



US 20110287423A1

(19) **United States**

(12) **Patent Application Publication**  
**MAHIET et al.**

(10) **Pub. No.: US 2011/0287423 A1**

(43) **Pub. Date: Nov. 24, 2011**

(54) **DIAGNOSIS OF VIRAL INFECTIONS BY  
DETECTION OF GENOMIC AND  
INFECTIOUS VIRAL DNA BY MOLECULAR  
COMBING**

(75) Inventors: **CHARLOTTE MAHIET, PARIS  
(FR); FABRICE SALVAIRE,  
BENZONS (FR); EMMANUEL  
CONSEILLER, PARIS (FR);  
SEBASTIEN BARRADEAU,  
PARIS (FR)**

(73) Assignee: **GENOMIC VISION, PARIS (FR)**

(21) Appl. No.: **13/092,210**

(22) Filed: **Apr. 22, 2011**

**Related U.S. Application Data**

(60) Provisional application No. 61/327,397, filed on Apr. 23, 2010.

**Publication Classification**

(51) **Int. Cl.**  
**C12Q 1/68** (2006.01)  
**C12N 7/00** (2006.01)  
**G01N 33/53** (2006.01)  
(52) **U.S. Cl. .... 435/6.11; 436/501; 435/235.1**

(57) **ABSTRACT**

A method for detecting in vitro the presence of a genome of a DNA virus or a viral derived DNA in an infected eukaryotic cell, tissue or biological fluid using Molecular Combing or other nucleic acid stretching methods together with probes, especially nucleic acid probes, having a special design. A method for monitoring in vitro the effects of anti-viral treatment by following the presence of genomic viral or viral derived DNA polynucleotides in a virus-infected cell, tissue or biological fluid. Detection of an infectious form of a virus using Molecular Combing and DNA hybridization. A kit comprising probes used to carry out these methods and a composition comprising the probes.

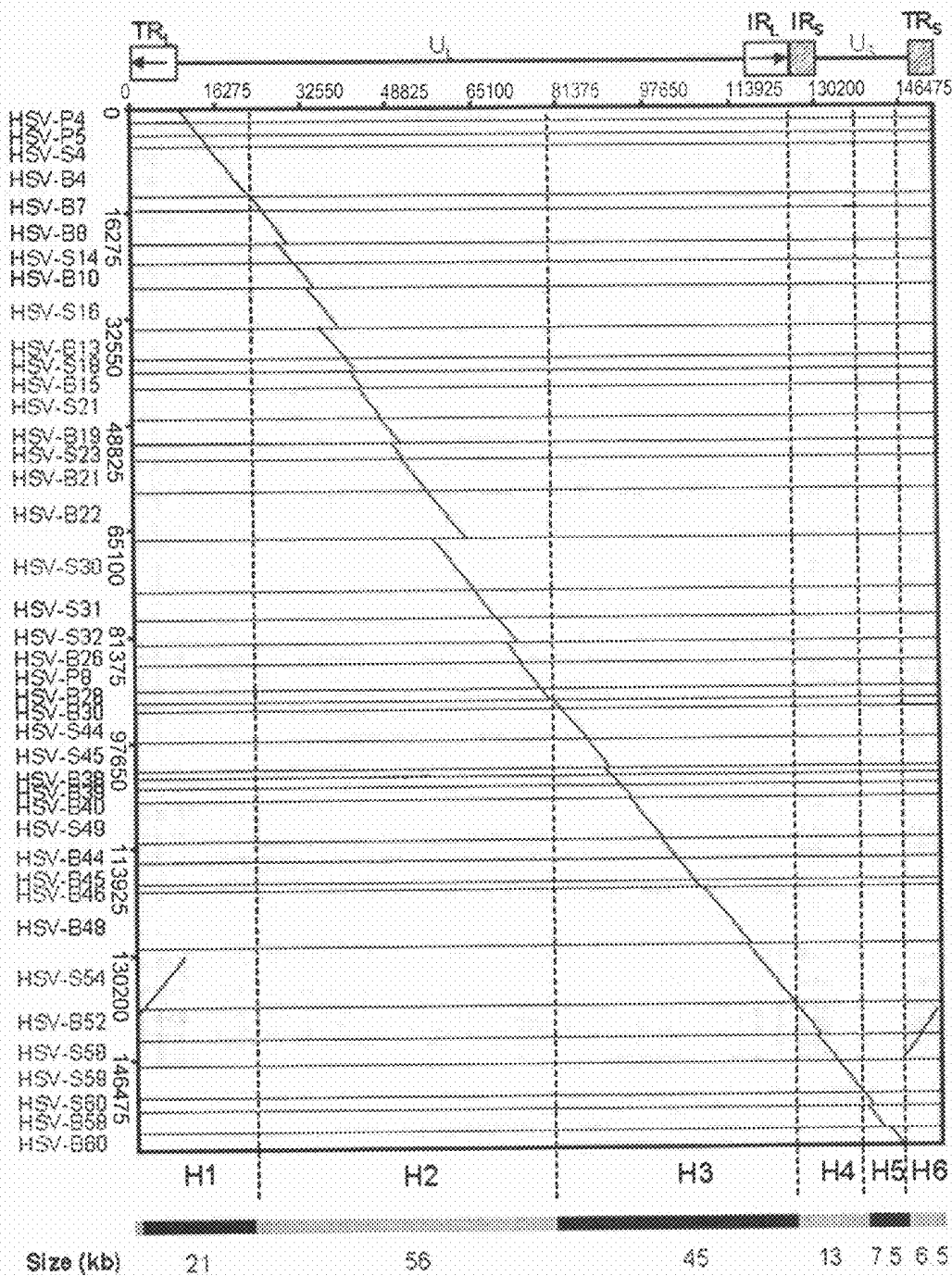
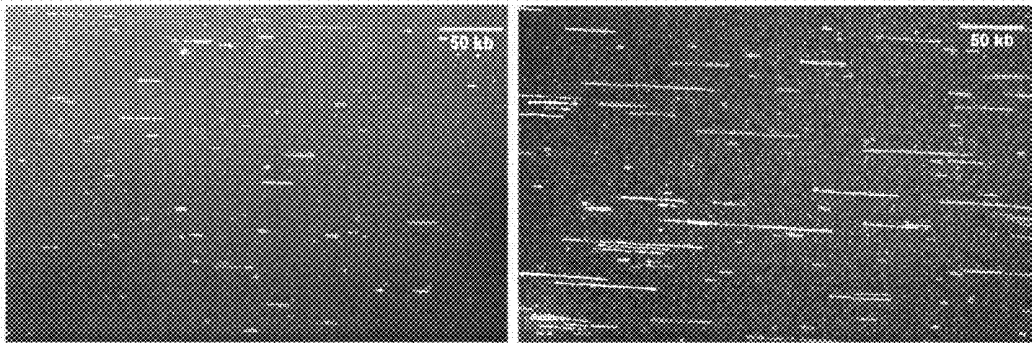


Figure 1

A



B

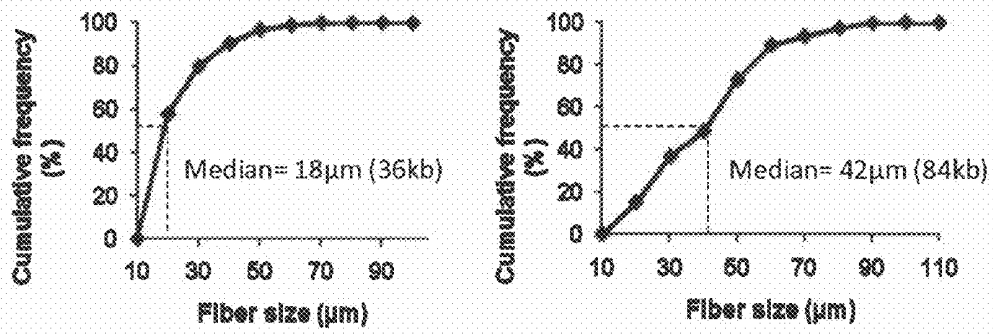


Figure 2

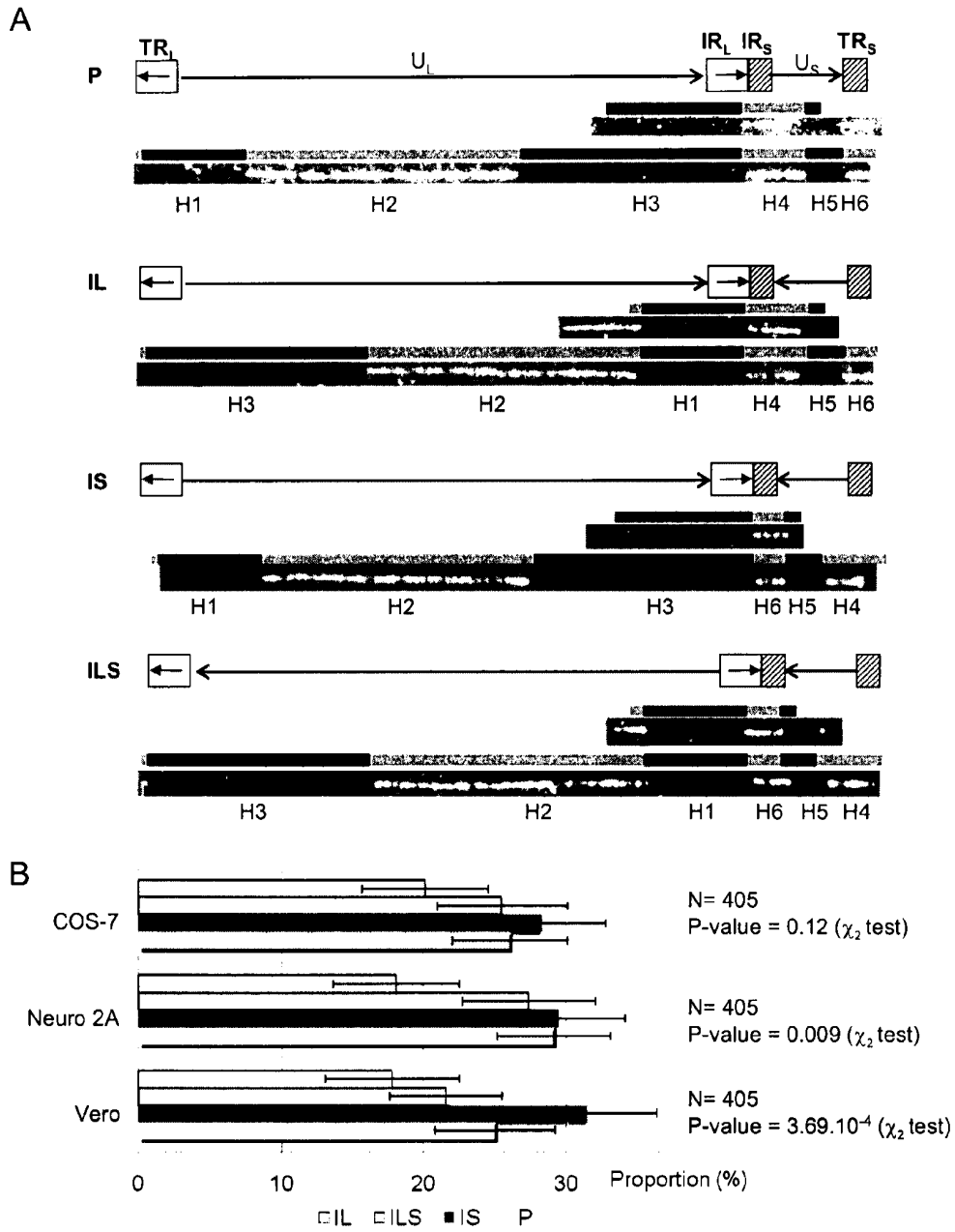


Figure 3

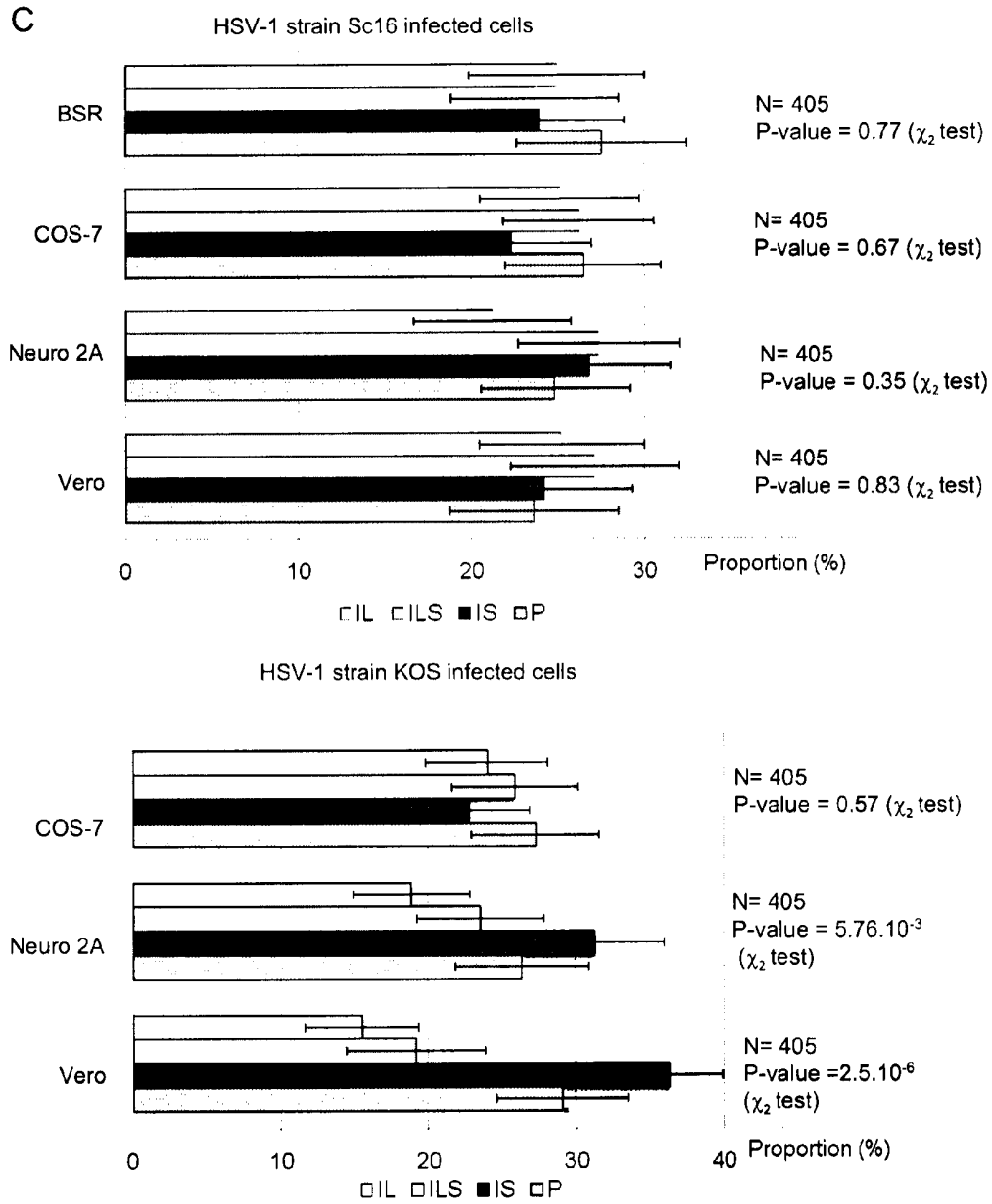
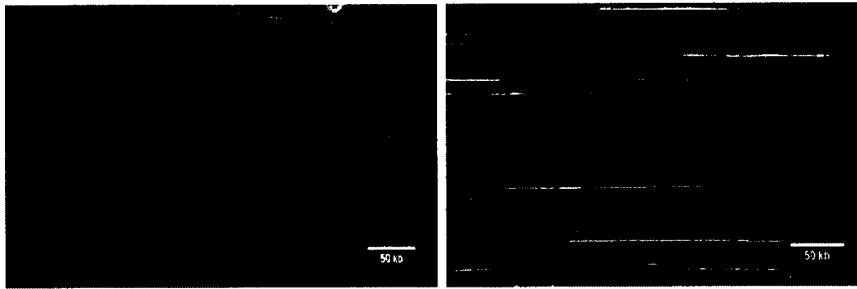


Figure 3

A



B

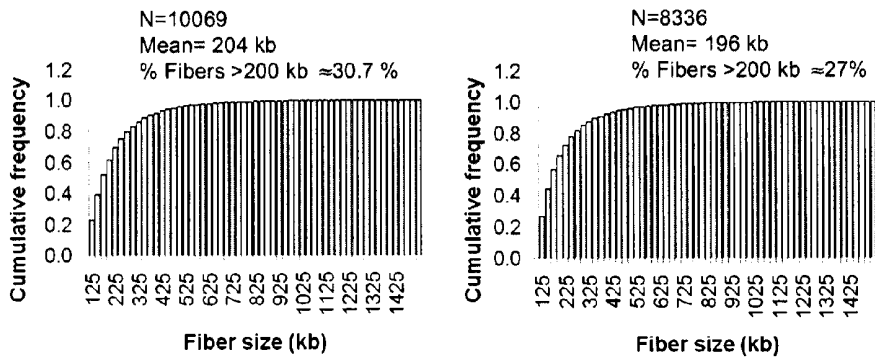


Figure 4

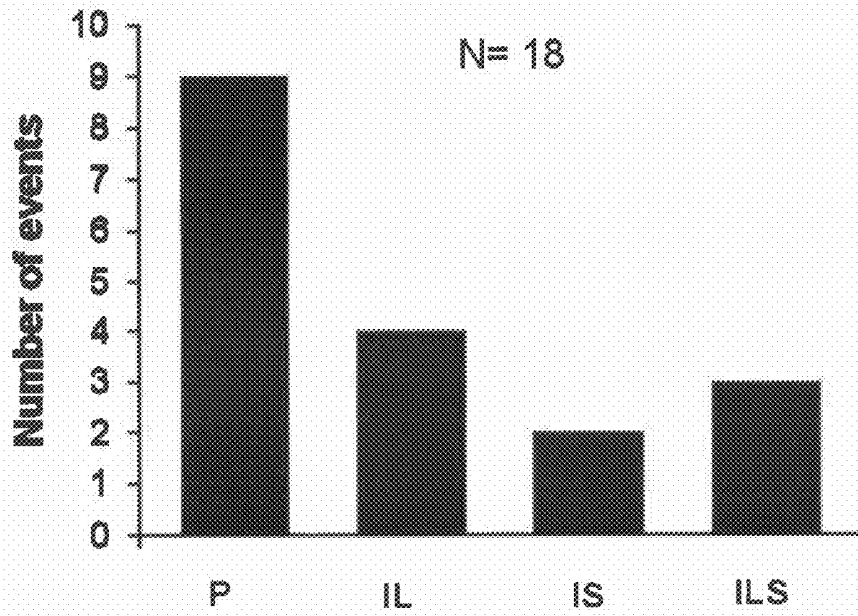


Figure 5

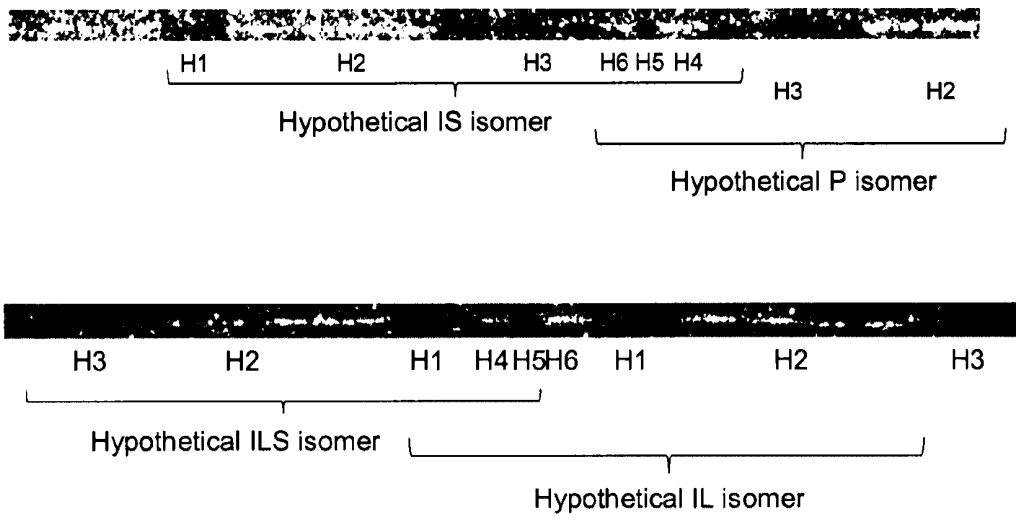


Figure 6

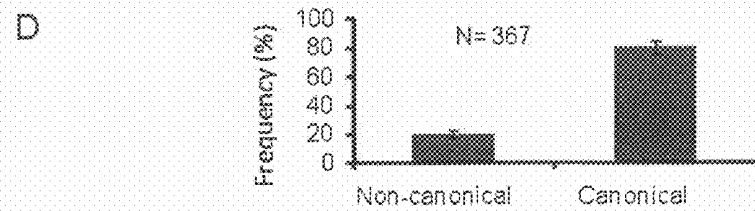
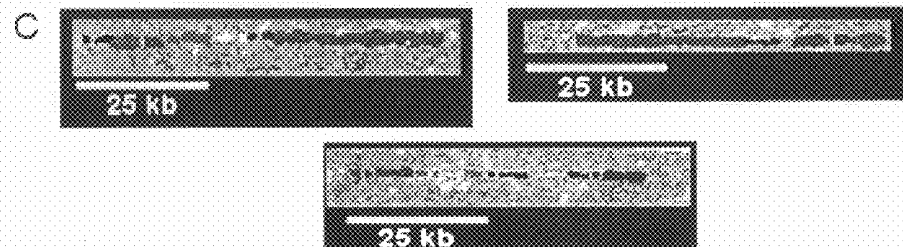
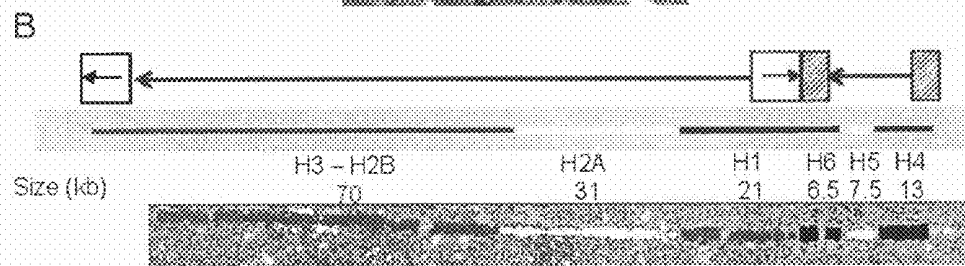
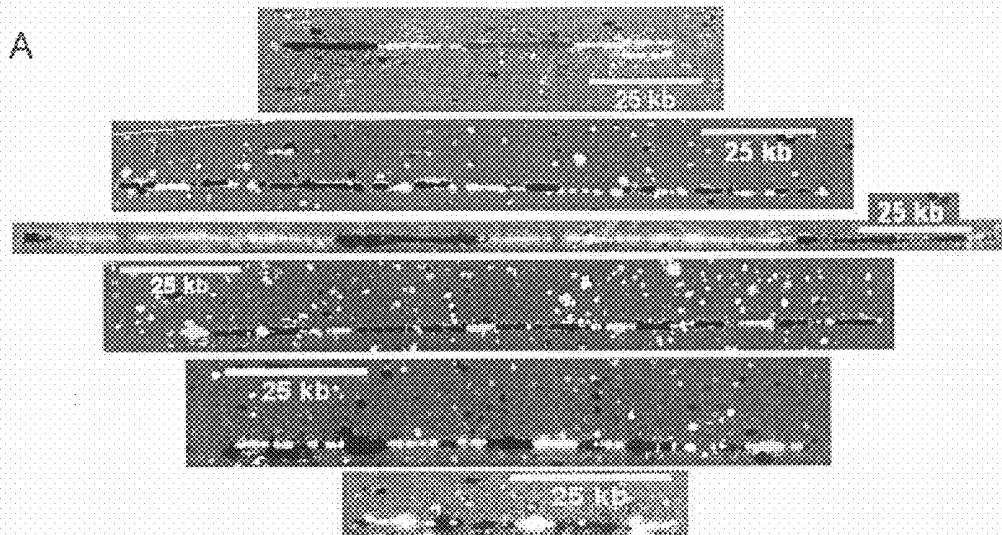
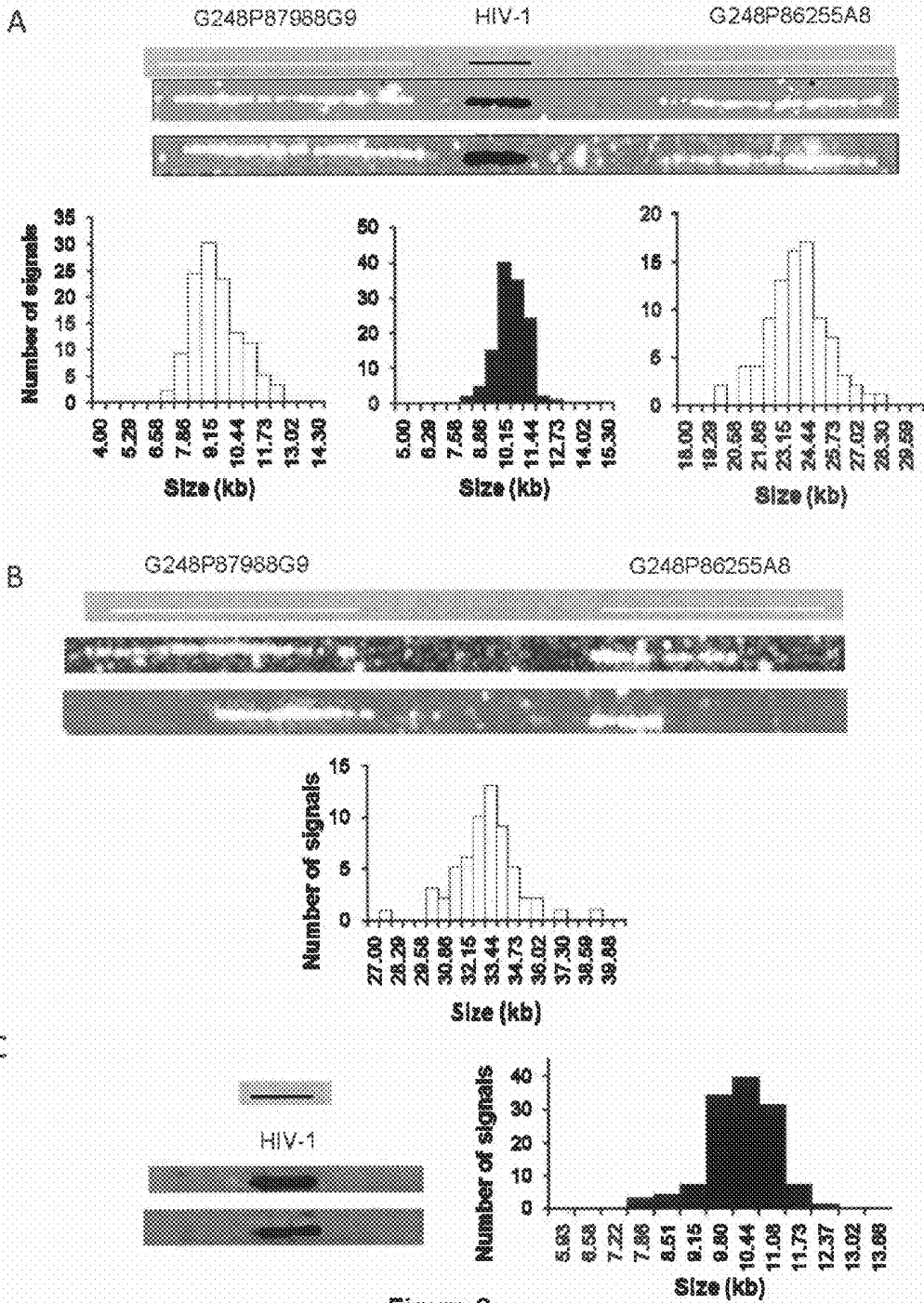


Figure 7



D

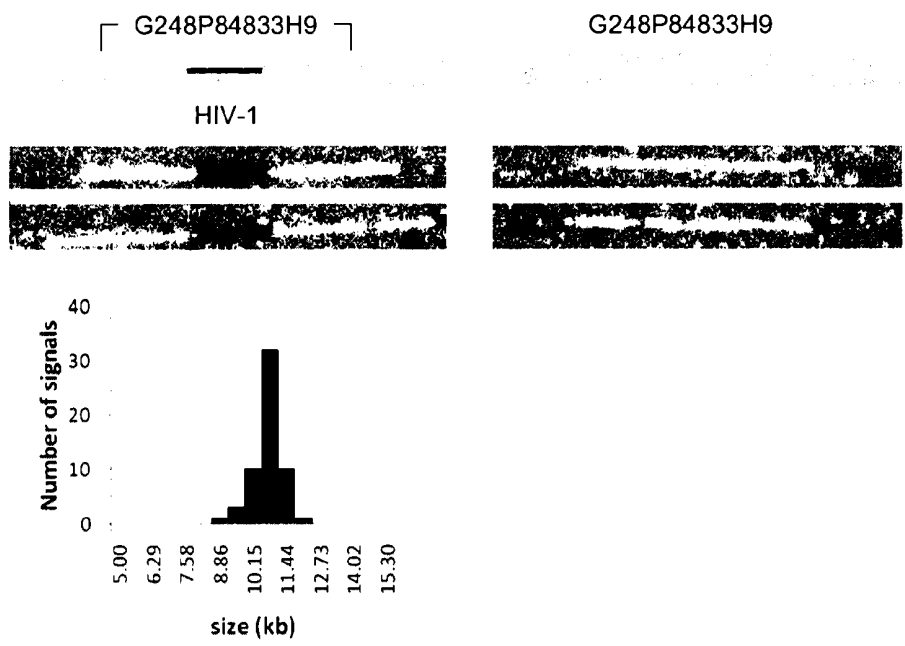


Figure 8

**DIAGNOSIS OF VIRAL INFECTIONS BY  
DETECTION OF GENOMIC AND  
INFECTIOUS VIRAL DNA BY MOLECULAR  
COMBING**

CROSS-REFERENCE TO RELATED  
APPLICATIONS

**[0001]** This application claims priority under 35 U.S.C. 119(e) to U.S. Provisional Application 61/327,397 filed Apr. 23, 2010 which is incorporated by reference in its entirety.

STATEMENT REGARDING FEDERALLY  
SPONSORED RESEARCH OR DEVELOPMENT

**[0002]** (None)

REFERENCE TO MATERIAL ON COMPACT  
DISK

**[0003]** (None)

BACKGROUND OF THE INVENTION

**[0004]** 1. Field of the Invention

**[0005]** A method for easily, rapidly and accurately detecting the presence of infectious viral DNA or other infectious or pathogenic DNA sequences in a viral host cell, tissue or biological fluid obtained from a subject or patient using Molecular Combing and/or DNA stretching in combination with specially designed probes for the infectious viral or pathogenic DNA sequence. A method for monitoring in vitro the effects of anti-viral treatment by following the presence of genomic viral polynucleotides in a virus-infected cell, tissue or a fluid of a patient to be tested as infected by a virus. Detection of infectious virus based on DNA hybridization. A kit comprising the probes used to carry out a method of the invention and a composition comprising said probes.

**[0006]** 2. Description of the Related Art

**[0007]** Molecular Combing

**[0008]** Molecular Combing technology has been disclosed in various patents and scientific publications, for example in U.S. Pat. No. 6,303,296, WO 9818959, WO 0073503, U.S. 2006/257910, U.S. 2004/033510, U.S. Pat. No. 6,130,044, U.S. Pat. No. 6,225,055, U.S. 6,054,327, WO 2008/028931, WO 2010/035140, and Michalet, Ekong et al. 1997, Herrick, Michalet et al. 2000, Herrick, Stanislawski et al. 2000, Gad, Aurias et al. 2001, Gad, Caux-Moncoutier et al. 2002, Gad, Klinger et al. 2002, Herrick, Jun et al. 2002, Pasero, Bensimon et al. 2002, Gad, Bieche et al. 2003, Lebofsky and Bensimon 2003, Herrick, Conti et al. 2005, Lebofsky and Bensimon 2005, Lebofsky, Heilig et al. 2006, Patel, Arcangioli et al. 2006, Rao, Conti et al. 2007 and Schurra and Bensimon 2009 (see citations below). The techniques, supplies and equipment described in these references, specifically those pertaining or relating to Molecular Combing, are hereby incorporated by reference to the publications cited above.

**[0009]** Bensimon, et al., U.S. Pat. No. 6,303,296, discloses DNA stretching procedures, Lebofsky, et al., WO 2008/028931 discloses detections of genomic sequences, and Lebofsky, et al., WO 2010/035140 A1, discloses a method for analysis of D4Z4 tandem repeat arrays on human chromosomes 4 and 10 based on stretching of nucleic acid and on Molecular Combing.

**[0010]** Molecular Combing is a technique enabling the direct visualization of individual nucleic acid molecules and

has been successfully used for the location of myc and certain Human Papillomavirus (HPV) sequences in cervical cancer cell lines as well as for the study of replication kinetics on previously characterized MYC/papillomavirus (HPV18) amplicon and particularly on its origin of replication (Herrick, et al., *Cancer Res.* 65(4): 1174-1179 2005). Nucleic acid rearrangements and amplifications in tumor cells were correlated with cancer outcome using this technique. Herrick, et al. probed combed nucleic acids from cultured tumor cell lines IC1, IC2, IC4 and IC5 with probes specific to N-myc and c-myc genes and probes to detect HPV nucleic acids containing full-length HPV-18 and HPV-44 (~7.8 kb). Conti, et al., *Genes, Chromosomes & Cancer* 46:724-734 (2007) studied origin activation in an HPV/MYC amplicon containing certain HPV 18 nucleic acid sequences and analyzed origin activity and fork movement on DNA stretched by Molecular Combing from the IC1 tumor cell line. Conti identifies gene rearrangements and amplifications as hallmarks of cancer cells focusing on the mechanisms involving both HPV and MYC nucleic acid sequences. Both these studies used cultured tumor cells in which HPV nucleic acid sequences act as oncogenes but not as infectious genomic viral nucleic acids.

**[0011]** Episomal and integrated Epstein-Barr virus DNA obtained from cultured Burkitt's lymphoma cell lines was characterized by FISH on combed DNA, (Reisinger, et al., *Int. J. Cancer* 118: 1603-1608 (2006). Reisinger's objective was to study the genomic organization of viral genomes and understand and attempt to discriminate between the episomal and integrated forms of EBV in infected eukaryotic cells. DNA from cultured Burkitt's lymphoma cell lines was stretched by Molecular Combing and probed with EBV-specific DNA probes that cover genomic sequences but not the region between the terminal repeat (TR) and the internal repeat 1 (IR1) sequences.

**[0012]** Reisinger et al. and Herrick et al. also do not describe methods for detecting rearrangements of viral genomes using probe sets tagged with different haptens that can generate a different color-fluorescence array that will allow the identification of rearrangements within this specific region of the viral DNA. These two references describe the detection of viral DNA in cultured cells but not in tissue or in biological fluids.

**[0013]** In contrast, in the present invention, the standard extraction procedure is modified to allow the extraction of viral DNA from viral particles in order to analyze by Molecular Combing the structure of the viral DNA in the infectious particles. Moreover, the invention develops a methodology of extraction of high molecular weight DNA from solid tissues such as cornea that is compatible with the analysis by Molecular Combing and that allows the detection of the presence of the infectious forms of HSV DNA in an infected cornea. These two methodologies have never been described before this invention.

**[0014]** Stretching nucleic acid, in particular viral or genomic DNA provides immobilized nucleic acids in linear and parallel strands and is preferably performed with a controlled stretching factor on an appropriate surface (e.g. surface-treated glass slides). After stretching, it is possible to hybridize sequence-specific probes detectable for example by fluorescence microscopy (Lebofsky, Heilig et al. 2006). Thus, a particular sequence may be directly visualized on a single molecule level. The length of the fluorescent signals and/or their number, and their spacing on the slide provides a direct reading of the size and relative spacing of the probes. How-

ever, up to now, it was not shown that this technology can be used to detect the infectious forms of viruses in an infected eukaryotic cell or viral sequences or genome of a derived mutated viral sequence infecting a cell or tissue or present in a fluid of mammals.

**[0015]** Virus Structure and Genome

**[0016]** The viruses that cause human diseases are placed into seven groups based upon the existence of an envelope, the type of the capsid, the nature of the genome (DNA or RNA, double-stranded or single-stranded), size, and target cells (International Committee on Taxonomy of Viruses, <http://www.ictvonline.org>; last accessed Apr. 7, 2011). The enveloped DNA viruses include the poxviruses, herpesviruses, and hepadnaviruses. The nonenveloped group includes the adenoviruses, papovaviruses, and parvoviruses.

**[0017]** An enormous variety of genomic structures can be seen among viral species. Indeed, there are millions of different types of viruses (Breitbart and Rohwer 2005), though only about 5,000 of them have been described in detail (Dimmock, Easton et al. 2007). A viral genome is either single-stranded or double-stranded. Single-stranded genomes consist of an unpaired nucleic acid, analogous to one-half of a ladder split down the middle. All RNA viruses are single-stranded except for the double-stranded reoviruses. Double-stranded genomes consist of two complementary paired nucleic acids, analogous to a ladder. All DNA viruses are double-stranded except for the parvoviruses, which have single-stranded DNA. Some viruses, such as those belonging to the Hepadnaviridae, contain a genome that is partially double-stranded and partially single-stranded (Collier and Oxford 2006)

**[0018]** Viral genomes are circular, as in the polyomaviruses, or linear, as in the adenoviruses (Fields, Knipe et al. 2007). The type of nucleic acid is irrelevant to the shape of the genome. Among RNA viruses, the genome is often divided up into separate parts within the virion and is called segmented. Each segment often codes for one protein and they are usually found together in one capsid. Every segment is not required to be in the same virion for the overall virus to be infectious, as demonstrated by the brome mosaic virus (Collier and Oxford 2006).

**[0019]** For viruses with RNA or single-stranded DNA, the strands are said to be either positive-sense (called the plus-strand) or negative-sense (called the minus-strand), depending on whether it is complementary to the viral messenger RNA (mRNA) (Fields, Knipe et al. 2007). Positive-sense viral RNA is identical to viral mRNA and thus can be immediately translated by the host cell. Negative-sense viral RNA is complementary to mRNA and thus must be converted to positive-sense RNA by an RNA polymerase before translation. DNA nomenclature is similar to RNA nomenclature, in that the coding strand for the viral mRNA is complementary to it (-), and the non-coding strand is a copy of it (+). Some single-stranded RNA viruses called retrovirus (which includes HIV) replicate in the host cell via the enzyme reverse transcriptase to produce DNA from its RNA genome. The DNA is then incorporated into the host's genome by an integrase enzyme. The virus genome, or provirus, remains in the genome of the infected cell and thereafter replicates as part of the host cell's DNA.

**[0020]** Genome size varies greatly amongst different kinds or species of viruses ranging from 1.7 kb for the hepatitis delta virus (HDV) to 309 kb for the Poxvirus. RNA viruses generally have smaller genome sizes than DNA viruses because of a higher error-rate when replicating, and have a maximum

upper size limit. Beyond this limit, errors in the genome when replicating render the virus useless or uncompetitive. To compensate for this, RNA viruses often have segmented genomes where the genome is split into smaller molecules, thus reducing the chance of error. In contrast, DNA viruses generally have larger genomes because of the high fidelity of their replication enzymes (Pressing and Reaney 1984).

**[0021]** Viruses can undergo genetic changes that occur through different mechanisms. The first process is called genetic drift where individual bases in the DNA or RNA mutate to other bases. Most of these point mutations are "silent" meaning that they do not change the protein that the gene encodes but others can confer evolutionary advantages such as resistance to antiviral drugs. Antigenic shift occurs when there is a major change in the genome of the virus (Pan, Li et al. 2007). This can be a result of recombination or reassortment. When this happens with influenza viruses (Hampson and Mackenzie 2006), pandemics might result. RNA viruses often exist as quasispecies or swarms of viruses of the same species but with slightly different genome nucleoside sequences. Such quasispecies are a prime target for natural selection (Metzner 2006). Secondly, segmented genomes confer also evolutionary advantages; different strains of a virus with a segmented genome can shuffle and combine genes and produce progeny viruses that have unique and new characteristics. This is called reassortment or viral sex (Goudsmit 1998). Finally, genetic recombination which is the process by which a strand of DNA is broken and then joined to the end of a different DNA molecule can arise. This can occur when viruses infect cells simultaneously and studies of viral evolution have shown that recombination has been rampant in the species studied (Worobey and Holmes 1999). Recombination is common to both RNA and DNA viruses (Umene 1999; Lukashov 2005).

**[0022]** The HSV are among the larger viruses with a diameter ranging from about 150 nm to 200 nm. They are enclosed within a loosely fitting envelope that contains glycoprotein spikes (Wildy, Russell et al. 1960). Like other enveloped viruses, herpes viruses are prone to deactivation by organic solvents or detergents and are unstable outside the host's body. The icosahedral capsid houses a core of double-stranded DNA that winds around a proteinaceous spindle in some viruses. The genome sequence of HSV-1 strain 17, first published in 1988, contains 152,261 residues in each strand (Genbank accession number: NC\_001806.1; last accessed Apr. 7, 2011) (McGeoch, Dalrymple et al. 1988). The HSV genome is regarded as being composed of two covalently joined segments, the long (L) and short (S) regions (Roizman 1979). The L region consists of a unique sequence ( $U_L$ ) flanked by a pair of oppositely oriented repeat elements (termed  $R_L$ , with the terminal and internal copies specifically referred to as  $TR_L$  and  $IR_L$ ) (Wadsworth, Jacob et al. 1975). The S region similarly consists of  $IR_S$ ,  $U_S$ , and  $TR_S$ . The sequences of the  $R_L$  and  $R_S$  elements are distinct, except that there is a 400 base pairs (bp) direct repeat at the genome termini, termed the  $\alpha$  sequence, and at least one further copy of this is found at the junction of L and S regions, in the orientation opposite to the terminal copies (Roizman 1979; Davison and Wilkie 1981). Each terminus possesses one overhanging residue, with the 3'-hydroxyl group free (Mocarski and Roizman 1982). Preparations of HSV DNA contain equimolar amounts of four sequence-orientation isomers, in which  $U_L$  and  $U_S$  each lie independently in one of two possible orientations with respect to the joint between L and S

sequences (Hayward, Jacob et al. 1975; Bataille and Epstein 1997). One isomer is defined as the prototype (Roizman 1979). HSV-1 DNA has a base composition of 68.3% G+C (McGeoch, Dalrymple et al. 1988). The G+C content is not constant throughout the genome; the 6.6-kbp  $R_S$  elements deviate most notably from the mean, with a G+C content of 79.5% (McGeoch, Dolan et al. 1986). The genome of HSV-2 has not been completely sequenced, but it closely resembles that of HSV-1, albeit with a slightly higher G+C content.

**[0023]** The genomic polynucleotide sequences of numerous viruses are well-known to those of skill in the art and are incorporated by reference to Fields, B. N., D. M. Knipe, et al. (2007), *Fields' Virology*. Philadelphia, Wolters kluwer/Lippincott Williams & Wilkins; and to the NCBI Virus Genomes database accessible at <http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=10239&type=5&name=Viruses> (last accessed Apr. 7, 2011).

**[0024]** Viral Pathogenesis

**[0025]** Many types of viral pathogens are recognized. For a recent review on viral infection, one can refer to Fields, et al. (Fields, Knipe et al. 2007) and references therein which are incorporated by reference. A viral infection occurs when a virus enters the body through such processes as breathing air contaminated with a virus, eating contaminated food, or by having sexual contact with a person who is infected with a virus. A viral infection may also be caused by an insect bite. In a viral infection, the virus invades the inside of the body's cells in order to reproduce. A virus then spreads to other cells and repeats the process. This process of viral infection results in a variety of symptoms that vary in character and severity depending on the type of viral infection and individual factors. Common symptoms of a viral infection include fatigue, flu-like symptoms and fever.

**[0026]** Many types of viral infections, such as a common cold, are self limiting in generally healthy people. This means that the viral infection causes illness for period of time, then it resolves and symptoms disappear. However, some people are at risk for developing serious complications of viral infection. In addition, certain types of viral infections, such as HIV/AIDS, are not self limiting and cause serious complications and/or can eventually be fatal. As an example of serious damage, some forms of measles virus infection lead to brain's inflammation, and consequently permanent brain injury.

**[0027]** There are many types of viruses that cause a wide variety of viral infections or viral diseases. For example, there are over 200 different viruses that can cause a cold or an upper respiratory infection. Other common viruses include the influenza virus which causes influenza or the flu. The Epstein-Barr virus and the cytomegalovirus cause infectious mononucleosis, the varicella zoster virus (VZV) causes shingles and chickenpox, and HIV causes AIDS.

**[0028]** Viruses have different mechanisms by which they produce disease in an organism, which largely depends on the viral species. Mechanisms at the cellular level primarily include cell lysis, the breaking open and subsequent death of the cell. In multicellular organisms, if enough cells die the whole organism will start to suffer the effects. Although viruses cause disruption of healthy homeostasis, resulting in disease, they can also exist relatively harmlessly within an organism. An example would include the ability of the herpes simplex virus (HSV), which causes cold sores, to remain in a dormant state within the human body. This is called latency (Margolis, Elfman et al. 2007) and is a characteristic of the all herpes viruses including the EBV, which causes glandular

fever, and the VZV. Most people have been infected with at least one of these types of HSV (Whitley and Roizman 2001). However, these latent viruses might sometimes be beneficial, as the presence of the virus can increase immunity against bacterial pathogens, such as *Yersinia pestis* (Barton, White et al. 2007). On the other hand, latent chickenpox infections return in later life as the disease called shingles.

**[0029]** Some viruses can cause life-long or chronic infections, where the viruses continue to replicate in the body despite the host's defense mechanisms (Bertoletti and Gehring 2007). This is common in hepatitis B and C virus infections. People chronically infected are known as carriers, as they serve as reservoirs of infectious virus. In populations with a high proportion of carriers, the disease is said to be endemic (Rodrigues, Deshmukh et al. 2001). In contrast to acute lytic viral infections this persistence implies compatible interactions with the host organism.

**[0030]** HSV-1 infection is one example of a viral infection. HSV was named for the tendency of some herpes infections to produce a creeping rash (Taylor, Brockman et al. 2002). It is the common name for a large family whose members include herpes simplex 1 and 2, the cause of fever blisters and genital infections; VZV, the cause of chickenpox and shingles; cytomegalovirus (CMV), which affects the salivary glands and other viscera; EBV, associated with infection of the lymphoid tissue; and some recently identified viruses (HSV-6, -7, and -8). Prominent features of the family are its tendency toward viral latency and recurrent infections.

**[0031]** HSV infections, often colloquially called herpes, usually target the mucous membranes. The virus enters cracks or cuts in the membrane surface and then multiplies in basal and epithelial cells in the immediate vicinity. This results in inflammation, edema, cell lysis, and formation of a characteristic thin-walled vesicle. The main diseases of HSV are facial herpes (oral, optic, and pharyngeal), genital herpes, neonatal herpes, and disseminated disease. Herpes labialis otherwise known as fever blisters or cold sores is the most common recurrent HSV-1 infection. Vesicles usually crop up on the muco-cutaneous junction of the lips or on adjacent skin. Herpetic keratitis (also called ocular herpes) is an infective inflammation of the eye in which a latent virus travels into the ophthalmic rather than the mandibular branch of the trigeminal nerve. Preliminary symptoms are a gritty feeling in the eye, conjunctivitis, sharp pain, and sensitivity to light. Some patients develop characteristic branched or opaque corneal lesions as well. In 25% to 50% of cases, keratitis is recurrent and chronic and can interfere with vision.

**[0032]** The primary infection, asymptomatic in more than 90% of cases (Liesegang 1989) takes place in the oral mucosae, as a consequence of contact with infected particles of saliva. Soon after infection, the linear viral genome circularizes and DNA replication initiates at an origin (for review: (Boehmer and Lehman 1997). DNA replication initially proceeds by a theta mechanism and subsequently switches to a sigma or rolling-circle mode to yield long head-to-tail concatemers. Multiple DNA replication forks that arise from homologous recombination, a sequence-mediated genome isomerization, and other events, lead to the formation of an extensive network of branched DNA intermediates. Finally, these structures are resolved into unit-length genomes and packaged into preassembled capsids. Nevertheless, DNA molecules shorter than the full-length standard HSV-1 viral DNA can become encapsidated within nuclear capsids provided they contain the cleavage/packaging signal (Vlazny,

Kwong et al. 1982). However, capsids containing HSV-1 genome significantly shorter than standard viral genome are not infective.

**[0033]** After replication in epithelial tissues, viruses propagate in neurons before becoming latent. Type 1 HSV enters primarily the trigeminal, or fifth cranial, nerve, which has extensive innervations in the oral region while type 2 HSV usually becomes latent in the ganglion of the lumbo-sacral spinal nerve trunk. Following various triggering factors such as fever, UV radiation, stress, or mechanical injury, the virus migrates to the body surface and produces a local skin or membrane lesion, often in the same site as a previous infection. Since the principal location of latent HSV-1 is the trigeminal ganglion, responsible for sensory innervation of the face, most recurrences are located in the eyes or the lips. The seroprevalence of HSV-1 in the general population ranges from 24.5% to 67%, with 30 to 70% of positive subjects experiencing recurrent herpes labialis (Liesegang 2001). Indeed, almost everyone may potentially exhibit an HSV-1 infection event since recent studies using Polymerase Chain Reaction (PCR) on post-mortem tissues showed that nearly 100% of people at least 60 years of age have several HSV-1 genome copies present in the trigeminal ganglion (Wang, Lau et al. 2005). The eye, and particularly the cornea, is the second most frequent location of HSV-1 infection. The prevalence of herpes keratitis is 149/100,000 (Liesegang, Melton et al. 1989), with more than 30 new events per 100,000 inhabitants annually (Labetoulle, Auquier et al. 2005). Once the eye has been clinically affected by an herpetic infection, the rate of relapse is 23% within 2 years and 40% within 5 years of follow-up (Wilhelmus, Coster et al. 1981; Liesegang 1989). The visual prognosis is poor, especially in deep corneal infections (stromal keratitis), with up to 60% of affected eyes reaching a visual acuity of less than 20/40 after 5 years of follow-up (Kabra, Lalitha et al. 2006). As a consequence, HSV-1 is a leading cause of blindness throughout the world. In less developed countries, HSV-1 ocular infection accounts for about 10% of patients attending corneal clinics (Pramod, Rajendran et al. 1999), and despite the use of topical or oral antiviral agents in the last three decades, herpes simplex keratitis also remains the most frequent cause of infectious corneal opacities in the most developed areas of the world (group 1998), accounting for about 10% of patients undergoing corneal transplantation (Leger, Larroque et al. 2001). Moreover, the natural risk of HSV-1 reactivation in recently grafted cornea is about 25% in the first year following surgery (Lomholt, Baggesen et al. 1995), and about 33% of primary graft failure (no clearing of the graft) have shown to be associated with the presence of the HSV-1 genome in the cornea (Cockerham, Bijwaard et al. 2000). These data explain why HSV-1 remains a major cause of corneal graft failure, accounting for about 22% of all cases of re-grafting.

**[0034]** The clinical complications of latency and recurrent infections become more severe with advancing age, cancer chemotherapy, or other conditions that compromise the immune defenses. The HSV are among the most common and serious opportunists among AIDS patients (Ramaswamy and Geretti 2007). Nearly 95% of this group will experience recurrent bouts of skin, mucous membrane, intestinal, and eye disease from these viruses.

**[0035]** Human immunodeficiency virus (HIV) is a lentivirus (a member of the retrovirus family) that causes acquired immunodeficiency syndrome (AIDS) (Weiss 1993; Douek, Roederer et al. 2009), a condition in humans in which the

immune system begins to fail, leading to life-threatening opportunistic infections. Infection with HIV occurs by the transfer of blood, semen, vaginal fluid, pre-ejaculate, or breast milk. Within these bodily fluids, HIV is present as both free virus particles and virus within infected immune cells. The four major routes of transmission are unsafe sex, contaminated needles, breast milk, and transmission from an infected mother to her baby at birth (vertical transmission). Screening of blood products for HIV has largely eliminated transmission through blood transfusions or infected blood products in the developed world.

**[0036]** HIV infection in humans is considered pandemic by the World Health Organization (WHO). From its discovery in 1981 to 2006, AIDS killed more than 25 million people. HIV infects about 0.6% of the world's population. In 2005 alone AIDS claimed an estimated 2.4-3.3 million lives of which more than 570,000 were children. Antiretroviral treatment reduces both the mortality and the morbidity of HIV infection.

**[0037]** HIV infects primarily vital cells in the human immune system such as helper T cells (to be specific, CD4<sup>+</sup> T cells), macrophages, and dendritic cells. HIV infection leads to low levels of CD4<sup>+</sup> T cells through three main mechanisms: first, direct viral killing of infected cells; second, increased rates of apoptosis in infected cells; and third, killing of infected CD4<sup>+</sup> T cells by CD8 cytotoxic lymphocytes that recognize infected cells. When CD4<sup>+</sup> T cell numbers decline below a critical level, cell-mediated immunity is lost, and the body becomes progressively more susceptible to opportunistic infections.

**[0038]** Most people infected with HIV-1 left untreated eventually develop AIDS and eventually die from opportunistic infections or malignancies associated with the progressive failure of the immune system (Lawn 2004). HIV progresses to AIDS at a variable rate affected by viral, host, and environmental factors; most will progress to AIDS within 10 years of HIV infection: some will have progressed much sooner, and some will take much longer (Buchbinder, Katz et al. 1994; 2000). Treatment with anti-retroviral drugs increases the life expectancy of people infected with HIV and even after HIV has progressed to diagnosable AIDS, the average survival time with antiretroviral therapy was estimated to be more than 5 years as of 2005 (Schneider, Gange et al. 2005). Without antiretroviral therapy, someone who has AIDS typically dies within a year (Morgan, Mahe et al. 2002).

**[0039]** Diagnosis of and Treatment of Viral Infections

**[0040]** Viral diseases are diagnosed by a variety of clinical and laboratory methods. These include a clinical assessment of symptoms, viral isolation in cell or animal culture, and serological testing for antibodies to a virus. Making a diagnosis of a viral infection begins with taking a thorough personal and family medical history, including symptoms, and completing a physical examination. Diagnosing some viral infections, such as seasonal influenza, may be made based on a history and physical. Blood tests, such as a complete blood count may be done. A complete blood count measures the numbers of different types of blood cells, including white blood cells (WBCs). Different types of WBCs increase in number in characteristic ways during an infectious process, such as viral infection. A culture test may also be performed. This test requires taking a small sample from the body area that is suspected to be infected with a virus and grows the sample in a lab to determine the type of microorganism causing illness. Common samples tested with a culture include

those from the throat, blood, and sputum from the lungs. Diagnostic tests may also include a lumbar puncture, also called a spinal tap, which involves withdrawing a small sample of cerebrospinal fluid (CSF) from the spine with a needle. The sample of CSF is tested for white blood cells and other indications of viral infection that may be in the spine or brain, such as viral meningitis. X-rays may be performed to assist in the diagnosis of some viral infections. This may include taking a chest X-ray for suspected cases of viral pneumonia. Additional tests may be performed in order to rule out or confirm other diseases that may accompany viral infection or cause similar symptoms, such as a secondary bacterial infection.

**[0041]** HSV infections are diagnosed by a variety of clinical and laboratory methods, including a clinical assessment of symptoms, isolation in cell or animal culture, and serological testing for antibodies to HSV. Small, painful, vesiculating lesions on the mucous membranes of the mouth or genitalia, lymphadenopathy, and exudate are typical diagnostic symptoms of herpes simplex. Further diagnostic support is available by examining scrapings from the base of such lesions stained with Giemsa, Wright (also called Tzanck preparation), or Papanicolaou (Pap) methods. The presence of multinucleate cells, giant cells, and intranuclear eosinophilic inclusion bodies can help establish herpes infection. This method, however, will not distinguish among HSV-1, HSV-2, or other herpesviruses, which require more specific subtyping. Moreover, laboratory culture and specific tests are essential for diagnosing immunosuppressed and neonatal patients with severe, disseminated herpes infection. A specimen of tissue or fluid is introduced into a primary cell line such as monkey kidney or human embryonic kidney tissue cultures and is then observed for cytopathic effects within 24 to 48 hours. Direct tests on specimens or cell cultures using fluorescent antibodies or detection of DNA using specific probes or amplification by Polymerase Chain Reaction (PCR) can differentiate among HSV-1, HSV-2, and closely related HSV, but are not useful tools to confirm that the genomes of said viruses found in a sample are intact or infectious. While serological analysis is useful for primary infection, it is inconclusive for recurrent illness because the antibody titer to HSV upon recurrence of the infection usually does not increase.

**[0042]** Several agents are available for treatment of HSV infections. Acyclovir is the most effective therapy developed to date that is nontoxic and highly specific to HSV. Famciclovir and valacyclovir are alternate drugs. Topical medications applied to genital and oral lesions cut the length of infection and reduce viral shedding. Systemic therapy is available for more serious complications such as herpes keratitis and disseminated herpes. New evidence indicates that a daily dose of oral acyclovir taken for a period of 6 months to one year can be effective in preventing recurrent genital herpes. Over-the-counter cold sore medications containing menthol, camphor, and local anesthetics lessen pain and may protect against secondary bacterial infections, but they probably do not affect the progress of the viral infection. Some protection in suppressing cold sores can be obtained from the amino acid lysine, taken orally in the earliest phases of recurrence. However, current treatments do not reduce the load of the DNA matrix, and thus are unable to reduce the risk of further viral reactivation. Ideally, an ultimate weapon against HSV-1 infection should be durably present in the cells, avoid questions of sensitivity and, if possible, reduce the load of viral

genomes. New antiviral treatments with the aim to eliminate the virus DNA by the action of endonucleases are in development.

**[0043]** Virus DNA detection of HSV in cornea have been performed from tear fluid (Fukuda, Deai et al. 2008), corneal scrapings (El-Aal, El Sayed et al. 2006) and corneal explants button from healthy and diseased patients (Crouse, Pflugfelder et al. 1990) using PCR-based methods (Llorente, Hidalgo et al. 1998; Gonzalez-Villasenor 1999; Kessler, Muhlbauer et al. 2000; O'Neill, Wyatt et al. 2003; Kimura, Ihira et al. 2005; Namvar, Olofsson et al. 2005; Strick and Wald 2006; Susloparov, Susloparov et al. 2006; Engelmann, Petzold et al. 2008; Sugita, Shimizu et al. 2008; Tanaka, Kogawa et al. 2009; Wada, Mizoguchi et al. 2009; Yu, Shi et al. 2009) or DNA/DNA hybridization on infected cells (Nago, Hayashi et al. 1988; Kotronias and Kapranos 1998). However, none of these methods allowed for both the detection and the nature of an infectious HSV genome in infected cells in a single analysis.

**[0044]** HSV is one type of viral disease for which a large spectrum of technologies for the diagnosis of HSV infection has been disclosed in various patents and publications as for example in EP0139416, EP0263025, WO0202131, WO2004036185, and in Matsumoto, Yamada et al. 1992, Kotronias and Kapranos 1998, Kessler, Muhlbauer et al. 2000, Namvar, Olofsson et al. 2005, Susloparov et al. 2006, Sugita, Shimizu et al. 2008, and Yu, Shi et al. 2009.

**[0045]** The most reliable method of diagnosis of the HSV infectious disease is to isolate the virus for determination, but this method required culture cells and requires several days for the determination. There are also immunological methods but they are mostly unreliable and difficult to perform, especially during the latency phase. A large number of PCR-based methods have been developed for enhanced sensitivity and faster time to result than is possible by conventional means but these methods do not allow analysis of the whole HSV genome and cannot predict the infectivity of the sample containing nucleotide sequences of the virus because the complete genome of the said virus is not tested. Direct methods of detection of the HSV genome in cells in situ hybridization (Nago, Hayashi et al. 1988; Kotronias and Kapranos 1998) or dot blot DNA-DNA hybridization (Matsumoto, Yamada et al. 1992) have been reported but these methods are not able to determine the type of HSV. There are instances in which rapid, sensitive, and specific diagnosis of HSV disease is imperative. There is therefore, a clinical need to develop a rapid and sensitive tool to aid in the diagnosis of HSV. There also remains a need for a tool for the typing of the HSV infection. Rapid identification of the specific etiological agent involved in a viral infection provides information which can be used to determine appropriate therapy within a short period of time.

**[0046]** Analysis of the configuration of HSV-1 genome following lytic or latent infection has been performed by the gel electrophoresis system of Gardella et al. (Gardella, Medveczky et al. 1984). In these gels circular molecules characteristically exhibit lower mobilities than the corresponding linear forms do and this has allowed the detection of episomal forms of latent herpes virus genomes. However, this technique does not discriminate between the isomer of HSV genome in infected cells or tissues. Other techniques like pulse-field gel electrophoresis (PFGE) have been used to provide insights into the mechanism of HSV DNA replication (Severini, Morgan et al. 1994). This family of tests is highly time-consuming

and in all cases requires southern blotting, implying manipulation of radioactivity, long migration and/or exposure times.

**[0047]** Another virus for which numerous detection and analytic methods have been reported is Human Immunodeficiency Virus, or HIV, including HIV-1 and HIV-2. HIV-1 load in blood plasma, as measured by the number of copies of HIV-1 RNA, is a major laboratory marker widely used in clinical practice. Higher virus loads are directly linked to more rapid progression to AIDS in HIV-1-infected individuals. The effectiveness of highly active antiretroviral therapy (HAART) is also assessed by measuring the HIV-1 load in plasma. A patient is considered to be successfully treated by HAART when HIV-1 load in plasma stays below the detection limit of commercial assays which is currently 50 copies of HIV-1 RNA per ml of plasma. However, in spite of its clinical success, HAART cannot eradicate the virus, mainly due to the persistence of various viral reservoirs including latently infected resting CD4<sup>+</sup> cells (Hermankova, Siliciano et al. 2003; Siliciano, Kajdas et al. 2003). Recent studies demonstrated that both virus replication and evolution do continue in some patients even when HIV-1 RNA in plasma is undetectable and therapy is otherwise considered to be successful (Frenkel, Wang et al. 2003; Havlir, Strain et al. 2003; Ramratnam, Ribeiro et al. 2004; Chun, Nickle et al. 2005; Tobin, Learn et al. 2005). HAART failure as a result of development of drug-resistant HIV-1 strains is a common problem (del Rio 2006). Thus, special attention should be given to characterizing HIV-1 residual replication by studying its molecular markers in peripheral blood mononuclear cells (PBMC). In particular, the amounts of cell-associated HIV-1 RNA and proviral DNA should be quantified. Of these, the amounts of proviral DNA may reflect the size of the pool of latently infected cells. However, systematic studies of the relationships between the cellular HIV-1 RNA/DNA levels and therapy outcome are hindered by the extremely low copy numbers of HIV-1 RNA/DNA in PBMC under HAART. Therefore, development of highly sensitive methods for quantification of cellular forms of HIV-1 RNA/DNA is essential.

**[0048]** Real-time reverse transcription-PCR (RT-PCR) is currently the preferred method for quantification of HIV-1 RNA/DNA in cells (Douek, Brenchley et al. 2002; Fischer, Joos et al. 2004). However, despite their accuracy and specificity, single-step real-time RT-PCR methods using the TaqMan detection chemistry are unable to reliably quantify <100 copies of HIV-1 RNA/DNA target per reaction in the context of total cellular RNA/DNA (Espy, Uhl et al. 2006). This evokes the possibility of yielding false-negative results when PBMC material from patients under HAART is studied, especially when limited amounts of clinical material are available for analysis. Methods that use SYBR green-based detection chemistry to detect HIV-1 RNA/DNA may be more sensitive (Espy, Uhl et al. 2006) but are prone to false-positive results because DNA binding dyes do not bind in a sequence-specific manner. With a theoretical detection limit of one molecule per reaction, nested PCR is considered a more sensitive method than real-time PCR. However, only semi-quantitative data can be produced with this method. In addition, it requires labor-intensive and time-consuming experimental procedures. In contrast as described below the inventors have discovered that Molecular Combing allows the detection of unique events and can be used as a highly sensitive method for the detection and quantification of the HIV proviral DNA in the infected cells of patients.

**[0049]** To diagnose viral infection, the use of antigen detection assays is on the increase, especially those based on monoclonal antibodies. Rapid nucleic acid detection using specific probes directed against DNA or RNA or amplification methods are other options for several viruses. These have the advantage of being so sensitive that they could conceivably detect the viral nucleic acid in a single infected cell. For example, PCR technique which allows the enzymatic amplification of minute quantities of DNA often undetectable by other methods, has been widely used for detection of part of genome of several viral agents such as HSV (Rowley, Whitley et al. 1990), human immunodeficiency virus (HIV) (Ou, Kwok et al. 1988) and human papilloma viruses (HPV) (Shibata, Arnheim et al. 1988).

**[0050]** A significant limitation of PCR is that it does not allow one to confirm the integrity of the complete genome detected and consequently cannot be a standard of characterization of complete infectious viral genome found in a tested sample. PCR thus also lacks the specificity required for testing the efficiency of an antiviral treatment against infectious virus particles because its results do not indicate whether the viral polynucleotides detected are infectious. Serological methods do not directly detect infectious viral polynucleotides, such as chromosomally-integrated viral genomic DNA, and require that a subject mount an immune response to a virus prior to detection or that a sufficient amount of viral antigen be present in a sample.

**[0051]** Detection of Viral Oncogenes and Activation of Proto-Oncogenes by Virus

**[0052]** Many cancers originate from a viral infection; this is especially true in animals such as birds, but also in humans. Worldwide, the WHO International Agency for Research on Cancer estimated that in the year 2002 20% of human cancers were caused by infection of which 10-15% are caused by one of seven different viruses (Carrillo-Infante, Abbadessa et al. 2007). However, only a minority of persons or animals will go on to develop cancers after infection. Tumor viruses come in a variety of forms: viruses with a DNA genome, such as HPV (cervical cancer), Hepatitis B virus (liver cancer), and EBV (a type of lymphoma), and viruses with an RNA genome, like the Hepatitis C virus (HCV) can cause cancers, as can retroviruses having both DNA and RNA genomes (Human T-lymphotropic virus and hepatitis B virus, which normally replicates as a mixed double and single-stranded DNA virus but also has a retroviral replication component).

**[0053]** A direct oncogenic viral mechanism involves either insertion of additional viral oncogenic genes into the host cell (acutely-transforming virus) or to enhance already existing oncogenic genes (proto-oncogenes) in the genome (slowly-transforming virus) (for review, (Parsonnet 1999)). In acutely-transforming viruses the viral particles carry a gene that encodes for an overactive oncogene called viral-oncogene (v-*onc*) and the infected cell is transformed as soon as v-*onc* is expressed. In contrast, in slowly-transforming viruses, the virus genome is inserted, especially as viral genome insertion is obligatory part of retroviruses, near a proto-oncogene in the host genome. The viral promoter or other transcription regulation elements, in turn, cause overexpression of that proto-oncogene, which, in turn, induces uncontrolled cellular proliferation. Because viral genome insertion is not specific to proto-oncogenes and the chance of insertion near that proto-oncogene is low, slowly-transform-

ing viruses have very long tumor latency compared to acutely-transforming virus, which already carries the viral-oncogene.

**[0054]** Some viruses are tumorigenic when they infect a cell and persist as circular episomes or plasmids, replicating separately from host cell DNA (Epstein-Barr virus and Kaposi's sarcoma-associated herpesvirus). Indirect viral oncogenicity also exists and involves chronic nonspecific inflammation occurring over decades of infection, as is the case for HCV-induced liver cancer. These two mechanisms differ in their biology and epidemiology: direct tumor viruses must have at least one virus copy in every tumor cell expressing at least one protein or RNA that is causing the cell to become cancerous. Foreign virus antigens are expressed in these tumors, consequently persons who are immunosuppressed such as AIDS or transplant patients are at higher risk for these types of cancers.

**[0055]** Gene or Cellular Therapy Based on the Use Viral Vectors

**[0056]** Gene therapy involves the insertion of genes into an individual's cells and tissues to treat diseases, such as hereditary diseases where deleterious mutant alleles of a gene are replaced with functional ones. Although the technology is still in its infancy, it has been used with some success. Viral vectors are a tool commonly used by scientific to deliver such genetic material into tissues (gene therapy) or cells (cellular therapy). These viral vectors are mainly derived from lentiviruses, retroviruses, adenoviruses or herpes viruses (for review, (Thomas, Ehrhardt et al. 2003)).

**[0057]** The genetic material in lentiviruses or retroviruses is in the form of RNA molecules, while the genetic material of their hosts is in the form of DNA. The genome of these classes of viruses can be modified by inserting a gene sequence of interest to be transferred in the infected tissue or cells. As the wild type virus, the recombinant viruses will introduce its RNA together with some enzymes, namely reverse transcriptase and integrase, into the cell. This RNA molecule from the recombinant virus must produce a DNA copy from its RNA molecule before it can be integrated into the genetic material of the host cell. After this DNA copy is produced and is free in the nucleus of the host cell, it must be incorporated into the genome of the host cell. That is, it must be inserted into the chromosomes in the cell by another enzyme carried in the virus called integrase. If this host cell divides later, its descendants will all contain the new genes.

**[0058]** One of the problems of gene therapy that can result from the use of lentiviruses or retroviruses is that the integrase enzyme can insert the genetic material of the virus into any arbitrary position in the genome of the host; it randomly shoves the genetic material into a chromosome. If genetic material happens to be inserted in the middle of one of the original genes of the host cell, this gene will be disrupted (insertional mutagenesis). If the gene happens to be one regulating cell division, uncontrolled cell division (i.e., cancer) can occur. This problem has recently begun to be addressed by utilizing single-chain homing endonucleases (Grizot, Smith et al. 2009) or by including certain sequences such as the locus control region to direct the site of integration to specific chromosomal sites (Zhou, Zhao et al. 2007).

**[0059]** Adenoviruses are viruses that carry their genetic material in the form of double-stranded DNA. They cause respiratory, intestinal, and eye infections in humans (especially the common cold). When these viruses infect a host cell they introduce their DNA molecule into the host. The genetic

material of the adenoviruses is not incorporated into the host cell's genetic material. The DNA molecule is left free in the nucleus of the host cell and the instructions in this extra DNA molecule are transcribed just like any other gene. The only difference is that these extra genes are not replicated when the cell is about to undergo cell division so the descendants of that cell will not have the extra gene. Adeno-associated viruses (AAV) from the parvovirus family are small viruses with a genome of single stranded DNA. The wild type AAV can insert genetic material at a specific site on chromosome 19 with near 100% certainty (Huser and Heilbronn 2003). But the recombinant AAV, which does not contain any viral genes and only the therapeutic gene, does not integrate into the genome. As a result, treatment with either adenoviral vector or AAV will require readministration in a growing cell population although the absence of integration into the host cell's genome should prevent the development of cancer (Douglas 2007).

**[0060]** Herpes viruses are currently used as gene transfer vectors due to their specific advantages over other viral vectors. Among the unique features of HSV derived vectors are the very high transgenic capacity of the virus particle allowing to carry long sequences of foreign DNA, the genetic complexity of the virus genome, allowing to generate many different types of attenuated vectors possessing oncolytic activity, and the ability of HSV vectors to invade and establish lifelong non-toxic latent infections in neurons from sensory ganglia from where transgenes can be strongly and long-term expressed. Three different classes of vectors can be derived from HSV: replication-competent attenuated vectors, replication-incompetent recombinant vectors and defective helper-dependent vectors known as amplicons. Replication-defective HSV vectors are made by the deletion of one or more immediate-early genes, e.g. ICP4, which is then provided in trans by a complementing cell line. Oncolytic HSV vectors are promising therapeutic agents for cancer. Such HSV based vectors have been tested in glioma, melanoma and ovarian cancer patients.

**[0061]** The inventors have surprisingly discovered that Molecular Combing and stretching techniques in combination with specially designed probes or sets of probes can be used to diagnose, detect or monitor subjects carrying infectious viral DNA and other pathogenic genes by means of Molecular Combing and/or DNA stretching techniques. These diagnostic applications are unknown in the prior art which generally performed Molecular Combing on DNA samples easily obtained from known cultured cell lines. While Molecular Combing has been previously applied for chromosomal analysis, it was not previously recognized that this methodology or adaptations of this methodology could be usefully applied to diagnosing viral infection or detecting other pathogenic polynucleotides in a sample. On the other hand, the inventors have surprisingly found that DNA obtained directly from biological samples from a virus-infected subject can be evaluated by Molecular Combing procedures and active, infectious viral DNA detected in these samples. These methods overcome many of the problems associated with existing diagnostic methods for virus infection and provide a convenient, rapid and accurate method for detection and monitoring of infectious viral polynucleotides in a virus-infected subject or patient.

#### BRIEF SUMMARY OF THE INVENTION

**[0062]** The invention provides a reliable, simple, fast, and inexpensive way to detect infectious or pathogenic poly-

nucleotide sequences, including viral genomic sequences, in a sample from a subject or patient using a Molecular Combing and/or DNA stretching techniques in combination with specific probes which are able to bind with at least 5%, 10%, 25%, 50%, 75%, especially sets of probes that bind to 80-100% of the infectious or pathogenic polynucleotide sequence.

**[0063]** The inventors have discovered that Molecular Combing is a powerful technique (i) to detect and quantify viral DNA or viral origin-DNA in infected mammalian cells, tissues or biological fluids, such as infectious herpes simplex virus DNA in infected cornea, (ii) follow viral replication and viral genomic rearrangements that occur in the infected mammalian cells tissues or biological fluids, and (iii) identify infectious virus polynucleotides or infectious virus polynucleotides present in infected cells, tissues or biological fluids.

**[0064]** Unlike prior art techniques for detecting virus in infected cells or tissues, the Molecular Combing techniques of the invention can be performed simultaneously and do not require multiple time consuming and expensive procedures. The methods of the invention are easily applicable as diagnostic tools for detecting viral contamination or infection in cells or tissues, useful for determining the efficacy of antiviral treatments by quantifying or otherwise evaluating the effects or efficacies of such treatments on the quantity of infectious virus or infectious viral polynucleotides or viral replication in infected mammalian cells or tissues. The inventors have found that Molecular Combing overcomes many of the problems associated with past methods of detecting or diagnosing a viral infection, including PCR-based methods, and offers an easy and rapid way to do so. Specific embodiments of the invention include the following.

**[0065]** A method for detecting a viral genome or infectious viral polynucleotide in a biological sample comprising isolating, extracting or otherwise obtaining a polynucleotide from said sample; Molecular Combing said polynucleotide to form a stretched polynucleotide; contacting the stretched polynucleotide with one or more probes that recognize the infectious viral or genomic viral polynucleotide sequence; detecting hybridization of the probes to the combed sample. The polynucleotide may be infectious genomic viral DNA or infectious non-genomic viral DNA, an oncogene or a protooncogene and may be integrated into a chromosome of a subject or may be episomal or transgenic DNA. Detection of an entire viral genome or entire infectious viral DNA is importantly correlated with infection. For example, regarding the HSV causative agents of herpetic keratitis, it is essential to be able to show the presence of complete infectious genomes in the cells which can then be partially or totally generated in the infected cells.

**[0066]** The biological sample is generally obtained directly from a subject or patient and may constitute tissue, a cell sample, or blood, CSF, or synovial fluid sample or other fluid biological sample. A DNA molecule or fiber for stretching or for Molecular Combing may be extracted from the biological sample. A sample may be obtained from a living or non-living subject. Subjects include animals susceptible to viral diseases including humans, non-human mammals, such as cattle, bovines, sheep, goats, horses, pigs, dogs, cats and non-human primates; birds, such as chicken, turkey, duck, goose, ostrich, emu, or other birds; reptiles, amphibians and other animals.

**[0067]** A stretched or combed DNA molecular is contacted with one or more probes suitable for identifying the infectious

or biologically active genomic, infectious viral, oncogene or protooncogene DNA. For detection of a DNA virus one or more probes, or preferably a set of probes covering 80-100% or any intermediate subrange or value of the viral genome are employed. The DNA virus may be one that is single or double stranded. Representative DNA viruses include herpes virus, such as Herpes Simplex Virus (HSV) types 1 and 2, papillomaviruses, and hepatitis viruses, such as hepatitis B virus. Probes that bind to retroviruses, such as HIV, preferably a retrovirus integrated as a provirus into a host chromosome are also selected in a similar manner and preferably a set of probes covering 80-100% of the proviral genome are employed.

**[0068]** The probes may correspond to at least 5% and preferably 80-100% of the genome of a virus strain already known to be infectious or may be designed to encompass the combination of genes known to be essential for replication, reproduction, or pathogenicity of a particular kind of virus.

**[0069]** Said probes can correspond to HSV-1 (NC\_001806.1)(SEQ ID NO: 15), HSV-2 (NC\_001798.1)(SEQ ID NO: 16), HIV1 (NC\_001802.1)(SEQ ID NO: 17), HIV2 (NC\_001722.1)(SEQ ID NO: 18), HBV (NC\_003977.1)(SEQ ID NO: 19), HPV16 (NC\_001526.2)(SEQ ID NO: 20), HPV18 (X05015.1)(SEQ ID NO: 21), HPV31 (J04353.1)(SEQ ID NO: 22), HPV33 (M12732.1)(SEQ ID NO: 23) and HPV45 (X74479.1)(SEQ ID NO: 24) which are also incorporated by reference to their accession numbers.

**[0070]** The one or more probes or probe set may comprise two, three or more different subsets of probes tagged with different labels. Preferably, such probes are used to determine the configurations of a genomic or infectious viral DNA or to identify different portions or segments of a viral genome or infectious viral DNA molecule or another pathogenic DNA molecule such as an oncogene or protooncogene.

**[0071]** In another embodiment the invention is directed to a method for detecting or identifying an infectious or genomic viral polynucleotide sequence in a mammalian cell, tissue or biological fluid comprising using Molecular Combing to detect the presence or the quantity of infectious viral polynucleotide or genomic viral polynucleotide in a cell, tissue or biological fluid. Still another embodiment constitutes a method for quantifying an infectious or genomic viral polynucleotide sequence comprising using Molecular Combing to detect quantity of infectious viral polynucleotide or genomic viral polynucleotide in a cell, tissue or biological fluid compared to an uninfected control sample or an otherwise similar sample obtained at a different point in time.

**[0072]** Methods for detecting the presence of a virus or detecting replication of a virus in a biological sample are also contemplated, as well as methods for longitudinally following virus replication or the rearrangement of a viral genome or infectious viral DNA especially in mammalian cells, tissue or biological fluid. Such methods may be performed by Molecular Combing DNA obtained from appropriate samples or samples taken over a specific time period for the presence of latent or replicating viral DNA or rearranged viral DNA in a cell or tissue.

**[0073]** The invention also pertains to method for evaluating the efficacy of anti-viral treatment using Molecular Combing to detect the presence, arrangement or quantity of infectious or genomic viral DNA in a sample obtained from a subject, treating the subject with an anti-viral agent or performing other anti-viral therapy, and monitoring or re-evaluating the

presence, arrangement or quantity of infectious or genomic viral DNA after treatment using Molecular Combing techniques of described herein.

**[0074]** A kit for performing molecule combing containing a Molecular Combing apparatus and/or reagents, one or more probes that bind to an infectious polynucleotide, and optionally one or more cell, tissue or biological fluid sample(s) as well as other conventional ingredients for performing Molecular Combing, DNA stretching or probe hybridization. In some embodiments the kits will contain a set of probes for detecting or identifying genomic viral DNA in a combed DNA molecule covering at least 5% of a viral genome, infectious viral nucleic acids, oncogene, or proto-oncogene, especially 80, 85, 90, 95, 99 or 100%. The minimal length of the probe or set probe is 1 kb. In some embodiments, the kit will further contain software for detecting or classifying the infectious sequences and in other embodiments the kit may additionally comprise instructions for detecting a viral genome or infectious viral nucleic acids according to the methods of the invention.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0075]** FIG. 1. Example of HSV-1 specific-probes that can be used according to the invention. The dot plot represents the comparison between the HSV-1 genome sequence (horizontal axe) and the HSV-1 specific probes (vertical axe) to map the different probes to each other. Some of the HSV-1 fragments (HSV-S54, -B52 and -S58) hybridize with the inverted repeated sequences and consequently are present in two copies. The HSV-1 fragments listed in Table A can be associated and revealed in different colours to detect the HSV-1 genome. In this example, the association of 41 probes (on the 63 probes available, Table A) as follow allows covering 99% of the HSV-1 genome. The HSV-S54, -P4, -P5, -S4 and -B4 fragments correspond to the apparent H1 21 kb probe. The apparent H2 56 kb probe consists of 19 overlapping probes (HSV-B7, -B8, -S14, -B10, -S16, -B13, -S18, -B15, -S21, -B19, -S23, -B21, -B22, -S30, -S31, -S32, -B26, -P8, and -B28 fragments). The 45 kb H3 probe is composed of the HSV-B30, -S44, -S45, -B38, -B39, -B40, -S49, -B44, -B45, -B46, -B48 and -S54 fragments. Due to the presence of the inverted repeated sequences, both H4 (13 kb) and H6 (6.5 kb) probes are composed of HSV-B52 and -S58 fragments; the HSV-S59 fragment is also present in the H4 probe. Finally, 3 overlapping fragments (HSV-S60, -B58 and -B60) formed the H5 7.5 kb probe. The localization probes which form a signature specific for a specific region of the HSV-1 genome is in red (digoxigenin labelling, black boxes) for H1, H3 and H5 probes and in green (biotin labelling, grey boxes) for the H2, H4 and H6 probes. The sizes (rounded to the nearest kb) of the localization probes are indicated respectively below each boxes.

**[0076]** FIG. 2. Comparison of two methods of HSV-1 DNA extraction from viral particles. A) Example of combed HSV-1 DNA on silanized surfaces. After combing of the HSV-1 DNA solution extracted with either the standard phenol chloroform (left picture) or the modified extraction protocol from agarose plug-embedded HSV-1 (right picture), the fibers are revealed with the intercalating agent YOYO-1 and observed on a epifluorescence microscope equipped with a camera. The figure shows two pictures that are characteristic of each extraction method. Scale is indicated as a bar. B) Histogram showing the cumulative frequency of the number DNA fiber sizes within a given length interval which were contained on

genomic DNA isolated from HSV-1 particles. Interval width is 10  $\mu\text{m}$ . Thus, for example, the third point represents the number of measures in the [30-40  $\mu\text{m}$ ] interval. Left panel shows the length of every HSV-1 DNA fibers that was recorded (654 measures) in the DNA solution extracted with the standard phenol:chloroform extraction. Median size of the fibers is 18  $\mu\text{m}$  that is, 36 kb. The right panel show the distribution of the length of DNA fibers (2322 measures) obtained with the alternative method. In that case, the median size of the distribution is 42  $\mu\text{m}$  that is, 84 kb.

**[0077]** FIG. 3. Detection of isomers of the HSV-1 genome in DNA solution extracted from viral particles and infected cells. A) Examples of FISH on combed HSV-1 DNA. Schematic representation of the different possible organization of the hybridization patterns corresponding to the different isomers of the HSV-1 genome are indicated (digoxigenin labelled-H1, H3 and H5 probes are represented in black boxes; biotin-labelled-H2, H4 and H6 probes are depicted in grey boxes). The minimal requirement hybridization patterns as defined in the "Analysis of HSV-1 detected signals" section are also indicated just above the complete signal. Four representative linear hybridization chains showing each complete isomer of HSV-1 genome (White: Texas Red/Alexa 594-fluorescence: H1, H3 and H5; Black: green Alexa 488-fluorescence: H2, H4 and H6). Schematic representations of each HSV-1 genome isomer are shown above the corresponding pictures. (B) Histogram of the distribution genome isomers of the HSV-1 KOS strain in viral particles produced in Vero, COS7 and Neuro2A cell lines. Hybridization signals were selected and analyzed as described in the "Examples" section. In this example, a total of 405 hybridization signals for each experiment were identified and classified. Each bar represents the number of each isomer of the HSV-1 genome. In this example, the distribution of the HSV-1 KOS strain isomers are equivalently distributed in viral particles from COS7 cells whereas the P and IS isomers are the more frequent isomer in the viral particles produced from the Neuro2A and Vero cell lines. This latter distribution is statistically different from an equimolar distribution (Chi-2 test). (C) Histograms of the distribution of the genome isomers of HSV-1 strains Sc16 and KOS in different infected cells (BSR, COS-7, Neuro 2A and Vero cells). The hybridization signals were selected and analyzed as described in the "Examples" section. In these examples, 405 signals from each production were selected and classified. The distribution of the HSV-1 strain Sc16 produced in BSR, COS-7, Neuro 2A and Vero cells and of the HSV-1 strain KOS produced in COS-7 is statistically equivalent to an equimolar distribution (Chi-2 test). In Neuro 2A and Vero infected cells, the P and IS isomers are more frequent than the IL and ILS isomers.

**[0078]** FIG. 4. Method of extraction of genomic DNA from mouse and rabbit cornea. A) Example of combed genomic DNA on silanized surfaces. After extraction from cornea with the protocol described in Examples section, the genomic DNA solution is combed and revealed with the intercalating agent YOYO-1 before observation on an epifluorescence microscope equipped with a camera. Left pictures show a representative picture of combed DNA extracted from a healthy mouse cornea and the right picture from a healthy rabbit cornea. Molecular Combing is performed at low density to allow measurement of the length of the genomic DNA fiber. Scale is indicated as bar. B) Histogram showing the cumulative frequency of the number DNA fiber sizes within a given length interval which were contained on genomic DNA

isolated from HSV particles. Interval width is 25 kb. Thus, for example, the fifth bar represents the number of measures in the [225-250 kb] interval. Left panel shows the length of every DNA fibers that was recorded (10069 measures) in the DNA solution extracted from the mouse cornea. Median size of the fibers is 204 kb. In this sample, around 31% of the genomic DNA fiber exhibited a size above 200 kb. The right panel show an example of the distribution of the length of DNA fibers (8336 measures) obtained from the rabbit cornea. In that case, the median size of the distribution is 196 kb with a proportion of fibers above 200 kb in length of 27%.

**[0079]** FIG. 5. Histogram of the distribution of isomers of HSV-1 genome in infected mouse cornea. Total DNA from HSV-1 infected mouse cornea was extracted, combed and hybridized with the HSV-1-specific probes. In this example, a total of 18 hybridization signals were identified and classified. Each bar represents the number of each isomer of the HSV-1 genome.

**[0080]** FIG. 6. Detection of replication intermediate of the HSV-1 in HSV-1 infected mouse cornea. The figure shows two examples of complex HSV-1 genome DNA corresponding to replication concatemers (White signal: Texas Red/Alexa 594-fluorescence: H1, H3 and H5; Black signal: green Alexa 488-fluorescence: H2, H4 and H6). The upper picture depicted a signal compose of at least two hypothetical overlapping HSV-1 genomes composed of an IS and a P isomer while the lower picture show a HSV-1 concatemer consisting of at least hypothetical ILS and IL signals.

**[0081]** FIG. 7. Detection of non canonical HSV-1 genomes in HSV-1 infected mouse cornea. A) Examples of non canonical HSV-1 from infected mouse cornea. The figure shows representative examples of hybridization patterns that do not correspond to one of the canonical isomers of the HSV-1 genome (White signal: Texas Red/Alexa 594-fluorescence: H1, H3 and H5; Black signal: green Alexa 488-fluorescence: H2, H4 and H6). Scale bar is indicated. B) Canonical hybridization pattern corresponding to the ILS isomer (biotin-labelled H1, H2B and H3 probes are represented in grey boxes and signal; digoxigenin-labelled H2A and H5 are depicted in white boxes and signals, and Alexa488-labelled H4 and H6 probes are depicted in black boxes and signals) obtained on combed DNA extracts of HSV-1 strain Sc16-infected Vero cells. This example shows HSV-1 specific probes H1 to H6 labelled to evaluate the proportion of non-canonical structures in the H4/H6 region. C) Non canonical H4/H6 on HSV-1 strain Sc16 in Vero cell extracts. The first hybridization pattern shows alternation of Alexa 488 fluorescence signal (black signals) of various sizes corresponding to the H4/H6 probes and Alexa 594 fluorescence (grey signal) corresponding to fragments H1, H2B or H3 probes or AMCA/Alexa 350 fluorescence signal (white signal) corresponding to part of H2A or H5 probes of a maximum of 10kb. The second example exhibit an alternation between Alexa 488 fluorescence signal (black signals) of various sizes and AMCA/Alexa 350 fluorescence signal (white signal) that is surrounded by Alexa 594 fluorescence signal (grey signal). The third example shows a unique repetition of an Alexa 488 fluorescence signal (black signals) surrounded by Alexa 594 fluorescence signal (grey signal). Scale bar is indicated. D) Histograms of the distribution between canonical and non canonical structure in the H4/H6 regions of HSV-1. 367 hybridization signals were selected and analysed as described in the "Examples" section. In this example, 80% of the H4/H6 probes correspond to the theoretical structure.

**[0082]** FIG. 8. Detection of proviral HIV-1 DNA in ACH-2 cells culture. All histograms are showing number of signals within a given length interval (0.644 kb/class) in function of one FISH signal length or gap size between two FISH signals, in kilobases. Thus, for example, the second bar of the first histogram represents the number of measures in the [7.86-8.50 kb] interval. A) Examples of proviral HIV-1 DNA located at chromosome 7p15 using the G248P87988G9 and G248P86255A8 fosmid. Schematic representation of organization of the hybridization pattern corresponding to the integrated proviral HIV-1 DNA is indicated (digoxigenin labelled-HIV-1 probes are represented in a black box and signal; biotin-labelled-fosmids are depicted in white boxes and signals). Histograms on the left and right showing the distribution of the gap size between fosmid G248P87988G9 green Alexa 488-fluorescence signal and HIV-1 Texas Red/Alexa 594-fluorescence signal, and between HIV-1 Texas Red/Alexa 594-fluorescence signal and fosmid G248P86255A8 green Alexa 488-fluorescence signal, respectively. Middle histogram shows the distribution of the HIV-1 Texas Red/Alexa 594-fluorescence signal size. B) Examples of normal allele of the 7p15 locus. Schematic representation of organization of the hybridization pattern corresponding to the normal allele is indicated (biotin-labelled-fosmids are depicted in grey boxes and signals). Histogram showing the distribution of the gap size between fosmid G248P87988G9 green Alexa 488-fluorescence signal and G248P86255A8 green Alexa 488-fluorescence signal. C) Isolated proviral HIV-1 DNA. Schematic representation of organization of the hybridization pattern corresponding to the isolated form of HIV-1 is indicated (digoxigenin labelled-HIV-1 probes are represented in a black box and signal). Histogram shows the distribution of HIV-1 Texas Red/Alexa 594-fluorescence signal size. D) Examples of proviral HIV-1 DNA located at chromosome 7p15 using the G248P84833H9 fosmids. Schematic representations of organization of the hybridization patterns corresponding to the integrated proviral HIV-1 DNA (digoxigenin labelled-HIV-1 probes are represented in a black box and signal; and the wild type locus (biotin-labelled-G248P84833H9 fosmids are depicted in white boxes and signals) are indicated. The histogram shows the distribution of the HIV-1 Texas Red/Alexa 594-fluorescence signal size.

#### DETAILED DESCRIPTION OF THE INVENTION

**[0083]** The invention enables a rapid, specific and sensitive detection of infectious viral polynucleotides or infectious viral origin DNA in a sample that avoids the significant constraints imposed by amplification methods like PCR or serological tests such as ELISA.

**[0084]** In contrast to prior art methods, the Molecular Combing techniques of the invention permit the successful detection of complete viral genomes or infectious or virulent portions of viral polynucleotide sequences leading to improved diagnosis of acute or latent viral infections.

**[0085]** The method of the invention enables to detect the type of HSV and the structure of its genome reliably, in a time- and cost-effective fashion, and with none of the constraints of manipulating radioactivity. Moreover, the method of the invention enables to follow the presence of the viral infection after antiviral treatment, whatever the type of treatment considered.

**[0086]** The present invention relates to a method for detecting in vitro the presence of a genome of DNA viruses in

infected cells of eukaryotic cells, in particular the detection of the HSV genome. Said method comprises a hybridization step of nucleic acid representative of given virus with at least a probe or a set of probes which is (are) specific for HSV DNA.

**[0087]** Molecular Combing techniques which may be used in accordance with the invention are disclosed by Bensimon, et al., U.S. Pat. No. 6,303,296 and by Lebofsky, et al., WO 2008/028931 the disclosure of which are hereby incorporated by reference. The inventors recognized that no technique can detect the presence of a complete or non-rearranged genome of infectious form of a virus except by traditional culture techniques and thus sought to apply and adapt these techniques to detection of infectious viral DNA. The combing method was never tested as an efficient tool to detect such viral DNA forms but only for the location of long sequences of DNA inserted in a cellular genome in view of understanding the genetic influence of such sequence in a specific genetic environment. Non-limiting examples of detection of infectious viral (HSV) DNA by the modified methods developed by the inventors are shown below.

**[0088]** The inventors modified the standard extraction protocol to isolate viral genomic DNA from viral particles. Generally, 0.1% SDS is used for the lysis of viral particles in the low-melting agarose plugs instead of 0.1% sarkosyl in the standard protocol. A methodology was also developed to extract genomic DNA from cornea to investigate the presence of HSV DNA in infected cornea for purposes of diagnosis.

**[0089]** The present invention relates to a method for detecting viral DNA, in particular HSV DNA and HIV DNA, contained on nucleic acid of infected cells, tissue or biological fluid. Said method comprises a hybridization step of nucleic acid representative of said viral DNA with specific probes or set of probes that cover the entire said viral DNA and that permit to identify rearrangements within the nucleic acid representative of said viral DNA.

**[0090]** The term “nucleic acid” and in particular “nucleic acid representative of viral DNA” as used herein designates one or several molecules of any type of nucleic acid capable of being attached to and stretched on a support as defined herein, and more particularly stretched by using Molecular Combing technology; nucleic acid molecules include DNA (in particular genomic DNA, especially viral DNA, or cDNA) and RNA (in particular mRNA). A nucleic acid molecule can be single-stranded or double-stranded.

**[0091]** “Nucleic acid representative of said viral DNA” means that said nucleic acid contains the totality of the genetic information or the essential information with respect to the purpose of the invention, which is present on said viral DNA. This term includes genomic viral DNA, such as that integrated into a host chromosome and which can produce infectious virus, infectious viral DNA which may lack certain genomic elements but can be infectious when expressed in particular host cells, and viral genes which exert a pathogenic effect on host cells, such as viral oncogenes.

**[0092]** A proto-oncogene includes those normal genes, which when altered by mutation can convert into oncogenes that causes a cell to grow or divide in an unregulated manner. Proto-oncogenes have diverse cellular functions, some provide signals for cell division, and others may play roles in apoptosis. Functional and structural characteristics of oncogenes, including their nucleic acid sequences, are well-known in the art and are incorporated by reference to *Human Cancer Viruses: Principles of Transformation and Pathogen-*

*esis* by J. Nicolas, et al., Karger Publishers (2010), ISBN3805585764, 9783805585767 (244 pages) which is incorporated by reference.

**[0093]** An oncogene encompasses a defective version of a proto-oncogene. A single copy of an oncogene can cause uncontrolled cell growth. Representative oncogenes include ras, myc, src, Her-2/neu, hTERT, and Bcl-2. Functional and structural characteristics of oncogenes, including their nucleic acid sequences, are well-known in the art and are incorporated by reference to *Oncogene: Gene, Mutation, Tumor, Apoptosis, Gene Expression, Protein, Cell Growth, Cellular Differentiation*, by L. M. Surhone, et al., Betascript Publishers (2010), ISBN6130361599, 9786130361594 (172 pages). Further description and identification of oncogenes is incorporated by reference to [http://\\_www.cancerquest.org/index.cfm?page=780](http://_www.cancerquest.org/index.cfm?page=780) (last accessed Apr. 11, 2011) and to Cooper G. *Oncogenes*. Jones and Bartlett Publishers, 1995 and Vogelstein B, Kinzler K W; *The Genetic Basis of Human Cancer*. McGraw-Hill: 1998, both of which are incorporated by reference. The process of activation of proto-oncogenes to oncogenes can include viral transduction or viral integration, point mutations, insertion mutations, gene amplification, chromosomal translocation and/or protein-protein interactions. Viruses that can induce activation of proto-oncogenes include HBV and HCV (hepatocellular carcinoma), HTLV (leukaemia), HPV (cervical, anal and penile cancer), HSV-8 (Kaposi's sarcoma), Merkel cell polyomavirus (Merkel cell carcinoma) and, EBV (Burkitt's lymphoma, Hodgkin's lymphoma, post-transplantation lymphoproliferative disease and Nasopharyngeal carcinoma)

**[0094]** In a particular embodiment, the nucleic acid sample used for stretching is genomic DNA, in particular total genomic DNA or more preferably chromosomal genomic DNA (nuclear genomic DNA) of infected cells or tissues, and/or fragments thereof. The term “nucleic acid” is in particular used herein to designate a nucleic acid representative of one or several chromosome(s) and/or of one or several fragment(s) of chromosomes. Said fragments can be of any size, the longest molecules reaching several megabases. Said fragment are generally comprised between 10 and 2000 kb, more preferably between 20 and 500 kb and are in average of about 300 kb.

**[0095]** The nucleic acid sample used in the method of the invention can be obtained from a biological fluid or from a tissue of biological origin, said sample or tissue being isolated for example from a human, a non human mammal or a bird.

**[0096]** As defined herein, a probe is a polynucleotide, a nucleic acid/polypeptide hybrid or a polypeptide, which has the capacity to hybridize to nucleic acid representative of virus DNA as defined herein, in particular to RNA and DNA. This term encompasses RNA (in particular mRNA) and DNA (in particular viral cDNA or viral genomic DNA) molecules, peptide nuclear acid (PNA), and protein domains. Said polynucleotide or nucleic acid hybrid generally comprises or consists of at least 100, 300, 500 nucleotides, preferably at least 700, 800 or 900 nucleotides, and more preferably at least 1, 2, 3, 4 or 5 kb. For example, probes of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 kb or more than 15 kb, in particular 30, 50, 100 or 150 kb can be used. Such a probe or a set of probes may correspond or cover 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 or 100% of a viral genome, especially that of an infectious virus. A nucleic acid probe may be a single or double-stranded polynucleotide or a modified polynucleotide. A set of continuous set of probes

will cover a particular section of the genome or the entire genome or overlap each other with respect to this section. A suitable probe may be based on a viral polynucleotide sequence, such as those disclosed herein, and may be identified or synthesized by any appropriate method. Probes may be labelled or tagged radioactively, fluorescently or by other means known in the art.

**[0097]** A polypeptide probe generally specifically binds to a sequence of at least 6 nucleotides, and more preferably at least 10, 15, or 20 nucleotides. As used herein, the sequence of a probe, when the probe is a polypeptide, should be understood as the sequence to which said polypeptide specifically binds.

**[0098]** By "a portion of" a particular region, it is meant herein consecutive nucleotides of the sequence of said particular region. A portion according to the invention can comprise or consist of at least 15 or 20 consecutive nucleotides, preferably at least 100, 200, 300, 500 or 700 consecutive nucleotides, and more preferably at least 1, 2, 3, 4 or 5 consecutive kb of said particular region. For example, a portion can comprise or consist of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 consecutive kb of said particular region.

**[0099]** In a particular embodiment, the probe used or at least one of the probes used is a nucleotide variant of the probe showing a complementary sequence of 100% to a portion of one strand of the target nucleic acid. The sequence of said variant can have at least 70, 80, 85, 90 or 95% complementarity to the sequence of a portion of one strand of the target nucleic acid. Said variant can in particular differ from the probe which is 100% identical or complementary by 1 to 20, preferably by 1 to 10, nucleotide deletion(s), insertion(s) and/or more preferably substitution(s), in particular by 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 nucleotide deletion(s), insertion(s) and/or more preferably substitution(s) in the original nucleotide sequence. In a particular embodiment, the variant keeps the capacity to hybridize, in particular to specifically hybridize, to the sequence of the nucleic acid target, similarly to the probe that is 100% identical or 100% complementary to a sequence of the nucleic acid target (in particular in the hybridization conditions defined herein).

**[0100]** The term "complementary sequences" in the context of the invention means "complementary" and "reverse" or "inverse" sequences, i.e. the sequence of a DNA strand that would bind by Watson-Crick interaction to a DNA strand with the said sequence.

**[0101]** In a particular embodiment of the invention, the probes or one or several probes used to carry out the invention are labelled with one or several hapten(s) (for example biotin and digoxigenin) and revealed with specific antibodies directed against these haptens. Use of different haptens for a given probe or set of probe will allow to detect rearrangements within a given viral DNA. Said probes can be labelled as defined herein and as described in patent application WO 2008/028931, which is incorporated by reference.

**[0102]** A set of probes as used herein comprises of at least two probes. For example, said set of probes can consist of 2 to 20 probes (2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 probes). The number of probes in a set does usually not exceed 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 probes depending on the sensitivity that is required and the length of the probe; a set of probes preferably consists of 2, 3, 4, 5, 6, 7, 8, 9 or 10 probes at the most.

**[0103]** The probes or the set of probes of the invention do not only allow the detection of the entire whole genomic viral

DNA but also to identify rearrangements that can occur within the genomic viral DNA. For example, the probe sets used in the Example 1 for HSV detection allow the detection of the different HSV genome isomers that are generated by homologous recombination between the inverted repeated sequences that surround the unique regions  $U_L$  and  $U_S$ . Moreover, each probes set (H1 to H6) for the detection of HSV are composed of several probes (3 to 19 different fragments) that are labeled with different two different haptens (digoxigenin for the fragments consisting of H1, H3 and H5 probes; biotin for the fragments consisting of H2, H4 and H6 probes). Changing the haptens of one or of several fragments by other haptens (whatever the nature of this hapten) within a given probe set permits the generation of a different color-fluorescence array that allows the identification of rearrangements within this specific region of the viral DNA.

**[0104]** Molecular Combing can also be a useful tool for an early diagnosis of patients susceptible of developing a cancer caused by a viral infection, for example, those at risk of conversion of a proto-oncogene into an oncogene. Indeed, Molecular Combing can discriminate between integrated and episomal viral genome. In the case where the viral DNA is integrated, by using a specific set of probes Molecular Combing will allow to determine whether the integrated viral DNA contains a complete viral oncogene or whether it is integrated in a close proximity of a cellular proto-oncogene.

**[0105]** The invention may also be used in conjunction with gene therapy. By using specific sets of probes for each transgene and/or viral vector, Molecular Combing provides a powerful tool for an evaluation of efficacy and safety of the viral vector-based gene therapy. Indeed, Molecular Combing can discriminate between integrated and episomal transgene. In the case where the transgene is integrated (lentiviruses and retrovirus based vectors), by using a specific set of probes Molecular Combing will allow to determine whether the integrated viral DNA is integrated in a gene (insertional mutagenesis) or in a close proximity of a cellular proto-oncogene. In the case the transgene is episomal (Adenoviral vector, AAV or HSV vector); Molecular Combing will be useful for the quantification of the transgene in the growing cell population and may help to define the right time for readministration of the transgene.

**[0106]** Specific embodiments of the invention include the following methods and products. A method for detecting an infectious viral polynucleotide in a biological sample comprising: separating, extracting or otherwise obtaining a polynucleotide from said sample, Molecular Combing said polynucleotide to form a stretched polynucleotide, contacting said stretched polynucleotide with one or more probes that recognize the infectious polynucleotide sequence, detecting hybridization of the probes to the combed sample. This method may be performed using a biological sample that is a tissue or cell sample obtained from a subject, for example, a blood, plasma, serum, CSF, synovial fluid sample or some other kind of biological fluid from the subject. A polynucleotide used in this method may be extracted from the tissue or cell sample or from components of a biological fluid obtained from the subject. Generally, the sample will be obtained from a living subject, but samples may also be obtained from deceased subjects. Samples can be obtained from humans or other mammals such as cattle, bovines, sheep, goats, horses, pigs, dogs, cats and non-human primates, or from other animals such as avian species such as a chicken, turkey, duck, goose, ostrich, emu, or other bird. The method can be prac-

ticed with polynucleotides which are DNAs which may also contain or comprise infectious or non-infectious genomic viral DNA or non-infectious non-genomic viral polynucleotides such as DNA or RNA. The polynucleotide detected or analyzed by this method may be integrated into the DNA of the subject or can be in episomal form. The polynucleotide to be detected can be contacted with one or more probes that bind to a DNA virus, including both single-stranded and double-stranded DNA viruses. For example, the polynucleotide may be contacted with one or more probes that bind to a herpes virus, such as herpes simplex virus (HSV). Other probes may be used which bind to other viruses like papilloma virus, hepatitis B virus, or retroviruses like HIV. In some embodiments the method will use a set of probes that bind to at least 80%, 90%, 95%, or 99% of the genomic or infectious DNA of a virus. In other embodiments the one or more probes used in the method will comprise at least two sets of probes that are tagged with different labels, for example, to permit the identification of different portions or segments of a viral genome to which the different probes bind or to permit the identification of rearrangements in a viral genome.

**[0107]** Other specific embodiments of the invention include:

**[0108]** A method for detecting, identifying or visualizing an infectious or genomic viral polynucleotide sequence in a mammalian cell, tissue or biological fluid comprising using Molecular Combing to detect the presence or the quantity of infectious viral polynucleotide or genomic viral polynucleotide in a cell, tissue or biological fluid.

**[0109]** A method for quantifying an infectious or genomic viral polynucleotide sequence comprising using Molecular Combing to detect quantity of infectious viral polynucleotide or genomic viral polynucleotide in a cell, tissue or biological fluid compared to an uninfected control sample or an otherwise similar sample obtained at a different point in time or from a different source or clinical sample.

**[0110]** A method for detecting or following viral presence or replication or viral genomic rearrangements in a mammalian cell comprising Molecular Combing for the presence of latent or replicating viral DNA or rearranged viral DNA in a cell, tissue or biological fluid.

**[0111]** A method for evaluating the efficacy of anti-viral treatment comprising detecting using Molecular Combing the presence, arrangement or quantity of infectious or genomic viral DNA in a sample obtained from a subject, treating said subject with an anti-viral agent, and re-evaluating using Molecular Combing the presence, arrangement or quantity of infectious viral or genomic viral DNA in said subject.

**[0112]** A set of probes covering 80-100% of the HSV, HIV, HBV or HPV genome.

**[0113]** A kit for performing molecule combing comprising a Molecular Combing apparatus and/or reagents, one or more probes that bind to an infectious polynucleotide, and optionally one or more cell, tissue or biological fluid sample(s). The kit which comprises a set of probes for detecting or identifying genomic viral DNA in a combed DNA molecule. The kit further comprising software for detecting or classifying the infectious sequences.

**[0114]** The following biological materials which have been deposited under the terms of the Budapest Treaty at the CNCM, Institut Pasteur 25 Rue du Docteur Roux, F-75724

Paris Cedex 15: HSV-B4 (CNCMI-4298), HSV-B19 (CNCMI-4299), HSV-Sc54 (CNCMI-4300), and HSV-P4 (CNCMI-4301).

## EXAMPLES

### Example 1

#### Herpes Simplex Virus Detection

**[0115]** Preparation of Embedded DNA Plugs from Viral Particles

**[0116]** HSV-1 DNA was extracted from viral particles by standard phenol:chloroform extraction (Ben-Zeev, Weinberg et al. 1974) or by a modified procedure described in Lebofsky et al. (Lebofsky, Heilig et al. 2006) which are both incorporated by reference. Briefly, HSV-1 particles were resuspended in 1× PBS at a concentration of  $5 \cdot 10^6$  viral particles/mL, and mixed thoroughly at a 1:1 ratio with a 1.2% w/v solution of low-melting point agarose (Nusieve GTG, ref. 50081, Cambrex) prepared in PBS, at 50° C. 90  $\mu$ L of the viral particles/agarose mix was poured in a plug-forming well (BioRad, ref. 170-3713) and left to cool at least 30 min at 4° C. Embedded viral particles were lysed in 0.1% SDS-0.5M EDTA (pH 8.0) solution at 50° C. for 30 minutes. After three washing steps in 0.5M EDTA (pH 8.0) buffer of 10 minutes at room temperature, plugs were digested by overnight incubation at 50° C. with 2 mg/mL Proteinase K (Eurobio code GEXPRK01, France) in 250  $\mu$ L digestion buffer (0.5M EDTA, pH 8.0). The use of 0.1% SDS instead of Sarkosyl was very productive and allows a very high quality of extracted viral DNA to be collected.

**[0117]** Preparation of Embedded DNA Plugs from Infected Cells

**[0118]** The extraction of HSV-1 DNA from infected cells culture (BSR, COS-7, Neuro 2A and Vero) was performed as previously described (Schurra and Bensimon 2009). Briefly, infected cells were pelleted by centrifugation at 5000 g for 5 minutes, resuspended at a concentration of  $2 \cdot 10^6$  cells/mL in 1× PBS buffer and mixed thoroughly at a 1:1 ratio with a 1.2% w/v solution of low-melting point agarose (Nusieve GTG, ref. 50081, Cambrex) prepared in 1× PBS at 50° C. 90  $\mu$ L of the cell/agarose mix was poured in a plug-forming well (BioRad, ref. 170-3713) and left to cool down at least 30 min at 4° C.

**[0119]** Lysis of cells in the blocks was performed as previously described (Schurra and Bensimon 2009). Briefly, Agarose plugs were incubated overnight at 50° C. in 250  $\mu$ L of a 0.5M EDTA (pH 8), 1% Sarkosyl, 250  $\mu$ g/mL proteinase K (Eurobio, code : GEXPRK01, France) solution, then washed twice in a Tris 10 mM, EDTA 1 mM solution for 30 in at room temperature.

**[0120]** Preparation of Embedded DNA Plugs from Infected Cornea

**[0121]** HSV-1 strain Sc16 infected mouse cornea was collected at a final stage of infection, and kept in Corneamax® (Eurobio code EYEMAX00, France) medium. After rinsing three times during 15 minutes at room temperature with 1× PBS solution, the entire cornea was cut into small pieces. Tissue lysis was carried out for up to 16 h at 37° C. in 0.3 mg/mL Collagenase type A (Roche, code 10 103 578 001), and 0.8 mg/mL GIBCO™ Dispase (Invitrogen, France, code 17105-041), both prepared in GIBCO™ 1× Hanks' Balanced Salt Solution HBSS buffer (Invitrogen, France, code 14060040). Lysates were pelleted by centrifugation at 5000g for 10 minutes, resuspended at a concentration of  $1 \cdot 10^6$  to  $2 \cdot 10^6$  cells/mL in 1× PBS buffer and mixed thoroughly at a 1:1

ratio with a 1.2% w/v solution of low-melting point agarose (Nusieve GTG, ref 50081, Cambrex) prepared in 1× PBS at 50° C. 90 µL of the cell/agarose mix was poured in a plug-forming well (BioRad, ref 170-3713) and left to cool down at least 30 min at 4° C.

**[0122]** Lysis of cells in the blocks was performed as previously described (Schurra and Bensimon 2009). Briefly, Agarose plugs were incubated overnight at 50° C. in 250 µL of a 0.5M EDTA (pH 8), 1% Sarkosyl, 250 µg/mL proteinase K (Eurobio, code: GEXPRK01, France) solution, then washed twice in a Tris 10 mM, EDTA 1 mM solution for 30 in at room temperature.

**[0123]** Final Extraction of DNA and Molecular Combing

**[0124]** Plugs of embedded DNA from viral particles or corneas were treated for combing DNA as previously described (Schurra and Bensimon 2009). Briefly, plugs were melted at 68° C. in a MES 0.5 M (pH 5.5) solution for 20 min, and 1.5 units of beta-agarase (New England Biolabs, ref. M0392S, MA, USA) was added and left to incubate for up to 16 h at 42° C. The DNA solution was then poured in a Teflon reservoir and Molecular Combing was performed using the Molecular Combing System (Genomic Vision S.A., Paris, France) and Molecular Combing coverslips (20 mm×20 mm, Genomic Vision S.A., Paris, France). The combed surfaces were dried for 4 hours at 60° C.

**[0125]** Synthesis and Labelling of HSV-1 Probes

**[0126]** The coordinates of all the probes relative to the Genbank sequence NC\_001806.1 are listed in Table A. Probe size ranges from 1110 to 9325 bp in this example.

**[0127]** The HSV-1 specific probes were produced by either *SacI* or *BspEI* (New England Biolabs Inc., Beverly, Mass., USA code R0156L and R0156L, respectively) enzymatic digestion of the HSV-1 sc16 strain obtained from the CNRS (Prof. Marc Labetoulle, laboratoire de virologie moléculaire et structurale, UMR CNRS 2472-INRA 1157, Gif-sur-Yvette, France) or by long-range PCR using LR Taq DNA polymerase (Roche, kit code: 11681842001) using the primers listed in table B and the DNA from HSV-1 sc16 as template DNA. *SacI* and *BspEI* HSV-1 fragments were ligated in *SacI* and *XmaI*-digested pNEB193 plasmid (New England Biolabs Inc., Beverly, Mass., USA, code N3051S), respectively. PCR products were ligated in the pCR®2.1 vector using the TOPO® TA cloning Kit (Invitrogen, France, code K455040). The two extremities of each probe were sequenced for verification purpose. The apparent H1 (21 kb), H2 (56 kb), H3 (45 kb), H4 (13 kb), H5 (7.5 kb) and H6 (6.5 kb) probes are mixes of several adjacent or overlapping probes listed in table A. For the visualisation of the non canonical structure implicated the H4 and H6 probes, the apparent H2 probe was split in two probes H2A (31 kb) and H2B (25 kb).

**[0128]** The labelling of the probes was performed using conventional random priming protocols. For biotin-11-dCTP labelling, the BioPrime® DNA kit (Invitrogen, code: 18094-011, CA, USA) was used according to the manufacturer's instruction, except the labelling reaction was allowed to proceed overnight. For 11-digoxigenin-dUTP and Alexa488-7-OBEA-dCTP, the dNTP mix from the kit was replaced by the mix specified in table C. 200 ng of each plasmid was labelled in separate reactions. For isomer classification, H1, H3, H5 probes were labelled with 11-digoxigenin-dUTP while H2, H4 and H6 probes were labelled with biotin-11-dCTP. For the visualisation of the non canonical structure implicated the H4 and H6 probes, H1, H2B and H3 probes were labelled with

biotin-11-dCTP, H2A and H5 probes with 11-digoxigenin-dUTP and H4, H6 with Alexa 488-7-OBEA-dCTP. The reaction products were visualized on an agarose gel to verify the synthesis of DNA.

**[0129]** Hybridization of HSV-1 Probes on Combed Viral DNA and Detection

**[0130]** Subsequent steps were also performed essentially as previously described in Schurra and Bensimon, 2009 (Schurra and Bensimon 2009) which is incorporated by reference. Briefly, a mix of labelled probes (250 ng of each probe, see below for details regarding probe synthesis and labelling) were ethanol-precipitated together with 10 µg hering sperm DNA and 2.5 µg Human Cot-1 DNA (Invitrogen, ref. 15279-011, CA, USA), resuspended in 20 µL of hybridization buffer (50% formamide, 2×SSC, 0.5% SDS, 0.5% Sarkosyl, 10 mM NaCl, 30% Block-aid (Invitrogen, ref. B-10710, CA,USA). The probe solution and probes were heat-denatured together on the Hybridizer (Dako, ref. S2451) at 90° C. for 5 min and hybridization was left to proceed on the Hybridizer overnight at 37° C. Slides were washed 3 times in 50% formamide, 2×SSC and 3 times in 2×SSC solutions, for 5 min at room temperature. Detection antibody layers and their respective dilution in Block-Aid are described in table D and E. For each layer, 20 µL of the antibody solution was added on the slide and covered with a combed coverslip and the slide was incubated in humid atmosphere at 37° C. for 20 min. The slides were washed 3 times in a 2×SSC, 1% Tween20 solution for 3 min at room temperature between each layer and after the last layer. For isomer classification, detection was carried out using a Texas Red coupled mouse anti digoxigenin (Jackson ImmunoResearch, France) antibody in a 1:25 dilution for H1, H3, and H5 probes, and an Alexa488-coupled streptavidin antibody (Invitrogen, France) in a 1:25 dilution for H2, H4 and H6 probes as primary antibodies. As second layer, an Alexa594-coupled goat anti mouse (Invitrogen, France) diluted at 1:25 and a biotinylated goat antistreptavidin (Vector Laboratories, UK) diluted at 1:50 were used. To amplify the Alexa488-fluorescence signal of H2, H4 and H6 probes, an additional detection layer was realized by using the same Alexa488 coupled-streptavidin used for the first layer at a 1:25 dilution. For the visualisation of the non-canonical structure implicated the H4 and H6 probes, detection was carried out using an AMCA-coupled mouse anti-digoxigenin (Jackson ImmunoResearch, France) antibody in a 1:25 dilution for the H2A and H5 probes, an Alexa 594-coupled streptavidin antibody (Invitrogen, France) in a 1:25 dilution for the H1, H2B and H3 and a rabbit anti Alexa 488 in a 1:25 dilution as primary antibodies. As second layer, an Alexa 594-coupled goat anti mouse (Invitrogen, France) diluted at 1:25, a biotinylated goat anti-streptavidin (Vector Laboratories, UK) diluted at 1:50 and an Alexa 488-coupled goat anti-mouse diluted at 1:25 were used. To amplify the Alexa 594-fluorescence signal of the H1, H2B and H3 probes and the AMCA/Alexa 350 signal of the H2A and H5 probes, an additional detection layer was realized by using the same Alexa 594 coupled-streptavidin used for the first layer at a 1:25 dilution and an Alexa350 coupled goat anti rat diluted at 1:25, respectively. After the last washing steps, all glass cover slips were dehydrated in ethanol and air dried.

**[0131]** Analysis of HSV-1 Detected Signals

**[0132]** For direct visualisation of combed HSV-1 fibers, cover slips were mounted with 20 µL of a Prolong (Invitrogen, France ref P36930)-YOYO-1 iodide (Molecular Probes,

code Y3601) mixture (1/1000 v/v) and scanned with inverted automated epifluorescence microscope, equipped with a 40× objective (ImageXpress Micro, Molecular Devices, USA). Length of the YOYO-1-stained DNA fibers were measured and converted to kb using an extension factor of 2 kb/μm (Schurra and Bensimon 2009), with an internal software Gvlab 04.2.1 (Genomic Vision S.A., Paris, France).

**[0133]** For isomers classification, hybridized-combed DNA from viral particles or cornea preparation were scanned without any mounting medium using an inverted automated epifluorescence microscope, equipped with a 40× objective (ImageXpress Micro, Molecular Devices, USA) and the signals can be detected visually or automatically by an in house software (Gvlab 0.4.2). Both FISH signals composed of a continuous signal of Texas Red/Alexa 594-fluorescence for H1, H3, H5, and Alexa 488-fluorescence for H2, H4 and H6, and signals composed of a continuous signal corresponding to one of the pattern described below were considered:

**[0134]** (a) A minimum of 28 kb long of Texas Red/Alexa 594-fluorescence signal, directly followed by an Alexa488-fluorescence signal corresponding to H4 and another Texas Red/Alexa594-fluorescence signal of a length minimal of 3 kb.

**[0135]** (b) A minimum of 28 kb long of Texas Red/Alexa 594-fluorescence signal, directly followed by an Alexa 488-fluorescence signal corresponding to H6 and another Texas Red/Alexa 594-fluorescence signal of a length minimal of 3 kb.

**[0136]** (c) A minimum of 3 kb an Alexa 488 fluorescence signal directly followed by a Texas Red/Alexa 594 fluorescence signal corresponding to H1, next to an Alexa 488-fluorescence signal corresponding to H4 and another Texas Red/Alexa 594-fluorescence signal of a length minimal of 3 kb.

**[0137]** (d) A minimum of 3 kb an Alexa 488-fluorescence signal directly followed by a Texas Red/Alexa 594-fluorescence signal corresponding to H1, next to an Alexa 488-fluorescence signal corresponding to H6 and another Texas Red/Alexa 594-fluorescence signal of a length minimal of 3 kb.

**[0138]** FISH signals selected were then classified depending of the pattern of the continuous FISH signals analysed:

**[0139]** The probe array composed of H1/H2/H3/H4/H5/H6 probes or the pattern (a) is classified as a Prototype (P) form of HSV-1 (Hayward, Jacob et al. 1975)

**[0140]** The pattern H1/H2/H3/H6/H5/H4 or the pattern (b) is classified as a Inverted Short (IS) genomic region form of HSV-1 (Hayward, Jacob et al. 1975)

**[0141]** The pattern H3/H2/H1/H4/H5/H6 or the pattern (c) is classified as a Inversed Long (IL) genomic region of HSV-1 (Hayward, Jacob et al. 1975)

**[0142]** The pattern H3/H2/H1/H6/H5/H4 or the pattern (d) is classified as a Inversed Long and Short (ILS) genomic region form of HSV-1 (Hayward, Jacob et al. 1975)

**[0143]** The hypothesis that the observed distribution differ significantly to the expected distribution (an equivalent number of events for the four isomers, (Bataille and Epstein 1997) was tested by a chi-square test, and accepted when the p-value observed was below 0.05.

**[0144]** For the visualisation of the non canonical structure implicated the H4 and H6 probes, hybridized-combed DNA from infected cells preparation were scanned without any mounting medium using an inverted automated epifluorescence microscope, equipped with a 40× objective (ImageX-

press Micro, Molecular Devices, USA) and the signals can be detected visually on an in house software (Gvlab 0.4.2). All signals composed of a continuous signal of Alexa 488-fluorescence for H4 and H6 were selected. Signals corresponding to one of the pattern below were classified as canonical structure:

**[0145]** (e) A continuous signal composed of a minimum of 3kb of an Alexa 594-fluorescence signal, followed of 13 kb of an Alexa 488-fluorescence signal corresponding to the H4 probe, 7.5 kb of an AMCA/Alexa 350-fluorescence signal corresponding to the H5 probe and 7 kb of an Alexa 488-fluorescence corresponding to the H6 probe.

**[0146]** (f) A continuous signal composed of a minimum of 3 kb of an Alexa 594-fluorescence signal, followed of 7 kb of an Alexa 488-fluorescence signal corresponding to the H6 probe, 7.5 kb of an AMCA/Alexa 350-fluorescence signal corresponding to the H5 probe and 13 kb of an Alexa 488-fluorescence corresponding to the H4 probe.

**[0147]** All other signals were classified as non canonical structure. The proportions of the canonical and non canonical structure were compared.

**[0148]** Extraction of HSV-1 DNA from Viral Particles

**[0149]** During sample preparation many DNA molecules are sheared at random location due to uncontrolled manipulation forces resulting in high variability in the size of DNA prepared. It has been showed that high molecular weight DNA can be stretched on by Molecular Combing using a glass coverslip when it is deproteinised in a molten agarose plug (Lebofsky and Bensimon 2003). Thus, the analyzed DNA molecules are of variable length, with an average of about 300 kb, the longest molecules reaching several megabases.

**[0150]** Since HSV-1 DNA has never been used for Molecular Combing, the inventors first evaluated the quality of the DNA fibers in terms of length extracted by two different methods: the standard phenol:chloroform extraction and the method described by (Lebofsky, Heilig et al. 2006) that have been slightly modified as described in the "Example" section.

**[0151]** As shown in FIG. 2, no fiber of size above 80 μm was detected. This is in concordance with the maximum expected size for the HSV-1 genome (152 kb) considering a constant elongation factor of 2 kb. With both methods, we do not detect only 152 kb long DNA fibers because there is still random DNA shearing due to the mechanical manipulation that cannot be avoided and because DNA molecules shorter than the full-length standard HSV-1 viral genome can become encapsidated within nuclear capsids (Vlazny, Kwong et al. 1982).

**[0152]** However, the median size of DNA fibers is 36 kb with 1.2% of fiber longer than 140 kb when extraction has been performed with the standard phenol:chloroform method. In contrast, the median size of HSV-1 DNA fiber is 84 kb with 2.5% of fiber longer than 140 kb when the extraction of DNA from agarose plug-embedded viral particles using our alternative protocol has been realized. Although the proportion of long molecules is low, there are a lot of long combed DNA molecules available for analysis since there are several ten thousand fibers combed on a glass coverslip.

**[0153]** These results indicate that the alternative method developed by the inventors improved the quality of combed DNA extracted from viral particles compared to standard method allowing analysis by Molecular Combing.

**[0154]** Structure of the HSV-1 Genome in Viral Particles and its Distribution

**[0155]** The inventors applied Molecular Combing to uniformly stretch the HSV-1 DNA extracted from viral particles and infected cells and hybridized the resulting combed HSV-1 DNA with labeled adjacent and overlapping HSV-1-specific DNA probes (FIG. 1; H1, H3, H5: red Texas Red/Alexa 594-fluorescence; H2, H4, H6: green Alexa 488-fluorescence) to determine the structure of the HSV-1 genome (FIG. 3A). Immunofluorescence microscopy (FIG. 3B.) exhibited 405 multicolor linear patterns for each production of HSV-1 KOS strain viral particles produced from COS7, Vero and Neuro2 A cell lines that fulfilled the criteria for evaluation (see "Examples" section). Classification of the signal showed that the distribution of the HSV-1 KOS strain isomers are equivalently distributed in viral particles from COS7 cells whereas the P and IS isomers are the more frequent isomer in the viral particles produced from the Neuro2A and Vero cell lines.

**[0156]** On hybridized combed HSV-1 infected cells, the distribution of the four isomers was compared between Sc16 and KOS HSV-1 strains produced in different cell lines (BSR, COS7, Neuro2A and Vero). 405 multicolor linear patterns corresponding to each production and that fulfilled the criteria for evaluations (see "Experimental procedures" section) were classified. The distribution between the four isomers was statistically equivalent (Chi2 test) in all production of HSV-1 Sc16 (in BSR, Vero, Neuro 2a and COS-7 cells). In the same way, the distribution was found equivalent for HSV-1 strain KOS produced in COS-7 cells. Strikingly, for the first time, the inventors have found that the IS and P isomers are the predominant forms of the HSV-1 DNA strain KOS preparation from Vero and Neuro2A cells while IL isomers is the less present isomers.

**[0157]** The inventors have found that the Molecular Combing techniques as described herein are powerful methods for analysis of the structure of the HSV-I genome DNA at the level of the unique molecule and to quantify its distribution in a biological sample.

**[0158]** Extraction of Genomic DNA from Mouse and Rabbit Cornea

**[0159]** Molecular Combing has been successfully performed with DNA solution from isolated cells including cultured cells (i.e., established cell strains, immortalized primary cells) or biological fluids (i.e., peripheral blood lymphocytes, amniotic cells) (Gad, Klinger et al. 2002; Caburet, Conti et al. 2005). However, the human cornea is a solid tissue with a complex structure composed of 5 layers: the corneal epithelium, the collagen-rich Bowman's membrane, the corneal stroma which consisting of regularly-arranged collagen fibers along with sparsely distributed interconnected keratocytes), the acellular Descemet's membrane and, the corneal endothelium. In order to extract genomic DNA from cornea, the inventors developed a specific method to isolate corneal cells before proceeding with the standard procedure. Different methods including mechanical disruption and enzymatic digestion of cornea were tested. The latter was given the best results and was optimized using different types of proteases (i.e., trypsin, collagenase A, dispase . . . ) at different concentration. FIG. 4A and B shows an example of results obtained with both mouse and rabbit cornea that were digested with 0.3 mg/ml collagenase A and 0.8 mg/ml Dispase for 16 h. As for standard extraction for Molecular Combing genomic DNA is broken at random locations. Thus, the analyzed genomic

DNA molecules are of variable length, with an average of about 200 kb, the longest molecules above 1 Mb (megabases) for both type of cornea. The size of DNA fiber extracted from cornea is slightly inferior to the typical size that can be obtained from isolated cells (an average of about 300 kb with the longest molecules reaching several megabases).

**[0160]** Detection of HSV-1 Infection in Mouse Cornea

**[0161]** The inventors therefore adapted and applied Molecular Combing on DNA extracted from HSV-1 infected mouse cornea and hybridized with the HSV-1 specific probes as described above. As shown in FIG. 5, this enabled detection all the types of isomers of HSV-1 genome in mouse infected cornea.

**[0162]** In addition to the detection of mature HSV-1 genome, the Molecular Combing procedures of the invention allow the detection of concatemers (FIG. 6) indicating that the virus is actively replicating in the cornea of the analyzed sample.

**[0163]** Detection of Non-Canonical Forms in Infected Cells and Mouse Cornea

**[0164]** The inventors detected non-canonical structure of the HSV-1 genome (FIG. 7A) that probably arises from recombination during the replication of the virus in infected mouse cornea extract and in infected cells extracts.

**[0165]** The labelled adjacent and overlapping HSV-1 specific probes were hybridized on combed DNA extracts from HSV-1 strain Sc16 infected Vero cells (FIG. 7B; H1, H2A, H3: red Alexa 594-fluorescence that appears in grey; H2B, H5: blue AMCA/Alexa 350-fluorescence that appears in white; H4, H6: green Alexa 488-fluorescence that appears in black) to evaluate the proportion of the non-canonical structures in the H4/H6 region. A total of 367 multicolour linear patterns were classified, 20% (73) are found to have non-canonical H4/H6 structure.

**[0166]** Molecular Combing enables the visualisation of non canonical structure and by their infinity of combination of barcode possible is a powerful method to analyse them.

TABLE A

Name	Start	End	Size (bp)
HSV-B1	1	1323	1323
HSV-B2	1324	7259	5936
HSV-P4	9237	11276	2004
HSV-P5	11090	13245	2156
HSV-Sc4	13088	14971	1884
HSV-P6	14554	17565	3065
HSV-B4	14827	22438	7595
HSV-Sc7	17853	20762	2910
HSV-B7	22953	25152	2200
HSV-Sc11	23400	25447	2048
HSV-B8	25153	29997	4845
HSV-Sc14	27944	31215	3272
HSV-B10	30903	34697	3795
HSV-Sc16	33044	39471	6428
HSV-B13	35891	40272	4382
HSV-Sc18	40315	42288	1974
HSV-B15	41017	43898	2882
HSV-Sc21	42621	47358	4738
HSV-P8	44192	46987	2795
HSV-B18	44682	47174	2493
HSV-B19	47175	50931	3757
HSV-Sc23	49040	51392	2353
HSV-B21	50959	56138	5180
HSV-Sc24	51393	53348	1956
HSV-Sc25	53349	56049	2701
HSV-B22	56139	63370	7232
HSV-Sc30	56775	64599	7825

TABLE A-continued

Name	Start	End	Size (bp)
HSV-Sc31	64600	69017	4418
HSV-B24/25	65757	66423	4740
HSV-Sc32	69018	72802	3785
HSV-B26	70497	73717	3221
HSV-P8	73229	77332	4103
HSV-B27	73718	77164	3447
HSV-B28	77165	79105	1941
HSV-B30	79937	81056	1120
HSV-Sc44	80991	85801	4811
HSV-B31	81507	83780	2724
HSV-B33	84298	85576	1279
HSV-Sc45	85802	90164	4363
HSV-B35	86304	87747	1444
HSV-B38	89249	90453	1205
HSV-B39	90453	92176	1723
HSV-B40	92177	94195	2019
HSV-Sc47	91545	93723	2179
HSV-Sc49	94122	100285	6164
HSV-B41	94196	96148	1953
HSV-B42	96149	99454	3306
HSV-B44	99492	102440	2949
HSV-B45	102441	106065	3625
HSV-B46	106066	107175	1110
HSV-B48	108009	116579	8571
HSV-P8c	109960	113148	3188
HSV-Sc54	115744	125068	9325
HSV-B52	125044	129901	4858
HSV-Sc56	125079	128601	3523
HSV-Sc58	129089	133046	3958
HSV-B55	130841	135542	4702
HSV-Sc59	133047	137945	4899
HSV-B56	135543	137690	2148
HSV-Sc60	137946	140155	2210
HSV-P8d	138148	139821	1673
HSV-B58	138757	141926	3170
HSV-B60	142840	145515	2276
HSV-Sc64	144918	149148	4231

TABLE B

Probes	Forward Primer	Reverse Primer
HSV-P4	TGG TTG TGT TAC TGG GCA AA (SEQ ID NO: 1)	TCG ATC GAC GAC ACC ATA AA (SEQ ID NO: 2)
HSV-P5	CAG ATA CGA CTC CCG CAG AT (SEQ ID NO: 3)	CGA CGA CCT CGA CGT TAT TT (SEQ ID NO: 4)
HSV-P6	CGT GAG GTC CAA AAT CAC CT (SEQ ID NO: 5)	GAC AGG CAA GCT CAA AGT CC (SEQ ID NO: 6)
HSV-P8	AGA TGT CCA CGA GCA CCA G (SEQ ID NO: 7)	CCT GAC TTT GTG GGG CTA AA (SEQ ID NO: 8)

[0167] Primers sequences used for the synthesis of probes by long-range PCR. An extract of DNA from HSV-1 strain Sc16 is used as template.

TABLE C

Labelling	Non-labelled dNTPs (Invitrogen, ref. 10297-018, CA, USA)	Labelled dNTP (Roche, ref. 11 558 706 910, France)
Dig-dUTP	dATP, dCTP, dGTP 40 $\mu$ M each dTTP 20 $\mu$ M	Dig-11-dUTP 20 $\mu$ M
Alexa488-7- OBEA-dCTP	dATP, dTTP, dGTP 40 $\mu$ M each dCTP 20 $\mu$ M	Alexa488-7-OBEA-dCTP 20 $\mu$ M

[0168] Mixes used in replacement of the dNTP mix of the random priming kit for labelling with dig-dUTP and Alexa488-7-OBEA-dCTP. The concentrations indicated are the final concentration in the labelling reaction. The non-labelled dNTPs and the labelled dNTP were added together in replacement of the provided dNTP mix intended for labelling with biotin-11-dCTP-.

TABLE D

Description	Abbreviation	Supplier
Streptavidin, coupled to Alexa Fluor 488	Strep/A488	Invitrogen (France, S11223)
Goat anti-streptavidin, coupled to biotin	G anti-strep/biotin	Vector Laboratories (France; BA-0500)
Mouse anti-dig, coupled to Texas Red	M anti-DIG/TR	Jackson Immuno Research (France; 200-072-156)
Goat anti-mouse, coupled to A594	G anti-M/A594	Invitrogen (France; A11005)
Streptavidin, coupled to Alexa Fluor 594	Strep/A594	Invitrogen (France, S11227)
Mouse anti-Dig, AMCA coupled	M anti-DIG/AMCA	Jackson Immuno Research (France; 200-152-156)
Rat anti-mouse, AMCA coupled	R anti-M/AMCA	Jackson Immuno Research (France; 415-155-166)
Goat anti-rat, Alexa 350 coupled	G anti-R/A350	Invitrogen (France, A21093)
Rabbit anti Alexa 488	R anti-A488	Invitrogen (France, A11094)
Goat anti mouse, coupled Alexa 488	G anti-M/A488	Invitrogen (France, A11001)

[0169] List of antibodies and other hapten-binding molecules used for the detection of probes.

TABLE E

	1 <sup>st</sup> layer	2 <sup>nd</sup> layer	3 <sup>rd</sup> layer
<u>1-color scheme</u>			
Biotin/green	strep/A488 (1/25)	Goat anti-strep/ biotin (1/50)	strep/A488 (1/25)
<u>2-color scheme</u>			
Biotin/green	strep/A488 (1/25)	Goat anti-strep/ biotin (1/50)	strep/A488 (1/25)
Dig/red	Mouse anti- DIG/TR (1/25)	Goat anti-M/ A594 (1/25)	—

TABLE E-continued

	1 <sup>st</sup> layer	2 <sup>nd</sup> layer	3 <sup>rd</sup> layer
3-color scheme			
Biotin/red	Strep/A594 (1/25)	G anti-strep/ biotin (1/50)	Strep/A594 (1/25)
Dig/blue	M anti- DIG/AMCA (1/25)	R anti-M/AMCA (1/25)	G anti-R/A350 (1/25)
A488/green	R anti-A488 (1/25)	G anti-M/A488 (1/25)	—

**[0170]** Composition of the 2 or 3 layers for the detection of probes by fluorescence. The dilution for each detection agent is indicated in brackets. The abbreviations refer to table D.

#### Example 2

##### Human Immunodeficiency Virus Detection

**[0171]** Preparation of Embedded DNA Plugs from ACH-2 Cells Culture

**[0172]** ACH-2 cell lines (Clouse, Powell et al. 1989) were cultivated according to the authors' instructions. DNA was extracted as described in (Schurra and Bensimon 2009). Briefly, cells were resuspended in 1× PBS at a concentration of 10<sup>7</sup> cells/mL mixed thoroughly at a 1:1 ratio with a 1.2% w/v solution of low-melting point agarose (Nusieve GTG, ref. 50081, Cambrex) prepared in 1× PBS at 50° C. 90 µL of the cell/agarose mix was poured in a plug-forming well (BioRad, ref. 170-3713) and left to cool down at least 30 min at 4° C. Agarose plugs were incubated overnight at 50° C. in 250 µL of a 0.5M EDTA (pH 8), 1% Sarkosyl, 250 µg/mL proteinase K (Eurobio, code: GEXPRK01, France) solution, then washed twice in a Tris 10 mM, EDTA 1 mM solution for 30 in at room temperature.

**[0173]** Final Extraction of DNA and Molecular Combing

**[0174]** Plugs of embedded DNA from ACH-2 cells were treated for combing DNA as previously described (Schurra and Bensimon 2009). Briefly, plugs were melted at 68° C. in a MES 0.5 M (pH 5.5) solution for 20 min, and 1.5 units of beta-agarase (New England Biolabs, ref. M0392S, MA, USA) was added and left to incubate for up to 16 h at 42° C. The DNA solution was then poured in a Teflon reservoir and Molecular Combing was performed using the Molecular Combing System (Genomic Vision S.A., Paris, France) and Molecular Combing coverslips (20 mm×20 mm, Genomic Vision S.A., Paris, France). The combed surfaces were dried for 4 hours at 60° C.

**[0175]** Synthesis and Labelling of HIV-1 Probes

**[0176]** The coordinates of the three probes relative to the Genbank sequence M19921.1 are listed in table F. Probe size ranges from 2927 to 3749 bp in this example.

**[0177]** The HIV specific probes were produced by long-range PCR using LR Taq DNA polymerase (Roche, kit code: 11681842001) using the primers listed in table G and the DNA from HIV pNL4-3 as template DNA. PCR products were ligated in the pCR®2.1 vector using the TOPO® TA cloning Kit (Invitrogen, France, code K455040). The two extremities of each probe were sequenced for verification purpose.

**[0178]** Two fosmids G248P87988G9 and G248P86255A8 flanking the insertion site of HIV-1 or one fosmid

G248P84833H9 encompassing the HIV-1 provirus insertion site in ACH-2 cells (Ishida, Hamano et al. 2006), according to Human March 2006 Assembly (NCBI Build 36.1), and the HIV-1 probes were labeled using conventional random priming protocols. For biotin-11-dCTP labelling, the BioPrime® DNA kit (Invitrogen, code: 18094-011, CA, USA) was used according to the manufacturer's instruction, except the labelling reaction was allowed to proceed overnight. For digoxigenin-11-dUTP, the dNTP mix from the kit was replaced by the mix specified in table C. 200 ng of each plasmid/fosmid was labelled in separate reactions. For entire HIV-1 detection, HIV-1 was labelled with digoxigenin-11-dUTP while fosmids were labelled with biotin-11-dCTP. The reaction products were visualized on an agarose gel to verify the synthesis of DNA.

**[0179]** Hybridization of HIV-1 Probes on Combed Viral DNA and Detection

**[0180]** Subsequent steps were also performed essentially as previously described in Schurra and Bensimon, 2009 (Schurra and Bensimon 2009). Briefly, a mix of labelled probes (250 ng of each probe) were ethanol-precipitated together with 10 µg herring sperm DNA and 2.5 µg Human Cot-1 DNA (Invitrogen, ref. 15279-011, CA, USA), resuspended in 20 µL of hybridization buffer (50% formamide, 2×SSC, 0.5% SDS, 0.5% Sarkosyl, 10 mM NaCl, 30% Block-aid (Invitrogen, ref. B-10710, CA, USA). The probe solution and probes were heat-denatured together on the Hybridizer (Dako, ref. S2451) at 90° C. for 5 min and hybridization was left to proceed on the Hybridizer overnight at 37° C. Slides were washed 3 times in 50% formamide, 2×SSC and 3 times in 2×SSC solutions, for 5 min at room temperature. Detection antibody layers and their respective dilution in Block-Aid are described in table D and E. For each layer, 20 µL of the antibody solution was added on the slide and covered with a combed coverslip and the slide was incubated in humid atmosphere at 37° C. for 20 min. The slides were washed 3 times in a 2×SSC, 1% Tween20 solution for 3 min at room temperature between each layer and after the last layer.

**[0181]** Detection of entire HIV-1 was carried out using a Texas Red coupled mouse anti-digoxigenin (Jackson ImmunoResearch, France) antibody in a 1:25 dilution for HIV-1 probes, and an Alexa488-coupled streptavidin antibody (Invitrogen, France) in a 1:25 dilution for fosmids as primary antibodies. As second layer, an Alexa594-coupled goat anti mouse (Invitrogen, France) diluted at 1:25 and a biotinylated goat antistreptavidin (Vector Laboratories, UK) diluted at 1:50 were used. To amplify the Alexa488-fluorescence signal of fosmids, an additional detection layer was realized by using the same Alexa488 coupled-streptavidin used for the first layer at a 1:25 dilution. After the last washing steps, all glass cover slips were dehydrated in ethanol and air dried.

**[0182]** Analysis of HIV-1 Detected Signals

**[0183]** Hybridized-combed DNA from ACH-2 cells preparation were scanned without any mounting medium using an inverted automated epifluorescence microscope, equipped with a 40× objective (ImageXpress Micro, Molecular Devices, USA) and the signals can be detected visually or automatically by an in house software (Gvlab 0.4.2):

**[0184]** Using the two fosmids G248P87988G9 and G248P86255A8 flanking the insertion site of HIV-1, FISH signals corresponding to one of the pattern as follow were considered and measured, using an extension factor of 2 kb/µm (Schurra and Bensimon 2009):

- [0185]** FISH signals composed of a continuous signal of Texas Red/Alexa 594-fluorescence for HIV-1. The entire signal that correspond of isolated HIV proviral DNA is measured
- [0186]** FISH signal array composed of signal chain of Texas Red/Alexa 594-fluorescence for HIV-1, flanked by two gaps, and two continuous signals of Alexa 488-fluorescence corresponding to the fosmid sequences. The entire Texas Red/Alexa 594 fluorescence signal and gaps length flanking this signal were measured and corresponds to HIV-1 proviral DNA integrated in the chromosome 7 at 7p15 (Ishida, Hamano et al. 2006).
- [0187]** FISH signal array with two Alexa 488-fluorescence signals separated by a gap. Measurement of the gap length that corresponds to the 7p15 locus without integration of HIV-1 proviral DNA was performed.
- [0188]** Using the fosmid G248P84833H9 encompassing the HIV-1 provirus insertion site, FISH signals corresponding to one of the pattern as follow were considered and measured, using an extension factor of 2 kb/ $\mu$ m (Schurra and Bensimon 2009):
- [0189]** FISH signals composed of a continuous signal of Texas Red/Alexa 594-fluorescence for HIV-1. The entire signal that correspond of isolated HIV proviral DNA is measured
- [0190]** FISH signal array composed of a continuous signal chain of Texas Red/Alexa 594-fluorescence for HIV-1, flanked by two signals of Alexa 488-fluorescence corresponding to the fosmid sequences. The entire Texas Red/Alexa 594 fluorescence signal was measured and corresponds to HIV-1 proviral DNA integrated in the chromosome 7 at 7p15 (Ishida, Hamano et al. 2006).
- [0191]** FISH signal with one long Alexa 488-fluorescence signals corresponds to the 7p15 locus without integration of HIV-1 proviral DNA.
- [0192]** Detection of HIV-1 in ACH-2 Cells Culture
- [0193]** The inventors have applied Molecular Combing to detect complete HIV-1 integrated provirus in ACH-2 cell lines (Clouse, Powell et al. 1989), which contain a unique integrated form of HIV-1 in its genome in NT5C3 (Cytosolic 5'-nucleotidase III) gene on 7p14.3 (Ishida, Hamano et al. 2006). Labeled fosmids flanking the insertion site (G248P87988G9 and G248P86255A8) were hybridized on combed ACH-2 DNA simultaneously than labelled HIV-1 probes. A mean size of 10.2 kb $\pm$ 0.8 kb was obtained from measurement of 124 HIV-1 FISH signals flanking by one or both fosmids, corresponding to the expected size of HIV-1 (9.7 kb) (FIG. 8A, HIV probes: red Texas Red/Alexa 594-fluorescence, Fosmids: green Alexa 488-fluorescence). Normal alleles of NT5C3 gene are detected and measurement of the gap length between fosmids G248P87988G9 and G248P86255A8 FISH signal leads to a mean size of 32.8 $\pm$ 1.8 kb, lightly superior to the expected size of 31 kb, according to Human March 2006 Assembly (NCBI Build 36.1). This result shows that Molecular Combing by its resolution may have bring some further information about this locus (FIG. 8B). Furthermore, 133 isolated HIV-1 FISH signals were measured in this combed ACH-2 DNA preparation, with a mean size of 10.05 kb $\pm$ 0.79 kb (FIG. 8C). This contrasts with the expected unique HIV-1 site of insertion described previously (Ishida, Hamano et al. 2006) and suggests that it exists in

ACH-2 genome another or others insertion site(s) of HIV-1, or that a non integrated form of HIV-1 is persistent in ACH-2 nucleus. Similar observations are performed when the labeled fosmid encompassing the insertion site (G248P84833H9) was hybridized on combed ACH-2 DNA simultaneously than labelled HIV-1 probes (FIG. 8D). A mean size of 10.4 kb $\pm$ 0.5 kb was obtained from measurement of 57 HIV-1 FISH signals within the fosmid signal, corresponding to the expected size of HIV-1 provirus (9.7 kb) (HIV probes: red Texas Red/Alexa 594-fluorescence, fosmid: green Alexa 488-fluorescence). Furthermore, 35 isolated HIV-1 FISH signals were also detected and measured in this combed ACH-2 DNA preparation, with a mean size of 10.03 kb $\pm$ 0.82 kb (not shown).

**[0194]** These results indicate that Molecular Combing is a powerful method to analyze the structure of the HIV genome DNA and to quantify its integration in genomic DNA at the level of the unique molecule and in any biological sample.

TABLE F

Name	Start	End	Size (bp)
HIV-S1	1	3026	3026
HIV-S2	3018	5944	2927
HIV-S3	5961	9709	3749

**[0195]** Coordinates of the three probes used in this example, relative to the Genbank sequence M19921.1

TABLE G

Probes	Forward Primer	Reverse Primer
HIV-S1	TGGAAGGGCTAATTT GGTC (SEQ ID NO : 9)	TATTGCTGGTGATCCT TTCC (SEQ ID NO : 10)
HIV-S2	CCAGCAATATTCCAGT GTAGC (SEQ ID NO : 11)	TGAACAAACTTGGCA ATGA (SEQ ID NO : 12)
HIV-S3	CATCTCCTATGGCAGG AAGA (SEQ ID NO : 13)	TGCTAGAGATTTTCCA CACTGA (SEQ ID NO : 14)

**[0196]** Primers sequences used for the synthesis of probes by long-range PCR. An extract of DNA from HIV-1 pNL-4 is used as template.

## Example 3

## Detecting an Oncogene by Molecular Combing

**[0197]** In a manner to analogous to the detection of HSV genomic DNA in Example 1, probes are designed to detect the presence of a viral oncogene. Probes are designed to complement 80-100% of the active viral oncogene of interest and Molecular Combing is performed. The results indicate the presence of an active oncogene in a subject, leading to diagnosis and therapeutic intervention.

## Example 4

## Detecting Rearrangements of Infectious Viral DNA Using Molecular Combing

**[0198]** In a manner to analogous to the detection of HSV genomic DNA in Example 1, probes are designed to detect the presence of different arrangements of infectious viral DNA.

Different sets of probes tagged with different haptens recognized by different colored fluorescent probes are designed to complement 80-100% of the active viral oncogene of interest and Molecular Combing is performed. Variations in the arrangement of viral genes in a subject's cells, tissue or biological fluid are used to diagnose or prognose the risks of progression of the viral disease or disease associated with rearrangement of the viral genome, such as the risk of or induction of a tumorigenic properties by conversion of proto-oncogenes into oncogenes.

#### Example 5

##### Monitoring Genetic Therapy by Molecular Combing

[0199] In a manner to analogous to the detection of proviral forms of HIV-1 in Example 2, probes are specially designed to complement 80-100% of an integrated therapeutic adenovirus vector or the transgenic sequence(s) it carries. Molecular Combing is performed using the specially designed probes to detect the presence of transgenic material integrated into a host chromosome and whether it is arranged in form that can actively express the transgene(s). The quantity of transgene(s) in the subject is followed longitudinally and a determination is made when and whether to re-administer the therapeutic adenovirus vector.

#### REFERENCES

[0200] Patents

[0201] U.S. Pat. No. 6,130,044: Surfaces for biological reactions, process for preparing them and process for their use Bensimon D., Bensimon A., Heslot F. Oct. 10, 2000

[0202] U.S. Pat. No. 6,303,296, US2006257910, U.S. Pat. No. 6,054,327: Process for aligning macromolecules by passage of a meniscus and applications, Bensimon D., Bensimon A., Heslot F., Oct. 16, 2001.

[0203] U.S. Pat. No. 6,225,055: Apparatus for the parallel alignment of macromolecules, and use thereof, Bensimon A., Bensimon D., May 22, 2002.

[0204] U.S. 2004033510: Method for the diagnosis of genetic diseases by molecular combing and diagnostic kit, Bensimon A., Bensimon D., Michalet X., Feb. 19, 2004

[0205] EP0263025: Fractions containing antigens specific for herpes virus type 1 (HSV-1) and type 2 (HSV-2), method for isolating these fractions and diagnostic method specific for HSV-1 and/or HSV-2 using these fractions, Markoulatos P., Jun. 30, 1993.

[0206] EP0139416: Molecularly cloned diagnostic product and method of use, Berman, P. W., Lasky L. A., Nov. 5, 2003.

[0207] WO9818959: Procédé de diagnostic de maladies génétiques par Peignage Moléculaire et coffret de diagnostic, Bensimon A., Bensimon D., Michalet X., 1998 May 7.

[0208] WO0073503: Utilisation du peignage dans l'identification des origines de répllication d'ADN, Bensimon A., Herrick J., Hyrien O., 2000 Dec. 7.

[0209] WO0202131 : Compositions et méthodes de diagnostic et de traitement de l'infection du virus de l'herpès. Hosken N. A., Day C. H., Dillon D. C., McGowan P., Sleath P. R., 2002 Jan. 10.

[0210] WO2004036185 Diagnostic et traitement des maladies du virus de l'herpes simplex (HSV), Kriesel J. D.; Lepfert M. F.; Spruance S. L., Otterud B. E M, Peiffer A. P., 2004 Apr. 29.

[0211] WO2008028931: Genomic Morse Code, Lebofsky R., Bensimon A., Walrafen P., 2008 Mar. 13.

#### Scientific Publications

[0212] Barton, E. S., D. W. White, et al. (2007). "Herpesvirus latency confers symbiotic protection from bacterial infection." *Nature* 447(7142): 326-329.

[0213] Bataille, D. and A. L. Epstein (1997). "Equimolar generation of the four possible arrangements of adjacent L components in herpes simplex virus type 1 replicative intermediates." *J Virol* 71(10): 7736-7743.

[0214] Ben-Zeev, A., E. Weinberg, et al. (1974). "Isolation of herpes simplex virus DNA from virus particles and infected cells by electrophoresis in polyacrylamide gels." *J Gen Virol* 25(1): 63-73.

[0215] Bertoletti, A. and A. Gehring (2007). "Immune response and tolerance during chronic hepatitis B virus infection." *Hepatology Res* 37 Suppl 3: S331-338.

[0216] Boehmer, P. E. and I. R. Lehman (1997). "Herpes simplex virus DNA replication." *Annu Rev Biochem* 66: 347-384.

[0217] Breitbart, M. and F. Rohwer (2005). "Here a virus, there a virus, everywhere the same virus?" *Trends Microbiol* 13(6): 278-284.

[0218] Caburet, S., C. Conti, et al. (2005). "Human ribosomal RNA gene arrays display a broad range of palindromic structures." *Genome Res* 15(8): 1079-1085.

[0219] Cockerham, G. C., K. Bijwaard, et al. (2000). "Primary graft failure : a clinicopathologic and molecular analysis." *Ophthalmology* 107(11): 2083-2090; discussion 2090-2081.

[0220] Collier, L. H. and J. S. Oxford (2006). *Human virology : a text for students of medicine, dentistry, and microbiology*. Oxford ; New York, Oxford University Press.

[0221] Crouse, C. A., S. C. Pflugfelder, et al. (1990). "Detection of herpes viral genomes in normal and diseased corneal epithelium." *Curr Eye Res* 9(6): 569-581.

[0222] Davison, A. J. and N. M. Wilkie (1981). "Nucleotide sequences of the joint between the L and S segments of herpes simplex virus types 1 and 2." *J Gen Virol* 55(Pt 2): 315-331.

[0223] Dimmock, N. J., A. J. Easton, et al. (2007). *Introduction to modern virology*. Malden, MA, Blackwell Pub.

[0224] El-Aal, A. M., M. El Sayed, et al. (2006). "Evaluation of herpes simplex detection in corneal scrapings by three molecular methods." *Curr Microbiol* 52(5): 379-382.

[0225] Fields, B. N., D. M. Knipe, et al. (2007). *Fields' virology*. Philadelphia, Wolters kluwer/Lippincott Williams & Wilkins.

[0226] Fukuda, M., T. Deai, et al. (2008). "Presence of a large amount of herpes simplex virus genome in tear fluid of herpetic stromal keratitis and persistent epithelial defect patients." *Semin Ophthalmol* 23(4): 217-220.

[0227] Gad, S., A. Aurias, et al. (2001). "Color bar coding the BRCA1 gene on combed DNA: a useful strategy for detecting large gene rearrangements." *Genes Chromosomes Cancer* 31(1): 75-84.

[0228] Gad, S., I. Bieche, et al. (2003). "Characterisation of a 161 kb deletion extending from the NBR1 to the BRCA1 genes in a French breast-ovarian cancer family." *Hum Mutat* 21(6): 654.

[0229] Gad, S., V. Caux-Moncoutier, et al. (2002). "Significant contribution of large BRCA1 gene rearrangements in 120 French breast and ovarian cancer families." *Oncogene* 21(44): 6841-6847.

- [0230] Gad, S., M. Klinger, et al. (2002). "Bar code screening on combed DNA for large rearrangements of the BRCA1 and BRCA2 genes in French breast cancer families." *J Med Genet* 39(11): 817-821.
- [0231] Gardella, T., P. Medveczky, et al. (1984). "Detection of circular and linear herpesvirus DNA molecules in mammalian cells by gel electrophoresis." *J Virol* 50(1): 248-254.
- [0232] Goudsmit, J. (1998). *Viral sex: the nature of AIDS*. New York ; Oxford, Oxford University Press.
- [0233] group, T. H. E. D. S. (1998). "Acyclovir for the prevention of recurrent herpes simplex virus eye disease. Herpetic Eye Disease Study Group." *N Engl J Med* 339(5): 300-306.
- [0234] Hampson, A. W. and J. S. Mackenzie (2006). "The influenza viruses." *Med J Aust* 185(10 Suppl): S39-43.
- [0235] Hayward, G. S., R. J. Jacob, et al. (1975). "Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components." *Proc Natl Acad Sci USA* 72(11): 4243-4247.
- [0236] Herrick, J., C. Conti, et al. (2005). "Genomic organization of amplified MYC genes suggests distinct mechanisms of amplification in tumorigenesis." *Cancer Res* 65(4): 1174-1179.
- [0237] Herrick, J., S. Jun, et al. (2002). "Kinetic model of DNA replication in eukaryotic organisms." *J Mol Biol* 320(4): 741-750.
- [0238] Herrick, J., X. Michalet, et al. (2000). "Quantifying single gene copy number by measuring fluorescent probe lengths on combed genomic DNA." *Proc Natl Acad Sci USA* 97(1): 222-227.
- [0239] Herrick, J., P. Stanislawski, et al. (2000). "Replication fork density increases during DNA synthesis in *X. laevis* egg extracts." *J Mol Biol* 300(5): 1133-1142.
- [0240] Kabra, A., P. Lalitha, et al. (2006). "Herpes simplex keratitis and visual impairment: a case series." *Indian J Ophthalmol* 54(1): 23-27.
- [0241] Kessler, H. H., G. Muhlbauer, et al. (2000). "Detection of Herpes simplex virus DNA by real-time PCR." *J Clin Microbiol* 38(7): 2638-2642.
- [0242] Kotronias, D. and N. Kapranos (1998). "Detection of herpes simplex virus DNA in human spermatozoa by in situ hybridization technique." *In Vivo* 12(4): 391-394.
- [0243] Labetoulle, M., P. Auquier, et al. (2005). "Incidence of herpes simplex virus keratitis in France." *Ophthalmology* 112(5): 888-895.
- [0244] Lebofsky, R. and A. Bensimon (2003). "Single DNA molecule analysis: applications of molecular combing." *Brief Funct Genomic Proteomic* 1(4): 385-396.
- [0245] Lebofsky, R. and A. Bensimon (2005). "DNA replication origin plasticity and perturbed fork progression in human inverted repeats." *Mol Cell Biol* 25(15): 6789-6797.
- [0246] Lebofsky, R., R. Heilig, et al. (2006). "DNA replication origin interference increases the spacing between initiation events in human cells." *Mol Biol Cell* 17(12): 5337-5345.
- [0247] Leger, J., B. Larroque, et al. (2001). "Influence of severity of congenital hypothyroidism and adequacy of treatment on school achievement in young adolescents: a population-based cohort study." *Acta Paediatr* 90(11): 1249-1256.
- [0248] Liesegang, T. J. (1989). "Epidemiology of ocular herpes simplex. Natural history in Rochester, Minn., 1950 through 1982." *Arch Ophthalmol* 107(8): 1160-1165.
- [0249] Liesegang, T. J. (2001). "Herpes simplex virus epidemiology and ocular importance." *Cornea* 20(1): 1-13.
- [0250] Liesegang, T. J., L. J. Melton, 3rd, et al. (1989). "Epidemiology of ocular herpes simplex. Incidence in Rochester, Minn., 1950 through 1982." *Arch Ophthalmol* 107(8): 1155-1159.
- [0251] Lomholt, J. A., K. Baggesen, et al. (1995). "Recurrence and rejection rates following corneal transplantation for herpes simplex keratitis." *Acta Ophthalmol Scand* 73(1): 29-32.
- [0252] Lukashev, A. N. (2005). "Role of recombination in evolution of enteroviruses." *Rev Med Virol* 15(3): 157-167.
- [0253] Margolis, T. P., F. L. Elfman, et al. (2007). "Spontaneous reactivation of herpes simplex virus type 1 in latently infected murine sensory ganglia." *J Virol* 81(20): 11069-11074.
- [0254] Matsumoto, T., O. Yamada, et al. (1992). "Rapid DNA diagnosis of herpes simplex virus serotypes." *J Virol Methods* 40(1): 119-125.
- [0255] McGeoch, D. J., M. A. Dalrymple, et al. (1988). "The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1." *J Gen Virol* 69(Pt 7): 1531-1574.
- [0256] McGeoch, D. J., A. Dolan, et al. (1986). "Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1." *Nucleic Acids Res* 14(4): 1727-1745.
- [0257] Metzner, K. J. (2006). "Detection and significance of minority quasispecies of drug-resistant HIV-1." *J HIV Ther* 11(4): 74-81.
- [0258] Michalet, X., R. Ekong, et al. (1997). "Dynamic molecular combing: stretching the whole human genome for high-resolution studies." *Science* 277(5331): 1518-1523.
- [0259] Mocarski, E. S. and B. Roizman (1982). "Structure and role of the herpes simplex virus DNA termini in inversion, circularization and generation of virion DNA." *Cell* 31(1): 89-97.
- [0260] Namvar, L., S. Olofsson, et al. (2005). "Detection and typing of Herpes Simplex virus (HSV) in mucocutaneous samples by TaqMan PCR targeting a gB segment homologous for HSV types 1 and 2." *J Clin Microbiol* 43(5): 2058-2064.
- [0261] Ou, C. Y., S. Kwok, et al. (1988). "DNA amplification for direct detection of HIV-1 in DNA of peripheral blood mononuclear cells." *Science* 239(4837): 295-297.
- [0262] Pan, X. P., L. J. Li, et al. (2007). "Differences of YMDD mutational patterns, precore/core promoter mutations, serum HBV DNA levels in lamivudine-resistant hepatitis B genotypes B and C." *J Viral Hepat* 14(11): 767-774.
- [0263] Pasero, P., A. Bensimon, et al. (2002). "Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus." *Genes Dev* 16(19): 2479-2484.
- [0264] Patel, P. K., B. Arcangioli, et al. (2006). "DNA replication origins fire stochastically in fission yeast." *Mol Biol Cell* 17(1): 308-316.
- [0265] Pramod, N. P., P. Rajendran, et al. (1999). "Herpes simplex keratitis in South India: clinico-virological correlation." *Jpn J Ophthalmol* 43(4): 303-307.
- [0266] Pressing, J. and D. C. Reaney (1984). "Divided genomes and intrinsic noise." *J Mol Evol* 20(2): 135-146.
- [0267] Ramaswamy, M. and A. M. Geretti (2007). "Interactions and management issues in HSV and HIV coinfection." *Expert Rev Anti Infect Ther* 5(2): 231-243.

- [0268] Rao, V. A., C. Conti, et al. (2007). "Endogenous gamma-H2AX-ATM-Chk2 checkpoint activation in Bloom's syndrome helicase deficient cells is related to DNA replication arrested forks." *Mol Cancer Res* 5(7): 713-724.
- [0269] Reisinger, J., S. Rumpler, et al. (2006). "Visualization of episomal and integrated Epstein-Barr virus DNA by fiber fluorescence in situ hybridization." *Int J Cancer* 118(7): 1603-1608.
- [0270] Rodrigues, C., M. Deshmukh, et al. (2001). "Significance of HBV DNA by PCR over serological markers of HBV in acute and chronic patients." *Indian J Med Microbiol* 19(3): 141-144.
- [0271] Roizman, B. (1979). "The structure and isomerization of herpes simplex virus genomes." *Cell* 16(3): 481-494.
- [0272] Rowley, A. H., R. J. Whitley, et al. (1990). "Rapid detection of herpes-simplex-virus DNA in cerebrospinal fluid of patients with herpes simplex encephalitis." *Lancet* 335(8687): 440-441.
- [0273] Schurra, C. and A. Bensimon (2009). "Combing genomic DNA for structural and functional studies." *Methods Mol Biol* 464: 71-90.
- [0274] Severini, A., A. R. Morgan, et al. (1994). "Study of the structure of replicative intermediates of HSV-1 DNA by pulsed-field gel electrophoresis." *Virology* 200(2): 428-435.
- [0275] Shibata, D. K., N. Arnheim, et al. (1988). "Detection of human papilloma virus in paraffin-embedded tissue using the polymerase chain reaction." *J Exp Med* 167(1): 225-230.
- [0276] Sugita, S., N. Shimizu, et al. (2008). "Use of multiplex PCR and real-time PCR to detect human herpes virus genome in ocular fluids of patients with uveitis." *Br J Ophthalmol* 92(7): 928-932.
- [0277] Susloparov, M. A., I. M. Susloparov, et al. (2006). "[Herpes simplex virus type 1 and 2 (HSV-1,2) DNA detection by PCR during genital herpes]." *Mol Gen Mikrobiol Virusol*(1): 38-41.
- [0278] Taylor, T. J., M. A. Brockman, et al. (2002). "Herpes simplex virus." *Front Biosci* 7: d752-764.
- [0279] Umene, K. (1999). "Mechanism and application of genetic recombination in herpesviruses." *Rev Med Virol* 9(3): 171-182.
- [0280] Vlazny, D. A., A. Kwong, et al. (1982). "Site-specific cleavage/packaging of herpes simplex virus DNA and the selective maturation of nucleocapsids containing full-length viral DNA." *Proc Natl Acad Sci USA* 79(5): 1423-1427.
- [0281] Wadsworth, S., R. J. Jacob, et al. (1975). "Anatomy of herpes simplex virus DNA. II. Size, composition, and arrangement of inverted terminal repetitions." *J Virol* 15(6): 1487-1497.
- [0282] Wang, K., T. Y. Lau, et al. (2005). "Laser-capture microdissection: refining estimates of the quantity and distribution of latent herpes simplex virus 1 and varicella-zoster virus DNA in human trigeminal Ganglia at the single-cell level." *J Virol* 79(22): 14079-14087.
- [0283] Whitley, R. J. and B. Roizman (2001). "Herpes simplex virus infections." *Lancet* 357(9267): 1513-1518.
- [0284] Wildy, P., W. C. Russell, et al. (1960). "The morphology of herpes virus." *Virology* 12: 204-222.
- [0285] Wilhelmus, K. R., D. J. Coster, et al. (1981). "Prognosis indicators of herpetic keratitis. Analysis of a five-year observation period after corneal ulceration." *Arch Ophthalmol* 99(9): 1578-1582.
- [0286] Worobey, M. and E. C. Holmes (1999). "Evolutionary aspects of recombination in RNA viruses." *J Gen Virol* 80 (Pt 10): 2535-2543.
- [0287] Yu, X., B. Shi, et al. (2009). "A random PCR screening system for the identification of type 1 human herpes simplex virus." *J Virol Methods* 161(1): 91-97.
- [0288] HIV-Related References
- [0289] (2000). "Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Concerted Action on SeroConversion to AIDS and Death in Europe." *Lancet* 355(9210): 1131-1137.
- [0290] Buchbinder, S. P., M. H. Katz, et al. (1994). "Long-term HIV-1 infection without immunologic progression." *AIDS* 8(8): 1123-1128.
- [0291] Carrillo-Infante, C., G. Abbadessa, et al. (2007). "Viral infections as a cause of cancer (review)." *Int J Oncol* 30(6): 1521-1528.
- [0292] Chun, T. W., D. C. Nickle, et al. (2005). "HIV-infected individuals receiving effective antiviral therapy for extended periods of time continually replenish their viral reservoir." *J Clin Invest* 115(11): 3250-3255.
- [0293] Clouse, K. A., D. Powell, et al. (1989). "Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone." *J Immunol* 142(2): 431-438.
- [0294] del Rio, C. (2006). "Current concepts in antiretroviral therapy failure." *Top HIV Med* 14(3): 102-106.
- [0295] Douek, D. C., J. M. Brenchley, et al. (2002). "HIV preferentially infects HIV-specific CD4+ T cells." *Nature* 417(6884): 95-98.
- [0296] Douek, D. C., M. Roederer, et al. (2009). "Emerging concepts in the immunopathogenesis of AIDS." *Annu Rev Med* 60: 471-484.
- [0297] Douglas, J. T. (2007). "Adenoviral vectors for gene therapy." *Mol Biotechnol* 36(1): 71-80.
- [0298] Espy, M. J., J. R. Uhl, et al. (2006). "Real-time PCR in clinical microbiology: applications for routine laboratory testing." *Clin Microbiol Rev* 19(1): 165-256.
- [0299] Fischer, M., B. Joos, et al. (2004). "Cellular viral rebound after cessation of potent antiretroviral therapy predicted by levels of multiply spliced HIV-1 RNA encoding nef." *J Infect Dis* 190(11): 1979-1988.
- [0300] Frenkel, L. M., Y. Wang, et al. (2003). "Multiple viral genetic analyses detect low-level human immunodeficiency virus type 1 replication during effective highly active antiretroviral therapy." *J Virol* 77(10): 5721-5730.
- [0301] Grizot, S., J. Smith, et al. (2009). "Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease." *Nucleic Acids Res* 37(16): 5405-5419.
- [0302] Havlir, D. V., M. C. Strain, et al. (2003). "Productive infection maintains a dynamic steady state of residual viremia in human immunodeficiency virus type 1-infected persons treated with suppressive antiretroviral therapy for five years." *J Virol* 77(20): 11212-11219.
- [0303] Hermankova, M., J. D. Siliciano, et al. (2003). "Analysis of human immunodeficiency virus type 1 gene expression in latently infected resting CD4+ T lymphocytes in vivo." *J Virol* 77(13): 7383-7392.

- [0304] Herrick, J., C. Conti, et al. (2005). "Genomic organization of amplified MYC genes suggests distinct mechanisms of amplification in tumorigenesis." *Cancer Res* 65(4): 1174-1179.
- [0305] Huser, D. and R. Heilbronn (2003). "Adeno-associated virus integrates site-specifically into human chromosome 19 in either orientation and with equal kinetics and frequency." *J Gen Virol* 84(Pt 1): 133-137.
- [0306] Ishida, T., A. Hamano, et al. (2006). "5' long terminal repeat (LTR)-selective methylation of latently infected HIV-1 provirus that is demethylated by reactivation signals." *Retrovirology* 3: 69.
- [0307] Lawn, S. D. (2004). "AIDS in Africa: the impact of coinfections on the pathogenesis of HIV-1 infection." *J Infect* 48(1): 1-12.
- [0308] Morgan, D., C. Mahe, et al. (2002). "HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries?" *AIDS* 16(4): 597-603.
- [0309] Parsonnet, J. (1999). *Microbes and malignancy: infection as a cause of human cancers*. New York; Oxford, Oxford University Press.
- [0310] Ramratnam, B., R. Ribeiro, et al. (2004). "Intensification of antiretroviral therapy accelerates the decay of the HIV-1 latent reservoir and decreases, but does not eliminate, ongoing virus replication." *J Acquir Immune Defic Syndr* 35(1): 33-37.
- [0311] Schneider, M. F., S. J. Gange, et al. (2005). "Patterns of the hazard of death after AIDS through the evolution of antiretroviral therapy: 1984-2004." *AIDS* 19(17): 2009-2018.
- [0312] Schurra, C. and A. Bensimon (2009). "Combing genomic DNA for structural and functional studies." *Methods Mol Biol* 464: 71-90.
- [0313] Siliciano, J. D., J. Kajdas, et al. (2003). "Long-term follow-up studies confirm the stability of the latent reservoir for HIV-1 in resting CD4+ T cells." *Nat Med* 9(6): 727-728.
- [0314] Thomas, C. E., A. Ehrhardt, et al. (2003). "Progress and problems with the use of viral vectors for gene therapy." *Nat Rev Genet* 4(5): 346-358.
- [0315] Tobin, N. H., G. H. Learn, et al. (2005). "Evidence that low-level viremias during effective highly active antiretroviral therapy result from two processes: expression of archival virus and replication of virus." *J Virol* 79(15): 9625-9634.
- [0316] Weiss, R. A. (1993). "How does HIV cause AIDS?" *Science* 260(5112): 1273-1279.
- [0317] Zhou, H. S., N. Zhao, et al. (2007). "Site-specific transfer of an intact beta-globin gene cluster through a new targeting vector." *Biochem Biophys Res Commun* 356(1): 32-37.

## INCORPORATION BY REFERENCE

[0318] Each document, patent, patent application or patent publication cited by or referred to in this disclosure is incorporated by reference in its entirety, especially with respect to the specific subject matter surrounding the citation of the reference in the text. However, no admission is made that any such reference constitutes background art and the right to challenge the accuracy and pertinence of the cited documents is reserved.

## SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 24

<210> SEQ ID NO 1  
 <211> LENGTH: 20  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Forward Primer HSV-P4

<400> SEQUENCE: 1

tggttggtggt actgggcaaa 20

<210> SEQ ID NO 2  
 <211> LENGTH: 20  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Reverse Primer HSV-P4

<400> SEQUENCE: 2

tcgatcgagc acaccataaa 20

<210> SEQ ID NO 3  
 <211> LENGTH: 20  
 <212> TYPE: DNA  
 <213> ORGANISM: Artificial  
 <220> FEATURE:  
 <223> OTHER INFORMATION: Forward Primer HSV-P5

<400> SEQUENCE: 3

---

-continued

---

cagatacgcac tcccgcagat 20

<210> SEQ ID NO 4  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial  
<220> FEATURE:  
<223> OTHER INFORMATION: Reverse Primer HSV-P5

<400> SEQUENCE: 4

cgacgcacctc gacgttattt 20

<210> SEQ ID NO 5  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial  
<220> FEATURE:  
<223> OTHER INFORMATION: Forward Primer HSV-P6

<400> SEQUENCE: 5

cgtgaggtcc aaaatcacct 20

<210> SEQ ID NO 6  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial  
<220> FEATURE:  
<223> OTHER INFORMATION: Reverse Primer HSV-P6

<400> SEQUENCE: 6

gacaggcaag ctcaaagtcc 20

<210> SEQ ID NO 7  
<211> LENGTH: 19  
<212> TYPE: DNA  
<213> ORGANISM: Artificial  
<220> FEATURE:  
<223> OTHER INFORMATION: Forward Primer HSV-P8

<400> SEQUENCE: 7

agatgtccac ggcaccag 19

<210> SEQ ID NO 8  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Artificial  
<220> FEATURE:  
<223> OTHER INFORMATION: Reverse Primer HSV-P8

<400> SEQUENCE: 8

cctgactttg tggggctaaa 20

<210> SEQ ID NO 9  
<211> LENGTH: 19  
<212> TYPE: DNA  
<213> ORGANISM: Artificial  
<220> FEATURE:  
<223> OTHER INFORMATION: Forward Primer HIV-S1

<400> SEQUENCE: 9

tggaagggct aatttggtc 19

-continued

---

```

<210> SEQ ID NO 10
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Reverse Primer HIV-S1

<400> SEQUENCE: 10

tattgctggt gatcctttcc                20

<210> SEQ ID NO 11
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Forward Primer HIV-S2

<400> SEQUENCE: 11

ccagcaatat tccagtgtag c                21

<210> SEQ ID NO 12
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Reverse Primer HIV-S2

<400> SEQUENCE: 12

tgaacaaaac ttggcaatga                20

<210> SEQ ID NO 13
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Forward Primer HIV-S3

<400> SEQUENCE: 13

catctcctat ggcaggaaga                20

<210> SEQ ID NO 14
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Reverse Primer HIV-S3

<400> SEQUENCE: 14

tgctagagat tttccacact ga            22

<210> SEQ ID NO 15
<211> LENGTH: 152261
<212> TYPE: DNA
<213> ORGANISM: Herpes simplex HSV-1 (NC_001806.1)

<400> SEQUENCE: 15

agcccggggc ccccggggc gcgcgcgcg gcaaaaaagg cggcgggcgg tccggggcggc    60
gtgcgcgcgcg cgcgcgggcg tggggggcgg gcccgcgga gcggggggag gagcgggggg    120
aggagcgggg ggaggagcgg ggggaggagc ggggggagga gcggggggag gagcgggggg    180
aggagcgggg ggaggagcgg ggggaggagc ggggggagga gcggggggag gagcgggggg    240
aggagcgggg ggaggagcgg ggggaggagc ggggggagga gcggggggag gagcgggggg    300

```

-continued

---

aggagcgggg ggaggagcgg ccagacgccc aaaacgggcc cccccaaaa cacaccccc	360
gggggtcgcg cgcggccctt taaagcgggt gcggcgggca gcccgggccc cccgcggcgg	420
agactagcga gttagacagg caagcactac tcgcctctgc acgcacatgc ttgcctgtca	480
aactctacca ccccggaacg ctctctgtct ccatggcccc cgcgcgcgc catcgcggcc	540
cccgcgccc cggccgccc gggcccaagg gcgcccctcc aaccgcacag tcccaggtaa	600
cctccacgcc caactcggaa cccgcggtea ggagcgcgcc cgcggcgcgc ccgcgcgcgc	660
ccccgcggg tgggcccccg cctctctgtt cgtctgtgct gcgccagtgg ctccacgttc	720
ccgagtccgc gtccgacgac gacgatgacg acgactggcc ggacagcccc ccgcccgagc	780
cgggcgcaga gggccggccc accgcgcgcg cccccggccc cggccccca ccgcccggcg	840
tgggcccggg gggcggggct gaccctccc acccccctc gcgcccttc cgccttcgcg	900
cgcgcctgc cctccgcctg cgcgtcacgc cggagcaact ggcggcctg cgcctgcgac	960
gcgcggggcg ggagggggcg ccggagcccc ccgcgacccc cgcgaccccc gcgacccccg	1020
cgacccccgc gacccccgcg cgggtgcgct tctcgcccc cgtccgggtg cgcacactgg	1080
tggctcgggc ctcggcgcgc cgcctggcgc gcccgggctc gtgggcccgc gageggggcg	1140
accgggctcg gttccggcgc cgggtggcgg aggcggaggc ggatccggg ccgtgcctgg	1200
ggcccagagg ccgtgcccgg gccctggccc gcggagccgg cccggcgaac tcgggtaac	1260
gttacaccgc aggcggcctg ggtcttcgcg ggagctcccc ggagctccgc accaagccgc	1320
tctccggaga gacgatggca ggagccgcgc atatatacgc ttggagccag cccgcctca	1380
cagggcgggc cgctcgggg gcgggactgg ccaatcggcg gccgccagcg cggcggggcc	1440
cggccaacca gcgtccgcgc agtcttcggg gcccgcccc ttgggcggga gttaccgccc	1500
aatgggcccg gcccccact tcccggatg gtaattaaaa acttgcaaga ggccttgctc	1560
cgtctcccgc tatggtaatt agaaactcat taatgggcgg ccccgccgc ccttcccgt	1620
tccggcaatt cccgcggccc ttaatgggca accccgggat tccccgcctc ccgcgcgcg	1680
cgtaaccact ccctgggggt tccgggttat gctaattgct tttttggcgg aacacacggc	1740
ccctcgcgca ttggcccgcg ggtcgcgcaa tgaaccgca ttggtcccct ggggttcggg	1800
gtatggtaat gagttctctc gggaaaggcg gaagccccg ggcaaccgac caggccaagc	1860
ccctgttcg tcggcgggag gggcatgcta atggggttct ttgggggaca ccgggttggg	1920
ccccaaaac gggggcggg ccgtgcacgc taatgatatt ctttgggggc gccgggttgg	1980
tccccgggga cggggcgcgc ccgcgggtgg cctgcctccc ctgggacgcg cggccattgg	2040
gggaatcgtc actgcccgc ctttggggag gggaaaggcg tggggataa gttagccctg	2100
gcccgacagt ctggtcgcac ttgcacctc gcaactcggag cgagacgcag cagccaggca	2160
gactcgggcc gcccccctc cgcatacaca cagaagcccc gcctacgttg cgaccccag	2220
ggaccctccg tccgcgacct tccagccgca tacgaccccc atggagcccc gcccccgagc	2280
gagtaccgcg cggcctgagg gccgccccca gcgcgaggtg aggggcggg cgcctatgct	2340
ggggcgccat attggggggc gccatattgg ggggcgccat gttgggggac ccccgaccct	2400
tacactggaa ccggccgcca tggttgggga ccccactca tacacgggag ccgggcgcca	2460
tgttggggcg ccatgttagg gggcgtggaa ccccgtagca ctatatatac agggaccggg	2520
ggcgcctatg taggggtg gcgaaccccc gaccctatat atacagggac cggggtcgcg	2580

-continued

---

ctgttggggg tcgccatgtg accccctgac tttatatata cagacccccca acacatacac	2640
atggcccctt tgactcagac gcagggcccc gggtcgccgt gggaccccct gactcataca	2700
cagagacacg cccccacaac aaacacacaa ggaccgggggt cgccgtgttg ggggcgtggt	2760
ccccactgac tcatacgcag gccccctta ctacacgcga tctagggggg tggggaggag	2820
ccgcccgcga tatttggggg acgccgtggg acccccgact ccggtgcgctc tggaggcgcg	2880
gagaagaggg aagaagaggg gtcgggatcc aaaggacgga ccagaccac ctttggttgc	2940
agacccttt ctccccctc ttccgaggcc agcagggggg caggactttg tgaggcgggg	3000
gggggagagg gggaactcgt ggggtgctgat tgacgcggga aatccccccc cattcttacc	3060
cgccccctt ttttccccct agcccccccc ggatgtctgg gtgtttccct gcgaccgaga	3120
cctgccggac agcagcgact ctgaggcgga gaccgaagtg ggggggcggg gggacgccga	3180
ccaccatgac gacgactccg cctccgaggg ggacagcacg gacacggaac tgttcgagac	3240
ggggctgctg gggccgcagg gcgtggatgg gggggcggtc tcggggggga gccccccccg	3300
cgaggaagac ccgcgcagtt cgcggggcgc cccccctcga gaggacgggg ggagcgacga	3360
gggacgctg tgcccgctgt gcacggatga gatcgcgccc cacctgcgct gcgacacct	3420
cccgtgcatg caccgcttct gcatccccgt catgaaaacc tggatgcaat tgcgcaaac	3480
ctgcccctg tgcaacgcca agctggtgta cctgatagtg ggcgtgacgc ccagcgggtc	3540
gttcagcacc atccccatcg tgaacgaccc ccagacccgc atggaggccg aggaggccgt	3600
cagggcgggc acggccgtgg actttatctg gacgggcaat cagcggttcg ccccgcggtg	3660
cctgaccctg ggggggcaca cggtagggc cctgtcgccc acccaccgag agcccaccac	3720
ggacgaggat gacgacgacc tggacgacgg tgaggcgggg ggcggcaagg accctggggg	3780
aggaggagga ggaggggggg ggagggagga ataggcgggc gggcgaggaa agggcgggcc	3840
ggggaggggg cgtaacctga tcgcgcccc cgttgtctct tgcagcagac tacgtaccgc	3900
ccgcccccg ccggacgccc cgcgcccc caccgagagg cgcgcccgcg cccccctgga	3960
cgggcggggc gtctcacgca gccccccagc cggccgcggc tcggacagcg cccccctcgg	4020
cgcccatcgg gccacacggc agcagtaaca ccaacaccac caccaacagc agcggcggcg	4080
gcggctccc ccagtcgca gcccgggcgc cgcggggggc gtctggccc tcgggggggg	4140
ttgggggttg ggttgggggt ttgaaagcgg aggcggggcg gccgaggggc cggacgggcc	4200
cccttgtaaa cagaccgccc ccccttgcaa acaacagaga ccccatagtg atcagcgact	4260
ccccccggc ctctccccac aggccccccg cggcgcccat gccaggetcc gccccccgccc	4320
ccggggcccc cgcgtcgcg gccgcgtcgg gacccgcgcg cccccgcgcg gccgtggccc	4380
cgtcgtgcg agcgcgcct cccgggcccc gccccgcgc cccggcccc ggggcgggagc	4440
cggcgcgcc ccccgggag gcgcgcctgt tgccccagtc gcaactcgtc ctggtcagg	4500
ccgcgaacca agaacagagt ctgtgccggg cgcgtgcgac ggtggcgcgc ggctcggggg	4560
ggccgggctt ggaggggtgg cacgggccc cccggggcgc cgcccccctc ggccgcgccc	4620
cgtccccctc cgcgcctct gtcgagcagg aggcggcggg gcgtccgagg aagaggcgcg	4680
ggtcgggcca ggaaaacccc tccccagc ccacgcgtcc cccccgcgc cgggacgggg	4740
ccaagagggc ggccagcgc cccccctcgc actcagggcc gggggggcgc ggccaggggtg	4800
ggcccgggac ccccctgacg tcctcggcgg cctccgcctc ttcctcctct gcctcttct	4860

-continued

---

cctcggcccc	gacccccg	ggggccgct	cttccgccc	cgggccg	tcctcctccg	4920
cttccgctc	ctcgggccc	gcccgcggt	cctggggag	gagacaag	gaaacctccc	4980
tcggcccc	cgctgttct	gggcgcggg	ggccgagg	gtgtgccc	aagacgcgcc	5040
acgcggag	ttcggggcc	gtccccg	gcggcctca	gcctacct	cccatctcgg	5100
gggtctctag	cgtggtcgcc	ctgtgcctt	acgtgaaca	gactatcac	ggggactgcc	5160
tgcccatcct	ggacatggag	acggggaaca	tcggggcgta	cgtggtcctg	gtggaccaga	5220
cgggaaacat	ggcagccc	ctgcgggccc	cggtcccc	ctggagccc	cgaccctgc	5280
tccccgag	cgccggta	cacgtgatc	cccccgagta	cccagcggc	cccgcgtcgg	5340
agtggaacag	cctctggatg	acccccgtg	ggaacatgct	gttcgaccag	ggcacccctag	5400
tgggcgcct	ggacttccc	agcctcgggt	ctcggcacc	gtggtccc	gagcaggggg	5460
cgctcgacc	ggcagggga	aaacaataag	ggcgccccc	cggtgttctg	gggagggggg	5520
ggtcgggccc	tgggtggtct	ctggccgccc	ccactacacc	agccaatccg	tgtcggggag	5580
gggaaaagt	aaagacacg	gcaccacaca	ccagcgggtc	ttttgtgtg	gcccataata	5640
aaaaaaaa	aggggatttt	tgtgtctgt	tgggaaata	aggtttactt	ttgtatcttt	5700
tccctgtctg	tgttgatgt	atcgcgggga	tgcgtgggag	tgggggtg	tgggagtggg	5760
ggtgcgtggg	agtgggggtg	cgtgggagtg	ggggtgcgtg	ggagtggggg	tgcgtgggag	5820
tgggggtg	tgggagtggg	ggtcgctggg	agtgggggtg	cgtgggagtg	ggggtgccat	5880
gttgggcagg	ctctgtgtt	aaccacagag	ccgcggccc	ggctgcctga	ccaccgatcc	5940
ccgaaagcat	cctgccactg	gcattggagcc	agaaccacag	tgggttgggt	gtgggtggtta	6000
agtttccg	agcgcctgcc	gcgccggact	gacctggcct	ctggcccca	caaagggcgg	6060
ggggggttaa	ctacactata	gggcaacaaa	ggatgggag	ggggcgggg	cgggacgggg	6120
cgcccaaa	ggggtcggcc	acaccacaga	cgtgggtgtt	gggggtggg	gcccaggggt	6180
ggggggggg	gagacagaaa	caggaacata	gttagaaaac	aagaatgcgg	tgcagccaga	6240
gaatcacagg	agacgagggg	atgggcgtgt	tggttaccaa	cccacacca	ggcatgctcg	6300
gtggtatgaa	ggagggggg	cggtgcttct	tagagaccgc	cgggggacgt	ggggtggtg	6360
tgcaaaggca	cgccgacccc	cgccggccag	tgggcccgtg	ctccatcccc	ccctcccccg	6420
acccttccca	cccccgctg	ccagagatca	ccccggtccc	ccggcacc	ccactcctcc	6480
gtatcctcgc	tttaggaaca	actttagggg	gggtacacac	gcgcccgtca	tttcttcca	6540
cacccccct	cccccgca	cccccccc	aggcagtaag	acccaagcat	agagagccag	6600
gcacaaaa	acagggcggg	tgggacacat	gccttcttgg	agtacgtggg	tcattggcgt	6660
ggggggttac	agcgcacccc	cccgaacccc	tggcgtctt	ccagccggcc	cttagataag	6720
ggggcagttg	gtggtcggac	gggtaagtaa	cagagtctga	ctaagggtg	gaggggggga	6780
aaagaacggg	ctggtgtgct	gtaaacagag	cccaccgccc	agtggcgtgg	ccgaccttag	6840
cctctggggc	gccccctg	gtttgggtcc	ccccctcta	ttggggagaa	gcaggtgtct	6900
aacctacctg	gaaacgcggc	gtctttgttg	aacgcacccc	gggcgcctc	gacgagtggg	6960
ataacggggg	aggaagggg	ggaggagggt	actgggggtg	aagaaggggg	gggggagaag	7020
cgagaacagg	aaaggggatg	gagcccgcca	gaacaccgag	gaaaaaaaa	ccacagccca	7080
tgcgccgggc	cgttgtgggg	ccccgggccc	gggccccctg	ggtccgcggg	ggccccgggc	7140

-continued

---

cgggccgcca cgggggcccg cegtggcgg taaccccag tgttcacetc aggccccggg	7200
ccgggaaccc ggaaaagcct ccggggggcc tttttcgcgt cgcgtgccgg cgagcgggtc	7260
cggacggggc ccggaccggc cgggtcgggg gccctcgtc ccgggcccga cgcggccttc	7320
gccccgtgag gggacagacg aacgaaacat tccggcagc gaaccgaaaa caccocagac	7380
gggttaaaga aacagaaacc gcaaccccc caacccccga aacggggaaa acgaaaaaac	7440
agaccagcgg ccggccggcg cttaggggga ggatgtcgc gacgcccctt ggccgccccg	7500
gctgcagggg ggccccgaga gcccgggcac ccggacgcgc ccgaaaagtc ttctgcacca	7560
cccgcgatcg gcacggcccg gcccccgtt ttataaaggc tgagatgacg cagcaaaaac	7620
aggccacagc accacgtggg taggtgatgt aattttatct tccctcgtctg cggcctaattg	7680
gatttcgggg cgcggtgcc cctgtctcag agcacttaac ggattgatct ctgcggggca	7740
cgcgcgcctc taatggaccg gcgcggggcg gggggccgga taccacacg gcgggggggg	7800
gggtgtcggc ggcgctctgc tggcccggcg ccacataaac aatgactctg ggcctttctg	7860
cctctgccgc ttgtgagtc gcgcgcggcg tctgcggtgt cggcgggcgc tcggcgggct	7920
ggggcgggcg ccgtgttcgg tctcggtagc cggccggcgg gtggactcgc ggggggcccg	7980
agggtggaag gcaggggggt gtaggatggg tatcaggact tccacttccc gtccttccat	8040
cccccgctcc cctcggttgt tccctgcctc ccccaacacc ccgcccgttt ccgttggggg	8100
tgattattgt gtcgggatcg tgcggggcgg gggctgcggc ggcagggggc ggggcgtggg	8160
cgggggtgct cgtcagtcga ccgggctcag tggggcgtg ggggtgggtg gagaaggcga	8220
ggagactggg gtgggggtgt cgggtgggtg ttgtttttg tggttgtttt tgtgtctgtt	8280
cccgtccccc gtcaccccc ccctcggctc cctcggctcc ccgctcggc gtgtttgtgt	8340
ttgtttatc cgacattggt ttatttaaat aaacacagc gttctgcgtg tctgttcttg	8400
cgtgtggctg ggggcttata tgtggggtcc cgggggcggg atggggttta gcggcggggg	8460
gcggcgccgc ggaacggggc ctggagataa cggccccggg ggaacggggg accggggctg	8520
ggatccccga ggtgggtggg tgggcggcgg tggccgggccc gggccgggccc gggccgggccc	8580
gggtggggcg ggtttgaaa aacgaggagg aggagagaaa ggcggggggg ggggagacgg	8640
gggaaaagca aggacacggc ccgggggggt ggagcgcggg ccggggccgct cgtaaagacc	8700
gcgacccggc cgcgggggag cgttgtcggc gtcggctctc cggccccgt cctcccttt	8760
tttgaccaac cagcgccccc cccccccctc accaccattc ctactaccac caccaccacc	8820
accaccgaca cctcccgcgc acccccggcc acatcccccc ccaacccgca ccaccagcac	8880
gggttggggg tagcagggga tcaaaggggg gcaaagcggc ggggcccgttc gggggggggg	8940
ggggggggcg gaaaccaag taggcccgcc catccgcggc ccctcccggc agccaagccc	9000
ccagcgtcgg gtgtcacggg gaaagagcag aggggagagg ggagaggggg ggagagggga	9060
gaggggggga gaggggagag ggggggagag gggagagggg gggagagggg agagggggg	9120
agaggggaga gggggggaga ggggagaggg ggggagaggg gagagggggg gagaggggag	9180
agggggggag aggggagagg gggggagagg gggtatataa accaacgaaa agcgcgggaa	9240
cggggatagc gggccttgtg ggcacgacgt cgtggttgtg ttaactggca aacacttggg	9300
gactgtaggt ttctgtgggt gccgacccta ggcgctatgg ggattttggg ttgggtcggg	9360
cttattgcgc ttgggggttt gtgtgtcggc gggggccttc cttcaaccga atatgttatt	9420

-continued

---

cggagtcggg tggctcgaga ggtgggggat atattaaagg tgccttgtgt gccgctcccg	9480
tctgacgac ttgattggcg ttacgagacc cctcgggcta taaactatgc tttgatagac	9540
ggtatatttt tgcgttatca ctgtcccgga ttggacacgg tcttggggga taggcatgcc	9600
cagaaggcat attgggttaa ccccttttta tttgtggcgg gttttttgga ggacttgagt	9660
taccccgct ttctgcca caccaggaa acagaaacgc gcttggccct ttataagag	9720
atagccagg cgtcggacag tcgcaagcag gccgccagcc acacacctgt gaaggctggg	9780
tgtgtgaact ttgactatc gcgcaccgc cgtctgttag ggcgacagga tttgggacct	9840
accaacggaa cgtctggacg gaccccggtt ctgccccgg acgatgaagc gggcctgcag	9900
ccgaagcccc tcaccacgccc gccgccatc atcgccactg cggacccac cccgcgacgg	9960
gacgcccga caaaaagcag acgcccagca cccactccc ggcgctcta acgatgcctc	10020
gacggaaacc cgtccgggtt cggggggcga accggccgccc tgcgctcgt cagggccggc	10080
ggcctcctc gcccccctag aggctggtcc cgtcgtgtg acgttttct cgtcccgcgc	10140
ccccgacct cccatggatt taacaaacgg gggggtgtcg cctcggcga cctcggcgc	10200
tctggactgg accacgtttc ggcgtgtgtt tctgatcgac gacgctggc ggcccctgat	10260
ggagcctgag ctggcgaacc ccttaaccgc ccacctcctg gccgaatata atcgtcgggtg	10320
ccagaccgaa gaggtgctgc cgcgcggga ggatgtgttt tcgtggactc gttattgcac	10380
ccccgacgag gtgcgcgttg ttatcatcgg ccaggacca taccaccacc cggccaggc	10440
gcacggactt gcgttttagc tgcgcgcgaa cgtgcgcct ccccgagtc ttcggaatgt	10500
cttggcggcc gtcaagaact gttatccga ggcacggatg agcggccacg gttgcctgga	10560
aaagtggcgc cgggacggcg tcctgttact aaacacgacc ctgaccgtca agcgggggc	10620
ggcggcgtcc cactctagaa tcggttggga ccgttctgt ggcggagtta tccgccggtt	10680
ggccgcgcgc cgcgccggcg tgggtgttat gctctggggc acacacgccc agaatgccat	10740
caggccggac cctcgggtcc attcgtcct caagttttcg caccgctgc cctctccaa	10800
ggtccgttc ggaacctgcc agcatttct cgtggcgaac cgatacctc agaccggtc	10860
gatttcaccc atcgactggt cggtttgaaa ggcacgacg tccggggttt ttgctgggtg	10920
gggcttttg gattttccga tgaataaaga cggttaatgg ttaaacctct ggtctcatac	10980
gggtcgtgta tgcggggcgt cgggggagag ggagtccct ctgcgctgc gattctagcc	11040
tcgtggggtt ggacgttcca caccgcaaac cagagtcgg ggatatcgc agatacgaact	11100
cccgcagatt ccattcgggg tgcgctgtg gcctcacctg accaaccttt acacgggggc	11160
ccggaacggg aggccacagc gccgtctttc tccccaacgc gcgcgatga cggcccgcgc	11220
tgtaccgacg gccctacgt gacgtttgat accctgttta tgggtcgtc gatcgacgaa	11280
ttaggcgtc gccagctcag ggacaccatc cgcaggacc tgcgggtgtc gctggccaag	11340
tttagcattg cgtgcaccaa gacctctcg ttttcgggaa acgccccgc ccaccacaga	11400
cgcggggcgt tccagcggcg cagcggggcg ccgcgcagca acaaaagcct ccagatgttt	11460
gtgttgtgca aacgcgcca cgcctcga gtgcgagagc agcttcgggt cgttattcag	11520
tcccgaagc cgcgcaagta ttacacgca tcttcggacg ggcggctctg cccgcctgc	11580
cccgtgttc tccagagtt cgtctcgtcc gagccaatgc gcctocaccg agataacgtc	11640
atgctggcct cgggggcgca gtaaccgccc ccccccatg ccacctcac tgcctcgc	11700

-continued

---

gcgtgtttga tgtaataaaa taacacataa atttggtctg ttgtttgttg tctttaatgg 11760  
 accgcccgca aggggggggg ggcatttcag tgtcgggtga cgagcgcgat ccggccggga 11820  
 tccataggacc ccaaaagtgt gtctgcgtat tccagggcgg ggctcagttg aatctccccg 11880  
 agcaacctcta ccagcaggtc cgcgggtgggc tggagaaact cggccgtccc ggggcaggcg 11940  
 gttgtcgggg gtggaggcgc ggcgccacc ccgtgtgccc cgctggcgt ctctctggg 12000  
 ggcgaccctg aaatggttgc agtgatgtaa atgggtgccc cggccagac cacggtcaaa 12060  
 atgcggcgcg tggcgctccg ggcgcttccg ccgcgcgagg agctgaccca ggagtcaaac 12120  
 ggatacgcgt acatatgggg gtcccccccg cgttcgagct tctggttgct gtccccgct 12180  
 ataaagcggg aggcacaaaa ttcgggcgca cagtcgataa tcaccaacag cccaatgggg 12240  
 gtgtgctgga taacaacgccc tccgcgcggc aggcggtcct ggcgctccc gccccgtacc 12300  
 atgatcgcgc ggtgcccgtc ctcaaaaaa tgcaccaact cgcggcgctc gggcagtgcg 12360  
 ctggtcagcg aggccctggc gtggcatagg ctatacgcga tggctgtctg tggattggac 12420  
 atctcgcggg gggtagtgag tccccgggc cgggttcggt ggaactgtaa ggggacggcg 12480  
 ggtaataga caatgaccac gttcggatcg cgcagagccc atagtatgtg ctactaatg 12540  
 acgtcatcgc gctcgtggcg ctccccgagc ggatttaagt tcatgcgaag gaattcggag 12600  
 gaggtggtgc gggacatggc cacgtacgcg ctggtgaggc gcagggtgcc gggcgtaaag 12660  
 cagatggcga ccttgtccag gctaaggccc tgggagcgcg tgatggctat ggcaagcttg 12720  
 gagctgatgc cgtagtcggc gtttatggcc atggccagct ccgtagagtc aatggactcg 12780  
 acaaaactgc tgatgttggg gttgacgacg gacatgaagc cgtgttggc ccgcaagacc 12840  
 acgtaaggca ggggggcctc tccagtaac tggccacgt tggccgtcgc gtgccgcctc 12900  
 cgcagctcgt ccgcaaagc aaacaccgt gcgtacgtgt atcccatgag cgtataattg 12960  
 tccgtctgca gggcgacgga catcagcccc ccgcgcggcg agccggtcag catctcgcag 13020  
 ccccggaaga taactgtgtc cacgtacgtg ctaaaggggg cgacttcaa tgcctccccg 13080  
 aagagctctt ggaggattcc gaatctcccc aggaaggccc gcttcagcag cgcaactgg 13140  
 gtgtgaaagg cggcggtggt ctccggttcc ccgggggtgt agtggcagta aaacacgtcg 13200  
 agctgttgtt cgtccagccc cgcgaaaata acgtcagagt cgtcgtcggg aaaatcgtcc 13260  
 gggccccctg cccgcggccc cagttgctta aaatcaaacg cacgctcgcg gggggcgcct 13320  
 gcgtcgccca ttaccgacgc ctgctcggc acccccgaag atttggggcg cagagacaga 13380  
 atctccgcg ttagttctcc catcggggcg taggcgaggg tcctctgggt cgcatecagg 13440  
 cccgggctgc gcagaaagt gtaaaaggag ataagcccgc taaatatgag ccgcgacagg 13500  
 aacctgtagg caaactccac cgaagtctcc ccctgagtct ttacaaagct gtcgtcacgc 13560  
 aacactgct cgaaggccc gaacgtccca ctaaacccaa aaaccagttt tcgaggcgc 13620  
 gcggtcaccg cgatctggct gttgaggacg taagtacgt cgttgccggc caccaccagc 13680  
 tgctgtttgc tgtgcacctc gcagcgcagc tgcgccgct cctggctctg gctctcgcag 13740  
 tagttggtga tgcggctggc gttggccgtg agccactttt caatcgtcag gccgggctgg 13800  
 tgtgtcagcc gtcggtatct gtcaaaactc ttgaccgaca cgaacgtaag cacggggagg 13860  
 gtgaacacga cgaactcccc ctcaacgggt accttcaggt aggcgtggag cttggccatg 13920  
 tacgcctca cctctttgtg ggaggagaac agccgcgtcc agccggggag gttggcgggg 13980

-continued

---

ttggtgatgt agttttccgg gacgacgaag cgatccacga actgcatgtg ctctcgggtg 14040  
 atgggcaggc cgtactccag caccttcatg aggttaccga actcgtgctc gacgcaccgt 14100  
 ttgttgtaa taaaaatggc ccagctatac gagaggcggg cgtactcgcg cagcgtgcgg 14160  
 ttgcagatga ggtacgtgag cacgttctcg ctctggcggg cggaaacaccg cagtttctgg 14220  
 tgctcgaagg tcgactccag ggacgcctgc tcgctcggcg agcccacaca caccaacacg 14280  
 ggccgcaggc gggccgcgta ctgggggggtg tggtagaggg cgttaatcat ccaccagcaa 14340  
 tacaccacgg ccgtgaggag gtgacgcccc aggagcccgg cctcgtcgat gacgatcacg 14400  
 ttgtgcggg taaaggccgg cagcgcctcg tgggtggcgg gggccaaaccg cgtcagggcg 14460  
 ccctcggcca accccagggt ccgttccagg gcggccaggg cgcgaaactc gttccgcaac 14520  
 tcctcgcctc cggaggcggc cagggcgcgc ttcgtgaggt ccaaaatcac ctcccagtag 14580  
 tacgtcagat ctcgtcgtg caggtcctcc agcggaggcg ggttgctggt caggggtgac 14640  
 gggtagctgc ccagttgggc ctggacgtga ttcccgcgaa acccaaattc atgaaagatg 14700  
 gtgttgatgg gtcggctgag aaaggcgcct gagagtctgg cgtacatggt ttgggcccga 14760  
 atgcgcgtgg cgcctgcac cacacagtcc aagacctcgt tgattgtctg cacgcacgtg 14820  
 ctctttccgg agccagcgtt gccggtgata agatacaccg cgaacggaaa ctccctgagg 14880  
 ggcaggcctg cgggggactc taaggccgcc acgtcccgga accactgcag atggggcaact 14940  
 tgctcctcgt cgagctggtt ttgagagagc tctcggatgc gcttaaggat tggctgcacc 15000  
 ccgtgcatag acgtaaaaat taaaaaggcc tcggccctcc ctggaacggc tggctcggcc 15060  
 ccgggttget gaaggtgcgg cgggcccgggt ttctgtccgt ctagtggcg cccccgcg 15120  
 gccgcgcca tgaccgcacc acgtcgcgg gcccccacta cgcgtgcgcg gggggacacg 15180  
 gaagcgtgt gctccccga ggacggctgg gtaaaggctc accccagccc cgttacgatg 15240  
 ctgttccgag agattctcca cgggcagctg gggatataccg agggccaggg ggtgtacaac 15300  
 gtcgtccggt ccagcgaggc gaccaccggc cagctgcagg cggcgatctt tcacgcgctc 15360  
 ctcaacgcca ccacttaccg ggacctcgag gcggactggc tcggccacgt ggcggcccg 15420  
 ggtctgcagc cccaaagggt ggttcgcccg tacaggaacg cccgggaggc ggatatcgcc 15480  
 ggggtggcgg agcgggtggt cgacacgtgg cggaaacacg ttaggacgac gctgctggac 15540  
 tttgcccacg ggttggtcgc ctgctttgcg cggggcggcc cgagcggccc gtcaggcttc 15600  
 cccaaatata tcgactggct gacgtgcctg gggtcggctc ccatattacg caagcgacaa 15660  
 gaaggggggtg tgacgcaggg tctgaggcg tttctcaagc agcaccgct gaccgcccag 15720  
 ctggccacgg tcgaggaggc cgcggagcgc gccggccccg ggtttttga gctggcgctg 15780  
 gccttcgact ccacgcgctg ggcggactac gaccgctgt atatctacta caaccaccgc 15840  
 cggggcgact ggtcgtgag agaccccacg agcgggcagc gcggagaatg tctggtgctg 15900  
 tggccccct tgtggaccgg ggacgctcgt gtcttcgatt cgcccgcca cgggctggtt 15960  
 cccgagatcg tcgctgtca ctccctccgg gaacacgcgc acgtctgccc gctgcgcaat 16020  
 acccgcctcg tcaagtgct gctggggcgc aagagcgaca gcgagcgcgg ggtggccggt 16080  
 gccgcgcccg tcgttaacaa ggtgttgggg gaggaacgac agaccaaggc cgggtcggcc 16140  
 gcctcgcgcc tcgtgagggt tatcatcaac atgaagggca tgcgccacgt aggcgacatt 16200  
 aacgacaccg tcgcttccca cctcgacgag gccggggggc acctgataga cgcgccggcc 16260

---

-continued

---

gtcgacggta ccctccctgg attcggaag ggcggaaca gcccggggc tgcgggccag 16320  
gaccaggggg ggcgggccc gcagcttcgc caggccttc gcacggccgt ggtaacaac 16380  
atcaacggcg tgttgagggt ctatataaat aacctgttg gaaccatcga gcgcctgcgc 16440  
gagaccaacg cgggcttggc gacccaattg caggagcgcg acccgagct ccggcgcga 16500  
acagcggggg ccctggagcg ccagcagcgc gcggccgacc tggcggccga gtcctgacc 16560  
ggtaggatcg gcagcccgc tgcggggggc gacctgctcc gggccgacta tgacattatc 16620  
gacgtcagca agtccatgga gcagcagcgc tacgtcgcca acagcttca gacccgtac 16680  
atcccttcgt acgcccagga cctggagcgc ctgtcgcgc tctgggagca cgagctgggtg 16740  
cgctgtttaa aaattctgtg tcaccgcaac aaccagggcc aagagacgtc gatctcgtac 16800  
tccagcgggg cgatgcggc attcgtcgc cctactttg agtcagtctc tcgggcccc 16860  
cgggtaggcg cgcgccatcg gggctccgat gtcacctcgg gggaggagga gttatgggat 16920  
gcggtgttta agaaaaccg cctgcaaac tacctgacag acatcgggc cctgttcgtc 16980  
gcgacgtcc agcagcagc gctgccccg cccccctcc cggtcggcgc cgatttcgg 17040  
ccccggcgt ccccgcgggg ccggtccaga tcgcggtcgc ccggaagaac tgcgcgaggc 17100  
gcgcggacc agggcggggg catcgggcac cgggatggcc gccgcagcgc ccgacgatga 17160  
ggggtcggcc gccaccatcc tcaagcaggc catcgcggg gaccgcagcc tggtcgaggc 17220  
ggccgaggcg attagccagc agacgtcgt ccgcctggcc tgcgaggtgc gccagtcgg 17280  
cgaccgccag ccgcggttta ccgccaccag catcgcgcgc gtcgacgtcg cgcctgggtg 17340  
ccggttcggg ttcgttctgg acgggagtc ccgagcgcg tatgtgacgt ccgaggatta 17400  
cttaagcgc tgcgcggcc agtcagttta tcgcggttc gcggtggcgg tctgacggc 17460  
caacgaggac cacgtgcaca gcctggcgt gccccctc gttctgctgc accggttctc 17520  
cctgttaac cccagggacc tccctggactt tgagcttgcc tgtctgctga tgtacctgga 17580  
gaactgcccc cgaagccacg ccaccccgc gaccttgcc aaggttctgg cgtggctcgg 17640  
ggtcgcgggt cgcgcacgt cccattcga acgcgttcgc tgcctttcc tccgcagttg 17700  
ccactgggtc ctaaacacac tcactgtcat ggtgtacgta aaacgctcg acgacgagtt 17760  
cgtcctgcc cactggtaca tggcccggta cctgctggcc aacaaccgc cccccgttct 17820  
ctcggccctg tctgtgcca cccgcagcag ctctcattc cggctgcccg gcccgcccc 17880  
ccgctccgac tgcgtggcct ataaccgc cgggatcatg gggagctgct gggcgtcgg 17940  
ggaggtgcgc gcgcctctgg tctattgggt gcttccggag accccaaaac gacagcgtc 18000  
gtcgtgttt tatcagttt gttgaattt aggaataaa cccggtttg tttctgtggc 18060  
ctcccagcgc atgcgcgtgt ccttactcgc tcttggtggg tgggtggctg tgtatggct 18120  
cccatctgtg cggggagggg ggcaagtccg cacgtattcg gacagactca agcacacag 18180  
ggggagcgt cttgtctcag ggcaatgtt ttattggtca aactcaggca aacagaaacg 18240  
acatcttgtc gtcгааggga tacacaaact tccccctc gccccatact cccgccagca 18300  
ccccggtaaa caccaactca atctcgcga ggatttcgc caggtgatga gcgcagtcca 18360  
cgggggggag cacaaggggc cgcgggtata gatcgcggg gacgcgacc gactccccgc 18420  
ctccgggaca gacacgcag acgcgcggc agtagtgctc tgcgtccagc aaggcggcgc 18480  
cgcggaaggc agtggggggc aaggggtcgc tggcctcaa gggggacacc cgaacgctcc 18540

-continued

---

```

ag tactccgc gtccaaccgt ttattaaacg cgtccaagat aaggcggtcg caggcgtcct 18600
ccataaggcc ccggggccgtg agtgcgctct cctccggcac gcatgccgtt gtcaggccca 18660
ggaccctcgc cagcgtgtcg cgtacgaccc ctgccccgt ggtgtaecgc ggccccgcga 18720
gaggaaatcc cccaagatgg tcagtgttgt cgcgggagtt ccagaaccac actcccgcct 18780
ggctccaggc gactgcgtgg gtgtagacgc cctcgagggc caggcacagt gggtgccgca 18840
gccggacggc gttggcccta agcacggctc ccacggccgt ctcgatggcc cgcggggcgt 18900
cctcgatcac cccggaagcc gcatccgcgt cttgggggtc cacgttaaag acaccccaga 18960
acgcaccccc atcgcccccg cagaccgcga acttcaccga gctggccgtc tcctcgatct 19020
gcaggcagac gggggccatt accccaccca ggagctgccg cagcgcaggg caggcgttgc 19080
acgtgtccgg gaccaggcgc tccaagacgg ccccggccca gggctctgag ggagcggcca 19140
ccaccagcgc gtcagctctt gctaggcccg tccggccgtg ggggtccgcc agcccgcctc 19200
ccccgaggtc ggcaggggcc gccaggagct gggcgcgaag tccggggaag caaaaccgcg 19260
ccgtccagac gggcccgcag gcccgggggc ggtctaacag ttggatgatt ttagtggcgg 19320
gatgccaccg cgcaccgcgc tcccgcaccg cgggcaggag gcatccggct gcccccagg 19380
ccacgcccgg ccaggctcgc ggggggagga cgacctggc ccccaccgcg gcccaggccc 19440
ccaggagcgc ggcgtaagcg gccgcggccc cgcgcaccag gtcccgtgcc gactcggccg 19500
tgcccgccac ggtgaacgtg ggccaacccc gaaaccccag gacggcaaag tacgggacgg 19560
gtcccccccg gacctcaaac tcgggcccga gaaaggcaaa gacggggggc agggcccgcg 19620
gggcggcgtg gaccgtggta tgccactgcc ggaaaagggc gacgagcgc gccgcggaga 19680
acttctcgcc ggcgcttaca aagtagtctg aatcgcgggg cagcagcacc cgtgccgtga 19740
ctcgttcggg gtgcccgcgt ggcgcaggc ccacctgca cacctgacc aggtcccga 19800
acgcgccctc cttcttgatc ggcggaaaac caagagtctg gtattcgcgc gaaatagcg 19860
cggttccggg ggtgatgta acggtcagcg aagcggcggga cgcgcactgg ggggtgtcgc 19920
gaatggccgc caggcgcgcc cacgccagc gcgcgtcggg atgctcggca acgcgcgccg 19980
ccagggccat agggctgatg tcaatgttg cctccgcgac caggagagcg gcgcgagggg 20040
cggcggggcg gccccacgac gctctctcaa ctttcaccac cagtcccgtg cgtgggtccg 20100
agccgatacg cagcggggcg aacagggcca ccggcccgt ctggcgctcc agggcccga 20160
ggacgcacgc gtacagcgc ccgccacagag tcgggttctc caggggctcc agcggggagg 20220
cggccggcgt cgtcgcggcg cgggcggccg ccacgacggc ctggacggag acgtccgcgg 20280
agccgtagaa atcccgcagc tccgtcgcgg tgacggagac ctccgcaaag cgcgcgcgac 20340
cctcccctgc ggcgttgcca catacaaaat acaccagggc gtggaagtac tcgcgagcgc 20400
gggggggcag ccataccgcg taaagggtaa tggcgtgac gctctcctcc accacaacga 20460
tatctcgggt gtccatcgca cggcccctaa ggatcacggg cggctctgtg gtccatgct 20520
gccgtgcctg gccgggcccg gtgggtcgcg gaaacgggtg acgggggggg gccggttttt 20580
ggggttgggg tgggggtggg aaacggcccc ggtccggggg ccaacttggc cctcgggtgc 20640
gttcggcaa cagcgcgcc ggteccggga cgaccacgta ccgaacgagt gcggteccga 20700
gacttatagg gtgctaaagt tcaccgcccc ctgcatcatg ggccaggcct cggtggggag 20760
ctccgacagc gccgcctcca ggatgatgct acggttgggg ttggcgtgg atgagtgcgt 20820

```

-continued

---

gcgcaaacag cgccccacg cgggcacgcg tagcttgaag cgcgcgcccg caaactcccg 20880  
 cttgtgggccc ataagcaggg cgtacagctg cctgtgggtc cggcagggcg tgtggtcgat 20940  
 gtgggtgggcg tccaacaacc ccaagattgt ctgtttggtg aggtttttaa cgcgccccgc 21000  
 cccgggaaac gtctgcgtgc ttttgccat ctgcacgcca aacagttcgc cccagattat 21060  
 cttgaacagc gccaccgcgt ggtccgtctc gctaaccggac cgcgcgggg gacagccgct 21120  
 tagggcgctg gcgacgcgtg tgacggcttc ctccgagagc agaagtccgt cggttacgtt 21180  
 acagtggccc agttcgaaca ccagctgcat gtagcggctg tagtgggggg tcagtaggtc 21240  
 cagcacgtca tcggggccga aggtcctccc agatcccccg gccgcgagt cccaatgcag 21300  
 cgcgcgccc atgggtcctgc acaggcaca cagctcccag acggggggtta cgttcagggt 21360  
 ggggggcagg gccacgagct ccagctctcc ggtgacgttg atcgtgggga tgacgcccgt 21420  
 ggcgtagtgg tcatagatcc gccgaaatat ggcgctgctg cgggtggcca tgggaacgcg 21480  
 gagacaggcc tccagcaacg ccaggtaaat aaaccgcgtg cgtcccatca ggctgttgag 21540  
 gttcgcgatg agcgcgacaa tttccgcccg cgcgacatcg gaccggaggt attttccgac 21600  
 gaaaagacc acctcctccg tctcggcggc ctgggcccgc agcgcgcct cgggatcccg 21660  
 gcaccgcagc tcccgtagat cgcgctgggc cctgagggcg tcgaaatgta cgcgccgcaa 21720  
 aaacagacag aagtcccttg gggtcagggt atcgtcgtgt cccagaaagc gcacgcgtat 21780  
 gcagtttagg gtcagcagca tgtgaaggat gttaaaggctg tccagagagc acgccagcgt 21840  
 gcctctctca aagtagtgtt tgtaacggaa tttgtgtgag atgcgcgacc cccgcccag 21900  
 cgacgtgctg catgccgacg cgtcacagcg ccccttgaac cggcgacaca gcaggtttgt 21960  
 gacctgggag aactgcgcgg gccactggcc gcaggaactg accacgtgat taaggagcat 22020  
 gggcgtaaag acgggctccg agcgcgcccc ggagccgtcc atgtaaatca gtagctcccc 22080  
 cttcgggagg gtcgcgaccc gtcccaggga ctggtacacg gacaccatgt cgggtccgta 22140  
 gttcatgggt ttcacgtagg cgaacatgcc atcaaagtgc aggggatcga agctgaggcc 22200  
 cacggttacg accgctcgtt atataaacac gcggtattgg ccccacgtgg tcacgtcccc 22260  
 gaggggggtg agcagatgaa gcaacagcac gcggtccgta aactgacggc agaaccgggc 22320  
 cacgatctcc gcgaaggaga ccgtcgacga aaaaatgcag atgttatcgc ccccgccaag 22380  
 cgcgccttcc agtcccccaa agaactgggc ccccgggccc tccggagagg cgtccggaga 22440  
 cgggcccgtc ggcggcccgg cggggcgacg ggcagcctgc aggagctcgg tcccagacg 22500  
 cgggagaaac aggcaccggc gcgcccgaaa cccgggcatg gcgtactcgc cgaccaccac 22560  
 atgcacgttt ttttcgcccc ggagaccgca caggaagtcc accaactcgc cgttggcgggt 22620  
 tgcgtccatg gcgatgatcc gaggacagat gcgcagcagg cgtagcatta acgcatccac 22680  
 gcgcccagat tctgcacatg ttggcgaata gagctggccc agcgtcgaca taacctcgtc 22740  
 cagaacgagg acgtcgtagt tgttcagaag gttggggccc acgcgatgaa ggctttccac 22800  
 ctggacgata agtcggtgga aggggcccgtc gttcataatg taattggtgg atgagaagta 22860  
 ggtgacaaag tcgaccaggc ctgactcagc gaaccgcgtc gctagggtct gggtaaaact 22920  
 ccgacgacag gagacgacga gcacactcgt gtccggagag tggatcgtt cccgcagcca 22980  
 cgggatcagc gcggtagttt ttcccagacc cattggcgcg cggaccacag tcacgcacct 23040  
 ggccgtcggg gcgctcgcgt tggggaagggt gacgggtccc tgctgctgcc gctcgtcgt 23100

-continued

---

tgttttcggg tgaaccggg gcacccattc ggccaaatcc ccccgtaca acatccgcgc	23160
tagcgatacg ctcgacgtgt actgttcgca ctcgctgtcc ccaatgggac gcccgcccc	23220
cagaggatct cccgactccg gcgccccac gaaaggcatg accggggcgc ggacggcgtg	23280
gtgggtctgg tgtgtgcagg tggcgacgtt tgtggtctct gcggtctgcg tcacggggt	23340
cctcgtcctg gcctctgtgt tccgggcacg gttccctgc ttttacgcca cggcgagctc	23400
ttatgcggg gtgaactcca eggccgaggt gcgcggggt gtacgcgtgc cctcaggtt	23460
ggacacgcag agccttgtgg gcacttatgt aatcacggcc gtgttgtgt tggcctggc	23520
cgtgtatgcc gtggteggcg ccgtgacctc ccgctacgac cgcgccctgg acgcgggcgc	23580
ccgtctggct gcggccccga tggccatgcc gcacgccacg ctgatcgccg gaaacgtctg	23640
ctcttggtg ctgcagatca ccgtcctggt gctggcccat cgcacagcc agctggccca	23700
cctggtttac gtectgcact ttgcgtgtct ggtgtatctt gcggcccatt tttgaccag	23760
gggggtcctg agcgggacgt atctgcgtca ggtgcacggc ctgatggagc tggccccgac	23820
ccatcatcgc gtcgtcggcc eggctcgcgc cgtgctgaca aacgcctgc tgttgggct	23880
cttctgtgc acggccgacg ccgcggtatc cctgaatacc atcgccgcgt tcaacttaa	23940
tttttcggcc ccgggcatgc tcaatcgcct gaccgtgctg ttccgcatc tcgtcgtatc	24000
gctgttgtg gtggteggag ggggtgtgtg tcaactcgtg cgcgtgttg tgggccccca	24060
cctgggggcc gtggccgcca cgggcatcgt ccgacctggc tgcgagcact attacacaa	24120
cggctactac gttgtggaga cgcagtggcc gggggctcag acgggagtc gcgtcgcct	24180
cgcctggtc gccgcctttg ccctcggcat ggccgtgctc cgtgcaccc gcgcctatct	24240
gtatcacagg cggcaccaca ccaaatctt tatgcgcatg cgcgacacgc gacaccgcgc	24300
acattccgcc ctcaagcgcg tacgcagttc catgcgcgga tcgcgagacg gccgccacag	24360
gcccgcaccc ggccagcccg ccgggattcc cgaatatgcg gaagaccct acgcatctc	24420
atcggcggc cagctcgacc ggtaaccgaga ttccgacggg gagccgattt acgacgaggt	24480
ggcggacgac caaacgcagc tattgtacgc caagatacaa caccgcggc acctgcccga	24540
cgacgatccc atctatgaca ccgttggggg gtacgacccc gagccgcgcg aggacccct	24600
gtacagcacc gtccgcctgt ggtagctgtt tggttccgtt ttaataaac gtttgtgtt	24660
aaccgcaccg tgggtgatgt ctggtgtgtg gcgtccgac ccgttactat caccgtcccc	24720
ccccccccct caaccgcgc gattgtgggt ttttataaaa cgacacgcgt gcgaccgat	24780
acagaacatt gttttggtt ttattcgcta tcggacatgg ggggtgaaa ctgggtggcg	24840
gggcaggcgc ctccgggggt ccgcccgtga gtgtggcgcg aggggggggc cgatgaacgc	24900
aggcgtgtc tccccggggc ccgcgtaacc ccgcatat ccgggggcac gtagaaatta	24960
ccttctctt cggactcgat atccacgacg taaaagtcgt gggcggtcag cgagacgacc	25020
tcccctcgt ccgtgatgag gacgttgtt ccgacgcgc agggccgggc cccggagAAC	25080
gagaggccca tagctcggg agcgtgtcgt cgaatgccag gcgctgctt cgtggatgg	25140
ccttatagat ctccgatcgt atgcggacgg ggtaatgat cagggcgatc ggaacggcct	25200
ggttcgggag aatggacgcc ttgtgggtc ctgcggcccc gagagcccc gcgcccct	25260
ccaggcggaa cgttacgcc tctcgcgcgc tgggtcgggt cctgcccgata aacgtcacca	25320
gatgcgggtg gggggggcag tcggggaagt ggctgtcag cacgtagccc tgcaccaaga	25380

-continued

---

tctgcttaaa gttcgggtga cgggggttcg cgaagacggg ctccggcgcg accagatccc	25440
cggagctcca ggacacgggg gagatggtgt ggcgtccgag gtcggggcg ccaaacagaa	25500
gcacctccga gacaacgccg ctatttaact ccaccaaggc cccgatccgc gcgagacc	25560
gccttttttc gcccgaggcg tgggctctg accaggcctg gtcttgctg acgagagcct	25620
cctccgggcc ggggacgcgc ccgggcgca agtatcgac gctgggcttc gggatcgacc	25680
ggataaatgc ccggaacgcc tccggggacc ggtgtgcat caagtcctcg tacggggagg	25740
ccgtggggtc gctggggtcc atggggtcga aagcgtactt ggcccggcat ttgacctgt	25800
aaaaggccag gggggtcttg gggactggg ccaggtagcc gtgaatgtcc cgaggacaga	25860
cgagaatgc cagggacgcc ccgacctcc ccgtgtgacc gtccatgagg accccacacg	25920
tatgcaegt ctcttcggcg aggtcgtgg gttcgtgaa gataaagcgc cgcgtgtcgg	25980
cgccggcctc gccgcgctg tccgcgcggc ccacgcagta gcgaaacagc aggcctcggg	26040
ccgtcggctc gttcaccgc ccgaacatca ccgccgaaga ctgtacatcc ggcgcaggc	26100
tggcgttggt cttcagccc tggggcgaga aacacggacc ctgggggccc cagcggaggg	26160
tggatgctgt cgtgagccc gcgccgagca gggccatag ctggcagtc gcctggttt	26220
gcgtggcgc ctctgtaaac cccatgagg gccggggcgc cacggcgtcc gcggcgccg	26280
ggggcccgcg gcgcgtcagg ccgataggt gccgaccgag tccgcggtcc accatacccg	26340
cctcctcgag gaccacggcc agggaaacaca gataatccag gcgggcccag aggggaccga	26400
tggccagagg ggcgcggcag ccgcgcagca acccgcgag gtggcgctcg aacgtctcgg	26460
ctagtatatg ggagggcagc gcgttggga tcaccgacgc cgaccacata gactcaaggt	26520
ccggggagtc gggatcggg tccgggtcgc gggcgtgggt gccccagga gatagcggaa	26580
tgtctggggt cggagccct gaggcgtcag aaagtgcgg cgacgcggcc cggggcttt	26640
cgtctcgggt gtcggtggcg tgctgatcac gtggggggtt aacgggcgaa tgggagctcg	26700
ggtccacagc tgatgtcgtc tgggggtggg ggggcagggg acggaagggt gttgtcagcg	26760
gaagactggt agggcgggg gccttgggg ggctgtcgg gccacaggg gtgtcctcgg	26820
ccagggccca gggacgctta gtcacggtgc gtcccggcg acatgctggg cctaccgtgg	26880
actccatttc cgagacgacg tggggggagc ggtggttag cgcgcccg ggtgaacgct	26940
gattctcacg acagcgcgtg ccgcgcgac gggttggtgt gacacaggcg ggacaccagc	27000
accaggagag gcttaagctc gggaggcagc gccaccgac acagtatcg cttgtgtgtg	27060
tgctggtaat ttataccgc atccgtaaac gcgcgccgaa tcttgggatt gcggaggtgg	27120
cgccggatgc cctctgggac gtcatacgc aggcgtggg tggtggtctc ggccgagttg	27180
acaaacaggg ctgggtgcag caccgagcga taggcgagca gggccagggc gaagtccgc	27240
gacagctggt tgttaaaata ctggtaaccg ggaaccggg tcacgggtac gccaggtc	27300
ggggcgagct acacgctaac caccaactcc agcagcgtct ggcccagggc gtacaggtca	27360
accgctaacc cgacgtcgtg cttcaggcgg tggttggtaa attcggccc ttcgttggtta	27420
aggtatttca ccaacagctc cgggggctgg ttataccgt gaccaccag ggtgtgaaag	27480
ttgctgtggt ttagggcggg gggcatgcca aacatccgg gggacttag gtcggctcc	27540
tggaggcaaa actgccccg ggcgatcgtg gagttggagt tgagggtgac gaggtaaaag	27600
tcggcgagga cggcccgcg gagcggagac gcgtccgacc gcagcatgac gaggatggtg	27660

-continued

---

gcgcaactga tatccagggtg gctgatcccc caggtggtgt ttaaaaacac aacggcgcgg 27720  
gccagctccg tgaagcactg gtggagggcc gtcgagaccg aggggtttgt tgtgcgccag 27780  
gacgccagtt ggccgata cttaccaggg tccatgtcgt acgcggggaa cactatctgt 27840  
cgttgttca gcgagaaccc gaggggcgcg atgaagccgc ggatgttggt ggtgcgcccg 27900  
gcgcgtagaa cgcactcccc gaccaacagg gtcgcatga gctcaacggc aaaccactcc 27960  
ttttccttta tggcttaac ggcaagctta tgttcgcgaa tcagttggac gtcaccgtat 28020  
ccccagacc ccccaagct tcgggccccg gggatctcga gggctgtgta gtgtagggcg 28080  
gggttgatgg cgaacacggg gctgcatagc ttgcgatgc gcgtgagggg gaggatgtgc 28140  
gagggggacg aggggggtgc ggttaacgcc gcctgggatc tgcgcagggg cgggcccgtc 28200  
agtttgccg ccgtaccggg cgtctcggg gacgcgcggc gatgagacga gcggctcatt 28260  
cgccatcggg atagtcccg gcgaagccgc tcgcgaggc cggatcgggt gcgggacccg 28320  
tgaggagagc gggagacggc ggcgtcctgg agagaggggc cgtggggcg cccggaggcc 28380  
ccgtgggggt tggagtgtac taggatgccc agccaatcct tgaaggaccg ttggcgtgca 28440  
ccttgggggc tgaggttagc tgccacatga ccagcaggtc gctgtctgcg ggactcatcc 28500  
atccttcggc caggtcgcgc tctcccaca gagaagcgtt ggtcgcctgc tcctcgagtt 28560  
gctcctcctg gtccgcaaga cgatcgtcca cggcgtccag gcgctcacca agcgcgggat 28620  
cgaggtaccg tcggtgtgcg gttagaaggt cacgacgcgc cgttctgctc tccacgcgaa 28680  
ttttaacaca ggtcgcgcgc tgctgcatca tctctaagcg cgcgogggac tttagccgcg 28740  
cctccaattc caagtgggccc gcctttgcag ccataaagcc gccaaacaaac cgaggatcct 28800  
gggtgctgac gccctcccgg tgcagctgca gggctctggtc cttgtaaatc tcggctcggg 28860  
ggtgcgtctc ggcacagcgt cggcgcaggg ccgctggggc ggcactctcg tccattccgc 28920  
caccctcgcg gcgaccggg ggggtgctctg atagtctcgc gtgcccaagg cccgtgatcg 28980  
gggtacttcg ccgcccgcag ccgcccaccg gtgtgcgcga tgtttggtca gcagctggcg 29040  
tccgacgtcc agcagtagct gagagcctc gagaaacaga ggcaactaa ggtgggcgcg 29100  
gacgagggct cggcgggctt caccatgggc ggcgatgccc tacgagtgcc ctttttagat 29160  
ttcgcgaccg cgaccccaaa gcgcccaccg accgtggtcc ctggcgtcgg gacgctccac 29220  
gactgctgcg agcactcgcg gctcttctcg gccgtggcgc ggcggctgct gtttaatagc 29280  
ctggtgcccg cgcaactaaa gggcgtgat ttcgggggcg accacacggc caagctggaa 29340  
ttcctggccc ccgagttggt acgggcgggtg gcgcgactgc ggtttaagga gtgcccgcgg 29400  
gcgagcgtgg tgccctcagc taacgcctac tatagcgttc tgaatacgtt tcaggccctc 29460  
caccgctccg aagcctttcg ccagctgggt cactttgtgc gggactttgc ccagctgctc 29520  
aaaaactcct tccgggcctc cagcctcagc gagaccaagg gccccccaa aaaacgggccc 29580  
aaggtggacg tggccacca cggccggacg tacggcacgc tggagctggt ccaaaaaatg 29640  
atccttatgc acgccacctt ctttctggcc gccgtgctcc tcggggacca cgcggagcag 29700  
gtcaaacagt tctgctgctc cgtgtttgag atccccctgt ttagcgcgcg gccctgccc 29760  
cacttcccgc agcgcgccac cgtgtttctc gtcccccgcc gccacggcaa gaectggttt 29820  
ctggtgcccc tcactcgcgt gtgctggccc tcttttcggg ggatcaagat ccgctacacg 29880  
gcgcacatcc gcaaggcgcg cgagccgggt tttgaggaga tcgacgcctg cctgcccggc 29940

-continued

---

tggttcgggtt	cgccccgagt	ggaccacggt	aaaggggaaa	ccatctcctt	ctcgtttccg	30000
gacgggtcgc	gcagtagcat	cgtgtttgcc	tccagccaca	acacaaactg	aagtcctctt	30060
ttctttcgca	tggctctccc	aaggggcccc	gggtcgaccc	gacccacacc	cacccacca	30120
catacacaca	caaccagacg	cgggaggaaa	gtctgccccg	tgggcaactga	tttttatctg	30180
ggatcgcttg	aggaggcccc	ggcaacggcc	cgggcaacgg	tggggcaact	cgtagcaaat	30240
aggcgactga	tgtacgaaga	gaagacacac	aggcgccacc	cggcgctggt	cggggggatg	30300
ttgtccgcgc	cgcaccgtcc	cccagcagcc	tcttgacagc	ggtccgtgat	gcaaggacgg	30360
cggggggcct	gcagcagggt	gaccgtatcc	acgggatggc	caaagagaag	cggacacagg	30420
ctagcatccc	cctggaccgc	cagggtacac	tgggcatctc	tggccacagc	acacggggcg	30480
acgcagggac	aggactccgt	tacgacggag	gagagccaca	gtcgttggc	ggaatcgatg	30540
tggggcgcg	gggcgcagga	ctcgcagccc	cccgggtggt	tgggtgacct	ggccaggagc	30600
catcccagat	ggcggggcct	gcttcccggg	ggacagagcg	accccaggtc	gctgtccatg	30660
gcccagcagt	agatctggcc	gctggggagg	tgccaccagg	ccccggggcc	caaggcgcag	30720
cacgcgcccc	gctccggggg	ggtcttcgcg	gggaccagat	acgcgccatc	cagctcggcg	30780
accactggct	cctccgcgag	ctgttcgggt	ggtgggtcgg	gggtttctct	cgggggggtg	30840
gccgcccgta	tgcgtgcgaa	cgtgaggggt	cacaggagcg	gggtcagggg	gtcgtcaccg	30900
ctccggaggt	ggacgatcgc	gcagtagcgg	cgtctcgggt	taaagaaaaa	gagggcaaaag	30960
aaggtgttcg	ggggcaaccg	cagcgccttg	gggcgcgtca	gatacagaaa	aatctcgcag	31020
aagagggcgc	gcccgggggtc	tgggttagga	agggccacct	gacacagagg	ctcggtaggg	31080
accgttagac	accgaaagat	cttagaccgc	tcttccgccc	gaacgacgcg	ccacacaaag	31140
acggagttag	caatgcgcgc	gatagagtcg	acgtccgtcc	ccaggtcgtc	gactctatct	31200
cgcgtgcccc	gagctccgcg	ccgggaatcc	ggccggggca	aggtecccg	gggaccaggc	31260
gggccaccgg	gcccgggggg	tcccagctgc	gccatgccgg	gggcgggggg	agggcaaac	31320
ccagaggcgg	gggccaaccg	cgcggggagg	agtgggtggg	caggtgggcc	gggggaaggc	31380
gcccgcctagc	gagaccggcg	gttcccggac	gacaccttgc	gacaaaacct	aaggacagcg	31440
gcccgcgcga	cgggggtccga	gaggctaagg	taggcccga	tgttaatggt	gaacgcaaaag	31500
ccgccgggaa	agacaactat	gccacagagg	cggcgattaa	accccaggca	gaggtaggcg	31560
tagctttccc	cgggcaggta	ttgctcgcag	accctgcgtg	gggctgtgga	ggggacggcc	31620
tccatgaagc	gacatttact	ctgctcgcgt	ttactgacgt	caccatccat	cgccacggcg	31680
attggacgat	tgttaagccg	cagcgtgtct	ccgcttgtgc	tgtagtagtc	aaaaacgtaa	31740
tggcctcgg	agtcggcaaa	cggggccggg	aggctcgtcg	cagcggggac	gacccgccgc	31800
ccccgaccgc	ccgctcccc	caggtgtgcc	aggacggcca	gggcatacgc	ggtgtgaaaa	31860
aaggcgtcgg	ggcgggtccc	ctcagcggcg	cgcacaggt	tctcaggag	aatggggaag	31920
cgcctggtca	cctcccccaa	ccacgcgcgt	tggtcggggc	caaagtcata	gcgcaggcgc	31980
tgtgagattc	gcgggcccgc	ctgaagcgcg	gcccggatgg	cctggcccag	ggcccggagg	32040
cacgccagat	gtatgcgcgc	ggtaaaggcg	acctcggcgg	cgatgtcaaa	gggcggcagg	32100
acggggcgcg	ggtggcgcag	gggcacctcg	agcgggggaa	agcgtagcag	cagctccgcc	32160
tgcccagcgg	gagacagctg	gtggggcgcg	acgacgcgtt	ctcggcgcca	ggcctcggtc	32220

-continued

---

agggccgtgg ccagcgcccga ggacagcagc ggagggcggg cgcgtcgecc gccccacgcc 32280  
 acggagttct cgtaggagac gacgacgaag cgtgcttgg ttccgtagtg gtggcgacgg 32340  
 accacggaga tagaacgacg gctccacagc cagtccggcc ggtegcgcc gccacgggct 32400  
 tcccatccgc gatccaaaca ctcgaccagc gaccgaggct ttgcgggtacc aggggtaagg 32460  
 gttagaacgt cgttcaggat gtctctgccc ccggggcccg ggggcgctgg gggcacaag 32520  
 cggccccgc cggggggctc cagaccgcc agcaccgcat ctgcgtcagc cggccccatg 32580  
 gcgccccgc tgacggcctg gtgaaccagg gcgccctggc gtagccccga tgcaacgcca 32640  
 caggccgcac gcccggtccg cgtcgggacc ggggtggcggc gggtgacgtc ctgcaactgc 32700  
 cgtgaacca acgcgaggat ctctctgttc tctgtgcga tggacacgtc ctgggcccgc 32760  
 gtcgtgtcgc cgcggggggc cgtcagctgc tctccggggc agatggggggc gtcggacgcc 32820  
 ccgacgatgg gcgggtctgc gggcgcccc gcgtggggcc gggccaaggc ctgcggacgc 32880  
 ggggacgcgc tttccccag acccatggac aggtgggccc cagcctcctt cgcggccgcg 32940  
 gggcgccgcg cgcgaagcag agcagcgtag cggcacaat gccgacagac gcgcatgatg 33000  
 cgcgtgctgt cggcccgta gcgcgtgttg ggggggacga gctcgtcgtg actaaacaga 33060  
 atcacgcggg cacagctcgc ccccgagccc caccgaagcc gcagcgcgc caccggcgtac 33120  
 gggcataga cgcctcgcgc gtcacacacc acgggcaggg agacgaacaa cccccggcg 33180  
 ctggacgcac gcggaaggag gccagggtgt gccggcacga cgggggccc aagctcccc 33240  
 accgatccg cgggcacgta ggcggcaaac gccgtgcacc acggggtaca gtcgcccgtg 33300  
 gcatgagccc gactctgat ttcgacctgg aagtttgcgg ccgtcccag tccggggcgg 33360  
 ccgcatca gggcgcccag agggattccc gcggccgcca ggcactcgtt ggatatgatg 33420  
 acgtgaacca aagaccgagg gccgaccgg gccgtggccc agatcgtctg gactcgttg 33480  
 gccaagtgcg cgttcattgt tcgggggtgg gtgtgggtgt gtaggcgatg cgggtcccc 33540  
 gactccgccc gaagggcgtg ggtttggcgc gcgtatgcgt attcgccaac ggaggcgtgc 33600  
 gtgcttatgc gcggcgcgtt tctctctct ctagggaatc cgaggccagg actttaacct 33660  
 gctctttgtc gacgaggcca actttattcg cccggatgct gtcacagcga ttatgggctt 33720  
 tctcaaccag gccaaactgca agattatctt cgtgtcgtcc accaacaccg ggaaggccag 33780  
 tacgagcttt ttgtacaacc tccgcggggc cgcagacgag cttctcaacg tggtagccta 33840  
 tatatgcgat gatcacatgc cgagggtggt gacgcacaca aacgccacgg cctgttcttg 33900  
 ttatatctc aacaagccc ttttcacac gatggacggg gcggttcgcc ggaccgcccga 33960  
 tttgtttctg gccgattcct tcatgcagga gatcatcggg gccaggcca gggagaccgg 34020  
 cgacgaccgg cccgttctga ccaagtctgc gggggagcgg tttctgtgt accgcccctc 34080  
 gaccaccacc aacagcggcc tcatggcccc cgattgtac gtgtacgtgg atcccgcgtt 34140  
 cacggccaac acccgagcct ccgggaccgg cgtcgtctgc gtcggggcgt acccgacga 34200  
 ttatatcacc ttcgcctcgg agcacttttt tctccgcgcg ctcacgggct cggccccgc 34260  
 cgacatgcc cgtcgcgtc tccacagtct gacgcaggtc ctggccctgc atcccggggc 34320  
 gtttcgcggc gtcggggtgg cggtcgagg aaatagcagc caggactcgg ccgtcgccat 34380  
 cgccacgcac gtgcacacag agatgcaccg cctactggcc tcggaggggg ccgacgcggg 34440  
 ctcgggcccc gagcttctct tctaccactg cgagcctccc gggagcgcgg tgctgtaccc 34500

-continued

---

```

cttttctctg ctcaacaaac agaagacgcc cgcctttgaa cactttatta aaaagttaa 34560
ctccgggggc gtcattggcct cccaggagat cgtttccgcg acggtgcgcc tgcagaccga 34620
cccggctgag tatctgctcg agcagctaaa taacctcacc gaaacctctt cccccaacac 34680
tgacgtccgt acgtattccg gaaaacggaa cggcgctcgc gatgacctta tggtcgccgt 34740
cattatggcc atctacctcg cggcccaggc cggacctccg cacacattcg ctctatcac 34800
acgcgtctcg tgagcgccca ataaacacac ccaggtagc tacgcacgac cacggtgctg 34860
tctgttaagg gggggggggg aaggggggtg tggcggaag cgtgggaaca cgggggattc 34920
tctcacgacc ggcaccagta ccaacccctt gtgaacacag aaacccaac ccaaatccca 34980
taaacatacg acacacaggc atattttgga atttcttagg tttttattta tttaggtatg 35040
ctggggtttc tcctggatg cccaccccca cccccctg ggtctagccg gcccttaggg 35100
atagcgtata acgggggcca tgtctccgga ccgcacaac gccgcgccgt caaagtgca 35160
cacccgaacc acgggagcca gggccaaggt gtctcctagt tggcccgctt gggtcagcca 35220
ggcgacgagc gctctgtaa cggcgaccct tcgctctcca tcctgcatca gggccggggc 35280
ttcgggggta atgagctggg cggcctccc cgtgacactc tgcattcga gtagagcgtt 35340
cacgtaccgc tcctgggacc ttagcgcaaa gagccggggg attagcgtaa ggatgatggt 35400
ggttccctcc gtgacgcagt aaaccatgtt aaggaccagc gatcgcagct cggcgtttac 35460
gggaccgagt tgttgacgt ccgcccagc cgagaggcga ctcccggtgt agtacagcac 35520
gttgaggtct ggcagccctc cgggggttct ggggctgggg ttcaggtccc ggatgccct 35580
ggccacgagc cgcgccacga tttcgcgcgc caggggcgat ggaagcggaa cgggaaaccg 35640
caacgtgagg tcacagcaat ccaggcgcac gtccgctcgt tggccctcga acacggggcg 35700
gacgaggctg atggggctcc cgttacagag atctacgggg gaggtgttgc gaaggtaac 35760
ggtgccggcg tgggtgaggc ccacgtccag ggggcaggcg acgattcgcg tgggaagcac 35820
ccgggtgatg accgcgggga agcgccttcg gtacgccagc aacaaccca acgtgtcggg 35880
actgacgctt ccggagacga aggattcgtg cgcacgctc gccagcgtca gttgccggcg 35940
gatggtcggc aggaatacca cccgcccttc gcagcgtcgc agcgcgcgc catcgggggc 36000
cgagatgcc gagggatcgc cgatgtcagt ttcaaagccg tccgccagca tggcgccgat 36060
ccacgcggca gggagtgca tgggtggttc ggtggcggga ggagcgcggt gggggtcagc 36120
ggcgtagcag agacgggcca ccaacctcgc ataggacggg ggggtgggtct tagggggtg 36180
ggaggcgaca gggaccccag agcatgcgcg gggaggtctg tcgggcccag acgcaccgag 36240
agcgaatccg tcgcgggagt cccggcttgg gttttatggg gcccgccctt cggaatcgcg 36300
gcttctcggc ggggacaaaag gggcgggggc taggggcttg cggaaacaga agacgcgtgg 36360
gataaaagaa tcgcactacc ccaaggaagg gcggggcggt ttattacaga gccagtccct 36420
tgagcgggga tgcgtcatag acgagatact gcgcgaagtg ggtctcccgc gcgtgggctt 36480
ccccgttgcg ggcactgcgg aggagggcgg ggtcgcctgc gcaggtagc gggtaggcct 36540
cctgaaacag gccacacggg tcctccacga gttcgcggca cccccggggg cgttaaaact 36600
gtacgtcgtt ggcggcggtg gccgtggaca ccgccgaacc cgtctccacg atcaggcgtt 36660
ccaggcagcg atgtttggcg cgcgatgcgg ccgacgtaaa gaacttaaag caggggctga 36720
gcaccggcga gggcccggtt aggtggtagg ccccgttata gagcaggtcc ccgtacgaaa 36780

```

-continued

---

atcgctgcga	cgcccacggg	ttggccctgg	ccgcgaaggc	ccgggacggg	tcgctctggc	36840
cgtggtcgta	catgagggcg	gtgacatccc	cctccttgtc	ccccgcgtaa	acgcccccg	36900
cggcgcgtcc	ccgggggttg	cagggccggc	ggaagtagtt	gacgtcggtc	gacacggggg	36960
tggcgataaa	ctcacacacg	gcgtcctggc	cgtgggtccat	ccctgcgcgc	cgcggaacct	37020
gggcgcaccc	gaacacgggg	acgggctggg	ccggccccag	gcggtttccc	gccacgaccg	37080
cgttccgcag	gtacacgggt	cccgcgttgt	ccaggagagg	gggagccccg	cgccccaggt	37140
aaaagttttg	gggaaggttg	cccattgctg	tgacggggtt	gcggacgggt	gccgtggcca	37200
cgacggcggg	gtagccccacg	cccagggtcca	cgttcgcgcg	cggtcgggtg	agcgtgaagt	37260
ttaccccccc	gccagtttctg	tgcggggcca	cctggagctg	gcccaggaag	tacgcctccg	37320
acgcgcgctc	cgagaacacg	acgttctcag	tcacaagcgc	gtcctgtcgg	acgacgggtg	37380
acccaaaacc	gggatggagg	cccgtcttga	gctgatgatg	caaggccacg	ggactgatct	37440
tgaagtacc	cgccatgagc	gcgtaggtca	gcgcgttctc	cccggccgcg	ctctcgcgga	37500
cgtgctgcac	gacgggctgt	cggtcgcacg	aaaagtagtt	ggccccaga	gccgggggga	37560
ccagggggac	ctgccgcgac	aggtcgcgca	gggccggggg	gaaattgggc	gcgttcgcca	37620
cgtggtcggc	cccggcgaac	agcgcgtgga	cggggagggg	gtaaaaatag	tcgccatttt	37680
ggatgggatg	gtccagatgc	tggggggcca	tcagcaggat	tccggcgtgc	aacgccccgt	37740
cgaatatgcg	catgttggtg	tgggacgcgg	tgttgccgcc	cgcgtcgggc	gccgcgcgagc	37800
agagcagcgc	cgttgtgcgt	tccggccatg	tgtgggccag	cacctgcagc	gtgagcatgg	37860
cgggcccgtc	cactaccacg	cgcccgttgt	gaaacatggc	gttgaccgtg	ttggccacca	37920
gattggcccg	gtcagggggg	tgcgcggggg	ccgtcacggg	gtcgtcgggg	cactcctcgc	37980
cgggggcgat	ctccgggacc	accatgttct	gcagggtggc	gtatacgcgg	tcgaagcgaa	38040
cccccgcggt	gcagcagcgg	ccccgcgaga	aggcgggcac	catcacgtag	tagtaaatct	38100
tgtggtgcac	ggtccagtc	gcccccggt	gcggccggtc	atccgcggcg	tccgcggtct	38160
gggcctgggt	ggtgtgcagc	agctggccgt	cgttgccggt	gaagtccgcg	gtcgcacagt	38220
tacatgccgc	cgcttacacg	gggtcgtggc	ccccgcgct	aaccgcgag	tcgcatggc	38280
ggtccagggc	cgccgcgcc	atcagggcgt	cacagtcca	cacgaggggt	ggcagcagcg	38340
ccgggtctcg	cattaggtga	ttcagctcgg	cttgccctg	cccgccagc	tccgggccgg	38400
tcagggtaaa	gtcatecaacc	agctgggcca	gggcctcgac	gtgcgccacc	aggtcccgg	38460
acacggccat	gcactcctcg	ggaaggtctc	ccccgaggt	ggtcacgacg	tacgagacca	38520
gcgagtagtc	ggtcacgaac	gcccgcgacc	gcgtgttgtt	ccagtagctg	gtgatgcaact	38580
ggaccacgag	ccggggccagg	gcgcagaaga	cgtgctcgt	gccgtgatg	gcggcctgca	38640
gcaggtaaaa	caccgcggg	tagttgcgg	cgtcgaacgc	cccgcgaacg	gcggcgatgg	38700
tggcgggggc	catggcgtgg	cgccccacc	ccagctccag	gccccggcg	tcccggaacg	38760
ccgcgggaca	tagcgcaggg	ggcaagttgc	cgttcaccac	gcgccaggtg	gcttgatct	38820
cccccgggcc	ggcgggggga	acgtcccccc	ccggcagctc	cacgtcggcc	acccccacaa	38880
agaagtcgaa	cgcggggtgc	agctcaagag	ccaggttggc	ggtgtcgggc	tgataaaact	38940
gctccgggg	catctggcct	tccgcgaccc	atcggaccgg	cccgtggggc	aggcgtgccc	39000
cccagcggtt	caaaaacacg	tgctgcatgt	ctgcggcggg	gccggccggg	gccgccacgt	39060

-continued

---

```

acgccccgta cggattggcg gcttcgacgg ggtcgcgggtt aaggcccccg accgcccgt 39120
caacgttcat cagcgaaggg tggcacacgg tcccgatcgc gtgttccaga gacaggcgca 39180
gcacctggcg gtccttcccc caaaaaaaca gctggcgggg cgggaaggcg cgggatccg 39240
ggtggccggg ggcggggact aggtccccg cgtgcgcggc aaaccgttcc atgaccggat 39300
tgaacaggcc caggggcagc acgaacgtca ggtccatggc gccaccagg gggtagggaa 39360
cgttgggtgc ggcgtagatg cgctctccca gggcctccag aaagaccagc ttctcgccga 39420
tggacaccag atccgcgcgc acgcgcgtcg tctggggggc gctctcgagc tcgtccagcg 39480
tctgccggtt caggteagc tgctcctcct gcatctccag caggtggcgg cccacgtcgt 39540
ccagacttcg caccgacctg cccatcacga gcgccgtgac caggttggcc ccgttcagga 39600
ccatctcgcc gtacgtcacc ggcacgtcgg cttcgggtgc ctccacttc aggaaggact 39660
gcaggaggcg ctgtttgatc ggggcgggtg tgacgagcac cccgtcgacc ggcgccccgc 39720
gcgtgtcggc atgcgtcaga cggggcacgg ccacggaggc ctgcgtggcc gtggtgaggt 39780
ccacgagcca ggcctcgagc gcctccccgc ggtggccccg cttgcccagg aaaaagctcg 39840
tctcgcagaa gcttcgcttt agctcggcga ccagggtcgc cggggccacc ctggtggcca 39900
ggcggccggt gtcagggtat cgttgcatcg gcaacaacaa agccaggggc ggcgcctttt 39960
ccagcagcac gtgcagcatc tggtcggcgc tgccgcgctc aaacgccccg aggacggcct 40020
ggacgttgcg agcgagctgt tggatggcgc gaaactggcg atgcgcgccg ataccctcc 40080
cgtccagggc ctccccctg agcagggcga tggcctcgtt ggccaggctg aaggcggcgt 40140
tcaggggccc gcggtcgata atcttggta tgaattgtg tgtgggttc tcgatgggt 40200
gcgggccgct cggggcaatc agcggctggt ggacctcgaa ctgtacgcgc cctcgttca 40260
tgtaggccag ctccgaaac ttggtacaca cgcacgccac cgacaaccg agctccagaa 40320
agcgcacgag cgacagggtg ttgcaatcgc accccagcag ggcgtcgaa tcgacgtcgt 40380
acaggctggt tgcacggag cgcacgcggg aaaaaaac aaacaggcgt cgatgcgagc 40440
ccacctcgat cgtgctaagg agggaccctg tcggcaccat ggcgcggcga taccggtatc 40500
ccggagggtc gcggttggga gcggccatgg ggtcgcgtgg agatcggctg tctctagcga 40560
tattggcccg gggaggctaa gatccacccc aacgcccgcg caccctgta cgtgcccagc 40620
ggcccaaggt ccaccgaaa acacgacggg cccggaccca aaaaggcggg ggatgctgtg 40680
tgagaggccc ggtgccggtc gggggggaaa ggcaccggga gaaggtgcg gcctcgttcc 40740
aggagaacct agtgteccca acagaccctg ggacgtggga tcccaggcct tatatacccc 40800
cccccccgc cccccctg tagaacgcga cgggtgcatt caagatggcc ctggtccaaa 40860
agcgtgccag gaagaaattg gcagaggcgg caaagctgtc gcgcgccc acccacatcg 40920
aggccccgc cgcgaggct atccccagg cccgtgtgcg caggggatcg gtggcgcgca 40980
gcatttgggt ggtggcgata aagtggaaaa gcccgcccgc actgaaggtc tcgtggcgcg 41040
cggcgaacaa ggcacacagc gccgtgcctc ccaaaaacac ggacatcccc caaaacacgg 41100
gcgcccagaa cggcagacga tccctcttga tgttaacgta caggaggagc gcccgaccg 41160
cccacgtaac gtagtagccg acgatggcgg ccagataca ggcggcgccc accacccttc 41220
cggtcagccc gtaatacatg cccgctgcca ccatctccaa cggcttcagg accaaaaacg 41280
accaaaggaa cagaatcacg cgctttgaaa agaccggctg ggtatggggc ggaagacgcg 41340

```

-continued

---

```

agtatgccga actgacaaaa aaatcagagg tgccgtacga ggacaatgaa aactgttcct 41400
ccagcggcag ttctccctcc tccccccga aggeggcctc gtcgaccaga tctcgatcca 41460
ccagaggaag gtcacccccg atggtcatgg ggtgtgcggt ggagggtggg agaccgaaac 41520
cgcaaagggt cgcttacgtc agcaggatcc cgagatcaaa gacacccggg ttcttgacaca 41580
aacaccaacc gggttgcac cgcggaggcg agtggtttga taaggccgtt ccgcgccctg 41640
atataacctt tgatgttag cacaaaaacc ggaatttacg cctacgcccc aatgccccag 41700
caagatgagg taggtaaccc ccccggtggg gtgacgttgc gtttagttca ttggaggcca 41760
aggggaaaaa tggggtgggg aggaaacgga aaaccagta ggccgtgtcg ggaacacgcc 41820
cggggttgc ctcaaaaggc aggggtccata ctacggaagc cgtcgttga ttcgagacct 41880
gcctgtgcaa cgcacgtcgg ggttgccctg gtccggttcg gccccaccg cgtgcgccac 41940
gcacgaggac gagtccgctg gctttattgg cgttccaagc gttgccctcc agtttctggt 42000
gtcgggttcc ccccatacc acgcccacat ccaccgtagg gggcctctgg gccgtgttac 42060
gtcgcgcccc gcgatggagc ttagctacgc caccaccatg cactaccggg acgttctggt 42120
ttacgtcaca acggaccgaa accgggccta ctttgtgtgc ggggggtgtg tttattccgt 42180
ggggcgccgc tgtgcctcgc agcccgggga gattgccaag tttggtctgg tcgttcgagg 42240
gacaggccca gacgaccgcy tggctgcca ctatgtacga agcgagctcc gacaacgcgg 42300
cctgcaggac gtgcgtccca ttggggagga cgaggtgttt ctggacagcg tgtgtcttct 42360
aaaccgaaac gtgagctccg agctggatgt gattaacacg aacgacgtgg aagtgtgga 42420
cgaatgtctg gccagtact gcaactcgtc gcgaaccagc ccgggtgtgc taatatccgg 42480
gctgcgctg cgggcgccag acagaatcat cgagttgttt gaacacccaa cgatagtcaa 42540
cgttctctcg cactttgtgt ataccccgtc cccatacgtg ttcgccctgg cccagcgca 42600
cctcccccg cteccgagct cgctggaggc cctggtgagc ggccgtttg acggcatccc 42660
cgccccacgc cagccacttg acgcccacaa ccgcgcacg gatgtggtta tcacgggccc 42720
ccgcgcccc cagcccatcg ccgggtcggg ggcggggtcg gggggcgccg gcgccaagcg 42780
ggccaccgtc agcgagtctg tgcaagtcaa acacattgac cgcgtgggccc ccgctggcgt 42840
ttcgcggcg cctccgcca acaaacccga ctcgagttcc ctggtgccc gggcccagga 42900
ttccgcccc ccggcccca cgctaaggga gctgtggtgg gtgttttatg ccgcagaccg 42960
ggcgctggag gagccccgcy ccgactctgg cctcaccgcy gaggaggtac gtgccgtacg 43020
tgggttccg gagcaggcgt ggaaactggt tggctccgcy ggggccccgc gggcgtttat 43080
cggggccgcy ttgggctgca gccccctca aaagctagcc gtttactact atatcatcca 43140
ccgagagagg cgctgtccc ccttccccgc gctagtcgcy ctcgtaggcc ggtacacaca 43200
gcgccacggc ctgtaagtc ctcggccga cgaccagtc ttggccgatg ccatcaacgg 43260
gctgtttcgc gacgcgctgg cggccggaac cacagccgag cagctcctca tgttcgacct 43320
tctccccca aaggacgtgc cgggtgggaa cgacgtgcag gccgacagca ccgctctgct 43380
gcgctttata gaatcgcaac gtctcgccgt ccccgggggg gtgatctccc ccgagcacgt 43440
cgcgtaacct ggtgcgttcc tgagcgtgct gtacgctggc cgcgggcca tgtccgcagc 43500
cacgcacacc gcgcgctga caggggtgac ctccctggtg ctagcgggtg gtgacgtgga 43560
ccgtcttcc gcgtttgacc cgcgagcggc gggcgcgccc agccgcaagc gggccgcccg 43620

```

-continued

---

```

gtacctggat gtgcttctta cegttcgtct cgctcgctcc caacacggac agtctgtgta 43680
aaagacccca ataaacgtat atcgctacta cacccttgtg tgtcaatgga cgcctctccg 43740
gggggggggg agggaaagca aagaggggct gggggagcgg caccaccggg gcctgaacaa 43800
acaaaccaca gacacgggta cagttttatc ggtcgggcgg agaaacggcc gaagccacgc 43860
ccactttatt cgcgtctcca aaaaaacggg aacttgtcc ggagaacctt taggatgcca 43920
gccagggcgg cggtaatcat aaccacgccc agcgcagagg cggccagaaa cccgggcgca 43980
attgcgccca cgggctgcgt gtcaaaagct agcaaatgaa tgacgggtcc gtttgaaat 44040
agcaacaagg ccgtggacgc cacgtcgcgc gaaaacacgc ttggggcgcc ctccgtcggc 44100
ccggcggcga tttgctgctg tgtgtgtgct gtatccacca gcaacacaga catgacctcc 44160
ccggccgggg ttagcgcgat aaacacggcc cccacgagcc ccaggtcgcg ctggttttgg 44220
gtgcgcacca gccgcttggc ctcgatatcc cgggtggagc ctccgatgt cgcggtgagg 44280
taggttagga acagtgggcg tcggacgtcg acgcccgtga gctttagacc gatcccccg 44340
ggcagagggg agtgggtgac gacgtagctg gcggttgggg tgatgggtac caggatccgt 44400
ggctcgacgt tggcagactg cccccgcac cgatgtgagg cctcagggac gaaggcgcgg 44460
atcagggcgt ttagtgtgct ccaacgcgct agggtcgagg cgaggccgtg ggtctgctgg 44520
gccaggactt cgaccggggt ctccggatcg gtggttgag ccagcgcgct caggataaac 44580
acgctctcgt ctatagcaaa gcgcaggag gccgcgatg gcgaaaagtg gtccggaagc 44640
caaaagaggg tttctggtg gtccggccgg gccagcggcg tccggaggtc ggcggtggtc 44700
gctgcggcga cgtcggacgt acacagggcc gaggctatca gaaggctccg gcgggcgcgt 44760
tcccgcgca ccgccgaggg gacgccagcc aagaacggct gccggaggac agccgaggcg 44820
taaaaatagc cccggtggac gaccggggtg gtcagcacgc gggcccctag aaactcggca 44880
tacagggcgt cgatgagatg ggctgcgctg ggcgccactg cgtcgtaacg cgaggggcta 44940
tccagcacga aggccagctg atagcccagc gcgtgtaatg ccaagctctg ttcgcgctcc 45000
agaatctcgg ccaccaggtg ctggagccga gcctctagct gcaggcgggc cgtgggatcc 45060
aagactgaca cattaaaaa cacagatcc gcggcacagc ccgcggcccc gcgggcggcc 45120
aacccgcaa gcgcgcgcga gtgggcaaaa aagcctagca ggtcggagag gcagaccgcg 45180
ccgtttcgt gggcggcggt caccgaaagca aaaccgagc tcgcgagcag ccccgttagg 45240
cgccagaaga gagggggcg cgggccctgc tcggcgcggc cgtccccga gaaaaactcc 45300
gcgtatgcc gcgacaggaa ctgggcgtag ttcgtgccct cctccgggta gccgcccacg 45360
cggcggaggg cgtccagcgc ggagccgttg tcggcccgcg tcagggaccc taggacaaag 45420
accgatacc gggggccgcc cggggggccc ggaagagccc ccggggggtt ttcgtccgcg 45480
gggtccccga ccgatctag cgtctggccc gcggggacca ccatcacttc caccgagggg 45540
ctgtcgtgca tggatatcac gagcccctag aattcccgcc cgtagcgcgc gcgcaccagc 45600
gcggcatcgc acccgagcac cagctcccc gtctccaga tgcccacggg ccacgtcag 45660
gccagcgggg agaaatcac gtacctacct ggggatctca acaggccccg ggtggccaac 45720
caggctcgtg acgctgtgtg caggctcgtg atgtccagct ccgctcgtcg gtgcccggcg 45780
gccccaacg gcggtcgggg gggcgggtga tcacgcccgc cgctcgggtg gctcgcgctc 45840
gccacgttgt cccccgcgg gaaacgtcag gcctcggggt cagggacggc cgaaacgtt 45900

```

-continued

---

acccaggccc gggaaacgcag caacacggag gcggtggat tgtgcaagag acccttaagg 45960  
 ggggagaccg aggggggagg ctgggaggtc ggctcgaccg tggggggggc gggcaggctc 46020  
 gcggtcgggg gccggccgag caggttagtc ttcgggatgt aaagcagctg gccgggggtcc 46080  
 cgcggaact cgcccggtgt gaccaatata aaacaaaagc gctcctcgta ccagcgaaga 46140  
 aggggcagag atgccgtagt caggtttagt tcgtccggcg gcgccagaaa tccgcgcggt 46200  
 ggtttttggg ggtcgggggt gtttggcagc cacagacgcc cgggtgtcgt gtcgcgccag 46260  
 tacatgctgt ccattgccag gccatccaaa aacctgggt ctgtctgctc agtccagctc 46320  
 tggacctgac cccacgcaac gcccaaaaata ataaccacca cgaaccataa accattcccc 46380  
 atgggggacc ccgtccctaa cccacggggc ccgtggctat ggcagggtt gccgccccga 46440  
 cggtggctgc gagccctggg ccttcaaccg aacttggggg ttgggggtgg gaaaaggaag 46500  
 aaacgcgggc gtattggtcc caatggggtc tcgggtgggt atcgacagag tccagccct 46560  
 gggaccgaac cccgcgttta tgaacaaaac acccaacacc cgtgcgtttt attctgtctt 46620  
 tttattgccc tcatagcgcg ggttccctcc ggtattgtct ccttccgtgt ttcagtttagc 46680  
 ctccccatc tccggggcaa acgtgcgcgc caggtcgagc atcgctcgta tggagcctgg 46740  
 ggtggtgacg tgggtctgga ccatcccgga ggtaagtgtc agcagggcgt cccggcagcc 46800  
 ggcgggggat tggctgtaat ccaggataaa gacatgcatg ggacggaggc gtttggccaa 46860  
 gacgtccaaa gcccaggcaa acacgttata caggtcgccg ttgggggcca gcaactcggg 46920  
 ggcccgaaac agggtaataa acgtgtcccc gatatggggt cgtgggcccg cgttctctg 46980  
 gggctcggca ccctggggcg gcacggccgc cccgaaaagc tgtcccaat cctcccgcca 47040  
 cgaccgccg ccctgcagat accgcaccgt attggcaagc agccataaa cgcggcgaat 47100  
 cgcgccagc atagccaggt caagccgctc gccggggcgc tggcgtttg ccaggcggtc 47160  
 gatgtgtctg tcctccgaa gggcccccac cactggttt gtgcccggca aggtcggcgg 47220  
 gatgagggcc acgaaaccca gcacggcctg gggggctatg ctgcccataa ggtatcgcc 47280  
 ggcgggtag cacaggagg gcggcagtg atggcggctc aagatgaggg tgagggcccg 47340  
 gggcggggca tgtgagctcc cagcctcccc cccgatatga ggagccagaa cggcgtcgt 47400  
 cacggcataa ggcattgcca ttgttatctg ggcgcttctc attaccacc cgcgctcccc 47460  
 ggcgatatc tcacctggt cgaggcgggt ttgtgtggtg tagatgtctc cgattgtctc 47520  
 ggaagcccc aacaccgcc agtaagtcat cggtcgggt acgtagacga tctcgtcgcg 47580  
 cgaaccagc gccaccagca gttgcgtggt ggtggttttc cccatccctg ggggaccgctc 47640  
 tataaaacc cgcagtagcg tgggcatttt ctgctccagg cggacttccg tggetttttg 47700  
 ttgcccggca gggcgcaacg ccgtacgtcg gttgttatgg ccgcgagaac gcgcagcctg 47760  
 gtcgaacgca gacgcgtggt gatggcaggg gtacgaagcc atacgcgctt ctacaaggcg 47820  
 ctggccgaag aggtgcggga gtttcacgcc accaagatct gcggcacgct gttgacgctg 47880  
 ttaagcgggt cgtgcaggg tcgctcggtt ttcgaggcca cacgcgtcac cttaatatgc 47940  
 gaagtggacc tgggaccgcg cgcggccgac tgcactcgcg tgttcgaatt cgccaatgac 48000  
 aagacgctg gcggggtttg tgctcatata gaactaaaga catgcaata tatttcttcc 48060  
 ggggacaccg ccagcaaacg cgagcaacgg gccacgggga tgaagcagct gcgccactcc 48120  
 ctgaagctcc tgcagctcct cgcgcctccg ggtgacaaga tagtgtaect gtgccccgctc 48180

-continued

---

ctggtgtttg tcgcccacg gacgctccg gtcagcccg tgacccggct cgtcccgcag 48240  
aaggtctccg gtaatatcac cgcagtcgtg cggatgctcc agagcctgtc cacgtatacg 48300  
gtccccattg agcctaggac ccacgagacc cgtcgcgcc gcggcggcgc cgcggggggg 48360  
tctgcgagca gaccgaaaag gtcacactct ggggcgcgcg acccgccoga gtcagcggcc 48420  
cgccagttac caccgcgca ccaaaccccc acctccacgg agggcggggg ggtgcttaag 48480  
aggatcgcgg cgtctttctg cgtgcccggt gccaccaaga ccaaaccccg agccgcctcc 48540  
gaatgagagt gtttcgttcc ttccccctcc ccccgctca gacaaacct aaccacgct 48600  
taagcggccc ccgagaggc cgaagactca tttggtacc gcgggagcca cccgacaaca 48660  
gcccccggt tttcccacgc cagacgcccg tccgctgtgc catcgcgcc cctcatcca 48720  
cccccatct tgtccccaaa taaaacaagg tctggtagt aggacaacga ccgagttct 48780  
cgtgtgttat tttcgtctc cgcctctcgc agatggacc gtactgcca tttgacgctc 48840  
tggacgtctg ggaacacagg cgttcatag tcgccgattc ccgaaactc atcaccgccg 48900  
agtcccccg ggacttttgg atgctccccg tctttaacct cccccgggag acggcggcgg 48960  
agcaggtggt cgtcctacag gccagcgcga cagcggctgc cgtgccttg gagaacgccg 49020  
ccatgcaggc ggccgagctc cccctcgata tcgagcggc gttacgccg atcgaacgga 49080  
acgtgcacga gatcgcaggc gccctggagg cgtggagac gcggcggcc gccgcggaag 49140  
aggcggatgc cgcgcgcgg gatgagccgg cgggtgggg gcagcgggg gcgcccccg 49200  
gtctggcgt cgcggagatg gaggtccaga tcgtgcgcaa cgaccgccg ctacgatacg 49260  
acacaaact ccccggtgat ctgctacaca tgggttacgc gggccgcgg gcgaccggct 49320  
cgtcgggggt ggtgttcggg acctggtacc gcactatcca ggaccgcacc atcaccgact 49380  
ttcccctgac caccgcagt gccgacttc gggacggcc tatgtccaag accttcatga 49440  
cggcgtggt actgtccctc caggcgtgcg gccggctgta tgtgggccag cgcactatt 49500  
ccgccttca gtcgcgcgtg ttgtgtctct acctgctgta ccgaaacacg cacggggcgg 49560  
ccgacgatag cgaccgcgt ccggtcacgt tcgggatct gctgggcgg ctgccccgct 49620  
acctggcgt cctggcgcg gtgatcggga ccgagggcg ccggccacag taccgtaac 49680  
gcgacgaaa gctccccaa agcagttcg cggccggcgg gggccgctac gaacacggag 49740  
cgtggcgtc gcacatcgt atcgcacgc tgatgacca cggggtgctc ccggcggccc 49800  
cgggggacgt cccccgggac gcgagtacc acgttaacc cgacggcgtg gcgaccacg 49860  
acgacataaa ccgcccgcg gcccgcttcc tcagccgggg ccacaaccta ttctgtggg 49920  
aggaccagc tctgctgcgg gcaaccgca acaccataac ggccctgggc gttatccagc 49980  
ggctcctcgc gaacggcaac gtgtacgcgg accgcctcaa caaccgctg cagctgggca 50040  
tgctgatccc cggagcgtc ccttcggagg ccatcgcgg tggggcctcc gggctcgaact 50100  
cgggggcat caagagcgg gacaacaatc tggagcgcct atgtgcaat tacgtgcttc 50160  
cgtgtaccg ggccgaccg gcggtcgcgc tgaccagct gtttcccggc ctggcgcgcc 50220  
tgtgtcttga cgcaggcgg gggcggccgg tcgggtcgac gcggcgggtg gtggatatgt 50280  
catcgggggc ccgccaggc gcctggtgc gcctcaccgc cctggaactc atcaaccgca 50340  
cccgcacaaa ccccccctt gtgggggagg ttatccacgc ccacgacgc ctggcgatcc 50400  
aatacgaaca ggggcttggc ctgctggcgc agcaggcacg cattggcttg ggtccaaca 50460

-continued

---

```

ccaagcgttt ctccgcttc aacgttagca gcgactacga catgttgtag tttttatgtc 50520
tggggttcat tccacagtag ctgtcggcgg tttagtgggt ggtgggagag gggggagggg 50580
gcattagga gaaaaaacia gagcctccgt tgggttttct ttgtgcctgt actcaaaagg 50640
tcataccccg taaacggcgg gctccagctc cgcccgggcg gttggcgtga acgcaacggc 50700
gggagctggg ttagcgttta gtttagcatt cgctctcgcc ttccgccccg ccccccgacc 50760
gttgccctt ttttttttcc gtccacaaa gtctctgtgg gtgcgcgat ggcagccgat 50820
gccccgggag accggatgga ggagcccctg cccgacaggg ccgtgcccat ttacgtggct 50880
gggtttttgg cctgtatga cagcggggac tcgggcgagt tggcattgga tccggatacg 50940
gtgcggcggg cctgcctcc ggataaccca ctcccagta acgtggacca ccgcgctggc 51000
tgcgaggtgg ggcgggtgct ggccgtggtc gacgaccccc gcgggcccgt ttttgtgggg 51060
ctgatcgct gcgtgcagct ggagcgcgtc ctcgagacgg ccgccagcgc tgcgatttcc 51120
gagcgcggcg ggcggcggct ctccggggag gagcgcctgt tgaactgat caccaactac 51180
ctgccctcgg tctccctggc cacaaaaagc ctggggggcg aggcgcaccc cgatcgcaag 51240
ctgttcgcgc acgtgcgct gtgcgcgatc gggcggcgc tcggcactat cgtcacctac 51300
gacaccggtc tcgacgccc catcgccccc ttccgccacc tgcgcggcgc gtctgcgag 51360
ggggcgcggc gactggccc cgagggcgag ctgcgctgt ccggggcgcac ctggggcggc 51420
ggcgtggagg cgctgaccca cacgctgctt tccaccggc ttaacaacat gatgctgcgg 51480
gaccgctgga gctgtgtggc cgagcggcgg cggcaggcgg ggatgcggg acacacctac 51540
ctccaggcga gcgaaaaatt caaatgtgg ggggcggagc ctgtttccgc gccggcgcgc 51600
gggtataaga acggggcccc ggagtccacg gacataccgc ccggctcgat cgctgccggc 51660
ccgcagggtg accggtgccc aatcgtccgt cagcgcgggg tcgccttgtc cccggtactg 51720
cccccatga acccggcttc gacatcgggc acccggggcc ccgcgcggcc cggcgcggg 51780
agctacctgt ggatcccgcc ctcccattac aaccagctcg tcggcggcca tgcgcggccc 51840
caacccagc cgcattccgc gtttggttcc ccggctgcgg cggggtcctg ggctatggg 51900
cctcaccgtg cgggttcttc ccagcattac cctcccacg tcgcccata gtatccggg 51960
gtgctgttct cgggaccag ccaactcgag gcgcagatag ccgcgttggg gggggccata 52020
gccgcggacc gccaggcggg cggtcagccc gccgcgggag accctggggg cggggggtcg 52080
ggaaagcgtc gccggtacga ggcggggccc tcggagtctt actgcgacca ggacgaaccg 52140
gacgcggact acccgtacta ccccggggag gctcgaggcg cgcgcgcgg ggtcgactcc 52200
cggcgcggcg cccgccattc tcccgggacc aacgagacca tcacggcgt gatggggggc 52260
gtgacgtctc tgcagcagga actggcgcac atgcgggctc ggaccagcgc cccctatgga 52320
atgtacagc cggtgccgca ctatgcctc cagggtgggg agccggaacc aacaacgacc 52380
caccggccc tttgtcccc ggaggccgtg tatcggcccc caccacacag cccccctac 52440
ggtctcccc agggctccgg gtcccatgcc cccactcccc cgtatgcccc agctgectgc 52500
ccgcaggcc cgcaccgcc cccatgtctt tccaccaga cgcgcggccc tctaccgag 52560
gagcccgct tccccccgc gcceaccgga tcccaaccgg aggcatacaa cgcggaggcc 52620
ggggcccttg tcaacgccag cagcgcagca cacgtggagc ttgacacggc ccgcgcggcc 52680
gatttgttcg tctctcagat gatggggggc cgtgatttcg ccccggtctt tggtaacctg 52740

```

-continued

---

ggatgtctta ctgtatatct ttttaataa accaggtaat accaaataag acccattggt 52800  
 gtatgttctt tttttattgg gaggcgcggg taggcgggta gctttacaat gcaaaagcct 52860  
 tcgacgtgga ggaagcgctg gggggggggg gaatcggcac tgaccaaggg ggtccgtttt 52920  
 gtcacgggaa aggaaagagg aaacaggcgc cggacaccgc ggggagtttg tgttcccttt 52980  
 tctttcttcc cacacacaca aaaggcgtac caaacaaca aaccaaaaga tgcacatgcg 53040  
 gtttaacacc cgtggttttt atttacaaca aacccccat cacaggctcg cctcgtcggc 53100  
 gtcaccgtct ttgttgggaa cttgggtgta gttggtggtg cggcgcttgc gcatgaccat 53160  
 gtcggtgacc ttggcgctga gcagcgcgct cgtgcccctc ttcttggcct tgtgttccgt 53220  
 gcgctccatg gcagacacca gggccatgta ccgtatcatc tcccgggect cggttagctt 53280  
 ggcctcgtca aagtcgccc cctcctcggc ctccccggac gcgtccgggt tggtaggggt 53340  
 cttgagctcc ttggtggtta gcgggtacag ggcctcatg gggttgctct gcagcgcgat 53400  
 gacgtagcga aaggcgaaga aggcgcgcgc caggccggcc aggaccaaca gaccacggc 53460  
 cagcgcccca aaggggttgg acatgaagga ggacacgccc gacacggcgc ataccacgcc 53520  
 gcccacgatg cccatcacca ccttgccgac cgcgcgcccc aggtcgccca tcccctcgaa 53580  
 gaacgcgccc aggcccgcaa acatggcggc gttggcgctg gcgtggatga ccgtgtcgat 53640  
 gtcggcgaag cgcaggctgt gcagctggtt gcggcgctgg acctccgtgt agtccagcag 53700  
 gccctgtcc ttgatctcgt ggcgggtgta caactccagg gggacaaact cgtgatcctc 53760  
 cagcatggtg atgttgaggt cgatgaaggt gctgacgggt gtgatgtcgg cgcggctcag 53820  
 ctggtgggag tacgcgtact cctcgaagta cacgtagccc ccaccgaagg tgaagtagcg 53880  
 ccggtgtccc acgggtgcaag gctcgcgcgc atcgcgcgct agccgcagct cgttgttctc 53940  
 cccagctgc cctcagacca acgggcccctg gtcttcgtac cgaaagctga ccagggggcg 54000  
 gctgtagcag gcccggggcc gcgagctgat gcgcatcgag ttttggacga tcacgttgtc 54060  
 cgcggcgacc ggcacgcacg tggagacggc catcacgtcg ccgagcatcc gcgcgctcac 54120  
 ccgcggccc acgggtggcc aggcgatggc gttggggttc agcttgcggg cctcgttcca 54180  
 cagggtcagc tcgtgattct gcagctcgca ccacgcgatg gcaacgcggc ccaacatctc 54240  
 gttgacatgg cgctgtatgt ggttgaactt aaactgcagc ctggcgaact cgatggagga 54300  
 ggtggtcttg atcgcctcca cggacgcggt ggcgctggcc ccgggcgggc ggggctggg 54360  
 gtttgggggc ttgcgctct cctcgcggag gtgttcccgc acgtacagct ccgcgagcgt 54420  
 gttgctgaga aggggtggt acgcgatcag aaagcccca ttggccaggt agtactgctg 54480  
 ctggcccacc ttgatgtgcg tcgcgttcta cctgcggggc aagatgcggc ccatggcgtc 54540  
 gcggcgctcc ttgccgatgc agtccccag gtccacgcgc gagagcgggt actcggtcag 54600  
 gttggtggtg aaggtggtg atatggcgtc ggaagagaat cggaaggagc cgcctactc 54660  
 ggagcgcagc atctcgtcca cctcctgcca cttggtcatg gtgcagaccg acgggcgctt 54720  
 tggaccaccag tcccaggcca cgggtgaactt gggggtcgtg agcaggttcc ggggtgctcg 54780  
 cgcctgggcc cgggcttgg ttggtgaggt gcgcgcgtag aagccgtcga cctgctttaa 54840  
 gcggtcggcg gcgtagctgg tgtgttcggt gtgcgacccc tcccggtagc cgtaaaacgg 54900  
 ggacatgtac acaagctgc cagtcgcca caaaaactcg tcgtacgggt acaccgagcg 54960  
 cgcgtccacc tcctcagca tgcagtttac cgtcgtcccg taccgggtgga acgcctccac 55020

-continued

---

ccgcgagggg ttgtacttga ggtegggtgt gtgccagccc cggctcgtgc gggtcgcggc 55080  
 gttggccggg ttcagctcca tgcctgtctc gtggctgtcc cggtgaaacg cggtggtctc 55140  
 caggttgttg cgcacgtact tggccgtgga ccgacagacc cccttgccgt tgatcttgtc 55200  
 gateacctcc tcgaagggga cgggggcgcg gtctctaaaag atccccataa actgggagta 55260  
 gcggtggccg aaccacacct gcgaaacggg gacgtctttg tagtacatgg tggccttgaa 55320  
 cttgtacggg gcgatgttct ccttgaagac caccgcgatg ccctccgtgt agttctgacc 55380  
 ctccggcccg gtcgggcagc ggcgcggctg ctcgaaactgc accaccgtgg cgccttgagg 55440  
 ggggtgggca acgtaaaaag ttgcatcggg gttctccgcc ttgatgtccc gcaggtgctc 55500  
 gcgcaggggt gcgtggcccg cggcgacggg cgcgttgtcg ccggcggggc gcggcggtt 55560  
 tgggggttcc ggttttctgt tcttcttcgg ttccgggtcc cccgttgggg gggcgccagg 55620  
 ggcgggcggc gcgggagtg caggcccccc gttcccgccc tgggtcgcgg ccgcgacccc 55680  
 aggcgtgccg ggggaaactc gagccgcgca gccaccagg acccccagcg tcaaccccaa 55740  
 gagcgcctat acgacgaacc accggccccc ccgcgcgggg gcgcccctgg gcctggcggg 55800  
 actacggggg cccctcgtgc ccccctcag gtagcctggg ggcgaggtgc tggaggaccg 55860  
 agtagaggat cgagaaaaag tctcggctgt agaccacgac cgaccggggg ccgatacagc 55920  
 cgtcgggggc gctctcgacg atggccacca gcggacagtc ggagtcgtac gtgagatata 55980  
 gcgcggcgcg gtaacggtaa gcaccttcgg aggtcgggcg gctgcagtcc gggcgcgcca 56040  
 actcgagctc ccgcgaccgg tagaccgagg caaagagtg ggtggcgata atcagctcgc 56100  
 gaatatctc ccagcggcg cgctgagtg gcgttattcc ggaaatgcc tcaaacagc 56160  
 aaaacctctg aaattcgtg acggcccaat cagcaccgca gccccccgccc cccatgatga 56220  
 accgggcgag ctctctctc aggtgcggca ggagccccc gttctcgac ctgtaataca 56280  
 gcgcgggtgt ggggggctgg gcgaagctgt ggggtgagtg atcaaagagg ggcctgttga 56340  
 cgagctcga gaagcgatgg gtgatgctgg ggagcagggc cgggtccacc tgggtcgcga 56400  
 ggagagacgc tcgcatgaac cggtgccgct cgaacacgcc cggcgccgag cggttgtcga 56460  
 tgaccgtgcc cgcgcccgcc gtcagggcgc agaagcgcgc gcgcgcccga aagccgttgg 56520  
 cgaccgcggc gaacgtcgcg ggcagcaact cgcctgggac gctgaccgc agcatcttct 56580  
 cgagctcccc gcgctgctc cggacgcagc gccccaggct ggccaacgac cgtctcgtca 56640  
 ggcggtcgc gtacagccgc cgtcgtccc gcacgtccgc ggcctgtgc gtggcgatgt 56700  
 cccccacgt ctggggcccc tgccccccg gcccccgggc acggtcttcg tctcgcgcc 56760  
 cgcgccggg agctcccaac ccccggtccc ctctctctac ggcgacacgg tccccgtgt 56820  
 cgtcggggcc cgcgcccgcc ttggggcgct ccgcccgcgc ccccgcccc atgcgcgcca 56880  
 gcacgcgac cagcgcctcc tcgtgcact gttcggggct gacgaggcgc cgaagagcg 56940  
 gcgtcgtcag gtgggtggtc tagcacgcgc ggatgagcgc ctcgatctga tcgtcgggtg 57000  
 acgtggcctg accgcggatt attagggcgt ccaccatata cagcgcgcc aggtggctcc 57060  
 cgaacgcgcg atcgaaatgc tccgcccgc gcccgaaacag cgcagttcc acggccaccg 57120  
 cggcggctc ctgctgcaac tcgcgcccgc ccagcgcggg caggttgcgt gcaaacgcgt 57180  
 ccatgggtgt ctggccggcg cggtcgcccg acgcgagcca gaatcgcaat tcgctgatgg 57240  
 cgtacaggcc gggcgtggtg gcctgaaaca cgtcgtgcgc ctccagcagg gcgtcggcct 57300

-continued

---

ccttgccgac cgagtcgttc tcgggcgacg ggtggggctg cccgtcgccc cccgcggctc 57360  
 gggccagcgc atggtccaac acggagagcg cccgcgcgcg gtcggcgctc gacagcccgg 57420  
 cggcgtgggg caggtaccgc cgcagctcgt tggcgtccag ccgcacctgc gcctgctggg 57480  
 tgacgtgggt acagatacgg tccgccaggc ggccggcgat cgtcgccccc tggttcgccg 57540  
 tcacacacag ttctcgaana cagaccgcgc aggggtggga cgggtcgcta agctccgggg 57600  
 ggacgataag gcccgacccc accgccccca ccataaactc ccgaacgcgc tccagcgcgg 57660  
 cgggtggcgc gcgcgagggg gtgatgaggt ggcagtagtt tagctgcttt agaaagtctt 57720  
 cgacgtcgtg caggaaacac agctccatat ggacggtccc gccataccta tccagcctga 57780  
 cccgttggtg atacggacag ggtcgggcca ggcccatggt ctccgtgaaa aacgcgcgca 57840  
 cgtctcccgc ggtcgcgaac gtctccaggc tgcaccaggc ccgctcgccc tcgcccacag 57900  
 cgtactctag cagcaactcc agggtgaccg acagcggggg gagaaaggcc cggcctgggg 57960  
 cctccaggcc cggcctcaga cgacgccgca gcgcccgcac ctgaagcgcg ttcagcttca 58020  
 gttgggggag cttccccctg ccgatgtggg ggtcgcaccg ccggagcagc tctatctgaa 58080  
 acacataggt ctgcacctgc ccgagcaggc ctaacaactt ttgacgggcc acggtgggct 58140  
 cggacaccgg ggccggccatc tcgcccgcgc gatctgtacc gcggccggag tatgcggtgg 58200  
 accgagcggg tccgtacgct acccggcgct tggctgagcc ccggggctcc cctcttcggg 58260  
 gcgcccctcc gcgggcccgc cgaccggcaa gccgggagtc ggccggcgcg gcgtttctgc 58320  
 tctattccca gacaccgcgg agaggaatca cggcccgcgc agagatatag acacggaaca 58380  
 caaacaagca cggatgtcgt agcaataatt tattttacac acattccccg ccccgcccta 58440  
 ggttccccca ccccccaacc cctcacagca tatccaactg caggtctccc tttttgtcgg 58500  
 ggggcccctc cccaaacggg tcatccccgt ggaacgcccg tttgcccgcg gcaaatgccg 58560  
 gtcccggggc ccccgggccc ccgaacggcg tcgctgtgtc gtcctcgccg ccaaaatccc 58620  
 caaagttaaa cacctccccg gcgttgccga gttggctgac tagggcctcg gcctcgtcgc 58680  
 ccacctccag ggccgcgctc tcgcacctc gcctgtgccc gcctccagg gcacgcgcgg 58740  
 tcagctccat catctctcgt cttaggtact cgtcctccag gagcgcagc cagtctcga 58800  
 tctgcagctg ctgggtgccc ggccccaggc ttttcacggt cgcacgaac acgctactgg 58860  
 cgacggccgc cccgcctcgt gagataatgc cccggagctg ctgcacagc gagctttcgt 58920  
 gcgctcccgc gccgaggtt gaggccgcgc acacaaaccg gccccgggga caggccagga 58980  
 cgaacttgcg ggtgcggtca aaaataagga gcgggcacgc gtttttgccg cccatcaggc 59040  
 tggcccagtt cccggcctga aacacacggt cgttgccggc catgocgtag tacttgctga 59100  
 tgctcaacc caacacgacc atggggcgcg ccgccatgac gggccgcagc aggttgccagc 59160  
 tggcgaacat ggacgtccac gcgcccggat gcgctccac ggcgtccatc agcgcgcggg 59220  
 ccccgccctc caggcccgcg ccgcctcgtg cggaccacgc ggccgcagcc tgcacgctgg 59280  
 ggggacggcg ggaccccgcg atgatggccc taagggtggt gatgaagtat gtcgagtgat 59340  
 cgcagtaccg cagaatctgg tttgccatgt agtacatcgc cagctcgtc acgttggtgg 59400  
 gggccaggtt aataaagttt atcgcgcctg agtccaggga aaacttttta atgaacgcga 59460  
 tggctcagat gtcctcgcgc gacaggagcc ggccgggaaag ctggttgctg tggagggccg 59520  
 tcagaacca ctgcgggttc ggctgggtgg accccggggg cttgccgttg gggagatgg 59580

---

-continued

---

ccgcgtggaa ctgcttcagc agaaagccca gcggtccgag gaggatgtcc acgcgcttgt 59640  
cgggcttctg gtaggcgctc tggaggctgg cgaccgcgc cttggcgcc tcggacgcgt 59700  
tgggcctcgc gcccgcaac aacacggcgc tcttgacgcg cagctccttg gaaacccca 59760  
gggtcacgcg ggcaacgtcg cctcgaagc tctctcggc gggggcgcgc tggccggcgc 59820  
ttaggctggg ggcgcagata gccgcccct ccgagagcgc gaccgtcagc gttttggccg 59880  
acagaaacc gttgttaaac atgtccatca cgcgcgcgcg cagcaccggt tggaaattgat 59940  
tgcgaaagt gcgcccctc accgactgcc cggcgaacac cccgtggcac tgactcaggg 60000  
ccaggtcctg gtacacggcg aggttggatc gccgcccag aagctgaagc agggggcacg 60060  
gcccgcacgc gtacgggtcc agcgtcagg acatggcgtg gttggcctc cccagaccgt 60120  
cgcgaaaact gaagttcctc cctccacca ggttgcgcat cagctgctcc acctcgcgt 60180  
ccacgacctg cctgaegtgt ttcaccaccg tatgcaggc ctcgcggtg gtgatgatgg 60240  
tctccagccg ccccatggcc gtggggacc cctggtccac gtactgcagg gtctcgagtt 60300  
cggccatgac gcgctcggc gccgcgcggt acgtctcctg catgatggc cgggcggtct 60360  
cggatccgc cgcgccttc agggccgaga aggcggcgt gtttccagc acgtcgcagt 60420  
cgctgtacat gctgttcacg gtccgaaga cgcgatggc tccgcggcg gcgctggcga 60480  
actttgatg gcgcgcccgg aggcgcacga gcgtcgtgt tacgcaggcg tggcgcgtgt 60540  
cgaaggtgca taggttacag ggcacgtcgg tctggttgg gtcgcgcagc taccgaaaca 60600  
cgtccatctc ctgggcgcc acgatcacgg cgcctcgca gcgctccagg taaaacagca 60660  
tcttggccag cagcgcggg gaaaaccac acagcatggc caggtgctc cggcaaat 60720  
cctgggttcc gccgcagag ggcgcggtg gccgaccctc gaaccgggc accacgtgc 60780  
cctcgcggtc cacctgtggg ttggcccca cgtgggtccc gggcacgagc aagaagcgt 60840  
aaaaggaggg tttgctgtg tccttgggt ccgcccggcc ggcgtcgtcc acctcgggtga 60900  
gatggagggc cgagttggtg ctaataacca tggccccac gagtcccgc gcgcgcgcca 60960  
ggtaaccccc gacggcgtg gcgcgggccc cggccgtgct ctggccctc aacagcggcc 61020  
acgcggagat gtcggtggg ggctcgtcaa agacggccat cgacacgata gactcagggg 61080  
ccaggcggc gtctccggcc atgacggagg ccaggcgtg ttcgaaccgc cccgcgcgc 61140  
ccttgcccgc gccgtcgcgc cgcgccgcg gggctctacc ctggctggct tcgaaggcgc 61200  
tgaacgtaat gtcggcggg agggcggcgc cctcgtggt ttcgtcaaac gccaggtggg 61260  
cggccgcgcg ggccaacggc tccacgttc gccatcgcag tgccaacggc gcgggtccca 61320  
cgaccgcctc gaacaggagg cgggtgagg ggcggttaa aaacggaagc gggtaggtaa 61380  
atttctccc gatcgatcgc tggttggcgt tgaacggctc tgcgatgaca cggctaaaat 61440  
ccgcatgaa cagctgcaac gggtagacgg gtatgcggtg cacctccgc ccgcctatgg 61500  
ttacctgtc cgagcctccc aggtgcagaa aggtgttgt gatgcacac gcctcctga 61560  
agcctcggg aacgaccaga tacaggagg cgcggtcgcg gtccaggcgc aggcgctcac 61620  
acagcgcctc ccccgctgc tctgtttga ggtcgcggg ccgggggggtg tagtccgaaa 61680  
agccaaaatg gcggcgtgccc cgtcgcgaaa gtcgcgtcag gttcggggcc tgggtgctgg 61740  
ggtccagggt ccggccgcgc tgaagacgt acacggacga gctgtagtgc gagggcgtca 61800  
gttccaggga caccgcggtc ccccagacc ccgtcgtcgc agaaccacg accacggcca 61860

-continued

---

cgttggcctc aaagccgctc tccacggtca ggcccacgac caggggcgcc acggcgacgt 61920  
 cggaatcgcc gctgcgtgcc gacagtaacg ccagaagctc gatgccttcg gacggacacg 61980  
 cgcgagcgta cacgtatccc aggggcccgg gggggacctt gatggtggtt gccgtcttgg 62040  
 gctttgtctc catgtccttt tgtaaatcgg tccgcgaacg gaggtaatcc cggcacgacg 62100  
 acggacgccc gacaaggtat gtctcccagc cgtcaaaatc cggggggggg cggcgacggt 62160  
 caaggggagg gttggagacc ggggttgggg aatgaatccc tccccttcac cgacaacccc 62220  
 cgggtaacc acggggtcgc cgtatgaacc cggcgccggg caacgcgggg tccctgcgag 62280  
 aggcacagat gcttacggtc aggtgctccg ggtcgggtgc gtctggtatg cggttggtat 62340  
 atgtacactt tacctggggg cgtgccggtc cggcccagcc cctcccacgc cccgcgcgtc 62400  
 atcagccggg gggcgtggcc gctattataa aaaaagtgag aacgcgaagc gttcgcactt 62460  
 tgtcctaata atatatatat tattaggaca aagtgcgaac gcttcgcggtt ctaacttttt 62520  
 ttataatagc ggccacgccc accggctacg tcaactctct gtcggccgcc ggcggtccat 62580  
 aagcccggcc ggcggggcgc acgcgaataa accgggcccgc cggccggggc gccgcgcagc 62640  
 agtcgcgcgc ccggatccgc cagacaaaca aggccttgc acatgccggc ccgggcgagc 62700  
 ctgggggtcc ggtaatTTTT ccatcccacc caagcggctt tttgggtttt tctctcccc 62760  
 cctcccaca tccccctct taggggttcg ggtgggaaca accgcgatgt tttccgggtg 62820  
 cggcggcccgc ctgtccccg gaggaaagtc ggcggccagg gcggcgtccg ggttttttgc 62880  
 gcccgccggc cctcggggg ccagccgggg acccccgcct tgtttgaggc aaaactttta 62940  
 caacccctac ctgccccag tcgggacgca acagaagccg accgggcca cccagcgcga 63000  
 tacgtactat agcgaatgag atgaatttcg attcatcgcc ccgcggtg gcgacgagga 63060  
 tgccccccgc gagaagcgcg ccggggtgca cgaacgtcac ctcaagcgcg ccccgaaggt 63120  
 gtactgcggg ggggacgagc gcgacgtcct ccgcgtcggg tcggcggct tctggccgcg 63180  
 gcgctcgcgc ctgtgggggc gcgtggacca cggcccggcg gggttcaacc ccaccgtcac 63240  
 cgtctttcac gtgtacgaca tccctggagaa cgtggagcac gcgtacgca tcgcgcgggc 63300  
 ccagttccac gcgcggttta tggacgccat cacaccgacg gggaccgtca tcacgtcct 63360  
 gggcctgact ccggaaggcc accgggtggc cgttcacgtt tacggcacgc ggcagtactt 63420  
 ttacatgaac aaggaggagg tcgacaggca cctacaatgc cgcgccccac gagatctctg 63480  
 cgagcgcagc gcccgggccc tgcgcgagtc cccgggcgcg tcgttccgcg gcattccgc 63540  
 ggaccacttc gaggcggagg tgggtggagc caaccgacgt tactactacg agacgcgcc 63600  
 cgctctgttt tacccgctct acgtccgaag cggcgtgtg ctgtcgtacc tgtgcgacaa 63660  
 cttctgccc gccaacaaga agtacgaggg tggggtcgac gccaccacc ggttcatcct 63720  
 ggacaacccc gggttcgtca ccttcggctg gtaccgtctc aaaccgggcc ggaacaacac 63780  
 gctagcccag ccggcggccc cgtatggcct cgggacatcc agcgcgctc agtttaactg 63840  
 tacggcggac aacctggcca tcgagggggg catgagcgac ctaccggcat acaagctcat 63900  
 gtgcttcgat atcgaatgca aggcgggggg ggaggacgag ctggccttc cgttgccgcg 63960  
 gcacccggag gaacctgtca tccagatata ctgtctgctc tacgacctgt ccaccaccgc 64020  
 cctggagcac gtctctctgt tttcgtcgg ttccctgcgac ctcccgaat ccacactgaa 64080  
 cgagctggcg gccaggggcc tgcccacgcc cgtggttctg gaattcgaca gcgaattcga 64140

-continued

---

gatgctgttg gccttcatga cccttgtgaa acagtacggc cccgagttcg tgaccgggta 64200  
caacatcadc aacttcgact ggcccttctt gctggccaag ctgacggaca tttacaaggt 64260  
ccccctggac gggtagcggc gcatgaaacgg ccggggcgtg tttcgcgtgt gggacatagg 64320  
ccagagccac ttcagaagc gcagcaagat aaaggtgaac ggcatggtga acatcgacat 64380  
gtacgggatt ataaccgaca agatcaagct ctcgagctac aagctcaacg ccgtggccga 64440  
agccgtcctg aaggacaaga agaaggacct gagctatcgc gacatccccc cctactacgc 64500  
cgccgggccc gcgcaacgcy gggtagcggc cgagtactgc atacaggatt cctgctggt 64560  
gggacagctg ttttttaagt ttttgcccca tctggagctc tcggccgtcg cgcgcttgcc 64620  
gggtattaac atcaccgcga ccatctacga cggccagcag atccgcgtct ttactgtcct 64680  
gctgcgcctg gccaccaga agggctttat tctgcccggc acccaggggc gatttagggg 64740  
cgccgggggg gaggcgcaca agcgtccggc cgcagcccgg gaggacgagg agcggccaga 64800  
ggaggagggg gaggacgagg acgaacgcga ggaggcggg ggcgagcggg agccggaggg 64860  
cgccgggggg accgcccggc ggcacgtggg gtaccagggg gccagggtcc ttgacccccc 64920  
ttccgggttt cacgtgaacc ccgtggtggt gttcgacttt gccagcctgt accccagcat 64980  
catccaggcc cacaacctgt gcttcagcac gctctccctg agggccgacg cagtggcgca 65040  
cctggaggcg ggcaaggact acctggagat cgagggtgggg gggcgacggc tgttcttctg 65100  
caaggctcac gtgcgagaga gcctcctcag catcctcctg cgggactggc tcgccatcgc 65160  
aaagcagatc cgctcgcgga ttccccagag cagccccgag gaggccgtgc tccggacaaa 65220  
gcagcaggcc gccatcaagg tcgtgtgtaa ctcggtgtac gggttcacgg gactgcagca 65280  
cggactcctg ccgtgcctgc acgttgcccg gacggtgacg accatcggcc gcgagatgct 65340  
gctcgcgacc cgcgagtacg tccacgcgcg ctgggcggcc ttcgaacagc tccgtgcccga 65400  
tttcccggag gcggccgaca tgccgcgccc cgggcccctat tccatgcgca tcatctacgg 65460  
ggacacggac tccatctttg tgctgtgccc cgccctcacg gccgcccggc tgacggccct 65520  
ggcgacaag atggcgagcc acatctcgcg cgcgctgttt ctgcccccca tcaaaactcga 65580  
gtgcgaaaag acgttccaca agctgctgct gatcgccaag aaaaagtaca tcggcgctcat 65640  
ctacgggggt aagatgctca tcaaggcgtg ggatctggtg cgcaaaaaca actgcgcgctt 65700  
tataaacccg acctccaggg ccctggtcga cctgctgttt tacgacgata ccgtctccgg 65760  
agcggcccgc gcgttagccg agcgcccccg ggaggagtgg ctggcgcgac cctgcccga 65820  
gggactgcag gcgttcgggg ccgtcctcgt agacgcccac cggcgcatca ccgaccggga 65880  
gagggacatc caggactttg tccctaccgc cgaactgagc agacaccgcg gcgctacac 65940  
caacaagcgc ctggcccacc tgacgggtga ttacaagctc atggcccggc gcgcgaggt 66000  
cccgtccatc aaggaccgga tcccgtacgt gatcgtggcc cagaccggcg aggtagagga 66060  
gacggtcgcg cggtggcccg cctccgcgca gctagacgcc gccgcccag gggacgagcc 66120  
cgcccccccc gcggccctgc cctcccggcg caagcgcgcc cgggagacgc cgtcgcctgc 66180  
cgaccccccg ggaggcgcgt ccaagccccg caagctgctg gtgtccgagc tggccgagga 66240  
tcccgcatac gccattgccc acggcgtcgc cctgaacacg gactattact tctcccacct 66300  
gttggggggc gcgtgctgta cattcaagc cctgtttggg aataacgcca agatcaccga 66360  
gagtctgtta aaaaggttta ttcccgaagt gtggcacccc ccggacgacg tggccgcccg 66420

-continued

---

gctccgacc gcagggttcg gggcggtggg tgccggcgct acggcggagg aaactcgtcg 66480  
aatgttgc atagcctttg atactctagc atgagcccc cgtcgaagct gatgtccctc 66540  
atthtacaat aaatgtctgc ggccgacacg gtcggaatct ccgctcctcg gggtttctct 66600  
gcggtgcgcc ggaccacgag cacaaacgtg ctctgccaca cgtgggcgac gaaccgggtac 66660  
cccgggacg cggtagcat ccggtctatg agccggtagt gcagggtggc ggacgtgccg 66720  
ggaaagatga cgtacagcat gtggccccg taagtggggt ccgggtaaaa caacagccgc 66780  
gggtcgacg ccccgctcc gcgcaggatc gtgtggacga aaaaaagctc gggttggcca 66840  
agaatcccgg ccaagaggtc ctggaggggg gcggttggtg ggtcggccaa caccaccaag 66900  
gaggccagga aggcgcatg ctccaatcgc gtgttgatct gctgcacgaa ggccaggatt 66960  
agggcctcgc ggtggtggc ggcgaaccgc ccgtctccc cgttgacgcg gggacagcaa 67020  
ccccgatgc ctaggtagta gcccatccc gagagggtea ggcagttgct ggccacggtc 67080  
tggtccagac agaagggcag cgagacggga gtggtcttca ccaggggac cgagagcgag 67140  
cgcacgatgg cgtctcctc ggagggcgct tgggcgaggg cggcgaaaag gccccgatag 67200  
cgtggcgct cgtgtaaca cagctcctgt ttgcggcgct gaggcggcag gctctcccg 67260  
gaggcccgc gcaccacgcc cagagtcgcc ccggccgcag aggagcgcga ccgcccggc 67320  
tccttgccgt gatagggccc gggccgggag ccgcccgcgt gggggtcggt gtcatacata 67380  
ggtacacagg gtgtgctcca gggcagggag cgagatcgag tggcgtctaa gcagcgcgcc 67440  
cgcctcagc acaaatgtgg cgagcgcggt gggctttggt acaatacct gatacgtctt 67500  
gaaggtgtag atgagggcac gcaacgctat gcagacacgc ccctcgaact cgttcccga 67560  
ggccagctt gccttgtaga gcagcagctc gtcgggatgg gtggcggggg gatggccgaa 67620  
cagaaccag ggtcaacct ccctcctcgt aatggcgcac atggggtcac agaactgtg 67680  
ctaaagatg gcctcggggc ccgcccggc aagcaggctc acaaacggc ccccgcccc 67740  
gggctgcgc tcggggtcag cctcagcgtg gtcgacgac ggtacgatac agtcgaagag 67800  
gctcgtgtt tttccagat agcggaccac ggaggcccgc agtctgcgca gggccagcca 67860  
gtaagcagc accagtaaca ggttacacag caggcattct ccgcccgtgc gcccgcccc 67920  
ccggccgtgt ttcagcacg tggccatcag agggcccagg tcgaggtcgg gctgggcatc 67980  
gggttcggta aactgcgcaa agcgcggagc cacgtcgcgc gtgcgtgccc cgcgatgcgc 68040  
tcccaggac tggcggacc tggcgcgac ggcctccgc gcagcgcga gctggggccc 68100  
cgactcccag acggcggggg tgcccggcag gagcaacagg accagatccg cgtacgccc 68160  
cgtatccgc gactcctcc gctcggggtc cccggcgacc gtctcgaatt ccccgttgcg 68220  
agcggcggcg cgagtacag agctgtcccc gccccgcgc cgaccctccg tgcagtccag 68280  
gagacgggcg caatccttc agttcatcag cgcggtggtg agcagcggct gctgcccga 68340  
tcccgcgcc gaccccgccc cctcctcgc cccggaggcc aaggttccga tgagggcccg 68400  
ggtggcagc tgcgccagga acgagtagtt ggagtactgc acctggcgg ctcggggga 68460  
gggcgagggc ttgggtgct tctgggcatg ccgcccggc accccgcgt cggtacggaa 68520  
gcagcagtg agaaaaagt gccggtggat gtcgtttatg gtgagggcaa agcgtgcgaa 68580  
ggagccgacc agggctcct tcttggtgcg cagaaagtgg cggtcacatga cgtacacaaa 68640  
ctcgaacgc gccacgaaga tgctagcggc gcagtggggc gccccaggc atttggcaca 68700

-continued

---

gagaaacgcg taatcggcca cccactgagg cgagaggcgg taggtttgct tgtacagctc	68760
gatggtgcgg cagaccagac agggccggtc cagcgcgaag gtgtcgatgg ccgcccgga	68820
aaagggcccg gtgtccaaaa gccccctccc acagggatcc gggggcgggt tgcggggctc	68880
tccgcgcccg ccgaacccc ctccgtcgcc cgccecccgc cgggcccttg agggggcgggt	68940
gaccacgtcg gcggcgacgt cctcgtcgag cgtaccgacg ggcggcacac ctatcacgtg	69000
actggccgtc aggagctcgg cgcagagagc ctctgtaaga gccaggaggc tgggatcgaa	69060
ggccacatac gcgcgctcga acgccccgc ctccagctg ctgcccgggg actcttcgca	69120
caccgcgacg ctgcaccagg ccccgggggg cgaagtgtcc atggtgggc gggaggggcg	69180
cacgcgccag cgaactttac gggacacaat ccccactg cgcgtgcgggt ccagaccct	69240
ggagagtcta gacgcgcgct acgtctcgcg agacggcgcg catgacgcgg ccgtctggtt	69300
cgaggatatg acccccgcgc agctggagggt tgtcttccc actacggacg ccaagctgaa	69360
ctacctgtcg cggacgcagc ggctggcctc cctcctgacg tacgccgggc ctataaaagc	69420
gcccgcgacg gccgcgcgcc cgcagacccc ggacaccgcg tgtgtgcacg gcgagctgct	69480
cgccccaag cgggaaagat tgcggcgggt cattaaccgg ttcttgacc tgcaccagat	69540
tctcggggc tgacgcgct cctgttgggc gggacggctc gcgaaccctt tgggtgggtt	69600
acgcgggca cgcacgtccc atcgcggggc ccatggcggg actgggcaag cctacaccg	69660
gccaccagg tgacgccttc gagggtctcg ttcagcgaat tcggttacc gtccatcta	69720
cgttgcgggg cggggacggg gaggcgggcc cctactctcc ctccagctc cctccaggt	69780
gogcctttca gtttcatggc catgacgggt ccgacgagtc gtttccatc gagtatgtac	69840
tgcggcttat gaacgactgg gccgaggctc cgtgcaacc ttacctgcgc atacagaaca	69900
ccggcgtgct ggtgctggtt caggggtttt ttcacgccc acacaacgcc cccggggggc	69960
cgattacgcc agagcggacc aatgtgatcc tgggtccac cgagacgacg gggctgtccc	70020
tccggacact ggacaccatc aagggggcgc tcggcctgga tgcccggccg atgatggcca	70080
gcattgtgat cagctgcttt gtgcgcatgc cccgcgtgca gctcgcgttt cggttcatgg	70140
gccccgaaga tgcggacgag acgagacgga tctgtgcgc gcgcccgag caggctatta	70200
cccgtcgcgc ccgaaccccg cggctcccgg aggcgtacgg ggcgcaggcc gggctggggg	70260
tggccggaac gggtttccgc gccagggggg acggttttgg cccgctccc ttgttaaccc	70320
aagggccctc ccgcccggtg caccaggccc tgcggggtct taagcaccta cggattggcc	70380
ccccgcgct cgttttggcg gcgggactcg tctgggggc cgtatttg tgggtgggtg	70440
gtgctggcgc gcgcctataa aaaaggacgc accgcgcgcc taatcgccag tgcgttccgg	70500
acgccttcgc cccacacagc cctcccagc gacacccca taccgcttc cgacctccgg	70560
tcccgatggc cgtcccgcaa tttcaccgcc ccagcaccgt taccaccgat agcgtccggg	70620
cgttggcat gcgcgggctc gtcttgcca ccaataactc tcagtttacc atggataaca	70680
accaccgca cccccaggc acccaagggg cegtgcggga gtttctccgc ggtcaggcgg	70740
cggcgtgac ggaccttgt ctggcccacg caaacaacac gtttaccgc cagcctatgt	70800
tgcggggcga cgccecgcc cctcgggtgc ggccecgctt tggcctgcgc cgcacctatt	70860
caccgtttgt cgttcgagaa ccttcgacgc cggggacccc gtgaggcccg gggagttcct	70920
tctggggtgt tttaatcaat aaaagaccac accaacgcac gagccttgcg tttaatgtcg	70980

-continued

---

tgttattca agggagtggg atagggttcg acggttcgaa acttaacaca ccaataatc	71040
gagcgcgtct agcccagtaa catgcgcacg tgatgtaggc tggtcagcac ggcgtcgctg	71100
tgatgaagca gcgcccggcg ggteccgtgt aactgctgtt gtaggcggtg acaggcgcgg	71160
atcagtaccg ccaggcgcgt acgaccgggt cgttgcaegt agcgtcgcga cagaactgcg	71220
tttgccgata cgggccccggg gccgaattgt aagcgcgtca cctcttggga gtcacggtcg	71280
gataacgcac tgaatgggtc gttggttatg ggggagtgtg gttccccagg gagtgggtcg	71340
agcgcctcgg cctcggaaac cgagaggaac aacgaggtgg cgtcggagtc ttcgtcgtca	71400
gagacataca gggctctgaag cagcgacacg ggcggggggg tagcgtcgat gtgtagcgcg	71460
agggaggatg cccacgaaga caccocagac aaggagctgc ccgtgcgtgg atttgtggaa	71520
gacgcggaag ccgggacgga tgggcggtt tgccgtgccc ggaaccgaac cgcgggatac	71580
tccccgggtg ctacatgccc gttttggggc tggggttggg gctggggttg gggctggggt	71640
tggggctggg gttggggctg gggttggggc tggggttggg gttggggttg gggctggggt	71700
tggggttggg gctggggctg gggctggggc tgggctggg gctggggctg gggctggggc	71760
tggggctggg gctggggctg gggctggggc tgggctggg gctggggttg gggcgcggac	71820
agggcgtgta cggtaaaatg ccccggggg cgcgcagatg tgggtggcgt ggccaccggc	71880
tgccgtgtag tggggcggcg ggaaccggg cctccggggc caacaccgcc ctccagcgtc	71940
aagtatgtgg ggggcggggc tgacgtcggg ggcggggcga cgggttggac cgcgggaggc	72000
gggggagagg gacctcggg agaggatgag gtcggctcgg ccgggttgcg gcctaaaaca	72060
ggggccgtgg ggtcggcggg gtcccagggt gaagggaggg attcccgcga ttcggacagc	72120
gacgcgcagc cggggcgcgt aaggcgcgcg tcggcccgc ctacgggaac cctggggggg	72180
gttggcgcgg gaccgaggt tagcgggggg cggcgtttt cgccccggg caaaaccgtg	72240
ccggttgcga ccggggggcg aacgggatcg atagggagag cgggagaagc ctggccggcg	72300
gcctggggcc cgagcgggag gggcacacca gacaccaaag cgtggggcgc tggctctggg	72360
ggtttgggag gggccggggg gcgcgcgaaa tcggtaaccc gggcgaccgt gtcggggagg	72420
gcaggcggcc gccaaacctg ggtggtcgcg gaagcctggg tggcgcgcgc cagggagcgt	72480
gcccggcggg gtcggcgcgc gcgcgacccg gacgaagaag cggcagaagc gcgggaggag	72540
gcgggggggc gggggcggtt ggcacgggg ggcgcgggg aactttgggg ggcggcgaag	72600
cgccggaagt cgtcgcgggg gcccacgggc gccggcgcg tgccttcggc cgggacgccc	72660
ggtcgtgctt cgcgagccgt gactgcccgc ccaggggggc gcggtgcaca ctgggacgtg	72720
gggacggact gatcggcggg gggcgaagg gggccgggg caaggagggg cgcggggcgg	72780
ccggagtcgt cagacgcgag ctctccagg ccgtgaatcc atgccacat gcgagggggg	72840
acgggctcgc cgggggtggt gtcgggtgaat agcgtggggg ccaggcttcc gggccccaac	72900
gagccctccg ccccaacaag gtccacaggg ccgggggtcg ggtttgggac cgaggggctc	72960
tggctcgtcg gggcgcgctg gtacaccgga tgccccggga atagctcccc cgacaggagg	73020
gaggcgtcga acgcccggc gaggatagct cgcgcgagga aggggtcctc gtcggtggcg	73080
ctggcggcga ggacgtcctc gcccccggc aaaaacggga gctcctcggg ggcctcgtg	73140
ccaaacaaac gcacgtcggg ggggcgggg gggccgggt tttccacaa caccgcgacc	73200
ggggtcatgg agatgtccac gacaccaga cacggcgggc cccgggcgag gggccgctcg	73260

---

-continued

---

gcgatgagcg cggacaggcg cgggagctgt gccgccagac acgcgttttc gatcgggttc 73320  
aggtcggcgt gcaggaggcg gacggcccac gtctcgatgt cggacgacac ggcacgcgc 73380  
aaggcggcgt ccggcccgcg agcgcgtgag tcaaacagcg tgagacacag ctccagctcc 73440  
gactcgcggg aaaaggccgt ggtgttcgcg agcgcacga cgcggggcgc gcccaggagc 73500  
actgcccca gcaccaggtc catggccgta acgcgcgcg cgggggtgcg gtgggtggcg 73560  
gcgcccggca cggcgacgtg ctggcccgtg ggccggtaga gggcgttggg gggagcgggg 73620  
ggtgacgcct cgcgcccccc cggggggctc agcgtctgcc cagattccag acgcgcggtc 73680  
agaaggcgt cgaactgtc atactctgtg tagtctccg gaaacatgca ggtccaaaga 73740  
gcgaccagag cgtgtcttgg gagacacatg cgcgcgagga cgctcaccgc cgcagcgc 73800  
tgggcgggac tcagctttcc cagcgcggcg ccgcgctcgg ttcccagctc ggggaccgag 73860  
cgccaggcgc ccagggggtc ggtttcggac aacttgccgc ggcgccagtc tgcagccgc 73920  
gtgccgaaca tgaggccccg ggtcggaggg cctccggccg aaaaocgtgg cagcacgcgg 73980  
atcggggcgt ctggatgccc ggtcaggcgc tgcacgaata gcattggaatc tctgctgttc 74040  
tgaaacgcac gggggagggt gagatgcatg tactcgtgtt ggcgaaccag atccaggcgc 74100  
caaaaggtgt aaatgtgttc cggggagctg gccaccagcg ccaaccagcag gtcgttctcg 74160  
ttaaaggaaa cgcggtgcct agtggagctc tggggtccga gcggcggccc cggggccgcc 74220  
gcgtcaccac cccattccag ctggcccag cgcacccaa actcgcgcgt gagagtgttc 74280  
gcgacgaggg cgcgtagatg ctgcgcccgc gcacccatcg agggccccca tctgcctgg 74340  
cgggtggcga caaagcgtcc gaagagctga aagttggcgg cctgggcgctc gctgagggcc 74400  
agctgaaacc ggttgatgac ggtgaggacg tacatggccg tgacggtcga gcccgactcc 74460  
aggggtgtcc tcggaagcgg ggggcgaatg catgcccctc cgggacacat cagcagcgcg 74520  
ccgagcttgt cgttcacggc cgggaagcag agcgcgtact gcagtggcgt tccatccggg 74580  
acaaaaaacg tggggcgcaa cggccgatcc agcgtactgg tggcctcgcg cagcaccagg 74640  
ggccccgggc ctccgctac tcgcaggtag gcctcgcgcc ggcggcgcag catctcgggg 74700  
tcggcctctt ggcgggtgg ggcggacgccc cgggcgcggg cgtctagggc gcgaagatec 74760  
acgagcaggg gcgcgggcgc ggcgcgcgcg cccgcgcccg tctggcctgt gcccttggcg 74820  
tacgcctat ataagcccat gcggcgttgg atgagctccc gcgcgcccgc gaaactcctcc 74880  
accgcccatg gggccaggtc cccggccacc gcgtcgaatt ccgccaacag gccccccagg 74940  
gtgtcaaagt tcactctcca ggcaccctt ggcaccacct cgtcccgcag ccgggcgctc 75000  
aggtcggcgt gttgggccac gcgcccccg agctcctcca cggccccggc ccgctcggcg 75060  
ctcttggcgc ccaggacgcc ctggtacttg gcgggaaggc gctcgtagtc ccgctgggct 75120  
cgcagccccg acacagtgtt ggtgggtgccc tgcagggcgc gaagctgctc gcacgcgcg 75180  
cgaaatccct cgggcgattt ccaggcccc cgcggaacgc ggccgaagcg accccatacc 75240  
tcgtcccact ccgcctcggc ctctcgcaga gacctccgca gggcctcgcg cgggcgacgg 75300  
gtgtcgaaga gcgcctgcag gcgcgcgccc tgtcgcgtca ggaggcccgg gccgtcgcg 75360  
ctggccgcgc ttagcgggtg cgtctcaaag gtacgctggg catgttccaa ccaggcgacc 75420  
gcctgcacgt cgagctcgcg gcctctctcc gtctggtcca ccagaatttc gacctgatcc 75480  
gcatctctcc ccgcccagcg gcctggtccc agcgtcttgg ccaaggctgc cgggacggcg 75540

-continued

---

accaccttca gcagggtctt cagattggcc agaccctcgg cctcagctg ggccccgcgc	75600
tcgcgcgcgg ccagcacctc ccgcagcccc gccgtgacct gctcgggtggc ttccggcgcgc	75660
tgctgtttgg cgcgcaccac ggcgtccttg gtatcggcca ggtcctgtcg ggtcacgaat	75720
gcgacgtagt gcgcgtaecg cgtgtccttc acggggctct ggtccacgcg ctccagcgc	75780
gccacgcacg ccaccagcgc gtccctcgtc gggcagggca gggtagcccc tgccccgaca	75840
agctcggcgg ccgcgcgcgg gtctgttcgc acccgggata tctcctcgcg ggccggcggc	75900
aggtccagcg ccacgcttcc gatcgcgcgc cgcgcgtcgg cccggagggc gtccaggcga	75960
tcgcggatat ccacgtactc ggcgtagccc ttttgaaaaa acggcacgta ctggcgcagg	76020
gccggcacgc cccccaaagtc ttccgacagg ttaggacgg cctcgtggta gtcgataaac	76080
ccgtcgttcg cctgggcccc ctccagcagc cccccgcca gccgcagaag ccgcgccagg	76140
ggctcggtgt ccaccgaaa catgtcggcg tacgtgtcgg ccgcggcccc gaaggccgcg	76200
ctccagtcga tcgggtgaat ggctgcgagc ggggggagca tgggggtggcg ctggttctcg	76260
ggggtgtatg ggttaaacgc aaggccctgc tccagggcaa gggtcaccgc cttggcgttg	76320
gttccacgcg cctgttcggc ccgttttcgg aagtccccgg ggtgtagcc gtgcgtgcc	76380
gccagcgctc gcaggcgacg gagctcgacc acgtcaaaact cggcacccgt ttccacgcgg	76440
tccagcacgg cctccacgtc ggcggcccag cgtcgtggc tactcggggc gcctggggcc	76500
gccatcttct ctctgaggtc ggcgggtggc gcctcaagtt cgtcggcgcg gcctcgcgtg	76560
gcgcgatga ctttccccc ctctcgcagg gcgcgcgcgc tgggggagtg gtccccggcc	76620
gtcccttcgg cgtgcaacag gccccgaac ctgccctcgt ggccccgcgag gctttccccg	76680
gcgcgggtgg tcgcgcgcgt cgcggcctgg atcagggagg catgctctcc ctccggttg	76740
ttggcggccc ggcgcacctg gacgacaagg tcggcggcag ccgacctaa ggtcgtgagc	76800
tgggcgatgg ccaccgcgcg gtccagggcc aaccgagtcg ccttgacgta tcccgcggcg	76860
ctgtcggcca tggccgctag gaaggccagg ggggaggccc ggtcgtggc ggccgcgcgc	76920
agggccgcca ccgcgtcagc caggacgcgg tcgcgccgca cggccgcac caccgtcagc	76980
gcagggtctg ccgttcgcag ggcggcgtcg ccggcgttga tggcgttcga gacggcgtgg	77040
gctatgatcg gggcgtgatc ggcgaagaac tgcaagagaa accgagctc tggggcgtcg	77100
gcgaacaggt tcttcagcac caccacgaag ctgggatgca agccagacag agccgtcgc	77160
gtgtccggag tcgggtgctc caggcatct cggtactgcc ccagcagccc ccacatgtcc	77220
gccccagcgc ccgcgtaac ctccaggggc gcccccgaa cggcctcggg gaggtccgac	77280
cagcccgcg gcagggaggc ccgcagggtc gccaggacgg ccggacaggc ctttagcccc	77340
acaaagtcag ggagggggcg caggaccccc tggagtttgt gcaagaact ctccccggcg	77400
tcgcgggcca ctttcgccc ctccccgcgt cctcagagca ttgcctccag ggagcgcgcg	77460
cgtccccgca aacgggcacg cgcacgggg gcgagctctg ccgtcagctt ggccgcatcc	77520
atggccccgc cctgcgcag cgttctctcg gccatgcgcg tggcctctgg cagacgccc	77580
ccgtcgtcgg ggtaggggcga cgcgccgggc gcaggaacaa aggcgcgctc gctgtccagc	77640
tgctggccca gggccgcatc tagggcgtcg aagcgcgcga gctcggccag acccagctg	77700
cggcgcgctc gttggtcgtt aatgtcgcgg atgctgcgcg ccagctcgtc cagtggcttg	77760
cgttctatca gcccttggtt ggcggcgtcc gtcaggacgg agagccaggc cgcagggtcc	77820

-continued

---

tcggggcgct ccagcgtctg gccccgctgg atcagatccc gcaacaggat ggccgtgggg 77880
ctggtcgcga tcggggggcg ggcgggaatg gcggcgcgct gcgcgatgtc ccgctgtgtc 77940
tggtcgaaga caggcaggga ctcgagcagc tggaccacgg gcaecagcgg gccccaagcc 78000
acgtgaaacc ggcggtcggt gttgtcgtg gcctgtagag ccttggcgct gtatacggcc 78060
ccccggtaaa agtactcctt aaccgcgccc tcgatcgccc gacgggcctg ggtccgcacc 78120
tcctccagcc gaacctgaac ggcctcgggg cccagggggg gtggcgcgcg agccccctgc 78180
ggggccgccc cggccggggc gggcattacg ccgagggggc cggcgtgctg tgagaccgcy 78240
tcgaccccg c gagcgagggc gtcgagggcc tcgcgcactc ggcgatcctc cgcctccacc 78300
ctaactcttt cggccagggc aaatttggcc agagcctgga ctctatacag aagcggttct 78360
gggtgcgtcg ggttggcggg ggcaaaaagg gtgtccgggt gggcctgcga gcgctccaga 78420
agccactcgc cagggcggtg atacagattg gccggcgggg ccgcgcgaag ctgcagctcc 78480
aggtccgcga gttccccgta aaaggcgtcc gtctcccga tgacatccct agccacaagg 78540
atcagcttcg ccagcggccg cgcagccgac agagagtttt cgtccagcac gtgctggagc 78600
agggcgagat gggcgccac gtcggccagg ctcagcgcg tggaggccag aaagtcccc 78660
acggccggtt tccagggcag catgttcagg gtaaaccca gcaggcgccg gcccgggccg 78720
gccaccccg cctgggtgtg cgtccgggcc ccgttctcga tgagaaaagg gaggaacgct 78780
tcaaaaaaa aaataacaca gagctccagc agccccggag agggccgata cggcgaccgt 78840
aaggcgctga tggtagccg cgaacacggc gcgacctcgc gggccagggc gccggagcac 78900
gcggtgaact taaccgccgt ggcggccacg tttgggtggg cctcgaacag ctggcgaggg 78960
tctgcgccc ggggctcggg cgagcggcga gtcttcagcg cctcaggggc ctgtgaggac 79020
gccggaaccg tgggcccgtc gtctcggccc gcctcggcga ccggcgccc gcccggtcg 79080
gggggtgccc aggcgaggac aggtcccgga acggaggcgg ggaccgcggc cccgacgggg 79140
gttttgctt tgggggtgga tttctcttg gttttggcag ggggggcca gcgtttcgtt 79200
ttctccccg aagtcaggtc ttcgacgctg gaaggcggag tccaggtggg tcggcgggcg 79260
ttgggaaggc cggccagta gcgtgcccgg tgcgaccaa ccgggacgac gcccatctcc 79320
aggacccgca tgtcgtcgtc atcttcttcg gccgcctctg ccgggggggt cttggggcg 79380
gagggaggcg gtgggtggat cgcggagggt gggtcggcgg agggtgggtc gccggagggg 79440
ggatccgtgg gtgggtacc cttcagggcc accgcccata catcgtcggg cgcgccatc 79500
ggcgcttg cctctggtt tgccgacgga ccggccgtcc cccgggatgt ctcggaggcc 79560
ctgtcgtcgc gacgggccc ggtcgggtgc ggcgactggg cggctgtggg cgggtgtggc 79620
ccccggcccc ctaccccc cggggggccc acgcccagc agggctcccc caggcccgcg 79680
atctcgcgcc gcaggggggt cgtgatggcc acgcgccgtt cgtgaaacg ttegtectgc 79740
aggtaaagtct cgtggcccc gtaaagatgc agagccgcgg ccgtcaagtc cgcaggagcc 79800
gcgggttccc gcccgacgg cagcaaaaac accatggctc ccgccaccg tacgtccggg 79860
cgatcgggg tgtaatacgt caggtatgga tacatgtccc ccgccgcac tttggcgatg 79920
aacgcggggg tgccctcgg aaggccgtgc gggtaaaaa ggtatgcggg gtcgcccgtc 79980
ctgaacagcc ccatccctag ggggccaatg gttaggagcg tgtaacgacag gggcgccagg 80040
gcccacgggc cggcgaagaa cgtgtgtgcy gggcattgtg tctccagcag gcccgccgcg 80100

-continued

---

```

ggctccccga agaagccac ctcgccgtat acgcgcgaga agacacagcg cagtccgccg 80160
cgcgccccctg ggtactcgag gaagttgggg agctcgacga tcgaacacat gcgcggcggc 80220
ccagggcccg cggtcgcgcg cgtccaactcg cccccctcga ccaaacatcc ctcgatggcc 80280
tccgcggaca ggacgtcgcg agggcccaca tcaaatatga ggctgagaaa ggacagcgcg 80340
gagcgcgatgc acgataccga cccccccggc tccaggtcgg gcgcgaactg gttccgagca 80400
ccggtgacca cgatgtcgcg atcccccccg cgttccatcg tggagtgcgg tggggtgccc 80460
gcgatcatat gtgccctgct gccagagac ccggcctggt tatggaccgg acccccgggg 80520
ttagtggtgt ttccgccacc catgcccccg taccatggcc cgggttcccc tgattaggct 80580
acgagtcgcg gtgatcgctt cccaaaaacc gagctcgtt tgtctgtctt ggtcttccac 80640
ccccccccc gcccgcccc acaccataac accgagaaca acacacgggg gtgggcgtaa 80700
cataataaag ctttattggt aactagttaa cggcaagtcc gtgggtggcg cgacggtgtc 80760
ctccgggatc atctcgtcgt cctcgacggg ggtggtgaa tgaggcggcc cctcgcggtc 80820
cgcttgccgt gggccgtgcc cataggcctc cggtctctgt cgtccatgg gcataggcgc 80880
ggggagactg tttccggcgt cgcggacctc caggtccctg ggagactccg gtcgggctaa 80940
cggacgaaac gcggaagcgc gaaacacgcc gtcggtgacc cgcaggagct cgttcatcag 81000
taaccaatcc atactcagcg taacggccag cccctggcga gacagatcca cggagtccgg 81060
aacgcggtc gtctggccca gggggccgag gctgtagtcc cccaggccc ctaggtcgcg 81120
acggctcgta agcacagcgc ggtcggccgc ggggctttgc gggggggcgt cctcgggcgc 81180
atgcgccatt acctctcgga tggccgcggc gcgctggtcg gccgagctga ccaaggcgc 81240
cacgaccaag gcgcgctccg tctgcaggcc ctccacgtg tcgtggagtt cctggacaaa 81300
ctcggccacg ggtcgggtc ccgcggccgc gcgcgcggct tgatagcagg ccgacagacg 81360
ccgcagcgc gctagaaact gaccatgaa gcaaaacccg gggacctggt ctccgacag 81420
cagcttcgac gcccgggcgt gaatgccga cacgacggac agaaaccctg gaatttcgcg 81480
ccggaccaag gccagcact gtctctcgtg cgacacctgg gccccagct cgtcgcacac 81540
ccccaggtgc gccgtggtt cggtgatgac ggaacgcagg ctgcgaggg acgcgaccag 81600
cgcgcgcttg gcgtcgtgat acatgctgca gtactgactc accgcgtccc ccatggcctc 81660
ggggggccag gggcccaggc ggtcgggcgt gtccccgacc accgcataca ggccggcggc 81720
gtcgtctcgc aaccgacact cgaaaaaggc ggagagcgtg cgcagtgtca gccgcagcag 81780
cacgatggcg tcctccagtt ggcaaatcag ggggtctgcg cgtcggcga ggtcctgcag 81840
cacccccgg gcgcccaggc cgtacatgct aatcaacagg aggctggtgc ccacctcggg 81900
ggcggggggg ggtgcagct gaccagggg ccgcagctgc tcgacggcac ccctggagat 81960
cacgtacagc tcccggagca gctgctctat gttgtcggcc atctgcatag tggggccgag 82020
gccgccccgg gcgcccgggt cgaggagggt gatcagcgcg cccagtttg tgcatggcc 82080
ctcaaccgtg gggagatagc ccagccaaa gtccccggcc caggccaaca cacgcagggc 82140
gaactcgacc gggcggggaa ggtaggccgc gctacacgtg gccctcagcg cgtccccgac 82200
caccagggcc agaacgtagg ggaacgaagc cgggtcggcg aggacgttg ggtgaatgcc 82260
ctcagggcg ggaagcggga tctgggtcgc cgcggccagg tggacagagg gggcgtggct 82320
gggctgccc acggggagaa gcgcggacag cggcgtggcc ggggtggtgg gggatgctc 82380

```

-continued

---

ccagtgggtc	tgaccataca	cgtegatcca	gatgagcgcc	gtctcgcgga	gaaggctggg	82440
ttgaccggaa	ctaaagcggc	gctcggccgt	ctcaaaactcc	cccacgagcg	cccgcgcgag	82500
gctcgcgaga	tgttccgctg	gcaecggccg	acctatgata	cgcgcacgcg	tctggcttag	82560
aacgcccccc	gacagggcca	cgcctcgcga	gagccgccc	tgcgtgtgct	cgtggcgc	82620
ctggaccocg	ctgaaagt	ttacgtagtt	ggcatagtag	ccgtattccc	gcgccaacc	82680
aaacacgttc	gacccccgca	gggcaatgca	cccaaagagc	tgctggactt	cgccgagtc	82740
gtggccgggt	ggcgtcgcg	cggggagccc	cgcgcgcca	aacccctcca	gggcccgaag	82800
gtagtgcgtg	cagtgcgagg	cgtggaacc	agcgtcgatc	agggtgtga	tcaccacgga	82860
gggcgaattg	gtattctgga	tcaacgtcca	cgtctgctgc	agcagagcca	gcagccgctg	82920
ctggcgccg	gcgagggct	gctccccgag	ctgcagcagg	ctggagacgg	caggctgga	82980
gactgccagt	gcccagcaac	tcaggaacgg	cacgtcggga	tcaaacacgg	ccacgtccgt	83040
ccgcacgcgc	gccattagcg	tccccgggg	cgcacaggcc	gagcgcgggc	tgacgcggct	83100
gagggccgtc	gacacgcgca	cctcctcgcg	gctgcgaacc	atcttgttgg	cctcagcgg	83160
cggaatcatt	atggccgggt	cgatctccc	cacggtgtgc	tgaaactgcg	ccaacagggg	83220
cggccggacc	acagccccc	gctcgggggt	cgtcaggtag	tcgtccacca	gggccaacgt	83280
aaagagggcc	cgtgtgaggg	gagtgagggt	cgcgtcgtct	atgcgctgga	ggtgcgccga	83340
gaacagcgtc	acccgattac	tcaccagggc	caagaaccgg	aggccctctt	gcacgaacgg	83400
ggcggggaag	agcaggctgt	acgcgggggt	ggttaaggttc	gcgctgggct	gccccaacgg	83460
gaccggcgcc	atcttgagcg	acgtctcccc	aagggcctcg	atggaggtcc	gcgggctcat	83520
ggccaagcag	ctcttggtga	cggtttgcca	gcggtctatc	cactccacgg	cgcactggcg	83580
gacgcggacc	ggccccaggg	cgcgcgcggt	gcgcaggccg	gcggaatcca	gcgcatggga	83640
cgtgtcggag	ccggtgaccg	cgaggatggt	gtccttgatg	acctccatct	cccggaaggc	83700
ctggtcgggg	gctcggggga	gagccaccac	caagcgggtg	acgagcaacc	cggggaggtt	83760
ctcggccaag	agcgcctct	ccggaagccc	gtgggcccgg	tggagcgcgc	acaggtgttc	83820
cagcagcggc	cgcacagcag	cccgcgcgct	tgccggggcg	atggccgttc	ccgacaacag	83880
aaacgcgcgc	atggcggcgc	gcagcttggc	cgtggccaga	aacgcgggt	cgtccgcccc	83940
gtttgcgctc	tcggccgctg	gggttggcgg	ttggcgaagg	ccggctaggc	tcgccaatag	84000
gcgctgcata	ggtccgtccg	agggcggacc	ggcgggtgag	gtcgtgacga	cgggggcctc	84060
ggacgggaga	ccgcggtctg	ccatgacgcc	cgcctcgcgt	gggtggggga	cagcgtagac	84120
caacgacgag	accgggcccgg	aatgactgtc	gtgcgctgta	gggagcggcg	aattatcgat	84180
ccccgcggc	cctccaggac	ccccgcaggc	ggtgcgagta	ccccgcctct	tcgcggggtg	84240
ttatacggcc	acttaagtc	cggcatcccc	ttcgcggacc	caggccccgg	ggattgtccg	84300
gatgtgcggg	cagccccgac	ggcgtgggtt	gcggacttct	tgccggggcg	cccaaatggc	84360
cctttaaacc	tgtgtatacg	gacgcgcggc	gccagtcggc	caacacaacc	caccggaggc	84420
ggtagccgcg	tttggctgtg	gggtgggtgg	ttccgccttg	cgtgagtgtc	ctttcgacct	84480
ccccccccct	ccctcccccg	ggtcttgcta	ggtcgcgata	tggggtcgca	atgaagacca	84540
atccgctacc	cgcaaccccc	tccgtgtggg	gcgggagtag	cgtggaactc	ccccccacaa	84600
cacgcgatac	cgcgggacag	ggcctgcttc	ggcgcgtcct	gcgccccccg	atctctcgcc	84660

-continued

---

gcgacggccc agggctcccc agggggtcgg gaccccggag ggcggccagc acgctgtggt 84720  
tgcttgccct ggacggcaca gacgcgcccc ctggggcgct gaccccaac gacgataccg 84780  
aacaggccct ggacaagatc ctgcccggca ccatgcgcgg gggggcgcc ctgatcggct 84840  
ccccgcgcca tcataaacc cgccaagtga tctgacgga tctgtgcaa cccaacgceg 84900  
atcgtgctgg gacgctgctt ctggcgctgc ggcaccccgc cgacctgect cacctggccc 84960  
accagcgcgc cccgccaggc cggcagaccg agcggctggg cgaggcctgg gcccagctga 85020  
tggaggcgac cgccctgggg tcggggcgag ccgagagcgg gtgcacgcgc gcgggcctag 85080  
tgtcgtttaa cttcctgggt gcggcgctgt ccgcctcgta cgacgcgcgc gacgcgcgcg 85140  
atgcggtagc ggcccacgtc accggcaact accgcgggac gcgggtgggg gcgcgcctgg 85200  
atcgtttttc cgagtgctct gcgcacatgg ttcacacgca cgtcttcccc caccaggtca 85260  
tgcgggtttt cggggggctg gtgtcgtggg tcaccagga cgagctagcg agcgtcaacc 85320  
ccgtgtgcgc cgggccccag gaggcggcgc acaccggcca cccgggcccg ccccgctcgg 85380  
ccgtgatcct cccggcatgt cgttctgtgg acctggacgc cgagctgggg ctggggggcc 85440  
cgggcgcggc gtttctgtac ctggtattca cttaccgcca gcgcccggac caggagctgt 85500  
gttgtgtgta cgtgatcaag agccagctcc ccccgcgcg gttggagccg gccctggagc 85560  
ggctgtttgg gcgcctccgg atcaccaca cgattcacgg caccgaggac atgacgcccc 85620  
cggccccaaa ccgaaacccc gacttcccc tcgcccgcct ggcccgaat ccccaaaccc 85680  
cgcgttgctc ggctggccc gtcacgaacc ccagttcgc cgacagctg taccgctggc 85740  
agccggacct gcgggggcgc cccaccgcac gcacctgtac gtacccgcc tttgcagagc 85800  
tcggcatgat gcccgaggat agtccccgct gcctgcaccg caccgagcgc tttggggcgg 85860  
tcagcgtccc cgttgtcatt ctggaaggcg tgggtgtggc ccccgccgag tggcgggcat 85920  
gcgcgtgagc gtacgaaacg ccccgcccac acaacgctcc gcccacaacc ccttccccgc 85980  
tgtcactcgt tgttcgttga cccgggctgc cgccaataa agccactgaa acccgaaacg 86040  
cgagtgtgt aacgtccttt gggcgggagg aagccacaaa atgcaaatgg gatacatgga 86100  
aggaacacac ccccgtagc caggacatcg gtgtgtcctt ttgggttca ctgaaactgg 86160  
ccccgcgcc acccctgcgc gatgtggata aaaagccagc ggggtgggt tagggtacca 86220  
cagggtgggt ctttggaaac ttgcccgtcg ccgtgctcct gtgagcttgc gtcctcccc 86280  
ggtttcttt gcgctcccc cttccggacc tgctctgcc tactcttctt tggtctcgg 86340  
tgcgattcgt caggcagcgg ccttgtcgaa tctcgacccc accactgcc ggacccgccc 86400  
acgtcccctc tcgagcccgc cgaaacccc cgcgctgtgt gaaatggcca gccgccagc 86460  
cgcacctct cccgtcgaag cgcgggcccc ggttggggga caggaggccg gcggccccag 86520  
cgcagccacc cagggggagg ccgcccgggc cctctcgc caccggcacc acgtgtactg 86580  
ccagcgagtc aatggcgtga tgggtcttcc cgacaagacg cccgggtccg cgtcctaccg 86640  
catcagcgat aacaactttg tccaatgtgg ttccaactgc accatgatca tcgacggaga 86700  
cgtggtgcgc gggcgcccc aggaaccggg ggccgcggca tccccgctc cctcgttgc 86760  
ggtgacaaac atcggagccg gcagcgacgg cgggaccgcc gtcgtggcat tcgggggaaac 86820  
cccacgtcgc tcggcgggga cgtctaccgg taccagacg gccgacgtcc ccaccgaggc 86880  
ccttgggggc cccctcctc ctccccgctt caccctgggt ggcggctgtt gttcctgctg 86940

-continued

---

cgacacacgg cgccgctctg cggatttcgg gggggagggg gatccagtcg gccccgcgga 87000  
 gttcgtctcg gacgaccggt cgtccgattc cgactcggat gactcggagg acacggactc 87060  
 ggagacgctg tcacacgcct cctcggagct gtccggcggg gccacgtacg acgacgccct 87120  
 tgactccgat tcgtcctcgg atgactccct gcagatagat gggcccgctg gtgcgccgtg 87180  
 gagcaatgac acccgcccc tggatgtttg ccccgggacc cccggcccgg gcgcccacgc 87240  
 cggtggtccc tcagcggtag acccacacgc gccgacgcca gaggccggcg ctggctcttc 87300  
 ggccgatccc gccgtggccc gggacgacgc ggaggggctt tcggaccccc ggccacgtct 87360  
 gggaacgggc acggcctacc ccgtccccc ggaactcacg cccgagaacg cggaggccct 87420  
 ggcgcgcttt ctgggagatg ccgtgaaccg cgaaccgcg ctcctgctgg agtacttttg 87480  
 ccggtgcgcc cgcgagaaa ccaagcgtgt ccccccagg acattcggca gccccctcg 87540  
 cctcacggag gacgactttg ggcttctcaa ctacgcgctc gtggagatgc agcgcctgtg 87600  
 tctggacgtt cctccggtcc cgcgaaacgc atacatgcc tattatctca gggagtatgt 87660  
 gacgcggctg gtcaacgggt tcaagccgct ggtgagccgg tccgctcgcc tttaccgcat 87720  
 cctggggggt ctggtgcacc tgcggatccg gaccgggag gccctcttg aggagtggct 87780  
 cggatccaag gaagtggccc tggattttg cctgacggaa aggcttcgcg agcacgaagc 87840  
 ccagctggg atcctggccc aggctctgga ccattacgac tgtctgatcc acagcacacc 87900  
 gcacacgctg gtcgagcggg ggctgcaatc ggccctgaag tatgaggagt tttacctaaa 87960  
 cgtttttggc gggcactaca tggagtcctt cttccagatg tacaccgca tcgcccgtt 88020  
 tttggcctgc cggccacgc cggcctcgc ccacatcgc ctggggcgag aggggtcgtg 88080  
 gtgggaaatg ttaagtctt tttccaccg cctctacgac caccagatcg taccgtcgac 88140  
 ccccgccatg ctgaacctgg ggacccgcaa ctactacacc tccagctgct acctggtaaa 88200  
 ccccaggcc accacaaaa aggcgaccct gcgggccatc accagcaacg tcagtgccat 88260  
 cctcgcgcc aacgggggca tcgggctatg cgtgcaggcg tttaacgact ccggccccgg 88320  
 gaccgccagc gtcatcccc ccctcaaggt ccttgactcg ctggtggcgg cgcacaacaa 88380  
 agagagcgcg cgtccgacc gcgcgctcgt gtacctggag ccgtggcaca ccgacgtgcg 88440  
 ggccgtgctc cggatgaagg gggctcctgc cggcgaagag gccagcgcct gcgacaatat 88500  
 cttcagcgc cctcggatgc cagacctgtt tttcaagcgc ctgattcgcc acctggacgg 88560  
 cgagaagaac gtcacatgga ccctgttcga ccgggacacc agcatgtcgc tcgcccactt 88620  
 tcacggggag gagttcgaga agctctacca gcacctcag gtcctgggt tcggcgagca 88680  
 gataccatc caggagctgg cctatggcat tgtgcgcagt gcggccacga ccgggagccc 88740  
 cttcgtcatg tcaaaagac cgggtgaacc ccactacatc tacgacaccc agggggcggc 88800  
 catcgccgc tcaacctct gcaccgagat cgtccatccg gcctccaagc gatccagtgg 88860  
 ggtctgcaac ctgggaagcg tgaatctggc ccgatgcgct tccaggcaga cgtttgactt 88920  
 tgggcggctc cgcgacgcg tgcaggcgtg cgtgctgatg gtgaacatca tgatcgacag 88980  
 cacgtacaa cccacgcccc agtgacccc cggcaacgac aacctgcggt ccatgggaat 89040  
 cggcatcgag ggccctgcaca cggcctgcct gaagctgggg ctggatctgg agtctgccga 89100  
 atttcaggac ctgaacaaac acatgcgca ggtgatgctg ctgtcggcga tgaagaccag 89160  
 caacgcgctg tcgcttcgcg gggcccgtcc cttcaaccac ttaagcgca gcatgtatcg 89220

---

-continued

---

cgccggccgc tttcactggg agcgccttcc ggacgcccgg ccgcggtacg agggcgagtg 89280  
ggagatgcta cgccagagca tgatgaaaca cggcctgcgc aacagccagt ttgtcgcgt 89340  
gatgcccacc gccgcctcgg cgcagatctc ggacgtcagc gagggtcttg cccccctgtt 89400  
caccaacctg ttcagcaagg tgaccocggga cggcgagacg ctgcgccccca acacgtctct 89460  
gctaaaggaa ctggaacgca cgtttagcgg gaagcgcctc ctggaggtga tggacagtct 89520  
cgacgccaag cagtggctcc tggcgcaggc gctcccgtgc ctggagccca cccaccccct 89580  
ccggcgattc aagaccgctg ttgactacga ccagaagttg ctgatcgacc tgtgtgcgga 89640  
ccgcgcccc tacgtcgacc atagccaatc catgaccctg tatgtcacgg agaaggcgga 89700  
cgggaccctc ccagctcca ccttggctcc ccttctggtc cacgcatata agcggggact 89760  
aaaaacaggg atgtactact gcaaggctcg caaggcgacc aacagcgggg tctttggcgg 89820  
cgacgacaac attgtctgca tgagctgcgc gctgtgaccg acaaaccccc tccgcgccag 89880  
gcccgcgcc actgtcgtcg ccgtcccacg ctctcccctg ctgcoatgga tccgcggcc 89940  
ccagccctct cccccctct gacggccctt acggaccaga gcgcgacggc ggacctggcg 90000  
atccagattc caaagtgcc cgacccccgag aggtacttct acacctccca gtgtcccgc 90060  
attaaccacc tgcgctccct cagcatcctt aaccgctggc tggaaaccga gcttgttttc 90120  
gtgggggacg aggaggacgt ctccaagctt tccgagggcg agctcagctt ttaccgcttc 90180  
ctctctcgtt tctgtcggc cgccgacgac ctggttacgg aaaacctggg cggcctctcc 90240  
ggcctgtttg agcagaagga cattctccac tactacgtgg agcaggaatg catcgaagtc 90300  
gtacactcgc gcgtgtacaa catcatccag ctggtgcttt tccacaacaa cgaccaggcg 90360  
cgccgcgagt acgtggccgg taccatcaac caccggcca tccgcgcaa ggtggactgg 90420  
ttggaagcgc ggtgtcggga atgcgcctcc gttccgaaa agttcattct catgatectc 90480  
atcgaggca tcttttttgc gcctcgttt gccgcatcg cctacctcg caccaacaac 90540  
cttctgctgg tcacctgcca gtcaaacgac ctcatcagcc gggacgagc cgtgcacacg 90600  
acggcctcgt gttacatcta caacaactac ctcgccgggc acgccaagcc cccgcccgc 90660  
cgcgtgtacg ggtgttccg ccaggcggtc gagatcgaga tgggattat ccgateccag 90720  
gcccgcagc acagccatat cctgagcccg gcggcgtcgg cggccatcga aaactacgtg 90780  
cgattcagc cggatcgctt gttgggcctt atccacatga agccactgtt tccgccccca 90840  
cccccgacg ccagctttcc gctgagcctc atgtccaccg acaaacacac caattttttc 90900  
gagtgtcgca gcacctccta cgccggggcg gtcgtcaacg atctgtgagt gtcgcgccgc 90960  
gcttctaccc gtgtttgccc ataataaac tctgaaccaa actttgggtc tcatttggat 91020  
tcttgcagg gacgcggggg tgggagagga taaaaggcgg cgcaaaaagc agtaaccagg 91080  
tccgtccaga ttctcggggc ataggatacc ataattttat tgggtgggtcg tttgtcggg 91140  
gacaagcgcg ctcgtctgac gtttgggcta ctcgtcccag aatttgcca ggacgtcctt 91200  
gtagaacgcg ggtggggggg cctgggtccc caactgctcc agaaacctgt cggcgatatc 91260  
agggggcgtg atatgcggg tcacgataga tcgcgccagg tttctcgcg ggatgtcctg 91320  
gtagatagc aggcgtttca gaagagtcca cgcccccg tccttggggc cgataagcga 91380  
tatgacgtac ttaatgtagc ggtgttccac cagctcgggt atggatcatgg gatcggggag 91440  
ccagtccagg gactctgggg cgctcgtggat gacgtggcgt cgcgggttgg ccacataact 91500

-continued

---

gcggtgctct tccagcagct gcgcgttcgg gacctggacg agctcgggcg gggtagtat 91560  
ctccgaggag gacgacctgg ggccgggggtg gcccccgta acgtcccggg gatccagggg 91620  
gaggctctcg tcgtcttcgt atccgceggc gatctgttgg gttagaattt cggccacga 91680  
gacgcgcgctc tcgggtccgc cggcgccggc cggcagaggg ggctggttt cgtggagcg 91740  
cgagctggtg tgtcccggc ggatggccc cgggtctga gagcactcg ggggggtcca 91800  
gtgacattcg cgcagacat cctccacgga ggcgtaggtg ttattgggat ggaggtcgg 91860  
gtggcagcgg acaagaggg ccaggaactg gggtagctc atcttaaagt actttagtat 91920  
atcgcgacag ttgactgtgg gaatgtagca ggcgctaata tccaacacaa tatcacagcc 91980  
catcaacagg aggtcagtg cctgtgtgta cacgtacgag accgtgttgg tgtgatagag 92040  
gttgccgag gcatcgtccg cctccagctg acccgagtta atgtaggcgt accccagggc 92100  
ccggagaacg cgaatacaga acagatgccc cagacgagg gccggcttcg agggcgcggc 92160  
ggacggcagc gcggctccgg acccggcgt cccccggtc cccgagcca gagaggtgcc 92220  
gcgccggcgc atgttggaaa aggcagagct gggctcggag tcggtgatgg gggaaaggcg 92280  
tgagagggcg tccacgtcac tggcctcctc gtccgtccgg cactgggccc tcgtgcgggc 92340  
caggatggcc ttggctccaa acacaaccgg ctccatacaa ttgaccgcc gatcggtaac 92400  
gaagatgggg aaaagggact tttgggtaa caccttaat aagcgacaga ggcagtgtag 92460  
cgtaatggcc tcgctgctgt aactgggta tcggcgtga tatttgacca ccaactgtga 92520  
catgacgttc cacaggtcca cggcaatggg ggtgaagtac ccggccgggg cccaaggcc 92580  
ccggcgttg accagatgg gtgtgtgggc aaacttcac atccgaaca aacctatgc 92640  
aggtaaatg taactcggga tcggcctaac taaggcgtgg ttggtgcgac ggtccgggac 92700  
accgagcct gtctctctgt gtatggtgac ccagacaaca acaccgacac aagaggacaa 92760  
taatccgtta ggggacgctc ttataattt cgatggcca actccacgcg gattggtgca 92820  
gcaccctgca tcgcccggtg cgggccaacc tccccccgc tcattgcctc ttccaaaagg 92880  
gtgtggccta acgagctggg ggcgtattta atcagctag cgcggcgggc ctgccctagt 92940  
ttctggctcg gtgagcagc gtccggttgc ttgggtcccc tggtgccaat caaaaccca 93000  
ccctcgcagc ggcatacggc ccctccgct cccgcaccgc agacccggc cggctgccc 93060  
tcaccaccga agcccacctc gtcactgtgg ggtgttccca gcccgcttg ggatgacgga 93120  
tccccctggc ggtgtggccc ccgcctccc cgtggaggac gcgtcggac cgtccctcg 93180  
gcagccggag gagggggcgc cctgccaggg ggtcctgcag ggcgccgaac ttaatggaat 93240  
cctacaggcg tttgccccgc tcgacacgag ccttctggac tcgcttctgg ttatgggca 93300  
ccggggcctc cttatccata acacgatctt tggggagcag gtgttcctgc ccctggaaca 93360  
ctcgcaatc agtcggtatc gctggcggc acccaaggcg gcgttcctgt ctctcgtgga 93420  
ccagaagcgc tccctcctga gcgtgttctg cccaaccag taccggacc tacgtcgggt 93480  
ggagttggcg atcaaggggc agggcccgtt tcgcaagctg gttcagcga tatggacgac 93540  
gacgtccgac ggcgagggcg ttgagctagc cagcagagcg ctgatgaag gcgaactgac 93600  
gagctttgtg gtgctggttc cccagggaac ccccagctt cagttgcgcc tgacgaggcc 93660  
gcagctcacc aaggtcctta acgcgaccgg ggcgatagt gccacgcca ccacgttcga 93720  
gctcgggggtt aacggcaaat tttccgtgtt caccacagat acctcgtca cctttgctgc 93780

-continued

---

ccgcgaggag ggcgtgtcgt ccagcaccag caccaggtc cagatcctgt ccaacgcgt 93840  
 caccaaggcg ggcagggcgg ccgccaacgc caagacggtg tacggggaaa ataccatcg 93900  
 caccttctct gtggtcgtcg acgattgcag catgcccggc gtgctccggc gactgcaggt 93960  
 cggcgggggc accctcaagt tcttctcac gaccccctc cccagtctgt gcgtaccgc 94020  
 caccggtccc aacgcggtat cggcgggtatt tctcctgaaa cccagaaga tttgcctgga 94080  
 ctggctgggt catagccagg ggtctcctc agccgggagc tcggcctccc gggcctctgg 94140  
 gagcgagcca acagacagcc aggactccgc gtcggacgcg gtcagccacg gcgatccgga 94200  
 agacctgat ggcgctgccc gggcgggaga ggcgggggccc ttgcatgect gtccgatgcc 94260  
 gtcgtcgacc acgcccgtca ctcccacgac caagcggggg cgctcggggg gcgaggatgc 94320  
 gcgcgaggac acggccctaa agaaacctaa gacggggctc cccaccgac ccccgcccgc 94380  
 agatccagtc cccctggaca cggaggacga ctccgatgcg gcggacggga cggcggcccg 94440  
 tcccgcgct ccagacgccc ggagcggaa cggttacgcg tgttactttc gcgacctccc 94500  
 gaccggagaa gcaagcccc gcgccttctc cgcttccgg gggggcccc aaaccccgt 94560  
 tggttttgga tccccctgac gggcgggggc cttggcggcc gcccaactct cgcacctcc 94620  
 cgggttaatg taaataaatg tggattgccc caacacttcc ccgctgtcgc cgtgtggttc 94680  
 atgtgtgtgc ctggcggccc caccctcggg ttctgtgatt tctttccct gtcttataa 94740  
 aagccgtatg tggggcgtga cggaaaccac ccgctgcca tcacggcca ggcgcgggat 94800  
 gctccgaac gacagccacc gggcgggtgc cccggaggac ggccaggac gggctgacga 94860  
 cggacggcca cacctcgcgt gcgtgggggc cctggcgcgg gggttcatgc atatctggct 94920  
 tcaggccgcc acgctgggtt ttgcccggatc ggtcgttatg tcgcgcgggc cgtacgcgaa 94980  
 tgccgcgtct gggcgttctc ccgtcgggtg cgcctgctg ggtttatgc gcgcacccc 95040  
 tccccctcgc cggcccaccg cgcggatata cgcctggctc aaactggcgg ccggtgggac 95100  
 ggcccttgtt ctgtggagtc tcggggagcc cggcacgcag ccggggggccc cggcccggg 95160  
 cccggccacc cagtgcctgg cactgggcgc cgcctatgcg gcctcctgg tgcctgcga 95220  
 tgacgtctat ccgctcttcc tctcgcgcc ggggcccctg ttctgctgca cctggggat 95280  
 ggtcgtcggc gggctgacga tcggaggcag cgcgcgctac tggtaggatc gtgggcccgc 95340  
 cgcggccgcc ctggcccggc cgggtgtggtc gggcccgggg gcgaccaccg ccagggactg 95400  
 ctttccagg gcttgcggcg accaccgccc cgtctgtgtc atcaccgcag gcgagtctct 95460  
 ttcccgcgc ccccggagg acccagagcg acccggggtt cccgggccc cgtccccccc 95520  
 gacccccca cgatcccacg ggcggccggc cgatgaggtc gcaccggcca gggctcgcgc 95580  
 gccgaaaac gctcgggtgc ccgtggctac ctttctgggg gcgggcgcgc ttgcccgtca 95640  
 gacggtgcga gaacatgccc ggggaacgcc gggcccgggc ctgcctgtgt ggcgccaggt 95700  
 gtttctcgga ggcctatggt cggtagccct gacggagctg tgcagggcgc tccgcccctg 95760  
 ggaccttacg gaccgctgc tgtttgttca cgcggactg caggtcatca acctcgggtt 95820  
 ggtgtttcgg ttttccagg ttgtcgtgta tgcggcgtca gggggtgccc tgtggatttc 95880  
 gttggcgcag gtgctggggc tccggcgtc cctgcacagg aaggacccc gggacggggc 95940  
 ccggttggcg gcgacgcttc ggggctctt cttctccgtg tacgcgctgg ggtttgggg 96000  
 ggggtgctg ctgtgccctc cggggtcaac gggcggggcg tcgggcgatt gatataat 96060

-continued

---

tcaataaaag gcattagtc cgaagaccgc cgggtgtgta tgatttcgcc ataacacca 96120
aaccccgat ggggccggg tataaattcc ggaaggggac acgggctacc ctcaactatcg 96180
agggcgcttg gtcgggaggg cgcacgaac gcacaccccc atccgggtgg cgggtggag 96240
gtcgttttca tgcccggtct cgccttgccg ggaacgctag ccgatccctc gcgaggggga 96300
ggcgtcgggc atggccccgg ggcgggtggg ccttgccgtg gtcctgtgga gcctgttgtg 96360
gctcggggcg ggggtgtccg ggggctcggg aactgcctcc accgggcccc cgatcacccg 96420
gggagcggtg acgaacgca gcgagggccc cacatcgggg tccccgggt cagccgccag 96480
cccggaggtc acccccacat cgaccccaaa cccaacaat gtcacacaaa aaaaaaccac 96540
ccccaccgag ccggccagcc cccaacaac cccaagccc acctccacgc caaaaagccc 96600
ccccactcc acccccagc ccaaacccaa gaacaacacc acccccgcca agtcggggcg 96660
ccccactaaa cccccggg cgggtgtggg cgaccgccc gaccattgg ccggtaecgg 96720
ctcgggggtg cagatccgat gccggtttcg gaattccacc cgcattgagt tccgcctcca 96780
gatattggct tactccatgg gtccgctccc cccaatcgt ccggctccc acctagagga 96840
ggtcctgacg aacatcacg cccacccgg gggactcctg gtgtacgaca gcgccccaa 96900
cctaaccgac cccacgtgc tctggggcga gggggccgg ccgggcccg acctccgtt 96960
gtattctgtc accgggcccg tgccgaccca gcggctgatt atcggcgagg tgacgcccgc 97020
gaccagggga atgtattact tggcctgggg ccggatggac agcccgcacg agtacgggac 97080
gtgggtgcgc gtcgcgatg tccgccccg gtctctgacc ctccagccc acgcggtgat 97140
ggagggtcag ccgttcaagg cgacgtgac gcccgccgc tactaccgc gtaaccccgt 97200
ggagttgtc tggttcagg acgaccacca ggtgtttaac ccgggcccaga tcgacacgca 97260
gacgcacgag caccgccg ggttcaccac agtctctacc gtgacctcc aggetgtcgg 97320
cggccaggtc ccccgccg ccttcacctg ccagatgacg tggcatcgcg actccgtgac 97380
gttctcgcga cgaatgcca ccgggtggc cctggtgctg ccggcgccaa ccatcacat 97440
ggaatttggg gtcgcgattg tggctcgcac gcccgctgc gtccccgagg gcgtgacgtt 97500
tgctgtgtc ctgggggacg acccctcacc ggcggetaag tcggccgta cggcccagga 97560
gtcgtgacg caccggggc tggctacggt ccggtccacc ctgcccattt cgtacgacta 97620
cagcgagtac atctgtcgg tgaccggata tccggccggg attcccgttc tagaacacca 97680
cggcagtcac cagccccac ccagggaccc caccgagcgg caggatgatc aggcgatcga 97740
gtgggtgggg attggaatcg gggttctcgc gccgggggtc ctggtcgtaa cggcaatcgt 97800
gtacgtcgtc cgcacatcac agtcgcgga gcgtcatcgg cggtaacgca agaccccc 97860
gttaccttt taatatctat atagtttgg ccccccteta tcccggccac cgtggggcgc 97920
tataaagccg ccacctctc tccctcagg tcatccttg tcgatccga acgacacag 97980
gcgtggagca aaacgcctcc cctgagccg ctttctacc aacacaacgg catgcctctg 98040
cgggcatcgg aacacgccta ccggcccctg ggcgccggg cacccccac gcgggctcgg 98100
ctccccccg cgccctgggt tggcgtcggg accatcatc ggggagttgt gatcattgcc 98160
gcgttggtcc tcgtccctc gcgggctcgc tgggcacttt cccatgca cagcgatgg 98220
cacgagttca acctcgggtg catatcctg gatccgacc ccatggagca cgagcaggcg 98280
gtcggcggtg gtacgcccc ggcgacctg atcccccg cggctgcaa acagctggcc 98340

---

-continued

---

gccgtcgcac gcgtccagtc ggcaagatcc tcgggctact ggtgggtgag cggagacggc 98400  
attcgggect gcctgcggct cgtcgcggc gtcggcggtt ttgaccagtt ttgcgaggag 98460  
cccgcccttc gcatatgcta ctatcccgc agtcccgggg gctttgttca gtttctaact 98520  
tcgacccgca acgcgctggg gctgocgtga ggcgcggtga ctgcggctctg tctcgtctcc 98580  
tcttctcccc ttccctcccc ctccgcattc caggatcaca ccggtcaacg aggggtgggg 98640  
gggtccggca cggacccaaa ataataaaca cacaatcacg tgcgataaaa agaacacggc 98700  
gtcccctgtg gtgttttttg ttatttttat taaatctcgt cgacaaaacg ggggaaaggg 98760  
gcgtggctca gcgacggcag cacggggcga ggcgttcacc ggctccggcg tccttcgctg 98820  
ttaagcttgg tcaggagggc gctcaggggc gcgacgttgg tcggggcctc gttggtcagg 98880  
gcgttgctc gatggcgggc gaggacgggc gaggggctca acggcggggg cgggggcccg 98940  
gtgcggcccg ggggggaaaa tagggcggtat cccccccagt cgtacagggg attttccgcc 99000  
tcaatgtacg gggaggcccg cgetgcattc gccgtgttca cgcagacgtt ttcgtagacc 99060  
cgcattccatg gtatttctcc gtagacacgc cccccctcct cgctcacctg ctcgatatatt 99120  
gactcgtcgt cctcgtaggg ggcgtgccgt tcgcccggcg aggcggcgtg ggtggctttg 99180  
cggcgggctg cgtcgtcgtc gtcgtcggcc gtcagatacg tggettccat ctggtcgggt 99240  
tctccctccg gggcgggtcc ccacaccctg ggccgatcga ggctcccag agacgcgcgc 99300  
cggacgagga gggggcacgt gcgcgcggc ggtcgcctgt cgggtcccgc gacggttacg 99360  
gccgggagc gcgggggac ctcccccatg tgcgtgtaat acgtggccgg ctgtgtggcc 99420  
gcagcggggg gctcggcgac cgggtcgttc gcatccgaa gcggggccc cgcgcctcc 99480  
gcgcggccc tcggaaacct ccgggtggac gcgggggtcg agtgtaggcg aggtcggggg 99540  
agggggggg gctcgttgc gcgcgcgc ctctgaaatct tttcccgaca ggtcccacc 99600  
cccgcgcgat gcccccccg gccgcgtggc atgtcgtccg ggggaggccc cgcggaccac 99660  
gtcgtccgc gagaccccac gagccgcagg atggactcgt agtggagcga cggcggcccg 99720  
ttcgggagca gatccgcgc cagggcggcc ccgaaccaag ccttgatgct caactccatc 99780  
cgggcccagc tggggcggt catcgtgggg aacagggggg cgggtggtccg acagaaaacg 99840  
tctggctgt ccaccgcgc ccgcagatac tcgttgttca ggctgtcggg gcccagacg 99900  
ccgtaccgc tgagggtcgc gttgatgata tactgggct ggtgatggac gatcgacaga 99960  
acctccaccg tgatacagc ggtatccac gtcccgtacg taccgccct ccgcttgcg 100020  
gtctgccaca ggttgctag gcgcgtcagg tggcccagga cgtcgtgac cgcgcctc 100080  
agcgcctgc actgcatgga gccggtcgtg ccgctgggac cccggtccag atggcgccg 100140  
aacgtttccg cgggcgcctc cgggctgcgc ccgagcggga ggaaccggcg attggaggga 100200  
ctcagccggt gacatacgtg cttgtccgtc gtccacagca tccaggacgc ccaccggtac 100260  
agcagcgaga cgtaggccag gagctcgtg agccgcagt cgggtgcggg gctggggcgg 100320  
cttgggtccg cggggcgc ataaagacatg tactgctgaa tccgatggag ggcgtcgcgc 100380  
aggccggcca cgttgccgc gtaactggcc gccacggccc cgtcttgaa cggggtgcgc 100440  
gccagcagct ttggcgccag ggtgggcgc agcagcacgt gaaggctggg gtcgcagtcg 100500  
cccacggggt cctcggggg gtccaggccc ctgggacca ccgtctgcag gtaactccag 100560  
tactcgtgta ggatggcgc gctcaactgg ccgcccggca gctccacctc gccagcgc 100620

-continued

---

tgggtggcgg ccgaagcgta gtgcccgatg tactcgtagt gcgggtcgct ggcgagcccg 100680  
tccacgatca aactctcggg aaccgtgttg tgttgccgcg cggccaaccg gacgctgcca 100740  
tcggtgcagg tcgaaaacgc cggctgcgcg tcgctcggagc gctgcccga ggcgcccacg 100800  
gcccgcgtaa ggagcccctc cgggggtggg agcagacacc cgcggaagat ggcgcccctc 100860  
ggaaaccccg cgttgtcgcc gcgatcagg ttggcaggcg tcaggcaccg cgcagcccgc 100920  
aggagctcg cgcgcgctc cgcgctgc atggtgacgc ccgttcggtc gggaccgcc 100980  
ggtcggagtt atgcccctc caggccatc ggggcgcttt ttatcgggag gagcttatgg 101040  
gcgtggcggg cctcccagcc cggctgcgcg cctcccagc acgtgcgccc gcaggcggc 101100  
ggccccctc tctcccata gcagtttct aaactgggac atgatgtcca ccacgcccgc 101160  
ccgcccgc aacacggacc cgcgcttac gggggcggg ggaaggct ccaggtcctt 101220  
gagaagaaag gcggggtctc ccgtcccga cacgggggccc cggggcgtg aggagcggg 101280  
gcccagatcc acgtgctccg cggcgcgcg gacgtccgc cagaactgg cgggggtggt 101340  
gcccgcgta aggggtcggg tcgctcggag gacgcacgcg tagcgcaggg ggggtgacgt 101400  
gcccacctc gggcccgtga atccccctc aaacgcggc agtgtcacgc acgccaccac 101460  
ggtgtcggca aagcccagca gccgctcag gacgagccc gcgccagaa tggcgcgct 101520  
ggcccgcg tcgtcccggc gccggtgcgc gtcccgcac gcccccggc actttaagg 101580  
cacggtcgc agggccgtgt gcagcgcgta caccgcagc cccagcagg cgttgagccc 101640  
gctgttggc agcagcccgc gcgctcgggt gtcgcccag ccctcgtgct cggcccccac 101700  
gaccgcggg cttcccagg gcagggcgc aaacagctc tcccgcgcca cgtccgcaa 101760  
ggcgggtg tgcaactgct ggtgcaggc cccccccag accaccgaga gccactggc 101820  
cgtctgctc gccatcacc ccagcacatc cagcagcgc cccaggaagg cggcctccc 101880  
cgtcaaacg caccggacg cgctcggatt gaagcggcg agcagggccc cgttgccag 101940  
gtacgtcatg cggcccgat agcggggcgc cacgcgacag tcgcggtcca gcagcgcg 102000  
cacccccggc cagtacagca gggaccccag cagctgcg gaaacccg cgtcggggc 102060  
ggattgggg gacactaac ccccgcgct cagtaacgc acgcccgcg ccccagcgg 102120  
acgcaacgc gtgaggtcgc cgaactgc cctcagctc gcccccctgt cgtccaggtc 102180  
agaccgcgc gctcccgt gaaggcgcgt cccgcacacc caccgttga tggccagccg 102240  
cacgagcga tcccacaaa agctcatcgc ctgggcggg ctggttttg ttcgacgatc 102300  
cgtcaggta agaateccat cgcctcgtat ataccaggc aacgcctgc cctgctgag 102360  
ggtttggcg aaaaacacc cggggtgtc gggggaggc aagtgcagta ccccacgcg 102420  
cgataaaccg aacgcgctat ccgacacgg gtaaaaccg gccgatgcc ccagggctag 102480  
ggcggagcgc acggaactc cccacacggc aaactgagg gccagtcgat ccaacgggaa 102540  
tgcccgcg agctcccggc ccggcagcgc tccctccaga acctccact tggcgggga 102600  
acgggcccgc cgcgctcct ccggcccag ggttcggg tagtcgtcct cctcgtactg 102660  
cagctcctc aggaacagc gcgacggc caccgcgaa ccgcccagcc gcccaaaat 102720  
agcccgcg tcgacgggac ccaggatcc cccctgcg gcctgcggag gaccgcccgg 102780  
aacctcatc tcactgtcca ggcgaccgc caccgactg ctacgggccc catcgggccc 102840  
ggggcgtgc cgggacgctc ggcgatggg tgtggcggg gcttcgacg cgcgcccctc 102900

---

-continued

---

tggggctcgc gggccttccc gtcgacggcg cacgggcggc tcgtcgcccg ccatctcctc 102960  
cagagcctct agctcgctgt cgtcatcccc gcggaacacc gcacgcaggt accccatgaa 103020  
ccccacccca tcgcccgcgtg gctcgtccgc cacgggcgag gcgcgggggc gggtgatgc 103080  
gcgctcctg cgcgccggg gttcgcgagc cgacatggtg gcgatagacg cgggttatcg 103140  
gatgtccgct accccccaaa aaagaaaaag accccacagc gcggatggag gccgggtag 103200  
gtgcccggg accccctcgc gatgggaatg gacgggagcg acggggcgg cgcaaaaaa 103260  
cgcagtatct cccgcgaagg ctaccgccg cccagcccc cggccaaatg cggaaacggt 103320  
cccgcgctct cgcctttata cgcgggccc cctgcgacac aatcaccgt cctgggttc 103380  
gaatctacac gacaggcccc cagacgcggc taacacacac gccggcaacc cagaccccag 103440  
tgggttggtt gcgcggtccc gtctcctggc tagttcttcc cccaccacc aaataatcag 103500  
acgacaaccg caggttttgt aatgtatgtg ctctgttcta ttgtggatac gaaccggtga 103560  
cgggagggga aaaccagac gggggatgcg ggtccggtcg cgcacctac ccaccgtact 103620  
cgtcaattcc aagggcacgt gtaaacatct gctcaaacct gaagtcggcc atatccagag 103680  
cgccttaggg ggcggagtcg tggggggtaa atcccgccc cggggaatcc ccgtcccca 103740  
acatgtccag atcgaaatcg tctagcgcgt cggcatgcgc catcgccacg tcctcgccgt 103800  
ctaagtggag ctctgcccc aggtgacat cggtcggggg ggccgtcgac agtctgcccg 103860  
tgtgtcccgc ggggagaaaag gacaggcgcg gagcccccag ccccgctct tcgggggct 103920  
cgtcgtccgg gagatcgagc aggcctcoga tggtagacc gtaattgttt tctgtacgcg 103980  
cgcggtgta cgcgtgttcc cgcataccg cctcggaggc cgaggtcgtg aagctggaat 104040  
acgagtccaa ctctgcccc atcaacacca taaagtacc agaggcgcgg gcttggttc 104100  
catgcagggt gggaggggc gtcaacggcg cccctggctc ctccgtagcc gcgtgcgca 104160  
ccagcgggag gttaaaggtc tcgcaaatgt ggtttagctc ccgcagccgg cgggcctcga 104220  
ttggcactcc ccggacgggt agcgtccgt tgacgaaat gaagggtgg aacagaccg 104280  
ccaactgacg ccagctctcc aggtcgcaac agaggcagtc aaacaggtcg ggccgcatca 104340  
tctgctcggc gtacgcggcc cataggatct cgcgggtcaa aaatagatac aaatgcaaaa 104400  
acaaaacacg cgcagacga gcggtctctc ggtagtacct gtccgcgac gtggcgcgca 104460  
gcatttctcc caggtcgca tcgctcgcg geatgtgcgc ctggcgggtg agctgcccga 104520  
cgtggcgcg caggtaccgg tacaggccc agcagaagt ggccaacacg gttcgatagc 104580  
tctcctccc cgcctgtagc tcggcgtgga agaaacgaga gacgcttcg tagtagagcc 104640  
cgaggccgtc gcgggtggcc ggaagcgtag ggaaggccac gtcgcccgtg gcgcgaatgt 104700  
cgatttgggt gcgttcgggg acgtacgcgt cccccattc caccacatcg ctgggcagcg 104760  
ttgataggaa ttacactcc cggtacaggc cggcgttggc cggtagcgc gaaaacagat 104820  
cctcgttcca ggtatcgagc atggtacata gcgcggggccc cgcgctaaag cccaagtcgt 104880  
cgaggagacg gttaaagagg gcggcggggg ggacgggcat ggggtggggag gccatgagct 104940  
gggcctggct caggcgcgcc gttcgtaca gcgggggggc cgcgggggtg tttttgggac 105000  
ccccggccg gcgggggggc ggtggcgaag cgcctcgcg gttcatgtcg gcaaacagct 105060  
cgtcgaccaa gaggtccatt ggggtggggt gatacgggaa agacgatatc gggcttttga 105120  
tgcgatcgtc cccgccccc cagagagtgt gggacgccc acggcgcggg aagagaaaa 105180

---

-continued

---

ccccaaacgc gtttagaggac cggacgcgacc ttatggggggg aagtgggcag cgggaacccc 105240  
gtccgttccc gaggaatgac agcccgtggt ccccaccacg catttaagca acccgcacgg 105300  
gccgccccgt acctcgtgac ttccccccac attggctcct gtcacgtgaa ggcgaaccga 105360  
ggcgcgctgt ccaaccacc ccccgcacc cagtcgccgt ccccgctgga ttgggaaaca 105420  
aaggcacgca acgccaacac cgaatgaacc cctgttggtg ctttattgtc tgggtacgga 105480  
agttttcact cgacggggcg tctggggcga gaagcggagc gggctggggc tccaggtcgc 105540  
tcggtggggc gcgacgccgc agaagccct cagtcgccg tggccgcgtc gacgtcctgc 105600  
accacgtctg gattcaccaa ctctgtggcg cgctgaagca ggttttgccc ctccgagacc 105660  
gtcacgcgga tgggtggtgat gccaaaggagt tctgtgaggt cttcgtctgt gcgcgagcgc 105720  
gacatgtccc agagctggac gcgcccaatc cgggcatgca tggccgcag gcgcccagacc 105780  
gcgcgcgaga agacgcgctt gttaaagccg gccaccggg gggccatgg ccgctcgggg 105840  
tttggggggg cgtgctaaa gtgcagcttt ctggccagcc cctgcgcggg tgtcttgat 105900  
cgggttgccg ccgtcgacgc gggggcgtct gggagtgcgg cggattctgg ctgggcccgat 105960  
ttctgcccgc ggtggtctc gcgcccggg gccgcgggg ccttagtcgc caccgctgg 106020  
gttcgggggg cccggggggc ggtggtgggt gtgcgtccgg cccctccgga cccagcgggt 106080  
ggcgaggcgc cccgcgcagg ccccgggccc gacaaaaccg ccccgaaac gggacgccgc 106140  
gtccggggga cctccgggtg ttcgtcgtct tcggatgacg agccccgta gagggcataa 106200  
tccgactcgt cgtactggac gaaacggacc tcgccctct ggcgcgagcg tgtctgtagg 106260  
gcgcccaggg gggaggtgtc aggcggacta tcgggactcg ccatacctga agacggggtg 106320  
tagtacagat cctcgtactc atcgcgcgga acctcccgc gaccgcactt caccgagcgg 106380  
cgagaggtca tggttccacg aacacgctag ggtcggatgc gcggacaatt aggcctgggt 106440  
tcggacggcg ggggtggtgc aggtgtggag aggtcgcgag ataggggcgg cccgggagag 106500  
aagagagggg ccgcaaaacc cactggggat gcgtgagtg ccctctgtgg gcggtggggg 106560  
agagtcttat aggaagtgca tataaccaca acctatgggt ctaaccaatc cccaggggcc 106620  
aagaaacaga cagccccca acggtctcgg tttccgcgag gaaggggaag tctgggaca 106680  
ccctccacc ccaccctca cccacacag ggcgggttca ggcgtgccc gcagccagta 106740  
gcctctggca gatctgacag acgtgtgcga taatacacac gcccatcgag gccatgccta 106800  
cataaaaggg caccagggcc cccggggcag acatttggcc agtgttttg gtctcgcacc 106860  
gcgcgcccc gatcccacg gcgccgccct cctcgcggg cggetcccc gcggggccc 106920  
cgtctcccgc cgtaaggcg acgagcaaga caaacaacag gcccgcccga cagaccctc 106980  
tgggggggccc catcgtccct aacaggaaga tgagtcagtg gggatccggg gcgacccctg 107040  
tccagccgga cagcttgggt ggggggtacg atggcgactg gcacacggcc gtcgtaactc 107100  
gcgggggcgg agtcgtgcaa ctgaacctgg tcaacaggcg cgcggtggct tttatgccga 107160  
aggttagcgg ggactccgga tggcccgctg ggcgcgtctc tctggacctg cgaatggcta 107220  
tgccggctga cttttgcgcg attattcacg cccccgcgt agccagcccc gggcaccacg 107280  
taatactggg tcttatcgac tcgggttacc gcggaaccgt tatggccgtg gtcgtagcgc 107340  
ctaaaaggac gcgggaattt gccccggga cctgcgggt cgacgtgacg ttctgggaca 107400  
tctggcgac cccccggcc ctaccgagc cgatttcct gcggcagttc ccgcaactgg 107460

---

-continued

---

cgcccccccc tccaaccggg gccgggatac gcgaagatcc ttggttgag ggggcgctcg 107520  
gggccccaaag cgtgactacg gccctaccgg cgcgacgccc agggcggtcc ctctctatg 107580  
ccgcgagact gacgccggtt cagacggaac acggggacgg cgtacgagaa gccatcgct 107640  
tccttccaaa acgcgaggag gatgocggtt tcgacattgt cgtccgtcgc ccggtcaacc 107700  
tccccgcaaa cggcaccacg gtcgtgcagc catccctccg catgctccac gcggacgccc 107760  
ggccccgggc ctgctatgtg ttggggcggt cgtcgtcaa cgccccgggc ctctgggtcg 107820  
ttctacgcg ctggctccc gggcacgtat gtgcgtttgt tgtttacaac cttacggggg 107880  
ttctgtgac cctcgaggcc ggcgccaaag tcgcccagct cctggttgcg gggcgggacg 107940  
ctcttcttg gatcccccg gacaactttc acgggaccaa agcgttcga aactaccca 108000  
ggggtgttcc ggactcaacc gccgaaccca ggaacccgcc gctcctggtg tttacgaacg 108060  
agtttgacgc ggaggcccc ccgagcgagc gcgggaccgg gggttttggt tctaccggtg 108120  
tttagccac agctttgggt tcgttccggg caataaaaa cgtttgtatc gcatcttcc 108180  
tgtgtgtagt tgtttatgtt ggatgcctgt gggctctatca caccgcccc tccatccac 108240  
aaacacaaaa cacacgggtt ggatgaaaac acgcatttat tgacccaaaa cacacggagc 108300  
tgctcgagat gggccagggc gaggtgaggc tggggaggct gtaggtctg gaacggacac 108360  
gcggggacac gattccggtt tggggtccgg gagggcgctc ccgtttcggg ccgacggcgc 108420  
cagcgtaacc tccggggggc gcgtgtgggg gtgccccaaag gagggcgct ccgtcacc 108480  
aatccccccc gaccgggtt ccccggaac ccggaaggcg gagaggcaa gggccggtc 108540  
ggcgatggcc acatcctcca tgaccacgct actctcgcc atgctccgaa tagcctggga 108600  
gacgagcaca tcccgaggat tgtagcggc ccccacggac atgtacatct gcaggatgt 108660  
ggccatacac gtgtccgca ggcggcgc atgtctctga tgggcccga ccgcccctc 108720  
gatcgtgggg gctcgaacc cggggtggtg gcgcccagc cgttctaggt tcaccatgca 108780  
ggcgtgttac gtcggggcca aggcggggc cttcacgagg cgtcgggtgt cgtccaggga 108840  
ccccaggcgc tcacgagcgc tgatgggggc gggaaagtag cgttaacga ccgccaggc 108900  
ctctgcagc cgcggctccg cctccgagg cggaaaggcc gcgcgatca tctcatattg 108960  
ttctcgggg cgcgctccc agccacatat agccccgaga agagaagcca tcgcgggcgg 109020  
gtactggccc ttgggcgcgc ggacgcaatg gggcaggaag acgggaaccg cggggagagg 109080  
cgggcggccc ggactcccgt ggaggtgacc gcgctttatg cgaccgacgg gtgcgttatt 109140  
acctcttca tcgcccctct cacaaaactc ctactggggg ccgagccggt ttatatattc 109200  
agctacgacg catacagca cgatggccgt gccgacgggc ccacggagca agacagggtc 109260  
gaagagagtc gggcgctcta ccaagcgtcg ggcgggctaa atggcgactc cttccgagta 109320  
accttttgtt tattggggac ggaagtgggt gggaccacc aggcccggc gcgaaccgca 109380  
cccatgttcg tctgtcgtt cgagcgagcg gacgacgtc ccgcttaca ggacgcccctg 109440  
gcgacggga ccccgctaca accggaccac atcgccgcca ccctggacgc ggaggccacg 109500  
ttcgcgctgc atcgcaacat gatcctggct ctcaccgtgg ccatcaacaa cgccagcccc 109560  
cgcaccggac gcgacgccc cgcggcgagc tatgatcagg gcgctccct acgctcgtc 109620  
gtggggcgca cgtccctggg acaacggcgc cttaccacgc tatacgtcca ccacagggtg 109680  
cgcgtgcttg ccgctaccg caggcgctat tatggaagc gcgagagtcc cttctggtt 109740

---

-continued

---

cttagcaaat tcgggcccga cgaaaaaagc ctggtgctca ccaactcgta ctacctgctt 109800  
caggcccagc gtctgggggg gcgccccggc acgtacgacc tgcaggccat caaggacatc 109860  
tgcgccacct acgcgattcc ccaecgcccc cgccccgaca cgcgcagcgc tgcgtccctg 109920  
acctcgtttg cgcceatcac gcggttctgt tgcacgagcc agtacgcccg cggggcccgcg 109980  
ggggccgggt ttccgcttta cgtggagcgc cgtattgcgg ccgacgtccg cgagaccagt 110040  
gcgctggaga agttcataac ccacgatcgc agttgcctgc gcgtgtccga ccgtgaattc 110100  
attacgtaca tctacctggc ccattttgag tgtttcagcc cccccgcct agccacgcat 110160  
cttcgggccc tgacgaccga cgaccccaac cccgcccga gcaaggagca gccctcgccc 110220  
ctgggcaggg aggcctgga acaatttttt tgtcacgtgc gcgccaact gaatatcggg 110280  
gagtagtca aacacaacgt gacccccgg gagaccgtcc tggatggcga tacggccaag 110340  
gcctacctgc gcgctcgc gtacgcgccc ggggcccga cgccccccc cgcgtattgc 110400  
ggggccgtgg actccgccac caaaatgatg gggcgtttg cggacgccga aaagctcctg 110460  
gtcccccg ggtggcccgc gtttgcccc gccagtcgcc gggaggacac ggcgggccc 110520  
acgccccccc cacagacctg cggaaattgc aagcgcctcc tgagactggc cgccacggaa 110580  
cagcagggcc ccacaccccc ggcgatcgc gcgcttatcc gtaatgcggc ggtgcagact 110640  
cccctgccc tctaccggat atccatggc cccacgggac aggcatttgc cgcgctggcc 110700  
tgggacgact gggcccgc atacgcggac gctcgcctgg ccgaagcggc cgtgtccgcc 110760  
gaagcggcgc cgcaccccga ccacggcgcg ctgggcaggc ggctcacgga tcgcatccgc 110820  
gcccagggcc ccgtgatgcc ccctggcggc ctggatgccg gggggcagat gtacgtgaat 110880  
cgcaacgaga tattcaacgg gcgctgcca atcacaaaca tcctcctgga tctcgacatc 110940  
gccctgaagg agcccgtccc ctttcgccc ctccacgagg ccctgggcca ctttaggcgc 111000  
ggggctctgg ctgcggttca gctcctggtt cccgcccgc gcgtggaacc cgacgatat 111060  
ccctgttatt ttttcaaaag cgcattgccc cccggcccgc cgtccgtggg ttccggcagc 111120  
ggactcggca acgacgacga cggggactgg tttccctgct acgacgacgc cgtgatgag 111180  
gagtgggcgc aggaaccggc gcceatggac acatcccacg atccccgga cgacgaggtt 111240  
gcctactttg acctgtgcca cgaagtgcgc cccacggcgc aacctcgcga aacggattcg 111300  
cccgtgtgtt cctgcaccga caagatcgga ctgcgggtgt gcatgccctg ccccgcccgc 111360  
tacgtcgtcc acggttctct aacgatcgcg ggggtggcac gggcatcca gcaggcgggt 111420  
ctgttgacc gagatthtgt ggaggccatc gggagctacg taaaaaactt cctgttgatc 111480  
gatacggggg tgtacgcccc cggccacagc ctgcgcttgc cgtatthtgc caaaatcgc 111540  
cccacgggc ctgcgtgccc aagcgtcgtc ccagtgcttg tgatcccccc cgctgcaaa 111600  
gacgttcggc cgtttgccc gcgcacgccc gaccgcccgc gcttcattt tcacgcccgc 111660  
cccacctatc tcgcttcccc cggggagatc cgtgtcctgc acagcctggg tggggactat 111720  
gtgagcttct ttgaaaggaa ggcgtcccgc aacgcgctgg aacactttgg gcgacgag 111780  
acctgacgg aggtcctggg tcggtacaac gtacagcccg atgcgggggg gaccgtcag 111840  
gggttcgcat cggaactgct ggggcccgata gtcgctgca tcgaaaccca ctttcccga 111900  
cacgcccggc aatatcaggc cgtatccgtc cggcggggcc tcagtaagga cgactgggtc 111960  
ctcctacagc tagtccccgc tcgcggtacc ctgcagcaaa gcctgctgtg tctgcgcttt 112020

---

-continued

---

aagcacggcc gggcagctcg cgcacggcg cggacattcg tcgctgag cgtcggggcc 112080  
aacaaccgcc tgtgcgtgtc cttgtgtcag cagtgtttg ccgccaatg cgacagcaac 112140  
cgcctgcaca cgtgtttac cattgaecgc ggcacgccat gctcgcctc cgttcctcgc 112200  
agcaacctc aaccgtcgtc ttgataacgg cgtacggcct cgtgctcgtg tggtaacccg 112260  
tcttcggtgc cagtccgctg caccgatgta ttacgcggt acgcccacc ggcaccaaca 112320  
acgacaccgc cctcgtgtgg atgaaaatga accagaccct attgtttctg ggggccccga 112380  
cgcaccccc caacgggggc tggcgaacc acgcccata ctgctacgcc aatcttatcg 112440  
cgggtagggt cgtgccctc caggtcccac ctgacgccat gaatcgtcgg atcatgaacg 112500  
tccacgaggc agttaactgt ctggagacc tatggtacac acgggtcgt ctggtggtcg 112560  
taggtggtt cctgtatctg cgttcctcgc cctccacca acgctgatgt atgtttggcg 112620  
tcgtgagtc cgcacacaag atggtggccc cggccaccta cctcttgaac tacgcaggcc 112680  
gcatcgtatc gagcgtgttc ctgcagtacc cctacacgaa aattaccgc ctgctctcgc 112740  
agctgtcggc ccagcggcaa aacctggtc agttgtttga gacggaccgc gtcacctct 112800  
tgtaccaccg ccccgccatc ggggtcctc taggtcgcga gttgatgcta cgtttgtgg 112860  
cgtgggtct catcgtcggc accgcttca tatccgggg ggcctgtcgc atcacatacc 112920  
ccctgtttct gaccatcacc acctggtgtt ttgtctccac catcggcctg acagagctgt 112980  
attgtattct gcggcggggc ccgccccca agaacgcaga caaggccgc gccccggggc 113040  
gatccaaggc gctgtcgggc gtctcgggc gctgctgttc catcctcctc tcgggcatcg 113100  
cagtgcgatt gtgttatatc gccgtggtg ccggggtggt gctcgtggcg cttcactacg 113160  
agcaggagat ccagagggcg ctgtttgatg tatgacgtca catccaggcc ggcggaacc 113220  
gtaacggcat atgcaaattg gaaactgtcc tgtcttgggg cccaccacc cgacgcgtca 113280  
tatgcaaatg aaaatcggtc ccccgaggcc acgtgtagcc tggatcccaa cgacccgcc 113340  
catgggtccc aattggcctc cccgttacca agaccaacc agccagcgtc tccaccccg 113400  
cccgggtccc cgcggaagc gaaacgggta tgtgatatgc taattaaata catgccactg 113460  
acttatggg tctgattgtt cctgtctgt gccggaggtg gggcgggggc cccgcccggg 113520  
gggcggaacg aggaggggtt tgggagagcc ggcgccgca ccacgggtat aaggacatcc 113580  
accacccgc cgtggtggt gtgcagccgt gttccaaaca cggtcacgct tcgggtcctc 113640  
tccccgattc gggcccggtc gctcgtacc ggtgcgccac caccagaggc catatccgac 113700  
acccagccc cgacggcagc cgacagccc gtcattggca ctgacattga tatgctaatt 113760  
gacctcggc tggacctctc cgacagcgt ctggacgagg accccccga gccggcggag 113820  
agccgcgcg acgacctgga atcggacagc agcggggagt gttcctcgtc ggacgaggac 113880  
atggaagacc cccacggaga ggacggacc gagccgatac tcgacgcgc tcgcccggcg 113940  
gtccgccgt ctcgtccaga agacccggc gtaccagca cccagacgc tcgtccgacg 114000  
gagcggcagg gcccacaaga tcctcaacca gcgccccaca gtgtgtggtc gcgcccggg 114060  
gcccggcag cgtcttctc ccccgagcag cacgggggca aggtggccc cctccaacc 114120  
ccacgacca aagcccagc tgcccgggc ggacgcctg ggcctcgcag ggtcgggg 114180  
cgcggtggtc ccgggctcg cgatggtttg tcggaacccc gccggcgtgc cccagaacc 114240  
aatcgcaacc ctgggggacc ccgccccgg gcggggtgga cggacggccc cggcgcgcc 114300

---

-continued

---

catggcgagg cgtggcgcgg cagtgagcag cccgaccac cggaggcca gcgacacgg 114360  
ggcgtgcgcc aagcaccccc cccgctaag acgctggcga ttgcccccc gcccgcgga 114420  
ccccgcgcc cggccccgga gcgaaaggcg cccgcccgcg acaccatcga cgcaccacg 114480  
cggttggtcc tgcgctccat ctccgagcgc gcgggcgtcg accgcacag cgagagcttt 114540  
ggccgcagcg cacaggtcat gcacgacccc ttggggggg agccgtttcc cgcgcgaat 114600  
agccccggg ccccggtgct ggggggcca ggaggccct ttgacgccga gaccagacgg 114660  
gtctcctggg aaaccttggc gcgccacggc ccgagcctct atcgacttt tgcggcaat 114720  
cctcgggccc catcgaccgc caaggccatg cgcgactgcg tgcgcccga agaaaatttc 114780  
atcgaggcgc tggcctccgc cgacgagacg ctggcgtggt gcaagatgtg catccaccac 114840  
aacctgcgc tgcgccccca ggaacccatt atcgggacga ccgcggtgt gctggataac 114900  
ctcgcacgc gctcggggc cttctccag tgcactcga aggcgcgagg cctgtcggc 114960  
ctggacgaac tgtgttcgcg gcgcgctcg gcggacatta aggacattgc atcctcgtg 115020  
tttgtcattc tggccaggct gcacaaccgc gtcgagcgtg gcgcccggga gatcgactac 115080  
gcgacccttg gtgtcggggc cggagagaag atgcatttct acctccccg gcctcgcag 115140  
gcgggcctga tcgaaatcct agacacgcac cgcaggaggt gttcagtcg tgtctcgcg 115200  
ttgacggcca gtcacatcgt cgccccccg tacgtgcacg gcaaatattt ttattgcaac 115260  
tccctgtttt aggtacaata aaaacaaaac atttcaaaac aatcgccct cgtgttgctc 115320  
ttctttgctc atggccggcg gggcgtgggt cacggcagat ggggggggtg ggcggcgt 115380  
acggcctggg tggcgggagg gaactaacc aacgtataaa tccgtccccg ttccaaggcc 115440  
gggtgcatag tgccttagg agcttcccg cggggcgcac ccccccttt gcaactatgac 115500  
agcgaccccc ctcaccaacc tgttcttaag ggcgccggac ataaccacg tggcccccc 115560  
ttactgcctc aacgccacct ggcaggccga aacggccatg cacaccagca aaacggactc 115620  
cgcttgctg ggcgtgcgga gttacctggt ccgcccctcc tgtgagacca gcggcacaat 115680  
ccactgctt ttctttcgcg tatacaagga caccaccac accctccgc tgattaccga 115740  
gctcgcgaac tttcgggacc tggttaacca cccgcccgtc ctacgcgaac tggaggataa 115800  
gcgcggggtg cggctcgggt gtgcgcccgc gtttagcgtc gggacgatta aggacgtctc 115860  
tgggtccggc gcgtcctcgg cgggagagta cacgataaac gggatcgtgt accactgcca 115920  
ctgtcggat ccgttctcaa aaacatgctg gatgggggccc tccgcccgc tacagacct 115980  
gcgtccatc agctccagcg gcatggccc cgcgcccgca gagcatcgac gcgtcaagat 116040  
taaaattaag gcgtgatctc caaccccc atgaatgtgt gtaaccccc ccaaaaaaat 116100  
aaagagccgt aacccaacca aaccaggcgt ggtgtgagtt tgggaccca aagccctcag 116160  
agacaacgc acaggccagt atggaccgtg atacttttat ttattaactc acaggggcgc 116220  
ttaccgccac aggaatacca gaataatgac caccacaatc gcgaccacc caatacagc 116280  
atggccccc accacgccac aacagccctg tcgcccgtat ggggcatgat cagacgagcc 116340  
gcgcccgcg cgttggccc gtacagctc gcgcgaattg accctaggag gcccccacgc 116400  
gcccagttt tgcgttcgct cctggctcgc gggcgccaaa gcccggacg gctgttcggt 116460  
cgaacgaac gccacgacag tggcataggt tgggggggtg tccgacatag cctcggcgtc 116520  
cgtcgggagg cccgacaaga ggtccctgt gatgtcgggt gggcccaaa gcctggttc 116580

---

-continued

---

cggaagaaac agggggggttg ccaataaacc gccagggccca aaactccggc gctgcgcacg 116640  
tcgttcggcg cggcgccggg cgcgcgcagc ggctcgctgg gcggcttggc gtgagcggcc 116700  
ccgctccgac gcctcgccct ctccggagga ggttgccgga attggcacgg acaacagggg 116760  
cccagcagag tacgggtggag gtgggtccgt gggggtgtcc agatcaataa cgacaaacgg 116820  
cccctcgttc ctaccagaca agctatcgta ggggggcccgg ggatcagcaa acgcttccc 116880  
cgcgctccat aaaccgcggt cgggttgccg cgctccgaa gccatggatg cgcgccaaag 116940  
ccacgactcc cgcgcgctag gtccttgggg taatggaaaa ggcctactc cccatccaag 117000  
ccagccaagt taacgggcta cgccttcggg aatgggactg gcacccggc ggattttgtt 117060  
gggctggcat gcgtcgccc accgagggcc gcgtccacgg gacgcgcctt ttataacccc 117120  
gggggtcatt cccaacgatc acatgcaatc taactggctc cctctcccc cctctcccc 117180  
tctccccccc tctcccctc cccccctct cccctctccc cctctctccc ctctcccc 117240  
ctctcccctc tccccccctc tcccctctcc ccccctctcc cctctcccc cctctcccct 117300  
ctccccccct ctcccctctc cccccctctc cctctctccc tctgtctttt ccccgtagca 117360  
cccagcgtg gggcgctggc tgccgggagg ggcgcggat gggcgggctt acttggtttc 117420  
cgcccccccc cccccccccc cgaaccgccc cgcgcgcttt gcccccttt gatccctgc 117480  
taccccaac ccgtgctggt ggtgcgggtt ggggggggat gtggcgggg gtgcgcggga 117540  
ggtgtcggtg gtggtggtg tggtggtagt aggaatggtg gtgagggggg gggggcgctg 117600  
gttggtcaaa aaagggagg cggggggcgc gcagaccgac ggcgacaacg ctcccggcg 117660  
gccgggtcgc ggctcttacg agcggcccgg cccgcgctcc cccccccgg gccgtgtcct 117720  
tgctttcccc cegtctcccc cccccccgccc ttctctctct cctctctggt ttccaaaacc 117780  
cgcaccaacc ggcccggccc ggcggggccc ggcgggcca cgcgcgcca cccaccacc 117840  
tcgggatacc cagccccggt cccccgtccc ccgggggccc ttatctccag cgcgccgtec 117900  
ggcgcgccc cccccccgccc taaaccccat cccgccccgg ggaccccaca tataagcccc 117960  
cagccacacg caagaacaga caccgagaac ggctgtgttt atttaataa accaatgtcg 118020  
gaataaaca acacaacac cgcgcgaggg gggacggagg ggacggagg agggggtgac 118080  
gggggacggg aacagacaca aaaaacaaca caaaaaaca ccaccaccg acaccaccac 118140  
cccagtctcc tcgcttctc caccaccacc caccgcccga ctgagcccgg tcgatcgacg 118200  
agcaccaccg cccacgccc cgcctctgcc ccggcgaccc cggcccgcga cgateccgac 118260  
aacaataaca accccaacgg aaagcggcgg ggtgttgggg gaggcgagga acaaccgagg 118320  
ggaaacgggg atggaaggac ggaagtgga agtcctgata cccatcctac acccccctgc 118380  
cttcaccct ccggccccc cgcagctccac ccgcccggcc gctaccgaga ccgaacacgg 118440  
cggcgcgcgc agccgcgcca gccgcgcgc acaccgcaga gccggcgcgc gactcaciaa 118500  
gocgcagagg cagaaaggcc cagagtcatt gtttatgtgg ccgcgggcca gcagacggcc 118560  
cgcgacaccc cccccccgccc cgtgtgggta tccggcccc cgcgccgccc cgttccatta 118620  
agggcgcgcg tccccgcgag atatcaatcc gttaatgtct ctgcagacag gggcaccgcg 118680  
cccggaaatc cattagccc cagacgagga aaataaaatt acatcaccta cccacgtggt 118740  
gctgtggcct gtttttgcgt cgtcatctca gcctttataa aagcgggggc gcggccgctgc 118800  
cgatcgcggg tggtgcgaaa gactttccgg gcgcgtccgg gtgcgcggc tctccgggccc 118860

-continued

---

ccccgcagc cggggcgcc aagggcgctc ggcgacatcc tccccctaag cgccggcccg 118920  
ccgctggctc gtttttctgt tttccccgtt tccgggggtg tgggggttgc ggtttctgtt 118980  
tctttaaacc gctcgggggtg tttttcgttc cgtcgcggga atgtttctgtt cgtctgtccc 119040  
ctcacggggc gaaggcccg ctagggcccg gacgaggggc ccccgaccgc ggcgggtcccg 119100  
gccccgtccg gacccgctcg ccggcaccgc acgcgaaaaa ggcccccccg aggtttttcc 119160  
gggttcccg cccggggcct gagatgaaca ctccggggtta ccgccaacgg ccggcccccg 119220  
tggcgcccg gcccggggc ccggcggacc caaggggccc cggccccggg ccccaaacg 119280  
gccccggca tgcgctgtgg ttttttttc ctccggtgtc tgcggggtc catcgccttt 119340  
cctgttctc cttctcccc cccctctctt cccccact accctctcc ctcccttct 119400  
ccccgctat cccactcgtc gagggcgccc cgtgtcgtt caacaagac gccgcgttcc 119460  
caggtaggt agacacctc ttctcccaa tagaggggg ggacccaac gacagggggc 119520  
gccccagagg ctaaggtcgg ccaccgact ccggggtgg ctcgtgttac agcacaccag 119580  
cccgttctt tccccccct ccacccttag tcagactctg ttacttacc gtcgaccac 119640  
caactgccc cttatctaag ggcggctgg aagaccgcca gggggtcggc cgtgtcgtc 119700  
gtaaccccc acgccaatga cccactact ccaagaagg atgtgtccc ccccgctgt 119760  
gtttttgtc ctggctctc atgctgggt cttactgct ggggggggg agtgccgggg 119820  
agggggggg tggaagaaa tgcaccgccc gtgtgtacc cccctaaagt tgttctaaa 119880  
gagaggtac ggaggagtg cgggtgccc gggaccggg tgatctctg cagcggggg 119940  
tgggaaggt cgggggagg ggggatggg taccgccc cctggcccg cgggtgccc 120000  
tgccttgc caccaacc cgtccccg cgggtctct agaagacc cccccctcc 120060  
ttcataccac cgagctgccc ggggtgtgg ttgtaacca acaccgcat cccctcgtc 120120  
cctgtgattc tctggctgca ccgactctt gttttctaac tatgttctg tttctgttc 120180  
cccccccc acccctccc cccaccccc aacaccacc tctgtggtg gcccgacc 120240  
ctttggggc ccccgctccc ccccgccc cctcccatc tttgttccc tatagtgtg 120300  
ttaaccccc ccgcccctt gggcgccc aggccaggc agtcggggg gccagggct 120360  
cgcgaaact taacaccac acccaacca ctgtggttct ggctccatg cagtggcagg 120420  
atgcttccg ggatcgggtg tcaggcagcc cggcccgcg ctctgtggtt aacaccagag 120480  
cctgccaac atggcacc cactcccac cccccact cccagcacc cccactcca 120540  
cgaccccc ctcccagca cccccact caccgcccc cactcccag cccccact 120600  
cccagcacc cccactcca cgaccccc ctcccagca tccccgat acatccaaca 120660  
cagacaggga aaagatacaa aagtaaacct ttatttcca acagacagca aaaatccc 120720  
gagtttttt ttattaggc caacacaaa gaccgctgg tgtgtgtgc ccgtgtctt 120780  
cacttttccc ctccccgaca cggattggt ggtgtagtgg gcgcccag agaccacca 120840  
gcgcccagc cccccctcc cacaacacg gggggcgtc cttattgtt tccctcgtc 120900  
cgggtcagc cccccgctc cccggacc ggggtcggg accgaggt gcggaagtcc 120960  
agggcgcaca ctagggtgct ctggtcgaac agcatgttc caccggggg catccagagg 121020  
ctgttccact ccgacgggg ggcctcggg tactcgggg gcacacgtg gttaccggc 121080  
gtctcgggga gcagggtgc cggctccag cggggacc cggcccag ccgggtcgc 121140

-continued

---

atgtttcccg tctgggccac caggaccacg tacgccccga tgttcccgt ctccatgtcc 121200  
aggatgggca ggcagtcccc cgtgatagtc ttgttcacgt aaggcgacag ggcgaccacg 121260  
ctagagaccc ccgagatggg caggtagcgc gtgagccgc ccgcggggac ggcgccgaa 121320  
gtctccgctt ggcgctcttt ccgggcacac ttctctggcc ccgcgggccc agaagcagcg 121380  
cgggggcca gggaggtttc ctctgtctc cctcccaggg caccgacggc ccgcccagag 121440  
gaggcggaag cggaggagga cgcggccccc gcgggggaag agggcgcccc cgcgggggtc 121500  
ggggccgagg aggaagagc agaggaggaa gaggcggagg ccgcccagga cgtcaggggg 121560  
gtcccgggcc caccctggcc gcgccccccc ggccctgagt cggagggggg gtgctctccc 121620  
gccctcttgg cccctcggcc gcgagggggg ggacgcgtgg actgggggga ggggttttcc 121680  
tggcccgacc cgcctcttct cctcggacgc accgcgcct cctgctcagc agaggcgcg 121740  
gaggggagcg ggcgcgccgc ggagggggcg gcgcccggg agggcccgtg cccaccctcc 121800  
acgcccggcc cccccagacc gcgcccacc gtgcacgcgc ccggcacag actctgttct 121860  
tggttcgcgg cctgagcccg ggacgagtg cactggggca caggcgcgcc gtccgcgggg 121920  
cggcgggccc gctccgcccc gggggccggg gcgcgggggc cgggccccgg agggcgct 121980  
cgcacgcacg gggccacggc gcgcccgggg cgcgcccggc ccgacgcggc cgcggacgcg 122040  
ggggggcccg ggcggggggg ggagcctggc atgggcgccc cggggggcct gtggggagag 122100  
gccggggggg agtcctgat cactatgggg tctctgttgt ttgcaagggg ggcgggtctg 122160  
ttgacaaggg ggcctctccg gccctcggc cgcctcggc ccgcttcaac aaccccaacc 122220  
ccaaaccaca ccccccgga gggccagacc gcccccgcc gcgcccggc tcgactgag 122280  
cgggagccgc cgcgcccgtc gctgttgggt gtggtgttgg tgtactgct gccgtgtggc 122340  
ccgatgggag ccgagggggg cgtgtctcca gcccgggccg gctggggggc tgcgtgagac 122400  
gccccgccc tcacgggggg gcgcccggc cctctcgtg ggggggcgcg ggcgctcccg 122460  
cggggggcgg ggcgtacgta gtctgctgca agagacaacg gggggcgcca tcaggttacg 122520  
ccccctccc ggcggccctt ttctcgcgcc gcccgctat tctcctccc cccccctcc 122580  
tctcctcct cccccagggt ccttgcggcc ccccgctca ccgtctcca ggtcgtctc 122640  
atcctcgtcc gtggtgggct ccgggtgggt gggcgacagg gccctcaccg tgtgcccccc 122700  
cagggtcagg taccggggg gaaaccgctg attgcccgc cagataaagt ccacggcct 122760  
gccccctct acggcctcct cggcctccat gcgggtctgg gggtcgtca cgtcgggat 122820  
gggtctgaac gaaccgctgg gcgtcacgcc cactatcagg tacaccagct tggcgttga 122880  
cagcgggag gtgttcgca attgcatcca ggttttcatg cacgggatgc agaagcggtg 122940  
catgcacggg aagggtctgc agcgcagggt gggcgcgatc tcacccgtgc acacggcgca 123000  
cacgtcgcct cgtctctccc ccccgctctc tcgagggggg gcgccccgc aactgcccgg 123060  
gtcttctctg cggggggggc tccccccga gaccgcccc ccatccagc cctgcccggc 123120  
cagcagcccc gtctcgaaca gttccgtgtc cgtgctgtcc gcctcggagg cggagtcgtc 123180  
gtcatggtg tcggcgtccc cccgcccccc cacttcggtc tccgctcag agtcgtgct 123240  
gtccggcagg tctcgtctgc agggaaacac ccagacatcc gggcggggct aaggggaaaa 123300  
aagggggggc ggtaaagatg gggggggatt tcccgcgtca atcagcacc cagagttccc 123360  
cctctcccc ccccgctca caaagctctg cccccctgct ggcctcggaa gaggggggag 123420

---

-continued

---

aaaggggtct gcaaccaaag gtggtctggg tccgtccttt ggatcccagc cccctcttctt 123480  
ccctcttctc ccgcccctca gacgcaccgg agtcgggggt cccacggcgt cccccaaata 123540  
tgggggggcg ctctctccca ccccctaga tgcgtgtgag taaggggggc ctgctatga 123600  
gtcagtgggg accaacgccc caacacggcg accccggtec ttgtgtgttt gttgtggggg 123660  
cgtgtctctg tgtatgagtc aggggggtcc acggcgaccc cgggcccctgc gtctgagtca 123720  
aaggggcat gtgtatgtgt tgggggtctg tatatataaa gtcaggggggt cacatggcga 123780  
cccccaacag ggcgaccccc gtccctgtat atataagggtc aggggggtcc gcacccccta 123840  
acatggcgcc cccggtcctt gtatatatag tgtcacgggg ttccacgccc cctaacatgg 123900  
cgccccaa ca tggcgcccgg ctcccgtgta tgagtggggg tcccccaaca tggcgcccgg 123960  
ttccagtga agggctgggg gtcccccaac atggcgcccc ccaatatggc gcccccaat 124020  
atggcgcccc agacatggcg cccggcccct cactcgcgcg tgggggcggc cctcaggcgg 124080  
gcggtactc gctccggggc ggggctccat gggggtcgtg tgcggctgga gggctcggga 124140  
cggagggtcc ctgggggtcg caacgtaggc ggggcttctg tgggtatgcg gagagggggc 124200  
ggcccagtc tgcctggctg ctgctctcgc ctccagtgcc cgagggtgca atgcgaccag 124260  
actgtcgggc cagggtaac ttataccca cgccttccc ctccccaaag gggcgccagt 124320  
gacgattccc ccaatggcgg cgcgtcccag gggaggcagg cccaccgagg ggcggccccg 124380  
tccccgggga ccaaccggc gccccc aaag aatataatta gcatacagg cccggcccc 124440  
gatttggggg ccaaccggc tgtccccaa agaaccat tagcatgccc ctcccgcga 124500  
cgcaacagg gcttggcctg cgtcgtgccc ccggggcttc ccgcctccc gaagaaactc 124560  
attaccatac ccggaacccc aggggaccaa tgcgggttca ttgagcgacc cgcgggcaa 124620  
tgcgagagg gcccgtgtgt ccgcaaaaa agcaattagc ataaccgga accccagggg 124680  
agtggttacg cgcggcgagg gaggcgggga ataccgggtg tgcattaa gggcgcggg 124740  
aattgccgga agcgggaagg ggcggcgggg ccgcccatta atgagttct aattaccata 124800  
ccgggaagcg gaacaaggc tcttgcaagt ttttaattac cataccggga agtggcgggc 124860  
ccgcccatt gggcgtaac tcccgcaca tgggcggggc cccgaagact cggcgagcgc 124920  
tgggtggcgg gggcccggc cgcgtggcgg ccgcatggt ccagtcggc ccccgaggcg 124980  
ggcccctcgt tggggcggg ctggctccaa gcgtatata gcgaggctcc tgccatcgtc 125040  
tctccggaga gcggcttggg gcggagctcc cgggagctcc gcggaagacc caggccgct 125100  
cgggtgtaac gttagaccga gttcgcggg ccggctcgcg gggccagggc cggggcacgg 125160  
gcctcggggc ccaggcagg cccgatgacc gcctcggcct ccgcccaccg gcgcccgaac 125220  
cgagcccggc cggcccgcgc cggggcccac gagccgaggc gcgcccaggc ggcggccgag 125280  
gcccagacca ccagggtggc caccggagc tggggcgaga agcgcaccgc cgcgggggtc 125340  
gccccgggtc cgggggtcgc gggggtcgcg ggggtcgcgg ggggtcggc cccccctcc 125400  
ccgcccgcgc gtcgaggcg caggcgcgc aggtgctccg cggtagcgcg caggcggagg 125460  
gcgagggcgc gcggaaggcg gaagggggcg gaggggggtg gggaggggtc agccccccc 125520  
ccccggcca cgcggggcgg tggggggcgg ggcggggggg cggcggggtt gggcggggcc 125580  
tctggcgccg gctcggggcg ggggctgtcc ggccagtcgt cgtcatcgtc gtcgtcggac 125640  
gcggactcgg gaacgtggag ccaatggcgc agcagcagc aacaagaagg cgggggcca 125700

-continued

---

ccggcggggg ggcggcgccg ggcggccgcg ggcgcgctcc tgaccgcggg ttccgagttg 125760  
ggcgtggagg ttacctggga ctgtgcggtt gggacggcgc ccgtgggccc gggcgcccg 125820  
ggcggggggg ggcgcgatg gcgcgccgcg cgggccatgg agacagagag cgtgcccggg 125880  
tggtagagtt tgacaggcaa gcatgtgctg gcagaggcga gtagtgcttg cctgtetaac 125940  
tcgctagtct cggcccgggg gggcccgggc tgcccgcgc caccgcttta aagggccgcg 126000  
cgcgaccccc ggggggtgtg ttttgggggg ggcccgtttt cggcgtctgg ccgctcctcc 126060  
ccccgctect cccccgctc cccccccgc tctcccccc gctcctcccc ccgctcctcc 126120  
ccccgctect cccccgctc cccccccgc tctcccccc gctcctcccc ccgctcctcc 126180  
ccccgctect cccccgctc cccccccgc tctcccccc gctcctcccc ccgctcctcc 126240  
ccccgctect cccccgctc cccccccgc tcccgcgcc ccgccccca cgcggccgc 126300  
gcgcgcgac gccgcccga ccgcccgcg ctttttttgc gcgcgcgcg gcccggggg 126360  
ggcccgggct gccacagtg aaaccaacag agcacggcgc actccgcacg tcacacgtca 126420  
cgteatccac cacacctgac caacaacaca actcacagcg acaactcacc gcgcaacaac 126480  
tcctgttct catccacag tcaccgcgca cctcccgtc ctcagacgt accccggcgc 126540  
aacacaccgc tctgteta caccaccgc cctccccag cccagccct cccagcccc 126600  
agccctccc gggcccagcc cccccggcc ccagccctcc ccggcccag cctccccgg 126660  
ccccagccct ccccgcccc agccctccc gggcccagcc cccccggcg cgtcccgcgc 126720  
tccctcgggg gggttcggg atctctacct cagtgcgcgc aatctcaggt cagagatcca 126780  
aacctccgg gggcgcccc gcaccaccac cgcctccgc cccctccgc cctccgccc 126840  
ctcccgcctc tcgccccctc ccgcccctcg ccccccccg cccctccgc cctccgccc 126900  
ctgccccct cccgcccctc gccccctcc gccctccgc cctccgccc cctgcccc 126960  
tcccgcctc cgcccccctc gcgccctcg cccctccgc cctccgccc ctcgcccc 127020  
tcgccccctc ccgcccctcg ccccccccg cccctccgc cctccgccc ctcgcccc 127080  
cccccccc cccccctcc cccccctcg cctccgccc cctgcccc tcccgcctc 127140  
cgaataaaca acgctactg aaaacttaa caggttgttg ccgtttattg cgtcttcggg 127200  
tctcacaagc gcccccccc gtcccgccc gttacagcac ccgctcccc tcgaacgcgc 127260  
cgcgctgctc ttgctcccag gcgcccctcc agtccacaac ttcccgcgc gggggcgtgg 127320  
ccaagccgc ctcgcccc agcacctcca cggccccgc cgcggcagc acggtgccc 127380  
tgcgcccct ggcgagggc cagcgaatcc cggggcgcg cggcggcag gccccgggc 127440  
cgtcgtcgtc gccgcgagc accaggggg gggcgtcgtc gtcgggctcc agcagggcg 127500  
ggcgcaaaa gtcctccgc ggcgcgcgc accgggccc gcccgcgcgc accgctcgc 127560  
gccccagcgc caegtacag ggcgcagcg gcgcccagc gcccagcgc gcgcaggcg 127620  
cgtgcgagtg ggcctcctc tcgcagaagt ccggcgcgcc gggcgccatg gcgtcgggtg 127680  
tccccaggc cgcgcccgg ccgtccagcg cggcagcac ggcggggcg tactcgcgcg 127740  
gggacatgg caccgcgctg tccgggccc agcgcgtgcg cacgcgtag cgcacgttg 127800  
cgcgcccga caggcgagc ggcggcgct cggggtacag gcgcgctgc gggcctcca 127860  
cgcgcccga gacccccgg ccgaacacgc ggcggggc cagcaccgtg cggcgaggt 127920  
ccccgcgcgc cggccagcgc ccggcgccact gcacggcggg cagcagctc cagccaggt 127980

-continued

---

aggcgtgctg ccgacgaccc cggggcccggt cggcgggcca gtcgacggcg cgcacgggtg 128040  
tgaccacgat gagccgcccgg tcgcccggcgc tggcgagcag cccagaaac tccacggccc 128100  
cggcgaaggc caggtccccg gtggacagca gcagcacgcc ctgtgccc agcgcgcgaca 128160  
cgtcgggggc gcccgtccaa ttgcccggcc aggcggccgt gtcgggcccg cacagccggt 128220  
tggccagggc cgcacgacag caggacagcc cgcgcgctc ggccgaccac tccggcggcc 128280  
ccccgaggc cccgcccggc gccaggtcct cgcggcggcag cggcggagtac agcaccacca 128340  
cgcgcacgtc ctcggggctg gggatctggc gcatccaggc cgcgatgccc cgcagcgggc 128400  
ccgaggcgcg cagggggcca aagaggcggc ccccggcggc cccgtggggg tgggggttat 128460  
cgtcgtcgtc gcccggcccg cagcggcctc gggcggcggg ggcgggcccg ggcacccgcg 128520  
cggcgatcga ggcaggggcc cgcgggtcaa acatgagggc cggtcgcccag gggacgggga 128580  
acagcgggtg gtcgctgagc tcggccacgg cgcgcgggga gcagtaggcc tccaggcggg 128640  
cggccgcggg cgcgcgctg tggctgggccc cgggggctg ccgcccagcc cgcggcggg 128700  
ggtcggggcc ctcggcgggc cggcgcgaca cggccacggg cgcgggcccg gctcgcgccc 128760  
cggcggcccg gggcgcgccc ggtcgggccc gggcgggctc gggcccggg ggcgtggagg 128820  
ggggcgcggg cgcggggagg gggcgcgccc cgtccgagcc gggggcgtcc ggcgcccctc 128880  
tcttcgtctt cgggggtcgc gggccgccc ctcggggcgg ccggggcccg cggggactct 128940  
tgccttgccg cccctcccgc ggcgcgccc aggcggcggc ggcgcgccc cgcctgcccg 129000  
cgtccggtgc gctggcccgc gccccagca gggggcgcag gctctggtt tcaaacagca 129060  
ggtccgcccg ggcggcggcc gcggagctc gcagggcggc gtcccggcc agcgcggggc 129120  
ccagggcccc ggcgacaggc ctcacggcgc gcacggcggc cagggcggcc tcgctgcccg 129180  
cggccacgcg caggtccccg cgcagggcga tgagcaccag cgcgtcgcgc acgaaccgca 129240  
gctcgcgcag ccacgcgccc aggcggggcg cgtcggcgtg cggcggcggc ggggaagcgg 129300  
ggcccgcggg tccctcccgc cgcggggggc tggcggggcc gggcccggcc agcccgggga 129360  
cggcgcgccc gtcgcccgtc aagccctcgg ccagcgcctc caggatccc cgcagggcgg 129420  
ccaggcactc gaaggccacg cggcggcctc gggcgcgccc cccggcgtc tcgctggcgt 129480  
cggcgtggcg ggcggcgtc gggctgctgc ccccgcggg ggaggcggg cgcggcggaca 129540  
gccgcccag ggcggcggag atccccgcgg cgcgtaacc ggcgggccc ggcgctcgc 129600  
ccggtgcccg ggcggcgcag cggcgcagcc cctcgtcctc tgcgcccggc ccggggctcc 129660  
ccgcccggcc cgtcagcgc cgtttctcgc gcgccaacag gggcgcgtag ggcggcgcga 129720  
ggctggtcag caggaagccc ttctgcccgc ggtcgtatc ggcgctcatg gccacggcgg 129780  
ccgcccgtg cgcaggccc cagccgaagc ggcggcccgc catggcgtag cccaggtggg 129840  
gcacggcccg cgcacgctg ccggtgatga aggagctgct gttcgcgccc ggcggcggaga 129900  
tccggaagca ggcctggtcc agcgcacgct ccccggggac cagcgcggg ttctggagcc 129960  
acccatggc ctcgcccgtc ggggtgtaca gcagcggcgt gatcaggggc tactgctgcg 130020  
cggcgtcgc cagctcggg gccccacagc ccgcccggg gcccgaggcc tcgaaccggc 130080  
gtcgcgccc ctcgcccctc ggcgcccggc agaggcccgg cggcgtgctc cccagcgcgc 130140  
cgtacagcac cgcggcccgg ggcggggggc cggcggcccg ccacggctcc ccgctgacgt 130200  
accgctcgc atagcgcgag tagaaggcgc cggaggtcgc gtcggcgtcc agctcagccc 130260

-continued

---

gcccgggctg cccggccgtg aagcggcccc tggcgtcgcg gccggccacc gccgcgcggg 130320  
cccggcggcg ctccgatgccc cccgcggagg ccgcgggggt cctcgcgcc gcccggggt 130380  
tgggcgcggc ctccgagagg gggggtggcc cggcggggg cggcgtccgc cccggggctg 130440  
ccggcgcgcg gtcgacgga ccccgcgca cggccgcgc ctccggtgcg tggcggccg 130500  
cgtcgttgc gtcgtcgtcc tcgtcctcgt cggacgacga ggacgaagag gatgcggacg 130560  
acgaggacga ggaccggag tccgacgagg tcgatgacgc cgatggccgc caccggcct 130620  
gacgacgtct cccggcggcg tggccggcg ggcgcggcga caggcggctc gtggggctcc 130680  
gatacgcgcc gcgtagcggg gcctcccgt cgcggcccc ggccggggcc cggtcgcccg 130740  
cggcgtcggc tgcgtcgtc tactcgtccc cgtcatcgtc gtcggctcga aaggcgggg 130800  
tcccggcggc cagggccgc gggcggggc tcgggatcgt ccggacggcc tcctctacca 130860  
tggaggccag cagagccagc tgcgcggcg agacggcgc cccggcgtcc tcgccggct 130920  
cggtgccgc cgcggggggc ctcccgtccc gccggcgcg tcgaggtcg tgggggtggt 130980  
cggggctcgt gtcggggctg tcccgcctt cctccgtctc cgcgccccac ccgagggccc 131040  
cccctcgtc gcggtctggg ctccgggtg gcggcggccc gtcggtgggg cccggggagc 131100  
cggggcgcgt cttgttctcc gacgccatcg ccgatcggg gcgatacctc ggggatacgg 131160  
ctgcgacggc ggacgtagca cggtaggtca cctacggact ctccgatggg ggagggggcg 131220  
agaccacgg accccgacga ccccgcctt cgcgcggaa ctacgcgga ccggtccgatg 131280  
cttgggtggg aaaaaggaca gggacggcgc atcccctcc cgcgcttctg ccgcgtatcg 131340  
gcgtcccgc gcggcgagcg tctgacggc tgtctctggc ggtcccgcg ccgggtcgtg 131400  
atccgtgctg gcagccgcgc tccgtgtgga cgateggggc gtcctcgggc tcatatagtc 131460  
ccagggccgc gcgggaagga ggagcagcgg aggcgcggc cccccgccc ccccgcggg 131520  
ccccccgca acggaattcc attatgcacg accccgcccc gacgcggca cgcggggggc 131580  
ccgtggcgc ggcccgttgg tcgaaccccc ggccccgccc atccgcgcca tctgcatgg 131640  
gcggggcgc agggcgggtg ggtccgcgcc ccgccccgca tggcatctca ttaccgccc 131700  
atccggcgtt tccgcttcc gttccgatg ctaacgagga acgggcaggg ggcggggccc 131760  
gggccccgac ttcccgttc ggcgtaatg agatacagac cccgcgcgcc cgttggccgt 131820  
ccccggccc cccggtccc cccgcccggc gccgggacca acgggacggc gggcgccca 131880  
agggccgccc gccttgcgc cccccattg gccggcgggc gggaccgccc caagggggcg 131940  
gggcgcggc gtaaaagaag tgagaacgcg aagcgttcgc acttcgtccc aatataata 132000  
tattattagg gcgaagtgc agcactggcg ccgtgcccga ctccgcgcc gcccggggg 132060  
cggccccggc cggcgggggg cgggtctctc cggcgcacat aaaggcccgc cgcgaccgac 132120  
gcccgcagac ggcgcggccc acgaacgacg ggagcggctg cggagcagc ggaccgggag 132180  
cgggagtcg agagggccgt cggagcggc ggcgtcggca tcgcgacgcc ccggctcggg 132240  
atcgggatcg catcggaaa ggcacgcggc acgccccggg gaaagaccg cccaccccac 132300  
ccacgaaaca caggggacgc accccgggg cctccgacga cagaaacca ccggtccgcc 132360  
tttttgcac gggtaagcac cttgggtggg cggaggggg ggggacgcgg gggcgaggga 132420  
ggggggacgc gggggcggag gaggggggc gccggggcg agggggggg acgccccggc 132480  
ggagggggg ggcgcgggg cgggaggag gggctcacc gcgttcgtgc cttcccgcag 132540

-continued

---

gaggaacgtc ctctgtagagg cgaccggcgg cgaccgttgc gtggaccgct tctgctcgt 132600  
cggcgggggg gaagccactg tggctctccg ggacgttttc tggatggccg acatttcccc 132660  
aggcgctttt gcgccttgtg taaaagecgg gcgtcccgtc ctcegatccc cgccccggg 132720  
cacgcgcaag cgcaagcgcc cttcccggcc cctctcatcg gagtctgagg tagaatccga 132780  
tacagccttg gagtctgagg tcaaatccga gacagcatcg gattcgaccg agtctgggga 132840  
ccaggatgaa gccccccgca tcggtggcgg tagggccccc cggaggcttg gggggcggtt 132900  
ttttctggac atgtcggcgg aatccaccac ggggacggaa acggatgcgt cgggtgctgga 132960  
cgaccccgac gacacgtccg actggtctta tgacgacatt cccccacgac ccaagcgggc 133020  
ccgggtaaac ctgcggtcga cgagctctcc cgatcggcgg gatgggggta tttttcctaa 133080  
gatggggcgg gtccggtcta cccgggaaac gcagccccgg gccccacccc cgtcggcccc 133140  
aagcccaaat gcaatgctac ggcgctcggg gcgccaggcc cagaggcgga gcagcgcacg 133200  
atggaccccc gacctgggct acatgcgcca gtgtatcaat cagctgtttc gggctctgcg 133260  
ggtcgccccg gacccccacg gcagtgccaa ccgcctgcgc cacctgatac gcgactgtta 133320  
cctgatggga tactgcccag cccgtctggc cccgcgcacg tggtgccgtt tgctgcaggt 133380  
gtccggcggg acctggggga tgcacctgcg caacaccata cgggaggtgg aggctcgatt 133440  
cgacgccacc gcggaacccg tgtgcaagct tccttgtttg gagaccagac ggtacggccc 133500  
ggagtgtgat cttagtaatc tcgagattca tctcagcggc acaagcagtg atgaaatctc 133560  
cgatgccacc gatctggagg ccgcccggtc ggaccacacg ctccggtccc agtccgacac 133620  
ggaggatgcc ccctcccccg ttacgctgga aaccccagaa ccccgcggtt ccctcgtctg 133680  
gcgtctggag gatgagtttg gggagtttga ctggaccccc caggagggct cccagccctg 133740  
gctgtctgcg gtcgtggcgg ataccagctc cgtggaacgc cggggcccat ccgattctgg 133800  
ggcgggtcgc gccgcagaag accgcaagtg tctggacggc tgccggaaaa tgcgcttctc 133860  
caccgcctgc cctatccgt gcagcgacac gtttctccgg ccgtgagtcg ggtcggcccc 133920  
acccccttgt atgtcccaa aataaaagac caaaatcaaa gcgtttgtcc cagcgtctta 133980  
atggcgggaa gggcgagagg aaacagacca ccgagcatg ggggggtgtt gggggtttat 134040  
tggcaccggg ggctaagggg tggtaaccgg atagcagatg tgaggaagtc ggggcccgtc 134100  
gcccggaacg gcgatcagag ggtcagtttc ttgcccagca cggcccggcg atgtgggttg 134160  
ctcgtctggg acctcgggca tgcccataca cgcacaacac ggacgcccga ccggatggga 134220  
cgctgtaagg gggcctgggg tagctgggtg gggtttgtgc agagcaatca gggaccgcag 134280  
ccagcgcata caatcgcgct cccgtccggt tgtcccgggc agtaccacgc cgtactggtta 134340  
ttcgtaccgg ctgagcaggg tctccagggg gtggttgggg gcccggggga acgggggtcca 134400  
cgccacggtc cactcgggca aaaaccgagt cggcacggcc cacggttctc ccaccaacgc 134460  
gtctggggtc ttgatggcga taaatcttac cccgagccgg attttttggg cgtattcgag 134520  
aaacggcaca cacagatccg ccgcgcctac caccacaag tggtagaggc gaggggggct 134580  
gggttggtct cggtcagca gtcggaagca cgcacggcgg tccacgacct cgggtgctctc 134640  
caaggggctg tctccgcaa acagggccgt ggtggtgttt ggggggcagc gacaggacct 134700  
agtgcgcacg atcggggcgg tgggtttggg taagtccatc agcggctcgg ccaaccgctg 134760  
aaggttggcc ggacgaacga cgaccggggt acccaggggt tctgatgcca aaatgcggca 134820

-continued

---

ctgcctaagc aggaagctcc acagggccgg gcttgctgctg acggaagtcc ggggcagggc 134880  
gttgttctgg tcaaggaggg tcattactgt gacgacaaca acgcccattg tggtatatta 134940  
caggcccgtg tccgatttgg ggcacttgca gatttgaag gccacgcacg gcggggagac 135000  
aggccgacgc gggggctgct ctaaaaattt aagggcccta cggtocacag accgccttc 135060  
ccgggggggc ccttgagcgc accggcagcg gaggcgtccg ggggagggga gggtgattta 135120  
cgggggggta ggtcaggggg tgggtcgtca aactgccgct ccttaaaacc ccggggcccg 135180  
tcgttcgggg tgctcgttgg ttggcaactca cggtgccggc aatggcctgt cgtaatgttt 135240  
gtcgcgttta cgggggacag ggcaggagga aggaggaggc cgtcccgcg gagacaaaagc 135300  
cgtcccgggt gtttctcat ggcccccttt ataccaccgc cgaggacgcg tcctggact 135360  
ccccgcccc ggagaccccc aaacottccc acaccacacc acccagcag gcccagcgc 135420  
tgtgtcatct gcaggagatc cttgccaga tgtacgaaa ccaggactac cccatagagg 135480  
acgaccccag cgcggatgcc gcggacgatg tcgacgagga cgcgccggac gacgtggcct 135540  
atccggagga atacgcagag gagcttttcc tgcccgggga cgcgaccggc ccccttatcg 135600  
gggccaacga ccacatccct ccccgtgtg gcgcactccc ccccgtata cgacgacgca 135660  
gccgggatga gattggggcc acgggattta ccgcggaaga gctggacgcc atggacaggg 135720  
aggcggctcg agccatcagc cgcggcgcca agccccctc gaccatggcc aagctgggta 135780  
ctggcatggg ctttacgatc cacggagcgc tcaccaccag atcggagggg tgtgtctttg 135840  
acagcagcca tccagattc cccaacggg taatcgtgaa ggcggggtgg tacacgagca 135900  
cgagccaaga ggcgcgactg ctgaggcgac tggaccacc gccgatcctg cccctcctgg 135960  
acctgcatgt cgtctccggg gtcacgtgtc tggctctccc caagtaccag gccgacctgt 136020  
atactatct gagtaggcgc ctgaaccac tgggacgccc gcagatcga gcggtctccc 136080  
ggcagctcct aagcgcgctt gactacatc accgccaggc cattatccac cgcgacatta 136140  
agaccgaaaa tatttttatt aacaccccc aggacattg cctgggggac tttggcgcg 136200  
cgtgcttctg gcagggttcc cgatcaagcc ccttccccta cggaatcgcc ggaaccatcg 136260  
acaccaacgc ccccagggtc ctggccgggg atccgtatac cacgaccgtc gacatttga 136320  
gocccggtct ggtgatcttc gagaactgccc tccacaacgc gtccttgctc tcggccccc 136380  
gcgccccca aaggggcccc tgcgacagtc agatcaccg catcatccga caggcccagg 136440  
tccacgttga cgagtttcc ccgcatccag aatcgcgct cacctcgcgc tacgctccc 136500  
gocgggcccga gaacaatcgc ccgccgtaca cccgaccggc ctggaccgc tactacaaga 136560  
tggacataga cgtcgaatat ctggtttgca aagccctcac cttcgacggc gcgcttcgcc 136620  
ccagcgcgc agagctgctt tgtttgccg tgtttcaaca gaaatgaccg cccctgggg 136680  
gcggtgctgt ttgcggttg gcacaaaaag accccgatcc gcgtctgtgg tgttttggc 136740  
atcatgtgc agggcgccat gcgtgcccgt gttcccatta tcccattcct tttggttctt 136800  
gtcgggtgat cgggggttcc caccaacgtc tctccacca cccaaccca actccagacc 136860  
accggtcgtc cctcgcatga agcccccaac atgaccaga ccggcaccac cgactctccc 136920  
accgcatca gccttaccac gcccgaccac acacccccca tgccaagtat tggactggag 136980  
gaggaggaag aggaggaggg ggcgggggac ggcgaacatc ttgagggggg agatgggacc 137040  
cgtgacaccc taccaccgtc cccgggcccc gccttcccgt tggctgagga cgtcgagaag 137100

---

-continued

---

gacaaaccca accgtcccgt agtcccaccc cccgatccca acaactcccc cgcgcgcccc 137160  
gagaaccagtc gccccaagac accccccacc attatcgggc cgctggcaac tgcceccacg 137220  
acccgactca cctcaaaggc acgacccttg gttccgacgc ctcaacatac cccgctgttc 137280  
tcgttcctca ctgcctcccc cgcctcggac accctcttcg tcgtcagcac cgtcatccac 137340  
accttatcgt ttttgtgat tgggtgcgat gcgacacacc tgtgtggcgg ttggtccaga 137400  
cgcggggcgc gcacacaccc tagcgtgcgt tacgtgtgcc tgccgtccga acgcgggtag 137460  
ggtatggggc gggggatggg gagagcccac atgcggaaag caagaacaat aaaggcgggtg 137520  
gtatctagtt gatatgcac tctgggtgtt tttgggggtg ggcggacgcg gggcggtcac 137580  
tggacggggc gcagttaaat acatgcccgg gacccatgaa gcatagcgca cttccggggc 137640  
tcagaaccca cccgaaacgg ccaacggacg tctgagccag gcctggctat cccgagaaac 137700  
agcacacgac ttggcgttct gtgtgtcgcg atgtctctgc gcgcagctcg gcactcgggg 137760  
cttttgggaa gcctcgtggg ggctgttctt gccgccaccc atcggggacc tgcggccaac 137820  
acaacggacc ccttaacgca gcceccagtg tcccctcacc ccagccccct ggggggcttt 137880  
gccgtcccc tcgtagtcgg tgggctgtgc gccgtagtc tgggggcggc atgtctgctt 137940  
gagctcctgc gtcgtacgtg ccgcgggtgg gggcgttacc atccctacat ggaccagtt 138000  
gtcgtataat tcccccccc ccccccttc tccgcgtggg tgatgtcggg tccaaactcc 138060  
cgacaccacc agctggcatg gtataaatca ccggtgcgcc ccccaaaaca tgtccggcag 138120  
ggggatgggg gggcaatgcg gagggcacc aacaacacc ggctaaccag gaaatccgtg 138180  
gccccgccc ccaataaaga tcgcggtagc ccggccgtgt gacactatcg tccataccga 138240  
ccacaccgac gaatccccca agggggaggg gccatcttac gaggaggagg ggtataacaa 138300  
agtctgtctt taaaaagcag gggttaggga gttgttcggt cataagcttc agcgcgaacg 138360  
accaactacc ccgatcatca gttatcctta aggtctcttt tgtgtggtgc gttccgggat 138420  
ggggggggct gccgccaggt tgggggccgt gattttgttt gtctcatag tgggctcca 138480  
tggggtcgcg agcaaatatg ccttggtgga tgctctctc aagatggccg accccaatcg 138540  
ctttcgcggc aaagaccttc cggctcctga ccagctgacc gaccctcggg gggtcggcg 138600  
cgtgtaccac atccaggcgg gcctaccgga cccgttcag cccccagcc tccgatcac 138660  
ggtttactac gccgtgttg agcgcgcctg ccgcagcgtg ctccctaaacg caccgtcggga 138720  
ggccccccag attgtcgcg gggcctccga agacgtccgg aaacaacct acaacctgac 138780  
catcgttg tttcggatgg gaggcaactg tgctatcccc atcacggtca tggagtacac 138840  
cgaatgctcc tacaacaagt ctctgggggc ctgtccatc cgaacgcagc cccgctggaa 138900  
ctactatgac agcttcagcg ccgtcagcga ggataacctg gggttcctga tgcaccccc 138960  
cgcgtttgag accgcgggca cgtacctgcg gctcgtgaag ataaacgact ggacggagat 139020  
tacacagttt atcctggagc accgagccaa gggctcctgt aagtagccc tcccgtgcg 139080  
catccccccg tcagctcgc tctccccca ggctaccag cagggggtga cgggtgacag 139140  
catcgggatg ctgccccgt ctatccccga gaaccagcgc accgtcggc tatacagctt 139200  
gaagatcgc ggggtggcag ggcceaaagg cccatacag agcaccctgc tgcceccgga 139260  
gctgtccgag accccaacg ccacgcagcc agaactcgc ccggaagacc ccgaggattc 139320  
ggccctcttg gaggaccccc tggggacggg ggcgcgcaa atcccacaa actgacacat 139380

---

-continued

---

accgtcgatc caggacgccc cgagcgcctta ccatcccccg gccaccccga acaacatggg 139440  
cctgatcgcc ggccgcggtg gcggcagctc cctggcagcc ctggtcattt gcggaattgt 139500  
gtactggatg cgccgcaca ctcaaaaagc cccaaagcgc ataegcctcc cccacatccg 139560  
ggaagacgac cagccgtctc cgcaccagcc cttgttttac tagatacccc ccttaatgg 139620  
gtgcgggggg gtcaggtctg cggggttggg atgggacctt aactccatat aaagcgagtc 139680  
tggaaggggg gaaaggtgga cagtcgataa gtcggtagcg ggggacgccc acctgttccg 139740  
cctgtcgcac ccacagcttt ttttgcaac cgtcccgttc cgggatgccg tgccgcccgt 139800  
tgcagggcct ggtgctcgtg ggcctctggg tctgtgccac cagcctggtt gtccgtggcc 139860  
ccacggtcag tctggtatca aactcatttg tggacgccgg gcccttgggg cccgacggcg 139920  
tagtgaggga agacctgctt attctcgggg agcttcgctt tgtgggggac caggctcccc 139980  
acaccaccta ctacgatggg ggcgtagagc tgtggcaacta ccccatggga cacaaatgcc 140040  
cacgggtcgt gcatgctgtc acggtgaccc cgtgccaccg tcgccccgcc gtggcattcg 140100  
ccctgtgtcg cgcgaccgac agcactcaca gccccgcata tcccacctg gagctcaatc 140160  
tggcccaaca gccgcttttg cgggtccaga gggcaacgcg ggactatgcc ggggtgtacg 140220  
tgttacgcgt atgggtcggg gacgcgcaaa acgccagcct gtttgcctg gggatggcca 140280  
tagccgccga agggactctg gcgtacaacg gctcggccta tggtcctgc gaccggaaac 140340  
tgcttccgtc ttccgccccg cgtctggccc cggcgagcgt ataccacccc gccctaacc 140400  
aggcctccac cccctcgacc accaactoca cccctcgac caccatcccc gctcctcga 140460  
ccaccatccc cgtcccccaa gcatcgacca cgccttccc cacgggagat ccaaaaccac 140520  
aacctcccgg ggtcaaccac gaacccccat ctaatgccac gcgagcgacc cgcgactcgc 140580  
gatacgctct aacggtgacc cagataatcc agatagccat ccccgctcc atcatagccc 140640  
tgggtgttct ggggagctgt atttgcttta tacacagatg tcaacgccc taccgacgct 140700  
cccgtcggcc gatttacagc ccccagatgc ccacgggcat ctcatgcgcg gtgaacgaag 140760  
cggccatggc ccgcctcgga gccgagctca aatcgcatcc gagcaccccc cccaaatccc 140820  
ggcgccggtc gtcacgcaac ccaatgccc cctgacggc catcgccgaa gagtcggagc 140880  
ccgctggggc ggtcgggctt ccgacgcccc ccgtggaccc cacgacaccc accccaacgc 140940  
ctcccctggt ggtataggtc cacggccact ggccgggagc accacataac cgaccgcagt 141000  
ccctgagttg ggaataaacc ggtattatct acctatatcc gtgtatgtcg atttctttcc 141060  
ccccctcccc ggaacccaaa gaaggaagca aagaatggat gggaggagtt caggaagccc 141120  
gggagagggc ccgcccgcga ttttaaggcgt tgttgtgtg actttgcctc ttctggcggg 141180  
ttggtcgggt gctgtttgtt gggctcccat tttaccgaa gatcggctgc tatccccggg 141240  
acatggatcg cggggcggtg gtggggtttc ttctcggtgt ttgtgtgta tctgcttgg 141300  
cgggaacgcc caaacgtcc tggagacggg tgagtgtcgg cgaggacgtt tcgttcttc 141360  
cagctcccgg gctacgggg cgcggcccga cccagaaact actatgggccc gtggaacccc 141420  
tggatgggtg cggcccctta caccogctg ggtctctgct gatccccccc aagcaggtgc 141480  
ccgagacggt cgtggatgcg gcgtgcatgc gcgctccggt cccgctggcg atggcgtaac 141540  
cccccccggc cccatctgcg accggggggtc tacgaacgga cttcgtgtgg caggagcggc 141600  
cggcctggt taaccggagt ctggttattc acggggctcc agagacggac agcggcctgt 141660

-continued

---

ataccctgtc cgtggcgac ataaaggacc cggctcgcca agtggcctcg gtggtcctgg 141720  
tggtgcaacc ggccccagtt ccgaccccc ccccgacccc agccgattac gacgaggatg 141780  
acaatgacga gggcgaggac gaaagtctcg ccggcactcc cgccagcggg accccccggc 141840  
tccccctcc ccccgcccc ccgaggcttt ggcccagcgc ccccgaagtc tcacatgtgc 141900  
gtggggtgac cgtgctatg gagactccgg aagctatcct gttttcccc ggggagacgt 141960  
tcagcacgaa cgtctccatc catgccatcg cccacgacga ccagacctac tccatggacg 142020  
tcgtctggtt gaggttcgac gtgcccact cgtgtgcccga gatgccaata tacgaatcgt 142080  
gtctgtatca cccgcagctc ccagaatgct tgtccccggc cgacgcgcgc tgcgccgcga 142140  
gtactgtgac gtctcgcctg gccgtccgca gctacgcggg gtgttccaga acaaaccccc 142200  
caccgcctg ttcggccgag gctcacatgg agcccgtccc ggggctggcg tggcaggcgg 142260  
cctcctcaa tctggagttc cgggacgct ccccacaaca ctccggcctg tatctgtgtg 142320  
tgggtgactc caacgacct attcacgct ggggccacat taccatcagc accgcggcgc 142380  
agtaccgaa cgcggtggtg gaacagcccc tcccacagcg cggcgcggat ttggccgagc 142440  
ccaccaccc gcacgtcggg gccctcccc acgcgcccc aaccacggc gccctcgggt 142500  
tagggcggt gatgggggccc gccctgctgc tgtctgact ggggttctgc gtgtgggct 142560  
gtatgacctg ttggcgcagg cgtgctggc gggcggttaa aagcagggcc tcgggtaagg 142620  
ggccccgta cattcgcctg gccgacagcg agctgtacgc ggactggagc tcggacagcg 142680  
agggagaac cgaccaggtc ccgtggctgg ccccccgga gagaccgac tctcctcca 142740  
ccaatgatc cggctttgag atcttatac caacgctcc gtctgtatac ccccgtagcg 142800  
atgggcatca atctcgcgc cagctcaca cctttggatc cgggaaggccc gatcgcctt 142860  
actcccaggc ctccgattcg tccgtctct ggtaaggcgc cccatcccga ggccccact 142920  
cggctgcccga actggcgag cgccggcgag gtggacgtcg gagacgagc aatcgcgatt 142980  
tccgacgaac gcggaccccc ccgacatgac cgcgccccc tcgccacgct gaccgcgccc 143040  
tcgccacacc cgcgaccccc gggctacacg gccgttctct ccccgatggc cctccaggct 143100  
gtcgacgccc cctcctgtt tgtcgcctgg ctggcgcctc ggtggctccg ggggcttcc 143160  
ggcctggggg ccgtcctgtg tgggattgct tggatgtga cgtcaattgc ccgaggcga 143220  
taaaggccg gtggctccc tagccgagc aaattaaaa tcgtgagtca cagcgaccgc 143280  
aacttccac ccggagcttt cttccggcct cgatgacgtc ccggctctcc gatcccaact 143340  
cctcagcgcg atccgacatg tccgtgccc tttatccac ggctcgcga gtttcggctc 143400  
aagcctaact ctccgaaagc gaagacgag cgcccaacga cttcctcgtc cgcattgggc 143460  
gccaacagtc ggtattaagg cgtcgacgca gacgcacccg ctgcgctggc atggtgatcg 143520  
cctgtctcct cgtggcctt ctgtcggcg gatttggggc gctcctgatg tggctgctcc 143580  
gctaaaagac cgcacgaca cgcgctcct tcttctcgtc tctcttccc cccatcacc 143640  
cgcaattgc acccagcctt taactacatt aaattgggtt cgattggcaa tgttctctcc 143700  
cgggtgattt ttgggtgggt ggggagtggt tgggtgggga gtgggtgggt ggggagtggt 143760  
tgggtgggga gtgggtgggt ggggagtggt tgggtgggga gtgggtgggt ggggagtggt 143820  
tgggtgggga gtgggtgggt ggggagtggt tgggtgggga gtggcaagga agaaacaagc 143880  
ccgaccacca gacagaaat gtaaccatc ccaaaccgac tctgggggct gtttggggg 143940

---

-continued

---

tcggaacat aggatgaaca aaccaccccg taccacccgc acccaagggt gcggtggctc 144000  
atcggcatct gtccggtatg ggttgttccc caccactcgc cgttcggacg tcttagaatc 144060  
atggcggttt tctatgccga catcggtttt ctccccgca ataagacacg atcgcataaa 144120  
atctgtttgt aaaatttatt aagggtacaa attgccttag cacaggggtg gggttagggc 144180  
cgggtcccca caccaaaacg caccaaaacg atgcaggcag tgggtcgagt acagccccgc 144240  
gtacgaacac gtcatgctgt gtgtcagaca gcaccagaaa gcacaggcca tcaacaggtc 144300  
gtgatgtgt cggtagggtt ggacgcgggg ggccatggtg gtgataaagt taatggccgc 144360  
cgtccgccag ggccacaggg gcgacgtctc ttggttggcc cggagccact ggggtgtggac 144420  
cagccgcgcg tggcgcccca acatggcccc ttagccgggg ggcgggggat cgcgcacgtt 144480  
tgcagcgcac atcgcagaca cctcagacc ggttcgaaag aaggcccggt ggtccgcggg 144540  
caacatcacc aggtgcgcaa gcgcccgggc gtccagaggg tagagccctg agtcatccga 144600  
ggttggctca tcgccccggg cttgccgcaa gtgcgtgtgg gttgggttc cggtagggcg 144660  
gacgcgaacc gcggtgtgga tcccagcgcg ggcccagcgc tatgctccat gttgtgggga 144720  
gaaggggtct gggctcgcga ggggggcata cttgccgggg ctatacagac ccgcgagccg 144780  
tacgtggttc gcgggggggt cgtgggttcc ggggctccc gggagaccgg ggtcccggg 144840  
gagaccgggg ctcccgggga gaccggggtt gtcgtggatc cctgggttca cgcggtacc 144900  
tggggtctct gggagctcgc ggtactctgg gttccctagg ttctcgggtt ggtcgcggaa 144960  
ccccgggctc ccggggaaca cgcggtgtcc tgggattgt tggcggctcg acggttcag 145020  
atggtctoga gatcgtagt tccgcaccga ctcgtagtag acccgaatct ccacattgcc 145080  
ccgcgcttg atcattatca ccccgttgcg ggggtccgga gatcatcgc ggtgtctctc 145140  
gaggtgctgt aacacctctg ggtgcatgc cggcggacgc cacgctttt aagtaaacat 145200  
ctgggtcgc ccgcccact gggcccggg gttgggtctg gctcatctc agagccacgg 145260  
gggggaacca cctccgccc agagactcgg gtgatggtc taccggggac tcaacgggtt 145320  
accgattac ggggactgtc ggtcacggtc ccgcccgttc ttcgatgtc cacacccaag 145380  
gatgcgttg gggcgatttc gggcagcgc ccgggagagc gcagcagggg acgctccggg 145440  
tcgtgcaagg cggttctggc cgcctccgg tccctacgcc ccttttatt gatctcatc 145500  
cgtacgtcgg cgtacgtctc gggcccaacc cgcattggtt ccaggaagggt gtccgccatt 145560  
tccagggccc acgacatgct ccccccgac gacaggaag cggtcacgc aacggtcgc 145620  
gccggtcgc tcgacgagga cgttctctc gcgggaaggc acgaacgcgg gtgagcccc 145680  
tctcgcgcc ccggtcccc cctcctcgc ccccggttc cccctctcc gccccgggt 145740  
ccccctctc ccgccccgc gtccccctc ctccgcccc gcgtcccccc tctcgcgcc 145800  
ccggtcccc cctcctccac ccccggttc cccctctc cgcacacca aggtgcttac 145860  
ccgtgcaaaa aaggcggacc ggtgggttct tgcgtcggg ggcggggg gtgcgtcccc 145920  
tgtgttctgt ggtgggggtg ggcgggtctt tccccccgc gtccgctgt ccttttcga 145980  
tgcgatccc atcccagcc gggcgctgc gatgccgac cgtccgctc cagcggccct 146040  
ctgcactcc cgtcccgggt ccgctgtctc cgcagccgt cccgtcttc gtggccggcg 146100  
ccgtctcgg gcgtcggctc gcgcccgcct ttatgtgcgc cggagagacc cccccccgc 146160  
cgccccggcc ccccccggg gccggcgcg agtcgggca ggcgccagt ctcgacttc 146220

---

-continued

---

gcccataataa tatatatata ttgggacgaa gtgcgaacgc ttcggttct cacttctttt 146280  
accggcgccgccc cccgccccct tggggcggtc ccgccccgccc gccaatgggg gggcggaag 146340  
gcgggcgccgccc cttggggccc ccgcccgtccc gttggtccc gctccggcg ggcgggaccg 146400  
ggggggcccgg ggacggccaa cgggcccggc gggctcgat ctattaccg ccgaaccggg 146460  
aagtcggggc ccggggcccc ccccctgccc gttcctcgtt agcatgcgga acggaagcgg 146520  
aaaccgcccg atcggcggtt aatgagatgc catgcccggc ggggcccgga cccaccgcc 146580  
ctcggcccc gcccatggca gatggcgccc atgggggggg ccgggggttc gaccaaccgg 146640  
ccgcccgcac gggccccggc cgtcccggc tggggcggg gtcgtgcata atggaattcc 146700  
gttcgggggtg gggcccggc gggggcgggg ggcggcgcc ctcgctgct cctcctccc 146760  
gcccggcccct gggactatat gagcccagg acgcccgat cgtccacacg gagcggcgt 146820  
gcccagaccg atccagacc cagcgggga ccccagaga cagaccgta gagcctcgc 146880  
gcccgggac gccgatacc ggacgaagc cgggagggg atcggccgc cctgtcctt 146940  
ttcccacca agcatcgacc ggtcccgcct agttcccgc cagcggcggg ggtcgtcgg 147000  
gtcctgggt ctcgcccc cccccatc agagtcgta ggtgacctac cgtgctacgt 147060  
ccgcccgcac agccgtatcc ccggaggatc gcccgcctc ggcgatggc tgggagaaca 147120  
agcagccc cggctccc gggcccacc acggggccc gcccccggc agcccagacc 147180  
gcccagacc gggggccc cgggagcgg ggagggggg gacgacccc 147240  
accagacc cagaccacc cagacctc acgacccc cggggacgg agggcccgg 147300  
cgggggccc cagcggccc gaggacccc gggacccc ctcgcccga cagctggct 147360  
tgctggctc catggttag gaggcccgc ggacgatcc gaccccggc cccgcccct 147420  
cggcccggc gaccccggc ttctgagccc acgacgatga cggggacgag tacgacgac 147480  
cagcccgcac cgcggcgcc cgggcccgg cccggggccc cgaacgggag gcccggctac 147540  
gcccggcgt tccggacccc accgaccgc tgtcccggc cccgcccggc cagcccggc 147600  
ggagacgtc taccgccc gggggccc cggcgtcct gacctcgtc gactcccgg 147660  
cctcgtcctc gtcgtccc tccctctcgt cctcgtcgc cagcgggac gaggacgac 147720  
accgcaaca cgcggccc cagcaccgc aggcggggc cgtcccggc ggtccgtcga 147780  
gcccggccc ggcagcccc gggcgggc cccccccc cgggcccacc cccctctcc 147840  
aggcccggc caagcccgg cggcgggc gaaccccgc ggcctccc ggcggcatc 147900  
agcccggc ggcggcgcc cgggtggcc gcccggacc cagggccc ttcacggcc 147960  
ggcagccc ggggtcgag ctggagccc acggacctc cggcgcttc tacggcgct 148020  
atcggcagg gtacgtcag ggggagcgt gcccggccc cgggcccgg ccccggggc 148080  
gggtgctgta cggcgccc ggcgacacc gcccgggct ctggggggc cccgagggc 148140  
aggagggcg acgcccgtt gaggcctc gcccgggc ggcctgtg gcccggagc 148200  
tggggcagc cgcgagcag taccccctga tcacgggct gctgtacacc cgggacggc 148260  
aggccatgg gtggctccag aaccggccc tggctccc ggacgtggc ctggaccag 148320  
cctgctccc gatctcggc gccgcccga acagcagct cttcaccacc ggcagcgtg 148380  
cggggggcgt gcccaccct ggctaccga tggggccc cggcttcggc tggggcctg 148440  
cgcagccc ggcggccc gcatgagcc gccgataca cggcgccag aagggtctc 148500

-continued

---

tgctgaccag cctgcgcccgc gcctacgcgc ccctgttggc gcgcgagaac gcggcgctga 148560  
cgggggcccgc ggggagcccc ggcgcggcg cagatgacga gggggtcgcc gccgtcgccg 148620  
ccgcgcacc gggcgagcgc gcggtgcccg ccgggtacgg cgcgcggggg atcctcgccg 148680  
ccctggggcg gctgtccgct gcgcccgcct cccccgggg gggcgacgac ccgcagcccg 148740  
cccccaacgc cgacgccgac gacgacgccg ggcgcgcgc ccaggccggc cgcgtggccg 148800  
tcgagtgcct ggcgcctgc gcggggatec tggaggcgct ggccgagggc ttcgacggcg 148860  
acctggcggc cgtcccgggg ctggccgggg cccggcccgc cagcccccg cggccggagg 148920  
gaccgcggg ccccgttcc ccgcgcgcgc cgcacgccga cgcgcccgct ctgcgcgct 148980  
ggctgcgcga gctgcggttc gtgcgcgacg cgtggtgct catgcgctg cgcggggacc 149040  
tgccgctggc cggcggcagc gaggccgcgc tggccgcct gcgcgcctg agcctggctg 149100  
cgggggccct gggcccgcgc ctgcgcgggg acccgcgct gccgagctcc gcggccgccc 149160  
ccgcgcgga cctgctgttt gacaaccaga gcctgcgccc cctgctggcg gcggcggcca 149220  
gcgcaccgga cgcgcgcgac gcgctggcgg ccgcgcgcgc ctccgcgcgc cgcggggagg 149280  
ggcgcaagc caagagtccc ggcggggccc ggcgcggcg aggcggcggc ccgcgacccc 149340  
cgaagacgaa gaagagcggc gcgacgccc ccggctcgga cgcgcccgcc cccctcccg 149400  
cgcgcgcgc cccctccacg cccccgggg ccgagcccgc cccgccccag cccgcggcgc 149460  
ccccggcgc cgcggcgcag gccgcgcgc gccccgtggc cgtgtcgcgc cggcccgcgc 149520  
agggcccga cccctgggg ggctggcggc ggcagcccc ggggcccagc cacacggcgg 149580  
cgcgcgcgc cgcgcgcctg gaggcctact gctcccgcgc cgcgctggcc gagctcacgg 149640  
accaccgct gttcccgcgc ccctggcgac cggcccctcat gttgacccg cgggcccctgg 149700  
cctcgatgc cgcgcggtgc gccggggccc cccccgcgc ccaggccgcg tgcggcggcg 149760  
gcgacgacga cgataacccc cacccccacg gggccgcggc gggccgcctc tttggcccc 149820  
tgccgcctc gggcccgcgc cgcgcctgg cggcctggat gcgccagatc cccgaccccg 149880  
aggacgtgc cgtggtggtg ctgtactgc cgtgccggg cagggacctg gccggcggcg 149940  
gggcctcggg ggggcgcgc gcgaggtcgc ccgagcgcgc cgggctgtcc tgcctgctgg 150000  
cggccctggc caaccggctg tgcgggcggc acacggcgc ctgggcgggc aattggaccg 150060  
gcgccccga cgtgtcggcg ctgggcgcac agggcgtgct gctgctgtcc acgcgggacc 150120  
tggcctcgc cggggcgcgc gagtttctgg ggctgctgc cagcgcggc gaccggcggc 150180  
tcacgtggt caacaccgtg cgcgcctgc actggcccgc cgaaggccc gcggtgtgc 150240  
ggcagcacgc ctacctggcg tgcgagctgc tgcccgcgt gcagtgccc gtgcgctggc 150300  
cggcggcgc ggcacctgc cgcacggctg tggcctcggg ccgcgtgttc ggcccggggg 150360  
tcttcgcgc cgtggaggcc gcgcacgcgc gcctgtaccc cgaagcgcgc ccgctgcgc 150420  
tgtgccgcgc cggcaacgtg cgtaccgcgc tgcgcacgc ctccggccc gacacgccg 150480  
tgcccatgct cccgcgcgag taccgcggg ccgtgctgccc ggcgctggac ggcggggcg 150540  
cggcctcggg gaccaccgac gccatggcgc ccggcgcgc ggacttctgc gagggaggag 150600  
cccactcga cgcgcctgc gcgcgctggg gcctgggcgc gccgctgcgc cccgtgtacg 150660  
tggcgtggg gcgcgaggcg gtgcgcgcgc gcccgcccgc gtggcgcggg ccgcggaggg 150720  
acttttgcgc ccgcgcctc ctggagccc acgacgacgc cccccgcgc gtgctgcgc 150780

-continued

```

gcgacgacga cggccccggg gccctgccc cgccgcccgc cgggattcgc tgggcctcgg 150840
ccacgggccc cagcggcacc gtgctggcgg cggcgggggc cgtggagggtg ctggggggcgg 150900
aggcgggctt ggccacgccc ccgcccgggg aagttgtgga ctgggaaggc gcctgggacg 150960
aagacgacgg cggcgcgttc gagggggacg ggggtgctgta acgggcccgg acggggcggg 151020
gcgcttgtag gacccgaaga cgcaataaac ggcaacaacc tgattaagtt ttgcagtagc 151080
gttgtttatt cgagggggcg gagggggcga gggggcggag gggcgagggg gggggagggg 151140
gcgagggggc ggagggggcg agggggcgga gggggcggag ggggggaggg ggcgaggggc 151200
gggagggggc gagggggcgg agggggcgag gggcgggagg gggcgagggg cgggaggggg 151260
cgagggggcg gagggggcga gggggcggag gggcgagggg gcgggagggg gcgagggggc 151320
ggagggggcg agggggcgga gggggcggag ggcgggaggg ggcgaggggc gggagggggc 151380
gagggggcgg agggggcgag gggcgggagg gggcgagggg cgggtggtggt gcgcccggc 151440
ccccggaggg tttggatctc tgacctgaga ttggcggcac tgaggtagag atgcccgaac 151500
cccccgagg gagcggggga cgcgcccggg agggctgggg ccggggaggg ctggggcggg 151560
ggagggctgg ggcgggggag ggtcggggcc ggggagggct ggggcccggg agggctgggg 151620
ccggggaggg ctggggctgg ggagggctgg ggtcggggag gggcggtgg tgtgtagcag 151680
gagcgggtg ttgcgcccgg gtacctctgg aggagcggga ggtgcgcggt gacgtgtgga 151740
tgaggaacag gagttgttgc cgggtgagtt gtcgctgtag gttgtgtgtg tgggcaggtg 151800
tggtggatga cgtgacgtgt gacgtgaggg gtgcgcccgt ctctgttggg ttaacctgtg 151860
gcagcccggg ccccccgcg ggcgcccgc gcgcaaaaaa ggcggggcgg ggtccggggc 151920
gctgcccgc gcgcccggg cgtggggggc gggggcggc gagcgggggg aggagcgggg 151980
ggagggcgg ggggaggagc ggggggagga gcggggggag gagcgggggg aggagcgggg 152040
ggagggcgg ggggaggagc ggggggagga gcggggggag gagcgggggg aggagcgggg 152100
ggagggcgg ggggaggagc ggggggagga gcggggggag gagcgggggg aggagcgggg 152160
ggagggcgg ggggaggagc ggcagacgc cgaaacggg cccccccaa aacaccccc 152220
cggggggtcg cgcgcccgc tttaaagcgg tggcggcggg c 152261

```

&lt;210&gt; SEQ ID NO 16

&lt;211&gt; LENGTH: 154746

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Herpes simplex HSV-2 (NC\_001798.1)

&lt;400&gt; SEQUENCE: 16

```

agtccccgtc ctgcccggc ggggcccggc cgggaaaaaa gccgcccggg ggcgcccggc 60
ggaaggcagc cccgcccggc gcggggggag gggcggcggc cgcgggggag cggcccggctc 120
cgggggaggg acggggaagg gggcggcggg ggtgcccgtg cgccccggcc gccgcccggc 180
cccgccttcg ccccccccc caaaaaaac ccccccggg ggttgactcc ccgggggaaa 240
agagggcggg cgggagtcgc gtcctgccc cgcgccctta agagggcccg caacacggcc 300
cgggtcggc acgcccggc ggcggggtag gttcgtagg caagcacgga ctggcgggta 360
cacgtgcatg cgtgcccagt gaaactctcc gcccccggc gctccggctc cgggcctacg 420
ccgagcccag cgccccggc tgtcccggc ccgggggtccc cgcggcccgg gtcgccggcg 480
cggcccggc cccgcccggc cagccgtgcc gcgcccggc gctccagccc tgcccggccc 540

```

-continued

---

cgggcgctc ccaaccgcag actcccaaat ggtccctgcg tacgactcgg gaaccgcggt	600
cgagagcgcg ccggcccggt cctcgctcct gcggcgctgg ctgctggtgc cccagggcga	660
cgacagcgac gacgcggact acgcgcgcaa cgacgacgca gagtgggcca acagccccc	720
gagcgagggc ggggggaaagg cgccggaggc ccgcgacgcc gcgcctgccc ccgcctgccc	780
cccgcgcgcg ccgcgcaagg agcgcggggc gcagcgcgcc cttccgcccc acctggcgct	840
acggctgcgc accacgacgg agtacctggc gcgcctgagc ctgcgcgggc ggcgggcccc	900
cgcgtccccc ccgcgggacg cgccgcgcgg gaaggtacgc ctccctccg acccctgac	960
gcccctccga ccccctgacg cccctccgac cccctgacgc ccctccgacc ccctgacgcc	1020
cctccgaccc cctgacgccc ctccgacccc ctgacgcccc tccgaccccc gtgtctcccc	1080
gcccgcaggt gtgcttctcg ccgcgcgtgc aggtgcgcca tctggtggcc tgggagacgg	1140
ccgcgcgctt ggcggcagcg gggctcctggc cgcgcgagcg ggccgacgcg gaccggttcc	1200
ggcgcccgct ggcggcgggc gaggcgggtca tcggaccgtg cctggagccc gaggcccag	1260
ctcgggcccc agccccgagc cgggcccacg aagacggcgg acccgcggag gaggaggagg	1320
cggcgccggc ggcgcgcggg tcctccgccc ccgcggggccc gggccgtcgg gcggtctagg	1380
gttgaaccgg cgagggcgcg ctcggccggc ggagccccgg agctccgaag gtctgcgca	1440
ggccgctctc cgaagagacg atggggagccc cgcgatata tccgcgaggg ccggcgcgcg	1500
ccccgcgct ccgcccgcgc cagggggcgg cgccggccaa ccgcgcgccc ccgcgcgggc	1560
ccggactccg ccccgcgacg cgccccgcgc cgcttcccgt gtatgtaaat tagaaacttt	1620
taataggcgg tcccggcccgc catccccgcg catggttaatt agcaactttt aatgggcccg	1680
cgttcccgcg cgcggttaatt agcagctttt aacgggcccgc cattcccgcg tatggttaatt	1740
aaaaacgttc ggacggcccc tcgctccccg cgtaattact cctcgggggt tccgggttat	1800
gctgattact ttcttggcag aacacgcaga gcctcgcgcg ccgcgggggt ggtgggctga	1860
tcggccccta ttggtcccct gggcttccta gtatgtaaat gaatttttcc ccggggcgcg	1920
gcaccaactca gggccgcgcc ggcggggcgc cggggggact cccatctgcg tcggcggggg	1980
gcggcgcatg ctaatgggggt tcttgagta caccgggttg gtccccgggg acggggcgcg	2040
cccgagaggg ggggattccc tcctccgccc ccgcggggg cgcgcgcta ttgggggaat	2100
cgtaaatgcc gcccctttgg gggagtgat aggcgcccgg tataaggcag ccccggtgta	2160
cggtcgggccc gcattcgcac cccggcactg cgagcgacgg agcggcgccc ccgcgggagg	2220
aggagaccgg gagagacaga gactaaaacc cgcaagaga gagaccgcgg gccgcccgtct	2280
cgagtctacc ctaccggcg tcattggaacc ccggcccggc acgagctccc gggcggaacc	2340
cggccccgag ccgcccgcgc ggcagacccc cggcacggtg agagggcgac ccccggtct	2400
caggcccccc cttttcccgc gaccaccggc ctgcgggttg ggggtggtcg ccggcggtgg	2460
gctcgggggc ggggacgctt gacggggccc acccccggcc cgcttaagcg gtcgggggac	2520
ccccgtgggc cgtgcgcgccc ccccgcaccc tctggggggg cgagggaggc agggaggagc	2580
ccgagagcgg gggacagggg gggagacgag gggtcggaat ccaaaggacg cagaccacct	2640
ttggttacgg acccctttct ccccccttc cgaacaaaaa gcagcgggcg gggggccggg	2700
gtgagggagg gacacggggg acacggcgcg ggggtcccgc ctcacgcccc gcgcccctta	2760
aatccccccc gttgctttgt caagcagccc gccgccccgc acgcctgggg gatgctcaac	2820

-continued

---

gacatgcagt ggctcgccag cagcgactcg gaggaggaga ccgaggtggg aatctctgac	2880
gacgaccttc accgcgactc cacctccgag gcgggcagca cggacacgga gatgttcgag	2940
gcgggctga tggacgcggc cacgccccg gcccgcccc cggccgagcg ccagggcagc	3000
cccacgcccg ccgacgcgca gggatcctgt gggggtgggc ccgtgggtga ggaggaagcg	3060
gaagcgggag gggggggcga cgtgtgtgcc gtgtgcacgg acgagatcgc cccgcccctg	3120
cgtgccaga gttttccctg cctgcacccc ttctgcatcc cgtgcatgaa gacctggatt	3180
ccgttgccga acacgtgtcc cctgtgcaac accccggtag cgtacctgat agtgggcgtg	3240
accgccagcg ggtcgttcag caccatcccc atagtgaacg acccccggac ccgctgggag	3300
gccgagggcg ccgtgcgggc cggcacggcc gtggacttta tctggacggg caaccgcggg	3360
acggccccgc gctccctgtc gctgggggga cacacggtcc gcgccctgtc gccaccccc	3420
ccgtggcccg gcacggacga cgaggacgat gacctggccg acggtgaggg cgggcggggg	3480
tcgggcgggg ggcgggcggg ggtcgggcgg gggtcgggcg ggggtcgggc gggggtcggg	3540
cgggggtcgg gcgggggtcg ggcgggggtc gggcgggggt cgggcggggg tcgggcgggg	3600
gtcgggcact aaccgggggc tcccgtctct gtctccctct gcagtgact acgtcccgc	3660
cgcccccgca agagcgcccc ggcgcggggg cgcggtgctg ggggcgaccc gcggaacctc	3720
ccagcccgcc gcgacccgag cggcgcccc tggcgcccc cggagcagca gcagcggcgg	3780
cgccccgtt cgggcggggg tgggatctgg gtctgggggc ggcctgccc tcgcgccgt	3840
cgtgccgaga gtggcctctc tccccctgc ggcgcgggg ggcgcgcgc aggcgcggcg	3900
ggtgggcgaa gaccccgcgg cggcggaggg caggacgccc cccgcgagac agccccgcgc	3960
ggccccaggag cccccatag tcatcagcga ctctccccg ccgtctccgc gcgcccccgc	4020
gggccccggg ccctctctct ttgtctctc ctctccgca caggtgtcct cgggccccgg	4080
ggggggaggt ctgccacagt cgtcggggcg cgcgcgcgc ccccgcgcg ccgtcgcccc	4140
gcgcgtccgg agtccgcccc gcgcgcgcgc cgcgccctg gtgtctgcga gcgaggacgc	4200
ggccgggccc gcgcccggc ccgtgcgggt ggacgcgcac cgcgcgccc ggtcgcgcat	4260
gacccaggct cagaccgaca cccaagcaca gagtctgggc cgggcaggcg cagaccgacgc	4320
gcgcgggtcg ggagggccgg gcgcggaggg aggaccggg gtcccccgcg gcaccaaac	4380
ccccggtgcc gcccccaag ccgcggaggg ggcggcgccc cgcccccgga agaggcggcg	4440
gtcggactcg ggcggcgcg cctcgtctc cgcctcttc tccgcccgc cgcgctcgc	4500
cctcgcccc cagggggtgg gggccaagag ggcggcgccc cgcgggccc cggactcggga	4560
ctcgggcgac cgcggccacg ggcgctcgc cccggcgtcc gcggggcgcg cgcgccctc	4620
ggcgtctccg tcgtcccagg ccgcggtcgc cgcgcctcc tctctctcc cctcctctc	4680
ctccgctcc tctctccg cctcctctc ctccgctcc tctctctcc cctcctctc	4740
ctccgctcc tctctccg cctctctc tcggggcggg gctggggga gcgtcgcgtc	4800
cgcgtccggc gctggggaga gacgagaaac ctccctcggc ccccgcgctg ctgcgcccg	4860
ggggccgagg aagtgtcca ggaagacgc ccacgcggag ggcggcccc agccccgggc	4920
ccgcgacccg gcgcccggc tcaecgcta cctgcccac gcgggggtct cgagcgtcgt	4980
ggccctggcg ccttacgtga acaagacggt cacgggggac tgcctgccc tctggacat	5040
ggagacgggc cacatagggg cctacgtggt cctcgtggac cagacgggga acgtggcgga	5100

-continued

---

cctgctgcgg gccgcggccc ccgcgtggag ccgccgcacc ctgctccccg agcacgcgcg	5160
caactgctg aggccccccg actacccgac gccccccgctcggagtgga acagcctctg	5220
gatgaccccc gtgggcaaca tgctctttga ccagggcacc ctggtgggctcgcctggactt	5280
ccacggcctc cgtctgcgcc acccgtggtc tgggagcag ggcgcgcccg cgcggcccg	5340
cgacgcccc gcgggcccacg gggagtaggg ggagctaaca ctcggttgc tccccgaagg	5400
aagccgcccc ccaccggacc accggccgag gcgcctcggg ggcaggggga ggtggggggg	5460
gggaaagacg gggagagagc aggaagtggg ggtgggagtg gggggggggg acggacacgg	5520
ccccgaacag caacacacac cagcattttg ttatggactt tctggccttg ttgaaaactt	5580
gaggaaaaaa aaaactttat atttataaaa attttacaat aaagttttgt gatgcttttg	5640
acacactttg ttgttgccct ttgatgcagc tccccgcgc aggggggccc gggatggggg	5700
ggaaggagg aggagagggg gggcgggca cgagaagccg cccccacccc cgaggcctgt	5760
tggtctttat catagaacag agccggggcc cggcctcgtt ctggctcctt gtcttggtg	5820
gtggcgggc tggctggcgg gtaaaaaag agtgtgtccg tgttgacagg gagggggcc	5880
cgatcgtgca gacacgcac gtctggcccg ccagaccctg ggggtggtg gcaggagtgg	5940
gagggcgcc ggctcgggga gggaggggg ggggggtcag ccgcaccacc ggcgcgaagc	6000
caggggccc ggaactttga tagagagggg ggaagtggg gcggggcgga gggcggttga	6060
atcacaacgc atgcacgccc tctgccccg gggacgggtg ggaggaagga ggaggagaa	6120
gagaagaacc gaggcctgca cccgcactta cggcctgccc cccccccc cggcgccca	6180
ccccgcccgc acacctgccc gccacgccc cccctcctca ccctggctgg gagaaaggag	6240
gaggagcagg aagaggagac ccgagggcatg caaccgcaact cccccacccc cgcgccaca	6300
cctgccccgc acgccccccc ctcttaacc tggctcggg gagactccca tggggcgag	6360
gggctcgcg cgttcgcaac accacaccac accacacggc ccaccacaac acggccccacc	6420
acgacacaac acgacacgac gcgttttgcg gggcatgcaa gtcgacacac cgcgcccgtg	6480
cctaccttc ctagcggcc ccggccccg gccctttcc ttcggccacc actaccacca	6540
cccccccgc cgcgcccaag cggtagagga aggggacgg cgccacacccc acggtgtg	6600
ccgggcaacg gcctttgggg ttgttgggg ggggtgaccg gcgcgtggg gcggtggcg	6660
tacgggccc acccgcgct cccccccc gaacgacgac ggggggggg gaaacgggg	6720
tgggtggaag ggaagaggaa ggagaaagg ggggtggatc cgaacacgcc ggatccgca	6780
aaataataac aaaacaaaca aaaacagaaa caaaaacaaa aacacctaga aaaaaaggat	6840
acgggttggc tcgcgggcgg tgcggctgac ctgcctgccc tttctgggac cccgcctcg	6900
tgtttcttga aagggggagg aagaacagtt ctcccccaac ccctgctctc ttctctctc	6960
cgcgccccc cccccctct ccccgccgc tcagcagaag ctcaoctgta cgacctaaa	7020
cctacctgag agaacgcgcg gcgttcgagg ggcgcgctct ctcaacgag acacacgag	7080
gcgccccccc ccccgggagc ctgggtcccc cggcggacgg ctcaacgggc gggcgctc	7140
ggtgggacgc gggcaagggg cggcggcggc ggggggggg gggggaaatg tgaggagagc	7200
gagacagaga gagagaagg agagggagg ggcgcggcgg gacgggggaa gacgaggaga	7260
agggaaagggg cgagggctcg gcccgggagc gggcgggccc gggagggaga agaaacggaa	7320
cgcgaaaacg ccgcccgcgc ggcgggggg cccggggccc ccgctcctc cggggggccc	7380

-continued

---

gggceggacc gccgggcggg ggaagccttc cgcccggcgc cgggceggeta cccgggaccc	7440
ccggccggga atcgaaaaaa gcctccgggg gcccctttcg cgcctttcgc gaacgcgcgg	7500
cgccggaggg ggcggccgcc gaggtgctgg ggcctctccg gccggggcgc acctcggcgg	7560
ccaagccccg gcccgccccg gggctcccga ggcaagaggc ggaccctcgg aggcgcggaa	7620
gaagacggga ggcgggggaa aaaaggggga agagaggggg aggtaggagagg gggagaggag	7680
aagggcgcgc cggtgccggg agcagccttc cttctccgga gtccctctcg atcggcggcg	7740
ggccctcgcg ttcgttgctg ccgcgcccc ggttttataa agacagggat gacgcagcag	7800
aaatgcccaac agcaaacacgc gggcggggct cgggctctcc ggcggcttaa tggatctccg	7860
ggcacggcgc ccgcaaccgc agagcactca gctggcgcgc cccccccaa cgtgggagtg	7920
tttaatggaa gggcgtgggg ccggccgcgc gatgcccgcg ggggcctaat gcggcgggag	7980
gcgtgggcgc ctggccgcgc ggcccgctcg ctggcccgcg gcccgctcgc tggcccggcg	8040
ccacgtaaac aatgacacag gggttctctc cgccgcggcc ggcgcggggc gttgcggcc	8100
cggcccggcc ccggagcccc cggcgcctgt cggtcgggc cgcgggctcc gggggctccg	8160
cactctgccc ggtcgcgcc gtcccccctc ttgetgcttt tccgcgcgc tetctttccc	8220
gttgctttcc ctctccccc cccccctct ctctctctct ctctctctct ccgccatcct	8280
cccgccggc cgcccactcc ccgctcggcc tctccggctg cggtgcttgg gtctccttcg	8340
tcggcggcgc gggggggggc gtcgggactc gcggagggcc ggagaatgga aggcgagggg	8400
atgcaggagg aggatcggga ctccccatct tctgcccctc catcctcctt ttttcgctt	8460
tccaaccgcg ccgccaccac cccccctcc ttgcgccgc cgcctcgcgc cggaccctc	8520
cccccgctgt tccccccatc gttcaaccac acgccccca ccgcgccttg gctggttggg	8580
gggtggcggc ggtggtcggc gtgctcggc aggetcggg cgcgggtag gtggtgggc	8640
gggtggtggg ggggggcccc gctgcgtctc gcccgatcc cgcgggtggg gcgcggcggc	8700
ggtcggggtg gggggagagt gtcgtgggtg tgtttctgtg tccccacca ccaactccac	8760
cccgaccgc gccgcgcccc cgtttctgcc gcccgcgcc tctgtgttg accccgggt	8820
ggcggcggg ggggggtgcc gtaggtgtg cggcggggc cggccgggg ccggggctcg	8880
ctggtccgc gaagtaaaga aaagatcgc accgtgtgtt cgtctgtgtg ttctgcgcg	8940
cgcggggcc cccctcggc gcggggcggt ggggggggt cggggtcgc gcggggaagg	9000
aaggaaaagc ccggaagcg ccgggaggg gcgccggcgc gacgcgggc gccggcggg	9060
ggcgcgcgc gcccgggcg ggcgcgcgc cggccgggc gggcgcgc gcggccgggc	9120
gggggcgcgc gccggccggg cgggggcgc cggcgcccg gcgggggcgc gcggcggcg	9180
ggcggggcg cgttttccc cgcgcgccc tcgggttccc aagacctatc acgtgtgcgc	9240
aggggagggg aggacgcgg ggaggggagg acgcggggg ggggaggagc cgggggatat	9300
ataaagcgg agaaaagcgc ggaatggca tattggacc gcgtgattcg gttgctcgc	9360
gttgtctgt ttggacgttt tttatgcggg aacaagggg cttaccggtt acactgtccg	9420
ctcgtatgg ggtctgtctg tctgtttgg cttgtcgtta tgggagcctg gggggcgtg	9480
ggtgggtcac aggcaaccga atatgttctt cgtagtgtta ttgccaaaga ggtgggggac	9540
atactaagag tgccctgcat gcggaccccc gcggacgatg tttcttggcg ctacgaggcc	9600
ccgtccgtta ttgactatgc ccgcatagac ggaatatctt ttcgctatca ctgccgggg	9660

-continued

---

ttggacacgt ttttgtggga taggcacgcc cagagggcgt atctgggtaa cccctttctc	9720
tttgccggcg gatttttggg ggacttgagt cactctgtgt ttccggccga caccagga	9780
acaacgacgc gccgggccc tttataagag atacgcgatg cgttgggagc tcgaaaacag	9840
gccgtcagcc acgcaccctg cagggccggg tgtgtaaact ttgactactc acgcactcgc	9900
cgtgcgtcg ggcgacgcga tttacggcct gccaacacca cgtcaacgtg ggaaccgcct	9960
gtgtcgtcgg acgatgaagc gagctcgcag tcgaagcccc tcgccacca gccgcccgtc	10020
ctgcacctt cgaacgcccc cccacggcgg gtctccccga cgcgaggtcg gcgcccgc	10080
actgcctcc gacgcaacta gccacgtctg catcgcaagc caccctgggt cgggagcagg	10140
acagccgacc cgtctagcgg ccgggtcggc tgtccagcgt cgtgcacctg gaggtgtcc	10200
gccgggctg atgttttccg catctacgac ccccgaacag cccctggggc tgtcgggcca	10260
tgcgacgcgc cccctgcgca cttccgtgcc cctggactgg gccgcgttc gcgcgcggtt	10320
tctgatcgac gacgcctggc ggcacctgtt ggagccggag ctgcggaacc ccctaaccgc	10380
gcgcctctc gcggagtatg accgtcgggt ccagaccgaa gaggtgctgc cgcgcgggga	10440
ggatgtgttc tcttgacgc ggtattgtac ccccgacgac gtgcgcgtgg ttatcatcgg	10500
gcaggacccc taccaccatc ccggccaggc gcacggcctg gcgtttagcg tgcgtcggga	10560
tgtgccggtg cctccgagtc tacggaacgt gctggcggcg gttaaaaatt gttaccccga	10620
cgcgcgatg agcggccgcg gctgcctgga aaagtgggct cgcgacggcg tgetgttgtt	10680
gaacacgacc ctgaccgta agcgcggggc ggcggcgtcc cactccaagc ttggatggga	10740
ccgttttggt ggcgggggtg tccaacggct ggcgcgcgc cgcggggcc tggcttttat	10800
gctctggggc gcccatgccc agaacgcgat cagggcccgc cctcgccaac actacgtcct	10860
caagttttct caccctcgc cctctccaa ggtcccgtt gggacgtgcc agcatttct	10920
cgcgcgaat cgtacacctg aaacccggga cattatgccg atcgactggt cggataaaga	10980
tgcgacatc cggggtcttg atttacgagg gggcaattaa taaagactgt tgatggttaa	11040
atctcgggtc tcataccggt ccgtgatgtc gggcgtgggg gaagagaggg tcccctctgc	11100
gtttactatc cttgcctcgt ggggctggac gtttgcaacc cagaacctg atcctggcgc	11160
gtcggcaat acgacgccc tagagtcgat tgcggggacc gcaccggacg cgcacgtggg	11220
gcctctcgc gcgagaccgg accgggatgc gatctcccgc cttacgtcga gcgtggccgg	11280
cgaccgccg ggggcccagc gccctacgt caccttgat actctgttta tggtatcttc	11340
gatcgacgaa ctggggcgcc gccagctcac ggatacgatc cgtaaggacc tgcggctgtc	11400
gctggccaag ttcagcatcg cgtgtacca gacctcgtc ttttcgggga cggccgcgcg	11460
ccagcgaag cgcggagcac gcgcgcaacg cacatgcgta ccacgcagca acaagagcct	11520
ccagatgttc gttttgtgca agcgcgccc cgcgcgcag gtgcgcgagc agctcggggc	11580
ggttattcgg tcgcgcaagc cgcgcaagta ttacacgcgg tccctcggatg ggcggctctg	11640
cccggccgtc cccgtgtttg tacacgagtt tgtttcgtcc gaacctatgc gctccatcg	11700
agataacgtc atgtgtctc cggaaaccaga ctaagcacc cgcgcgtccc ctttcttttc	11760
cccctacct tccccgtta ctgatgtgt gtacgtttca ataaataaca cgtagcttat	11820
tttgttgat gatggattga ttgattttat tgaccgttcg ttcccccggc ggtgcgctcg	11880
ccgcgcgcag agggaatatg caacggggc ggggtggggag gaaagaaggt ttcaggtcc	11940

-continued

---

gggggttggg tctgcgctcg ccagggtggg gctgatctga atttcccga gaacctcgac 12000  
 cagtaggtct gttgtgtttg ctgggaactc gcccgcggtt ggggatacgg gggcgggggg 12060  
 tgtggtcggg cggacgtcca ggggtgctt atcgacccc cgcgcgcct cgggggcccgt 12120  
 cccgtagatc gttgcggtga tgtagatggt gtccggggtc cacaccaccg tcaggatgcc 12180  
 ggccgtcgc caaccggacgc ttccgctgct cgatgagctg acccaggagt caaaggggta 12240  
 cgcgtacata tgggctccc accagcctc cagcctctgg gtactagcgc gtccataaaa 12300  
 gcggtatgcg caaaattcgg cagcagctc gataatcacc agcagcccga tgggggtgtg 12360  
 ttgtatcacc acgcctccgc ggggcaggcg gtccctggcg gctcgacccc gcgtcagaac 12420  
 cgcgcgcgtc cctgactcaa acacgtgac cacctgtgcc gcgtccggca gcgcgctcgt 12480  
 tagcgaecgc ctgggtgat gtaggctgta cgcgatggtc gtctgggggt tcccctatgc 12540  
 tcgggggggt gggggtgaat gtcacccggc cgggtgctgg tgggaacgcg agggaatgga 12600  
 gggtaatag acaatgacca cattcggatc gcgtagagca gatagtatgt gctcgctaat 12660  
 gacgtcatcg cgttcctggc gctcccggag cgggtttaga ttcattgtca ggaactcggga 12720  
 tgagggtgtg cgggacatgg ctacgtacgc gctgttagg cgcaggttc cgggcgtgaa 12780  
 gcatatggcg acctgtcca gactgagccc ctgggagcgc gtgatggtca tcgaggttt 12840  
 ggagctgatg ccgtagtcgg cgttgatggc catggccagc tccgtggagt cgatcgactc 12900  
 gacaaactca ctgatgttg tattgacgac agacatgaag ccgtgctggt cccgcaggac 12960  
 gatgtagggc aggggggact cctccaagaa ctccggccag cggccgctcg cgtgcgcgcg 13020  
 ccgcagctcc tcccgaaacg cgaacacccg ggtgtacgtg taccocatca gcgtgtagtt 13080  
 gtccgtctgc agggccacgg acatcagccc cccgcgcggc gagccggtca gcagctcga 13140  
 gcccggaaa atgacattgt ccacgtaggt gctgaagggg gcgctctcaa acacctccc 13200  
 gaagagctcc cgtaggataa ggtatcgc cagaaaggcc ctcttcagga gcccactg 13260  
 ggctggacg gccgcggtg tctcaggctc ttcgagggcg tagtggcagt agaacacgtc 13320  
 cagctgctgt tcgtccagcc cggcgaagat aacgtcaagg tcgtcgtcgg ggaagtctgc 13380  
 cgggcccccg tcccggggc ccagggtgctt aaaattgaac gcacgctccc cgggagagcg 13440  
 gtccgtggtg tcggcggccc tggttgcca tcgcgcggcg gcgtcccggc gtagcgacag 13500  
 gagttctgcc gtcagctccc ctaggcggcc gtaggccagg gtcctctggg tcgctccag 13560  
 gccggggcgc tggagaaagt tgtaaaagt aatcagccc cgaacatga gccgcgacag 13620  
 gaaccggtag gcgaactcca ccgaggtctc cccctgggtc ttcacgaagc tgcctcgcg 13680  
 cagcacagcc tcgaaggtcc gaaaactccc gtcgaaccca aacaccatct ttcggaggcg 13740  
 cgcggtcacc gcgacctggc tggtgaggac gtacgtgatg tcgttccggg ccacgactag 13800  
 ctggtgcttg ctgtgcaect cacagcgcac gtgcccgcg tcctggctct gactctggga 13860  
 gtagttggtg atgcgactgg cgttgccgct gatccacttt tccatggtca gcgtggggtg 13920  
 ctgctgagc cgtcgatact cgtcaaacctc tttgaccgac acaaacgtga gcacggggag 13980  
 ggtaaacaca acaaacctcc cctcgcgagt cacctttagg taggcgtgga gcttgccat 14040  
 gtacgcgctg acctccttgt gggacgagaa cagccgcgtc caccocgaa ggttgcccg 14100  
 gttggtgatg taactttccg ggacgacaaa gcggtccaca aactgcatgt gctcctcgg 14160  
 gatgggaagg ccgtactcca gcaccttcat gaggttcccg aactcgtgct ccacacatcg 14220

---

-continued

---

cttgttgta atgaaaatgg cccagctgtg cgagaggcgc gtgtactcgc gtaggggtgcg 14280  
gttgccagatg aggtacgtga gcacgttttc gctctgccgg acggagcatc gcagtttttg 14340  
gtgttcgaag gtggactcca cgcagggcgt ctgggtcggc gaccccacgc acaccagcac 14400  
cggccgcagg cggcccgcgt actgggggggt gtgggtacagg gcgttaatca tccaccagca 14460  
atacaccacg gtcgtgagta ggtgccgccc caggagcccg gcctcgtcga tgacgataat 14520  
gttgctcggg gtgaaagccg gcagcgcgcc gtgtgtgacc gaggccaggc gcgtgagggc 14580  
accctggccc agcccccagg tctgctctag ggcggtgagg gcgtggaact cgtttcgcgc 14640  
gtcttcgccc ccgtgcgccc ccagggcccc cttggtgatg tcgaggatca cctcccagta 14700  
gtacgtcagg tctcgcgctg gcaggtcttc cagcggggcg gggctgctgg ccaggggtga 14760  
cgggtgctgc cccagctggg cctggaactg attcccgcga aaccogaact cgtgaaagat 14820  
gggtgtgatg ggtcgactca gaaacgcccc cgagagctta acgtacatgt tctgcgcccgc 14880  
gattcgcgtg gcgcccgtga ccacgcagtc caggacctcg ttgagggctc gcacgcacgt 14940  
actctttccg gatccggcgt tgcccgtgat gagatacgcg gcgaacggaa actcccggag 15000  
cggcaggccc gtcgggacct ccaaggcccc cacgtcccgg aaccactgca ggcgcggcac 15060  
ctgcgtgacg tcgagctgct gctcgcgagag ctctcgggatg cgtgcgatga ttggttgac 15120  
cccgtgcatg gacgtaaaat ttaaaaacgc ctcgtccctg aaccgcacgg cgggtctggc 15180  
cccgggctgc tgtggggggc gacctggtgc ccggaactcc cgcgagccct ccccgcggga 15240  
cgcccctatg gccgcacagc gcgcgcgggc gccggcgatg cggacgcggg gcggcgacgc 15300  
ggcgtatgc gccccgagg acggctgggt gaaggttca cccacccccg ggacgatgtt 15360  
gttcgcgag attctcctcg ggcagatggg gtacaccgag ggtcaggggg tgtacaact 15420  
cgtccggtcc agcagggccg ccccccgaca gctgcaggcg gcgatcttc acgcgctcct 15480  
caacgccacg acgtaccggg acctggagga ggactggcgc cgcacgtgg tggcccgcgg 15540  
cctccagccg cagcggctgg ttccgagga ccggaacgcc cgggagggcg atatcgccgg 15600  
gggtggccgag cgggtgttcc acacgtggcg atgcacgctc aggaacgcgc tgetggactt 15660  
tgcccacggg gtggtagact gctttgcgcc gggcgccca agcggaccca ccagcttccc 15720  
caaatatc gactggctga cgtgtctggg gctggtccc atattgcga agacgcgca 15780  
gggggaggcg acgcagccgc tggggggcgt tctcaggcag cacacgctgc cccggcagct 15840  
ggccacggtc gccggggccg cggagcgcgc cggcccggg cttctggatc tggcctcgc 15900  
gttcgactcc acgcgcatgg cggaaatcga ccgctgcac atctactaca accatcgccg 15960  
gggggagtg ctgggtcgcg acccggtcag cgggcagcgc ggcgagtgcc tgggtctgtg 16020  
cccccccctg tggaccggcg acccctgggt cttcgattcg cccgttcagc ggetgtgccc 16080  
cgagatcgtc gcgtgccacg ccctccggga acacgcgca atctgccgtc tgcgaaacac 16140  
cgcgtccgtc aaggtgctgt tggggcgcaa gagcgacagc gagcgcgggg tggctggcgc 16200  
cgcgcgggtc gtcaataaag cgctggggga ggatgacgag acgaaggccg gctcggccgc 16260  
ctcgcgtctc gtgcgctca tcatcaacat gaagggcatg cggcactgg gcgacatcaa 16320  
cgacacggta cgcgctact tggacgaggc gggggggcac ctgatcga ccccgcctg 16380  
cgaccacacc ctcccgggt tcggcaaggg cgcaaccggc cgcgggtcgc gccccagga 16440  
cccggggggc cgaccgcagc agcttcgcca ggcgtttcag acggccgtgg tcaacaacat 16500

---

-continued

---

caacggcatg ctggagggct atatcaataa tctctttgga accatagaac gcctgcgaga 16560  
gacgaacgcg ggtctggcga cccagctgca ggcgcgcgac cgcgagctgc ggcgcgcca 16620  
ggcggggcgc ctggagcggg agcagcgcgc ggcggaccgg gcgcccgggg gaggcgcggg 16680  
cgcgccggcg gaggcggatc ttctccgggc cgactacgac attatcgacg tcagcaagtc 16740  
catggacgac gacacgtacg tggccaacag ttccagcac cagtacatcc ccgctacgg 16800  
ccaggacctc gagcgcctgt cgcgcctctg ggagcacgag ctggtgcgct gcttcaagat 16860  
tctgcgccac cgcaacaagc agggccagga aacgtcgtac tcgtactcta gcggggcgat 16920  
cgctccttc gtggccccgt atttcagta cgtgcttcgc gcccccgag cgggcgcgct 16980  
catcaccggc tccgatgtca tcctagggga ggaggagta tgggagcgg tctttaagaa 17040  
aaccgcctg cagacgtacc tgacagacgt cgcggccctg ttcgtggcgg acgtacagca 17100  
cgcggctctg ccccgcccc cctcccaac ccccgccgat ttccgggca gcgcgtcccc 17160  
gcggggcggg tcccgtccc ggaccggac ccgatcccgg tcgcccggga gaacgcgag 17220  
gggtgcgccc gaccaggggt ggggcgctca acgcaggat ggcgcacccc acgcccgcg 17280  
atgaggaac ggcgcgcgc atcctcaaac aggcacgcgc cggggaccgc agtctggtcg 17340  
aggtggcga ggggatcagc aaccaggcgc tgcgcgcat ggcctgcgag gtgcgccagg 17400  
tcagcgatcg ccagccgccc tttaccgca ccagcgtcct gcgcggtgac gtcacccca 17460  
gggggcggt gcggttcggt ctggacggga gttccgacga cgcgtaegtgc gcgctggagg 17520  
attacttta gcgctcggg gaccagcga cgtatcgcgg ttttgcggtc gtcgtectca 17580  
cggccaacga ggaccacgtg cacagcctgg ccgtgcccc cctcgttctg ctgcaccgca 17640  
tctccttgt tcgccccacg gacctccggg acttcgagct cgtctgctg ctgatgtacc 17700  
tgagaaactg tcccggagc cacgccacgc cctcgtggt cgtcaagggtg tcgctggtg 17760  
tgggggtcgt ggcgccccac gcgtctcctc tcgagcgcgt ccgctgcctt ctectccca 17820  
gctgccactg gatcctgaac acgctaagt gcatggcggg cgtgaagccc ttcgacgacg 17880  
agctagtct gcccactgg tacatggccc actacctgct ggccaacaat cgccccccg 17940  
tcctctcggc cctgttttgc gccacccgc agagctctgc gttgcagttg cccgggccc 18000  
tccccgcac ggactgtgtg gcctataacc cggccggcgt catgggaagc tgctggaatt 18060  
ccaaggacct gcgttcggct ctggtgtatt ggtggcttc ggggagcccc aaacgcgga 18120  
cctcgtcgt tttctatcgg ttttctaac tccgaaaat aaacgtgttt tttatggaac 18180  
gttccccacc tgcgtgtca tctctcggg gatggtggtg ggcctgtgtg tgtgtctgt 18240  
gcaccgaagg aggaaagtgg gggggtggtg gtgctggtgg tggaaagaca tgatagagg 18300  
aacaagaaa tagaagaaa ccacaaccgg cgcgtgccag taaatacga cgcgcgcaca 18360  
cgcggggggt aagttggagc acggggcccc ggtttattga ccaaatcag ggaacagaa 18420  
accgaatctt ttcacgaaa gggtaacaa agctcccgc ctcgcccac acgcctcca 18480  
gaacccccgt aaacaccagt tgaatctgc gcaggatctc gcgcaggtga tgggcgcagt 18540  
ccacgggggg gagcaccagc ggcgcgggt acagatccac ggggacgccc accgactccc 18600  
cgccccggg acatacgcgc acgacgcgtc tccagtattg ctccgcgtcc agcaggcgc 18660  
ctcccgga ggcggtttgg ggcaggggt cgtcggcctc gcctgggggg gtcagaacgc 18720  
tccagtactc cgcgtccaga cgcctccga aggcattccg gacaaagcgg tcacaggcgt 18780

-continued

---

cctccatgac gcccccggcc gcgcacacgg cctcctccgg cgggccggcg gccggccgcc 18840  
 ggaggattcg tctcagcgcg tcgcgcataa cctcggcccg cgcggcgtag gcggccccgc 18900  
 ggagaggaaa tcctctcagg aagtcgggtg catcgcggga gttccagaac caccccccgg 18960  
 tctggctcca ggtgacgacg tgggtgtaga cgcctctcgg cgcagggag gggcgagggc 19020  
 gcggcgatg gccgttggcc gaaagtagcg gcgcacgga cgcctcagag gcccgcgggg 19080  
 cgtcctggat cgcgcgctgc cgcgcgtccg cgtccccggg gtccaagtgg aacagcccc 19140  
 agaacgcagc cccggtgccg ccgcagaccg caaacttcac cgagctggcc gtctgctcga 19200  
 tctgcaggca gacggcggcc atgaccccg cgcagcagct cggagcgcg gggcagggct 19260  
 cgcacgcgtc cgcgaccagg cgctccagca cggccccggg ccagggtctc gagggggcgg 19320  
 ccgcaaccag cgcgtccagc ctttcagggc ccgccccccc cggggtctcc gccagccccg 19380  
 cctccccgag gcccgcgagg gcggccagga gctgggctcg gagcccgag aaacaaaacc 19440  
 gcgcgctcca gaccggcccc acggcccgcc gggggctcag tagttgatg gtggtggccg 19500  
 tggggtgcca ccgcgcgacc gcttccccga aggcgggag gagcgggccg gccgcctccg 19560  
 aggccacggc cggccatgcc cgcgggggca ggacgacct ggcgccacc gccggccagg 19620  
 ccccaggca cgcggcatgg gtggcccgcg cgcgccgac caggtcacgc gccgactcgg 19680  
 cggcggcggc ggccggcagc gtaaacgtgg gccagccccg aaatcccagc acggcaagt 19740  
 attggacggg cctccccggc acctcaaac cgggccccag aaaagcgaag acgggggcca 19800  
 gggctccggg ggccggcgtg accgtggtat gccactgccc gaagagggcg accagcgcg 19860  
 gggcggagaa cccgtcgcg gcgctcagca agtagtcgta gccgcggcg agcagcacc 19920  
 gcgcgctgac ccgctcgggg gtgcccgggg gccgcaggcc gacctcgac acctcgacca 19980  
 ggtccgcgaa ggcgccctcc ttctggtcg gcgaaacgc cagggtggtg tattcgcgcg 20040  
 caaaaacgcg ggtcctcgtc gtgatggtga cgcgcagcga ggcggaggac gcgcaactgg 20100  
 ggctgtcgcg aatggcggcc aggcgcgccc acgccaaccg cgcgcggggg tgcctggcga 20160  
 cgcgcgcgga cagggccagc gggctcagct cgaccttggc ctccactcc aggagggcgg 20220  
 cgcgaggagc ggccggcggg ccccacgacg cctttcgac cctcagacc agaccgctt 20280  
 gcgggtccc acccagggc agcgggaaga agagggcca cgcgcccgtc tggcgtcca 20340  
 gggccgccag aacgcacgca tacagcggcc gccacagggt cgggtcccc aggggtcca 20400  
 gcggggaggc gcccggggcc gtcgcggcgc gggcgccgc gacggcccc gggcccgaga 20460  
 cgtcggggga gccgtagaag tcctgcaggt cggacgaacc aacggacacc tccgcaagc 20520  
 gcgcgcgcgc ctcccccgcg cgcgtcgcag agaccagata cagcagggcg tggaggcagt 20580  
 cgcgcgtgcg cgggggcagc catacccgct ataggtaat ggcgtgacg ctctcctcca 20640  
 cccaaacgat gccgggggct tccatgccac gacgcccggg ggttgccgtg tatcgaacga 20700  
 gcgcggcccc agacttatag ggtgctaaag ttcaccgccc cctgcatcat gggccaggcc 20760  
 tcggtgggaa gctccgacag agccgcctcg agaatgatgt cagtgttggg ctgggcgcg 20820  
 gagcgtgcg tgcgaagca gcgccccac gcgggcgcgc gcagcttga gcgcgcgccc 20880  
 gcaaacctcc gcttatgggc catcagcagc gcgtacagct gtctgtgct cggcagggc 20940  
 ctgtggtcga tgcggtgggc gtccagcagc tccacgatgg ctgccttggg gaggtttta 21000  
 acgcgccccg ccccggaaaa cgtctgctg ctcttggcca gctgcacccc gaacagtctg 21060

-continued

---

ccccagatga tcttgaacag cgacagcgcg tgctccgtct cgctcacgga cccgcgcggg 21120  
gggcagccgc tcagggcgct gccacgcgc ttaaccgct cctccgacag caagggggcg 21180  
tcggtcacgt tacagtggcc cagttcgaac accagctgca tgtagcggtc gtagtggggg 21240  
ttcagcagct ccagcacgct ctcggggcta aagggtcgcc cggaccccc ggccatcgag 21300  
tcccactgca ggcacgcggc catggtgctg cacagacgga acagctcca gacggggcg 21360  
acgtttaggg tggggtgtag ggcacaagc tccagctctc cggcggcgtt gatcgtgggg 21420  
atgacgcccg tggcgtagt gtcgtaaagc cgcggaaga tggcgtgct atgggcggcc 21480  
atggggacgc gaagacaggc ctccagcgc accaggtaga tgaaccgct gcggccgacc 21540  
aggtgttga ggcgcgcgat gagcgcgacc acctcggcgc gcgcgacgct cggccggagg 21600  
tacttttoga cgaaggcc cacctctcc gtctcggcgc cctgggcca cagggacgtg 21660  
tcggggtcct ggcagcgcag ctccgcaga tccgctggg ccctcagggc atcaaatgt 21720  
atccccgca aaaacagaca aaagttcctc ggggtcagc cggcgtcgtg gcccagAAC 21780  
cgcacgtgca tgcagttgag ggtcagaagc atgtggagga tgtaagact gtccgcgagg 21840  
cacgccagc tgcacctctc gaagtagtgc ttgtaccgga attgctgta gatgcgcgac 21900  
ccccgcgct gcgcgcgctc ggcgtgcgac gcgtcgcag gccctttaa cggcggcac 21960  
aacaggttcg tcacctggga aaactgtgcc ggccactgcc cgctggcgt caccacgtgg 22020  
ttgacgagca tgggcgtaaa gacgggctcc gagcgcgccc cggacccgct catgtagatc 22080  
agcagctccc ccttgcggag agtcggtacc cgcgccagc actggtacac ggacaccatg 22140  
tccggcccgt agttcatggg ttccacgtag gcgaacatgc tgtcaaagt cggcggatcg 22200  
aagtaaggc ccaccgtcac gaccgtgtg tagatgacca cccggtaccg gcccacgtg 22260  
gtcacgtcgc cgggcggggg gacgagtgag agcagcagca cgcggtccgt aaactgccg 22320  
cagaacctgg caacgacctc cgcgaaggag accgtcgcag agaagatgca gacgttatct 22380  
ccgcggcca ggcgcgctc cagctcccc aagaagggtg cgtccggggg ggcgtccggg 22440  
gggggcgccc cgcgcccg ccccgcgcg cgcagggcgc cctgcaggac ctcgggccc 22500  
aggcgcggga gaaacagaca acggcgcgcc gaaaatccgg gcatggcgta ctcccgatg 22560  
accacgtgaa cgttcttttc gccccggagg ctgcacagaa agtccaccag ctgcgcgttg 22620  
gcggtggcgt ccatggcgat gatcccggg cacgtgcgca gcaggcgcag catcaacgcg 22680  
tcgacgcggc ccagctgctg catcgtcggc gactacagt ggcccaact cgcacatgact 22740  
tcgtccagga cgagcacgct gtagttgttc aacaggttcg ggcccacgc atgaagactt 22800  
tccacctgca cgatgagagc gtggaagggg cggtcgttca tgatgtaatt ggtggatgag 22860  
aagtaggtga cgaagtccgg caaccctgac tcagcgaacc gcgtcccgag ggtctgagta 22920  
aaactccgac gacaggagac gaccagcaca ctctgtccg gagagtggat cgttcccc 22980  
aaccagcggc tcagcgcggt agtttttccc gagccattg gcgcgcggac cacagttacg 23040  
caccggcccg tcggggcgct cgcgtccggg aagggtgacgg gtccgtgttg ctgcgcctcg 23100  
atcgttgttt tcgggtggac cgggggaacc cactcggcca aatccccccc gtaaagcatc 23160  
cgcgccagc atacactcga cgtgtaactc tcgcactcgt cateccgat gggacgcccg 23220  
gccccaggg gatccccga ggccgcgccc ggccgcgac tcgcgcccgg ggcggggcg 23280  
gcgtgggtgg tctggtgtgt gcaggtggcg acgttcatc tctcgccat ctgcgtcgtg 23340

-continued

---

```

gggetcctgg tgtggcctc tgtgttccgg gacaggtttc cctgccttta cgcccccgcg 23400
acctcttatg cgaaggcgaa cgccacggtc gaggtgcgcg ggggtgtagc cgtccccctc 23460
cgggttgaca cgagagcct gctggccacg tacgcaatta cgtetaegct gttgctggcg 23520
gcgcccggtg acgcccggg gggcgcggtg acctcgcgct acgagcgcg gctggatgcg 23580
gcccgtcgcc tggcggcggc ccgatggcg atgccacacg ccacgctaat cgccggaaac 23640
gtctgcgcg ggtgtttgca gatcacagtc ctgctgctgg cccaccgcat cagccagctg 23700
gcccacctta tctacgtcct gcactttgcg tgctcgtgt atctcgcggc ccatttttgc 23760
accagggggg tctgagcgg gacgtacctg cgtcagggtc acggcctgat tgaccggcg 23820
ccgacgcacc atcgtatcgt cggtcgggtg cgggcagtaa tgacaaaacgc cttattactg 23880
ggcacctcc tgtgcacggc gcgcgcgcg gtctcgttga acacgatcgc cgccctgaac 23940
ttcaactttt cgccccgag catgctcctc tgctgaega cgtgttcgc cctgcttgc 24000
gtgtcgtgt tgttgggtgt cgaggggggt ctgtgtcact acgtgcgcg gttggtgggc 24060
ccccacctcg gggccatcg cgccaccggc atcgtcgggc tggcctgcca gcactaccac 24120
accggtggtt actacgtgtt ggagcagcag tggccggggg cccagacggg agtcccgctc 24180
gcccgtggcg tegtgcggc ctttgccctc gccatggccg tgettccggtg cagcgcgccc 24240
tacctgtatc accggcgaca ccacactaaa tttttcgtgc gcctgcgcca caccggcac 24300
cgccccatt cgcgcttcg acgcgtaacg agctccatgc gcggttctag gcgtggcggg 24360
ccgccccgag acccgggcta cgcgaaaacc cctacgcca gcgtgtccca ccacggcgag 24420
atcgaccggt atggggatcc cgacggggac ccgatctacg acgaagtggc ccccgaccac 24480
gaggccgagc tctacgcccc agtgcaacgc cccgggcctg tgcccagcgc cgagcccatt 24540
tacgacaccg tggagggtta tgcgccaagg tccgcggggg agccggtgta cagcacgctt 24600
cggcgatggt agccgtttcg ttcgttttaa taaaccgacg ttgtgcgttt caccatactt 24660
cggcgcgcg gttgtgtgtt tttttttttt gtgggtttta ttttcccccc accccttct 24720
tttctttcgg ccaccacccc cctcctcccc cgtactatac aacaaaaaat accacacata 24780
cgacaaaata cggacaatca tttctgtctt tattegctat cagagagtgg gggcgtgagc 24840
gtggcaggag ggcggggccac gtcgggggtc cgccgtctgg tgtgacgcca tgggggggtc 24900
gatgcgcgcc ggtactgggg ccccgggccc cgggtgacca cgcgcacgtc ggggggcaag 24960
tagaagtac cctcttcttc ggactcgatg tccacgacgt caaatctgtg ggcggtcagc 25020
gagacgacct ccccgccgct ggtggtgatg acgttgtgtc ggcagcagca gggccgccc 25080
ccggagaacg cgaggcccat aacttggcga gcgtatcgtc gaaggccagg cggtgtttc 25140
gccgatgctc ccggtagatc cccggctcga cgcgacggg ggtgatgatc agggcgatcg 25200
gaacggcctg gtcggggagg atcgatgctt tggcgggtcc gggggcccgc ccaggcccgg 25260
cggcgctcc gcggccgtcc tccaggcgga acgtcacgccc ctctccgcg cccgcgcggt 25320
gcctgccgag gaacgtcacc aggtgcggtt gcagggggca gtcgggaaag tggctgtcga 25380
ggacgtatcc ctgcaccaag atctgtttga agttcgggtg cggggggttg gcgaagatgg 25440
gctcgcgggc aaccagctcc ccggagctcc aggccacggg agagatggtg cgacgctcaa 25500
ggtcggggac gccaacaga agcacctccg agacaacgcc gctatttaac tccaccagcg 25560
cccgatccgg ggcggagcat cgcctttttt cgccggcggc gcgggaaatcg agccagtccc 25620

```

-continued

---

ggtcttgggt gacgagcgcc tctctcgggc ccggaacgcg cccgggcgcg aagtagcgca 25680  
 cgccgggggt ggggatggac cggatgaacg cccggaacg ctcggcgat cgccgcgcca 25740  
 tcaggtctct gtacgcgag gcccggggg cgccggggtc cgcggggtcg aacgctact 25800  
 tggctcggca cttaacctcg tagaaggcca ggggggtctg gggggcggg gccaggtagc 25860  
 cgtgagggtc cctggggcac acgaggatgt ccagggacgc ccccaccatg cccgtgtggc 25920  
 cgtccatgag gaccccgac gcgtgcacgt tctctcggc gaggtcccg ggttggtaga 25980  
 agacgaagcg cccggcgctg gcgtcgtctg tgaccccg gtccgcgcg cccacgcagt 26040  
 agcgaacag caggtttcgg gccctcggct cgttcaccg cccgaacatc accgccgacg 26100  
 actgggctc cagcccgag ctggcgttgt gggtagcca ctgggacgag aagcacggac 26160  
 cctgcgccc ccaccgcag gtggaggcgg tcgtcaggcc ccgccaagc agggcccaga 26220  
 gctggcagtc ggccctggtt tgcgtcggc cctcgtaaaa tcccataagc gggcgggggg 26280  
 cgacggctc ggcggcgac gggggggcgc ggcgctcag gcgccagag tgccggccga 26340  
 gcccgcgtc caccatgccc gccgcctcca gcgacacgac gagggagcac agatagtcca 26400  
 ggcgagccca caggggccc atggccagag gggagcggac gccgcgacg aggcccgca 26460  
 ggtggcgctc gaacgtttcc gccaaagat gggggggcag tgcgttggg atcgccgacg 26520  
 ccgaccacat cgggtcgggg tccgggggac cggggctgca gtccgggtcg atggcgtgtg 26580  
 cgcccccg cgagagggga atgtcgggg ttggcgggccc ggatgaggcc tcagagaggg 26640  
 ccggggaacg gggccggggc ttttcgccc gggcccgcg gtcgggttgc ccactgggg 26700  
 ggctctggg ccaatgggaa cccggggccc ccggtgacgt ggggcggggt ggggcggggc 26760  
 ggggcccaca gacggtcggc agatctaggc tgttgggtcg gggccgctc gggggactat 26820  
 cggggtcgg ggcggggtcc gcggggcgt tggcgccgg tgttgcggc gccgccattt 26880  
 ttacgagcag ccgaagagct cgagggcgga agggatcctc acgacagaga gtggcgcgcg 26940  
 gccgggttg cgtgacagag gcggggagacc agcaccagca gcggcctcag ctccggcggc 27000  
 agcgacacc acgacaggac ggccttctgc gtgcgctggt aattatata ctgctccgtg 27060  
 aacgcgccc gaactctggg attcgaagg tggcgccgga tgcctcggc cactcatac 27120  
 gccaggccgt ggggttggc ctcggccgag ttgacaaaga gggcggggtg cagaacgcag 27180  
 cgataggcga ggagggccc ggcaaaagtc ggcgagagct ggttgttaa gtactggtag 27240  
 cccgggacgc ggtcacggg gacgccagc ctcggggcca cgtacacgt aaccagcagc 27300  
 tccagcagc tctgcccag ggcgtagaga tcgaccgcca gcccgacgct gtgcttcagg 27360  
 gggcggttgt taaactcggc ccgctcgtt ttgaggtact ttaccgagag ctccgggtggc 27420  
 tggttgtacc cgtgccccac cagagtgtga aagttggcgg tggtcagggc ggcgggcatc 27480  
 ccaaaccccc ggggggactc gaggtccggc tctggaggc aaaactggcc ccgggatatc 27540  
 gtggagttg agttcagggt caccaggcta aagtcggcca ggacggccgg ccggagcgac 27600  
 acccgctcc atcgagcat cagaggacg ttggcgact tgatgtccag gtggctgac 27660  
 ccgcacctg tgttcaggaa caccacggc gcgccaggt ctgtgaagca gtggtggagg 27720  
 gccgtcgcga cggaggggg ggtcgcgccc agggacgcca gctggccgat gtacttgcgg 27780  
 aggtccatgt cgtacgggg gaacacgac tggcgctgct gcagcgagaa cccgagcggg 27840  
 gtgataaagc cgcggatgct gtgggtcggc ccgcccggaa gagcgactc ccccacgagc 27900

-continued

---

agggtcgca	cgagctccac	ggcaaacac	tctttttccc	ggatggctt	cacggcgagc	27960
ttgtgttcgc	gaatcaactg	cacctcgccg	tacccccccg	agccccgaa	gctgccccgc	28020
ccgggatct	ccagggctgc	gtagcggagg	gcggggttga	cggcgaatac	ggggatgcat	28080
agcttgtgga	tgcgcgcgag	ggacaggatg	tgcgaggggg	gcgacggggg	cgaggtcatg	28140
gccgtctcgg	acctgcgcag	ggcgggggc	cttagcttgg	ccgcagggcc	gggggcctcg	28200
ggggacgagc	ggcagcagag	cgagcggctc	actcgccatc	gggacagtcc	cgcgcaagc	28260
cgctccccga	agctggatcg	gcggcgggac	ccggggcggg	ctccggagac	ggcgccgtct	28320
cggggggagg	ggcgccttgg	gcgtccggac	gcccgcggc	tgagggagtg	tatgtaggac	28380
gcgagccagg	ccttgaaggag	gcgtcgggtg	gcacctgggg	ggctgatgtc	agctgccaca	28440
tgactagcag	gtcgtgtcgc	cccggactca	tccatccgtc	cgccaggtcg	ccgtccccc	28500
acagagacgc	gttcgcgcg	gcctcttcga	gctgtctctc	ctggctccga	agacgatcgt	28560
ccgcgcgctc	caggcgcctc	ctaagcgcg	gatcagagta	ccgtcgggtg	gcggttagaa	28620
aatcacgtcg	cgccgcttgc	tcttccacgc	gaattttaac	acaggtcgtc	cgctgtcga	28680
tcctctctaa	gcgcgcgcgg	gactttagcc	gcgcctccaa	ttccaagtgg	gccgccttgg	28740
cgccataaaa	ggcgcacaaa	aacctaggat	cttgtgtact	cacgccctcc	cggtgtagct	28800
gcagggctcg	gtccctgtac	acctcggccc	ggagggtgct	ctcggccaaa	cgtcggcga	28860
gggcgcgctg	gctggcgtct	cggtcatct	cgccgcccc	gcgcgcgcc	gacgtcggac	28920
tccttcgccc	cgaccccc	gacctcagcc	gcccccgct	cgccgcgat	gtttgccag	28980
cagctggcgt	ccgacgtgca	gcagtacctg	gagcgcctgg	agaaacagag	gcaacagaag	29040
gtgggcgctg	acgagggcgt	ggcgggcctg	acgctcggcg	gcgatgcgct	gcgcgtccct	29100
tttttgatt	ttgccaccgc	gacgcccag	cgccaccaga	ccgtggctcc	ggcgctcggg	29160
acgctccacg	actgctgcga	gcactcgccg	ctcttctcgg	ccgtcgcgcg	gcggttgcgt	29220
tttaataagc	tggtgcccgg	gcaactcagg	ggcggtgact	ttggggcgca	ccacacggcc	29280
aagctggagt	tcctggcccc	cgagctgggt	cgggcgggtg	cgccctcgcg	gtttcgggag	29340
tgcgcgcg	aggacgcgct	gccccaacgc	aaagcctact	acagcgtcct	gaacacgttt	29400
caggccctgc	accgctccga	agcctttcgg	cagttggctc	acttcgtcgc	ggacttcgcc	29460
cagttgttga	aaacctcgtt	ccgggcctct	agtctcgcgg	agactacggg	ccccccgaag	29520
aaacgggcca	aggtggacgt	ggccaccac	gggcagacgt	acggcacctt	ggagctcttc	29580
cagaaaaatga	tactaatgca	cgcgacctac	tttctggccg	ccgtgctgct	cggggaccac	29640
gcggagcagg	tcaacacgtt	cctgcggctc	gtgttcgaga	tccccctgtt	tagcgacacg	29700
gccgtgcggc	acttcgcgca	gcgcgccacc	gtgtttctag	tcccaggcgc	ccacggaaaag	29760
acctggtttt	tggtgcccc	catcgcgctg	tcgctcgcgt	ccttcggggg	gatcaagata	29820
ggctacacgg	cccacatccg	caaggcgacc	gagcccgtgt	ttgatgagat	cgacgcctgc	29880
ctgcggggct	ggtttggctc	gtcccgggtg	gaccacgtca	aggggaaac	catctcgttc	29940
tcgttccccg	acggctcgcg	cagcacgata	gtgtttgctc	ccagccacaa	cacgaacgta	30000
agtacccctt	cctccccgcg	tgcctgtttc	cccgggtgccc	ccctccccga	gatcgaccga	30060
cagacaaaaca	cagccagacg	cgagtggtgg	acgacacgcc	cgacgcccc	ccccgccat	30120
ggcggggggg	aagccttact	gtttatttgt	aatcggacga	tgaggctctg	gccacggccc	30180

-continued

---

g c g c g a c c g c g g g g c a g c t c g t t g c a a a c a g g c g g t g g t a t a c g a t g a c a g a a c g c a g a 30240  
g g c g c c a c c c g g c g t g g t c g g g c g g a t g a c g t t t c c g c c c g t c c c g c c a c g a c g a 30300  
c c t c g t g c a g g t g g g c c g t g a t g c g c g g g c g g c g g g t c g c t g c c g a g g a t a a c c g c g t 30360  
c c a c g g g g t g c c g a a g a g g a g c t g a c a c a g g c t c g c g t c g c t c c c c g g a c g g c c a g g g t g c 30420  
g c t g g g c c a t a t t g g a c c a c a t g c a c g g g g c g a c g c a g g g a c a g g c c t c c g c c a c g g c g g 30480  
g g g c g c g c c a c a g c g c g t g c g g a a t c g a t g t g g g c c g t c g g g g c g c a g g c g c c t c 30540  
c t c c c g g g g g g t g c g g t a a t c t g g a t a g c a g c c a t c t a a t g g c g g g c c c g g t g c c c g 30600  
g g g g a c a g a g c g a c c c c a g g t c a t c a t c e c a t g g c c c a g c a g t a t a t g c g g c c g c g g g g a 30660  
g g t g c c a c c a g g c c c c c g g a c c c a g g g c a c a g c a c g c c c c g g a t t c g g g g g c c g t g t c c g 30720  
t g g g t a c c a g g t a g g c g c g c g t c g a g c t c g t g g c c a c g g g c t c g t c c g c g a g c t g t t c g g 30780  
c g g c g g g g t c g g g g t t t c t c c g g g g g g a g g c a g c t t c a g g t g g c c g a a g g c t a g g g 30840  
t g c a c a g c a g c g g g g t c c g g g t g c g g t g c g t t a c g t g c g g a g t g g a c g g t g c g c a g t a g c 30900  
g g c g c t c g c g g t t a a a g a a g a a a a a t g g c a a a g a a c g t g t t c g a a g g c a g g c g c a g c g c t 30960  
t g g g c c g c g t c a g g t a c a g g a a g a t c t c g c a g a a a a g g g c a c g c t c g g g t c c g 31020  
g a a g g g c c a c c t g g c a c a g c g g c t c g g t g a g g a c c g t g a g g a c c g a a a a t c t t a a g c c 31080  
g c t c g t c c c c c g a a c g a c g c g c c a c a c g a a g a c a g a g t t g g c g a t g c g c g c g a c g a g g t 31140  
c g g c t t c g g g c c c c g g g t c g g g g c g c g c g t c g g g g g g g g c g c c c c g g t g a c c c g g c g 31200  
g g g c g c g c g c t c c c g g g g g c c t g g c g t c g c c t g g g a c g c a g a g t g c c c g t g t g c c a 31260  
g g t t g g t g g t g g g a a g g g a c c g g a g a c g c a c c a a a a g c a g a g g g g c c a g c g t g t a t g 31320  
a g t t g g g g g g g g g t g g g t a g c g g t g g a a c a a a g c a c g c g t c a g c g g a c a a g g c c g g g 31380  
t c c c g t a g c c g c c c c g a c a g a a c c g g a g t c c g a c g g c a c g c g c g a c g g g g t c t g c g a g 31440  
g c t g a g g t a c g c c g c g g t g t a a t g g t a a a c g a a a g c c t c c c g a a a g a c c a c t a g c c c 31500  
g c a g a g g c g g c g a t t g a a c c a a g g c a g a g t a c g c g t a g c t c t c t c c c g a a g g t a t t g 31560  
c t c g a g a c c c t g t g t g g g g a g t g g a g g g g t g c c c t c c a t g a a g c g a c a t t a c t c t g 31620  
c t c g c g t c c a t t a c g t e a c c g t e a a t c a c a c t g c g a t t g g a c g g t t g g t a g g c g c a g 31680  
c g t g t c t c c g c t g g t g c t g t a g t a g t c a a a c g c g t a g t g g c g t c g g a g t c g g c g a a g c g 31740  
g g c g g g g a t g t e g t e g c t g a g a g g a c g a g c c g c c g c g c c c c g a c c g c c t g g c c 31800  
g c c c a g a t g c g c c a c a c g g c a a g g g c g t a c g c g t g t g a a a a a c g c g t c g g g g c g g t 31860  
c c c c t c g a g g c g c g c a t c a g g t t c t c c a g g a g c a c g g g a a g c g c c g c g t a c c c t c c c c 31920  
t a g c c a c t c g c t c t g t g g g g g c c a a a g t c g t a g c g a g g c g t g g a a g a t g c g c g g g g c c 31980  
g c c t t g g a g c g c g c c c g g a t a g a g t g g c c a g g g c c c g c a g a c a c g c a t c t g g a t g c g 32040  
c g c a c a a a g g c a c c a c t c g c c g c g a t g t c a a g g g c t g c a g c a g g g c g c g g t g g c g 32100  
c a g g g g t c c c t c g a g c g c g g a a a g c a c a c a g c a g c a g c g c c g t c t g g g c c g c g g g g a c a g 32160  
c t g g t g g g g g c g c a c a c a c g c t c g g c g g c a c a g g c c t c c g t c a g g g c c g t g g c c a g t c 32220  
g g a g g a c a g c g c g g g g g c g g g c g c g t c c c g c c c c a c g c c a c c g a a t t c t c g t a g g a 32280  
g a c a c a c a c g a a g c g t g t t g t c c c g t a g t g a t g g c g a g g a c c a c g g a g a t g g a g c g 32340  
a c g g c t c c a c a g c a g t c g g c c g g t c g c c c g g c a g a g t t c c a c c c g c g g t c c a g 32400  
c c a c t c g a c c a g c a t c g c g c t t g g c g g t c c c c g g c a c g a g g t g a g c a c g t c g t t g a g 32460

-continued

---

```

gacgtcctcg cccgcggccc gggggccccc ccggctggca aagcgccccc cgccgggcgg 32520
ctccaggccc gccagcaccg cctccgcgtc cgacgcgccc agggctcccc cgtgacggc 32580
ctggtggacc agggcgccct ggcggagccc cgaggcgacg ccggaggccg cgtgcttggg 32640
gcgcgcgagg accgggtggc ggcgggtgac gtectgcacg gcccgctgga ccagcgcgag 32700
gatctcctcg ttctcttgcg tgatggacac gtccctcccg gtggccgtgt cgctccccg 32760
ggcctgagc tgctctccg gggagatggg ggggtctggg gtgccgacaa cggccggccc 32820
ggccccgccc gagaccgagg acgcctgggg agtgggggtg ccgctttccc ccatccccag 32880
ggacaggtgg gccgcgcct ccgtcgcggc ggcgggagcc gcggccccca gccgcgcgac 32940
gtagcgacaa aagtggcggc agaggcgcac gaggcgcgcg ccgtcggccg cgtatcgcgt 33000
gtttgggggg acgagctcgt cgtaactgaa caggagcacg cgggcacagg tgcgccacgg 33060
gccccacgcc aggcgcagcg ccgcgacagt gtacgggtcg tacacgcctt ggcgctcga 33120
cgcgaccggc agggagacga acagcccgcc cgcgctgggg acgcgcggca ggaggtccgg 33180
gtgcgcggg atgacggggg ctaggatcgc ccccaccgca tccgcggca cgtaggcggc 33240
aaacgccgaa cgcaccgggg tgcagtcgcc ggtcgcgtgg gccccgggtc gggtttcgac 33300
ccggaagtcc gccgcgcct caccgtcggg gcggcgcgc acgagggcgg acagcgggac 33360
ccccgccgc gccaggcact cgctggagat gatgacgtga atcagcgagg cggggtgct 33420
cgggtcccg gtgagatcgt attggacctc gttggcaaag tgcgcgtca tggcccggcc 33480
ggcggtgca gcccttccc gtgcgggaa gggcgtggg ggggggtgcg tgtgcgcgc 33540
ctcggggccc gcgggcgcac gtgccttat acgctgtgtg tttcgtctgt ccccaggaa 33600
tccggggcca ggactttaac ctgcttttcg tcgacgaggc caactttatt cgcgggatg 33660
cggtcagac gattatggg tttctcaatc aggccaaact caagatcacc ttcgtctcgt 33720
cgaccaacac cgggaaggcc agcacgagct tttgtgacaa cctccgcggg gccgcgcgac 33780
agctgctcaa cgtggtcacc tataatgcg acgaccacat gccgcgggtg gtgacgcaca 33840
ccaacgccac ggcctgttcc tgctatatcc tgaacaaacc cgtgtttatc acgatggacg 33900
gcgcgcttc cggagcggc gatctgttcc tgcgcgactc ctteatgag gagateatcg 33960
gggggcaggc ccgcgagacc ggcgacgacc gcccgtcct aacaaagtcc gcgggggagc 34020
ggtttctgct gtaccgccc tccaccacca ccaacagcgg cctgatggcc cccgagctgt 34080
acgtgtacgt ggaccggcg ttcacggcca acacgcgcgc ctccggcacc ggcacgcgg 34140
tcgtcgggag gtaccgcgac gatttcatta tcttcgcctt ggagcacttt ttcctcccg 34200
cgctcacgg atcggcccc gcggacatcg ccgctgctgt cgtgcacagc ctgcgccagg 34260
tgctggcgt gcaccccggg cgttttcgca gcgttcgctt ggcggtcgag ggaacagca 34320
gccaggactc ggcgtggcc atcgccacac acgtgcatac cgagatgac cgcatectgg 34380
cctcggcggg ggccaacggc ccggggccc agctcctct ctatactgc gagccgccg 34440
gcggcgggt attgtacccc ttctttctgc tcaacaaaca gaagacgccc gccttegaat 34500
actttatcaa aaagttaac tccggggggc tcatggcgtc ccaggagctc gtctccgtga 34560
cgggtgcgct gcagaccgac ccggtcagat atctgtccga gcagctcaac aacctcagc 34620
aaaccgtctc tcccacacc gacgtccgca tgaactccgg aaaacgcaac ggtgccggcg 34680
acgacctcat ggtcgcggtc atcatggcca tttacctggc ggcgccgacc gggatcccc 34740

```

-continued

---

cggccttttt tccgatcaag cgcacgtctt gactctttct tgccgtttct tttgtttctc 34800  
 tttctttccc cccctctctc cgcaataaac gccttcccgg aactgtgttt cccccctac 34860  
 aacagtgttg tccgttggtt ggggtggttg ggtgcggggg tgggcggggg aagcaagaaa 34920  
 acggtcggcg aacacaacat cgggaaaaac gattcccga cgtgcgtctt ccagattcg 34980  
 acacacacac cccccctctc cttaataaaa cacaaaccac acgctcgttg gttggttaat 35040  
 gccagcgtt tatttacctc ttgttttttt tgcgtttcct ccgcggttcc ctccccaa 35100  
 cgctgcccc cgcctcaggg gtagcggata accggggcca tgcgcggga ttgcacaacg 35160  
 gcggcgccgt cgaacgtaca caccogaacc gccggggcca gggccaggat gtccccgagt 35220  
 tggcccgcgt gcgccagcca ggcgaccagc gcctcgtaaa gcggcagcct gcgttcgccc 35280  
 tctctcatca gcattggggc ttcggggttg atgagctggg cggtttctcg cgtgacgctc 35340  
 tgcattctga ggagcgcgtt cacgtatccg tctggggcgc tcagcgcgag cagccggggg 35400  
 atgagcgtga ggatgagggg ggttctctcg gttatggagt agaccatggt gaggacgagc 35460  
 gaccgcagct cgtgttttcc ggagggcagt tgcgtgagct cggccacgag cgagagacgg 35520  
 gccccgttgt aatacagcac gttgaggtcg gggagctccc cgggcgtccc ggggtcgggg 35580  
 ttgaggtccc ggatgccccg ggcgaccagc cgcgcgacta tctcgcgggc caggggcgtt 35640  
 gggagcggga ccggaaaccc cagcgtgagg tccagcgact ccaggcgcac gtccgtccc 35700  
 tggccctcga agcggggcgg gacgagctg acgggatccc cgttgacagag gtcgacgggg 35760  
 gaggtgttgc ggagattgac ggtgccggcg tgcgtgagcc ccaggtccac ggggcaggcg 35820  
 acgattcgcg tgggcagcac ccgcgtgatt accgcgggga agcgcctgcg gtacgccagc 35880  
 aacaacccca acgtgtcggg actaactcct ccggagacga acgattcgtg cgcacgtcc 35940  
 gcgagccca gctggcggcg gatggtcggc agaaagacca ctgaccctc gcaccgctgc 36000  
 agcggccgcg catcggggcg cgagataccc gaggggatcg cgatgtctgc ttcgaaacaa 36060  
 tccgtgatca tggcgcgggg ccgcgagaca ccggaacgcg ggggtgcggg agggccggaa 36120  
 agcgaacgc aaccgggacg atgatgaaac agagatgggg ggcaccgacc gttggtggaga 36180  
 gggggcgggg cagggctcag cagcagcagc ggggaggtct gtcgtgcgca ggagcccag 36240  
 gtgagaatca gtcccccgga gctcgggtct gggttttatt gggacctgcc ctcggaatcg 36300  
 cggctcccag tccaagcccc cctggggggg gcggggacag ggggtgtgtg tgggtaaaag 36360  
 caacgtcggg aatcaaac caatgcccc aacagaaaa aaaaagacgg gcgggtggag 36420  
 ggaaagctgg ggaagaagaa gccaatttta cagagacagg ccctttagcg gggaggcgtc 36480  
 gtatagaga tactgcgtaa agtgggtctc tcgcgcgtgg gcctcccctc cggggcgct 36540  
 gcgtagcagg gcggggctcg tggcgcaggt gatcgggtag gcttctgaa acaggccgca 36600  
 cgggtcttcc acgagctcgg ggcaccccgg cgggcgctta aactgcacgt cgtggcagc 36660  
 ggtggccgtg gataccgccc atcccgtttc cagatgaga cgtccaggc agcgtatggt 36720  
 ggccgtgatg tcggccgccc tgaagaactt gaagcagggg ctgaggacgg gcgaggcccc 36780  
 gttgaggtga taggccccgt gtacagcag gtccccgtac gagaaccgct gcgacgccc 36840  
 cgggttgccc gtggcccgca agggcccgcc cgggtcgtct tggccgtggt cgtacatgag 36900  
 ggctatgacg tccccctctc tgtccccgc gtacacgccc ccggccgccc gtccccgccc 36960  
 gttgcagggc cggcgaagt agttgatgct cgtggccacg ggggtggcga tgaactcaca 37020

-continued

---

caccgcatcc tgcctgtggt ccatgccggc gcgcccggc acctgggccc agccaaagac 37080  
cgggaggggc tgggcccggc ccagccggtt tcccgccacg acccggttgc gcaggtacac 37140  
ggcggcccgc ttgtctagca cggggggggc cccgcccggc aggtaaaagt tttgggggag 37200  
gttccccatg tccgtaacgg ggttgcggac ggtgccctg gcccgacggc cgtgttagcc 37260  
cacaccagg tccacgttcc cgcgcggctg ggtgagcgtg aagttgacc ccccgcccgt 37320  
ttcgtggcgg gccacctgga gctggcccag aaagtacgcc tccgacgcgc gctcggaaaa 37380  
cagcacgttc tccgtcacga agcggctcctg ccgcacgacg gtgaaccgca acccgggggtg 37440  
gaggcccgtc ttgagctggt gatacagggc cacggggctc atcttgaagt accccgccat 37500  
gagcgcgtag gtcagcgcgt tctccccgc cgcgctctcg cgggctgct gcaccacggg 37560  
ctggcggatg gaggagaagt agttggcccc cagggccggg gggaccaggg ggacgtcgcg 37620  
cgccaggtcg cgcaggggc gggggaagt gggcgcgttg gccacgtggt cggcgcgcc 37680  
aaacagcgcg tggacgggca ggacgtagaa gtattcgcca ttttgatgg tgtggtccag 37740  
gtgctggggg gccatgagca gcacgccggc gtgcagcgc ccgtcgaaga tgcgcatgtt 37800  
ggcctcgcac gcggtgttgg cgcgccgcgc gggcgcgcgc gagcacagca gcgcccctgt 37860  
gcgctcggcc atgttgtgct ccagcacctg cagcgtgagc atggcgggccc cgtcgcgac 37920  
gacgcgcccg ttgtggaaca tgcgcttgac cgtgttggcc accagattgg cgggatgcag 37980  
cgggtggggc gggctcggtca cgggatcgtc cgggcactcc tcaccggggg cgatctccgg 38040  
gaccaccatg tctgcagcg tggcgtacac gcggtcgaag cggacccccg cgtgcagca 38100  
gcgccccgcg gagaaggccg gcaccagcac gtaatagtag attttgggt ggacggtcca 38160  
gtcggcccgc cgtgcccggc ggtcgtcggc gcgctcggcc cgcgggccc ggtgtttgtg 38220  
cagcagccgg ccgtcgttgc ggttaaagtc ggccgtcgc acgttgacg ccgcccgtta 38280  
gacgggctcg tcccccccgc cgtcaatccg gcagctcctg tggcgggtcca gggcccctg 38340  
tcgcataagg ccgtcgcagt cccacacgag gggcggcagc agcgcgggt cgcgcatcag 38400  
gtgattcagc tccgctgag cctccccgc cagctccggg cccggcaggg taaagctcgc 38460  
caccagctgg gccagggct gcagctgggc caccaggtec cgatacacgg ccatgcactc 38520  
ctcggggagg tccccccga ggtaggtcac gatgtacgag accagcaggt agtcgttcc 38580  
gaacgcccgc catcgcgtgt gttccagta gctggtgatg cactgagtca cgagcccgc 38640  
cagggcgcag aacacgtgct cgttcccgtg aatcgcggct tgcagcaggt aaaacaccgc 38700  
cgggtagctg cgtcctcga acgccccgc gacggcggct atggtagccg gcgcatggc 38760  
gtggcggcca acgcccagct ccagggcccg ggcgtcacga aacgccaccg gacacagcgc 38820  
caggggcagg ttcccgttga ccacgcgcca ggtggcctgg atccccccg gaccggcccg 38880  
ggggacttcg ccgcccggaa gctcgcagtc ggccacgccc gcgaagaagt cgaacgggg 38940  
gtgcagctcc agaccaggt tggcgttgc gggctcagc aactgctcc cgtcatctg 39000  
gcactcggcg acccaccgga cccggcccgtg ggcgagggc tgcgccagg cgttcagaaa 39060  
acgtgctgc atgtcccgc cggggcccgc cggggcccgc acgtacccc cgtacggatt 39120  
cgcggcctcg acggggtcgt ggttcacgcc cccgacggcc gcgctgatgt tcatgagcga 39180  
aggatgacac acgggtcccga ccgcttctc catggacagc cgcagaaact ggtggtcctt 39240  
tccccaaaa aacagctgcc ggggagggaa cgcgcggggc tccgggtggc cggggcggg 39300

---

-continued

---

caccagggtcc cccggcgtgcg cggcggaagcg ctccatggcc gggttgaaca gccccagggg 39360  
caggacgaac gtcagggtcca tggcgcgccac cagggggtag ggcacgttgg tggcggcgta 39420  
gatgcgtctc tccagggcct ccaggaagac cagcctgtcg cctatggcca ccagatccgc 39480  
gcgcacgcgc gttgtctggg gggcgctttc gagttcctcc agcgtctccc ggttcgcctc 39540  
gagttgctcc tctgcatac ccagcaggtg gcggcccacg tcgtccaggc tccgcacggc 39600  
cttgcctacc accagcgcgc tgacgaggtt ggccccgttc aagaccatct cgcctaggt 39660  
caccggcacg tcggcctcgg tgcctccac ctccaggaag gactgcagga ggcgctggtt 39720  
gatggcggcg gtggtgacca gcaacccgcg gaccggccgc ccgcgcgtgt cggcgtgct 39780  
caggcggggc acggccacgg acggctgcgt cgcctggtc aggtccacga gccaggcctc 39840  
gatggcctcg cggcgtggtc ccgccttgcc caggaagaag ctcgtgtcgc aaaagctccg 39900  
cttcagctcg gcgaccaggg tcgcccgggc aaccctggtc gccaggcgc cgttctcgag 39960  
atatcgttgc atgggcaaca gcagggccag gggagcgcgc ttctccaaca gcacgtgcag 40020  
catctggtcg gccgtgcgc gctcaaacgc ccccaggacg gcctggacgt tgcgcgcgag 40080  
ctgctggatg gcgcgcagct ggcgatgcag gctaagccc gtcccgtcca gggcctccc 40140  
cgtgagcagg gcaatggcct cgggtggccag gctgaaggcg gcgttcaggg cccggcggtc 40200  
gatgacctc gtcagttaat tatgcacggg ctgctcgacg ggggtcgggc cgtcgcgggc 40260  
gatgaggggc tgggtgacct cgaactgcac acgcccctcg ttcagttaag ccagctcccg 40320  
gaacttggcg cacacgcacg ccacggacag gccgagctcc agaaagcgc caagcgcag 40380  
ggtgttgca taggaccca gcagggcgc aaactctacg tcatacaggc tgtttctgc 40440  
ggagcgcacg cggcgcaaaa aatcaaaag tctgcgggtg gacgccacct cgatcgtact 40500  
caggatggag ccggtgggca ggatggccgc ggcgtaccgg taaccggggg ggtcgcgggc 40560  
aggagcggcc attgggttcc ttgggggatt cgcaggctcc atcaagccga gctcgggaag 40620  
gccaaagccc tcccgcacaa cgcctcaccg ccggcggacg cgactaaca cccacgggccc 40680  
gcaaaaacc caaggggcaa cccgaccaac aacagcgcag gggaggaag gcgtaaggg 40740  
ggcgttggga ggcaaaaaga aagaaaacac ccagacgtag ccccaggac cggccggcgt 40800  
cctctgtccc cgagcaccga ctgtgccaa caggcacggg ggcgagctgc ccctgcctta 40860  
tataaccccc cgcacacccc ccgttagaac gcgacgggtg cctcaagat ggccctggtc 40920  
caaaagcgtg ctagaaaaaa gttggtaaag gcggcaaac agtccgccc cgcacccac 40980  
atggcggcgc cggccgcgca ggcgattccc agagaacggg cgcggagggg atccgtgcgg 41040  
ggcagcagct ggtggcggg gatccaatgg aaaagcccgt cgggactgaa cgtctcatgg 41100  
gcgcccgcga ccagggcgcga cagggcgcgc ccgcccata taacgcacaa cccccaaaac 41160  
acgggtggcg acaacggcag gcgatcccgt ttgatgtca cgtacaggag gagcgcctc 41220  
gccagccacg tgacatagta ggcgagacg gcggctataa tacatgcccg cgcaccgcgc 41280  
cgtccggtcc acccgttaata catgcccgcg gccaccagct ccagcggctt gaggaccagg 41340  
aacgaccaag caaacatcac caccgccttg gaaaagaccg gctgggtgtg gggcggaaga 41400  
cgcgagtagg ccgaactgac aaaaaaatca gacgtgcccgt acgaggacag cgaaaactgt 41460  
tcatcgagcg gcagttctcc gtccctcccc ccacacgcgg cctcgtctac cagctcgcga 41520  
tccaacaaag gaacatcatc ccgattgtc atggtcgggt cggggagccc gcgaggcagc 41580

-continued

---

aaaaccgaaa gtagtgctgg cggcgcgggc ccgggtccgg acccaagctt cagggatggg 41640  
 gggcggaggg caaaatcaaa caagcaccgc gcgggttcta cacacaaccc ccaccgggt 41700  
 agtatccgcg gatgagagtg cctggcgaag tcacgtccca gcaggatata aacctcggcc 41760  
 gttgggcccg gaaccccga aattcacacc cacgcctga cgcccaaatc atgggtggat 41820  
 gtggttcgcg agccgcacat ccgtgcgtcc gccctcccc gcgggtgat gacgtggcgg 41880  
 ttagtcagtg ggaaggcagg gggaaagatg ggttggggga ggaaacgaag aaaacacca 41940  
 gagggccacg tcgggaatgc gcccgaggtt gtccttaaaa ggccggccgt gcgtgacgga 42000  
 agccgtcgtt tgcccaagca ccgacgcgcg gatccacagt ggggggagtt cctccgtccg 42060  
 gccacaaccc tacgcgcggg gggcacgcgc gagagcaacc cacgggtccc gttcgcgcca 42120  
 ccgcagccc ttgctcccac caccctcctc ccaccacccc actattcccc cccccaagtc 42180  
 cgccccgttg ctgcgcggcc atggagctca gctatgccac caccctgcac caccgggacg 42240  
 ttgtgtttta cgtcacggca gacagaaacc gcgcctactt tgtgtgcggg gggtcggttt 42300  
 attccgtagg gcggcctcgg gattctcagc cgggggaaat tgccaagttt ggccctgggtg 42360  
 tccgggggac agggcccaaa gaccgcatgg tcgccaaacta cgtacgaagc gagctccgcc 42420  
 agcgcggcct gcgggacgtg cggcccgtgg gggaggacga ggtgttcctg gacagcgtgt 42480  
 gtctgctaaa ccggaacgtg agctccgagc gagacgtgat taataccaac gacgttgaag 42540  
 tgctggaaga atgcctggcc gaactactga cctcgtcgcg aaccagcccg ggggtgctgg 42600  
 tgaccggggt gcgcgtgcgc gcgcgagaca gggtcatega gctatgtgag caccgggcca 42660  
 tcgtcaacat ttctcgcgc ttcgcgtaca cccctcccc ctacgtattc gccctggccc 42720  
 aggcgcacct ccccggtc ccgagctcgc tggagcccct ggtgagcggc ctgtttgacg 42780  
 gcattcccgc cccgcgccag cccctggacg cccgcgaccg gcgcacggat gtcgtgatca 42840  
 cgggcacccc cgccccaga ccgatggcgc ggaccggggc cgggggcgcg ggggccaagc 42900  
 gggccaccgt cagcgagttc gtgcaagtga agcaccatga ccgtgttggtg tccccgagcg 42960  
 tctcttcgcg cccccgcgc agcgccccc acgcgagctc gcccccccg gggctccagg 43020  
 aggcgccccc gcggggcccc ccgctcaggg agctgtggtg ggtgtttctac gccggcgacc 43080  
 gggcgtgga ggagccccc gccgagtcgg gattgacgcg cgaggaggtc cgcgcctgtc 43140  
 atgggttccg ggagcaggcg tggaaactgt ttgggtcggg gggggctccg cgggcgtttc 43200  
 tcggggccgc gctggccctg agcccgaccc aaaagctcgc cgtctactac tatctcatcc 43260  
 accgggagcg gcgcatgtcc ccctccccg cgctcgtgcg gctcgtcggg cggtaacatc 43320  
 agcgcacagg cctgtacgtt cccgcgccc acgaaaccgac gttggccgat gccatgaacg 43380  
 ggctgttccg cgacgcgctg gcggccggga ccgtggccga gcagctcctc atgttcgacc 43440  
 tcctcccgcc caaggacgtg ccggtgggga gcgacgcgcg ggccgacagc gccgcctg 43500  
 tgcgctttgt ggactcgcaa cgctgaacc cgggggggtc cgtctcgcgc gagcacgtca 43560  
 tgtacctcgg cgcttccctg ggcgtgtgt acgcccggcca cggacgcctg gccggggcca 43620  
 cgcataccgc gcgcctgacg ggcgtgacgt ccctggtcct gaccgtgggg gacgtcgacc 43680  
 ggatgtccgc gtttgaccgc gggccggcgc gggcggctgg ccgcacgcga accgcccggg 43740  
 acctggacgc gctgctaac gtttgctcgg ctgcgcccc gcaccggccag tctgtgtgag 43800  
 atatcccaat aaagtgcagt cgttttctaa cccacggatg ccgttgtatg cctatacggg 43860

-continued

---

ggactatggg gggggaaagg aaaggaaca ggaatggaga agggaaagga acagaggcgg 43920
tagcggacgc acggcggaca caataacaaa cagaccgagg acacggaggg agtcggttg 43980
gttggcggtg gacgccgctg cgctccacaca cccgtttatt cgcgtctcca caaaaatggg 44040
acgcacgttc ggaccacct aaggatgccc gccagggcgg cggtaatcat aacgaccccc 44100
agcggcgacg cggccagaaa cccggggggc atggtggcga tgggcagcgt gtcaaaggcc 44160
agcagatgaa tcacagttcc gttggggaac aacaacaggg ccacggacgg cacgtcgctg 44220
gaaaaacagt tcggggtgcc gcgccccggc ccctgggcca gctgctgttg ggtggcatcc 44280
gtgtccacca gcagcaccga catgacctcc ccggccgggg tgtagcgag aaacacggcc 44340
cccacgaggc cgaggtcgcg ccggttttcg gtgcgcacca gccgcttcgg ctcaatctcc 44400
cgcgcgtgcc cttcgcaggt ggcggtgaga taggtgataa acacggggcg gcggacgtca 44460
acgcccgtaa gcttगततcc gatcccgggg ggcaaggggg tgggggtgac gacgtagctg 44520
gcgttgtggg tgatgggac gaggatccgg ggctccgctg tgtgcgacgg gccgctacac 44580
tgggtgggtg cctccgggga gaaggcggcg atcagggcgt tgtagtgcg ccagcgcgtg 44640
agaacggagg ccacgcggcg ggtctgtgt gccatgacgt ccgcccggat gtcggatcgg 44700
gtggccatgg ccagcgcgtc caggatgaac ccgccctcgg cgagatcga gcgcagggaa 44760
gctgcgcatg gggaaaagt gtccgggagc cagaagagg tttctgggtg gtcggctcctg 44820
gctagcggcg cccggagatc ggcgtgggct gccgcggcga cgtcggacgt acacagggcc 44880
gtggttatga ggaggcccc gcggggcgct tcccgtgct cggccgaggg cgcgcccgcc 44940
aggaacggcg cccggaggac ggccgtggcg taaaacagcg ctcggcggac catcggggcg 45000
gttagcgcgc ggccggcggg aaactcggcg tacagggcgt cgatcaggcg ggccgcgctc 45060
ggggccaccg cgcctatagg cgcggggcgt tccaacagca acgcccagctg atagcccagc 45120
gcgtgcgcca ccaggtctct ctctcgctcg aggatcggcg ccaccagatg cccgagggcg 45180
gcctccagcc gcaggcgggg ccgccgggtc aacacggaca cgttcaggaa caccgagctg 45240
gccgcgacg ccgctgctcc ccggggcgcc aggccggcca gcacgcgca gtgggccaaa 45300
aagcccagca ggtcggagag gcgaatcggc tcgtgggctg gggccgcgtt gacgaacgca 45360
aaccccagc aggcgagcag ccccgagagg cgcagaaaca gggacggacg cgcgtccgtg 45420
ccggagcccc ggtcctcccc caaaaactcc gcataggccc gcgacatata ctgggcgtag 45480
ttcgtgctct cctcggggtg gccggccacc cgcggagagg cgtccagcgc cagccggtg 45540
tcggcggcg tcggggcccc caggacaaag acgcgatacc tggggccggc cggaggcccc 45600
gggagcaccg cggggcggtt ttcgtcggtc ggatttccga cccgagcgg ggtcttgtcc 45660
gcaggcacca ctatgatctc ggccggaggg ctgtcccga tcgatatac gagccccatg 45720
aagcccttc cgtatcgcg gcgcacgagc gcggcgctgc acccgaaagc cagcccggcc 45780
gtcgtccaga cccccaggg ccacgtcgag gccgacgggg agaggtacac gtaccgacc 45840
ggagtccgta gcaggcccc ggcggccagc caggtaacgg atcggtgtg cagatcgcg 45900
atgctcaggt tcgtcgctcg atgcctcggg gtcccccggg cggccccgg gggcgcgcg 45960
ttcgtcgcc cgtccgggtg cctctcggtc gcccgctcgt ccccccggg gaacgtaagc 46020
ccctcgggt cggcgcgggc cgcgaatgt acccagggccc gggaccgaa cagcggcgag 46080
gcgccggggt tgtgcgacag tccctgagc tgggtcacct cggcgggggg acgggacgtg 46140

-continued

---

```

ggccccgect cggggagctc gggcagctc gcgttccgag gccggccgag cagataggtc 46200
tttgggatgt aaagcagctg cccggggtcc cgaggaaact cggccgtggt gaccaacacg 46260
aaacaaaagc gctcggcgta ccacgaagc atgggcacgg atgccgtagt caggttgagt 46320
tcgcccgggg gcgccaagcg tccgcgctgg gggtcgctgg cgtcgggggt gttgggcaac 46380
cacagacgcc cgtggtttgt gtcgcgccag tacgtgcggg ccaaccccag accgtgcaaa 46440
aacaccgggt cgatttgcct cgtccagtag gtgtcatggc ccccggaac gccaccagg 46500
acccccatca cccccacag accggggccc atggtcgtcg tcccggctgc cagtccgag 46560
atgggggggg gtgtccgtag ccacggccca aagaggctcc gcacctcgga ggctatcgga 46620
ggccctttgt tgcgtaagc gcgggcaaaa ggatggggtg gggtaggggt aaaagcacia 46680
aggagtagc agaccgaaaa caaggacgga tcggcccctc cgtttttcgt gtaggggtgct 46740
gatacggtag cagccctggc cccgaacccc cgcgcttatg gacacaccac acgacaacaa 46800
tgccttttat tctgttcttt tattgcgctc atcgcgggga ggcttccgt tcgggcttcc 46860
gtgtttgaac taaactcccc ccacctcgcg ggcaaacgtg cgcgccagggt cgcgtatctc 46920
ggcgatggac ccggcggttg tgacgcgggt tgggatcatc ccggcggtga ggcgcaacag 46980
ggcgtctcga caccgcagcg gcgactgac gttaatccagg acaaatagat gcatcggaag 47040
gaggcggtcg gccaaagcgt ccaagaccca ggcaaaaatg tggtaacaagt ccccgttggg 47100
ggccagcagc tcgggaacgc ggaacagggc aaacagcgtg tcctcgatgc ggggcagaga 47160
ccccgcgccc tcctcggggg cggggcgcgg ggtcgcgccc gcgacccccg tcagccggcc 47220
ccagtcctcc cgcacacctc cgcgcgctg caggtaccgc accgtgttgg cgagtagatc 47280
gtagacacgg cgaatggcgg acagcatggc caggtaaacg cgtcgcgccg ggcgttggcg 47340
tctggccagg cgttcggcgt gtccggcctc cggaaggaca cccaggacca ggtcgtgccc 47400
gggcgcggtc gggggcatga gggccacgaa cgccaacacg gcctgggggg tcatgcttcc 47460
catgaggtac cgcgcggccc ggtagcacag cagggaggcg atagggtgcc ggtcgaaaac 47520
aaggtgagg gccggggggc gggcttgcgg gccacagcc tccccccga tatgaggagc 47580
caaaaaggcg tcctcgcgca cataaggcgt gctcattgtt atctgggcgc tggcattac 47640
caccgcgcc tccccggccg atatctcgcc gcggtccaga cggtagctgc tgtttagat 47700
gttcgtcagg gtcctggagg ccccagcac ctgccagtaa gtcacggct cggggacgta 47760
gacgatattg tcgcccggcc ccagggcctc catcagctgc cggaggtgg tggcttccc 47820
cacccttg ggctcgtcta tataaacccc cagcagcgtg ggcagctccg gatccccgcg 47880
ggctcggag gcccccggc gatggctagg acgggacgcc gcgcgccgt cggtaggccc 47940
gctcgcaaga gcagcctgac cgaacgcagg cgcgtgctgt tggccggcgt gagaagccat 48000
accgcttct acaaggcgtt cgcocgagag gtgcgggagt tcaacgccac caggatttgt 48060
ggaacgctgc tgacgctgat gagcgggtcg ctgcagggtc gctcgtggt cgaggccacg 48120
cgcgtcact taatatgca agtggaacctc gggccgcgcc gccacagctg catctgcgtg 48180
ttcgaattcg ccaatgacaa aacgttggga ggtgtgtgcg tcatcctgga gctaaagaca 48240
tgcaaatcga tttcttccgg ggacacggcc agcaaacgcg aacagcggac cacgggcatg 48300
aagcagctgc gccactcccc gaagctgctg cagtcgctcg cgcctccggg ggacaaggct 48360
gtctacctgt gtctatttt ggtgtttgct gcgcagcgtg cgtcgcgct cagcccgctg 48420

```

-continued

---

```

accggctcg tcccgaaaa gatctccggc aacatcaccg cggccgtgcg gatgctccaa 48480
agcctgtcca cgtatgccgt gccgcggaa ccgcagaccg ggcggtcgcg gcgccgggtc 48540
gccgcagacc ccagaccgca aaggccccc tcccgcacac gtgacccgga aggcacggcg 48600
ggtcacccgg cccaccaga gagcgcaccg cctcccccag gggctgtagg cgtcgtgcg 48660
gagggggggg gtgtgcttca gaaaatcgcg gcgctttttt gcgtgccggg gcccgccaag 48720
agcagacccc ggacaaaaac cgagttaggt tctgtgtgtt gttttttttt ttttttttcc 48780
tcgttttgtt ttctcttctt tccccccccc ctcccccgct tctggccaag cactctcacc 48840
tgcttaagcg gaaccgcggg gcgcgcgggg actcatttgt cgcggcgac acccaccgca 48900
caacagcccc tgggtgtgca ccgctgtcgc ccccgctctgt cgcctctccc ttttttcccc 48960
ccctcaaaga acgtggtgtt gggcgcggcg caattcttcc cggagcgcgg tcgtcgcccg 49020
cccgcggccc tcgaacatgg acccgtaact ccttttcgac gcgctggacg tttgggaaca 49080
caggcgttc atcgtcgcg actccaggag ctctcacc cccgagttcc cccgggactt 49140
ctggatgtg cccgtgttca acatcccccg ggagacggcg gcggagcggg cggcagtgtc 49200
gcaggcccag cgcaccgcgg ccgcggcgcc cctggagaac gcccccctcc aggcgccgca 49260
gctgcccgtc gacatcgagc gccggatagc cccgatcgag cagcaggtgc atcacatcgc 49320
cgacgccctg gaggcgctgg agaccgcggc ggccgcggcc gaagaggcgg atgccgcgcg 49380
ggacgccgag gcgagggggg agggcgcgtc ggacggggca gcgccgtcgc ccaccgcggg 49440
ccccgccccc gcggagatgg aggttcagat cgtacgcaac gaccgcgcgc tacgatacga 49500
taccaacctc cccgtggatc tgetacacat ggtgtacgcg ggccgcgggg ccgcgggttc 49560
gtcgggagtc gtctttggta cctggtagcc cagatccag gaacgcacca tcgcccactt 49620
ccccctgacc acccgcagcg ccgacttctg agacgggcgc atgtccaaga ccttcatgac 49680
cgcctggtc ctgtctctgc agtcgtcgcg ccggctgtac gtgggccagc gccactatc 49740
cgccttcgag tgcgcgctgc tgtgtctgta tctgctgtac cgaaccacc accgagctctc 49800
ccccgatcgc gatcgcgctc ccgttgcgtt cggggacctg ctggcccgcc tcgccgctc 49860
cctggcgcgt ctggccgcgg taatcggcga cgagagcggc cgcgcagat accgctaccg 49920
cgacgacaag ctgcccgaag cgcagttcgc ggccggccgc ggccgctacg agcacggggc 49980
cctggccacc cagctcgtga tcgccacggt ggtgcgccac ggggtgctac cggcggcccc 50040
ggcgacgctt ccccgagaca ccagcaccgc cgtgaacccc gacgacgtgg cccaccgcga 50100
cgacgtcaac cgcgcgcggc ccgcgttttt ggcacgcggc cacaacctct tcctgtggga 50160
ggaccagacg ctgctcgggg cgaccgcca caccattacg gccctggcgg tgettggcg 50220
gctcctcgcg aacggcaacg tgtagcggga ccgcctcgc aaccgcctgc agctgggcat 50280
gctgatcccg ggagccgctc cggcggaggc catcctcgg ggggctcgg gattggactc 50340
gggcgcata aaaagcggcg acaacaacct ggaggcgtg tgcgttaact atgtacttcc 50400
gctgtatcag gcagacccca cggctcagct gaccagttg tttccggggc tggcccct 50460
gtgcctggac gccagggcg ggcggcact ggcgtcgcg aggcgcgtgg tggatatgtc 50520
gtcggggccc cgcagggcg cgcctcgcg cctcaccgcg ctggagctca tcaaccgcac 50580
ccgcacaaac accaccctg tgggggagat tattaacgc cagatgctc tggggataca 50640
atcgaacag gggcctggg tgcctgcccc gcaggcacgc atcggcttg cgtcaaac 50700

```

---

-continued

---

caagcgattc gccacgttca acgtgggcag cgactacgac ctgttgact ttttgtgtct 50760  
cgggttcatt ccccagtaac tgtccgtggc ctagggaagg gtgggggtgg tgggtgtggg 50820  
gtgtttttct gttgttgttt ctggtccgcc tggtcacaaa aggcacggcg ccccgaaacg 50880  
cgggcttttag tcccggcccg gacgtcggcg gacacgcaac aacggcgggc cccgtgggtg 50940  
ggtaagtgg ttccgggggca tccgtgtatt cccttgcccg ctccacccc cccccccctt 51000  
cccgttttgt ttgtttgtgc gggtgcccat ggcgtcggcg gaaatgcgcg agcggttgga 51060  
ggcgectctg cccgacgggg cggtgcccat ctacgtggcc gggtttttgg cctgtacga 51120  
cagcggggac ccgggcgagc tggccctgga cccagacacg gtgcgtgagg cctgcctcc 51180  
ggagaacccc ctgccgatca acgtagacca ccgcgctcgg tgcgaggtgg gccgggtgct 51240  
cgccgtggtc aacgaccctc gggggccggt ttttgtgggg ctgatcgcgt gcgtgcagct 51300  
ggagcgcgtc ctgcgagcgg ccgcagcgcg cgtatTTTT gacgcgcgcg gaccgcgct 51360  
ctcccgggag gagcgtctgc tgaactgat caccaactac ctgccatcgg tctcgtgtc 51420  
cacaaaacgc cgggggggag aggttccgcc cgaccgcacc ctgtttgcgc acgtggccct 51480  
gtgcgccatc gggcgcgcc ctggaacct cgtcacctac gacaccagcc tagacgggc 51540  
catcgtccg tttcgcacc tggaccggc gacgcgcgag ggggtgcgac gcgaggccgc 51600  
cgaggccgag ctccgctcgg ccggggcgac ctgggcccc ggcggtggagg cgtcacaca 51660  
cacgtcctc tccaccgccc tcaacaacat gatgctcgt gaccgctgga gccttgtggc 51720  
cgagcggcgg cggcaggccc ggatgcggc acacacgtac cttcaggcga gcgaaaaatt 51780  
taaaatatgg ggggcccagt ctgcccctgc gccggagcgc gggataaaa ccggcgcgcc 51840  
gggtgccatg gacacatccc ccgcgcgag cgttcccgc ccgcaggtcg ccgtccgtgc 51900  
gcgtcaagtc gcgtcgtctt cttcttctc tcttttccg gcaccggccc atatgaacce 51960  
cgttccgca tccggcggcc cggccccccc gccgcccggc gacgggagtt atttgtggat 52020  
ccccgcctct cattacaatc agctcgtcac cgggcaatcc gcgccccgc acccgccgct 52080  
gaccgcgtgc ggcctcggc ccgcggggac ggtggcctac ggacacccc gcgcccggcc 52140  
gtccccgac taccgcctc ctcccgcaca cccgtaccgc ggtatgctgt tccggggccc 52200  
cagtcccctg gagcccaga tccgcgcgct ggtgggggccc atcgcgcgcg accgccaggc 52260  
gggtgggctt ccggcggccc ccggagacca cgggatccgg gggtcggcga agcgcgcgcg 52320  
acacgaggtg gagcagccc agtacgactg cggcctgac gagccggacc gggacttccc 52380  
gtattacccg ggcgagggcc gccccgagcc gcgcccggtc gactcccggc gcgcccgcg 52440  
ccaggcttcc gggccccac aaaccatcac ggcgctggtg ggggagggtga cgtccctgca 52500  
gcaggaaactg gcgcacatgc gcgcgcgtac ccacgcccc tacgggcccgt atccgcgggt 52560  
ggggccctac caccacccc acgcagacac ggagacccc gcccaaccac cccgtacccc 52620  
cgccaaggcc gtctatctgc cgcgcgcgca catcgccccc ccggggcctc ctctatccgg 52680  
ggcggtcccc ccaccctcgt atccccaggt tgcggttacc cccggteccg ctccccgct 52740  
acatcagccc tccccgcac acccccacc ccctccgcg ccgcggggac ccacgcctcc 52800  
ccccgcgcg agcttaccac aaccgaggc gcccggcgcg gagccggcg ccttagttaa 52860  
cgccagcagc gcggcccacg tgaacgtgga cacggcccgg gccgcgcatc tgtttgtgtc 52920  
acagatgatg gggtcgccgt aactgcctc caggatccgg acttgggggg ggtgtgtgtt 52980

-continued

---

```

ttcatatatt ttaataaac aaacaaccgg acaaaagtat acccacttcg tgtgcttg 53040
tttttgtttg agaggggggg gtggagtggg ggggaaagtg ggccgaatga cacaaaaatt 53100
aggtcggagg ggtgaggggg gggggctagg agccgaaccg atggcccca cacgcgacgg 53160
aaggcccgga agactaccac ggggaggggg tgtgaaagc gaccggtcgc agggagacgg 53220
ggttggtttg gggttggttt ggggttggtt ttcccgttag cacatgtctg catttgtttt 53280
tctagtacac ccccccccc cccccaaata aaaaccaagg caaaacaata ccagaagtca 53340
tgtgtatttt tgaacatcgg tgtcttttta ttatacaca agcccagctc cctcccctc 53400
ccttagagct cgtcttcgct tccggcctcg tccctggtt ggagcggaga gtacctggct 53460
ttgttgctgc tgcgcagaac catgttggtg accttgagc tgagcagggc gctcgtgccc 53520
ttctttctgg ccttggttcc cgtgcgctcc atggccgaca ccaaagccat atacgggatc 53580
atctctcggg cctcggccaa cttggcctcg tcaaacccgc cccctcgcg gcttctctc 53640
cctcccgcg ccacgcccc ggggtcggaa gtcttgagtt ccttggtggt gagcggatac 53700
agggccttca tgggatttgg ttgcagtgc aggacgtagc ggaaggcga gaaggccgcg 53760
accaggcccg ccaggaccag cagccccacg gcaagcgcgc cgaaggggtt ggacataaag 53820
gaggacacgc ccgagacggc cgacaccacg cccccacta ctcccatgac taccttgccg 53880
accgcgccc ccaagtcccc catcccctcg aagaacgcgc acagccccgc gaacatggcg 53940
gcgttgccgt cggcgcggat gaccgtgctc atgtcggcaa agcgcaggtc gtgcagctgg 54000
ttgcggcctc ggacctcctg ttagtccagc aggcctctgt ccttgatctc gtggcgcgtg 54060
tagacctcca ggggcacaaa ctctggttcc tccagcatgg tgatgttcag gtcgatgaag 54120
gtgctgacgg tggtagcctc ggcgcgactc agctggtgag agtacgcgta ctctcgaag 54180
tacacgtagc cccgcgcgaa gatgaagtag cgcgggtggc ccacggtgca cgctcagac 54240
gcgtcgcggg taggcgcagc ctctgttctc tgcgccagct gccctcgcg cagcgggccc 54300
tggctctcgt accgaaagt gaccaggggg cggctgtagc acgtccccgg ccgcgagctg 54360
acgcgcacgc agttctgcac gatcacgttg tccggggcga cgggcaegca cgtggagacg 54420
gccatgacgt ctccgagcgt gcgcgcgctc acccgcggc cgacgggtggc ggaggcagtg 54480
gcgttggggg tagacttgcg ggctcgttc cagagagtca gctcgtggtt ctgcagctcg 54540
caccacgcga cggcgatgcg cccagcatg tcattcacgt ggcgctgtat gtggttatac 54600
gtaaactgca gccggcgcaa ctcgatcgag gagtgggtct tgatgcgctc cacggacgcg 54660
ttggcgttgg ggcctcccg cagtggcgcg ggcgtggcat tccggggctt ggcgtcctgc 54720
tcccgcagct actcccgcac gtacagctcg gcgagcgtgt tgctgaggag gggctggtag 54780
gcgatgagga agccccctg ggcaggttag tactgcggct ggcccacctt gatgtgcgtg 54840
gcgttgact tgcgcgcaaa catgcggtcg atggcctcgc gggcatcccg gccgatgcag 54900
tcgcccaggt cgacgcgcga gagcagtagc tcggtcaggt tgggtggtgaa ggtggtcgag 54960
atggcgtcgg aggagaagcg gaaggagccc ccgtactcgg cgcggagcat ctctccacc 55020
tccctgccact tggatcatgt gcagaccgcc ggtcgtctcg gcaccagtc ccaggccacg 55080
gtaaacttgg ggtctctcag caagttcggc gtcgtcggcg acgtggcccg ggcctcctg 55140
gtgaggtcgc gcgcgtagaa cccgtcgacc tgcttgaagc ggtcggcggc gtagctggtg 55200
tgctcgtgtg gcgaccctc ccggtagcgc taaaacgggg acatgtacac aaagtcgccc 55260

```

-continued

---

gtcgccagca caaactcatt gtaacgggtac accgaccgag cgtccacctc ctgcacgatg 55320  
 cagttgaccg tcgtgccgta ccgatggaac gcctccaccc gcgaggggtt gtacttgagg 55380  
 tcggtggtgt gccaccccc gctcgtgcgc gtggcgacct tcgcgggctt gagctccatg 55440  
 tcggtctcgt ggtcgtccc gtgaaacgag gtggtctcca tgttggtccg cagctacttg 55500  
 gccgtggagc ggcagacccc cttggtgta atcttgctga tcacctctc gaagggaacg 55560  
 ggggcgaggc cctcgaatat cccataaac tgggagtagc ggtggccgaa ccacacctgc 55620  
 gacacggta cgtctttgta gtacatgggt gccttgaatt tgtacggggc gatgttctcc 55680  
 ttgaagacca ccgcatgccc ctccgtgtag ttctgcccct ccgggcgagc cgggcagcgg 55740  
 cgcggctgct caaactgcac caccgtggcg ccgctcgggg gcgggcacac gtaaaactgg 55800  
 gcacggcgtc tctcgcacct gatttccgc aggtgcgcgc gcagcgtggc gtggccggcg 55860  
 gcgacggctc cgttgccgct gggggggcgg gtgcgctcgg gccgcttggg cggctttttg 55920  
 gtttccgct tccgggccc ggtggtcgcg gggctcggga cggggggcgg ccgggagcgg 55980  
 ggacccccgt tcgcccgcac ggtcgcggcc acccgcggcc ccacgaccag cgcgcaaatc 56040  
 gcccccggg ccgcccgcgc caccgcggcc accagcggcc ccacgaccag cgcgcaaatc 56100  
 aagccccccc cgcgcatggc gggcctcagg gggcgcgctc ctcccgcgc ccgctagtct 56160  
 gggggcgagg tgcgtagga ccgagtagag gatggaaaa acgtctcggc cgtaaaccac 56220  
 gaccgagcgg ggtccgatgc agccgtcggg gccgctctcg acgatggcca ccagcggaca 56280  
 gtcggagttg tacgtgaggt acacgcccgg cgggtagcgg tacagacctt cggaggtcgg 56340  
 gcggtgcag tcggggcggc gcaactcaag ctcccgcac cggtagaccg acgcaagag 56400  
 tgtggtggcg ataagtagct cgcgaatata tcgccaggcg gcgctcggg tgggctgat 56460  
 tccggaaca ccgtcaaac agtagaactt ttgaaactc ctgacggccc aatcagcgc 56520  
 cgaaccccc cgcgccatga tgaagcgggc gaggctctcc ttgaggtcgc gcaggagccc 56580  
 cacgttctcg acgctgtag acagcgggt gttggggggc tgggcgaagc tgtgggtgga 56640  
 gtggtcgaac agggggccc tgaagcgcgc gaagaagcga tgggtgatgc tggggagcag 56700  
 ggcgggttc acctggtggc gcagcagcga cgtcgcagc aaccggtcgc cgtcaaacac 56760  
 gcccccggc gcgcccgtgt cgatgacct gcccgccc gccgtcaggc cgcagaagcg 56820  
 cgcgcccgc gcgaagcctt tggcagacc gccgaaggc gcgggcagca cctcgcctg 56880  
 gacgtgacc cgcagatct tctcgcgct cccgcgctc tcgcccagc agcgcggcag 56940  
 gctggccagc gaaccgttgg tcaggcggtc cgcgtacagc cgcggcggct cccgcagctc 57000  
 cgcggcggc cgcgtcggga tgcgcccc gctctcggc ccctcggccc ctggctcggg 57060  
 gccgcgctc ccgtctcgc tcgcccggc ccccgccc cgcctcggc cccctcctc 57120  
 cgcggcggc cggggtctt cctcctcggc ccccgggc gcgcccggc ccccagcgg 57180  
 cgcagcagc cggcgcagc cctcctcgc gactgctc gggctgacga gccgcccag 57240  
 cagcggcgtc gtcagggtgt ggtcgtagca cgcgctatc agcgcctcga tctgatcgtc 57300  
 gggcagcgc cctggcccgc cgtgatcag ggcgtccacc atgtccagc ccgcccagg 57360  
 cccccgaac gcgcatgca agtgctcgc ccgcccggc aacagcgcga gctccagcgc 57420  
 caccgcccgc gctcctgct gcagctcgc ctgcgccc gcgttcagg tctcggcga 57480  
 ggcgtccatg gtcagggtgg gggcgcgac gccgagccc agccagaagc gcagctcgt 57540

-continued

---

gatggcgtagc aggccggggcg tagtggcctg aaacacgtca tgcgcctcca gcagggcgctc 57600  
ggcctcctcg cggacagaag agctatcggc gggcgggggg ccggccctgg ccccgccgcc 57660  
cgcccgggtc cgcgccagcg cctggctccag cacacagagc gctcgcgcgc gggcgggcgctc 57720  
cgacagcccg gcggcgtagg gcaggtaccg tcgcagctcg ttggcgtcca gccgcacctg 57780  
ggcctgttgg gtgacgtggg tacagatgcg gtccgccagg cggcgggcga ttgtcgcccc 57840  
ttggttcgcg gtgacgcaca gctcctcgaa acagaccgcg cacgggtggg acgggtcgct 57900  
cagctccggg ggcacgatga gggccgaccc caccgcccgc accataaact cccggagcgcg 57960  
ctccagcgcg gccgtggcgc cgctcggggg ggtgatgagg tggcagtagt tcagctgctt 58020  
gagaaaattc tcgacatcat gcaggaagca cagctccatg cggacgtccc cgcgctacgt 58080  
ctgcagcccg atctgtggtt ggtacggaca gggtcggggc agaccatgg tctcggtgaa 58140  
aaaggcagag acgtcaccgc ttggtcgcgaa cgtttccagg tggcccagga gccgctcccc 58200  
ctcgcgccac gcgtactcca ggagcaactc cagggtgacc gacagcgggg tgagaaaggc 58260  
ggcggcctga gctccagcc cggccgcgag gtgcccccgc agcacgcgca cctggagcgc 58320  
gttgagtttt agctggggca gcttccccag gccgatctgg gggtcgcatc gtcgaagcag 58380  
ctctagctga aaaacgtacg tctgtacctg cccgagcagg gccaacagtt tctgtcgggc 58440  
cgcagtgggc tcgaaaaccg cggccggggg cgcggccgcc atggcgagtc acccgccct 58500  
gctgtggttt agttaaggtt tggggggggg tgggtcagag gcgcgccccg cgcggactga 58560  
tgcggcgcg gcgccctgac atcccctctt tatgccctc gcccgcccgc ccgccccgc 58620  
gggtgtccgt gattcgcgga gtcggggcct tgtgtttctt tctttcccc ccgaatccgt 58680  
tctttcttcc tcaccccccc ctccccacac acccaccag gactcgccac cacaaggagg 58740  
cgagagcccc tcgctaacc aaagacacag tcacgagaca cgatatcgac tgtagtgcg 58800  
atcgtttatt ttatacacia caccaacctt tccttcgacc cccccaccc ccgcccctag 58860  
agcatatcca acgtcaggtc ctttttctcc ggtggctcct ccccaaacgg atcgtcgccg 58920  
tgaaacgccc gctttcgggc gacccggcc gcccccgcg ccgcgcgcaa accgcccgaac 58980  
gacgcccgtt ggtcatcctc gtcgcccgaa tccccaaagt taaacacctc cccggcgggc 59040  
ccgagctggc tgaccagggc ctccgcctcg tgggccacct ccagggcgcg gtcggtcgac 59100  
cactcgccat gcccgcgctc cagggcgcg gtggtaaaact ccatcatttc ctgctcagg 59160  
tactcgctct ccagcagcgc cagccagtc tcgatctgca gctgctgggt gcggggggcc 59220  
aggetcttga cggtcgccc aaacacgctg ctggcgaccg ccgccccgc ctcgcaatg 59280  
atgccccgga gctgctcgca cagcgaatgc tcgtggggcc cgcgcccgag actcgacgcc 59340  
gcgcacacia acccgccctt gggcgaggcc aggacaaact tgcgggtgcg gtcгааagatc 59400  
agcagcgggc acgctttttt gccgcccagc aggctggccc agttcccgc ctgaaacacg 59460  
cggctcgttc cggccatgcc gtagtatttg ctgatgctga ggcccagcac gaccatcggg 59520  
cgcgcgcca tcacggggcg cagcaggtt cagctcgcga acatggacgt ccagggcgcg 59580  
gggtgcgct cgagggagtc catcagcgc cggggcccgc cctccaggcc cgcgcccgc 59640  
tgcggggccc agcggcgcc gcctgcacg ctggggggac ggcgggaccc ggcgatgacg 59700  
gccgtgaggg tgtttatgaa gtacgtcgag ttgctgcagt acctcaagat ctggttggcc 59760  
atgtagtaca tggccagttc gctcacgtta ttgggggcca ggttgataaa gttaatcgcg 59820

-continued

---

ccgtagtcca gggagaacct cttaatgaac gcgatggtct ctatgtcctc ggcgacaag 59880  
 agccgggctg ggagctggtt gcgctggagg gcggtccaga accactgcgg gttcggctgg 59940  
 ttcgaccccg ggggcttgcc gttgggaaa atgaccgcgt ggaactgctt cagcaggaag 60000  
 cccagcggtc cgaggaggat gtccacgcgc ttgtcgggct tctggtagge gctctggagg 60060  
 ctggcgacc gcgccttggc ggctcggac gcgttggcgc tcgccccgc gaacaacacg 60120  
 cggctcttga cgcgagttc cttgggaaa ccaaggtca cgcgggcaac gtcgcctcg 60180  
 aagctgctct cggcgggggc cgtctggccg gccgttaggc tgggggcgca gatagccgc 60240  
 ccctccgaga ggcgaccgt cagcgtcttc gccgacagga acccgttgtt gaacaggctc 60300  
 atgacgcgcc gccgcagcac cggttggaat tgattcggaa agttgcgcc ctcgaccgac 60360  
 tgcccgcgca acacccctg cactggctc agggccaggt cctggtacac ggcgaggttg 60420  
 gaccgccgcg cgaggagctg cagcaggggg cacggcccgc aggtgtacgg gtcagcgcac 60480  
 agcgacatgg cgtggttggc ctcggccaga ccgtcgcgga acttaaagtt gcgccctcg 60540  
 atcaggttgc gcatcagctg ttccacctcg cgatccacca gctgcttgat gttgttcacc 60600  
 accgtgtgca gggcctcgcg gttgccgata atcgtctcca gccctcccag ggcgtgggc 60660  
 accgcctggt ccactgactg cagggcctcg agctcggcca tgaacgcctc ggtggccgcg 60720  
 cggtaactct cctgcatgat ggtccgggtg ttctcggacc cgtccgcgcg cttcagggcc 60780  
 gagaaggcgg cgtagtccc cagcacgtcg cagtcgctgt acgcgctgtt catcgttccg 60840  
 aagaccccaa tggcccccg ggcggcgtc gcgaacttgg ggtggcgggc ccgagccgc 60900  
 atcagcgtcg tgtgcgcgca ggcgtggcgg gtctcgaagg tacacaggtt gcagggcacg 60960  
 tcggtctggc ccgagtcgcg gacgtagcga aacacgtcca tctcctggcg ccgacgatg 61020  
 actccgcgt cgcagcgtc caggtaaaa agcatcttgg ccagcagggc cggagagAAC 61080  
 ccgacagca tggccagggt ctcgccggcg aactcctggg ttccgccgac gaggggcgcc 61140  
 gtgggggcc cctcgtacc gggcaccacg tggcctcgc ggtccagctg cgggttggcc 61200  
 gccacgtgcg tgcgggac gagaaaag cggtaaaagg agggcttctt gtggtccttg 61260  
 ggtccgcgc gcccgcgctc gtccacctg gtcaggtgga gggccgaatt ggtgctgAAC 61320  
 accatggcgc ccacgagggc cgcggcgcgc gccaggtacg ccccgacggc gccggcgcgg 61380  
 gcccgggcg tttcctggcc ctcaagcagg ggccaagtgg tgatgtcggg ggcggctcg 61440  
 tcaaaagacc ccacgacac gatggactcc agggccaggg cggcgtcgc ccgccatcacc 61500  
 gaggccaggc gctgctcaaa cccgccgcg gggccttgt tcccgcgctc gcgcgcgcc 61560  
 cgctggggct taccctggct ggcctcgaag gccgtgaacg taatgtcggc ggggagggcc 61620  
 gcgcctcgt ggtttctcgt gaacgccagg tggcgggcgc cgcgggccac ggcgtccacg 61680  
 ttccggcgc gcagggccg ggcggcgggc ccgacgacc cctcgaacag caggcggcg 61740  
 agggggcggg tgaaaacgg aaggggtag ttgaaattct ccccgatcga tcggtggttg 61800  
 cagttaaacg gatcggcgat gaccggcta aaatccggca taaacatctg cagcggatac 61860  
 acggggatgc ggtgaacct cgcgtcccc atggttacct tgtccatccc gccagatgc 61920  
 aggaaggtgt tctgatgca cacggcctcc cggaagccct ccgtgatcac cagatacagc 61980  
 aaggccgggt ccgggtccag tccgagccgc tcgcacagcg cgtccccgt cgtctcgtgc 62040  
 tttaggtcgc agggccgggg cgcgtagtcc gcgaagccaa aatgcggcg cgcgcctcg 62100

-continued

---

cagagccgcg tcaggttggg ggccctgggtg ctggggggcca ggtggcggcc gccgtgaaag 62160  
acgtaaaacgg acgggctgta gtgocgagggc ataagcttga gggacaccgc ggtcccccca 62220  
aggcccgctcg tgcgggaccc gacgaccgcg gccacgttgg cctcaaaccc gctctccacg 62280  
gtcaggccga cgatgagggg cgcgacggcg acgtccgctg cgccgctgcg cgccgacagt 62340  
agcgacagca gctccaggcc ttcggccgga caggcgcggc catacacgta ccccatcggc 62400  
cccggaggaa ccttgacggt ggtcgctggt ttgggcttgg tgtccatggc ttccgggaga 62460  
tcggcgaccg gcaggaacgg gggcccggca agacgaccgg gggcagacgg gggaggccgc 62520  
gcgtggtcga cggctgctcg ccgctcgctg ctctccgatg gggtcgaatg ccggcgctgg 62580  
gggtggggtc tacaccgcc cgcttcgcca gcggcccctg gtgggggtgg gatgggtggg 62640  
atgggggtgg cgagaatgg ccgccaccgg atcgcccgcg acgggggggg ccgggggttg 62700  
gcaaggtttg ggcgcaaggc tccagcggcg attcgagagg cctcgggatg gcggcccaga 62760  
gctgggtatg ctccggccgg gcggccgcta tatgtacggc gtgctgggag gggcgcgctc 62820  
gggccccgcc caggtccgc caccgccgc gcgtcatcgg cagggggcgt ggccgccctt 62880  
ctaaaaaag tgagaacgag aagcgttcgc actttgtcct aataatata atactattag 62940  
gacaaagtgc gaacgcttcg cgttctcact ttttttagaa gggcggccac gccccccttg 63000  
acgtcacgct caccggggcg gccggccgcc cataagcggc gcctgcccgg ccgataaaaa 63060  
gaaaccgagg cgcccccgcg gacaccacac actggtcttc gaaccccgga cgcgcagaag 63120  
ggaccccggc gcgggtccgc cggtaagagc cggggggaac atcgccaccg ccatcccacc 63180  
ccgagctggt ggggtggcgg gtgggggggc tggtagggcg gtggtgggag gggcgcgctg 63240  
atagcaggac aacgaccggc ggcgatgttt tgtgcgcggc gcggcccggc ttccccggg 63300  
gggaagccgg cggtccggcg ggcgtctggg ttttttgccc cccacaaccc ccggggagcc 63360  
accagacggc caccgcgcgc ttgcccggcg cagaacttct acaaccccca cctcgctcag 63420  
accggaacgc agccaaaggc cctcgggccc gctcagccgc atacgtacta cagcgagtgc 63480  
gacgaatttc gatttatcgc ccccgcttcg ctggacgagg acgccccgc ggagcagcgc 63540  
accgggttc acgacggcgc cctccggcgc gccctaagg tgtaactgcg gggggacgag 63600  
cgcgacgtcc tccgctggg cccggagggc ttctggccc gtcgcttgcg cctgtggggc 63660  
ggtgaggacc atgccccga ggggttcgac cccaccgca ccgtcttcca cgtgtacgac 63720  
atcctggagc acgtggaaca cgcgtacagc atgcgcgcgc cccagctcca cgagcgattt 63780  
atggacgcca tcacgcccgc cgggaccgct atcacgcttc tgggtctgac cccgaaaggc 63840  
catcgctcg ccgttcacgt ctacggcacg cggcagtact tttacatgaa caaggcggag 63900  
gtggatcggc acctgcagt ccgtgcccc cgcgatctct gcgagcgctt ggcggcggcc 63960  
ctgcgcgagt cgcggggggc gtcgctccgc ggcattccg cggaccactt cgaggcggag 64020  
gtggtggagc gcgcccagct gtactattac gaaacgcgcc cgaccctgta ctaccgctc 64080  
ttcgtgcgaa gcgggcgcgc gctggcctac ctgtgcgaca acttttgccc cgcgatcagg 64140  
aagtacgagg gggcgctcga gcgccaccac cggtttatcc tggacaaccc ggggtttgtc 64200  
acctcggct ggtaccgct caagcccggc cgcgggaaac gcgccgcca accgcgcccc 64260  
ccgacggcgt tcggaacctc gagcgcgctc gagtttaact gcacggcggg caacctggcc 64320  
gtcagggggg ccatgtgtga cctgccggcc tacaagctca tgtgcttoga tatcgaatgc 64380

-continued

---

```

aaggccgggg gggaggacga gctggccttt ccggtcgcgg aacgcccga agacctcgtc 64440
atccagatct cctgtctgct ctacgacctg tccaccaccg ccctcgagca catcctcctg 64500
tttctcctcg gatcctgcga cctcccagag tcccacctca gcgatctcgc ctccaggggc 64560
ctgcccggccc ccgtcgtcct ggagtttgac agcgaattcg agatgctgct ggccttcctg 64620
accttcgtca agcagtagcg ccccaggttc gtgaccgggt acaacatcat caacttcgac 64680
tggcccttcg tctgaccaa gctgacggag atctacaagg tcccgcctga cgggtacggg 64740
cgcataaacg gccgggggtg gttccgcgtg tgggacatcg gccagagcca ctttcagaag 64800
cgcagcaaga tcaaggtgaa cgggatggg aacatcgaca tgtacggcat catcaccgac 64860
aagggtcaaac tctccagcta caagctgaac gccgtcgcgg aggcctcctt gaaggacaag 64920
aagaaggatc tgagctaccg cgacatcccc gcctactacg cctccgggccc cgcgcagcgc 64980
gggggtgatcg gcgagtattg tgtgcaggac tcgctgctgg tcgggcagct gttcttcaag 65040
tttctgcgcg acctggagct ttccgcgctc gcgcgcctgg cgggatcaa catcaccgccc 65100
accatctacg accggccagca gatccgcgtc ttcaegtgcc tcctgcgcct tcggggccag 65160
aagggcttca tctgcgcgga caccagggg cggtttcggg gcctcgacaa ggaggcggccc 65220
aagcgcggcg ccgtgcctcg gggggaaggg gagcggccgg gggacgggaa cggggacgag 65280
gataaggacg acgacagagga cggggacgag gacggggacg agcgcgagga ggtcgcgcgc 65340
gagaccgggg gccgcgacgt tgggtaccag ggggcccggg tcctcgaccc caectccggg 65400
tttcactcgc accccgtggt ggtgtttgac tttgcagccc tgtaccccag catcatccag 65460
gcccacaacc tgtgcttcag tacgctctcc ctgcggcccg aggcctcgcg gcacctggag 65520
gcggaccggg actacctgga gatcgagggt gggggccgac ggctgttctt cgtgaaggcc 65580
cacgtacgcg agagctgct gagcatcctg ctgcgcgact ggctggccat gcgaaagcag 65640
atccgctcgc ggtaccccc aagcaccccc gaggagggcc tcctcctcga caagcaacag 65700
gccgccatca agtggtgtg caactcgggt tacgggttca cgggggtgca gcacctcctt 65760
ctgcctgccg tgcactgggc gccaccgtg acgaccatcg gccgcgagat gctcctcgcg 65820
acgcgcgctg actgcaacg gcgctggggc gagttcgatc agctgctggc cgactttccg 65880
gaggcggccc gcatgcgcgc ccccgttcg tactccatgc gcatcatcta cggggacacg 65940
gactccattt tegtthttg cgcgggcctc acggcgcggg gcctgggtggc catgggagac 66000
aagatggcga gccacatctc gcgcgcgctg ttctcctccc cgatcaagct cagtgccgaa 66060
aaaacgttca ccaagctgct gctcatcgcc aagaaaaagt acatcgcgct catctgcggg 66120
ggcaagatgc tcatcaaggg cgtggatctg gtgcgcaaaa acaactgcgc gtttatcaac 66180
cgcacctcca gggccctggg agacctcgtg ttttacgacg ataccgtatc cggagcggcc 66240
gccgcgtag cagagcggcc cgcagaggag tggctggcgc gaccctgcc caggggactg 66300
caggcgttcg gggccgtcct cgtagacgcc catcggcgca tcaccgaccc ggagagggac 66360
atccaggact ttgtctcac gccccaactg agcagacacc cgcgcgcgta caccaacaag 66420
cgctggccc acctgacggg gtattacaag ctcatggccc gccgcgcgca ggtcccgtcc 66480
atcaaggacc ggtacccgta cgtgatcgtg gccagacccc gcgaggtaga ggagacggtc 66540
gcgcggctgg ccgcccctcg cgagctagac gccgcggccc caggggacga gcccgcccc 66600
ccagcggccc tgcccctccc ggccaagcgc ccccgggaga cgcctcgcga tgccgacccc 66660

```

---

-continued

---

ccgggaggcg cgtccaagcc ccgcaagctg ctggtgtccg agctggcggg ggatcccggg 66720  
tacgccatcg cccggggcgt tccgctcaac acggactatt acttctcgca cctgctgggg 66780  
gcgccctgcg tgacgttcaa ggcctgtttt ggaataaac ccaagatcac cgagagtctg 66840  
ttaaagaggt ttattcccca gacgtggcac cccccggacg acgtggcgcg gcggtcagg 66900  
gcccgggggt tcgggcccgc gggggccggc gctacggcgg aggaaactcg tcgaatgttg 66960  
catagagcct ttgatactct agcatgagcc ccccgtcgaa gctgatgtcc cgcactcttc 67020  
aataaatgtc tgcggccgac acggtcggaa ttcccgctc cgctggttcc tctgcttgc 67080  
gtctgaccac gagcacaac gtgctctgcc acacgtgggc ggcgaaccgg tagccggggc 67140  
acgcggtcag catccgatcg atgagccggt agtgcagggt ggccgacgtg ccggggaaga 67200  
tgacgtacag catgtggccc ccgtacgtgg ggtccgggta aaaaagaac cgggggtcgc 67260  
acgccccccc tccgcccagg atcgtgtgca cgaaaaagag ctcgggctgg ccgagcgtat 67320  
cggccaggag gtccctggagg ggggtgctgt ggcggtcggc cagcacgacc agggaggcca 67380  
gaaaggtgcg gtgctcaaag atcgtattga tctgctgcac gaaggccagg atgaggcct 67440  
cgcggctgac ggtggccagc gcgccgtcgc ccgctgca cgcggggcag cagccccga 67500  
tcccaggta gtgcccctg cccgagaggg tcaggcagtt gtcggccacg gtctggtcca 67560  
ggctgaaggg gagcgcacag ggggtcgtct tcaccagggg cacggagagc gagcgcacga 67620  
tgccgatctc ctccggggg gtctgggcca gggcggcga gaagcccgcg tagcgcggc 67680  
gctcgtcag gcagagctcc agcctgcgcg cgtgcgacgg caggctcttg cgggaggccc 67740  
ggcgtccac gccggggttc ccggcggcgg aaaagcgcga ccgcccggg gtcttgcgc 67800  
ggccgggccc gggccgggag ccgagcgcac gggggcgat gtcatacata ggtacagagg 67860  
gtgtgctcca gggacaggag agagatcag tgctgtctga gcagcgcgc gccctcgcgg 67920  
acaaatgtgg ccagcgcggt gggcttcggc acaaatcct ggtacgtctt gaaggtgtag 67980  
atgagggccc gcagggctat acagaccgc cctcgaact cgttgcgca ggccaacttg 68040  
gccttgtgaa gctgcagctc gtcgcgatgg tcggcgcggg ggtggccaaa caggaccag 68100  
gggtcgaact ccatctcgt gatggcgcac atcggatcgc agaacatgtg cttgaagatg 68160  
gcctcggggc ccgcccggc aagcagctc acgaaccggc ccccgcccc gggctcgcgc 68220  
tcggggtccc cctcagctg gtccaacgac ggcactatgc agtcgaagag gctggtggtg 68280  
ttctccgagt agcggacgac ggacgcctc aggcgtcga tggccagcca gtaggcccgc 68340  
accagcaaca gattgcacag caggcattcc ccgcccgtgc gcccgcccc ccggccgtgc 68400  
ttcagcacgg tggccatcag cgggcccagg tccaggtcgg gctggggctg gggctcggcg 68460  
aactgcgcaa aacgcggggc gcgctcgcgc atgcgcgccc cgcggtgcgc tcccaggac 68520  
tcgctgaccg cggcgcggcg ggcgtccgcg gcggcgcgca gccggggccc cgactcccag 68580  
acggcggggg tgcccggcag cagcagcagg atcaggtcgg cgtacgcca cgtctccggc 68640  
tcacccccct gcgccagcgc cccggcggcg gcctcgaact cccggttgcg ggcggcggcg 68700  
cgcgtgcagc agctgtctcc gccccgcgc ttgccctcgg tgcagtcgag caggcggggc 68760  
cagtccttcc agttcatcag ggcgggtggt agggagggtt gcgttccga gcccccggcc 68820  
gcccccgccc ccgcccgtc atcgcggcg gaggccaggg tcccgatgag ggcgggggtt 68880  
cggactgcg cgaggaagga atagttggag tactgcacct tggcggcgc ccgggagggc 68940

---

-continued

---

gtcggcctgg gttgcttctg ggcgtggcgc ccgggcaccc cgcctcgggt ccggaagcag 69000  
cagtggagaa agaaatgccg gtggatgtcg ttgatggta gggcgaagcg cgcgaaggag 69060  
ccgacaaggg tcgccttctt ggtgcgcagg aagtgggtgt ccatgacgta gacgaactcg 69120  
aaggcggcca cgaagatgct cgcggcgcag tggggcgcgc ccaggcactt ggcgcagagg 69180  
aacgcgtaat cggccacca ctggggcgag aggcggtagg cctgcttcta cagctcgatg 69240  
gtcgggcaga ccagacaggg cgcgtccagc gcgaagggtg cgacggacgc cgcggcgaag 69300  
ggccccgtgt ccaagagtcc ctctgccgtg gggctctcgg cggggccgcg ggcggacccc 69360  
ggccccgcc cccccgaagc ctccgcgcgc gcccccgcgc gccgcggggg ggcgggcgcg 69420  
acgtcgctct ccacgtcctc gtcgagcgcg ctccggggcg gcacgcctac cacgtgacag 69480  
gccccagga gctcggcgca cagggcctcg ttaagagcca gaaggtcggg atcgaaggcc 69540  
acatacggac gctcgaagc gccctcctc cagctgctgc ccggcgactc ttcgcgcacg 69600  
gcgcgctcg acggcacc ccggggcgag gtcgccatgg ccggtcagc ggggcgcacg 69660  
cgtccgcgaa cgttaaggga cgcgacccc gactgcgcgc tgggttccca gaccctgaa 69720  
agtctagac cgcgctacgt ctccgcgagc ggcgcggggg acgcggccgt ctggttcgag 69780  
gacatgacc ccgccgaact agaggttata ttcccacca cggacgcaa gctgaactac 69840  
ctctcgcgga cgcagcgggt ggcctccctc ctgacgtac ccgggcctat aaaagcggcc 69900  
gacggccccg cggccccaca tacgcaggac accgcgtgcg tgcacggcga gctgctcgc 69960  
cgaaagcgcg aacggttcgg ggcggctatt aaccggttcc tggacctgca ccagatcctg 70020  
cggggctgac gcgcgcttcg gcggggcacc ggcaccggga ccgacttgtt ttacataaca 70080  
gtaggggggt ggggaacgcg cacccttgcc cggtcgcgat ggcggggatg gggaaagcct 70140  
acggcggccg cccgggggag gcgttcgagg gtctcgttca gcgcacagg ctcatgttc 70200  
ccgccacgct gcgcggcggg ggtggggagt cggggcccta ctcccatcc aaccgccct 70260  
cgagatgtgc ctccagttc caccggcagg atgggtccga cgaggcctc ccgacgagt 70320  
acgtcctcgc gctcatgaac gactgggccc atgtgccctg caaccctac ctgcgcgtgc 70380  
agaacaccg cgtttcgggt ctgtttcagg ggttttttaa ccggccccac ggcgccccgg 70440  
ggggcgcgat cacggcggag cagaccaacg tgattctgca ctccaccgag acgacgggac 70500  
tgtccctcgg agacctggag gacgtcaagg ggcgcctcgg cctggacgcc cggccgatga 70560  
tggccagcat gtggatcagc tgccttctgc gcatgccccg ggtgcagctc gcgtttcgg 70620  
tcatgggccc cgaggacgcc gttcgcacgc ggcggatcct gtgtcgcgc gccgagcagg 70680  
ccctcgcgcg tcgcgcgcgg tccaggcgggt ccagagatga ctacggggcg gtgggtgggtg 70740  
cggcggcgc caactcttc ggagcgcgcg ggcgggggt cgcgcctcgc gcccccag 70800  
cgcgcgcgcg acggggacgc gcccgctcgt ggcacaggc cgtgcagttg ttcggggccc 70860  
cgcgtccggg cccccggcg cttctgttcg tggcggcggg gctgtttctg ggggcgcgta 70920  
tctggtgggc ggttggcgcg cgcctatgaa agggggcgag ccaccgtccc gcccgccagt 70980  
gcatcccaga cgcgcgcgag ccgcacatcc cctccgctcc cgcctccggc ccgattctta 71040  
cggcgcgacc caaggtcccc atggccgcc cgcagtttca ccgccccagc accattaccg 71100  
ccgacaacgt ccgggcgctc ggcagcgcgc ggtcgtgtt ggccaccaac aacgctcagt 71160  
tcatcatgga taacagctac ccgcatccg acggaacgca ggggtcgggtg cgagagtctc 71220

-continued

---

ttcgcgggca	ggccgcggcg	ctgacggacc	tcggggtgac	ccacgccaac	aacacgttcg	71280
ccccgcagcc	tatggtcgcg	ggcgacgccc	cgcccgaaatg	gctgcggccc	tcgttcggtc	71340
ttaagcgcac	gtattccccc	tttgcgttc	gcgaccccaa	gacccccagc	acccccgtgag	71400
tcctcggcgg	gtccctccgc	ggccgtctct	cgttgcccc	ctttcccct	tcccgggtgg	71460
ttcaataaaa	aacaccaaca	tacgatattc	gcgtttgata	cgtttattgg	ggggggtgta	71520
gggcccacg	atcggcgatt	aacaacacca	aacaatcgag	cgctctaac	ccagtaacat	71580
gcgcacgtga	tgtagctgg	tcagcacggc	gttgctgcgc	tgaaacagcg	ccctgcgggt	71640
ccgctgcagc	tgtgtgtgta	tgcggcggca	tcgcgggatc	aaaaccgcca	gggcgctaag	71700
accggtgctt	cgtacgtagc	gtcgcgacaa	gacggcattt	gcctgtacgg	gcaagggggc	71760
aaattgcgag	tgtggtgact	ggaggtggtc	ggcggccaat	gggcggggtg	gttcgctcggc	71820
ggggggcaag	tgcggttcgc	gtgggagggg	gtcgcgccc	tcggtatcat	ccgagtcgca	71880
gaaacgcagg	gagctctcgt	cggagtgctt	atcatcggag	gagatgtgca	gcgtctgaag	71940
cagcgatgcg	ggtggggggc	cggagtcgac	gtgaagcgcg	agagaggaag	cccacgaagt	72000
cacagcggac	actgggaggt	gggtgtttgt	atgtgtggga	gactcggggc	tcgggaccga	72060
gtctcggctc	tggggtgtaa	gcgtccgagt	tacgggcggc	agggcgggct	ggggcagggg	72120
cggtcggggc	aggggcggct	ggggcagggg	cggtcggggc	agggcgggct	ggggcagggg	72180
cggtcggggc	aggggcggct	ggggcagggg	cggtcggggc	agggcgggct	ggggcagggg	72240
cggtcggggc	aggggcggct	ggggcagggg	cggtcggggc	agggcgggct	ggggcagggg	72300
tttggtcgcg	ggtgactggg	gtggggggcg	gcgggcaacc	gggcctccgg	gcacgacca	72360
accgcacaaa	ggctcgcctc	gggcaaccgg	gcctggggcc	aaaggcgggg	ggctggtctg	72420
gacggcggag	gtcggggggc	caaggcccgc	agaagcggc	actgccccc	ctgcggcggg	72480
aaccgcggcc	gcgtggtcgc	ctgggtcccc	gggagagggg	agggagtca	acgagccga	72540
gagcgcggcg	accgcggggc	gcgtgagggc	ccggggtggg	ccggccgcgg	ggccccgggg	72600
gggtgctggc	gagggaccgc	ctgtgtctc	gcggcggccg	cggcgggcgg	cgccccgggg	72660
gacgaccgct	ccttcggcgc	gcggagggcg	gatgggcgcg	agcgtggggg	cgggaaaggc	72720
cccgcgagcc	gagcgggggc	cgggcggaag	gggcaaagca	gaaacccaag	ccgggggcgc	72780
ggactccggg	gtgggcggct	ggtcgggagg	acgcgcggaa	cgggcgaccg	gggcgaccgg	72840
ggcggggagt	gccggcggac	gccaccctc	ggggggggcg	gaggcccggg	gcgcgcgcga	72900
tttggcacgc	gtccggcggg	acctgcgcac	gcgcggcacg	gcggcggaga	aagcggcggc	72960
agagccggaa	aaggccgggg	gaggaagcgc	ggcatccgcg	gggggactcg	gtgtgggtgg	73020
cgagggccgt	gggtcgtcgc	gaggggcccac	gggcacgcgc	cccgtgtttt	gttgaggcgg	73080
gacactcggg	cgtgttccgc	gagccgtagc	tcggcggccc	atgggcggcg	gtgcgtactg	73140
ggacgtgggg	acggactgat	cggtggcggg	ggggggaaga	agggccgggg	ccggattggg	73200
cgtggggcgg	ccggcgtcgt	cggacgccag	ctcctccagg	ccgtggatcc	aggcccacat	73260
gcgagggggg	acgggctcgc	cggtggtggc	gtcggtgagg	agagtggggg	cgaggacccc	73320
cgggtccgcc	tgcctgcggg	ggggggcagc	ggggtcctcg	ggaaccgatc	cgccatcccc	73380
ccccgcaagg	tcccgcgggt	gcgcgggcgc	ggtcggggca	gagggacctg	cctcgtcggc	73440
gagggggcgc	tggtaaaccg	ggtgtcccg	gaacagctcc	cccgtcagga	gggagcgtc	73500

-continued

---

gaagggccgc cccaggatgg cccgcgcgaa gaaggggtcc gcgtcggcgg cgtccgccgc	73560
gagaacgtcc cccgcggtag ccacaaaacgg aagctcctcg gtggcctcgc tgcccacaaa	73620
ccgcacgtca ggggggcccgg ggggtccgg ggcttcccac aagaccgca cgggggtcat	73680
ggagatgtcc acgaggacca ggcacggggg cccgtcggcg agagggcgt cggcgatgag	73740
cgccgacagg cgcgggagct gcgcgccag acacgcgttt tcgatcgggt tgagatcgg	73800
gtggaggagg ccgacggccc acgtctcgat gtcggacgac acgacgtcgc gcagggcggc	73860
gtccggcccgc cggggcggcg agtcgaagag cgtcaggcac agttccagtt ccgactcgcg	73920
ggagaaggcc gtggtgttgc ggagcggccac cacgacgggc gcgcgagga gcaccgcggc	73980
cagaaccagg tccatggccc taacgcgcgc ggcgggggtg cggtaggtcg cggcgccag	74040
cacggccacg tctggcccgc tgggtcggta gagggcgtgg ggggcctcgg ggagggacgc	74100
ctcgcgccc cccgcggggc cgagcgtctg gccagactcc aggcgtgccc ccaggagggc	74160
gtcgaagctg tctactcgg ttagtctgctc gggaaacatg caggtccaca gcgcggccaa	74220
agcggcgtc ggccagacaca tgcgcccag gacgtcacc gccccaggg cctgggcccg	74280
actgagcttc ccgagcggc ggacgtccc gcctgggtc ccgagctcca agccgagcg	74340
ccagggcgc agcgggtcgg ttctggacag cttgcccgc gcgccagtc ccagccgct	74400
gccgaacagg agccccggg tccggggggc tccgtccaaa aacgtcggca acacgcggat	74460
gcgggcgtc ggatgcgggg tcaggcgtg gacgaacagc atggactcc ctgctcctc	74520
gaacgcgctc tcgagggtga ggtgcctgta ctctgctgg cgaacaggt ccaggcgcca	74580
gaagttagt atgtgttccg gaacgcggc caccagcgc accagcagct cgttctcgtt	74640
gaaggcagc cagtggcgtt gggaccccc ggggcccggc ggcggacgc gcgcccccgc	74700
tccggacgcc cagcccagct gggcccagc acacccaaa tcgcgcgtga gggtaggtgc	74760
gacgaggcg acgtacagct cggcgcgcg gtccatcgag gcgccccag tcgcctggcg	74820
atggcgcacg aagcgaccga acagctgaaa gttggcggc tgggcgtcgc tgagggccag	74880
ctggagccg ttcacgacgc tcagcacgta catggcctg accgtcggg ccgattcag	74940
gacgtccgtc ggaagcggg gccgcacgca ggcgcctcg ggacgcatca gcagcgcgc	75000
gagttgtcgt gtgacggccg ggaagcatag cgcgtactgc agcggcgtt cgtccggggc	75060
caaaaagctg gtggcgaac gcagatccag agcgtgacg gcctcacgca gcaccagggg	75120
ccccgggtct ccgcggcgc gcagatacgc ctcgcccgg cggcgcagca gctgcgggtc	75180
gacctcgtg cctcggggg aagaagagg ccgggcgcg gcgtcaggg ccggaagatc	75240
aacgagcagg ggcggggcg cggactcgc gccgcgccc gtctggcgc cggcctggc	75300
gtacgcgcta tataagccca tgcggtattg gatgagttcc cgcgcgccc ggaactcctc	75360
caccgcccac ggggccaggt ccgcggcgc cgcgtcgaac tccgcagca ggcccccag	75420
ggcgtcaaag ttcactctcc agggcaccct gcgcaccacc tcatcccga gccgggcgca	75480
cagggcgggtg tcttgggtga gcgcgcgccc cagctcctcc acggcctcc cgcgtcggc	75540
gcccttggcg cccagacgc cctggtaacct ggcggaaagg cgtcgtagg ccggctgggc	75600
ccgcagcccc gacaccgtgt tggtaggtgc ctgcaggggc cgcagctgct cgtgcatggc	75660
gcggaacccc tcgggggact tccagggccc cccccggag cggccaaagc gaccccagac	75720
ctcgtcccac tccgctcgg cctcctccag ggacctcgc agggcgtcga cgcggcggc	75780

-continued

---

agtatcaaag agcgccccca ggcggccggc gtgcccgcgc agggggccgg ggcctgcgc 75840  
 gcggggcg cttagcgggt gcgtctcgaa ggtgcgctgg gcgtgctcta gccagataac 75900  
 cgcgggcacg tcgagctcgc gcgtttttc ggtctgatcc aacagaaacct cgacctggtc 75960  
 ggcgatctcc gccaccgagc gcgcctggtc gagcgtcttg gccacggtcg ccgggacggc 76020  
 gaccaccttc agcatggtct tgaggttggc caggccctcg gcctcgatct gggcccggcg 76080  
 ctgcgcgcgc gccacgcct cccgcaggcc cgccatgacc cgctcggtgg cctccgcgcg 76140  
 ctgctgtttg gcgcgcacca ctgcgtcctt ggtctcggcc gtgtcctgcc gggtcacgaa 76200  
 ggcacatac tcggcgtaag ccgtgttctt cacggggctc tggtecacgc gctccaacgc 76260  
 cgccgcgcac gcgaccagcg cgtcctcgct gggacacggc agggtgaccc cgtccggac 76320  
 cagctccgcg gtggcctccg ggtcattccg ggccgcggat atctgctccg ccggcgccgc 76380  
 caggctccag ggcacgcgc cgagcgcgc gtgcacgtcg gcccggtgg cgtccaggcg 76440  
 atcgcggagc tccacgtagt cggcgtagcc atgttgaag aacggcacgt accggcgagc 76500  
 gccgggacg ctgcctatgt cgtccgcag gcgccccacg gcctcgtggt agtcgataaa 76560  
 cccgtgcgcc gcctgggcca tttccaggag cccctccgcg atgcgcagca gccgcgccag 76620  
 gggctcggcg tcgaccgaa acatgtcggc gtaggtttcg gcggcgcgct ggaacgccgc 76680  
 gctccagccg agcggtgga tggcgcgag ccgggggagc atgggggtggc gctggttctc 76740  
 gggggtgtag gggtaaacg cgaaggcct atccaggcg agggtgacc cctcgcgctt 76800  
 ggccgcgagc gcctgctcgg gcgccttgcg gaagtcccg gggttgtagc cgtgcgtgcc 76860  
 cgccagcgc tcgaggcgcc gcagctcag cacgtcgaac tcggcgcggt tctcgacgcg 76920  
 gtccagcgc gcctcgacgc cggcgccca gcctcgtcg ctgccccggg cgcgctgggc 76980  
 cgccatcttc gccgtcaggt cggcgacggc ggctcaagt tcgtcgccgc ggcgtcgcgt 77040  
 ggcgccgatg acctgcccc gctcctcag gcgcgcgcc ctgggggaat ggtccccgc 77100  
 cgtcccttcg gcgtgcagca ggccccgaa cccagcctcg tgccccgca ggtttcccg 77160  
 agcagcggtc gtcgcgcgg ccgcggcacc gatgaggcg gcctgggtcc cctccggctg 77220  
 ggcgcaggcc cgcgcgctg ggaactaccg gtcggcgcc gccgaccca gggctgtag 77280  
 ctgcctgatg gcccccgcg cctccaggc cagccgagtc gcctttacat accccgcgcg 77340  
 gctatcggcc agcaccgca ggaaggacag gggcgaggcc gggtcgcggg ccgcccgcgc 77400  
 caggggcgc accgcgtcc caggggcgc atgcgcgcgc acggcccgct ccaccgtcgc 77460  
 cgcgggactt gccgtcgcga cggcgcgct cccggcgctg atggcgcttg acacggcttt 77520  
 ggcgattgtg gggcgctgat cggaaaagaa ctgcacgagg accggcgtct cggggcgctc 77580  
 ggcgaacagg gtcttcagca ccaccacgaa ggcgggatgc agggcgcca gagccgtcgc 77640  
 ggtatccggg gtcgggtgtt ccagggcctc ccggtactgc cccagcagcc cccacaggtc 77700  
 cgcccgcagc gccccgtga cttccgggg ggggcccccg acggcctcgg ccaggctcgt 77760  
 ccaccccgcg ggcaggagg cccgcagggt cgccagcag gccggacacg cctttagccc 77820  
 cacaaagtcc gggaggggcc gcaggacccc ttggagtgtg tgcaggaact tctcccgggc 77880  
 gtcgtgggcc accttgggc gctcccgcgc gtcggtgagc atgcctcca gggcgtgggc 77940  
 gcgctcccga agcggggagc gcgcctccgg agcagctcc gccgtcatct tggccgctc 78000  
 catggccctc gcctgcgca gcgcgtcttc ggccatgcgc gtggcctcgg gggacagccc 78060

-continued

---

gccccctcg acgtacggcg cggggccggg cgcgggacg aaggcccggt cgctgtccag	78120
ctgctgctcg agcgcggcggt cgagggcgct gaagcgctgc agttcggcca gccccgagct	78180
gcgcgcggcc tctggtgctg tgatgcccgt gatgctgcgc gccagctctt ccaggggctt	78240
gcgttcgatg agccccggg tcgcggcgct ggtcaggacc gagagccagg ccgccaggtc	78300
ctcgggggca tctaggggct gggcccctg gagcaggctc cgcagcagga tggcctgggg	78360
gctggtggcg agggggggcg ggggggggag cgcggcgcgc tgagcagcgt ccgcgctgtg	78420
ttggtcaaag gccggtagcg attccagcaa ctggaccatg ggcacgaccg cggccgaggc	78480
cacgtgaaac cgacagtcgt ggctgtcgct ggctgcagg gccttcgcgc tgtatacggc	78540
tccccggtgg aagtactcct tgaccgcgct ctcgatcgcc cggcgggctt ggatccgcac	78600
gtcctccagc cgcgcctgga tggcctcggg gccagggcg ggcgggcacg gggccctgcc	78660
gccggcgccc gggggggcg gcacgggcat cacggtcagg ggcggcgcg cgtgcgagac	78720
cgagtcgacc ccgcgggca gggcgtctaa ggcttcgcgc atctcgggg cctccgcctc	78780
gacccgcatc tcttcgcccc gggcaaacct ggccagcgcc tggatccgat ggagaagcgg	78840
ctccgggtgc gtcgggggtg cggggggcga cagggtgttc ggggtggcg cgcagcgctc	78900
caggagccac tctccgaggt gtgcgtacag attggccggc gggcgggcg gcagctgcag	78960
atccaggtcc gcgaggtccc cgtaaaaggc gtccgtctcc cgaataacgt cctggcgac	79020
caggaccagc tttagcgagg ccagggcgccc gatctgcgaa ttttcgtcca gcacgtgctg	79080
gatgaggggc cgtgggggcg ccacgtccgc caggctcatg cgcgtggacg ccaggaagtc	79140
cccagcggcc gttttcggg gcagcatgcg cagggtgaag tccagcaggg ccgcggcccg	79200
gccggccaacc ccggcctgct tatgctgctg ggcgccgttc tcgatcaaaa aggcgaggac	79260
gcgctcaaag aagaagatga cgcagagctc caacagcccc ggggtgcccg ggtacggcga	79320
ccgcaggcg ttgatggtga gctgcgaaca cgcggccacc tcgcgggcca gggcgcatc	79380
gcgcgcccg agcgggaccg ccgtggcggc cacattgggg tggacctcga acagctgcgc	79440
caggtcggcg ccggggggct ccggggggcg gcgggcccc agcgtctcga gcacggacgg	79500
cgacgacggg ctgcggggcc cgtcgtcgcc gccgcctgc ccgactgcg ggggggtatc	79560
cggtgcggga gggaccgtgg cggctatggg cgtcggggag gaggcgggga cctcggcggc	79620
gacgggggccc tcttctcttg gcgcggactt cttcttggcc ttggcggggc gggccttggg	79680
ggcgggctc tcgcccaggt tcagatcctc cacgctggac ggtggggtcc aggtgggccc	79740
gcggcgcttg ggcaagccgg tagaatagcg cgcgggtgg cgacccaacc gactgcccc	79800
cacctccagg acccgcaggt cctcggttc ttccggcccg tcccggcggt gtgtctgctg	79860
ggcgggggcg gcgtgcgggt gaccgagggc cgcggcgctc gggcccgagg gcttcgcggg	79920
cgggggcccc tccagggctg ctgcccacac atcctcgggg gggcggtttg ggtgccccgc	79980
ctgcggtgtg tcgggtgggc ccgagggccc ccggggggcc tcggggggcc ggtcggcccc	80040
aggggtctgg acgtgggtgg gcgcggggag cgcggggacg accgggcccg agccttctcc	80100
gtccccctg gggaccacac gcacaaagag cgcgccgagc ccccgatct cgcgccgag	80160
ggggtgggtg atggccacgc gccgctcgac gaacggttcg tcctgcaggt aagtctcgt	80220
ggccccgtg aggtgcaggc ccgcggcggt caggctccgc ggcgccacgg cccccgggccc	80280
ggagggcaca aaaaacacca tggcgcccc ccaccgacc ttggggcggt cgtggcgcta	80340

-continued

---

atacgtcagg tacgggtaca cgtcgcccgc ccgcacctg gcgataaacg cgggcgttcc	80400
cgcgggcagg ccgtgcgggt caaacagata ggccgtgtcg ccgtcccggg agagccccc	80460
gcccaggggg ccgatgggta ggagcgtgta ggacagcggc cgcattggccc aggggcccggc	80520
gaagaacgtg tgcgcggggc attgctctc cagcagcccc gccgtgggct ccccgaagaa	80580
gcccacctcg ccgtacaccc gcgaaaaac gcaacgcagg ccgcgcgcgc ccgccgggta	80640
ctccaggaag ttggggagct cgataatgga acacatgcgc ggccggcccgg agcccgcggc	80700
cgcgcgcgtc cactcgcccc cctccaccag acatccctcg atggcctccg cggacagcac	80760
gtcgcggggc cccacgtcga aaagaagact gagaaacgac agggacgagc gcatgcacga	80820
taccgacccc ccgggtcca gatcggtcgc gaactgggtc cgaacaccgg tgaccacgat	80880
atcgcgatcc ccttggcgtc tcatcgtggg gtgaggtagc gcggccggaa tcatgtgtgc	80940
cgcgcccgc acgagcgggg cctgtttatg ggccgggctg cccgatgagt actgtgttt	81000
ccgcgcgcgc aaccccccg cccatcaacc gcctgttcgt ccccctaacc acacaccgg	81060
tatcgcgtgt gtgtggtttc ccgggaagac acatcccacc ccatgaagtt ttgcccttt	81120
tttcgctcc gactacgcc acctttccac cccccccaa aaaaaaaca accaactccc	81180
agatggatgg gtgcgataa aaagctttat tattgtttaa ccaaaggcga gtectacggg	81240
tgtaccggg gtgtctcctg cggcgtcctc tcgtcgtcct ccacgggggt gttgggccc	81300
gggaccgtct cgcgcccgc cgggcgcgtc gacggcgcgc gggcctgcgt gtectgtggg	81360
ccgggtgtcg tgggttcggg ggtgctaccg ccggcattt gggcctccag gtcccgggg	81420
gccccgggc cgcgcgaagg ccgaaacgc gaggcgcgaa acacgcgcgc ggtgacctgc	81480
aggagctcgt ttattaatag ccagtcctat ctcagcgtag cggccagccc ctggggagac	81540
aggtccacgg agtccggaac caccgtcggc tgaccacagg gccccagget gtagtcccc	81600
caggccccca ggtcatgacg gttcgtgagc acgacgaggt ctgcggccgg gctggggggc	81660
gcgtcctcgg tcgcgtgggc catcacctcc tgaatggctg cgggtgcgtg atcggccgag	81720
ctggcgaagc gcgccacgac cagcgcgcgc tccgtctgca ggccttcca cgtgtcgtgg	81780
agttcctgaa cgaactcggc caccgcctcg gggccctgg ccgcgcgtgc ggctgatag	81840
ccggccgaga ggcgcgcga gcgcgcagg aactgactca tgtaacagaa cccggggacc	81900
tggtcccccg acatcaactt tgacgcctc gcgtggatgc ccgacacgat gccagggaac	81960
ccgtgattt cccgcgcac gacggccagc acgttaccct cgtgcgagac ctgggcccgc	82020
agctcgtcgc ataccctcag gtgcgcgcgc gtctcgggta cgacggaccg cagcccgcgc	82080
agggacgcga ccagcgcgcg cttggcgtcg tgatacatgc cgcagtactg gctcaccgcg	82140
tcgcccattg cctcggggcg ccagggcccc aggcgctcgt gggcgtctgc gaccacggcg	82200
tacaggcggg gccctcgtct ctcgaaaccg cactcaaaga aggcggcgag cgtgcgcattg	82260
tgacgcgca gcagcagat cgcgtcctcc agctggcggc ccagggggtc ggcgcgctcg	82320
gcgagctcct gcagcaccgc ccggggccgc agggcgtaaa tgctgatcag cagcaggctg	82380
ctgccacct cgggaggtg ggggggaggc agctggaccg cgggcccag ctgctcagc	82440
gccccctgg cgatcacgta cagctcgcgc agcagctgct cgatgtgtgc ggccatctgc	82500
atcgtgggcc cgacgcggc ccgggtggcc ggttcgagga ggggtgatcag cgcgccaat	82560
tttgtcggg gccctcagc ggtggggaga tagcccagc cgaagtcgcg cgcaccggcc	82620

-continued

---

```

agcaccgccga gggcaaacctc gatggggcgg ggcaggtagg cagcgttgca cgtggccctc 82680
agcgcgtccc cgaccaccag ggcagcacg taagggacga accccgggtc ggcgaggacg 82740
ttgggggtga tgccctccag ggcgggaaag cggatcttgg tggccgcggc caggtgaacc 82800
gagggggcgt ggctaggcgg cccgacgggg agcagcggcg acagcggcgt ggcgggggtg 82860
gtgggggtca ggtcccagtg ggtctggccg tacacgtcga gccagatgag cgcctctctg 82920
cgcaggaggc tgggctggcc ggcgctgaag cggcgtcggc ccgtctcaa cccccccag 82980
agcgtgcgcc gcaggctcgc caggtgttcc gtcggcacgg cggggcccat gatgcgcgcc 83040
agcgtctggc tgaggacgcc gcccgacagg ccgaccgct cagagagccg cccgtgcgtg 83100
tgctcgtcgg cgcctcggat ccgcccgaac gttttcacgt agccggcgtg gtgcccgtac 83160
tcccgcgcga gcccaaac gttcgcctcc gcaagggcaa tgcacccaaa gagctgctgg 83220
atctcgtcga gccctggccc gggggggcgtc cgcgagggca cccccccac caaaaacccc 83280
tccagggcgc atatgtactg ggtgcagtgc gcgggcgtga acccgcgtc ggttaagcgtg 83340
ttgatcacca cggagggcga gttgctgttc tggaccaaag cccacgtctg ctgcagcagc 83400
gcgaggagcc gttgctgggc cccggcggag ggcggctccc ctactgtcag cagcccggtg 83460
acggccggac ggaagatggc cagcgcggac gcactcagaa acggcacgtc ggggtcgaag 83520
acggccgcgt ccgtccgcac gcgcgccatc agcgtccccg gggggcgcga cgcggaccgc 83580
gggctgacgc ggttagggc ggtcgacacg cgcacctcct cgcgactgcg aaccattttg 83640
gtggcctcga ggggggggat catgatagcc gggctgatct cccgcaccgt gtgctgaaac 83700
tgggcccagca gcggcggcgg gaccaccgcg ccccgatcgg gggctcgtcag gtagtcgtcc 83760
accagcgcga cgcgtaaacg ggcggcgggtg aggggggtca gggcgcgctc gtcgatgcgc 83820
tgtaggtcgc ccgagaacag cgtcacccaa ttgctgacca gggccaagaa ccggagaccc 83880
tcttgacaga tcggggacgg gaagagcagg ctgtacgccc ggggtggtcag gttggcggcg 83940
ggttgccccg ggggaaccgc ggacatctta agcgacatct ccccaggggc ctccagggag 84000
gtcccggggt tcattggccc gcagctctgg gtgacggctc gccagcggtc gatccactcc 84060
acggcacact ggcggacgcg caccggcccc agggcggcgg tgggtgcgag cccggcggcc 84120
tccagcgcgt gggctcgtgc ggagccgggt atcgccagga ccgtgtcctt gatgacgtcc 84180
atctcccga aggcgcctc ggggggtctc gggagcggca ccgccatgcg gtgcaccagc 84240
agcccgggga ggttctcggc caagagcgcg gtctccggaa gcccggtggc ccggtgcaag 84300
gcgcacagtt gctccaggag cgggtgccag cagccccgcg cctccgcgg gcccaccgcc 84360
gcgcccgaca acagaaacgc gcgcgtggcg gcgtgcagtt tggccgcgga cagaaacgcc 84420
ggctcgtccg cgtgccccg cggctcgtc gagggggagg gcgcccggcg gaggttggtc 84480
aggctccccg acaggacctg caacggctcg tttgggggtg gagcggacgg ggggtcatg 84540
ccggcggggc ccgggacctg gagcgcgtg tccgacatgg cgaccggcgt gcgctcctcg 84600
cgacgcggcg cggagaccgc gggcccaaac gggaatgact gccgcggccc tatacggagg 84660
ggctaagtat cccccgggga ccttcgaaa ccccgggcgt gtcgcaagta cgcgcggaag 84720
gcgcccgtg ttatacggcg cgttatgtcc cggcattccg ttcgtgggtt cgggcccggg 84780
tgctgtcggg tgggagtgtg tgtggggggg ggcggcgcga cggcggcccg gaccaagtgt 84840
atcgcgccg ttcctggggg cggcccaaca ggcctttaa acatttgcgt atgcaccggc 84900

```

-continued

---

```

ccagccagtc ggacaccgga acccaccaga ggcggaagcc gccttcgccc gtgagggtgc 84960
gtgtgttttc tggtagcgtg tttttccttt cgcctctct cctcctccac ctccaccacc 85020
cccccccaca actcgcctgt tggcgatcgg cgggaaaacc atgaaaacca agccactccc 85080
gacagccccg atggcgtggg ccgagagtgc cgtggaaaacc accaccagcc cgcgcgagct 85140
cgcggggccac gccccgctcc ggcgcgtcct gcgcccgcct atcgtctgcc gcgacggccc 85200
ggtgcttttg ggggacaggg ccccaggag gacggccagt acgatgtggc tctgagggat 85260
cgacccccgc gagtctctc cgggaacgcg cgtaccgga gacgataccg agcaggccgt 85320
ggacaagatc ctacggggag cccggcgcg gggagggtg accgtccccg gcgccccccg 85380
ctatcacctg acccgcagg taacctgac ggatctctgc caacaaaacg cggagccggc 85440
cggggcgctc cttttggccc tgcggcacc caccgacct ccccacctg cccgcatcgc 85500
ggctccgccc ggcggcaga ccgagcact ggcggaggcc tggggccagc tctggaggc 85560
ctccgccctg ggtccgggc gggccgagag cggctgcgcg cgcgcgggccc ttgtgtcgtt 85620
taactttctg gtggccgct gcgcccgcct ctacgatgc cgcgacgcc cggaggcgtt 85680
cggggccacc atcacgacca actacggcgg gacgcgggccc ggggcccggc tggaccggtt 85740
ttccgaatgc ctgcgcgcca tggtcacac gcacgtgttt ccccacgagg tcatgcggtt 85800
tttcgggggg ctagtgtcgt gggtcacaca ggacgagctg gctagcgtca ccgccctctg 85860
cagcggacc caggaggcca cacacaccgg ccaccgggccc aggcctcgtt cggccgttac 85920
catcccggcc tgcgctctg tggacctgga cgcgagctg tgcctggggg gccctggggg 85980
ggcgttctg tacttggctc tcacctacc acagtgccgg gaccaagagc tctgttgcgt 86040
gtacgtggtc aagagccag tccccccgc cggactggag gcgcccctgc agcggctgtt 86100
cgggcccctc cgataacca acacgattca cggggccgag gacatgacgc cccctcccc 86160
gaaccgaaac gttgactttc cgctcgcct cccggcccgc agctcgcaat ccccgcggtg 86220
ctcggcgagc caagtacga acccccagtt tgtcgacagg ctgtaccgct ggcagccgga 86280
tctcggggg cgcctaccg cacgacctg cacatacgc gccttcgag agctgggtgt 86340
catgccagac gacagcccc cgtgtctgca ccgcaaccg cggtttgggg cggtcggcgt 86400
tccggttgc atcctggagg gcgtggtg gcgcccggcg gggtaggggg cctgcgcgtg 86460
atcgtctatt gacgagggcc gcccaaccgc agcgacctc cctcctccact tcccccccc 86520
tacacaccaa ctccgcccgc ccgtctctgg ccgtgcgcgg ccccgctcgt ccgtctcaat 86580
aaagccagg taaatccgtg acgtggtg tttggcgtgt gtctctgaaa tggcggaaac 86640
cgacatgcaa atgggattca tggacatggt acacccccct gactcaggag ataggcataat 86700
cctccttaga ttgactcagc acacgatcgc accccacccc tgtgtgccgg ggataaaagc 86760
caacgcgggc ggtctgggtt accacaacag gtgggtgctt cggggacttg acggtcgcca 86820
ctctcctgc agccctcagc tcttcgccc cggattcctg ttgcgttct gtcggccggt 86880
gctgtcctgt cgacagattg ttggcgactg cccgggtgat tgcgcccgc gtcgctcctt 86940
tcggtcgtac cgcaccccc gcctcccacg ggcgcgcgc tgtttccgtt cagcgcgtcc 87000
gagccaccgt caecttggtt ccaatggcca accgcccctc cgcacccgcc ctgcgaggag 87060
cgcggtctcc gtccgaacga caggaaacccc gggagcccga ggtcgcctcc cctggcggcg 87120
accacgtggt ttgaggaaa gtcagcggcg tgatggtgct ttccagcgt cccccggccc 87180

```

-continued

---

ccgcggccta ccgcattagc gacagcagct ttgttcaatg cggctccaac tgcagtatga	87240
taatcgacgg agacgtggcg cgcggtcatt tgcgtgacct cgagggcgct acgtccaccg	87300
gcgccttcgt cgcgatctca aacgtcgcag ccggcgggga tggccgaacc gccgtcgtgg	87360
cgctcggcgg aacctcgggc ccgtccgcga ctacatccgt ggggaccacg acgtccgggg	87420
agttcctcca cgggaaccca aggacccccg aacccaagg accccaggct gtccccccgc	87480
cccctcctcc cccctttcca tggggccaag agtgctgcgc ccgtccgcgat gccaggggcg	87540
gcgcgcgagaa ggacgtcggg gccgcggagt catggtcaga cggcccgtcg tccgactccg	87600
aaacggagga ctccgactcc tcggacgagg atacgggttc ggagacgctg tctcgatcct	87660
cttcgatctg ggcgcgaggg gcgactgacg acgatgacag cgactccgac tcgcggtcgg	87720
acgactccgt gcaccccagc gttgtcgttc gtcgcagatg gagegaacgg cccgcccccg	87780
tggcctttcc caagcccccg gcccccggcg actcccccg aaacccccgc ctgggcgcgcg	87840
gcacccggcc gggctccgcg acggaccocg gcgcgtcggc cgactccgat tccgcggccc	87900
acgccgcgcg accccaggcg gacgtggcgc cggttctgga cagccagccc actgtgggaa	87960
cggacccccg ctaccagtc cccctagaac tcacgcccga gaacgcggag gcggtggcgc	88020
ggtttctggg ggacgcctgc gaccgcgagc ccgcgctcat gctggagtac ttctgtcggg	88080
gcgcccgcga ggagagcaag cgcgtgcccc cacgaacctt cggcagcgc ccccgctca	88140
cggaggacga ctttgggctc ctgaactacg cgcctcgtga gatgcgacgc ctgtgcctgg	88200
accttcccc ggtcccccc aacgcataca cgcctatca tctgaggag tatgcgacgc	88260
ggctggtaa cgggttcaaa cccctggtgc ggcggtccc cgcctgtat cgcacctcgg	88320
gggttctggt ccacctgcgc atccgtaccc gggaggcctc ctttgaggaa tggatgcgct	88380
ccaaggaggt ggacctggac ttcgggctga cggaaaggct tcgcgaacac gaggcccagc	88440
taatgatcct ggcccaggcc ctgaacctc acgactgtct gatccacagc acccgaaca	88500
cgctcgtcga gcgggggctg cagtcggcgc tgaagtacga agagttttac ctcaagcgt	88560
tcggcgggca ctacatggag tccgtcttcc agatgtacac ccgcatcgcc gggtttctgg	88620
cgtgccgggc gaccgcggc atgcgccaca tcgcctggg gcgacagggg tcgtggtggg	88680
aaatgttcaa gttcttttcc caccgcctct acgaccacca gatcgtgccg tccacccccg	88740
ccatgctgaa cctcggaaac cgcaactact acacgtccag ctgctacctg gtaaaccccc	88800
aggccaccac taaccaggcc accctccggg ccatcaccgg caacgtgagc gccatcctcg	88860
cccgcaacgg gggcatcggg ctgtgcatgc agcgttcaa cgacgccagc cccggcacccg	88920
ccagcatcat gccggccctg aaggctcctg actcctggt ggcgggcac aacaaacaga	88980
gcacgcgccc caccggggcg tgcgtgtacc tggaaacctg gcacagcgac gttcgggcccg	89040
tgctcagaat gaaggcgctc ctgcgccggc aggaggccca gcctgcgac aacatcttca	89100
gcgcctctg gatccggac ctgttcttca agcgcctgat ccgccacctc gacggcgaga	89160
aaaacgtcac ctggctccctg ttcgaccggg acaccagcat gtcgctcgcc gactttcacg	89220
gcgaggagtt cgagaagctg tacgagcacc tcgaggccat ggggttcggc gaaacgatcc	89280
ccatccagga cctggcgtag gccatcgtgc gcagcgcggc caccaccgga agcccttca	89340
tcatgtttaa ggacgcggtg aaccgccaact acatctacga cacgcaaggg gcggccatcg	89400
ccgctccaa cctctgcacc gagatcgtcc acccggcctc caagcgatcc agtggggtct	89460

---

-continued

---

gcaacctggg aagcgtgaat ctggcccgat gcgtctccag gcagacgttt gactttgggc 89520  
ggctccgcga cgccgtgcag gcgtgcgtgc tgatggtgaa catcatgatc gacagcacgc 89580  
tacaaccac gccccagtgc acccgcgga acgacaacct gcggtccatg ggcattggca 89640  
tgcagggcct gcacacggcg tgcctcaaga tgggctgga tctggagtcg gccgagttcc 89700  
gggacctgaa cacacacatc gccgaggtga tgctgctcgc ggccatgaag accagtaacg 89760  
cgctgtcgtc tcgcccggcg cgtccctca gccacttta gcgcagcatg taccggggcg 89820  
gccgctttca ctgggagcgc ttttcgaac ccagcccgcg gtacgagggc gagggggaga 89880  
tgctacgcca gagcatgat aaacacggcc tgcgcaacag ccagttcatc gcgctcatgc 89940  
ccaccgccgc ctccggcccag atctcggacg tcagcgaggg ctttgcccc ctgttcacca 90000  
acctgttcag caaggtgacc agggacggcg agacgctcgc ccccaaacg ctcttgctga 90060  
aggaactcga gcgcacgttc ggcgggaagc ggctcctgga cgcgatggac gggctcgagg 90120  
ccaagcagtg gtctgtggcc caggccctgc cttgctgga ccccgcccac cccctccggc 90180  
ggttcaagac ggccttcgac tacgaccagg aactgctgat cgacctgtgt gcagaccgcg 90240  
ccccatgt tgatcacag caatccatga ctctgtatgt cacagagaag gcggacggga 90300  
cgctccccgc ctccaccctg tcccgcttc tcgtccacgc atataagcgc ggctgaaga 90360  
cggggatgta ctactgcaag gttcgaagg cgaccaacag cgggggttc gccggcgacg 90420  
acaacatcgt ctgcacaagc tgcgcgctgt aagcaacagc gctccgatcg gggtcaggcg 90480  
tcgctctcgg tcccgcatac cgccatggat ccgcgcgtct ccccgcgag caccgacccc 90540  
ctagataccc acgcgtcggg ggcggggcg gccccgatc cgggtgtccc cccccgag 90600  
cggtacttct acacctccca gtgcccgcac atcaaccacc ttcgctccct cagcatcctg 90660  
aacgctggc tggagaccga gctcgtgttc gtgggggacg aggaggacgt ctccaagctc 90720  
tccgaggcg agctcggtt ctaccgctt ctgtttgcct tcctgtcggc cgcggacgac 90780  
ctggtgacgg aaaacctggg cggcctctcc ggctcttcg aacagaagga cattcttcac 90840  
tactacgtgg agcaggaatg catcgaggtc gtccactcgc gcgtctaaa catcatccag 90900  
ctggtgctct ttcacaacaa cgaccaggcg cgcgcgct atgtggccc caccatcaac 90960  
caccggcca ttcggtcaa ggtggactgg ctggaggcgc ggggtcggga atgcgactcg 91020  
atcccgaga agttcatcct catgatcctc atcgagggcg tctttttgc cgctcgctc 91080  
gccccatcg cgtacctcg caccaacaac ctctcggg tcacctgcca gtcgaacgac 91140  
ctcatcagcc gcgacgagc cgtgcatac acagcctcgt gctacatcta caacaactac 91200  
ctcggggcc acgccaagc cgaggcgcg cgcgtgtacc ggctgtttcg ggaggcggtg 91260  
gatatcgaga tccggttcat ccgatcccag gccccgacg acagctctat cctgagttcg 91320  
ggggccctg cgccatcga gaactacgtg cgattcagc cggatcgcct gctgggctg 91380  
atccatagc agcccctgta ttccgcccc gccccgacg ccagctttcc cctcagcctc 91440  
atgtccaccg acaaacacac caacttcttc gagtgcgca gcacctgta cgcggggcc 91500  
gtcgtcaac atctgtgagg gtctgggccc cctgttagc atgtctaacc gaataaagg 91560  
ggtcgaaac gactgttggg tctccggtg gattattac caggggagg ggggtgccc 91620  
tggggaaag gaaggaaacg ccgaaaccag agaaaaggac caaaaggaa acgcgtccaa 91680  
ccgataaatc aagcggcag cagaacccc agatgcataa taacgatattt attactctta 91740

-continued

---

ttattaacag gtcgggcacg gggaggggat gggggcgcgc gtttctccg ttccggctac 91800  
 tcgtcccaga atttagccag gacgtccttg taaaacgcgg gcggggggcgc gtggggccac 91860  
 agctgcgcca gaaaccggtc ggcgatgtcc ggggcgggtga tatgccgagt cacgatggag 91920  
 cgcgctaaat cttcgtcggc gaggtcctga tagatgggca gtcttttttag aagagtccag 91980  
 ggtccccgct ccttggggct gataagcgat atgacgtact tgacgtatct gtgctccacc 92040  
 agctcggcga tggtcacggc atccgggcagc cagtccaggc cctccggggc gtctgtggatg 92100  
 acgtggcggc gacgtccggc gacatagccg cgggtgtccg cgaccgcctg cgcgttgggg 92160  
 acctgcacga gctcggggcg ggtgagtatc tccgaggagg acgaccgggc gccgtcgcgc 92220  
 ggcccaccgg cgacgtccgg gggctggagg ggggggtctt cttcgtagtc gtctctgccc 92280  
 gcgatctggt gggccagaat ttcggctcac gagatgcgcg tctcggaggcc gaccggggcc 92340  
 gcggtcagcg taggcatgct ctccaggagg cgcgagttgg cgcgctcccg ccgggcccgc 92400  
 cggcgggect gggatcggct cggggcggtc cagtgacact cgcgcagcac gtctctgacg 92460  
 gacgcgtagg tgttattggg gtgcaggctc gtgtggcagc ggacgaacag cgcaggaac 92520  
 tgccggtaac tcactctgaa gtaactgcagc aggtcgcggc agtgaatcgt ccgaatgtag 92580  
 ccggtgctga tgtccaacac gatctcgcag cccatcagca ggagatcggc atccgtggta 92640  
 tgacagtagc cgaccgtgtt ggtatgatag aggttcgcgc aggcgtcgtc ggcctccagc 92700  
 tgacccgagt tgatgtaggc gtaccccagc gccccagaaa cgcggataca gaacaggtga 92760  
 gccaggcga gggccggctt cgagggcgcg ccgagggggc ccgcccggcc tgggcccggc 92820  
 gcccgcttc cccggtcccc cggggcgaag gcgtgcccgc ggcggcgcac gttggaag 92880  
 gcgaaactgg gctcggagtc ggtgatgggg gaaggcggcg gcgagggcgc tacgtcactg 92940  
 gctctctctg ccgtgcggca ctgggcccgc gtgcgggcca ggatgcctt ggccccgaac 93000  
 acaaccggct cggtacactc gacccccgca tcggtcacga agatggggaa cagggacttt 93060  
 tgggtaaaaca cccgtaaac actacagaga cagtgtagcg tgattgcctc gcggtcgtaa 93120  
 cttgggtagc ggcgctgata ttttaaccacc aggtatata tgacattcca caggtccacg 93180  
 gcgatggggg taaagtagcc ctccggggcc cggaggcccc ggcgcttccac cagatggtga 93240  
 gtctgggcaa acttcatcat gccaaacaga cccattccgg cacgattgta ggtgcggata 93300  
 ggtctctcta cagagctgta taggtgtgac ggtccgggac acccaagccc gccgcccctg 93360  
 tgtacagtgg ctgcggcgac gacccccctc caacaagacg ctatcccggg aaaggcacgc 93420  
 tctttataat tcttttttat ttcccatcta cgtgcgggatt ggtgcaaccg ccggcgcgcg 93480  
 ccggtgcagg ccgaccatct ctctcttccc cccctcccc tccccgagc cctcaaagag 93540  
 ggtgtggcct aactagcggg aggcgtatct aaccagacta gggcggcggg tccgcccgtag 93600  
 tccttggtc gggtagccac tgctctgtgg ctgggtccc ccggcccccc taacccccat 93660  
 ccggtcccgc tcattcccgc cctccgcctg cgacacaaac ggccgcgect ccgggcccgg 93720  
 tgacacgacg cgcctcgtct ctgcggattg tcccgggagc gtcgcgcat ggctcatctt 93780  
 cccggcgggt cggcccccgc cccctttctg gaggacgcga tcccgtcgc gcgcgagcgg 93840  
 acggaagact ggcgcctctg ccagatagtg ctgcaggggc ccgagctgaa cgggatcctg 93900  
 caggcctttg cgcgcttcg cacgagcctt ttggaactgc tcctggctct gggcgaccga 93960  
 ggcacccctg tacataacgc gattttcggc gagcaggtgt ttctgcccct cgaccattcg 94020

-continued

---

cagttcagtc gctatcgatg gggcggacc accgcggcgt tccgtctct cgtggaccag 94080  
aagcgatccc tgcctgagcgt gtttcgcgcc aaccagtacc ctgacctgcg gcgggtggag 94140  
ctgacggtca cgggccaggc cccgtttcgc acgctggtgc agcgcataatg gacgaccgcg 94200  
tccgacggag aggcctgga gcttgccagc gagacgctca tgaaacgca gttgacgagc 94260  
ttcgcggtac tactccccc gggcgacccc gacgtccagc tgcgectcac gaagccccag 94320  
ctcacgaagg tggtaaacgc cgtcggggac gagaccgcca aaccaccac gttecgagtc 94380  
ggccccaacg gcaagtttc cgtgtttaac gcgcgcacct gcgtcacctt tgcgcccgc 94440  
gaggagggcg cgtcgtccag caccagcgc caggteccaga ttctgaccag cgcgctgaag 94500  
aaggcgggccc aagcggccc caacgccaag acggtctacg gggaaaacac acaccgcaca 94560  
ttctcgttg tcgtcgacga ctgcagcatg cggcgggtcc tccggcggct ccaggtcggc 94620  
ggggggaacc tcaagtctt cctcacggcc gacgtcccca gcgtgtgtgt caccgccacc 94680  
ggccccaacg cgtgtcggc ggtgtttctt ttaaaacccc agcgggtctg cctgaactgg 94740  
ctcggccgga gcccggttc ctcgaccggg agcttggcgt cccaggactc tggggccggc 94800  
ccgaccgaca gccaggactc ctctccgag ccggacgcgg gcgaccgcgg cgcaccagaa 94860  
gaagaaggcc tcgaggcca gccccgggta ccgcccgct tcccgaacc gccgggaacc 94920  
aagcggaggc accccggggc gaaagtgtc ccgcgggacg acgcccacaa gcgcccgaag 94980  
acggcgtgc ccgccccc caccgagcgc gagtcgccc ccctctcgc gagatacggg 95040  
cccaggcggc cggagggtg ggggacggc ggccgctacg cgtgctactt tcgagacctc 95100  
cagaccggcg acgagagccc cagccccctc tccgcttcc ggggtcccca aagacccca 95160  
tacggcttg ggttgcctc acgcaacgg gtggtggcgg aacgcctcac cgcgcccggg 95220  
cacgcggggt gcgttgtgt aaaaaataa ataatgggg tagtgtgtcc cccccctc 95280  
caaccaatat ggtgtcgtg tgtggtccg ggttgcgct ccgctcttc caccccctt 95340  
ccccctctt tttgttttg cgtgcgctta taagagcggg cccggggccc ttcgagctt 95400  
caccgagagc gccgtcggc cccgggtgcg ggatgtgtcg cggggacagc cccggggtcg 95460  
cgggcgggag cggcgaacac tgcctcggg gggatgatgg ggacgacggg cgcgccgc 95520  
tcgctcgtg ggtgcccac gctcgggggt tcgcgcatct ctggctccag gccgccacgc 95580  
tgggcttctg ggggtctgtc gttctgtcgc gcggcccga tgcggacgcc atgtcggggg 95640  
cgttcgtgat cgggagacc ggcctgggt tctcgcgcg ccccccgcg ttcgcccgc 95700  
cgccgacgc tgtgtcgcg tggctgagc tggtcggcg gggagcggc gtggccctgt 95760  
ggagcctcgg ggaggcggc gcgcctccgg ggttccggg cccggcgacc cagtgcctgg 95820  
cgctcggggc cgcctacgc gcgctgctg tctgcccga cgaagtcac ccccttttc 95880  
tctcgcgcc gcggcccctg tttgtcgca cctgggggt tctcgtcggc gggctgacga 95940  
taggcggcag tgcgcgctac tggatgatc accccgcgc gcgcgggccc ctgacggcgg 96000  
cgggtgtggt gggcctcggg acaaccgccc cggggacag cttttccaag gectgtcccc 96060  
gccaccgccc cttttcgtc gtctcgcgg tcgagctcc ccccccga tacgcccgg 96120  
aggacgcga gcggccaaca gaccacggc cctgttacc gtcgacgcac caccagcgat 96180  
ctccgcggt ctgcggcgc gggcgcgcac gggccgaaa catctgggtt cccgtggtga 96240  
cctttcgggg cgcgctcgc ctggcccct gcgcgcgcg aggtctgac gcggtcctg 96300

-continued

---

caggccccgt cctgccgctg tggccccagg tgtttgtcgg gggccacgcg gcggcggggc 96360  
tgacggagct gtgtcagacc ctccgcccc gggacctcac ggacccgctg ctgtttgctg 96420  
acgtcggatt ccaggctctg aaccacgggc tgatgtttgt ggtccccgac atccgctgat 96480  
acgcgatgct ggggggcgcc gtgtggatct cgtgacgca ggtgcttggg ctccggcgcc 96540  
gccttcacaa ggaccagac gccgggccc gggcgccgc gaccctgctg ggctctttt 96600  
tctccgtcta cgcattgggg tttcgccgcg ggggtgctgt cggcccgctg atggcggcga 96660  
gccggcggtc ggggtgatcg ccatttcaaa taaaaggcac gagttccccg aataccaccg 96720  
gcgtgtgatg atttcgccct accgctccga tccccggggg gaggggggaa ggaatgggg 96780  
cgggggggtg cgtggacggg tataaaggcc aggggggcag cggggcccat cactgttagg 96840  
gtgttaggtt gggaggtggc acaaaaagcg aactcccgt gttgtagtgt tccgcgggag 96900  
cgggtggttt cggcaaaccc tctcgtctgc gccgggcgcg cccaccgctc ctccgcgggg 96960  
gccggggctc ttctggtcat ggccttggga cgggtgggccc tagccgtggg cctgtggggc 97020  
ctgctgtggg tgggtgtggt cgtggtgctg gccaatgcct cccccggacg cagcataacg 97080  
gtgggccccg gggggaacgc gagcaatgcc gccccctccg cgtccccgcg gaacgcctcc 97140  
gccccccgaa ccacaccac gcccccccaa ccccgcaagg cgacgaaaag taaggcctcc 97200  
accgccaac cggccccgccc cccaagacc gggccccgga agacatcctc ggagcccgtg 97260  
cgatgcaacc gccacgacc gctggccccg tacggtctgc gggtgcaaat ccgatcccgg 97320  
tttcccaact ccaccgcgc ggagtccgc ctccagatct ggcgttatgc cagggcgacg 97380  
gacccgaga tcggaacggc gcctagctta gaggaggtga tggtaaacgt gtcggccccg 97440  
cccgggggccc aactggtgta tgacagcgc cccaaccgaa cggaccgca cgtgatctgg 97500  
cggaggggcg ccggcccggg cgcagcccc cggtgtact cggctcgtcg gccgctgggt 97560  
cggcagcggc tcctatcgc agagctgacc ctggagacc agggcatgta ctactgggtg 97620  
tggggccgga cggaccgccc gtccgctac gggacctggg tgcgcttccg cgtgttccgc 97680  
cctccgtcgc tgaccatcca cccccacgcg gtgctggagg gccagccgtt taaggcgacg 97740  
tgcaacggcg ccacctacta cccgggcaac cgcgaggagt tegtctggtt caggacgggt 97800  
cgccgggtgt tcgatccggc ccagatacac acgcagacgc aggagaacc cagcggcttt 97860  
tccaccgtct ccaccgtgac ctcccgggcc gtcggcggcc agggcccccc gcgcaacctc 97920  
acctgccagc tgactggca ccgcgactcc gtgtcgttct ctccggcga cgcagcggc 97980  
acggcatcgg tctgcccgc gccaacctc accatggagt ttaccggcga ccatgcggtc 98040  
tgcaacggcg gctgtgtgcc cgaggggggt acgtttgcct ggttctcggg ggacgactcc 98100  
tcgcccggcg agaaggtggc cgtcgcgtcc cagacatcgt cggggcggcc cggcaccgccc 98160  
acgatccgct ccaccctgac ggtctcgtac gagcagacc agtacatctg ccggtggcg 98220  
ggataccggc acggaattcc ggtcctagag caccacggca gccaccagcc cccgcccgcg 98280  
gacccaccg agcggcaggt gatccggcg gtggaggggg cggggatcgg agtggctgtc 98340  
cttgtcggcg tggttctggc cgggaccgcg gtagtgtacc tcaccacgc ctctcgggtg 98400  
cgctatcgtc ggtgcccgtg actccggggc cgggcccggc cgcgggtgtt ctctctttcc 98460  
accccttccg tccccgtac ccaccacacc ccaccacacc cccccgctg cccccggcg 98520  
ttataagcgc ccgactcgc ttttccacc ggaatcct cggcccgatc cgaacggcgc 98580

---

-continued

---

acgcgcgctg ggctccaaac gcctccggaa gagagcgccc cgccccgata ttcaagcccg 98640  
cgggtggtgct atggetttcc gtgcttcggg acccgccctac cagccccctcg cccccgcggc 98700  
ctccccggcg cgggctcgtg ttccggccgt ggcttgatc ggcgtcggag cgatcgtcgg 98760  
ggcctttgcg ctgctgcccg cgttggttct cgtacccctc cggctcctcgt ggggactctc 98820  
gccgtgagac agcggctggc aggaattcaa cgcgggatgc gtcgctggg accccacccc 98880  
cgtcagcac gagcaggcgg tcggcggctg cagcgcgccc gccaccctta tccccgctgc 98940  
ggccgccaag cacctggccc ctctgacacg cgtccaggcg gagagatcgt cgggttactg 99000  
gtgggtgaac ggagacggca tccggacctg tctgagactc gtcgacagcg tcagtggcat 99060  
cgacgagttt ttcgaggagc tcgcatccg catatgctac taccacgaa gccccggcgg 99120  
gtttgtccgc ttcgttaactt cgatacgtaa cgccttgggg ttgctgtag gcgctcgtcc 99180  
gacggctccc cttctgccc ctctcttccc cccaccccac ccacgacca acgacggcgt 99240  
ttggccaata cctcctttt ttctttttct ctccccccc ccccaaaaa aacaataaac 99300  
agctaattgc gtacgacaaa ccatgcggaa ctgctggtt tttttctctg tttgttactt 99360  
tttattgaaa cagacatacg gggaaagggg ccgaaaccg agcgggtggg gccggcggtc 99420  
gcattttttt aatggctctg gtgctggccc cgtttgagct tcgcaacag ggcgctgagg 99480  
gcggcgacgt tcgctggccc gtcgttggcc agcgcgttgg tccgggggcg ggcgggcatg 99540  
ggcgacaggc ttagtcccgg gtccggggcg cgtgtggccc gccgagggga gaagagggca 99600  
gaccgcccc agtctgacag gggattttcc gcctcagatg acggggagtc cggggcgtct 99660  
ccccgcaggg cggccccgccc ggcaagaagc cggcgagggc agatgttttc gtatacccga 99720  
accagggga tctcctcgtg gacgcgcccc ccacccctcgc ccaccgactc gtaaatggaa 99780  
ctcgcgtcct cggagggggc gcggggggcg tggtttcgg ccggccaggc ggcggcggcg 99840  
gtggtgtcgg cggcgggggt ggcgccaagc ccgacgcccg cgggcatggc ggcgtcctcg 99900  
tcgggcagca gatacgtgtt ttccatctgg tccggttcgg cctccgctgc cggccccag 99960  
gtccgaccc cgtcgtagac cccggcggcc tcgctgtag ccgagagcgg gcgctcggcg 100020  
gctgcccggc gctgctcggg gggcgcgggg ttgctggggg ggagggcggg gggcgcggcg 100080  
gccatagcgt tgaatacgt ggccggccgg ccggcgcagg gctcgggacc ccggtcggcc 100140  
gcgtcagcgt gcgggggctc ggggaggtcc tcgctggtgc gcctgcacct ccgaggggccc 100200  
gcgggggtcg agtggggggc agccccgggg agcggcgggg gtgctgtgtc gcgcccggtc 100260  
cgttgatct tgtcccggca gctcccggc accgcgcccg ggccccccg tgggcccggc 100320  
gccgcgaggc gcaggatgga ctcgtagtgg ggcgacgggg ttccgctccg aagcaggtcc 100380  
ggggccaggc cggccccgaa ccaggacttg atgctgagtt ccacccgggc ccagctcggg 100440  
gcggtcctcg tgggaaacag gggggcggcg gtctcgcaga agcgtcctg gctgtccacc 100500  
gccgccgtaa ggtactcgtt gttcagctg tcggaggccc agacgacata cccggttaagc 100560  
gtcgcgttaa ttatatactg ggcgtggtgg tggactatgg atagaacctc gacggctcag 100620  
acgatggcgt ccacgatccc gtacgtgccc ccgctgctgt tgcgggtctc ccacaggtgg 100680  
gccaggcgcg tcagggtggc caggacgtcg ctgaccgccc cccgagggc catgactcgc 100740  
atcgagcccg tggctcggct gggcccggcg tccaggtggc gcgcaaacgt ctccgagggc 100800  
gcctccagac tcccgtgag gcgccgaac cggcgtcggc cggggcccag gcggcgacac 100860

-continued

---

acgtacttgt ccgcccgtcca cagcatccac gagggcccaat ggtacaacac ggagacgtag 100920  
gccaggagct cgtcagcccg cagtgcgggtg tccgtgctcg gccggctcgg gtctgcgggg 100980  
cgcataaaga acatgtactg ctggagcctg tgggcccgt cgcgcaaccc cgcacccgcg 101040  
gcgcgctact tggcccgccg gcccccgtc ttgaacgggg cgcgcaccag cagcttcggg 101100  
agcagggtgg gcccgagcag cacgtgcagg ctggggctcg agtcgcccgc cgggtcgtcg 101160  
gggatgtcca ggcgctggg cagcaccgtc tggaggtact tccagtactg cgtaggatg 101220  
gcgcgctca gctggccgcc cgcagctcc acctcgcga gcgcctgctt ggcggccgac 101280  
gcgtagtgcc gtagtagtc ttagtgccgg tcgctggcga gcccgctac gatcaggctc 101340  
tcggggacgg tgttatgggt ccgcccgcgc agccggacgc tgcgatcggc gccggtcaga 101400  
aaccccggct gcaggtcgtc ggcgcgctgc cgcaggacgc ccaaggccgc gctgaggagc 101460  
ccctccgggg tggggagcag acaccccggc aagatgcgcc gctcggggac gcccgcggtg 101520  
gcgccgaggc tgaggttggc cggcgtcagg caccgcccga gccgcaggga gctcgcgccg 101580  
cgcgcccggc gttgatcggc ggagaccgtt cggtcggggg ccccgccggt cggaggtatg 101640  
ccgctcccgg gatatagggt ttgcttttta tggggaggcg cctatggggc tggcggggcg 101700  
cccagcccgg tcgcccgcct cccggacacg tgcgcccgga gggcggcggc ctectcgtcg 101760  
cccagcagcc gtttccgaaa ctgcccagc atgtccacga cgcggaccgc cggcccagc 101820  
acggactcgc tattcagggg ggcggggggg aaggcccga ggtcttcgag caggaaggcg 101880  
gggtctgcgc tcccgtcacc gggcgcggcg ggcgcgagg acgcccggcg aaggtccacg 101940  
tgttccgccc cggcgcgccc gtcccacaa aatttgccgg ggggtggtccg cgcgtacagg 102000  
ggctgggtcg cgcggaggac gcacgcgtag cgcagggggg tgtacgtgcc cactcgggg 102060  
gccgtcgacc cgcgctcaaa cgcggccagg gccacgcacg cgcaccaccg gtcggccagg 102120  
cccagcagcc gctgcaggat gagccccgc gccagcacgg cgcgcccggc cgcgcccgtc 102180  
tccctgccc gccgcccgtc cccgagccc agggcgatt tccaggtaac ggtcggccagg 102240  
gccgtgtgca ggcgctacac ggcgcccgc agcacggcgt tcagcccgt ggtggcgagc 102300  
aggcggcgcg ccgcccgtgc gcccccgcgc tctgtctcgg cgcgcccagc cccggggctg 102360  
cccaggggca gggcgcgaaa cagcgcctcc tgctccacgt ccgcaaacgc ggggtggggc 102420  
gagtgcgggt gcaggcgcgc ccccacgacc accgagagcc actggaccgt ctgctcccgc 102480  
aggaccgcca gcacgtccag gacgcccgc gcaaacgcgg cctcccgcgg gagcacgcat 102540  
ttgacggcgc cggggtttaa gcgggagc agagcccggc tggcgatgta cgtcatgccc 102600  
cccgcgtagc gggcggccc gcgacagtcg cgcggcaggc gcgcccagc cccggggcag 102660  
tacagcaggg acccagcga actgcgaaa acccggcgt cggggccggg gtgggggggc 102720  
gcgcccctc ccgcccgtg cagcggcagc gcggcggccc ccaaggccgc caaccccgtg 102780  
aggctcgcga actcccgtcg gagctcggc gccctgtcgt cgagctccga gccgcccgc 102840  
tccgtgtgca ggcgctccc ccagaccac ccggtgatc ccacccgcac gatggcgtcc 102900  
accagaaaac ccatcgcgc ggaggggctg gttttgccc gccgatccgt caggtcgagg 102960  
atcgcgtcgc ccgtgacgta ccagggcagc gcctcgcct gctgcagcgt ctggcggaaa 103020  
aacaccttg ggtcggccc ggaggcaaa tgcatgacc ccacgcccga cagcccgaac 103080  
gcgctatccg gacacgggta gaaccccggc ggatgtccca gggccagggc cgcgcccagc 103140

---

-continued

---

gactcgtecc acgcggcgac tcggggggtc aggcgggtcca gggggaatgc cgectgcagc 103200  
tccgggcccc acacggggcc cgcgagaatc tcgaccgtcg cggaaaggcc cgccccgggg 103260  
ccgtcatcgt gcgcgacggc ggcggggtag tcgtcctcct cgtagttgag ctctgccagg 103320  
aacagcggcg agggcaccac ccgcgaaccg cccaccgcgc ccaaaacgtc gcgtgggtcc 103380  
atcgggcccc ggtagcctcc ccgcggggcc cgcgtgatgg cgtgtcccc gcgtccgcga 103440  
acggactggc tctggcctg aacggacctg gggcgcgaa aggacgccc gcgggggggc 103500  
gcccgcgcc gggcctcgga cgcgcgtcgg gaccgggggt gaccgggggc ctcccggcga 103560  
cggcgcgggg gcggctcttc gctcgcctc tccccgcgg cctcgacctc gctgtcgtcg 103620  
tccacgttaa acaccgccc caggtacccc attaacccga ctccaccgcc ctccgggctcg 103680  
tctccaacgg gcgagtcgga gcgatgcgga gacggggcat gggaccgggt ggaggcgcgc 103740  
ctccggccta cggcatgccc gcgcacggac atggtggcgg gaggcccgat tttttacaca 103800  
cgccctcccc gcagacggac gaggaaagg gtggtgagag gggggaggcc caaacgggga 103860  
ggtggggggg agggggcggt cccagggagc ggggggtagg aaccggcacg acgggaacag 103920  
agaaaacgca accgctccaa caagggtggg ggggtgggcc tcgtccccc gcagaccgc 103980  
gggcaaatgc gagaacggga cccgcgcgc tcctttata cgcggacccc agcaccacga 104040  
gccgttctgt gacgcgaatc tacacgaccg cgggctcgta ggcgcgacta acgccaacc 104100  
caacggcaca cccccccac cccgcgcgta acccatttc tttcatggtc ccgtaataaa 104160  
cagccaacgc acgcccgcga tgatgagttg cttgccaatg tttattgctg tggttgcgaa 104220  
ccctctatcg cgatacagac ggaggtgagg cggggcgggt gtgggggggg ggcgcgccgc 104280  
ccggtcgcac atcctacccc ccaaagtcgt caatgcccat ggcatcggta aacatctgtt 104340  
caaaactcaa atcgtccacg tccaaagccc catacagac ggggtcgtgg gtcattccc 104400  
gggaggggga ctccacgtcc cccagcatct ccaagtcgaa gtcgtccagg gcgtcggcgg 104460  
gcgtcatatc cacctcctcg ccgtccaggc ggagttcgtc tcccaggctg acgtcggtaa 104520  
tggggggcgg ggtggacagt ctgcgggggc gttgtcccgc ggagagaaac gacatgcgcg 104580  
gcgccaccag cccggcctcc gcaggagcgt catcgtcgtc cgggaggtcg agcaggccct 104640  
cgattgtcga tccgtaattg tttctggtcc gcccgcggtc atacgcgtgc tcccgcata 104700  
cggactcgcc ctccgaggtc gcgacgctgg agtacgagtc caacttggcc cggatcagca 104760  
gcataaaagta cccagaggag cgggcctggt tgccctgcag gacgggcggg gtcgtgagg 104820  
gcgccccggg ttcctcgcgc gccgcacttc gcaccagcgg gaggttcagg tgctcgcgaa 104880  
tgtggttag ctcccgcagt cgcggggcct ccacgggaac tccccgcacg gtgagcgatc 104940  
cgttgataaa catcaggggc tgaaacagac acgccaactg gcgccagctc tccaggtcgc 105000  
agcagaggcc gtggaacaga tcgggcccga tcatctgctc ggcgtacgcg gcccatagga 105060  
tctcgcggct cagaaagagg tatagatgca gaaacaggac gcgcgccagg cgcgcggtct 105120  
cgcggtagta cctgtccgcg atcgtggtgc gcagcatctc ccgaggtcg cggttgcggc 105180  
cccgcagtgt tgccctggcg tgtagctgcc gaacgctggc gcgcaggtag cggtagcagg 105240  
ccgagcaaaa atttgccaac acggtccggg agctctcctc ccgcgccgc agctaccgc 105300  
ggaaaaactg cgcctaggcc tcgtagtacg aaggcagctc gtcgcggtg gcgggcagg 105360  
tggggaacgc cacgtcgcgg tgggcgcgaa tgtcgatcgg ggagcgtcg gggacgtcgc 105420

-continued

---

catccccca gtcgatcaag tcgctgggca gcgctgacag aaacttgca tcccggatca 105480  
tgtcggcggt ggtcgggaac ccagagaaca ggtcctcgtt ccaggtatct agcatgggtac 105540  
acagcgcggg acccgcgctg aagcccagat cgtcagaggag acggttaaac agggccgcgg 105600  
gggggacggg catggcgggc gagggcatca gctgggcctg actcagccga ccggtggcgt 105660  
acagcggagg ggcggctggg gtgttcttgg gaccccggc tggcctgggg ggcggtgggc 105720  
aaacccgctc cgcgtccgca aacagatcgt cgaccaacag gtccatgggg gcggttgggt 105780  
ccgggaataa cgatctcgag agggcaatga gacgtgcccg agcggccggc ggcggagagg 105840  
gggggagggg tccgggacct gcgacagaaa aaggccgggg ccctcgcgaa gggaaatccc 105900  
gggggtgccc tgcgtccccg aggactgaca tctcgcgtcc accaccccgc atttaagtat 105960  
caccacagtg ccccccacaa cctcgtgact tccccaccgc tccgggcggc ccgtcccccg 106020  
cgctcggaag ggaggcgtgt ccttctctcc gccctcccgc cccctcccgc cctcccgc 106080  
cctcccgc ctcccgcct tcccgcctc cccgcctc cgcctccc cgcctccc 106140  
gccctccc cccctccc ccctcggcac aaacgcgtgc tgacagcga gtggttaaat 106200  
cgaccgtgat gctttattgt ctgtcgtctg aacgcggctg gggctcgtac tcgagggggc 106260  
ggcggggacg ggaagccgag cgggcggggg cccgtcgggt cgcggcgca cgcggcgcg 106320  
ggcggccccc ggccggccggt gtcgcgtcga cgtcctgcgc cgcgtcggga ttcaccaact 106380  
cgttcgcgcg ctcgagggag ttcttgcct cgcagaccgt caccgcaatg gtggtgaggt 106440  
cgaggagctc gttgaggtct tcgtcgggtg gcggccgcca catgtcccac agctgtaccg 106500  
ccgccagccc ggcgtgctgt gcccccaggc gcccgaccgc ggcgcagaag acgcgcttgt 106560  
tgaacccggc caccgggggg gtccacggcg cctgggggct cgggtggggc gtgctgaagt 106620  
gcagcttctt ggcagctcc tgggcgggtg tcttggttct tcccaggcc gtgggagcgg 106680  
ggcgtctag gagcacggcg gaggcggcct gggcgggtcg cctcccgcgg gcggggtcgg 106740  
tcgcccgggt cgcggaggcc ttagggccc cgcgctcat tttgggggtc cgcgggggag 106800  
ggcgtcgcga gcgcccgcgg gcgcccacgg ggcgggggg ggggtggagg gcgcgcgcg 106860  
ggcggggggt gtgagagccc gcgacggagc ccgaaagcgc cggctcgcgg gtatcccggg 106920  
actcgtcgtc gtccgaagac gaggcccggt agagggcata cccagcctcg tcataatgga 106980  
gaaagcgaac ctcgcccctc gggcgcgcgc gcacggggcc agcggccggc cggaaatcgt 107040  
cgcgaggact cctcgggtcc gccggggaga ccgggacata gtacagctcc tcgtgggtcc 107100  
cgcgcgcgcc ttcggcgga cagcacttga cggagcggcg agaggcatg gtctatcgga 107160  
gacaccgggg acgcccgtgc ggatcacagg gaaggcgtcg gcgaaaggag cagagagcgt 107220  
cggaaaggcg cgagggaggg aaagaggag accggcgggg tacgggagag cagcgagggc 107280  
ctgcgtaacc cagggggggt gcgggagtg ctcctcgcgg gttcgggggg agagtttata 107340  
ggaagtggat ataaccgag gcgacgggac taaccaatcc cggggggggc aacggacaga 107400  
cacgccccga acaggcccga ctccgcgag gaagcaaagg ccggggggcc cccaacgaca 107460  
cgcccacccc tcccacacag ggcgggctca ggctgaccgc gcggccagtg cccgctgaca 107520  
tatctgatac acgtgcgca tcatacata gccatcgag gtcctgcta gataaaagg 107580  
caccaggacc cccgggagcg acaccacacc ggcgctgctg ccccgccatt gcgctcccc 107640  
gataacgccc cgtgcgcctg ccgcttccg cggctcccc ggcaagccc gcagagcgc 107700

---

-continued

---

gacgaacaac agcaccaccc agcggcccag tcttgccggg ttccccgtca tcgccggcgat 107760  
gagtcagtg gggcccagg cgatccttgt ccagacggac agcaccaacc ggaatgccga 107820  
tggggactgg caagcggccc tagctattcg cgggggcccga gtcgttcaac tgaacatggg 107880  
caacaaaacg gccgtggatt ttaccccggc agaatgccgg gactocgaat gggccgtggg 107940  
ccgcgtctct ctgggacctg gaatggcaat gccgcgggac ttctgcgca ttattcacgc 108000  
ccccgggta tccggcccc ggccccacgt gatgctcggg ctgcctgact cgggctaccg 108060  
cggaaaccgc ctggccctgg tcgtagcccc gaacgggacg cgcgggttg cccccggggc 108120  
cctccgggtc gacgtgacgt ttctggacat ccgggcccacc ccccgacc tcaccgagcc 108180  
gagctccctg caccggtttc cgcagttggc gccgtcccc ctggcagggt tacgagaaga 108240  
tccttggttg gacggggcgc tcgcgaccgc cggggggcgc gtggccctgc cggccagacg 108300  
gcccggggga tcgctggtct acgcgggcca gctaacgcag gtgaccaccg agcacggcga 108360  
ctgcgtgcac gaggcgccc cctttctgcc aaagcgcgag gaggacgcag gctttgacat 108420  
tctcatccac cgagccgtga ccgtccccgc caacggcgc accggtcacc agccgtccct 108480  
ccgcgtattg cgcgcggccg accgaccaga ggctctgat gtgctggggc ggtcgtcgt 108540  
caatgccagg ggccctcctg tcatgcctac gcgctggccc tccgggcacg cctgtgcgtt 108600  
tgttgatgt aacctgaccg gaggccccgt gaccctacaa gccgggtcca aggtcgcca 108660  
gctgctcgtc cgggggacc accgccccct ctggatcccc cccgacaaca tccacgagga 108720  
cggcgcattc cgggctacc ccagaggggt tccggacgcg accgccacc cccgagacc 108780  
gccgattttg gtgtttacga acgagtttga cgcggacgc cccccagca agcggggggc 108840  
cgggggggtt ggctccactg gcatctagac cgcgcctcgc gtcgggcccag atggggcccc 108900  
ggtaataaaa gagctctgtt tcgcatatgc cctgggtgtg cgggtttttt tttgtgtct 108960  
gtctgcccgg cgtcgggtt tccgttctgt cgtcgtatc acatacgcac aaacacacgg 109020  
gtagagtga accgaaaccg gtcgacgttt attcaccaca cagaacaca agctaagcga 109080  
gaaggagggg ggccctcgtc gacgaggcct gccgtttggg ggcggacgtg cgtgacgtg 109140  
ggtcgggtgt agggctcggg gggggcaccg gccggggcgc aacgggggat ctgctgcccg 109200  
cgtgggtgac tgggaccgac gcaacctccg gggcttctgc cctcgtaggc cggggggggg 109260  
cctcggctgc tccaagcccc gcggtgcggg tccctccggc cagagccgag gtcgagagac 109320  
caagggccc ctcgcgcat gccacgtcct ccatgaccac gtcgctctcg gccatgctc 109380  
gaatggcctg ggagacgagc acgtccgcg acttgtccgc ggccccacc gacatgtaca 109440  
tctgcaggat ggtggccatg cactgtctcg ccaggcggcg catctgtcc cgtgcccgg 109500  
caacggcccc gtcgatggtg gagccctcga gtcccgggtg gtggcgcgcc agcctctcga 109560  
ggttgacct gcaggcgtg tatgtcggg ccaggcgcgc cgccttcacg aggcgcccgg 109620  
tgtcgtccag cgactctagg gcgtctcga gcgtgatggg ggcgggcaaa agcgcattga 109680  
ccaccgccag ggccctcctg agccgcggct ccgcctcga gggcgagcc gggccccgaa 109740  
tcctctcata ttgtgttcc tcggggcgcg ttccccacc gcacagacc ccgagcaggg 109800  
accatcccc ggaacacgc gcggtctg cgcggctt ccccccccc acccctccg 109860  
ggttcgcagg ggcgatggg accggaagact gcgatcaca agggcggtc gttgcccgtc 109920  
ccgtggaggg tacggcgtg tatgcagcc accgggtcgt taccacctc tcgctcggc 109980

---

-continued

---

tcctcacaaa ctgcctgctg ggggccgagc cgttgatat attcagctac gacgcgtacc 110040  
ggtcogatgc gcccattggc cccacgggcg cccccaccga acaggagagg ttcgagggga 110100  
gccggggcct ctaccgggat cggggggggc taaatggcga ttcatttcgg gtgacctttt 110160  
gtttattggg gacggaagtg ggcgtgaccc accaccgaa agggcgcacc cggcccatgt 110220  
ttgtgtgccc cttcgagcga ggcgacgacg tcgccgtgct ccaagacgcc ctgggcccgcg 110280  
ggacccatt gctcccggcc cagctcacag caactctgga cttggaggcg acgtttgccc 110340  
tccacgctaa catcatcatg gctctcaccg tggccatcgt ccacaacgcc cccgcccga 110400  
tcggcagcgg cagcacccgc cccctgatg agcccggcga atcgatgcgc tcggctgctg 110460  
ggcgcgatgc cctggggcag cgcggcctca ccacgctgtt cgtgcaccac gaggcgcgcg 110520  
tgctgggggc gtaccggcgg cgtattatg ggagcgcga aagcccctt tggtttctga 110580  
gcaaatcgg cccggcagaa aagagcctgg tgctggcgc taggtactac ctactccagg 110640  
ctccgcctt ggggggcgccc ggagccacgt acgatctgca ggccgtgaaa gacatctgcg 110700  
cgacctacgc aatccccccac gaccacgccc ccgacaccct cagtgcgcgc tccttgacct 110760  
cgttcgcgcg catcactcgg ttctgttgca cgagccagta cccccgcggg gccgcggcgcg 110820  
ctgggtttcc gctgtatgtg gaggcgcga tcgccgcga cgtacgcgag accggcgcgc 110880  
tggagaagt catcgcccac gatcgacgt gctgcgcgt gtcggaccgg gaattcatta 110940  
cgtacateta cctggcccac tttgagtget tcagccccc gcgcctggcc acgcatctcc 111000  
gggcccgtgac caccacgac cccagccccg cggccagcac ggagcagccc tcgcccctgg 111060  
gtcgggaggc ggtggaacag ttctccggc acgtgcgcgc ccagctgaac atccgcgagt 111120  
acgtaaagca aaacgtcacc cccagggaaa ccgccctggc gggagacgcg gccgcgcct 111180  
acctgcgcgc gcgcacgtat gccccggcgg cctcaccgcc cgcccccgcg tactgcgggg 111240  
tcgcagactc gtccacaaa atgatgggac gtctggcgga agcagaaagg ctctagtc 111300  
cccacggctg gcccgcttc gcaccaaaa cccccgggga cgacgcgggg ggcggcactg 111360  
ccgcccccca gacctgcgga atcgtcaagc gcctcctcaa gctggccgcc acggagcagc 111420  
agggcacgac gccccggcgg atcgcggctc tcatgcagga cgcgtcggtc caaaccccc 111480  
tgcccgtgta caggattacc atgtcccga ccggccaggc gtttgccgcg gcgcccgggg 111540  
acgactgggc ccgcgtgacg cgggacgcgc gcccgccgga agcagaccgtg gtcgcccagc 111600  
cggcggcggc gcccgagccc ggcgcgctcg gccggcggct cagcgcgcgc atttgccc 111660  
ggggccccgc gctccccccg ggcggcctgg ccgtcggggg ccagatgtac gtgaaccgca 111720  
acgagatctt caacgcgcgc ctggcctta cgaacatcat cctggatctg gacatcgccc 111780  
tgaaggagcc cgtccccttt ccccgctcc acgagccct gggtcacttt aggcgcgggg 111840  
cgtggcggc ggttcagctg ttgtttccc cggcccggct agaccccgac gctatccct 111900  
gttattttt caaaagcgc tgctggcccc gcgcgcgcgc cgtctgtgcg ggcgacgggc 111960  
cctcggcggc tggcgcagc ggcgacgggg actggttccc cgacgcgggt ggtcccggcg 112020  
acgaggagtg ggaggaggac acggaccccc tggacacgac ccacggcccc ctcccggacg 112080  
acgaggccgc gtacctcag ctgctacacg aacagatacc agcggcgcgc cccagcgaac 112140  
cggactccgt cgtgtgttcc tgccgcgaca agatcgggct gcgcgtgtgc ctaccggctc 112200  
ccgccccgta cgttgtgacg ggctccctga cgatgcgtgg ggtggcagag gtgatccagc 112260

---

-continued

---

aggcgggtgct gttggaccgc gacttcgtgg aggccgtagg gagccacgta aagaactttt 112320  
tgctgatcga tacgggcggtg tacgcccacg gccacagcct gcgcttgccg tatttcgcca 112380  
agatcggccc cgacggctcc gcgtgcggcc ggttattgcc cgtettctgt atccccccg 112440  
cgtgcgagga cgttcggggtg ttctctgccc cgcacgccga cccgcggcgc ttcaacttcc 112500  
acgccccgcc catgttttcc gcggccccgc gggagatccg cgtcctccac agcctgggcg 112560  
gggactatgt cagcttttcc gagaagaagg cgtcgcgcaa cgccttgag cactttgggc 112620  
gacgcgagac cctgacggag gttctgggccc gctacgatgt gcggccccgac gccggggaga 112680  
ccgtggaggg gttcgcgctca gaactgtgg ggcgaatagt cgcgtgcac gaggctcact 112740  
ttccccgagca cgcgcgggaa taccaggccc tgtccgttcc cggggccgtc attaaggacg 112800  
actgggtcct gctgcagctg atccccggcc gcgggcgccct gaaccaaagc ctctcgtgtc 112860  
tgcgcttcaa gcacggcagg gcaagtccgc cgcacggccc gacctttctc gcgctgagcg 112920  
tcgggaccaa caaccgccta tgcgcgtccc tgtgtcagca gtgctttgcc actaaatgcg 112980  
ataacaaccg cctgcacacg ctgtttaccg tcgatcgggg caccatgc tcgcggtccg 113040  
ctccctccag cacctcacga ccgtcatctt cataacggcc tacggcctcg tgctcgcgtg 113100  
gtacatcgtc tttggtgcca gtccgcctca ccgatgtatt tacgcggtgc gccccgcccg 113160  
ggcgcacaac gataccgccc tcgtgtggat gaagataaac cagacgctgt tgtttctggg 113220  
cccgcgacc gcccccccg gcggggcatg gacccccac gccccgctct gctacgcaa 113280  
tatcatcgaa ggtcgggccc tgctccctcc ggccatccc ggcccatga gccgcgggt 113340  
catgaacgtg cacgagggcg taaactgctt ggaggccctc tgggacacc agatgcccct 113400  
ggtggctgct ggttggttcc tgtatctagc gttcgtccc cttaccaac gacgatgcat 113460  
gttcggcgtc gtgagtcgg cgcacagcat ggtggccccg gcgacctatc ttttgaacta 113520  
cgccggccgc atagtgtcga gcgtgttctt gcaatacccc tacacgaaa tcaccgcct 113580  
cctctgcgag ctatccgttc aacgccagac cctggtgcag ctgttcgagg cggatccggt 113640  
caccttcttg taccaccgcc cggccattgg cgtcatcgtg ggctgcgagc tgetgetccg 113700  
cttcgtggcc ctcggtctca tcgtcggcac cgtctctatc tcccggggcg cctgcgcat 113760  
cacacacccc ctgtttctaa caatcaccac ctggtgttcc gtgtccatca tcgccctgac 113820  
ggagctgtat ttcactctgc ggcggggctc ggccccaaa aacgcggaac cagcggcccc 113880  
cagggggcgc tccaaagggt ggtcgggctc ctgcgggccc tgetgttcca tcatcctctc 113940  
cggtatcgcc gtgcgcctgt gctatatcgc cgtcgtggcc ggggtggtgc tcgtggcct 114000  
tcgctacgaa caggagatcc agcgggcctt gtttgatctg tgacgtaacg cctcttccgt 114060  
tggaagaggc ggaccagtc gccatacaa attaaataca cgaccgcct cgggcctacg 114120  
caccctcgca cgtcgcctgc aaattaaaat cgtgcacaga gccgatccgg cctcgggtct 114180  
gcttgcctc cccccggccc agcacaggca ggctcgtccc acttcgcat acaccacc 114240  
ctaccgctg cttccgcacc cccgcctacg cgtgtacgcg aaggcggacc cagacctgcc 114300  
gtatgctaat taaatacata aaaccacccc tcggtgtccc attggttctt ggggacggcg 114360  
ggggcggggg cggtagcgc cgcacggggg ggacaaggag gagtttcgga aagccggccc 114420  
cggctcgtgc ggtataaggc cagccaccgg cccactgggc gctgtgtgct gccgtgtgcc 114480  
gaccccggtt gcgcgtcggg gccctcctc gattcggacc cggccactct cttccgacac 114540

---

-continued

---

gcgccccctc ggaggacacc cgccatccca gccccggcga cctacaacat ggctaccgac 114600  
attgatatgc taatcgacct aggattggac ctgtccgaca gcgagctcga ggaggacgct 114660  
ctggagcggg acgaggaggg ccgcccggac gaccccagat ccgacagcag cggggagtggt 114720  
tcctcgtcgg acgaggacat ggaagacccc tgcggagacg gaggggcgga ggccatcgcac 114780  
gcgcgattc ccaaaggctc cccggcccgc cccgaggacg ccggcacccc cgaagcctcg 114840  
acgcctcgcc cggcagcgcg gcggggagcc gacgatccgc caccgcgac caccggcgtg 114900  
tggtcgcgcc tcgggaccag cgggtcggct tccccccggg aaccgcacgg ggggaaggtg 114960  
gcccgcaccc aacccccgtc gaccaaggca ccgcatcccc gaggcgggcg gcgaggtcgc 115020  
cgccggggcc ggggtcgata cggccccggc ggcgccgact ccacacaaa accccgcccg 115080  
cgcgtctcca gaaacgcccc caaccaaggg ggtcgccacc ccgctcggc gcggacggac 115140  
ggccccggcg ccaccacgcg cgaggcgcgg cgcggagggg agcagctcga cgtctccggg 115200  
ggcccgcggc cacgaggcac gcgccaggcc cccccctcgc tgatggcgt gtcctgacc 115260  
cccccgcacg cggacggcgc gcccccggtc ccggagcgaagg aggcgcctc tgcgacacc 115320  
atcgacccc cgttcggggc ggttctgcga tccatccg agcgcgcggc ggtcgagcgc 115380  
atcagcgaag gctttggacg cagtgcctg gtcacgcaag accccttgg cgggatgccc 115440  
tttcccgcg cgaacagccc ctgggctccc gtgctggcca cccaagcggg ggggtttgac 115500  
gccgagacc gtcgggtttc ctgggaaacc ctggtcgtc acgcccggag cctctaccgc 115560  
acattcgcag ccaaccgcg ggcgcgctc acagccaagg ccatgcgcga ctgctgctg 115620  
cgccaggaag atctcatcga ggcctcggc tccgcggatg agacgctggc gtggtgcaag 115680  
atgtgcattc accacaatct cccgctccgc ccccaggacc ctatcatcgg aacggcggcc 115740  
gcccgtgctg aaaacctcgc cagcgcctg cgcctcttc tgcagtgcta cctgaaggcc 115800  
cgaggcctgt gcgggctgga cgacctgtgc tcgcccgcgc gcctgtcggc cattaaggat 115860  
attgctcctc ttgtgttggc catcctggcc cgcctcgcca acccgcgcga gcggcggctg 115920  
tcggagatcg actacacgac cgtggggggt ggggcccggc agacgatgca cttttacatc 115980  
ccgggggccc gcatggcggg tctcattgaa atactggaca cgcaccgcca ggagtgctcc 116040  
agtcgcgtgt gcgagctgac ggcagtcac actatcgcgc ccttatatgt gcacggcaaa 116100  
tacttctact gcaactcctc attttaggca agaataaaca tattgacgtc aacccaagtg 116160  
gttcctgtgt atgttcttgg cgcgcgcggc ggggtggggc gagactccgg gcgcatgccc 116220  
gcgtgcgcgt gggaggaggg cgatgaccca ccggataaat gtggggcccc ggcccggccc 116280  
gcttcacagc gcgtccagga actcacggca gacgcgtatt caccgacccc cccctcgcaa 116340  
catgacaacg acgcccctct cgaacctggt ttacgggccc ccggacatca cccacgctgc 116400  
cccccgctac tgtctgaatg ccacgtggca ggcccgaaac gccctgcaca cgacaaaaac 116460  
ggaccccgcg tcctcggcgc cgcggagtta tttagtccc gcctcctgct cgaccagcgg 116520  
ccccatccac tgtttttctc ttgcccgtga caaggactcg cagcactccc ttcgctggt 116580  
taccgagctc cgcaactcgc cggacctggt caaccaccgc cccgtcttgc gcgaactaga 116640  
ggataagcgt ggggggcccg tcgcggtcac gggcccatc agctgcggaa ccatcaagga 116700  
cgtctccggt gcatccccgc cgggggaata cacgataaac ggtatcgtgt accactgtca 116760  
ctgtcgtgat ccgttctcca aaacctgctg gctcggggca tcccggccc tacaacacct 116820

-continued

---

tcgctctata agctcaagcg gcacggccgc tcgcgcgga gaacagcgac gccacaaaat 116880  
caaaatcaaa atcaaggtat aaccaccccc cttccctccg agtccgtatg caacctcatt 116940  
aataaagagt gagaaccaac caaaacagac gcggtgtgag tttgtgggtt ataggaaccc 117000  
ggtaaatacc acgcgacgaa ccagcatgtg tgttaacgca acttttattc gttgtatcgc 117060  
gggagggggg aagcttaccg ccaaaggaag gccaaatga taacgacgac caccgcgacc 117120  
acccccaaaa ccgcatgacg acacgtcccg ccacaccacc ctggggcttg gggcgtgctg 117180  
gagctcgacg cacagcgggc cgcgcgttg gcccggtaca gctctcgcga attgacaagc 117240  
gggggtcgcc acgtgcgcga gctttgcacg cggggttggt cggccggccc caccgacccg 117300  
cccgggtggc cgtcggaca tgcggccatg accatggcgt aggtgggggg gcgatccgag 117360  
gtcgcctctg cgtaaagtag gaggcccgac gggaggtcgc ctcccacgcc aggggtggcc 117420  
ccaatcatag tttccggtag aaacaggggg gtctccacaa acaaccccc tgggccaaaag 117480  
ctccggcgcc gcgcccgtcg ttcggcgcgg cgctggcgc gccgagcggc ccgccaggcg 117540  
gcgcgggcgc agcggccacg ctcacacacc tcgcccgtcac cggaagaagc cggtgaaca 117600  
agcccaaccg gcgacgtccc tgcagagtac ggtggaggcg agtccgtggg ggtgtcgata 117660  
tcaataacga caaactggcc cgcgctcgcg ccggccacac tctcgatgg gggcggggcg 117720  
tcaatcacgc tatcatctcc gtcatccctg catgcgtggg catgccagc ccccaacgcc 117780  
atggtgggga ttcgcgctc agaagcctgc atgtcgtgtg gtcggtcgta gtccaacgtg 117840  
cctcccccc ccaccacaca gccgggtccc acgcccagca ctagaccgca gacgtcgccc 117900  
aacggaggtc cccgtgcaca gaccgcgctc tttatagccc caggggttgc taattaacgc 117960  
acgcatcgag acgcaattta ttttgcctcc ccgcgtcctc ccctcccctg cgcacacgtg 118020  
ataggtcttg ggaaccgag gggcgacgcg gggaaagcgc gcccccgccc ggcgcgcgcg 118080  
cgcccccgcc cggccgcccgc gcgccccgc ccggcccgcc cgcgcccccg cccggccgcc 118140  
gcgccccccc gcccgcccgc cgcgcccccc cgcggggccc ccgcgcccc ccgccggccc 118200  
gccccgctc cgcggcgccc ccctcccggc gcttcggggg tcttctctc cttcccgcgc 118260  
gcgacccga ccccgcccca ccgccccgc cggcaggggg gccccggcgc cgcgcagaa 118320  
acacagacga acacacggtg gcgatctttt ctttacttcg gcggaccagc gagccccgpc 118380  
cccggccccg gcccccgcgc cacaccacag gcaccccccc ccgcccacca ccccggggtc 118440  
cacacaggag cgcgcggggc gcagaaaacg gggcgcggcg gcggtcgggg tgggagtgg 118500  
ggtgggggac acgaaaacac acccacgaca ctctcccccc accccgaccg ccgcccgcgc 118560  
ccaccggcgg gatcggggcg agacgcagcc gggccccccc ccaccacccg cccaccacc 118620  
tacccccgcg ccgcagcctc cggcagcagc ccgaccaccg ccgccacccc ccaaacagcc 118680  
aaggcgcggt gggggcggtg gtggtgaacg atggggggaa cacggggggg aggggtcccg 118740  
ggcgaggcgg gcggggcgaag gaaggggggg tgggtggcgg gccggtggaa agcggaaaa 118800  
cggaggatgg aagggcagaa gatggggagt cccgatcctc ctcctgcac cctcgcctt 118860  
ccattctccg gccctcgcg agtcccagc ccccccccc gccgcccgac gaaggagacc 118920  
caagcaccgc agcggagag gccgagcgg gagtgggcg cggggcggga ggatggcgga 118980  
gagagagaga gagagagaga gagggggggg gggggagagg gaaagcaac ggaaagagag 119040  
gcgcgcgga aagcagcaag aggggggacg gggcgagccg ggcagagtgc ggagccccg 119100

---

-continued

---

gagcccgcgg ccgcagcccga gcagcgcgcg gggctccggg gccgggcccgg gccggcaacg 119160  
ccccgcgcgg gccgcgggggg agagaacccc tgtgtcattg tttacgtggc cgcggggccag 119220  
cagaaggggc gcggggccagc agaaggggcg cggcgccagc ggcccaagcc tcccgcgcga 119280  
ttaggccccc gcggggcacc ggccggccggc cccacgcctt tccattaaac actcccacgt 119340  
tggggggggg cgcgccagct gagtgtcttg cggttgcggg cgcctgccc ggagatccat 119400  
taagccgcgg gagagcccga gcccgcgcgg cgtgttgctg tgggcatttc tctgcgctca 119460  
tccctgtctt tataaaaccg ggggcgcggc agcaacgaac gcaggggccc gccgcggatc 119520  
gagagggact ccggagaagg aaggtctctc cgcgcaccgg cgcgcccttc tccctcccc 119580  
tccctacctc cccctctctt cccctctttt tccccgcct cccgtctctt tccgcgcctc 119640  
cgagggtccg cctcttgctt cggggaaccc cggggggggc ggggcttggc cgcagaggtg 119700  
cgccccggcc ggagggggccc ccgcacctcg gcggcgcgcc cctccggcgc cgcgcgttcg 119760  
cgaaaggcgc gaaagggggc cccggaggct tttttcgatt cccggcccgg ggtcccgggt 119820  
agccgcggcg cgcggggggg aaggcgtccc cgcggggggc gtcgggcccg ggcggcccggc 119880  
ggagcgcggg ggcggggggg ccccgggccc cgcggggggc gtttcgcgtt tccgtttctt 119940  
ctccctcccg ggcgcggccc ctccggggc cgaccctcgc ccttccctt ctctctctt 120000  
tcccccgcc cgcgcggccc ctccctctt cctctctctt ctctgtctc ctctctcac 120060  
atttcccc ccccccccg ccgcgcgcgc cctttgccc cgtcccacc agacgcgcgg 120120  
ccgcgtgagc cgtccgcggg gggaccagg ctccgggggg gggggggccc tgcgtgtgtc 120180  
tcgtgtgaga gagcgcgccc ctcaaacgc gcgcgttctc gcaggtaggt ttagggtcgt 120240  
acaggtagc tctcgtgag gcgcggggga gagggggggg gggcgggcgg aagagagaag 120300  
agagcagggg ttgggggaga actgttctt ctccccctt caagaaacac gaggcggggg 120360  
tcccagaaag ggcaggcagc tcagccgcac cgcggcgag ccaaccgta tccctttttt 120420  
ctagggtgtt ttgtttttgt ttctgtttt gttgttttg ttattatttt cgcggatccg 120480  
gcgtgttcg atccacccc ccttctctt tctcttccc tccaccac ccccgtttcc 120540  
cccccccc cgtcgttcc cgggggggca ggccggggc gggcccgta gccaccgccc 120600  
cccacgcgc ggtcaccccc cccaacaac cccaaggcg cgtgccggc cacagccgtg 120660  
ggtgtggcgc ccgtccctt cctctaccgc gtgggcggcg gcgggggggt ggtggtagtg 120720  
gtggcggaag gaaacggggc gggggccggg gccgtaggg aaaggtaggc acgcgcgcgg 120780  
tgtgtcgact tgcattcccc gcaaaacgcg tcgtgtcgtg ttgtgtcgtg gtgggcccgtg 120840  
ttgtggtggg ccgtgtggtg tgggtggtg ttgcgaacgc gcgagcccc tcgccccgat 120900  
gggagtctcc ccgcagccag ggtaaaggag ggcgggcgtg gcgggcagggt gtgcgggcgg 120960  
ggtgggtga gtgcgggttc atgcctcggg tctctcttc ctgctctcc tctttctcc 121020  
cagccagggt gagagggggc gggcgtggcg ggcagggtg cgggcgggggt gggcgccggg 121080  
gcgggggtgg gcacgggggt aagtgcgggt geatgcctcg ggtctctct tctccctcc 121140  
cctctctccc acccgcccc gggggcagag ggcgtgcag cgtgtgtatt caaccgccc 121200  
cgccccgcc ccactttccc cctctctat caaagtccc tggccccgg cttcgcgcgg 121260  
gtggtgcggc tgaccccc cctctctcc ccccgagcca ggcgccctc cactctgcc 121320  
caccacccc agggctcggc cggccagacg tgcgtctct gcaagatcgg gccccctcc 121380

---

-continued

---

ctgtcaaacac ggacacactc tttttttacc cgccagccag cccgccacc caccaagaca 121440  
gggagccaga acgaggcccg gccccggctc tgttctatga taaagaccaa caggcctcgg 121500  
gggtgggggc ggtctctcgt gcccccctc cctcctcctc ctcccttccc ccccatcccc 121560  
ggccccctg cgcgggggag ctgcatcaaa ggccaacaac aaagtgtgtc aaaagcatca 121620  
caaaacttta ttgtaaaatt tttataaata taaagtttt ttttctcctca agttttcaac 121680  
aaggccagaa agtccataac aaaatgctgg tgtgtgttgc tgttcggggc cgtgtccgtc 121740  
ccccccccc actcccacc ccaacttctg tctcctccc gtctttcccc cccccacct 121800  
ccccctgccc ccgaggcgcc tcggccgggt gtccgggtgg gggcggttc cttcgggcag 121860  
caagccaggt gttagctccc cctactcccc gtggcccgcg ggggctcgc cggccggcgc 121920  
gggcgcgccc tgcctccgag accacgggtg gcgcgaccgg aggcctgga agtccagcgc 121980  
gcccaccagg gtgcctcgtt caaagagcat gttgccacc ggggtcatcc agaggtgtt 122040  
cactccgac gcggggggcg tcggtagtc gggggcctc acgcagtgc gcgcgtgctc 122100  
ggggagcagg gtgcggcgcc tccacgcggg ggccgcggcc cgcagcaggt ccgccacct 122160  
ccccgtctgg tccacgagga ccacgtaggc ccctatgtgg cccgtctcca tgtccaggac 122220  
gggcaggcag tccccctga ccgtcttgtt cacgtaaggc gccagggcca cgacgctcga 122280  
gacccccgc atgggcaggt agcgcgtgag gccggcgccc gggtcgcggg ccccgggctc 122340  
ggggccgccc tccgcgtggt gcgtcttctt ggcacacttc ctcggcccc gcggcgcagc 122400  
agcgcggggg ccgaggggag tttctcgtct ctccccagcg ccggacgcgg acgcgagcgt 122460  
cccaccagcc ccgcccgcag aggaagaggc ggaggaggag gaggcggagg aggaggaggc 122520  
ggaggaggag gaggcggagg aggaggaggc ggaggaggag gaggcggagg aggaggaggc 122580  
ggaggaggag gaggcggcgg cgaccgcggc ctgggacgac ggagacggc acggggggcgc 122640  
ggcggccgcg gacggccggg cgagcggccc gtggcccgcg tcgcccaggt ccgagtcggg 122700  
ggcccgggcg ggcggccccc tcttgcccc caccctctgg ggggcgaggg gcgagcggcg 122760  
ggcggcggag gaagaggcgg aggacgaggc cgcggggccc gagtccgacc cgcgcctctt 122820  
ccggggcgcg gccggccccc cctccgcggc gtggggggcg gcaccggggg tgttgggtgc 122880  
gcgggggacc ccgggtcctc cctccgcgccc cggccctccc gaccgcgcg cgtcggctgc 122940  
gcctgcccgg ccagactct gtgcttgggt gtccgtctga gcctgggtca tgcgggaccg 123000  
gggcgcgcg tgcgctcca ccggcacggc gggcggcgcg ggcggggccc cgtccgct 123060  
cgcagacacc acggggggcg cggcggcgcg gggcggactc cggacgcgcg gggcgcggc 123120  
cgcgcggggg cgcgcggcgc gcccgcgaca ctgtggcaga cctccccccc cggggccgga 123180  
ggacacctgt gcggaggagg aggagacaaa ggagagcggc ccggggcccc cggggcggcg 123240  
cggagacggc gggggagagt cgtgatgac tatggggggc tcctgggccc cgcggggctg 123300  
tctcgcgggg ggcgtcctgc cctccgcgccc cgcggcgtct tcgcccacc gccgcgctg 123360  
cgcgcgcccc ccgcccggcg cagggggaag agaggccact ctcggcacga cggccgcgac 123420  
ggcagggccc ccccagacc cagatcccac ccccggccc aacggggcgc cgcgctgct 123480  
gctgctccgc ggggcccag gggcggcgg tcgggtcgc gcgggctggg aggttccgcg 123540  
ggtcggcccc gaccggccc ccccgcgccc gggcgtctt cggggggcgg gcgggacgta 123600  
gtccactgca gagggagaca gagacgggag cccccggta gtgccgacc cccgcccgac 123660

-continued

---

ccccgccga cccccgccg accccgcc cccccgcc cccccgcc cccgcccg 123720  
ccccgcccc gcccgcccc cccccgcccc cccccgcc cccccgcc cccccgcc 123780  
cctcaccgtc ggcaggtca tcgtcctcgt cgtccgtgcc gggccaeggg ggggtggcg 123840  
acaggggcgc gaccgtgtgt cccccagcg acagggagcg cggggccgtc cgggggttg 123900  
ccgtccagat aaagtccacg gccgtgccg cccgcacgpc cgcctcggc tccacgcggg 123960  
tccgggggtc gttcactatc gggatggtgc tgaacgacce gctggcggtc acgcccacta 124020  
tcaggtaacg caccgggggtg ttgcacaggg gacacgtgtt gcgcaacgga atccaggtct 124080  
tcatgcacgg gatgcagaag ggggtgcaggc agggaaaact ctggcagcgc agggggcggg 124140  
cgatctcgtc cgtgcacacg gcacacacgt ccccccccc tcccgcttc gettctctct 124200  
caccacggg cccaccccca caggatccct gcgcgtcggc gggcggtggg ctgccctggc 124260  
gctcggcggg gggcggggc gggggcggtg cgcgctccat caggccccgc tcgaaacatct 124320  
ccgtgtccgt gctgcccgcc tcggaggtgg agtcgcggtg aaggtcgtcg tcagagattc 124380  
ccaectcggc ctctctctcc gagtcgctgc tggcgagcca ctgcatgtcg ttgagcatcc 124440  
cccaggcgtg cggggcgggg ggtgcttga caaagcaacg ggggggattt agaggcgcg 124500  
ggggtgagg cgggaccccc gcgcgctgtc ccccggtgtc ctccctcacc cgggcccccc 124560  
gcccgtgct tttgttcgg aaggggggga gaaagggggt cgtaacccaa ggtggtctgc 124620  
gtcctttgga ttcgacccc tcgtctcccc cctgtcccc cgtctcggg ctctctctctg 124680  
cctccctcgc cccccagag ggtcgggggg cggcgcaecg cccacggggg tccccgacc 124740  
gcttaagcgg gccggggggtc gggcccgta agcgtcccc ccccgagcc caccgcccc 124800  
gaccaccccc aaccgcagc cgggtggtcc ggggaaaagg gggggcctga gaccgggggg 124860  
tcgcccctc accgtgccgg ggtctgccc cggcgccgc tcggggcccg ggtccgccc 124920  
ggagctcgtc cggggcgggg gttccatgag cggggtagg gtagactcga gacggcgggc 124980  
cggggtctct ctcttcggg gttttagct ctgtctctcc gggctctctc ctcccgccgg 125040  
gccgcccgc cgtcgtcgc agtcgcccgg tcgcaatgcg gcccgaccgt cacacggggc 125100  
tgccttatac cgggcgccta tccactcccc caaaggggcg gcatttaaga tcccccaat 125160  
agccgcgcgc cccggcgggg gcggagggag ggaatcccc cctctcgggg cggccccgct 125220  
cccggggacc aaccgggtgt actccaagaa cccattagc atgcgcccgc ccccgccgac 125280  
gcagatggga gtcccccccg ccccccccg gcgcggccct gagggtggtc ccccccggg 125340  
gaaaaattca ttagcatact aggaagccca ggggaccaat aggggccgat cagcccacc 125400  
accggcgggc gcgcgaggct ctgctgttcc tgccaagaaa gtaatcagca taaccggaa 125460  
ccccgaggga gtaattacgc ggggagcgag gggcgtccg aacgtttta attaccataa 125520  
cggggaatgg cggcccgta aaagctgcta attaccgca gcgggaaacg cggcccatta 125580  
aaagttgcta attaccatgc gcgggatgg cggccgggac cgcctattaa aagtttctaa 125640  
ttaccatacc gggaaagcgg cgcggggcgg tcgcccgggg ggagtcgggg cccgcccggc 125700  
ggcgcgcggt tggccggcgc cccccctgg ggcggcgga gcggcggggg ggcgcccggc 125760  
cctcgggat atatacggg ggtcccatc gtctcttcgg agagcgccct cgcgcagacc 125820  
ttcggagctc cggggctccg ccggccgagg ccgcccctgc cggttcaacc ctgaccgccc 125880  
cgacggcccg gcccgcggc ggcggaggac ccgcgcgcgc ccgcgcgcgc ctctctctcc 125940

---

-continued

---

tccgcgggtc cgccgtcttc gtgggcccgg gctcgggctc gggcccagc tccggcctcg 126000  
ggctccaggc acgggtccgat gaccgcctcg gccgcgccca cgcggcgccg gaaccggctcg 126060  
cggctcggccc gctcgcgcgc ccaggacccc cgtcgggcca ggcgcgcggc cgtctcccag 126120  
gccaccagat ggcgcacctg caccgcgcgc gagaagcaca cctgcggggc gggagacacg 126180  
ggggtcggag gggcgtcagg gggtcggagg ggcgtcaggg ggtcggaggg gcgtcagggg 126240  
gtcggagggg cgtcaggggg tcggaggggc gtcagggggt cggaggggcg tcagggggtc 126300  
ggaggggagg cgtaccttcc cgcgcggcgc gtcgcggg cgggacgcgg ggggccgcgc 126360  
ccggcgcagg ctcaggcgcg ccaggtaact cgtcgtggtg cgcagccgta gcgccaggtg 126420  
gggcggaagg gggcgtcgcg gcccgcgctc cttgcgcggc ggcggcgggg ggcaggcggc 126480  
ggcaggcgcg gcgtcggggg cctcgcgcgc cttccccccg cctcgcctcg ggggctgtt 126540  
cgcccactct gcgtcgtcgt tgccggcgta gtcgcgctcg tcgctgtcgt ccgcctgggg 126600  
caccagcagc cagcgcgccca ggagcagga cgcggccggc gcgctctcga ccgcggttcc 126660  
cgagtcgtac gcagggacca tttgggagtc tcgggttggg agcgcgcggg ggcgcggcac 126720  
ggctggagcg ccggggcgcg gcacggctgg agcgcggggg cgcggccggc gccggggacc 126780  
ccggcgcggg ggaccccggc ggcgggacat ggcggcgggc tgggctcggc gtaggcccgg 126840  
agccggagcg cgtcggggcg ggagagttca ctcggcaacg atgcacgtgt aaccgccagt 126900  
ccgtgcttgc ctaggcaact caccgcctcc ggctggcgtg cgcagcccg gccgtgttgc 126960  
gggcccctct aaggggcgcg ggcaggacgg ggaactccgc cccgcctctt tcccccggg 127020  
gagtcaacc ccgggggggg tgttttttgg gggggggcgc gaaggcggg ggcggcgcg 127080  
ggcggggcgc agggcagccc cgcgcgcccc cttccccgtc cctcccccg agccggcgc 127140  
tcccccgcg gcgcgcgccc tcccccgcg cgcgcggggg ctgccttccc gcgggcgccc 127200  
ccgcgcggct tttttccgc gccccgcccc gcgcggcagg acggggacta gcaggctgtg 127260  
ccgcagacca ccacacactc ccaagctccc cgcgcccccg aagacgccag tcgcaccacc 127320  
gctcgcctc gcagaccaga cagtgcacc aagcaccgc cgcgccgac acggttcccc 127380  
gccacccct cctccctc catcccgcg agctcgcggc agccctccc ccccgcgcg 127440  
cacggggctg cgtcccgcg gccgcctccc ccgcggccgc ctccccccg ccccgccccg 127500  
ggggcttccc ccgcccctcc ccccgcgccc gcggccccga gctcgcagca gccctcccct 127560  
cccgccccg gtgccttccc tcccgctcct gcgggggggc tcgggcccac tgacctcgt 127620  
aacctgcact caggtcagag cccagaccc cccgcggggc cgggagacgt gccgccgccc 127680  
cgacccccgc ccgcccgacc cccgcccgc cgacccccgc ccgcccgacc cccgcccgc 127740  
cgacccccgc ccgcccgacc cccgcccgc cgacccccgc ccgcccgacc cccgcccgc 127800  
cgacccccgc ccgcccgacc cccgcccgc cgacccccgc ccgcccgacc cccgcccgc 127860  
cgacccccgc ccgcccgacc cccgcccgc cgacccccgc ccgcccgacc cccgaataaa 127920  
ccacacaagg cgttacgtt tcgtctgtct cgttctttat ttctcacaca cgcgcggggc 127980  
catcgcggc tetgtcttaa aggcgcacag acgcccgatt cctccccct ctcccctct 128040  
ccccctccc ccgctccgg aagtttccc ccccgtaact cccaaaacag tccgctcgtc 128100  
tcgtcctcca gctcccgct catgtccac ggctcgcgc tcggcggcgt gccagcccc 128160  
gcggcggtcc ccaccacct caccgcgcgc cccgcccggc ccagcacct ccccgcgcg 128220

---

-continued

---

cccgcggcgc acgcccagcg tatctgcggg ggcgggcccg cgtccgcgtc gtcgcgcagc 128280  
accagcgggg gcgcgtcgcc gtcgggctcg agcagcggcc gcgcgcagaa ctcccgcgc 128340  
ggcccgcgca gctccgcccg gccccgcgcg acggcgtcgc gcccagcgc cactagacg 128400  
ggccgcagcg gcgcgcccag gcccagcgc gcgcaggcgc ggtgcgagtg cgcctcgtec 128460  
tcgcagaagt ccggcgcgcc gggcgccatg gcgtcgcgcc cgcccagggc ggcggcccgcg 128520  
ccgtccagcg ccgggagcac ggcgcggcgg tactcgcgcg gggacatggg caccagcgtg 128580  
tcggggccga agcgcgtgcg cacgcggtac cgcacgttgg ccccgcggca gaggcgcagc 128640  
ggcggcgcgt cggggtacag gcgcgcgtgc gcggcctcca cgcgcgcgaa gaccccgcg 128700  
ccgaacacgc ggcgggagcg cagcacgggt gcgcgcaggt cccgcgcgcg cggccagcgc 128760  
acggcgcact gcacggcggg cagcacctcg caggccaggt aggcgtgctg ccgcgagacc 128820  
acgggcccgt cggcgggcca gtccgcggcg cgcacggcgt tgacgacgat gaggcggcgg 128880  
tcgcaggcgc cgccagcag ccccaggaac tccacggcgc cggcgaaggc caggtcccgc 128940  
gtggacagca gcagcacgcc ctgcgcgcc agcgcgcgaga cgtcgggggc gccggtccag 129000  
ttgccgccc aggcggccgt ggcgggcccg cagagccggt tgcccagggc cgcagcagg 129060  
caggacagcc cgcgcgcgtc ggcggaccac tccggggggg gcccccgcg ggcgcggccc 129120  
gcggccaggt cctcgcccgg cagcggcgag tagaggatca ccacgcgcac gtcctccggg 129180  
tcgggcaect ggcgcctcca ggcgcgcgcg cggcgcagcg gcccagggc gcgcagcggg 129240  
ccgaaggcgg cggcgcgcgc cccggggggc ggggcggcgc agcgcgcggc cagcagggcc 129300  
agcgcgcgcg ggtcgaacat gagggccggg cgcacgcgcg cggggaagag cgggtggtcc 129360  
gtgagctcgg ccacggcccg cggggcgcag taggcctcca gggcggcggc cagggcgcgc 129420  
ggcgtgtggc tggcccccg cggctggcgg cgcacgcgc cctcggggtc ggggcctcgc 129480  
gcgggcgcgc gggtcagcgc cgcggggcgc ggcggccgcg gcgcgcgcgt cggcggggcg 129540  
gggggcgcgc ccccgcggg agggggcggc gcggggcggg gggcgtcgcg gcggtcttc 129600  
ttcggggggc gcggggcgc ccccgggcgc gccctggcgc gggcggggct cttgcgcttg 129660  
cgcgcctccc gcggcgcgga ggcgggcgcg gcgagcaggt cggcgcggc gaagggtgctg 129720  
gccagcaggg ggcgcaggct ctggttctgg aagagcaggt ccgcggcggc ggcggcggcg 129780  
gagctcagca ggcgcgggct ccgcggcagc gccgggcccga gggcccgcg gaccaggctc 129840  
acggcgcgca cggcggccac ggcggcctcg ctgccgcggc ccacgcgcag gtcccgcgc 129900  
aggcgcacat gcaccagcgc gtcgcgcacg aaaccgagct cgcgcagcca ggcgcgcagg 129960  
cggggcgcgt cggcgtgccc cggggcggcc gcgccgcgg gcccccggcg cggggcgcgc 130020  
gcgggcggg ctccggccag ccccggcacg gccgcaggt cgcgctgaa gccctccgc 130080  
agcgcctcca ggatcccgc gcaggcggcc aggcactcca cggccacgcg gccgcctec 130140  
gcgcgcggc cgcgcgccac accgcggcg ccgtcgtcgt cgtcgtcgtc gtcggccccg 130200  
gccggcgcgg aggcggggcg ggcgctcagg cgcgccagg cggcagcac ccccgcggcg 130260  
ccgtagcccg cgggcaaccg gcgctcgtc gccggcgcg ccgcgcgca cggcaaccgg 130320  
gcggcggcgg cggcgggctt cccgcggcg tcgtcgcgt cgtggcggtt ggcgtcgcg 130380  
ccgtcgtcgg gggttcgcgc cccggtcagc gccgcgttct cgcgcgccag caggggcgcg 130440  
taggcgcggc gcaggctggt cagcaggaag cccttctgcg cgcggtcgtg gcggcggctc 130500

---

-continued

---

atggccacgg cgcccccac gtgcgccagg ccccagccga agcggcccgc cgccatggcg 130560  
taccocaggt ggggcacggc ccgcccacg ctgccggaga tgaaggagct gctggtgccc 130620  
gccgcgcccg agatccggaa gcaggcctgg tccagcgcca cgtcccggg cgccacggcg 130680  
gggttctgga gccaccccat cgcctccggc tccggcgtgt acagcagccg cgtgatcagg 130740  
gcgtactgct gcgcccgcgc gccagctcg ggcgcccaca cgggcgccgg ggcgcccag 130800  
gcctcgaacc gggcccgcgc ctctccgcc tggggcgccc cccagaggcc gggggcgctg 130860  
tcgcccagcc cgcctacag cacgcgcccc gggggcgggg ggcggcccc gggccacggc 130920  
tcccgcgta cgtaccgcgc gcgtagcgc gcgtagaagg cgcggaggc cgcgtcggcg 130980  
tccagctcga cccgcgggg ccgcccggcc gtgaagcggc ccgtggcgtc gcggccggcc 131040  
accgccgcgc gggcccggcg gcgctccagg cggcccgcgc tcgcgcggg ggtccggccc 131100  
ggggcgggct cggccctggg ggggctcggc cggggcgccc ccccggggc cctcggggc 131160  
acccccgct cctcgtcgc cgcgccagg gtcccggcc gggcgtggtc tgcggcgtg 131220  
gcggggcgcc gggcgcgctc tcgctcgtc tcgctcgtc acgaggaggc ggatgcagac 131280  
gaggaggagg aggcggagg gagggcgagg gacgcccagc acgaggatcc ggattttgat 131340  
gagtcagagg cggccgagcg ccgcccgggg gcgcccggc ggcggtggtg gtggtggtgg 131400  
tgggtcggc ggggcgccgg gggctcggc gacaggctgg ccattggggtc cgggtacgcc 131460  
ccgcggacc cggacgtcgt ctccggtcgc cggaccagc ggcgccgctc gcggctcgtc 131520  
tcctcgtcgt cgtcgtcgc tcgctcgtc tcctcggcc aatcggcgcg catggagggg 131580  
gtccgcggcg gagaaggcga gcgggcccgt tcttcttgc gcgcgtcgc ctccgggggg 131640  
ggcgacggga tcgtcgaac ggcctcgtcc accatcgagg ccagcagggc cagctgcgc 131700  
ggcgagacga cgcctccgc ggcaggctcg tcgacggcct ccccggacc cggggccgc 131760  
tcgtcggcat cggcatcggc ggcggcgtc tcggcctcgt ctctcgtc ctccggccca 131820  
ccgtgccacc cgaaccgggg ccgcccggcg gggcgacggt ccggggtcgg ggtggggcg 131880  
ggtccgtcgg ctggatccgg agatccgggg ccgcccgtcg tctccgcgc ggcgggaga 131940  
cgtcccgcgt cctcgtcgc cctcggacc tcggccccgc ggcctcgtc cgtcgtcgc 132000  
gtctctctct tcttccgctg ctccgccgac atcgcctcc accggggtgt gcgggggggg 132060  
ggtctctctc tctctctca ggggcggcag tggggggggg tggttggcag tctctctccc 132120  
ccccgtcgg tcgctcgtg tcgctgtgc ttttcgctc tccgcgccga tcgggtagat 132180  
cctggcgccc gcgtcggtag ccgctcgc tgtggacgat cgcgccgctc cctggctgat 132240  
atagtcctcg gggcgccggc ggcgggggga aaggaggagg acgcccggga ggagcgatcg 132300  
acgcgccgc gcccccgtc gccggggtc cgcgccagc tggaaaccga ttatgcggcg 132360  
ccccccccg acgcccgcgc gtcgcgtcc gtggcgggcg ccggttggtc gcgcccgc 132420  
cgtccgcc gcgcccgc tcattagcgc ccggcgggg cggttcgc tcgcccccgc 132480  
gatgctaagt agaccctcgt cgcgggggg ctccgtcccc tgccttcgc ggttcgtggt 132540  
aatgagatgc cggcccgcgc ctcccgttg ccccgcgg ccccaagg gcccggcagg 132600  
tcgcccgtt ggtcggcggg cggtccgcc ccaaggggg cggggcgcca gggtaaaaga 132660  
agtgagaac cgaagcgttc gcacttcgtc ctaatagtat atatattatt agggcaagt 132720  
gcgagcgtg gcgcccgc cggggcccgc gtcaccccgc gtcgccccc aaaggggcg 132780

---

-continued

---

gggcccagcagg gtaaaagaag tgagaacgcg aagcgttcgc acttcgtcct aatagtatat 132840  
atattattag ggcaaaagtgc gagcactggc gccctgcccc gggcccgcgt catcccgcgg 132900  
gctccgcccc gaggcggggc cggacggggg gcgggcccgt cctcgcgcac ataaagggcc 132960  
ggcgtcccgg tcgcccgcgc accaggggca caccggctgc gcggcggaga ccgggacggc 133020  
agcggcggca tcgcaagggg ggcacagcgc agacagagac gccggcggcg agcggggcac 133080  
cgacgcaccc ggatcgggatc ggatacacag acgcggggcg atcggttcct ttctgttctg 133140  
cctttccctc ccccccccc cccccaccc tgtacgtacc gcgaggaccc atccaccac 133200  
tgcagcctta tcgcaggtac ggtgaccggg ggggcccggc ggggggacgg gcgggggacg 133260  
gggggacggg ccggggggag gggcccgggg gacgggcccgg ggggacgggc ccgggggacg 133320  
ggcccggggg acgggcccgg gggacgggccc ggggggacgg gccgggggga ccggcccggg 133380  
ggcccggggg ccggggggcc ggggggcccgg ggggacgggg ggacgggggg acggggggac 133440  
ggggggacgg ggggacgggg ggacgggggg acggggggac ggggggacgg ggggacgggg 133500  
ggacgggggg acggggggac ggggggacgg ggggacgggg ggacggggcc gggggacggg 133560  
gggacgggccc ggggggacgg ggggacgggc cggggggacg gggggacggg ccggggggac 133620  
ggggggacgg gccgggggga cggggccccg atcccaacat ccgcgcttc tcgcaggccg 133680  
ggcgcgcctc tcgtggacgg gacaccgggtg tggttaactgg cgacaaggcg ttgccactat 133740  
ggcagacatc cccccggacc gccccgcgct caacacgacg cctgcgaatc atgctcccc 133800  
atccccaccc ccgggttcac ggaagcgcag acgcccgcgc cccccagct cgtcggaatc 133860  
tgagggttaag cccgacacag aatcggaatc ctctcgacc gagtcgtccg aggatgaggc 133920  
gggagacctc cgcggggggc gccgtgcctc cccgcccggg ctcgggggga ggtattttt 133980  
ggatctgtgc gcagaatcga ccacggggac ggaatcggag ggaacggggc cgtcggacga 134040  
cgatgatgat gatgcgtcag acggtcgggt ggttgacaca ccccccgca aatccaagcg 134100  
accccgaaac aacctcgat taacgagctc ccccgaccgg cgtgcgggtg tggttttccc 134160  
cgagggtgag agaagcgaca gacctatccg cgcggcgcaa cccagggccc cggccagtct 134220  
tccggggatc gcgcacgcgc accggcgctc tgctgcaccg gccagatgc ggagcggagc 134280  
cgctggacg cttgatctgc attacatag ccagtgctgc aaccagctct ttcggatcct 134340  
gcgtgccccc ccgaacccc cggcagcgc caaccgcctg cgcacacctg tgcgagactg 134400  
ctacctcatg ggctactgcc ggaccccctc ggggcccgcg acgtggggcc gcctgctgca 134460  
gatctcgggc ggaacctggg acgtgcgctc gcgaaacgca atccgggag tcgaggcgca 134520  
ttttgaacct gccgcccgag ccgtgtgcga gctgcctgt ctgaaaccca ggcgttacgg 134580  
ccccgagtgt gatggttgca atctcgagac caacggcggc tcgacgagcg atgatgagat 134640  
atcggatgag acggactcgg acgataccct cgcgtcccat tccgacacgg agggggggcc 134700  
ctccccggcc gccggggaga acccggaatc cgcgtccggc ggggctatcg cgctcggct 134760  
ggagtgtgag tttgggacgt ttgactggac gtccgaggag ggctcccagc cctggctgctc 134820  
cgcggtggtc gccgatacca gctccgcca acgctctggc ctaccgccc cgggcgcgtg 134880  
tcgcgcaacg gaagccccag aacgcgagga cgggtgccga aaaatgcgct tccccgccc 134940  
ctgcccctat cctgcccgc acacatttct ccggccatga gcgcgggacc ccagcccgg 135000  
tgtgtttgcc aaacgaaaaa taaacgccc acaagaaagc tttgtgtctc gactgtctg 135060

---

-continued

---

tttttctggg ggtggaggaa ggaacgacaa aaaaaagaaa caaacgcgac accgctcgta 135120  
cgtgtaatgg ggccgcagtg ttttttatta gcatcggggg gggttagagg ttggtgattg 135180  
gatagcaaac gtgggatgac ggagggcact cgtcgccaac ggccagcggg ggccccgggt 135240  
tctgggggtc atcgtecccc gtctgccagg agggctcctc gggaatctcg ggtcgcccca 135300  
tgcacgtaaa acacgggccc tgcgtggggg gggtcgccgg atgcgggccc gatgatgcgg 135360  
ggcggggttt gttgtgagga gccacgaggg accgtagcca gcgaagacag ctgctgtccc 135420  
ggtcgccggg caccaccacg ccgtattggt attcgtatcg gctaaggaga ttttcagggg 135480  
ggtgattagg cgtgcggggg aacgggggtcc acgacacggt ccgctcgggc aaaaaccgat 135540  
cgggcagggg ccacggttcc cccaccacag cgtcgttggt cttcgtggcg atgaagcga 135600  
accccagccg ggttttttgg tgcgtactga aaaacggcac acacaggtcc gccgccccga 135660  
ccaccacag gtggtatagc cgggtggggc cggggcgctc ttgatgcagg agccgaaaac 135720  
acgcaggggc atccagaatc tcgatgcttt ccagggggtc gtctccgca aacaggcccg 135780  
tcgtggtgtt tgggggacag gcacaggagc gggttcgcac gatcggtcgg gtgaatttgg 135840  
gcaagtccat cagaggctcg gccagcctgc gaaggtcgc cgggccaacc accaccgggg 135900  
ttcccagagg ctcgagggcc aggatccggc attgccgaag cagaaaactc cacagagccg 135960  
ggcttgctc agcggaaagtc cgcggcaggg cgtttcgttg gtctaggagg gtaaccacac 136020  
ttacaacaac aacgccatg tcggtatatt aggccctggg tccgatcttc actcaetcgc 136080  
ctgtctcgg acctatgac ggggggacgg cgcgcggacc cgggggggct gcttctatc 136140  
acacggcccg ttcgcacggt cgattttttc agccttgttt ggttggttag gtatcccga 136200  
taatctgacg ttcggatat agggggcggg gggtagtggg ggggtgtctc acaaaactgc 136260  
gcttcttaaa acaccggggc ccgtcctcgc gggtgctcgt tggttggcac gcgcgacgcg 136320  
gcaaatggcc tgtcgtaaat tctgtggggg ctaccgtaga ccgacaaga gacaggaggc 136380  
gtccgtcccg ccggagacaa acacggcccc ggccctcccg gcgagcact tttatacccc 136440  
cgcggaggat gcgtacctgg ccccggggc cccgaaaacc atccaccctt cccgcccacc 136500  
gtccccggc gaggtgcgc gcctgtgtca gctgcaggag atcttgccc agatgcacag 136560  
cgacgaggac taccatcgc tggacgcgc ggggtcggag gaggaagac aggccgacga 136620  
tgacgccccg gatgacgtgg cctaccggga ggactacgcg gaggggctt tctgtccat 136680  
ggtttcggc gccccctgc ccggagccag cggccatcct cctgttccgg gccgcgcagc 136740  
ccccccgac gtcggacct gcgacacggg taagggtggg gccacgggt tcaccccgga 136800  
agagctcgc accatggacc gggaggcact tcgggcatc agccgggggt gcaagcccc 136860  
ttcgaccctg gcaaaactgg tgaccgggct gggattcgcg atccacggag cgtcatccc 136920  
ggggtcggag ggggtgtct ttgatagcag ccaccgaaac taccctcctc gggtaactct 136980  
caaggcgggg tggtaaccca gcacgagcca cgaggcgcgg ctgctgagac gcctgaacca 137040  
ccccgcgac ctaccctcc tggacctgca cgtcgtttct ggggtcacgt gtctggtcct 137100  
ccccaaatg actgcgacc tgtataccta tctgagcaag cgcctctc cgttgggcca 137160  
cctacagata accgcggtct cccggcagct cttgagcgc atcgactacg tccactgcaa 137220  
aggcatcctc caccgcgata ttaagaccga gaacatcttc atcaacaccc ccgagaacat 137280  
ctgtctgggg gactttgggg cggcgtgctt tgtgcgcggg tgcgatcga gcccttcca 137340

-continued

---

ttacgggatc gcaggcacca tcgatacaaa cgcccccgag gtccctggccg gggatccgta 137400  
caccacaggta atcgacatct ggagcgcggg cctgggtgatc tttgagaccg ccgtccacac 137460  
cgcgctccttg ttctcggccc cgcgcgaccc cgaaggcgcg ccgtgcgaca accagatcgc 137520  
gcgcatcctc cgacaggccc aggtacacgt cgacgagttt ccgacgcacg cggaatcgcg 137580  
cctcaccgcg cactaccgct cgcgggcggc cgggaacaat cgtccggcgt ggacccgacc 137640  
ggcgtggacc cgctactaca agatccacac agacgtcgaa tatctcatat gcaaagccct 137700  
tacctttgac gcggcgctcc gcccgaagcgc cgcggagttg ctgcccctgc cgtatttca 137760  
ccctaagtga ccccgctccc cccggggggg gtggaggggg gggctggtg gatgttttg 137820  
cacaaaaaga cgcgccctc gggttttggg gtttttgga ccttgccgc ccggtcctg 137880  
cacgccatcg ctcccagggt gcttctctt tttgtcttt ctggtcttcc ggggacacgc 137940  
ggcgggtcgg gtgtccccg accaattaat ccccccaaca gcgatgtgt tttccggga 138000  
ggttccccg tggtcaata ttgttatgcc tatccccggt tggacgatcc cgggcccttg 138060  
ggttcccgcg acgccccggg gcaagacctg cccccgcgcg tcgtccgta cgagccccg 138120  
ggccgctcgt tctcaccggg ggggctgggt ttgetggcgc cgccggtacg cggatttggc 138180  
gcaaccaacg caacgtatgc gcccgtgtg acgtactacc ggtcacccg cgectgccg 138240  
cagcccatcc tcttcggca gtatggaggg tgtcggggc gcgagccgc gtccccaaag 138300  
acgtcggggt cgtacacgta cacgtaccag ggcggcgggc ctccgaccg gtacgctctc 138360  
gtaaatgctt cctgctgtg gccgatctgg gaccgcggcg cggagacatt cgagtaccg 138420  
atcgaactcg gcggcgagct gcaactgggt ctgttggtgg tagaggtggg cggggagggc 138480  
ccccgcccc ccccccccc acaggcgggc cgtgcgagg gcggccccg cgtccccccg 138540  
gtcccccgcg gcccccctg gcgctcggg cccccggtat ggtattccgc ccccaacccc 138600  
gggtttcgtg gctcgcgtt cgggagcgc tgtctgcccc cacagacgc cgccgcccc 138660  
agcgacctac cagcgtcgc ttttgcctcc cagagcctgc tgggtgggat tacgggcccgc 138720  
acgtttatc ggatggcag cccacggaa gacgtcgggg tctcgcgc ccattgggc 138780  
ccccgggccc tagatgacg tccgtacgc ccttccac cccgccccg gtttcgacgc 138840  
gccctgcgga cagaccccga gggggtcgac cccgacgttc gggcccccg aaccggcgcg 138900  
cgcctcatgg ccttgaccga ggacacgtcc tccgattcgc ctacgtccgc tccggagaag 138960  
acgccccctc ctgtgtcggc caccgccatg gcaccctcag tcgaccaag cgcggaaccg 139020  
accgccccg caaccactac tcccccgac gagatggcca cacaagcgc aacggtcgc 139080  
gttacgccg aggaaacggc agtcgcctcc ccgcccgcga ctgcatcgt ggagtctcg 139140  
cactccccg ccgcgcgcg ggcaacgccc ggggcccggg acacgaaac cagcagccc 139200  
tccgcagcga aaacgcccc caccacacca gccccacga ccccccgcc caegtctacc 139260  
cacgcgaacc cccgccccac gactccggg ccccaacaa cccctcccg acccgcaacc 139320  
ccgggtccg tgggcccctc cgccgcgccc acggccgatt cccccctcag cgectcgc 139380  
cccctaccg cgccggggc ctccggcgcc aacgttctg tcgcccgcac caccgccag 139440  
cccgaaacc ggggcccgc ccgtaccccc ccaacggacc caaagacgca cccaccgga 139500  
cccgcggacg ctcccccg ctccgacgc ccccccccc ccgaacatcg cggcgacc 139560  
gaggagtgtt agggcgccgg ggaaggcgaa ccccccgagg acgacgacg cgcaccggc 139620

---

-continued

---

ctcgecttcc gaactccgaa cccaacaaa ccacccccg cgcgccccg gcccatccgc 139680  
cccacgctcc cgcaggaat tcttgggccc ctcgccccca acacgcctcg cccccccgc 139740  
caagtcctcc ctaaggacat gccctcgggc cccacacccc aacacatccc cctgttctgg 139800  
ttcctaaccg cctcccctgc tctagatata ctctttatca tcagcaccac catccacacg 139860  
ggggcgttcg tttgtctggt cgccttgcca gcacaacttt ggcgcggccg ggcggggcgc 139920  
aggcgatacg cgcacccgag cgtgcgttac gtatgtctgc caccggagcg ggattagggg 139980  
gtgggggttg ggggcgagaa acgatgaagg acgggaaagg gaacagcgac caaatgtcac 140040  
gataagaaca ataaacctgt gacgtcaatc agatatgtga gtttggttgt gttttgtggg 140100  
actggggggc ggggggtggg ggtatcagtg ggtgacagag tcttttaaaa gacgtgtccc 140160  
ggggccctcg agatgcgcaa cttttggcca cacagagaaa ggccccaga cgaagtccac 140220  
cgggtccccc aacaaaaaca aaaacctga ccgcccgcgg ggggcgtgcc tgtgttttg 140280  
gtctcaatgg atcggtatgc cgttcggacc tgggggattg tgggaatcct cgggtgtgct 140340  
gctgttgggg ccgcacccc cggccccgcg tccgatacaa caaacgcgac cgcacgcctc 140400  
cccacgcacc ccccactcat ccgttccggg ggctttgccc tccccctcat cgtggggggg 140460  
ctgtgtctca tgattctggg gatggcgtgt ctactcgagg tcctgcctcg cctgggtcgc 140520  
gagttggcga ggtgctgccc ccacgcgggc caatttgcgc catgattttt cgcctttctg 140580  
gccttgcccc caccceatcg ccccgattgt gtgtcgggtg cccgggttac agcagetatg 140640  
gagcggctcg taataataat ttggttctgc ccacacgccc cgtgcccggc atgggttctg 140700  
cgggaaggac gaaataatcc ggcgatcccc aagcgtacca actggggggg gggggggggg 140760  
ggaaaagaaa ctaaaaacac atcaagccca caacccatcc cacaatgggg gttatggcgg 140820  
accaccgca ccaccatact ccgattcgcac cacatatgca accaaatcac ccccagaggg 140880  
gaggttccat ttttacgagg aggaggagta taatagagtc tttgtgttta aaacccgggg 140940  
tcgggtggtt gttcggctcat aagctgcatt gcgaacgact agtcgccgtt tttcgtgtgc 141000  
atcgcgtatc acggcatggg cgccttgacc tccggcgtcg ggaecggcgc cctgctagtt 141060  
gtcgcggttg gactccgcgt cgtctcgcgc aaatacgctc tagcagaccc ctcgcttaag 141120  
atggccgatc ccaatcgatt tcgcgggaag aaccttccgg ttttgacca gctgaccgac 141180  
ccccccggg tgaagcgtgt ttaccacatt cagccgagcc tggaggaccc gttccagccc 141240  
cccagcatcc cgatcactgt gtactacgca gtgetggaac gtgcctgccg cagcgtgctc 141300  
ctacatgccc catcggaggc ccccagatc gtgcgcgggg cttcggaaga ggcccgaaa 141360  
cacacgtaca acctgacct cgcctgggat cgcctgggag acaattgcgc tatccccatc 141420  
acggttatgg aatacaccca gtgcccctac aacaagtcgt tgggggtctg ccccatccga 141480  
acgcagcccc gctggageta ctatgacagc tttagcgcgc tcagcgagga taacctggga 141540  
ttcctgatgc acgccccgc cttcgagacc gcgggtactg acctgcggct agtgaagata 141600  
aacgactgga cggagatcac acaatttata ctggagcacc gggcccgcgc ctctgcaag 141660  
tacgtctccc cctgcgcat cccccggca gcgtgcctca cctcgaaggc ctaccaacag 141720  
ggcgtgacgg tcgacagcat cgggatgcta cccccttta tccccgaaa ccagcgcacc 141780  
gtcgcctat acagcttaaa aatcgcggg tggcagggcc ccaagcccc gtacaccagc 141840  
acctgctgc cgcggagct gtccgacacc accaacgcca cgaacccga actcgttccc 141900

-continued

---

gaagaccocg aggactcggc cctcttagag gatccccgcg ggacgggtgc ttcgagatc 141960  
ccccaaaact ggcacatccc gtcgatccag gacgtcgcgc cgcaccacgc ccccgccgc 142020  
cccagcaacc cgggcctgat catcgccgcg ctggccggca gtaccctggc ggtgctggtc 142080  
atcgccggta ttgcgttttg ggtacgcgc cgcgctcaga tggcccccaa ggcctacgt 142140  
ctccccaca tccgggatga cgacgcgcc ccctcgcacc agccattgtt ttactagagg 142200  
agtttcccg ctccegtgta cctctggccc cgtgtgggag ggtggctggg gtatttgggt 142260  
gggacttga ctcgcataa agggagtctc gaaggaggga aactaggaca gttcatagc 142320  
cgggagcgtg gggcgcgcac cgcctgccc acgattagcc accgcgcca cagccacctc 142380  
gaccgcctc atccccgtat gcccgccgc tcgctgcagg gcctggcgat cctgggcctg 142440  
tgggtctgcy ccaccggcct ggtcgtccgc ggccccacgc tcagtctggt ctgagactca 142500  
ctcgtgatg cggggccgt ggggccccag ggctcgtgg aagaggacct gcgtgttttc 142560  
ggggagctt attttgggg ggcccagtc cccatacaa actactacga cggcatcatc 142620  
gagctgttc actaccccc ggggaaccac tgccccgcg ttgtacacgt ggtcacactg 142680  
accgatgcc cccgcgcgcc cgcctgggg ttcacctgt gtcctcgc gcaccacgc 142740  
cacagcccc cctatecgac cctggagctg ggtctggcgc ggcagccgt tctgegggtt 142800  
cgaacggcaa cgcgcgacta tgccggtctg tatgtcctgc gcgtatgggt cggcagcgcg 142860  
acgaacgcca gccggtttgt tttgggggtg gcgctctctg ccaacgggac gtttgtgat 142920  
aacggctcg actacggctc ctgcgatccg gcgcagcttc ccttttcggc cccgcgcctg 142980  
ggaccctega gcgtatacac ccccgagcc tcccagcca cccctccacg gacaacgaca 143040  
ccccgcctc cccccgaga cccgaccccc gccccgggg acacagggac gcccgcgccc 143100  
gcgagcggcg agatagcccc gcccaattcc acgcgatcgg ccagcgaatc gagacacagg 143160  
ctaaccgtag cccaggtaat ccagatcgcc ataccggcgt ccatcatcgc ctttgtggtt 143220  
ctgggcagct gtatctgctt catccataga tgccagcgc gatacaggcg cccccgcggc 143280  
cagatttaca accccggggg cgtttcctgc gcggtcaacg aggcggccat ggcgccctc 143340  
ggagccgagc tgcgatccca cccaaacacc cccccaaa cccgacgcg ttctgctgctg 143400  
tccaagacca tgccttccct aacgtcgata gctgaggaat cggagccagg tccagtcgtg 143460  
ctgctgtccg tcagtcctgc ccccccaggt ggcccagcgg ccccccaaga ggtctaggtc 143520  
caagcgggcc gttcggcagg cccgccccac cgcctccatc gtggttattt ccccccccc 143580  
cccccaata aaccgatgtt atttgcctat atgcgtgtgt tggatccctt tgtgatcgtt 143640  
cgtcattccc cggatggcat gggaggcggg taatggatgg gcggggcccg gggggaggaa 143700  
aaagaataaa gggggtagtg tcggagaggg ccgcccgcga tttaggaggt cgcgccccg 143760  
actctgtgc ttgggtgac ttggtgcgc gccgctcagct agtctccgat ctgccccgac 143820  
cgacggctcc tgccaccga acatggtcgc cggggccggg ttggtgtttt ttgttggagt 143880  
ttgggtcgtg tctgctcctg cggcagcacc cagaacgtcc tggaaacggg taacctcggg 143940  
cgaggacgtg gtgttcttc cggcgcgcc ggaacgcacc cgggcccaca aactactgtg 144000  
ggcccgga cccctggatg cctcgcgtcc cctgcgccc tcgtgggtgg cgtgtggcc 144060  
ccccgcagc gtgctcgaga cggctgtgga tgccgctgc atgcgcgcc cggaaaccgt 144120  
cgccatagca tacagtcgcc cgttccccgc gggcgacgag ggactgtatt cggagtggc 144180

---

-continued

---

gtggcgcgat cgcgtagccg tggccaacga gagtctggtc atctacgggg cccctggagac 144240  
ggacagcggg ctgtacaccc tgtccgtggg cggcctaagc gacgaggcgc gccaaagtggc 144300  
gtcgggtggt ctggctgtgg agcccgcgcc tgtgccgacc cggacccccg acgactacga 144360  
cgaagaagac gacgcggggc tgacgaaacg acgcccgtca gcgttcccc cccaaacccc 144420  
cccccgctgt cccccgctg ccccccgac gcaccctcgt gttatcccc aggtgtccca 144480  
cgtgcgcggg gtaacggctc atatggagac cctggaggcc attctgtttg cccccgggga 144540  
gacgtttggg acgaacgtct ccatccacgc cattgcccac gacgacggtc cgtacgccat 144600  
ggacgtcgtc tggatcggg ttgacgtgcc gtccctcgtc gccgatatgc ggatctacga 144660  
agcttgtctg tatcacccgc agcttcacga gtgtctatct ccggccgacg cgcctgtgctc 144720  
cgtaagtcc tgggcgtacc gcctggcggg ccgcagctac gccggctggt ccaggactac 144780  
gccccgcgcg cgtggtttg ccgaggctcg catggaaccg gtcccggggt tggcgtggct 144840  
ggcctccacc gtcaatctgg aattccagca gcctcccc cagcacgccg gcctctacct 144900  
gtgcgtggg tacgtggagc atcatatcca gcctggggc cacatgacca tcagcacccg 144960  
ggcgcagtac cgaacgcgc tgggtgaaaca gcacctccc cagcgcaccg ccgagcccgt 145020  
cgagcccacc cgcgcgcacg tgagagcccc ccatcccgcg ccctccgcgc gcggcccgt 145080  
gcgcctcggg gcgggtcctg gggcggccct gttgctggcc gccctcgggc tgtccgcctg 145140  
ggcgtgatg acctgtggt gcaggcgtc ctggcgggcg gttaaaagcc gggcctcggc 145200  
gacgggcccc acttacatc gcgtggcgga cagcagctg tacgaggact ggagttcgga 145260  
cagcaggggg gagcgcgacg ggtccctgtg gcaggaccct ccggagagac ccgactctcc 145320  
ctccacaat ggatccgggt ttgagatctt atcaccaacg gctccgtctg tataacccca 145380  
tagcgagggg cgtaaatctc gcccccgt caccacctt ggttcgggaa gcccgggccg 145440  
tcgtcactcc caggcctct atccgtccgt cctctggtaa ggcgtcttcc gacgacgcgg 145500  
acgtcggcga tgaactgatt gccatcgcgg acgcacgcgg ggaccgccca gagaccctgc 145560  
ccccgcgcgc gggcggcgc ccgcgccgt gccgcagacc acctcgggc ggtcccccg 145620  
cggcctttcc cgtggccctc cagcgcgtg acgccccct ccaattcgtc acctggctcg 145680  
ccgtgcctg gctgcggggg gcgggtgggc tcggggccgt cctgtgcggg attgcgtttt 145740  
acgtgacgtc aatcgcgcga ggccataaaa ggtccggcgg ccaacccgcc gcagctcata 145800  
aaaatcgtga gtcacggcaa cccaccttc gcctccgcc tccgcagcg ccttccgcg 145860  
tcccgatga cctcccggcc gcgcgaccaa gactcgggtc gttccagcgc gtcgggtgccc 145920  
ctttaccccg cggcctcgc cgtcccgga gaagcctact actcgaaaag cgaagacgag 145980  
gccccaacg acttccctg gcgcatggg ccgcagcagt cggctcctaa gcgcccagcg 146040  
cggcgcacgc ggtgcgtcgg gctgggtatc gcctgtctcg tcgtggccct cctatctgga 146100  
gggttcgggg cacttttggg gtggctgctc gcctaaatga gcctcgatg tatggcgcct 146160  
tcttcgcccc caccctcgc gcgacccac gtccgtatgt taattgcaat aaagtgggtg 146220  
attgtcatta cgtctacta ggtgtcttt ttttttggg ggggggggag gaaatgcaga 146280  
aaagggtaag aaattctcgg aatttcaccc ccgggggggg gcaagtgcag taaccagtt 146340  
cctcagtggt tgggaaatct attgaactct ccggctcct ccgtgttagg gaagtctctt 146400  
gggaaatct attgacctc gcccccccc ccccaggag gggggcagtg cagtaaccca 146460

-continued

---

gttctccgt gctgggaaa tctctctgcc gggtacgggc tccagacgaa ggaccatac 146520  
atttcccat ccgcaccca catctggcgt tctagagtca cgacgcattt gcccccgcc 146580  
ccgcagcaac acacaaagcg atttcaattt tcacgatttt attattaatt acaccaacca 146640  
ccctgtcccc gggacgtggt caggaccggg ggtccgcacc caaacgcacg aaacaaatgc 146700  
tggcagtgtg ccgaatataa ccccgctag gaacacgtcg acgcgtgccc caaacagcac 146760  
cagaaggcgc atgccatcag caggtcgtgc atatggcgat gtgtttggac gcagggcgca 146820  
gccgcggcga taaaattcat ggcggccgct cgcacgggcc acagcggcga ggactccctg 146880  
ttggcccga gccattgggt atgaaccagc tgcgcctcct gtcgcacctt ggctcccgc 146940  
agcggggcg gtgggtcgtg ggtgttgaga gcacacaggc gggacacctc gatcacctc 147000  
cgaaaaaag cccggtggtc gcggggcagc atctgcaggt gcgcagggc ctgggcgttg 147060  
agaggtaca actcggagcg gggggactcc gggggcgggt ccgcgcggtg ccgcgagttg 147120  
gcacgcttg gggcccgggt gtcggacgcg ggcgcgttat ggatcccgc gcggggcaga 147180  
acgtacgtgc gttggcggcg cgatgagggg tccggcgtgc cgaggggggc gtaggggacc 147240  
gggctaggca agcccgggg ttgcggggg ttcccgtggg ggtctaggct cctggggcac 147300  
ccgtgggggt cgtgggggtc gcgggtccct gggatgcgc gggaccctgg gttctctggg 147360  
agatcgtgga actcgcgggt cccctgggctc tgggggaacc cggggctccc tggggacacg 147420  
tgggtccctg ggaattcttg atggtcggac ggcttcagat ggcttcggga tcgagagggc 147480  
cgacagact cgtagtagac ccgaatctcc acgtttccc gccgcggat catggtcgc 147540  
gcccgggtgc gggggcccgt cggtcggaag cgagtccct tcaagcgtgt ccgctcctct 147600  
gggtgcctg ccgtcggatg ggggtccctt taaggaaagg tctcggctgc ccgcccac 147660  
cggggtttg ggtggggcg gggaaacccc ggatgccatg ggggggtcac accctaagcg 147720  
ccggcgcgt ggttgggtgg gggtagaggg gaggccccg tcgacgagat cgtatcaagg 147780  
ggccagcacg cgatcctgcc gctcgttcga tctagcacac ccacgggtct gctgtgtggg 147840  
attcgaatc gcgggatccg atcgcacgtc cggaggacac agcagcggga gctccgggtc 147900  
ggtaaccgca gttctggcgt cctctcggtc cteccgttc cttttatgga tctccgcga 147960  
gacatcgcca tacgtccggt gtgtgcaccc cgaagaatcc agaaacatgt ccgtcgtttt 148020  
cagggcccaa gacatgggtg cccgtccacg aaggcggcgc ccggcctcgc agaaagcgcg 148080  
gatgttggga tcggggcccc gtccccccg cccgtcccc cgtccccccg gcccgteccc 148140  
ccgtcccccc ggcctcctcc cccgtcccc cggccccgtc ccccgteccc ccggcccgtc 148200  
cccccgtecc cccgtcccc cgtccccccg tcccccgte ccccgteccc cccgtcccc 148260  
cgtecccccg tcccccgte cccccgtecc cccgtcccc cgtecccccg tcccccgte 148320  
cccccgtecc ccggcccccc cggccccccg gccccccggt cccccggccc gtecccccg 148380  
cccgtcccc ccggcccgtc cccccggccc tccccccggt cccgtcccc ggcccgtccc 148440  
cccggcccgt cccccggccc cgtecccccg gcccgtcccc cccgtcccc cccgtcccc 148500  
cggccggccc cccgggtcac cgtacctcgc ataagcgtgc agtgggtgga tgggtcctcg 148560  
cggtagctac aggggtgggg gggggggggg ggaggaaag gcagaacgaa aaggaaccga 148620  
tgcccccg cctctgtatc cgatccgac cgggtcgtc ggtgccccgc tcgcccggg 148680  
cgtctctgct tcgctgtggc cccctcgcgc atgcccgcgc tgcccgtccc gtctccgccc 148740

---

-continued

---

cgcagccggt gtgcccctgg tgcggcggcg accgggacgc cggcccttta tgtgcgcgag 148800  
gaacggcccc cccccctcc gggcccgcct cggggcggag cccgcgggat gacgcgggccc 148860  
ccgggcaggg cgcagtgct cgcactttgc cctaataata tataactat taggacgaag 148920  
tggaacgct tcgcgttctc acttctttta cctgcggccc cgcgccctt tggggcggag 148980  
cgcgggatga cgcgggcccc gggcagggcg ccagcgcctc cactttgcc taataatata 149040  
tatactatta ggacgaagtg gaaacgcttc gcgttctcac ttcttttacc ctgcggcccc 149100  
gcccccttg gggcggagcc ccccgcggac caacggggcg acctgcgccc cccctttggg 149160  
gcccgcgggg gccaacggga gcgcggggcc ggcatctcat taccacgaac ccggaaggcc 149220  
aggggagcga gcccgcgccc gacgagggtc tcattagcat cgcgggcgga agcgggaagcc 149280  
gcccgcgccc ggcgctaagt agatgcgccc cgggcggagc ggcggcggcg cgaccaacgg 149340  
gcccgcgcca cggacgcgga cgcgcggggc tggggcggg gccgcgata atcggttcc 149400  
acctgggggc ggaacccccg cgagccgggg cgcggcggcg tcgatcgtc ctctccgcg 149460  
tcctcctcct tccccccgc cccgcgccc ccgaggacta taccagccag gcgacggggc 149520  
gatcgtccac acggagcgcg gctaccgacg cggcccagc gatctaccg atcggcgcgg 149580  
agaggcgaag agacacagc acacgcacgc accgcacggg ggggagagag actgccaacc 149640  
acccccccc actgcccgcc ctgaagaaga agaagaagc cccccccc caccaccccg 149700  
tcggaggcga tgcgcccga gcagcggaa agaaagaaga gcagcagcag gacgcagggc 149760  
cgcggggcgc aggtcgcgat ggcggacgag gacgggggac gtctccgggc cgcggcggag 149820  
acgacccgcg gccccggatc tccggatcca gccgacggac cgcgcccac ccgaacccg 149880  
gaccgtgcc ccgcccgcg gccccgggtc ggggtggcac gtgggcccga ggagaacgaa 149940  
gacgagcccg acgaccccgc cgcgatgcc gatgcccagc agggcggccc ggcgtccggg 150000  
gaggccgctc acgagcctgc cgcggacggc gtcgtctcgc cgcggcagct ggccctgctg 150060  
gcctcgatgg tggacgagc cgttcgacg atcccgtcgc ccccccgga gcgagcggc 150120  
gccaagaag aagcggcccc ctcgcctct ccgcccggga cccctccat gcgcccgat 150180  
tatggcgagg agaacgacga gcagcagcag gacgacgat acgacgacc gcgcccggc 150240  
cgtgggtcc gcggaccgga gacgacgtc gcggtccgcg gggcgtacc ggacccatg 150300  
gccagcctgt cgcgcgacc cccggcggcc cgcgacacc accaccacca ccaccaccg 150360  
cgcggcgcg cccccgcg gcgctcggc gcctctgact catcaaatc cggatcctc 150420  
tcgtcggcgt cctccgcctc ctctccgc tcctcctcct cgtctgcat cgcctcctc 150480  
tctgacgac acgacgacga gcagcccgc cgcgccccg ccagcgcgc agaccacgc 150540  
gcgggcggga ccctcggcgc ggaacgacgag gaggcggggg tgcccgcgag ggcccgggg 150600  
cgcgcccgc ggcgagccc ccccagggcc gagcccgc cggcccggac ccccgcggc 150660  
accgcgggccc gcctggagc cgcgcgggccc cgcgcggcgg tggccggccc gcagcccacg 150720  
ggccgcttca cggcggggc gcccccggcg gtcgagctgg acgcccagc ggccctccggc 150780  
gccttctac gcgctaccg gcaggggtac gtcagcgggg agcctgggccc cggggcggc 150840  
ccccgcgccc cggggcgcgt gctgtacggc gggctgggcg acagcccgc cggcctctgg 150900  
gggggcgccc aggcggagga ggcgcgggccc cggttcgagg cctcgggccc cccggcggc 150960  
gtgtgggccc ccgagctggg gcagcgggccc cagcagtagc ccctgatcac gcggctgctg 151020

-continued

---

tacacgccgg acgccggaggc gatgggggtgg ctccagaacc cgcgcgtggc gcccggggac 151080  
gtggcgctgg accaggcctg ettccggatc tccggcgccg cgcgcaacag cagctccttc 151140  
atctccggca gcgtggcgcg ggccgtgccc cacctggggt acgccatggc ggccggccgc 151200  
ttcggctggg gcctggcgca cgtggcgccc gccgtggcca tgagccgccc ctacgaccgc 151260  
gcgcagaagg gcttctctgct gaccagcctg cgcgcgcct acgcgccct gctggcgcg 151320  
gagaacgcgg cgtgaccgg ggccgcaacc cccgacgacg cgcgacgcgc caaccgccac 151380  
gacggcgacg acgcccgcgg gaagcccgcc gccccgcgg ccccgttgcc gtcggcgggc 151440  
gcgtcgccgg ccgacgagcg cgcggtgccc gccggctaag cgcgcgcggg ggtgctcgcc 151500  
gccctggggc gcctgagcgc cgcgcccgcc tccgcgccgg ccggggccga cgacgacgac 151560  
gacgacgacg gcgcccggcg tgggtggcgg gcccgcgcg cggaggcggg ccgctgccc 151620  
gtggagtgcc tggccgccc cgcgggatac ctggaggcgc tggcgagggg cttcgacgcg 151680  
gacctggcgg ccgtgcccgg gctggccgga gcccgcccc cgcgccccc gcgcccgggg 151740  
cccgcggggc cggccgcccc gccgcacgcc gacgcgcccc gcctgcgcgc ctggctgccc 151800  
gagctgcccgt tcgtgcgcga cgcgctgggt ctgatgcgcc tgccggggga cctgcgcgtg 151860  
gccggcgcca gcgaggccc cgtggcccgc gtgcgcgccc tgagcctggt cgcggggccc 151920  
ctggggcccg cgtgcgccc gagcccgcgc ctgctgaget ccgcccgcgc cgcgcccgcg 151980  
gacctgctct tccagaacca gagcctgcgc cccctgctgg ccgacacgct cgcgcccgcc 152040  
gactcgctcg ccgcccgcgc ctccgcgccc cgggaggcgc gcaagcgaag gagcccgcgc 152100  
ccggccaggg ccgcccgggg cggcgcggcc cccccccga agaagagccc cgcggacgcc 152160  
cccgcgccc cggccgcccc tcccgccggg gcccgcccc ccgcccgcgc gacgcccgcg 152220  
ccgcccgcgc cgcgcccgcg ggcgctgacc cgcgcccgc cgcaggggccc cgaccccgcg 152280  
ggcggtggc gccgcccagc gccggggccc agccacacgc cggcgccctc ggccgcccgc 152340  
ctggaggcct actgcgcccc gcggggcgtg gccgagctca cggaccaccc gctcttcccc 152400  
gcgcccgtgg gcccgcccct catgttcgac ccgcccgcgc tggcctcctt ggccgcccgc 152460  
tgcccccgcc cgcgcccgcg cggcgcccgc gccgcttcg gcccgctgcg cgcctcgggc 152520  
ccgctgcgcc gcgcccggcg ctggatgcgc caggtgccc acccggagga cgtgcgcgtg 152580  
gtgatcctct actcgcgct gccggggcag gacctggccc cgggcccgc cggggcgggg 152640  
cccccccgg agtggctcgc cagcgcggc gggtgtcct gcctgctggc ggccctgggc 152700  
aacggctct gcgggcccgc cacggcccgc tgggggggca actggaccgg cgcgcccgc 152760  
gtctcgccgc tggcgccgca gggcgctgct ctgctgtcca cgcgggacct ggccctcgcc 152820  
ggcggcctgg agttcctggg gctgctggcc gccgctgccc accgcccct catcgtcgtc 152880  
aacgcccgtg gcgcccggga ctggcccgc gacgggccc tggctcgcgc gcagcagccc 152940  
tacctggcct gcgagggtct gcccccgtg cagtgcgccc tgccgtggcc ggccgcccgc 153000  
gacctgccc gcaccgtgct ggcccccgc cgcgtgttcg ggccgggggt cttcgcccgc 153060  
gtggaggccc cgcacgccc cctgtacccc gacgcccgc cgtgcccct ctgcccgggg 153120  
gccaacgtg ggtaccgct gcgacgccc ttcggcccc acacgctggt gcccatgtcc 153180  
ccgcccaggt accgcccgc cgtgctccc gcgctggac gccgggccc cgcctcgggc 153240  
gcggcgacg ccattggccc cggcgcgccc gacttctgcg aggacgaggc gacctcgccc 153300

-continued

```

cgcgccctgcg cgcgctgggg cctggggcgcg ccgctgcggc cgttctacgt ggcgctgggg 153360
cgcgacgccc tgccggggcg cccggcgagg ctgcgagggc cgcggcgga gttctgcgcg 153420
cgggcgctgc tcgagcccca cggcgacgcg cccccgctgg tgcgcgcgca cgacgaggac 153480
gcgggccccc ccccgcagat acgctggggc tggcgcgcg gcccgcgggg gacggtgctg 153540
gcccgggcgg gcggcgggct ggaggtgggt gggaccgccc cggggctggc caccgccc 153600
aggcgcgagc ccgtggacat ggacgaggag ctggaggagc acgacgagc actgtttggg 153660
gagtgcgggg gggggaact tccgggagcg ggggaggggg gagatgggga gagggggaag 153720
gaatcgggcg tctgtgcgcc ttaagacag acgcgcgat ggcgcgcgc gtgtgtgaga 153780
aataaagaac gagacagag aaaaactacc gccttgtgtg gtttattcgg ggtcggggcg 153840
ggcgggggtc gggcgggcgg gggcggggc ggcgggggtc gggcgggcgg ggtcggggcg 153900
ggcgggggtc gggcgggcgg gggcggggc ggcgggggtc gggcgggcgg ggtcggggcg 153960
ggcgggggtc gggcgggcgg gggcggggc ggcgggggtc gggcgggcgg ggtcggggcg 154020
ggcgggggtc gggcgggcgg gggcggggc ggcgggggtc gggcgggcgg cactctccc 154080
gcccccgcg ggggtctggg gctctgacct gactgcaggt tacgaagtc agtgggccc 154140
agccccccc caggagcggg agggaaggca cggggcgcg gagggagggg ctgctgcgag 154200
ctcggggcgg cggcgcgggg gggagggggc ggggaagccc cggggcgggg gcgcggggga 154260
ggcgcccgcg ggggagggcg ccgcgggacc gcagccccgt ggcgcgcggg ggggaggggc 154320
tgcccgagc tcggcgggg gagggggagg gaggggtgg cggggaaccg tgtgcggcg 154380
ggcggggtct tggtgcaact gtctggtctg cgagggcgag cgggtggtcg actggcgtct 154440
tcgggggggc ggggagcctg ggagtgtgtg gtggtctgcg gcacagcctg ctagtcccc 154500
tcctgccgcg cggggcgggg cgcgggaaaa aagcccgcg gggcgcccc cgggaaggca 154560
gcccccgggc gcgcgggggg agggcgggc ccccggggg agcggccggc tccgggggag 154620
ggacggggaa gggggcgcg ggggctgccc tgccgcccgc ccgcccgcgc cgcgccctt 154680
cgcgcccccc ccaaaaaaac acccccccg ggggttgact ccccggggga aaagagggcg 154740
ggcggg 154746

```

&lt;210&gt; SEQ ID NO 17

&lt;211&gt; LENGTH: 9181

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Human Immunodeficiency Virus HIV1 (NC\_001802.1)

&lt;400&gt; SEQUENCE: 17

```

ggtctctctg gtttagaccag atctgagcct gggagctctc tggctaacta gggaaaccac 60
tgcttaagcc tcaataaagc ttgccttgag tgcttcaagt agtgtgtgcc cgtctgttgt 120
gtgactctgg taactagaga tccctcagac ccttttagtc agtgtggaaa atctctagca 180
gtggcgccc aacagggacc tgaaagcgaa agggaaacca gaggagctct ctgcacgag 240
gactcggcct gctgaagcgc gcacggcaag aggcgagggg cggcgactgg tgagtacgcc 300
aaaaattttg actagcggag gctagaagga gagagatggg tgcgagagcg tcagtattaa 360
gcggggggaga attagatcga tgggaaaaaa ttcggttaag gccaggggga aagaaaaaat 420
ataaattaaa acatatagta tgggcaagca gggagctaga acgattcgca gttaatcctg 480
gcctgttaga aacatcagaa ggctgtagac aaatactggg acagctacaa ccatcccttc 540

```

-continued

---

agacaggatc agaagaactt agatcattat ataatacagt agcaaccctc tattgtgtgc	600
atcaaaggat agagataaaa gacaccaagg aagctttaga caagatagag gaagagcaaa	660
acaaaagtaa gaaaaaagca cagcaagcag cagctgacac aggacacagc aatcagggtca	720
gccaaaatta cccatagtg cagaacatcc aggggcaaat ggtacatcag gccatatcac	780
ctagaacttt aatgcatgg gtaaaagtag tagaagagaa ggctttcagc ccagaagtga	840
taccatggtt ttcagcatta tcagaaggag ccaccccaca agatttaaac accatgctaa	900
acacagtggg gggacatcaa gcagccatgc aaatgttaaa agagaccatc aatgaggaag	960
ctgcagaatg ggatagagtg catccagtgc atgcagggcc tattgcacca ggcagatga	1020
gagaaccaag gggaaagtac atagcaggaa ctactagtac ccttcaggaa caaataggat	1080
ggatgacaaa taatccacct atcccagtag gagaaattha taaaagatgg ataactcctgg	1140
gattaaataa aatagtaaga atgtatagcc ctaccagcat tctggacata agacaaggac	1200
caaaggaacc ctttagagac tatgtagacc ggttctataa aactctaaga gccgagcaag	1260
cttcacagga ggtaaaaaat tggatgacag aaaccttgtt ggtccaaaat gcgaaccag	1320
attgtaagac tatttttaaa gcattgggac cagcggctac actagaagaa atgatgacag	1380
catgtcaggg agtaggagga cccggccata aggcaagagt tttggctgaa gcaatgagcc	1440
aagtaacaaa ttcagctacc ataatgatgc agagaggcaa ttttaggaac caaagaaaga	1500
ttgttaagtg tttcaattgt ggcaagaag ggcacacagc cagaaatgc agggccccta	1560
ggaaaaaggg ctgttgaaaa tgtggaagg aaggacacca aatgaaagat tgtactgaga	1620
gacaggctaa ttttttaggg aagatctggc cttcctacaa gggaaggcca ggaattttc	1680
ttcagagcag accagagcca acagccccac cagaagagag cttcaggtct ggggtagaga	1740
caacaactcc cctcagaag caggagccga tagacaagga actgtatcct ttaactccc	1800
tcaggtcact ctttgcaac gaccctcgt cacaataaag ataggggggc aactaaagga	1860
agctctatta gatcaggag cagatgatac agtattagaa gaaatgagtt tgccaggaag	1920
atgaaacca aaaatgatg ggggaattgg aggttttacc aaagtaagac agtatgatca	1980
gatactcata gaaatctgtg gacataaagc tataggtaca gtattagtag gacctaccc	2040
tgtaacata attggaagaa atctgttgac tcagattggt tgcactttaa atttcccat	2100
tagccctatt gagactgtac cagtaaaatt aaagccagga atggatggcc caaaagttaa	2160
acaatggcca ttgacagaag aaaaaataaa agcattagta gaaatttga cagagatgga	2220
aaaggaaggg aaaatttcaa aaattgggcc tgaatatcca tacaatactc cagtatttgc	2280
cataaagaaa aaagacagta ctaaatggag aaaattagta gatttcagag aacttaataa	2340
gagaactcaa gacttctggg aagttcaatt aggaatacca catcccagc ggttaaaaaa	2400
gaaaaaatca gtaacagtac tggatgtggg tgatgcata ttttcagttc ctttagatga	2460
agacttcagg aagtatactg catttaccat acctagtata aacaatgaga caccagggat	2520
tagatatcag tacaatgtgc tccacaggg atggaagga tcaccagcaa tattccaaag	2580
tagcatgaca aaaatcttag agccttttag aaaacaaaat ccagacatag ttatctatca	2640
atacatggat gatttctatg taggatctga cttagaataa gggcagcata gaacaaaaat	2700
agaggagctg agacaacatc tgttgaggtg gggacttacc acaccagaca aaaaacatca	2760
gaaagaacct ccattccttt ggatgggtta tgaactccat cctgataaat ggacagtaca	2820

-continued

---

gcctatagtg	ctgccagaaa	aagacagctg	gactgtcaat	gacatacaga	agttagtggg	2880
gaaattgaat	tgggcaagtc	agatttacc	agggattaaa	gtaaggcaat	tatgtaaact	2940
ccttagagga	accaaagcac	taacagaagt	aataccacta	acagaagaag	cagagctaga	3000
actggcagaa	aacagagaga	ttctaaaaga	accagtacat	ggagtgtatt	atgacccatc	3060
aaaagactta	atagcagaaa	tacagaagca	ggggcaaggc	caatggacat	atcaaattta	3120
tcaagagcca	tttaaaaatc	tgaaaacagg	aaaatatgca	agaatgaggg	gtgcccacac	3180
taatgatgta	aaacaattaa	cagaggcagt	gcaaaaaata	accacagaaa	gcatagtaat	3240
atggggaaag	actcctaaat	ttaaactgcc	catacaaaag	gaaacatggg	aaacatggtg	3300
gacagagtat	tggcaagcca	cctggattcc	tgagtgggag	tttgttaata	cccctccctt	3360
agtgaaatta	tgttaccagt	tagagaaaga	acctatagta	ggagcagaaa	ccttctatgt	3420
agatggggca	gctaacaggg	agactaaatt	aggaaaagca	ggatatgtta	ctaatagagg	3480
aagacaaaaa	gttgtcacc	taactgacac	aacaaatcag	aagactgagt	tacaagcaat	3540
ttatctagct	ttgcaggatt	ggggattaga	agtaaacata	gtaacagact	cacaatatgc	3600
attaggaatc	attcaagcac	aaccagatca	aagtgaatca	gagttagtca	atcaaataat	3660
agagcagtta	ataaaaaagg	aaaaggctca	tctggcatgg	gtaccagcac	acaagggaat	3720
tggaggaaat	gaacaagtag	ataaattagt	cagtgtctga	atcaggaaag	tactattttt	3780
agatggaata	gataaggccc	aagatgaaca	tgagaaatat	cacagtaatt	ggagagcaat	3840
ggctagtgat	tttaacctgc	cacctgtagt	agcaaaagaa	atagtagcca	gctgtgataa	3900
atgtcagcta	aaaggagaag	ccatgcatgg	acaagtagac	tgtagtccag	gaatatggca	3960
actagattgt	acacatttag	aaggaaaagt	tatcctggta	gcagttcatg	tagccagtgg	4020
atatatagaa	gcagaagtta	tccagcaga	aacagggcag	gaaacagcat	attttctttt	4080
aaaattagca	ggaagatggc	cagtaaaaac	aatacatact	gacaatggca	gcaatttcac	4140
cggtgctacg	gttagggccg	cctgttggg	ggcgggaatc	aagcaggaat	ttggaattcc	4200
ctacaatccc	caaagtcaag	gagtagtaga	atctatgaat	aaagaattaa	agaaaattat	4260
aggacaggta	agagatcagg	ctgaacatct	taagacagca	gtacaaatgg	cagtattcat	4320
ccacaatttt	aaaagaaaag	gggggattgg	ggggtacagt	gcaggggaaa	gaatagtaga	4380
cataatagca	acagacatac	aaactaaaga	attacaaaaa	caaattacaa	aaattcaaaa	4440
ttttcggggt	tattacaggg	acagcagaaa	tccactttgg	aaaggaccag	caaagctcct	4500
ctggaaaagg	gaaggggcag	tagtaataca	agataatagt	gacataaaag	tagtgccaag	4560
aagaaaagca	aagatcatta	gggattatgg	aaaacagatg	gcaggtgatg	attgtgtggc	4620
aagtagacag	gatgaggatt	agaacatgga	aaagttagt	aaaacacat	atgtatgttt	4680
cagggaaagc	taggggatgg	ttttatagac	atcactatga	aagcctcat	ccaagaataa	4740
gttcagaagt	acacatccca	ctaggggatg	ctagattggt	aataacaaca	tattggggtc	4800
tgcatacagg	agaaagagac	tggcatttgg	gtcagggagt	ctccatagaa	tggaggaaaa	4860
agagatatag	cacacaagta	gacctgaac	tagcagacca	actaattcat	ctgtattact	4920
ttgactgttt	ttcagactct	gctataagaa	agccttatt	aggacacata	gttagcccta	4980
gggtgtgaata	tcaagcagga	cataacaagg	taggatctct	acaatacttg	gcactagcag	5040
cattaataac	acaaaaaag	ataaagccac	ctttgcctag	tgttacgaaa	ctgacagagg	5100

-continued

---

atagatggaa caagccccag aagaccaagg gccacagagg gagccacaca atgaatggac	5160
actagagcct ttagaggagc ttaagaatga agctgttaga cttttccta ggatttgget	5220
ccatggctta gggcaacata tctatgaaac ttatggggat acttgggagc gagtgggaagc	5280
cataataaga attctgcaac aactgctggt tatccatctt cagaattggg tgcgcacata	5340
gcagaatagg cgttactcga cagaggagag caagaaatgg agccagtaga tcctagacta	5400
gagccctgga agcatccagg aagtcagcct aaaactgctt gtaccaattg ctattgtaaa	5460
aagtgttctt ttcattgcca agtttgtttc ataacaaaag ccttaggcat ctctatggc	5520
aggaagaagc ggagacagcg acgaagagct catcagaaca gtcagactca tcaagcttct	5580
ctatcaaacg agtaagtagt acatgtaatg caacctatac caatagtagc aatagtagca	5640
ttagtagtag caataataat agcaatagtt gtgtgggtcca tagtaatcat agaataatag	5700
aaaatattaa gacaaagaaa aatagacagg ttaattgata gactaataga aagagcagaa	5760
gacagtggca atgagagtga aggagaaata tcagcacttg tggagatggg ggtggagatg	5820
gggaccatg ctccctggga tgttgatgat ctgtagtgct acagaaaaat tgtgggtcac	5880
agtctattat ggggtacctg tgtggaagga agcaaccacc actctatctt gtgcatcaga	5940
tgctaaagca tatgatagc aggtacataa tgtttgggcc acacatgctt gtgtaccac	6000
agacccaac ccacaagaag tagtattggt aaatgtgaca gaaaatttta acatgtggaa	6060
aaatgacatg gtagaacaga tgcacatgga tataatcagt ttatgggatc aaagcctaaa	6120
gccatgtgta aaattaacc cactctgtgt tagtttaaag tgcactgatt tgaagaatga	6180
tactaatacc aatagtagta gggggagaat gataatggag aaaggagaga taaaaactg	6240
ctctttcaat atcagcaca gacataagagg taaggtgcag aaagaatatg cttttttta	6300
taaacttgat ataataccaa tagataatga tactaccagc tataagttga caagttgtaa	6360
cacctcagtc attacacagc cctgtccaaa ggtatccttt gagccaattc ccatacatta	6420
ttgtgccccg gctggttttg cgattctaaa atgtaataat aagacgttca atggaacagg	6480
accatgtaca aatgtcagca cagtacaatg tacacatgga attaggccag tagtatcaac	6540
tcaactgctg ttaaatggga gtctagcaga agaagaggta gtaattagat ctgtcaattt	6600
cacggacaat gctaaaacca taatagtaca gctgaacaca tctgtagaaa ttaattgtac	6660
aagaccaac aacaatacaa gaaaaagaat ccgtatccag agaggaccag ggagagcatt	6720
tgttacaata ggaaaaatag gaaatatgag acaagcacat tgtaacatta gtagagcaaa	6780
atggaataac actttaaaac agatagctag caaattaaga gaacaatttg gaaataataa	6840
aacaataatc ttaagcaat cctcaggagg ggaccagaa attgtaacgc acagttttaa	6900
ttgtggaggg gaatttttct actgtaattc aacacaactg tttaatagta cttggtttaa	6960
tagtacttgg agtactgaag ggtcaaataa cactgaagga agtgacacaa tcacctccc	7020
atgcagaata aaacaatta taaacatgtg gcagaaagta ggaaaagcaa tgtatgcccc	7080
tcccatcagt ggacaaatta gatgttcatc aaatattaca gggctgctat taacaagaga	7140
tggtggtaat agcaacaatg agtccgagat cttcagacct ggaggaggag atatgaggga	7200
caattggaga agtgaattat ataaatataa agtagtaaaa attgaacct taggagtagc	7260
accaccaag gcaaaagaaa gagtgggtgca gagagaaaa agagcagtg gaaataggagc	7320
ttgttctctt gggttcttgg gagcagcagg aagcactatg ggcgcagcct caatgacgct	7380

-continued

---

```

gacggtacag gccagacaat tattgtctgg tatagtgcag cagcagaaca atttgctgag 7440
ggctattgag gcgcaacagc atctgttgca actcacagtc tggggcatca agcagctcca 7500
ggcaagaatc ctggctgtgg aaagatacct aaaggatcaa cagctcctgg ggatttgggg 7560
ttgctctgga aaactcattt gcaccactgc tgtgccttgg aatgctagtt ggagtaataa 7620
atctctggaa cagatttggc atcacacgac ctggatggag tgggacagag aaattaacaa 7680
ttacacaagc ttaatacact ccttaattga agaatcgcaa aaccagcaag aaaagaatga 7740
acaagaatta ttggaattag ataatgggc aagtttggg aattggttta acataacaaa 7800
ttggctgtgg tatataaaat tattcataat gatagtagga ggcttggtag gtttaagaat 7860
agtttttctg gtactttcta tagtgaatag agttaggcag ggatattcac cattatcgtt 7920
tcagaccac ccaccaacc cgaggggacc cgacaggccc gaaggaatag aagaagaagg 7980
tggagagaga gacagagaca gatccattcg attagtgaac ggatccttgg cacttatctg 8040
ggacgatctg cggagcctgt gcctcttcag ctaccaccgc ttgagagact tactcttgat 8100
tgtaacgagg attgtggaac ttctgggacg caggggggtg gaagccctca aatattggtg 8160
gaatctccta cagtattgga gtcaggaact aaagaatagt gctgttagct tgctcaatgc 8220
cacagccata gcagtagctg aggggacaga tagggttata gaagtagtac aaggagcctg 8280
tagagctatt cgccacatac ctagaagaat aagacagggc ttggaaagga ttttgcata 8340
agatgggtgg caagtggcga aaaagtagtg tgattggatg gcctactgta agggaaagaa 8400
tgagacgagc tgagccagca gcagataggg tgggagcagc atctogagac ctggaaaaac 8460
atggagcaat cacaagtagc aatacagcag ctaccaatgc tgcttgtgcc tggctagaag 8520
cacaagagga ggaggagggt ggttttccag tcacacctca ggtacctta agaccaatga 8580
cttacaaggc agctgtagat cttagccact ttttaaaga aaagggggga ctggaagggc 8640
taattcactc ccaagaaga caagatatcc ttgatctgtg gatctaccac acacaaggct 8700
acttccctga ttagcagaac tacacaccag ggccaggggt cagatatcca ctgaccttg 8760
gatggtgcta caagctagta ccagttgagc cagataagat agaagaggcc aataaaggag 8820
agaacaccag cttgttacac cctgtgagcc tgcattggat ggatgaccgc gagagagaag 8880
tgtagagtg gaggtttgac agccgcctag catttcatca cgtggcccga gagctgcatc 8940
cggagtactt caagaactgc tgacatcgag cttgctacaa gggactttcc gctggggact 9000
ttccaggag gcgtggcctg ggcgggactg gggagtggcg agccctcaga tctgcatat 9060
aagcagctgc tttttgcctg tactgggtct ctctggttag accagatctg agcctgggag 9120
ctctctggct aactagggaa cccactgctt aagcctcaat aaagcttgc ttagtgctt 9180
c 9181

```

```

<210> SEQ ID NO 18
<211> LENGTH: 10359
<212> TYPE: DNA
<213> ORGANISM: Human Immunodeficiency Virus HIV2 (NC_001722.1)

```

```

<400> SEQUENCE: 18

```

```

tgcaagggat gttttacagt aggaggagac atagaatcct agacatatac ctagaaaaag 60
aggaagggat aataccagat tggcagaatt atactcatgg gccaggagta aggtacccaa 120
tgacttcgg gtggctgtgg aagctagat cagtagaact ctcacaagag gcagaggaag 180

```

-continued

---

atgaggccaa ctgcttagta caccagcac aaacaagcag acatgatgat gagcatgggg	240
agacattagt gtggcagttt gactccatgc tggcctataa ctacaaggcc ttcactctgt	300
accagaaga gtttgggcac aagtcaggat tgccagagaa agaatggaag gcaaaactga	360
aagcaagagg gataccatat agtgaataac aggaacaacc atacttggtc aaggcaggaa	420
gtagtacta agaaacagct gaggtgcag ggactttcca gaaggggctg taaccaaggg	480
aggacatgg gaggagctgg tggggaacgc cctcactt actgtataaa tgtaccgct	540
tcttgactg tattcagtcg ctctgaggag aggetggcag atcgagcct gagagttct	600
ctccagcact agcaggtaga gcctgggtgt tccctgctgg actctacca gtacttgcc	660
ggactgggc agacggctcc acgcttgctt gcttaagac ctctcaata aagctgccag	720
ttagaagcaa gttaagtgtg tgttccatc tctcctagtc gccgcctgg cattcgggtg	780
tcatctgagt aacaagacc tggctgtgta ggaccttct cgcttgggga atccaaggca	840
ggaaaatccc tagcaggttg gcgcccgaac agggacttga agaggactga gaagcctgg	900
aactcgctg agtgaaggca gtaagggcgg caggaacaaa ccacgacgga gtgctcctag	960
aaagggcgg gccaggtac caaagggcgc gtgtggagcg ggagtgaag aggcctccgg	1020
gtgaagtaa gtacctacac caaaaactgt agccagaaaa ggctgttat cctacctta	1080
gacaggtaga agattgtggg agatgggccc gagaaactcc gtcttgagag ggaaaaagc	1140
agacgaatta gaaaaagtta ggttacggcc cggcggaaag aaaaagtaca ggttaaaaca	1200
tattgtgtgg gcagcgaatg aattggataa attcggattg gcagagagcc tgttgagtc	1260
aaaagaaggt tgccaaaaga ttctcagagt tttagatcca ttagtaccaa cagggtcaga	1320
aaatttaaaa agccttttta ataccgtctg cgtcatttgg tgcttgcacg cagaagagaa	1380
agtgaagat actgaggaag caaagaaact agcacagaga catctagtgg cagaaactgg	1440
aactgcagag aaaatgccc atacaagtag accaacagca ccactagtg gaaaaagagg	1500
aaactacccc gtgcaacaag cgggtggcaa ctatgtccat gtgccactga gccccgaac	1560
tctaaatgca tgggtaaaat tagtggagga aaagaagttc ggggcagaag tagtgccagg	1620
attcaggca ctctcagaag gctgcacgcc ctatgatatt aatcaaatgc ttaattgtgt	1680
ggcgatcac caagcagcta tgcaataat cagagagatt attaataag aagcagcaga	1740
ctgggattcg cagcacccaa taccaggccc cttaccagca ggacagctca gagaccaag	1800
agggctgac atagcaggaa caacaagcac agtagatgaa cagatccagt ggatgtatag	1860
gccacaaaat cccgtaccgg tagggaacat ctacagaaga tggatccaaa tagggctgca	1920
aaagtgtgtc agaaagtaca acccaactaa catcttagac ataaacagg gacaaaaga	1980
accgttccaa agctatgtag acaggttcta caaaagcttg agggcagaac aaacagaccc	2040
agcagtaaaa aattggatga cccaaacgct gctaatacag aatgccaacc cagactgcaa	2100
gtagtacta aaaggactgg ggatgaatcc caccctagaa gagatgctaa ccgcctgcca	2160
gggggtaggc ggaccaggcc agaaagccag gctaattggct gaagccctaa aagaggctat	2220
gggaccaagc cctatcccat ttgcagcagc ccaacaaaga aaggcaatta ggtattgaa	2280
ctgtgaaag gagggacact cggcaagaca gtgccgagca cctagaagac agggctgctg	2340
gaagtgtggc aagccaggac acatcatggc aaactgcccc gaaagacagg caggttttt	2400
agggttgggc ccacgggaa agaagcctcg caactcccc gtgaccaag cccctcaggg	2460

-continued

---

gctgatacca acagcacctc cggcagatcc agcagcggaa ctggtggaga gatatatgca	2520
gcaagggaga aagcagaggg agcagagggga gagaccatac aaagaggtga cggaggactt	2580
gctgcacctc gagcagagag agacacctca cagagaggag acagaggact tgctgcacct	2640
caattctctc tttgaaaaag accagtagtc acagcgtaca tcgaggatca gccggtagaa	2700
gtcttactag acacaggggc tgatgactca atagtagcag gaatagaatt aggggacaat	2760
tacactccaa aaatagtagg gggaataggg ggatttataa acaccaaaga atacaaaaat	2820
gtagaataa aagtactaaa taaaagagta agagccacca taatgacagg agatacccca	2880
atcaacatct ttggcagaaa tattctgaca gccttaggca tgtcattaaa tttaccagtt	2940
gccaaagatag agccaataaa agtaacattg aagccagggga aagatggacc aaggctgaaa	3000
caatggcccc taacaaaaga gaaaatagaa gcactaaaag agatctgtga aaaaatggaa	3060
aaagagggcc agctagaaga ggcacctcca actaatcctt ataatacccc cacatttgca	3120
attaagaaaa aggacaagaa caaatggagg atgctgatag attttagaga actaataaag	3180
gtgactcaag atttcacaga aattcagcta ggaattccac acccggcagg actagccaaa	3240
aagaaaagga tctctatatt agatgtaggg gatgcctatt tttccatacc actacatgaa	3300
gattttaggg agtatactgc atttacccta ccagcagtaa acaatatgga accaggaaaa	3360
agatataat ataaagtctt gccacaagga tggaaagggat caccagcaat tttcaatac	3420
acaatgaggg aagtcttaga acctttcaga aaagcaaacc cagatgtcat tctcatccag	3480
tacatggatg atattctaat agctagtgc aggacaggtt tagagcatga caaagtggtc	3540
ctgcagctaa aagaacttct aaatggccta gggttttcta ctccagatga gaagttccaa	3600
aaagaccctc catttcaatg gatgggtgt gaactatggc caactaaatg gaagctgcag	3660
aaactacaac tgccccagaa agacatatgg acagtcaatg acatccaaaa gctagtggga	3720
gtcttaaatt gggcggcaca aatctattca ggaataaaaa ccaaacactt atgtagacta	3780
attagaggaa aaatgacact cacagaagaa gtgcagtggga cagaactagc agaagcagag	3840
ctagaagaaa acaaaattat cttgagccag gaacaagaag gatattatta ccaagaagaa	3900
aaagaattag aggcaacaat ccaaaaaagc caaggacatc aatggacata caaaatacac	3960
caggaagaga aaatcctaaa agtaggaaag tatgcaaaga taaaaatac ccataccaat	4020
ggggtcagat tactagcaca ggtagtccag aaaataggaa aagaggcact agtcatttgg	4080
ggacggatac caaaatttca cctgccagtg gagagagaga cctgggagca gtggtgggat	4140
aactactggc aagtgcacatg gatccagag tgggactttg tatctacccc accactggtc	4200
aggttaacat ttaacctagt agggatcct ataccaggcg cagagacctt ctacacagat	4260
ggatcatgca atagacagtc aaaagagggga aaagcaggat atgtaacaga tagaggaaaa	4320
gacaaagtaa aagtattaga acaaaactacc aatcagcagg cagaattaga agtccttcgg	4380
atggcactgg cagactcagg cccaaagggt aatatcatag tagattcaca gtatgtaatg	4440
gggatagtag caggccagcc aacagagtca gaaaatagaa tagtgaacca gatcatagaa	4500
gaaatgataa agaaggaagc agtctatggt gcatgggtcc cagcccataa aggcatagga	4560
ggaaaccagg aagtagacca tttagtaagt caaggcatca gacaagtatt attcctggaa	4620
aagatagagc ccgctcaaga ggaacatgaa aaatatcata gcattataaa agaactaacc	4680
cataaatttg gaataccctt tctagtagca agacagatag taaactcatg tgcccaatgc	4740

-continued

---

caacagaaag gagaagccat acatgggcaa gtaaatgcag aaataggcgt ttggcaaatg	4800
gactacacac acttagaagg aaaaatcatt atagtagcag tacatgttgc aagtggattc	4860
atagaagcag aagtcattccc acaggaatca ggaaggcaga cagcactctt cctattaaaa	4920
ctggccagta ggtggccaat aacgcacttg cacacagaca atggcccaaa cttaacttca	4980
caggaagtga agatggtggc atggtgggta ggtatagaac aatcctttgg agtaccttac	5040
aaccacaaaa gccagggagt agtagaagca atgaatcacc acctaaagaa tcagataagt	5100
agaattagag aacaggcaaa tacaatagaa acaatagtac tgatggcagt tcattgcatg	5160
aattttaaaa gaaggggagg aataggggat atgacccag cagaaagact aatcaacatg	5220
attaccacag aacaagaaat acaattcctc caaagaaaa attcaaattt taaaaatttc	5280
caggtctatt acagagaagg cagagatcag ctgtggaaag gacctggtga actactgtgg	5340
aagggagaag gagcagtcag agtcaaggta gggacagaca taaaagtagt accaagaagg	5400
aaggccaaga ttatcagga ctatggagga agacaggaac tggatagtag tccccactg	5460
gagggtgcca gggaggatgg agaaatggca tgccttgtc aagtacctga aatacagaac	5520
aaaagaccta gaggaggtgc gctatgttcc ccaccacaag gtgggatggg catggtggac	5580
ttgcagcagg gtaatatcc cactacaagg aaaaagtcag ctagaatac aggcattatg	5640
gaacctaaac ccagaaaaag gatggctctc ctctcatgca gtaagattaa cctggtatac	5700
agaaaagttc tggacagatg ttaccccaga ctgtgcagac atcctaatac atagcactta	5760
tttctcttgc tttacggcag gtgaagtaag aagagccatc agaggggaaa agttattgtc	5820
ctgtgcaac tatcccacag ctcataaagc acaggtacca tcaactcaat acctagccct	5880
agtagtagta caacaaaatg acagacccca gagaaagggt accgccagga aacagtggag	5940
aagagacatc tggagaggcc ttcgagtggc tagagaggac catagaagcc ttaaacaggg	6000
aggcagtga ccatctgccc cgagagctca tttccagggt gtggcaagg tcctggagat	6060
attggcatga tgaacaaggg atgtcagcaa gctacacaaa gtatagatat ttgtgcctaa	6120
tgcaaaaagc tatatttaca catttcaaga gaggggtgcac ttgctggggg gaggacatgg	6180
gccgggaagg attggaagc caaggacctc cccctctctc ccctccaggt ctagtetaat	6240
gactgaagca ccaacagagt ttccccaga agatgggacc ccacggaggg acttagggag	6300
tgactgggta atagaaactc tgagggaaat aaaggaagaa gccttaagac attttgatcc	6360
ccgcttgcta attgctcttg gctactatat ccataataga catggagaca cccttgaagg	6420
cgccagagag ctcattaaaa ccctacaacg agccctcttc gtgcactca gagcgggatg	6480
taaccgctca agaattggct aaacaaggag aagaactcct tgcccagctg caccgacctc	6540
tagagccatg cactaacaaa tgctattgta agcagatgcag tttccattgc cagctgtgtt	6600
tctcgaaaaa ggggctcgga atatcatatg agcgaagggt cagacgaaga aggactccaa	6660
ggaaaactaa gactccttcg cctctctgac cagacaagtg agtatggagc ctggtaggaa	6720
tcagctgttt gttgtcattt tactaacaag tgcttgctta gtatattgta gccagtatgt	6780
gactgttttc tatggcatac ccgctgggaa aaatgcactc attcccttat tttgtgcaac	6840
taaaaataga gacacttggg ggaccatata gtgcttgcca gacaatgatg attatcagga	6900
aataatttta aatgtgacag aggcttttga tgcattggaat aatacagtga cagaacaagc	6960
agtagaagat gtctggcatc tatttgagac atcaataaaa ccattgtgtca agtaaacacc	7020

-continued

---

tctatgtgtg gcaatgaatt gtagcagggg tcaaggggat accacgaccc cgaatcccag	7080
gacctcgagt tccacaacct cgagaccacc cacatccgca gcctccataa taaatgaaac	7140
ttctaactgc atagaaaaca acacatgcgc aggattaggg tatgaggaga tgatgcaatg	7200
tgagttcaat atgaaggggt tagaacaaga taagaaaagg aggtataagg acacatggta	7260
tttagaagat gtggtttgtg acaacacaac agctggcaca tgttacatga gacattgcaa	7320
cacatcaatc atcaaagagt catgtgataa gcactattgg gatgctatga ggtttagata	7380
ctgtgacca ccgggctttg ccctattaag atgtaatgat accaactatt caggctttga	7440
acctaagtgc actaaagtag tagctgcttc atgcacaagg atgatggaaa cgcaaacttc	7500
tacttggttt ggctttaatg gcactagagc agaaaataga acatatactc attggcatgg	7560
cagagataat aggactatca ttagcttaaa caagtattat aatctcaca tgctgtgtaa	7620
gagaccagga aataagacag ttttaccat aacacttatg tcaggattag tgtttcactc	7680
tcagccaatc aacacaaggc ctaggcaggc atggtgccgg tttggaggca gatggaggga	7740
agccatgcag gaggtgaagc aaacccctgt acaacatccc agatacaaag gaatcaatga	7800
tacagggaaa attaaactta cgaaaccggg agcaggctca gaccgggaag tggcatttat	7860
gtggactaac tgacagggag aatttctcta ctgtaacatg acttgggtcc tcaattgggt	7920
agaagacaag aaccaaaccac ggcgcaacta ttgccatata aagcagataa ttaatacctg	7980
gcataaagta gggaaaatg tatatttgcc tcctagggaa ggggagtggt cctgtgaatc	8040
aacagtaacc agcataattg ctaacattga catagataaa aatcggactc ataccaacat	8100
taccttagt gcagaagtgg cagaactgta ccgattagaa ctgggagact acaaatat	8160
agaaaataca ccaattggct tcgcacctac agatcagaga aggtactcct caactccagt	8220
gaggaacaaa agaggtgtgt tcgtgctagg gttcttgggt tttctcgcga cagcaggtc	8280
tgcaatgggc gcgcggtccc tgacgctgtc agcccagtc cggactttac tggccgggat	8340
agtgcagcaa cagcaacagc tgttgacgt agtcaagaga caacaagaaa tgttgcgact	8400
gaccgtctgg ggaacgaaaa acctccaggc aagagtcact gctatcgaga agtacctaaa	8460
gcacagggca cagctaaatt catggggatg tgcgtttaga caggtctgcc aactactgt	8520
accgtgggta aatgactctt tatcgcctga ctggaaaaat atgacatggc aggagtggga	8580
gaaacaagtc cgtacactag aggcaaatat cagtcaaagt ttagaagaag cccaaattca	8640
acaagaaaag aatagtatg aattacaaaa attaaatagc tgggatattc ttggcaactg	8700
gtttgactta acctcctggg tcaagtatat tcaatatgga gtgcatatag tagtgggaat	8760
aatagcttta agaatagcaa tctatgtagt gcaattgtta agtagattta gaaagggcta	8820
taggcctggt ttctcttccc cccccggta tctccaacag atccatatcc acaaggaccg	8880
gggacagcca gccaacgaag gaacagaaga agacgtcggg gggcagctg gttacgactt	8940
gtggccttgg ccaataaact atgtgcagtt cctgatccac ctactgactc gcctcttgat	9000
cgggctatac aacatctgca gagacttact atccaagaac tccccgaccc gccgactgat	9060
ctcccagagt ctaacagcaa tcagggactg gctgagactt aaggcggccc aactgcaata	9120
tgggtgagag tggatccaag aagctttcca agcattcgcg aggactacga gagagactct	9180
tgcgggagcg tggggatggt tatgggaagc agcgcgagc atcgggaggg gaatactcgc	9240
agttccaaga agaactcaggc agggagcaga actcgcctc ctgtgaggga cagcagtatc	9300

-continued

---

```

agcagggaga gtacatgaac agcccatgga gaaacccagc aacagaaaga cagaaagatt 9360
tgtataggca gcaaaatgat gatgatgtag attctgatga tgatgaccta ataggagttc 9420
ctgttacacc aagagtacca cggagagaaa tgacctataa attggcaata gatatgtcac 9480
atthtataaa agaaaaaggg ggactgcaag ggatgtttta cagtaggagg agacatagaa 9540
tcctagacat atacctagaa aaagaggaag ggataatacc agattggcag aattatactc 9600
atgggcccagg agtaaggtag ccaatgtact tcgggtggct gtggaagcta gtatcagtag 9660
aactctcaca agaggcagag gaagatgagg ccaactgctt agtacacca gcacaaacaa 9720
gcagacatga tgatgagcat ggggagacat tagtgtggca gtttgactcc atgctggcct 9780
ataactacaa ggcccttcaact ctgtaccag aagagtttgg gcacaagtca ggattgccag 9840
agaagaatg gaaggcaaaa ctgaaagcaa gagggatacc atatagttaa taacaggaac 9900
aaccatactt ggtcaaggca ggaagtact actaagaaac agctgaggct gcagggactt 9960
tcagaaggg gctgtaacca agggaggagc atgggaggag ctggtgggga acgccctcat 10020
acttactgta taaatgtacc cgcttcttgc attgtattca gtcgctctgc ggagaggctg 10080
gcagatcgag cctgagagg ttctctccag cactagcagg tagagcctgg gtgttccctg 10140
ctggactctc accagtactt ggccggtact gggcagacgg ctccacgctt gcttgcttaa 10200
agacctctc aataaaactg ccagttagaa gcaagttaag tgtgtgttcc catctctcct 10260
agtcgccgcc tggtcattcg gtgttcatct gagtaacaag accctggtct gttaggacct 10320
ttctcgcttt gggaatccaa ggcaggaaaa tccttagca 10359

```

&lt;210&gt; SEQ ID NO 19

&lt;211&gt; LENGTH: 3215

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Hepatitis B Virus HBV (NC\_003977.1)

&lt;400&gt; SEQUENCE: 19

```

ctccacaaca tccaccaag ctctgctaga tcccagagt aggggctat attttctctg 60
tgggtgctcc agttccgaa cagtaaaacc tgttccgact actgcctcac ccatatcgtc 120
aatcttctcg aggactgggg accctgcacc gaacatggag agcacaacat caggattcct 180
aggaccctg ctctgtttac aggcgggggt tttcttgttg acaagaatcc tcacaatacc 240
acagagtcta gactcgtggt ggacttctct caatthtcta gggggagcac ccacgtgtcc 300
tggccaaaat tcgcagtccc caacctccaa tcaactacca acctcttctc ctccaactg 360
tcctggctat cgtggatgt gtctgcggcg ttttatcata ttcctctca tcctgctgct 420
atgcctcacc ttcttgttgg ttcttctgga ctaccaaggt atgttgcccg tttgtcctct 480
acttcagga acatcaacta ccagcacggg accatgcaga acctgcacga ttctgctca 540
aggaaacctc atgtttcctt cttgttctg tacaaaaact tcggacggaa actgcacttg 600
tattccatc ccatcatcct gggctttctc aagattccta tgggagtggg cctcagtcctg 660
tttctctggt ctacgtttac tagtgccatt tgttcagtag ttcgtagggc tttccccac 720
tgtttggctt tcagctatat ggatgatgtg gtattggggg ccaagtctgt acaacatctt 780
gagtcctctt ttacctctat taccattht cttttgtctt tgggtatata tttgaacctt 840
aataaaacca aacgttgggg ctactccctt aacttcatgg gatatgtaat tggagttgg 900
ggtactttac cgcaggaaca tattgtacaa aaactcaagc aatgtttctg aaaattgctt 960

```

-continued

---

gtaaatagac ctattgattg gaaagtatgt caaagaattg tgggtctttt gggccttctg	1020
gcccccttta cacaatgtgg ctatcctgcc ttgatgcctt tatatgcatg tatacaatct	1080
aagcaggctt tcactttctc gccaaacttac aaggccttct tgtgtaaaca atatctaaac	1140
ctttaccocg ttgcccggca acggtcagggt ctctgccaaag tgtttgctga cgcaaccccc	1200
acgggttggg gcttggccat agggccatcgg cgcctgcctg gaacctttgt ggctcctctg	1260
ccgatccata ctgcccgaact cctagcagct tgttttgctc gcagccgggc tggagcgaaa	1320
cttatcggaa ccgacaactc agttgtcctc tctcggaaat acacctcctt tccatggctg	1380
ctaggctgtg ctgccaactg gatcctgcgc gggacgtcct ttgtctacgt cccgtcggcg	1440
ctgaatcccg cggacgacct gtctcggggc cgtttggggc tctaccgtcc ccttcttcat	1500
ctgcccgtcc ggccgaccac ggggcgcacc tctctttacg cggctcctcc gtctgtgcct	1560
tctcctctgc cggaccgtgt gcacttcgct tcacctctgc acgtagcatg gagaccaaccg	1620
tgaacgcccc ccaggtcctg cccaaggtct tacacaagag gactccttga ctctcagcaa	1680
tgtcaacgac cgaccttgag gcatacttca aagactgttt gtttaaagac tgggaggagt	1740
tgggggagga gattaggtta aaggtccttg tactaggagg ctgtaggcat aaattggtct	1800
gttcaccagc accatgcaac tttttcccct ctgcctaato atctcatgtt catgtcctac	1860
tgttcaagcc tccaagctgt gccttgggtg gctttggggc atggacattg acccgataaa	1920
agaatttggg gcttctgtgg agttactctc ttttttgctt tctgacttct ttccttctat	1980
tgcagatctc ctgcacaacc cctctgctct gtatcgggag gccttagagt ctccggaaca	2040
ttgttcacct caccatacag cactcaggca agctattctg tgttggggtg agttgatgaa	2100
tctggccacc tgggtgggaa gtaatttggg agaccagca tccagggaaat tagtagtcag	2160
ctatgtcaat gttaatatgg gcctaaaaat tagacaacta ttgtggttcc acatttctctg	2220
ccttactttt ggaagagaaa ctgctcctga gtatttgggt tcttttggag tgtggattcg	2280
cactcctccc gcttacagac caccaaatgc cctatctta tcaaaccttc cggaaactac	2340
tgttggttaga cgacgaggca ggtcccctag aagaagaact ccctcgcctc gcagacgaag	2400
gtctcaatcg ccgcgtcgca gaagatctca atctcgggaa tctcaatgtt agtatecctt	2460
ggactcataa ggtgggaaac tttactgggc tttattcttc tactgtacct gtctttaatc	2520
ctgattggaa aactccctcc tttcctcaca ttcatttaca ggaggacatt attaatagat	2580
gtcaacaata tgtgggccct ctgacagtta atgaaaaaag gagattaaaa ttaattatgc	2640
ctgctagggt ctatcctaac cttaccaaat atttgcctt ggacaaaaggc attaaaccgt	2700
attatcctga atatgcagtt aatcattact tcaaaactag gcattattta cactactctgt	2760
ggaaggctgg cattctatat aagagagaaa ctacacgcag cgcctcattt tgtgggtcac	2820
catattcttg ggaacaagag ctacagcatg ggaggttggc cttccaaaacc tcgacaaggc	2880
atggggacga atctttctgt tcccatacct ctgggattct tcccogatca ccagttggac	2940
cctgcgttcg gagccaactc aaacaatcca gattgggact tcaaccccaa caaggatcac	3000
tggccagagg caaatcaggt aggagcggga gcatttggtc cagggttcac cccaccacac	3060
ggaggccttt tgggggtggg ccctcaggct cagggcatat tgacaacact gccagcagca	3120
cctcctcctg cctccaccaa tcggcagtca ggaagacagc ctactcccat ctctccacct	3180
ctaagagaca gtcatectca ggccatgcag tggaa	3215

---

-continued

---

<210> SEQ ID NO 20  
<211> LENGTH: 7905  
<212> TYPE: DNA  
<213> ORGANISM: Human Papilloma Virus HPV16 (NC\_001526.2)

<400> SEQUENCE: 20

actacaataa ttcattgtata aaactaaggg cgtaaccgaa atcgggttgaa ccgaaaccgg 60  
ttagtataaa agcagacatt ttatgcacca aaagagaact gcaatgttcc aggaccacca 120  
ggagcgaccc agaaagtac cacagttatg cacagagctg caaacaacta tacatgatat 180  
aatattagaa tgtgtgtact gcaagcaaca gttactgcga cgtgaggtat atgacttttc 240  
ttttcgggat ttatgcatag tatatagaga tgggaatcca tatgctgtat gtgataaatg 300  
tttaaagttt tattctaaaa ttatgagta tagacattat tgttatagtt tgtatggaac 360  
aacattagaa cagcaataca acaaacgggt gtgtgatttg ttaattaggt gtattaactg 420  
tcaaaagcca ctgtgtcctg aagaaaagca aagacatctg gacaaaaagc aaagattcca 480  
taataataag ggtcggtgga ccggtcgatg tatgtcttgt tgcagatcat caagaacacg 540  
tagagaaacc cagctgtaat catgcatgga gatacaccta cattgcatga atatatgtta 600  
gatttgcaac cagagacaac tgatctctac tgttatgagc aattaaatga cagctcagag 660  
gaggaggatg aaatagatgg tccagctgga caagcagaac cggacagagc ccattacaat 720  
attgtaacct tttgttgcaa gtgtgactct acgcttcggt tgtgcgtaca aagcacacac 780  
gtagacattc gtactttgga agacctgtta atgggcacac taggaattgt gtgcccacac 840  
tgttctcaga aaccataatc taccatggct gatcctgcag gtaccaatgg ggaagagggt 900  
acgggatgta atggatggtt ttatgtagag gctgtagtgg aaaaaaaaaac aggggatgct 960  
atatcagatg acgagaacga aaatgacagt gatacaggtg aagatttgggt agattttata 1020  
gtaaatgata atgattatct aacacaggca gaaacagaga cagcacatgc gttgtttact 1080  
gcacaggaag caaaacaaca tagagatgca gtacagggtc taaaaagaaa gtattttgta 1140  
gtccacttag tgatattagt ggatgtgtag acaataatat tagtctcaga ttaaaagcta 1200  
tatgtataga aaaaacaaag agagctgcaa aaaggagatt atttgaaagc gaagacagcg 1260  
ggatggcaa tactgaagtg gaaactcagc agatgttaca ggtagaaggc cccatgaga 1320  
ctgaaacacc atgtagtacg tatagtgggt gaagtggggg tggttgcagt cagtacagta 1380  
gtggaagtgg gggagagggt gttagtgaac gacacactat atgccaaca ccaattacaa 1440  
atattttaaa tgtactaaaa actagtaatg caaaggcagc aatgttagca aaatttaag 1500  
agttatacgg ggtgagtttt tcagaattag taagaccatt taaaagtaat aaatcaactg 1560  
gttgcgattg gtgtattgct gcatttggac ttacaccag tatagctgac agtataaaaa 1620  
cactattaca acaatattgt ttatatttac acattcaaag tttagcatgt tcatggggaa 1680  
tgggtgtgtt actattagta agatataaat gtggaaaaaa tagagaaaca attgaaaaat 1740  
tgctgtctaa actattatgt gtgtctccaa tgtgtatgat gatagagcct ccaaaattgc 1800  
gtagtacagc agcagcatta tattgtgata aaacaggtat atcaaatatt agtgaagtgt 1860  
atggagacac gccagaatgg atacaaagac aaacagtatt acaacatagt tttaatgatt 1920  
gtacatttga attatcacag atggtacaat gggcctacga taatgacata gtagacgata 1980  
gtgaaattgc atataaatat gcacaattgg cagacactaa tagtaatgca agtgccttcc 2040

---

-continued

---

taaaaagtaa ttcacaggca aaaattgtaa aggattgtgc aacaatgtgt agacattata 2100  
aacgagcaga aaaaaaacia atgagtatga gtcaatggat aaaatataga tgtgataggg 2160  
tagatgatgg aggtgattgg aagcaaatg ttatgttttt aaggtatcaa ggtgtagagt 2220  
ttatgtcatt ttttaactgca ttaaaaagat ttttgcaagg catacctaaa aaaaattgca 2280  
tattactata tgggtgcagct aacacaggta aatcattatt tgggtatgagt ttaatgaaat 2340  
ttctgcaagg gtctgtaata tggtttgtaa attctaaaag ccatttttgg ttacaacct 2400  
tagcagatgc caaaataggt atgttagatg atgctacagt gccctgttgg aactacatag 2460  
atgacaattt aagaaatgca ttggatggaa atttagtttc tatggatgta aagcatagac 2520  
cattggtaca actaaaatgc cctccattat taattacatc taacattaat gctggtacag 2580  
attctaggtg gccattattha cataatagat tgggtggtgt tacatttctc aatgagtttc 2640  
catttgacga aaacggaaat ccagtgatg agcttaatga taagaactgg aaatcctttt 2700  
tctcaaggac gtggtccaga ttaagtttgc acgaggacga ggacaaggaa aacgatggag 2760  
actcttggc aacgtttaa tgtgtgtcag gacaaaatc taacacatta tgaatatgat 2820  
agtacagacc tacgtgacca tatagactat tggaaacaca tgcgcctaga atgtgctatt 2880  
tattacaagg ccagagaaat gggattttaa catattaacc accaagtgg gccaacactg 2940  
gctgtatcaa agaataaagc attacaagca attgaaactgc aactaacgtt agaaacaata 3000  
tataactcac aatataagta tgaaaaagtg acattacaag acgttagcct tgaagtgtat 3060  
ttaactgcac caacaggatg tataaaaaaa catggatata cagtggaaat gcagtttgat 3120  
ggagacatat gcaatacaat gcattatata aactggacac atatatatat ttgtgaagaa 3180  
gcatcagtaa ctgtggtaga gggccaagtt gactattatg gtttatatta tgttcatgaa 3240  
ggaatacga catattttgt gcagttttaa gatgatgcag aaaaatatag taaaaataaa 3300  
gtatgggaag ttcattcggg tggtcaggta atattatgct ctacatctgt gtttagcagc 3360  
aacgaagtat cctctcctga aattattagg cagcacttgg ccaaccacc cgcgcgacc 3420  
cataccaaag ccgtcgcctt gggcaccgaa gaaacacaga cgactatcca gcgaccaaga 3480  
tcagagccag acaccggaaa cccctgccac accactaagt tggtagcag agactcagtg 3540  
gacagtgtc caatcctcac tgcatttaac agctcacaca aaggacggat taactgtaat 3600  
agtaacacta caccatagc acattttaa ggtgatgcta atactttaa atgtttaaga 3660  
tatagattta aaaagcattg tacattgtat actgcagtgt cgtctacatg gcattggaca 3720  
ggacataatg taaaacataa aagtgaatt gttacactta catatgatag tgaatggcaa 3780  
cgtgaccaat ttttgtctca agttaaata ccaaaaacta ttacagtgtc tactggattt 3840  
atgtctatat gacaaatctt gatactgcat ccacaacatt actggcgtgc tttttgcttt 3900  
gcttttgtgt gcttttgtgt gtctgcctat taatacgtcc gctgcttttg tctgtgtcta 3960  
catacacatc attaataata ttggatttac tattgtggat aacagcagcc tctgcgttta 4020  
gggtttttat tgtatatatt atatttgttt atataccatt atttttaata catacacatg 4080  
cacgtttttt aattacataa tgtatatgta cataatgtaa ttgttacata taattgttgt 4140  
ataccataac ttactatttt ttctttttta ttttcatata taattttttt ttttgtttgt 4200  
ttgtttgttt tttataaac tgttattact taacaatgcg acacaaactg tctgcaaac 4260  
gcacaaaacg tgcacggctt acccaacttt ataaaactg caaacaggca ggtacatgct 4320

-continued

---

cacctgacat tatacctaag gttgaaggca aaactattgc tgaacaaata ttacaatatg	4380
gaagtatggg tgtatTTTTT ggtgggtag gaattggaac agggtcgggt acaggcggac	4440
gcactgggta tattccattg ggaacaaggc ctcccacagc tacagataca cttgctcctg	4500
taagaccccc ttaacagta gatcctgtgg gcccttctga tccttctata gtttctttag	4560
tggaagaaac tagttttatt gatgctggg caccaacatc tgtaccttc attccccag	4620
atgtatcagg atttagtatt actacttcaa ctgataccac acctgctata ttagatatta	4680
ataaactgt tactactgtt actacacata ataatccac tttcactgac ccatctgtat	4740
tgcagcctcc aacacctgca gaaactggag ggcattttac actttcatca tccactatta	4800
gtacacataa ttatgaagaa attcctatgg atacatttat tgttagcaca aaccetaaca	4860
cagtaactag tagcacacc ataccagggt ctgcgccagt ggcagccta ggattatata	4920
gtcgcacaa acaacagggt aaagtgttag acctgcttt tgaaccact cccactaaac	4980
ttattacata tgataatcct gcatatgaag gtatagatgt ggataatata ttatTTTT	5040
ctagtaatga taatagtatt aatatagctc cagatcctga ctttttggat atagttgctt	5100
tacataggcc agcattaacc tctaggccta ctggcattag gtacagtaga attggaata	5160
aacaaacact acgtactcgt agtgaaaaat ctatagggtc taaggtaacat tattattatg	5220
atttaagtac tattgatcct gcagaagaaa tagaattaca aactataaca ccttctacat	5280
atactaccac ttcacatgca gcctcaccta cttctattaa taatggatta tatgatattt	5340
atgcagatga ctttattaca gatacttcta caacccgggt accatctgta cctctacat	5400
ctttatcagg ttatattcct gcaatacaaa caattccttt tgggtgtgca tacaatattc	5460
ctttagtatc aggtcctgat ataccatta atataactga ccaagctcct tcattaatc	5520
ctatagttcc agggctcca caatatacaa ttattgctga tgcaggtgac ttttatttac	5580
atcctagtta ttacatgcta cgaaaacgac gtaaacgttt accatatttt ttttcagatg	5640
tctctttggc tgcttagtga ggccactgtc tacttgctc ctgtoccagt atctaagggt	5700
gtaagcagg atgaatatgt tgcacgcaca aacatatatt atcatgcagg aacatccaga	5760
ctacttcag ttggacatcc ctattttcct attaaaaaac ctaacaataa caaaatatta	5820
gttcctaaag tatcaggatt acaatacagg gtatttagaa tacatttacc tgacccaat	5880
aagtttgggt ttctgacac ctcattttat aatccagata cacagcggct ggtttgggcc	5940
tgtgtagggt ttgagtagg tcgtggtcag ccattagggt tgggcattag tggccatcct	6000
ttattaaata aattggatga cacagaaaat gctagtgtt atgcagcaaa tgcaggtgtg	6060
gataatagag aatgtatatt tatggattac aaacaaacac aattgtgttt aattggttgc	6120
aaaccaccta taggggaaca ctggggcaaa ggcacccat gtaccaatgt tgcagtaaat	6180
ccagggtatt gtccaccatt agagttaata aacacagtta ttcaggatgg tgataggtt	6240
catactggct ttggtgctat ggactttact acattacagg ctaacaaaag tgaagtcca	6300
ctggatattt gtacatctat ttgcaaatat ccagattata ttaaaatggt gtcagaacca	6360
tatggcgaca gcttattttt ttatttacga agggaacaaa tgtttgttag acatttattt	6420
aatagggtgt gtactgttgg tgaaaaatgta ccagacgatt tatacattaa aggctctggg	6480
tctactgcaa atttagccag ttcaaatatt tttcctacac ctagtgggtc tatggttacc	6540
tctgatgccc aaatattcaa taaaccttat tggttacaac gagcacaggg ccacaataat	6600

-continued

---

```

ggcatttggt ggggtaacca actatttggt actgttggtg atactacacg cagtacaaat 6660
atgtcattat gtgctgccat atctacttca gaaactacat ataaaaatac taactttaag 6720
gagtacctac gacatgggga ggaatatgat ttacagttta tttttcaact gtgcaaaata 6780
accttaactg cagaagttat gacatacata cattctatga attccactat tttggaggac 6840
tggaattttg gtctacaacc tccccagga ggcacactag aagatactta taggtttgta 6900
accaggcaa ttgcttgta aaaacatata cctccagcac ctaagaaga tgatccccctt 6960
aaaaaataca ctttttggga agtaaattha aaggaaaagt tttctgcaga cctagatcag 7020
tttcttttag gacgcaaat tttactaca gcaggattga aggccaaacc aaaatttaca 7080
ttaggaaaac gaaaagctac acccaccacc tcatctacct ctacaactgc taaacgcaaa 7140
aacgtaagc tgtaagtatt gtatgtatgt tgaattagtg ttgtttgttg tgtatatgtt 7200
tgtatgtgct tgtatgtgct tgtaaatatt aagttgatg tgtgtttgta tgtatggat 7260
aataaacacg tgtgtatgtg tttttaaag cttgtgtaac tattgtgtca tgcaacataa 7320
ataaacttat tgtttcaaca cctactaatt gtgttggtg tattcattgt atataaacta 7380
tatttgctac atcctgtttt tgttttatat atactatatt ttgtagcgcc aggccattt 7440
tgtagcttca accgaattg gttgcatgct ttttggcaca aaatgtgttt ttttaaatag 7500
ttctatgtca gcaactatgg tttaaacttg tacgtttcct gcttgccatg cgtgccaaat 7560
ccctgttttc ctgacctgca ctgcttgcca accattccat tgttttttac actgactat 7620
gtgcaactac tgaatcaact tgtaacattg gtcataataa ataaatcact atgcgccaac 7680
gcctacata ccgctgtag gcacatattt ttggcttgtt ttaactaacc taattgcata 7740
tttggcataa ggtttaaact tctaaggcca actaaatgct accctagtgc atacatgaac 7800
tgtgtaaagg ttatgcatac attgttcatt tgtaaaactg cacatgggtg tgtgcaaac 7860
gattttgggt tacacattta caagcaactt atataataat actaa 7905

```

&lt;210&gt; SEQ ID NO 21

&lt;211&gt; LENGTH: 7857

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Human Papilloma Virus HPV18 (X05015.1)

&lt;400&gt; SEQUENCE: 21

```

attaactctt ttaacaattg tagtatataa aaaaggagat aaccgaaaac ggtcgggacc 60
gaaaaagggt tatataaaag atgtgagaaa cacaccacaa tactatggcg cgctttgagg 120
atccaacacg gcgacctac aagetacctg atctgtgcac ggaactgaac acttcactgc 180
aagacataga aataacctgt gtatattgca agacagtatt ggaacttaca gaggtatttg 240
aatttgcat taaagattta tttgtggtg atagagacag tataccccat gctgcatgcc 300
ataaatgtat agatttttat tctagaatta gagaattaag acattattca gactctgtgt 360
atggagacac attgaaaaa ctaactaaca ctgggttata caatttatta ataaggtgcc 420
tgcggtgcca gaaaccgttg aatccagcag aaaaacttag acaccttaat gaaaaacgac 480
gatttcacaa catagctggg cactatagag gccagtcca ttcgtgctgc aaccgagcac 540
gacaggaacg actccaacga cgcagagaaa cacaagtata atattaagta tgcattggac 600
taaggcaaca ttgcaagaca ttgtattgca tttagagccc caaaatgaaa ttccggttga 660
cctctatgt cagcagcaat taagcgactc agaggaagaa aacgatgaaa tagatggagt 720

```

-continued

---

taatcatcaa catttaccag cccgacgagc cgaaccacaa cgtcacacaa tgttgtgtat	780
gtgttgtaag tgtgaagcca gaattgagct agtagtagaa agctcagcag acgaccttcg	840
agcattccag cagctgtttc tgaacacct gtcctttgtg tgcctgtggt gtgcatccca	900
gcagtaagca acaatggctg atccagaagg tacagacggg gagggcacgg gttgtaacgg	960
ctggttttat gtacaagcta ttgtagacaa aaaaacagga gatgtaatat cagatgacga	1020
ggacgaaaat gcaacagaca cagggtcggg tatggttagat tttattgata cacaaggaac	1080
atthttgtgaa caggcagagc tagagacagc acaggcattg ttccatgcgc aggaggtcca	1140
caatgatgca caagtgttgc atgtttttaa acgaaagtth gcaggaggca gcacagaaaa	1200
cagtccatta ggggagcggc tggaggtgga tacagagtha agtccacggt tacaagaaat	1260
atctttaaat agtgggcaga aaaaggcaaa aaggcggctg tttacaatat cagatagtgg	1320
ctatggctgt tctgaagtgg aagcaacaca gattcagtha actacaaatg gcgaacatgg	1380
cggcaatgta tgtagtggcg gcagtacgga ggctatagac aacgggggca cagagggcaa	1440
caacagcagt gtacagcggta caagtgacaa tagcaatata gaaaaatgaa atccacaatg	1500
taccatagca caatataaag acttgthtaa agtaaacaaat aaacaaggag ctatgttagc	1560
agtatthtaa gacacatatg ggctatcatt tacagattha gthtagaaat ttaaaagtga	1620
taaaaccacg tgtacagatt gggttacagc tatatttgga gtaaaccaa caatagcaga	1680
aggatthtaa acactaatac agccatttat attatatgcc catattcaat gtctagactg	1740
taaatgggga gtattaatat tagcctgtt gcgttacaaa tgtggtaaga gtagactaac	1800
agttgctaaa ggtthaagta cgttgthaca cgtacctgaa acttgatgt taattcaacc	1860
acaaaattg cgaagtgtg ttgcagcact atattggtat agaacaggaa tatcaaatat	1920
tagtgaagta atgggagaca cacctgagtg gatacaaaaga cttactatta tacaacatgg	1980
aatagatgat agcaatthtg atthgtcaga aatggtacaa tgggcatttg ataagtgtg	2040
gacagatgaa agcgatatgg catttgaata tgccttatta gcagacagca acagcaatgc	2100
agctgccttt ttaaaaagca atthccaagc taaatthta aaagattgtg ccacaatgtg	2160
caaacattat aggcgagccc aaaaacgaca aatgaatag tccagtgga tacgatttag	2220
atgttcaaaa atagatgaag ggggagattg gagaccaata gtgcaattcc tgcgatacca	2280
acaaatagag thtataacat thttaggagc cthaaatca thtttaaaag gaaccccaaa	2340
aaaaaattgt thagtatttt gtggaccagc aaatacagga aatcatatt thggaatgag	2400
thttatacac thtatacaag gagcagthaat atcattthtg aattccacta gtcattthtg	2460
gttgaaccg thaacagata ctaaggtggc catgttagat gatgcaacga ccacgtgttg	2520
gacatacttt gatacctata tgagaaatgc gthtagatggc aatccaataa gtattgatag	2580
aaagcacaaa ccattaatac aactaaaatg tctccaata ctactaacca caaatataca	2640
tccagcaaa gataatagat ggccatattt agaaagtaga ataacagtat thgaatttcc	2700
aaatgcattt ccatttgata aaaaaggcaa tccagtatat gaaataaatg acaaaaattg	2760
gaaatgtttt thtgaagga catggtccag attagatttg cacaggaag aggaagatgc	2820
agacaccgaa ggaaccctt tcggaacgth taagthgctg gcaggacaaa atcatagacc	2880
actatgaaaa tgacagtaaa gacatagaca gccaataca gtattggcaa ctaatacgtt	2940
gggaaaatgc aatattcttt gcagcaaggg aacatggcat acagacatta aaccaccag	3000

-continued

---

tgggtgccagc ctataacatt tcaaaaagta aagcacataa agctattgaa ctgcaaatgg	3060
ccctacaagg ccttgcacaa agtcgatata aaaccgagga ttggacactg caagacacat	3120
gcgaggaact atggaataca gaacctactc actgctttaa aaaagggtggc caaacagtac	3180
aagtatatth tgatggcaac aaagacaatt gtatgaccta thtagcatgg gacagtgtgt	3240
attatatgac tgatgcagga acatgggaca aaaccgctac ctgtgtaagt cacaggggat	3300
tgtattatgt aaaggaaggg tacaacacgt tttatataga atttaaaagt gaatgtgaaa	3360
aatatgggaa cacaggtacg tgggaagtac attttgggaa taatgtaatt gattgtaatg	3420
actctatgtg cagtaccagt gacgacacgg tatccgctac tcagcttggt aaacagctac	3480
agcacacccc ctcaccgtat tccagcaccg tgtccgtggg caccgcaaag acctacggcc	3540
agacgtcggc tgctacacga cctggacact gtggactcgc ggagaagcag catttgtggac	3600
ctgtcaaccc acttctcggg gcagctacac ctacaggcaa caacaaaaga cggaaactct	3660
gtagtggtaa cactacgcct ataatacatt taaaagggtga cagaacagct ttaaaatgtt	3720
tacggtagcag attgcgaaaa catagcgacc actatagaga tataatcatcc acctggcatt	3780
ggacaggtgc aggcaatgaa aaaacaggaa tactgactgt aacataccat agtgaaacac	3840
aaagaacaaa atttttaaat actgttgcaa ttccagatag tgtacaaaata ttgggtgggat	3900
acatgacaat gtaatacata tgctgtagta ccaatatgtt atcacttatt tttttatth	3960
gcttttgtgt atgcatgtat gtgtgctgcc atgtcccgt tttgccatct gtctgtatgt	4020
gtgcgtatgc atgggtattg gtatttgtgt atattgtggt aataacgtcc cctgccacag	4080
cattcacagt atatgtatth tgtttttat tgccatggt actattgcat atacatgcta	4140
tattgtctth acagtaattg tataggttgt tttatacagt gtattgtaca ttgtatattt	4200
tgttttatac cttttatgct ttttgtatth ttgtaataaa agtatggtat cccaccgtgc	4260
cgcacgacgc aaacgggctt cggtaactga cttatataaa acatgtaaac aatctggtag	4320
atgtccacct gatgttgttc ctaaggtgga gggcaccacg ttagcagata aaatattgca	4380
atggccaagc cttggtatath ttttgggtgg acttggcata ggtactggca gtggtacagg	4440
gggtcgtaca ggttacatth cattgggtgg gcgttccaat acagtggtgg atgttggtec	4500
tacacgtccc ccagtggtta ttgaaactgt gggccccaca gaccatcta ttgttacatt	4560
aatagaggac tccagtgtgg ttacatcagg tgcacctagg cctacgttta ctggcacgtc	4620
tgggtttgat ataacatctg cgggtacaac tacacctgcg gttttggata tcacacctc	4680
gtctacctct gtgtctatth ccacaaccaa ttttaccat cctgcattth ctgatecgtc	4740
cattattgaa gttccacaaa ctggggaggt ggcaggtaat gtatttgttg gtaccctac	4800
atctggaaca catgggtatg aggaaatacc tttacaaaca tttgcttct ctggtacggg	4860
ggaggaaccc attagtagta cccattgcc tactgtgcgg cgtgtagcag gtecccgct	4920
ttacagtagg gcctaccaac aagtgtcagt ggctaaccct gagtttctta cacgtccatc	4980
ctctttaat acatagaca acccgccctt tgagcctgtg gacactacat taacatttga	5040
tcctcgtagt gatgttctc attcagatth tatggatath atccgtctac ataggcctgc	5100
tttaacatcc aggcgtggga ctgttcgctt tagtagatta ggtcaacggg caactatgtt	5160
taccgcagc ggtacacaaa taggtgctag ggttcactth tatcatgata taagtcctat	5220
tgcaccttc ccagaatata ttgaaactgca gcctttagta tctgccacgg aggacaatga	5280

-continued

---

cttgtttgat	atatatgcag	atgacatgga	ccctgcagtg	cctgtacat	cgcgttctac	5340
tacctccttt	gcatttttta	aatattcgcc	cactatatct	tctgcctctt	cctatagtaa	5400
tgtaacggtc	cctttaacct	cctcttggga	tgtgcctgta	tacacgggtc	ctgatattac	5460
attaccatct	actacotctg	tatggcccat	tgtatcaccc	acggccctg	cctctacaca	5520
gtatattggt	atacatggta	cacattatta	tttgtggcca	ttatattatt	ttattcctaa	5580
gaaacgtaaa	cgtgttccct	atTTTTttgc	agatggcttt	gtggcggcct	agtgacaata	5640
ccgtatatct	tcacctctct	tctgtggcaa	gagttgtaaa	taccgatgat	tatgtgactc	5700
ccacaagcat	atTTTTatcat	gctggcagct	ctagattatt	aactgttggg	aatccatatt	5760
ttagggttcc	tgcaggtggt	ggcaataagc	aggatattcc	taaggtttct	gcataccaat	5820
atagagtatt	tagggtgca	ttacctgacc	caaataaatt	tggtttacct	gatactagta	5880
tttataatcc	tgaaacacaa	cgtttagtgt	ggcctgtgct	tggagtggaa	attggccgtg	5940
gtcagccttt	aggtgttggc	cttagtgggc	atccatttta	taataaatta	gatgacactg	6000
aaagttccca	tgcgccacag	tctaattgtt	ctgaggacgt	tagggacaat	gtgtctgtag	6060
attataagca	gacacagtta	tgtatTTTgg	gctgtgcccc	tgetattggg	gaacactggg	6120
ctaaaggcac	tgcttgtaaa	tcgcgtcctt	tatcacaggg	cgattgcccc	cctttagaac	6180
ttaaaaacac	agTTTTggaa	gatggtgata	tggtagatac	tggataggt	gccatggact	6240
ttagtacatt	gcaagatact	aaatgtgagg	taccattgga	tatttgcag	tctatttcta	6300
aatatcctga	ttatttacia	atgtctgcag	atccttatgg	ggattccatg	tttttttct	6360
tacggcgtga	gcagctTTTT	gctaggcatt	tttggaaatg	agcaggtact	atgggtgaca	6420
ctgtgcctca	atccttatat	attaaggcca	caggatgccc	tgcttcacct	ggcagctgtg	6480
tgtattctcc	ctctccaagt	ggctctattg	ttacctctga	ctcccagttg	tttaataaac	6540
catattgggt	acataaggca	cagggtcata	acaatgggtg	ttgctggcat	aatcaattat	6600
ttgttactgt	ggtagatacc	actcccagta	ccaatttaac	aatatgtgct	tctacacagt	6660
ctcctgtacc	tgggcaatat	gatgctacca	aatttaagca	gtatagcaga	catggtgagg	6720
aatatgattt	gcagtttatt	tttcagttgt	gtactattac	tttaactgca	gatggtatgt	6780
cctatattca	tagtatgaat	agcagtatTT	tagaggattg	gaactttggg	gttccccccc	6840
ccccaaactac	tagtttgggtg	gatacatatc	gttttgtaca	atctgttct	attacctgtc	6900
aaaaggatgc	tgcaccggct	gaaaataagg	atccctatga	taagttaaag	ttttggaatg	6960
tggatttaaa	ggaagaagttt	tcttttagact	tagatcaata	tccccttggg	cgtaaatTTT	7020
tggttcaggc	tggattgctg	cgcaagcccc	ccataggccc	tcgcaaacgt	tctgctccat	7080
ctgccactac	gtcttctaaa	cctgccaaagc	gtgtgcgtgt	acgtgccagg	aagtaatatg	7140
tgtgtgtgta	tatatatata	catctattgt	tgtgtttgta	tgtcctgtgt	ttgtgtttgt	7200
tgtatgattg	cattgtatgg	tatgtatggg	tgtgtttgta	tgtgtatgt	tactatattt	7260
gttggatgt	ggcattaaat	aaaatatggt	ttgtggttct	gtgtgttatg	tggttgcgcc	7320
ctagttagta	acaactgtat	ttgtgtttgt	ggatggggtg	ttgcttgttg	ggctatatat	7380
tgtcctgtat	ttcaagttat	aaaactgcac	accttacagc	atccatttta	tcctacaatc	7440
ctccattttg	ctgtgcaacc	gatttcgggt	gcctttggct	tatgtctgtg	gttttctgca	7500
caatacagta	cgtggcact	attgcaaac	ttaatctttt	gggcactgct	cctacatatt	7560

-continued

---

ttgaacaatt ggcgcgcctc tttggcgcac ataaggcgca cctgggatta gtcattttcc	7620
tgtccagggt cgctacaaca attgcttgca taactatata cactccctaa gtaataaaac	7680
tgcttttagg cacatatatt agtttggttt tacttaagct aattgcatac ttggcttgta	7740
caactacttt catgtccaac attctgtota cccttaacat gaactataat atgactaagc	7800
tgtgcataca tagtttatgc aaccgaaata ggttgggcag cacatactat acttttc	7857

&lt;210&gt; SEQ ID NO 22

&lt;211&gt; LENGTH: 7912

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Human Papilloma Virus HPV31 (J04353.1)

&lt;400&gt; SEQUENCE: 22

taataataat aatcttagta taaaaaagta gggagtgacc gaaagtgggtg aaccgaaaac	60
ggttgggtata taaagcacat agtatattgt gcaaacctac agacgccatg ttcaaaaatc	120
ctgcagaaag acctcggaaa ttgcatgaac taagctcggc attggaaata ccctacgatg	180
aactaagatt gaattgtgtc tactgcaaaag gtcagttaac agaaacagag gtattagatt	240
ttgcatctac agatthaaca atagtatata gggacgacac accacacgga gtgtgtacaa	300
aatgtttaag attttattca aaagtaagtg aatttagatg gtatagatat agtgtgtatg	360
gaacaacatt agaaaaattg acaaaaaag gtatatgtga tttgttaatt aggtgtataa	420
cgtgtcaaag accgttgtgt ccagaagaaa aacaaagaca tttggataaa aagaaacgat	480
tccacaacat aggaggaagg tggacaggac gttgcatagc atgttgagaga agacctcgtg	540
ctgaaaccca agtgtaaaca tgcgtggaga aacacctacg ttgcaagact atgtgttaga	600
tttgcaacct gaggcaactg acctccactg ttatgagcaa ttacctgaca gctcagatga	660
ggaggatgtc atagacagtc cagctggaca agcagaaccg gacacatcca attacaatat	720
cgttaccttt tgttgcagc gtaagtctac acttcgtttg tgtgtacaga gcacacaagt	780
agatattcgc atattgcaag agctgttaat gggctcattt ggaatcgtgt gcccactg	840
ttctactaga ctgtaactac aatggctgat ccagcaggta cagatgggga ggggacggga	900
tgcaatgggt ggttttatgt agaagcagta attgacagac agacagggga caacatttca	960
gaggacgaaa atgaagacag tagtgatact ggggaggata tgggtgactt tattgacaat	1020
tgtaatgtat acaacaatca ggcagaagca gagacagcac aggcattggt tcatgcacag	1080
gaagcggagg aacatgcaga ggctgtgcag gttctaaaac gaaagtatgt aggtagtctt	1140
ttaagtgata ttagtagttg tgtggattat aatattagtc cacggttaaa agctatatgc	1200
atagaaaata acagtaaaac agcaaaacga agactctttg aacttccaga cagcgggtat	1260
ggcaatactg aagtggaaaac gcagcagatg gtacaggtag aggagcaaca aacaacatta	1320
agttgtaatg gtatgtacgg gacacatagt gaacgagaga atgaaactcc aacacgtaat	1380
atattgcaag tgtaaaaaac tagcaatggt aaagctgcta tgttaggtaa atttaagaa	1440
ttatatgggt taagttttat ggaactaatt aggccatttc aaagcaataa aagcacatgt	1500
actgattggt gtgtagctgc gtttggagtt acaggtagcag ttgcagaagg atttaaaacc	1560
ctattgcaac catattgttt gtattgcat ttacaaaagt tagcatgttc ctggggcatg	1620
gttatgttaa tgcttgtgag atttaaatgt gcaaaaaata gaataacaat tgaaaaatta	1680
ttagaaaaat tattgtgtat atctacaaat tgtatgttaa ttcagccacc caaattacgt	1740

-continued

---

agcacagctg cagcattata ttggtacaga acaggaatgt caaacattag cgatgtatat	1800
ggtgaaacac cagaatggat agaaagacaa acagtattac agcatagttt taatgacaca	1860
acatttgatt tgtcccaaat ggtacaatgg gcatatgaca atgatgttat ggatgatagt	1920
gaaattgect ataatatgac acaattagct gacagtgata gtaatgcatg tgcattttta	1980
aaaagtaatt cgcaggcaaa aatagttaaa gattgtggaa caatgtgtag acattataaa	2040
cgagcagaaa aacgacaaat gtccatggga cagtggatta aaagtagatg tgacaaagtt	2100
agtgcagaag gtgactggag ggacatagta aagtttttaa gatatcaaca aatagaattt	2160
gtgcatttt tatctgcatt aaagctgttt ttaaaaggag tgccaaagaa aaactgtatt	2220
ttaatacatg gtgcacctaa tacaggtaaa tcatattttg gaatgagcct tattagcttt	2280
ttacaaggat gtataatato atatgcaaat tcaaaaagtc atttttggtt acaaccactg	2340
gctgatgcta aaataggeat gttagatgat gctacaacgc catggtggca ttatatagac	2400
aattacctac gaaatgcact agatggcaac cctgtatcta tagatgtaaa gcataaagct	2460
ttaatgcagt taaatgtcc tcctttattg attacatcta atataaatgc aggtaaggat	2520
gacagatggc catacctaca tagcagactg gtggttttta catttccaaa tccatttcca	2580
tttgacaaaa acggaaatcc agtatatgaa ttaagtata aaaactggaa atcctttttc	2640
tcaaggacgt ggtgcagatt aaatttgcac gaggaagagg acaagaaaa cgatggagac	2700
tcctttctca cgtttaaatg tgtgtcagga caaaatatta gaacattatg aaaatgatag	2760
taaocgactt tgtgatcata tagactattg gaaacatatt cgacttgaat gtgtattaat	2820
gtataaagca agagaatgg gaatacacag tattaaccac caggtggtgc cagcgttgtc	2880
agtatcaaag gccaaagcct tacaagctat tgaactacaa atgatgttg aaacattaaa	2940
taactactga tacaanaatg aggactggac aatgcagcaa acaagtcttg aactgtattt	3000
aactgcacct acaggggtgt taaaaaaca tggatatact gtagagggtc aatttgatgg	3060
tgatgtacac aacaccatgc attatactaa ctggaaattt atatacctat gtatagatgg	3120
ccaatgtact gttgtggaag ggcaagttaa ttgtaagggc atttattatg tacatgaagg	3180
acataataca tattttgtaa attttacaga agaggcaaaa aaatatggga ctggtaaaaa	3240
atgggaagtg catgcgggtg gtcaggtaat tgtttttcct gaatctgtat ttagcagtga	3300
cgaaaatcc tttgctggga ttgttacaaa gctaccaaca gccacaaca ccaccacatc	3360
gaattccaaa acctgcgcct tgggcaccag tgaaggtgtg cggcgggcca cgactctac	3420
taagcgacca agaacagagc cagagcacag aaacaccac cacccaaca agttgttgcg	3480
aggcgactcc gtggacagtg tcaactgtgg ggttatcagt gcagctgcat gcacaaacca	3540
aacaagggtc gtcagttgtc ctgcaactac acctataata cacttaaaag gtgatgcaaa	3600
tatatataaa tgtttaagat ataggctgtc aaaaataaaa caattgtatg aacaagtgtc	3660
atctacatgg cattggacat gtacagatgg aaacataaaa aatgctattg taacctaac	3720
atatataagt acatcacaaa gagacgattt tttaaatact gtaaaaatac ctaacacagt	3780
atcagtgta acaggatata tgaactatta gcctaagat tgaactaaat atttctacag	3840
taagcattgt gctatgcttt ttgctttgct tttgtgtgct actatgtgtg tgcctgtca	3900
tacgtccact tgtgctgtct gtgtcgggat atgcaacact actattatta attgtgattt	3960
tatgggttat tgcaacctct ccattacggt gtttttgtat atatgtgtg tttatatata	4020

-continued

---

ttccattatt	tgtaattcat	acacatgcat	cttttttaag	tcaacagtaa	cttttttact	4080
tgtgtatact	gttgtttga	ttggtattgg	tattggtatt	ggtattggta	ttggtataat	4140
aaactttttt	actttttttt	tattattacc	atgcggtcca	aacgctctac	aaaacgcact	4200
aaacgtgcgt	ctgctacaca	attatatcaa	acatgtaaag	cagcaggtae	ttgtccatca	4260
gacgttatac	ctaaaataga	acatactacc	attgcagacc	aaatattaag	gtatggtagt	4320
atgggtgttt	tttttgggtg	gttgggtatt	gggtccggct	ctggtactgg	gggtcgcact	4380
ggatagtcc	ctcttagtac	acgtcctct	acagtatctg	aggcaagtat	acctattaga	4440
ccaccagtta	gcattgacct	tgtaggtccc	ttggacctcc	ctatagtaag	tcttgttgaa	4500
gaactctgga	ttgttgatgt	tggtgcccct	gctcctatac	cacaccctcc	tacaacatct	4560
gggtttgaca	ttgctacaac	tgcagacaca	acacctgcaa	ttttagatgt	aacaagtgtt	4620
agcacacatg	aaaatcctac	ttttactgat	ccatctgfat	tgcagcctcc	tacaacctgca	4680
gaaacatcag	gtcatttact	actttcatca	tcacttatta	gcacacataa	ttatgaggaa	4740
atacctatgg	atacatttat	tgtttctact	aataatgaaa	acataacaag	tagcacaccc	4800
attccagggg	tgcgccgtcc	tgcacgttta	gggttatata	gtaaggctac	acaacaagta	4860
aaagtatttg	atccaacgtt	tcttagtgct	ccaaaacagc	taattacata	tgaaaaccct	4920
gcctatgaaa	ctgtaaatgc	tgaagaatct	ttatactttt	ccaatacctc	gcataatata	4980
gccctgatc	ccgactttct	agatattata	gcattacata	ggcctgccc	tacctcacgt	5040
aggaacactg	ttagatatag	tagactaggt	aataaacaac	ctttgcgcac	tcgtagtgg	5100
gctactattg	gtgcaagggt	gcattattat	tatgatatta	gtagtattaa	tcctgcaggt	5160
gaaagtattg	aaatgcaacc	tttagggggc	tctgcaacta	ctacttctac	tttaaatgat	5220
ggcttatatg	acatttatgc	agacactgat	tttactgtgg	atacactctc	cacacataat	5280
gtttcccctt	ctactgctgt	acagtccaca	tctgctgtgt	ctgctatgt	acctacaaat	5340
accactgtgc	cactaagtac	aggttttgac	attcccata	tttctgggcc	tgatgtacct	5400
atagagcatg	cacctacaca	ggttttccca	tttcctttgg	cccctacaac	gccacaagtg	5460
tctatttttg	ttgatggggg	tgatttttat	ttgcacccta	gttattatat	gttaaaacgt	5520
cgacgtaaac	gtgtatcata	tttttttaca	gatgtctctg	tggcggccta	gagaggctac	5580
tgtctactta	ccacctgtcc	cagtgtctaa	agttgtaagc	acggatgaat	atgtaacacg	5640
aaccaacata	tattatcacg	caggcagtcg	taggtgctt	acagtaggcc	atccatatta	5700
ttccatacct	aaatctgaca	atcctaaaaa	aatagttgta	ccaaagggtg	caggattaca	5760
atatagggta	tttaggggtc	gtttaccaga	tccaaacaaa	tttggatttc	ctgatacctc	5820
tttttataat	cctgaaactc	aacgcttagt	ttgggctgtg	gttggtttag	aggtaggtcg	5880
cgggcagcca	ttagggtgag	gtattagtgg	tcctccatta	ttaataaat	ttgatgacac	5940
tgaaaactct	aatagatatg	ccggtgggtc	tggcactgat	aataggaat	gtatatcaat	6000
ggattataaa	caaacacaac	tgtgtttact	tggttgcaaa	ccacctattg	gagagcattg	6060
gggtaaagg	agtccttcta	gtaacaatgc	tattaccctc	ggtgattgtc	ctccattaga	6120
attaaaaaat	tcagttatac	aagatgggga	tatggttgat	acaggctttg	gagctatgga	6180
ttttactgct	ttacaagaca	ctaaaagtaa	tgttcctttg	gacatttgta	attctatttg	6240
taaatatoca	gattatctta	aaatgggtgc	tgagccatat	ggcgatacat	tattttttta	6300

-continued

---

```

tttacgtagg gaacaaatgt ttgtaaggca tttttttaat agatcaggca cggttggtga 6360
atcggtcocct actgacttat atattaaagg ctccggttca acagctactt tagctaacag 6420
tacatacttt cctacaccta gcggtcccat ggttacttca gatgcacaaa tttttaataa 6480
accatattgg atgcaacgtg ctccaggaca caataatggt atttggtggg gcaatcagtt 6540
atgtgttact gtggtagata ccacacgtag taccaatag tctgtttgtg ctgcaattgc 6600
aaacagtgat actacattta aaagtagtaa ttttaagag tatttaagac atggtaggga 6660
atgtgattta caatttatat ttcagttatg caaaataaca ttatctgcag acataatgac 6720
atatattcac agtatgaatc ctgctatttt ggaagattgg aattttggat tgaccacacc 6780
tcocctcaggt tctttggagg atacctatag gtttgcacc tcacaggcca ttacatgtca 6840
aaaaactgcc ccccaaaagc ccaaggaaga tccatttaaa gattatgtat tttgggaggt 6900
taatttmetaa gaaaagtgtt ctgcagattt agatcagttt ccactgggtc gcaaatmttt 6960
attacaggca ggatataggg cacgtcctaa atttaagca ggtaaacgta gtgcaccctc 7020
agcatctacc actacaccag caaaacgtaa aaaaactaaa aagtaatgga tgtgtatgta 7080
atacatgtgt ctgtatgtgt atgtgctgtt gctgtattgt atatgtgtgt gtttgtgtgt 7140
tatatatggt atatgtatgt ttatgtatgc gtgtgtactt gtatatatgt atagtatgtt 7200
atgtgtgtat gtatgtatg tatgttaata aatatgtgta tacctgtgtg tgttgtgtat 7260
gttgcctta tatacacctt attagtaaca tactattact atttataaaa ctattgttcc 7320
tacttgttcc tacttgttcc tgctcctccc aatagtcctg tacttatttc tgcctataat 7380
ttagggtgca cgccatagta aaagtgtgac acccggctccg ttttttgcaa ctaaagctac 7440
tccatmttg a ttttatgcag ccattmttaaa tccctaaccg ttttcggttg cattgtmtta 7500
acatgctagt acaactatgc tgatgcagta gttctgctgt ttttggtttc ctgaatacta 7560
gtttttgcca acattctggc ttgtagtmtt ctgcctaaca caccttgcca acatataatc 7620
cagtccaact ttgcaattat actatgaatc atgtttgttt aaatacaact gtagttcaac 7680
tatgtgtcat gcacatatat tatattatcc tacacacctt aaactgcttt taggcacata 7740
ttttgtagat tatctatat cttgattgca gtgctggctt ttgcacatgt ttaaaactgcc 7800
aaggttgtgt catgcattat aaataagttg tatgttactc atataattaa ttgcatatag 7860
gtattacacc gttttcggtt acagttttac aagcaattgt tctttttata ct 7912

```

&lt;210&gt; SEQ ID NO 23

&lt;211&gt; LENGTH: 7909

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Human Papillomavirus HPV33 (M12732.1)

&lt;400&gt; SEQUENCE: 23

```

gtaaaactata atgccaagtt ttaaaaaagt aggggtgtaac cgaaagcggg tcaaccgaaa 60
acgggtgcata tataaagcaa acattttgca gtaaggtagt gcacgactat gtttcaagac 120
actgaggaaa aaccacgaac attgcatgat ttgtgccaag cattggagac aactatacac 180
aacattgaac tacagtgcgt ggaatgcaaa aaacctttgc aacgatctga ggtatatgat 240
tttgcatttg cagatttaac agttgtatat agagagggaa atccatttgg aatatgtaaa 300
ctgtgtttgc ggttcttata taaaattagt gaatatagac attataatta ttctgtatat 360
ggaaatacat tagaacaacac agttaaaaaa cctttaaatg aaatattaat taggtgtatt 420

```

-continued

---

atatgtcaaa gacctttgtg tctcaagaa aaaaaacgac atgtggattt aaacaaacga	480
tttcataata tttcgggtcg ttgggcaggg cgctgtgcgg cgtgttgag gtccccagct	540
agagaaactg cactgtgacg tgtaaaaacg ccatgagagg acacaagcca acgttaaagg	600
aatatgtttt agatttatat cctgaaccaa ctgacctata ctgctatgag caattaagtg	660
acagctcaga tgaggatgaa ggcttgacc gccagatgg acaagcaca ccagccacag	720
ctgattacta cattgtaacc tgtgtgcaca cttgtaacac cacagttcgt ttatgtgtca	780
acagtacagc aagtgcacta cgaaccatac agcaactact tatgggcaca gtgaatattg	840
tgtgccctac ctgtgcacaa caataaacat catctacaat ggccgatcct gaaggtaaca	900
atggggctgg gatgggggtg actggttggt ttgaggtaga agcagtcata gagagaagaa	960
caggagataa tatttcagaa gatgaggatg aaacagcaga tgacagtggc acggatttac	1020
tagagtttat agatgattct atggaaaata gtatacaggc agacacagag gcagcccggg	1080
cattgtttaa tatacaggaa ggggaggatg atttaaatgc tgtgtgtgca ctaaacgaa	1140
agtttgccgc atgttcacaa agtgctgcgg aggacgttgt tgatcgtgct gcaaacccgt	1200
gtagaacgtc tattaataaa aataaagaat gcacatacag aaaacgaaaa atagatgagc	1260
tagaagacag cggatatggc aatactgaag tggaaactca gcagatgta caacaggtag	1320
aaagtcaaaa tggcgacaca aacttaaatg acttagaatc tagtggggtg ggggatgatt	1380
cagaagtaag ctgtgagaca aatgtagata gctgtgaaaa tgttacgttg caggaaatta	1440
gtaatgttct acatagtagt aatacaaaaag caaatatatt atataaatTT aaagaggcct	1500
atggaataag ttttatggaa ttagtaagac cttttaaag tgataaaaca agctgtacag	1560
attgggtgat aacaggatat ggaattagtc catcagtagc agaaagtta aaagtattaa	1620
ttaaacagca tagtttgat actcatttac aatgtttaac ttgcataga ggaataataa	1680
tattattggt aattagattt aggtgtagca aaaacaggtt aacagtagca aaactaatga	1740
gtaatttatt atcaatacct gaaacatgta tggttataga gccacaaaa ttacggagcc	1800
aaacatgtgc attgtattgg tttagaacag caatgtcaaa cattagtgat gtacaaggta	1860
caacacctga atggatagat agactaactg ttttacaaca tagctttaat gataatatat	1920
ttgatttaag tgaatggta cagtgggcac atgataacga gttaacggac gatagtgaca	1980
ttgcatatta ttatgcacaa cttgcagatt caaatagtaa tgctgctgca tttttaaaa	2040
gtaactcaca agcaaaaata gtaaaggact gtggaataat gtgtagacat tataaaaaag	2100
cagaaaaacg taaaatgtca ataggacaat ggatacaaaag tagatgtgaa aaaacaaatg	2160
atggaggaaa ttggagacca atagtacagt tgttaagata tcaaaacatt gaatttacag	2220
catttttagg tgcatttaaa aagtttttaa aaggataacc aaaaaaacg tgtatgctaa	2280
tttgtggacc agcaataca ggaagtcac attttggaa gagtttaata cagtttttaa	2340
aagggtgtgt tatatcatgt gtaaattcta aaagtcactt ttggttgac ccattatcag	2400
atgcaaaaat aggaatgata gatgatgtaa cgccaataag ttggacatat atagatgatt	2460
acatgagaaa tgcgttagat ggaatgaaa tttcaataga tgtgaaacat agggcattag	2520
tgcaattaaa atgtccacca ctgcttctta cctcaaatc aaatgcaggc acagactcta	2580
gatggccata tttacatagt agattaacag tatttgaatt taaaaatcca ttcccatttg	2640
atgaaaatgg taaccagtg tatgcaataa atgatgaaaa ttggaatcc tttttctcaa	2700

-continued

---

ggacgtggtg caaattagat ttaatagagg aagaggacaa ggaaaacat ggaggaata	2760
tcagcacgtt taaatgcagt gcaggagaaa atactagatc tttacgaagc tgataaaact	2820
gatttaccat cacaaattga acattggaaa ctgatacgca tggagtgtgc tttattgtat	2880
acagccaaac aaatgggatt ttcacattta tgccaccagg tgggtgccttc tttgttagca	2940
tcaaagacca aagcatttca agtaattgaa ctacaaatgg cattagagac attaagtaaa	3000
tcacagtata gtacaagcca atggacattg caacaaacaa gcttagaggt gtggctttgt	3060
gaaccaccaa aatgttttaa aaaacaagga gaaacagtaa ctgtgcaata tgacaatgac	3120
aaaaaaaaata caatggatta tacaactgg ggtgaaatat atattataga ggaagataca	3180
tgtactatgg ttacagggaa agtagattat ataggatgtg attatataca taactgtgaa	3240
aaggatatatt ttaaatattt taaagaggat gctgcaaagt attctaaac acaaatgtgg	3300
gaagtacatg tgggtggtca ggtaattggt tgcctacgt ctatatctag caaccaata	3360
tccactactg aaactgctga catacagaca gacaacgata accgaccacc acaagcagcg	3420
gccaaacgac gacgacctgc agacaccaca gacaccgccc agccccttac aaagctgttc	3480
tgtgcagacc ccgccttggg caatagaaca gcacgtactg caactaactg cacaaacaag	3540
cagcggactg tgtgtagtto taacgltgca cctatagtgc atttaaaagg tgaatcaaat	3600
agtttaaaat gtttaagata cagattaaaa ccttataaag agttgtatag ttctatgtca	3660
tccacctggc attggaccag tgacaacaaa aatagtaaaa atggaattgt aactgtaaca	3720
tttgtaaactg aacagcaaca acaaatgttt ttaggtaccg taaaaatacc acctactgtg	3780
caaataagta ctggatttat gacattataa gtgtacatca caagccaata tgtgctgcta	3840
attgtatata acctatgatat ttgtttttgt attatgtttt atattgtttt tatgcttacc	3900
cttattatta cgtcctttaa tactttccat ttctacctat gcttgggtgc tgggtgtggt	3960
attgctgctt tgggtgtttg tgggatctcc tttaaaaatt ttttttgcct atttgttgtt	4020
tttatattta ccaatgatgt gtattaattt tcatgcacag catatgacac aacaagagta	4080
atgtatatac atgtatatac tgtttgtata tatgtgcaca tgggtgggtgt ttaacattgt	4140
tgtgtgtatt ttatgttttt tttttttgta ttactaataa atacctttat atttttagcag	4200
tgtattatta tgagacacaa acgatctaca aggcgcaagc gtgcactgc aacacaacta	4260
taccaaaacat gcaagccac aggcacctgc ccaccgatg ttattcctaa agtgggaagga	4320
agtaccatag cagatcaaat tcttaaatat ggcagtttag gggttttttt tgggtgttta	4380
ggatttgga caggctctgg ttcagggtga aggactggct atgtacctat tggactgac	4440
ccacctacag ctgcaatccc cttgcagcct atacgtcctc cggttactgt agacctgtt	4500
ggacctttag actcgtctat agtgtcatta atagaagaaa caagttttat agaggcaggt	4560
gcaccagccc catctatccc tacacatca ggttttgatg ttactacatc tgcagatact	4620
acacctgcaa ttattaatgt ttcatctggt ggggagtcac ctattcaaac tatttctaca	4680
catttaaac ccacatttac tgaacctct gtactacacc ctccagcgcc tgcagaagcc	4740
tctggacatt ttatattttc ttcccctact gttagcacac aaagtatga aaacatacca	4800
atggatacct ttgttttttc cacagacagt agtaatgtaa catcaagcac gccattcca	4860
gggtctcgcc ctgtggcacg ccttggttta tatagtcgca ataccaaca ggttaaggtt	4920
ggtgacctg cttttttaac atcgccctcat aaacttataa catatgataa tctgcattt	4980

-continued

---

gaaagctttg accctgaaga cacattacaa tttcaacata gtgatatac acctgctcct	5040
gatcctgact ttctagatat tattgcatta cataggcctg ctattacatc tcgtagacat	5100
actgtgcgctt tttagtagag aggtcaaaaa gccacactta aaactcgagc tggtaaacaa	5160
attggagcta gaatacatta ttatcaggat ttaagtccta ttgtgccttt agaccacacc	5220
gtgccaaaatg aacaatatga attacagcct ttacatgata cttctacatc gtcttatagt	5280
attaatgatg gtttgtatga tggttatgct gacgatgtgg ataatgtaca caccccaatg	5340
caacactcat acagtagctt tgcaacaaca cgtaccagca atgtgtctat acctttaaat	5400
acaggatttg atactcctgt tatgtctggc cctgatatac cttccccttt atttcccaca	5460
tctagcccat ttgttccat ttgcctcttt tttccttttg acaccattgt ttagtagcgt	5520
gctgactttg ttttacatcc tagttatctt attttacgct gcaggcgtaa acgttttcca	5580
tattttttta cagatgtccg tgtggcggcc tagtgaggcc acagtgtacc tgcctcctgt	5640
acctgatctc aaagtgtgca gcaactgatga atagtgtctc gcacaagca tttattatta	5700
tgtgtgtagt tccagacttc ttgctgttgg ccatccatat tttctatta aaaatcctac	5760
taacgctaaa aaattattgg tacccaaagt atcaggcttg caatataggg ttttaggggt	5820
cogtttacca gatcctaata aatttggatt tccctgacacc tccctttata acctgatac	5880
acaacgatta gtatgggcat gtgtaggcct tgaatatagg agagggcagc cattaggcgt	5940
tggcataagt ggtcatcctt tattaacaaa atttgatgac actgaaaccg gtaacaagta	6000
tccctggacaa ccgggtgctg ataatatggga atgtttatcc atggattata aacaaacaca	6060
gttatgttta cttggatgta agcctccaac aggggaacat tggggtaaag gtgttgcttg	6120
tactaatgca gcacctgcca atgattgtcc acctttagaa cttataaata ctattattga	6180
ggatggtgat atggtggaca caggatttgg ttgcatggat tttaaaacat tgcaggctaa	6240
taaaagtgat gttcctattg atatttgtgg cagtacatgc aaatatccag attattttaa	6300
aatgactagt gagccttatg gtgatagttt atttttcttt cttcgacgtg aacaaatgtt	6360
tgtaagacac ttttttaata gggctggtac attaggagag gctgttcccg atgacctgta	6420
cattaaaggt tcaggaacta ctgcctctat tcaaagcagt gcttttttcc cactcctag	6480
tggatcaatg gttacttccg aatctcagtt atttaataag ccatattggc tacaacgtgc	6540
acaaggtcat aataatggta tttgttgggg caatcaggta tttgttactg tggtagatac	6600
cactcgcagt actaatatga ctttatgcac acaagtaact agtgacagta catataaaaa	6660
tgaaaatctt aaagaatata taagacatgt tgaagaatat gatctacagt ttgtttttca	6720
actatgcaaa gttaccttaa ctgcagaagt tatgacatat attcatgcta tgaatccaga	6780
tatttttagaa gattggcaat ttggtttaac acctcctcca tctgctagtt tacaggatac	6840
ctataggttt gttacctctc aggcatttac gtgtcaaaaa acagtacctc caaaggaaaa	6900
ggaagacccc ttaggtaaat atacattttg ggaagtggat ttaaaggaaa aattttcagc	6960
agatttagat cagtttcctt tgggacgcaa gtttttatta caggcaggtc ttaaagcaaa	7020
acctaaaact aaacgtgcag cccccacatc caccgcaca tcgtctgcaa aacgcaaaaa	7080
ggttaaaaaa taacactttg tgtaattgtg ttatgttgtt gttttgttct gtctatgtac	7140
tttgtgtgtg tgtgtgtgtg tgtgtttgt tttttgtgta tgtgttaca tgtatgttat	7200
gtttagtgtt actgtgtttg ttttatgtgt acttgtttgt gtgcatgttc tatgtacttg	7260

-continued

---

```

tcagtttct gtttgtgat atgtaataa aacatttgt gtatttgta aactatttgt 7320
atgtatgta tgtatatggg tgtacctata tgagtaagga gttgtattgc ttgccctacc 7380
ctgcattgca atgtacctac ctttatttcc ctatatttgt agtaacctaca tgttttagtat 7440
tgctttacct tttgacatac tagtgtccat attgtacaat ttcctccatt ttgtatgect 7500
aaccgttttc ggttacttgg catacatacc ctatgacatt ggcagaacag ttaatccttt 7560
tctttcctgc actgtgtttg tctgtacttg ctgcattggc atacataccc tatgacattg 7620
gcagaacagt taatcctttt ctttctgca ctgtgtttgt ctgtacttgc tgcattgact 7680
catatataca tgcagtgcaa ttgcaaaaata cttaattgta ctaatagttt acacatgctt 7740
ttaggcacat atttttactt tactttcaaa ccttaagtgc agttttggct tacacaattg 7800
ctttgatgac caaactatgc cttgtaaaag tgagtcacta cctgtttatt accaggtgtg 7860
gactaacctt tttaggteat attggtcatt tataatcctt tatataata 7909

```

&lt;210&gt; SEQ ID NO 24

&lt;211&gt; LENGTH: 7858

&lt;212&gt; TYPE: DNA

&lt;213&gt; ORGANISM: Human Papillomavirus HPV45 (X74479.1)

&lt;400&gt; SEQUENCE: 24

```

aatactttta acaattatac tacataaaaa aggggtgtaac cgaaaacggg tgcaacccaaa 60
aacgggtgcat ataaaaagctt tgtggaaaag tgcattacag gatggcgcgc tttgacgatc 120
caaagcaacg accctacaag ctaccagatt tgtgcacaga attgaatata tcaactacaag 180
acgtatctat tgcctgtgta tattgcaaag caacattgga acgcacagag gtatatcaat 240
ttgcttttaa agatttatgt atagtgtata gagactgtat agcatatgct gcatgccata 300
aatgtataga cttttattcc agaattagag aattaagata ttattcaaac tctgtatatg 360
gagagacact ggaaaaaata actaatacag agttgtataa tttgttaata aggtgcctgc 420
ggtgccagaa accattgaac ccagcagaaa aacgtagaca ccttaaggac aaacgaagat 480
ttcacagcat agctggacag taccgagggc agtgtaatac atggtgtgac caggcacggc 540
aagaaagact tgcgacagct agggaaaacac aagtatagca ataagtatgc atggaccccg 600
ggaaacactg caagaattg tattgcattt ggaacctcag aatgaattag atcctgttga 660
cctgttgtgt tacgagcaat taagcgagtc agaggaggaa aacgatgaag cagatggagt 720
tagtcatgca caactaccag cccgacgagc cgaaccacag cgtcacaaaa ttttgtgtgt 780
atggtgtaag tgtgacggca gaattgagct tacagtagag agctcggcag aggaccttag 840
aacactacag cagctgtttt tgagcacctt gtcctttgtg tgtccgtggg gtgcaactaa 900
ccaataatct acaatggcgg atccagaagg taccgacggg gagggaaacgg ggtgtaattg 960
ctggttcttt gtagaaacaa ttgtagagaa aaaaacaggg gatgtaatat cagatgatga 1020
ggatgaaact gcaacagata cagggtcgga tatggtagat tttattgaca cacaattatc 1080
catttgtgaa caggcagagc aagagacagc acaggcattg ttccatgcgc aggaagtcca 1140
gaatgatgca caggtgttgc atcttttaaa acgaaagttt gcaggaggca gcaaggaaaa 1200
cagtcatta ggggagcagc taagtgtgga tacggatcta agtccacggg tacaagaaat 1260
ttcattaat agtgggcaca aaaaagcaaa acgacggttg tttacaatat cagatagttg 1320
ctatggctgt tctgaagtgg aagctgcaga gactcaggtg actgtaaaaa ctaatgcccga 1380

```

-continued

---

aaatggcggc agtgtacata gtacacaaag tagtgggtggg gatagtagtg acaatgcaga	1440
aaatgtagat ccgcattgca gtattacaga actaaaggag ctattacaag caagtaacaa	1500
aaaggctgca atgctggcag tatttaaaga catatatggg ctgtcattta cggatttggg	1560
tagaaatfff aaaagtgata aaacaacatg tacagattgg gtaatggcta tatttggagt	1620
taatccaacg gtagcagaag gctttaaagc attaattaaa ccagcaacgt tatacgcca	1680
tatccaatgt ttagattgta aatggggagt attaatatta gctttattaa gatataaatg	1740
tggcaaaaat agactaactg ttgcaaaagg cttaagcaca ttgttgacg tacctgaaac	1800
atgtatgta attgaaccac caaaattgag aagtagtggt gcagcattat actggtatag	1860
aacaggtata tccaatatta gtgaagtaag tggagacaca cctgagtgga taaaaagact	1920
gacaattatt caacatggta ttgacgatag taattttgat ttgtcagaca tggtgcaatg	1980
ggcatttgat aatgacctta cagatgaaag tgatattgca tttcaatag cccaattagc	2040
agactgcaac agtaatgcag ctgcattttt aaaaagtaac tgccaagcca aatatttaa	2100
agattgtgct gtaatgtgta gacattataa aagagcaca aaacgcaaaa tgaatatgtc	2160
tcaatggatt aaatatagat gttccaaaat agatgaaggt ggggattgga gacctatag	2220
acaattccta agatatcagg gagtagaatt tattagcttt ttaaggcac taaaggaatt	2280
tcttaaagga acacaaaaa aaaattgtat actgttatat ggacctgca atacaggaaa	2340
atcgtatfff ggaatgagtt ttatacattt cctacaaggt gcaataatat catttga	2400
ttcaaacagc catttttggg tagaacctgt agcagatact aaggtagcca tgttgatga	2460
tgccacacac acgtgttggc catattttga taattatag agaaatgcat tagatggtaa	2520
tcctataagt atagacagaa agcataaac attattacag ctaaaatgct ctccaatcct	2580
attaacatcc aatattgatc cagcaaaaag taataaatgg ccatatttag aaagtagggt	2640
gacggatfff acatttccac atgcatttcc atttgataaa aatggtaac cagtatatga	2700
aataaatgat aaaaattgga aatgtttttt tgaaggaca tggtcagat tagatttga	2760
cgaggacgat gaagatgcag acaccgaagg aatcccttcc ggaacgttta agtgcgttac	2820
aggcaaaaat actagaccac tatgaaaatg acagtaaaag cataaacagc caaataagtt	2880
attggcaact tatacgtttg gaaaatgcaa tactatttac agcaaggaa catggtatta	2940
ccaaactaaa ccaccaggtg gtgcctccta ttaacatttc aaaaagcaaa gcacataaag	3000
ctattgaact gcaaatggcc ttaaagggcc ttgcacaaag caagtataac aatgaggaat	3060
ggacactgca agatacatgc gaggaactat ggaatacaga accgtcgcag tgttttaaaa	3120
aaggcggtaa aaccgtgacg gtatactttg atggcaacaa ggacaactgt atgaaactag	3180
tagtatggga cagtatatat tataatactg agacagggat atgggacaaa acagcagcat	3240
gtgttagcta ttgggtgta tattatataa aagatggaga taccacatat tatgtacaat	3300
ttaaagcga atgtgagaaa tatggaata gtaatacgtg ggaagtacaa tatgggggca	3360
atgtaattga ttgtaatgac tctatgtgca gtaccagtga cgacacggta tccgtactc	3420
agattgttag acagctacaa cacgctcca cgtcgacccc caaaaccgca tccgtgggca	3480
ccccaaaacc ccacatccag acgcccggta ctaagcgacc tagacagtgt ggactcacag	3540
agcagacca cggacgtgtc aacaccacg tgcacaaccc gctcctgtgt tcaagtacaa	3600
gtaacaacaa aagaaggaaa gtgtgtagtg gtaaacctac gcctataata cacttaaaag	3660

-continued

---

gtgacaaaa cagtttgaaa tgtttaagat ataggctacg caaatatgca gaccattact	3720
cagaaatata cttccacctgg cattggacag gttgtaataa aaacactggt atattaactg	3780
taacataata tagtgaggta caaagaaata cctttttgga thtagttact attcctaaca	3840
gtgtacaaat ctgggtggga tacatgacta tatgaatctg tatattgtat acagtatgta	3900
acattactat gctatcttta gtgtttttat tgtgcttttc tgtgtgcctt tatgtgtgct	3960
gcaatgtccc gcttgtgcag tctgtctatg tgtgtgcttt tgcttggttg ttgggttttc	4020
ttttatagtg tgttattaca tccccattaa cagcatttgc tgtatacatt tgttgcatt	4080
tactacctat gtttgtatta catatgcatg ctttacacac catacaataa ttactataat	4140
gtacagtaca gtgtaacata cctgtgatgt gcatgttgtt gtatttttgt atttttgtat	4200
ttttgtattt ttgtatttta tatgtttaat aaaccatggt atcccaccgt gcagcacgtc	4260
gcaagcgggc ctctgcaact gacttatata gaacatgtaa gcaatccggt acgtgcccc	4320
ctgatgttat taacaaagtg gaaggcacia ccttagctga taaaatttta cagtgttcta	4380
gccttgggat atttttgggt ggccttgcca ttggtagcgg cagtggttct ggaggccgta	4440
cgggctatgt acccttaggg ggcaggctca atactgttgt ggatgttggc cccactaggc	4500
cacctgtggt tattgaacct gtagggccta ctgatccatc tattgttacg ttggtagagg	4560
attccagtgt tgttgctctt ggtgctccgg tccccacatt taccggaacc tctgggtttg	4620
aaattacgtc ttctggtact accacaccag ctgtgttggga catcacacct accgtggact	4680
ctgtttctat ttctgcaact agttttacaa atcctgcatt ttctgatccc tctattattg	4740
aggtgcccc aacaggggag gtatcaggta atattttgt ttgtacacca acatcgggca	4800
gccatggata tgaggaataa cctttacaaa cttttgcac tcttggttca ggtacggaac	4860
ccattagtag taccctctc cctactgtgc ggcgggtacg gggtccccgc ctgtatagta	4920
gggctaatac acaggtccgt gtgtccacct cacagttttt aacacatccc tcatcgttgg	4980
ttacatttga taatccagct tatgagcccc tggacaccac actatccttt gacctacca	5040
gtaatgttcc tgattccgat tttatggata ttattcgttt gcataggcca gcattatcct	5100
ctagacgttg cactgttaga ttttagtagat tgggtcaaag ggcaaccatg tttacacgta	5160
gtggtaaaca aatagggggt agggtagatt tttaccatga tataagcccc attgctgcta	5220
cagaggaaat tgaattgcag cctttaatta gtgctacaaa tgatagtgac ctgtttgatg	5280
tatatgcaga cttccacct cctgcgtcca ctacacctag cactatacac aaatcattta	5340
catatccaaa gtattccttg accatgcctt ctactgctgc atcctcttac agtaatgtta	5400
cagtaccatt aacatctgca tgggatgtac ctatatatac tggcccggaac atttatattgc	5460
catcccatac tcctatgtgg cctagtagat ctctaccaa tgcttccacc accacctata	5520
taggtattca tggcacacia tattatttat ggccatggta ttattatttt cctaaaaaac	5580
gtaaactgat tccctatttt tttgcagatg gctttgtggc ggctagtga cagtacggt	5640
tatcttccac caccttctgt ggccagagtt gtcagcaact atgattatgt gtctgcaca	5700
agcatatttt atcatgcagg cagttcccga ttattaactg taggcaatcc atattttagg	5760
gtgttaccta atgggtcagg taataaacag gctgttccta aggtatccgc atatcagtat	5820
aggggtttta gagtagcttt acccgatcct aataaatttg gattacctga ttctactata	5880
tataatcctg aaacacaacg tttgggttgg gcatgtgtag gtatggaaat tggctcgtggg	5940

-continued

---

cagcctttag gtattggcct aagtggccat ccattttata ataaattgga tgatacagaa	6000
agtgtcatg cagctacagc tgttattacg caggatgta gggataatgt gtcagttgat	6060
tataagcaaa cacagctgtg tatttttaggt tgtgtacctg ctattggtga gcactgggcc	6120
aagggcacac tttgtaaacc tgcacaattg caacctgggtg actgtcctcc tttggaactt	6180
aaaaacacca ttattgagga tggatgatg gtggatacag gttatggggc aatggatttt	6240
agtacattgc aggatacaaa gtgcgaggtt ccattagaca tttgtcaatc catctgtaa	6300
tatccagatt atttgcaaat gtctgctgat ccctatgggg attctatggt tttttgccta	6360
cgccgtgaac aactgtttgc aagacatttt tggaaatagg caggtgttat gggtgacaca	6420
gtacctacgg acctatata taaaggcact agcgtaata tgcgtgaaac cctggcagt	6480
tgtgtgtatt ccccttctcc cagtggctct attattactt ctgattctca attattta	6540
aagccatatt ggttacataa gggccagggc cataacaatg gtatttgttg gcataatcag	6600
ttgtttgta ctgtagtga cactaccgc agtactaatt taacattatg tgctctaca	6660
caaaatcctg tgccaagtac atatgacct actaagtta agcagtatag tagacatgtg	6720
gaggaatag atttacagtt tatttttcag ttgtgacta ttactttaac tgcagaggtt	6780
atgcatata tccatagat gaatagtagt atattagaaa attggaattt tgggtgccct	6840
ccaccaccta ctacaagttt ggtggatata tctgttttg tgcaatcagt tctgtttacc	6900
tgtcaaaagg atactacacc tccagaaaag caggatccat atgataaatt aaagttttgg	6960
actgttgacc taaaggaaaa attttctcc gatttggatc aatatccct tggtcgaaag	7020
tttttagttc aggctgggtt acgtcgtagg cctaccatag gacctcgtaa gcgtcctgct	7080
gcttccacgt ctactgcato tactgcatct aggcctgcca aacgtgtacg tatacgtagt	7140
aagaaataat atgttagcac atatatgtat gtttgtatgt atggttttgt atgttgatg	7200
tatgtatgta tttgtgtgat atattactgt attttgtttg tttgcgtgcg tgtatgatg	7260
aatgtgcctt gtggcatgta tgggtgtact gtacataatt gtggtattaa ataaagtatg	7320
ctaatagtgt tgtgtagggt tgcacccttg tgagtaacaa tactatttgt gtgtatgtgt	7380
attgctttgt accctatatt ctttctgta tttcaagtta taaacttga tactacacag	7440
catccatttt acttataatc ctccattttg ctgtgcaacc gatttcggtt gcctgtggct	7500
tatatgtgac cttttaaaca taatacctaa actggcacat ttacaacccc tacatagttt	7560
aacctactgg cgcgccttct tggcgtaacat gtggcacacc tggatttagt cattttcctg	7620
tccaggtgta ctaaaacaat ggctgcaca actgtatcca caccctatgt aataaaactg	7680
cttttaggca catattttag tctgttttta cctgtgctaa ttgtataatt ggcgtgtaga	7740
accactttct tatccaacaa tctgtctact tgttacataa actataaact gactcactta	7800
tacatacata gtttatgcaa ccgaaaaagg ttgggccta taacacatac cttttctt	7858

---

1. A method for detecting an infectious viral polynucleotide in a biological sample comprising:
  - separating, extracting or otherwise obtaining a polynucleotide from said sample,
  - Molecular Combing said polynucleotide to form a stretched polynucleotide,
  - contacting said stretched polynucleotide with one or more probes that recognize an infectious viral polynucleotide sequence,
  - detecting hybridization of the probes to the combed sample; thereby detecting an infectious viral polynucleotide.
2. The method of claim 1, wherein said biological sample is a tissue sample, cell(s), serum, blood, CSF, or synovial fluid sample obtained from a human, a non-human mammal or a bird.
3. The method of claim 1, wherein said polynucleotide is an infectious or a non-infectious viral DNA extracted from tissue, cell(s), serum, blood, CSF, or synovial fluid sample of a human, a non-human mammal, or a bird.
4. The method of claim 1, wherein said polynucleotide is integrated into the DNA or episomal or a biological fluid of a human or a non-human mammal or a bird.
5. The method of claim 1, wherein said polynucleotide is detected with one or more probes that bind to a viral DNA of a double stranded DNA virus.
6. The method of claim 1, wherein said polynucleotide is detected with one or more probes that bind to DNA of a herpes virus, a papilloma virus, a hepatitis virus or a retrovirus.
7. The method of claim 1, wherein said one or more probes comprises a set of probes that bind to at least 1 kb of the genomic DNA of a virus.
8. The method of claim 1, wherein said one or more probes comprises at least two sets of probes that are tagged with different labels selected so as to permit the identification of different portions or segments or rearranged fragments of a viral genome.
9. A method for detecting or identifying an infectious or genomic viral polynucleotide sequence in a mammalian cell, tissue or biological fluid comprising using Molecular Comb-

ing to detect the presence or the quantity of infectious viral polynucleotide or genomic viral polynucleotide in a cell, tissue or biological fluid.

10. A method for quantifying an infectious or genomic viral polynucleotide sequence comprising using Molecular Combing to detect quantity of infectious viral polynucleotide or genomic viral polynucleotide in a cell, tissue or biological fluid compared to an uninfected control sample or an otherwise similar sample obtained at a different point in time.

11. A method for detecting or following viral presence or virus replication or viral genomic rearrangements in a mammalian cell comprising Molecular Combing for the presence of latent or replicating viral DNA or rearranged viral DNA in a cell, tissue or biological fluid.

12. A method for evaluating the efficacy of an anti-viral treatment comprising:

detecting using Molecular Combing the presence, arrangement or quantity of infectious or genomic viral DNA in a sample obtained from a subject,

treating said subject with an anti-viral agent, and re-evaluating using Molecular Combing the presence, arrangement or quantity of infectious viral or genomic viral DNA in said subject.

13. A set of probes covering 5-100% of the HSV or HIV genome.

14. A kit for performing Molecular Combing comprising a Molecular Combing apparatus and/or reagents, one or more probes that bind to an infectious viral polynucleotide for detecting or identifying genomic viral DNA in a combed DNA molecule, and

optionally one or more cell, tissue or biological fluid sample(s), and

optionally software for detecting or classifying the infectious sequences.

15. Biological materials HSV-B4 (CNCM I-4298), HSV-B19 (CNCM I-4299), HSV-Sc54 (CNCM I-4300), or HSV-P4 (CNCM I-4301).

\* \* \* \* \*

专利名称(译)	通过分子梳理检测基因组和感染性病毒DNA来诊断病毒感染		
公开(公告)号	<a href="#">US20110287423A1</a>	公开(公告)日	2011-11-24
申请号	US13/092210	申请日	2011-04-22
[标]申请(专利权)人(译)	基因组影像公司		
申请(专利权)人(译)	基因组VISION		
当前申请(专利权)人(译)	基因组VISION		
[标]发明人	MAHIET CHARLOTTE SALVAIRE FABRICE CONSEILLER EMMANUEL BARRADEAU SEBASTIEN		
发明人	MAHIET, CHARLOTTE SALVAIRE, FABRICE CONSEILLER, EMMANUEL BARRADEAU, SEBASTIEN		
IPC分类号	C12Q1/68 C12N7/00 G01N33/53		
CPC分类号	C12Q1/705 C12Q1/703 C12Q1/6816 C12Q2533/101 A61P31/20 C12Q1/6883 C12Q1/708 C12Q2600/156		
优先权	61/327397 2010-04-23 US		
外部链接	<a href="#">Espacenet</a> <a href="#">USPTO</a>		

摘要(译)

一种使用Molecular Combing或其他核酸延伸方法与探针，尤其是核酸探针一起在体外检测感染的真核细胞，组织或生物液体中DNA病毒或病毒衍生DNA的基因组的存在的方法。特殊的设计。一种通过在病毒感染的细胞，组织或生物液体中存在基因组病毒或病毒衍生的DNA多核苷酸来体外监测抗病毒治疗的效果的方法。使用Molecular Combing和DNA杂交检测病毒的感染形式。包含用于实施这些方法的探针的试剂盒和包含该探针的组合物。

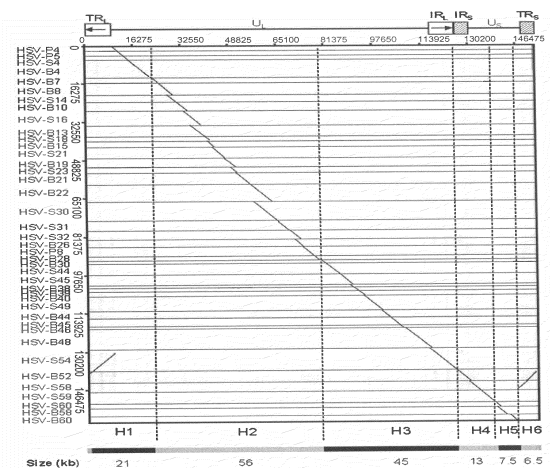


Figure 1