Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2872867 C 2020/09/15

(11)(21) 2 872 867

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) **Date de dépôt PCT/PCT Filing Date:** 2013/05/10
(87) **Date publication PCT/PCT Publication Date:** 2013/11/14
(45) **Date de délivrance/Issue Date:** 2020/09/15
(85) **Entrée phase nationale/National Entry:** 2014/11/06
(86) **N° demande PCT/PCT Application No.:** AU 2013/000481
(87) **N° publication PCT/PCT Publication No.:** 2013/166558
(30) **Priorité/Priority:** 2012/05/11 (US61/646,174)

(51) **Cl.Int./Int.Cl.** ***C12Q 1/6809*** (2018.01), ***C12Q 1/68*** (2018.01), ***C12Q 1/6886*** (2018.01)

(72) **Inventeurs/Inventors:**
MOLLOY, PETER, AU;
LAPOINTE, LAWRENCE, AU;
PEDERSEN, SUSANNE, AU

(73) **Propriétaires/Owners:**
CLINICAL GENOMICS PTY LTD, AU;
COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION, AU

(74) **Agent:** RIDOUT & MAYBEE LLP

(54) **Titre : GROUPE DE MARQUEURS GENETIQUES DIAGNOSTIQUES DESTINE AU CANCER COLORECTAL**
(54) **Title: DIAGNOSTIC GENE MARKER PANEL FOR COLORECTAL CANCER**

(57) **Abrégé/Abstract:**
The present invention relates generally to a method of screening for the onset, predisposition to the onset and/or progression of a neoplasm. More particularly, the present invention relates to a method of screening for the onset, predisposition to the onset and/or progression of a neoplasm by screening for changes to the methylation levels of a panel of gene markers including BCATI, IKZFI, IRF4, GRASP and/or CAHM. The method of the present invention is useful in a range of applications including, but not limited to, those relating to the diagnosis and/or monitoring of colorectal neoplasms, such as colorectal adenocarcinosis.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 November 2013 (14.11.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/166558 A1

(51) International Patent Classification:
C12Q 1/68 (2006.01) *C12N 15/11* (2006.01)

(21) International Application Number:
PCT/AU2013/000481

(22) International Filing Date:
10 May 2013 (10.05.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/646,174 11 May 2012 (11.05.2012) US

(71) Applicants: **CLINICAL GENOMICS PTY LTD** [AU/AU]; Riverside Life Sciences Building, 11 Julius Avenue, North Ryde, New South Wales 2113 (AU). **COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION** [AU/AU]; Limestone Avenue, Campbell, Australian Capital Territory 2612 (AU).

(72) Inventors: **MOLLOY, Peter**; 41 Bellevue Street, Chatswood, New South Wales 2067 (AU). **LAPOINTE, Lawrence**; 66 Range Road, West Pennant Hills, New South Wales 2125 (AU). **PADERSEN, Susanne**; 30 Michael Street, North Ryde, New South Wales 2113 (AU).

(74) Agent: **DAVIES COLLISON CAVE**; 1 Nicholson Street, Melbourne, Victoria 3000 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2013/166558 A1

(54) Title: DIAGNOSTIC GENE MARKER PANEL FOR COLORECTAL CANCER

(57) Abstract: The present invention relates generally to a method of screening for the onset, predisposition to the onset and/or progression of a neoplasm. More particularly, the present invention relates to a method of screening for the onset, predisposition to the onset and/or progression of a neoplasm by screening for changes to the methylation levels of a panel of gene markers including BCAT1, IKZF1, IRF4, GRASP and/or CAHM. The method of the present invention is useful in a range of applications including, but not limited to, those relating to the diagnosis and/or monitoring of colorectal neoplasms, such as colorectal adenocarcinosis.

DIAGNOSTIC GENE MARKER PANEL FOR COLORECTAL CANCER

FIELD OF THE INVENTION

The present invention relates generally to a method of screening for the onset, predisposition to the onset and/or progression of a neoplasm. More particularly, the present invention relates to a method of screening for the onset, predisposition to the onset and/or progression of a neoplasm by screening for changes to the methylation levels of a panel of gene markers. The method of the present invention is useful in a range of applications including, but not limited to, those relating to the diagnosis and/or monitoring of colorectal neoplasms, such as colorectal adenocarcinosis.

BACKGROUND OF THE INVENTION

Colorectal cancer includes cancerous growths in the colon, rectum and appendix. With 655,000 deaths worldwide per year, it is the fourth most common form of cancer in the United States and the third leading cause of cancer-related death in the Western world. Colorectal cancers arise from adenomatous polyps in the colon. These mushroom-shaped growths are usually benign, but some develop into cancer over time. Localized colon cancer is usually diagnosed through colonoscopy.

Invasive cancers that are confined within the wall of the colon (stages I and II) are curable with surgery. If untreated, they spread to regional lymph nodes (stage III), where up to 73% are curable by surgery and chemotherapy. Cancer that metastasizes to distant sites (stage IV) is usually not curable, although chemotherapy can extend survival, and in rare cases, surgery and chemotherapy together have seen patients through to a cure (Markowitz and Bertagnolli, 2009, *N. Engl. J. Med.* 361(25): 2449–60). Radiation is used with rectal cancer.

Colorectal cancer is preceded by adenomas. Adenomas are benign tumours, or neoplasms, of epithelial origin which are derived from glandular tissue or exhibit clearly defined glandular structures. Some adenomas show recognisable tissue elements, such as fibrous tissue (fibroadenomas) and epithelial structure, while others, such as bronchial adenomas, produce active compounds that might give rise to clinical syndromes.

Adenomas may progress to become an invasive neoplasm and are then termed adenocarcinomas. Accordingly, adenocarcinomas are defined as malignant epithelial tumours arising from glandular structures, which are constituent parts of many organs of the body. The term adenocarcinoma is also applied to tumours showing a glandular growth pattern. These

tumours may be sub-classified according to the substances that they produce, for example mucus secreting and serous adenocarcinomas, or to the microscopic arrangement of their cells into patterns, for example papillary and follicular adenocarcinomas. These carcinomas may be solid or cystic (cystadenocarcinomas). Each organ may produce tumours showing a variety of histological types, for example the ovary may produce both mucinous and cystadenocarcinoma.

Adenomas in different organs behave differently. In general, the overall chance of carcinoma being present within an adenoma (i.e. a focus of cancer having developed within a benign lesion) is approximately 5%. However, this is related to size of an adenoma. For instance, in the large bowel (colon and rectum specifically) occurrence of a cancer within an adenoma is rare in adenomas of less than 1 centimetre. Such a development is estimated at 40 to 50% in adenomas which are greater than 4 centimetres and show certain histopathological change such as villous change, or high grade dysplasia. Adenomas with higher degrees of dysplasia have a higher incidence of carcinoma. In any given colorectal adenoma, the predictors of the presence of cancer now or the future occurrence of cancer in the organ include size (especially greater than 9mm) degree of change from tubular to villous morphology, presence of high grade dysplasia and the morphological change described as "serrated adenoma". In any given individual, the additional features of increasing age, familial occurrence of colorectal adenoma or cancer, male gender or multiplicity of adenomas, predict a future increased risk for cancer in the organ – so-called risk factors for cancer. Except for the presence of adenomas and its size, none of these is objectively defined and all those other than number and size are subject to observer error and to confusion as to precise definition of the feature in question. Because such factors can be difficult to assess and define, their value as predictors of current or future risk for cancer is imprecise.

Once a sporadic adenoma has developed, the chance of a new adenoma occurring is approximately 30% within 26 months.

The symptoms of colorectal cancer depend on the location of tumor in the bowel, and whether is has metastasised. Unfortunately, many of the symptoms may occur in other diseases as well, and hence symptoms may not be conclusively diagnostic of colorectal cancer.

Local symptoms are more likely if the tumor is located closer to the anus. There may be a change in bowel habit (new-onset constipation or diarrhea in the absence of another cause), a feeling of incomplete defecation and reduction in diameter of stools. Tenesmus and change in stool shape are both characteristic of rectal cancer. Lower gastrointestinal bleeding,

including the passage of bright red blood in the stool, may indicate colorectal cancer, as may the increased presence of mucus. Melena, black stool with a tarry appearance, normally occurs in upper gastrointestinal bleeding (such as from a duodenal ulcer), but is sometimes encountered in colorectal cancer when the disease is located in the beginning of the large bowl.

A tumor that is large enough to fill the entire lumen of the bowel may cause bowel obstruction. This situation is characterized by constipation, abdominal pain, abdominal distension and vomiting. This occasionally leads to the obstructed and distended bowel perforating and causing peritonitis.

Certain local effects of colorectal cancer occur when the disease has become more advanced. A large tumor is more likely to be noticed on feeling the abdomen, and it may be noticed by a doctor on physical examination. The disease may invade other organs, and may cause blood or air in the urine or vaginal discharge.

If a tumor has caused chronic occult bleeding, iron deficiency anaemia may occur. This may be experienced as fatigue, palpitations and noticed as pallor. Colorectal cancer may also lead to weight loss, generally due to a decreased appetite.

More unusual constitutional symptoms are an unexplained fever and one of several paraneoplastic syndromes. The most common paraneoplastic syndrome is thrombosis, usually deep vein thrombosis.

Colorectal cancer most commonly spreads to the liver. This may go unnoticed, but large deposits in the liver may cause jaundice and abdominal pain (due to stretching of the capsule). If the tumor deposit obstructs the bile duct, the jaundice may be accompanied by other features of biliary obstruction, such as pale stools.

Colorectal cancer can take many years to develop and early detection of colorectal cancer greatly improves the prognosis. Even modest efforts to implement colorectal cancer screening methods can result in a drop in cancer deaths. Despite this, colorectal cancer screening rates remain low. There are currently several different tests available for this purpose:

- Digital rectal exam: The doctor inserts a lubricated, gloved finger into the rectum to feel for abnormal areas. It only detects tumors large enough to be felt in the distal part of the rectum but is useful as an initial screening test.
- Faecal occult blood test: a test for blood in the stool. Two types of tests can be used for detecting occult blood in stools i.e. guaiac based (chemical test) and immunochemical. The sensitivity of immunochemical testing is superior to that of chemical testing

without an unacceptable reduction in specificity (Weitzel JN (December 1999). "Genetic cancer risk assessment. Putting it all together". *Cancer* 86 (11 Suppl): 2483–92).

- Endoscopy:
  - Sigmoidoscopy: A lit probe (sigmoidoscope) is inserted into the rectum and lower colon to check for polyps and other abnormalities.
  - Colonoscopy: A lit probe called a colonoscope is inserted into the rectum and the entire colon to look for polyps and other abnormalities that may be caused by cancer. A colonoscopy has the advantage that if polyps are found during the procedure they can be removed immediately. Tissue can also be taken for biopsy.
- Double contrast barium enema (DCBE): First, an overnight preparation is taken to cleanse the colon. An enema containing barium sulfate is administered, then air is insufflated into the colon, distending it. The result is a thin layer of barium over the inner lining of the colon which is visible on X-ray films. A cancer or a precancerous polyp can be detected this way. This technique can miss the (less common) flat polyp.
- Virtual colonoscopy replaces X-ray films in the double contrast barium enema (above) with a special computed tomography scan and requires special workstation software in order for the radiologist to interpret. This technique is approaching colonoscopy in sensitivity for polyps. However, any polyps found must still be removed by standard colonoscopy.
- Standard computed axial tomography is an x-ray method that can be used to determine the degree of spread of cancer, but is not sensitive enough to use for screening. Some cancers are found in CAT scans performed for other reasons.
- Blood tests: Measurement of the patient's blood for elevated levels of certain proteins can give an indication of tumor load. In particular, high levels of carcinoembryonic antigen (CEA) in the blood can indicate metastasis of adenocarcinoma. While these tests are frequently false positive or false negative, and are not recommended for screening, they can be useful to assess disease recurrence. CA19-9 and CA 242 biomarkers can indicate e-selectin related metastatic risks, help follow therapeutic progress, and assess disease recurrence. Recently, an assay for detection in plasma of methylated sequences of the Septin 9 gene has also become available to assist in diagnosis of colorectal cancer.

- Positron emission tomography (PET) is a 3-dimensional scanning technology where a radioactive sugar is injected into the patient, the sugar collects in tissues with high metabolic activity, and an image is formed by measuring the emission of radiation from the sugar. Because cancer cells often have very high metabolic rates, this can be used to differentiate benign and malignant tumors. PET is not used for screening and does not (yet) have a place in routine workup of colorectal cancer cases.
- Stool DNA testing is an emerging technology in screening for colorectal cancer. Premalignant adenomas and cancers shed DNA markers from their cells which are not degraded during the digestive process and remain stable in the stool. Capture, followed by PCR amplifies the DNA to detectable levels for assay.
- High C-Reactive Protein levels as risk marker

Despite the existence of these tests, diagnosis remains problematic. Most of the more sensitive tests are quite invasive and expensive and therefore uptake by patients is low. There is therefore an ongoing need to develop simpler and more informative diagnostic protocols or aids to diagnosis that enable one to direct colonoscopy at people more likely to have developed adenomas or carcinomas. A simple and accurate screening test would enable much more widely applicable screening systems to be set up.

To this end, more recently there have been identified genetic markers which are modulated, in terms of their expression levels, in individuals who have developed a neoplasm of the large intestine. Some of these markers are upregulated in terms of their level of expression, while others are downregulated. However, a feature common to most of these markers, and the use of genetic marker expression levels in general as a diagnostic, is that they often exhibit only moderate levels of sensitivity and specificity. The development of diagnostic protocols which provide high levels of sensitivity and specificity is highly sought after.

In work leading up to the present invention, it has been unexpectedly determined that the gene markers BCAT1, IKZF1 IRF4, GRASP and CAHM, although each individually being one of a number of gene markers known to exhibit utility in terms of diagnosing colorectal neoplasia development, have in fact been determined to collectively enable a significantly higher level of sensitivity or specificity to be achieved than where any one of these gene markers is analysed either alone or together with other unrelated gene markers. More specifically, screening for an increase in the level of methylation of any two or more of these

five specific markers can be designed so as to achieve a level of specificity or sensitivity not previously achievable.

Bearing in mind the large number of gene markers which, in differential gene expression analysis studies, have been shown to exhibit modulated expression levels in neoplasia, the identification that five specific markers can in fact collectively provide improved diagnostic outcomes relative either to these markers individually or to other groups of markers is both unexpected and unpredictable.

SUMMARY OF THE INVENTION

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

As used herein, the term "derived from" shall be taken to indicate that a particular integer or group of integers has originated from the species specified, but has not necessarily been obtained directly from the specified source. Further, as used herein the singular forms of "a", "and" and "the" include plural referents unless the context clearly dictates otherwise.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

The subject specification contains nucleotide sequence information prepared using the programme PatentIn Version 3.5, presented herein after the bibliography. Each nucleotide sequence is identified in the sequence listing by the numeric indicator <210> followed by the sequence identifier (e.g. <210>1, <210>2, etc). The length, type of sequence (DNA, etc) and source organism for each sequence is indicated by information provided in the numeric indicator fields <211>, <212> and <213>, respectively. Nucleotide sequences referred to in the specification are identified by the indicator SEQ ID NO: followed by the sequence identifier (e.g. SEQ ID NO:1, SEQ ID NO:2, etc.). The sequence identifier referred to in the specification correlates to the information provided in numeric indicator field <400> in the sequence listing, which is followed by the sequence identifier (e.g. <400>1, <400>2, etc). That is SEQ ID NO:1 as detailed in the specification correlates to the sequence indicated as <400>1 in the sequence listing.

One aspect of the present invention is directed to a method of screening for the onset or

predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of a DNA region selected from:

(i) the region, including 2kb upstream of the transcription start site, defined by any two or more of Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443;

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of any two or more of:

(1) BCAT1 (2) IKZF1 (3) IRF4 (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In another aspect said method is directed to identifying biological samples in which any one of said DNA regions exhibits a higher level of methylation.

In still another aspect said method is directed to identifying biological samples in which two or more of said DNA regions exhibit a higher level of methylation.

In yet another aspect there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393; and

(2) chr7:50344378...50472798;

and optionally one or more of (3) chr6:391739..411443, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1; and

(2) IKZF1;

and optionally one or more of (3) IRF4, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large

intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In one embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798; and

(3) chr6:391739..411443; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1 and (3) IRF4.

In another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798 and

(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1; and (4) GRASP.

In still another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798 and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1 and (5) CAHM

In still yet another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; and

(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1, (3) IRF4 and (4) GRASP.

In yet still another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1, (3) IRF4 and (5) CAHM.

In a further embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(4) chr12:52400748..52409671 and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1, (4) GRASP and (5) CAHM.

In yet another further embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443

(4) chr12:52400748..52409671 and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1, (3) IRF4, (4) GRASP and (5) CAHM.

In another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393; and
(5) chr6:163834097..163834982;
and optionally one or more of (2) chr7:50344378...50472798, (3) chr6:391739..411443; and (4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1; and
(5) CAHM;
and optionally one or more of (2) IKZF1, (3) IRF4 and (4) GRASP

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In still another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(2) chr7:50344378..50472798; and
(5) chr6:163834097..163834982;
and optionally one or more of (1) chr12:24962958..25102393, (3) chr6:391739..411443 and (4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:
(2) IKZF1; and
(5) CAHM;
and optionally one or more of (1) BCAT1, (3) IRF4 and (4) GRASP

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In a further embodiment there is provided a method of screening for the onset or

predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393; and

(3) chr6:391739..411443;

and optionally one or more of (2) chr7:50344378...50472798, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1; and

(3) IRF4;

and optionally one or more of (2) IKZF1, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In yet another further embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393; and

(4) chr12:52400748..52409671;

and optionally one or more of (2) chr7:50344378...50472798, (3) chr6:391739..411443 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1; and

(4) GRASP;

and optionally one or more of (2) IKZF1, (3) IRF4 and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In still another further embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an

individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(2) chr7:50344378..50472798; and
(3) chr6:391739..411443;
and optionally one or more of (1) chr12:24962958..25102393, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(2) IKZF1; and
(3) IRF4;
and optionally one or more of (1) BCAT1, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In yet still another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(2) chr7:50344378..50472798; and
(4) chr12:52400748..52409671;
and optionally one or more of (1) chr12:24962958..25102393, (3) chr6:391739..411443 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(2) IKZF1; and
(4) GRASP;
and optionally one or more of (1) BCAT1, (3) IRF4 and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In still yet another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(3) chr6:391739..411443; and
(4) chr12:52400748..52409671;
and optionally one or more of (1) chr12:24962958..25102393, (2) chr7:50344378...50472798 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(3) IRF4; and
(4) GRASP;
and optionally one or more of (1) BCAT1, (2) IKZF1 and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(3) chr6:391739..411443; and
(5) chr6:163834097..163834982;
and optionally one or more of (1) chr12:24962958..25102393, (2) chr7:50344378...50472798 and (4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:
(3) IRF4; and
(5) CAHM;
and optionally one or more of (1) BCAT1, (2) IKZF1 and (4) GRASP

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

The subregions which have been determined to exhibit particular utility are listed below with reference to the gene and chromosomal region within which they are found:

(1) BCAT subregions chr12:25101992-25102093 (SEQ ID NO:1 or corresponding minus strand) and chr12:25101909-25101995 (SEQ ID NO:2 or corresponding minus strand)

(2) IKZF1 subregions: chr7:50343867-50343961 (SEQ ID NO:3 or corresponding minus strand) and chr7:50343804-5033895 (SEQ ID NO:4 or corresponding minus strand)

(3) IRF4 subregions chr6:392036-392145 (SEQ ID NO:5 or corresponding minus strand)

(4) GRASP subregions: chr12:52399672-52399922, chr12:52400821-52401051 (SEQ ID NO:6 or corresponding minus strand), chr12:52401407-52401664 (SEQ ID NO:7 or corresponding minus strand) chr12:52400866-52400973 and Chr12:52401107-52401664.

(5) CAHM subregions: chr6:163834295-163834500 (SEQ ID NO:8 or corresponding minus strand), chr6:163834621-163834906, chr6:163834393-163834455 and chr6:163834393-163834519.

To the extent that the method of the present invention includes analysing the methylation of GRASP, the subject residues are:

chr12:52399713	chr12:52399731	chr12:52399749	chr12:52399783
chr12:52399796	chr12:52399808	chr12:52399823	chr12:52399835
chr12:52399891			
chr12:52400847	chr12:52400850	chr12:52400859	chr12:52400866
chr12:52400869	chr12:52400873	chr12:52400881	chr12:52400886
chr12:52400893	chr12:52400895	chr12:52400899	chr12:52400902
chr12:52400907	chr12:52400913	chr12:52400919	chr12:52400932
chr12:52400938	chr12:52400958	chr12:52400962	chr12:52400971
chr12:52400973	chr12:52400976	chr12:52400998	chr12:52401008
chr12:52401010	chr12:52401012	chr12:52401016	chr12:52401019
chr12:52401025	chr12:52401041	chr12:52401044	chr12:52401053
chr12:52401060	chr12:52401064	chr12:52401092	chr12:52401118
chr12:52401438	chr12:52401448	chr12:52401460	chr12:52401465
chr12:52401474	chr12:52401477	chr12:52401479	chr12:52401483
chr12:52401504	chr12:52401514	chr12:52401523	chr12:52401540
chr12:52401553	chr12:52401576	chr12:52401588	chr12:52401595
chr12:52401599	chr12:52401604	chr12:52401606	chr12:52401634
chr12:52401640	chr12:52401644	chr12:52401659	
chr12:52401160	chr12:52401165	chr12:52401174	chr12:52401177
chr12:52401179	chr12:52401183	chr12:52401204	chr12:52401215
chr12:52401223	chr12:52401240	chr12:52401253	chr12:52401288
chr12:52401295	chr12:52401299	chr12:52401304	chr12:52401334
chr12:52401340	chr12:52401344	chr12:52401359	

or a corresponding cytosine at position n+1 on the opposite DNA strand.

To the extent that the method of the present invention includes analysing CAHM, the subject residues are:

chr6:163834330	chr6:163834332	chr6:163834357
chr6:163834373	chr6:163834384	chr6:163834390
chr6:163834392	chr6:163834406	chr6:163834412
chr6:163834419	chr6:163834443	chr6:163834448
chr6:163834452	chr6:163834464	chr6:163834483
chr6:163834653	chr6:163834660	chr6:163834672
chr6:163834675	chr6:163834678	chr6:163834681
chr6:163834815	chr6:163834824	chr6:163834835
chr6:163834840	chr6:163834853	chr6:163834855
chr6:163834858	chr6:163834863	chr6:163834869
chr6:163834872		

or a corresponding cytosine at position n+1 on the opposite DNA strand.

To the extent that the method of the present invention includes analysing IKZF1, the subject residues are:

chr7:50343869	chr7:50343872	chr7:50343883
chr7:50343889	chr7:50343890	chr7:50343897
chr7:50343907	chr7:50343909	chr7:50343914
chr7:50343934	chr7:50343939	chr7:50343950
chr7:50343959	chr7:50343805	chr7:50343822
chr7:50343824	chr7:50343826	chr7:50343829
chr7:50343831	chr7:50343833	chr7:50343838
chr7:50343847	chr7:50343850	chr7:50343858
chr7:50343864	chr7:50343869	chr7:50343872
chr7:50343890		

or a corresponding cytosine at position n+1 on the opposite DNA strand.

To the extent that the method of the present invention includes analysing IRF4, the subject residues are:

chr6:392036	chr6:392047	chr6:392049
chr6:392057	chr6:392060	chr6:392066
chr6:392080	chr6:392094	chr6:392102
chr6:392131		

or a corresponding cytosine at position n+1 on the opposite DNA strand.

Another aspect of the present invention is directed to a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the level of expression of a DNA region selected from:

(i) the region, including 2kb upstream of the transcription start site, defined by any two or more of Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; or

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of any two or more of:

(1) BCAT1 (2) IKZF1 (3) IRF4 (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a lower level of expression of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a neoplastic state.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the genomic sequence of GRASP.

Figure 2 depicts the genomic sequence of CAHM.

Figure 3 depicts the genomic sequence of IRF4.

Figure 4 depicts the genomic sequence of BCAT1.

Figure 5 depicts the genomic sequence of IKZF1.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is predicated, in part, on the determination that the change in methylation pattern of a panel of gene markers can provide a higher level of either diagnostic specificity or diagnostic sensitivity when tested collectively than if any one of these markers is tested either in isolation or together with markers other than those specified herein. This finding has therefore now facilitated the development of an improved test for diagnosing, prognosing or monitoring neoplasms of the large intestine based on assessing the methylation of two or more of the gene markers BCAT1, IKZF1, CAHM, GRASP and IRF4.

In accordance with the present invention, it has been determined that certain specific panels of genes are modulated, in terms of differential changes to their levels of methylation, depending on whether or not the cell in issue is neoplastic or not. It should be understood that the genes in issue are described herein both by reference to their name and their chromosomal coordinates. The chromosomal coordinates are consistent with the human genome database version Hg19 which was released in February 2009 (herein referred to as "Hg19 coordinates").

Accordingly, one aspect of the present invention is directed to a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of a DNA region selected

from:

(i) the region, including 2kb upstream of the transcription start site, defined by any two or more of Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443;

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of any two or more of:

(1) BCAT1 (2) IKZF1 (3) IRF4 (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In one embodiment, said method is directed to identifying biological samples in which any one of said DNA regions exhibits a higher level of methylation.

In another embodiment, said method is directed to identifying biological samples in which two or more of said DNA regions exhibit a higher level of methylation.

Without limiting the present invention to any one theory or mode of action, the panel of DNA regions (markers) specified herein provide not only an improved diagnostic outcome relative to prior art methods but, in addition, enable the development of a screening method which can be designed to focus on providing either a high level of diagnostic specificity or a high level of diagnostic sensitivity. As would be understood by the person of skill in the art, in the context of diagnostics "sensitivity" defines the proportion of positive results which are correctly identified, that is, the percentage of individuals correctly identified as having the disease in issue. "Specificity", however, defines the proportion of negative results which are correctly identified, that is, the percentage of individuals correctly identified as not having the disease in issue.

In the context of the present invention it has been determined that screening a patient sample for the methylation status of the specified panel of markers can be designed to provide either a diagnostic result which exhibits a high level of specificity or a diagnostic result which exhibits a high level of sensitivity.

Where a high level of sensitivity is sought, the screening method is designed to identify samples in which any one of the markers of the panel exhibits increased methylation relative to

control levels. That is, not all of the markers are required to exhibit hypermethylation in order to define the result as positive. Where a higher level of sensitivity is sought, the level of specificity is inherently reduced. However, if it is desired to pursue a higher level of specificity (which may reduce the level of sensitivity) then the method is designed to identify samples in which two or more of the panel of DNA regions exhibit increased methylation.

Accordingly, to the extent that the present invention is directed to embodiments of the method in which "any one of" the DNA regions (markers) of the specified panel exhibits a higher level of methylation, these embodiments are designed to achieve results exhibiting increased sensitivity based on any one of the markers in the panel being hypermethylated. It should be understood that it need not be the same marker which is hypermethylated in each sample. Rather, it is simply that one of the markers which forms part of the panel is hypermethylated. It should also be understood that in relation to some samples, two or more of the markers of the panel may be hypermethylated. These samples should be understood to nevertheless fall within the scope of this embodiment of the invention since this embodiment is not excluding the situation where multiple markers are hypermethylated but is merely including within its scope all those samples where as few as one marker are hypermethylated. Overall these data will provide increased sensitivity but reduced specificity.

To the extent that the present invention is directed to embodiments of the method in which "two or more" of the DNA regions (markers) of the specified panel exhibit a higher level of methylation, these embodiments are designed to achieve a higher level of specificity. This is achieved by virtue of the fact that at least two, if not more, of the specified panel of markers are determined to be hypermethylated.

Reference to "large intestine" should be understood as a reference to a cell derived from one of the eight anatomical regions of the large intestine, which regions commence after the terminal region of the ileum, these being:

(i) the cecum;
(ii) the ascending colon;
(iii) the transverse colon;
(iv) the descending colon;
(v) the sigmoid colon;
(vi) the rectum;
(vii) the splenic flexure; and
(viii) the hepatic flexure.

Reference to "neoplasm" should be understood as a reference to a lesion, tumour or other encapsulated or unencapsulated mass or other form of growth which comprises neoplastic cells. A "neoplastic cell" should be understood as a reference to a cell exhibiting abnormal growth. The term "growth" should be understood in its broadest sense and includes reference to proliferation. In this regard, an example of abnormal cell growth is the uncontrolled proliferation of a cell. Another example is failed apoptosis in a cell, thus prolonging its usual life span. The neoplastic cell may be a benign cell or a malignant cell. In a preferred embodiment, the subject neoplasm is an adenoma or an adenocarcinoma. Without limiting the present invention to any one theory or mode of action, an adenoma is generally a benign tumour of epithelial origin which is either derived from epithelial tissue or exhibits clearly defined epithelial structures. These structures may take on a glandular appearance. It can comprise a malignant cell population within the adenoma, such as occurs with the progression of a benign adenoma or benign neoplastic legion to a malignant adenocarcinoma.

Preferably, said neoplastic cell is an adenoma or adenocarcinoma and even more preferably a colorectal adenoma or adenocarcinoma.

Reference to "DNA region" should be understood as a reference to a specific section of genomic DNA. These DNA regions are specified either by reference to a gene name or a set of chromosomal coordinates. Both the gene names and the chromosomal coordinates would be well known to, and understood by, the person of skill in the art. As detailed hereinbefore, the chromosomal coordinates correspond to the Hg19 version of the genome. In general, a gene can be routinely identified by reference to its name, via which both its sequences and chromosomal location can be routinely obtained, or by reference to its chromosomal coordinates, via which both the gene name and its sequence can also be routinely obtained.

Reference to each of the genes/DNA regions detailed above should be understood as a reference to all forms of these molecules and to fragments or variants thereof. As would be appreciated by the person of skill in the art, some genes are known to exhibit allelic variation between individuals or single nucleotide polymorphisms. SNPs encompass insertions and deletions of varying size and simple sequence repeats, such as dinucleotide and trinucleotide repeats. Variants include nucleic acid sequences from the same region sharing at least 90%, 95%, 98%, 99% sequence identity i.e. having one or more deletions, additions, substitutions, inverted sequences etc. relative to the DNA regions described herein. Accordingly, the present invention should be understood to extend to such variants which, in terms of the present diagnostic applications, achieve the same outcome despite the fact that minor genetic

variations between the actual nucleic acid sequences may exist between individuals. The present invention should therefore be understood to extend to all forms of DNA which arise from any other mutation, polymorphic or allelic variation.

It should be understood that the "individual" who is the subject of testing may be any human or non-human mammal. Examples of non-human mammals includes primates, livestock animals (e.g. horses, cattle, sheep, pigs, donkeys), laboratory test animals (e.g. mice, rats, rabbits, guinea pigs), companion animals (e.g. dogs, cats) and captive wild animals (e.g. deer, foxes).

Preferably the mammal is a human.

The panel of genes which has been identified demonstrates increased methylation in large intestine neoplastic cells relative to corresponding non-neoplastic cells. This increased methylation is, in some cases, localised to a few specific CpG sites within the DNA region while in other cases it is observed across a wide range of CpG sites. However, although specific regions from all genes exhibit increased methylation and are therefore useful diagnostic markers, it is the analysis of these specific markers as a panel which has provided an independent and very significant improvement over the sensitivity and specificity could be obtainable over these markers are analysed individually. This improvement is also significant even when considered against the sensitivity and specificity obtainable using other markers known to exhibit increased methylation as a marker of the onset of a large intestine neoplasm. Accordingly, the method of the present invention has now provided a means of achieving a higher level of sensitivity and specificity than has been achievable to date.

In one aspect there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393; and
(2) chr7:50344378...50472798;
and optionally one or more of (3) chr6:391739..411443, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1; and
(2) IKZF1;

and optionally one or more of (3) IRF4, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In one embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798; and
(3) chr6:391739..411443; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1 and (3) IRF4.

In another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798 and
(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1; and (4) GRASP.

In still another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798 and
(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1 and (5) CAHM

In still yet another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393
(2) chr7:50344378...50472798
(3) chr6:391739..411443; and
(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1, (3) IRF4 and (4) GRASP.

In yet still another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798
(3) chr6:391739..411443; and
(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1, (3) IRF4 and (5) CAHM.

In a further embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798
(4) chr12:52400748..52409671 and
(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1, (4) GRASP and (5) CAHM.

In yet another further embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798
(3) chr6:391739..411443
(4) chr12:52400748..52409671 and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1, (3) IRF4, (4) GRASP and (5) CAHM.

In another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393; and

(5) chr6:163834097..163834982;

and optionally one or more of (2) chr7:50344378...50472798, (3) chr6:391739..411443; and (4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1; and

(5) CAHM;

and optionally one or more of (2) IKZF1, (3) IRF4 and (4) GRASP

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In still another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(2) chr7:50344378..50472798; and

(5) chr6:163834097..163834982;

and optionally one or more of (1) chr12:24962958..25102393, (3) chr6:391739..411443 and (4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(2) IKZF1; and

(5) CAHM;

and optionally one or more of (1) BCAT1, (3) IRF4 and (4) GRASP

in a biological sample from said individual wherein a higher level of methylation of at least

one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In a further embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393; and
(3) chr6:391739..411443;
and optionally one or more of (2) chr7:50344378...50472798, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1; and
(3) IRF4;
and optionally one or more of (2) IKZF1, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In yet another further embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393; and
(4) chr12:52400748..52409671;
and optionally one or more of (2) chr7:50344378...50472798, (3) chr6:391739..411443 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1; and
(4) GRASP;
and optionally one or more of (2) IKZF1, (3) IRF4 and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large

intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In still another further embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(2) chr7:50344378..50472798; and
(3) chr6:391739..411443;
and optionally one or more of (1) chr12:24962958..25102393, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(2) IKZF1; and
(3) IRF4;
and optionally one or more of (1) BCAT1, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In yet still another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(2) chr7:50344378..50472798; and
(4) chr12:52400748..52409671;
and optionally one or more of (1) chr12:24962958..25102393, (3) chr6:391739..411443 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(2) IKZF1; and
(4) GRASP;
and optionally one or more of (1) BCAT1, (3) IRF4 and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In still yet another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(3) chr6:391739..411443; and

(4) chr12:52400748..52409671;

and optionally one or more of (1) chr12:24962958..25102393, (2) chr7:50344378...50472798 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(3) IRF4; and

(4) GRASP;

and optionally one or more of (1) BCAT1, (2) IKZF1 and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In another embodiment there is provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(3) chr6:391739..411443; and

(5) chr6:163834097..163834982;

and optionally one or more of (1) chr12:24962958..25102393, (2) chr7:50344378...50472798 and (4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(3) IRF4; and

(5) CAHM;

and optionally one or more of (1) BCAT1, (2) IKZF1 and (4) GRASP

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

In accordance with these aspects and embodiments, in yet another embodiment said

control level is a non-neoplastic level.

In one embodiment, said method is directed to identifying biological samples in which any one of said DNA regions exhibits a higher level of methylation.

In another embodiment, said method is directed to identifying biological samples in which two or more of said DNA regions exhibits a higher level of methylation.

Still further, in another embodiment, said large intestine tissue is preferably colorectal tissue.

In still another embodiment, the neoplasm is malignant, such as a carcinoma.

In a further embodiment, the neoplasm is non-malignant, such as an adenoma.

In terms of screening for the methylation of these gene regions, it should be understood that the assays can be designed to screen either the specific regions listed herein (which correspond to the "plus" strand of the gene) or the complementary "minus" strand. It is well within the skill of the person in the art to choose which strand to analyse and to target that strand based on the chromosomal coordinates provided herein. In some circumstances, assays may be established to screen both strands.

It should be understood that one may screen for the specified panel of markers exclusively or one may elect to additionally screen for other markers, such as other DNA hypermethylation markers, other RNA expression level markers or other protein markers. These other markers may, for example, provide additional information in relation to the health status of the patient in issue.

Without limiting the present invention to any one theory or mode of action, although measuring the methylation levels across these DNA regions is diagnostic of a large intestine neoplastic condition, it has been determined that discrete subregions are particularly useful in this regard since these subregions contain a high density of CpG dinucleotides which are frequently hypermethylated in large intestine neoplasias, such as colorectal cancers. This finding renders these subregions a particularly useful target for analysis since it both simplifies the screening process due to a shorter more clearly defined region of DNA requiring analysis and, further, the fact that the results from these regions will provide a significantly more definitive result in relation to the presence, or not, of hypermethylation than would be obtained if analysis was performed across the DNA region as a whole. This finding therefore both simplifies the screening process and increases the sensitivity and specificity of large intestine neoplasia diagnosis.

The subregions which have been determined to exhibit particular utility are listed below with reference to the gene and chromosomal region within which they are found:

(1) BCAT subregions chr12:25101992-25102093 (SEQ ID NO:1 or corresponding minus strand) and chr12:25101909-25101995 (SEQ ID NO:2 or corresponding minus strand)

(2) IKZF1 subregions: chr7:50343867-50343961 (SEQ ID NO:3 or corresponding minus strand) and chr7:50343804-5033895 (SEQ ID NO:4 or corresponding minus strand)

(3) IRF4 subregions chr6:392036-392145 (SEQ ID NO:5 or corresponding minus strand)

(4) GRASP subregions: chr12:52399672-52399922, chr12:52400821-52401051 (SEQ ID NO:6 or corresponding minus strand), chr12:52401407-52401664 (SEQ ID NO:7 or corresponding minus strand) chr12:52400866-52400973 and Chr12:52401107-52401664.

(5) CAHM subregions: chr6:163834295-163834500 (SEQ ID NO:8), chr6:163834621-163834906, chr6:163834393-163834455 and chr6:163834393-163834519.

Without limiting the present invention to any one theory or mode of action, the skilled person may screen one or more subregions for each gene marker.

In one embodiment, the methylation marker subregions tested for each selected gene marker are:

(1) The BCAT subregion defined by SEQ ID NO:1 or SEQ ID NO:2 or corresponding minus strand;

(2) The IKZF1 subregion defined by SEQ ID NO:3 or SEQ ID NO:4 or corresponding minus strand;

(3) The IRF4 subregion defined by SEQ ID NO:5 or corresponding minus strand;

(4) The GRASP subregion defined by SEQ ID NO:6 or 7 or corresponding minus strands; and

(5) The CAHM subregion defined by SEQ ID NO:8 or corresponding minus strand.

Without limiting the present invention to any one theory or mode of action, DNA methylation is universal in bacteria, plants, and animals. DNA methylation is a type of chemical modification of DNA that is stable over rounds of cell division but does not involve changes in the underlying DNA sequence of the organism. Chromatin and DNA modifications are two important features of epigenetics and play a role in the process of cellular differentiation, allowing cells to stably maintain different characteristics despite containing the same genomic material. In eukaryotic organisms DNA methylation occurs only at the number

5 carbon of the cytosine pyrimidine ring. In mammals, DNA methylation occurs mostly at the number 5 carbon of the cytosine of a CpG dinucleotide. CpG dinucleotides comprise approximately 1% human genome.

70-80% of all CpGs are methylated. CpGs may be grouped in clusters called "CpG islands" that are present in the 5' regulatory regions of many genes and are frequently unmethylated. In many disease processes such as cancer, gene promoters and/or CpG islands acquire abnormal hypermethylation, which is associated with heritable transcriptional silencing. DNA methylation may impact the transcription of genes in two ways. First, the methylation of DNA may itself physically impede the binding of transcriptional proteins to the gene, thus blocking transcription. Second, methylated DNA may be bound by proteins known as Methyl-CpG-binding domain proteins (MBDs). MBD proteins then recruit additional proteins to the locus, such as histone deacetylases and other chromatin remodelling proteins that can modify histones, thereby forming compact, inactive chromatin termed silent chromatin. This link between DNA methylation and chromatin structure is very important. In particular, loss of Methyl-CpG-binding Protein 2 (MeCP2) has been implicated in Rett syndrome and Methyl-CpG binding domain protein 2 (MBD2) mediates the transcriptional silencing of hypermethylated genes in cancer.

In humans, the process of DNA methylation is carried out by three enzymes, DNA methyltransferase 1, 3a and 3b (DNMT1, DNMT3a, DNMT3b). It is thought that DNMT3a and DNMT3b are the *de novo* methyltransferases that set up DNA methylation patterns early in development. DNMT1 is the proposed maintenance methyltransferase that is responsible for copying DNA methylation patterns to the daughter strands during DNA replication. DNMT3L is a protein that is homologous to the other DNMT3s but has no catalytic activity. Instead, DNMT3L assists the *de novo* methyltransferases by increasing their ability to bind to DNA and stimulating their activity. Finally, DNMT2 has been identified as an "enigmatic" DNA methylstransferase homolog, containing all 10 sequence motifs common to all DNA methyltransferases; however, DNMT2 may not methylate DNA but instead has been shown to methylate a small RNA.

"Methylation status" should therefore be understood as a reference to the presence, absence and/or quantity of methylation at a particular nucleotide, or nucleotides, within a DNA region. The methylation status of a particular DNA sequence (e.g. DNA region as described herein) can indicate the methylation state of every base in the sequence or can indicate the methylation state of a subset of the base pairs (e.g., of cytosines or the methylation state of one

or more specific restriction enzyme recognition sequences) within the sequence, or can indicate information regarding regional methylation density within the sequence without providing precise information of where in the sequence the methylation occurs. The methylation status can optionally be represented or indicated by a "methylation value." A methylation value can be generated, for example, by quantifying the amount of intact DNA present following restriction digestion with a methylation dependent restriction enzyme. In this example, if a particular sequence in the DNA is quantified using quantitative PCR, an amount of template DNA approximately equal to a mock treated control indicates the sequence is not highly methylated whereas an amount of template substantially less than occurs in the mock treated sample indicates the presence of methylated DNA at the sequence. Accordingly, a value, i.e., a methylation value, for example from the above described example, represents the methylation status and can thus be used as a quantitative indicator of the methylation status. This is of particular use when it is desirable to compare the methylation status of a sequence in a sample to a threshold value.

The method of the present invention is predicated on the comparison of the level of methylation of specific DNA regions of a biological sample with the control methylation levels of these DNA regions. The "control level" is the "normal level", which is the level of methylation of the DNA region of a corresponding large intestine cell or cellular population which is not neoplastic or in another biological sample from which DNA may be isolated for assay.

The normal (or "non-neoplastic") methylation level may be determined using non-neoplastic tissues derived from the same individual who is the subject of testing. However, it would be appreciated that this may be quite invasive for the individual concerned and it is therefore likely to be more convenient to analyse the test results relative to a standard result which reflects individual or collective results obtained from individuals other than the patient in issue. This latter form of analysis is in fact the preferred method of analysis since it enables the design of kits which require the collection and analysis of a single biological sample, being a test sample of interest. The standard results which provide the normal methylation level may be calculated by any suitable means which would be well known to the person of skill in the art. For example, a population of normal tissues can be assessed in terms of the level of methylation of the genes of the present invention, thereby providing a standard value or range of values against which all future test samples are analysed. It should also be understood that the normal level may be determined from the subjects of a specific cohort and for use with

respect to test samples derived from that cohort. Accordingly, there may be determined a number of standard values or ranges which correspond to cohorts which differ in respect of characteristics such as age, gender, ethnicity or health status. Said "normal level" may be a discrete level or a range of levels. An increase in the methylation level of the subject genes relative to normal levels is indicative of the tissue being neoplastic.

The term "methylation" shall be taken to mean the presence of a methyl group added by the action of a DNA methyl transferase enzyme to a cytosine base or bases in a region of nucleic acid, e.g. genomic DNA. As described herein, there are several methods known to those skilled in the art for determining the level or degree of methylation of nucleic acid.

By "higher level" is meant that there are a higher number of methylated CpG dinucleotides in the subject diagnosed than in a control sample, that is, either the proportion of DNA molecules methylated at a particular CpG site is higher or there are a higher number of separate CpG sites methylated in the subject. It should be understood that the terms "enhanced" and "increased" are used interchangeably with the term "higher".

In relation to detecting a "higher level" of methylation, it should be understood that in some situations the normal/control level will in fact correspond to the absence of any detectable methylation while the neoplastic level will correspond to the presence of methylation, *per se*. In this situation the diagnostic method is relatively simple since one need only screen for the mere presence of methylation (i.e. a qualitative assessment only), rather than assessing the methylation levels relative to a control level of methylation, which analysis necessarily involves a measure of quantification. Without limiting the present invention in any way, it is observed in blood-derived samples, for example, that in the context of some markers the change in methylation of that marker upon the onset of neoplasia is a shift from undetectable levels of methylation to the presence of detectable methylation. In these situations a relatively simple qualitative assessment is enabled where one need only screen a test sample to determine the presence or not of methylation. In the context of the definitions provided herein, reference to "higher level" encompasses both a relative increase in the level of methylation of a marker or the onset of methylation where previously none was evident. As detailed hereinbefore, the control level may be newly assessed for each patient or there may be a standard result against which all test samples are assessed. Where it is known that methylation is not present on the marker of interest, one need only screen for the presence or not of methylation since the control level is the absence of methylation and the "higher level" is thereby the presence of any amount of methylation.

The present invention is not to be limited by a precise number of methylated residues that are considered to be diagnostic of neoplasia in a subject, because some variation between patient samples will occur. The present invention is also not limited by positioning of the methylated residue.

Nevertheless, a number of specific cytosine residues which undergo hypermethylation within these subregions have also been identified. In another embodiment, therefore, a screening method can be employed which is specifically directed to assessing the methylation status of one or more of either these residues or the corresponding cytosine at position n+1 on the opposite DNA strand.

To this end, detailed in Table 1 are the cytosine residues which have been identified in this regard. It should be appreciated by the person of skill in the art that these individual residues are numbered by reference to Hg19, which also corresponds to the numbering of the specific subregions listed hereinbefore and which can be further identified when the coordinate numbering for each subregion is applied to the corresponding subregion sequences which are provided in the sequence listing. It should be understood that these residues have been identified in the context of the subregion DNA. However, there are other residues which are hypermethylated outside the subregions themselves but within the larger DNA region from which the subregions derive. Accordingly, these specified residues represent a particularly useful subset of individual cytosine residues which undergo hypermethylation within the context of the DNA regions and subregions herein disclosed. These individual residues are grouped below according to the DNA region within which they occur. These DNA regions are identified by reference to both the Hg19 chromosomal coordinates and the gene region name.

To the extent that the method of the present invention includes analysing the methylation of GRASP, the subject residues are:

chr12:52399713	chr12:52399731	chr12:52399749	chr12:52399783
chr12:52399796	chr12:52399808	chr12:52399823	chr12:52399835
chr12:52399891			
chr12:52400847	chr12:52400850	chr12:52400859	chr12:52400866
chr12:52400869	chr12:52400873	chr12:52400881	chr12:52400886
chr12:52400893	chr12:52400895	chr12:52400899	chr12:52400902
chr12:52400907	chr12:52400913	chr12:52400919	chr12:52400932
chr12:52400938	chr12:52400958	chr12:52400962	chr12:52400971
chr12:52400973	chr12:52400976	chr12:52400998	chr12:52401008
chr12:52401010	chr12:52401012	chr12:52401016	chr12:52401019
chr12:52401025	chr12:52401041	chr12:52401044	chr12:52401053
chr12:52401060	chr12:52401064	chr12:52401092	chr12:52401118

chr12:52401438	chr12:52401448	chr12:52401460	chr12:52401465
chr12:52401474	chr12:52401477	chr12:52401479	chr12:52401483
chr12:52401504	chr12:52401514	chr12:52401523	chr12:52401540
chr12:52401553	chr12:52401576	chr12:52401588	chr12:52401595
chr12:52401599	chr12:52401604	chr12:52401606	chr12:52401634
chr12:52401640	chr12:52401644	chr12:52401659	

chr12:52401160	chr12:52401165	chr12:52401174	chr12:52401177
chr12:52401179	chr12:52401183	chr12:52401204	chr12:52401215
chr12:52401223	chr12:52401240	chr12:52401253	chr12:52401288
chr12:52401295	chr12:52401299	chr12:52401304	chr12:52401334
chr12:52401340	chr12:52401344	chr12:52401359	

or a corresponding cytosine at position n+1 on the opposite DNA strand.

To the extent that the method of the present invention includes analysing CAHM, the subject residues are:

chr6:163834330	chr6:163834332	chr6:163834357
chr6:163834373	chr6:163834384	chr6:163834390
chr6:163834392	chr6:163834406	chr6:163834412
chr6:163834419	chr6:163834443	chr6:163834448
chr6:163834452	chr6:163834464	chr6:163834483
chr6:163834653	chr6:163834660	chr6:163834672
chr6:163834675	chr6:163834678	chr6:163834681
chr6:163834815	chr6:163834824	chr6:163834835
chr6:163834840	chr6:163834853	chr6:163834855
chr6:163834858	chr6:163834863	chr6:163834869
chr6:163834872		

or a corresponding cytosine at position n+1 on the opposite DNA strand.

To the extent that the method of the present invention includes analysing IKZF1, the subject residues are:

chr7:50343869	chr7:50343872	chr7:50343883
chr7:50343889	chr7:50343890	chr7:50343897
chr7:50343907	chr7:50343909	chr7:50343914
chr7:50343934	chr7:50343939	chr7:50343950
chr7:50343959	chr7:50343805	chr7:50343822
chr7:50343824	chr7:50343826	chr7:50343829
chr7:50343831	chr7:50343833	chr7:50343838
chr7:50343847	chr7:50343850	chr7:50343858
chr7:50343864	chr7:50343869	chr7:50343872
chr7:50343890		

or a corresponding cytosine at position n+1 on the opposite DNA strand.

To the extent that the method of the present invention includes analysing IRF4, the subject residues are:

chr6:392036	chr6:392047	chr6:392049
chr6:392057	chr6:392060	chr6:392066
chr6:392080	chr6:392094	chr6:392102
chr6:392131		

or a corresponding cytosine at position n+1 on the opposite DNA strand.

The detection method of the present invention can be performed on any suitable biological sample. To this end, reference to a "biological sample" should be understood as a reference to any sample of biological material derived from an animal such as, but not limited to, cellular material, biofluids (e.g. blood), faeces, tissue biopsy specimens, surgical specimens or fluid which has been introduced into the body of an animal and subsequently removed (such as, for example, the solution retrieved from an enema wash). The biological sample which is tested according to the method of the present invention may be tested directly or may require some form of treatment prior to testing. For example, a biopsy or surgical sample may require homogenisation prior to testing or it may require sectioning for *in situ* testing of the qualitative expression levels of individual genes. Alternatively, a cell sample may require permeabilisation prior to testing. Further, to the extent that the biological sample is not in liquid form, (if such form is required for testing) it may require the addition of a reagent, such as a buffer, to mobilise the sample.

To the extent that the DNA region of interest is present in a biological sample, the biological sample may be directly tested or else all or some of the nucleic acid present in the biological sample may be isolated prior to testing. In yet another example, the sample may be partially purified or otherwise enriched prior to analysis. For example, to the extent that a biological sample comprises a very diverse cell population, it may be desirable to enrich for a sub-population of particular interest. It is within the scope of the present invention for the target cell population or molecules derived therefrom to be treated prior to testing, for example, inactivation of live virus. It should also be understood that the biological sample may be freshly harvested or it may have been stored (for example by freezing) prior to testing or otherwise treated prior to testing (such as by undergoing culturing).

The choice of what type of sample is most suitable for testing in accordance with the method disclosed herein will be dependent on the nature of the situation. Preferably, said sample is a faecal (stool) sample, enema wash, surgical resection, tissue biopsy or blood sample (e.g. whole blood, serum or plasma).

More preferably, said biological sample is a blood sample, biopsy sample or stool sample.

As detailed hereinbefore, the present invention is designed to screen for a neoplastic cell or cellular population, which is located in the large intestine. Accordingly, reference to "cell or cellular population" should be understood as a reference to an individual cell or a group of cells. Said group of cells may be a diffuse population of cells, a cell suspension, an encapsulated population of cells or a population of cells which take the form of tissue.

Reference to the "onset" of a neoplasm, such as adenoma or adenocarcinoma, should be understood as a reference to one or more cells of that individual exhibiting dysplasia. In this regard, the adenoma or adenocarcinoma may be well developed in that a mass of dysplastic cells has developed. Alternatively, the adenoma or adenocarcinoma may be at a very early stage in that only relatively few abnormal cell divisions have occurred at the time of diagnosis. The present invention also extends to the assessment of an individual's predisposition to the development of a neoplasm, such as an adenoma or adenocarcinoma. Without limiting the present invention in any way, changed methylation levels may be indicative of that individual's predisposition to developing a neoplasia, such as the future development of an adenoma or adenocarcinoma or another adenoma or adenocarcinoma.

Although the preferred method is to assess methylation levels for the purpose of diagnosing neoplasia development or predisposition thereto, the detection of converse changes in the levels of said methylation may be desired under certain circumstances, for example, to monitor the effectiveness of therapeutic or prophylactic treatment directed to modulating a neoplastic condition, such as adenoma or adenocarcinoma development. For example, where elevated levels of methylation indicate that an individual has developed a condition characterised by adenoma or adenocarcinoma development, screening for a decrease in the levels of methylation subsequently to the onset of a therapeutic treatment regime may be utilised to indicate successful clearance of the neoplastic cells. In another example, one can use this method to test the tissue at the margins of a tumour resection in order to determine whether the full margin of the tumour has been removed.

The present method can therefore be used in the diagnosis, prognosis, classification, prediction of disease risk, detection of recurrence of disease, selection of treatment of a number of types of neoplasias and monitoring of neoplasias. A cancer at any stage of progression can be detected, such as primary, metastatic, and recurrent cancers.

The present invention provides methods for determining whether a mammal (e.g., a human) has a neoplasia of the large intestine, whether a biological sample taken from a mammal contains neoplastic cells or DNA derived from neoplastic cells, estimating the risk or

likelihood of a mammal developing a neoplasm, monitoring the efficacy of anti-cancer treatment, or selecting the appropriate anti-cancer treatment in a mammal with cancer. Such methods are based on the determination that neoplastic cells have a different methylation status than normal cells in the DNA regions described herein. Accordingly, by determining whether or not a cell contains differentially methylated sequences in the DNA regions as described herein, it is possible to determine that a cell is neoplastic .

The method of the invention can be used to evaluate individuals known or suspected to have a neoplasia or as a routine clinical test, i.e., in an individual not necessarily suspected to have a neoplasia. Further diagnostic assays can be performed to confirm the status of neoplasia in the individual.

Further, the present methods may be used to assess the efficacy of a course of treatment. For example, the efficacy of an anti-cancer treatment can be assessed by monitoring DNA methylation of the sequences described herein over time in a mammal having cancer. For example, a reduction or absence of methylation in any of the diagnostic sequences of the invention in a biological sample taken from a mammal following a treatment, compared to a level in a sample taken from the mammal before, or earlier in, the treatment, indicates efficacious treatment.

The method of the present invention is therefore useful as a one-time test or as an on-going monitor of those individuals thought to be at risk of neoplasia development or as a monitor of the effectiveness of therapeutic or prophylactic treatment regimes directed to inhibiting or otherwise slowing neoplasia development. In these situations, mapping the modulation of methylation levels in any one or more classes of biological samples is a valuable indicator of the status of an individual or the effectiveness of a therapeutic or prophylactic regime which is currently in use. Accordingly, the method of the present invention should be understood to extend to monitoring for increases or decreases in methylation levels in an individual relative to their normal level (as hereinbefore defined), or relative to one or more earlier methylation levels determined from a biological sample of said individual.

The methods for detecting neoplasia can comprise the detection of one or more other cancer-associated polynucleotide or polypeptides sequences. Accordingly, detection of methylation by the method of the invention can be used either alone, or in combination with other screening methods for the diagnosis or prognosis of neoplasia.

Any method for detecting DNA methylation can be used in the methods of the present invention. A number of methods are available for detection of differentially methylated DNA

at specific loci in either primary tissue samples or in patient samples such as blood, urine, stool or saliva (reviewed in Kristensen and Hansen *Clin Chem.* 55:1471-83, 2009; Ammerpohl *et al. Biochim Biophys Acta.* 1790:847-62, 2009; Shames *et al. Cancer Lett.* 251:187-98, 2007; Clark *et al. Nat Protoc.* 1:2353-64, 2006). For analysis of the proportion or extent of DNA methylation in a target gene, DNA is normally treated with sodium bisulfite and regions of interest amplified using primers and PCR conditions that will amplify independently of the methylation status of the DNA. The methylation of the overall amplicon or individual CpG sites can then be assessed by sequencing, including pyrosequencing, restriction enzyme digestion (COBRA) or by melting curve analysis. Alternatively ligation-based methods for analysis of methylation at specific CpG sites may be used. Detection of aberrantly methylated DNA released from tumours and into bodily fluids is being developed as a means of cancer diagnosis. Here, in the case of hypermethylated sequences, it is necessary to use sensitive methods that allow the selective amplification of the methylated DNA sequence from a background of normal cellular DNA that is unmethylated. Such methods based on bisulfite-treated DNA, for example; include methylation selective PCR (MSP), Heavymethyl PCR, Headloop PCR and Helper-dependent chain reaction (PCT/AU2008/001475).

Briefly, in some embodiments, methods for detecting methylation include randomly shearing or randomly fragmenting the genomic DNA, cutting the DNA with a methylation-dependent or methylation-sensitive restriction enzyme and subsequently selectively identifying and/or analyzing the cut or uncut DNA. Selective identification can include, for example, separating cut and uncut DNA (e.g., by size) and quantifying a sequence of interest that was cut or, alternatively, that was not cut. See, e.g., U.S. Pat. No. 7,186,512. Alternatively, the method can encompass amplifying intact DNA after restriction enzyme digestion, thereby only amplifying DNA that was not cleaved by the restriction enzyme in the area amplified. See, e.g., U.S. patent application Ser. Nos. 10/971,986; 11/071,013; and 10/971,339. In some embodiments, amplification can be performed using primers that are gene specific. Alternatively, adaptors can be added to the ends of the randomly fragmented DNA, the DNA can be digested with a methylation-dependent or methylation-sensitive restriction enzyme, intact DNA can be amplified using primers that hybridize to the adaptor sequences. In this case, a second step can be performed to determine the presence, absence or quantity of a particular gene in an amplified pool of DNA. In some embodiments, the DNA is amplified using real-time, quantitative PCR.

In some embodiments, the methods comprise quantifying the average methylation

density in a target sequence within a population of genomic DNA. In some embodiments, the method comprises contacting genomic DNA with a methylation-dependent restriction enzyme or methylation-sensitive restriction enzyme under conditions that allow for at least some copies of potential restriction enzyme cleavage sites in the locus to remain uncleaved; quantifying intact copies of the locus; and comparing the quantity of amplified product to a control value representing the quantity of methylation of control DNA, thereby quantifying the average methylation density in the locus compared to the methylation density of the control DNA.

The quantity of methylation of a locus of DNA can be determined by providing a sample of genomic DNA comprising the locus, cleaving the DNA with a restriction enzyme that is either methylation-sensitive or methylation-dependent, and then quantifying the amount of intact DNA or quantifying the amount of cut DNA at the DNA locus of interest. The amount of intact or cut DNA will depend on the initial amount of genomic DNA containing the locus, the amount of methylation in the locus, and the number (i.e., the fraction) of nucleotides in the locus that are methylated in the genomic DNA. The amount of methylation in a DNA locus can be determined by comparing the quantity of intact DNA or cut DNA to a control value representing the quantity of intact DNA or cut DNA in a similarly-treated DNA sample. The control value can represent a known or predicted number of methylated nucleotides. Alternatively, the control value can represent the quantity of intact or cut DNA from the same locus in another (e.g., normal, non-diseased) cell or a second locus.

By using at least one methylation-sensitive or methylation-dependent restriction enzyme under conditions that allow for at least some copies of potential restriction enzyme cleavage sites in the locus to remain uncleaved and subsequently quantifying the remaining intact copies and comparing the quantity to a control, average methylation density of a locus can be determined. A methylation-sensitive enzyme is one which cuts DNA if its recognition sequence is unmethylated while a methylation-dependent enzyme cuts DNA if its recognition sequence is methylated. If the methylation-sensitive restriction enzyme is contacted to copies of a DNA locus under conditions that allow for at least some copies of potential restriction enzyme cleavage sites in the locus to remain uncleaved, then the remaining intact DNA will be directly proportional to the methylation density, and thus may be compared to a control to determine the relative methylation density of the locus in the sample. Similarly, if a methylation-dependent restriction enzyme is contacted to copies of a DNA locus under conditions that allow for at least some copies of potential restriction enzyme cleavage sites in the locus to remain uncleaved, then the remaining intact DNA will be inversely proportional to

the methylation density, and thus may be compared to a control to determine the relative methylation density of the locus in the sample. Such assays are disclosed in, e.g., U.S. patent application Ser. No. 10/971,986.

Kits for the above methods can include, e.g., one or more of methylation-dependent restriction enzymes, methylation-sensitive restriction enzymes, amplification (e.g., PCR) reagents, probes and/or primers.

Quantitative amplification methods (e.g., quantitative PCR or quantitative linear amplification) can be used to quantify the amount of intact DNA within a locus flanked by amplification primers following restriction digestion. Methods of quantitative amplification are disclosed in, e.g., U.S. Pat. Nos. 6,180,349; 6,033,854; and 5,972,602, as well as in, e.g., Gibson *et al.*, *Genome Research* 6:995-1001 (1996); DeGraves, *et al.*, *Biotechniques* 34(1):106-10, 112-5 (2003); Deiman B, *et al.*, *Mol. Biotechnol.* 20(2):163-79 (2002). Amplifications may be monitored in "real time."

Additional methods for detecting DNA methylation can involve genomic sequencing before and after treatment of the DNA with bisulfite. See, e.g., Frommer *et al., Proc. Natl. Acad. Sci. USA* 89:1827-1831 (1992). When sodium bisulfite is contacted to DNA, unmethylated cytosine is converted to uracil, while methylated cytosine is not modified.

In some embodiments, restriction enzyme digestion of PCR products amplified from bisulfite-converted DNA is used to detect DNA methylation. See, e.g., Sadri & Hornsby, *Nucl. Acids Res.* 24:5058-5059 (1996); Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534 (1997).

In some embodiments, a methylation-specific PCR ("MSP") reaction is used alone or in combination with other methods to detect DNA methylation. An MSP assay entails initial modification of DNA by sodium bisulfite, converting all unmethylated, but not methylated, cytosines to uracil, and subsequent amplification with primers specific for methylated verses unmethylated DNA. See, Herman *et al. Proc. Natl. Acad. Sci. USA* 93:9821-9826 (1996); U.S. Pat. No. 5,786,146.

In some embodiments, a MethyLight assay is used alone or in combination with other methods to detect DNA methylation (see, Eads *et al., Cancer Res.* 59:2302-2306 (1999)). Briefly, in the MethyLight process genomic DNA is converted in a sodium bisulfite reaction (the bisulfite process converts unmethylated cytosine residues to uracil). Amplification of a DNA sequence of interest is then performed using PCR primers that hybridize to CpG dinucleotides. By using primers that hybridize only to sequences resulting from bisulfite conversion of methylated DNA, (or alternatively to unmethylated sequences) amplification can

indicate methylation status of sequences where the primers hybridize. Furthermore, the amplification product can be detected with a probe that specifically binds to a sequence resulting from bisulfite treatment of a unmethylated DNA. If desired, both primers and probes can be used to detect methylation status. Thus, kits for use with MethyLight can include sodium bisulfite as well as primers or detectably-labelled probes (including but not limited to Taqman or molecular beacon probes) that distinguish between methylated and unmethylated DNA that have been treated with bisulfite. Other kit components can include, e.g., reagents necessary for amplification of DNA including but not limited to, PCR buffers, deoxynucleotides; and a thermostable polymerase.

In some embodiments, a Ms-SNuPE (Methylation-sensitive Single Nucleotide Primer Extension) reaction is used alone or in combination with other methods to detect DNA methylation (see, Gonzalgo & Jones, *Nucleic Acids Res.* 25:2529-2531 (1997)). The Ms-SNuPE technique is a quantitative method for assessing methylation differences at specific CpG sites based on bisulfite treatment of DNA, followed by single-nucleotide primer extension (Gonzalgo & Jones, *supra*). Briefly, genomic DNA is reacted with sodium bisulfite to convert unmethylated cytosine to uracil while leaving 5-methylcytosine unchanged. Amplification of the desired target sequence is then performed using PCR primers specific for bisulfite-converted DNA, and the resulting product is isolated and used as a template for methylation analysis at the CpG site(s) of interest.

Typical reagents (e.g., as might be found in a typical Ms-SNuPE-based kit) for Ms-SNuPE analysis can include, but are not limited to: PCR primers for specific gene (or methylation-altered DNA sequence or CpG island); optimized PCR buffers and deoxynucleotides; gel extraction kit; positive control primers; Ms-SNuPE primers for a specific gene; reaction buffer (for the Ms-SNuPE reaction); and detectably-labelled nucleotides. Additionally, bisulfite conversion reagents may include: DNA denaturation buffer; sulfonation buffer; DNA recovery regents or kit (e.g., precipitation, ultrafiltration, affinity column); desulfonation buffer; and DNA recovery components.

Additional methylation detection methods include, but are not limited to, methylated CpG island amplification (see, Toyota *et al.*, *Cancer Res.* 59:2307-12 (1999)) and those described in, e.g., U.S. Patent Publication 2005/0069879; Rein, *et al. Nucleic Acids Res.* 26 (10): 2255-64 (1998); Olek, *et al. Nat. Genet.* 17(3): 275-6 (1997); and PCT Publication No. WO 00/70090.

More detailed information in relation to several of these generally described methods is

provided below:

(a) Probe or Primer Design and/or Production

Several methods described herein for the diagnosis of a neoplasia use one or more probes and/or primers. Methods for designing probes and/or primers for use in, for example, PCR or hybridization are known in the art and described, for example, in Dieffenbach and Dveksler (Eds) (In: PCR Primer: A Laboratory Manual, Cold Spring Harbor Laboratories, NY, 1995). Furthermore, several software packages are publicly available that design optimal probes and/or primers for a variety of assays, e.g. Primer 3 available from the Center for Genome Research, Cambridge, Mass., USA.

Clearly, the potential use of the probe or primer should be considered during its design. For example, should the probe or primer be produced for use in a methylation specific PCR or ligase chain reaction (LCR) assay the nucleotide at the 3' end (or 5' end in the case of LCR) should preferably correspond to a methylated nucleotide in a nucleic acid.

Probes and/or primers useful for detection of a sequence associated with a neoplasia are assessed, for example, to determine those that do not form hairpins, self-prime or form primer dimers (e.g. with another probe or primer used in a detection assay). Furthermore, a probe or primer (or the sequence thereof) is often assessed to determine the temperature at which it denatures from a target nucleic acid (i.e. the melting temperature of the probe or primer, or Tm). Methods for estimating Tm are known in the art and described, for example, in Santa Lucia, *Proc. Natl. Acad. Sci. USA*, 95: 1460-1465, 1995 or Bresslauer *et al.*, *Proc. Natl. Acad. Sci. USA*, 83: 3746-3750, 1986.

Methods for producing/synthesizing a probe or primer of the present invention are known in the art. For example, oligonucleotide synthesis is described, in Gait (Ed) (In: Oligonucleotide Synthesis: A Practical Approach, IRL Press, Oxford, 1984). For example, a probe or primer may be obtained by biological synthesis (e.g. by digestion of a nucleic acid with a restriction endonuclease) or by chemical synthesis. For short sequences (up to about 100 nucleotides) chemical synthesis is preferable.

For longer sequences standard replication methods employed in molecular biology are useful, such as, for example, the use of M13 for single stranded DNA as described by Messing, *Methods Enzymol*, 101, 20-78, 1983. Other methods for oligonucleotide synthesis include, for example, phosphotriester and phosphodiester methods (Narang, *et al. Meth. Enzymol* 68: 90, 1979) and synthesis on a support (Beaucage, *et al. Tetrahedron Letters* 22:1859-1862, 1981) as

well as phosphoramidate technique, Caruthers, M. H., *et al.*, *Methods in Enzymology*, Vol. 154, pp. 287-314 (1988), and others described in "Synthesis and Applications of DNA and RNA," S. A. Narang, editor, Academic Press, New York, 1987, and the references cited therein. Probes comprising locked nucleic acid (LNA) are synthesized as described, for example, in Nielsen *et al. J. Chem. Soc. Perkin Trans.*, 1:3423, 1997; Singh and Wengel, *Chem. Commun.* 1247, 1998. While, probes comprising peptide-nucleic acid (PNA) are synthesized as described, for example, in Egholm *et al.*, *Am. Chem. Soc.*, 114:1895, 1992; Egholm *et al.*, *Nature*, 365:566, 1993; and Orum *et al.*, *Nucl. Acids Res.*, 21:5332, 1993.

(b) Methylation-Sensitive Endonuclease Digestion of DNA

In one example, the increased methylation in a sample is determined using a process comprising treating the nucleic acid with an amount of a methylation-sensitive restriction endonuclease enzyme under conditions sufficient for nucleic acid to be digested and then detecting the fragments produced. Exemplary methylation-sensitive endonucleases include, for example, HhaI or HpaII. Preferably, assays include internal controls that are digested with a methylation-insensitive enzyme having the same specificity as the methylation-sensitive enzyme employed. For example, the methylation-insensitive enzyme MspI is an isoschizomer of the methylation-sensitive enzyme HpaII.

Hybridization Assay Formats

In one example, the digestion of nucleic acid is detected by selective hybridization of a probe or primer to the undigested nucleic acid. Alternatively, the probe selectively hybridizes to both digested and undigested nucleic acid but facilitates differentiation between both forms, e.g., by electrophoresis. Suitable detection methods for achieving selective hybridization to a hybridization probe include, for example, Southern or other nucleic acid hybridization (Kawai *et al.*, *Mol. Cell. Biol.* 14:7421-7427, 1994; Gonzalgo *et al.*, *Cancer Res.* 57:594-599, 1997).

Suitable hybridization conditions are determined based on the melting temperature (Tm) of a nucleic acid duplex comprising the probe. The skilled artisan will be aware that optimum hybridization reaction conditions should be determined empirically for each probe, although some generalities can be applied. Preferably, hybridizations employing short oligonucleotide probes are performed at low to medium stringency. In the case of a GC rich probe or primer or a longer probe or primer a high stringency hybridization and/or wash is preferred. A high stringency is defined herein as being a hybridization and/or wash carried out in about 0.1 x SSC buffer and/or about 0.1% (w/v) SDS, or lower salt concentration, and/or at

a temperature of at least 65°C., or equivalent conditions. Reference herein to a particular level of stringency encompasses equivalent conditions using wash/hybridization solutions other than SSC known to those skilled in the art.

In accordance with the present example, a difference in the fragments produced for the test sample and a negative control sample is indicative of the subject having a neoplasia. Similarly, in cases where the control sample comprises data from a tumor, cancer tissue or a cancerous cell or pre-cancerous cell, similarity, albeit not necessarily absolute identity, between the test sample and the control sample is indicative of a positive diagnosis (i.e. cancer).

Amplification Assay Formats

In an alternative example, the fragments produced by the restriction enzyme are detected using an amplification system, such as, for example, polymerase chain reaction (PCR), rolling circle amplification (RCA), inverse polymerase chain reaction (iPCR), in situ PCR (Singer-Sam *et al.*, *Nucl. Acids Res.* 18:687, 1990), strand displacement amplification (SDA) or cycling probe technology.

Methods of PCR are known in the art and described, for example, by McPherson *et al.*, PCR: A Practical Approach. (series eds, D. Rickwood and B. D. Hames), IRL Press Limited, Oxford. pp 1-253, 1991 and by Dieffenbach (ed) and Dveksler (ed) (In: PCR Primer: A Laboratory Manual, Cold Spring Harbour Laboratories, NY, 1995. Generally, for PCR two non-complementary nucleic acid primer molecules comprising at least about 18 nucleotides in length, and more preferably at least 20-30 nucleotides in length are hybridized to different strands of a nucleic acid template molecule at their respective annealing sites, and specific nucleic acid molecule copies of the template that intervene the annealing sites are amplified enzymatically. Amplification products may be detected, for example, using electrophoresis and detection with a detectable marker that binds nucleic acids. Alternatively, one or more of the oligonucleotides are labelled with a detectable marker (e.g. a fluorophore) and the amplification product detected using, for example, a lightcycler (Perkin Elmer, Wellesley, Mass., USA, Roche Applied Science, Indianapolis, IN, USA).

Strand displacement amplification (SDA) utilizes oligonucleotide primers, a DNA polymerase and a restriction endonuclease to amplify a target sequence. The oligonucleotides are hybridized to a target nucleic acid and the polymerase is used to produce a copy of the region intervening the primer annealing sites. The duplexes of copied nucleic acid and target

nucleic acid are then nicked with an endonuclease that specifically recognizes a sequence at the beginning of the copied nucleic acid. The DNA polymerase recognizes the nicked DNA and produces another copy of the target region at the same time displacing the previously generated nucleic acid. The advantage of SDA is that it occurs in an isothermal format, thereby facilitating high-throughput automated analysis.

Cycling Probe Technology uses a chimeric synthetic primer that comprises DNA-RNA-DNA that is capable of hybridizing to a target sequence. Upon hybridization to a target sequence the RNA-DNA duplex formed is a target for RNaseH thereby cleaving the primer. The cleaved primer is then detected, for example, using mass spectrometry or electrophoresis.

For primers that flank or are adjacent to a methylation-sensitive endonuclease recognition site, it is preferred that such primers flank only those sites that are hypermethylated in neoplasia to ensure that a diagnostic amplification product is produced. In this regard, an amplification product will only be produced when the restriction site is not cleaved, i.e., when it is methylated. Accordingly, detection of an amplification product indicates that the CpG dinucleotide/s of interest is/are methylated.

As will be known to the skilled artisan, the precise length of the amplified product will vary depending upon the distance between the primers. Clearly this form of analysis may be used to determine the methylation status of a plurality of CpG dinucleotides provided that each dinucleotide is within a methylation sensitive restriction endonuclease site. In these methods, one or more of the primers may be labelled with a detectable marker to facilitate rapid detection of amplified nucleic acid, for example, a fluorescent label (e.g. Cy5 or Cy3) or a radioisotope (e.g. ^{32}P).

The amplified nucleic acids are generally analyzed using, for example, non-denaturing agarose gel electrophoresis, non-denaturing polyacrylamide gel electrophoresis, mass spectrometry, liquid chromatography (e.g. HPLC or dHPLC), or capillary electrophoresis. (e.g. MALDI-TOF). High throughput detection methods, such as, for example, matrix-assisted laser desorption/ionization time of flight (MALDI-TOF), electrospray ionization (ESI), mass spectrometry (including tandem mass spectrometry, e.g. LC MS/MS), biosensor technology, evanescent fiber-optics technology or DNA chip technology (e.g., WO98/49557; WO 96/17958; Fodor *et al.*, *Science* 767-773, 1991; U.S. Pat. No. 5,143,854; and U.S. Pat. No. 5,837,832), are especially preferred for all assay formats described herein. Alternatively, amplification of a nucleic acid may be continuously monitored using a melting curve analysis method as described herein and/or in, for example, U.S. Pat. No. 6,174,670.

(c) Other Assay Formats

In an alternative example, the increased methylation in a sample is determined by performing a process comprising treating chromatin containing the nucleic acid with an amount of DNaseI under conditions sufficient for nucleic acid to be digested and then detecting the fragments produced. This assay format is predicated on the understanding that chromatin containing methylated DNA, e.g., hyper methylated DNA, has a more tightly-closed conformation than non-hyper methylated DNA and, as a consequence, is less susceptible to endonuclease digestion by DNase I.

In accordance with this method , DNA fragments of different lengths are produced by DNase I digestion of methylated compared to non-methylated DNA. Such different DNA fragments are detected, for example, using an assay described earlier. Alternatively, the DNA fragments are detected using PCR-SSCP essentially as described, for example, in Gregory and Feil, *Nucleic Acids Res.*, 27, e32i-e32iv, 1999. In adapting PCR-SSCP to the present invention, amplification primers flanking or comprising one or more CpG dinucleotides in a nucleic acid that are resistant to DNase I digestion in a neoplasia sample but not resistant to DNase I digestion in a healthy/normal control or healthy/normal test sample are used to amplify the DNase I-generated fragments. In this case, the production of a specific nucleic acid fragment using DNase I is diagnostic of neoplasia, because the DNA is not efficiently degraded. In contrast, template DNA from a healthy/normal subject sample is degraded by the action of DNase I and, as a consequence, amplification fails to produce a discrete amplification product. Alternative methods to PCR-SSCP, such as for example, PCR-dHPLC are also known in the art and contemplated by the present invention.

(d) Selective Mutagenesis of Non-Methylated DNA

In an alternative method the increased methylation in a sample is determined using a process comprising treating the nucleic acid with an amount of a compound that selectively mutates a non-methylated cytosine residue within a CpG dinucleotide under conditions sufficient to induce mutagenesis.

Preferred compounds mutate cytosine to uracil or thymidine, such as, for example, a salt of bisulfite, e.g., sodium bisulfite or potassium bisulfite (Frommer *et al.*, 1992, *supra*). Bisulfite treatment of DNA is known to distinguish methylated from non-methylated cytosine

residues, by mutating cytosine residues that are not protected by methylation, including cytosine residues that are not within a CpG dinucleotide or that are positioned within a CpG dinucleotide that is not subject to methylation.

Sequence Based Detection

In one example, the presence of one or more mutated nucleotides or the number of mutated sequences is determined by sequencing mutated DNA. One form of analysis comprises amplifying mutated nucleic acid using an amplification reaction described herein, for example, PCR. The amplified product is then directly sequenced or cloned and the cloned product sequenced. Methods for sequencing DNA are known in the art and include for example, the dideoxy chain termination method or the Maxam-Gilbert method (see Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989) or Zyskind *et al.*, Recombinant DNA Laboratory Manual, (Acad. Press, 1988)).

As the treatment of nucleic acid with a compound, such as, for example, bisulfite results in non-methylated cytosines being mutated to uracil (and hence thymidine after an amplification process), analysis of the sequence determines the presence or absence of a methylated nucleotide. For example, by comparing the sequence obtained using a control sample or a sample that has not been treated with bisulfite, or the known nucleotide sequence of the region of interest with a treated sample facilitates the detection of differences in the nucleotide sequence. Any thymine residue detected at the site of a cytosine in the treated sample compared to a control or untreated sample may be considered to be caused by mutation as a result of bisulfite treatment. Suitable methods for the detection of methylation using sequencing of bisulfite treated nucleic acid are described, for example, in Frommer *et al.*, 1992, *supra* or Clark *et al.*, *Nucl. Acids Res.* 22:2990-2997, 1994.

In another method, the presence of a mutated or non-mutated nucleotide in a bisulfite treated sample is detected using pyrosequencing, such as, for example, as described in Uhlmann *et al.*, Electrophoresis, 23: 4072-4079, 2002. Essentially this method is a form of real-time sequencing that uses a primer that hybridizes to a site adjacent or close to the site of a cytosine that is methylated. Following hybridization of the primer and template in the presence of a DNA polymerase each of four modified deoxynucleotide triphosphates are added separately according to a predetermined dispensation order. Only an added nucleotide that is complementary to the bisulfite treated sample is incorporated and inorganic pyrophosphate (PPi) is liberated. The PPi then drives a reaction resulting in production of detectable levels of

light. Such a method allows determination of the identity of a specific nucleotide adjacent to the site of hybridization of the primer.

Methods of solid phase pyrosequencing are known in the art and reviewed in, for example, Landegren *et al.*, *Genome Res.*, 8(8): 769-776, 1998. Such methods enable the high-throughput detection of methylation of a number of CpG dinucleotides.

A related method for determining the sequence of a bisulfite treated nucleotide is methylation-sensitive single nucleotide primer extension (Me-SnuPE) or SNaPmeth. Suitable methods are described, for example, in Gonzalgo and Jones, 1997, *supra*, or Uhlmann *et al.*, *Electrophoresis*, 23:4072-4079, 2002. An oligonucleotide is used that hybridizes to the region of a nucleic acid adjacent to the site of a cytosine that is methylated. This oligonucleotide is then used in a primer extension protocol with a polymerase and a free nucleotide diphosphate or dideoxynucleotide triphosphate that corresponds to either or any of the possible bases that occur at this site following bisulfite treatment (i.e., thymine or cytosine). Preferably, the nucleotide-diphosphate is labelled with a detectable marker (e.g. a fluorophore). Following primer extension, unbound labelled nucleotide diphosphates are removed, e.g. using size exclusion chromatography or electrophoresis, or hydrolyzed, using for example, alkaline phosphatase, and the incorporation of the labelled nucleotide to the oligonucleotide is detected, indicating the base that is present at the site.

Clearly other high throughput sequencing methods are encompassed by the present invention. Such methods include, for example, solid phase minisequencing (as described, for example, in Southern *et al.*, *Genomics*, 13:1008-1017, 1992), or minisequencing with FRET (as described, for example, in Chen and Kwok, *Nucleic Acids Res.* 25:347-353, 1997).

Restriction Endonuclease-Based Assay Format

In one method, the presence of a non-mutated sequence is detected using combined bisulfite restriction analysis (COBRA) essentially as described in Xiong and Laird, 2001, *supra*. This method exploits the differences in restriction enzyme recognition sites between methylated and unmethylated nucleic acid after treatment with a compound that selectively mutates a non-methylated cytosine residue, e.g., bisulfite.

Following bisulfite treatment a region of interest comprising one or more CpG dinucleotides that are methylated and are included in a restriction endonuclease recognition sequence is amplified using an amplification reaction described herein, e.g., PCR. The amplified product is then contacted with the restriction enzyme that cleaves at the site of the

CpG dinucleotide for a time and under conditions sufficient for cleavage to occur. A restriction site may be selected to indicate the presence or absence of methylation. For example, the restriction endonuclease TaqI cleaves the sequence TCGA, following bisulfite treatment of a non-methylated nucleic acid the sequence will be TTGA and, as a consequence, will not be cleaved. The digested and/or non-digested nucleic acid is then detected using a detection means known in the art, such as, for example, electrophoresis and/or mass spectrometry. The cleavage or non-cleavage of the nucleic acid is indicative of cancer in a subject. Clearly, this method may be employed in either a positive read-out or negative read-out system for the diagnosis of a cancer.

Positive Read-Out Assay Format

In one embodiment, the assay format of the invention comprises a positive read-out system in which DNA from a sample that has been treated, for example, with bisulfite is detected as a positive signal. Preferably, the non-hypermethylated DNA from a healthy or normal control subject is not detected or only weakly detected.

In a preferred embodiment, the increased methylation in a subject sample is determined using a process comprising:

(i) treating the nucleic acid with an amount of a compound that selectively mutates a non-methylated cytosine residue under conditions sufficient to induce mutagenesis thereby producing a mutated nucleic acid;

(ii) hybridizing a nucleic acid to a probe or primer comprising a nucleotide sequence that is complementary to a sequence comprising a methylated cytosine residue under conditions such that selective hybridization to the non-mutated nucleic acid occurs; and

(iii) detecting the selective hybridization.

In this context, the term "selective hybridization" means that hybridization of a probe or primer to the non-mutated nucleic acid occurs at a higher frequency or rate, or has a higher maximum reaction velocity, than hybridization of the same probe or primer to the corresponding mutated sequence. Preferably, the probe or primer does not hybridize to the non-methylated sequence carrying the mutation(s) under the reaction conditions used.

Hybridization-Based Assay Format

In one embodiment, the hybridization is detected using Southern, dot blot, slot blot or other nucleic acid hybridization means (Kawai *et al.*, 1994, *supra*; Gonzalgo *et al.*, 1997,

supra). Subject to appropriate probe selection, such assay formats are generally described herein above and apply mutatis mutandis to the presently described selective mutagenesis approach.

Preferably, a ligase chain reaction format is employed to distinguish between a mutated and non-mutated nucleic acid. Ligase chain reaction (described in EP 320,308 and U.S. Pat. No. 4,883,750) uses at least two oligonucleotide probes that anneal to a target nucleic acid in such a way that they are juxtaposed on the target nucleic acid. In a ligase chain reaction assay, the target nucleic acid is hybridized to a first probe that is complementary to a diagnostic portion of the target sequence (the diagnostic probe) e.g., a nucleic acid comprising one or more methylated CpG dinucleotide(s), and with a second probe that is complementary to a nucleotide sequence contiguous with the diagnostic portion (the contiguous probe), under conditions wherein the diagnostic probe remains bound substantially only to the target nucleic acid. The diagnostic and contiguous probes can be of different lengths and/or have different melting temperatures such that the stringency of the hybridization can be adjusted to permit their selective hybridization to the target, wherein the probe having the higher melting temperature is hybridized at higher stringency and, following washing to remove unbound and/or non-selectively bound probe, the other probe having the lower melting temperature is hybridized at lower stringency. The diagnostic probe and contiguous probe are then covalently ligated such as, for example, using T4 DNA ligase, to thereby produce a larger target probe that is complementary to the target sequence, and the probes that are not ligated are removed by modifying the hybridization stringency. In this respect, probes that have not been ligated will selectively hybridize under lower stringency hybridization conditions than probes that have been ligated. Accordingly, the stringency of the hybridization can be increased to a stringency that is at least as high as the stringency used to hybridize the longer probe, and preferably at a higher stringency due to the increased length contributed by the shorter probe following ligation.

In another example, one or both of the probes is labelled such that the presence or absence of the target sequence can be tested by melting the target-probe duplex, eluting the dissociated probe, and testing for the label(s). Where both probes are labelled, different ligands are used to permit distinction between the ligated and unligated probes, in which case the presence of both labels in the same eluate fraction confirms the ligation event. If the target nucleic acid is bound to a solid matrix e.g., in a Southern hybridization, slot blot, dot blot, or microchip assay format, the presence of both the diagnostic and contiguous probes can be

determined directly.

Methylation specific microarrays (MSO) are also useful for differentiating between a mutated and non-mutated sequence. A suitable method is described, for example, in Adorjan *et al. Nucl. Acids Res.*, 30: e21, 2002. MSO uses nucleic acid that has been treated with a compound that selectively mutates a non-methylated cytosine residue (e.g., bisulfite) as template for an amplification reaction that amplifies both mutant and non-mutated nucleic acid. The amplification is performed with at least one primer that comprises a detectable label, such as, for example, a fluorophore, e.g., Cy3 or Cy5.

To produce a microarray for detection of mutated nucleic acid oligonucleotides are spotted onto, for example, a glass slide, preferably, with a degree of redundancy (for example, as described in Golub *et al.*, *Science*, 286:531-537, 1999). Preferably, for each CpG dinucleotide analyzed two different oligonucleotides are used. Each oligonucleotide comprises a sequence N_2-16CGN_2-16 or N_2-16TGN_2-16 (wherein N is a number of nucleotides adjacent or juxtaposed to the CpG dinucleotide of interest) reflecting the methylated or non-methylated status of the CpG dinucleotides.

The labelled amplification products are then hybridized to the oligonucleotides on the microarray under conditions that enable detection of single nucleotide differences. Following washing to remove unbound amplification product, hybridization is detected using, for example, a microarray scanner. Not only does this method allow for determination of the methylation status of a large number of CpG dinucleotides, it is also semi-quantitative, enabling determination of the degree of methylation at each CpG dinucleotide analyzed. As there may be some degree of heterogeneity of methylation in a single sample, such quantification may assist in the diagnosis of cancer.

Amplification-Based Assay Format

In an alternative example, the hybridization is detected using an amplification system. In methylation-specific PCR formats (MSP; Herman *et al. Proc. Natl. Acad. Sci. USA* 93:9821-9826, 1992), the hybridization is detected using a process comprising amplifying the bisulfite-treated DNA. Accordingly, by using one or more probe or primer that anneals specifically to the unmutated sequence under moderate and/or high stringency conditions an amplification product is only produced using a sample comprising a methylated nucleotide. Alternate assays that provide for selective amplification of either the methylated or the unmethylated component from a mixture of bisulfite-treated DNA are provided by Cottrell *et*

al., Nucl. Acids Res. 32: e10, 2003 (HeavyMethyl PCR), Rand *et al. Nucl. Acids Res.* 33:e127, 2005 (Headloop PCR), Rand *et al. Epigenetics* 1:94-100, 2006 (Bisulfite Differential Denaturation PCR) and PCT/AU07/000389 (End-specific PCR).

Any amplification assay format described herein can be used, such as, for example, polymerase chain reaction (PCR), rolling circle amplification (RCA), inverse polymerase chain reaction (iPCR), in situ PCR (Singer-Sam *et al.*, 1990, *supra*), strand displacement amplification, or cycling probe technology. PCR techniques have been developed for detection of gene mutations (Kuppuswamy *et al.*, *Proc. Natl. Acad. Sci. USA* 88:1143-1147, 1991) and quantitation of allelic-specific expression (Szabo and Mann, *Genes Dev.* 9: 3097-3108, 1995; and Singer-Sam *et al.*, *PCR Methods Appl.* 1: 160-163, 1992). Such techniques use internal primers, which anneal to a PCR-generated template and terminate immediately 5' of the single nucleotide to be assayed. Such as format is readily combined with ligase chain reaction as described herein above. The use of a real-time quantitative assay format is also useful. Subject to the selection of appropriate primers, such assay formats are generally described herein above and apply mutatis mutandis to the presently described selective mutagenesis approach.

Methylation-specific melting-curve analysis (essentially as described in Worm *et al.*, *Clin. Chem.*, 47:1183-1189, 2001) is also contemplated by the present invention. This process exploits the difference in melting temperature in amplification products produced using bisulfite treated methylated or unmethylated nucleic acid. In essence, non-discriminatory amplification of a bisulfite treated sample is performed in the presence of a fluorescent dye that specifically binds to double stranded DNA (e.g., SYBR Green I). By increasing the temperature of the amplification product while monitoring fluorescence the melting properties and thus the sequence of the amplification product is determined. A decrease in the fluorescence reflects melting of at least a domain in the amplification product. The temperature at which the fluorescence decreases is indicative of the nucleotide sequence of the amplified nucleic acid, thereby permitting the nucleotide at the site of one or more CpG dinucleotides to be determined. As the sequence of the nucleic acids amplified using the present invention

The present invention also encompasses the use of real-time quantitative forms of PCR, such as, for example, TaqMan (Holland *et al., Proc. Natl. Acad. Sci. USA*, 88:7276-7280, 1991; Lee *et al., Nucleic Acid Res.* 21:3761-3766, 1993) to perform this embodiment. For example, the MethylLight method of Eads *et al., Nucl. Acids Res.* 28: E32, 2000 uses a modified TaqMan assay to detect methylation of a CpG dinucleotide. Essentially, this method

comprises treating a nucleic acid sample with bisulfite and amplifying nucleic acid comprising one or more CpG dinucleotides that are methylated in a neoplastic cell and not in a control sample using an amplification reaction, e.g., PCR. The amplification reaction is performed in the presence of three oligonucleotides, a forward and reverse primer that flank the region of interest and a probe that hybridizes between the two primers to the site of the one or more methylated CpG dinucleotides. The probe is dual labelled with a 5' fluorescent reporter and a 3' quencher (or vice versa). When the probe is intact, the quencher dye absorbs the fluorescence of the reporter due to their proximity. Following annealing of to the PCR product the probe is cleaved by 5' to 3' exonuclease activity of, for example, Taq DNA polymerase. This cleavage releases the reporter from the quencher thereby resulting in an increased fluorescence signal that can be used to estimate the initial template methylation level. By using a probe or primer that selectively hybridizes to unmutated nucleic acid (i.e. methylated nucleic acid) the level of methylation is determined, e.g., using a standard curve.

Alternatively, rather than using a labelled probe that requires cleavage, a probe, such as, for example, a Molecular Beacon is used (see, for example, Mhlanga and Malmberg, *Methods* 25:463-471, 2001). Molecular beacons are single stranded nucleic acid molecules with a stem-and-loop structure. The loop structure is complementary to the region surrounding the one or more CpG dinucleotides that are methylated in a neoplastic sample and not in a control sample. The stem structure is formed by annealing two "arms" complementary to each other, which are on either side of the probe (loop). A fluorescent moiety is bound to one arm and a quenching moiety that suppresses any detectable fluorescence when the molecular beacon is not bound to a target sequence is bound to the other arm. Upon binding of the loop region to its target nucleic acid the arms are separated and fluorescence is detectable. However, even a single base mismatch significantly alters the level of fluorescence detected in a sample. Accordingly, the presence or absence of a particular base is determined by the level of fluorescence detected. Such an assay facilitates detection of one or more unmutated sites (i.e. methylated nucleotides) in a nucleic acid.

Fluorescently labelled locked nucleic acid (LNA) molecules or fluorescently labelled protein-nucleic acid (PNA) molecules are useful for the detection of nucleotide differences (e.g., as described in Simeonov and Nikiforov, *Nucleic Acids Research*, 30(17):1-5, 2002). LNA and PNA molecules bind, with high affinity, to nucleic acid, in particular, DNA. Fluorophores (in particular, rhodomine or hexachlorofluorescein) conjugated to the LNA or PNA probe fluoresce at a significantly greater level upon hybridization of the probe to target

nucleic acid. However, the level of increase of fluorescence is not enhanced to the same level when even a single nucleotide mismatch occurs. Accordingly, the degree of fluorescence detected in a sample is indicative of the presence of a mismatch between the LNA or PNA probe and the target nucleic acid, such as, in the presence of a mutated cytosine in a methylated CpG dinucleotide. Preferably, fluorescently labelled LNA or PNA technology is used to detect at least a single base change in a nucleic acid that has been previously amplified using, for example, an amplification method known in the art and/or described herein.

As will be apparent to the skilled artisan, LNA or PNA detection technology is amenable to a high-throughput detection of one or more markers by immobilizing an LNA or PNA probe to a solid support, as described in Orum *et al.*, *Clin. Chem.* 45:1898-1905, 1999.

Alternatively, a real-time assay, such as, for example, the so-called HeavyMethyl assay (Cottrell *et al.*, 2003, *supra*) is used to determine the presence or level of methylation of nucleic acid in a test sample. Essentially, this method uses one or more non-extendible nucleic acid (e.g., oligonucleotide) blockers that bind to bisulfite-treated nucleic acid in a methylation specific manner (i.e., the blocker/s bind specifically to unmutated DNA under moderate to high stringency conditions). An amplification reaction is performed using one or more primers that may optionally be methylation specific but that flank the one or more blockers. In the presence of unmethylated nucleic acid (i.e., non-mutated DNA) the blocker/s bind and no PCR product is produced. Using a TaqMan assay essentially as described supra the level of methylation of nucleic acid in a sample is determined.

Other amplification based methods for detecting methylated nucleic acid following treatment with a compound that selectively mutates a non-methylated cytosine residue include, for example, methylation-specific single stranded conformation analysis (MS-SSCA) (Bianco *et al.*, *Hum. Mutat.*, 14:289-293, 1999), methylation-specific denaturing gradient gel electrophoresis (MS-DGGE) (Abrams and Stanton, *Methods Enzymol.*, 212:71-74, 1992) and methylation-specific denaturing high-performance liquid chromatography (MS-DHPLC) (Deng *et al. Chin. J. Cancer Res.*, 12:171-191, 2000). Each of these methods use different techniques for detecting nucleic acid differences in an amplification product based on differences in nucleotide sequence and/or secondary structure. Such methods are clearly contemplated by the present invention.

As with other amplification-based assay formats, the amplification product is analyzed using a range of procedures, including gel electrophoresis, gel filtration, mass spectrometry, and in the case of labelled primers, by identifying the label in the amplification product. In an

alternative embodiment, restriction enzyme digestion of PCR products amplified from bisulfite-converted DNA is performed essentially as described by Sadri and Hornsby, *Nucl. Acids Res.* 24:5058-5059, 1996; and Xiong and Laird, *Nucl. Acids Res.* 25:2532-2534, 1997), to analyze the product formed.

High throughput detection methods, such as, for example, matrix-assisted laser desorption/ionization time of flight (MALDI-TOF), electrospray ionization (ESI), Mass spectrometry (including tandem mass spectrometry, e.g. LC MS/MS), biosensor technology, evanescent fiber-optics technology or DNA chip technology, can also be employed.

As with the other assay formats described herein that utilize hybridization and/or amplification detection systems, combinations of such processes as described herein above are particularly contemplated by the selective mutagenesis-based assay formats of the present invention. In one example, the increased methylation is detected by performing a process comprising:

(i) treating the nucleic acid with an amount of a compound that selectively mutates a non-methylated cytosine residue within a CpG dinucleotide under conditions sufficient to induce mutagenesis thereby producing a mutated nucleic acid;

(ii) hybridizing the nucleic acid to two non-overlapping and non-complementary primers each of which comprises a nucleotide sequence that is complementary to a sequence in the DNA comprising a methylated cytosine residue under conditions such that hybridization to the non-mutated nucleic acid occurs;

(iii) amplifying nucleic acid intervening the hybridized primers thereby producing a DNA fragment consisting of a sequence that comprises a primer sequence;

(iv) hybridizing the amplified DNA fragment to a probe comprising a nucleotide sequence that corresponds or is complementary to a sequence comprising a methylated cytosine residue under conditions such that hybridization to the non-mutated nucleic acid occurs; and detecting the hybridization.

Negative Read-Out Assays

In another example, the assay format comprises a negative read-out system in which reduced methylation of DNA from a healthy/normal control sample is detected as a positive signal and preferably, methylated DNA from a neoplastic sample is not detected or is only weakly detected.

In a preferred embodiment, the reduced methylation is determined using a process

comprising:

(i) treating the nucleic acid with an amount of a compound that selectively mutates a non-methylated cytosine residue within a CpG island under conditions sufficient to induce mutagenesis thereby producing a mutated nucleic acid;

(ii) hybridizing the nucleic acid to a probe or primer comprising a nucleotide sequence that is complementary to a sequence comprising the mutated cytosine residue under conditions such that selective hybridization to the mutated nucleic acid occurs; and

(iii) detecting the selective hybridization.

In this context, the term "selective hybridization" means that hybridization of a probe or primer to the mutated nucleic acid occurs at a higher frequency or rate, or has a higher maximum reaction velocity, than hybridization of the same probe or primer to the corresponding non-mutated sequence. Preferably, the probe or primer does not hybridize to the methylated sequence (or non-mutated sequence) under the reaction conditions used.

Hybridization-Based Assay Format

In one embodiment the hybridization is detected using Southern, dot blot, slot blot or other nucleic acid hybridization means (Kawai *et al.*, 1994, *supra*; Gonzalgo *et al.*, 1997, *supra*). Subject to appropriate probe selection, such assay formats are generally described herein above and apply mutatis mutandis to the presently described selective mutagenesis approach. Preferably, a ligase chain reaction format is employed to distinguish between a non-mutated and mutated nucleic acid. In this respect, the assay requirements and conditions are as described herein above for positive read-out assays and apply mutatis mutandis to the present format. However the selection of probes will differ. For negative read-out assays, one or more probes are selected that selectively hybridize to the mutated sequence rather than the non-mutated sequence.

Preferably, the ligase chain reaction probe(s) have 3'-terminal and/or 5'-terminal sequences that comprise a CpG dinucleotide that is not methylated in a healthy control sample, but is hypermethylated in cancer, such that the diagnostic probe and contiguous probe are capable of being ligated only when the cytosine of the CpG dinucleotide is mutated to thymidine e.g., in the case of a non-methylated cytosine residue.

As will be apparent to the skilled artisan the MSO method described supra is amenable to either or both positive and/or negative readout assays. This is because the assay described detects both mutated and non-mutated sequences thereby facilitating determining the level of

methylation. However, an assay detecting only methylated or non-methylated sequences is contemplated by the invention.

Amplification-Based Assay Format

In an alternative example, the hybridization is detected using an amplification system using any amplification assay format as described herein above for positive read-out assay albeit using primers (and probes where applicable) selectively hybridize to a mutated nucleic acid.

In adapting the HeavyMethyl assay described supra to a negative read-out format, the blockers that bind to bisulfite-treated nucleic acid in a methylation specific manner bind specifically to mutated DNA under moderate to high stringency conditions. An amplification reaction is performed using one or more primers that may optionally be methylation specific (i.e. only bind to mutated nucleic acid) but that flank the one or more blockers. In the presence of methylated nucleic acid (i.e., mutated DNA) the blocker/s bind and no PCR product is produced.

In one example, the reduced methylation in the normal/healthy control subject is detected by performing a process comprising:

(i) treating the nucleic acid with an amount of a compound that selectively mutates non-methylated cytosine residues under conditions sufficient to induce mutagenesis thereby producing a mutated nucleic acid;

(ii) hybridizing the nucleic acid to two non-overlapping and non-complementary primers each of which comprises a nucleotide sequence that is complementary to a sequence in the DNA comprising a mutated cytosine residue under conditions such that hybridization to the mutated nucleic acid occurs;

(iii) amplifying nucleic acid intervening the hybridized primers thereby producing a DNA fragment consisting of a sequence that comprises a primer sequence;

(iv) hybridizing the amplified DNA fragment to a probe comprising a nucleotide sequence that corresponds or is complementary to a sequence comprising a mutated cytosine residue under conditions such that hybridization to the mutated nucleic acid occurs; and

(v) detecting the hybridization.

As will be apparent to the skilled artisan a negative read-out assay preferably includes a suitable control sample to ensure that the negative result is caused by methylated nucleic acid

rather than a reaction failing.

This invention also provides kits for the detection and/or quantification of the diagnostic sequences of the invention, or expression or methylation thereof using the methods described herein.

For kits for detection of methylation, the kits of the invention can comprise at least one polynucleotide that hybridizes to at least one of the diagnostic sequences of the invention and at least one reagent for detection of gene methylation. Reagents for detection of methylation include, e.g., sodium bisulfite, polynucleotides designed to hybridize to sequence that is the product of a biomarker sequence of the invention if the biomarker sequence is not methylated (e.g., containing at least one C→U conversion), and/or a methylation-sensitive or methylation-dependent restriction enzyme. The kits may also include control natural or synthetic DNA sequences representing methylated or unmethylated forms of the sequence. The kits can provide solid supports in the form of an assay apparatus that is adapted to use in the assay. The kits may further comprise detectable labels, optionally linked to a polynucleotide, e.g., a probe, in the kit. Other materials useful in the performance of the assays can also be included in the kits, including test tubes, transfer pipettes, and the like. The kits can also include written instructions for the use of one or more of these reagents in any of the assays described herein.

As detailed hereinbefore, hypermethylation is associated with transcriptional silencing. Accordingly, in addition to the increased level of methylation of these genes providing a basis upon which to screen for the predisposition to or onset of a large intestine neoplasm, the downregulation in the level of expression of these genes is also diagnostically valuable. In accordance with this aspect of the present invention, reference to a gene "expression product" or "expression of a gene" is a reference to either a transcription product (such as primary RNA or mRNA) or a translation product such as protein. In this regard, one can assess changes to the level of expression of a gene either by screening for changes to the level of expression product which is produced (i.e. RNA or protein), changes to the chromatin proteins with which the gene is associated, for example the presence of histone H3 methylated on lysine at amino acid position number 9 or 27 (repressive modifications) or changes to the DNA itself which acts to downregulate expression, such as changes to the methylation of the DNA. These genes and their gene expression products, whether they be RNA transcripts, changes to the DNA which act to downregulate expression or encoded proteins, are collectively referred to as "neoplastic markers".

Accordingly, another aspect of the present invention is directed to a method of

screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the level of expression of a DNA region selected from:

(i) the region, including 2kb upstream of the transcription start site, defined by any two or more of Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; or

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of any two or more of:

(1) BCAT1 (2) IKZF1 (3) IRF4 (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a lower level of expression of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a neoplastic state.

In one embodiment, said method is directed to identifying biological samples in which any one of said DNA regions exhibits a higher level of methylation.

In another embodiment, said method is directed to identifying biological samples in which two or more of said DNA regions exhibits a higher level of methylation.

The method of this aspect of the present invention is predicated on the comparison of the level of the neoplastic markers of a biological sample with the control levels of these markers. The "control level" may be either a "normal level", which is the level of marker expressed by a corresponding large intestine cell or cellular population which is not neoplastic.

As detailed hereinbefore, the normal (or "non-neoplastic") level may be determined using tissues derived from the same individual who is the subject of testing. However, it would be appreciated that this may be quite invasive for the individual concerned and it is therefore likely to be more convenient to analyse the test results relative to a standard result which reflects individual or collective results obtained from individuals other than the patient in issue.

There is more particularly provided a method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the level of expression of one or more genes or transcripts selected from:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393; and
(2) chr7:50344378...50472798;
and optionally one or more of (3) chr6:391739..411443, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream or:
(1) BCAT1; and
(2) IKZF1;
and optionally one or more of (3) IRF4, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a lower level of expression of at least one of the DNA regions of group (i) and/or group (ii) relative to control levels is indicative of a neoplastic large intestine neoplasm or a predisposition to the onset of a neoplastic state.

Preferably, said control level is a non-neoplastic level.

In one embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798; and
(3) chr6:391739..411443; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1 (2) IKZF1 and (3) IRF4.

In another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798; and
(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1 (2) IKZF1 and (4) GRASP.

In still another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by

Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1 (2) IKZF1 and (5) CAHM

In still yet another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; and

(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1 (2) IKZF1 (3) IRF4 and (4) GRASP.

In yet still another embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1 (2) IKZF1 (3) IRF4 and (5) CAHM.

In a further embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of any one or more of:

(1) BCAT1 (2) IKZF1 (4) GRASP and (5) CAHM.

In yet another further embodiment of this aspect, the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; or

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of any one or more of:

(1) BCAT1 (2) IKZF1 (3) IRF4 (4) GRASP and (5) CAHM.

In one embodiment, said method is directed to identifying biological samples in which any one of said DNA regions exhibits a higher level of methylation.

In another embodiment, said method is directed to identifying biological samples in which two or more of said DNA regions exhibits a higher level of methylation.

As detailed hereinbefore, the present invention is designed to screen for a neoplastic cell or cellular population, which is located in the large intestine. Accordingly, reference to "cell or cellular population" should be understood as a reference to an individual cell or a group of cells. Said group of cells may be a diffuse population of cells, a cell suspension, an encapsulated population of cells or a population of cells which take the form of tissue.

Reference to "expression" should be understood as a reference to the transcription and/or translation of a nucleic acid molecule. Reference to "RNA" should be understood to encompass reference to any form of RNA, such as primary RNA or mRNA or non-translated RNA (e.g. miRNAs etc.). Without limiting the present invention in any way, the modulation of gene transcription leading to increased or decreased RNA synthesis may also correlate with the translation of some of these RNA transcripts (such as mRNA) to produce a protein product. Accordingly, the present invention also extends to detection methodology which is directed to screening for modulated levels or patterns of the neoplastic marker protein products as an indicator of the neoplastic state of a cell or cellular population. Although one method is to screen for mRNA transcripts and/or the corresponding protein product, it should be understood that the present invention is not limited in this regard and extends to screening for

any other form of neoplastic marker expression product such as, for example, a primary RNA transcript.

In terms of screening for the downregulation of expression of a marker it would also be well known to the person of skill in the art that changes which are detectable at the DNA level are indicative of changes to gene expression activity and therefore changes to expression product levels. Such changes include but are not limited to, changes to DNA methylation. Accordingly, reference herein to "screening the level of expression" and comparison of these "levels of expression" to control "levels of expression" should be understood as a reference to assessing DNA factors which are related to transcription, such as gene/DNA methylation patterns. These have, in part, been described in detail hereinbefore.

It would also be known to a person skilled in the art that changes in the structure of chromatin are indicative of changes in gene expression. Silencing of gene expression is often associated with modification of chromatin proteins, methylation of lysines at either or both positions 9 and 27 of histone H3 being well studied examples, while active chromatin is marked by acetylation of lysine 9 of histone H3. Thus association of gene sequences with chromatin carrying repressive or active modifications can be used to make an assessment of the expression level of a gene.

Reference to "nucleic acid molecule" should be understood as a reference to both deoxyribonucleic acid molecules and ribonucleic acid molecules and fragments thereof. The present invention therefore extends to both directly screening for mRNA levels in a biological sample or screening for the complementary cDNA which has been reverse-transcribed from an mRNA population of interest. It is well within the skill of the person of skill in the art to design methodology directed to screening for either DNA or RNA. As detailed above, the method of the present invention also extends to screening for the protein product translated from the subject mRNA or the genomic DNA itself.

In one preferred embodiment, the level of gene expression is measured by reference to genes which encode a protein product and, more particularly, said level of expression is measured at the protein level.

In another particularly preferred embodiment, said gene expression is assessed by the association of DNA with chromatin proteins carrying repressive modifications, for example, methylation of lysines 9 or 27 of histone H3.

The present invention should be understood to encompass methods of detection based on identifying both proteins and/or nucleic acid molecules in one or more biological samples.

This may be of particular significance to the extent that some of the neoplastic markers of interest may correspond to genes or gene fragments which do not encode a protein product. Accordingly, to the extent that this occurs it would not be possible to test for a protein and the subject marker would have to be assessed on the basis of transcription expression profiles or changes to genomic DNA.

The term "protein" should be understood to encompass peptides, polypeptides and proteins (including protein fragments). The protein may be glycosylated or unglycosylated and/or may contain a range of other molecules fused, linked, bound or otherwise associated to the protein such as amino acids, lipids, carbohydrates or other peptides, polypeptides or proteins. Reference herein to a "protein" includes a protein comprising a sequence of amino acids as well as a protein associated with other molecules such as amino acids, lipids, carbohydrates or other peptides, polypeptides or proteins.

The proteins encoded by the neoplastic markers of the present invention may be in multimeric form meaning that two or more molecules are associated together. Where the same protein molecules are associated together, the complex is a homomultimer. An example of a homomultimer is a homodimer. Where at least one marker protein is associated with at least one non-marker protein, then the complex is a heteromultimer such as a heterodimer.

Reference to a "fragment" should be understood as a reference to a portion of the subject nucleic acid molecule or protein. This is particularly relevant with respect to screening for modulated RNA levels in stool samples since the subject RNA is likely to have been degraded or otherwise fragmented due to the environment of the gut. One may therefore actually be detecting fragments of the subject RNA molecule, which fragments are identified by virtue of the use of a suitably specific probe.

Although the preferred method is to detect the expression product or DNA changes of the neoplastic markers for the purpose of diagnosing neoplasia development or predisposition thereto, the detection of converse changes in the levels of said markers may be desired under certain circumstances, for example, to monitor the effectiveness of therapeutic or prophylactic treatment directed to modulating a neoplastic condition, such as adenoma or adenocarcinoma development. For example, where reduced expression of the subject markers indicates that an individual has developed a condition characterised by adenoma or adenocarcinoma development, for example, screening for an increase in the levels of these markers subsequently to the onset of a therapeutic regime may be utilised to indicate reversal or other form of improvement of the subject individual's condition. The method of the present

invention is therefore useful as a one off test or as an on-going monitor of those individuals thought to be at risk of neoplasia development or as a monitor of the effectiveness of therapeutic or prophylactic treatment regimes directed to inhibiting or otherwise slowing neoplasia development.

Means of assessing the subject expressed neoplasm markers in a biological sample can be achieved by any suitable method, which would be well known to the person of skill in the art. To this end, it would be appreciated that to the extent that one is examining either a homogeneous cellular population (such as a tumour biopsy or a cellular population which has been enriched from a heterogeneous starting population) or a tissue section, one may utilise a wide range of techniques such as *in situ* hybridisation, assessment of expression profiles by microassays, immunoassays and the like (hereinafter described in more detail) to detect the absence of or downregulation of the level of expression of one or more markers of interest. However, to the extent that one is screening a heterogenous cellular population or a bodily fluid in which heterogeneous populations of cells are found, such as a blood sample, the absence of or reduction in level of expression of a particular marker may be undetectable due to the inherent expression of the marker by non-neoplastic cells which are present in the sample. That is, a decrease in the level of expression of a subgroup of cells may not be detectable. In this situation, a more appropriate mechanism of detecting a reduction in a neoplastic subpopulation of the expression levels of one or more markers of the present invention is via indirect means, such as the detection of epigenetic changes.

Methods of detecting changes to gene expression levels (in addition to the methylation analyses hereinbefore described in detail), particularly where the subject biological sample is not contaminated with high numbers of non-neoplastic cells, include but are not limited to:

(i) *In vivo* detection.

Molecular Imaging may be used following administration of imaging probes or reagents capable of disclosing altered expression of the markers in the intestinal tissues.

Molecular imaging (Moore *et al.*, *BBA*, 1402:239-249, 1988; Weissleder *et al.*, *Nature Medicine* 6:351-355, 2000) is the *in* vivo imaging of molecular expression that correlates with the macro-features currently visualized using "classical" diagnostic imaging techniques such as X-Ray, computed tomography (CT), MRI, Positron Emission Tomography (PET) or endoscopy.

(ii) Detection of downregulation of RNA expression in the cells by Fluorescent *In Situ* Hybridization (FISH), or in extracts from the cells by technologies such as Quantitative

Reverse Transcriptase Polymerase Chain Reaction (QRTPCR) or Flow cytometric qualification of competitive RT-PCR products (Wedemeyer *et al.*, *Clinical Chemistry* 48:9 1398-1405, 2002).

(iii) Assessment of expression profiles of RNA, for example by array technologies (Alon *et al.*, *Proc. Natl. Acad. Sci. USA*: 96:6745-6750, June 1999).

A "microarray" is a linear or multi-dimensional array of preferably discrete regions, each having a defined area, formed on the surface of a solid support. The density of the discrete regions on a microarray is determined by the total numbers of target polynucleotides to be detected on the surface of a single solid phase support. As used herein, a DNA microarray is an array of oligonucleotide probes placed onto a chip or other surfaces used to amplify or clone target polynucleotides. Since the position of each particular group of probes in the array is known, the identities of the target polynucleotides can be determined based on their binding to a particular position in the microarray.

DNA microarray technology make it possible to conduct a large scale assay of a plurality of target nucleic acid molecules on a single solid phase support. U.S. Pat. No. 5,837,832 (Chee *et al.*) and related patent applications describe immobilizing an array of oligonucleotide probes for hybridization and detection of specific nucleic acid sequences in a sample. Target polynucleotides of interest isolated from a tissue of interest are hybridized to the DNA chip and the specific sequences detected based on the target polynucleotides' preference and degree of hybridization at discrete probe locations. One important use of arrays is in the analysis of differential gene expression, where the profile of expression of genes in different cells or tissues, often a tissue of interest and a control tissue, is compared and any differences in gene expression among the respective tissues are identified. Such information is useful for the identification of the types of genes expressed in a particular tissue type and diagnosis of conditions based on the expression profile.

(iv) Measurement of altered neoplastic marker protein levels in cell extracts, for example by immunoassay.

Testing for proteinaceous neoplastic marker expression product in a biological sample can be performed by any one of a number of suitable methods which are well known to those skilled in the art. Examples of suitable methods include, but are not limited to, antibody screening of tissue sections, biopsy specimens or bodily fluid

samples. To the extent that antibody based methods of diagnosis are used, the presence of the marker protein may be determined in a number of ways such as by Western blotting, ELISA or flow cytometry procedures. These, of course, include both single-site and two-site or "sandwich" assays of the non-competitive types, as well as in the traditional competitive binding assays. These assays also include direct binding of a labelled antibody to a target.

(v) Determining altered expression of protein neoplastic markers on the cell surface, for example by immunohistochemistry.

(vi) Determining altered protein expression based on any suitable functional test, enzymatic test or immunological test in addition to those detailed in points (iv) and (v) above.

A person of ordinary skill in the art could determine, as a matter of routine procedure, the appropriateness of applying a given method to a particular type of biological sample.

A related aspect of the present invention provides a molecular array, which array comprises a plurality of:

(i) nucleic acid molecules comprising a nucleotide sequence corresponding to any two or more of the neoplastic marker DNA hereinbefore described or a sequence exhibiting at least 80% identity thereto or a functional derivative, fragment, variant or homologue of said nucleic acid molecule; or

(ii) nucleic acid molecules comprising a nucleotide sequence capable of hybridising to any one or more of the sequences of (i) under medium stringency conditions or a functional derivative, fragment, variant or homologue of said nucleic acid molecules; or

(iii) nucleic acid probes or oligonucleotides comprising a nucleotide sequence capable of hybridising to any two or more of the sequences of (i) under medium stringency conditions or a functional derivative, fragment, variant or homologue of said nucleic acid molecule; or

(iv) probes capable of binding to any two or more of the proteins encoded by the nucleic acid molecules of (i) or a derivative, fragment or, homologue thereof

wherein the level of expression of said marker genes of (i)-(iii) or proteins of (iv) is indicative of the neoplastic state of a cell or cellular subpopulation derived from the large intestine.

Preferably, said percent identity is at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%.

"Hybridization" refers to the process by which a nucleic acid strand joins with a complementary strand through base pairing. Hybridization reactions can be sensitive and

selective so that a particular sequence of interest can be identified even in samples in which it is present at low concentrations. Stringent conditions can be defined by, for example, the concentrations of salt or formamide in the prehybridization and hybridization solutions, or by the hybridization temperature, and are well known in the art. For example, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature, altering the time of hybridization, as described in detail, below. In alternative aspects, nucleic acids of the invention are defined by their ability to hybridize under various stringency conditions (e.g., high, medium, and low), as set forth herein.

Reference herein to a low stringency includes and encompasses from at least about 0 to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridization, and at least about 1 M to at least about 2 M salt for washing conditions. Generally, low stringency is at from about 25-30°C to about 42°C. The temperature may be altered and higher temperatures used to replace formamide and/or to give alternative stringency conditions. Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridization, and at least about 0.5 M to at least about 0.9 M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01 M to at least about 0.15 M salt for hybridization, and at least about 0.01 M to at least about 0.15 M salt for washing conditions. In general, washing is carried out $T_m = 69.3 + 0.41$ (G+C)% (Marmur and Doty, *J. Mol. Biol.* 5:109, 1962). However, the T_m of a duplex DNA decreases by 1°C with every increase of 1% in the number of mismatch base pairs (Bonner and Laskey, *Eur. J. Biochem.* 46:83, 1974). Formamide is optional in these hybridization conditions. Accordingly, particularly preferred levels of stringency are defined as follows: low stringency is 6 x SSC buffer, 0.1% w/v SDS at 25-42°C; a moderate stringency is 2 x SSC buffer, 0.1% w/v SDS at a temperature in the range 20°C to 65°C; high stringency is 0.1 x SSC buffer, 0.1% w/v SDS at a temperature of at least 65°C.

Where nucleic acids of the invention are defined by their ability to hybridize under high stringency, these conditions comprise about 50% formamide at about 37°C to 42°C. In one aspect, nucleic acids of the invention are defined by their ability to hybridize under reduced stringency comprising conditions in about 35% to 25% formamide at about 30°C to

35°C. Alternatively, nucleic acids of the invention are defined by their ability to hybridize under high stringency comprising conditions at 42°C in 50% formamide, 5X SSPE, 0.3% SDS, and a repetitive sequence blocking nucleic acid, such as cot-1 or salmon sperm DNA (e.g., 200 n/ml sheared and denatured salmon sperm DNA). In one aspect, nucleic acids of the invention are defined by their ability to hybridize under reduced stringency conditions comprising 35% formamide at a reduced temperature of 35°C.

Preferably, the subject probes are designed to bind to the nucleic acid or protein to which they are directed with a level of specificity which minimises the incidence of non-specific reactivity. However, it would be appreciated that it may not be possible to eliminate all potential cross-reactivity or non-specific reactivity, this being an inherent limitation of any probe based system.

In terms of the probes which are used to detect the subject proteins, they may take any suitable form including antibodies and aptamers.

A library or array of nucleic acid or protein probes provides rich and highly valuable information. Further, two or more arrays or profiles (information obtained from use of an array) of such sequences are useful tools for comparing a test set of results with a reference, such as another sample or stored calibrator. In using an array, individual probes typically are immobilized at separate locations and allowed to react for binding reactions. Primers associated with assembled sets of markers are useful for either preparing libraries of sequences or directly detecting markers from other biological samples.

A library (or array, when referring to physically separated nucleic acids corresponding to at least some sequences in a library) of gene markers exhibits highly desirable properties. These properties are associated with specific conditions, and may be characterized as regulatory profiles. A profile, as termed here refers to a set of members that provides diagnostic information of the tissue from which the markers were originally derived. A profile in many instances comprises a series of spots on an array made from deposited sequences.

A characteristic patient profile is generally prepared by use of an array. An array profile may be compared with one or more other array profiles or other reference profiles. The comparative results can provide rich information pertaining to disease states, developmental state, receptiveness to therapy and other information about the patient.

Another aspect of the present invention provides a diagnostic kit for assaying biological samples comprising one or more agents for detecting one or more neoplastic markers and reagents useful for facilitating the detection by said agents. Further means may

also be included, for example, to receive a biological sample. The agent may be any suitable detecting molecule.

The present invention is further described by reference to the following non-limiting examples.

5

EXAMPLE 1

CAHM: Colorectal Adenocarcinoma HyperMethylated

Single-plex performance:

- 74 normals: 0.8 pg hypermethylated/mL plasma (95% CI: 0.2; 1.4)
- 73 adenomas: 9.1 pg hypermethylated/mL plasma (95% CI: -8; 27)
- 73 cancers: 1788 pg hypermethylated/mL plasma (95% CI: 231; 3344).

The CAHM assay (threshold cut of 3pg/mL plasma) is 55% sensitive for colorectal cancer with a 93% specificity (Figure 1).

GRASP: General Receptor for Phosphoinositides 1-Associated Scaffold Protein.

Single plex performance:

- 34 normals: 1.3 pg hypermethylated/mL plasma (95% CI: 0.2; 2.5)
- 33 adenomas: 0.1 pg hypermethylated/mL plasma (95% CI: -0.1; 0.4)
- 33 cancers: 670.8 pg hypermethylated/mL plasma (95% CI: -470.7; 1812)

The GRASP assay is 58% sensitive for colorectal cancer with a 100% specificity using a threshold cut of 20pg methylated/mL plasma (Figure 2).

IRF4: Interferon Regulatory Factor 4

Single plex performance:

- 24 normals: 6 pg hypermethylated/mL plasma (95% CI: 2.9; 9.0)
- 23 adenomas: 3.7 pg hypermethylated/mL plasma (95% CI: -0.4; 8.0)
- 33 cancers: 5340 pg hypermethylated/mL plasma (95% CI: -5369; 16049)

The IRF4 assay is 57% sensitive for colorectal cancer with a 96% specificity using a threshold cut of 20pg methylated/mL plasma (Figure 3).

BCAT1: Branched chain amino acid transaminase 1, cytosolic

Single plex performance summary

- 14 normals: 0 pg hypermethylated/mL plasma (95% CI: na)
- 13 adenomas: 0.4 pg hypermethylated/mL plasma (95% CI: -0.4; 1.3)
- 13 cancers: 4088 pg hypermethylated/mL plasma (95% CI: -4459; 10539)

The BCAT1 assay is 62% sensitive for colorectal cancer with a 100% specificity using a threshold cut of 3pg methylated/mL plasma (Figure 4).

IKZF1: IKAROS family zinc finger 1

Single plex performance summary

- 14 normals: 0 pg hypermethylated/mL plasma (95% CI: na)
- 13 adenomas: 0 pg hypermethylated/mL plasma (95% CI: na)
- 13 cancers: 113 pg hypermethylated/mL plasma (95% CI: -43; 270)

The IKZF1 assay is 54% sensitive for colorectal cancer with a 100% specificity using a threshold cut of 3pg methylated/mL plasma (Figure 4).

The methylation levels of the five genes were measured in the same 14 normals, 13 adenomas and 13 cancers. The table below demonstrates the improvement in clinical utility when measuring the methylation in **TWO of the genes**:

(%sens;%spec)	$CAHM_{>6pg/mL}$	$GRASP_{>20pg/mL}$	$IRF4_{>20pg/mL}$	$BCAT_{>3pg/mL}$	$IKZF1_{>3pg/mL}$
$CAHM_{>6pg/mL}$	54% ; 100%	62% ; 100%	69% ; 100%	69% ; 100%	69% ; 100%
$GRASP_{>20pg/mL}$	-	54% ; 100%	69% ; 100%	77% ; 100%	77% ; 100%
$IRF4_{>20pg/mL}$	-	-	54% ; 100%	85% ; 100%	69% ; 100%
$BCAT_{>3pg/mL}$	-	-	-	62% ; 100%	77% ; 100%
$IKZF1_{>3pg/mL}$	-	-	-	-	54% ; 100%

Combining the methylation levels measured in **any three of the five genes** resulted in:

$CAHM_{6pg/mL}$ - $GRASP_{20pg/mL}$ - $IRF4_{20pg/mL}$: 77% sensitivity and 100% specificity

$CAHM_{6pg/mL}$ - $GRASP_{20pg/mL}$ - $BCAT_{3pg/mL}$: 77% sensitivity and 100% specificity

$CAHM_{6pg/mL}$ - $GRASP_{20pg/mL}$ - $IKZF1_{3pg/mL}$: 77% sensitivity and 100% specificity

$CAHM_{6pg/mL}$ - $IKZF1_{3pg/mL}$ - $BCAT_{3pg/mL}$: 77% sensitivity and 100% specificity

$GRASP_{20pg/mL}$ - $IKZF1_{3pg/mL}$ - $BCAT_{3pg/mL}$: 85% sensitivity and 100% specificity

Combining the methylation levels from all five genes results in a 92% sensitivity and 100% specificity:

Targeted amplicon sequences for GRASP, CAHM, IRF4, IKZF1 and BCAT1

The wild-type DNA sequence of the CAHM MSP amplicon is located on Chromosome 6, plus strand; 163,834,393 ⇒ 163,834,455 (Hg19)

GAAGGAAGCA TTTCGAGCAC GACTGACGCT CCCCTTATTA TTTGCTAAGC CGCTGCGCTC GGG

The wild-type DNA sequence of the GRASP MSP amplicon is located on Chromosome 12, plus strand; 52,400,886 ⇒ 52,400,973 (Hg19)

cggaagtcgc gcccgccgct ccggtcccga ccccgggacc ccctgccgca gccgccaccc ctgggccccc agcggacgag ctgtacgc

The wild-type DNA sequence of the IKZF1 MSP amplicon is located on Chromosome 7, plus strand; 50,343,867 ⇒ 50,343,961 (Hg19)

gacgacgcac cctctccgtg tcccgctctg cgcccttctg cgcgccccgc tccctgtacc ggagcagcga tccgggaggc ggccgagagg tg

The wild-type DNA sequence of the BCAT1 MSP amplicon is located on Chromosome 12, minus strand; 25,101,992 ⇒ 25,102,093 (Hg19)

gtcttcctgc tgatgcaatc CGctaggtcg cgagtctccg ccgcgagagg gccggtctgc aatccagccc gccacgtgta ctcgccgccg cctcg

The wild-type DNA sequence of the IRF4 MSP amplicon is located on Chromosome 6, minus strand; 392,036 ⇒ 392,145 (Hg19)

tgggtgcctt ggacggcccc gcctcagcca ctcctggggc cccgacagtc cggttagctc atcccgtcca gcttgtggcg accccgtcgc aggagcgcgg agggcaggcg

PCR protocols used to measure the methylation levels across GRASP, CAHM, IRF4, IKZF1 and BCAT1 in plasma specimens

Real-time PCR protocols GRASP, CAHM, IRF4, BCAT1, IKZF1

	CAHM	GRASP	IRF4	BCAT1	IKZF1
FWD Primer	5'GAAGGAA GTATTTCGA GTACGATTG AC 200nM final	5'CGGAAGT CGCGTTCGT C 200nM final	5'TGGGTGT TTTGGACG GTTTC 400nM final	5'GTTTTTTT GTTGATGTA ATTCGTTAG GTC 200 nM final	5'GACGACG TATTTTTTT CGTGTTTC 200 nM final
REV primer	5'CCCGAACG CAACGACTT AA 200nM final	5'GCGTACA ACTCGTCCG CTAA 200nM final	5'CGCCTAC CCTCCGCG 400nM final	5'CAATACC CGAAACGA CGACG 200nM final	5'GCGCACC TCTCGACCG 200nM final
Probe	na	[HEX] TTCGATTTC GGGATTTTT TGTCGTAGT C [BHQ1] 100nM final	[HEX] TCGTTTAGT TTGTGGCG ATTTCGTCG [BHQ1] 200nM final	[HEX] TTCGTCGCG AGAGGGTC GGTT [BHQ] 200nM final	na
1x PCR	Platinum Taq DNA polymerase (Invitrogen)	Platinum Taq DNA polymerase (Invitrogen)	Platinum Taq DNA polymerase (Invitrogen)	Platinum Taq DNA polymerase (Invitrogen)	GoTaq Hot Start buffer . 1x buffer contains 2mM MgC12 (Promega)
Final $MgCL_2$ conc. [mM]	4.0	4.0	3.0	4.0	+1.0
dNTPs	0.2mM	0.2mM	0.2mM	0.2mM	-
SYBR	1/120,000	na	na	na	1/120,000
LC480 cycling conditions					
1x	95 °C; 2 min	95 °C; 2 min	95 °C; 2 min	95 °C; 2 min	95 °C; 2 min
3x	92 °C; 15 sec 62 °C; 15 sec 72 °C; 20 sec	92 °C; 15 sec 64 °C; 15 sec 72 °C; 20 sec	**50 cycles:** 95 °C; 10 sec 61 °C; 20 sec (w/quant)	**50 cycles:** 95 °C; 15 sec 62 °C; 30 sec 72 °C; 30 sec (w/quant)	**50 cycles:** 95 °C; 15 sec 62°C; 30 sec 72 °C; 30 sec (w/quant)
47x	82 °C; 15 sec 62 °C; 15 sec 72 °C; 20 sec (w/quant)	85 °C; 15 sec 64 °C; 15 sec 72 °C; 20 sec (w/quant)			
melt curve analysis	95 °C; 5 sec 65 °C; 1 min 97 °C continuous 0.11deg/sec	na	na	na	95 °C; 5 sec 65 °C; 1 min 97 °C continuous 0.11deg/sec
Correct Melt Temp (°C)	< 80 °C	na	na	na	82.9-83.2°C

TABLE 1

Gene	Genomic Strand	Current sub-region(s) of interest (genomic sequence)	Coordinates of sub-region(s)	Resulting bisulphite converted sequence (strands no longer complementary)	Oligonucleotide sequences for measurement of methylation levels
BCAT1	top strand	5'-cagtgccCGaggCGgCGgCGagtacaCGtggC GggctggattgcagacCGgccctctCGCGgCGgagactCGC GacctagCGgattgcatcagcaggaagac (SEQ ID NO. 1)	25,101,992 - 25,102,093	5'-tagtgttCGaggCGgCGgCGagtataCGtggCGggttgga ttgtagatCGgttttttCGCGgCGgagattCGCGatttagCGgatt gtattagtaggaagat	
	minus strand	3'-gtcacggGCtccGCcGCcGCtcatgtGCaccGCc cgacctaacgtctgGCcgggagaGCGCcGCctctgaGCGCt ggatcGCctaacgtagtcgtccttctg		3'-gttatggGCttlGCtGCtGCttatgtGCattGCttgatttaatg tttgGCtgggagaGCGCtGCttttgaGCGCtggattGCttaatgt agttgtttttttg	5'-gttttttgttgatgtaattcgttaggtc 5'-caataccegaaacgacgacg 5'-ttcgtcgcgagagggtcggtt 5'-tttttgttgatgtaattcgttaggtc 5'-attacaaaccgaccctctcg
	top strand	5'- agatcccaagggtCGtagcccctggcCGtgtggacCGg gtctgCGgctgcagagCGCGgtccCGgctgcagcaagacctgg ggcagt (SEQ ID NO. 2)	25,101,909 - 25,101,995	This sequence is for measuring CpG methylation levels using methylation sensitive restriction enzymes (e.g. HbaII, HhaI (underlined)	5'-agatcccaagggtcgtagc 5'-actgccccaggtcttgct
	minus strand	3'- tctagggttcccaGCatcggggacccgGCacacctgGCc cagacGCcgacgtctcGCGCcaggGCcgacgtcgttctggacc ccgtca			
IKZF1	top strand	5'-gaCGaCGcaccctctcCGtgtccCGctctgCGcccct ctgCGCGcccCGctcccctgtacCGgagcagCGatcCGggag gCGgcCGagaggtgCGc (SEQ ID NO. 3)	50,343,867-50,343,961	5'-gaCGaCGtatttttttCGtgtttCGttttgCGtttttttgCGCG ttlCGutttttgtatCGgagtagCGattCGggaggCGgtCGagagg tgCGt	5'-gacgacgtatttttttcgtgtttc 5'-gcgcaccictcgaccg 5'-tttgtatcggagtagcgattcgggag
	minus strand	3'-ctGCtGCgtgggagagGCacaggGCgagacGCgg gaagacGCGCgggGCgagggacatgGCctcgtcGCtagGC cctccGCcgGCtctccacGCg		3'-ttGCtGCgtgggagagGCataggGCgagatGCgggaa gatGCGCgggGCgagggatatgGCtttgttGCtagGCtttttGCt gGCttttatGCg	
	top strand	5'- cCGgagttgCGgctgagaCGCGCGcCGCGCG agcCGggggactCGgCGaCGgggCGgggaCGgaCGa CGcaccctctcCGtgtccCGctct (SEQ ID NO. 4)	50,343,804 - 50,343,895	This sequence is for measuring CpG methylation levels using methylation sensitive restriction enzymes (e.g. HbaII, HhaI (underlined)	5'- ggagttgcggctgagac 5'- agagcgggacacggaga

Gene	Genomic Strand	Current sub-region(s) of interest (genomic sequence)	Coordinates of sub-region(s)	Resulting bisulphite converted sequence (strands no longer complementary)		Oligonucleotide sequences for measurement of methylation levels
	minus strand	3'- gGCctcaacGCcgactcGCcGCgGCcGCGC tcgGCccccgaGCccGCtGCcccGCcctGCccGCtGC gtggagagGCacaggGCgaga				
IRF4	top strand	5'- CGcctgcctcCGCGctcctgCGaCGggtCGcc acaagcggaCGgagagctaacCGgactgtCGgggccccag gagtggctgaggCGgggCGtccaaggcaccca (SEQ ID NO. 5)	392,036 - 392,145	5'-CGttggttttCGCGttttgCGaCGgggtCGttataagttg gaCGggatgagttaatCGgattgtCGggttttaggagtggttgagg CGgggtCGtttaaggtattta		5'- gttttgcgacggggtc 5'- taaaaccccgacaatccg
	minus strand	3'-GCggacgggagGCGCgaggacGCtGCcccaGC ggtgttcgacctGCcctactgcgattgGCctgacaGCcccgggctc ctcaccgactccGCcccgGCaggttccgtgggt		3'-GCggatgcggagGCCCgaggatGCtGCttaGCggtgttg attGCttattgattgGCttgataGCtttgggttttattgatttGCttt gGCaggttttgtgggt		5'- tgggtgttttggacggttc 5'- tagttatttttgggttcgatagttc 5'- cgcctaccctccgcg 5'- tcgtttagtttgtggcgatttcgtcg
GRASP	top strand	5'-caggaagctgcagcagaaggaggaggCGgCGgcca cccCGgacccCGcCGccCGgactccCGactCGgaagtCG CGccCGcCGctcCGgtccCGacccCGggaccccctgcCG cagcCGccaccccgggccccagCGgaCGagctgtaCGCG gCGctggaggactatcaccctgcCGagctgtacCGCGCGctC GcCGtgtcCGgggcaccctgcccCGcCGaaaggtgCGtcc ccCGccCGcttcaggatctgctcagccccctctcCGactccctaca ggccctgctgactccCG (SEQ ID NO. 6)	52,400,821 - 52,401,051	5'-taggaagttgtagtagaagaggaggagCGgCGgttatttCGga tttCGtCGttCGattttCGattCGgaagtCGtCGttCGtCGtt CGgtttCGatttCGgatttttgtCGtagtCGttattttgggttttag CGgaCGagttgtaCGCGgCGttggaggattattatttgtCGagtt gtatCGCGCGCGttCGtCGtgttCGgggtatttttgttCGtCGaaa ggtgCGttttttCGttCGttttaggatttgrtagttttttttCGatttttata gggttttgttgatttCG		5'-cggaagtcgcgttcgtc 5'-gcgtacaactcgtccgctaa 5'- ttcgatttcgggatttttgtcgtagtc 5'-cggatttttcgatttcggaagt
	minus strand	3'-gtccttcgacgtcgtcttcctcctccGCcGCcggtgggG CctgggGCgCGgGCctgaggGCtgaCCtcaGCGCGgg GCgGCgagGCcagGCtGggGCcctgggggacgGCgtcg GCggtgggaccccgggggtcGCctCCtcgacatGCGCcGC gacctcctgatagtgggacgCtcgacatgGCGCGCgaGCg GCacagGCccccgtggacgggCgCgGCttccacGCagggg GCggCGgaagtcctagacgagtcgggagagGCtgagggatg tcccggacgactgagGC		3'-gttttttgatgtgttttttttttttCGCtCgtgggGCttggGC gGCggGCttgaggGCtgaCGCttaCGCGGgGCgGCgagG CtaggGCtgggGCtttgggggatgGCgttgGCggtgggaattgg gggttGCtCGCttgatatGCGCtCGCgatttttgatatgtggatgGC ttgatatgGCGCGCGaGCgGCataggGCttttgtgggatgggGC gGCttttttatGCagggggGCggCGgaagttttagatgagtgggga gagGCtgagggatgttttggatgattgagGC		5'- ggtagggtgttttcggatac 5'- aacgaacgaactatacgcgac

Gene	Genomic Strand	Current sub-region(s) of interest (genomic sequence)	Coordinates of sub-region(s)	Resulting bisulphite converted sequence (strands no longer complementary)	Oligonucleotide sequences for measurement of methylation levels
CAHM	top strand	5'-gacagagacagcccgagcaagttgaaggtcCGagagc cccCGgtggagaagCGggcCGgtgctgCGcCGCGCGtgC GttctcactctgaggaaagtgCGtgggagcCGctgactccCGgata gcacaccttcCGagggactcccCGattcctggctggggcct gcCGcctggccccaCGtctgaCGtaCGggCGCGagggcc actgctccctgaacttctgtCGgaaccCGgaCGagtggaggggt CGcagg (SEQ ID NO. 7)	52,401,407 - 52,401,664	5'-gatagagatagttttaggtaagttgaaggttCGagagtttCGgtgg gagaagCGggtCGgtggttgCGtCGCGtgCGttttatttgagg aagtgCGtgggagatCGttgatttCGgatagtatatttttCGagggg attttCGatttttggttgggagttgtCGtttggttttaCGttgaCGta CGgggCGCGaggttattgttttttggattttgtCGgaatCGgaC GtagtggagggggtCGtagg	
	minus strand	3'-ctgtctctgtcgggtccgttcaacttccagGCtctcgggg GCcacctcttcGCccgGCcaccgacGCgGCGCacGCaa gagtgagactccttcacGCaccctcgGCgactgagGCctatcgt gtggaagGCtccctgagggGCtaaggacccgaccccggacg GCggaccgggtGCagactGCatGCccGCGCtcccggtg acgaggacctgaagacaGCcttgGCctGCgtcacccccccaG Cgtcc		3'-tgttttgttgtgggttgttaatttttagGCtttggggGCtatttttt tGCttgGCtattgatGCgGCGCatGCaagagtgagatttttatG Cattttt gGCgattgagGCttatgtgtggggaagGCttttttgaggGG CtaaggatttgatttttggatgGCggattgggtGCagattGCatGCt ttGCGCttttggtgatgaggggatttgaagataGCttgGCtcGCgttat tttttttaGCgttt	5'-cggagttagcggttttttacg 5'-cgataaaaaaaaacgaaccga 5'-agagtgagaacgtacgcggc
	top strand	5'-gacagagacagccccgagcaagttgaaggtcCGagag cccgCGgtggagaagCGggcCGgtggctgCGcCGCGtg CGttctcactctgaggaaagtgCGtgggagcCGctgactccCGga tagcacaccctteCGagggactcccCGattcctggctggggcct tgcCGcctggccccaCGtctgaCGtaCGggCGCGaggcc cactgtccctggacttctgtCGgaacCGgaCGagtggaggg gtCGcagg		This sequence is for measuring CpG methylation levels using methylation sensitive restriction enzymes (e.g. HhaII, HhaI (underlined)	5'-caagttgaaggtccgagagc 5'-cgcacttcctcagagtgaga
	minus strand	3'-ctgtctctgtcgggtccgttcaacttccagGCtctcgggg GCcacctcttcGCccgGCcaccgacGCgGCGCacGCaa gagtgagactccttcacGCaccctcgGCgactgagGCctatcgt gtggaagGCtccctgaggGCtaaggacccgaccccggacg GCggaccgggtGCagactGCatGCccGCGCtcccggtg acgaggacctgaagacaGCcttgGCctGCgtcacccccccaG Cgtcc			
	top strand	5'- atctgtaaaaatgttgacttctgcttttcagactaCGCGcac agcctctttattcctactgCGgcttcattccctccaCGgaaactgaC GccatCGCGaaggaagcatttCGagcaCGactgaCGctcccc ttattatttgctaagcCGctgCGctCGgtctggctaCGatttgctttt cagaataaCGggaaggtgcaacaaga (SEQ ID NO. 8)	163,834,295 163,834,500	5'-atttgtaaaaatgttgatttttgttttttagattaCGCGtatagttttttatt ttttattgCGgtttatttttttaCGgaaarttgaCGttatCGCGaagga agtatttCGagtaCGattgaCGttttttttattattgttaagtCGtgCGt tCGggtttggttaCGatttgttttagaataaCGggaaggtgtaataaga	5'- gaaggaagtatttcgagtacgattgaccc 5'- cccgaacgcaacgacttaa 5'- gcctctaaaaaaacgatcttattacacc

Gene	Genomic Strand	Current sub-region(s) of interest (genomic sequence)	Coordinates of sub-region(s)	Resulting bisulphite converted sequence (strands no longer complementary)	Oligonucleotide sequences for measurement of methylation levels
	minus strand	3'-tagacattttacaactgaagacgaaaagtctgatGCGCg tgtcggagaaataaaggatgacGCgaagtaaggaagtGCcttgtg actGCggtaGCGCttccttcgtaaaGCtcgtGCtgactGCgag gggaataataaacgattcgGCgacGCgaGCcagaccgatGCt aaacgaaagtcttattGCccttccacgttgttct	163,834,621 163,834,906	3'-tagatatttttaataattgaagatgaaaagtttgatGCGCgtgtggag aactaaggaatgatgatGCtgaagtaaggaagtGCttgtgattGCggta GCGCttttttgtaaaGCttgtGCtgattGCgagggaataataaatg atttgGCgatGCgaGCttagattgatGCtaaatgaaagttttattGCtt ttttatgtgtttt	5'-gaaacactaacgccatcg 5'-cgtagttagattcgagcgtag 5'-aggggagcgttagtcgtgttcgaaa
	minus strand	3'-cgtCacgacgaaaggtcggagagtcgtttagtGCtgt gGCttttcttcggtGCcGCcGCtGCccttcccGCaGCGCG CacgaaggagaGCccGCtgttccGCcctcgGCccGCGCgGC cgGCtcccGCggGCcGCgtctcaggGCgtctccGCctGCg GCGCcgtCGCGgaGCtttcggagtttgagaataggaGCcg agaggGCgggtggagGCgggGCgtcggttctggGCGCgG CaccGCccggGCtGCcggttcctttcgggtcggtcgaagGC tgGCac		tgGCatgatgaaaggttggagagttgtttagtGCtgtgGCttttttggt GCtGCtGCtGCttttttttGCaGCGCGCGCatgaaggaGCtGC tgtttGCtttgGCttGCGCgGCtgGCtttGCggGCtGCgttt taggGCgttttGCttGCgGCGCtgtGCGCggaGCttttggag tttgagaaaggaGCtgagaggGCgggtggagGCgggGCgttg gttttggGCGCgGCatGCtggGCtGCtggttttttgggtggttg ggagGC tgGCat	5'-gttttttcggcgataaagc 5'-cgcctctacgaaactctacg 5'-cgtcggtcgaggcgttc

BIBLIOGRAPHY

Abrams and Stanton, *Methods Enzymol.*, 212:71-74, 1992

Adorjan *et al. Nucl. Acids Res.*, 30: e21, 2002

Alon *et al., Proc. Natl. Acad. Sci. USA*: 96:6745-6750, June 1999

Ammerpohl *et al. Biochim Biophys Acta.* 1790:847-62, 2009

Beaucage, *et al. Tetrahedron Letters* 22:1859-1862, 1981

Bianco *et al., Hum. Mutat.*, 14:289-293, 1999

Bonner and Laskey, *Eur. J. Biochem.* 46:83, 1974

Bresslauer *et al., Proc. Natl. Acad. Sci. USA*, 83: 3746-3750, 1986

Caruthers, M. H., *et al., Methods in Enzymology*, Vol. 154, pp. 287-314 (1988)

Chen and Kwok, *Nucleic Acids Res.* 25:347-353, 1997

Clark *et al. Nat Protoc.* 1:2353-64, 2006

Clark *et al., Nucl. Acids Res.* 22:2990-2997, 1994

Cottrell *et al., Nucl. Acids Res.* 32: e10, 2003

DeGraves, *et al., Biotechniques* 34(1):106-10, 112-5 (2003)

Deiman B, *et al., Mol. Biotechnol.* 20(2):163-79 (2002)

Deng *et al. Chin. J. Cancer Res.*, 12:171-191, 2000

Dieffenbach (ed) and Dveksler (ed) (In: PCR Primer: A Laboratory Manual, Cold Spring Harbour Laboratories, NY, 1995

Eads *et al., Nucl. Acids Res.* 28: E32, 2000

Egholm *et al., Am. Chem. Soc.*, 114:1895, 1992

Egholm *et al., Nature*, 365:566, 1993

Fodor *et al., Science* 767-773, 1991

Frommer *et al., Proc. Natl. Acad. Sci. USA* 89:1827-1831 (1992)

Gibson *et al., Genome Research* 6:995-1001 (1996)

Golub *et al., Science*, 286:531-537, 1999

Gonzalgo & Jones, *Nucleic Acids Res.* 25:2529-2531 (1997)

Gonzalgo *et al., Cancer Res.* 57:594-599, 1997

Gregory and Feil, *Nucleic Acids Res.*, 27, e32i-e32iv, 1999

Havelange *et al., Blood* 2011, 118:2827

Herman *et al. Proc. Natl. Acad. Sci. USA* 93:9821-9826 (1996)

Holland *et al., Proc. Natl. Acad. Sci. USA*, 88:7276-7280, 1991

Javierre *et al.*, *Mol. Cancer Res.* 9(8):1139-51, 2011

Kawai *et al.*, *Mol. Cell. Biol.* 14:7421-7427, 1994

Kibriya *et al.*, *BMC* 2011, 4:50

Kristensen and Hansen *Clin Chem.* 55:1471-83, 2009

Kuppuswamy *et al.*, *Proc. Natl. Acad. Sci. USA* 88:1143-1147, 1991

Landegren *et al.*, *Genome Res.*, 8(8): 769-776, 1998

Lee *et al.*, *Nucleic Acid Res.* 21:3761-3766, 1993

Markowitz and Bertagnolli, 2009, *N. Engl. J. Med.* 361(25): 2449–60

Marmur and Doty, *J. Mol. Biol.* 5:109, 1962

Martinez *et al.*, *Am. J. Surg Pathol.* 2012, 36:296

McPherson *et al.*, PCR: A Practical Approach. (series eds, D. Rickwood and B. D. Hames), IRL Press Limited, Oxford. pp 1-253, 1991

Messing, *Methods Enzymol*, 101, 20-78, 1983

Mhlanga and Malmberg, *Methods* 25:463-471, 2001

Moore *et al.*, *BBA*, 1402:239-249, 1988

Narang, *et al. Meth. Enzymol* 68: 90, 1979

Nevrivy *et al. JBC* 2000, 275(22):16827-36

Nielsen *et al. J. Chem. Soc. Perkin Trans.*, 1:3423, 1997

Olek, *et al. Nat. Genet.* 17(3): 275-6 (1997)

Orum *et al.*, *Clin. Chem.* 45:1898-1905, 1999

Orum *et al.*, *Nucl. Acids Res.*, 21:5332, 1993

Oster *et al.*, *Int. J. Cancer* 2011, 129:2855

Pathak *et al. PLoS One* 2011, 6:e22628

Rand *et al. Epigenetics* 1:94-100, 2006

Rand *et al. Nucl. Acids Res.* 33:e127, 2005

Rein, *et al. Nucleic Acids Res.* 26 (10): 2255-64 (1998)

Sadri & Hornsby, *Nucl. Acids Res.* 24:5058-5059 (1996)

Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (2nd Ed., CSHP, New York 1989)

Santa Lucia, *Proc. Natl. Acad. Sci. USA*, 95: 1460-1465, 1995

Shames *et al. Cancer Lett.* 251:187-98, 2007

Simeonov and Nikiforov, *Nucleic Acids Research*, 30(17):1-5, 2002

Singer-Sam *et al.*, *Nucl. Acids Res.* 18:687, 1990

Singer-Sam *et al.*, *PCR Methods Appl.* 1: 160-163, 1992

Singh and Wengel, *Chem. Commun.* 1247, 1998

Slattery *et al.*, *Carcinogenesis* 2011, 32:160

Southern *et al.*, *Genomics*, 13:1008-1017, 1992

Szabo and Mann, *Genes Dev.* 9: 3097-3108, 1995

Toyota *et al.*, *Cancer Res.* 59:2307-12 (1999)

Uhlmann *et al.*, *Electrophoresis*, 23:4072-4079, 2002

Wedemeyer *et al.*, *Clinical Chemistry* 48:9 1398-1405, 2002

Weissleder *et al.*, *Nature Medicine* 6:351-355, 2000

Weitzel JN (December 1999), *Cancer* 86 (11 Suppl): 2483–92

Worm *et al.*, *Clin. Chem.*, 47:1183-1189, 2001

Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534 (1997)

Yamashita *et al.* (*Cancer Sci.* 2010, 101:1708

Zyskind *et al.*, Recombinant DNA Laboratory Manual, (Acad. Press, 1988)

CLAIMS

1. A method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the methylation status of:

(i) the DNA region, including 2kb upstream of the transcription start site of Hg19 coordinates:

(1) chr12:24962958..25102393

in combination with one or more of the following DNA regions, including 2kb upstream of the transciption start site of Hg19 coordinates:

(2) chr7:50344378...50472798;

(3) chr6:391739..411443;

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1

in combination with one or more of the following gene regions, including 2kb upstream of:

(2) IKZF1 (3) IRF4 (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

2. The method according to claim 1 wherein said method comprises assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393; and

(2) chr7:50344378...50472798;

and optionally one or more of (3) chr6:391739..411443, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1; and

(2) IKZF1;

and optionally one or more of (3) IRF4, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

3. The method according to claim 2 wherein the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798; and

(3) chr6:391739..411443; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1 and (3) IRF4.

4. The method according to claim 2 wherein the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798 and

(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1; and (4) GRASP.

5. The method according to claim 2 wherein the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798 and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1 and (5) CAHM.

6. The method according to claim 2 wherein the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; and

(4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1, (3) IRF4 and (4) GRASP.

7. The method according to claim 2 wherein the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(3) chr6:391739..411443; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1, (2) IKZF1, (3) IRF4 and (5) CAHM.

8. The method according to claim 2 wherein the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393

(2) chr7:50344378...50472798

(4) chr12:52400748..52409671 and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1, (4) GRASP and (5) CAHM.

9. The method according to claim 2 wherein the gene marker panel which is screened for is:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393
(2) chr7:50344378...50472798
(3) chr6:391739..411443
(4) chr12:52400748..52409671 and
(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1, (2) IKZF1, (3) IRF4, (4) GRASP and (5) CAHM.

10. The method according to claim 1 wherein said method comprises assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:
(1) chr12:24962958..25102393; and
(5) chr6:163834097..163834982;
and optionally one or more of (2) chr7:50344378...50472798, (3) chr6:391739..411443; and (4) chr12:52400748..52409671; or

(ii) the gene region, including 2kb upstream of:
(1) BCAT1; and
(5) CAHM;
and optionally one or more of (2) IKZF1, (3) IRF4 and (4) GRASP

in a biological sample from said individual wherein a higher level of methylation of said DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

11. The method according to claim 1 wherein said method comprises assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393; and

(3) chr6:391739..411443;

and optionally one or more of (2) chr7:50344378...50472798, (4) chr12:52400748..52409671 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1; and

(3) IRF4;

and optionally one or more of (2) IKZF1, (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

12. The method according to claim 1 wherein said method comprises assessing the methylation status of:

(i) the DNA regions, including 2kb upstream of the transcription start site, defined by Hg19 coordinates:

(1) chr12:24962958..25102393; and

(4) chr12:52400748..52409671;

and optionally one or more of (2) chr7:50344378...50472798, (3) chr6:391739..411443 and (5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1; and

(4) GRASP;

and optionally one or more of (2) IKZF1, (3) IRF4 and (5) CAHM

in a biological sample from said individual wherein a higher level of methylation of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a large intestine neoplastic state.

13. The method according to any one of claims 1 to 12 wherein any one of said DNA regions exhibits a higher level of methylation.

14. The method according to any one of claims 1 to 12 wherein any two or more of said DNA regions exhibit a higher level of methylation.

15. The method according to any one of claims 1 to 14 wherein said neoplastic cell is an adenoma or an adenocarcinoma.

16. The method according to any one of claims 1 to 15 wherein said control level is a non-neoplastic level.

17. The method according to any one of claims 1 to 16 wherein said control level is the level of a previously screened biological sample from said individual.

18. The method according to claim 17 wherein a decrease in the level of methylation relative to said control level is indicative of the clearing of the neoplasm.

19. The method according to any one of claims 1 to 18 wherein said neoplasm is a colorectal neoplasm.

20. The method according to any one of claims 1 to 19 wherein said biological sample is a faecal sample, enema wash, surgical resection, tissue biopsy or blood sample.

21. The method according to claim 20 wherein said blood sample is whole blood, serum or plasma.

22. The method according to any one of claims 1 to 21 wherein said control level is determined from a corresponding biological sample.

23. The method according to any one of claims 1 to 22 wherein said methylation is assessed in one or more chromosomal subregions selected from:

(1) BCAT subregions chr12:25101992-25102093 (SEQ ID NO:1 or corresponding minus strand) and chr12:25101909-25101995 (SEQ ID NO:2 or corresponding minus strand)

(2) IKZF1 subregions: chr7:50343867-50343961 (SEQ ID NO:3 or corresponding minus strand) and chr7:50343804-5033895 (SEQ ID NO:4 or corresponding minus strand)

(3) IRF4 subregions chr6:392036-392145 (SEQ ID NO:5 or corresponding minus strand)

(4) GRASP subregions: chr12:52399672-52399922, chr12:52400821-52401051 (SEQ ID NO:6 or corresponding minus strand), chr12:52401407-52401664 (SEQ ID NO:7 or corresponding minus strand) chr12:52400866-52400973 and Chr12:52401107-52401664.

(5) CAHM subregions: chr6:163834295-163834500 (SEQ ID NO:8 or corresponding minus strand), chr6:163834621-163834906, chr6:163834393-163834455 and chr6:163834393-163834519.

24. The method according to claim 23 wherein said subregion is selected from SEQ ID NO.1, SEQ ID NO.2, SEQ ID NO.3, SEQ ID NO.4, SEQ ID NO.5, SEQ ID NO.6, SEQ ID NO.7 or SEQ ID NO.8 or corresponding minus strands.

25. The method according to claim 23 or 24, said method comprising assessing the methylation of one or more cytosine residues selected from:

(GRASP)

chr12:52399713	chr12:52399731	chr12:52399749	chr12:52399783
chr12:52399796	chr12:52399808	chr12:52399823	chr12:52399835
chr12:52399891			
chr12:52400847	chr12:52400850	chr12:52400859	chr12:52400866
chr12:52400869	chr12:52400873	chr12:52400881	chr12:52400886
chr12:52400893	chr12:52400895	chr12:52400899	chr12:52400902
chr12:52400907	chr12:52400913	chr12:52400919	chr12:52400932
chr12:52400938	chr12:52400958	chr12:52400962	chr12:52400971
chr12:52400973	chr12:52400976	chr12:52400998	chr12:52401008
chr12:52401010	chr12:52401012	chr12:52401016	chr12:52401019
chr12:52401025	chr12:52401041	chr12:52401044	chr12:52401053
chr12:52401060	chr12:52401064	chr12:52401092	chr12:52401118

chr12:52401438	chr12:52401448	chr12:52401460	chr12:52401465
chr12:52401474	chr12:52401477	chr12:52401479	chr12:52401483
chr12:52401504	chr12:52401514	chr12:52401523	chr12:52401540
chr12:52401553	chr12:52401576	chr12:52401588	chr12:52401595
chr12:52401599	chr12:52401604	chr12:52401606	chr12:52401634
chr12:52401640	chr12:52401644	chr12:52401659	

chr12:52401160	chr12:52401165	chr12:52401174	chr12:52401177
chr12:52401179	chr12:52401183	chr12:52401204	chr12:52401215
chr12:52401223	chr12:52401240	chr12:52401253	chr12:52401288
chr12:52401295	chr12:52401299	chr12:52401304	chr12:52401334
chr12:52401340	chr12:52401344	chr12:52401359	

(CAHM)

chr6:163834330	chr6:163834332	chr6:163834357
chr6:163834373	chr6:163834384	chr6:163834390
chr6:163834392	chr6:163834406	chr6:163834412
chr6:163834419	chr6:163834443	chr6:163834448
chr6:163834452	chr6:163834464	chr6:163834483
chr6:163834653	chr6:163834660	chr6:163834672
chr6:163834675	chr6:163834678	chr6:163834681
chr6:163834815	chr6:163834824	chr6:163834835
chr6:163834840	chr6:163834853	chr6:163834855
chr6:163834858	chr6:163834863	chr6:163834869
chr6:163834872		

(IKZF1)

chr7:50343869	chr7:50343872	chr7:50343883
chr7:50343889	chr7:50343890	chr7:50343897
chr7:50343907	chr7:50343909	chr7:50343914
chr7:50343934	chr7:50343939	chr7:50343950
chr7:50343959	chr7:50343805	chr7:50343822
chr7:50343824	chr7:50343826	chr7:50343829
chr7:50343831	chr7:50343833	chr7:50343838
chr7:50343847	chr7:50343850	chr7:50343858
chr7:50343864	chr7:50343869	chr7:50343872
chr7:50343890		

(IRF4)

chr6:392036	chr6:392047	chr6:392049
chr6:392057	chr6:392060	chr6:392066
chr6:392080	chr6:392094	chr6:392102
chr6:392131		

or a corresponding cytosine at position n+1 on the opposite DNA strand.

26. A method of screening for the onset or predisposition to the onset of or monitoring a large intestine neoplasm in an individual, said method comprising assessing the level of expression of:

(i) the DNA region, including 2kb upstream of the transcription start site of Hg19 coordinates:

(1) chr12:24962958..25102393

in combination with one or more of the following DNA regions, including 2kb upstream of the transcription start site of Hg19 coordinates:

(2) chr7:50344378...50472798;

(3) chr6:391739..411443;

(4) chr12:52400748..52409671; and

(5) chr6:163834097..163834982; or

(ii) the gene region, including 2kb upstream of:

(1) BCAT1

in combination with one or more of the following gene regions, including 2kb upstream of:

(2) IKZF1 (3) IRF4 (4) GRASP and (5) CAHM

in a biological sample from said individual wherein a lower level of expression of at least one of the DNA regions of group (i) and/or (ii) relative to control levels is indicative of a large intestine neoplasm or a predisposition to the onset of a neoplastic state.

27. The method according to claim 26 wherein said level of expression is mRNA expression or protein expression.

28. The method according to claim 27 wherein said neoplastic cell is an adenoma or an adenocarcinoma.

29. The method according to any one of claims 26 to 28 wherein said control level is a non-neoplastic level.

30. The method according to any one of claims 26 to 28 wherein said control level is the level

of a previously screened biological sample from said individual.

31. The method according to claim 30 wherein an increase in the level of expression relative to said control level is indicative of the clearing of the neoplasm.

32. The method according to any one of claims 26 to 31 wherein said neoplasm is a colorectal neoplasm.

33. The method according to any one of claims 26 to 32 wherein said biological sample is a faecal sample, enema wash, surgical resection, tissue biopsy or blood sample.

34. The method according to claim 33 wherein said blood sample is whole blood, serum or plasma.

35. The method according to any one of claims 26 to 34 wherein said control level is determined from a corresponding biological sample.

36. The method according to any one of claims 1 to 25 wherein said DNA region is the promoter region.

37. The method according to any one of claims 1 to 36 wherein said individual is a human.

FIGURE 1

>hg19_refGene_NM_181711 range=chr12:52398748-52409671 5'pad=0 3'pad=0 strand=+ repeatMasking=none

GRASP resides on the plus strand, thus the top strand in the below is the representative 5'-3'- sequence of GRASP.

YELLOW: sub-regions of interest towards which three bisulphite conversion and methylation specific assays are given as examples of designs targeting both strands

GREEEN font: sub-region of interest given as an example of a methylation sensitive restriction enzymes (hhaI/HbaII) PCR amplicon

```
   1  CCCACCCCCCACAGATGCTGTCTGTCAGGAGCAGGCAACAGACGGTCCTGGGGTCCATCT
   1  GGGTGGGGGGTGTCTACGACAGACAGTCCTCGTCCGTTGTCTGCCAGGACCCCAGGTAGA

  61  GTGCTAATGCTTCACAATAGAGTCCCATGTTTCCCCACTGACACCCCCTCGGCCCCTCAG
  61  CACGATTACGAAGTGTTATCTCAGGGTACAAAGGGGTGACTGTGGGGGAGCCGGGGAGTC

 121  CTGATGGCCACTTGGTGGTGCCAAGGCGGGAGGTGGGAGGAGGGAGAAGAAGGCACTAAG
 121  GACTACCGGTGAACCACCACGGTTCCGCCCTCCACCCTCCTCCCTCTTCTTCCGTGATTC

 181  AGAGCTACCTCTTAGCTCCTGGCAGTCCTCAATCCACCCCGGCCCCCCACCCAACAAGCA
 181  TCTCGATGGAGAATCGAGGACCGTCAGGAGTTAGGTGGGGCCGGGGGGTGGGTTGTTCGT

 241  TCCTGCCATCTGGACTTGTGCAATCACTGAGGGGCGGAAAAGCACCCTCTTCCACCCACA
 241  AGGACGGTAGACCTGAACACGTTAGTGACTCCCCGCCTTTTCGTGGGAGAAGGTGGGTGT

 301  CCTCTTGTAGGGGATGGGGGCCTAGAGGACTGGGGGTGGGGAGGAGAACACAGAGTCAAG
 301  GGAGAACATCCCCTACCCCCGGATCTCCTGACCCCCACCCCTCCTCTTGTGTCTCAGTTC

 361  GAGACTAGAGAAGAAGACTGAGCCAGGCGCAAGAACTGAGACAGGCAGGAGGCAGAAAGT
 361  CTCTGATCTCTTCTTCTGACTCGGTCCGCGTTCTTGACTCTGTCCGTCCTCCGTCTTTCA

 421  CTTTCCTGGCCTCGCAGGTGGACGTGGCCATTGCCCCTCTGTGCTCCTTATGCCACGTGT
 421  GAAAGGACCGGAGCGTCCACCTGCACCGGTAACGGGGAGACACGAGGAATACGGTGCACA

 481  CTAACACAGCACATGTGCAAGGTACCACTCCCTACCAGGCCAAGAGCCTCTGTAACCCCA
 481  GATTGTGTCGTGTACACGTTCCATGGTGAGGGATGGTCCGGTTCTCGGAGACATTGGGGT

 541  GTGCCCAGCCAGTACCTGGCACACAGCAGGGCTTAAATGTTGGACAACATCACAGAGTGG
 541  CACGGGTCGGTCATGGACCGTGTGTCGTCCCGAATTTACAACCTGTTGTAGTGTCTCACC

 601  TTAGATTGCAGGCTCTGGAACCAGGCTGCCTGGCTTTGAATCTCAGCTCTGCCGCTGAGT
 601  AATCTAACGTCCGAGACCTTGGTCCGACGGACCGAAACTTAGAGTCGAGACGGCGACTCA

 661  GACTTAGGGCAAATTACTTATCTTCTCTGGGCCTCAGTTTCCTCATCTGTAAGGGAGGAT
 661  CTGAATCCCGTTTAATGAATAGAAGAGACCCGGAGTCAAAGGAGTAGACATTCCCTCCTA

 721  AATGGTGCTTATTTCGTAGGGTTGTTATGAAGACCAAGTGAGTTAATGCATGTATGTAAA
 721  TTACCACGAATAAAGCATCCCAACAATACTTCTGGTTCACTCAATTACGTACATACATTT

 781  AGGACACACAGAACAAACAGTGTGCATAGCACATGCTAAGTGCTCAATAAATGTTAACTG
 781  TCCTGTGTGTCTTGTTTGTCACACGTATCGTGTACGATTCACGAGTTATTTACAATTGAC

 841  TGAAGGCATAACAGATTTGGGAAGACTGGGGGAGGAGGGAGGGGACAGACAAGGGCACCG
 841  ACTTCCGTATTGTCTAAACCCTTCTGACCCCCTCCTCCCTCCCCTGTCTGTTCCCGTGGC

 901  CATTGCCGCCTCTGCTTCTTGGCCGATGGGTGTTGGGACATGGAGTCCCCAGCCTGTGCT
 901  GTAACGGCGGAGACGAAGAACCGGCTACCCACAACCCTGTACCTCAGGGGTCGGACACGA

 961  GCTGACGCCCTGTGGTGGCTAGCCGCTCTTTAGACACCAGACGCTGCCTTCCACTTCCTC
 961  CGACTGCGGGACACCACCGATCGGCGAGAAATCTGTGGTCTGCGACGGAAGGTGAAGGAG

1021  CTCTCACTTCTTATTCGGGCACCACTGACGACAGAGACTTCGGCTGGGCCATTCCCGCCT
1021  GAGAGTGAAGAATAAGCCCGTGGTGACTGCTGTCTCTGAAGCCGACCCGGTAAGGGCGGA
```

FIGURE 1 (cont'd)

```
1081  CTCCTACCGCCTGGCAGCCATGGTAACAGCCTGGGGGATGGTCCCTAAGATGGGGTCAGA
1081  GAGGATGGCGGACCGTCGGTACCATTGTCGGACCCCCTACCAGGGATTCTACCCCAGTCT

1141  GAACGGCTGAGCATGGCCATCCCCCCAGCCCTGAACTCCCACTCAGCCTGGGCTCATCCC
1141  CTTGCCGACTCGTACCGGTAGGGGGGTCGGGACTTGAGGGTGAGTCGGACCCGAGTAGGG

1201  AACCTCCTGAGTCAAGTCCACAGTAAGCGTCTGTCTCCTGGGAAGGGGGATCTGACCCCG
1201  TTGGAGGACTCAGTTCAGGTGTCATTCGCAGACAGAGGACCCTTCCCCCTAGACTGGGGC

1261  GCATCAGGAGCTGAGGAAAGGTGTTGCCAGCAGGGATCACTCAAGCCCCTTGGAGGGGAC
1261  CGTAGTCCTCGACTCCTTTCCACAACGGTCGTCCCTAGTGAGTTCGGGGAACCTCCCCTG

1321  AGGCCAGAACTAAGCCTACTGCTTTCAGGGTTTCCACCTCTTCCCTTCTTCCTAGACGTG
1321  TCCGGTCTTGATTCGGATGACGAAAGTCCCAAAGGTGGAGAAGGGAAGAAGGATCTGCAC

1381  CGAACAGGGGGCAGCCTCTTTCAGCTCTTCTGGGGCAGCTGCGTCAAGGGAAAATCCTGG
1381  GCTTGTCCCCCGTCGGAGAAAGTCGAGAAGACCCCGTCGACGCAGTTCCCTTTTAGGACC

1441  GCCCTCTCCCTCCCTGGGGGCTTCTGCGCAGTGAGTTCAGGGAGTCTCCTCCCTCCTCCA
1441  CGGGAGAGGGAGGGACCCCCGAAGACGCGTCACTCAAGTCCCTCAGAGGAGGGAGGAGGT

1501  TCGGGCCTCCACCCCTACTCCTGCGCAGCCCCGCTGCCGCTCTCCCTGCCCTGGAGTCTC
1501  AGCCCGGAGGTGGGGATGAGGACGCGTCGGGGCGACGGCGAGAGGGACGGGACCTCAGAG

1561  CTGCAGCCCCTTGGATCTCCGTGACTCTCCCTACCTCCCCGACTCCCCAGGCTTCTTACA
1561  GACGTCGGGGAACCTAGAGGCACTGAGAGGGATGGAGGGGCTGAGGGGTCCGAAGAATGT

1621  GTGACCTCTTACCGTGCCCCACTCCATGAATCGCCAGAGCTATTCGTCCCTAAATTTCAA
1621  CACTGGAGAATGGCACGGGGTGAGGTACTTAGCGGTCTCGATAAGCAGGGATTTAAAGTT

1681  ACCTTGCGCAATGTCCCTTCACAGACCCCTCCAGGTATCACGCAGCCCCGAGCCCCGAGC
1681  TGGAACGCGTTACAGGGAAGTGTCTGGGGAGGTCCATAGTGCGTCGGGGCTCGGGGCTCG

1741  CCCGCCCCGGGGGCCTCATCCCGCCCCTTCGCGTCCGCGGCTCGTTTTCCCCCACTGAGC
1741  GGGCGGGGCCCCCGGAGTAGGGCGGGGAAGCGCAGGCGCCGAGCAAAAGGGGGTGACTCG

1801  GCCCAGCTCCCGCAGTTTCCCCGGCCGTCGAGCGCCGTGGGCGGGGCTCCAGGGCGGCGG
1801  CGGGTCGAGGGCGTCAAAGGGGCCGGCAGCTCGCGGCACCCGCCCCGAGGTCCCGCCGCC

1861  CGCCTCGCGGGGAGGGTCCTCCGTGCTGGGGGCGAGGCCACCCGAGGCAGCTCCCCGCCC
1861  GCGGAGCGCCCCTCCCAGGAGGCACGACCCCCGCTCCGGTGGGCTCCGTCGAGGGGCGGG

1921  GCCCCCAACCCCGCCCCGCTCTCGGAGCCTATAAAGGGAGGCGACCCGCGGCCCGCCCGG
1921  CGGGGGTTGGGGCGGGGCGAGAGCCTCGGATATTTCCCTCCGCTGGGCGCCGGGCGGGCC

1981  CTGGCATCCCCCAGCCGCCGCCAGCCCCGCCGAGGGGAGCCAGCGCCGTCTCTGAGGGGC
1981  GACCGTAGGGGGTCGGCGGCGGTCGGGGCGGCTCCCCTCGGTCGCGGCAGAGACTCCCCG

2041  GTCCGGCGCCGGAGCCATGACCCTCCGCCGACTCAGGAAGCTGCAGCAGAAGGAGGAGGC
2041  CAGGCCGCGGCCTCGGTACTGGGAGGCGGCTGAGTCCTTCGACGTCGTCTTCCTCCTCCG

2101  GGCGGCCACCCCGGACCCCGCCGCCCGGACTCCCGACTCGGAAGTCGCGCCCGCCGCTCC
2101  CCGCCGGTGGGGCCTGGGGCGGCGGGCCTGAGGGCTGAGCCTTCAGCGCGGGCGGCGAGG

2161  GGTCCCGACCCCGGGACCCCCTGCCGCAGCCGCCACCCCTGGGCCCCCAGCGGACGAGCT
2161  CCAGGGCTGGGGCCCTGGGGGACGGCGTCGGCGGTGGGGACCCGGGGGTCGCCTGCTCGA

2221  GTACGCGGCGCTGGAGGACTATCACCCTGCCGAGCTGTACCGCGCGCTCGCCGTGTCCGG
2221  CATGCGCCGCGACCTCCTGATAGTGGGACGGCTCGACATGGCGCGCGAGCGGCACAGGCC

2281  GGGCACCCTGCCCCGCCGAAAGGTGCGTCCCCCGCCCGCCTTCAGGATCTGCTCAGCCCC
2281  CCCGTGGGACGGGGCGGCTTTCCACGCAGGGGGCGGGCGGAAGTCCTAGACGAGTCGGGG

2341  TCTCCGACTCCCTACAGGGCCTGCTGACTCCGCAGTGCCCTCTCCTCGGCGTCCGCGGAG
2341  AGAGGCTGAGGGATGTCCCGGACGACTGAGGCGTCACGGGAGAGGAGCCGCAGGCGCCTC
```

FIGURE 1 (cont'd)

```
2401  TCCCCCACCTTCTTCCCCGGCCCGCTGGGTGCCTCGACTCCCCGCGTTCCCCGCTGCTGC
2401  AGGGGGTGGAAGAAGGGGCCGGGCGACCCACGGAGCTGAGGGGCGCAAGGGGCGACGACG

2461  GAAGGCCGTGGCCCTCGCCTGCACACCGCGCCCAGGCTCGGTGGCTCTTAACTCCGCGCC
2461  CTTCCGGCACCGGGAGCGGACGTGTGGCGCGGGTCCGAGCCACCGAGAATTGAGGCGCGG

2521  CCATGCACGCCCCCTCTCTCCCTCCTTGACTCCTCCCAGCACCCCCCTTCTCCTACCCGC
2521  GGTACGTGCGGGGGAGAGAGGGAGGAACTGAGGAGGGTCGTGGGGGGAAGAGGATGGGCG

2581  TCCATCTGGCTTTCTGCCCCCCATGCCCCGCCTCCCCGTGGCCAGGTGTCCTGGGTCCCC
2581  AGGTAGACCGAAAGACGGGGGGTACGGGGCGGAGGGGCACCGGTCCACAGGACCCAGGGG

2641  AGGAGCCCCTCGCCCGAGGGACAGAGACAGCCCCAGGCAAGTTGAAGGTCCGAGAGCCCC
2641  TCCTCGGGGAGCGGGCTCCCTGTCTCTGTCGGGGTCCGTTCAACTTCCAGGCTCTCGGGG

2701  CGGTGGGAGAAGCGGGCCGGTGGCTGCGCCGCGTGCGTTCTCACTCTGAGGAAGTGCGTG
2701  GCCACCCTCTTCGCCCGGCCACCGACGCGGCGCACGCAAGAGTGAGACTCCTTCACGCAC

2761  GGGAGCCGCTGACTCCGGATAGCACACCCTTCCGAGGGGACTCCCCGATTCCTGGGCTGG
2761  CCCTCGGCGACTGAGGCCTATCGTGTGGGAAGGCTCCCCTGAGGGGCTAAGGACCCGACC

2821  GGGCCTGCCGCCTGGCCCCACGTCTGACGTACGGGGCGCGAGGGCCACTGCTCCCTGGAC
2821  CCCGGACGGCGGACCGGGGTGCAGACTGCATGCCCCGCGCTCCCGGTGACGAGGGACCTG

2881  TTCTGTCGGAACCGGACGCAGTGGGAGGGGTCGCAGGGCGCCCGCGGGGCAGGAAGGATG
2881  AAGACAGCCTTGGCCTGCGTCACCCTCCCCAGCGTCCCGCGGGCGCCCCGTCCTTCCTAC

2941  CGGGCCGCGCCCACCTCTGAGTCCCCTCTGCCAGCCTCTTCCTCTGGCCCCAGGAGACCT
2941  GCCCGGCGCGGGTGGAGACTCAGGGGAGACGGTCGGAGAAGGAGACCGGGGTCCTCTGGA

3001  GAGGCTCAGAACCTACACAACACCAGGTTAAGAAGAGGGGCCTGGTGGCCTTTCCTCACC
3001  CTCCGAGTCTTGGATGTGTTGTGGTCCAATTCTTCTCCCCGGACCACCGGAAAGGAGTGG

3061  CAGCCGCCCTCCTTCGCCCCGGCCCCCAGCTAGCCCCCACACAATGAACAGCTTGTTGAG
3061  GTCGGCGGGAGGAAGCGGGGCCGGGGGTCGATCGGGGGTGTGTTACTTGTCGAACAACTC

3121  AATTTGCATTTTATGAAAATCATGTTGAAAGACAAAGGGGTCTCTCTGTGCTGCCCAGTC
3121  TTAAACGTAAAATACTTTTAGTACAACTTTCTGTTTCCCCAGAGAGACACGACGGGTCAG

3181  CTTCCTCCCTGGCCGTTTGGGAACTGTCCCCACCCCTGAGGCCAATCTGGTTCTGAACCT
3181  GAAGGAGGGACCGGCAAACCCTTGACAGGGGTGGGGACTCCGGTTAGACCAAGACTTGGA

3241  TCTCTTCCTTGCCTTGGGCAGCTTTGGGGGAGGGTTAGCAAAGGCACAGAACAGAAAGGC
3241  AGAGAAGGAACGGAACCCGTCGAAACCCCCTCCCAATCGTTTCCGTGTCTTGTCTTTCCG

3301  CCTGGGCTGTGCAGGCTCCAAAGAAAAGGGCTGCTCTGGGACTGGACCTCCTCCCAGGAC
3301  GGACCCGACACGTCCGAGGTTTCTTTTCCCGACGAGACCCTGACCTGGAGGAGGGTCCTG

3361  CAAAAAGTTAGGGAGGGTGAGAGACTACTTTAGTTTATCAAGGACCCTGAAGAGACAGGA
3361  GTTTTTCAATCCCTCCCACTCTCTGATGAAATCAAATAGTTCCTGGGACTTCTCTGTCCT

3421  ACCTTCATCTCTCATCCTTCCTCCACACCCCCCACCACCACCCCTAAAGAACTCCCAGTC
3421  TGGAAGTAGAGAGTAGGAAGGAGGTGTGGGGGGTGGTGGTGGGGATTTCTTGAGGGTCAG

3481  TCGGTCCTTTAGTGAGACTTGCTGACAAGTTTGACATCTAAGATGTTTTGTCCCAGAAAG
3481  AGCCAGGAAATCACTCTGAACGACTGTTCAAACTGTAGATTCTACAAAACAGGGTCTTTC

3541  CACAAAATATATGGCAATGGAGAGAGAGACCCAGGTATAGCTGGGCACAGCTGGTCACCT
3541  GTGTTTTATATACCGTTACCTCTCTCTCTGGGTCCATATCGACCCGTGTCGACCAGTGGA

3601  GCAGCTGGGATCCACAACTGGTCCTTGAAACGGCCTGTACCTTAGGAGACCTGGCACCTC
3601  CGTCGACCCTAGGTGTTGACCAGGAACTTTGCCGGACATGGAATCCTCTGGACCGTGGAG

3661  TGACCCCACATTCTGGCTGGGATTGCCAGCCCTTCGGGACAGGGTCCCTGACCCCAGCCC
3661  ACTGGGGTGTAAGACCGACCCTAACGGTCGGGAAGCCCTGTCCCAGGGACTGGGGTCGGG
```

FIGURE 1 (cont'd)

```
3721  TCCCAAGCCACTGTCTGTAGCTGAGAAATTAGATGGAGAGACCCAACTGGCAGAAGGTCC
3721  AGGGTTCGGTGACAGACATCGACTCTTTAATCTACCTCTCTGGGTTGACCGTCTTCCAGG

3781  TCGGAGCACCTTGATAGGTGAGCCCAGGGGACACCTATCCTCTGAATCCCACTGGGGAAG
3781  AGCCTCGTGGAACTATCCACTCGGGTCCCCTGTGGATAGGAGACTTAGGGTGACCCCTTC

3841  AGCCCCTGCCTCAGCTTTGGGAGTCTGGATGGCCTGAGCCTCTACAGACATGGGCCCTAG
3841  TCGGGGACGGAGTCGAAACCCTCAGACCTACCGGACTCGGAGATGTCTGTACCCGGGATC

3901  GGGTGGAGACCATTTTAGAATAATGATCTCCCCACCTGCTGCCAGTGTGGAAACCAGCAA
3901  CCCACCTCTGGTAAAATCTTATTACTAGAGGGGTGGACGACGGTCACACCTTTGGTCGTT

3961  GGGCTTAGAGGTTCATGGATCTGGAACCCAGGAGGATGGTTGTGTCCCTGCAGTCCCAGG
3961  CCCGAATCTCCAAGTACCTAGACCTTGGGTCCTCCTACCAACACAGGGACGTCAGGGTCC

4021  TATGAAGGTAAGGCCTGTAGAGAAAGTTGAGGAGTGGTGCCAGTAGTGGCACTTGGCAAA
4021  ATACTTCCATTCCGGACATCTCTTTCAACTCCTCACCACGGTCATCACCGTGAACCGTTT

4081  TAGTTCCACCAGCACCAAAACAGGTGTGGATGTGGAGCTGGGAAGGGGCAGACAGGAAGT
4081  ATCAAGGTGGTCGTGGTTTTGTCCACACCTACACCTCGACCCTTCCCCGTCTGTCCTTCA

4141  GGGTCGCTGTGTCCAGGGTAACCTCTCAGTGTCTGCTCAAGGACAGTCCCGCCTTACCTG
4141  CCCAGCGACACAGGTCCCATTGGAGAGTCACAGACGAGTTCCTGTCAGGGCGGAATGGAC

4201  CTCCTTCTGACCATCTTACTGCCCAGGGCTCAGGATTCCGCTGGAAGAATCTCAGCCAGA
4201  GAGGAAGACTGGTAGAATGACGGGTCCCGAGTCCTAAGGCGACCTTCTTAGAGTCGGTCT

4261  GTCCTGAACAGCAGCGGTGAGTCACCAACACCCAGCCCCTGCCATGGTCCAAAGGGGTGA
4261  CAGGACTTGTCGTCGCCACTCAGTGGTTGTGGGTCGGGGACGGTACCAGGTTTCCCCACT

4321  GGTGCTGGGTTGGGGGGTGCCAGGACAGCATCTCAGCCTAGCGAGGGTAGTCATTCTCCT
4321  CCACGACCCAACCCCCCACGGTCCTGTCGTAGAGTCGGATCGCTCCCATCAGTAAGAGGA

4381  AGCCTTTAAACATGGCTCCCCCGCTGGGAGTGAGGGCAGGTGGGGGTGAGGGCTGTGTGA
4381  TCGGAAATTTGTACCGAGGGGGCGACCCTCACTCCCGTCCACCCCCACTCCCGACACACT

4441  AGGGGGTGGCAGAGAAAGGAGGTAGGGACTTTTGAGTCATTAGGTGCTCTTCAACACACC
4441  TCCCCCACCGTCTCTTTCCTCCATCCCTGAAAACTCAGTAATCCACGAGAAGTTGTGTGG

4501  CCCAGTCCCTGCCAGTTACCCCTCCCTGGGGAGATCTGGATGAACGGATGTGGGAGTTGG
4501  GGGTCAGGGACGGTCAATGGGGAGGGACCCCTCTAGACCTACTTGCCTACACCCTCAACC

4561  GAGGGGGTGTAGGCAAAGTCTATGGGGAACGTCCTAACCCAGGCCTCCTGGGCTCCCCTG
4561  CTCCCCCACATCCGTTTCAGATACCCCTTGCAGGATTGGGTCCGGAGGACCCGAGGGGAC

4621  AGCCCTCCTAGCCTTGGCAAACCCCACGGGCCCAGACCTTAGCCAGGCTGTGTCCAGCTG
4621  TCGGGAGGATCGGAACCGTTTGGGGTGCCCGGGTCTGGAATCGGTCCGACACAGGTCGAC

4681  CTTGGGGCTGGCTGCCCCTGCCTCCAGAATGTCAGTCCTCTTCTGGTCCCAGCCTGTGGC
4681  GAACCCCGACCGACGGGGACGGAGGTCTTACAGTCAGGAGAAGACCAGGGTCGGACACCG

4741  AGCCCTCTTAGGAGGGATCTGAGCGTGGGCAGAAGTAGGTGCTGACTCCAGGCCCTGGGA
4741  TCGGGAGAATCCTCCCTAGACTCGCACCCGTCTTCATCCACGACTGAGGTCCGGGACCCT

4801  CCCCACAGTTTCCCTTCTCTACTCATGTCCCATCCCTGATTTGGGCTTGACTTTCTCTAA
4801  GGGGTGTCAAAGGGAAGAGATGAGTACAGGGTAGGGACTAAACCCGAACTGAAAGAGATT

4861  AATGGATCAGCAAACTGACCTGCTAGCCAGATTGGCCTGTTTGGCCCTTGAGCTAATAAT
4861  TTACCTAGTCGTTTGACTGGACGATCGGTCTAACCGGACAAACCGGGAACTCGATTATTA

4921  TTTTAAAGAGTTGTAAAAATAAAAACAAATAACAATATATGACAGAGATCATAAGGGGCC
4921  AAAATTTCTCAACATTTTTATTTTTGTTTATTGTTATATACTGTCTCTAGTATTCCCCGG

4981  TGCAAAGAAAGCCTAAAGTATTTACCATTTGGCCCTTTGCAGAAAAAGTTTGCTGACCCT
4981  ACGTTTCTTTCGGATTTCATAAATGGTAAACCGGGAAACGTCTTTTTCAAACGACTGGGA
```

FIGURE 1 (cont'd)

```
5041  TCCAAAAAACCCTAGAACATTGGGGTAGACGCGGAACTCCCAGCCACCCTGCTCCCTGCC
5041  AGGTTTTTTGGGATCTTGTAACCCCATCTGCGCCTTGAGGGTCGGTGGGACGAGGGACGG

5101  CCAATAGTACAGAGAGGGGAAAGGCCTGCTGTTCAGAGTGAGAGGGGCTCGAGCTGGAGT
5101  GGTTATCATGTCTCTCCCCTTTCCGGACGACAAGTCTCACTCTCCCCGAGCTCGACCTCA

5161  GGGGAGGTGCTGCTCAGCATTTGGGGGAATCTCTGGGTCAGGCTGGAAGCGGAGAAGCCT
5161  CCCCTCCACGACGAGTCGTAAACCCCCTTAGAGACCCAGTCCGACCTTCGCCTCTTCGGA

5221  GGTCTCCAGGGCCTTTCCGAGGCTGAGCTACTTGAAGGGGCCTTTGGAGGCTGCTTTAAC
5221  CCAGAGGTCCCGGAAAGGCTCCGACTCGATGAACTTCCCCGGAAACCTCCGACGAAATTG

5281  TGTGCCTCTGGCTGGGCGGGAAGGAAGGGGGTGGGAGTGGGCAGAGGAAACTCTGGGCTC
5281  ACACGGAGACCGACCCGCCCTTCCTTCCCCCACCCTCACCCGTCTCCTTTGAGACCCGAG

5341  CCCCAGCAGGCGAGGATCTGGAGCAGCTCCAGACCATTGTGTCCAGCGGGCAGGCCTTCA
5341  GGGGTCGTCCGCTCCTAGACCTCGTCGAGGTCTGGTAACACAGGTCGCCCGTCCGGAAGT

5401  GTGCGGGAAGGGGCGGCCTCGAGGCTCCCCTCCCCCAGCCCCCACATCTGGTGGGCTGGC
5401  CACGCCCTTCCCCGCCGGAGCTCCGAGGGGAGGGGGTCGGGGGTGTAGACCACCCGACCG

5461  CCCAGCATAGCTGGGAGGAGCAGCTGTGGTCTTGCTGAGCCTCGTGACTGGCCTCTGGGG
5461  GGGTCGTATCGACCCTCCTCGTCGACACCAGAACGACTCGGAGCACTGACCGGAGACCCC

5521  GTGGGGCCAGTCCTCTCTCCAAAGCTGTGGAACAGAGGAGGCTCCAGGCTGTGGCTGAAT
5521  CACCCCGGTCAGGAGAGAGGTTTCGACACCTTGTCTCCTCCGAGGTCCGACACCGACTTA

5581  TTCGGGCCTTAGCTAGTCAGTAAGTGGTACTGTAGGGCTCACTGGAAAGCTGGGATGGGG
5581  AAGCCCGGAATCGATCAGTCATTCACCATGACATCCCGAGTGACCTTTCGACCCTACCCC

5641  CTAAAAAACTCAGCCGGCTCATCCTTCAGGGGACCGGCCCTTCTCTGTGCTTCCCTCCCA
5641  GATTTTTTGAGTCGGCCGAGTAGGAAGTCCCCTGGCCGGGAAGAGACACGAAGGGAGGGT

5701  CCACTAGAGGCTGGATTGGTTCTCTGTGGCTCCATGAGGCTGATTTCAATTGGGTATGGG
5701  GGTGATCTCCGACCTAACCAAGAGACACCGAGGTACTCCGACTAAAGTTAACCCATACCC

5761  AAAGACATTCCAATAATCTGTGATGAGACTGAGCTGTCACTGGAGACAGAGTCCTGCTGG
5761  TTTCTGTAAGGTTATTAGACACTACTCTGACTCGACAGTGACCTCTGTCTCAGGACGACC

5821  AATGTTCTAGACCAGTTGCCAATATCCAGGGAGGCCCAAGCTATGGGACTTCTACACCTT
5821  TTACAAGATCTGGTCAACGGTTATAGGTCCCTCCGGGTTCGATACCCTGAAGATGTGGAA

5881  TTAATCCTAATTGTCTTGACCCCTGTGTCTCTTGCAGGAAAGTGCTGACGTTGGAGAAGG
5881  AATTAGGATTAACAGAACTGGGGACACAGAGAACGTCCTTTCACGACTGCAACCTCTTCC

5941  AGGATAACCAGACCTTCGGCTTTGAGATCCAGGTGGGAGAAGCTGCACACAGGGGTCAGG
5941  TCCTATTGGTCTGGAAGCCGAAACTCTAGGTCCACCCTCTTCGACGTGTGTCCCCAGTCC

6001  GGGGTTGGATGACCAGCCTGAGGGATGAACGGACTTGTCCCAACCCTGGGCTGAGGGTCC
6001  CCCCAACCTACTGGTCGGACTCCCTACTTGCCTGAACAGGGTTGGGACCCGACTCCCAGG

6061  CCTTGTCCATTACAGACTTATGGCCTTCACCACCGGGAGGAGCAGCGTGTGGAAATGGTG
6061  GGAACAGGTAATGTCTGAATACCGGAAGTGGTGGCCCTCCTCGTCGCACACCTTTACCAC

6121  ACCTTTGTCTGCCGAGTTCATGAGTCTAGCCCTGCCCAGCTGGCTGGGCTCACACCAGGT
6121  TGGAAACAGACGGCTCAAGTACTCAGATCGGGACGGGTCGACCGACCCGAGTGTGGTCCA

6181  GGGGCCTGAGCCCAGGACACCCAGGTCTGGGAAGGGGATATGACCTTACTCCCAAGCAAA
6181  CCCCGGACTCGGGTCCTGTGGGTCCAGACCCTTCCCCTATACTGGAATGAGGGTTCGTTT

6241  GGGGGTGAGCAATCTCTCCTGAAATCAATTCCTCTTCCTTTTCCTTCTTTGAGAAGGCCA
6241  CCCCCACTCGTTAGAGAGGACTTTAGTTAAGGAGAAGGAAAAGGAAGAAACTCTTCCGGT

6301  GAAAGAAGAATGGATAGAATCTGGGCTTTGGATCTAGGCAAATTTGCCATGTACTGTGTG
6301  CTTTCTTCTTACCTATCTTAGACCCGAAACCTAGATCCGTTTAAACGGTACATGACACAC
```

FIGURE 1 (cont'd)

```
6361  ATCTTGCACAGCCCCATCTACAAAATGAGGGTAATAATGCATCCAAACATCACAGTGCGG
6361  TAGAACGTGTCGGGGTAGATGTTTTACTCCCATTATTACGTAGGTTTGTAGTGTCACGCC

6421  CAAGGGGATTCCCTGGGCACACTGCCAGGGCCTAATTAATGGTGGATGATGCTGCTGCTG
6421  GTTCCCCTAAGGGACCCGTGTGACGGTCCCGGATTAATTACCACCTACTACGACGACGAC

6481  CTCTGATTCCTCCCAGCAACCCTGGCAGTCAGCATGGGCAGGAGCCAGGGAAGAAGCAAC
6481  GAGACTAAGGAGGGTCGTTGGGACCGTCAGTCGTACCCGTCCTCGGTCCCTTCTTCGTTG

6541  ATTCCATTAAGTCTGTTTGATATTGGGGATCAGGCCAATCCTGCCCCAAAATGGGCCCAG
6541  TAAGGTAATTCAGACAAACTATAACCCCTAGTCCGGTTAGGACGGGGTTTTACCCGGGTC

6601  TGCTGAGGAACCGATGTTACTCCCTTTTAAAAAATTAGAAACTTTTTTTTTTTTAAGAG
6601  ACGACTCCTTGGCTACAATGAGGGAAAATTTTTTAATCTTTGAAAAAAAAAAAAATTCTC

6661  ACAGGGCCTCAGTCTGTCACCCAGGCTAGAGTGCAGTGGCGTAATCATAACTCACTATAA
6661  TGTCCCGGAGTCAGACAGTGGGTCCGATCTCACGTCACCGCATTAGTATTGAGTGATATT

6721  CCTTGAACTCCTGGGCTCAAGCGATCCTCCTCTTGCCTCTGCCTCCCAAAGCAATATGTT
6721  GGAACTTGAGGACCCGAGTTCGCTAGGAGGAGAACGGAGACGGAGGGTTTCGTTATACAA

6781  ACTTCCTCTAACAAGGAAATTATGCTTCAGCAGGAGATCCCTGGATTGAGCAGATCTAGA
6781  TGAAGGAGATTGTTCCTTTAATACGAAGTCGTCCTCTAGGGACCTAACTCGTCTAGATCT

6841  GTCCCCAGGTTCCAGGAAGGGCAGCCTGAAACTGTATGAATCAATCCCCCCTCCACCATC
6841  CAGGGGTCCAAGGTCCTTCCCGTCGGACTTTGACATACTTAGTTAGGGGGGAGGTGGTAG

6901  TTTGCCCCTAAGCCCCTACCTCCTTCCCACTTACCAGCAGCCCGTGCTAGCTATCTTAGT
6901  AAACGGGGATTCGGGGATGGAGGAAGGGTGAATGGTCGTCGGGCACGATCGATAGAATCA

6961  CCATTTTCTGTTACCATAACAGAATACCTGATACTGGGTAATTATAAAGAAAAGAGGTTT
6961  GGTAAAAGACAATGGTATTGTCTTATGGACTATGACCCATTAATATTTCTTTTCTCCAAA

7021  ATTTAGCTCATGGTTCTGGAGGCTGGGAAGTTCAAGACTGGGTGGCCGCATCAGGTGAGG
7021  TAAATCGAGTACCAAGACCTCCGACCCTTCAAGTTCTGACCCACCGGCGTAGTCCACTCC

7081  GCCTCATGCTGCGTTCTGACATGATGGATGGCATCTCATGGCAGGAACGCCTGCAAGAGT
7081  CGGAGTACGACGCAAGACTGTACTACCTACCGTAGAGTACCGTCCTTGCGGACGTTCTCA

7141  GGTGAGTAGGCACACGCAAAAGAGACAAAACATGGGTGATCTTGCTTTATAGCAACCCAC
7141  CCACTCATCCGTGTGCGTTTTCTCTGTTTTGTACCCACTAGAACGAAATATCGTTGGGTG

7201  TCGCCAGGTAACTAAACCAGTCCTACCAGAGCAAGAACTCACTCCCCAAAAACAGCATGG
7201  AGCGGTCCATTGATTTGGTCAGGATGGTCTCGTTCTTGAGTGAGGGGTTTTTGTCGTACC

7261  ATCCCTTCACTGGGCAGATCCTTCATGGCCCAAATGCCACTTTTTTTGAGATGGAGTTTC
7261  TAGGGAAGTGACCCGTCTAGGAAGTACCGGGTTTACGGTGAAAAAAACTCTACCTCAAAG

7321  GCTCTTGTTGCCCAGGCTGGAGTGCAATGGCATGATCTCAGCTCACTGCAGCCTCTGCCT
7321  CGAGAACAACGGGTCCGACCTCACGTTACCGTACTAGAGTCGAGTGACGTCGGAGACGGA

7381  CCCAGATTCAAGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACTGG
7381  GGGTCTAAGTTCGTTAAGAGGACGGAGTCGGAGGGCTCATCGACCCTAATGTCCGTGACC

7441  CCACCAAGCCCAGCTCATTTTTGTATTTTTAGTAGAGATGGGGTTTTGCCTTGTTGGCCA
7441  GGTGGTTCGGGTCGAGTAAAAACATAAAAATCATCTCTACCCCAAAACGGAACAACCGGT

7501  GGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGG
7501  CCGACCAGAGCTTGAGGACTGGAGTCCACTAGGTGGGCGGAGCCGGAGGGTTTCACGACC

7561  GATTACAGGTGTGAGCCACGGCGCTCGGCCCCAAATGCCTCTTAAAGGTCCTACCATCTC
7561  CTAATGTCCACACTCGGTGCCGCGAGCCGGGGTTTACGGAGAATTTCCAGGATGGTAGAG

7621  TCAGCACTGTTACGTTGGGGATGAAGCCTCAACATGAGTTTTGGGGACAAACAATATTTA
7621  AGTCGTGACAATGCAACCCCTACTTCGGAGTTGTACTCAAAACCCCTGTTTGTTATAAAT
```

FIGURE 1 (cont'd)

```
7681  AATGGTAGCAGTAGCCAAGTGGTATACTACAACTTCTCAAAGTAGTTTCAAATCCACTCT
7681  TTACCATCGTCATCGGTTCACCATATGATGTTGAAGAGTTTCATCAAAGTTTAGGTGAGA

7741  CCCTCCCCATCATCCCTGAAGCATCCACAGCAGGGATCCATACCCCTTTTTACAGATAGG
7741  GGGAGGGGTAGTAGGGACTTCGTAGGTGTCGTCCCTAGGTATGGGGAAAAATGTCTATCC

7801  AATGGGGCCCTGACATGGGGCACACTGTGCTTGAGGTCAGGAAGCTCTCCAGTGGTGCAG
7801  TTACCCCGGGACTGTACCCCGTGTGACACGAACTCCAGTCCTTCGAGAGGTCACCACGTC

7861  GGTAGCAGAACTATCCCTCTGGGGGCCAGTCATTCCCTGTGCTTCTCTCCCACCACCAGA
7861  CCATCGTCTTGATAGGGAGACCCCCGGTCAGTAAGGGACACGAAGAGAGGGTGGTGGTCT

7921  GGCTGGACTGATTATCTGTGGATCCATGAGGCCATGCCATGCCTGGGGCCAGGGTCGGTT
7921  CCGACCTGACTAATAGACACCTAGGTACTCCGGTACGGTACGGACCCCGGTCCCAGCCAA

7981  CCCAGCTGGGTGCTGCTCACCTGCTCTTCCCTGAATTGACTGGGTCTTATGGCCAGCGTG
7981  GGGTCGACCCACGACGAGTGGACGAGAAGGGACTTAACTGACCCAGAATACCGGTCGCAC

8041  GACTGGACAGAAATAACCAAATCTGAGGGTAGCCCAGGGTCCTGGGTGGGCTTAGCTTTG
8041  CTGACCTGTCTTTATTGGTTTAGACTCCCATCGGGTCCCAGGACCCACCCGAATCGAAAC

8101  GGACAGAACTTCTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGG
8101  CCTGTCTTGAAGAAAAAAAAAAAAAAAAAAAACTCTGCCTCAGAGCGAGACAGCGGGTCC

8161  CTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCAT
8161  GACCTCACGTCACCGCGCTAGAGCCGAGTGACGTTCGAGGCGGAGGGCCCAAGTGCGGTA

8221  TCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCTGCTACCACGCCCGGCTA
8221  AGAGGACGGAGTCGGAGGGCTCATCGACCCTGATGTCCGCGGACGATGGTGCGGGCCGAT

8281  ATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCT
8281  TAAAAAACATAAAAATCATCTCTGCCCCAAAGTGGCACAATCGGTCCTACCAGAGCTAGA

8341  CCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC
8341  GGACTGGAGCACTAGGCGGGCGGAGCCGGAGGGTTTCACGACCCTAATGTCCGCACTCGG

8401  ACCGCGCCCGGCCGGACAGAACTTCTTGAGGGCAGAGTGAGGGTGTTGGGTGTGTGAGGT
8401  TGGCGCGGGCCGGCCTGTCTTGAAGAACTCCCGTCTCACTCCCACAACCCACACACTCCA

8461  CCCCACCCTTTGGCTGGGGTGCTGGGCGTGGACAACCTGGTAGTCACTACAGGCCCACAA
8461  GGGGTGGGAAACCGACCCCACGACCCGCACCTGTTGGACCATCAGTGATGTCCGGGTGTT

8521  GAATGGCTGTGGTTCTGTGTGTTTGGATCGAGTATGAGAGATGTCAGAGGAGATCTACAG
8521  CTTACCGACACCAAGACACACAAACCTAGCTCATACTCTCTACAGTCTCCTCTAGATGTC

8581  AGAACCATGAGGAGTTGGAGGAGGGAGTTCAGGAGTATTCCCTGGAGTTACTACCCACCC
8581  TCTTGGTACTCCTCAACCTCCTCCCTCAAGTCCTCATAAGGGACCTCAATGATGGGTGGG

8641  TTCTCCCCTTTCTGGCTAAGGTAGGAGGCTGGTATTCTAGCATGCCCCATGGAGCAAATC
8641  AAGAGGGGAAAGACCGATTCCATCCTCCGACCATAAGATCGTACGGGGTACCTCGTTTAG

8701  TAACCCCCTTGTCTGCCTGGCAGGGGACACCATCGCCAGCGTCAATGGCCTGAATGTGGA
8701  ATTGGGGGAACAGACGGACCGTCCCCTGTGGTAGCGGTCGCAGTTACCGGACTTACACCT

8761  AGGCATCCGGCATCGAGAGATTGTGGACATCATTAAGGCGTCAGGCAATGTTCTCAGGTA
8761  TCCGTAGGCCGTAGCTCTCTAACACCTGTAGTAATTCCGCAGTCCGTTACAAGAGTCCAT

8821  TGTCTGGGAGCCGAGGTGCCTGAATTCCTGAGCTCAGCCTCTTGGTATTTCCTCAGCCTG
8821  ACAGACCCTCGGCTCCACGGACTTAAGGACTCGAGTCGGAGAACCATAAAGGAGTCGGAC

8881  TGGCTCACTCAGGCTTTGATTCTCCAACCTCAGACTGGAAACTCTATATGGGACATCAAT
8881  ACCGAGTGAGTCCGAAACTAAGAGGTTGGAGTCTGACCTTTGAGATATACCCTGTAGTTA

8941  TCGGAAGGCAGAACTGGAGGCTCGTCTGCAGTACCTGAAGGTAGGGGAACCTAGATAACG
8941  AGCCTTCCGTCTTGACCTCCGAGCAGACGTCATGGACTTCCATCCCCTTGGATCTATTGC
```

FIGURE 1 (cont'd)

```
9001  TCCAGCCTCCACCCTCCTCTTCCCAAGCCTCTGCCCTGTGGGGGGTCACTTACAGCTGAA
9001  AGGTCGGAGGTGGGAGGAGAAGGGTTCGGAGACGGGACACCCCCCAGTGAATGTCGACTT

9061  TCTCCTGTAACTGAAGATTTCTTTTGTCTACTCCCTGATCCCTCGTCTGTACCCACGTCC
9061  AGAGGACATTGACTTCTAAAGAAAACAGATGAGGGACTAGGGAGCAGACATGGGTGCAGG

9121  TGCTAGTTTCCTGCTAGCTTGTGACAGTGGGGTGGGGACTGTCTCTTGCAGCAAACCCTG
9121  ACGATCAAAGGACGATCGAACACTGTCACCCCACCCCTGACAGAGAACGTCGTTTGGGAC

9181  TATGAGAAGTGGGGAGAGTACAGGTCCCTAATGGTGCAGGAGCAGCGGCTGGTGCATGGT
9181  ATACTCTTCACCCCTCTCATGTCCAGGGATTACCACGTCCTCGTCGCCGACCACGTACCA

9241  GAGTAGATCCCGGGGTGTGAGGGGCCACTTGTTCTGCTACAGACACCCCATCTGCGCTTC
9241  CTCATCTAGGGCCCCACACTCCCCGGTGAACAAGACGATGTCTGTGGGGTAGACGCGAAG

9301  CCCCTCAGAACTGGCGGGTTCTAGTAAAGAACGGTTTACTAGTAAACCTCCACAGTAAAG
9301  GGGGAGTCTTGACCGCCCAAGATCATTTCTTGCCAAATGATCATTTGGAGGTGTCATTTC

9361  CAAGGTTTGTTGAGCTACTACTACTTTGGGTACTGCCGCAGCCACATTTCTGGTTATTCT
9361  GTTCCAAACAACTCGATGATGATGAAACCCATGACGGCGTCGGTGTAAAGACCAATAAGA

9421  CACACTAGCCCTCTGCAGGTAAGTAACAGATAGTGTCATCTTCACTTTACAGAGGGGAAC
9421  GTGTGATCGGGAGACGTCCATTCATTGTCTATCACAGTAGAAGTGAAATGTCTCCCCTTG

9481  ACCAGGGCTCAAAGAGTTTGTGTCAATTATTTCCGAGGCCCGTTTGTGCTAAGCGGCCCC
9481  TGGTCCCGAGTTTCTCAAACACAGTTAATAAAGGCTCCGGGCAAACACGATTCGCCGGGG

9541  CATAAGGATCTGAGTGATAATTCCTGATTTGCGGGTGAGGGAAGTAGGGTCTCAGGTTTG
9541  GTATTCCTAGACTCACTATTAAGGACTAAACGCCCACTCCCTTCATCCCAGAGTCCAAAC

9601  TGCCTGGCCAGGGCCACAGAGATGGTCAGCAGGACGGAGCGGGGACGCCCCGCCCATGCC
9601  ACGGACCGGTCCCGGTGTCTCTACCAGTCGTCCTGCCTCGCCCCTGCGGGGCGGGTACGG

9661  CTCCGACCCTTTGCAGAGGCCACCACGGTCCAGGCCTGACCCGCCCCCTACCTCTCCCGT
9661  GAGGCTGGGAAACGTCTCCGGTGGTGCCAGGTCCGGACTGGGCGGGGGATGGAGAGGGCA

9721  CTCTGCGCAGGCCTGGTGGTGAAGGACCCCAGCATCTACGACACGCTGGAGTCGGTGCGC
9721  GAGACGCGTCCGGACCACCACTTCCTGGGGTCGTAGATGCTGTGCGACCTCAGCCACGCG

9781  TCCTGCCTCTACGGCGCGGGCCTGCTCCCGGGCTCGCTGCCCTTCGGGCCTCTGCTCGCC
9781  AGGACGGAGATGCCGCGCCCGGACGAGGGCCCGAGCGACGGGAAGCCCGGAGACGAGCGG

9841  GTGCCCGGGCGTCCCCGCGGAGGCGCCCGACGGGCCAGGGGCGACGCCGACGACGCCGTC
9841  CACGGGCCCGCAGGGGCGCCTCCGCGGGCTGCCCGGTCCCCGCTGCGGCTGCTGCGGCAG

9901  TACCACACGTGCTTCTTCGGGGACTCCGAGCCGCCGGCGCTGCCGCCCCCGCCGCCCCCG
9901  ATGGTGTGCACGAAGAAGCCCCTGAGGCTCGGCGGCCGCGACGGCGGGGGCGGCGGGGGC

9961  GCCCGCGCCTTCGGCCCGGGCCCCGCCGAGACCCCTGCCGTGGGGCCGGGCCCTGGGCCG
9961  CGGGCGCGGAAGCCGGGCCCGGGGCGGCTCTGGGGACGGCACCCCGGCCCGGGACCCGGC

10021  CGGGCCGCGCTGAGCCGCAGCGCCAGTGTGCGGTGCGCGGGCCCTGGCGGGGGCGGAGGC
10021  GCCCGGCGCGACTCGGCGTCGCGGTCACACGCCACGCGCCCGGGACCGCCCCCGCCTCCG

10081  GGGGGCGCGCCGGGCGCGCTCTGGACTGAGGCTCGCGAGCAGGCCCTATGCGGCCCCGGC
10081  CCCCCGCGCGGCCCGCGCGAGACCTGACTCCGAGCGCTCGTCCGGGATACGCCGGGGCCG

10141  CTGCGCAAAACCAAGTACCGCAGCTTCCGCCGGCGGCTGCTCAAGTTCATCCCCGGACTC
10141  GACGCGTTTTGGTTCATGGCGTCGAAGGCGGCCGCCGACGAGTTCAAGTAGGGGCCTGAG

10201  AACCGCTCCCTGGAGGAGGAGGAGAGCCAGCTGTAGGGGCGGGGGCGGGCAGGGAGGTAT
10201  TTGGCGAGGGACCTCCTCCTCCTCTCGGTCGACATCCCCGCCCCCGCCCGTCCCTCCATA

10261  TTATTTATTTATTCGCAACAGCCAGCGCTAAAAGAGGGGGAGGCCGAGCCAAGAGGACCC
10261  AATAAATAAATAAGCGTTGTCGGTCGCGATTTTCTCCCCCTCCGGCTCGGTTCTCCTGGG
```

FIGURE 1 (cont'd)

```
10321  CAGGAGCCCAGAGCAGCGGGAGAGGGTCCTTCCTAGCCTCGGCCCGCCGGGTCGGTTCCT
10321  GTCCTCGGGTCTCGTCGCCCTCTCCCAGGAAGGATCGGAGCCGGGCGGCCCAGCCAAGGA

10381  GGCTGGTGTCTGCTGAGGGAGTGGGGGGCCCAGCCCCTTCTCTTCTCCCCCGCCAAACCA
10381  CCGACCACAGACGACTCCCTCACCCCCCGGGTCGGGGAAGAGAAGAGGGGGCGGTTTGGT

10441  CAGTGGGAGCTGGGGCAGGGGGAGAGCCAGGCAATCGGGGGCCAAAGATGGGGGTGCTCG
10441  GTCACCCTCGACCCCGTCCCCCTCTCGGTCCGTTAGCCCCCGGTTTCTACCCCCACGAGC

10501  CCTACAGTCTGCATCTGTAGTGCCTTGTGGGGTATCCAGGAACACCCTCCCAGCAGGGGA
10501  GGATGTCAGACGTAGACATCACGGAACACCCCATAGGTCCTTGTGGGAGGGTCGTCCCCT

10561  TGGGAACCCTGTCCCATGAAGCCCTCTCCTCAGCTTTACTTGCTCCCCCGCCCTTAGCCT
10561  ACCCTTGGGACAGGGTACTTCGGGAGAGGAGTCGAAATGAACGAGGGGGCGGGAATCGGA

10621  TGGGGAGAAATGGCCCGTGGTGGGCTGACCCCCCACCCTCCACACACACAGTTCCATGAC
10621  ACCCCTCTTTACCGGGCACCACCCGACTGGGGGGTGGGAGGTGTGTGTGTCAAGGTACTG

10681  CCAGCGGGCCCCCAGGGGCATCAGGTGCTGGTCCTCCTCCCTCCTGGCCTCGACCCCTAA
10681  GGTCGCCCCGGGGGTCCCCGTAGTCCACGACCAGGAGGAGGGAGGACCGGAGCTGGGGATT

10741  GGGCTTCGCCCCTCCCAGGGGCCTGTAACTAAGTCGGGTCCTGCCAGGCAGGGGGCCTGT
10741  CCCGAAGCGGGGAGGGTCCCCGGACATTGATTCAGCCCAGGACGGTCCGTCCCCCGGACA

10801  GTTCTGTGCCCCTTGGGAGACAGGAACTGGCGAGTTCAGGTGGGGTGGGGACAGCACAGA
10801  CAAGACACGGGGAACCCTCTGTCCTTGACCGCTCAAGTCCACCCCACCCCTGTCGTGTCT

10861  CTGTTCCACCGTTGTGCATATTGTTGCTTCTGAACCACAAACTGTATAAATGGATGGTTT
10861  GACAAGGTGGCAACACGTATAACAACGAAGACTTGGTGTTTGACATATTTACCTACCAAA

10921  TTTG
10921  AAAC
```

FIGURE 2

>hg19_refGene_NR_037593 range=chr6:163834097-163836982 5'pad=0 3'pad=0 strand=- repeatMasking=none

CAHM resides on minus strand, thus the top strand in the below is the representative 5'- to 3'- sequence of CAHM

Yellow: sub-regions of interests towards which three bisulphite conversion and methylation specific assays are given as an example. Note this is also an example on assays running on opposite strands

```
   1  AGCATAATTTCGGTCTCCTACGGAATTTCTTTAAACTAACGAGCCCCACACTGACTTCCA
   1  TCGTATTAAAGCCAGAGGATGCCTTAAAGAAATTTGATTGCTCGGGGTGTGACTGAAGGT

  61  CTTAGTATTGAACAGCTGCACCAACTCCCCAACCACGAAACCCCGCTTTAGTCCTCACGG
  61  GAATCATAACTTGTCGACGTGGTTGAGGGGTTGGTGCTTTGGGGCGAAATCAGGAGTGCC

 121  GAGGGGAGAGCTCCCGAACGCGGGGCGGTGGGGGTCCGCGCAGAGCAGGCTCGGGAGGCC
 121  CTCCCCTCTCGAGGGCTTGCGCCCCGCCACCCCCAGGCGCGTCTCGTCCGAGCCCTCCGG

 181  GAGCGGAGGAGCGGGGCCCGCGCCGGGCACACGCGTCCTCCAGGCCGGCGGCGCGGCGTG
 181  CTCGCCTCCTCGCCCCGGGCGCGGCCCGTGTGCGCAGGAGGTCCGGCCGCCGCGCCGCAC

 241  GGCCTGCTCTCGGGAGGGCTCGGCGAGGACCACCGAGTCCGCGCAGCCCCCGCCGACCTC
 241  CCGGACGAGAGCCCTCCCGAGCCGCTCCTGGTGGCTCAGGCGCGTCGGGGGCGGCTGGAG

 301  TCGGAAATGTGACCGCCGAAGCTAGAGGGGCCGCAGGGCTCACCCCGCAGCCCTCCGCAA
 301  AGCCTTTACACTGGCGGCTTCGATCTCCCCGGCGTCCCGAGTGGGGCGTCGGGAGGCGTT

 361  CGCGCCAACTTGGACGCCAGCGGCGCGGGGTCCCGGCCGGGCAGGGGGCGTCCGGCACCG
 361  GCGCGGTTGAACCTGCGGTCGCCGCGCCCCAGGGCCGGCCCGTCCCCCGCAGGCCGTGGC

 421  GCGGGTCCACCCCGGCCCGCCCGCCCGCCGCAAGCCCGGCCCGGCACCTCCTTTGTGCGC
 421  CGCCCAGGTGGGGCCGGGCGGGCGGGCGGCGTTCGGGCCGGGCCGTGGAGGAAACACGCG

 481  CCTCCTGGCCTCCCAGTCCCCCCGGCGCTCGGTCCCGCCCGGCCGTGACCTTCCCGCCCT
 481  GGAGGACCGGAGGGTCAGGGGGGCCGCGAGCCAGGGCGGGCCGGCACTGGAAGGGCGGGA

 541  CCCCACCATCCCAAGCGGGGAAAGGGGCCGCCCCGGCGGGGGTCGGGCCGGGGCCGGGGC
 541  GGGGTGGTAGGGTTCGCCCCTTTCCCCGGCGGGGCCGCCCCCAGCCCGGCCCCGGCCCCG

 601  CCTGGAGACGCTCACCTTCGTCCAGCAGCCGCTCGAGGTGGTTGAAGATCCCGCAGAAGT
 601  GGACCTCTGCGAGTGGAAGCAGGTCGTCGGCGAGCTCCACCAACTTCTAGGGCGTCTTCA

 661  TGGGCAGGCTGCTCATGAGCTTCTTGTCGTTCATCAGCTGCATCAGGTAATCTGGGGTGG
 661  ACCCGTCCGACGAGTACTCGAAGAACAGCAAGTAGTCGACGTAGTCCATTAGACCCCACC

 721  GCTTCGGCTTCTCCTTCGTTTCCATTTCCCCGACCATATTCCAGGCTCCGCAGCTCACTC
 721  CGAAGCCGAAGAGGAAGCAAAGGTAAAGGGGCTGGTATAAGGTCCGAGGCGTCGAGTGAG

 781  CGCCCGCCGCCGCCGCCGCCGCCGGAGAGGAGGGAGGGGCGGGGGCGAGCCCCGGCC
 781  GCGGGCGGCGGCGGCGGCGGCGGCCTCTCCTCCCTCCCCGCCCCCGCTCGGGGCCGG

 841  GCGGGCCGCTCGGGACCCGGCGTCCCGCTCCGCGCCGGCTCCCGCTCTGGCTCCCGCGCC
 841  CGCCCGGCGAGCCCTGGGCCGCAGGGCGAGGCGCGGCCGAGGGCGAGACCGAGGGCGCGG

 901  GAGCCCCGGGGCGGCCGGCACCGCGCGCTCGCCCGCCCCCCGCAGGCACTTTCCGCCCGG
 901  CTCGGGGCCCCGCCGGCCGTGGCGCGCGAGCGGGCGGGGGGCGTCCGTGAAAGGCGGGCC

 961  GCCCGACGCGCTCCCAGCCGCCGCCGCTCTCGGCCGCCCGGCTCGGCGCGTCCCTCACGA
 961  CGGGCTGCGCGAGGGTCGGCGGCGGCGAGAGCCGGCGGGCCGAGCCGCGCAGGGAGTGCT

1021  GGCCGGCGGGCGGGCGGGAGGCGGGTCTCGGCTGGCCGAGTGCGGGCGCGGGGCGCGGGC
1021  CCGGCCGCCCGCCCGCCCTCCGCCCAGAGCCGACCGGCTCACGCCCGCGCCCCGCGCCCG

1081  TGCTGGCGCCGGACGAGGAGGAGCCGGGGCCGCCGCTCGCTCGGCTGGTCGAGCCCGCTC
1081  ACGACCGCGGCCTGCTCCTCCTCGGCCCCGGCGGCGAGCGAGCCGACCAGCTCGGGCGAG
```

FIGURE 2 (cont'd)

```
1141  AGCGCCGCCGCCGCCTGTGCCGCGGGCTCGGGGGGTCTCTCTGGGCTGGTCCCCGCCGCG
1141  TCGCGGCGGCGGCGGACACGGCGCCCGAGCCCCCCAGAGAGACCCGACCAGGGGCGGCGC

1201  GCGCCTCTGCCGCTGCCGCCGCGCTCCGACTGCGCTCCTCCTCTTCCTCCTCCTCCTTCT
1201  CGCGGAGACGGCGACGGCGGCGCGAGGCTGACGCGAGGAGGAGAAGGAGGAGGAGGAAGA

1261  CCTCCTCCTCCTCCTCACTCACTTGGCGGCGGAGTTCGCCTCAGTTCAGGCGGCGCTGCC
1261  GGAGGAGGAGGAGGAGTGAGTGAACCGCCGCCTCAAGCGGAGTCAAGTCCGCCGCGACGG

1321  AGCGGCGGCGGCAGCAGCGGCGGCGGCGGGGGGAGGGGCAGCGGCGCGGGGAGGGCCGGG
1321  TCGCCGCCGCCGTCGTCGCCGCCGCCGCCCCCCTCCCCGTCGCCGCGCCCCTCCCGGCCC

1381  GGCGGGCGGCGGGCGGGGCCGCGCAGGGCGGCCGTTAGCGGGTCCCGGCCGGCGCGCTCC
1381  CCGCCCGCCGCCCGCCCCGGCGCGTCCCGCCGGCAATCGCCCAGGGCCGGCCGCGCGAGG

1441  CGCTGCCTGCTGCCCGCCGCCCCGCGACCGGGGCAGCCGCCGGCGCCGCGCCCCCTCCTC
1441  GCGACGGACGACGGGCGGCGGGGCGCTGGCCCCGTCGGCGGCCGCGGCGCGGGGGAGGAG

1501  CCTGCCTCCCCGGCGGGGCGGGCGACGGGGGCAACGGGAGTGTGACGGACTGCGGCGCTC
1501  GGACGGAGGGGCCGCCCCGCCCGCTGCCCCCGTTGCCCTCACACTGCCTGACGCCGCGAG

1561  GGCGGGCCTGGGCGCAAGGACTGGGTCGCCGCCGTGCGGGGCGCCCGGGGCGGGTGGGTG
1561  CCGCCCGGACCCGCGTTCCTGACCCAGCGGCGGCACGCCCCGCGGGCCCCGCCCACCCAC

1621  TGGGCTGCGGGGCCGACGGCCGCGGGGGGGGCGCGGGGCTTTCCACGGCACGGGGCGGGT
1621  ACCCGACGCCCCGGCTGCCGGCGCCCCCCCCGCGCCCCGAAAGGTGCCGTGCCCCGCCCA

1681  GGGGCTGCGGGCGCGGTCGGGGCTGGCGCGGCAGGCGAGAGGAATTCCAGCGGGACGCGC
1681  CCCCGACGCCCGCGCCAGCCCCGACCGCGCCGTCCGCTCTCCTTAAGGTCGCCCTGCGCG

1741  CAGCGCTGCCTCCGCCTCGCCTCCCGTTACTATAGTTTTTTCAGAACCTCTTATCCACCC
1741  GTCGCGACGGAGGCGGAGCGGAGGGCAATGATATCAAAAAAGTCTTGGAGAATAGGTGGG

1801  CGCAGAAGAGAAATTTTTTTTTAACGTCCTGTTGCTTCACGGGGGACTTTCGGGGTGGGC
1801  GCGTCTTCTCTTTAAAAAAAAATTGCAGGACAACGAAGTGCCCCCTGAAAGCCCCACCCG

1861  ATCACGTGGCCCGACACGTAGCCGTGGGGCTCCAGAGCGCGTGCGCGCGTCGAGCCCGGT
1861  TAGTGCACCGGGCTGTGCATCGGCACCCCGAGGTCTCGCGCACGCGCGCAGCTCGGGCCA

1921  CGCCACGCCCCTTCCGAGGGTCGCTCCGCCCCCGTCGGGCCCCGCGAGCGCGCGGTGGGG
1921  GCGGTGCGGGGAAGGCTCCCAGCGAGGCGGGGGCAGCCCGGGGCGCTCGCGCGCCACCCC

1981  GAGGGGAGGGCCCCGGAGCGCGCCTGCGTGGGGCGGGGGCGGCAGCCGACTAGGGGCTGG
1981  CTCCCCTCCCGGGGCCTCGCGCGGACGCACCCCGCCCCCGCCGTCGGCTGATCCCCGACC

2041  GTCTGGCCGTTTAGGGCCGGGTCTTGGCCCGTCGCCCACGGTGCGGAGGGCTGGTGGGCT
2041  CAGACCGGCAAATCCCGGCCCAGAACCGGGCAGCGGGTGCCACGCCTCCCGACCACCCGA

2101  TTCCTTGGCCGTCGGGCCCGCCACGGCGCGGGTCTTGGCTGCGGGGCGGAGGTGGGGCGG
2101  AAGGAACCGGCAGCCCGGGCGGTGCCGCGCCCAGAACCGACGCCCCGCCTCCACCCCGCC

2161  GAGAGCCGAGGATAAGAGTTTGAGGCTTTTCGAGGCGCGTGCCGCGGCGTCCGCCTCTGC
2161  CTCTCGGCTCCTATTCTCAAACTCCGAAAAGCTCCGCGCACGGCGCCGCAGGCGGAGACG

2221  GGGACTCTGCGCCGGGCGCCCTCGGCCGGCGCGCCCGGCTCCCGCTTTGTCGCCGAGGGA
2221  CCCTGAGACGCGGCCCGCGGGAGCCGGCCGCGCGGGCCGAGGGCGAAACAGCGGCTCCCT

2281  AGCACGCGCGACGCCCCTCCCGTCGCCGCCGTGGCTTCTTTCGGTGTTCGTGATTTGCTG
2281  TCGTGCGCGCTGCGGGGAGGGCAGCGGCGGCACCGAAGAAAGCCACAAGCACTAAACGAC

2341  AGAGGCTGGAAAGCAGCACGGCGGAGAGGAGCCTTGCACTCGCCAGGCGGGAAGCCTGCG
2341  TCTCCGACCTTTCGTCGTGCCGCCTCTCCTCGGAACGTGAGCGGTCCGCCCTTCGGACGC

2401  CGGACACGCGTGCGCACCCACGGGGCGGCGGGCGGGCGTGGGGGGTCCGGGCCACGCGGG
2401  GCCTGTGCGCACGCGTGGGTGCCCCGCCGCCCGCCCGCACCCCCCAGGCCCGGTGCGCCC
```

FIGURE 2 (cont'd)

```
2461  CGACGCGCCTCTAGGGAAGCGATCTTGTTGCACCTTCCCGTTATTCTGAAAGCAAATCGT
2461  GCTGCGCGGAGATCCCTTCGCTAGAACAACGTGGAAGGGCAATAAGACTTTCGTTTAGCA

2521  AGCCAGACCCGAGCGCAGCGGCTTAGCAAATAATAAGGGGAGCGTCAGTCGTGCTCGAAA
2521  TCGGTCTGGGCTCGCGTCGCCGAATCGTTTATTATTCCCCTCGCAGTCAGCACGAGCTTT

2581  TGCTTCCTTCGCGATGGCGTCAGTGTTCCGTGAGGGAATGAAGCCGCAGTAGGAAATAAA
2581  ACGAAGGAAGCGCTACCGCAGTCACAAGGCACTCCCTTACTTCGGCGTCATCCTTTATTT

2641  GAGGCTGTGCGCGTAGTCTGAAAAGCAGAAGTCAACATTTTTACAGATGAAGAAAGAATA
2641  CTCCGACACGCGCATCAGACTTTTCGTCTTCAGTTGTAAAAATGTCTACTTCTTTCTTAT

2701  CGGAGGCAAGAGGTCTTTCTCTGCAGTTTGGTGGATTTCCAACATTTAGACTTGTTTGGA
2701  GCCTCCGTTCTCCAGAAAGAGACGTCAAACCACCTAAAGGTTGTAAATCTGAACAAACCT

2761  AGAATTTCCTCAGCTGCACCAATGAAGTCCTTGATCTATAGAAGTCGGCAGTCCCTAAAT
2761  TCTTAAAGGAGTCGACGTGGTTACTTCAGGAACTAGATATCTTCAGCCGTCAGGGATTTA

2821  CTACGTCTGCATTTTGTTGCAAATCCTTTATAACATTCCATTAAAATAATGCAGAGTTAT
2821  GATGCAGACGTAAAACAACGTTTAGGAAATATTGTAAGGTAATTTTATTACGTCTCAATA

2881  TTAATA
2881  AATTAT
```

FIGURE 3

>hg19_refGene_NM_001195286 range=chr6:389739-411443 5'pad=0 3'pad=0 strand=+ repeatMasking=none

The IRF4 gene resides on the plus strand. The top strand in the below is the representative 5'-3' gene sequence of IRF4

Yellow: Sub-region of interest towards which a bisulphate conversion and methylation specfic assay is given as an example

```
   1  AGATTAGCAATTACATTTAAAAACAAGGTAGTGGTGTGAAGGCTCAGAGTTTCACACAAC
   1  TCTAATCGTTAATGTAAATTTTTGTTCCATCACCACACTTCCGAGTCTCAAAGTGTGTTG

  61  GACTGTTGTTCTGAACTAAAAGCAGAGGATGTGGATGAGTGAGGGGAGGGTATGTTGAAA
  61  CTGACAACAAGACTTGATTTTCGTCTCCTACACCTACTCACTCCCCTCCCATACAACTTT

 121  ATGACTAAACCCGGTGCAGAGGACACGTATACGAAAGTGCAAAAAGCTCTCATTTGCCAT
 121  TACTGATTTGGGCCACGTCTCCTGTGCATATGCTTTCACGTTTTTCGAGAGTAAACGGTA

 181  TTTTTGTCCTCAGCAAACTCCTCTCTGGTATGTTTGTGAACGTTTGAGGGAAAGGTAACA
 181  AAAAACAGGAGTCGTTTGAGGAGAGACCATACAAACACTTGCAAACTCCCTTTCCATTGT

 241  CTGATTAAAATATAATGCAACAAGAAGGCAACGATCTTTGCTTTGTTCACTGATGTGTTC
 241  GACTAATTTTATATTACGTTGTTCTTCCGTTGCTAGAAACGAAACAAGTGACTACACAAG

 301  CAAGTGCCTAGGACGGTTCCTGACAAGTAGCAGGTGCTCAAAATACATTTGTCAAGTGAA
 301  GTTCACGGATCCTGCCAAGGACTGTTCATCGTCCACGAGTTTTATGTAAACAGTTCACTT

 361  TAAATCTGTGAAAGGAAAGGAAGAAAAACAGGTTGTGACCAGAACTTTTTGTGAACAAAA
 361  ATTTAGACACTTTCCTTTCCTTCTTTTTGTCCAACACTGGTCTTGAAAAACACTTGTTTT

 421  CCCTGTTGTTTATTACATAGACTTTGGTAATGGAAAACTTACTTGGGTGAAATAAAAAAT
 421  GGGACAACAAATAATGTATCTGAAACCATTACCTTTTGAATGAACCCACTTTATTTTTTA

 481  GGCATCCTGTTAATTTCTGAGAACGTCTGTTGGTTTTAGGAAAAGCATACCTATCCATAC
 481  CCGTAGGACAATTAAAGACTCTTGCAGACAACCAAAATCCTTTTCGTATGGATAGGTATG

 541  CTCACTACCTGTTTTGCTCTAGCCTAAACATAGATTCTTCTTTGGTTAGGAGACATTGAC
 541  GAGTGATGGACAAAACGAGATCGGATTTGTATCTAAGAAGAAACCAATCCTCTGTAACTG

 601  CCAATATATAAGATTAAAACAAAAACTGAGAAATAAGTGAAATCAAAAGCTTTTCAAAAG
 601  GGTTATATATTCTAATTTTGTTTTTGACTCTTTATTCACTTTAGTTTTCGAAAAGTTTTC

 661  TTGCAGGTTGACCTACGGTGGTGGTGGTCTAAAATCATATAGAAACCAGAGATCTGATGG
 661  AACGTCCAACTGGATGCCACCACCACCAGATTTTAGTATATCTTTGGTCTCTAGACTACC

 721  TTTACCTGAAAATGCGAAGTAAATGCACATACTCAGCATCTCAGACGATCGAAAGCTCAA
 721  AAATGGACTTTTACGCTTCATTTACGTGTATGAGTCGTAGAGTCTGCTAGCTTTCGAGTT

 781  CGGGTGAAAGCTCAGGGGGTTAATTTTGCGATTAAGGACATCTTGGAAAAGTATGTAAAA
 781  GCCCACTTTCGAGTCCCCCAATTAAAACGCTAATTCCTGTAGAACCTTTTCATACATTTT

 841  TCCCTGGTCCACTTAAGTATTATTCCTATTTTGGGCTTTTATTTATTTATTTTGAGATGG
 841  AGGGACCAGGTGAATTCATAATAAGGATAAAACCCGAAAATAAATAAATAAAACTCTACC

 901  AGTCTTGCTCTGTCGCCCACCTCGGCTCACTGCAACTTCAGCCTCCCGGGTTCAAGCGAT
 901  TCAGAACGAGACAGCGGGTGGAGCCGAGTGACGTTGAAGTCGGAGGGCCCAAGTTCGCTA

 961  TCTCGTGGCTCGTCCTCTCAAGTAGCTGGGGCCACGCCCGGCTAATTTTTGTATTTTTAG
 961  AGAGCACCGAGCAGGAGAGTTCATCGACCCCGGTGCGGGCCGATTAAAAACATAAAAATC

1021  TAGAGATGGGGTTTCACCGTGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGCGATC
1021  ATCTCTACCCCAAAGTGGCACAACCGGTCCGACCAGAGCTTGAGGACTGGAGTCCGCTAG

1081  CACCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCTTGAGCCACAGCGCCTGGCCTAT
1081  GTGGCGGAGCCGGAGGGTTTCACGACCCTAATGTCCGAACTCGGTGTCGCGGACCGGATA
```

FIGURE 3 (cont'd)

```
1141  TTTGGGCTTTTATACCCCACTGGTAAACTGCTTTCCTCCAGGTTGAGGTTAAAACGACAT
1141  AAACCCGAAAATATGGGGTGACCATTTGACGAAAGGAGGTCCAACTCCAATTTTGCTGTA

1201  CATTTTAAGGTGAACTGAAGTCTGGAAGTGATTAAGCACTTGGATCCTTAGGGAGCCTCT
1201  GTAAAATTCCACTTGACTTCAGACCTTCACTAATTCGTGAACCTAGGAATCCCTCGGAGA

1261  CCCCGCCCCCATCTCTTTCATGCTAAGATAATTAAAACTTCGGGGCCGGGGCATTGTCTG
1261  GGGGCGGGGGTAGAGAAAGTACGATTCTATTAATTTTGAAGCCCCGGCCCCGTAACAGAC

1321  AGTCACTTCAATTCACCAGCCTAATAGATGCAAAAGGATGTAAGCATGTCAGACACGCAG
1321  TCAGTGAAGTTAAGTGGTCGGATTATCTACGTTTTCCTACATTCGTACAGTCTGTGCGTC

1381  AGACAGTATTTGAATCAAGCTTAATAGCTCAAGGGAGCTGGGCCATTTCCTATTTTCTTT
1381  TCTGTCATAAACTTAGTTCGAATTATCGAGTTCCCTCGACCCGGTAAAGGATAAAAGAAA

1441  TTAGTGAGTGCGATGTTCTCTAAACACCGCGGAGAGGCAGGGTTCCCGGTGATGGCCTTG
1441  AATCACTCACGCTACAAGAGATTTGTGGCGCCTCTCCGTCCCAAGGGCCACTACCGGAAC

1501  CCGAGGGTGCTCCCGCAACCTCCACCTCCAGTTCTCTTTGGACCATTCCTCCGTCTTCCG
1501  GGCTCCCACGAGGGCGTTGGAGGTGGAGGTCAAGAGAAACCTGGTAAGGAGGCAGAAGGC

1561  TTACACGCTCTGCAAAGCGAAGTCCCCTTCGCACCAGATTCCCGCTACTACACGCCCCCC
1561  AATGTGCGAGACGTTTCGCTTCAGGGGAAGCGTGGTCTAAGGGCGATGATGTGCGGGGGG

1621  ATTTCCCGCCCTGGCCACATCGCTGCAGTTTAGTGATTGACTGGCCTCCTGAGGTCCTGG
1621  TAAAGGGCGGGACCGGTGTAGCGACGTCAAATCACTAACTGACCGGAGGACTCCAGGACC

1681  CGCAAAGGCGAGATTCGCATTTCGCACCTCGCCCTTCGCGGGAAACGGCCCCAGTGACAG
1681  GCGTTTCCGCTCTAAGCGTAAAGCGTGGAGCGGGAAGCGCCCTTTGCCGGGGTCACTGTC

1741  TCCCCGAAGCGGCGCGCGCCCGGCTGGAGGTGCGCTCTCCGGGCGCGGCGCGCGGAGGGT
1741  AGGGGCTTCGCCGCGCGCGGGCCGACCTCCACGCGAGAGGCCCGCGCCGCGCGCCTCCCA

1801  CGCCAAGGGCGCGGGAACCCCACCCCGGCCGCGGCAGCCCCCAGCCTTCACGCCGGCCCT
1801  GCGGTTCCCGCGCCCTTGGGGTGGGGCCGGCGCCGTCGGGGGTCGGAAGTGCGGCCGGGA

1861  GAGGCTCGCCCGCCCGGCCGGCCCCGGCTCTCGGCTTGCAAAGTCCCTCTCCCCAGTCCA
1861  CTCCGAGCGGGCGGGCCGGCCGGGGCCGAGAGCCGAACGTTTCAGGGAGAGGGGTCAGGT

1921  ACCCCCGGCCCCCACAGGCCTCGGCGCCCCGCCCCGCCCCAGGCCCCGCCCCAGAGAGTT
1921  TGGGGGCCGGGGGTGTCCGGAGCCGCGGGGCGGGGCGGGGTCCGGGGCGGGGTCTCTCAA

1981  CTATAAAGTTCCTCTTTCCCACCTCGCACTCTCAGTTTCACCGCTCGATCTTGGGACCCA
1981  GATATTTCAAGGAGAAAGGGTGGAGCGTGAGAGTCAAAGTGGCGAGCTAGAACCCTGGGT

2041  CCGCTGCCCTCAGCTCCGAGTCCAGGGCGAGGTAAGGGCTGGAGTCGGGCAGGAGGAGGG
2041  GGCGACGGGAGTCGAGGCTCAGGTCCCGCTCCATTCCCGACCTCAGCCCGTCCTCCTCCC

2101  GTGTGAGGCTGATACCAGAGAGGACCCGGAGCGCGAACCAGAGGTTCGACCTCCAGGGCA
2101  CACACTCCGACTATGGTCTCTCCTGGGCCTCGCGCTTGGTCTCCAAGCTGGAGGTCCCGT

2161  GCGCAGGGTACCCCGGCTTCGGAGCGGGAAGGGAGCGCGCCCCGTCCTGGAGCTCCGACT
2161  CGCGTCCCATGGGGCCGAAGCCTCGCCCTTCCCTCGCGCGGGGCAGGACCTCGAGGCTGA

2221  CCCACCCCATCTGCGCTGAGCCGGAGGCGCTGGTTTGGGCTCCAAGGCCCGCCTCCTTGG
2221  GGGTGGGGTAGACGCGACTCGGCCTCCGCGACCAAACCCGAGGTTCCGGGCGGAGGAACC

2281  CTCTGCCCGAGCCTCCCCGCCTGCCCTCCGCGCTCCTGCGACGGGGTCGCCACAAGCTGG
2281  GAGACGGGCTCGGAGGGGCGGACGGGAGGCGCGAGGACGCTGCCCCAGCGGTGTTCGACC

2341  ACGGGATGAGCTAACCGGACTGTCGGGGCCCCAGGAGTGGCTGAGGCGGGGCCGTCCAAG
2341  TGCCCTACTCGATTGGCCTGACAGCCCCGGGGTCCTCACCGACTCCGCCCCGGCAGGTTC

2401  GCACCCACACAAGACGGCACAACTGCCTGCGAGAAACAGGCCCGGCCCTGTGGACCCCAA
2401  CGTGGGTGTGTTCTGCCGTGTTGACGGACGCTCTTTGTCCGGGCCGGGACACCTGGGGTT
```

FIGURE 3 (cont'd)

```
2461  TCCGAGGCTCCTTCCCCTGCTCTTCGTTCCTAAGGGGCCCAAGCTCACGGCGGCCTCCGG
2461  AGGCTCCGAGGAAGGGGACGAGAAGCAAGGATTCCCCGGGTTCGAGTGCCGCCGGAGGCC

2521  CGCGGTGCTCACCCGCTGGCGCAGGAGGAGGAGGAGCTCCACATTTGGGTCGCTCCGAGC
2521  GCGCCACGAGTGGGCGACCGCGTCCTCCTCCTCCTCGAGGTGTAAACCCAGCGAGGCTCG

2581  CTTGCGTGCGGTGGCCTAGCCGGCCTGGCGCGGTCCCTGCCTCCCAGGCTCCGCAGCTGT
2581  GAACGCACGCCACCGGATCGGCCGGACCGCGCCAGGGACGGAGGGTCCGAGGCGTCGACA

2641  CGTCGCCCTCTCCCGCGCCCTCCCCGCCTCCGCTCTCCCGGGCCTGCTCCGGGGTCCGGC
2641  GCAGCGGGAGAGGGCGCGGGAGGGGCGGAGGCGAGAGGGCCCGGACGAGGCCCCAGGCCG

2701  GGACGCTCTGCGCGCGGAATCCCCCGTACTGGGGCTGCAGCCCCCGCGTCTGCGCCACTT
2701  CCTGCGAGACGCGCGCCTTAGGGGGCATGACCCCGACGTCGGGGGCGCAGACGCGGTGAA

2761  GTCGTTTGCAGAGCCCACTTAGTGCGCGCTAGCTGGGCAGGGATAGGGGTCCTATTCGGG
2761  CAGCAAACGTCTCGGGTGAATCACGCGCGATCGACCCGTCCCTATCCCCAGGATAAGCCC

2821  GCGAAGGGTCTGGATGCGAGCAGAGAAAGCGGAGGGTGGAGGAACCCGGGGCTGCGCCCC
2821  CGCTTCCCAGACCTACGCTCGTCTCTTTCGCCTCCCACCTCCTTGGGCCCCGACGCGGGG

2881  TGGAACGCCCGGCCGCAGGCGAGGTCCTCCGCGCGTGGAGGCCGCCAGGGGAGTGGAAAC
2881  ACCTTGCGGGCCGGCGTCCGCTCCAGGAGGCGCGCACCTCCGGCGGTCCCCTCACCTTTG

2941  TGACAGAGTCGCGGGGAAGGGGCGAGAAGCGGGTTGGGAGTGAGCGAAGGCAAGCGAGAG
2941  ACTGTCTCAGCGCCCCTTCCCCGCTCTTCGCCCAACCCTCACTCGCTTCCGTTCGCTCTC

3001  CTGCGAGTGAGTGCGGAAGGAGGGCCAGGAGGGGTGGCGGCTGGGTGGGGAGAGAGGGTG
3001  GACGCTCACTCACGCCTTCCTCCCGGTCCTCCCCACCGCCGACCCACCCCTCTCTCCCAC

3061  CAAGACGAGCGGCGCGTGTCGGGAGCCTTTGGGCTGCGGGTGCGTTACAGGAGAGCAGGC
3061  GTTCTGCTCGCCGCGCACAGCCCTCGGAAACCCGACGCCCACGCAATGTCCTCTCGTCCG

3121  GGGTAGGAGCCTTCGCGGGGGCCGAGCTCGGAAGGCGGACGGCTGTGCCCGCCCAGGGGA
3121  CCCATCCTCGGAAGCGCCCCCGGCTCGAGCCTTCCGCCTGCCGACACGGGCGGGTCCCCT

3181  TGCGCCCGGGCCGGCCGCGAAGGTGCCTTCTTCCGGGGGCCCGGACGACCCTGACACGGC
3181  ACGCGGGCCCGGCCGGCGCTTCCACGGAAGAAGGCCCCCGGGCCTGCTGGGACTGTGCCG

3241  ACGCGCGCGCTTCGCAGCCTCAAAGACTCCGGGGCCTCGTGGTCACTGGCGCAGGGGATC
3241  TGCGCGCGCGAAGCGTCGGAGTTTCTGAGGCCCCGGAGCACCAGTGACCGCGTCCCCTAG

3301  GGGGCGGGGTGCCCGGAGTGCGGTGCCTCGTGGCTGAAGGGCAGCTCTTCTCCCCGCAGT
3301  CCCCGCCCCACGGGCCTCACGCCACGGAGCACCGACTTCCCGTCGAGAAGAGGGGCGTCA

3361  GCAGAGCAGAGCGGGCGGAGGACCCCGGGCGCGGGCGCGGACGGCACGCGGGGCATGAAC
3361  CGTCTCGTCTCGCCCGCCTCCTGGGGCCCGCGCCCGCGCCTGCCGTGCGCCCCGTACTTG

3421  CTGGAGGGCGGCGGCCGAGGCGGAGAGTTCGGCATGAGCGCGGTGAGCTGCGGCAACGGG
3421  GACCTCCCGCCGCCGGCTCCGCCTCTCAAGCCGTACTCGCGCCACTCGACGCCGTTGCCC

3481  AAGCTCCGCCAGTGGCTGATCGACCAGATCGACAGCGGCAAGTACCCCGGGCTGGTGTGG
3481  TTCGAGGCGGTCACCGACTAGCTGGTCTAGCTGTCGCCGTTCATGGGGCCCGACCACACC

3541  GAGAACGAGGAGAAGAGCATCTTCCGCATCCCCTGGAAGCACGCGGGCAAGCAGGACTAC
3541  CTCTTGCTCCTCTTCTCGTAGAAGGCGTAGGGGACCTTCGTGCGCCCGTTCGTCCTGATG

3601  AACCGCGAGGAGGACGCCGCGCTCTTCAAGGTCTCCGGCCTCGGGAGCCGGCGGGGGCGC
3601  TTGGCGCTCCTCCTGCGGCGCGAGAAGTTCCAGAGGCCGGAGCCCTCGGCCGCCCCCGCG

3661  GCCGGGGAGGGCCCAGAGACAGAGCCCGGGGTCCCCGGCGCCGCCTCCGAGGCGAGCCCA
3661  CGGCCCCTCCCGGGTCTCTGTCTCGGGCCCCAGGGGCCGCGGCGGAGGCTCCGCTCGGGT

3721  GGGGACCGCGCGGGGCGGACGGGCGGGCGGCGGAGGCATCAGGTGGCGTCGCCGGAGCCG
3721  CCCCTGGCGCGCCCCGCCTGCCCGCCCGCCGCCTCCGTAGTCCACCGCAGCGGCCTCGGC
```

FIGURE 3 (cont'd)

```
3781  CAGGAGGAGGAAAGGAGGCCTCGGCTCTCAGCGGGACCGCGGGGGCCGGGAGCCGGGTCC
3781  GTCCTCCTCCTTTCCTCCGGAGCCGAGAGTCGCCCTGGCGCCCCCGGCCCTCGGCCCAGG

3841  TGGGCGCGTGGAGGCTGCAGGGAAACCGCTGAAGGCCCGGCCGGGCCCGGGGAAGGGCGG
3841  ACCCGCGCACCTCCGACGTCCCTTTGGCGACTTCCGGGCCGGCCCGGGCCCCTTCCCGCC

3901  CCAAAGGCTTGAGGGGTTTTGCGCGTTCGTCCGTGCGTTCTCGTTTCCACGCAAGCCTCC
3901  GGTTTCCGAACTCCCCAAAACGCGCAAGCAGGCACGCAAGAGCAAAGGTGCGTTCGGAGG

3961  CGCCCTTCCTCCGGGCTCCCGTCTGCCGCCTCCGTCCGTGGGTCCCCCTCGCCCTCTCCG
3961  GCGGGAAGGAGGCCCGAGGGCAGACGGCGGAGGCAGGCACCCAGGGGGAGCGGGAGAGGC

4021  TGCGTCCGCGCCTGTGCCGGCGGCTGTTTTCGTCTCTCACCGCGTCTCTGTTTCTCTTTT
4021  ACGCAGGCGCGGACACGGCCGCCGACAAAAGCAGAGAGTGGCGCAGAGACAAAGAGAAAA

4081  CGCTGCTTTTCTCTCTGAGTCTCTCTCTCTCCATGTTTTTCCTGAGGTCAGCCTCTCTTC
4081  GCGACGAAAAGAGAGACTCAGAGAGAGAGAGGTACAAAAAGGACTCCAGTCGGAGAGAAG

4141  TCGCTCCTGCTAGCTCTCTGCGGGTACTCCCACCTCTGTCTTTCTCTTTGTGTGTCTCTG
4141  AGCGAGGACGATCGAGAGACGCCCATGAGGGTGGAGACAGAAAGAGAAACACACAGAGAC

4201  TCTCTCTCTTTCCCCCATCGCAGTGGAACTCAGGGCCTCTGTCTAGAGCTGTCTCCCTTG
4201  AGAGAGAGAAAGGGGGTAGCGTCACCTTGAGTCCCGGAGACAGATCTCGACAGAGGGAAC

4261  CCCTTTGCGCGAGTGCACACACGTGTGTCGTTGTTACGATTGTTCTCCCCTAAGGCAGTT
4261  GGGAAACGCGCTCACGTGTGTGCACACAGCAACAATGCTAACAAGAGGGGATTCCGTCAA

4321  TACCCAGAGAACTACGTGTCTGGGCCCAGCCCCCACCTGTGGGCAGAGCAGGGGAAGGGG
4321  ATGGGTCTCTTGATGCACAGACCCGGGTCGGGGGTGGACACCCGTCTCGTCCCCTTCCCC

4381  ACTTCCTCCGGGAATTTGGTCTCAATTTGCTCTCAGAGTGCCTCAGCTGTGCTGCCATCC
4381  TGAAGGAGGCCCTTAAACCAGAGTTAAACGAGAGTCTCACGGAGTCGACACGACGGTAGG

4441  AGATGTCTCCTGTGGGTGACAGCTCACACCACAGCTGTCCTTAGTCCTAGGCAAGCTCAC
4441  TCTACAGAGGACACCCACTGTCGAGTGTGGTGTCGACAGGAATCAGGATCCGTTCGAGTG

4501  TCAGACACTGGGTGGGTAGAGCCCCCTCAAGGAACCTCCGCATCTCACTCTACCGGTTAT
4501  AGTCTGTGACCCACCCATCTCGGGGGAGTTCCTTGGAGGCGTAGAGTGAGATGGCCAATA

4561  AAATACCCAGAAAATGTGTCTTCAACTTGGCAGTGATAGGGTCCAAGATGGAAAATCATT
4561  TTTATGGGTCTTTTACACAGAAGTTGAACCGTCACTATCCCAGGTTCTACCTTTTAGTAA

4621  TTCCTAGAGCATAGCTGGGGTCTTTCAGTTTACGTTCTGAGCAACGGTGTAAATCTGAAG
4621  AAGGATCTCGTATCGACCCCAGAAAGTCAAATGCAAGACTCGTTGCCACATTTAGACTTC

4681  GACCTATGCGCCATTCTTTCTTTTAGAAATACAATTTCAAGATTCCTTATAAAATATCTG
4681  CTGGATACGCGGTAAGAAAGAAAATCTTTATGTTAAAGTTCTAAGGAATATTTTATAGAC

4741  TATTTTGGGTTTAGCATGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAAGATGAGGT
4741  ATAAAACCCAAATCGTACCACCGAGTGTGGACATTAGGGTCGTGAAACCCTTCTACTCCA

4801  GGAAGGATGCCTATGCAACATAGTGAGACCCTGCCTCTGCAAAAAATAATAATTAAAAAA
4801  CCTTCCTACGGATACGTTGTATCACTCTGGGACGGAGACGTTTTTTATTATTAATTTTTT

4861  ATTAGCCAGGCATGGTGGTGAACACCTGTAGGCCGAGGGCCCAGCTACTCTAGAGGCTGA
4861  TAATCGGTCCGTACCACCACTTGTGGACATCCGGCTCCCGGGTCGATGAGATCTCCGACT

4921  GATAGGAGGACCCTAGAACTCAGGAGTTCGATGCTGCGGTGAGCTATCATCGTGCCACTG
4921  CTATCCTCCTGGGATCTTGAGTCCTCAAGCTACGACGCCACTCGATAGTAGCACGGTGAC

4981  TACTCTAGCCTGGGCAGCAGAGCAGGACTCTGATTCCAAAAAGAAAGCTAATAAACCAGA
4981  ATGAGATCGGACCCGTCGTCTCGTCCTGAGACTAAGGTTTTTCTTTCGATTATTTGGTCT

5041  CATGTATTTTGACTTTTCGTTCTCTTCATTCTTTCCCACCAGGCTTGGGCACTGTTTAAA
5041  GTACATAAAACTGAAAAGCAAGAGAAGTAAGAAAGGGTGGTCCGAACCCGTGACAAATTT
```

FIGURE 3 (cont'd)

```
5101  GGAAAGTTCCGAGAAGGCATCGACAAGCCGGACCCTCCCACCTGGAAGACGCGCCTGCGG
5101  CCTTTCAAGGCTCTTCCGTAGCTGTTCGGCCTGGGAGGGTGGACCTTCTGCGCGGACGCC

5161  TGCGCTTTGAACAAGAGCAATGACTTTGAGGAACTGGTTGAGCGGAGCCAGCTGGACATC
5161  ACGCGAAACTTGTTCTCGTTACTGAAACTCCTTGACCAACTCGCCTCGGTCGACCTGTAG

5221  TCAGACCCGTACAAAGTGTACAGGATTGTTCCTGAGGGAGCCAAAAAAGGTAGGGGCTCT
5221  AGTCTGGGCATGTTTCACATGTCCTAACAAGGACTCCCTCGGTTTTTTCCATCCCCGAGA

5281  CCTGAATTTGGGTCACCTAACAGAGGCAGCCAGATCCTTGAGGCACCTTAACTTCATTCT
5281  GGACTTAAACCCAGTGGATTGTCTCCGTCGGTCTAGGAACTCCGTGGAATTGAAGTAAGA

5341  GAGCATCACTTTCTAGCTTTCCTTTTGTATTGCCTGCCTGCCTGCCTTCTGCCTCACAGT
5341  CTCGTAGTGAAAGATCGAAAGGAAAACATAACGGACGGACGGACGGAAGACGGAGTGTCA

5401  GAAGCCCAGCCTCCTTGCCTGAGTTATGTGGGTCACAAGTTGAAATTCCTGTGGGGTTCA
5401  CTTCGGGTCGGAGGAACGGACTCAATACACCCAGTGTTCAACTTTAAGGACACCCCAAGT

5461  GCATAGGAGAATTAATCACACTGTATGCCTCAATGTATATGGGGGGGGGTGCATTGAATA
5461  CGTATCCTCTTAATTAGTGTGACATACGGAGTTACATATACCCCCCCCCACGTAACTTAT

5521  TGTGTTTTCCAGTTAGTCTTTCAAAAAAAAAAAGAAAAAAGAAAACTGGATAACTTGTTA
5521  ACACAAAAGGTCAATCAGAAAGTTTTTTTTTTCTTTTTCTTTTGACCTATTGAACAAT

5581  AAATTTACTTATCTAATAGTCTTTAGTGAGTGACTATTTTGCAGAAAAGGAGTAATGAGA
5581  TTTAAATGAATAGATTATCAGAAATCACTCACTGATAAAACGTCTTTTCCTCATTACTCT

5641  GGAATATCTCATCACTATGCATGGTAATAAAACTGTCAACTTTTAAAGTGTAACATGGTC
5641  CCTTATAGAGTAGTGATACGTACCATTATTTTGACAGTTGAAAATTTCACATTGTACCAG

5701  AGCTGAGGTGTTTATTAAAAGTCAGATTCCTGGACCCCTCTTTTCAAGATTCTGACCAGT
5701  TCGACTCCACAAATAATTTTCAGTCTAAGGACCTGGGGAGAAAAGTTCTAAGACTGGTCA

5761  GTGTCTGGAAGAGAGCTCAGGAATCTGCATCTGGAACAAATCCTTCAGGTGACCTCACTT
5761  CACAGACCTTCTCTCGAGTCCTTAGACGTAGACCTTGTTTAGGAAGTCCACTGGAGTGAA

5821  GGGAAACATCTGGCTGGCGGGGTCCTTTATGTCACAAGGCACATTCGCATGCCATCTGTC
5821  CCCTTTGTAGACCGACCGCCCCAGGAAATACAGTGTTCCGTGTAAGCGTACGGTAGACAG

5881  ATGTATTTCTTTACCTGTCCTGTACAACTCAGACCTTGTAGAAAAAAATCAGTTTAATAG
5881  TACATAAAGAAATGGACAGGACATGTTGAGTCTGGAACATCTTTTTTTAGTCAAATTATC

5941  CATGAAAAAATAACTTCCCAAGGGAGTTGCTGAAAAATACCTTATTTGTCAACACCGTGT
5941  GTACTTTTTTATTGAAGGGTTCCCTCAACGACTTTTTATGGAATAAACAGTTGTGGCACA

6001  TATGCATTCTAAGAATTAAATGCTCAGGTATTTTTACAAGATTTGACATTTAGTTAGGTC
6001  ATACGTAAGATTCTTAATTTACGAGTCCATAAAAATGTTCTAAACTGTAAATCAATCCAG

6061  AGTTCCTGTTTTTACGTTGTGCCATTTCCCTTTTCCCCAAACATGTAGGAGCCAAGCAGC
6061  TCAAGGACAAAAATGCAACACGGTAAAGGGAAAAGGGGTTTGTACATCCTCGGTTCGTCG

6121  TCACCCTGGAGGACCCGCAGATGTCCATGAGCCACCCCTACACCATGACAACGCCTTACC
6121  AGTGGGACCTCCTGGGCGTCTACAGGTACTCGGTGGGGATGTGGTACTGTTGCGGAATGG

6181  CTTCGCTCCCAGCCCAGGTATGGTGGAGGGCACTGGGCTCCCTGAGGGCGAGGCTGTGTG
6181  GAAGCGAGGGTCGGGTCCATACCACCTCCCGTGACCCGAGGGACTCCCGCTCCGACACAC

6241  GGCCAGCTGCCCACATGGCCAGAGAACCACAGCAGCCCAGACAGCAGAACTTGCCATTTG
6241  CCGGTCGACGGGTGTACCGGTCTCTTGGTGTCGTCGGGTCTGTCGTCTTGAACGGTAAAC

6301  CTATGGCTGCTCCAACAGCCCAGAAAAACCCCAGGTCACTGAACGAATGTCTCACTTTCC
6301  GATACCGACGAGGTTGTCGGGTCTTTTTGGGGTCCAGTGACTTGCTTACAGAGTGAAAGG

6361  ACACGGTGCTGCCATTGGTGTGGATTTTAAGTTGGGGAGGGTCGGGCGTGTCCGCCTGTT
6361  TGTGCCACGACGGTAACCACACCTAAAATTCAACCCCTCCCAGCCCGCACAGGCGGACAA
```

FIGURE 3 (cont'd)

```
6421  GGAATATGCTTCTCAGGTCTTCTGGGAAACAGATGTTTTGTGGAAGTGGAAGATTTTGGA
6421  CCTTATACGAAGAGTCCAGAAGACCCTTTGTCTACAAAACACCTTCACCTTCTAAAACCT

6481  AGTAGTGCCTTATCATGTGAAACCACAGGGCAGCTGATCTCTTCAGGCTTTCTTGATGTG
6481  TCATCACGGAATAGTACACTTTGGTGTCCCGTCGACTAGAGAAGTCCGAAAGAACTACAC

6541  AATGACAGCTTTGTTTCATCCACTTTGGTGGGTAAAAGAAGGCAAATTCCCCTGTGGTAC
6541  TTACTGTCGAAACAAAGTAGGTGAAACCACCCATTTTCTTCCGTTTAAGGGGACACCATG

6601  TTTTGGTGCCAGGTTTAGCCATATGACGAAGCTTTACATAAAACAGTACAAGTATCTCCA
6601  AAAACCACGGTCCAAATCGGTATACTGCTTCGAAATGTATTTTGTCATGTTCATAGAGGT

6661  TTGTCCTTTATGATCCTCCATGAGTGTTTTCACTTAGTCTGATGAAGGGTTCACTCCAGT
6661  AACAGGAAATACTAGGAGGTACTCACAAAAGTGAATCAGACTACTTCCCAAGTGAGGTCA

6721  CTTTTCGGATGATAAAATGCTTCGGCTGTCAGTCTAATAAGGGATTCCCTGAGGAGTTTG
6721  GAAAAGCCTACTATTTTACGAAGCCGACAGTCAGATTATTCCCTAAGGGACTCCTCAAAC

6781  GAGGCTGTAAGAGCACCCCCCGTCTCAATGCCAGTGCTTCTTATCTCAGCCTCTCCTGCA
6781  CTCCGACATTCTCGTGGGGGGCAGAGTTACGGTCACGAAGAATAGAGTCGGAGAGGACGT

6841  CTCCTTTACCCCCGTCTCGATGCCAGTGCTTCCTATCTCAGCCTCTCCTGCACTCCTTTA
6841  GAGGAAATGGGGGCAGAGCTACGGTCACGAAGGATAGAGTCGGAGAGGACGTGAGGAAAT

6901  CCCCCGTCTCGATGCCAGTGCTTCCTATCTCAGCCTCTCCTGCACTCCTTTACCCCCGTC
6901  GGGGGCAGAGCTACGGTCACGAAGGATAGAGTCGGAGAGGACGTGAGGAAATGGGGGCAG

6961  TCGATGCCAGTGCTTCTTATCTCAGCCTCTCCTGCACTCCTTTACCCCCGTCTCAATGCC
6961  AGCTACGGTCACGAAGAATAGAGTCGGAGAGGACGTGAGGAAATGGGGGCAGAGTTACGG

7021  AGTGCTTCCTATCTCAGCCTCTCCTGCACTCCTTTAACCCCGTCTCGATGCCAGTGCTTC
7021  TCACGAAGGATAGAGTCGGAGAGGACGTGAGGAAATTGGGGCAGAGCTACGGTCACGAAG

7081  TTATCTCAGCCTCTCCTGCACTCCTTTACCCCCGTCTCAATGCCAGTGCTTCCTATCTCA
7081  AATAGAGTCGGAGAGGACGTGAGGAAATGGGGGCAGAGTTACGGTCACGAAGGATAGAGT

7141  GCCTCTCCTGCACTCCTTTACCCCCGTCTCAATGCCAGTGCTTCCTATCTCAGCCTCTCC
7141  CGGAGAGGACGTGAGGAAATGGGGGCAGAGTTACGGTCACGAAGGATAGAGTCGGAGAGG

7201  TGCACTCCTTTACCCCCGTCTCGATGCCAGTGCTTCTTATCTCAGCCTCTCCTGCACTCC
7201  ACGTGAGGAAATGGGGGCAGAGCTACGGTCACGAAGAATAGAGTCGGAGAGGACGTGAGG

7261  TTTACCCCCGTCTCGATGCCAGTGCTTCCTATCTCAGCCTCTCCTGCACTCCTTTACCCC
7261  AAATGGGGGCAGAGCTACGGTCACGAAGGATAGAGTCGGAGAGGACGTGAGGAAATGGGG

7321  CGTCTCAATGCCAGTGCTTCTTATCTCAGCCTCTCCTGCACTCCTTTAGCAGGTTCACAA
7321  GCAGAGTTACGGTCACGAAGAATAGAGTCGGAGAGGACGTGAGGAAATCGTCCAAGTGTT

7381  CTACATGATGCCACCCCTCGACCGAAGCTGGAGGGACTACGTCCCGGATCAGCCACACCC
7381  GATGTACTACGGTGGGGAGCTGGCTTCGACCTCCCTGATGCAGGGCCTAGTCGGTGTGGG

7441  GGAAATCCCGTACCAATGTCCCATGACGTTTGGACCCCGCGGCCACCACTGGCAAGGCCC
7441  CCTTTAGGGCATGGTTACAGGGTACTGCAAACCTGGGGCGCCGGTGGTGACCGTTCCGGG

7501  AGCTTGTGAAAATGGTAAGGAGGATACCAGTGCAGGAAATAGAAGAGCTAATTGCTAATG
7501  TCGAACACTTTTACCATTCCTCCTATGGTCACGTCCTTTATCTTCTCGATTAACGATTAC

7561  TGGCCATGGGCCATGGCGAATCCTGGTCTGTCCTGGGCAGCACCAAAGCTCTTTCCCCTT
7561  ACCGGTACCCGGTACCGCTTAGGACCAGACAGGACCCGTCGTGGTTTCGAGAAAGGGGAA

7621  CTTAGAGGCTGCGTGTGCAGCTCGGGGAGAGGGGGGTTTTCTCACTCCTGTGGGATGGTG
7621  GAATCTCCGACGCACACGTCGAGCCCCTCTCCCCCCAAAAGAGTGAGGACACCCTACCAC

7681  GCATCCCACAGCCAAGTTTACTCTCAGGATCCATGCTAGGCACTGCCCTTCGTGGGATCT
7681  CGTAGGGTGTCGGTTCAAATGAGAGTCCTAGGTACGATCCGTGACGGGAAGCACCCTAGA
```

FIGURE 3 (cont'd)

```
7741  TATTTAAACACCACAGAGATCACTCCAGGCAGTGGAAAACAGCCCCCCACTTTTTATATA
7741  ATAAATTTGTGGTGTCTCTAGTGAGGTCCGTCACCTTTTGTCGGGGGGTGAAAAATATAT

7801  GAGAGGAGATTGTGGAAGCTCATAAATAATAAGTAGGGATGCTGGGAGATTATTGAGCTG
7801  CTCTCCTCTAACACCTTCGAGTATTTATTATTCATCCCTACGACCCTCTAATAACTCGAC

7861  GTATATTTCTCTCGAATAGTGGTGTGTTCATTGTTGCTATTGTTGGAATTAGAGGTGATA
7861  CATATAAAGAGAGCTTATCACCACACAAGTAACAACGATAACAACCTTAATCTCCACTAT

7921  CTAAAAAAGGAACGGAAAAATACCATGGCAAACACATAAAACGTCTCTACAGGTGCCTAC
7921  GATTTTTTCCTTGCCTTTTTATGGTACCGTTTGTGTATTTTGCAGAGATGTCCACGGATG

7981  CTACCACTGGGTCCCAGAAAGACTTGCTGATTGTGGATTAGATGTGGTCCCCTGCCCTCG
7981  GATGGTGACCCAGGGTCTTTCTGAACGACTAACACCTAATCTACACCAGGGGACGGGAGC

8041  GGAGTTCCCCAGACATCAAAATGAGCTCCCTCAGCAGAAGCTCTCATTAGCTGGGTGCAG
8041  CCTCAAGGGGTCTGTAGTTTTACTCGAGGGAGTCGTCTTCGAGAGTAATCGACCCACGTC

8101  GGAGAATCTAGCAATGCGGTAACTCAGGCCTTCAGGACTTTTGAGTTAGTTACTTAAATG
8101  CCTCTTAGATCGTTACGCCATTGAGTCCGGAAGTCCTGAAAACTCAATCAATGAATTTAC

8161  GAAGTCCTGAGACAGGTCAACACACTGAACCCTTTGGATGTTTTAGGAATATTTGTTCCT
8161  CTTCAGGACTCTGTCCAGTTGTGTGACTTGGGAAACCTACAAAATCCTTATAAACAAGGA

8221  TTGCAATTTGGCCTTTTTTTCCCCAGTCATGTTCAGGTTATCCTTGGCTGCAGGTAGTGA
8221  AACGTTAAACCGGAAAAAAAGGGGTCAGTACAAGTCCAATAGGAACCGACGTCCATCACT

8281  AGAGGTGTACCACTCTTTAATGTTTGTACTGAGTGGCATGTTCACAAAAAAGCTCAGAAT
8281  TCTCCACATGGTGAGAAATTACAAACATGACTCACCGTACAAGTGTTTTTTCGAGTCTTA

8341  TCAGTTGCATTGGCTTTTCAAGAGGGATTTGCAGCAAGGGCTGGTTTCCAATATGGTAGT
8341  AGTCAACGTAACCGAAAAGTTCTCCCTAAACGTCGTTCCCGACCAAAGGTTATACCATCA

8401  GGAGGAATTGGCTAACATAAAACTAGCTCCTGAAATTCAGACTTGCGTTTACTGCTCTGG
8401  CCTCCTTAACCGATTGTATTTTGATCGAGGACTTTAAGTCTGAACGCAAATGACGAGACC

8461  CTCTATGGAATTAGTAACCATGGGAAAGAGGCAATGCTTAGAACTGTGTAAAGTGATGGA
8461  GAGATACCTTAATCATTGGTACCCTTTCTCCGTTACGAATCTTGACACATTTCACTACCT

8521  GTAGGTAGAGACCTGTCTTGAAGCATCGGTAACAGTAAATGAAGCAGCAGCAAGATCAGG
8521  CATCCATCTCTGGACAGAACTTCGTAGCCATTGTCATTTACTTCGTCGTCGTTCTAGTCC

8581  TGTTCTGTGCCACATTCTTTTTCCATTCATGTTTCATACAACATTGTCTTCAGGAACATG
8581  ACAAGACACGGTGTAAGAAAAAGGTAAGTACAAAGTATGTTGTAACAGAAGTCCTTGTAC

8641  GTTCAGTGTTCTCCCAACCCACAAGAGACAAGCAAAATAAAAATGGACTAAATAATTTGC
8641  CAAGTCACAAGAGGGTTGGGTGTTCTCTGTTCGTTTTATTTTACCTGATTTATTAAACG

8701  TAGTTTTTAATTGGGCTCTTTGGCACAGACTCAAGTAGACCTGAACTCTCAGCTCTGTTG
8701  ATCAAAAATTAACCCGAGAAACCGTGTCTGAGTTCATCTGGACTTGAGAGTCGAGACAAC

8761  TGAACCCGCTTTGTGAACAAGGACAAAGCATGTCATCTGTCTCTGAGCCTCAGTCTCCCT
8761  ACTTGGGCGAAACACTTGTTCCTGTTTCGTACAGTAGACAGAGACTCGGAGTCAGAGGGA

8821  GTTAGTCAGTTACAGAGAGAGCGCTCCAAGTTTCCTTCCAAAGCTGTGCACTGGCTCGTT
8821  CAATCAGTCAATGTCTCTCTCGCGAGGTTCAAAGGAAGGTTTCGACACGTGACCGAGCAA

8881  TCTCCATCAGCTTTGGTTTTCCAGTTGGTCTCCAGCTCATCTTTGCCTCACAAGAGCGTG
8881  AGAGGTAGTCGAAACCAAAAGGTCAACCAGAGGTCGAGTAGAAACGGAGTGTTCTCGCAC

8941  CCCTGCTAGTTGCTGGTGCTGGGAGTGGGTGGGAAGGTGATTGGGCGCCAGCCCCTTCCT
8941  GGGACGATCAACGACCACGACCCTCACCCACCCTTCCACTAACCCGCGGTCGGGGAAGGA

9001  TCCCAGGCTTCACACACACACCCCAGGAAGCCCCGCGGTGCGTCGGACTCTCTGTCTAGA
9001  AGGGTCCGAAGTGTGTGTGTGGGGTCCTTCGGGGCGCCACGCAGCCTGAGAGACAGATCT
```

FIGURE 3 (cont'd)

```
 9061  CATCATCTGATTTTTATTTGCAAATGCAGGTTGCCAGGTGACAGGAACCTTTTATGCTTG
 9061  GTAGTAGACTAAAAATAAACGTTTACGTCCAACGGTCCACTGTCCTTGGAAAATACGAAC

 9121  TGCCCCACCTGAGTCCCAGGCTCCCGGAGTCCCCACAGAGCCAAGCATAAGGTCTGCCGA
 9121  ACGGGGTGGACTCAGGGTCCGAGGGCCTCAGGGGTGTCTCGGTTCGTATTCCAGACGGCT

 9181  AGCCTTGGCGTTCTCAGGTGAGTGCAGGGTTTGCTCCTGGAGGCACCGCAGGAGGCCAGC
 9181  TCGGAACCGCAAGAGTCCACTCACGTCCCAAACGAGGACCTCCGTGGCGTCCTCCGGTCG

 9241  CTCTGCTGCCAGCTCTGTCATCTTTGGGCAGATTCGATGGGACTTTAGACACTTGCTTTG
 9241  GAGACGACGGTCGAGACAGTAGAAACCCGTCTAAGCTACCCTGAAATCTGTGAACGAAAC

 9301  CTCCCTCTGGGGTCTGGAGTAGATGTAGACACATCCTGTGTGTGAGGTGACCAGGGTGAT
 9301  GAGGGAGACCCCAGACCTCATCTACATCTGTGTAGGACACACACTCCACTGGTCCCACTA

 9361  TTAGGAGCACCATTAGAAAACCTGACATCACTGCTTGTGGTTCTGCTGACCGTTTCAGCC
 9361  AATCCTCGTGGTAATCTTTTGGACTGTAGTGACGAACACCAAGACGACTGGCAAAGTCGG

 9421  ACTGGCTTGAATGGAGTCATTTTGGCTTCTTCACTGGCACCTCTCTGAATTTCTAGGAAT
 9421  TGACCGAACTTACCTCAGTAAAACCGAAGAAGTGACCGTGGAGAGACTTAAAGATCCTTA

 9481  GTGCCTTTACCTTTACCGAGGGCCCCTCTTCAGCCAACATTCTCACGATGTGGAATAATT
 9481  CACGGAAATGGAAATGGCTCCCGGGGAGAAGTCGGTTGTAAGAGTGCTACACCTTATTAA

 9541  GCTTGGAAGTGTAGAAGGGCTTCTCATTTTGAGAAGCTGATCATCCTTCCAGGTTGAGCC
 9541  CGAACCTTCACATCTTCCCGAAGAGTAAAACTCTTCGACTAGTAGGAAGGTCCAACTCGG

 9601  ACAAATAAGTCCTCCTCCTCTACTCCCTGGGGACATTAGTTCTGGTCCCTCATCTCTAAA
 9601  TGTTTATTCAGGAGGAGGAGATGAGGGACCCCTGTAATCAAGACCAGGGAGTAGAGATTT

 9661  ACATTGATGTGCCTAAGAGTAATACACATTTTGGTCTTCCTCTGAACTTTAATATAGCTT
 9661  TGTAACTACACGGATTCTCATTATGTGTAAAACCAGAAGGAGACTTGAAATTATATCGAA

 9721  GCAAACAAATATGGATTCAATCTGATTTTTAAAGTTTTATTTCTAAAAAAAAAAAAAAAA
 9721  CGTTTGTTTATACCTAAGTTAGACTAAAAATTTCAAAATAAAGATTTTTTTTTTTTTTTT

 9781  TCCCTGCACCATGGAGATCTTACCTACTATAAAGAAGGCACCTCTAGGCTTGGCAAGCAC
 9781  AGGGACGTGGTACCTCTAGAATGGATGATATTTCTTCCGTGGAGATCCGAACCGTTCGTG

 9841  ACGTGCTATATGTATATTTATTTTTCAGATAATATTTTGGATTGTTTTTAAATGGGATTT
 9841  TGCACGATATACATATAAATAAAAAGTCTATTATAAAACCTAACAAAAATTTACCCTAAA

 9901  GTTTTTATATTAAAACCAAATAGCTTAAGGTTTGGAATTCTCATCTTGCCCTCTGGCATC
 9901  CAAAAATATAATTTTGGTTTATCGAATTCCAAACCTTAAGAGTAGAACGGGAGACCGTAG

 9961  TTTAAAAATCAGTGATGAAAAATACTAACTAAATTCTGAAGGTTTCAGGGAGGCGAGAGT
 9961  AAATTTTTAGTCACTACTTTTTATGATTGATTTAAGACTTCCAAAGTCCCTCCGCTCTCA

10021  TGTGGCACTTTTGCTGCTCAGAGGGGAGCTGGAGTTTGACCTACCAGCTCTTTCTAGTTG
10021  ACACCGTGAAAACGACGAGTCTCCCCTCGACCTCAAACTGGATGGTCGAGAAAGATCAAC

10081  TGAATGAGGCTTGCACCTTTTTTTCTGAGGTGCTCGTGAGTAACTGAGGATGCCCTTTGG
10081  ACTTACTCCGAACGTGGAAAAAAAGACTCCACGAGCACTCATTGACTCCTACGGGAAACC

10141  GAGGAGGTGCTTCTGAGCAGGAAGGCTTGTGTTTGTTTTTAGAAACTTTCAAACCCTTGT
10141  CTCCTCCACGAAGACTCGTCCTTCCGAACACAAACAAAAATCTTTGAAAGTTTGGGAACA

10201  CTTGAACACCTAAGACTTGTGTGGGTGCCTGAAGAGTAGGAAATAAACAGCTATTTATAT
10201  GAACTTGTGGATTCTGAACACACCCACGGACTTCTCATCCTTTATTTGTCGATAAATATA

10261  CTCGGCAACCTCGTGATTTCTGATGACATTAAATGAAATGAAACCTGCCCCGAGAATCAC
10261  GAGCCGTTGGAGCACTAAAGACTACTGTAATTTACTTTACTTTGGACGGGGCTCTTAGTG

10321  CTCGAATGGCCAACACCCACTCTCTTTGGGGCGCACTGTCTGACTCGCTTTCAGACAGTG
10321  GAGCTTACCGGTTGTGGGTGAGAGAAACCCCGCGTGACAGACTGAGCGAAAGTCTGTCAC

10381  TGTTGAAGCAGAGAATTGAGACAGTGATGTGGGTTAAGTCTCAAAACCTGCTGTTGAGGA
10381  ACAACTTCGTCTCTTAACTCTGTCACTACACCCAATTCAGAGTTTTGGACGACAACTCCT
```

FIGURE 3 (cont'd)

```
10441  ATAAGATAGTTTTGTGGGTTTCTTTTTTTGGTATGCATGATACAGATTAATTATCAGCCA
10441  TATTCTATCAAAACACCCAAAGAAAAAAACCATACGTACTATGTCTAATTAATAGTCGGT

10501  TGAGCACATTTATTAGATAACTGTGATTCCCATTGATTTTGGGGTTCCATACAGTTACCA
10501  ACTCGTGTAAATAATCTATTGACACTAAGGGTAACTAAAACCCCAAGGTATGTCAATGGT

10561  GTGGCAGCTGCCTCCTGTCTGTGCCACTCCCCTGCGGTCTGGCTGAGGAGCTTGTGCAAA
10561  CACCGTCGACGGAGGACAGACACGGTGAGGGGACGCCAGACCGACTCCTCGAACACGTTT

10621  CCCAGCTTGGTTCTTTCAACTCATGGGCGATATTTTAAAGAGGCTTCTGAAATCAGCAAA
10621  GGGTCGAACCAAGAAAGTTGAGTACCCGCTATAAAATTTCTCCGAAGACTTTAGTCGTTT

10681  ACGTTTGTAAATGGATTCATCTTGTTTAAAGTTTCAGAGAGCGCAGTTTTAGAGTTCTTT
10681  TGCAAACATTTACCTAAGTAGAACAAATTTCAAAGTCTCTCGCGTCAAAATCTCAAGAAA

10741  TCAAACTTAAACACAAAGCCCAGTGGGAATTTCAGCAGACTTTACATGAACTTTTAGAGT
10741  AGTTTGAATTTGTGTTTCGGGTCACCCTTAAAGTCGTCTGAAATGTACTTGAAAATCTCA

10801  ATTTTTCTCATTAATTTCTTTCTTTTCACAAGAATTAATTTTGTGATAAAAACACCACAT
10801  TAAAAAGAGTAATTAAAGAAAGAAAAGTGTTCTTAATTAAAACACTATTTTTGTGGTGTA

10861  ATCCAGTGTGGAAAACTAATTTAAAAATCAGAAAAACAAAAATGGAAAATAAAAATATCA
10861  TAGGTCACACCTTTTGATTAAATTTTTAGTCTTTTTGTTTTTACCTTTTATTTTTATAGT

10921  CATTCTACCACCCAGAAATCACCATTAAGGCCTCAGTGCTTACCCTTTAATGTGTGTATG
10921  GTAAGATGGTGGGTCTTTAGTGGTAATTCCGGAGTCACGAATGGGAAATTACACACATAC

10981  TGAGTGTGTTTATATATAAATTATATAGACATAAATAATTAAGCACCTAATTCTATATAT
10981  ACTCACACAAATATATATTTAATATATCTGTATTTATTAATTCGTGGATTAAGATATATA

11041  AATTTATATATGTACACACACACACACACACACACACATTTGCTTTTTACCTAAATGATG
11041  TTAAATATATACATGTGTGTGTGTGTGTGTGTGTGTGTAAACGAAAAATGGATTTACTAC

11101  TCATGTTATTATTTTGGGGACGTTTCAGGTATGTGGACCGACTTGCTTCTTTCTGAAACA
11101  AGTACAATAATAAAACCCCTGCAAAGTCCATACACCTGGCTGAACGAAGAAAGACTTTGT

11161  CCGTTATATTTTTTGCTAGTATAAATTAAACTTAGGGTTCATTTCCTCCAAAGTAAAAAT
11161  GGCAATATAAAAAACGATCATATTTAATTTGAATCCCAAGTAAAGGAGGTTTCATTTTTA

11221  TGATTTGCTTAAAGATGAATTTTCATTCAGTGACACATGACTTAAATGACATCACCACTA
11221  ACTAAACGAATTTCTACTTAAAAGTAAGTCACTGTGTACTGAATTTACTGTAGTGGTGAT

11281  TCATCATCATAATTCACAAATAAGACTTAGCAAGTTCTTAATAGTGATGCTCTATAATAG
11281  AGTAGTAGTATTAAGTGTTTATTCTGAATCGTTCAAGAATTATCACTACGAGATATTATC

11341  GAATTCAACTTACATTTCCAGAAAATTAATGAAAGATACACTTGTTCCTTAAAAAAAGAG
11341  CTTAAGTTGAATGTAAAGGTCTTTTAATTACTTTCTATGTGAACAAGGAATTTTTTTCTC

11401  TCGGTGGCCACATACGTTTGTAATATGCTGGTTTAGACAGATTAAATGGGTTTCTTTCCA
11401  AGCCACCGGTGTATGCAAACATTATACGACCAAATCTGTCTAATTTACCCAAAGAAAGGT

11461  CGGGACTTCTCCAAGTGTTGACCGAGCTCATGCTGTCTGTGGATCTCTGGGGGAGACAAG
11461  GCCCTGAAGAGGTTCACAACTGGCTCGAGTACGACAGACACCTAGAGACCCCCTCTGTTC

11521  GGCACGCTGTGTTTCCCAGGCCTCCTTGACGCGGGAACCTTTTGTTCCCCCACGGGACAT
11521  CCGTGCGACACAAAGGGTCCGGAGGAACTGCGCCCTTGGAAAACAAGGGGGTGCCCTGTA

11581  GTGGGACACTAGAGTTCCACCACAGGTGCTTGGCTCTGTGGAGTCGTTGGCCTCGAGGTG
11581  CACCCTGTGATCTCAAGGTGGTGTCCACGAACCGAGACACCTCAGCAACCGGAGCTCCAC

11641  GTGTCCTTGGCCCCCAGCACTGACTCCGGAGCTCTGTGTTTGCAGACTGCCGGCTGCACA
11641  CACAGGAACCGGGGGTCGTGACTGAGGCCTCGAGACACAAACGTCTGACGGCCGACGTGT

11701  TCTGCCTGTACTACCGGGAAATCCTCGTGAAGGAGCTGACCACGTCCAGCCCCGAGGGCT
11701  AGACGGACATGATGGCCCTTTAGGAGCACTTCCTCGACTGGTGCAGGTCGGGGCTCCCGA
```

FIGURE 3 (cont'd)

```
11761  GCCGGATCTCCCATGGACATACGTATGACGCCAGCAACCTGGACCAGGTCCTGTTCCCCT
11761  CGGCCTAGAGGGTACCTGTATGCATACTGCGGTCGTTGGACCTGGTCCAGGACAAGGGGA

11821  ACCCAGAGGACAATGGCCAGAGGAAAAACATTGAGAAGCTGCTGAGCCACCTGGAGAGGG
11821  TGGGTCTCCTGTTACCGGTCTCCTTTTTGTAACTCTTCGACGACTCGGTGGACCTCTCCC

11881  GCGTGGTCCTCTGGATGGCCCCCGACGGGCTCTATGCGAAAAGACTGTGCCAGAGCAGGA
11881  CGCACCAGGAGACCTACCGGGGGCTGCCCGAGATACGCTTTTCTGACACGGTCTCGTCCT

11941  TCTACTGGGACGGGCCCCTGGCGCTGTGCAACGACCGGCCCAACAAACTGGAGAGAGACC
11941  AGATGACCCTGCCCGGGGACCGCGACACGTTGCTGGCCGGGTTGTTTGACCTCTCTCTGG

12001  AGACCTGCAAGCTCTTTGACACACAGCAGTTCTTGTCAGGTAAGGCACCTCGTTATCTGT
12001  TCTGGACGTTCGAGAAACTGTGTGTCGTCAAGAACAGTCCATTCCGTGGAGCAATAGACA

12061  TAGAATGGAGGTGGTGATGGCCTGCCTGCCACAGGGGTCAGAAACCACAGGGTCCCTCCC
12061  ATCTTACCTCCACCACTACCGGACGGACGGTGTCCCCAGTCTTTGGTGTCCCAGGGAGGG

12121  ACCCCAGGCTGAGGTCTTCCTCCTGTTGACTTCGGCGCCCACTGGGCTTGGGGCTTGACT
12121  TGGGGTCCGACTCCAGAAGGAGGACAACTGAAGCCGCGGGTGACCCGAACCCCGAACTGA

12181  CCAGTAGAGATCTTCTGTCTGGCTCTGTTGTGGGCAGAGAGTCCGGATCAGATGGTCCAG
12181  GGTCATCTCTAGAAGACAGACCGAGACAACACCCGTCTCTCAGGCCTAGTCTACCAGGTC

12241  GTGGACCATGGCTCAGGCTGTTCTCTTGCAATTCTGGTTTTCAGCAGTCGCCTCTTAAAA
12241  CACCTGGTACCGAGTCCGACAAGAGAACGTTAAGACCAAAAGTCGTCAGCGGAGAATTTT

12301  CAGTTCCATTTCATTGCACTATGATCTCGACAAATACGTCTGCTGATGTGGCGGCTCTTT
12301  GTCAAGGTAAAGTAACGTGATACTAGAGCTGTTTATGCAGACGACTACACCGCCGAGAAA

12361  TCAGGTTCCCAGAAAAATCATAAAGGTTGGAAGCATATTGTTAACATTCTTGGTTGTCAG
12361  AGTCCAAGGGTCTTTTTAGTATTTCCAACCTTCGTATAACAATTGTAAGAACCAACAGTC

12421  GTGTGTCACCCTCATGCAGATGCTGGCCCACGAGTGAGACGGAGGGTGGAGTGGACTGTA
12421  CACACAGTGGGAGTACGTCTACGACCGGGTGCTCACTCTGCCTCCCACCTCACCTGACAT

12481  CTGGTTTTGCACAGGAGAAGTTATAGGAGAGTGCTATAAGTTGGCTCTGTGATATGAACA
12481  GACCAAAACGTGTCCTCTTCAATATCCTCTCACGATATTCAACCGAGACACTATACTTGT

12541  TTTCTAAGCATATTTTTCAGAACCAGTGAGTTTTGTTTTTCTACCTTGGATTTCATGACA
12541  AAAGATTCGTATAAAAAGTCTTGGTCACTCAAAACAAAAAGATGGAACCTAAAGTACTGT

12601  ATTCAGCGAACCTCTGGCTGGCTTCTGTGAAGCCCCTCGCCCTGTCATGCTGGGCTTCCT
12601  TAAGTCGCTTGGAGACCGACCGAAGACACTTCGGGGAGCGGGACAGTACGACCCGAAGGA

12661  CACTGGCTTTGTGGGGCTGTCCAGGTGGAGGGAGCCTCCCGCGTGGGAGGGAGCTGCTCT
12661  GTGACCGAAACACCCCGACAGGTCCACCTCCCTCGGAGGGCGCACCCTCCCTCGACGAGA

12721  ACCCGCCACTCGCCACCCCACCGCCAACTCTTCCTCTGCCTTTTCCCAGTCCTGTCCCTT
12721  TGGGCGGTGAGCGGTGGGGTGGCGGTTGAGAAGGAGACGGAAAAGGGTCAGGACAGGGAA

12781  CCTGAGCCCCTTCAGAAACTCCTGTGTCTTTGATTTCTCATGTGCCTCCCCTTTTCTTGA
12781  GGACTCGGGGAAGTCTTTGAGGACACAGAAACTAAAGAGTACACGGAGGGGAAAAGAACT

12841  ATGTGCTTCTTCATTCAACTTATTTAAAATAAAAATATGACACGTCTGTGTGCTCTGGCA
12841  TACACGAAGAAGTAAGTTGAATAAATTTTATTTTTATACTGTGCAGACACACGAGACCGT

12901  CTGTTGACATGTGCTCGTACTTCTGGGTGTAAGTGACCCAAGTTTTTGTAAGCTCATCCA
12901  GACAACTGTACACGAGCATGAAGACCCACATTCACTGGGTTCAAAAACATTCGAGTAGGT

12961  GATTTTCTTTTGGTGCCCAAAGCAAATATGTCCCCAGATAACTAAGCTTCAGGCCAGGTG
12961  CTAAAAGAAAACCACGGGTTTCGTTTATACAGGGGTCTATTGATTCGAAGTCCGGTCCAC

13021  TGGTGGCCCATGCCTGTAATCCCAGCACTTTAGGAGGCTGAGGTGGGCAGATCACTTGAG
13021  ACCACCGGGTACGGACATTAGGGTCGTGAAATCCTCCGACTCCACCCGTCTAGTGAACTC

13081  GCCAGGAGTTCAAGACCAGCCTGGCCAACATGGCAAAACCCCGTCTCTACTAAAGATACA
13081  CGGTCCTCAAGTTCTGGTCGGACCGGTTGTACCGTTTTGGGGCAGAGATGATTTCTATGT
```

FIGURE 3 (cont'd)

```
13141  AAAATTAGCTGGGCGTGATGGTGGGCGCCTGTAATCCCAGCTACTCAGGAGGATGAGGCA
13141  TTTTAATCGACCCGCACTACCACCCGCGGACATTAGGGTCGATGAGTCCTCCTACTCCGT

13201  CAAGAATCGCTTGAACCTGTGGGAGGCGGAGGTTGCAGTGAGCTGAGATCGCACCACGGT
13201  GTTCTTAGCGAACTTGGACACCCTCCGCCTCCAACGTCACTCGACTCTAGCGTGGTGCCA

13261  GCTGCAGTTTGCAGGACAGAGCGAGACTCCATTGCCCCCCGCACCACAAAAAAATAACTA
13261  CGACGTCAAACGTCCTGTCTCGCTCTGAGGTAACGGGGGGCGTGGTGTTTTTTTATTGAT

13321  AGCTCACACAGAGGCTGACCATAGGAAGAGGGAGAACTGTGCTGGTTCCCCGAGGCGGCG
13321  TCGAGTGTGTCTCCGACTGGTATCCTTCTCCCTCTTGACACGACCAAGGGGCTCCGCCGC

13381  GCAAGTGAGGAAGTCACCTGGGGCTGGACACCTCTCATTCTAAGGGCAGAACCATCCCCA
13381  CGTTCACTCCTTCAGTGGACCCCGACCTGTGGAGAGTAAGATTCCCGTCTTGGTAGGGGT

13441  AGCCCTGGCCAGGGCGGTCCCATCCTTCTTGGGGTTGGGCATGGTGTTCATGGGGGACCA
13441  TCGGGACCGGTCCCGCCAGGGTAGGAAGAACCCCAACCCGTACCACAAGTACCCCCTGGT

13501  AGGAAATTGAAGGGCAACTGGGGGGACCCCCCCTCCCCTGTCAGCATCTCCTGTGTCTGC
13501  TCCTTTAACTTCCCGTTGACCCCCCTGGGGGGGAGGGGACAGTCGTAGAGGACACAGACG

13561  CTTGGGATCTGGAAGAGCTCACATGGGCCAGGAAGGCTAAGGCCCACTGGGCCTGGATTT
13561  GAACCCTAGACCTTCTCGAGTGTACCCGGTCCTTCCGATTCCGGGTGACCCGGACCTAAA

13621  CTGAAGACTCTGGACCTTGGTGCCAGTGGATTCAGAAGATACCACAAGGTGAGGGCTTTC
13621  GACTTCTGAGACCTGGAACCACGGTCACCTAAGTCTTCTATGGTGTTCCACTCCCGAAAG

13681  TAATGAAAGTGTCACAAGGAACTGGCACTGGCTTTTCTGAGTGCCTCTTGCTGGGCGTTT
13681  ATTACTTTCACAGTGTTCCTTGACCGTGACCGAAAAGACTCACGGAGAACGACCCGCAAA

13741  TTAGGGAGGTAGAGCCCCCGTGGTGACTAAGCTGGAAGGTGCACATTGAGTCACAGGTGC
13741  AATCCCTCCATCTCGGGGGCACCACTGATTCGACCTTCCACGTGTAACTCAGTGTCCACG

13801  ACTGCGTGAGAGACAGCACAGGCAGGGGGTGGACGCCTGTGAGTGTCCTGGGGCTGTAGA
13801  TGACGCACTCTCTGTCGTGTCCGTCCCCCACCTGCGGACACTCACAGGACCCCGACATCT

13861  CTTGCGCGACTAGAACTTACTATTAATCTGTGAGCAAGAGCTGGTCTTGGCTTTCATTCC
13861  GAACGCGCTGATCTTGAATGATAATTAGACACTCGTTCTCGACCAGAACCGAAAGTAAGG

13921  TTCCTCTGTAACCAAGGGCTGTGCTCTTTGCCCACTGCAGCCTCTCACCTCAAAGTGTTC
13921  AAGGAGACATTGGTTCCCGACACGAGAAACGGGTGACGTCGGAGAGTGGAGTTTCACAAG

13981  ACTGAGGTTCAAGGAGGATGCTGTGAACGTAAAAACTGCAGCTGTGCCAACCAGCTTCTG
13981  TGACTCCAAGTTCCTCCTACGACACTTGCATTTTTGACGTCGACACGGTTGGTCGAAGAC

14041  CATAATTAAGGATTCCCACCAAACACTCTCATGTTATCTAGGGTTGGAGCCATGCTTTCT
14041  GTATTAATTCCTAAGGGTGGTTTGTGAGAGTACAATAGATCCCAACCTCGGTACGAAAGA

14101  CAGAGAATTGTGCCAACTCGCCATTCTGATTAGCCTGTGTAGGTGTAGTCTCAGATCACG
14101  GTCTCTTAACACGGTTGAGCGGTAAGACTAATCGGACACATCCACATCAGAGTCTAGTGC

14161  GCAGTGTGAATGTATTTTACAGATTCTGACTAAGTCATTTGGGTTTGATTTGAATTCTGG
14161  CGTCACACTTACATAAAATGTCTAAGACTGATTCAGTAAACCCAAACTAAACTTAAGACC

14221  AAAAAAAAAAAAGCAGGAAGTCAAATAGTCCTGTAAGTTAGCTAGAAACTTCTGTTCAGT
14221  TTTTTTTTTTTTCGTCCTTCAGTTTATCAGGACATTCAATCGATCTTTGAAGACAAGTCA

14281  TGAAGAGAACAGTGGAGATCTTTGATATCTTCCTATTCAGGTCTGCACAGCACTAGGGAC
14281  ACTTCTCTTGTCACCTCTAGAAACTATAGAAGGATAAGTCCAGACGTGTCGTGATCCCTG

14341  AGCCCCCAGGGCCCGGCCCGAGGGTGTGTATCTGATAAGGACGCGTCTGTTCTGCAGAAG
14341  TCGGGGGTCCCGGGCCGGGCTCCCACACATAGACTATTCCTGCGCAGACAAGACGTCTTC

14401  CTGTAGCGGCTCCTGTGTCAACTCGTTTCTTTTGCGGGCGTGACATTTTATTGTAGCTAC
14401  GACATCGCCGAGGACACAGTTGAGCAAAGAAAACGCCCGCACTGTAAAATAACATCGATG
```

FIGURE 3 (cont'd)

```
14461  AGGCAGAAGATTTGTCTTGTGTACAGGGGAGGGAGCATGGGCTAAAGTCAGGGGATGGGC
14461  TCCGTCTTCTAAACAGAACACATGTCCCCTCCCTCGTACCCGATTTCAGTCCCCTACCCG

14521  ACTTGTCTTTCCAACAATGTTCCAGTCCGTTTTGTATCTTTTGGGTAGGTGTGGTGGCTT
14521  TGAACAGAAAGGTTGTTACAAGGTCAGGCAAAACATAGAAAACCCATCCACACCACCGAA

14581  CAGGATGTACATGTGTATATCCACAAGCGGGAGGCCAGGGGGGCTGCCCCACTCCTGTTC
14581  GTCCTACATGTACACATATAGGTGTTCGCCCTCCGGTCCCCCCGACGGGGTGAGGACAAG

14641  TCAAGTCAAGTTACTGTTCCATCCCTGGAGACAGGCAAGAAGTCTACTCAGAGTTACACA
14641  AGTTCAGTTCAATGACAAGGTAGGGACCTCTGTCCGTTCTTCAGATGAGTCTCAATGTGT

14701  GTTCAGGCATAGTGACAGTGGGACTCAGCTGCCGATGGCTGGTGCTCAGTCACCACGTTG
14701  CAAGTCCGTATCACTGTCACCCTGAGTCGACGGCTACCGACCACGAGTCAGTGGTGCAAC

14761  ACGTTACATATTTTCTGTGGTGCCTGAGTTACGTGGATGTCCGCAGTAGCACAGATACTG
14761  TGCAATGTATAAAAGACACCACGGACTCAATGCACCTACAGGCGTCATCGTGTCTATGAC

14821  GATTATGTGGGCTCTGTAGTGAGGGGATGAGGTTTTAAAATGATCCTGGGATGTCGAAAC
14821  CTAATACACCCGAGACATCACTCCCCTACTCCAAAATTTTACTAGGACCCTACAGCTTTG

14881  ATTCTGATTTTTTAAATGAAAACTTGTCCATGGCATAAATTGGTCTTTTGCCATTGCGTG
14881  TAAGACTAAAAAATTTACTTTTGAACAGGTACCGTATTTAACCAGAAAACGGTAACGCAC

14941  ACTACGTTTCTTGTTCTTTATCTCCCGACTTCACTGTGGTCTACTACCTTTTGAATCTTC
14941  TGATGCAAAGAACAAGAAATAGAGGGCTGAAGTGACACCAGATGATGGAAAACTTAGAAG

15001  GTGGGTCCAATGCTGCAAAGCAGTGTTCTTTAGCTGTCGACTAGTTCCTCTTGAAGAATC
15001  CACCCAGGTTACGACGTTTCGTCACAAGAAATCGACAGCTGATCAAGGAGAACTTCTTAG

15061  GAGGGAGACCGAGGGCCCTGGGGGGAAAGCACCCAAAGGAATGCATACATGCTATTTTGT
15061  CTCCCTCTGGCTCCCGGGACCCCCCTTTCGTGGGTTTCCTTACGTATGTACGATAAAACA

15121  ATCTGAGATGTTCACATCAAGAGCCCCACTCAGCGGAGTAAGAGTGCTCATTCCTCTTGC
15121  TAGACTCTACAAGTGTAGTTCTCGGGGTGAGTCGCCTCATTCTCACGAGTAAGGAGAACG

15181  AGTGTTCAGAATCACAGTAAGCTCTTGCTTCCTGTTTAACCTGGTGTGTTCGGTGATGAG
15181  TCACAAGTCTTAGTGTCATTCGAGAACGAAGGACAAATTGGACCACACAAGCCACTACTC

15241  GGTTTCTGAACATGGTCTCTTTCTTGTGGGCTTCTACAGAGCTGCAAGCGTTTGCTCACC
15241  CCAAAGACTTGTACCAGAGAAAGAACACCCGAAGATGTCTCGACGTTCGCAAACGAGTGG

15301  ACGGCCGCTCCCTGCCAAGATTCCAGGTGACTCTATGCTTTGGAGAGGAGTTTCCAGACC
15301  TGCCGGCGAGGGACGGTTCTAAGGTCCACTGAGATACGAAACCTCTCCTCAAAGGTCTGG

15361  CTCAGAGGCAAAGAAAGCTCATCACAGCTCACGTGAGTCCTCAGTTACACTCCTACCATA
15361  GAGTCTCCGTTTCTTTCGAGTAGTGTCGAGTGCACTCAGGAGTCAATGTGAGGATGGTAT

15421  GTGGCTTCCTGTTCTTTGTAAAGGCCAGAGTTTCATTTAGAAAAGTCCCAAATGAAAAGT
15421  CACCGAAGGACAAGAAACATTTCCGGTCTCAAAGTAAATCTTTTCAGGGTTTACTTTTCA

15481  AAATGTCAAATGACCTGGAAAAATAAGCGTAACCCTTAAACTAGTGAGGGAGGAAGCATG
15481  TTTACAGTTTACTGGACCTTTTTATTCGCATTGGGAATTTGATCACTCCCTCCTTCGTAC

15541  GTCCAACGAGCAGCACAGTCTGAGGACTCACGTGCTCCTCCCAGACTTGAGATCTGCTCA
15541  CAGGTTGCTCGTCGTGTCAGACTCCTGAGTGCACGAGGAGGGTCTGAACTCTAGACGAGT

15601  TCAAAGAACCAGGGAGGACAGCCTCAAAGGAGGGTGCCTCTTTCCCCATCTTTTTTATTT
15601  AGTTTCTTGGTCCCTCCTGTCGGAGTTTCCTCCCACGGAGAAAGGGGTAGAAAAAATAAA

15661  TTCAGGAAAGTTGTCGTATTTCCATTTTATAGGGATTGAAAAAGATTGATGGTTAAAGTT
15661  AAGTCCTTTCAACAGCATAAAGGTAAAATATCCCTAACTTTTTCTAACTACCAATTTCAA

15721  GCCCTTTAAAATTCTCCAGGTTAAGATTGCTGTAAGAATGCTATCTAGCTAGTGGATCTT
15721  CGGGAAATTTTAAGAGGTCCAATTCTAACGACATTCTTACGATAGATCGATCACCTAGAA

15781  CATTCAATGGAAAAGCTTTTCCCAAATGAGAAACTATCCATTCCCTGGAGGCATTTTGTA
15781  GTAAGTTACCTTTTCGAAAAGGGTTTACTCTTTGATAGGTAAGGGACCTCCGTAAAACAT
```

FIGURE 3 (cont'd)

```
15841  GGTCTCTGCAGCTGTGTCCTAGCACCTCTTTTATTTCTGGAATTTTAGAAATTATTTAAT
15841  CCAGAGACGTCGACACAGGATCGTGGAGAAAATAAAGACCTTAAAATCTTTAATAAATTA

15901  TATTGGGCATATGATATTTGAAAGAGGCTGAATCTTTCAAGGAATTCAAGCAAATCAGAC
15901  ATAACCCGTATACTATAAACTTTCTCCGACTTAGAAAGTTCCTTAAGTTCGTTTAGTCTG

15961  CCTCTCGTAATGTTCTCTAGCATAGCTCAAGATGGGTTCAACTGTGGCAAGTAACACTAA
15961  GGAGAGCATTACAAGAGATCGTATCGAGTTCTACCCAAGTTGACACCGTTCATTGTGATT

16021  AAGGGTAGGGTTAGGATTAGGGTTAGAGTTTTCTTTTTTTATGAGAAAAGGTTTCCATAGA
16021  TTCCCATCCCAATCCTAATCCCAATCTCAAAAGAAAAAATACTCTTTTCCAAAGGTATCT

16081  TAACATGAAGCATTAATTTGCCATGGGAAACAGAATCTTGAACCCTTAAGACTTTTGAAT
16081  ATTGTACTTCGTAATTAAACGGTACCCTTTGTCTTAGAACTTGGGAATTCTGAAAACTTA

16141  TTGAAAACTTCACATTGCTACATACTTGCAAGCATTATTAACGGGCTTGCATGCTGAGCA
16141  AACTTTTGAAGTGTAACGATGTATGAACGTTCGTAATAATTGCCCGAACGTACGACTCGT

16201  ATACCAGGTAAAAGTTACACCCCCTAACATCAAAGTTCTCATCATGTTCTAGCTCTGATG
16201  TATGGTCCATTTTCAATGTGGGGGATTGTAGTTTCAAGAGTAGTACAAGATCGAGACTAC

16261  CTATTTCTCATTGAGAACTGAAATGAAACATGGTTTAATCTTGAACATACAACCCCCCTT
16261  GATAAAGAGTAACTCTTGACTTTACTTTGTACCAAATTAGAACTTGTATGTTGGGGGGAA

16321  CTTGATTTTAAAAACAAAGACCAGCCAACCAACCAGTATATAATCCCATAGACTCAGGAG
16321  GAACTAAAATTTTTGTTTCTGGTCGGTTGGTTGGTCATATATTAGGGTATCTGAGTCCTC

16381  TTTTCTCATGAGTTCTCTCCAGTGATTGACTTAATAATTGATATGATTGTTGAAATAATA
16381  AAAAGAGTACTCAAGAGAGGTCACTAACTGAATTATTAACTATACTAACAACTTTATTAT

16441  ACTGATACATTGTTAACAACAGGAATTGCTAAATGACTAAATAAACTTGGCATTGATAAG
16441  TGACTATGTAACAATTGTTGTCCTTAACGATTTACTGATTTATTTGAACCGTAACTATTC

16501  CAGCATTTAAGAAGTTGATGATCGGCCGGGCACAGTGGCTCACGCCTGTAATCCCGGCAC
16501  GTCGTAAATTCTTCAACTACTAGCCGGCCCGTGTCACCGAGTGCGGACATTAGGGCCGTG

16561  TTTGGGAGGCTGAGGCGGGCAGATCATGTGGTCAGGAGTTCGAGACCAGCCTGACCAACA
16561  AAACCCTCCGACTCCGCCCGTCTAGTACACCAGTCCTCAAGCTCTGGTCGGACTGGTTGT

16621  TGGTGAAACCCCCTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGGTGCGCACCT
16621  ACCACTTTGGGGGAGAGATGATTTTTATGTTTTTAATCGGCCCGCACCACCACGCGTGGA

16681  GTAATCCCAGCTACTCGGGAGGCTGAAGCAAGGGAGTCACTTGAACCCAGGAGTCAGAGC
16681  CATTAGGGTCGATGAGCCCTCCGACTTCGTTCCCTCAGTGAACTTGGGTCCTCAGTCTCG

16741  TTGCAGTGAGCCGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGTGAGACTCTGT
16741  AACGTCACTCGGCTCTAACGTGGTGACGTGAGGTCGGACCCACTGTCTCACTCTGAGACA

16801  CTCAAAAAAAAAACAAAAAAAGAAGTTGATGATCAGCTCAGCTATCACATCAATCCAGTT
16801  GAGTTTTTTTTTTGTTTTTTTCTTCAACTACTAGTCGAGTCGATAGTGTAGTTAGGTCAA

16861  TTAAAGTTGACATGGATGCAAGTCACTGTCCTCTAAGTCAGGGAAAGATGGATCCCCAGA
16861  AATTTCAACTGTACCTACGTTCAGTGACAGGAGATTCAGTCCCTTTCTACCTAGGGGTCT

16921  TGACCCCCTCATGCTGAAGACGAGGCTGCGTGGTCCATAAAATGAATAAACACACGTGTA
16921  ACTGGGGGAGTACGACTTCTGCTCCGACGCACCAGGTATTTTACTTATTTGTGTGCACAT

16981  CACCTATGAGTTCCACTTTTAAAAATTATTAATTTAATTCATTTCAGGTACTTAGGAGTT
16981  GTGGATACTCAAGGTGAAAATTTTTAATAATTAAATTAAGTAAAGTCCATGAATCCTCAA

17041  TAGCTTAAGTCGTCATATATAAAATACAGTCTCTAATATCATGATTGATGGAGAGCATTC
17041  ATCGAATTCAGCAGTATATATTTTATGTCAGAGATTATAGTACTAACTACCTCTCGTAAG

17101  AGCTTGCCTTAGATGCTGTAAATTCAGGAAAAGCTGAACATTAAGGCGTTTTTGTTTTTT
17101  TCGAACGGAATCTACGACATTTAAGTCCTTTTCGACTTGTAATTCCGCAAAAACAAAAAA
```

FIGURE 3 (cont'd)

```
17161  GGTGAGTGTGATCCACGCTAATAACCAGACAGTATAAATTTGAGGCCAGTTAGCTTCTCT
17161  CCACTCACACTAGGTGCGATTATTGGTCTGTCATATTTAAACTCCGGTCAATCGAAGAGA

17221  TTTTCCACAATAGTAGTTTTGTTGTTGTTTAACTTTACTTTTTATTTCTTTTACCTCTGG
17221  AAAAGGTGTTATCATCAAAACAACAACAAATTGAAATGAAAAATAAAGAAAATGGAGACC

17281  TTGAGAACTATACAGATTACCTAGAAATCAAATCATTCAGGATGAGCTCCTTGTTTCTTT
17281  AACTCTTGATATGTCTAATGGATCTTTAGTTTAGTAAGTCCTACTCGAGGAACAAAGAAA

17341  CAAACTGCAGTTGTGGAAAAACAAAATCATTGGCCTAAATTACCCTGCAGATTCCCTCAG
17341  GTTTGACGTCAACACCTTTTTGTTTTAGTAACCGGATTTAATGGGACGTCTAAGGGAGTC

17401  ACTTCCTTTTAATTTGATCACTTTACTTGATTTGATGCAGTTTTAGTTAAATGTACATTT
17401  TGAAGGAAAATTAAACTAGTGAAATGAACTAAACTACGTCAAAATCAATTTACATGTAAA

17461  TAAGTGGCAGTCGATTTGAAAGTAACATCTTTAACCCAGAATTAAGTCACTTTGGCTGTT
17461  ATTCACCGTCAGCTAAACTTTCATTGTAGAAATTGGGTCTTAATTCAGTGAAACCGACAA

17521  TTTTTACATGTTGATCTATGGAAGGACTAACCAAAAAATTGCTTCTTTAGATGGCTCCAA
17521  AAAAATGTACAACTAGATACCTTCCTGATTGGTTTTTTAACGAAGAAATCTACCGAGGTT

17581  ATATGAAATGTTTTGATGTGTTCTAGGATGTAACTTTGGGCTTTACGTTACTGTCTCCTT
17581  TATACTTTACAAAACTACACAAGATCCTACATTGAAACCCGAAATGCAATGACAGAGGAA

17641  AGAATCTGAGTGCTGTTTAATAGTGAGCCAGTTGCAGGATATCTCAGTAATGTCTTTTTA
17641  TCTTAGACTCACGACAAATTATCACTCGGTCAACGTCCTATAGAGTCATTACAGAAAAAT

17701  AAATCTCTTATTAAAGGTAGAACCTCTGCTAGCCAGACAACTATATTATTTTGCTCAACA
17701  TTTAGAGAATAATTTCCATCTTGGAGACGATCGGTCTGTTGATATAATAAAACGAGTTGT

17761  AAACAGTGGACATTTCCTGAGGGGCTACGATTTACCAGAACACATCAGCAATCCAGAAGA
17761  TTTGTCACCTGTAAAGGACTCCCCGATGCTAAATGGTCTTGTGTAGTCGTTAGGTCTTCT

17821  TTACCACAGATCTATCCGCCATTCCTCTATTCAAGAATGAAAAATGTCAAGATGAGTGGT
17821  AATGGTGTCTAGATAGGCGGTAAGGAGATAAGTTCTTACTTTTTACAGTTCTACTCACCA

17881  TTTCTTTTTCCTTTTTTTTTTTTTTTTTTGATACGGGGATACGGGGTCTTGCTCTGTCT
17881  AAAGAAAAAGGAAAAAAAAAAAAAAAAAACTATGCCCCTATGCCCCAGAACGAGACAGA

17941  CCCAGGCTGGAGTGCAGTGACACAATCTCAGCTCACTGTGACCTCCGCCTCCTGGGTTCA
17941  GGGTCCGACCTCACGTCACTGTGTTAGAGTCGAGTGACACTGGAGGCGGAGGACCCAAGT

18001  AGAGACTCTCCTGCCTCAGCCTCCCTGGTAGCTGGGATTACAGGTGTGAGCCACTGCACC
18001  TCTCTGAGAGGACGGAGTCGGAGGGACCATCGACCCTAATGTCCACACTCGGTGACGTGG

18061  CACCCAAGACAAGTGATTTTCATTGTAAATATTTGACTTTAGTGAAAGCGTCCAATTGAC
18061  GTGGGTTCTGTTCACTAAAAGTAACATTTATAAACTGAAATCACTTTCGCAGGTTAACTG

18121  TGCCCTCTTACTGTTTTGAGGAATTCAGAAGTGGAGATTTCAGTTCAGCGGTTGAGGAGA
18121  ACGGGAGAATGACAAAACTCCTTAAGTCTTCACCTCTAAAGTCAAGTCGCCAACTCCTCT

18181  ATTGCGGCGAGACAAGCATGGAAAATCAGTGACATCTGATTGGCAGATGAGCTTATTTCA
18181  TAACGCCGCTCTGTTCGTACCTTTTAGTCACTGTAGACTAACCGTCTACTCGAATAAAGT

18241  AAAGGAAGGGTGGCTTTGCATTTCTTGTGTTCTGTAGACTGCCATCATTGATGATCACTG
18241  TTTCCTTCCCACCGAAACGTAAAGAACACAAGACATCTGACGGTAGTAACTACTAGTGAC

18301  TGAAAATTGACCAAGTGATGTGTTTACATTTACTGAAATGCGCTCTTTAATTTGTTGTAG
18301  ACTTTTAACTGGTTCACTACACAAATGTAAATGACTTTACGCGAGAAATTAAACAACATC

18361  ATTAGGTCTTGCTGGAAGACAGAGAAAACTTGCCTTTCAGTATTGACACTGACTAGAGTG
18361  TAATCCAGAACGACCTTCTGTCTCTTTTGAACGGAAAGTCATAACTGTGACTGATCTCAC

18421  ATGACTGCTTGTAGGTATGTCTGTGCCATTTCTCAGGGAAGTAAGATGTAAATTGAAGAA
18421  TACTGACGAACATCCATACAGACACGGTAAAGAGTCCCTTCATTCTACATTTAACTTCTT

18481  GCCTCACACGTAAAAGAAATGTATTAATGTATGTAGGAGCTGCAGTTCTTGTGGAAGACA
18481  CGGAGTGTGCATTTTCTTTACATAATTACATACATCCTCGACGTCAAGAACACCTTCTGT
```

FIGURE 3 (cont'd)

```
18541  CTTGCTGAGTGAAGGAAATGAATCTTTGACTGAAGCCGTGCCTGTAGCCTTGGGGAGGCC
18541  GAACGACTCACTTCCTTTACTTAGAAACTGACTTCGGCACGGACATCGGAACCCCTCCGG

18601  CATCCCCCACCTGCCAGCGGTTTCCTGGTGTGGGTCCCTCTGCCCCGCCCTCCTTCCCAT
18601  GTAGGGGGTGGACGGTCGCCAAAGGACCACACCCAGGGAGACGGGGCGGGAGGAAGGGTA

18661  TGGCTTTCTCTCCTTGGCCTTTCCTGGAAGCCAGTTAGTAAACTTCCTATTTTCTTGAGT
18661  ACCGAAAGAGAGGAACCGGAAAGGACCTTCGGTCAATCATTTGAAGGATAAAAGAACTCA

18721  CAAAAAACATGAGCGCTACTCTTGGATGGGACATTTTTGTCTGTCCTACAATCTAGTAAT
18721  GTTTTTTGTACTCGCGATGAGAACCTACCCTGTAAAAACAGACAGGATGTTAGATCATTA

18781  GTCTAAGTAATGGTTAAGTTTTCTTGTTTCTGCATCTTTTTGACCCTCATTCTTTAGAGA
18781  CAGATTCATTACCAATTCAAAAGAACAAAGACGTAGAAAAACTGGGAGTAAGAAATCTCT

18841  TGCTAAAATTCTTCGCATAAAGAAGAAGAAATTAAGGAACATAAATCTTAATACTTGAAC
18841  ACGATTTTAAGAAGCGTATTTCTTCTTCTTTAATTCCTTGTATTTAGAATTATGAACTTG

18901  TGTTGCCCTTCTGTCCAAGTACTTAACTATCTGTTCCCTTCCTCTGTGCCACGCTCCTCT
18901  ACAACGGGAAGACAGGTTCATGAATTGATAGACAAGGGAAGGAGACACGGTGCGAGGAGA

18961  GTTTGTTTGGCTGTCCAGCGATCAGCCATGGCGACACTAAAGGAGGAGGAGCCGGGGACT
18961  CAAACAAACCGACAGGTCGCTAGTCGGTACCGCTGTGATTTCCTCCTCCTCGGCCCCTGA

19021  CCCAGGCTGGAGAGCACTGCCAGGACCCACCACTGGAAGCAGGATGGAGCTGACTACGGA
19021  GGGTCCGACCTCTCGTGACGGTCCTGGGTGGTGACCTTCGTCCTACCTCGACTGATGCCT

19081  ACTGCACACTCAGTGGGCTGTTTCTGCTTATTTCATCTGTTCTATGCTTCCTCGTGCCAA
19081  TGACGTGTGAGTCACCCGACAAAGACGAATAAAGTAGACAAGATACGAAGGAGCACGGTT

19141  TTATAGTTTGACAGGGCCTTAAAATTACTTGGCTTTTTCCAAATGCTTCTATTTATAGAA
19141  AATATCAAACTGTCCCGGAATTTTAATGAACCGAAAAAGGTTTACGAAGATAAATATCTT

19201  TCCCAAAGACCTCCACTTGCTTAAGTATACCTATCACTTACATTTTTGTGGTTTTGAGAA
19201  AGGGTTTCTGGAGGTGAACGAATTCATATGGATAGTGAATGTAAAAACACCAAAACTCTT

19261  AGTACAGCAGTAGACTGGGGCGTCACCTCCAGGCCGTTTCTCATACTACAGGATATTTAC
19261  TCATGTCGTCATCTGACCCCGCAGTGGAGGTCCGGCAAAGAGTATGATGTCCTATAAATG

19321  TATTACTCCCAGGATCAGCAGAAGATTGCGTAGCTCTCAAATGTGTGTTCCTGCTTTTCT
19321  ATAATGAGGGTCCTAGTCGTCTTCTAACGCATCGAGAGTTTACACACAAGGACGAAAAGA

19381  AATGGATATTTTAAATTCATTCAACAAGCACCTAGTAAGTGCCTGCTGTATCCCTACATT
19381  TTACCTATAAAATTTAAGTAAGTTGTTCGTGGATCATTCACGGACGACATAGGGATGTAA

19441  ACACAGTTCAGCCTTTATCAAGCTTAGTGAGCAGTGAGCACTGAAACATTATTTTTTAAT
19441  TGTGTCAAGTCGGAAATAGTTCGAATCACTCGTCACTCGTGACTTTGTAATAAAAAATTA

19501  GTTTAAAAAGTTTCTAATATTAAAGTCAGAATATTAATACAATTAATATTAATATTAACT
19501  CAAATTTTTCAAAGATTATAATTTCAGTCTTATAATTATGTTAATTATAATTATAATTGA

19561  ACAGAAAAGACAAACAGTAGAGAACAGCAAAAAAATAAAAAGGATCTCCTTTTTTCCCAG
19561  TGTCTTTTCTGTTTGTCATCTCTTGTCGTTTTTTTATTTTTCCTAGAGGAAAAAAGGGTC

19621  CCCAAATTCTCCTCTCTAAAAGTGTCCACAAGAAGGGGTGTTTATTCTTCCAACACATTT
19621  GGGTTTAAGAGGAGAGATTTTCACAGGTGTTCTTCCCCACAAATAAGAAGGTTGTGTAAA

19681  CACTTTTCTGTAAATATACATAAACTTAAAAAGAAAACCTCATGGAGTCATCTTGCACAC
19681  GTGAAAAGACATTTATATGTATTTGAATTTTTCTTTTGGAGTACCTCAGTAGAACGTGTG

19741  ACTTTCATGCAGTGCTCTTTGTAGCTAACAGTGAAGATTTACCTCGTTCTGCTCAGAGGC
19741  TGAAAGTACGTCACGAGAAACATCGATTGTCACTTCTAAATGGAGCAAGACGAGTCTCCG

19801  CTTGCTGTGGAGCTCCACTGCCATGTACCCAGTAGGGTTTGACATTTCATTAGCCATGCA
19801  GAACGACACCTCGAGGTGACGGTACATGGGTCATCCCAAACTGTAAAGTAATCGGTACGT
```

FIGURE 3 (cont'd)

```
19861  ACATGGATATGTATTGGGCAGCAGACTGTGTTTCGTGAACTGCAGTGATGTATACATCTT
19861  TGTACCTATACATAACCCGTCGTCTGACACAAAGCACTTGACGTCACTACATATGTAGAA

19921  ATAGATGCAAAGTATTTTGGGGTATATTATCCTAAGGGAAGATAAAGATGATATTAAGAA
19921  TATCTACGTTTCATAAAACCCCATATAATAGGATTCCCTTCTATTTCTACTATAATTCTT

19981  CTGCTGTTTCACGGGGCCCTTACCTGTGACCCTCTTTGCTGAAGAATATTTAACCCCACA
19981  GACGACAAAGTGCCCCGGGAATGGACACTGGGAGAAACGACTTCTTATAAATTGGGGTGT

20041  CAGCACTTCAAAGAAGCTGTCTTGGAAGTCTGTCTCAGGAGCACCCTGTCTTCTTAATTC
20041  GTCGTGAAGTTTCTTCGACAGAACCTTCAGACAGAGTCCTCGTGGGACAGAAGAATTAAG

20101  TCCAAGCGGATGCTCCATTTCAATTGCTTTGTGACTTCTTCTTCTTTGTTTTTTTAAATA
20101  AGGTTCGCCTACGAGGTAAAGTTAACGAAACACTGAAGAAGAAGAAACAAAAAAATTTAT

20161  TTATGCTGCTTTAACAGTGGAGCTGAATTTTCTGGAAAATGCTTCTTGGCTGGGGCCACT
20161  AATACGACGAAATTGTCACCTCGACTTAAAAGACCTTTTACGAAGAACCGACCCCGGTGA

20221  ACCTCCTTTCCTATCTTTACATCTATGTGTATGTTGACTTTTTAAAATTCTGAGTGATCC
20221  TGGAGGAAAGGATAGAAATGTAGATACACATACAACTGAAAAATTTTAAGACTCACTAGG

20281  AGGGTATGACCTAGGGAATGAACTAGCTATGAAATACTCAGGGTTAGGAATCCTAGCACT
20281  TCCCATACTGGATCCCTTACTTGATCGATACTTTATGAGTCCCAATCCTTAGGATCGTGA

20341  TGTCTCAGGACTCTGAAAAGGAACGGCTTCCTCATTCCTTGTCTTGATAAAGTGGAATTG
20341  ACAGAGTCCTGAGACTTTTCCTTGCCGAAGGAGTAAGGAACAGAACTATTTCACCTTAAC

20401  GCAAACTAGAATTTAGTTTGTACTCAGTGGACAGTGCTGTTGAAGATTTGAGGACTTGTT
20401  CGTTTGATCTTAAATCAAACATGAGTCACCTGTCACGACAACTTCTAAACTCCTGAACAA

20461  AAAGAGCACTGGGTCATATGGAAAAAATGTATGTGTCTCCCAGGTGCATTTCTTGGTTTA
20461  TTTCTCGTGACCCAGTATACCTTTTTTACATACACAGAGGGTCCACGTAAAGAACCAAAT

20521  TGTCTTGTTCTTGAGATTTTGTATATTTAGGAAAACCTCAAGCAGTAATTAATATCTCCT
20521  ACAGAACAAGAACTCTAAAACATATAAATCCTTTTGGAGTTCGTCATTAATTATAGAGGA

20581  GGAACACTATAGAGAACCAAGTGACCGACTCATTTACAACTGAAACCTAGGAAGCCCCTG
20581  CCTTGTGATATCTCTTGGTTCACTGGCTGAGTAAATGTTGACTTTGGATCCTTCGGGGAC

20641  AGTCCTGAGCGAAAACAGGAGAGTTAGTCGCCCTACAGAAAACCCAGCTAGACTATTGGG
20641  TCAGGACTCGCTTTTGTCCTCTCAATCAGCGGGATGTCTTTTGGGTCGATCTGATAACCC

20701  TATGAACTAAAAAGAGACTGTGCCATGGTGAGAAAAATGTAAAATCCTACAGTGAAATGA
20701  ATACTTGATTTTTCTCTGACACGGTACCACTCTTTTTACATTTTAGGATGTCACTTTACT

20761  GCAGCCCTTACAGTATTGTTACCACCAAGGGCAGGTAGGTATTAGTGTTTGAAAAAGCTG
20761  CGTCGGGAATGTCATAACAATGGTGGTTCCCGTCCATCCATAATCACAAACTTTTTCGAC

20821  GTCTTTGAGCGAGGGCATAAATACAGCTAGCCCCAGGGGTGGAACAACTCTGGGAGTCTT
20821  CAGAAACTCGCTCCCGTATTTATGTCGATCGGGGTCCCCACCTTGTTGAGACCCTCAGAA

20881  GGGTACTCGCACCTCTTGGCTTTGTTGATGCTCCGCCAGGAAGGCCACTTGTGTGTGCGT
20881  CCCATGAGCGTGGAGAACCGAAACAACTACGAGGCGGTCCTTCCGGTGAACACACACGCA

20941  GTCAGTTACTTTTTTAGTAACAATTCAGATCCAGTGTAAACTTCCGTTCATTGCTCTCCA
20941  CAGTCAATGAAAAAATCATTGTTAAGTCTAGGTCACATTTGAAGGCAAGTAACGAGAGGT

21001  GTCACATGCCCCCACTTCCCCACAGGTGAAAGTTTTTCTGAAAGTGTTGGGATTGGTTAA
21001  CAGTGTACGGGGGTGAAGGGGTGTCCACTTTCAAAAAGACTTTCACAACCCTAACCAATT

21061  GGTCTTTATTTGTATTACGTATCTCCCGAAGTCCTCTGTGGCCAGCTGCATCTGTCTGAA
21061  CCAGAAATAAACATAATGCATAGAGGGCTTCAGGAGACACCGGTCGACGTAGACAGACTT

21121  TGGTGCGTGAAGGCTCTCAGACCTTACACACCATTTTGTAAGTTATGTTTTACATGCCCC
21121  ACCACGCACTTCCGAGAGTCTGGAATGTGTGGTAAAACATTCAATACAAAATGTACGGGG

21181  GTTTTTGAGACTGATCTCGATGCAGGTGGATCTCCTTGAGATCCTGATAGCCTGTTACAG
21181  CAAAAACTCTGACTAGAGCTACGTCCACCTAGAGGAACTCTAGGACTATCGGACAATGTC
```

FIGURE 3 (cont'd)

```
21241  GAATGAAGTAAAGGTCAGTTTTTTTTTGTATTGATTTTCACAGCTTTGAGGAACATGCAT
21241  CTTACTTCATTTCCAGTCAAAAAAAAACATAACTAAAAGTGTCGAAACTCCTTGTACGTA

21301  AAGAAATGTAGCTGAAGTAGAGGGGACGTGAGAGAAGGGCCAGGCCGGCAGGCCAACCCT
21301  TTCTTTACATCGACTTCATCTCCCCTGCACTCTCTTCCCGGTCCGGCCGTCCGGTTGGGA

21361  CCTCCAATGGAAATTCCCGTGTTGCTTCAAACTGAGACAGATGGGACTTAACAGGCAATG
21361  GGAGGTTACCTTTAAGGGCACAACGAAGTTTGACTCTGTCTACCCTGAATTGTCCGTTAC

21421  GGGTCCACTTCCCCCTCTTCAGCATCCCCCGTACCCCACTTTCTGCTGAAAGAACTGCCA
21421  CCCAGGTGAAGGGGGAGAAGTCGTAGGGGGCATGGGGTGAAAGACGACTTTCTTGACGGT

21481  GCAGGTAGGACCCCAGAGGCCCCCAAATGAAAGCTTGAATTTCCCCTACTGGCTCTGCGT
21481  CGTCCATCCTGGGGTCTCCGGGGGTTTACTTTCGAACTTAAAGGGGATGACCGAGACGCA

21541  TTTGCTGAGATCTGTAGGAAAGGATGCTTCACAAACTGAGGTAGATAATGCTATGCTGTC
21541  AAACGACTCTAGACATCCTTTCCTACGAAGTGTTTGACTCCATCTATTACGATACGACAG

21601  GTTGGTATACATCATGAATTTTTATGTAAATTGCTCTGCAAAGCAAATTGATATGTTTGA
21601  CAACCATATGTAGTACTTAAAAATACATTTAACGAGACGTTTCGTTTAACTATACAAACT

21661  TAAATTTATGTTTTTAGGTAAATAAAAACTTTTAAAAATTTGTTA
21661  ATTTAAATACAAAAATCCATTTATTTTTGAAAATTTTTAAACAAT
```

FIGURE 4

>hg19_refGene_NM_001178092 range=chr12:24962958-25102393

The BCAT1 gene resides on the minus strand, thus the top strand in the below is the representative 5'- to 3'- sequence of BCAT1

YELLOW: Sub-region of interest towards which a bisulphite conversion and methylation specific assay IS GIVEN AS AN EXAMPLE

Green: Sub-region of interest where a methylation specific restriction enzyme (HHaI/HBAII) amplicon is given as an example

```
   1  AGTAGGGAGGTGGGCAGGAGCCAGTGATGACGGAATGGCAATCACATTTGACCTCTGATC
   1  TCATCCCTCCACCCGTCCTCGGTCACTACTGCCTTACCGTTAGTGTAAACTGGAGACTAG

  61  TGTTTATTTCCTCCTCCTTGACGTCTCCATATAAATGTTACACGGGCATCCCCACACTCG
  61  ACAAATAAAGGAGGAGGAACTGCAGAGGTATATTTACAATGTGCCCGTAGGGGTGTGAGC

 121  GATACGCACCCACAGTGGCTGATTCGGGGGTAACCGTGTCATTTGCTTGCAACACTGGCA
 121  CTATGCGTGGGTGTCACCGACTAAGCCCCCATTGGCACAGTAAACGAACGTTGTGACCGT

 181  CCTCTGCCCTGCACCCCGGGAGTGAGCAGTGAGTGAGGCTCGGGTCTGGGCGCTGGCTCC
 181  GGAGACGGGACGTGGGGCCCTCACTCGTCACTCACTCCGAGCCCAGACCCGCGACCGAGG

 241  GAATCTTCGGGCTGGGAGAGACTCCACCATCTGGGGGCGGCCTGGGGGAGCAGCCTTAGT
 241  CTTAGAAGCCCGACCCTCTCTGAGGTGGTAGACCCCCGCCGGACCCCCTCGTCGGAATCA

 301  GTCTTCCTGCTGATGCAATCCGCTAGGTCGCGAGTCTCCGCCGCGAGAGGGCCGGTCTGC
 301  CAGAAGGACGACTACGTTAGGCGATCCAGCGCTCAGAGGCGGCGCTCTCCCGGCCAGACG

 361  AATCCAGCCCGCCACGTGTACTCGCCGCCGCCTCGGGCACTGCCCCAGGTCTTGCTGCAG
 361  TTAGGTCGGGCGGTGCACATGAGCGGCGGCGGAGCCCGTGACGGGGTCCAGAACGACGTC

 421  CCGGGACCGCGCTCTGCAGCCGCAGACCCGGTCCACACGGCCAGGGGCTACGACCCTTGG
 421  GGCCCTGGCGCGAGACGTCGGCGTCTGGGCCAGGTGTGCCGGTCCCCGATGCTGGGAACC

 481  GATCTGCCCTCCGCTCAGCTCGAGCTTCCCTCGTGGCCGACGGAACAATGAAGGTAACTA
 481  CTAGACGGGAGGCGAGTCGAGCTCGAAGGGAGCACCGGCTGCCTTGTTACTTCCATTGAT

 541  CTTATGGTTTTGTCCGTGTTTTACAAAAATGTGTGCGTGAATCGAACCGGCGATTTCTCC
 541  GAATACCAAAACAGGCACAAAATGTTTTTACACACGCACTTAGCTTGGCCGCTAAAGAGG

 601  AAGAAACATAGTTGGCAGGGAGGGGAGGAAGGCGAGACAACCATGGCTTATATCCCCCGC
 601  TTCTTTGTATCAACCGTCCCTCCCCTCCTTCCGCTCTGTTGGTACCGAATATAGGGGGCG

 661  AAACGTCTCAGTATCTTCTTTATCAATCGTAGTTTGCGGGGACCGTGCATTCTGTTCAGA
 661  TTTGCAGAGTCATAGAAGAAATAGTTAGCATCAAACGCCCCTGGCACGTAAGACAAGTCT

 721  TTTCGGTTTAACCTCCACTCGCAGGACGTGCCTTCTCGGACTTTTTCACATTCGCTTTTG
 721  AAAGCCAAATTGGAGGTGAGCGTCCTGCACGGAAGAGCCTGAAAAAGTGTAAGCGAAAAC

 781  GGAACGGAGGTGAAAGTCTGCTACAGCTCCCTCCCCTGCTTGTGAAGTTTGGAAAGGAAG
 781  CCTTGCCTCCACTTTCAGACGATGTCGAGGGAGGGGACGAACACTTCAAACCTTTCCTTC

 841  TGAGGGCTTCTCTCAGTTTCTCCTATGCACAGGAGGTGGGGAATTTTGGAGAGGAGGTCT
 841  ACTCCCGAAGAGAGTCAAAGAGGATACGTGTCCTCCACCCCTTAAAACCTCTCCTCCAGA

 901  GGGGATGTCCCGGGCTGTAAATGCGCTTTCCTGCAGCGTGTGTTCGTGATGCAGGAGGGA
 901  CCCCTACAGGGCCCGACATTTACGCGAAAGGACGTCGCACACAAGCACTACGTCCTCCCT

 961  GCGGCTGGAAGAGTTGACCCGGGTGGAGAGGGGAGGGAGAACGATCTTTTCACTGTTAAA
 961  CGCCGACCTTCTCAACTGGGCCCACCTCTCCCCTCCCTCTTGCTAGAAAAGTGACAATTT

1021  AGCAGAAGGCCCCTCTTATTTTGTTTCTTGTGGAAATTAAAGCTTCTAGGAGTTACAAAT
1021  TCGTCTTCCGGGGAGAATAAAACAAAGAACACCTTTAATTTCGAAGATCCTCAATGTTTA

1081  GATAGACGCCTAAGGGCATTCATTTGTTAAACCAGTGAGCACCCTAATGGCTCTGATCTT
1081  CTATCTGCGGATTCCCGTAAGTAAACAATTTGGTCACTCGTGGGATTACCGAGACTAGAA
```

FIGURE 4 (cont'd)

```
1141  GCTAATCAAGGATAATTAGCCTAGCTGCCTATTATTTCTGACTATTTAGGTGTAGGGATG
1141  CGATTAGTTCCTATTAATCGGATCGACGGATAATAAAGACTGATAAATCCACATCCCTAC

1201  TACACCGTTGAGTGTTTTGTTTTCCTAGTTTCACACACTGGAAAACTGCAATCGCATATA
1201  ATGTGGCAACTCACAAAACAAAAGGATCAAAGTGTGTGACCTTTTGACGTTAGCGTATAT

1261  AAACCCGAAGAGCAACTTTTTCCAAGGAAAGTGAAGGATGGGCAGATATTCTTGCCATCT
1261  TTTGGGCTTCTCGTTGAAAAAGGTTCCTTTCACTTCCTACCCGTCTATAAGAACGGTAGA

1321  AACATTTTGGAATTGAGATGACCATAATGCAATCTGAACTTCTGGTAGGAATAAAGCCTC
1321  TTGTAAAACCTTAACTCTACTGGTATTACGTTAGACTTGAAGACCATCCTTATTTCGGAG

1381  GTTTCAGAGTTTCTAAGGAGGGAGATGAAGCGTGGTTCTCAACTGCTCTTATCTACAACC
1381  CAAAGTCTCAAAGATTCCTCCCTCTACTTCGCACCAAGAGTTGACGAGAATAGATGTTGG

1441  CATCCCAGACCATAGTTTCACTTAGAAGAAATAATATGTAAATAGCGCCTTTGTTGCCAA
1441  GTAGGGTCTGGTATCAAAGTGAATCTTCTTTATTATACATTTATCGCGGAAACAACGGTT

1501  ACAAAAACTTTCACTACAATTAAACAGTAACCGAGTGGAAACTATGCGTTTTTGTATTCA
1501  TGTTTTTGAAAGTGATGTTAATTTGTCATTGGCTCACCTTTGATACGCAAAAACATAAGT

1561  ACAGAGATCACACAACTTTTCTTCTATAGTTAGAGGTTCTATTTCTGGCTTCCTCATGCT
1561  TGTCTCTAGTGTGTTGAAAAGAAGATATCAATCTCCAAGATAAAGACCGAAGGAGTACGA

1621  TGTTTCATTTCAGATGAAAAACACACTTTAGTTGATTATAGCTTTAGTAAGTAAAAGGAG
1621  ACAAAGTAAAGTCTACTTTTTGTGTGAAATCAACTAATATCGAAATCATTCATTTTCCTC

1681  AAAACTAAGAATGAACAGTCTTGCTTTTCCAGCTTTTTCCTGTCCATGTGCAGTTGTGGT
1681  TTTTGATTCTTACTTGTCAGAACGAAAAGGTCGAAAAAGGACAGGTACACGTCAACACCA

1741  TAATTACGCGTTTTTGTAGGTCCAAAAACCCCTGTCCCTTCCACAAGATACTTTTATAAC
1741  ATTAATGCGCAAAAACATCCAGGTTTTTGGGGACAGGGAAGGTGTTCTATGAAAATATTG

1801  ATGGGGAACTAGGTAAGATATCTACAAGTAGTCTCTTTTCTGTTTCCTGATATTTCCTAA
1801  TACCCCTTGATCCATTCTATAGATGTTCATCAGAGAAAAGACAAAGGACTATAAAGGATT

1861  GATTAAGAAAGAAATGATTGTATTCTAGGTATGATCGAAGAAGTATGGAAGGTTCCATTA
1861  CTAATTCTTTCTTTACTAACATAAGATCCATACTAGCTTCTTCATACCTTCCAAGGTAAT

1921  AGTTAGTTTTATCACAATATTTCAAACATCTTTCCACTTATCTCTTAAATATGTATTGCT
1921  TCAATCAAAATAGTGTTATAAAGTTTGTAGAAAGGTGAATAGAGAATTTATACATAACGA

1981  ATCACAAGCTTGATTTTTTAAAAAAATGTTATTTGCACATGGCATCTTGATTGCCTTTGT
1981  TAGTGTTCGAACTAAAAAATTTTTTTACAATAAACGTGTACCGTAGAACTAACGGAAACA

2041  CCTCAGCAAATTCTCCGGGCTTAAGTCATGCATGAGCTTGTATTTTAAGACAGACACATG
2041  GGAGTCGTTTAAGAGGCCCGAATTCAGTACGTACTCGAACATAAAATTCTGTCTGTGTAC

2101  CAGAATGGGCAACTTGGATAAGAAACATAGGATATACACTGAGTAGCATGTGTACAAAAA
2101  GTCTTACCCGTTGAACCTATTCTTTGTATCCTATATGTGACTCATCGTACACATGTTTTT

2161  TGGTAAGTGAGTATTATGACTTAATAAATTTAGATGGGGATTTAAAAATGATACACATGT
2161  ACCATTCACTCATAATACTGAATTATTTAAATCTACCCCTAAATTTTTACTATGTGTACA

2221  AAATGGGATTGGAGGCAAGGGAAGTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG
2221  TTTACCCTAACCTCCGTTCCCTTCAGGCACACACACACACACACACACACACACACAC

2281  TGTGTGTGTGTGGAGGGAAGACGGGAGGCCTAGTGGGGAGAGAGGTTTTACATTCCAGAC
2281  ACACACACACACCTCCCTTCTGCCCTCCGGATCACCCCTCTCTCCAAAATGTAAGGTCTG

2341  ACAGGATGTGGCCCCTCAGGCTAGGCTCAGAGATGTAAAGCTAGTTAAGCTGGCCTAACA
2341  TGTCCTACACCGGGGAGTCCGATCCGAGTCTCTACATTTCGATCAATTCGACCGGATTGT

2401  AAAGGGTGGTGGTTAGAAAAAGGAAAAAAAGAGGTGGGGGTTGGGGAGGAAAAGAATTTA
2401  TTTCCCACCACCAATCTTTTTCCTTTTTTTCTCCACCCCCAACCCCTCCTTTTCTTAAAT
```

FIGURE 4 (cont'd)

```
2461  ATGCTGGCCAATATTTTCATCTGATAAGTTTCAGTAATGAGTAATTTCACTGGATTTCAG
2461  TACGACCGGTTATAAAAGTAGACTATTCAAAGTCATTACTCATTAAAGTGACCTAAAGTC

2521  AACTACAGAGAGAAAAGGGGAAAAATATTTTTCCTTTTTTACGCAAGAATTTCTCTTGCT
2521  TTGATGTCTCTCTTTTCCCCTTTTTATAAAAAGGAAAAAATGCGTTCTTAAAGAGAACGA

2581  TTTCAGCTCATTAGAACCAGATCTAATTATTTGAAATGAGTTTAAAACCTTATTATGAAA
2581  AAAGTCGAGTAATCTTGGTCTAGATTAATAAACTTTACTCAAATTTTGGAATAATACTTT

2641  GGAAGTTGAAAACATTCGATCACAATTGAGTTTTGTACTATGGACTAAAACCAAGATGTG
2641  CCTTCAACTTTTGTAAGCTAGTGTTAACTCAAAACATGATACCTGATTTTGGTTCTACAC

2701  CTTTCTGTCAGTAGCTGTGCAACTCCCTGAAAAAAGAAAAATAGCGCAGAGCGATCTGAG
2701  GAAAGACAGTCATCGACACGTTGAGGGACTTTTTTCTTTTTATCGCGTCTCGCTAGACTC

2761  CTTTGTGAAGGCTACACAATTTATCAGGCCCAGAGAGATAGGAGTGTGAGACTTCAGTCA
2761  GAAACACTTCCGATGTGTTAAATAGTCCGGGTCTCTCTATCCTCACACTCTGAAGTCAGT

2821  CGTCCCCCCGTACGCATGCGCAGGAGTAATTATTTAAAGACATTTTGTTCCTGACTAACT
2821  GCAGGGGGGCATGCGTACGCGTCCTCATTAATAAATTTCTGTAAAACAAGGACTGATTGA

2881  GCCTCACTCATTATGTTCCTGGAATTTGTGATACAAAAAAACAATGTTTAGCCTATCAAT
2881  CGGAGTGAGTAATACAAGGACCTTAAACACTATGTTTTTTTGTTACAAATCGGATAGTTA

2941  AGTTTATATAGTTTATGTTATTTTAATATAAATTCTTCTCAACTTAGAAACTGCCTCTTT
2941  TCAAATATATCAAATACAATAAAATTATATTTAAGAAGAGTTGAATCTTTGACGGAGAAA

3001  TTTCTTTTAAAAACCCACTTGTAACTGCTGTATATATTCAGGGCAACTTGATTTGGTGTT
3001  AAAGAAAATTTTTGGGTGAACATTGACGACATATATAAGTCCCGTTGAACTAAACCACAA

3061  TCGTTGGTTGCAGTCCTCAAATTTGACCTCAATATACTCTCTACTTATATAAATTTTGTC
3061  AGCAACCAACGTCAGGAGTTTAAACTGGAGTTATATGAGAGATGAATATATTTAAAACAG

3121  TCAGTTTTGTCCGTAGGCTGACATTCCAATTGTTTTCCTAATTATATATTTTGTTCCTTA
3121  AGTCAAAACAGGCATCCGACTGTAAGGTTAACAAAAGGATTAATATATAAAACAAGGAAT

3181  TTTCTCATGTCTAGAAATAAAGCTAGGGAAGTAAACAGCTGGAAAACTGATAGTGGTGTT
3181  AAAGAGTACAGATCTTTATTTCGATCCCTTCATTTGTCGACCTTTTGACTATCACCACAA

3241  ATTAACCTAACTGAGAGGGCTGGGCTTATGAGTAAATTATTCCTGTGCATAAATGAAAGA
3241  TAATTGGATTGACTCTCCCGACCCGAATACTCATTTAATAAGGACACGTATTTACTTTCT

3301  TTGTTTCTAAATGGCAGCTATGTTGAGCCTTTGTATTAGTTACATCTTTTGGGGTTTCAG
3301  AACAAAGATTTACCGTCGATACAACTCGGAAACATAATCAATGTAGAAAACCCCAAAGTC

3361  GTTGCCAGCTTCTTTAAATGGCCTCTTCTCAACTGGTATAGGAAGTTGGAAAATAAAGAA
3361  CAACGGTCGAAGAAATTTACCGGAGAAGAGTTGACCATATCCTTCAACCTTTTATTTCTT

3421  CTTTCTTTATTTCCCTTTAGTGTTGGTAACCACAGTTTCTTTAAATAACAGATGCCCACC
3421  GAAAGAAATAAAGGGAAATCACAACCATTGGTGTCAAAGAAATTTATTGTCTACGGGTGG

3481  CCAGGGAAACCCAGTGGCATGATCAGCCTTATGCTCAGGAAGGTTTGGTTGCTATACTGG
3481  GGTCCCTTTGGGTCACCGTACTAGTCGGAATACGAGTCCTTCCAAACCAACGATATGACC

3541  AGAGTGAGAATCTGGGACCTTCCATGATAGACATAATTGTGCCAGCTATCCTCTACTTTC
3541  TCTCACTCTTAGACCCTGGAAGGTACTATCTGTATTAACACGGTCGATAGGAGATGAAAG

3601  TATATGATGCTATCATTCTTTTTTCTTTTCTTTTTTTTGACAGAATCTCACTGTGTTGTT
3601  ATATACTACGATAGTAAGAAAAAAGAAAAGAAAAAAAACTGTCTTAGAGTGACACAACAA

3661  CAGGCTGGAGTGCCGCAGAGTGATCTCTACTCACTGCAACCTCTGCCTCCTAGGTTCAAG
3661  GTCCGACCTCACGGCGTCTCACTAGAGATGAGTGACGTTGGAGACGGAGGATCCAAGTTC

3721  CAATTCTTGTGCCTCAGCCTCCTGAGAAGCTGGGATTACCACCATGCCCAGCTAATTTTC
3721  GTTAAGAACACGGAGTCGGAGGACTCTTCGACCCTAATGGTGGTACGGGTCGATTAAAAG
```

FIGURE 4 (cont'd)

```
3781  GTATTTTTTTGTAAAGACGGAGTTTTGCCATGTTGGCCAGGCTGGTCTGGAACTCCTGGC
3781  CATAAAAAAACATTTCTGCCTCAAAACGGTACAACCGGTCCGACCAGACCTTGAGGACCG

3841  CTCAAGCGATCTGCCCATCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGTGCTATCA
3841  GAGTTCGCTAGACGGGTAGAGCCGGAGGGTTTCACGACCCTAATGTCCGTACACGATAGT

3901  TTCTTTTTGAGCACAGGAGGAAACTATGGCTGATCTGCTCATTTTACTGTCAAAAATATA
3901  AAGAAAAACTCGTGTCCTCCTTTGATACCGACTAGACGAGTAAAATGACAGTTTTTATAT

3961  GGCAGCTTTTGTATTGTAACATTATCCCCTTAACAGCTGTCCAACTGGTCTATACATTTC
3961  CCGTCGAAAACATAACATTGTAATAGGGGAATTGTCGACAGGTTGACCAGATATGTAAAG

4021  ACCTAACATATGGTAAATGTTCCTACTGTTGGCCTTTCCTGAGCCTCTGTCCATTCAACT
4021  TGGATTGTATACCATTTACAAGGATGACAACCGGAAAGGACTCGGAGACAGGTAAGTTGA

4081  AAGGCCACTGCAATGGGATTGTATGCAGTTCAGCCAAGAAATAAATGTGAGGTGCAACTC
4081  TTCCGGTGACGTTACCCTAACATACGTCAAGTCGGTTCTTTATTTACACTCCACGTTGAG

4141  ACATTTTTTAAATGTTTCAAAATTAAGCATTTAGGTGCTTTTATTCATAAACATGACTTT
4141  TGTAAAAAATTTACAAAGTTTTAATTCGTAAATCCACGAAAATAAGTATTTGTACTGAAA

4201  GCAAATATTGGTAAACAGTTTCATCATAAGACCTTGTCAGTGACTCACAATGTCAATGGT
4201  CGTTTATAACCATTTGTCAAAGTAGTATTCTGGAACAGTCACTGAGTGTTACAGTTACCA

4261  GAACGCAAAGGAGAATCTGTCTTTTCTGGAACACCAGGTTCAGGTATCAGTTTCAACTAT
4261  CTTGCGTTTCCTCTTAGACAGAAAAGACCTTGTGGTCCAAGTCCATAGTCAAAGTTGATA

4321  GCTTATAATTTGTGCTCTAAAGTCCACAAAGCTAAATTATCTTTCCCCTAAAAATTGAGT
4321  CGAATATTAAACACGAGATTTCAGGTGTTTCGATTTAATAGAAAGGGGATTTTTAACTCA

4381  ACTCATATCCTTAAAATAGCTTGGAAATTAGATTACCTTATGAATACAGCATAGCATGTC
4381  TGAGTATAGGAATTTTATCGAACCTTTAATCTAATGGAATACTTATGTCGTATCGTACAG

4441  ACATAAACGTTTTTCTACTGTTATACCTTTTTGGTGGGAAAGACTATTTTGCAAAACATT
4441  TGTATTTGCAAAAAGATGACAATATGGAAAAACCACCCTTTCTGATAAAACGTTTTGTAA

4501  TTTCTTGTGTTCAGTAATTAATGGGAAAAACAAACTCTAAAAGAAAGAGACAGTTATGTC
4501  AAAGAACACAAGTCATTAATTACCCTTTTTGTTTGAGATTTTCTTTCTCTGTCAATACAG

4561  CCAAAAATACAGGAAAATGGAAGAATGGAAACAGGGAAGACCAATAAAGAAAAAAATAGA
4561  GGTTTTTATGTCCTTTTACCTTCTTACCTTTGTCCCTTCTGGTTATTTCTTTTTTTATCT

4621  GGAAATTACAAGTAATAAACTGAAGTCATACTGAAAAACATATGGCTCTTTCTTATTGAA
4621  CCTTTAATGTTCATTATTTGACTTCAGTATGACTTTTTGTATACCGAGAAAGAATAACTT

4681  AGTTGTGAAATACTTGCAAAACTCAGCTAGTTCCATTAAGAATAGTCTTTAATCTAAAAA
4681  TCAACACTTTATGAACGTTTTGAGTCGATCAAGGTAATTCTTATCAGAAATTAGATTTTT

4741  GAACAAATTCATTTGTCAGTTCTTTATTTCGGTGTGTAGATTTGAATGTTTCATCAATTG
4741  CTTGTTTAAGTAAACAGTCAAGAAATAAAGCCACACATCTAAACTTACAAAGTAGTTAAC

4801  CATTTACGGTTTATCCCTCAGTGTGAGAACTGTTTTTGCTGTTTTGACTATAAATGGATA
4801  GTAAATGCCAAATAGGGAGTCACACTCTTGACAAAAACGACAAAACTGATATTTACCTAT

4861  TCTGCTTTAGGAAGACTTAGGAAGCCTGGGACATAGGATTACAGATCAAATTCTATTTGA
4861  AGACGAAATCCTTCTGAATCCTTCGGACCCTGTATCCTAATGTCTAGTTTAAGATAAACT

4921  CTCCATAGAAGTTGAGACGGAAGGAGCACCTGTTGGGATATTCTAGACCAAACATTCCAC
4921  GAGGTATCTTCAACTCTGCCTTCCTCGTGGACAACCCTATAAGATCTGGTTTGTAAGGTG

4981  AGCTGTCTCTTAACGTGACTGACAAAGAGACACAGTTGAGACGAAGTTTGGGTATATCCT
4981  TCGACAGAGAATTGCACTGACTGTTTCTCTGTGTCAACTCTGCTTCAAACCCATATAGGA

5041  GGGATGAGGCCAAGGGTGGATTTGCACAGGATGGTCACAGGGCAGTTGTTAGCATTCCAC
5041  CCCTACTCCGGTTCCCACCTAAACGTGTCCTACCAGTGTCCCGTCAACAATCGTAAGGTG
```

FIGURE 4 (cont'd)

```
5101  CACTACATAAAAGGGTGATTTTGCCAAGTGCATTTCATTTCATTTCAGTTGATTTAGATA
5101  GTGATGTATTTTCCCACTAAAACGGTTCACGTAAAGTAAAGTAAAGTCAACTAAATCTAT

5161  CCAAGTTATATTGTTTTGACTTTAGGACTAGATGTGAGTTCAGGAATTTGAAATCATAAT
5161  GGTTCAATATAACAAAACTGAAATCCTGATCTACACTCAAGTCCTTAAACTTTAGTATTA

5221  AAATGTTTCCTTGAACTTTGAAGATAAAAGGTAGAGTTAAAAGACAGCTGCGAAGTCTAC
5221  TTTACAAAGGAACTTGAAACTTCTATTTTCCATCTCAATTTTCTGTCGACGCTTCAGATG

5281  ATGGGGAAGCATAATACATGACTCAGGGTTCCATGGTTTTCAGTGTCGGTGAGTGCTCTG
5281  TACCCCTTCGTATTATGTACTGAGTCCCAAGGTACCAAAAGTCACAGCCACTCACGAGAC

5341  TACACTTATAAGAATAGTGCTGGTGTTTATAGAAGAAACACATGCAGGTTTGAGGCAAGA
5341  ATGTGAATATTCTTATCACGACCACAAATATCTTCTTTGTGTACGTCCAAACTCCGTTCT

5401  CTTCCTTGCTTAATTGGTGCTTTTGCTTTCCATTTTCTAAATGCCATGCAAACTACTTTT
5401  GAAGGAACGAATTAACCACGAAAACGAAAGGTAAAAGATTTACGGTACGTTTGATGAAAA

5461  TTTTTTTTCTTTTGTTGAGACGGAGTCTTGCTCTGTCACCCAGGCTGGAGTACAGTGGTG
5461  AAAAAAAAGAAAACAACTCTGCCTCAGAACGAGACAGTGGGTCCGACCTCATGTCACCAC

5521  CGATCTCGGCTCACTGCAAGCTCCGCCTCCTGGGTTCACGCCATTCTCCTGCCTCAGCTT
5521  GCTAGAGCCGAGTGACGTTCGAGGCGGAGGACCCAAGTGCGGTAAGAGGACGGAGTCGAA

5581  CCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCGGCTAATTTTTTGTATTTTTT
5581  GGACTCATCGACCCTGATGTCCGCGGGCGGTGGTGCGGGCCGATTAAAAAACATAAAAAA

5641  GTAGAGACGGGGTTTCACCATGTTAGCCAGGAAGGTCTTGATCTCCTGACCTCGTGATCT
5641  CATCTCTGCCCCAAAGTGGTACAATCGGTCCTTCCAGAACTAGAGGACTGGAGCACTAGA

5701  GCCCGCCTCGGCCACCCAAGGTGCTGGGATTACAGGCGTGAGCCACCGCTCCCGGCCTCA
5701  CGGGCGGAGCCGGTGGGTTCCACGACCCTAATGTCCGCACTCGGTGGCGAGGGCCGGAGT

5761  AACTACTTTAACTGAAAACATTCTTCTAAGGAAAGCATCAGAAAACTTACTAGAATTGCT
5761  TTGATGAAATTGACTTTTGTAAGAAGATTCCTTTCGTAGTCTTTTGAATGATCTTAACGA

5821  TAAGAATAAATAAATAATACAAATAAAAGAGCAAAATGAGAAAGTAGACAGACTCCTAGT
5821  ATTCTTATTTATTTATTATGTTTATTTTCTCGTTTTACTCTTTCATCTGTCTGAGGATCA

5881  CTATTGTGCTATTGATCTCTCCAGGTTCTAGGCTTTCTTGGGTCTGACACCACTTTGTAA
5881  GATAACACGATAACTAGAGAGGTCCAAGATCCGAAAGAACCCAGACTGTGGTGAAACATT

5941  ATTTTTTTTTTTTTTTTTTTTTTTTCCAGATAGGGTCTCATTCTGTCACCTAGGCTGGA
5941  TAAAAAAAAAAAAAAAAAAAAAAAAGGTCTATCCCAGAGTAAGACAGTGGATCCGACCT

6001  GTGCAGTGGTACAATCATGGCTCACTTCAGCTTCAACCTCCTGGGCTCAAGTGATCTTCC
6001  CACGTCACCATGTTAGTACCGAGTGAAGTCGAAGTTGGAGGACCCGAGTTCACTAGAAGG

6061  TGCCTCAGCTCCCTGAGTAGCTGGGACCACAGGCATGTGCCACCACGCCCGGTGAATTTT
6061  ACGGAGTCGAGGGACTCATCGACCCTGGTGTCCGTACACGGTGGTGCGGGCCACTTAAAA

6121  TGTATTTTTTGTAGAGATGGGGTTTCTCCATGTTGCACTGGCTGGTCTTGAACTCCTGGG
6121  ACATAAAAAACATCTCTACCCCAAAGAGGTACAACGTGACCGACCAGAACTTGAGGACCC

6181  CTCCAGTGATCCACCCACCTTGGCCTCCAAAGTGCTGGGATTACAGGCATCCGTCACTGC
6181  GAGGTCACTAGGTGGGTGGAACCGGAGGTTTCACGACCCTAATGTCCGTAGGCAGTGACG

6241  ACTGGGCCAGTTTGTAAATCTTATGAAAAGTATGGCCAATCTTTTGCAAAATGTGCAAAC
6241  TGACCCGGTCAAACATTTAGAATACTTTTCATACCGGTTAGAAAACGTTTTACACGTTTG

6301  ACATATAGACACAAAGTTTTATATCTAAATTATTGGTTATTAAAGCCAGGACAATAGAGT
6301  TGTATATCTGTGTTTCAAAATATAGATTTAATAACCAATAATTTCGGTCCTGTTATCTCA

6361  GCATATAGCATGGAAGACAGGTAGAAAAGACTGAAATGGTTTAGTTATTCAGAGCAGTGG
6361  CGTATATCGTACCTTCTGTCCATCTTTTCTGACTTTACCAAATCAATAAGTCTCGTCACC
```

FIGURE 4 (cont'd)

```
6421  AGGGCTATTGGCTCTTGCCTTTGCAGTCAGCTGAGATATGAAACTATATAAGCAGAAGCT
6421  TCCCGATAACCGAGAACGGAAACGTCAGTCGACTCTATACTTTGATATATTCGTCTTCGA

6481  TCACGCTGCCATTTGAATATTGAAAATCTCCACTTTTAGTTCAAGAGATAAACAACCAAT
6481  AGTGCGACGGTAAACTTATAACTTTTAGAGGTGAAAATCAAGTTCTCTATTTGTTGGTTA

6541  GGTTTGTTTGTTCACAGCACATAAGTTATCACTTTCTTCTCTGTTTCATCCCATCAAGAT
6541  CCAAACAAACAAGTGTCGTGTATTCAATAGTGAAAGAAGAGACAAAGTAGGGTAGTTCTA

6601  CTAAGGTGTATAGTTACAGAATAGAACTCGGTTCATACCATATGCACCCTAATTTCTTCG
6601  GATTCCACATATCAATGTCTTATCTTGAGCCAAGTATGGTATACGTGGGATTAAAGAAGC

6661  ATCTTCAACTAGTTTGGAAGTGTGTGTGGGTGTTTTCTTAGCCAAGGTTAATATGCAATT
6661  TAGAAGTTGATCAAACCTTCACACACACCCACAAAAGAATCGGTTCCAATTATACGTTAA

6721  TCTTAGTGGAGTGTAGGAGTTTGGGGTGCCTAGCCTTCCTACTTCCCACTTAGATAACTT
6721  AGAATCACCTCACATCCTCAAACCCCACGGATCGGAAGGATGAAGGGTGAATCTATTGAA

6781  CCCCTACCTGTTGTACCTGTCAAAGGAGAAGGCCAGTATGGCCATGTTCTTTGCAGGTGA
6781  GGGGATGGACAACATGGACAGTTTCCTCTTCCGGTCATACCGGTACAAGAAACGTCCACT

6841  CTCCACCTTAGCCACAGCTGATTGTAACAAGGGTGTATATTTGGCCCAAGTCCAGCCATC
6841  GAGGTGGAATCGGTGTCGACTAACATTGTTCCCACATATAAACCGGGTTCAGGTCGGTAG

6901  TTTTTTTTTTTTTTTTGAGGCGGAGTCTTGTTCTGTCACCCACGCTGGAGTGCAGTGGTG
6901  AAAAAAAAAAAAAAAACTCCGCCTCAGAACAAGACAGTGGGTGCGACCTCACGTCACCAC

6961  TGATTTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGCCT
6961  ACTAAAACCGAGTGACGTTGGAGACGGAGGGCCCAAGTTCGTTAAGAGGACGGAGTCGGA

7021  CCTGAGTAGCTGGGATTACAGACACCCACGACCACGCTTGGCTTTGTATTTTTGTATTTT
7021  GGACTCATCGACCCTAATGTCTGTGGGTGCTGGTGCGAACCGAAACATAAAAACATAAAA

7081  TAGTAGAGACGGGGTTTCTCCATGTTGGCCAGGCTGGTCTAGAACTCTTGACCTCAAGTG
7081  ATCATCTCTGCCCCAAAGAGGTACAACCGGTCCGACCAGATCTTGAGAACTGGAGTTCAC

7141  ATCCGCCCGCCTTGGCCTCCCAAACTACTGGGATTACATGCCCAGCCAGCCATCTTTTAC
7141  TAGGCGGGCGGAACCGGAGGGTTTGATGACCCTAATGTACGGGTCGGTCGGTAGAAAATG

7201  TTCTTTCTTATTCTCTTATTCCACATATTCAAGCAGGTACCACGTCCTGTGGACTCTGTC
7201  AAGAAAGAATAAGAGAATAAGGTGTATAAGTTCGTCCATGGTGCAGGACACCTGAGACAG

7261  TTCTCTGAATCTATAGGTTCTCTCTTTCCCATTGCCCTGCATACTGGTTTATGCTCTCAT
7261  AAGAGACTTAGATATCCAAGAGAGAAAGGGTAACGGGACGTATGACCAAATACGAGAGTA

7321  TACTTCTCGCTTGAATTATTCCAAAAGTCTCCCAACTGGTCTGAAGCTTAATTTTCATTA
7321  ATGAAGAGCGAACTTAATAAGGTTTTCAGAGGGTTGACCAGACTTCGAATTAAAAGTAAT

7381  TGTGCTGTTCGATGTCATTCCACCACGCAAAATCTATGGCTTCTCCATTGTCTCCTGAGT
7381  ACACGACAAGCTACAGTAAGGTGGTGCGTTTTAGATACCGAAGAGGTAACAGAGGACTCA

7441  TTAGTTTAGTGCATGAGACCGTTTAAAATCTACCTACACCTCTTCTATGTGCATGCATTT
7441  AATCAAATCACGTACTCTGGCAAATTTTAGATGGATGTGGAGAAGATACACGTACGTAAA

7501  CTCTAGTCATAAGAGAGGATTTGATCTATCTCAAAATGTGTTCTTATATTTGCATTTACT
7501  GAGATCAGTATTCTCTCCTAAACTAGATAGAGTTTTACACAAGAATATAAACGTAAATGA

7561  GGTTTTTAACTCCCTCCATCCTATCTATAATTTGCACCATGTTGCTCACCATGCAGCCCA
7561  CCAAAAATTGAGGGAGGTAGGATAGATATTAAACGTGGTACAACGAGTGGTACGTCGGGT

7621  CTCCCCAGACACCTTGTGAGAACAGTGATCATGTTTTTGCAGTGGAATGAACATGGGCGA
7621  GAGGGGTCTGTGGAACACTCTTGTCACTAGTACAAAAACGTCACCTTACTTGTACCCGCT

7681  AAGGTGGGAAGGTATGCATAGATTTACATTTTGAAGTCAGGAAATAAGATACTGATCAAT
7681  TTCCACCCTTCCATACGTATCTAAATGTAAAACTTCAGTCCTTTATTCTATGACTAGTTA
```

FIGURE 4 (cont'd)

```
7741  GCTTTAATTCGACTGTGCTTTTTACCTTTTCCTCTGTTGTAGCATGAACCTTGCCCCAAA
7741  CGAAATTAAGCTGACACGAAAAATGGAAAAGGAGACAACATCGTACTTGGAACGGGGTTT

7801  AAGATATTGATCATATGCTGCAACTCCTCTGCACTTACCTTTTCTTCTACCATGGCAGGA
7801  TTCTATAACTAGTATACGACGTTGAGGAGACGTGAATGGAAAAGAAGATGGTACCGTCCT

7861  ACCTTGCACCATTGTTGGTGCTTTACAGATATTGGTTGAGTTAAATCGAGAGCCTATTTC
7861  TGGAACGTGGTAACAACCACGAAATGTCTATAACCAACTCAATTTAGCTCTCGGATAAAG

7921  TGGACTTTCAAGATTCAGATTAATTGATTCTGTTTTATACTTCTCTGTTTTTGAATGGCA
7921  ACCTGAAAGTTCTAAGTCTAATTAACTAAGACAAAATATGAAGAGACAAAAACTTACCGT

7981  CATACATCTTTCCACACTTGGATTTCTTTGAAGTTAAAGCTTTAATTGACTTCCAAAAGA
7981  GTATGTAGAAAGGTGTGAACCTAAAGAAACTTCAATTTCGAAATTAACTGAAGGTTTTCT

8041  CAGAGGATTATAAAAACACAGAGAACAGTAAAGTATTTACACGGTCTTAATGCCTATTGC
8041  GTCTCCTAATATTTTTGTGTCTCTTGTCATTTCATAAATGTGCCAGAATTACGGATAACG

8101  CAAACACATATAGCTGCTATTAATGCCTATTGCCAGACTCATATAGCTGCAAACTCATGT
8101  GTTTGTGTATATCGACGATAATTACGGATAACGGTCTGAGTATATCGACGTTTGAGTACA

8161  AGCTTAATGCCCATTGCCAAACTCATATAGCTGCTATATATGTCTACTGGAAAACTCACA
8161  TCGAATTACGGGTAACGGTTTGAGTATATCGACGATATATACAGATGACCTTTTGAGTGT

8221  TAGCTGGTTGATTTTAAATTGATTGTAGGAAAATTCATACAAAAAGATGCTACATAGGAA
8221  ATCGACCAACTAAAATTTAACTAACATCCTTTTAAGTATGTTTTTCTACGATGTATCCTT

8281  TATAAACTGAACACCACAAAAATCGTAATGAACAAACTCAGTAATGCCAGCCTGTTCAAT
8281  ATATTTGACTTGTGGTGTTTTTAGCATTACTTGTTTGAGTCATTACGGTCGGACAAGTTA

8341  CACTAGTTTTCAGATTAAGGCAGTATGCAAAGGGAAACATCCTTGGAATGGTCTTAGAAT
8341  GTGATCAAAAGTCTAATTCCGTCATACGTTTCCCTTTGTAGGAACCTTACCAGAATCTTA

8401  ATAGAACAAAGTCATGTGTTTGGGAATCTTAACCAATTCCAATACTGTTCAACATGTTTT
8401  TATCTTGTTTCAGTACACAAACCCTTAGAATTGGTTAAGGTTATGACAAGTTGTACAAAA

8461  TCTTTGTTGCAAACAATGCAGTTTTTAAGCAAACTAATGAATGCATACCCTGACAAACAA
8461  AGAAACAACGTTTGTTACGTCAAAAATTCGTTTGATTACTTACGTATGGGACTGTTTGTT

8521  ATCCAACAAAAGAGGTACTTAGAGCAGTGTTCATTGAATCCTAGAATTTTGAAAGAAAAA
8521  TAGGTTGTTTTCTCCATGAATCTCGTCACAAGTAACTTAGGATCTTAAAACTTTCTTTTT

8581  ATCTATTAGGTCATTTTGGGCTTTTACTTCTGCTGAAGAAGGATTGTTCTTTTGAGTTTA
8581  TAGATAATCCAGTAAAACCCGAAAATGAAGACGACTTCTTCCTAACAAGAAAACTCAAAT

8641  TTTTTCTAGCTTTCAGTTCTATTCCATTTAAATGCCTCGAGAAATAGGACTTCTCAAATT
8641  AAAAAGATCGAAAGTCAAGATAAGGTAAATTTACGGAGCTCTTTATCCTGAAGAGTTTAA

8701  CCCTGTGGAGAGTTCTGGCACCTACTATATTCACTTTTGAAAGTCTCCTGATAACTGATT
8701  GGGACACCTCTCAAGACCGTGGATGATATAAGTGAAAACTTTCAGAGGACTATTGACTAA

8761  TCCTTTTTCTTTTGGTTATACACATCAACTATTCATGGTTGGCTCTGCTGATAGTTTTCC
8761  AGGAAAAAGAAAACCAATATGTGTAGTTGATAAGTACCAACCGAGACGACTATCAAAAGG

8821  ATGATTCTTTATGCTTATTTAAAAGAAATCATATAACATTAATCATATAACATATAACCA
8821  TACTAAGAAATACGAATAAATTTTCTTTAGTATATTGTAATTAGTATATTGTATATTGGT

8881  TTACTTGGTTATAAAGCATGATATTAACTAAACTGGCTTTTAAAAATTCAATAACATTAG
8881  AATGAACCAATATTTCGTACTATAATTGATTTGACCGAAAATTTTTAAGTTATTGTAATC

8941  AGCAAATATTTTATTAATATTGTCAGTAGTTCCATGAATATATAGATGTAGGATGTCAAT
8941  TCGTTTATAAAATAATTATAACAGTCATCAAGGTACTTATATATCTACATCCTACAGTTA

9001  AAAGTTTATGCCCCTTAATAACTGAGTGTGTTAATGTTGTATTTGCTGGTTAACTGGCAC
9001  TTTCAAATACGGGGAATTATTGACTCACACAATTACAACATAAACGACCAATTGACCGTG
```

FIGURE 4 (cont'd)

```
9061  CTTCCCTCCTTTTGATCAATGTTCTTAGAACAGTATGTACAGGAGTACTTTAGTCTACCA
9061  GAAGGGAGGAAAACTAGTTACAAGAATCTTGTCATACATGTCCTCATGAAATCAGATGGT

9121  CTTGATATTCAAGGACAACACTTCGAAGCTTGAAATGAGGGATTGTTTTAACCATATTAA
9121  GAACTATAAGTTCCTGTTGTGAAGCTTCGAACTTTACTCCCTAACAAAATTGGTATAATT

9181  CGTAACTGCATAGTTTGCAATAGATTTAATGATTCAACATCATTGTTAATATTTAGTTTA
9181  GCATTGACGTATCAAACGTTATCTAAATTACTAAGTTGTAGTAACAATTATAAATCAAAT

9241  ATTAAAAGTGCCTTATATTAACTTGTACACATGTGAAATGACTTATCAACAAGATTATTT
9241  TAATTTTCACGGAATATAATTGAACATGTGTACACTTTACTGAATAGTTGTTCTAATAAA

9301  ATTGCAGCATTGTTTGTGATTACAAAGAGTGGAAGTCATTTGCATAGTCATTAGGCGTAA
9301  TAACGTCGTAACAAACACTAATGTTTCTCACCTTCAGTAAACGTATCAGTAATCCGCATT

9361  AAAATAATGACATTAAAAGAAAACAATTGAGGAAAGAACCAGTAAGCTCTCAGGTAGAGA
9361  TTTTATTACTGTAATTTTCTTTTGTTAACTCCTTTCTTGGTCATTCGAGAGTCCATCTCT

9421  TAAAAAGAAAGTTACCGTGAGCCTTTTCACTGTATATATTTATATATTTGATTTTTGACA
9421  ATTTTTCTTTCAATGGCACTCGGAAAAGTGACATATATAAATATATAAACTAAAAACTGT

9481  CGTGAGTTTATTATAATCAAGATTGAATGAAAAAAATTGAAGCCTTTTAATGAAATTAGC
9481  GCACTCAAATAATATTAGTTCTAACTTACTTTTTTTAACTTCGGAAAATTACTTTAATCG

9541  TTGACACCAGAGTCATTATCAAGAAAATCTGAAATACTGGTACCTATGCACATTTGATTT
9541  AACTGTGGTCTCAGTAATAGTTCTTTTAGACTTTATGACCATGGATACGTGTAAACTAAA

9601  TATTATTTTTAGTTTTATTATTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGACTGGAG
9601  ATAATAAAAATCAAAATAATAAAAAACTCTGTCTCAGAACGAGACAGCGGGTCTGACCTC

9661  TGAAGTGGTGTGATCTCAGCTCACTGCAATCTCCAGCCTCCCAGGTTCAAGCGATTCTCC
9661  ACTTCACCACACTAGAGTCGAGTGACGTTAGAGGTCGGAGGGTCCAAGTTCGCTAAGAGG

9721  TCAGCCTCCCGAGTAGCTGGGATTACAGATGCCTGCCACCATGCTTGGCTAGTTTTTGTA
9721  AGTCGGAGGGCTCATCGACCCTAATGTCTACGGACGGTGGTACGAACCGATCAAAAACAT

9781  TTTTTAGTAGTGACGGAGTTTTGCCATGCTGGCCAGGCTGATCTCAAACTCCTGACCTCA
9781  AAAAATCATCACTGCCTCAAAACGGTACGACCGGTCCGACTAGAGTTTGAGGACTGGAGT

9841  GGTGGTCCACCTGCCTCAGCCTCCCAAAGTGTTGGGATTACAGGCGTGAGCCACTGCACC
9841  CCACCAGGTGGACGGAGTCGGAGGGTTTCACAACCCTAATGTCCGCACTCGGTGACGTGG

9901  CGGCCATTTAAGCAAATTTTAGTTAAGACCAGTAACATATTTTTGACAGTATTTCCGAAG
9901  GCCGGTAAATTCGTTTAAAATCAATTCTGGTCATTGTATAAAAACTGTCATAAAGGCTTC

9961  TGTTGTATTTAATTTAACTAATTGAAAATTATCTATTAAGAAAGCAAGTTAAATGTAATG
9961  ACAACATAAATTAAATTGATTAACTTTTAATAGATAATTCTTTCGTTCAATTTACATTAC

10021 TGTATACTTCATAGTATACACATTATAATGATTCATAGTAACCATTATATGAAGAGTGTC
10021 ACATATGAAGTATCATATGTGTAATATTACTAAGTATCATTGGTAATATACTTCTCACAG

10081 TAAACTTTTCTTCAAATCTACCATTTATATAGTACAATAATGACCAGCAAACAATGGCTA
10081 ATTTGAAAAGAAGTTTAGATGGTAAATATATCATGTTATTACTGGTCGTTTGTTACCGAT

10141 AGTTTTACTTAATTAGACATTTGTGTAGCTCCTACCAAGGCTACCAGACTTTGTACCCAG
10141 TCAAAATGAATTAATCTGTAAACACATCGAGGATGGTTCCGATGGTCTGAAACATGGGTC

10201 TAACAGGAAAACAGAAATAATTCAAGCATTATTCTCACTCTTGAGAAACAATTCTATTTA
10201 ATTGTCCTTTTGTCTTTATTAAGTTCGTAATAAGAGTGAGAACTCTTTGTTAAGATAAAT

10261 GGGAGATATACATAAAAAGCAAAGGTGGCTGGGTGCAGTGGCTCATGCCTGTAATCCCAG
10261 CCCTCTATATGTATTTTTCGTTTCCACCGACCCACGTCACCGAGTACGGACATTAGGGTC

10321 CACTTTGGGAGGCCAAGGCAGGCAGATCACTTGAGGCCAGGAGTTCGAGACTAGCCTGAC
10321 GTGAAACCCTCCGGTTCCGTCCGTCTAGTGAACTCCGGTCCTCAAGCTCTGATCGGACTG
```

FIGURE 4 (cont'd)

```
10381  CAACATGGTGAAACCCTGTTTCTACTAAACATACAAAAATTAGCTGGGTGTGGTGGCATG
10381  GTTGTACCACTTTGGGACAAAGATGATTTGTATGTTTTTAATCGACCCACACCACCGTAC

10441  CACCTGCAATCCCAGCTACTGGGCAGGCTGAGGCAGATAATTGCTTGAACCTGGGAGGTG
10441  GTGGACGTTAGGGTCGATGACCCGTCCGACTCCGTCTATTAACGAACTTGGACCCTCCAC

10501  GAGGCTACAGTGAGCTGAGATTGCACCACTGTACTCCAGCCTGGGTGACAGAGCGAGACT
10501  CTCCGATGTCACTCGACTCTAACGTGGTGACATGAGGTCGGACCCACTGTCTCGCTCTGA

10561  CCATCTCAATGAATAAATAAATAAATAAATAAATAGCAAAGGCAATAATGTGCATATAGT
10561  GGTAGAGTTACTTATTTATTTATTTATTTATTTATCGTTTCCGTTATTACACGTATATCA

10621  AGGTGCTACAGTAGAAGTCTGCACAAGGTTCTGTGTATGTGGAAAAGACAAGTCAGAAAA
10621  TCCACGATGTCATCTTCAGACGTGTTCCAAGACACATACACCTTTTCTGTTCAGTCTTTT

10681  ATTAGGAAAGGCTTTATTGGGAGGTGACATTTAAGCAGAATCACGAAAGAAGACCTGATG
10681  TAATCCTTTCCGAAATAACCCTCCACTGTAAATTCGTCTTAGTGCTTTCTTCTGGACTAC

10741  AACATGGGGGTAAGAGCACTTGTACACAACCCCATACTCCTGGGCTCAGACTCACTTGCT
10741  TTGTACCCCCATTCTCGTGAACATGTGTTGGGGTATGAGGACCCGAGTCTGAGTGAACGA

10801  CATTCCCTAATCATAGAAGACCTTGAGGAAAAGCATGCTGGGGTTCCCTTCCAGCCCACT
10801  GTAAGGGATTAGTATCTTCTGGAACTCCTTTTCGTACGACCCCAAGGGAAGGTCGGGTGA

10861  GGTAAATACAGAGAGGCCTAAGCACCCAACATCCCCCAGGGAATACACCACTGGGACTTA
10861  CCATTTATGTCTCTCCGGATTCGTGGGTTGTAGGGGGTCCCTTATGTGGTGACCCTGAAT

10921  AAGCATTGGCAACCCAGCACCCAGTTAAATTCCCCTTGTCACTGTTTCCACATTGTCCCC
10921  TTCGTAACCGTTGGGTCGTGGGTCAATTTAAGGGGAACAGTGACAAAGGTGTAACAGGGG

10981  ATCACCTTCACTCTCTTGATGCTCTGCATTTTCTCTCTTAACTCGACCCACAGTAGACCC
10981  TAGTGGAAGTGAGAGAACTACGAGACGTAAAAGAGAGAATTGAGCTGGGTGTCATCTGGG

11041  TCCCACTCAAATCTGCCCCCAATAACCTTTGCAACCAATATTACCGCACTACACTTTATC
11041  AGGGTGAGTTTAGACGGGGGTTATTGGAAACGTTGGTTATAATGGCGTGATGTGAAATAG

11101  TTCCCTAAGGGTTTCCTGCTCCTCCTGGTCTTAGGTGAGGTCATTTCTCTGCCAGCCTTT
11101  AAGGGATTCCCAAAGGACGAGGAGGACCAGAATCCACTCCAGTAAAGAGACGGTCGGAAA

11161  AAAGTGGAAGCTGCTCATTTTCTCTGTCCCTCATCCAGGGCTGGAGTGGGACTGATGCCC
11161  TTTCACCTTCGACGAGTAAAAGAGACAGGGAGTAGGTCCCGACCTCACCCTGACTACGGG

11221  TTCACCCTGCCCAGTACTGTTTCCACATGAGACTCCTTGCATAGCATCTGCTCTCTGCTA
11221  AAGTGGGACGGGTCATGACAAAGGTGTACTCTGAGGAACGTATCGTAGACGAGAGACGAT

11281  CTCAGACCAGGAGTCTGTGCTCTCTCTTCCTCTAACCATCATTGGAAGACACCCCAGTCC
11281  GAGTCTGGTCCTCAGACACGAGAGAGAAGGAGATTGGTAGTAACCTTCTGTGGGGTCAGG

11341  CTCTGGCCCCATTATCCACTGTAGATTTTAGCACTTACCTCCTCTCTTCCCTCCCATGTG
11341  GAGACCGGGGTAATAGGTGACATCTAAAATCGTGAATGGAGGAGAGAAGGGAGGGTACAC

11401  CTGTTTTCCCATTTATTTTAAGGTGCATGTGGATGGCCCATTCCCCTTCCTGGCATCTTA
11401  GACAAAAGGGTAAATAAAATTCCACGTACACCTACCGGGTAAGGGGAAGGACCGTAGAAT

11461  ATCCCTTTAATTTCTCATTTCCATTAATTTGCTACATTCCTCATATGCCAGCTACCCCTA
11461  TAGGGAAATTAAAGAGTAAAGGTAATTAAACGATGTAAGGAGTATACGGTCGATGGGGAT

11521  TGGCCATATCCCAGACCTTTGAATCACACACCAAATTCCATCAAGTCCTATGAGTCATTT
11521  ACCGGTATAGGGTCTGGAAACTTAGTGTGTGGTTTAAGGTAGTTCAGGATACTCAGTAAA

11581  TTGGTCCAAGACCTATCTTACATCTATGCAGTGCTATCTATTTTAACAACCATGTGCCCA
11581  AACCAGGTTCTGGATAGAATGTAGATACGTCACGATAGATAAAATTGTTGGTACACGGGT

11641  GTTTAGGCCACTGTCTTCTCTTCACTGAATATATGTGTTCCTGTTTTTTCTCATCCCCAT
11641  CAAATCCGGTGACAGAAGAGAAGTGACTTATATACACAAGGACAAAAAAGAGTAGGGGTA
```

FIGURE 4 (cont'd)

```
11701  TTGTTTCCCAGACACAATTTGTTCTCCAGAGCAGAGTGATCACATCAGTCACCTGCTATA
11701  AACAAAGGGTCTGTGTTAAACAAGAGGTCTCGTCTCACTAGTGTAGTCAGTGGACGATAT

11761  AAAACTCCTCAACCCGGTGGGGCATGGTGACTCACAGCCCCAGCACTGTAATCCCAGCAC
11761  TTTTGAGGAGTTGGGCCACCCCGTACCACTGAGTGTCGGGGTCGTGACATTAGGGTCGTG

11821  TTTGGGAGGCTGAGGTGGGAGGATCACTTGAGCACAGGAGGAGCTGAAGTGATCTGTGAT
11821  AAACCCTCCGACTCCACCCTCCTAGTGAACTCGTGTCCTCCTCGACTTCACTAGACACTA

11881  CACGTCACTGCACTCCAGCCTGGGCTACAGAGTGACACCCTATCTCAAAAACAAGCAAAC
11881  GTGCAGTGACGTGAGGTCGGACCCGATGTCTCACTGTGGGATAGAGTTTTTGTTCGTTTG

11941  AAACAAACAAACAAAAACCAACAAAAAATTTTTTATTAAGCAAAAAACCTCTTCAACCTC
11941  TTTGTTTGTTTGTTTTTGGTTGTTTTTTAAAAAATAATTCGTTTTTTGGAGAAGTTGGAG

12001  TTCCTGTTATGTTTTGAATAAAAGCTAAAGTCCCCACCATGCTCCAGAAGTTCCCAAGGG
12001  AAGGACAATACAAAACTTATTTTCGATTTCAGGGGTGGTACGAGGTCTTCAAGGGTTCCC

12061  ATCTGGTCCTTCTCATCTTTACAGCCTCATCTTGTTTTGCCGTTTCCACCTCCCTCATCC
12061  TAGACCAGGAAGAGTAGAAATGTCGGAGTAGAACAAAACGGCAAAGGTGGAGGGAGTAGG

12121  CCCAAACTCTTTTTCTTCAGGGCCTTTGCACTTGCTGATGCCTCTGCCTGGAAGGTCTTC
12121  GGGTTTGAGAAAAAGAAGTCCCGGAAACGTGAACGACTACGGAGACGGACCTTCCAGAAG

12181  TTCTGATTCCCCATCCAGGGGTGCCTTATCACCCTGCCCTGCCAACCTCAGATTGGATGC
12181  AAGACTAAGGGGTAGGTCCCCACGGAATAGTGGGACGGGACGGTTGGAGTCTAACCTACG

12241  TACTTCCTCAAAGAGACCCTTGGGACTCCCATGTTTCCCAGGCTTTCAGTGCCCTCATCT
12241  ATGAAGGAGTTTCTCTGGGAACCCTGAGGGTACAAAGGGTCCGAAAGTCACGGGAGTAGA

12301  CAACTTGTAATGATTGTATTGGGGTGATTATTTGTATGACTTTGGGATTGATATCTGACT
12301  GTTGAACATTACTAACATAACCCCACTAATAAACATACTGAAACCCTAACTATAGACTGA

12361  CTTCCACTAGACCATACATTTCTTGAGAGCAGAGAGCATATCCAGTTTTTTTACCACTGT
12361  GAAGGTGATCTGGTATGTAAAGAACTCTCGTCTCTCGTATAGGTCAAAAAAATGGTGACA

12421  GCAACCATATGTCGTCCAATGCTTGGCCCATAATAGGGGGTTGGTGAATGTGGTAGGCTG
12421  CGTTGGTATACAGCAGGTTACGAACCGGGTATTATCCCCCAACCACTTACACCATCCGAC

12481  AGTCATAGCCCCCAAGGATGCCCATAGCCTCATCCCCAGAAGCTTTCATGGTAAAAGGGA
12481  TCAGTATCGGGGGTTCCTACGGGTATCGGAGTAGGGGTCTTCGAAAGTACCATTTTCCCT

12541  TTTTGCAGATTGTGATTAAGTTATGGATCTTGAGATGGAGAGATTATCCTTGATTATCTA
12541  AAAACGTCTAACACTAATTCAATACCTAGAACTCTACCTCTCTAATAGGAACTAATAGAT

12601  AATAGGTCCAATATAATCAGTAGCGTCTTTATAAGAGAGAGGCAGAGGGTCAAAAGCAGA
12601  TTATCCAGGTTATATTAGTCATCGCAGAAATATTCTCTCTCCGTCTCCCAGTTTTCGTCT

12661  ATCAGAGAAGGAGATGCCACAGTGGAAGCAAAGGGGTGTGTGTGTATGTGTGTGTGTATG
12661  TAGTCTCTTCCTCTACGGTGTCACCTTCGTTTCCCCACACACACATACACACACACATAC

12721  TGTGTGTGTGTGTGTGAGAGAGAGAGATTGAGATTTAAAGAGCTCCATTACTTGATCTGA
12721  ACACACACACACACACTCTCTCTCTCTAACTCTAAATTTCTCGAGGTAATGAACTAGACT

12781  GGATAAGGATGAGCCAAAGAATGCAGGTGGCTTGTAGAAGCTGGAGAAGCCAAGAAAACA
12781  CCTATTCCTACTCGGTTTCTTACGTCCACCGAACATCTTCGACCTCTTCGGTTCTTTTGT

12841  TTCTCACTCTAGAGCCTCCAGAAGGAATGCAGCTCTGCTGACACCTTGATTTGGAACTTC
12841  AAGAGTGAGATCTCGGAGGTCTTCCTTACGTCGAGACGACTGTGGAACTAAACCTTGAAG

12901  CAGAACCATAAGATAATAATTTTGTGTAGTTTTGAGGCTGGGCTCATGCCTGTAATTCCA
12901  GTCTTGGTATTCTATTATTAAAACACATCAAAACTCCGACCCGAGTACGGACATTAAGGT

12961  GCACTTTGGGAGGCCAAGGTGGGTGGATCACTTGAGGTCAGGAGTTTGAGACTAGCCTGG
12961  CGTGAAACCCTCCGGTTCCACCCACCTAGTGAACTCCAGTCCTCAAACTCTGATCGGACC
```

FIGURE 4 (cont'd)

```
13021  CCAACATGGTGAAACCCTGTATGTACTAAATATACAAAAATTAGCCAGGTGTGGTGGTGC
13021  GGTTGTACCACTTTGGGACATACATGATTTATATGTTTTTAATCGGTCCACACCACCACG

13081  ACGCCTATAGACCCAACTACTTGGGAAGCTGGGGCAGGAGAATCACTTGAACCCGGGAGG
13081  TGCGGATATCTGGGTTGATGAACCCTTCGACCCCGTCCTCTTAGTGAACTTGGGCCCTCC

13141  TAGAGATTGCAGTAAGCTAAGATTGTGCCACTGCACTCCAGCCTGGGCTACAGCGTGAGA
13141  ATCTCTAACGTCATTCGATTCTAACACGGTGACGTGAGGTCGGACCCGATGTCGCACTCT

13201  CTTCGTCAGAAAAAAAGAATAATTTTGTATAATTTCAGGCCACAAAGTCTGCAGTCATCT
13201  GAAGCAGTCTTTTTTTCTTATTAAAACATATTAAAGTCCGGTGTTTCAGACGTCAGTAGA

13261  GTTACAACAGCAGTAGTACACGAATACAGATTTTGGTACCTGGAAATGGTGTGCTACTGT
13261  CAATGTTGTCGTCATCATGTGCTTATGTCTAAAACCATGGACCTTTACCACACGATGACA

13321  AACAGATACCTAAACAATGTGGAAGTGGCTTTTGAATTGGGTGTTGGGCAGCATGTGGAA
13321  TTGTCTATGGATTTGTTACACCTTCACCGAAAACTTAACCCACAACCCGTCGTACACCTT

13381  GAATTTTGAGAAGCATGATAGGAAAAGCCTAGCTTGCCTTAGAAAGACTGTTAGTAGAAG
13381  CTTAAAACTCTTCGTACTATCCTTTTCGGATCGAACGGAATCTTTCTGACAATCATCTTC

13441  TAGGGATGTTAAAGGCGCTGCTGGTGAGGACTTGGAAGTGAAGAGCATGGTAAAGAAAAT
13441  ATCCCTACAATTTCCGCGACGACCACTCCTGAACCTTCACTTCTCGTACCATTTCTTTTA

13501  GTGTTTCATTTTAGAGAATATCAAAGTCATTGTAGAAGGTTGGCAGAAGTGTAGGTTTTC
13501  CACAAAGTAAAATCTCTTATAGTTTCAGTAACATCTTCCAACCGTCTTCACATCCAAAAG

13561  ATAATGTGTTGTTGGTGAGGGCTGAGAAAGAAATGAGGACCATGTTTAGGGAAACTGGAG
13561  TATTACACAACAACCACTCCCGACTCTTTCTTTACTCCTGGTACAAATCCCTTTGACCTC

13621  GAAAAAGATCCTTGTCATACAGTGGAATGAAACTGTGGAATTGTGTCTTACAGTTATATG
13621  CTTTTTCTAGGAACAGTATGTCACCTTACTTTGACACCTTAACACAGAATGTCAATATAC

13681  GGAAGCAGAATTTGTTTGAACTTGGATATTTGGCTTAGCTTTCCAAGTAAAGTATTGAAG
13681  CCTTCGTCTTAAACAAACTTGAACCTATAAACCGAATCGAAAGGTTCATTTCATAACTTC

13741  GTGTAGCCTGGTTTTCTCTTGCTGTTTTATAGTAAAATGCCCGAGAAAAAGGGATAAATT
13741  CACATCGGACCAAAAGAGAACGACAAAATATCATTTTACGGGCTCTTTTTCCCTATTTAA

13801  GAGGTAGGAACTGTTAAGAAAAAGGAACAAAGACATGATGACTTGGGAAATTCTCAGGAT
13801  CTCCATCCTTGACAATTCTTTTTCCTTGTTTCTGTACTACTGAACCCTTTAAGAGTCCTA

13861  ACCTGGTTTGTAAAAGATGCTAGAATTAGAAGATTCATTGTCAGGAAAGTGTGCTCTGGA
13861  TGGACCAAACATTTTCTACGATCTTAATCTTCTAAGTAACAGTCCTTTCACACGAGACCT

13921  GAGGAAGCCAAGGGTGAGTTCCTTGTGGCGTAGGTTAGGGGCTGTGGCAGGGATGCAGGG
13921  CTCCTTCGGTTCCCACTCAAGGAACACCGCATCCAATCCCCGACACCGTCCCTACGTCCC

13981  AAGAGATGATGGTGACTTGGACTAGGGGTCTGACGGGGTATGGGAAAATAGAGAGGAATG
13981  TTCTCTACTACCACTGAACCTGATCCCCAGACTGCCCCATACCCTTTTATCTCTCCTTAC

14041  AGTTGATTCTGTTGGGACTTTGTATTGAGTTCCTTGATTCCATAAGTGGAAACACATGTA
14041  TCAACTAAGACAACCCTGAAACATAACTCAAGGAACTAAGGTATTCACCTTTGTGTACAT

14101  GAGCTCTCTTTCCTGGCCACCTAAGCTCTCTGACCATATCCCTGACTCTCTCCAGAATGC
14101  CTCGAGAGAAAGGACCGGTGGATTCGAGAGACTGGTATAGGGACTGAGAGAGGTCTTACG

14161  CTTGGCTTCATGTCCACGTGAGGACAAATCATATTTTTCTGCATTTTGCGTGCTGGATTC
14161  GAACCGAAGTACAGGTGCACTCCTGTTTAGTATAAAAAGACGTAAAACGCACGACCTAAG

14221  CTCTCTGGACTAAAGATCCAATGAGAGGTGGTTTCTAGCTTTCTGCTCTTTCCCACAACA
14221  GAGAGACCTGATTTCTAGGTTACTCTCCACCAAAGATCGAAAGACGAGAAAGGGTGTTGT

14281  CTCCGTCAGTGAGTCTTCTGGATGTCCTATAGGCAACTCACTGGATCCACTTGGTCTGTG
14281  GAGGCAGTCACTCAGAAGACCTACAGGATATCCGTTGAGTGACCTAGGTGAACCAGACAC
```

FIGURE 4 (cont'd)

```
14341  GAAGAATCACTACCACATCCATTTCACCTTGGGCACAGACTAGAAGGTGAGGCAGTCATG
14341  CTTCTTAGTGATGGTGTAGGTAAAGTGGAACCCGTGTCTGATCTTCCACTCCGTCAGTAC

14401  TCCTGCAGTCACTGGGTAGCTTACAAGTGCTAATACAGCAATAGCTTAATTAGAGAGAGA
14401  AGGACGTCAGTGACCCATCGAATGTTCACGATTATGTCGTTATCGAATTAATCTCTCTCT

14461  GAGAAGAAGAAGAATGTTATGGAGCCTGTTTCTAGAGAGGAAACCACAATATCTTGATCA
14461  CTCTTCTTCTTCTTACAATACCTCGGACAAAGATCTCTCCTTTGGTGTTATAGAACTAGT

14521  AGAGACCCTAAAACTGGAAGAACCAGAGCAGCTGGTATCAAAATGTGGAAAAAACTGAGA
14521  TCTCTGGGATTTTGACCTTCTTGGTCTCGTCGACCATAGTTTTACACCTTTTTTGACTCT

14581  TGTGGCAAGTGGTACAACTGATAGCCAAGGAGGAAAAATTTGTATCCCCCTGTAATCCCA
14581  ACACCGTTCACCATGTTGACTATCGGTTCCTCCTTTTTAAACATAGGGGGACATTAGGGT

14641  GCACTTTGGGAGGCTGAAACAGGTGGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGG
14641  CGTGAAACCCTCCGACTTTGTCCACCTAGTGGACTCCAGTCCTCAAGCTCTGGTCGGACC

14701  CCAATATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCTGGGCGTGGTGGTGT
14701  GGTTATACCACTTTGGGACAGAGATGATTTTTATGTTTTTAATCGACCCGCACCACCACA

14761  GCACCTATAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAAAATTGCTTGAACCTGGGAGC
14761  CGTGGATATCAGGGTCGATGAACCCTCCGACTCCGTCCTTTTAACGAACTTGGACCCTCG

14821  AGAGGTTGCAGTGAGCCAAGATCGCACCACTGCACTCCAGCCTGGGTGACAAGAGTTAAA
14821  TCTCCAACGTCACTCGGTTCTAGCGTGGTGACGTGAGGTCGGACCCACTGTTCTCAATTT

14881  CTCTGTCTCAAAAAAATAAAAAGAAAAGTTTGTACCCTCTTCTTTTCTCTTGCCTCAGGA
14881  GAGACAGAGTTTTTTTATTTTTCTTTTCAAACATGGGAGAAGAAAAGAGAACGGAGTCCT

14941  TAACTGAGATTCTTATAATGCATTTATTGAACAAATATTTGAGGAGTACTCATTATGTAA
14941  ATTGACTCTAAGAATATTACGTAAATAACTTGTTTATAAACTCCTCATGAGTAATACATT

15001  GCATCAGACACTGCTCAAGGTCTTGAGGCTGTGAGGGTGTATTAGTCCGTTTTCACACTG
15001  CGTAGTCTGTGACGAGTTCCAGAACTCCGACACTCCCACATAATCAGGCAAAAGTGTGAC

15061  CTATAAAGAACTACCTGAGACTGGGTGATTTATAAAGAAAAGAGGTTTAACTGAATTACA
15061  GATATTTCTTGATGGACTCTGACCCACTAAATATTTCTTTTCTCCAAATTGACTTAATGT

15121  GTTCATCATGGCCAGAAAGGCCTCAGGAAACTTACAATCATGGCAGAAGGTGAAGAGGAA
15121  CAAGTAGTACCGGTCTTTCCGGAGTCCTTTGAATGTTAGTACCGTCTTCCACTTCTCCTT

15181  GCAAGGCACATCTTACATGGTGGCAGGAGACACAGTGAGAAGGGAGAAGTGCCACACTTT
15181  CGTTCCGTGTAGAATGTACCACCGTCCTCTGTGTCACTCTTCCCTCTTCACGGTGTGAAA

15241  TAAACCATCAGATCTTATGATGATATAGGAGTTAAGAAGAAATCACTTAGGTAGATAGTG
15241  ATTTGGTAGTCTAGAATACTACTATATCCTCAATTCTTCTTTAGTGAATCCATCTATCAC

15301  AGGGTACGGGAGTCCTTGGTAAGGCTTTTCTTTTAAATGAAAAGCAGCCCCACACATAGA
15301  TCCCATGCCCTCAGGAACCATTCCGAAAAGAAAATTTACTTTTCGTCGGGGTGTGTATCT

15361  CAAGCAAGCTGGGAGCTTGCACGGGTGAATGCCGGCAGAAACTAGGAACCAGACACGTTA
15361  GTTCGTTCGACCCTCGAACGTGCCCACTTACGGCCGTCTTTGATCCTTGGTCTGTGCAAT

15421  AGATGGCGGCTCCATTTTCCCTTCTCTGCCAGCCAAGTGTACAGTAAGGAGCAGACAAGC
15421  TCTACCGCCGAGGTAAAAGGGAAGAGACGGTCGGTTCACATGTCATTCCTCGTCTGTTCG

15481  TGGTGCCGGCCAAGAGGAGAATTCATGTGCATAATAAGATTAGGGTGGGGTGACCAGCCT
15481  ACCACGGCCGGTTCTCCTCTTAAGTACACGTATTATTCTAATCCCACCCCACTGGTCGGA

15541  TCCCAGTGCGCTATGTGAACATCATACCTGATTGAACCAATCTGTGAGCCCTATGTAAAT
15541  AGGGTCACGCGATACACTTGTAGTATGGACTAACTTGGTTAGACACTCGGGATACATTTA

15601  CAGACACCACCTCCTCAAGCCTGACTATAAAATCTGGTGCATTCGCCACCCGCCGGTGTT
15601  GTCTGTGGTGGAGGAGTTCGGACTGATATTTTAGACCACGTAAGCGGTGGGCGGCCACAA
```

FIGURE 4 (cont'd)

```
15661  TCCTCTCAGAAGTCCCCTCTCTCTCATTAGAGAGAGGGCTGTTTTCCTTTCTCCTTCTTC
15661  AGGAGAGTCTTCAGGGGAGAGAGAGTAATCTCTCTCCCGACAAAAGGAAAGAGGAAGAAG

15721  TGCCTATTAACCTGTGCTCCTAAACTCCCTGTGTGTGTCTGTTCTGAATTTTCCCGGCAT
15721  ACGGATAATTGGACACGAGGATTTGAGGGACACACACAGACAAGACTTAAAAGGGCCGTA

15781  GAGAGGATGAACCTGGGGTATATACCCCAGAAAACATAGCCACTTAAGTGAGAACTCACT
15781  CTCTCCTACTTGGACCCCATATATGGGGTCTTTTGTATCGGTGAATTCACTCTTGAGTGA

15841  CACTATTCACCCCCAAGATCCAATCACCTTCGACCAGGTTCCTCCCCTGGCGTGAGGGGA
15841  GTGATAAGTGGGGGTTCTAGGTTAGTGGAAGCTGGTCCAAGGAGGGGACCGCACTCCCCT

15901  TTAAAATTCCACATGAGGTTTGAGTAGGGACACAGAGCCAAACCAAACCATATCAGAGGG
15901  AATTTTAAGGTGTACTCCAAACTCATCCCTGTGTCTCGGTTTGGTTTGGTATAGTCTCCC

15961  TGAAGAGCAGGTAAAGTTGCAGCTCTCCTGCAACTTGCAGGGAAAGGCCATATAACACAT
15961  ACTTCTCGTCCATTTCAACGTCGAGAGGACGTTGAACGTCCCTTTCCGGTATATTGTGTA

16021  GGTTATGAGCACAAACTGGGACTGGACTGCAGGTGTGTGAGCCCCAGCTTTGCTATTGAG
16021  CCAATACTCGTGTTTGACCCTGACCTGACGTCCACACACTCGGGGTCGAAACGATAACTC

16081  TGACCACCAGCAAGTTATTAAACTGTGCCTCAGTTTTTTTGTTTATGAAATGGTGGTATT
16081  ACTGGTGGTCGTTCAATAATTTGACACGGAGTCAAAAAAACAAATACTTTACCACCATAA

16141  AGTACAACCCACCTCTCATGGATGATGTGAGGATTAAATAAAATAACATGTTGCCAGGCA
16141  TCATGTTGGGTGGAGAGTACCTACTACACTCCTAATTTATTTTATTGTACAACGGTCCGT

16201  CAGTGGTGCACACCTGTAATCTCGGCACATTGGGAGGCTGAGGCCCGAGGGTTACTTGAG
16201  GTCACCACGTGTGGACATTAGAGCCGTGTAACCCTCCGACTCCGGGCTCCCAATGAACTC

16261  GTTAGAAGCCTGAGACCTGCCTGGCCAACATAGCGAAACCTCATCTCTGCTTAAAATACA
16261  CAATCTTCGGACTCTGGACGGACCGGTTGTATCGCTTTGGAGTAGAGACGAATTTTATGT

16321  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCCAGGCGTGGTGGCTCATGCTTGTAATC
16321  TTTTTTTTTTTTTTTTTTTTTTTTTTTTCCGGTCCGCACCACCGAGTACGAACATTAG

16381  CCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACAAGGTCAGGAGTTTGAGACCAGCCTG
16381  GGTCGTGAAACCCTCCGGCTCCGCCCACCTAGTGTTCCAGTCCTCAAACTCTGGTCGGAC

16441  GCCAATATGGTGAATCCCCGTCTCTACTAAAAATACAAAAATTAGCCAAGTGTGGTGGTG
16441  CGGTTATACCACTTAGGGGCAGAGATGATTTTTATGTTTTTAATCGGTTCACACCACCAC

16501  GGCACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAG
16501  CCGTGGACATCAGGGTCGATGAGCCCTCCGACTCCGTCCTCTTAGTGAACTTGGACCCTC

16561  GCAGAAGTTGCAGTGAGCCGAGGTTGCGCCACTGCACTCCAGCCTGGGCGACGGAGCGAG
16561  CGTCTTCAACGTCACTCGGCTCCAACGCGGTGACGTGAGGTCGGACCCGCTGCCTCGCTC

16621  ACTCCATCTCAAAACAAAACAAAAAAACCAAAAAAATTAACCAGGCATGGTGGCACACAC
16621  TGAGGTAGAGTTTTGTTTTGTTTTTTTGGTTTTTTTAATTGGTCCGTACCACCGTGTGTG

16681  CTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGTGGA
16681  GACATTAGGGTCGATGAGTCCTCCGACTCCGTCCTCTTAGTGAACTTGGGTCCTCCACCT

16741  GGTTGCAGTGAGTTAAGATTGTGCCACTGCACTCCAGCCTGTGCGATAGAAGGAGACAGT
16741  CCAACGTCACTCAATTCTAACACGGTGACGTGAGGTCGGACACGCTATCTTCCTCTGTCA

16801  CTGCAACAAAGATCCAAAAACAAATATAAAGCACTTATAACAGGGTCTGGTCTGAAATAA
16801  GACGTTGTTTCTAGGTTTTTGTTTATATTTCGTGAATATTGTCCCAGACCAGACTTTATT

16861  GTACTATGTAAGTACTGTAAGTATTAACTAATATTATTACAATTTATCCTCATATAGTAA
16861  CATGATACATTCATGACATTCATAATTGATTATAATAATGTTAAATAGGAGTATATCATT

16921  CATATCAAATATTTTCAGTCAGTGATGCAATATGAATAAAATGAAATGATGCAGTAGTGA
16921  GTATAGTTTATAAAAGTCAGTCACTACGTTATACTTATTTTACTTTACTACGTCATCACT
```

FIGURE 4 (cont'd)

```
16981  CAGAACTGAGAGAAAGCTGGTACAAATTTATACTGACTGGTCAGGGAAAGCTTCACAGAA
16981  GTCTTGACTCTCTTTCGACCATGTTTAAATATGACTGACCAGTCCCTTTCGAAGTGTCTT

17041  AAGGTGAACTTTCAGCTGACACCTGAGTACCATGGTCTGGAGGTAAGAGTGTCCCATGCA
17041  TTCCACTTGAAAGTCGACTGTGGACTCATGGTACCAGACCTCCATTCTCACAGGGTACGT

17101  TATGGAACAGCAAATGAAACTTCCCAGAGCATAATCTGAGAGGGGAGATTTCAGGAGTGG
17101  ATACCTTGTCGTTTACTTTGAAGGGTCTCGTATTAGACTCTCCCCTCTAAAGTCCTCACC

17161  AGCCACATCATTGAAGTTCCTTAGTCCTGGTAAGTGTGGATTTTATTCTAAGTGTGGTGA
17161  TCGGTGTAGTAACTTCAAGGAATCAGGACCATTCACACCTAAAATAAGATTCACACCACT

17221  AATGCCCCTGAGAGGTTTTGAGTAGTAGATGGAATTCCATTTTTAAAAGTCACTGGTGAG
17221  TTACGGGGACTCTCCAAAACTCATCATCTACCTTAAGGTAAAAATTTTCAGTGACCACTC

17281  TCAAGTAGTGTACTATGGTTACATTTTTTTCCCCTTATACCTCATTTTGTTTTTCCTGCA
17281  AGTTCATCACATGATACCAATGTAAAAAAAGGGGAATATGGAGTAAAACAAAAAGGACGT

17341  GTTAATTCTTGCATTATTTCCTTTTATTTACTTACCTTACTATATATTTATTATGAATTC
17341  CAATTAAGAACGTAATAAAGGAAAATAAATGAATGGAATGATATATAAATAATACTTAAG

17401  TTTCCAATTCCCAACTTCCTGACAGCCTCTCAATGTAATTTTACTTAGGTTAAATGTATC
17401  AAAGGTTAAGGGTTGAAGGACTGTCGGAGAGTTACATTAAAATGAATCCAATTTACATAG

17461  TGACAGTTCATTTTCTTTTTTGTCTCTGTACTCTTTCTGCTACCATTCTCATTCTTGCAC
17461  ACTGTCAAGTAAAAGAAAAAACAGAGACATGAGAAAGACGATGGTAAGAGTAAGAACGTG

17521  TAATTTGTGCTGGTTTCTCTCTGAGCCAGCTGCACAGCTGCCTGCCTAAGACTTCCTCTT
17521  ATTAAACACGACCAAAGAGAGACTCGGTCGACGTGTCGACGGACGGATTCTGAAGGAGAA

17581  GGCATCTTTTACCCTCTCTCTGTGTTACATCCCCGGTTTTCTGAATCGTGTGGCGTTCTC
17581  CCGTAGAAAATGGGAGAGAGACACAATGTAGGGGCCAAAAGACTTAGCACACCGCAAGAG

17641  CTTTTTACTTTACTTCCTCATTTGTGTTGCACGTATGCTCTAGTTGTTTTTAGATAAAGA
17641  GAAAAATGAAATGAAGGAGTAAACACAACGTGCATACGAGATCAACAAAAATCTATTTCT

17701  GTGCATCAGAGGTAGATGTTTTGATTCCTTACATGCCTGACTTTTTGTTCTACTTTTGCT
17701  CACGTAGTCTCCATCTACAAAACTAAGGAATGTACGGACTGAAAAACAAGATGAAAACGA

17761  CTAGGTTAATACTTGATTTTGAATAATATCCTAGGTTAGAAATGATTTTTCGTCAGAATT
17761  GATCCAATTATGAACTAAAACTTATTATAGGATCCAATCTTTACTAAAAAGCAGTCTTAA

17821  TTAAGGGAATTGCTTTTCTGTTGTCTAGTTTCCTGTGTTACTGATAGGAAGTCCAAAGGG
17821  AATTCCCTTAACGAAAAGACAACAGATCAAAGGACACAATGACTATCCTTCAGGTTTCCC

17881  GCTCGTATTTCAGGTCCTTTTTATGTGTAACCTGTTTTTTGTTCCTTGAGGTATTCAGGG
17881  CGAGCATAAAGTCCAGGAAAAATACACATTGGACAAAAAACAAGGAACTCCATAAGTCCC

17941  TCTCTTTTTATTGCTAATTTGGGGAGCTTTATCATAACATGGGTCCTTAGTCATTCATCC
17941  AGAGAAAAATAACGATTAAACCCCTCGAAATAGTATTGTACCCAGGAATCAGTAAGTAGG

18001  TAGAAAATTTTCCTTTCCTATATCTTTGATGATTTCGTCTCCTCTGCTTTCCCTCTTTTC
18001  ATCTTTTAAAAGGAAAGGATATAGAAACTACTAAAGCAGAGGAGACGAAAGGGAGAAAAG

18061  TCATTCTTAAACTTCTATTTGTTTGACATTAGACCTCCTTGTTTGATCTTCTAATTTAAT
18061  AGTAAGAATTTGAAGATAAACAAACTGTAATCTGGAGGAACAAACTAGAAGATTAAATTA

18121  TTTTCTCTCCTAGTGTCCAATTTTTGAGGTGGGAGAAAGATCTCATCAATGTTATCTTTC
18121  AAAAGAGAGGATCACAGGTTAAAAACTCCACCCTCTTTCTAGAGTAGTTACAATAGAAAG

18181  AGCCTTTCCCCTGAATTTTAAAATTTTACTCTTCATATTTCTTCTGATTAATCCTTTAAA
18181  TCGGAAAGGGGACTTAAAATTTTAAAATGAGAAGTATAAAGAAGACTAATTAGGAAATTT

18241  AAAATAGCATCCTGGGCCAGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGAAGG
18241  TTTTATCGTAGGACCCGGTCCACGTCACCGAGTGTGGACATTAGGGTCGTGAAACCTTCC
```

FIGURE 4 (cont'd)

```
18301  CTGAGCGGGGAGTGGATCACCTGAGGTCAGGAGCTCAAGACCAGTCTGGCCAACATGATG
18301  GACTCGCCCCTCACCTAGTGGACTCCAGTCCTCGAGTTCTGGTCAGACCGGTTGTACTAC

18361  AAACCCTGTCTCTACTAAAAATACAAAAACTTAGCCAGGCGTGGTGGTGGGCTCCTGTAG
18361  TTTGGGACAGAGATGATTTTTATGTTTTTGAATCGGTCCGCACCACCACCCGAGGACATC

18421  TCCCAGCTACTTGGGAGGCTGAGACAGGAGAATCACCAGAACCTGGGAGGTGGAGGTTGT
18421  AGGGTCGATGAACCCTCCGACTCTGTCCTCTTAGTGGTCTTGGACCCTCCACCTCCAACA

18481  AGTGAGCTGAGATTGTGTGCCATTGTACTCCAACCTGGGTGACAGAGCAAGACTCTATCT
18481  TCACTCGACTCTAACACACGGTAACATGAGGTTGGACCCACTGTCTCGTTCTGAGATAGA

18541  CAAAAAAAAAAAAAAAAAAAAAAGGGCCCTGTTCCACTGAGGCTTCAGTGCGGAGCATGG
18541  GTTTTTTTTTTTTTTTTTTTTTTCCCGGGACAAGGTGACTCCGAAGTCACGCCTCGTACC

18601  AGGGATACTACCTCCAAATTCCAAACACGTTTGACACATCTAAAGAACCTTCTAAAGGAT
18601  TCCCTATGATGGAGGTTTAAGGTTTGTGCAAACTGTGTAGATTTCTTGGAAGATTTCCTA

18661  GACTAAACTAATTCTGCAAGCTTTAGCCTCTAACTTATCTCCAATTTATATTTAATAGAA
18661  CTGATTTGATTAAGACGTTCGAAATCGGAGATTGAATAGAGGTTAAATATAAATTATCTT

18721  GGCTGGAAATTGAACAGAAAGGAAATAAGATTATTTGGAAATAAAAAATTGTCAGCCAAG
18721  CCGACCTTTAACTTGTCTTTCCTTTATTCTAATAAACCTTTATTTTTTAACAGTCGGTTC

18781  CAAGGAGAGCTTGGCATTTCTGGAAACTTCTTTACCCTGGAAGAGGTAAAAAAGTGATTC
18781  GTTCCTCTCGAACCGTAAAGACCTTTGAAGAAATGGGACCTTCTCCATTTTTTCACTAAG

18841  TCCCATTCCCCCAACAGGGCATTCACTAAAAATTAAGTGAGAAATGCTTTGCCTAGAGGA
18841  AGGGTAAGGGGGTTGTCCCGTAAGTGATTTTTAATTCACTCTTTACGAAACGGATCTCCT

18901  TCTGTAGCTTGCGAACACTTGACTTGTGCCACTGGTGTACGGAACATAGCAGTGTCAGCT
18901  AGACATCGAACGCTTGTGAACTGAACACGGTGACCACATGCCTTGTATCGTCACAGTCGA

18961  CTGCCCTGGGACAGTCAAGGCTGGCTGGTGCTGCCTGTGGTTAGAGTCAAGAGAGGGGTC
18961  GACGGGACCCTGTCAGTTCCGACCGACCACGACGGACACCAATCTCAGTTCTCTCCCCAG

19021  TTCAGTGGGTGCAGTTTGCAGTTCCAGAGTTTAGTAGAGCAAAGACGTCATGAACTATGT
19021  AAGTCACCCACGTCAAACGTCAAGGTCTCAAATCATCTCGTTTCTGCAGTACTTGATACA

19081  GAACTCTGAGGAAAGGTGTCATAGAAACACCTCAAAAGGACACTGTCTAAGAAGGCTACT
19081  CTTGAGACTCCTTTCCACAGTATCTTTGTGGAGTTTTCCTGTGACAGATTCTTCCGATGA

19141  TGCTGGGGCAAGGAGACCCCAGGGGTCAGAGGCTGTGTAGGGAATAATGCCAGAGCTGGG
19141  ACGACCCCGTTCCTCTGGGGTCCCCAGTCTCCGACACATCCCTTATTACGGTCTCGACCC

19201  GTTGAAGGAGAATCCTTAGAGGTGAGACAGAATGTGCATCATTGCAGTGGGGCTGTGCAA
19201  CAACTTCCTCTTAGGAATCTCCACTCTGTCTTACACGTAGTAACGTCACCCCGACACGTT

19261  GGCAGTGGTCTCCACAGGAGGTGAGGACAAGGGGGAGGGGAGGTTGTGAAGAATCCGTAA
19261  CCGTCACCAGAGGTGTCCTCCACTCCTGTTCCCCCTCCCCTCCAACACTTCTTAGGCATT

19321  GTGACCATGATGCAATAGAAAGCCAGAATTTGGGCATTTGCCATAGGAGAGGGCCCCCAG
19321  CACTGGTACTACGTTATCTTTCGGTCTTAAACCCGTAAACGGTATCCTCTCCCGGGGGTC

19381  TGCCAGATTATAATGGTCACAGACAAATGAGATTTTTAACCATTTCCTCTAATTCATGGT
19381  ACGGTCTAATATTACCAGTGTCTGTTTACTCTAAAAATTGGTAAAGGAGATTAAGTACCA

19441  CCACACTTAACCCCGTGGGTCCTGGAAACAGCAGCTGGAGAATGGAGAGGAAAGTGATGT
19441  GGTGTGAATTGGGGCACCCAGGACCTTTGTCGTCGACCTCTTACCTCTCCTTTCACTACA

19501  GGGAGAGAGAAAACTTTCTTTACCCTTTTCATAACTGAGTTTATGAAGTAAACTGACAGT
19501  CCCTCTCTCTTTTGAAAGAAATGGGAAAAGTATTGACTCAAATACTTCATTTGACTGTCA

19561  AGAAAGATAAATAGGAGAAAAGACATAAATTTATTTTGTGCGTACACATGGGAGTCCCAT
19561  TCTTTCTATTTATCCTCTTTTCTGTATTTAAATAAAACACGCATGTGTACCCTCAGGGTA
```

FIGURE 4 (cont'd)

```
19621  AAAATATGAGACTCAAAGAAAGGCCAGATGATTGAAGTTTATATAGTATTCCTGAGCTAC
19621  TTTTATACTCTGAGTTTCTTTCCGGTCTACTAACTTCAAATATATCATAAGGACTCGATG

19681  AGAAAAGGAATCAGGGCTTGGGGCTTATTGAGGGGAGGTGGTGACAGGTTATAAGAGGGT
19681  TCTTTTCCTTAGTCCCGAACCCCGAATAACTCCCCTCCACCACTGTCCAATATTCTCCCA

19741  GAGGAGAGGAAGTGTATAGCAAGCATAGGTTGTCTTGTTATGCAGATAAAAAGTCTCTCA
19741  CTCCTCTCCTTCACATATCGTTCGTATCCAACAGAACAATACGTCTATTTTTCAGAGAGT

19801  GGTGATAAAAGTTGTCTCAGAGTCCCCCATCAGAAGAATAGGTAATAGCCTGTGACAAGG
19801  CCACTATTTTCAACAGAGTCTCAGGGGGTAGTCTTCTTATCCATTATCGGACACTGTTCC

19861  TCTACCTGTCCAATCTTCTCCCTGGTGATAAAATCTTCCCTGGTTGATGAGATTCTTATC
19861  AGATGGACAGGTTAGAAGAGGGACCACTATTTTAGAAGGGACCAACTACTCTAAGAATAG

19921  AAAGTTTTTTTCTTTATAGATATAAATTTTGTTTACAAAAAGACAGTTTTTCAGAGCTAC
19921  TTTCAAAAAAAGAAATATCTATATTTAAAACAAATGTTTTTCTGTCAAAAAGTCTCGATG

19981  TCCTGTGTCTACAGTGTCTCAGAATAACTAACTGGAAATACATCAAATAAGCATATCCTG
19981  AGGACACAGATGTCACAGAGTCTTATTGATTGACCTTTATGTAGTTTATTCGTATAGGAC

20041  ATGTGGCATATTCTGGTCATATTTTGGGATGGTGCGTCCTGAGCTCCAATGCTGAAAAAA
20041  TACACCGTATAAGACCAGTATAAAACCCTACCACGCAGGACTCGAGGTTACGACTTTTTT

20101  CAGGAACCCCATTTCTGTTGGATGTGCCAACCTAAAATAATCAAAAGGGTCAGAATCTAG
20101  GTCCTTGGGGTAAAGACAACCTACACGGTTGGATTTTATTAGTTTTCCCAGTCTTAGATC

20161  TTTAAAGATACTTTATTCAAGTGCAAAGTTTAAGGACAGTGCTAGGGAAACACGGATCCC
20161  AAATTTCTATGAAATAAGTTCACGTTTCAAATTCCTGTCACGATCCCTTTGTGCCTAGGG

20221  AAAGAATGGAAGTCAGCGTTCTGAACTTTAGAAGTTTGGGATTGTTTATAGAGAGAGAGT
20221  TTTCTTACCTTCAGTCGCAAGACTTGAAATCTTCAAACCCTAACAAATATCTCTCTCTCA

20281  TTAGGGAAGCTTAACAGAATTTCAACATCTTTCTGTGTAAGGCTTAATGCCTAGTTACAA
20281  AATCCCTTCGAATTGTCTTAAAGTTGTAGAAAGACACATTCCGAATTACGGATCAATGTT

20341  TGATCTGATTAGTCAAGGTGGTCACTTTCATTTGAGAAAGGCATATTTAACATTTCACGG
20341  ACTAGACTAATCAGTTCCACCAGTGAAAGTAAACTCTTTCCGTATAAATTGTAAAGTGCC

20401  GCTGGGAGCGGTGGTTCATGCCTGTAATCTCAGCACTTTGGGAGGCCGAGGCGGGCAGAT
20401  CGACCCTCGCCACCAAGTACGGACATTAGAGTCGTGAAACCCTCCGGCTCCGCCCGTCTA

20461  CATCTGAGGTCAGGAGTTCGAGGCCAGCCTGGCCAATATGGGGAAACCCCTTCTCTACTA
20461  GTAGACTCCAGTCCTCAAGCTCCGGTCGGACCGGTTATACCCCTTTGGGGAAGAGATGAT

20521  AAAATACAATACAATAAAATAAAATAAAATAAAAAAAATTAGCCGGGCATGGTGCACGTG
20521  TTTTATGTTATGTTATTTTATTTTATTTTATTTTTTTAATCGGCCCGTACCACGTGCAC

20581  CCTGTAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGCAG
20581  GGACATTAGGGTCGATGAACCCTTCGACTCCGTCCTCTTAACGAACTTGGACCCTCCGTC

20641  AGGTTGTGGTGAGCTGAGATTGCACCATTGCACTCCAGCCTGGGCAACAAGAATGAAACT
20641  TCCAACACCACTCGACTCTAACGTGGTAACGTGAGGTCGGACCCGTTGTTCTTACTTTGA

20701  CCATCTCAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAACATT
20701  GGTAGAGTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTTGTAA

20761  TCACACTGAAGGTGTAACAGTCATGGGTTTGTATTTTCGTTTTTCTTGGTTTTTTTTTTT
20761  AGTGTGACTTCCACATTGTCAGTACCCAAACATAAAAGCAAAAAGAACCAAAAAAAAAAA

20821  TTTTTTTGAGACTGTCTCGCTCTGTATTGCAGGCTGGCATGGCACAATCACAGCTCGCTG
20821  AAAAAAACTCTGACAGAGCGAGACATAACGTCCGACCGTACCGTGTTAGTGTCGAGCGAC

20881  CAGCTTCGATTTTGTGGGCTCAAGTGATCCTACCACCTCAGTCTCCCGACTAGCTGGGAC
20881  GTCGAAGCTAAAACACCCGAGTTCACTAGGATGGTGGAGTCAGAGGGCTGATCGACCCTG
```

FIGURE 4 (cont'd)

```
20941  CACAGTCATGTGCCACCACACCCGGCTAATTTATTTTTTATTTCTTATAGAGACCAGATC
20941  GTGTCAGTACACGGTGGTGTGGGCCGATTAAATAAAAAATAAAGAATATCTCTGGTCTAG

21001  CCCCTTTTGCCCAGGCTGGTCTCAAACTCCTGGACTCAAACAATTCTCCTGCCTCAGCCT
21001  GGGGAAAACGGGTCCGACCAGAGTTTGAGGACCTGAGTTTGTTAAGAGGACGGAGTCGGA

21061  CCCAAAGTGCCTGGATTACAGGCATGGGCTACTGTGCTCATCCCATGGGGTCTTTTGTGC
21061  GGGTTTCACGGACCTAATGTCCGTACCCGATGACACGAGTAGGGTACCCCAGAAAACACG

21121  CAACTGGTGTGAGTTAGGTACAGGACAATAAAGGAGGCAGTTAATCCATTAATCCATAAC
21121  GTTGACCACACTCAATCCATGTCCTGTTATTTCCTCCGTCAATTAGGTAATTAGGTATTG

21181  AAAGATCAGTGAAGGGGGAAGGTCTGGTCTCTGGTCTCTCCTAGTCATTTACAGAACTAG
21181  TTTCTAGTCACTTCCCCCTTCCAGACCAGAGACCAGAGAGGATCAGTAAATGTCTTGATC

21241  AACGATGAGAAGAGAGTGAATCTATAACCTAAGAAGCAGAATTGCAAACATGCTATGTGA
21241  TTGCTACTCTTCTCTCACTTAGATATTGGATTCTTCGTCTTAACGTTTGTACGATACACT

21301  CTCAGTTTCCAGTACTTAACTTCTCCCTTGGCATAATAAATAATTTTAGAGAGTTTTAAG
21301  GAGTCAAAGGTCATGAATTGAAGAGGGAACCGTATTATTTATTAAAATCTCTCAAAATTC

21361  AATTTTGTTTTCTTTTACAGATGTTAACAACCTCAGCTTGCTGGGGGCTGGTGGTGGTGG
21361  TTAAAACAAAAGAAAATGTCTACAATTGTTGGAGTCGAACGACCCCCGACCACCACCACC

21421  CTGCTCAGACAGTAAAGCATTGAATGCAATGTGAGGTCAAAGTTTTATTACAGATCATAC
21421  GACGAGTCTGTCATTTCGTAACTTACGTTACACTCCAGTTTCAAAATAATGTCTAGTATG

21481  CGGTCTTTTAAGTACTTGCAGCAAGATTTTCCTAATACTTAAAAGTGACTGGAAACTCAA
21481  GCCAGAAAATTCATGAACGTCGTTCTAAAAGGATTATGAATTTTCACTGACCTTTGAGTT

21541  TAGGATCAGTGTGATCTTCTCCAGATAGGGATGGTAGATTTTTGAAAGTGGGTGTGAAGG
21541  ATCCTAGTCACACTAGAAGAGGTCTATCCCTACCATCTAAAAACTTTCACCCACACTTCC

21601  GCAGTAATGAGAGAAAAATAAAACTATTTCTGGTTGTACTCTCCCAAGTCAAGGTCATTC
21601  CGTCATTACTCTCTTTTTATTTTGATAAAGACCAACATGAGAGGGTTCAGTTCCAGTAAG

21661  AATGAATGGTTATATAGATTTGCATCAAAAGTCCCTGTTTTCAGCCTCACCTTTGGAAAC
21661  TTACTTACCAATATATCTAAACGTAGTTTTCAGGGACAAAAGTCGGAGTGGAAACCTTTG

21721  ACCTGGATTCTCTATTCTGAGACTTTCTGAGATGAGGGAAAGGTGGCTTCTCATCAGTAT
21721  TGGACCTAAGAGATAAGACTCTGAAAGACTCTACTCCCTTTCCACCGAAGAGTAGTCATA

21781  CTCCTTCTTTGTGCAATACAGTTTCAGTTCTTTCTAATCTACAGAGTTAGTTCCTTCTCT
21781  GAGGAAGAAACACGTTATGTCAAAGTCAAGAAAGATTAGATGTCTCAATCAAGGAAGAGA

21841  ACTGTCAGCTTTCTACCTTCCAAAATGTGTTAATGTCACTCATTATCATGTTTCCCATTC
21841  TGACAGTCGAAAGATGGAAGGTTTTACACAATTACAGTGAGTAATAGTACAAAGGGTAAG

21901  TCTTTCCCTTATGGTTTAAACTTTAAAAGACAATGGTATTTGTTTAAGTGAATATCTAAA
21901  AGAAAGGGAATACCAAATTTGAAATTTTCTGTTACCATAAACAAATTCACTTATAGATTT

21961  AAGATGTGGAGAAAGATTCATGTGTTCAGTCTACCATGTTTAATTGGAGTCACTGGTAGA
21961  TTCTACACCTCTTTCTAAGTACACAAGTCAGATGGTACAAATTAACCTCAGTGACCATCT

22021  TTTTTAAGGAAAAATAACATTCTAGGTTTTTCCAGTATTGCTATGTAGCTTGTCTATGTA
22021  AAAAATTCCTTTTTATTGTAAGATCCAAAAAGGTCATAACGATACATCGAACAGATACAT

22081  GATCACCTTGCATTCTTCTGTTTCTCTCTTATAACTTGATTTTTTTCTTTTCTTTTCTTT
22081  CTAGTGGAACGTAAGAAGACAAAGAGAGAATATTGAACTAAAAAAAGAAAAGAAAAGAAA

22141  TCTTTTTTTTTCGAGACAGAGTCTCACTTTGTCACCCAGGCTGGAGTACAGTGGCACAAT
22141  AGAAAAAAAAAGCTCTGTCTCAGAGTGAAACAGTGGGTCCGACCTCATGTCACCGTGTTA

22201  CTCGTCTCATTGCTGCCTCCATCTCCTGGGTTCAAGTGATTCTCCCACCTCACCTTCCCA
22201  GAGCAGAGTAACGACGGAGGTAGAGGACCCAAGTTCACTAAGAGGGTGGAGTGGAAGGGT
```

FIGURE 4 (cont'd)

```
22261  AATAGCCGGGATTACAGACATCTTATAACTTGATTTTTACTCATTTCATTATCATCCAAG
22261  TTATCGGCCCTAATGTCTGTAGAATATTGAACTAAAAATGAGTAAAGTAATAGTAGGTTC

22321  AGATGAAATTTAACTCAAAGCAAATGACTATTGTTTCAAATACAAATCTTAATCAGGGAG
22321  TCTACTTTAAATTGAGTTTCGTTTACTGATAACAAAGTTTATGTTTAGAATTAGTCCCTC

22381  AAAAGATTATAACTTATAGTTGAACTGGGGTTTGGAGATTATATGTGTTTTTTAAAATTA
22381  TTTTCTAATATTGAATATCAACTTGACCCCAAACCTCTAATATACACAAAAAATTTTAAT

22441  TTCTGTCCCACATGTAACATACAAGTGAGGATTTTAAAAATTTGTCTGCAGTTTTTTTTT
22441  AAGACAGGGTGTACATTGTATGTTCACTCCTAAAATTTTTAAACAGACGTCAAAAAAAAA

22501  TTTTAATTGATCATTCTTGGGTGTTTCTCGCAGAGGGGGATTTGGCAGGGTCATAGGACA
22501  AAAATTAACTAGTAAGAACCCACAAAGAGCGTCTCCCCCTAAACCGTCCCAGTATCCTGT

22561  ATAGTGGAGGGAAGGTCAGCAGATAAACGAGTGAACAAAGGTCTCTGGTTTTCCTAGGCA
22561  TATCACCTCCCTTCCAGTCGTCTATTTGCTCACTTGTTTCCAGAGACCAAAAGGATCCGT

22621  GAGGACCCTGCGGCCTTCCGCAGTGTTTGTGTCCCTGGGTACTTGAGATTAGGGAGTGGT
22621  CTCCTGGGACGCCGGAAGGCGTCACAAACACAGGGACCCATGAACTCTAATCCCTCACCA

22681  GATGACTCTTAAAGAGCATGCTGCCTTCAAGCATCTGTTTAACAAAGCACATCTTGCACC
22681  CTACTGAGAATTTCTCGTACGACGGAAGTTCGTAGACAAATTGTTTCGTGTAGAACGTGG

22741  GCCCTTAATCCATTTAACCCTGAGTGGACACAGCACATGTTTCAGAGAGCACCGGGTTGG
22741  CGGGAATTAGGTAAATTGGGACTCACCTGTGTCGTGTACAAAGTCTCTCGTGGCCCAACC

22801  GGGTAAGGTCATAGATCAACAGCATCCCAAGGCAGAAGAATTTTTCTTAGTACAGAACAA
22801  CCCATTCCAGTATCTAGTTGTCGTAGGGTTCCGTCTTCTTAAAAAGAATCATGTCTTGTT

22861  AATGGAGTCTCCTATGTCTACTTCTTTCTACACAGACACAGCAACAATCTGATTTCTCTA
22861  TTACCTCAGAGGATACAGATGAAGAAAGATGTGTCTGTGTCGTTGTTAGACTAAAGAGAT

22921  TCTTTTCCCCACATTTCCCCCTTTTCTATTCCACAAAACCGCCATTGTCATCATGGCCCG
22921  AGAAAAGGGGTGTAAAGGGGGAAAAGATAAGGTGTTTTGGCGGTAACAGTAGTACCGGGC

22981  TTCTCAATGAGCTGTTGGGTACACCTCCCAGACGGGGTGGCGGCCGGGCAGAGGGGCTCC
22981  AAGAGTTACTCGACAACCCATGTGGAGGGTCTGCCCCACCGCCGGCCCGTCTCCCCGAGG

23041  TCACTTCCCAGAAGGGGTGGCCTGGCAGAGGCACCCCCAACCTCCCTCCCGGACGGGGCG
23041  AGTGAAGGGTCTTCCCCACCGGACCGTCTCCGTGGGGGTTGGAGGGAGGGCCTGCCCCGC

23101  GCTGGCCAGGCGGGGGCTGCCCCCAACTTCCCAGACGGGGTGGCTGCTGGGCGGAGGGGC
23101  CGACCGGTCCGCCCCCGACGGGGGTTGAAGGGTCTGCCCCACCGACGACCCGCCTCCCCG

23161  TCTTTACTTCTCAGATGGGGCGGCTGCCGGGTGGAGGGGCTCCTCACTTCTCAGACGGGG
23161  AGAAATGAAGAGTCTACCCCGCCGACGGCCCACCTCCCCGAGGAGTGAAGAGTCTGCCCC

23221  CGGCCGGGCAGAGGCGCTCCTCACCTCCCAGACGGGGCAGCAGGGCAGAGGCGCTCCCCA
23221  GCCGGCCCGTCTCCGCGAGGAGTGGAGGGTCTGCCCCGTCGTCCCGTCTCCGCGAGGGGT

23281  CATCTCAGACGATGGGCGGCCGGGCAGAGACGCTCCTCACTTCCTAGACGGGATGGCAGC
23281  GTAGAGTCTGCTACCCGCCGGCCCGTCTCTGCGAGGAGTGAAGGATCTGCCCTACCGTCG

23341  CGGGAAGAGGCGCTCCTCACTTCCCAGACTGGGCAGCCAGGCAGAGGGGCTCCTCACATC
23341  GCCCTTCTCCGCGAGGAGTGAAGGGTCTGACCCGTCGGTCCGTCTCCCCGAGGAGTGTAG

23401  CCAGACGATGGGCAGCCAGGCAGAGAAGCTCCTCACTTCCCAGATGGGGTGGCGGCCGGG
23401  GGTCTGCTACCCGTCGGTCCGTCTCTTCGAGGAGTGAAGGGTCTACCCCACCGCCGGCCC

23461  CAGAGGCTGCAATCTCGGCACTTTGGGAGGCCAAGGCAGGCGGCTGGGAGGTGGAGGTTG
23461  GTCTCCGACGTTAGAGCCGTGAAACCCTCCGGTTCCGTCCGCCGACCCTCCACCTCCAAC

23521  TAGCGAGCCGAGATCACGCCACTACACTCCAGCTTGGGCAACATTGAGCACTGAGTGAAC
23521  ATCGCTCGGCTCTAGTGCGGTGATGTGAGGTCGAACCCGTTGTAACTCGTGACTCACTTG
```

FIGURE 4 (cont'd)

```
23581  GAGACTCCGTCTGCAATCCCGGCACCTCGGGAGGCCGAGGCTGGCAGATCACTCCCGGTT
23581  CTCTGAGGCAGACGTTAGGGCCGTGGAGCCCTCCGGCTCCGACCGTCTAGTGAGGGCCAA

23641  AGGAGCTGGAGACCAGCCCGGCCAACACAGCGAAACCCCGTCTCCACCAAAAAAATACGA
23641  TCCTCGACCTCTGGTCGGGCCGGTTGTGTCGCTTTGGGGCAGAGGTGGTTTTTTTATGCT

23701  AAACCAGTCAGGCATGGTGGCGCGCGCCTGCAATCACAGGCACTCGGCAGGCTGAGGCAG
23701  TTTGGTCAGTCCGTACCACCGCGCGCGGACGTTAGTGTCCGTGAGCCGTCCGACTCCGTC

23761  GACAATCAGGCAGGGAGGTTGCAGTGAGCAGAGATGGCAGCAGTACAGTCCAGCTTCGGC
23761  CTGTTAGTCCGTCCCTCCAACGTCACTCGTCTCTACCGTCGTCATGTCAGGTCGAAGCCG

23821  TTGGCATCAGAGGGAGACCGTGGAAAGAGAGGGAGAGGGAGACCGTGGGGAGAGGGAGAG
23821  AACCGTAGTCTCCCTCTGGCACCTTTCTCTCCCTCTCCCTCTGGCACCCCTCTCCCTCTC

23881  GGGGAGGGGGAGCTGCAATTTTTTATTTAGCAATTTCCAAACCTATTGTGTTAGGCAGGG
23881  CCCCTCCCCCTCGACGTTAAAAAATAAATCGTTAAAGGTTTGGATAACACAATCCGTCCC

23941  TTCTCCAGGGAAACAGAACCAATAGGGCATATCTATATCTATATTTATATCTTTATCTAT
23941  AAGAGGTCCCTTTGTCTTGGTTATCCCGTATAGATATAGATATAAATATAGAAATAGATA

24001  ATCTATATCTATAATCTTTCTATCTATCTATCAAGAGATCGAGAGAGAGAGATTTATTTT
24001  TAGATATAGATATTAGAAAGATAGATAGATAGTTCTCTAGCTCTCTCTCTCTAAATAAAA

24061  AAGGAATTGGTTCATGTGATTGTGGGAGCTGGCAAGTCTGAAATCCATAGGGCAAGTTGG
24061  TTCCTTAACCAAGTACACTAACACCCTCGACCGTTCAGACTTTAGGTATCCCGTTCAACC

24121  TCAGCTGGAAATTCTGGTAAGAGTTGATGTTGCAGTCGTGAGTCTGAAATTTGCAGGGCA
24121  AGTCGACCTTTAAGACCATTCTCAACTACAACGTCAGCACTCAGACTTTAAACGTCCCGT

24181  GGTCAGGTACAGTATCTATTGCAATTGCAGTCCTGAGGCCAAATTCCTTCTTCAGGAAAC
24181  CCAGTCCATGTCATAGATAACGTTAACGTCAGGACTCCGGTTTAAGGAAGAAGTCCTTTG

24241  CTCCGTCTTTTTTGCTGTTAAGGCCTTCAACTAATTGGATGAAGCCCTCCCACATTGAGG
24241  GAGGCAGAAAAAACGACAATTCCGGAAGTTGATTAACCTACTTCGGGAGGGTGTAACTCC

24301  GCAATCTGCTTTACTCAAAGTCTGCTGATTTAAATGGTAATCACATCTAAAAAATACTTT
24301  CGTTAGACGAAATGAGTTTCAGACGACTAAATTTACCATTAGTGTAGATTTTTTATGAAA

24361  CACAGCAACATTTGAACTTGTGTTGTCTAAGCAATTGGGCACTATAGCTTAGCCAAGCAG
24361  GTGTCGTTGTAAACTTGAACACAACAGATTCGTTAACCCGTGATATCGAATCGGTTCGTC

24421  ACACATGAAATTAACCATCACTCTGTAGAGATGAAAGAACTACTGTTTCAAAAAACAACT
24421  TGTGTACTTTAATTGGTAGTGAGACATCTCTACTTTCTTGATGACAAAGTTTTTTGTTGA

24481  GTTTTTGAATCCGTCATTTATCAACGGATTCACCCAATGGTTAACCTTTCTATATGCTTT
24481  CAAAAACTTAGGCAGTAAATAGTTGCCTAAGTGGGTTACCAATTGGAAAGATATACGAAA

24541  CTCTTTCTCCCAAGTACTGAATCATATGAAAATAAATTATGGTTGTGATATTTCCTATCT
24541  GAGAAAGAGGGTTCATGACTTAGTATACTTTTATTTAATACCAACACTATAAAGGATAGA

24601  AAATACTTCATATGGGACATTCTCCTACATAGCACACCTAAGAAGATTAGTATTAATTCA
24601  TTTATGAAGTATACCCTGTAAGAGGATGTATCGTGTGGATTCTTCTAATCATAATTAAGT

24661  GAGAGATCATGTCATATCTGGTTCATATATAATTTTCCCTAGCTGACCTCAAAATGTGTT
24661  CTCTCTAGTACAGTATAGACCAAGTATATATTAAAAGGGATCGACTGGAGTTTTACACAA

24721  TGTTTTGTTTACATATTACTCTTTATAGTTATGTTTCTTTTAATACGGATCAATATTTGT
24721  ACAAAACAAATGTATAATGAGAAATATCAATACAAAGAAAATTATGCCTAGTTATAAACA

24781  GTGTTTTTGCTGCTATTTTTTTTAATGACATTGTTTTGAATAATTCAGGCTCATTGGCTA
24781  CACAAAAACGACGATAAAAAAAATTACTGTAACAAAACTTATTAAGTCCGAGTAACCGAT

24841  TTTCACATCCTGAATTTCTCTAATTTTATCTTTATGATTAGAGCCAGGCTCTAATTCTTG
24841  AAAGTGTAGGACTTAAAGAGATTAAAATAGAAATACTAATCTCGGTCCGAGATTAAGAAC
```

FIGURE 4 (cont'd)

```
24901  GCAAGAATAACACTGAAGTGAGGTTGTGTACTCCTAATTGCCCTGCATCAAAGACACATG
24901  CGTTCTTATTGTGACTTCACTCCAACACATGAGGATTAACGGGACGTAGTTTCTGTGTAC

24961  ATATGCATTGGTTCCTGCATTAGTCATGTTAAGTTTGATTAGCTGGTGAAGGTGGTCGCC
24961  TATACGTAACCAAGGACGTAATCAGTACAATTCAAACTAATCGACCACTTCCACCAGCGG

25021  ACCATTTCTCACTCTTGTAAAGGGACATTTTCTCCTTTTAAATTAATATGTAATTTGTGG
25021  TGGTAAAGAGTGAGAACATTTCCCTGTAAAAGAGGAAAATTTAATTATACATTAAACACC

25081  CACAAGACTTTGAGACTGTGTAAATATCCCATTCTCTGACAACCTTTCTTCTGATGACTT
25081  GTGTTCTGAAACTCTGACACATTTATAGGGTAAGAGACTGTTGGAAAGAAGACTACTGAA

25141  TAGCATTAACTGATCTAATTATATCATTCTTGGCATTGAGTAGCATTCTTGTGTAAAGAA
25141  ATCGTAATTGACTAGATTAATATAGTAAGAACCGTAACTCATCGTAAGAACACATTTCTT

25201  GGGGTTCTCTTTCTCTCCCTCCTCTCATTATAATTTGTTACTGTTATTTTCTTACAAGTT
25201  CCCCAAGAGAAAGAGAGGGAGGAGAGTAATATTAAACAATGACAATAAAAGAATGTTCAA

25261  ACTAAGAATTACTCAAAGGCAGTATTTTTCAATAAATCATTGGTAATTTATGGTCTTCTC
25261  TGATTCTTAATGAGTTTCCGTCATAAAAAGTTATTTAGTAACCATTAAATACCAGAAGAG

25321  ACTCTAGAAAAGGGAGGATGGGTCAGAAAGCTGAGGGGAAAAAAAGGGAGGAAACGAGAA
25321  TGAGATCTTTTCCCTCCTACCCAGTCTTTCGACTCCCCTTTTTTCCCTCCTTTGCTCTT

25381  TGTATATTATCGACATAATCGGTTGAAAATGTCTAGTGCTGTAATTATGCCATCATGCAG
25381  ACATATAATAGCTGTATTAGCCAACTTTTACAGATCACGACATTAATACGGTAGTACGTC

25441  ATTAAAAAGCCACCTGTGAAAAATCAACTTGTATTGTTTTGTTTTGTTTGATTTTTAATA
25441  TAATTTTTCGGTGGACACTTTTTAGTTGAACATAACAAAACAAAACAAACTAAAAATTAT

25501  AATCCTTATAGTGTATGAATAAAAATTAAATTAATTGATCAGTATCTTCACTAATGGCAG
25501  TTAGGAATATCACATACTTATTTTTAATTTAATTAACTAGTCATAGAAGTGATTACCGTC

25561  AGAAGGGGATGGACTGAGCAGATTCTGCTTTTCACTCTGATATTTGTTAATTGCCCTGTG
25561  TCTTCCCCTACCTGACTCGTCTAAGACGAAAAGTGAGACTATAAACAATTAACGGGACAC

25621  TGGTCTGAAGACCTGGGTTCTAGTACCAGCTCTCCCTCTGCAGACAGGTGTGTTACTTTG
25621  ACCAGACTTCTGGACCCAAGATCATGGTCGAGAGGGAGACGTCTGTCCACACAATGAAAC

25681  GTCAAGGCAATCAATCTTCTTGATACTCAGTTTGGTTCTCTCTGATTTGAGCATAAGAAA
25681  CAGTTCCGTTAGTTAGAAGAACTATGAGTCAAACCAAGAGAGACTAAACTCGTATTCTTT

25741  TATACTTGAAGGCCGGGCATGGTGGCTTATGCCTGTAATTCCAGCACTATGGGAAGCCGA
25741  ATATGAACTTCCGGCCCGTACCACCGAATACGGACATTAAGGTCGTGATACCCTTCGGCT

25801  GGTGGGAGGATCGCTTGGGCTCAGGAGTTTGAGACCAGGCTGGGCAACATGGCAAAACCT
25801  CCACCCTCCTAGCGAACCCGAGTCCTCAAACTCTGGTCCGACCCGTTGTACCGTTTTGGA

25861  CAACTCTACAAAAAACAAACAAAAAATTAGCCAGGCATGGTGGTGCACGCCTGTAGTCCC
25861  GTTGAGATGTTTTTTGTTTGTTTTTTAATCGGTCCGTACCACCACGTGCGGACATCAGGG

25921  AGCTACTCAGGAGGTTGAGGCAGGAGAATCCTGGGAGGCAGAGGTTGCAGTGATCTGAGA
25921  TCGATGAGTCCTCCAACTCCGTCCTCTTAGGACCCTCCGTCTCCAACGTCACTAGACTCT

25981  TTGTGCCACTGCACTCCAGCCTGGGAGCCTGGGTGACAGAGAGAGACCCTGTCTAAAATA
25981  AACACGGTGACGTGAGGTCGGACCCTCGGACCCACTGTCTCTCTCTGGGACAGATTTTAT

26041  TATATATACACACACTTGAATGCTTCATGTCCCTTGCTCAGTGACAGCCCCGTCCTCTGC
26041  ATATATATGTGTGTGAACTTACGAAGTACAGGGAACGAGTCACTGTCGGGGCAGGAGACG

26101  TTCTGCACAGTTGACCAAGGCTTCTCAGTCTACTCCTGTGTAGAAAAGAATTAAAGCAGG
26101  AAGACGTGTCAACTGGTTCCGAAGAGTCAGATGAGGACACATCTTTTCTTAATTTCGTCC

26161  CCTGAGGCTGCTAGTCTTAGAAAGGCTTGCAAAGTTGGCCCTTGTCTGGTGTCTGGGAAC
26161  GGACTCCGACGATCAGAATCTTTCCGAACGTTTCAACCGGGAACAGACCACAGACCCTTG
```

FIGURE 4 (cont'd)

```
26221  TTGAATTTCTGGAGGGTTCTCACCATTCCCTGAAAAGAATGGCTCACGATGCCTAAAATA
26221  AACTTAAAGACCTCCCAAGAGTGGTAAGGGACTTTTCTTACCGAGTGCTACGGATTTTAT

26281  TTTGTACAAACAATGAAGTTTCTGCTGAACATCTGCTTCCTTCTGGGAGTCTGGAATTTT
26281  AAACATGTTTGTTACTTCAAAGACGACTTGTAGACGAAGGAAGACCCTCAGACCTTAAAA

26341  AGGATGTGCTAGGTAGGGGTGTTTACATGACAGCCCCACATAAAAACCTTGGGCACTGAG
26341  TCCTACACGATCCATCCCCACAAATGTACTGTCGGGGTGTATTTTTGGAACCCGTGACTC

26401  TGTCTAATGAGATTCCCTAGTAAGTAATATTTTACACATATTATCACAACTCATTGCTGG
26401  ACAGATTACTCTAAGGGATCATTCATTATAAAATGTGTATAATAGTGTTGAGTAACGACC

26461  AGGAGTTAAGTGCATCCTGTGTGACTCCACTGGGAGAGGACTCTGGGAAGCACATGTCTG
26461  TCCTCAATTCACGTAGGACACACTGAGGTGACCCTCTCCTGAGACCCTTCGTGTACAGAC

26521  ATTTCCTTTGTCCCATGTGCCTTTTCCTTTTGCTGATTTTTGCCTTGTAAATTTTTTTTT
26521  TAAAGGAAACAGGGTACACGGAAAAGGAAAACGACTAAAAACGGAACATTTAAAAAAAAA

26581  TTTTTTTTGAGAGACAGAGTCTTACTCTGTTGCCCAGACTGGAGTTCAGTGGCCCCATGT
26581  AAAAAAAACTCTCTGTCTCAGAATGAGACAACGGGTCTGACCTCAAGTCACCGGGGTACA

26641  CGGCTCACTGCAGCCTCTGCCTCCTGGATTCAAGCAATTCTCCTGCCTCAGCCTCCCGAG
26641  GCCGAGTGACGTCGGAGACGGAGGACCTAAGTTCGTTAAGAGGACGGAGTCGGAGGGCTC

26701  TAGCTGGGACTACAGGCGCGCGCTGCCACACCGGCTAATTTTTGTATTTTTAGTAGAGAC
26701  ATCGACCCTGATGTCCGCGCGCGACGGTGTGGCCGATTAAAAACATAAAAATCATCTCTG

26761  GGGGTTTCACCATGTTGGCCAGGATGGTCTTGATCTCCTGACCTCGTGATCCGCCTGCCT
26761  CCCCAAAGTGGTACAACCGGTCCTACCAGAACTAGAGGACTGGAGCACTAGGCGGACGGA

26821  CGGCCTCCCAAAGTGCTGGAATTATAGGCCTGGGCCACTGTGCCTGGCCATAATTTTTAT
26821  GCCGGAGGGTTTCACGACCTTAATATCCGGACCCGGTGACACGGACCGGTATTAAAAATA

26881  CCTGTAATAAATTATAGTCGAAAGTATTGAGTCCTGTGAATCAACTTAGAGAATCACCAA
26881  GGACATTATTTAATATCAGCTTTCATAACTCAGGACACTTAGTTGAATCTCTTAGTGGTT

26941  ACTGCTTCCAATTTTCTCATATGTAAAATGGGTTGCTATAATAATGATGATATCTACCTT
26941  TGACGAAGGTTAAAAGAGTATACATTTTACCCAACGATATTATTACTACTATAGATGGAA

27001  ATAGGGTTGTTTTAAGTATGAAAAATATGGTGAAATATGTGTGAAACTAGGGGCCGGATG
27001  TATCCCAACAAAATTCATACTTTTTATACCACTTTATACACACTTTGATCCCCGGCCTAC

27061  TAGGATGAGTCATACAGTGGGTGGTTATTATTACGGATCATATTAACTCTTCAGGTAAAA
27061  ATCCTACTCAGTATGTCACCCACCAATAATAATGCCTAGTATAATTGAGAAGTCCATTTT

27121  ACCATCTTGTAGCTGGGCAAGGTGGCACATGCCTGTAGTCCCAGCTACTTGGGGGACTGA
27121  TGGTAGAACATCGACCCGTTCCACCGTGTACGGACATCAGGGTCGATGAACCCCCTGACT

27181  GGCAGGTGGATCGCTTGAGCCTAGGAGTTCAAGGCTGTAGTACGCTATGATCACGTCTAT
27181  CCGTCCACCTAGCGAACTCGGATCCTCAAGTTCCGACATCATGCGATACTAGTGCAGATA

27241  CAATAGCCACTGGACTCCAGCCTGGACAATGTAATGAGAGTCTGTCTCAAAAAAAAGGAA
27241  GTTATCGGTGACCTGAGGTCGGACCTGTTACATTACTCTCAGACAGAGTTTTTTTTCCTT

27301  CAATCTTGTGTTTGCGTGTGTGCTCATGTCTTTGTTGGAATAATACTGACACAGTCACTC
27301  GTTAGAACACAAACGCACACACGAGTACAGAAACAACCTTATTATGACTGTGTCAGTGAG

27361  AAATTAACTCTCAAATCTTTGATTATTCATTTGATGGGAACAACAGGTTTATGTGCCAGC
27361  TTTAATTGAGAGTTTAGAAACTAATAAGTAAACTACCCTTGTTGTCCAAATACACGGTCG

27421  TTATCAGGACCCCTAGTTCTTTCATTAAAGTGTGTATGTGTGTGTATGTGTGTATGTG
27421  AATAGTCCTGGGGATCAAGAAAGTAATTTCACACATACACACACACATACACACATACAC

27481  TGCACATGTGTGCTAGTTTTCTTATCTAACGACGAAGTCCTTTGGGAGGTTTACTTTTTC
27481  ACGTGTACACACGATCAAAAGAATAGATTGCTGCTTCAGGAAACCCTCCAAATGAAAAAG
```

FIGURE 4 (cont'd)

```
27541  ATTACTATTTTTTTCATGCTTTCCTTTGATGTGCACCTGGCAATTTATAGAGCCAATGTG
27541  TAATGATAAAAAAAAGTACGAAAGGAAACTACACGTGGACCGTTAAATATCTCGGTTACAC

27601  TACTACATATTCTGCTGCAGAAAATGACAGAGGTGAGGAAGGAGCAGGAAGCAGATTTAG
27601  ATGATGTATAAGACGACGTCTTTTACTGTCTCCACTCCTTCCTCGTCCTTCGTCTAAATC

27661  AAAGGAATGTTACAGTTTATCAAGTTGTGACTAATGGAAGCTATTAAAACAAAAGGTACC
27661  TTTCCTTACAATGTCAAATAGTTCAACACTGATTACCTTCGATAATTTTGTTTTCCATGG

27721  AACCCATCTTAAATATTCACTGCCTCTACAACTATTTTTCTATAACTGTTGTTCCTTGTC
27721  TTGGGTAGAATTTATAAGTGACGGAGATGTTGATAAAAAGATATTGACAACAAGGAACAG

27781  AGCTCTGGGGTGGAGGGAATATTCCTGAATGTTTTATTAAGTAATTCAAAGATCTTCCTA
27781  TCGAGACCCCACCTCCCTTATAAGGACTTACAAAATAATTCATTAAGTTTCTAGAAGGAT

27841  GGTATGTGTATCAAAGATCCTATGATGTATCTGTCTTGTAAATTGAGACTGGCATATATT
27841  CCATACACATAGTTTCTAGGATACTACATAGACAGAACATTTAACTCTGACCGTATATAA

27901  GGGCTTACATTACTGGGAAAAATGTATGTCATCATCATTGGTGAAAAAGAGTTTCCATAG
27901  CCCGAATGTAATGACCCTTTTTACATACAGTAGTAGTAACCACTTTTTCTCAAAGGTATC

27961  CAGTTTTTTCTAAGTAGAGAGATAAATGAATCAGACTTTTGGAGGAGGTATGTTTTAACT
27961  GTCAAAAAAGATTCATCTCTCTATTTACTTAGTCTGAAAACCTCCTCCATACAAAATTGA

28021  CACTGGAGGATCACACCAAGTCTGAACATCTAAGGAACCCTAGTGGAAAACTGGTTTTCC
28021  GTGACCTCCTAGTGTGGTTCAGACTTGTAGATTCCTTGGGATCACCTTTTGACCAAAAGG

28081  TGTATGCTGCTCTAGCGATCTGAGGAAAGGCCTAAAAAAAGCCAGTGTCCCTGTGATATT
28081  ACATACGACGAGATCGCTAGACTCCTTTCCGGATTTTTTTCGGTCACAGGGACACTATAA

28141  GTGAAATACATATATGTGGTCTTTGTTCCTGTTTCCTGGCATACAAGTACTAAAAATCCT
28141  CACTTTATGTATATACACCAGAAACAAGGACAAAGGACCGTATGTTCATGATTTTTAGGA

28201  TGGAATCTCCAAAGTGCTGTCTTTTTGTATATTAGTGTTGACTGATAGCTTCAGAGTGGG
28201  ACCTTAGAGGTTTCACGACAGAAAAACATATAATCACAACTGACTATCGAAGTCTCACCC

28261  GCTGGTCAAGTTGATCACCAATGGCCAGTGGTTTAATCAGTCATGCCTATGTAATGAAGC
28261  CGACCAGTTCAACTAGTGGTTACCGGTCACCAAATTAGTCAGTACGGATACATTACTTCG

28321  CTCCATGAAAATCCAAAGGGAATGGACTGGGGAGCTTCTAAGAAGTGTTTTTTACCTTCC
28321  GAGGTACTTTTAGGTTTCCCTTACCTGACCCCTCGAAGATTCTTCACAAAAAATGGAAGG

28381  CGTGAGCCTCCTGTGCTGAACACGTGGAGGGTCACGGGAAGGTAAGCAAAAATTCATTCA
28381  GCACTCGGAGGACACGACTTGTGCACCTCCCAGTGCCCTTCCATTCGTTTTTAAGTAAGT

28441  CATGCTGGGAGGGTGGCACATCCCAACTCCACGAGGACAAAAGCTCCTGTGCTCAGAACC
28441  GTACGACCCTCCCACCGTGTAGGGTTGAGGTGCTCCTGTTTTCGAGGACACGAGTCTTGG

28501  CTTCCAGACTTGGACCTTGTATCTCTTCATTGGCTTTTATTTGTATTTTAAAAAATATCC
28501  GAAGGTCTGAACCTGGAACATAGAGAAGTAACCGAAAATAAACATAAAATTTTTTATAGG

28561  TGAGTGATAAACTGGTAGATGTAAGTGTCTTCCTCAGTTCTGTGAACCACGCTAGCAAAT
28561  ACTCACTATTTGACCATCTACATTCACAGAAGGAGTCAAGACACTTGGTGCGATCGTTTA

28621  TAATCAAACCGGAAGAGGGGGTCATGGCAACCCCAACTTGAAACCTGTTGGTCAGCAGTT
28621  ATTAGTTTGGCCTTCTCCCCCAGTACCGTTGGGGTTGAACTTTGGACAACCAGTCGTCAA

28681  CCAGAGACCTGAACTTATGACTGGTTGGAAGGACGGGGTGCTCTTGTGGGACTGAGCCCT
28681  GGTCTCTGGACTTGAATACTGACCAACCTTCCTGCCCCACGAGAACACCCTGACTCGGGA

28741  CAACTTGTGGATCTGATGCTATCTCTGGGTAGACAGTGTTGGAGCTGAATTGGAGGACAC
28741  GTTGAACACCTAGACTACGATAGAGACCCATCTGTCACAACCTCGACTTAACCTCCTGTG

28801  CTAGCTTGTGTCCACTGCAGAACTGATTGCTTGCTTCCTGCTGGGGAGAAATCCCTATAT
28801  GATCGAACACAGGTGACGTCTTGACTAACGAACGAAGGACGACCCCTCTTTAGGGATATA
```

FIGURE 4 (cont'd)

```
28861  ATTTTGGGGTCCAGAAGTCTTCCGTGTTGATTGTTTTTGTGTGAGAGCAGAGGAAAAATG
28861  TAAAACCCCAGGTCTTCAGAAGGCACAACTAACAAAAACACACTCTCGTCTCCTTTTTAC

28921  GTTTGAGAGTTTTTCGGGAAACAGCCCCTCTTAGTATAGTGGTAGAAACACTCCAAAAAT
28921  CAAACTCTCAAAAAGCCCTTTGTCGGGGAGAATCATATCACCATCTTTGTGAGGTTTTTA

28981  AAAATATGCTACCCCAAAAATAAAAAATGAAAATAAAAATAAAAATCTGGAGTGGTCAGT
28981  TTTTATACGATGGGGTTTTTATTTTTTACTTTTATTTTTATTTTTAGACCTCACCAGTCA

29041  TTGCCAAGAGGATTTATGGAGCATCTAAGGTCAAGCATTACTTCTCAGCCCTGCAGCAGG
29041  AACGGTTCTCCTAAATACCTCGTAGATTCCAGTTCGTAATGAAGAGTCGGGACGTCGTCC

29101  CCTCTAATTAAGCCATGCCACAGAGTGCGGTTTGCAGAGAGGTGGAGTTCCTGTGGATTC
29101  GGAGATTAATTCGGTACGGTGTCTCACGCCAAACGTCTCTCCACCTCAAGGACACCTAAG

29161  ATTCTTCTCTCTAAAATTATAAATAACGCAGCTTCCAAACATGCCATATTACCTCACAAG
29161  TAAGAAGAGAGATTTTAATATTTATTGCGTCGAAGGTTTGTACGGTATAATGGAGTGTTC

29221  GTCTAAAGACTCAATATGAATCTTCAACTTAACAATTATAACTGACACAGATTGAGATCT
29221  CAGATTTCTGAGTTATACTTAGAAGTTGAATTGTTAATATTGACTGTGTCTAACTCTAGA

29281  TCCTGGATGCTGAAATAGGCTTTACATGATTTATCTCATTTAGTTCTCATAAAACTCTAT
29281  AGGACCTACGACTTTATCCGAAATGTACTAAATAGAGTAAATCAAGAGTATTTTGAGATA

29341  TATTAATATTATTTTACAGATAGAGAAATGGAAGCACAGAGAGGGTAAGTGTATTAGTCC
29341  ATAATTATAATAAAATGTCTATCTCTTTACCTTCGTGTCTCTCCCATTCACATAATCAGG

29401  ATTTTTATGCTGCTAATAAAGACATACCCAAGACAGGGTAATTTATACAGAAAAAGAGGT
29401  TAAAAATACGACGATTATTTCTGTATGGGTTCTGTCCCATTAAATATGTCTTTTTCTCCA

29461  TTAATGGACTCACATTTCCACATGGCTGGGAGGCCTCACAATCATGGCAGAAGGCAAAAG
29461  AATTACCTGAGTGTAAAGGTGTACCGACCCTCCGGAGTGTTAGTACCGTCTTCCGTTTTC

29521  GCACTTCTTACCTGGTGGTGGCAAGACAGAAAATGAGAACCAAGCGAAAGGGGTTTCCCC
29521  CGTGAAGAATGGACCACCACCGTTCTGTCTTTTACTCTTGGTTCGCTTTCCCCAAAGGGG

29581  TTATAAAACCATCAGGTAACATGAGACAATTATTCACTACCACAGAAACTATATGGGGGA
29581  AATATTTTGGTAGTCCATTGTACTCTGTTAATAAGTGATGGTGTCTTTGATATACCCCCT

29641  AACCGCCCCCATGATTCAGTTATCTTCCATTGGGTCCCTCCCACAACACGTGGAGCTACA
29641  TTGGCGGGGGTACTAAGTCAATAGAAGGTAACCCAGGGAGGGTGTTGTGCACCTCGATGT

29701  ATTCAAGATGAGATTTGTGTGGAGACATAGCTAACCCATATCAGTAAGTAATTAGACAAA
29701  TAAGTTCTACTCTAAACACACCTCTGTATCGATTGGGTATAGTCATTCATTAATCTGTTT

29761  CCTTGCAAAGCTAGTAAGTGGTATTGCTGGAATTAAATCTGCCAATGTCACTCCAGAGTC
29761  GGAACGTTTCGATCATTCACCATAACGACCTTAATTTAGACGGTTACAGTGAGGTCTCAG

29821  TATGTGATTTACTGGTATATTATAGGGTCTACCTATAAGCTTGATAACCTTATGTTAGAA
29821  ATACACTAAATGACCATATAATATCCCAGATGGATATTCGAACTATTGGAATACAATCTT

29881  CAAATAACCATGACATATTCTGTTCAGTAGTTTGCATTTCCTTTTGACTCCTACAGGGGA
29881  GTTTATTGGTACTGTATAAGACAAGTCATCAAACGTAAAGGAAAACTGAGGATGTCCCCT

29941  GGAGGCCGAGGTTCAGGGACATTAAAGTGGGGTTTAGAGACATTAGGTGGTTTACCCAAG
29941  CCTCCGGCTCCAAGTCCCTGTAATTTCACCCCAAATCTCTGTAATCCACCAAATGGGTTC

30001  CTTTATCTACTAGATACTAGATAATTGTCTATCTGACAATAAAGCTAGAACTTACACCCA
30001  GAAATAGATGATCTATGATCTATTAACAGATAGACTGTTATTTCGATCTTGAATGTGGGT

30061  GCTGGACAATGGAGCTACAGCTTAAAACAAAATCTTCTAATGAGCTCAGAACTCATGGCT
30061  CGACCTGTTACCTCGATGTCGAATTTTGTTTTAGAAGATTACTCGAGTCTTGAGTACCGA

30121  TAATAAGATTTGGGAAAAGTCCTGCGGACATTTCACTGTCTATCCAGAAATGCGAGGAAG
30121  ATTATTCTAAACCCTTTTCAGGACGCCTGTAAAGTGACAGATAGGTCTTTACGCTCCTTC
```

FIGURE 4 (cont'd)

```
30181  TTAGGAGAAAAGGCAAAGTATCTCATAGACAGATAAAGGCAGTTCAGTAAGTAGCCTGAA
30181  AATCCTCTTTTCCGTTTCATAGAGTATCTGTCTATTTCCGTCAAGTCATTCATCGGACTT

30241  TAAATAGCCATATTTAAAATCTAATTTGGTTCTTTTGTGCCTAACACAATTGACTTTATG
30241  ATTTATCGGTATAAATTTTAGATTAAACCAAGAAAACACGGATTGTGTTAACTGAAATAC

30301  CTGGTATCTATTTTCTGCTCTGTCTCTCTTTCTGTCTCTACTTATGTGTGTATGCATGTG
30301  GACCATAGATAAAAGACGAGACAGAGAGAAAGACAGAGATGAATACACACATACGTACAC

30361  TTATAAAATATATCAGCCCATTGATCAATTTTTAATTTAAGCAAAAACATGGTAGATTAA
30361  AATATTTTATATAGTCGGGTAACTAGTTAAAAATTAAATTCGTTTTTGTACCATCTAATT

30421  TATATTCAATCAATTTATTGAGTAGATATTTATCGAGGCTCACACATTTATCTGACTGTG
30421  ATATAAGTTAGTTAAATAACTCATCTATAAATAGCTCCGAGTGTGTAAATAGACTGACAC

30481  TACTAAGCACTGGAAACCCAAAGAGAATGAAGTCACTGCCTAGACAGTTGGAGGCTGACA
30481  ATGATTCGTGACCTTTGGGTTTCTCTTACTTCAGTGACGGATCTGTCAACCTCCGACTGT

30541  CATAAACAAGCAGGGTAATAATCAAGTGTCATAACTAGTCTATTACATGCAATGTTAAAC
30541  GTATTTGTTCGTCCCATTATTAGTTCACAGTATTGATCAGATAATGTACGTTACAATTTG

30601  TAGCATCTCTACCATGCAAAACTTGTGCTCTTGGGGATACAGGAAGACTTTCCAGAGAAG
30601  ATCGTAGAGATGGTACGTTTTGAACACGAGAACCCCTATGTCCTTCTGAAAGGTCTCTTC

30661  TACACAGGACAGGAATGTGTTAGGAAACCAGTCAGTTCCCAGGTCCTCGATCAGGTCACA
30661  ATGTGTCCTGTCCTTACACAATCCTTTGGTCAGTCAAGGGTCCAGGAGCTAGTCCAGTGT

30721  GTGACTTGCGATTCTATTTTGAACATAGGTCTTAAATCTTTATTTTGATAGCAAGGGTTT
30721  CACTGAACGCTAAGATAAAACTTGTATCCAGAATTTAGAAATAAAACTATCGTTCCCAAA

30781  TAAACTTTTGATATATGATATACTTCTGAAAATCAAATTTCATATTCTAAAATTAAGTAT
30781  ATTTGAAAACTATATACTATATGAAGACTTTTAGTTTAAAGTATAAGATTTTAATTCATA

30841  TTTTGCCCTTAACTAACTTGAACATAATACTATAATACATTAACATATATTATTAATAAT
30841  AAAACGGGAATTGATTGAACTTGTATTATGATATTATGTAATTGTATATAATAATTATTA

30901  AAATAGGACCCCTGGAAGTCCAAGAGAGACATGAGGCTTATTTGGTATGTTAAAATCATA
30901  TTTATCCTGGGGACCTTCAGGTTCTCTCTGTACTCCGAATAAACCATACAATTTTAGTAT

30961  CAGGAAGTGTTGTTAAATAAGAAATGGTGTTACACTTTCTTTGAGTTATATTTATATGGA
30961  GTCCTTCACAACAATTTATTCTTTACCACAATGTGAAAGAAACTCAATATAAATATACCT

31021  TATGTTGTTAATATGTGTTCCAACATTGTATGAGATTCCTAAAAATCTGATATGGACATC
31021  ATACAACAATTATACACAAGGTTGTAACATACTCTAAGGATTTTTAGACTATACCTGTAG

31081  CAGTGTGGTTGTTGGGTCCTCTATGAGCCTAGGGAGCACTGGCCGGTTACCTTGCATCCA
31081  GTCACACCAACAACCCAGGAGATACTCGGATCCCTCGTGACCGGCCAATGGAACGTAGGT

31141  TTCATGTCAGCGTTGGTAGTGTCAGGCACCCGATGGTTCTTAGAAGAGTACTTAGCTCTA
31141  AAGTACAGTCGCAACCATCACAGTCCGTGGGCTACCAAGAATCTTCTCATGAATCGAGAT

31201  CTCTGCAGTGTCTGAAGTTGCTGGATATATACCTGCTGTATAAACTGCTGACTAGGGCGC
31201  GAGACGTCACAGACTTCAACGACCTATATATGGACGACATATTTGACGACTGATCCCGCG

31261  TGCAGTTTACTTACTGTCTTCTCATGGGGACCTTCCCCTTCAACTCTTTCCTCTTGGGTT
31261  ACGTCAAATGAATGACAGAAGAGTACCCCTGGAAGGGGAAGTTGAGAAAGGAGAACCCAA

31321  TCACATCTTGGGTGGGGAGTTTCATCCTAGTAGTTTGCCTGAGAATACAGGTGAACCCAC
31321  AGTGTAGAACCCACCCCTCAAAGTAGGATCATCAAACGGACTCTTATGTCCACTTGGGTG

31381  AGAACAAAGTGGACTTCGAAAGCATCTCCTTGGAGTGAGACTTTGCTTCTTTCTCTTTGC
31381  TCTTGTTTCACCTGAAGCTTTCGTAGAGGAACCTCACTCTGAAACGAAGAAAGAGAAACG

31441  TAGCACCATCTTGTACCTTATCGTCATGAACTTTGTTGGCTGAATCATTCCCATTTACTT
31441  ATCGTGGTAGAACATGGAATAGCAGTACTTGAAACAACCGACTTAGTAAGGGTAAATGAA

31501  AATTGAAGAGTAAGAGACTGGAACAATGCTCACTTTGATTTTCCTGGATAAGAGTTAAGA
31501  TTAACTTCTCATTCTCTGACCTTGTTACGAGTGAAACTAAAAGGACCTATTCTCAATTCT
```

FIGURE 4 (cont'd)

```
31561  GTTCTTGAGATGGCAGCTTGTTTGACACATGAATTTTCTTCAAATTTGTGCTTACTACCA
31561  CAAGAACTCTACCGTCGAACAAACTGTGTACTTAAAAGAAGTTTAAACACGAATGATGGT

31621  ACTGATTTGGTGTGGAGGAGAGCCCAGAGAAGTTCCCTCTCTCTGTCAGAACAACTTTGT
31621  TGACTAAACCACACCTCCTCTCGGGTCTCTTCAAGGGAGAGAGACAGTCTTGTTGAAACA

31681  AACATTTATTAACCTGACTTCTGCCTTCAATTAACTGTAACCTTTTGCCTTCCAAATTAA
31681  TTGTAAATAATTGGACTGAAGACGGAAGTTAATTGACATTGGAAAACGGAAGGTTTAATT

31741  AAAGTTCCACATTACTCCAAAAATAAAAAGTGAAAATAAAAATAAAAGGCTGATATGTCA
31741  TTTCAAGGTGTAATGAGGTTTTTATTTTTCACTTTTATTTTTATTTTCCGACTATACAGT

31801  TGGTATATGCAGTAAATTGCTTTATTCTGATTTTTTTTTCTAAAAGCCTTTTGCAAATCC
31801  ACCATATACGTCATTTAACGAAATAAGACTAAAAAAAAAAGATTTTCGGAAAACGTTTAGG

31861  TGAAGTGCTGTGTATTCAAGGAAATTCATGAAAGGGACCTTGACAAGTAGTCTTAAATAC
31861  ACTTCACGACACATAAGTTCCTTTAAGTACTTTCCCTGGAACTGTTCATCAGAATTTATG

31921  AGGTTTCTGATAACTTTGGAGATCACACCTTTGGACTAGGTAAAAACATCCAAAACTCGT
31921  TCCAAAGACTATTGAAACCTCTAGTGTGGAAACCTGATCCATTTTTGTAGGTTTTGAGCA

31981  CTGGGCGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGCGGATC
31981  GACCCGCGTCACCGAGTGCGGACATTAGGGTCGTGAAACCCTCCGGTTCCGTCCGCCTAG

32041  ACGAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCATCTCTACTAAAA
32041  TGCTCCAGTCCTCTAGCTCTGGTAGGACCGATTGTGCCACTTTGGGGTAGAGATGATTTT

32101  ACACAAAAACAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTTCCAGCTACCCGGCA
32101  TGTGTTTTTGTTTTAATCGGCCCGCACCACCGCCCGCGGACATCAAGGTCGATGGGCCGT

32161  GGCTGAGGCGGAAGAATGGCGCGAAGCGGCAAGACGGAGCTTGCAGTGAGCCGAGATCGC
32161  CCGACTCCGCCTTCTTACCGCGCTTCGCCGTTCTGCCTCGAACGTCACTCGGCTCTAGCG

32221  GCCACTGCACTCCAGCCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGTCCAAAGCTCTAATG
32221  CGGTGACGTGAGGTCGGAGTTTTTTTTTTTTTTTTTTTTTTTTCAGGTTTCGAGATTAC

32281  AAAAACTGATGCATTCATGCAGATTGCTAACCCAACATCAAGGAGAACGAGAATTACATG
32281  TTTTTGACTACGTAAGTACGTCTAACGATTGGGTTGTAGTTCCTCTTGCTCTTAATGTAC

32341  GGACTGAGTGAATGAAGGATTGAAATTACTATTTTAAAGGGTTTTTGTTTGAAACATTAC
32341  CCTGACTCACTTACTTCCTAACTTTAATGATAAAATTTCCCAAAAACAAACTTTGTAATG

32401  TGATTTTTTTTATGTTTTGTTTTCCAGACTCAAGAAAATATTTTTTTTCTTTTGAGCTAT
32401  ACTAAAAAAAATACAAAACAAAAGGTCTGAGTTCTTTTATAAAAAAAAGAAAACTCGATA

32461  TTATAGCTTACAGCATTGGGGTAAAGTATACTTCTGTGAGCAAAATGGAAACATTTACCT
32461  AATATCGAATGTCGTAACCCCATTTCATATGAAGACACTCGTTTTACCTTTGTAAATGGA

32521  TTCTTTCTACCTGAATTCTTCAGAATTTGGAAACTATTCATTAGTGTTCTTATTTTCTGG
32521  AAGAAAGATGGACTTAAGAAGTCTTAAACCTTTGATAAGTAATCACAAGAATAAAAGACC

32581  CAATATAGTTATTTGCATAAGTTCAATAAGAACCTATTTTCGGCTGTGCATGGTGGCTCA
32581  GTTATATCAATAAACGTATTCAAGTTATTCTTGGATAAAAGCCGACACGTACCACCGAGT

32641  TGCCTGTAATCCCAGCACCTTGGGAGGGCGAGGCCGGTGGATCACAAGGTCAGGAGTTTG
32641  ACGGACATTAGGGTCGTGGAACCCTCCCGCTCCGGCCACCTAGTGTTCCAGTCCTCAAAC

32701  AGAACAGCCTGGCCAAGATGGTGAAACCCCGTCTCTACTAAAACTACAAAGATTAACCAG
32701  TCTTGTCGGACCGGTTCTACCACTTTGGGGCAGAGATGATTTTGATGTTTCTAATTGGTC

32761  GCGCGGTGGCAGGCGCCTGTAATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCTT
32761  CGCGCCACCGTCCGCGGACATTAGGGTCGATGAGCCCTCCGACTCCGTCCTCTTATCGAA

32821  GAACCTTGGGTGGCAGAGGTTGCAGTGAGCCGAGATCGCACCACTGCGCTCCAGCCTAGG
32821  CTTGGAACCCACCGTCTCCAACGTCACTCGGCTCTAGCGTGGTGACGCGAGGTCGGATCC
```

FIGURE 4 (cont'd)

```
32881  GGACAAGAGCGAGACTCTATCTCAAAACAAACAAACAAAAAAACAAAAAGCAAACAAACA
32881  CCTGTTCTCGCTCTGAGATAGAGTTTTGTTTGTTTGTTTTTTTGTTTTTCGTTTGTTTGT

32941  AAAAAAGAGCTTATTTTCTTTTGTAACTGGACACATGGAAACAGTGGTTATTTTACCAAG
32941  TTTTTTCTCGAATAAAAGAAAACATTGACCTGTGTACCTTTGTCACCAATAAAATGGTTC

33001  GCTTTGACTGGAATGTCTTTTTTTTGAGTGCAACCGGACTGCTTCGAGGGATTGAGGTTG
33001  CGAAACTGACCTTACAGAAAAAAAACTCACGTTGGCCTGACGAAGCTCCCTAACTCCAAC

33061  ACTTTATACAGCCAGTAGACTTGGAAAAAGACACACCTGGTACCTTGCCCACACAATTCC
33061  TGAAATATGTCGGTCATCTGAACCTTTTTCTGTGTGGACCATGGAACGGGTGTGTTAAGG

33121  TTTATAGCGTTTCTAACTTTGCAATAAGTAAAGAATGTCACTTTCTGACAGGCCCAGGGA
33121  AAATATCGCAAAGATTGAAACGTTATTCATTTCTTACAGTGAAAGACTGTCCGGGTCCCT

33181  CCTCAAGATATTTTGGAGACGTTGAAAAGAGAGCAATTCACCCAATTCATACAGGTTTTA
33181  GGAGTTCTATAAAACCTCTGCAACTTTTCTCTCGTTAAGTGGGTTAAGTATGTCCAAAAT

33241  CAGGCACAGTCTGATGGCGAATCTTTGGCTTGGCTTCATTGCTTCAAGAGTTTGTTTTTT
33241  GTCCGTGTCAGACTACCGCTTAGAAACCGAACCGAAGTAACGAAGTTCTCAAACAAAAAA

33301  TGACACGGAATCTCACTCTGTCACCCATTCTGGAGTGCAGTGGTGAGATCTTGGCTCACT
33301  ACTGTGCCTTAGAGTGAGACAGTGGGTAAGACCTCACGTCACCACTCTAGAACCGAGTGA

33361  GAAACCTCTGCTTCCTGGATTCAAGTGATTCTCCTGCTTCAGCCTCCCGAGTAGCTGGGA
33361  CTTTGGAGACGAAGGACCTAAGTTCACTAAGAGGACGAAGTCGGAGGGCTCATCGACCCT

33421  TTACAGGCATGTGTTGCTACTCCCAGCTAATTTTTTGTATTTTTGGTAGAGATGAGGTTT
33421  AATGTCCGTACACAACGATGAGGGTCGATTAAAAAACATAAAAACCATCTCTACTCCAAA

33481  CACCATGTTGCCCAGGGTGGTCTTGAACTCCTGACCTCAGGTGATCCACCCACCTCAGCC
33481  GTGGTACAACGGGTCCCACCAGAACTTGAGGACTGGAGTCCACTAGGTGGGTGGAGTCGG

33541  TCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGTGCCCAGCCCTTGAGAGGTTTCAAA
33541  AGGGTTTCACGACCCTAATGTCCGCACTCGGTGACACGGGTCGGGAACTCTCCAAAGTTT

33601  AGTCTTAATCTGAGATTCCTTATGAAAAAATTCCAACAAAGCCAACTTAAGAGCCATCCC
33601  TCAGAATTAGACTCTAAGGAATACTTTTTTAAGGTTGTTTCGGTTGAATTCTCGGTAGGG

33661  TGTAGAAAATATATCTATTTAAATATCGTCTGTTTTATGCCACAATGTTTGTTTTGCGTT
33661  ACATCTTTTATATAGATAAATTTATAGCAGACAAAATACGGTGTTACAAACAAAACGCAA

33721  GCAAGCCGTTATCTCCTAGGCATGGGTAATCCACTTTGAGAGGAAAATCCAGACATCATT
33721  CGTTCGGCAATAGAGGATCCGTACCCATTAGGTGAAACTCTCCTTTTAGGTCTGTAGTAA

33781  GTGCCTAATTTGCTGCCACCTGTGTCTGTGCTGGATCTTACCAGGGTTAGCTTTATTTTC
33781  CACGGATTAAACGACGGTGGACACAGACACGACCTAGAATGGTCCCAATCGAAATAAAAG

33841  TTTCAACTGGCAAAATTCTGATAAAGCATGTAGCTTCACGGTAGAGCAAAAAGTGTAACT
33841  AAAGTTGACCGTTTTAAGACTATTTCGTACATCGAAGTGCCATCTCGTTTTTCACATTGA

33901  TAATTGAAAGTCATGACTTTCAGGAGTCTAATGATGCTAATATATTTTCTTCTAGAGAGG
33901  ATTAACTTTCAGTACTGAAAGTCCTCAGATTACTACGATTATATAAAAGAAGATCTCTCC

33961  TGTCTTTATTGTATGTAATTCCCTGAAGAAATGTAAAGATTTATCTAATTCCTATTGCTC
33961  ACAGAAATAACATACATTAAGGGACTTCTTTACATTTCTAAATAGATTAAGGATAACGAG

34021  TGACAGGCGTGCTCAATAGAAGCATTGAAGTAAAATAAAAATAACCTTTTTAATCTGTGT
34021  ACTGTCCGCACGAGTTATCTTCGTAACTTCATTTTATTTTTATTGGAAAAATTAGACACA

34081  ATTTCTCTTGCCCACCAAGATATTTAATTTTTTTCTTGATCTTTGCTTTTGTTTATGTGA
34081  TAAAGAGAACGGGTGGTTCTATAAATTAAAAAAAGAACTAGAAACGAAAACAAATACACT

34141  AGATATTTATGTCTATATCATTTCTGTATATATCACAGAAAAGTTTTCACTTGCTTGACA
34141  TCTATAAATACAGATATAGTAAAGACATATATAGTGTCTTTTCAAAAGTGAACGAACTGT

34201  TTTTATTCTACAATGATTTTTAGTTTAAGTTTAAAGAATGGCTACATTCAGATTTATTTT
34201  AAAATAAGATGTTACTAAAAATCAAATTCAAATTTCTTACCGATGTAAGTCTAAATAAAA
```

FIGURE 4 (cont'd)

```
34261  CACAACTATAGTCCTGCTAAACTGCCTTTGTTGAAAACTCAGTTTTTATAATAAATAAAA
34261  GTGTTGATATCAGGACGATTTGACGGAAACAACTTTTGAGTCAAAAATATTATTTATTTT

34321  AATATCAAAGCAAAAAAAAAAGAGCCTTCATGATAAATCACCATTCTTACTGCACTTTAA
34321  TTATAGTTTCGTTTTTTTTTTCTCGGAAGTACTATTTAGTGGTAAGAATGACGTGAAATT

34381  GCAGATAATCATGCCAGAGACTAAACTGACTTTGCAAATAAATTAGTTTTATTATTATTA
34381  CGTCTATTAGTACGGTCTCTGATTTGACTGAAACGTTTATTTAATCAAAATAATAATAAT

34441  TCTTTGTGGGTAAAAGTTGGGATGACTGTAGAAGAGAAAATTATATTTCAGAAGAAAATT
34441  AGAAACACCCATTTTCAACCCTACTGACATCTTCTCTTTTAATATAAAGTCTTCTTTTAA

34501  ATACCATACTTGTTATTAGATTCTAGTCTTGATAATTGTTTTTTTGAGTTTTTGTTATTT
34501  TATGGTATGAACAATAATCTAAGATCAGAACTATTAACAAAAAAACTCAAAAACAATAAA

34561  GCCTACAATTTGGGCTGAATCCTGAATTCCTTCCTGGGTACAAGTCTCCAAACTAATGTT
34561  CGGATGTTAAACCCGACTTAGGACTTAAGGAAGGACCCATGTTCAGAGGTTTGATTACAA

34621  TCCTAATTTTTCTACCATTTTTCTGTCTTGGAATCACCAAGCCCTGCAGGCTTTAGCTAG
34621  AGGATTAAAAAGATGGTAAAAAGACAGAACCTTAGTGGTTCGGGACGTCCGAAATCGATC

34681  ACAACTTGATATAAACTTTGGAAGAAATAATGACAGCAACTTAATATATGAGCAGTGTTC
34681  TGTTGAACTATATTTGAAACCTTCTTTATTACTGTCGTTGAATTATATACTCGTCACAAG

34741  ATGTCAGCTGATGTATGGATTACTCAGAAGGTTTACTTGAACACCTGATTCAAACTATAA
34741  TACAGTCGACTACATACCTAATGAGTCTTCCAAATGAACTTGTGGACTAAGTTTGATATT

34801  TCCAGAAATATCTGTAAGATTGCCACTGCAATCTGAAGCTACTTCAGAGACTCTAGGAAA
34801  AGGTCTTTATAGACATTCTAACGGTGACGTTAGACTTCGATGAAGTCTCTGAGATCCTTT

34861  ATTAGTTTATAGACAACTCCAGATGTTAATCTTTGTTTTTCTTTTGTTTTCATGGAAATG
34861  TAATCAAATATCTGTTGAGGTCTACAATTAGAAACAAAAAGAAAACAAAAGTACCTTTAC

34921  CCTTTTATTAAACATCTGATTGCTTGATTCATAGAGGCCTGACTTTGGTGGAAGCAACAC
34921  GGAAAATAATTTGTAGACTAACGAACTAAGTATCTCCGGACTGAAACCACCTTCGTTGTG

34981  CACTGCCTGAAATGAGATATAACTGTTAACCGTTTAATTGGAGCGGCCTGTTCTCACGAC
34981  GTGACGGACTTTACTCTATATTGACAATTGGCAAATTAACCTCGCCGGACAAGAGTGCTG

35041  TGCAAGACTAGTTCAATGGCTTATAAAATAATCCACCCGCTGGGCACAGTAGTTCACACC
35041  ACGTTCTGATCAAGTTACCGAATATTTTATTAGGTGGGCGACCCGTGTCATCAAGTGTGG

35101  TGTAATCTCAGCACTTTGGGAGGCTGAGGTGGGCAAATTGCTTGAGCTCAGGAGTTCGAG
35101  ACATTAGAGTCGTGAAACCCTCCGACTCCACCCGTTTAACGAACTCGAGTCCTCAAGCTC

35161  ACCAGCCAGAGCAACATGGTGAAACCCTGTCTCTACCAAAAATATAAAAAATTAGCCGGG
35161  TGGTCGGTCTCGTTGTACCACTTTGGGACAGAGATGGTTTTTATATTTTTTAATCGGCCC

35221  CATAGTAGCATGTGCCTGTGGTCCCAGCTCCTCGGGAGGTGGATGTGGGAGGATTGCCTG
35221  GTATCATCGTACACGGACACCAGGGTCGAGGAGCCCTCCACCTACACCCTCCTAACGGAC

35281  AGCCTGGGAGGCAGAGCTTGCAGTGAGCTGAGATTATGACACTGCACTCTAGCCTGGGCA
35281  TCGGACCCTCCGTCTCGAACGTCACTCGACTCTAATACTGTGACGTGAGATCGGACCCGT

35341  ATACAGTGAGACCCCATCTCAAAAAAAAAAAAAAAAAGATAATCTACCAACCCAATTTCTC
35341  TATGTCACTCTGGGGTAGAGTTTTTTTTTTTTTTTCTATTAGATGGTTGGGTTAAAGAG

35401  AACTGGGGAATGTTGGGCTCCTGCTCTCCTGTCTCTGACCTTTCATTCTCCCTTGAGGCT
35401  TTGACCCCTTACAACCCGAGGACGAGAGGACAGAGACTGGAAAGTAAGAGGGAACTCCGA

35461  AGCCATCAAGACTACAACCCCTCTTCCCCAAAGTGGGCCATAGAAACAAAACCCCTTTTC
35461  TCGGTAGTTCTGATGTTGGGGAGAAGGGGTTTCACCCGGTATCTTTGTTTTGGGGAAAAG

35521  CCCAAAGCCAACCATAAAACCTAAAAATAGTACTCTAACTTCCACCACCCCCAGCCTGTC
35521  GGGTTTCGGTTGGTATTTTGGATTTTTATCATGAGATTGAAGGTGGTGGGGGTCGGACAG
```

FIGURE 4 (cont'd)

```
35581  TGTGTAAGGCCATAAAGAAATTATCTGACTCACTTTGTTTGACTGAAGGTTATAAGACCC
35581  ACACATTCCGGTATTTCTTTAATAGACTGAGTGAAACAAACTGACTTCCAATATTCTGGG

35641  CCATTCCAGAGAGGGTCCAGCACCACACCCAGAAGGAAGAAATGCATGCTCAGAGAGGCA
35641  GGTAAGGTCTCTCCCAGGTCGTGGTGTGGGTCTTCCTTCTTTACGTACGAGTCTCTCCGT

35701  AGAAGGGTCCAGACAGACAGGCTGTGCTGGATTTCCCTGCTCAGTCTATTAGCATTAGAC
35701  TCTTCCCAGGTCTGTCTGTCCGACACGACCTAAAGGGACGAGTCAGATAATCGTAATCTG

35761  CATACCTTTTTTGTCCAATCATATTTCTACGTGGCTCTCCATACTTTGTTTAACCTAACC
35761  GTATGGAAAAAACAGGTTAGTATAAAGATGCACCGAGAGGTATGAAACAAATTGGATTGG

35821  TAAGGGGGCTCAGGGTTTTGGCCTGGGCATTGTCTGTTGGGGAATAGAGTGAGTCATCCC
35821  ATTCCCCCGAGTCCCAAAACCGGACCCGTAACAGACAACCCCTTATCTCACTCAGTAGGG

35881  CAGCTCATGGGTTTGCATCCAGTTCTTGTTGTAAAAGGCCCAAAGCCTGTTGGATGGCAA
35881  GTCGAGTACCCAAACGTAGGTCAAGAACAACATTTTCCGGGTTTCGGACAACCTACCGTT

35941  CCCTGAGCCATCGTGGAAGGGGGGTTCCAGTTTCACAAATAGATACTCAGACAGCCAAAC
35941  GGGACTCGGTAGCACCTTCCCCCCAAGGTCAAAGTGTTTATCTATGAGTCTGTCGGTTTG

36001  TACCATCTACTTGGTGCCAACGTTTGCACTGTGGTCAAAGACTTACCTAGCACAGACTGA
36001  ATGGTAGATGAACCACGGTTGCAAACGTGACACCAGTTTCTGAATGGATCGTGTCTGACT

36061  ACAAATCTTCCCATCTGTCACATAAATGTCCCCAAGCAATGTTGAAGCACATGCCAGGAT
36061  TGTTTAGAAGGGTAGACAGTGTATTTACAGGGGTTCGTTACAACTTCGTGTACGGTCCTA

36121  CGGCCTTGCTAGGTCCTGCTTGGTAGATAAGCAAATGGCTTCCTTGTGGTGTTTTTATTC
36121  GCCGGAACGATCCAGGACGAACCATCTATTCGTTTACCGAAGGAACACCACAAAAATAAG

36181  TATTTTGTCTCATTAACACTACAACTTTGTGTTATCTACTTGATAATCTGTAATTGTAAA
36181  ATAAAACAGAGTAATTGTGATGTTGAAACACAATAGATGAACTATTAGACATTAACATTT

36241  TACATACAGGATTATGTAATTTGTGTAAATACATAATGACAGACTTCTGAAAACTGGTAT
36241  ATGTATGTCCTAATACATTAAACACATTTATGTATTACTGTCTGAAGACTTTTGACCATA

36301  TTTTTACTTGGTGTCCATTTGAAGCTGTTTTAGGTTCTGTGGCACATGAACCAATGGCAT
36301  AAAAATGAACCACAGGTAAACTTCGACAAAATCCAAGACACCGTGTACTTGGTTACCGTA

36361  TTAAAAAGCATACTCTCCTTGCTCTTAAAAAAACAAACAAAACAAAACAAAAACTGTCCA
36361  AATTTTTCGTATGAGAGGAACGAGAATTTTTTTGTTTGTTTTGTTTTGTTTTTGACAGGT

36421  TTAAAATTTTTGTATTTTTGTAAAGACAGGGTTTCACCATAAAAATGGACAGTTTCCCTT
36421  AATTTTAAAAACATAAAAACATTTCTGTCCCAAAGTGGTATTTTTACCTGTCAAAGGGAA

36481  ATATCTTTGGGTCTTCGTTCTGTAGGCTCTCGTGTCACATAAAATTATGATCAAATAAAT
36481  TATAGAAACCCAGAAGCAAGACATCCGAGAGCACAGTGTATTTTAATACTAGTTTATTTA

36541  TTGTATGACTTTTCTCCTATTAATCTGCCCCTTGTCAATGATTTTGAGCAACCTTCTGAA
36541  AACATACTGAAAAGAGGATAATTAGACGGGGAACAGTTACTAAAACTCGTTGGAAGACTT

36601  GAAAGTGATGGGGAAGTATTCCCTTTGCTGCCACAAGGTATATATCAGATAAGTCAGTCT
36601  CTTTCACTACCCCTTCATAAGGGAAACGACGGTGTTCCATATATAGTCTATTCAGTCAGA

36661  ATTTTCTTGAACATATATCACATTCATAGGAAAAACACTTTTTCCAAAAGATTACTGTCT
36661  TAAAAGAACTTGTATATAGTGTAAGTATCCTTTTTGTGAAAAAGGTTTTCTAATGACAGA

36721  TCCCAAATACCCACGTTCTCCATTTTTTTTTTTTTTTTTTAGAGATGGGGTCTCACTGTG
36721  AGGGTTTATGGGTGCAAGAGGTAAAAAAAAAAAAAAAAAATCTCTACCCCAGAGTGACAC

36781  TTGTCCAGGCTAGTCTCAAGCTTCTGGACTCAAGCAATCCTCTCACCTCAGCTTCCCAAA
36781  AACAGGTCCGATCAGAGTTCGAAGACCTGAGTTCGTTAGGAGAGTGGAGTCGAAGGGTTT

36841  GTGCTGTGATTACAGACATAAGCCATCACGCCCGGCCCCATTTTCTCCTTTGGGTGTATG
36841  CACGACACTAATGTCTGTATTCGGTAGTGCGGGCCGGGGTAAAAGAGGAAACCCACATAC

36901  TATATTGCTCATTGTCAGGCATCAACCATTTCTCTTCTCTGAGATGGAGTCTCACTCTGT
36901  ATATAACGAGTAACAGTCCGTAGTTGGTAAAGAGAAGAGACTCTACCTCAGAGTGAGACA
```

FIGURE 4 (cont'd)

```
36961  CACCCAGGCTGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAAGCTCTGCCTCCCTGGTT
36961  GTGGGTCCGACCTCACGTCACCGCGTTAGAACCGAGTGACGTTCGAGACGGAGGGACCAA

37021  CACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGCCCACCACTACG
37021  GTGCGGTAAGAGGACGGAGTCGGAGGGCTCATCGACCCTGATGTCCACGGGTGGTGATGC

37081  CCCGGCTAATTTTTTTGTATTTCTTTTTTGGTAGAGACGGGGTTTCACCGTGTTAGACAG
37081  GGGCCGATTAAAAAAACATAAAGAAAAAACCATCTCTGCCCCAAAGTGGCACAATCTGTC

37141  GATGGTCTCAATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGAT
37141  CTACCAGAGTTAGAGGACTGGAGCACTAGGCGGGCGGAGCCGGAGGGTTTCACGACCCTA

37201  TACAGGTGTGAGCCACTGCGCCCAGCCCTCTTCTAATTTTTTATATTTATCATGTATTTA
37201  ATGTCCACACTCGGTGACGCGGGTCGGGAGAAGATTAAAAAATATAAATAGTACATAAAT

37261  CTAGAACTGGTTTATTTTCGTGAGCATATGTTGTGGATTCTACTATAAATAATGCACTTA
37261  GATCTTGACCAAATAAAAGCACTCGTATACAACACCTAAGATGATATTTATTACGTGAAT

37321  ATAATCATTAATAAAACTTAAATCCTTTCCGTCTTGATTGTGGCCTTTTCTTTTTGCTAT
37321  TATTAGTAATTATTTTGAATTTAGGAAAGGCAGAACTAACACCGGAAAAGAAAAACGATA

37381  TAATTTCCTCAGCATAGGTTAACCTGAACTCTTAAAAGCTATTTAGGTACTGGTATAAAA
37381  ATTAAAGGAGTCGTATCCAATTGGACTTGAGAATTTTCGATAAATCCATGACCATATTTT

37441  CAAAATTTTCTCCAAAGCAATCCAGTATATAATATCTGGATTTTAAAAAAGCAGCCTCAT
37441  GTTTTAAAAGAGGTTTCGTTAGGTCATATATTATAGACCTAAAATTTTTTCGTCGGAGTA

37501  TTAATTAGGATTAATTTCTCCTTTCAGAAGTTAAGTGATTCTGTTTCTGTAAAACGCAAA
37501  AATTAATCCTAATTAAAGAGGAAAGTCTTCAATTCACTAAGACAAAGACATTTTGCGTTT

37561  ATTATAAGCTTAATTGTTTTCCACACAGGAAATGAAATTCCTAATCTTTTTATCCAGAAT
37561  TAATATTCGAATTAACAAAAGGTGTGTCCTTTACTTTAAGGATTAGAAAAATAGGTCTTA

37621  CTTCATGTGTTTGTCTCTCTGTCCATCTGTCACTCCTTTTTAAGGTACCTGGAGATGACC
37621  GAAGTACACAAACAGAGAGACAGGTAGACAGTGAGGAAAAATTCCATGGACCTCTACTGG

37681  TAGAGGCTTGCCACGCTTAGGGTTTCCTCCCTGTGCAGAAGAAGATGGGAAGTGCAAATG
37681  ATCTCCGAACGGTGCGAATCCCAAAGGAGGGACACGTCTTCTTCTACCCTTCACGTTTAC

37741  AGGACGAGGGGTGGGGAGGGAAGAGGGAGTCAGGTGATTCAGCAAGTGCAGGAGCCCACA
37741  TCCTGCTCCCCACCCCTCCCTTCTCCCTCAGTCCACTAAGTCGTTCACGTCCTCGGGTGT

37801  GACTAAGCTCAGGAATTTGCACTTTGTCTGCTCGCAGAGTGAGTTTCATCGTGGTGTGTA
37801  CTGATTCGAGTCCTTAAACGTGAAACAGACGAGCGTCTCACTCAAAGTAGCACCACACAT

37861  TATATTGAGATATGCAACAGACTCCACTGGAATGTCAGAAGAACTTATGTTTATTTTTCC
37861  ATATAACTCTATACGTTGTCTGAGGTGACCTTACAGTCTTCTTGAATACAAATAAAAAGG

37921  TCTAACATGTGGGAAAGATTAAACTGTATTTGTATTTAACTCACATAGATTGGCACGTGC
37921  AGATTGTACACCCTTTCTAATTTGACATAAACATAAATTGAGTGTATCTAACCGTGCACG

37981  TCATTCTGTGTGTCACCCTGCCACGCTGTCACAAACTTCGAGCAGGGTCCTGAGGGGAGA
37981  AGTAAGACACACAGTGGGACGGTGCGACAGTGTTTGAAGCTCGTCCCAGGACTCCCCTCT

38041  GGAGGGCATAGCTCCTGCTCCCAGGGGGTTGCCTGTGTAGCCCTGCGCATTCATTCTCTT
38041  CCTCCCGTATCGAGGACGAGGGTCCCCCAACGGACACATCGGGACGCGTAAGTAAGAGAA

38101  CCAGAGGGTTGTGATTTGTGGCGGTCTGTGTGAGCCTGGTTAAGGAGAGTTACAGATTAT
38101  GGTCTCCCAACACTAAACACCGCCAGACACACTCGGACCAATTCCTCTCAATGTCTAATA

38161  GTCATCTGGTGTTAACTGATAATCTTTGCAAAAGGGACATTGGCTTTAAGAGATGTCTGC
38161  CAGTAGACCACAATTGACTATTAGAAACGTTTTCCCTGTAACCGAAATTCTCTACAGACG

38221  AAAAAAAACTCTAAACAAACATAAATAAAAATTTAAAAGGATAAAGCCAGAATAATAATG
38221  TTTTTTTTGAGATTTGTTTGTATTTATTTTTAAATTTTCCTATTTCGGTCTTATTATTAC
```

FIGURE 4 (cont'd)

```
38281  AAAATATTAGTAATGTAAAACAAGGAAATGTGTCATGTTTTTCATGATAACTAATATTTT
38281  TTTTATAATCATTACATTTTGTTCCTTTACACAGTACAAAAAGTACTATTGATTATAAAA

38341  TTCAATTTTTTTTTTTTTTTAAGACAGTCTCACTCTGTCACCAGGCTGGAGTCAGTAGTG
38341  AAGTTAAAAAAAAAAAAAAAATTCTGTCAGAGTGAGACAGTGGTCCGACCTCAGTCATCAC

38401  CGATCTTGGCTCACTGCAACCTCCACCTCCCAGGTTCACACAATTCTCATGTCTCAGCCT
38401  GCTAGAACCGAGTGACGTTGGAGGTGGAGGGTCCAAGTGTGTTAAGAGTACAGAGTCGGA

38461  CCTGAGTAGCTGGGTTTACAGGCACTCGCCACCATGCCCAGCTAATTTTTGTATTTTTAG
38461  GGACTCATCGACCCAAATGTCCGTGAGCGGTGGTACGGGTCGATTAAAAACATAAAAATC

38521  TAAAGACAGGGTTTTACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTGATC
38521  ATTTCTGTCCCAAAATGGTACAACCGGTCCGACCAGAGTTTGAGGACTGGAGTCCACTAG

38581  GGCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGTGCCACTGTGCCTGGCCCCA
38581  CCGGCGGAACCGGAGGGTTTCACGACCCTAATGTCCGCACACGGTGACACGGACCGGGGT

38641  GTTCTTATTATAAGTTCTCCATTAACCATATCATGAGATTAAAAAAACAACAATGATTTC
38641  CAAGAATAATATTCAAGAGGTAATTGGTATAGTACTCTAATTTTTTTGTTGTTACTAAAG

38701  ATCGGAACTGACAAATAGTGGGTGACACAATTTTGATCATCAAGAAGACTTTAAAAGAAG
38701  TAGCCTTGACTGTTTATCACCCACTGTGTTAAAACTAGTAGTTCTTCTGAAATTTTCTTC

38761  GATTTGTATCACTGTCATTACTGATTACTTGGCCCTTGTGTATACTGGGCCTTTAGCTAG
38761  CTAAACATAGTGACAGTAATGACTAATGAACCGGGAACACATATGACCCGGAAATCGATC

38821  GGCTGGTACAATAAAATAACCATAGTTAGAAAGATATTTTAAAGGATTACTTATTTCATG
38821  CCGACCATGTTATTTTATTGGTATCAATCTTTCTATAAAATTTCCTAATGAATAAAGTAC

38881  TTTATCTCACCCTTTTATCCCTATTTTTGGGCATGATTTATTATTATTAGGTTTTTAAAT
38881  AAATAGAGTGGGAAAATAGGGATAAAAACCCGTACTAAATAATAATAATCCAAAAATTTA

38941  AAATAATGAAATAAGCATGTTAGGTTTGCATGCTGCTTTGGTTAACATTGCTATATAACC
38941  TTTATTACTTTATTCGTACAATCCAAACGTACGACGAAACCAATTGTAACGATATATTGG

39001  AACCACCCCAAACTTTGTGGCCTTAAAACAATAATTTATTTATTTATTTATTTATTTATT
39001  TTGGTGGGGTTTGAAACACCGGAATTTTGTTATTAAATAAATAAATAAATAAATAAATAA

39061  TATTTATTTATTTATTTATTTTTGAGACAGAGTCTCGCTATGTCACCCAGGCTGGAGTGC
39061  ATAAATAAATAAATAAATAAAAACTCTGTCTCAGAGCGATACAGTGGGTCCGACCTCACG

39121  AGCGGCATGATCTCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGTGATCCTCCTGCC
39121  TCGCCGTACTAGAGCCGAGTGACGTTGGAGGTGGAGGACCCAAGTTCACTAGGAGGACGG

39181  TCAGCCTCCCGTAGAGCTGGGATTACAGGCACCCGCCATCAAGCCCAGCTAATTTTTCCT
39181  AGTCGGAGGGCATCTCGACCCTAATGTCCGTGGGCGGTAGTTCGGGTCGATTAAAAAGGA

39241  TTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGTTGGTCTCGAACTCCTGATC
39241  AATAAAAATCATCTCTGCCCCAAAGTGGTACAACCGGTCCAACCAGAGCTTGAGGACTAG

39301  TCAGGTGATATGCCTGCTTTGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACTGT
39301  AGTCCACTATACGGACGAAACCGGAGGGTTTCACGACCCTAATGTCCACACTCGGTGACA

39361  GCATGGCCAACAATAATTTATTTTTATCTCTCATGGTCTCTAGATTGCCTGGGCCCGGCT
39361  CGTACCGGTTGTTATTAAATAAAAATAGAGAGTACCAGAGATCTAACGGACCCGGGCCGA

39421  GAGCAGTCTGAGTCAGCCTTATGGTCTCTCATGTGTTTTTAATCAGTCATGAGGGGCCGG
39421  CTCGTCAGACTCAGTCGGAATACCAGAGAGTACACAAAAATTAGTCAGTACTCCCCGGCC

39481  CATCATCTGAAGGCTTGACTGGGCTGGATGTTTAAGACAGCCACTCACAGGGCTGACAGT
39481  GTAGTAGACTTCCGAACTGACCCGACCTACAAATTCTGTCGGTGAGTGTCCCGACTGTCA

39541  GATGCTGGATGTTGCACTGTTGACTAATGAGCCTGCACGTGGCTCCTCCATGTGATTTGG
39541  CTACGACCTACAACGTGACAACTGATTACTCGGACGTGCACCGAGGAGGTACACTAAACC
```

FIGURE 4 (cont'd)

```
39601  GCTATTTATAGCTGGGGTGACTGGGTTCTAGGTGGGGTAGTTCCAAGATCCTTCCAAGAG
39601  CGATAAATATCGACCCCACTGACCCAAGATCCACCCCATCAAGGTTCTAGGAAGGTTCTC

39661  GAAGAAAGCAGAAACTGCCTGGCCAGTTAGTGTATATGCTTAGTCTTAGAACAGTGTCAC
39661  CTTCTTTCGTCTTTGACGGACCGGTCAATCACATATACGAATCAGAATCTTGTCACAGTG

39721  TTCCATATTTCTCCATAGTCCATTAATCAAAGCAGTTACAGGAAGCAGACAGACAGGAAG
39721  AAGGTATAAAGAGGTATCAGGTAATTAGTTTCGTCAATGTCCTTCGTCTGTCTGTCCTTC

39781  GGAAAACACAATGCAACAGAAGGAGAGATTGAAGTGTTGGCAGCTGCAAGCCAAGAAACA
39781  CCTTTTGTGTTACGTTGTCTTCCTCTCTAACTTCACAACCGTCGACGTTCGGTTCTTTGT

39841  CCTGGGTTACCAGAAGCTAGAAGAAGAAAGGTAGGATTCTCCCCTACAGGTTCCAAAGGA
39841  GGACCCAATGGTCTTCGATCTTCTTCTTTCCATCCTAAGAGGGGATGTCCAAGGTTTCCT

39901  AGCTGGGCCTGTCCACACCTCACCTCTGGAACTATGAGACAATCACTTTCTGTGGGTTTA
39901  TCGACCCGGACAGGTGTGGAGTGGAGACCTTGATACTCTGTTAGTGAAAGACACCCAAAT

39961  AGCTACCCAGTTTGTGGCACTTTGTTACTACCGTCCTAGGCAAATGATACAGTAACTACA
39961  TCGATGGGTCAAACACCGTGAAACAATGATGGCAGGATCCGTTTACTATGTCATTGATGT

40021  TACACAGTACATAGAATTGCTCATTTTACTTCGTTGAATGGGACTTGGCATTGAATAAGC
40021  ATGTGTCATGTATCTTAACGAGTAAAATGAAGCAACTTACCCTGAACCGTAACTTATTCG

40081  CCAAATGAGAATTCTTTTGTCAAGATTTTTTTATGGCTCACTTTTCCAGAGACGTGACAT
40081  GGTTTACTCTTAAGAAAACAGTTCTAAAAAAATACCGAGTGAAAAGGTCTCTGCACTGTA

40141  GGCAGGAGTAAGGGGAGGGTTTGGCAGGTAGGAATTATACTCATCTGTCATTGTCTTCCC
40141  CCGTCCTCATTCCCCTCCCAAACCGTCCATCCTTAATATGAGTAGACAGTAACAGAAGGG

40201  AATACCAACCAGGAAGCATGTAAATAATAGCTGTGCAGTTTGTTTTCTGAACCAAACTGG
40201  TTATGGTTGGTCCTTCGTACATTTATTATCGACACGTCAAACAAAAGACTTGGTTTGACC

40261  GCCTTATTTTTTTTAAAGAAACTTTATTAAACAAACAAAAACCTAGAAACACTAGATGTA
40261  CGGAATAAAAAAAATTTCTTTGAAATAATTTGTTTGTTTTTGGATCTTTGTGATCTACAT

40321  TATTTTTAAGGTTTCTCTCACATAATATCCTAATCATCAGTATACCGTAATCCAAGGGCT
40321  ATAAAAATTCCAAAGAGAGTGTATTATAGGATTAGTAGTCATATGGCATTAGGTTCCCGA

40381  ACATATAACTGAAAAATAAATGGGTCTTAAAAAAATTAAAACATGGTATATAACACTAGC
40381  TGTATATTGACTTTTTATTTACCCAGAATTTTTTTAATTTTGTACCATATATTGTGATCG

40441  TAACATGTACTGAGCAGTTACTAAGTGCTAAGTACTGCTAAGTCCTTTATAGATATCTTC
40441  ATTGTACATGACTCGTCAATGATTCACGATTCATGACGATTCAGGAAATATCTATAGAAG

40501  ATGTTGAGTGTCGTCACTGCCCATTACTTACTCTGAGTTTAAATATGAGACTCAAAAGAG
40501  TACAACTCACAGCAGTGACGGGTAATGAATGAGACTCAAATTTATACTCTGAGTTTTCTC

40561  GTTAGTAATTTGCCAAAAGTCATAGGGCTTATAAGTGATAGAACCAGGATAACGAATCTC
40561  CAATCATTAAACGGTTTTCAGTATCCCGAATATTCACTATCTTGGTCCTATTGCTTAGAG

40621  AACCAAGGGACTTCAATGCCCAAACCGGTGCTCAGACTGTAGCCCCATGTAGTTAGCTAT
40621  TTGGTTCCCTGAAGTTACGGGTTTGGCCACGAGTCTGACATCGGGGTACATCAATCGATA

40681  TTACATTTGAAAAAATTTTTAATGTTTATTATTTTTTTTAAGACAGGGTCTTGCTCTGTC
40681  AATGTAAACTTTTTTAAAAATTACAAATAATAAAAAAAATTCTGTCCCAGAACGAGACAG

40741  ACCTAGGCTGGAGTGGAGTGGTGTGATCTTGGCTCACTGCAACCTCCACCTCCTGGGCTC
40741  TGGATCCGACCTCACCTCACCACACTAGAACCGAGTGACGTTGGAGGTGGAGGACCCGAG

40801  AAGTGATCCTCTCATCTCAGCCTCCCAAGTAGCTGGGACCACAGGTGTATGGCACCACGC
40801  TTCACTAGGAGAGTAGAGTCGGAGGGTTCATCGACCCTGGTGTCCACATACCGTGGTGCG

40861  CCAACTAATTTGTTGTGTTCTTGGTAGAGGTGAGTTTCACCATGTTGCCCAGGCTGGTCT
40861  GGTTGATTAAACAACACAAGAACCATCTCCACTCAAAGTGGTACAACGGGTCCGACCAGA
```

FIGURE 4 (cont'd)

```
40921  TGAACTCCTGAGCTCAAGTGACCTGCCCACTTGCCTCACCCTCCTAAAGTGGGGGGATTA
40921  ACTTGAGGACTCGAGTTCACTGGACGGGTGAACGGAGTGGGAGGATTTCACCCCCCTAAT

40981  CAGGTGTGAGCCACCATGCCCGGCCCAGCTATTTACATTTAAATATAAATTAAAATGAAA
40981  GTCCACACTCGGTGGTACGGGCCGGGTCGATAAATGTAAATTTATATTTAATTTTACTTT

41041  CAAGATTAAAAATTCGGTTTCTCAATCACACTAACCACACTTGAAGTGCTCAGTAGCTAC
41041  GTTCTAATTTTTAAGCCAAAGAGTTAGTGTGATTGGTGTGAACTTCACGAGTCATCGATG

41101  ATGTGGCTTAGTGGCTACCTGGCAGATATAGAACGTTTCCATCAGCACAGAAAGTTCTAC
41101  TACACCGAATCACCGATGGACCGTCTATATCTTGCAAAGGTAGTCGTGTCTTTCAAGATG

41161  TGGAAAGCATTGATCCAATCTCTTTGAACAGTCACAATGCATATAAAGTCCTTATTATGA
41161  ACCTTTCGTAACTAGGTTAGAGAAACTTGTCAGTGTTACGTATATTTCAGGAATAATACT

41221  CTGTACAAGGGAAGACAACGAGTAGAGGAAGGCAGAAAATAGGAAAAAAGGAAGAAGAAA
41221  GACATGTTCCCTTCTGTTGCTCATCTCCTTCCGTCTTTTATCCTTTTTTCCTTCTTCTTT

41281  GATTTAGGGTGAGGGAGAGCAGTTCTATGAAGAAAAGAAGCTCATATCTAGGAGTGGAGT
41281  CTAAATCCCACTCCCTCTCGTCAAGATACTTCTTTTCTTCGAGTATAGATCCTCACCTCA

41341  GGACCCTGAGAACATGTTAGTCTATTATATTCTAAACTTTGGCAGTTAAAATTCCAAAGT
41341  CCTGGGACTCTTGTACAATCAGATAATATAAGATTTGAAACCGTCAATTTTAAGGTTTCA

41401  ATATTAAGTACATGCTTTCTAAAATATTAGAGAGCTGCAGTTCAACAGATTTTTAACTCT
41401  TATAATTCATGTACGAAAGATTTTATAATCTCTCGACGTCAAGTTGTCTAAAAATTGAGA

41461  CCACAGCATGTCCGGCACTGGGTGAGGTGGGATGGCTAGGTGAATAATATAGGCATTGTC
41461  GGTGTCGTACAGGCCGTGACCCACTCCACCCTACCGATCCACTTATTATATCCGTAACAG

41521  ATGTGTGTTCGTGTGAAGAGACCACCAAACAGGCTTTGTGTGAGCAATAAAGCTTTTTAA
41521  TACACACAAGCACACTTCTCTGGTGGTTTGTCCGAAACACACTCGTTATTTCGAAAAATT

41581  TCACCTGGGTACAGGCGGGCCGAGTCTGAAAAGAGAGTCAGCAAAGGGTGGTGGGATTAT
41581  AGTGGACCCATGTCCGCCCGGCTCAGACTTTTCTCTCAGTCGTTTCCCACCACCCTAATA

41641  AATTAGTTCTTAGAGGTTTGGGATAGGCGGTGGAGTAGGAACAATTTTTTGTGGGCAGGG
41641  TTAATCAAGAATCTCCAAACCCTATCCGCCACCTCATCCTTGTTAAAAAACACCCGTCCC

41701  GATGGATCTTACAAAGTACATTCTCAAGGGCGAGGAGAATATTACAAAGTACCTTAAGGG
41701  CTACCTAGAATGTTTCATGTAAGAGTTCCCGCTCCTCTTATAATGTTTCATGGAATTCCC

41761  CAGGGAAGGATATTACAAAGTACCTTCTCAAGGGTGGGGAGGGTGTATCATACAAAGTGC
41761  GTCCCTTCCTATAATGTTTCATGGAAGAGTTCCCACCCCTCCCACATAGTATGTTTCACG

41821  ATTCACAAGGGTGGGGAAATATCACAAAGTACATTATCGCAAGGGCAGGGAGGGTGTATT
41821  TAAGTGTTCCCACCCCTTTATAGTGTTTCATGTAATAGCGTTCCCGTCCCTCCCACATAA

41881  GTCACAAAGTCAATTGATCGGTTAGGGTGGGGCAGGAACAAATCACAACAGTGGAATGTC
41881  CAGTGTTTCAGTTAACTAGCCAATCCCACCCCGTCCTTGTTTAGTGTTGTCACCTTACAG

41941  ATCAGTTAAGGCAGGAACCTGCTATTTTCACTTTTGTTGATCTTCAGTTGCTTCAGGCCA
41941  TAGTCAATTCCGTCCTTGGACGATAAAAGTGAAAACAACTAGAAGTCAACGAAGTCCGGT

42001  TCTGGATGTATACGTGCAGGTCACAGGGGATATGATGGATTAGCTTGGGCTCAGAGGCCT
42001  AGACCTACATATGCACGTCCAGTGTCCCCTATACTACCTAATCGAACCCGAGTCTCCGGA

42061  GACAGACATGAACTTCCCATCATAAAATTAGAAGAGCATTAATCAAATGATTAGAATAAA
42061  CTGTCTGTACTTGAAGGGTAGTATTTTAATCTTCTCGTAATTAGTTTACTAATCTTATTT

42121  ATATTACCATTATTTTGAAATGGCAAGCATTGTATGTAAATAAACAGCAGAGGAACCCAC
42121  TATAATGGTAATAAAACTTTACCGTTCGTAACATACATTTATTTGTCGTCTCCTTGGGTG

42181  TGTAGCCTAGACAGGGTTCAGAGACGTTCCTGCCAAGGACATGGCACAGGGACTGAACAA
42181  ACATCGGATCTGTCCCAAGTCTCTGCAAGGACGGTTCCTGTACCGTGTCCCTGACTTGTT

42241  GGAAGGATGACAGTGCCTGGCAAAGGCAGGGAATGTTACAGGTATTTGGGCACCTGTGAG
42241  CCTTCCTACTGTCACGGACCGTTTCCGTCCCTTACAATGTCCATAAACCCGTGGACACTC
```

FIGURE 4 (cont'd)

```
42301  AAGGCCTGATGTCCGAAAGGACTGAAATAAATTAAATGTGTGGGGGGAGAGAGAAAGAAA
42301  TTCCGGACTACAGGCTTTCCTGACTTTATTTAATTTACACACCCCCCTCTCTCTTTCTTT

42361  GAGCCTAGCAAAAGGTGAGCTGGAGAGTCAGAGGCAGAAGGCAGAATCTGGGATGGGCTG
42361  CTCGGATCGTTTTCCACTCGACCTCTCAGTCTCCGTCTTCCGTCTTAGACCCTACCCGAC

42421  GGAGCTTGCCAATCATTTTGTCAAATGGGTGATGTGAAGCCTTGAAGGTTTTTTGACCAA
42421  CCTCGAACGGTTAGTAAAACAGTTTACCCACTACACTTCGGAACTTCCAAAAAACTGGTT

42481  GAGGAGATGTCTTCAGATTTGCGTTTGAGACATTTTCACTGGTTGTCATGCTAAGAATGG
42481  CTCCTCTACAGAAGTCTAAACGCAAACTCTGTAAAAGTGACCAACAGTACGATTCTTACC

42541  CTTAAAGAGAAGCAAGGCTGGGGTCAGAGGAGCTGCTTGGAAGCTGTTTGGTACTAGGAT
42541  GAATTTCTCTTCGTTCCGACCCCAGTCTCCTCGACGAACCTTCGACAAACCATGATCCTA

42601  GCGGATAACATGGTTGGAGAAATGGCAACTGTCAGGGTATTTGGAGCAAGAATGGATAGA
42601  CGCCTATTGTACCAACCTCTTTACCGTTGACAGTCCCATAAACCTCGTTCTTACCTATCT

42661  CATTGGTGATCATCACTTGGATGCATTTGAGGGACAACTGCTATGTTTCTACTAGAGCAA
42661  GTAACCACTAGTAGTGAACCTACGTAAACTCCCTGTTGACGATACAAAGATGATCTCGTT

42721  CTACATAGAGCAGGAGTAAGCAACTTACAGTCCACAGACCACATTTGCTCTGTCCCGTTT
42721  GATGTATCTCGTCCTCATTCGTTGAATGTCAGGTGTCTGGTGTAAACGAGACAGGGCAAA

42781  TTCAAATACAGTTTAATGGAAACAGAGCCATGCCCGTTCTTTCACCTATTCTGCATGGCT
42781  AAGTTTATGTCAAATTACCTTTGTCTCGGTACGGGCAAGAAAGTGGATAAGACGTACCGA

42841  GACATCAGGCATGGTTCTAAGTAATATACATAAGTTTCATATATTAATGCTCATCTATCC
42841  CTGTAGTCCGTACCAAGATTCATTATATGTATTCAAAGTATATAATTACGAGTAGATAGG

42901  CAGTTGACAAACATGAAGTGTTGCTATTGTCTCATTTTACAGCGGAGAAAACTGACATAG
42901  GTCAACTGTTTGTACTTCACAACGATAACAGAGTAAAATGTCGCCTCTTTTGACTGTATC

42961  AGATTCTCTTTCGCACCCATAGCTTCAGTTACCACCTCTATGCCAATGATTCAATAATAC
42961  TCTAAGAGAAAGCGTGGGTATCGAAGTCAATGGTGGAGATACGGTTACTAAGTTATTATG

43021  CTCTAATCCAGACCTCTTCTCTGAGCTTCAGACAGGCAGATCTCTCACAGCCAGTCTTGG
43021  GAGATTAGGTCTGGAGAAGAGACTCGAAGTCTGTCCGTCTAGAGAGTGTCGGTCAGAACC

43081  AGACACTTCAGAAATCATAGGTCCCAAACCAGACTCCTCCTCATCCCCTCAAACACATTC
43081  TCTGTGAAGTCTTTAGTATCCAGGGTTTGGTCTGAGGAGGAGTAGGGGAGTTTGTGTAAG

43141  TTTTTCAACATTCCTTTTTTCTTAGTGAACATCACCAAATAGATAAGCAAGTCAGATCCA
43141  AAAAAGTTGTAAGGAAAAAAGAATCACTTGTAGTGGTTTATCTATTCGTTCAGTCTAGGT

43201  AGGAGGCCTCCTTGGCACTTCTCTCATCCTTGTCCCACCTTATATCATCTATTTTTCTTT
43201  TCCTCCGGAGGAACCGTGAAGAGAGTAGGAACAGGGTGGAATATAGTAGATAAAAAGAAA

43261  TTCTTTTTTTTTGAGATGGTGTCTCACTGTGTCACCCAGGCAAGTGCAGTGGTGTGATCC
43261  AAGAAAAAAAAACTCTACCACAGAGTGACACAGTGGGTCCGTTCACGTCACCACACTAGG

43321  TGGCCCACTGCAGCCTCAACCTCACCAAGTTCAGGTAATCCTCTCACCTTAATCCCGTGA
43321  ACCGGGTGACGTCGGAGTTGGAGTGGTTCAAGTCCATTAGGAGAGTGGAATTAGGGCACT

43381  GTAGCTGGGACTACAGGCACATGCCACCGAGCCCGGCTAATTTTTGTATTTTTACCAGAG
43381  CATCGACCCTGATGTCCGTGTACGGTGGCTCGGGCCGATTAAAAACATAAAAATGGTCTC

43441  ACAGGGTTTTGCCATGTTGCCCAGGCTGCCACCTTATCTCCATTCTTGCTGAAATCTGGC
43441  TGTCCCAAAACGGTACAACGGGTCCGACGGTGGAATAGAGGTAAGAACGACTTTAGACCG

43501  TGTTTCTCTCCACTTCTATTACCAACCACCAGCCAAATCTGACTTATTAAAAGGATCATC
43501  ACAAAGAGAGGTGAAGATAATGGTTGGTGGTCGGTTTAGACTGAATAATTTTCCTAGTAG

43561  TCGTATGTGGGTTCTGTAATAGCTTCTTCAAAATCCATTCCTCTCCACCCAGACCCTCTT
43561  AGCATACACCCAAGACATTATCGAAGAAGTTTTAGGTAAGGAGAGGTGGGTCTGGGAGAA
```

FIGURE 4 (cont'd)

```
43621  CCAATCTCCAAACTGCAGCCAATACCTTCAATTCTGTCACCGCTCCTGCATACTCCACCT
43621  GGTTAGAGGTTTGACGTCGGTTATGGAAGTTAAGACAGTGGCGAGGACGTATGAGGTGGA

43681  AATCCTTAACCTTTCAATGGTTTCTCTTGCCCCTAGGTTAAGCTAAATTGAAAACTGCTA
43681  TTAGGAATTGGAAAGTTACCAAAGAGAACGGGGATCCAATTCGATTTAACTTTTGACGAT

43741  AAGCTACAGCTTCTTAGTGTAACTAGTAAGTACAGTGACTACTATGTCTGGTTTACTTCA
43741  TTCGATGTCGAAGAATCACATTGATCATTCATGTCACTGATGATACAGACCAAATGAAGT

43801  GCTGTTGCTGACTAATAGCACTTCTTGAATGAAGAAATTATATTACCGTACTGTAGGGAA
43801  CGACAACGACTGATTATCGTGAAGAACTTACTTCTTTAATATAATGGCATGACATCCCTT

43861  AAATTTGTACTTCAATGGAAAACTGAATTATTTAGAGCAAACCTTTAAGGTTATCTGATT
43861  TTTAAACATGAAGTTACCTTTTGACTTAATAAATCTCGTTTGGAAATTCCAATAGACTAA

43921  CTGATTCCATGGGCCCTTTCATTGGGTCTTACATTGTTTCATTGTTTTTGAGTTATTTTA
43921  GACTAAGGTACCCGGGAAAGTAACCCAGAATGTAACAAAGTAACAAAAACTCAATAAAAT

43981  CAGATGACTGATAGAAATGCAAAATAATTATTGCATGCAGTCATGCAAGTAAAATCTGAA
43981  GTCTACTGACTATCTTTACGTTTTATTAATAACGTACGTCAGTACGTTCATTTTAGACTT

44041  GGTCAGGGTTCAATTAACTTCTCTCCCCTACTGTGGAAGAAGCCAATTTCATTTGCATAT
44041  CCAGTCCCAAGTTAATTGAAGAGAGGGGATGACACCTTCTTCGGTTAAAGTAAACGTATA

44101  TCATTTCAATATTTCTGACCTATAAACATGCCTGTTCTTTGAACCAGTTTTTATATCTGC
44101  AGTAAAGTTATAAAGACTGGATATTTGTACGGACAAGAAACTTGGTCAAAAATATAGACG

44161  CTGGTAACTTCATTAAATGATTTAAATAAACTCTATAATTTAGTTTATTTTATACCTGTA
44161  GACCATTGAAGTAATTTACTAAATTTATTTGAGATATTAAATCAAATAAAATATGGACAT

44221  TGACAGTCAGGCTTTTACCTTATTTTTGCTTTTTAAAATTATTCTTTAGCTCTCCCTAAG
44221  ACTGTCAGTCCGAAAATGGAATAAAAACGAAAAATTTTAATAAGAAATCGAGAGGGATTC

44281  GCACTACTTGGCTGGCCCTTTGGCTACTCACTAGCCCAAGCTCCATCCTCTTCTGTCTAC
44281  CGTGATGAACCGACCGGGAAACCGATGAGTGATCGGGTTCGAGGTAGGAGAAGACAGATG

44341  CGCAGCCCCAGGCTCTCTTTTGGCCCTTTGTTCCAGCCCTCCTCAGAGAAGTTCTTTTGT
44341  GCGTCGGGGTCCGAGAGAAAACCGGGAAACAAGGTCGGGAGGAGTCTCTTCAAGAAAACA

44401  CCTGGAACTCACTTTCATACCAGGTGCTCTTCATCCCTTCTTCCTTTAGTAAATTTCTTT
44401  GGACCTTGAGTGAAAGTATGGTCCACGAGAAGTAGGGAAGAAGGAAATCATTTAAAGAAA

44461  AATCTTTCAGATTTCTGCTCAAGGATCACTTGCTGAAAGCCTTTCCAGGCCTTTCTGAGT
44461  TTAGAAAGTCTAAAGACGAGTTCCTAGTGAACGACTTTCGGAAAGGTCCGGAAAGACTCA

44521  CAATACAATCCATTATTAGAGGCTCTTAGAGCACAGGCAGCACTTGTTATAATTTTACAT
44521  GTTATGTTAGGTAATAATCTCCGAGAATCTCGTGTCCGTCGTGAACAATATTAAAATGTA

44581  CTGTCTGTGGGATTCCTTAACACTTTCTAACACAATTAACACTTTTCTCTCTGGCTGGGG
44581  GACAGACACCCTAAGGAATTGTGAAAGATTGTGTTAATTGTGAAAAGAGAGACCGACCCC

44641  GCAATGCATGATAACAGGATGATGGTATGCTCTCAACCCTCCCCCAGCAACTAGCCCAGT
44641  CGTTACGTACTATTGTCCTACTACCATACGAGAGTTGGGAGGGGGTCGTTGATCGGGTCA

44701  GCTTCGAACATGGTGGATAGATACACCGTTTTTATTTTAGACAGATTTTTGAGTAGCAAG
44701  CGAAGCTTGTACCACCTATCTATGTGGCAAAAATAAAATCTGTCTAAAAACTCATCGTTC

44761  TACCTACTGCTCAGGAGACAGATGGACTGAAAAAATATATTTGGGAGTCATCAGCACATA
44761  ATGGATGACGAGTCCTCTGTCTACCTGACTTTTTTATATAAACCCTCAGTAGTCGTGTAT

44821  GGTGCAGCCATGAAAATGCACGAGAAGCAAAACCATTGGCAGGAACAAAATGAAGGCTAT
44821  CCACGTCGGTACTTTTACGTGCTCTTCGTTTTGGTAACCGTCCTTGTTTTACTTCCGATA

44881  TAGAGACCTTATTTAGGTGGAATAATTGGACAGAAGCCCTTAATGGGGAGGAAATAGAAG
44881  ATCTCTGGAATAAATCCACCTTATTAACCTGTCTTCGGGAATTACCCCTCCTTTATCTTC

44941  CAGTTAATACAGATGGAGGCCGGGCGCCGTGGCTCACGCCTGTAATCCTAGCACTTTGGG
44941  GTCAATTATGTCTACCTCCGGCCCGCGGCACCGAGTGCGGACATTAGGATCGTGAAACCC
```

FIGURE 4 (cont'd)

```
45001  AGGCTGAGATAGGTGGGTGGCTTGAGTCCAGGAGTTTGAGAGCAGCCTGGGCAACATAGC
45001  TCCGACTCTATCCACCCACCGAACTCAGGTCCTCAAACTCTCGTCGGACCCGTTGTATCG

45061  AAAACCCGGTCTCTACAAAAAATACAAAAATTAGCCCGGCGTGGTGGCGCAGGCCTGTAG
45061  TTTTGGGCCAGAGATGTTTTTTATGTTTTTAATCGGGCCGCACCACCGCGTCCGGACATC

45121  TCCCGGCTACTGGGGAGGCTGAGATGGGAGAATTGCTTGAACCGGGGAGGCGGATGTTGC
45121  AGGGCCGATGACCCCTCCGACTCTACCCTCTTAACGAACTTGGCCCCTCCGCCTACAACG

45181  AATTAGACGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTCA
45181  TTAATCTGCTCTAGCGCGGTGACGTGAGGTCGGACCCGCTGTCTCGCTCTGAGGCAGAGT

45241  AAACAACAAGGACAAAACTTATCTTATAAGCGACCCTGGACTTCCAGGGACACTCGCCGC
45241  TTTGTTGTTCCTGTTTTGAATAGAATATTCGCTGGGACCTGAAGGTCCCTGTGAGCGGCG

45301  TCTATGGTGGTACGTGACCTGCCTGGACTCTGTGGTCTTTGGAGCTTGGGACCCCACCCC
45301  AGATACCACCATGCACTGGACGGACCTGAGACACCAGAAACCTCGAACCCTGGGGTGGGG

45361  AGCGTAAGAGGCCTTTCTGTGCCTTGTATAAAGGAGAGGAAAAGTGTAGGTTTTCTGAGT
45361  TCGCATTCTCCGGAAAGACACGGAACATATTTCCTCTCCTTTTCACATCCAAAAGACTCA

45421  TGATAAGATTAGAGCTACAGTATATTATCAGTAGCGTTGTATGAAACCATTTAAAAAACC
45421  ACTATTCTAATCTCGATGTCATATAATAGTCATCGCAACATACTTTGGTAAATTTTTTGG

45481  AAACCGGATTTAACCGCTTTCACCCAGAATAAACTGCCCTCCCCCTCTTTCCAAGATCAA
45481  TTTGGCCTAAATTGGCGAAAGTGGGTCTTATTTGACGGGAGGGGGAGAAAGGTTCTAGTT

45541  GTCAGCCCCAGTTACCACTTCCTATTAATTTAAAAACCAGGAAGGCTCATAATTTAGGCT
45541  CAGTCGGGGTCAATGGTGAAGGATAATTAAATTTTTGGTCCTTCCGAGTATTAAATCCGA

45601  GACGTTGCCCCAGGCGAACACTGCATATAAGACAATAGCTATTCCAATCCTCCTTTCCCT
45601  CTGCAACGGGGTCCGCTTGTGACGTATATTCTGTTATCGATAAGGTTAGGAGGAAAGGGA

45661  TTCTAAACCGAATAGTACTTCTAACATCACTTTTCCAGTCCTCTTTCTAAATATTGCGAT
45661  AAGATTTGGCTTATCATGAAGATTGTAGTGAAAAGGTCAGGAGAAAGATTTATAACGCTA

45721  CACGAGCCTACATCGGAAAACTATGAAAACAGTATGATTGGAGCTCTTGATGGTCTGCTT
45721  GTGCTCGGATGTAGCCTTTTGATACTTTTGTCATACTAACCTCGAGAACTACCAGACGAA

45781  TGCCTGAAAACGATGAAAGGCATTGAGATCTTTGGGCAGAAATGGCCCTTAGGGAATGCG
45781  ACGGACTTTTGCTACTTTCCGTAACTCTAGAAACCCGTCTTTACCGGGAATCCCTTACGC

45841  AGCCTGGTAGCCCCATAACACCCCGAGTTACAAATTCGGTTTGGGTAGGTGGCGCAGGGA
45841  TCGGACCATCGGGGTATTGTGGGGCTCAATGTTTAAGCCAAACCCATCCACCGCGTCCCT

45901  AATCTAAAAGCTGGATTCACGTTCTCCCCCTTCAGTTTGCACATAGCGGTCAGTGGCAGA
45901  TTAGATTTTCGACCTAAGTGCAAGAGGGGGAAGTCAAACGTGTATCGCCAGTCACCGTCT

45961  CAAGAGCAGACGCACAAATGTCAATCCCCGCCCGCAAACTCTCGAGGGCAGAAACTGCTT
45961  GTTCTCGTCTGCGTGTTTACAGTTAGGGGCGGGCGTTTGAGAGCTCCCGTCTTTGACGAA

46021  TTTGTCAGTTGGATTTGGAGTCGGTGGAAAAGCTGCCCTAATGTAGTTTCTCCGAGGTTT
46021  AAACAGTCAACCTAAACCTCAGCCACCTTTTCGACGGGATTACATCAAAGAGGCTCCAAA

46081  CCTTAGCCTTGACCGAGGGGCGCTTCCCGGCCATTCACCTAGAGGTTGTTTAATAAACAA
46081  GGAATCGGAACTGGCTCCCCGCGAAGGGCCGGTAAGTGGATCTCCAACAAATTATTTGTT

46141  GGATGCTCGCGAAATATCTGCGCTTGGAAAGGCGCTCGTTCGTGGCGCGCATTCTCGGGC
46141  CCTACGAGCGCTTTATAGACGCGAACCTTTCCGCGAGCAAGCACCGCGCGTAAGAGCCCG

46201  CTCCGCAAGCGACCCCGGTGACAGGGACAACCGCTTCGGTTTTAGCGACTGCAGACAGAC
46201  GAGGCGTTCGCTGGGGCCACTGTCCCTGTTGGCGAAGCCAAAATCGCTGACGTCTGTCTG

46261  TGGGACGAGACGGTTGGAGGCTCCTCCCCAAGGGATGCTGGAGGGGTTGCGTCGTACCCT
46261  ACCCTGCTCTGCCAACCTCCGAGGAGGGGTTCCCTACGACCTCCCCAACGCAGCATGGGA
```

FIGURE 4 (cont'd)

```
46321  GCGCCTGGCCCTGGCGCGCGGCCCCAGGTCGTGGTACCCAGCGCCCTATGGGCCGTGCGC
46321  CGCGGACCGGGACCGCGCGCCGGGGTCCAGCACCATGGGTCGCGGGATACCCGGCACGCG

46381  CGGGGCTTGGCCACACCGCCTGCTTTCGCTTCCAGCCGCGCGCTCCGTGCCACTGCCGCT
46381  GCCCCGAACCGGTGTGGCGGACGAAAGCGAAGGTCGGCGCGCGAGGCACGGTGACGGCGA

46441  CTCTGCAGCCCCGCGTCCCCGCAGCCTCCCCATGGCCAGCCCGCTTCGCTCCGCTGCGGC
46441  GAGACGTCGGGGCGCAGGGGCGTCGGAGGGGTACCGGTCGGGCGAAGCGAGGCGACGCCG

46501  CCTTGCCCGCCAGGTACCTCGAACCCGGGCGTTTGCGGAAGGGGGGAGGATTGGAACCCG
46501  GGAACGGGCGGTCCATGGAGCTTGGGCCCGCAAACGCCTTCCCCCCTCCTAACCTTGGGC

46561  GGTCTCGGTAGCTCGCGGGCCTGGCCGGGCGCCTTGTCGCCGTTTCCTGCACCATCCTCC
46561  CCAGAGCCATCGAGCGCCCGGACCGGCCCGCGGAACAGCGGCAAAGGACGTGGTAGGAGG

46621  TTCGCCTTGCCCTCCATTCCGCCTCCAGCGAGGCGTCTTCCCTTCCCCGCATCCCTGCCC
46621  AAGCGGAACGGGAGGTAAGGCGGAGGTCGCTCCGCAGAAGGGAAGGGGCGTAGGGACGGG

46681  GAAATCTGGAGTCCCAGCCTGCAATCTCCACCTCTTCGAGGTTCCCGCTGCCCAGGTCTA
46681  CTTTAGACCTCAGGGTCGGACGTTAGAGGTGGAGAAGCTCCAAGGGCGACGGGTCCAGAT

46741  GCACCCTCATGGGTAACCCGCTCCGGAGCGTGGCGAGGACCGCCACGGGGGACGTGAGGG
46741  CGTGGGAGTACCCATTGGGCGAGGCCTCGCACCGCTCCTGGCGGTGCCCCCTGCACTCCC

46801  TAGCTATGGACTCGCTCTGAGGGAGGAGGCGGGAGCTGAATCTCTGGGCTGCCAGAACCC
46801  ATCGATACCTGAGCGAGACTCCCTCCTCCGCCCTCGACTTAGAGACCCGACGGTCTTGGG

46861  ACAGCCACATCCTACGTGACTCTGCCACCCCAAAATATTTTGACCGCAGCCTTCTGCCTC
46861  TGTCGGTGTAGGATGCACTGAGACGGTGGGGTTTTATAAAACTGGCGTCGGAAGACGGAG

46921  CTTGGATCTCTTCCTTCCCCACCCCCACCCCCGTAGTTATTTAGCAGATTACGCATTAAA
46921  GAACCTAGAGAAGGAAGGGGTGGGGGTGGGGGCATCAATAAATCGTCTAATGCGTAATTT

46981  ACAAATGTCTGCAGGTTTTCCCAATTAGTCCCGCTTCCCTGTGTCTTTATCTTTTAAATT
46981  TGTTTACAGACGTCCAAAAGGGTTAATCAGGGCGAAGGGACACAGAAATAGAAAATTTAA

47041  GCCCACTAATACCATGAGGTTTAAGGTGTGGGGTGGATGCTGCGGCATCGGAGGACCCTG
47041  CGGGTGATTATGGTACTCCAAATTCCACACCCCACCTACGACGCCGTAGCCTCCTGGGAC

47101  CTGGTGGAGGAAATGGTTCACGCCCGTCCCCGTTCCCTTTGCAGGCTTGCTATTGTGCGT
47101  GACCACCTCCTTTACCAAGTGCGGGCAGGGGCAAGGGAAACGTCCGAACGATAACACGCA

47161  CTGTGATTGACAAGACCACGAGGCTGAGCGCGCCCTGGAGATTTTTCTATAAATGGCTTA
47161  GACACTAACTGTTCTGGTGCTCCGACTCGCGCGGGACCTCTAAAAAGATATTTACCGAAT

47221  ACACCCCAGTCTAGACTATTTGCTCGGATATAAGGGAGACAATTGTTTTTTTGTTCTTTG
47221  TGTGGGGTCAGATCTGATAAACGAGCCTATATTCCCTCTGTTAACAAAAAAACAAGAAAC

47281  CCGGCGAACCCTGGCTCTGTAGGGCTGACCTGGAATTTAACCAGTCTTCCCTGAGCCGGC
47281  GGCCGCTTGGGACCGAGACATCCCGACTGGACCTTAAATTGGTCAGAAGGGACTCGGCCG

47341  GGAGGAGGACAAAAACCGCCGCGACCCCGGCAGGGTGGGAAGTGCAGGGCAGCGCTCCCA
47341  CCTCCTCCTGTTTTTGGCGGCGCTGGGGCCGTCCCACCCTTCACGTCCCGTCGCGAGGGT

47401  AGACACGCTTGTTGGAGGTTCGGGCCTGGGTGCTTGGTTGTCTGAGCCTCCTTTTTTGTG
47401  TCTGTGCGAACAACCTCCAAGCCCGGACCCACGAACCAACAGACTCGGAGGAAAAAACAC

47461  TTTGCCTGGGTCCTGGAGAGGAGCGCACGGTATCATGGTGAGCGTCACGTAGGTTACCCC
47461  AAACGGACCCAGGACCTCTCCTCGCGTGCCATAGTACCACTCGCAGTGCATCCAATGGGG

47521  GGGTCCCGCTTACCCACCTGCATTTACTTAATGGTGGTTTAATTCTTCTTTAAGGATTGC
47521  CCCAGGGCGAATGGGTGGACGTAAATGAATTACCACCAAATTAAGAAGAAATTCCTAACG

47581  AGTAACGGATGCTCCGCAGAGTGTACCGGAGAAGGAGGATCAAAAGAGGTGGTGGGGACT
47581  TCATTGCCTACGAGGCGTCTCACATGGCCTCTTCCTCCTAGTTTTCTCCACCACCCCTGA

47641  TTTAAGGTAAGTTGCTTGCCAGAGGCTTAGTGTCTAAAGCAACGTTCAACTGAATAAATC
47641  AAATTCCATTCAACGAACGGTCTCCGAATCACAGATTTCGTTGCAAGTTGACTTATTTAG
```

FIGURE 4 (cont'd)

```
47701  CATTTCTTGTTGAGATTCGGGGAAATTTTCACTGTGTGGGTTACCCAGTTCCTAGGTTGT
47701  GTAAAGAACAACTCTAAGCCCCTTTAAAAGTGACACACCCAATGGGTCAAGGATCCAACA

47761  GTTGCCAGACTTTTAAAATTCCACTCTTAGAGGAAAAAAAGCCATTTAGAGAGTGACAGT
47761  CAACGGTCTGAAAATTTTAAGGTGAGAATCTCCTTTTTTTCGGTAAATCTCTCACTGTCA

47821  ACATTCATCACAACAAAAATTATCTAGCAGCTAATTGCTGATTTTATTTTGAGTCATATA
47821  TGTAAGTAGTGTTGTTTTTAATAGATCGTCGATTAACGACTAAAATAAAACTCAGTATAT

47881  TTTTTCTTGAAAAATAGTCAATGCGTAAATATTAAACTAGTTGTTTATCAGATTCCATAA
47881  AAAAAGAACTTTTTATCAGTTACGCATTTATAATTTGATCAACAAATAGTCTAAGGTATT

47941  TACGTACAAACCTTCAGTCTACTTGTAGAAAAAGTTAAACTGTGATTTTTCTCTTTCAGT
47941  ATGCATGTTTGGAAGTCAGATGAACATCTTTTTCAATTTGACACTAAAAAGAGAAAGTCA

48001  AGCACATAACACTCTACAGTGAAAGTGTTGCATTTTCTTCTTTTCTATCTGAATTGCGGG
48001  TCGTGTATTGTGAGATGTCACTTTCACAACGTAAAAGAAGAAAAGATAGACTTAACGCCC

48061  GTTTATATAGTAAATAGCTCATACTTCAAACAGTGAGAAACAAAGCACGTTGAAAACTGT
48061  CAAATATATCATTTATCGAGTATGAAGTTTGTCACTCTTTGTTTCGTGCAACTTTTGACA

48121  ACAAATACAAATGTAAGATGTAATATTATTTTTTCTATAAGTCTATCTCCTACCCTCCCT
48121  TGTTTATGTTTACATTCTACATTATAATAAAAAAGATATTCAGATAGAGGATGGGAGGGA

48181  CCTCGTGGAGGTTTTAGCTTCTTCTTTTTATGCCTAGTGCTGCTCTTCTGCCTGTTTGAG
48181  GGAGCACCTCCAAAATCGAAGAAGAAAAATACGGATCACGACGAGAAGACGGACAAACTC

48241  TCACTGTCAATTGTCATCAGACAATTGACACAGAGTGTCCTAATGATCTCACATCCTGGC
48241  AGTGACAGTTAACAGTAGTCTGTTAACTGTGTCTCACAGGATTACTAGAGTGTAGGACCG

48301  GGTTCAGAGAGGGAAGGAAACTTGACTCATCAGTTTCTCACACATTGAGCTCTCTCACAG
48301  CCAAGTCTCTCCCTTCCTTTGAACTGAGTAGTCAAAGAGTGTGTAACTCGAGAGAGTGTC

48361  TGTGCTGGAGACTGTGCCATGCATTGGACATCCAAAGGTGAATTCAAGATTAGGTCTCTA
48361  ACACGACCTCTGACACGGTACGTAACCTGTAGGTTTCCACTTAAGTTCTAATCCAGAGAT

48421  CTTGCAAGGAATTTTTATTTTATTTTTCTTTTTCCTTCTCTTGCTGTCATCTTAGCAAGG
48421  GAACGTTCCTTAAAAATAAAATAAAAAGAAAAAGGAAGAGAACGACAGTAGAATCGTTCC

48481  ATTTTTTAAAAAGCAAAATTAGAATATTGCAATTTCTTTTCTGTTTTTTGAGACAGGGTC
48481  TAAAAAATTTTTCGTTTTAATCTTATAACGTTAAAGAAAAGACAAAAAACTCTGTCCCAG

48541  TCGCTCTGTCACCCAGGCTGGAGTGCAGTGGCACGATCTCAGCTCACTGCAATCTCTGCC
48541  AGCGAGACAGTGGGTCCGACCTCACGTCACCGTGCTAGAGTCGAGTGACGTTAGAGACGG

48601  TCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGAGATTACAGGCATGT
48601  AGGGTCCAAGTTCACTAAGAGGACGGAGTCGGAGGACTCATCGACTCTAATGTCCGTACA

48661  GCCATCACGCCCAGCTAATTTTGTATTTTTAGTAGAGATGGGGTTATCACCATGTTAGCC
48661  CGGTAGTGCGGGTCGATTAAAACATAAAAATCATCTCTACCCCAATAGTGGTACAATCGG

48721  AGGCTGGTCTCGAACTCCTGGCCTCATGCAATCCACCTCCCAAAGTGCTGGGATTGCAGG
48721  TCCGACCAGAGCTTGAGGACCGGAGTACGTTAGGTGGAGGGTTTCACGACCCTAACGTCC

48781  CGTGAGCCACCAAGCCTAGCCTGAAGAATTCTTTGCAAAGAAAAATATCAACTCAAATTT
48781  GCACTCGGTGGTTCGGATCGGACTTCTTAAGAAACGTTTCTTTTTTATAGTTGAGTTTAAA

48841  AAATAATCTTTTCCTTTCTTTTTCTTTTTTTTGAGACAGAGTCTTGTTCTGTTGCCCAGA
48841  TTTATTAGAAAAGGAAAGAAAAAGAAAAAAAACTCTGTCTCAGAACAAGACAACGGGTCT

48901  CTGAAGTGCAGTGGCACCATCAATTAGGTATCACTTTCAGCCTTGATCTCTCAGACTCAA
48901  GACTTCACGTCACCGTGGTAGTTAATCCATAGTGAAAGTCGGAACTAGAGAGTCTGAGTT

48961  GTGATCCTCTTAGCCTCTCAAGTACCTGAGACTATCGGCTTGCATGCCATCAGGCCCAAC
48961  CACTAGGAGAATCGGAGAGTTCATGGACTCTGATAGCCGAACGTACGGTAGTCCGGGTTG
```

FIGURE 4 (cont'd)

```
49021  TAATTTCGTCATTTTTTGTAGAGATGGGATCTCCCTGTGTTGCCCAGACTGATCTCCAAC
49021  ATTAAAGCAGTAAAAAACATCTCTACCCTAGAGGGACACAACGGGTCTGACTAGAGGTTG

49081  TCCTGGGCTTGTGCAACCCTCCCACCTCGGCCTCCCAAAATGCTGGGTCACAGGCATAAG
49081  AGGACCCGAACACGTTGGGAGGGTGGAGCCGGAGGGTTTTACGACCCAGTGTCCGTATTC

49141  CCACAATGCCTGGCCATAAATAATCTTTTTTATTTGTTAATTTAGATTTCTGTATATTAG
49141  GGTGTTACGGACCGGTATTTATTAGAAAAAATAAACAATTAAATCTAAAGACATATAATC

49201  GTTTTATTTAGGGGAATTGCCTGTTTACACTCATGTATTTACACATCAAAACATGAGTTA
49201  CAAAATAAATCCCCTTAACGGACAAATGTGAGTACATAAATGTGTAGTTTTGTACTCAAT

49261  TGGTGGATGTATTCACAGCACATATGTAATAAGGCATGTGTAAAAATGATATTGATAATT
49261  ACCACCTACATAAGTGTCGTGTATACATTATTCCGTACACATTTTTACTATAACTATTAA

49321  TAGGGAATAAAAATGGTAGATTTTATAGACCATCTTTGTTTTTGATATACTTTTCTCTGT
49321  ATCCCTTATTTTTACCATCTAAAATATCTGGTAGAAACAAAAACTATATGAAAAGAGACA

49381  CTCAAAAAACAAAAACAAAAACAAACAAACAAACAAAATATATATATATAGCCTACATTG
49381  GAGTTTTTTGTTTTTGTTTTTGTTTGTTTGTTTTATATATATATATCGGATGTAAC

49441  TTGTTAACTATAATCACCCTACTCTGCTATGAACAGTAGACTTTATCGCTTCTATCTAAC
49441  AACAATTGATATTAGTGGGATGAGACGATACTTGTCATCTGAAATAGCGAAGATAGATTG

49501  TGTATGTTTGTACTCATTAACCAACCTCTCTTCCTCCACCTACACCCTTCCCACCCCCTG
49501  ACATACAAACATGAGTAATTGGTTGGAGAGAAGGAGGTGGATGTGGGAAGGGTGGGGGAC

49561  CCATCTTCTGATATCTATCATTCTATGAGTACTCTCTTCCTCCATGAGATCAACTTTTTT
49561  GGTAGAAGACTATAGATAGTAAGATACTCATGAGAGAAGGAGGTACTCTAGTTGAAAAAA

49621  TTTTAGCTCTCACGTATGAGTGAGAACATGCAATATTTGTTTTTCTGGGTTTTTTTTTTT
49621  AAAATCGAGAGTGCATACTCACTCTTGTACGTTATAAACAAAAAGACCCAAAAAAAAAAA

49681  TTCTGTGTTTGGCTTATTTCGCTTACCATAATGACTTCCAGTTCCATCCATGTTGCTACA
49681  AAGACACAAACCGAATAAAGCGAATGGTATTACTGAAGGTCAAGGTAGGTACAACGATGT

49741  AATGACACGATTTTACTCTTTTTTAATGGCTGAGTAGTATTCCATTGTGTATATGCACCA
49741  TTACTGTGCTAAAATGAGAAAAAATTACCGACTCATCATAAGGTAACACATATACGTGGT

49801  CATTTTATTTGTTCATTGATGGACACTTAGGTGATTCCATGTCTTGACTATTGTGAATAG
49801  GTAAAATAAACAAGTAACTACCTGTGAATCCACTAAGGTACAGAACTGATAACACTTATC

49861  TGCTGTAATAAACATGGGACTGCAGATATCCCTTTGATATACTGATTTCCTTTCCTTTGG
49861  ACGACATTATTTGTACCCTGACGTCTATAGGGAAACTATATGACTAAAGGAAAGGAAACC

49921  AAAATACCCAGTAGTAGGATTGTTGGTTTATATGGTCATTCTACTTTTAGTTTTTTTTTT
49921  TTTTATGGGTCATCATCCTAACAACCAAATATACCAGTAAGATGAAAATCAAAAAAAAAA

49981  GGAAAGCTTCATACCATTTTCCACAGTAGATAAACTAATTTACATTCTTACCAACAGTGT
49981  CCTTTCGAAGTATGGTAAAAGGTGTCATCTATTTGATTAAATGTAAGAATGGTTGTCACA

50041  ATAAGCGTTCTCTTTTCTCTGTATCCTCATCAGCATCTGTTATTTTTTGTCTTTCTAATA
50041  TATTCGCAAGAGAAAAGAGACATAGGAGTAGTCGTAGACAATAAAAAACAGAAAGATTAT

50101  ATAACCATTCTAACTGTGGTTGCTTTTGACATATTTCTCTACCTTTTTGTTAATCTACAT
50101  TATTGGTAAGATTGACACCAACGAAAACTGTATAAAGAGATGGAAAAACAATTAGATGTA

50161  CATTTATAACATTTCATATGGCCTTTTGGGTCTCTCTTGGGTAGTTTTCGTATTTCTTCT
50161  GTAAATATTGTAAAGTATACCGGAAAACCCAGAGAGAACCCATCAAAAGCATAAAGAAGA

50221  CCCCTCCCCATTTATGATGCAGTCATATTACATGAGGATGCAGTTTTGCAGCAGGGCCTA
50221  GGGGAGGGGTAAATACTACGTCAGTATAATGTACTCCTACGTCAAAACGTCGTCCCGGAT

50281  GCAATGAGATACGTTGTTACACATCCTGGCGAATCTTTCTGTGGACTCTAGGATCCTCTT
50281  CGTTACTCTATGCAACAATGTGTAGGACCGCTTAGAAAGACACCTGAGATCCTAGGAGAA

50341  CATTTGCAAGAGGTCAGCCATCTATTTCTTCCAATTCCTCCAGCTTCACATATCTCCACA
50341  GTAAACGTTCTCCAGTCGGTAGATAAAGAAGGTTAAGGAGGTCGAAGTGTATAGAGGTGT
```

FIGURE 4 (cont'd)

```
50401  GAATGCAACACTGCAAATCTGCAGTTCTCTTCATGATCCTCTTTATTAGATTTTAAAAGA
50401  CTTACGTTGTGACGTTTAGACGTCAAGAGAAGTACTAGGAGAAATAATCTAAAATTTTCT

50461  GTATCTCCAAAAGATGAACATAGTGCCTAAGGCTTTCCAAATGGAGTAAGAATAATGCAC
50461  CATAGAGGTTTTCTACTTGTATCACGGATTCCGAAAGGTTTACCTCATTCTTATTACGTG

50521  AAGAACCTAGAATAAAGTTATTAAATGTCTGATACAGTTGTAATACAGACTTGATTTTTG
50521  TTCTTGGATCTTATTTCAATAATTTACAGACTATGTCAACATTATGTCTGAACTAAAAAC

50581  AAAAATGTGGCCAATGAAGAGAAGAGCAAATGTGCTATACGATGGTGAGCAAAAGGTGTT
50581  TTTTTACACCGGTTACTTCTCTTCTCGTTTACACGATATGCTACCACTCGTTTTCCACAA

50641  GACAGGCCTGGCATGGTGACTCACCCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGG
50641  CTGTCCGGACCGTACCACTGAGTGGGGACATTAGGGTCGTGAAACCCTCCGACTCCACCC

50701  TGGATCATTTGAGGTCAGGAGTTTGAGACTAGCCTGGCCAACATGGTGAAACCCTGTTTG
50701  ACCTAGTAAACTCCAGTCCTCAAACTCTGATCGGACCGGTTGTACCACTTTGGGACAAAC

50761  TATTAAAAATACAAAAATTAGCCAGGTGCGATGACACATGCCTGTAATCCCAGCTACCTG
50761  ATAATTTTTATGTTTTTAATCGGTCCACGCTACTGTGTACGGACATTAGGGTCGATGGAC

50821  GGAGGCTGAGGCTTGAGACTTGCTTGAACCGGGGAGGCGGAGGTTGCAGCGAGCCGAGAT
50821  CCTCCGACTCCGAACTCTGAACGAACTTGGCCCCTCCGCCTCCAACGTCGCTCGGCTCTA

50881  TGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAG
50881  ACACGGTGACGTGAGGTCGGACCCGTTGTCTCGTTTTTTTTTTTTTTTTTTTTTTTTC

50941  TGTTGACTGAACCCTGGCTGAAACAGTTTTCAGGTGCTTTCAACAAATTTACACTTCCCT
50941  ACAACTGACTTGGGACCGACTTTGTCAAAAGTCCACGAAAGTTGTTTAAATGTGAAGGGA

51001  GCCCTGAGCCTCCCTATTTCAGAACAAATGATTACCCAACTGCAAACTTTGAATTAGGAA
51001  CGGGACTCGGAGGGATAAAGTCTTGTTTACTAATGGGTTGACGTTTGAAACTTAATCCTT

51061  GTTGTTTTTAATATGCCCTGTTTTTCTCAATATTTCTTGCCATTTCACAATCAGTTATTG
51061  CAACAAAAATTATACGGGACAAAAAGAGTTATAAAGAACGGTAAAGTGTTAGTCAATAAC

51121  TGGCATTCCAATTCTTCCACAAACATATGATGTGAAACATCCTGACTTCGATGTCATCAG
51121  ACCGTAAGGTTAAGAAGGTGTTTGTATACTACACTTTGTAGGACTGAAGCTACAGTAGTC

51181  GGAATGTGGAGCATTGGAAAGAGACAGATTAATTATGCAGAAGAAAGTATCATTTGACAA
51181  CCTTACACCTCGTAACCTTTCTCTGTCTAATTAATACGTCTTCTTTCATAGTAAACTGTT

51241  AATATTATTTTTATTTATCTTTGGAAAGGAAATCTAGAAAAAGAAAATACTTTTTAAAGA
51241  TTATAATAAAAATAAATAGAAACCTTTCCTTTAGATCTTTTTCTTTTATGAAAAATTTCT

51301  ACACTGATAAAAATTGTTTATAAAATTTGTTTCCTTTAGGAAACAAATCGTTCCCTTAAA
51301  TGTGACTATTTTTAACAAATATTTTAAACAAAGGAAATCCTTTGTTTAGCAAGGGAATTT

51361  ACTACATTTCATATCTGCCCCAAACGAGTCACATGTCTTCAGCTAATATTGCTTGTTTAC
51361  TGATGTAAAGTATAGACGGGGTTTGCTCAGTGTACAGAAGTCGATTATAACGAACAAATG

51421  GTTTGAACTTTATTCAATTCTATCACCTAGTAATTTAGTATTCTTAGGTAATTTCTCTTT
51421  CAAACTTGAAATAAGTTAAGATAGTGGATCATTAAATCATAAGAATCCATTAAAGAGAAA

51481  GAGCTCAGGTTAAAAGTTGTTATTTTTAGTTTTCAGAGTTGCAGAATTTTTTTTTCAAGT
51481  CTCGAGTCCAATTTTCAACAATAAAAATCAAAAGTCTCAACGTCTTAAAAAAAAAGTTCA

51541  AGAGAAAAGTTCATGATGAGATTCTGAAGCCACCTAATCGTATATTACTTTGAAAAACTG
51541  TCTCTTTTCAAGTACTACTCTAAGACTTCGGTGGATTAGCATATAATGAAACTTTTTGAC

51601  GATATATGCCGGGCGCGGTAGCTCACACCTGTAATCCCAGCACTTTCAGAGGCCGAGGTG
51601  CTATATACGGCCCGCGCCATCGAGTGTGGACATTAGGGTCGTGAAAGTCTCCGGCTCCAC

51661  GGTGGGTCACCTGAGGTCAGGAGTTCGAGACCAGCCTGGTCAACATGGTGAAACCCCATC
51661  CCACCCAGTGGACTCCAGTCCTCAAGCTCTGGTCGGACCAGTTGTACCACTTTGGGGTAG
```

FIGURE 4 (cont'd)

```
51721  TCTACCAAAAATACAAAACTTAGCCGGGTGTTGTGGCACATGCCTGTAGTCCCAGCTACT
51721  AGATGGTTTTTATGTTTTGAATCGGCCCACAACACCGTGTACGGACATCAGGGTCGATGA

51781  CAGGAGGCTGAAGCAGGAGAATCACTTGAACCTGGGAGGCGGAGGTTGCAGTGAGCCCAG
51781  GTCCTCCGACTTCGTCCTCTTAGTGAACTTGGACCCTCCGCCTCCAACGTCACTCGGGTC

51841  ATCACATCACTGCTCTTCCGCCTGGGTGACAGAGTGAGACTCCATCTAAAAGAAAAAGCA
51841  TAGTGTAGTGACGAGAAGGCGGACCCACTGTCTCACTCTGAGGTAGATTTTCTTTTTCGT

51901  AAACAAAACAAAACTGGATACAGGGCAGTGCTTAGGTACAAATCAGAAAGTTAGAGGTTG
51901  TTTGTTTTGTTTTGACCTATGTCCCGTCACGAATCCATGTTTAGTCTTTCAATCTCCAAC

51961  CTATTCATATCATTAAGAAATAATGTGTCTTACGAAAGAATGTAGCAGTTTACTGAAATA
51961  GATAAGTATAGTAATTCTTTATTACACAGAATGCTTTCTTACATCGTCAAATGACTTTAT

52021  GAAGTCATACCTAGAAATGTAATAATTTATCTTTAATTGTCAGAAATTTAACTGATATTA
52021  CTTCAGTATGGATCTTTACATTATTAAATAGAAATTAACAGTCTTTAAATTGACTATAAT

52081  TGAAGTTCTAATTATAAGAACACTTATTTGAGGTTAAAAATTACCTTATTCTATCCCTGC
52081  ACTTCAAGATTAATATTCTTGTGAATAAACTCCAATTTTTAATGGAATAAGATAGGGACG

52141  TTTGCCGCTAATTTGATCTCTCTGGATTAGTGGGTTATTGTGGGTCCTATGGGTATTTTA
52141  AAACGGCGATTAAACTAGAGAGACCTAATCACCCAATAACACCCAGGATACCCATAAAAT

52201  AGGAGAGGTCTCAGAAAACTAAAAATTATATATACTGGATCATAATCATTCTTGTCACCC
52201  TCCTCTCCAGAGTCTTTTGATTTTTAATATATATGACCTAGTATTAGTAAGAACAGTGGG

52261  ACTCTCATCTTTCTGTTTTTGTTTTGTGTTTTACTACCCTTCCCCTTTTACTAAGTAGTC
52261  TGAGAGTAGAAAGACAAAAACAAAACACAAAATGATGGGAAGGGGAAAATGATTCATCAG

52321  TTAGTAGAAGAAAATTTGCATTAAGTTTGTCGTACTGTATTGATATGGCTCCAAATGTTT
52321  AATCATCTTCTTTTAAACGTAATTCAAACAGCATGACATAACTATACCGAGGTTTACAAA

52381  CTTAGCCCAGGGTCCTCCCCATTCACTGATAAAAGCTAACTAAGGTAAATCACAGGTGTT
52381  GAATCGGGTCCCAGGAGGGGTAAGTGACTATTTTCGATTGATTCCATTTAGTGTCCACAA

52441  TATTGTGCAAGAATACTGTGAAACTATTCATGTAATAAAAAGTTTGTGTTAGAATTTCAC
52441  ATAACACGTTCTTATGACACTTTGATAAGTACATTATTTTTCAAACACAATCTTAAAGTG

52501  TCTCTTTAAATCCTGGACATCTTCTGTGAGTGTATCTGAAAGCAGCCAGACATCTTCAGT
52501  AGAGAAATTTAGGACCTGTAGAAGACACTCACATAGACTTTCGTCGGTCTGTAGAAGTCA

52561  ATTTAAGAAACCCTTTACTATTCTGTGTGCTAGGCACTGTGTTAGCTTTGGAATTCAACT
52561  TAAATTCTTTGGGAAATGATAAGACACACGATCCGTGACACAATCGAAACCTTAAGTTGA

52621  GTGAATACTCTAGCCAGACTTCTTGTCTAAATAGTCCAGAATAGGGGTTGCTAAACTTTC
52621  CACTTATGAGATCGGTCTGAAGAACAGATTTATCAGGTCTTATCCCCAACGATTTGAAAG

52681  TTTCTTTTTTTTTGAGATGGAGTTCCACTCTTATTGCCCAGGCTGGAGTGCACTGGTGTG
52681  AAAGAAAAAAAAACTCTACCTCAAGGTGAGAATAACGGGTCCGACCTCACGTGACCACAC

52741  TCTCTGCTGACTGCAACCTCTGCCTCCCAGGTTCAAGCAGTTCTCCTGCCTCAGCCTCCC
52741  AGAGACGACTGACGTTGGAGACGGAGGGTCCAAGTTCGTCAAGAGGACGGAGTCGGAGGG

52801  AAGTAGCTGGGATTACAGGCGCCCACCACTACACTCAGTTAATTTTTGTATTTTTAGTAG
52801  TTCATCGACCCTAATGTCCGCGGGTGGTGATGTGAGTCAATTAAAAACATAAAAATCATC

52861  AGACAGGGTTGTGCTGTTTGCCAGGTTGATCTCAAACTCCTGACCTTTGGTGATCCACCC
52861  TCTGTCCCAACACGACAAACGGTCCAACTAGAGTTTGAGGACTGGAAACCACTAGGTGGG

52921  GCCTCGGCCTCCCAAAGTGTTGGGATTCAGGGTGTGCCACTGTGCCCGGCCTGCTAAACT
52921  CGGAGCCGGAGGGTTTCACAACCCTAAGTCCCACACGGTGACACGGGCCGGACGATTTGA

52981  TTCTTAAATGGCTAGATGATAAATATTTTAGGCTTGAGGGCCATCAGGTTTCTGTTACAA
52981  AAGAATTTACCGATCTACTATTTATAAAATCCGAACTCCCGGTAGTCCAAAGACAATGTT

53041  CTAGTTGGCTGTGCTGCCACTGAAGATACAAGAATGGCGTGACTATGTTCCAATAACATT
53041  GATCAACCGACACGACGGTGACTTCTATGTTCTTACCGCACTGATACAAGGTTATTGTAA
```

FIGURE 4 (cont'd)

```
53101  TTAATTTCAAAACAGGCCAGCTGGCTGCATTTGGTCTGCAAGTAAAGTCTGCCTGACCCC
53101  AATTAAAGTTTTGTCCGGTCGACCGACGTAAACCAGACGTTCATTTCAGACGGACTGGGG

53161  CTACCTAGTCCAGAGGACAACGGGAGAGAAAAGGGATTCAAAGATAAAAATAATTTAGCT
53161  GATGGATCAGGTCTCCTGTTGCCCTCTCTTTTCCCTAAGTTTCTATTTTTATTAAATCGA

53221  GGAATATATTTCTTTAAATAAACTTATTATACTTAGTAAAAAGTCTTAAATTAACGTTTT
53221  CCTTATATAAAGAAATTTATTTGAATAATATGAATCATTTTTCAGAATTTAATTGCAAAA

53281  TACTTGAATTAAATAGTGGTAAAACAGGCTGGGCACAGTGGCTCATGCCTGTAATCTTAG
53281  ATGAACTTAATTTATCACCATTTTGTCCGACCCGTGTCACCGAGTACGGACATTAGAATC

53341  AACTTTGGGAGGTCGAGGCAGGCAGATTGCTTAAGCCCAGGAGTTCGAGACCAGCCTGGG
53341  TTGAAACCCTCCAGCTCCGTCCGTCTAACGAATTCGGGTCCTCAAGCTCTGGTCGGACCC

53401  CAACATAGTGAGTCCCTATCTCTAAAAAAAATACAAAATAATAATAATAATGGTAGAACA
53401  GTTGTATCACTCAGGGATAGAGATTTTTTTTATGTTTTATTATTATTATTACCATCTTGT

53461  AAGTCAATTTTTTATTGAAACTTGGCATTTTATTGGCATATTACAAAGTAGCATTACTAG
53461  TTCAGTTAAAAAATAACTTTGAACCGTAAAATAACCGTATAATGTTTCATCGTAATGATC

53521  ACTAGCCTGAACATTACAGTAATATTCTGTCCTTAATGCCTTTTGTGTCACTGTAATATG
53521  TGATCGGACTTGTAATGTCATTATAAGACAGGAATTACGGAAAACACAGTGACATTATAC

53581  AGCTTTCTGTTTGTTTTGGAATATGTTTTCAGCCTGAATATTATTCTAAAAATACAATTT
53581  TCGAAAGACAAACAAAACCTTATACAAAAGTCGGACTTATAATAAGATTTTTATGTTAAA

53641  ACTATCATTCATAATATATCAAATGACTCACGTAGCTGCAAGGCAGTGAATTAAACAGAA
53641  TGATAGTAAGTATTATATAGTTTACTGAGTGCATCGACGTTCCGTCACTTAATTTGTCTT

53701  TTAGATCATTTTAAAAATAATGATAAGAGCTGATGATGTAAAGTGAAGGTCTGTATTTAA
53701  AATCTAGTAAAATTTTTATTACTATTCTCGACTACTACATTTCACTTCCAGACATAAATT

53761  TTTGGGAGGGAAAAGACTTTTTTTGTATATTCAAAGAGATGGAGTTTGACATCCCTTCAC
53761  AAACCCTCCCTTTTCTGAAAAAAACATATAAGTTTCTCTACCTCAAACTGTAGGGAAGTG

53821  AACCTGTTACAGGTTGAACTTGCCCCTGTTTTTTTAACTGATCAAAAAGTTCCTTCTCTT
53821  TTGGACAATGTCCAACTTGAACGGGGACAAAAAAATTGACTAGTTTTTCAAGGAAGAGAA

53881  GTGTTTCTTTTTTTCTTCTCTTTTCTCTTCTTTTCTTTTCTTTTTTTTTTTTTCTGAGACA
53881  CACAAAGAAAAAAAGAAGAGAAAAGAGAAGAAAAGAAAAGAAAAAAAAAAAAAGACTCTGT

53941  GAGTCTTGCTCTGTCTCCCAGGCTGGAGTGCAGTGGTGCGATCTTGACTCTCTGCAACCT
53941  CTCAGAACGAGACAGAGGGTCCGACCTCACGTCACCACGCTAGAACTGAGAGACGTTGGA

54001  CCGCCTCCCGGGTTCAAGTGATTTCCTGCCTCAGCCTCCTGAGTAGCTGCGACTACAGG
54001  GGCGGAGGGCCCAAGTTCACTAAAAGGACGGAGTCGGAGGACTCATCGACGCTGATGTCC

54061  TGCATGCCACCACGCCTGGCTAATTTTTAGTATTTTTAGTAGAGACGGGGTTTCACCATG
54061  ACGTACGGTGGTGCGGACCGATTAAAAATCATAAAAATCATCTCTGCCCCAAAGTGGTAC

54121  TTAGCTAGGATGGTCTCAATCTCCAAACCTTGTGATCTGCATGCCTCGGCTTCCCAAAGT
54121  AATCGATCCTACCAGAGTTAGAGGTTTGGAACACTAGACGTACGGAGCCGAAGGGTTTCA

54181  GCTGGTCTCCTGTGTTTTTGTCAGCATTCCACAGATGCTATAAAGTTAGTGGTGGTATCG
54181  CGACCAGAGGACACAAAAACAGTCGTAAGGTGTCTACGATATTTCAATCACCACCATAGC

54241  CATGCAATTTGTACGCTTAATGTTTGGCCATTTGCAGTGGGCAAACTGGCTATGTCGGGA
54241  GTACGTTAAACATGCGAATTACAAACCGGTAAACGTCACCCGTTTGACCGATACAGCCCT

54301  GCTGGTGAGATGCTGCTTGTGGGGAGTTGTTTCCTATTACTCTGATTTATCTTTGTTTAA
54301  CGACCACTCTACGACGAACACCCCTCAACAAAGGATAATGAGACTAAATAGAAACAAATT

54361  ATAAGAGTTCCTTTGTTCACTGTTTCCTCCACACAGAACAGAGGATTTATATCTTAGTGT
54361  TATTCTCAAGGAAACAAGTGACAAAGGAGGTGTGTCTTGTCTCCTAAATATAGAATCACA
```

FIGURE 4 (cont'd)

```
54421  TCATTCTCAACTTTCCTGTTTTCAAAGAGGAAAACATTTGCTTCACTTTGCTCATTTCTT
54421  AGTAAGAGTTGAAAGGACAAAAGTTTCTCCTTTTGTAAACGAAGTGAAACGAGTAAAGAA

54481  CTCCAGTGCACCCAGAGGATTTTGAAGGCTTATGCCAACAATTCTTTGGGGCAGTTGGAG
54481  GAGGTCACGTGGGTCTCCTAAAACTTCCGAATACGGTTGTTAAGAAACCCCGTCAACCTC

54541  TGTAACCATTGCTAAATAGCAGTTCAGTACTTAATGACATTCATTTTAGTTTAACAAATC
54541  ACATTGGTAACGATTTATCGTCAAGTCATGAATTACTGTAAGTAAAATCAAATTGTTTAG

54601  ATTGTGAGTGTTGCCATTTTTTATTAGGGAAAACATCTCTTTCATCTTTCCCAAATACTC
54601  TAACACTCACAACGGTAAAAAATAATCCCTTTTGTAGAGAAAGTAGAAAGGGTTTATGAG

54661  AACATGAGTCCTATGGTGAGAGTGAAAAAGGGTTGTATTCTTTTTTTGCATCATCTTACC
54661  TTGTACTCAGGATACCACTCTCACTTTTTCCCAACATAAGAAAAAAACGTAGTAGAATGG

54721  CCATCTGCTATTTTGTCCCTCTTCTATATATCACTGAAATTTGGTTCTGATATTCTATTA
54721  GGTAGACGATAAAACAGGGAGAAGATATATAGTGACTTTAAACCAAGACTATAAGATAAT

54781  GGCAGTGTACATTGCAGAGATTATGAATAACTGGCCGAATTCTGAATGCAGGCATGATTT
54781  CCGTCACATGTAACGTCTCTAATACTTATTGACCGGCTTAAGACTTACGTCCGTACTAAA

54841  GTTCAGCTCAGAGTGTTAAATTTCTGTGTTAGCTGCTAACCTTTAAAAGTCAATTGATTT
54841  CAAGTCGAGTCTCACAATTTAAAGACACAATCGACGATTGGAAATTTTCAGTTAACTAAA

54901  CCTTTTTTTTTTCCCCTCTGTATAATCTAGCCAGCGAATGCTCAGTAGCTTTTCTGTAA
54901  GGAAAAAAAAAAAGGGGAGACATATTAGATCGGTCGCTTACGAGTCATCGAAAAGACATT

54961  ATAGTGATGATTTTTCTTCCCCAGGCTAAAGACCTAATAGTCACACCAGCTACCATTTTA
54961  TATCACTACTAAAAAGAAGGGGTCCGATTTCTGGATTATCAGTGTGGTCGATGGTAAAAT

55021  AAGGAAAAACCAGACCCCAATAATCTGGTTTTTGGAACTGTGTTCACGGATCATATGCTG
55021  TTCCTTTTTGGTCTGGGGTTATTAGACCAAAAACCTTGACACAAGTGCCTAGTATACGAC

55081  ACGGTGGAGTGGTCCTCAGAGTTTGGATGGGAGAAACCTCATATCAAGCCTCTTCAGAAC
55081  TGCCACCTCACCAGGAGTCTCAAACCTACCCTCTTTGGAGTATAGTTCGGAGAAGTCTTG

55141  CTGTCATTGCACCCTGGCTCATCAGCTTTGCACTATGCAGTGGAAGTAAGTACATGGGAA
55141  GACAGTAACGTGGGACCGAGTAGTCGAAACGTGATACGTCACCTTCATTCATGTACCCTT

55201  TTAAAGAGAGTGACATGCTTGCACTTCACTGTGGGTACTAAGTAGCTTCTCCTGCTAAAA
55201  AATTTCTCTCACTGTACGAACGTGAAGTGACACCCATGATTCATCGAAGAGGACGATTTT

55261  TAGCTCTTGGGCAGTATGGGCTTTATCATATTTACCAGAGGAAGCCGAAAATGTCTTTAA
55261  ATCGAGAACCCGTCATACCCGAAATAGTATAAATGGTCTCCTTCGGCTTTTACAGAAATT

55321  AAGTTTTCCTATAGACAGTTGGATTGAAGATAAATCTTAGCAGCTAATTTTTCAGATAAT
55321  TTCAAAAGGATATCTGTCAACCTAACTTCTATTTAGAATCGTCGATTAAAAAGTCTATTA

55381  CTTGCATTTTCTGTGACGTCAATAAAAAAATTATTTAAAGCAAATAGAAAGGACAGCTGG
55381  GAACGTAAAAGACACTGCAGTTATTTTTTTAATAAATTTCGTTTATCTTTCCTGTCGACC

55441  TGGGAAAATTGGCAATTAAAAAGCCAGCACTGTTGATAATTTAATAAGTCACCTTTCTTC
55441  ACCCTTTTAACCGTTAATTTTTCGGTCGTGACAACTATTAAATTATTCAGTGGAAAGAAG

55501  TTAGATCTTCTAATAATTGAGAAAGGGTCATTGAGAACAATATTATGATTTGATTTTAAT
55501  AATCTAGAAGATTATTAACTCTTTCCCAGTAACTCTTGTTATAATACTAAACTAAAATTA

55561  GCTAACTCCACATTTTTATGACCAGTAATAATGTCTAGTTATGCATGTTAAGGGACATGT
55561  CGATTGAGGTGTAAAAATACTGGTCATTATTACAGATCAATACGTACAATTCCCTGTACA

55621  AATTAAATGTCCTCCTTCAGCCATGAGTTGAAATAATTTTAAGCATAAAATGTTTAAAGA
55621  TTAATTTACAGGAGGAAGTCGGTACTCAACTTTATTAAAATTCGTATTTTACAAATTTCT

55681  ATTCTTTCTCTGGTCCTTTTCAGGGTGCACAATTAACCAGATGCACAGTATACTTCCCAT
55681  TAAGAAAGAGACCAGGAAAAGTCCCACGTGTTAATTGGTCTACGTGTCATATGAAGGGTA

55741  TGTCTTGATTCTCATTGTATTTTACTTTTGAATGATTTGGACATCTTTCCCAGAATAATA
55741  ACAGAACTAAGAGTAACATAAAATGAAAACTTACTAAACCTGTAGAAAGGGTCTTATTAT
```

FIGURE 4 (cont'd)

```
55801  GGGGTGCTGGAAATAATAATAGTCTTCCTATTAAGTGCTAAAATATGGTAATAGTAAAAT
55801  CCCCACGACCTTTATTATTATCAGAAGGATAATTCACGATTTTATACCATTATCATTTTA

55861  ATTAATAGATTAATGATAAAATAACAACTTTTGTTTACTGAATTCCTGCAATACGTTCTG
55861  TAATTATCTAATTACTATTTTATTGTTGAAAACAAATGACTTAAGGACGTTATGCAAGAC

55921  TATTATGTTCCATATATGAATCATCTCATTTAATTACCTTAAGTGAGAAGGCATGTGATT
55921  ATAATACAAGGTATATACTTAGTAGAGTAAATTAATGGAATTCACTCTTCCGTACACTAA

55981  AAGCATTCAGGCTTTGAAGTCAGACATCCTGGGCTCAAATTCGGTCTCATTTGCTCCATA
55981  TTCGTAAGTCCGAAACTTCAGTCTGTAGGACCCGAGTTTAAGCCAGAGTAAACGAGGTAT

56041  GGAACCTTTATGACTTGAGGAATTTAATCATTTGTGCTTTAATTTCTTTATCTATAAAAT
56041  CCTTGGAAATACTGAACTCCTTAAATTAGTAAACACGAAATTAAAGAAATAGATATTTTA

56101  TGGGGTAGTTATAATATCTATCTTACAGTGATCTCATAAGTATAAAATAAAGTAATTTAT
56101  ACCCCATCAATATTATAGATAGAATGTCACTAGAGTATTCATATTTTATTTCATTAAATA

56161  CTAAGCATATAGAGTCATATCTGGTGCATAGCGAGAGCCCAGTTGTTAGCAATAATAATT
56161  GATTCGTATATCTCAGTATAGACCACGTATCGCTCTCGGGTCAACAATCGTTATTATTAA

56221  AACTCTTGAAGCAGTGGTCTCATCCATTATAAAAGGAGGCTCAGAAGAGTTGTTTCTTGA
56221  TTGAGAACTTCGTCACCAGAGTAGGTAATATTTTCCTCCGAGTCTTCTCAACAAAGAACT

56281  CTAATATTACACAGCAGTAAGGGTGGAGCTAAGGTTCCAATGAGAAGCTTCTTTCTGTTA
56281  GATTATAATGTGTCGTCATTCCCACCTCGATTCCAAGGTTACTCTTCGAAGAAAGACAAT

56341  TATAGCCCCTACATGTATACACCAAAATTACTAGGATGCTTTTGCAATTAACTCCTTCAA
56341  ATATCGGGGATGTACATATGTGGTTTTAATGATCCTACGAAAACGTTAATTGAGGAAGTT

56401  AATTATAAATTCATATAATTTTGGAAAGATATTCACTGTCCATATGTATCCACAAAATGT
56401  TTAATATTTAAGTATATTAAAACCTTTCTATAAGTGACAGGTATACATAGGTGTTTTACA

56461  TGAATTTAGGAATCTAAGTTTCTGTTCCCAGCTCTTCCTGTTATGTCAATTTGAATGAGT
56461  ACTTAAATCCTTAGATTCAAAGACAAGGGTCGAGAAGGACAATACAGTTAAACTTACTCA

56521  TCTTTCATTATTTGCCACAAGCCCTCACCATGTCATAAGGTATAATAAAGAACAGGTGAA
56521  AGAAAGTAATAAACGGTGTTCGGGAGTGGTACAGTATTCCATATTATTTCTTGTCCACTT

56581  ATGAAGTGCTTTTATTTATTTATTTTTCAGACAGATTCTTGCTCTGTCACCCAGGCTGGA
56581  TACTTCACGAAAATAAATAAATAAAAAGTCTGTCTAAGAACGAGACAGTGGGTCCGACCT

56641  GTGCAGTGGCACGATCGTGGCTCACTGCAGCCTCAAACTCCTGGGCTCAAGCAAACCTCG
56641  CACGTCACCGTGCTAGCACCGAGTGACGTCGGAGTTTGAGGACCCGAGTTCGTTTGGAGC

56701  TACCTCAGCCTCCTGAGTAGCTGGGACCACAGACACATGTCACCATGTCTGGCTAATTTT
56701  ATGGAGTCGGAGGACTCATCGACCCTGGTGTCTGTGTACAGTGGTACAGACCGATTAAAA

56761  TACATATTTTGTTGAGACTGGGTTTTGCCATGTTGCCCAGGCTGGTCTCAAACTCCTGGG
56761  ATGTATAAAACAACTCTGACCCAAAACGGTACAACGGGTCCGACCAGAGTTTGAGGACCC

56821  CTCAAGCAATCCTCTTGCCTTGGCTTCTTCAAGTGCTGGGATGACAGCCATGAGCCACTG
56821  GAGTTCGTTAGGAGAACGGAACCGAAGAAGTTCACGACCCTACTGTCGGTACTCGGTGAC

56881  CTCCTGGCTATGTGCTTTTAAATTTGAAAGGAATAATATGCCCAAGCCCAGGTCACTGGT
56881  GAGGACCGATACACGAAAATTTAAACTTTCCTTATTATACGGGTTCGGGTCCAGTGACCA

56941  TCTCAAACTTAATGTGCATCAGAATCAGCTGGCTACATCTCCAGAGTTTCTCATTTCTAA
56941  AGAGTTTGAATTACACGTAGTCTTAGTCGACCGATGTAGAGGTCTCAAAGAGTAAAGATT

57001  CAAGTGTCCAGATCATGTTAATTTTGTCGCTCCAGGGACCACATTTTGGGAATTGCTGGC
57001  GTTCACAGGTCTAGTACAATTAAAACAGCGAGGTCCCTGGTGTAAAACCCTTAACGACCG

57061  CTATGGGAAGTAAATGAGAAATGATTTAATGTCATTTTTGGAAGTTAAAAAAGATTTTGG
57061  GATACCCTTCATTTACTCTTTACTAAATTACAGTAAAAACCTTCAATTTTTTCTAAAACC
```

FIGURE 4 (cont'd)

```
57121  TGGTCACTGGTTTTTGAAAACTCTTTGAAGGCTGGGTGGGGTGGTTCACATCTGTAATTT
57121  ACCAGTGACCAAAAACTTTTGAGAAACTTCCGACCCACCCCACCAAGTGTAGACATTAAA

57181  CAGCAGTTTCGGGGAGATGGGGTTCTTGGGAAATGAGCCCTCCACCCATGGGATCTGATG
57181  GTCGTCAAAGCCCCTCTACCCCAAGAACCCTTTACTCGGGAGGTGGGTACCCTAGACTAC

57241  CTATTTCCACGTAGATAATGTCAGAGTCGTTCAAACCAGAGTGACTCCATCCTGAATAGG
57241  GATAAAGGTGCATCTATTACAGTCTCAGCAAGTTTGGTCTCACTGAGGTAGGACTTATCC

57301  GGCTGGGTAAAATAAGGCTGAGACCTACTGGGCTGCAACATGATGAGACCCTGACTCTAT
57301  CCGACCCATTTTATTCCGACTCTGGATGACCCGACGTTGTACTACTCTGGGACTGAGATA

57361  TGAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAAACGGCCAGGCGAGGTGGCT
57361  ACTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTTTGCCGGTCCGCTCCACCGA

57421  CACGCCTGTAATACCAGCACTTTGGGAGGCCAAGGCGGGCGGATCACGAGGTCAGGAGAT
57421  GTGCGGACATTATGGTCGTGAAACCCTCCGGTTCCGCCCGCCTAGTGCTCCAGTCCTCTA

57481  CGAGACCATCCTGGCTAACATGGTGAAACCCCATCTCTACTAAAAATAGAAAAAAAAAAA
57481  GCTCTGGTAGGACCGATTGTACCACTTTGGGGTAGAGATGATTTTTATCTTTTTTTTTTT

57541  AAAATTAGCCGGGCGTGGTGGCAGGCGCCTGTAGTCCAAGCTACTTAGGAGGCTGAGGCA
57541  TTTTAATCGGCCCGCACCACCGTCCGCGGACATCAGGTTCGATGAATCCTCCGACTCCGT

57601  GGAAAATGACGTGAACCCAGGAGGCGGAGTTTGCAGTGAGCCGAGATCACGCCGCTGCAC
57601  CCTTTTACTGCACTTGGGTCCTCCGCCTCAAACGTCACTCGGCTCTAGTGCGGCGACGTG

57661  TCCACACTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAACAAAAACCTCTTTGG
57661  AGGTGTGACCCACTGTCTCGTTCTGAGGTAGAGTTTTTTTTTTTTGTTTTTGGAGAAACC

57721  CCATCTATTTAGATGAGTTGGAATGCTACTTTGATTGGTTTTGCCAATTTAGATTTTTGT
57721  GGTAGATAAATCTACTCAACCTTACGATGAAACTAACCAAAACGGTTAAATCTAAAAACA

57781  AATTAGGGTAGTTTAAAAAAACTCTCAAATTACTGTTATCCTTTGTTAAAAACCAAAATC
57781  TTAATCCCATCAAATTTTTTTGAGAGTTTAATGACAATAGGAAACAATTTTTGGTTTTAG

57841  TTTTATTATTAGTCTTCTGCTTACAATATAAGTAAAATAAATGTTTTGTATTAATAACTG
57841  AAAATAATAATCAGAAGACGAATGTTATATTCATTTTATTTACAAAACATAATTATTGAC

57901  GATTGTATATTGCATGGTTACCACTATGTAACAAAATTAATAGCTTAAAATTACAAGCGT
57901  CTAACATATAACGTACCAATGGTGATACATTGTTTTAATTATCGAATTTTAATGTTCGCA

57961  CATTTTAAAATTATTTTTCATGTTTCTGGGGTTGACTAGGATGGCTAAGTGGTTTTTGCT
57961  GTAAAATTTTAATAAAAAGTACAAAGACCCCAACTGATCCTACCGATTCACCAAAAACGA

58021  TGGGGTCCGTCACACGGTTGCTGTCAGTGGCTGGGGTTAGAGTCAGCTCACAGGCTTCCT
58021  ACCCCAGGCAGTGTGCCAACGACAGTCACCGACCCCAATCTCAGTCGAGTGTCCGAAGGA

58081  AACTCATATGGTGCTGGAGGGAGTGTCTCTTAGCATGCTAATGCTTCATAATTTGTGTAT
58081  TTGAGTATACCACGACCTCCCTCACAGAGAATCGTACGATTACGAAGTATTAAACACATA

58141  AATGAACAGTGAGGATGACCAGAGGTCACTTTCATTGCCATCTGGGCTGCAAAACTCAAG
58141  TTACTTGTCACTCCTACTGGTCTCCAGTGAAAGTAACGGTAGACCCGACGTTTTGAGTTC

58201  GAGTTGAAGGCTCTTCCGGCATTACTTGTTCTCCATGTGGTCTCCTCAGGATGGCAGTCA
58201  CTCAACTTCCGAGAAGGCCGTAATGAACAAGAGGTACACCAGAGGAGTCCTACCGTCAGT

58261  CAGAGTAGGCAGCCTTCTTATGTGCTGACCTAGGGCCCAAAATCTGTATGTGTCAGTTGA
58261  GTCTCATCCGTCGGAAGAATACACGACTGGATCCCGGGTTTTAGACATACACAGTCAACT

58321  CTGAGCCCAGATGGAAGCTCTATTGCTGTTTACGCCCTAGCCTCAAAACTACTCTTTCAA
58321  GACTCGGGTCTACCTTCGAGATAACGACAAATGCGGGATCGGAGTTTTGATGAGAAAGTT

58381  GAAAAATTCTCCAACCGAGTTTTTCCTCTACTCTCACACCACCACAACAATAGTCAACAT
58381  CTTTTTAAGAGGTTGGCTCAAAAAGGAGATGAGAGTGTGGTGGTGTTGTTATCAGTTGTA

58441  AGGAGACTTCTATGATCAAAGGTGTAGGGATTTCTCCCCACCACCAAGCAGTGAACCCCA
58441  TCCTCTGAAGATACTAGTTTCCACATCCCTAAAGAGGGGTGGTGGTTCGTCACTTGGGGT
```

FIGURE 4 (cont'd)

```
58501  GCTGGGTGTCCTCCAATTCAATTCTCACACTGTTATTACGGGAAAGGTGTCCCAATCCAG
58501  CGACCCACAGGAGGTTAAGTTAAGAGTGTGACAATAATGCCCTTTCCACAGGGTTAGGTC

58561  TCCCCAAGAGGTTCTTGGATGTTGCACAAGAAAGAATTTGGGGAGAGTCCACAAAGTAAA
58561  AGGGGTTCTCCAAGAACCTACAACGTGTTCTTTCTTAAACCCCTCTCAGGTGTTTCATTT

58621  ATAAAATTTATTAAGAAAGTAAAGGAAAAAAAGAATGGCTACTCCGTAGACAGAGGAGCT
58621  TATTTTAAATAATTCTTTCATTTCCTTTTTTTCTTACCGATGAGGCATCTGTCTCCTCGA

58681  CCGAGGGCTGCTGGTTGGCTATTTTAATGGTTATTTCTTGATCATTTGCTAAAGAAGGGG
58681  GGCTCCCGACGACCAACCGATAAAATTACCAATAAAGAACTAGTAAACGATTTCTTCCCC

58741  TAGATTATTCATAAATTTTCCAGAAAAGGGATGGGCAATTCCTGGAACTGAGGGTTCCTC
58741  ATCTAATAAGTATTTAAAAGGTCTTTTCCCTACCCGTTAAGGACCTTGACTCCCAAGGAG

58801  CCCTTTTTAGACCTTGTAGGTTAACTTCCCACCATTGCCATGGCCTCTGTAAATTGTCGT
58801  GGGAAAAATCTGGAACATCCAATTGAAGGGTGGTAACGGTACCGGAGACATTTAACAGCA

58861  GACACTGGTAGGAGTGTCTGTTAGCATGCTAATGCATTATAATTAGCTTATAATGAGCAG
58861  CTGTGACCATCCTCACAGACAATCGTACGATTACGTAATATTAATCGAATATTACTCGTC

58921  CGAGGACGGCCAGACATCCCTTTCATTGCCATATGGGATTTGGCAGGTTTTGGCTGGCTT
58921  GCTCCTGCCGGTCTGTAGGGAAAGTAACGGTATACCCTAAACCGTCCAAAACCGACCGAA

58981  CTTTACCACATCTTGTTTTATCAGTGGGGTCTTTGTAACCTGTATGTTGTGCCGACCTCC
58981  GAAATGGTGTAGAACAAAATAGTCACCCCAGAAACATTGGACATACAACACGGCTGGAGG

59041  TATCTCATCCTGTGACCTAGAATACTAACCTCCTGAGAATGCAGCCCAGTAGGTTTCAGC
59041  ATAGAGTAGGACACTGGATCTTATGATTGGAGGACTCTTACGTCGGGTCATCCAAAGTCG

59101  CTTATTTTACCCAGCCCCTAATCAGGATGGAGTCGCTCTGGTCCAAACACCTCTGACATT
59101  GAATAAAATGGGTCGGGGATTAGTCCTACCTCAGCGAGACCAGGTTTGTGGAGACTGTAA

59161  ATCCACCTGGAAATAGTGTCAGATCCCATGGGTTAAGGGCTCATTCCCCAAGACCCTAAC
59161  TAGGTGGACCTTTATCACAGTCTAGGGTACCCAATTCCCGAGTAAGGGGTTCTGGGATTG

59221  CACCCTCACCTGGCCCAGACACAAGTCACAAGTCCTGTCCTCTGGAATGTCTGACTGAGA
59221  GTGGGAGTGGACCGGGTCTGTGTTCAGTGTTCAGGACAGGAGACCTTACAGACTGACTCT

59281  GGCATCAAGTTGGAGTTCCCATGACCCCCTCTTCTGGTTTGATTAATTTACTGGAGCAGC
59281  CCGTAGTTCAACCTCAAGGGTACTGGGGGAGAAGACCAAACTAATTAAATGACCTCGTCG

59341  TCACAAAACTCAGGGAAACACTTAGGTTTAGTGGTTTATTATAAAGGATACTGCAAAGTG
59341  AGTGTTTTGAGTCCCTTTGTGAATCCAAATCACCAAATAATATTTCCTATGACGTTTCAC

59401  TATAGATGAAGCGATGTGTAGGGTGAGGTATCGGGGGATAGAGTGCAGGGCTTCCATGCC
59401  ATATCTACTTCGCTACACATCCCACTCCATAGCCCCCTATCTCACGTCCCGAAGGTACGG

59461  CTCCCGGGGTGTGCCACCCTTTAGGAACCTTCATGTGTTCTGCTATCCGGAAGCTCACTG
59461  GAGGGCCCCACACGGTGGGAAATCCTTGGAAGTACACAAGACGATAGGCCTTCGAGTGAC

59521  AACTCATTCTTTTGGGTTTTTATGGAAGCTTCACGTCAATATTCCTTCCCCTGGGGTATG
59521  TTGAGTAAGAAAACCCAAAAATACCTTCGAAGTGCAGTTATAAGGAAGGGGACCCCATAC

59581  AGGCAGGACCTTCACCGGAGAGGGTCTTTTATTTTATTTCATTATTTTTAAAATATAAGA
59581  TCCGTCCTGGAAGTGGCCTCTCCCAGAAAATAAAATAAAGTAATAAAAATTTTATATTCT

59641  TGGGGTCTTGCTATGTAGCCTAGGGTGGTATTGAATTCCTGGGCCCAAGTAAGCCTCCCG
59641  ACCCCAGAACGATACATCGGATCCCACCATAACTTAAGGACCCGGGTTCATTCGGAGGGC

59701  CCTCTGCCTCCGAAAGTGCTGGGACTACAGGCATGAGCCACCCAGCCCCTTTGGAGAGGA
59701  GGAGACGGAGGCTTTCACGACCCTGATGTCCGTACTCGGTGGGTCGGGGAAACCTCTCCT

59761  TCTTAAGACTCACAGTTAGAAAGGAGGGCAGGAGAAGGTCAGAGAGATTCTGCTTCCTGC
59761  AGAATTCTGAGTGTCAATCTTTCCTCCCGTCCTCTTCCAGTCTCTCTAAGACGAAGGACG
```

FIGURE 4 (cont'd)

```
59821  CCCTGAGGCCTAACACACCCAGCATTATAACAAAAGACTGTAACAAGGGCTATGGGAGTT
59821  GGGACTCCGGATTGTGTGGGTCGTAATATTGTTTTCTGACATTGTTCCCGATACCCTCAA

59881  ACCAGCCAGGAACTGTAAATCACACAGTATTATTTCTGCCTTACTCTTATCAGTCATGAT
59881  TGGTCGGTCCTTGACATTTAGTGTGTCATAATAAAGACGGAATGAGAATAGTCAGTACTA

59941  GGTTACAAAGGCCTGTCCAAGTTCAAAAGGAGGAAAAAAGCACTCCACTTCTTGATGAGG
59941  CCAATGTTTCCGGACAGGTTCAAGTTTTCCTCCTTTTTTCGTGAGGTGAAGAACTACTCC

60001  AAGGTCAGGGTTCTGGAAGAGCATGGGGGACTGGAAATATTGTTCGGCCATTTCATGGAA
60001  TTCCAGTCCCAAGACCTTCTCGTACCCCCTGACCTTTATAACAAGCCGGTAAAGTACCTT

60061  AATACAGTCTGTAATAGACATCAGTGATAAATGGGGAGATATGGTGGATGTATTGTTTTC
60061  TTATGTCAGACATTATCTGTAGTCACTATTTACCCCTCTATACCACCTACATAACAAAAG

60121  CTCTGAGTGCCAAATTGGGTTCTAAGATTTCACAGGCCAGATATTATTGAAGTGGCTACG
60121  GAGACTCACGGTTTAACCCAAGATTCTAAAGTGTCCGGTCTATAATAACTTCACCGATGC

60181  TTGTCTGGGGTATATACCTTGGGGTTCGTCATTTTGTGCCAGGAAAATTTAGGACAAGGA
60181  AACAGACCCCATATATGGAACCCCAAGCAGTAAAACACGGTCCTTTTAAATCCTGTTCCT

60241  CTCACACGAGGAGTTTAGGAGTGGAGGTTTAATAGGCAGAAGAGAAACAGAAACAGTTCT
60241  GAGTGTGCTCCTCAAATCCTCACCTCCAAATTATCCGTCTTCTCTTTGTCTTTGTCAAGA

60301  TTCTGTAGGGAGAGAGGGATCTTCGAGGGAAAAGACCATCCAGCAGCAGATGAGCTGGAT
60301  AAGACATCCCTCTCTCCCTAGAAGCTCCCTTTTCTGGTAGGTCGTCGTCTACTCGACCTA

60361  TTTATAGTCAGGTTTGAGGAGGAAGTGTCTGATTTACTTAGGGCTAACAGGTTGGTTTGA
60361  AAATATCAGTCCAAACTCCTCCTTCACAGACTAAATGAATCCCGATTGTCCAACCAAACT

60421  TCAGGTATGACGTTTACATAGCACACCGGGAAGTCTGGTTGCCCCACCCTAATCCTTTTA
60421  AGTCCATACTGCAAATGTATCGTGTGGCCCTTCAGACCAACGGGGTGGGATTAGGAAAAT

60481  TGCAAATGGACTTTCCTAGTTGATTGGAGCTATCTTGTCTGCTCCTTACTGTACAGGTGG
60481  ACGTTTACCTGAAAGGATCAACTAACCTCGATAGAACAGACGAGGAATGACATGTCCACC

60541  CTGACAAAGAGAAGGGAAGATGGAGCCGCCATCTTGAACATGATTGACGCAGCTGTGGAT
60541  GACTGTTTCTCTTCCCTTCTACCTCGGCGGTAGAACTTGTACTAACTGCGTCGACACCTA

60601  ATCTATGTCTGCAGCTTGATTTTACAGGCTGCTCTTCTTTAGAAAGGGGCTGCTTTTCAT
60601  TAGATACAGACGTCGAACTAAAATGTCCGACGAGAAGAAATCTTTCCCCGACGAAAAGTA

60661  TAAAACAAAAACCTTACTGAGGACTCTTGTACCCTCACTATCTGCAGGTGATTTCTTAAC
60661  ATTTTGTTTTTGGAATGACTCCTGAGAACATGGGAGTGATAGACGTCCACTAAAGAATTG

60721  TCCTGTATCATTATTATAGTGTCATAGGAATCTAGGAAGTGTGATTGCCTTCTGGAGAGA
60721  AGGACATAGTAATAATATCACAGTATCCTTAGATCCTTCACACTAACGGAAGACCTCTCT

60781  ACTTCAGGGTCCCTGGATTTTCTCACTAATCATATTCATTTCATTCCCTGTTTACCTGTT
60781  TGAAGTCCCAGGGACCTAAAAGAGTGATTAGTATAAGTAAAGTAAGGGACAAATGGACAA

60841  TAATAAATGTTGAGTGGCTACTATGTGCCCAGGGCTGTGCAAGGTGCCTGGAATAGGATG
60841  ATTATTTACAACTCACCGATGATACACGGGTCCCGACACGTTCCACGGACCTTATCCTAC

60901  GCAAGCAAACCAACCATGACCTTGCTCTCAAAGAATTGTTTTCTTTCTTTCTTTTTTTTC
60901  CGTTCGTTTGGTTGGTACTGGAACGAGAGTTTCTTAACAAAAGAAAGAAAGAAAAAAAAG

60961  TGAGACAAGGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCCTGATCTCAGCTCACT
60961  ACTCTGTTCCAGAACGAGACAGCGGGTCCGACCTCACGTCACCGGACTAGAGTCGAGTGA

61021  GCAACTTCCACCTCCTGGCTTCAACAGATTCTCCTCCCTCAGCCTCCCAGGTAGCTGGGA
61021  CGTTGAAGGTGGAGGACCGAAGTTGTCTAAGAGGAGGGAGTCGGAGGGTCCATCGACCCT

61081  TTATAGGTGCCCACCACCATGCCCGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTC
61081  AATATCCACGGGTGGTGGTACGGGCCGATTAAAAACATAAAAATCATCTCTGTCCCAAAG

61141  ATCATGTTGGCCAGGCTAGTCTCAAACTCCTGACCTCAAGTGATCCGCCTGCCTCAGTGC
61141  TAGTACAACCGGTCCGATCAGAGTTTGAGGACTGGAGTTCACTAGGCGGACGGAGTCACG
```

FIGURE 4 (cont'd)

```
61201  CGAGATTACAGGCATGAGCCACTGCACCTGGCCTGTTTTCTGTCTTTTTTTTTTAACCCC
61201  GCTCTAATGTCCGTACTCGGTGACGTGGACCGGACAAAAGACAGAAAAAAAAAATTGGGG

61261  CACATGGAAAGTTAGGCAGGGATTTAATTTCAAAGAGTTTCTTCCCAAATATTGTGGTGA
61261  GTGTACCTTTCAATCCGTCCCTAAATTAAAGTTTCTCAAAGAAGGGTTTATAACACCACT

61321  AACATCTAGTCCATTTATGTTTCAGGCTTTTTCATAGATATTACTGTTTCTCTTAGACTG
61321  TTGTAGATCAGGTAAATACAAAGTCCGAAAAAGTATCTATAATGACAAAGAGAATCTGAC

61381  ACAGACTTAGGACATCACATGATTGAATGTCAACCTGACTTTTAAGTTCAATAATAGTTT
61381  TGTCTGAATCCTGTAGTGTACTAACTTACAGTTGGACTGAAAATTCAAGTTATTATCAAA

61441  CATTTTTTAAATGCAAAATTAAATAGGAGTAATATCCTTGAACATGTTTCTCCTTCACTC
61441  GTAAAAAATTTACGTTTTAATTTATCCTCATTATAGGAACTTGTACAAAGAGGAAGTGAG

61501  ATCGGACCTATTGAGTCTAGAATCCATTTTATTTTGAGAGCTCCAACTACATGTGGTTTC
61501  TAGCCTGGATAACTCAGATCTTAGGTAAAATAAAACTCTCGAGGTTGATGTACACCAAAG

61561  TGTTAATATCTGAAAGTAATTTTTAATATCTGAAAGTTTCAAATAGTGATAATATTTGAA
61561  ACAATTATAGACTTTCATTAAAAATTATAGACTTTCAAAGTTTATCACTATTATAAACTT

61621  GGTTTCAAATAGTGGAGGAGAGACGCAAACTGCGAGTAGGGAACGCGGATTGAAGGAAAT
61621  CCAAAGTTTATCACCTCCTCTCTGCGTTTGACGCTCATCCCTTGCGCCTAACTTCCTTTA

61681  GAAAAGTTGATAATAGGCACTTCATTTTCAGATTATAACAAATATTTAGTGAAATAAAGC
61681  CTTTTCAACTATTATCCGTGAAGTAAAAGTCTAATATTGTTTATAAATCACTTTATTTCG

61741  ATAATTTTAAAAAATGGATAAGACATGTATGTACTATTTAATGTGTACTAATAAATGAAT
61741  TATTAAAATTTTTTACCTATTCTGTACATACATGATAAATTACACATGATTATTTACTTA

61801  GCTCATATTTGCCATCCAGATAAAGAAATAGAACAGTTGAACATCACAGAAGCCCTCGCG
61801  CGAGTATAAACGGTAGGTCTATTTCTTTATCTTGTCAACTTGTAGTGTCTTCGGGAGCGC

61861  TGTCCCATTCAGATCACAAACTCTATCTTTCTACTGGTAACCATTATTCTTACTTTTGTG
61861  ACAGGGTAAGTCTAGTGTTTGAGATAGAAAGATGACCATTGGTAATAAGAATGAAAACAC

61921  ATAACCTTTCTTTGCTTTCCTTTATGGAGTGTGTAGATTCTTAACACTATAGTTTTGCCT
61921  TATTGGAAAGAAACGAAAGGAAATACCTCACACATCTAAGAATTGTGATATCAAAACGGA

61981  CTTTTTGAACTTCATGTAAATGAAAAAATGTATATATTTTGTGTTTGTGTGCATGTGTGT
61981  GAAAAACTTGAAGTACATTTACTTTTTTACATATATAAAACACAAACACACGTACACACA

62041  GTGTGTGCTTGCACACGTGTGAATACCCCTTATTGTTTGAGCTTAGTATTATGTAAGATC
62041  CACACACGAACGTGTGCACACTTATGGGGAATAACAAACTCGAATCATAATACATTCTAG

62101  TAACCAGGAAGGACCAACCTAACATTTGTGGGGCCTGGGACCTAAGTATAGACAGAGATG
62101  ATTGGTCCTTCCTGGTTGGATTGTAAACACCCCGGACCCTGGATTCATATCTGTCTCTAC

62161  ACATGGTCCACTCCATATCTGTTCCCATCCCAGTTCTGCCCTGTGCTACAAGGGGTCTTC
62161  TGTACCAGGTGAGGTATAGACAAGGGTAGGGTCAAGACGGGACACGATGTTCCCCAGAAG

62221  CAGACACTGCATAGGTACCCATACCACCTGCAGGCAGATGACCTTTGCCTGTCCTTTGGA
62221  GTCTGTGACGTATCCATGGGTATGGTGGACGTCCGTCTACTGGAAACGGACAGGAAACCT

62281  GTTAGGCCTGGGTACCCTACTAGTGGTGTCTGCTCTTGGGATGATAAACCCGAGGGGCTT
62281  CAATCCGGACCCATGGGATGATCACCACAGACGAGAACCCTACTATTTGGGCTCCCCGAA

62341  ATCAGCTTTATATATATATATATATATATATATATATATTTTTTTTTTTTTTTTTTTAG
62341  TAGTCGAAATATATATATATATATATATATATATATATAAAAAAAAAAAAAAAAAAATC

62401  CTAGCATCTCTCTCGTTCTTCCAGGCTGGAGAGCAGTGGTGTGATCTTGGCTCACTGTAA
62401  GATCGTAGAGAGAGCAAGAAGGTCCGACCTCTCGTCACCACACTAGAACCGAGTGACATT

62461  TCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGAGATTAC
62461  AGAGGTGGAGGACCCAAGTTCGCTAAGAGGACGGAGTCGGAGGGTTCATCGACTCTAATG
```

FIGURE 4 (cont'd)

```
62521  AGGCACGTGCCACTATGCCTGGCTAATTTTTGTATTTTTAGTAGAGACTGGGTTTCACCA
62521  TCCGTGCACGGTGATACGGACCGATTAAAAACATAAAAATCATCTCTGACCCAAAGTGGT

62581  TGTTGGCATGTGAGCCACTATGCCCGGCCTCTTTTTTTTTTTTTTTTAAGAGACAGGGTC
62581  ACAACCGTACACTCGGTGATACGGGCCGGAGAAAAAAAAAAAAAAAATTCTCTGTCCCAG

62641  TTACTATGTTGCCCAGGCTGGTGTTGAACTCCTGGCCTCAATAGATCCTCTGTCCTTGGT
62641  AATGATACAACGGGTCCGACCACAACTTGAGGACCGGAGTTATCTAGGAGACAGGAACCA

62701  CTCCCAAAGTGCTGGCTCAGTCATTATGTTCTTTACTGCCATCCATTGCAGTATAAAAAG
62701  GAGGGTTTCACGACCGAGTCAGTAATACAAGAAATGACGGTAGGTAACGTCATATTTTTC

62761  AATAACAGTTTTCCCTCAGAATATTCTTGGTGAAACAGTAACAATTATTAATTTTATTAA
62761  TTATTGTCAAAAGGGAGTCTTATAAGAACCACTTTGTCATTGTTAATAATTAAAATAATT

62821  ATCTTGGCTCTTGAGCAAATGTCATTTTAATATTCTGTGTGATAAAATAGCAATGTCACA
62821  TAGAACCGAGAACTCGTTTACAGTAAAATTATAAGACACACTATTTTATCGTTACAGTGT

62881  TAAAGTGCTTTTACATACTTAAGTACTATGGTTGTCTGAAGAAAAAGGAGTTGTGTGATT
62881  ATTTCACGAAAATGTATGAATTCATGATACCAACAGACTTCTTTTTCCTCAACACACTAA

62941  GTCTTGAGTTTGTAAGTGGAACTACTTTTTTCATCGAGCACCGTTTTTACTTGAAATAAC
62941  CAGAACTCAAACATTCACCTTGATGAAAAAAGTAGCTCGTGGCAAAAATGAACTTTATTG

63001  AAAGTATGAATATGAATGCAGATTTGGCCATTTGGCAGATTTTTATTTATTTATTTGAGA
63001  TTTCATACTTATACTTACGTCTAAACCGGTAAACCGTCTAAAAATAAATAAATAAACTCT

63061  CAAGGTCTTGCTGTGTCACCCAGACTGGGGTGCAGCTCACTGCAGCCTTTACCTCCTGGG
63061  GTTCCAGAACGACACAGTGGGTCTGACCCCACGTCGAGTGACGTCGGAAATGGAGGACCC

63121  CTCAAGTGATCCTCCTGCCTCAGCCTCCTTAGTTGCTGGGACCACAGGTGTGCAGCACCA
63121  GAGTTCACTAGGAGGACGGAGTCGGAGGAATCAACGACCCTGGTGTCCACACGTCGTGGT

63181  TGCTCAGCTAATTTTTAAATTTTTTGTGGAGACAAGGTCTGACCATGTTGCCCAGGCTGG
63181  ACGAGTCGATTAAAAATTTAAAAAACACCTCTGTTCCAGACTGGTACAACGGGTCCGACC

63241  TCTCGAACTCCTGGGCTCCAGCAATCCTCTCGCTTCAGCTTCTCAAAGTGCTGGAATTAC
63241  AGAGCTTGAGGACCCGAGGTCGTTAGGAGAGCGAAGTCGAAGAGTTTCACGACCTTAATG

63301  AGGCGTGAGCCACGGTTCCCAGCTCGCCTTTTCATATGTTTATTGGGTTGCTAAATATGT
63301  TCCGCACTCGGTGCCAAGGGTCGAGCGGAAAAGTATACAAATAACCCAACGATTTATACA

63361  TCTTATTTGAAGTGCCCACTTGTTCAAGTTTTTTGTTCACTTTTTTATTGCTGTCTTTAT
63361  AGAATAAACTTCACGGGTGAACAAGTTCAAAAAACAAGTGAAAAAATAACGACAGAAATA

63421  AATTTTTTGATGACAAGAAATATGTTTAAATTAAAGGCAGTTTTGTCAATCTGCCTTATG
63421  TTAAAAAACTACTGTTCTTTATACAAATTTAATTTCCGTCAAAACAGTTAGACGGAATAC

63481  TTTGGAATTTGTTGTGTCTTTAAAGACTCTTTTCCTGCCCTTATATTATAAAGATATCTT
63481  AAACCTTAAACAACACAGAAATTTCTGAGAAAAGGACGGGAATATAATATTTCTATAGAA

63541  ATAGTATTGATATTGTCATTCACAGTTAGGTCTTTAATCCTAGAAGTTGATTTTTGTATG
63541  TATCATAACTATAACAGTAAGTGTCAATCCAGAAATTAGGATCTTCAACTAAAAACATAC

63601  AAGTAAGGGCCCAATTTATTTTTTTCCACCTATGTATACCCAGTTTTCCCAGCACCATTT
63601  TTCATTCCCGGGTTAAATAAAAAAAGGTGGATACATATGGGTCAAAAGGGTCGTGGTAAA

63661  AAACGGGCGGTACTTTCCCCTCTGATATGTCATGATGACACCTTTGTCATGTACACATGA
63661  TTTGCCCGCCATGAAAGGGGAGACTATACAGTACTACTGTGGAAACAGTACATGTGTACT

63721  GTTTGCTTCTGGATCTCTATTTGTGTTCAGTTGTTTTACTTGTCTGCATAAATGCCACAG
63721  CAAACGAAGACCTAGAGATAAACACAAGTCAACAAAATGAACAGACGTATTTACGGTGTC

63781  TCTTGGTTTCTCAGTCATTATAATGAGCCTTGCCTGCTAGGAAGGGTCCTTCTACTTTGT
63781  AGAACCAAAGAGTCAGTAATATTACTCGGAACGGACGATCCTTCCCAGGAAGATGAAACA
```

FIGURE 4 (cont'd)

```
63841  TCTCCAAGAGTGCCTTGGATATTCTTGGTCATTTTCTTTCCATATAATTTTTAGATTTGG
63841  AGAGGTTCTCACGGAACCTATAAGAACCAGTAAAAGAAAGGTATATTAAAAATCTAAACC

63901  TTTATCAAGTCGATAAACTTTTGTCATGCAGTTTTTTAAAAGAAAATAAAGTGCCCTAAT
63901  AAATAGTTCAGCTATTTGAAAACAGTACGTCAAAAAATTTTCTTTTATTTCACGGGATTA

63961  GTGATTTTTTTTAATTTGGATTACATTGAATACATAAATCAGTTTGGGGAAAATCAACAT
63961  CACTAAAAAAAAATTAAACCTAATGTAACTTATGTATTTAGTCAAACCCCTTTTAGTTGTA

64021  CTTTATAATATTGAGGCTTTTTATCCATGAGCATGAAATTGCTTTCTACATTAATAACTT
64021  GAAATATTATAACTCCGAAAAATAGGTACTCGTACTTTAACGAAAGATGTAATTATTGAA

64081  GGTAAAGTTTTATATTTTCCTTATGGATCTTTCACTTCTTTTGTTAGATTTTTCTCCACC
64081  CCATTTCAAAATATAAAAGGAATACCTAGAAAGTGAAGAAAACAATCTAAAAAGAGGTGG

64141  TAAGTACTGTATCTTTTAATTCTACTGTAAATAGTGTTTTTATTTTTTAACTGCTGCTAT
64141  ATTCATGACATAGAAAATTAAGATGACATTTATCACAAAAATAAAAAATTGACGACGATA

64201  TACATAGGAATATAATTTTGCATCTGGCCAGCTTGCCTAAACAGATGATTTCAAAATATT
64201  ATGTATCCTTATATTAAAACGTAGACCGGTCGAACGGATTTGTCTACTAAAGTTTTATAA

64261  ATCTGTCAGATTTTCTGTGTTCTCTATCTCTGTAATCATATCATCTATGTTTTATGATCT
64261  TAGACAGTCTAAAAGACACAAGAGATAGAGACATTAGTATAGTAGATACAAAATACTAGA

64321  ATGACAGTTATATTTTTTCATCTCCAAGCTGATATTTTAAATTTCATTATTCTTGCTTTA
64321  TACTGTCAATATAAAAAAGTAGAGGTTCGACTATAAAATTTAAAGTAATAAGAACGAAAT

64381  TACTGGTTGTACATATTCAGTTAGAGTGATTGCAGTAGGCAATTTGCGATGTTCCCAATT
64381  ATGACCAACATGTATAAGTCAATCTCACTAACGTCATCCGTTAAACGCTACAAGGGTTAA

64441  TGATAAGGAAAGCTTTTACATTTTACCATTAAGTATAACGTTTGCGAGCTTATTATAGCT
64441  ACTATTCCTTTCGAAAATGTAAAATGGTAATTCATATTGCAAACGCTCGAATAATATCGA

64501  ATCCTTTGTCACTTAGGGAAGTTATCTTTGTTTTTTAGTTATCTAGGAAACTAAAAAATC
64501  TAGGAAACAGTGAATCCCTTCAATAGAAACAAAAAATCAATAGATCCTTTGATTTTTTAG

64561  AGAAAAGTTAGTTTGATTTTATCAGATACATCTGCTACATCTAATGAGATGATCATATAG
64561  TCTTTTCAATCAAACTAAAATAGTCTATGTAGACGATGTAGATTACTCTACTAGTATATC

64621  TTTTTTTCCTCCCATAAACTGTGAATATAATGAATCACATTAATTGATTTTTTTAAATGT
64621  AAAAAAAGGAGGGTATTTGACACTTATATTACTTAGTGTAATTAACTAAAAAAATTTACA

64681  TAAACCAATCTCCATTTCTGGGATAAATTCAACTTAGTCATGTTGTATTAGCCTTTTGTA
64681  ATTTGGTTAGAGGTAAAGACCCTATTTAAGTTGAATCAGTACAACATAATCGGAAAACAT

64741  TATCTTATATCTATAATACTGAATGAGATTATCATACTCTTTACATGTATTTGAAGCCAT
64741  ATAGAATATAGATATTATGACTTACTCTAATAGTATGAGAAATGTACATAAACTTCGGTA

64801  GTCATTAAGTGCATACAATTTTTATTAAATTTTTAATTGACAAATAATTGTATTTATTTA
64801  CAGTAATTCACGTATGTTAAAAATAATTTAAAAATTAACTGTTTATTAACATAAATAAAT

64861  TGGGGTACGATGTAGTGTTTTGATATATGTTTACATTGTAGAGTGACTAAATCAAGCTAA
64861  ACCCCATGCTACATCACAAAACTATATACAAATGTAACATCTCACTGATTTAGTTCGATT

64921  TTAGCATACTCATCACTGAACATATTTATCATTTCTTGTGGTGAGAACATTTGAAATTTA
64921  AATCGTATGAGTAGTGACTTGTATAAATAGTAAAGAACACCACTCTTGTAAACTTTAAAT

64981  ATCTCTTAGCAATTTTGAAATATATGATACATTATTATTAACTACAGTCACCATGCTGTG
64981  TAGAGAATCGTTAAAACTTTATATACTATGTAATAATAATTGATGTCAGTGGTACGACAC

65041  CGATAGGTCTCAAAAACTTATTCCTCCTAATTGAAATTTTGTTCCCTTTGACCAACATCT
65041  GCTATCCAGAGTTTTTGAATAAGGAGGATTAACTTTAAAACAAGGGAAACTGGTTGTAGA

65101  TCCAATTCCCCAACCCCTAGCCTCCTCTGGTAACCACCATTCTACTCTCTACTTCCATTA
65101  AGGTTAAGGGGTTGGGGATCGGAGGAGACCATTGGTGGTAAGATGAGAGATGAAGGTAAT
```

FIGURE 4 (cont'd)

```
65161  GTTTGACTTTTTCAGTCCACATATGAGTGAAATCGTGCAGTATTTGTCTTTCTGTACTTG
65161  CAAACTGAAAAAGTCAGGTGTATACTCACTTTAGCACGTCATAAACAGAAAGACATGAAC

65221  TCTTGTTTCACTTAGCATAATGTCCAACGGGTTCCCCATGTCATTGTAAATGATGGAATT
65221  AGAACAAAGTGAATCGTATTACAGGTTGCCCAAGGGGTACAGTAACATTTACTACCTTAA

65281  TTCTTTTTTATAGCTAAATAACATTCCATCATGTATACATACCACATTTTCTTTATCGAT
65281  AAGAAAAAATATCGATTTATTGTAAGGTAGTACATATGTATGGTGTAAAAGAAATAGCTA

65341  TCATCTGTTGATGGACAGTTAGGTTGCTTCTGTATCCTGGCTATTGTGAATAATGCTACA
65341  AGTAGACAACTACCTGTCAATCCAACGAAGACATAGGACCGATAACACTTATTACGATGT

65401  ATAAACATGAGAGCACAGATACCTCATGCTGATTTCAATTCCTTTGGATAAGGGTCCCCA
65401  TATTTGTACTCTCGTGTCTATGGAGTACGACTAAAGTTAAGGAAACCTATTCCCAGGGGT

65461  AACCCCAGGCTATGGAACAGCAGTGGTATGTGTGGCCTGTTAGGTACTGGATCGCACTGT
65461  TTGGGGTCCGATACCTTGTCGTCACCATACACACCGGACAATCCATGACCTAGCGTGACA

65521  AGGAGGTGAGCAGCTGCGAGTGAGCATTACCACCTGAGCTCCATCTCATCTCAGATCAGT
65521  TCCTCCACTCGTCGACGCTCACTCGTAATGGTGGACTCGAGGTAGAGTAGAGTCTAGTCA

65581  GGCAGCATTAGGTTCTCATAGGAGTGTGAACCCTATTGTGAACTGTGCATTGCAGGAATC
65581  CCGTCGTAATCCAAGAGTATCCTCACACTTGGGATAACACTTGACACGTAACGTCCTTAG

65641  TAGGTGGTGCACTTTTTATGAGAATCTAATGCTTGATGATCTGAGGTAGAATGATTGCAC
65641  ATCCACCACGTGAAAAATACTCTTAGATTACGAACTACTAGACTCCATCTTACTAACGTG

65701  TCTGCAAGACTCCCCACGCTCATCCCTGTGGAAAAATTATCTTCCACAAAACTGGTCCCT
65701  AGACGTTCTGAGGGGTGCGAGTAGGGACACCTTTTTAATAGAAGGTGTTTTGACCAGGGA

65761  GGTGCCAAAAAGGTTGGGGACTGCTGCTTTTGGATATATACCCAGAAGTGGGATTGCTGA
65761  CCACGGTTTTTCCAACCCCTGACGACGAAAACCTATATATGGGTCTTCACCCTAACGACT

65821  AGCATATGGTAGTTTAATTTTAATTTTAATTTAATTTAGTTTAATTAATTTATTTTTAGC
65821  TCGTATACCATCAAATTAAAATTAAAATTAAATTAAATCAAATTAATTAAATAAAAATCG

65881  ACCTAGGCTGGAATGCAGCAGTGCAATCATAGCTCACTGCAGCTTCCAATTCCTGACCTC
65881  TGGATCCGACCTTACGTCGTCACGTTAGTATCGAGTGACGTCGAAGGTTAAGGACTGGAG

65941  AAGCCATCCTCCCAACTCAGCCTCCCAAATTTCTGGAACTACAGGAATGAGCCATCTCAC
65941  TTCGGTAGGAGGGTTGAGTCGGAGGGTTTAAAGACCTTGATGTCCTTACTCGGTAGAGTG

66001  TCAGCCCTATTTTTAGTGTTGGGAAGAATCTCAAAAAGTAATGACTGTACTAATTTACAC
66001  AGTCGGGATAAAAATCACAACCCTTCTTAGAGTTTTTCATTACTGACATGATTAAATGTG

66061  CGCTACCCTCAGTATACAAGGGTTCCATATCCTCACTAATACTTATCTTTTATCTTTTTG
66061  GCGATGGGAGTCATATGTTCCCAAGGTATAGGAGTGATTATGAATAGAAAATAGAAAAAC

66121  ATAGTAGCCATTCTAAGTATGTATAGTTTTAAAATTATGAAGTTTAAATTGAACCTTAGG
66121  TATCATCGGTAAGATTCATACATATCAAAATTTTAATACTTCAAATTTAACTTGGAATCC

66181  AAGTTGCTACTCTGTCACCAGGCTGGAGTGCAGCAGCGTGATCTTGGTTCACCACAATCT
66181  TTCAACGATGAGACAGTGGTCCGACCTCACGTCGTCGCACTAGAACCAAGTGGTGTTAGA

66241  CCGCCTCCTGGGTTTAAGCAATTCTCCTGGCTCAGCCTCCCAAGTAGCTGGGACTACAGG
66241  GGCGGAGGACCCAAATTCGTTAAGAGGACCGAGTCGGAGGGTTCATCGACCCTGATGTCC

66301  CACATGCTACCATGCCTGGCTCATTTTTATATTTTTGGTAGAGATGGGTTTCACCATGTT
66301  GTGTACGATGGTACGGACCGAGTAAAAATATAAAAACCATCTCTACCCAAAGTGGTACAA

66361  GTCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCCAACAGCCTTGGCCTCCCAACGT
66361  CAGGTCCGACCAGAGCTTGAGGACTGGAGTTCACTAGGTTGTCGGAACCGGAGGGTTGCA

66421  GTTGGGATTACAGACGTGAGCCATCACACCCAGCCTTGAAGTTGCCTTTTTTATCTTCAG
66421  CAACCCTAATGTCTGCACTCGGTAGTGTGGGTCGGAACTTCAACGGAAAAAATAGAAGTC

66481  TAGTGCTTCTTGGCTTAAATTCTTTTGTTTTATAAATAGTAATGGAAGCTTGCCAGTTTT
66481  ATCACGAAGAACCGAATTTAAGAAAACAAAATATTTATCATTACCTTCGAACGGTCAAAA
```

FIGURE 4 (cont'd)

```
66541  ATTTTAGAAAGTATTTGCAGGCCAGGTGCAGTGGCTCATGCCTGTAATCCCAGCACTTTG
66541  TAAAATCTTTCATAAACGTCCGGTCCACGTCACCGAGTACGGACATTAGGGTCGTGAAAC

66601  GGAGGCCGAGGCAGGCACATCACTTGAGCCCAGGAGTTCGAGACCAGCCTAGGCAACATG
66601  CCTCCGGCTCCGTCCGTGTAGTGAACTCGGGTCCTCAAGCTCTGGTCGGATCCGTTGTAC

66661  GCAAAACCTCGTCTCTACAAAAAATACAAAAAATTAACTTGGGCGTGGTGGTGCGCGCCT
66661  CGTTTTGGAGCAGAGATGTTTTTTATGTTTTTTAATTGAACCCGCACCACCACGCGCGGA

66721  GTAGTCCCAGCTACTCTGGAGGCTGAGGTGGGAGGATTGCTTGAACCCAGGGGGTAGAGG
66721  CATCAGGGTCGATGAGACCTCCGACTCCACCCTCCTAACGAACTTGGGTCCCCCATCTCC

66781  TTGCAGTGAGCCAAGATCTTGCCACTGAACTTCATTCAGCCTGGGTGACAGAGAAAAACA
66781  AACGTCACTCGGTTCTAGAACGGTGACTTGAAGTAAGTCGGACCCACTGTCTCTTTTTGT

66841  ACTCTGTCTCAAAAAAAAAAAATAAATAATAAAAAATAAAGTATTTGCATGATATATCTT
66841  TGAGACAGAGTTTTTTTTTTTATTTATTATTTTTTATTTCATAAACGTACTATATAGAA

66901  TTTCTGTCTTTTTGATTCAGATATTGTCAGCTGAGTCCAATCTAACAAACTTATATTTTA
66901  AAAGACAGAAAAACTAAGTCTATAACAGTCGACTCAGGTTAGATTGTTTGAATATAAAAT

66961  ACCGAAGCTTTAGTACTGTTATATTTAATAAACTAATAATATATTTGCACTTAAATACAA
66961  TGGCTTCGAAATCATGACAATATAAATTATTTGATTATTATATAAACGTGAATTTATGTT

67021  CATCTAATTTTATGCTTTTTAATTATACATGATTTGTATTTCTTTTTTTCCCATGTAAGT
67021  GTAGATTAAAATACGAAAAATTAATATGTACTAAACATAAAGAAAAAAAGGGTACATTCA

67081  TTTTAGGGTATACATCTTTAAAGGCCTTACTTTTCATTATCGTTCAGTTTAACTTATTTT
67081  AAAATCCCATATGTAGAAATTTCCGGAATGAAAAGTAATAGCAAGTCAAATTGAATAAAA

67141  CTAATTTCCATTTTTCTTCTGTGACCTATGGTTTAGTTAGTGGTGGATTTATTTTCCAAA
67141  GATTAAAGGTAAAAAGAAGACACTGGATACCAAATCAATCACCACCTAAATAAAAGGTTT

67201  ATATGGATGGTGTCCAGTCAAACTTTTGGTATTGATTTCTGGCTTATCTTGTGTTTGGAA
67201  TATACCTACCACAGGTCAGTTTGAAAACCATAACTAAAGACCGAATAGAACACAAACCTT

67261  AATATGTGAATTCTACAGCTGGTCTGTGTAAGAGTTCTCTGTATCAGTTAGGGAAGTTTG
67261  TTATACACTTAAGATGTCGACCAGACACATTCTCAAGAGACATAGTCAATCCCTTCAAAC

67321  TTCTTCCTGTTGTTCATATCTTGCATATCTTTACTAAGGTTTAGCCTGTTTGTTCTATAA
67321  AAGAAGGACAACAAGTATAGAACGTATAGAAATGATTCCAAATCGGACAAACAAGATATT

67381  ATTACTTAGAGAAGCTGTGTTAAATTTTCCCACTAAAATTATTGATTTTTGTCTTTCTCT
67381  TAATGAATCTCTTCGACACAATTTAAAAGGGTGATTTTAATAACTAAAAACAGAAAGAGA

67441  GTATTTATCAGTATGTATCTACTGAAGTCTTTTCAAATCCCTGATCAGTGGTTTGTTCAG
67441  CATAAATAGTCATACATAGATGACTTCAGAAAAGTTTAGGGACTAGTCACCAAACAAGTC

67501  GACAGCACTTTGATTGAGGGAGGTGACGTGGCTTAACATAGTATGGTCTTGAGCTGCCTG
67501  CTGTCGTGAAACTAACTCCCTCCACTGCACCGAATTGTATCATACCAGAACTCGACGGAC

67561  GTTTCTGTGATTTAGGGACTGAACCAGCTCTGGTCACATCTGGATTATCTACTATCCAGC
67561  CAAAGACACTAAATCCCTGACTTGGTCGAGACCAGTGTAGACCTAATAGATGATAGGTCG

67621  ATTCAAATGAACCAGATTACATGAAAGGATATGGGGGTACTCTTTGGGGGCCACGGTGCC
67621  TAAGTTTACTTGGTCTAATGTACTTTCCTATACCCCCATGAGAAACCCCCGGTGCCACGG

67681  ACAAGGCCTTTGGTTTCGATCTCCATTGCCTAGGGAATTAACCTGTGCTGTGTGATTCTA
67681  TGTTCCGGAAACCAAAGCTAGAGGTAACGGATCCCTTAATTGGACACGACACACTAAGAT

67741  GATGTTTTCTTCCTGCTCCCTTAAATAACACCTAAATGTGTTATTCAGCAGTGTGGCTCC
67741  CTACAAAAGAAGGACGAGGGAATTTATTGTGGATTTACACAATAAGTCGTCACACCGAGG

67801  TGCTAAGCAGAGGAGAACTCGGAATTCCTACCAAATTCTGATAATAAAATAACTTTCCTC
67801  ACGATTCGTCTCCTCTTGAGCCTTAAGGATGGTTTAAGACTATTATTTTATTGAAAGGAG
```

FIGURE 4 (cont'd)

```
67861  CACCTCAACCATTGTTTTTACCAGTCCTGTCAACTGTTAGCTTTCCTGGATTTAATGATC
67861  GTGGAGTTGGTAACAAAAATGGTCAGGACAGTTGACAATCGAAAGGACCTAAATTACTAG

67921  TCAAGTTGAGGTCAATGTCAACAGCTTTATGGGATGAATAAGATGAACGCCCATAGGAAA
67921  AGTTCAACTCCAGTTACAGTTGTCGAAATACCCTACTTATTCTACTTGCGGGTATCCTTT

67981  GTCAGTCTCTTGAACACGGGTTGTTTTCTTTTGGTTACCCTCTAAGATTTTTTCTCTTTG
67981  CAGTCAGAGAACTTGTGCCCAACAAAAGAAAACCAATGGGAGATTCTAAAAAAGAGAAAC

68041  ATCTCCAGTTATTTGAAGGATTGAAGGCATTTCGAGGAGTAGATAATAAAATTCGACTGT
68041  TAGAGGTCAATAAACTTCCTAACTTCCGTAAAGCTCCTCATCTATTATTTTAAGCTGACA

68101  TTCAGCCAAACCTCAACATGGATAGAATGTATCGCTCTGCTGTGAGGGCAACTCTGCCGG
68101  AAGTCGGTTTGGAGTTGTACCTATCTTACATAGCGAGACGACACTCCCGTTGAGACGGCC

68161  TATGTAAGTGTAGGAGTTTCTTTTGTGTTTCTTTTAATGTAACGGGTCACGGTGTTGACC
68161  ATACATTCACATCCTCAAAGAAAACACAAAGAAAATTACATTGCCCAGTGCCACAACTGG

68221  ACCAAATAAATAATTTCTGAATAGTTAGATAAATATTCAGTACAAACCATATGAACATTA
68221  TGGTTTATTTATTAAAGACTTATCAATCTATTTATAAGTCATGTTTGGTATACTTGTAAT

68281  ACATTATTTCTGATCTACTTTAGTTAGGTAAAAATACAAGAAAATGGCCGGGCACGGTGG
68281  TGTAATAAAGACTAGATGAAATCAATCCATTTTTATGTTCTTTTACCGGCCCGTGCCACC

68341  TTCACTCCTATAATCCCAGCACTCTGGGAGACTGAGGTGGGTGGATCACTTGAAACCAGG
68341  AAGTGAGGATATTAGGGTCGTGAGACCCTCTGACTCCACCCACCTAGTGAACTTTGGTCC

68401  AATTGGAGACAAGCCTGGCCAACATGGTGAAACACCATCTCTGCAAAAAATACAAAAATT
68401  TTAACCTCTGTTCGGACCGGTTGTACCACTTTGTGGTAGAGACGTTTTTTATGTTTTTAA

68461  AGCCGGGTTTGGTGGCACATGCCTGTGGTCCCATCGGTGTGGCTGAGGCATGAGAATTGC
68461  TCGGCCCAAACCACCGTGTACGGACACCAGGGTAGCCACACCGACTCCGTACTCTTAACG

68521  TTGAACCTGGCAGGTGGAGGCTGCAGTGAGCCAAGATAATCACATCACTGCCCTCCAGCC
68521  AACTTGGACCGTCCACCTCCGACGTCACTCGGTTCTATTAGTGTAGTGACGGGAGGTCGG

68581  TAAGCCGCAAAGTGAGACTCTGTAAAAAAAAAAAAACAAAAAACAAAAAACCCAAGAAAAT
68581  ATTCGGCGTTTCACTCTGAGACATTTTTTTTTTTTGTTTTTTGTTTTTTGGGTTCTTTTA

68641  AATATCATTTGGCAATTCCACAATTTTTTAAAATCAGTGTCTGTGGGGGATTCATATGCA
68641  TTATAGTAAACCGTTAAGGTGTTAAAAAATTTTAGTCACAGACACCCCCTAAGTATACGT

68701  AGTGAGAGATGAAAGAGTAAGGAAGAAAAATGTATCTCAACCTAAAATAATATTTGAAGA
68701  TCACTCTCTACTTTCTCATTCCTTCTTTTTACATAGAGTTGGATTTTATTATAAACTTCT

68761  CCACATGAAATTAATATTGAAAATGGAAAATATGGTGTGCAGCACAGGTATTTTGCGGAT
68761  GGTGTACTTTAATTATAACTTTTACCTTTTATACCACACGTCGTGTCCATAAAACGCCTA

68821  ATTTTGCCTAATTGCTTTAGATGCTAGCAAGATATAGTAACTTAATTATTTTGATATTTT
68821  TAAAACGGATTAACGAAATCTACGATCGTTCTATATCATTGAATTAATAAAACTATAAAA

68881  ATCTTAAGTAATGTTTTGAGATAAAAAGATATCAAAATGGTTTCTCATGGCCATTGAATC
68881  TAGAATTCATTACAAAACTCTATTTTTCTATAGTTTTACCAAAGAGTACCGGTAACTTAG

68941  GTGGCTGTCACAATGTTTCTTATGGAGCAGAAATCATTGTAAGATGGGTTAGATTTTTTT
68941  CACCGACAGTGTTACAAAGAATACCTCGTCTTTAGTAACATTCTACCCAATCTAAAAAAA

69001  TTTGAGAGAGAGTCTCACTCTGTCACCCAAGCTATAGTGCAGTGGCACAATCTTGGCTCA
69001  AAACTCTCTCTCAGAGTGAGACAGTGGGTTCGATATCACGTCACCGTGTTAGAACCGAGT

69061  CTACAACCTCCACCTCCGGGGTTCAAGTGATTCTTGTGCCTCAGCCTCCTTAGTATCTGG
69061  GATGTTGGAGGTGGAGGCCCCAAGTTCACTAAGAACACGGAGTCGGAGGAATCATAGACC

69121  GATTACAAGTGCACGCAGCCATGCCCAGCTTAATTTTTTTAAAAATTATTTTTAGTGGAG
69121  CTAATGTTCACGTGCGTCGGTACGGGTCGAATTAAAAAAATTTTTAATAAAAATCACCTC
```

FIGURE 4 (cont'd)

```
69181  ACATGATTTCACCATGTTAGCCAGGCTGGACTTGAACTTCTGGCCTCAAGTGATCTGCCT
69181  TGTACTAAAGTGGTACAATCGGTCCGACCTGAACTTGAAGACCGGAGTTCACTAGACGGA

69241  GCCTCAGCCTCCCAAAGTTCTGGGATTACAGGCATGAGCCACCGTGCCCAGCTTCAGGGA
69241  CGGAGTCGGAGGGTTTCAAGACCCTAATGTCCGTACTCGGTGGCACGGGTCGAAGTCCCT

69301  TTTAGCCCAACCTATGCATTCAGGAAGCCTGAGGACGGCTAAATCCAAGTGTTCTTCATT
69301  AAATCGGGTTGGATACGTAAGTCCTTCGGACTCCTGCCGATTTAGGTTCACAAGAAGTAA

69361  CCTCTGCCTATTTTTGTAACCCGCATTTCACCCATCTCCTTTGGAGGCCATAATATAGTA
69361  GGAGACGGATAAAAACATTGGGCGTAAAGTGGGTAGAGGAAACCTCCGGTATTATATCAT

69421  GCAGAGGGGAAATTAAGGTAAATTCTTTGGACTAGGGCATAGTTTCGCTCTTCTTCTCAT
69421  CGTCTCCCCTTTAATTCCATTTAAGAAACCTGATCCCGTATCAAAGCGAGAAGAAGAGTA

69481  TACAGCTAGATCACTGGGGAAAAGTCTTTAAAATGGGGCTGGAAGGAAAGGGAAGTAATA
69481  ATGTCGATCTAGTGACCCCTTTTCAGAAATTTTACCCCGACCTTCCTTTCCCTTCATTAT

69541  TTTGTTAGACATCAATTTTGCTACCAAACTGTGTTATGTAGTTTCACATTCCTTATCTGT
69541  AAACAATCTGTAGTTAAAACGATGGTTTGACACAATACATCAAAGTGTAAGGAATAGACA

69601  GTAATCACAGCAGTAAGGCGTTGCTAAGAAGATTATTTGTTGTACAGTTTATTAGTGCCT
69601  CATTAGTGTCGTCATTCCGCAACGATTCTTCTAATAAACAACATGTCAAATAATCACGGA

69661  TTGTACAGAGAAGAAAACTTGAGTCCAGAGAGATTAAATGACTTGTCTGAAGTCACAAAG
69661  AACATGTCTCTTCTTTTGAACTCAGGTCTCTCTAATTTACTGAACAGACTTCAGTGTTTC

69721  AGCAAAAGGACAGGGCAAGAAATTAGATCCAGGTGTTTTTATTCCTGGTCAAGTTCTCTT
69721  TCGTTTTCCTGTCCCGTTCTTTAATCTAGGTCCACAAAAATAAGGACCAGTTCAAGAGAA

69781  CTCACTGAATTCTGTCTCTAGACAGCAATACCAGGGGCTGAGGCAGGTGCTACAGGGTTC
69781  GAGTGACTTAAGACAGAGATCTGTCGTTATGGTCCCCGACTCCGTCCACGATGTCCCAAG

69841  ACCTATATGATTGGCAGTCCCAATTTCTATAACATTACTACATTATTATTGCTGTTACAG
69841  TGGATATACTAACCGTCAGGGTTAAAGATATTGTAATGATGTAATAATAACGACAATGTC

69901  GTCAATATTTAAATAGACAAATTCTGACTTCATGAGCTTTACTCAGATGCTTATGACTGA
69901  CAGTTATAAATTTATCTGTTTAAGACTGAAGTACTCGAAATGAGTCTACGAATACTGACT

69961  CTCATTTCTACCTTAGGTGACCCATCACATCTCTGAAAGTATGTTTCTCTACTTTGCAAG
69961  GAGTAAAGATGGAATCCACTGGGTAGTGTAGAGACTTTCATACAAAGAGATGAAACGTTC

70021  TGAGCAGAGACTAAATTATTTTATGAAGTATCACCCATCTAGTAGATGTAATTTTTTTTT
70021  ACTCGTCTCTGATTTAATAAAATACTTCATAGTGGGTAGATCATCTACATTAAAAAAAAA

70081  CTTTAGAAATAAAGTGCTCTTTGCTTTTGCCCATTTTCCTATTGGGTTATCTATCTCTTC
70081  GAAATCTTTATTTCACGAGAAACGAAAACGGGTAAAAGGATAACCCAATAGATAGAGAAG

70141  ATTGACTTTGGGGATATCCAAATATATATACCAGTTAAAATTCTTTGTTGATCATGCTAC
70141  TAACTGAAACCCCTATAGGTTTATATATATGGTCAATTTTAAGAAACAACTAGTACGATG

70201  CAATAACTTCACCCAGTTTGTTGCTTTACATCTTTTGAAGAACGGAAGTCTTTAAGTTTC
70201  GTTATTGAAGTGGGTCAAACAACGAAATGTAGAAAACTTCTTGCCTTCAGAAATTCAAAG

70261  AAATAGTCACATTTTATCCATTTTTTTCCTTCATGATTTGTGTTTTACATCTCCTACTTA
70261  TTTATCAGTGTAAAATAGGTAAAAAAAGGAAGTACTAAACACAAAATGTAGAGGATGAAT

70321  AGAATCCTTTCCTACCTTAAGGTCACAGAGATATTTTGTACTTTTTAAATAAAGTCTTAG
70321  TCTTAGGAAAGGATGGAATTCCAGTGTCTCTATAAAACATGAAAAATTTATTTCAGAATC

70381  TATTTTGTCACATTTAAGTATATCATCTAATTAGAATTGATTGGGGGGTATGGTAAGATG
70381  ATAAAACAGTGTAAATTCATATAGTAGATTAATCTTAACTAACCCCCCATACCATTCTAC

70441  TAGGGATCTAATCTTATCTTTCACATGTGGATAGCTCGTCATCCCATCACAATTTTTGGG
70441  ATCCCTAGATTAGAATAGAAAGTGTACACCTATCGAGCAGTAGGGTAGTGTTAAAAACCC
```

FIGURE 4 (cont'd)

```
70501  GGGAAAATCAGAAAAAATTGTTTGGTTTTATTATTACTTTAAAATTTTTTTTATTGGTAC
70501  CCCTTTTAGTCTTTTTTAACAAACCAAAATAATAATGAAATTTTAAAAAAAATAACCATG

70561  CATAAATTATACATATTTATGAGGTACACTGATGTTTTCATGCATGCATATAATGTGTAA
70561  GTATTTAATATGTATAAATACTCCATGTGACTACAAAAGTACGTACGTATATTACACATT

70621  TGATTAGGACAAATAAGAAAGAAATAAATAAAAATTAGGATAATTAGGATATCCATCACC
70621  ACTAATCCTGTTTATTCTTTCTTTATTTATTTTTAATCCTATTAATCCTATAGGTAGTGG

70681  TCAAACATTTCTTATTTTTTTGTGTTGGGATTATTAATTTTTTAACAGTTTAAATGCTTC
70681  AGTTTGTAAAGAATAAAAAAACACAACCCTAATAATTAAAAAATTGTCAAATTTACGAAG

70741  AGTGAGGATCATTTATTATTGCCACATTGTGAGAAAATTCTGTCATTTATTATGTTTTTC
70741  TCACTCCTAGTAAATAATAACGGTGTAACACTCTTTTAAGACAGTAAATAATACAAAAAG

70801  ATTCTTTCAGGTATTTGACAAAGAAGAGCTCTTAGAGTGTATTCAACAGCTTGTGAAATT
70801  TAAGAAAGTCCATAAACTGTTTCTTCTCGAGAATCTCACATAAGTTGTCGAACACTTTAA

70861  GGATCAAGAATGGGTCCCATATTCAACATCTGCTAGTCTGTATATTCGTCCTACATTCAT
70861  CCTAGTTCTTACCCAGGGTATAAGTTGTAGACGATCAGACATATAAGCAGGATGTAAGTA

70921  TGGAACTGAGGTGCAAACTGACTCTTTATTTTGGGGTACTTTGCTGGGCAAGTTATTATT
70921  ACCTTGACTCCACGTTTGACTGAGAAATAAAACCCCATGAAACGACCCGTTCAATAATAA

70981  GTTTATTGTTGTTTGAAACTTCTAGCATATAGTTTGCAGTAGCCCCCCAGTAGCATCATT
70981  CAAATAACAACAAACTTTGAAGATCGTATATCAAACGTCATCGGGGGGTCATCGTAGTAA

71041  AATGACTTCTTCCTTCTGATTGTCAAGATTAGAGGTAGCAAACATTTTCTGTGAAGCGCC
71041  TTACTGAAGAAGGAAGACTAACAGTTCTAATCTCCATCGTTTGTAAAAGACACTTCGCGG

71101  AATCAGTAAATATTTTTTGGCTTTGTGGCCATATAGTTTCTAATAAGAAATTAGATCCCT
71101  TTAGTCATTTATAAAAAACCGAAACACCGGTATATCAAAGATTATTCTTTAATCTAGGGA

71161  ACCTCTTACCATACCCCAAAATCAATTATAATTCGATTATATACTTAAATGTGACAAAAT
71161  TGGAGAATGGTATGGGGTTTTAGTTAATATTAAGCTAATATATGAATTTACACTGTTTTA

71221  ACTAAGACGTTTATTAAGAAATACAAAATATCTCTGTGACCTTAGGTAGGAAATGATTTT
71221  TGATTCTGCAAATAATTCTTTATGTTTTATAGAGACACTGGAATCCATCCTTTACTAAAA

71281  TTTTTTTTGAGGTCTCTCTCTGGTTTTTTTTTTTTGCCCACGTTGTCGCTTAGGCTGGGG
71281  AAAAAAAACTCCAGAGAGAGACCAAAAAAAAAAAAACGGGTGCAACAGCGAATCCGACCCC

71341  TGCAGTGATGCAATCACGGCTGACTGCAGCTTTGACCTCCTGGGCTCAAGTGATCCTCCT
71341  ACGTCACTACGTTAGTGCCGACTGACGTCGAAACTGGAGGACCCGAGTTCACTAGGAGGA

71401  GCCTCAGCTTCCCAAGTAGCTAGTACTACAGGCATGTGCTACCATGCCTAGCTAATTTTT
71401  CGGAGTCGAAGGGTTCATCGATCATGATGTCCGTACACGATGGTACGGATCGATTAAAAA

71461  AAATTTTTTGTAGAGATGATGTCTCACTCTGTTGTCCTGGCTGGTCTCAAACTCCTGGCC
71461  TTTAAAAAACATCTCTACTACAGAGTGAGACAACAGGACCGACCAGAGTTTGAGGACCGG

71521  TCTAGCAATCCTCCCACCTTGGCCTCCCAAAGTCCTGGGATTACAGGCATGAGCCACTTC
71521  AGATCGTTAGGAGGGTGGAACCGGAGGGTTTCAGGACCCTAATGTCCGTACTCGGTGAAG

71581  ACCTGACCGGAGAGGATTCTTAAACTCTGCCATTTTAGTACTAAAGCAGCCATAGACAGC
71581  TGGACTGGCCTCTCCTAAGAATTTGAGACGGTAAAATCATGATTTCGTCGGTATCTGTCG

71641  ACATAAATGAGTAAATGCAGTGCTGTTTGCCTTGCCCAATTAAGACTAGTCATTATTTTC
71641  TGTATTTACTCATTTACGTCACGACAAACGGAACGGGTTAATTCTGATCAGTAATAAAAG

71701  TAAGTTTACAAGTGCAGTGCCTCCCATTGTGGCAGTACTTCTCTGAGGTAGAAATGGTTT
71701  ATTCAAATGTTCACGTCACGGAGGGTAACACCGTCATGAAGAGACTCCATCTTTACCAAA

71761  TAAAATTTAGGGGAGGGAGAGCTGCTAATTTTCATTTTCATGAGGTCTTGGGATTACTTG
71761  ATTTTAAATCCCCTCCCTCTCGACGATTAAAAGTAAAAGTACTCCAGAACCCTAATGAAC

71821  GCATGGAGTGAGAATGGAGGTTCCCCTTCCCACATAGAAAAATGGTGAAAATAAAGTTCA
71821  CGTACCTCACTCTTACCTCCAAGGGGAAGGGTGTATCTTTTTACCACTTTTATTTCAAGT
```

FIGURE 4 (cont'd)

```
71881  CAGTGGGAGTTAAGGAAGGAAATTGAAGAGCTGGGATGGTAGGTTTAAATAACATGACCT
71881  GTCACCCTCAATTCCTTCCTTTAACTTCTCGACCCTACCATCCAAATTTATTGTACTGGA

71941  TATGATTTCCTGGTGGGTCACAGTCTCTTGGTTGGGTGCCCTTGTGAATAAAGGCATGGA
71941  ATACTAAAGGACCACCCAGTGTCAGAGAACCAACCCACGGGAACACTTATTTCCGTACCT

72001  TTGGGCATGCCTGGAGGGGCCAAGGCCTTTTGAGCGGACGAGTATGTTTATATGAAGACC
72001  AACCCGTACGGACCTCCCCGGTTCCGGAAAACTCGCCTGCTCATACAAATATACTTCTGG

72061  CTGTCACCTGTCTCTGGTATCTGCTAGGCAGCTGTGGCTTGAGAGTGTGATATTGAGGGA
72061  GACAGTGGACAGAGACCATAGACGATCCGTCGACACCGAACTCTCACACTATAACTCCCT

72121  TAATTGATGGAAGGAGATTGAATGTGGCTGGCAGGTAAAGAGATAGGGAACATATGGAGA
72121  ATTAACTACCTTCCTCTAACTTACACCGACCGTCCATTTCTCTATCCCTTGTATACCTCT

72181  TATTTAGAATAAGAGAGTGTGAGATAGAGCGCTTTTGCTATAGCCTAGCCATTCTGTTCT
72181  ATAAATCTTATTCTCTCACACTCTATCTCGCGAAAACGATATCGGATCGGTAAGACAAGA

72241  CTACTAAGTGTTGCGGGGTTACCAGGTTTTCTACTAGGCTGTATGAATGCAAAAGGTAGA
72241  GATGATTCACAACGCCCCAATGGTCCAAAAGATGATCCGACATACTTACGTTTTCCATCT

72301  GAGATATATAGGGCCATGGAATAAATACCGAAGTGTTTTCTTTGGTATCCTTGATACCTT
72301  CTCTATATATCCCGGTACCTTATTTATGGCTTCACAAAAGAAACCATAGGAACTATGGAA

72361  TCGAAAGCCGTGTAAGGATCTAGATAATGAATTTCGTGAATTGCAGTCTCTCTGACAAGT
72361  AGCTTTCGGCACATTCCTAGATCTATTACTTAAAGCACTTAACGTCAGAGAGACTGTTCA

72421  TTTAAAGAATAGGAACAATTTTAAAAAGTCATCTACTTCCATAAGTTTTTTGACATTTTA
72421  AAATTTCTTATCCTTGTTAAAATTTTTCAGTAGATGAAGGTATTCAAAAAACTGTAAAAT

72481  AGATAATTGAATTTTTTTTTTTCAGCTTTTAAGTTCCAGGGTACATGTGCAGGTAGGTGT
72481  TCTATTAACTTAAAAAAAAAAAGTCGAAAATTCAAGGTCCCATGTACACGTCCATCCACA

72541  GTTATATAGGTAAATGTAAGCCATGGTGGTTTGCTGTACAAATCAACCCATCACCTAGGT
72541  CAATATATCCATTTACATTCGGTACCACCAAACGACATGTTTAGTTGGGTAGTGGATCCA

72601  ATTAAGCCCAGCATCCATTAGCTATTCTTCCTGATGCTCTTCCTCCTCCCGCCCCTTCCG
72601  TAATTCGGGTCGTAGGTAATCGATAAGAAGGACTACGAGAAGGAGGAGGGCGGGGAAGGC

72661  ACAGACCCTAGTATGTGCTGCTCCCTGCATGTGCCCATGTGTTCTCATTGTTCAGCTCCC
72661  TGTCTGGGATCATACACGACGAGGGACGTACACGGGTACACAAGAGTAACAAGTCGAGGG

72721  ACTTATAAGAGAACATGCAGTGTTTGGTTTTCTTTTAATATTATTTTTAACTTTATGATA
72721  TGAATATTCTCTTGTACGTCACAAACCAAAAGAAAATTATAATAAAAATTGAAATACTAT

72781  GGAAGGAGTACTTTTTAAAGTTGGTTGAAACATTGTTACTGATCTCAAAGCACTACCTTC
72781  CCTTCCTCATGAAAAATTTCAACCAACTTTGTAACAATGACTAGAGTTTCGTGATGGAAG

72841  CAGTTATTCTGAGATACTTTTTCAGTCTAAGATCTGATACTCCTATTCCTGGTTTGTTGA
72841  GTCAATAAGACTCTATGAAAAAGTCAGATTCTAGACTATGAGGATAAGGACCAAACAACT

72901  GAGTTTTTGTCATGAATGGATGTTGGATTTTGTCAATTGCTTTTTTTGAATCTCTTGACA
72901  CTCAAAAACAGTACTTACCTACAACCTAAAACAGTTAACGAAAAAAACTTAGAGAACTGT

72961  TGATTATGTGGTTTTGTACTTTATTACATTATTAATGAATGAATTATATCAATTTATTTT
72961  ACTAATACACCAAAACATGAAATAATGTAATAATTACTTACTTAATATAGTTAAATAAAA

73021  TGGATGTTAAATCAACCTTATTTTCCTAAGCTAAATCCTACTTGGTCATTGTGAATAATC
73021  ACCTACAATTTAGTTGGAATAAAAGGATTCGATTTAGGATGAACCAGTAACACTTATTAG

73081  CATTTTTTTATCTTGCTATATTCAGTTTGCTAATATTTTAAGAATTTTTGTTTCACTGTT
73081  GTAAAAAAATAGAACGATATAAGTCAAACGATTATAAAATTCTTAAAAACAAAGTGACAA

73141  CATGAAATATATTGATCTAGTTTTCATTTTTCTTTTCTTTCTTTTTTTTTTTTTTTTTA
73141  GTACTTTATATAACTAGATCAAAAGTAAAAAGAAAAGAAAGAAAAAAAAAAAAAAAAAT
```

FIGURE 4 (cont'd)

```
73201  CATCTCCCTCCTCTTCCTCTTGATCTAGTTTTCTAATACAGCTGCACTCCAACTTACGAT
73201  GTAGAGGGAGGAGAAGGAGAACTAGATCAAAAGATTATGTCGACGTGAGGTTGAATGCTA

73261  GGTCCTACTTAGGAGTTTTTGACTGTATGATGTTTGAAAACAATATGAGTTTAGTAGAAA
73261  CCAGGATGAATCCTCAAAAACTGACATACTACAAACTTTTGTTATACTCAAATCATCTTT

73321  CTTACTTTGAGTACACATACCATTCTGTTTTTTTGCTTTTAGTACAGTATTCAATAAATT
73321  GAATGAAACTCATGTGTATGGTAAGACAAAAAAACGAAAATCATGTCATAAGTTATTTAA

73381  ATGTGAGATAACCAATACTTTATTATAAAATAATACACTTTTAGATGATTCTGCCCTACC
73381  TACACTCTATTGGTTATGAAATAATATTTTATTATGTGAAAATCTACTAAGACGGGATGG

73441  ATAGGCTAATGTAATTGTTCTGAGCACATTTAAGATAGGATAAGCTAAGATACGATGGAT
73441  TATCCGATTACATTAACAAGACTCGTGTAAATTCTATCCTATTCGATTCTATGCTACCTA

73501  GTTTGGTAGGTTAGGTATATTAAATGCATTTTCACCTTAAGGATATTTTCAACTTAGGAT
73501  CAAACCATCCAATCCATATAATTTACGTAAAAGTGGAATTCCTATAAAAGTTGAATCCTA

73561  GAGTTTATTGAGTTGTAACCCCATCACAAGTTGAGGAACATCTATGTATCTAATTCCATG
73561  CTCAAATAACTCAACATTGGGGTAGTGTTCAACTCCTTGTAGATACATAGATTAAGGTAC

73621  ATTCAGTAAATTCTTTAGACTGCCAAAATCCTGTAATCTAGAGTTAGACAAGATCTTAAA
73621  TAAGTCATTTAAGAAATCTGACGGTTTTAGGACATTAGATCTCAATCTGTTCTAGAATTT

73681  TCAGGTTTTGTGGTCCATCTTGTAGTCATGATCATTCTCACTATTAGGGAATTCTTAACT
73681  AGTCCAAAACACCAGGTAGAACATCAGTACTAGTAAGAGTGATAATCCCTTAAGAATTGA

73741  TTTTGCATTTGAATCACCGAGGTGTTATTTAAAATACACATATGGAATCCAAGCTCCAGA
73741  AAAACGTAAACTTAGTGGCTCCACAATAAATTTTATGTGTATACCTTAGGTTCGAGGTCT

73801  GACTCTGAGTAGATCTGGTGAGTGATTGTGATGCCTGCCCACATTCAAGAACCTTAGTAC
73801  CTGAGACTCATCTAGACCACTCACTAACACTACGGACGGGTGTAAGTTCTTGGAATCATG

73861  AAAAGCCCTGCTGAATGGTTCTCTTTCACTATACACTCTATCAACTTATGTACTTGATCA
73861  TTTTCGGGACGACTTACCAAGAGAAAGTGATATGTGAGATAGTTGAATACATGAACTAGT

73921  CACTTTGGACCTCAAAATCAATAATTTTTAAAAATCTACTTCCTAAGTTGTGAGGTACGA
73921  GTGAAACCTGGAGTTTTAGTTATTAAAAATTTTTAGATGAAGGATTCAACACTCCATGCT

73981  ATTGAAGGAAATAATGAATACTATTTGGGGAGTCAACATTATTTATTAACGAATACCATT
73981  TAACTTCCTTTATTACTTATGATAAACCCCTCAGTTGTAATAAATAATTGCTTATGGTAA

74041  TGGGGAGCCGACATTATTTATATTTGGTTGGTATATTTGTATCAACTATCATTTTCTAAT
74041  ACCCCTCGGCTGTAATAAATATAAACCAACCATATAAACATAGTTGATAGTAAAAGATTA

74101  TAATGGTATAGGCAGAATAAACCATCACTTGTGATTCAGAAAGGAACATGATACAAATAA
74101  ATTACCATATCCGTCTTATTTGGTAGTGAACACTAAGTCTTTCCTTGTACTATGTTTATT

74161  CATCCAATAATGTATATGGATTTAAATAGATCTGCCCCCTACAGATGTTACGCAATTCAT
74161  GTAGGTTATTACATATACCTAAATTTATCTAGACGGGGGATGTCTACAATGCGTTAAGTA

74221  GTTGGTAATGATCATTCATTAGGTTTTCGTGGTGGCTACATTTGATATGTGAAGTACTGG
74221  CAACCATTACTAGTAAGTAATCCAAAAGCACCACCGATGTAAACTATACACTTCATGACC

74281  CTGGAAGACTGGTGCCAGTAGCTTCAGGTTTGCTGGGATTCTTTCATCAGAGATGACTAG
74281  GACCTTCTGACCACGGTCATCGAAGTCCAAACGACCCTAAGAAAGTAGTCTCTACTGATC

74341  CTTTACTTATCTTTTTATGTTGCATGTTTCTATGATGGGCCTATCCTGTCATTGTCTTCA
74341  GAAATGAATAGAAAAATACAACGTACAAAGATACTACCCGGATAGGACAGTAACAGAAGT

74401  TAGTATAATTGCAAGCTCAACTCTGTCAGTTATGTGATTTTTTTTTTTTTTTGGTCCTG
74401  ATCATATTAACGTTCGAGTTGAGACAGTCAATACACTAAAAAAAAAAAAAAAACCAGGAC

74461  TCTTGCTTTTGGCTGACTTCTATATGTTTTTATTTCATGTCCTTTTCATAGTTAGATTAA
74461  AGAACGAAAACCGACTGAAGATATACAAAAATAAAGTACAGGAAAAGTATCAATCTAATT

74521  CTGGGTCTTATGTCCAAGAGATATTTTGATTTCTTTGAATGAATAAATCTATGTTAGGCT
74521  GACCCAGAATACAGGTTCTCTATAAAACTAAAGAAACTTACTTATTTAGATACAATCCGA
```

FIGURE 4 (cont'd)

```
74581  ACTGTATTAGGTTTTATAGACACCTAGTTGATCACTTGTGTCCAGTCTCTATTCCCTGGT
74581  TGACATAATCCAAAATATCTGTGGATCAACTAGTGAACACAGGTCAGAGATAAGGGACCA

74641  ATGGTAGTTGCTGTGTTTATAGAACATTGAGGATTCTACAATGGACACATTGGTGTTGGG
74641  TACCATCAACGACACAAATATCTTGTAACTCCTAAGATGTTACCTGTGTAACCACAACCC

74701  CCCCAGCCCAGACCCAAATCGAAATTTCTAGAGCCAGGATCAGAGACTCTGCATTCAACA
74701  GGGGTCGGGTCTGGGTTTAGCTTTAAAGATCTCGGTCCTAGTCTCTGAGACGTAAGTTGT

74761  AACTCTCCAGAGATTTCATAACCACTTTTAAAGGACTCACCATCCCAAACCTTGTTAAGG
74761  TTGAGAGGTCTCTAAAGTATTGGTGAAAATTTCCTGAGTGGTAGGGTTTGGAACAATTCC

74821  AATGAAATATATGTCTGCCAGTAATTTCAGTGCTTTAAGTCATTGAAAGCTGAACTTTTA
74821  TTACTTTATATACAGACGGTCATTAAAGTCACGAAATTCAGTAACTTTCGACTTGAAAAT

74881  ATGGCACCAAAGTGGAAATATTTATATAATTGAAAGCTAAACTTCATCCTAATGGCAGTG
74881  TACCGTGGTTTCACCTTTATAAATATATTAACTTTCGATTTGAAGTAGGATTACCGTCAC

74941  TTGAGTTCACTTCCATAGGCCCTGATATAACAGTCTAATATTATTGTTTTTCACATAAAA
74941  AACTCAAGTGAAGGTATCCGGGACTATATTGTCAGATTATAATAACAAAAAGTGTATTTT

75001  TAAGATTCAGTCACTGCCATATAATTAGGGAGTGGCTCTGCTATTGGATGGACTTAAAAT
75001  ATTCTAAGTCAGTGACGGTATATTAATCCCTCACCGAGACGATAACCTACCTGAATTTTA

75061  ATTTCATTATGTTTGCCAAGCTCTGTCTCTTTACATGTTTTATAATTAACTGATCCTTCA
75061  TAAAGTAATACAAACGGTTCGAGACAGAGAAATGTACAAAATATTAATTGACTAGGAAGT

75121  TTTTTATTTTGTTTTATTTTATTTTTGAGATAGAGTCTCGCTTGGTTGCCCAGGCTGGAG
75121  AAAAATAAAACAAAATAAAATAAAAACTCTATCTCAGAGCGAACCAACGGGTCCGACCTC

75181  TGCAATGGTGCAATCTTGGCTCATTGCAACCTCTGCTTCCCAGGTTCAAGTGATTCTCCT
75181  ACGTTACCACGTTAGAACCGAGTAACGTTGGAGACGAAGGGTCCAAGTTCACTAAGAGGA

75241  GCCTCAGCCTCCTGATTAGCGGGGACTACAGGTGCGTGCCACCATGCCCAGCTAATTTTT
75241  CGGAGTCGGAGGACTAATCGCCCCTGATGTCCACGCACGGTGGTACGGGTCGATTAAAAA

75301  GTACTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCCCAAACTCCTGACC
75301  CATGAAAATCATCTCTGCCCCAAAGTGGTACAACCGGTCCGACCAGGGTTTGAGGACTGG

75361  TCAGGTGTTCTGCCCACCTCGGCCTCCCAAAGTGCTGGATTACAGGCATGAGCCACCATG
75361  AGTCCACAAGACGGGTGGAGCCGGAGGGTTTCACGACCTAATGTCCGTACTCGGTGGTAC

75421  CCCGGCCTCCTTCATTTTTAATAATCTTCATAGATACATACTTGTTGTACTAATGTTTTT
75421  GGGCCGGAGGAAGTAAAAATTATTAGAAGTATCTATGTATGAACAACATGATTACAAAAA

75481  GTGCAGTTCCTCATGATAAATAATTAAATCAATTACTAAAATCCTTATAGATTTATGATG
75481  CACGTCAAGGAGTACTATTTATTAATTTAGTTAATGATTTTAGGAATATCTAAATACTAC

75541  ATAATAAATGGAAGTGTGTATTATTGAAACATGTTTAATTAACTGTTCATATATAACCAG
75541  TATTATTTACCTTCACACATAATAACTTTGTACAAATTAATTGACAAGTATATATTGGTC

75601  TGAGTAGGGATTGGCAATTAAAAATACCAGGCTACCAGTGCATGGAAATTTGAAGAAAAC
75601  ACTCATCCCTAACCGTTAATTTTTATGGTCCGATGGTCACGTACCTTTAAACTTCTTTTG

75661  AACAACCTGAAATTCAGAATACTCCAAAATCTGAAACTTTTTGAGCACTGACATGACTAA
75661  TTGTTGGACTTTAAGTCTTATGAGGTTTTAGACTTTGAAAAACTCGTGACTGTACTGATT

75721  AAATGCTCATTGGAGCATTTCAGATTTCATTTTTTCAGATTAGGGATGCTGAATTGGTAA
75721  TTTACGAGTAACCTCGTAAAGTCTAAAGTAAAAAAGTCTAATCCCTACGACTTAACCATT

75781  GCATAATGCAGATTAATCCAAAATCTGAAAAAAATCCTAAATCAAAAACAGTTTTGGTCA
75781  CGTATTACGTCTAATTAGGTTTTAGACTTTTTTTAGGATTTAGTTTTTGTCAAAACCAGT

75841  GAAGCTTTGTGGATAAGGGATAGTCCAAGCTTTTTGGACTGAAGCTTTTTGGATAAGGTA
75841  CTTCGAAACACCTATTCCCTATCAGGTTCGAAAAACCTGACTTCGAAAAACCTATTCCAT
```

FIGURE 4 (cont'd)

```
75901  TAGCCTATAATACATAAATACATAATACAGGATAATCACTGACCTGAATGTATAGCATCA
75901  ATCGGATATTATGTATTTATGTATTATGTCCTATTAGTGACTGGACTTACATATCGTAGT

75961  GTATGGTACACAAATATTTAAATCCATGTACTTTGTTGGATATTATGTTATATAAACTTC
75961  CATACCATGTGTTTATAAATTTAGGTACATGAAACAACCTATAATACAATATATTTGAAG

76021  TTGTAAAAAGTTTAGAGAGTAAAATCATTCACATTACCAAAAGATTTCTCATATTTATAA
76021  AACATTTTTCAAATCTCTCATTTTAGTAAGTGTAATGGTTTTCTAAAGAGTATAAATATT

76081  AAAGGATTATAGCAGCATTTTAATCAGTATCTCCTAAAGTATTTATTTATGGGTTGTTAA
76081  TTTCCTAATATCGTCGTAAAATTAGTCATAGAGGATTTCATAAATAAATACCCAACAATT

76141  ATCTAAAAGAGAAACTACAGTCTGGAGGGCTCTAAACTACTGCTGATTACATATCAGTTG
76141  TAGATTTTCTCTTTGATGTCAGACCTCCCGAGATTTGATGACGACTAATGTATAGTCAAC

76201  GGAATTAAGAGAAAATCTCAGATCATAATATTTAAATAAAAACATTTACATGAATGATTT
76201  CCTTAATTCTCTTTTAGAGTCTAGTATTATAAATTTATTTTTGTAAATGTACTTACTAAA

76261  CTAAGTTTTCACTTTACTAGACAAGAAACTAAAATTATATGACTGTCATGTTCAATATTG
76261  GATTCAAAAGTGAAATGATCTGTTCTTTGATTTTAATATACTGACAGTACAAGTTATAAC

76321  GATACTTAACAACACTACATCCAAGGAAGCCAGGTTTTGCTGATTTTAACCTCGTCCTGT
76321  CTATGAATTGTTGTGATGTAGGTTCCTTCGGTCCAAAACGACTAAAATTGGAGCAGGACA

76381  TGAGTAACTCCATGCAGGTACAGTGGTACAAACACATGTGGACACACAGACCCAGACCTT
76381  ACTCATTGAGGTACGTCCATGTCACCATGTTTGTGTACACCTGTGTGTCTGGGTCTGGAA

76441  TATTTCCTCGGGGATTCTGGTAAGTGTAGAGGTGTCTTGACCATAGTCAATAGCAGAAAT
76441  ATAAAGGAGCCCCTAAGACCATTCACATCTCCACAGAACTGGTATCAGTTATCGTCTTTA

76501  ATGCACAACCACAAGCTCTGAAATTCCATTTGATTGTTGCTCTTAGCCTTCCACTTTTTA
76501  TACGTGTTGGTGTTCGAGACTTTAAGGTAAACTAACAACGAGAATCGGAAGGTGAAAAAT

76561  GGCATATGTCAGTCCCATTTCATCCCACCCACTCTCTCCATTTGGGAGGTGGTAGAGTGA
76561  CCGTATACAGTCAGGGTAAAGTAGGGTGGGTGAGAGAGGTAAACCCTCCACCATCTCACT

76621  GACAATACTGAACTATACCAGGGCTGAGAATAAGTGTCTTTGAAAAGGGTCGAATCTGTG
76621  CTGTTATGACTTGATATGGTCCCGACTCTTATTCACAGAAACTTTTCCCAGCTTAGACAC

76681  ATGCTAGAGACAGCCCTGGAAACTGTCAGGAGTTAGCCCAACCAGGACTGTGACCAGATC
76681  TACGATCTCTGTCGGGACCTTTGACAGTCCTCAATCGGGTTGGTCCTGACACTGGTCTAG

76741  CCAGGCTGTCCACCACAGATTTGCAGGCTACCTGCAGATGTGACTTTTTCAAATGTACTC
76741  GGTCCGACAGGTGGTGTCTAAACGTCCGATGGACGTCTACACTGAAAAAGTTTACATGAG

76801  AAAGACAAGCTGACTTTGGCTAATGCAGAAATTAATTAATTCCTCTTTTTCAGTAAACTT
76801  TTTCTGTTCGACTGAAACCGATTACGTCTTTAATTAATTAAGGAGAAAAAGTCATTTGAA

76861  TATCCTCCTACACTTAGATTGTCTGTAAGTACATCTCTGGAGTTGAGCTCAACTAGAAAG
76861  ATAGGAGGATGTGAATCTAACAGACATTCATGTAGAGACCTCAACTCGAGTTGATCTTTC

76921  AAGCTGTCTGCATTTAGTTAATACCAGCATTGGATGCACACAATATAATTTGAATAATGT
76921  TTCGACAGACGTAAATCAATTATGGTCGTAACCTACGTGTGTTATATTAAACTTATTACA

76981  AGAGCAACATTTTAATCATCTCTCATAAGCAGTTTCCCCATTTGCAACTTCCTGGGAAAT
76981  TCTCGTTGTAAAATTAGTAGAGAGTATTCGTCAAAGGGGTAAACGTTGAAGGACCCTTTA

77041  CTTACATTATGAAATGTTCATTAGAACTCCCATTTTAAAACCCTGCTGCCATTTTGTGAG
77041  GAATGTAATACTTTACAAGTAATCTTGAGGGTAAAATTTTGGGACGACGGTAAAACACTC

77101  TGGGCAGAAGAAAGGCTGGAGATGAGATTTGTCTGTCTCAAATGATTCACTCCTTCATTT
77101  ACCCGTCTTCTTTCCGACCTCTACTCTAAACAGACAGAGTTTACTAAGTGAGGAAGTAAA

77161  CTGGTTCACTCCACAGGGAGCTTAAGGACATGTGCGTTTCATAGACATGCCTTTCTTTTC
77161  GACCAAGTGAGGTGTCCCTCGAATTCCTGTACACGCAAAGTATCTGTACGGAAAGAAAAG

77221  CTTTTATCACCTTGTTAAATCTTAAGATTGAACATACAATTCAGCCTGTAATCATATCAA
77221  GAAAATAGTGGAACAATTTAGAATTCTAACTTGTATGTTAAGTCGGACATTAGTATAGTT
```

FIGURE 4 (cont'd)

```
77281  TTCCCATGAAAAATATTTTCCTTATTCACTGCCTCTATTTGTGCTTATACATCTGATCAT
77281  AAGGGTACTTTTTATAAAAGGAATAAGTGACGGAGATAAACACGAATATGTAGACTAGTA

77341  CTTTTTCCTTTACTTTACAATTAGTTTCTATGAACAGAAGAATCTGAAAGATTTAAAGAC
77341  GAAAAAGGAAATGAAATGTTAATCAAAGATACTTGTCTTCTTAGACTTTCTAAATTTCTG

77401  TAGTTAGCACAGCAGCAATTTATAAATCATTAATTTGGGTACTGTCTATACATTTAATGT
77401  ATCAATCGTGTCGTCGTTAAATATTTAGTAATTAAACCCATGACAGATATGTAAATTACA

77461  GTTCTTTTTCATAGTCCTATGTGAAATACATATACATATATATATATTCCACTGTTTATA
77461  CAAGAAAAAGTATCAGGATACACTTTATGTATATGTATATATATATAAGGTGACAAATAT

77521  AACAATATTCAGTTTTATTTGCATAACTCAATATATTTTATATTAACAGCTCCGTTTTTA
77521  TTGTTATAAGTCAAAATAAACGTATTGAGTTATATAAAATATAATTGTCGAGGCAAAAAT

77581  GTGTCTTAAGACATCGAGTCTGTTTTTAGAGATACAGTCTATTTGCCTTGATTTTAGCCT
77581  CACAGAATTCTGTAGCTCAGACAAAAATCTCTATGTCAGATAAACGGAACTAAAATCGGA

77641  GTTATAATCAGCTTATTATTATCTGGCCACTCTTACCAGTGAGTATCATCTTGAGGCTGT
77641  CAATATTAGTCGAATAATAATAGACCGGTGAGAATGGTCACTCATAGTAGAACTCCGACA

77701  TCTTGCTCATTTCAACTTGTCACAGGGGATGTTTTTTACTCTTTGGAGTTTAGGTACACC
77701  AGAACGAGTAAAGTTGAACAGTGTCCCCTACAAAAAATGAGAAACCTCAAATCCATGTGG

77761  AGGTTATTAATTATGCTCCTGATTCATGGTACTTCCATCATGGGGCATATAAGGTCCCTG
77761  TCCAATAATTAATACGAGGACTAAGTACCATGAAGGTAGTACCCCGTATATTCCAGGGAC

77821  TCTTTGTTTTATTCCCTTCTATCTTCATTCACCACTCTCATTTCCTCTCTCTCCACTAGA
77821  AGAAACAAAATAAGGGAAGATAGAAGTAAGTGGTGAGAGTAAAGGAGAGAGAGGTGATCT

77881  AATAATTCTCAGGAATCAAATTAATATTATTCTGGCTTATATGTGTTTTTAAAAATCTCA
77881  TTATTAAGAGTCCTTAGTTTAATTATAATAAGACCGAATATACACAAAAATTTTTAGAGT

77941  TCGCTTTTTATACATGTAATTTTAATTTATGTAAGAGACTTTGTCCTACAGATTTCATTG
77941  AGCGAAAAATATGTACATTAAAATTAAATACATTCTCTGAAACAGGATGTCTAAAGTAAC

78001  TGTTTCTTCTCTTTTTCGCCCAGCACTGTGTGTTCCATTTCTATCCAGTCTACTATGTGT
78001  ACAAAGAAGAGAAAAAGCGGGTCGTGACACACAAGGTAAAGATAGGTCAGATGATACACA

78061  ACATCAAGTCCATCATTTCTAACTGTTGCATAGTTCATGATGCGAGGCATCTTCCACATT
78061  TGTAGTTCAGGTAGTAAAGATTGACAACGTATCAAGTACTACGCTCCGTAGAAGGTGTAA

78121  GTATCCCAATTCCCTAACTGGAATCCCAGTTCCCTGCTACCTCACACAATGCTCCAGTGA
78121  CATAGGGTTAAGGGATTGACCTTAGGGTCAAGGGACGATGGAGTGTGTTACGAGGTCACT

78181  ACAGCTGGAAAATCATACAACCCAAGGAATTATTGAAAGTGATTATATCCCAGCAAAAGG
78181  TGTCGACCTTTTAGTATGTTGGGTTCCTTAATAACTTTCACTAATATAGGGTCGTTTTCC

78241  CCCAGGCCCATCTTACACTTCCACCCCAAGAAGGTACTGCTACTTTTCGGGAATGGCCTG
78241  GGGTCCGGGTAGAATGTGAAGGTGGGGTTCTTCCATGACGATGAAAAGCCCTTACCGGAC

78301  GTCTCCTTGGCGAAATTGGCATTTGATCTGTTCTCATCACTGTCGTCTGGCATCAGAAGT
78301  CAGAGGAACCGCTTTAACCGTAAACTAGACAAGAGTAGTGACAGCAGACCGTAGTCTTCA

78361  GACCCCACTTGTCAGGCACTGTACTCAGTGCTTTCCATGCACTAACTCTGATAACAATCT
78361  CTGGGGTGAACAGTCCGTGACATGAGTCACGAAAGGTACGTGATTGAGACTATTGTTAGA

78421  TTTTTTTTTCTTTTTTTGAGACAGAGTCTTGCTCTGTTGCCCAGGCTGGAGTGCAGTTGT
78421  AAAAAAAAAGAAAAAAACTCTGTCTCAGAACGAGACAACGGGTCCGACCTCACGTCAACA

78481  GTGATCTCAGCTCACTGCAACCTCCACCTCCCAAATTCAAGCGATTCTCCTGCCTCAGCC
78481  CACTAGAGTCGAGTGACGTTGGAGGTGGAGGGTTTAAGTTCGCTAAGAGGACGGAGTCGG

78541  TCCCCAGTAGCTGGGACTACAGGCGTGCACCACTACACCTGGCTAATTTTTGTATTTTTA
78541  AGGGGTCATCGACCCTGATGTCCGCACGTGGTGATGTGGACCGATTAAAAACATAAAAAT
```

FIGURE 4 (cont'd)

```
78601  GTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAAATCCTGACCTCAGGTGAT
78601  CATCTCTACCCCAAAGTGGTACAACCGGTCCGACCAGAACTTTAGGACTGGAGTCCACTA

78661  CCACTTGACCTCCCAAAGTGCTGGAATTGCAGGCAAGAGCCACTGAGCTTGGCCTGATAA
78661  GGTGAACTGGAGGGTTTCACGACCTTAACGTCCGTTCTCGGTGACTCGAACCGGACTATT

78721  CAATCTTGTGAGCAGGTGCTTATGCAGTATTTTAGGGATGTGAAAGTAAGGTCCAAAGCG
78721  GTTAGAACACTCGTCCACGAATACGTCATAAAATCCCTACACTTTCATTCCAGGTTTCGC

78781  GTGAAATAAGAGCCTGTATTCTCTTAACCTCCATCCCGCAACATCTTTTATGTGTCAGTC
78781  CACTTTATTCTCGGACATAAGAGAATTGGAGGTAGGGCGTTGTAGAAAATACACAGTCAG

78841  CCTGCACTAAGTCTAGGGATACTGTGCTGAGCAAAAACTATCTTCTTGCTTTCACAGAGC
78841  GGACGTGATTCAGATCCCTATGACACGACTCGTTTTTGATAGAAGAACGAAAGTGTCTCG

78901  TCAGTGTCTGGCAGGGGGTCAGACTTTCATAATATAAACCCACTAAATGAGTGCATATTG
78901  AGTCACAGACCGTCCCCCAGTCTGAAAGTATTATATTTGGGTGATTTACTCACGTATAAC

78961  ATGGCGGTCACTGCTGTGAAGAAGTGGGAATGGTTAGCATGGAAACATAGCCCCAATGGG
78961  TACCGCCAGTGACGACACTTCTTCACCCTTACCAATCGTACCTTTGTATCGGGGTTACCC

79021  GGCGACTAAGTTCTTTCCCATCTCATTCTACTATGTCTACTTTCATCTCTACTTTTTTCT
79021  CCGCTGATTCAAGAAAGGGTAGAGTAAGATGATACAGATGAAAGTAGAGATGAAAAAAGA

79081  TTTCTGAGCTATCAGCTTTCTTGCTCTCAACTATATTTGATGTATTTAACCAGATTTTAT
79081  AAAGACTCGATAGTCGAAAGAACGAGAGTTGATATAAACTACATAAATTGGTCTAAAATA

79141  ATATTGTATTATGTATTTTATAACATGTATATATATCATATAACATACACATATACATAC
79141  TATAACATAATACATAAAATATTGTACATATATATAGTATATTGTATGTGTATATGTATG

79201  ATACACACAGCCAATTAAGTCTCAAGTTCAATTACAATGCTTCTAAGTTTAGGACTATTC
79201  TATGTGTGTCGGTTAATTCAGAGTTCAAGTTAATGTTACGAAGATTCAAATCCTGATAAG

79261  ATTTTCATTTTGGCAGAGAAATTATCTGGTAAAAAATTCAAGTAGAACCAAAGGTATACA
79261  TAAAAGTAAAACCGTCTCTTTAATAGACCATTTTTTAAGTTCATCTTGGTTTCCATATGT

79321  ATAGATGGGAAGTTTCCTTTTCACTCTCACACTCTCATTTCCATCCAGGGTTCTACAATT
79321  TATCTACCCTTCAAAGGAAAAGTGAGAGTGTGAGAGTAAAGGTAGGTCCCAAGATGTTAA

79381  CATCTCCTTCCAAGCAATTCCTATTGCCAGCTTTTCTTATACCCTTTCAGAAATACACAC
79381  GTAGAGGAAGGTTCGTTAAGGATAACGGTCGAAAAGAATATGGGAAAGTCTTTATGTGTG

79441  ACACACACACACACATATATACACACACACACACACCCTTACATGTGTTTGTACATTA
79441  TGTGTGTGTGTGTGTATATATGTGTGTGTGTGTGTGTGGGAATGTACACAAACATGTAAT

79501  TCTCTCTTTTTACAAAATTGGCAGTGATTCTTTGTCCTGTTTCTGCCACTACTCAAAAAT
79501  AGAGAGAAAAAATGTTTTAACCGTCACTAAGAAACAGGACAAAGACGGTGATGAGTTTTTA

79561  TTTAAAACAAATTACACACCTATAAACATACAATTTAAAGAAAAGAGTGAAAGTGAAAAT
79561  AAATTTTGTTTAATGTGTGGATATTTGTATGTTAAATTTCTTTTCTCACTTTCACTTTTA

79621  CTAAGGACCCAGCACATGGTTAAGATGAGAGCATGATGGGTTCAATGGAAAGCCTTTGTG
79621  GATTCCTGGGTCGTGTACCAATTCTACTCTCGTACTACCCAAGTTACCTTTCGGAAACAC

79681  TACCTTTCTCCTTTGCCTCCGTTCCCTCTCCAATAGAGTCAACCACGTTGGAGGTTATAG
79681  ATGGAAAGAGGAAACGGAGGCAAGGGAGAGGTTATCTCAGTTGGTGCAACCTCCAATATC

79741  CACTTCCTTCTTTATGTCCCCCCATCTCCCATCTGGTATGGTATTCTCTTTTACTGCAGA
79741  GTGAAGGAAGAAATACAGGGGGGTAGAGGGTAGACCATACCATAAGAGAAAATGACGTCT

79801  TGGACATTTTGGCAGACATATGTAGTTGCTTCTCTTTGGTATTTACCTAGGAGTTGAAGT
79801  ACCTGTAAAACCGTCTGTATACATCAACGAAGAGAAACCATAAATGGATCCTCAACTTCA

79861  GCTGGGTCACAGGGTGTATGTATATTTCACTTTAGTAGATACTGCTGTTGTAGGACTTTC
79861  CGACCCAGTGTCCCACATACATATAAAGTGAAATCATCTATGACGACAACATCCTGAAAG

79921  TCCTTAGTTTAGCTAAAGATGGGGTCCTTGTCCCAAGGCCATGAAAAATTAGGCTCGCAG
79921  AGGAATCAAATCGATTTCTACCCCAGGAACAGGGTTCCGGTACTTTTTAATCCGAGCGTC
```

FIGURE 4 (cont'd)

```
79981  ACAATTTGAAAGGTGAGAATAATGGAACTTATTGGGAAAAAAGGGAAACAGGGACTCCCC
79981  TGTTAAACTTTCCACTCTTATTACCTTGAATAACCCTTTTTTCCCTTTGTCCCTGAGGGG

80041  ACAAAGCCAGGGTCCTGCTAGCATGCTTCCTGCCTTGCAGATTGAATCCCAGGTACCACC
80041  TGTTTCGGTCCCAGGACGATCGTACGAAGGACGGAACGTCTAACTTAGGGTCCATGGTGG

80101  CAGGAAGAGGAGGGGCCAGGCTCCTCCCCACTGTGAATGGTGTGACCGTCTGTGGCTCCC
80101  GTCCTTCTCCTCCCCGGTCCGAGGAGGGGTGACACTTACCACACTGGCAGACACCGAGGG

80161  CAGTGTGCACTCCTCCCAGTGTGCAGGCCGGTTGGAGTTTCTCTGGGGTCTTCTTCCCAC
80161  GTCACACGTGAGGAGGGTCACACGTCCGGCCAACCTCAAAGAGACCCCAGAAGAAGGGTG

80221  TTGGCTGTCTCACTGCCAAACTCCCAAAGTGCTTATACCAGTTTACTCTCCCATCAGTGG
80221  AACCGACAGAGTGACGGTTTGAGGGTTTCACGAATATGGTCAAATGAGAGGGTAGTCACC

80281  TGTATGTATTTGCTTCATTTCTTTGCCAATACTGATACTGTCTTATTAACTTTCCTCATT
80281  ACATACATAAACGAAGTAAAGAAACGGTTATGACTATGACAGAATAATTGAAAGGAGTAA

80341  CTAATGTATACAGTCTTGTGTTATTTGTGGTTTAATTTCCCTGAGAACCAATTATGTTGA
80341  GATTACATATGTCAGAACACAATAAACACCAAATTAAAGGGACTCTTGGTTAATACAACT

80401  GTACCTTTTCATAGGTTTACCAGCTACAGAAGTCCTCTTGTGAAGTGCCTTTGCAAGTCT
80401  CATGGAAAAGTATCCAAATGGTCGATGTCTTCAGGAGAACACTTCACGGAAACGTTCAGA

80461  TTTGTCTTTTTAATTTTTTGCATCTTTGTATATTCTAGGTATAAATCCTTTGATACATAT
80461  AAACAGAAAAATTAAAAAACGTAGAAACATATAAGATCCATATTTAGGAAACTATGTATA

80521  ATCACAAATAACTTCCTAGTCTATGGCAAATCCTATGTTGGATAAACGTACTGCAAATAG
80521  TAGTGTTTATTGAAGGATCAGATACCGTTTAGGATACAACCTATTTGCATGACGTTTATC

80581  TTTCCCAGTCTGTGGCTAGCTTCTGTGTCTGTGTCTTGCTTTGTCACTTTCTTACTGATA
80581  AAAGGGTCAGACACCGATCGAAGACACAGACACAGAACGAAACAGTGAAAGAATGACTAT

80641  TCTTTTGATGAACAGAAGTACTTAGTTTTAAAGTAGTTAAAGTTTTCTTTATGTTTTATA
80641  AGAAAACTACTTGTCTTCATGAATCAAAATTTCATCAATTTCAAAAGAAATACAAAATAT

80701  CTTTTTATGCATTTTAAAGAAATTTTTTGTTACTCCAAGGTCATACAGGTTAAATAAATT
80701  GAAAAATACGTAAAATTTCTTTAAAAAACAATGAGGTTCCAGTATGTCCAATTTATTTAA

80761  CTTACTCCAAGAACATTCTTTTAAAGTTTTGCTTTTTTTCATTTAAATGAAATCCATATG
80761  GAATGAGGTTCTTGTAAGAAAATTTCAAAACGAAAAAAAGTAAATTTACTTTAGGTATAC

80821  AAAGTGAACTTTGTTATCAAGTAACATATACAATTTCTTTTATTTTTCTGCACGGAGACC
80821  TTTCACTTGAAACAATAGTTCATTGTATATGTTAAAGAAAATAAAAAGACGTGCCTCTGG

80881  AAGTTGTCCCATCACCATTTATTGAGTGGTCTATCCTTTCCTCACTGATTACAGTGCCAG
80881  TTCAACAGGGTAGTGGTAAATAACTCACCAGATAGGAAAGGAGTGACTAATGTCACGGTC

80941  TTTTGATATGTCATGTTTCCATAGATGAAGAATATTTTGGGGCTCTCTATCTGTACCTTC
80941  AAAACTATACAGTACAAAGGTATCTACTTCTTATAAAACCCCGAGAGATAGACATGGAAG

81001  CTTTTTAAAAAATATAAATAGTGGTGATACTATGAACAATTTTGTTTGTATTATTGTTTT
81001  GAAAAATTTTTTATATTTATCACCACTATGATACTTGTTAAAACAAACATAATAACAAAA

81061  TTGCACTTAAAATTTGTCTTGGAGATCTTTTCATTTTGGCACATACGGATACACCCCTAT
81061  AACGTGAATTTTAAACAGAACCTCTAGAAAAGTAAAACCGTGTATGCCTATGTGGGGATA

81121  TTTTTTTTCCATTTTCCTAGATGTGGGGTATCCATTGGACATATGTACCGTACTGAATTT
81121  AAAAAAAAGGTAAAAGGATCTACACCCCATAGGTAACCTGTATACATGGCATGACTTAAA

81181  AATATAGAATGTCATAAATTACAAGATGTAACATTATTTTATACACTGAGATGGAGAATC
81181  TTATATCTTACAGTATTTAATGTTCTACATTGTAATAAAATATGTGACTCTACCTCTTAG

81241  AAACTGCAACGCAGTACTTCCCTGATCATCCTGAGTGATTCATGAGTTAGTTATACTAGC
81241  TTTGACGTTGCGTCATGAAGGGACTAGTAGGACTCACTAAGTACTCAATCAATATGATCG
```

FIGURE 4 (cont'd)

```
81301  CTCTTATTTCTTTGAGGTGTACGTTTTCTGAGGGAATCTGCCATTTCTCCTGCCATGAGT
81301  GAGAATAAAGAAACTCCACATGCAAAAGACTCCCTTAGACGGTAAAGAGGACGGTACTCA

81361  TGCATTGCTTGGCTCGTGATAGGCATTTTTATTTTTGCATGTAATTTTGTAACAAATGCT
81361  ACGTAACGAACCGAGCACTATCCGTAAAAATAAAAACGTACATTAAAACATTGTTTACGA

81421  AACACTGCTTTATCTTCTTTTTGTTTTTGTTTATTGAGACGGAGTCTTGCTCTGTTTCCC
81421  TTGTGACGAAATAGAAGAAAAACAAAAACAAATAACTCTGCCTCAGAACGAGACAAAGGG

81481  AGGCTGGAGTGCAGTGGTGCAATCTCGGCTCACTGCAACCTCTGTCTCCCAGGTTCAAGC
81481  TCCGACCTCACGTCACCACGTTAGAGCCGAGTGACGTTGGAGACAGAGGGTCCAAGTTCG

81541  GATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGTGCCACCATGCTCGG
81541  CTAAGAGGACGGAGTCGGAGGACTCATCGACCCTAATGTCCGTACACGGTGGTACGAGCC

81601  CTACTTTTTGTATTTTTTGTAGAGATGGGGTTTCACCATGTTGGTCAGGCTTGTCTTGAA
81601  GATGAAAAACATAAAAAACATCTCTACCCCAAAGTGGTACAACCAGTCCGAACAGAACTT

81661  CTCCTGACCTCAGGTGATCTGCCCCACTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTG
81661  GAGGACTGGAGTCCACTAGACGGGGTGAGTCGGAGGGTTTCACGACCCTAATGTCCACAC

81721  AGCCACCAGCCAGCTTTATCTTTTTAAAGGTTTCTTCCTATTGTGGGTTTCATTAGAGGG
81721  TCGGTGGTCGGTCGAAATAGAAAAATTTCCAAAGAAGGATAACACCCAAAGTAATCTCCC

81781  CTTGGTTGTTGCTTTGCAAGAAAACATTGACTTGCTTTTGTTTAATAGTCTTATAAATAT
81781  GAACCAACAACGAAACGTTCTTTTGTAACTGAACGAAAACAAATTATCAGAATATTTATA

81841  TTGTTTCACTAACATCAAATACATGCCATGCTCCTGCTTTTGAGCTTTTCTGTGCAGTAA
81841  AACAAAGTGATTGTAGTTTATGTACGGTACGAGGACGAAAACTCGAAAAGACACGTCATT

81901  CTTTTCTTCTCAGTGCTAAATCATAAGTTTTTAAAGACATTTAACATGGCACACTCAACA
81901  GAAAAGAAGAGTCACGATTTAGTATTCAAAAATTTCTGTAAATTGTACCGTGTGAGTTGT

81961  TGTGTCACCAATGCCCATAGCCACTGGAAGTATCATCAGTGTCAACAGCTATGACGTTTG
81961  ACACAGTGGTTACGGGTATCGGTGACCTTCATAGTAGTCACAGTTGTCGATACTGCAAAC

82021  GCAATGACAGCTTTGCCATAACTGCTGCTTAGCCAATAACAATTGTGGAATGCTGAGTCC
82021  CGTTACTGTCGAAACGGTATTGACGACGAATCGGTTATTGTTAACACCTTACGACTCAGG

82081  AGATCTCAAACATGTTAAAGGTAAAAAAAAAAAAAAAAAAAAAGATATTTTCGAGTCTGTA
82081  TCTAGAGTTTGTACAATTTCCATTTTTTTTTTTTTTTTTTTCTATAAAAGCTCAGACAT

82141  AAACATGGTAGCGTATCCGGAATTGGTGGGTTCTTGGTCTCACTGACTTCAAGAATGAAG
82141  TTTGTACCATCGCATAGGCCTTAACCACCCAAGAACCAGAGTGACTGAAGTTCTTACTTC

82201  CCATGGACCCTCGCGGTAAGTGTTACAGTTCTTAAAGGTGTCGTGTCTGGAGTTTGTTCC
82201  GGTACCTGGGAGCGCCATTCACAATGTCAAGAATTTCCACAGCACAGACCTCAAACAAGG

82261  TTCTGATGTTCAGATGTGTTCGGAGTTTCTTCCTTCTGGTAGGTTCGTGGTCTCGTTGGC
82261  AAGACTACAAGTCTACACAAGCCTCAAAGAAGGAAGACCATCCAAGCACCAGAGCAACCG

82321  TCAGGAGTGAAGCTGCGGACCTTTGCGGTGAGTGTTGCAGCTCTTAAGGCGGCATGTCTG
82321  AGTCCTCACTTCGACGCCTGGAAACGCCACTCACAACGTCGAGAATTCCGCCGTACAGAC

82381  GAGTTGTTCGTTCCTTGCGGTGGGTTCGTGGTCTCGCTGGCTTCAGGAGTGAAGCTGCAG
82381  CTCAACAAGCAAGGAACGCCACCCAAGCACCAGAGCGACCGAAGTCCTCACTTCGACGTC

82441  ACCTTCCCAGTGAGTGTTACAGCTCATAAAGGCAGTGTGGACCCAAAGAGTGAGCAGCAG
82441  TGGAAGGGTCACTCACAATGTCGAGTATTTCCGTCACACCTGGGTTTCTCACTCGTCGTC

82501  CAAGATTTACTGCAAAGAGCGAGAGAACAAAGCTTCCACAGTATGGAAGGGGACCCCAGC
82501  GTTCTAAATGACGTTTCTCGCTCTCTTGTTTCGAAGGTGTCATACCTTCCCCTGGGGTCG

82561  AGGTTGCCACTGCTGGCTCAGGCAGCCTGCTTTTATTCTCTTATCTGGCCCCACCCACAT
82561  TCCAACGGTGACGACCGAGTCCGTCGGACGAAAATAAGAGAATAGACCGGGGTGGGTGTA

82621  CCTGCTGATTGGTCCATTTTACAGAGAGCTGATTGGTCTGTTTTGTCAGGGTGCTGATTG
82621  GGACGACTAACCAGGTAAAATGTCTCTCGACTAACCAGACAAAACAGTCCCACGACTAAC
```

FIGURE 4 (cont'd)

```
82681  GTGCATTTACAATCCCTGAGCTAGACACAAAAGTTCTCCACCTCCCCACTGGATTAGCTA
82681  CACGTAAATGTTAGGGACTCGATCTGTGTTTTCAAGAGGTGGAGGGGTGACCTAATCGAT

82741  GATAAAGAGTGTTGATTGATGTATTTACAAACCCTGAGCTAGACACAGAGTGCTGACTGG
82741  CTATTTCTCACAACTAACTACATAAATGTTTGGGACTCGATCTGTGTCTCACGACTGACC

82801  TGCATTTACAAACCTTGAGCTAGATACAGAGCGCCGATTGGTGTATTTACAATCTCTTAG
82801  ACGTAAATGTTTGGAACTCGATCTATGTCTCGCGGCTAACCACATAAATGTTAGAGAATC

82861  CCAGACGTAAAGATTCTCCAAGTCCCCACTAGACTCAGGAGCCCAGCTGGCTTCACCCAG
82861  GGTCTGCATTTCTAAGAGGTTCAGGGGTGATCTGAGTCCTCGGGTCGACCGAAGTGGGTC

82921  TGTATCTTGCACCAGGGCTGCAGGTGGAGCTGCCTGCCAGTCCCGCGCCGTGCACCCGCA
82921  ACATAGAACGTGGTCCCGACGTCCACCTCGACGGACGGTCAGGGCGCGGCACGTGGGCGT

82981  CTCCTCAGCCCTCGGGCTGTTGATGGGACCAGGCACTGTGGAGTGGGGGTGGCACTCGTC
82981  GAGGAGTCGGGAGCCCGACAACTACCCTGGTCCGTGACACCTCACCCCCACCGTGAGCAG

83041  AGGGAGACTCAGGCCGCACAGGATCCCACAGCGGTGGCGGGGGGCGGGAGGCTCAGGCAT
83041  TCCCTCTGAGTCCGGCGTGTCCTAGGGTGTCGCCACCGCCCCCCGCCCTCCGAGTCCGTA

83101  GGCAGGCTGCAGGTCCCGAGCCCTGCCCTGCAGGGAGGCAGCTAAGGCCTGGCGAGAAAT
83101  CCGTCCGACGTCCAGGGCTCGGGACGGGACGTCCCTCCGTCGATTCCGGACCGCTCTTTA

83161  CTGGCCCGGCGAGAAATCTAGCGCGGCGCCGGTGGGCCAGCACTGCTGGGGGACCCGGTG
83161  GACCGGGCCGCTCTTTAGATCGCGCCGCGGCCACCCGGTCGTGACGACCCCCTGGGCCAC

83221  CACCCTCCGCAGCTGCTGGCCTGGGTGCTAAGCCCCTCACTGCCCGGGGCTGGCAGGGCC
83221  GTGGGAGGCGTCGACGACCGGACCCACGATTCGGGGAGTGACGGGCCCCGACCGTCCCGG

83281  AGCCGGTCGCTCCGAATGTGGGGCCTGCCAAGCCCACTCCCACCCGGAACTCCAGCTGGC
83281  TCGGCCAGCGAGGCTTACACCCCGGACGGTTCGGGTGAGGGTGGGCCTTGAGGTCGACCG

83341  CCACAAGCACCATGCGCAGCCCGGGTTCCCGCCCATGCCTCTCCCTCCACACCTCCCTGC
83341  GGTGTTCGTGGTACGCGTCGGGCCCAAGGGCGGGTACGGAGAGGGAGGTGTGGAGGGACG

83401  AAACTGAGGGAGCCAGCTCCGGCCTTGGCCAGCCCAGAGAATGGCTCCCACAGTGCAGCG
83401  TTTGACTCCCTCGGTCGAGGCCGGAACCGGTCGGGTCTCTTACCGAGGGTGTCACGTCGC

83461  GCAGGCTGAAGGGCTCCTCAAGCGCGGTCAGAATGGGCGCCGAGGCCGAGGAGGCACCGA
83461  CGTCCGACTTCCCGAGGAGTTCGCGCCAGTCTTACCCGCGGCTCCGGCTCCTCCGTGGCT

83521  GAGCGAGTGAGGGCTGGGAGGGCTGCCAGCACTGTCATCTCTCAGTAGCATAATTTATAT
83521  CTCGCTCACTCCCGACCCTCCCGACGGTCGTGACAGTAGAGAGTCATCGTATTAAATATA

83581  AACCAGTCTCTTGTTGATTAACCATTATATTATTGATACAACAAATATAAGAACTGTTAC
83581  TTGGTCAGAGAACAACTAATTGGTAATATAATAACTATGTTGTTTATATTCTTGACAATG

83641  TAGTAAATTGAGTAAAGCAGGAAAATTAAATTTGGTCTCTTACTGATCAACCATTATTAG
83641  ATCATTTAACTCATTTCGTCCTTTTAATTTAAACCAGAGAATGACTAGTTGGTAATAATC

83701  ATTGTTGATAACAACGGTTATAATGTTGATCAACCGTTATATTGTTGATACAATTATATT
83701  TAACAACTATTGTTGCCAATATTACAACTAGTTGGCAATATAACAACTATGTTAATATAA

83761  TGATGTATTTAACCAGATTTTATATATTATATTCTATATTTTATAACATGTATATGTATC
83761  ACTACATAAATTGGTCTAAAATATATAATATAAGATATAAAATATTGTACATATACATAG

83821  ATATAACATACACATATACATACACACACAGCCAATATAATTTCACTCTCAGTTTCTT
83821  TATATTGTATGTGTATATGTATGTGTGTGTCGGTTATATTAAAGTGAGAGTCAAAGAA

83881  GTTAATCAACTATTATATTGTTGATACAACAAATATAAGAACTGTTACTAGTAAATTGAA
83881  CAATTAGTTGATAATATAACAACTATGTTGTTTATATTCTTGACAATGATCATTTAACTT

83941  TAAAGCAGGAAAATTACATTTGGGAGCCTTAAAGATGTAGGTTTTCTTCAAGAAATTTAT
83941  ATTTCGTCCTTTTAATGTAAACCCTCGGAATTTCTACATCCAAAAGAAGTTCTTTAAATA
```

FIGURE 4 (cont'd)

```
84001  GTTATAGTTATATCCCCTGTTGCATGTGAAGACATACATTCAAGGATACTTACTGCAGCT
84001  CAATATCAATATAGGGGACAACGTACACTTCTGTATGTAAGTTCCTATGAATGACGTCGA

84061  CACACAATTTGCTGATAGGTTGGATGTGGGTGTGAGAGAAGAGATGCTGTTTTCCTTTTG
84061  GTGTGTTAAACGACTATCCAACCTACACCCACACTCTCTTCTCTACGACAAAAGGAAAAC

84121  GGCGATGGCAACTGGATAGTAATCAGTTACTAAGATGGAGAAGAATGAGAAAGAAGTCTT
84121  CCGCTACCGTTGACCTATCATTAGTCAATGATTCTACCTCTTCTTACTCTTTCTTCAGAA

84181  TTTTTGTTTGTTTGTTGGAGTGATATGGGTAGAATAAAGAGTTAAAACAGTCATGCATGT
84181  AAAAACAAACAAACAACCTCACTATACCCATCTTATTTCTCAATTTTGTCAGTACGTACA

84241  AGGTACTATAGATGTGAAAGTCTATTATTCAGGGGAGAGTTTGGAATTTAAGGGATAGAT
84241  TCCATGATATCTACACTTTCAGATAATAAGTCCCCTCTCAAACCTTAAATTCCCTATCTA

84301  TCTTGTCTTTAAGCATAAAAATGCCTTTGAAGGTGAGAAACTAGATTACATCTTTAAGGG
84301  AGAACAGAAATTCGTATTTTTACGGAAACTTCCACTCTTTGATCTAATGTAGAAATTCCC

84361  GAGAACTGTAGATAGAGATGAAAAGAAGGCAAGGACTAACCCCTGGGGCAGGTCAGCATT
84361  CTCTTGACATCTATCTCTACTTTTCTTCCGTTCCTGATTGGGGACCCCGTCCAGTCGTAA

84421  TAGAAGAAAGGGAGTTTGTAAAATGCTAAGCTGTCTGAGAAGTAGCAGCCTGCAAGGTAG
84421  ATCTTCTTTCCCTCAAACATTTTACGATTCGACAGACTCTTCATCGTCGGACGTTCCATC

84481  AAAGAAAAACAAGAGAGGTTTAGTGGACTGGAAACCAAGTGAACCTGGAAACCACAAGCA
84481  TTTCTTTTTGTTCTCTCCAAATCACCTGACCTTTGGTTCACTTGGACCTTTGGTGTTCGT

84541  GCTGTGGCAAATCCTGCCACAGTGAGATGAGGCCTGAGAACTGAGTATTGGATTTGTCAA
84541  CGACACCGTTTAGGACGGTGTCACTCTACTCCGGACTCTTGACTCATAACCTAAACAGTT

84601  TGTGGAAGGCCTTGGTGATCTTGATAAGCATGTGTTAGTAGGGTGGTGGGGTTGCAGTGT
84601  ACACCTTCCGGAACCACTAGAACTATTCGTACACAATCATCCCACCACCCCAACGTCACA

84661  TTGGGATTTGAGAGAGAATGAGAGAATGGGTAGCTGTGATGCTGAATATGTATAAATGTC
84661  AACCCTAAACTCTCTCTTACTCTCTTACCCATCGACACTACGACTTATACATATTTACAG

84721  AGGAGGAATTCTGCTGCACAGTGGGACAGAGAATGGACGCCTCCCTGGAATGGGATAGAT
84721  TCCTCCTTAAGACGACGTGTCACCCTGTCTCTTACCTGCGGAGGGACCTTACCCTATCTA

84781  GCTAAAGGGAGTGTGTGTCTTAGTTGGGAGATATTATGGCATATCTCTAAGCTAATAGTA
84781  CGATTTCCCTCACACACAGAATCAACCCTCTATAATACCGTATAGAGATTCGATTATCAT

84841  ATTAGCCAGAAGGAAGGGAAAACTGATAGTTAAAATGTTCAAAAGATTTCTCACATTCCT
84841  TAATCGGTCTTCCTTCCCTTTTGACTATCAATTTTACAAGTTTTCTAAAGAGTGTAAGGA

84901  ATCATAGTTCTCCTTCTAAGGGTGACCACGTAAGCAGGCAGGGACATGTGACTAAGTAGA
84901  TAGTATCAAGAGGAAGATTCCCACTGGTGCATTCGTCCGTCCCTGTACACTGATTCATCT

84961  GCAGCATCCAACCCCCACCTTGCACGGTCACACACATACTGCTCCATTGACACCATAGTG
84961  CGTCGTAGGTTGGGGGTGGAACGTGCCAGTGTGTGTATGACGAGGTAACTGTGGTATCAC

85021  TCTTGGTGTGAGCTCTTCATTACTTGACTCAATATCTTCTTCCTCATCTTCTCAGCTGGG
85021  AGAACCACACTCGAGAAGTAATGAACTGAGTTATAGAAGAAGGAGTAGAAGAGTCGACCC

85081  GTTGTTGGGTATATCTTCATTGCAGATATTAGCATCCTTTGTGGTCTTCTTTGTCTTTCT
85081  CAACAACCCATATAGAAGTAACGTCTATAATCGTAGGAAACACCAGAAGAAACAGAAAGA

85141  ATTCTAGGGCCAGCACAGTGGGAGGCCTAGTCAACTGAGCCAAACTGGATATAACCACAA
85141  TAAGATCCCGGTCGTGTCACCCTCCGGATCAGTTGACTCGGTTTGACCTATATTGGTGTT

85201  GGTCAAAGGCAGGAGCCAGTGAGGAGATCTTGGTTTTGAGGTCATTCCTGGTGCGATAAT
85201  CCAGTTTCCGTCCTCGGTCACTCCTCTAGAACCAAAACTCCAGTAAGGACCACGCTATTA

85261  AACCATCCTGTAGCATTTTCTTGTCACTTAGAGAAGGCATTAGCAAACTTTTTTTGTAAT
85261  TTGGTAGGACATCGTAAAAGAACAGTGAATCTCTTCCGTAATCGTTTGAAAAAAACATTA

85321  GGTCTTAATAGTAAATATTTTTGACTTTGTAGGCTATATAGCTACTCACCTCTGCTTTTA
85321  CCAGAATTATCATTTATAAAAACTGAAACATCCGATATATCGATGAGTGGAGACGAAAAT
```

FIGURE 4 (cont'd)

```
85381  TAGTACAAAAGCAGCCATAGACAATTCATAAATAAATGAACATGTTCCAATAAAAGCAAT
85381  ATCATGTTTTCGTCGGTATCTGTTAAGTATTTATTTACTTGTACAAGGTTATTTTCGTTA

85441  AAAAACAGGCAGCAGTCTGAAGTTTGTCGACTCCTGATTTGGAATCATTCCAGAAAGTGA
85441  TTTTTGTCCGTCGTCAGACTTCAAACAGCTGAGGACTAAACCTTAGTAAGGTCTTTCACT

85501  AATAAAAGCATGGCTTCTTTTTTTTTTTTGAGACAAGGTCCTGCTCTGTCACCCAGGCTG
85501  TTATTTTCGTACCGAAGAAAAAAAAAAAACTCTGTTCCAGGACGAGACAGTGGGTCCGAC

85561  GAGTACAGTGGTGTCATCATAGCTCACTACAGCTTTGACCTGCTGGGCTCAAGTAATCCT
85561  CTCATGTCACCACAGTAGTATCGAGTGATGTCGAAACTGGACGACCCGAGTTCATTAGGA

85621  CCCACCTCAGCCTCCTGAGTGTCTAGGATTACAGGCATGTGACATCACACCTGGCTGATT
85621  GGGTGGAGTCGGAGGACTCACAGATCCTAATGTCCGTACACTGTAGTGTGGACCGACTAA

85681  TTTAAATTTTTTGTATTGATGGGGTCTTGCTATGTTGCCAAGGCTCACCTCAAACTCCTG
85681  AAATTTAAAAAACATAACTACCCCAGAACGATACAACGGTTCCGAGTGGAGTTTGAGGAC

85741  GTCTCAAGTGATTCTCCTGCCCCAGGATTCCCAAAGTGCTGGAATTAAAGGTGGGGGCCA
85741  CAGAGTTCACTAAGAGGACGGGGTCCTAAGGGTTTCACGACCTTAATTTCCACCCCCGGT

85801  CTGTGCCCAGCCAAAGCATGGCCTTTCTAAGGGTGACTCACTGTAGCACCTCCTGAGTAC
85801  GACACGGGTCGGTTTCGTACCGGAAAGATTCCCACTGAGTGACATCGTGGAGGACTCATG

85861  TTTCCTATCAAGTATCCTGGACCACTTTTTAAATAATTCATTCTAAAATCCATATTTAGG
85861  AAAGGATAGTTCATAGGACCTGGTGAAAAATTTATTAAGTAAGATTTTAGGTATAAATCC

85921  ATTGATTCCAAATTTGCCAACCTGTATTTCCCAAATAACTGGCATGATAATTCTTGTCTT
85921  TAACTAAGGTTTAAACGGTTGGACATAAAGGGTTTATTGACCGTACTATTAAGAACAGAA

85981  TTTCCCAAATGCTCATTTTCCACTCTTCAAACGTTGAGTGTTGACTGAAGTTCTGTGGTC
85981  AAAGGGTTTACGAGTAAAAGGTGAGAAGTTTGCAACTCACAACTGACTTCAAGACACCAG

86041  ATGTTTACAGTCATTTTTATGCTCTTTGTATTGTACATATGGAATGTATTTACCATTTCT
86041  TACAAATGTCAGTAAAAATACGAGAAACATAACATGTATACCTTACATAAATGGTAAAGA

86101  GTTAGCAATGATTTCTTGCAAAAACAATAACACTTGCTAGTTTTGCGTGCTTAAAAATAT
86101  CAATCGTTACTAAAGAACGTTTTTGTTATTGTGAACGATCAAAACGCACGAATTTTTATA

86161  CGAGCCCTGGCGCTCTCTACATATTATCTAATGTAATCCTTATATTAACCCTGTGAGTAG
86161  GCTCGGGACCGCGAGAGATGTATAATAGATTACATTAGGAATATAATTGGGACACTCATC

86221  TCATTATCACTTCTTTTACACAGGAGGAAGTTAGAGAAGAGTGTTAAGTGTTCTTTTAGT
86221  AGTAATAGTGAAGAAAATGTGTCCTCCTTCAATCTCTTCTCACAATTCACAAGAAAATCA

86281  CTTTTCCACTTTGTCTTTAGTTTTCTCTTAGCTGTTCTTTTAAATTCTAAGGTCATTCTT
86281  GAAAAGGTGAAACAGAAATCAAAAGAGAATCGACAAGAAAATTTAAGATTCCAGTAAGAA

86341  ATCTAAAAACTTAAAAGAGTTGATCAGCCTTTTTATCAAGTGTTAAATTTATACCATCTT
86341  TAGATTTTTGAATTTTCTCAACTAGTCGGAAAAATAGTTCACAATTTAAATATGGTAGAA

86401  TTTTCTTTGAACAGGGCTGGCAAATAAACCTGGATTTATCCCTTAGCCCTGTTACAGCAA
86401  AAAAGAAACTTGTCCCGACCGTTTATTTGGACCTAAATAGGGAATCGGGACAATGTCGTT

86461  GTCTCAGAAAATTTGGACTTCAGCTTCCCCAGTGATACATTTTCAAATGCCTATGTATGG
86461  CAGAGTCTTTTAAACCTGAAGTCGAAGGGGTCACTATGTAAAAGTTTACGGATACATACC

86521  ATTGCCCATTAGTTTATTGGCTAGAACTAAAGGAAATCAGGCAAGAGACAAAGGACATTT
86521  TAACGGGTAATCAAATAACCGATCTTGATTTCCTTTAGTCCGTTCTCTGTTTCCTGTAAA

86581  TCTCTAGGGTGTCAGTCAATGCAAAATGTTAGGACTAGCATCAGATTTTTTTTTTTTAC
86581  AGAGATCCCACAGTCAGTTACGTTTTACAATCCTGATCGTAGTCTAAAAAAAAAAAAATG

86641  AAATGAAAATAATTGAAGTGTGGTAGAAGTTCCTCTAGTTCTGGCAGTCCTAAATATTGT
86641  TTTACTTTTATTAACTTCACACCATCTTCAAGGAGATCAAGACCGTCAGGATTTATAACA
```

FIGURE 4 (cont'd)

```
86701  TAACTGAAGAATGATGGGGCTCATAAATTTGGAAAGGAGAGAGAGGGGAAAAGAGAACGG
86701  ATTGACTTCTTACTACCCCGAGTATTTAAACCTTTCCTCTCTCTCCCCTTTTCTCTTGCC

86761  GAATTTATGCTGAGCAGGGTGGCTAAGTATATATATTCAACAGGCTATAGGAGGAGCTAT
86761  CTTAAATACGACTCGTCCCACCGATTCATATATATAAGTTGTCCGATATCCTCCTCGATA

86821  TAGTATTCACGAAGGGGTGGCACACACATACATAGTAGGCTAACGTGTATGCAGCATGCA
86821  ATCATAAGTGCTTCCCCACCGTGTGTGTATGTATCATCCGATTGCACATACGTCGTACGT

86881  TCACAAGTCACTTTGGGGTGGAGCCTTAACATTTAATGTATTACAGTTAGGGCCTATACA
86881  AGTGTTCAGTGAAACCCCACCTCGGAATTGTAAATTACATAATGTCAATCCCGGATATGT

86941  TCAAAAGGTGAAGCAGAGGACACGAAAGCCCTCTGTGTTCAGCCTCTGTAGACTGGCCAG
86941  AGTTTTCCACTTCGTCTCCTGTGCTTTCGGGAGACACAAGTCGGAGACATCTGACCGGTC

87001  AACCACTCTGTGGTCAGAGGTCTCTTATCAGGAAAGAATTCTTGTTGTTGTGTTGAAACT
87001  TTGGTGAGACACCAGTCTCCAGAGAATAGTCCTTTCTTAAGAACAACAACACAACTTTGA

87061  GCAAAAGGGAGGGGCAGTGTCAGGCAGGTGGTTGATACCAGTCTTTCAGAAGCACTGGTT
87061  CGTTTTCCCTCCCCGTCACAGTCCGTCCACCAACTATGGTCAGAAAGTCTTCGTGACCAA

87121  TCTGTTTAACTTTTAGAGAAGAAAGCTTAATTGGAGTGTTACTGGAAAGGGATACCAGTC
87121  AGACAAATTGAAAATCTCTTCTTTCGAATTAACCTCACAATGACCTTTCCCTATGGTCAG

87181  TAGACCACAAGAGAGGATTCTTGGATCTCACACAAGAAAGAATTCAGGGTGAGTCCACAG
87181  ATCTGGTGTTCTCTCCTAAGAACCTAGAGTGTGTTCTTTCTTAAGTCCCACTCAGGTGTC

87241  AGTAAGATCAAATCAAGTTTATTAGAGAAGTAGAGAAAGAAAAGAATGGCTGCTCCATAA
87241  TCATTCTAGTTTAGTTCAAATAATCTCTTCATCTCTTTCTTTTCTTACCGACGAGGTATT

87301  GCAGAGCAGCACTGAAGGCTTCTGGTTGGCTATTTTTATGGTTATTTCTTGATTATATGC
87301  CGTCTCGTCGTGACTTCCGAAGACCAACCGATAAAAATACCAATAAAGAACTAATATACG

87361  TAAGCAAGGGGTGGATTATTCATGAATTTTCCTGGAAAGGGGTGGGGAATTCCCAGAACT
87361  ATTCGTTCCCCACCTAATAAGTACTTAAAAGGACCTTTCCCCACCCCTTAAGGGTCTTGA

87421  GAGGGTTTCTCTTTCTTTTAGACCATATAGGGTAAGTTCTGGATGTTGCCATGGCATTTG
87421  CTCCCAAAGAGAAAGAAAATCTGGTATATCCCATTCAAGACCTACAACGGTACCGTAAAC

87481  TAAACTGCCATGGCACTGGTGGGAATGTCTTTTAGCATGCTAATGCATTATAAATAGCCT
87481  ATTTGACGGTACCGTGACCACCCTTACAGAAAATCGTACGATTACGTAATATTTATCGGA

87541  AAGATAAGCTGTGAGGATGACCAGAGGTCAATTTCCTCACCATCTTAGTTTTGGCTGGCT
87541  TTCTATTCGACACTCCTACTGGTCTCCAGTTAAAGGAGTGGTAGAATCAAAACCGACCGA

87601  TCTTTACTGCTTCCTGTTTTTTCAGTGGGTTCTTTGTGACCTGTATCTGTCTTGTAATGT
87601  AGAAATGACGAAGGACAAAAAAGTCACCCAAGAAACACTGGACATAGACAGAACATTACA

87661  CCTATCTCATCCTGTAACTAAGAATGCCTGACCTCTTAGGAATGGAGCACAGGTGGTCTC
87661  GGATAGAGTAGGACATTGATTCTTACGGACTGGAGAATCCTTACCTCGTGTCCACCAGAG

87721  AGCTCATTTTATCTATCCCCTATTCAAGATGGAGTTGCTCTGGTTTAAATAGTTCTGACA
87721  TCGAGTAAAATAGATAGGGGATAAGTTCTACCTCAACGAGACCAAATTTATCAAGACTGT

87781  TATTTTCCCCAATACCCCCAAAGGGACCCTTAATCCCAAAATTTGCAGACAGATGAAGAT
87781  ATAAAAGGGGTTATGGGGGTTTCCCTGGGAATTAGGGTTTTAAACGTCTGTCTACTTCTA

87841  CCATCTTCTGTAGCTTCTTCTGGCTGAATACAGGTGATGATATTTTTGCGTAACTATTCA
87841  GGTAGAAGACATCGAAGAAGACCGACTTATGTCCACTACTATAAAAACGCATTGATAAGT

87901  GGGTAGAGAGGAGCTCAGTGAGAAAGCATCAGTATGGTGAAGGCCATTCCTAACTCCCGA
87901  CCCATCTCTCCTCGAGTCACTCTTTCGTAGTCATACCACTTCCGGTAAGGATTGAGGGCT

87961  GTTCTCACAAAAGGTGATATCTGGAAGATTAATAAGTGTTCAATTTAAGAAAACGTTGAG
87961  CAAGAGTGTTTTCCACTATAGACCTTCTAATTATTCACAAGTTAAATTCTTTTGCAACTC

88021  TAGTCTTATCCTGCATTGCTACACAAAGAGTACCACAGCAATATATTCTACAACAGTAAA
88021  ATCAGAATAGGACGTAACGATGTGTTTCTCATGGTGTCGTTATATAAGATGTTGTCATTT
```

FIGURE 4 (cont'd)

```
88081  GCAAAATAAGTATAATTATCCCAACTAAACTAAATAACAAGCCTTTCCATGAACTAGGCA
88081  CGTTTTATTCATATTAATAGGGTTGATTTGATTTATTGTTCGGAAAGGTACTTGATCCGT

88141  GTTGTTGGAACCAAGCTTATATGGGGTTGCTAGCCAATTCCAACACATGTTCAGAATTAA
88141  CAACAACCTTGGTTCGAATATACCCCAACGATCGGTTAAGGTTGTGTACAAGTCTTAATT

88201  AATACTGATCCAGATATTTATGTTACCCTTCTGTTTCTTCTGAGCAGCAGCTAGAGATCA
88201  TTATGACTAGGTCTATAAATACAATGGGAAGACAAAGAAGACTCGTCGTCGATCTCTAGT

88261  CTGGTTGGTTCATAGGAACAAACAGGGTCAATCTAAATGGCAGAAAAAACTCGAAACAAT
88261  GACCAACCAAGTATCCTTGTTTGTCCCAGTTAGATTTACCGTCTTTTTTGAGCTTTGTTA

88321  GAATGGGACTAGAGTTGAATAACAAGTATACCATAGTTTCTGAAACATAATTTTTCTCTC
88321  CTTACCCTGATCTCAACTTATTGTTCATATGGTATCAAAGACTTTGTATTAAAAAGAGAG

88381  TCCAGTCTCCCATTTCTATTGAAAACAAATCATGGTAGGACTGATTTGTTTGCACAGTAA
88381  AGGTCAGAGGGTAAAGATAACTTTTGTTTAGTACCATCCTGACTAAACAAACGTGTCATT

88441  GCTTTAGTCTTATTATGCTTGGCCTGGTTATTTGTATAAAGCATAGCAATAATAATTATT
88441  CGAAATCAGAATAATACGAACCGGACCAATAAACATATTTCGTATCGTTATTATTAATAA

88501  TGCCACGTAGGCTTTTTAAAAATTGACTTTGATGTAACTTTGTTCCATAAGAAATCTCAG
88501  ACGGTGCATCCGAAAAATTTTTAACTGAAACTACATTGAAACAAGGTATTCTTTAGAGTC

88561  ATTAGACTTTCTAAAGCCTTGAGCTGAGACACAGATTTATCTGTGCCTGCAAATACTTGT
88561  TAATCTGAAAGATTTCGGAACTCGACTCTGTGTCTAAATAGACACGGACGTTTATGAACA

88621  ATGAGTTGTGTGAATTCTCCTTTTGAGATCCCAAGATAACTGGAGCTCCTAGGCCTGTCA
88621  TACTCAACACACTTAAGAGGAAAACTCTAGGGTTCTATTGACCTCGAGGATCCGGACAGT

88681  GAAAGTGACATTCTTTACTTACCACAGGTCAGGAACCCTGTAAAGGAATTGAATAGACAA
88681  CTTTCACTGTAAGAAATGAATGGTGTCCAGTCCTTGGGACATTTCCTTAACTTATCTGTT

88741  GGTATGAGGCCAGCTTTTCCCAGAGGCTTTCATCAGTTCTGTAAGTCAACTTTGATTCCT
88741  CCATACTCCGGTCGAAAAGGGTCTCCGAAAGTAGTCAAGACATTCAGTTGAAACTAAGGA

88801  TAAAGCAATCTGTTTATATTTGAAAGAATGCCATTCCAGTCAAAGCCCTGGTAAAACAGG
88801  ATTTCGTTAGACAAATATAAACTTTCTTACGGTAAGGTCAGTTTCGGGACCATTTTGTCC

88861  CAGTGTCTCTAATTGTGTCCTGTTACAAAGGAAAACAGATTTTTATTGCACATATGCAAA
88861  GTCACAGAGATTAACACAGGACAATGTTTCCTTTTGTCTAAAAATAACGTGTATACGTTT

88921  TTACTATGCTGCCATAAGTTAAGAATACTCACAAATAGTTTTCAAATTCTTGAGAAATCA
88921  AATGATACGACGGTATTCAATTCTTATGAGTGTTTATCAAAAGTTTAAGAACTCTTTAGT

88981  GGTAGAAAGAAATATGCTCCAAAGTTTTCTCATAGGAGTATAATTTACTCAATTTTTAAC
88981  CCATCTTTCTTTATACGAGGTTTCAAAAGAGTATCCTCATATTAAATGAGTTAAAAATTG

89041  AACTGTAAATAGCTCAAAAAAAAAGGTTTCTTGACTCTGAAAAATGAAACAAAGGATCAG
89041  TTGACATTTATCGAGTTTTTTTTTCCAAAGAACTGAGACTTTTTACTTTGTTTCCTAGTC

89101  CAACATTTTAAGCAAAAAGTCACTAGAAGGTTATTTTGTTCTTTTATTAGTTTAGTCCAT
89101  GTTGTAAAATTCGTTTTTCAGTGATCTTCCAATAAAACAAGAAAATAATCAAATCAGGTA

89161  GCTGTTAATTCCTGTTTGCTCAATATTTATGAACATATTGGTTTCCCAAGGGAGACTCTT
89161  CGACAATTAAGGACAAACGAGTTATAAATACTTGTATAACCAAAGGGTTCCCTCTGAGAA

89221  GAAAGTTTTTTTCCTCTCTATCTTAATGGCACACTTTACAAAATTTTTCAGAAACCTGCA
89221  CTTTCAAAAAAAGGAGAGATAGAATTACCGTGTGAAATGTTTTAAAAAGTCTTTGGACGT

89281  TTTAAGAGCCCTCTATCTGATTATAAACCATCTTTTAAAGAGGATCAAAACAAGACAACA
89281  AAATTCTCGGGAGATAGACTAATATTTGGTAGAAAATTTCTCCTAGTTTTGTTCTGTTGT

89341  ATTTTCTGTGGATAACATAAAGTCTTAGGAAAGCCGTGGTTAAAGACACAATTGACTAGA
89341  TAAAAGACACCTATTGTATTTCAGAATCCTTTCGGCACCAATTTCTGTGTTAACTGATCT
```

FIGURE 4 (cont'd)

```
89401  AATTTTGGTTACTTCTGTGGCATACAACAATTTTACATAACAGTTATAATTATTACTGAT
89401  TTAAAACCAATGAAGACACCGTATGTTGTTAAAATGTATTGTCAATATTAATAATGACTA

89461  AACATAAACTAAGTCATATCAGAATTTAGGAGTTTCCCATAATTTTGGAGCACATACCAA
89461  TTGTATTTGATTCAGTATAGTCTTAAATCCTCAAAGGGTATTAAAACCTCGTGTATGGTT

89521  TAACATATTTATACAAATACAACTCAAAGAAAGCTAAACACCATTTCATATTTGACAGTG
89521  ATTGTATAAATATGTTTATGTTGAGTTTCTTTCGATTTGTGGTAAAGTATAAACTGTCAC

89581  CTTCCTGTATGATTTTAATATACCAAGTAAGCCAAATATGTCATTTTTGGACTTAAGGGG
89581  GAAGGACATACTAAAATTATATGGTTCATTCGGTTTATACAGTAAAAACCTGAATTCCCC

89641  ACCTCATATCTAAAGTATTGACCAGATCAGGAAAAGGCATAATATAGAATTTGATTTTGG
89641  TGGAGTATAGATTTCATAACTGGTCTAGTCCTTTTCCGTATTATATCTTAAACTAAAACC

89701  AAAGTTTGTCAAATATCAAAGATTTAAAACACTTGATATTATAAAATCGAATCCCAGGTC
89701  TTTCAAACAGTTTATAGTTTCTAAATTTTGTGAACTATAATATTTTAGCTTAGGGTCCAG

89761  ACTGTAAAGTCATTTATTTAGCCAAAATGATAATACAAAGATTTTCCAAAAGCAAAATCC
89761  TGACATTTCAGTAAATAAATCGGTTTTACTATTATGTTTCTAAAAGGTTTTCGTTTTAGG

89821  TTTTTTTTTTTTTTTTTTTTGAGACGAGTCTGGCTCTGTCGCCCAGCTGGAGTTCAGTG
89821  AAAAAAAAAAAAAAAAAAAAACTCTGCTCAGACCGAGACAGCGGGTCGACCTCAAGTCAC

89881  GCGCCATCTCGGCTCACTGCAAGCTCCGTCTCCCGGGTTCACACCATTCTCCTGCCTCAG
89881  CGCGGTAGAGCCGAGTGACGTTCGAGGCAGAGGGCCCAAGTGTGGTAAGAGGACGGAGTC

89941  CCTCCGGGGTAGCTGGGACTACAGGCACCCGCCACTACGCTTGGCTAATTTTTTTGTGTT
89941  GGAGGCCCCATCGACCCTGATGTCCGTGGGCGGTGATGCGAACCGATTAAAAAAACACAA

90001  TTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTCCATCTTCTGACCTCGTG
90001  AAATCATCTCTGCCCCAAAGTGGTACAATCGGTCCTACCAGAGGTAGAAGACTGGAGCAC

90061  ATCCACCCGCCTCAGCCTCCCAAAGTGCTGGTATTACAGGCGTGAGCCACTGCGCCCAGC
90061  TAGGTGGGCGGAGTCGGAGGGTTTCACGACCATAATGTCCGCACTCGGTGACGCGGGTCG

90121  AAGCAAAATCCTTTACTCATTGATAGAGGGAAGACAGCTTTCCAAACAATGTCTCACTTT
90121  TTCGTTTTAGGAAATGAGTAACTATCTCCCTTCTGTCGAAAGGTTTGTTACAGAGTGAAA

90181  TTCTTCTTTATTTTGTTGTTTATTCAAAAGGCAAACAAAAATCTTCCATGATCTTTTAAT
90181  AAGAAGAAATAAAACAACAAATAAGTTTTCCGTTTGTTTTTAGAAGGTACTAGAAAATTA

90241  ATTACATGAAAATTTTGTTCAAGAGAGAAAGCCAAATTTCATCTTTGCATTAATGAATGT
90241  TAATGTACTTTTAAAACAAGTTCTCTCTTTCGGTTTAAAGTAGAAACGTAATTACTTACA

90301  CAAATCCAATTCTTAAGAAAACCTTGTAGACAAATTATTCAATCTTAATCAATTTTACCA
90301  GTTTAGGTTAAGAATTCTTTTGGAACATCTGTTTAATAAGTTAGAATTAGTTAAAATGGT

90361  TGAGATAAGATTCTCATAAACCTTTTGTAGTCCTTTACAATTTTTTTTTGTTGTTGTTAA
90361  ACTCTATTCTAAGAGTATTTGGAAAACATCAGGAAATGTTAAAAAAAAACAACAACAATT

90421  AGAGCAGATTAATGCTCTAAGAAAACCCTGTTGTGCTTTTATTCCAATGTTCTATTTATG
90421  TCTCGTCTAATTACGAGATTCTTTTGGGACAACACGAAAATAAGGTTACAAGATAAATAC

90481  GAAAAACTGAATAATACCCTTTAACTTTAGCTAATATTTTAACACACAGAATTTCTTTTA
90481  CTTTTTGACTTATTATGGGAAATTGAAATCGATTATAAAATTGTGTGTCTTAAAGAAAAT

90541  CAAGATTAATTTTCACTGACCTCCCACAACTTACTCAAACCTTTAGCTTTATCCTATGTA
90541  GTTCTAATTAAAAGTGACTGGAGGGTGTTGAATGAGTTTGGAAATCGAAATAGGATACAT

90601  ACTTAAAACAGTCTTTTAACTCTGTAAACTAGGCAATAAGACCACATTCCCATGCCTTCT
90601  TGAATTTTGTCAGAAAATTGAGACATTTGATCCGTTATTCTGGTGTAAGGGTACGGAAGA

90661  TATAATCTTTTACCAAAAACACATTCTATTTTCCTTACACGCCTTGCAGGTAACACTGTT
90661  ATATTAGAAAATGGTTTTTGTGTAAGATAAAAGGAATGTGCGGAACGTCCATTGTGACAA

90721  TCTCTGGTAGCCTCAATTACATGTGCTACAATGTTAACTCTTAGCAACTTTTATATTTGG
90721  AGAGACCATCGGAGTTAATGTACACGATGTTACAATTGAGAATCGTTGAAAATATAAACC
```

FIGURE 4 (cont'd)

```
90781  TGAAAAGTCTGATAAATAAGTGGTTTTAATTATGTACCAGGTGTGGAGCCTAGGACGCCA
90781  ACTTTTCAGACTATTTATTCACCAAAATTAATACATGGTCCACACCTCGGATCCTGCGGT

90841  GACAGAAGTGCAGATAAGGTCTGAGTCTTTCCAGCATAGCTAGGGGGCATGGCCAACTCC
90841  CTGTCTTCACGTCTATTCCAGACTCAGAAAGGTCGTATCGATCCCCCGTACCGGTTGAGG

90901  ACATGTCCCCAGGCCCTACCTAGAATCTAATGGCTCCAAAGCAGGTGAGTTGAACAATTA
90901  TGTACAGGGGTCCGGGATGGATCTTAGATTACCGAGGTTTCGTCCACTCAACTTGTTAAT

90961  TCAAGTTAAAGAAGCAGTTTATGGCATTGAAGCATTCAGCAAATCTAATTTAATCTGACC
90961  AGTTCAATTTCTTCGTCAAATACCGTAACTTCGTAAGTCGTTTAGATTAAATTAGACTGG

91021  TAATTTAGACCAGATGTCTAAATTTTGAAGACATTTTTATTTTACCAATATATTTGATTC
91021  ATTAAATCTGGTCTACAGATTTAAAACTTCTGTAAAAATAAAATGGTTATATAAACTAAG

91081  TCTTTATTTCCCAAAGATTATTAAAGTCCCATGAATTAAAAGGTGTTAAAGTTTTTATTT
91081  AGAAATAAAGGGTTTCTAATAATTTCAGGGTACTTAATTTTCCACAATTTCAAAAATAAA

91141  TTCTGACAAAATATTCAATTTAAGTGCTTATTTTTCAAGCCAATTAATTAGAGCTCTTTT
91141  AAGACTGTTTTATAAGTTAAATTCACGAATAAAAAGTTCGGTTAATTAATCTCGAGAAAA

91201  CTATGCATATCACACATACAACACACAAAAATACACAGACAGAAGACCCAGTAGCTGTTA
91201  GATACGTATAGTGTGTATGTTGTGTGTTTTTATGTGTCTGTCTTCTGGGTCATCGACAAT

91261  ATTTTTTCATTTGCCAGCTTTTAAGAATCTTAATTGGATTACTGGCTTCAGGGTGGAGCA
91261  TAAAAAAGTAAACGGTCGAAAATTCTTAGAATTAACCTAATGACCGAAGTCCCACCTCGT

91321  ACGTGTGGGGACAGAGCCGGGAAAACATGCAGTTTCTGGGGCCTAATAAACAGGTGCAGT
91321  TGCACACCCCTGTCTCGGCCCTTTTGTACGTCAAAGACCCCGGATTATTTGTCCACGTCA

91381  TAGAAGGCAAAACAGATTCCCCAAAATTACAGATCTCATTTTTATATTGGATTCTGGATC
91381  ATCTTCCGTTTTGTCTAAGGGGTTTTAATGTCTAGAGTAAAAATATAACCTAAGACCTAG

91441  CCAAAAAGAGGGAATCAGCCCATCCCCTCTGGGAGTTTTGGAAGGTGTAGAAGAATGTTT
91441  GGTTTTTCTCCCTTAGTCGGGTAGGGGAGACCCTCAAAACCTTCCACATCTTCTTACAAA

91501  CCATACCTTCTAGGTGGCCAAGAGCATGCTTCTCTGATCCAAACGTGCACAGAGTGGGGT
91501  GGTATGGAAGATCCACCGGTTCTCGTACGAAGAGACTAGGTTTGCACGTGTCTCACCCCA

91561  ATTCCCCTATAACTGCTATTAGCCTTCCCTTAAAGCATATTTCCTGCCTAGTTATTACAC
91561  TAAGGGGATATTGACGATAATCGGAAGGGAATTTCGTATAAAGGACGGATCAATAATGTG

91621  ACCAAGGTTAAAAGCTCTCCCATAATGCAAAGTAATTTCTGATACCCCTACAAGTAAAAA
91621  TGGTTCCAATTTTCGAGAGGGTATTACGTTTCATTAAAGACTATGGGGATGTTCATTTTT

91681  ACATTAGGTAACACAATGCAAATCAGAGCAATGCCTTGGATTTTGAGAGGGATTTGCTTG
91681  TGTAATCCATTGTGTTACGTTTAGTCTCGTTACGGAACCTAAAACTCTCCCTAAACGAAC

91741  CCTTCAATTCCTGGGGTTCCATGAGGAAAACAGATTTTTCCCAAAATGGAGCCTGTGGTG
91741  GGAAGTTAAGGACCCCAAGGTACTCCTTTTGTCTAAAAAGGGTTTTACCTCGGACACCAC

91801  CTCCCTCAGTTTTTTCTAAGGAAGCCCAGGCTGTTAGAAATGATCTTAGGTCCTCTCAAT
91801  GAGGGAGTCAAAAAAGATTCCTTCGGGTCCGACAATCTTTACTAGAATCCAGGAGAGTTA

91861  GTGGGCATCAAGAGTGGCAAGAAAACAAAATGAGAAAAACAATTCAGTTGACTGAGAAGA
91861  CACCCGTAGTTCTCACCGTTCTTTTGTTTTACTCTTTTTGTTAAGTCAACTGACTCTTCT

91921  AAAAACTTTTATCTAGAAAAAAATCAATATAAAACCTGCCAGATAGATAGATAGATAGAT
91921  TTTTTGAAAATAGATCTTTTTTTAGTTATATTTTGGACGGTCTATCTATCTATCTATCTA

91981  AATCTTGGACATCACTTTTAATTAAGCTGACTTTTAACCAAATCTCTTATTATCAGACTC
91981  TTAGAACCTGTAGTGAAAATTAATTCGACTGAAAATTGGTTTAGAGAATAATAGTCTGAG

92041  TAGCCAGGACAAACAGCTAATATTTCTGGCTTTTGAACTTTACCAAAAGTAACCTCCCAG
92041  ATCGGTCCTGTTTGTCGATTATAAAGACCGAAAACTTGAAATGGTTTTCATTGGAGGGTC
```

FIGURE 4 (cont'd)

```
92101  GTGAAACCAATAAGCCTTAACTAAGGTTATGATTTAACCACAGGTGTGGGAGGTATTTTC
92101  CACTTTGGTTATTCGGAATTGATTCCAATACTAAATTGGTGTCCACACCCTCCATAAAAG

92161  AAAGAGGTGGTAAGCAGTTTTTACAAGATCTGGAACCTCCAAACATAGCTCAGAGAAAGG
92161  TTTCTCCACCATTCGTCAAAAATGTTCTAGACCTTGGAGGTTTGTATCGAGTCTCTTTCC

92221  AAGATTCAAGACAGGGAGTCAGAAGTTGTTCATGAGGGGAAGATAATTAATAAATGGCAA
92221  TTCTAAGTTCTGTCCCTCAGTCTTCAACAAGTACTCCCCTTCTATTAATTATTTACCGTT

92281  AGGTCACATAGATATCAAACCAGAAAGGACTCATTCCCGAAGCTAGGAATTGAATCCAGG
92281  TCCAGTGTATCTATAGTTTGGTCTTTCCTGAGTAAGGGCTTCGATCCTTAACTTAGGTCC

92341  CCACCACTGTGAAATAGCAAAGCCTTAGTTACTGAGCTGCAGTACTGCACAGTCTCCATT
92341  GGTGGTGACACTTTATCGTTTCGGAATCAATGACTCGACGTCATGACGTGTCAGAGGTAA

92401  GCTCCTCCGAGAAGGAGCCTAGAGCAGCCAATTTTGAGCTTGCAAAGGCTTTTAACTGCT
92401  CGAGGAGGCTCTTCCTCGGATCTCGTCGGTTAAAACTCGAACGTTTCCGAAAATTGACGA

92461  CAAGATAATTTTTAGGACTAATTGACAGGAACCCCAAAATTCACACCCTCTGGATGGCAG
92461  GTTCTATTAAAAATCCTGATTAACTGTCCTTGGGGTTTTAAGTGTGGGAGACCTACCGTC

92521  AGATCAGGAGAAAGTACCCCCACGTGGTTACAAGGTCAAGCTTTCAAGGACATAAAACAA
92521  TCTAGTCCTCTTTCATGGGGGTGCACCAATGTTCCAGTTCGAAAGTTCCTGTATTTTGTT

92581  GACGAGAGGGAAACTTCATCCAGTTTTTGTTTCAGGGACCTGCAGAAAAGTTTGTAACTG
92581  CTGCTCTCCCTTTGAAGTAGGTCAAAAACAAAGTCCCTGGACGTCTTTTCAAACATTGAC

92641  ACCAGTTTGCGGGCCAGCTTGAAGACCAGATTTGTAGGATTCCTAGGCCTGCATTCTATC
92641  TGGTCAAACGCCCGGTCGAACTTCTGGTCTAAACATCCTAAGGATCCGGACGTAAGATAG

92701  CTGTGGCTAGCTAGGGCTGCCATAATAGAATATTACAGACTAGGCAGCTTCAACCACAGG
92701  GACACCGATCGATCCCGACGGTATTATCTTATAATGTCTGATCCGTCGAAGTTGGTGTCC

92761  TGTTTATTTTCTCACTGCTCCAAAGTCTAGAAGTCCAAGATGAGGGTGCCTGTAAAGCTG
92761  ACAAATAAAAGAGTGACGAGGTTTCAGATCTTCAGGTTCTACTCCCACGGACATTTCGAC

92821  ATTTCTGGTAAGAGCTCTCCTGGCTTGTAGATGGCCACCTTCTCCCTATGTCCTCTCTGG
92821  TAAAGACCATTCTCGAGAGGACCGAACATCTACCGGTGGAAGAGGGATACAGGAGAGACC

92881  GCCTTTCCTCTATCTGGCCAAGGGAGCTATCTCTGGTGTCTTTTCCTCTTCTTATAAGGA
92881  CGGAAAGGAGATAGACCGGTTCCCTCGATAGAGACCACAGAAAAGGAGAAGAATATTCCT

92941  CAGCAGGCCTATGGATTAAGGCCTCACTTTCATGTCTTCATTTAAGTTTAATTACCTCTT
92941  GTCGTCCGGATACCTAATTCCGGAGTGAAAGTACAGAAGTAAATTCAAATTAATGGAGAA

93001  TAAAGGCCCGATGTCCAAATACAGTCACACCGTGGGTTAGGATTTCAACATAGGAATTTT
93001  ATTTCCGGGCTACAGGTTTATGTCAGTGTGGCACCCAATCCTAAAGTTGTATCCTTAAAA

93061  AGGGACACAATTCAACCCATAATAGGAGGGCATTGGAAGGAGAGTGTTCCAGGCAGAGGA
93061  TCCCTGTGTTAAGTTGGGTATTATCCTCCCGTAACCTTCCTCTCACAAGGTCCGTCTCCT

93121  AGTAGTGCAGTAGAGCATATGGAAGTAGGAAAGTGAGGAGCAATGGAAGGGGTTCAGTAT
93121  TCATCACGTCATCTCGTATACCTTCATCCTTTCACTCCTCGTTACCTTCCCCAAGTCATA

93181  GGTTTGAGCGTGGGGTGTGTAGAGGGTGAAGGCTATTTCACCCAGTGCAGTAAGCTCTAT
93181  CCAAACTCGCACCCCACACATCTCCCACTTCCGATAAAGTGGGTCACGTCATTCGAGATA

93241  CATCCACACCGGGGTTTTACAGTGGGAGAAAGGAGGGTGTTTATTTGCAGGTCAAGAAGC
93241  GTAGGTGTGGCCCCAAAATGTCACCCTCTTTCCTCCCACAAATAAACGTCCAGTTCTTCG

93301  AAACAGAATCAGGGAGCTCATGCTTAAGACCTGAACTCCACAGTGGCTTACATGTAAGGT
93301  TTTGTCTTAGTCCCTCGAGTACGAATTCTGGACTTGAGGTGTCACCGAATGTACATTCCA

93361  GCAGAGGTTACAGGCAAAGGCATCAATCAACACATGGAAGCCATATATTGGTTTGGCCTA
93361  CGTCTCCAATGTCCGTTTCCGTAGTTAGTTGTGTACCTTCGGTATATAACCAAACCGGAT

93421  AAAAGGTGGGATATCCTGAAGTAGGGGCTTACAGGTTACAGGTAAATCCAAATATTTTCT
93421  TTTTCCACCCTATAGGACTTCATCCCCGAATGTCCAATGTCCATTTAGGTTTATAAAAGA
```

FIGURE 4 (cont'd)

```
93481  GATTTGCACATTGTTTAAGGAAGGGGATATTTGTCTAAAGATGGGATCAGCAGGCCAGGT
93481  CTAAACGTGTAACAAATTCCTTCCCCTATAAACAGATTTCTACCCTAGTCGTCCGGTCCA

93541  GTGGTGGCTCATGCTTGTAACCTCAGCACTTTGAAAGGCTGAAGAAGGAGGGTAACTTGA
93541  CACCACCGAGTACGAACATTGGAGTCGTGAAACTTTCCGACTTCTTCCTCCCATTGAACT

93601  GGCCAGGATTTCAAGACCAGCCTAGTCAACATGGCAAGACTCTGTCTCTAAAAAAAATAA
93601  CCGGTCCTAAAGTTCTGGTCGGATCAGTTGTACCGTTCTGAGACAGAGATTTTTTTTATT

93661  AAAAGAAAAAAAGAAAAAGAAAAAAAATAAAGATGGAATCAGCAGAAAGGAATGTTAGGT
93661  TTTTCTTTTTTTCTTTTTCTTTTTTTTATTTCTACCTTAGTCGTCTTTCCTTACAATCCA

93721  CTGGCCCATGGGCGGACTTCCTCCAGACTTCTCAGAAAGAAATTTAAGACAAAGAATGGT
93721  GACCGGGTACCCGCCTGAAGGAGGTCTGAAGAGTCTTTCTTTAAATTCTGTTTCTTACCA

93781  GCCAGGCGTGGTAGTTCACGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGAGGGTGGAT
93781  CGGTCCGCACCATCAAGTGCGGACATTAGGGTCGTGAAACCCTCCGACTCCTCCCACCTA

93841  CACTTGAGGTCAGGGATTTGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTA
93841  GTGAACTCCAGTCCCTAAACTCTGGTCGGACCGGTTGTACCACTTTGGGGTAGAGATGAT

93901  AAAATACAAAAACTAGCCGGGTGTGGTGGTGCACGCCTGTAATCCCAGCTCCTCAGGAGG
93901  TTTTATGTTTTTGATCGGCCCACACCACCACGTGCGGACATTAGGGTCGAGGAGTCCTCC

93961  CTGAGGCAGGAGAATCGCTTGAACCTGGGAGGTGGAGGTTGCAGTGAGCCAAGATTGCGC
93961  GACTCCGTCCTCTTAGCGAACTTGGACCCTCCACCTCCAACGTCACTCGGTTCTAACGCG

94021  CACTGCACTCCACCCTGGGCAAGAGAGCGAGACTTGAGACTCCATCTAAAAAAACAAAAT
94021  GTGACGTGAGGTGGGACCCGTTCTCTCGCTCTGAACTCTGAGGTAGATTTTTTTGTTTTA

94081  GGAAAACAAAGAATGGTGATAATTCAGTCCTCAATTCCTTTTCTTTTTTTTTTCTTTCTT
94081  CCTTTTGTTTCTTACCACTATTAAGTCAGGAGTTAAGGAAAAGAAAAAAAAAAGAAAGAA

94141  TTTTTTTTTTTTTAACTGCTCCTGCGGAGCAGGGCTACCTCATAGCCAGTATGTCCAGAG
94141  AAAAAAAAAAAAATTGACGAGGACGCCTCGTCCCGATGGAGTATCGGTCATACAGGTCTC

94201  TTGTCTCAGTTCCCTCGTTATCTGTGGTCTGGAGGATCCATATGATGGTGATCCAGGTTC
94201  AACAGAGTCAAGGGAGCAATAGACACCAGACCTCCTAGGTATACTACCACTAGGTCCAAG

94261  TGAAAAACAACTCTGGAACATCTGTTAAGATGTTATCTTTAGTTTCTATAGGAAACCAAA
94261  ACTTTTTGTTGAGACCTTGTAGACAATTCTACAATAGAAATCAAAGATATCCTTTGGTTT

94321  CATCTTCTGACTCTGACCTCCTTGGTGATTATTTTAAGCTACTGTTTCCCTTTTTGCTTA
94321  GTAGAAGACTGAGACTGGAGGAACCACTAATAAAATTCGATGACAAAGGGAAAAACGAAT

94381  TCGAGTTGCTCATTTACTTCTCAAGGCTAGCAGGATGACTGGACTTTCTCTTGAAGGAAC
94381  AGCTCAACGAGTAAATGAAGAGTTCCGATCGTCCTACTGACCTGAAAGAGAACTTCCTTG

94441  TCAAGCTGTTCCTTTATTTTTATGCTCAGGGGCGGGGGGCACTCAGCAGGTCCCTAAGAG
94441  AGTTCGACAAGGAAATAAAAATACGAGTCCCCGCCCCCCGTGAGTCGTCCAGGGATTCTC

94501  GTATCCCCGCTGTGTCTCACTATGGGTGAAGGCTCAGCGGGAAGAAAGCTAGATACCAGT
94501  CATAGGGGCGACACAGAGTGATACCCACTTCCGAGTCGCCCTTCTTTCGATCTATGGTCA

94561  TGGAACTCTCTTAAAGTAAGGAGATCTTGGGTACTTTATTAATATTTTTTCAAGCTTTG
94561  ACCTTGAGAGAATTTCATTCCTCTAGAACCCATGAAATAATTATAAAAAAAGTTCGAAAC

94621  GTAAAGTATGAAATAACTCTGGATTAAGAGGAGAAGAACTAGAACTTAATAACCATGACA
94621  CATTTCATACTTTATTGAGACCTAATTCTCCTCTTCTTGATCTTGAATTATTGGTACTGT

94681  TTTGGGCAAATTGTTTGAAGCCTCAGTTTCCTCAACCTAAAAATAGGAATTATAACTTTT
94681  AAACCCGTTTAACAAACTTCGGAGTCAAAGGAGTTGGATTTTTATCCTTAATATTGAAAA

94741  TGAATTGTTCTGTTGTATGAGATAATGACAGTGAAATATTTTGGAGCCAAAAAGCAATGC
94741  ACTTAACAAGACAACATACTCTATTACTGTCACTTTATAAAACCTCGGTTTTTCGTTACG
```

FIGURE 4 (cont'd)

```
94801  ATACTATTGGGGGTTATTATTATTTTTTGATCTTTCAGCAAGCCAGCTATAGTAGCTTTT
94801  TATGATAACCCCCAATAATAATAAAAAACTAGAAAGTCGTTCGGTCGATATCATCGAAAA

94861  CCTATATAAACACAGTATGAATTCAGTGCCCTCGCTTTCCTGACCCTCACAGCATTCAAG
94861  GGATATATTTGTGTCATACTTAAGTCACGGGAGCGAAAGGACTGGGAGTGTCGTAAGTTC

94921  CCCTGTATGGACCGCTGAGCTCTCCCTTGAAGCACCCCTTAGTCCAGTATTCTTTACTAT
94921  GGGACATACCTGGCGACTCGAGAGGGAACTTCGTGGGGAATCAGGTCATAAGAAATGATA

94981  TTTCCCCTTTAACAATTTAAACTTCTTTCAAAACAACTGGCAGTAGTTTTTTAAAAAAGA
94981  AAAGGGGAAATTGTTAAATTTGAAGAAAGTTTTGTTGACCGTCATCAAAAAATTTTTTCT

95041  AAACAGGCTACTCTTCAAATTAAAAAAAAAATTGTCTTGCTTTAGGAAAAATGTTCTTCT
95041  TTTGTCCGATGAGAAGTTTAATTTTTTTTTTAACAGAACGAAATCCTTTTTACAAGAAGA

95101  GTCTTTTCCAATTATCTTATTTCTTCTGCCTCTCATCTGCCTTGCCTAGGCAGAATTATT
95101  CAGAAAAGGTTAATAGAATAAAGAAGACGGAGAGTAGACGGAACGGATCCGTCTTAATAA

95161  CTCAAGCCTGCACATGCTGCTGTGCCTCTCTGGCTTCTAGGCCAAAACATAGGCTAATAG
95161  GAGTTCGGACGTGTACGACGACACGGAGAGACCGAAGATCCGGTTTTGTATCCGATTATC

95221  TACTCTTCTCAGCTCACTGAGCTCTTTCCTGCCGTCATTCAAATCACCTCTGAAATCCTG
95221  ATGAGAAGAGTCGAGTGACTCGAGAAAGGACGGCAGTAAGTTTAGTGGAGACTTTAGGAC

95281  CCTTTTGATAAAGTAATTTCTGACTGATCAGAGGAAAGTGATGATTTTCTCTGTGACATG
95281  GGAAAACTATTTCATTAAAGACTGACTAGTCTCCTTTCACTACTAAAAGAGACACTGTAC

95341  TCTTTTTATTTTCCATCACACTCAGTGTCTTGTCCTTCTATGCAACATATCCATGTATTT
95341  AGAAAAATAAAAGGTAGTGTGAGTCACAGAACAGGAAGATACGTTGTATAGGTACATAAA

95401  AAAAACCCTTGTATATGTATTTGCCATGTATTTGTGAGGTATAGCTCGTGAAGTAAAAAA
95401  TTTTTGGGAACATATACATAAACGGTACATAAACACTCCATATCGAGCACTTCATTTTTT

95461  AGTCTTTATTTTTTTAAACACATCATCATTTCAGCTTCAAAGTCAATGTGGTCTCTTGAG
95461  TCAGAAATAAAAAAATTTGTGTAGTAGTAAAGTCGAAGTTTCAGTTACACCAGAGAACTC

95521  AAGTTAAATATGCTGCATTTGAAATTTTATGTTTCTCATTTTTCATGTTTTTTTTTAAAG
95521  TTCAATTTATACGACGTAAACTTTAAAATACAAAGAGTAAAAAGTACAAAAAAAAATTTC

95581  CCCCTTTTAATTCCTCATGTATTTGAATATTCCAGGAAGCTAGTAGAGTAAAAAGCCAAA
95581  GGGGAAAATTAAGGAGTACATAAACTTATAAGGTCCTTCGATCATCTCATTTTTCGGTTT

95641  GTGTTAAAGTGTTTATAAGGAATGTGACTCATCGAAATTGACTCACACAATTCATCAGCT
95641  CACAATTTCACAAATATTCCTTACACTGAGTAGCTTTAACTGAGTGTGTTAAGTAGTCGA

95701  ATTGTCATCCATATGTTCTCATTATTTATGTGGCTTTTGCTTTATTTAATCAGATGACTG
95701  TAACAGTAGGTATACAAGAGTAATAAATACACCGAAAACGAAATAAATTAGTCTACTGAC

95761  AACTTATCATCGTGGGATTTTGCTTGTAGTTGTTTAATATTATATAATAAAATACTCTTG
95761  TTGAATAGTAGCACCCTAAAACGAACATCAACAAATTATAATATATTATTTTATGAGAAC

95821  TGTGAAACAAGTTAATTGACTTTTTTTAATTAAAAAAAATTTATGTGAGAGTCAGGAGTT
95821  ACACTTTGTTCAATTAACTGAAAAAAATTAATTTTTTTTAAATACACTCTCAGTCCTCAA

95881  ACATGTGCAAGTTTGTTACATGGGTATATGTTGTGATACAAAGATTTTAGCTTCTAAGGA
95881  TGTACACGTTCAAACAATGTACCCATATACAACACTATGTTTCTAAAATCGAAGATTCCT

95941  TCTCATTGCCCAAGTAGTGAACGTAGTACACTATAGGTAGTATTTCTAAACACTTTTCCC
95941  AGAGTAACGGGTTCATCACTTGCATCATGTGATATCCATCATAAAGATTTGTGAAAAGGG

96001  CTCCCCCTCTTTTGGAATCCCTAGTGTTTATTGTTCCCATCTTTGTGTCTATATGTACCT
96001  GAGGGGGAGAAAACCTTAGGGATCACAAATAACAAGGGTAGAAACACAGATATACATGGA

96061  AATGCTTAGCTCTCACTTATAAGTGCGGACGTGTGATATTTTGGTTTTCTGTTTCTGTGT
96061  TTACGAATCGAGAGTGAATATTCACGCCTGCACACTATAAAACCAAAAGACAAAGACACA

96121  TAATTCATTTAGGATAGAGAGACTCCAGCTGCAGCTGTGTTGCTACAAATGACATGAGTT
96121  ATTAAGTAAATCCTATCTCTCTGAGGTCGACGTCGACACAACGATGTTTACTGTACTCAA
```

FIGURE 4 (cont'd)

```
96181  CATTCTTTTTTATGGCTGCATAGTATTCCATGGTTTATATGTATCACATTTTCTTTATTC
96181  GTAAGAAAAAATACCGACGTATCATAAGGTACCAAATATACATAGTGTAAAAGAAATAAG

96241  AGTCCACCATTGATGGTCACCTGGGTTGATTCCATGTCTTTGTTGTTGTGGACAGGGCTG
96241  TCAGGTGGTAACTACCAGTGGACCCAACTAAGGTACAGAAACAACAACACCTGTCCCGAC

96301  TGATAAACATATGAGTGCAGTTACCTTTTTTTGTATAACAGTTCATAATAAAATACTCTT
96301  ACTATTTGTATACTCACGTCAATGGAAAAAAACATATTGTCAAGTATTATTTTATGAGAA

96361  AACAGTAGCACAGAATTTAGATTGCTTTTCAGGCTAGAAGGGATAATTGAAACATAGAAA
96361  TTGTCATCGTGTCTTAAATCTAACGAAAAGTCCGATCTTCCCTATTAACTTTGTATCTTT

96421  AAAATAATGACTCCTAAGCATAGAAACAGACCCAATCATAAAGCTCTCTTCCTCTCTTTC
96421  TTTTATTACTGAGGATTCGTATCTTTGTCTGGGTTAGTATTTCGAGAGAAGGAGAGAAAG

96481  TGCCTAGTGTTAGATCTTATTCATTGTCTTAAAGCACTGTGACCATTTTCAGTGAAATCA
96481  ACGGATCACAATCTAGAATAAGTAACAGAATTTCGTGACACTGGTAAAAGTCACTTTAGT

96541  CAGTTTTATTCATGAAAAGACAATGAATTGGTGTAAGTAATCATTTGAGGATGAAGGTGA
96541  GTCAAAATAAGTACTTTTCTGTTACTTAACCACATTCATTAGTAAACTCCTACTTCCACT

96601  CTAGTCTGTGATCAACACCAGTACTTGAAGTACTTGTAAAACATGAATTGACCTTGGTTA
96601  GATCAGACACTAGTTGTGGTCATGAACTTCATGAACATTTTGTACTTAACTGGAACCAAT

96661  AATTGCAGAAAATAAAGTATTAGTATACATAAACACATTTTCTCTCAGCATTGAAAAGTT
96661  TTAACGTCTTTTATTTCATAATCATATGTATTTGTGTAAAAGAGAGTCGTAACTTTTCAA

96721  ATAGTCTATAATTTACCTGCTTGCTAAGAATAAAAATATTACTTAGAGGTTTATCAAGAC
96721  TATCAGATATTAAATGGACGAACGATTCTTATTTTTATAATGAATCTCCAAATAGTTCTG

96781  CTTTTTTCTACAAATGGATTTAATGTACAGAATTATTCCATTTTCTATCTGTGCTCTGTA
96781  GAAAAAAGATGTTTACCTAAATTACATGTCTTAATAAGGTAAAAGATAGACACGAGACAT

96841  TACTACCACCTAGTGGAATAAAATGCAAATTTACCTGTCAAAATCAGGAGAAGAAAATGA
96841  ATGATGGTGGATCACCTTATTTTACGTTTAAATGGACAGTTTTAGTCCTCTTCTTTTACT

96901  ACCAGGGGAAATTTATTTCAGGAAAAAATTTTCATAACACGTTATTGTTACTTCTTTTAG
96901  TGGTCCCCTTTAAATAAAGTCCTTTTTTAAAAGTATTGTGCAATAACAATGAAGAAAATC

96961  TATAATGTAGTTTGCATAAAATACATTAGTTTGGAGTGAAAAAAGTCCCAGCCATTTTTG
96961  ATATTACATCAAACGTATTTTATGTAATCAAACCTCACTTTTTTCAGGGTCGGTAAAAAC

97021  TGATAGCTATTAAAATGCTATTAGTTTGTCCACAAATAAGCTATTCTATTCAGAAACTTA
97021  ACTATCGATAATTTTACGATAATCAAACAGGTGTTTATTCGATAAGATAAGTCTTTGAAT

97081  GAAGCTAGGTAAATTATTGAGATGATGCCAGCATATTATATCTCTAAATTATGTTAAGAA
97081  CTTCGATCCATTTAATAACTCTACTACGGTCGTATAATATAGAGATTTAATACAATTCTT

97141  GCTTCTCTAAATTCCTCATCACTAGGGTGTATTAGTCCATTTTCACACTGCTATAAAGAA
97141  CGAAGAGATTTAAGGAGTAGTGATCCCACATAATCAGGTAAAAGTGTGACGATATTTCTT

97201  CTACCTGAGACTGGGTAATTTATGAAGAAAAGAGGTTTAATTGACTCACAGTTCTGCAGG
97201  GATGGACTCTGACCCATTAAATACTTCTTTTCTCCAAATTAACTGAGTGTCAAGACGTCC

97261  CTTAACAGGAAGCATGACTGGGAGACCTCAGGAAACATACAGTGGAAGAAGGTGAAAGGG
97261  GAATTGTCCTTCGTACTGACCCTCTGGAGTCCTTTGTATGTCACCTTCTTCCACTTTCCC

97321  AAGCAAGGACCTTCTCATGGCAGCAGGAGAGGGAGAGCACAGGGGAAGTGCCACACACTT
97321  TTCGTTCCTGGAAGAGTACCGTCGTCCTCTCCCTCTCGTGTCCCCTTCACGGTGTGTGAA

97381  TTAAACCATCAGATCTCAGGAGAACTCACTCACTCTCACAAGCACAGCAGGGAGGAAATT
97381  AATTTGGTAGTCTAGAGTCCTCTTGAGTGAGTGAGAGTGTTCGTGTCGTCCCTCCTTTAA

97441  CACCCCCACGATCCTATCATCTCACACTGGGCCCCTCCTCCAATTTGACATGAGATATGG
97441  GTGGGGGTGCTAGGATAGTAGAGTGTGACCCGGGGAGGAGGTTAAACTGTACTCTATACC
```

FIGURE 4 (cont'd)

```
97501  GCAGGGACACAAATCCAAACCGTATCATAGGATTTGTTAAAAAAATCTAACATGTTCTGC
97501  CGTCCCTGTGTTTAGGTTTGGCATAGTATCCTAAACAATTTTTTTAGATTGTACAAGACG

97561  TTCAAGGCACCCATGAAATGAAGGTGTCATTTCTTAACTGAGAGGTTAAATAAAAGTGCA
97561  AAGTTCCGTGGGTACTTTACTTCCACAGTAAAGAATTGACTCTCCAATTTATTTTCACGT

97621  GAGATAAGATGGTTAAGTTTTATCGTAGCTGGGAAGAAAAGCGGGTGTGTGTATTATGCT
97621  CTCTATTCTACCAATTCAAAATAGCATCGACCCTTCTTTTCGCCCACACACATAATACGA

97681  TAGGTTTTTCCTCCCAAACAATGAAACCCAAAAAGGCAAATGAAGCTTTAAGGCGCTTTA
97681  ATCCAAAAAGGAGGGTTTGTTACTTTGGGTTTTTCCGTTTACTTCGAAATTCCGCGAAAT

97741  TCCTTTACTGGTAGGATGGTCCTCCTTTAATCTTCAGTCTCTGTTAAAGAACAAACACAA
97741  AGGAAATGACCATCCTACCAGGAGGAAATTAGAAGTCAGAGACAATTTCTTGTTTGTGTT

97801  GCTAACAATCAAAACCTGTTGTTGCTTGCTGACGGATTTCTCTAACATAGTCTTCCATCA
97801  CGATTGTTAGTTTTGGACAACAACGAACGACTGCCTAAAGAGATTGTATCAGAAGGTAGT

97861  ACCATCATAGGTTTGTTGTGGCCACCCAGGCCCTTTGGTCGTCATGTTCCCATGGAATGC
97861  TGGTAGTATCCAAACAACACCGGTGGGTCCGGGAAACCAGCAGTACAAGGGTACCTTACG

97921  TAATGAAATTACCTTCCCATTTCCTAAGGTGACATGCCTTTCAGATGGCATTGGTGTGTG
97921  ATTACTTTAATGGAAGGGTAAAGGATTCCACTGTACGGAAAGTCTACCGTAACCACACAC

97981  CTTGTGTAAACACATTTTTTTGTGCTCATGAATATTTTACCTGCCTGATTTTGGTGATCT
97981  GAACACATTTGTGTAAAAAAACACGAGTACTTATAAAATGGACGGACTAAAACCACTAGA

98041  GAGTGTTGAGCAGTGGGTGAGGGATGGTGGGGCAATGGGTAGTAAGGGGAAGTTTAATTG
98041  CTCACAACTCGTCACCCACTCCCTACCACCCCGTTACCCATCATTCCCCTTCAAATTAAC

98101  AAGTTCAGATAAAAATATCCCTCCAGAAACATTACAAAGCCACAAAATCAGCTGAAGTTG
98101  TTCAAGTCTATTTTTATAGGGAGGTCTTTGTAATGTTTCGGTGTTTTAGTCGACTTCAAC

98161  AGTCTTCTATTGTAAAGATTCTCTTTCTTCTCCTGACAGATCCCCATTCCTATCCAGGTA
98161  TCAGAAGATAACATTTCTAAGAGAAAGAAGAGGACTGTCTAGGGGTAAGGATAGGTCCAT

98221  GACCAGGCTCTAGTGAAGATTGGACACAAGTGGTTTATATCCCGAAGCTCTTAACCTAAG
98221  CTGGTCCGAGATCACTTCTAACCTGTGTTCACCAAATATAGGGCTTCGAGAATTGGATTC

98281  AGTCCTGCCTTGTTTTTTTACTTCATGTGGTCTTATATTGCTGAGGATAGCCCAGTGATC
98281  TCAGGACGGAACAAAAAAAATGAAGTACACCAGAATATAACGACTCCTATCGGGTCACTAG

98341  ATTTCTACCTTAATATGTAGCCTTCAAAACGTTGGGTCTCTCAGTTTTCTTCCTTAATAA
98341  TAAAGATGGAATTATACATCGGAAGTTTTGCAACCCAGAGAGTCAAAAGAAGGAATTATT

98401  ACAATGAACCCTAATTATTTATAACTGCTCTTGTGATGGAGGAATTGGCATGATTTTGCA
98401  TGTTACTTGGGATTAATAAATATTGACGAGAACACTACCTCCTTAACCGTACTAAAACGT

98461  TTGTGGGCTTCTGGCTTTACAGTGTTTGTGGGAAAGATGGTCTTCTATCTTTGTTATTTA
98461  AACACCCGAAGACCGAAATGTCACAAACACCCTTTCTACCAGAAGATAGAAACAATAAAT

98521  TAAGTCATTCATTCCCTGTTTTTTTGAAGATATTGTTAAATAAAAAAACTTCAGCTGAAT
98521  ATTCAGTAAGTAAGGGACAAAAAAACTTCTATAACAATTTATTTTTTTGAAGTCGACTTA

98581  TAAATTTAAAAGTTTTTAATTGAGCAAAGAACCATTCTCAATTCGGGCAGCCTTCCCATC
98581  ATTTAAATTTTCAAAAATTAACTCGTTTCTTGGTAAGAGTTAAGCCCGTCGGAAGGGTAG

98641  CAGAGTAGTCCCTGTGACTTCAGTGCAGCCACATGGTGGAAGAGGATTTATGGACAGAGG
98641  GTCTCATCAGGGACACTGAAGTCACGTCGGTGTACCACCTTCTCCTAAATACCTGTCTCC

98701  AAGGAAAGTGACATACAGGAAACAGAAAAGAGATTCAGAAACAGCTGAATTGGTTATAAC
98701  TTCCTTTCACTGTATGTCCTTTGTCTTTTCTCTAAGTCTTTGTCGACTTAACCAATATTG

98761  TCAGCGTTTGCCTTATTTGAACATGGTTTGAACAGTTGGCCACGTTTGATTGGCCAAAAC
98761  AGTCGCAAACGGAATAAACTTGTACCAAACTTGTCAACCGGTGCAAACTAACCGGTTTTG
```

FIGURE 4 (cont'd)

```
 98821  TCAGTAATTGGCACAAGAGTAGGCTACCATCTGTTTATAATGTCCACTTAGGTTATAGTT
 98821  AGTCATTAACCGTGTTCTCATCCGATGGTAGACAAATATTACAGGTGAATCCAATATCAA

 98881  CATGATGTGCAAAGAAACCTTTAAGCTGAGCTTAAAATGTAATGAGGCAGCTGTAGGCTA
 98881  GTACTACACGTTTCTTTGGAAATTCGACTCGAATTTTACATTACTCCGTCGACATCCGAT

 98941  AACTTGATTTAACAGTACTTATCCCTCTGTTGAAATTGCACTCTTGCTACCTGCTAGTTG
 98941  TTGAACTAAATTGTCATGAATAGGGAGACAACTTTAACGTGAGAACGATGGACGATCAAC

 99001  ATGTGGGCAGGTGTGCTAATGACACATTCGCCTGCACCTACCACCTGCTGAAGACCCTGC
 99001  TACACCCGTCCACACGATTACTGTGTAAGCGGACGTGGATGGTGGACGACTTCTGGGACG

 99061  CATGTCCTGAAGCAGTCACTTAGGCTATGAAACAGAAGGATGAGTTAGAGAAGTAAATAT
 99061  GTACAGGACTTCGTCAGTGAATCCGATACTTTGTCTTCCTACTCAATCTCTTCATTTATA

 99121  TATATTTTACCCCCCAAAGCCTAAAACCATGTCTAGGTTCTCCCCTTTCTCAATTCCCAA
 99121  ATATAAAATGGGGGGTTTCGGATTTTGGTACAGATCCAAGAGGGGAAAGAGTTAAGGGTT

 99181  TCATGCAACCTTGAAAGGTGTTATATAGATAGTTGTTAGAAATATGTCTGTCCTCCAAAA
 99181  AGTACGTTGGAACTTTCCACAATATATCTATCAACAATCTTTATACAGACAGGAGGTTTT

 99241  CCCTACTAACACGCAAAAATGGACAAATCTAGCTTTTTTGCTAAAATCCTAACAACTCAG
 99241  GGGATGATTGTGCGTTTTTACCTGTTTAGATCGAAAAAACGATTTTAGGATTGTTGAGTC

 99301  AAAGCAAAGCAAAATGAAATGTATGTACTCCATTCTCAGGATTTCCTTCTTAAAAACCAC
 99301  TTTCGTTTCGTTTTACTTTACATACATGAGGTAAGAGTCCTAAAGGAAGAATTTTTGGTG

 99361  TAGGCTTTGTGTAACTCAATTTTAAAAGTATGGTAATAGCTCCTGTAATTCGATGGCTTC
 99361  ATCCGAAACACATTGAGTTAAAATTTTCATACCATTATCGAGGACATTAAGCTACCGAAG

 99421  TTCTGGGAGACAGATATAAATCTATTCTAGACTTTAAGACTGGTGTCCACATTTTAAGTT
 99421  AAGACCCTCTGTCTATATTTAGATAAGATCTGAAATTCTGACCACAGGTGTAAAATTCAA

 99481  ACAGTTGTATGTACTTCTACTTTGTTGCAGCCTTCTCTTGGAGTCAAGAAGCCTACCAAA
 99481  TGTCAACATACATGAAGATGAAACAACGTCGGAAGAGAACCTCAGTTCTTCGGATGGTTT

 99541  GCCCTGCTCTTTGTACTCTTGAGCCCAGTGGGACCTTATTTTTCAAGTGGAACCTTTAAT
 99541  CGGGACGAGAAACATGAGAACTCGGGTCACCCTGGAATAAAAAGTTCACCTTGGAAATTA

 99601  CCAGTGTCCCTGTGGGCCAATCCCAAGTATGTAAGAGCCTGGAAAGGTGGAACTGGGGAC
 99601  GGTCACAGGGACACCCGGTTAGGGTTCATACATTCTCGGACCTTTCCACCTTGACCCCTG

 99661  TGCAAGATGGGAGGGTAAGTAAATCTGTGTCTCTGTCTAAAGGAAAGACACCTTCATGAC
 99661  ACGTTCTACCCTCCCATTCATTTAGACACAGAGACAGATTTCCTTTCTGTGGAAGTACTG

 99721  CATTTAGTTAGTGCCATTTCACTTCAGTATGAACACATATAATACTTTTCAAACTTTAAT
 99721  GTAAATCAATCACGGTAAAGTGAAGTCATACTTGTGTATATTATGAAAAGTTTGAAATTA

 99781  GCTCATTGTGTTTCATGGTTAATTTTCAGTTTAATCTTTCAGCAAGCCAGCTGTAGTAGC
 99781  CGAGTAACACAAAGTACCAATTAAAAGTCAAATTAGAAAGTCGTTCGGTCGACATCATCG

 99841  TTTTCCTATATAAACGCAGTATGAATTCAGTGCCCTCGCTTTTAGATGTGCTAGTTTAAA
 99841  AAAAGGATATATTTGCGTCATACTTAAGTCACGGGAGCGAAAATCTACACGATCAAATTT

 99901  GGATAGCCCATGGCCCTTTGTAAGGGAGAAACACTGGATATGATATCTAGGGAAGAACTG
 99901  CCTATCGGGTACCGGGAAACATTCCCTCTTTGTGACCTATACTATAGATCCCTTCTTGAC

 99961  AAGTAAGAAACTACAACAAGCATCACATCGGTGCTTTGGATAAACTTCCCCTGGGCGCTG
 99961  TTCATTCTTTGATGTTGTTCGTAGTGTAGCCACGAAACCTATTTGAAGGGGACCCGCGAC

100021  GTCTCTGGCTGCAGGCTGACGGCTGGCATTAAAAAGCATCTGGTCCTGGGACCCAGGCTC
100021  CAGAGACCGACGTCCGACTGCCGACCGTAATTTTTCGTAGACCAGGACCCTGGGTCCGAG

100081  TCAGTTGGCTCCAGGGAGTCATGTCTGGGGAAGGTGCCCGCTTCACTGAAGTGCTGTCTT
100081  AGTCAACCGAGGTCCCTCAGTACAGACCCCTTCCACGGGCGAAGTGACTTCACGACAGAA
```

FIGURE 4 (cont'd)

```
100141  CATTGCCACAGATTGTGGCCACACCTGCTGCCAACTGGCTGCGGGACAGCCCTGCTTTGG
100141  GTAACGGTGTCTAACACCGGTGTGGACGACGGTTGACCGACGCCCTGTCGGGACGAAACC

100201  CTCTGTCAGGCCAGTCTCATAGGCCATATGTAGCTGGCAGGAAATAAAAGAGCATTTGAA
100201  GAGACAGTCCGGTCAGAGTATCCGGTATACATCGACCGTCCTTTATTTTCTCGTAAACTT

100261  AAGTGTTGGACGATTTCAGGGGTGGACTTGGCCATGAGTTGGGAATGCCCCTATGGGTGA
100261  TTCACAACCTGCTAAAGTCCCCACCTGAACCGGTACTCAACCCTTACGGGGATACCCACT

100321  ACGAGAATAACTGGTTTAAGCAGCATGGGTTCTTGCTTTGCATGTATCAAATGGTGCTGT
100321  TGCTCTTATTGACCAAATTCGTCGTACCCAAGAACGAAACGTACATAGTTTACCACGACA

100381  TTGCATACTCAGGGCAGCAGCTGTCCTAGAAGGTGACTCCATGGCAGTCTTGTGGGCCCT
100381  AACGTATGAGTCCCGTCGTCGACAGGATCTTCCACTGAGGTACCGTCAGAACACCCGGGA

100441  GGTAGGACATTTAGGTCCAGGAGGCAGGGGAGAGTCACATGACGTGCTGCTGGGGTATTA
100441  CCATCCTGTAAATCCAGGTCCTCCGTCCCCTCTCAGTGTACTGCACGACGACCCCATAAT

100501  GAGGCAATGACCACTGGCTTTGACCCCTCTGTAGTCTTGCTGTATTTCCTTTTTCTTTTC
100501  CTCCGTTACTGGTGACCGAAACTGGGGAGACATCAGAACGACATAAAGGAAAAAGAAAAG

100561  TCTCCTCACCTACCCAAATTAAAAAAGAACGGTTATAAAAGTCATCCCCACTGGCAACTT
100561  AGAGGAGTGGATGGGTTTAATTTTTTCTTGCCAATATTTTCAGTAGGGGTGACCGTTGAA

100621  TAGCAACTGCATTCCCTCAATGGGTAGAGAAATCCAAATGGAGTGCATTAAAACTAGGAA
100621  ATCGTTGACGTAAGGGAGTTACCCATCTCTTTAGGTTTACCTCACGTAATTTTGATCCTT

100681  GTCTATTCTCCTTCTGGCTTAGAGCACACCTTCCTGTGTCCTAGGCTAGCCAGGTTGAAA
100681  CAGATAAGAGGAAGACCGAATCTCGTGTGGAAGGACACAGGATCCGATCGGTCCAACTTT

100741  ACTCCTTTTTTTGACTCTTGATTTCTTTTTTTACTGGTTCAAACTTCTTCGCAGAGCATG
100741  TGAGGAAAAAAACTGAGAACTAAAGAAAAAAATGACCAAGTTTGAAGAAGCGTCTCGTAC

100801  TGCCTATGTACCACTGCATACCATATATATACCATATATAGAGAGTAAATAAACAGTCTT
100801  ACGGATACATGGTGACGTATGGTATATATATGGTATATATCTCTCATTTATTTGTCAGAA

100861  GATGAAAAGTAATAAGGAAAAAGTTTCCCTGTTTAAAATATTGCTTTCTAGCTCCCCGAC
100861  CTACTTTTCATTATTCCTTTTTCAAAGGGACAAATTTTATAACGAAAGATCGAGGGGCTG

100921  CTGTTTAGGAAAAAGATAACCTGGGGACACAGAAATTGCAGGGCAATCTTTATTAACCGG
100921  GACAAATCCTTTTTCTATTGGACCCCTGTGTCTTTAACGTCCCGTTAGAAATAATTGGCC

100981  TATTTGGGCAGGTTGTAATTTATTTTGTAAACAATGATTTAATGCACTATTGCAGAATAC
100981  ATAAACCCGTCCAACATTAAATAAAACATTTGTTACTAAATTACGTGATAACGTCTTATG

101041  ATGATCACTGTAGGCCAGTTAAAAATTAATGTTTTCAGTACCACACAGAGTGAACCCTAA
101041  TACTAGTGACATCCGGTCAATTTTTAATTACAAAAGTCATGGTGTGTCTCACTTGGGATT

101101  TGTAAACTGTGGACTTGTGGACCACACTAATGCAAGATGATAATAAGAAGAAAAACTGAG
101101  ACATTTGACACCTGAACACCTGGTGTGATTACGTTCTACTATTATTCTTCTTTTTGACTC

101161  AGGGTGTTCGGGAAGGGGTATAATAGGGGAATTCTGTTACCTTTTGCTCAATTTCTCTGT
101161  TCCCACAAGCCCTTCCCCATATTATCCCCTTAAGACAATGGAAAACGAGTTAAAGAGACA

101221  GAATCTAAAAGTACTCTAAAATAATAAAATCTATTATAAAAATATTTGATTCATTAGGCT
101221  CTTAGATTTTCATGAGATTTTATTATTTTAGATAATATTTTTATAAACTAAGTAATCCGA

101281  CAGGTGGACTATGATCCATGGAAATTGACTTGAAATCTCTTAGGAAAGGAAGAGACCGGA
101281  GTCCACCTGATACTAGGTACCTTTAACTGAACTTTAGAGAATCCTTTCCTTCTCTGGCCT

101341  GAAGAAAGTATTGAACTCATAAATTAAAAATGGCAAGAGGGAAGGTGGTAACCAAAAGGA
101341  CTTCTTTCATAACTTGAGTATTTAATTTTTACCGTTCTCCCTTCCACCATTGGTTTTCCT

101401  TGGGCATTTAACCTATATGCCTATCACTTCCAAAAAGTTCTAGGGCCGTCTCCTCCTATG
101401  ACCCGTAAATTGGATATACGGATAGTGAAGGTTTTTCAAGATCCCGGCAGAGGAGGATAC

101461  CTTCTGACCTGGAGAGGTACAAGGGGAACAAAATATTCCGCATTTAATATTATCCTGTTG
101461  GAAGACTGGACCTCTCCATGTTCCCCTTGTTTTATAAGGCGTAAATTATAATAGGACAAC
```

FIGURE 4 (cont'd)

```
101521  AGTTGGCAAAGACATTTGGAGTCAACTTTAGTTTTGTGTATTATTACATCTGTAAACCTA
101521  TCAACCGTTTCTGTAAACCTCAGTTGAAATCAAAACACATAATAATGTAGACATTTGGAT

101581  ATTAAGTCAACAGTTACCCTTGACACCTGTAGTCAACTGGTATTATTAAATAACTATTGA
101581  TAATTCAGTTGTCAATGGGAACTGTGGACATCAGTTGACCATAATAATTTATTGATAACT

101641  GTCTCATAAAGTCAAGAACTTTCTGCTTTCTATCTAAGTGAATCTAGGGCAGAGACTGAG
101641  CAGAGTATTTCAGTTCTTGAAAGACGAAAGATAGATTCACTTAGATCCCGTCTCTGACTC

101701  CATCAGGATTCTTTAGCATGCAGCCAAGCTTGAGAAGCACTGAGCTAGAGAAAGGTCTTG
101701  GTAGTCCTAAGAAATCGTACGTCGGTTCGAACTCTTCGTGACTCGATCTCTTTCCAGAAC

101761  CACTGTAGAAACAGCTTTAGTTCCAGGTGATTCCTTGGTTAATTTGGGAGGAATAAGGCT
101761  GTGACATCTTTGTCGAAATCAAGGTCCACTAAGGAACCAATTAAACCCTCCTTATTCCGA

101821  TTCTGTCACTCTCTCTCTCTTTCTCTCTTTTTCTCTCTCTCTTGCACATGCACACGTGCA
101821  AAGACAGTGAGAGAGAGAGAAAGAGAGAAAAAAGAGAGAGAGAACGTGTACGTGTGCACGT

101881  CACACACACACAAATTATCTCTCTTAGTTTTTTCTTTTTTCCTTGCATTCTGATTCTGCA
101881  GTGTGTGTGTGTTTAATAGAGAGAATCAAAAAAGAAAAAAGGAACGTAAGACTAAGACGT

101941  TTTAAGAGTACAAGCTTTGAAGTCAGAAAGACCTGGCTTAAAAACCTAACTGTCATAATG
101941  AAATTCTCATGTTCGAAACTTCAGTCTTTCTGGACCGAATTTTTGGATTGACAGTATTAC

102001  GTGCGTAAATGTCGTTATAGATTTGTCCAAACCCATAGAATGTACACCAGGAATGAACAC
102001  CACGCATTTACAGCAATATCTAAACAGGTTTGGGTATCTTACATGTGGTCCTTACTTGTG

102061  TAATGAAAACTCTAGACTTCAGGTGATAATGGTATGTCAATGTGGTTCATTAATTTCAGC
102061  ATTACTTTTGAGATCTGAAGTCCACTATTACCATACAGTTACACCAAGTAATTAAAGTCG

102121  AAATGTACCACTCTGGTGTGAGGCATTGGTAATGGTAGAAACTATTCATTTGGATGGGGA
102121  TTTACATGGTGAGACCACACTCCGTAACCATTACCATCTTTGATAAGTAAACCTACCCCT

102181  GAAAGGGATATGGGAAATCTTTGTACTTGCCTTTCAGTTTTGGTGGGAACCTAAAACTGC
102181  CTTTCCCTATACCCTTTAGAAACATGAACGGAAAGTCAAAACCACCCTTGGATTTTGACG

102241  TCTGAAAAAAACCTGAAGTCTTTAAAAACAAACCAACCTAATTGCCTTGGGCATGTTACT
102241  AGACTTTTTTTGGACTTCAGAAATTTTTGTTTGGTTGGATTAACGGAACCCGTACAATGA

102301  TAACTGCTCTATGTATCAGTTTCCTCATTTTTAAAGTTGAGGCAATAATACCTTGCACAG
102301  ATTGACGAGATACATAGTCAAAGGAGTAAAAATTTCAACTCCGTTATTATGGAACGTGTC

102361  ATTAAGTGAGATAATAGAAAGTGCCTAGGATAGTACTTGGCATATAGAAAGCTCTCAGCA
102361  TAATTCACTCTATTATCTTTCACGGATCCTATCATGAACCGTATATCTTTCGAGAGTCGT

102421  AATGCAGTATGATAACCTAATGGTTAAAAGCTTGGGTTTAGGAGTCAAATGTCTGGGAGT
102421  TTACGTCATACTATTGGATTACCAATTTTCGAACCCAAATCCTCAGTTTACAGACCCTCA

102481  TACACACTTAACTATCTGTGAGATCTTCTACAACTGAATCCTCTCTCCAAGTCTTCCAAT
102481  ATGTGTGAATTGATAGACACTCTAGAAGATGTTGACTTAGGAGAGAGGTTCAGAAGGTTA

102541  GAACAATAAAGTCTAGGGTTGTTTTTAGATGACTCAATGAAATAATGTAGTTAAGTACTT
102541  CTTGTTATTTCAGATCCCAACAAAAATCTACTGAGTTACTTTATTACATCAATTCATGAA

102601  AGCACCCAGTCTGGCAGGTATGGGTTGCATATCTCTTATCTGAAATGCTTGGGACCAGAT
102601  TCGTGGGTCAGACCGTCCATACCCAACGTATAGAGAATAGACTTTACGAACCCTGGTCTA

102661  GTGTTTTGGATTTCAAATTTTGGAGTATTTCTCTTTGGGCATCCCTAATCCAAAAAATTC
102661  CACAAAACCTAAAGTTTAAAACCTCATAAAGAGAAACCCGTAGGGATTAGGTTTTTTAAG

102721  GAAATCCAAATTGAGCATCATGTTGGCACCGAAAAAGTTTTAGATTTTAGAGCATTTCAA
102721  CTTTAGGTTTAACTCGTAGTACAACCGTGGCTTTTTCAAAATCTAAAATCTCGTAAAGTT

102781  AGTTTTAAATTAGGGATCCTCAACCTGTAGTAAACTCTCACAGTCACTGATGTTGTGAGG
102781  TCAAAATTTAATCCCTAGGAGTTGGACATCATTTGAGAGTGTCAGTGACTACAACACTCC
```

FIGURE 4 (cont'd)

```
102841  TTTTCTCACTTTCTCATTTCATTTCCTGGACTCTCTTTATCTTAGAAGAAATCAGTTAAA
102841  AAAAGAGTGAAAGAGTAAAGTAAAGGACCTGAGAGAAATAGAATCTTCTTTAGTCAATTT

102901  TTTTACATAATAGAATTTATTCTTTTGGAGGGATGTATTTAATATGTCTATTCTCCCAGA
102901  AAAATGTATTATCTTAAATAAGAAAACCTCCCTACATAAATTATACAGATAAGAGGGTCT

102961  TAACCTATTTAAAACCATATTGAAATGTTTCTATGTCAACAGCTTTCTTACTTTTTCATT
102961  ATTGGATAAATTTTGGTATAACTTTACAAAGATACAGTTGTCGAAAGAATGAAAAAGTAA

103021  TTTCTGTTTCCTCCAAGAACATGTTTCTCTTATGAATATTGGCAGCACATGGGCGGGAAC
103021  AAAGACAAAGGAGGTTCTTGTACAAAGAGAATACTTATAACCGTCGTGTACCCGCCCTTG

103081  CCAACTATATTACCATGGTATCAGTTATCCTGTGGACATTAATAATACAGCTTACACCTG
103081  GGTTGATATAATGGTACCATAGTCAATAGGACACCTGTAATTATTATGTCGAATGTGGAC

103141  TGAGAGAACTCTGGGAAAATTAAAACTTGTTTGGGGGAATAAAGACCATCTACAGAATAT
103141  ACTCTCTTGAGACCCTTTTAATTTTGAACAAACCCCCTTATTTCTGGTAGATGTCTTATA

103201  AATAATACATCAATCAATCAATTAAAACTTAAACATAAAAAACCCTTTTCTTTTCAGTGC
103201  TTATTATGTAGTTAGTTAGTTAATTTTGAATTTGTATTTTTTGGGAAAAGAAAAGTCACG

103261  ATTGAAATAGTCACTGTCTTTAATCCTCTTTCCTTGCTCAGTGCTTCATGAAGCTGAGGA
103261  TAACTTTATCAGTGACAGAAATTAGGAGAAAGGAACGAGTCACGAAGTACTTCGACTCCT

103321  GCCATGGTTTTAATTCAGCTACACTTTCAGAAATGTGTATGCATTTGTATTTAGAAGCTT
103321  CGGTACCAAAATTAAGTCGATGTGAAAGTCTTTACACATACGTAAACATAAATCTTCGAA

103381  CTTTGAAAGGCTAAATATTCATCCTGTAAGAATTCTCTATGTGTTTTTTAATACAGTCTA
103381  GAAACTTTCCGATTTATAAGTAGGACATTCTTAAGAGATACACAAAAAATTATGTCAGAT

103441  AAGGGCATTTTGTATATGAAAAAATAATTTATTTTCAGGTAGGGTAAAACTTCACTGCTT
103441  TTCCCGTAAAACATATACTTTTTTATTAAATAAAAGTCCATCCCATTTTGAAGTGACGAA

103501  CACTACTTCGTAGTTGTGTCTACGAATGAGAAAAAATTCAAATAAAATAGTTATTATTAT
103501  GTGATGAAGCATCAACACAGATGCTTACTCTTTTTTAAGTTTATTTTATCAATAATAATA

103561  TAGTTTTTAGAGATGCAGTCTTGCTGTGTTGCCCAGGCTGGTCTGGGACTTCTGGCCTCA
103561  ATCAAAAATCTCTACGTCAGAACGACACAACGGGTCCGACCAGACCCTGAAGACCGGAGT

103621  AGCCATCCTCCTGCCTCGGTGTCCTGAGTCACTGGGATTACAGGCATGAGCCACCACACC
103621  TCGGTAGGAGGACGGAGCCACAGGACTCAGTGACCCTAATGTCCGTACTCGGTGGTGTGG

103681  CCACTGAAATAGTTATTTTTTAAAAAACTTAATTTCATTAGGATGTAACAGTCTCATGTA
103681  GGTGACTTTATCAATAAAAAATTTTTTGAATTAAAGTAATCCTACATTGTCAGAGTACAT

103741  ACTCATCCTTTGGCTTTTATTGTTATTATTTATTTTTTCCAAGCAACCCAATGAGATATA
103741  TGAGTAGGAAACCGAAAATAACAATAATAAATAAAAAAGGTTCGTTGGGTTACTCTATAT

103801  TAAAGCAGATATTTTTATCTGCCAGATAATTAAATGAGGGAAAAATAAGGACTAACTTGC
103801  ATTTCGTCTATAAAAATAGACGGTCTATTAATTTACTCCCTTTTTATTCCTGATTGAACG

103861  TTAAGTTCACATATACCAAATAGTGGAAATGCAAGAGCTGTCCTCAGGTCTGTTAACAAC
103861  AATTCAAGTGTATATGGTTTATCACCTTTACGTTCTCGACAGGAGTCCAGACAATTGTTG

103921  AAACCTGTGGTCTTTCTACCAGACGTCCCCTGCAATCTGGTTTATACTCATTATTAAAAA
103921  TTTGGACACCAGAAAGATGGTCTGCAGGGGACGTTAGACCAAATATGAGTAATAATTTTT

103981  GCATTTGTAATTTAACTCAGTTATCTATTTACTATCAGTTATGAATAAGGATTATTTGAC
103981  CGTAAACATTAAATTGAGTCAATAGATAAATGATAGTCAATACTTATTCCTAATAAACTG

104041  TTTTTATGGTCATTAAACAATAGCTTCCTATCTGTTGTCTTTGTCTTTTTAATTTTTATT
104041  AAAAATACCAGTAATTTGTTATCGAAGGATAGACAACAGAAACAGAAAAATTAAAAATAA

104101  TATTTTATTGTTTTTTTAATAGCTTCCTATCTGTATACCAGACTTGCTTCATCTCCACAT
104101  ATAAAATAACAAAAAAATTATCGAAGGATAGACATATGGTCTGAACGAAGTAGAGGTGTA

104161  TGACCAAGAATTTATATTTTGTCATCTGTATTTAGAAGCATATGTTTTAAGCGTCTTTTA
104161  ACTGGTTCTTAAATATAAAACAGTAGACATAAATCTTCGTATACAAAATTCGCAGAAAAT
```

FIGURE 4 (cont'd)

```
104221  ATGGCTCTTTTTTTTTTTTTTGCTGGGAGGGGACAGTCTCGCTCTGTCTCCTAGGCTGGAG
104221  TACCGAGAAAAAAAAAAAAAACGACCCTCCCCTGTCAGAGCGAGACAGAGGATCCGACCTC

104281  TGCAGTGGCACCATCTCAGCTCACTGCAACCTCCGCCTCCCTGGTGTGAGTGGTGCTCCT
104281  ACGTCACCGTGGTAGAGTCGAGTGACGTTGGAGGCGGAGGGACCACACTCACCACGAGGA

104341  GCCTCAGCCTCCCGAGTAGGTGGGATTACAGGCGTGGTCCACCAAACCCAGCTAGTTTTT
104341  CGGAGTCGGAGGGCTCATCCACCCTAATGTCCGCACCAGGTGGTTTGGGTCGATCAAAAA

104401  TATATTTTTGGTAGAGACAGGGTTTCACCATGTTAGCCAGGCTGGTCTGGAACTCCTGAC
104401  ATATAAAAACCATCTCTGTCCCAAAGTGGTACAATCGGTCCGACCAGACCTTGAGGACTG

104461  CTCAAGTAACTTGCCCACCTCAGCCTCCCAAAGTGCTGAGATTACAGGTATGAGCCACTG
104461  GAGTTCATTGAACGGGTGGAGTCGGAGGGTTTCACGACTCTAATGTCCATACTCGGTGAC

104521  TGCCTGGCCCTTTTAAGGACTCTTTTTTTTTTTTTTTTTTTTTTGAGACAGTCTCACTC
104521  ACGGACCGGGAAAATTCCTGAGAAAAAAAAAAAAAAAAAAAAAAACTCTGTCAGAGTGAG

104581  TGTCGCTCAGACTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCACCTCCCGG
104581  ACAGCGAGTCTGACCTCACGTCACCGCGCTAGAGCCGAGTGACGTTCGAGGTGGAGGGCC

104641  GTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCAGCCACC
104641  CAAGTGCGGTAAGAGGACGGAGTCGGAGGGCTCATCGACCCTGATGTCCGCGGTCGGTGG

104701  ACGCCCGGATAATTTTTAGTATTTTTAGTAGAGACGAGGTTTCACCGTGTTAGCCAGGAT
104701  TGCGGGCCTATTAAAAATCATAAAAATCATCTCTGCTCCAAAGTGGCACAATCGGTCCTA

104761  GGTCTCAATCTCCTGACCTCATGATCCGGCCGTCTCGCCTCCCAAAGTGCTGGGATTACA
104761  CCAGAGTTAGAGGACTGGAGTACTAGGCCGGCAGAGCGGAGGGTTTCACGACCCTAATGT

104821  GGCGTGAGCCACCGCGCCTGGCCCTTTTAAGGACTCTTAACATTACCTTGGTGTGAGTCA
104821  CCGCACTCGGTGGCGCGGACCGGGAAAATTCCTGAGAATTGTAATGGAACCACACTCAGT

104881  ACCTCATCGTGCTCCAATGGCAATCACCATAGTATGATGTCTATTTTCCACAGCTTACTG
104881  TGGAGTAGCACGAGGTTACCGTTAGTGGTATCATACTACAGATAAAAGGTGTCGAATGAC

104941  GAATATTTCAGCTATTAGTCAATATTCTTTCAGTTGCAAGTCAGAAAACCAACTTGAAAT
104941  CTTATAAAGTCGATAATCAGTTATAAGAAAGTCAACGTTCAGTCTTTTGGTTGAACTTTA

105001  TATTATTTATTTATTTATTTATTTTTTTGAGACGGAATTTTGCTCTTGTTGCCCAGGCTG
105001  ATAATAAATAAATAAATAAATAAAAAAACTCTGCCTTAAAACGAGAACAACGGGTCCGAC

105061  GAGTGCCATGGCACAATCTCGGCTCACTGCAACCTCCGCCTCCCGAGTTCAAGCAATTCT
105061  CTCACGGTACCGTGTTAGAGCCGAGTGACGTTGGAGGCGGAGGGCTCAAGTTCGTTAAGA

105121  CCTGCCTCAGCCTCCAGAGTAGCTGGGATTACAGGCATGGGCCACCACTCCCGGCTAATT
105121  GGACGGAGTCGGAGGTCTCATCGACCCTAATGTCCGTACCCGGTGGTGAGGGCCGATTAA

105181  TTTTTGTTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGGATGGTCTCAAACTC
105181  AAAAACAATAAAAATCATCTCTGCCCCAAAGAGGTACAACCAGTCCTACCAGAGTTTGAG

105241  CCAACCTCAGGTGATCTGCCCACCTCAGCCTCCCAAAGTGCTGGGATTACAGGCATGAGC
105241  GGTTGGAGTCCACTAGACGGGTGGAGTCGGAGGGTTTCACGACCCTAATGTCCGTACTCG

105301  CACCGTGCCCGGCCTGGAAATGATCTTAAGCAGACAAGGGCTTATGAAACTAGGATGCTA
105301  GTGGCACGGGCCGGACCTTTACTAGAATTCGTCTGTTCCCGAATACTTTGATCCTACGAT

105361  GGGGCTAAAGCAATTATTTCCAAGTTTCTTGTGTCCTTTGCACATCCCCCTTTTTCTTCC
105361  CCCCGATTTCGTTAATAAAGGTTCAAAGAACACAGGAAACGTGTAGGGGGAAAAAGAAGG

105421  CTTCTTCCTCTGCCTCTGGTCAGAGCTTCTGGCTCCATGTTGGCCTAGGGCTCTCCCACT
105421  GAAGAAGGAGACGGAGACCAGTCTCGAAGACCGAGGTACAACCGGATCCCGAGAGGGTGA

105481  TCAGCCGCTTTCTTCTGTGAAGTGTGGTGATCATGGACACAGCCGACCCCAGTTTTACAG
105481  AGTCGGCGAAAGAAGACACTTCACACCACTAGTACCTGTGTCGGCTGGGGTCAAAATGTC
```

FIGURE 4 (cont'd)

```
105541  CATACCAAGTAAGCAACCCCAGAAGGGAGAAAATGCCTGTATTTCAGGTTATAGACATTA
105541  GTATGGTTCATTCGTTGGGGTCTTCCCTCTTTTACGGACATAAAGTCCAATATCTGTAAT

105601  AAATGAGTGAATGCACCCAAAAAAGGAATCATGTGACTGTCTTAGGTTATATGCCCACCC
105601  TTTACTCACTTACGTGGGTTTTTTCCTTAGTACACTGACAGAATCCAATATACGGGTGGG

105661  TTTAACTCAGCTGCAGCAGGGCAAGATGCTATATTGTCATCAGTTGTGGTAGGGCAAGAT
105661  AAATTGAGTCGACGTCGTCCCGTTCTACGATATAACAGTAGTCAACACCATCCCGTTCTA

105721  GCTATATTATCAGTCCAAGCAGAATGACATAGCTCTATAGAAAGACTATGTTTTTATAAG
105721  CGATATAATAGTCAGGTTCGTCTTACTGTATCGAGATATCTTTCTGATACAAAAATATTC

105781  AGTATGGAATATAGTGCAGATGAAAACAGTAGATATACACTAACATGTGGCTCTAATGTT
105781  TCATACCTTATATCACGTCTACTTTTGTCATCTATATGTGATTGTACACCGAGATTACAA

105841  TAATTTTTATTTTTTTTGAGACAGAGTCTCACTGTGTCACCCAGGCTGGAGTGCAGTGGT
105841  ATTAAAAATAAAAAAAAACTCTGTCTCAGAGTGACACAGTGGGTCCGACCTCACGTCACCA

105901  GCTATCTTGGCTCACTGCAACTTTTGCCTCCCGAGTTAAGATGATTCTCCTCCCTCAGCC
105901  CGATAGAACCGAGTGACGTTGAAAACGGAGGGCTCAATTCTACTAAGAGGAGGGAGTCGG

105961  TCCCTAGCAGCTGAGATTATAGGCACCCGCCACCACGCCCAGCTAATTCTTGTATTTTCA
105961  AGGGATCGTCGACTCTAATATCCGTGGGCGGTGGTGCGGGTCGATTAAGAACATAAAAGT

106021  GTAGAGACATGGTTTCACCATGTTGGCCGGGCTGGTCTCGAACTCCTGACCTCAGTGATC
106021  CATCTCTGTACCAAAGTGGTACAACCGGCCCGACCAGAGCTTGAGGACTGGAGTCACTAG

106081  CGCCCATCTTAGGCTCCTAAAGTGCTAGGATTACAGACATAAGCCACCGCGCCTGGCCTC
106081  GCGGGTAGAATCCGAGGATTTCACGATCCTAATGTCTGTATTCGGTGGCGCGGACCGGAG

106141  ATGCTCAACTTTTATTTAGACAACAGTACAAATTGTAATTTTACCCAGATGTAAAAGAGT
106141  TACGAGTTGAAAATAAATCTGTTGTCATGTTTAACATTAAAATGGGTCTACATTTTCTCA

106201  CTAAGAGACCTAAATAGATATTTTGTGGTAGATTACATTAATGTAATGTAATGTAATGTA
106201  GATTCTCTGGATTTATCTATAAAACACCATCTAATGTAATTACATTACATTACATTACAT

106261  GGAGTTTTACATTACTACATTTGTAGTTGTTTTAGTTGTTTTTCATTGTTTGTTTCTGTG
106261  CCTCAAAATGTAATGATGTAAACATCAACAAAATCAACAAAAAGTAACAAACAAAGACAC

106321  TATTTCTGGGGATTTCATGGTGCAACAGGAGGACTAGAGTTTATTCCCGATCCCTGACTG
106321  ATAAAGACCCCTAAAGTACCACGTTGTCCTCCTGATCTCAAATAAGGGCTAGGGACTGAC

106381  TCTTACGGCAGTACTTCTCAGACCTACTGAGTTATACAACCAGGGGTGTAGGGAGGCAAG
106381  AGAATGCCGTCATGAAGAGTCTGGATGACTCAATATGTTGGTCCCCACATCCCTCCGTTC

106441  CAGTCATCACATTTATTTGAGGACTCTCATGTTTTGTCTAAAATATTCATCCTTATTTTA
106441  GTCAGTAGTGTAAATAAACTCCTGAGAGTACAAAACAGATTTTATAAGTAGGAATAAAAT

106501  TAATATACTCCATAATTTTCCCATATAAGTTTAGATAAAAGATGATACCTTAAAAGTATT
106501  ATTATATGAGGTATTAAAAGGGTATATTCAAATCTATTTTCTACTATGGAATTTTCATAA

106561  ATAACACTAAATTGTCATTCCGGGAAAATGTGCTGTGTTCCTTTCATTTTAAATAATCTG
106561  TATTGTGATTTAACAGTAAGGCCCTTTTACACGACACAAGGAAAGTAAAATTTATTAGAC

106621  AGATGAATGGTGCAGGCTGAGAAAAATAGGCTTTTGAAATACTTTTTGGTCAACATCGAG
106621  TCTACTTACCACGTCCGACTCTTTTTATCCGAAAACTTTATGAAAAACCAGTTGTAGCTC

106681  ACTTTTAACATTTGTCCTATCTTTTAGGGGAATTTAGACATTGGGCATTATGTATTCCTT
106681  TGAAAATTGTAAACAGGATAGAAAATCCCCTTAAATCTGTAACCCGTAATACATAAGGAA

106741  GGATACATTAAAAGGTGTTAGATAAGAGCGCTCACACTTGCTGCTCTTCTATTGCAACTT
106741  CCTATGTAATTTTCCACAATCTATTCTCGCGAGTGTGAACGACGAGAAGATAACGTTGAA

106801  GACAGAAGTTACTTCTCTAAACCTTAGTTTCTCATCTGAACAATGGAAATAATGATAGTG
106801  CTGTCTTCAATGAAGAGATTTGGAATCAAAGAGTAGACTTGTTACCTTTATTACTATCAC

106861  TCTATCTCATAGTATTGGTGTGAGATTTAAATGAAATAATTCATGTAAAATTGCTTGGTG
106861  AGATAGAGTATCATAACCACACTCTAAATTTACTTTATTAAGTACATTTTAACGAACCAC
```

FIGURE 4 (cont'd)

```
106921  TGTGGAAACATAGTAACTTTTAACTGTTCTCATTAACTTGGCTATTTACCTGTGTTCTTC
106921  ACACCTTTGTATCATTGAAAATTGACAAGAGTAATTGAACCGATAAATGGACACAAGAAG

106981  TATCATTACATTTCAAGTTAAAATTTCAAGGTTTCACGTTGTTCTCTCTCCAATGAATGC
106981  ATAGTAATGTAAAGTTCAATTTTAAAGTTCCAAAGTGCAACAAGAGAGAGGTTACTTACG

107041  AGTTAAATCACTTAGATTTTCTAATGAGTTTTCCTTTAATGTTATGCAGTTGACTAGAGA
107041  TCAATTTAGTGAATCTAAAAGATTACTCAAAAGGAAATTACAATACGTCAACTGATCTCT

107101  GTGTATATATATTAAGAACATAATACCTTCACTATGTCCACATGTGAATTAAGAGCTTGA
107101  CACATATATATAATTCTTGTATTATGGAAGTGATACAGGTGTACACTTAATTCTCGAACT

107161  GGCTATCAGAGGGTAGAAGATGAGGGAGTGGGGATGAAGTATGTTTGTAACCTGTGTATT
107161  CCGATAGTCTCCCATCTTCTACTCCCTCACCCCTACTTCATACAAACATTGGACACATAA

107221  TTCTCTTTCTTAAAGGAATTACGGCTCATCTCTTTTTGCCCAATGTGAAGCAGTAGATAA
107221  AAGAGAAAGAATTTCCTTAATGCCGAGTAGAGAAAAACGGGTTACACTTCGTCATCTATT

107281  TGGGTGTCAGCAGGTCCTGTGGCTCTATGGAGAGGACCATCAGATCACTGAAGTGGGAAC
107281  ACCCACAGTCGTCCAGGACACCGAGATACCTCTCCTGGTAGTCTAGTGACTTCACCCTTG

107341  TATGAATCTTTTTCTTTACTGGATAAATGAAGATGGAGGTAATCCACTCAGATTTCACTG
107341  ATACTTAGAAAAAGAAATGACCTATTTACTTCTACCTCCATTAGGTGAGTCTAAAGTGAC

107401  TGTGCATTTCTCAGTCATTTTCAACAGGACCAGAAAGACAAGAGCTAGAAACTTAGTACA
107401  ACACGTAAAGAGTCAGTAAAAGTTGTCCTGGTCTTTCTGTTCTCGATCTTTGAATCATGT

107461  AGGGAGTAAGTAGCACAGTTCCAAAGGCTAAGTAAGGACTACTGTGCTTTTGCTGCCCAA
107461  TCCCTCATTCATCGTGTCAAGGTTTCCGATTCATTCCTGATGACACGAAAACGACGGGTT

107521  ATAGACACAAGTTTCTTTTTTTTCTTTCTTTCTTTTTTAAAGAGGTGGAGTCTCACTATG
107521  TATCTGTGTTCAAAGAAAAAAAAGAAAGAAAGAAAAAATTTCTCCACCTCAGAGTGATAC

107581  TCACCCAGGCTGCAGTGCAATGGTGTGATCTCAGCTCACTGCACCCTGGTTCAAGCAATT
107581  AGTGGGTCCGACGTCACGTTACCACACTAGAGTCGAGTGACGTGGGACCAAGTTCGTTAA

107641  CTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTGCAGGAGCCTGCCACCATGCCCAGCTAA
107641  GAGGACGGAGTCGGAGGGCTCATCGACCCTAACGTCCTCGGACGGTGGTACGGGTCGATT

107701  TTTTTGTATTTTTTTTTTTAGTAGAGATAGGGTTTCACCATGTTGGTCAGGCTGGTCTTG
107701  AAAAACATAAAAAAAAAAAATCATCTCTATCCCAAAGTGGTACAACCAGTCCGACCAGAAC

107761  AACTCCTGACCTCAGGTGATCCACCGGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGTG
107761  TTGAGGACTGGAGTCCACTAGGTGGCCGGAACCGGAGGGTTTCACGACCCTAATGTCCAC

107821  TGAGCCACTGCACCCGGCCGAGTTTCGATTTAAAGAAGCTCTAGAGGTTTTTTGTTTTTC
107821  ACTCGGTGACGTGGGCCGGCTCAAAGCTAAATTTCTTCGAGATCTCCAAAAAACAAAAAG

107881  GTTTTTTGTTTTTCTGGCTCTAGATGTACTGATTATTGAGATACTGAGATGTCACATTAA
107881  CAAAAAACAAAAAGACCGAGATCTACATGACTAATAACTCTATGACTCTACAGTGTAATT

107941  GGGATTTAAGTTACCATGATTGGAAGATTTGCTTCCTTATCCCTTCTGCAGTTCCCCAAC
107941  CCCTAAATTCAATGGTACTAACCTTCTAAACGAAGGAATAGGGAAGACGTCAAGGGGTTG

108001  AATTTTCCTATCTGAACACCTTGCATTGAGACACAGGTCTCAGATCAGGTCAATGGTCAT
108001  TTAAAAGGATAGACTTGTGGAACGTAACTCTGTGTCCAGAGTCTAGTCCAGTTACCAGTA

108061  TCACCATATACCGTATATAAAATGCTAGAGTAATAAATGCAGTGTACTGATTCTACTGGA
108061  AGTGGTATATGGCATATATTTTACGATCTCATTATTTACGTCACATGACTAAGATGACCT

108121  AGGATTTTGTGTACTCTCTGGCTGTAATTGTTTTAACAATTGTTGCCTTTTCTACATATT
108121  TCCTAAAACACATGAGAGACCGACATTAACAAAATTGTTAACAACGGAAAAGATGTATAA

108181  TTGAAAATGTATCACAATCTCTCCAAATTTACTCTAGAACTCAATCTGATTAAAGAGCAC
108181  AACTTTTACATAGTGTTAGAGAGGTTTAAATGAGATCTTGAGTTAGACTAATTTCTCGTG
```

FIGURE 4 (cont'd)

```
108241  CATTTTAAAAGATAAATCAATGTATCTTGGGATTGTAACTAGATCCTGGTTCATTCAAGA
108241  GTAAAATTTTCTATTTAGTTACATAGAACCCTAACATTGATCTAGGACCAAGTAAGTTCT

108301  CAATTTCATATATCCACTATAGATATCAATATTTCTCTATTATCTTTAATATCCTGAGTA
108301  GTTAAAGTATATAGGTGATATCTATAGTTATAAAGAGATAATAGAAATTATAGGACTCAT

108361  CTCATTACTTTTCTTTGCTGTTTTTTTTTATTAGTTGGCAGAGACCATAGGGGTTCAGTG
108361  GAGTAATGAAAAGAAACGACAAAAAAAAATAATCAACCGTCTCTGGTATCCCCAAGTCAC

108421  AAGTGTTGGCAACGTGTATGATATTGAATTAAATGGCCAACGGTGGGAGAAAAACAGGCA
108421  TTCACAACCGTTGCACATACTATAACTTAATTTACCGGTTGCCACCCTCTTTTTGTCCGT

108481  TTTATTTCTATTAGAATTTTGATAATCATACCAAATTTAGAAGCGTATTTTTAGACTTGT
108481  AAATAAAGATAATCTTAAAACTATTAGTATGGTTTAAATCTTCGCATAAAAATCTGAACA

108541  TTACATTCTTAGGTTGTTCATTTCACAGTAGTAATTATAGAGACTAGAAATATAAATGAT
108541  AATGTAAGAATCCAACAAGTAAAGTGTCATCATTAATATCTCTGATCTTTATATTTACTA

108601  GTTTCCAAATACTTATGAAGTCTCAGTGTTAGTAGTTTCAATTAAATTTTTTAAAATGCA
108601  CAAAGGTTTATGAATACTTCAGAGTCACAATCATCAAAGTTAATTTAAAAAATTTTACGT

108661  GTTTATATTTGAAAGTAGCTACAGTAGCCTCTCTTGCTCTAATGTTATTTGGCTACAAAA
108661  CAAATATAAACTTTCATCGATGTCATCGGAGAGAACGAGATTACAATAAACCGATGTTTT

108721  AGAGGAAAATGGTAATTTAGATGTGGCTGAATTCAGAGCTAAGTACTAATTTTAGTGTTA
108721  TCTCCTTTTACCATTAAATCTACACCGACTTAAGTCTCGATTCATGATTAAAATCACAAT

108781  GTTGCAGAGACTAACAAGATGAGTCTTATACTGCAGTCAGAGGTAGAATAAACAGAAGAT
108781  CAACGTCTCTGATTGTTCTACTCAGAATATGACGTCAGTCTCCATCTTATTTGTCTTCTA

108841  ATGGTGAGTAAGCCTTGTCCTGGGCTTGGCCATTGGTAGTCCTTGGCATATGAAAAACCA
108841  TACCACTCATTCGGAACAGGACCCGAACCGGTAACCATCAGGAACCGTATACTTTTTGGT

108901  GATACAGATCATTCTCATCCTTCCATTCCTGCTGTTTGTACTTGGCACTGTTTGAAAATA
108901  CTATGTCTAGTAAGAGTAGGAAGGTAAGGACGACAAACATGAACCGTGACAAACTTTTAT

108961  AAATTTATACAGGAATATTTATGATTTATAAATTTGAGTTACTGTATGTGACATATAACA
108961  TTTAAATATGTCCTTATAAATACTAAATATTTAAACTCAATGACATACACTGTATATTGT

109021  AAAGCTACACTTGGAAGATTGTAGCCAGTGTTGACCTCTGCACTTAGAAGATAATAGGCC
109021  TTTCGATGTGAACCTTCTAACATCGGTCACAACTGGAGACGTGAATCTTCTATTATCCGG

109081  GATTCAACCCTTCTTCCCATCTGCCCACTCTTACCCCCACCCCAAACTCCACCAAAAAAC
109081  CTAAGTTGGGAAGAAGGGTAGACGGGTGAGAATGGGGGTGGGGTTTGAGGTGGTTTTTTG

109141  CACACAGCTCAAAACAAAAACAGTTACATCATTCAGAGGTGGTACCATCATTATTTTTAT
109141  GTGTGTCGAGTTTTGTTTTTGTCAATGTAGTAAGTCTCCACCATGGTAGTAATAAAAATA

109201  GGATGGGAGATGTCAGAATTTAAGTACACTCTGAATCAGAGCTACATGATTTAACCAGAA
109201  CCTACCCTCTACAGTCTTAAATTCATGTGAGACTTAGTCTCGATGTACTAAATTGGTCTT

109261  ACACCAGAAACTGTCCTGGTGGAAAATTGTTCTTTATCCAAACAATAATCATGAAACATT
109261  TGTGGTCTTTGACAGGACCACCTTTTAACAAGAAATAGGTTTGTTATTAGTACTTTGTAA

109321  TCCACCATGGCTTTTAAAGGTCAGAATAGAAATGAGTGAGTTTGTCTTCTGCCTCTTGCA
109321  AGGTGGTACCGAAAATTTCCAGTCTTATCTTTACTCACTCAAACAGAAGACGGAGAACGT

109381  AAATATACAAGACCTTAAGACTGTTAGTTACTATTCAGATTATGGTTGCCTGGGATAACT
109381  TTTATATGTTCTGGAATTCTGACAATCAATGATAAGTCTAATACCAACGGACCCTATTGA

109441  TTAGTTTTAATATACTCCTTCACTTTGGGGATGTTCATCTAACAGAACAAATAAAATTCA
109441  AATCAAAATTATATGAGGAAGTGAAACCCCTACAAGTAGATTGTCTTGTTTATTTTAAGT

109501  GAATAAGGAGAAAGGAAATAGAATATATATTTGAAGTAAATAGAATGTTAGGAAAAACCT
109501  CTTATTCCTCTTTCCTTTATCTTATATATAAACTTCATTTATCTTACAATCCTTTTTGGA

109561  TTGGGCTAAAAAAAAAGACATGCAATTCTGAAGTAACTTTGATTTATATCATATGATTTT
109561  AACCCGATTTTTTTTTCTGTACGTTAAGACTTCATTGAAACTAAATATAGTATACTAAAA
```

FIGURE 4 (cont'd)

```
109621  AAGAAAATACAATGAACAAACGAGTTGAAAATCAGCACAATACTGATTTTCACAAACGTT
109621  TTCTTTTATGTTACTTGTTTGCTCAACTTTTAGTCGTGTTATGACTAAAAGTGTTTGCAA

109681  ACCGATAGGAGGTATAAACAGAGTTCTTTCGGTAATAGCCTGCCGTCTTTTCATGAGCGC
109681  TGGCTATCCTCCATATTTGTCTCAAGAAAGCCATTATCGGACGGCAGAAAAGTACTCGCG

109741  CCACACACACTGCTCCTGTTGGTTGAGATGGCTGGTGCAGGAAGGTGGATGGGGTGGTGG
109741  GGTGTGTGTGACGAGGACAACCAACTCTACCGACCACGTCCTTCCACCTACCCCACCACC

109801  ATAGTGTGTGTTCAAAGCACAGTGCTGTGCACAGTCACTATGTGTGGCACTCAGCTTGTA
109801  TATCACACACAAGTTTCGTGTCACGACACGTGTCAGTGATACACACCGTGAGTCGAACAT

109861  ACTCATGGGAGTTTTGATTTCAGTAACAGCACCCATGGTGGCAAGATCACACATCATCAC
109861  TGAGTACCCTCAAAACTAAAGTCATTGTCGTGGGTACCACCGTTCTAGTGTGTAGTAGTG

109921  TCTGTTGATGAGTAATTTAACCCATACGCTTGCTATTCACCATTTGAACAAAGCATTGTA
109921  AGACAACTACTCATTAAATTGGGTATGCGAACGATAAGTGGTAAACTTGTTTCGTAACAT

109981  TCATATCCTGGACCCTGGAGATACCCTGAAAAATAAGACCCAGGCCTTGTGGTTGAGTTC
109981  AGTATAGGACCTGGGACCTCTATGGGACTTTTTATTCTGGGTCCGGAACACCAACTCAAG

110041  ACAGTCACTTTGGGAAAACAACCACAGCAGTGAATCTCGGTATTGCTCAGGCCACATCAT
110041  TGTCAGTGAAACCCTTTTGTTGGTGTCGTCACTTAGAGCCATAACGAGTCCGGTGTAGTA

110101  GAGTCTTCTCGGACACGTCAGGCGTGTATACGGACTCGACTTGGAGCACTCTTCCTCAGA
110101  CTCAGAAGAGCCTGTGCAGTCCGCACATATGCCTGAGCTGAACCTCGTGAGAAGGAGTCT

110161  TGTCTGTACCGCTTGCTCTTTGACATCCTCCAAATCTTTGCTTACAAATCCCATTCTCAG
110161  ACAGACATGGCGAACGAGAAACTGTAGGAGGTTTAGAAACGAATGTTTAGGGTAAGAGTC

110221  TCTGGCTTTCCTTGAGCTCTTAATCAAAATCATAAACTCCATTTCCCATTTCTCCTCAGC
110221  AGACCGAAAGGAACTCGAGAATTAGTTTTAGTATTTGAGGTAAAGGGTAAAGAGGAGTCG

110281  ATTCCAGCATTCTCCATTGCTTTTCCTTCTTTGTTTCCCCCTGTGTAGTCTAACATACTA
110281  TAAGGTCGTAAGAGGTAACGAAAAGGAAGAAACAAAGGGGGACACATCAGATTGTATGAT

110341  TCAACACAATAAATGTTTACCTGTGTCTATTAGAAAGTAAGCAGGGAGATATGTCAGTTT
110341  AGTTGTGTTATTTACAAATGGACACAGATAATCTTTCATTCGTCCCTCTATACAGTCAAA

110401  TGTTCACTGCTGTATTCATAGAGCTTAAAACGGTAAACAGCACTCAATATGTATTTGTAG
110401  ACAAGTGACGACATAAGTATCTCGAATTTTGCCATTTGTCGTGAGTTATACATAAACATC

110461  AATAAATGGCTAATTGTGCTGATGGTGTTTAGGGCACTTATGCAGACCTACATTCCTGGT
110461  TTATTTACCGATTAACACGACTACCACAAATCCCGTGAATACGTCTGGATGTAAGGACCA

110521  TTCCATTAGTGTTTGACTTTGGGAAGTCACAGATGAATCCTGCTGTTATTGCTTCTAAAT
110521  AAGGTAATCACAAACTGAAACCCTTCAGTGTCTACTTAGGACGACAATAACGAAGATTTA

110581  TCCTCATCTCCTAAAAATGACTTCTTTGTCAAAAACAAGCTGCATTTGTAGGTTCTGTTT
110581  AGGAGTAGAGGATTTTTACTGAAGAAACAGTTTTTGTTCGACGTAAACATCCAAGACAAA

110641  TCCTCATGTTGGGCAATCTGGAGTGATTTCTGGCTCTTGTAAGTATGTTTGTGGTTGCAG
110641  AGGAGTACAACCCGTTAGACCTCACTAAAGACCGAGAACATTCATACAAACACCAACGTC

110701  AGGAAAATGAATATTTCTCACCCAAATGTTGTCATGTCATCATCTCTGTCTGGCAACCAG
110701  TCCTTTTACTTATAAAGAGTGGGTTTACAACAGTACAGTAGTAGAGACAGACCGTTGGTC

110761  AGAAGCAGATGTTCTGTTCCTTGACCTCAGGCTAGTGGTTACAGTTCTGGGTGTTAAACA
110761  TCTTCGTCTACAAGACAAGGAACTGGAGTCCGATCACCAATGTCAAGACCCACAATTTGT

110821  TGTTTCACAAAACAGCACTTGCACATGCGGGGCTGCTGTTTTCATTTAACTTCATGTAGT
110821  ACAAAGTGTTTTGTCGTGAACGTGTACGCCCCGACGACAAAAGTAAATTGAAGTACATCA

110881  CTCAGGGCTCAGCAAGCTTTAGGAAAGGTCCGGTTTGGGGGCACTGGAAGTGTGTGTGTG
110881  GAGTCCCGAGTCGTTCGAAATCCTTTCCAGGCCAAACCCCCGTGACCTTCACACACACAC
```

FIGURE 4 (cont'd)

```
110941  TGTGTACATATATATACACACACAAATAGATCCATACATTTATGTATCTAGTTGGCTAGC
110941  ACACATGTATATATATGTGTGTGTTTATCTAGGTATGTAAATACATAGATCAACCGATCG

111001  TATTGCAAAACACATGTTTATTTTGAGATCATCAACACTTCTAAACATGACAAAGAGTAG
111001  ATAACGTTTTGTGTACAAATAAAACTCTAGTAGTTGTGAAGATTTGTACTGTTTCTCATC

111061  CCCTAAGAACATTGCCAAGAACAAATTTTCTCCTTTCCAGAAAGCAATTTTCAGCAGATG
111061  GGGATTCTTGTAACGGTTCTTGTTTAAAAGAGGAAAGGTCTTTCGTTAAAAGTCGTCTAC

111121  ATTTTCAGAAACAAAATTCAAAAAAGCCATACTGCTGTATCCTAGCAAAGATAACTTATT
111121  TAAAAGTCTTTGTTTTAAGTTTTTTCGGTATGACGACATAGGATCGTTTCTATTGAATAA

111181  CGTTTATTTTTTACTCAAAAACATTTATTATATACACTCTGGGCTAGGCACAATGCCCAG
111181  GCAAATAAAAAATGAGTTTTTGTAAATAATATATGTGAGACCCGATCCGTGTTACGGGTC

111241  TTTTGAGGATAGAAGGATGAAGATAGAGTTTGGTATTCCAGATAGAGCTTGATTGTTGGG
111241  AAAACTCCTATCTTCCTACTTCTATCTCAAACCATAAGGTCTATCTCGAACTAACAACCC

111301  CATGTCTGATGTAACCCCAGTGGAGTGGGCGAGCCTAGAACCACCATGTATTGAGGGAGA
111301  GTACAGACTACATTGGGGTCACCTCACCCGCTCGGATCTTGGTGGTACATAACTCCCTCT

111361  GGGTGGTGGCCTCAGTGTCAGGACTCACTAGATTCAGACTCCTGGGTGTCAGAAGTGGGG
111361  CCCACCACCGGAGTCACAGTCCTGAGTGATCTAAGTCTGAGGACCCACAGTCTTCACCCC

111421  AGGAAAGGGAGATGCTGAGTGCCAGAGGGTACTGTAAGGAGAGCTTGTGCAGCTGATGGA
111421  TCCTTTCCCTCTACGACTCACGGTCTCCCATGACATTCCTCTCGAACACGTCGACTACCT

111481  TCATGCATGATGCCTGTGAACAATTGTCCTACTAGACCTCACAGTAGTTAGGGATGTCTG
111481  AGTACGTACTACGGACACTTGTTAACAGGATGATCTGGAGTGTCATCAATCCCTACAGAC

111541  CAGTCATCTCACAGTAGAGCAGTAGCTCAGCTTTGGAGTTTTGTTTCTTGAATGTCTTCT
111541  GTCAGTAGAGTGTCATCTCGTCATCGAGTCGAAACCTCAAAACAAAGAACTTACAGAAGA

111601  TTTTTAACCGTACACGTGCTTGAGCTCACTCTGATAAGCCCATTACATTTGAATCTCAGA
111601  AAAAATTGGCATGTGCACGAACTCGAGTGAGACTATTCGGGTAATGTAAACTTAGAGTCT

111661  GAGCAAAGGCAGCATCTCCATACTTGTTGCTGATGCTCCACCAGGTTTGAAAATCACTAA
111661  CTCGTTTCCGTCGTAGAGGTATGAACAACGACTACGAGGTGGTCCAAACTTTTAGTGATT

111721  AGATAGACTCTTAGGCTGGGTGCGGTGGCTCACATCTGTAATCTCAGCACTTTGGGAGGC
111721  TCTATCTGAGAATCCGACCCACGCCACCGAGTGTAGACATTAGAGTCGTGAAACCCTCCG

111781  CGAAGTGGGCAGATCATCTGAGCTCAGGAGTTCGAGACCAGCCTGACCAACATGGTGAAA
111781  GCTTCACCCGTCTAGTAGACTCGAGTCCTCAAGCTCTGGTCGGACTGGTTGTACCACTTT

111841  CCCCGTCTCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGCGCGTACCTGTAATCCC
111841  GGGGCAGAGATGATTTTTATGTTTTTAATCGGTCCGCACCACCGCGCATGGACATTAGGG

111901  AGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGACCCTGGGAGGCAGAGGGTGCAGTG
111901  TCGATGAGTCCTCCGACTCCGTCCTCTTAGTGAACTGGGACCCTCCGTCTCCCACGTCAC

111961  AGCTGAGATGGTGCCACTGCACTCCAGCCTGGGTAACCTAAAAAAAAAAAAAAAAAAAAA
111961  TCGACTCTACCACGGTGACGTGAGGTCGGACCCATTGGATTTTTTTTTTTTTTTTTTTT

112021  AAATTAGACTTTCCAAGGGCAGCGTGGAAGCAGTGACCAGCCTGAGTAGGCACGCACACT
112021  TTTAATCTGAAAGGTTCCCGTCGCACCTTCGTCACTGGTCGGACTCATCCGTGCGTGTGA

112081  TCTTGGGGATAGCCTCCTCCATAGCATTCTTAGAAGGTTTGGTAATGACCAGTTTTTCTG
112081  AGAACCCCTATCGGAGGAGGTATCGTAAGAATCTTCCAAACCATTACTGGTCAAAAAGAC

112141  AGTAAGACCCTATCTACATTATTCTAATTAGTAACCACCTTTCATCTGCTAAGGTAAATT
112141  TCATTCTGGGATAGATGTAATAAGATTAATCATTGGTGGAAAGTAGACGATTCCATTTAA

112201  TTCACTTCAGAATATTCCTCTTAAGACCATGCAATAACTCTTACTGTTTGTTGGTGGAAA
112201  AAGTGAAGTCTTATAAGGAGAATTCTGGTACGTTATTGAGAATGACAAACAACCACCTTT
```

FIGURE 4 (cont'd)

```
112261  ACCATCTCATTATTTTACTGGTTAAGGGTAAACTGAAGATGATTAGATAACTTTCCTTCC
112261  TGGTAGAGTAATAAAATGACCAATTCCCATTTGACTTCTACTAATCTATTGAAAGGAAGG

112321  TTCCCTCCCCCCTTCCTCCCTTCCTTCTCTTTCTTCCTTCCTTTCTCCCTCCCTCCCTCC
112321  AAGGGAGGGGGGAAGGAGGGAAGGAAGAGAAAGAAGGAAGGAAAGAGGGAGGGAGGGAGG

112381  CTCTCTCTTTCTTTCCTTCTCTCTTTCTTTTCTTTTCTTTTTCTTTCTTTCTTTCTTTT
112381  GAGAGAGAAAGAAAGGAAGAGAGAAAGAAAAGAAAAGAAAAAAGAAAGAAAGAAAGAAAA

112441  TTCTTTCTTCCTTCCTTCGTTCCTTCCTCCCTCCCTCCCTCTCTCTCTTTCTTTCTTTCT
112441  AAGAAAGAAGGAAGGAAGCAAGGAAGGAGGGAGGGAGGGAGAGAGAGAAAGAAAGAAAGA

112501  CTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTCTCTTTCTT
112501  GAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAAAGAA

112561  TTCTTTCTTTCTTTCTTTCTCTTTCTCTTTCCTTCCTTCCTTCCTTTCTTTCTCTCTCTC
112561  AAGAAAGAAAGAAAGAAAGAGAAAGAGAAAGGAAGGAAGGAAGGAAAGAAAGAGAGAGAG

112621  TCCTTCCTTCCTTCCTTCTTTCCTTCCTTCCTTCTTTCCTTCCTTCCTTCTTTCTTCTTT
112621  AGGAAGGAAGGAAGGAAGAAAGGAAGGAAGGAAGAAAGGAAGGAAGGAAGAAAGAAGAAA

112681  TTTTAAATAAGGGTCTACCTGAAGTGTGAACTGATTTGGGAAATTTGATCATGCAGAACA
112681  AAAATTTATTCCCAGATGGACTTCACACTTGACTAAACCCTTTAAACTAGTACGTCTTGT

112741  GTAAAATTTTCTTTATGTTTGTTGCTAGCAAAATTGTGATGGTTTTTTAAAAAAATGGCT
112741  CATTTTAAAAGAAATACAAACAACGATCGTTTTAACACTACCAAAAAATTTTTTTACCGA

112801  ATTGATAATAAGGCATGCCTAATGTAGTGAAAGTTTGAAAATGTCCCAATTTCTGATTGA
112801  TAACTATTATTCCGTACGGATTACATCACTTTCAAACTTTTACAGGGTTAAAGACTAACT

112861  CAGAAGAAGAACTGGCAACTCCTCCACTAGATGGCATCATTCTTCCAGGAGTGACAAGGC
112861  GTCTTCTTCTTGACCGTTGAGGAGGTGATCTACCGTAGTAAGAAGGTCCTCACTGTTCCG

112921  GGTGCATTCTGGACCTGGCACATCAGTGGGTGGGTGCCTTTGATATGAACAACTTTTGTA
112921  CCACGTAAGACCTGGACCGTGTAGTCACCCACCCACGGAAACTATACTTGTTGAAAACAT

112981  AGCCTGAAATAAAAAATAATCAAAATAAGCCCAGCCTAGTGTTGAAATAATATCCTGTGG
112981  TCGGACTTTATTTTTTATTAGTTTTATTCGGGTCGGATCACAACTTTATTATAGGACACC

113041  TTCCAACCAATGTTATCTAATCTTTAAATTCTAACGTGAATGAAAAGTATTCTTCTAGAA
113041  AAGGTTGGTTACAATAGATTAGAAATTTAAGATTGCACTTACTTTTCATAAGAAGATCTT

113101  GGGCTTTCTTGGCAAGACTATTAAAATAAGTATTGTTGATAGCAAGGGCAGAGAGAAACC
113101  CCCGAAAGAACCGTTCTGATAATTTTATTCATAACAACTATCGTTCCCGTCTCTCTTTGG

113161  AGTTTAGCAAGATGTCATAGTGGAGAAGTATTCTAGTAGTAGTCTCAGTCTACCAGGTAT
113161  TCAAATCGTTCTACAGTATCACCTCTTCATAAGATCATCATCAGAGTCAGATGGTCCATA

113221  CTCCCACCAACTTTTGTCTGGTTCATACCTGATGTGCTTCCATCCAGACATTCCTTAACT
113221  GAGGGTGGTTGAAAACAGACCAAGTATGGACTACACGAAGGTAGGTCTGTAAGGAATTGA

113281  TCCCAAGTCTGTGTTGCTTCCAAATGGCAATGATACTTCTTCCTTTCCTTTATTTATACC
113281  AGGGTTCAGACACAACGAAGGTTTACCGTTACTATGAAGAAGGAAAGGAAATAAATATGG

113341  AGTGTAACCATGCTTGCCGTAGAAATGAAATAATTTCACTGACATATAAATGTTTACTAA
113341  TCACATTGGTACGAACGGCATCTTTACTTTATTAAAGTGACTGTATATTTACAAATGATT

113401  TATAACATTATTGGTTTGACTGCCTTCTTTTTCTCTCTGTCCTACCCCCATCCTCCAAAA
113401  ATATTGTAATAACCAAACTGACGGAAGAAAAAGAGAGACAGGATGGGGGTAGGAGGTTTT

113461  TTCATATACATGCCTATTTTAAAATCAGAACATGCAAGCAAATGTAAAGTTGTTTTCAGT
113461  AAGTATATGTACGGATAAAATTTTAGTCTTGTACGTTCGTTTACATTTCAACAAAAGTCA

113521  AGACTTTCAAATTCGTGTGACACAGGCAAGCATATTAAACAATTTGGTGGACCGGGTGCA
113521  TCTGAAAGTTTAAGCACACTGTGTCCGTTCGTATAATTTGTTAAACCACCTGGCCCACGT
```

FIGURE 4 (cont'd)

```
113581  GTGGTTCATGCCTATAATTCTAGCATTTTGGGAGGCCAAGGCAGGAGGATCACTTAAGCC
113581  CACCAAGTACGGATATTAAGATCGTAAAACCCTCCGGTTCCGTCCTCCTAGTGAATTCGG

113641  CAGGAATTCGAGACCAGCCTGGGCAACATAGTGAGACCCCGAATCTACAAAAAAATCAAA
113641  GTCCTTAAGCTCTGGTCGGACCCGTTGTATCACTCTGGGGCTTAGATGTTTTTTTAGTTT

113701  AACATTAGCTGAACATGGTGGTACATGCCTGTAGTCCCAGCTACTTGGAGGGGCTGAGGT
113701  TTGTAATCGACTTGTACCACCATGTACGGACATCAGGGTCGATGAACCTCCCCGACTCCA

113761  GGGAGGATGGCATAAGATCTGGAGGTCAAGGTCGCAGTGAGCTATGATCAAGCCACTGCA
113761  CCCTCCTACCGTATTCTAGACCTCCAGTTCCAGCGTCACTCGATACTAGTTCGGTGACGT

113821  CTACAGCCTGAATACAGGGTGAGACCATGTCTCTAAAAGTAAATTAAATAAATAAATAAA
113821  GATGTCGGACTTATGTCCCACTCTGGTACAGAGATTTTCATTTAATTTATTTATTTATTT

113881  TAAATAAATAAATAAATAAATAAATAAAATAAAATTTAAAAAGGTTTGGTTTTGGAAATG
113881  ATTTATTTATTTATTTATTTATTTATTTTATTTTAAATTTTTCCAAACCAAAACCTTTAC

113941  ACGTAATGTATTCAAGACAAATTATTGAATGAATTGGGTGTCTGGTAAAGGTTCTTGGAT
113941  TGCATTACATAAGTTCTGTTTAATAACTTACTTAACCCACAGACCATTTCCAAGAACCTA

114001  GGAACTTTTGACCCAGGTGTTAAAATGGTCTCCAAAGAACATGAAAAGGCTTTGAGCTTA
114001  CCTTGAAAACTGGGTCCACAATTTTACCAGAGGTTTCTTGTACTTTTCCGAAACTCGAAT

114061  CTGTGGTCAAAGAGTACTAGTAGTAATAGGACAAATGTTTAGAAACACACACGTTATCTA
114061  GACACCAGTTTCTCATGATCATCATTATCCTGTTTACAAATCTTTGTGTGTGCAATAGAT

114121  GGAGCATACCAGCATGGTTTCTCTGGAAAAGACTTTCAAAATAGATGAAATGTGCTATAA
114121  CCTCGTATGGTCGTACCAAAGAGACCTTTTCTGAAAGTTTTATCTACTTTACACGATATT

114181  ATAATTCTACAAAACCCTGAAGTAGACAAGACCCAACTGCTACATGCTGGCTGGCGCATT
114181  TATTAAGATGTTTTGGGACTTCATCTGTTCTGGGTTGACGATGTACGACCGACCGCGTAA

114241  GGGCATGACATTTACAAACCGTTGGTATAATTTTCAAACAGCTTGCCCTGTCCTAAATAT
114241  CCCGTACTGTAAATGTTTGGCAACCATATTAAAAGTTTGTCGAACGGGACAGGATTTATA

114301  GTTTGCAATTGGAAAGATATTTTTGAAGCTCAGTAATTTTTGTTTTATTAAAGAAAAATA
114301  CAAACGTTAACCTTTCTATAAAAACTTCGAGTCATTAAAAACAAAATAATTTCTTTTTAT

114361  TCAACCAAGTACAGAAGGTTCGACGACAAGTTGCAAGTGGAAATACTTTCCCACACTAAT
114361  AGTTGGTTCATGTCTTCCAAGCTGCTGTTCAACGTTCACCTTTATGAAAGGGTGTGATTA

114421  GATTTGCTAGCATGAAAAATAGGTTATTATGGAGGAAATTTTTTCTTGACTTCATAGCTG
114421  CTAAACGATCGTACTTTTTATCCAATAATACCTCCTTTAAAAAAGAACTGAAGTATCGAC

114481  TAATTTATTTTAAAAGCTATTTTAAAACATGAGACGTGAAATGCCAAGAGGCATGTAGGG
114481  ATTAAATAAAATTTTCGATAAAATTTTGTACTCTGCACTTTACGGTTCTCCGTACATCCC

114541  GGTCACCACTATTTGCCAGATAGGGGAAATTGTCCAAAGATAAGTAAGGGTGAGCAGAGC
114541  CCAGTGGTGATAAACGGTCTATCCCCTTTAACAGGTTTCTATTCATTCCCACTCGTCTCG

114601  AAAAACAATTCTGCTTATTTACAAAAGAGTTTTCTTTTCTTTTCTTTTTCAGGACACAGA
114601  TTTTTGTTAAGACGAATAAATGTTTTCTCAAAAGAAAAGAAAAGAAAAAGTCCTGTGTCT

114661  ACTCAGCTTGTTTTCAATTAATTTGCCTGATTTTCTGCAGTTCATTTACTTTTGAACAAC
114661  TGAGTCGAACAAAAGTTAATTAAACGGACTAAAAGACGTCAAGTAAATGAAAACTTGTTG

114721  ATAATTGCAATTGTAGACTGAGAGAAATTGAAACTTTCAAAGAGCCATATTTCTATTGCA
114721  TATTAACGTTAACATCTGACTCTCTTTAACTTTGAAAGTTTCTCGGTATAAAGATAACGT

114781  GATATATTTTCCTGCTCTTCCAAATCTACTTACAGCATGAGTTCTTCTTTTAAATATTCA
114781  CTATATAAAAGGACGAGAAGGTTTAGATGAATGTCGTACTCAAGAAGAAAATTTATAAGT

114841  AATATTTTGAATATTGCCAAGAGCTTTGATTTCCATTTTTATCTCTTGTGGGTTTATAAA
114841  TTATAAAACTTATAACGGTTCTCGAAACTAAAGGTAAAAATAGAGAACACCCAAATATTT

114901  TTAAGAAAAAATACTCATCTTATTTTTTTAAACCTCTCTATTTTTATTGCCCTTTATTCA
114901  AATTCTTTTTTATGAGTAGAATAAAAAAATTTGGAGAGATAAAAATAACGGGAAATAAGT
```

FIGURE 4 (cont'd)

```
114961  AATAACTTGTTGACAAACTTTGAACTTGAACCACTGAGGTAAAAGAACAAGAATTAAACA
114961  TTATTGAACAACTGTTTGAAACTTGAACTTGGTGACTCCATTTTCTTGTTCTTAATTTGT

115021  GATAGTTTAAACACATAGCTTAAAAGGATCTTTTTCCCATTTCCTATCCTTGAGCAAAGA
115021  CTATCAAATTTGTGTATCGAATTTTCCTAGAAAAAGGGTAAAGGATAGGAACTCGTTTCT

115081  ATATATTCAAACACTTTGGCAGAAGTCAATGAGGTTATACCACTAATTCCATGATGAAAA
115081  TATATAAGTTTGTGAAACCGTCTTCAGTTACTCCAATATGGTGATTAAGGTACTACTTTT

115141  TCAACTGAATGTGATACTGAAAGAGAAGGAAGAGAATTGTCACTGTAAAGTCAACTGTTA
115141  AGTTGACTTACACTATGACTTTCTCTTCCTTCTCTTAACAGTGACATTTCAGTTGACAAT

115201  GTCATATTAGGAAAAAAAATACATACAATACAATTTCTCAAATAAAGTCCAAATATACAT
115201  CAGTATAATCCTTTTTTTTATGTATGTTATGTTAAAGAGTTTATTTCAGGTTTATATGTA

115261  TCAACGTTTAAAAATAATGAGTATTTCAGATATTTGAACTCAGTCTGTTCTTTATTCCAT
115261  AGTTGCAAATTTTTATTACTCATAAAGTCTATAAACTTGAGTCAGACAAGAAATAAGGTA

115321  AAAAGATATAGGTAAGCCGTGCACAGTGGCTCACAACTATAATCCCAGCACTTTGGCACT
115321  TTTTCTATATCCATTCGGCACGTGTCACCGAGTGTTGATATTAGGGTCGTGAAACCGTGA

115381  TTGGGAGGCTGAGGTGGGAGGATCACATGAGCCCAGCCTGGGCAACATAGGGAGACCCCT
115381  AACCCTCCGACTCCACCCTCCTAGTGTACTCGGGTCGGACCCGTTGTATCCCTCTGGGGA

115441  ATCTTTACAAAATAAAATATAAAATATAAAACCTAGTTGGGCATGGCAGCATACACCTGT
115441  TAGAAATGTTTTATTTTATATTTTATATTTTGGATCAACCCGTACCGTCGTATGTGGACA

115501  AGTCCCAGGTGCTCAGGAGACTGAGACAGGAGGATCACTTGGGCCTGGGAGGTCGAGGCT
115501  TCAGGGTCCACGAGTCCTCTGACTCTGTCCTCCTAGTGAACCCGGACCCTCCAGCTCCGA

115561  GCAATGAGCCAAGATTATGCCACTGCATTCCAGCCTGGGTGACAGGGCAAGACCCTGTCT
115561  CGTTACTCGGTTCTAATACGGTGACGTAAGGTCGGACCCACTGTCCCGTTCTGGGACAGA

115621  TAAAAAAAAAAAAAAAAAGAAAGAGACAAACGTAGAAAGAGTACCAAATTTGGAGCCAATC
115621  ATTTTTTTTTTTTTTTTCTTTCTCTGTTTGCATCTTTCTCATGGTTTAAACCTCGGTTAG

115681  AGATCAGACCTGGGTTCAAATTGAGTTTCTGTCACTTACTTCCTTTGTGGCCCTGGAGAT
115681  TCTAGTCTGGACCCAAGTTTAACTCAAAGACAGTGAATGAAGGAAACACCGGGACCTCTA

115741  GCCCTTCAACCTCCTTGAACCTTAGTTTCCTCATCTGTAAAATGGGGATGATAGTACTCA
115741  CGGGAAGTTGGAGGAACTTGGAATCAAAGGAGTAGACATTTTACCCCTACTATCATGAGT

115801  CCCTTAAAATTACAGAACTTGGCATAATAGCGGGTGCTTGATAAATGGTAGCTATAATTA
115801  GGGAATTTTAATGTCTTGAACCGTATTATCGCCCACGAACTATTTACCATCGATATTAAT

115861  TTAATATTATATAATACATTGTTCTCCTTTATGGTAAAGTAGGTTTTAAGACACTGATTT
115861  AATTATAATATATTATGTAACAAGAGGAAATACCATTTCATCCAAAATTCTGTGACTAAA

115921  TTTTGTTTTGTTTTGTTTTTGGAGACAGAGTCTCACTCCGTCACCCAGGCTGGACTGCAG
115921  AAAACAAAACAAAACAAAAACCTCTGTCTCAGAGTGAGGCAGTGGGTCCGACCTGACGTC

115981  TGACGTGATCTTGGTTCACTGCAACCTCTGCCTCCTGAGTTCAAGCGATTCTCCTGCCTC
115981  ACTGCACTAGAACCAAGTGACGTTGGAGACGGAGGACTCAAGTTCGCTAAGAGGACGGAG

116041  AGCCTCCTGAGTAGCTGGGATTACAGGCACCCGCCACCATGCCTGGCTAATTTTTGTATT
116041  TCGGAGGACTCATCGACCCTAATGTCCGTGGGCGGTGGTACGGACCGATTAAAAACATAA

116101  TTTAGTAGAGACAGGGTTTTACCAGGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAAG
116101  AAATCATCTCTGTCCCAAAATGGTCCAACCGGTCCGACCAGAGTTTGAGGACTGGAGTTC

116161  TAATCTGCCTGCCTCAGCCTCCCCAAATGCAGAGATTACAGGCATGAGCCACCGTGCCCG
116161  ATTAGACGGACGGAGTCGGAGGGGTTTACGTCTCTAATGTCCGTACTCGGTGGCACGGGC

116221  GCCTAAGACACTGTTAACACGGAGCATGCATCAAGCAGCATTTGGGATGCATCAATTTTG
116221  CGGATTCTGTGACAATTGTGCCTCGTACGTAGTTCGTCGTAAACCCTACGTAGTTAAAAC
```

FIGURE 4 (cont'd)

```
116281  TGAGACTTTATACAAGTTAACAAATAAGAAGGCAAAAAAGAAGGCATATGTTAAAGTATA
116281  ACTCTGAAATATGTTCAATTGTTTATTCTTCCGTTTTTTCTTCCGTATACAATTTCATAT

116341  CCATGTGTCTTGTGGTGATAATCTTATATTTTATATTCTGATAATCTCATATATAAATCT
116341  GGTACACAGAACACCACTATTAGAATATAAAATATAAGACTATTAGAGTATATATTTAGA

116401  GTGACCTTGAAGTAGGTTACTTAACACCTCGATGTGTTCACCAGTAAGATGGTGTGCCAC
116401  CACTGGAACTTCATCCAATGAATTGTGGAGCTACACAAGTGGTCATTCTACCACACGGTG

116461  TTCCAGCTTTCCCTTAAAGGATGATTGTGACCTTGTCTAAGTGCAATCAGAATGTTAAGC
116461  AAGGTCGAAAGGGAATTTCCTACTAACACTGGAACAGATTCACGTTAGTCTTACAATTCG

116521  AGAACAGTAAGTATTGCAATTGTTTTGCATGCCTTTTCCTTTAAAATATACTTGTTATTT
116521  TCTTGTCATTCATAACGTTAACAAAACGTACGGAAAAGGAAATTTTATATGAACAATAAA

116581  TTTATATTCCATGCAGGGTGAATTTAAGGTGTCAGAGAGATACCTCACCATGGATGACTT
116581  AAATATAAGGTACGTCCCACTTAAATTCCACAGTCTCTCTATGGAGTGGTACCTACTGAA

116641  GACAACAGCCCTGGAGGGGAACAGAGTGAGAGAGATGTTTGGCTCTGGTACAGCCTGTGT
116641  CTGTTGTCGGGACCTCCCCTTGTCTCACTCTCTCTACAAACCGAGACCATGTCGGACACA

116701  TGTTTGCCCAGTTTCTGATATACTGTACAAAGGCGAGGTACGACAAGTATTTCCTCATTT
116701  ACAAACGGGTCAAAGACTATATGACATGTTTCCGCTCCATGCTGTTCATAAAGGAGTAAA

116761  CCTATCATTTCCAATCACATTAAATTACATTTAATTTAAGTGTATGAATAAAATTGTATC
116761  GGATAGTAAAGGTTAGTGTAATTTAATGTAAATTAAATTCACATACTTATTTTAACATAG

116821  TGGACCCAGACCCAAGATGCAAAAGTTTTGTCGTTGTTGTTGTTTTTCTTTATTTTAAAG
116821  ACCTGGGTCTGGGTTCTACGTTTTCAAAACAGCAACAACAACAAAAAGAAATAAAATTTC

116881  AGATGGAGTCTCGCTGTGTTGCCCAGGCTGGAGTGCAGTGGCTATTCACTGGTGTGATCA
116881  TCTACCTCAGAGCGACACAACGGGTCCGACCTCACGTCACCGATAAGTGACCACACTAGT

116941  TAGCTCACTGCAGCCTCAAACTCCTGGGCTCACGCAATCCTTCTGCTTCAGCCTCCTGAG
116941  ATCGAGTGACGTCGGAGTTTGAGGACCCGAGTGCGTTAGGAAGACGAAGTCGGAGGACTC

117001  TTGCTGGGATTACAGGAGTGCGCCACCGCACCATGCTAGGTCTTTTTTGTTGTTGTTTTT
117001  AACGACCCTAATGTCCTCACGCGGTGGCGTGGTACGATCCAGAAAAAACAACAACAAAAA

117061  GTTTTTTAACTTGGCACTACAAAAAAAATAACTGGGACAGTTTATCCAGATGAAACTGTA
117061  CAAAAAATTGAACCGTGATGTTTTTTTTATTGACCCTGTCAAATAGGTCTACTTTGACAT

117121  TTAGCATATATCTACTATGTTTCCAGTGATTTCATGTAAGCTTAAACATGGTGGGAAGTT
117121  AATCGTATATAGATGATACAAAGGTCACTAAAGTACATTCGAATTTGTACCACCCTTCAA

117181  AGTTTTGTCCTTGGACACGTGTAAGATCTGTCTTACCTACATGTCGGGATTGGAGCTATG
117181  TCAAAACAGGAACCTGTGCACATTCTAGACAGAATGGATGTACAGCCCTAACCTCGATAC

117241  CCGACAGTCTTCTACTCTTCTCTCTTCCTTCCTACCAGCTCACTAGGGAAGGCTCTATAG
117241  GGCTGTCAGAAGATGAGAAGAGAGAAGGAAGGATGGTCGAGTGATCCCTTCCGAGATATC

117301  GTAGTTACATTTCAGTTTCTGGTTAACCAAAAGTACTTCCAAACCACTGGATTAGGAATG
117301  CATCAATGTAAAGTCAAAGACCAATTGGTTTTCATGAAGGTTTGGTGACCTAATCCTTAC

117361  GCAAATAGTTATGGAAAATGCATTCGAGTGGTCAATCCCTTGATTTTTTTTTACTATTAA
117361  CGTTTATCAATACCTTTTACGTAAGCTCACCAGTTAGGGAACTAAAAAAAAAATGATAATT

117421  TGCTGTTATCATTAATAGTAAGTAGACCACTTTACAAGAATTATTCATTGGCAAACCTGG
117421  ACGACAATAGTAATTATCATTCATCTGGTGAAATGTTCTTAATAAGTAACCGTTTGGACC

117481  TGCATTTACTAAAATCTTACAAAAGGAATTTGAGGATTTGGTTGGAAGAAATGTTCTCCT
117481  ACGTAAATGATTTTAGAATGTTTTCCTTAAACTCCTAAACCAACCTTCTTTACAAGAGGA

117541  CTTAGTAAATTAGTCTGATTTTATATGTGCCATACTTTTGTTTTCTTGTCAAATTGTTTT
117541  GAATCATTTAATCAGACTAAAATATACACGGTATGAAAACAAAAGAACAGTTTAACAAAA

117601  GCTTTTCACTTTATAACAAAGCTCCAATACTTATGATTAACATTGGCATCTAACTTTATA
117601  CGAAAAGTGAAATATTGTTTCGAGGTTATGAATACTAATTGTAACCGTAGATTGAAATAT
```

FIGURE 4 (cont'd)

```
117661  ACTTAAAGGCATGTAAGTATTAGTCTTATAATGTGAACTGGCTTAAGTTATTCATATTAT
117661  TGAATTTCCGTACATTCATAATCAGAATATTACACTTGACCGAATTCAATAAGTATAATA

117721  AGCAGATCCTTGAACAACACAGGGATTAGGGGTGTCAACCCCCATGCAGTTGAAAAATCC
117721  TCGTCTAGGAACTTGTTGTGTCCCTAATCCCCACAGTTGGGGGTACGTCAACTTTTTAGG

117781  ACATACAACTTTTTTTGTTTGTTTGTTTTTTTAAACGGGGTCTTTCTCTGTCACCTGTAC
117781  TGTATGTTGAAAAAAACAAACAAACAAAAAAATTTGCCCCAGAAAGAGACAGTGGACATG

117841  TGGAGTGCAGGGGCACAATCTCGGCTCACTGCAACCTCCGCCTCCCAAGCTCAAGCGATC
117841  ACCTCACGTCCCCGTGTTAGAGCCGAGTGACGTTGGAGGCGGAGGGTTCGAGTTCGCTAG

117901  CTCTCACCTCAGTCTCCCAAGTAGCTGGGACCACAGGTGTGCACCACCACACCCAGCTAA
117901  GAGAGTGGAGTCAGAGGGTTCATCGACCCTGGTGTCCACACGTGGTGGTGTGGGTCGATT

117961  TTTTAGTATTTTTTGTAGAGACGGGGTTTCAAGATGTTTCACAGGCTGGTCTCAAACTCT
117961  AAAATCATAAAAAACATCTCTGCCCCAAAGTTCTACAAAGTGTCCGACCAGAGTTTGAGA

118021  TGAGTTCAAGTGATTCACCTGCCTTGACCTCCCAAAGTGCTGGTATTACAGGCGTGAGTC
118021  ACTCAAGTTCACTAAGTGGACGGAACTGGAGGGTTTCACGACCATAATGTCCGCACTCAG

118081  ACCATGCCCGGCCCACATATAACTTTTGACTTCCCAAAAACTTAACTACTAGTAGCCTGC
118081  TGGTACGGGCCGGGTGTATATTGAAAACTGAAGGGTTTTTGAATTGATGATCATCGGACG

118141  TGTTCACCGGAAGCCTTATCAACAACAAAAACAGTCGATTGGCACATATTTTGTATGTTA
118141  ACAAGTGGCCTTCGGAATAGTTGTTGTTTTTGTCAGCTAACCGTGTATAAAACATACAAT

118201  TGTATGTTATATATGTATTCTTACAATAAAGTAAGTTACGAAAATCATGGGGAAGGGAAA
118201  ACATACAATATATACATAAGAATGTTATTTCATTCAATGCTTTTAGTACCCCTTCCCTTT

118261  ATATATTTACTATTCATTACATGGAAGTGGATCATCATAAAGGTCTTCATCCTCATTGTC
118261  TATATAAATGATAAGTAATGTACCTTCACCTAGTAGTATTTCCAGAAGTAGGAGTAACAG

118321  CATCATGTTGAGTAGGCTGAGGAGGAGGAGGAGGAGGAGGGATTGGTCTTGCTGTTTTAG
118321  GTAGTACAACTCATCCGACTCCTCCTCCTCCTCCTCCCTAACCAGAACGACAAAATC

118381  GAGTGGCAGAGGCAGAAGAAAGTCCACGTATAAGTGAACTTGTGTAGTTCAAACCCGTGT
118381  CTCACCGTCTCCGTCTTCTTTCAGGTGCATATTCACTTGAACACATCAAGTTTGGGCACA

118441  TGTTCAAGTGTCAACCGTAAATGCAAACTCCACTGGTTTTCAAATTTTACTTTAGAGTTA
118441  ACAAGTTCACAGTTGGCATTTACGTTTGAGGTGACCAAAAGTTTAAAATGAAATCTCAAT

118501  TTTTCCCTCAGCTGATTGGAGTTTTGCTCAACTCCTCTGACTTTAATAATGGCTGCTTCA
118501  AAAAGGGAGTCGACTAACCTCAAAACGAGTTGAGGAGACTGAAATTATTACCGACGAAGT

118561  TTTGTTGAGGTTTTAGAACTTGCTTTTAAGGGATTGCTCTTGACATAATTGCATCAACAG
118561  AAACAACTCCAAAATCTTGAACGAAAATTCCCTAACGAGAACTGTATTAACGTAGTTGTC

118621  ATAGACTGTTTCATGGGAGAAGAATTGTTCTGTTTCATGTCAATACTTACATTGATAAGA
118621  TATCTGACAAAGTACCCTCTTCTTAACAAGACAAAGTACAGTTATGAATGTAACTATTCT

118681  ATGCATCAAGGTAATGTGATCCATCCACTGAAGTTGTCATTAGGCCAGATTTTTTTTAAG
118681  TACGTAGTTCCATTACACTAGGTAGGTGACTTCAACAGTAATCCGGTCTAAAAAAAATTC

118741  TTGAATTATTATTTCACTTCTAGAAATAAGAAGTTTATAAAGTTTAGCATAATTTTAATG
118741  AACTTAATAATAAAGTGAAGATCTTTATTCTTCAAATATTTCAAATCGTATTAAAATTAC

118801  ATCCAGGCTTAGGAAGCATTTACAGCCAAAAAGTAAGTATATATAATTATCATATTATAC
118801  TAGGTCCGAATCCTTCGTAAATGTCGGTTTTTCATTCATATATATTAATAGTATAATATG

118861  TAAAGGGTATAATATGAAAAGTGTTTGAGAGCACAGACTATCTCCTAAAATCAAACTATC
118861  ATTTCCCATATTATACTTTTCACAAACTCTCGTGTCTGATAGAGGATTTTAGTTTGATAG

118921  AAAACCTGCAAGCCAGCAAGGCATGAGTAGGCATGACTCAGATACAAATTAGAAGCATGA
118921  TTTTGGACGTTCGGTCGTTCCGTACTCATCCGTACTGAGTCTATGTTTAATCTTCGTACT
```

FIGURE 4 (cont'd)

```
118981  AATCTCTTTCAGCCATGCATTTAATCATGCTATCCAATATTGTGGAGGTAGTGGCCCCTG
118981  TTAGAGAAAGTCGGTACGTAAATTAGTACGATAGGTTATAACACCTCCATCACCGGGGAC

119041  GAAGTTTACCTTGGGCCAGAGATAGGTATAGAAGCCCCTTTTATGAAGTCATAGACTTGC
119041  CTTCAAATGGAACCCGGTCTCTATCCATATCTTCGGGGAAAATACTTCAGTATCTGAACG

119101  TTTAATTTATGCTCCCAGAATGACTTGCATTAAACTGTGGCAAAATATCTTTGAGTTCTC
119101  AAATTAAATACGAGGGTCTTACTGAACGTAATTTGACACCGTTTTATAGAAACTCAAGAG

119161  ATATTAGCCAGAGGGTTAATATGTAGAAGACAATCATTCATCCTTACAGATTCTCAGAAT
119161  TATAATCGGTCTCCCAATTATACATCTTCTGTTAGTAAGTAGGAATGTCTAAGAGTCTTA

119221  CTTGGAAGAGGTAGGGTTCTTTAAGTTTAGCCACTATTTTAGGTAACTGTAGCTCCTCCT
119221  GAACCTTCTCCATCCCAAGAAATTCAAATCGGTGATAAAATCCATTGACATCGAGGAGGA

119281  TGGTTCTTCAAATGATCCTGAAGGCTGGTATACAAACCAGTTATGCCCTAAAATTAATTG
119281  ACCAAGAAGTTTACTAGGACTTCCGACCATATGTTTGGTCAATACGGGATTTTAATTAAC

119341  GTAGGTAGGCAATGTATAACCTGTCATTGGCTAGTTTTTATTCAAATTTTGGATATGAAA
119341  CATCCATCCGTTACATATTGGACAGTAACCGATCAAAAATAAGTTTAAAACCTATACTTT

119401  GTCTTTGGCATAAGGTGCTAGCACTCAGAGTCAGTAGGTTTTAGTGCTATTCCTTAACTT
119401  CAGAAACCGTATTCCACGATCGTGAGTCTCAGTCATCCAAAATCACGATAAGGAATTGAA

119461  CAGTGGAATTGCCTTAGTATAGCAGAATTGCTTGAGTGTCTTATCTGTTATCACCATATA
119461  GTCACCTTAACGGAATCATATCGTCTTAACGAACTCACAGAATAGACAATAGTGGTATAT

119521  CATGAGTACCCTCAAATTATCTCATTTCCCTTTCTTTTGTAGACAATACACATTCCAACT
119521  GTACTCATGGGAGTTTAATAGAGTAAAGGGAAAGAAAACATCTGTTATGTGTAAGGTTGA

119581  ATGGAGAATGGTCCTAAGCTGGCAAGCCGCATCTTGAGCAAATTAACTGATATCCAGGTA
119581  TACCTCTTACCAGGATTCGACCGTTCGGCGTAGAACTCGTTTAATTGACTATAGGTCCAT

119641  AAGCTTTTCTTTTCTTTTCTTTTCTTTCCTTTTCTTTTCTTTTTTTTATGTCACCTTATT
119641  TTCGAAAAGAAAAGAAAAGAAAAGAAAGGAAAAGAAAAGAAAAAAAATACAGTGGAATAA

119701  TCTGTCAGAGCTTACAACTAAATTATATTTCAATGAAGAGGATTAAATAATACAATTTGA
119701  AGACAGTCTCGAATGTTGATTTAATATAAAGTTACTTCTCCTAATTTATTATGTTAAACT

119761  GATCAATAAGTTAAATTTAAGGAAGATATTCCCAACACAATGTGTGGGCATGAGCAAAGC
119761  CTAGTTATTCAATTTAAATTCCTTCTATAAGGGTTGTGTTACACACCCGTACTCGTTTCG

119821  AGAAATGGGAAAGAAGATGTATGTTTTAAAAAAAGAAAGAAAATATTGTTCAGGACTCCA
119821  TCTTTACCCTTTCTTCTACATACAAAATTTTTTTCTTTCTTTTATAACAAGTCCTGAGGT

119881  CATTACTTAACATTAAAAAATAGATGTAATTTTTGGTAACATTTAAAAATCTCATTAAGA
119881  GTAATGAATTGTAATTTTTTATCTACATTAAAAACCATTGTAAATTTTTAGAGTAATTCT

119941  AGGATAACATAACTTTTGTTAAAAAAAGGAAGAAAAAATTTGAAAGTAGAAGCAACTTCA
119941  TCCTATTGTATTGAAAACAATTTTTTTCCTTCTTTTTTAAACTTTCATCTTCGTTGAAGT

120001  GGAAGAAAATTACAGTGCATGTGATTCCAACACTGAGATATAATCACTGTTAATTTTTTA
120001  CCTTCTTTTAATGTCACGTACACTAAGGTTGTGACTCTATATTAGTGACAATTAAAAAAT

120061  TGTCTTTTTATTTTTGAGCTGGAGTTTCACTCTTGTTGCCCAGGCTGGAGTACGATGGCA
120061  ACAGAAAAATAAAAACTCGACCTCAAAGTGAGAACAACGGGTCCGACCTCATGCTACCGT

120121  CGATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAATTTTCCTACCTCAGCCT
120121  GCTAGAACCGAGTGACGTTGGAGACGGAGGGCCCAAGTTCGTTAAAAGGATGGAGTCGGA

120181  CCCGAGTAGCTGGGATTATAGGTGTGCGCCACCATGCCCAGCTAATTTTTGTGTTTTTAG
120181  GGGCTCATCGACCCTAATATCCACACGCGGTGGTACGGGTCGATTAAAAACACAAAAATC

120241  TAGAGATGGGATTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTAAGGTAATC
120241  ATCTCTACCCTAAAGTGGTACAACCGGTCCGACCAGAACTTGAGGACTGGATTCCATTAG

120301  CACCTGCCTTGGCCTCCTAAATTGTTGGGATTACAGGCATGAACCACTGTGCCCTGCCAT
120301  GTGGACGGAACCGGAGGATTTAACAACCCTAATGTCCGTACTTGGTGACACGGGACGGTA
```

FIGURE 4 (cont'd)

```
120361  TTTTTGATGTCTTAACCTTCAATATTTTGTAAGCATTTATATACTTTATTTATTTATTTA
120361  AAAAACTACAGAATTGGAAGTTATAAAACATTCGTAAATATATGAAATAAATAAATAAAT

120421  TTTATTTATTTATTTATTTACTTATTTATTTTTGAGACTAAGTCTCGCTGTGTCACCCAG
120421  AAATAAATAAATAAATAAATGAATAAATAAAAACTCTGATTCAGAGCGACACAGTGGGTC

120481  GCTGGAGAACAGTGGCATGATCTCAGCTCACTACAACCTCCACCTCCTGGGTTCAAGCGA
120481  CGACCTCTTGTCACCGTACTAGAGTCGAGTGATGTTGGAGGTGGAGGACCCAAGTTCGCT

120541  TTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGTCTCATGCCACCGTTCCCGGC
120541  AAGAAGACGGAGTCGGAGGGCTCATCGACCCTAATGTCCAGAGTACGGTGGCAAGGGCCG

120601  TAATTTTTGTATTTTTAGTAGAGATGAGGTTTTGCCCTGTTGGCCAGGCTGGTCTTGAAC
120601  ATTAAAAACATAAAAATCATCTCTACTCCAAAACGGGACAACCGGTCCGACCAGAACTTG

120661  TCTTGACCTCAGGTGATCCACCCGCCTTGGCTTCCCAAAGTGCTGGGATTACAGCCGTGA
120661  AGAACTGGAGTCCACTAGGTGGGCGGAACCGAAGGGTTTCACGACCCTAATGTCGGCACT

120721  GCCACCATGCCCAGCCAGTATACTTTTAGATAGTTATTTTTTCCTAACAAAATGGGATCA
120721  CGGTGGTACGGGTCGGTCATATGAAAATCTATCAATAAAAAAGGATTGTTTTACCCTAGT

120781  TATTGTACATAATTTTGTTACATGTTTTTTCACTGAATAAATGCTTTGCTTTTTTATATC
120781  ATAACATGTATTAAAACAATGTACAAAAAAGTGACTTATTTACGAAACGAAAAAATATAG

120841  AATAAGTAGTTTTCTCCAACACCATTTATAAAGTTTAATTGAATTCCATTATATTGATGT
120841  TTATTCATCAAAAGAGGTTGTGGTAAATATTTCAAATTAACTTAAGGTAATATAACTACA

120901  ACCATAATTTTTTTTAAAAAATTATATTATTGAACATTTATGTTGTTGGCAATTTTTTTTT
120901  TGGTATTAAAAAAATTTTTTAATATAATAACTTGTAAATACAACAACCGTTAAAAAAAAA

120961  CACTAGTATGGACAGTACATCCTCGATTACTTCCTAAGGATAAATCTGGGATCAAAGAGT
120961  GTGATCATACCTGTCATGTAGGAGCTAATGAAGGATTCCTATTTAGACCCTAGTTTCTCA

121021  ATACACGTCTATTTTTTCATACTGCTAACTTTGTTTTGCCTATTTTTTTTCTGAGCAATA
121021  TATGTGCAGATAAAAAAGTATGACGATTGAAACAAAACGGATAAAAAAAAGACTCGTTAT

121081  ATCTACAGTAATGCCTCATCAGGCAGTTAAGCAGGCATGCTAAATACCAGATAACACTGG
121081  TAGATGTCATTACGGAGTAGTCCGTCAATTCGTCCGTACGATTTATGGTCTATTGTGACC

121141  GCTACTCAGAAAAGAGATTATCCTGTAATTTGCAAGCATTTCACAAGTTAGAAGACTTTG
121141  CGATGAGTCTTTTCTCTAATAGGACATTAAACGTTCGTAAAGTGTTCAATCTTCTGAAAC

121201  TTGAAGCCAGTTACTATCCTTGTCAATTAGTAAAAATATCCAAGAGGGCCTGAGACTTAC
121201  AACTTCGGTCAATGATAGGAACAGTTAATCATTTTTATAGGTTCTCCCGGACTCTGAATG

121261  ATTGAAAAGTAAAATAAAGTTTGATCATTTGAGTGATGTCGGATGGCATTAAGTCATGAA
121261  TAACTTTTCATTTTATTTCAAACTAGTAAACTCACTACAGCCTACCGTAATTCAGTACTT

121321  ACTAAGTGAAACTTAAAGGTGGCTGAGAATGCAACAGGCCTCCAGAGAGCAATGGAGGTC
121321  TGATTCACTTTGAATTTCCACCGACTCTTACGTTGTCCGGAGGTCTCTCGTTACCTCCAG

121381  CTTGGGTAGTGATAAAAATGGCATTTCTGGCTGGGTGCAGTGGCTCACACCTATAATCCC
121381  GAACCCATCACTATTTTTACCGTAAAGACCGACCCACGTCACCGAGTGTGGATATTAGGG

121441  AACACTATGGGAGGCCAAGACAGGAGGATCGCTTGAGGCCAGGTGTTCAAGACTAGCCTA
121441  TTGTGATACCCTCCGGTTCTGTCCTCCTAGCGAACTCCGGTCCACAAGTTCTGATCGGAT

121501  GGCAACATAGTGAGAACCTTCCTCTAAAAAAAGGGAAATAAAATAAAAATAAAAGTTTAA
121501  CCGTTGTATCACTCTTGGAAGGAGATTTTTTTCCCTTTATTTTATTTTTATTTTCAAATT

121561  AAGAGAGTTTAAAAGGCAGTTTTAAAAGAGACTTACTCAGCAAGTTTTCTGCCAAATTTA
121561  TTCTCTCAAATTTTCCGTCAAAATTTTCTCTGAATGAGTCGTTCAAAAGACGGTTTAAAT

121621  CAGTATTAAAAAGAAAGAAAGTGTGTTTTTGTTTCTGTTTGTTTTTTTGAGACAGAGTCT
121621  GTCATAATTTTTCTTTCTTTCACACAAAAACAAAGACAAACAAAAAAACTCTGTCTCAGA
```

FIGURE 4 (cont'd)

```
121681  TACTCTGTTGCTCAGGCTGGGGTGCAGTGGTGCAATCTCAGCTCACTGCAACCCCTGCCT
121681  ATGAGACAACGAGTCCGACCCCACGTCACCACGTTAGAGTCGAGTGACGTTGGGGACGGA

121741  CCTGGGTTCAAGTGATTCTTGTGCCTCAGCCTCCCACACAGCAGGGATTATGGGCATGCA
121741  GGACCCAAGTTCACTAAGAACACGGAGTCGGAGGGTGTGTCGTCCCTAATACCCGTACGT

121801  CCATCATGCCTGGCTAATTTTTGTATTTTTAGTAGAGTCGAGGTTTCGCCATGTTGGCCA
121801  GGTAGTACGGACCGATTAAAAACATAAAAATCATCTCAGCTCCAAAGCGGTACAACCGGT

121861  GGCTGGTCTCAAACTCCTAGCCTCAAGTGATTTGCCCACCTTAGCCTCCCAAAGTGCTGA
121861  CCGACCAGAGTTTGAGGATCGGAGTTCACTAAACGGGTGGAATCGGAGGGTTTCACGACT

121921  GATTACAGGCATGAGCCACCGCACCCTGGCCTTGATTGTGGTTTTTTAATCAATGAGCTG
121921  CTAATGTCCGTACTCGGTGGCGTGGGACCGGAACTAACACCAAAAAATTAGTTACTCGAC

121981  CCATACTTCTGCTCTCCAGCAAAGGTTTTCTATGAGAATTTTCTGTAAGCTAAGTAAAGG
121981  GGTATGAAGACGAGAGGTCGTTTCCAAAAGATACTCTTAAAAGACATTCGATTCATTTCC

122041  AATATCTTAATATTGGAAGAAATAAAATTTAGTCAGAGTTGTTTTCTACCTTTGTACACA
122041  TTATAGAATTATAACCTTCTTTATTTTAAATCAGTCTCAACAAAAGATGGAAACATGTGT

122101  GCAACACTTACAAGTAATTTGAAGAGTAGCTAATAAAAATTTCTGCATGGATGGCTTAGG
122101  CGTTGTGAATGTTCATTAAACTTCTCATCGATTATTTTTAAAGACGTACCTACCGAATCC

122161  CATTTTAATTTTTTTAAGGTTGCCTGAGTAAATCTGAAAATAATAAATTATTACAGTTTT
122161  GTAAAATTAAAAAAAATTCCAACGGACTCATTTAGACTTTTATTATTTAATAATGTCAAAA

122221  GCTCAAATAAAGAACACAAATAGAACAAAAAATACATAACTGGCTAGATTACATGAAAGC
122221  CGAGTTTATTTCTTGTGTTTATCTTGTTTTTTATGTATTGACCGATCTAATGTACTTTCG

122281  CTTTTTCCAAAACAGAGGTTTAGAGGCCAGGGCAAATAGGCAGGAAAAATAAATAAAGGG
122281  GAAAAAGGTTTTGTCTCCAAATCTCCGGTCCCGTTTATCCGTCCTTTTTATTTATTTCCC

122341  CATCCAAATTGGAAAAGAAGAAGTTAAATTAGCTTTCTTAGCAGATGACATTATCTTACA
122341  GTAGGTTTAACCTTTTCTTCTTCAATTTAATCGAAAGAATCGTCTACTGTAATAGAATGT

122401  CCTAGAAAAACCTAAAGACTCCCTCAAAAAACCTGTAGAACTGATAAATGAATTCAGTAA
122401  GGATCTTTTTGGATTTCTGAGGGAGTTTTTTGGACATCTTGACTATTTACTTAAGTCATT

122461  AGTTGCAGGATATAAAATCAACATATAAAAATCAGTAGAATTTATATATGCTAACAGTGA
122461  TCAACGTCCTATATTTTAGTTGTATATTTTTAGTCATCTTAAATATATACGATTGTCACT

122521  ACAATCTGAAAAAGAAATCAAGAAAGCAATCCCATTTATAATAGCCACAAAAAGTATAAA
122521  TGTTAGACTTTTTCTTTAGTTCTTTCGTTAGGGTAAATATTATCGGTGTTTTTCATATTT

122581  ATACCTAGGAATCAATCTAACCAAAGACATGAGAGATCTATACAAGGAAAAGTATAAAAC
122581  TATGGATCCTTAGTTAGATTGGTTTCTGTACTCTCTAGATATGTTCCTTTTCATATTTTG

122641  TATGATTTAAAAAATTGAAGAAAACACACAAAGGTGGAAAGATATTCCATATTCATGGAT
122641  ATACTAAATTTTTTAACTTCTTTTGTGTGTTTCCACCTTTCTATAAGGTATAAGTACCTA

122701  TGGAAGAATTAATTTTGTTAAAATGACTATACTACACAAAGCAATTTACAGATTCAATGC
122701  ACCTTCTTAATTAAAACAATTTTACTGATATGATGTGTTTCGTTAAATGTCTAAGTTACG

122761  AATCCCTATTAAAATACTGATGACTTTCTTCACAGAAATAGAAAAAACAATCCTAAAATG
122761  TTAGGGATAATTTTATGACTACTGAAAGAAGTGTCTTTATCTTTTTTGTTAGGATTTTAC

122821  TATATGGAGCTAAAAAGACCCCAAATAGCCAAAGCAATCTTGAGCAAAAAGAACAAAGTT
122821  ATATACCTCGATTTTTCTGGGGTTTATCGGTTTCGTTAGAACTCGTTTTTCTTGTTTCAA

122881  AGAATCACCACACTACCTTACTTCAAAATTTATTATAAAACTATAATAACCAAAACAGCA
122881  TCTTAGTGGTGTGATGGAATGAAGTTTTAAATAATATTTTGATATTATTGGTTTTGTCGT

122941  TAGTACTGATATAAAAACAGACATAGTAGACCAGTGGGACAGAATAGAGAACCCAGATAC
122941  ATCATGACTATATTTTTGTCTGTATCATCTGGTCACCCTGTCTTATCTCTTGGGTCTATG

123001  CTAAGTCCACATGTTTACAACCAACTCATCTTTGACAAAGGTGCCAAGAACTGTAACGGG
123001  GATTCAGGTGTACAAATGTTGGTTGAGTAGAAACTGTTTCCACGGTTCTTGACATTGCCC
```

FIGURE 4 (cont'd)

```
123061  GAAGACAGTCTTTTCAATAAATGGTGCTGGGAAAATTGGATAATTATATGCAGAAGAATG
123061  CTTCTGTCAGAAAAGTTATTTACCACGACCCTTTTAACCTATTAATATACGTCTTCTTAC

123121  AAACTAGATTCGCTTCTCTTACCATACACAGAAATCAAATGAAAGTATATTAATGACTTG
123121  TTTGATCTAAGCGAAGAGAATGGTATGTGTCTTTAGTTTACTTTCATATAATTACTGAAC

123181  AATCTTAGACCTGAAACTATGAAATTACTTGAAGAAAACATTAAGGAAATGCTTCAGGAC
123181  TTAGAATCTGGACTTTGATACTTTAATGAACTTCTTTTGTAATTCCTTTACGAAGTCCTG

123241  ATTGGTCTAAGCAAAGATCTCATTCTGTCACTCAGGCTGGAGTGTAGTGGCATGAACCTG
123241  TAACCAGATTCGTTTCTAGAGTAAGACAGTGAGTCCGACCTCACATCACCGTACTTGGAC

123301  GCTCATTGTGGCATTGACCTCCTGGGCTCAAGCAATTCTCCCACCTTGGCTTTCCATGTA
123301  CGAGTAACACCGTAACTGGAGGACCCGAGTTCGTTAAGAGGGTGGAACCGAAAGGTACAT

123361  GCTAGAACCACAGGTGCATGCCACCACATTCCACTAATTTTTAAATTTTTTGTAGAGATG
123361  CGATCTTGGTGTCCACGTACGGTGGTGTAAGGTGATTAAAAATTTAAAAAACATCTCTAC

123421  GAGTCTCACCATCTTGCCCAGGCTGGTCTCGAACTCCTGAACTCAAGGGATCCTCCTGTC
123421  CTCAGAGTGGTAGAACGGGTCCGACCAGAGCTTGAGGACTTGAGTTCCCTAGGAGGACAG

123481  TCAGCCTCCCAAAGTGCTGAGATTACTAGCATAAGCCACTGTCCCTGGCCAGCGAAGATT
123481  AGTCGGAGGGTTTCACGACTCTAATGATCGTATTCGGTGACAGGGACCGGTCGCTTCTAA

123541  TCTTATGTAAGACCTGAAAAAGCATAGGCTATCAAAGCAAAAATAGGCAATTGGGATTAT
123541  AGAATACATTCTGGACTTTTTCGTATCCGATAGTTTCGTTTTTATCCGTTAACCCTAATA

123601  GTCAAGCTGGAAAGCTTCTCCACAGCAAAGAAAACAATCCACAAAGGGAATAGATTACCC
123601  CAGTTCGACCTTTCGAAGAGGTGTCGTTTCTTTTGTTAGGTGTTTCCCTTATCTAATGGG

123661  AAAGAATGGATGAAAATATTTGCAAACTATCAATCTGACAAGGGATTAATAACCAGAGTC
123661  TTTCTTACCTACTTTTATAAACGTTTGATAGTTAGACTGTTCCCTAATTATTGGTCTCAG

123721  TATAAGGAGCTCAAACAACTGTATAGGAAAAATCTAGTAATCTGGTTTTAAAATGGGCAT
123721  ATATTCCTCGAGTTTGTTGACATATCCTTTTTAGATCATTAGACCAAAATTTTACCCGTA

123781  GAGACCTGATTAGAAAACTCAACAGAAAAAAAAATTTGATTAAAAATGGATAAAAGATCT
123781  CTCTGGACTAATCTTTTGAGTTGTCTTTTTTTTTAAACTAATTTTTACCTATTTTCTAGA

123841  GAATGGACATTTCTCAAAAGAAGACAGACAAATGGCCAGGAGGTACATGAAAAAATGCGC
123841  CTTACCTGTAAAGAGTTTTCTTCTGTCTGTTTACCGGTCCTCCATGTACTTTTTTACGCG

123901  AACATCACTAATTATCAGAGAAATGCAAAACAAAACCACAATGTAAAATCCACTTGCCCT
123901  TTGTAGTGATTAATAGTCTCTTTACGTTTTGTTTTGGTGTTACATTTTAGGTGAACGGGA

123961  AGTTAAAATGGCTTGTATTAAAAAAAAAAAAAAAAAACAGGCAATAACAGATGCTGGGAGG
123961  TCAATTTTACCGAACATAATTTTTTTTTTTTTTTTGTCCGTTATTGTCTACGACCCTCC

124021  ATGTGGAGAAAGGGAACCATCATATATTGTTGGTTGAAATATAAATAGTACAGCCACTAT
124021  TACACCTCTTTCCCTTGGTAGTATATAACAACCAACTTTATATTTATCATGTCGGTGATA

124081  GGAGAACAGTATGGAGGGTCCTCCAAAAATAAAAGTAGAACTAGAACTTCCGTGTGATCC
124081  CCTCTTGTCATACCTCCCAGGAGGTTTTTATTTTCATCTTGATCTTGAAGGCACACTAGG

124141  AGCAATTCTACCACCGGGTATATATCCAAAAGAAAAAAATTAGTATATCAAAGAGATATC
124141  TCGTTAAGATGGTGGCCCATATATAGGTTTTCTTTTTTTAATCATATAGTTTCTCTATAG

124201  TGCACTCCCATGTTTATTGCAATAGTATTCACAATAGCCAAGATATGGAATTAACCTAAA
124201  ACGTGAGGGTACAAATAACGTTATCATAAGTGTTATCGGTTCTATACCTTAATTGGATTT

124261  TGTCTGTCAACAGATAAATGGATAAAGAGAATGTGGCATAAAGATATATATATATATATA
124261  ACAGACAGTTGTCTATTTACCTATTTCTCTTACACCGTATTTCTATATATATATATATAT

124321  TATACACACACACACACACACACACACAATGTGGTACTATTCAGCCATAAAAAAGAATGA
124321  ATATGTGTGTGTGTGTGTGTGTGTGTTACACCATGATAAGTCGGTATTTTTTCTTACT
```

FIGURE 4 (cont'd)

```
124381  AATCCTGTCATTTGCAGCAACATGTATGGAACTGGAGGCCATTTTGTTAAGTGAAATAAG
124381  TTAGGACAGTAAACGTCGTTGTACATACCTTGACCTCCGGTAAAACAATTCACTTTATTC

124441  CGAGGGACCGAAAGACAAACATCACATGTTCTCATGCAGGTGCTAAAAAAGTAGATCCTA
124441  GCTCCCTGGCTTTCTGTTTGTAGTGTACAAGAGTACGTCCACGATTTTTTCATCTAGGAT

124501  TGAAGACAGAGAGTAGATTAGAGGTTACCAGAGGCCGAGAAGGGGAGTGGGAAGTAGAGG
124501  ACTTCTGTCTCTCATCTAATCTCCAATGGTCTCCGGCTCTTCCCCTCACCCTTCATCTCC

124561  ATAAAGGAAAAAAACAAGAATATAAATGTATTTATTACCATTAAACTGAACATTTAAAAA
124561  TATTTCCTTTTTTTGTTCTTATATTTACATAAATAATGGTAATTTGACTTGTAAATTTTT

124621  TTGTAAAGATGGTAAATTATATATGTGTATTTTATCTTAATTAAAATTTTAAAAATCAGG
124621  AACATTTCTACCATTTAATATATACACATAAAATAGAATTAATTTTAAAATTTTTAGTCC

124681  AGTTTAGAAATTTTTAAGAGAAGTGTGCCCATAGTGATAAGAAAGAAAGACCCTTTTGGA
124681  TCAAATCTTTAAAAATTCTCTTCACACGGGTATCACTATTCTTTCTTTCTGGGAAAACCT

124741  AAATTTTCTTGAACAAATTATACAGTATTTTCTTAAACAAATTGTACATTCTGTAGCAGT
124741  TTTAAAAGAACTTGTTTAATATGTCATAAAAGAATTTGTTTAACATGTAAGACATCGTCA

124801  TAATACACCAAATAAGATTTCAGGGTTACCTCTTGGAGATTGTATGAAATAAAAACATTT
124801  ATTATGTGGTTTATTCTAAAGTCCCAATGGAGAACCTCTAACATACTTTATTTTTGTAAA

124861  AGGTGGGCATGATGGCTCATTCCTGTAATCCTAAAACTTTGGGAGGCTGATGCAGGAACA
124861  TCCACCCGTACTACCGAGTAAGGACATTAGGATTTTGAAACCCTCCGACTACGTCCTTGT

124921  ATACTTGAGGCTAGTACTTCAAGACGAGCCTAGGCAACATAGTAAGACCCAGTCTCCACA
124921  TATGAACTCCGATCATGAAGTTCTGCTCGGATCCGTTGTATCATTCTGGGTCAGAGGTGT

124981  AAACAACCAAAAAAAATAGCAAGTGTGGTGGCACACACCTGTAGTTCCAGCTACTCAGGA
124981  TTTGTTGGTTTTTTTTATCGTTCACACCACCGTGTGTGGACATCAAGGTCGATGAGTCCT

125041  GGCTGAGGCAGGAGGAACACTTGAGGCCAGGAATGTGAGACTGCAGTGGGCTATGATTGT
125041  CCGACTCCGTCCTCCTTGTGAACTCCGGTCCTTACACTCTGACGTCACCCGATACTAACA

125101  GCCACAGCACTCCACTCCAGCCTTCACTGTAGAAAGAGACCCTGTCTCGAAAGGAAGGAG
125101  CGGTGTCGTGAGGTGAGGTCGGAAGTGACATCTTTCTCTGGGACAGAGCTTTCCTTCCTC

125161  GGAGGGAGGGAGGGAGGGAGGGAGGGAGGGAATGAAGGAAGGAAGGAAGGAAGGAAGGAA
125161  CCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTTACTTCCTTCCTTCCTTCCTTCCTTCCTT

125221  GGAAGGAAATGTAAACTCATTTGATTTGGCTGTAAAGGATTAATAGAAAAGGAAAACGTA
125221  CCTTCCTTTACATTTGAGTAAACTAAACCGACATTTCCTAATTATCTTTTCCTTTTGCAT

125281  GTCACTAGGCTGAGAAAAATTGAACTCTATTGAATTGGTATAAAGATCTAAGTGGAGGTG
125281  CAGTGATCCGACTCTTTTTAACTTGAGATAACTTAACCATATTTCTAGATTCACCTCCAC

125341  TTTGGCAGACATGATCATGAAGCTTTCACTGAGGAAGGAAACTCTAGGAGGCAAAAGCAG
125341  AAACCGTCTGTACTAGTACTTCGAAAGTGACTCCTTCCTTTGAGATCCTCCGTTTTCGTC

125401  AAATGCAACTCATCAAAATTGAAATGGTGACAAAATATGTGCATTCAAACCTGAAAGTGA
125401  TTTACGTTGAGTAGTTTTAACTTTACCACTGTTTTATACACGTAAGTTTGGACTTTCACT

125461  CTGAAAGAGAAGCCCAAACACACAATTTTGGGGTAAGATGGGAAACCATTACCAAAGTTC
125461  GACTTTCTCTTCGGGTTTGTGTGTTAAAACCCCATTCTACCCTTTGGTAATGGTTTCAAG

125521  TCAGAAAGAGCATTTTTAAAAGTAAGATTATAGAAGCCAAAAGGATATGTAGTAGTCTGT
125521  AGTCTTTCTCGTAAAAATTTTCATTCTAATATCTTCGGTTTTCCTATACATCATCAGACA

125581  TCTCACACTGCTGTAACGAACTACCTGAGACTGGGTAGTTTGTAAAGAAAAGAGGTTTTA
125581  AGAGTGTGACGACATTGCTTGATGGACTCTGACCCATCAAACATTTCTTTTCTCCAAAAT

125641  TTGACGCTCAGTTCTGCAGGCCATACAGGAGACATGGCTGGGGAGGCCTCAGGAAACTTA
125641  AACTGCGAGTCAAGACGTCCGGTATGTCCTCTGTACCGACCCCTCCGGAGTCCTTTGAAT

125701  CAGTCATGGCGGAAGGCAAAGGGGAAGCCAGTACATATTACATGGTGGGAACAGGAGGAG
125701  GTCAGTACCGCCTTCCGTTTCCCCTTCGGTCATGTATAATGTACCACCCTTGTCCTCCTC
```

FIGURE 4 (cont'd)

```
125761  GAGAGGGAGGGTGGGAGGTGCCACTCACTTTTTAAAACAAACAGATCTCCTTAGAACTCA
125761  CTCTCCCTCCCACCCTCCACGGTGAGTGAAAAATTTTGTTTGTCTAGAGGAATCTTGAGT

125821  CTCACTATCATGAGAACAGCAAGGGGGAGGTTTGCTCCCATGATGCAATTACCTGACACC
125821  GAGTGATAGTACTCTTGTCGTTCCCCCTCCAAACGAGGGTACTACGTTAATGGACTGTGG

125881  AGGCCCTTCCTCCAACATGGGGGATTACAATTCCACATGAGATTTGGGTGGGGACACAAA
125881  TCCGGGAAGGAGGTTGTACCCCCTAATGTTAAGGTGTACTCTAAACCCACCCCTGTGTTT

125941  TTCAAACCATATCAGAATGTTAGTGACTAAAGTAAATGTTCTTGTTTAACTCTTTAATGT
125941  AAGTTTGGTATAGTCTTACAATCACTGATTTCATTTACAAGAACAAATTGAGAAATTACA

126001  GTATATTGAGTGCAAGCCATATGCAGGGGCTATTCTAGGTACTGGTGGTTTACTCAAGAA
126001  CATATAACTCACGTTCGGTATACGTCCCCGATAAGATCCATGACCACCAAATGAGTTCTT

126061  AAAGGGGGCAAAAAAATGAAGAAGGTGGCCAGGCATGGTGGCTTACGCCTGTAATCCCAA
126061  TTTCCCCCGTTTTTTTACTTCTTCCACCGGTCCGTACCACCGAATGCGGACATTAGGGTT

126121  CACTTTGGGAGACTGAGACGTGAGAATTGCTTGAGCCCAGGAGTTTGAGATCAGCCTGGT
126121  GTGAAACCCTCTGACTCTGCACTCTTAACGAACTCGGGTCCTCAAACTCTAGTCGGACCA

126181  AACATGGCAAATCCCTGTCTCTACAAAAAAATGCAAACAAACAAACAAACAAACAAACAA
126181  TTGTACCGTTTAGGGACAGAGATGTTTTTTTACGTTTGTTTGTTTGTTTGTTTGTTTGTT

126241  ATTAATCGGGCATGGTGTCACATGTCTGTAGTCCCAGCTACAGGTTGAGGTGGGAAGATG
126241  TAATTAGCCCGTACCACAGTGTACAGACATCAGGGTCGATGTCCAACTCCACCCTTCTAC

126301  CCTTAAGCCTAGGGAGGTTGAGGCTGCAGTCAGCCATGATAGCACCACTGCACTCCAGCC
126301  GGAATTCGGATCCCTCCAACTCCGACGTCAGTCGGTACTATCGTGGTGACGTGAGGTCGG

126361  TGGGCAACAGAGTGAGACCCCATCTGAAAAAAAAAAGAAGTCAAAACAAATGAAAAAAAAA
126361  ACCCGTTGTCTCACTCTGGGGTAGACTTTTTTTTTCTTCAGTTTTGTTTACTTTTTTTTT

126421  AAAAAGTCAAAACAAATATTCAGTGTAAATGTGAGACCATGTAAAAAAGTTTAAAAGTAA
126421  TTTTTCAGTTTTGTTTATAAGTCACATTTACACTCTGGTACATTTTTTCAAATTTTCATT

126481  AATATATAGTTTAAGCAGAGGATGGGAGACGGGGAAAAAACCTTTTCTCCCTAACTTTAT
126481  TTATATATCAAATTCGTCTCCTACCCTCTGCCCCTTTTTTGGAAAAGAGGGATTGAAATA

126541  ATACTCCCTCTATAAATTAGTGCTTATGGCCAATAAGAAATGAATAATTTCTTTTTTCAC
126541  TATGAGGGAGATATTTAATCACGAATACCGGTTATTCTTTACTTATTAAAGAAAAAAGTG

126601  CGTTGGTAGTAAAGTATAATGGTAGTTAACACAAAAAACTAAATGATATATTCAGAACCT
126601  GCAACCATCATTTCATATTACCATCAATTGTGTTTTTTGATTTACTATATAAGTCTTGGA

126661  CATACTAAACAAAGAAACAAAGGCATGCAATCACAGGACAGGTGGAAACAGTGGTTCTGA
126661  GTATGATTTGTTTCTTTGTTTCCGTACGTTAGTGTCCTGTCCACCTTTGTCACCAAGACT

126721  ACTAACCACAAACCAGAACTACCAAATGTGCTACAAATTGTTTAGTAACAAAGGACCATC
126721  TGATTGGTGTTTGGTCTTGATGGTTTACACGATGTTTAACAAATCATTGTTTCCTGGTAG

126781  CCTTCAGAATTGAAGAGCACGTCTCAGTTATCTCATGCCTATGTGCATGGAACTTAAAGT
126781  GGAAGTCTTAACTTCTCGTGCAGAGTCAATAGAGTACGGATACACGTACCTTGAATTTCA

126841  CTCAAAATTACTCTTACAAGATGGCAAATTGAGGCTGTCAGCTAAGCTCCTAGATTGTCT
126841  GAGTTTTAATGAGAATGTTCTACCGTTTAACTCCGACAGTCGATTCGAGGATCTAACAGA

126901  GTTTCTCCAAGAAAAATAAATTTCAGTAATTATTAAGGAAAAAGCATAGAATTGCATCAT
126901  CAAAGAGGTTCTTTTTATTTAAAGTCATTAATAATTCCTTTTTCGTATCTTAACGTAGTA

126961  GTTTTACATTTCATTACTTTTGGCATGACATGGGAGCATGGATGTTTTCTACAAAGTGCT
126961  CAAAATGTAAAGTAATGAAAACCGTACTGTACCCTCGTACCTACAAAAGATGTTTCACGA

127021  CTGATAAATTGAAGAAGAGGCCAACAACAGCCACCGAAAAAACCTTCAAGATTCTTGCCA
127021  GACTATTTAACTTCTTCTCCGGTTGTTGTCGGTGGCTTTTTTGGAAGTTCTAAGAACGGT
```

FIGURE 4 (cont'd)

```
127081  GTATGGCTCAGGGGAAAAATTCCCTCATAACCACAGATTTGACTTTTCAACTTCTGTACT
127081  CATACCGAGTCCCCTTTTTAAGGGAGTATTGGTGTCTAAACTGAAAAGTTGAAGACATGA

127141  TCTGCATATGAGTAATAATCTCTGAGGTTAAGTATTCATCCCATTTAGCTAATTGTCCAC
127141  AGACGTATACTCATTATTAGAGACTCCAATTCATAAGTAGGGTAAATCGATTAACAGGTG

127201  ATAACTCCCTGGGTCTTAGATTGCTCTATAAATACCTCCTAAGAGAAGAATTTTTCAAAT
127201  TATTGAGGGACCCAGAATCTAACGAGATATTTATGGAGGATTCTCTTCTTAAAAAGTTTA

127261  GACAGTTTCGTTTTTCCTTTTTCTTCTTTTAAAGCAACAAAGAAAAGCACGTCTGTGTTT
127261  CTGTCAAAGCAAAAAGGAAAAAGAAGAAAATTTCGTTGTTTCTTTTCGTGCAGACACAAA

127321  ATATTTAACCAAACTTAAACTTGGAGAGCATTGGCCGTGTTGTCGACGTGTTTCTTTTCT
127321  TATAAATTGGTTTGAATTTGAACCTCTCGTAACCGGCACAACAGCTGCACAAAGAAAAGA

127381  CTACTGTTAAATTCCTTACTGTTCCAGTTACTCACCATCATACTCACCTTTGCCTCTGCC
127381  GATGACAATTTAAGGAATGACAAGGTCAATGAGTGGTAGTATGAGTGGAAACGGAGACGG

127441  ACTTGGGGTCCTCTGTTCTGCAGTTGTGTTCTGCTCTCTGATGCCTCCAGGTTTTGCTGC
127441  TGAACCCCAGGAGACAAGACGTCAACACAAGACGAGAGACTACGGAGGTCCAAAACGACG

127501  ACCTTCTAGAAATTGCTCACCTGGATAAGGCCTTGGCTGTCCTTTATGACTCATCTCGTT
127501  TGGAAGATCTTTAACGAGTGGACCTATTCCGGAACCGACAGGAAATACTGAGTAGAGCAA

127561  CAGAAGCCTTTTGAATCTCTCTTGAAGAATATTCCCTTTGCTCAGTCCAAATACACATTC
127561  GTCTTCGGAAAACTTAGAGAGAACTTCTTATAAGGGAAACGAGTCAGGTTTATGTGTAAG

127621  ACTCACTTGAATGCTGCTTCTTCCTCCAGTTTCACAGTGCTGGGGATGATCCCGTAAGGC
127621  TGAGTGAACTTACGACGAAGAAGGAGGTCAAAGTGTCACGACCCCTACTAGGGCATTCCG

127681  CTTTAAATTACATTATAATTAGTCATGCACTTCTCTCTGCTCTGGGGCTGTTTGAATCGT
127681  GAAATTTAATGTAATATTAATCAGTACGTGAAGAGAGACGAGACCCCGACAAACTTAGCA

127741  TCACTCTCTTTATACCTCCAGTGTTCAGCACAGAGCCTAGGGTATAATGCGTATGCAGTA
127741  AGTGAGAGAAATATGGAGGTCACAAGTCGTGTCTCGGATCCCATATTACGCATACGTCAT

127801  ATGTTTCTTGAAGGAGTGAATGGGCAAGTGAAACACCAAACAAATACTAAAGAGCAAAAC
127801  TACAAAGAACTTCCTCACTTACCCGTTCACTTTGTGGTTTGTTTATGATTTCTCGTTTTG

127861  ATGTATTGCTCTTGTATATTAACACATTCCGAGAAGAGATGGGAGGAAGGTCTTGAATAA
127861  TACATAACGAGAACATATAATTGTGTAAGGCTCTTCTCTACCCTCCTTCCAGAACTTATT

127921  TATCAAATTATTTCTTAAGAGATCTTGATTTAGATTTTTTTCAAAATGAGATACATGTTG
127921  ATAGTTTAATAAAGAATTCTCTAGAACTAAATCTAAAAAAAGTTTTACTCTATGTACAAC

127981  ACACTGCTGTCAACAAATAATCACATAAAAGCTAGAGCCTGGCTTTTAAGTATCAGGGGG
127981  TGTGACGACAGTTGTTTATTAGTGTATTTTCGATCTCGGACCGAAAATTCATAGTCCCCC

128041  AGCATGTGTGAATAATCAGCATTGCTGCTGTTGAAGCTCCATAAAGTGGGAAGCAGCGAC
128041  TCGTACACACTTATTAGTCGTAACGACGACAACTTCGAGGTATTTCACCCTTCGTCGCTG

128101  TGCAAAGATCCAGAGATGGACATCTGAAGCATAGGAAAGGGGCAGCTTGAATGAGATAGG
128101  ACGTTTCTAGGTCTCTACCTGTAGACTTCGTATCCTTTCCCCGTCGAACTTACTCTATCC

128161  AGCAGGATCTGTATTGATTGCAGGTCACCAGATAGAATTAAAAAAATCTATGTTCTCACT
128161  TCGTCCTAGACATAACTAACGTCCAGTGGTCTATCTTAATTTTTTTAGATACAAGAGTGA

128221  AGTGGCTACAGTTTTCAAGGACAAATGACTATTAGTGGTTATTGTATTTTTCTTGTTTTC
128221  TCACCGATGTCAAAAGTTCCTGTTTACTGATAATCACCAATAACATAAAAAGAACAAAAG

128281  AATGGAGCAGCTCCATGACTTAGTTACATAAGAAGTTAAAGGCTCCACCCCCTTCCCTAC
128281  TTACCTCGTCGAGGTACTGAATCAATGTATTCTTCAATTTCCGAGGTGGGGGAAGGGATG

128341  CAATGTCTTAAGCTGTCCTCCGTGTGCATACAAACTTTGTTTCCAGACTGCCCTCGTGTG
128341  GTTACAGAATTCGACAGGAGGCACACGTATGTTTGAAACAAAGGTCTGACGGGAGCACAC

128401  ACTCAGTCCCTGGACCAGCTTCCCAGCCCCTTCTGTTTCCAGCACATCCTCCTCCACGTG
128401  TGAGTCAGGGACCTGGTCGAAGGGTCGGGGAAGACAAAGGTCGTGTAGGAGGAGGTGCAC
```

FIGURE 4 (cont'd)

```
128461  TTACAGGATTACAAAGCTTTGCACAGCCTTTCAATAGGGTTTAGAGCAAGAATGACAGAT
128461  AATGTCCTAATGTTTCGAAACGTGTCGGAAAGTTATCCCAAATCTCGTTCTTACTGTCTA

128521  TGCTACCTGGTAGGCAGCACTGACTATAGTTTTATTGGCGTTGCGTAGAGTGACTGCATG
128521  ACGATGGACCATCCGTCGTGACTGATATCAAAATAACCGCAACGCATCTCACTGACGTAC

128581  ACATGCAGGTTGTGCTTTTGTTTTGCCAGGAATTTAAAGGCATTTAGCAAAGCCGTATGA
128581  TGTACGTCCAACACGAAAACAAAACGGTCCTTAAATTTCCGTAAATCGTTTCGGCATACT

128641  AATTATCTGCCTTATGTACCTCGAATGTGAGGTAATGAACTTGTCATTAGTCGTGTTTTA
128641  TTAATAGACGGAATACATGGAGCTTACACTCCATTACTTGAACAGTAATCAGCACAAAAT

128701  GTTTAGTGTTCTCGCTGGAATTTGCAAAAGATTAAAATTCTGCACATTTCATTAAGTCTT
128701  CAAATCACAAGAGCGACCTTAAACGTTTTCTAATTTTAAGACGTGTAAAGTAATTCAGAA

128761  CCACATGGGCTCAAAATACCATCCCGCAGGCTCGACCTGACTTGCCACGTTGCTTTGACA
128761  GGTGTACCCGAGTTTTATGGTAGGGCGTCCGAGCTGGACTGAACGGTGCAACGAAACTGT

128821  CCATCTCTGATTCCAGTGAACAAAACTAAGTAAGTGTGACTTACAGCCGCCCTACCAAAC
128821  GGTAGAGACTAAGGTCACTTGTTTTGATTCATTCACACTGAATGTCGGCGGGATGGTTTG

128881  ATTGAAGGTGTGAGCCTGCACTGTTGTTTTGTAATGACAATGACAAGAGTATAATGACAA
128881  TAACTTCCACACTCGGACGTGACAACAAAACATTACTGTTACTGTTCTCATATTACTGTT

128941  GAGTATAATGACATCCTAGTCCATGAAGACATGTTTTCTCTTTTCCCTCCCTCTATGTAT
128941  CTCATATTACTGTAGGATCAGGTACTTCTGTACAAAAGAGAAAAGGGAGGGAGATACATA

129001  ACTTACTTACTCTTAAGGAAATAATATAATTTCTATTTCTGCTTAAATCTGCCATCTCTA
129001  TGAATGAATGAGAATTCCTTTATTATATTAAAGATAAAGACGAATTTAGACGGTAGAGAT

129061  CTTAGAATTATTTTCCCTTGAAAAATTAAATTGCATTATCCATCAGCCAAATGGAAAATC
129061  GAATCTTAATAAAAGGGAACTTTTTAATTTAACGTAATAGGTAGTCGGTTTACCTTTTAG

129121  TGATTACTAAATGCCTACTTCTAGACATTATGAATTACCTGTAATATCACATTTCTTGGG
129121  ACTAATGATTTACGGATGAAGATCTGTAATACTTAATGGACATTATAGTGTAAAGAACCC

129181  CACTTCTCAGAATGCTATGATTTTACATAAAAGCATAAATTACAGAATGAGAAGAGTAAA
129181  GTGAAGAGTCTTACGATACTAAAATGTATTTTCGTATTTAATGTCTTACTCTTCTCATTT

129241  GCTTAGTGCATATTATGCAGGTTCGCAGAGTGATATGTGGCCTACGACTGTCATTGTAAT
129241  CGAATCACGTATAATACGTCCAAGCGTCTCACTATACACCGGATGCTGACAGTAACATTA

129301  CAAAATTGATTTTTGATCATTTATAAAACTAAAACAACTAATTGAAAGCAAATTTACTTT
129301  GTTTTAACTAAAAACTAGTAAATATTTTGATTTTGTTGATTAACTTTCGTTTAAATGAAA

129361  AACTGATGAAATTCAAGATGAAGTGACAGTAGTCATGAATCACAGTTCTGTAGATCTGCT
129361  TTGACTACTTTAAGTTCTACTTCACTGTCATCAGTACTTAGTGTCAAGACATCTAGACGA

129421  AAAATGCCAATGAACTGTTATTAATGGTGAAAGTGTTTAATATATTAATGGGAATCCAAG
129421  TTTTACGGTTACTTGACAATAATTACCACTTTCACAAATTATATAATTACCCTTAGGTTC

129481  TAAGGAAAATAAAACACATATTTTGCCCTTACCTTAAGTCATGTTGACATCATACATAAA
129481  ATTCCTTTTATTTTGTGTATAAAACGGGAATGGAATTCAGTACAACTGTAGTATGTATTT

129541  TCAGGTAGAATTTGAAGATGTTTCATCCAGCATTCATCATGAAGCAACAGAGTCAATAGC
129541  AGTCCATCTTAAACTTCTACAAAGTAGGTCGTAAGTAGTACTTCGTTGTCTCAGTTATCG

129601  GGGCAGTTACAGTCTGTAGGCAATGCTTTCTGGTGTGGATTATGCAAGAAACTTACAGTC
129601  CCCGTCAATGTCAGACATCCGTTACGAAAGACCACACCTAATACGTTCTTTGAATGTCAG

129661  ATTTGGCTTCCTTTATTCCCCTTTGCATGTAAACTACCAGGTGGTAGGTAGACCCATTCT
129661  TAAACCGAAGGAAATAAGGGGAAACGTACATTTGATGGTCCACCATCCATCTGGGTAAGA

129721  TGGCTGCCAGTATTTTTACCACTTGGATAGATTCTGCTCTCTGAGGTCTTTAAGACAAAC
129721  ACCGACGGTCATAAAAATGGTGAACCTATCTAAGACGAGAGACTCCAGAAATTCTGTTTG
```

FIGURE 4 (cont'd)

```
129781  ACAGCACATGTACATTGTGATATAGGAGAAAGAACATCAGACAGAGGCAGAGAACTGGGC
129781  TGTCGTGTACATGTAACACTATATCCTCTTTCTTGTAGTCTGTCTCCGTCTCTTGACCCG

129841  TTGTCAAAATTAAGGGTGACCTAGAGCCTTAGTGTCCTTATATGTAAGATGAGAAGCTTG
129841  AACAGTTTTAATTCCCACTGGATCTCGGAATCACAGGAATATACATTCTACTCTTCGAAC

129901  GATTAAAATTCATTGCTCCTTACACTTGATTTTTATGCAGCAGAACTTTCTTCAAAAGAG
129901  CTAATTTTAAGTAACGAGGAATGTGAACTAAAAATACGTCGTCTTGAAAGAAGTTTTCTC

129961  ATCACTCCTGGAATCCCAATATACAAAATGGATAAAAGTAGGATTGTTCTGAAGTAAAGA
129961  TAGTGAGGACCTTAGGGTTATATGTTTTACCTATTTTCATCCTAACAAGACTTCATTTCT

130021  AGTAAAATCCAGAGCCCCACTTATTTGAGTCATCTCCTACCTTTTAGATGGTTATTAACA
130021  TCATTTTAGGTCTCGGGGTGAATAAACTCAGTAGAGGATGGAAAATCTACCAATAATTGT

130081  TTGTTTCTAGTTCAGAAATTCTGATGTTTTATATTAATTTAAACTAAAGTTAAATTAGGA
130081  AACAAAGATCAAGTCTTTAAGACTACAAAATATAATTAAATTTGATTTCAATTTAATCCT

130141  CTAGAAGAGCCATTTCGTGTGAAAATAGACATAGTCTGAGTGATTTAGGCAAACCCATAG
130141  GATCTTCTCGGTAAAGCACACTTTTATCTGTATCAGACTCACTAAATCCGTTTGGGTATC

130201  TGACAGATTTAGTGGCCTGATACACTGGCTTGGAAAGATCTGTTCCAAGTCAGTCTCAAT
130201  ACTGTCTAAATCACCGGACTATGTGACCGAACCTTTCTAGACAAGGTTCAGTCAGAGTTA

130261  ACACAAGGCTCGGTGAACACATGGCGAGCAGTCTTGGTGGTGGTCAAGAACATGGAATTT
130261  TGTGTTCCGAGCCACTTGTGTACCGCTCGTCAGAACCACCACCAGTTCTTGTACCTTAAA

130321  GCTGTGAGAGAGAGCTTGGTTAGGGTCTGGGCTCTTCCATCTGTTAACTGCAAACTTGAG
130321  CGACACTCTCTCTCGAACCAATCCCAGACCCGAGAAGGTAGACAATTGACGTTTGAACTC

130381  CAAGTTATTCAATCTCCCTAATATTCAGTTTTCTCATTCATAAAATGGATTGATAATGGT
130381  GTTCAATAAGTTAGAGGGATTATAAGTCAAAAGAGTAAGTATTTTACCTAACTATTACCA

130441  ACCTTCTATGTACAGTTACTGTGGGTATTAAATGAGGTAATAAGCCTGTAAAACTCCTGG
130441  TGGAAGATACATGTCAATGACACCCATAATTTACTCCATTATTCGGACATTTTGAGGACC

130501  CATGATGACATTGGCCTACATTAATTGCTCACAGAGTGAGAACTTATTATAAACACAGGC
130501  GTACTACTGTAACCGGATGTAATTAACGAGTGTCTCACTCTTGAATAATATTTGTGTCCG

130561  ATTTTTGTTGAATGACTGACTGATTGAATGAATGAATATAGAACTAGCTTTCTATTTGTG
130561  TAAAAACAACTTACTGACTGACTAACTTACTTACTTATATCTTGATCGAAAGATAAACAC

130621  ATGTATGGAATATCTCTATTTCATTTAATTTCATTTTGAATATTAAAATTCAGGCTTGCT
130621  TACATACCTTATAGAGATAAAGTAAATTAAAGTAAAACTTATAATTTTAAGTCCGAACGA

130681  TTTAAACCAAAGTATTTGCTGATGCAAAGAATGAAATTCTGACTATGATACAGTTACATA
130681  AAATTTGGTTTCATAAACGACTACGTTTCTTACTTTAAGACTGATACTATGTCAATGTAT

130741  ACTTTTGAAGAAATAAGAAAAATGCTAACACAGCATGAACAACTTTGATTTAGAGAACAG
130741  TGAAAACTTCTTTATTCTTTTTACGATTGTGTCGTACTTGTTGAAACTAAATCTCTTGTC

130801  AGTCTCTGTTAAAATCCATGGGACTTGTCAACTTCATATTTTTGTTTATGAGCATAAAAC
130801  TCAGAGACAATTTTAGGTACCCTGAACAGTTGAAGTATAAAAACAAATACTCGTATTTTG

130861  ACTCTTTAGTGCATGTCGGAATCATTTGATGCTTGCTGAATCACAGGTTGCTGGAACCCA
130861  TGAGAAATCACGTACAGCCTTAGTAAACTACGAACGACTTAGTGTCCAACGACCTTGGGT

130921  TCCCCAGATTCTGATTCAGTCTGAGATGAGGTCTGGGAATTTGCATTTTTAACTTGTTGT
130921  AGGGGTCTAAGACTAAGTCAGACTCTACTCCAGACCCTTAAACGTAAAAATTGAACAACA

130981  CTGGTGATGCTGATGCTAGTGGTCTGAGAATACACTTTGAGGATCACTTTTTCACAGTAA
130981  GACCACTACGACTACGATCACCAGACTCTTATGTGAAACTCCTAGTGAAAAAGTGTCATT

131041  AAGTCACAGTCATTAAGCAAGTTAACTCAATGAGAAATGATAATGAAAAGGCATGTACCT
131041  TTCAGTGTCAGTAATTCGTTCAATTGAGTTACTCTTTACTATTACTTTTCCGTACATGGA

131101  AAATAGATAAAGGACCTATACCAGCTATGAATTTGAGTTCAGCAATAATTGTAAAATTTG
131101  TTTATCTATTTCCTGGATATGGTCGATACTTAAACTCAAGTCGTTATTAACATTTTAAAC
```

FIGURE 4 (cont'd)

```
131161  TAATTCTTTTCAGGTATTTGGTTTTTAAAAAACTCCCTATGTATGCTTTTATCTTATGTT
131161  ATTAAGAAAAGTCCATAAACCAAAAATTTTTTGAGGGATACATACGAAAATAGAATACAA

131221  TTATATGGTGTCCCTCTTTTCCTCTCTTTCTCTCTAATCCAGTTTTTAATGTGAACATGT
131221  AATATACCACAGGGAGAAAAGGAGAGAAAGAGAGATTAGGTCAAAAATTACACTTGTACA

131281  TTAATGTGAAAATGTTGTTCAAAGCAGAAGCGAACCTTTTGTGTAACCTTGGCGGTATTA
131281  AATTACACTTTTACAACAAGTTTCGTCTTCGCTTGGAAAACACATTGGAACCGCCATAAT

131341  AGTTTGTTTGTAAGCTATTTCTGCCCTGTTAGCTTCTGTACTGAAACACGTTTTCTTGCT
131341  TCAAACAAACATTCGATAAAGACGGGACAATCGAAGACATGACTTTGTGCAAAAGAACGA

131401  TTTGTTGCAGTATGGAAGAGAAGAGAGCGACTGGACAATTGTGCTATCCTGAATGGAAAA
131401  AAACAACGTCATACCTTCTCTTCTCTCGCTGACCTGTTAACACGATAGGACTTACCTTTT

131461  TAGAGGATACAATGGAAAATAGAGGATACCAACTGTATGCTACTGGGACAGACTGTTGCA
131461  ATCTCCTATGTTACCTTTTATCTCCTATGGTTGACATACGATGACCCTGTCTGACAACGT

131521  TTTGAATTGTGATAGATTTCTTTGGCTACCTGTGCATAATGTAGTTTGTAGTATCAATGT
131521  AAACTTAACACTATCTAAAGAAACCGATGGACACGTATTACATCAAACATCATAGTTACA

131581  GTTACAAGAGTGATTGTTTCTTCATGCCAGAGAAAATGAATTGCAATCATCAAATGGTGT
131581  CAATGTTCTCACTAACAAAGAAGTACGGTCTCTTTTACTTAACGTTAGTAGTTTACCACA

131641  TTCATAACTTGGTAGTAGTAACTTACCTTACCTTACCTAGAAAAACATTAATGTAAGCCA
131641  AAGTATTGAACCATCATCATTGAATGGAATGGAATGGATCTTTTTGTAATTACATTCGGT

131701  TATAACATGGGATTTTCCTCAATGATTTTAGTGCCTCCTTTTGTACTTCACTCAGATACT
131701  ATATTGTACCCTAAAAGGAGTTACTAAAATCACGGAGGAAAACATGAAGTGAGTCTATGA

131761  AAATAGTAGTTTATTCTTTAATATAAGTTACATTCTGCTCCTCAAACAAATGCAATTTTT
131761  TTTATCATCAAATAAGAAATTATATTCAATGTAAGACGAGGAGTTTGTTTACGTTAAAAA

131821  TGTGTGTGTTTGAAAGCTAATTTGAGAAAATTTCATAGGTTACATTTCCTGCAGCCTATC
131821  ACACACACAAACTTTCGATTAAACTCTTTTAAAGTATCCAATGTAAAGGACGTCGGATAG

131881  TTTATCCACAGAAAGTGTTTTCTTTTTTTTAAATCAAGACTTTTAAAACTGGATTTCCTC
131881  AAATAGGTGTCTTTCACAAAAGAAAAAAAATTTAGTTCTGAAAATTTTGACCTAAAGGAG

131941  CCATCACTGTTTTTTGAAGGTCCTCCAAGTCCGTGTTAAGGTAAATATCTGTTTTCTTCC
131941  GGTAGTGACAAAAAACTTCCAGGAGGTTCAGGCACAATTCCATTTATAGACAAAAGAAGG

132001  TGATGTCACAGCCTGAGCATACTCTGTGCATTAGGAAGACCTGAGTGCATTTCCCACCAT
132001  ACTACAGTGTCGGACTCGTATGAGACACGTAATCCTTCTGGACTCACGTAAAGGGTGGTA

132061  TGTCCTTTCCACATTATGTTGTAGCTGGCTGGCTGTCAGGCGACTACAAGACTGAGGGTC
132061  ACAGGAAAGGTGTAATACAACATCGACCGACCGACAGTCCGCTGATGTTCTGACTCCCAG

132121  TTGTGCCTTATAGATCTTTGTATCCCCCATGGCTGACATATAGTAGGTACTCAGTAAATG
132121  AACACGGAATATCTAGAAACATAGGGGGTACCGACTGTATATCATCCATGAGTCATTTAC

132181  GTTTTATAATGAATCAGTGAACATTTTGCTTCTATAGAAGTGTACCTTCTTTGTTTCTAT
132181  CAAAATATTACTTAGTCACTTGTAAAACGAAGATATCTTCACATGGAAGAAACAAAGATA

132241  ATTATGAAACCTCTTTATTAGAATTTGTGATTGATTCTGACAGTGTATAGATTTACCTTA
132241  TAATACTTTGGAGAAATAATCTTAAACACTAACTAAGACTGTCACATATCTAAATGGAAT

132301  TATTGTCTTTATTTTCCATGAGCTACTAAGTCATTAGAGATACTCTGAAGCATAGTTAGT
132301  ATAACAGAAATAAAAGGTACTCGATGATTCAGTAATCTCTATGAGACTTCGTATCAATCA

132361  TTAGGAAATCACTTCATATTGATTGTATTAGAATTATCTTGGAATTGAAGATATATCCCT
132361  AATCCTTTAGTGAAGTATAACTAACATAATCTTAATAGAACCTTAACTTCTATATAGGGA

132421  AGAGCAGGGGACCCCAACCCCCAGGCCATGGGCCACACAGCAGGAAGAGGTGAGTGGTGG
132421  TCTCGTCCCCTGGGGTTGGGGGTCCGGTACCCGGTGTGTCGTCCTTCTCCACTCACCACC
```

FIGURE 4 (cont'd)

```
132481  GCCATTGAGGAGCTTCATCTGTATTTATGGCTACTTCCCATCACTCGAATTACCACCTGA
132481  CGGTAACTCCTCGAAGTAGACATAAATACCGATGAAGGGTAGTGAGCTTAATGGTGGACT

132541  ACTCCACCTCTTGTCAGCTCAGTGGCAGCATTAGATTCTCATAGGAGCACAAATCCTATT
132541  TGAGGTGGAGAACAGTCGAGTCACCGTCGTAATCTAAGAGTATCCTCGTGTTTAGGATAA

132601  GTGAACTCTGCATGCAAGGGATCTAGGCTATGCGCTCCTTATGAGAATCTAATGCTTGAT
132601  CACTTGAGACGTACGTTCCCTAGATCCGATACGCGAGGAATACTCTTAGATTACGAACTA

132661  GACCTGAGGTGTAACAGTTTCATCCTGAAACCACCCTTCACCCTGCAGTCTGTGGAAAAA
132661  CTGGACTCCACATTGTCAAAGTAGGACTTTGGTGGGAAGTGGGACGTCAGACACCTTTTT

132721  TTGTCTTCCACAAAACTGGTCCCTGGTGCCAAAAATGTTGGGGACCACTGCTCTAGAGAG
132721  AACAGAAGGTGTTTTGACCAGGGACCACGGTTTTTACAACCCCTGGTGACGAGATCTCTC

132781  AGGTCATGATATCATACCAACCAAATGGAAATGACAAATGTTTTATGTCAAGTGTTAATT
132781  TCCAGTACTATAGTATGGTTGGTTTACCTTTACTGTTTACAAAATACAGTTCACAATTAA

132841  GCAGAAATAAATCTTTTTTTTTTTTTTTTGGTAGAAAACAAAGAGGCATACTCTGATTTT
132841  CGTCTTTATTTAGAAAAAAAAAAAAAAAACCATCTTTTGTTTCTCCGTATGAGACTAAAA

132901  TATACTCTGTTTTTGCAGGTGCTCTTTTCTTTGAATGGAGATTTGATGAGCAAGTGGTTA
132901  ATATGAGACAAAAACGTCCACGAGAAAAGAAACTTACCTCTAAACTACTCGTTCACCAAT

132961  GGATGCAGGGAGAGCTACTATGGGTGATATTTTCCTTGTTTAGGAGCTGTGAGTTAAAAT
132961  CCTACGTCCCTCTCGATGATACCCACTATAAAAGGAACAAATCCTCGACACTCAATTTTA

133021  TGTATCCTTTGTGGTTTATCTAAGGAAAGTCAAATCTTGACAGAAAACATTTTTCCTTGG
133021  ACATAGGAAACACCAAATAGATTCCTTTCAGTTTAGAACTGTCTTTTGTAAAAAGGAACC

133081  AAGGTCAACTCTCAGACATTGTATTTTGGTTTCCCTCAGTCCTCATAACTTCCTTCTTGC
133081  TTCCAGTTGAGAGTCTGTAACATAAAACCAAAGGGAGTCAGGAGTATTGAAGGAAGAACG

133141  TGAACATATTTTATTCTCTTTTCAGAGAAGGAAAATAAAAAGGATTCTAAAAGTTTGATG
133141  ACTTGTATAAAATAAGAGAAAAGTCTCTTCCTTTTATTTTTCCTAAGATTTTCAAACTAC

133201  CATTGGAAAAATTTCCTTGAGGCATTTAGCAACACATAGAAAATGGGCTTTGATTCTTTT
133201  GTAACCTTTTTAAAGGAACTCCGTAAATCGTTGTGTATCTTTTACCCGAAACTAAGAAAA

133261  CCAAAACTTTTAGCCATAGGGTCTTTTATAGACAGGGATAGTAAAATGAAAATTGAGAAA
133261  GGTTTTGAAAATCGGTATCCCAGAAAATATCTGTCCCTATCATTTTACTTTTAACTCTTT

133321  TATAAGATGAAAAGGAATGATAAAAATATCTTTTAGGGGGCTTTTAATTGGTGATCTGAA
133321  ATATTCTACTTTTCCTTACTATTTTTATAGAAAATCCCCCGAAAATTAACCACTAGACTT

133381  ATCTTGGGAGAAGCTGTTCTTTTCAGGCCTGAGGTGCTCTTGACTGTCGCCTGCGCACTG
133381  TAGAACCCTCTTCGACAAGAAAAGTCCGGACTCCACGAGAACTGACAGCGGACGCGTGAC

133441  TGTACCCCGAGCAACATTCTAAGGGTGTGCTTTCGCCTTGGCTAACTCCTTTGACCTCAT
133441  ACATGGGGCTCGTTGTAAGATTCCCACACGAAAGCGGAACCGATTGAGGAAACTGGAGTA

133501  TCTTCATATAGTAGTCTAGGAAAAAGTTGCAGGTAATTTAAACTGTCTAGTGGTACATAG
133501  AGAAGTATATCATCAGATCCTTTTTCAACGTCCATTAAATTTGACAGATCACCATGTATC

133561  TAACTAAATTTCTATTCCTATGAGAAATGAGAATTATTTATTTGCCATCAACACATTTTA
133561  ATTGATTTAAAGATAAGGATACTCTTTACTCTTAATAAATAAACGGTAGTTGTGTAAAAT

133621  TACTTTGCATCTCCAAATTTATTGTGGCGAGACTTGTCCATTGTGAAAGTTAGAGAACAT
133621  ATGAAACGTAGAGGTTTAAATAACACCGCTCTGAACAGGTAACACTTTCAATCTCTTGTA

133681  TATGTTTGTATCATTTCTTTCATAAAACCTCAAGAGCATTTTTAAGCCCTTTTCATCAGA
133681  ATACAAACATAGTAAAGAAAGTATTTTGGAGTTCTCGTAAAAATTCGGGAAAAGTAGTCT

133741  CCCAGTGAAAACTAAGGATAGATGTTTAAAAACTGGAGGTCTCCTGATAAGGAGAACACA
133741  GGGTCACTTTTGATTCCTATCTACAAATTTTTGACCTCCAGAGGACTATTCCTCTTGTGT

133801  ATCCACCATTGTCATTTAAGTAATAAGACAGGAAATTGACCTTGACGCTTTCTTGTTAAA
133801  TAGGTGGTAACAGTAAATTCATTATTCTGTCCTTTAACTGGAACTGCGAAAGAACAATTT
```

FIGURE 4 (cont'd)

```
133861  TAGATTTAACAGGAACATCTGCACATCTTTTTTCCTTGTGCACTATTTGTTTAATTGCAG
133861  ATCTAAATTGTCCTTGTAGACGTGTAGAAAAAAGGAACACGTGATAAACAAATTAACGTC

133921  TGGATTAATACAGCAAGAGTGCCACATTATAACTAGGCAATTATCCATTCTTCAAGACTT
133921  ACCTAATTATGTCGTTCTCACGGTGTAATATTGATCCGTTAATAGGTAAGAAGTTCTGAA

133981  AGTTATTGTCACACTAATTGATCGTTTAAGGCATAAGATGGTCTAGCATTAGGAACATGT
133981  TCAATAACAGTGTGATTAACTAGCAAATTCCGTATTCTACCAGATCGTAATCCTTGTACA

134041  GAAGCTAATCTGCTCAAAAAGATCAACAAATTAATATTGTTGCTGATATTTGCATAATTG
134041  CTTCGATTAGACGAGTTTTTCTAGTTGTTTAATTATAACAACGACTATAAACGTATTAAC

134101  GCTGCAATTATTTAATGTTTAATTGGGTTGATCAAATGAGATTCAGCAATTCACAAGTGC
134101  CGACGTTAATAAATTACAAATTAACCCAACTAGTTTACTCTAAGTCGTTAAGTGTTCACG

134161  ATTAATATAAACAGAACTGGTGGCACTTAAAATGATAATGATTAACTTATATTGCATGTT
134161  TAATTATATTTGTCTTGACCACCGTGAATTTTACTATTACTAATTGAATATAACGTACAA

134221  CTCTTCCTTTCACTTTTTTCAGTTTCTACATTTCAGACCGAGCTTGTCAGCTTTTTTGAA
134221  GAGAAGGAAAGTGAAAAAAGTCAAAGATGTAAAGTCTGGCTCGAACAGTCGAAAAAACTT

134281  AACACATCAGTAGAAACCAAGATTTTAAAATGAAGTGTCAAGACAAAGGCAAAACCTGAG
134281  TTGTGTAGTCATCTTTGGTTCTAAAATTTTACTTCACAGTTCTGTTTCCGTTTTGGACTC

134341  CAGTTCCTAAAAAGATTTGCTGTTAGAAATTTTCTTTGTGGCAGTCATTTATTAAGGATT
134341  GTCAAGGATTTTTCTAAACGACAATCTTTAAAAGAAACACCGTCAGTAAATAATTCCTAA

134401  CAACTCGTGATACACCAAAAGAAGAGTTGACTTCAGAGATGTGTTCCATGCTCTCTAGCA
134401  GTTGAGCACTATGTGGTTTTCTTCTCAACTGAAGTCTCTACACAAGGTACGAGAGATCGT

134461  CAGGAATGAATAAATTTATAACACCTGCTTTAGCCTTTGTTTTCAAAAGCACAAAGGAAA
134461  GTCCTTACTTATTTAAATATTGTGGACGAAATCGGAAACAAAAGTTTTCGTGTTTCCTTT

134521  AGTGAAAGGGAAAGAGAAACAAGTGACTGAGAAGTCTTGTTAAGGAATCAGGTTTTTTCT
134521  TCACTTTCCCTTTCTCTTTGTTCACTGACTCTTCAGAACAATTCCTTAGTCCAAAAAAGA

134581  ACCTGGTAAACATTCTCTATTCTTTTCTCAAAAGATTGCTGTAAGAAAAAATGTAAGACA
134581  TGGACCATTTGTAAGAGATAAGAAAAGAGTTTTCTAACGACATTCTTTTTTACATTCTGT

134641  AAAAAAAAAAAAAAAAAACAAACAGAGGCAGAGGCAGGCAGTAGCAAGAAAGCAGAGCGTA
134641  TTTTTTTTTTTTTTTTTGTTTGTCTCCGTCTCCGTCCGTCATCGTTCTTTCGTCTCGCAT

134701  ACATCAGCTAGATGGTAACATGCAATGTCAGCTCTCTTGAAGACATGGGAAACCTAAGTT
134701  TGTAGTCGATCTACCATTGTACGTTACAGTCGAGAGAACTTCTGTACCCTTTGGATTCAA

134761  ACACCTTGGGTTAAAATTCTTCACCATATTAGTTTTGTTGCTTCATAAAATTTACCTAAG
134761  TGTGGAACCCAATTTTAAGAAGTGGTATAATCAAAACAACGAAGTATTTTAAATGGATTC

134821  CAAGTGGTCTTGCTTGCCTCAAATCCAAGCAGTCTTGAACACTTGGAGGCAATTAATGAG
134821  GTTCACCAGAACGAACGGAGTTTAGGTTCGTCAGAACTTGTGAACCTCCGTTAATTACTC

134881  TATATCTTAGTCAAAAGAATTGTTGGAGCTTTTTATTAAAGCTACAGTTTCAGTTCTGCT
134881  ATATAGAATCAGTTTTCTTAACAACCTCGAAAAATAATTTCGATGTCAAAGTCAAGACGA

134941  TTTGGGGAATTGTGCTATGAAAGCAGCTGCCAAAATAAGCTCATTTATTTTCTTCAATCC
134941  AAACCCCTTAACACGATACTTTCGTCGACGGTTTTATTCGAGTAAATAAAAGAAGTTAGG

135001  CACTCAGTGCTCAGTCACTATATTCTGTTTCCTTTTTTTTTTTCAAGTTGCATATTTGGT
135001  GTGAGTCACGAGTCAGTGATATAAGACAAAGGAAAAAAAAAAAAGTTCAACGTATAAACCA

135061  TTCCCCTTATGATTGGGAAAGATGAATTTTCAGCAGAAAACATTGTTTGTTCACTTTCAA
135061  AAGGGGAATACTAACCCTTTCTACTTAAAAGTCGTCTTTTGTAACAAACAAGTGAAAGTT

135121  AGAGTGATAGTTTCTAAAACATTTAGAGCAATAAATATTCATCAGAGGTACCAAGTAAGC
135121  TCTCACTATCAAAGATTTTGTAAATCTCGTTATTTATAAGTAGTCTCCATGGTTCATTCG
```

FIGURE 4 (cont'd)

```
135181  CGGCAGAAGAGTTAAGGGTTAGAGAAATCCCTTATTTCATGTCTTGACTCTAAAATTATC
135181  GCCGTCTTCTCAATTCCCAATCTCTTTAGGGAATAAAGTACAGAACTGAGATTTTAATAG

135241  AAAGTACTTTTCCTTGTAATGTGGATTTCTTCTTATGCGGATATGCAAAAACTTCAGTTA
135241  TTTCATGAAAAGGAACATTACACCTAAAGAAGAATACGCCTATACGTTTTTGAAGTCAAT

135301  TACGTAGTAATGCTAGCAGGTAATTTTAGTAGACATTTTATAACAACTGTCACTTTGTTT
135301  ATGCATCATTACGATCGTCCATTAAAATCATCTGTAAAATATTGTTGACAGTGAAACAAA

135361  CGCCACATGTAGAGTTTGTTCAGCTATTTTCCAGATATCTCCCCACAAAAGGAGGCAAAG
135361  GCGGTGTACATCTCAAACAAGTCGATAAAAGGTCTATAGAGGGGTGTTTTCCTCCGTTTC

135421  GGTACCAGCTTTTCAATGAGCATTACCTATTACTTGGCAAAGATGATGAAGACTCTATTA
135421  CCATGGTCGAAAAGTTACTCGTAATGGATAATGAACCGTTTCTACTACTTCTGAGATAAT

135481  ATAGTTCATTTGATAAATGTTGACATAACCAACAATAGAGATTAGGAAGTTAGTTTTAAG
135481  TATCAAGTAAACTATTTACAACTGTATTGGTTGTTATCTCTAATCCTTCAATCAAAATTC

135541  AAATCAATGGCATATAGACATTACCCTCATGGAGTTTGTATTCTACTACTTGAACTGATT
135541  TTTAGTTACCGTATATCTGTAATGGGAGTACCTCAAACATAAGATGATGAACTTGACTAA

135601  GTAGCTATAAAAGCATAGTTAGATAGCTGAATAGTTAGATCATAAGCAAAGAAGGCCAGA
135601  CATCGATATTTTCGTATCAATCTATCGACTTATCAATCTAGTATTCGTTTCTTCCGGTCT

135661  ACACATCTCTTATCAAGAAATCAATGAATAGTTTATCTCATTTTTAAAGCAACTTTATCC
135661  TGTGTAGAGAATAGTTCTTTAGTTACTTATCAAATAGAGTAAAAATTTCGTTGAAATAGG

135721  TTCTTTAATTCCTTCCTTTCTTCTAGTGCAAAACTACTTAATAAGGTTGGTGTTTAGGTT
135721  AAGAAATTAAGGAAGGAAAGAAGATCACGTTTTGATGAATTATTCCAACCACAAATCCAA

135781  AGTGTTCACACCATTCCTCATCTGGTGTGAATTACCTTCTCTTTCTTTACTATTTACTAC
135781  TCACAAGTGTGGTAAGGAGTAGACCACACTTAATGGAAGAGAAAGAAATGATAAATGATG

135841  CAACCTAGTACATGTGTTGACTGAATTCTTTTCAAACAATGTTGAGTTATCATGGTGCAC
135841  GTTGGATCATGTACACAACTGACTTAAGAAAAGTTTGTTACAACTCAATAGTACCACGTG

135901  CTAATAAATTAACACCACAGATTACAGCATCCTTGCTGATTTTCTCAGCAAAGCCAGATT
135901  GATTATTTAATTGTGGTGTCTAATGTCGTAGGAACGACTAAAAGAGTCGTTTCGGTCTAA

135961  AGATGGAAATAAACAAAGAAAATGATCCTAGAGTGAATTTTTCTAGAAAATATCTATTAT
135961  TCTACCTTTATTTGTTTCTTTTACTAGGATCTCACTTAAAAAGATCTTTTATAGATAATA

136021  GAACCATGCTGTTTAAAGTATTAGCTTGAAGGTGATGGATCCAGCTATTCAGAAAATAAC
136021  CTTGGTACGACAAATTTCATAATCGAACTTCCACTACCTAGGTCGATAAGTCTTTTATTG

136081  TTTCATATAACCATGATTTTGCACAGTATGAGGTCTTAAATGTGTGGAAAGAGATAAATT
136081  AAAGTATATTGGTACTAAAACGTGTCATACTCCAGAATTTACACACCTTTCTCTATTTAA

136141  TTTTATCATTACCACAAACCCCTTTTAAAGATTCAAAGGTGGAAGAAAGTGATTTATTTT
136141  AAAATAGTAATGGTGTTTGGGGAAAATTTCTAAGTTTCCACCTTCTTTCACTAAATAAAA

136201  TTCTCTTCAGCATACATATATAAAAGACTTGTCAGATGTTTAATTTGGGGAGGTTGATAA
136201  AAGAGAAGTCGTATGTATATATTTTCTGAACAGTCTACAAATTAAACCCCTCCAACTATT

136261  TGAAACATATCAACAGAGTATAGTAGTTATAGTAGTGTTTGTGGGTAAATAATTTCCTGG
136261  ACTTTGTATAGTTGTCTCATATCATCAATATCATCACAAACACCCATTTATTAAAGGACC

136321  GGTCAGACATATATAAACATATTTGCTTCAAAATGATAAAGGCATGAAATCAGTCTTAAA
136321  CCAGTCTGTATATATTTGTATAAACGAAGTTTTACTATTTCCGTACTTTAGTCAGAATTT

136381  AATTGAAATGGGGGTGATGGGGGAGAAAAAGAAGAACAAATTTGAAGTGCCCTTTCAAAT
136381  TTAACTTTACCCCCACTACCCCCTCTTTTTCTTCTTGTTTAAACTTCACGGGAAAGTTTA

136441  CTGCTGGATACAAGTATTGAAGTTTTAAGTCATCTTATTCTGTCTGAAAGTGTATTTTTC
136441  GACGACCTATGTTCATAACTTCAAAATTCAGTAGAATAAGACAGACTTTCACATAAAAAG

136501  ATTCTACAATAGACCCAATCAACAAGACGTATAACTTGAGTTGCATGATGTTCAGTTTAT
136501  TAAGATGTTATCTGGGTTAGTTGTTCTGCATATTGAACTCAACGTACTACAAGTCAAATA
```

FIGURE 4 (cont'd)

```
136561  GTAATCTACTGTTGGGATGGTAAGAATTGATGTAGGCTGTGGTGTAAGAATGAATTAAAA
136561  CATTAGATGACAACCCTACCATTCTTAACTACATCCGACACCACATTCTTACTTAATTTT

136621  TATAGTTTCACTGGCTTTTCTCTACATATCCACTATCACAATGGCTAGGTTTCCTGTTGC
136621  ATATCAAAGTGACCGAAAAGAGATGTATAGGTGATAGTGTTACCGATCCAAAGGACAACG

136681  TCACTATTGGATTCTGGAGAAAAATTTAATGAAAGATGATATCAGAGGAAGAATAAGTGG
136681  AGTGATAACCTAAGACCTCTTTTTAAATTACTTTCTACTATAGTCTCCTTCTTATTCACC

136741  AGGTAGAGAAGAAAGGAATGATAGAGGAGGGGAAAAAAACAAAACATATTTTTGTGTTAT
136741  TCCATCTCTTCTTTCCTTACTATCTCCTCCCCTTTTTTTGTTTTGTATAAAAACACAATA

136801  CCAAAGGAGCTTTTTCCTTATTCTGTCAAGCATTGAGATCTTCTTCAGCTTTCAATGTAG
136801  GGTTTCCTCGAAAAAGGAATAAGACAGTTCGTAACTCTAGAAGAAGTCGAAAGTTACATC

136861  TTGCTAAATACAAATAATGCTACTAGGTAGTGACTAAATATAGCAAACACTTCATCAGAT
136861  AACGATTTATGTTTATTACGATGATCCATCACTGATTTATATCGTTTGTGAAGTAGTCTA

136921  ATTAGAATTAGGTCACACTATTGAGGTTATAATCTGAAGGTTGTGTTACATAGAAACCAC
136921  TAATCTTAATCCAGTGTGATAACTCCAATATTAGACTTCCAACACAATGTATCTTTGGTG

136981  TTTAGATTATTATCAACTTGGACTAGGCTTTATTTTATAATAGCATAGTAAGTAATATCT
136981  AAATCTAATAATAGTTGAACCTGATCCGAAATAAAATATTATCGTATCATTCATTATAGA

137041  ATTGTGTCATTTCTTCAACCATTTTATTCTAAGATCCATGAAGCTTCTTGAGGCCAAATA
137041  TAACACAGTAAAGAAGTTGGTAAAATAAGATTCTAGGTACTTCGAAGAACTCCGGTTTAT

137101  AAATAATAAGTTTAGACAAGAAGTAGATTGTGACTTTTTTCCCTTAGAGATACTATTTAC
137101  TTTATTATTCAAATCTGTTCTTCATCTAACACTGAAAAAAGGGAATCTCTATGATAAATG

137161  TATCTCCTATCCTGATAGGTGGAAGGTTTACTGAATTGGAAATTGGTTGACTATTAGTTT
137161  ATAGAGGATAGGACTATCCACCTTCCAAATGACTTAACCTTTAACCAACTGATAATCAAA

137221  TTAACTAAAATGTGCAATAACACATTGCAGTTTCCTCAAACTAGTTTCCTATGATCATTA
137221  AATTGATTTTACACGTTATTGTGTAACGTCAAAGGAGTTTGATCAAAGGATACTAGTAAT

137281  AACTCATTCTCAGGGTTAAGAAAGGAATGTAAATTTCTGCCTCAATTTGTACTTCATCAA
137281  TTGAGTAAGAGTCCCAATTCTTTCCTTACATTTAAAGACGGAGTTAAACATGAAGTAGTT

137341  TAAGTTTTTGAAGAGTGCAGATTTTTAGTCAGGTCTTAAAAATAAACTCACAAATCTGGA
137341  ATTCAAAAACTTCTCACGTCTAAAAATCAGTCCAGAATTTTTATTTGAGTGTTTAGACCT

137401  TGCATTTCTAAATTCTGCAAATGTTTCCTGGGGTGACTTAACAAGGAATAATCCCACAAT
137401  ACGTAAAGATTTAAGACGTTTACAAAGGACCCCACTGAATTGTTCCTTATTAGGGTGTTA

137461  ATACCTAGCTACCTAATACATGGAGCTGGGGCTCAACCCACTGTTTTTAAGGATTTGCGC
137461  TATGGATCGATGGATTATGTACCTCGACCCCGAGTTGGGTGACAAAAATTCCTAAACGCG

137521  TTACTTGTGGCTGAGGAAAAATAAGTAGTTCGAGGAAGTAGTTTTTAAATGTGAGCTTAT
137521  AATGAACACCGACTCCTTTTTATTCATCAAGCTCCTTCATCAAAAATTTACACTCGAATA

137581  AGATAGAAACAGAATATCAACTTAATTATGAAATTGTTAGAACCTGTTCTCTTGTATCTG
137581  TCTATCTTTGTCTTATAGTTGAATTAATACTTTAACAATCTTGGACAAGAGAACATAGAC

137641  AATCTGATTGCAATTACTATTGTACTGATAGACTCCAGCCATTGCAAGTCTCAGATATCT
137641  TTAGACTAACGTTAATGATAACATGACTATCTGAGGTCGGTAACGTTCAGAGTCTATAGA

137701  TAGCTGTGTAGTGATTCTTGAAATTCTTTTTAAGAAAAATTGAGTAGAAAGAAATAAACC
137701  ATCGACACATCACTAAGAACTTTAAGAAAAATTCTTTTTAACTCATCTTTCTTTATTTGG

137761  CTTTGTAAATGAGGCTTGGCTTTTGTGAAAGATCATCCGCAGGCTATGTTAAAAGGATTT
137761  GAAACATTTACTCCGAACCGAAAACACTTTCTAGTAGGCGTCCGATACAATTTTCCTAAA

137821  TAGCTCACTAAAAGTGTAATAATGGAAATGTGGAAAATATCGTAGGTAAAGGAAACTACC
137821  ATCGAGTGATTTTCACATTATTACCTTTACACCTTTTATAGCATCCATTTCCTTTGATGG
```

FIGURE 4 (cont'd)

```
137881  TCATGCTCTGAAGGTTTTGTAGAAGCACAATTAAACATCTAAAATGGCTTTGTTACACCA
137881  AGTACGAGACTTCCAAAACATCTTCGTGTTAATTTGTAGATTTTACCGAAACAATGTGGT

137941  GAGCCATCTGGTGTGAAGAACTCTATATTTGTATGTTGAGAGGGCATGGAATAATTGTAT
137941  CTCGGTAGACCACACTTCTTGAGATATAAACATACAACTCTCCCGTACCTTATTAACATA

138001  TTTGCTGGCAATAGACACATTCTTTATTATTTGCAGATTCCTCATCAAATCTGTAATTAT
138001  AAACGACCGTTATCTGTGTAAGAAATAATAAACGTCTAAGGAGTAGTTTAGACATTAATA

138061  GCACAGTTTCTGTTATCAATAAAACAAAAGAATCCTGTTTGTGTGGTTTCATGAAATCAG
138061  CGTGTCAAAGACAATAGTTATTTTGTTTTCTTAGGACAAACACACCAAAGTACTTTAGTC

138121  CATTGTTGAATGCATGAAGTAATAATGCTAAATTAACATTTTTATGATGTCTCAAGGTTT
138121  GTAACAACTTACGTACTTCATTATTACGATTTAATTGTAAAAATACTACAGAGTTCCAAA

138181  CTGGTCAAGGGAAGTAAATGTAGGATAGTATTTTTACACCAAAATGACACAGAGAGAATT
138181  GACCAGTTCCCTTCATTTACATCCTATCATAAAAATGTGGTTTTACTGTGTCTCTCTTAA

138241  GAGCACACCAGAAAGACCAGAAACCACACCACTGGATAGAGATTCAATATGTTTCTTTTT
138241  CTCGTGTGGTCTTTCTGGTCTTTGGTGTGGTGACCTATCTCTAAGTTATACAAAGAAAAA

138301  CAAACATTTGGACAAGAAAAAAATGGGCATTTAAAAATTCTTCCTTCCCCTGGTTATGGA
138301  GTTTGTAAACCTGTTCTTTTTTTACCCGTAAATTTTTAAGAAGGAAGGGGACCAATACCT

138361  TTTATCTGTAGTAAAACTTAGCTTTGTCGTTTGAGATTTGCACAGAATGGGGGGAGTAGA
138361  AAATAGACATCATTTTGAATCGAAACAGCAAACTCTAAACGTGTCTTACCCCCCTCATCT

138421  TTACCTCTTCCCATTCATTGTCATAATGGATCTACATCACTGATAAACACCATACTTCTA
138421  AATGGAGAAGGGTAAGTAACAGTATTACCTAGATGTAGTGACTATTTGTGGTATGAAGAT

138481  TATGTGGTTAACTAGCTTTAGAATAAAAGACACTTTAAAAAGTAAAAGGCCTAGAGATCT
138481  ATACACCAATTGATCGAAATCTTATTTTCTGTGAAATTTTTCATTTTCCGGATCTCTAGA

138541  TCAATGAAGTGTCCCTTTTAGCCAAACCAGGCCTTGCAGAAATTGTCCTCAAAAGCACCA
138541  AGTTACTTCACAGGGAAAATCGGTTTGGTCCGGAACGTCTTTAACAGGAGTTTTCGTGGT

138601  AGGGAGAACAAAGCCAAGTGCAGAATTACCCAAGGGTCACACATTTTGTATTCATTTTCT
138601  TCCCTCTTGTTTCGGTTCACGTCTTAATGGGTTCCCAGTGTGTAAAACATAAGTAAAAGA

138661  TATAATTTGCCCAAATGATTTGAAGTACAGCAAAACCTGAAGTTTTGCAAAGAGTTACTG
138661  ATATTAAACGGGTTTACTAAACTTCATGTCGTTTTGGACTTCAAAACGTTTCTCAATGAC

138721  TAAACTGAACTTAAAAATGTCAGAGTGCTCGGTGACCCACTTCTCCAGACCCTGTCAACC
138721  ATTTGACTTGAATTTTTACAGTCTCACGAGCCACTGGGTGAAGAGGTCTGGGACAGTTGG

138781  TGTGAAAATATTGGCCTTTATTCAGATCTCTCAAGAAGTTACGCGCAGAGTTTGGAAGGT
138781  ACACTTTTATAACCGGAAATAAGTCTAGAGAGTTCTTCAATGCGCGTCTCAAACCTTCCA

138841  CTAGGCAAAGGTTATTAGTCAAGTGTTCTTACAGTGTCAACGCTCAATTCCCACAAGCGT
138841  GATCCGTTTCCAATAATCAGTTCACAAGAATGTCACAGTTGCGAGTTAAGGGTGTTCGCA

138901  GAGAAAGAGAGACCTGTCATTCCTGAGGGTGATGACATACATTTACTGGAGCTTATATAA
138901  CTCTTTCTCTCTGGACAGTAAGGACTCCCACTACTGTATGTAAATGACCTCGAATATATT

138961  TTTATCAGATAAGACAGCAGTTTCCTTCAGGGTAGAAAGTGTGTTTTCTACATTGATTTA
138961  AAATAGTCTATTCTGTCGTCAAAGGAAGTCCCATCTTTCACACAAAAGATGTAACTAAAT

139021  GTACAAAACAAAAAGAAAAGGGGATATTTCAAATTTTATAATTATTTTTCTGCTAAGCTG
139021  CATGTTTTGTTTTTCTTTTCCCCTATAAAGTTTAAAATATTAATAAAAAGACGATTCGAC

139081  ATTCAGTGTGATTTAAGCATATTTTTCAAATCATGAATCTGATTCCATATACATATGTGC
139081  TAAGTCACACTAAATTCGTATAAAAAGTTTAGTACTTAGACTAAGGTATATGTATACACG

139141  CTTATATTGTGATAATTTATTTTTAAGTGAAATATGCTATCATAGCCTGACTTTATGTAT
139141  GAATATAACACTATTAAATAAAAATTCACTTTATACGATAGTATCGGACTGAAATACATA

139201  AGTGGTGACAACTTGCAGGATCGCATTTCTGTAACCAAACGGCCGACAGCTGAGGTGTAG
139201  TCACCACTGTTGAACGTCCTAGCGTAAAGACATTGGTTTGCCGGCTGTCGACTCCACATC
```

FIGURE 4 (cont'd)

```
139261  ATGCTTCCTATGGTCTGTAGAATAATCACTGGGCTTGTTCTCTAGCTATGTCTGTATGCA
139261  TACGAAGGATACCAGACATCTTATTAGTGACCCGAACAAGAGATCGATACAGACATACGT

139321  ATCGCGACAGTGTTGATCAAAACCCATGATCATTCTCTCCAAAGGTCTTTGTCACTAAGC
139321  TAGCGCTGTCACAACTAGTTTTGGGTACTAGTAAGAGAGGTTTCCAGAAACAGTGATTCG

139381  CACTAGGGATTCTGAAAAGTTCGTGAGCTGAAACAAATAAATTGAGTTGGAAGATT
139381  GTGATCCCTAAGACTTTTCAAGCACTCGACTTTGTTTATTTAACTCAACCTTCTAA
```

FIGURE 5

>hg19_refGene_NM_001220776 range=chr7:50342378-50472798 5'pad=0 3'pad=0 strand=+ repeatMasking=none

IKZF1 runs on plus strand.

YELLOW: SUB-REGION OF INTEREST TOWARDS WHICH AN EXAMPLE OF A BISULPHITE CONVERSION AND METHYLATION SPECIFIC ASSAY IS GIVEN

```
   1  CCTGTCTCTGGTGTGCCCTGGCCCCCGACTTGGAGGCCTCCTGGGCCAGGCCAAGACCTT
   1  GGACAGAGACCACACGGGACCGGGGGCTGAACCTCCGGAGGACCCGGTCCGGTTCTGGAA

  61  CCCCGGCAGCGATGGTCTCCAGCCACACTCAACTGCCCTGAAGGGACATTTCCTGCTTAT
  61  GGGGCCGTCGCTACCAGAGGTCGGTGTGAGTTGACGGGACTTCCCTGTAAAGGACGAATA

 121  TCCCTTGCCCGGCTGTGTCCTCCACCCGGAAGGCCTGTGCCTTCTTCGCCTGCATGTCCT
 121  AGGGAACGGGCCGACACAGGAGGTGGGCCTTCCGGACACGGAAGAAGCGGACGTACAGGA

 181  ACCCTGAGGAGGCTCCCTTGGTCTTTCATCGCTCTCCCTATGGGTCTTCACGCCTTCCCG
 181  TGGGACTCCTCCGAGGGAACCAGAAAGTAGCGAGAGGGATACCCAGAAGTGCGGAAGGGC

 241  AACCACCGCGCCCAAGCAGGAGCACGTTCTCGGCCCTCTTCACAGGGCGCTCCTCCTCAC
 241  TTGGTGGCGCGGGTTCGTCCTCGTGCAAGAGCCGGGAGAAGTGTCCCGCGAGGAGGAGTG

 301  AGGGGTGCCCGGGATTTTTATTCTGTGCCTTCCTGGTGGCTCCTACAAGTCTGGAAGGGC
 301  TCCCCACGGGCCCTAAAAATAAGACACGGAAGGACCACCGAGGATGTTCAGACCTTCCCG

 361  AGGAGGCGCATCTCACTCCTCTGGGTCCCCTCCCCTAGCGCCTGGCGGGAGCCCAGGCTG
 361  TCCTCCGCGTAGAGTGAGGAGACCCAGGGGAGGGGATCGCGGACCGCCCTCGGGTCCGAC

 421  CATTTGTGGAATTCATGACTTTTTCTCTCCTGCTCAAGCTGAACACATTGCTGGCTCCTG
 421  GTAAACACCTTAAGTACTGAAAAAGAGAGGACGAGTTCGACTTGTGTAACGACCGAGGAC

 481  CTCGGGTGGAGCCCGGCTAATTAGAGTGAGGGGCTCCCCGTAGGGCGAAGGGGTGCGCTG
 481  GAGCCCACCTCGGGCCGATTAATCTCACTCCCCGAGGGGCATCCCGCTTCCCCACGCGAC

 541  TCAGATGTGGCATTCCCGTTTTACGGAGACACACGGTGTCTTACACGCCAGGGAGAGGTC
 541  AGTCTACACCGTAAGGGCAAAATGCCTCTGTGTGCCACAGAATGTGCGGTCCCTCTCCAG

 601  TGAGACGCAAAGAGCCGTCGAGCGGGCTGCGGGATTGCTTCGCTGTCACCTCCGCCTGCA
 601  ACTCTGCGTTTCTCGGCAGCTCGCCCGACGCCCTAACGAAGCGACAGTGGAGGCGGACGT

 661  GCCACCCTTCCGCACGCACTTGTGTGTGCACCCAGGCCAACATGGAAGGCGCCATCCTAA
 661  CGGTGGGAAGGCGTGCGTGAACACACACGTGGGTCCGGTTGTACCTTCCGCGGTAGGATT

 721  CTTCTGCCGTGAGCAGGTGGGAGGGAAGAGAGACGAGAGGTATTCCATTGGTTGTCTGGG
 721  GAAGACGGCACTCGTCCACCCTCCCTTCTCTCTGCTCTCCATAAGGTAACCAACAGACCC

 781  AAAATGAATTGCACCTTCCCCTCCCTTGCGGAGGATCAACTTTTCCCACCCCCTCGGGTG
 781  TTTTACTTAACGTGGAAGGGGAGGGAACGCCTCCTAGTTGAAAAGGGTGGGGGAGCCCAC

 841  GGCACTCGCATCCTGGGGCCGGAGCCTGAACCCGGGAGCCAAGGGGCCCCAGTTCCAGGG
 841  CCGTGAGCGTAGGACCCCGGCCTCGGACTTGGGCCCTCGGTTCCCCGGGGTCAAGGTCCC

 901  ACGTGAAGCTGAGCGTACAGCGGGCGCTCCCAGACACTGGGGAAAGTGCTTTACGATGTC
 901  TGCACTTCGACTCGCATGTCGCCCGCGAGGGTCTGTGACCCCTTTCACGAAATGCTACAG

 961  CCGAGTCCCTCCAGTCTCGCCAGCGGGGCGAGCGTGAGGGTGCCCCGACCGACCAGCGGC
 961  GGCTCAGGGAGGTCAGAGCGGTCGCCCCGCTCGCACTCCCACGGGGCTGGCTGGTCGCCG

1021  CCCGGGTGCAGGGTGGCGGGCCCGGCGGCGCGCGTCCCCCTCCCCCTCCTGGCGGCCCGC
1021  GGGCCCACGTCCCACCGCCCGGGCCGCCGCGCGCAGGGGGAGGGGGAGGACCGCCGGGCG

1081  ACGTGTCGCCCGCGCCGCGCCCCCACGGGTTACGCGCGGGTCCCGCAGCGCCGCGGCCGA
1081  TGCACAGCGGGCGCGGCGCGGGGGTGCCCAATGCGCGCCCAGGGCGTCGCGGCGCCGGCT

1141  GCCGGGCTGCCCGGCCCGCGGACACAGCGCCGGCCGCCGCATCCCGTGCGGGGCCGCGGC
1141  CGGCCCGACGGGCCGGGCGCCTGTGTCGCGGCCGGCGGCGTAGGGCACGCCCCGGCGCCG
```

```
1201  GCGATGCTGCGCTGGAATGAGGAAGCGCGGCGGCGAGGGGAGGGCCCGGGCGCGGTGCGC
1201  CGCTACGACGCGACCTTACTCCTTCGCGCCGCCGCTCCCCTCCCGGGCCCGCGCCACGCG

1261  GCGGGGGTGGCGGCGGCGCGCCGAGCGGGCCCGGCGCGGGCGAGCGGGCTGCAGCCGGCG
1261  CGCCCCCACCGCCGCCGCGCGGCTCGCCCGGGCCGCGCCCGCTCGCCCGACGTCGGCCGC

1321  GCGGCGCCAGCAGGTACGGCCCGCACCCGCCGCCGCCCCGGCGGCCTTTGGGGGCTGAGC
1321  CGCCGCGGTCGTCCATGCCGGGCGTGGGCGGCGGCGGGGCCGCCGGAAACCCCCGACTCG

1381  CGGAGCCCGGCGCGATTGCAAAGTTTTCGTGCGCGGCCCCTCTGGCCCGGAGTTGCGGCT
1381  GCCTCGGGCCGCGCTAACGTTTCAAAAGCACGCGCCGGGGAGACCGGGCCTCAACGCCGA

1441  GAGACGCGCGCCGCGCGAGCCGGGGGACTCGGCGACGGGGCGGGGACGGGACGACGCACC
1441  CTCTGCGCGCGGCGCGCTCGGCCCCCTGAGCCGCTGCCCCGCCCCTGCCCTGCTGCGTGG

1501  CTCTCCGTGTCCCGCTCTGCGCCCTTCTGCGCGCCCCGCTCCCTGTACCGGAGCAGCGAT
1501  GAGAGGCACAGGGCGAGACGCGGGAAGACGCGCGGGGCGAGGGACATGGCCTCGTCGCTA

1561  CCGGGAGGCGGCCGAGAGGTGCGCGCGGGGCCGAGCCGGCTGCGGGGCAGGTCGAGCAGG
1561  GGCCCTCCGCCGGCTCTCCACGCGCGCCCCGGCTCGGCCGACGCCCCGTCCAGCTCGTCC

1621  GACCGCCAGCGTGCGTCACCCCAAAGTTTGCGGGGTGGCAGGGCGCGCGCTCTGGCCACC
1621  CTGGCGGTCGCACGCAGTGGGGTTTCAAACGCCCCACCGTCCCGCGCGCGAGACCGGTGG

1681  CGCCGCTCTGGGCGGCAGCTGGTGGCAACGCAAGGGCGCGGCGGGGGCGGCCGGCGCGGA
1681  GCGGCGAGACCCGCCGTCGACCACCGTTGCGTTCCCGCGCCGCCCCCGCCGGCCGCGCCT

1741  GGGGGCCAGGTACGGGGCCCGCGGGCGGCGCTGTGCGCGCGGGGCAGCCGGTCGGCCGGG
1741  CCCCCGGTCCATGCCCCGGGCGCCCGCCGCGACACGCGCGCCCCGTCGGCCAGCCGGCCC

1801  AGCGCGAAAGCCTGGTCTGAGCCGGCTGGGGGCGGGGAGTGTGGCGGAGAAATGGGGAAC
1801  TCGCGCTTTCGGACCAGACTCGGCCGACCCCCGCCCCTCACACCGCCTCTTTACCCCTTG

1861  AATGCGAGTGAGCAACTTCAGGAAGTCATTGTGAAAGAAAGCTGGGAAGAGCTCCGCGGC
1861  TTACGCTCACTCGTTGAAGTCCTTCAGTAACACTTTCTTTCGACCCTTCTCGAGGCGCCG

1921  CAAGTTAGCAGGACACTCTAACAAGTGACTGCGCGGCCCGCGCCCGGGGCGGTGACTGCG
1921  GTTCAATCGTCCTGTGAGATTGTTCACTGACGCGCCGGGCGCGGGCCCCGCCACTGACGC

1981  GCAAGCCCCCTGGGTCCCCGCGCGGCGCATCCCAGCCTGGGCGGGACGCTCGGCCGCGGC
1981  CGTTCGGGGGACCCAGGGGCGCGCCGCGTAGGGTCGGACCCGCCCTGCGAGCCGGCGCCG

2041  GAGGCGGGCAAGCCTGGCAGGGCAGAGGGAGCCCCGGCTCCGAGGTTGCTCTTCGCACCC
2041  CTCCGCCCGTTCGGACCGTCCCGTCTCCCTCGGGGCCGAGGCTCCAACGAGAAGCGTGGG

2101  GAGGATCAGTCTTGGCCCCAAAGCGCGACGCACAAATCCACGTGAGTGTTTTCAAATTGA
2101  CTCCTAGTCAGAACCGGGGTTTCGCGCTGCGTGTTTAGGTGCACTCACAAAAGTTTAACT

2161  ATTTCAATAGGAAAACTTGGGGTAACTGGTGAATTTAAAAAAAAAAAAAACACAGTAAAGA
2161  TAAAGTTATCCTTTTGAACCCCATTGACCACTTAAATTTTTTTTTTTTGTGTCATTTCT

2221  AAAGCGGTAAGGTTGGTAGACCCTGGTGTCGCTCAGGTCCGCCTCTCTTTTCTGAGGACA
2221  TTTCGCCATTCCAACCATCTGGGACCACAGCGAGTCCAGGCGGAGAGAAAAGACTCCTGT

2281  GTGAGAGAGTTCACTTCTGTCAAGCGTCTGTTGCTCTGCACTGTGCCAGCAGGTGCAGGA
2281  CACTCTCTCAAGTGAAGACAGTTCGCAGACAACGAGACGTGACACGGTCGTCCACGTCCT

2341  CCAGGCCGACATGGGACACTTCTGAGCAGCCCCGCTGTCACCAGGAGAGGAGTTCTAGCT
2341  GGTCCGGCTGTACCCTGTGAAGACTCGTCGGGGCGACAGTGGTCCTCTCCTCAAGATCGA

2401  CCCAACCATATTTAAATTTATGTAGACCTACATATACCCACGGAAGTCAGCCTTTATAAA
2401  GGGTTGGTATAAATTTAAATACATCTGGATGTATATGGGTGCCTTCAGTCGGAAATATTT

2461  GTCGTGTGTAAAGAGTTTTCCTTATATTTGAGCCGGGAGCTTTCTTTTTATACTATAAAT
2461  CAGCACACATTTCTCAAAAGGAATATAAACTCGGCCCTCGAAAGAAAAATATGATATTTA

2521  ATGATGAGATCGAGTCTGAACTTAATTTCTGCAAGAGAGGAATTATCCCGGCTTTGAAAA
2521  TACTACTCTAGCTCAGACTTGAATTAAAGACGTTCTCTCCTTAATAGGGCCGAAACTTTT

2581  GTTAGTCCTTTTGCTGACCGCAGGTTTGACGCTCAAGTCACCAAACCTTCTCAGGAAAAC
2581  CAATCAGGAAAACGACTGGCGTCCAAACTGCGAGTTCAGTGGTTTGGAAGAGTCCTTTTG

2641  CCTTAGTAATATTAAGGCATCAGGTTACTTGCGGTTATATTTGAAATGTATTTTAAATAT
```

```
2641  GGAATCATTATAATTCCGTAGTCCAATGAACGCCAATATAAACTTTACATAAAATTTATA

2701  TTGTCAAGCATCGCTGCTGATGCCTAAGGAACCTCGTGAGGGCTTGTTTTTCCTTCTAAT
2701  AACAGTTCGTAGCGACGACTACGGATTCCTTGGAGCACTCCCGAACAAAAAGGAAGATTA

2761  TTGGAGGCATCTAATGACCGAAAACCGTAGCGATTCCATAGGGTCTGACCAGGCACAGCT
2761  AACCTCCGTAGATTACTGGCTTTTGGCATCGCTAAGGTATCCCAGACTGGTCCGTGTCGA

2821  TTCAAATGCAGCTTCCCTCTCTCTAGGGACTGCAGCCCACCCAGACTGAATTTCAATGCG
2821  AAGTTTACGTCGAAGGGAGAGAGATCCCTGACGTCGGGTGGGTCTGACTTAAAGTTACGC

2881  GTGCGCTTTGCTTAGGTTACCCACTCACAATTTCCCACTGCGCCGCAGGCAGTATATTTC
2881  CACGCGAAACGAATCCAATGGGTGAGTGTTAAAGGGTGACGCGGCGTCCGTCATATAAAG

2941  AGCTTTGAGATACCTTGTTTTAAAATTCCAGACAAAATGGTGTTGAGGAAATGTCTCCTT
2941  TCGAAACTCTATGGAACAAAATTTTAAGGTCTGTTTTACCACAACTCCTTTACAGAGGAA

3001  ACTAGTCCCATCAACTTCTGTTAAAAGAGGAAAATTTATGGAATTTGAAAATACTGCGTA
3001  TGATCAGGGTAGTTGAAGACAATTTTCTCCTTTTAAATACCTTAAACTTTTATGACGCAT

3061  TGATATTTAAACTTTCATAGACATTCAAATGCTTTTAAGGCCAGGTTCAATTTGGTTATG
3061  ACTATAAATTTGAAAGTATCTGTAAGTTTACGAAAATTCCGGTCCAAGTTAAACCAATAC

3121  AGTCGAGGGGTGGGGGGGACCCACATAGAAATGTCCTGGGTCCTCTTGAGTTTATTTCTT
3121  TCAGCTCCCCACCCCCCCTGGGTGTATCTTTACAGGACCCAGGAGAACTCAAATAAAGAA

3181  TGTTTGAAGATGTTTGTTCAATGAGTTTTATTGTACTCATCTTTTATATGGAATTTTAAA
3181  ACAAACTTCTACAAACAAGTTACTCAAAATAACATGAGTAGAAAATATACCTTAAAATTT

3241  AAGTAACAATTTCAGTATTATTTATATTAGAATGTGTCAGAATTATTTCCGTGACAAATC
3241  TTCATTGTTAAAGTCATAATAAATATAATCTTACACAGTCTTAATAAAGGCACTGTTTAG

3301  AGATCATTTGGGCTATGGCTTAAAATGTACACGAGGCAAATATTCATGACAAGAAGATTC
3301  TCTAGTAAACCCGATACCGAATTTTACATGTGCTCCGTTTATAAGTACTGTTCTTCTAAG

3361  ACCTTCTTACGCTGGCATCTTGTAAAATGCAGAACAAGTTAAAGAAATAATGTGTACACA
3361  TGGAAGAATGCGACCGTAGAACATTTTACGTCTTGTTCAATTTCTTTATTACACATGTGT

3421  TACAAATAATGATGTCACATTAAAAATACTACACTATTCTTGCTTGATGGAATGTATCTG
3421  ATGTTTATTACTACAGTGTAATTTTTATGATGTGATAAGAACGAACTACCTTACATAGAC

3481  ATTTCCAATTTCACCATGAACATATTTCATACATTTTTTACATGAAAAAAAACGTGACTC
3481  TAAAGGTTAAAGTGGTACTTGTATAAAGTATGTAAAAAATGTACTTTTTTTTGCACTGAG

3541  TTAAGTCTCACAGTCAATCAGAGCTGGTGACCAGAACATTTTATTGAACTAAATGGTCAT
3541  AATTCAGAGTGTCAGTTAGTCTCGACCACTGGTCTTGTAAAATAACTTGATTTACCAGTA

3601  GTTTTCTTCCCCTTTTGTTTCACGGTGAGAGTTGAAGGAAGGAGTTTAGAAACTCTCCAG
3601  CAAAAGAAGGGGAAAACAAAGTGCCACTCTCAACTTCCTTCCTCAAATCTTTGAGAGGTC

3661  TACTTGTTTAATTCATCAGTGTTCTAATTAGAGTGGTACCTCTTGGAAAACTACACACCC
3661  ATGAACAAATTAAGTAGTCACAAGATTAATCTCACCATGGAGAACCTTTTGATGTGTGGG

3721  CCCTAATGCAGAAACATCATAGCAATAATCACCCACCCTCAGGGTCTCCAGGAGACCACA
3721  GGGATTACGTCTTTGTAGTATCGTTATTAGTGGGTGGGAGTCCCAGAGGTCCTCTGGTGT

3781  AGGGCTGCAGATAAAAGTCTGGATGTGTTAGGTTTGACCCTTTCGAAGAGTTTTACACAG
3781  TCCCGACGTCTATTTTCAGACCTACACAATCCAAACTGGGAAAGCTTCTCAAAATGTGTC

3841  GCTCCTAAAGAGAAGATCAGCTGTGGCCGTTTGTAGCCATTTCCTTTGTCGAAAAACTAA
3841  CGAGGATTTCTCTTCTAGTCGACACCGGCAAACATCGGTAAAGGAAACAGCTTTTTGATT

3901  GATCGCAGTGAATGTATTAGCCAAGAGGTCTAAAGCCCTGTTGTACTGCAGGCCACTGTC
3901  CTAGCGTCACTTACATAATCGGTTCTCCAGATTTCGGGACAACATGACGTCCGGTGACAG

3961  TTCCTTGTTTGACTAGAGACTTGGAGTTTGAGAACAGTGGTTCTTTGGTTTGGATACATT
3961  AAGGAACAAACTGATCTCTGAACCTCAAACTCTTGTCACCAAGAAACCAAACCTATGTAA

4021  TTTTGTTCTTGATTTGGATGTGTGTGTTTCATGCGTGGTTAATATAGCATATTTTCAATA
4021  AAAACAAGAACTAAACCTACACACACAAAGTACGCACCAATTATATCGTATAAAAGTTAT

4081  TAAATGTCAAAAATTTTGAAATAGGAAAGAACTCTCTATATATTAATGTACTTATACACA
4081  ATTTACAGTTTTTAAAACTTTATCCTTTCTTGAGAGATATATAATTACATGAATATGTGT
```

```
4141  CACTTCAAGATTATGCATTTATTAACAGATACATGAAATAAATTCCATGTGCATATGCAC
4141  GTGAAGTTCTAATACGTAAATAATTGTCTATGTACTTTATTTAAGGTACACGTATACGTG

4201  ATATGCACACAGAGCGTGCACACACACAGCATGCACACAGCGTGGAGTGAGAGGCATGGG
4201  TATACGTGTGTCTCGCACGTGTGTGTGTCGTACGTGTGTCGCACCTCACTCTCCGTACCC

4261  GCAGTGTGGAAGAGTTTTAACATCAAACAGACCTGAAATGAGTATTAAAGGCCCCCTTTA
4261  CGTCACACCTTCTCAAAATTGTAGTTTGTCTGGACTTTACTCATAATTTCCGGGGGAAAT

4321  TTTTTAAACTTTTACTAAAACAAGATGGATTTCCCTATGTTATATAATGGTGAATTTTAG
4321  AAAAATTTGAAAATGATTTTGTTCTACCTAAAGGGATACAATATATTACCACTTAAAATC

4381  GCATAAATAACGTTTTTTGAGTGTTGCATAATTGTACGTATTAATGTAATGTAACTGTGG
4381  CGTATTTATTGCAAAAAACTCACAACGTATTAACATGCATAATTACATTACATTGACACC

4441  TTAACGAAGAATTCATCAAGGATATCACTGTTTTGTGGCATTTTTTTTTTCCTCCTCTAA
4441  AATTGCTTCTTAAGTAGTTCCTATAGTGACAAAACACCGTAAAAAAAAAAGGAGGAGATT

4501  TCTTTGGACTTGTGAAATAATTTCACTATGAAATAAATGTTGGTTCTTGTCATATTCTAA
4501  AGAAACCTGAACACTTTATTAAAGTGATACTTTATTTACAACCAAGAACAGTATAAGATT

4561  GGGAGATTGATGTAAGTGGCTCCACTCCAGCTTACAGAAGGTAAACCACGACCTTTTTGC
4561  CCCTCTAACTACATTCACCGAGGTGAGGTCGAATGTCTTCCATTTGGTGCTGGAAAAACG

4621  GTTCTCTGAAAACGCTTGTCTTCCGATGCCTCTGTTTCTAAGACTGACAAGCACTCTGGG
4621  CAAGAGACTTTTGCGAACAGAAGGCTACGGAGACAAAGATTCTGACTGTTCGTGAGACCC

4681  GGCACTGTGACGCCTGCTTCTAGCGGCAGAGTTGCTGCAGCTCCTGTCCTGGCTGTGAAC
4681  CCGTGACACTGCGGACGAAGATCGCCGTCTCAACGACGTCGAGGACAGGACCGACACTTG

4741  ATTGTTCTCTCTCTGGTGTCTCTATGTTCATAACTACAGAGACTTCAGCTCTATTCCATT
4741  TAACAAGAGAGAGACCACAGAGATACAAGTATTGATGTCTCTGAAGTCGAGATAAGGTAA

4801  TCATATTTGTGCTGAATAATCATTCCATTTTATGGGAGAAAACACAAGATGTAAAAGCAA
4801  AGTATAAACACGACTTATTAGTAAGGTAAAATACCCTCTTTTGTGTTCTACATTTTCGTT

4861  CAAGTGACCCATCCTTTGAAGCTTACAAGAAGAGAAACATTAATCTATTTCACGTCTTGA
4861  GTTCACTGGGTAGGAAACTTCGAATGTTCTTCTCTTTGTAATTAGATAAAGTGCAGAACT

4921  AAACAGATCAGTTTTATTTTGCTCAAAAAGGGCACATGTACATTTTTGATCTAGGTCTTA
4921  TTTGTCTAGTCAAAATAAAACGAGTTTTTCCCGTGTACATGTAAAAACTAGATCCAGAAT

4981  GAAACGTAGAGTTTCAGAGGATCAGCATTATACACACTGTCACACACACACACACTTAAA
4981  CTTTGCATCTCAAAGTCTCCTAGTCGTAATATGTGTGACAGTGTGTGTGTGTGTGAATTT

5041  ATTCAGATGAGGAACAAGATAGGAATGAGGTTTTGTTAGGGACGCAGAGCACCTAAAACC
5041  TAAGTCTACTCCTTGTTCTATCCTTACTCCAAAACAATCCCTGCGTCTCGTGGATTTTGG

5101  AAAGGATATCGACAGTAACAAAGCTGTTTTTACTGTAGTGCTGACTGAACACTCATGCTG
5101  TTTCCTATAGCTGTCATTGTTTCGACAAAAATGACATCACGACTGACTTGTGAGTACGAC

5161  GTGTCTTCATGTGGACCATGGCTTTCTTGTATTTCTTTGCAGTTTAATAAATGACTTCAT
5161  CACAGAAGTACACCTGGTACCGAAAGAACATAAAGAAACGTCAAATTATTTACTGAAGTA

5221  ATCTCAGGTTACCTTTCCACATCTCCTGGAATATATGTTTATGTCCTTAAAGTTTCAGTG
5221  TAGAGTCCAATGGAAAGGTGTAGAGGACCTTATATACAAATACAGGAATTTCAAAGTCAC

5281  TCGTCACTTTAGTAGCTTTAGTTTGAGTTTTTAAATGTTTGGTAATATTCCAACAAATAT
5281  AGCAGTGAAATCATCGAAATCAAACTCAAAAATTTACAAACCATTATAAGGTTGTTTATA

5341  TTTTTAAGACATTATGAAACCTTATGAAGTGCCATATATTACAAGTGAGATAAAACAGCA
5341  AAAAATTCTGTAATACTTTGGAATACTTCACGGTATATAATGTTCACTCTATTTTGTCGT

5401  AGCAAAAGAAGGTTTGCAGAAGGTTTTTAAGTGGCGAAGTGCGGGCCTGCCCATTTTGGT
5401  TCGTTTTCTTCCAAACGTCTTCCAAAAATTCACCGCTTCACGCCCGGACGGGTAAAACCA

5461  GTCTCCTTGGTGGTTACTCCTGAGAAGGGCCTGGAGGAAGAGCAACTGAGGCCTAATCTA
5461  CAGAGGAACCACCAATGAGGACTCTTCCCGGACCTCCTTCTCGTTGACTCCGGATTAGAT

5521  CAGGCAACTGCCAAATTGTTTCAGTTGACGTTTTTCCCTCTCATGTTTGACTATAATAAA
5521  GTCCGTTGACGGTTTAACAAAGTCAACTGCAAAAAGGGAGAGTACAAACTGATATTATTT
```

```
5581  TAGGTAGTTGCCAGTGGAGCCTTCAGCCAACCACCTGGTAATAAACTGTTAAAAATGGTG
5581  ATCCATCAACGGTCACCTCGGAAGTCGGTTGGTGGACCATTATTTGACAATTTTTACCAC

5641  CAAACCCTAGGTCACAGGTGTGGGGGCCATTTGTCTTGCCTGTTAACAGGCCTGGCCTTA
5641  GTTTGGGATCCAGTGTCCACACCCCCGGTAAACAGAACGGACAATTGTCCGGACCGGAAT

5701  ATTCTTTTCTCCCATGGCCATTTCTGCCTTTGGGGAACTCACAATTCCTGTTGACTAAAA
5701  TAAGAAAAGAGGGTACCGGTAAAGACGGAAACCCCTTGAGTGTTAAGGACAACTGATTTT

5761  GAGCACCCTTTTCCACCACAAGCCTGACAAATCAGACGTCCACATAATTTCTGAACTCGT
5761  CTCGTGGGAAAAGGTGGTGTTCGGACTGTTTAGTCTGCAGGTGTATTAAAGACTTGAGCA

5821  TTTGGTTAGGACAGGAAGCACAGGCTCCCTTCCTGTCTGTGTTTTCCTAAGAGAAAACGG
5821  AAACCAATCCTGTCCTTCGTGTCCGAGGGAAGGACAGACACAAAAGGATTCTCTTTTGCC

5881  TCTTCCCTCCTTTTTTGCATATTTGGCAAGTGGTTCCACCTTTCTCTGCACCCTGGTGGA
5881  AGAAGGGAGGAAAAAACGTATAAACCGTTCACCAAGGTGGAAAGAGACGTGGGACCACCT

5941  GTGTGAAGGCAGCAGAGGAACCTTTTGGAGGAGGAAGAGGACACAGAGGCCCTGTAGCCA
5941  CACACTTCCGTCGTCTCCTTGGAAAACCTCCTCCTTCTCCTGTGTCTCCGGGACATCGGT

6001  GGCACCAAGATCCCTCCCAGGTGGCTGGGTCTGAGGGGAACTCCGAGCAGCCCTAGGTCC
6001  CCGTGGTTCTAGGGAGGGTCCACCGACCCAGACTCCCCTTGAGGCTCGTCGGGATCCAGG

6061  TCAAAGTCTGGATTTGTGTGGAAAAGGCAGCTCTCACTTGGCCTTGGCGAGGCCTCGGTT
6061  AGTTTCAGACCTAAACACACCTTTTCCGTCGAGAGTGAACCGGAACCGCTCCGGAGCCAA

6121  GGTTGGTGAGTGCCACACGGTTTCTTTGTGTGCTTGCATGGATTGGAATAGCCATTGTGT
6121  CCAACCACTCACGGTGTGCCAAAGAAACACACGAACGTACCTAACCTTATCGGTAACACA

6181  TCTTCCGTCTTCCCTGCTGGTGTTTCCACAGTGGGTGGCCTGAGCCCAGAGCAGCTCCCC
6181  AGAAGGCAGAAGGGACGACCACAAAGGTGTCACCCACCGGACTCGGGTCTCGTCGAGGGG

6241  ATATCCCTGTGCAGGCCACCTGTCTCGGGTGATGGAGAGCATCATTATGCTCCGTCTGAA
6241  TATAGGGACACGTCCGGTGGACAGAGCCCACTACCTCTCGTAGTAATACGAGGCAGACTT

6301  CGCTCTGCTTTCGGATGGCCCCATGCTCCACCTCCTGATAGCTCGTGGCGCGGGGCCACG
6301  GCGAGACGAAAGCCTACCGGGGTACGAGGTGGAGGACTATCGAGCACCGCGCCCCGGTGC

6361  GCTTAACAAATGGCTGAAAATGGGTCCTAATTAGTGGAAAAGTGCTTTCTTCATATTTTC
6361  CGAATTGTTTACCGACTTTTACCCAGGATTAATCACCTTTTCACGAAAGAAGTATAAAAG

6421  TCACTCGAGTGTGCAGTGATTCATTTTTCTTCTGCAATCAGCTCACTGCTAAAGTAAATC
6421  AGTGAGCTCACACGTCACTAAGTAAAAAGAAGACGTTAGTCGAGTGACGATTTCATTTAG

6481  TGACTCTCTTCCCGCCATTGCACACCAAAAGTTAACTCTAATGGGTAGGAGGTTAGGTTT
6481  ACTGAGAGAAGGGCGGTAACGTGTGGTTTTCAATTGAGATTACCCATCCTCCAATCCAAA

6541  GTTGAGAGAGCAATGCAGTAAAAAGAGGGGATCCAATGTGGTCTTGTCTGTCTGGTCTTC
6541  CAACTCTCTCGTTACGTCATTTTTCTCCCCTAGGTTACACCAGAACAGACAGACCAGAAG

6601  CTTTCTTCGTTTTTTCCTCCCTTGTCTTCTCTGTCATTCCCTTCCCTCCATTTGCCTTGC
6601  GAAAGAAGCAAAAAAGGAGGGAACAGAAGAGACAGTAAGGGAAGGGAGGTAAACGGAACG

6661  CTTTCCTGTCCTTCCCTTCCCTTCCTTCCCCTCTTTCTTTCTATAATTGGTGGGGGGTTT
6661  GAAAGGACAGGAAGGGAAGGGAAGGAAGGGGAGAAAGAAAGATATTAACCACCCCCCAAA

6721  GCACAGACTGCCAAAACACTAAGAACTGTGTAAAGTGTTTTTGAATGGCCTTACACATAT
6721  CGTGTCTGACGGTTTTGTGATTCTTGACACATTTCACAAAAACTTACCGGAATGTGTATA

6781  TGAAGTAGATTTTTATGCTCCATTTTTGAGATCACACACTAAAATCTATACCTTTAAAGC
6781  ACTTCATCTAAAAATACGAGGTAAAAACTCTAGTGTGTGATTTTAGATATGGAAATTTCG

6841  ATTTTCTGTTAGTTTGAAACTATTTGAAAATGAACAATGTGGTTTAGATTAGAGTCCTGT
6841  TAAAAGACAATCAAACTTTGATAAACTTTTACTTGTTACACCAAATCTAATCTCAGGACA

6901  TCTGAAGCTAGGAGTTCCACTATGAATATTGATTTATCAGTTTTTGACAAATTTTTGTTG
6901  AGACTTCGATCCTCAAGGTGATACTTATAACTAAATAGTCAAAAACTGTTTAAAAACAAC

6961  TTATACCAGATTTTCACTGGCAAACCTAGAGCAAATAAAATTCCACATAAGATACTTCCC
6961  AATATGGTCTAAAAGTGACCGTTTGGATCTCGTTTATTTTAAGGTGTATTCTATGAAGGG

7021  TAGACCTAATGGGAAAAATGTTTAATTTAGAGTCTTTAGGAGAAATGAGAATGAGGAATT
```

```
7021  ATCTGGATTACCCTTTTTACAAATTAAATCTCAGAAATCCTCTTTACTCTTACTCCTTAA

7081  GACCTTTTGTAAGCTTACTTCTGAGGCACTCTGAAGTGTGTTCCAGTGCTTTTAATGGAA
7081  CTGGAAAACATTCGAATGAAGACTCCGTGAGACTTCACACAAGGTCACGAAAATTACCTT

7141  ACTAGAGAGAGCCAGCAACCCCCTAGTGTGAGCCCCACTTTTAACCGGAAAAAGTGACCT
7141  TGATCTCTCTCGGTCGTTGGGGGATCACACTCGGGGTGAAAATTGGCCTTTTTCACTGGA

7201  TTTCCTCCTCCTTTGTGCTGAGTTTTGCGTAGGGCAGAAAATTAAGCTGATATTCAAAGA
7201  AAAGGAGGAGGAAACACGACTCAAAACGCATCCCGTCTTTTAATTCGACTATAAGTTTCT

7261  GATTCACTGCAAAAACATATTGATAAATCGTATATTCTATTTCATTAAATTAAAACCATA
7261  CTAAGTGACGTTTTTGTATAACTATTTAGCATATAAGATAAAGTAATTTAATTTTGGTAT

7321  CTGCTAATTATCTCAGGTTGTTAAACATAAGGCAATTAATTATCATTTTAAAAGTTGGTA
7321  GACGATTAATAGAGTCCAACAATTTGTATTCCGTTAATTAATAGTAAAATTTTCAACCAT

7381  GGAAGTTGTGAGTACTTTTGCAGTATGAGTGTTTTCCCGCTTTAGTATGAGGTTGTGTAT
7381  CCTTCAACACTCATGAAAACGTCATACTCACAAAAGGGCGAAATCATACTCCAACACATA

7441  GTTTGCTTGAATTTACAGAATTTTCACTTTAAGAGCAGACAATGTTTTGTTAAAGAAATG
7441  CAAACGAACTTAAATGTCTTAAAAGTGAAATTCTCGTCTGTTACAAAACAATTTCTTTAC

7501  AAATTTGCTAAAAAGGAGCATGTAAAGTGAAACATTAAAAATAAATAATTTCAACTTACT
7501  TTTAAACGATTTTTCCTCGTACATTTCACTTTGTAATTTTTATTTATTAAAGTTGAATGA

7561  TAAGAGCTGCAGAAAAATCTGATTGCTGTGTTTAAAATGAATTTTCCCACATTTCGCTCT
7561  ATTCTCGACGTCTTTTTAGACTAACGACACAAATTTTACTTAAAAGGGTGTAAAGCGAGA

7621  CTTATGGACAGGAGCATTTTCTGTCAGGTTATAAATAAAGACATGCCCATTTTTTGTACC
7621  GAATACCTGTCCTCGTAAAAGACAGTCCAATATTTATTTCTGTACGGGTAAAAAACATGG

7681  CCCACAAATGAGGAAGTTGTAAGCTCTCTGAGGTTTTACTGATGAGCCCCCTCCCCCTGG
7681  GGGTGTTTACTCCTTCAACATTCGAGAGACTCCAAAATGACTACTCGGGGGAGGGGGACC

7741  GTTTGCATGAAGAGATCATAGGCCACAAATAAAGGACTACAAAATGGGGTCTAAACTATC
7741  CAAACGTACTTCTCTAGTATCCGGTGTTTATTTCCTGATGTTTTACCCCAGATTTGATAG

7801  CTGGTGGGGCCTGATACCCACGTTTCGCATGGACCTTACGATGTGATGAATGGTTTTGGC
7801  GACCACCCCGGACTATGGGTGCAAAGCGTACCTGGAATGCTACACTACTTACCAAAACCG

7861  ATGAGTGTCTTAAGAATGCTTCCAGATTCGGGTTACAGGACAGCCAGCGCTGAGCTCCCT
7861  TACTCACAGAATTCTTACGAAGGTCTAAGCCCAATGTCCTGTCGGTCGCGACTCGAGGGA

7921  ATTGCAGAACAAAGTAGGAATCTAGAACTTTCTTGCTAACAGGATCCAGCTAAAACACCA
7921  TAACGTCTTGTTTCATCCTTAGATCTTGAAAGAACGATTGTCCTAGGTCGATTTTGTGGT

7981  AGTTAGATTCTTAAATGATGTTCTTTTCTGTCATTATTTGATTGTTGTCAGTAGCAGTAA
7981  TCAATCTAAGAATTTACTACAAGAAAAGACAGTAATAAACTAACAACAGTCATCGTCATT

8041  TTGTTACCAAGCCATTGATGCTTCTATTCTTCCCTTTGCCCTTCTGAGACACAGCTCATT
8041  AACAATGGTTCGGTAACTACGAAGATAAGAAGGGAAACGGGAAGACTCTGTGTCGAGTAA

8101  TTGACTTCAGTGGAACCCCTCGAAGGTGGGGTGATGAGCAAGGTGAATTTTCAAAGTAAA
8101  AACTGAAGTCACCTTGGGGAGCTTCCACCCCACTACTCGTTCCACTTAAAAGTTTCATTT

8161  GCTACTAAGAGACCAAACTACAATTTAAGGAACCTGATTTTTGAATCAAATTCCATATAC
8161  CGATGATTCTCTGGTTTGATGTTAAATTCCTTGGACTAAAAACTTAGTTTAAGGTATATG

8221  TGTGGGTATAGTTCAACATAGATTAATTTCTTATAGTTATTATGAAAAAAATCTCATCTT
8221  ACACCCATATCAAGTTGTATCTAATTAAAGAATATCAATAATACTTTTTTTAGAGTAGAA

8281  GATGATAGCTGATAATTTTGTGGGTGTCGTAAACAAAACAGAGGTCAGAATTCAGTCCCT
8281  CTACTATCGACTATTAAAACACCCACAGCATTTGTTTTGTCTCCAGTCTTAAGTCAGGGA

8341  TGGGGAAAATTTCCAATTAGTAGGAAACCAAGTGGCCTACCTTAGTTTGAAGACACCCAT
8341  ACCCCTTTTAAAGGTTAATCATCCTTTGGTTCACCGGATGGAATCAAACTTCTGTGGGTA

8401  CAGGATGTCTGCACCTTTTCATCCTCTCTGGAGGAAAGACTAAATACCCATTATTGTATA
8401  GTCCTACAGACGTGGAAAAGTAGGAGAGACCTCCTTTCTGATTTATGGGTAATAACATAT

8461  TAGGTCAGGCCAAAGCAGCCTTTTATATTGCAAGGAATAAGAGGTAAATAGATATATGTG
8461  ATCCAGTCCGGTTTCGTCGGAAAATATAACGTTCCTTATTCTCCATTTATCTATATACAC
```

8521 CAACAATGAATCCCCTAATGTGTTTACTCTAGAACACATGTTCTTTCTGTATTTATATGT
8521 GTTGTTACTTAGGGGATTACACAAATGAGATCTTGTGTACAAGAAAGACATAAATATACA

8581 AGATTTTGTAGATCTTGTCTTACCACCTGCTAATGGTAGATACTGTATCTAAATAAGTTG
8581 TCTAAAACATCTAGAACAGAATGGTGGACGATTACCATCTATGACATAGATTTATTCAAC

8641 AGGAAAATTTATAGTACCTAGGAATGTGTCCTCAGTGGGCCAATCAATCAATCATGACTT
8641 TCCTTTTAAATATCATGGATCCTTACACAGGAGTCACCCGGTTAGTTAGTTAGTACTGAA

8701 CAGGTTATTTTTAATAAATATACACGTATGGGTTCATAAACAATGGGATGTTCTTGTGAA
8701 GTCCAATAAAAATTATTTATATGTGCATACCCAAGTATTTGTTACCCTACAAGAACACTT

8761 GATCTAAATAATTTTACTTCTTTGGGACTAAATAAAATATAGCTTTTGCCAAATAAACTC
8761 CTAGATTTATTAAAATGAAGAAACCCTGATTTATTTTATATCGAAAACGGTTTATTTGAG

8821 ACACAAGCACTTATTTTAATAGAAGTCAAATGGCTTTGCAGAAACTTCAGTTTTACAGGT
8821 TGTGTTCGTGAATAAAATTATCTTCAGTTTACCGAAACGTCTTTGAAGTCAAAATGTCCA

8881 GCATTGTTTGAAATGTTACGGGTATACAAGTGGATTTCTCTATTATGTACAGTGTTAAGT
8881 CGTAACAAACTTTACAATGCCCATATGTTCACCTAAAGAGATAATACATGTCACAATTCA

8941 TTGAGTTTCAAAATGTCCACCTGAAATGATTTACTTGTACGTTAAGATAATTTAACTGCT
8941 AACTCAAAGTTTTACAGGTGGACTTTACTAAATGAACATGCAATTCTATTAAATTGACGA

9001 AAGAAGGCAAGATAAAGCATTCTTTGTGACACCATATGGCCTTGCTGAGGGAAAAACTTA
9001 TTCTTCCGTTCTATTTCGTAAGAAACACTGTGGTATACCGGAACGACTCCCTTTTTGAAT

9061 CTGTTATAAGTTTGTGTTTATCTCTCTTTTTAAAAAAAAATGAAGAAAAAAACGTTTAAA
9061 GACAATATTCAAACACAAATAGAGAGAAAAATTTTTTTTTACTTCTTTTTTTGCAAATTT

9121 ATAATGGGAACACAGCAGTTCCTGGGGTCCTCTGTCTCTTTATCTTATTATAGTAAATTA
9121 TATTACCCTTGTGTCGTCAAGGACCCCAGGAGACAGAGAAATAGAATAATATCATTTAAT

9181 CCAAAAAAATAATGACCTGGGGCATGTCTGTGTGGACCCTTCTTTTAGAGGCAGTTTCTG
9181 GGTTTTTTTATTACTGGACCCCGTACAGACACACCTGGGAAGAAAATCTCCGTCAAAGAC

9241 TGTTTTGTAAAGCTGTAGGTTCTATTTTCATTGCACTTCATATTGCTGCACAGCTCCTGA
9241 ACAAAACATTTCGACATCCAAGATAAAAGTAACGTGAAGTATAACGACGTGTCGAGGACT

9301 CCATGCATGAAGGTCCTCTGAAATCGGTAAGAGGGCAGAAGAAAATGATTCTAAACTTAG
9301 GGTACGTACTTCCAGGAGACTTTAGCCATTCTCCCGTCTTCTTTTACTAAGATTTGAATC

9361 ATTTTTTTAACTTAAGTGATGAAGTGTGAAACGCCATTTATATTTGAGGAAGCTACCTAG
9361 TAAAAAAATTGAATTCACTACTTCACACTTTGCGGTAAATATAAACTCCTTCGATGGATC

9421 GAAGTGGCTCATGTCGATGGCCCAAATCAGAAGAGGGCCTGTAAAAGCTTCTATCAATTT
9421 CTTCACCGAGTACAGCTACCGGGTTTAGTCTTCTCCCGGACATTTTCGAAGATAGTTAAA

9481 TGACTGTGTATGCTTCTACCATGGCGGCTCAATAAACAGCAGTATTAGTTTAAGAGTGGA
9481 ACTGACACATACGAAGATGGTACCGCCGAGTTATTTGTCGTCATAATCAAATTCTCACCT

9541 TGGTACAGTAGTATAGACGGGAAGCCTCTCCTCTCCGTGTGAACCGTGCACCCCTATGAG
9541 ACCATGTCATCATATCTGCCCTTCGGAGAGGAGAGGCACACTTGGCACGTGGGGATACTC

9601 AGGGTAGAGACAATACAATATGCCTGTAACGTCAGGACAGACAGTCATGGCCAGCTTGAA
9601 TCCCATCTCTGTTATGTTATACGGACATTGCAGTCCTGTCTGTCAGTACCGGTCGAACTT

9661 CTCCAGCCCTGGGCTTCTTGCAGCAACAAACGTGAACACAGAGGACTGTCTCCAACTCCA
9661 GAGGTCGGGACCCGAAGAACGTCGTTGTTTGCACTTGTGTCTCCTGACAGAGGTTGAGGT

9721 CTTTCTCTATTTTTAAAACAACTTTTTGAATACAGTATCTGCCATCTTTTCTTATACCTC
9721 GAAAGAGATAAAAATTTTGTTGAAAAACTTATGTCATAGACGGTAGAAAAGAATATGGAG

9781 ACTTTGAAACAGGTGGCTCCACTGTGGCATTTAAAATGTTCTGTTTCTTTTCCCTCTGTA
9781 TGAAACTTTGTCCACCGAGGTGACACCGTAAATTTTACAAGACAAAGAAAAGGGAGACAT

9841 TCAAATACCTCTTTACCAAGAAAACATTCAAACAGCATAGTTTTTAACTGTATTTTGAAA
9841 AGTTTATGGAGAAATGGTTCTTTTGTAAGTTTGTCGTATCAAAAATTGACATAAAACTTT

9901 GGTTTCCTTAGTTCCCTTTGACCCTTCCTCTTTTGCATATCAGTTCCTGGCCATAAAAAT
9901 CCAAAGGAATCAAGGGAAACTGGGAAGGAGAAAACGTATAGTCAAGGACCGGTATTTTTA

```
 9961  AAAAAATGCTAGGACAGAATTGCACATCTGAGCTGATTTGCCCTCAAAAAGTTTCACAGT
 9961  TTTTTTACGATCCTGTCTTAACGTGTAGACTCGACTAAACGGGAGTTTTTCAAAGTGTCA

10021  GGAACAAACCGCAGGAGGAGTTTTCTGTGGCTCAGTTAAATGTCGGGGGAGGGTGGTGTG
10021  CCTTGTTTGGCGTCCTCCTCAAAAGACACCGAGTCAATTTACAGCCCCCTCCCACCACAC

10081  AAAGCCAAATTGGATTCCTGCTTTCCTGTTTAAATCTTGTTTTTCATTGTTATTTGCACC
10081  TTTCGGTTTAACCTAAGGACGAAAGGACAAATTTAGAACAAAAAGTAACAATAAACGTGG

10141  AGCAATACTCTGTGGAATAATCATGAAAATGTGTAGATTGGCAGCTAATTTTTGAAAAAT
10141  TCGTTATGAGACACCTTATTAGTACTTTTACACATCTAACCGTCGATTAAAAACTTTTTA

10201  GAAAAGAATCAGAAATGAAATAAGAGTGCTCGGAAGTTTTTATGTTCTCTCAACCTGTTT
10201  CTTTTCTTAGTCTTTACTTTATTCTCACGAGCCTTCAAAAATACAAGAGAGTTGGACAAA

10261  TGTCAAATTGTTACGAAAACCTATAAGGTCTCTTTGACTAGATACAAAGACTTTGCACAT
10261  ACAGTTTAACAATGCTTTTGGATATTCCAGAGAAACTGATCTATGTTTCTGAAACGTGTA

10321  TGCCTTAGCTTTCTCTTGAAGCATTTCCTTTTTTAAAATACAGTGTAATTCACAGTGATA
10321  ACGGAATCGAAAGAGAACTTCGTAAAGGAAAAAATTTTATGTCACATTAAGTGTCACTAT

10381  TGATAGATTTGCAAAAGTAAAATCTACCAGTCTGAAGATGAAAGGACTTGTCTCTTAGCA
10381  ACTATCTAAACGTTTTCATTTTAGATGGTCAGACTTCTACTTTCCTGAACAGAGAATCGT

10441  GGAATAATGGGTTTTATTAAAGAGGTCTGTGACCTAAGGCATTTTAAATAAATTACAGGC
10441  CCTTATTACCCAAAATAATTTCTCCAGACACTGGATTCCGTAAAATTTATTTAATGTCCG

10501  TTGGTCCCTGTCTCCCCCATGTATCTACTCCCTTCAATATAAGCATCATTGAGTATTTAA
10501  AACCAGGGACAGAGGGGGTACATAGATGAGGGAAGTTATATTCGTAGTAACTCATAAATT

10561  GGAAATAACCCCAAATGTAACTCTAGTGTAGCTTCACTTGTCAGGGAGGAAAAAGTAAAT
10561  CCTTTATTGGGGTTTACATTGAGATCACATCGAAGTGAACAGTCCCTCCTTTTTCATTTA

10621  AGCATACATTTGGCCAAATAACCAGAACTTTACTGTAGAAGTTTTATGATGAAATTTGCC
10621  TCGTATGTAAACCGGTTTATTGGTCTTGAAATGACATCTTCAAAATACTACTTTAAACGG

10681  TTTAGTGCAGAGTATTACAAAGATCATGTTTAGTTTCTAGCAGTATATAAGTAGCATCCA
10681  AAATCACGTCTCATAATGTTTCTAGTACAAATCAAAGATCGTCATATATTCATCGTAGGT

10741  TCCTTATCTGTCATGCATTTGGAGTGTGCGACCCCTGCACTGGGCTGCAACATTCTGATG
10741  AGGAATAGACAGTACGTAAACCTCACACGCTGGGGACGTGACCCGACGTTGTAAGACTAC

10801  GGCAAGAGTGCTAGGGAGAAAGAGGCATCACCATCAGACTGCACGGGTTCAAGTGTCAGC
10801  CCGTTCTCACGATCCCTCTTTCTCCGTAGTGGTAGTCTGACGTGCCCAAGTTCACAGTCG

10861  TCTGTGGTTGATTAGCTGTGTGACCTGGGGAAAGCTATTTCTCTTAGCCTTGGTTCTCTC
10861  AGACACCAACTAATCGACACACTGGACCCCTTTCGATAAAGAGAATCGGAACCAAGAGAG

10921  ATCTATAAAATGGAGATAATGATGCAGATGCCTTGGGTTTAATTGGGAGAGTTAAAGACA
10921  TAGATATTTTACCTCTATTACTACGTCTACGGAACCCAAATTAACCCTCTCAATTTCTGT

10981  CATTTACATATTTAGCAAGTAGGTGTTGAATTCTAGCTCTACATTGGACACTATGCCAGG
10981  GTAAATGTATAAATCGTTCATCCACAACTTAAGATCGAGATGTAACCTGTGATACGGTCC

11041  TGCTCAAATAAACAAGTGGACAAGACAGACAACACCCATGGTCTTATGAGGCTTAACCAT
11041  ACGAGTTTATTTGTTCACCTGTTCTGTCTGTTGTGGGTACCAGAATACTCCGAATTGGTA

11101  TTGCCTCTTCAATGCCAGAAACTTAGTAGGTTGATTAGATAAAGCCAGTGAGTACCAGTA
11101  AACGGAGAAGTTACGGTCTTTGAATCATCCAACTAATCTATTTCGGTCACTCATGGTCAT

11161  TCCTTTTCTTTGCAGCCTTTTCCTGGCACACTAAAAATACTCAGTACATATGAAATATCA
11161  AGGAAAAGAAACGTCGGAAAAGGACCGTGTGATTTTTATGAGTCATGTATACTTTATAGT

11221  CTGGACAAAGAATCCCCCTTAGAGTACCAGTGGAGAAGGAAGGCATTTGCTTAAAAGCAA
11221  GACCTGTTTCTTAGGGGGAATCTCATGGTCACCTCTTCCTTCCGTAAACGAATTTTCGTT

11281  ACCAACAGAAAGACATTGTAAGGCAGTTGTTTAAGTCTCAGAGACCTATAATTTTTTTCT
11281  TGGTTGTCTTTCTGTAACATTCCGTCAACAAATTCAGAGTCTCTGGATATTAAAAAAAGA

11341  TTTTTCTTTTTTTTCATCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCACAATCTCAG
11341  AAAAAGAAAAAAAAAGTAGAGCGAGACAGCGGGTCCGACCTCACGTCACCGTGTTAGAGTC

11401  CTCACTGCAAGCTCCACCTTCCGGGTTCATGCCATTCTTCTGCCTCAGCCTCCCAAGTAG
```

```
11401  GAGTGACGTTCGAGGTGGAAGGCCCAAGTACGGTAAGAAGACGGAGTCGGAGGGTTCATC

11461  CAGAGACTACAGGCGCCCGCCACCACACCTGGCTAATTTTTTGTATTTTTAGTGGAGACG
11461  GTCTCTGATGTCCGCGGGCGGTGGTGTGGACCGATTAAAAAACATAAAAATCACCTCTGC

11521  GGGTTTCGCCGTGTTAGCCAGGATGGTCTTGATCTCCTGACCTCATGATCCGCCTGCCTC
11521  CCCAAAGCGGCACAATCGGTCCTACCAGAACTAGAGGACTGGAGTACTAGGCGGACGGAG

11581  GGCCTCCCAAAGTGCTGGGATTACTGGCATGAGCCACCACGCCCGGCAACTACAATTGTT
11581  CCGGAGGGTTTCACGACCCTAATGACCGTACTCGGTGGTGCGGGCCGTTGATGTTAACAA

11641  CTTAAAGCTTGTAGAATTACTGTGTGCTACCAACAGACAGGCTAATTTTGAGTGACCCTC
11641  GAATTTCGAACATCTTAATGACACACGATGGTTGTCTGTCCGATTAAAACTCACTGGGAG

11701  AGTACTTTGTACAGTTAATTTGGCACGCTGTGTACTTAGTGGCTTTTTAACAGCTATAAA
11701  TCATGAAACATGTCAATTAAACCGTGCGACACATGAATCACCGAAAAATTGTCGATATTT

11761  TTTGGGCTGCTAGAAAAGTAGTAAAGTTGTGATTCTTGACAGGCATCTATCTGCATTTTC
11761  AAACCCGACGATCTTTTCATCATTTCAACACTAAGAACTGTCCGTAGATAGACGTAAAAG

11821  ATTTTTACTTCATTTGTCTAGACTCAGCTTGTCAGAATTATGGAAGAGACTCCTTGTGTC
11821  TAAAAATGAAGTAAACAGATCTGAGTCGAACAGTCTTAATACCTTCTCTGAGGAACACAG

11881  AGGGCAAGCACTGTGAAGAGAGGTATTCACTGTCAGAAAAGAGAGGGGAGCTGGAGGCAG
11881  TCCCGTTCGTGACACTTCTCTCCATAAGTGACAGTCTTTTCTCTCCCCTCGACCTCCGTC

11941  CTCAGAGGCCTGAGACCCGCCTCCACAGGAGCCCCAGCAGGTTCGGTGGAGCTCTGGCCA
11941  GAGTCTCCGGACTCTGGGCGGAGGTGTCCTCGGGGTCGTCCAAGCCACCTCGAGACCGGT

12001  CACTCTCCTTTGGGATGCTGAAGTCAGAATGAGTTCACTTCCCAGCCAGTCTTGCCAAGG
12001  GTGAGAGGAAACCCTACGACTTCAGTCTTACTCAAGTGAAGGGTCGGTCAGAACGGTTCC

12061  CTCCTCACCTGGAAGCAGCAACTGCCCAGGGCTGTTGGATGTTTCTCCCCAGGGGACAGC
12061  GAGGAGTGGACCTTCGTCGTTGACGGGTCCCGACAACCTACAAAGAGGGGTCCCCTGTCG

12121  CAGGTCCCAGTCCCGCCTCGGTGTGGAAGGAGGAAAGGCAGGGTCCAGGAAGCTGTTTCA
12121  GTCCAGGGTCAGGGCGGAGCCACACCTTCCTCCTTTCCGTCCCAGGTCCTTCGACAAAGT

12181  GGACAGGCCCAAGGTCCCCCAGGGATGCCTTTCAGGGTCAGCGGAGGCTGTAAATCAGCA
12181  CCTGTCCGGGTTCCAGGGGGTCCCTACGGAAAGTCCCAGTCGCCTCCGACATTTAGTCGT

12241  GGGCCCACACGGCCTGGAAGAGGCCCCTGTGCTGTCGGCTTGCCCGGCTTGCCCGGCTCC
12241  CCCGGGTGTGCCGGACCTTCTCCGGGGACACGACAGCCGAACGGGCCGAACGGGCCGAGG

12301  TAGTCCGGCTTCTGCTCCTCCTTTGTAAAGTTATGGATATGCTAATAGTTTCCAACTGAG
12301  ATCAGGCCGAAGACGAGGAGGAAACATTTCAATACCTATACGATTATCAAAGGTTGACTC

12361  ACTAGGAAAGTAAGTCCTACTTGACACTGTTTGGTCAGAAAGAGGGAGAGAAAGGAGAAG
12361  TGATCCTTTCATTCAGGATGAACTGTGACAAACCAGTCTTTCTCCCTCTCTTTCCTCTTC

12421  GACAGAGAGAGACTGAGAGAGAGACAGTCTCAGACAAAGGGAGACGGAGGGAGGGAGGGA
12421  CTGTCTCTCTCTGACTCTCTCTCTGTCAGAGTCTGTTTCCCTCTGCCTCCCTCCCTCCCT

12481  GAGACAGAGAAAGAGATGGGAGGTAGGTGTGGGAGGAGGGAGAGATGCAGAAGGCAGAGG
12481  CTCTGTCTCTTTCTCTACCCTCCATCCACACCCTCCTCCCTCTCTACGTCTTCCGTCTCC

12541  AAAGACAGACAGAGATTTAGACCTCCCAAGTCAGTGAGCAGTCCAGAGTTGGAGTGGAGG
12541  TTTCTGTCTGTCTCTAAATCTGGAGGGTTCAGTCACTCGTCAGGTCTCAACCTCACCTCC

12601  GTGCCTGGTGGCTTGTGACTGCAGACTCCACTCCCCGCTCCTAGAGGCACAGCCATGGAC
12601  CACGGACCACCGAACACTGACGTCTGAGGTGAGGGGCGAGGATCTCCGTGTCGGTACCTG

12661  AGCTTCTGTCACGTTGGCCCTGCACTTATCTCTGCATCTATTTCCCCTTGTGCAAGATTC
12661  TCGAAGACAGTGCAACCGGGACGTGAATAGAGACGTAGATAAAGGGGAACACGTTCTAAG

12721  AGAACTGCATGCTCCAAAAAAACAATAAAAGCATTCATGTTCATAAGAATTGCACAGGTA
12721  TCTTGACGTACGAGGTTTTTTTGTTATTTTCGTAAGTACAAGTATTCTTAACGTGTCCAT

12781  AAAGGTAGTTTGCTGATATTGTTGTATTTTTTACTATCGCTTCTTTTAGGTCTTGCCTGA
12781  TTTCCATCAAACGACTATAACAACATAAAAAATGATAGCGAAGAAAATCCAGAACGGACT

12841  AATTGTTTGGGTTTCCCAGGCAAAGTAGAAAACTGCGGTACGTTTCTGTGAAATAATTAT
12841  TTAACAAACCCAAAGGGTCCGTTTCATCTTTTGACGCCATGCAAAGACACTTTATTAATA
```

12901 TCCTTCTGGCATCTCCCTTTACAGACCTACTGATCTTGATTTTTCATTTAGGTGAAAGTT
12901 AGGAAGACCGTAGAGGGAAATGTCTGGATGACTAGAACTAAAAAGTAAATCCACTTTCAA

12961 TGTGAAAACATGCCATTAGCTTGCTTTGTGATTAACTCCTTTTACTGAATGTGAGCTCCT
12961 ACACTTTTGTACGGTAATCGAACGAAACACTAATTGAGGAAAATGACTTACACTCGAGGA

13021 TTTAAATTGAGGCCATATCAAGCTTAAATTCCATATTTTACCCGGCACTCTGCATTTCTT
13021 AAATTTAACTCCGGTATAGTTCGAATTTAAGGTATAAAATGGGCCGTGAGACGTAAAGAA

13081 CCATGTGGGAGAGGAGGGGCTCAGTAAGTGCTTTGTAAAATACACAGCCGAAGTGATGCA
13081 GGTACACCCTCTCCTCCCCGAGTCATTCACGAAACATTTTATGTGTCGGCTTCACTACGT

13141 CGTGCTAACAAAGGAGTGTGACAGGACTTAAGTGCCCTTCTAGACACTTCAGGCTCCCCT
13141 GCACGATTGTTTCCTCACACTGTCCTGAATTCACGGGAAGATCTGTGAAGTCCGAGGGGA

13201 TTGTAAGCTGTCTTGGAAGAGGCCACATTTCCTTTCCCTCAAACAGTTTCTCATTGTTTG
13201 AACATTCGACAGAACCTTCTCCGGTGTAAAGGAAAGGGAGTTTGTCAAAGAGTAACAAAC

13261 ATTATTCTTTTAGCCTTTCTCTGGAAGCAAAGCCACTTTTACGAGAAAGTCACTGCTTTT
13261 TAATAAGAAAATCGGAAAGAGACCTTCGTTTCGGTGAAAATGCTCTTTCAGTGACGAAAA

13321 TCATCTCAAGAGATGCAAGTTTGGAGTTTGGGGAAGTTTTCAGGTGCCCGTCAAGTCATC
13321 AGTAGAGTTCTCTACGTTCAAACCTCAAACCCCTTCAAAAGTCCACGGGCAGTTCAGTAG

13381 CTTTATGATGTCAGACGAGTCAGGCCACAGAATTCACAGGGCTCAGTGCAGACCGAAAAC
13381 GAAATACTACAGTCTGCTCAGTCCGGTGTCTTAAGTGTCCCGAGTCACGTCTGGCTTTTG

13441 TTGAGGCCTCTTGTTCAGAAATTATTAAAAATTTTGGTGAACATCACCCCAAGCAAAGAG
13441 AACTCCGGAGAACAAGTCTTTAATAATTTTTAAAACCACTTGTAGTGGGGTTCGTTTCTC

13501 ATCCCCTAAGCACCAGCCCCCAAGCAACTGCACTCATAAGCCCATGAAGCCCCCTGCTGT
13501 TAGGGGATTCGTGGTCGGGGGTTCGTTGACGTGAGTATTCGGGTACTTCGGGGGACGACA

13561 CAGAAACAATGTGGTTGAAATTGTGTATGCACTTGGAAGTGAGATGGATTGCAAAACACA
13561 GTCTTTGTTACACCAACTTTAACACATACGTGAACCTTCACTCTACCTAACGTTTTGTGT

13621 GGTCTCCATGCTGGGGCAGGAGTGGTGATAGGGCATGGAGTGGAAATGTCCAGCAGGCCC
13621 CCAGAGGTACGACCCCGTCCTCACCACTATCCCGTACCTCACCTTTACAGGTCGTCCGGG

13681 ACGTGCGAAAATGCAGAGCTCTCTGGCTCTTGCAGACTTGGCTGCTGACAATAGACGCGC
13681 TGCACGCTTTTACGTCTCGAGAGACCGAGAACGTCTGAACCGACGACTGTTATCTGCGCG

13741 TCCAGGAAGGTGCTCGCTGTGGTGTGATCTGCTGCCCACCCCTAGCTCCCTCCAGGAGAC
13741 AGGTCCTTCCACGAGCGACACCACACTAGACGACGGGTGGGGATCGAGGGAGGTCCTCTG

13801 TGGTGCGGGGACTGTTTGCAAATGACTGCAAAAGTAAGAAGGTTCCCACAGAGCAGAGCT
13801 ACCACGCCCCTGACAAACGTTTACTGACGTTTTCATTCTTCCAAGGGTGTCTCGTCTCGA

13861 TGATTTGGGGACCAGCCGAGGGCAGTTTGTCAGGATTCCGGCTTGAAACTGTTCTCACAT
13861 ACTAAACCCCTGGTCGGCTCCCGTCAAACAGTCCTAAGGCCGAACTTTGACAAGAGTGTA

13921 CTCACCGCCTGAAAGGACGAGTGTGTCCAGAGGACTTAGCATTGATCACCTCTGTCTCCA
13921 GAGTGGCGGACTTTCCTGCTCACACAGGTCTCCTGAATCGTAACTAGTGGAGACAGAGGT

13981 TGCAGCAAACTCAGAGGCTCAGCCCGCATTCCACTGGAAGGGCGTTTGCCAGTGGTGTTG
13981 ACGTCGTTTGAGTCTCCGAGTCGGGCGTAAGGTGACCTTCCCGCAAACGGTCACCACAAC

14041 GTTGGAAGAGCCTTGACTTTGCCTTAGGAAACATCTTTTTTTAAGAATTGAAAATAACTT
14041 CAACCTTCTCGGAACTGAAACGGAATCCTTTGTAGAAAAAAATTCTTAACTTTTATTGAA

14101 GAGTATGCAACAGTAGGGCATTTGTTATATAAATTAGTTGACTAGTGTGTAGCCAGTAAA
14101 CTCATACGTTGTCATCCCGTAAACAATATATTTAATCAACTGATCACACATCGGTCATTT

14161 ATGATGATGGTGGTGTGTATTTGTTAAATAAAAAGATATGTGTGGTATTAAATTAAAAAA
14161 TACTACTACCACCACACATAAACAATTTATTTTTCTATACACACCATAATTTAATTTTTT

14221 TATTTTAAAACAACATATTTGTAATCTGTTTAGTGTCCTCTTTTTGTAAAAAGTACAGAA
14221 ATAAAATTTTGTTGTATAAACATTAGACAAATCACAGGAGAAAAACATTTTTCATGTCTT

14281 ATAAATATACAGAAAAAATAGTAGTCCTAAGTGGTAGAAATTATGAGCATTTTCTTGCCT
14281 TATTTATATGTCTTTTTTATCATCAGGATTCACCATCTTTAATACTCGTAAAAGAACGGA

```
14341  TTAAAAAAAGTTGTAAAAGATTGTATCATTTATGTAGCAAAAAGTTTTAAGTCAGCATTC
14341  AATTTTTTTCAACATTTTCTAACATAGTAAATACATCGTTTTTCAAAATTCAGTCGTAAG

14401  TAAAAATTTCGTGTTGTTATAGTTGCTGTGACAAGATTTAACTTCTGTATGCTTCACCAA
14401  ATTTTTAAAGCACAACAATATCAACGACACTGTTCTAAATTGAAGACATACGAAGTGGTT

14461  TCAATACAGAGGTATTTAAGACCCGGTGTGTGATAGGCCGCGCTAAAATACTATACACAT
14461  AGTTATGTCTCCATAAATTCTGGGCCACACACTATCCGGCGCGATTTTATGATATGTGTA

14521  CTTCAGAAAACTAGAGAACTAACTTCTAACTTCCTATATTAGTGTGGCACGGCTGTTACA
14521  GAAGTCTTTTGATCTCTTGATTGAAGATTGAAGGATATAATCACACCGTGCCGACAATGT

14581  AAGATTTTTCTCATTTGAGTCTATCTTGCTTCTTTATCATTGTTTTGACAGTTTCAGAAG
14581  TTCTAAAAAGAGTAAACTCAGATAGAACGAAGAAATAGTAACAAAACTGTCAAAGTCTTC

14641  AATCGTGGCTTTTCCCCTTTTTTACAGTAAAGGTACCTGAGACTCTTGACGTATTGCTTT
14641  TTAGCACCGAAAAGGGGAAAAAATGTCATTTCCATGGACTCTGAGAACTGCATAACGAAA

14701  TTGGAAATGCTTGTGCTGGTCACATGCTTGCATCTGGGCTAGTGTGTCTGGCTTCCGTGT
14701  AACCTTTACGAACACGACCAGTGTACGAACGTAGACCCGATCACACAGACCGAAGGCACA

14761  GCTGGTGGATGCTTACTCTGTTTTCTGAAATACTTTTTCTGTACAGTGGCCACTAGCTGT
14761  CGACCACCTACGAATGAGACAAAAGACTTTATGAAAAAGACATGTCACCGGTGATCGACA

14821  ACTCCTAAGCCACACACCTACCTTGAAAATTCATGTCACTTTTAGAAATAGATAAAAGCC
14821  TGAGGATTCGGTGTGTGGATGGAACTTTTAAGTACAGTGAAAATCTTTATCTATTTTCGG

14881  CCTCCCATCCAGAAAAAGTGACTATCATGTATATCCTCATCATGACTAATACTGATATTC
14881  GGAGGGTAGGTCTTTTTCACTGATAGTACATATAGGAGTAGTACTGATTATGACTATAAG

14941  CTGAAATTGAAAATACATATTCCATATGTACCATAAAAGGTATTAAAGATATATGGAGTG
14941  GACTTTAACTTTTATGTATAAGGTATACATGGTATTTTCCATAATTTCTATATACCTCAC

15001  ATAGATATATTATATATAACACTTCTACCCTCACAGTTTTCAGCCTAATTGAGAGGGTAA
15001  TATCTATATAATATATATTGTGAAGATGGGAGTGTCAAAAGTCGGATTAACTCTCCCATT

15061  GATCCCTGAATCATCCATCAGTTTTTCAGGTCTCTGCTGAAAGCAGGCCACAGCTCAGAT
15061  CTAGGGACTTAGTAGGTAGTCAAAAAGTCCAGAGACGACTTTCGTCCGGTGTCGAGTCTA

15121  CCACACATCTGAACCAGAGACAGAGGTGGCCAAAAATAAAAAGGGGGACAGGGGGACAAC
15121  GGTGTGTAGACTTGGTCTCTGTCTCCACCGGTTTTTATTTTTCCCCCTGTCCCCCTGTTG

15181  CTGGTTTAGAGTCAACAAATAGACTGCATTTTCTGGTTAGTGAAGGAGCTCTCCTGAAAG
15181  GACCAAATCTCAGTTGTTTATCTGACGTAAAAGACCAATCACTTCCTCGAGAGGACTTTC

15241  TCATATACCAGAGCATAAATGAGCAGATTTCCTTGAGGTCACCTTCTGCTGGCCATAGCT
15241  AGTATATGGTCTCGTATTTACTCGTCTAAAGGAACTCCAGTGGAAGACGACCGGTATCGA

15301  TTCTTATCTGTGGAGCTGCCAGCTGTCATCCACTTTGGGGCACCTGAGACTGCCGAGCGG
15301  AAGAATAGACACCTCGACGGTCGACAGTAGGTGAAACCCCGTGGACTCTGACGGCTCGCC

15361  CAGGCCAGGACCCAAGTGCGAAAACACAGAACACCTTTTTGTTTCTACTCCACTGATGCT
15361  GTCCGGTCCTGGGTTCACGCTTTTGTGTCTTGTGGAAAAACAAAGATGAGGTGACTACGA

15421  GGGGTTCTCTCCCTGGTGTTTGTGGCTCGTAGTACACTCTGTGGAACATTCACTATGGTC
15421  CCCCAAGAGAGGGACCACAAACACCGAGCATCATGTGAGACACCTTGTAAGTGATACCAG

15481  ATCGAAGGGCAGCATCTTCCCAGTTGTTTCTTTCTTTTCTTTTTTTTTTTTAATTTAAAC
15481  TAGCTTCCCGTCGTAGAAGGGTCAACAAAGAAAGAAAAGAAAAAAAAAAAATTAAATTTG

15541  CGATCTGAGAAGCCAGCCATCTGTCAGCAAAACAGGAAGGCTCGGGCTGTCTCCTGGGCT
15541  GCTAGACTCTTCGGTCGGTAGACAGTCGTTTTGTCCTTCCGAGCCCGACAGAGGACCCGA

15601  CGTTTTGCTGCCGTAGTGAGCGTCACTTCTCCCCGTGTAAGAGTGCTGGTGAAGGCTGAG
15601  GCAAAACGACGGCATCACTCGCAGTGAAGAGGGGCACATTCTCACGACCACTTCCGACTC

15661  GCAAGGGCCCAGAAAGATTGAGGGACAAAGACAGGAGCGCCCGCATTGCCCATCTGCCAG
15661  CGTTCCCGGGTCTTTCTAACTCCCTGTTTCTGTCCTCGCGGGCGTAACGGGTAGACGGTC

15721  GCTGGAGGTGTATTCATTATTGATGGAGGTAGTGCAGTTGCTGCTCAGATATGCAGCCCT
15721  CGACCTCCACATAAGTAATAACTACCTCCATCACGTCAACGACGAGTCTATACGTCGGGA

15781  GCCTGGGTAAATGAGACATTCTTCAGCAAATTGCTTCGTTTTTTGATTGCTGATTGTACG
```

```
15781  CGGACCCATTTACTCTGTAAGAAGTCGTTTAACGAAGCAAAAAACTAACGACTAACATGC

15841  CGTGTCACCAAGCTGACTCAAGGTTCATCGATGCATGCTCAGTAAATTAGAAAGAACATA
15841  GCACAGTGGTTCGACTGAGTTCCAAGTAGCTACGTACGAGTCATTTAATCTTTCTTGTAT

15901  ACTATGGATCAGCCAAGAGAATGAATTCTGTGCCTACAATGACCCAGGGCCATTTAATTT
15901  TGATACCTAGTCGGTTCTCTTACTTAAGACACGGATGTTACTGGGTCCCGGTAAATTAAA

15961  TCTGCTTAATTTTGTTGCAGTCAGTTTGCATTTTGGGTTATTATGCAGTAGGAAATTAAC
15961  AGACGAATTAAAACAACGTCAGTCAAACGTAAAACCCAATAATACGTCATCCTTTAATTG

16021  AATAAATAACAAATTTGGTCCTCCTGTGCTTGTAATGATATTTTTATAAATCTTTGTAAT
16021  TTATTTATTGTTTAAACCAGGAGGACACGAACATTACTATAAAAATATTTAGAAACATTA

16081  GCTGTTTTTAAAAGGATCAAGGTCTGTGCCAGTCTGATACTCCAGCAAGTATGTGAGGAG
16081  CGACAAAAATTTTCCTAGTTCCAGACACGGTCAGACTATGAGGTCGTTCATACACTCCTC

16141  GAAAATGCATTATTCTTGCTAGATAACCTTGTTGTTAAATAGCATAGGGGTTCTTTATCT
16141  CTTTTACGTAATAAGAACGATCTATTGGAACAACAATTTATCGTATCCCCAAGAAATAGA

16201  CTCTCTCTTTCTCATATCTTATTAGTATTTTTGCTTTAAACTAAAATCCCTTCCTCTCTT
16201  GAGAGAGAAAGAGTATAGAATAATCATAAAAACGAAATTTGATTTTAGGGAAGGAGAGAA

16261  TCTCAGATAACCTGAGGACCATGGATGCTGATGAGGGTCAAGACATGTCCCAAGTTTCAG
16261  AGAGTCTATTGGACTCCTGGTACCTACGACTACTCCCAGTTCTGTACAGGGTTCAAAGTC

16321  GTGAGACCTTATGAGATAGCTGTGTGGGAAGTTCATGAGAAAAGCTTCCCTGGGGCCGGA
16321  CACTCTGGAATACTCTATCGACACACCCTTCAAGTACTCTTTTCGAAGGGACCCCGGCCT

16381  AGTCACAGTGCTTGGTATGCTCATGGGGGAGGAATAGGGGCTATTCTGCAAAAGAAAAGA
16381  TCAGTGTCACGAACCATACGAGTACCCCCTCCTTATCCCCGATAAGACGTTTTCTTTTCT

16441  CCATGATGGAATTTGCCTGAGTGTTTCCTTCACCTGTTACAAATTATCTCACTTTGAGCT
16441  GGTACTACCTTAAACGGACTCACAAAGGAAGTGGACAATGTTTAATAGAGTGAAACTCGA

16501  GAACAGAAAGCCTCCAAGATGAAATTAGTTTTACTGTTAAACTTCAGGAAAAAAAAACGG
16501  CTTGTCTTTCGGAGGTTCTACTTTAATCAAAATGACAATTTGAAGTCCTTTTTTTTTGCC

16561  GAAGAGTTAAATACATTTTTGTACTGTTGGAAGGAAAAATGGCTGATTGGTTTAAAACCC
16561  CTTCTCAATTTATGTAAAAACATGACAACCTTCCTTTTTACCGACTAACCAAATTTTGGG

16621  AAACACATGCCAATGATGGTACTTAAAGAGAGAGAGAGAGAGAAGCTTGAAAAACATAAT
16621  TTTGTGTACGGTTACTACCATGAATTTCTCTCTCTCTCTCTTCGAACTTTTTGTATTA

16681  TGTTGGGCACAGTCATGACTGTTTGTTCATTAAGCATGGACACAACATTGCTCCCCTTTG
16681  ACAACCCGTGTCAGTACTGACAAACAAGTAATTCGTACCTGTGTTGTAACGAGGGGAAAC

16741  CCATATATCTTTTCAAGCCGTATTGGATATAGCTCTTCTCATCCAGGAGACCCAGGAAGT
16741  GGTATATAGAAAAGTTCGGCATAACCTATATCGAGAAGAGTAGGTCCTCTGGGTCCTTCA

16801  GGAGAAGTCTGTAGTAGGAAAAGCCTAAGGGTAGGTCACAGACTGTGACCATTTGGCAGC
16801  CCTCTTCAGACATCATCCTTTTCGGATTCCCATCCAGTGTCTGACACTGGTAAACCGTCG

16861  ACTGAGGGTGGACGGCGAGCCAGTCCAACAAAACCGCACAGTTCCCCAGTGCATGGACAT
16861  TGACTCCCACCTGCCGCTCGGTCAGGTTGTTTTGGCGTGTCAAGGGGTCACGTACCTGTA

16921  AGGAAGACAGCTTTCTATCTGGCCCTGTATCCAGAGGCGTCAGCCCCAGTAGCAGCTTTC
16921  TCCTTCTGTCGAAAGATAGACCGGGACATAGGTCTCCGCAGTCGGGGTCATCGTCGAAAG

16981  ATGGACTTTGGGGTTTTCGGTATTTCATATTTTTGAGCCTCACAGACTCACAGCCAGCCC
16981  TACCTGAAACCCCAAAAGCCATAAAGTATAAAAACTCGGAGTGTCTGAGTGTCGGTCGGG

17041  CAGAGGCTGACTTATATTTGAGAAAGTTCTCAGTGGCACCTTGCCTTGGCTGAGCGCCCT
17041  GTCTCCGACTGAATATAAACTCTTTCAAGAGTCACCGTGGAACGGAACCGACTCGCGGGA

17101  CGTGTTTTGAAGTTTCTATGGGATTCTACAAGTTGGTGCTCCTGATGAAGACCAGGACCT
17101  GCACAAAACTTCAAAGATACCCTAAGATGTTCAACCACGAGGACTACTTCTGGTCCTGGA

17161  ATGTGTGGCTGCTCCCCTGCTTGGTGGTTTCCCTGGGGAAGGTGCAGGAGAGGATCTTCT
17161  TACACACCGACGAGGGGACGAACCACCAAAGGGACCCCTTCCACGTCCTCTCCTAGAAGA

17221  GAGTTCCATGGAACTGGAGATAGATCTGCCAATCACAGGCTTCCTTCTCCACCACTCCTC
17221  CTCAAGGTACCTTGACCTCTATCTAGACGGTTAGTGTCCGAAGGAAGAGGTGGTGAGGAG
```

17281 AGCCGCTCTATTCATGTTTCAGATTTTGGACTTAAACTCTCCCAGGTGCAAAGAACAAAC
17281 TCGGCGAGATAAGTACAAAGTCTAAAACCTGAATTTGAGAGGGTCCACGTTTCTTGTTTG

17341 AAAAGGCTAGCTTATTTTTCTTTTAGAGTGAGGCTTCGTATTTATTACAATATAATTGCC
17341 TTTTCCGATCGAATAAAAAGAAAATCTCACTCCGAAGCATAAATAATGTTATATTAACGG

17401 ACATTCTTTGTGTAATTCTCACATTTATATCTTAAATATAATTCTCATGAATGAGAATTA
17401 TGTAAGAAACACATTAAGAGTGTAAATATAGAATTTATATTAAGAGTACTTACTCTTAAT

17461 TATAATTCTCTTTTTGTATATCATTGAATATTTTCACTTAATTTTTAATTTTTTTAATCG
17461 ATATTAAGAGAAAAACATATAGTAACTTATAAAAGTGAATTAAAAATTAAAAAAATTAGC

17521 TCACAAAATAATTGTGTACATAGACACAAAATAATTGGGTACATAGTGATGTTGTGATAT
17521 AGTGTTTTATTAACACATGTATCTGTGTTTTATTAACCCATGTATCACTACAACACTATA

17581 ATACAATGTATAGTAATCGGATCAGGTAAATCAGCATATTCATCATCTCAAACATTTATC
17581 TATGTTACATATCATTAGCCTAGTCCATTTAGTCGTATAAGTAGTAGAGTTTGTAAATAG

17641 GTTTCTTTGTATTAGGAACATTCGACATCTTCCTTCTAGCTATTTGAAACTATATATTAT
17641 CAAAGAAACATAATCCTTGTAAGCTGTAGAAGGAAGATCGATAAACTTTGATATATAATA

17701 TGTTGACTACAGTCATCCTGCAATGGTGTAGAACACTAGAACTTATTCTTCCTACCTAGC
17701 ACAACTGATGTCAGTAGGACGTTACCACATCTTGTGATCTTGAATAAGAAGGATGGATCG

17761 TGTAATTTTGTCTCCTTTAACAAATCTCTCCCTATCTTCCACTCCCCCGACCTTTCCAGC
17761 ACATTAAAACAGAGGAAATTGTTTAGAGAGGGATAGAAGGTGAGGGGGCTGGAAAGGTCG

17821 CTCTATTAGCCTCTGTCCTACTTTCTACTTATAATGATGACAGCAGCATTTGTTAGTTTC
17821 GAGATAATCGGAGACAGGATGAAAGATGAATATTACTACTGTCGTCGTAAACAATCAAAG

17881 CACATGTGAGTGAGAACATGTGGCTTTTTAACTTTTAGAATGTGGTATTCAGGCACTTCA
17881 GTGTACACTCACTCTTGTACACCGAAAAATTGAAAATCTTACACCATAAGTCCGTGAAGT

17941 TGGTACAGTTGGTAAAAGTGAAAATGTGTCCAAAAGTTTGTGATTATCTATATAAACAAA
17941 ACCATGTCAACCATTTTCACTTTTACACAGGTTTTCAAACACTAATAGATATATTTGTTT

18001 AATGGTATAAATACAAATATCAATTTTGCATTGAAGAACTTACCTTAGAGGTATATTCTC
18001 TTACCATATTTATGTTTATAGTTAAAACGTAACTTCTTGAATGGAATCTCCATATAAGAG

18061 ACAAGTGCACAGAGCATTTAAGCATTTGTTCACTGCAGCATTGTTATCAGTATTTTAAAA
18061 TGTTCACGTGTCTCGTAAATTCGTAAACAAGTGACGTCGTAACAATAGTCATAAAATTTT

18121 CTATGGTACATCCATGTACTTCCACATACAGCTCTTAAAAATAAGGAGGATATGAATGAA
18121 GATACCATGTAGGTACATGAAGGTGTATGTCGAGAATTTTTATTCCTCCTATACTTACTT

18181 CTAGTATGAAAAGAAGTCCAAATACATGTGAAAGTGAGAATAGCATGGTTCTGGATGGTA
18181 GATCATACTTTTCTTCAGGTTTATGTACACTTTCACTCTTATCGTACCAAGACCTACCAT

18241 TGCAAAGTATGATCTCGTTCTTTTAAAAGAAAATAAATTACATACACATACATATTTTCT
18241 ACGTTTCATACTAGAGCAAGAAAATTTTCTTTTATTTAATGTATGTGTATGTATAAAAGA

18301 ATATGCTTGCCCATAACGTTTAGGAAAATTCTTGGGTGATATTTATTAACCTGGACTTCC
18301 TATACGAACGGGTATTGCAAATCCTTTTAAGAACCCACTATAAATAATTGGACCTGAAGG

18361 TCTTGGAAGACTGATGGTAGAAGGAAGGGGACGAGTTAGGGAAGAGGAGGAGAAGGAAAA
18361 AGAACCTTCTGACTACCATCTTCCTTCCCCTGCTCAATCCCTTCTCCTCCTCTTCCTTTT

18421 CTTTGCTTTTCATCTTCTACCTTTTAGCATTATTTGAATTTATTTTCCTTAAGCGTTTAC
18421 GAAACGAAAAGTAGAAGATGGAAAATCGTAATAAACTTAAATAAAAGGAATTCGCAAATG

18481 TTTGTTTCGTAAACAAAAAAGCACAAAAACAAAAAACGAGTTAAATGGGAAAAAAAGCAG
18481 AAACAAAGCATTTGTTTTTTCGTGTTTTTGTTTTTTGCTCAATTTACCCTTTTTTTCGTC

18541 TTTAGCTCTTTATAGCCTCTCATTTGGCTTCGCCAGCCTCTCACTGCAGCCTCAGAGAGC
18541 AAATCGAGAAATATCGGAGAGTAAACCGAAGCGGTCGGAGAGTGACGTCGGAGTCTCTCG

18601 TGGTCTGGGAAACACTGGTAGATGAGGACTGTAATCCTCACTCATGGAAGAGGATCTCAT
18601 ACCAGACCCTTTGTGACCATCTACTCCTGACATTAGGAGTGAGTACCTTCTCCTAGAGTA

18661 TCACTGGGTTTGCTGACTGTGACTAGAAGTGATTAGGGTGTCAAAAAACCCAAGCATGTT
18661 AGTGACCCAAACGACTGACACTGATCTTCACTAATCCCACAGTTTTTTGGGTTCGTACAA

18721 AAAAATTTCCAGAGGCCAAAAAGATGCTTTCATTGTTCTGCTCTTCTTTTCCTTGTCGCT
18721 TTTTTAAAGGTCTCCGGTTTTTCTACGAAAGTAACAAGACGAGAAGAAAAGGAACAGCGA

18781 TTCACTTTGGGTAGCTTCTAAATTGGTATTTTGCATGGTGCATTTAAAGAAAATGAGACC
18781 AAGTGAAACCCATCGAAGATTTAACCATAAAACGTACCACGTAAATTTCTTTTACTCTGG

18841 CCTTTGGCCAATGCAGGAGTCTACACTCTGATATTCTAGAGTCAAAGCTGAATGCTGACA
18841 GGAAACCGGTTACGTCCTCAGATGTGAGACTATAAGATCTCAGTTTCGACTTACGACTGT

18901 CCTAGGAATTCATCTCTAGAATGTTTATATAAGGAATAGCCCCTCAGTATTCCGATCTCG
18901 GGATCCTTAAGTAGAGATCTTACAAATATATTCCTTATCGGGGAGTCATAAGGCTAGAGC

18961 TATCTTAGTAACGAAACTAACAAAAGCCTGATTCTCCTCTGGTAGTTTTCTTGTCTTTAC
18961 ATAGAATCATTGCTTTGATTGTTTTCGGACTAAGAGGAGACCATCAAAAGAACAGAAATG

19021 CATAATACAAAATAAGTAATTTGTTCTGCACCCTGACTGTTCAAAGGATAGGGTAGCTGG
19021 GTATTATGTTTTATTCATTAAACAAGACGTGGGACTGACAAGTTTCCTATCCCATCGACC

19081 GGGCGGGGACAAGAATGGAGACCTTATTACATAAGACTTCCTGAAAAAGGAAACTCTGTT
19081 CCCGCCCCTGTTCTTACCTCTGGAATAATGTATTCTGAAGGACTTTTTCCTTTGAGACAA

19141 TTTGTTTGAAATGATTTGGTCTGAAATTTAGTTTGTGTACACTTACCAAAGGGATTCCTA
19141 AAACAAACTTTACTAAACCAGACTTTAAATCAAACACATGTGAATGGTTTCCCTAAGGAT

19201 TTTCTAAAACACTCATACTGCTTTTGATTCCTGTTAACCTTTGAGCACTCTACGTAATGA
19201 AAAGATTTTGTGAGTATGACGAAAACTAAGGACAATTGGAAACTCGTGAGATGCATTACT

19261 TGAGAGCACTTAAAGAGTCATGTCACTTTTAGTAAAGAATCAAAGGATACTTTTTCTACT
19261 ACTCTCGTGAATTTCTCAGTACAGTGAAAATCATTTCTTAGTTTCCTATGAAAAAGATGA

19321 TCTTCGAGTTTGATCTCTGCTTCTCCAGTTAAAACCAGTATTTGTTTTTTTCATTTCTAA
19321 AGAAGCTCAAACTAGAGACGAAGAGGTCAATTTTGGTCATAAACAAAAAAAGTAAAGATT

19381 AGTTGGAAGAAATGACAGTTAGTTATGGCATAAGGATGTACATTTAACCAAATAGGAGTT
19381 TCAACCTTCTTTACTGTCAATCAATACCGTATTCCTACATGTAAATTGGTTTATCCTCAA

19441 GACATTCTTGGTAAGAAATCTTACCAAGATTATGTTATAGATTATAAGAAATCTTAACAA
19441 CTGTAAGAACCATTCTTTAGAATGGTTCTAATACAATATCTAATATTCTTTAGAATTGTT

19501 GAATATGTTCCTAAATCATCCTCTTTTCCCATAAAATATTAAAGTATCAGCAATTTCATA
19501 CTTATACAAGGATTTAGTAGGAGAAAAGGGTATTTTATAATTTCATAGTCGTTAAAGTAT

19561 GGATTCAACCTAATGTATGCGAAATGCTAGATAAACAGATAAATACTTAATATCTGGCTT
19561 CCTAAGTTGGATTACATACGCTTTACGATCTATTTGTCTATTTATGAATTATAGACCGAA

19621 TTTTTCAAAGCACTGGGTTATTTGTTCCTTGAGATTTATCCTAAATGTGGGCTATACCCT
19621 AAAAAGTTTCGTGACCCAATAAACAAGGAACTCTAAATAGGATTTACACCCGATATGGGA

19681 GGTTTACAGTGTCTCACAGATGTGTAGTAGTAGACACTCCATAAGTGTTTACTGACTTGA
19681 CCAAATGTCACAGAGTGTCTACACATCATCATCTGTGAGGTATTCACAAATGACTGAACT

19741 ATCCACAGGGTACTGAGAAAATGCTACTGATAGACTTGGAGGAGAGCATATCTAAAGCAA
19741 TAGGTGTCCCATGACTCTTTTACGATGACTATCTGAACCTCCTCTCGTATAGATTTCGTT

19801 GCTACCCTTTCCTTTAGGGCACGTCTCACTAATTCTTTGGGTAAAGCGTATTTTTCTTCC
19801 CGATGGGAAAGGAAATCCCGTGCAGAGTGATTAAGAAACCCATTTCGCATAAAAAGAAGG

19861 TTTTGTGTTTTTGGCAGTCTTTCCAAAAATACGTGTTATACCTATGCATTATTTTTTGGT
19861 AAAACACAAAAACCGTCAGAAAGGTTTTTATGCACAATATGGATACGTAATAAAAAACCA

19921 TTGGTTTCTAAAGAAAGAGTCAGCCGGTGGGAAAGTGAAGGATGTGGGAACTGAGAGATC
19921 AACCAAAGATTTCTTTCTCAGTCGGCCACCCTTTCACTTCCTACACCCTTGACTCTCTAG

19981 TGCATCAGCATCCCACCTCTACCTCCCACGATGGGACCTGAGACAGTTATTTTTGCCTCC
19981 ACGTAGTCGTAGGGTGGAGATGGAGGGTGCTACCCTGGACTCTGTCAATAAAAACGGAGG

20041 TGGACCACTATAGTATCATCTGTAACAGGAGGGACTTGAGCCAGTTGATCTCTAAGGTTC
20041 ACCTGGTGATATCATAGTAGACATTGTCCTCCCTGAACTCGGTCAACTAGAGATTCCAAG

20101 CTCTGGCACCTGTGACCCTAAATAGATATTGGATATTGGTTTAATGCTATTTGTAGTGTG
20101 GAGACCGTGGACACTGGGATTTATCTATAACCTATAACCAAATTACGATAAACATCACAC

20161 TTTTTTTGGGGATATGGAAACCAGAAGTTTGTTTCCATAAACATAAACATAAACTGTATA

20161 AAAAAAACCCCTATACCTTTGGTCTTCAAACAAAGGTATTTGTATTTGTATTTGACATAT

20221 TATCTAAAGGATATGGAAACCTTTAGATATATATAATCTGCTTACGTAAAGAAGGTTTGT
20221 ATAGATTTCCTATACCTTTGGAAATCTATATATATTAGACGAATGCATTTCTTCCAAACA

20281 ATATATTGCAGTGTCAATGGGAATATTTTATCAAGTTAAGCATAGTAAATCACATTGATT
20281 TATATAACGTCACAGTTACCCTTATAAAATAGTTCAATTCGTATCATTTAGTGTAACTAA

20341 AAATGCTTTGTATTTACCAAACATTACCCAAAGTGTTTTCTCCTTTCAACCTCACAAGGA
20341 TTTACGAAACATAAATGGTTTGTAATGGGTTTCACAAAAGAGGAAAGTTGGAGTGTTCCT

20401 CCCACAGAAGAAAATACAGTTATCATTTCCAACCTGCAGGGAGCTGAGACACAGAGAATT
20401 GGGTGTCTTCTTTTATGTCAATAGTAAAGGTTGGACGTCCCTCGACTCTGTGTCTCTTAA

20461 TAAGCAACTGACCGGAAGTCCAACAGGGAGTCAGAGATTGCTCTGGGGTGTGATCCCCAC
20461 ATTCGTTGACTGGCCTTCAGGTTGTCCCTCAGTCTCTAACGAGACCCCACACTAGGGGTG

20521 TTGGACCCTAGAGTGGAAGCTTCTCCACTACTTTATAGAGTTGAGATTCTATATTTTGAG
20521 AACCTGGGATCTCACCTTCGAAGAGGTGATGAAATATCTCAACTCTAAGATATAAAACTC

20581 CTTGTATTTACCCAGAGAATTATATCCTCTTGGGCAATTGTGTATAATAAAACCTCATGC
20581 GAACATAAATGGGTCTCTTAATATAGGAGAACCCGTTAACACATATTATTTTGGAGTACG

20641 ATTTAGGAGAGGCGGGATGACAGAACTTTGTTGAGTGAATTATAATCTACTTGAGAAATT
20641 TAAATCCTCTCCGCCCTACTGTCTTGAAACAACTCACTTAATATTAGATGAACTCTTTAA

20701 ATTTGCTTACATTTTATAAGCTAATTATACCATATCTCATCCAGTTTTCCCAGAACACTT
20701 TAAACGAATGTAAAATATTCGATTAATATGGTATAGAGTAGGTCAAAAGGGTCTTGTGAA

20761 CTCATAGGTAATGCTTTATTTGAAACATAGGCCATAGGTAAGTTAAGTGTAAATGTGTAT
20761 GAGTATCCATTACGAAATAAACTTTGTATCCGGTATCCATTCAATTCACATTTACACATA

20821 TTTTATAATTTAACCAGAAGTTTATTTCATTTTTCTAAATAAGTGAAATTGTATTGCATC
20821 AAAATATTAAATTGGTCTTCAAATAAAGTAAAAAGATTTATTCACTTTAACATAACGTAG

20881 TTCTAAATTATTCTATTTAAACACTTGATGTCTTGCTGTCTCCGTCTCTGTGTGTTTGCA
20881 AAGATTTAATAAGATAAATTTGTGAACTACAGAACGACAGAGGCAGAGACACACAAACGT

20941 TGTCATTGTACATGTTCTTAGGAAAAGTGTGGGAGCTTGACGCAATATATACCTTATGTT
20941 ACAGTAACATGTACAAGAATCCTTTTCACACCCTCGAACTGCGTTATATATGGAATACAA

21001 TCTATGTGCATATAGTTTACCAAATAATACCATAAGTTTACTTAGCATATTAGAATCCAT
21001 AGATACACGTATATCAAATGGTTTATTATGGTATTCAAATGAATCGTATAATCTTAGGTA

21061 GCACATTATTTTTATTTTATCTTCACCGCAACCCTGTGGGATAGACCAAAATCATGCTTT
21061 CGTGTAATAAAAATAAAATAGAAGTGGCGTTGGGACACCCTATCTGGTTTTAGTACGAAA

21121 TCAGCCTCCTTTTTCCACTTGAGGAAAGGAGTCTTAAAAAAGGGACCAGTCTCATGTTCC
21121 AGTCGGAGGAAAAAGGTGAACTCCTTTCCTCAGAATTTTTTCCCTGGTCAGAGTACAAGG

21181 CATTCGTCTTACAACTAATTGGTCAAGCCAGAAAGCCAGAACTATGTCCTGGGTCACTAA
21181 GTAAGCAGAATGTTGATTAACCAGTTCGGTCTTTCGGTCTTGATACAGGACCCAGTGATT

21241 CTCCTAGTCACTGTGTGTTAGTATTTGAGATGCCTGTTGGCTTGATTTAGTCATTTATTT
21241 GAGGATCAGTGACACACAATCATAAACTCTACGGACAACCGAACTAAATCAGTAAATAAA

21301 TTTAGTGTTTTATAATCCTTGCATACTTTTACATTTTAAATGGTTAACCAGGCAAATTGG
21301 AAATCACAAAATATTAGGAACGTATGAAAATGTAAAATTTACCAATTGGTCCGTTTAACC

21361 TTTAAAATCAGTGCATAAAAATACTGTGCCTATCATGATGGGTTTCATGAAGTGATAACT
21361 AAATTTTAGTCACGTATTTTTATGACACGGATAGTACTACCCAAAGTACTTCACTATTGA

21421 TTTCATCATGGAGATCCTCAGCTGTCACAGAAGATGAGGGGCCCTGGGTACAGAGGCTCA
21421 AAAGTAGTACCTCTAGGAGTCGACAGTGTCTTCTACTCCCCGGGACCCATGTCTCCGAGT

21481 CGTGAGGGATGAAAGTCTCAGCAGCCCGGACTTACACTTTGGGGCTTTTAGGCAAATCAG
21481 GCACTCCCTACTTTCAGAGTCGTCGGGCCTGAATGTGAAACCCCGAAAATCCGTTTAGTC

21541 ACAACCTCTTAAGAACTATCACTGAGTTCAGGCAAGGCGAGCTTGAATTAACACAGGGCC
21541 TGTTGGAGAATTCTTGATAGTGACTCAAGTCCGTTCCGCTCGAACTTAATTGTGTCCCGG

21601 CTTGGTGGGCATGTGAATATATCTCACTTCACTACCATCCAGTTCTGACTCTTTACTAGA
21601 GAACCACCCGTACACTTATATAGAGTGAAGTGATGGTAGGTCAAGACTGAGAAATGATCT

```
21661  TGCCCCTGTACATACCAAGACTGATTTTTTATTCTCCCTTCTCCCCATGTGGTTTCTTCT
21661  ACGGGGACATGTATGGTTCTGACTAAAAAATAAGAGGGAAGAGGGGTACACCAAAGAAGA

21721  GCATAGAGAGTTCCTATTGATCAGTCTGACCCATGGTATTTTAGAATTGCGATCCCTACT
21721  CGTATCTCTCAAGGATAACTAGTCAGACTGGGTACCATAAAATCTTAACGCTAGGGATGA

21781  GTTTCATTATTCCTTTTTCTCCCCCATGTTGAAAAAAATAAATGTCCTGAGATGCAAGAT
21781  CAAAGTAATAAGGAAAAAGAGGGGGTACAACTTTTTTTATTTACAGGACTCTACGTTCTA

21841  CAGGGACACTGGAGCACTGACATTTAGTTCAGTGCAGGAACTGAAGGCAGATGTAATTCT
21841  GTCCCTGTGACCTCGTGACTGTAAATCAAGTCACGTCCTTGACTTCCGTCTACATTAAGA

21901  TAAGAAGCGTACCTGTTATTATGAACCATCCTCAACAAATTGTAGTGGATCTTGTTTTCT
21901  ATTCTTCGCATGGACAATAATACTTGGTAGGAGTTGTTTAACATCACCTAGAACAAAAGA

21961  CATAGATACAGCAGTTAAATTTTTTAATAAAAGTAACTAAGAGTTATTTGGATGTATTTT
21961  GTATCTATGTCGTCAATTTAAAAAATTATTTTCATTGATTCTCAATAAACCTACATAAAA

22021  AGCATGCACTGAGCGGAAAGTACGACATTTCTTCATTGGGTAAGTCCTGATTCTTTATGA
22021  TCGTACGTGACTCGCCTTTCATGCTGTAAAGAAGTAACCCATTCAGGACTAAGAAATACT

22081  TCCTCACTTGGTTCCAGGGCCCCATGCATCTAAGGGTGTCTCAGAGCATCCTGCAGTGCT
22081  AGGAGTGAACCAAGGTCCCGGGGTACGTAGATTCCCACAGAGTCTCGTAGGACGTCACGA

22141  CCAGCATGATCGCAGGGAAAAGCTATAGGAGGAAAAGAGTCAATAAAGTTTAGTTTCTCA
22141  GGTCGTACTAGCGTCCCTTTTCGATATCCTCCTTTTCTCAGTTATTTCAAATCAAAGAGT

22201  ACCTCCCACCTCCACCCCATAATAATGACAGCTGGTTAATCATGAGACGCGTGCACACCC
22201  TGGAGGGTGGAGGTGGGGTATTATTACTGTCGACCAATTAGTACTCTGCGCACGTGTGGG

22261  CACACGCCCTGTACATGTTTACTCATTGGGATAGCATGTCAGGCCAGAAGGCTCCATGGT
22261  GTGTGCGGGACATGTACAAATGAGTAACCCTATCGTACAGTCCGGTCTTCCGAGGTACCA

22321  CATTTCTATGAAGGTACTTTAGCAGGTCTTCAAGAAGGCAAGTGGCCTGGGTCCCTGCCT
22321  GTAAAGATACTTCCATGAAATCGTCCAGAAGTTCTTCCGTTCACCGGACCCAGGGACGGA

22381  CCCCAAATTGCAAGCTCCCTGCTTTATGTAGGAGACCTATGTGTATATTACAGTTCTGTG
22381  GGGGTTTAACGTTCGAGGGACGAAATACATCCTCTGGATACACATATAATGTCAAGACAC

22441  TAAGATTATTTTGTTATTCTTACCCCCACACCCACCCCCCAACCCCCCGCTGCCACCAAA
22441  ATTCTAATAAAACAATAAGAATGGGGGTGTGGGTGGGGGGGTTGGGGGGCGACGGTGGTTT

22501  AAAAAAAAAAAAAAAATTCCTCTGACAACCTTCATAAAGTCCTGGGAGTTTGAACACCATT
22501  TTTTTTTTTTTTTTTAAGGAGACTGTTGGAAGTATTTCAGGACCCTCAAACTTGTGGTAA

22561  GCTCTAGGAAGTCATCTTATACAAAAATAAGAGTTGTGAGGTGGTTCATATACCTCCTGC
22561  CGAGATCCTTCAGTAGAATATGTTTTTATTCTCAACACTCCACCAAGTATATGGAGGACG

22621  GTTCTCCTATTTGGAGTTTTTCCCCATTTATGAAAGAGGTGAAAACGCTAAGATATTTAG
22621  CAAGAGGATAAACCTCAAAAAGGGGTAAATACTTTCTCCACTTTTGCGATTCTATAAATC

22681  CAATTATTACTTTAAACATTTTCTATTTATAGGCCGGGCGCAGTGGCTCATGCCTGTAAT
22681  GTTAATAATGAAATTTGTAAAAGATAAATATCCGGCCCGCGTCACCGAGTACGGACATTA

22741  CCCAGCGCTTGGGAGGCCAAGGCAGGCAGATCACGAGGTCAGGAGATCGAGACCATCTTG
22741  GGGTCGCGAACCCTCCGGTTCCGTCCGTCTAGTGCTCCAGTCCTCTAGCTCTGGTAGAAC

22801  GCTAACACGGTGAAACCCTGTCTCTACTAAAAAAATACAAAAATTTAGTTGGGCATGGTG
22801  CGATTGTGCCACTTTGGGACAGAGATGATTTTTTTATGTTTTTAAATCAACCCGTACCAC

22861  GGGGATACCTGTGGTCCCAGCTACTCGGGAGGCTGAGACAGGAGAATGGCTTGAACCTGG
22861  CCCCTATGGACACCAGGGTCGATGAGCCCTCCGACTCTGTCCTCTTACCGAACTTGGACC

22921  GAGGCCGAGCTTGCAGTGAGCCAAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGC
22921  CTCCGGCTCGAACGTCACTCGGTTCTAGCGCGGTGACGTGAGGTCGGACCCACTGTCTCG

22981  AAGACTCTGTCTCGAAAAAAAAAAAAAAATTCTATTTACAGCAGTGAAAATAGTAGTGACT
22981  TTCTGAGACAGAGCTTTTTTTTTTTTTAAGATAAATGTCGTCACTTTTATCATCACTGA

23041  TAATGCACATTGCCAAGGCTTTAGCATAACATGAACACTTTCACTCAATGTCTCTCTGGC
23041  ATTACGTGTAACGGTTCCGAAATCGTATTGTACTTGTGAAAGTGAGTTACAGAGAGACCG
```

```
23101  CTTTTGTTTTTCCTTGGGAAATTCTTATAATCCTGCTCCGTCTTTAACTATTCATTTTGT
23101  GAAAACAAAAAGGAACCCTTTAAGAATATTAGGACGAGGCAGAAATTGATAAGTAAAACA

23161  ATTGGCTATCCAAATATACCCAATAATGCTCTTTCTGAAAATATGCCAATTGTGGTAATT
23161  TAACCGATAGGTTTATATGGGTTATTACGAGAAAGACTTTTATACGGTTAACACCATTAA

23221  ACAGCTAAGCTGGAATATTAAATTGTGATGTCTGTTTTCCAGAGAATGAAGTAGTATTCC
23221  TGTCGATTCGACCTTATAATTTAACACTACAGACAAAAGGTCTCTTACTTCATCATAAGG

23281  CCAGAGCATAGGCTTGGTGCCTGTGCAGGTTCTATTTTAAATATTCCAGGAAGGGTTGTT
23281  GGTCTCGTATCCGAACCACGGACACGTCCAAGATAAAATTTATAAGGTCCTTCCCAACAA

23341  TTATATACTGAGGATGATTTTACTGGTCTTGCCAGTCGTCTGAAATGCTGGTATTACTCT
23341  AATATATGACTCCTACTAAAATGACCAGAACGGTCAGCAGACTTTACGACCATAATGAGA

23401  TGTGGAAGGTTTATTCAAACAAACAAGGACATTTCACACAATACCTAGTCATGTTTTTCA
23401  ACACCTTCCAAATAAGTTTGTTTGTTCCTGTAAAGTGTGTTATGGATCAGTACAAAAAGT

23461  GACATTTTAATGTTTGGTTCATCATTTGCACACACTCTCAAAAATCTAGGTTTGTCTATG
23461  CTGTAAAATTACAAACCAAGTAGTAAACGTGTGTGAGAGTTTTTAGATCCAAACAGATAC

23521  TGTTCATATCATTTTGCCTGTTGCCAGCTCAGTCAGCAGGCACACTCTCCCAGGCTGTTG
23521  ACAAGTATAGTAAAACGGACAACGGTCGAGTCAGTCGTCCGTGTGAGAGGGTCCGACAAC

23581  CTGTTTTGTTAGACTTCTTCAGGACCTTCATCTAAAATGGTCTTCCACACGTAGCTATAC
23581  GACAAAACAATCTGAAGAAGTCCTGGAAGTAGATTTTACCAGAAGGTGTGCATCGATATG

23641  TGCATAAGTTCACATCATCTGTTTCTTGCATGTGGGTTGTGTCTCAACTCAAGTTTAAGT
23641  ACGTATTCAAGTGTAGTAGACAAAGAACGTACACCCAACACAGAGTTGAGTTCAAATTCA

23701  TAGATTTGGAAGGGCGGAAACTATAGGAGTTGCAGCTTCAGTGGAGAAAAGAGCATTTCC
23701  ATCTAAACCTTCCCGCCTTTGATATCCTCAACGTCGAAGTCACCTCTTTTCTCGTAAAGG

23761  TACTAGTTATGGCTTCCCAAGGAAGGTTAGATTCCTCAGAGTAGGAGTGATTCCCCAATG
23761  ATGATCAATACCGAAGGGTTCCTTCCAATCTAAGGAGTCTCATCCTCACTAAGGGGTTAC

23821  CTAGAACCTTTGGTCAAATATAATTCTAATCCAGTCAAAATAAATACAGGTATTCTGTAA
23821  GATCTTGGAAACCAGTTTATATTAAGATTAGGTCAGTTTTATTTATGTCCATAAGACATT

23881  AACCCGATTTCATTTTGTAAATCCTACTTTGTATAGTATAAGCAATTTTTGTATTTGTGT
23881  TTGGGCTAAAGTAAAACATTTAGGATGAAACATATCATATTCGTTAAAAACATAAACACA

23941  GGATTATATTTTATTTTCCTATTTCAAAGAGAAGAATTTGTATTAGCAGACTCCCTTTGC
23941  CCTAATATAAAATAAAAGGATAAAGTTTCTCTTCTTAAACATAATCGTCTGAGGGAAACG

24001  ATGCGGAGAGGGGATCATTTTCCCAGTAGGCATGGGGTTCCCTTCCATTCCTTGTCCAGT
24001  TACGCCTCTCCCCTAGTAAAAGGGTCATCCGTACCCCAAGGGAAGGTAAGGAACAGGTCA

24061  CTTCTTTTCCCCACTAAGTTAAGTCAAACTAAGCAGCTGGTAAGATATTCCCTGGTTCTT
24061  GAAGAAAAGGGGTGATTCAATTCAGTTTGATTCGTCGACCATTCTATAAGGGACCAAGAA

24121  GCAAAGAAAGTGAGCAGATGGCAGAATGTATAGCTCTAAGCAGAATACCTGGTGTGGTAT
24121  CGTTTCTTTCACTCGTCTACCGTCTTACATATCGAGATTCGTCTTATGGACCACACCATA

24181  CCTCAAACACAAATTGACAGGAGGGTGTGGTGTGGCAAGCTCATTGTGGGGGTAAATTGG
24181  GGAGTTTGTGTTTAACTGTCCTCCCACACCACACCGTTCGAGTAACACCCCCATTTAACC

24241  AATAAGCTTACAGGGGGAAGAGTTGACAAAAGATAGGAAGAACCTTAAAAATATAGATGC
24241  TTATTCGAATGTCCCCCTTCTCAACTGTTTTCTATCCTTCTTGGAATTTTTATATCTACG

24301  CTTTTATGCAGTGATAAAATGTCTAGATATTTATACTGTGGTGATTATTAGGAATATGTG
24301  GAAAATACGTCACTATTTTACAGATCTATAAATATGACACCACTAATAATCCTTATACAC

24361  CAAAGATTGGCTATTAGGATGTTCATTACAGTGTTGTTTAATAATTATAAAAGGACAGAA
24361  GTTTCTAACCGATAATCCTACAAGTAATGTCACAACAAATTATTAATATTTTCCTGTCTT

24421  AGCAATGTGGACTCAAAAATAGGAAAAGAATTTAAATAAATCCTAGTGTACCCGTTATAC
24421  TCGTTACACCTGAGTTTTTATCCTTTTCTTAAATTTATTTAGGATCACATGGGCAATATG

24481  ATGAAATTATGGAAATATGACCCTGAGCATGGAAATATGTACATGAGAATGTCTAAAAGC
24481  TACTTTAATACCTTTATACTGGGACTCGTACCTTTATACATGTACTCTTACAGATTTTCG

24541  TAGTTCATTTTGAAAAACAAAATAATGTCACCTCATATTATTTATAGTATATAAAGATGA
```

```
24541  ATCAAGTAAAACTTTTTGTTTTATTACAGTGGAGTATAATAAATATCATATATTTCTACT

24601  TTTTAAGAGTGGCAGTGTCTGGGATTATAGGTGATTGTATTTCTTCCCTTTTGCACATCT
24601  AAAATTCTCACCGTCACAGACCCTAATATCCACTAACATAAAGAAGGGAAAACGTGTAGA

24661  ATGTTCTCTCATTTGTATTGTGTGGGGAGAAGTGACTTTTTTTATAAAAAGAAAAAGGTA
24661  TACAAGAGAGTAAACATAACACACCCCTCTTCACTGAAAAAAATATTTTTCTTTTTCCAT

24721  TATGCATCCCAGCAGAGAAGCACTGGCTCCACCCAGTACCTGCCTCCTCATGCCACCCTC
24721  ATACGTAGGGTCGTCTCTTCGTGACCGAGGTGGGTCATGGACGGAGGAGTACGGTGGGAG

24781  TCAAGCCAAAAGCCGGGGGAAGCCCAGGCACCTTGACCATGACCGCCCGAGACTCACACT
24781  AGTTCGGTTTTCGGCCCCCTTCGGGTCCGTGGAACTGGTACTGGCGGGCTCTGAGTGTGA

24841  TCTTCTTTCTCATCAGGGAAGGAAAGCCCCCCTGTAAGCGATACTCCAGATGAGGGCGAT
24841  AGAAGAAAGAGTAGTCCCTTCCTTTCGGGGGGACATTCGCTATGAGGTCTACTCCCGCTA

24901  GAGCCCATGCCGATCCCCGAGGACCTCTCCACCACCTCGGGAGGACAGCAAAGCTCCAAG
24901  CTCGGGTACGGCTAGGGGCTCCTGGAGAGGTGGTGGAGCCCTCCTGTCGTTTCGAGGTTC

24961  AGTGACAGAGTCGTGGGTAAGTGGGTCACCAGCGGCCTCTGTGCCTGTGAAACCTTTATC
24961  TCACTGTCTCAGCACCCATTCACCCAGTGGTCGCCGGAGACACGGACACTTTGGAAATAG

25021  TCTTTGTATTTTTCCAAGACAGTGATGAAGGGATGCAAGTCATTTTATCCATTGTGTTCC
25021  AGAAACATAAAAAGGTTCTGTCACTACTTCCCTACGTTCAGTAAAATAGGTAACACAAGG

25081  CTCAACTGGCATATTAAAGAAATATGGCACAAAGATCAGCAGGATGGGGGTCCTCTGGTG
25081  GAGTTGACCGTATAATTTCTTTATACCGTGTTTCTAGTCGTCCTACCCCCAGGAGACCAC

25141  TGTGGGAGGATGGACACTCACAGGCCAGCATGGCCGTGAGAGCCACACACCCCGCAAAAT
25141  ACACCCTCCTACCTGTGAGTGTCCGGTCGTACCGGCACTCTCGGTGTGTGGGGCGTTTTA

25201  GTCCAAGTTGAGGAGCAATCCTGCCCAGGGACGCGTCTCTGTCACTGTCCTCTGTCCTCA
25201  CAGGTTCAACTCCTCGTTAGGACGGGTCCCTGCGCAGAGACAGTGACAGGAGACAGGAGT

25261  CTGCACTTGCAGGAATATCAAATGTTATGGATTGTAGATCGTGAAAATTACACACTTACG
25261  GACGTGAACGTCCTTATAGTTTACAATACCTAACATCTAGCACTTTTAATGTGTGAATGC

25321  TGTTTTGGCAACAGTGCTTTTCAGTGTTTGGGGTTAGAACAAGCCACATCTGGCCATTTT
25321  ACAAAACCGTTGTCACGAAAAGTCACAAACCCCAATCTTGTTCGGTGTAGACCGGTAAAA

25381  ATGTTATCCCTCTAATCTCTAGTTCTTAAATTTCAGATTTAACTGAAAATAGAAAGTTTC
25381  TACAATAGGGAGATTAGAGATCAAGAATTTAAAGTCTAAATTGACTTTTATCTTTCAAAG

25441  ATAAATGGTAATTTTTAAATGTTAGAAAAAAATCAAAGCACACATTTACATACCTTTTCC
25441  TATTTACCATTAAAAATTTACAATCTTTTTTTAGTTTCGTGTGTAAATGTATGGAAAAGG

25501  ATAAAAAAGTGTGTGTGAAACCATCTGCAACTCCACTGGTTAACTTACACATGCGACAGT
25501  TATTTTTTCACACACACTTTGGTAGACGTTGAGGTGACCAATTGAATGTGTACGCTGTCA

25561  TTGTGTTTGGAAAACTTCACCCACCCTCTTTATTTCAAGAAACCAAAGACGGAGAGAAGT
25561  AACACAAACCTTTTGAAGTGGGTGGGAGAAATAAAGTTCTTTGGTTTCTGCCTCTCTTCA

25621  TCAGTGAATTAACCAAGGTGAAATGACAGAGTCAGAATTGAACTGAATTTCTGACTGAAA
25621  AGTCACTTAATTGGTTCCACTTTACTGTCTCAGTCTTAACTTGACTTAAAGACTGACTTT

25681  AACTAGTTCTCGCTCCATTATACCATGTTCTAATGAGCTAATGAGTCAACAGTTCCCTGG
25681  TTGATCAAGAGCGAGGTAATATGGTACAAGATTACTCGATTACTCAGTTGTCAAGGGACC

25741  AATACCTGTTTTCTTCTTTAAAAAGAGAAAGAGCTGTGATACTGAGAGCTACCTATTGGC
25741  TTATGGACAAAAGAAGAAATTTTTCTCTTTCTCGACACTATGACTCTCGATGGATAACCG

25801  ATAAAAGAATATGAAGGACATACAGGTAAAATGAGGTAGACTGGATCAATGTGGGTCAGC
25801  TATTTTCTTATACTTCCTGTATGTCCATTTTACTCCATCTGACCTAGTTACACCCAGTCG

25861  TACATAAGCCCCCATAGCCAAGGAGGGAACATGTATGTAAAATTCTCAGGCTTATGAGAA
25861  ATGTATTCGGGGGTATCGGTTCCTCCCTTGTACATACATTTTAAGAGTCCGAATACTCTT

25921  TCACAACTGTATGTCTATAGGACATCCTAATGCAATAATGGGAAAATGTGATAACTGAAT
25921  AGTGTTGACATACAGATATCCTGTAGGATTACGTTATTACCCTTTTACACTATTGACTTA

25981  TTTATCTACATGTATAGAAAACATATTTTGGGCCGGGCATGGTGGCTCAAGCTTGTAATC
25981  AAATAGATGTACATATCTTTTGTATAAAACCCGGCCCGTACCACCGAGTTCGAACATTAG
```

```
26041  CCAGCACTTTGGGAGGCCGAGGCTGGCGGATCACAAGGTCAGGAAATCGAGACCATCCTG
26041  GGTCGTGAAACCCTCCGGCTCCGACCGCCTAGTGTTCCAGTCCTTTAGCTCTGGTAGGAC

26101  GCTAACACGGTGAAACCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGGCGTAGTGGT
26101  CGATTGTGCCACTTTGGGACAGAGATGATTTTTATGTTTTTTAATCGGCCCGCATCACCA

26161  GGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGCA
26161  CCCGCGGACATCAGGGTCGATGAGCCCTCCGACTCCGTCCTCTTACCGCACTTGGGCCGT

26221  GGCGGAGCTTGCAGTGAGCCGAGATTGCGCCACTGCACTCCAGCCTGGATGACAGAGCCA
26221  CCGCCTCGAACGTCACTCGGCTCTAACGCGGTGACGTGAGGTCGGACCTACTGTCTCGGT

26281  GACTCCATCTCAAAAAAAAAAAAAAAAAAGAAAAAGTAAAGAAAACATATTTTGTATGTT
26281  CTGAGGTAGAGTTTTTTTTTTTTTTTTTCTTTTTCATTTCTTTTGTATAAAACATACAA

26341  TTTTATGTGTCTTAGAAACCAAGTGGACTAGGTATAATGGTTCTTTTTTTTTTTAACTCA
26341  AAAATACACAGAATCTTTGGTTCACCTGATCCATATTACCAAGAAAAAAAAAAATTGAGT

26401  GCAGCACAGTTAATGGCACAAACAGATAAAGATGGTGGCTAAAGCCATTTGGTGTAGATG
26401  CGTCGTGTCAATTACCGTGTTTGTCTATTTCTACCACCGATTTCGGTAAACCACATCTAC

26461  AATCCAGAAAACAAGTACATTTATGAATTATTTTACCCTATTGGAAAATGGCTACTATAT
26461  TTAGGTCTTTTGTTCATGTAAATACTTAATAAAATGGGATAACCTTTTACCGATGATATA

26521  GTGGAAAGTAGAGGGGAAGAGAAATGAACTTGTGAATTATCTACTCTATTCTACCAAAGA
26521  CACCTTTCATCTCCCCTTCTCTTTACTTGAACACTTAATAGATGAGATAAGATGGTTTCT

26581  AAGATACATATATCATTTCAGGGTTTTCCAAAGTTGGGATGCAAAGGATCATGTTTCATT
26581  TTCTATGTATATAGTAAAGTCCCAAAAGGTTTCAACCCTACGTTTCCTAGTACAAAGTAA

26641  TTAATTTATGTATTTATTTTTGAGATTGTCTTCTATGTATTGCAAGTGATACTGGTTTTA
26641  AATTAAATACATAAATAAAAACTCTAACAGAAGATACATAACGTTCACTATGACCAAAAT

26701  CCATGCACTGTACCACTACGAGGGTTCCTTTCAAAACACATTCATATAAATTAAAAAATC
26701  GGTACGTGACATGGTGATGCTCCCAAGGAAAGTTTTGTGTAAGTATATTTAATTTTTAG

26761  AGTTTAAATAAAAACATCAAGTCAATAGTAGTACATAGATATGATAAAAGTTTGCAAAAG
26761  TCAAATTTATTTTTGTAGTTCAGTTATCATCATGTATCTATACTATTTTCAAACGTTTTC

26821  TGGCACAGGAATAGTCATTGGGTGGTGATGTTCCCTGATTTTTTTAGTCCCTGAAGCACC
26821  ACCGTGTCCTTATCAGTAACCCACCACTACAAGGGACTAAAAAAATCAGGGACTTCGTGG

26881  CAGCGCCAAGGGTACTGTTCTTCCTGCTTAACACAGGAGCAAACAGATGCTCACACCCTG
26881  GTCGCGGTTCCCATGACAAGAAGGACGAATTGTGTCCTCGTTTGTCTACGAGTGTGGGAC

26941  CCCAAGAACACAGAGGCTCAGGGACTGAACCAGAAGAAATGATCCATTAACCATGTAAAT
26941  GGGTTCTTGTGTCTCCGAGTCCCTGACTTGGTCTTCTTTACTAGGTAATTGGTACATTTA

27001  GAACAGAGAGCCTGGATGTCAGAGTGAGGAGGCCAGCTCGGGTCGCTGTGGAGAAAGAGA
27001  CTTGTCTCTCGGACCTACAGTCTCACTCCTCCGGTCGAGCCCAGCGACACCTCTTTCTCT

27061  TGGGCTCTCCCTGGAATCCCTGCAGCTGCCCTTAGAGGCTGTGAGCCATCGTGGCTAGGG
27061  ACCCGAGAGGGACCTTAGGGACGTCGACGGGAATCTCCGACACTCGGTAGCACCGATCCC

27121  CATCTTTAGAGAACTGAAAGTCAGGCCGATTCCTGGAGAGTGAAGAATAGCACACGTGCC
27121  GTAGAAATCTCTTGACTTTCAGTCCGGCTAAGGACCTCTCACTTCTTATCGTGTGCACGG

27181  TTCAAAACAGTGTCAGTATCCTGTCCCTTACGGACGTCATATTGTTGGTGGGTTTTCCTG
27181  AAGTTTTGTCACAGTCATAGGACAGGGAATGCCTGCAGTATAACAACCACCCAAAAGGAC

27241  GGTGGGAGATCACCCAGTATCTGGTTGGCATTACTGATGGGTTCCCAGAGCTGTCTTTTG
27241  CCACCCTCTAGTGGGTCATAGACCAACCGTAATGACTACCCAAGGGTCTCGACAGAAAAC

27301  GTTCAAAGTCCTGCCCTCGGTCCCAGCTGCCCCCAGCACAGGTGTCACCTTGGAGTCAGG
27301  CAAGTTTCAGGACGGGAGCCAGGGTCGACGGGGGTCGTGTCCACAGTGGAACCTCAGTCC

27361  GGCGGGTTTGTCATAGCCCATGGGTCCTGGCCCTGCCCTTTATCATGTTTGACATGATTT
27361  CCGCCCAAACAGTATCGGGTACCCAGGACCGGGACGGGAAATAGTACAAACTGTACTAAA

27421  TCAGTCTACTTTTCCTTCAGGATGTAATTTAGAAAATATGATTTCAGAATTGTGAAATTT
27421  AGTCAGATGAAAAGGAAGTCCTACATTAAATCTTTTATACTAAAGTCTTAACACTTTAAA
```

```
27481  AAAAACCCTAGAAAACCTCTTTGCTTTGCTTGTAAATAGCAGTTGATGCAGCCCTGTGGA
27481  TTTTTGGGATCTTTTGGAGAAACGAAACGAACATTTATCGTCAACTACGTCGGGACACCT

27541  CAGGCGGGAACAGCCATGTGACGTGGGCACTGAGATGATTTGGGAGATTGAATGCTCCTC
27541  GTCCGCCCTTGTCGGTACACTGCACCCGTGACTCTACTAAACCCTCTAACTTACGAGGAG

27601  AGTGGGGCTAGAGGCCACAGGAGTGATAGGTTACTGTAAAGAAAGCAAAAAAGATGCAAA
27601  TCACCCCGATCTCCGGTGTCCTCACTATCCAATGACATTTCTTTCGTTTTTTCTACGTTT

27661  AGCCTACATGACAAACCAGGCACGACACAAAGGAGAGGCTTGAGAAGAGGAAGTGAAGGA
27661  TCGGATGTACTGTTTGGTCCGTGCTGTGTTTCCTCTCCGAACTCTTCTCCTTCACTTCCT

27721  GGGAGAGGATTCCACAAGACCGAAGTGCCCAAGGAGACATGGAGGGAAGAGAGGGCTCTT
27721  CCCTCTCCTAAGGTGTTCTGGCTTCACGGGTTCCTCTGTACCTCCCTTCTCTCCCGAGAA

27781  CACAGAGCTGAGTTTAGAGACCACCCTCCATGGTTCTGGAAGGAAACCAGCCAGCTTTCA
27781  GTGTCTCGACTCAAATCTCTGGTGGGAGGTACCAAGACCTTCCTTTGGTCGGTCGAAAGT

27841  TATGTGGTGACAGGCTGTTCATTAATTTGTAGTGTACAGCATTAAATAAATGAAGTATGT
27841  ATACACCACTGTCCGACAAGTAATTAAACATCACATGTCGTAATTTATTTACTTCATACA

27901  AGAGAACATTAGGAAACAATATAAAAAGGTAGGGGAATTCGTGGATAAAGGGAATAACCT
27901  TCTCTTGTAATCCTTTGTTATATTTTTCCATCCCCTTAAGCACCTATTTCCCTTATTGGA

27961  GGGTACCATATTTCATCACATTATTCCATGGGAATAAGTCATCAGTGCAAAGGACTGTAA
27961  CCCATGGTATAAAGTAGTGTAATAAGGTACCCTTATTCAGTAGTCACGTTTCCTGACATT

28021  GGAGTGCACAGCAGGAAATGAAAAGACAGAAGCAGTGGGGCCGAGAGGAGGAAATGGAGT
28021  CCTCACGTGTCGTCCTTTACTTTTCTGTCTTCGTCACCCCGGCTCTCCTCCTTTACCTCA

28081  TGCTGAAATGCAGGGTGAGCAGCAGGGGCGGAGAGTGCCCTGTGGGTAAGGCACACCTCA
28081  ACGACTTTACGTCCCACTCGTCGTCCCCGCCTCTCACGGGACACCCATTCCGTGTGGAGT

28141  GAAGGGACAGGTCCAGGCCCTCTCCTCCCTGGCCTGGAGCTTCTGGCAAGGCTTTTCTTC
28141  CTTCCCTGTCCAGGTCCGGGAGAGGAGGGACCGGACCTCGAAGACCGTTCCGAAAAGAAG

28201  TCTGAGCCTCACTTTCCTCTTTTGAAGGTGGCAATAGTAACTGAACCTGCATCANNNNNN
28201  AGACTCGGAGTGAAAGGAGAAAACTTCCACCGTTATCATTGACTTGGACGTAGTNNNNNN

28261  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28261  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28321  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28321  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28381  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28381  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28441  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28441  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28501  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28501  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28561  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28561  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28621  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28621  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28681  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28681  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28741  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28741  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28801  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28801  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28861  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
28861  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

28921  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
```

28921 NNNN

30421 NNNN

31861 NNNN

33301 NNNN

34801 NNNN

36241 NNNN

37681 NNNN

39181 NNNN

40621 NNNN

42061 NNNN

43561 NNNN

45001 NNNN

46441 NNNN

47941 NNNN

49381 NNNN

50821 NNNN

52321 NNNN

53761 NNNN

55201 NNNN

56701 NNNN

58141 NNNN

59581 NNNN

61081 NNNN

62521 NNNN

63961 NNNN

65461 NNNN

66901 NNNN

68341 GAAAGAAGAAGGAAGGAAGGAAAGAAGGAAAAAGAAAAATTAAAAGGAGGAAAGAGAAAG

68401 TCCTTTCTTTCCTTCCTTCTTTCCTCTTTTTTTCTTTCCTCCCTCCCTCCCTTTCTTTCT
68401 AGGAAAGAAAGGAAGGAAGAAAGGAGAAAAAAAGAAAGGAGGGAGGGAGGGAAAGAAAGA

68461 TCCTTCCTTTCTTTCTTCCCTCCTTCCCTCCCCACCCTCCCTCCCTTCCCCTCGTCTCCC
68461 AGGAAGGAAAGAAAGAAGGGAGGAAGGGAGGGGTGGGAGGGAGGGAAGGGGAGCAGAGGG

68521 TTCCCCTCGTCTCCCTTCCCCTCCCCTCCCTCCCTCCCTTCCTCCTTTTTCTTTCCCTTT
68521 AAGGGGAGCAGAGGGAAGGGGAGGGGAGGGAGGGAGGGAAGGAGGAAAAAGAAAGGGAAA

68581 CCTCCTTCCTTCCTTCCTCTCTTCTTTTCTCCATCCCTTTCTTCCTTTTAACATTTGATC
68581 GGAGGAAGGAAGGAAGGAGAGAAGAAAAGAGGTAGGGAAAGAAGGAAAATTGTAAACTAG

68641 ATGGGTACCATTTCATCAAGATTCTGTTACTTAGGTCCAAGGACTAGTGACCAAATTGTT
68641 TACCCATGGTAAAGTAGTTCTAAGACAATGAATCCAGGTTCCTGATCACTGGTTTAACAA

68701 TTTCCCAGGAGTAAACTTTTCTACAAGACATAGAGAAATGACTGCCAGTGCCATCAGGCA
68701 AAAGGGTCCTCATTTGAAAAGATGTTCTGTATCTCTTTACTGACGGTCACGGTAGTCCGT

68761 CACCATCCTGTGTGATAGGTAACCTAAAATAATGGTACAGCGGGGAGTAACTGGAAGAGT
68761 GTGGTAGGACACACTATCCATTGGATTTTATTACCATGTCGCCCCTCATTGACCTTCTCA

68821 CACAACTACACAAATCTCTTGCATTTTCCTGATCATTAAATGAATTAGGGATTTCCTGGT
68821 GTGTTGATGTGTTTAGAGAACGTAAAAGGACTAGTAATTTACTTAATCCCTAAAGGACCA

68881 TTGAGATTTGTGGACAAGGAGAACTACTTCTTTGTACAAGAGGCTTGTGCCAGGAGAAGA
68881 AACTCTAAACACCTGTTCCTCTTGATGAAGAAACATGTTCTCCGAACACGGTCCTCTTCT

68941 CACGCTGTGGTTACAATGTTACTGTCTCTAGCTATGGTCATTTCATGTCTTTAACTGAGT
68941 GTGCGACACCAATGTTACAATGACAGAGATCGATACCAGTAAAGTACAGAAATTGACTCA

69001 GTACCTAGAGTGTAATACAGTGCAGCAGTGATGGGCTGAAGGCCTAATGCACCAGCACAC
69001 CATGGATCTCACATTATGTCACGTCGTCACTACCCGACTTCCGGATTACGTGGTCGTGTG

69061 ATGCCTGGAAGACCTAACGTTGAGAATCCGCATTTTTTTCATCACTCACATTTATTTTGC
69061 TACGGACCTTCTGGATTGCAACTCTTAGGCGTAAAAAAAAGTAGTGAGTGTAAATAAAACG

69121 ATCTCCCTGTTTTGATTACCTTCAGAGACCCAAGCACTTGCTTTTGTATATTTGAGTCTG
69121 TAGAGGGACAAAACTAATGGAAGTCTCTGGGTTCGTGAACGAAAACATATAAACTCAGAC

69181 GGTTTTATTTTCCTTTTAGAAATAGTCCAAAATCCTTTTAGAGCCTAGACCTTTGAATAA
69181 CCAAAATAAAAGGAAAATCTTTATCAGGTTTTAGGAAAATCTCGGATCTGGAAACTTATT

69241 AGTGCTAACTTAACCTGAGGAATGCTGTTTTTTTGTTCACAAATAAGGCTTAATTGTGAA
69241 TCACGATTGAATTGGACTCCTTACGACAAAAAAACAAGTGTTTATTCCGAATTAACACTT

69301 GTCGTGAGAATTGCTTTCTGAGCTGACTTATAAATTTACCCACCTCTTTTTCCCCAAACC
69301 CAGCACTCTTAACGAAAGACTCGACTGAATATTTAAATGGGTGGAGAAAAAGGGGTTTGG

69361 CCCAAAAAATCCTCCAGAAGGTGAGCCCAGCTTGTTGTAAAAGATTTTGAAGCAGAAGTC
69361 GGGTTTTTTAGGAGGTCTTCCACTCGGGTCGAACAACATTTTCTAAAACTTCGTCTTCAG

69421 CTGGGTTTAAGCTTCAGTTCTCTGCCTTACCAGAATCGGGATGTGTGGGGCAGGCCTACC
69421 GACCCAAATTCGAAGTCAAGAGACGGAATGGTCTTAGCCCTACACACCCCGTCCGGATGG

69481 TCCTCTTTCTTCATCTGCAAAATACATTGAATGAAATCATATTTTTTCCTAAGTATATGA
69481 AGGAGAAAGAAGTAGACGTTTTATGTAACTTACTTTAGTATAAAAAAGGATTCATATACT

69541 CATGTAAAGTGGTATTTCAAAAACTCATTAATCTGCATGACCAGGTCAGAAAAGACACAC
69541 GTACATTTCACCATAAAGTTTTTGAGTAATTAGACGTACTGGTCCAGTCTTTTCTGTGTG

69601 CTTTTGCCTTGTGTCTAGTGTGTGACTGTCCTCACATGTAGATGACCCAGGCCACATCCG
69601 GAAAACGGAACACAGATCACACACTGACAGGAGTGTACATCTACTGGGTCCGGTGTAGGC

69661 CTGATGTTGTCTTCCTTTACTTCTTTGTCCCTTCCAATGACTCCTTTATGAAGTCAATCA
69661 GACTACAACAGAAGGAAATGAAGAAACAGGGAAGGTTACTGAGGAAATACTTCAGTTAGT

69721 GCAGTGTTTCTAAGCCAGGTTCCTTCTCTTCACACATCCCCAACCCACAGGGCAAGTCAT
69721 CGTCACAAAGATTCGGTCCAAGGAAGAGAAGTGTGTAGGGGTTGGGTGTCCCGTTCAGTA

69781 CCACATTTTGCCTCTCTATTTTTTGTGCTTTGGGAAACACATTTTACAACTTCCAAACTA
69781 GGTGTAAAACGGAGAGATAAAAAACACGAAACCCTTTGTGTAAAATGTTGAAGGTTTGAT

```
69841  TTTGCCTAAGTTCATCTGAGCACATCTAGAATTATTCCACTCCTCTGCATCAGTGCTAAA
69841  AAACGGATTCAAGTAGACTCGTGTAGATCTTAATAAGGTGAGGAGACGTAGTCACGATTT

69901  AGTACAATCATGTATAAGCTGTTATACTCCCGTTATTTTGTACCTCATGTGCAGGTAGAT
69901  TCATGTTAGTACATATTCGACAATATGAGGGCAATAAAACATGGAGTACACGTCCATCTA

69961  TGCAAGTTCCTTGCAGGCAGGTACTACATCTACTTGAATTGCTTCACGATCACATAACAC
69961  ACGTTCAAGGAACGTCCGTCCATGATGTAGATGAACTTAACGAAGTGCTAGTGTATTGTG

70021  AATGCAGTGTTAAAAGTAGAAAAGTAATTGACATCTTGACTAACAGAAACACTAAATTCA
70021  TTACGTCACAATTTTCATCTTTTCATTAACTGTAGAACTGATTGTCTTTGTGATTTAAGT

70081  GTGTATTGTGATGATACCTTTTTAAGAGATAGGCTATAGAAATAGAAAATCATCTTAGTT
70081  CACATAACACTACTATGGAAAAATTCTCTATCCGATATCTTTATCTTTTAGTAGAATCAA

70141  GGAATTGAAAGTCAAGAAGCAACATTTTAAGTAAGAGATGAAAATACTTAAAAGGGAAAG
70141  CCTTAACTTTCAGTTCTTCGTTGTAAAATTCATTCTCTACTTTTATGAATTTTCCCTTTC

70201  GAGCAGTGCTTACACAGATGAGGGGAAAAAATAAAATGTATCATCTAAAAATTTAAATCA
70201  CTCGTCACGAATGTGTCTACTCCCCTTTTTTATTTTACATAGTAGATTTTTAAATTTAGT

70261  AATATATTAGGTTGGTGCAAAAGTAACTGCGATTTTTGCCATTAAAAGTAATAGCAAAAA
70261  TTATATAATCCAACCACGTTTTCATTGACGCTAAAAACGGTAATTTTCATTATCGTTTTT

70321  CCACGGTTACTTTTGCACCAACCTAATAGATGGGAAAAATAAGAGGAATGAATATTTTAA
70321  GGTGCCAATGAAAACGTGGTTGGATTATCTACCCTTTTTATTCTCCTTACTTATAAAATT

70381  GCTTTGCTATATAATTAAAATATTCTTAGAAGTCTGGAGTCTGTGAAGGTCACACCCTCT
70381  CGAAACGATATATTAATTTTATAAGAATCTTCAGACCTCAGACACTTCCAGTGTGGGAGA

70441  GGTCTTCTCCCAGCCCATAGGGTATAAATAATCTGAATTGACGGCATCCAGGGATCTCAG
70441  CCAGAAGAGGGTCGGGTATCCCATATTTATTAGACTTAACTGCCGTAGGTCCCTAGAGTC

70501  AAATTATTAGTACATCCCACAGTGAATTACCACCTTACTAAAATATTCATGGGTATATAC
70501  TTTAATAATCATGTAGGGTGTCACTTAATGGTGGAATGATTTTATAAGTACCCATATATG

70561  TATGGATTTGTTTTATCCTATTTAGTCTTAAAAACTATAAAGAAATCTGCAGGCTTATTA
70561  ATACCTAAACAAAATAGGATAAATCAGAATTTTTGATATTTCTTTAGACGTCCGAATAAT

70621  ACATATTACTCAGAATCATATTGTCTCCAAAGCACAAACTGAATCAGTTACAAGATATTG
70621  TGTATAATGAGTCTTAGTATAACAGAGGTTTCGTGTTTGACTTAGTCAATGTTCTATAAC

70681  GACTAGAGATCATGGCAAATCAGAGGTACATAAGACCTAGTTCCGTTGTGGAGCTAAACA
70681  CTGATCTCTAGTACCGTTTAGTCTCCATGTATTCTGGATCAAGGCAACACCTCGATTTGT

70741  AACTGCAGAGACCTAAAGGGAAGCCTTGCACCACACTCTAGGTTTGGAGCTCAGGTTTTG
70741  TTGACGTCTCTGGATTTCCCTTCGGAACGTGGTGTGAGATCCAAACCTCGAGTCCAAAAC

70801  AGTGGTGTCAGCACTCCAGAACACATGGGATCCCCGGGAGGTGGAAATTGAGCCGTCTTT
70801  TCACCACAGTCGTGAGGTCTTGTGTACCCTAGGGGCCCTCCACCTTTAACTCGGCAGAAA

70861  GGAGAATCAGCTAATGAGACAGATGCATGTTAAATGTCTGTTGTGGCCCAGGCACTCTGC
70861  CCTCTTAGTCGATTACTCTGTCTACGTACAATTTACAGACAACACCGGGTCCGTGAGACG

70921  TAGGCAGAGGGGTGAACCAGAAGAATGAGATTCATGGGGCCAAAGAATTTGCCTTCTGGT
70921  ATCCGTCTCCCCACTTGGTCTTCTTACTCTAAGTACCCCGGTTTCTTAAACGGAAGACCA

70981  GTAAGAAAAGATGGAGGCAGCTTGGCAGAAAGAAAAAAAAGGTAAAAGATAGAAATGAAA
70981  CATTCTTTTCTACCTCCGTCGAACCGTCTTTCTTTTTTTTCCATTTTCTATCTTTACTTT

71041  TACAGAATAATCTATCTCCTCATCCCACAAGCAATTCTGCCTGGTTTGTGGCTCACGCCT
71041  ATGTCTTATTAGATAGAGGAGTAGGGTGTTCGTTAAGACGGACCAAACACCGAGTGCGGA

71101  GTAATCCCAGCACTTTGGGAGGCCGAGGCAGGAGGATCACTTGAGGTCAGGAGTTCGAGA
71101  CATTAGGGTCGTGAAACCCTCCGGCTCCGTCCTCCTAGTGAACTCCAGTCCTCAAGCTCT

71161  CTAGCCTGGACCACAAGGTGAAACCTCATCTCTACTAAAAATGCAAAAAAATTTAGGGCC
71161  GATCGGACCTGGTGTTCCACTTTGGAGTAGAGATGATTTTTACGTTTTTTTAAATCCCGG

71221  GGGTGCAGTGGCAGTTCTGCCTGGTTTGAACATCTGGTTCAGTTAAAATACTTCATGAAA
71221  CCCACGTCACCGTCAAGACGGACCAAACTTGTAGACCAAGTCAATTTTATGAAGTACTTT
```

71281 ATATTTAGCTTGATGCACTGAAATCCATCACTGCTTTTGTGGGATGCACTGACTGCAGTC
71281 TATAAATCGAACTACGTGACTTTAGGTAGTGACGAAAACACCCTACGTGACTGACGTCAG

71341 TGACTGTTCTCCTAGTGGAAAGGGGAGCAATAGCTTCTGTCTGCATGTGCATGAAACAGA
71341 ACTGACAAGAGGATCACCTTTCCCCTCGTTATCGAAGACAGACGTACACGTACTTTGTCT

71401 TGGAAATTTAGAAGAGATTCTAGTGCAGACAACTTCATCCCACCCCCGCTGCCCACAGAG
71401 ACCTTTAAATCTTCTCTAAGATCACGTCTGTTGAAGTAGGGTGGGGGCGACGGGTGTCTC

71461 CTAGCCATGTGCACGGTGTGTTTGAACAATGTTATGTTGCTGCTGTGCACAGAGATCCCC
71461 GATCGGTACACGTGCCACACAAACTTGTTACAATACAACGACGACACGTGTCTCTAGGGG

71521 AGTCTCATAAAACATACTGGATGGAAGGCTGTCGAGAGATCATCAGGCCCCAGTGCATAA
71521 TCAGAGTATTTTGTATGACCTACCTTCCGACAGCTCTCTAGTAGTCCGGGGTCACGTATT

71581 GTTCTTCCTAAGCATCCTTAACAGGTGATCACCTCGCCTATTGAGAACCCTCAGGAGGCA
71581 CAAGAAGGATTCGTAGGAATTGTCCACTAGTGGAGCGGATAACTCTTGGGAGTCCTCCGT

71641 GCCCACTTACGGAAGCACTGATTGTTACAAAAGCCCTGAACTAAATCTGCTTGTCTCTAA
71641 CGGGTGAATGCCTTCGTGACTAACAATGTTTTCGGGACTTGATTTAGACGAACAGAGATT

71701 AGCCAGCCCTGCAGATGTCTGCAGTGTAACTACTTTCTCTTCTCTCAAGCAGTTCAGCCC
71701 TCGGTCGGGACGTCTACAGACGTCACATTGATGAAAGAGAAGAGAGTTCGTCAAGTCGGG

71761 TTCAAATCATTGACAACAGCTCTACAAAGCCCTCTAAATAGCCCGTTCTCCTGGCTAAAG
71761 AAGTTTAGTAACTGTTGTCGAGATGTTTCGGGAGATTTATCGGGCAAGAGGACCGATTTC

71821 TGTGTTGGTTTCCTCAGCATATGATGAAATAAACACCATGATGCCTTTGCCTTGCTTTTT
71821 ACACAACCAAAGGAGTCGTATACTACTTTATTTGTGGTACTACGGAAACGGAACGAAAAA

71881 CTTGCTGGTCTCACTCTTCCCTGGCATTCTTCAATGTGCCCAAGTCCTTCCTAGAACATT
71881 GAACGACCAGAGTGAGAAGGGACCGTAAGAAGTTACACGGGTTCAGGAAGGATCTTGTAA

71941 GCACCCAAAGGGCCTTCCTGGTCTTTCCAGGGAATGCAGAATGAAAAAGAAATCCCATCC
71941 CGTGGGTTTCCCGGAAGGACCAGAAAGGTCCCTTACGTCTTACTTTTTCTTTAGGGTAGG

72001 CCTGCCTCGAAAGGAGCTCTTCTGTGTGTAAGTGTACCTTCCCCCTCAGGGCATAATGGG
72001 GGACGGAGCTTTCCTCGAGAAGACACACATTCACATGGAAGGGGGAGTCCCGTATTACCC

72061 AAACAGTTTGCCAAAGCTAACTGTGTAATTTCACATGACTCACTGCCGTTAGTCTGAGAA
72061 TTTGTCAAACGGTTTCGATTGACACATTAAAGTGTACTGAGTGACGGCAATCAGACTCTT

72121 AGCATGGCATATTAAACTCATTTTCCATGCTTCCCAACTTTATTTACTATTCCATTTCTG
72121 TCGTACCGTATAATTTGAGTAAAAGGTACGAAGGGTTGAAATAAATGATAAGGTAAAGAC

72181 AGGGGCAGAGAGCTGAGCAGGAATAACACCTTCCTTGCTCCTCCTAATGGGGCACATGTG
72181 TCCCCGTCTCTCGACTCGTCCTTATTGTGGAAGGAACGAGGAGGATTACCCCGTGTACAC

72241 CACATATAAAATTGTGTTCGATCTTTGCAGCCACGACCCATGTTAACTGTCAACTGAGAA
72241 GTGTATATTTTAACACAAGCTAGAAACGTCGGTGCTGGGTACAATTGACAGTTGACTCTT

72301 CTCATGAGTCAGTTTCTCAGGAACAGCCCTGAAGACATTATCTCCTGCATCTTGTATTTG
72301 GAGTACTCAGTCAAAGAGTCCTTGTCGGGACTTCTGTAATAGAGGACGTAGAACATAAAC

72361 TGCCTTTGATATTTGGAGGGAAGGGTAGGATCTGGCAATTATCGTTATCTCATTTGGTTC
72361 ACGGAAACTATAAACCTCCCTTCCCATCCTAGACCGTTAATAGCAATAGAGTAAACCAAG

72421 AGTGTATTTCATTTCAGCCTGTTAAGATCTTTTGGAAATACTGTATGTGCTAGTTACAGG
72421 TCACATAAAGTAAAGTCGGACAATTCTAGAAAACCTTTATGACATACACGATCAATGTCC

72481 GACAGTGATAAATAAGACAGCATTCCTGTCCTTGAGATGCTATTTCCACAGCTCCATGAG
72481 CTGTCACTATTTATTCTGTCGTAAGGACAGGAACTCTACGATAAAGGTGTCGAGGTACTC

72541 ATGCCTAGTTAAAAACAGAGCCTTTTTTGGAAACAGCCTACTAATCTGTTAAAATAATAA
72541 TACGGATCAATTTTTGTCTCGGAAAAAACCTTTGTCGGATGATTAGACAATTTTATTATT

72601 TGGAAAAAGATAAAGTTAGCATACCTGAGCTGCAAGGGCTGCCTCCTGGTCTGATCTCTG
72601 ACCTTTTTCTATTTCAATCGTATGGACTCGACGTTCCCGACGGAGGACCAGACTAGAGAC

72661 ATAAGGATCAGATCCTAGAGCCTCTGAGATCCCTGTCCTCCCTGTCTGCACAAGCACTCG
72661 TATTCCTAGTCTAGGATCTCGGAGACTCTAGGGACAGGAGGGACAGACGTGTTCGTGAGC

72721 CAGAAGGAGAAACAGTTTACGGTGGTTCATCATGACTTTGAACCAAGCTTAAAGTCAAAG

72721 GTCTTCCTCTTTGTCAAATGCCACCAAGTAGTACTGAAACTTGGTTCGAATTTCAGTTTC

72781 TCATACTCTTTAAACCATTGGAAACCAAGTTTTTGCGAGTTGCCTAAAGTGGGCAAAAAT
72781 AGTATGAGAAATTTGGTAACCTTTGGTTCAAAAACGCTCAACGGATTTCACCCGTTTTTA

72841 CCCACAATGCACCTGAGCACAAGCAGGAAAACATGCCATTGTCTCCCCAGGAGCCCTCCT
72841 GGGTGTTACGTGGACTCGTGTTCGTCCTTTTGTACGGTAACAGAGGGGTCCTCGGGAGGA

72901 TGACTTCTACTTTACTTTATCCATTGACACTAGTCTTATAAGTAGCTTTGTCTGTCCAAT
72901 ACTGAAGATGAAATGAAATAGGTAACTGTGATCAGAATATTCATCGAAACAGACAGGTTA

72961 TTTTAATTGAATTTCGTTTTTATTTTCGTGAGTGGAAATATCACTTTGAAAGAAATCAGT
72961 AAAATTAACTTAAAGCAAAAATAAAAGCACTCACCTTTATAGTGAAACTTTCTTTAGTCA

73021 CCTCTCCTGAAATCCAGAATTCCTGGAGCTCAGTTTACATGTTTGACGCGTGTCACATGA
73021 GGAGAGGACTTTAGGTCTTAAGGACCTCGAGTCAAATGTACAAACTGCGCACAGTGTACT

73081 TTCCACAAGTCACTCAAGGGCAGGAGGATTACCATCGATACGGAAAGATTTCTTAGAAAG
73081 AAGGTGTTCAGTGAGTTCCCGTCCTCCTAATGGTAGCTATGCCTTTCTAAAGAATCTTTC

73141 CTTAAGTGAAAGGGAAACCAGGGAGAAAGTGTGCTTCGTGAGAATAGGTATGCGGACGGC
73141 GAATTCACTTTCCCTTTGGTCCCTCTTTCACACGAAGCACTCTTATCCATACGCCTGCCG

73201 TCTTTGTGACGCTCTGTGCAACCCTCGTGTTTGCATTGAACCAAGCTGTGCTGCAGATGG
73201 AGAAACACTGCGAGACACGTTGGGAGCACAAACGTAACTTGGTTCGACACGACGTCTACC

73261 ACACCCATCCCATTTCCCCGACTCATACCAGGGAGGCCCACTTTGCAAGGTACACAGAGG
73261 TGTGGGTAGGGTAAAGGGGCTGAGTATGGTCCCTCCGGGTGAAACGTTCCATGTGTCTCC

73321 GGACCACAGAGCAGGGGGCCATGCAGGGGACCAAGGGACATTTTAGTGTAACCAAAGTGT
73321 CCTGGTGTCTCGTCCCCCGGTACGTCCCCTGGTTCCCTGTAAAATCACATTGGTTTCACA

73381 GAAGTATGCCCTTTATTTCAAAAGAATAAAGAAAATCACAGGTTTTCCCGCTGATATGCC
73381 CTTCATACGGGAAATAAAGTTTTCTTATTTCTTTTAGTGTCCAAAAGGGCGACTATACGG

73441 AGGGACATTTCCAAGAGAATTCCTTTTTTGAGAGAAATCTCCTTTGATATCCCATCAGTC
73441 TCCCTGTAAAGGTTCTCTTAAGGAAAAAACTCTCTTTAGAGGAAACTATAGGGTAGTCAG

73501 AGCCATACTGCATAATTGTTAGATAGTGAAAGAAATTCATTTTTTAAGTTTGTCAGAAAA
73501 TCGGTATGACGTATTAACAATCTATCACTTTCTTTAAGTAAAAAATTCAAACAGTCTTTT

73561 ATAAATTCTTTGAAGTCTTAAATTGTTGCCTCCGGACATGACATATCTGGCTCTTCCAGA
73561 TATTTAAGAAACTTCAGAATTTAACAACGGAGGCCTGTACTGTATAGACCGAGAAGGTCT

73621 ATCCATGTTAGTCCTAGCTGAGGAAGAAGGAGAAGGAGAGGGGCGTTTGTTGATTATTGA
73621 TAGGTACAATCAGGATCGACTCCTTCTTCCTCTTCCTCTCCCCGCAAACAACTAATAACT

73681 TTTTGTAAGATGCCCCACACGTTGGCTATTAGCAGAATTCTCACTTCTAAAAGGAAAATG
73681 AAAACATTCTACGGGGTGTGCAACCGATAATCGTCTTAAGAGTGAAGATTTTCCTTTTAC

73741 AGTGTGAGCTATGTTCAATGAGAAACAGTTATTTTTGGGACATTCTTTGAGGTAAAACAC
73741 TCACACTCGATACAAGTTACTCTTTGTCAATAAAAACCCTGTAAGAAACTCCATTTTGTG

73801 CTCCTTAAGATGCTGCTTCCTTATTGCTATGGGACCAAGATTAGAGCAAGAACATAGTGG
73801 GAGGAATTCTACGACGAAGGAATAACGATACCCTGGTTCTAATCTCGTTCTTGTATCACC

73861 TTTTCAGACCCTGGACATCATCCACAGCCGCAGCAGAGGCCCTGCCCACTTGAACAATGA
73861 AAAAGTCTGGGACCTGTAGTAGGTGTCGGCGTCGTCTCCGGGACGGGTGAACTTGTTACT

73921 GACAGGCCACCATTTTGTTTCGAGAATGAGCAAAGTGAACCACCATGCCAGATATTGTTA
73921 CTGTCCGGTGGTAAAACAAAGCTCTTACTCGTTTCACTTGGTGGTACGGTCTATAACAAT

73981 AGTCAGCAACTCTTCTGAAGATGGACATAGTAAAAAATAAACAAACAAGGCAGCACTTGA
73981 TCAGTCGTTGAGAAGACTTCTACCTGTATCATTTTTTATTTGTTTGTTCCGTCGTGAACT

74041 GATGGTCATGGCAGAGCAATCTCAACAAGCGATTTGTTATTTTGCACAGTGATGTACCCA
74041 CTACCAGTACCGTCTCGTTAGAGTTGTTCGCTAAACAATAAAACGTGTCACTACATGGGT

74101 CTCATTGAATAAAATGCACCAGAACCATGCACTACAGATGCTAAGGAGAGTTGTCCTTCA
74101 GAGTAACTTATTTTACGTGGTCTTGGTACGTGATGTCTACGATTCCTCTCAACAGGAAGT

74161 ACAAAGGAGATAGGCCCCACTGGCCGTGGGGCAGTTTATGTTATTTGGTCCTTGTCATGG
74161 TGTTTCCTCTATCCGGGGTGACCGGCACCCCGTCAAATACAATAAACCAGGAACAGTACC

74221 GCAGGCATGGCCCTTTCTATGCTTCACAGATGAGGAAGGTCCCTGGCACAGGTCCAGTCT
74221 CGTCCGTACCGGGAAAGATACGAAGTGTCTACTCCTTCCAGGGACCGTGTCCAGGTCAGA

74281 CCAGCATGGCCTAGAGGTGGCAGGTGCTACTTAGCATCGCCAGCCTCCTGCTTGGTCATG
74281 GGTCGTACCGGATCTCCACCGTCCACGATGAATCGTAGCGGTCGGAGGACGAACCAGTAC

74341 GGGTCAGCCAGTTTATAACACGAACGGAGGTTAATGAACTGATCTTCCCATCGCACAACT
74341 CCCAGTCGGTCAAATATTGTGCTTGCCTCCAATTACTTGACTAGAAGGGTAGCGTGTTGA

74401 GGTATGAACCCACATCTTCTGATTATAAATCTTTTGCTCTTTAACTCTTAGTCATTACCA
74401 CCATACTTGGGTGTAGAAGACTAATATTTAGAAAACGAGAAATTGAGAATCAGTAATGGT

74461 CTGTCTAGTGTAGGCCTGTGTGTTCATGGCCTTTGTTGCCACCAAAAGATCAACTATTAG
74461 GACAGATCACATCCGGACACACAAGTACCGGAAACAACGGTGGTTTTCTAGTTGATAATC

74521 CTGAATAACATACTGAGACATGTTGGTGTTGTGTTCTAAATAGCACTAGTAAACTGTTAG
74521 GACTTATTGTATGACTCTGTACAACCACAACACAAGATTTATCGTGATCATTTGACAATC

74581 GGAAATCTGACTATATAGCTACTATCCAGTCTAGTTTTTCTTGCGAAGTGTTTGTTGAGT
74581 CCTTTAGACTGATATATCGATGATAGGTCAGATCAAAAAGAACGCTTCACAAACAACTCA

74641 GTGTAATGAGGAGTAAGGAAGGTGAATAAGAAATGGCCTGAGTCTTTATAAAATAAGCAA
74641 CACATTACTCCTCATTCCTTCCACTTATTCTTTACCGGACTCAGAAATATTTTATTCGTT

74701 GGAGAGAAAACAGTTTTGATGAGAAGCGCATGGAATTTTTAGAAGATAGGACGTGTATTA
74701 CCTCTCTTTTGTCAAAACTACTCTTCGCGTACCTTAAAAATCTTCTATCCTGCACATAAT

74761 TGTACCTATAAGAATGGGTAGGATTTTAGAAGAGATGGATGGGGAAAAGGAACAGGGAGT
74761 ACATGGATATTCTTACCCATCCTAAAATCTTCTCTACCTACCCCTTTTCCTTGTCCCTCA

74821 GGGAACAAAACGTGGACCAAGGAAGAGCAGGTTTAGCCATGGAAGCCTCACCGCCGGCTT
74821 CCCTTGTTTTGCACCTGGTTCCTTCTCGTCCAAATCGGTACCTTCGGAGTGGCGGCCGAA

74881 TGGTTTATCGTGGGCCAAGGGGACAGACCCTGTGGGGAGGGCTGGCAGCAGGGAGGTCTT
74881 ACCAAATAGCACCCGGTTCCCCTGTCTGGGACACCCCTCCCGACCGTCGTCCCTCCAGAA

74941 CCAGAAGTCTATCCTGCAGGCAGTAACAGCCACCCAGTCTATAAGCTGAGTGGGCATGGG
74941 GGTCTTCAGATAGGACGTCCGTCATTGTCGGTGGGTCAGATATTCGACTCACCCGTACCC

75001 GGTGATGGGGAATGGGTGGGGAGTTATTGGGGTAACTTACCCCAAAATGATAGCTAGCTG
75001 CCACTACCCCTTACCCACCCCTCAATAACCCCATTGAATGGGGTTTTACTATCGATCGAC

75061 GAACCATTTATTCTATTGCATTTTATCAATAAATCTTATAGGAAGTACCATCCTAGTGAA
75061 CTTGGTAAATAAGATAACGTAAAATAGTTATTTAGAATATCCTTCATGGTAGGATCACTT

75121 AACCCTGTCACATTGAGGGCATTCATGCTTATGTTTTAAAACATGTTATTGGGTCTATGA
75121 TTGGGACAGTGTAACTCCCGTAAGTACGAATACAAAATTTTGTACAATAACCCAGATACT

75181 AAAATAAGGCTGAAACCTATGAGCACCCTCCATGCAAAGTTTCAGTCAAGACTTTGGAAA
75181 TTTTATTCCGACTTTGGATACTCGTGGGAGGTACGTTTCAAAGTCAGTTCTGAAACCTTT

75241 CAAGACAGTGTCTTACTCACTTTATAAATTCATTCAGAAAGCCGTAGGTTTGAAAATTCC
75241 GTTCTGTCACAGAATGAGTGAAATATTTAAGTAAGTCTTTCGGCATCCAAACTTTTAAGG

75301 AAACTTAGATGTAAGAAGCTCTGAGAAACACATGAAATCACCCCCACATCAGTAGAGATG
75301 TTTGAATCTACATTCTTCGAGACTCTTTGTGTACTTTAGTGGGGGTGTAGTCATCTCTAC

75361 TCTCAGCAGACATGGGAAGAGGGGCAGCAGGGTGTAGGGAGGTGGGGCAGCCCCGGGGTG
75361 AGAGTCGTCTGTACCCTTCTCCCCGTCGTCCCACATCCCTCCACCCCGTCGGGGCCCCAC

75421 GGCCTTGCAGGCTGGGCTTGAATCCCATGGCCACCACCGTCCGCTGGGAGGCCTGGAGCC
75421 CCGGAACGTCCGACCCGAACTTAGGGTACCGGTGGTGGCAGGCGACCCTCCGGACCTCGG

75481 GGCTGCCCACTCTCTGACCAGCACATGTTGATGCTGTATCCTTGAAGGGACCGTGGTCTG
75481 CCGACGGGTGAGAGACTGGTCGTGTACAACTACGACATAGGAACTTCCCTGGCACCAGAC

75541 ACATCCTGTGATGCAGACCTGAATCCAGCACCCACAGGGTCTGCACATTCCCTCTTTGAG
75541 TGTAGGACACTACGTCTGGACTTAGGTCGTGGGTGTCCCAGACGTGTAAGGGAGAAACTC

75601 GTGGAGCCCAGCTCCAGAGGCTGGTCCCTGACTCTGTTTCTCAAGAAGCCTGTACAGATG
75601 CACCTCGGGTCGAGGTCTCCGACCAGGGACTGAGACAAAGAGTTCTTCGGACATGTCTAC

```
75661  TTCCCCTCACCACTGTTTCCAGTCACCTTTGGCTTTCACGGTGCAGATGCTAAGTTTGAT
75661  AAGGGGAGTGGTGACAAAGGTCAGTGGAAACCGAAAGTGCCACGTCTACGATTCAAACTA

75721  TTTCAGAGCCCATCTGGGAATTTAGTGAACTGAACAGGTAGCATTTCTGAACCCACCCAT
75721  AAAGTCTCGGGTAGACCCTTAAATCACTTGACTTGTCCATCGTAAAGACTTGGGTGGGTA

75781  AACCCATGCCCTCCCCACTGATTTTGAAAGAGAGTTTGCTGCAGGTGACTTTGCAGCTGG
75781  TTGGGTACGGGAGGGGTGACTAAAACTTTCTCTCAAACGACGTCCACTGAAACGTCGACC

75841  GTAGAGAATCTTGGGGCAGGATTCCGAGGCAGGCAGATGAGTGAGGATAAATTGGGTTCT
75841  CATCTCTTAGAACCCCGTCCTAAGGCTCCGTCCGTCTACTCACTCCTATTTAACCCAAGA

75901  GACGGCACGTTACACCAGTGGACTCTAACGACGACCTCACCTCGTGCACAGATAATTCTG
75901  CTGCCGTGCAATGTGGTCACCTGAGATTGCTGCTGGAGTGGAGCACGTGTCTATTAAGAC

75961  CCTTGTGCTTAACCGTTAGAAATGTGTCACTGAAGTGTGAACATATTATGCTGTTAGATT
75961  GGAACACGAATTGGCAATCTTTACACAGTGACTTCACACTTGTATAATACGACAATCTAA

76021  TCCCATCATTTCTGTTCTTTCATTCCCTCTCATTTGCATTGGTTACTCATAAATGTAGAT
76021  AGGGTAGTAAAGACAAGAAAGTAAGGGAGAGTAAACGTAACCAATGAGTATTTACATCTA

76081  CTTTGGTATGATTTGTACAACTGCCGGGTGTCAATCTGTGAAAGAAATAGCAGAGCAAGC
76081  GAAACCATACTAAACATGTTGACGGCCCACAGTTAGACACTTTCTTTATCGTCTCGTTCG

76141  TGGGCTCTGGTAGCGCTTTATCCCTGCGTGCTGGCTTGCCCGGGTTGACTCAGAGGCAGT
76141  ACCCGAGACCATCGCGAAATAGGGACGCACGACCGAACGGGCCCAACTGAGTCTCCGTCA

76201  CTCACATTCAGCTGCGCTGGGGCCAAGGACCCAGGGAGCCAAGTGTGTTTCTGTTTTCTG
76201  GAGTGTAAGTCGACGCGACCCCGGTTCCTGGGTCCCTCGGTTCACACAAAGACAAAAGAC

76261  TATTTAGCAATTTAAGACCTGCGTTTAAATACTAGCTATGCATTCTAGCAAAAGAGGTTT
76261  ATAAATCGTTAAATTCTGGACGCAAATTTATGATCGATACGTAAGATCGTTTTCTCCAAA

76321  ATATTTTAACACAGTAACTCTTAATTGTTTAATTCAGTTCGTGTGTTACCTCTCGGAATA
76321  TATAAAATTGTGTCATTGAGAATTAACAAATTAAGTCAAGCACACAATGGAGAGCCTTAT

76381  AAATAGTGGAAGCCAATTAACTATAGACTTCATTAGTTTGGATTTAAGATCACCAAAACA
76381  TTTATCACCTTCGGTTAATTGATATCTGAAGTAATCAAACCTAAATTCTAGTGGTTTTGT

76441  TTTACCACATCTCAATTGTTCATTACATGCTCTCTCTTTTTTAATGCAGTTTTTATAATA
76441  AAATGGTGTAGAGTTAACAAGTAATGTACGAGAGAGAAAAAATTACGTCAAAAATATTAT

76501  TGGGGAGTGGGGGTGTGGATTTAACCATTATATTTTTATGATATTGGGAGTAGATTTAAC
76501  ACCCCTCACCCCCACACCTAAATTGGTAATATAAAAATACTATAACCCTCATCTAAATTG

76561  CATTATGTAAATTGGATTTTTTAATTTTAAAAAGTCACTTATCTTGATGTAAAATCATGT
76561  GTAATACATTTAACCTAAAAAATTAAAATTTTTCAGTGAATAGAACTACATTTTAGTACA

76621  CTTAGTAACCTTGATAAACTAAGTTTTGCATGATTACACCTTAAGGTTAAAACATATTTC
76621  GAATCATTGGAACTATTTGATTCAAAACGTACTAATGTGGAATTCCAATTTTGTATAAAG

76681  ATTTCATCATTTCCCAGAAGGGGCACTGAATTCACCCATTTCCTGTTTTTCATCTCAGAA
76681  TAAAGTAGTAAAGGGTCTTCCCCGTGACTTAAGTGGGTAAAGGACAAAAAGTAGAGTCTT

76741  TGTTCTGTGTTTCCTCCATATCTACTTTCCCGGCAGCAGGACCCTGGAAGCAGTCACACC
76741  ACAAGACACAAAGGAGGTATAGATGAAAGGGCCGTCGTCCTGGGACCTTCGTCAGTGTGG

76801  AACCTCATTTACCCCACCTGAGATTTGTGCGCTTTGAACATAGTTGCAATCAAATCAACA
76801  TTGGAGTAAATGGGGTGGACTCTAAACACGCGAAACTTGTATCAACGTTAGTTTAGTTGT

76861  ATATCTTTGCTCAGAAATGGATATGTGAAGTAAAATGTGCTGCCCTGTTCATGTATGAAA
76861  TATAGAAACGAGTCTTTACCTATACACTTCATTTTACACGACGGGACAAGTACATACTTT

76921  TCATTCAGCTAGCTGGCCAGTGAGCCCTTCATTGCAACAAAGATTTTTCTAGAGCCCCTG
76921  AGTAAGTCGATCGACCGGTCACTCGGGAAGTAACGTTGTTTCTAAAAAGATCTCGGGGAC

76981  CACTATCTGGGGCTATGTCAGGCCTCACACTCCTGCTCACACGTTTGGAGGCTACCTTGG
76981  GTGATAGACCCCGATACAGTCCGGAGTGTGAGGACGAGTGTGCAAACCTCCGATGGAACC

77041  CCAGTATTACCTTAATCCAGCATTTAGGGGAAGGAGCATTTCAAGACTAAATTTTCTAAA
77041  GGTCATAATGGAATTAGGTCGTAAATCCCCTTCCTCGTAAAGTTCTGATTTAAAAGATTT

77101  CTGCTCAAGCCTACCTCATTTTATTTCTTGTGTATTTTAACACTTTTGGATGAGGACTCT
```

77101 GACGAGTTCGGATGGAGTAAAATAAAGAACACATAAAATTGTGAAAACCTACTCCTGAGA

77161 TCCTAGAACCTACTAACAATTCCCCCCCGCCCCCATGCCAAGATTCTTTAAAGACTTTCT
77161 AGGATCTTGGATGATTGTTAAGGGGGGGCGGGGGTACGGTTCTAAGAAATTTCTGAAAGA

77221 TGAAAACGCTTCAGTCTTTTCTTCTTAGCTCAAAAGTACTATCCTAAATACTAGCTCTGG
77221 ACTTTTGCGAAGTCAGAAAAGAAGAATCGAGTTTTCATGATAGGATTTATGATCGAGACC

77281 CATTACAGGGAGTTAATTTGTGGGCACACACAGTAAATTATAAACCCCTTAAGCAGAGAA
77281 GTAATGTCCCTCAATTAAACACCCGTGTGTGTCATTTAATATTTGGGGAATTCGTCTCTT

77341 GATTGTATTATCAGTTTATTAGTTTTCTTATTCATTCCTTCAGCACACATTTCTGTTGCC
77341 CTAACATAATAGTCAAATAATCAAAAGAATAAGTAAGGAAGTCGTGTGTAAAGACAACGG

77401 TCCAATGCAGCAGAGAAATTGACTTCTACAGTTTCCACAACAGTCCAGATTTCAACTGTG
77401 AGGTTACGTCGTCTCTTTAACTGAAGATGTCAAAGGTGTTGTCAGGTCTAAAGTTGACAC

77461 TAGTGCTCTTTCAGCCAGAAAGTACACTTGTGTTTCCCTGGCCATAGGTCCTGAACCTCA
77461 ATCACGAGAAAGTCGGTCTTTCATGTGAACACAAAGGGACCGGTATCCAGGACTTGGAGT

77521 CTTCTGAAAAGTCATTGTGCATAGAGAGCTAATAGCTGTACCCTAAATGATCCTGGCTTT
77521 GAAGACTTTTCAGTAACACGTATCTCTCGATTATCGACATGGGATTTACTAGGACCGAAA

77581 GAATTCTCTTATCTGCTTGGATAGTATTATCTGTCTCTTCCTCTGCATTCTAATTTGCTA
77581 CTTAAGAGAATAGACGAACCTATCATAATAGACAGAGAAGGAGACGTAAGATTAAACGAT

77641 CTTCTAATCTGCTGGGAATTACAATAAGAAAGAACCATTTAATCATTTTTACAACTGTGC
77641 GAAGATTAGACGACCCTTAATGTTATTCTTTCTTGGTAAATTAGTAAAAATGTTGACACG

77701 CTAAAGAGAGTGTGTGTAAGTGCCGAGAGAGTGTATGAGGGACTTGCCCATGAGTAAATG
77701 GATTTCTCTCACACACATTCACGGCTCTCTCACATACTCCCTGAACGGGTACTCATTTAC

77761 CATGAATTTTAGGTCAAGGGTTTTTTGCTTCTCTTTTGGTTGATTACCTCAGAGATCAGT
77761 GTACTTAAAATCCAGTTCCCAAAAAACGAAGAGAAAACCAACTAATGGAGTCTCTAGTCA

77821 TTTACTTTCTTTCTCATTCTTGACCTATCATCACTAGCTGATATGGATGATGTGTACAAC
77821 AAATGAAAGAAAGAGTAAGAACTGGATAGTAGTGATCGACTATACCTACTACACATGTTG

77881 TTCTGAGTAAGAATAATGTCAATGGGACGGGATGGGATTGGCTGGTGATTCTGTTGATCT
77881 AAGACTCATTCTTATTACAGTTACCCTGCCCTACCCTAACCGACCACTAAGACAACTAGA

77941 TAAAGTTTATATATTTTAAGTTTAGTGTTTCAGAATGAGACCAAAGCGGTGACATTTTCA
77941 ATTTCAAATATATAAAATTCAAATCACAAAGTCTTACTCTGGTTTCGCCACTGTAAAAGT

78001 ACCTCTTCGGTCTCTCTTCAGTTTTTTCATTTTAAGTTTTTGTTGTGCTTCTATCACTTA
78001 TGGAGAAGCCAGAGAGAAGTCAAAAAAGTAAAATTCAAAAACAACACGAAGATAGTGAAT

78061 AAGGAAGCCTCCAAGTTGAAATCAAATACTAATGACATTTTTATCTAATGTATAAATGTG
78061 TTCCTTCGGAGGTTCAACTTTAGTTTATGATTACTGTAAAAATAGATTACATATTTACAC

78121 TTTTTATTATTTATTAGGAAATTTATTTTACTTGGCCCTCAGCCATGACATATCATGGCA
78121 AAAAATAATAAATAATCCTTTAAATAAAATGAACCGGGAGTCGGTACTGTATAGTACCGT

78181 TAATCACTATCCTAAATTTGTATATCTTATCCTTGCATAAGGTAAGACTTTCTATGAATT
78181 ATTAGTGATAGGATTTAAACATATAGAATAGGAACGTATTCCATTCTGAAAGATACTTAA

78241 ACACATATTTGTATTTTCCTCTCTTACATATTTTAAGAATTTTTTTTATACTTTGTTTTCT
78241 TGTGTATAAACATAAAAGGAGAGAATGTATAAAATTCTTAAAAAAATATGAAACAAAAGA

78301 GCAAATGAAATATTGCTCATAAGCCAGGTGATGGCTGTCCACGCCTCTTTTCCCCTGCTT
78301 CGTTTACTTTATAACGAGTATTCGGTCCACTACCGACAGGTGCGGAGAAAAGGGGACGAA

78361 ACCCTTGCTTAGATCTTATGGTAGAATCTTTTCATAGAAGACACAGAAAGACATGAAAGA
78361 TGGGAACGAATCTAGAATACCATCTTAGAAAAGTATCTTCTGTGTCTTTCTGTACTTTCT

78421 AAGAGCTGGAGAAGCCTGAGGGGCTGCCCAGTGAGTGGTCTGCTAGGATGCTGTGCCACA
78421 TTCTCGACCTCTTCGGACTCCCCGACGGGTCACTCACCAGACGATCCTACGACACGGTGT

78481 GCCCAGGCACAGGAGGCAGGGAGAGCAATGGGGCCCTTCCCTTCCACCAACATTCAGCAG
78481 CGGGTCCGTGTCCTCCGTCCCTCTCGTTACCCCGGGAAGGGAAGGTGGTTGTAAGTCGTC

78541 AATTTGCAGCTCCATGTTTCCAAAGCTTCCAGGGCACTTGCATTTAGAGAGAGAGAGCAA
78541 TTAAACGTCGAGGTACAAAGGTTTCGAAGGTCCCGTGAACGTAAATCTCTCTCTCTCGTT

78601 GCAAGCTGCCTTTCCTCTTCCTCAGTTCTGCCAGCCACACTCTTGCCATGATGAGCAGTT
78601 CGTTCGACGGAAAGGAGAAGGAGTCAAGACGGTCGGTGTGAGAACGGTACTACTCGTCAA

78661 TCAGCCAAAAGCTCCTTCCCCTGCCCTAACACCTCCTGCAAGGCCTGAGGTCTGGAAGCC
78661 AGTCGGTTTTCGAGGAAGGGGACGGGATTGTGGAGGACGTTCCGGACTCCAGACCTTCGG

78721 ACCTGCGCCTGCTGGCCCCTCCTTTCTTGTTTCTGCAATGGATGTTGTGGCCCTGTGAGG
78721 TGGACGCGGACGACCGGGGAGGAAAGAACAAAGACGTTACCTACAACACCGGGACACTCC

78781 GAGAAGAGAAAAAGAAGTTGCCCTCCTCCCTCTATCCTCACCTCCTGCCATGCTGTCACC
78781 CTCTTCTCTTTTTCTTCAACGGGAGGAGGGAGATAGGAGTGGAGGACGGTACGACAGTGG

78841 CTTATAAGAGAGAAGGGCTAACCATCCAGGCTAATCCTCCAGTGATGCAGGAGAGGACAT
78841 GAATATTCTCTCTTCCCGATTGGTAGGTCCGATTAGGAGGTCACTACGTCCTCTCCTGTA

78901 CCTGGCCGGAAGAGTCAGAGCTTCCAGGTGAGCTCAGGTGGGTCAGCCCCCGAGGCTGTG
78901 GGACCGGCCTTCTCAGTCTCGAAGGTCCACTCGAGTCCACCCAGTCGGGGGCTCCGACAC

78961 AAGAGCCCAGGGGCCAGTAGATGCCACTTTTGCTCCAGGAAGAATCTTCAACTGTGTCCT
78961 TTCTCGGGTCCCCGGTCATCTACGGTGAAAACGAGGTCCTTCTTAGAAGTTGACACAGGA

79021 TTTTATTCAAGGGGCTCTCTTTTCAGCGAATCTCTAGATGTACTAGTCACAAACACTCGC
79021 AAAATAAGTTCCCCGAGAGAAAAGTCGCTTAGAGATCTACATGATCAGTGTTTGTGAGCG

79081 ATTTATTAAAATGTATCCATGATCCCACAATCCTTTTACATATTTCTCTCCAGGAGTATC
79081 TAAATAATTTTACATAGGTACTAGGGTGTTAGGAAAATGTATAAAGAGAGGTCCTCATAG

79141 ACATTTCTGAGGGCCCTGTGCCTGTTTTCTGCAGGAAGTTGCTGTTGTCCCGGGTCCCCC
79141 TGTAAAGACTCCCGGGACACGGACAAAAGACGTCCTTCAACGACAACAGGGCCCAGGGGG

79201 TGCCCCCAGCACCTCTGTTACAAGAAGCAGACCCTTCATGCCACACTGGGACCCAGGGAG
79201 ACGGGGGTCGTGGAGACAATGTTCTTCGTCTGGGAAGTACGGTGTGACCCTGGGTCCCTC

79261 GCCCCAAGCCAGGATGCTGGGATCTTAGTAAAGGTTGGAATGATGTCAGAACATAGAGGA
79261 CGGGGTTCGGTCCTACGACCCTAGAATCATTTCCAACCTTACTACAGTCTTGTATCTCCT

79321 GGCAGAATTCCCCCCATAGCATCATCCTGGAGGGCGCTGATTTGTGTGCTCTGCCAGGTT
79321 CCGTCTTAAGGGGGGTATCGTAGTAGGACCTCCCGCGACTAAACACACGAGACGGTCCAA

79381 CATCTGTGACTCAGGATTTAAAAGCCCCAGGTGTGGTGTCCCTTCTGTGCCTGCAAGGTG
79381 GTAGACACTGAGTCCTAAATTTTCGGGGTCCACACCACAGGGAAGACACGGACGTTCCAC

79441 CGTCTTTAGCAGTTCTCCCTGGTGTGGAAGGATCAGTGGTTCTTGCAGCCTAGGCACCCT
79441 GCAGAAATCGTCAAGAGGGACCACACCTTCCTAGTCACCAAGAACGTCGGATCCGTGGGA

79501 CCACAGCAAGCCCAACACAGGTGCTGTGAGCAGCTGGTTCATGAACGGTGATCCTGGGGA
79501 GGTGTCGTTCGGGTTGTGTCCACGACACTCGTCGACCAAGTACTTGCCACTAGGACCCCT

79561 GAAGAGGAGGATGAAAATGGAAAACCAGTGCAAAGGTGTGTCATCCAGTTGGTTACTGCT
79561 CTTCTCCTCCTACTTTTACCTTTTGGTCACGTTTCCACACAGTAGGTCAACCAATGACGA

79621 GTGGGTGCCCGGGCTTCCATCCCACCAGAGCCCCCACTAGCCAGGGCAGCCTGGCGGAGG
79621 CACCCACGGGCCCGAAGGTAGGGTGGTCTCGGGGGTGATCGGTCCCGTCGGACCGCCTCC

79681 TGAGCATCATTCTGCTTTCTGGCTGCACCTGTGTGGACAGAGCCCGTTCCAAAGCCTCCC
79681 ACTCGTAGTAAGACGAAAGACCGACGTGGACACACCTGTCTCGGGCAAGGTTTCGGAGGG

79741 AGGCACAGGGCAGACATCAAGCAAGGGGCAGAGCTGGGGCGAGGGCTCTGGGTGCCCCCG
79741 TCCGTGTCCCGTCTGTAGTTCGTTCCCCGTCTCGACCCCGCTCCCGAGACCCACGGGGGC

79801 TTATTGAACACAGGCCTGAAGGGAGCCTAAGAGGTGGCTGGCAGGAGATACTACAAGGCA
79801 AATAACTTGTGTCCGGACTTCCCTCGGATTCTCCACCGACCGTCCTCTATGATGTTCCGT

79861 GAAAGACAGGGCAGCAGTCTTTGTACTGTACGTTTTTGTTTTTTAAGAGAGAACAAAAAG
79861 CTTTCTGTCCCGTCGTCAGAAACATGACATGCAAAAACAAAAAATTCTCTCTTGTTTTTC

79921 TCAAGGTTGAGAGATGAGAAATACTCTTGCCAAAAAGAGAACACCAAAATCCTGGAGTCA
79921 AGTTCCAACTCTCTACTCTTTATGAGAACGGTTTTTCTCTTGTGGTTTTAGGACCTCAGT

79981 CGGGTCTTCCATTACCCTCTCTAGCATTTCTGCTTGTCTTTTCCCTAGTATTGGTGAAGA
79981 GCCCAGAAGGTAATGGGAGAGATCGTAAAGACGAACAGAAAAGGGATCATAACCACTTCT

80041 ATTTTTAAAGTAACCAACCCATACCGTTGGTTACAGGCCCTGTGGTCACCAGGCTGTCTC
80041 TAAAAATTTCATTGGTTGGGTATGGCAACCAATGTCCGGGACACCAGTGGTCCGACAGAG

80101 CCCATGATGGGGATGGAGAGTGGTTAGTGCAGAAACTTAGACCTCCCCTCCAGCTTGTTG
80101 GGGTACTACCCCTACCTCTCACCAATCACGTCTTTGAATCTGGAGGGGAGGTCGAACAAC

80161 AATGCCCTGAAGTTTATCTAGAAGGGAAGATACTCAAGCGTAGGATTTCAGTGATGCATT
80161 TTACGGGACTTCAAATAGATCTTCCCTTCTATGAGTTCGCATCCTAAAGTCACTACGTAA

80221 TGGACAGCATCAGTACTATGCTTCAGCGTCAAAAACGTCACTTTGGGTAGGAACAATACA
80221 ACCTGTCGTAGTCATGATACGAAGTCGCAGTTTTTGCAGTGAAACCCATCCTTGTTATGT

80281 TAATGCGTGGCAATGCCTTTTGTGACACCTGTTCAGGAGATTCCCATAGGAAGCTTCTGA
80281 ATTACGCACCGTTACGGAAAACACTGTGGACAAGTCCTCTAAGGGTATCCTTCGAAGACT

80341 GGCAGAGTCCTCAGGTGAGGAGGGGGACAGGCCTGGGTCTTGGGGGAAGGTGGAGATGAC
80341 CCGTCTCAGGAGTCCACTCCTCCCCCTGTCCGGACCCAGAACCCCCTTCCACCTCTACTG

80401 CAGCCTCATGCCTCTCTCCCCAGTGGCCTCAGTCTCCATCAGGCAAATCTTGAGAAGCCT
80401 GTCGGAGTACGGAGAGAGGGGTCACCGGAGTCAGAGGTAGTCCGTTTAGAACTCTTCGGA

80461 CCTCCATTCTGCAGGCAAATGACTGAGATGTGTGAGCTCTGCTTCCCAACAGGCGGAAAT
80461 GGAGGTAAGACGTCCGTTTACTGACTCTACACACTCGAGACGAAGGGTTGTCCGCCTTTA

80521 TCACATGGGGAAGGGCACCTGTGAATGGCCTTCTAGAACTATCAGGAAGTTCTGGATTTA
80521 AGTGTACCCCTTCCCGTGGACACTTACCGGAAGATCTTGATAGTCCTTCAAGACCTAAAT

80581 GTACCCTGTGAGAGCAGATGGTCCTGGGTGCACTCGGATGATCCTGCCTAGCCAGCTGGA
80581 CATGGGACACTCTCGTCTACCAGGACCCACGTGAGCCTACTAGGACGGATCGGTCGACCT

80641 TTGCAGAGTGGTCATCATTATCAGTGAGCTATTGAGGGGTTGAGAGCAGTCAGTGTCATG
80641 AACGTCTCACCAGTAGTAATAGTCACTCGATAACTCCCCAACTCTCGTCAGTCACAGTAC

80701 TATAAGGATGGGTTGGGGTACCGGCCAACGCCCGCCTTGAGGCCTGGCCGCCGGATGGGA
80701 ATATTCCTACCCAACCCCATGGCCGGTTGCGGGCGGAACTCCGGACCGGCGGCCTACCCT

80761 GGCGAGAGCACAAGAGAGTCCTGGCTAGGTGTGCCCCGCTCTGCAGGGCCACACCATCCA
80761 CCGCTCTCGTGTTCTCTCAGGACCGATCCACACGGGGCGAGACGTCCCGGTGTGGTAGGT

80821 CTGGAAACCCTGTGCCCACTGGGCTGCGTGCTAAGAGGCTGTGGCCAATGCTCTTGTCCA
80821 GACCTTTGGGACACGGGTGACCCGACGCACGATTCTCCGACACCGGTTACGAGAACAGGT

80881 AATTTTACTGAATCTGCAGTCTCTCTTAATTCACTCCAAAGTTTAATGTGTTAGCTTGCG
80881 TTAAAATGACTTAGACGTCAGAGAGAATTAAGTGAGGTTTCAAATTACACAATCGAACGC

80941 AATAAATAAATAAATAAATAAATAAATAAATAAATGGAAAGAAACAGTCCTCAGAAAGTC
80941 TTATTTATTTATTTATTTATTTATTTATTTACCTTTCTTTGTCAGGAGTCTTTCAG

81001 CCAGTCAAATTTTAATTCCAACAGATATTCAGCAGTTTCCTCTAAGAACAATGAGAGTTC
81001 GGTCAGTTTAAAATTAAGGTTGTCTATAAGTCGTCAAAGGAGATTCTTGTTACTCTCAAG

81061 TGGGGCTAGCCAGTGTTTCTCTTGGAAAATAAGGAAGAGGGAAAGCGGTGCATTCATTTA
81061 ACCCCGATCGGTCACAAAGAGAACCTTTTATTCCTTCTCCCTTTCGCCACGTAAGTAAAT

81121 AAAACCTGCCCTAGGGAGGCAGCGTTCCCTGTGAACTCCGAGGCCACAGTGAGCAGAGCA
81121 TTTTGGACGGGATCCCTCCGTCGCAAGGGACACTTGAGGCTCCGGTGTCACTCGTCTCGT

81181 GGCTGGAGGCCTCCCGGCTGTCCTGCCCCTCCTGCCACATGCCTGTGAAAATGCTAGATA
81181 CCGACCTCCGGAGGGCCGACAGGACGGGGAGGACGGTGTACGGACACTTTTACGATCTAT

81241 AGGACTTTTTCTCAGCAACTTCCACGCTCCTTCAGTGGGGATGTCTTTGACTCAGAGCTC
81241 TCCTGAAAAAGAGTCGTTGAAGGTGCGAGGAAGTCACCCCTACAGAAACTGAGTCTCGAG

81301 TGCCACTGGTTATCTCCACGAACAGAAAATGCCACAGATGGGTTAATTCACTGTGTTGTT
81301 ACGGTGACCAATAGAGGTGCTTGTCTTTTACGGTGTCTACCCAATTAAGTGACACAACAA

81361 CTCATTTTCCCTCAGTTTCAGGCTTTTCTCTCCTTGCCTGTTTTCCTCGCTTAAAAAATG
81361 GAGTAAAAGGGAGTCAAAGTCCGAAAAGAGAGGAACGGACAAAAGGAGCGAATTTTTTAC

81421 ATGTGGGGGTCCCTAAACGCATCTACCCCGATAGATTTATGTTTTCTTTTCCATTAGTCC
81421 TACACCCCCAGGGATTTGCGTAGATGGGGCTATCTAAATACAAAAGAAAAGGTAATCAGG

81481 ACTTTGCGTCTCAGCCCTAAAATTAAGTTTTTGATTATAATGTAAGGAAGTTTTACCATA

```
81481  TGAAACGCAGAGTCGGGATTTTAATTCAAAAACTAATATTACATTCCTTCAAAATGGTAT

81541  TTTTAACTCTGGCTTTTTAAATACAAAAGAAAAATAACAGAATGGCCTTCTAGAACTATC
81541  AAAATTGAGACCGAAAAATTTATGTTTTCTTTTTATTGTCTTACCGGAAGATCTTGATAG

81601  AGGAACTTCTGGATCTAGTACCCTGTGGGGACAGATGGTCTGGAGCCAGCGCACTTGGAT
81601  TCCTTGAAGACCTAGATCATGGGACACCCCTGTCTACCAGACCTCGGTCGCGTGAACCTA

81661  GATCTCTCTCAGCCATCTGGACCATAGAGTGAACTTGTTACTGTCCCTCCAAGGCTGACA
81661  CTAGAGAGAGTCGGTAGACCTGGTATCTCACTTGAACAATGACAGGGAGGTTCCGACTGT

81721  ACTATGAGCTCATACTCTTCTAGCAGTTTAATTTAGACCCAAGAAAGGCTGTGTGTGTGT
81721  TGATACTCGAGTATGAGAAGATCGTCAAATTAAATCTGGGTTCTTTCCGACACACACACA

81781  GAATTTGTGAGAGTGTGTGAGTGTGTCTTTGTGCCTGTGTTTCTGTGTGTGTCTGTGTGT
81781  CTTAAACACTCTCACACACTCACACAGAAACACGGACACAAAGACACACACAGACACACA

81841  ACATGTGTGTCTGTGGGGTGTGTCTGTGTGTGCACGGCTGTGTGTCTGTGCGTGTGCATG
81841  TGTACACACAGACACCCCACACAGACACACACGTGCCGACACACAGACACGCACACGTAC

81901  GCTTTGTCTGTGTGTGTGCGCACGCACTCTGCTAAGCTACTCAATGCAATTCCTTACTCT
81901  CGAAACAGACACACACACGCGTGCGTGAGACGATTCGATGAGTTACGTTAAGGAATGAGA

81961  TACTTCCCTTTCTGTCACTTCTCCATAATTCTTTGTATTTCGGTTGGGCTGGTATCTCGC
81961  ATGAAGGGAAAGACAGTGAAGAGGTATTAAGAAACATAAAGCCAACCCGACCATAGAGCG

82021  GGCGGCTTCCTCTTTTCCTGCATTCCTATATTTCATTATTTGCTCTTGTTCCTCTTCTAG
82021  CCGCCGAAGGAGAAAAGGACGTAAGGATATAAAGTAATAAACGAGAACAAGGAGAAGATC

82081  GGCTTTTACAATATAGCCAGGAGGATGTGGAAACCCAGTTACAAGATGACACACAAGCAC
82081  CCGAAAATGTTATATCGGTCCTCCTACACCTTTGGGTCAATGTTCTACTGTGTGTTCGTG

82141  AGTGTCACAATCGCTGTGCTTGGGCCCTACCTCCTGGAGACCAGGGTGGGACGGTGTCTC
82141  TCACAGTGTTAGCGACACGAACCCGGGATGGAGGACCTCTGGTCCCACCCTGCCACAGAG

82201  TGGATATGGAGGGAAGGGCGGAAGATCACGGGGTGGCAGAAGGCCCGACTGTCCGGTCTT
82201  ACCTATACCTCCCTTCCCGCCTTCTAGTGCCCCACCGTCTTCCGGGCTGACAGGCCAGAA

82261  CTGGAAGCTGGGGTCTCGTGGTCGCTGGGCTGGTGGTGCTCTAAAACCTAGAGTAACCGA
82261  GACCTTCGACCCCAGAGCACCAGCGACCCGACCACCACGAGATTTTGGATCTCATTGGCT

82321  GAGCAGGATGACTGCACCCCTGACTGCCGACTCACCGGGCTTCGGAGGGCCATGGGTGTG
82321  CTCGTCCTACTGACGTGGGGACTGACGGCTGAGTGGCCCGAAGCCTCCCGGTACCCACAC

82381  TGATTCCACTTGTGATCCTCCTACTTCCCCAGCCCACCCTGGTGGTCTCCCACTTCTCTC
82381  ACTAAGGTGAACACTAGGAGGATGAAGGGGTCGGGTGGGACCACCAGAGGGTGAAGAGAG

82441  TCTCTGCCTGTGGAGTGTTATAACCGACTCCTTCTTCTAGGCATCTCCATCCCAACCCCA
82441  AGAGACGGACACCTCACAATATTGGCTGAGGAAGAAGATCCGTAGAGGTAGGGTTGGGGT

82501  GACCAGTTCCCTTAAACCACCTCCCACCACAGGACCCCCACAAGCAGGGACCTCTCACAG
82501  CTGGTCAAGGGAATTTGGTGGAGGGTGGTGTCCTGGGGGTGTTCGTCCCTGGAGAGTGTC

82561  TTCCTGTTGCCTGGGGACTTCATCTCAGCTTCCAAGGCCTCTGGAACCAGGCCCTGCATC
82561  AAGGACAACGGACCCCTGAAGTAGAGTCGAAGGTTCCGGAGACCTTGGTCCGGGACGTAG

82621  CTGGGCATGACTGTCCCCATGATTCTCTGAGCCAGCCACCACCCCACCACCCCCCCACCG
82621  GACCCGTACTGACAGGGGTACTAAGAGACTCGGTCGGTGGTGGGGTGGTGGGGGGGTGGC

82681  CCAAGTCCACTGCTGACCCTGCTGTAGCCCCCCACCTCCACTTGTTGAGATTCCCTCTTC
82681  GGTTCAGGTGACGACTGGGACGACATCGGGGGGGTGGAGGTGAACAACTCTAAGGGAGAAG

82741  CTGCTTCTCATTTTGAATGTCACCCTTCTGTGACTTCTCACAACCCCCCTGCCACCCCTC
82741  GACGAAGAGTAAAACTTACAGTGGGAAGACACTGAAGAGTGTTGGGGGGACGGTGGGGAG

82801  CCTCCAGTTGGCATGGCTTAGTCTTCCTAAGAACATGTGCTGCTTCCCAGTACCCAGGGC
82801  GGAGGTCAACCGTACCGAATCAGAAGGATTCTTGTACACGACGAAGGGTCATGGGTCCCG

82861  TGCTGTGCACCGTGGCACGTGAGTTTCGGTGCATCTTCCTCCTGCTGTTTATAAATTGCA
82861  ACGACACGTGGCACCGTGCACTCAAAGCCACGTAGAAGGAGGACGACAAATATTTAACGT

82921  ACATCCCTGGACAGATCTTGGATTTTCCCTCTATAATTTCTCTGGAGACTCCCTCGGGGT
82921  TGTAGGGACCTGTCTAGAACCTAAAAGGGAGATATTAAAGAGACCTCTGAGGGAGCCCCA
```

```
82981  TTGCACACACATCTGTGGAATGCCCAAGTTGGAGATGGTCCTCCCTGATGTGTAGTGTGG
82981  AACGTGTGTGTAGACACCTTACGGGTTCAACCTCTACCAGGAGGGACTACACATCACACC

83041  GCTATTCAGAGGGCCTCATTTCATAAACTCATTAACAGCTAGGACGCGAAGACTTGGAGA
83041  CGATAAGTCTCCCGGAGTAAAGTATTTGAGTAATTGTCGATCCTGCGCTTCTGAACCTCT

83101  AGTTAGCTCAGGACACACGGAGTGGGAGGGAGGCAACGAGTTAAGAAATGGGCTTTGGAG
83101  TCAATCGAGTCCTGTGTGCCTCACCCTCCCTCCGTTGCTCAATTCTTTACCCGAAACCTC

83161  TGCCTCGAATTTGGCCTCCACTTTTCTTTTCTACTGTCTTACCTGGTCTTAGGCAGGTCG
83161  ACGGAGCTTAAACCGGAGGTGAAAAGAAAAGATGACAGAATGGACCAGAATCCGTCCAGC

83221  GGTGACCATCTGGAGTCGGAGTTTCTCTCATTTGTAAAGCGGGGGTGATATTTTTCTCAC
83221  CCACTGGTAGACCTCAGCCTCAAAGAGAGTAAACATTTCGCCCCCACTATAAAAAGAGTG

83281  AGGGCTGCTTCAAGGGTTAACAAAGTTAATGTCTAGGAAGGACATAGCATAGCACCTTCA
83281  TCCCGACGAAGTTCCCAATTGTTTCAATTACAGATCCTTCCTGTATCGTATCGTGGAAGT

83341  CCTGGTTGCAAGCCCCTGGAGTCACATGGCAACAGCTGTGGAACCAGGTAAATGACTTGA
83341  GGACCAACGTTCGGGGACCTCAGTGTACCGTTGTCGACACCTTGGTCCATTTACTGAACT

83401  CTTACATGTTACACGCTCTTAGTTGTTCGATTTGTAAATGGGAAGAGTGTGTGACCCAAA
83401  GAATGTACAATGTGCGAGAATCAACAAGCTAAACATTTACCCTTCTCACACACTGGGTTT

83461  TTAGTCGTTTCCAGATATTTCCTGGTAAATGGTTGTTGAATCATAAAACTAGAAAGATGG
83461  AATCAGCAAAGGTCTATAAAGGACCATTTACCAACAACTTAGTATTTTGATCTTTCTACC

83521  GAAAGAAAGGGAGGAACCCCCCTTACTTCGAAGAAGTGTTGTTGATGTGAAGGACAGGAC
83521  CTTTCTTTCCCTCCTTGGGGGGAATGAAGCTTCTTCACAACAACTACACTTCCTGTCCTG

83581  TTTTTGAGACACTCTCATCCTAAGAAACAGCTGATATCTACTAGGAAAAAACATGACATT
83581  AAAAACTCTGTGAGAGTAGGATTCTTTGTCGACTATAGATGATCCTTTTTTGTACTGTAA

83641  TGAAGTTTCTTCCTAAGAGATGTGAGGTTTTGACAGAAGTGCCAGGAAGCCAGAGGTGTG
83641  ACTTCAAAGAAGGATTCTCTACACTCCAAAACTGTCTTCACGGTCCTTCGGTCTCCACAC

83701  TGAGTAGTGGGCCAATCTCGCGTGAAGGCTGTGGGACGGGGCAGGAGAGGGGCAGCCAGC
83701  ACTCATCACCCGGTTAGAGCGCACTTCCGACACCCTGCCCCGTCCTCTCCCCGTCGGTCG

83761  ATTTCCTGAGCGCCAGCTAATGGCGGGGACCTGGCACCTTTGCTCTGTGACTCTCCAGCT
83761  TAAAGGACTCGCGGTCGATTACCGCCCCTGGACCGTGGAAACGAGACACTGAGAGGTCGA

83821  GTATGGTGACATGAGGCTGGCGTTTTCTTTGGTTTCATACGTGAGGAAGCTGAGGCTTGG
83821  CATACCACTGTACTCCGACCGCAAAAGAAACCAAAGTATGCACTCCTTCGACTCCGAACC

83881  TGATGTCCTGTGCATGGCTCGGACTCCAGAGCTACTCAGAGGCAGAGGCCGTCTGTAAAC
83881  ACTACAGGACACGTACCGAGCCTGAGGTCTCGATGAGTCTCCGTCTCCGGCAGACATTTG

83941  ACCTTCAGGTGGTTGGAGGCAGGGTGCTGCAGCCAATGCTGATCACTCACAAGCCAAGGC
83941  TGGAAGTCCACCAACCTCCGTCCCACGACGTCGGTTACGACTAGTGAGTGTTCGGTTCCG

84001  TTTATTGTAAGGGCTAAATAAAATAACATATGCCTAGCATTTATACAGCACGGGGCTTTT
84001  AAATAACATTCCCGATTTATTTTATTGTATACGGATCGTAAATATGTCGTGCCCCGAAAA

84061  TAGGAAATAGTGAGACGACCATGGAGAGGTGGAGGAAGAGAGAAGGGAAAGAAGAAAAAG
84061  ATCCTTTATCACTCTGCTGGTACCTCTCCACCTCCTTCTCTCTTCCCTTTCTTCTTTTTC

84121  AAAAAAAAAGCCTTAAAGAGTTTCTTAAGAGAACTATATATTACAAAGTCCTTGGAGTTT
84121  TTTTTTTTTCGGAATTTCTCAAAGAATTCTCTTGATATATAATGTTTCAGGAACCTCAAA

84181  TTTTTCCCTGGTTAGACATAAGTTTACATGAAACGTTAAACGGCTTAAAGTAGCCACATC
84181  AAAAAGGGACCAATCTGTATTCAAATGTACTTTGCAATTTGCCGAATTTCATCGGTGTAG

84241  TCCTTCCTTATGTCTTCTCACAGCCACTTGTGAGTTTCTTATCAAGATACAGAATTATGG
84241  AGGAAGGAATACAGAAGAGTGTCGGTGAACACTCAAAGAATAGTTCTATGTCTTAATACC

84301  CCAAGTGCGGTAGCTCACACGTGTAATCCCAGCACTTTGAGAGACTGAGGTGGAAGGATC
84301  GGTTCACGCCATCGAGTGTGCACATTAGGGTCGTGAAACTCTCTGACTCCACCTTCCTAG

84361  ATGTGTGCTCAGGAGTTTCAGACCAACCTAGGCAATATGGTAAAACCCTGTCTCTACAAA
84361  TACACACGAGTCCTCAAAGTCTGGTTGGATCCGTTATACCATTTTGGGACAGAGATGTTT
```

84421 AAAATACAAAAATTAGCTGGGCATGGTGGCATATTCCAATAGTTCCAGGCACTCGGGAGG
84421 TTTTATGTTTTTAATCGACCCGTACCACCGTATAAGGTTATCAAGGTCCGTGAGCCCTCC

84481 CTGAGACAGGAGGATTGCTTGAGCCCAAGAGGTCAAGACTGCAGTGAGCTGTGATCGAGC
84481 GACTCTGTCCTCCTAACGAACTCGGGTTCTCCAGTTCTGACGTCACTCGACACTAGCTCG

84541 CACTGCACTACAGCCTGGGCCACAGAGCAAGATCCTGTCTCAAAAAATAAAATAAAATAA
84541 GTGACGTGATGTCGGACCCGGTGTCTCGTTCTAGGACAGAGTTTTTTATTTTATTTTATT

84601 AATAAATTCTGCAACAAGTCACAATTCCTTGCTCAGAAATCTCTACAGGTGTGCTTCTTT
84601 TTATTTAAGACGTTGTTCAGTGTTAAGGAACGAGTCTTTAGAGATGTCCACACGAAGAAA

84661 GTTAAGAAATGGAGGAAACATGAATTTCATCCTGACTTCTGAGTCTTGTAGAGACCAGAT
84661 CAATTCTTTACCTCCTTTGTACTTAAAGTAGGACTGAAGACTCAGAACATCTCTGGTCTA

84721 CTGACTCCTGCAGCCTTCTGCCCACAGAGCTGGCAGGCGGGGGAATGTGCGCAGAGTGAG
84721 GACTGAGGACGTCGGAAGACGGGTGTCTCGACCGTCCGCCCCCTTACACGCGTCTCACTC

84781 GAGGAGCTGATCTGACATTCCGAACAGTGGAGACTGCTGCCTATTTGGCCGCACCACTGT
84781 CTCCTCGACTAGACTGTAAGGCTTGTCACCTCTGACGACGGATAAACCGGCGTGGTGACA

84841 CCTTCTGCATGGAAAAGTCACAGTAAATAATGGACTCTCTACCTGTGAGTCGATTTTACT
84841 GGAAGACGTACCTTTTCAGTGTCATTTATTACCTGAGAGATGGACACTCAGCTAAAATGA

84901 GAGCTGCCTTTTTAAATATATGTGTAGAAAGGGCGATCTCTGTGTGTCATTTGCACATGT
84901 CTCGACGGAAAAATTTATATACACATCTTTCCCGCTAGAGACACACAGTAAACGTGTACA

84961 ACATACACATGTACACACGTGCACACGTGGTCACTCATGTACACATGTGTGTGCACCAGT
84961 TGTATGTGTACATGTGTGCACGTGTGCACCAGTGAGTACATGTGTACACACACGTGGTCA

85021 GCACCAGACCAGATGCACCCACCCAATTCCACGCTCCTCAGCCCCTTGCTTTCCTCTGCA
85021 CGTGGTCTGGTCTACGTGGGTGGGTTAAGGTGCGAGGAGTCGGGGAACGAAAGGAGACGT

85081 GACAGCCAGTTCAGCTACAGTTATGCCAGATTGCCTGCACTCTGATCTTTACCATCACCA
85081 CTGTCGGTCAAGTCGATGTCAATACGGTCTAACGGACGTGAGACTAGAAATGGTAGTGGT

85141 GACTCTCAGCTTGAGCTCATGCCTGAGAAATCCACTTTCTGGGAGGAAGGCAGAAGAAAC
85141 CTGAGAGTCGAACTCGAGTACGGACTCTTTAGGTGAAAGACCCTCCTTCCGTCTTCTTTG

85201 TGCCAAGGCAGGACTGAAGACTTGACCTCCACCTGTAGTGGGGTCAGGTCACCAGAGGCT
85201 ACGGTTCCGTCCTGACTTCTGAACTGGAGGTGGACATCACCCCAGTCCAGTGGTCTCCGA

85261 GGCTGACTCCCCAGGATTTCTCAGCAGAGTATTACACAAATACCCCTTCCTTAAAGATTA
85261 CCGACTGAGGGGTCCTAAAGAGTCGTCTCATAATGTGTTTATGGGGAAGGAATTTCTAAT

85321 GCAACCACTTAAAGACACAAAGTGTGCAGGACTGTACCCATGGCCGAGAGCGGCGGGATT
85321 CGTTGGTGAATTTCTGTGTTTCACACGTCCTGACATGGGTACCGGCTCTCGCCGCCCTAA

85381 AAGAGGGCAACTGAGTTGTTCTTCCCATCCCTCCTCCTCTGTGCTGAGTCACCTTCTGCT
85381 TTCTCCCGTTGACTCAACAAGAAGGGTAGGGAGGAGGAGACACGACTCAGTGGAAGACGA

85441 TGGAATTCTGACAGGGACCCGTTGGGCTTCTGGAAGGGAGGGACAGCAGAGGATCCACCC
85441 ACCTTAAGACTGTCCCTGGGCAACCCGAAGACCTTCCCTCCCTGTCGTCTCCTAGGTGGG

85501 TCTGTGTGTCCTGGGGGAGATTACTTATCTCTGGCCTCCCTGAAGCAGGGCCTGGGCTTC
85501 AGACACACAGGACCCCCTCTAATGAATAGAGACCGGAGGGACTTCGTCCCGGACCCGAAG

85561 TGGAATCTTTGAGGCTGACACCCTGCCAGCCCTGGGGATGAGAGGGAATGGCCGTGTCTG
85561 ACCTTAGAAACTCCGACTGTGGGACGGTCGGGACCCCTACTCTCCCTTACCGGCACAGAC

85621 TCTGCAGAGCCTGAGGAGGAGCTGAGCACAGCCTCCAGAGTTCTCTTCAGTTGATCGCTT
85621 AGACGTCTCGGACTCCTCCTCGACTCGTGTCGGAGGTCTCAAGAGAAGTCAACTAGCGAA

85681 AGGTGGACAAAGGCCACAGAAATGGATTTAAACTCCTCAGCCCTTTCTTTGCATGTTTGT
85681 TCCACCTGTTTCCGGTGTCTTTACCTAAATTTGAGGAGTCGGGAAAGAAACGTACAAACA

85741 TCTCATTTGAAGTCAGAAGTGATATGTCCTACACCTCAAGAACGTGTGAAATGCACATAC
85741 AGAGTAAACTTCAGTCTTCACTATACAGGATGTGGAGTTCTTGCACACTTTACGTGTATG

85801 AATAACCCCATTTCAGGAAGCCAAGTCCAGCTTAACAGTCAAAACATTTCCTTCAGTCTT
85801 TTATTGGGGTAAAGTCCTTCGGTTCAGGTCGAATTGTCAGTTTTGTAAAGGAAGTCAGAA

85861 TAGTCCTTCACTTTGCCGAACTCCCTTTTACACCGGCAGCAACAGTTTAACCTGTTGCTT

```
85861  ATCAGGAAGTGAAACGGCTTGAGGGAAAATGTGGCCGTCGTTGTCAAATTGGACAACGAA

85921  CTGTAAGAGTGTGCTACTGGGAAAACCACATCTAAAACACGTGTGCAGTTACATCAGCTA
85921  GACATTCTCACACGATGACCCTTTTGGTGTAGATTTTGTGCACACGTCAATGTAGTCGAT

85981  GAGCACATGCTAAACAGTTGATCAAAGGCTCTTGCCTTGTGGCCCACGCTGCAGACACTC
85981  CTCGTGTACGATTTGTCAACTAGTTTCCGAGAACGGAACACCGGGTGCGACGTCTGTGAG

86041  TGACGACTGCCGAGCTCCGCAGCCCCATGTCGTCCCTTCCGCACTGCCTGCTGTGCCTCT
86041  ACTGCTGACGGCTCGAGGCGTCGGGGTACAGCAGGGAAGGCGTGACGGACGACACGGAGA

86101  CCTCTCCATGTGGCAGGGAACACAGCCAGTCATCACCATGTGGCTCTGCCCGGCGCTGCC
86101  GGAGAGGTACACCGTCCCTTGTGTCGGTCAGTAGTGGTACACCGAGACGGGCCGCGACGG

86161  CCAGCATGTCCTGACAGGGCCTAGATATGGAAAGGTGGCTCTCCATGCACACACCCCAAG
86161  GGTCGTACAGGACTGTCCCGGATCTATACCTTTCCACCGAGAGGTACGTGTGTGGGGTTC

86221  CCCCTCCTGCCCGCCGTGTGACCCACACTCTTATGGGCAGCCCAGTTATTTTGTAGCATT
86221  GGGGAGGACGGGCGGCACACTGGGTGTGAGAATACCCGTCGGGTCAATAAAACATCGTAA

86281  TCCCTTCCTTATCATTTTGGCCCAGTGATCCAGCACAAATCTCCCTTATTAGAATAAAAT
86281  AGGGAAGGAATAGTAAAACCGGGTCACTAGGTCGTGTTTAGAGGGAATAATCTTATTTTA

86341  TTGGAATGACAAAATTAAATTTCATTTTTCACTTATATTGAGGACCTCACACTCTTCACC
86341  AACCTTACTGTTTTAATTTAAAGTAAAAAGTGAATATAACTCCTGGAGTGTGAGAAGTGG

86401  CCTGCCACGATCCCTGACAAGAGCCTTCCTATCTAATCATTGTTCCTCCAGCCCTCTTAG
86401  GGACGGTGCTAGGGACTGTTCTCGGAAGGATAGATTAGTAACAAGGAGGTCGGGAGAATC

86461  TTTTCTTCAGCCTTTCTTGATTGCCTGAATGTCCCTTTCCCTTCTCCTTTTAAAGCATGA
86461  AAAAGAAGTCGGAAAGAACTAACGGACTTACAGGGAAAGGGAAGAGGAAAATTTCGTACT

86521  ACCAAGCTTTCTTACCCCGTTCTCATTATCATTTTTGCATTTTCTTCTTTGCATATGATT
86521  TGGTTCGAAAGAATGGGGCAAGAGTAATAGTAAAAACGTAAAAGAAGAAACGTATACTAA

86581  CTCCTTAAATTATAAAACTTGGGGGTAATTTCTAGAGGTGCCATCATAGTGCTTCTGTCT
86581  GAGGAATTTAATATTTTGAACCCCCATTAAAGATCTCCACGGTAGTATCACGAAGACAGA

86641  ACTCAGTGTCTTTAGAATCAGCAAATATCATTTTTACAAAAAAGTAGTATTTCTTCCAAA
86641  TGAGTCACAGAAATCTTAGTCGTTTATAGTAAAAATGTTTTTTCATCATAAAGAAGGTTT

86701  AAAGAGTAAGCAAGAAGGTTACAACACTGGGAAAATATCCCTAAGCCTGTTCTTCAACCT
86701  TTTCTCATTCGTTCTTCCAATGTTGTGACCCTTTTATAGGGATTCGGACAAGAAGTTGGA

86761  GTTGAATGTTTTCCCCTAAATTGTTATATGGAGATCCTGGACCCGGAAGTTGGCTGACAT
86761  CAACTTACAAAAGGGGATTTAACAATATACCTCTAGGACCTGGGCCTTCAACCGACTGTA

86821  GAAACAGGCCTAGCAGGGCAGCTGAGGAAATGCTCCAACCTCAGGATCCAGGAAGATTGC
86821  CTTTGTCCGGATCGTCCCGTCGACTCCTTTACGAGGTTGGAGTCCTAGGTCCTTCTAACG

86881  ACATGGCACCAAAACAATCATTTAAAAGCTAATCCTGGCCAGGCACAGTGGCTCACGCCT
86881  TGTACCGTGGTTTTGTTAGTAAATTTTCGATTAGGACCGGTCCGTGTCACCGAGTGCGGA

86941  GTAATCCCAGCACTTTGGGAGGCCGAGGTGGGCTGATCACAAGGTCAAGAGATCGAGACC
86941  CATTAGGGTCGTGAAACCCTCCGGCTCCACCCGACTAGTGTTCCAGTTCTCTAGCTCTGG

87001  ATCCTGACCAACATGGTGAAACCCCATCTCTACCAAAAGTACAAATATTAGCTGGGCATG
87001  TAGGACTGGTTGTACCACTTTGGGGTAGAGATGGTTTTCATGTTTATAATCGACCCGTAC

87061  GTGGTGCGCGCCTGTAGTCCCAGCTACCCGGGAGGCTGAGGCAGGAGAATCACTTGAACC
87061  CACCACGCGCGGACATCAGGGTCGATGGGCCCTCCGACTCCGTCCTCTTAGTGAACTTGG

87121  CGGGAGGAGGAGGTTGCAGTGAGCCAAGATCATGCCACTGCACTCCAGCCTGGTGACAGA
87121  GCCCTCCTCCTCCAACGTCACTCGGTTCTAGTACGGTGACGTGAGGTCGGACCACTGTCT

87181  GCAAGACTCCATCTCAAAAAAGAAAAAAAAAGCTAATCCCATAAAAGAACCATCATTTTG
87181  CGTTCTGAGGTAGAGTTTTTTCTTTTTTTTTCGATTAGGGTATTTTCTTGGTAGTAAAAC

87241  AACCTGCTGTTTCCTTTCTTGTACTCTTTAGAAGTACCCATTTCTCCCTTTCTAAGCCAT
87241  TTGGACGACAAAGGAAAGAACATGAGAAATCTTCATGGGTAAAGAGGGAAAGATTCGGTA

87301  AGGTGTATTAATTGGAGCTTTTTCTATCTTAATTAGTTCACTGTGAACAATTAAAATTGT
87301  TCCACATAATTAACCTCGAAAAAGATAGAATTAATCAAGTGACACTTGTTAATTTTAACA
```

87361 GTTAATAAAACAAAAACAAAAATAGGACTGGTGCCTAGTTGTACTACATGAAGAGAGAAA
87361 CAATTATTTTGTTTTTGTTTTTATCCTGACCACGGATCAACATGATGTACTTCTCTCTTT

87421 GGGCCAGACATTGGTTTTTCCTAATCTTCTGTCAAGTTCTCCAAATCATCTGCTCTGGAA
87421 CCCGGTCTGTAACCAAAAAGGATTAGAAGACAGTTCAAGAGGTTTAGTAGACGAGACCTT

87481 GGTAGCTCCAACAGCTGGGATTTGAAGTAAAGCATAGTGACTTTGGCCATCACTGACATG
87481 CCATCGAGGTTGTCGACCCTAAACTTCATTTCGTATCACTGAAACCGGTAGTGACTGTAC

87541 TCCCATTTGAAGCAACCAAAACTGTCCCTCGACCTGACACTCATCCCTGAAACACCATGA
87541 AGGGTAAACTTCGTTGGTTTTGACAGGGAGCTGGACTGTGAGTAGGGACTTTGTGGTACT

87601 GGGTAAGTGAGGCGCTTCGGAAGGTCCACTCAACCCCATTGCCAGATAGAGTAAGTGTCT
87601 CCCATTCACTCCGCGAAGCCTTCCAGGTGAGTTGGGGTAACGGTCTATCTCATTCACAGA

87661 GCCAGGGGCATTTGGAGCTGAAGGGAAGACTGTACAGACTCACGGTTTGCAGGCACTGAA
87661 CGGTCCCCGTAAACCTCGACTTCCCTTCTGACATGTCTGAGTGCCAAACGTCCGTGACTT

87721 GGCGTTTTCCTGCCTTCTTTTTCACCTTCAGTGGACTTGCAAAGCACTAGCCCTATTGTT
87721 CCGCAAAAGGACGGAAGAAAAAGTGGAAGTCACCTGAACGTTTCGTGATCGGGATAACAA

87781 TTTCCCATCTGGGAAACATGGTGAATGTTGGTTGGTTGTAGCTAATCCTATGGGTCTTGA
87781 AAAGGGTAGACCCTTTGTACCACTTACAACCAACCAACATCGATTAGGATACCCAGAACT

87841 GGTCTTTGTTGACAAGAAGGTAGATGTTATCTTTATCTGCGTGTGGCTTTCTACTAAAAC
87841 CCAGAAACAACTGTTCTTCCATCTACAATAGAAATAGACGCACACCGAAAGATGATTTTG

87901 ATGAGCTACAGGGCTCTCTTTTTTGTTTTAAAGCATTTTTCCATAAGGTTCACCCTTACT
87901 TACTCGATGTCCCGAGAGAAAAAACAAAATTTCGTAAAAAGGTATTCCAAGTGGGAATGA

87961 ATTGCTTATCTGAATAATATTACCTGCTGAGAAGTTTATTCATTGCTCACCAGTTGTAGG
87961 TAACGAATAGACTTATTATAATGGACGACTCTTCAAATAAGTAACGAGTGGTCAACATCC

88021 GAGATTTTGACACAGGACTGGAGGATTTTTTTCTCATCGTAACAGTGCAGACCCATGGAA
88021 CTCTAAAACTGTGTCCTGACCTCCTAAAAAAAGAGTAGCATTGTCACGTCTGGGTACCTT

88081 AGCTTGGAAGCAGTTGTGACCGGATAAGAGCAGGTTGAGGATGATAATCTTAGGGCAATG
88081 TCGAACCTTCGTCAACACTGGCCTATTCTCGTCCAACTCCTACTATTAGAATCCCGTTAC

88141 AGCAGGTTGATTGAGAGGGGTGCCTGAAAGCAAGGTCCCCATGATGCAAGCAAACAAACT
88141 TCGTCCAACTAACTCTCCCCACGGACTTTCGTTCCAGGGGTACTACGTTCGTTTGTTTGA

88201 CACATGCCAGACGGTGGACAGGAAACACAGGCAGGATGTGGTGCAGGCTGGGCTGCCTTG
88201 GTGTACGGTCTGCCACCTGTCCTTTGTGTCCGTCCTACACCACGTCCGACCCGACGGAAC

88261 GCTGGGCAAAGGCAGGGGCTTCATGGCCATGCTGGGATGTGACGTGGTTGGAAAACAATA
88261 CGACCCGTTTCCGTCCCCGAAGTACCGGTACGACCCTACACTGCACCAACCTTTTGTTAT

88321 AGAATGGAGACCTCTCCAAAGACCTCAGACATAATTTGGCCATGAGAGGCAAGGGTAGAG
88321 TCTTACCTCTGGAGAGGTTTCTGGAGTCTGTATTAAACCGGTACTCTCCGTTCCCATCTC

88381 GCGAGCCCCCTGTGGGTTGGAATGTATGGAATCCAGATGAAGCACCATCAACATGCATGG
88381 CGCTCGGGGGACACCCAACCTTACATACCTTAGGTCTACTTCGTGGTAGTTGTACGTACC

88441 GCTACATAGGAGAGCTGGGTCTGGGGGAAAAGATGGGATTTGGGTTTGGGTGTGCATGTT
88441 CGATGTATCCTCTCGACCCAGACCCCCTTTTCTACCCTAAACCCAAACCCACACGTACAA

88501 CCAGGCATGGAGGGAAGAATGGGCAGAATGACGAGGATCAAGCCCTGGTGGAGCCCCCAG
88501 GGTCCGTACCTCCCTTCTTACCCGTCTTACTGCTCCTAGTTCGGGACCACCTCGGGGGTC

88561 ACCAGGGCCCTGGGCATATTTCCACTTGCCTTCTTCTGCAAACTTTTGTGTGCCTGTTGG
88561 TGGTCCCGGGACCCGTATAAAGGTGAACGGAAGAAGACGTTTGAAAACACACGGACAACC

88621 GTGGAAGGTACTATGCTGGGGCCATCCAGGATCCGGAGCAGTGTAAGACATCATGGCCTG
88621 CACCTTCCATGATACGACCCCGGTAGGTCCTAGGCCTCGTCACATTCTGTAGTACCGGAC

88681 CATGGTGTGACCCTTGACATCTGAAAAGAACGAGCTGGTGGAGTGGGAAGGAGACAAGTG
88681 GTACCACACTGGGAACTGTAGACTTTTCTTGCTCGACCACCTCACCCTTCCTCTGTTCAC

88741 GCAGCAGAGTTGACATCATCAGGACTGACTGCAGGACAGGCGGTGAAACGAGCCACTGAG
88741 CGTCGTCTCAACTGTAGTAGTCCTGACTGACGTCCTGTCCGCCACTTTGCTCGGTGACTC

```
88801  GAGTTACCCAGTGCCCTGGGGAGCTTAGATGTGAGACTTACACCACCCAGCACTCCCCAG
88801  CTCAATGGGTCACGGGACCCCTCGAATCTACACTCTGAATGTGGTGGGTCGTGAGGGGTC

88861  CTCTGTGTCCCTGGGTCCCACTGAACTTGATTGTATTTTAGTTGCCTCTTTTCTAAGACC
88861  GAGACACAGGGACCCAGGGTGACTTGAACTAACATAAAATCAACGGAGAAAAGATTCTGG

88921  GGAGCAACAATACCTTCCTTGAGTGTGCATGTGCAGTCCATAGATGGCAGCCAGTGTTCT
88921  CCTCGTTGTTATGGAAGGAACTCACACGTACACGTCAGGTATCTACCGTCGGTCACAAGA

88981  CTTCCCATGAGAAGACTCTGAGCCTCCTCACTAGAGGGTGGTCTAAAAACAAATGGATCC
88981  GAAGGGTACTCTTCTGAGACTCGGAGGAGTGATCTCCCACCAGATTTTTGTTTACCTAGG

89041  ATCCACAGAAGGTCTGAAAGGTTTTCATAGAATCTTGACTTGGTAGCACTTCAGGATCTA
89041  TAGGTGTCTTCCAGACTTTCCAAAAGTATCTTAGAACTGAACCATCGTGAAGTCCTAGAT

89101  ACTCGTGCTTTCCAGAAGTGTGGCCTGTGGTGTTCTGTCTGGAGTATTTGGTCCTCCCAT
89101  TGAGCACGAAAGGTCTTCACACCGGACACCACAAGACAGACCTCATAAACCAGGAGGGTA

89161  CCACCTCCACCTTGCTGTCCTGTCCTGTGACTCTCACTAGCTCAGTCACCGGGTCAGGAA
89161  GGTGGAGGTGGAACGACAGGACAGGACACTGAGAGTGATCGAGTCAGTGGCCCAGTCCTT

89221  GCCCTGCCATGAAATGCTAGATTTGGAGGCATTGCCAAGGTACCACAAACCATACTCAAA
89221  CGGGACGGTACTTTACGATCTAAACCTCCGTAACGGTTCCATGGTGTTTGGTATGAGTTT

89281  CAAGGTCATAATAGACTAGTGATCTCACTGTGGCTAACAGGTTCATGTGATTTAAAGAAT
89281  GTTCCAGTATTATCTGATCACTAGAGTGACACCGATTGTCCAAGTACACTAAATTTCTTA

89341  GACAGCATTTTTTCTAAATTTTATAAATTTTATAAATTTCTCCTAAATACAAATCTGAAT
89341  CTGTCGTAAAAAAGATTTAAAATATTTAAAATATTTAAAGAGGATTTATGTTTAGACTTA

89401  GTGTTTACCTTTAGACAGATTTTCCCAGAAAATTGTCAGTGTTCTCGAATTTGAAAGTAG
89401  CACAAATGGAAATCTGTCTAAAAGGGTCTTTTAACAGTCACAAGAGCTTAAACTTTCATC

89461  ATCAGATCTCCTCCCTCTAGTTTTGAAAATCTCAAGTAGGTTTAGCCTCTCCCCAGAGTG
89461  TAGTCTAGAGGAGGGAGATCAAAACTTTTAGAGTTCATCCAAATCGGAGAGGGGTCTCAC

89521  ACAACGGCACAAGAAGTTTATAATTTTATGCACAGCTAAATGCAAACTGAAAAGTGTGGT
89521  TGTTGCCGTGTTCTTCAAATATTAAAATACGTGTCGATTTACGTTTGACTTTTCACACCA

89581  TTGTGGGTATTTCATTTTTCCCATAATGTTCATCAATTTGGCAATTCAAATAGGATTCAG
89581  AACACCCATAAAGTAAAAAGGGTATTACAAGTAGTTAAACCGTTAAGTTTATCCTAAGTC

89641  AATCCAAGGTGTTCGCAGCATTTTATATGAAAATGTGCCTGAGAGTTTGGGGAATAGAAA
89641  TTAGGTTCCACAAGCGTCGTAAAATATACTTTTACACGGACTCTCAAACCCCTTATCTTT

89701  TTCTTACCAAATAAGTTAAATTTGCTCTTGATAAGAATATATTTTGTAAAATGTAGAAAA
89701  AAGAATGGTTTATTCAATTTAAACGAGAACTATTCTTATATAAAACATTTTACATCTTTT

89761  CTGAGAAAAAAAGTTTGACACCTGTTTTATCTGCATCCTTTTATTAAACAGTGGCCTGAA
89761  GACTCTTTTTTTCAAACTGTGGACAAAATAGACGTAGGAAAATAATTTGTCACCGGACTT

89821  TTTACTAAAGGACAGAGAACAGGGTTGAGAGGTAATCACTGAATAATAGCTTCGCTAATC
89821  AAATGATTTCCTGTCTCTTGTCCCAACTCTCCATTAGTGACTTATTATCGAAGCGATTAG

89881  CCACAGTGACTGCTTCTTATGACAAGCAAGAGGGTCAACTTAGAGGCAGGACTGTCTTCA
89881  GGTGTCACTGACGAAGAATACTGTTCGTTCTCCCAGTTGAATCTCCGTCCTGACAGAAGT

89941  GGGGCAAAAGCAAGTCAGCACCTACAAACCTCTCCAAGAGAGTGAGAAAAAGCAGGTTGT
89941  CCCCGTTTTCGTTCAGTCGTGGATGTTTGGAGAGGTTCTCTCACTCTTTTTCGTCCAACA

90001  ATGGCTCATTTTGCAGCTTGCCAGAAGCAATGTGGGAATATTGTCAGACTGCACAGACCT
90001  TACCGAGTAAAACGTCGAACGGTCTTCGTTACACCCTTATAACAGTCTGACGTGTCTGGA

90061  GGGGCCAGCTGGCTCCTAGAGTGCGAGGAGCTCTCAGTGACTGCTGTGCATACTCACCTG
90061  CCCCGGTCGACCGAGGATCTCACGCTCCTCGAGAGTCACTGACGACACGTATGAGTGGAC

90121  CTCTGTGACTTTGTTCAGTCCCTACCCTGGATGTGTGGTGAAGGGTGCCAGGCTTAGCTG
90121  GAGACACTGAAACAAGTCAGGGATGGGACCTACACACCACTTCCCACGGTCCGAATCGAC

90181  CTGGACCAGTGAGGTCCCTGAGGGGGCAGGGCTAAGAGGTGTAGCAGTCGCACTTGACCA
90181  GACCTGGTCACTCCAGGGACTCCCCCGTCCCGATTCTCCACATCGTCAGCGTGAACTGGT

90241  GTCATTGATTCACTGGTTGGGAATTTGTTCATTGAGGGGACACCATTACTCCCCATGCCT
```

```
90241  CAGTAACTAAGTGACCAACCCTTAAACAAGTAACTCCCCTGTGGTAATGAGGGGTACGGA

90301  CACGTGCAGGTGAGTGTGTGTATATTTGCATTAACACACAAACACACTTATGTGTGTCCC
90301  GTGCACGTCCACTCACACACATATAAACGTAATTGTGTGTTTGTGTGAATACACACAGGG

90361  TGTATGTGCAGCTGCAGTACCTTACACATCTTCCAGTGCCTCTCAAGCATCATAAAAACA
90361  ACATACACGTCGACGTCATGGAATGTGTAGAAGGTCACGGAGAGTTCGTAGTATTTTTGT

90421  GCTTCCACAAAACCTCTACTACCCACCTGGCATCCACAGAGCTCCCAGTGATTGGCTCCC
90421  CGAAGGTGTTTTGGAGATGATGGGTGGACCGTAGGTGTCTCGAGGGTCACTAACCGAGGG

90481  AAAAGATACTGAACTTTGGCATGGAACAGAGTGAGTTTCCACATTTGCTGTTGTACATAA
90481  TTTTCTATGACTTGAAACCGTACCTTGTCTCACTCAAAGGTGTAAACGACAACATGTATT

90541  TTACCTGTCCAGCTAGAATAGTACACACTTCTTTTACTGAGCCCAGAAGTTTATGTGCCT
90541  AATGGACAGGTCGATCTTATCATGTGTGAAGAAAATGACTCGGGTCTTCAAATACACGGA

90601  GAACACGGATGGTACAGGAACAAGGGTGAGCGCTTATCAATTCAGCATTACTCTGGAGAC
90601  CTTGTGCCTACCATGTCCTTGTTCCCACTCGCGAATAGTTAAGTCGTAATGAGACCTCTG

90661  ATGAAACAAACTAATAGGTACAGCTATATTTTGTTTTTTAAACAATTTTGTGAACCACAT
90661  TACTTTGTTTGATTATCCATGTCGATATAAAACAAAAAATTTGTTAAAACACTTGGTGTA

90721  ATTTGGGACAAAAATCAACAAGTAAATTTGAATTACACTAGGCAAAGGACTGATTCTTTT
90721  TAAACCCTGTTTTTAGTTGTTCATTTAAACTTAATGTGATCCGTTTCCTGACTAAGAAAA

90781  GAAACGTGATTGACATTCAACGTCTAACTTTCCAAAAGAAGACATACATGTGGCCCACAA
90781  CTTTGCACTAACTGTAAGTTGCAGATTGAAAGGTTTTCTTCTGTATGTACACCGGGTGTT

90841  TCATATGAAATAAAGTTCATCATCATTGATCATTAGAGAAATGCAAATCAAAACCACAAT
90841  AGTATACTTTATTTCAAGTAGTAGTAACTAGTAATCTCTTTACGTTTAGTTTTGGTGTTA

90901  GAGATACCATCTCACACCAGTCAGAATGGTTATTAATAAAAAGTCAAAGAATGACAGTTG
90901  CTCTATGGTAGAGTGTGGTCAGTCTTACCAATAATTATTTTTCAGTTTCTTACTGTCAAC

90961  CTGGTGAGGTTGTAGAGAAAAAGGAACGCTTATACACTGTTGAGGGGATTGTAAATTAGT
90961  GACCACTCCAACATCTCTTTTTCCTTGCGAATATGTGACAACTCCCCTAACATTTAATCA

91021  TCAACCATGTGGAAGACAGTGTGGCAATTCCTCAAAGACCTATATACCATTTGACCCAGC
91021  AGTTGGTACACCTTCTGTCACACCGTTAAGGAGTTTCTGGATATATGGTAAACTGGGTCG

91081  AATCCCATTACTGCATATATACCCAAAGGAATATGAGTCATTCTACCGTAAGACACATGC
91081  TTAGGGTAATGACGTATATATGGGTTTCCTTATACTCAGTAAGATGGCATTCTGTGTACG

91141  ACGCATTTGTTTGTTGCAGCACTATTCACGATAGCAAAAACATGGAATCAACCTACATGC
91141  TGCGTAAACAAACAACGTCGTGATAAGTGCTATCGTTTTTGTACCTTAGTTGGATGTACG

91201  CCATCAATGATAGACTGGATAAAGAAAATGTGGTACATATACCCCATGGAATACTATGCA
91201  GGTAGTTACTATCTGACCTATTTCTTTTACACCATGTATATGGGGTACCTTATGATACGT

91261  GCCATAAGAAAAAGAATGAGATCATGCCCTTTGCAGGAACATGGATGGAGCTGGAGGCCA
91261  CGGTATTCTTTTTCTTACTCTAGTACGGGAAACGTCCTTGTACCTACCTCGACCTCCGGT

91321  TTATCCTTAGCAAACTAACCCAGGAACAGAAAACCAAGTGCAGATGTTCTCACTTTTAAG
91321  AATAGGAATCGTTTGATTGGGTCCTTGTCTTTTGGTTCACGTCTACAAGAGTGAAAATTC

91381  TGTGAGCTGAATGAGGAGAATACATGGGCACATAGAGGGGCCTATCAGAGGGTGGAGGGT
91381  ACACTCGACTTACTCCTCTTATGTACCCGTGTATCTCCCCGGATAGTCTCCCACCTCCCA

91441  GGGAGAAGGAAGAGTATCAGAAAAAATAACAGGTAGTAGATTTAATACCGGATGACAAAA
91441  CCCTCTTCCTTCTCATAGTCTTTTTTATTGTCCATCATCTAAATTATGGCCTACTGTTTT

91501  CAATCTGTATAACAAACCCCCATGACATGAGTTTACCTATATAACAAACCCGCACATGTA
91501  GTTAGACATATTGTTTGGGGGTACTGTACTCAAATGGATATATTGTTTGGGCGTGTACAT

91561  CTCCTGAACTTAAAAGTTAAATTAAAAAAAAATTAATGTCTAATAATATATTACAGTATT
91561  GAGGACTTGAATTTTCAATTTAATTTTTTTTTAATTACAGATTATTATATAATGTCATAA

91621  CTTCATATTCAATGGCAATGTGTGAAGTGGGAAGTGTCTTGACAGAATTCGGTGTTTCAA
91621  GAAGTATAAGTTACCGTTACACACTTCACCCTTCACAGAACTGTCTTAAGCCACAAAGTT

91681  GGCTTACACTTTGATGCCCAAGACTGCACAAGGCTACATTTTCTACTGGTGAGACAAATT
91681  CCGAATGTGAAACTACGGGTTCTGACGTGTTCCGATGTAAAAGATGACCACTCTGTTTAA
```

```
91741  CCAGACGCATTGCATTCAGATCTAATCTCTTAGCTCCTTAATCTTCAGGGTACTGGTAAA
91741  GGTCTGCGTAACGTAAGTCTAGATTAGAGAATCGAGGAATTAGAAGTCCCATGACCATTT

91801  CATGAAGACCTCCCCAGTGCTGTAGTCATCATGATATGTACAGCAGGTGGCTGAGCTCTG
91801  GTACTTCTGGAGGGGTCACGACATCAGTAGTACTATACATGTCGTCCACCGACTCGAGAC

91861  GATGTAGACTGCAGGGATATATTAGGAAGTTAATTCTCAAGGCAAGTCATCTTCAAGCAC
91861  CTACATCTGACGTCCCTATATAATCCTTCAATTAAGAGTTCCGTTCAGTAGAAGTTCGTG

91921  CATATCAGCATGATCAGCAATATAAGTAGTATCTCAGTGCTTTGTTGTTTAGTCAGAGTT
91921  GTATAGTCGTACTAGTCGTTATATTCATCATAGAGTCACGAAACAACAAATCAGTCTCAA

91981  TTGTACTCTATCACCCATTGTAATGTTCCTATTTGCAAAAGGTAATACATACCCTTTAAA
91981  AACATGAGATAGTGGGTAACATTACAAGGATAAACGTTTTCCATTATGTATGGGAAATTT

92041  ACATCTTTGCTTTTTCTCCCATTATCGAGATGCTAGCAGCTTCATAAAGCAGAATAACTA
92041  TGTAGAAACGAAAAAGAGGGTAATAGCTCTACGATCGTCGAAGTATTTCGTCTTATTGAT

92101  AGGGCAAACAGATTATATAAAGGGTTGGAGCTCAATGAAGACAACAAGAACAGCAAAGGT
92101  TCCCGTTTGTCTAATATATTTCCCAACCTCGAGTTACTTCTGTTGTTCTTGTCGTTTCCA

92161  TATTGTAAAACTGGCTGCTTGCAGGCCAACAAGCACATCCATATGGAGGCAATCAGTTTA
92161  ATAACATTTTGACCGACGAACGTCCGGTTGTTCGTGTAGGTATACCTCCGTTAGTCAAAT

92221  TGCTACCTCTGTCTGTTTGATGGGATTCATAATATTGACTTTATCCATTAGATTTGGACT
92221  ACGATGGAGACAGACAAACTACCCTAAGTATTATAACTGAAATAGGTAATCTAAACCTGA

92281  ACCAGGGAATAAAATAAGCAGATGGAGAGTAAGGATTTGCTAGGAAATAATTCAGCCAGT
92281  TGGTCCCTTATTTTATTCGTCTACCTCTCATTCCTAAACGATCCTTTATTAAGTCGGTCA

92341  CACTTTGAAAGCTGTTCAAGAAACAGCTTTCAAAGTGTCTCTCAAACTATGTTTGCCCAT
92341  GTGAAACTTTCGACAAGTTCTTTGTCGAAAGTTTCACAGAGAGTTTGATACAAACGGGTA

92401  TATCCCAATAATTTATTTCCCAATAATTTCATGGGAAAAGAAGGAAGTTCTGTGGTCAGA
92401  ATAGGGTTATTAAATAAAGGGTTATTAAAGTACCCTTTTCTTCCTTCAAGACACCAGTCT

92461  TAAATCTGGAAAACACTGGTTTAAGCAAAGTTCAGTAGGTCTGCTTCCCTGCAGGTCACC
92461  ATTTAGACCTTTTGTGACCAAATTCGTTTCAAGTCATCCAGACGAAGGGACGTCCAGTGG

92521  TCAGAGTCTTTACTCTGCTAACCTAGGAACTCATCCAACAAGTTTAATTTAACAGCTACA
92521  AGTCTCAGAAATGAGACGATTGGATCCTTGAGTAGGTTGTTCAAATTAAATTGTCGATGT

92581  CTGTGTACGTCACTTTAACAGTCACTGAGCTGTGACTCTTGGGGGAAAGATTGTGCGTGT
92581  GACACATGCAGTGAAATTGTCAGTGACTCGACACTGAGAACCCCCTTTCTAACACGCACA

92641  GTGTGTGTGTGTGTACACATGTGTGCACATGTGCAGAATCTACCAAATCTTAAGAGAAAG
92641  CACACACACACACATGTGTACACACGTGTACACGTCTTAGATGGTTTAGAATTCTCTTTC

92701  GAACATGCTGGGAAACTGTCCTGTGAAAGAGAATAGAAACCTGAAGATTTGAGGCAGTGA
92701  CTTGTACGACCCTTTGACAGGACACTTTCTCTTATCTTTGGACTTCTAAACTCCGTCACT

92761  TAGCATTTATGAAAGCAGCAGATAAGGACTAATCACCAAAAGGGGTAGCTCTTTTGTTGG
92761  ATCGTAAATACTTTCGTCGTCTATTCCTGATTAGTGGTTTTCCCCATCGAGAAAACAACC

92821  TTGGGGAAAACAGGAATTTTTCCCCCACCCAATGTGCTGCATTTTCTAATTTTCTATGAA
92821  AACCCCTTTTGTCCTTAAAAAGGGGGTGGGTTACACGACGTAAAAGATTAAAAGATACTT

92881  CACTTCCTAAGAAAAAGCTGAATGAAGAACATTTGCGATGCAATCAGCTCATTAAGAAAC
92881  GTGAAGGATTCTTTTTCGACTTACTTCTTGTAAACGCTACGTTAGTCGAGTAATTCTTTG

92941  ACGCACTTTTGTGGAGATACGTGCTGTCCCAGGAGATGCTCTGCGAGGAGCCGAGTGTTT
92941  TGCGTGAAAACACCTCTATGCACGACAGGGTCCTCTACGAGACGCTCCTCGGCTCACAAA

93001  GGACTGGAGCTGCTGAATGGTTTCTCACAGTTCTAGAATGTTTGGGGCTGCACCCTCTAA
93001  CCTGACCTCGACGACTTACCAAAGAGTGTCAAGATCTTACAAACCCCGACGTGGGAGATT

93061  GATGTTGAACCCATCAGTAATTGCTCCAAACCACTTTATGGGATATAATGCTGTGAGTTG
93061  CTACAACTTGGGTAGTCATTAACGAGGTTTGGTGAAATACCCTATATTACGACACTCAAC

93121  ACACCTGAGGGGATTGTGGTCCTGTTCATGAGTAATTACTTTTCTGTTGCCTATAGAAGG
93121  TGTGGACTCCCCTAACACCAGGACAAGTACTCATTAATGAAAAGACAACGGATATCTTCC
```

93181 GCCAGCAATAGCAGATGAGTAGCTGAACAGTGGTTTTGAGTAATAAAACGTTCTTTTTTA
93181 CGGTCGTTATCGTCTACTCATCGACTTGTCACCAAAACTCATTATTTTGCAAGAAAAAAT

93241 AAAAAAAGTAATGCTTTCTGTTAAACTCTGACTATACTCTCTCCTGGTATCACAACCCAG
93241 TTTTTTTCATTACGAAAGACAATTTGAGACTGATATGAGAGAGGACCATAGTGTTGGGTC

93301 CTTTCTTTTTGCCTTCTTTATTGCAGTTACATATGGGGCTGATGACTTTAGGGATTTCCA
93301 GAAAGAAAAACGGAAGAAATAACGTCAATGTATACCCCGACTACTGAAATCCCTAAAGGT

93361 TGCAATAATTCCCAAATCTTTCTCTCGTAAGTATATGCCTTGCTTCTGGAAAACAAAAGC
93361 ACGTTATTAAGGGTTTAGAAAGAGAGCATTCATATACGGAACGAAGACCTTTTGTTTTCG

93421 ATGCCTTCATCTCCTATCATGTAAATATCGTACGTGCATGTTCCTTCATCAACCCCCGAG
93421 TACGGAAGTAGAGGATAGTACATTTATAGCATGCACGTACAAGGAAGTAGTTGGGGGCTC

93481 ATACATTAAATATTCACTGTTCTATTCGTTAGACACCTACCATATCATTTTTGGGTATTT
93481 TATGTAATTTATAAGTGACAAGATAAGCAATCTGTGGATGGTATAGTAAAAACCCATAAA

93541 ATACCATAAAGTGCAAAACGAAGGTCTAGGCAGTTGTGCGGTGTCCTGGGAGCATGGCAG
93541 TATGGTATTTCACGTTTTGCTTCCAGATCCGTCAACACGCCACAGGACCCTCGTACCGTC

93601 TGGAGTAACAGTAAGGGTTGGAGTCACAGTAGCACTGATGGGATTGTTACTTCGCAGATG
93601 ACCTCATTGTCATTCCCAACCTCAGTGTCATCGTGACTACCCTAACAATGAAGCGTCTAC

93661 CTGCTGGACAGCTTTGAGATTACTCCTGTAATCCTTTTTATTGAGTGTGAGGATAAAATG
93661 GACGACCTGTCGAAACTCTAATGAGGACATTAGGAAAAATAACTCACACTCCTATTTTAC

93721 GCCATATTTTCTGTCACCAACAAAGAACTGAACATGGTTTCAAGATAATTTTTACAATTT
93721 CGGTATAAAAGACAGTGGTTGTTTCTTGACTTGTACCAAAGTTCTATTAAAAATGTTAAA

93781 TGTGGGACAGGAGAAGGAGATATGCTCTTCATCTGTGGTTCTGGAGTATTAATTTCAGAA
93781 ACACCCTGTCCTCTTCCTCTATACGAGAAGTAGACACCAAGACCTCATAATTAAAGTCTT

93841 CTGAGAAGAGAAAATTGCAGGCCTCCAGGCCGTTATTCTGTGGTGCATCGCTGCTGGAGC
93841 GACTCTTCTCTTTTAACGTCCGGAGGTCCGGCAATAAGACACCACGTAGCGACGACCTCG

93901 CAACGTACTGTTCCTGCTTGGTGCCCGCAGGACCTTCTCTGATCTGAGCTGTTAAACCAA
93901 GTTGCATGACAAGGACGAACCACGGGCGTCCTGGAAGAGACTAGACTCGACAATTTGGTT

93961 AAACAGATGTGAATTTCATTTTTCTTTTTTAGATATTGAGCTGATAAATGATAAATATAT
93961 TTTGTCTACACTTAAAGTAAAAAGAAAAAATCTATAACTCGACTATTTACTATTTATATA

94021 TTGGGTAACATCCTCAAATGAGTACTTTAAAGAACAGAGAGTGCTTTTGAAAATGTATAG
94021 AACCCATTGTAGGAGTTTACTCATGAAATTTCTTGTCTCTCACGAAAACTTTTACATATC

94081 AGAGCTGAATATTTAAAGCCATACAAGTTGATAATCCCCTTAGGACCTGTATTAGTAAAT
94081 TCTCGACTTATAAATTTCGGTATGTTCAACTATTAGGGGAATCCTGGACATAATCATTTA

94141 CCTCTAATACTATAAAGCAACATCAGAAAACAGGTCGCATTTGTTAACACTGTGACAATT
94141 GGAGATTATGATATTTCGTTGTAGTCTTTTGTCCAGCGTAAACAATTGTGACACTGTTAA

94201 ATAATTATAGGCTCTAATTTGAACAGTGCCAGTAACAGTTAAAGACAGGTGTCTGACATC
94201 TATTAATATCCGAGATTAAACTTGTCACGGTCATTGTCAATTTCTGTCCACAGACTGTAG

94261 CTGGTTATCAAAAACTATCTGGTGTCTAAGGAATTTAACAGTATCAAAGTACATCTTTTT
94261 GACCAATAGTTTTTGATAGACCACAGATTCCTTAAATTGTCATAGTTTCATGTAGAAAAA

94321 GCTCACAAACTAGATTGGCTACTTTTCCAGCCTATAGAAGTTACAGAAATCTTTCTTTTA
94321 CGAGTGTTTGATCTAACCGATGAAAAGGTCGGATATCTTCAATGTCTTTAGAAAGAAAAT

94381 CTGATATTCCTTTGATTCCACTTTAAAGCAATAGCTTGATACCTACTTTTTGAGATGTAT
94381 GACTATAAGGAAACTAAGGTGAAATTTCGTTATCGAACTATGGATGAAAAACTCTACATA

94441 ATGTATGTACTTGTAAATGCTTATATATGTGTATTTGTTTACACGTGAGCTTATATGAAC
94441 TACATACATGAACATTTACGAATATATACACATAAACAAATGTGCACTCGAATATACTTG

94501 ATATAAAACATTATGGTATTGGACAGAGAAATGATATTGATTTGAAGAATCATATTGTTG
94501 TATATTTTGTAATACCATAACCTGTCTCTTTACTATAACTAAACTTCTTAGTATAACAAC

94561 AAGTATTTTGAGAAGTCAAATGCACTTGAGAGTAAGCTAATGCATCCTAAAACATGTTTC
94561 TTCATAAAACTCTTCAGTTTACGTGAACTCTCATTCGATTACGTAGGATTTTGTACAAAG

94621 TGTGAAGTCTGAGGAAGGTGTCAGCCAGAGCATATGTTGAAGCAGACCCTTCAGTGAAGC

94621 ACACTTCAGACTCCTTCCACAGTCGGTCTCGTATACAACTTCGTCTGGGAAGTCACTTCG

94681 CTTCAGTCTGTAAGAATCGCTGCATGTGAAGATGTGTTAATGTTATGATAGTTACAGTTT
94681 GAAGTCAGACATTCTTAGCGACGTACACTTCTACACAATTACAATACTATCAATGTCAAA

94741 TTATAAAAGAGATGATATACACTTGGATATGCTTTCTGTCTATATTTATGCAAATGTGTC
94741 AATATTTTCTCTACTATATGTGAACCTATACGAAAGACAGATATAAATACGTTTACACAG

94801 CATAAGGTATTGGTGTCTCTCTTTCTCTCCCACTCTCCCCAGTGTTGGAATTGTGACTAT
94801 GTATTCCATAACCACAGAGAGAAAGAGAGGGTGAGAGGGGTCACAACCTTAACACTGATA

94861 CTTCTCACACAAGCGGCTACTTGGTCTTGATGCCTTCCCCCGCAAAACAGCAACCAAACT
94861 GAAGAGTGTGTTCGCCGATGAACCAGAACTACGGAAGGGGGCGTTTTGTCGTTGGTTTGA

94921 GTTCTGGGCCAATATCACCACCTTGTGGTCATGATGAAGAATTGCCCCCTTTGCCCTCAA
94921 CAAGACCCGGTTATAGTGGTGGAACACCAGTACTACTTCTTAACGGGGGAAACGGGAGTT

94981 CACCTCTTTTCTTCTTGAAAATTAAAAACAACCCCTTTCACCCCCTCTACTGTCCTTATT
94981 GTGGAGAAAAGAAGAACTTTTAATTTTTGTTGGGGAAAGTGGGGGAGATGACAGGAATAA

95041 CCAGTTTGTGTCCGTAGTTGCTGGAGGAAAGAAAATGCCTATTGCTCTTTTTTTTATTCT
95041 GGTCAAACACAGGCATCAACGACCTCCTTTCTTTTACGGATAACGAGAAAAAAAATAAGA

95101 CTTTATCTGTTTACTCTTCTGCCTTTCTTTTCTCTCCCTGACTTCATTCCATTTAAAGGC
95101 GAAATAGACAAATGAGAAGACGGAAAGAAAAGAGAGGGACTGAAGTAAGGTAAATTTCCG

95161 CTTAAATTGCAAATAACGAAAAAAGTATTTTCTAATAATATACCCTAGTGGGACGAAAAC
95161 GAATTTAACGTTTATTGCTTTTTTCATAAAAGATTATTATATGGGATCACCCTGCTTTTG

95221 AACCTTCTACATTTTAAATGATATTAAGTTATAATAAATGGTTCCATGAACTTTAAAACG
95221 TTGGAAGATGTAAAATTTACTATAATTCAATATTATTTACCAAGGTACTTGAAATTTTGC

95281 CTTAAAAGTTTTATAAATCTCTTTTAGAGGCAACACAAAAATATCTATATATATATCTTC
95281 GAATTTTCAAAATATTTAGAGAAAATCTCCGTTGTGTTTTTATAGATATATATATAGAAG

95341 TCTTCACCTCAAATTTTCATTTAATTACCGTGATTCAAGATATATCTTAAATATTTCCAT
95341 AGAAGTGGAGTTTAAAAGTAAATTAATGGCACTAAGTTCTATATAGAATTTATAAAGGTA

95401 TTCATTTGATTTTTACTTGCTAGTCACCAGTGTTGCAAGAATATTTCTGTACCATTTTTC
95401 AAGTAAACTAAAAATGAACGATCAGTGGTCACAACGTTCTTATAAAGACATGGTAAAAAG

95461 AAGTAAAATATTGAGCATTTCAAGCTTACATGGCCTAAACTCACACTAGAGTGACTACTT
95461 TTCATTTTATAACTCGTAAAGTTCGAATGTACCGGATTTGAGTGTGATCTCACTGATGAA

95521 CGGACTATCTTTTCAATGGAAAAATGCGTTCTGAGTACCAAACATTTACTTTTTCAGTAA
95521 GCCTGATAGAAAAGTTACCTTTTTACGCAAGACTCATGGTTTGTAAATGAAAAAGTCATT

95581 ATATGAAGACACATTCTAAAAAGGAAAGGGAAATGAAGAGGAAGCTTAATTCATTGGTCA
95581 TATACTTCTGTGTAAGATTTTTCCTTTCCCTTTACTTCTCCTTCGAATTAAGTAACCAGT

95641 GGTGGGGACTAACAGTGTACATTATACCTATGCATTAAAAAATGTTTAATTATCCATGAG
95641 CCACCCCTGATTGTCACATGTAATATGGATACGTAATTTTTTACAAATTAATAGGTACTC

95701 CACATACAGTGCCTGGCACAGATTCAGGCCTCTTGGCTCCTCTTGGCATCTGAGTTCCTT
95701 GTGTATGTCACGGACCGTGTCTAAGTCCGGAGAACCGAGGAGAACCGTAGACTCAAGGAA

95761 GGATTCTCCAGATTCTTCAATCCAAACTTGTCTATCTACCAAAGTCCACCTTCCATACAG
95761 CCTAAGAGGTCTAAGAAGTTAGGTTTGAACAGATAGATGGTTTCAGGTGGAAGGTATGTC

95821 GGGTGTGGGGTCTCTCTTAGTACTGTTGGGTGCCAATAACATAGAAGCCCTCCAAGATAG
95821 CCCACACCCCAGAGAGAATCATGACAACCCACGGTTATTGTATCTTCGGGAGGTTCTATC

95881 GTCAATGAGATTTTAATTTTTCCCAACTTTAAATAGCATCTGGGAAGGAAAGCAGCTTTT
95881 CAGTTACTCTAAAATTAAAAAGGGTTGAAATTTATCGTAGACCCTTCCTTTCGTCGAAAA

95941 GACTAGGAGTTTGAAAACATCCACTCTTCATATTTATTGTCACATTTATTGAGTAATAAA
95941 CTGATCCTCAAACTTTTGTAGGTGAGAAGTATAAATAACAGTGTAAATAACTCATTATTT

96001 ATTGGGCTCTTGTCCATTATCTCTCAGTCAATGAAAGAGCCCACGGTGGGCATGAGCCAG
96001 TAACCCGAGAACAGGTAATAGAGAGTCAGTTACTTTCTCGGGTGCCACCCGTACTCGGTC

96061 AGTGTTCCGGAGAAAGAGAGGACCAGTCTCCATTTCATGTTCTATATTTAATCAGTTCAA
96061 TCACAAGGCCTCTTTCTCTCCTGGTCAGAGGTAAAGTACAAGATATAAATTAGTCAAGTT

96121 TTCAGCACCTTACTGAGTTTGACATTTTTCTAAGTAATGGGGTGGGGGGTTACAAAATAT
96121 AAGTCGTGGAATGACTCAAACTGTAAAAAGATTCATTACCCCACCCCCCAATGTTTTATA

96181 TTATAACTATCTTTCCTATCTTCAGGGAGATGTACCATTTCACTGAAGCAATACCACAGC
96181 AATATTGATAGAAAGGATAGAAGTCCCTCTACATGGTAAAGTGACTTCGTTATGGTGTCG

96241 TAAGCAGTTCCAGCTCCGAGTAACAGAGGAAGGAGTGAGCGATTTAGGAACTCACAGAAG
96241 ATTCGTCAAGGTCGAGGCTCATTGTCTCCTTCCTCACTCGCTAAATCCTTGAGTGTCTTC

96301 GGCACGGCCAGGCAACTCGGGCCCTGGCAAATGTTCCATAGAGACAGCATCCATGGGGTT
96301 CCGTGCCGGTCCGTTGAGCCCGGGACCGTTTACAAGGTATCTCTGTCGTAGGTACCCCAA

96361 GAACCAAGCACAATGCCTGCAAATCGGGCAGGCGGGTTCATGGAATTGATTCTAGGGCAG
96361 CTTGGTTCGTGTTACGGACGTTTAGCCCGTCCGCCCAAGTACCTTAACTAAGATCCCGTC

96421 GTATGATCAGCACCATTTAAGGAGTGGCAGTTGTTAAGAAACTGAAGGGATTTCTCATAT
96421 CATACTAGTCGTGGTAAATTCCTCACCGTCAACAATTCTTTGACTTCCCTAAAGAGTATA

96481 GGCATTTTCAACAGAATTCAGCAGGAGCCTTGCATATCCCTTGGGATGGGTTCCGGGGTA
96481 CCGTAAAAGTTGTCTTAAGTCGTCCTCGGAACGTATAGGGAACCCTACCCAAGGCCCCAT

96541 TTATGGAGCTGCCAGGAGGTGCTCAGCTTCAGTAAGTGCCAGTCATTCCTTCCCAACCTC
96541 AATACCTCGACGGTCCTCCACGAGTCGAAGTCATTCACGGTCAGTAAGGAAGGGTTGGAG

96601 CTCCTTCATTAGTGGAAGTATCAGCTGGTGTCATTGCCTTCTTCAGGAGGATGGATTTCA
96601 GAGGAAGTAATCACCTTCATAGTCGACCACAGTAACGGAAGAAGTCCTCCTACCTAAAGT

96661 AAATGAAGGGCCAACAAAACTCACATTCTCGCAGAGCCTCCGTCACAACTGCATGTGTTC
96661 TTTACTTCCCGGTTGTTTTGAGTGTAAGAGCGTCTCGGAGGCAGTGTTGACGTACACAAG

96721 ACTGCCATCAACAGGCCAACACCCACTTGTTCTTTTCTTCTAGGTAAAGGAAAAGGTCTG
96721 TGACGGTAGTTGTCCGGTTGTGGGTGAACAAGAAAAGAAGATCCATTTCCTTTTCCAGAC

96781 TGGGACAAAGGCCCTGAACTCCTCACTCCTTGAATGGGAGGCATATTGGTGGAGGCTGTC
96781 ACCCTGTTTCCGGGACTTGAGGAGTGAGGAACTTACCCTCCGTATAACCACCTCCGACAG

96841 ACACAACTACAGGTGGCAGGAACCTGCCATCAGGGTTCCCATCAGCCTCCAGGGCTTGCT
96841 TGTGTTGATGTCCACCGTCCTTGGACGGTAGTCCCAAGGGTAGTCGGAGGTCCCGAACGA

96901 TCTCTGCCCTGTGTGTGGACACCTTCCATCACTGGAAACCCCTCTCTCAGAAGAACCACT
96901 AGAGACGGGACACACACCTGTGGAAGGTAGTGACCTTTGGGGAGAGAGTCTTCTTGGTGA

96961 TCTTGCCCTCATCTCCCACTTGCCTCTCCTCGGCACCTTCTCCACCCTCAGGGGCCAAGC
96961 AGAACGGGAGTAGAGGGTGAACGGAGAGGAGCCGTGGAAGAGGTGGGAGTCCCCGGTTCG

97021 TGAGGAAGTGTGCATCCTGTGCCACCAACCTTGCTCGCAGGAATTAGTTTCAGTCCCCAG
97021 ACTCCTTCACACGTAGGACACGGTGGTTGGAACGAGCGTCCTTAATCAAAGTCAGGGGTC

97081 TGACTGACTCTTCCAAACCTGCACCAGCCTATGCTGCCCTTTTGAAAGGAGGCGTGCGTA
97081 ACTGACTGAGAAGGTTTGGACGTGGTCGGATACGACGGGAAAACTTTCCTCCGCACGCAT

97141 CAGCCTGGCTGGTACAGAAGATTCCAGATTTCAGTGACTACAGGACGCTGTCACCCCTTC
97141 GTCGGACCGACCATGTCTTCTAAGGTCTAAAGTCACTGATGTCCTGCGACAGTGGGGAAG

97201 TGTTTTTCTTTTTTCTGTCCACATCTCTGATTGAGCATACCTTCAGGGAAAACCCCAGAA
97201 ACAAAAAGAAAAAAGACAGGTGTAGAGACTAACTCGTATGGAAGTCCCTTTTGGGGTCTT

97261 ACCTTTTTCACACTTACATGGGAAGCCAAGACACTCTCATCCTGGGCTTGTGTTTTTAAA
97261 TGGAAAAAGTGTGAATGTACCCTTCGGTTCTGTGAGAGTAGGACCCGAACACAAAAATTT

97321 ATCTATATTTTAATTTCACAGGTCCTTACATTTATGACTAATAGATTTCAGATTTTGAGA
97321 TAGATATAAAATTAAAGTGTCCAGGAATGTAAATACTGATTATCTAAAGTCTAAAACTCT

97381 TACAATCTGTAGTTGCAACTTATGAAGATAGTTTCAAGCCCTGATTTCTGTCATCTATGA
97381 ATGTTAGACATCAACGTTGAATACTTCTATCAAAGTTCGGGACTAAAGACAGTAGATACT

97441 AAACAGTAAGCAAGTTGCTTGTATTCTCCTCCTAGCTGGATAAGAGACCCCCGTTTTCAT
97441 TTTGTCATTCGTTCAACGAACATAAGAGGAGGATCGACCTATTCTCTGGGGGCAAAAGTA

97501 GGACAGCCTCCTTGGGTGGACAGAAGTCTCCTGATCTTGTCTATGTAACCAGCACCCACT
97501 CCTGTCGGAGGAACCCACCTGTCTTCAGAGGACTAGAACAGATACATTGGTCGTGGGTGA

97561 TTGCATTTTTCCGCAAAGAAAAGAGGAGCTTAACCCCATTTACGGGAAATCCACGCACTG
97561 AACGTAAAAAGGCGTTTCTTTTCTCCTCGAATTGGGGTAAATGCCCTTTAGGTGCGTGAC

97621 CTTGAATCTTACGCGCCCCGTGGTGACGGTGTTGCCCACTCTGTTCCCTTCCCACCAAAT
97621 GAACTTAGAATGCGCGGGGCACCACTGCCACAACGGGTGAGACAAGGGAAGGGTGGTTTA

97681 CCTGCCAGCACCTGGCAGTCATCCAGTCCTTGTCTTAGACTTCACCAACCTTTCCCGTCA
97681 GGACGGTCGTGGACCGTCAGTAGGTCAGGAACAGAATCTGAAGTGGTTGGAAAGGGCAGT

97741 TAACATGGCTTAGAAGTTGTCCGAATAGTGCTAGAAGCTTTCAGCCCTGGAATGCTCAGC
97741 ATTGTACCGAATCTTCAACAGGCTTATCACGATCTTCGAAAGTCGGGACCTTACGAGTCG

97801 ATCTGCCATCTAACTTCCATTTAAAACGTTTTGAGCATTTGCTGATACAGGTTCCTCTGT
97801 TAGACGGTAGATTGAAGGTAAATTTTGCAAAACTCGTAAACGACTATGTCCAAGGAGACA

97861 CCTTTCCCCATTATTTTCAGTTCCTTGGCCTTGCCCATAGTGATCTGTGGATATTGTCTG
97861 GGAAAGGGGTAATAAAAGTCAAGGAACCGGAACGGGTATCACTAGACACCTATAACAGAC

97921 CATACTCTTTTTTAGAAACCTAGAGCATCGCAGTGTCCTTAATTTTCCTCTTCAAACTCT
97921 GTATGAGAAAAAATCTTTGGATCTCGTAGCGTCACAGGAATTAAAAGGAGAAGTTTGAGA

97981 ATATATTCCTGAGCTGATCATCCCTACTGCTGTTCCTACAGAACTGTGTGACTAATCTTC
97981 TATATAAGGACTCGACTAGTAGGGATGACGACAAGGATGTCTTGACACACTGATTAGAAG

98041 TATATATGTCCCTACATCTAGCCATCTAGGTGGAGATCTAGGTACCTATCTCCCTAAAAT
98041 ATATATACAGGGATGTAGATCGGTAGATCCACCTCTAGATCCATGGATAGAGGGATTTTA

98101 TGTACTTGGTATCCACATTGACCTCAGGACAGTATGTGATAGGCTCTTGTGCCTTTTCGA
98101 ACATGAACCATAGGTGTAACTGGAGTCCTGTCATACACTATCCGAGAACACGGAAAAGCT

98161 CTAGGAGTTTGAAAACATCCACTCTTCATATTTATTGTCACATTTGTTGAATAATAAAAT
98161 GATCCTCAAACTTTTGTAGGTGAGAAGTATAAATAACAGTGTAAACAACTTATTATTTTA

98221 TGGGCTCTTGTCCATTATGTCTCAGTCAGTGAAAGAGCCCAGGGTGGGCCTGAGCCAGAG
98221 ACCCGAGAACAGGTAATACAGAGTCAGTCACTTTCTCGGGTCCCACCCGGACTCGGTCTC

98281 TGTTTGTGTCCCTAGTTCTCTTTTGGTTCTGTTACAGCTCTCAGTGACAGTGTCATTGAG
98281 ACAAACACAGGGATCAAGAGAAAACCAAGACAATGTCGAGAGTCACTGTCACAGTAACTC

98341 GCTAACAAGGACAGAGTACTGCTTTCAGCCACCATTTGTCCAATGAGTGGCTGATCTCCA
98341 CGATTGTTCCTGTCTCATGACGAAAGTCGGTGGTAAACAGGTTACTCACCGACTAGAGGT

98401 GGCCTCTGGCTTTGAGAACATCTGTGATATCTAAGGCAGCATCCATTGTGGGCTTTCCCC
98401 CCGGAGACCGAAACTCTTGTAGACACTATAGATTCCGTCGTAGGTAACACCCGAAAGGGG

98461 CATGCTTCTGTTTCCTTCCTGTCACATAGCTTTGCCTCCTCCTGCAAGCAGCCTGCTGTA
98461 GTACGAAGACAAAGGAAGGACAGTGTATCGAAACGGAGGAGGACGTTCGTCGGACGACAT

98521 GCAGAACCGGTGTTCCTGAAGCCAGAAACCCAAAGGTCGTGTCCAATGCTTCCCTGCTGT
98521 CGTCTTGGCCACAAGGACTTCGGTCTTTGGGTTTCCAGCACAGGTTACGAAGGGACGACA

98581 TCTGCTCCCCACCTGCAAGCGCCCACACACTGATCAACAGCACACGCCATAGAGCATGCC
98581 AGACGAGGGGTGGACGTTCGCGGGTGTGTGACTAGTTGTCGTGTGCGGTATCTCGTACGG

98641 AAAGAATCCCAGAAATACTCATTCTTAATGATCCAGAAGAACAAGTGATAGCTTCTGGCA
98641 TTTCTTAGGGTCTTTATGAGTAAGAATTACTAGGTCTTCTTGTTCACTATCGAAGACCGT

98701 AGTTCTAATAGGCCATATGTGCCCATGGGAAAGCAAGAGTCAATTCTCTCTAGCTGGTTG
98701 TCAAGATTATCCGGTATACACGGGTACCCTTTCGTTCTCAGTTAAGAGAGATCGACCAAC

98761 CCTCTGAATTTATGAAGTCAAGCCGCGTAGGGGAACATGCCAAGAGGATTATCCCAATGC
98761 GGAGACTTAAATACTTCAGTTCGGCGCATCCCCTTGTACGGTTCTCCTAATAGGGTTACG

98821 CTGTTCTGCGTGGTGTCCTTTTTCTAGCCTCGGAAGCACTAGGCCTACACCTTCGCTATA
98821 GACAAGACGCACCACAGGAAAAAGATCGGAGCCTTCGTGATCCGGATGTGGAAGCGATAT

98881 GCCCTCGTCTTAGCGCCTGTATTTGAAAAACTTGCCACACCAATCTTAATTTGCTCCATT
98881 CGGGAGCAGAATCGCGGACATAAACTTTTTGAACGGTGTGGTTAGAATTAAACGAGGTAA

98941 GTGTTCTAATCTCATTTTAAAAACCACAAGGTAAATGATTAAAAATAATCTTAGATTAAG
98941 CACAAGATTAGAGTAAAATTTTTGGTGTTCCATTTACTAATTTTTATTAGAATCTAATTC

99001 GAGTACACAGATCTTTGCAGCCTCATTGTGTTTCCAGAGCGGGCTAGTGTAGACACTGGT

```
 99001  CTCATGTGTCTAGAAACGTCGGAGTAACACAAAGGTCTCGCCCGATCACATCTGTGACCA

 99061  GAACAAGGAGGGCCTAGAGAATTAGCTTTCTCCCAGAAAGATGCACAGCCTCTACCTGAG
 99061  CTTGTTCCTCCCCGGATCTCTTAATCGAAAGAGGGTCTTTCTACGTGTCGGAGATGGACTC

 99121  AGTACCAATTCCAGAGAACTCAATGGTCATTAAGCACCATTGCCTCGACAGCCAGCCAGC
 99121  TCATGGTTAAGGTCTCTTGAGTTACCAGTAATTCGTGGTAACGGAGCTGTCGGTCGGTCG

 99181  CTCACTTGCCTGCCTATCTCCTTTATTTTCCAAGTATCTTGTCCCTGGCAGTGGGGAGAG
 99181  GAGTGAACGGACGGATAGAGGAAATAAAAGGTTCATAGAACAGGGACCGTCACCCCTCTC

 99241  GTTAGCAGGAGGCTGCTCAGATGCTCTCGGTCTCTGATCTTCAGGATCTGAAGGGGAGAG
 99241  CAATCGTCCTCCGACGAGTCTACGAGAGCCAGAGACTAGAAGTCCTAGACTTCCCCTCTC

 99301  CATTTGAAGAATCCCCATTGCTGGATTTCTCAGGACAATCCTGCATAATGCCAGGATCTG
 99301  GTAAACTTCTTAGGGGTAACGACCTAAAGAGTCCTGTTAGGACGTATTACGGTCCTAGAC

 99361  ATGGAGGAGACAGGCAGCTCTGTTAATCCTCTGGTGCATCCTCACTTCTGTGGTCTCCAA
 99361  TACCTCCTCTGTCCGTCGAGACAATTAGGAGACCACGTAGGAGTGAAGACACCAGAGGTT

 99421  GTCCACCATGTCCCAGTTAATTCATTTCATTATTCATCTGTAAGTCATTTCCCTAAAGAG
 99421  CAGGTGGTACAGGGTCAATTAAGTAAAGTAATAAGTAGACATTCAGTAAAGGGATTTCTC

 99481  CTAATAAGAAAACACTGGCAGTACAATCAGCCCTCCATATCCATGGGTTCAGCCTCTGGA
 99481  GATTATTCTTTTGTGACCGTCATGTTAGTCGGGAGGTATAGGTACCCAAGTCGGAGACCT

 99541  AATTCAACCCACTGTGGCTTGAAAATACAGTATTCAAGAGACATGGAACCTGTGGATATA
 99541  TTAAGTTGGGTGACACCGAACTTTTATGTCATAAGTTCTCTGTACCTTGGACACCTATAT

 99601  GAAGGCAGATTTTTCGTATCCATAGGTTCTATAGGGCCATTTTAGGGACTTGCACATCTA
 99601  CTTCCGTCTAAAAAGCATAGGTATCCAAGATATCCCGGTAAAATCCCTGAACGTGTAGAT

 99661  CAGACCTTCAGGAGTCTCAGAACTGATGCACTGCAGATACCAAGAGATGACTGTAAGTGG
 99661  GTCTGGAAGTCCTCAGAGTCTTGACTACGTGACGTCTATGGTTCTCTACTGACATTCACC

 99721  TTAGGAATTTGGATGCTGGGGCCAGGCTGCCTGCAGTTACCTTCCAGCTCTGCCACTTGC
 99721  AATCCTTAAACCTACGACCCCGGTCCGACGGACGTCAATGGAAGGTCGAGACGGTGAACG

 99781  CAGCTATGTGACCTTAGCAAGTTGTTCAACCTCTCTGTGCCTTGGCTTCTTCAACTGTAA
 99781  GTCGATACACTGGAATCGTTCAACAAGTTGGAGAGACACGGAACCGAAGAAGTTGACATT

 99841  AATAGGATAATGATAGCACTTCCCTTATGGAGTCCTTGTGAAGGTTAAATGGCAGAGTAC
 99841  TTATCCTATTACTATCGTGAAGGGAATACCTCAGGAACACTTCCAATTTACCGTCTCATG

 99901  AATTAATGTGCTGTGTGCCCAGCGTGTGGTATTGGGGTTAGGTGAAAGACTGTACTGGGA
 99901  TTAATTACACGACACACGGGTCGCACACCATAACCCCAATCCACTTTCTGACATGACCCT

 99961  TCACTGGGGGTGAGCCTCCATGTGCCATGCCCCAACCACTATCCTCAGCTAACTTCTTGT
 99961  AGTGACCCCCACTCGGAGGTACACGGTACGGGGTTGGTGATAGGAGTCGATTGAAGAACA

100021  CTGTCAGTGTTAGGCTGGTGCTATAATAATTTCCATTTTCATGGATGAGGAAATCAAGGC
100021  GACAGTCACAATCCGACCACGATATTATTAAAGGTAAAAGTACCTACTCCTTTAGTTCCG

100081  ACAGAGAAGTTACATGACTTGCCTAAGATCTCAGTGCTTTTAAATAGTGCAGCTAAGATT
100081  TGTCTCTTCAATGTACTGAACGGATTCTAGAGTCACGAAAATTTATCACGTCGATTCTAA

100141  CCAGACCAGGTATTTTTATTTCAGTGTCTGGACTGTAGATCTCTAAATCGAGAGGAACTC
100141  GGTCTGGTCCATAAAAATAAAGTCACAGACCTGACATCTAGAGATTTAGCTCTCCTTGAG

100201  CCTGAATAAAATAAGCTTGGAGTGCTGTTAACTAAGTTGGTTATTTAAGACAGTTGTTCT
100201  GGACTTATTTTATTCGAACCTCACGACAATTGATTCAACCAATAAATTCTGTCAACAAGA

100261  CAGTCCTCAGTGTACATTAGAATCTTCTGGAGGGTATGTTAAAACAGACTGCTGGGCCCA
100261  GTCAGGAGTCACATGTAATCTTAGAAGACCTCCCATACAATTTTGTCTGACGACCCGGGT

100321  CCGAGAGTTCCTGAGCCTGCAACTCTGGGTTGAGCCTGAGAGTCTGTGCTTCTAGTAAGT
100321  GGCTCTCAAGGACTCGGACGTTGAGACCCAACTCGGACTCTCAGACACGAAGATCATTCA

100381  TCTCAGGTGATGCTTATGCTGCTTCTCCCAAATTTTGAGAACCATTGATTGAAAATGTTG
100381  AGAGTCCACTACGAATACGACGAAGAGGGTTTAAAACTCTTGGTAACTAACTTTTACAAC

100441  ATCAAAAATTATGTGGTCTAGGCTGAGCACAGTGGCTCACAACTGTAATCTCAACACTTT
100441  TAGTTTTTAATACACCAGATCCGACTCGTGTCACCGAGTGTTGACATTAGAGTTGTGAAA
```

```
100501  GGGAGGCCAAGTCAGGTGGATCACTGAGCTCAGGAGTTCGAGAGCAGCCTGGTCAACATG
100501  CCCTCCGGTTCAGTCCACCTAGTGACTCGAGTCCTCAAGCTCTCGTCGGACCAGTTGTAC

100561  ACAAGACCCATCTCTACAAAAAACACCAAAAAAAAATGGCTGGATGTGGTGGTGCACACC
100561  TGTTCTGGGTAGAGATGTTTTTTGTGGTTTTTTTTTACCGACCTACACCACCACGTGTGG

100621  TGTAGTCACAACTACTGGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCGGGAGACAGAG
100621  ACATCAGTGTTGATGACCCCTCCGACTCCACCCTCCTAGTGAACTCGGGCCCTCTGTCTC

100681  CAAGACCCTGTCTCAAAAAAGAAACATTTAGTCTTGATTGTCATCTATCTCATTGATCAT
100681  GTTCTGGGACAGAGTTTTTTCTTTGTAAATCAGAACTAACAGTAGATAGAGTAACTAGTA

100741  TTTACTTGGCAAAATTTACCTACTTACTCTTATTAGTCTGTAAAAATGGTTATTAAAATG
100741  AAATGAACCGTTTTAAATGGATGAATGAGAATAATCAGACATTTTTACCAATAATTTTAC

100801  GTGGCTTTCAGTGTAACTACAAATTCTCTTAGTCATTATAGTGTTGGATTCACCAATGTA
100801  CACCGAAAGTCACATTGATGTTTAAGAGAATCAGTAATATCACAACCTAAGTGGTTACAT

100861  TCATCAGTCAGTGTTTCTGAAGTGTGTAAATGAAATAACTCCTACTTTCTTTAGGAAGAA
100861  AGTAGTCAGTCACAAAGACTTCACACATTTACTTTATTGAGGATGAAAGAAATCCTTCTT

100921  AAATATTCAAATGACATAATTACCTTGTGATGTGTGACTTAAAGGTAACAGTAATGCAGA
100921  TTTATAAGTTTACTGTATTAATGGAACACTACACACTGAATTTCCATTGTCATTACGTCT

100981  GTCAATAGTGGTCATTGTATTCAAGACACTAAAAGTTCACCTCCCACCCCCCCCCCCCAC
100981  CAGTTATCACCAGTAACATAAGTTCTGTGATTTTCAAGTGGAGGGTGGGGGGGGGGGGTG

101041  CCACCACCCAACAAACACACAAGCTTCTTTCCCTTTGGAAAAAAAAAGCTCTTCCAGATAC
101041  GGTGGTGGGTTGTTTGTGTGTTCGAAGAAAGGGAAACCTTTTTTTTCGAGAAGGTCTATG

101101  CTACATTCATAAACTATCCCAATTAACCCTTCAGCAAGTGGAAGAAGTGTAAGAAAGGAT
101101  GATGTAAGTATTTGATAGGGTTAATTGGGAAGTCGTTCACCTTCTTCACATTCTTTCCTA

101161  ACCTCTTCTTTAGAACACAGGGTTTGTTTTTATGTTATTTAAGATAAACAGGAATTCAAA
101161  TGGAGAAGAAATCTTGTGTCCCAAACAAAAATACAATAAATTCTATTTGTCCTTAAGTTT

101221  TGGTCATGTACCAAAGCAACACAAAGAACTTCCGGAAATCTGAAAGGGAACTGTGGTCAG
101221  ACCAGTACATGGTTTCGTTGTGTTTCTTGAAGGCCTTTAGACTTTCCCTTGACACCAGTC

101281  AACTCTGAGCATTTTTATGTTTACTGAGTTTTGTCCCAAAGTTTATTAATGTTTACATGC
101281  TTGAGACTCGTAAAAATACAAATGACTCAAAACAGGGTTTCAAATAATTACAAATGTACG

101341  CACAAGGAAAGGTAGCATCACAATAAGAGACGTTTTTCAGGCTTGATAACCACTTATTAG
101341  GTGTTCCTTTCCATCGTAGTGTTATTCTCTGCAAAAAGTCCGAACTATTGGTGAATAATC

101401  GTATTTTGCCAAACAAGTTCACACATCCTAGAGAGCTGGATTGTGTGACCCAGAACCCAC
101401  CATAAAACGGTTTGTTCAAGTGTGTAGGATCTCTCGACCTAACACACTGGGTCTTGGGTG

101461  CCTCTAGGGCAAGGTGCCCATCTGATGGGTAGGGTGTAGGAGTAGGCCTCAGACCACTCC
101461  GGAGATCCCGTTCCACGGGTAGACTACCCATCCCACATCCTCATCCGGAGTCTGGTGAGG

101521  TGACGTGAACCTGCTTAAAGTGAGGGCCCAATTCTAAAGTGGGAACTATGTAAATACCTT
101521  ACTGCACTTGGACGAATTTCACTCCCGGGTTAAGATTTCACCCTTGATACATTTATGGAA

101581  TCTAGTGCATTTTCAGATAATCGCCACTGGGCCTATGGATGGAGAGGGCTGGCAGATCTC
101581  AGATCACGTAAAAGTCTATTAGCGGTGACCCGGATACCTACCTCTCCCGACCGTCTAGAG

101641  CCTGTAACCCCAGGTGCATCCCGAGGCCTGCCACCGAAGCCCACTCAAGGCTGAATGCAC
101641  GGACATTGGGGTCCACGTAGGGCTCCGGACGGTGGCTTCGGGTGAGTTCCGACTTACGTG

101701  GGCGAGCTCAGGCTGCTCTCCCCTTGGTATTTGCTAAGAACTTCTGTTTAGTAGCTCTCC
101701  CCGCTCGAGTCCGACGAGAGGGGAACCATAAACGATTCTTGAAGACAAATCATCGAGAGG

101761  ACACCTATTTGATTGTCTTTTTGCTGCTGTGTTGTTTTGTTGAGTTTTTTTTTTGCAATG
101761  TGTGGATAAACTAACAGAAAAACGACGACACAACAAAACAACTCAAAAAAAAAACGTTAC

101821  ACACTGAGTGGCCTCCTGTATTGTTTCTTTCAGCCAGTAATGTTAAAGTAGAGACTCAGA
101821  TGTGACTCACCGGAGGACATAACAAAGAAAGTCGGTCATTACAATTTCATCTCTGAGTCT

101881  GTGATGAAGAGAATGGGCGTGCCTGTGAAATGAATGGGGAAGAATGTGCGGAGGATTTAC
101881  CACTACTTCTCTTACCCGCACGGACACTTTACTTACCCCTTCTTACACGCCTCCTAAATG
```

```
101941  GAATGCTTGATGCCTCGGGAGAGAAAATGAATGGCTCCCACAGGGACCAAGGCAGCTCGG
101941  CTTACGAACTACGGAGCCCTCTCTTTTACTTACCGAGGGTGTCCCTGGTTCCGTCGAGCC

102001  CTTTGTCGGGAGTTGGAGGCATTCGACTTCCTAACGGAAAACTAAAGTGTGATATCTGTG
102001  GAAACAGCCCTCAACCTCCGTAAGCTGAAGGATTGCCTTTTGATTTCACACTATAGACAC

102061  GGATCATTTGCATCGGGCCCAATGTGCTCATGGTTCACAAAAGAAGCCACACTGGTAAGG
102061  CCTAGTAAACGTAGCCCGGGTTACACGAGTACCAAGTGTTTTCTTCGGTGTGACCATTCC

102121  CCTGGCTCAGTTTTTCCTTTAGTGGCCTGGAGAAGGTGCATGGGGTTTGAAGGAGGAAAG
102121  GGACCGAGTCAAAAAGGAAATCACCGGACCTCTTCCACGTACCCCAAACTTCCTCCTTTC

102181  CATCCTGTCTTCCTTGTGTTCTGAGCATGTTTCTAATTGACTGGTAGCTCAGTTGTTGCA
102181  GTAGGACAGAAGGAACACAAGACTCGTACAAAGATTAACTGACCATCGAGTCAACAACGT

102241  AGCGATTGGTTCCAAGTGGTACCGAGTCATAGAGTCCTTGTTCTGGTACAGCCTTGTAAA
102241  TCGCTAACCAAGGTTCACCATGGCTCAGTATCTCAGGAACAAGACCATGTCGGAACATTT

102301  GGACTTCTCAACACGTACCAATTCCACCCTATAAATAAAACAAGGGAAAAGTGAACAGCA
102301  CCTGAAGAGTTGTGCATGGTTAAGGTGGGATATTTATTTTGTTCCCTTTTCACTTGTCGT

102361  TCACATGAGAGGCTTGGCGAGGGCTGCTATTATAGTAACACATACTAAGTAGTCTCAGCT
102361  AGTGTACTCTCCGAACCGCTCCCGACGATAATATCATTGTGTATGATTCATCAGAGTCGA

102421  GAGCCCTCAGGGTACGTGTGCTGAGTGGTCACCCTCCACAAAACAAAAAATCCTGATACA
102421  CTCGGGAGTCCCATGCACACGACTCACCAGTGGGAGGTGTTTTGTTTTTTAGGACTATGT

102481  CCAAAACTTACTCCTCAAAGTTTCCACTGAGAAACCATGAGTAAAATGTGTGTTTAAATT
102481  GGTTTTGAATGAGGAGTTTCAAAGGTGACTCTTTGGTACTCATTTTACACACAAATTTAA

102541  GTATCCAAACTAACAGGGTTTGATGTTTAGAAACAATAATCAATGATGGAATAGCAGCAA
102541  CATAGGTTTGATTGTCCCAAACTACAAATCTTTGTTATTAGTTACTACCTTATCGTCGTT

102601  AATCGAGTTTTCAGAAAGACCTCAGATGAGCTTTCAAATGGCTTGCCCTCTAACAGGAAA
102601  TTAGCTCAAAAGTCTTTCTGGAGTCTACTCGAAAGTTTACCGAACGGGAGATTGTCCTTT

102661  GACTTTGAATCAGATGCTTCTATTGCCACTGGTTATCAGCTCAAATTCCTAAAGAACTTC
102661  CTGAAACTTAGTCTACGAAGATAACGGTGACCAATAGTCGAGTTTAAGGATTTCTTGAAG

102721  ATCCCAAATACCCTGTCTTGCTGAAAGGTTTACTGGAAGTATAAGAGAATGTCATGTTCT
102721  TAGGGTTTATGGGACAGAACGACTTTCCAAATGACCTTCATATTCTCTTACAGTACAAGA

102781  GTGTCCAGAAAGGAAGGAACACGGGCACCCTAGTGTCAGCGAGTTGTGCTCAGGCTCACA
102781  CACAGGTCTTTCCTTCCTTGTGCCCGTGGGATCACAGTCGCTCAACACGAGTCCGAGTGT

102841  GGATGCCCCCTCACCAGAGGCGTGGGAACACGGCGAGCCCCAGCTGGCCGCGCTCTGCCA
102841  CCTACGGGGGAGTGGTCTCCGCACCCTTGTGCCGCTCGGGGTCGACCGGCGCGAGACGGT

102901  CTGTTTCTAATAGCCGGTCACGTTGATGGAAGTGTCACAGAGTTGTCCAACAGAACTGTC
102901  GACAAAGATTATCGGCCAGTGCAACTACCTTCACAGTGTCTCAACAGGTTGTCTTGACAG

102961  CAGTCAGAAAACCACCACTCATGGTGCTGGAGTGTCTTAGAAGTAAAATATGAATAACAC
102961  GTCAGTCTTTTGGTGGTGAGTACCACGACCTCACAGAATCTTCATTTTATACTTATTGTG

103021  ACACTTATTTAACTATAGGCAGGAGGTTTCTTATGAATTGAAGAGAAACTTTTCTTTGCG
103021  TGTGAATAAATTGATATCCGTCCTCCAAAGAATACTTAACTTCTCTTTGAAAAGAAACGC

103081  TGGGAAGCTGTTCTAAAGTTGGGTAAAACACAATAGATCCACCACCTCTAGCAGCCACTG
103081  ACCCTTCGACAAGATTTCAACCCATTTTGTGTTATCTAGGTGGTGGAGATCGTCGGTGAC

103141  ATAGCTGAAACGTGAAACATAGAGACCCTAAGCTATCACTGCCTCTGAGCTGGCATTGTT
103141  TATCGACTTTGCACTTTGTATCTCTGGGATTCGATAGTGACGGAGACTCGACCGTAACAA

103201  AGGTCATCATAAAGCTAGCGTCTCCCACTGCAAAACCCAAGAGGAAAAAAATAGTTGAAA
103201  TCCAGTAGTATTTCGATCGCAGAGGGTGACGTTTTGGGTTCTCCTTTTTTTATCAACTTT

103261  ATCCTATTTTAAAGGCCTAGCAGATTCTATAAGATACCTTGGGAAAATGATGACGATGAC
103261  TAGGATAAAATTTCCGGATCGTCTAAGATATTCTATGGAACCCTTTTACTACTGCTACTG

103321  TTGAAATCAGACCCTCGTATGCTGCTTCCGTGGGGGCGAACAAAAATATGTTCATCAAAA
103321  AACTTTAGTCTGGGAGCATACGACGAAGGCACCCCCGCTTGTTTTTATACAAGTAGTTTT

103381  ATTAAGCAGAAATGCAGAAAATTTTGTGAGCCAAAAAAGCTGTGTTATCAGGAAATGCAG
```

```
103381  TAATTCGTCTTTACGTCTTTTAAAACACTCGGTTTTTTCGACACAATAGTCCTTTACGTC

103441  ATATTTGTGGGGTTGTAATTTTTTATATTTGAATCGGGCGGTTTTCAAAATGATCTATTC
103441  TATAAACACCCCAACATTAAAAAATATAAACTTAGCCCGCCAAAAGTTTTACTAGATAAG

103501  CATTTGTAGTGTATCTGAAAACCTATAAAAATAAGTTGATATCAATAGATATCCATCTTC
103501  GTAAACATCACATAGACTTTTGGATATTTTTATTCAACTATAGTTATCTATAGGTAGAAG

103561  CATAAAAATTCAACTTCTAAAATTAAGCAAACTTTGCTTTTTCTAATGGCCCCTTTATCC
103561  GTATTTTTAAGTTGAAGATTTTAATTCGTTTGAAACGAAAAAGATTACCGGGGAAATAGG

103621  TCAAATTACCCACTGAAATAGACGGATCACACTCAGCCCTAAGTGAAGCAAGCGTGCATG
103621  AGTTTAATGGGTGACTTTATCTGCCTAGTGTGAGTCGGGATTCACTTCGTTCGCACGTAC

103681  AGAGTAGTCCCAGCCTCGCCTTTGTAATGAGGTGGAAATTAACATGAAGGTAGGCTACCC
103681  TCTCATCAGGGTCGGAGCGGAAACATTACTCCACCTTTAATTGTACTTCCATCCGATGGG

103741  TGTGATAGACACTTAACAGGATACTCGGGGACCCATGGTAATACATCCCTGATAAGGAAT
103741  ACACTATCTGTGAATTGTCCTATGAGCCCCTGGGTACCATTATGTAGGGACTATTCCTTA

103801  AGACCTCACAAATGAACTACTTGCCTGTTAATTCATTTAAAGCCTGACTGTACAGTGAAA
103801  TCTGGAGTGTTTACTTGATGAACGGACAATTAAGTAAATTTCGGACTGACATGTCACTTT

103861  ATTCTCTTAAATAAATATTTGATAAGTGAATCAAATTCTGGCTTACTAAATTGCCAAAAT
103861  TAAGAGAATTTATTTATAAACTATTCACTTAGTTTAAGACCGAATGATTTAACGGTTTTA

103921  ATAATGACTGCCTGCCTTGATAAAAAGAATAATTACATTTAATGAATAAACCTGCCAAGT
103921  TATTACTGACGGACGGAACTATTTTTCTTATTAATGTAAATTACTTATTTGGACGGTTCA

103981  ACAGATATGCCAGGTGGCACCTGCTGTTTGCTGCTCACTTCTCCCAACAAATAGCAGCTA
103981  TGTCTATACGGTCCACCGTGGACGACAAACGACGAGTGAAGAGGGTTGTTTATCGTCGAT

104041  GGTGCCACCTGCACACCAATAAGCTGAGTTTTTCACTTACGGAACAAATAACTTTCAGAA
104041  CCACGGTGGACGTGTGGTTATTCGACTCAAAAAGTGAATGCCTTGTTTATTGAAAGTCTT

104101  GTCAATTTTATAGTTTCTGCCTTGCCCTTTGTTAAAAAAATACACACACTTGAAGCAATG
104101  CAGTTAAAATATCAAAGACGGAACGGGAAACAATTTTTTTATGTGTGTGAACTTCGTTAC

104161  AACTCATGAATTGTTTTCATCATGCATTTCCACTTGCCTAAATATAAAGGGTCCCAGGTA
104161  TTGAGTACTTAACAAAAGTAGTACGTAAAGGTGAACGGATTTATATTTCCCAGGGTCCAT

104221  GATACAGAAATACCTGGTTTGGCCAAACTTGGTTTGAATAACTAGACATGCTAGAAAAGG
104221  CTATGTCTTTATGGACCAAACCGGTTTGAACCAAACTTATTGATCTGTACGATCTTTTCC

104281  TTTTCATTTTCCTGGGATCTGAGTGGAATATGTTAGAAAAGGCATGCTTTCTGAATTCTC
104281  AAAAGTAAAAGGACCCTAGACTCACCTTATACAATCTTTTCCGTACGAAAGACTTAAGAG

104341  TATGCTTAAAACATTTCTAGAGCAGTGCTTCTCAAACTTGAAAGAGCATGTAAATCACCT
104341  ATACGAATTTTGTAAAGATCTCGTCACGAAGAGTTTGAACTTTCTCGTACATTTAGTGGA

104401  GAGATCTCGTGAACATACAACTCTGCTTCAGACCTGGGGGGAGGCAGGAGAGCAAGCATT
104401  CTCTAGAGCACTTGTATGTTGAGACGAAGTCTGGACCCCCCTCCGTCCTCTCGTTCGTAA

104461  CCTAACAAGCTCCCAGCTTGTGGAGCACAGAGCCGCCTCAGGCAGCAGGGCAGTGCAGCT
104461  GGATTGTTCGAGGGTCGAACACCTCGTGTCTCGGCGGAGTCCGTCGTCCCGTCACGTCGA

104521  GAGCAGTGACAGTGAACAGGGCCATTCAGAGGGCTCTCCACCTGGGGCTTAGGTACGACG
104521  CTCGTCACTGTCACTTGTCCCGGTAAGTCTCCCGAGAGGTGGACCCCGAATCCATGCTGC

104581  GGAATCCCCACTGGACAGGCTAGGACTTGCACTGTGGCCATTGTTCCTCCTCCTGCCCAT
104581  CCTTAGGGGTGACCTGTCCGATCCTGAACGTGACACCGGTAACAAGGAGGAGGACGGGTA

104641  GGCTGAGTCAGCTTCTCAGTCCCTTCCAGGATACACTGAGAGGATTCAGGGGCGGTCTCG
104641  CCGACTCAGTCGAAGAGTCAGGGAAGGTCCTATGTGACTCTCCTAAGTCCCCGCCAGAGC

104701  CCTCTGCCCATATCCCCACGTTGGTGAAGTACCACTGGCGCCATTTTCTAATCAGCTCAT
104701  GGAGACGGGTATAGGGGTGCAACCACTTCATGGTGACCGCGGTAAAAGATTAGTCGAGTA

104761  CGGCACCAGCACGTCACTTCCCCTGTTGTGCAGCTCAGTTTTATATTTCTGACTCGGGTG
104761  GCCGTGGTCGTGCAGTGAAGGGGACAACACGTCGAGTCAAAATATAAAGACTGAGCCCAC

104821  TAATAGTAACATCTTCCCCACCAATGTCTCAGGGTTATTGTGAGAACCAGAGAGGAGAGA
104821  ATTATCATTGTAGAAGGGGTGGTTACAGAGTCCCAATAACACTCTTGGTCTCTCCTCTCT
```

```
104881  GGTAGGAAGGAAGCAGGAAATGCAGGGTGCAAATACACCACATTATTATTCTAAATAATG
104881  CCATCCTTCCTTCGTCCTTTACGTCCCACGTTTATGTGGTGTAATAATAAGATTTATTAC

104941  GGATTTTTAATAACAAAGACCAAGAAGATTGTCTGTGCCTATCTAGTTCCCATCTGTAAG
104941  CCTAAAAATTATTGTTTCTGGTTCTTCTAACAGACACGGATAGATCAAGGGTAGACATTC

105001  ATAAAAAGGTGCAGTCCCTGAGCCTCCTAAAAACAGCACCCTAATGGCACTGCTCCTCCC
105001  TATTTTTCCACGTCAGGGACTCGGAGGATTTTTGTCGTGGGATTACCGTGACGAGGAGGG

105061  AGGTATGATTGCAGTAGACCTACGTCAGTTCTTGCAGTTCAAACACATTTTTTCACTTTC
105061  TCCATACTAACGTCATCTGGATGCAGTCAAGAACGTCAAGTTTGTGTAAAAAAGTGAAAG

105121  TTCTGAAATGCCCTCAAATTCTATGACATAGTATCCATGAAGACTTCCCTTTTTACCCAG
105121  AAGACTTTACGGGAGTTTAAGATACTGTATCATAGGTACTTCTGAAGGGAAAAATGGGTC

105181  TGAAGTTACTAGAGGGGGTTCAAACCTTCTTACAGAGGGTTCCTAAAATGGGAGATCTGT
105181  ACTTCAATGATCTCCCCCAAGTTTGGAAGAATGTCTCCCAAGGATTTTACCCTCTAGACA

105241  GGAGTCATTGGGATTAAAAATTAAATTTCAAATTCATGACATCAATGTCCTCATCCCAGT
105241  CCTCAGTAACCCTAATTTTTAATTTAAAGTTTAAGTACTGTAGTTACAGGAGTAGGGTCA

105301  TAAATAAAAGTAAATACAACCTTGAAGAAATTTGACCTACTCAGCTGCAATAGCTGTTGA
105301  ATTTATTTTCATTTATGTTGGAACTTCTTTAAACTGGATGAGTCGACGTTATCGACAACT

105361  CCTGCAGAAGGCAAATGATGGACCCCGGCCACGCAAGAACACACATGGCGTGGGCTGCGT
105361  GGACGTCTTCCGTTTACTACCTGGGGCCGGTGCGTTCTTGTGTGTACCGCACCCGACGCA

105421  TCTTGTGCCCCAGCTCCATGGACGTGTCTTCTTAGCTGTGGACCCCTCCAAAGGAAACTG
105421  AGAACACGGGGTCGAGGTACCTGCACAGAAGAATCGACACCTGGGGAGGTTTCCTTTGAC

105481  CTTGAGTGGAGTGCTCTGTATTGTCACTCCTGGGAATGCTCTCTTAACACCCCCATGCCT
105481  GAACTCACCTCACGAGACATAACAGTGAGGACCCTTACGAGAGAATTGTGGGGGTACGGA

105541  TGTGCTATGTGTTCATACCAGCATAGGCACTCAGAAGCAAAAAAGGTTTCCAGTGAAATT
105541  ACACGATACACAAGTATGGTCGTATCCGTGAGTCTTCGTTTTTTCCAAAGGTCACTTTAA

105601  TGTTTGCAGGCAGGTATTAAGACATCAGACACTGCAGGATCCAAGTGAACCCCTGCCCGC
105601  ACAAACGTCCGTCCATAATTCTGTAGTCTGTGACGTCCTAGGTTCACTTGGGGACGGGCG

105661  CTGCCTGCCTGCCCTGCTCTGCTCTGCAGAAGGGCATGTGTCTGTAGCAATCTGCCCTTT
105661  GACGGACGGACGGGACGAGACGAGACGTCTTCCCGTACACAGACATCGTTAGACGGGAAA

105721  CTGTCTGCGATGGGCAGAGACCCTGGCTCTGCCTCCTGTTTTCCCTGTAGCCCTCCAGAC
105721  GACAGACGCTACCCGTCTCTGGGACCGAGACGGAGGACAAAAGGGACATCGGGAGGTCTG

105781  CTGGGGAGAGACATGGGACTCCAGCACCTGCCCAGCACGGAAGGGGGTCTTTCCCAGGTG
105781  GACCCCTCTCTGTACCCTGAGGTCGTGGACGGGTCGTGCCTTCCCCCAGAAAGGGTCCAC

105841  TAAGTGTAGGGCTGCATTCACTAATTTAGGGTCATACCACAAACACTGTTGTGTCACCAG
105841  ATTCACATCCCGACGTAAGTGATTAAATCCCAGTATGGTGTTTGTGACAACACAGTGGTC

105901  GATGTTGTCACATGCTTATTGTGAGCATTTTTGTGACATGAGCTTTCTCCTGAGAGGCCA
105901  CTACAACAGTGTACGAATAACACTCGTAAAAACACTGTACTCGAAAGAGGACTCTCCGGT

105961  CCCTGTCCAAATGAGAGCTGTAGTTGGAGGCACATCCGTCTCATGCTGCCTGGGTCTCCC
105961  GGGACAGGTTTACTCTCGACATCAACCTCCGTGTAGGCAGAGTACGACGGACCCAGAGGG

106021  TGCACCAGCTCACTCTACTTTTCCTTTTTATCCACTTGGAGTTAAATGCACACATTTATG
106021  ACGTGGTCGAGTGAGATGAAAAGGAAAAATAGGTGAACCTCAATTTACGTGTGTAAATAC

106081  CACAAGTGCACAGAAGGCCCAGAACTGTTCCTAATAGATAATTAAATGATCTCTCTCCAG
106081  GTGTTCACGTGTCTTCCGGGTCTTGACAAGGATTATCTATTAATTTACTAGAGAGAGGTC

106141  TGGAAACTCTCCTGGGCTACCTGGCTGCTAGCTGTTTTAGCCTCTTCATCCAGTATTGAT
106141  ACCTTTGAGAGGACCCGATGGACCGACGATCGACAAAATCGGAGAAGTAGGTCATAACTA

106201  TTACATGCATCGGTTTGTAACTCAGACAGTATAGTTAGTTAGTTTGTTTCACCATCACCT
106201  AATGTACGTAGCCAAACATTGAGTCTGTCATATCAATCAATCAAACAAAGTGGTAGTGGA

106261  TCCAACCTGCTTACGTTTAAACTCCTTCTCTCTGCTTAGCTAATACAGATTTCTGTAAGG
106261  AGGTTGGACGAATGCAAATTTGAGGAAGAGAGACGAATCGATTATGTCTAAAGACATTCC
```

106321 ATTTGGAGAAATACCAGATACCTCCTGGTTAGTATCAGGACCTCTCCATTGGCCTATAAT
106321 TAAACCTCTTTATGGTCTATGGAGGACCAATCATAGTCCTGGAGAGGTAACCGGATATTA

106381 TGAACATCACTGTAAATTAGTTTCAGCTTTTATAGTAATGCATCTAACCCGCCAGCTTAA
106381 ACTTGTAGTGACATTTAATCAAAGTCGAAAATATCATTACGTAGATTGGGCGGTCGAATT

106441 AAAAAAAGTGATGATAATCATGAAATTCAAGTGTCAATATAGACTTTTAAATCTGGTGTG
106441 TTTTTTTCACTACTATTAGTACTTTAAGTTCACAGTTATATCTGAAAATTTAGACCACAC

106501 GATGTTAATAGTGCCAGATATGCTTCTGTGTAATCATTTTTGAGAAGTGCTGTCTTCTTT
106501 CTACAATTATCACGGTCTATACGAAGACACATTAGTAAAAACTCTTCACGACAGAAGAAA

106561 ACCATAATCACATTAATCAATCGGGAAATGGTTAAATAGACATTTTAATACATTGTTTTT
106561 TGGTATTAGTGTAATTAGTTAGCCCTTTACCAATTTATCTGTAAAATTATGTAACAAAAA

106621 TCCTGAGAAAAAAAAGTGACTCCAGGAAATATGTTGTTACAATATCTCAACACCAGGAATT
106621 AGGACTCTTTTTTTCACTGAGGTCCTTTATACAACAATGTTATAGAGTTGTGGTCCTTAA

106681 GTGGTCACATAGCATTCTTCTTCTAAATAAAGCATTTGTTTTACAAAAGAGTTTCATGTT
106681 CACCAGTGTATCGTAAGAAGAAGATTTATTTCGTAAACAAAATGTTTTCTCAAAGTACAA

106741 TACTTAAACGTGACAAATGTATCACTCTTCGAAGTGAGAGTTAAATGAAATGTTCCCTGA
106741 ATGAATTTGCACTGTTTACATAGTGAGAAGCTTCACTCTCAATTTACTTTACAAGGGACT

106801 TTTAGATTCCAAGACCTGTGACTACTCAGTAGGACAATAGCATATTTCTTTTACTCAGAT
106801 AAATCTAAGGTTCTGGACACTGATGAGTCATCCTGTTATCGTATAAAGAAAATGAGTCTA

106861 ACCTGTTTATCCCTTTGTCAGCTCCAGAATGGAGACACAGGTGCTGCCACACAGAGGAGA
106861 TGGACAAATAGGGAAACAGTCGAGGTCTTACCTCTGTGTCCACGACGGTGTGTCTCCTCT

106921 GGGTCAGGTGTTTTCTTAGTCACAAATGCAGCAGTCTCACCCCTGAAAGAATTCCTTCAT
106921 CCCAGTCCACAAAAGAATCAGTGTTTACGTCGTCAGAGTGGGGACTTTCTTAAGGAAGTA

106981 AACAACTAAATTGCACCTCTTAGAATACCTGTACTTAAAATATAATTTGTCTTCTATGGA
106981 TTGTTGATTTAACGTGGAGAATCTTATGGACATGAATTTTATATTAAACAGAAGATACCT

107041 TATAGACAAATAATAGCTTAAATGTTTTAGGATTTGTAAAGAAAGAAAAAGCTAGAAATC
107041 ATATCTGTTTATTATCGAATTTACAAAATCCTAAACATTTCTTTCTTTTTCGATCTTTAG

107101 TCAATGCAACTCTGACTTGTATGTCCTGCCATAGGAAGGTTTGGGACTTTTCTTTGGAAG
107101 AGTTACGTTGAGACTGAACATACAGGACGGTATCCTTCCAAACCCTGAAAAGAAACCTTC

107161 ATGGGAGGCTTCCTTCACAGTTTCTTAACTTCTCATTTTATTCAATCTTTCTCCATCTCT
107161 TACCCTCCGAAGGAAGTGTCAAAGAATTGAAGAGTAAAATAAGTTAGAAAGAGGTAGAGA

107221 CTCTCTCTTTCTCTCATCCTCACGCTCTCTGTCTGTCTCTCTCCTGATTTGTCTGCTTAG
107221 GAGAGAGAAAGAGAGTAGGAGTGCGAGAGACAGACAGAGAGAGGACTAAACAGACGAATC

107281 GGACAGACAGGTAATGGAGAAGAAATTGAAAGGGAAAATACAGGGAAAATGTAAATGAAA
107281 CCTGTCTGTCCATTACCTCTTCTTTAACTTTCCCTTTTATGTCCCTTTTACATTTACTTT

107341 AGAAGTTATTATAACTGTTATAATGAAGTGCCTTTCATTTAAATATAAGAATGTAACGCT
107341 TCTTCAATAATATTGACAATATTACTTCACGGAAAGTAAATTTATATTCTTACATTGCGA

107401 GAAATGAACTTGTGATCCTGAAATGCATTTTTAATGAGTTTCCTTTTTATTTTGCTGCTT
107401 CTTTACTTGAACACTAGGACTTTACGTAAAAATTACTCAAAGGAAAAATAAAACGACGAA

107461 CAGTGGCATATTTTAAAGACCCTTTGAAAAAAGCCACATTAAAAAACCCCACCAAATGCC
107461 GTCACCGTATAAAATTTCTGGGAAACTTTTTTTCGGTGTAATTTTTTGGGGTGGTTTACGG

107521 ACAACACGCAAAATTGGATATTGGTTTATTCCAAAACTGCTGATCAGGAAGAGTGGCTCT
107521 TGTTGTGCGTTTTAACCTATAACCAAATAAGGTTTTGACGACTAGTCCTTCTCACCGAGA

107581 TCATCACAAAATGTTGCAGTCACTTTATTTAAATGAGTTTGAATTTGCATTGTGATGTGC
107581 AGTAGTGTTTTACAACGTCAGTGAAATAAATTTACTCAAACTTAAACGTAACACTACACG

107641 CATTCTGATTTAGGGAAAAAAAATTAGATTTTCAGATCAAATTGACCCAGCCAGTGAAGC
107641 GTAAGACTAAATCCCTTTTTTTTAATCTAAAAGTCTAGTTTAACTGGGTCGGTCACTTCG

107701 GTTAAGGAGCTGGCAGGTTTAGTCTGAAAGCCTCATGTTCTATCAAACCTGCAGCCGGTG
107701 CAATTCCTCGACCGTCCAAATCAGACTTTCGGAGTACAAGATAGTTTGGACGTCGGCCAC

107761 GAAGTCACCAGGCCCCCGTGGGAAACAACTTTCTCGTAGCATCGTCCTCATGTCCCCACG

```
107761  CTTCAGTGGTCCGGGGGCACCCTTTGTTGAAAGAGCATCGTAGCAGGAGTACAGGGGTGC

107821  CTGAGTTTAGTTCTCACCAGCTCTCCTCTCTCCGTCCCAGGAGAACGGCCCTTCCAGTGC
107821  GACTCAAATCAAGAGTGGTCGAGAGGAGAGAGGCAGGGTCCTCTTGCCGGGAAGGTCACG

107881  AATCAGTGCGGGGCCTCATTCACCCAGAAGGGCAACCTGCTCCGGCACATCAAGCTGCAT
107881  TTAGTCACGCCCCGGAGTAAGTGGGTCTTCCCGTTGGACGAGGCCGTGTAGTTCGACGTA

107941  TCCGGGGAGAAGCCCTTCAAATGCCACCTCTGCAACTACGCCTGCCGCCGGAGGGACGCC
107941  AGGCCCCTCTTCGGGAAGTTTACGGTGGAGACGTTGATGCGGACGGCGGCCTCCCTGCGG

108001  CTCACTGGCCACCTGAGGACGCACTCCGGTAGGTCCCCTGGATGCAGTCCGGGGCTGTCT
108001  GAGTGACCGGTGGACTCCTGCGTGAGGCCATCCAGGGGACCTACGTCAGGCCCCGACAGA

108061  GGGTGTCCCGGGATTCCTCCACTCTGCCCGCCTGGGTCCCGGATTGTGTCCTTGCTGGCT
108061  CCCACAGGGCCCTAAGGAGGTGAGACGGGCGGACCCAGGGCCTAACACAGGAACGACCGA

108121  AACCCTGAGTCCCTCCCAGCTCGCAGTCCTGCATCGGGTGTGAGCTGTTGCTTCTTTGAC
108121  TTGGGACTCAGGGAGGGTCGAGCGTCAGGACGTAGCCCACACTCGACAACGAAGAAACTG

108181  ATTGCCCCATCCCCCCTCCCCATATTCTCCTTTTCCCACTGCACCAGGGAGATTGGGCGC
108181  TAACGGGGTAGGGGGGAGGGGTATAAGAGGAAAAGGGTGACGTGGTCCCTCTAACCCGCG

108241  AGGACCCCCATGCACATACACACACAGAAGTCTCACCTTAGGTAGCATGTTTCAGAACAG
108241  TCCTGGGGGTACGTGTATGTGTGTGTCTTCAGAGTGGAATCCATCGTACAAAGTCTTGTC

108301  GCCTCCTCATCCCGTTTGTCAGGCTGGTATGGTTTCCCTAAAGGATTTTGCGGGAAATGT
108301  CGGAGGAGTAGGGCAAACAGTCCGACCATACCAAAGGGATTTCCTAAAACGCCCTTTACA

108361  TTTCACCAAGTGTACTGCTAAGACCCAAAACGTTTTCAAGTACAATTCTTTCTGTTGTTA
108361  AAAGTGGTTCACATGACGATTCTGGGTTTTGCAAAAGTTCATGTTAAGAAAGACAACAAT

108421  AGTCCCGCTTTGGAGTGTTTTACACAGGTGTAATGGATAGCTTTTTTGCAGAGCTGGTAG
108421  TCAGGGCGAAACCTCACAAAATGTGTCCACATTACCTATCGAAAAAACGTCTCGACCATC

108481  AGAAGGGTGATTTAGGGCGCACCCACCCAGACTGAGCCCCGTGTGGCTCTCACACCAAAA
108481  TCTTCCCACTAAATCCCGCGTGGGTGGGTCTGACTCGGGGCACACCGAGAGTGTGGTTTT

108541  ACCAGCCAAGGCCACAGTTACAGAGGCCAGTCTGGGGCTGTTACCAGATTTTAGACAGCA
108541  TGGTCGGTTCCGGTGTCAATGTCTCCGGTCAGACCCCGACAATGGTCTAAAATCTGTCGT

108601  GCCTTTCTCTTTGAATTAGACAGTTAAAGTACAACCCACATAATCTGGAGTCTTGACAAG
108601  CGGAAAGAGAAACTTAATCTGTCAATTTCATGTTGGGTGTATTAGACCTCAGAACTGTTC

108661  ATCATCATAGGCATAAACGCTCTATCATTCTCAAAACAGTCTCAGCCTGCAAAGTTCAAA
108661  TAGTAGTATCCGTATTTGCGAGATAGTAAGAGTTTTGTCAGAGTCGGACGTTTCAAGTTT

108721  TCCACTAAAGTTTGGTTAGATCCTCTGCCTCCTGAGAAATGGTCCTGGGTGTTTCATTAT
108721  AGGTGATTTCAAACCAATCTAGGAGACGGAGGACTCTTTACCAGGACCCACAAAGTAATA

108781  CCAGCAGTCCCATAATTCTACAGGGCAGAGGAAGAGAGGGCTCTTGGCCGGCCTGTCATG
108781  GGTCGTCAGGGTATTAAGATGTCCCGTCTCCTTCTCTCCCGAGAACCGGCCGGACAGTAC

108841  GATCATGTTTGCCTACAGTGTGGTCTATACAACATGACATGGCACAGGTCTCCTTCATAC
108841  CTAGTACAAACGGATGTCACACCAGATATGTTGTACTGTACCGTGTCCAGAGGAAGTATG

108901  CGTCCAGTTGGGGATATTTGCTGTAGCATACTGCATGAGACTTCGGAGGCGAAAGGTTGA
108901  GCAGGTCAACCCCTATAAACGACATCGTATGACGTACTCTGAAGCCTCCGCTTTCCAACT

108961  TGGCTTTTGTCCTTCCCCTCAAGGAGCTTCCTGCCCCAGCCAGCCACCAGCCAGGGCCCT
108961  ACCGAAAACAGGAAGGGGAGTTCCTCGAAGGACGGGGTCGGTCGGTGGTCGGTCCCGGGA

109021  TCTCCAAGCAGCAGCCTCTCTAAGCCGGTGCCCTGGGGGATGGCAATGCCTCAGACCAGC
109021  AGAGGTTCGTCGTCGGAGAGATTCGGCCACGGGACCCCCTACCGTTACGGAGTCTGGTCG

109081  TACTCCTCACCCACCCTGGGTAGCAGGATAAGGAGGAGCCTCCCTCAGGGAGGCAGACGT
109081  ATGAGGAGTGGGTGGGACCCATCGTCCTATTCCTCCTCGGAGGGAGTCCCTCCGTCTGCA

109141  GTGTTCTTTGTGAAATATCTGCAGCGGCCCAGGCCATCTCTTCCCAAATGTGATGCGTGT
109141  CACAAGAAACACTTTATAGACGTCGCCGGGTCCGGTAGAGAAGGGTTTACACTACGCACA

109201  ATTTGATGGTTGAGGGTTTTAGAGGCTGCTCATTGTGTCCATCTCTTTATCACACATTTA
109201  TAAACTACCAACTCCCAAAATCTCCGACGAGTAACACAGGTAGAGAAATAGTGTGTAAAT
```

```
109261  CTGAGCACCCCTGGGTGCCCATTCTGCACAGAAGACCTCAGGTGCACACAAGGAAAGCAC
109261  GACTCGTGGGGACCCACGGGTAAGACGTGTCTTCTGGAGTCCACGTGTGTTCCTTTCGTG

109321  CAATGTGTTACAGGAGGCATGACAGAGTGTCTGAGGGGACAGTGGGAGCATAAGGAAGGG
109321  GTTACACAATGTCCTCCGTACTGTCTCACAGACTCCCCTGTCACCCTCGTATTCCTTCCC

109381  GACAGTGGGTCCTGGGGTGGAAAGTCAGGAGAGACTTCCCATAGAAGGGAGGGGCTTCTG
109381  CTGTCACCCAGGACCCCACCTTTCAGTCCTCTCTGAAGGGTATCTTCCCTCCCCGAAGAC

109441  AAGTACATGCTGGAAGGCGAGTGTCTGAGGTTTGCTTTAGAGACATTCTAAACTAGTGTG
109441  TTCATGTACGACCTTCCGCTCACAGACTCCAAACGAAATCTCTGTAAGATTTGATCACAC

109501  TGAGTCGGCAGTAATCCCAGAGAGGGGTGCGGCACCCATGTTGGCAAGAACTCGGTCTCA
109501  ACTCAGCCGTCATTAGGGTCTCTCCCCACGCCGTGGGTACAACCGTTCTTGAGCCAGAGT

109561  CTCCTGGCAGGTGGTCCCTGCAGTAGTTCACCCAGTTGGCTGGAGACAAAAAGGAGAGGA
109561  GAGGACCGTCCACCAGGGACGTCATCAAGTGGGTCAACCGACCTCTGTTTTTCCTCTCCT

109621  GTCACAGGGGCTGGTGCATCACCTCCTCTCCATCCTGACCTCCTTCCTGGCGGTGCACAT
109621  CAGTGTCCCCGACCACGTAGTGGAGGAGAGGTAGGACTGGAGGAAGGACCGCCACGTGTA

109681  GGAGAAGGATCCCCACATGCTCGCCATCAGAAAATTGTCATGATTTGGGTGATTTGACTT
109681  CCTCTTCCTAGGGGTGTACGAGCGGTAGTCTTTTAACAGTACTAAACCCACTAAACTGAA

109741  CCTAAAAACTTCCAAGAAAGGAGCTATACCAAGCTGAGGAGTTGCTTGCCCGAGAGGGCG
109741  GGATTTTTGAAGGTTCTTTCCTCGATATGGTTCGACTCCTCAACGAACGGGCTCTCCCGC

109801  GGCAGATCAGCAGGGGGGCCATGGAGTGGGACACTTGCCTTGTTACAGAGGGACAGGGAG
109801  CCGTCTAGTCGTCCCCCCGGTACCTCACCCTGTGAACGGAACAATGTCTCCCTGTCCCTC

109861  AACAGGTGGTTCCCCTGAGATAGGAAGACAGTAGCATGACGGTGAGCATCTTTAAATGTT
109861  TTGTCCACCAAGGGGACTCTATCCTTCTGTCATCGTACTGCCACTCGTAGAAATTTACAA

109921  CACATGGGTGATGAGCATGTATGGTGGTCCCACCTGGAGACAAAGGTGGTACCTGGGCCT
109921  GTGTACCCACTACTCGTACATACCACCAGGGTGGACCTCTGTTTCCACCATGGACCCGGA

109981  GATAAGCAAGGCCCATGGAGGCCTCAGCTAAGAGAACTGTGCAGTCTGTCTGCCATGCTA
109981  CTATTCGTTCCGGGTACCTCCGGAGTCGATTCTCTTGACACGTCAGACAGACGGTACGAT

110041  ACCAGTGGTCTAGGAAGAAGCTTCACATCTGTTGTTTCTAAGGAGCCATGTAGACATAGA
110041  TGGTCACCAGATCCTTCTTCGAAGTGTAGACAACAAAGATTCCTCGGTACATCTGTATCT

110101  CACAGTTGGCAGCAATATTGCCTTGGGTTAGTGGGAAGCACCACTTCCCATGTCAGTGGA
110101  GTGTCAACCGTCGTTATAACGGAACCCAATCACCCTTCGTGGTGAAGGGTACAGTCACCT

110161  AAACACAGACACTCGTTCTTAGGTTGACACCAACCATAGGTCATCATCACTCCTAGATTA
110161  TTTGTGTCTGTGAGCAAGAATCCAACTGTGGTTGGTATCCAGTAGTAGTGAGGATCTAAT

110221  CTCTGATTGCAATGGAATCTCTCATGCCCATCAGTCCTTAAAAGCAGAGCTGCTCTCTTT
110221  GAGACTAACGTTACCTTAGAGAGTACGGGTAGTCAGGAATTTTCGTCTCGACGAGAGAAA

110281  CCTTTCTTGTTGTGGAGTTAGGAACATGTAGTGGAGTATTTTATGTTATTCTGGCTCCAG
110281  GGAAAGAACAACACCTCAATCCTTGTACATCACCTCATAAAATACAATAAGACCGAGGTC

110341  GCCTATGTATTTCATTTGAGGCATGATGGTTTCCTATTCTAAGTACCACTAAGTATTGAG
110341  CGGATACATAAAGTAAACTCCGTACTACCAAAGGATAAGATTCATGGTGATTCATAACTC

110401  TAGTTATAAAATGGTGGTAGACTTTGGACTGTGTATTTCTATTCTGGATCCACGGATGAG
110401  ATCAATATTTTACCACCATCTGAAACCTGACACATAAAGATAAGACCTAGGTGCCTACTC

110461  CAAGTGGAGAAAACCCATCATGGATGCAGGGCTGCTGGGGATCTGTCTGTCCGAAAATGC
110461  GTTCACCTCTTTTGGGTAGTACCTACGTCCCGACGACCCCTAGACAGACAGGCTTTTACG

110521  TATAGAGACACACACTGCTCTTGGGGAAAAAAACAGTGATGATTTCTTCAAAATGATTTC
110521  ATATCTCTGTGTGTGACGAGAACCCCTTTTTTTTGTCACTACTAAAGAAGTTTTACTAAAG

110581  ACTGAAGTGAAATAAGGACTTTGGCCTTGCTCTGGCAAATTCAAATGTGGCCACACTCTA
110581  TGACTTCACTTTATTCCTGAAACCGGAACGAGACCGTTTAAGTTTACACCGGTGTGAGAT

110641  CTCCATTGTGGAATGGTGACTCCCCTGTCCCATGAAAAACCAGGATCAACATCACAGCTT
110641  GAGGTAACACCTTACCACTGAGGGGACAGGGTACTTTTTGGTCCTAGTTGTAGTGTCGAA
```

110701 TCTCTAGCAACTGGTGGTGGCCTGAGGTCTTATGAACTAGCATGTTAGCAGGGATAGAGT
110701 AGAGATCGTTGACCACCACCGGACTCCAGAATACTTGATCGTACAATCGTCCCTATCTCA

110761 TAGGGAATTTGTCCGATGCTGACAGCAAGAGGGTGCTGGGGGAGCTGGAGGGGTGGAGAA
110761 ATCCCTTAAACAGGCTACGACTGTCGTTCTCCCACGACCCCCTCGACCTCCCCACCTCTT

110821 GTGATCTCCCTTCCTGCCTCATGGCTTCCTTTGGAAGTTGCAAGCTTAGAATTTCTTTCC
110821 CACTAGAGGGAAGGACGGAGTACCGAAGGAAACCTTCAACGTTCGAATCTTAAAGAAAGG

110881 CAAAGAATATTGGACTATGCTTCAAGAGACACTTGAGTTCAGTTGCTTACAGTGAACATA
110881 GTTTCTTATAACCTGATACGAAGTTCTCTGTGAACTCAAGTCAACGAATGTCACTTGTAT

110941 CTTCTTATGTTCCAGAGTTAAACTCAATCATATTTTCAGAAAAATAGATATTCAGACATC
110941 GAAGAATACAAGGTCTCAATTTGAGTTAGTATAAAAGTCTTTTTATCTATAAGTCTGTAG

111001 ATGTCATATTTACAGGAAGCTATTTGATCATGGAATTTACTGAAGTGACAGCTTTTTCAG
111001 TACAGTATAAATGTCCTTCGATAAACTAGTACCTTAAATGACTTCACTGTCGAAAAAGTC

111061 GAAAGGGTATCGAGTGGTTATCTATATCCTATTAATTTCTAAAAGCAATGCTCAATAAAA
111061 CTTTCCCATAGCTCACCAATAGATATAGGATAATTAAAGATTTTCGTTACGAGTTATTTT

111121 TGTGCAATTTTAGGCAATATTCTGTGTTTCAACATTATTTCCTTATGTTGGGGGTACATC
111121 ACACGTTAAAATCCGTTATAAGACACAAAGTTGTAATAAAGGAATACAACCCCCATGTAG

111181 ACTTTATATAGATAAAGAAAGATCCTTTTCATTCTCTATTAAATGGTCGAGAAGTCAAAT
111181 TGAAATATATCTATTTCTTTCTAGGAAAAGTAAGAGATAATTTACCAGCTCTTCAGTTTA

111241 TTTGGCTTTCATGTTGGTCTTCTTACTGTAGCTTCTGTTCACCTAAGAAAGAGATTTAAC
111241 AAACCGAAAGTACAACCAGAAGAATGACATCGAAGACAAGTGGATTCTTTCTCTAAATTG

111301 CAACATAAGCTTTGGCTAAACCATTGTACCAAACATTTCTATTTGGATGCTCTTAACTTT
111301 GTTGTATTCGAAACCGATTTGGTAACATGGTTTGTAAAGATAAACCTACGAGAATTGAAA

111361 TATACAATACTTTGAAATGTGTTCAATAAAACAACATGAGAAAGAAGTAGAAGATTCTGC
111361 ATATGTTATGAAACTTTACACAAGTTATTTTGTTGTACTCTTTCTTCATCTTCTAAGACG

111421 TATTAGGTTGGAAATTGATATTCCTTACATAAGTGATTAAAGATACTTTATTCAGTTTAT
111421 ATAATCCAACCTTTAACTATAAGGAATGTATTCACTAATTTCTATGAAATAAGTCAAATA

111481 CCCTTAAAATGGTATCTACTGAAGTCCTTAAACTTAGGCATGATATTAATTCCTGGCTAG
111481 GGGAATTTTACCATAGATGACTTCAGGAATTTGAATCCGTACTATAATTAAGGACCGATC

111541 TCTTTCTTTTTCCTAAAATATCACTCAATAAAATCCAAGAAGGACTAAAACTCATTATAT
111541 AGAAAGAAAAAGGATTTTATAGTGAGTTATTTTAGGTTCTTCCTGATTTTGAGTAATATA

111601 AGAACTGCACTGCCCAATACAGCAGCCACTAGCCACATGCAGTATTTTTAATTATAATAA
111601 TCTTGACGTGACGGGTTATGTCGTCGGTGATCGGTGTACGTCATAAAAATTAATATTATT

111661 GTTAAAATTAAACAAAATTAAAACTTCAGCCCTCAGTGGCACTAGCCACACTTTTTGTGT
111661 CAATTTTAATTTGTTTTAATTTTGAAGTCGGGAGTCACCGTGATCGGTGTGAAAAACACA

111721 CCAACAGCTACATATGGCTGGTGGCTGCCATATTGGACAGCACTAATAGAGAGCATTTCT
111721 GGTTGTCGATGTATACCGACCACCGACGGTATAACCTGTCGTGATTATCTCTCGTAAAGA

111781 ATCATCATAGGAAGTTCTGTTGAACAAGGCTGATCTGGAAGCTACACCATCCTAAAGCTC
111781 TAGTAGTATCCTTCAAGACAACTTGTTCCGACTAGACCTTCGATGTGGTAGGATTTCGAG

111841 TTTGGCAAGAAAGCAGAGCTTTTCCATAAGCTCAGTTTATGAACAATATATTTTGTATTT
111841 AAACCGTTCTTTCGTCTCGAAAAGGTATTCGAGTCAAATACTTGTTATATAAAACATAAA

111901 TCTATGTATATGTATATGAGAATACGTATACATAGAAATATATGAAAATACATATACATA
111901 AGATACATATACATATACTCTTATGCATATGTATCTTTATATACTTTTATGTATATGTAT

111961 TAAATATACATATACATAAAATAGAAATACATATACATAGAAAATACAAAATATTTTGTA
111961 ATTTATATGTATATGTATTTTATCTTTATGTATATGTATCTTTTATGTTTTATAAAACAT

112021 TTACATATACATATGCATGGAAATACAAAATATTTTGTATTATATAATGCAAATGTAAGA
112021 AATGTATATGTATACGTACCTTTATGTTTTATAAAACATAATATATTACGTTTACATTCT

112081 TATATTTATATTATGTAAGAAAATACATTTGTGCAATAAAACGCTTATTTCATTCAAGCA
112081 ATATAAATATAATACATTCTTTTATGTAAACACGTTATTTTGCGAATAAAGTAAGTTCGT

112141 TATTTGTCCAGCATGGTTTCTTAGAGACTGAAGTTACCCAGTTTAGCACCTGGAACTGAA

```
112141  ATAAACAGGTCGTACCAAAGAATCTCTGACTTCAATGGGTCAAATCGTGGACCTTGACTT

112201  TTCACTCTGGCTGTGTTAGATGTAGCACCAGCGTCATGACAGATCCAGCCCACAGCATAG
112201  AAGTGAGACCGACACAATCTACATCGTGGTCGCAGTACTGTCTAGGTCGGGTGTCGTATC

112261  ACTGAGGCCTACTCTGGAGGCAAATCTGCAGACAATGTGGCCAACACATGGGCCTTCCTC
112261  TGACTCCGGATGAGACCTCCGTTTAGACGTCTGTTACACCGGTTGTGTACCCGGAAGGAG

112321  ACATCTCTGTCCAGTCGCTGGCATCCACGTTATAACACAGCAGTGGCTCGTTTTCTCCAG
112321  TGTAGAGACAGGTCAGCGACCGTAGGTGCAATATTGTGTCGTCACCGAGCAAAAGAGGTC

112381  TTAAAGGCTAGGCTTCCGTGATTGAAGCAGGAGAGAATTTGGTAGTCAGTAAAGTAGACA
112381  AATTTCCGATCCGAAGGCACTAACTTCGTCCTCTCTTAAACCATCAGTCATTTCATCTGT

112441  AACTAAGACAAAACAAGAGTCGCATGAGTTCTTGGTACAATTAAAGAAACCCTTGGCCAC
112441  TTGATTCTGTTTTGTTCTCAGCGTACTCAAGAACCATGTTAATTTCTTTGGGAACCGGTG

112501  CAACGTTTTTAAATTCAGGAATTTCACCAAGTCCGTAACAGTTTTAGCTCTCAAGGGTAG
112501  GTTGCAAAAATTTAAGTCCTTAAAGTGGTTCAGGCATTGTCAAAATCGAGAGTTCCCATC

112561  AATTTTTTTTTTAACTTTTTTGGTAATAATTGTATTGCATGCATTCCCCTTACACAGAAG
112561  TTAAAAAAAAAAATTGAAAAAACCATTATTAACATAACGTACGTAAGGGGAATGTGTCTTC

112621  GCTGGCATTTAATTGGGGTCTTGAACTCAATTGTGTTTTCTGCAGTTGGTAAACCTCACA
112621  CGACCGTAAATTAACCCCAGAACTTGAGTTAACACAAAAGACGTCAACCATTTGGAGTGT

112681  AATGTGGATATTGTGGCCGAAGCTATAAACAGCGAAGCTCTTTAGAGGAACATAAAGAGC
112681  TTACACCTATAACACCGGCTTCGATATTTGTCGCTTCGAGAAATCTCCTTGTATTTCTCG

112741  GCTGCCACAACTACTTGGAAAGCATGGGCCTTCCGGGCACACTGTACCCAGGTAAGCGCT
112741  CGACGGTGTTGATGAACCTTTCGTACCCGGAAGGCCCGTGTGACATGGGTCCATTCGCGA

112801  GCTGCTCGGAGGCCAGCCTGGTGGGCTCTCCCCCCAGCACGGTGGGGAAGGAGGGCGCTC
112801  CGACGAGCCTCCGGTCGGACCACCCGAGAGGGGGGTCGTGCCACCCCTTCCTCCCGCGAG

112861  TGCATGCAGCCTTAGGAGCAGAGCCTTGGGCCTGCTTCCTGCCGGGGCTAGGAGGGAGGG
112861  ACGTACGTCGGAATCCTCGTCTCGGAACCCGGACGAAGGACGGCCCCGATCCTCCCTCCC

112921  AAGTTTTTGGCCAATAGCATCAGTTTCACCAGAAGCACGTTGTGCTTCCCAGCTTTCTAG
112921  TTCAAAAACCGGTTATCGTAGTCAAAGTGGTCTTCGTGCAACACGAAGGGTCGAAAGATC

112981  GTCCTCATCTGACCAGAGAGAGCTTGATTTTAAAACCCTTCCCACTTCCAATCGGGAGAA
112981  CAGGAGTAGACTGGTCTCTCTCGAACTAAAATTTTGGGAAGGGTGAAGGTTAGCCCTCTT

113041  ACTCCTAGGATAGCAGTGACCTTGAAAGTTTTGGGGTTGTTTTTTGGATGTTGGTGATTT
113041  TGAGGATCCTATCGTCACTGGAACTTTCAAAACCCCAACAAAAAACCTACAACCACTAAA

113101  TAAAACAACAACAAAAAAAACACCTCAAGTAGTGATATTTCTTTGTAAACAAAATAAAAT
113101  ATTTTGTTGTTGTTTTTTTTGTGGAGTTCATCACTATAAAGAAACATTTGTTTTATTTTA

113161  GTAAAATATTGTTTTGAAACAATTTTTTAACAAAGTTGATCAAATAAGAACTTTCAGGCT
113161  CATTTTATAACAAAACTTTGTTAAAAAATTGTTTCAACTAGTTTATTCTTGAAAGTCCGA

113221  GTGATTTTAAGCATCAGTTCAGACTGAGGGCAGGGTATCCACTCCTGCATGGAGTGTGAG
113221  CACTAAAATTCGTAGTCAAGTCTGACTCCCGTCCCATAGGTGAGGACGTACCTCACACTC

113281  GTTGAAGTTTGTTGGAAGCCCCTGGTAGTTGATAATATTGAGGTTGTTGAAGGATGAGAG
113281  CAACTTCAAACAACCTTCGGGGACCATCAACTATTATAACTCCAACAACTTCCTACTCTC

113341  GAAAGGTCGCTTCTTTCTTAGGAGCTGTGATGCTCCACTGTATTGCATAACGAGATAGCA
113341  CTTTCCAGCGAAGAAAGAATCCTCGACACTACGAGGTGACATAACGTATTGCTCTATCGT

113401  CTTGATTCAGACCCCAGAGCCTGTAGAATAAATGATGTGGAGAGAGTATTGAAAAAGGCT
113401  GAACTAAGTCTGGGGTCTCGGACATCTTATTTACTACACCTCTCTCATAACTTTTTCCGA

113461  GGATTTATCTGGGGGAATGGCAATGTTATAAAAGGTTCTTCCCTAATTTTTTAAGAGATA
113461  CCTAAATAGACCCCCTTACCGTTACAATATTTTCCAAGAAGGGATTAAAAAATTCTCTAT

113521  AGCAATTTATAAAGAGATCAAATAAAACTAATCTCTTTCAAAACAGTATTCATTTAGCAC
113521  TCGTTAAATATTTCTCTAGTTTATTTTGATTAGAGAAAGTTTTGTCATAAGTAAATCGTG

113581  TACCCCTGGAAAAATCATCACTAGACTGTGAGTCGAAATTTAACGAATGGAGGAAAATTT
113581  ATGGGGACCTTTTTAGTAGTGATCTGACACTCAGCTTTAAATTGCTTACCTCCTTTTAAA
```

113641 ATTAAGCCTCATTCACCTTTCGGACTTTAACAAATGGTTCAAAAGGGAAAATTCACCAGA
113641 TAATTCGGAGTAAGTGGAAAGCCTGAAATTGTTTACCAAGTTTTCCCTTTTAAGTGGTCT

113701 GTACCCTCAGATTATTAATGGCAGATGAGAGCAGGAAAAAAGAATGTTGAGAATGGCTAA
113701 CATGGGAGTCTAATAATTACCGTCTACTCTCGTCCTTTTTTCTTACAACTCTTACCGATT

113761 AAATTGATTCCAATTAGTCTTGGCATGGAGAGGCAACATCTCACCTGGCCAGAGCCCTGC
113761 TTTAACTAAGGTTAATCAGAACCGTACCTCTCCGTTGTAGAGTGGACCGGTCTCGGGACG

113821 AGCCAGAGTGCTGTCTCCTCTGCCAGTCAGCCAGGGATCTGGGGCTTACAGGCGATTCCT
113821 TCGGTCTCACGACAGAGGAGACGGTCAGTCGGTCCCTAGACCCCGAATGTCCGCTAAGGA

113881 CCAGGCCTCAATCTTCTCACCTGTGAAATGGGAAGATACCATTAAATTGTCATTTGGCTT
113881 GGTCCGGAGTTAGAAGAGTGGACACTTTACCCTTCTATGGTAATTTAACAGTAAACCGAA

113941 CTTTCAAAATCTCAGGTCTAGAATGGAAAGGCATTGAAGGTGAGTGGAAGAGAAGAAACT
113941 GAAAGTTTTAGAGTCCAGATCTTACCTTTCCGTAACTTCCACTCACCTTCTCTTCTTTGA

114001 GGATGTTAAAATAATAATAATGTAACAGTTATTAATTCACATGGCCAGACCCCAGGGCAG
114001 CCTACAATTTTATTATTATTACATTGTCAATAATTAAGTGTACCGGTCTGGGGTCCCGTC

114061 AAGACTAATGGACGAGAGGATGATTATGTTCCTTTGAAATAAAATGTCAACATTCAAATA
114061 TTCTGATTACCTGCTCTCCTACTAATACAAGGAAACTTTATTTTACAGTTGTAAGTTTAT

114121 TTGTTTTTTTCTGTAAAAAATAAGTGAATGGGTCTTGGAAGGAAGTGCTCCTAATGATGT
114121 AACAAAAAAAAGACATTTTTTATTCACTTACCCAGAACCTTCCTTCACGAGGATTACTACA

114181 AGTTTACATATTAATGCTTAATAAACCATTTTATTTACTCAACAAATATTCACTGAGGAC
114181 TCAAATGTATAATTACGAATTATTTGGTAAAATAAATGAGTTGTTTATAAGTGACTCCTG

114241 TTATAATCAATAATACCAAGCCCTATTGAATACAGTAATCTAATATACACATAGAATTGA
114241 AATATTAGTTATTATGGTTCGGGATAACTTATGTCATTAGATTATATGTGTATCTTAACT

114301 AAAGACAAAATGAACAAAAGAAACAGACAGTATATCATTTCTGTTTCCTCCACTGTTTTC
114301 TTTCTGTTTTACTTGTTTTCTTTGTCTGTCATATAGTAAAGACAAAGGAGGTGACAAAAG

114361 ATTACTAAAGAAACAAGCTGCAGTACACAGCCGGCTGAGAGCTGCTTCCGGGGGAGCAGG
114361 TAATGATTTCTTTGTTCGACGTCATGTGTCGGCCGACTCTCGACGAAGGCCCCCTCGTCC

114421 AGAAGCAGCACTCATTTCACTGAGACCTGGGTTTGAGTTCTCTGTTTTTGAATGTGACCC
114421 TCTTCGTCGTGAGTAAAGTGACTCTGGACCCAAACTCAAGAGACAAAAACTTACACTGGG

114481 AGAGCTCATCGCTTAAATGCGACTGATAATTCGTGCCTGGCCACCCTCATAGCAGCGCTC
114481 TCTCGAGTAGCGAATTTACGCTGACTATTAAGCACGGACCGGTGGGAGTATCGTCGCGAG

114541 TGAAAGAAACATAAAGGAGTGGATACACGTGTGAGCAGGAGCCTAGACATGTGTCAGGTA
114541 ACTTTCTTTGTATTTCCTCACCTATGTGCACACTCGTCCTCGGATCTGTACACAGTCCAT

114601 CTTACCTAAACATGGATGTTTAGGTCACTGACATACTGACACCATCAAGCTAATTCTTCT
114601 GAATGGATTTGTACCTACAAATCCAGTGACTGTATGACTGTGGTAGTTCGATTAAGAAGA

114661 CTATTGGTCTCTATGTCAGTAAGGCTACGTGCCAGTAATGTGTATTAAATATTACCCTGT
114661 GATAACCAGAGATACAGTCATTCCGATGCACGGTCATTACACATAATTTATAATGGGACA

114721 GGTGAGTATCAATTTCACCCTAAAATGTTCATTTATTCCCTTTGGTGGTTAAGTTACTTG
114721 CCACTCATAGTTAAAGTGGGATTTTACAAGTAAATAAGGGAAACCACCAATTCAATGAAC

114781 GACTACCACAGTGATGCCTCTATGTCCCCACACTCAGGAGGTATGAAAGCGTCTCTAGGA
114781 CTGATGGTGTCACTACGGAGATACAGGGGTGTGAGTCCTCCATACTTTCGCAGAGATCCT

114841 AGGAGCAGTGGTCTGGTCTCAGTGTGGCAGCAGACAAAGCCAGGCTCGCACTTAATTACA
114841 TCCTCGTCACCAGACCAGAGTCACACCGTCGTCTGTTTCGGTCCGAGCGTGAATTAATGT

114901 TTGAATTATTTCTGTCTGGTATGGCCTCAAAAGGTGCAAAATGATCATAAATATCTTTCG
114901 AACTTAATAAAGACAGACCATACCGGAGTTTTCCACGTTTTACTAGTATTTATAGAAAGC

114961 ACAGTCACCTTTGGTTTTTGCCCTTGCCAGCACCTGGCCCTTGTAGGTGCTGCTGGTTTA
114961 TGTCAGTGGAAACCAAAAACGGGAACGGTCGTGGACCGGGAACATCCACGACGACCAAAT

115021 AGGAACTTGGAGTATTAAGGGGATCCATAAAAGGACTTTGACAGAAGTCTTCCTGAACTG
115021 TCCTTGAACCTCATAATTCCCCTAGGTATTTTCCTGAAACTGTCTTCAGAAGGACTTGAC

115081 TGCTTAGGGCCAGGAAGGCAGGCAGGGGCTTGGTCATTTCAGCTGTGCTGTACCCGGATG
115081 ACGAATCCCGGTCCTTCCGTCCGTCCCCGAACCAGTAAAGTCGACACGACATGGGCCTAC

115141 AGTGGAAACTGTGCCTTCTGATGCATCTCTCCTCCTCTGTTCCCCTCTAAGAACCCCTTG
115141 TCACCTTTGACACGGAAGACTACGTAGAGAGGAGGAGACAAGGGGAGATTCTTGGGGAAC

115201 TCTGTGCTCAGATGGATGAGCAAGTAGTCTCTTCTCCCACCCCAAAAACCTCCTTAGATG
115201 AGACACGAGTCTACCTACTCGTTCATCAGAGAAGAGGGTGGGGTTTTTGGAGGAATCTAC

115261 GGGTGGATTCAGTAAAATCAAAGCCTCAGTTAGGCAATATAGGAAAAGATGAATTTTTGT
115261 CCCACCTAAGTCATTTTAGTTTCGGAGTCAATCCGTTATATCCTTTTCTACTTAAAAACA

115321 GAATATAATTTTCTTTAGTAAAAGGCTTTTGGTCACCTCTTATAATTTCTCAATGTCTTT
115321 CTTATATTAAAAGAAATCATTTTCCGAAAACCAGTGGAGAATATTAAAGAGTTACAGAAA

115381 ATGGAAAAAATAAGTTTCTCTTGAGCCTTTGGTGCACAAACAGAGTCACCAGCCCTTGGA
115381 TACCTTTTTTATTCAAAGAGAACTCGGAAACCACGTGTTTGTCTCAGTGGTCGGGAACCT

115441 GGTGCATGCATCTAATGACAGAGCAGCCTGTGACTCGCGAGCATCAGGCTCACTTGTTCC
115441 CCACGTACGTAGATTACTGTCTCGTCGGACACTGAGCGCTCGTAGTCCGAGTGAACAAGG

115501 TTGCTGTGATTGACTCTGGGGGGTGGAGGCTTTTGGGAGGCCCTGCGCCATGGTGCAGAG
115501 AACGACACTAACTGAGACCCCCCACCTCCGAAAACCCTCCGGGACGCGGTACCACGTCTC

115561 GAAGACCCCTGACCTCTGGGAGCCCTGGGCAGGCTGTGGGTGTGGCTTGGCAGAGAAGGA
115561 CTTCTGGGGACTGGAGACCCTCGGGACCCGTCCGACACCCACACCGAACCGTCTCTTCCT

115621 TGGACATGTCAGGAAGCATATTAGCTTAATAGATGAAGTGTGTGACCAGTGTAATTGGGT
115621 ACCTGTACAGTCCTTCGTATAATCGAATTATCTACTTCACACACTGGTCACATTAACCCA

115681 TTTATGGATTTATATTTATCTTGCATATAAATCAGGGTGTGGCCTATATATGTTTGCCCA
115681 AAATACCTAAATATAAATAGAACGTATATTTAGTCCCACACCGGATATATACAAACGGGT

115741 TTTAATAAAAAACACATATTCCAAATGTTAAAATATAAAATATACCAGCAACAATAGCCA
115741 AAATTATTTTTTGTGTATAAGGTTTACAATTTTATATTTTATATGGTCGTTGTTATCGGT

115801 ATTTGAAATATATGTTTTCAAAGTAACATTCATAATAGCAATGAAGACATAAATACTTAG
115801 TAAACTTTATATACAAAAGTTTCATTGTAAGTATTATCGTTACTTCTGTATTTATGAATC

115861 AAATGAAGAAAACCAGCAAGCCATAGTGAGGCAAGAAAGAAAACATGGGAGATTGGAGAG
115861 TTTACTTCTTTTGGTCGTTCGGTATCACTCCGTTCTTTCTTTTGTACCCTCTAACCTCTC

115921 GTAGAGAGGTCTGGATGCCGTCCATGGGCCTCTGCCCAAATTAGTGTATCATTTCAACAT
115921 CATCTCTCCAGACCTACGGCAGGTACCCGGAGACGGGTTTAATCACATAGTAAAGTTGTA

115981 AAAACTGGCTTGTAAATTTTACTGGAATCTGGTAACCATGTATCATAAAGTATCTGGATA
115981 TTTTGACCGAACATTTAAAATGACCTTAGACCATTGGTACATAGTATTTCATAGACCTAT

116041 TATACATGAGATAATTGGCAACTTTTTAGAAGAAGAATAATGTCAAGTGATGTATGAAAA
116041 ATATGTACTCTATTAACCGTTGAAAAATCTTCTTCTTATTACAGTTCACTACATACTTTT

116101 TGTAGCTGAAATGGATGGAACTCTTTGAAGAATAGACAAACAAATCAATAATAGAGATGG
116101 ACATCGACTTTACCTACCTTGAGAAACTTCTTATCTGTTTGTTTAGTTATTATCTCTACC

116161 ATACCCTACTGAAGACCCCATGATAAATCAGCATGAAATATCCGAGGCAGGAGGAAGCAT
116161 TATGGGATGACTTCTGGGGTACTATTTAGTCGTACTTTATAGGCTCCGTCCTCCTTCGTA

116221 AAGGCAGGAACGGTATTGGAAGAATTGATCAGGCATTTAAGATAATCAGTTTACATTATC
116221 TTCCGTCCTTGCCATAACCTTCTTAACTAGTCCGTAAATTCTATTAGTCAAATGTAATAG

116281 ATCTCCTTATACACTAGGATAAATTGCCGGGAAATGAAAGAGTTATATTTTTATAACTCC
116281 TAGAGGAATATGTGATCCTATTTAACGGCCCTTTACTTTCTCAATATAAAAATATTGAGG

116341 AACCACAATAACTGATGTAAATATAGATCCAGCTATGCAGGATACAAAGTTGTATCTTCA
116341 TTGGTGTTATTGACTACATTTATATCTAGGTCGATACGTCCTATGTTTCAACATAGAAGT

116401 ATTATGGAAAGAGATGACATAGAACAATTGAGAATGACTAGAAATGGCCCACACATCAGC
116401 TAATACCTTTCTCTACTGTATCTTGTTAACTCTTACTGATCTTTACCGGGTGTGTAGTCG

116461 ATTTACCTCCATAATTTCTAAATTTTCTACAGCAGTCAGGGGTGCTTTGATTTTCAGGAG
116461 TAAATGGAGGTATTAAAGATTTAAAAGATGTCGTCAGTCCCCACGAAACTAAAAGTCCTC

116521 TTTGTTGTATGCTCAGGGATAAGGGGCTCTGGGTTTTTCCTCCACATTGCCCAGATCTGC

116521 AAACAACATACGAGTCCCTATTCCCCGAGACCCAAAAAGGAGGTGTAACGGGTCTAGACG

116581 TGTAGGCTCTTGGGGGCCACATTTTTGTATCTCCGTTCACTGTGCTGGTCCATGAAGTCA
116581 ACATCCGAGAACCCCCGGTGTAAAAACATAGAGGCAAGTGACACGACCAGGTACTTCAGT

116641 GATGTTAGATCACACAGTTCTTTGCGTCTCAGAGGGCTGCTGTTAGGGACTGACTGAGGT
116641 CTACAATCTAGTGTGTCAAGAAACGCAGAGTCTCCCGACGACAATCCCTGACTGACTCCA

116701 AGCAGATTTGGGAGTATTATGAAAAGTTAAAAGGACTTGTCAAAGGGACATCAATACTTA
116701 TCGTCTAAACCCTCATAATACTTTTCAATTTTCCTGAACAGTTTCCCTGTAGTTATGAAT

116761 ATCCCAGTCCAGCCTCTGGTGATGGTGGTGGTGCCTGTGCATTTGGACTTAAAGAGGATG
116761 TAGGGTCAGGTCGGAGACCACTACCACCACCACGGACACGTAAACCTGAATTTCTCCTAC

116821 ATGTGTGGTGAGCAGCTGTTGGAGGAAGAAGCATTTGCAGGTGGCTGGTCATTGGCATTG
116821 TACACACCACTCGTCGACAACCTCCTTCTTCGTAAACGTCCACCGACCAGTAACCGTAAC

116881 CTCCGGCTGCCTCTGCCTGTCTGGAAGTGTTGCTGGGAAGATTAGAATTAATCTCTAGGA
116881 GAGGCCGACGGAGACGGACAGACCTTCACAACGACCCTTCTAATCTTAATTAGAGATCCT

116941 AGGGCCTGGCTCTTGTAGGCACTTAACAAATGTCAGATTTAACATTGGACGCGACTGAAC
116941 TCCCGGACCGAGAACATCCGTGAATTGTTTACAGTCTAAATTGTAACCTGCGCTGACTTG

117001 CCTTTAAACATAAGCCTTTCTAAACTGGCCTCTCTGTCTTTGACTTTAGTCATTAAAGAA
117001 GGAAATTTGTATTCGGAAAGATTTGACCGGAGAGACAGAAACTGAAATCAGTAATTTCTT

117061 GAAACTAATCACAGTGAAATGGCAGAAGACCTGTGCAAGATAGGATCAGAGAGATCTCTC
117061 CTTTGATTAGTGTCACTTTACCGTCTTCTGGACACGTTCTATCCTAGTCTCTCTAGAGAG

117121 GTGCTGGACAGACTAGCAAGTAACGTCGCCAAACGTAAGAGCTCTATGCCTCAGAAATTT
117121 CACGACCTGTCTGATCGTTCATTGCAGCGGTTTGCATTCTCGAGATACGGAGTCTTTAAA

117181 CTTGGTAAGAGTTAAATGTTTGCTGTCTCTTAAAAAAAAACTATGTGGGTGTTTTAGATG
117181 GAACCATTCTCAATTTACAAACGACAGAGAATTTTTTTTTGATACACCCACAAAATCTAC

117241 CAAGTAGAAATGAGTTGAGGGTGGAAGAAAGGGAAAAAAATCTTATTTTTTCAAAAGGAA
117241 GTTCATCTTTACTCAACTCCCACCTTCTTTCCCTTTTTTTAGAATAAAAAAGTTTTCCTT

117301 AAATTGGTAAGCTTAACATTCCTTAAATATCTTAGAATTTTTTCCAATAAGTATCTTAAA
117301 TTTAACCATTCGAATTGTAAGGAATTTATAGAATCTTAAAAAAGGTTATTCATAGAATTT

117361 AATAACAAACCTCCCATCAGTTTTTCCTAGATTTGATTTTGCAGCATCTGGGGCCTGCCC
117361 TTATTGTTTGGAGGGTAGTCAAAAAGGATCTAAACTAAAACGTCGTAGACCCCGGACGGG

117421 TGTGATCTGCCTGTGGACATCGCTCTTAGGGGCGGCTGCACCAGCGTGCACAGGGTGGAG
117421 ACACTAGACGGACACCTGTAGCGAGAATCCCCGCCGACGTGGTCGCACGTGTCCCACCTC

117481 AGTTTGGGCCTGGCTCGTCCGGGGGACACCACACTGCAGGACACTCCAGGCCTGGCCGGC
117481 TCAAACCCGGACCGAGCAGGCCCCCTGTGGTGTGACGTCCTGTGAGGTCCGGACCGGCCG

117541 TTCTCAGAGCTTCAGATCCTCATTTTTCATATGAAGCTCCTAATGCTCCCCTTATGGGGG
117541 AAGAGTCTCGAAGTCTAGGAGTAAAAAGTATACTTCGAGGATTACGAGGGGAATACCCCC

117601 ACTCTGAAGGGTTAATGGGAGGAATCATACAGTGACTGACCCCTGAGAAGTGTCCAGTGA
117601 TGAGACTTCCCAATTACCCTCCTTAGTATGTCACTGACTGGGGACTCTTCACAGGTCACT

117661 AGACAGGGCTTAGCTAGGATTGCTGTTTTGCCTAATGCTCTGCGGGATTAAAAAAAAAGA
117661 TCTGTCCCGAATCGATCCTAACGACAAAACGGATTACGAGACGCCCTAATTTTTTTTTCT

117721 AGAAGAAGAACAAGACCATTCGTCTCTCTAGGAGCATTGCCCAGAGTAGGTATTAGACAC
117721 TCTTCTTCTTGTTCTGGTAAGCAGAGAGATCCTCGTAACGGGTCTCATCCATAATCTGTG

117781 ACCAACACCACCATCCAGCCAGACGCTGCAGGGACAGTGAGCCAGGGTCCGAGTGGAAAG
117781 TGGTTGTGGTGGTAGGTCGGTCTGCGACGTCCCTGTCACTCGGTCCCAGGCTCACCTTTC

117841 GCGCTAGGCTTGGGAACCAGCTCAGAGTCAATACAGAGCCACCGCCACTCACCAACTCTG
117841 CGCGATCCGAACCCTTGGTCGAGTCTCAGTTATGTCTCGGTGGCGGTGAGTGGTTGAGAC

117901 TCAGCTTAGTAAAATGGCTCTGCCCCTAGAGCCCTGGTTCCATCCTTTAGTATCTCACAG
117901 AGTCGAATCATTTTACCGAGACGGGGATCTCGGGACCAAGGTAGGAAATCATAGAGTGTC

117961 GGTGATTGTGAATATCCCATGACTCCAAGATTGAGAAAACGTTTAGAATCCCTCGGTGTG
117961 CCACTAACACTTATAGGGTACTGAGGTTCTAACTCTTTTGCAAATCTTAGGGAGCCACAC

```
118021  AAGGTTAACTCTGTCCGGAAAGAGGACCAGTAAAAGCTTCATGAGGCTGAGATGCACTTT
118021  TTCCAATTGAGACAGGCCTTTCTCCTGGTCATTTTCGAAGTACTCCGACTCTACGTGAAA

118081  GGAAGAGGAATAGAGTTTCAGCACATTCTAGGTGTTGGAGGAATGGGGGAATCTAGGCAG
118081  CCTTCTCCTTATCTCAAAGTCGTGTAAGATCCACAACCTCCTTACCCCCTTAGATCCGTC

118141  ATGTTTAAAATCAATGAGAAACCAGAATGCTGACCATGAGGGTTGGAGTGGGGGCCTAAG
118141  TACAAATTTTAGTTACTCTTTGGTCTTACGACTGGTACTCCCAACCTCACCCCCGGATTC

118201  GACATGACGGAGGAGCAGGGTGTGTTCCCAGCTTAACTCAGGTACCCATGGGGAAGCAGG
118201  CTGTACTGCCTCCTCGTCCCACACAAGGGTCGAATTGAGTCCATGGGTACCCCTTCGTCC

118261  AAAAGTGAAGGTGTCCTAGGCAGCTCTGCCACAGGATGAATGGCTTCAGATGCCAGGTGA
118261  TTTTCACTTCCACAGGATCCGTCGAGACGGTGTCCTACTTACCGAAGTCTACGGTCCACT

118321  GCGAGGGACCCTTCATTCAGTCAGCAGGAAAGAAGCACTGGCATATTTTTTATGAGAACA
118321  CGCTCCCTGGGAAGTAAGTCAGTCGTCCTTTCTTCGTGACCGTATAAAAAATACTCTTGT

118381  AAGGCTAGGATAGTAAAGACAGCAAGTACCAAAAAATGACTGGAAAAGGGAGACTGTGGA
118381  TTCCGATCCTATCATTTCTGTCGTTCATGGTTTTTTACTGACCTTTTCCCTCTGACACCT

118441  GGCAGTGGCAGCAGGCATGGAAAGAAGGGCTTGTGAAGGGGAAGGGGTGGTGTCAGAGGA
118441  CCGTCACCGTCGTCCGTACCTTTCTTCCCGAACACTTCCCCTTCCCCACCACAGTCTCCT

118501  ACATAGGGCTGGGGGCAGGGATTAGGTGAGGGAAACCATGAGTCACACTGATTCTAGAGT
118501  TGTATCCCGACCCCCGTCCCTAATCCACTCCCTTTGGTACTCAGTGTGACTAAGATCTCA

118561  AGTGTGCCCTTGATGAAAAGGATAACACCAGGTTCTAGGAAAAGATGGGGTTCTGTTTTT
118561  TCACACGGGAACTACTTTTCCTATTGTGGTCCAAGATCCTTTTCTACCCCAAGACAAAAA

118621  GACGTGTTGATTTTCAAGGACTTCTGGTGTTTGTGACACATGGGGAAATTGTGGTGGGAG
118621  CTGCACAACTAAAAGTTCCTGAAGACCACAAACACTGTGTACCCCTTTAACACCACCCTC

118681  AGAGGTGGGGCCAGAACAGGGGCTGGTGAGGCCAAGGGTCCCAGAGGGCACCTGTTGACC
118681  TCTCCACCCCGGTCTTGTCCCCGACCACTCCGGTTCCCAGGGTCTCCCGTGGACAACTGG

118741  TGCAGGATGACATGAAGGGGGAAGGACAGAGGCAAGGCCAAGTCCTGGGCACCAGCCTCC
118741  ACGTCCTACTGTACTTCCCCCTTCCTGTCTCCGTTCCGGTTCAGGACCCGTGGTCGGAGG

118801  CTCTTGCAGCTTCAAATAGGGCTCCATTTTGACCTTTTGATTAATTAGAGGTTTGTCATA
118801  GAGAACGTCGAAGTTTATCCCGAGGTAAAACTGGAAAACTAATTAATCTCCAAACAGTAT

118861  GGTTGGGGGTTGAGAGGAGCAAGGGAGAGAAGGATTCAGTGTACAAAAAGAATGAAAGCC
118861  CCAACCCCCAACTCTCCTCGTTCCCTCTCTTCCTAAGTCACATGTTTTTCTTACTTTCGG

118921  ACTGGCTGAGCCAGTGGGGAGTTGTCCACACACACATGAGCCTTTGGACCATGAGAACGA
118921  TGACCGACTCGGTCACCCCTCAACAGGTGTGTGTGTACTCGGAAACCTGGTACTCTTGCT

118981  GGGAGGCCTTGCCTTCCTGAACGGAGTAGGAGTGAGGTCCTGTGCTGAGCGTAAGCAGTG
118981  CCCTCCGGAACGGAAGGACTTGCCTCATCCTCACTCCAGGACACGACTCGCATTCGTCAC

119041  GGATTCCCACAGCACTGGGCACAGAGCCCACGGGCTGCCTCCTGAGCAGCCAGCATCTGC
119041  CCTAAGGGTGTCGTGACCCGTGTCTCGGGTGCCCGACGGAGGACTCGTCGGTCGTAGACG

119101  CTGGGGTGGACACAGTGACAGAGAGATGGGTGGTGACTGGGGTATGGGCAGAGATAAGGC
119101  GACCCCACCTGTGTCACTGTCTCTCTACCCACCACTGACCCCATACCCGTCTCTATTCCG

119161  AGCAAGTGTGTGCAAGGGGAGTGAAGGGTTACTGACCTTAAGAAGCAGGGATGGCGTCCT
119161  TCGTTCACACACGTTCCCCTCACTTCCCAATGACTGGAATTCTTCGTCCCTACCGCAGGA

119221  CTGTCAGGTGAGGAGCCTGGAGAATGCTTTGGTGAATGAACGTTTGCAGCCCCTTTTAGC
119221  GACAGTCCACTCCTCGGACCTCTTACGAAACCACTTACTTGCAAACGTCGGGGAAAATCG

119281  TTTTGGAGACTTGAAACCAAAGGAGAGATTCATCTGTGAAACTCTACTGGAGCCACTCCC
119281  AAAACCTCTGAACTTTGGTTTCCTCTCTAAGTAGACACTTTGAGATGACCTCGGTGAGGG

119341  CAACCCCCACCCTTGTGAGACCACAATGTGGGCGTTGGCTTGAGATGCTTCTGTGTTAGT
119341  GTTGGGGGTGGGAACACTCTGGTGTTACACCCGCAACCGAACTCTACGAAGACACAATCA

119401  AGAAGAAATAAACAACACAGTGCTCTGATGAGGCAAAGCGAAGATGAAAAAGGAGTTCCC
119401  TCTTCTTTATTTGTTGTGTCACGAGACTACTCCGTTTCGCTTCTACTTTTTCCTCAAGGG
```

119461 AGGGGACATAGTAGGAACAGTGGACGAGGGTAGCAGAAGAGGAGTTTGGAGCAAAAGACT
119461 TCCCCTGTATCATCCTTGTCACCTGCTCCCATCGTCTTCTCCTCAAACCTCGTTTTCTGA

119521 CACAAGCAGCTGCATAATCTGTTGGTGCTTGGCAGTTCATTTGTAAAAATGATGCCTCTT
119521 GTGTTCGTCGACGTATTAGACAACCACGAACCGTCAAGTAAACATTTTTACTACGGAGAA

119581 CCTGCCCTAAAATACCTACCTTACCCCCGCTTCAACTTGATGAGATTTCCATCAGTCACT
119581 GGACGGGATTTTATGGATGGAATGGGGGCGAAGTTGAACTACTCTAAAGGTAGTCAGTGA

119641 CCCAATGTGTCACAGCTTCTGCAGCCCTAAAATTAAAAGGTGAGTGAGTCTCTGAGGCCC
119641 GGGTTACACAGTGTCGAAGACGTCGGGATTTTAATTTTCCACTCACTCAGAGACTCCGGG

119701 CTCTCCACTTCTCGGATGCTGAGTTTAGCCTTCATGTGAATGTGGAAAGACTAGGAATAC
119701 GAGAGGTGAAGAGCCTACGACTCAAATCGGAAGTACACTTACACCTTTCTGATCCTTATG

119761 AGCTGTTATCACACAAGCTGGCCCAATAGTGGTTCAGTTGAGAGAGCCCCATCCTTCAGA
119761 TCGACAATAGTGTGTTCGACCGGGTTATCACCAAGTCAACTCTCTCGGGGTAGGAAGTCT

119821 GTCAGCTCCAGCTAGGAGTGACTGGTGGCCTTGAGCATGGTGCTGGGCTTAGTGTTGCCA
119821 CAGTCGAGGTCGATCCTCACTGACCACCGGAACTCGTACCACGACCCGAATCACAACGGT

119881 TCTGTGGAATGGGTGTGGGTCTGTTGCCCTGCCTCCTCCCAGAGCTATTCTGAGGCTCAG
119881 AGACACCTTACCCACACCCAGACAACGGGACGGAGGAGGGTCTCGATAAGACTCCGAGTC

119941 AAGGGGTGATGGATGTGATGGTGCTCCCAACACTAGAAAGCATCTTAAGAATGTAAGATT
119941 TTCCCCACTACCTACACTACCACGAGGGTTGTGATCTTTCGTAGAATTCTTACATTCTAA

120001 TTCATGATGACTGTTGCTCAGAGTGGCTATTATAGTTTTGCTTTATTGTTCTATAACCTA
120001 AAGTACTACTGACAACGAGTCTCACCGATAATATCAAAACGAAATAACAAGATATTGGAT

120061 TGATTAAAATTTTTACCTTAAACTTTGACGTGAGTGTGAATAAGTATTTGTTTTGCCAGC
120061 ACTAATTTTAAAAATGGAATTTGAAACTGCACTCACACTTATTCATAAACAAAACGGTCG

120121 AACATTCCTCACCACTGGGGCCATTAAAGATCTCCCCCTCTGAGACCATCAAATACAGGT
120121 TTGTAAGGAGTGGTGACCCCGGTAATTTCTAGAGGGGGAGACTCTGGTAGTTTATGTCCA

120181 CAACAGGACTGATTAATCTAATTAGAAAAGGGCTTGTATTAAATAGCAATGATAATTGTT
120181 GTTGTCCTGACTAATTAGATTAATCTTTTCCCGAACATAATTTATCGTTACTATTAACAA

120241 GTTTTTAGTCTGTCTGGTGTTTGACTTGGGAACGTTTTTAAAATAGAGAAAAGCACAAAG
120241 CAAAAATCAGACAGACCACAAACTGAACCCTTGCAAAAATTTTATCTCTTTTCGTGTTTC

120301 AGGAAAACAACAATTACCAATATTCCTGCTACCCATTATAATTATCTAGGTATATTTTCT
120301 TCCTTTTGTTGTTAATGGTTATAAGGACGATGGGTAATATTAATAGATCCATATAAAAGA

120361 TCTTTTGTAAGAAAAAGAAACCCTGTTATATTGTTAAAATAACACAAAGTTAATATAAAG
120361 AGAAAACATTCTTTTTCTTTGGGACAATATAACAATTTTATTGTGTTTCAATTATATTTC

120421 AATTTTAATGCAAAGATTAATGTTTTCAAATCACCACAAAACCCAACATCCAGAAATTAC
120421 TTAAAATTACGTTTCTAATTACAAAAGTTTAGTGGTGTTTTGGGTTGTAGGTCTTTAATG

120481 CAATATTAAAAGTAGAAAAGTATCATTCTAAATATTTTCTGTTGCATATGTATGTGAGTG
120481 GTTATAATTTTCATCTTTTCATAGTAAGATTTATAAAAGACAACGTATACATACACTCAC

120541 GATAGGCTGATGAATTAGGTGGATTGATGGATAGGTAAATATGAAATAAATACTTTCATA
120541 CTATCCGACTACTTAATCCACCTAACTACCTATCCATTTATACTTTATTTATGAAAGTAT

120601 AATATTCCAACTTATCATACATGCCTTAAATTCAAGAGGTGAAAAAAGACCCAAACAAAA
120601 TTATAAGGTTGAATAGTATGTACGGAATTTAAGTTCTCCACTTTTTTCTGGGTTTGTTTT

120661 CTAGAGAAGCGGCTTATTTTAAATATCCTCTGACATAAAGGAATATTATATTTAAAGGAT
120661 GATCTCTTCGCCGAATAAAATTTATAGGAGACTGTATTTCCTTATAATATAAATTTCCTA

120721 CCTCTAAGATTAAAAATATGTACTATGAAAAACATTAAGAAATTTGAATTTTTTTTAATC
120721 GGAGATTCTAATTTTTATACATGATACTTTTTGTAATTCTTTAAACTTAAAAAAAATTAG

120781 CATTTGTTTCAATTTAAGCAGCATCTACTGGCTCACTGCTTTGAAAAATAAGGACAGTAT
120781 GTAAACAAAGTTAAATTCGTCGTAGATGACCGAGTGACGAAACTTTTTATTCCTGTCATA

120841 TCCAGTTCACATTCAGTGTTCCAGTGTTCACATTATCTTATTATTTTTACATTGTCCAGC
120841 AGGTCAAGTGTAAGTCACAAGGTCACAAGTGTAATAGAATAATAAAAATGTAACAGGTCG

120901 TTTGTAATATTCACATTCTATTCTGTAATCATAATTCATAGTAGTTTAGTTATTTATTAC

120901 AAACATTATAAGTGTAAGATAAGACATTAGTATTAAGTATCATCAAATCAATAAATAATG

120961 TAACTCTATTTAAATAGATTCAAGGATCAGACCCTGCCCTTTTCTTCTTATTTATGTTTA
120961 ATTGAGATAAATTTATCTAAGTTCCTAGTCTGGGACGGGAAAAGAAGAATAAATACAAAT

121021 TTTTGATTAATCTCTTAATTGATTGGACTTTACATTCAAGCAACTTTTTTAAAAAAAAGT
121021 AAAACTAATTAGAGAATTAACTAACCTGAAATGTAAGTTCGTTGAAAAAATTTTTTTTCA

121081 TTCTATAGATGTTCTATTTCTATCATTGTATTGTTTTTGAGGATGTTGGCCTGTTGCCTT
121081 AAGATATCTACAAGATAAAGATAGTAACATAACAAAAACTCCTACAACCGGACAACGGAA

121141 TGTATTTGATGAGCATTTTGACAGAGTCTATGGTCTTGGGCCACTCTTTCTTTTTCTCCC
121141 ACATAAACTACTCGTAAAACTGTCTCAGATACCAGAACCCGGTGAGAAAGAAAAAGAGGG

121201 TTGAGAACTTTTTAGATTTTGCTGATGGCATTGCTTGTTGAATGTTGCTGTGGAAACATC
121201 AACTCTTGAAAAATCTAAAACGACTACCGTAACGAACAACTTACAACGACACCTTTGTAG

121261 AAGTCTAGTGTAACTGTTTCTTCTTCAAGGTGATTTGCATTTTATTCCTGAATGCCTGAG
121261 TTCAGATCACATTGACAAAGAAGAAGTTCCACTAAACGTAAAATAAGGACTTACGGACTC

121321 GGTTCTTTATTTAACCTTGAAGTTAAATACCCTAATTAGGATGTATCTTGGTCTATTCAT
121321 CCAAGAAATAAATTGGAACTTCAATTTATGGGATTAATCCTACATAGAACCAGATAAGTA

121381 TCGGAATAAAAAATTCCTGCCATTTTGTCTAGAGAGTCCCTTTTTTTTCTCTTTATTTCT
121381 AGCCTTATTTTTTAAGGACGGTAAAACAGATCTCTCAGGGAAAAAAAAGAGAAATAAAGA

121441 GGGAAATTCTCTTTTATATAAATATGTTTTGTTCCATCTATTGTGATCTCTGTTGAGGGA
121441 CCCTTTAAGAGAAAATATATTTATACAAAACAAGGTAGATAACACTAGAGACAACTCCCT

121501 TACCAGTTGTCCATATGTTAGATAATTTGTCTTCCATATCTGTTAACAGTTCTTAAAGTT
121501 ATGGTCAACAGGTATACAATCTATTAAACAGAAGGTATAGACAATTGTCAAGAATTTCAA

121561 TTTTGTTTATTTCTTTGTCTATTTTTACATTTACTCACTGTTCTCTTGTGGTTTTCCTCT
121561 AAAACAAATAAAGAAACAGATAAAAATGTAAATGAGTGACAAGAGAACACCAAAAGGAGA

121621 GTCAGTAATTTAATTTTTAGTAGTTCCTGTTCTATTACTTGCTATTTTTAATCCATGCAT
121621 CAGTCATTAAATTAAAAATCATCAAGGACAAGATAATGAACGATAAAAATTAGGTACGTA

121681 TAATTTTATAATAATATTATTTTGCTCCTTATTTTGTCTCCTGAGACCCGAAATCTCTTT
121681 ATTAAAATATTATTATAATAAAACGAGGAATAAAACAGAGGACTCTGGGCTTTAGAGAAA

121741 TTTCCTCTTACTCTGTTGCTTTTGCATTTTATTTTGAATACTTTTAAAATTGATTCCATG
121741 AAAGGAGAATGAGACAACGAAAACGTAAAATAAAACTTATGAAAATTTTAACTAAGGTAC

121801 TTATGAAGCAATTATGAGGCATTTCCTCTCTTGTTGGAATTAACGATTTTTTCCCCTAGG
121801 AATACTTCGTTAATACTCCGTAAAGGAGAGAACAACCTTAATTGCTAAAAAAGGGGATCC

121861 AGGGACTCTATGGTCTGTGTTTTACTTCCTTTCTTCCCCTGTATTTCTAGAAAATATTTT
121861 TCCCTGAGATACCAGACACAAAATGAAGGAAAGAAGGGGACATAAAGATCTTTTATAAAA

121921 CCTAGTAACCCTGACATTTCTTTTCATCTTGCTTATTCTAGTTGGTCTGATATAGCTTGA
121921 GGATCATTGGGACTGTAAAGAAAAGTAGAACGAATAAGATCAACCAGACTATATCGAACT

121981 TTGACATTTCAGCCTTCTTCCCACTATATTTTTTTTTCCTGTGAGAGCTATTGGGTTTTC
121981 AACTGTAAAGTCGGAAGAAGGGTGATATAAAAAAAAAGGACACTCTCGATAACCCAAAAG

122041 TAAATCCTGAAAGAATGCCAAAGATGGGGTTGGAGGAGTTTGGTGAGGCAAAGTGCAGCC
122041 ATTTAGGACTTTCTTACGGTTTCTACCCCAACCTCCTCAAACCACTCCGTTTCACGTCGG

122101 TTTGTTAAAATACTTTTCCTTTGCTCTCTCTCCCTCATCTGAAATTTAGTTAAATACCCT
122101 AAACAATTTTATGAAAAGGAAACGAGAGAGAGGGAGTAGACTTTAAATCAATTTATGGGA

122161 AAGCCATCAGCACTGTACCTAGTTGGGGAATGCTTTCATCCCCACAGGAGATTCTCTGGG
122161 TTCGGTAGTCGTGACATGGATCAACCCCTTACGAAAGTAGGGGTGTCCTCTAAGAGACCC

122221 GCTTTGGGCCATCTTCCCCTTCAGTGTAGACCACAGAGGACTTTGCTTCTGTCCCAGGGA
122221 CGAAACCCGGTAGAAGGGGAAGTCACATCTGGTGTCTCCTGAAACGAAGACAGGGTCCCT

122281 GCCCGCAGGGGCTCACTTCTCCATGTTCATCTGATTCTTGTCAGCCAAGGTTTCAAATGC
122281 CGGGCGTCCCCGAGTGAAGAGGTACAAGTAGACTAAGAACAGTCGGTTCCAAAGTTTACG

122341 TTTTCTGATCAGAACAGGGAAAAGATACCTATCTGAATCATGTCTTTATAGATATGAGGC
122341 AAAAGACTAGTCTTGTCCCTTTTCTATGGATAGACTTAGTACAGAAATATCTATACTCCG

```
122401  TATGAGGGAAAATTCTGAGGTTATTCTTGACTCACACCTAAAGATTTGGAAATGAGATTA
122401  ATACTCCCTTTTAAGACTCCAATAAGAACTGAGTGTGGATTTCTAAACCTTTACTCTAAT

122461  GCAGCAAAGCTTTGCCCTACATCTCATGTCAGAATTTTCTGTTTCTTTCTAGTCTTTGAG
122461  CGTCGTTTCGAAACGGGATGTAGAGTACAGTCTTAAAAGACAAAGAAAGATCAGAAACTC

122521  TGTATGTGTGTTCTCACACACGCCATAATGAAATGCATATTATATATAATTATGTGTATA
122521  ACATACACACAAGAGTGTGTGCGGTATTACTTTACGTATAATATATTTAATACACATAT

122581  TATAATATTCTATGACTATACATGACATGTTCCTTTAGCTGATTGCTGTTAAGAGAAATT
122581  ATATTATAAGATACTGATATGTACTGTACAAGGAAATCGACTAACGACAATTCTCTTTAA

122641  TATAGGTTTTTATTTTTCTTGTTTTGTTGGGTATTAAGGAAGAGAAATTCTATGGTAATT
122641  ATATCCAAAAATAAAAAGAACAAAACAACCCATAATTCCTTCTCTTTAAGATACCATTAA

122701  TTCATGTGGCACAGTAATCTGGCATATATGTTGATTTTTTTCCTACACCCATTTGTTGTG
122701  AAGTACACCGTGTCATTAGACCGTATATACAACTAAAAAAAGGATGTGGGTAAACAACAC

122761  ATACCAAGTTTGAAAACAACAGATTTCAGTGGTTGCTTGGGAAACCACAGAACCATGACT
122761  TATGGTTCAAACTTTTGTTGTCTAAAGTCACCAACGAACCCTTTGGTGTCTTGGTACTGA

122821  TGGGGAGAGACAGGATGATTAGGTGGGAAAGCACCCTTTTGGTGGGGCTGTAAAGACTTT
122821  ACCCCTCTCTGTCCTACTAATCCACCCTTTCGTGGGAAAACCACCCCGACATTTCTGAAA

122881  TATATTTAGCAAAATTGGCTACAAAGTCCATTCCCCTCCTTTTCTTGCCTTGATTTGGTA
122881  ATATAAATCGTTTTAACCGATGTTTCAGGTAAGGGGAGGAAAAGAACGGAACTAAACCAT

122941  GAGGGATAGACTTGGATACAAACTAGAATGGATTCATTCTTCTCTGGAGTTAGTGTAACA
122941  CTCCCTATCTGAACCTATGTTTGATCTTACCTAAGTAAGAAGAGACCTCAATCACATTGT

123001  AGACATTTAGCTGCTCAACACAAAAACAGAAACAAAAAAATTGTGTGGTTTCAGCAGTGC
123001  TCTGTAAATCGACGAGTTGTGTTTTTGTCTTTGTTTTTTTAACACACCAAAGTCGTCACG

123061  TATACAATTACTTTTTCTGACCTTTAATGGAGAGAAACACCACTTCTTTGGTCCCTACCA
123061  ATATGTTAATGAAAAAGACTGGAAATTACCTCTCTTTGTGGTGAAGAAACCAGGGATGGT

123121  TCAGCTTCATAGGGTTTTCATCCTGTTCTGTTTCTGGGAGGGCGTAACTGGCCATGCACA
123121  AGTCGAAGTATCCCAAAAGTAGGACAAGACAAAGACCCTCCCGCATTGACCGGTACGTGT

123181  AGTTTTTTTTCTCTAATCAGAGTATGTGCCACTTCTGACCACCAGTAGATGAAAACGAAT
123181  TCAAAAAAAAGAGATTAGTCTCATACACGGTGAAGACTGGTGGTCATCTACTTTTGCTTA

123241  GGAAACCAGGCTATTATATGATACATATCCATTACAAAATAAGACATGAAACTCAAAGGT
123241  CCTTTGGTCCGATAATATACTATGTATAGGTAATGTTTTATTCTGTACTTTGAGTTTCCA

123301  ACTTTATGGTATAATGGGGCATATATTCCTGGACAATTCTTAATGGTCACAGATTTTATA
123301  TGAAATACCATATTACCCCGTATATAAGGACCTGTTAAGAATTACCAGTGTCTAAAATAT

123361  AAAGGACTATTAGTAAATGTATGAATTACAGAGTAATTTATCCTTCTGTTAGTAAGAACC
123361  TTTCCTGATAATCATTTACATACTTAATGTCTCATTAAATAGGAAGACAATCATTCTTGG

123421  AGCTGATGACCTCAGTGTCAGGTGCATCGTGGAAGGTGTTGGGACCTTCCCTTGCCACCA
123421  TCGACTACTGGAGTCACAGTCCACGTAGCACCTTCCACAACCCTGGAAGGGAACGGTGGT

123481  CCCTCACCAGCCATCATCAGCCATAACCTGCACATTGGGGAAGTTTTGACTTATCCCTCA
123481  GGGAGTGGTCGGTAGTAGTCGGTATTGGACGTGTAACCCCTTCAAAACTGAATAGGGAGT

123541  CTTTTGCCCCTCTTCAAGCTGTTCTTTCCACAGTGAATGAGAAGGCCACTTCTTCCTTCA
123541  GAAAACGGGGAGAAGTTCGACAAGAAAGGTGTCACTTACTCTTCCGGTGAAGAAGGAAGT

123601  AACCTTTCAGTGGTTTCCATTTTCCTTTAGACAAAGTCTCTGCCTAGCTGGCCTCTGCCT
123601  TTGGAAAGTCACCAAAGGTAAAAGGAAATCTGTTTCAGAGACGGATCGACCGGAGACGGA

123661  GCCCCTCCTGCCTACCTCTCGAGCACTGCCCCCACCTAGGGCTCTGGTTCCCCAACCTTC
123661  CGGGGAGGACGGATGGAGAGCTCGTGACGGGGGTGGATCCCGAGACCAAGGGGTTGGAAG

123721  ACTCGGTCCTGCCACACCTCCCAGCCCCTTCTCCCTTCAGAACTTTCCTTCTTGTTGTCC
123721  TGAGCCAGGACGGTGTGGAGGGTCGGGGAAGAGGGAAGTCTTGAAAGGAAGAACAACAGG

123781  CCAACACTGGGACACAAAACCCTCCTTATCAACCCTCCTTATCTGGCTGACTCTTACAAG
123781  GGTTGTGACCCTGTGTTTTGGGAGGAATAGTTGGGAGGAATAGACCGACTGAGAATGTTC
```

123841 ATCAGAAACCTGTGTAATGCTCTCATGGCACGCTCCCCTTGTCTTCGTGGATTTCTCAGA
123841 TAGTCTTTGGACACATTACGAGAGTACCGTGCGAGGGGAACAGAAGCACCTAAAGAGTCT

123901 TGGGAAGGAATTATCCATGCAATCACACATAAACTTCTACCTACCCTCCCCTAGTAGCTG
123901 ACCCTTCCTTAATAGGTACGTTAGTGTGTATTTGAAGATGGATGGGAGGGGATCATCGAC

123961 TCTGCTGCTAAGGATGGGGACCATTCTCACTTACTCACTGTTCTGTCCCTCTGCCCAGTC
123961 AGACGACGATTCCTACCCCTGGTAAGAGTGAATGAGTGACAAGACAGGGAGACGGGTCAG

124021 CAGATGTGTTGAAGGATGGAAATATACAGAGTAGTGGTAAAATATAAACCGTTCAGACAT
124021 GTCTACACAACTTCCTACCTTTATATGTCTCATCACCATTTTATATTTGGCAAGTCTGTA

124081 TCCAAGGATGGGCTCATGTGCTTTGACTCATTAATGTACCACTGCTGAAAACAGAACACA
124081 AGGTTCCTACCCGAGTACACGAAACTGAGTAATTACATGGTGACGACTTTTGTCTTGTGT

124141 GCCGCAGTCTTGCCAGTAAGAGTGCAGTTACTGTAATTAATGAATTTGCTAATTAAGCCA
124141 CGGCGTCAGAACGGTCATTCTCACGTCAATGACATTAATTACTTAAACGATTAATTCGGT

124201 TGATTTCATACTGAACTTATGACCAACATATTGAGAAGGTGTGTCTTCAAGAAAATTTAT
124201 ACTAAAGTATGACTTGAATACTGGTTGTATAACTCTTCCACACAGAAGTTCTTTTAAATA

124261 TTTTTGTATTAAGATATTTACTCCAAAGCTAATTGAAGAAGCCAAATCTAGGCTCTGGTT
124261 AAAAACATAATTCTATAAATGAGGTTTCGATTAACTTCTTCGGTTTAGATCCGAGACCAA

124321 TCACCATTGCCAGGGAAATGAGCTCATGGACTCCTATGAACTGATGATGTTAGATCAGAA
124321 AGTGGTAACGGTCCCTTTACTCGAGTACCTGAGGATACTTGACTACTACAATCTAGTCTT

124381 GTTTCTCAAGGCCAGGGCCCAATCACTGCTGAGGCGTCAACAGTAGTTCCTTGTACATCA
124381 CAAAGAGTTCCGGTCCCGGGTTAGTGACGACTCCGCAGTTGTCATCAAGGAACATGTAGT

124441 ATAATTCTCATTACTTTTAAAAAATAACAGATGAATAGCAACTATTTTCCCTGTAGCTCC
124441 TATTAAGAGTAATGAAAATTTTTTATTGTCTACTTATCGTTGATAAAAGGGACATCGAGG

124501 CTTGCTGTGCCTCCTACCCTCCACCACATGTTTCTGGGGAGCCCTGCTTCGGGCCTGCCA
124501 GAACGACACGGAGGATGGGAGGTGGTGTACAAAGACCCCTCGGGACGAAGCCCGGACGGT

124561 ACTACAGAGAATTACTTTTGAGTATCCCTTCCACTCTCATCTCAAGACAGAGTTCATCTA
124561 TGATGTCTCTTAATGAAAACTCATAGGGAAGGTGAGAGTAGAGTTCTGTCTCAAGTAGAT

124621 CCTTTGGGTTATTTGTCAAAAATGTGTCATTTTATTACAAAAAATATACAATCATCATGT
124621 GGAAACCCAATAAACAGTTTTTACACAGTAAAATAATGTTTTTTATATGTTAGTAGTACA

124681 ATTTTGATTAAATTTTACACTAGATTATTAAAATTATTAAATACAATTATTAAAATTAAT
124681 TAAAACTAATTTAAAATGTGATCTAATAATTTTAATAATTTATGTTAATAATTTTAATTA

124741 AATTTAACATATCACATATTTTAAATATATTGTATATAATGAATAATAATATAATTATTG
124741 TTAAATTGTATAGTGTATAAAATTTATATAACATATATTACTTATTATTATATTAATAAC

124801 TCTATTTTAATTCAATAAATGTATAGTAAGTTAGCCAGTTGTAAATTACTGAGAACACTC
124801 AGATAAAATTAAGTTATTTACATATCATTCAATCGGTCAACATTTAATGACTCTTGTGAG

124861 TACTGAAAAAGCATCATTTCAAATACACTATTTAAAATATTAAATGAAATACAATAACAT
124861 ATGACTTTTTCGTAGTAAAGTTTATGTGATAAATTTTATAATTTACTTTATGTTATTGTA

124921 AATTAAACTAATCTTTGGTTCCCCTATTTATGTATTCATTTATCCAACAAAATCTCCTTA
124921 TTAATTTGATTAGAAACCAAGGGGATAAATACATAAGTAAATAGGTTGTTTTAGAGGAAT

124981 AGTGCTTATAATGGGTAGGTCCTGGCTCGGTGTCCCCTAGACAGACGCATGGGCCTTCCC
124981 TCACGAATATTACCCATCCAGGACCGAGCCACAGGGGATCTGTCTGCGTACCCGGAAGGG

125041 CCAGCCCGTCAGTATGGTGCAGGTGTGATGTGTCCGCAGGTGTGTGTGTATGTGTGCAGG
125041 GGTCGGGCAGTCATACCACGTCCACACTACACAGGCGTCCACACACACATACACACGTCC

125101 TGTGGGGTCCGCAGGCGTGCTGGGCCCCCAGGCCGTGTTCCCCTTCCCCTCCCCGGTTGT
125101 ACACCCCAGGCGTCCGCACGACCCGGGGGTCCGGCACAAGGGGAAGGGGAGGGGCCAACA

125161 AGATTTCAGCTGTTGCTGCCAGACCTGACCGGTTCCGGAGGTGGCCGCGCCCCACTCACT
125161 TCTAAAGTCGACAACGACGGTCTGGACTGGCCAAGGCCTCCACCGGCGCGGGGTGAGTGA

125221 GTCGCCTGCTTTCCACAGGGGACAAGGGCCTGTCCGACACGCCCTACGACAGCAGCGCCA
125221 CAGCGGACGAAAGGTGTCCCCTGTTCCCGGACAGGCTGTGCGGGATGCTGTCGTCGCGGT

125281 GCTACGAGAAGGAGAACGAAATGATGAAGTCCCACGTGATGGACCAAGCCATCAACAACG

125281 CGATGCTCTTCCTCTTGCTTTACTACTTCAGGGTGCACTACCTGGTTCGGTAGTTGTTGC

125341 CCATCAACTACCTGGGGGCCGAGTCCCTGCGCCCGCTGGTGCAGACGCCCCCGGGCGGTT
125341 GGTAGTTGATGGACCCCCGGCTCAGGGACGCGGGCGACCACGTCTGCGGGGGCCCGCCAA

125401 CCGAGGTGGTCCCGGTCATCAGCCCGATGTACCAGCTGCACAAGCCGCTCGCGGAGGGCA
125401 GGCTCCACCAGGGCCAGTAGTCGGGCTACATGGTCGACGTGTTCGGCGAGCGCCTCCCGT

125461 CCCCGCGCTCCAACCACTCGGCCCAGGACAGCGCCGTGGAGAACCTGCTGCTGCTCTCCA
125461 GGGGCGCGAGGTTGGTGAGCCGGGTCCTGTCGCGGCACCTCTTGGACGACGACGAGAGGT

125521 AGGCCAAGTTGGTGCCCTCGGAGCGCGAGGCGTCCCCGAGCAACAGCTGCCAAGACTCCA
125521 TCCGGTTCAACCACGGGAGCCTCGCGCTCCGCAGGGGCTCGTTGTCGACGGTTCTGAGGT

125581 CGGACACCGAGAGCAACAACGAGGAGCAGCGCAGCGGTCTCATCTACCTGACCAACCACA
125581 GCCTGTGGCTCTCGTTGTTGCTCCTCGTCGCGTCGCCAGAGTAGATGGACTGGTTGGTGT

125641 TCGCCCCGCACGCGCGCAACGGGCTGTCGCTCAAGGAGGAGCACCGCGCCTACGACCTGC
125641 AGCGGGGCGTGCGCGCGTTGCCCGACAGCGAGTTCCTCCTCGTGGCGCGGATGCTGGACG

125701 TGCGCGCCGCCTCCGAGAACTCGCAGGACGCGCTCCGCGTGGTCAGCACCAGCGGGGAGC
125701 ACGCGCGGCGGAGGCTCTTGAGCGTCCTGCGCGAGGCGCACCAGTCGTGGTCGCCCCTCG

125761 AGATGAAGGTGTACAAGTGCGAACACTGCCGGGTGCTCTTCCTGGATCACGTCATGTACA
125761 TCTACTTCCACATGTTCACGCTTGTGACGGCCCACGAGAAGGACCTAGTGCAGTACATGT

125821 CCATCCACATGGGCTGCCACGGCTTCCGTGATCCTTTTGAGTGCAACATGTGCGGCTACC
125821 GGTAGGTGTACCCGACGGTGCCGAAGGCACTAGGAAAACTCACGTTGTACACGCCGATGG

125881 ACAGCCAGGACCGGTACGAGTTCTCGTCGCACATAACGCGAGGGGAGCACCGCTTCCACA
125881 TGTCGGTCCTGGCCATGCTCAAGAGCAGCGTGTATTGCGCTCCCCTCGTGGCGAAGGTGT

125941 TGAGCTAAAGCCCTCCCGCGCCCCCACCCCAGACCCCGAGCCACCCCAGGAAAAGCACAA
125941 ACTCGATTTCGGGAGGGCGCGGGGGTGGGGTCTGGGGCTCGGTGGGGTCCTTTTCGTGTT

126001 GGACTGCCGCCTTCTCGCTCCCGCCAGCAGCATAGACTGGACTGGACCAGACAATGTTGT
126001 CCTGACGGCGGAAGAGCGAGGGCGGTCGTCGTATCTGACCTGACCTGGTCTGTTACAACA

126061 GTTTGGATTTGTAACTGTTTTTTGTTTTTTGTTTGAGTTGGTTGATTGGGGTTTGATTTG
126061 CAAACCTAAACATTGACAAAAAACAAAAAACAAACTCAACCAACTAACCCCAAACTAAAC

126121 CTTTTGAAAAGATTTTTATTTTTAGAGGCAGGGCTGCATTGGGAGCATCCAGAACTGCTA
126121 GAAAACTTTTCTAAAAATAAAAATCTCCGTCCCGACGTAACCCTCGTAGGTCTTGACGAT

126181 CCTTCCTAGATGTTTCCCCAGACCGCTGGCTGAGATTCCCTCACCTGTCGCTTCCTAGAA
126181 GGAAGGATCTACAAAGGGGTCTGGCGACCGACTCTAAGGGAGTGGACAGCGAAGGATCTT

126241 TCCCCTTCTCCAAACGATTAGTCTAAATTTTCAGAGAGAAATAGATAAAACACGCCACAG
126241 AGGGGAAGAGGTTTGCTAATCAGATTTAAAAGTCTCTCTTTATCTATTTTGTGCGGTGTC

126301 CCTGGGAAGGAGCGTGCTCTACCCTGTGCTAAGCACGGGGTTCGCGCACCAGGTGTCTTT
126301 GGACCCTTCCTCGCACGAGATGGGACACGATTCGTGCCCCAAGCGCGTGGTCCACAGAAA

126361 TTCCAGTCCCCAGAAGCAGAGAGCACAGCCCCTGCTGTGTGGGTCTGCAGGTGAGCAGAC
126361 AAGGTCAGGGGTCTTCGTCTCTCGTGTCGGGGACGACACACCCAGACGTCCACTCGTCTG

126421 AGGACAGGTGTGCCGCCACCCAAGTGCCAAGACACAGCAGGGCCAACAACCTGTGCCCAG
126421 TCCTGTCCACACGGCGGTGGGTTCACGGTTCTGTGTCGTCCCGGTTGTTGGACACGGGTC

126481 GCCAGCTTCGAGCTACATGCATCTAGGGCGGAGAGGCTGCACTTGTGAGAGAAAATACTA
126481 CGGTCGAAGCTCGATGTACGTAGATCCCGCCTCTCCGACGTGAACACTCTCTTTTATGAT

126541 TTTCAAGTCATATTCTGCGTAGGAAAATGAATTGGTTGGGGAAAGTCGTGTCTGTCAGAC
126541 AAAGTTCAGTATAAGACGCATCCTTTTACTTAACCAACCCCTTTCAGCACAGACAGTCTG

126601 TGCCCTGGGTGGAGGGAGACGCCGGGCTAGAGCCTTTGGGATCGTCCTGGATTCACTGGC
126601 ACGGGACCCACCTCCCTCTGCGGCCCGATCTCGGAAACCCTAGCAGGACCTAAGTGACCG

126661 TTTGCGGAGGCTGCTCAGATGGCCTGAGCCTCCCGAGGCTTGCTGCCCCGTAGGAGGAGA
126661 AAACGCCTCCGACGAGTCTACCGGACTCGGAGGGCTCCGAACGACGGGGCATCCTCCTCT

126721 CTGTCTTCCCGTGGGCATATCTGGGGAGCCCTGTTCCCCGCTTTTTCACTCCCATACCTT
126721 GACAGAAGGGCACCCGTATAGACCCCTCGGGACAAGGGGCGAAAAAGTGAGGGTATGGAA

126781 TAATGGCCCCCAAAATCTGTCACTACAATTTAAACACCAGTCCCGAAATTTGGATCTTCT
126781 ATTACCGGGGGTTTTAGACAGTGATGTTAAATTTGTGGTCAGGGCTTTAAACCTAGAAGA

126841 TTCTTTTTGAATCTCTCAAACGGCAACATTCCTCAGAAACCAAAGCTTTATTTCAAATCT
126841 AAGAAAAACTTAGAGAGTTTGCCGTTGTAAGGAGTCTTTGGTTTCGAAATAAAGTTTAGA

126901 CTTCCTTCCCTGGCTGGTTCCATCTAGTACCAGAGGCCTCTTTTCCTGAAGAAATCCAAT
126901 GAAGGAAGGGACCGACCAAGGTAGATCATGGTCTCCGGAGAAAAGGACTTCTTTAGGTTA

126961 CCTAGCCCTCATTTTAATTATGTACATCTGTTTGTAGCCACAAGCCTGAATTTCTCAGTG
126961 GGATCGGGAGTAAAATTAATACATGTAGACAAACATCGGTGTTCGGACTTAAAGAGTCAC

127021 TTGGTAAGTTTCTTTACCTACCCTCACTATATATTATTCTCGTTTTAAAACCCATAAAGG
127021 AACCATTCAAAGAAATGGATGGGAGTGATATATAATAAGAGCAAAATTTTGGGTATTTCC

127081 AGTGATTTAGAACAGTCATTAATTTTCAACTCAATGAAATATGTGAAGCCCAGCATCTCT
127081 TCACTAAATCTTGTCAGTAATTAAAAGTTGAGTTACTTTATACACTTCGGGTCGTAGAGA

127141 GTTGCTAACACACAGAGCTCACCTGTTTGAAACCAAGCTTTCAAACATGTTGAAGCTCTT
127141 CAACGATTGTGTGTCTCGAGTGGACAAACTTTGGTTCGAAAGTTTGTACAACTTCGAGAA

127201 TACTGTAAAGGCAAGCCAGCATGTGTGTCCACACATACATAGGATGGCTGGCTCTGCACC
127201 ATGACATTTCCGTTCGGTCGTACACACAGGTGTGTATGTATCCTACCGACCGAGACGTGG

127261 TGTAGGATATTGGAATGCACAGGGCAATTGAGGGACTGAGCCAGACCTTCGGAGAGTAAT
127261 ACATCCTATAACCTTACGTGTCCCGTTAACTCCCTGACTCGGTCTGGAAGCCTCTCATTA

127321 GCCACCAGATCCCCTAGGAAAGAGGAGGCAAATGGCACTGCAGGTGAGAACCCCGCCCAT
127321 CGGTGGTCTAGGGGATCCTTTCTCCTCCGTTTACCGTGACGTCCACTCTTGGGGCGGGTA

127381 CCGTGCTATGACATGGAGGCACTGAAGCCCGAGGAAGGTGTGTGGAGATTCTAATCCCAA
127381 GGCACGATACTGTACCTCCGTGACTTCGGGCTCCTTCCACACACCTCTAAGATTAGGGTT

127441 CAAGCAAGGGTCTCCTTCAAGATTAATGCTATCAATCATTAAGGTCATTACTCTCAACCA
127441 GTTCGTTCCCAGAGGAAGTTCTAATTACGATAGTTAGTAATTCCAGTAATGAGAGTTGGT

127501 CCTAGGCAATGAAGAATATACCATTTCAAATATTTACAGTACTTGTCTTCACCAACACTG
127501 GGATCCGTTACTTCTTATATGGTAAAGTTTATAAATGTCATGAACAGAAGTGGTTGTGAC

127561 TCCCAAGGTGAAATGAAGCAACAGAGAGGAAATTGTACATAAGTACCTCAGCATTTAATC
127561 AGGGTTCCACTTTACTTCGTTGTCTCTCCTTTAACATGTATTCATGGAGTCGTAAATTAG

127621 CAAACAGGGGTTCTTAGTCTCAGCACTATGACATTTTGGGCTGACTACTTATTTGTTAGG
127621 GTTTGTCCCCAAGAATCAGAGTCGTGATACTGTAAAACCCGACTGATGAATAAACAATCC

127681 CGGGAGCTCTCCTGTGCATTGTAGGATAATTAGCAGTATCCCTGGTGGCTACCCAATAGA
127681 GCCCTCGAGAGGACACGTAACATCCTATTAATCGTCATAGGGACCACCGATGGGTTATCT

127741 CGCCAGTAGCACCCCGAATTGACAACCCAAACTCTCCAGACATCACCAACTGTCCCCTGC
127741 GCGGTCATCGTGGGGCTTAACTGTTGGGTTTGAGAGGTCTGTAGTGGTTGACAGGGGACG

127801 GAGGAGAAATCACTCCTGGGGGAGAACCACTGACCCAAATGAATTCTAAACCAATCAAAT
127801 CTCCTCTTTAGTGAGGACCCCCTCTTGGTGACTGGGTTTACTTAAGATTTGGTTAGTTTA

127861 GTCTGGGAAGCCCTCCAAGAAAAAAAAATAGAAAAGCACTTGAAGAATATTCCCAATATTC
127861 CAGACCCTTCGGGAGGTTCTTTTTTTTATCTTTTCGTGAACTTCTTATAAGGGTTATAAG

127921 CCGGTCAGCAGTATCAAGGCTGACTTGTGTTCATGTGGAGTCATTATAAATTCTATAAAT
127921 GGCCAGTCGTCATAGTTCCGACTGAACACAAGTACACCTCAGTAATATTTAAGATATTTA

127981 CAATTATTCCCCTTCGGTCTTAAAAATATATTTCCTCATAAACATTTGAGTTTTGTTGAA
127981 GTTAATAAGGGGAAGCCAGAATTTTTATATAAAGGAGTATTTGTAAACTCAAAACAACTT

128041 AAGATGGAGTTTACAAAGATACCATTCTTGAGTCATGGATTTCTCTGCTCACAGAAGGGT
128041 TTCTACCTCAAATGTTTCTATGGTAAGAACTCAGTACCTAAAGAGACGAGTGTCTTCCCA

128101 GTGGCATTTGGAAACGGGAATAAACAAAATTGCTGCACCAATGCACTGAGTGAAGGAAGA
128101 CACCGTAAACCTTTGCCCTTATTTGTTTTAACGACGTGGTTACGTGACTCACTTCCTTCT

128161 GAGACAGAGGATCAAGGGCTTTAGACAGCACTCCTTCAATATGCAATCACAGAGAAAGAT
128161 CTCTGTCTCCTAGTTCCCGAAATCTGTCGTGAGGAAGTTATACGTTAGTGTCTCTTTCTA

```
128221  GCGCCTTATCCAAGTTAATATCTCTAAGGTGAGAGCCTTCTTAGAGTCAGTTTGTTGCAA
128221  CGCGGAATAGGTTCAATTATAGAGATTCCACTCTCGGAAGAATCTCAGTCAAACAACGTT

128281  ATTTCACCTACTCTGTTCTTTTCCATCCATCCCCCTGAGTCAGTTGGTTGAAGGGAGTTA
128281  TAAAGTGGATGAGACAAGAAAAGGTAGGTAGGGGGACTCAGTCAACCAACTTCCCTCAAT

128341  TTTTTTCAAGTGGAATTCAAACAAAGCTCAAACCAGAACTGTAAATAGTGATTGCAGGAA
128341  AAAAAAGTTCACCTTAAGTTTGTTTCGAGTTTGGTCTTGACATTTATCACTAACGTCCTT

128401  TTCTTTTCTAAACTGCTTTGCCCTTTCCTCTCACTGCCTTTTATAGCCAATATAAATGTC
128401  AAGAAAAGATTTGACGAAACGGGAAAGGAGAGTGACGGAAAATATCGGTTATATTTACAG

128461  TCTTTGCACACCTTTTGTTGTGGTTTTATATTGTAACACCATTTTTCTTTGAAACTATTG
128461  AGAAACGTGTGGAAAACAACACCAAAATATAACATTGTGGTAAAAAGAAACTTTGATAAC

128521  TATTTAAAGTAAGGTTTCATATTATGTCAGCAAGTAATTAACTTATGTTTAAAAGGTGGC
128521  ATAAATTTCATTCCAAAGTATAATACAGTCGTTCATTAATTGAATACAAATTTTCCACCG

128581  CATATCATGTACCAAAAGTTGCTGAAGTTTCTCTTCTAGCTGGTAAAGTAGGAGTTTGCA
128581  GTATAGTACATGGTTTTCAACGACTTCAAAGAGAAGATCGACCATTTCATCCTCAAACGT

128641  TGACTTCACACTTTTTTTGCGTAGTTTCTTCTGTTGTATGATGGCGTGAGTGTGTGTCTT
128641  ACTGAAGTGTGAAAAAAACGCATCAAAGAAGACAACATACTACCGCACTCACACACAGAA

128701  GGGTACCGCTGTGTACTACTGTGTGCCTAGATTCCATGCACTCTCGTTGTGTTTGAAGTA
128701  CCCATGGCGACACATGATGACACACGGATCTAAGGTACGTGAGAGCAACACAAACTTCAT

128761  AATATTGGAGACCGGAGGGTAACAGGTTGGCCTGTTGATTACAGCTAGTAATCGCTGTGT
128761  TTATAACCTCTGGCCTCCCATTGTCCAACCGGACAACTAATGTCGATCATTAGCGACACA

128821  CTTGTTCCGCCCCCTCCCTGACACCCCAGCTTCCCAGGATGTGGAAAGCCTGGATCTCAG
128821  GAACAAGGCGGGGGAGGGACTGTGGGGTCGAAGGGTCCTACACCTTTCGGACCTAGAGTC

128881  CTCCTTGCCCCATATCCCTTCTGTAATTTGTACCTAAAGAGTGTGATTATCCTAATTCAA
128881  GAGGAACGGGGTATAGGGAAGACATTAAACATGGATTTCTCACACTAATAGGATTAAGTT

128941  GAGTCACTAAAACTCATCACATTATCATTGCATATCAGCAAAGGGTAAAGTCCTAGCACC
128941  CTCAGTGATTTTGAGTAGTGTAATAGTAACGTATAGTCGTTTCCCATTTCAGGATCGTGG

129001  AATTGCTTCACATACCAGCATGTTCCATTTCCAATTTAGAATTAGCCACATAATAAAATC
129001  TTAACGAAGTGTATGGTCGTACAAGGTAAAGGTTAAATCTTAATCGGTGTATTATTTTAG

129061  TTAGAATCTTCCTTGAGAAAGAGCTGCCTGAGATGTAGTTTTGTTATATGGTTCCCCACC
129061  AATCTTAGAAGGAACTCTTTCTCGACGGACTCTACATCAAAACAATATACCAAGGGGTGG

129121  GACCATTTTTGTGCTTTTTTCTTGTTTTGTTTTGTTTTGACTGCACTGTGAGTTTTGTAG
129121  CTGGTAAAAACACGAAAAAAGAACAAAACAAAACAAAACTGACGTGACACTCAAAACATC

129181  TGTCCTCTTCTTGCCAAAACAAACGCGAGATGAACTGGACTTATGTAGACAAATCGTGAT
129181  ACAGGAGAAGAACGGTTTTGTTTGCGCTCTACTTGACCTGAATACATCTGTTTAGCACTA

129241  GCCAGTGTATCCTTCCTTTCTTCAGTTCCAGCAATAATGAATGGTCAACTTTTTTAAAAT
129241  CGGTCACATAGGAAGGAAAGAAGTCAAGGTCGTTATTACTTACCAGTTGAAAAAATTTTA

129301  CTAGATCTCTCTCATTCATTTCAATGTATTTTTACTTTAAGATGAACCAAAATTATTAGA
129301  GATCTAGAGAGAGTAAGTAAAGTTACATAAAAATGAAATTCTACTTGGTTTTAATAATCT

129361  CTTATTTAAGATGTACAGGCATCAGAAAAAAGAAGCACATAATGCTTTTGGTGCGATGGC
129361  GAATAAATTCTACATGTCCGTAGTCTTTTTTCTTCGTGTATTACGAAAACCACGCTACCG

129421  ACTCACTGTGAACATGTGTAACCACATATTAATATGCAATATTGTTTCCAATACTTTCTA
129421  TGAGTGACACTTGTACACATTGGTGTATAATTATACGTTATAACAAAGGTTATGAAAGAT

129481  ATACAGTTTTTTATAATGTTGTGTGTGGTGATTGTTCAGGTCGAATCTGTTGTATCCAGT
129481  TATGTCAAAAAATATTACAACACACACCACTAACAAGTCCAGCTTAGACAACATAGGTCA

129541  ACAGCTTTAGGTCTTCAGCTGCCCTTCTGGCGAGTACATGCACAGGATTGTAAATGAGAA
129541  TGTCGAAATCCAGAAGTCGACGGGAAGACCGCTCATGTACGTGTCCTAACATTTACTCTT

129601  ATGCAGTCATATTTCCAGTCTGCCTCTATGATGATGTTAAATTATTGCTGTTTAGCTGTG
129601  TACGTCAGTATAAAGGTCAGACGGAGATACTACTACAATTTAATAACGACAAATCGACAC

129661  AACAAGGGATGTACCACTGGAGGAATAGAGTATCCTTTTGTACACATTTTGAAATGCTTC
```

129661 TTGTTCCCTACATGGTGACCTCCTTATCTCATAGGAAAACATGTGTAAAACTTTACGAAG

129721 TTCTGTAGTGATAGAACAAATAAATGCAACGAATACTCTGTCTGCCCTATCCCGTGAAGT
129721 AAGACATCACTATCTTGTTTATTTACGTTGCTTATGAGACAGACGGGATAGGGCACTTCA

129781 CCACACTGGCGTAAGAGAAGGCCCAGCAGAGCAGGAATCTGCCTAGACTTTCTCCCAATG
129781 GGTGTGACCGCATTCTCTTCCGGGTCGTCTCGTCCTTAGACGGATCTGAAAGAGGGTTAC

129841 AGATCCCAATATGAGAGGGAGAAGAGATGGGCCTCAGGACAGCTGCAATACCACTTGGGA
129841 TCTAGGGTTATACTCTCCCTCTTCTCTACCCGGAGTCCTGTCGACGTTATGGTGAACCCT

129901 ACACATGTGGTGTCTTGATGTGGCCAGCGCAGCAGTTCAGCACAACGTACCTCCCATCTA
129901 TGTGTACACCACAGAACTACACCGGTCGCGTCGTCAAGTCGTGTTGCATGGAGGGTAGAT

129961 CAACAGTGCTGGACGTGGGAATTCTAAGTCCCAGTCTTGAGGGTGGGTGGAGATGGAGGG
129961 GTTGTCACGACCTGCACCCTTAAGATTCAGGGTCAGAACTCCCACCCACCTCTACCTCCC

130021 CAACAAGAGATACATTTCCAGTTCTCCACTGCAGCATGCTTCAGTCATTCTGTGAGTGGC
130021 GTTGTTCTCTATGTAAAGGTCAAGAGGTGACGTCGTACGAAGTCAGTAAGACACTCACCG

130081 CGGGCCCAGGGCCCTCACAATTTCACTACCTTGTCTTTTACATAGTCATAAGAATTATCC
130081 GCCCGGGTCCCGGGAGTGTTAAAGTGATGGAACAGAAAATGTATCAGTATTCTTAATAGG

130141 TCAACATAGCCTTTTGACGCTGTAAATCTTGAGTATTCATTTACCCTTTTCTGATCTCCT
130141 AGTTGTATCGGAAAACTGCGACATTTAGAACTCATAAGTAAATGGGAAAAGACTAGAGGA

130201 GGAAACAGCTGCCTGCCTGCATTGCACTTCTCTTCCCGAGGAGTGGGGTAAATTTAAAAG
130201 CCTTTGTCGACGGACGGACGTAACGTGAAGAGAAGGGCTCCTCACCCCATTTAAATTTTC

130261 TCAAGTTATAGTTTGGATGTTAGTATAGAATTTTGAAATTGGGAATTAAAAATCAGGACT
130261 AGTTCAATATCAAACCTACAATCATATCTTAAAACTTTAACCCTTAATTTTTAGTCCTGA

130321 GGGGACTGGGAGACCAAAAATTTCTGATCCCATTTCTGATGGATGTGTCACACCTTTTCT
130321 CCCCTGACCCTCTGGTTTTTAAAGACTAGGGTAAAGACTACCTACACAGTGTGGAAAAGA

130381 GTCAAAATAAAATGTCTTGGAGGTTATGACTCCTTGGTGAA
130381 CAGTTTTATTTTACAGAACCTCCAATACTGAGGAACCACTT