAGGTCTCCGCGTTGCTCAGCGCGAACGGGGTCAT
CTTGGGCACTTGCCGCCCCAGGAAGGGCGCGTGCCCCGGTTTCGAGTTGCAGTCGCAGCGCAGC
GGGATCAGCAGGTGCCCCGTGCCCCGACTCGGCGTTGGGGTACAGCGCGCGCATGAAGGCCTGCA
5 TCTGGCGGAAGGCCATCTGGGCCCTTGCGGCCCTCCGAGAAGAACATGCCGCAGGACTTGCCCCGA
GAACTGGTTTTGCGGGGCAGCTGGCGTCGTGCAGGCAGCAGCGCGCGTCGGTGTGGCGATCTGC
ACCACGTTGCGCCCCCACCAGTTCTTCACGATCTTGGCCTTGGACGATTGCTCCTTCAGCGCGCG
CTGCCCCGTTCTCGCTGGTACATCCATCTCGATCATATGTTCCCTTGTTACCATGCTGCTGCCGTG
CAGACACTTCAGCTCGCCCTCCGTCTCGGTGCAGCGGTGCTGCCACAGCGCGCAGCCCCGTGGGC
10 TCGAAAGACTTGTAGGTCACCTCCGCGAAGGACTGCAGGTACCCCTGAAAAAGCGGCCCATCA
TGGTCACGAAGGTCTTGTGCTGCTGAAGGTCAGCTGCAGCCCGCGGTGCTCCTCGTTCAGCCAG
GTCTTGACACGCGCCGCCAGCGCCTCCACCTGGTCGGGCAGCATCTTGAAGTTCACCTTCAGCTC
ATTCTCCACGTGGTACTTGTCCATCAGCGTGCGCGCCGCTCCATGCCCTTCTCCCAGGCCGACA
CCAGCGGCAGGCTCACGGGGTTCTTACCATACCGTGCCGCGCGCTCCGCCGCGCTTTCGCTT
15 TCCGCCCCGCTGTTCTCTTCTCTCTCTCTCTCTCTCTCGCCGCCGCCACTCGCAGCCCCCGCACC
ACGGGGTCGTCTTCTGTCAGGCGCTGCACCTTGCGCTTGCCGTTGCGCCCCCTGCTTGATGCGCAC
GGGCGGGTTGCTGAAGCCCACCATCACCAGCGCGGCCTTCTTGTGCTCGTCCTCGCTGTCCAGAA
TGACCTCCGGGGAGGGGGGGTTGGTCATCTCAGTACCGAGGCACGCTTCTTTTTCTTCTGCGG
GCGTTGCCAGCTCCGCGGCTGCGGCCGCTGCCGAGGTCGAAGGCCGAGGGCTGGGCGTGCGCG
20 GCACCAGCGCGTCTGCGAGCCGTCTCGTCTCTCTCGGACTCGAGACGGAGGCGGGCCCCGCTT
CTTCGGGGGCGCGCGGGGCGGCGGAGGCGGCGGCGGCGACGGAGACGGGGACGAGACATCGTC
CAGGGTGGGTGGACGGCGGGCCGCGCCGCTCCGCGCTCGGGGGTGGTCTCGCGCTGGTCTCT
TCCCGACTGGCCATCTCCCACTGCTCCTTCTCTCTATAGGCAGAAAGAGATCATGGAGTCTCTCAT
GCGAGTCGAGAAGGAGGAGGACAGCCTAACC GCCCTCTGAGCCCTCCACCACCGCCGCCACC
25 ACCGCCAATGCCGCCGCGGACGACGCGCCACCGAGACCACCGCCAGTACCACCTCCCCAGCG
ACGCACCCCCGCTCGAGAATGAAGTGCTGATCGAGCAGGACCCGGGTTTTGTGAGCGGAGAGG
AGGATGAGGTGGATGAGAAGGAGAAGGAGGAGGTCGCCGCTCAGTGCCAAAAGAGGATAAA
AAGCAAGACCAGGACGACGCAGATAAGGATGAGACAGCAGTCGGGCGGGGGAACGGAAGCCA
TGATGCTGATGACGGCTACCTAGACGTGGGAGACGACGTGCTGCTTAAGCACCTGCACCGCCAG
30 TGCGTCATCGTCTGCGACGCGCTGCAGGAGCGCTGCGAAGTGCCCCTGGACGTGGCGGAGGTCA
GCCGCGCTACGAGCGGCACCTCTTCGCGCCGCACGTGCCCCCAAGCGCCGGGAGAACGGCAC
CTGCGAGCCCCAACCCGCGTCTCAACTTCTACCCGGTCTTCGCGGTACCCGAGGTGCTGGCCACCT
ACCACATCTTTTTCCAAAAGTGAAGATCCCCCTCTCTGCGCGCCAACCGCACCCGCGCCGAC
AAAACCTGACCCTGCGGCAGGGCGCCACATACCTGATATCGCCTCTCTGGAGGAAGTGCCCA
35 AGATCTTCGAGGGTCTCGGTGCGGACGAGAAACGGGCGGCGAACGCTCTGCACGGAGACAGCG
AAAACGAGAGTCACTCGGGGGTGTGGTGGAGCTCGAGGGCGACAACGCGCGCCTGGCCGTAC
TCAAGCGCAGCATAGAGGTACCCACTTTGCCTACCCGCGCTCAACCTGCCCCCAAGGTTCAT
GAGTGTGGTCATGGGCGAGCTCATCATGCGCCGCGCCAGCCCCTGGCCGCGGATGCAAACTTG
CAAGAGTCTCCGAGGAAGGCCTGCCCCGCGTCAGCGACGAGCAGCTGGCGCGCTGGCTGGAG
40 ACCCGCGACCCCGCGCAGCTGGAGGAGCGGCGCAAGCTCATGATGGCCGCGGTGCTGGTCACCG
TGGAGCTCGAGTGTCTGCAGCGCTTCTTCGCGGACCCCGAGATGCAGCGCAAGCTCGAGGAGAC

CCTGCACTACACCTTCCGCCAGGGCTACGTGCGCCAGGCCTGCAAGATCTCCAACGTGGAGCTC
TGCAACCTGGTCTCCTACCTGGGCATCCTGCACGAGAACCGCCTCGGGCAGAACGTCCTGCACT
CCACCCTCAAAGGGGAGGCGCGCCGCGACTACATCCGCGACTGCGCCTACCTCTTCCTCTGCTAC
ACCTGGCAGACGGCCATGGGGGTCTGGCAGCAGTGCCTGGAGGAGCGCAACCTCAAGGAGCTG
5 GAAAAGCTCCTCAAGCGCACCTCAGGGACCTCTGGACGGGCTTCAACGAGCGCTCGGTGGCCG
CCGCGCTGGCGGACATCATCTTTCCCGAGCGCCTGCTCAAGACCCTGCAGCAGGGCCTGCCCGA
CTTCACCAGCCAGAGCATGCTGCAGAACTTCAGGACTTTCATCCTGGAGCGCTCGGGCATCCTGC
CGGCCACTTGCTGCGCGCTGCCAGCGACTTCGTGCCATCAAGTACAGGGAGTGCCCCGCCGCC
GCTCTGGGGCCACTGCTACCTCTTCAGCTGGCCAACTACCTCGCTACCACTCGGACCTCATGG
10 AAGACGTGAGCGGCGAGGGCCTGCTCGAGTGCCACTGCCGCTGCAACCTCTGCACGCCCCACCG
CTCTCTAGTCTGCAACCCGCGAGCTGCTCAGCGAGAGTCAGATTATCGGTACCTTCGAGCTGCAGG
GTCCCTCGCTGACGAGAAGTCCGCGGCTCCAGGGCTGAAACTCACTCCGGGGCTGTGGACTTC
CGCCTACCTACGCAAATTTGTACCTGAGGACTACCACGCCCCAGAGATCAGGTTCTACGAAGAC
CAATCCCGCCCCGCCAAGGCGGAGCTCACCGCCTGCGTCATCACCCAGGGGCACATCCTGGGCC
15 AATTGCAAGCCATCAACAAAGCCCGCCGAGAGTTCTTGCTGAAAAAGGGTCGGGGGGTGTACCT
GGACCCCCAGTCCGGCGAGGAGCTAAACCCGCTACCCCCGCCGCCGCCAGCAGCGGGACCTT
GCTTCCCAGGATGGCACCCAGAAAGAAGCAGCAGCCGCCGCCGCCGCCAGCCATACATGCTT
CTGGAGGAAGAGGAGGAGGACTGGGACAGTCAGGCAGAGGAGGTTTCGGACGAGGAGCAGGA
GGAGATGATGGAAGACTGGGAGGAGGACAGCAGCCTAGACGAGGAAGCTTCAGAGGCCGAAG
20 AGGTGGCAGACGCAACACCATCGCCCTCGGTGCGAGCCCCCTCGCCGGGGCCCCCTGAAATCCTC
CGAACCCAGCACCAGCGCTATAACCTCCGCTCCTCCGGCGCCGGCGCCACCCGCCCGCAGACCC
AACCGTAGATGGGACACCACAGGAACCGGGGTGCGTAAGTCCAAGTGCCCCGCCGCCGCCACCG
CAGCAGCAGCAGCAGCAGCGCCAGGGCTACCGCTCGTGGCGCGGGCACAAGAACGCCATAGTC
GCCTGCTTGCAAGACTGCGGGGGCAACATCTCTTTCCGCCGCCGCTTCCTGCTATTCCACCACGG
25 GGTGCGCTTTCCCCGCAATGTCCTGCATTACTACCGTCATCTCTACAGCCCCCTACTGCAGCGGCG
ACCCAGAGGCGGCAGCGGCAGCCACAGCGGCGACCAACCTAGGAAGATATCCTCCGCGGGC
AAGACAGCGGCAGCAGCGGCCAGGAGACCCGCGGCAGCAGCGCGGGAGCGGTGGGCGCACT
GCGCCTCTCGCCCAACGAACCCCTCTCGACCCGGGAGCTCAGACACAGGATCTTCCCCACTTTGT
ATGCCATCTTCCAACAGAGCAGAGGCCAGGAGCAGGAGCTGAAAATAAAAAACAGATCTCTGC
30 GCTCCCTCACCCGCGAGCTGTCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAGGA
CGCGGAGGCACTCTTCAGCAAATACTGCGCGCTCACTCTTAAAGACTAGCTCCGCGCCCTTCTCG
AATTTAGGCGGGAGAAAACTACGTCACTCGCCGGCCGCCGCCAGCCCCGCCAGCCGAGATGAGC
AAAGAGATTCCCACGCCATACATGTGGAGCTACCAGCCGAGATGGGACTCGCGGCGGGAGCG
GCCAGGACTACTCCACCCGCATGAACTACATGAGCGCGGGACCCACATGATCTCACAGGTCA
35 ACGGGATCCGCGCCCAGCGAAACCAAATACTGCTGGAACAGGCGGCCATCACCGCCACGCCCC
GCCATAATCTCAACCCCCGAAATTGGCCCCGCCGCCCTCGTGTACCAGGAAACCCCTCCGCCAC
CACCGTACTACTTCCGCGTGACGCCCAGGCCGAAGTCCAGATGACTAACTCAGGGGCGCAGCTC
GCGGGCGGCTTTCGTACGGGGCGCGGCCGCTCCGACCAGGTATAAGACACCTGATGATCAGAG
GCCGAGGTATCCAGCTCAACGACGAGTCCGTGAGCTCTTCGCTCGGTCTCCGTCCGGACGGAAC
40 TTTCCAGCTCGCCGGATCCGGCCGCTCTTCGTTACGCCCCGCCAGGCGTACCTGACTCTGCAGA
CCTCGTCTCGGAGCCCCGCTCCGGCGGCATCGGAACCCTCCAGTTCGTGGAGGAGTTCGTGCCC

TCGGTCTACTTCAACCCCTTCTCGGACCTTCCCGAGCTACCCCGACGATTCATTCGGAACCT
TGACGCGGTGAAGGACTCGGCGGACGGCTACGACTGAATGTCAGGTGTCGAGGCAGAGCAGCT
TCGCCTGAGACACCTCGAGCACTGCCGCCGCCACAAGTGCTTCGCCCCGCGTTCTGGTGAGTTCT
GCTACTTTCAGCTACCCGAGGAGCATACCGAGGGGCCGGCGCACGGCGTCCGCTGACCACCCA
5 GGGCGAGGTTACCTGTTCCCTCATCCGGGAGTTTACCCTCCGTCCCCTGCTAGTGGAGCGGGAGC
GGGGTCCCTGTGTCCTAACTATCGCCTGCAACTGCCCTAACCTGGATTACATCAAGATCTTTGC
TGTCATCTCTGTGCTGAGTTTAATAAACGCTGAGATCAGAATCTACTGGGGCTCCTGTGCGCCATC
CTGTGAACGCCACCGTCTTCACCCACCCCGACCAGGCCAGGCGAACCTCACCTGCGGTCTGCA
TCGGAGGGCCAAGAAGTACCTCACCTGGTACTTCAACGGCACCCCTTTGTGGTTTACAACAGCT
10 TCGACGGGGACGGAGTCTCCCTGAAAGACCAGCTCTCCGGTCTCAGCTACTCCATCCACAAGAA
CACCACCCTCCAACCTTCCCTCCCTACCTGCCGGGAACCTACGAGTGCGTCACCGGCCGCTGCA
CCCACCTCACCCGCTGATCGTAAACCAGAGCTTTCGGGAACAGATAACTCCCTCTTCCCCAGA
ACAGGAGGTGAGCTCAGGAACTCCCCGGGGACCAGGGCGGAGACGTACCTTCGACCCTTGTG
GGGTTAGGATTTTTTATTACCGGGTTGCTGGCTCTTTAATCAAAGTTTCCTTGAGATTTGTTCTT
15 TCCTTCTACGTGTATGAACACCTCAACCTCCAATAACTCTACCCTTTCTTCGGAATCAGGTGACTT
CTCTGAAATCGGGCTTGGTGTGCTGCTTACTCTGTTGATTTTTTTCCTTATCATACTCAGCCTTCT
GTGCCTCAGGCTCGCCGCTGCTGCGCACACATCTATATCTACTGCTGGTTGCTCAAGTGCAGGG
GTCGCCACCCAAGATGAACAGGTACATGGTCCTATCGATCCTAGGCCTGCTGGCCCTGGCGGCC
TGCAGCGCCGCCAAAAAAGAGATTACCTTTGAGGAGCCCGCTTGCAATGTAACCTTTCAAGCCCG
20 AGGGTGACCAATGCACCACCTCGTCAAATGCGTTACCAATCATGAGAGGCTGCGCATCGACTA
CAAAAACAAAACCTGGCCAGTTTGCGGTCTATAGTGTGTTTACGCCCGGAGACCCCTCTAACTACT
CTGTCACCGTCTTCCAGGGCGGACAGTCTAAGATATTCAATTACACTTTCCTTTTTATGAGTTAT
GCGATGCGGTCATGTACATGTCAAAACAGTACAACCTGTGGCCTCCCTCTCCCCAGGCGTGTGTG
GAAAATACTGGGTCTTACTGCTGTATGGCTTTCGCAATCACTACGCTCGCTCTAATCTGCACGGT
25 GCTATACATAAAAATTCAGGCAGAGGCGAATCTTTATCGATGAAAAGAAAATGCCCTTGATCGCTA
ACACCGGCTTTCTATCTGCAGAATGAATGCAATCACCTCCCTACTAATCACCACCACCTCCTTG
CGATTGCCCATGGGTTGACACGAATCGAAGTGCCAGTGGGGTCCAATGTCACCATGGTGGGGCC
CGCCGGCAATTCACCCCTCATGTGGGAAAAATTTGTCCGCAATCAATGGGTTCATTTCTGCTCTA
ACCGAATCAGTATCAAGCCAGAGCCATCTGCGATGGGCAAAATCTAACTCTGATCAATGTGCA
30 AATGATGGATGCTGGGTACTATTACGGGCAGCGGGGAGAAATCATTAACTTGCGGACCCAC
AAGGACTACATGCTGCATGTAGTCGAGGCACTTCCCACTACCACCCCACTACCACCTCTCCAC
CACCACCACCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT
CCCGCAAAAGCACCATGATTAGCACAAAGCCCCCTCGTGCTCACTCCACGCGCGGGGCCCAT
CGGTGCGACCTCAGAAACCACCGAGCTTTGCTTCTGCCAATGCACTAACGCCAGCGCTCATGAA
35 CTGTGTGACCTGGAGAATGAGGATGTCCAGCAGCTCCGCTTGCCCTGACCCAGGAGGCTGTGG
AGCCCGTTGCCCTGAAGCAGATCGGTGATTCAATAATTGACTCTTCTTCTTTTGCCACTCCCGAA
TACCCTCCCGATTCTACTTTCCACATCACGGGTACCAAAGACCCCTAACCTCTCTTTCTACCTGATG
CTGCTGCTCTGTATCTCTGTGGTCTCTTCCGCGCTGATGTTACTGGGGATGTTCTGCTGCCTGATC
TGCCGCAGAAAGAGAAAAGCTCGCTCTCAGGGCCAACCACTGATGCCCTTCCCTACCCCCCGG
40 ATTTTGAGATAACAAGATATGAGCTCGCTGCTGACACTAACCGCTTTACTAGCCTGCGCTCTAA
CCCTTGTCGCTTGCGACTCGAGATTCCACAATGTACAGCTGTGGCAGGAGAAAATGTTACTTTC

AACTCCACGGCCGATACCCAGTGGTCGTGGAGTGGCTCAGGTAGCTACTTAACTATCTGCAATA
GCTCCACTTCCCCCGGCATATCCCCAACCAAGTACCAATGCAATGCCAGCCTGTTACCCCTCATC
AACGCTTCCACCCTGGACAATGGACTCTATGTAGGCTATGTACCCTTTGGTGGGCAAGGAAAGA
CCCACGCTTACAACCTGGAAGTTCGCCAGCCCAGAACCACTACCCAAGCTTCTCCACCACCAC
5 CACCACCACCACCATCACCAGCAGCAGCAGCAGCAGCAGCCACAGCAGCAGCAGCAGATTATT
GACTTTGGTTTTGGCCAGCTCATCTGCCGCTACCCAGGCCATCTACAGCTCTGTGCCCCGAAACCA
CTCAGATCCACCGCCCAGAAACGACCACCGCCACCACCCTACACACCTCCAGCGATCAGATGCC
GACCAACATCACCCCTTGGCTCTTCAAATGGGACTTACAAGCCCCACTCCAAAACCAGTGGAT
GCGGCCGAGGTCTCCGCCCTCGTCAATGACTGGGCGGGGCTGGGAATGTGGTGGTTCGCCATAG
10 GCATGATGGCGCTCTGCCTGCTTCTGCTCTGGCTCATCTGCTGCCTCCACCGCAGGCGAGCCAGA
CCCCCATCTATAGACCCATCATTGTCCTGAACCCCGATAATGATGGGATCCATAGATTGGATGG
CCTGAAAAACCTACTTTTTTCTTTACAGTATGATAAATTGAGACATGCCTCGCATTTTCTTGATC
ATGTTCTTCTCCACCTTTTCTGGGGTGTCTACGCTGGCCGCTGTGTCTCACCTGGAGGTAGAC
TGCCTCTCACCTTCACTGTCTACCTGCTTTACGGATTGGTCAACCTCACTCTCATCTGCAGCCTA
15 ATCACAGTAATCATCGCCTTCATCCAGTGCATTGATTACATCTGTGTGCGCCTCGCATACTTCAG
ACACCACCCGCAGTACCGAGACAGGAACATTGCCCAACTTCTAAGACTGCTCTAATCATGCATA
AGACTGTGATCTGCCTTCTGATCCTCTGCATCCTGCCACCCTCACCTCCTGCCAGTACACCACA
AAATCTCCGCGCAAAAGACATGCCTCCTGCCGCTTACCCAACCTGTGGAATATACCCAAATGCT
ACAACGAAAAGAGCGAGCTCTCCGAAGCTTGGCTGTATGGGGTCATCTGTGTCTTAGTTTTCTGC
20 AGCACTGTCTTTGCCCTCATAATCTACCCCTACTTTGATTTGGGATGGAACGCGATCGATGCCAT
GAATTACCCACCTTTCCCGCACCCGAGATAATTCCACTGCGACAAGTTGTACCCGTTGTGCTTA
ATCAACGCCCCCATCCCTACGCCCACTGAAATCAGCTACTTTAACCTAACAGGCGGAGATGA
CTGACGCCCTAGATCTAGAAATGGACGGCATCAGTACCGAGCAGCGTCTCCTAGAGAGGCGCAG
GCAGGCGGCTGAGCAAGAGCGCCTCAATCAGGAGCTCCGAGATCTCGTTAACCTGCACCAGTGC
25 AAAAGAGGCATCTTTTGTCTGGTAAAGCAGGCCAAAGTCACCTACGAGAAGACCGGCAACAGC
CACCGCCTCAGTTACAAATTGCCACCCAGCGCCAGAAGCTGGTGTCTATGGTGGGTGAGAATC
CCATCACCGTCACCCAGCACTCGGTAGAGACCGAGGGGTGTCTGCACTCCCCCTGTGCGGGTCC
AGAAGACCTCTGCACCCTGGTAAAGACCCTGTGCGGTCTCAGAGATTTAGTCCCCTTTAACTAAT
CAAACACTGGAATCAATAAAAAGAATCACTTACTTAAAATCAGACAGCAGGTCTCTGTCCAGTT
30 TATTCAGCAGCACCTCCTTCCCCCTCCCAACTCTGGTACTCCAAACGCCTTCTGGCGGCAAAAC
TTCCTCCACACCCTGAAGGGAATGTCAGATTCTTGCTCCTGTCCCTCCGCACCCACTATCTTCATG
TTGTTGCAGATGAAGCGCACAAAACGTCTGACGAGAGCTTCAACCCCGTGTACCCCTATGACA
CGGAAAGCGGCCCTCCCTCCGTCCCTTTCTCACCCCTCCCTTCGTGTCTCCCGATGGATTCCAA
GAAAGTCCCCCGGGTCTGTCTCTGAACCTGGCCGAGCCCCCTGGTCACTTCCCACGGCATGCT
35 CGCCCTGAAAATGGGAAGTGGCCTCTCCCTGGACGACGCTGGCAACCTCACCTCTCAAGATATC
ACCACCGCTAGCCCTCCCCCTCAAAAAAACCAAGACCAACCTCAGCCTAGAAACCTCATCCCCC
TAACTGTGAGCACCTCAGGCGCCCTACCGTAGCAGCCGCGCTCCCCTGGCGGTGGCCGGCAC
CTCCCTCACCATGCAATCAGAGGCCCCCTGACAGTACAGGATGCAAAACTCACCTGGCCACC
AAAGGCCCCCTGACCGTGTCTGAAGGCAAACCTGGCCTTGCAAACATCGGCCCCGCTGACGGCCG
40 CTGACAGCAGCACCTCACAGTCAGTGCCACACCACCCCTTAGCACAAGCAATGGCAGCTTGGG
TATTGACATGCAAGCCCCCATTTACACCACCAATGGAAAACCTAGGACTTAACTTTGGCGCTCCCC

TGCATGTGGTAGACAGCCTAAATGCACTGACTGTAGTTACTGGCCAAGGTCTTACGATAAACGG
AACAGCCCTACAACTAGAGTCTCAGGTGCCCTCAACTATGACACATCAGGAAACCTAGAATTG
AGAGCTGCAGGGGGTATGCGAGTTGATGCAAATGGTCAACTTATCCTTGATGTAGCTTACCCATT
TGATGCACAAAACAATCTCAGCCTTAGGCTTGACAGGGACCCCTGTTTGTTAACTCTGCCCACA
5 ACTTGATGTTAACTACAACAGAGGCCTCTACCTGTTACATCTGGAAATACCAAAAAGCTAGA
AGTTAATATCAAAACAGCCAAGGGTCTCATTTATGATGACACTGCTATAGCAATCAATGCGGGT
GATGGGCTACAGTTTGACTCAGGCTCAGATACAAATCCATTAAAACTAACTTGGATTAGGAC
TGGATTATGACTCCAGCAGAGCCATAATTGCTAACTGGGAACTGGCCTAAGCTTTGACAACAC
AGGTGCCATCACAGTAGGCAACAAAAATGATGACAAGCTTACCTTGTGGACCACACCAGACCCA
10 TCCCCTAACTGTAGAATCTATTAGAGAAAGATGCTAAATTCACACTTGTTTTGACTAAATGCGG
CAGTCAGGTGTTGGCCAGCGTTTCTGTTTTATCTGTAAAAGGTAGCCTTGCGCCCATCAGTGGCA
CAGTAAGTGTGCTCAGATTGTCCTCAGATTGATGAAAATGGAGTTCTACTAAGCAATTCTTCC
CTTGACCCTCAATACTGGAACACAGAAAAGGTGACCTTACAGAGGGCACTGCATATACCAACG
CAGTGGGATTTATGCCCAACCTCACAGCATACCCAAAAACACAGAGCCAACTGCTAAAAGCAA
15 CATTGTAAGTCAGGTTTACTTGAATGGGGACAAATCCAAACCCATGACCCTCACCATTACCCTCA
ATGGAACATAATGAAACAGGAGATGCCACAGTAAGCACTTACTCCATGTCATTCTCATGGAACG
GAATGGAAGTAATTACATTAATGAAACGTTCCAAACCAACTCCTTCACCTTCTCCTACATCGCCC
AAGAATAAAAAGCATGACGCTGTTGATTGATTCAATGTGTTTCTGTTTTATTTCAAGCACAAC
AAAATCATTCAAGTCATTCTTCCATCTTAGCTTAATAGACACAGTAGCTTAATAGACCCAGTAGT
20 GCAAAGCCCCATTCTAGCTTATACTAGTGGAGAAGTACTCGCCTACATGGGGGTAGAGTCATA
ATCGTGCATCAGGATAGGGCGGTGGTGTGTCAGCAGCGCGGAATAAACTGCTGCCGCCGCCG
TCCGTCCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGCGATGATTTCGACCGCCCGCAGCA
TAAGGCGCCTTGTCCTCCGGGCACAGCAGCGCACCCCTGATCTCACTTAAATCAGCACAGTAACT
GCAGCACAGCACCACAATATTGTTCAAAATCCACAGTGCAAGGCGCTGTATCCAAAGCTCATG
25 GCGGGGACCACAGAACCACGTGGCCATCATACCACAAGCGCAGGTAGATTAAGTGGCGACCC
CTCATAAACACGCTGGACATAAACATTACCTCTTTTGGCATGTTGTAATTCACCACCTCCCGGTA
CCATATAAACCTCTGATTAAACATGGCGCCATCCACCACCATCCTAAACCAGCTGGCCAAAACC
TGCCCGCCGGCTATACACTGCAGGGAACCGGGACTGGAACAATGACAGTGGAGAGCCCAGGAC
TCGTAACCATGGATCATCATGCTCGTCATGATATCAATGTTGGCACAACACAGGCACACGTGCA
30 TACACTTCCTCAGGATTACAAGCTCCTCCCGCGTTAGAACCATATCCCAGGGAACAACCCATTCC
TGAATCAGCGTAAATCCCACACTGCAGGGAAGACCTCGCACGTAACTCACGTTGTGCATTGTCA
AAGTGTTACATTGCGGCAGCAGCGGATGATCCTCCAGTATGGTAGCGCGGGTTTCTGTCTCAA
AGGAGGTAGACGATCCCTACTGTACGGAGTGCGCCGAGACAACCGAGATCGTGTTGGTCGTAGT
GTCATGCCAAATGGAACGCCGGACGTAGTCATATTTCTGAAGTCTTAGATCTCTCAACGCAGC
35 ACCAGCACCAACACTTCGCAGTGTAAGGCCAAGTGCCGAGAGAGTATATATAGGAATAAAA
AGTGACGTAAACGGGCAAAGTCCAAAAAACGCCAGAAAAACCGCACGCGAACCTACGCCCCG
AAACGAAAGCCAAAAAACAAGTACACTCCCTTCCGGCGTCAACTTCCGCTTTCCACGCTACG
TCACTTGCCCCAGTCAAACAACTACATATCCCGAACTTCCAAGTCGCCACGCCCCAAACACCG
CCTACACCTCCCCGCCCGCCGGCCCCGCCCCAAACCCGCTCCCGCCCCGCGCCCCGCGCG
40 CGCCCATCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTATTGATGATGGTTTAAAC
GGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGAGTATTCTATAGTGTACCTAAATAGCT

TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAAC
ATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAA
TTGCGTTGCGCTCACTGCCCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATC
GGCCAACGCGAACCCCTTGCGGCCGCCGGCCGTCGACCAATTCTCATGTTTGACAGCTTATCA
5 TCGAATTTCTGCCATTTCATCCGCTTATTATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAGGG
CACCAATAACTGCCTTAAAAAAATTACGCCCGCCCTGCCACTCATCGCAGTACTGTTGTAATTC
ATTAAGCATTCTGCCGACATGGAAGCCATCACAAACGGCATGATGAACCTGAATCGCCAGCGGC
ATCAGCACCTTGTCGCCTTGCGTATAATATTTGCCCATGGTGAAAACGGGGGCGAAGAAGTTGT
CCATATTGGCCACGTTTAAATCAAACTGGTGAACTCACCCAGGGATTGGCTGAGACGAAAAA
10 CATATTCTCAATAAACCCCTTAGGGAAATAGGCCAGGTTTTACCGTAACACGCCACATCTTGCG
AATATATGTGTAGAACTGCCGAAATCGTCGTGGTATTCCTCCAGAGCGATGAAAACGTTTC
AGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACACTATCCCATATCACCAGCTCACCGTCT
TTCATTGCCATACGGAATTCGGGATGAGCATTTCATCAGGCGGGCAAGAATGTGAATAAAGGCCG
GATAAACTTGTTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCTGAACGGTC
15 TGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGA
TATATCAACGGTGTTATATCCAGTGATTTTTTCTCCATTTAGCTTCCTTAGCTCCTGAAAACT
CGATAACTCAAAAAATACGCCCGGTAGTGATCTTATTTTATTATGGTGAAAGTTGGAACCTCTTA
CGTGCCGATCAACGTCTCATTTTCGCCAAAAGTTGGCCAGGGCTTCCCGGTATCAACAGGGAC
ACCAGGATTTATTTATTCTGCGAAGTGATCTTCCGTCACAGGTATTTATTCGCGATAAGCTCATG
20 GAGCGGCGTAACCGTCGCACAGGAAGGACAGAGAAAGCGCGGATCTGGGAAGTGACGGACAG
AACGGTCAGGACCTGGATTGGGGAGGCGGTTGCCGCCGCTGCTGCTGACGGTGTGACGTTCTCT
GTTCCGGTCACACCACATACGTTCCGCCATTCCATGCGATGCACATGCTGTATGCCGGTATACC
GCTGAAAGTTCTGCAAAGCCTGATGGGACATAAGTCCATCAGTTCAACGGAAGTCTACACGAAG
GTTTTTGCGCTGGATGTGGCTGCCCGGCACCGGGTGCAGTTTGCGATGCCGGAGTCTGATGCGGT
25 TGCGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGCCCTTATATGGAAATGTGGAACTG
AGTGATATGCTGTTTTGTCTGTAAACAGAGAAGCTGGCTGTTATCCACTGAGAAGCGAACG
AAACAGTCGGGAAAATCTCCATTATCGTAGAGATCCGCATTATTAATCTCAGGAGCCTGTGTA
GCGTTTATAGGAAGTAGTGTTCTGTCTATGATGCCTGCAAGCGGTAACGAAAACGATTTGAATAT
GCCTTCAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTGGAGCGGATTATGTCAGCA
30 ATGGACAGAACAACCTAATGAACACAGAACCATGATGTGGTCTGTCCTTTTACAGCCAGTAGTG
CTCGCCGAGTCGAGCGACAGGGCGAAGCCCTCGAGTGAGCGAGGAAGCACCAGGGAACAGCA
CTTATATATTCTGCTTACACACGATGCCTGAAAAAACTTCCCTTGGGGTTATCCACTTATCCACG
GGGATATTTTTATAATTATTTTTTTTATAGTTTTTAGATCTTCTTTTTTAGAGCGCCTTGAGGCCT
TTATCCATGCTGGTTCTAGAGAAGGTGTTGTGACAAATTGCCCTTTCAGTGTGACAAATCACCTT
35 CAAATGACAGTCCTGTCTGTGACAAATTGCCCTTAACCCTGTGACAAATTGCCCTCAGAAGAAG
CTGTTTTTTCACAAAGTTATCCCTGCTTATTGACTCTTTTTTATTTAGTGTGACAATCTAAAACT
TGTCACACTTCACATGGATCTGTCTATGGCGGAAACAGCGGTTATCAATCACAAGAAACGTAAAA
ATAGCCCGCAATCGTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCA
AAAACGTATGCTGTATCTGTTGCTTACCAGATCAGAAAAATCTGATGGCACCTACAGGAACAT
40 GACGGTATCTGCGAGATCCATGTTGCTAAATATGCTGAAATATTCGGATTGACCTCTGCGGAAG
CCAGTAAGGATATACGGCAGGCATTGAAGAGTTTCGCGGGGAAGGAAGTGGTTTTTTATCGCCC

TGAAGAGGATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCTTGGTTTATCAAACGTGCGCAC
AGTCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCTCATTCCCTTCTTTATCGGGTT
ACAGAACCGGTTTACGCAGTTTCGGCTTAGTGAAACAAAAGAAATCACCAATCCGTATGCCATG
CGTTTATACGAATCCCTGTGTACAGTATCGTAAGCCGGATGGCTCAGGCATCGTCTCTCTGAAAAT
5 CGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGCCTGACTTCCGCCGCC
GCTTCCTGCAGGTCTGTGTTAATGAGATCAACAGCAGAACTCCAATGCGCCTCTCATACATTGAG
AAAAAGAAAGGCCGCCAGACGACTCATATCGTATTTTCCTTCCGCGATATCACTTCCATGACGA
CAGGATAGTCTGAGGGTTATCTGTACAGATTTGAGGGTGGTTCGTACATTTGTTCTGACCTAC
TGAGGGTAATTTGTACAGTTTGTCTGTTTCCTTCAGCCTGCATGGATTTTCTCATACTTTTTGAA
10 CTGTAATTTTTAAGGAAGCCAAATTTGAGGGCAGTTTGTACAGTTGATTTCCCTTCTTTCCCTT
CGTCATGTGACCTGATATCGGGGGTTAGTTCGTATCATTGATGAGGGTTGATTATCACAGTTTA
TTACTCTGAATTGGCTATCCGCGTGTGTACCTCTACCTGGAGTTTTTCCCACGGTGGATATTTCTT
CTTGCGCTGAGCGTAAGAGCTATCTGACAGAACAGTTCTTCTTTGCTTCCTCGCCAGTTCGCTCG
CTATGCTCGGTTACACGGCTGCGGCGAGCGCTAGTGATAATAAGTGAAGTGAAGTATGTGCTCTTC
15 TTATCTCCTTTTGTAGTGTTGCTCTTATTTTAAACAACCTTTCGGGTTTTTGTGACTTTGCGATTT
TGTTGTTGCTTTGCAGTAAATTGCAAGATTTAATAAAAAAACGCAAAGCAATGATTAAAGGATG
TTCAGAATGAACTCATGGAAACACTTAACCAGTGCATAAACGCTGGTCATGAAATGACGAAGG
CTATCGCCATTGCACAGTTTAATGATGACAGCCCGGAAGCGAGGAAAATAACCCGGCGCTGGAG
AATAGGTGAAGCAGCGGATTTAGTTGGGGTTTCTTCTCAGGCTATCAGAGATGCCGAGAAAGCA
20 GGGCGACTACCGCACCCGGATATGGAAATTCGAGGACGGGTTGAGCAACGTGTTGGTTATACAA
TTGAACAAATTAATCATATGCGTGATGTGTTTGGTACGCGATTGCGACGTGCTGAAGACGATTTT
CCACCGGTGATCGGGGTTGCTGCCCATAAAGGTGGCGTTTACAAAACCTCAGTTTCTGTTTCATCT
TGCTCAGGATCTGGCTCTGAAGGGGCTACGTGTTTTGCTCGTGGAAGGTAACGACCCCCAGGGA
ACAGCCTCAATGTATCACGGATGGGTACCAGATCTTCATATTCATGCAGAAGACACTCTCCTGCC
25 TTTCTATCTTGGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGCTGGCCGGGGCTTG
ACATTATTCCTTCTGTCTGGCTCTGCACCGTATTGAAACTGAGTTAATGGGCAAATTTGATGAA
GGTAAACTGCCCACCGATCCACACCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCATGACTA
TGATGTCATAGTTATTGACAGCGCGCTAACCTGGGTATCGGCACGATTAATGTGCTATGTGCTG
CTGATGTGCTGATTGTTCCACGCCTGCTGAGTTGTTTACTACACCTCCGCACTGCAGTTTTTCG
30 ATATGCTTCGTGATCTGCTCAAGAACGTTGATCTTAAAGGGTTCGAGCCTGATGTACGTATTTTG
CTTACCAAATACAGCAATAGTAATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAATTCGGGATG
CCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGGATGAAGTTGGTAAAGGTCAGAT
CCGGATGAGAACTGTTTTTGAACAGGCCATTGATCAACGCTCTTCAACTGGTGCCTGGAGAAAT
GCTCTTTCTATTTGGGAACCTGTCTGCAATGAAATTTTCGATCGTCTGATTAAACCACGCTGGGA
35 GATTAGATAATGAAGCGTGCGCCTGTTATTCCAAAACATACGCTCAATACTCAACCGGTTGAAG
ATACTTCGTTATCGACACCAGCTGCCCCGATGGTGGATTGTTAATTGCGCGCGTAGGAGTAATG
GCTCGCGGTAATGCCATTACTTTGCCTGTATGTGGTCGGGATGTGAAGTTTACTCTTGAAGTGCT
CCGGGGTGATAGTGTTGAGAAGACCTCTCGGGTATGGTCAGGTAATGAACGTGACCAGGAGCTG
CTTACTGAGGACGCACTGGATGATCTCATCCCTTCTTTTCTACTGACTGGTCAACAGACACCGGC
40 GTTCGGTGAAGAGTATCTGGTGTACATAGAAATTGCCGATGGGAGTCGCCGTCGTAAAGCTGCT
GCACTTACCGAAAGTGATTATCGTGTTCTGGTTGGCGAGCTGGATGATGAGCAGATGGCTGCAT

TATCCAGATTGGGTAACGATTATCGCCCAACAAGTGCTTATGAACGTGGTCAGCGTTATGCAAG
 CCGATTGCAGAAATGAATTTGCTGGAAATATTTCTGCGCTGGCTGATGCGGAAAATATTTACGTA
 AGATTATTACCCGCTGTATCAACACCGCCAAATTGCCTAAATCAGTTGTTGCTCTTTTTCTCACC
 CCGGTGAACTATCTGCCCCGTGAGGTGATGCACTTCAAAAAGCCTTTACAGATAAAGAGGAATT
 5 ACTTAAGCAGCAGGCATCTAACCTTCATGAGCAGAAAAAAGCTGGGGTGATATTTGAAGCTGAA
 GAAGTTATCACTCTTTTAACTTCTGTGCTTAAAACGTCATCTGCATCAAGAACTAGTTTAAAGCTC
 ACGACATCAGTTTGCTCCTGGAGCGACAGTATTGTATAAGGGCGATAAAAATGGTGCTTAACCTG
 GACAGGTCTCGTGTTCCTCAACTGAGTGTATAGAGAAAATTGAGGCCATTCTTAAAGAACTTGAAA
 AGCCAGCACCTGATGCGACCACGTTTATGCTACGTTTATCTGTCTTTACTTAATGTCCTTTGTT
 10 ACAGGCCAGAAAGCATAACTGGCCTGAATATTCTCTCTGGGCCACTGTTCCACTTGTATCGTCG
 GTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACC
 ACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGG
 TCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCA
 TGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGT
 15 CTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCAC
 GGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGATCCCACTCGTGTGTGCGGTC
 TGATTATCGGTCTGGGACCACGGTCCCACTTGTATTGTCGATCAGACTATCAGCGTGAGACTACG
 ATTCCATCAATGCCTGTCAAGGGCAAGTATTGACATGTCGTCGTAACCTGTAGAACGGAGTAAC
 CTCGGTGTGCGGTTGTATGCCTGCTGTGGATTGCTGCTGTGTCCTGCTTATCCACAACATTTTGC
 20 CACGGTTATGTGGACAAAATACCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATC
 CGATGCAAGTGTGTCGCTGTGACGAGCTCGCGAGCTCGGACATGAGGTTGCCCGTATTCAGT
 GTCGCTGATTTGTATTGTCTGAAGTTGTTTTACGTTAAGTTGATGCAGATCAATTAATACGATA
 CCTGCGTCATAATTGATTATTTGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATG
 ATAATCATTATCACTTTACGGGTCTTTCCGGTGATCCGACAGGTTACGGGGCGGCGACCTCGCG
 25 GGTTCCTGCTATTTATGAAAATTTTCCGGTTTAAAGGCGTTTCCGTTCTTCTTCGTCATAACTTAAT
 GTTTTTATTTAAAATACCTCTGAAAAGAAAGGAAACGACAGGTGCTGAAAGCGAGCTTTTTGG
 CCTCTGTCGTTTCCTTTCTCTGTTTTGTCCGTGGAATGAACAATGGAAGTCCGAGCTCATCGCTA
 ATAACTTCGTATAGCATACATTATACGAAGTTATATTCGATGCGGCCGCAAGGGGTTTCGCGTCA
 GCGGGTGTGGCGGGTGTGCGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAG
 30 TGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCC
 ATTCGCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGC
 CAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAAGTTGGGTAACGCCAGGGTTTCCAGT
 CACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTCGAGCT
 CGGTACCCGGGGATCCTCGTTTAAAC

35

SEQ ID NO: 9 - Polynucleotide sequence encoding ChAd155#1375 backbone construct

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGATGGGCGGCGGG
 GCGGGGCGCGGGGCGGAGGCGGGTTTGGGGCGGGCCGGCGGGCGGGGCGGTGTGGCGGAA
 GTGGACTTTGTAAGTGTGGCGGATGTGACTTGCTAGTGCCGGGCGCGGTAAAAGTGACGTTTT
 40 CGTGCGGACAACGCCCCCGGAAGTGACATTTTCCCGCGTTTTTACCGGATGTTGTAGTGAA
 TTTGGGCGTAACCAAGTAAGATTTGGCCATTTTCGCGGGAAAACGAAACGGGGAAGTGAAATC

TGATTAATTTTTCGTTAGTCATACCGCGTAATATTTGTCTAGGGCCGAGGGACTTTGGCCGATTA
CGTGGAGGACTCGCCAGGTGTTTTTTGAGGTGAATTTCCGCGTTCCGGGTCAAAGTCTGCGTTT
TATTATTATAGGATATCCCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCT
CATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACG
5 GGGTCATTAGTTCATAGCCCATATATGGAGTTCGCGGTTACATAACTTACGGTAAATGGCCCCGCC
TGGCTGACCGCCCAACGACCCCCGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCC
TGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
10 CATCGCTATTACCATGGTGTATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
CACGGGGATTTCOAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAA
CGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCATGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAG
AGATCGTCGACGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTG
15 ACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGC
GGATTCCCCGTGCCAAGAGTGAGATCTTCCGTTTATCTAGGTACCGGGCCCCCCTCGAGGTCTGA
CGGTATCGATAAGCTTCACGCTGCCGCAAGCACTCAGGGCGCAAGGGCTGCTAAAGGAAGCGG
AACACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGCTA
TCTGGACAAGGGAAAACGCAAGCGCAAAGAGAAAAGCAGGTAGCTTGCAGTGGGCTTACATGGC
20 GATAGCTAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTC
TGTAAGGTTGGGAAGCCCTGCAAAGTAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGG
CGCAGGGGATCAAGATCTAACCAGGAGCTATTTAATGGCAACAGTTAACCAGCTGGTACGCAAA
CCACGTGCTCGCAAAGTTGCGAAAAGCAACGTGCCTGCGCTGGAAGCATGCCCCGAAAAACGTG
GCGTATGTACTCGTGTATATACTACCACTCTAAAAAACCGAACTCCGCGCTGCGTAAAGTATGC
25 CGTGTTCTGCTGACTAACGGTTTCGAAGTGACTTCCTACATCGGTGGTGAAGGTCACAACCTGCA
GGAGCACTCCGTGATCCTGATCCGTGGCGGTGCTGTTAAAGACCTCCCGGGTGTTCGTTACCACA
CCGTACGTGGTTCGCTTACTGCTCCGGCGTTAAAGACCGTAAGCAGGCTCGTTCCAAGTATGG
CGTGAAGCGTCCTAAGGCTTAATGGTAGATCTGATCAAGAGACAGGATGACGGTCGTTTCGCAT
GCTTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGCTAT
30 GACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGC
GCCCCGTTCTTTTTGTCAAGACCGACCTGTCCGGTGGCCTGAATGAACTGCAGGACGAGGCAGC
GCGGCTATCGTGGCTGGCCACGACGGGCGTTTCCTTGCGCAGCTGTGCTCGACGTTGTCAGTAA
GCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTATCTACCTTG
CTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGC
35 TACCTGCCCATTTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTAATCGGATGGAAGCC
GGTCTTGTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAAGTTCG
CCAGGCTCAAGGCGCGCATGCCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTT
GCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTG
GCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAAT
40 GGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTTCGACGCGCATCGCCTTCTAT
CGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCC

CAACCTGCCATCACGAGATTTGATTCCACCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCG
TTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCAC
CCCGGGCTCGATCCCCTCGGGGGGAATCAGAATTCAGTCGACAGCGCCGCGATCTGCTGTGCC
TTCTAGTTGCCAGCCATCTGTTGTTTGGCCCTCCCCCGTGCCTTCCTTGACCCCTGGAAGGTGCCAC
5 TCCCACTGTCCTTTCCTAATAAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTA
TTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATG
CTGGGGATGCGGTGGGCTCTATGGCCGATCAGCGATCGCTGAGGTGGGTGAGTGGGCGTGGCCT
GGGGTGGTCATGAAAATATATAAGTTGGGGGTCTTAGGGTCTCTTTATTTGTGTTGCAGAGACCG
CCGGAGCCATGAGCGGGAGCAGCAGCAGCAGTAGCAGCAGCGCCTTGGATGGCAGCATCG
10 TGAGCCCTTATTTGACGACGCGGATGCCCCACTGGGCCGGGGTGCCTCAGAATGTGATGGGCTC
CAGCATCGACGGCCGACCCGTCTTGCCCGCAAATTCGCCACGCTGACCTATGCGACCGTCGCG
GGGACGCCGTTGGACGCCACCGCCGCCGCCGCCACCGCAGCCGCCCTCGGCCGTGCGCAGCC
TGGCCACGGACTTTGCATTCTTGGGACCACTGGCGACAGGGGCTACTTCTCGGGCCGCTGCTGCC
GCCGTTGCGGATGACAAGCTGACCGCCCTGCTGGCGCAGTTGGATGCGCTTACTCGGGAAGTGG
15 GTGACCTTTCTCAGCAGGTCATGGCCCTGCGCCAGCAGGTCTCCTCCCTGCAAGCTGGCGGGAAT
GCTTCTCCCAAAAATGCCGTTTAAGATAAAATAAAACCAGACTCTGTTTGGATTAAAGAAAAGTA
GCAAGTGCATTGCTCTCTTTATTTTCATAATTTTCCGCGCGCGATAGGCCCTAGACCAGCGTTCTC
GGTCGTTGAGGGTGCAGGTGTATCTTCTCCAGGACGTGGTAGAGGTGGCTCTGGACGTTGAGATA
CATGGGCATGAGCCCGTCCCGGGGGTGGAGGTAGCACCCTGCAGAGCTTCATGCTCCGGGGTG
20 GTGTTGTAGATGATCCAGTCGTAGCAGGAGCGCTGGGCATGGTGCCTAAAAATGTCCTCAGCA
GCAGGCCGATGGCCAGGGGGAGGCCCTTGGTGTAAGTGTTTACAAAACGGTTAAGTTGGGAAG
GGTGCATTGCGGGAGAGATGATGTGCATCTTGGACTGTATTTTGTAGATTGGCGATGTTTCCGCCC
AGATCCCTTCTGGGATTCATGTTGTGCAGGACCACCAGTACAGTGTATCCGGTGCCTTGGGGA
ATTTGTGATGCAGCTTAGAGGGAAAAGCGTGGAAGAACTTGGAGACGCCTTGTGGCCTCCCAG
25 ATTTTCCATGCATTGTCATGATGATGGCAATGGGCCCGCGGAGGCAGCTTGGGCAAAGATA
TTTCTGGGGTTCGCTGACGTCGTAGTTGTGTTCCAGGGTGAAGTGCATAGGCCATTTTTACAAA
GCGCGGGCGGAGGGTGGCCGACTGGGGGATGATGGTCCCCTCTGGCCCTGGGGCGTAGTTGCC
TCGCAGATCTGCATTTCCCAGGCCTTAATCTCGGAGGGGGGAATCATATCCACCTGCGGGGCGA
TGAAGAAAACGGTTTCCGGAGCCGGGGAGATTAACTGGGATGAGAGCAGGTTTCTAAGCAGCT
30 GTGATTTTCCACAACCGGTGGGCCCATAAATAACACCTATAACCGGTTGCAGCTGGTAGTTTAG
AGAGCTGCAGCTGCCGTCGTCCCGAGGAGGGGGGCCACCTCGTTGAGCATGTCCCTGACGCGC
ATGTTCTCCCCGACCAGATCCGCCAGAAGGCGCTCGCCGCCAGGGACAGCAGCTCTTGCAAGG
AAGCAAAGTTTTTCAGCGGCTTGAGGCCGTCCGCCGTGGGCATGTTTTTCAGGGTCTGGCTCAGC
AGCTCCAGGCGGTCCAGAGCTCGGTGACGTGCTCTACGGCATCTCTATCCAGCATATCTCCTCG
35 TTTCGCGGGTTGGGGCGACTTTCGCTGTAGGGCACCAAGCGGTGGTTCGTCCAGCGGGGCCAGAG
TCATGTCCTTCCATGGGCGCAGGGTCTCTCGTCAGGGTGGTCTGGGTACGGTGAAGGGGTGCGC
TCCGGGCTGAGCGCTTGCCAAGGTGCGCTTGAGGCTGGTTCTGCTGGTGCTGAAGCGCTGCCGG
TCTTCGCCCTGCGCGTCGGCCAGGTAGCATTGACCATGGTGTGTCATAGTCCAGCCCCCTCCGCGGC
GTGTCCCTTGGCGCGCAGCTTGCCCTTGGAGGTGGCGCCGCACGAGGGGCAGAGCAGGCTCTTG
40 AGCGCGTAGAGCTTGGGGGCGAGGAAGACCGATTGCGGGGAGTAGGCGTCCGCGCCGCAGACC
CCGCACACGGTCTCGCACTCCACCAGCCAGGTGAGCTCGGGGCGCGCCGGGTCAAAAACAGGT

TTCCCCCATGCTTTTTGATGCGTTTCTTACCTCGGGTCTCCATGAGGTGGTGTCCCCGCTCGGTGA
CGAAGAGGGCTGTCCGTGTCTCCGTAGACCGACTTGAGGGGTCTTTTCTCCAGGGGGGTCCCTCGG
TCTTCCTCGTAGAGGAACTCGGACCACTCTGAGACGAAGCCCCGCTCCAGGCCAGGACGAAGG
AGGCTATGTGGGAGGGGTAGCGGTCGTTGTCCACTAGGGGGTCCACCTTCTCCAAGGTGTGAAG
5 ACACATGTGCGCTTCCCTCGGCGTCCAGGAAGGTGATTGGCTTGTAGGTGTAGGCCACGTGACCG
GGGGTTCCTGACGGGGGGGTATAAAAGGGGGTGGGGGCGCGCTCGTCTGCTACTCTCTTCCGCAT
CGCTGTCTGCGAGGGCCAGCTGCTGGGGTGAGTATTCCCTCTCGAAGGCGGGCATGACCTCCGC
GCTGAGGTTGTGAGTTTCCAAAACGAGGAGGATTTGATGTTACCTGTCCCGAGGTGATACCTT
TGAGGGTACCCGCGTCCATCTGGTCAGAAAACACGATCTTTTATTGTCCAGCTTGGTGCGAAC
10 GACCCGTAGAGGGCGTTGGAGAGCAGCTTGGCGATGGAGCGCAGGGTCTGGTTCTTGTCCCTGT
CGGCGCGCTCCTTGGCCGCGATGTTGAGCTGCACGTA CTGCGCGCGACGCAGCGCCACTCGGG
GAAGACGGTGGTGCCTCGTCGGGCACAGGCGCACGCGCCAGCCGCGTTGTGCAGGGTGAC
CAGGTCCACGCTGGTGGCGACCTCGCCGCGCAGGCGCTCGTTGGTCCAGCAGAGACGGCCGCC
TTGCGCGAGCAGAAGGGGGGCAGGGGGTTCGAGCTGGGTCTCGTCCGGGGGGTCCGCGTCCACG
15 GTGAAAACCCCGGGGCGCAGGCGCGCTCGAAGTAGTCTATCTTGCAACCTTGCATGTCCAGCG
CCTGCTGCCAGTCGCGGGCGGCGAGCGCGCTCGTAGGGGTTGAGCGGCGGGCCCCAGGGCAT
GGGGTGGGTGAGTGCGGAGGCGTACATGCCGCAGATGTCATAGACGTAGAGGGGCTCCCGCAG
GACCCCGATGTAGGTGGGGTAGCAGCGGCCCGCGGATGCTGGCGCGCACGTAGTCATACAGC
TCGTGCGAGGGGGCGAGGAGGTCGGGGCCCAGGTTGGTGCGGGCGGGGCGCTCCGCGCGGAAG
20 ACGATCTGCCTGAAGATGGCATGCGAGTTGGAAGAGATGGTGGGGCGCTGGAAGACGTTGAAG
CTGGCGTCTGCAGGCCGACGGCGTCGCGCACGAAGGAGGCGTAGGAGTCGCGCAGCTTGTGTA
CCAGCTCGGCGGTGACCTGCACGTCGAGCGCGCAGTAGTCGAGGGTCTCGCGGATGATGTCATA
TTAGCCTGCCCCCTTCTTTTCCACAGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTCCAGTA
CTCTTGATCGGGAACCGTCCGGTTCCGAACGGTAAGAGCCTAGCATGTAGA ACTGGTTGACG
25 GCCTGGTAGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGTAGGCCTGCGCGGCCCTTGGCGAGCG
AGGTGTGGGTCAGGGCGAAGGTGTCCCTGACCATGACTTTGAGGTA CTGGTGTGTAAGTCGGA
GTCGTGCGAGCCGCCCCGCTCCCAGAGCGAGAAGTCGGTGCCTTCTTGGAGCGGGGGTTGGGC
AGAGCGAAGGTGACATCGTTGAAGAGGATTTTGGCCGCGCGGGGCATGAAGTTGCGGGTGATGC
GGAAGGGCCCCGGCACTTCAGAGCGGTTGTTGATGACCTGGGCGGCGAGCACGATCTCGTCGAA
30 GCCGTTGATGTTGTGGCCACGATGTAGAGTTCCAGGAAGCGGGGCCGCCCCCTTACGGTGGGC
AGCTTCTTTAGCTCTTCGTAGGTGAGCTCCTCGGGCGAGGCGAGGCCGTGCTCGGCCAGGGCCC
AGTCCGCGAGGTGCGGGTTGTCTCTGAGGAAGGACTTCCAGAGGTGCGGGGCCAGGAGGGTCTG
CAGGCGGTCTCTGAAGGTCCTGAACTGGCGGCCACGCCCATT TTTTCGGGGGTGATGCAGTAG
AAGGTGAGGGGGTCTTGCTGCCAGCGGTCCCAGTCGAGCTGCAGGGCGAGGTGCGCGCGGCG
35 GTGACCAGGCGCTCGTCGCCCCGAATTTTCATGACCAGCATGAAGGGCACGAGCTGCTTCCGA
AGGCCCCCATCCAAGTGTAGGTCTCTACATCGTAGGTGACAAAGAGGCGCTCCGTGCGAGGATG
CGAGCCGATCGGGAAGAACTGGATCTCCCGCCACCAGTTGGAGGAGTGGCTGTTGATGTGGTGG
AAGTAGAAGTCCCGTCGCGGGGCCGAACACTCGTGCTGGCTTTTGTAAAAGCGAGCGCAGTACT
GGCAGCGCTGCACGGGTGTACCTCATGCACGAGATGCACCTTTCGCCCCGCGCACGAGGAAGCC
40 GAGGGGAAATCTGAGCCCCCGCCTGGCTCGCGGCATGGCTGGTTCTCTTCTACTTTGGATGCGT
GTCCGTCTCCGTCTGGCTCCTCGAGGGGTGTTACGGTGGAGCGGACCACCACGCCGCGGAGCC

GCAGGTCCAGATATCGGCGCGCGGCGGTTCGGAGTTTGATGACGACATCGCGCAGCTGGGAGCTG
TCCATGGTCTGGAGCTCCCGCGGCGGCGGCAGGTCAGCCGGGAGTTCTTGCAGGTTACCTCGC
AGAGTCGGGCCAGGGCGCGGGGCAGGTCTAGGTGGTACCTGATCTCTAGGGGCGTGTTGGTGGC
GGCGTCGATGGCTTGCAGGAGCCCGCAGCCCCGGGGGGCGACGACGGTGCCCCGCGGGGTGGT
5 GGTGGTGGTGGCGGTGCAGCTCAGAAGCGGTGCCGCGGGCGGGCCCCCGAGGTAGGGGGGGC
TCCGGTCCCGCGGGCAGGGGCGGCAGCGGCACGTTCGGCGTGAGCGCGGGCAGGAGTTGGTGC
TGTGCCCCGAGGTTGCTGGCGAAGGCGACGACGCGGCGGTTGATCTCCTGGATCTGGCGCCTCT
GCGTGAAGACGACGGCCCCGGTGAGCTTGAACCTGAAAGAGAGTTTCGACAGAATCAATCTCGG
TGTCATTGACCGCGGCCTGGCGCAGGATCTCCTGCACGTCTCCCGAGTTGTCTTGGTAGGCGATC
10 TCGGCCATGAACTGCTCGATCTCTTCTCCTGGAGGTCTCCGCGTCCGGCGCGTTCCACGGTGGC
CGCCAGGTCGTTGGAGATGCGCCCCATGAGCTGCGAGAAGGCGTTGAGTCCGCCCTCGTTCCAG
ACTCGGCTGTAGACCACGCCCCCTGGTCATCGCGGGCGCGCATGACCACCTGCGCGAGGTTGA
GCTCCACGTGCCGCGCAAGACGGCGTAGTTGCGCAGACGCTGGAAGAGGTAGTTGAGGGTGG
TGGCGGTGTGCTCGGCCACGAAGAAGTTCATGACCCAGCGGCGCAACGTGGATTCTGTTGATGTC
15 CCCC AAGGCCTCCAGCCGTTCCATGGCCTCGTAGAAGTCCACGGCGAAGTTGAAAACTGGGAG
TTGCGCGCCGACACGGTCAACTCCTCCTCCAGAAGACGGATGAGCTCGGCGACGGTGTGCGCGA
CCTCGCGCTCGAAGGCTATGGGGATCTCTTCTCCGCTAGCATCACACCTCCTCCTCTTCTCCT
CTTCTGGCACTTCCATGATGGCTTCTCCTCTTCGGGGGGTGGCGGCGGCGGCGGTGGGGGAGG
GGGCGCTCTGCGCCGGCGGCGGCGCACCGGGAGGCGGTCCACGAAGCGCGCGATCATCTCCCCG
20 CGGCGGCGGCGCATGGTCTCGGTGACGGCGCGGCCGTTCTCCCGGGGGCGCAGTTGGAAGACGC
CGCCGGACATCTGGTGTGGGGCGGGTGGCCGTGAGGCAGCGAGACGGCGCTGACGATGCATCT
CAACAATTGCTGCGTAGGTACGCCGCCGAGGGACCTGAGGGAGTCCATATCCACCGGATCCGAA
AACCTTTCGAGGAAGGCGTCTAACCAGTCGCAGTCGCAAGGTAGGCTGAGCACCGTGCGGGGCG
GCGGGGGGTGGGGGGAGTGTCTGGCGGAGGTGCTGCTGATGATGTAATTGAAGTAGGCGGACTT
25 GACACGGCGGATGGTTCGACAGGAGACCATGTCCTTGGGTCCGGCCTGCTGGATGCGGAGGCGG
TCGGCTATGCCCCAGGCTTCGTTCTGGCATCGGCGCAGGTCTTGATAGTAGTCTTGCATGAGCCT
TTCCACCGGCACCTCTTCTCCTTCTTCTGCTTCTTCCATGTCTGCTTCGGCCCTGGGGCGGCG
CCGCGCCCCCTGCCCCCATGCGCGTGACCCGAACCCCTGAGCGGTTGGAGCAGGGCCAGG
TCGGCGACGACGCGCTCGGCCAGGATGGCCTGCTGCACCTGCGTGAGGGTGGTTTGAAGTCAT
30 CCAAGTCCACGAAGCGGTGGTAGGCGCCCGTGTGATGGTGTAGGTGCAGTTGGCCATGACGGA
CCAGTTGACGGTCTGGTGGCCCGGTTGCGACATCTCGGTGTACCTGAGTCGCGAGTAGGCGCGG
GAGTCGAAGACGTAGTCGTTGCAAGTCCGCACCAAGTACTGGTAGCCACCAGGAAGTGCGGCG
GCGGCTGGCGGTAGAGGGGCCAGCGCAGGGTGGCGGGGGCTCCGGGGGCCAGGTCTTCCAGCA
TGAGGCGGTGGTAGGCGTAGATGTACCTGGACATCCAGGTGATACCCGCGGCGGTGGTGGAGGC
35 GCGCGGGAAGTCGCGCACCCGTTCCAGATGTTGCGCAGGGGCAGAAAGTGCTCCATGGTAGGC
GTGCTCTGTCCAGTCAGACGCGCGCAGTCGTTGATACTCTAGACCAGGGAACGAAAGCCGGT
CAGCGGGCACTCTTCCGTGGTCTGGTGAATAGATCGCAAGGGTATCATGGCGGAGGGCCTCGGT
TCGAGCCCCGGTCCGGGCCGGACGGTCCGCCATGATCCACGCGGTTACCGCCCGCGTGTCGAA
CCCAGGTGTGCGACGTACACAACGGTGGAGTGTTCTTTTGGCGTTTTTCTGGCCGGGCGCCGG
40 CGCCGCGTAAGAGACTAAGCCGCGAAAGCGAAAGCAGTAAGTGGCTCGCTCCCCGTAGCCGGA
GGGATCCTTGCTAAGGGTTGCGTTGCGGCGAACCCCGGTTTCAATCCCGTACTCGGGCCGGCCG

GACCCGCGGCTAAGGTGTTGGATTGGCCTCCCCCTCGTATAAAGACCCCGCTTGCGGATTGACTC
CGGACACGGGGACGAGCCCCCTTTATTTTGTCTTTCCCCAGATGCATCCGGTGCTGCGGCAGATG
CGCCCCCGCCCCAGCAGCAGCAACAACACCAGCAAGAGCGGCAGCAACAGCAGCGGGAGTCA
TGCAGGGCCCCCTCACCCACCTCGGCGGGCCGGCCACCTCGGCGTCCGCGGCCGTGTCTGGCG
5 CCTGCGGCGGCGGGGGGCCGGCTGACGACCCCGAGGAGCCCCCGGGCGCAGGGCCAGAC
ACTACCTGGACCTGGAGGAGGGCGAGGGCTGGCGCGGCTGGGGGCGCCGTCTCCCGAGCGCC
ACCCGCGGGTGACGCTGAAGCGCGACTCGCGCGAGGCGTACGTGCCTCGGCAGAACCTGTTTCA
GGACCGCGGGGCGAGGAGCCCGAGGAGATGCGGGACAGGAGGTTTACGCGCAGGGCGGGAGC
TGCGGCAGGGGCTGAACCGCGAGCGGCTGCTGCGCGAGGAGGACTTTGAGCCCGACGCGCGGA
10 CGGGGATCAGCCCCGCGCGCGCACGTGGCGGCCGCGACCTGGTGACGGCGTACGAGCAGA
CGGTGAACCAGGAGATCAACTTCCAAAAGAGTTTCAACAACCACGTGCGCACGCTGGTGGCGCG
CGAGGAGGTGACCATCGGGCTGATGCACCTGTGGGACTTTGTAAGCGCGCTGGTGCAGAACCCC
AACAGCAAGCCTCTGACGGCGCAGCTGTTTCTGATAGTGCAGCACAGCAGGACAACGAGGCG
TTTAGGGACGCGCTGCTGAACATCACCGAGCCCGAGGGTCGGTGGCTGCTGGACCTGATTAACA
15 TCCTGCAGAGCATAGTGGTGCAGGAGCGCAGCCTGAGCCTGGCCGACAAGGTGGCGGCCATCA
ACTACTCGATGCTGAGCCTGGGCAAGTTTTACGCGCGCAAGATCTACCAGACGCCGTACGTGCC
CATAGACAAGGAGGTGAAGATCGACGGTTTTTACATGCGCATGGCGCTGAAGGTGCTCACCTG
AGCGACGACCTGGGCGTGTACCGCAACGAGCGCATCCACAAGGCCGTGAGCGTGAGCCGGCGG
CGCGAGCTGAGCGACCGCGAGCTGATGCACAGCCTGCAGCGGGCGCTGGCGGGCGCCGGCAGC
20 GCGACAGGGAGGCGGAGTCTACTTCGATGCGGGGGCGGACCTGCGCTGGGCGCCAGCCGG
CGGGCCCTGGAGGCCGCGGGGGTCCGCGAGGACTATGACGAGGACGGCGAGGAGGATGAGGAG
TACGAGCTAGAGGAGGGCGAGTACCTGGACTAAACCGCGGGTGGTGTTCGCGTAGATGCAAG
ACCCGAACGTGGTGGACCCGCGCTGCGGGCGGCTCTGCAGAGCCAGCCGTCCGGCCTTAATC
CTCAGACGACTGGCGACAGGTCATGGACCGCATCATGTGCTGACGGCGGTAACCCGGACGCG
25 TTCCGGCAGCAGCCGAGGCCAACAGGCTCTCCGCCATCCTGGAGGCGGTGGTGCCTGCGCGCT
CGAACCCACGCACGAGAAGGTGCTGGCCATAGTGAACGCGCTGGCCGAGAACAGGGCCATCC
GCCCCGACGAGGCCGGGCTGGTGTACGACGCGCTGCTGCAGCGCGTGGCCCGCTACAACAGCG
GCAACGTGCAGACCAACCTGGACCGGCTGGTGGGGGACGTGCGCGAGGCGGTGGCGCAGCGCG
AGCGCGCGGATCGGCAGGGCAACCTGGGCTCCATGGTGGCGCTGAATGCCTTCTGAGCACGCA
30 GCCGGCCAACGTGCCGCGGGGGCAGGAAGACTACACCAACTTTGTGAGCGCGCTGCGGCTGATG
GTGACCGAGACCCCCAGAGCGAGGTGTACAGTCCGGGCCGACTACTTCTTCCAGACCAGCA
GACAGGGCCTGCAGACGGTGAACCTGAGCCAGGCTTTCAAGAACCTGCGGGGGCTGTGGGGCG
TGAAGGCGCCACCGGCGACCGGGCGACGGTGTCCAGCCTGCTGACGCCAACTCGCGCCTGCT
GCTGCTGCTGATCGCGCCGTTACGGACAGCGGCAGCGTGTCCCGGGACACCTACCTGGGGCAC
35 CTGCTGACCCTGTACCGCGAGGCCATCGGGCAGGCGCAGGTGGACGAGCACACCTTCCAGGAGA
TCACCAGCGTGAGCCGCGCGCTGGGGCAGGAGGACACGAGCAGCCTGGAGGCGACTCTGAAC
ACCTGCTGACCAACCGGCGGCAGAAGATTCCCTCGCTGCACAGCCTGACCTCCGAGGAGGAGCG
CATCTTGCCTACGTGCAGCAGAGCGTGAGCCTGAACCTGATGCGCGACGGGGTGACGCCAGC
GTGGCGCTGGACATGACCGCGCGCAACATGGAACCGGGCATGTACCGCGCGACCGGCCCTTACA
40 TCAACCGCCTGATGGACTACCTGCATCGCGCGGCGGCCGTGAACCCCGAGTACTTTACCAACGC
CATCCTGAACCCGCACTGGCTCCCGCCCGCGGGTTCTACAGCGGGGGCTTCGAGGTCCCGGAG

ACCAACGATGGCTTCCTGTGGGACGACATGGACGACAGCGTGTCTCCCCGCGGCCGCAGGCGC
TGGCGGAAGCGTCCCTGCTGCGTCCCAAGAAGGAGGAGGAGGAGGAGGCGAGTCGCCGCCGCG
GCAGCAGCGGCGTGGCTTCTGTCCGAGCTGGGGGCGGCAGCCGCCGCGCGCCCCGGGTCCCT
GGGCGGCAGCCCCCTTCCGAGCCTGGTGGGGTCTCTGCACAGCGAGCGCACCCGCCCTCGG
5 CTGCTGGGCGAGGACGAGTACCTGAATAACTCCCTGCTGCAGCCGGTGCGGGAGAAAAACCTGC
CTCCCGCCTTCCCCAACACGGGATAGAGAGCCTGGTGGACAAGATGAGCAGATGGAAGACCT
ATGCGCAGGAGCACAGGGACGCGCCTGCGCTCCGGCCGCCACGCGGCGCCAGCGCCACGACC
GGCAGCGGGGGCTGGTGTGGGATGACGAGGACTCCGCGGACGATAGCAGCGTGTCTGGACCTGG
GAGGGAGCGGCAACCCGTTGCGGCACCTGCGCCCCCGCCTGGGGAGGATGTTTTAAAAA
10 AAAAAAAGCAAGAAGCATGATGCAAAAATTAAATAAAACTCACCAAGGCCATGGCGACCGAG
CGTTGGTTTCTTGTGTTCCCTTCAGTATGCGGCGCGCGGCGATGTACCAGGAGGGACCTCCTCCC
TCTTACGAGAGCGTGGTGGGCGCGGCGGCGGCGGCCCTCTTCTCCCTTTGCGTCGCAGCTGCT
GGAGCCCGCTACGTGCCTCCGCGCTACCTGCGGCCTACGGGGGGGAGAAACAGCATCCGTTAC
TCGGAGCTGGCGCCCCGTGTTGACACACCACCCGGGTGTACCTGGTGGACAACAAGTCGGCGGACG
15 TGGCCTCCCTGAACTACCAGAACGACCACAGCAATTTTTTGACCACGGTCATCCAGAACAATGA
CTACAGCCCGAGCGAGGCCAGCACCCAGACCATCAATCTGGATGACCGGTCGCACTGGGGCGGG
GACCTGAAAACCATCCTGCACACCAACATGCCAACGTGAACGAGTTCATGTTACCAATAAGT
TCAAGGCGCGGGTGATGGTGTGCGGCTCGCACACCAAGGAAGACCGGGTGGAGCTGAAGTACG
AGTGGGTGGAGTTCGAGCTGCCAGAGGGCAACTACTCCGAGACCATGACCATTGACCTGATGAA
20 CAACGCGATCGTGGAGCACTATCTGAAAGTGGGCAGGCAGAACGGGGTCTGGAGAGCGACAT
CGGGGTCAAGTTCGACACCAGGAACCTCCGCTGGGGCTGGACCCCGTGACCGGGCTGGTTATG
CCCGGGGTGTACACCAACGAGGCCTTCCATCCCGACATCATCCTGCTGCCCCGCTGCGGGGTGG
ACTTCACTTACAGCCGCCTGAGCAACCTCCTGGGCATCCGCAAGCGGCAGCCCTTCCAGGAGGG
CTTACAGGATCACCTACGAGGACCTGGAGGGGGCAACATCCCCGCGCTCCTCGATGTGGAGGCC
25 TACCAGGATAGCTTGAAGGAAAATGAGGCGGGACAGGAGGATACCGCCCCCGCCGCTCCGCC
GCCGCCGAGCAGGGCGAGGATGCTGCTGACACCGCGGCCGCGGACGGGGCAGAGGCCGACCCC
GCTATGGTGGTGGAGGCTCCCGAGCAGGAGGAGGACATGAATGACAGTGCGGTGCGCGGAGAC
ACCTTCGTACCCGGGGGGAGGAAAAGCAAGCGGAGGCCGAGGCCGCGGCCGAGGAAAAGCA
ACTGGCGGCAGCAGCGGCGGCGGCGGCGTGGCCGCGGCGGAGGCTGAGTCTGAGGGGACCAA
30 GCCCCCAAGGAGCCCGTGATTAAGCCCCTGACCGAAGATAGCAAGAAGCGCAGTTACAACCT
GCTCAAGGACAGCACCAACACCGCGTACCGCAGCTGGTACCTGGCCTACAACTACGGCGACCCG
TCGACGGGGGTGCGCTCCTGGACCTGCTGTGCACGCCGACGTGACCTGCGGCTCGGAGCAGG
TGTA CTGGTCTGCTGCCCACATGATGCAAGACCCCGTGACCTTCCGCTCCACGCGGCAGGTCAG
CAACTTCCCGGTGGTGGGCGCCGAGCTGCTGCCCCTGCACTCCAAGAGCTTCTACAACGACCAG
35 GCCGTCTACTCCAGCTCATCCGCCAGTTCACCTCTCTGACCCACGTGTTCAATCGCTTTCCTGAG
AACCAGATTCTGGCGCGCCCGCCCGCCCCACCATCACCAACCGTCAGTGAAAACGTTTCTGCTCT
CACAGATCACGGGACGCTACCGCTGCGCAACAGCATCGGAGGAGTCCAGCGAGTGACCGTTACT
GACGCCAGACGCCGACCTGCCCCCTACGTTTACAAGGCCTTGGGCATAGTCTCGCCGCGCGTCT
TTCCAGCCGCACTTTTTGAGCAACACCACCATCATGTCCATCCTGATCTCACCCAGCAATAACTC
40 CGGCTGGGGACTGCTGCGCGCGCCAGCAAGATGTTCCGAGGGGCGAGGAAGCGTTCCGAGCA
GCACCCCGTGCGCGTGC GCGGGCACTTCCGCGCCCCCTGGGGAGCGCACAAACGCGGCCGCGCG

GGGCGCACCAACCGTGGACGACGCCATCGACTCGGTGGTGGAGCAGGCGCGCAACTACAGGCCC
GCGGTCTCTACCGTGGACGCGGCCATCCAGACCGTGGTGCGGGGCGCGCGGCGGTACGCCAAGC
TGAAGAGCCGCCGAAGCGCGTGGCCCGCCGCCACCGCCCGGACCCGGGGCCGCCGCCAAAC
GCGCCCGCCGCGGCCCTGCTTCGCCGGGCCAAGCGCACGGGCGCCGCGCCGCCATGAGGGCCGC
5 GCGCCGCTTGGCCGCCCGGCATCACCGCCGCCACCATGGCCCCCGTACCCGAAGACGCGCGGCC
GCCGCCCGCCGCGGCCCATCAGTGACATGGCCAGCAGGCGCCGGGGCAACGTGTACTGGGTGC
GCGACTCGGTGACCGGCACGCGCGTGGCCGTGCGCTTCCGCCCCCGCGGACTTGAGATGATGT
GAAAAAACAACACTGAGTCTCCTGCTGTTGTGTGTATCCCAGCGGCGGCGCGCGCAGCGTC
ATGTCCAAGCGCAAAATCAAAGAAGAGATGCTCCAGGTGCTCGCGCCGAGATCTATGGGCCCC
10 CGAAGAAGGAAGAGCAGGATTCAAGCCCCGCAAGATAAAGCGGGTCAAAAAGAAAAAGAAA
GATGATGACGATGCCGATGGGGAGGTGGAGTTCCTGCGCGCCACGGCGCCAGGCGCCCGGTGC
AGTGGAAGGGCCGGCGCGTAAAGCGCGTCTGCGCCCCGGCACCGCGGTGGTCTTACGCCCCG
CGAGCGCTCCACCCGACTTTCAAGCGCGTCTATGACGAGGTGTACGGCGACGAAGACCTGCTG
GAGCAGGCCAACGAGCGCTTCGGAGAGTTTGCTTACGGGAAGCGTCAGCGGGCGCTGGGGAAG
15 GAGGACCTGCTGGCGCTGCCGCTGGACCAGGGCAACCCCAACCCCACTCTGAAGCCCGTGACCC
TGCAGCAGGTGCTGCCGAGCAGCGCACCTCCGAGGCGAAGCGGGGTCTGAAGCGCGAGGGCG
GCGACCTGGCGCCACCGTGCAGCTCATGGTGCCCAAGCGGCAGAGGCTGGAGGATGTGCTGGA
GAAAATGAAAGTAGACCCCGGTCTGCAGCCGGACATCAGGGTCCGCCCCATCAAGCAGGTGGC
GCCGGGCCTCGGCGTGCAGACCGTGGACGTGGTCATCCCCACCGGCAACTCCCCCGCCGCCGCC
20 ACCACTACCGCTGCCTCCACGGACATGGAGACACAGACCGATCCCGCCGACCCGACGCCGAG
CCGCCGCCGCGACCTCCTCGGCGGAGGTGCAGACGGACCCCTGGCTGCCGCCGGCGATGTCAGC
TCCCCGCGCGCTCGCGGGCGCAGGAAGTACGGCGCCGCCAACCGCGTCTGCCCCGAGTACGCC
TTGCATCCTTCCATCGCGCCACCCCGGCTACCGAGGCTATACCTACCGCCCGCGAAGAGCCA
AGGGTTCCACCCGCCGTCCCCGCCGACGCGCCGCCACCACCCGCCGCCGCCGCCGAGACG
25 CCAGCCCGCACTGGCTCCAGTCTCCGTGAGGAAAGTGGCGCGCGACGGACACACCCTGGTGCTG
CCCAGGGCGCGCTACCACCCAGCATCGTTTAAAAGCCTGTTGTGGTTCTTGAGATATGGCCCT
CACTTGCCGCCTCCGTTTCCCGGTGCCGGGATACCGAGGAGGAAGATCGCGCCGACGAGGGGT
CTGGCCGGCCGCGGCCTGAGCGGAGGCAGCCGCCGCGCGCACCGGCGGGCGACGCGCCACCAGC
CGACGCATGCGCGGCGGGGTGCTGCCCCTGTTAATCCCCCTGATCGCCGCGGCGATCGGCGCCG
30 TGCCCGGGATCGCCTCCGTGGCCTTGCAAGCGTCCCAGAGGCATTGACAGACTTGCAAACCTTGC
AAATATGGAAAAAAACCCCAATAAAAAAGTCTAGACTCTCACGCTCGCTTGGTCTGTGACT
ATTTTGTAGAATGGAAGACATCAACTTTGCGTCGCTGGCCCCGCGTCACGGCTCGCGCCCGTTCC
TGGGACACTGGAACGATATCGGCACCAGCAACATGAGCGGTGGCGCCTTCAAGTTGGGGCTCTCT
GTGGAGCGGCATTAAAAGTATCGGGTCTGCCGTTAAAAATTACGGCTCCCGGGCCTGGAACAGC
35 AGCACGGGCCAGATGTTGAGAGACAAGTTGAAAGAGCAGAACTTCCAGCAGAAGGTGGTGGAG
GGCCTGGCCTCCGGCATCAACGGGGTGGTGGACCTGGCCAACCAGGCCGTGCAGAATAAGATCA
ACAGCAGACTGGACCCCCGCGCGCCGGTGGAGGAGGTGCCGCCGGCGCTGGAGACGGTGTCCC
CCGATGGGCGTGGCGAGAAGCGCCCGCGGCCCGATAGGGAAGAGACCACTCTGGTACGCGAGA
CCGATGAGCCGCCCCCGTATGAGGAGGCCCTGAAGCAAGGTCTGCCACCACGCGGCCCATCGC
40 GCCCATGGCCACCGGGGTGGTGGGCCGCCACACCCCGCCACGCTGGACTTGCCCTCCGCCCGCC
GATGTGCCGACGAGCAGAAGGCGGCACAGCCGGGGCCCGCCCGCGACCGCCTCCCGTTCCCTCG

CCGGTCCTCTGCGCCGCGCGGCCAGCGGCCCCCGCGGGGGGGTTCGCGAGGCACGGCAACTGGC
AGAGCACGCTGAACAGCATCGTGGGTCTGGGGTTCGCGTCCGTGAAGCGCCGCCGATGCTACTG
AATAGCTTAGCTAACGTGTTGTATGTGTGTATGCGCCCTATGTCGCCGCCAGAGGAGCTGCTGAG
TCGCCGCCGTTTCGCGCGCCACCACCACCGCCACTCCGCCCTCAAGATGGCGACCCCATCGAT
5 GATGCCGCAGTGGTCGTACATGCACATCTCGGGCCAGGACGCCTCGGAGTACCTGAGCCCCGGG
CTGGTGCAGTTTCGCCCCGCGCCACCGAGAGCTACTTCAGCCTGAGTAACAAGTTTAGGAACCCCA
CGGTGGCGCCACGCACGATGTGACCACCGACCGGTCTCAGCGCCTGACGCTGCGGTTTATTCC
CGTGACCGCGAGGACACCGCGTACTCGTACAAGGCGCGGTTACCCCTGGCCGTGGGCGACAAC
CGCGTGCTGGACATGGCCTCCACCTACTTTGACATCCGCGGGGTGCTGGACCGGGGTCCCCTTT
10 CAAGCCCTACTCTGGCACCGCCTACAACTCCCTGGCCCCCAAGGGCGCTCCCAACTCCTGCGAGT
GGGAGCAAGAGGAACTCAGGCAGTTGAAGAAGCAGCAGAAGAGGAAGAAGAAGATGCTGAC
GGTCAAGCTGAGGAAGAGCAAGCAGCTACCAAAAAGACTCATGTATATGCTCAGGCTCCCCTTT
CTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAACGGACGCTACAGCTACAGAACAAA
AACCTATTTATGCAGACCCTACATTCCAGCCCGAACCCCAAAATCGGGGAGTCCCAGTGGAATGA
15 GGCAGATGCTACAGTCGCCGGCGGTAGAGTGCTAAAGAAATCTACTCCCATGAAACCATGCTAT
GGTTCCTATGCAAGACCCACAAATGCTAATGGAGGTCAGGGTGTACTAACGGCAAATGCCCAGG
GACAGCTAGAATCTCAGGTTGAAATGCAATTCTTTTCAACTTCTGAAAACGCCCCGTAACGAGGC
TAACAACATTCAGCCCAAATTGGTGCTGTATAGTGAGGATGTGCACATGGAGACCCCGGATACG
CACCTTTCTTACAAGCCCGCAAAAAGCGATGACAATTCAAAAATCATGCTGGGTCAGCAGTCCA
20 TGCCCAACAGACCTAATTACATCGGCTTCAGAGACAACTTTATCGGCCTCATGTATTACAATAGC
ACTGGCAACATGGGAGTGCTTGCAGGTCAGGCCTCTCAGTTGAATGCAGTGGTGGACTTGCAAG
ACAGAAACACAGAACTGTCCTACCAGCTCTTGCTTGATTCCATGGGTGACAGAACCAGATACTT
TTCCATGTGGAATCAGGCAGTGGACAGTTATGACCCAGATGTTAGAATTATTGAAAATCATGGA
ACTGAAGACGAGCTCCCCAACTATTGTTTCCCTCTGGGTGGCATAGGGGTAAGTACACTTACCA
25 GGCTGTAAAAACCAACAATGGCAATAACGGGGGCCAGGTGACTTGGACAAAAGATGAACTTTT
GCAGATCGCAATGAAATAGGGGTGGGAAACAATTCGCTATGGAGATCAACCTCAGTGCCAACC
TGTGGAGAACTTCCTGTACTCCAACGTGGCGCTGTACCTACCAGACAAGCTTAAGTACAACCC
CTCCAATGTGGACATCTCTGACAACCCCAACACCTACGATTACATGAACAAGCGAGTGGTGGCC
CCGGGGCTGGTGGACTGCTACATCAACCTGGGCGCGCGCTGGTCGCTGGACTACATGGACAACG
30 TCAACCCCTTCAACCACCACCGCAATGCGGGCCTGCGCTACCGCTCCATGCTCCTGGGCAACGG
GCGCTACGTGCCCTTCCACATCCAGGTGCCCCAGAAGTTCTTTGCCATCAAGAACCTCCTCCTCC
TGCCGGGCTCCTACACCTACGAGTGGAACCTCAGGAAGGATGTCAACATGGTCCTCCAGAGCTC
TCTGGGTAAACGATCTCAGGGTGGACGGGGCCAGCATCAAGTTCGAGAGCATCTGCCTCTACGCC
ACCTTCTTCCCCATGGCCCACAACACGGCCTCCACGCTCGAGGCCATGCTCAGGAACGACACCA
35 ACGACCAGTCCCTCAATGACTACCTCTCCGCCGCAACATGCTCTACCCCATACCCGCCAACGCC
ACCAACGTCCCCATCTCCATCCCCTCGCGCAACTGGGCGGCCTTCCGCGGCTGGGCCTTACCCCG
CCTCAAGACCAAGGAGACCCCTCCCTGGGCTCGGGATTTCGACCCCTACTACACCTACTCGGGC
TCCATTCCCTACCTGGACGGCACCTTCTACCTCAACCACACTTCAAGAAGGTCTCGGTCACCTT
CGACTCCTCGGTCAGCTGGCCGGGCAACGACCGTCTGCTCACCCCCAACGAGTTCGAGATCAAG
40 CGCTCGGTTCGACGGGGAGGGCTACAACGTGGCCCAGTGCAACATGACCAAGGACTGGTTCCTGG
TCCAGATGCTGGCCAACTACAACATCGGCTACCAGGGCTTCTACATCCCAGAGAGCTACAAGGA

CAGGATGCTACTCTTCTTCAGGAACCTTCCAGCCCATGAGCCGCGAGGTGGACCAAGCAAG
TACAAGGACTACCAGGAGGTGGGCATCATCCACCAGCACAACTCGGGCTTCGTGGGCTACC
TCGCCCCCACCATGCGCGAGGGACAGGCCTACCCCGCCAACTTCCCCTATCCGCTCATAGGCAA
GACCGCGGTGCGACAGCATCACCAGAAAAAGTTCCTCTGCGACCGCACCCCTCTGGCGCATCCCC
5 TTCTCCAGCAACTTCATGTCCATGGGTGCGCTCTCGGACCTGGGCCAGAACTTGCTCTACGCCAA
CTCCGCCCACGCCCTCGACATGACCTTCGAGGTGACCCCATGGACGAGCCCACCCTTCTCTATG
TTCTGTTTCAAGTCTTTGACGTGGTCCGGGTCCACCAGCCGCACCGCGGCGTCATCGAGACCGTG
TACCTGCGTACGCCCTTCTCGGCCGGCAACGCCACCACCTAAAGAAGCAAGCCGCGAGTCATCGC
CGCCTGCATGCCGTGGGTTCACCGAGCAAGAGCTCAGGGCCATCGTCAGAGACCTGGGATGC
10 GGGCCCTATTTTTTGGGCACCTTCGACAAGCGCTTCCCTGGCTTTGTCTCCCCACACAAGCTGGC
CTGCGCCATCGTCAACACGGCCGGCCGCGAGACCGGGGGCGTGCACTGGCTGGCCTTCGCCTGG
AACCCGCGCTCCAAAACATGCTTCCCTTTGACCCCTTCGGCTTTTCGGACCAGCGGCTCAAGCA
AATCTACGAGTTCGAGTACGAGGGCTTGCTGCGTCGCAGCGCCATCGCCTCCTCGCCCGACCGCT
GCGTCACCCTCGAAAAGTCCACCCAGACCGTGCAGGGGGCCGACTCGGCCGCGCTGCGGTCTCTT
15 CTGCTGCATGTTTCTGCACGCCTTTGTGCACTGGCCTCAGAGTCCCATGGACCGCAACCCCCACCA
TGAAGTTGCTGACGGGGGTGCCCCAACTCCATGCTCCAGAGCCCCCAGGTGAGCCCCACCCTGCG
CCGCAACCAGGAGCAGCTCTACAGCTTCCCTGGAGCGCCACTCGCCTTACTTCCGCCGCCACAGC
GCACAGATCAGGAGGGCCACCTCCTTCTGCCACTTGCAAGAGATGCAAGAAGGGTAATAACGAT
GTACACACTTTTTTTCTCAATAAATGGCATCTTTTTATTTATACAAGCTCTCTGGGGTATTCATTT
20 CCCACCACCACCCGCCGTTGTGCGCATCTGGCTCTATTTAGAAATCGAAAGGGTCTTGCCGGGAG
TCGCCGTGCGCCACGGGCAGGGACAGTTGCGATACTGGTAGCGGGTGCCCCACTTGAAGTTCGG
GCACCACCAGGCGAGGCAGCTCGGGGAAGTTTTCGCTCCACAGGCTGCGGGTCAGCACCAGCGC
GTTTCATCAGGTGCGGCGCCGAGATCTTGAAGTCGCAGTTGGGGCCGCGGCCCTGCGCGCGCGAG
TTGCGGTACACCGGGTTGCAGCACTGGAACACCAACAGCGCCGGGTGCTTCACGCTGGCCAGCA
25 CGTGTGCGGTGCGAGATCAGCTCGGCGTCCAGGTCCCTCCGCGTTGCTCAGCGCGAACGGGGTCAT
CTTGGGCACTTGCCGCCCCAGGAAGGGCGCGTGCCCCGGTTTCGAGTTGCAGTCGCAGCGCAGC
GGGATCAGCAGGTGCCCCGTGCCCGGACTCGGCGTTGGGGTACAGCGCGCGCATGAAGGCCTGCA
TCTGGCGGAAGGCCATCTGGGCCCTTGCGCCCTCCGAGAAGAACATGCCGCAGGACTTGCCCGA
GAACTGGTTTTCGGGGCAGCTGGCGTCGTGCAGGCAGCAGCGCGCGTGGTGTGGCGATCTGC
30 ACCACGTTGCGCCCCCACCAGTTCTTCACGATCTTGGCCTTGGACGATTGCTCCTTCAGCGCGCG
CTGCCCCGTTCTCGCTGGTCACATCCATCTCGATCACATGTTTCCTTGTTACCATGCTGCTGCCGTG
CAGACACTTCAGCTCGCCCTCCGTCTCGGTGCAGCGGTGCTGCCACAGCGCGCAGCCCCGTGGGC
TCGAAAGACTTGTAGGTACCTCCGCGAAGGACTGCAGGTACCCCTGCAAAAAGCGGCCCATCA
TGGTCACGAAGGTCTTGTGCTGCTGAAGGTGAGTGCAGCCCCGCGGTGCTCCTCGTTACGCCAG
35 GTCTTGACACAGGCCGCCAGCGCACTCCACCTGGTCGGGCAGCATCTTGAAGTTACAGCTTCAGCTC
ATTCTCCACGTGGTACTTGTCCATCAGCGTGCGCGCCGCTCCATGCCCTTCTCCCAGGCCGACA
CCAGCGGCAGGCTCACGGGGTTCTTACCATACCCGTGGCCGCCGCTCCGCCGCGCTTTCGCTT
TCCGCCCCGCTGTTCTCTTCTCTTCTCTCTCTCTCTCGCCGCCGCCACTCGCAGCCCCCGCACC
ACGGGGTGGTCTTCTTCTGACGGCGCTGCACCTTGCGCTTGCCGTTGCGCCCCCTGCTTGATGCGCAC
40 GGGCGGGTTGCTGAAGCCCACCATCACCAGCGCGGCCTCTTCTTGCTCGTCCTCGCTGTCCAGAA
TGACCTCCGGGGAGGGGGGGTTGGTTCATCCTCAGTACCGAGGCACGCTTCTTTTTCTTCTTGGGG

GCGTTCGCCAGCTCCGCGGCTGCGGCCGCTGCCGAGGTCGAAGGCCGAGGGCTGGGCGTGCGCG
GCACCAGCGCGTCTTGCAGGCGTCTCGTCTCCTCGGACTCGAGACGGAGGCGGGCCCGCTT
CTTCGGGGGCGCGCGGGGCGGCGGAGGCGGCGGCGGACGGAGACGGGGACGAGACATCGTC
CAGGGTGGGTGGACGGCGGGCCGCGCCGCTCCGCGCTCGGGGGTGGTCTCGCGCTGGTCCTCT
5 TCCCCACTGGCCATCTCCCCTGCTCCTTCTCCTATAGGCAGAAAGAGATCATGGAGTCTCTCAT
GCGAGTCGAGAAGGAGGAGGACAGCCTAACCGCCCCCTTGAGCCCTCCACCACCGCCGCCACC
ACCGCCAATGCCGCCGCGGACGACGCGCCACCGAGACCACCGCCAGTACCACCCTCCCCAGCG
ACGCACCCCCGCTCGAGAATGAAGTGCTGATCGAGCAGGACCCGGGTTTTGTGAGCGGAGAGG
AGGATGAGGTGGATGAGAAGGAGAAGGAGGAGGTCGCCGCCTCAGTGCCAAAAGAGGATAAA
10 AAGCAAGACCAGGACGACGAGATAAGGATGAGACAGCAGTCGGGCGGGGAACGGAAGCCA
TGATGCTGATGACGGCTACCTAGACGTGGGAGACGACGTGCTGCTTAAGCACCTGCACCGCCAG
TGCGTCATCGTCTGCGACGCGCTGCAGGAGCGCTGCGAAGTGCCCCCTGGACGTGGCGGAGGTCA
GCCGCGCCTACGAGCGGCACCTCTTCGCGCCGACGTGCCCCCAAGCGCCGGGAGAACGGCAC
CTGCGAGCCCAACCCGCGTCTCAACTTCTACCCGGTCTTCGCGGTACCCGAGGTGCTGGCCACCT
15 ACCACATCTTTTTCCAAAAGTCAAGATCCCCCTCTCCTGCCGCGCCAACCGCACCCGCGCCGAC
AAAACCTTGACCCTGCGGCAGGGCGCCACATACCTGATATCGCCTCTCTGGAGGAAGTGCCCA
AGATCTTCGAGGGTCTCGGTGCGGACGAGAAACGGGCGGCGAACGCTCTGCACGGAGACAGCG
AAAACGAGAGTCACTCGGGGGTGTGGTGGAGCTCGAGGGCGACAACGCGCGCCTGGCCGTAC
TCAAGCGCAGCATAGAGGTCACCCACTTTGCCTACCCGGCGCTCAACCTGCCCCCAAGGTCAT
20 GAGTGTGGTCATGGGCGAGCTCATCATGCGCCGCGCCAGCCCCCTGGCCGCGGATGCAAACTTG
CAAGAGTCCTCCGAGGAAGGCCTGCCCCGCGTCAGCGACGAGCAGCTGGCGCGCTGGCTGGAG
ACCCGCGACCCCGCGCAGCTGGAGGAGCGGCGCAAGCTCATGATGGCCGCGGTGCTGGTCACCG
TGGAGCTCGAGTGTCTGCAGCGCTTCTTCGCGGACCCCGAGATGCAGCGCAAGCTCGAGGAGAC
CCTGCACTACACCTTCCGCCAGGGCTACGTGCGCCAGGCCTGCAAGATCTCCAACGTGGAGCTC
25 TGCAACCTGGTCTCCTACCTGGGCATCCTGCACGAGAACCGCCTCGGGCAGAACGTCTGCACT
CCACCCTCAAAGGGGAGGCGCGCCGCGACTACATCCGCGACTGCGCCTACCTCTTCCTCTGCTAC
ACCTGGCAGACGGCCATGGGGGTCTGGCAGCAGTGCCTGGAGGAGCGCAACCTCAAGGAGCTG
GAAAAGCTCCTCAAGCGCACCTCAGGGACCTCTGGACGGGCTTCAACGAGCGCTCGGTGGCCG
CCGCGCTGGCGGACATCATCTTTCCCGAGCGCCTGCTCAAGACCCTGCAGCAGGGCCTGCCCCA
30 CTTACCCAGCCAGAGCATGCTGCAGAACTTCAGGACTTTTCATCCTGGAGCGCTCGGGCATCCTGC
CGGCCACTTGCTGCGCGCTGCCAGCGACTTCGTGCCATCAAGTACAGGGAGTGCCCGCCGCC
GCTCTGGGGCCACTGCTACCTCTTCAGCTGGCCAACTACCTCGCCTACCACTCGGACCTCATGG
AAGACGTGAGCGGCGAGGGCCTGCTCGAGTGCCACTGCCGCTGCAACCTCTGCACGCCCCACCG
CTCTCTAGTCTGCAACCCGCGAGCTGCTCAGCGAGAGTCAGATTATCGGTACCTTCGAGCTGCAGG
35 GTCCCTCGCCTGACGAGAAGTCCGCGGCTCCAGGGCTGAACTCACTCCGGGGCTGTGGACTTC
CGCCTACCTACGCAAATTTGTACCTGAGGACTACCACGCCCACGAGATCAGGTTCTACGAAGAC
CAATCCCGCCCCGCCAAGGCGGAGCTACCGCCTGCGTCATCACCCAGGGGCACATCCTGGGCC
AATTGCAAGCCATCAACAAAGCCCGCCGAGAGTTCTTGCTGAAAAAGGGTCGGGGGGTGTACCT
GGACCCCCAGTCCGGCGAGGAGCTAAACCCGCTACCCCGCCGCCGCCAGCAGCGGGACCTT
40 GCTTCCAGGATGGCACCCAGAAAGAAGCAGCAGCCGCCGCCGCCGCCGAGCCATACATGCTT
CTGGAGGAAGAGGAGGAGGACTGGGACAGTCAGGCAGAGGAGGTTTCGGACGAGGAGCAGGA

GGAGATGATGGAAGACTGGGAGGAGGACAGCAGCCTAGACGAGGAAGCTTCAGAGGCCGAAG
AGGTGGCAGACGCAACACCATCGCCCTCGGTGCGAGCCCCCTCGCCGGGGCCCCCTGAAATCCTC
CGAACCCAGCACCAGCGCTATAACCTCCGCTCCTCCGGCGCCGGCGCCACCCGCCCCGAGACCC
AACCCTAGATGGGACACCACAGGAACCGGGGTTCGGTAAGTCCAAGTGCCCCGCCGCCACCG
5 CAGCAGCAGCAGCAGCAGCGCCAGGGCTACCGCTCGTGGCGCGGGCACAAGAACGCCATAGTC
GCCTGCTTGCAAGACTGCGGGGGCAACATCTCTTTTCGCCCCCGCTTCCTGCTATTCCACCACGG
GGTCGCCTTTCCCCGCAATGTCCTGCATTACTACCGTCATCTCTACAGCCCCCTACTGCAGCGGCG
ACCCAGAGGCGGCAGCGGCAGCCACAGCGGCGACCACCACCTAGGAAGATATCCTCCGCGGGC
AAGACAGCGGCAGCAGCGGCCAGGAGACCCGCGGCAGCAGCGGCGGGAGCGGTGGGCGCACT
10 GCGCCTCTCGCCAACGAACCCCTCTCGACCCGGGAGCTCAGACACAGGATCTTCCCCACTTTGT
ATGCCATCTTCCAACAGAGCAGAGGCCAGGAGCAGGAGCTGAAAATAAAAAACAGATCTCTGC
GCTCCCTCACCCGAGCTGTCTGTATCACAAAAGCGAAGATCAGCTTCGGCGCACGCTGGAGGA
CGCGGAGGCACTCTTCAGCAAATACTGCGCGCTCACTCTTAAAGACTAGCTCCGCGCCCTTCTCG
AATTTAGGCGGGAGAAAACTACGTCATCGCCGGCCGCCGCCAGCCCGCCAGCCGAGATGAGC
15 AAAGAGATTCCACGCCATACATGTGGAGCTACCAGCCGCAGATGGGACTCGCGGCGGGAGCG
GCCCAGGACTACTCCACCCGCATGAACTACATGAGCGCGGGACCCACATGATCTCACAGGTCA
ACGGGATCCGCGCCCAGCGAAACCAAATACTGCTGGAACAGGCGGCCATCACCGCCACGCCCC
GCCATAATCTCAACCCCCGAAATTGGCCCGCCGCCCTCGTGTACCAGGAAACCCCTCCGCCAC
CACCGTACTACTTCCGCGTGACGCCCAGGCCGAAGTCCAGATGACTAACTCAGGGGCGCAGCTC
20 GCGGGCGGCTTTCGTACGGGGCGCGGCCGCTCCGACCAGGTATAAGACACCTGATGATCAGAG
GCCGAGGTATCCAGCTCAACGACGAGTCGGTGAGCTCTTCGCTCGGTCTCCGTCCGACGGAAC
TTTCAGCTCGCCGGATCCGGCCGCTCTTCGTTACGCCCCGCCAGGCGTACCTGACTCTGCAGA
CCTCGTCTCGGAGCCCCGCTCCGGCGGCATCGGAACCCCTCCAGTTCGTGGAGGAGTTCGTGCCC
TCGGTCTACTTCAACCCCTTCTCGGGACCTCCCGGACGCTACCCCGACCAGTTCATTCCGAACCT
25 TGACGCGGTGAAGGACTCGGCGGACGGCTACGACTGAATGTCAGGTGTCGAGGCAGAGCAGCT
TCGCCTGAGACACCTCGAGCACTGCCGCCGCCACAAGTGCTTCGCCCCGCGTTCTGGTGAGTTCT
GCTACTTTTCAGTACCCGAGGAGCATACCGAGGGGCCGGCGCACGGCGTCCGCCTGACCACCCA
GGGCGAGGTTACCTGTTCCCTCATCCGGGAGTTTACCCTCCGTCCCCTGCTAGTGGAGCGGGAGC
GGGGTCCCTGTGTCTAACTATCGCCTGCAACTGCCCTAACCTGGATTACATCAAGATCTTTGC
30 TGTCATCTCTGTGCTGAGTTTAATAAACGCTGAGATCAGAATCTACTGGGGCTCCTGTGCCATC
CTGTGAACGCCACCGTCTTCACCCACCCCGACCAGGCCAGGCGAACCTCACCTGCGGTCTGCA
TCGGAGGGCCAAGAAGTACCTCACCTGGTACTTCAACGGCACCCCCCTTTGTGGTTTACAACAGCT
TCGACGGGGACGGAGTCTCCCTGAAAGACCAGCTCTCCGGTCTCAGCTACTCCATCCACAAGAA
CACCACCTCCAACCTTCCCTCCCTACCTGCCGGGAACCTACGAGTGCGTCACCGGCCGCTGCA
35 CCCACCTACCCGCCTGATCGTAAACCAGAGCTTTCCGGGAACAGATAACTCCCTCTTCCCCAGA
ACAGGAGGTGAGCTCAGGAAACTCCCCGGGGACCAGGGCGGAGACGTACCTTCGACCCTTGTG
GGGTTAGGATTTTTTATTACCGGGTTGCTGGCTCTTTAATCAAAGTTTCCTTGAGATTTGTTCTT
TCCTTCTACGTGTATGAACACCTCAACCTCCAATAACTCTACCCTTTCTTCGGAATCAGGTGACTT
CTCTGAAATCGGGCTTGGTGTGCTGCTTACTCTGTTGATTTTTTCTTATCATACTCAGCCTTCT
40 GTGCCTCAGGCTCGCCGCTGCTGCGCACACATCTATATCTACTGCTGGTTGCTCAAGTGCAGGG
GTCGCCACCCAAGATGAACAGGTACATGGTCTATCGATCCTAGGCCTGCTGGCCCTGGCGGCC

TGCAGCGCCGCCAAAAAAGAGATTACCTTTGAGGAGCCCGCTTGCAATGTAACCTTCAAGCCCG
AGGGTGACCAATGCACCACCCTCGTCAAATGCGTTACCAATCATGAGAGGCTGCGCATCGACTA
CAAAAACAAAACCTGGCCAGTTTGCGGTCTATAGTGTGTTACGCCCCGAGACCCCTCTAACTACT
CTGTCACCGTCTTCCAGGGCGGACAGTCTAAGATATTCAATTACACTTTCCCTTTTTATGAGTTAT
5 GCGATGCGGTCATGTACATGTCAAAACAGTACAACCTGTGGCCTCCCTCTCCCCAGGCGTGTGTG
GAAAATACTGGGTCTTACTGCTGTATGGCTTTCGCAATCACTACGCTCGCTCTAATCTGCACGGT
GCTATACATAAAATTCAGGCAGAGGCGAATCTTTATCGATGAAAAGAAAATGCCTTGATCGCTA
ACACCGGCTTTTCTATCTGCAGAATGAATGCAATCACCTCCCTACTAATCACCACCACCCTCCTTG
CGATTGCCCATGGGTTGACACGAATCGAAGTGCCAGTGGGGTCCAATGTCACCATGGTGGGCCC
10 CGCCGGCAATTCCACCCTCATGTGGGAAAAATTTGTCCGCAATCAATGGGTTCAATTTCTGCTCTA
ACCGAATCAGTATCAAGCCCAGAGCCATCTGCGATGGGCAAAATCTAACTCTGATCAATGTGCA
AATGATGGATGCTGGGTACTATTACGGGCAGCGGGGAGAAATCATTAACTTGGCGACCCAC
AAGGACTACATGCTGCATGTAGTCGAGGCACCTCCCACTACCACCCCACTACCACCTCTCCAC
CACCACCACCACTACTACTACTACTACTACTACTACTACTACTACCCTGCCCCGCCATA
15 CCGCAAAAGCACCATGATTAGCACAAAGCCCCCTCGTGCTCACTCCACGCGCGGGGCCCAT
CGGTGCGACCTCAGAAACCACCGAGCTTTGCTTCTGCCAATGCACTAACGCCAGCGCTCATGAA
CTGTTTCGACCTGGAGAATGAGGATGTCCAGCAGAGCTCCGCTTGCTGACCCAGGAGGCTGTGG
AGCCCGTTGCCCTGAAGCAGATCGGTGATTCAATAATTGACTCTTCTTCTTTGCCACTCCCGAA
TACCCTCCCGATTCTACTTTCCACATCACGGGTACCAAAGACCCTAACCTCTCTTTCTACCTGATG
20 CTGCTGCTCTGTATCTCTGTGGTCTCTTCCGCGCTGATGTTACTGGGGATGTTCTGCTGCCTGATC
TGCCGCAGAAAGAGAAAAAGCTCGCTCTCAGGGCCAACCACTGATGCCCTTCCCTACCCCCCGG
ATTTTGAGATAACAAGATATGAGCTCGCTGCTGACACTAACCGCTTACTAGCCTGCGCTCTAA
CCCTTGTCGCTTGCGACTCGAGATTCCACAATGTCACAGCTGTGGCAGGAGAAAAATGTTACTTTC
AACTCCACGGCCGATACCCAGTGGTCTGGAGTGGCTCAGGTAGCTACTTAACTATCTGCAATA
25 GCTCCACTTCCCCCGGCATATCCCCAACCAAGTACCAATGCAATGCCAGCCTGTTACCCCTCATC
AACGCTTCCACCCTGGACAATGGACTCTATGTAGGCTATGTACCCCTTGGTGGGCAAGGAAAGA
CCCACGCTTACAACCTGGAAGTTCGCCAGCCCAGAACCACTACCCAAGCTTCTCCACCACCAC
CACCACCACCACCATCACCAGCAGCAGCAGCAGCAGCAGCCACAGCAGCAGCAGCAGATTATT
GACTTTGGTTTTGGCCAGCTCATCTGCCGCTACCCAGGCCATCTACAGCTCTGTGCCCGAAACCA
30 CTCAGATCCACCGCCCAGAAACGACCACCGCCACCACCCTACACACCTCCAGCGATCAGATGCC
GACCAACATCACCCCTTGGCTCTTCAAATGGGACTTACAAGCCCCACTCCAAAACAGTGGAT
GCGGCCGAGGTCTCCGCCCTCGTCAATGACTGGGCGGGGCTGGGAATGTGGTGGTTCGCCATAG
GCATGATGGCGCTCTGCCTGCTTCTGCTCTGGCTCATCTGCTGCCTCCACCGCAGGCGAGCCAGA
CCCCCATCTATAGACCCATCATTGTCCTGAACCCCGATAATGATGGGATCCATAGATTGGATGG
35 CCTGAAAAACCTACTTTTTTCTTTTACAGTATGATAAATTGAGACATGCCTCGCATTTTCTTGTA
ATGTTCTTCTCCACCTTTTCTGGGGTGTCTACGCTGGCCGCTGTGTCTCACCTGGAGGTAGAC
TGCTCTCACCTTCACTGTCTACCTGCTTTACGGATTGGTCACCCTCACTCTCATCTGCAGCCTA
ATCACAGTAATCATCGCCTTCATCCAGTGCATTGATTACATCTGTGTGCGCCTCGCATACTTCAG
ACACCACCCGAGTACCGAGACAGGAACATTGCCCAACTTCTAAGACTGCTCTAATCATGCATA
40 AGACTGTGATCTGCCTTCTGATCCTCTGCATCCTGCCACCCTCACCTCCTGCCAGTACACCACA
AAATCTCCGCGCAAAAGACATGCCTCCTGCCGCTTACCCAACTGTGGAATATACCCAAATGCT

ACAACGAAAAGAGCGAGCTCTCCGAAGCTTGGCTGTATGGGGTCATCTGTGTCTTAGTTTTCTGC
AGCACTGTCTTTGCCCTCATAATCTACCCCTACTTTGATTTGGGATGGAACGCGATCGATGCCAT
GAATTACCCACCTTTCCCGCACCCGAGATAATTCCACTGCGACAAGTTGTACCCGTTGTCTGTTA
ATCAACGCCCCCATCCCCCTACGCCCACTGAAATCAGCTACTTTAACCTAACAGGCGGAGATGA
5 CTGACGCCCTAGATCTAGAAATGGACGGCATCAGTACCGAGCAGCGTCTCCTAGAGAGGCGCAG
GCAGGCGGCTGAGCAAGAGCGCCTCAATCAGGAGCTCCGAGATCTCGTTAACCTGCACCAGTGC
AAAAGAGGCATCTTTTGTCTGGTAAAGCAGGCCAAAGTCACCTACGAGAAGACCGGCAACAGC
CACCGCCTCAGTTACAAATTGCCACCCAGCGCCAGAAAGCTGGTGCTCATGGTGGGTGAGAATC
CCATCACCGTCACCCAGCACTCGGTAGAGACCGAGGGGTGTCTGCACTCCCCCTGTCTGGGGTCC
10 AGAAGACCTCTGCACCCTGGTAAAGACCCTGTGCGGTCTCAGAGATTTAGTCCCCTTTAACTAAT
CAAACACTGGAATCAATAAAAAGAATCACTTACTTAAAATCAGACAGCAGGTCTCTGTCCAGTT
TATTCAGCAGCACCTCCTTCCCCCTCCTCCCAACTCTGGTACTCCAAACGCCTTCTGGCGGCAAAC
TTCCTCCACACCCTGAAGGGAATGTCAGATTCTTGCTCCTGTCCCTCCGCACCCACTATCTTCATG
TTGTTGAGATGAAGCGCACAAAACGTCTGACGAGAGCTTCAACCCCGTGTACCCCTATGACA
15 CGGAAAGCGGCCCTCCCTCCGTCCCTTTCCTCACCCCTCCCTTCGTGTCTCCCGATGGATTCCAA
GAAAGTCCCCCGGGGTCTGTCTCTGAACCTGGCCGAGCCCCCTGGTCACTTCCCACGGCATGCT
CGCCCTGAAAATGGGAAGTGGCCTCTCCCTGGACGACGCTGGCAACCTCACCTCTCAAGATATC
ACCACCGCTAGCCCTCCCTCAAAAAAACCAAGACCAACCTCAGCCTAGAAACCTCATCCCCC
TAACTGTGAGCACCTCAGGCGCCCTACCGTAGCAGCCGCCGCTCCCCTGGCGGTGGCCGGCAC
20 CTCCCTCACCATGCAATCAGAGGCCCCCTGACAGTACAGGATGCAAAACTCACCTTGGCCACC
AAAGGCCCCCTGACCGTGTCTGAAGGCAAACCTGGCCTTGCAAACATCGGCCCCGCTGACGGCCG
CTGACAGCAGCACCTCACAGTCAGTGCCACACCACCCCTTAGCACAAGCAATGGCAGCTTGGG
TATTGACATGCAAGCCCCCATTTACACCACCAATGGAAAACCTAGGACTTAACTTTGGCGCTCCCC
TGCATGTGGTAGACAGCCTAAATGCACTGACTGTAGTTACTGGCCAAGGTCTTACGATAAACGG
25 AACAGCCCTACAACTAGAGTCTCAGGTGCCCTCAACTATGACACATCAGGAAACCTAGAATTG
AGAGCTGCAGGGGGTATGCGAGTTGATGCAAAATGGTCAACTTATCCTTGATGTAGCTTACCCATT
TGATGCACAAAACAATCTCAGCCTTAGGCTTGGACAGGGACCCCTGTTTGTTAACTCTGCCCACA
ACTTGGATGTAACTACAACAGAGGCCTCTACCTGTTACATCTGGAAATACCAAAAAGCTAGA
AGTTAATATCAAAACAGCCAAGGGTCTCATTTATGATGACACTGCTATAGCAATCAATGCGGGT
30 GATGGGCTACAGTTTGACTCAGGCTCAGATACAAATCCATTAAAACTAACTTGGATTAGGAC
TGGATTATGACTCCAGCAGAGCCATAATTGCTAAACTGGGAACTGGCCTAAGCTTTGACAACAC
AGGTGCCATCACAGTAGGCAACAAAAATGATGACAAGCTTACCTTGTGGACCACACCAGACCCA
TCCCCTAACTGTAGAATCTATTAGAGAAAGATGCTAAATTCACACTTGTTTTGACTAAATGCGG
CAGTCAGGTGTTGGCCAGCGTTTCTGTTTTATCTGTAAAAGGTAGCCTTGCGCCCATCAGTGGCA
35 CAGTAACTAGTGCTCAGATTGTCTCAGATTGATGAAAATGGAGTTCTACTAAGCAATTCTTCC
CTTGACCCTCAATACTGGAACCTACAGAAAAGGTGACCTTACAGAGGGCACTGCATATACCAACG
CAGTGGGATTTATGCCCAACCTCACAGCATACCAAAAACACAGAGCCAAACTGCTAAAAGCAA
CATTGTAAGTCAGGTTTACTTGAATGGGGACAAATCCAAACCCATGACCCTCACCATTACCCTCA
ATGGAATAATGAAACAGGAGATGCCACAGTAAGCACTTACTCCATGTCATTCTCATGGAAGT
40 GAATGGAAGTAATTACATTAATGAAACGTTCCAAACCAACTCCTTCACCTTCTCCTACATCGCCC
AAGAATAAAAAGCATGACGCTGTTGATTTGATTCAATGTGTTTCTGTTTTATTTCAAGCACAAC

AAAATCATTCAAGTCATTCTTCCATCTTAGCTTAATAGACACAGTAGCTTAATAGACCCAGTAGT
GCAAAGCCCCATTCTAGCTTATAGATCAGACAGTGATAATTAACCACCACCACCACCATACCTTT
TGATTACAGGAAATCATGATCATCACAGGATCCTAGTCTTCAGGCCGCCCCCTCCCTCCCAAGACA
CAGAATACACAGTCCTCTCCCCCGACTGGCTTTAAATAACACCATCTGGTTGGTCACAGACATG
5 TTCTTAGGGGTTATATTCCACACGGTCTCCTGCCGCGCCAGGCGCTCGTCGGTGATGTTGATAAA
CTCTCCCGGCAGCTCGCTCAAGTTCACGTCGCTGTCCAGCGGCTGAACCTCCGGCTGACGCGATA
ACTGTGCGACCGGCTGCTGGACGAACGGAGGCCGCGCCTACAAGGGGGTAGAGTCATAATCCTC
GGTCAGGATAGGGCGGTGATGCAGCAGCAGCGAGCGAAACATCTGCTGCCGCCGCCGCTCCGTC
CGGCAGGAAAACAACACGCCGGTGGTCTCCTCCGCGATAATCCGCACCGCCCGCAGCATCAGCT
10 TCCTCGTTCTCCGCGCGCAGCACCTCACCTTATCTCGCTCAAATCGGCGCAGTAGGTACAGCAC
AGCACCACGATGTTATTCATGATCCACAGTGCAGGGCGCTGTATCCAAAGCTCATGCCGGGAA
CCACCGCCCCACGTGGCCATCGTACCACAAGCGCACGTAAATCAAGTGTCGACCCCTCATGAA
CGCGCTGGACACAAACATTACTTCTTGGGCATGTTGTAATTCACCACCTCCCGGTACCAGATAA
ACCTCTGGTTGAACAGGGCACCTTCCACCACCATCTGAACCAAGAGGCCAGAACCTGCCACC
15 GGCTATGCACTGCAGGGAACCCGGGTTGGAACAATGACAATGCAGACTCCAAGGCTCGTAACCG
TGGATCATCCGGCTGCTGAAGGCATCGATGTTGGCACAACACAGACACACGTGCATGCACTTTC
TCATGATTAGCAGCTCTTCCCTCGTCAGGATCATATCCCAAGGAATAACCCATTCTTGAATCAAC
GTAAACCCACACAGCAGGGAAGGCCTCGCACATAACTCACGTTGTGCATGGTCAGCGTGTTCG
20 ATTCCGGAAACAGCGGATGATCCTCCAGTATCGAGGCGCGGGTCTCCTTCTCACAGGGAGGTAA
AGGGTCCCTGCTGTACGGAAGTGCAGGCGGACGACCGAGATCGTGTTGAGCGTAGTGTCATGGAA
AAGGGAACGCCGACGTGGTCATACTTCTTGAAGCAGAACCAGGTTTCGCGCGTGGCAGGCCTCC
TTGCGTCTGCGGTCTCGCCGTCTAGCTCGCTCCGTGTGATAGTTGTAGTACAGCCACTCCCGCAG
AGCGTCGAGGCGCACCCCTGGCTTCCGGATCTATGTAGACTCCGTCTTGACCGCGGCCCTGATAA
TATCCACCACCGTAGAATAAGCAACACCCAGCCAAGCAATACACTCGCTCTGCGAGCGGCAGAC
25 AGGAGGAGCGGGCAGAGATGGGAGAACCATGATAAAAACTTTTTTAAAGAATATTTTCCAAT
TCTTCGAAAGTAAGATCTATCAAGTGGCAGCGCTCCCTCCACTGGCGCGGTCAAACCTCTACGG
CCAAAGCACAGACAACGGCATTCTAAGATGTTCTTAATGGCGTCCAAAAGACACACCGCTCT
CAAGTTGCAGTAACTATGAATGAAAACCCATCCGGCTGATTTTCCAATATAGACGCGCCGGCA
GCGTCCACCAAAACCCAGATAATTTTCTTCTCTCCAGCGGTTTACGATCTGTCTAAGCAAATCCCT
30 TATATCAAGTCCGACCATGCCAAAAATCTGCTCAAGAGCGCCCTCCACCTTCATGTACAAGCAG
CGCATCATGATTGCAAAAATTCAGGTTCTTCAGAGACCTGTATAAGATTCAAAATGGGAACATT
AACAAAAATTCCTCTGTGCGCGAGATCCCTTCGCAGGGCAAGCTGAACATAATCAGACAGGTCC
GAACGGACCAAGTGAGGCCAAATCCCCACCAGGAACCAGATCCAGAGACCCTATACTGATTATGA
CGCGCATACTCGGGGCTATGCTGACCAGCGTAGCGCCGATGTAGGCGTGCTGCATGGGCGGCGA
35 GATAAAATGCAAAGTGCTGGTTAAAAAATCAGGCAAAGCCTCGCGCAAAAAGCTAACACATC
ATAATCATGCTCATGCAGGTAGTTGCAGGTAAGCTCAGGAACCAAAACGGAATAACACACGATT
TTCCTCTCAAACATGACTTCGCGGATACTGCGTAAAACAAAAAATTATAAATAAAAAATTAATT
AAATAACTTAAACATTGGAAGCCTGTCTCACAACAGGAAAAACCACTTTAATCAACATAAGACG
GGCCACGGGCATGCCGGCATAGCCGTAAAAAAATTTGGTCCCCGTGATTAACAAGTACCACAGAC
40 AGCTCCCCGGTCATGTGCGGGGTCATCATGTGAGACTCTGTATACACGTCTGGATTGTGAACATC
AGACAAACAAAGAAATCGAGCCACGTAGCCCGGAGGTATAATCACCCGCAGGCGGAGGTACAG

CAAAACGACCCCCATAGGAGGAATCACAAAATTAGTAGGAGAAAAAATACATAAACACCAGA
AAAACCCTGTTGCTGAGGCAAAATAGCGCCCTCCCGATCCAAAACAACATAAAGCGCTTCCACA
GGAGCAGCCATAACAAAGACCCGAGTCTTACCAGTAAAAGAAAAAAGATCTCTCAACGCAGCA
CCAGCACCAACACTTCGCAGTGTAAGGCAAGTGCCGAGAGAGTATATATAGGAATAAAAA
5 GTGACGTAAACGGGCAAAGTCCAAAAACGCCAGAAAAACCGCACGCGAACCTACGCCCCGA
AACGAAAGCCAAAAAACAAGTACACTCCCTTCCGGCGTCAACTTCCGCTTTCCACGCTACGT
CACTTGCCCCAGTCAAACAACTACATATCCCGAACTTCCAAGTCGCCACGCCCCAAACACCGC
CTACACCTCCCCGCGCCGCGCCGCCCCAAACCGCCTCCCGCCCCGCGCCCCGCGCCGCGC
10 GCCATCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTATTGATGATGGTTTAAACG
GATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGAGTATTCTATAGTGTCACCTAAATAGCTT
GGCGTAATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACA
TACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
GCCAACGCGAACCCCTTGCGGCCGCGCGGGCGTCGACCAATTCTCATGTTTGACAGCTTATCAT
15 CGAATTTCTGCCATTCATCCGCTTATTATCACTTATTCAGGCGTAGCAACCAGGCGTTTAAAGGC
ACCAATAACTGCCTTAAAAAATTACGCCCCGCGCTGCCACTCATCGCAGTACTGTTGTAATTCA
TTAAGCATTCTGCCGACATGGAAGCCATCACAAACGGCATGATGAACCTGAATCGCCAGCGGCA
TCAGCACCTTGTCGCTTGCGTATAATATTTGCCATGGTGAAAACGGGGGCGAAGAAGTTGTC
CATATTGGCCACGTTTAAATCAAACTGGTGAACTCACCCAGGGATTGGCTGAGACGAAAAAC
20 ATATTCTCAATAAACCCCTTAGGGAAAATAGGCCAGGTTTTACCGTAACACGCCACATCTTGCGA
ATATATGTGTAGAACTGCCGGAATCGTCGTGGTATTCACTCCAGAGCGATGAAAACGTTTCA
GTTTGCTCATGGAACGGTGTAACAAGGTGAACACTATCCCATATCACCAGCTCACCCTCTTT
CATTGCCATACGGAATTCCGGATGAGCATTATCATCAGGCGGGCAAGAATGTGAATAAAGGCCGGA
TAAAACTTGTCGCTTATTTTCTTTACGGTCTTAAAAAAGGCCGTAATATCCAGCTGAACGGTCTG
25 GTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGATA
TATCAACGGTGTTATATCCAGTGATTTTTTCTCCATTTAGCTTCCTTAGCTCCTGAAAATCTCG
ATAACTCAAAAAATACGCCCCGTAGTGATCTTATTTTATTATGGTGAAAGTTGGAACCTCTTACG
TGCCGATCAACGTCCTATTTTCGCCAAAAGTTGGCCAGGGCTTCCCGGTATCAACAGGGACAC
CAGGATTTATTTATTCTGCGAAGTGATCTTCCGTCACAGGTATTTATTCGCGATAAGCTCATGGA
30 GCGGCGTAACCGTCGCACAGGAAGGACAGAGAAAGCGCGGATCTGGGAAGTGACGGACAGAAC
GGTCAGGACCTGGATTGGGGAGGCGGTTGCCGCCGCTGCTGCTGACGGTGTGACGTTCTCTGTTT
CGGTCACACCACATACGTTCCGCCATTCTATGCGATGCACATGCTGTATGCCGGTATACCGCTG
AAAGTTCTGCAAAGCCTGATGGGACATAAGTCCATCAGTTCAACGGAAGTCTACACGAAGGTTT
TTGCGCTGGATGTGGCTGCCCCGCCACCGGGTGCAGTTTGCGATGCCGGAGTCTGATGCGGTTGC
35 GATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGCCTTTATATGGAAATGTGGAAGTGAAGT
GGATATGCTGTTTTTGTCTGTAAACAGAGAAGCTGGCTGTTATCCACTGAGAAGCGAACGAAA
CAGTCGGGAAAATCTCCATTATCGTAGAGATCCGCATTATTAATCTCAGGAGCCTGTGTAGCGT
TTATAGGAAGTAGTGTTCTGTCTATGATGCCTGCAAGCGGTAACGAAAACGATTGAATATGCCTT
CAGGAACAATAGAAATCTTCGTGCGGTGTTACGTTGAAGTGGAGCGGATTATGTCAGCAATGGA
40 CAGAACAACCTAATGAACACAGAACCATGATGTGGTCTGTCTTTTACAGCCAGTAGTGCTCGC
CGCAGTCGAGCGACAGGGCGAAGCCCTCGAGTGAGCGAGGAAGCACCAGGGAACAGCACTTAT

ATATTCTGCTTACACACGATGCCTGAAAAAAGTTCCCTTGGGGTTATCCACTTATCCACGGGGAT
ATTTTTATAATTATTTTTTTTATAGTTTTTAGATCTTCTTTTTTAGAGCGCCTTGTAGGCCTTTATC
CATGCTGGTTCTAGAGAAGGTGTTGTGACAAATTGCCCTTTCAGTGTGACAAATCACCCCTCAAAT
GACAGTCCTGTCTGTGACAAATTGCCCTTAACCCCTGTGACAAATTGCCCTCAGAAGAAGCTGTTT
5 TTTACAAAGTTATCCCTGCTTATTGACTCTTTTTTATTTAGTGTGACAATCTAAAAAGTTGTCAC
ACTTCACATGGATCTGTCTATGGCGGAAACAGCGGTTATCAATCACAAGAAACGTAAAAATAGCC
CGCGAATCGTCCAGTCAAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCAAAAACG
TATGCTGTATCTGTTTCGTTGACCAGATCAGAAAATCTGATGGCACCCCTACAGGAACATGACGGT
ATCTGCGAGATCCATGTTGCTAAATATGCTGAAATATTCGGATTGACCTCTGCGGAAGCCAGTA
10 AGGATATACGGCAGGCATTGAAGAGTTTCGCGGGGAAGGAAGTGGTTTTTTATCGCCCTGAAGA
GGATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCTTGGTTTATCAAACGTGCGCACAGTCCAT
CCAGAGGGCTTTACAGTGTACATATCAACCCATATCTCATTCCCTTCTTTATCGGGTTACAGAAC
CGGTTTACGCAGTTTCGGCTTAGTGAAACAAAAGAAATCACCAATCCGTATGCCATGCGTTTATA
CGAATCCCTGTGTCTAGTATCGTAAGCCGGATGGCTCAGGCATCGTCTCTCTGAAAATCGACTGG
15 ATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTATGCCTGACTTCCGCCGCCGCTTCCT
GCAGGTCTGTGTTAATGAGATCAACAGCAGAACTCCAATGCGCCTCTCATACATTGAGAAAAAG
AAAGGCCGCCAGACGACTCATATCGTATTTTCCTTCCGCGATATCACTTCCATGACGACAGGATA
GTCTGAGGGTTATCTGTCTACAGATTTGAGGGTGGTTCGTCTACATTTGTTCTGACCTACTGAGGGT
AATTTGTCTACAGTTTGTCTGTTTCCTTCAGCCTGCATGGATTTTCTCATACTTTTGAAGTGAAT
20 TTTTAAGGAAGCCAAATTTGAGGGCAGTTTGTCTACAGTTGATTTCCCTTCTCTTTCCCTTCGTCATG
TGACCTGATATCGGGGGTTAGTTCGTCTATCATTGATGAGGGTTGATTATCACAGTTTATTACTCT
GAATTGGCTATCCGCGTGTGTACCTCTACCTGGAGTTTTTCCACGGTGGATATTTCTTCTTGCGC
TGAGCGTAAGAGCTATCTGACAGAACAGTTCTTCTTTGCTTCCTCGCCAGTTCGCTCGCTATGCT
CGGTTACACGGCTGCGGCGAGCGCTAGTGATAATAAGTGACTGAGGTATGTGCTCTTCTTATCTC
25 CTTTTGTAGTGTTGCTCTTATTTTAAACAAGTTTTCGCGTTTTTTGATGACTTTGCGATTTTGTGTT
GCTTTGCAGTAAATTGCAAGATTTAATAAAAAAACGCAAAGCAATGATTAAAGGATGTTTCAGAA
TGAAACTCATGGAAACACTTAACCAAGTGCATAAACGCTGGTTCATGAAATGACGAAGGCTATCGC
CATTGCACAGTTTAAATGATGACAGCCCGGAAGCGAGGAAAATAACCCGGCGCTGGAGAATAGG
TGAAGCAGCGGATTTAGTTGGGGTTTCTTCTCAGGCTATCAGAGATGCCGAGAAAGCAGGGCGA
30 CTACCGCACCCGGATATGGAATTCGAGGACGGGTTGAGCAACGTGTTGGTTATACAATTGAAC
AAATTAATCATATGCGTGATGTGTTTGGTACGCGATTGCGACGTGCTGAAGACGTATTTCCACCG
GTGATCGGGGTTGCTGCCATAAAGGTGGCGTTTACAAAACCTCAGTTTCTGTTTCATCTTGCTCA
GGATCTGGCTCTGAAGGGGCTACGTGTTTTGCTCGTGGAAGGTAACGACCCCCAGGGAACAGCC
TCAATGTATCACGGATGGGTACCAGATCTTCATATTCATGCAGAAGACACTCTCCTGCCTTTCTA
35 TCTTGGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGCTGGCCGGGGCTTGACATT
ATTCCTTCCTGTCTGGCTCTGCACCGTATTGAACTGAGTTAATGGGCAAATTTGATGAAGGTAA
ACTGCCACCGATCCACACCTGATGCTCCGACTGGCCATTGAACTGTTGCTCATGACTATGATG
TCATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCACGATTAATGTCGTATGTGCTGCTGAT
GTGCTGATTGTTCCACGCCTGCTGAGTTGTTTGACTACACCTCCGCACTGCAGTTTTTCGATATG
40 CTTCTGTATCTGCTCAAGAAGCTTGATCTTAAAGGGTTCGAGCCTGATGTACGTATTTTGCTTAC
CAAATACAGCAATAGTAATGGCTCTCAGTCCCCGTGGATGGAGGAGCAAATTCGGGATGCCTGG

GGAAGCATGGTTCTAAAAAATGTTGTACGTGAAACGGATGAAGTTGGTAAAGGTCAGATCCGGA
TGAGAACTGTTTTTGAACAGGCCATTGATCAACGCTCTTCAACTGGTGCCTGGAGAAATGCTCTT
TCTATTTGGGAACCTGTCTGCAATGAAATTTTCGATCGTCTGATTAAACCACGCTGGGAGATTAG
ATAATGAAGCGTGCGCCTGTTATTCCAAAACATACGCTCAATACTCAACCGGTTGAAGATACTTC
5 GTTATCGACACCAGCTGCCCCGATGGTGGATTTCGTTAATTGCGCGCGTAGGAGTAATGGCTCGC
GGTAATGCCATTACTTTGCCTGTATGTGGTCGGGATGTGAAGTTTACTCTTGAAGTGCTCCGGGG
TGATAGTGTGAGAAGACCTCTCGGGTATGGTCAGGTAATGAACGTGACCAGGAGCTGCTTACT
GAGGACGCACTGGATGATCTCATCCCTTCTTTTCTACTGACTGGTCAACAGACACCGGCGTTCCG
TCGAAGAGTATCTGGTGTATAGAAAATTGCCGATGGGAGTCGCCGTCGTAAAGCTGCTGCACTT
10 ACCGAAAGTGATTATCGTGTCTGGTTGGCGAGCTGGATGATGAGCAGATGGCTGCATTATCCA
GATTGGGTAACGATTATCGCCCAACAAGTGCTTATGAACGTGGTCAGCGTTATGCAAGCCGATT
GCAGAATGAATTTGCTGGAAAATTTCTGCGCTGGCTGATGCGGAAAATATTTACGTAAGATT
ATTACCCGCTGTATCAACACCGCCAAATTCCTAAATCAGTTGTTGCTCTTTTTTCTCACCCCGGT
GAACTATCTGCCCGGTCAGGTGATGCACTTCAAAAAGCCTTTACAGATAAAGAGGAATTACTTA
15 AGCAGCAGGCATCTAACCTTCATGAGCAGAAAAAAGCTGGGGTGATATTTGAAGCTGAAGAAG
TTATCACTCTTTAACTTCTGTGCTTAAAACGTCATCTGCATCAAGAACTAGTTAAAGCTCACGA
CATCAGTTTGCTCCTGGAGCGACAGTATTGTATAAGGGCGATAAAATGGTGCTTAACCTGGACA
GGTCTCGTGTTCCAAGTGTATAGAGAAAATTGAGGCCATTCTTAAGGAACTTGAAAAGCC
AGCACCTGATGCGACCACGTTTTAGTCTACGTTTATCTGTCTTTACTTAATGTCCTTTGTTACAG
20 GCCAGAAAGCATAACTGGCCTGAATATTCTCTCTGGGCCACTGTTCCACTTGTATCGTCGGTCT
GATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACG
GTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCT
GATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCATGG
TCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTG
25 ATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGT
CCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGATCCCACTCGTGTTGTCTGGTCTGA
TTATCGGTCTGGGACCACGGTCCCACTTGTATTGTCGATCAGACTATCAGCGTGAGACTACGATT
CCATCAATGCCTGTCAAGGGCAAGTATTGACATGTGCTCGTAACCTGTAGAACGGAGTAACCTC
GGTGTGCGGTTGTATGCCTGCTGTGGATTGCTGCTGTGTCCTGCTTATCCACAACATTTTGC
30 CGGTTATGTGGACAAAATACCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCG
ATGCAAGTGTGTCGCTGTGACGAGCTCGCGAGCTCGGACATGAGGTTGCCCGTATTCAGTGT
CGCTGATTTGTATTGTCTGAAGTTGTTTTACGTTAAGTTGATGCAGATCAATTAATACGATACCT
GCGTCATAATTGATTATTTGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGATA
ATCATTATCACTTTACGGGTCCTTTCCGGTGATCCGACAGGTTACGGGGCGGCGACCTCGCGGGT
35 TTTCGCTATTTATGAAAATTTCCGGTTTAAGGCGTTTCCGTTCTTCTTCGTCATAACTTAATGTTT
TTATTTAAAATACCTCTGAAAAGAAAGGAAACGACAGGTGCTGAAAGCGAGCTTTTTGGCCTC
TGTCGTTTCTTTCTCTGTTTTGTCCGTGGAATGAACAATGGAAGTCCGAGCTCATCGCTAATA
ACTTCGTATAGCATACATTATACGAAGTTATATTCGATGCGGCCGCAAGGGGTTTCGCTCAGCG
GGTGTGGCGGGTGTCCGGGCTGGCTTAACATGCGGCATCAGAGCAGATTGTACTGAGAGTGC
40 ACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCCATT
GCCATTAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAG

CTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCAC
GACGTTGTAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGG
TACCCGGGGATCCTCGTTTAAAC

5

SEQ ID NO: 10 - Polynucleotide sequence encoding wild type ChAd155

CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGATGGGCGGCGCGGG
GCGGGAGGCGGGTCCGGGGGCGGGCCGGCGGGCGGGGCGGTGTGGCGGAAGTGGACTTTGTAA
GTGTGGCGGATGTGACTTGCTAGTGCCGGGCGCGGTAAAAGTGACGTTTTCCGTGCGCGACAAC
10 GCCCACGGGAAGTGACATTTTTCCGCGGTTTTTACCGGATGTTGTAGTGAATTTGGGCGTAACC
AAGTAAGATTTGGCCATTTTCGCGGGAAAACGAAACGGGGAAGTGAAATCTGATTAATTTTCG
GTTAGTCATACCGCGTAATATTTGTCGAGGGCCGAGGGACTTTGGCCGATTACGTGGAGGACTC
GCCCAGGTGTTTTTTGAGGTGAATTTCCGCGTTCCGGGTCAAAGTCTCCGTTTTATTATTATAGTC
AGCTGACGCGGAGTGATTTATACCTCTGATCTCGTCAAGTGGCCACTCTTGAGTGCCAGCGAG
15 TAGAGTTTTCTCCTCTGCCGCTCTCCGCTCCGCTCCGCTCGGCTCTGACACCGGGGAAAAAATGA
GACATTTACCTACGATGGCGGTGTGCTCACCGGCCAGCTGGCTGCTGAAGTCTGGACACCCT
GATCGAGGAGGTATTGGCCGATAATTATCCTCCCTCGACTCCTTTTGAGCCACCTACACTTCACG
AACTCTACGATCTGGATGTGGTGGGGCCAGCGATCCGAACGAGCAGGCGGTTTCCAGTTTTTTT
CCAGAGTCCATGTTGTTGGCCAGCCAGGAGGGGGTCGAACCTGAGACCCCTCCTCCGATCGTGG
20 ATTCCCCCGATCCGCCGAGCTGACTAGGCAGCCGAGCGCTGTGCGGGACCTGAGACTATGCC
CCAGCTGCTACCTGAGGTGATCGATCTCACCTGTAATGAGTCTGGTTTTCCACCCAGCGAGGATG
AGGACGAAGAGGGTGAGCAGTTTGTGTTAGATTCTGTGGAACAACCCGGGCGAGGATGCAGGT
CTTGTCAATATCACCGGAAAAACACAGGAGACTCCCAGATTATGTGTTCTCTGTGTTATATGAAG
ATGACCTGTATGTTATTTACAGTAAGTTTATCATCTGTGGGCAGGTGGGCTATAGTGTGGGTGG
25 TGGTCTTTGGGGGGTTTTTAATATATGTGTCAGGGGTTATGCTGAAGACTTTTTTATTGTGATTTTT
AAAGGTCCAGTGTCTGAGCCCGAGCAAGAACCTGAACCGGAGCCTGAGCCTTCTCGCCCCAGGA
GAAAGCCTGTAATCTTAAGTAGACCCAGCGCACCGGTAGCGAGAGGCCTCAGCAGCGCGGAGA
CCACCGACTCCGGTGCTTCCTCATACCCCCGAGATTACCCCCCTGGTGCCCCCTGTGTCCCGTT
AAGCCCGTTGCCGTGAGAGTCAGTGGGCGGCGGTCTGCTGTGGAGTGCATTGAGGACTTGCTTT
30 TTGATTCACAGGAACCTTTGGACTTGAGCTTGAAACGCCCCAGGCATTAAACCTGGTCACCTGG
ACTGAATGAGTTGACGCCTATGTTTGCTTTTGAATGACTTAATGTGTATAGATAATAAAGAGTGA
GATAATGTTTTAATTGCATGGTGTGTTAACTTGGGCGGAGTCTGCTGGGTATATAAGCTTCCCT
GGGCTAAACTTGGTTACACTTGACCTCATGGAGGCCTGGGAGTGTTTGGAGAACTTTGCCGGAG
TTCGTGCCCTTGCTGGACGAGAGCTCTAACAATACCTCTTGGTGGTGGAGGTATTTGTGGGGCTCT
35 CCCCAGGGCAAGTTAGTTTGTAGAATCAAGGAGGATTACAAGTGGGAATTTGAAGAGCTTTTGA
AATCCTGTGGTGAGCTATTGGATTCTTTGAATCTAGGCCACCAGGCTCTCTCCAGGAGAAGGTC
ATCAGGACTTTGGATTTTTCCACACCGGGGCGCATTGCAGCCGCGGTTGCTTTTCTAGCTTTTTTG
AAGGATAGATGGAGCGAAGAGACCCACTTGAGTTCGGGCTACGTCCTGGATTTTCTGGCCATGC
AACTGTGGAGAGCATGGATCAGACACAAGAACAGGCTGCAACTGTTGTCTTCCGTCCGCCCCGTT
40 GCTGATTCCGGCGGAGGAGCAACAGGCCGGGTGAGGAGCCGGGCCGTCGGGATCCGGAGGA
GAGGGCACCGAGGCCGGGCGAGAGGAGCGCGCTGAACCTGGGAACCGGGCTGAGCGGCCATCC

ACATCGGGAGTGAATGTCGGGCAGGTGGTGGATCTTTTCCAGAACTGCGGCGGATTTTACTA
TTAGGGAGGATGGGCAATTTGTTAAGGGTCTTAAGAGGGAGAGGGGGGCTTCTGAGCATAACG
AGGAGGCCAGTAATTTAGCTTTTAGCTTGATGACCAGACACCGTCCAGAGTGCATCACTTTTCAG
CAGATTAAGGACAATTGTGCCAATGAGTTGGATCTGTTGGGTCAGAAGTATAGCATAGAGCAGC
5 TGACCACTTACTGGCTGCAGCCGGGTGATGATCTGGAGGAAGCTATTAGGGTGTATGCTAAGGT
GGCCCTGCGGCCCGATTGCAAGTACAAGCTCAAGGGGCTGGTGAATATCAGGAATTGTTGCTAC
ATTTCTGGCAACGGGGCGGAGGTGGAGATAGAGACCGAAGACAGGGTGGCTTTCAGATGCAGC
ATGATGAATATGTGGCCGGGGGTGCTGGGCATGGACGGGGTGGTGATTATGAATGTGAGGTTCA
CGGGGCCCAACTTTAACGGCACGGTGTTTTTGGGGAACACCAACCTGGTCCTGCACGGGGTGAG
10 CTTCTATGGGTTTAACAACACCTGTGTGGAGGCCTGGACCGATGTGAAGGTCCGCGGTTGCGCCT
TTTATGGATGTTGGAAGGCCATAGTGAGCCGCCCTAAGAGCAGGAGTTCATTAAGAAATGCTT
GTTTGAGAGGTGCACCTTGGGGATCCTGGCCGAGGGCAACTGCAGGGTGCGCCACAATGTGGCC
TCCGAGTGCGGTTGCTTCATGCTAGTCAAGAGCGTGGCGGTAATCAAGCATAATATGGTGTGCG
GCAACAGCGAGGACAAGGCCTCACAGATGCTGACCTGCACGGATGGCAACTGCCACTTGCTGAA
15 GACCATCCATGTAACCAGCCACAGCCGGAAGGCCTGGCCCGTGTTCGAGCACAACCTTGCTGACC
CGCTGCTCCTTGATCTGGGCAACAGGCGGGGGGTGTTCTGCCCTATCAATGCAACTTTAGTCA
CACCAAGATCTTGCTAGAGCCCGAGAGCATGTCCAAGGTGAACTTGAACGGGGTGTGTTGACATG
ACCATGAAGATCTGGAAGGTGCTGAGGTACGACGAGACCAGGTCCCGGTGCAGACCCTGCGAG
TGCGGGGGCAAGCATATGAGGAACCAGCCCGTGATGCTGGATGTGACCGAGGAGCTGAGGACA
20 GACCACTTGGTTCTGGCCTGCACCAGGGCCGAGTTTGGTTCTAGCGATGAAGACACAGATTGAG
GTGGGTGAGTGGGCGTGGCCTGGGGTGGTCATGAAAATATATAAGTTGGGGGTCTTAGGGTCTC
TTTATTTGTGTTGCAGAGACCGCCGAGCCATGAGCGGGAGCAGCAGCAGCAGCAGTAGCAGCA
GCGCCTTGATGGCAGCATCGTGAGCCCTTATTTGACGACGCGGATGCCCCACTGGGCCGGGGT
GCGTCAGAATGTGATGGGCTCCAGCATCGACGGCCGACCCGTCCTGCCCCGAAATTCGCCACG
25 CTGACCTATGCGACCGTCGCGGGGACGCCGTTGGACGCCACCGCCGCCGCCGCCGCCACCGCAG
CCGCCTCGGCCGTGCGCAGCCTGGCCACGGACTTTGCATTCTGGGACCACTGGCGACAGGGGC
TACTTCTCGGGCCGCTGCTGCCGCCGTTTCGCGATGACAAGCTGACCGCCCTGCTGGCGCAGTTGG
ATGCGCTTACTCGGGAACCTGGGTGACCTTCTCAGCAGGTCATGGCCCTGCGCCAGCAGGTCTCC
TCCCTGCAAGCTGGCGGGAATGCTTCTCCACAAATGCCGTTTAAGATAAATAAAACCAGACTC
30 TGTTTGGATTAAAGAAAAGTAGCAAGTGCATTGCTCTCTTTATTTTCATAATTTTCCGCGCGCGAT
AGGCCCTAGACCAGCGTTCTCGGTGCTTGAGGGTGCGGTGTATCTTCTCCAGGACGTGGTAGAG
GTGGCTCTGGACGTTGAGATACATGGGCATGAGCCCGTCCCGGGGGTGGAGGTAGCACCCTGC
AGAGCTTCATGCTCCGGGGTGGTGTGTAGATGATCCAGTCGTAGCAGGAGCGCTGGGCATGGT
GCCTAAAAATGTCCTTCAGCAGCAGGCCGATGGCCAGGGGGAGGCCCTTGGTGTAAGTGTTTAC
35 AAAACGGTTAAGTTGGGAAGGGTGCATTGCGGGAGAGATGATGTGCATCTTGGACTGTATTTT
AGATTGGCGATGTTTCCGCCAGATCCCTTCTGGGATTCATGTTGTGCAGGACCACCAGTACAGT
GTATCCGGTGCACCTTGGGGAATTTGTCATGCAGCTTAGAGGGAAAAGCGTGGAAGAACTTGGAG
ACGCCTTTGTGGCCTCCCAGATTTTCCATGCATTTCGTCCATGATGATGGCAATGGGCCCGCGGGA
GGCAGCTTGGGCAAAGATATTTCTGGGGTTCGTGACGTCGTAGTTGTGTTCCAGGGTGAGGTG
40 TCATAGGCCATTTTACAAAGCGCGGGCGGAGGGTGCCCCACTGGGGGATGATGGTCCCCTCTG
GCCCTGGGGCGTAGTTGCCCTCGCAGATCTGCATTTCCAGGCCTTAATCTCGGAGGGGGGAAT

CATATCCACCTGCGGGGCGATGAAGAAAACGGTTTCCGGAGCCGGGAGATTAACCTGGGATGA
GAGCAGGTTTCTAAGCAGCTGTGATTTTCCACAACCGGTGGGCCCATAAATAACACCTATAACC
GGTTGCAGCTGGTAGTTTAGAGAGCTGCAGCTGCCGTCGTCCCGGAGGAGGGGGGCCACCTCGT
TGAGCATGTCCCTGACGCGCATGTTCTCCCCGACCAGATCCGCCAGAAGGCGCTCGCCGCCAG
5 GGACAGCAGCTCTTGCAAGGAAGCAAAGTTTTTCAGCGGCTTGAGGCCGTCCGCCGTGGGCATG
TTTTTCAGGGTCTGGCTCAGCAGCTCCAGGCGGTCCAGAGCTCGGTGACGTGCTCTACGGCATC
TCTATCCAGCATATCTCCTCGTTTCGCGGGTTGGGGCGACTTTCGCTGTAGGGCACCAAGCGGTG
GTCGTCCAGCGGGGCCAGAGTCATGTCCTTCCATGGGCGCAGGGTCCTCGTCAGGGTGGTCTGG
GTCACGGTGAAGGGGTGCGCTCCGGGCTGAGCGCTTGCCAAGGTGCGCTTGAGGCTGGTTCTGC
10 TGGTGCTGAAGCGCTGCCGGTCTTCGCCCTGCGCGTCGGCCAGGTAGCATTTGACCATGGTGTCA
TAGTCCAGCCCCCTCCGCGGCGTGTCCCTTGCGCGCAGCTTGCCCTTGAGGTTGGCGCCGCACG
AGGGGCAGAGCAGGCTCTTGAGCGCGTAGAGCTTGGGGGCGAGGAAGACCGATTTCGGGGGAGT
AGGCGTCCGCGCCGCAGACCCCGCACACGGTCTCGCACTCCACCAGCCAGGTGAGCTCGGGGCG
CGCCGGGTCAAAAACAGGTTTCCCCCATGCTTTTTGATGCGTTTCTTACCTCGGGTCTCCATGA
15 GGTGGTGTCCCCGCTCGGTGACGAAGAGGCTGTCCGTGTCTCCGTAGACCGACTTGAGGGGTCT
TTTCTCCAGGGGGGTCCCTCGGTCTTCCTCGTAGAGGAACTCGGACCACTCTGAGACGAAGGCC
CGCGTCCAGGCCAGGACGAAGGAGGCTATGTGGGAGGGGTAGCGGTCTGTCCACTAGGGGG
TCCACCTTCTCCAAGGTGTGAAGACACATGTCGCCTTCCTCGGCGTCCAGGAAGGTGATTGGCTT
GTAGGTGTAGGCCACGTGACCGGGGGTTCCTGACGGGGGGGTATAAAAGGGGGTGGGGGCGCG
20 CTCGTGCTCACTCTCTTCCGCATCGTGTCTGCGAGGGGCCAGCTGCTGGGGTGAGTATTCCTCT
CGAAGGCGGGCATGACCTCCGCGCTGAGGTTGTGAGTTTCCAAAACGAGGAGGATTTGATGTT
CACCTGTCCCGAGGTGATACCTTTGAGGGTACCCGCGTCCATCTGGTCAGAAAACACGATCTTTT
TATTGTCCAGCTTGGTGGCGAACGACCCGTAGAGGGCGTTGGAGAGCAGCTTGCGGATGGAGCG
CAGGGTCTGGTCTTGTCCCTGTCCGCGCGCTCCTTGGCCGCGATGTTGAGCTGCACGTACTCGC
25 GCGCGACGCAGCGCCACTCGGGGAAGACGGTGGTGCCTCGTCGGGCACCAGGCGCACGCGCC
AGCCGCGGTTGTGCAGGGTGACCAGGTCCACGCTGGTGGCGACCTCGCCGCGCAGGCGCTCGTT
GGTCCAGCAGAGACGGCCGCCCTTGCGCGAGCAGAAGGGGGGCAGGGGGTCGAGCTGGGTCTC
GTCCGGGGGGTCCGCGTCCACGGTGAAAACCCCGGGGCGCAGGCGCGCGCTCGAAGTAGTCTATC
TTGCAACCTTGCATGTCCAGCGCCTGCTGCCAGTCGCGGGCGGCGAGCGCGCGCTCGTAGGGGT
30 TGAGCGGCGGGCCCCAGGGCATGGGGTGGGTGAGTGCAGGAGCGTACATGCCGCAGATGTCAT
AGACGTAGAGGGGCTCCCGCAGGACCCCGATGTAGGTGGGGTAGCAGCGGCCGCCGCGATGC
TGGCGCGCACGTAGTCATACAGCTCGTGCGAGGGGGCGAGGAGGTGCGGGGCCAGGTTGGTGC
GGGCGGGGCGCTCCGCGCGGAAGACGATCTGCCTGAAGATGGCATGCGAGTTGGAAGAGATGG
TGGGGCGCTGGAAGACGTTGAAGCTGGCGTCTGCAGGCCGACGGCGTCGCGCACGAAGGAGG
35 CGTAGGAGTCGCGCAGCTTGTGTACCAGCTCGGCGGTGACCTGCACGTGAGCGCGCAGTAGTC
GAGGGTCTCGCGGATGATGTCATATTTAGCCTGCCCTTCTTTTTCCACAGCTCGCGGTTGAGGA
CAAACCTCTTCGCGTCTTTCCAGTACTCTTGATCGGGAAACCGTCCGGTTCGGAACGGTAAGAG
CCTAGCATGTAGAACTGGTTGACGGCCTGGTAGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGT
AGGCCTGCGCGGCCTTGCGGAGCGAGGTGTGGGTGAGGGCGAAGGTGTCCCTGACCATGACTTT
40 GAGGTACTGGTGCTTGAAGTCGGAGTCGTGCGAGCCGCCCCGCTCCAGAGCGAGAAGTCGGTG
CGCTTCTTGAGCGGGGGTTGGGCAGAGCGAAGGTGACATCGTTGAAGAGGATTTTGCCCGCGC

GGGGCATGAAGTTGCGGGTGATGCGGAAGGGCCCCGGCACTTCAGAGCGGTTGTTGATGACCTG
GGCGGCGAGCACGATCTCGTCGAAGCCGTTGATGTTGTGGCCACGATGTAGAGTTCCAGGAAG
CGGGGCCGGCCCTTTACGGTGGGCAGCTTCTTTAGCTCTTCGTAGGTGAGCTCCTCGGGCGAGGC
GAGGCCGTGCTCGGCCAGGGCCCAGTCCGCGAGGTGCGGGTTGTCTCTGAGGAAGGACTTCCAG
5 AGGTCGCGGGCCAGGAGGGTCTGCAGGCGGTCTCTGAAGGTCTGAACTGGCGGGCCACGGCCA
TTTTTTCGGGGGTGATGCAGTAGAAGGTGAGGGGGTCTTGCTGCCAGCGGTCCCAGTCGAGCTG
CAGGGCGAGGTCGCGCGCGGCGGTGACCAGGCGCTCGTCGCCCCGAATTTTCATGACCAGCATG
AAGGGCACGAGCTGCTTTCCGAAGGCCCCCATCCAAGTGTAGGTCTCTACATCGTAGGTGACAA
AGAGGCGTCCGTGCGAGGATGCGAGCCGATCGGAAGAAGTGGATCTCCCGCCACCAGTTGG
10 AGGAGTGGCTGTTGATGTGGTGAAGTAGAAGTCCCGTCGCCGGGCCGAACACTCGTGCTGGCT
TTTGTAAGCGAGCGCAGTACTGGCAGCGCTGCACGGGCTGTACCTCATGCACGAGATGCACC
TTTCGCCCCGCGCACGAGGAAGCCGAGGGGAAATCTGAGCCCCCGCTGGCTCGCGGCATGGCT
GGTTCTCTTCTACTTTGGATGCGTGTCCGTCTCCGTCTGGCTCCTCGAGGGGTGTTACGGTGGAG
CGGACCACCACGCCGCGGAGCCGAGGTCCAGATATCGGCGCGCGGCGGTGCGAGTTTGATGA
15 CGACATCGCGCAGCTGGGAGCTGTCCATGGTCTGGAGCTCCCGCGGCGGCGGCAGGTCAGCCGG
GAGTTCTTGACAGGTTACCTCGCAGAGTCGGGCCAGGGCGCGGGGCAGGTCTAGGTGGTACCTG
ATCTCTAGGGGCGTGTGGTGGCGGCGTCGATGGCTTGACAGAGCCCGCAGCCCCGGGGGCGA
CGACGGTGCCCCGCGGGGTGGTGGTGGTGGTGGCGGTGCAGCTCAGAAGCGGTGCCGCGGGCG
GGCCCCCGGAGGTAGGGGGGGCTCCGGTCCCGCGGGCAGGGGCGGCAGCGGCACGTGCGCGTG
20 GAGCGCGGGCAGGAGTTGGTGCTGTGCCCGGAGGTTGCTGGCGAAGGCGACGACGCGGCGGTT
GATCTCCTGGATCTGGCGCCTCTGCGTGAAGACGACGGGCCCGGTGAGCTTGAACCTGAAAGAG
AGTTCGACAGAATCAATCTCGGTGTCATTGACCGCGGCTGGCGCAGGATCTCCTGCACGTCTCC
CGAGTTGTCTTGGTAGGCGATCTCGGCCATGAACTGCTCGATCTCTTCTCCTGGAGGTCTCCGC
GTCCGGCGCGTTCCACGGTGGCCGCCAGGTGCTTGGAGATGCGCCCCATGAGCTGCGAGAAGGC
25 GTTGAGTCCGCCCTCGTTCCAGACTCGGCTGTAGACCACGCCCCCTGGTCATCGCGGGCGCGCA
TGACCACCTGCGCGAGGTTGAGCTCCACGTGCCGCGGAAGACGGCGTAGTTGCGCAGACGCTG
GAAGAGGTAGTTGAGGGTGGTGGCGGTGTGCTCGGCCACGAAGAAGTTCATGACCCAGCGGCG
CAACGTGGATTGTTGATGTCCCCCAAGGCCTCCAGCCGTTCCATGGCCTCGTAGAAGTCCACGG
CGAAGTTGAAAACTGGGAGTTGCGCGCCGACACGGTCAACTCCTCCTCCAGAAGACGGATGAG
30 CTCGGCGACGGTGTGCGGCACCTCGCGCTCGAAGGCTATGGGGATCTCTTCTCCTCCGCTAGCATCA
CCACCTCCTCCTCTTCTCCTCTTCTGGCACTTCCATGATGGCTTCTCCTCTTTCGGGGGGTGGCG
GCGGCGGCGGTGGGGGAGGGGGCGCTCTGCGCCGGCGGCGGCGCACCGGGAGGCGGTCCACGA
AGCGCGGATCATCTCCCCGCGGCGGCGGCGCATGGTCTCGGTGACGGCGCGGCCGTTCTCCCG
GGGGCGCAGTTGGAAGACGCCGCCGACATCTGGTGCTGGGGCGGGTGGCCGTGAGGCAGCGA
35 GACGGCGCTGACGATGCATCTCAACAATTGCTGCGTAGGTACGCCGCCGAGGGACCTGAGGGAG
TCCATATCCACCGGATCCGAAAACCTTTCGAGGAAGGCGTCTAACCAGTCGCAGTCGCAAGGTA
GGCTGAGCACCGTGGCGGGCGGCGGGGGGTGGGGGGAGTGTCTGGCGGAGGTGCTGCTGATGA
TGTAATTGAAGTAGGCGGACTTGACACGGCGGATGGTCGACAGGAGCACCATGTCTTGGGTCC
GGCCTGCTGGATGCGGAGGCGGTGCGCTATGCCCCAGGCTTCGTTCTGGCATCGGCGCAGGTCC
40 TTGTAGTAGTCTTGATGAGCCTTCCACCGGCACCTCTTCTCCTTCTCTTCTGCTTCTTCCATGT
CTGCTTCGGCCCTGGGGCGGCGCCGCGCCCCCTGCCCCCATGCGCGTGACCCCGAACCCCTG

AGCGGTTGGAGCAGGGCCAGGTCGGCGACGACGCGCTCGGCCAGGATGGCCTGCTGCACCTGC
GTGAGGGTGGTTTGGAAAGTCATCCAAGTCCACGAAGCGGTGGTAGGCGCCCGTGTGATGGTGT
AGGTGCAGTTGGCCATGACGGACCAGTTGACGGTCTGGTGGCCCGGTTGCGACATCTCGGTGTA
CCTGAGTCGCGAGTAGGCGCGGGAGTCGAAGACGTAGTCGTTGCAAGTCCGCACCAGGTACTGG
5 TAGCCCACCAGGAAGTGCGGCGGGGCTGGCGGTAGAGGGGCCAGCGCAGGGTGGCGGGGGCT
CCGGGGGCCAGGTCTTCCAGCATGAGGCGGTGGTAGGCGTAGATGTACCTGGACATCCAGGTGA
TACCCGCGGCGGTGGTGGAGGCGCGCGGGAAGTCGCGCACCCGGTTCCAGATGTTGCGCAGGG
GCAGAAAGTGCTCCATGGTAGGCGTGCTCTGTCCAGTCAGACGCGCGCAGTCGTTGATACTCTA
GACCAGGGAAAACGAAAGCCGGTCAGCGGGCACTCTTCCGTGGTCTGGTGAATAGATCGCAAG
10 GGTATCATGGCGGAGGGCCTCGGTTTCGAGCCCCGGTCCGGGCCGGACGGTCCGCCATGATCCA
CGCGGTTACCGCCCGCGTGTCTGAACCCAGGTGTGCGACGTCAGACAACGGTGGAGTGTTCCTTT
TGCGGTTTTTCTGGCCGGGCGCCGGCGCCGCGTAAGAGACTAAGCCGCGAAAGCGAAAGCAGTA
AGTGGCTCGCTCCCCGTAGCCGGAGGGATCCTTGCTAAGGGTTGCGTTGCGGCGAACCCCGGTT
CGAATCCCGTACTCGGGCCGGCGGACCCGCGGCTAAGGTGTTGGATTGGCCTCCCCCTCGTAT
15 AAAGACCCCGCTTGCGGATTGACTCCGGACACGGGGACGAGCCCCCTTTATTTTTGCTTTCCCCA
GATGCATCCGGTGCTGCGGCAGATGCGCCCCCGCCCCAGCAGCAGCAACAACACCAGCAAGA
GCGGCAGCAACAGCAGCGGGAGTCATGCAGGGCCCCCTACCCACCCTCGGCGGGCCGGCCAC
CTCGGCGTCCGCGGCCGTGTCTGGCGCCTGCGGCGGCGGCGGGGGCCGGCTGACGACCCCGAG
GAGCCCCCGCGGCGCAGGGCCAGACACTACCTGGACCTGGAGGAGGGCGAGGGCCTGGCGCGG
20 CTGGGGGCGCCGTCTCCCGAGCGCCACCCGCGGGTGACGCTGAAGCGCGACTCGCGCGAGGCGT
ACGTGCCTCGGCAGAACCTGTTTCAGGGACCGCGCGGGCGAGGAGCCCGAGGAGATGCGGGACA
GGAGGTTTCAGCGCAGGGCGGGAGCTGCGGCAGGGGCTGAACCGCGAGCGGCTGCTGCGCGAGG
AGGACTTTGAGCCCGACGCGCGGACGGGGATCAGCCCCGCGCGCGCGCACGTGGCGGCCGCCG
ACCTGGTGACGGCGTACGAGCAGACGGTGAACCAGGAGATCAACTTCCAAAAGAGTTTCAACA
25 ACCACGTGCGCACGCTGGTGGCGCGCGAGGAGGTGACCATCGGGCTGATGCACCTGTGGGACTT
TGTAAGCGCGCTGGTGCAGAACCCCAACAGCAAGCCTCTGACGGCGCAGCTGTTCTGATAGTG
CAGCACAGCAGGGACAACGAGGCGTTTAGGGACGCGCTGCTGAACATCACCGAGCCCGAGGGT
CGGTGGCTGCTGGACCTGATTAACATCCTGCAGAGCATAGTGGTGCAGGAGCGCAGCCTGAGCC
TGGCCGACAAGGTGGCGGCCATCAACTACTCGATGCTGAGCCTGGGCAAGTTTTACGCGCGCAA
30 GATCTACCAGACGCCGTACGTGCCCATAGACAAGGAGGTGAAGATCGACGGTTTTTACATGCGC
ATGGCGCTGAAGGTGCTCACCTGAGCGACGACCTGGGCGTGTACCGCAACGAGCGCATCCACA
AGGCCGTGAGCGTGAGCCGGCGGCGCGAGCTGAGCGACCCGCGAGCTGATGCACAGCCTGCAGC
GGGCGCTGGCGGGCGCCGGCAGCGGCGACAGGGAGGCGGAGTCCTACTTCGATGCGGGGGCGG
ACCTGCGCTGGGCGCCCAGCCGGCGGGCCCTGGAGGCCGCGGGGTCCGCGAGGACTATGACG
35 AGGACGGCGAGGAGGATGAGGAGTACGAGCTAGAGGAGGGCGAGTACCTGGACTAAACCGCG
GGTGGTGTTCGCGTAGATGCAAGACCCGAACGTGGTGGACCCGGCGCTGCGGGCGGCTCTGCA
GAGCCAGCCGTCCGGCCTTAACCTCAGACGACTGGCGACAGGTCATGGACCGCATCATGTGCG
CTGACGGCGCGTAACCCGGACGCGTTCCGGCAGCAGCCGCGAGGCCAACAGGCTCTCCGCCATCC
TGGAGGCGGTGGTGCCTGCGCGCTCGAACCCACGCACGAGAAGGTGCTGGCCATAGTGAACGC
40 GCTGGCCGAGAACAGGGCCATCCGCCCGGACGAGGCCGGGCTGGTGTACGACGCGCTGCTGCA
GCGCGTGGCCCGCTACAACAGCGGCAACGTGCAGACCAACCTGGACCGGCTGGTGGGGGACGT

GCGCGAGGCGGTGGCGCAGCGCGAGCGCGGATCGGCAGGGCAACCTGGGCTCCATGGTGGC
GCTGAATGCCTTCTGAGCACGCAGCCGGCCAACGTGCCGCGGGGGCAGGAAGACTACACCAA
CTTTGTGAGCGCGCTGCGGCTGATGGTGACCGAGACCCCCAGAGCGAGGTGTACCAGTCGGGC
CCGGACTIONTTCTTCCAGACCAGCAGACAGGGCCTGCAGACGGTGAACCTGAGCCAGGCTTTCA
5 AGAACCTGCGGGGGCTGTGGGGCGTGAAGGCGCCCCACCGGCGACCGGGCGACGGTGTCCAGCC
TGCTGACGCCAACTCGCGCCTGCTGCTGCTGCTGATCGCGCCGTTACGGACAGCGGCAGCGT
GTCCCGGGACACCTACCTGGGGCACCTGCTGACCCTGTACCGCGAGGCCATCGGGCAGGCGCAG
GTGGACGAGCACACCTTCCAGGAGATCACCAGCGTGAGCCGCGCGCTGGGGCAGGAGGACACG
AGCAGCCTGGAGGCGACTCTGAACCTACCTGCTGACCAACCGGCGGCAGAAGATTCCCTCGCTGC
10 ACAGCCTGACCTCCGAGGAGGAGCGCATCTTGCGCTACGTGCAGCAGAGCGTGAGCCTGAACCT
GATGCGCGACGGGGTGACGCCAGCGTGCGCTGGACATGACCGCGCGCAACATGGAACCGGG
CATGTACGCCGCGCACCGGCCTTACATCAACCGCCTGATGGACTACCTGCATCGCGCGGCGGCC
GTGAACCCCGAGTACTTTACCAACGCCATCCTGAACCCGCACTGGCTCCCGCCCGCCGGGTTCTA
CAGCGGGGGCTTCGAGGTCCCGGAGACCAACGATGGCTTCTGTGGGACGACATGGACGACAG
15 CGTGTCTCTCCCGCGGCCCGCAGGCGCTGGCGGAAGCGTCCCTGCTGCGTCCCAAGAAGGAGGAG
GAGGAGGAGGCGAGTCGCCGCCGCGGCAGCAGCGGCGTGGCTTCTGTCCGAGCTGGGGGCG
GCAGCCGCCGCGCGCCCCGGGTCCCTGGGCGGCAGCCCCCTTCCGAGCCTGGTGGGGTCTCTGC
ACAGCGAGCGCACCAACCGCCCTCGGCTGCTGGGCGAGGACGAGTACCTGAATAACTCCCTGCT
GCAGCCGGTGCGGGAGAAAAACCTGCCTCCCGCCTTCCCCAACAACGGGATAGAGAGCCTGGTG
20 GACAAGATGAGCAGATGGAAGACCTATGCGCAGGAGCACAGGGACGCGCCTGCGCTCCGGCCG
CCCACGCGGCGCCAGCGCCACGACCGGCAGCGGGGGCTGGTGTGGGATGACGAGGACTCCGCG
GACGATAGCAGCGTGCTGGACCTGGGAGGGAGCGGCAACCCGTTGCGGCACCTGCGCCCCCGCC
TGGGGAGGATGTTTTAAAAAAGCAAGAAGCATGATGCAAAAATTAATAAAAA
CTACCAAGGCCATGGCGACCGAGCGTTGGTTTCTTGTGTTCCCTTCAGTATGCGGCGCGCGGCG
25 ATGTACCAGGAGGGACCTCCTCCCTCTTACGAGAGCGTGGTGGGCGCGGCGGCGGCGGCCCT
CTTCTCCCTTTGCGTCGCAGCTGCTGGAGCCGCCGTACGTGCCTCCGCGCTACCTGCGGCCTACG
GGGGGGAGAAACAGCATCCGTTACTCGGAGCTGGCGCCCCTGTTGACACCACCGGGTGTACC
TGGTGGACAACAAGTCGGCGGACGTGGCCTCCCTGAACTACCAGAACGACCACAGCAATTTTTT
GACCACGGTCATCCAGAACAATGACTACAGCCGAGCGAGGCCAGCACCAGACCATCAATCTG
30 GATGACCGGTGCGACTGGGGCGGCGACCTGAAAACCATCCTGCACACCAACATGCCAACGTGA
ACGAGTTCATGTTACCAATAAGTTCAAGGCGCGGGTGATGGTGTGCGGCTCGCACACCAAGGA
AGACCGGGTGGAGCTGAAGTACGAGTGGGTGGAGTTCGAGCTGCCAGAGGGCAACTACTCCGA
GACCATGACCATTGACCTGATGAACAACGCGATCGTGAGGACTATCTGAAAGTGGGCAGGCAG
AACGGGGTCCCTGGAGAGCGACATCGGGGTCAAGTTCGACACCAGGAACTTCCGCCTGGGGCTGG
35 ACCCGTGACCGGGCTGGTTATGCCCGGGGTGTACACCAACGAGGCCTTCCATCCCGACATCAT
CCTGCTGCCCCGCTGCGGGGTGGACTTCACTTACAGCCGCCTGAGCAACCTCCTGGGCATCCGC
AAGCGGCAGCCCTTCCAGGAGGGCTTCAGGATCACCTACGAGGACCTGGAGGGGGGCAACATC
CCCGCGCTCCTCGATGTGGAGGCCTACCAGGATAGCTTGAAGGAAAATGAGGCGGGACAGGAG
GATACCGCCCCCGCCCTCCGCCGCCGCGAGCAGGGCGAGGATGCTGCTGACACCGCGGCCG
40 CGGACGGGGCAGAGGCCGACCCGCTATGGTGGTGGAGGCTCCCGAGCAGGAGGAGGACATGA
ATGACAGTGCGGTGCGCGGAGACACCTTCGTCACCCGGGGGAGGAAAAGCAAGCGGAGGCCG

AGGCCGCGGCCGAGGAAAAGCAACTGGCGGCAGCAGCGCGGCCGCGCGCTTGGCCGCGGCGG
AGGCTGAGTCTGAGGGGACCAAGCCCGCCAAGGAGCCCGTGATTAAGCCCCTGACCGAAGATA
GCAAGAAGCGCAGTTACAACCTGCTCAAGGACAGCACCAACACCGCGTACCGCAGCTGGTACCT
GGCCTACAACCTACGGCGACCCGTCGACGGGGGTGCGCTCCTGGACCCTGCTGTGCACGCCGGAC
5 GTGACCTGCGGCTCGGAGCAGGTGTACTGGTCGCTGCCCGACATGATGCAAGACCCCGTGACCT
TCCGCTCCACGCGGCAGGTCAGCAACTTCCCGGTGGTGGGCGCCGAGCTGCTGCCCGTGCACTC
CAAGAGCTTCTACAACGACCAGGCCGTCTACTCCCAGCTCATCCGCCAGTTCACCTCTCTGACCC
ACGTGTTCAATCGCTTTTCTGAGAACCAGATTCTGGCGCGCCCGCCCGCCCCACCATCACCACC
GTCAGTGAAAACGTTCTGCTCTCACAGATCACGGGACGCTACCGCTGCGCAACAGCATCGGAG
10 GAGTCCAGCGAGTGACCGTTACTGACGCCAGACGCCGCACCTGCCCTACGTTTACAAGGCCTT
GGGCATAGTCTCGCCGCGCGTCTTTCCAGCCGCACTTTTTGAGCAACACCACCATCATGTCCAT
CCTGATCTCACCCAGCAATAACTCCGGCTGGGGACTGCTGCGCGCGCCAGCAAGATGTTCCGA
GGGGCGAGGAAGCGTTCCGAGCAGCACCCCGTGCGCGTGCGCGGGCACTTCCGCGCCCCCTGGG
GAGCGCACAAACGCGGCCGCGCGGGGCGCACCAACCGTGACGACGCCATCGACTCGGTGGTGG
15 AGCAGGCGCGCAACTACAGGCCCGCGGTCTCTACCGTGACGCGGCCATCCAGACCGTGGTGCG
GGGCGCGCGGCGGTACGCCAAGCTGAAGAGCCCGCGGAAGCGCGTGCCCCGCGCCACCGCCG
CCGACCCGGGGCCCGCCGCAAAACGCGCCGCGCGGCCCTGCTTCGCCGGGCCAAGCGCACGGGC
CGCCGCGCCGCCATGAGGGCCGCGCGCCGCTTGCCGCGCGCATCACCGCCGCCACCATGGCCC
CCCGTACCCGAAGACGCGCGGCCCGCCCGCCGCCGCCGCCATCAGTGACATGGCCAGCAGGCG
20 CCGGGGCAACGTGTACTGGGTGCGCGACTCGGTGACCGGCACGCGCGTGCCCGTGCGCTTCCGC
CCCCCGCGGACTTGAGATGATGTGAAAAACAACACTGAGTCTCCTGCTGTTGTGTGTATCCCA
GCGGCGGCGGCGCGCGCAGCGTCATGTCCAAGCGCAAAATCAAAGAAGAGATGCTCCAGGTGCG
TCGCGCCGGAGATCTATGGGCCCCGAAGAAGGAAGAGCAGGATTCGAAGCCCCGCAAGATAA
AGCGGGTCAAAAAGAAAAAGAAAGATGATGACGATGCCGATGGGGAGGTGGAGTTCCTGCGCG
25 CCACGGCGCCAGGCGCCCGGTGCAGTGGAAGGGCCGGCGGTAAAGCGCGTCTGCGCCCCG
GCACCGCGGTGGTCTTACGCCCCGCGAGCGCTCCACCCGGACTTTCAAGCGCGTCTATGACGA
GGTGTACGGCGACGAAGACCTGCTGGAGCAGGCCAACGAGCGCTTCGGAGAGTTTGCTTACGGG
AAGCGTACGCGGGCGCTGGGGAAGGAGGACCTGCTGGCGCTGCCGCTGGACCAGGGCAACCCC
ACCCCCAGTCTGAAGCCCGTGACCTGCAGCAGGTGCTGCCGAGCAGCGCACCCCTCCGAGGCGA
30 AGCGGGGTCTGAAGCGCGAGGGCGGCGACCTGGCGCCACCGTGACGCTCATGGTGCCCAAGC
GGCAGAGGCTGGAGGATGTGCTGGAGAAAATGAAAGTAGACCCCGGTCTGCAGCCGGACATCA
GGGTCCGCCCCATCAAGCAGGTGGCGCCGGGCCCTCGGCGTGACAGCCGTGGACGTGGTTCATCCC
CACCGGCAACTCCCCCGCCGCCGCCACCCTACCGCTGCCTCCACGGACATGGAGACACAGACC
GATCCC GCCGAGCCGACGCCGAGCCGCCGCCGCCGCGACCTCCTCGGCGAGGTGCAGACGGACC
35 CCTGGCTGCCGCCGGCGATGTCAGTCCCCGCGCGCGTCGCGGGCGCAGGAAGTACGGCGCCGC
CAACGCGCTCCTGCCCCGAGTACGCTTGCATCCTTCCATCGCGCCACCCCGGCTACCGAGGCT
ATACCTACCGCCCGCAAGAGCCAAGGGTTCCACCCGCCGTCCCCGCCGACGCGCGCCGCCAC
CACCCGCCGCCGCCGCCGAGACGCCAGCCGCACTGGCTCCAGTCTCCGTGAGGAAAGTGGCG
CGCGACGGACACACCCTGGTGCTGCCAGGGCGCGCTACCACCCAGCATCGTTTAAAAGCCTG
40 TTGTGGTCTTGCAGATATGGCCCTCACTTGCCGCTCCGTTTCCCGGTGCCGGGATACCGAGGA
GGAAGATCGCGCCGAGGAGGGGTCTGGCCGGCCGCGGCCTGAGCGGAGGCAGCCGCCGCGCG

CACCGGCGGCGACGCGCCACCAGCCGACGCATGCGCGGCGGGGTGCTGCCCCTGTTAATCCCC
TGATCGCCGCGGCGATCGGCGCCGTGCCCAGGATCGCCTCCGTGGCCTTGCAAGCGTCCCAGAG
GCATTGACAGACTTGCAAACCTTGCAAATATGGAAAAAAACCCCAATAAAAAAGTCTAGACTC
TCACGCTCGCTTGGTCCTGTGACTATTTTGTAGAATGGAAGACATCAACTTTGCGTCGCTGGCCC
5 CGCGTCACGGCTCGCGCCCCGTTCTTGGGACACTGGAACGATATCGGCACCAGCAACATGAGCGG
TGGCGCCTTCAGTTGGGGCTCTCTGTGGAGCGGCATTAAAAAGTATCGGGTCTGCCGTTAAAAATT
ACGGCTCCCCGGGCCTGGAACAGCAGCACGGGCCAGATGTTGAGAGACAAGTTGAAAGAGCAGA
ACTTCCAGCAGAAGGTGGTGGAGGGCCTGGCCTCCGGCATCAACGGGGTGGTGGACCTGGCCAA
CCAGGCCGTGCAGAATAAGATCAACAGCAGACTGGACCCCCGGCCGCCGGTGGAGGAGGTGCC
10 GCCGGCGCTGGAGACGGTGTCCCCCGATGGGCGTGGCGAGAAGCGCCCCGCGGCCCGATAGGGA
AGAGACCACTCTGGTCACGCAGACCGATGAGCCGCCCCCGTATGAGGAGGCCCTGAAGCAAGG
TCTGCCCACCACGCGGCCCATCGCGCCCATGGCCACCGGGTGGTGGGCCGCCACACCCCCGCC
ACGCTGGACTTGCTCCGCCCCGCGATGTGCCGCAGCAGCAGAAGGCGGCACAGCCGGGCCCGC
CCGCGACCGCCTCCCGTTCCCTCCGCGGTCTCTGCGCCGCGCGGCCAGCGCCCCCGCGGGGG
15 GGTGCGGAGGCACGGCAACTGGCAGAGCACGCTGAACAGCATCGTGGGTCTGGGGGTGCGGTC
CGTGAAGCGCCGCCGATGCTACTGAATAGCTTAGCTAACGTGTTGTATGTGTGTATGCGCCCTAT
GTCGCCGCCAGAGGAGCTGCTGAGTCGCCGCCGTTTCGCGCGCCACCACCACCGCCACTCCGCC
CCTCAAGATGGCGACCCCATCGATGATGCCGCAGTGGTCGTACATGCACATCTCGGGCCAGGAC
GCCTCGGAGTACCTGAGCCCCGGGCTGGTGCAGTTCGCCCCGCGCCACCGAGAGCTACTTCAGCC
20 TGAGTAACAAGTTTAGGAACCCACCGTGGCGCCACGCACGATGTGACCACCGACCGGTCTCA
GCGCCTGACGCTGCGGTTTATTCCCGTGGACCGCGAGGACACCGCGTACTCGTACAAGGCGCGG
TTCACCCTGGCCGTGGGCGACAACCGCGTGTGGACATGGCCTCCACCTACTTTGACATCCGCGG
GGTGCTGGACCGGGGTCCCACTTTCAAGCCCTACTCTGGCACCGCCTACAACCTCCCTGGCCCCCA
AGGGCGCTCCCAACTCCTGCGAGTGGGAGCAAGAGGAAACTCAGGCAGTTGAAGAAGCAGCAG
25 AAGAGGAAGAAGAAGATGCTGACGGTCAAGCTGAGGAAGAGCAAGCAGCTACCAAAAAGACT
CATGTATATGCTCAGGCTCCCCCTTCTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAA
CGGACGCTACAGCTACAGAACAAAAACCTATTTATGCAGACCCTACATTCCAGCCCGAACCCCA
AATCGGGGAGTCCAGTGAATGAGGCAGATGCTACAGTCGCCGCGGTAGAGTGCTAAAGAA
ATCTACTCCCATGAAACCATGCTATGGTTCCTATGCAAGACCCACAAATGCTAATGGAGGTCAG
30 GGTGTACTAACGGCAAATGCCCAGGGACAGCTAGAATCTCAGGTTGAAATGCAATTCTTTTCAA
CTTCTGAAAACGCCCGTAACGAGGCTAACAACATTCAGCCCAAATTGGTGCTGTATAGTGAGGA
TGTGCACATGGAGACCCCGGATACGCACCTTTCTTACAAGCCCGCAAAAAGCGATGACAATTCA
AAAATCATGCTGGGTGAGCAGTCCATGCCAACAGACCTAATTACATCGGCTTCAGAGACAAC
TTATCGGCCTCATGTATTACAATAGCACTGGCAACATGGGAGTGCTTGCAGGTCAGGCCTCTCAG
35 TTGAATGCAGTGGTGGACTTGCAAGACAGAAACACAGAACTGTCCTACCAGCTCTTGCTTGATT
CCATGGGTGACAGAACCCAGATACTTTTCCATGTGGAATCAGGCAGTGGACAGTTATGACCCAGA
TGTTAGAATTATTGAAAATCATGGAATGAAGACGAGCTCCCCAACTATTGTTTCCCTCTGGGTG
GCATAGGGGTAACCTGACACTTACCAGGCTGTAAAAACCAACAATGGCAATAACGGGGGCCAGG
TGACTTGGACAAAAGATGAAACTTTTGCAGATCGCAATGAAATAGGGGTGGGAAACAATTCG
40 TATGGAGATCAACCTCAGTGCCAACCTGTGGAGAACTTCCTGTACTCCAACGTGGCGCTGTAC
CTACCAGACAAGCTTAAGTACAACCCCTCCAATGTGGACATCTCTGACAACCCCAACACCTACG

ATTACATGAACAAGCGAGTGGTGGCCCCGGGGCTGGTGGACTGCTACATCAACCTGGGCGCGCG
CTGGTCGCTGGACTACATGGACAACGTCAACCCCTTCAACCACCACCGCAATGCGGGCCTGCGC
TACCGCTCCATGCTCCTGGGCAACGGGCGCTACGTGCCCTTCCACATCCAGGTGCCCCAGAAGTT
CTTTGCCATCAAGAACCTCCTCCTGCGGGCTCCTACACCTACGAGTGGAACCTTCAGGAAGG
5 ATGTCAACATGGTCTCCAGAGCTCTCTGGGTAACGATCTCAGGGTGGACGGGGCCAGCATCAA
GTTTCGAGAGCATCTGCCTCTACGCCACCTTCTTCCCCATGGCCCACAACACGGCCTCCACGCTCG
AGGCCATGCTCAGGAACGACACCAACGACCAGTCCTTCAATGACTACCTCTCCGCCGCCAACAT
GCTCTACCCCATACCCGCCAACGCCACCAACGTCCCCATCTCCATCCCCCTCGCGCAACTGGGCGG
CCTTCCGCGGGCTGGGCCTTCACCCGCCTCAAGACCAAGGAGACCCCTCCTGGGCTCGGGATTG
10 GACCCCTACTACACCTACTCGGGCTCCATTCCCTACCTGGACGGCACCTTCTACCTCAACCACAC
TTTCAAGAAGGTCTCGGTACCTTCGACTCCTCGGTGAGCTGGCCGGGCAACGACCGTCTGCTCA
CCCCAACGAGTTTCGAGATCAAGCGCTCGGTGACGGGGAGGGCTACAACGTGGCCAGTGCA
ACATGACCAAGGACTGGTTCCTGGTCCAGATGCTGGCCAACTACAACATCGGTACACAGGGCTT
CTACATCCCAGAGAGCTACAAGGACAGGATGTACTCCTTCTTCAGGAACCTTCAGCCCATGAGC
15 CGGCAGGTGGTGGACCAGACCAAGTACAAGGACTACCAGGAGGTGGGCATCATCCACCAGCAC
AACAACTCGGGCTTCGTGGGTACCTCGCCCCACCATGCGCGAGGGACAGGCCTACCCCGCCA
ACTTCCCCTATCCGCTCATAGGCAAGACCGCGGTGACAGCATCACCCAGAAAAAGTTCTCTG
CGACCGCACCTCTGGCGCATCCCCCTTCTCCAGCAACTTCATGTCCATGGGTGCGCTCTCGGACC
TGGGCCAGAACTTGCTCTACGCCAACTCCGCCACGCCCTCGACATGACCTTCGAGGTGACCCC
20 ATGGACGAGCCACCTTCTCTATGTTCTGTTTGAAGTCTTTGACGTGGTCCGGGTCCACCAGCC
GCACCGCGGCGTCATCGAGACCGTGTACCTGCGTACGCCCTTCTCGGCCGGCAACGCCACCACC
TAAAGAAGCAAGCCGAGTCATCGCCGCCTGCATGCCGTGCGGTTCCACCGAGCAAGAGCTCAG
GGCCATCGTCAGAGACCTGGGATGCGGGCCCTATTTTTTGGGCACCTTCGACAAGCGCTTCCCTG
GCTTTGTCTCCCCACACAAGCTGGCCTGCGCCATCGTCAACACGGCCGGCCGCGAGACCGGGGG
25 CGTGCATGGCTGGCCTTCGCCTGGAACCCGCGCTCCAAAACATGCTTCCTCTTTGACCCCTTCG
GCTTTTCGGACCAGCGGCTCAAGCAAATCTACGAGTTCGAGTACGAGGGCTTGCTGCGTGCAG
CGCCATCGCCTCCTCGCCGACCGCTGCGTCACCCTCGAAAAGTCCACCCAGACCGTGCAGGGG
CCCGACTCGGCCGCTGCGGTCTCTTCTGCTGCATGTTTCTGCACGCCTTTGTGCACTGGCCTCAG
AGTCCCATGGACCGCAACCCACCACATGAACCTGCTGACGGGGGTGCCAACTCCATGCTCCAGA
30 GCCCCAGGTGAGGCCACCCCTGCGCCGCAACCAGGAGCAGCTCTACAGCTTCTGGAGCGCCA
CTCGCCTTACTTCCGCCGCCACAGCGCACAGATCAGGAGGGCCACCTCCTTCTGCCACTTGCAAG
AGATGCAAGAAGGGTAATAACGATGTACACACTTTTTTTCTCAATAAATGGCATCTTTTTATTTA
TACAAGCTCTCTGGGGTATTCATTTCCACCACCACCCGCCGTTGTGCCATCTGGCTCTATTTAG
AAATCGAAAGGGTTCTGCCGGGAGTCGCCGTGCGCCACGGGCAGGGACACGTTGCGATACTGGT
35 AGCGGGTGCCCCACTTGAACTCGGGCACCACAGGCGAGGCAGCTCGGGGAAGTTTTGCTCCA
CAGGCTGCGGGTCAGCACAGCGCTTCATCAGGTGCGGCGCCGAGATCTTGAAGTCGCAGTTG
GGGCGCGCCCTGCGCGCGCGAGTTGCGGTACACCGGGTTGCAGCACTGGAACACCAACAGCG
CCGGGTGCTTACGCTGGCCAGCACGCTGCGGTGCGAGATCAGCTCGGCGTCCAGGTCTCCGC
GTTGCTCAGCGCGAACGGGGTCATCTTGGGCACTTGCCGCCCCAGGAAGGGCGCGTGCCCCGGT
40 TTCGAGTTGCAGTCGAGCGCAGCGGGATCAGCAGGTGCCCCGTGCCCGACTCGGCGTTGGGGT
ACAGCGCGCGCATGAAGGCCTGCATCTGGCGGAAGGCCATCTGGGCCCTGGCGCCCTCCGAGAA

GAACATGCCGACAGACTTGCCGAGAACTGGTTTGCGGGGCAGCTGGCGTCGTCGACGGCAGCAGC
CGCGCGTCGGTGTGTGGCGATCTGCACCACGTTGCGCCCCACCGGTTCTTCACGATCTTGGCCTT
GGACGATTGCTCCTTCAGCGCGCGCTGCCCGTTCTCGCTGGTCACATCCATCTCGATCAGATGTT
CCTTGTTACACCATGCTGCTGCCGTGCAGACACTTCAGCTCGCCCTCCGTCTCGGTGCAGCGGTGC
5 TGCCACAGCGCGCAGCCCGTGGGCTCGAAAGACTTGTAGGTACCTCCGCGAAGGACTGCAGGT
ACCCCTGCAAAAAGCGGCCCATCATGGTCACGAAGGTCTTGTTGCTGCTGAAGGTCAGCTGCAG
CCGCGGTGCTCCTCGTTCAGCCAGGTCTTGACACGCGCCGCGAGCGCTCCACCTGGTCGGGCA
GCATCTTGAAGTTCACCTTCAGCTCATTCTCCACGTGGTACTTGTCCATCAGCGTGCAGCGCCGCC
TCCATGCCCTTCTCCCAGGCCGACACCAGCGGCAGGCTCACGGGGTTCTTCACCATCACCGTGGC
10 CGCCGCCTCCGCCGCGCTTTCGCTTTCGCCCCGCTGTTCTCTTCTCTTCTCTCTCTCTCGCCG
CCGCCACTCGCAGCCCCCGCACCACGGGGTCTGTTCTTCTGACGGCGCTGCACCTTGCCTTGCC
GTTGCGCCCTGCTTGATGCGCACGGGCGGGTTGCTGAAGCCCACCATCACAGCGCGGCCTCTT
CTTGCTCGTCTCTGCTGTCCAGAATGACCTCCGGGGAGGGGGGTTGGTCATCTCAGTACCGA
GGCACGCTTCTTTTTCTTCTTGGGGGCGTTCCGCAGCTCCGCGGCTGCGGCCGCTGCCGAGGTGC
15 AAGGCCGAGGGCTGGGCGTGCAGCGCACACAGCGCTCTGCGAGCGCTCTCGTCTCTCTCGGA
CTCGAGACGGAGGCGGGCCCGCTTCTTCGGGGGCGCGCGGGGCGGCGGAGGCGGCGGCGGCGA
CGGAGACGGGGACGAGACATCGTCCAGGGTGGGTGGACGGCGGGCCGCGCCGCTCCGCGCTC
GGGGGTGGTCTCGCGCTGGTCTCTTCCGACTGGCCATCTCCCACTGCTCCTTCTCCTATAGGC
AGAAAGAGATCATGGAGTCTCTCATGCGAGTCGAGAAGGAGGAGGACAGCCTAACCGCCCCCT
20 CTGAGCCCTCCACCACCGCCGCCACCACCGCCAATGCCGCCGCGGACGACGCGCCACCGAGAC
CACCGCCAGTACCACCCTCCCCAGCGACGCACCCCGCTCGAGAATGAAGTGCTGATCGAGCAG
GACCCGGGTTTTGTGAGCGGAGAGGAGGATGAGGTGGATGAGAAGGAGAAGGAGGAGGTGCC
GCCTCAGTGCCAAAAGAGGATAAAAAGCAAGACCAGGACGACGAGATAAGGATGAGACAGC
AGTCGGGCGGGGAACGGAAGCCATGATGCTGATGACGGCTACCTAGACGTGGGAGACGACGT
25 GCTGCTTAAGCACCTGCACCGCCAGTGCGTCATCGTCTGCGACGCGCTGCAGGAGCGCTGCGAA
GTGCCCCCTGGACGTGGCGGAGGTCAGCCGCGCCTACGAGCGGCACCTCTTCGCGCCGCACGTGC
CCCCAAGCGCCGGGAGAACGGCACCTGCGAGCCCAACCCGCGTCTCAACTTCTACCCGGTCTT
CGCGGTACCCGAGGTGCTGGCCACCTACCACATCTTTTTTCAAAAAGTGAAGATCCCCCTCTCT
GCCGCGCAACCGCACCCGCGCCGACAAAACCTGACCCTGCGGCAGGGCGCCACATACCTGA
30 TATCGCCTCTCTGGAGGAAGTGCCCAAGATCTTCGAGGGTCTCGGTGCGGACGAGAAACGGGCG
GCGAACGCTCTGCACGGAGACAGCGAAAACGAGAGTCACTCGGGGGTGGTGGTGGAGCTCGAG
GGCGACAACGCGCGCCTGGCCGTACTCAAGCGCAGCATAGAGGTACCCACTTTGCCTACCCGG
CGTCAACCTGCCCCCAAGGTATGAGTGTGGTCATGGGCGAGCTCATCATGCGCCGCGCCCA
GCCCCCTGGCCGCGGATGCAAACTTGCAAGAGTCCTCCGAGGAAGGCCTGCCGCGGTGACGCGAC
35 GAGCAGCTGGCGCGCTGGCTGGAGACCCGCGACCCGCGCAGCTGGAGGAGCGGCGCAAGCTC
ATGATGGCCGCGGTGCTGGTCACCGTGGAGCTCGAGTGTCTGCAGCGCTTCTTCGCGGACCCCG
AGATGCAGCGCAAGCTCGAGGAGACCCTGCACTACACCTTCCGCCAGGGCTACGTGCGCCAGGC
CTGCAAGATCTCCAACGTGGAGCTCTGCAACCTGGTCTCCTACCTGGGCATCCTGCACGAGAAC
CGCCTCGGGCAGAACGTCTGCACTCCACCCTCAAAGGGGAGGCGCGCCGCGACTACATCCGCG
40 ACTGCGCCTACCTCTTCTCTGCTACACCTGGCAGACGGCCATGGGGGTCTGGCAGCAGTGCCTG
GAGGAGCGCAACCTCAAGGAGCTGGAAAAGCTCCTCAAGCGCACCTCAGGGACCTCTGGACG

GGCTTCAACGAGCGCTCGGTGGCCGCCGCGCTGGCGGACATCATCTTTCCCGAGCGCCTGCTCA
AGACCCTGCAGCAGGGCCTGCCCCACTTCACCAGCCAGAGCATGCTGCAGAACTTCAGGACTTT
CATCCTGGAGCGCTCGGGCATCCTGCCGGCCACTTGCTGCGCGCTGCCCAGCGACTTCGTGCCCA
TCAAGTACAGGGAGTGCCCCGCCGCGCTCTGGGGCCACTGCTACCTCTTCCAGCTGGCCAACTA
5 CCTCGCCTACCACTCGGACCTCATGGAAGACGTGAGCGGCGAGGGCCTGCTCGAGTGCCACTGC
CGCTGCAACCTCTGCACGCCCCACCGCTCTCTAGTCTGCAACCCGAGCTGCTCAGCGAGAGTCA
GATTATCGGTACCTTCGAGCTGCAGGGTCCCTCGCCTGACGAGAAGTCCGCGGCTCCAGGGCTG
AAACTACTCCGGGGCTGTGGACTTCCGCCTACCTACGCAAAATTTGTACCTGAGGACTACCACGC
CCACGAGATCAGGTTCTACGAAGACCAATCCCCCCCCGCCAAGGCGGAGCTACCGCCTGCGTC
10 ATCACCAGGGGACATCCTGGGCCAATTGCAAGCCATCAACAAAGCCCCGCCGAGAGTTCTTGC
TGAAAAAGGGTCGGGGGGTGTACCTGGACCCCCAGTCCGGCGAGGAGCTAAACCCGCTACCCCC
GCCGCCGCCAGCAGCGGGACCTTGCTTCCCAGGATGGCACCCAGAAAGAAGCAGCAGCCGC
CGCCGCCGCCGAGCCATACATGCTTCTGGAGGAAGAGGAGGAGGACTGGGACAGTCAGGCAG
AGGAGGTTTCGGACGAGGAGCAGGAGGAGATGATGGAAGACTGGGAGGAGGACAGCAGCCTA
15 GACGAGGAAGCTTCAGAGGCCGAAGAGGTGGCAGACGCAACACCATCGCCCTCGGTTCGAGCC
CCCTCGCCGGGGCCCCCTGAAATCCTCCGAACCCAGCACCAGCGCTATAACCTCCGCTCCTCCGGC
GCCGGCGCCACCCGCCGCAGACCCAACCGTAGATGGGACACCACAGGAACCGGGGTCTGGTAA
GTCCAAGTGCCCGCCGCCGCCACCGCAGCAGCAGCAGCAGCAGCGCCAGGGCTACCGTCTGTGG
CGCGGGCACAAGAACGCCATAGTCGCCTGCTTGCAAGACTGCGGGGGCAACATCTCTTTGCCC
20 GCCGCTTCTGTATTCCACCACGGGGTCGCCTTTCCCCGCAATGTCCTGCATTACTACCGTCATC
TCTACAGCCCCTACTGCAGCGGCGACCCAGAGGCGGCAGCGGCAGCCACAGCGGCGACCACCA
CCTAGGAAGATATCCTCCGCGGGCAAGACAGCGGCAGCAGCGGCCAGGAGACCCGCGGCAGCA
GCGGCGGGAGCGGTGGGCGCACTGCGCCTCTCGCCCAACGAACCCCTCTCGACCCGGGAGCTCA
GACACAGGATCTTCCCCACTTTGTATGCCATCTTCCAACAGAGCAGAGGCCAGGAGCAGGAGCT
25 GAAAATAAAAAACAGATCTCTGCGCTCCCTCACCCGAGCTGTCTGTATCACAAAAGCGAAGAT
CAGCTTCGGCGCACGCTGGAGGACGCGGAGGCACTCTTCAGCAAATACTGCGCGCTCACTCTTA
AAGACTAGCTCCGCGCCCTTCTCGAATTTAGGCGGGAGAAAACACTACGTCATCGCCGGCCGCCGC
CCAGCCCGCCAGCCGAGATGAGCAAAGAGATTCCACGCCATACATGTGGAGCTACCAGCCGC
AGATGGGACTCGCGGCGGGAGCGGCCAGGACTACTCCACCCGCATGAACTACATGAGCGCGG
30 GACCCACATGATCTCACAGGTCAACGGGATCCGCGCCCAGCGAAACCAAATACTGCTGGAACA
GGCGGCCATCACCGCCACGCCCCGCCATAATCTCAACCCCCGAAATTGGCCCGCCGCCCTCGTG
TACCAGGAAACCCCTCCGCCACCACCGTACTACTTCCGCGTGACGCCAGGCCGAAGTCCAGA
TGAATAACTCAGGGGCGCAGCTCGCGGGCGGCTTTCGTACAGGGGCGCGGCCGCTCCGACCAGG
TATAAGACACCTGATGATCAGAGGCCGAGGTATCCAGCTCAACGACGAGTCGGTGAGCTCTTCG
35 CTCGGTCTCCGTCCGACGGAACCTTCCAGCTCGCCGATCCGGCCGCTCTTCGTTACGCCCCG
CCAGGCGTACCTGACTCTGCAGACCTCGTCTCGGAGCCCCGCTCCGGCGGCATCGGAACCTC
CAGTTCGTGGAGGAGTTCGTGCCCTCGGTCTACTTCAACCCCTTCTCGGGACCTCCCGGACGCTA
CCCCGACCAGTTCATTCCGAACCTTGACGCGGTGAAGGACTCGGCGGACGGCTACGACTGAATG
TCAGGTGTCGAGGCAGAGCAGCTTCGCCTGAGACACCTCGAGCACTGCCGCCGCCACAAGTGCT
40 TCGCCCGCGGTTCTGGTGAGTTCTGCTACTTTCAGCTACCCGAGGAGCATAACCGAGGGGCGCGC
GCACGGCGTCCGCCTGACCACCCAGGGCGAGGTTACCTGTTCCCTCATCCGGGAGTTTACCCTCC

GTCCCCCTGCTAGTGGAGCGGGAGCGGGGTCCCTGTGTCTAACTATCGCCTGCAACTGCCCTAAC
CCTGGATTACATCAAGATCTTTGCTGTCATCTCTGTGCTGAGTTTAATAAACGCTGAGATCAGAA
TCTACTGGGGCTCCTGTCGCCATCCTGTGAACGCCACCGTCTTCACCCACCCCGACCAGGCCAG
GCGAACCTCACCTGCGGTCTGCATCGGAGGGCCAAGAAGTACCTCACCTGGTACTTCAACGGCA
5 CCCCCCTTGTGGTTTACAACAGCTTCGACGGGGACGGAGTCTCCCTGAAAGACCAGCTCTCCGGT
CTCAGCTACTCCATCCACAAGAACACCACCCTCCAACCTCTCCCTCCCTACCTGCCGGGAACCTA
CGAGTGCGTCACCGGCCGCTGCACCCACCTACCCGCCTGATCGTAAACCAGAGCTTTCCGGGA
ACAGATAACTCCCTCTTCCCCAGAACAGGAGGTGAGCTCAGGAAACTCCCCGGGGACCAGGGCG
GAGACGTACCTTCGACCCTTGTGGGGTTAGGATTTTTTATTACCGGGTTGCTGGCTCTTTTAATCA
10 AAGTTTCCTTGAGATTTGTTCTTTCTTCTACGTGTATGAACACCTCAACCTCCAATAACTCTACC
CTTCTTCGGAATCAGGTGACTTCTCTGAAATCGGGCTTGGTGTGCTGCTTACTCTGTTGATTTTT
TTCCTTATCATACTCAGCCTTCTGTGCCCTCAGGCTCGCCGCCTGCTGCGCACACATCTATATCTAC
TGCTGGTTGCTCAAGTGCAGGGGTGCGCACCCAAGATGAACAGGTACATGGTCCATCGATCCT
AGGCCTGCTGGCCCTGGCGGCCTGCAGCGCCGCCAAAAAAGAGATTACCTTTGAGGAGCCCGCT
15 TGCAATGTAACCTTCAAGCCCGAGGGTGACCAATGCACCACCCTCGTCAAATGCGTTACCAATC
ATGAGAGGGCTGCGCATCGACTACAAAAACAAAACCTGGCCAGTTTGCGGTCTATAGTGTGTTTAC
GCCCCGAGACCCCTCTAACTACTCTGTACCGTCTTCCAGGGCGGACAGTCTAAGATATTCAATT
ACACTTTCCCTTTTTATGAGTTATGCGATGCGGTATGTACATGTCAAAACAGTACAACCTGTGG
CCTCCCTCTCCCCAGGCGTGTGTGGAAAAATACTGGGTCTTACTGCTGTATGGCTTTCGCAATCAC
20 TACGCTCGCTCTAATCTGCACGGTGCTATACATAAAATTACAGGCAGAGGCGAATCTTTATCGATG
AAAAGAAAATGCCTTGATCGCTAACACCGGCTTTCTATCTGCAGAATGAATGCAATCACCTCCCT
ACTAATCACCAACCACCCTCCTTGCGATTGCCATGGGTTGACACGAATCGAAGTGCCAGTGGGG
TCCAATGTCACCATGGTGGGCCCCGCCGCAATTCCACCCTCATGTGGGAAAAATTTGTCCGCA
ATCAATGGGTTCATTTCTGCTCTAACCGAATCAGTATCAAGCCCAGAGCCATCTGCGATGGGCA
25 AAATCTAACTCTGATCAATGTGCAAATGATGGATGCTGGGTACTATTACGGGCAGCGGGGAGAA
ATCATTAATTACTGGCGACCCACAAGGACTACATGCTGCATGTAGTCGAGGCACTTCCCCTACTAC
CACCCCCACTACCACCTCTCCCACCACCACCACCCTACTACTACTACTACTACTACTACTACTA
CTACCACTACCGCTGCCCCGCATACCCGCAAAAGCACCATGATTAGCACAAAGCCCCCTCGTGC
TCACTCCCACGCCGGCGGGCCCATCGGTGCGACCTCAGAAACCACCGAGCTTTGCTTCTGCCAAT
30 GCACTAACGCCAGCGCTCATGAACTGTTGACCTGGAGAATGAGGATGTCCAGCAGAGCTCCGC
TTGCCTGACCCAGGAGGCTGTGGAGCCCGTTGCCCTGAAGCAGATCGGTGATTCAATAATTGAC
TCTTCTCTTTTGCCACTCCCGAATACCCCTCCCGATTCTACTTTCCACATCACGGGTACCAAAGAC
CCTAACCTCTCTTTCTACCTGATGCTGCTGCTCTGTATCTCTGTGGTCTCTTCCGCGCTGATGTTA
CTGGGGATGTTCTGCTGCCTGATCTGCCGCAGAAAGAGAAAAGCTCGCTCTCAGGGCCAACCAC
35 TGATGCCCTTCCCCTACCCCCCGATTTTGAGATAACAAGATATGAGCTCGCTGCTGACACTAA
CCGCTTTACTAGCCTGCGCTCTAACCTTGTGCGTTGCGACTCGAGATTCCACAATGTCACAGCT
GTGGCAGGAGAAAATGTTACTTTCAACTCCACGGCCGATACCCAGTGGTCTGTGGAGTGGCTCAG
GTAGCTACTTAACTATCTGCAATAGCTCCACTTCCCCCGGCATATCCCCAACCAAGTACCAATGC
AATGCCAGCCTGTTACCCCTCATCAACGCTTCCACCCTGGACAATGGACTCTATGTAGGCTATGT
40 ACCCTTTGGTGGGCAAGGAAAGACCCACGCTTACAACCTGGAAGTTGCCAGCCAGCCAGAACCACT
ACCCAAGCTTCTCCCACCACCACCACCACCACCACCATCACCAGCAGCAGCAGCAGCAGCAGC

ACAGCAGCAGCAGCAGATTATTGACTTTGGTTTTGGCCAGCTCATCTGCCGCTACCCAGGCCATC
TACAGCTCTGTGCCCCGAAACCACTCAGATCCACCGCCCAGAAACGACCACCGCCACCACCCTAC
ACACCTCCAGCGATCAGATGCCGACCAACATCACCCCCCTGGCTCTTCAAATGGGACTTACAAG
CCCCACTCCAAAACCACTGGATGCGGCCGAGGTCTCCGCCCTCGTCAATGACTGGGCGGGGCTG
5 GGAATGTGGTGGTTCGCCATAGGCATGATGGCGCTCTGCCTGCTTCTGCTCTGGCTCATCTGCTG
CCTCCACCGCAGGCGAGCCAGACCCCCCATCTATAGACCCATCATTGTCCTGAACCCCGATAAT
GATGGGATCCATAGATTGGATGGCCTGAAAAACCTACTTTTTTCTTTTACAGTATGATAAATTGA
GACATGCCTCGCATTTTCTTGTACATGTTCTTCTCCACCTTTTCTGGGGTGTCTACGCTGGCC
GCTGTGTCTCACCTGGAGGTAGACTGCCTCTCACCTTCACTGTCTACCTGCTTTACGGATTGGTC
10 ACCCTCACTCTCATCTGCAGCCTAATCACAGTAATCATCGCCTTCATCCAGTGCATTGATTACAT
CTGTGTGCGCCTCGCATACTTCAGACACCACCCGCGAGTACCGAGACAGGAACATTGCCCAACTT
CTAAGACTGCTCTAATCATGCATAAGACTGTGATCTGCCTTCTGATCCTCTGCATCCTGCCACC
CTCACCTCCTGCCAGTACACCACAAAATCTCCGCGCAAAAGACATGCCTCCTGCCGCTTCACCCA
ACTGTGGAATATACCCAAATGCTACAACGAAAAGAGCGAGCTCTCCGAAGCTTGGCTGTATGGG
15 GTCATCTGTGTCTTAGTTTTCTGCAGCACTGTCTTTGCCCTCATAATCTACCCCTACTTTGATTTG
GGATGGAACGCGATCGATGCCATGAATTACCCACCTTTCCCGCACCCGAGATAATTCCACTGC
GACAAGTTGTACCCGTTGTCGTTAATCAACGCCCCCATCCCCTACGCCCACTGAAATCAGCTAC
TTTAACCTAACAGGCGGAGATGACTGACGCCCTAGATCTAGAAATGGACGGCATCAGTACCGAG
CAGCGTCTCCTAGAGAGGCGCAGGCAGGCGGCTGAGCAAGAGCGCCTCAATCAGGAGCTCCGA
20 GATCTCGTTAACCTGCACCAGTGCAAAAGAGGCATCTTTTGTCTGGTAAAGCAGGCCAAAGTCA
CCTACGAGAAGACCGGCAACAGCCACCGCCTCAGTTACAAATTGCCACCCAGCGCCAGAAGCT
GGTGTCTATGGTGGGTGAGAATCCCATCACCGTCACCCAGCACTCGGTAGAGACCGAGGGGTGT
CTGCACTCCCCCTGTCGGGGTCCAGAAGACCTCTGCACCCTGGTAAAGACCCTGTGCGGTCTCAG
AGATTTAGTCCCCCTTAACCTAATCAAACACTGGAATCAATAAAAAGAATCACTTACTTAAAATC
25 AGACAGCAGGTCTCTGTCCAGTTTATTCAGCAGCACCTCCTTCCCCCTCCTCCCACTCTGGTACT
CCAAACGCCTTCTGGCGGCAAACTTCCTCCACACCCTGAAGGGAATGTCAGATTCTTGCTCCTGT
CCCTCCGCACCCACTATCTTCATGTTGTTGCAGATGAAGCGCACCAAAACGTCTGACGAGAGCTT
CAACCCCGTGTAACCCCTATGACACGGAAGCGGCCCTCCCTCCGTCCCTTTCCTCACCCCTCCCT
TCGTGTCTCCCGATGGATTCCAAGAAAGTCCCCCGGGTCTGTCTCTGAACCTGGCCGAGCCC
30 CTGGTCACTTCCCACGGCATGCTCGCCCTGAAAATGGGAAGTGGCCTCTCCCTGGACGACGCTG
GCAACCTCACCTCTCAAGATATCACACCGCTAGCCCTCCCTCAAAAAACCAAGACCAACCT
CAGCCTAGAAACCTCATCCCCCTAACTGTGAGCACCTCAGGCGCCCTCACCGTAGCAGCCGCC
GCTCCCCTGGCGGTGGCCGGCACCTCCCTCACCATGCAATCAGAGGCCCCCTGACAGTACAGG
ATGCAAAACTCACCTGGCCACCAAAGGCCCCCTGACCGTGTCTGAAGGCAAACTGGCCTTGCA
35 AACATCGGCCCCGCTGACGGCCGCTGACAGCAGCACCTCACAGTCAGTGCCACACCACCCCTT
AGCACAAGCAATGGCAGCTTGGGTATTGACATGCAAGCCCCATTTACACCACCAATGGAAAAC
TAGGACTTAACTTTGGCGCTCCCCTGCATGTGGTAGACAGCCTAAATGCACTGACTGTAGTTACT
GGCCAAGGTCTTACGATAAACGGAACAGCCCTACAACTAGAGTCTCAGGTGCCCTCAACTATG
ACACATCAGGAAACCTAGAATTGAGAGCTGCAGGGGGTATGCGAGTTGATGCAAAATGGTCAACT
40 TATCCTTGATGTAGCTTACCCATTTGATGCACAAAACAATCTCAGCCTTAGGCTTGGACAGGGAC
CCCTGTTTGTAACTCTGCCACAACCTGGATGTAACTACAACAGAGGCCTCTACCTGTTCACA

TCTGGAAATACCAAAAAGCTAGAAGTTAATATCAAAACAGCCAAGGGTCTCATTTATGATGACA
CTGCTATAGCAATCAATGCGGGTGATGGGCTACAGTTTGACTCAGGCTCAGATACAAATCCATT
AAAACTAACTTGGATTAGGACTGGATTATGACTCCAGCAGAGCCATAATTGCTAACTGGGA
ACTGGCCTAAGCTTTGACAACACAGGTGCCATCACAGTAGGCAACAAAATGATGACAAGCTTA
5 CCTTGTGGACCACACCAGACCCATCCCCTAACTGTAGAATCTATTTCAGAGAAAGATGCTAAATT
CACACTTGTTTTGACTAAATGCGGCAGTCAGGTGTTGGCCAGCGTTTCTGTTTTATCTGTAAAAG
GTAGCCTTGCGCCCATCAGTGGCACAGTAACTAGTGCTCAGATTGTCCTCAGATTTGATGAAAAT
GGAGTTCTACTAAGCAATTCTTCCCTTGACCCTCAATACTGGAACCTACAGAAAAGGTGACCTTAC
AGAGGGCACTGCATATACCAACGCAGTGGGATTTATGCCCAACCTCACAGCATACCCAAAAACA
10 CAGAGCCAAACTGCTAAAAGCAACATTGTAAGTCAGGTTTACTTGAATGGGGACAAATCCAAAC
CCATGACCCTCACCATTACCCTCAATGGAACCTAATGAAACAGGAGATGCCACAGTAAGCACTTA
CTCCATGTCAATTCTCATGGAACCTGGAATGGAAGTAATTACATTAATGAAACGTTCCAAACCAACT
CCTTCACCTTCTCCTACATCGCCCAAGAATAAAAAGCATGACGCTGTTGATTTGATTCAATGTGT
TTCTGTTTTATTTTCAAGCACAAACAAATCATTCAAGTCATTCTTCCATCTTAGCTTAATAGACAC
15 AGTAGCTTAATAGACCCAGTAGTGCAAAGCCCCATTCTAGCTTATAGATCAGACAGTGATAATT
AACCACCACCACCACCATACCTTTTTGATTCAGGAAATCATGATCATCACAGGATCCTAGTCTTCA
GGCCGCCCCCTCCCTCCCAAGACACAGAATACACAGTCCTCTCCCCCGACTGGCTTTAAATAAC
ACCATCTGGTTGGTCACAGACATGTTCTTAGGGGTTATATTCCACACGGTCTCCTGCCGCGCCAG
GCGCTCGTCGGTGATGTTGATAAACTCTCCCGGCAGCTCGCTCAAGTTCACGTCGCTGTCCAGCG
20 GCTGAACCTCCGGCTGACGCGATAACTGTGCGACCGGCTGCTGGACGAACGGAGGCCGCGCCTA
CAAGGGGGTAGAGTCATAATCCTCGGTCAGGATAGGGCGGTGATGCAGCAGCAGCGAGCGAAA
CATCTGCTGCCGCCGCCGCTCCGTCCGGCAGGAAAACAACACGCCGGTGGTCTCCTCCGCGATA
ATCCGCACCGCCCGCAGCATCAGCTTCCTCGTTCTCCGCGCGCAGCACCTCACCTTATCTCGCT
CAAATCGGCGCAGTAGGTACAGCACAGCACCACGATGTTATTCATGATCCCACAGTGCAGGGCG
25 CTGTATCCAAAGCTCATGCCGGGAACACCGCCCCACGTGGCCATCGTACCACAAGCGCACGT
AAATCAAGTGTGACCCCTCATGAACGCGCTGGACACAAACATTACTTCTTGGGCATGTTGTA
ATTACACCTCCCGGTACCAGATAAACCTCTGGTTGAACAGGGCACCTTCCACCACCATCCTGA
ACCAAGAGGCCAGAACCTGCCCACCGGCTATGCACTGCAGGGAACCCGGGTTGGAACAATGAC
AATGCAGACTCCAAGGCTCGTAACCGTGGATCATCCGGCTGCTGAAGGCATCGATGTTGGCACA
30 ACACAGACACACGTGCATGCACTTTCTCATGATTAGCAGCTCTTCCCTCGTCAGGATCATATCCC
AAGGAATAACCCATTCTTGAATCAACGTAAAACCCACACAGCAGGGAAGGCCTCGCACATAACT
CACGTTGTGCATGGTCAGCGTGTTGCATTCCGGAACACGCGGATGATCCTCCAGTATCGAGGCG
CGGGTCTCCTTCTCACAGGGAGGTAAAGGGTCCCTGCTGTACGGAAGTGCAGGACGACCGGAG
ATCGTGTTGAGCGTAGTGTCATGGAAAAGGGAACGCCGGACGTGGTCATACTTCTTGAAGCAGA
35 ACCAGGTTGCGCGGTGGCAGGCCTCCTTGCGTCTGCGGTCTCGCCGTCTAGCTCGCTCCGTGTGA
TAGTTGTAGTACAGCCACTCCCGCAGAGCGTCGAGGCGCACCTGGCTTCCGGATCTATGTAGA
CTCCGTCTTGACCGCGGCCCTGATAATATCCACCACCGTAGAATAAGCAACACCCAGCCAAGC
AATACACTCGCTCTGCGAGCGGCAGACAGGAGGAGCGGGCAGAGATGGGAGAACCATGATAAA
AAACTTTTTTTAAAGAATATTTTCCAATTCTTCGAAAGTAAGATCTATCAAGTGGCAGCGCTCCC
40 CTCCACTGGCGCGGTCAAACCTCTACGGCCAAAGCACAGACAACGGCATTCTAAGATGTTCCCT
AATGGCGTCCAAAAGACACACCGCTCTCAAGTTGCAGTAAACTATGAATGAAAACCCATCCGGC

TGATTTTCCAATATAGACGCGCCGGCAGCGTCCACCAAACCCAGATAATTTTCTTCTCTCCAGCG
 GTTTACGATCTGTCTAAGCAAATCCCTTATATCAAGTCCGACCATGCCAAAAATCTGCTCAAGAG
 CGCCCTCCACCTTCATGTACAAGCAGCGCATCATGATTGCAAAAATTCAGGTTCTTCAGAGACCT
 GTATAAGATTCAAAATGGGAACATTAACAAAAATTCCTCTGTTCGCGCAGATCCCTTCGCAGGGC
 5 AAGCTGAACATAATCAGACAGGTCCGAACGGACCAGTGAGGCCAAATCCCCACCAGGAACCAG
 ATCCAGAGACCCTATACTGATTATGACGCGCATACTCGGGGCTATGCTGACCAGCGTAGCGCCG
 ATGTAGGCGTGCTGCATGGGCGGCGAGATAAAATGCAAAGTGCTGGTTAAAAATCAGGCAAA
 GCCTCGCGCAAAAAAGCTAACACATCATAATCATGCTCATGCAGGTAGTTGCAGGTAAGCTCAG
 GAACCAAAACGGAATAACACACGATTTTCTCTCAAACATGACTTCGCGGATACTGCGTAAAAC
 10 AAAAAATTATAAATAAAAAATTAATTAATAAATTAAACATTGGAAGCCTGTCTCACAACAGGA
 AAAACCACTTTAATCAACATAAGACGGGCCACGGGCATGCCGGCATAGCCGTAAAAAAATTGGT
 CCCCCTGATTAAACAAGTACCACAGACAGCTCCCCGGTCATGTCTGGGGGTCATCATGTGAGACTC
 TGTATACACGTCTGGATTGTGAACATCAGACAAACAAAGAAATCGAGCCACGTAGCCCCGAGGT
 ATAATCACCCGCAGGCGGAGGTACAGCAAAACGACCCCATAGGAGGAATCACAAAATTAGTA
 15 GGAGAAAAAATACATAAACACCAGAAAAACCCTGTTGCTGAGGCAAAATAGCGCCCTCCCGA
 TCCAAAACAACATAAAGCGCTTCACAGGAGCAGCCATAACAAAGACCCGAGTCTTACCAGTAA
 AAGAAAAAAGATCTCTCAACGCAGCACCAGCACCAACACTTCGCAGTGTAAGGCCAAGTGC
 CGAGAGAGTATATATAGGAATAAAAAGTGACGTAAACGGGCAAGTCCAAAAAACGCCAGAA
 AAACCGCACGCGAACCTACGCCCCGAAACGAAAGCCAAAAAACTAGACACTCCCTTCCGGC
 20 GTCAACTTCCGCTTTCCACGCTACGTCACTTCCCCCGGTCAAACAAACTACATATCCCGAACTT
 CCAAGTCGCCACGCCCCAAACACCGCCTACACCTCCCCGCGCGGCCGCCCCCGGACCCGC
 CTCCCCGCGCGCGCCCATCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATATTATTG
 ATGATG

25 **SEQ ID NO: 11 - Polynucleotide sequence encoding ChAd155/RSV**
 CATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGATGGGCGGCGCGGG
 GCGGGGCGCGGGGCGGGAGGCGGGTTTGGGGGCGGGCCGGCGGGCGGGGCGGTGTGGCGGAA
 GTGGACTTTGTAAGTGTGGCGGATGTGACTTGCTAGTGCCGGGCGCGGTAAAAGTGACGTTTTT
 CGTGC GCGACAACGCCCCCGGAAGTGACATTTTTCCCGCGTTTTTACCGGATGTTGTAGTGAA
 30 TTTGGGCGTAACCAAGTAAGATTTGGCCATTTTCGCGGGAACACTGAAACGGGGAAGTGAAATC
 TGATTAATTTTTCGTTAGTCATACCGCGTAATATTTGTCTAGGGCCGAGGGACTTTGGCCGATTA
 CGTGAGGACTCGCCAGGTGTTTTTTGAGGTGAATTTCCCGGTTCCGGGTCAAAGTCTGCGTTT
 TATTATTATAGGATATCCCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCT
 CATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACG
 35 GGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCC
 TGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG
 CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTACGGTAAACTGCCCACTTGGCAGT
 ACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCC
 TGGCATTATGCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT
 40 CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACT
 CACGGGGATTTCGAAGTCTCCACCCCATGACGTCAATGGGAGTTTGTGTTTGGCACCAAAATCAA

CGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTAC
GGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAG
AGATCGTCGACGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTG
ACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGC
5 GGATTCCCCGTGCCAAGAGTGAGATCTTCCGTTTATCTAGGTACCAGATATCGCCACCATGGAAC
TGCTGATCCTGAAGGCCAACGCCATCACCACCATCCTGACCGCCGTGACCTTCTGCTTCGCCAGC
GGCCAGAACATCACCGAGGAATTCTACCAGAGCACCTGTAGCGCCGTGAGCAAGGGCTACCTGA
GCGCCCTGAGAACCGGTGGTACACCAGCGTGATCACCATCGAGCTGAGCAACATCAAAGAAA
ACAAGTGCAACGGCACCGACGCCAAAGTGAAGCTGATCAAGCAGGAAGTGGACAAGTACAAGA
10 ACGCCGTGACCGAGCTGCAGCTGCTGATGCAGAGCACCCCCGCCACCAACAACCGGGCCAGAC
GGGAGCTGCCCCGTTTCATGAACTACACCCTGAACAACGCCAAAAAGACCAACGTGACCCTGAG
CAAGAAGCGGAAGCGGCGGTTCTTGGGCTTTCTGCTGGGCGTGCGGCGCCATTGCCAGCGGC
GTGGCCGTGTCTAAGGTGCTGCACCTGGAAGGCGAAGTGAACAAGATCAAGAGCGCCCTGCTGA
GCACCAACAAGGCCGTGGTGTCCCTGAGCAACGGCGTGAGCGTGCTGACCAGCAAGGTGCTGG
15 ATCTGAAGAACTACATCGACAAGCAGCTGCTGCCCATCGTGAACAAGCAGAGCTGCAGCATCAG
CAACATCGAGACAGTGATCGAGTTCCAGCAGAAGAACAACCGGCTGCTGGAATCACC CGGGA
GTTTCAGCGTGAACGCCGGCGTGACCACCCCTGTGTCCACCTACATGCTGACCAACAGCGAGCTG
CTGAGCCTGATCAACGACATGCCATCACCACGACAGAAAAAGCTGATGAGCAACAACGTG
CAGATCGTGCGGCAGCAGAGCTACTCCATCATGTCCATCATCAAAGAAGAGGTGCTGGCCTACG
20 TGGTGCAGCTGCCCCGTGACGGCGTGATCGACACCCCTGCTGGAAGCTGCACACCAGCCCCCT
GTGCACCACCAACACCAAAGAGGGCAGCAACATCTGCCTGACCCGGACCGACAGAGGCTGGTA
CTGCGACAACGCCGGCAGCGTGTCATTCTTTCCACAGGCCGAGACATGCAAGGTGCAGAGCAAC
CGGGTGTCTGCGACACCATGAACAGCCTGACCCTGCCCTCCGAAGTGAACCTGTGCAACGTGG
ACATCTTCAACCCCAAGTACGACTGCAAGATCATGACCTCCAAGACCGACGTGTCCAGCTCCGT
25 GATCACCTCCCTGGGCGCCATCGTGTCTGCTACGGCAAGACCAAGTGCACCGCCAGCAACAAG
AACCGGGGCATCATCAAGACCTTCAGCAACGGCTGCGACTACGTGTCCAACAAGGGGGTGGAC
ACCGTGTCCGTGGGCAACACCCTGTACTACGTGAACAAACAGGAAGGCAAGAGCCTGTACGTGA
AGGGCGAGCCCATCATCAACTTCTACGACCCCTGGTGTTCCTCCAGCGACGAGTTCGACGCCAG
CATCAGCCAGGTGAACGAGAAGATCAACCAGAGCCTGGCCTTCATCCGGAAGTCCGACGAGCTG
30 CTGCACAATGTGAATGCCGGCAAGTCCACCACCAACCGGAAGCGGAGAGCCCCTGTGAAGCAG
ACCCTGAACTTCGACCTGCTGAAGCTGGCCGGCGACGTGGAGAGCAATCCCGGCCCTATGGCCC
TGAGCAAAGTGAAGTGAACGATACACTGAACAAGGACCAGCTGCTGTCCAGCAGCAAGTACA
CCATCCAGCGGAGCACCGGCGACAGCATCGATACCCCAACTACGACGTGCAGAAGCACATCA
ACAAGCTGTGCGGCATGCTGCTGATCACAGAGGACGCCAACCACAAGTTCACCGGCCTGATCGG
35 CATGCTGTACGCCATGAGCCGGCTGGGCGGGAGGACACCATCAAGATCCTGCGGGACGCGGC
TACCACGTGAAGGCCAATGGCGTGACGTGACCACACACCGGCAGGACATCAACGGCAAAGAA
ATGAAGTTCGAGGTGCTGACCCTGGCCAGCCTGACCACCGAGATCCAGATCAATATCGAGATCG
AGAGCCGGAAGTCTACAAGAAAATGCTGAAAGAAATGGGCGAGGTGGCCCCCGAGTACAGAC
ACGACAGCCCCGACTGCGGCATGATCATCTGTGTATCGCCGCCCTGGTGATCACAAAGCTGGC
40 CGCTGGCGACAGATCTGGCCTGACAGCCGTGATCAGACGGGCCAACAAATGTGCTGAAGAACGA
GATGAAGCGGTACAAGGGCCTGCTGCCCAAGGACATTGCCAACAGCTTCTACGAGGTGTTTCGAG

AAGTACCCCCACTTCATCGACGTGTTTCGTGCACTTCGGCATTGCCAGAGCAGCACCAGAGGCG
GCTCCAGAGTGAGGGCATCTTCGCCGGCCTGTTTCATGAACGCCTACGGCGCTGGCCAGGTGAT
GCTGAGATGGGGCGTGCTGGCCAAGAGCGTGAAGAACATCATGCTGGGCCACGCCAGCGTGCA
GGCCGAGATGGAACAGGTGGTGGAGGTGTACGAGTACGCCAGAAGCTGGGCGGAGAGGCCGG
5 CTTCTACCACATCCTGAACAACCCTAAGGCCTCCCTGCTGTCCCTGACCCAGTTCCCCCACTTCTC
CAGCGTGGTGCTGGGAAATGCCGCCGACTGGGCATCATGGGCGAGTACCGGGGCACCCCCAG
AAACCAGGACCTGTACGACGCCGCCAAGGCCTACGCCGAGCAGCTGAAAGAAAACGGCGTGAT
CAACTACAGCGTGCTGGACCTGACCGCTGAGGAACTGGAAGCCATCAAGCACCAGCTGAACCCC
AAGGACAACGACGTGGAGCTGGGAGGCGGAGGATCTGGCGGCGGAGGCATGAGCAGACGGAA
10 CCCCTGCAAGTTCGAGATCCGGGGCCACTGCCTGAACGGCAAGCGGTGCCACTTCAGCCACAAC
TACTTCGAGTGGCCCCCTCATGCTCTGCTGGTGCGGCAGAACTTCATGCTGAACCGGATCCTGAA
GTCCATGGACAAGAGCATCGACACCCTGAGCGAGATCAGCGGAGCCGCCGAGCTGGACAGAAC
CGAGGAATATGCCCTGGGCGTGGTGGGAGTGCTGGAAAGCTACATCGGCTCCATCAACAACATC
ACAAAGCAGAGCGCCTGCGTGGCCATGAGCAAGCTGCTGACAGAGCTGAACAGCGACGACATC
15 AAGAAGCTGAGGGACAACGAGGAACTGAACAGCCCCAAGATCCGGGTGTACAACACCGTGATC
AGCTACATTGAGAGCAACCGCAAGAACAACAAGCAGACCATCCATCTGCTGAAGCGGTGCC
GCCGACGTGCTGAAAAAGACCATCAAGAACACCCTGGACATCCACAAGTCCATCACCATCAACA
ATCCCAAAGAAAGCACCGTGTCTGACACCAACGATCACGCCAAGAACAACGACACCACCTGAT
GAGCGGCCGCGATCTGCTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCCTCCCCCGTGCCT
20 TCCTTGACCCTGGAAGGTGCCACTCCCCTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCA
TTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGAT
TGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCCGATCAGCGATCGCTGA
GGTGGGTGAGTGGGCGTGGCCTGGGGTGGTCATGAAAATATATAAGTTGGGGGTCTTAGGGTCT
CTTTATTTGTGTTGCAGAGACCGCCGAGCCATGAGCGGGAGCAGCAGCAGCAGCAGTAGCAGC
25 AGCGCCTTGATGGCAGCATCGTGAGCCCTTATTTGACGACGCGGATGCCCCACTGGGCGGGG
TGCGTCAGAATGTGATGGGCTCCAGCATCGACGGCCGACCCGTCTGCCCCGAAATTCCGCCAC
GCTGACCTATGCGACCGTCGCGGGGACGCCGTTGGACGCCACCGCCGCCGCCGCCGCCACCGCA
GCCGCTCGGCCGTGCGCAGCCTGGCCACGGACTTTGCATTCTGCGGACCACTGGCGACAGGGG
CTACTTCTCGGGCCGCTGCTGCCGCCGTTGCGGATGACAAGCTGACCGCCCTGCTGGCGCAGTTG
30 GATGCGCTTACTCGGGAACCTGGGTGACCTTTCTCAGCAGGTCATGGCCCTGCGCCAGCAGGTCTC
CTCCCTGCAAGCTGGCGGGAATGCTTCTCCACAAATGCCGTTTAAGATAAAATAAACAGACT
CTGTTTGGATTAAAGAAAAGTAGCAAGTGCATTGCTCTCTTTATTTTCATAATTTCCGCGCGCGA
TAGGCCCTAGACCAGCGTTCTCGGTGCTTGAGGGTGCGGTGTATCTTCTCCAGGACGTGGTAGA
GGTGGCTCTGGACGTTGAGATACATGGGCATGAGCCCGTCCCGGGGGTGGAGGTAGCACCCTG
35 CAGAGCTTCATGCTCCGGGGTGGTGTGTAGATGATCCAGTCGTAGCAGGAGCGCTGGGCATGG
TGCCTAAAAATGTCCTTCAGCAGCAGGCCGATGGCCAGGGGGAGGCCCTTGGTGTAAAGTGTATA
CAAAACGGTTAAGTTGGGAAGGGTGCATTGCGGGAGAGATGATGTGCATCTTGGACTGTATTTT
TAGATTGGCGATGTTTCCGCCAGATCCCTTCTGGGATTCATGTTGTGCAGGACCACCAGTACAG
TGTATCCGGTGAACCTTGGGGAATTTGTCATGCAGCTTAGAGGGAAAAGCGTGGAAGAACTTGA
40 GACGCCTTTGTGGCCTCCAGATTTTCCATGCATTCGTCCATGATGATGGCAATGGGCCCCGCGG
AGGCAGCTTGGGCAAAGATATTTCTGGGGTCTGCTGACGTCGTAGTTGTGTTCCAGGGTGAGGTC

GTCATAGGCCATTTTACAAAGCGCGGGCGGAGGGTGCCCGACTGGGGGATGATGGTCCCTCT
GGCCCTGGGGCGTAGTTGCCCTCGCAGATCTGCATTTCCAGGCCCTAATCTCGGAGGGGGGAA
TCATATCCACCTGCGGGGCGATGAAGAAAACGGTTTCCGGAGCCGGGGAGATTAACCTGGGATGA
GAGCAGGTTTCTAAGCAGCTGTGATTTCCACAACCGGTGGGCCCATAAATAACACCTATAACC
5 GGTTCAGCTGGTAGTTTAGAGAGCTGCAGCTGCCGTCGTCCCGGAGGAGGGGGGCCACCTCGT
TGAGCATGTCCCTGACGCGCATGTTCTCCCCGACCAGATCCGCCAGAAGGCGCTCGCCGCCAG
GGACAGCAGCTCTTGCAAGGAAGCAAAGTTTTTCAGCGGCTTGAGGCCGTCCGCCGTGGGCATG
TTTTTCAGGGTCTGGCTCAGCAGCTCCAGGCGGTCCAGAGCTCGGTGACGTGCTCTACGGCATC
TCTATCCAGCATATCTCCTCGTTTCGCGGGTTGGGGCGACTTTCGCTGTAGGGCACCAAGCGGTG
10 GTCGTCCAGCGGGGCCAGAGTCATGTCCTTCCATGGGCGCAGGGTCTCGTCAGGGTGGTCTGG
GTCACGGTGAAGGGGTGCGCTCCGGGCTGAGCGCTTGCCAAGGTGCGCTTGAGGCTGGTTCTGC
TGGTGCTGAAGCGCTGCCGGTCTTCGCCCTGCGCGTCGGCCAGGTAGCATTTGACCATGGTGTCA
TAGTCCAGCCCCCTCCGCGGCGTGTCCCTTGCGCGCAGCTTGCCCTTGAGGTTGGCGCCGCACG
AGGGGCAGAGCAGGCTCTTGAGCGCGTAGAGCTTGGGGGCGAGGAAGACCGATTTCGGGGGAGT
15 AGGCGTCCGCGCCGCAGACCCCGCACACGGTCTCGCACTCCACCAGCCAGGTGAGCTCGGGGCG
CGCCGGGTCAAAAACAGGTTTCCCCCATGCTTTTGATGCGTTTCTTACCTCGGGTCTCCATGA
GGTGGTGTCCCCGCTCGGTGACGAAGAGGCTGTCCGTGTCTCCGTAGACCGACTTGAGGGGTCT
TTTCTCCAGGGGGTCCCTCGGTCTTCCTCGTAGAGGAACTCGGACCACTCTGAGACGAAGGCC
CGCGTCCAGGCCAGGACGAAGGAGGCTATGTGGGAGGGGTAGCGGTCTGTCCACTAGGGGG
20 TCCACCTTCTCCAAGGTGTGAAGACACATGTGCGCTTCCTCGGCGTCCAGGAAGGTGATTGGCTT
GTAGGTGTAGGCCACGTGACCGGGGGTTCCTGACGGGGGGGTATAAAGGGGGTGGGGGCGCG
CTCGTCTCACTCTCTCCGCATCGTGTCTGCGAGGGGCCAGCTGCTGGGGTGAGTATCCCTCT
CGAAGGCGGGCATGACCTCCGCGCTGAGGTTGTCAGTTTCCAAAACGAGGAGGATTTGATGTT
CACCTGTCCCGAGGTGATACCTTTGAGGGTACCCGCGTCCATCTGGTCAGAAAACACGATCTTTT
25 TATTGTCCAGCTTGGTGGCGAACGACCCGTAGAGGGCGTTGGAGAGCAGCTTGGCGATGGAGCG
CAGGGTCTGGTCTTGTTCCCTGTGCGCGCGCTCCTTGCGCGCGATGTTGAGCTGCACGTA CTGC
GCGCGACGCAGCGCCACTCGGGGAAGACGGTGGTGCCTCGTTCGGGACCAAGGCGCACGCGCC
AGCCGCGGTTGTGCAGGGTGACCAAGTCCACGCTGGTGGCGACCTCGCCGCGCAGGCGCTCGTT
GGTCCAGCAGAGACGGCCGCCCTTGCGCGAGCAGAAGGGGGGCAGGGGGTCGAGCTGGGTCTC
30 GTCCGGGGGGTCCGCGTCCACGGTGAAAACCCCGGGGCGCAGGCGCGCGTCAAGTAGTCTATC
TTGCAACCTTGATGTCCAGCGCCTGCTGCCAGTCGCGGGCGGCGAGCGCGCGCTCGTAGGGGT
TGAGCGGCGGGCCCCAGGGCATGGGGTGGGTGAGTGCGGAGGCGTACATGCCGCAGATGTCAT
AGACGTAGAGGGGCTCCCGCAGGACCCCGATGTAGGTGGGGTAGCAGCGGCCCGCGCGGATGC
TGGCGCGCACGTAGTCATACAGCTCGTGCGAGGGGGCGAGGAGGTGCGGGGCCAGGTTGGTGC
35 GGGCGGGGCGCTCCGCGCGGAAGACGATCTGCCTGAAGATGGCATGCGAGTTGGAAGAGATGG
TGGGGCGCTGGAAGACGTTGAAGCTGGCGTCTGCAAGGCCGACGGCGTCGCGCACGAAGGAGG
CGTAGGAGTCGCGCAGCTTGTGTACCAGCTCGGCGGTGACCTGCACGTGAGCGCGCAGTAGTC
GAGGGTCTCGCGGATGATGTCATATTTAGCCTGCCCCTTCTTTTCCACAGCTCGCGGTTGAGGA
CAAACCTCTTCGCGTCTTTCCAGTACTCTTGATCGGGAAACCGTCCGGTTCCGAACGGTAAGAG
40 CCTAGCATGTAGAACTGGTTGACGGCCTGGTAGGCGCAGCAGCCCTTCTCCACGGGGAGGGCGT
AGGCCTGCGCGGCCTTGCGGAGCGAGGTGTGGGTGAGGGCGAAGGTGTCCCTGACCATGACTTT

GAGGTACTGGTGCTTGAAGTCGGAGTCGTCGCAGCCGCCCCGCTCCCAGAGCGAGAAGTCGGTG
CGCTTCTTGAGCGGGGTTGGGCAGAGCGAAGGTGACATCGTTGAAGAGGATTTTGCCCGCGC
GGGGCATGAAGTTGCGGGTGATGCGGAAGGGCCCCGGCACTTCAGAGCGGTTGTTGATGACCTG
GGCGGCGAGCACGATCTCGTCGAAGCCGTTGATGTTGTGGCCACGATGTAGAGTTCCAGGAAG
5 CGGGGCCGGCCCTTTACGGTGGGCAGCTTCTTTAGCTCTTCGTAGGTGAGCTCCTCGGGCGAGGC
GAGGCCGTGCTCGGCCAGGGCCCAGTCCGCGAGGTGCGGGTTGTCTCTGAGGAAGGACTTCCAG
AGGTGCGGGGCCAGGAGGGTCTGCAGGCGGTCTCTGAAGGTCTTGAAGTGGCGGGCCACGGCCA
TTTTTTCGGGGGTGATGCAGTAGAAGGTGAGGGGGTCTTGCTGCCAGCGGTCCCAGTCGAGCTG
CAGGGCGAGGTGCGCGCGGCGGTGACCAGGCGCTCGTCGCCCCGAATTTTCATGACCAGCATG
10 AAGGGCACGAGCTGCTTTCCGAAGGCCCCCATCCAAGTGTAGGTCTCTACATCGTAGGTGACAA
AGAGGCGCTCCGTGCGAGGATGCGAGCCGATCGGGAAGAACTGGATCTCCCGCCACCAGTTGG
AGGAGTGGCTGTTGATGTGGTGGAAGTAGAAGTCCCGTCGCCGGGGCCGAACACTCGTGCTGGCT
TTTGTAAGCGAGCGCAGTACTGGCAGCGCTGCACGGGCTGTACCTCATGCACGAGATGCACC
TTTCGCCCCGCGCACGAGGAAGCCGAGGGGAAATCTGAGCCCCCGCTGGCTCGCGGCATGGCT
15 GGTCTCTTCTACTTTGGATGCGTGTCCTGCTCCGTCTGGCTCCTCGAGGGGTGTTACGGTGGAG
CGGACCACCACGCCGCGGAGCCGAGGTCCAGATATCGGCGCGCGGCGGTGCGAGTTTGATGA
CGACATCGCGCAGCTGGGAGCTGTCCATGGTCTGGAGCTCCCGCGGCGGCGGCAGGTCAGCCGG
GAGTTCTTGACGTTTACCTCGCAGAGTCGGGCCAGGGCGCGGGGCAGGTCTAGGTGGTACCTG
ATCTCTAGGGGCGTGTTGGTGGCGGCGTCGATGGCTTGCAGGAGCCCGCAGCCCCGGGGGGCGA
20 CGACGGTGCCCCGCGGGGTGGTGGTGGTGGTGGCGGTGCAGCTCAGAAGCGGTGCCGCGGGCG
GGCCCCCGGAGGTAGGGGGGGCTCCGGTCCCGCGGGCAGGGGCGGCAGCGGCACGTGCGCGTG
GAGCGCGGGCAGGAGTTGGTGCTGTGCCCGAGGTTGCTGGCGAAGGCGACGACGCGGCGGTT
GATCTCCTGGATCTGGCGCCTCTGCGTGAAGACGACGGGCCCGGTGAGCTTGAACCTGAAAGAG
AGTTCGACAGAAATCAATCTCGGTGTCATTGACCGCGCCTGGCGCAGGATCTCCTGCACGTCTCC
25 CGAGTTGTCTTGGTAGGCGATCTCGGCCATGAACTGCTCGATCTCTTCTCCTGGAGGTCTCCGC
GTCCGGCGCGTTCCACGGTGGCCGCCAGGTGCTTGGAGATGCGCCCCATGAGCTGCGAGAAGGC
GTTGAGTCCGCCCTCGTTCCAGACTCGGCTGTAGACCACGCCCCCTGGTCATCGCGGGCGCGCA
TGACCACCTGCGCGAGGTTGAGCTCCACGTGCCGCGCGAAGACGGCGTAGTTGCGCAGACGCTG
GAAGAGGTAGTTGAGGGTGGTGGCGGTGTGCTCGGCCACGAAGAAGTTCATGACCCAGCGGCG
30 CAACGTGGATTTCGTTGATGTCCCCCAAGGCCTCCAGCCGTTCCATGGCCTCGTAGAAGTCCACGG
CGAAGTTGAAAACTGGGAGTTGCGCGCCGACACGGTCAACTCCTCCTCCAGAAGACGGATGAG
CTCGGCGACGGTGTGCGGCACCTCGCGCTCGAAGGCTATGGGGATCTCTTCTCCTCCGCTAGCATCA
CCACCTCCTCCTTCTCCTCCTTCTGCGCACTTCCATGATGGCTTCTCCTCCTTCGGGGGGTGGCG
GCGGCGGCGGTGGGGGAGGGGGCGCTCTGCGCCGGCGGCGGCGCACCGGGAGGCGGTCCACGA
35 AGCGCGGATCATCTCCCCGCGGCGGCGGCGCATGGTCTCGGTGACGGCGCGGCGGTTCTCCCG
GGGGCGCAGTTGGAAGACGCCCGCGGACATCTGGTGCTGGGGCGGGTGGCCGTGAGGCAGCGA
GACGGCGCTGACGATGCATCTCAACAATTGCTGCGTAGGTACGCCGCCGAGGGACCTGAGGGAG
TCCATATCCACCGGATCCGAAAACCTTTCGAGGAAGGCGTCTAACCAGTCGCAGTCGCAAGGTA
GGCTGAGCACCGTGGCGGGCGGCGGGGGTGGGGGGAGTGTCTGGCGGAGGTGCTGCTGATGA
40 TGTAATTGAAGTAGGCGGACTTGACACGGCGGATGGTCGACAGGAGCACCATGTCTTGGGTCC
GGCCTGCTGGATGCGGAGGCGGTGCGCTATGCCCCAGGCTTCGTTCTGGCATCGGCGCAGGTCC

TTGTAGTAGTCTTGCATGAGCCTTCCACCGGCACCTCTTCTCCTTCCTCTTCTGCTTCTTCCATGT
CTGCTTCGGCCCTGGGGCGGCGCCGCGCCCCCTGCCCCCATGCGCGTGACCCCGAACCCCTG
AGCGGTTGGAGCAGGGCCAGGTCGGCGACGACGCGCTCGGCCAGGATGGCCTGCTGCACCTGC
GTGAGGGTGGTTTGAAGTCATCCAAGTCCACGAAGCGGTGGTAGGCGCCCGTGTGATGGTGT
5 AGGTGCAGTTGGCCATGACGGACCAGTTGACGGTCTGGTGGCCCGGTTGCGACATCTCGGTGTA
CCTGAGTCGCGAGTAGGCGCGGGAGTCGAAGACGTAGTCGTTGCAAGTCCGCACCAGGTACTGG
TAGCCCACCAGGAAGTGCGGCGGCGGCTGGCGGTAGAGGGGCCAGCGCAGGGTGGCGGGGGCT
CCGGGGGCCAGGTCTTCCAGCATGAGGCGGTGGTAGGCGTAGATGTACCTGGACATCCAGGTGA
TACCCGCGGCGGTGGTGGAGGCGCGCGGAAGTCGCGCACCCGGTTCAGATGTTGCGCAGGG
10 GCAGAAAGTGCTCCATGGTAGGCGTGCTCTGTCCAGTCAGACGCGCGCAGTCGTTGATACTCTA
GACCAGGGAAAACGAAAGCCGGTCAGCGGGCACTCTTCCGTGGTCTGGTGAATAGATCGCAAG
GGTATCATGGCGGAGGGCCTCGGTTTCGAGCCCCGGGTCCGGGCGGACGGTCCGCCATGATCCA
CGCGGTTACCGCCCGCTGTGCAACCCAGGTGTGCGACGTCAGACAACGGTGGAGTGTTCTTTT
TGCGTTTTTCTGGCCGGGCGCGCGCGCGTAAGAGACTAAGCCGCGAAAGCGAAAGCAGTA
15 AGTGCTCGCTCCCCGTAGCCGAGGGATCCTTGCTAAGGGTTGCGTTGCGGCGAACCCCGGTT
CGAATCCCGTACTCGGGCCGGCCGACCCGCGGCTAAGGTGTTGGATTGGCCTCCCCCTCGTAT
AAAGACCCCGCTTGCGGATTGACTCCGGACACGGGGACGAGCCCCTTTATTTTTGCTTTCCCCA
GATGCATCCGGTGCTGCGGCAGATGCGCCCCCGCCCCAGCAGCAGCAACAACACCAGCAAGA
GCGGCAGCAACAGCAGCGGGAGTCATGCAGGGCCCCCTCACCCACCCTCGGCGGGCCGGCCAC
20 CTCGGCGTCCGCGGCCGTGTCTGGCGCCTGCGGCGGCGGCGGGGGCCGGCTGACGACCCGAG
GAGCCCCCGCGGCGCAGGGCCAGACACTACCTGGACCTGGAGGAGGGCGAGGGCCTGGCGCGG
CTGGGGGCGCGCTCTCCCGAGCGCCACCCGCGGGTGACGCTGAAGCGCGACTCGCGCGAGGCGT
ACGTGCCTCGGCAGAACCTGTTTCAGGGACCGCGCGGGCGAGGAGCCCGAGGAGATGCGGGACA
GGAGGTTTCAGCGCAGGGCGGGAGCTGCGGCAGGGGCTGAACCGCGAGCGGCTGCTGCGCGAGG
25 AGGACTTTGAGCCCGACGCGCGGACGGGGATCAGCCCCGCGCGCGCGCACGTGGCGGCCGCCG
ACCTGGTGACGGCGTACGAGCAGACGGTGAACCAGGAGATCAACTTCCAAAAGAGTTTCAACA
ACCACGTGCGCACGCTGGTGGCGCGCGAGGAGGTGACCATCGGGCTGATGCACCTGTGGGACTT
TGTAAGCGCGCTGGTGCAGAACCCCAACAGCAAGCCTCTGACGGCGCAGCTGTTCTGTAGTAGT
CAGCACAGCAGGGACAACGAGGCGTTTAGGGACGCGCTGCTGAACATCACCGAGCCCCAGGGT
30 CGGTGGCTGCTGGACCTGATTAACATCCTGCAGAGCATAGTGGTGCAGGAGCGCAGCCTGAGCC
TGGCCGACAAGGTGGCGGCCATCAACTACTCGATGCTGAGCCTGGGCAAGTTTTACGCGCGCAA
GATCTACCAGACGCCGTACGTGCCCATAGACAAGGAGGTGAAGATCGACGGTTTTTACATGCGC
ATGGCGCTGAAGGTGCTCACCTGAGCGACGACCTGGGCGTGTACCGCAACGAGCGCATCCACA
AGGCCGTGAGCGTGAGCCGGCGGCGGAGCTGAGCGACCGCGAGCTGATGCACAGCCTGCAGC
35 GGGCGCTGGCGGGCGCCGCGCAGCGGCGACAGGGAGGCGGAGTCCTACTTCGATGCGGGGGCGG
ACCTGCGCTGGGCGCCCAGCCGGCGGGCCCTGGAGGCGCGGGGGTCCGCGAGGACTATGACG
AGGACGGCGAGGAGGATGAGGAGTACGAGCTAGAGGAGGGCGAGTACCTGGACTAAACCGCG
GGTGGTGTTCGCGTAGATGCAAGACCCGAACGTGGTGGACCCGGCGCTGCGGGCGGCTCTGCA
GAGCCAGCCGTCCGGCCTTAACCTCAGACGACTGGCGACAGGTCATGGACCGCATCATGTGCG
40 CTGACGGCGCGTAACCCGGACGCGTTCGGCGAGCAGCCGAGGCCAACAGGCTCTCCGCCATCC
TGGAGGCGGTGGTGCCTGCGCGCTCGAACCCACGCACGAGAAGGTGCTGGCCATAGTGAACGC

GCTGGCCGAGAACAGGGCCATCCGCCCCGACGAGGCCGGGCTGGTGTACGACGCGCTGCTGCA
GCGCGTGGCCCCGCTACAACAGCGGCAACGTGCAGACCAACCTGGACCGGCTGGTGGGGGACGT
GCGCGAGGCGGTGGCGCAGCGCAGCGCGCGGATCGGCAGGGCAACCTGGGCTCCATGGTGGC
GCTGAATGCCTTCTTGAGCACGCAGCCGGCCAACGTGCCGCGGGGGCAGGAAGACTACACCAA
5 CTTTGTGAGCGCGCTGCGGCTGATGGTGACCGAGACCCCCAGAGCGAGGTGTACCAGTCGGGC
CCGGACTACTTCTTCCAGACCAGCAGACAGGGCCTGCAGACGGTGAACCTGAGCCAGGCTTTCA
AGAACCTGCGGGGGCTGTGGGGCGTGAAGGCGCCCCACCGGCGACCGGGCGACGGTGTCCAGCC
TGCTGACGCCAACTCGCGCCTGCTGCTGCTGCTGATCGCGCCGTTACGGACAGCGGCAGCGT
GTCCCGGGACACCTACCTGGGGCACCTGCTGACCCTGTACC GCGAGGCCATCGGGCAGGCGCAG
10 GTGGACGAGCACACCTTCCAGGAGATCACCAGCGTGAGCCGCGCGCTGGGGCAGGAGGACACG
AGCAGCCTGGAGGCGACTCTGAACTACCTGCTGACCAACCGGCGGCAGAAGATTCCCTCGCTGC
ACAGCCTGACCTCCGAGGAGGAGCGCATCTTGCGCTACGTGCAGCAGAGCGTGAGCCTGAACCT
GATGCGCGACGGGGTGACGCCAGCGTGCGCTGGACATGACCGCGCGCAACATGGAACCGGG
CATGTACGCCGCGCACCGGCCTTACATCAACCGCCTGATGGACTACCTGCATCGCGCGGCGGCC
15 GTGAACCCCGAGTACTTTACCAACGCCATCCTGAACCCGCACTGGCTCCCGCCCGCCGGGTTCTA
CAGCGGGGGCTTCGAGGTCCCGGAGACCAACGATGGCTTCTGTGGGACGACATGGACGACAG
CGTGTCTTCCCCGCGGCCGAGGCGCTGGCGGAAGCGTCCCTGCTGCGTCCCAAGAAGGAGGAG
GAGGAGGAGGCGAGTCGCCGCCGCGGCAGCAGCGGCGTGGCTTCTCTGTCCGAGCTGGGGGCG
GCAGCCGCCGCGCGCCCCGGGTCCCTGGGCGGCAGCCCCTTCCGAGCCTGGTGGGGTCTCTGC
20 ACAGCGAGCGCACCAACCCGCCCTCGGCTGCTGGGCGAGGACGAGTACCTGAATAACTCCCTGCT
GCAGCCGTGCGGGAGAAAAACCTGCCTCCCGCCTTCCCCAACAACGGGATAGAGAGCCTGGTG
GACAAGATGAGCAGATGGAAGACCTATGCGCAGGAGCACAGGGACGCGCCTGCGCTCCGGCCG
CCCACGCGGCCAGCGCCACGACCGGCAGCGGGGGCTGGTGTGGGATGACGAGGACTCCGCG
GACGATAGCAGCGTGCTGGACCTGGGAGGGAGCGGCAACCCGTTGCGGCACCTGCGCCCCCGCC
25 TGGGGAGGATGTTTTAAAAAAGCAAGAAGCATGATGCAAAATTAATAAAAA
CTACCAAGGCCATGGCGACCGAGCGTTGGTTTCTTGTGTTCCCTTCAGTATGCGGCGCGCGGCG
ATGTACCAGGAGGGACCTCCTCCCTCTTACGAGAGCGTGGTGGGCGCGGCGGCGGCGCCCT
CTTCTCCCTTTGCGTCGACGTGCTGGAGCCGCCGTACGTGCCTCCGCGCTACCTGCGGCCTACG
GGGGGGAGAAACAGCATCCGTTACTCGGAGCTGGCGCCCCTGTTGACACCACCCGGGTGTACC
30 TGGTGGACAACAAGTCGGCGGACGTGGCCTCCCTGAACTACCAGAACGACCACAGCAATTTTTT
GACCACGGTCATCCAGAACATGACTACAGCCCGAGCGAGGCCAGCACCCAGACCATCAATCTG
GATGACCGGTCGCACTGGGGCGGCGACCTGAAAACCATCCTGCACACCAACATGCCCAACGTGA
ACGAGTTCATGTTACCAATAAGTTCAAGGCGCGGGTGATGGTGTGCGGCTCGCACACCAAGGA
AGACCGGGTGGAGCTGAAGTACGAGTGGGTGGAGTTCGAGCTGCCAGAGGGCAACTACTCCGA
35 GACCATGACCATTGACCTGATGAACAACGCGATCGTGAGCACTATCTGAAAGTGGGCAGGCAG
AACGGGGTCCCTGGAGAGCGACATCGGGGTCAAGTTCGACACCAGGAACTTCCGCCTGGGGCTGG
ACCCCGTGACCGGGCTGGTTATGCCCGGGGTGTACACCAACGAGGCCTTCCATCCCAGCATCAT
CCTGCTGCCCCGCTGCGGGGTGGACTTCACTTACAGCCGCTGAGCAACCTCCTGGGCATCCGC
AAGCGGCAGCCCTTCCAGGAGGGCTTCAGGATCACCTACGAGGACCTGGAGGGGGGCAACATC
40 CCCGCGCTCCTCGATGTGGAGGCCTACCAGGATAGCTTGAAGGAAAATGAGGCGGGACAGGAG
GATACCGCCCCCGCCGCTCCGCCGCCCGGAGCAGGGCGAGGATGCTGCTGACACCGCGGCCG

CGGACGGGGCAGAGGCCGACCCGCTATGGTGGTGGAGGCTCCCGAGCAGGAGGAGGACATGA
ATGACAGTGCAGTGCAGGAGACACCTTCGTACCCGGGGGAGGAAAAGCAAGCGGAGGCCG
AGGCCGCGGCCGAGGAAAAGCAACTGGCGGCAGCAGCGCGCGCGCGCGCTTGGCCGCGCGC
AGGCTGAGTCTGAGGGGACCAAGCCCGCCAAGGAGCCCGTGATTAAGCCCTGACCGAAGATA
5 GCAAGAAGCGCAGTTACAACCTGCTCAAGGACAGCACCAACACCGCGTACCGCAGCTGGTACCT
GGCCTACAACCTACGGCGACCCGTCGACGGGGGTGCGCTCCTGGACCCTGCTGTGCACGCCGAC
GTGACCTGCGGCTCGGAGCAGGTGTACTGGTTCGCTGCCCCGACATGATGCAAGACCCCGTGACCT
TCCGCTCCACGCGCAGGTCAGCAACTTCCCGGTGGTGGGCGCCGAGCTGCTGCCCCGTGACTC
CAAGAGCTTCTACAACGACCAGGCCGTCTACTCCAGCTCATCCGCCAGTTCACCTCTCTGACCC
10 ACGTGTTCAATCGCTTTCCTGAGAACCAGATTCTGGCGCGCCCGCCCGCCCCACCATCACCACC
GTCAGTGAACCGTTCCTGCTCTCACAGATCACGGGACGCTACCGCTGCGCAACAGCATCGGAG
GAGTCCAGCGAGTGACCGTTACTGACGCCAGACGCCGCACCTGCCCTACGTTTACAAGGCCTT
GGGCATAGTCTCGCCGCGCGTCTTTCCAGCCGCACTTTTTGAGCAACACCACCATCATGTCCAT
CCTGATCTCACCCAGCAATAACTCCGGCTGGGGACTGCTGCGCGCGCCAGCAAGATGTTCCGA
15 GGGGCGAGGAAGCGTTCGAGCAGCACCCCGTGCGCGTGCGCGGGCACTTCCGCGCCCCCTGGG
GAGCGCACAAACGCGGCCGCGCGGGGCGCACACCCTGGACGACGCCATCGACTCGGTGGTGG
AGCAGGCGCGCAACTACAGGCCCCGCGTCTCTACCGTGACGCGGCCATCCAGACCGTGGTGGC
GGGCGCGCGCGGTACGCCAAGCTGAAGAGCCGCGGAAGCGCGTGCCCCGCGCCACC GCCG
CCGACCCGGGGCCCGCCGCAAACGCGCCGCCGCGGCCCTGCTTCGCCGGGCCAAGCGCACGGGC
20 CGCCGCGCCGCCATGAGGGCCGCGCGCCGCTTGCCGCGCGCATCACCGCCGCCACCATGGCCC
CCCGTACCCGAAGACGCGCGGCCGCCGCCGCCGCCGCCATCAGTGACATGGCCAGCAGGCG
CCGGGGCAACGTGTACTGGGTGCGCGACTCGGTGACCGGCACGCGCGTGCCCGTGCGCTTCCGC
CCCCGCGGACTTGAGATGATGTGAAAAACAACACTGAGTCTCCTGCTGTTGTGTGTATCCCA
GCGGCGCGCGCGCGCAGCGTCATGTCCAAGCGCAAAATCAAAGAAGAGATGCTCCAGGTG
25 TCGCGCCGGAGATCTATGGGCCCCGAAGAAGGAAGAGCAGGATTCGAAGCCCCGCAAGATAA
AGCGGGTCAAAAAGAAAAAGAAAGATGATGACGATGCCGATGGGGAGGTGGAGTTCCTGCGCG
CCACGGCGCCAGGCGCCCGGTGCAGTGGAAGGGCCGCGCGTAAAGCGCGTCTGCGCCCCG
GCACCGCGGTGGTCTTACGCCCCGCGAGCGCTCCACCCGGACTTTCAAGCGCGTCTATGACGA
GGTGTACGGCGACGAAGACCTGCTGGAGCAGGCCAACGAGCGCTTCGGAGAGTTTGCTTACGGG
30 AAGCGTCAGCGGGCGCTGGGGAAGGAGACCTGCTGGCGCTGCCGCTGGACCAGGGCAACCCC
ACCCCACTGTAAGCCCGTGACCTGCAGCAGGTGCTGCCGAGCAGCGCACCTCCGAGGCGA
AGCGGGGTCTGAAGCGCGAGGGCGGCGACCTGGCGCCACCGTGACGCTCATGGTGCCCAAGC
GGCAGAGGCTGGAGGATGTGCTGGAGAAAAATGAAAGTAGACCCCGGTCTGCAGCCGGACATCA
GGGTCCGCCCCATCAAGCAGGTGGCGCCGGGCTCGGCGTGACAGCCGTGGACGTGGTCATCCC
35 CACCGGCAACTCCCCCGCGCCGCCACCACTACCGCTGCCTCCACGGACATGGAGACACAGACC
GATCCCGCCGACCGCGACCCGCGACCCGCCGCGACCTCCTCGGCGGAGGTGCAGACGGACC
CCTGGCTGCCCGCGCGATGTACGCTCCCCGCGCGCGTTCGCGGGCGCAGGAAGTACGGCGCCG
CAACGCGCTCCTGCCCCGAGTACGCTTGCATCCTTCCATCGCGCCACCCCGGCTACCGAGGCT
ATACCTACCGCCCGCAAGAGCCAAGGGTTCCACCCCGCGTCCCGGCCGACGCGCGCCGCCAC
40 CACCCGCGCGCCGCCGCGCAGACGCCAGCCCGCACTGGCTCCAGTCTCCGTGAGGAAAGTGGCG
CGCGACGGACACACCCTGGTGCTGCCCAGGGCGCGCTACCACCCAGCATCGTTTAAAAGCCTG

TTGTGGTTCTTGCAGATATGGCCCTCACTTGCCGCCTCCGTTTCCCAGGTGCCGGGATACCGAGGA
GGAAGATCGCGCCGAGGAGGGGTCTGGCCGGCCGCGGCCTGAGCGGAGGCAGCCGCCGCGCG
CACCGGCGGCGACGCGCCACCAGCCGACGCATGCGCGGCGGGGTGCTGCCCCTGTTAATCCCC
TGATCGCCGCGGCGATCGGCGCCGTGCCCAGGATCGCCTCCGTGGCCTTGCAAGCGTCCCAGAG
5 GCATTGACAGACTTGCAAACCTTGCAAATATGGAAAAAAACCCCAATAAAAAAGTCTAGACTC
TCACGCTCGCTTGGTCCTGTGACTATTTTGTAGAATGGAAGACATCAACTTTGCGTCGCTGGCCC
CGCGTCACGGCTCGCGCCCGTTCTTGGGACACTGGAACGATATCGGCACCAGCAACATGAGCGG
TGGCGCCTTCAGTTGGGGCTCTCTGTGGAGCGGCATTAAAAGTATCGGGTCTGCCGTTAAAAATT
ACGGCTCCCGGGCCTGGAACAGCAGCACGGGCCAGATGTTGAGAGACAAGTTGAAAGAGCAGA
10 ACTTCCAGCAGAAGGTGGTGGAGGGCCTGGCCTCCGGCATCAACGGGGTGGTGGACCTGGCCAA
CCAGGCCGTGCAGAATAAGATCAACAGCAGACTGGACCCCCGCGCCCGGTGGAGGAGGTGCC
GCCGGCGCTGGAGACGGTGTCCCCGATGGGCGTGGCGAGAAGCGCCCCGCGGCCCGATAGGGA
AGAGACCACTCTGGTCACGCAGACCGATGAGCCGCCCCCGTATGAGGAGGCCCTGAAGCAAGG
TCTGCCACCACGCGGCCCATCGCGCCCATGGCCACCGGGGTGGTGGGCCGCCACACCCCCGCC
15 ACGCTGGACTTGCTCCGCCCCCGATGTGCCGCAGCAGCAGAAGGCGGCACAGCCGGGCCCGC
CCGCGACCGCCTCCCGTTCCCTCCGCGGTCTCTGCGCCGCGCGGCCAGCGCCCCCGCGGGGG
GGTCGCGAGGCACGGCAACTGGCAGAGCACGCTGAACAGCATCGTGGGTCTGGGGGTGCGGTC
CGTGAAGCGCCGCCGATGCTACTGAATAGCTTAGCTAACGTGTTGTATGTGTGTATGCGCCCTAT
GTCGCCGCCAGAGGAGCTGCTGAGTCGCCGCCGTTCGCGCGCCACCACCACCGCCACTCCGCC
20 CCTCAAGATGGCGACCCCATCGATGATGCCGCAGTGGTCGTACATGCACATCTCGGGCCAGGAC
GCCTCGGAGTACCTGAGCCCCGGGTGGTGCAGTTTCGCCCCGCGCCACCGAGAGCTACTTCAGCC
TGAGTAACAAGTTTAGGAACCCACCGTGGCGCCACGCACGATGTGACCACCGACCGGTCTCA
GCGCCTGACGCTGCGGTTTATTCCCGTGGACCGCGAGGACACCGCGTACTCGTACAAGGCGCGG
TTCACCCTGGCCGTGGGCGACAACCGCGTGTGGACATGGCCTCCACCTACTTTGACATCCGCGG
25 GGTGCTGGACCGGGGTCCCCTTTCAAGCCCTACTCTGGCACCGCCTACAACCTCCCTGGCCCCCA
AGGGCGCTCCCAACTCCTGCGAGTGGGAGCAAGAGGAAACTCAGGCAGTTGAAGAAGCAGCAG
AAGAGGAAGAAGAAGATGCTGACGGTCAAGCTGAGGAAGAGCAAGCAGCTACCAAAAAGACT
CATGTATATGCTCAGGCTCCCCCTTCTGGCGAAAAAATTAGTAAAGATGGTCTGCAAATAGGAA
CGGACGCTACAGCTACAGAACAAAACCTATTTATGCAGACCCTACATTCCAGCCCCGAACCCCA
30 AATCGGGGAGTCCAGTGAATGAGGCAGATGCTACAGTCGCCGCGGTAGAGTGCTAAAGAA
ATCTACTCCCATGAAACCATGCTATGGTTCCTATGCAAGACCCACAAATGCTAATGGAGGTCAG
GGTGTACTAACGGCAAATGCCAGGGACAGCTAGAATCTCAGGTTGAAATGCAATTCTTTTCAA
CTTCTGAAAACGCCCCGTAACGAGGCTAACACATTTCAGCCCCAAATTGGTGTGTATAGTGAGGA
TGTGCACATGGAGACCCCGGATACGCACCTTTCTTACAAGCCCGCAAAAAGCGATGACAATTCA
35 AAAATCATGTGGGTCAGCAGTCCATGCCAACAGACCTAATTACATCGGCTTCAGAGACAACCT
TTATCGGCCTCATGTATTACAATAGCACTGGCAACATGGGAGTGCTTGCAGGTCAGGCCTCTCAG
TTGAATGCAGTGGTGGACTTGCAAGACAGAAACACAGAACTGTCTACCAGCTCTTGCTTGATT
CCATGGGTGACAGAACCAGATACTTTTCCATGTGGAATCAGGCAGTGGACAGTTATGACCCAGA
TGTTAGAATTATTGAAAATCATGGAACGAAGACGAGCTCCCCAACTATTGTTTCCCTCTGGGTG
40 GCATAGGGGTAACCTGACACTTACCAGGCTGTAAAAACCAACAATGGCAATAACGGGGGCCAGG
TGACTTGGACAAAAGATGAAACTTTTGCAGATCGCAATGAAATAGGGGTGGGAAACAATTCGCG

TATGGAGATCAACCTCAGTGCCAACCTGTGGAGAACTTCCTGTACTCCAACGTGGCGCTGTAC
CTACCAGACAAGCTTAAGTACAACCCCTCCAATGTGGACATCTCTGACAACCCCAACACCTACG
ATTACATGAACAAGCGAGTGGTGGCCCCGGGGCTGGTGGACTGCTACATCAACCTGGGCGCGCG
CTGGTCGCTGGACTACATGGACAACGTCAACCCCTTCAACCACCACCGCAATGCGGGCCTGCGC
5 TACCGCTCCATGCTCCTGGGCAACGGGCGCTACGTGCCCTTCCACATCCAGGTGCCCCAGAAGTT
CTTTGCCATCAAGAACCTCCTCCTGCGGGCTCCTACACCTACGAGTGGAACCTTCAGGAAGG
ATGTCAACATGGTCCCTCCAGAGCTCTCTGGGTAACGATCTCAGGGTGGACGGGGCCAGCATCAA
GTTTCGAGAGCATCTGCCTCTACGCCACCTTCTTCCCCATGGCCCACAACACGGGCCTCCACGCTCG
AGGCCATGCTCAGGAACGACACCAACGACCAGTCCCTTCAATGACTACCTCTCCGCCGCCAACAT
10 GCTCTACCCCATACCCGCCAACGCCACCAACGTCCCCATCTCCATCCCCCTCGCGCAACTGGGCGG
CCTTCCGCGGCTGGGCCTTACCCGCCTCAAGACCAAGGAGACCCCTCCTGGGCTCGGGATTG
GACCCCTACTACACCTACTCGGGCTCCATTCCCTACCTGGACGGCACCTTCTACCTCAACCACAC
TTTCAAGAAGGTCTCGGTACCTTCGACTCCTCGGTGAGCTGGCCGGGCAACGACCGTCTGCTCA
CCCCCAACGAGTTCGAGATCAAGCGCTCGGTGACGGGGAGGGCTACAACGTGGCCCAGTGCA
15 ACATGACCAAGGACTGGTTCCTGGTCCAGATGCTGGCCAACTACAACATCGGCTACCAGGGCTT
CTACATCCCAGAGAGCTACAAGGACAGGATGTACTCCTTCTTCAGGAACCTTCCAGCCCATGAGC
CGGCAGGTGGTGGACCAGACCAAGTACAAGGACTACCAGGAGGTGGGCATCATCCACCAGCAC
AACAACTCGGGCTTCGTGGGTACCTCGCCCCACCATGCGCGAGGGACAGGCCTACCCCGCCA
ACTTCCCCTATCCGCTCATAGGCAAGACCGCGGTGACAGCATCACCCAGAAAAAGTTCTCTG
20 CGACCGCACCTCTGGCGCATCCCCCTTCTCCAGCAACTTCATGTCCATGGGTGCGCTCTCGGACC
TGGGCCAGAACTTGCTCTACGCCAACTCCGCCCACGCCCTCGACATGACCTTCGAGGTGACCCC
ATGGACGAGCCACCCTTCTCTATGTTCTGTTGGAAGTCTTTGACGTGGTCCGGGTCCACCAGCC
GCACCGCGGCGTCATCGAGACCGGTGTACCTGCGTACGCCCTTCTCGGCCGGCAACGCCACCACC
TAAAGAAGCAAGCCGAGTCATCGCCGCCTGCATGCCGTGCGGTTCCACCGAGCAAGAGCTCAG
25 GGCCATCGTCAGAGACCTGGGATGCGGGCCCTATTTTTTGGGCACCTTCGACAAGCGCTTCCCTG
GCTTTGTCTCCCCACACAAGCTGGCCTGCGCCATCGTCAACACGGCCGGCCGCGAGACCGGGGG
CGTGCACTGGCTGGCCTTCGCCTGGAACCCGCGCTCCAAAACATGCTTCCTCTTTGACCCCTTCG
GCTTTTCGGACCAGCGGCTCAAGCAAATCTACGAGTTCGAGTACGAGGGCTTGCTGCGTGCAG
CGCCATCGCCTCCTCGCCCGACCGCTGCGTCACCCTCGAAAAGTCCACCCAGACCGTGCAGGGG
30 CCCGACTCGGCCGCTGCGGTCTCTTCTGCTGCATGTTTCTGCACGCCTTTGTGCACTGGCCTCAG
AGTCCCATGGACCGCAACCCCAACCATGAACTTGCTGACGGGGGTGCCCAACTCCATGCTCCAGA
GCCCCCAGGTGAGACCCACCCTGCGCCGCAACCAGGAGCAGCTCTACAGCTTCCTGGAGCGCCA
CTCGCCTTACTTCCGCCGCCACAGCGCACAGATCAGGAGGGCCACCTCCTTCTGCCACTTGCAAG
AGATGCAAGAAGGGTAATAACGATGTACACACTTTTTTTCTCAATAAATGGCATCTTTTTATTTA
35 TACAAGCTCTCTGGGGTATTCAATTTCCCAACCACCCCGCCGTTGTCGCCATCTGGCTCTATTTAG
AAATCGAAAGGGTTCTGCCGGGAGTCGCCGTGCGCCACGGGCAGGGACACGTTGCGATACTGGT
AGCGGGTGCCCCACTTGAACCTCGGGCACCACCAGGCGAGGCAGCTCGGGGAAGTTTTGCTCCA
CAGGCTGCGGGTCAGCACCAGCGGTTTCATCAGGTGCGGCGCCGAGATCTTGAAGTCGCAGTTG
GGGCCGCCGCCCTGCGCGCGCGAGTTGCGGTACACCGGGTTGCAGCACTGGAACACCAACAGCG
40 CCGGGTGCTTCACGCTGGCCAGCACGCTGCGGTGCGAGATCAGCTCGGCGTCCAGGTCTCCGC
GTTGCTCAGCGCGAACGGGGTCATCTTGGGCACTTGCCGCCCCAGGAAGGGCGCGTGCCCCGGT

TTCGAGTTGCAGTCGCAGCGCAGCGGGATCAGCAGGTGCCCCGTGCCCGGACTCGGCGTTGGGGT
ACAGCGCGCGCATGAAGGCCTGCATCTGGCGGAAGGCCATCTGGGCCCTGGCGCCCTCCGAGAA
GAACATGCCGCAGGACTTGCCCGAGAACTGGTTTTCGGGGCAGCTGGCGTCGTGCAGGCAGCAG
CGCGCGTCGGTGTGGCGATCTGCACCACGTTGCGCCCCACCGGTTCTTCACGATCTTGGCCTT
5 GGACGATTGCTCCTTCAGCGCGCGCTGCCCGTTCTCGCTGGTCACATCCATCTCGATCACATGTT
CCTTGTTACCATGCTGCTGCCGTGCAGACACTTCAGCTCGCCCTCCGTCTCGGTGCAGCGGTGC
TGCCACAGCGCGCAGCCCGTGGGCTCGAAAGACTTGTAGGTCACCTCCGCGAAGGACTGCAGGT
ACCCCTGCAAAAAGCGGCCCATCATGGTCACGAAGGTCCTTGTTGCTGCTGAAGGTCAGCTGCAG
CCCGCGGTGCTCCTCGTTCAGCCAGGTCTTGACACAGGCCGCCAGCGCCTCCACCTGGTCGGGCA
10 GCATCTTGAAGTTCACCTTCAGCTCATTCTCCACGTGGTACTTGTCCATCAGCGTGCAGCGCCGCC
TCCATGCCCTTCTCCCAGGCCGACACCAGCGGCAGGCTCACGGGGTTCTTACCATCACCGTGGC
CGCCGCCTCCGCCGCGCTTTCGCTTTCGCCCGCGTGTCTCTTCTCTTCTCTTCTCTCTCGCCG
CCGCCCCACTCGCAGCCCCCGCACCCAGGGGTGCTCTTCTGTCAGGCGCTGCACCTTGCGCTTGCC
GTTGCGCCCCGTGTTGATGCGCACGGGCGGGTTGCTGAAGCCCACCATCACCAGCGCGGCCTCTT
15 CTTGCTCGTCCTCGCTGTCCAGAATGACCTCCGGGGAGGGGGGGTTGGTCATCCTCAGTACCGA
GGCACGCTTCTTTTTTCTTCTGCGGGCGTTTCGCCAGCTCCGCGGCTGCGGCCGCTGCCGAGGTCG
AAGGCCGAGGGCTGGGCGTGCAGCGCACCCAGCGCTCCTGCGAGCCGTCCTCGTCTCTCTCGGA
CTCGAGACGGAGGCGGGCCCGCTTCTTCGGGGCGCGCGGGGCGGCGGAGGCGGCGGCGGCGA
CGGAGACGGGGACGAGACATCGTCCAGGGTGGGTGGACGGCGGGCCGCGCCGCTCCGCGCTC
20 GGGGGTGGTCTCGCGCTGGTCCTTCTCCGACTGGCCATCTCCCACTGCTCCTTCTCCTATAGGC
AGAAAGAGATCATGGAGTCTCTCATGCGAGTCGAGAAGGAGGAGGACAGCCTAACCGCCCCCT
CTGAGCCCTCCACCACCGCCGCCACCACCGCCAATGCCCGCGGACGACGCGCCACCGAGAC
CACCGCCAGTACCACCCTCCCCAGCGACGCACCCCGCTCGAGAATGAAGTGCTGATCGAGCAG
GACCCGGGTTTTGTGAGCGGAGAGGAGGATGAGGTGGATGAGAAGGAGAAGGAGGAGGTCCGC
25 GCCTCAGTGCCAAAAGAGGATAAAAAGCAAGACCAGGACGACGACGATAAGGATGAGACAGC
AGTCGGGCGGGGGAACGGAAGCCATGATGCTGATGACGGCTACCTAGACGTGGGAGACGACGT
GCTGCTTAAGCACCTGCACCGCCAGTGCGTCATCGTCTGCGACGCGCTGCAGGAGCGCTGCGAA
GTGCCCTTGACGTGGCGGAGGTCAGCCGCGCCTACGAGCGGCACCTCTTCGCGCCGCACGTGC
CCCCAAGCGCCGGGAGAACGGCACCTGCGAGCCCAACCCGCGTCTCAACTTCTACCGGTCTT
30 CGCGGTACCCGAGGTGCTGGCCACCTACCACATCTTTTTTCCAAAAGTCAAGATCCCCCTCTCCT
GCCGCGCAACCGCACCCGCGCCGACAAAACCTGACCCTGCGGCAGGGCGCCACATACCTGA
TATCGCCTCTCTGGAGGAAGTGCCCAAGATCTTCGAGGGTCTCGGTGCGGACGAGAAACGGGCG
GCGAACGCTCTGCACGGAGACAGCGAAAACGAGAGTCACTCGGGGGTGTGTTGGAGCTCGAG
GGCGACAACGCGCGCCTGGCCGTACTCAAGCGCAGCATAGAGGTCACCCACTTTGCCTACCCGG
35 CGCTCAACCTGCCCCCAAGGTATGAGTGTGGTCATGGGCGAGCTCATCATGCGCCGCGCCA
GCCCCTGGCCGCGGATGCAAACTTGCAAGAGTCTCCGAGGAAGGCCTGCCGCGGTACGCGAC
GAGCAGCTGGCGCGCTGGCTGGAGACCCGCGACCCCGCGCAGCTGGAGGAGCGGCGCAAGCTC
ATGATGGCCGCGGTGCTGGTCACCGTGGAGCTCGAGTGTCTGCAGCGCTTCTTCGCGGACCCCG
AGATGCAGCGCAAGCTCGAGGAGACCCTGCACTACACCTCCGCCAGGGCTACGTGCGCCAGGC
40 CTGCAAGATCTCCAACGTGGAGCTCTGCAACCTGGTCTCCTACCTGGGCATCCTGCACGAGAAC
CGCCTCGGGCAGAACGTCTGCACTCCACCCTCAAAGGGGAGGCGCGCCGCGACTACATCCGCG

ACTGCGCCTACCTCTTCTCTGCTACACCTGGCAGACGGCCATGGGGGTCTGGCAGCAGTGCCTG
GAGGAGCGCAACCTCAAGGAGCTGGAAAAGCTCCTCAAGCGCACCTCAGGGACCTCTGGACG
GGCTTCAACGAGCGCTCGGTGGCCGCCGCGCTGGCGGACATCATCTTTCCCGAGCGCCTGCTCA
AGACCCTGCAGCAGGGCCTGCCCCACTTCACCAGCCAGAGCATGCTGCAGAACTTCAGGACTTT
5 CATCCTGGAGCGCTCGGGCATCCTGCCGGCCACTTGCTGCGCGCTGCCCAGCGACTTCGTGCCCA
TCAAGTACAGGGAGTGCCCGCCGCGCTCTGGGGCCACTGCTACCTCTTCCAGCTGGCCAACTA
CCTCGCCTACCACTCGGACCTCATGGAAGACGTGAGCGGCGAGGGCCTGCTCGAGTGCCACTGC
CGCTGCAACCTCTGCACGCCCCACCGCTCTCTAGTCTGCAACCCGCAGCTGCTCAGCGAGAGTCA
GATTATCGGTACCTTCGAGCTGCAGGGTCCCTCGCCTGACGAGAAGTCCGCGGCTCCAGGGCTG
10 AAACCTACTCCGGGGCTGTGGACTTCCGCCTACCTACGCAAAATTTGTACCTGAGGACTACCACGC
CCACGAGATCAGGTTCTACGAAGACCAATCCCCCCCCGCCAAGGCGGAGCTACCCGCTGCGTC
ATCACCCAGGGGCACATCCTGGGCCAATTGCAAGCCATCAACAAAGCCCGCCGAGAGTTCTTGC
TGAAAAGGGTCGGGGGGTGTACCTGGACCCCCAGTCCGGCGAGGAGCTAAACCCGCTACCCCC
GCCGCGCCCCCAGCAGCGGGACCTTGCTTCCCAGGATGGCACCCAGAAAGAAGCAGCAGCCGC
15 CGCCGCGCCCGCAGCCATACATGCTTCTGGAGGAAGAGGAGGAGGACTGGGACAGTCAGGCAG
AGGAGGTTTCGGACGAGGAGCAGGAGGAGATGATGGAAGACTGGGAGGAGGACAGCAGCCTA
GACGAGGAAGCTTCAGAGGCCGAAGAGGTGGCAGACGCAACACCATCGCCCTCGGTGCGAGCC
CCCTCGCCGGGGCCCCCTGAAATCCTCCGAACCCAGCACCCAGCGCTATAACCTCCGCTCCTCCGGC
GCCGGCGCCACCCGCCCCGAGACCCAACCGTAGATGGGACACCACAGGAACCGGGGTCTGGTAA
20 GTCCAAGTGCCCGCCGCCGCCACCGCAGCAGCAGCAGCAGCAGCGCCAGGGCTACCGCTCGTGG
CGCGGGCACAAGAACGCCATAGTCGCCTGCTTGCAAGACTGCGGGGGCAACATCTCTTTCCGCC
GCCGCTTCTGCTATTCCACCACGGGGTCGCCTTTCCCCGCAATGTCCTGCATTACTACCGTCATC
TCTACAGCCCCTACTGCAGCGGCGACCCAGAGGCGGCAGCGGCAGCCACAGCGGCGACCACCA
CCTAGGAAGATATCCTCCGCGGGCAAGACAGCGGCAGCAGCGGCCAGGAGACCCGCGGCAGCA
25 GCGGGCGGAGCGGTGGGCGCACTGCGCCTCTCGCCCAACGAACCCCTCTCGACCCGGGAGCTCA
GACACAGGATCTTCCCCACTTTGTATGCCATCTTCCAACAGAGCAGAGGCCAGGAGCAGGAGCT
GAAAATAAAAAACAGATCTCTGCGCTCCCTCACCCGAGCTGTCTGTATCACAAAAGCGAAGAT
CAGCTTCGGCGCACGCTGGAGGACGCGGAGGCACTCTTCAGCAATACTGCGCGCTCACTCTTA
AAGACTAGCTCCGCGCCCTTCTCGAATTTAGGCGGGAGAAAACCTACGTCATCGCCGGCCGCGC
30 CCAGCCCGCCAGCCGAGATGAGCAAAGAGATTCCACGCCATACATGTGGAGCTACCAGCCGC
AGATGGGACTCGCGGCGGGAGCGGCCAGGACTACTCCACCCGCATGAACTACATGAGCGCGG
GACCCACATGATCTCACAGGTCAACGGGATCCGCGCCCAGCGAAACCAATACTGCTGGAACA
GGCGGCCATACCGCCACGCCCCGCCATAATCTCAACCCCCGAAATTGGCCCGCCGCCCTCGTG
TACCAGGAAACCCCTCCGCCACCACCGTACTACTTCCGCGTGACGCCAGGCCGAAGTCCAGA
35 TGAATAACTCAGGGGCGCAGCTCGCGGGCGGCTTTTCGTACGGGGCGCGGCCGCTCCGACCAGG
TATAAGACACCTGATGATCAGAGGCCGAGGTATCCAGCTCAACGACGAGTCGGTGAGCTCTTCG
CTCGGTCTCCGTCCGACGGAACCTTCCAGCTCGCCGATCCGCGCGCTCTTCGTTACGCCCCG
CCAGGCGTACCTGACTCTGCAGACCTCGTCCTCGGAGCCCCGCTCCGGCGGCATCGGAACCCCTC
CAGTTCTGTGGAGGAGTTCGTGCCCTCGGTCTACTTCAACCCCTTCTCGGGACCTCCCGGACGCTA
40 CCCCCAGCAGTTCATTCCGAACCTTTGACGCGGTGAAGGACTCGGCGGACGGCTACGACTGAATG
TCAGGTGTCGAGGCAGAGCAGCTTCGCCTGAGACACCTCGAGCACTGCCGCGCCACAAGTGCT

TCGCCCCGCGTTCTGGTGAGTTCTGCTACTTTCAGCTACCCGAGGAGCATAACCGAGGGGCGCGGC
GCACGGCGTCCGCCTGACCACCCAGGGCGAGGTTACCTGTTCCCTCATCCGGGAGTTTACCCTCC
GTCCCCTGCTAGTGGAGCGGGAGCGGGGTCCCTGTGTCTAACTATCGCCTGCAACTGCCCTAAC
CCTGGATTACATCAAGATCTTTGCTGTCATCTCTGTGCTGAGTTTAATAAACGCTGAGATCAGAA
5 TCTACTGGGGCTCCTGTGCGCCATCCTGTGAACGCCACCGTCTTCACCCACCCCGACCAGGCCAG
GCGAACCTCACCTGCGGTCTGCATCGGAGGGCCAAGAAGTACCTCACCTGGTACTTCAACGGCA
CCCCCTTTGTGGTTTACAACAGCTTCGACGGGGACGGAGTCTCCCTGAAAGACCAGCTCTCCGGT
CTCAGCTACTCCATCCACAAGAACACCACCCTCCAACCTTTCCCTCCCTACCTGCCGGGAACCTA
CGAGTGCGTCACCGGCCGCTGCACCCACCTCACCCGCTGATCGTAAACCAGAGCTTTCGGGA
10 ACAGATAACTCCCTCTTCCCCAGAACAGGAGGTGAGCTCAGGAAACTCCCCGGGGACCAGGGCG
GAGACGTACCTTCGACCCTTGTGGGGTTAGGATTTTTTATTACCGGGTTGCTGGCTCTTTAATCA
AAGTTTCCTTGAGATTTGTTCTTTCCCTTCTACGTGTATGAACACCTCAACCTCCAATAACTCTACC
CTTCTCTCGGAATCAGGTGACTTCTCTGAAATCGGGCTTGGTGTGCTGCTTACTCTGTGATTTTT
TTCCTTATCATACTCAGCCTTCTGTGCCCTCAGGCTCGCCGCTGCTGCGCACACATCTATATCTAC
15 TGCTGGTTGCTCAAGTGCAGGGGTCGCCACCCAAGATGAACAGGTACATGGTCCTATCGATCCT
AGGCCTGCTGGCCCTGGCGGCCTGCAGCGCCGCCAAAAAAGAGATTACCTTTGAGGAGCCCGCT
TGCAATGTAACCTTCAAGCCCCGAGGGTGACCAATGCACCACCCTCGTCAAATGCGTTACCAATC
ATGAGAGGCTGCGCATCGACTACAAAAACAAAACCTGGCCAGTTTTCGGTCTATAGTGTGTTTAC
GCCCCGAGACCCCTCTAACTACTCTGTCACCGTCTTCCAGGGCGGACAGTCTAAGATATTCAATT
20 ACACTTTCCCTTTTTATGAGTTATGCGATGCGGTTCATGTACATGTCAAAACAGTACAACCTGTGG
CCTCCCTCTCCCCAGGCGTGTGTGGAAAATACTGGGTCTTACTGCTGTATGGCTTTCGCAATCAC
TACGCTCGCTCTAATCTGCACGGTGCTATACATAAAATTACAGGCAGAGGCGAATCTTTATCGATG
AAAAGAAAATGCCTTGATCGCTAACACCGGCTTTCTATCTGCAGAATGAATGCAATCACCTCCCT
ACTAATCACCAACCACCTCCTTGCGATTGCCATGGGTTGACACGAATCGAAGTGCCAGTGGGG
25 TCCAATGTCAACATGGTGGGCCCCGCCGGAATTCCACCCTCATGTGGGAAAAATTTGTCCGCA
ATCAATGGGTTTCATTTCTGCTCTAACCGAATCAGTATCAAGCCCAGAGCCATCTGCGATGGGCA
AAATCTAACTCTGATCAATGTGCAAATGATGGATGCTGGGTACTATTACGGGCAGCGGGGAGAA
ATCATTAACTACTGGCGACCCCAAGGACTACATGCTGCATGTAGTCGAGGCACTTCCCACTAC
CACCCCCACTACCACCTCTCCCAACCACCACCACCTACTACTACTACTACTACTACTACTA
30 CTACCACTACCGCTGCCCCGCCATACCCGCAAAAGCACCATGATTAGCACAAAGCCCCCTCGTGC
TCACTCCCACGCCGCGGGCCCATCGGTGCGACCTCAGAAACCACCGAGCTTTGCTTCTGCCAAT
GCACTAACGCCAGCGCTCATGAACTGTTGACCTGGAGAATGAGGATGTCCAGCAGAGCTCCGC
TTGCCTGACCCAGGAGGCTGTGGAGCCCGTTGCCCTGAAGCAGATCGGTGATTCAATAATTGAC
TCTTCTCTTTTGCCACTCCCGAATACCCTCCCGATTCTACTTTCCACATCACGGGTACCAAAGAC
35 CCTAACCTCTCTTTCTACCTGATGCTGCTGCTGTATCTCTGTGGTCTCTTCCGCGCTGATGTTA
CTGGGGATGTTCTGCTGCCTGATCTGCCGAGAAAAGAGAAAAGCTCGCTCTCAGGGCCAACCAC
TGATGCCCTTCCCTACCCCCCGGATTTTGCAGATAACAAGATATGAGCTCGCTGCTGACACTAA
CCGCTTTACTAGCCTGCGCTCTAACCTTGTGCTTGGACTCGAGATTCCACAATGTCACAGCT
GTGGCAGGAGAAAATGTTACTTTCAACTCCACGGCCGATACCCAGTGGTCGTGGAGTGGCTCAG
40 GTAGCTACTTAACTATCTGCAATAGCTCCACTTCCCCCGGCATATCCCCAACCAAGTACCAATGC
AATGCCAGCCTGTTACCCCTCATCAACGCTTCCACCCTGGACAATGGACTCTATGTAGGCTATGT

ACCCTTTGGTGGGCAAGGAAAGACCCACGCTTACAACCTGGAAGTTCGCCAGCCCAGAACCCT
ACCCAAGCTTCTCCACCACCACCACCACCACCACCATCACCAGCAGCAGCAGCAGCAGCAGCC
ACAGCAGCAGCAGCAGATTATTGACTTTGGTTTTGGCCAGCTCATCTGCCGCTACCCAGGCCATC
TACAGCTCTGTGCCCCGAAACCACTCAGATCCACCGCCCAGAAACGACCACCGCCACCACCTAC
5 ACACCTCCAGCGATCAGATGCCGACCAACATCACCCCTTGGCTCTTCAAATGGGACTTACAAG
CCCCACTCCAAAACAGTGGATGCGGCCGAGGTCTCCGCCCTCGTCAATGACTGGGCGGGGCTG
GGAATGTGGTGGTTCGCCATAGGCATGATGGCGCTCTGCCTGCTTCTGCTCTGGCTCATCTGCTG
CCTCCACCGCAGGCGAGCCAGACCCCCATCTATAGACCCATCATTGTCCTGAACCCCGATAAT
GATGGGATCCATAGATTGGATGGCCTGAAAAACCTACTTTTTTCTTTTACAGTATGATAAATTGA
10 GACATGCCTCGCATTTTCTTGTACATGTTCTTCTCCACCTTTTCTGGGGTGTCTACGCTGGCC
GCTGTGTCTCACCTGGAGGTAGACTGCCTCTCACCTTCACTGTCTACCTGCTTTACGGATTGGTC
ACCCTCACTCTCATCTGCAGCCTAATCACAGTAATCATCGCCTTCATCCAGTGCATTGATTACAT
CTGTGTGCGCCTCGCATACTTCAGACACCACCCGAGTACCGAGACAGGAACATTGCCCCAATT
CTAAGACTGCTCTAATCATGCATAAGACTGTGATCTGCCTTCTGATCCTCTGCATCCTGCCACC
15 CTCACCTCCTGCCAGTACACCACAAAATCTCCGCGCAAAAGACATGCCTCCTGCCGCTTCACCCA
ACTGTGGAATATACCCAAATGCTACAACGAAAAGAGCGAGCTCTCCGAAGCTTGGCTGTATGGG
GTCATCTGTGTCTTAGTTTTCTGCAGCACTGTCTTTGCCCTCATAATCTACCCCTACTTTGATTTG
GGATGGAACGCGATCGATGCCATGAATTACCCACCTTTCCCGCACCCGAGATAATTCCACTGC
GACAAGTTGTACCCGTTGTCGTTAATCAACGCCCCCATCCCCTACGCCCCTGAAATCAGCTAC
20 TTTAACCTAACAGGCGGAGATGACTGACGCCCTAGATCTAGAAATGGACGGCATCAGTACCGAG
CAGCGTCTCCTAGAGAGGCGCAGGCAGGCGGCTGAGCAAGAGCGCCTCAATCAGGAGCTCCGA
GATCTCGTTAACCTGCACCAGTGCAAAAGAGGCATCTTTTGTCTGGTAAAGCAGGCCAAAGTCA
CCTACGAGAAGACCGGCAACAGCCACCGCCTCAGTTACAAATTGCCACCCAGCGCCAGAAGCT
GGTGTCTATGGTGGGTGAGAATCCCATCACCGTCACCCAGCACTCGGTAGAGACCGAGGGGTGT
25 CTGCACTCCCCCTGTCGGGGTCCAGAAGACCTCTGCACCCTGGTAAAGACCCTGTGCGGTCTCAG
AGATTTAGTCCCCCTTAATACTAATCAAACACTGGAATCAATAAAAAGAATCACTTACTTAAATC
AGACAGCAGGTCTCTGTCCAGTTTATTCAGCAGCACCTCCTTCCCCCTCCTCCCACTCTGGTACT
CCAAACGCCTTCTGGCGGCAAACTTCCTCCACACCCTGAAGGGAATGTCAGATTCTTGCTCCTGT
CCCTCCGCACCCACTATCTTCATGTTGTTGCAGATGAAGCGCACCAAAACGTCTGACGAGAGCTT
30 CAACCCCGTGTACCCCTATGACACGGAAAGCGGCCCTCCCTCCGTCCCTTTCTCACCCCTCCCT
TCGTGTCTCCCGATGGATTCCAAGAAAGTCCCCCGGGTCTGTCTGAACCTGGCCGAGCCC
CTGGTCACTTCCCACGGCATGCTCGCCCTGAAAATGGGAAGTGGCCTCTCCCTGGACGACGCTG
GCAACCTCACCTCTCAAGATATCACCAACCGCTAGCCCTCCCTCAAAAAAACCAAGACCAACCT
CAGCCTAGAAACCTCATCCCCCTAACTGTGAGCACCTCAGGCGCCCTCACCGTAGCAGCCGCC
35 GCTCCCCTGGCGGTGGCCGGCACCTCCCTACCATGCAATCAGAGGCCCCCTGACAGTACAGG
ATGCAAAACTCACCTGGCCACCAAAGGCCCCCTGACCGTGTCTGAAGGCAAACTGGCCTTGCA
AACATCGGCCCCGCTGACGGCCGCTGACAGCAGCACCTCAGTCAAGTGCACACCAACCCCTT
AGCACAAGCAATGGCAGCTTGGGTATTGACATGCAAGCCCCCATTTACACCACCAATGGAAAAC
TAGGACTTAACTTTGGCGCTCCCCTGCATGTGGTAGACAGCCTAAATGCACTGACTGTAGTTACT
40 GGCCAAGGTCTTACGATAAACGGAACAGCCCTACAACTAGAGTCTCAGGTGCCCTCAACTATG
ACACATCAGGAAACCTAGAATTGAGAGCTGCAGGGGGTATGCGAGTTGATGCAAATGGTCAACT

TATCCTTGATGTAGCTTACCCATTTGATGCACAAAACAATCTCAGCCTTAGGCTTGGACAGGGAC
CCCTGTTTGTAACTCTGCCACAACCTGGATGTAACTACAACAGAGGCCTCTACCTGTTTACA
TCTGGAAATACCAAAAAGCTAGAAGTTAATATCAAAACAGCCAAGGGTCTCATTATGATGACA
CTGCTATAGCAATCAATGCGGGTGATGGGCTACAGTTTGACTCAGGCTCAGATACAAATCCATT
5 AAAAACTAACTTGGATTAGGACTGGATTATGACTCCAGCAGAGCCATAATTGCTAAACTGGGA
ACTGGCCTAAGCTTTGACAACACAGGTGCCATCACAGTAGGCAACAAAAATGATGACAAGCTTA
CCTTGTGGACCACACCAGACCCATCCCCTAACTGTAGAATCTATTCAGAGAAAGATGCTAAATT
CACACTTGTTTTGACTAAATGCGGCAGTCAGGTGTTGGCCAGCGTTTCTGTTTTATCTGTAAAAG
GTAGCCTTGCGCCCATCAGTGGCACAGTAACTAGTGCTCAGATTGTCCTCAGATTTGATGAAAAT
10 GGAGTTCTACTAAGCAATTCTTCCCTTGACCCTCAATACTGGAACCTACAGAAAAGGTGACCTTAC
AGAGGGCACTGCATATACCAACGCAGTGGGATTTATGCCCAACCTCACAGCATACCCAAAAACA
CAGAGCCAAACTGCTAAAAGCAACATTGTAAGTCAGGTTTACTTGAATGGGGACAAATCCAAAC
CCATGACCCTCACCATTAACCTCAATGGAACCTAATGAAACAGGAGATGCCACAGTAAGCACTTA
CTCCATGTCAATTCTCATGGAACCTGGAATGGAAGTAATTACATTAATGAAACGTTCCAAACCAACT
15 CCTTCACCTTCTCCTACATCGCCCAAGAATAAAAAGCATGACGCTGTTGATTTGATTCAATGTGT
TTCTGTTTTATTTTCAAGCACAACAAAATCATTCAAGTCATTCTTCCATCTTAGCTTAATAGACAC
AGTAGCTTAATAGACCCAGTAGTGCAAAGCCCCATTCTAGCTTATAACTAGTGGAGAAGTACTC
GCCTACATGGGGGTAGAGTCATAATCGTGCATCAGGATAGGGCGGTGGTGTCTGCAGCAGCGCGC
GAATAAACTGCTGCCGCCGCCGCTCCGTCCTGCAGGAATACAACATGGCAGTGGTCTCCTCAGC
20 GATGATTTCGACCGCCCGCAGCATAAGGCGCCTTGTCCTCCGGGCACAGCAGCGCACCCCTGATC
TCACTTAAATCAGCACAGTAACTGCAGCACAGCACCACAATATTGTTCAAAATCCCACAGTGCA
AGGCGCTGTATCCAAAGCTCATGGCGGGGACCACAGAACCCACGTGGCCATCATACCACAAGCG
CAGGTAGATTAAGTGGCGACCCCTCATAAACACGCTGGACATAAACATTACCTCTTTTGGCATGT
TGTAATTCACCACCTCCCGGTACCATATAAACCTCTGATTAAACATGGCGCCATCCACCACCATC
25 CTAAACCAGCTGGCCAAAACCTGCCCGCCGGCTATACACTGCAGGGAACCGGGACTGGAACAAT
GACAGTGGAGAGCCCAGGACTCGTAACCATGGATCATCATGCTCGTCATGATATCAATGTTGGC
ACAACACAGGCACACGTGCATACACTTCCTCAGGATTACAAGCTCCTCCCGCGTTAGAACCATA
TCCCAGGGAACAACCCATTCTGAATCAGCGTAAATCCCACACTGCAGGGAAGACCTCGCACGT
AACTCACGTTGTGCATTGTCAAAGTGTTACATTGGGGCAGCAGCGGATGATCCTCCAGTATGGTA
30 GCGCGGGTTTCTGTCTCAAAAGGAGGTAGACGATCCCTACTGTACGGAGTGCGCCGAGACAACC
GAGATCGTGTGGTCGTAGTGTTCATGCCAAATGGAACGCCGGACGTAGTCATATTTCTGAAGT
CTTAGATCTCTCAACGCAGCACCAGCACCAACACTTCGCAGTGTAAGGAGGCAAGTGCCGAGAG
AGTATATATAGGAATAAAAAGTGACGTAAACGGGCAAAGTCCAAAAAACGCCAGAAAAACCG
CACGCGAACCTACGCCCCGAAACGAAAGCCAAAAAACACTAGACACTCCCTTCCGGCGTCAACT
35 TCCGCTTTCCACGCTACGTCACTTGCCCCAGTCAAACAACTACATATCCCGAACTTCCAAGTC
GCCACGCCCCAAACACCGCCTACACCTCCCCGCCCGCCGCCCCGCCCCAAACCCGCTCCCGC
CCCGCGCCCCGCCCGCGCCGCCCCATCTCATTATCATATTGGCTTCAATCCAAAATAAGGTATAT
TATTGATGATG

40

SEQ ID NO: 12 -

Polynucleotide sequence encoding the CASI promoter

GGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCA
ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGT
ACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCCGCTGGCATTATGCCAGTACATGACCTT
5 ATGGGACTTTCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTCGAGGTGA
GCCCCACGTTCTGCTTCACTCTCCCCATCTCCCCCCCCCTCCCCACCCCCAATTTTGTATTTATTTAT
TTTTTAATTATTTTGTGCAGCGATGGGGGCGGGGGGGGGGGGGGGCGCGGCCAGGCGGGGCG
GGGCGGGGCGAGGGGCGGGGCGGGGCGAGGCGGAGAGGTGCGGCGGCAGCCAATCAGAGCGG
CGCGCTCCGAAAGTTTCCTTTTATGGCGAGGCGGGCGGCGGCGGCCCTATAAAAAGCGAAGC
10 GCTCCCTATCAGTGATAGAGATCTCCCTATCAGTGATAGAGATCGTCGACGAGCTCGCGGCGGG
CGGGAGTCGCTGCGCGCTGCCTTCGCCCCGTGCCCCGCTCCGCCGCCGCTCGCGCCGCCGCC
CGGCTCTGACTGACCGCGTTACTAAAAACAGGTAAGTCCGGCCTCCGCGCCGGGTTTTGGCGCCTC
CCGCGGGCGCCCCCTCCTCACGGCGAGCGTGCCACGTCAGACGAAGGGCGCAGCGAGCGTCC
TGATCCTTCCGCCCGACGCTCAGGACAGCGGCCCGCTGCTCATAAGACTCGGCCCTAGAACCC
15 CAGTATCAGCAGAAGGACATTTTAGGACGGGACTTGGGTGACTCTAGGGCACTGGTTTTCTTTCC
AGAGAGCGGAACAGGCGAGGAAAAAGTAGTCCCTTCTCGGCGATTCTGCGGAGGGATCTCCGTG
GGGCGGTGAACGCCGATGATGCCTCTACTAACCATGTTTCATGTTTCTTTTTTTTCTACAGGTCC
TGGGTGACGAACAG

20 **SEQ ID NO: 13 - Ad5orf6 primer 1 polynucleotide sequence**

ATACGGACTAGTGGAGAAGTACTCGCCTACATG

SEQ ID NO: 14 - Ad5orf6 primer 2 polynucleotide sequence

ATACGGAAGATCTAAGACTTCAGGAAATATGACTAC

25

SEQ ID NO: 15 - BAC/CHAd155 ΔE1_TetO hCMV RpsL-Kana primer 1 polynucleotide sequence

ATTCAGTGTACAGGCGCGCCAAAGCATGACGCTGTTGATTGATTC

30 **SEQ ID NO: 16 - BAC/CHAd155 ΔE1_TetO hCMV RpsL-Kana (#1375) primer 2 polynucleotide sequence**

ACTAGGACTAGTTATAAGCTAGAATGGGGCTTTGC

SEQ ID NO: 17 - 1021-FW E4 Del Step1 primer polynucleotide sequence

TTAATAGACACAGTAGCTTAATAGACCCAGTAGTGCAAAGCCCCATTCTAGCTTATAACCCCTAT
35 TTGTTTATTTTCT

SEQ ID NO: 18 - 1022-RW E4 Del Step1 primer polynucleotide sequence

ATATATACTCTCTCGGCACTTGGCCTTTTACACTGCGAAGTGTTGGTGCTGGTGCTGCGTTGAGA
GATCTTTATTTGTAACTGTAAATTGTC

40

SEQ ID NO: 19 - 1025-FW E4 Del Step2 primer polynucleotide sequence

TTAATAGACACAGTAGCTTAATA

SEQ ID NO: 20 - 1026-RW E4 Del Step2 primer polynucleotide sequence

GGAAGGGAGTGTCTAGTGTT

5

SEQ ID NO: 21 - 91-SubMonte FW primer polynucleotide sequence

CAATGGGCGTGGATAGCGGTTTGAC

SEQ ID NO: 22 - 90-BghPolyA RW primer polynucleotide sequence

10 CAGCATGCCTGCTATTGTC

SEQ ID NO: 23 - CMVfor primer polynucleotide sequence

CATCTACGTATTAGTCATCGCTATTACCA

15 **SEQ ID NO: 24 - CMVrev primer polynucleotide sequence**

GACTTGGAAATCCCCGTGAGT

SEQ ID NO: 25 - CMVFAM-TAMRA qPCR probe polynucleotide sequence

ACATCAATGGGCGTGGATAGCGGTT

20

**SEQ ID NO: 26 - Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE)
polynucleotide sequence**

TAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTT
TACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCAT
25 TTTCTCCTCCTTGATAAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCA
ACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCT
GTCAGCTCCTTTCCGGGACTTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACATCGCCGCCT
GCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTGTCGGG
GAAATCATCGTCCTTTTCCTTGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTT
30 CTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGC
GGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCG
CCT

SEQ ID NO: 27 - ChAd3 fiber amino acid sequence

35 MKRTKTSDESNPVYPYDTESGPPSPVFLTPPFVSPDGFQESPPGVLSLNLAEPVLTSHGMLALKMGS
GLSLDDAGNLTSDITTASPLKKTKTNLSLETSSPLTVSTSGALTVAAPLAVAGTSLTMQSEAPL
TVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSLTLSATPPINVSSGSLGLDMEDPMYTHDGK
LGIRIGGPLRVVDSLHTLVVTGNGLTVDNNALQTRVTGALGYDTSGNLQLRAAGGMRIDANGQLI
LNVAYPFDAQNNLSRLGQGPLYINTDHNLDLNCNRGLTTTTNTNKKLETKISSGLDYDTNGAVIHK
40 LGTGLSFDNTGALTVGNTGDDKLTWTPDPSPNCRIHSDKDCKFTLVLTCKGSQILASVAALAVSG
NLASITGTVASVTIFLRFDQNGVLMENSSLDRQYWNFRNGNSTNAAPYTNAVGFMPNLAAYPKTQS

QTAKNNIVSQVYLNQDKSKPMTLTITLNGTNESESQVSHYMSFTWAWESGQYATETFATNSFTFSYIAEQ

SEQ ID NO: 28- PanAd3 fiber amino acid sequence

5 MKRAKTSDETFNPVYPYDTENGPPSVPFLLTPPFVSPDGFQESPPGVLSLRLSEPLVTSHGMLALKMGN
GLSLDDAGNLTSDQDVTTPPLKKTKTNLSLQTSAPLTVSSGSLTVAAAAPLAVAGTSLTMQSQA
TVQDAKLGLATQGPLTVSEGKLTQLTSAPLTAADSSTLTGTTTPPISVSSGSLGLDMEDPMYTHDGK
LGIRIGGPLQVVDSLHTLTVVVTGNGITVANNALQTKVAGALGYDSSGNLELRAAGGMRINTGGQLIL
DVAYPFDAQNLSLRLGQGPLYVNTNHNLDLNCNRGLTTTTSSNTTKLETKIDSGLDYNANGAIIAK
10 LGTGLTFDNTGAITVGNTGDDKLTWTTDPDPSPCRIHADKDKFTLVLTCKGSQILASVAALAVSGN
LSSMTGTVSSVTIFLRFDQNGVLMENSSLDKEYWNFRNGNSTNATPYTNAVGFMPNLSAYPKTQSQ
TAKNNIVSEVYLHGDKSKPMILTITLNGTNESESQVSHYMSFTWSWDSGKYATETFATNSFTFSY
IAEQ

15 SEQ ID NO: 29 - ChAd17 fiber amino acid sequence

MKRTKTSDESFPVYPYDTESGPPSVPFLLTPPFVSPDGFQESPPGVLSLNLAEPLVTSHGMLALKMGS
GLSLDDAGNLTSDITSTTPPLKKTKTNLSLETSSPLTVSTSGALTVAAAAPLAVAGTSLTMQSEAPL
AVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSTLTVSSTPPISVSSGSLGLDMEDPMYTHDGK
LGIRIGGPLRVVDSLHTLTVVVTGNGLTVDNNALQTRVTGALGYDTSGNLQLRAAGGMRIDANGQLI
20 LDVAYPFDAQNLSLRLGQGPLYVNTDHNLDLNCNRGLTTTTTNNTKKLETKISSGLDYDTNGAVII
KLGTGLSFDNTGALTVGNTGDDKLTWTTDPDPSPCRIHSDKDKFTLVLTCKGSQILASVAALAVS
GNLASITGTVASVTIFLRFDQNGVLMENSSLDKQYWNFRNGNSTNAAPYTNAVGFMPNLAAYPKTQ
SQTAKNNIVSQVYLNQDKSKPMTLTITLNGTNESESQVSHYMSFTWAWESGQYATETFATNSFT
FSYIAEQ

25

SEQ ID NO: 30 - ChAd19 fiber amino acid sequence

MKRTKTSDDKSFNPVYPYDTENGPPSVPFLLTPPFVSPDGFQESPPGVLSLNLAEPLVTSHGMLALKMGS
GLSLDDAGNLTSDQDVTTPPLKKTKTNLSLETSAPLTVSTSGALTLAAAAPLAVAGTSLTMQSEAPL
TVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSTLTVSATPPISVSSGSLGLDMEDPMYTHDGK
30 LGIRIGGPLRVVDSLHTLTVVVTGNGIAVDNNALQTRVTGALGYDTSGNLQLRAAGGMRIDANGQLIL
DVAYPFDAQNLSLRLGQGPLYVNTDHNLDLNCNRGLTTTTTNNTKKLETKIGSGLDYDTNGAVIIK
LGTGVSFDSTGALSFGNTGDDKLTWTTDPDPSPCRIHSDKDKFTLVLTCKGSQILASVAALAVSG
NLASITGTVSSVTIFLRFDQNGVLMENSSLDKQYWNFRNGNSTNATPYTNAVGFMPNLAAYPKTQS
QTAKNNIVSQVYLNQDKSKPMTLTITLNGTNESESQVSHYMSFTWAWESGQYATETFATNSFTFSYIAEQ
35

SEQ ID NO: 31 - ChAd24 fiber amino acid sequence

MKRTKTSDESFPVYPYDTENGPPSVPFLLTPPFVSPDGFQESPPGVLSLNLAEPLVTSHGMLALKMGS
GLSLDDAGNLTSDQDVTTPPLKKTKTNLSLETSAPLTVSTSGALTLAAAAPLAVAGTSLTMQSEAPL
40 TVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSTLTVSATPPINVSSGSLGLDMENPMYTHDGK
LGIRIGGPLRVVDSLHTLTVVVTGNGIAVDNNALQTRVTGALGYDTSGNLQLRAAGGMRIDANGQLIL
DVAYPFDAQNLSLRLGQGPLYVNTDHNLDLNCNRGLTTTTTNNTKKLETKIGSGLDYDTNGAVIIK

LGTGVSFDSTGALSVGNTGDDKLT LWTPDPSPNCRIHSDKDCKFTLVLT KCGSQILASVAALAVSG
 NLASITGTVSSVTIFLRF DQNGVLMENSSLDKQYWNFRNGNSTNATPYTNAVGFMPNLAAYPKTQS
 QTAKNNIVSQVYLN GDKSKPMILTITLNGTNESETSQVSHYSMSFTWAWESGQYATETFATNSFTFS
 YIAEQ

5

SEQ ID NO: 32 - ChAd11 fiber amino acid sequence

MKRTKTSDESFPVYPYDTENGPPSV PFLT PPFVSPDGFQESPPGVLSLNLA EPLVTSHGMLALKMGS
 GLSLDDAGNLTSQDVTTTTPPLKKT KTNLSLETSAPLTVSTSGALT LAAAVPLAVAGTSLTMQSEAPL
 TVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSLTISATPPLSTSNGLGIDMQAPIYTTNGKLG
 10 LNFAGAPLHVVDLSNALT VVTGQGLTINGTALQTRVSGALNYDSSGNLELRAAGGM RVDANGKLILD
 VAYPFDAQN NLSRLGQG PLFVNSAHNLDVNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAIN
 PGDGLEFGSGSDTNPLKTKLGLGLE YDSSRAIIAKLGTGLSFDNTGAITVGNKNDDKLT LWTPDPSP
 NCRIYSEKDAKFTLVLT KCGSQVLASVSVLSVKGSLAPISGTVTSAQILRFDENG VLLSNSSLD PQYW
 NYRKGDLT EGTA YTN AVGFMPNLTA YPKTQSQTAKSNIVSQVYLN GDKSKPMILTITLNGTNETGD
 15 ATVSTYSMSFSWNWNGSNYINETFQ TNSFTFSYIAEQ

SEQ ID NO: 33 - ChAd20 fiber amino acid sequence

MKRTKTSDESFPVYPYDTESGPPSV PFLT PPFVSPDGFQESPPGVLSLNLA EPLVTSHGMLALKMGS
 GLSLDDAGNLTSQDITTASPLKKT KTNLSLETSSPLTVSTSGALT VAAAAPLAVAGTSLTMQSEAPL
 20 TVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSLT V SATPPLSTSNGLGIDMQAPIYTTNGKL
 GLNFGAPLHVVDLSNALT VVTGQGLTINGTALQTRVSGALNYDTSGNLELRAAGGM RVDANGQLIL
 DVAYPFDAQN NLSRLGQG PLFVNSAHNLDVNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAI
 NAGDGLQFDSGSDTNPLKTKLGLGLDYDSSRAIIAKLGTGLSFDNTGAITVGNKNDDKLT LWTPDP
 SPNCRIYSEKDAKFTLVLT KCGSQVLASVSVLSVKGSLAPISGTVTSAQIVLRFDENG VLLSNSSLD PQ
 25 YWNYRKGDLT EGTA YTN AVGFMPNLTA YPKTQSQTAKSNIVSQVYLN GDKSKPMTLTITLNGTNET
 GDATVSTYSMSFSWNWNGSNYINETFQ TNSFTFSYIAEQ

SEQ ID NO: 34 - ChAd31 fiber amino acid sequence

MKRTKTSDESFPVYPYDTESGPPSV PFLT PPFVSPDGFQESPPGVLSLNLA EPLVTSHGMLALKMGS
 30 GLSLDDAGNLTSQDITTASPLKKT KTNLSLETSSPLTVSTSGALT VAAAAPLAVAGTSLTMQSEAPL
 TVQDAKLTLATKGPLTVSEGKLALQTSAPLTAADSSLT V SATPPLSTSNGLGIDMQAPIYTTNGKL
 GLNFGAPLHVVDLSNALT VVTGQGLTINGTALQTRVSGALNYDTSGNLELRAAGGM RVDANGQLIL
 DVAYPFDAQN NLSRLGQG PLFVNSAHNLDVNYNRGLYLFTSGNTKKLEVNIKTAKGLIYDDTAIAI
 NAGDGLQFDSGSDTNPLKTKLGLGLDYDSSRAIIAKLGTGLSFDNTGAITVGNKNDDKLT LWTPDP
 35 SPNCRIYSEKDAKFTLVLT KCGSQVLASVSVLSVKGSLAPISGTVTSAQIVLRFDENG VLLSNSSLD PQ
 YWNYRKGDLT EGTA YTN AVGFMPNLTA YPKTQSQTAKSNIVSQVYLN GDKSKPMTLTITLNGTNET
 GDATVSTYSMSFSWNWNGSNYINETFQ TNSFTFSYIAEQ

SEQ ID NO: 35 - PanAd1 fiber amino acid sequence

MKRAKTSDETFPVYPYDTENGPPSV PFLT PPFVSPDGFQESPPGVLSRLSEPLVTSHGMLALKMGN
 40 GLSLDDAGNLTSQDVTTVTPPLKKT KTNLSLQTSAPLTVSSGSLT VAAAAPLAVAGTSLTMQSQA PL
 TVQDAKLGLATQG PLTVSEGKLTLQTSAPLTAADSSLT V SATPPLSTSNGLSIDMQAPIYTTNGKL

ALNIGAPLHVVDLNLALT VVTGQGLTINGRALQTRVTGALS YDTEGNIQLQAGGGMRIDNNGQLILN
 VAYPFDAQNNLSRLGQGPLIVNSAHNLDLNLNRGLYLFTSGNTKKLEVNIKTAKGLFYDGTAIIN
 AGDGLQFGSGSDTNPLQTKLGLGLEYSNKAIITKLGTGLSFDNTGAITVGNKNDDKLTWTPDPDS
 PNCRINSEKDAKLTVLTKCGSQVLASVSVLSVKGSLAPISGTVTSAQIVLRFDENGVLSSNSLDPQY
 5 WNYRKGDSTEGTAYTNAVGFMPNLTA YPKTQSQTAKSNIVSQVYLN GDKTKPMTLTITLNGTNETG
 DATVSTYSMSFSWNWNGSNYINDTFQ TNSFTFSYIAQE

SEQ ID NO: 36 - PanAd2 fiber amino acid sequence

MKRAKTSDEFNPPVYPYDTENGPPSPVFLTPPFVSPDGFQESPPGVLSRLSEPLVTSHGMLALKMGN
 10 GLSLDDAGNLT SQDVTTPPLKKTKTNLSLQTSAPLTVSSGSLTVAAA PLAVAGTSLTMQSQA PL
 TVQDAKLGLATQGPLTVSEGKLT LQTSAPLTAADSSLT V SATPPLSTSNGLSIDMQAPIYTTNGKL
 ALNIGAPLHVVDLNLALT VVTGQGLTINGRALQTRVTGALS YDTEGNIQLQAGGGMRIDNNGQLILN
 VAYPFDAQNNLSRLGQGPLIVNSAHNLDLNLNRGLYLFTSGNTKKLEVNIKTAKGLFYDGTAIIN
 AGDGLQFGSGSDTNPLQTKLGLGLEYSNKAIITKLGTGLSFDNTGAITVGNKNDDKLTWTPDPDS
 15 PNCRINSEKDAKLTVLTKCGSQVLASVSVLSVKGSLAPISGTVTSAQIVLRFDENGVLSSNSLDPQY
 WNYRKGDSTEGTAYTNAVGFMPNLTA YPKTQSQTAKSNIVSQVYLN GDKTKPMTLTITLNGTNETG
 DATVSTYSMSFSWNWNGSNYINDTFQ TNSFTFSYIAQE

SEQ ID NO: 37 - RSV FΔTM amino acid sequence

MELLILKANAITILTAVTFCFASGQNITEEFYQSTCSAVSKGYLSALRTGWYTSVITIELSNIKENKCN
 20 GTDAKVKLIKQELDKYKNAVTELQLLMQSTPATNNRARELPRFMNYTLNNAKKTNVTL SKKRKR
 RFLGFLLVGSAIASGVA VSKVLHLEGEVNIKSALLSTNKAVVSLSNGVSVLTSKVLDLKNYIDKQ
 LPIVNVKQSCSISNIETVIEFQQKNNRLL EITREFSVNAGVTTTPVSTYMLTNSSELLSLNDMPITNDQKKL
 MSNNVQIVRQQSYSIMSIIKEEV LAYVVQLPLYGVIDTPCWKLHTSPLCTTNTKEG SNICLTRTDRGW
 25 YCDNAGSVSFFPQAETCKVQSNRV FCDTMNSLTLPSEVNLCNVDFNPKYDCKIMTSKTDVSSSVITS
 LGAIVSCYGKTKCTASNKNRGIKTF SNGCDYVSNGVDTVSVGNTLYYV NKQEGKSLYVKGEPIIN
 FYDPLVFP SDEFDASISQVNEKINQSLAFIRKSDELLHNVNAGKSTNRKRRAPVKQTLNFDLLKLAG
 DVESNPGPMALSKVKLNDTLNKDQL LSSSKYTIQRSTGDSIDTPNYDVQKHINKLCGMLLITEDANH
 KFTGLIGMLYAMSRLGREDTIKILRDAGYHVKANGVDVTTHRQDINGKEMKFEVLT LASLTTEIQINI
 30 EIESRKSYYKKMLKEMGEVAPEYRHDS PDCGMILCIAALVITKLAAGDRSGLTAVIRRANNVLKNEM
 KRYKGLLPKDIANS
 FYEVFEKYPHFIDVFVHF GIAQSSTRGGS RVEGIFAGLFMNAYGAGQVMLRWGV LAKSVKNIMLGH
 ASVQAEME QVVEVYEYAQKL GGEAGFYHILNPNKASLLSLTQFPHFSSVVLGNAAGLGIMGEYRGT
 PRNQDLYDAAKAYAEQLKENG VINYSVLDLTAEEL EAIKHQLNPKDNDVELGGGGSGGGGMSRRN
 35 PCKFEIRGHCLNGKRCHFSHNYFEWPPHALLVRQNFMLNRILKSMDKSIDTLSEISGAAELDRTEEYA
 LGVVGVLESYIGSINNITKQSACVAMSKLL TELNSDDIKLRDNEELNSPKIRVYNTVISYIESNRKNN
 KQTIHLLKRLPADVLKKTIKNTLDI HKSITINNPKESTVSDTNDHAKNNDTT

SEQ ID NO: 38 - HIV Gag polynucleotide sequence

40 ATGGGTGCTAGGGCTTCTGTGCTGTCTGGTGGTGAGCTGGACAAGTGGGAGAAGATCAGGCTGA
 GGCCTGGTGGCAAGAAGAAGTACAAGCTAAAGCACATTGTGTGGGCCTCCAGGGAGCTGGAGA
 GGTGTGCTGTGAACCTGGCCTGCTGGAGACCTCTGAGGGGTGCAGGCAGATCCTGGGCCAGCT
 CCAGCCCTCCCTGCAAACAGGCTCTGAGGAGCTGAGGTCCCTGTACAACACAGTGGCTACCCTG
 TACTGTGTGACCAGAAGATTGATGTGAAGGACACCAAGGAGGCCCTGGAGAAGATTGAGGAG

GAGCAGAACAAAGTCCAAGAAGAAGGCCAGCAGGCTGCTGCTGGCACAGGCAACTCCAGCCAG
GTGTCCCAGAACTACCCCATTTGTGCAGAACCTCCAGGGCCAGATGGTGCACCAGGCCATCTCCC
CCCGGACCCTGAATGCCTGGGTGAAGGTGGTGGAGGAGAGGCCTTCTCCCCTGAGGTGATCCCC
ATGTTCTCTGCCCTGTCTGAGGGTGCCACCCCCAGGACCTGAACACCATGCTGAACACAGTGG
5 GGGGCCATCAGGCTGCCATGCAGATGCTGAAGGAGACCATCAATGAGGAGGCTGCTGAGTGGG
ACAGGCTGCATCCTGTGCACGCTGGCCCCATTGCCCCGGCCAGATGAGGGAGCCCAGGGGCTC
TGACATTGCTGGCACCACCTCCACCCTCCAGGAGCAGATTGGCTGGATGACCAACAACCCCCC
ATCCCTGTGGGGGAAATCTACAAGAGGTGGATCATCCTGGGCCTGAACAAGATTGTGAGGATGT
ACTCCCCACCTCCATCCTGGACATCAGGCAGGGCCCCAAGGAGCCCTTCAGGGACTATGTGGA
10 CAGGTTCTACAAGACCCTGAGGGCTGAGCAGGCCTCCCAGGAGGTGAAGAACTGGATGACAGA
GACCCTGCTGGTGCAGAAATGCCAACCCTGACTGCAAGACCATCCTGAAGGCCCTGGGCCCTGCT
GCCACCCTGGAGGAGATGATGACAGCCTGCCAGGGGGTGGGGGGCCCTGGTCACAAGGCCAGG
GTGCTGGCTGAGGCCATGTCCCAGGTGACCAACTCCGCCACCATCATGATGCAGAGGGGCAACT
TCAGGAACCAGAGGAAGACAGTGAAGTGCTTCAACTGTGGCAAGGTGGGCCACATTGCCAAGA
15 ACTGTAGGGCCCCCAGGAAGAAGGGCTGCTGGAAGTGTGGCAAGGAGGGCCACCAGATGAAGG
ACTGCAATGAGAGGCAGGCCAACTTCCTGGGCAAAATCTGGCCCTCCCACAAGGGCAGGCCTGG
CAACTTCCTCCAGTCCAGGCCTGAGCCCACAGCCCCTCCCGAGGAGTCCTTCAGGTTTGGGGAG
GAGAAGACCACCCCCAGCCAGAAGCAGGAGGCCATTGACAAGGAGCTGTACCCCTGGCCTCCC
TGAGGTCCCTGTTTGGCAACGACCCCTCCTCCCAGTAA
20

CLAIMS

1. An isolated recombinant adenoviral polynucleotide, wherein the polynucleotide encodes a polypeptide comprising:
 - (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 3;
 - (b) a polypeptide having the amino acid sequence according to SEQ ID NO: 5; and
 - (c) a polypeptide selected from the group consisting of
 - (i) a polypeptide having the amino acid sequence according to SEQ ID NO: 1; and
 - (ii) a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1;wherein the polynucleotide comprises or comprised at least one gene of an adenoviral genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and E4 and a mutation or deletion which renders the at least one gene non-functional or absent.
2. A recombinant adenovirus comprising:
 - (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3;
 - (b) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 5; and
 - (c) a polynucleotide selected from the group consisting of
 - (i) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1 and
 - (ii) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1;wherein the polynucleotide comprises or comprised at least one gene of an adenoviral genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and E4 and a mutation or deletion which renders the at least one gene non-functional or absent.
3. A composition comprising:
 - (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3;
 - (b) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 5; and
 - (c) a polynucleotide selected from the group consisting of

- (i) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1; and
 - (ii) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1;

and a pharmaceutically acceptable excipient,

wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1; wherein the polynucleotide in (a), (b) and (c) comprises a recombinant polynucleotide; and wherein the polynucleotide comprises or comprised at least one gene of an adenoviral genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and E4 and a mutation or deletion which renders the at least one gene non-functional or absent.
- 4. A cell comprising:
 - (a) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 3;
 - (b) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 5; and
 - (c) a polynucleotide selected from the group consisting of
 - (i) a polynucleotide which encodes a polypeptide having the amino acid sequence according to SEQ ID NO: 1; and
 - (ii) a polynucleotide which encodes a functional derivative of a polypeptide having the amino acid sequence according to SEQ ID NO: 1, wherein the functional derivative has an amino acid sequence which is at least 80% identical over its entire length to the amino acid sequence of SEQ ID NO: 1;

wherein the cell is a host cell that expresses at least one adenoviral gene selected from the group consisting of E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 and L5, and wherein the polynucleotide in (a), (b) and (c) comprises a recombinant polynucleotide.
- 5. The host cell according to claim 4 wherein the host cell is able to grow in suspension.
- 6. The polynucleotide according to claim 1, wherein the polynucleotide comprises at least one of the following:
 - (a) an adenoviral 5' inverted terminal repeat;
 - (b) an adenoviral E1A region, or a fragment thereof selected from among the E1A_280R and E1A_243R regions;
 - (c) an adenoviral E1B or IX region, or a fragment thereof selected from among the group consisting of the E1B_19K, E1B_55K or IX regions;
 - (d) an adenoviral E2b region; or a fragment thereof selected from among the group consisting of the E2B_pTP, E2B_Polymerase and E2B_IVa2 regions;

- (e) an adenoviral L1 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L1_13.6k protein, L1_52k and L1_Illa protein;
 - (f) an adenoviral L2 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L2_penton protein according to claim 1, L2_pVII, L2_V, and L2_pX protein;
 - (g) an adenoviral L3 region, or a fragment thereof, said fragment encoding an adenoviral protein selected from the group consisting of the L3_pVI protein, L3_hexon protein according to claim 1 and L3_protease;
 - (h) an adenoviral E2A region;
 - (i) an adenoviral L4 region, or a fragment thereof said fragment encoding an adenoviral protein selected from the group consisting of the L4_100k protein, the L4_33k protein and protein L4_VIII;
 - (j) an adenoviral E3 region, or a fragment thereof selected from the group consisting of E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, and E3 ORF9;
 - (k) an adenoviral L5 region, or a fragment thereof said fragment encoding the L5_fiber fiber protein according to claim 1;
 - (l) an adenoviral E4 region, or a fragment thereof selected from the group consisting of E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, and E4 ORF1; and/or
 - (m) an adenoviral 3'-end, preferably an adenoviral 3' inverted terminal repeat.
7. The polynucleotide according to claim 6, further comprising an adenoviral VAI or VAII RNA region.
 8. The polynucleotide according to claim 7, wherein the VAI or VAII RNA region is from Ad5.
 9. The polynucleotide according to claim 1, wherein the genomic regions are E1A and/or E1B.
 10. The recombinant adenovirus according to claim 2, wherein the recombinant adenovirus is replication-incompetent.
 11. The replication-incompetent recombinant adenovirus of claim 10, wherein the adenovirus comprises a heterologous nucleic acid sequence encoding a protein and wherein the heterologous nucleic acid sequence is operatively linked to one or more sequences which direct expression of the protein in a host cell.

12. The replication-incompetent recombinant adenovirus according to claim 11, wherein the protein is an antigenic protein or a fragment thereof.
13. The recombinant adenovirus according to claim 2, wherein the adenovirus has a seroprevalence of less than 10% in human subjects.
14. The recombinant adenovirus according to claim 13, wherein the adenovirus has a seroprevalence of less than 5% in human subjects.
15. The composition of claim 3, further comprising a heterologous nucleic acid sequence encoding an antigenic protein or a fragment thereof, and wherein the heterologous nucleic acid sequence is operatively linked to one or more sequences that direct expression of the antigenic protein or a fragment thereof in a host cell.
16. The composition of claim 3, wherein the antigenic protein comprises a respiratory syncytial virus (RSV) antigenic protein.
17. The composition of claim 15 or claim 16 for use in inducing an immune response in a subject.
18. Use of the composition of claim 16 in the manufacture of a medicament for inducing an immune response against respiratory syncytial virus (RSV) in a subject.
19. An isolated recombinant adenoviral polynucleotide, wherein the polynucleotide encodes a polypeptide comprising:
 - (a) a polypeptide having the amino acid sequence according to SEQ ID NO: 1;
 - (b) a polypeptide having the amino acid sequence according to SEQ ID NO: 3; and
 - (c) a polypeptide having the amino acid sequence according to SEQ ID NO: 5;wherein the polynucleotide comprises or comprised at least one gene of an adenoviral genomic region selected from the group consisting of E1A, E1B, E2A, E2B, E3 and E4 and a mutation or deletion which renders the at least one gene non-functional or absent.
20. The isolated recombinant adenoviral polynucleotide of claim 19 further comprising a heterologous nucleic acid sequence encoding a protein and wherein the heterologous nucleic acid sequence is operatively linked to one or more sequences that direct expression of the protein in a host cell.