



US006528066B1

(12) **United States Patent**
Grose et al.

(10) **Patent No.:** **US 6,528,066 B1**
(45) **Date of Patent:** **Mar. 4, 2003**

- (54) **VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE**
- (75) Inventors: **Charles F. Grose**, Iowa City, IA (US); **Richard Santos**, Iowa City, IA (US)
- (73) Assignee: **University of Iowa Research Foundation**, Iowa City, IA (US)
- (*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 47 days.

- (21) Appl. No.: **09/661,596**
- (22) Filed: **Sep. 14, 2000**

Related U.S. Application Data

- (60) Provisional application No. 60/153,779, filed on Sep. 14, 1999.
- (51) **Int. Cl.**⁷ **A61K 39/245**; A61K 39/12
- (52) **U.S. Cl.** **424/229.1**; 424/230.1; 424/204.1; 435/91.1; 435/91.33; 435/89; 536/23.72
- (58) **Field of Search** 435/89, 91.1, 91.33; 424/229.1, 230.1, 204.1; 536/23.72

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,985,615 A	10/1976	Kubo
5,462,734 A	10/1995	Letchworth, III et al.
5,595,890 A	1/1997	Newton et al.
5,849,476 A	12/1998	Shiraki et al.
5,952,174 A	9/1999	Nikiforov et al.
6,087,170 A	* 7/2000	Kemble 435/368

OTHER PUBLICATIONS

A. M. Arvin et al., "Live Attenuated Varicella Vaccine," *Annu. Rev. Microbiol.*, 50, 59-100 (1996).

A. M. Arvin et al., "Immunity to Varicella-Zoster Viral Glycoproteins, gp I (gp 90/58) and gp III (gp 118), and to a Nonglycosylated Protein, p. 170," *J. Immunol.*, 137, 1346-1351 (1986).

ATCC CCL-171, "Homo sapiens (human)," [online]. Retrieved on May 16, 2001. Retrieved from the Internet:<URL: <http://phage.atcc.org/cgi-bin/searchengine/longview.cgi?view=ce,317407,CCL-171&text=CCL-.1,3> pages.

ATCC VR-568, "Varicella-Zoster," [online]. Retrieved on Jun. 5, 2001. Retrieved from the Internet:<URL: <http://phage.atcc.org/cgi-bin/searchengine/longview.cgi?view=av,343894,VR-586&text=VR-586>>, 2 pages.

ATCC VR-795, "Varicella-Zoster deposited as Varicella," [online]. Retrieved on May 16, 2001. Retrieved from the Internet:<URL: <http://phage.atcc.org/cgi-bin/searchengine/longview.cgi?view=av,476976,VR-795&text=VR-7.9>>, 2 pages.

ATCC VRL-1367, "Varicella-Zoster," [online]. Retrieved on Jun. 5, 2001. Retrieved from the Internet:<URL: <http://phage.atcc.org/cgi-bin/searchengine/longview.cgi?view=av,871705,VR-1367&text=VR-1367>>, 2 pages.

M. F. Bachman et al., "The influence of Antigen Organization on B Cell Responsiveness," *Science*, 262, 1448-1451 (1993).

R. E. Bergen et al., "Human T Cells Recognize Multiple Epitopes of an Immediate Early/Tegument Protein (IE62) and Glycoprotein I of Varicella Zoster Virus," *Viral Immunol.*, 4, 151-166 (1991).

J.I. Cohen et al., "Generation of varicella-zoster virus (VZV) and viral mutants from cosmid DNAs: VZV thymidylate synthetase is not essential for replication in vitro," *Proc. Natl. Acad. Sci. USA*, 90, 7376-7380 (1993).

N. L. Cole et al., "Colchicine treatment in the preparation of varicella-zoster virus inocula," *J. Virol. Methods*, 36, 111-118 (1992).

A.J. Davison et al., "The Complete DNA Sequence of Varicella-Zoster Virus," *J. Gen. Virol.*, 67, 1759-1816 (1986).

K. S. Dingwell et al., "Herpes simplex virus glycoproteins E and I facilitate cell-to-cell spread in vivo across junctions of cultured cells," *J. Virol.*, 68 834-845 (1994).

K. M. Duus et al., "Cell Surface Expression and Fusion by the Varicella-Zoster Virus gH:gL Glycoprotein Complex: Analysis by Laser Scanning Confocal Microscopy," *Virology*, 210, 429-440 (1995).

K. M. Duus et al., "Multiple Regulatory Effects of Varicella-Zoster Virus (VZV) gL on Trafficking Patterns and Fusogenic Properties of VZV gH," *J. Virol.*, 70, 8961-8971 (1996).

J. R. Ecker et al., "Varicella zoster virus DNA exists as two isomers," *Proc. Natl. Acad. Sci. USA*, 79, 156-160 (1982).

Frank et al., "Brainvox: An Interactive, Multimodal Visualization and Analysis System for Neuroanatomical Imaging," *Neuroimage*, 5, 13-30 (1997).

T. Gojobori et al., "Rates of evolution of the retroviral oncogene of Maloney murine sarcoma virus and of its cellular homologues," *Proc. Natl. Acad. Sci. USA*, 82, 4198-4201 (1985).

P. R. Kinchington et al., "Regulated Nuclear Localization of the Varicella-Zoster Virus Major Regulatory Protein, IE62," *J. Infect. Dis.*, 178(Suppl. 1), S16-21 (1998).

(List continued on next page.)

Primary Examiner—Ali R. Salimi
(74) *Attorney, Agent, or Firm*—Mueing, Raasch & Gebhardt, P.A.

(57) **ABSTRACT**

The present invention provides methods directed to detecting antibodies that specifically bind to a varicella zoster polypeptide, detecting the presence of a varicella zoster virus in an animal, diagnosing a disease caused by varicella zoster virus, and detecting a varicella zoster virus having a single nucleotide polymorphism in ORF68. The present invention also provides a vaccine composition, a method for producing a modified attenuated varicella zoster virus, isolated polynucleotides, and isolated polypeptides, and viruses.

3 Claims, 10 Drawing Sheets

OTHER PUBLICATIONS

- P. R. Kinchington et al., "Molecular basis for a geographic variation of varicella-zoster virus recognized by a peptide antibody," *Neurology*, 45 (Suppl 8), S13-14 (1995).
- P. LaRussa et al., "Restriction fragment length polymorphism of polymerase chain reaction products from vaccine and wild-type varicella-zoster virus isolates," *J. Virol.*, 66, 1016-1020 (1992).
- V. Litwin et al., "Receptor Properties of Two Varicella-Zoster Virus Glycoproteins, gpI and gpIV, Homologous to Herpes Simplex Virus gE and gI," *J. Virol.*, 66, 3643-3651 (1992).
- V. Litwin et al., "Cell Surface Expression of the Varicella-Zoster Virus Glycoproteins and Fc Receptor," *Virology*, 178, 263-272 (1990).
- S. Mallory et al., "Mutational Analysis of the Role of Glycoprotein I in Varicella-Zoster Virus Replication and Its Effects on Glycoprotein E Conformation and Trafficking," *J. Virol.*, 71, 8279-8288 (1997).
- D. J. McGeoch et al., "Molecular phylogeny of the alpha-herpesvirinae subfamily and a proposed evolutionary timescale," *J. Mol. Biol.*, 238, 9-22 (1994).
- D. J. McGeoch et al., "Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses," *J. Mol. Biol.*, 257, 443-458 (1995).
- J. F. Moffat et al., "Tropism of Varicella-Zoster Virus for Human CD4⁺ and CD8⁺ T Lymphocytes and Epidermal Cells in SCID-hu Mice," *J. Virol.*, 69, 5236-5242 (1995).
- J. F. Moffat et al., "Attenuation of the Vaccine Oka Strain of Varicella-Zoster Virus and Role of Glycoprotein C in Alpha-herpesvirus Virulence Demonstrated in the SCID-hu Mouse," *J. Virol.*, 72, 965-974 (1998).
- E. A. Montalvo et al., "Assembly and Processing of the Disulfide-Linked Varicella-Zoster Virus Glycoprotein gpII(140)," *J. Virol.*, 61, 2877-2884 (1987).
- National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Genbank X04370. Accession No. X04370 M14891 M16612, "The complete DNA sequence of varicella-zoster virus," [online]. *J. Gen. Virol.*, 67 (Pt 9) 1759-1816 (1986), [retrieved on May 29, 2001]. Retrieved from the Internet: <URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=Nucleotide&list_uids=599.8=GenBank>, 39 pages.
- T. I. Ng et al., "Phosphorylation of Varicella-Zoster Virus Open Reading Frame (ORF) 62 Regulatory Product by Viral ORF 47-Associated Protein Kinase," *J. Virol.*, 68, 1350-1359 (1994).
- N. Nishimura et al., "A Di-Acidic Signal Required for Selective Export from the Endoplasmic Reticulum," *Science*, 277, 556-560 (1997).
- J. K. Olson et al., "Complex formation facilitates endocytosis of the varicella-zoster virus gE:gI Fc receptor," *J. Virol.*, 72, 1542-1551 (1998).
- J. K. Olson et al., "Varicella-Zoster Virus Fc Receptor gE Glycoprotein: Serine/Threonine and Tyrosine Phosphorylation of Monomeric and Dimeric Forms," *J. Virol.*, 71, 110-119 (1997).
- J. K. Olson et al., "Endocytosis and Recycling of Varicella-Zoster Virus Fc Receptor Glycoprotein gE: Internalization Mediated by a YXXL Motif in the Cytoplasmic Tail," *J. Virol.*, 71, 4042-4054 (1992).
- J. A. Padilla et al., "High-resolution immuno-scanning electron microscopy using a non-coating method: study of herpes simplex virus glycoproteins on the surface of virus particles and infected cells," *J. Elect. Microscopy*, 46, 171-180 (1997).
- B. Rentier, "Introduction," *Neurol.*, 45(Suppl 8), S8 (1995).
- J. E. Rodriguez et al., "Entry and egress of varicella virus blocked by same anti-gH monoclonal antibody," *Virology*, 196, 840-844 (1993).
- J. Sambrook et al, *Molecular Cloning: A Laboratory Manual., Second Edition*, Cold Spring Harbor Laboratory Press (1989), (Cover Page, publication page and table of contents).
- R. A. Santos et al., "Varicella-Zoster Virus gE Escape Mutant VZV-MSP Exhibits an Accelerated Cell-to-Cell Spread Phenotype in both Infected Cell Cultures and SCID-hu Mice," *Virology*, 275, 306-317 (Sep., 2000).
- R. A. Santos et al., "Antigenic Variation of Varicella Zoster Virus Fc Receptor gE: Loss of a Major B Cell Epitope in the Ectodomain," *Virology*, 249, 21-31 (1998).
- T. Shioda et al., "Small amino acid changes in V3 hyper-variable region of gp120 can affect the T-cell-line and macrophage tropism of human immunodeficiency virus type 1," *Proc. Natl. Acad. Sci. USA*, 89, 9434-9438 (1992).
- F. I. Smith et al., "Variation in influenza virus genes epidemiological, pathogenic and evolutionary consequences", *The Influenza Viruses*, (R.M. Krug, Ed.), 319-359. Plenum Press, New York (1989).
- P. G. Spear, "Glycoproteins Specified by Herpes Simplex Viruses," *The Herpesviruses*, 3, (B. Roizman, Ed.), 315-356, Plenum Press, New York (1985).
- E. Szomolanyi-Tsuda et al., "T cell-independent antibody-mediated clearance of polyoma virus in T cell-deficient mice," *J. Exp. Med.*, 183, 403-411 (1996).
- M. Takahashi et al., "Development of a Live Attenuated Varicella Vaccine," *Biken J.*, 18, 25-33 (1975).
- T. A. Tatusova, et al. "Blast 2 Sequences, a new tool for comparing protein and nucleotide sequences," *FEMS Microbiol Lett*, 174, 247-250 (1999), and available at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>.
- U.S. National Institutes of Health, "About NIH Image," [online]. Retrieved on May 29, 2001. Retrieved from the Internet: <URL: <http://rsb.info.nih.gov/ni-image/about.html>, 2 pages.
- K. A. Weigle et al., "Common expression of varicella-zoster viral glycoprotein antigens in vitro and in chickenpox and zoster vesicles," *J. Infect. Dis.*, 148, 630-638 (1983).
- T. H. Weller, "Serial propagation in vitro of agents producing inclusion bodies derived from varicella and herpes zoster," *Proc. Soc. Exp. Biol. Med.*, 83, 340-346 (1953).
- S. A. Wharton et al., "Structure, function, and antigenicity of the hemagglutinin of influenza viruses," *The Influenza Viruses*, (R. M. Krug, Ed.), 153-171. Plenum Press, New York (1989).
- M. Yang et al., "Retrograde, Transneuronal Spread of Pseudorabies Virus in Defined Neuronal Circuitry of the Rat Brain is Facilitated by gE Mutations that Reduce Virulence," *J. Virol.*, 73, 4350-4359 (May, 1999).
- Z. Yao et al., "Varicella-Zoster Virus Glycoprotein gpI/gpIV Receptor: Expression, Complex Formation, and Antigenicity within the Vaccinia Virus-T7 RNA polymerase Transfection System," *J. Virol.*, 67, 305-314 (1993).

- Z. Zhu et al., "Targeting of glycoprotein I (gE) of varicella-zoster virus to the trans-Golgi network by an AyRV sequence and an acidic amino acid-rich patch in the cytosolic domain of the molecule," *J. Virol.*, *70*, 6563-6575 (1996).
- Z. Zhu et al., "Envelopment of varicella-zoster virus: targeting of viral glycoproteins to the trans-Golgi network," *J. Virol.*, *69* 7951-7959 (1995).
- F. L. Graham et al., "Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5," *J. Gen. Virol.*, *36*, 59-74 (1977).
- C. Grose et al., "Computer modeling of prototypic and aberrant nucleocapsids of varicella-zoster virus," *Virology*, *214*, 321-329 (1995).
- C. Grose, "Glycoproteins Encoded by Varicella-Zoster Virus: Biosynthesis, Phosphorylation, and Intracellular Trafficking," *Annu.Rev. Microbiol.*, *44*, 59-80 (1990).
- C. Grose, "Pathogenesis of Infection with Varicella Vaccine," *Infect. Dis. Clinics NA*, *10*, 489-505 (1996).
- C. Grose et al., "Varicella-Zoster Virus: Isolation and Propagation in Human Melanoma Cells at 36 and 32° C.," *Infect. Immun.*, *19*, 199-203 (1978).
- C. Grose et al., "Monoclonal Antibodies Against Three Major Glycoproteins of Varicella-Zoster Virus," *Infect. Immun.*, *1*, 381-388 (1983).
- C. Grose, "The Synthesis of Glycoproteins in Human Melanoma Cells Infected with Varicella-Zoster Virus," *Virology*, *101*: 1-9 (1980).
- B. H. Hahn et al., "Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS," *Science*, *232*, 1548-1553 (1986).
- R. Harson et al., "Egress of Varicella-Zoster Virus from the Melanoma Cell: a Tropism for the Melanocyte," *J. Virol.*, *69*, 4994-5010 (1995).
- C. Hatfield et al., "Epitope Mapping and Tagging by Recombination PCR Mutagenesis," *BioTechniques*, *22*, 332-337 (1997).
- J. Holland et al., "Rapid evolution of RNA genomes," *Science*, *215*, 1577-1585 (1982).
- M. Ito et al., "Human Leukocytes Kill Varicella-Zoster Virus-Infected Fibroblasts in the Presence of Murine Monoclonal Antibodies to Virus-Specific Glycoproteins," *J. Virol.*, *54*, 98-103 (1985).
- D. H. Jones et al., "A Rapid Method for Site-Specific Mutagenesis and Directional Subcloning by Using the Polymerase Chain Reaction to Generate Recombinant Circles," *BioTechniques*, *8*, 178-183 (1990).
- F. Jones et al., "Role of cytoplasmic vacuoles in varicella-zoster virus glycoprotein trafficking and virion envelopment," *J. Virol.*, *62*, 2701-2711 (1988).
- P. R. Kinchington et al., "Regulated Nuclear Localization of the Varicella-Zoster Virus Major Regulatory Protein, IE62," *J. Infect. Dis.*, *178*(Suppl. 1), S16-21 (1998).
- P. R. Kinchington et al., "Molecular basis for a geographic variation of varicella-zoster virus recognized by a peptide antibody," *Neurology*, *45* (Suppl 8), S16-21 (1995).
- P. LaRussa et al., "Restriction fragment length polymorphism of polymerase chain reaction products from vaccine and wild-type varicella-zoster virus isolates," *J. Virol.*, *66*, 1016-1020 (1992).
- V. Litwin et al., "Receptor Properties of Two Varicella-Zoster Virus Glycoproteins, gpI and gpIV, Homologous to Herpes Simplex Virus gE and gI," *J. Virol.*, *66*, 3643-3651 (1992).
- V. Litwin et al., "Cell Surface Expression of the Varicella-Zoster Virus Glycoproteins and Fc Receptor," *Virology*, *178*, 263-272 (1990).
- S. Mallory et al., "Mutational Analysis of the Role of Glycoprotein E Confirmation and Trafficking," *J. Virol.*, *71*, 8279-8288 (1997).
- D. J. McGeoch et al., "Molecular phylogeny of the alpha-herpesvirinae subfamily and a proposed evolutionary timescale," *J. Mol. Biol.*, *238*, 9-22 (1994).
- D. J. McGeoch et al., "Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses," *J. Mol. Biol.*, *247*, 443-458 (1995).
- J. F. Moffat et al., "Tropism of Varicella-Zoster Virus for Human CD4⁺ and CD8⁺ T Lymphocytes and Epidermal Cells in SCID-hu Mice," *J. Virol.*, *69*, 5236-5242 (1995).
- J. F. Moffat et al., "Attenuation of the Vaccine Oka Strain of Varicella-Zoster Virus and Role of Glycoprotein C in Alpha-herpesvirus Virulence Demonstrated in the SCID-hu Mouse," *J. Virol.*, *72*, 965-974 (1998).
- E. A. Montalvo et al., "Assembly and Processing of the Disulfide-Linked Varicella-Zoster Virus Glycoprotein gpII(140)," *J. Virol.*, *61*, 2877-2884 (1987).
- National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Genbank X04370. Accession Number X04370 M14891 M16612, "The complete DNA sequence of varicella-zoster virus," [online]. *J. Gen. Virol.*, *67* (Pt 9) 1759-1816 (1986), [retrieved on May 29, 2001]. Retrieved from the Internet: <URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=Nucleotide&list_uids=599.8>=GenBank, 39 pages.
- T. I. Ng et al., "Phosphorylation of Varicella-Zoster Virus Open Reading Frame (ORF) 62 Regulatory Product by Viral ORF 47-Associated Protein Kinase," *J. Virol.*, *68*, 1350-1359 (1994).
- N. Nishimura et al., "A Di-Acidic Signal Required for Selective Export from the Endoplasmic Reticulum," *Science*, *277*, 556-560 (1997).
- J. K. Olson et al., "Complex formation facilitates endocytosis of the varicella-zoster virus gE:gI Fc receptor," *J. Virol.*, *72*, 1542-1551 (1998).
- J. K. Olson et al., "Varicella-Zoster Virus Fc Receptor gE Glycoprotein: Serine/Threonine and Tyrosine Phosphorylation of Monomeric and Dimeric Forms," *J. Virol.*, *71*, 110-119 (1997).
- J. K. Olson et al., "Endocytosis and Recycling of Varicella-Zoster Virus Fc Receptor Glycoprotein gE: Internalization Mediated by a YXXL Motif in the Cytoplasmic Tail," *J. Virol.*, *71*, 4042-4054 (1992).
- J. A. Padilla et al., "High-resolution immuno-scanning electron microscopy using a non-coating method: study of herpes simplex virus glycoproteins on the surface of virus particles and infected cells," *J. Elect. Microscopy*, *46*, 171-180 (1997).
- B. Rentier, "Introduction," *Neurol.*, *45*(Suppl. 8), S8 (1995).
- J. E. Rodriguez et al., "Entry and egress of varicella virus blocked by same anti-gH monoclonal antibody," *Virology*, *196*, 840-844 (1993).
- J. Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press (1989), (Cover Page, publication page and table of contents).

- R. A. Santos et al., "Varicella-Zoster Virus gE Escape Mutant VZV-MSP Exhibits and Accelerated Cell-to-Cell Spread Phenotype in both Infected Cell Cultures and SCID-hu Mice," *Virology*, 275, 306-317 (Sep., 2000).
- R. A. Santos et al., "Antigenic Variation of Varicella Zoster Virus Fc Receptor gE: Loss of a Major B Cell Epitope in the Ectodomain," *Virology*, 249, 21-31 (1998).
- T. Shioda et al., "Small amino acid changes in V3 hyper-variable region of gp120 can affect the T-cell-line and macrophage tropism of human immunodeficiency virus type 1," *Proc. Natl. Acad. Sci. USA*, 89, 9434-9438 (1992).
- F. I. Smith et al., "Variation in influenza virus genes epidemiological, pathogenic, and evolutionary consequences", *The Influenza Viruses*, (R.M. Krug. Ed.), 319-359. Plenum Press, New York (1989).
- P. G. Spear, "Glycoproteins Specified by Herpes Simplex Viruses," *The Herpesviruses*, 3, (B. Roizman, Ed.), 315-356. Plenum Press, New York (1985).
- E. Szomolanyi-Tsuda et al., "T cell-independent antibody-mediated clearance of polyoma virus in T cell-deficient mice," *J. Exp. Med.*, 183, 403-411 (1996).
- M. Takahashi et al., "Development of a Live Attenuated Varicella Vaccine," *Biken J.*, 18, 25-33 (1975).
- T. A. Tatusova, et al., "BLAST 2 SEQUENCES, a new tool for comparing protein and nucleotide sequences," *FEMS Microbial Lett*, 174, 247-250 (1999), and available at <http://www.ncbi.nlm.nih.gov/gorf/b12.html>.
- U.S. National Institutes of Health, "About NIH Image," [online]. Retrieved on May 29, 2001. Retrieved from the Internet:>URL: <http://rsb.info.nih.gov/nih-image/about.html>, 2 pages.
- K. A. Weigle et al., "Common expression of varicella-zoster viral glycoprotein antigens in vitro and in chickenpox and zoster vesicles," *J. Infect. Dis.*, 148, 630-638 (1983).
- T. H. Weller, "Serial propagation in vitro of agents producing inclusion bodies derived from varicella and herpes zoster," *Proc. Soc. Exp. Biol. Med.*, 83340-346 (1953).
- S. A. Wharton et al., "Structure, function, and antigenicity of the hemagglutinin of influenza viruses," *The Influenza Viruses*, (R. M. Krug. Ed.), 153-171. Plenum Press, New York (1989).
- M. Yang et al., "Retrograde, Transneuronal Spread of Pseudorabies Virus in Defined Neuronal Circuitry of the Rat Brain is Facilitated by gE Mutations that Reduce Virulence," *J. Virol.*, 73, 4350-4359 (May, 1999).
- Z. Yao et al., "Varicella-Zoster Virus Glycoprotein gpI/gpIV Receptor: Expression, Complex Formation, and Antigenicity within the Vaccinia Virus-T7 RNA polymerase Transfection System," *J. Virol.*, 67, 305-314 (1993).
- Z. Zhu et al., "Targeting of glycoprotein I (gE) of varicella-zoster virus to the trans-Golgi network by an AYRV sequence and an acidic amino acid-rich patch in the cytosolic domain of the molecule," *J. Virol.*, 70, 6563-6575 (1996).
- Z. Zhu et al., "Envelopment of varicella-zoster virus: targeting of viral glycoproteins to the trans-Golgi network," *J. Virol.*, 69 7951-7959 (1995).

* cited by examiner

Fig. 1A

(SEQ ID NO: 72)

1 MGTV~~N~~KP~~V~~VG VLMGFGIITG TLRITNPVRA SVLRYDDFHT DEDKLD~~T~~NSV

51 YEPYHSDHA ESSWVNRGES SRKAYDHNSP YIWPRNDYDG FLENAHEHHG

101 VYNQGRGIDS GERLMQPTQM SAQEDLGDDT GIHVIPTLNG DDRHKIVNVD

151 QRQYGDVFKG DLNPKPQGQR LIEVSVEENH PFTLRAPIQR IYGVRYTETW

201 SFLPSLTCTG DAAPAIQHIC LKHTTCFQDV VVDVDC~~A~~ENT KEDQLAEISY

251 RFQGGK~~E~~ADQ PWIVVNTSTL FDELELDPPE IEPGV~~L~~KVLR TEKQYLG~~V~~YI

301 WNM~~R~~SGSDGTS TYATFLVTWK GDEKTRNPTP AVT~~P~~QPRGAE FHMWN~~Y~~HSHV

Fig. 1B

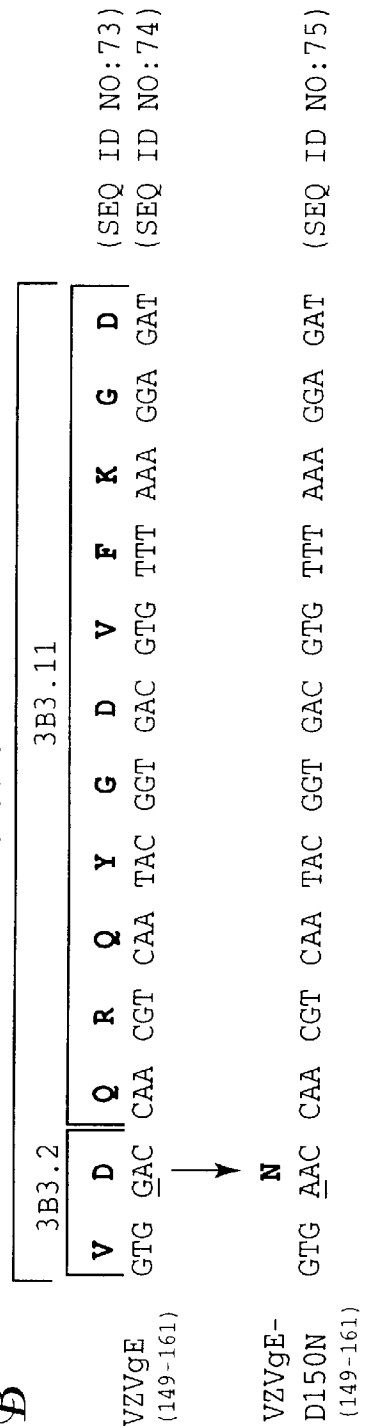


Fig. 2

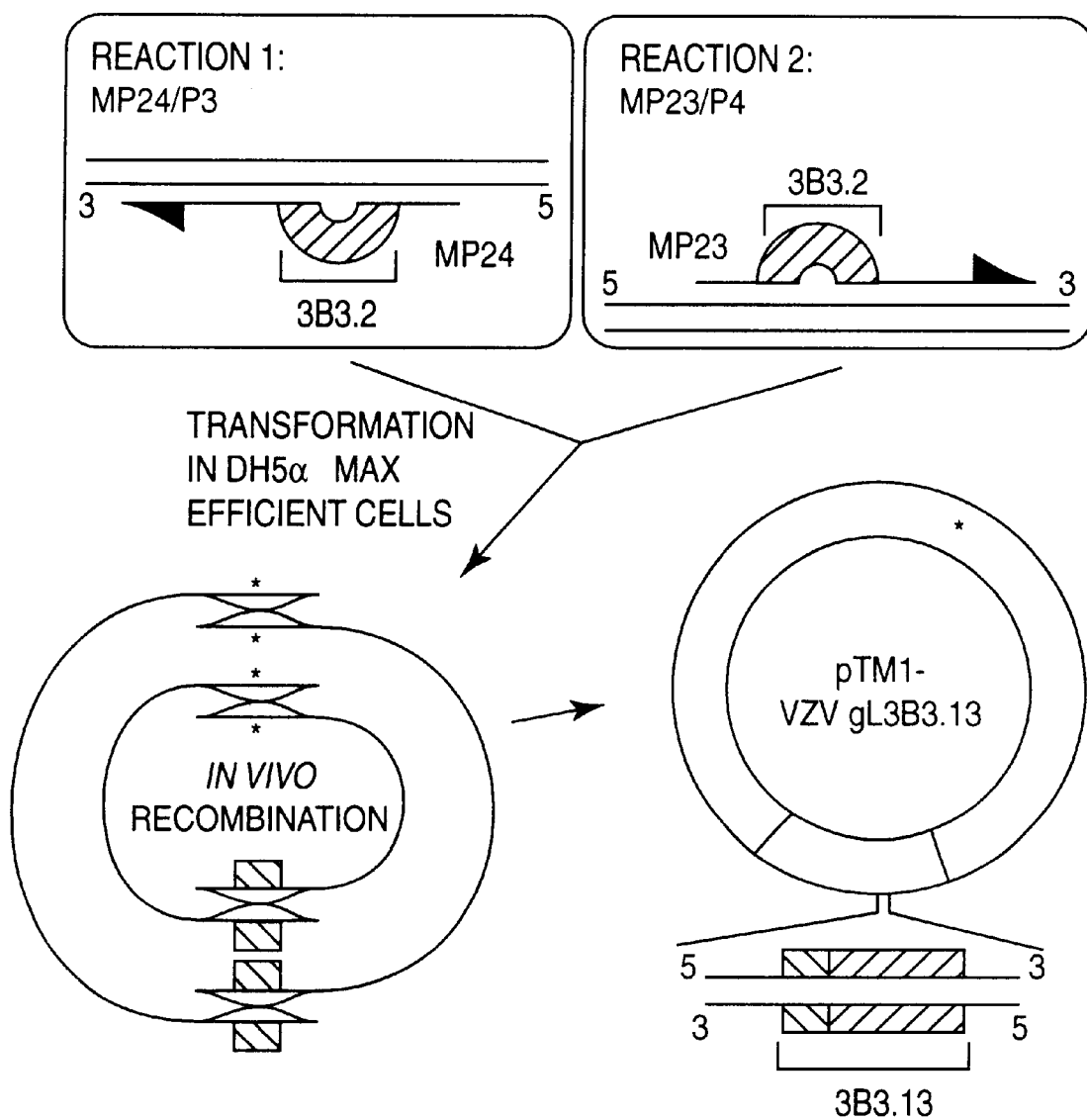


Fig. 3

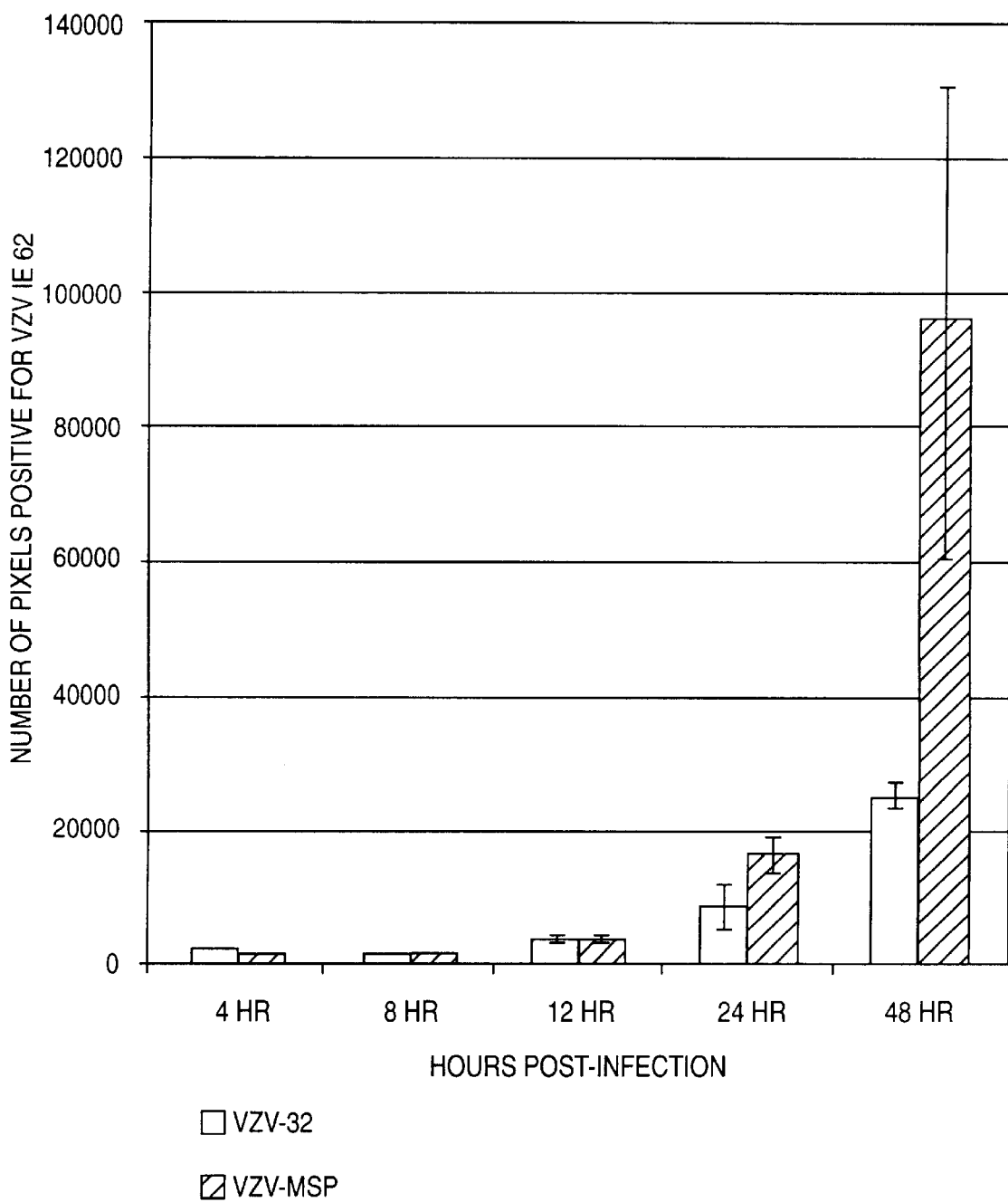


Fig. 4

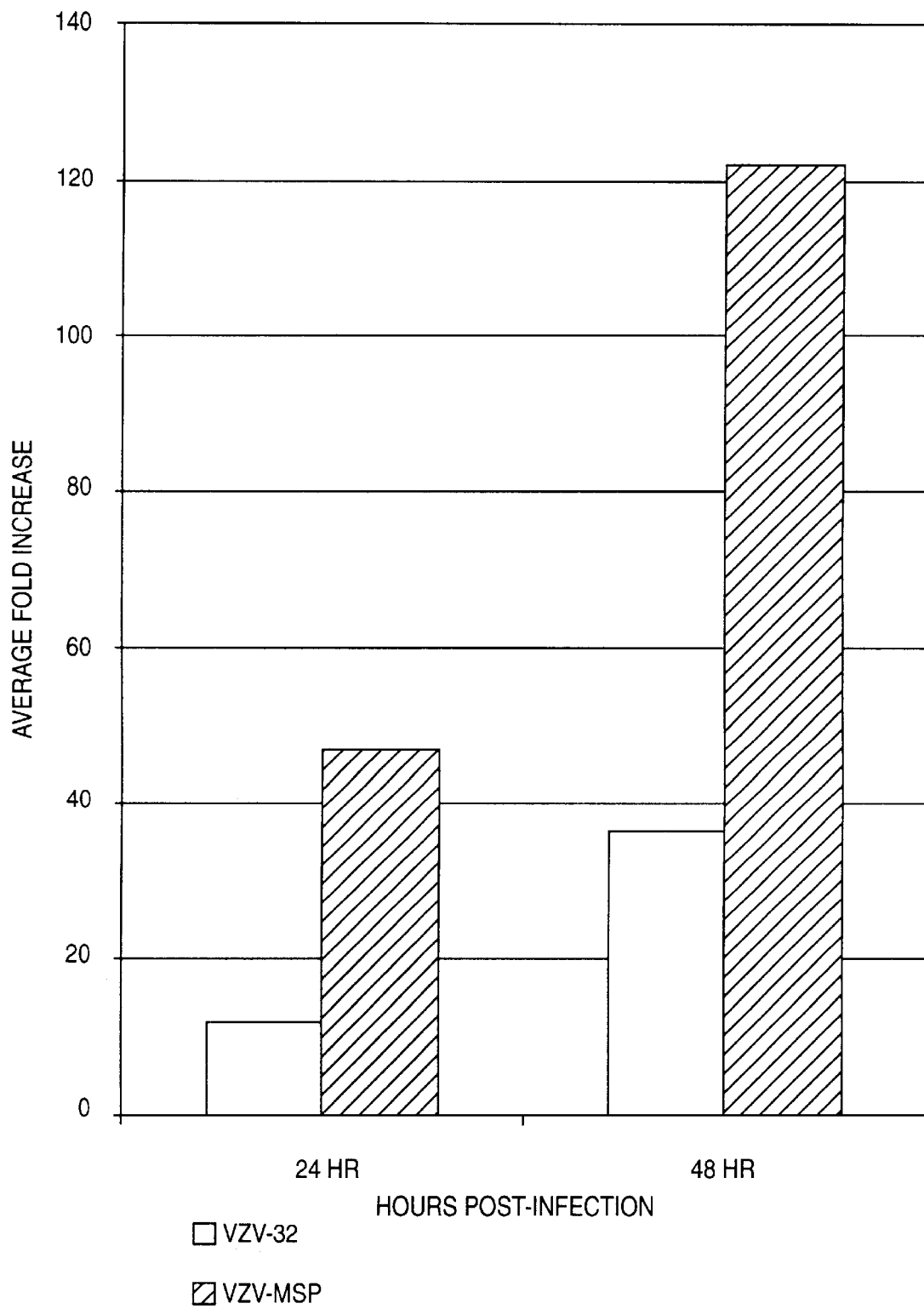
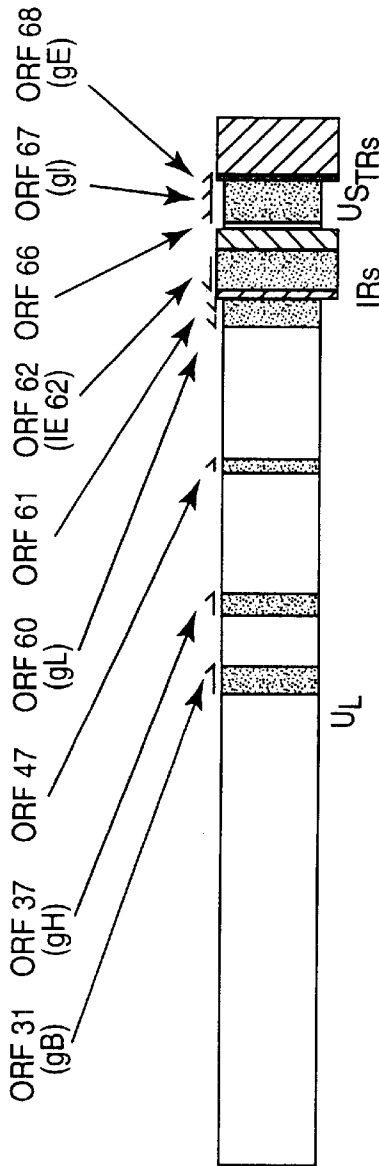


Fig. 5



□ - SEQUENCED REGIONS= 16% OF VIRAL GENOME

NUCLEOTIDES	SIZE OF REGION	ORF(s)	SUBSTITUTIONS
56767-57007	241	5' OF ORF 31 (gB)	NONE
57008-59614	2607	ORF 31 (gB)	NONE
66074-68599	2526	ORF 37 (gH)	c66879t(P269L)
83168-84697	1530	ORF 47	NONE
101170-101649	480	ORF 60 (gL)	NONE
101650-103081	1432	5' OF ORF 60 (gL)	a102203g a102575g c102617i a102969g NONE
103082-104485	1404	ORF 61	a104898g
104486-104936	450	5' OF ORF 61	c109044g (A90A)
105201-109133	3933	ORF 62 (IE62)	NONE
109134-109659	525	5' OF ORF 62	a114140g (S368S)
113037-114218	1182	ORF 66	NONE
114219-114495	277	5' of ORF 67 (gI)	NONE
114496-115560	1065	ORF 67 (gI)	NONE
115561-115807	247	5' OF ORF 68 (gE)	NONE
115808-117679	1872	ORF 68 (gE)	g116255a (D150N)

Fig. 6

Polymorphisms in gE

bp	119	448	660	1606	1808
aa	T>I	D>N	silent	L>I	G>D
Dumas	— CACACC T GAT—	— GTGGACC A AA—	— ATATG T TTA—	— CTTC T ACCA—	— TTTGG T AAC—
MSP	— CACACC T GAT—	— GTGGACC A AA—	— ATATG T TTA—	— CTTC T ACCA—	— TTTGG T AAC—
Ellen	— CACA T CGAT—	— GTGGACC A AA—	— ATATG C TTA—	— CTT A TACCA—	— TTTGG T AAC—
Iceland	— CACA T CGAT—	— GTGGACC A AA—	— ATATG C TTA—	— CTT A TACCA—	— TTTGG T AAC—
80-2	— CACA T CGAT—	— GTGGACC A AA—	— ATATG C TTA—	— CTT A TACCA—	— TTTGG T AAC—
Oka	— CACA T CGAT—	— GTGGACC A AA—	— ATATG T TTA—	— CTTC T ACCA—	— TTTGG T AAC—
VSD	— CACACC T GAT—	— GTGGACC A AA—	— ATATG T TTA—	— CTTC T ACCA—	— TTTG A TAA C —
32	— CACACC T GAT—	— GTGGACC A AA—	— ATATG T TTA—	— CTTC T ACCA—	— TTTGG T AAC—
VIA	— CACACC T GAT—	— GTGGACC A AA—	— ATATG T TTA—	— CTTC T ACCA—	— TTTGG T AAC—

Fig. 7

Polymorphisms in g1

bp	15	546
aa	5 Q>H	silent
Dumas	—ATCCAATGT	—TCTCCGTCT—
MSP	—ATCCAATGT	—TCTCCGTCT—
Ellen	—ATCCAATGT	—TCTCCGTCT—
Iceland	—ATCCAATGT	—TCTCCGTCT—
80-2	—ATCCAATGT	—TCTCCGTCT—
Oka	—ATCCAATGT	—TCTCC A TCT—
VSD	—ATCCAATGT	—TCTCCGTCT—
32	—ATCCA C TGT	—TCTCCGTCT—
VIA	—ATCCA C TGT	—TCTCCGTCT—

Fig. 8

Polymorphisms in gH		215	573	806	1254
bp	39	76	269		
aa	silent	R>K	silent	P>L	silent
Dumas	--- CCTCTTTGG-/-GATAGAAA-/-ATTCTGGAA-/-GGACCACCG-/-AACACTATA-----				
MSP	--- CCTCTTTGG-/-GATAGAAA-/-ATTCTGGAA-/-GGAC T ACCG-/-AACACTATA-----				
Ellen	--- CCTCT G TGG-/-GATA A AAAA-/-ATTCTGGAA-/-GGAC T ACCG-/-AACAC G ATA-----				
Iceland	--- CCTCT G TGG-/-GATA A AAAA-/-ATTCTGGAA-/-GGAC T ACCG-/-AACACTATA-----				
80-2	--- CCTCT G TGG-/-GATA A AAAA-/-ATTCTGGAA-/-GGAC T ACCG-/-AACACTATA-----				
Oka	--- CCTCTTTTGG-/-GATAGAAA-/-ATTCT T GAA-/-GGAC T ACCG-/-AACACTATA-----				
VSD	--- CCTCTTTTGG-/-GATAGAAA-/-ATTCTGGAA-/-GGACCACCG-/-AACACTATA-----				
32	--- CCTCTTTTGG-/-GATAGAAA-/-ATTCTGGAA-/-GGAC T ACCG-/-AACACTATA-----				
VIA	--- CCTCTTTTGG-/-GATAGAAA-/-ATTCTGGAA-/-GGAC T ACCG-/-AACACTATA-----				
bp	208	2099	2181	2445	
aa	silent	700 R>K	silent	silent	
Dumas	--- AAACCTCAA-/-AGCAGGGAT-/-TATTGCGGA-/-CTGGCGGTA---				
MSP	--- AAACCTCAA-/-AGCAGGGAT-/-TATTGCGGA-/-CTGGCGGTA---				
Ellen	--- AAACCCCAA-/-AGCAG A AGAT-/-TATTG T GGA-/-CTGGCGGTA---				
Iceland	--- AAACCCCAA-/-AGCAG A AGAT-/-TATTG T GGA-/-CTGGCGGTA---				
80-2	--- AAACCCCAA-/-AGCAG A AGAT-/-TATTG T GGA-/-CTGGCGGTA---				
Oka	--- AAACCTCAA-/-AGCAGGGAT-/-TATTGCGGA-/-CTGGCGGTA---				
VSD	--- AAACCTCAA-/-AGCAGGGAT-/-TATTGCGGA-/-CTGGCGGTA---				
32	--- AAACCTCAA-/-AGCAGGGAT-/-TATTGCGGA-/-CTGGC A GTA---				
VIA	--- AAACCTCAA-/-AGCAGGGAT-/-TATTGCGGA-/-CTGGC A GTA---				

VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE

CONTINUING APPLICATION DATA

This application claims the benefit of U.S. Provisional Application Ser. No. 60/153,779, filed Sep. 14, 1999, which is incorporated by reference herein.

GOVERNMENT FUNDING

The present invention was made with government support under Grant No. AI 22795, awarded by the National Institutes of Health. The Government has certain rights in this invention.

BACKGROUND

Varicella-zoster virus (VZV) is an ancient virus. Estimations of its origins have established that the modern herpesviruses arose some 60–80 million years ago. VZV is a member of the alphaherpesvirus subfamily of herpesviridae. It is the etiologic agent of chickenpox in childhood, after which the virus enters a latent state in the dorsal root ganglia; decades later, the same virus reactivates and causes the disease shingles (herpes zoster). The entire sequence of the 125 kbp VZV genome has been published (see Davison et al., *J. Gen. Virol.*, 67:1759–1816 (1986)). With the subsequent publication of sequence data from other herpesviruses, the alphaherpesviruses have now been subdivided into two genera called Simplexvirus and Varicellovirus. VZV is considered to have one of the most stable genomes of all herpesviruses. The Oka strain of varicella vaccine derived from a Japanese child with chickenpox has a few minor genomic differences from North American strains, but to date no antigenic variation has been discovered amongst the major surface immunogens of the virion (Arvin et al., *Annu. Rev. Microbiol.*, 50:59–100 (1996)).

Based on their extensive analyses of herpesviral molecular evolutionary history, it has been estimated that herpesvirus DNA sequences mutate 10–100 times faster than the equivalent classes of sequences on the host genome. For glycoprotein gB, a highly conserved open reading frame (ORF) among all herpesviruses, it has been calculated that nonsynonymous substitutions have occurred at a rate of 2.7×10^{-8} substitutions per site per year and synonymous substitutions at 10^{-7} substitutions per site per year. Convincing arguments have been made in favor of the concept of cospeciation; in other words, herpesvirus lineages arise by way of co-evolution with their specific host. In the case of VZV, the progenitor virus most likely arose 60–70 million years before the present.

Of all the human herpesviruses, VZV may undergo the fewest replication cycles during the lifetime of the infected host. Based on a probable schema of pathogenesis, the virus actively replicates for a period of 10–14 days after infection of the human host. During a bout of chickenpox, therefore, VZV has at most 20 replication cycles. Based on the current understanding of VZV latency and reactivation, no further replication occurs unless the individual develops herpes zoster in late adulthood. Because of the above scenario, the genetic stability of the VZV genome has been presumed.

VZV contains the smallest genome of the human herpesviruses, containing about 70 ORFs within the complete VZV-Dumas sequence. Of these ORFs, at least seven code for glycoproteins, of which glycoprotein B (gB), glycoprotein E (gE), glycoprotein H (gH), and glycoprotein I (gI) are present on the exterior of the virion. VZV gE, in

complex with glycoprotein I (gI), acts as a human Fc receptor on the surface of infected cells (Litwin et al., *J. Virol.*, 66:3643–51 (1992), Litwin et al., *Virology*, 178:263–72 (1990)). The cytoplasmic tails of both gE and gI contain endocytosis motifs, allowing internalization and recycling of the complex to and from the cell (Olson et al., *J. Virol.*, 71:110–119 (1997), Olson et al., *J. Virol.*, 71:4042–4054 (1992)). The gE and gI cytoplasmic tails also are modified by both serine/threonine and tyrosine phosphorylation motifs. The fact that gE cannot be deleted suggests that it is essential (Cohen et al., *Proc. Natl. Acad. Sci. USA*, 90:7376–7380 (1993), Mallory et al., *J. Virol.*, 71:8279–88 (1997)).

In VZV infection in humans, VZV gE is the most abundantly produced viral glycoprotein during infection. VZV gE is a major antigenic determinant to which numerous humoral and cytolytic responses are observed (Arvin et al., *J. Immunol.*, 137:1346–1351 (1986); Bergen et al., *Viral Immunol.*, 4:151–166 (1991); and Ito et al., *J. Virol.*, 54:98–103 (1985)). Recently, an immunodominant B-cell epitope was demarcated in the gE ectodomain; the epitope is defined by murine monoclonal antibody (MAb) 3B3 (Duus et al., *J. Virol.*, 70:8961–8971 (1996); Hatfield et al., *Bio-Techniques* 22:332–337 (1997); and Grose, U.S. Pat. No. 5,710,348).

It has long been believed that varicella zoster virus exists in nature as a single serotype (Rentier, *Neurol.*, 45(Suppl. 8), S8 (1995), and that all varicella zoster viruses had essentially the same immunological properties. The first strain of varicella zoster virus that was sequenced was VZV-Dumas. Following the publication of this sequence, it was further believed that all varicella zoster viruses had essentially the same genetic properties as VZV-Dumas.

Significant progress has been made in the diagnosis and vaccination against the sole VZV serotype that is believed to exist and cause disease in the United States. However, the production of the reagents used in diagnosis and vaccination of VZV is time consuming and expensive due to the slow growth rate of the strain grown to produce antigens for diagnostic and vaccine use.

SUMMARY OF THE INVENTION

The present invention represents a significant advance in the art of detecting and preventing varicella zoster virus infection and disease. During the characterization of a varicella zoster virus isolated from a patient, the surprising and unexpected observation was made that the virus had a different serotype. This strain was designated VZV-MSP. The molecular basis of the different serotype was found to be a single nucleotide polymorphism in the genome between VZV-Dumas and VZV-MSP. It was also determined that this single nucleotide polymorphism resulted in the loss of an epitope that is the epitope to which most protective antibody is produced upon vaccination with most currently used vaccines.

Typically, varicella zoster virus isolates can be divided into two groups with respect to growth rate in tissue culture cells. Some isolates, for instance VZV-Oka and VZV-Ellen, grow at a rate that results in complete lysis of a monolayer in about 5 to 7 days. Clinical isolates typically grow at a rate that results in complete lysis of a monolayer in about 4 to 5 days. Further investigation revealed that the new strain, VZV-MSP, unexpectedly and surprisingly had by in vitro tissue culture a growth rate that was significantly higher than previously characterized isolates, and was able to lyse a monolayer in about 2 days.

The present invention provides a method for detecting antibodies that specifically bind to a varicella zoster polypeptide. A biological sample that includes an antibody is contacted with a preparation that includes a varicella zoster polypeptide, for instance an isolated varicella zoster polypeptide or fragment thereof, to form a mixture. The varicella zoster polypeptide includes a polymorphism and can be encoded by a polymorphism of ORF37. The polymorphism in the polypeptide encoded by the polymorphic ORF37 can be due to a single amino acid polymorphism, which can be present in the polypeptide as a leucine at amino acid 269. Alternatively, the varicella zoster polypeptide includes a polymorphism and can be encoded by a polymorphism of ORF68. The polymorphism in the polypeptide encoded by the polymorphic ORF68 can be due to a single amino acid polymorphism, which can be present in the polypeptide as an asparagine at amino acid 150. The mixture is incubated under conditions to allow the antibody to specifically bind the polypeptide to form a polypeptide:antibody complex. The presence or absence of the polypeptide:antibody complex is then detected. Detecting the polypeptide:antibody complex indicates the presence of antibodies that specifically bind to a varicella zoster polypeptide.

The preparation can include whole varicella zoster virus, for instance VZV-MSP or a modified varicella zoster virus, where the modified virus has the ATCC designation VR-795 wherein the nucleotide sequence of the virus has been modified to comprise the polymorphism of ORF37 or ORF68. The biological sample can be blood, vesicle fluid, bone marrow, brain tissue, or combinations thereof. Also provided are kits for detecting antibodies that specifically bind to a varicella zoster polypeptide. This kits include a whole varicella zoster virus.

In another aspect, the present invention provides a method for detecting the presence of a varicella zoster virus in an animal. The method includes detecting the presence of an antibody to a varicella zoster virus polypeptide encoded by a polymorphic ORF of GenBank X04370. The ORF can be ORF37 or ORF68, where the encoded polypeptide includes a single amino acid polymorphism. When the polypeptide is encoded by ORF37, the single amino acid polymorphism present in the polypeptide can be a leucine at amino acid 269. When the polypeptide is encoded by ORF68, the single amino acid polymorphism present in the polypeptide can be an asparagine at amino acid 150. Optionally, the antibody that is detected does not specifically bind to the varicella zoster polypeptide encoded by ORF37 of GenBank Accession X04370 or ORF68 of GenBank Accession X04370.

The present invention is also directed to a method for diagnosing a disease, for instance chicken pox and shingles, caused by varicella zoster virus. The method includes contacting a polynucleotide, optionally an isolated polynucleotide, of a subject suspected of having a disease caused by varicella zoster virus with a primer pair. This is incubated under conditions suitable to form a detectable amplification product, and the primer pair will not form a detectable amplification product when incubated with a polynucleotide having the nucleotide sequence of GenBank Accession X04370. An amplification product is detected, where the detection indicates that the subject has a disease caused by varicella zoster virus. The polynucleotide of the subject can be present in a biological sample, including blood, vesicle fluid, bone marrow, brain tissue, or combinations thereof.

The polynucleotide that is amplified to result in a detectable amplification product can include a single nucleotide

polymorphism relative to the nucleotide sequence of GenBank Accession X04370 (SEQ ID NO:76). The primer pair can include a first primer that includes nucleotides that hybridize with a polynucleotide of GenBank Accession X04370, and a second primer comprising nucleotides that hybridize with a polynucleotide of GenBank Accession X04370, with the proviso that the 3' nucleotide of the second primer hybridizes to the single nucleotide polymorphism relative to the nucleotide sequence of GenBank Accession X04370 and does not hybridize with the corresponding nucleotide present in the nucleotide sequence of GenBank Accession X04370. The single nucleotide polymorphism can be present in ORF37, and the single nucleotide polymorphism can be present at nucleotide 806 of ORF37. The nucleotide at nucleotide 806 can be a thymine. The single nucleotide polymorphism can be present in ORF68, and the single nucleotide polymorphism can be present at nucleotide 448 of ORF68. The nucleotide at nucleotide 448 can be an adenine. An example of a primer pair is CGATGACAGACATAAAAATTGTAAATGTGA (SEQ ID NO:1) and CACCCAAGTATTGTTTTTCTGTCCG (SEQ ID NO:2). nucleotide of the second primer hybridizes to the single nucleotide polymorphism relative to the nucleotide sequence of GenBank Accession X04370 and does not hybridize with the corresponding nucleotide present in the nucleotide sequence of GenBank Accession X04370. The single nucleotide polymorphism can be present in ORF37, and the single nucleotide polymorphism can be present at nucleotide 806 of ORF37. The nucleotide at nucleotide 806 can be a thymine. The single nucleotide polymorphism can be present in ORF68, and the single nucleotide polymorphism can be present at nucleotide 448 of ORF68. The nucleotide at nucleotide 448 can be an adenine. An example of a primer pair is

CGATGACAGACATAAAAATTGTAAATGTGA (SEQ ID NO:1) and
CACCCAAGTATTGTTTTTCTGTCCG (SEQ ID NO:2).

The present invention further provides a method for detecting a varicella zoster virus, for instance VZV-MSP, having a single nucleotide polymorphism in ORF68. The method includes contacting a polynucleotide with a primer pair and incubating under conditions suitable to form a detectable amplification product. The primer pair amplifies a portion of ORF68 of GenBank Accession X04370 and/or a polymorphism thereof, that includes nucleotide 448 of ORF68. The amplification product is exposed to a restriction endonuclease having nucleotide 448 in its recognition sequence. Examples of restriction endonuclease include AflI, AsuI, AvaII, CfrI3I, Eco47I, NspIV, PshAI, Sau96I, and SmaI. The amplification product is then detected. The presence of an amplification product that is not cleaved by the restriction endonuclease indicates the presence of a varicella zoster virus having a single nucleotide polymorphism in ORF68. The polynucleotide can be present in a biological sample, including, for instance, blood, vesicle fluid, bone marrow, brain tissue, or combinations thereof. Optionally, the polynucleotide can be isolated. An example of a primer pair is

GGCATACTACCAATGACACG (SEQ ID NO:12) and
AAGCTCCAAGTCTCGGTGTACC (SEQ ID NO:71).

The present invention is directed to a vaccine composition that includes a modified attenuated varicella zoster virus. The modified attenuated virus has the ATCC designation VR-795, and the nucleotide sequence of the virus has been modified to contain a single nucleotide polymorphism. The

single nucleotide polymorphism can be present in the coding sequence encoding glycoprotein H. For instance, the single nucleotide polymorphism in the virus can be present at nucleotide 806 of the coding sequence encoding glycoprotein H. The nucleotide present at nucleotide 806 can be a thymine. The single nucleotide polymorphism can be present in the coding sequence encoding glycoprotein E. For instance, the single nucleotide polymorphism in the virus is present at nucleotide 448 of the coding sequence encoding glycoprotein E. The nucleotide present at nucleotide 448 can be an adenine.

Also provided by the present invention is a method for producing a modified attenuated varicella zoster virus. The method includes growing the virus in a tissue culture preparation. The virus has the ATCC designation VR-795, and the nucleotide sequence of the virus has been modified to contain a single nucleotide polymorphism. The single nucleotide polymorphism can be present in the coding sequence encoding glycoprotein H. For instance, the single nucleotide polymorphism in the virus can be present at nucleotide 806 of the coding sequence encoding glycoprotein H. The nucleotide present at nucleotide 806 can be a thymine. The single nucleotide polymorphism can be present in the coding sequence encoding glycoprotein E. The single nucleotide polymorphism in the virus can be present at nucleotide 448 of the coding sequence encoding glycoprotein E. The nucleotide present at nucleotide 448 can be an adenine. The modified attenuated virus can have an in vitro growth rate that is greater than the in vitro growth rate of a second varicella zoster virus. The second varicella zoster virus can be, for instance, VZV-32, ATCC VR-586, ATCC VR-1367, or ATCC VR-795. The growth rate of the modified varicella virus can be at least about 4-fold greater than the second varicella zoster virus at 48 hours post infection. Optionally, the modified varicella virus can be isolated.

The present invention further provides isolated polynucleotides, including an isolated polynucleotide having the nucleotide sequence of nucleotides 66,074 to 68,599 of GenBank Accession X04370, with the proviso that nucleotide 66,879 is a thymine; and an isolated polynucleotide having the nucleotide sequence of nucleotides 115,808 to 117,679 of GenBank Accession X04370, with the proviso that nucleotide 116,255 is an adenine. Also provided are the isolated polypeptides encoded by each of the above two polynucleotides. The polynucleotide can be isolated from a varicella zoster virus.

Also provided are viruses having the designation VZV-MSP, VZV-VSD, VZV-VIA, or VZV-Iceland.

Definitions

As used herein, an antibody that can “specifically bind” a polypeptide is an antibody that interacts only with the epitope of the antigen that induced the synthesis of the antibody, or interacts with a structurally related epitope. “Epitope” refers to the site on an antigen to which specific B cells and/or T cells respond so that antibody is produced. As used herein, the term “polypeptide:antibody complex” refers to the complex that results when an antibody specifically binds to a polypeptide.

“Polypeptide” as used herein refers to a polymer of amino acids and does not refer to a specific length of a polymer of amino acids. Thus, for example, the terms peptide, oligopeptide, protein, and enzyme are included within the definition of polypeptide. This term also includes post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like. Coding sequence, coding region, and open reading frame are used interchangeably and refer to a polynucleotide that

encodes a polypeptide, usually via mRNA, when placed under the control of appropriate regulatory sequences. The boundaries of the coding region are generally determined by a translation start codon at its 5' end and a translation stop codon at its 3' end.

An “ORF” followed immediately by a number, for instance ORF37 or ORF68, refers to a specific open reading frame of varicella zoster virus. The approximately 70 individual open reading frames of varicella zoster virus are known to the art, and are described in Davison et al. (*J. Gen. Virol.*, 67:1759–1816 (1986)) and at GenBank Accession X04370. GenBank Accession X04370 is also referred to herein as SEQ ID NO:76. For instance, ORF37 is the open reading frame encoded by nucleotides 66,074 to 68,599 of the nucleotide sequence at GenBank Accession X04370, and ORF68 is the open reading frame encoded by nucleotides 115,808 to 117,679 of the nucleotide sequence at GenBank Accession X04370. A “polymorphic ORF” followed immediately by a number, for instance polymorphic ORF37 or polymorphic ORF68, refers to an open reading frame of varicella zoster virus that has a nucleotide sequence similar to the appropriate nucleotide sequence of GenBank X04370, but includes a single nucleotide polymorphism. Moreover, a polymorphic ORF may contain an insertion or deletion of nucleotides, preferably an insertion of 3 nucleotides or a deletion of 3 nucleotides. When referring to a specific nucleotide of an ORF, the first nucleotide of the start codon is considered to be nucleotide 1, with the following amino acids labeled consecutively. When referring herein to a specific amino acid of a polypeptide encoded by an ORF, the first methionine (prior to any post-translational modification that may occur) is considered to be amino acid 1, with the following amino acids labeled consecutively.

As used herein, the term “polynucleotide” refers to a polymeric form of nucleotides of any length, either ribonucleotides or deoxynucleotides, and includes both double- and single-stranded DNA and RNA. A polynucleotide may include nucleotide sequences having different functions, including for instance coding sequences, and non-coding sequences. A polynucleotide can be obtained directly from a natural source, for instance from a virus, or can be prepared with the aid of recombinant, enzymatic, or chemical techniques. A polynucleotide can be linear or circular in topology. A polynucleotide can be, for example, a portion of a vector, such as an expression or cloning vector, or a fragment.

An “isolated” polypeptide or polynucleotide means a polypeptide or polynucleotide that has been either removed from its natural environment, produced using recombinant techniques, or chemically or enzymatically synthesized. Preferably, a polypeptide or polynucleotide of this invention is purified, i.e., essentially free from any other polypeptide or polynucleotide and associated cellular products or other impurities. An “isolated” varicella zoster virus means a varicella zoster virus has been removed from its natural environment, e.g., the cell that produced the virus.

As used herein, the term “whole varicella zoster virus” refers to a varicella zoster virus particle or virion. The particle can be infective, i.e., be able to reproduce when introduced to an appropriate tissue culture cell under the appropriate conditions, or the particle can be inactive, i.e., incapable of reproducing.

As used herein, a “biological sample” refers to a sample of tissue or fluid isolated from a subject, including but not limited to, for example, blood, plasma, serum, lymph tissue and lymph fluid, cerebrospinal fluid, bone marrow, brain tissue, samples of the skin, external secretions of the skin

including vesicle fluid from a pox, organs, biopsies and also samples of in vitro cell culture constituents including but not limited to conditioned media resulting from the growth of cells and tissues in culture medium, and cell components, or combinations thereof. A "subject" is an animal, including, for instance, a mouse or a human, preferably a human.

As used herein, the term "whole varicella zoster virus particle" refers to an intact varicella zoster virus, for instance a varicella zoster virus that has been produced by a cell and not manipulated to cause the polypeptides that make up the envelop to disassociate from one another.

As used herein, a "primer pair" refers to two single stranded polynucleotides that can be used together to amplify a region of a polynucleotide, preferably by a polymerase chain reaction (PCR). The polynucleotide that results from amplifying a region of a polynucleotide is referred to as an "amplification product." The phrase "under conditions suitable to form a detectable amplification product" refers to the reactions conditions that result in an amplification product. For instance, in the case of a PCR, the conditions suitable to form a detectable amplification product include the appropriate temperatures, ions, and enzyme.

As used herein, the term "hybridize" refers to the ability of two complementary single stranded polynucleotides to base pair with each other, where an adenine of one polynucleotide will base pair to a thymine of a second polynucleotide and a cytosine of one polynucleotide will base pair to a guanine of a second polynucleotide. When the term "hybridize" is used to describe the interaction between a primer and a polynucleotide, hybridization requires that the 3' nucleotide of a primer be able to base pair with the corresponding nucleotide of the polynucleotide that is to be amplified. Typically, the inability of the 3' nucleotide of a primer to base pair with the polynucleotide that is to be amplified results in no amplification (see Newton et al., U.S. Pat. No. 5,595,890).

As used herein, the term "in vitro growth rate" refers to the rate at which a varicella zoster virus spreads from an infected tissue culture cell to an adjacent uninfected tissue culture cell. A tissue culture cell is a cell that replicate in vitro in a nutritive media. The in vitro growth rate of a varicella zoster virus can be measured as described herein.

As used herein, the term "vaccine composition" refers to a pharmaceutical composition containing an antigen, where the composition can be used to prevent or treat a disease or condition in a subject. "Vaccine composition" thus encompasses both subunit vaccines, as described below, as well as compositions containing whole killed, attenuated or inactivated virus. "Subunit vaccine composition" refers to a composition containing at least one immunogenic polypeptide, but not all antigens, derived from a varicella zoster virus. Such a subunit vaccine composition is substantially free of intact virus particles. Thus, a "subunit vaccine composition" is prepared from an isolated, preferably purified, immunogenic polypeptide from the virus. A subunit vaccine composition can comprise the subunit antigen or antigens of interest isolated from other antigens or polypeptides from the pathogen.

As used herein, an "attenuated varicella zoster virus" refers to a varicella zoster virus that is less virulent in humans and preferably, when introduced to a human in the appropriate manner, causes a protective immunological response such that resistance to infection will be enhanced and/or the clinical severity of the disease reduced.

As used herein, a "single nucleotide polymorphism" and a "single amino acid polymorphism" refers to a specific type of polymorphism in a polynucleotide and a polypeptide, respectively, and are described in greater detail herein.

As used herein, the term "recognition sequence" refers to the site on a polynucleotide to which a restriction endonuclease binds prior to cleaving the polynucleotide.

Unless otherwise specified, "a," "an," "the," and "at least one" are used interchangeably and mean one or more than one.

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1. Sequence of the ectodomain of VZV gE. (A) The deduced amino acid sequence of the N-terminal 400 of 623 codons of wild-type VZV gE. The previously defined MAb 3B3 epitope is underlined. The aspartic acid residue altered in VZV-MSP is designated by an arrowhead at codon 150. The silent mutation is indicated by an arrowhead at codon 341. (B) Nucleotides 149–161 and deduced amino acid sequence of the MAb 3B3 epitope in the wild type VZV-32 strain (designated VZV gE) and the mutant VZV-MSP strain gE gene (designated VZV gE-D150N). The altered nucleotides (G to A) and amino acids (D to N) are underlined and marked with the arrow. The two additional codons inserted into the expression plasmid described in FIG. 4 are designated 3B3.2, while the original 3B3 epitope is designated 3B3.11.

FIG. 2. Recombination PCR mutagenesis. Two additional codons (3B3.2) were inserted into the MAb 3B3 epitope to produce plasmid gL 3B3.13 from plasmid gL 3B3.11.

FIG. 3. Quantitative analysis of VZV IE 62 by confocal microscopy. VZV-MSP or VZV-32 infected monolayers were examined by confocal microscopy at increasing times post-infection at $\times 4$ magnification. The total number of pixels positive for VZV IE 62 within each image was quantitated with the Brainvox tal_support programs (University of Iowa) as described herein. The graph summarizes the results from four separate images. Error bars: ± 1 S.D.

FIG. 4. Infectious center assays of VZV-MSP and VZV-32.

FIG. 5. Summary of genetic analysis of VZV-MSP (A) Schematic diagram showing regions of the VZV-MSP genome where the nucleotide sequence has been determined. ORFs 31, 37, 47, 60, 61, 62, 66, 67, and 68 are shown. The polypeptide encoded by the ORF is shown in parentheses under the appropriate ORF. The horizontal arrows above the schematic of the VZV-MSP genome represent the location of the ORFs and the direction of transcription. U_L , unique long; IRs, internal repeat short; U_S , unique short; TRs, terminal repeat short; shaded boxes, regions of the VZV-MSP that have been sequenced; hatched boxes, repeat sequences. (B) Summary of results of sequence analysis of amplified fragments. All mutations discovered are listed by the nucleotide number of the Dumas strain. Any substitutions within open reading frames are followed by the predicted amino acid expressed by VZV-MSP. Nucleotides, nucleotides that were sequenced (the numbering system used is that described in Davison et al., (*J. Gen. Virol.*, 67:1759–1816 (1986))); Size of Region, number of nucleotides sequenced; ORF(s), the ORF or region near an ORF that was sequenced; Substitutions, locations and nature of single nucleotide polymorphism. If the single nucleotide polymorphism encodes a mutation in the resulting polypeptide, the location and nature of the mutation is shown in parentheses. For instance, at position 269 of glycoprotein gH, the proline has been replaced with a leucine.

FIG. 6. Comparative sequence analysis of the VZV gE. VZV ORF 68 was amplified from the viral DNA of eight

VZV strains, including VZV-MSP. Each sequence was compared to the prototype VZV-Dumas genotype. The location of each detected polymorphism is designated by nucleotide number (bp) of the gE gene. Any resulting single amino acid polymorphism that results in gE (e.g., T>I) is noted below the location of the appropriate detected polymorphism. Silent, the single nucleotide polymorphism did not result in a single amino acid polymorphism; asterisk, location of the single nucleotide polymorphism.

FIG. 7. Comparative sequence analysis of VZV gI. VZV ORF 67 was amplified from viral DNA of eight VZV strains. Each sequence was compared to prototype VZV-Dumas genotype. Any resulting single amino acid polymorphism that results in gI (e.g., Q>H) is noted below the location of the appropriate detected polymorphism. Silent, the single nucleotide polymorphism did not result in a single amino acid polymorphism.

FIG. 8. Comparative sequence analysis of VZV gH. VZV ORF 37 was amplified from viral DNA of eight VZV strains. Each DNA sequence was compared to the prototype VZV-Dumas genotype. Nucleotide variations from the VZV-Dumas genotype were tabulated. The P269L mutation originally discovered in VZV-MSP was also present in six other VZV strains, including VZV-32. A total of nine polymorphisms within ORF 37 were discovered among the eight tested strains. Any resulting single amino acid polymorphism that results in gH (e.g., R>K) is noted below the location of the appropriate detected polymorphism. Silent, the single nucleotide polymorphism did not result in a single amino acid polymorphism.

FIG. 9. Comparative sequence analysis of VZV gL. VZV ORF 60 was amplified from eight VZV strains. Each sequence was compared to the VZV Dumas genotype. Only the gL gene of VZV Oka differed from the gL sequence of VZV Dumas.

FIG. 10. Comparative sequence analysis of the VZV IE 62 regulatory gene. VZV ORF 62 was amplified from eight VZV strains. Each sequence was compared to the VZV-Dumas genotype. A total of 38 polymorphisms were detected among the eight VZV strains. VZV-MSP contained only a silent substitution within codon 30 when compared to the VZV-Dumas gene. Any resulting single amino acid polymorphism that results in IE 62 (e.g., S>A) is noted below the location of the appropriate detected polymorphism. Asterisk, location of single nucleotide polymorphism that did not result in a single amino acid polymorphism

DETAILED DESCRIPTION OF THE INVENTION

Polynucleotides

The present invention provides polynucleotides, preferably isolated polynucleotides, having a (at least one) single nucleotide polymorphism. An isolated polynucleotide has a nucleotide sequence that is identical to a nucleotide sequence present in VZV-Dumas, but there is a nucleotide that is polymorphic between the isolated polynucleotide of the present invention and the corresponding polynucleotide present in VZV-Dumas. The nucleotide sequence of VZV-Dumas is present at GenBank Accession X04370. The term "polymorphism" refers to the coexistence of at least two different forms (i.e., at least two different nucleotide sequences or at least two different amino acid sequences) of a polynucleotide or a polypeptide in members of *Varicella zoster*. The polymorphism can be due to a single nucleotide that is different (a single nucleotide polymorphism). In contrast, a polymorphism can be due to several consecutive nucleotides, for instance 2, 3, or 4 consecutive nucleotides,

that are different, which is not within the scope of the present definition of "single nucleotide polymorphism." In the isolated polynucleotides of the present invention, a polymorphism is due to, in increasing order of preference, the presence of 1 single nucleotide polymorphism, at least 1, at least 2, at least 3, most preferably at least 4 single nucleotide polymorphisms in the polynucleotide. Preferably, the isolated polynucleotides of the present invention include no greater than 4 single nucleotide polymorphisms. A single nucleotide polymorphism could be separated from another single nucleotide polymorphism by only one nucleotide.

An example of a polynucleotide of the present invention is a polymorphic ORF37 or the complement thereof, where a polymorphism is at nucleotide 806 of ORF37 (see FIG. 8). The nucleotide at nucleotide 806 can be a guanine, adenine, or thymine, preferably a thymine.

Another example of a polynucleotide of the present invention has the nucleotide sequence of nucleotides 101, 650 to 103,081 (i.e., the region upstream of ORF60) of GenBank Accession X04370 or the complement thereof, but contains a single nucleotide polymorphism at nucleotide 102,203, or a single nucleotide polymorphism at nucleotide 102,575, or a single nucleotide polymorphism at nucleotide 102,617, or a single nucleotide polymorphism at nucleotide 102,969. The nucleotide at nucleotide 102,203 can be a guanine, cytosine, or thymine, preferably a guanine. The nucleotide at nucleotide 102,575 can be a guanine, cytosine, or thymine, preferably a guanine. The nucleotide at nucleotide 102,617 can be a guanine, adenine, or thymine, preferably a thymine. The nucleotide at nucleotide 102,969 can be a guanine, cytosine, or thymine, preferably a guanine.

A further example of a polynucleotide of the present invention has the nucleotide sequence of nucleotides 104, 468 to 104,936 (i.e., the region upstream of ORF6 1) of GenBank Accession X04370 or the complement thereof, but contains a single nucleotide polymorphism at nucleotide 104,898. The nucleotide at nucleotide 104,898 can be a guanine, cytosine, or thymine, preferably a guanine.

Another example of a polynucleotide of the present invention is a polymorphic ORF62 or the complement thereof, where a polymorphism is at nucleotide 90 of ORF62 (see FIG. 10). The nucleotide at nucleotide 90 can be a guanine, adenine, or thymine, preferably a guanine.

Another example of a polynucleotide of the present invention is a polymorphic ORF66 or the complement thereof, where a polymorphism is at nucleotide 1,104 of ORF66. The nucleotide at nucleotide 1,104 can be a guanine, cytosine, or thymine, preferably a guanine.

Another example of a polynucleotide of the present invention is a polymorphic ORF68 or the complement thereof, where a polymorphism is at nucleotide 448 of ORF68 (see FIG. 6). The nucleotide at nucleotide 448 can be an adenine, cytosine, or thymine, preferably an adenine.

Other examples of polynucleotides of the present invention are shown in FIGS. 5-10.

The polynucleotides of the present invention can be obtained by recombinant techniques known to the art including, for instance, cloning from a member of *Varicella zoster* or mutagenizing a polynucleotide so that it has the nucleotide sequence of a polynucleotide of the present invention. Alternatively, a polynucleotide of the present invention can be chemically or enzymatically synthesized by, for instance, an oligonucleotide synthesizer or PCR.

The present invention further includes polynucleotides that are similar to polynucleotides of the present invention as described above, including nucleotides of a polymorphic ORF37 or a polymorphic ORF68, or the complements

thereof. The similarity is referred to as structural similarity and is determined by aligning the residues of the two polynucleotides (i.e., the nucleotide sequence of the candidate polynucleotide and the nucleotide sequence of a preferred polynucleotide of the invention) to optimize the number of identical nucleotides along the lengths of their sequences; gaps in either or both sequences are permitted in making the alignment in order to optimize the number of shared nucleotides, although the nucleotides in each sequence must nonetheless remain in their proper order. Moreover, the nucleotide at the position of the single nucleotide polymorphism (e.g., the thymine at nucleotide 806 in a polymorphic ORF37) is invariant in the candidate polynucleotide. A candidate polynucleotide is the polynucleotide being compared to a preferred polynucleotide of the present invention. Preferably, two nucleotide sequences are compared using the Blastn program, version 2.0.14, of the BLAST 2 search algorithm, as described by Tatusova, et al. (*FEMS Microbiol Lett* 1999, 174:247-250), and available at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. Preferably, the default values for all BLAST 2 search parameters are used, including reward for match=1, penalty for mismatch=-2, open gap penalty=5, extension gap penalty=2, gap x_dropoff=50, expect=10, wordsize=11, and filter on. In the comparison of two nucleotide sequences using the BLAST search algorithm, structural similarity is referred to as "identities." Preferably, a polynucleotide includes a nucleotide sequence having a structural similarity with a preferred polynucleotide of the present invention of at least about 98%, more preferably at least about 99%, most preferably at least about 99.5% identity.

The present invention further includes isolated polynucleotide fragments. A polynucleotide fragment is a portion of an isolated polynucleotide as described herein, where the portion is preferably at least about 15, more preferably at least about 20, most preferably at least about 25 consecutive nucleotides and includes at least one single nucleotide polymorphism. The single nucleotide polymorphism can be at any location in the polynucleotide fragment, and preferably is the nucleotide at one of the 3' ends of the fragment (when the polynucleotide fragment is double stranded) or the nucleotide at the 3' end of the fragment (when the polynucleotide fragment is single stranded).

A polynucleotide of the invention can be inserted in a vector. Construction of vectors containing a polynucleotide of the invention employs standard ligation techniques known in the art. See, for instance, Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989). A vector can provide for further cloning (amplification of the polynucleotide), i.e., a cloning vector, or for expression of the polypeptide encoded by the coding sequence, i.e., an expression vector. The term vector includes, but is not limited to, plasmid vectors, viral vectors, cosmid vectors, or artificial chromosome vectors. Typically, a vector is capable of replication in a bacterial host, for instance *E. coli*. Preferably the vector is a plasmid.

Selection of a vector depends upon a variety of desired characteristics in the resulting construct, such as a selection marker, vector replication rate, and the like. Suitable host cells for cloning or expressing the vectors herein are prokaryote or eukaryotic cells. Preferably the host cell secretes minimal amounts of proteolytic enzymes. Suitable prokaryotes include eubacteria, such as gram-negative or gram-positive organisms. Preferably, *E. coli* is used.

Suitable host cells for the expression of the polypeptides of the invention, preferably encoded by a polymorphic ORF37 or a polymorphic ORF68 as described herein and

containing a single amino acid polymorphism can be derived from multicellular organisms. Such host cells are capable of complex processing and glycosylation activities. Vertebrate or invertebrate culture can be used. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda*, *Aedes aegypti*, *Aedes albopictus*, *Drosophila melanogaster*, *Trichoplusia ni*, and *Bombyx mori* are known to the art.

Vertebrate cells can also be used as hosts. Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (CAS-7, ATCC CRL-1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., *J. Gen. Virol.*, 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO); CHO-K1 (ATCC CCL-61); CHO-D; mouse sertoli cells (TM4); monkey kidney cells (CV1, ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (WI 38, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL 51); TRI cells; MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Suitable plasmids for expression in *E. coli*, for example, include pUC(X), pKK223-3, pKK233-2, pTrc99A, and pET-(X) wherein (X) denotes a vector family in which numerous constructs are available. pUC(X) vectors can be obtained from Pharmacia Biotech (Piscataway, NH) or Sigma Chemical Co. (St. Louis, Mo.). pKK223-3, pKK233-2 and pTrc99A can be obtained from Pharmacia Biotech. pET-(X) vectors can be obtained from Promega (Madison, Wis.). Stratagene (La Jolla, Calif.) and Novagen (Madison, Wis.). To facilitate replication inside a host cell, the vector preferably includes an origin of replication (known as an "ori") or replicon. For example, ColE1 and P15A replicons are commonly used in plasmids that are to be propagated in *E. coli*.

Suitable plasmids for expression in eukaryotic cells, for example, include the EPITAG vectors available from Invitrogen (Carlsbad, Calif.) for mammalian cells. Examples of suitable EPITAG vectors include pcDNA3.1/myc-His and pEF1/myc-His. Other plasmids that can be used in mammalian cells include, for example, pRc/RSV (Invitrogen) and pSecTag2 (Invitrogen). Suitable plasmids for expression in insect cells include, for instance, pIZ/V5-His (Invitrogen), and pBlueBac4.5 (Invitrogen).

An expression vector optionally includes regulatory sequences operably linked to the coding sequence. The invention is not limited by the use of any particular promoter, and a wide variety are known. Promoters act as regulatory signals that bind RNA polymerase in a cell to initiate transcription of a downstream (3' direction) coding sequence. The promoter used in the invention can be a constitutive or an inducible promoter. It can be, but need not be, heterologous with respect to the host cell. Preferred promoters for bacterial transformation include lac, lacUV5, tac, trc, T7, SP6 and ara.

Promoter sequences are known for eukaryotes. Most eukaryotic coding sequences have an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is the CXCAAT region where X may be any nucleotide. At the 3' end of most eukaryotic genes is an AATAAA sequence that may be a signal for addition of the poly A tail to the 3'

end of the coding sequence. All these sequences are suitably inserted into eukaryotic expression vectors.

Transcription of a coding sequence encoding a polypeptide of the present invention in mammalian host cells can be controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, and Hepatitis-B virus.

Transcription of a coding sequence encoding a polypeptide of the present invention by eukaryotes can be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually having about 10 to 300 bp, that act on a promoter to increase its transcription. Enhancers are relatively orientation- and position-independent, having been found 5' and 3' to coding sequences, within an intron as well as within the coding sequence itself. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, alpha-fetoprotein, and insulin). Enhancers from eukaryotic cell viruses are also known and include the SV40 enhancer on the late side of the replication origin, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the coding sequence encoding a polypeptide of the present invention, but is preferably located at a site 5' of the promoter.

An expression vector can optionally include a ribosome binding site (a Shine Dalgarno site for prokaryotic systems or a Kozak site for eukaryotic systems) and a start site (e.g., the codon ATG) to initiate translation of the transcribed message to produce the enzyme. It can also include a termination sequence to end translation. A termination sequence is typically a codon for which there exists no corresponding aminoacyl-tRNA, thus ending polypeptide synthesis. The polynucleotide used to transform the host cell can optionally further include a transcription termination sequence. The *rrnB* terminators, which is a stretch of DNA that contains two terminators, T1 and T2, is an often used terminator that is incorporated into bacterial expression systems. Transcription termination sequences in vectors for eukaryotic cells typically include a polyadenylation signal 3' of the coding sequence.

The polynucleotide used to transform the host cell optionally includes one or more marker sequences, which typically encode a molecule that inactivates or otherwise detects or is detected by a compound in the growth medium. For example, the inclusion of a marker sequence can render the transformed cell resistant to an antibiotic, or it can confer compound-specific metabolism on the transformed cell. Examples of a marker sequence are sequences that confer resistance to kanamycin, ampicillin, chloramphenicol, tetracycline, neomycin, and formulations of phleomycin D1 including, for example, the formulation available under the trade-name ZEOCIN (Invitrogen).

Polypeptides

The present invention is also directed to polypeptides, preferably isolated polypeptides, encoded by polynucleotides of the present invention. A polypeptide has an amino acid sequence that is identical to an amino acid sequence encoded by a coding sequence present in VZV-Dumas, but there is an amino acid that is polymorphic between the polypeptide of the present invention and the corresponding polypeptide encoded by VZV-Dumas. The polymorphism can be due to a single amino acid that is different (a single amino acid polymorphism). In contrast, a polymorphism can

be due to several consecutive amino acids, for instance 2, 3, or 4 consecutive amino acids, that are polymorphic, which is not within the scope of the present definition of "single amino acid polymorphism." In the polypeptides of the present invention, a polymorphism is due to, in increasing order of preference, the presence of 1 single amino acid polymorphism, at least 1, at least 2, at least 3, most preferably at least 4 single amino acid polymorphisms in the polypeptide. Preferably, the isolated polypeptides of the present invention include no greater than 4 single amino acid polymorphisms. A single amino acid polymorphism could be separated from another single amino acid polymorphism by only one amino acid.

Preferably, a polypeptide of the present invention has immunogenic activity. "Immunogenic activity" refers to an amino acid sequence which elicits an immunological response in a subject. An immunological response to a polypeptide is the development in a subject of a cellular and/or antibody-mediated immune response to the polypeptide fragment. Usually, an immunological response includes but is not limited to one or more of the following effects: the production of antibodies, B cells, helper T cells, suppressor T cells, and/or cytotoxic T cells, directed specifically to an epitope or epitopes of the polypeptide fragment.

The polypeptides of the present invention can be obtained from, for instance, a biological sample from a subject infected with a varicella zoster virus that encodes the polypeptide. The polypeptide can be obtained from tissue culture cells that have, for instance, been infected with a varicella zoster virus that encodes the polypeptide or contain a recombinant polynucleotide, preferably a polynucleotide of the invention, that encodes the polypeptide of the invention. Alternatively, the polypeptide can be obtained from a prokaryotic cell, for instance *Escherichia coli*, that contains a recombinant polynucleotide, preferably a polynucleotide of the invention, that encodes the polypeptide of the invention. The polypeptides of the present invention can also be obtained by chemical synthesis.

An example of a polypeptide of the present invention is encoded by a polymorphic ORF37, where a polymorphism is at amino acid 269 (see FIG. 8). When referring herein to a specific amino acid of a polypeptide encoded by an open reading frame, the first methionine is considered to be amino acid 1, with the following amino acids labeled consecutively. The polypeptide encoded by ORF37 is referred to in the art as glycoprotein H, gH, and gpIII. Preferably, the amino acid at position 269 in the polypeptide encoded by a polymorphic ORF37 is an amino acid other than proline, more preferably, the amino acid is a nonpolar (hydrophobic) amino acid, for instance alanine, leucine, isoleucine, valine, phenylalanine, tryptophan, or tyrosine, most preferably, the amino acid is leucine.

An example of a polypeptide of the present invention is encoded by a polymorphic ORF68, where a polymorphism is at amino acid 150 (see FIG. 6). The polypeptide encoded by nucleotides ORF68 is referred to in the art as glycoprotein E, gE, and gpI. Preferably, the amino acid at position 150 in the polypeptide encoded by a polymorphic ORF68 is an amino acid other than aspartic acid, more preferably, the amino acid is asparagine, lysine, histidine, or glutamic acid, most preferably, the amino acid is asparagine.

Other examples of polypeptides of the present invention are shown in FIGS. 5-10.

The present invention further includes polypeptides having similarity with the polypeptides of the present invention as described above, including the polypeptide encoded by a polymorphic ORF37 or a polymorphic ORF68. The simi-

larity is referred to as structural similarity and is generally determined by aligning the residues of the two amino acid sequences (i.e., a candidate amino acid sequence and the amino acid sequence of a preferred polypeptide of the present invention) to optimize the number of identical amino acids along the lengths of their sequences; gaps in either or both sequences are permitted in making the alignment in order to optimize the number of identical amino acids, although the amino acids in each sequence must nonetheless remain in their proper order. Moreover, the amino acid at the position of the single amino acid polymorphism (e.g., the leucine at amino acid 269 in the polypeptide encoded by a polymorphic ORF37) is invariant in the candidate polypeptide. A candidate amino acid sequence is the amino acid sequence being compared to an amino acid sequence present in a preferred polypeptide of the present invention. Preferably, two amino acid sequences are compared using the Blastp program, version 2.0.14, of the BLAST 2 search algorithm, as described by Tatusova et al. (*FEMS Microbiol. Lett.*, 174:247-250 (1999)), and available at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. Preferably, the default values for all BLAST 2 search parameters are used, including matrix=BLOSUM62; open gap penalty=11, extension gap penalty=1, gap x_dropoff=50, expect=10, wordsize=3, and filter on. In the comparison of two amino acid sequences using the BLAST search algorithm, structural similarity is referred to as "identities." Preferably, a polypeptide includes an amino acid sequence having a structural similarity with a preferred polypeptide of the present invention of, in increasing order of preference, at least about 96%, at least about 97%, at least about 98%, and most preferably, at least about 99% identity.

The present invention further includes polypeptide fragments. A polypeptide fragment is a portion of a polypeptide as described herein, where the portion includes at least one single amino acid polymorphism. Preferably, the polypeptide fragment has immunogenic activity. Preferably, a polypeptide fragment is at least about 8, more preferably at least about 12, most preferably at least about 20 amino acids in length.

Viruses

The present invention further provides isolated varicella zoster viruses. Preferably, the genome of an isolated varicella zoster virus of the present invention includes, in increasing order of preference, 1, at least 1, at least 2, at least 3, most preferably at least 4 single nucleotide polymorphisms when compared to the nucleotide sequence of GenBank Accession X04370. Preferably, the genome of an isolated varicella zoster virus of the present invention includes no greater than 4 single nucleotide polymorphisms. Examples of isolated varicella zoster viruses of the present invention include VZV-MSP, VZV-VSD, VZV-VIA, VZV-Iceland. Alternatively, the isolated varicella zoster virus can be a modified varicella zoster virus, preferably a modified attenuated varicella zoster virus. Modified varicella zoster viruses are described in greater detail herein.

A single nucleotide polymorphism can be present in a coding sequence where it can result in the encoded polypeptide containing a single amino acid polymorphism when compared to the polypeptides encoded by the nucleotide sequence of GenBank X04370. Alternatively, a single nucleotide polymorphism can be silent, i.e., not alter the amino acid sequence of a polypeptide encoded by a coding sequence. A single nucleotide polymorphism can be present in a region of the genome that is not a coding sequence. In an isolated varicella zoster virus of the present invention that encodes a polypeptide having a single amino acid

polymorphism, the varicella virus may have a serotype that is different than the serotype known to the art. Preferably, the serotype of an isolated varicella zoster virus of the invention is one that does not contain the epitope to which the monoclonal antibody 3B3 binds. Monoclonal antibody is available from the ATCC (accession number HB-12377). An example of a varicella zoster virus having this serotype is VZV-MSP.

Preferably, the isolated varicella zoster viruses of the present invention have the ability to spread from one cell to another at a rate that is greater than previously characterized varicella zoster viruses. This phenotype, which is also referred to herein as in vitro growth rate and cell-to-cell spread, can be measured by methods that are known to the art, including, for instance, the methods described in Example 2 (i.e., laser scanning confocal microscopy combined with pixel intensity measurement, infectious center assays, and replication in the SCID-hu mouse). Examples of previously characterized varicella zoster viruses that can be used as a baseline for measuring the in vitro growth rate of an isolated varicella virus of the present invention include VZV-32, Oka strain (see Kubo, U.S. Pat. No. 3,985,615), or the varicella zoster viruses having the designations ATCC VR-586, ATCC VR-1367, or ATCC VR-795. Examples of tissue culture cells that can be used include human melanoma cells (including, for instance, MeWo cells), lung fibroblasts (including, for instance, MRC-5 cells, which have the ATCC designation CCL-171), cells derived from human embryos, simian cells, or guinea pig cells.

Preferably, when the infectious center assay is used to measure the in vitro growth rate of a varicella zoster virus, tissue culture cells are added to the well of a 35-mm tissue culture plate and grown until they form a substantially confluent monolayer. The well is inoculated with between about 300 infectious centers to about 700 infectious centers, preferably about 500 infectious centers, i.e., an aliquot of the appropriate varicella zoster virus to result in the initial infection of about 500 cells. The resulting number of infectious centers in the well is measured at 24 hours after inoculation and at 48 hours after inoculation. Preferably, the number of infectious centers of a varicella zoster virus at 48 hours after inoculation is at least about 1.5-fold greater, more preferably at least about 2-fold greater, most preferably at least about 3-fold greater than a previously characterized varicella zoster virus.

Preferably, when laser scanning confocal microscopy combined with pixel intensity measurement is used to measure the in vitro growth rate of a varicella zoster virus, tissue culture cells are inoculated with the varicella zoster virus to be measured. These infected cells are then used to inoculate uninfected cells at a 1:8 ratio of infected to uninfected cells. The spread of the varicella zoster virus is then determined at 24 hours after inoculation at the 1:8 ratio and at 48 hours after inoculation at the 1:8 ratio. The spread of the varicella zoster virus away from a single cell that initially contained the virus can be measured by assaying for evidence of virus in adjacent cells. For instance, the presence of viral nucleic acid or a viral encoded polypeptide can be measured. Preferably, the presence of a viral encoded polypeptide is measured. Preferably, the viral encoded polypeptide is IE62. Preferably, the spread of a varicella zoster virus at 24 hours after inoculation at the 1:8 ratio is at least about 1.5-fold greater, more preferably at least about 2-fold greater than a previously characterized varicella zoster virus. Preferably, the spread of a varicella zoster virus at 48 hours after inoculation at the 1:8 ratio is at least about 2-fold greater, more preferably at least about 4-fold greater than a previously characterized varicella zoster virus.

The present invention is also directed at modifying a varicella zoster virus so that it has an in vitro growth rate that is greater than the in vitro growth rate prior to modification. A varicella zoster virus can be modified by altering the genome of the varicella zoster virus. Preferably, the genome is modified to contain, in increasing order of preference, 1 single nucleotide polymorphism, at least 1, at least 2, at least 3, most preferably, at least 4 single nucleotide polymorphisms. Preferably, the genome is modified to include no greater than 4 single nucleotide polymorphisms. The single nucleotide polymorphisms that could be incorporated into the genome of a varicella zoster virus are described herein. Methods of modifying a genome of a varicella zoster virus are known to the art (see, for instance, Cohen et al., *Proc. Natl. Acad. Sci. USA*, 90:7376-7380 (1993)). Preferably, recombinant DNA techniques are used to make the modification. Preferably, the single nucleotide polymorphisms that could be incorporated into a varicella zoster virus include nucleotide 806 of ORF37, where the single nucleotide polymorphism is a thymine, and/or nucleotide 448 of ORF68, where the single nucleotide polymorphism is an adenine. Examples of varicella zoster viruses that could be modified include a clinical isolate, Oka strain (see Kubo, U.S. Pat. No. 3,985,615), ATCC VR-586, ATCC VR-1367, or ATCC VR-795, preferably ATCC VR-795. It is expected that varicella zoster virus that is presently used to produce, for instance, antigen for diagnostic assays or whole virus for use in vaccine compositions, can be modified by this method. Diagnostic assays and vaccine compositions are described in greater detail herein. The modified virus will grow at a faster rate and result in lowered production costs.

Another aspect of the present invention is directed to methods for producing a varicella zoster virus that has a high in vitro growth rate. Preferably, the varicella zoster virus has an in vitro growth rate that is greater than the in vitro growth rate of a second varicella zoster virus, including, for instance, a clinical isolate, Oka strain (see Kubo, U.S. Pat. No. 3,985,615), VZV-32, ATCC VR-586, ATCC VR-1367, or ATCC VR-795. The method can further include isolation of the varicella virus that has the high in vitro growth rate.

Methods of Use

The present invention provides methods for detecting a varicella zoster virus. These methods are useful in, for instance, detecting a varicella zoster virus in an animal, diagnosing a disease caused by a varicella zoster virus, and detecting a varicella zoster virus having a single nucleotide polymorphism. Preferably, such diagnostic systems are in kit form. Kits are described in greater detail herein. In some aspects of the invention, preferably the varicella zoster virus detected is one having a serotype that is different than VZV-32, or the varicella zoster viruses having the designations ATCC VR-586, ATCC VR-1367, or ATCC VR-795, or having a single nucleotide polymorphism when compared to the nucleotide sequence of GenBank Accession X04370. Preferably, the varicella zoster virus detected is one to which the monoclonal antibody 3B3 does not bind. In some aspects of the invention, detecting a varicella zoster virus includes detecting antibodies that specifically bind to a varicella zoster polypeptide. Whether an antibody specifically binds a polypeptide or non-specifically binds a polypeptide can be determined using methods that are known in the art. Preferably, the polypeptide is gE, gH, gB, or IE62, most preferably gE. The methods include contacting an antibody with a preparation that includes a varicella zoster polypeptide to result in a mixture. Preferably, the antibody is present in a biological sample, more preferably blood, vesicle fluid, bone marrow, or brain tissue.

In this aspect of the invention the varicella zoster virus polypeptide contains a polymorphism. Such polypeptides are described herein. Preferably, the varicella zoster virus polypeptide is encoded by a polymorphic ORF68, and the encoded polypeptide includes an asparagine at amino acid 150. Alternatively and optionally, the varicella zoster virus polypeptide is encoded by a polymorphism of ORF37, and the encoded polypeptide includes a leucine at amino acid 269. The varicella zoster polypeptide in the preparation can be an isolated varicella zoster polypeptide or fragment thereof. Alternatively, preparation can further include whole varicella zoster virus, preferably VZV-MSP, VZV-VSD, VZV-VIA, or VZV-Iceland, more preferably, VZV-MSP.

The method farther includes incubating the mixture under conditions to allow the antibody to specifically bind the polypeptide to form a polypeptide:antibody complex. The preparation that includes the varicella zoster virus may also include reagents, for instance a buffer, that provide conditions appropriate for the formation of the polypeptide:antibody complex. The polypeptide:antibody complex is then detected. The detection of antibodies is known in the art and can include, for instance, immunofluorescence and peroxidase.

The methods for detecting the presence of antibodies that specifically bind to a varicella zoster polypeptide can be used in various formats that have been used to detect antibody to varicella zoster virus, including complement fixation, indirect fluorescent antibody, fluorescent antibody to membrane antigen, neutralization, indirect hemagglutination, immune adherence hemagglutination, radioimmunoassay, latex agglutination, and enzyme-linked immunosorbent assay.

Other methods for detecting a varicella zoster virus include the amplification of a polynucleotide, preferably by PCR. The polynucleotide can be one that is, for instance, isolated from a subject, preferably a subject suspected of having a disease caused by varicella zoster virus. Preferably, the polynucleotide is from a subject, for instance a biological sample, preferably blood, vesicle fluid, bone marrow, or brain tissue. In some aspects of the invention, the method includes contacting a polynucleotide, preferably an isolated polynucleotide, with a primer pair, incubating under conditions suitable to form a detectable amplification product, and detecting the amplification product. Detection indicates that the subject has a disease caused by varicella zoster virus.

The primer pair is one that will not form a detectable amplification product when incubated with a polynucleotide having the nucleotide sequence of GenBank Accession X04370, and preferably will form a detectable amplification product with a polynucleotide containing a single nucleotide polymorphism described herein. Preferably, one of the primers of the primer pair has a nucleotide sequence that hybridizes to a nucleotide sequence of GenBank Accession X04370; however, the 3' nucleotide of the primer corresponds to a single nucleotide polymorphism present in the varicella zoster virus that is to be detected. This method is known to the art as amplification refractory mutation system (ARMS; see Newton et al., U.S. Pat. No. 5,595,890). For instance, a primer pair could be CGATGACAGACAT-AAAATTGTAATGTGA (SEQ ID NO:1), where the underlined nucleotide corresponds to the single nucleotide polymorphism present in VZV-MSP in the coding sequence of nucleotides 115,808 to 117,679 (i.e., the polymorphic ORF68 coding sequence), and CACCCAAGTAT-TGTTTTCTGTCCG (SEQ ID NO:2). Optionally, an additional amplification can be done to detect a varicella zoster virus that does not have the single nucleotide polymorphism

by using, for instance, a primer pair that will form a detectable amplification product when incubated with a polynucleotide having the nucleotide sequence of GenBank Accession X04370. An example of such a primer pair is CGATGACAGACATAAAAATTGTAATGTGG (SEQ ID NO:3), and CACCCAAGTATTGTTTTTCTGTCCG (SEQ ID NO:2). Other primer pairs can be designed using methods known to the art to detect other single nucleotide polymorphisms described herein.

In another aspect of the invention that involves detecting a varicella zoster virus by amplification of a polynucleotide, preferably by PCR, the method is directed to detecting a varicella zoster virus having a single nucleotide polymorphism, preferably at nucleotide 448 of ORF68. Preferably, in the varicella zoster virus to be detected, nucleotide 448 of ORF68 is a cytosine, thymine, or adenine, more preferably an adenine. The method includes contacting a polynucleotide with a primer pair and incubating under conditions suitable to form a detectable amplification product. The amplification product is then exposed to a restriction endonuclease, preferably one that has the recognition sequence that includes nucleotide 448 and is no longer able to cleave when that nucleotide of the recognition sequence is not a guanine. Examples of such restriction endonuclease are AflII, AsuI, AvaII, CfrI 3I, Eco47I, NspIV, PshAI, Sau96I, and SniI. In VZV-Dumas and other varicella zoster viruses, the nucleotide at position 448 of ORF68 in the viral genome is a guanine, and is cleaved by the above-identified restriction endonucleases. When the nucleotide at position 448 of a polymorphic ORF68 is a cytosine, thymine, or adenine, more preferably an adenine, the restriction endonuclease is no longer able to cleave the amplification product. Thus, the method further includes detecting the amplification product after exposure to the restriction endonuclease. The presence of an amplification product that is not cleaved by, for instance, Avall, indicates the presence of a varicella zoster virus having a single nucleotide polymorphism at nucleotide 448.

The primer pair that is used in this aspect of the invention must amplify a region of varicella zoster virus genomic DNA that includes nucleotide 116,255. With out intending to be limiting, an example of a primer pair includes GGCAI-ACTACCAATGACACG (SEQ ID NO:12) and AAGCTC-CAAGTCTCGGTGTACC (SEQ ID NO:71), as well as some of the primers listed in Table 1. Other primers can be designed using methods known in the art.

The methods that involve detecting a varicella zoster virus by amplification of a polynucleotide, preferably by PCR, can also be used to determine the percentage of a population that has a particular single nucleotide polymorphism. Methods of screening populations for the presence of a single nucleotide polymorphism are known to the art. For instance, PCR is sensitive enough to allow samples from a large number of subjects to be pooled and assayed for the presence of a varicella zoster virus having a single nucleotide polymorphism.

The present invention also provides a kit for detecting a varicella zoster virus. The kit includes a varicella zoster polypeptide as described herein (when detecting antibody to varicella zoster virus) or a primer pair as described herein (when amplifying a polynucleotide) in a suitable packaging material in an amount sufficient for at least one assay. Optionally, other reagents such as buffers and solutions needed to practice the invention are also included. Instructions for use of the packaged polypeptide or primer pair are also typically included.

As used herein, the phrase "packaging material" refers to one or more physical structures used to house the contents

of the kit. The packaging material is constructed by well known methods, preferably to provide a sterile, contaminant-free environment. The packaging material has a label which indicates that the polypeptide or primer pair can be used for detecting a varicella zoster virus. In addition, the packaging material contains instructions indicating how the materials within the kit are employed to detect a varicella zoster virus. As used herein, the term "package" refers to a solid matrix or material such as glass, plastic, paper, foil, and the like, capable of holding within fixed limits a polypeptide or a primer pair. Thus, for example, a package can be a glass vial used to contain milligram quantities of a primer pair, or it can be a microtiter plate well to which microgram quantities of a polypeptide have been affixed. "Instructions for use" typically include a tangible expression describing the reagent concentration or at least one assay method parameter, such as the relative amounts of reagent and sample to be admixed, maintenance time periods for reagent/sample admixtures, temperature, buffer conditions, and the like.

The present invention is also directed to vaccines. In one aspect, the present invention is directed to vaccine compositions. Preferably, the subject receiving the vaccine composition will display a protective immunological response such that resistance to infection will be enhanced and/or the clinical severity of the disease reduced. A vaccine composition can include a modified varicella zoster virus, more preferably a modified attenuated varicella zoster virus. A varicella zoster virus can be modified as described above under "Viruses." In other alphaherpesviruses, for instance, pseudorabies virus (PRV), it has been found that spread of the virus in an infected animal is facilitated by gE mutations that reduce virulence (Yang et al., *J. Virol.*, 73:4350 (1999)), it has been found that increasing the in vitro growth rate does not result in an increased virulence of the virus. It is expected that the varicella zoster viruses used as a source of viral antigen for vaccination can be modified to have an increased in vitro growth rate, and not have a increase in virulence. Preferably, the varicella zoster virus that is modified to have an increased in vitro growth rate is Oka strain (see Kubo, U.S. Pat. No. 3,985,615), or ATCC VR-795. The modified varicella zoster virus of the vaccine composition can a live virus, or an inactivated whole virus preparation. The virulence of a varicella zoster virus modified to have a higher in vitro growth rate can be determined using methods known in the art, for instance by using human volunteers.

In another aspect, the vaccine composition can include an isolated varicella zoster virus polypeptide of the present invention or a fragment thereof. Varicella zoster virus polypeptides of the present invention are described herein.

The vaccine composition includes polypeptide or modified varicella zoster viruses having immunogenic activity. Immunogenic carriers can be used to enhance the immunogenicity of the polypeptide or modified varicella zoster viruses. Such carriers include but are not limited to other polypeptides, polysaccharides, liposomes, and bacterial cells and membranes. Polypeptide carriers may be joined to the polypeptides or modified varicella zoster viruses of the present invention to form fusion polypeptides by recombinant or synthetic means or by chemical coupling. Useful carriers and means of coupling such carriers to polypeptide antigens are known in the art.

The vaccine compositions may be formulated by means known in the art. The formulations include those suitable for parental (including subcutaneous, intramuscular, intraperitoneal, and intravenous administration. They are typically prepared as injectables, either as liquid solutions or

suspensions. Solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. The composition may also, for example, be emulsified, or the polypeptide or modified varicella zoster virus encapsulated in liposomes. Where mucosal immunity is desired and the vaccine includes a polypeptide or an inactivated varicella virus, the vaccine compositions may advantageously contain an adjuvant such as the nontoxic cholera toxin B subunit (see, e.g., U.S. Pat. No. 5,462,734). Cholera toxin B subunit is commercially available, for example, from Sigma Chemical Company, St. Louis, Mo. Other suitable adjuvants are available and may be substituted therefor.

The polypeptide or modified varicella zoster virus can be mixed with pharmaceutically acceptable excipients or carriers. Suitable excipients include but are not limited to water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the vaccine compositions may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the vaccine. Such additional formulations and modes of administration as are known in the art may also be used.

The present invention is illustrated by the following examples. It is to be understood that the particular examples, materials, amounts, and procedures are to be interpreted broadly in accordance with the scope and spirit of the invention as set forth herein.

EXAMPLE 1

Identification of Single Nucleotide Polymorphisms in ORF 68 (the gE gene) of VZV-MSP

This example demonstrates the presence of a single nucleotide polymorphism in gE. Because of the important functions of gE, this discovery was completely unexpected. This example provides a more complete characterization of the altered biological properties and genetic composition of this contemporary variant in VZV evolution. For the first time, a VZV variant virus has been discovered which has a cell-to-cell spread phenotype clearly distinguishable from previously characterized VZV strains.

Materials and Methods

Viruses and Cells

The mutant VZV was isolated from a 6 year old boy with leukemia, who contracted chickenpox and was hospitalized for intravenous acyclovir treatment in late 1995. The child's illness responded to treatment and no unusual sequelae were observed. The child's vesicle fluid was inoculated onto MRC-5 cells in glass tubes. The isolate was designated VZV-MSP because the child lived in Minnesota. The VZV-32 laboratory strain was isolated in Texas in 1976 from an otherwise healthy child with chickenpox (Grose, *Virology*, 101:1-9 (1980)). This virus has never been passaged more than 20 times. The VZV Oka strain was isolated from a Japanese child with chickenpox and attenuated by M. Takahashi in Japan in the 1970s (Takahashi et al., *Biken J.*, 18:25-33 (1975)). All viruses were subcultured in either MRC-5 cells or human melanoma cells (MeWo strain, available from C. Grose, University of Iowa, Iowa). MeWo cells are highly permissive for VZV replication but no infectious virus is released into the culture medium (Grose, *Virology*, 101:1-9 (1980)). Therefore, transfer of infectivity is carried out by trypsin dispersion of infected cells and relayering of infected cells onto an uninfected monolayer at a ratio of 1:8 (infected to uninfected cells). The HSV-1 Miyama strain was propagated in the FL line derived from

human amnion cells (Padilla et al., *J. Elect. Microscopy*, 46:171-180 (1997)).

Antibodies and Immunodetection by Confocal Microscopy

MAB 3B3 was produced in this laboratory (Grose et al., *Infect. Immun.*, 40:381-388 (1983)). The antigen for mouse immunization was VZV-32 infected cells. MAB 3B3 attaches to VZV gE even under stringent conditions of buffers containing 1% SDS. Other monoclonal antibodies to VZV gE (MAB 711), VZV gI (MAB 6B5) and VZV gH (MAB 206) were also produced and characterized in this laboratory and are described in Grose (*Annu. Rev. Microbiol.*, 44: 59-80 (1990)). Conditions for immunodetection of VZV proteins by laser scanning confocal microscopy have been outlined by Duus et al., (*J. Virol.*, 70:8961-8971 (1996)). Immunoblotting was performed with the above antibodies as described (Grose, *Annu. Rev. Microbiol.*, 44:59-80 (1990)).

Epitope Mapping by Recombination PCR Mutagenesis

The technique of recombination PCR mutagenesis has been adapted to investigate epitope mapping and tagging (Yao et al., *J. Virol.*, 67:305-314 (1993), Hatfield et al., *BioTechniques*, 22:332-337 (1997)). By this methodology, the epitope of MAB 3B3 was initially defined between amino acids 151-161 in the ectodomain of the 623-amino acid gE glycoprotein. The methodology for producing plasmid pTM1-VZV gL 3B3.11 is described in detail by Hatfield et al., (*BioTechniques*, 22:332-337 (1997)). As part of the investigation in Results (FIG. 2), an additional two codons were inserted at the N-terminus of the 11-amino acid 3B3 epitope, to produce a 13-amino acid epitope tag in the VZV gL protein. The mutating primers included the following: MP23 (sense) CAT ACT GTG TCG ACC AAA GGC AAT ACG GTG ACG TG (SEQ ID NO:4) and MP24 (antisense) TTG CCT TTGOGTC GAC ACA GTA TGC GAT TGT GAT AG (SEQ ID NO:5). PCR amplification was performed under the following parameters: 94° C. denaturation for 30 seconds, 50° C. annealing for 30 seconds, 72° C. extension for 5 minutes; after 25 cycles, there was a final extension at 72° C. for 7 minutes. PCR products were combined and transformed into cells where the overlapping regions underwent recombination to yield a plasmid containing the mutagenized insert. Plasmid purification was performed with a Qiagen Maxi Kit. The newly designated pTM1-VZV gL 3B3.13 plasmid was partially sequenced at the University of Iowa DNA Core Facility to confirm the authenticity of the insertional mutagenesis. This PCR mutagenesis protocol is very reliable with a error frequency of less than 0.025% (Jones et al., *BioTechniques*, 8:178-183 (1990)).

Subcloning of VZV-MSP gE

The subcloning of wild type VZV gE has been described previously (Yao et al., *J. Virol.*, 67:305-314 (1993)). Briefly, two flanking PCR primers were utilized which amplified the VZV-MSP gE ORF directly from VZV-MSP infected melanoma cells. These primers also created a Sac I and a Spe I restriction enzyme site at the 5' and 3' ends, respectively. The primers were the following: Nco gPI (sense) CGA CCC GGG GAG CTC CCA TGG GGA CAG TTA ATA AAC C (SEQ ID NO:6) and IP2 (anti-sense) CGC TCT AGA ACT AGT GGA TCC CCC GGG GAA TTT GTC ACA GGC TTT T (SEQ ID NO:7). PCR amplification was performed using AmpliTaq (Applied Biosystems, Foster City, Calif.) under the following conditions: 94° C. denaturation for 40 seconds, 50° C. annealing for 40 seconds, 72° C. extension for 4 minutes; after 35 cycles, there was a final extension at 72° C. for 7 minutes. After amplification, the PCR fragment was digested with Sac I and Spe I before cloning into the multiple cloning site of the expression vector pTM1.

Imaging of Viral Particles

Clean coverslips were coated with carbon, hydrophilized, irradiated with ultraviolet light for 12 hours, followed by a glow discharge for a few seconds and then sterilization by dry heat (160° C. overnight). MeWo cells were cultivated on the glass coverslips. When the cells became confluent, they were co-cultivated with either VZV-32 or VZV-MSP infected cells at a 1:8 ratio and incubated at 32° C. At the designated times after infection, the cells were prefixed with 1% glutaraldehyde in PBS at 4° C. for 1 hour, rinsed in chilled PBS followed by postfixation with 1% osmium tetroxide in PBS at 4° C. for 1 hour, and then dehydrated in a graded ethanol series. Finally, after two changes in 100% ethanol, the specimens were subjected to a critical point drying method. In the case of HSV-1 samples, FL cells prepared on coverslips were inoculated with HSV-1 Miyama strain at an MOI of 10. Twenty four hours after virus inoculation, coverslips were washed with PBS and processed as described for VZV. Subsequently, the specimens were mounted onto aluminum plates and observed with either a Hitachi S4000 or a Hitachi S-900 SEM. Imaging was performed at the University of Iowa Central Microscopy Research Facility and the University of Wisconsin-Madison Integrated Microscopy Facility.

Results

Analysis of the VZV Isolate by Confocal Microscopy

The virus designated VZV-MSP was initially isolated in human fibroblast monolayers. The cytopathic effect (CPE) was compatible with VZV, but the isolate was poorly reactive with antibodies in a commercial VZV diagnostic kit. Since the isolate did not stain with antibodies to herpes simplex virus (HSV) types 1 and 2 nor did its CPE resemble that of HSV, the virus isolate was further analyzed. When the isolate (VZV passage 1) was received, the infected cell monolayer was trypsin-dispersed and inoculated onto human melanoma cells (VZV passage 2). When CPE was apparent in 5 days, the infected cell monolayer was trypsin-dispersed one more time and inoculated onto 35 mm monolayers for examination by laser scanning confocal microscopy (VZV passage 3). The low passage VZV-32 strain was included in separate dishes as a control virus. When CPE covered about 70% of each monolayer, the infected monolayers were probed with MAb 3B3 against gE and MAb 6B5 against gI and examined by confocal microscopy. In prior studies, it has been shown that these two MAbs do not cross-react with other viral or cellular proteins (Grose, *Annu. Rev. Microbiol.*, 44:59–80 (1990)). As expected from numerous published experiments, MAb 3B3 and MAb 6B5 reacted with the laboratory strain VZV-32. In marked contrast, the anti-gE MAb 3B3 did not attach to cells infected with the VZV-MSP strain, even though MAb 6B5 did bind the infected cells strongly. As an additional control, the anti-gH MAb 206 was added to cultures individually infected with both VZV strains; all VZV-infected cultures were positive in this assay.

The next question addressed was whether the VZV-MSP strain failed to express the entire glycoprotein or whether it has lost an epitope on the glycoprotein. Monoclonal antibodies produced in the epitope-mapping of gE were used. It was previously established that the 3B3 epitope consisted of at least 11 amino acids 151–161 of gE (Duus et al., *J. Virol.*, 70:8961–8971 (1996)). Another monoclonal antibody, MAb 711, attaches to another as yet undefined epitope on the ectodomain of gE. This epitope does not overlap with the 3B3 epitope. Therefore, the above experiment was repeated with MAb 711 as the immunoprobe of VZV-32 and VZV-MSP infected monolayers. Both monolayers stained

positively, a result which indicated that gE was expressed in VZV-MSP infected cells but appeared to have lost either a small segment of its ectodomain or just the 3B3 epitope.

Sequence Analysis of VZV-MSP gE

To further investigate the nature of the gE mutation, we used PCR amplification techniques to first determine whether a full-length gE gene (VZV ORF 68) was present in the mutant strain. The full length gene was amplified. Thereafter, primers were used to amplify overlapping portions of the gE gene, each overlapping portion about 300 bases in size, beginning at the upstream region of ORF 68. Each fragment was subjected to DNA sequencing and each sequence was compared with the published Davison and Scott (Davison et al., *J. Gen. Virol.*, 67:1759–1816 (1986)) sequence of the Dumas strain (FIG. 1A; Genbank Accession number X04370). After analysis of the first 337 codons of VZV-MSP gE ORF, we found the first and most important base change in codon 150 (FIG. 1B, arrow); the substitution involved a replacement of a guanine by an adenine. Of great interest, this point mutation led to a change in amino acid from aspartic acid to asparagine (FIG. 1B). Since this alteration in gE occurred one amino acid away from the deduced 3B3 epitope, which is underlined in FIG. 1A, the sequence data strongly suggested that amino acid 150 was a previously unrecognized contributor to the 3B3 epitope. Further sequencing of VZV-MSP gE revealed one silent mutation in codon 341 of VZV-MSP gE. All other codons were identical to those in the gE sequence of the Dumas strain.

Epitope Mapping of VZV-MSP gE

In an earlier experiment, we had evaluated the 3B3 epitope by inserting the 11-amino-acid sequence into the unrelated VZV ORF 60, namely, the gL glycoprotein (Duus et al., *J. Virol.*, 70:8961–8971 (1996)). The epitope tag within gL was recognized by MAb 3B3 when observed by laser scanning confocal microscopy. To evaluate the contribution of the aspartic acid residue to formation of the epitope, the gL epitope mapping and tagging experiment was repeated in order to insert the aspartic acid residue in its correct location at the N-terminus of the 3B3 epitope. In order to obtain the proper parameters for the mutagenesis primers, one additional codon was inserted along with an aspartic acid (FIG. 1B). The pTM-1 expression plasmids, including gL-3B3.11 and gL-3B3.13, were transfected into HeLa cells and observed by confocal microscopy after labeling with MAb 3B3. Cells transfected with the gL-3B3.11 plasmid were positive in a restricted cytoplasmic pattern, as previously described by Duus et al., (*J. Virol.*, 70:8961–8971 (1996)). Cells transfected with the gL-3B3.13 were not only more intensely stained, the pattern was more widely distributed throughout the cytoplasm. Cells transfected with the pTM-1 gL plasmid alone were negative.

Subcloning the VZV-MSP gE ORF

After completion of the above experiment, we sought to confirm the epitope experiments by amplifying the entire gE gene from VZV-MSP DNA and inserting it into a pTM-1 expression vector. We had previously cloned wild-type gE into the same expression vector; the same primers were used for the second cloning experiment (Yao et al., *J. Virol.*, 67:305–314 (1993)). After transient transfection with these two forms of VZV gE as well as the pTM-1 vector as a control, the cell lysates were solubilized and subjected to electrophoresis followed by transfer to membranes. Additional control samples for the transfection immunoblotting experiments included MeWo cell monolayers infected with three VZV strains: VZV-32, VZV-Oka and VZV-MSP. Unin-

ected MeWo cells served as a negative control. All samples were blotted with MAb 3B3 followed by detection using chemiluminescence. The MAb attached to VZV-32, VZV Oka and VZV gE wild type, but not to VZV-MSP, VZV MSP gE or the vector and uninfected cell controls. When VZV-MSP gE was subsequently immunoblotted with a polyclonal monospecific antibody to gE, the result was positive. Thus, these results confirmed that VZV-MSP gE by itself was expressed but lacked the 3B3 epitope.

Alterations in Topography of Egress of Viral Particles

In previously published studies, it was shown that the egress of wild type VZV particles onto the surface of infected cells occurs in a distinctive pattern which was termed "viral highways" (Harson et al., *J. Virol.*, 69:4994-5010 (1995)). The viral highways are composed of thousands of viral particles which emerge in long rows across the surface of the syncytia. When the distribution of VZV-32 and VZV-MSP particles were compared at a low magnification level by scanning electron microscopy (SEM), wild type virions were again arranged in a pattern consistent with viral highways. Cells infected with the VZV-Oka strain show a similar pattern of viral highways (Grose, *Infect. Dis. Clinics NA*, 10: 489-505 (1996)). In contrast, no such topographical pattern was observed on samples infected with VZV-MSP; instead, viral particles were distributed more uniformly over the cell surface. After observation of numerous monolayers by SEM, it appeared that the number of VZV wild type virions present on the cell surface was less than those of VZV-MSP. The topographical arrangement of VZV-MSP particles also exhibited a high degree of similarity with that of HSV-1, in which thousands of particles covered the cell surface.

In a VZV-infected monolayer, cytopathic effect follows the longitudinal axis of the cells. In VZV-infected human melanoma cell cultures, individual syncytial foci enlarge and eventually merge until the entire monolayer has become a single syncytium. Virions only emerge after syncytia are formed but the virions are never released into the culture medium. If syncytial formation is blocked by adding anti-gH antibody into the culture medium, virions do not egress until the antibody is removed. Based on the imaging studies in this as well as previous reports, it is postulated that virions exit at the leading edge of syncytial foci which are merging. Further, VZV egress (gE/gI mediated) and VZV-induced cell-to-cell fusion (gH/gL mediated) are separate but interdependent events. The mutation on VZV-MSP gE appears to lessen that interdependence.

EXAMPLE 2

Identification of Single Nucleotide Polymorphisms in other ORFs of VZV-MSP and Assessment of Cell-to-cell Spread

This example provides a more complete characterization of the altered biological properties and genetic composition of this contemporary variant in VZV evolution. For the first time, a varicella zoster virus variant is described which has a cell-to-cell spread phenotype clearly distinguishable from previously characterized varicella zoster virus strains.

Cells, viruses and transfer of infectivity. VZV-MSP was isolated in Minnesota in late 1995. VZV-32 was isolated in Texas in the 1970s (Grose et al., *Infect. Immun.*, 19:199-203 (1978)). Reserve stocks of VZV-32 and VZV-MSP were prepared; thus low passages (<20) were used in all experiments in this report. VZV-Oka was isolated in Japan and attenuated in the 1970s (Takahashi et al., *Biken J.*, 18:25-33 (1975)). VZV-Dumas was isolated in Holland and

sequenced in its entirety by Davison et al., (Davison et al., *J. Gen. Virol.*, 67:1759-816 (1986)). All strains were propagated in human melanoma cells (MeWo strain). MeWo cells are highly permissive for VZV replication with no release of infectious virus into the culture medium (Grose, *Virology* 101:1-9 (1980)). Transfer of infectivity was carried out by inoculation of trypsin-dispersed infected cells onto an uninfected monolayer at a 1:8 ratio of infected:uninfected cells unless otherwise noted. Similarly, infectious center assays were carried out by described methods; these assays included both melanoma cell and human neonatal foreskin cell substrates (Cole et al., *J. Virol. Methods*, 36:111-8 (1992), Grose et al., *Infect. Immun.*, 19:199-203 (1978)).

Imaging by confocal microscopy. Replicate 35-mm monolayers of MeWo cells were overlaid with VZV-infected cells at a 1:8 ratio of infected:uninfected cells. At 4, 8, 12, 24, and 48 hours post-infection, the monolayers were fixed and permeabilized with 0.5 ml 2% paraformaldehyde with 0.05% Triton X-100. Cells were probed with an anti-IE 62 mouse monoclonal ascites (MAb 5C6, available from C. Grose, University of Iowa, Iowa) at a dilution of 1:1,000 (Ng et al., *J. Virol.*, 68:1350-1359 (1994)). The secondary antibody was goat anti-mouse IgG F(ab')₂ conjugated to Alexa 488 at a dilution of 1:2,500 (Molecular Probes, Eugene, Oregon). Cell nuclei were stained with TOTO-3 (Molecular Probes), a dimeric cyanine nucleic acid stain, at a dilution of 1:10,000. Samples were examined with a BioRad 1024 laser scanning confocal microscope, as described (Duus et al., *Virology* 210:429-440 (1995)).

Quantitative Analysis of Confocal Images. Confocal images were converted to TIFF format images (Confocal Assistant, v. 4.02, Bio-Rad Laboratories (Hercules, Calif.), available from ftp://ftp.genetics.bio-rad.com/Public/confocal/cas) and transferred to a Silicon Graphics Indy workstation in order to produce the color prints by Showcase software program (Silicon Graphics, Mountain View, Calif.). Confocal images also were analyzed with the Brainvox tal_support programs (Frank et al., *Neuroimage*, 5:13-30 (1997)). Similar analyses can be performed with the public domain NIH Image program, which was developed at the U.S. National Institutes of Health (<http://rsb.info.nih.gov/nih-image/>).

Replication in the SCID-hu mouse. The SCID-hu mouse has been established as an animal model for VZV replication (Moffat et al., *J. Virol.*, 69:5236-42 (1995)). In this model, C.B-17 scid/scid mice were implanted with fetal skin tissue subcutaneously as full thickness dermal grafts. Human fetal tissues were obtained with informed consent according to federal and state regulations and were screened for human immunodeficiency virus. The general care of the experimental animals used for this study was in accordance with the National Institutes of Health guidelines for laboratory animals and in compliance with the Animal Welfare Act (Public Law 94-279) as well as the Stanford University Administrative Panel on Laboratory Animal Care. Animal inoculations were performed according to the previously described protocol (Moffat et al., *J. Virol.*, 69:5236-42 (1995)), viz., an aliquot of infected cell suspension containing 10⁵ infectious centers was injected into each implant. Mock-infected implants were injected with human cells alone. Skin implants were harvested at 7, 14 and 21 days post-inoculation. The implants were fixed in 4% paraformaldehyde, paraffin-embedded, cut into 3- μ m sections, and stained with hematoxylin and eosin. Tissue sections were examined on a Leitz Diaplan light microscope, and digital images were acquired with an Optronics DEI 750 digital camera (Optronics Engineering, Goleta, Calif.). Digital images were formatted as described above.

Isolation of viral DNA. For all viral strains, a 25 cm² monolayer of MeWo cells was infected as described above. After development of 80–100% cytopathology, the infected monolayer was washed thrice with 0.5 ml of 0.01 M phosphate buffered saline (PBS), pH 7.4. Infected cells were then harvested by dislodging into 0.5 ml of PBS. Viral DNA was collected with a DNA easy Kit following the Blood and Body Fluid Protocol (Qiagen Inc, Valencia, Calif.). Following DNA easy protocol, DNA was placed onto a Microcon 50 filter (Millipore, Bedford, Mass.) and washed twice with 0.5 ml of Nanopure water (Bamstead/ThermoLyne, Dubuque, Iowa). Viral DNA was resuspended in 100 μ L of Nanopure water. DNA concentration was assessed visually after 1% agarose gel electrophoresis.

PCR amplification and sequencing of VZV genes. For each ORF, a pair of flanking primers was designed to amplify the gene of interest. PCR amplifications were performed with the Expand High-Fidelity PCR System

(Boehringer Mannheim, Indianapolis, Ind.). This system utilizes both Taq DNA and Pwo DNA polymerases, with the 3'–5' proofreading activity of Pwo DNA polymerase allowing increased fidelity (8.5×10^{-6} per bp error rate) (Boehringer Mannheim). After amplification, the PCR product was sequenced by using the dye terminator cycle sequencing chemistry with AmpliTaq DNA polymerase, FS enzyme (Perkin Elmer Applied Biosystems, Foster City, Calif.). Sequencing reactions were performed on and analyzed with an Applied Biosystems Model 373A stretch fluorescent automated sequencer (Perkin Elmer) at the University of Iowa DNA facility. All genes were PCR amplified twice and each PCR fragment was sequenced at least twice to confirm reported mutations. Each DNA sequence was compared to the prototype VZV-Dumas sequence. The accession number for the complete VZV-Dumas sequence is X04370. The primers used for amplification and/or sequencing the amplified fragments are shown in Table 1.

TABLE 1

VZV sequencing and amplification primers				
Protein ¹	Primer ²	bp ³	Seq/Amp ⁴	Sequence ⁵
gB				
(ORF 31)	Scp 1 (S)	-163 to -136	Seq/Amp	GGCGTTTTTCATAACCTCCGTTACGGGGG (SEQ ID NO:8)
	Scp 2 (AS)	2685 to 2658	Seq/Amp	CCCTGTGATGCGTAATGGAGACACATGA (SEQ ID NO:9)
	Sp 1 (A)	309 to 328	Seq	CTTTGTAATATACCGTCGCC (SEQ ID NO:10)
	Sp 2 (S)	201 to 220	Seq	CGTACGATTAGAACCAACTC (SEQ ID NO:11)
	Sp 3 (S)	569 to 588	Seq	GGCATACTACCAATGACACG (SEQ ID NO:12)
	Sp 4 (S)	929 to 948	Seq	AGTGCGGTGAGGTTGAAGAC (SEQ ID NO:13)
	Sp 5 (S)	1264 to 1283	Seq	CACCCGACTCGAAATACCAG (SEQ ID NO:14)
	Sp 6 (S)	1615 to 1634	Seq	TCTGGTAGTACTACGCGTTG (SEQ ID NO:15)
	Sp 7 (S)	1948 to 1970	Seq	GACTACAGTGAAATCAACGCCG (SEQ ID NO:16)
	Sp 8 (S)	2259 to 2279	Seq	CCCGATGAAGGCATTATATCC (SEQ ID NO:17)
Sp 9 (A)	1418 to 1437	Amp	TAGCTGGCACCACGACGAGG (SEQ ID NO:18)	
Sp 10 (A)	2565 to 2584	Amp	TGCGAACACGGGAGTATCCT (SEQ ID NO:19)	
gH				
(ORF 37)	Sp 2 (S)	104 to 131	Seq	CTGCTCTTCTACGAGAATATTCGACCG (SEQ ID NO:20)
	Sp 3 (A)	223 to 199	Seq	CGTGTTTTCTATCATTTCCCCAGTG (SEQ ID NO:21)
	Sp 4 (S)	448 to 471	Seq	ACTACGTTCCCAACCAACCCCTTG (SEQ ID NO:22)
	Sp 5 (S)	850 to 873	Seq	GCGGTTACAAGCGACACCACATGG (SEQ ID NO:23)
	Sp 6 (S)	1165 to 1191	Seq	CTGTTAGATGAGATCGTAGATGTTTCAG (SEQ ID NO:24)
	Sp 7 (S)	1498 to 1514	Seq	GCTACAGAGAGGCAGGCT (SEQ ID NO:25)
	Sp 8 (S)	1816 to 1840	Seq	TTGCATACCAACTAGACGAATCTG (SEQ ID NO:26)
	Sp 9 (S)	2129 to 2152	Seq	TAGAGACGGTCCGACTGCCCATC (SEQ ID NO:27)
	Sp 10 (S)	-32 to -5	Seq/Amp	CGGTGATATTGTAGCGCAAGTAACAGC (SEQ ID NO:28)
	Sp 11 (A)	2605 to 2580	Seq/Amp	CCCAAAGGTAGTGTATTAITTCGCG (SEQ ID NO:29)
	Sp 13 (A)	223 to 198	Seq	CGTCTCCTTCGTGTGTTG (SEQ ID NO:30)
Sp 14 (A)	1005 to 988	Seq	ATCCAAACTCTCTTCGGG (SEQ ID NO:31)	
Sp 15 (A)	2218 to 2199	Seq	TCGCCCCGTTGGTTAGATAC (SEQ ID NO:32)	
gE				
(ORF 68)	Ip2 (A)	2061 to 2039	Seq/Amp	CGCTCTAGAACTAGTGGATCCCCGGGGAITTTGACAGGCTTTT (SEQ ID NO:7)
	NcogpI (S)	1 to 11	Amp	CGACCCGGGGAGCTCCCATGGGGACAGITAATAAACC (SEQ ID NO:6)
	Sp 1 (S)	1452 to 1470	Seq	GCAITGTTGAAGCCGTAGCA (SEQ ID NO:33)
	Sp 3 (S)	199 to 217	Seq	ATGCGCGGCTCCGATGGTA (SEQ ID NO:34)
	Sp 4 (A)	505 to 486	Seq	GGCCTTGGGGTTTTGGAITA (SEQ ID NO:35)
	Sp 6 (S)	-70 to -48	Seq	GTCCATGGTTTITAGACCTCGGG (SEQ ID NO:36)
	Sp 7 (S)	543 to 561	Seq	GTTTACTTTACGCGACCG (SEQ ID NO:37)
	Sp 8 (S)	823 to 840	Seq	GAATTAGACCCCCCGAG (SEQ ID NO:38)

TABLE 1-continued

VZV sequencing and amplification primers				
Protein ¹	Primer ²	bp ³	Seq/Amp ⁴	Sequence ⁵
	Sp 9 (S)	1127 to 1146	Seq	TAGAGTGGTGTGTAIGTCCCC (SEQ ID NO:39)
	Sp 10 (S)	1601 to 1619	Seq	CACITCTACGATATGCCGC (SEQ ID NO:40)
	Sp 12 (A)	846 to 827	Seq	TTCAATCTCGGGGGGGTCTA (SEQ ID NO:41)
	Sp 13 (A)	1790 to 1772	Seq	TCCGTAGATTCCGAGTCT (SEQ ID NO:42)
gL (ORF 60)	P1 (S)	-1 to 16	Seq/Amp	CAAGCGCCATGGCATCACATAAAT (SEQ ID NO:43)
	P2 (A)	913 to 886	Amp	AAACACTAGTCCATGTGCATGTCCCGC (SEQ ID NO:44)
	Sp 1 (A)	548 to 530	Seq	GCTTGCGGGTTTTTTTGGT (SEQ ID NO:45)
	Sp 2 (A)	320 to 302	Seq	GCCAGCCCCTTTAAGGTGA (SEQ ID NO:46)
	Sp 4 (S)	473 to 494	Seq	GCCAATGAAATGAAACTATCGG (SEQ ID NO:47)
gI (ORF 67)	Scp 1 (S)	-60 to -33	Seq/Amp	CGGCTCACAGAGCTGCTCTTCGGTGTAG (SEQ ID NO:48)
	Scp 2 (A)	1154 to 1137	Seq/Amp	TAATCCTTCCCCTCATATCACAACGCGT (SEQ ID NO:49)
	Sp 1 (A)	995 to 978	Seq	GCGGCCTCCAACATCACA (SEQ ID NO:50)
	Sp 2 (A)	383 to 365	Seq	CCAGCATCCGGCTCTGTG (SEQ ID NO:51)
	Sp 4 (S)	314 to 337	Seq	CGTGTAGGTACAAACATTCGTGGC (SEQ ID NO:52)
ORF 47	ScP 1 (S)	-349 to -320	Amp	ATCTAATCCGTGGGGGTGCGAGTGTACAAG (SEQ ID NO:53)
	ScP 2 (A)	103 to 132	Amp	TCCATGTTCCGGTGTGATGCTCTGTAGGCGTG (SEQ ID NO:54)
	Sp 1 (S)	3 to 22	Seq	GGATGCTGACGACACACCCC (SEQ ID NO:55)
	Sp 2 (A)	163 to 182	Seq	GTTGCAGTTGACGGATTGGC (SEQ ID NO:56)
	Sp 3 (S)	341 to 360	Seq	TTGTACCAATCTCACGCTC (SEQ ID NO:57)
IE62 (ORF 62)	Sp 4 (S)	743 to 762	Seq	ATCGAACGTGCGGCTGACC (SEQ ID NO:58)
	Sp 5 (S)	1135 to 1154	Seq	CTTCCCGGACAACTGCCCAT (SEQ ID NO:59)
	Scp 1 (S)	-79 to -50	Seq/Amp	CACCAACCGCAATCGCAATCCTTTGAAGGC (SEQ ID NO:60)
	Scp 2 (AS)	3990 to 3961	Seq/Amp	TATAACAACAAACAGTCCGCGGCCAGTG (SEQ ID NO:61)
	Sp 1 (S)	318 to 342	Seq	GCCGAGGTCTTCCACACCCGATCT (SEQ ID NO:62)
	Sp 2 (S)	749 to 773	Seq	TTTGAAGGTTAAGGTCCCACTCCCG (SEQ ID NO:63)
	Sp 3 (S)	1142 to 1166	Seq	TACAGGCAGCAGGTCCGGACGCGAA (SEQ ID NO:64)
	Sp 4 (S)	1511 to 1535	Seq	TTACGAGGCCTCAACGGAACCCGTG (SEQ ID NO:65)
	Sp 5 (S)	1925 to 1949	Seq	GATCTCCCGGGTCAACCCTTCTCCA (SEQ ID NO:66)
	Sp 6 (S)	2309 to 2333	Seq	CAAGGCGTACTGTACCCCGGAAACC (SEQ ID NO:67)
	Sp 7 (S)	2685 to 2709	Seq	ACTCATGCCTGGGCGGGAACCTGGA (SEQ ID NO:68)
	Sp 8 (S)	3098 to 3122	Seq	CGTCGCATACACCGTGTGTACCCGC (SEQ ID NO:69)
	Sp 9 (S)	3500 to 3524	Seq	TGCCCTCCCCCGATTCCCAGAGTA (SEQ ID NO:70)

¹The protein that is encoded by the listed ORF.

²The designation of the primer.

³The base pairs to which the primer hybridizes. The numbering used denotes the transcriptional start site as +1.

⁴Seq, the primer was used as a primer in a sequencing reaction; Amp, the primer was used as a primer in PCR. Some primers were used in both sequencing and amplification.

⁵The nucleotide sequence of the primer, listed 5' to 3'.

Results

Cell-to-cell spread phenotype of VZV-MSP. During the initial assessment of VZV-MSP, it was observed that the egress of VZV-MSP particles in cell culture differed from that of other typical VZV laboratory strains, such as VZV-32 and VZV-Oka. Like other VZV strains, however, infectious virus was not released into the culture medium. The fact that greater numbers of VZV-MSP particles were present on the surface suggested that cell-to-cell spread may be increased. Cell-to-cell spread has been assessed in both herpes simplex

virus type 1 (HSV-1) and pseudorabies virus (PRV) through measurement of plaque size in permissive cells or number of cells infected within a typical plaque. Neither technique is easily applicable in the VZV system given the formation of irregularly shaped syncytia by VZV in cell culture and the absence of cell-free virus in tissue culture. Therefore, a technique was developed to assess cell-to-cell spread through confocal microscopic examination.

The spread of low-passage VZV-MSP was compared to the low-passage laboratory strain VZV-32. Human mela-

noma cells were inoculated with either VZV-32 or VZV-MSP at a 1:8 ratio of infected:uninfected cells. At increasing intervals post-infection, the infected monolayers were probed with an antibody against the VZV immediate early protein 62 (IE 62). The number of cells expressing IE 62 and the intracellular localization of the protein were determined. This assay was based on the observation by Kinchington et al., (*J. Infect. Dis.*, 178 Suppl 1:S16–21 (1998)) that VZV IE 62 is present in the nuclei of infected cells during early stages of infection, but then appears in the cytoplasm during later stages.

At the time point of four hours post-infection, comparing VZV-32 and VZV-MSP by confocal microscopy represents the number of infected cells originally overlaid onto each monolayer. Both monolayers contained similar levels of IE 62 positive cells. At twenty-four hours post-infection the difference in extent of spread between VZV-MSP and VZV-32 was apparent. For VZV-32 at 24 hours, 6–8 cells with advanced infection were present in each focus; EE 62 was present in both nucleus and cytoplasm. A few scattered cells adjacent to the infectious focus contained nuclei with IE 62 concentrated near the membrane; these cells represented a recent transfer of infectivity from the central focus. When the VZV-MSP culture was examined at 24 hours, large syncytia had already formed. A typical syncytium contained 20–30 nuclei, an infectious focus 3 to 4 fold larger than that seen with VZV-32. This experiment was repeated four times with equivalent results.

Quantitative analysis of confocal images of VZV IE 62. To quantify the difference in cell-to-cell spread, multiple confocal images were analyzed with the image analysis programs called Brainvox tal-support programs. Each confocal image is made up of 512x512 pixels, for a total of 262,144 pixels. The green fluorescence channel representing the presence of VZV IE 62 within each confocal image was analyzed. The image analysis program initially assigns a relative signal intensity within each pixel of the confocal image. Then, a threshold of signal intensity is calculated to remove background signals. This analysis facilitated quantification of all pixels within each confocal image that contained IE 62. Thus, confocal microscopy of IE 62 coupled with image analysis facilitated a comparison of the extent of viral spread between two different VZV strains.

The results for this analysis are shown in FIG. 3. There was no major difference between the extent of IE 62 spread for VZV-32 and VZV-MSP at 4, 8, and 12 hours post-infection. The results at both 4 and 8 hours, in particular, demonstrated that each monolayer was infected with a similar inoculum of infected cells. Also, the lack of a difference between the three time points (4 hours–12 hours) confirmed that the replication cycle for both VZV strains was greater than 12 hours, in agreement with previous studies. However, at twenty-four hours post-infection, there was a noticeable difference between spread of VZV-MSP and VZV-32. The extent of VZV-MSP spread was at least two-fold greater than VZV-32. At forty-eight hours post-infection, this difference increased further, as VZV-MSP spread was four-fold greater than VZV-32. Thus, image analysis provided a new method by which to measure differences in cell-to-cell spread between VZV strains. Again, this methodology is particularly suited for VZV because cell free virus is not released spontaneously from infected cell cultures; even after sonic disruption of infected monolayers, most viral particles remain attached to remnants of outer cellular membranes.

VZV infectious center assays. Because the results in the previous experiments represented a new application of con-

focal microscopy, a traditional method was also used to confirm the differences in VZV cell-to-cell spread, namely, infectious center assays. For these titrations, the initial virus inocula were replicate samples of VZV-infected cells frozen and stored in liquid nitrogen. After freezing, one aliquot was thawed and titrated from each lot. The inoculum for each 35-mm tissue culture dish was 500 infectious centers. Two dishes were harvested and assayed at each of the following time points: immediately after inoculation (0 hour), 24 hours, and 48 hours post-inoculation (FIG. 4). When comparing the average fold increase of VZV-MSP infectious centers to VZV-32 infectious centers, the spread of VZV-MSP was consistently greater than the spread of VZV-32 over both the first 24 hour period (24 hr pi) and the second 24 hr period (48 hr pi). Otherwise stated, at 48 hours post-infection, the cytopathic effect of VZV-MSP was complete, while numerous infectious center titrations with VZV-32 demonstrated that a 60–72 hr interval was required for similar spread (Grose et al., *Infect. Immun.*, 19:199–203 (1978)). Furthermore, the rapidity of VZV-32 spread was not altered over the initial 20 passages. VZV-Oka and VZV-Allen titrations exhibited a time course similar to that of VZV-32. Therefore, results from both quantitative confocal microscopic image analyses and infectious center assays documented that the spread of VZV-MSP was 3–4 fold greater than the spread of VZV-32.

Growth of VZV-MSP in the SCID-hu mouse. The SCID-hu mouse has provided the first reproducible animal model of VZV pathogenesis. Published studies have documented the pathology of viral infection in human thymus/liver and skin implants after inoculation with parental and vaccine Oka strains as well as low passage wild type virus (Moffat et al., *J. Virol.*, 69:5236–42 (1995), Moffat et al., *J. Virol.*, 72:965–974 (1998)). To assess whether VZV-MSP showed enhanced pathology in the SCID-hu mouse model, skin implants infected with VZV-MSP were harvested at 7, 14, and 21 days after infection. At 7 days post-infection, numerous foci of infection were visible in the epidermis of the human skin implant. By 14 days the foci coalesced into large necrotic lesions with histopathology typical of varicella vesicles. These vesicles were characterized by epidermal hyperplasia, balloon cells, and the separation of the keratin roof from the epidermis. After 21 days, the infection had spread into the dermis and destroyed the entire implant. Samples of the implants also were examined by previously described electron microscopy methods; the virion formation closely resembled that shown in FIG. 2 of the report by Moffat et al., (*J. Virol.*, 72:965–974 (1998)). Mock infected skin implants showed normal skin structure consisting of a thin layer of keratinocytes above the dermis and hair follicles.

Prior published studies had not shown such a rapid progression of pathology in the SCID-hu mouse infected with VZV. To further assess this aspect of VZV infection in the animal model, another experiment was performed with a clinical isolate passaged even fewer times than VZV-32; in addition, the low passage parent VZV-Oka strain was included. Again, the skin samples were collected and examined at days 7, 14, and 21 post-inoculation. When all the specimens were reviewed, the histopathology of the clinical isolate and parental VZV-Oka at 21 days post-inoculation were similar, and for both strains the histopathology was approaching that caused by VZV-MSP at 14 days post-inoculation. Even at 21 days, however, the former two viral strains never caused the total destruction seen after VZV-MSP infection. In short, the progression of VZV-MSP through the skin implant was noticeably more extensive than seen with other viral strains tested in the SCID-hu animal model.

Genetic analysis of other major glycoproteins of VZV-MSP. After documenting the enhanced cell-to-cell spread of VZV-MSP, it was determined whether mutations were present in ORFs other than gE that may be contributing to this phenotype. Specifically, ORFs 31, 37, 60, and 67 coding for VZV gB, gH, gL, and gI, respectively, were analyzed. The ORFs were amplified from the VZV-MSP viral genome and sequenced, then compared to the published VZV-Dumas sequence. Neither gI nor gL gene contained any nucleotide differences when compared to the nucleotide sequences of VZV-Dumas. Further, the gB sequence was identical to that of VZV-Dumas. However, VZV-MSP gH contained a single point mutation within codon 269 (CCA→CTA), converting a proline residue in the predicted VZV-Dumas peptide sequence to a leucine residue in VZV-MSP gH.

Given the presence of mutations within VZV-MSP gE and gH, a similar genetic analyses of VZV-32 was performed. As expected, VZV-32 lacked the D150N mutation within gE. VZV-32 gH, however, revealed the identical point mutation found within codon 269 of VZV-MSP gH. Thus, the mutation within VZV-MSP gH cannot account for the VZV-MSP cell-spread phenotype. VZV-32 contained one additional mutation within ORF 67 (gI) which would lead to a Q5H substitution (CAA→CAT). This substitution was within the probable leader sequence of VZV gI and thus would not be present in mature gI (Davison et al., *J. Gen. Virol.*, 67:1759–816 (1986)). Altogether, within five major glycoprotein ORFs, VZV-MSP contained two point mutations which caused amino acid substitutions when compared to VZV-Dumas: D150N in gE, and P269L in gH (FIG. 5).

In addition to 5 ORFs, we sequenced major portions of the 5' untranslated regions of ORFs 31, 60, 67 and 68. All regions were identical to VZV-Dumas except for that of ORF 60. The latter region contained four polymorphisms; these ranged from 554 to 1320 nucleotides from the ORF 60 initiation codon (FIG. 5). It is very unlikely that these polymorphisms will alter the expression of gL since they are located over 500 nucleotides upstream of the gL start site.

Genetic analysis of VZV-MSP regulatory proteins and kinases. Although viral glycoproteins are the most likely candidates for mediating the cell-to-cell spread phenotype of VZV-MSP, we considered the possibility that an alteration in immediate early (IE) regulatory events may contribute to this enhanced cell-to-cell spread phenotype. VZV expresses one predominant species, IE 62, which acts as the major regulatory protein for viral gene expression. This protein contains a potent acidic activation domain at its N-terminus and is a component of the virus particle. Therefore, the IE 62 gene of VZV-MSP was sequenced, but detected only one silent polymorphism within codon 30 when compared to the Dumas strain (GCG→GCC) (FIG. 5). Thus, the peptide sequence of VZV-MSP IE 62 was identical to the predicted VZV-Dumas sequence. Further, we sequenced the 5' untranslated region containing 525 nucleotides and this region was identical to VZV-Dumas. In addition, we sequenced the adjacent VZV-MSP ORF 61, which encodes the functional homolog of HSV-1 ICP0. Again, the nucleotide sequence was identical to VZV-Dumas.

Previous studies have shown that the viral protein kinase VZV ORF 47 can phosphorylate IE 62 (Ng, et al., 1994). Also, VZV ORF 66 encodes a protein kinase which has been shown to affect the intracellular localization and transactivation function of IE 62. Based upon these results, we wanted to determine whether mutations in these viral kinases could affect the function of IE 62 within VZV-MSP infected cells. Therefore, we sequenced both protein kinase genes within the VZV-MSP genome and found both to be

identical to the prototype VZV-Dumas sequence (FIG. 5). Thus, there was no genetic evidence of polymorphisms within either of two regulatory ORFs or either of two viral protein kinase ORFs. In short, after sequence analysis of over 15% of the VZV-MSP genome, the main impression was a striking similarity with VZV-Dumas except for the notable exceptions mentioned earlier.

EXAMPLE 3

Single Nucleotide Polymorphisms in Major Open Reading Frames of other Varicella Zoster Viruses

Materials and Methods

Viruses. VZV-MSP was isolated in Minnesota in late 1995. VZV-32 was isolated in Texas in the 1970s. VZV-Oka was isolated in Japan and attenuated in the 1970s. VZV-VSD was a wildtype virus collected in South Dakota in the 1980s. VZV-VIA was isolated in Iowa from a child with chickenpox in the 1990s. VZV-Iceland was isolated in Iceland from vesicle fluid of a child with chickenpox in the 1990s. VZV-Ellen was originally isolated in Georgia from a child with chickenpox in the 1960s and obtained from the American Type Culture Collection. VZV 80-2 was originally isolated in Pennsylvania from an adult with herpes zoster in the 1980s.

Propagation of viruses. All viral strains except VZV 80-2 virus were propagated in human melanoma cells (MeWo strain). Transfer of infectivity was performed by inoculation of trypsin dispersed infected cells onto an uninfected monolayer at a 1:8 ratio. Each 25 cm² VZV- infected monolayer was allowed to incubate until cytopathology reached 80%. The monolayer was then washed thrice with 5 ml of 0.01 M phosphate buffered saline (PBS) of pH 7.4. Cells were dislodged by scraping into 0.5 ml PBS. Viral DNA was collected from the cells using the DNeasy Tissue Kit following the DNeasy Protocol for Cultured Animal Cells (Qiagen Inc). Collected DNA was cleaned by placing on a Microcon 50 filter and washing twice with 0.5 ml of Nanopure water. (Bamstead/Thermolyne). Viral DNA was resuspended in 100 μ l of Nanopure water. The VZV 80-2 viral genome DNA was present in two cloned restriction enzyme libraries prepared by Ecker et al., (*Proc. Natl. Acad. Sci. USA*, 79:156–160 (1982)).

PCR amplification and sequencing of viral DNA. PCR amplification was performed with primers flanking the region of interest (Table 1). The Expand High Fidelity PCR System was used in the PCR amplification procedure (Roche Molecular Biochemicals). This system includes Taq DNA and Pwo DNA polymerases, with the 3'–5' proofreading activity of the Pwo DNA polymerase to increase the fidelity (Roche Molecular Biochemicals). Electrophoresis of each sample was carried out in a 1% agarose gel to determine concentration. The DNA sequencing reactions were performed using dye terminator cycle sequencing chemistry with AmpliTaq DNA polymerase, FS enzyme (PE Applied Biosystems, Foster City, Calif.). Reactions were run and subsequently analyzed with an Applied Biosystems Model 373A stretch fluorescent automated sequencer at the University of Iowa DNA Facility. Sequences were further analyzed using the program DNASIS V2.0 (Hitachi Software Engineering Co.). Any region of a VZV genome which differed in sequence from that of the prototypic VZV Dumas was re-amplified in a second PCR step and subjected to a second sequencing analysis.

Results

Polymorphisms in the VZV gE Gene

The VZV gE gene was of greatest interest because of the discovery of the gE mutant strain VZV-MSP. Surprisingly, six gE polymorphisms were found among the eight tested strains and isolates, four of which caused amino acid substitutions (FIG. 6). However, none of the tested strains contained the D150N mutation within the 3B3 epitope of VZV-MSP. VZV-Ellen, VZV-Iceland, and VZV 80-2 had three identical polymorphisms. One was a synonymous mutation within codon 220. Two non-synonymous mutations in these three strains caused amino acid substitutions within codons 40 (T→I) and 536 (L→I). The vaccine strain VZV-Oka also contained the mutation within codon 40, but lacked the other two mutations found within VZV-Ellen, VZV-Iceland, and VZV 80-2. VZV-VSD was the only strain tested which contained a polymorphism within the gE cytoplasmic domain of gE. Interestingly, this change within codon 603 (G→D) inserted an additional acidic amino acid adjacent to the acidic casein kinase II phosphorylation site of gE. VZV-32 and VZV-VIA were the only strains tested that did not contain gE substitutions when compared to the Dumas strain. Since the mutations previously found in VZV-MSP gE were not discovered in any other strain, VZV-MSP gE retained a unique sequence among all currently tested strains and isolates.

Polymorphisms in the VZV gI Gene

The discovery of several polymorphisms in the gE gene of the 8 strains was unexpected. Since VZV gE and gI proteins are commonly found in a complex in the infected cell culture, the gI gene was the next obvious candidate for further genetic analysis. Sequencing of ORF 67 led to the discovery of two changes from the published Dumas sequence (FIG. 7). VZV-32 had an A to C substitution at bp 15 of the ORF that resulted in a glutamine to histidine substitution. VZV Oka also had a silent change of G to A at bp 546. The number of gI sequence variants was less than that seen with gE and may suggest that gI function requires a more rigid amino acid sequence.

Polymorphisms in the VZV gH Gene

Next to the VZV gE:gI complex, the gH:gI complex has been most extensively studied because of its role in cell-to-cell fusion. Overall, the eight strains contained nine polymorphisms, three of which caused amino acid substitutions (FIG. 8). Again, VZV-Ellen, VZV-Iceland, and VZV 80-2 were remarkably similar, with identical changes within codons 76 (R→K) and 700 (R→K) as well as silent substitutions within codons 13, 676 and 727. VZV-Ellen possessed a unique polymorphism within codon 418 that allowed differentiation from VZV-Iceland and VZV 80-2. VZV-32 and VZV-VIA both contained a silent change within codon 815 not present in any other tested strain. VZV-MSP contained only the P269L polymorphism shared by seven strains, including VZV-32.

Polymorphisms in the VZV gL Gene

In a manner similar to gE and gI, gL is invariably linked with the gH protein in VZV-infected cultures. For this reason, the gL ORF was a another candidate gene for sequence analysis. The result was striking: only VZV-Oka gL gene differed from prototype Dumas gL gene (FIG. 9). The first change included the insertion of a methionine codon between amino acids 9 and 10. Secondly, there was a G to A substitution at bp 320 of VZV-Dumas that resulted in a glycine to aspartic acid change in the protein. The fact that the gL genes from the 7 other strains were identical to the prototype gL gene may suggest, as with the gI protein, that gL function requires a protein that is restricted in its genetic variability.

Polymorphisms in the VZV gB Gene

The gB gene is one of the most conserved genes among all herpesviruses. In the case of VZV, this protein is important in the infectious cycle of the virus, based on the previous evidence that addition of anti-gB monoclonal antibody to an infected culture inhibits the progression of infection (Montalvo et al., *J. Virol.*, 61:2877-2884 (1987)). Of the eight strains in which gB was sequenced, 7 were identical to the prototype Dumas gB gene. As with the gL gene, only the Oka gB gene was different from Dumas. The Oka gB had three alterations: an A to C change at bp 217, which led to an amino acid substitution of a threonine to a proline at residue 73; a G to T change at bp 391 resulting in a aspartic acid to a tyrosine substitution at amino acid 131; finally, a silent change of A to C at bp 294. In short, the sequence of gB was highly constrained.

Polymorphisms in the VZV IE62 Gene

The fact that polymorphisms were easily discernible in the major glycoprotein structural genes led to examining whether a similar situation existed in the major VZV regulatory gene called IE62. Overall, 38 polymorphisms were found in the IE 62 gene of the eight tested strains, when compared to the VZV-Dumas IE 62 gene (FIG. 10). As was the case with the gE and gH gene sequences, VZV-Ellen, VZV-Iceland, and VZV 80-2 showed remarkable similarity in their IE 62 genes. Identical silent substitutions within IE 62 were present in these strains at codons 61, 129, and 1071. Also, these three strains contained an identical polymorphism at codon 703 (V→A). VZV-Oka was similar to these three strains, as it shared the silent IE 62 substitutions within codons 61 and 129, but lacked the change at codon 1071. VZV-Oka shared additional polymorphisms with VZV-Ellen at codons 341 and 958 (R→G) that were not present in VZV-Iceland and VZV 80-2. VZV-Ellen revealed nine unique substitutions in addition to those shared with VZV-Iceland, VZV 80-2 and VZV-Oka.

VZV-32 and VZV-VIA both contained identical polymorphisms within IE 62 at codons 516,1057 (Q→R), 1072 (Q→R), 1080, and 1241. VZV-32 also contained five unique changes as well as containing the R958G mutation found in VZV-Ellen and VZV-Oka. VZV-VIA revealed four unique polymorphisms, which combined with the unique changes within VZV-32 allow differentiation of VZV-32 and VZV-VIA. VZV-VSD contained one unique substitution (A→V at codon 602) but otherwise was identical to the Dumas strain. Overall, numerous mutations were found within IE 62 among the tested VZV strains. However, VZV-MSP IE 62 contained no polymorphisms, other than the synonymous substitution found within codon 30, when compared to the Dumas strain. Thus, there was no genetic evidence that IE62 contributed to the VZV-MSP phenotype

Polymorphisms in the VZV Protein Kinase Gene

The VZV genome encodes at least two putative protein kinases, one in the UL region (ORF 47) and the second in the US region of the genome (ORF 66). The ORF 47 kinase is known to phosphorylate the IE62 gene product. Sequencing of the ORF47 protein revealed two variations from Dumas. VZV Ellen, Oka and Iceland shared a silent A to G transition at nucleotide 1449. VZV Ellen had a unique transversion of T to C at nucleotide 913, which caused an amino acid change from serine to proline. The VZV 80-2 gene 47 was not sequenced because neither cloned VZV DNA library contained a complete ORF (Ecker et al., *Proc. Natl. Acad. Sci. USA*, 79:156-160 (1982)).

Transitions Versus Transversions

Mutations that result in the substitution of a pyrimidine for a pyrimidine or a purine for a purine are called transi-

-continued

<211> LENGTH: 35
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 5

 ttgcctttgg tgcacacagt atgcgattgt gatag 35

<210> SEQ ID NO 6
 <211> LENGTH: 37
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 6

 cgacccgggg agctcccatg gggacagtta ataaacc 37

<210> SEQ ID NO 7
 <211> LENGTH: 46
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 7

 cgctctagaa ctagtggatc ccccggggaa tttgtcacag gctttt 46

<210> SEQ ID NO 8
 <211> LENGTH: 28
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 8

 ggcgttttca taacctccgt tacggggg 28

<210> SEQ ID NO 9
 <211> LENGTH: 28
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 9

 ccctgtgatg cgtaatggag acacatga 28

<210> SEQ ID NO 10
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 10

 ctttgtaata taccgtcgcc 20

<210> SEQ ID NO 11
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 11

-continued

cgtacgatta gaaccaactc 20

 <210> SEQ ID NO 12
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 12

 ggcatactac caatgacacg 20

 <210> SEQ ID NO 13
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 13

 agtggcgtga ggttgaagac 20

 <210> SEQ ID NO 14
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 14

 cacccgactc gaaataccag 20

 <210> SEQ ID NO 15
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 15

 tctggtagta ctacgcgttg 20

 <210> SEQ ID NO 16
 <211> LENGTH: 23
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 16

 gactacagtg aaattcaacg ccg 23

 <210> SEQ ID NO 17
 <211> LENGTH: 21
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 17

 cccgatgaag gcattatatt c 21

 <210> SEQ ID NO 18
 <211> LENGTH: 20
 <212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 18
 tagctggcac cacgacgagg 20

<210> SEQ ID NO 19
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 19
 tgcgaacacg ggagtatcct 20

<210> SEQ ID NO 20
 <211> LENGTH: 28
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 20
 ctgctcttct acgagaatat tccgaccg 28

<210> SEQ ID NO 21
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 21
 cgtgttttct atcatttccc cagtg 25

<210> SEQ ID NO 22
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 22
 actacgttcc caccaaacc ccttg 25

<210> SEQ ID NO 23
 <211> LENGTH: 24
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 23
 gcggttaca gcgacaccac atgg 24

<210> SEQ ID NO 24
 <211> LENGTH: 27
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 24
 ctgtagatg agatcgtaga tgttcag 27

-continued

<210> SEQ ID NO 25
 <211> LENGTH: 18
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 25

 gctacagaga ggcaggct 18

<210> SEQ ID NO 26
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 26

 ttgcataccc aactagacga atctg 25

<210> SEQ ID NO 27
 <211> LENGTH: 24
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 27

 tagagacggt cgcactgccc catc 24

<210> SEQ ID NO 28
 <211> LENGTH: 27
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 28

 cggtgatatt gtagcgcaag taacagc 27

<210> SEQ ID NO 29
 <211> LENGTH: 26
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 29

 cccaaaggta gtgtgtatta ttcgcg 26

<210> SEQ ID NO 30
 <211> LENGTH: 18
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 30

 cgtctccttc gtgtgttg 18

<210> SEQ ID NO 31
 <211> LENGTH: 18
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:

-continued

<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 31

atccaaactc tcttcggg 18

<210> SEQ ID NO 32
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 32

tcgccccgt ggtagatac 20

<210> SEQ ID NO 33
 <211> LENGTH: 19
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 33

gcattgtgaa gccgtagca 19

<210> SEQ ID NO 34
 <211> LENGTH: 19
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 34

atgcgaggct ccgatgta 19

<210> SEQ ID NO 35
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 35

ggccttggg ttttgatta 20

<210> SEQ ID NO 36
 <211> LENGTH: 22
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 36

gtccatggtt ttagacctcg gg 22

<210> SEQ ID NO 37
 <211> LENGTH: 19
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 37

gtttacttta cgcgaccg 19

-continued

<210> SEQ ID NO 38
 <211> LENGTH: 18
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 38

 gaattagacc cccccgag 18

<210> SEQ ID NO 39
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 39

 tagagtgggtt gtatgtcccc 20

<210> SEQ ID NO 40
 <211> LENGTH: 19
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 40

 cacttctacg atatgccgc 19

<210> SEQ ID NO 41
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 41

 ttcaatctcg ggggggtcta 20

<210> SEQ ID NO 42
 <211> LENGTH: 19
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 42

 tccgtagatt ccgagtcct 19

<210> SEQ ID NO 43
 <211> LENGTH: 24
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 43

 caagcgccat ggcacacat aaat 24

<210> SEQ ID NO 44
 <211> LENGTH: 27
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

-continued

<400> SEQUENCE: 44
aaacactagt ccatgtgcat gtcccg 27

<210> SEQ ID NO 45
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 45
gcttgcggtg ttttttggt 19

<210> SEQ ID NO 46
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 46
gccagcccct ttaaggtga 19

<210> SEQ ID NO 47
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 47
gccaatgaaa tgaaactatc gg 22

<210> SEQ ID NO 48
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 48
cggctcacag agctgctctt cgggtgtag 28

<210> SEQ ID NO 49
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 49
taatccttcc cctcatatca caacgcgt 28

<210> SEQ ID NO 50
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 50
gcggcctcca acatcaca 18

<210> SEQ ID NO 51
<211> LENGTH: 19

-continued

<212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 51

 ccagcatccg gctctgttg 19

<210> SEQ ID NO 52
 <211> LENGTH: 24
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 52

 cgtgtaggta caaacattcg tggc 24

<210> SEQ ID NO 53
 <211> LENGTH: 30
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 53

 atctaataccg tgggggtgcg agtgtacaag 30

<210> SEQ ID NO 54
 <211> LENGTH: 31
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 54

 tccatgttcg gtgatgtcct ctgtaggcgt g 31

<210> SEQ ID NO 55
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 55

 ggatgctgac gacacacccc 20

<210> SEQ ID NO 56
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 56

 gttgcagttg acggattggc 20

<210> SEQ ID NO 57
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 57

-continued

ttgtacccat cttcacgctc	20
<210> SEQ ID NO 58 <211> LENGTH: 20 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Oligonucleotide primer <400> SEQUENCE: 58	
atcgaacgtg cggcctgacc	20
<210> SEQ ID NO 59 <211> LENGTH: 20 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Oligonucleotide primer <400> SEQUENCE: 59	
cttcccggac aactgcccac	20
<210> SEQ ID NO 60 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Oligonucleotide primer <400> SEQUENCE: 60	
caccaaccgc aatcgcaatc ctttgaaggc	30
<210> SEQ ID NO 61 <211> LENGTH: 30 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Oligonucleotide primer <400> SEQUENCE: 61	
tattaacaac aaacagtccg cgcgccagtg	30
<210> SEQ ID NO 62 <211> LENGTH: 25 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Oligonucleotide primer <400> SEQUENCE: 62	
gccgaggtct tccacaccg attct	25
<210> SEQ ID NO 63 <211> LENGTH: 25 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Oligonucleotide primer <400> SEQUENCE: 63	
tttgaaggtt aaggccccac tcccc	25
<210> SEQ ID NO 64 <211> LENGTH: 25 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence	

-continued

<220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 64

 tacaggcagc aggtccggac gcgaa 25

 <210> SEQ ID NO 65
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 65

 ttacgaggcc tcaacggaac ccgtg 25

 <210> SEQ ID NO 66
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 66

 gatctccgc ggtcacccct ctcca 25

 <210> SEQ ID NO 67
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 67

 caaggcgtac tgtaccccg aaacc 25

 <210> SEQ ID NO 68
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 68

 actcatgcct gggccgggaa ctgga 25

 <210> SEQ ID NO 69
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 69

 cgtcgcatac accgtgtgta cccgc 25

 <210> SEQ ID NO 70
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

 <400> SEQUENCE: 70

 tgccctcccc cggattccca gagta 25

-continued

<210> SEQ ID NO 71
 <211> LENGTH: 22
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 71

aagctccaag tctcgggtga cc

22

<210> SEQ ID NO 72
 <211> LENGTH: 350
 <212> TYPE: PRT
 <213> ORGANISM: VARICELLA ZOSTER

<400> SEQUENCE: 72

Met Gly Thr Val Asn Lys Pro Val Val Gly Val Leu Met Gly Phe Gly
 1 5 10 15
 Ile Ile Thr Gly Thr Leu Arg Ile Thr Asn Pro Val Arg Ala Ser Val
 20 25 30
 Leu Arg Tyr Asp Asp Phe His Thr Asp Glu Asp Lys Leu Asp Thr Asn
 35 40 45
 Ser Val Tyr Glu Pro Tyr Tyr His Ser Asp His Ala Glu Ser Ser Trp
 50 55 60
 Val Asn Arg Gly Glu Ser Ser Arg Lys Ala Tyr Asp His Asn Ser Pro
 65 70 75 80
 Tyr Ile Trp Pro Arg Asn Asp Tyr Asp Gly Phe Leu Glu Asn Ala His
 85 90 95
 Glu His His Gly Val Tyr Asn Gln Gly Arg Gly Ile Asp Ser Gly Glu
 100 105 110
 Arg Leu Met Gln Pro Thr Gln Met Ser Ala Gln Glu Asp Leu Gly Asp
 115 120 125
 Asp Thr Gly Ile His Val Ile Pro Thr Leu Asn Gly Asp Asp Arg His
 130 135 140
 Lys Ile Val Asn Val Asp Gln Arg Gln Tyr Gly Asp Val Phe Lys Gly
 145 150 155 160
 Asp Leu Asn Pro Lys Pro Gln Gly Gln Arg Leu Ile Glu Val Ser Val
 165 170 175
 Glu Glu Asn His Pro Phe Thr Leu Arg Ala Pro Ile Gln Arg Ile Tyr
 180 185 190
 Gly Val Arg Tyr Thr Glu Thr Trp Ser Phe Leu Pro Ser Leu Thr Cys
 195 200 205
 Thr Gly Asp Ala Ala Pro Ala Ile Gln His Ile Cys Leu Lys His Thr
 210 215 220
 Thr Cys Phe Gln Asp Val Val Val Asp Val Asp Cys Ala Glu Asn Thr
 225 230 235 240
 Lys Glu Asp Gln Leu Ala Glu Ile Ser Tyr Arg Phe Gln Gly Lys Lys
 245 250 255
 Glu Ala Asp Gln Pro Trp Ile Val Val Asn Thr Ser Thr Leu Phe Asp
 260 265 270
 Glu Leu Glu Leu Asp Pro Pro Glu Ile Glu Pro Gly Val Leu Lys Val
 275 280 285
 Leu Arg Thr Glu Lys Gln Tyr Leu Gly Val Tyr Ile Trp Asn Met Arg
 290 295 300
 Gly Ser Asp Gly Thr Ser Thr Tyr Ala Thr Phe Leu Val Thr Trp Lys
 305 310 315 320

-continued

Gly Asp Glu Lys Thr Arg Asn Pro Thr Pro Ala Val Thr Pro Gln Pro
 325 330 335

Arg Gly Ala Glu Phe His Met Trp Asn Tyr His Ser His Val
 340 345 350

<210> SEQ ID NO 73
 <211> LENGTH: 13
 <212> TYPE: PRT
 <213> ORGANISM: VARICELLA ZOSTER

<400> SEQUENCE: 73

Val Asp Gln Arg Gln Tyr Gly Asp Val Phe Lys Gly Asp
 1 5 10

<210> SEQ ID NO 74
 <211> LENGTH: 39
 <212> TYPE: DNA
 <213> ORGANISM: VARICELLA ZOSTER

<400> SEQUENCE: 74

gtggaccaac gtcaatacgg tgacgtgttt aaaggagat 39

<210> SEQ ID NO 75
 <211> LENGTH: 39
 <212> TYPE: DNA
 <213> ORGANISM: VARICELLA ZOSTER

<400> SEQUENCE: 75

gtgaaccaac gtcaatacgg tgacgtgttt aaaggagat 39

<210> SEQ ID NO 76
 <211> LENGTH: 124884
 <212> TYPE: DNA
 <213> ORGANISM: Varicella zoster

<400> SEQUENCE: 76

aggccagccc tctcgcggcc ccctcgagag agaaaaaaaa aagcgacccc acctccccgc 60
 gcgtttgcgg ggcgaccatc ggggggggatg ggattttttg cgggaaacc cccccccgcc 120
 agcctttaac aaaaccgcg ccttttgctt ccaccctcgt tttactgctc ggatggcgac 180
 cgtgcactac tcccgcggac ctgggacccc gccggtcacc ctcacgtcgt cccccagcat 240
 ggatgacggt gcgaccccca tcccctacct accacatac gccgaggcgg tggcagacgc 300
 gccccccctt tacagaagcc gcgagagtct ggtgttctcc ccgcctcttt ttctcactgt 360
 ggagaatggc accaccaac agtcttacga ttgcctagac tgcgcttatg atggaatcca 420
 cagacttcag ctggcctttc taagaattcg caaatgctgt gtaccggcct ttttaattct 480
 ttttggtatt ctaccctta ctgctgctgt ggtcgccatt gttgccgttt ttcccaggga 540
 acctccaac tcaactacat gaaactactg tccggaaggg gaaggatatt attctcgctt 600
 gcagcttgct gcgctgtgat gcacaacaaa agctatatat gtcaccaaag ccaacgtcgc 660
 catctggagt actacacca gtacgttgca taacctgtcc atttgattt tcagttgcgc 720
 ggacgccttt ctccgggatc gtggccttgg gacatcaacc agtgaataa gaaccgccgg 780
 tggctctggt tgaacgacga gtggcgacgc gttgttctgc ataagctctg tatgctgata 840
 cataaacaca gagtctgtat cgctatcaga ttccogaaca ccttccggta cccatactc 900
 cgataccctg gacattgogg atoccaaaaa tataatatta acaggatttg ottatacttt 960
 gctacagcct atataaattt atgtgcgata catcttaagt gcatccgtac gttatttata 1020

-continued

cattgcctgt	cacgtgaaaa	gactgtgtta	cccaataaag	gttctacaaa	aaatgcttta	1080
ttgggtgttt	gtttaatagc	tattatcgta	accacccccc	gtaaaatcat	aaaatgcatg	1140
taatttctga	gacacttgca	tatgggcatg	ttcccgcat	tattatgggc	tccactctgg	1200
tgcgctccag	tttaaacgcc	accgcccagg	aaaatcccgc	gtcagaaacg	cgatgtttat	1260
tacgagtgtc	tgcggggaga	actgtagacc	tgccaggcgg	aggaaactga	cacattacct	1320
gtacaaaaac	ctatgtaatt	attggcaaat	atagcaaacc	cggcgaacgt	cttagccttg	1380
cccgtcta	agggcgtgca	atgacccctg	gaggtgcaag	gacatttatt	attttggcga	1440
tgaagaaaa	gcgatccaca	acgcttgggt	atgaatgtgg	tacgggcttg	catttactgg	1500
ctccatctat	gggtacattt	ctccgcacac	acggtttaag	taacagagat	ctctgtttat	1560
ggcggggtaa	tatttatgat	atgcatatgc	aacgtcttat	gttttgggag	aatatcgcg	1620
aaaataccac	tgaaacacct	tgtataacgt	cgacgttaac	atgcaacttg	acagaagact	1680
ctggtgaagc	cgcaacttacc	acgtcagacc	gaccactct	cccaacccta	acagccaag	1740
gaagaccaac	agtttccaac	attcgtggaa	tattgaaag	atccccctg	caacagccgg	1800
tctgtcacgc	ggttagat	gcccgaacct	cgaggggcgt	attgatgtaa	tactaaata	1860
aaataccct	ttttcgtatt	gtacgtat	ttatttaaat	gtgtagtcca	tagtccg	1920
acagccgctc	gggttttcc	cccacataca	acatgatcgt	atgcctcgg	tgaccgctc	1980
caacactcgc	ccgagaagg	ggatttaca	tgacagtgat	acccaatagc	cgccagatgt	2040
acaccagct	gtccgactc	cagcatcatc	tgctgagttg	cgccgctgaa	gggtgcatcg	2100
cataggggtg	tataatagc	catttccggt	aacagtcgtt	gggaatttag	gaggctgcaa	2160
aacgctgta	ggtaacata	cattggggat	tcagatggtt	tatctcgacg	tccaagtcca	2220
atcaaaaaag	cggtgaaatc	atcagcccgg	ccgcatgttg	ctcgaagagc	acataacctc	2280
ttaaacacct	acagagggga	tggcgtcgtt	gcattgtgag	tggcagggca	tgtccacctt	2340
gtttccaacg	ccagtgccgg	tataacttgt	gtaaaccgac	ccaacgggtc	aggtttaaga	2400
ttcactcgga	tgggttgact	gctttcggaa	gctcccgttg	tatccattaa	ttaaaccgtc	2460
ggtacacgct	tgggtgtgtg	tttaccggaa	tcagagacgg	aattgcaaag	atattggttt	2520
gaaagcaatg	taatcccgcc	catatatccc	caacgtcgcc	ttaaaaactc	ccacaatatt	2580
acatttttat	tagtctttta	ttaatataga	atcacataaa	caattgataa	aatcaagggg	2640
tgggtgataa	tgattaaaaa	tataaattga	tatgttttac	aagcatgaaa	taggtattta	2700
ctatttctac	aggtaaatat	gcttaatgat	taaaaataca	aattagtatg	ttttgacaag	2760
catgaaaaag	gtatttttta	ttttagcagt	taaaggact	acacttaaaa	tatttaccgt	2820
atggacgggc	gtcagaaaga	tgcccggccc	aagttgagag	ggtacattca	acacgaccac	2880
actcgcgttg	gtgggtgatt	agggcctcta	aaacaccggc	cagacatgac	ccgggtgtat	2940
attctgtgaa	cactgaaacg	ttacaactga	tatcatcata	ttccacaaat	ttagagccac	3000
ggacaactat	attagcaatg	cgggcaatca	taacaaacat	ataagtagta	atacagctga	3060
tatcactaaa	acgttgctgg	cgcaacagtt	cggggagagt	acgagacccc	aaatcgttgt	3120
ccctgtttag	aagaagacat	cttacaaaag	gccccagctt	taactttaa	ttctcaaaa	3180
gtgacttcga	ggttgcaaca	atgggattat	ttgtgtagat	gggcaagttt	tttgccgcta	3240
acattttaat	ccacgttaac	agttcatccg	cagactccaa	cgcttcaatc	aaagattctc	3300
cacgtatgac	tctctcacgc	aacgcggggg	caatcgtga	gtccatttta	tatgactcaa	3360

-continued

aggtacgata	aagttcatgt	ccgtacaaca	tcaactccgg	ccaagatgtg	ttttgtttta	3420
tccccgaaa	acatccaccg	gaagcccatg	aatcaccctc	ttgtattgtg	gcatacggga	3480
ctaccagttt	ttcaattggt	tcatctaaat	ggcgtaccga	gtcaatggtc	acgtggctc	3540
ccgcggtgga	gacgacttca	atagcacggc	ccgtaattcg	atcgaccggg	atatcatact	3600
cttttcgaat	acgtctcctg	cgggcgtctc	tcttgaaaa	tcgcaacctg	tacgattcgt	3660
catgtgtctg	atcatttctt	tctccctggg	tcattgcagg	aggcgttgta	ggacgccgtc	3720
ttcgatttga	cagggatcga	tcacggtggt	ttcttgaact	ttgagtgtta	taagatctgg	3780
atgatcgtcg	atgtccccgt	tcgatgcgtg	catatccagt	ctccacgtct	cttcctccat	3840
gatggtttga	atcgggtaat	acaacaacca	aagttttcgg	gcgattgtgg	tggtagcttt	3900
cacgccttcc	gtgccttctg	ttggaatacc	gtggattata	tgctgtatct	gcagtacgct	3960
ccacatacac	agttctagac	gttgtggagt	cctcgcctgg	agtggagcca	atagcttcat	4020
catttgccca	atcgggtgact	tccaatgcaa	agtcatccga	aggttcgtct	ggtagcaaat	4080
tcataaagtc	ttcacaata	gtagacacgt	ctgggtcggg	tggaattgaa	gcagaggcca	4140
tggtctgcaa	atatctgaca	attgcgtggt	tgcaagttgc	tgatcttcc	gccaatgttg	4200
tagaatttat	aggctcacc	aaccocgcaa	tggcgtggtt	tagtcacatg	attaatgctt	4260
ctgggagttt	tcactttccc	caaacaagct	tacctgcacc	ctttgttctg	aatgcataaa	4320
aataaccact	gctatagcaa	atatgacgat	ataaaaacat	tttatagcaa	ggccggacat	4380
tactgtagcg	caacatggtg	tgcatatacc	acgtattccc	cccgatttga	tatgatttaa	4440
atgattatoc	ttggttggtt	ttggtctaac	ataagatata	agcttacta	tagcgagcgt	4500
gcatacaaca	accagggcca	gaatccgaat	gtatgtgggg	tataataacg	cgcattggtg	4560
atatgcaacg	ccaagcgtta	aaagcacaat	acatccagat	gatatatgag	cgataacctc	4620
caaaagcacc	aataacgtaa	caactttatg	catatataaa	aaacttatag	ggtcagcatt	4680
aaatacttta	ctcataccat	cccgtcgcac	ggaaacatca	cataacaacc	ttgccaactt	4740
tgatattggg	taaccaagaa	gaatgttcga	aataaccctg	gttacgtaat	tcagtgaata	4800
tgatgtgggg	gatattaact	cacaggatga	tcggaatggc	caaacatac	gacgtattcg	4860
tcgaaattgt	aaatacatac	catatacaaa	ccatgcaaaa	aaaatcattt	ttagctgcac	4920
gcaccaaaaa	taagcgtgac	aattacgtgt	tcccagaaca	attcgaattt	tgctcatgcaa	4980
agggtgtaga	atagcggttt	ttaccatagt	atctcctgat	aatagatttt	cccgcagct	5040
gtaatcgtat	ccagatagcg	catccaaaaa	cgttgagtgg	tttacaacag	ttacatatat	5100
aagagagttg	ttataagacc	cccatacaac	cggtccacca	ttaatcaccg	tggttgcata	5160
cacacactca	tggtcaaaact	ttacacgagc	ggtataccat	agggtaaaaa	cagcatgtcc	5220
gctaagtaga	cacataatta	taaaatgttc	tgtcttgatt	cctaaagcct	gcatgaccgg	5280
tggaagatgg	caattcaagc	acgatgtagt	atcacacggt	tggtgttaac	tcgaagttaa	5340
atgttgataa	ttaggtactt	ctagagtaaa	gattgtatgc	atgogattgc	tatcgcactt	5400
tgtagcaaaa	cattgtttgtg	caagcgaat	acacaaaacg	ttgtgatgat	ccactcgcag	5460
agacacaaat	gtccggggag	ccgttcttcc	tccgcgatgg	ggatatcgaa	gacaagttaa	5520
ccctttttgt	ccgcataatg	gctgaaataa	caccagctcc	cttttgatgg	cgatacactt	5580
tgatgatggt	aaggtatatt	cgcgatcagc	cccggggaaa	tgaacagcaa	tatgctccac	5640
aatagattct	aatattgtgc	tgctgacaaa	ggcctccagt	gtaaatgcgt	ccagacaagt	5700
taccocgcgc	tcttttagag	cctttgttaa	agatatttgc	ggggagctaa	atatttgttt	5760

-continued

attacgcgca	accttacggt	caaaaaactc	tgcgatttcc	ccccaaggt	tatgtaaaat	5820
aaattgcact	ggaacattcg	actgcggtct	tgaatgaaa	tgaagtttg	ccgggtttct	5880
atgtgatgtc	acaaacgcta	atataatcaat	acactgctca	ggtacaacat	aaaatgggag	5940
tagttgtcca	accgcccgtc	ctgtggttgt	tactttggag	aaaaaaggca	gtcttaaaact	6000
atgtccgtgg	ctataaacac	cagtatctat	aaacgaaaag	tcccgtaaat	acggaccaat	6060
atattcaaca	aattcccgtt	ccagcaaac	cgcttgctgt	aatatttgg	caaaccctt	6120
taaagtggaa	gaccccacta	acgcataggg	atttgggatt	ggtacgcata	cctgaaacc	6180
tattttctct	ttacagttac	agggtagagt	ttcatgcaag	ttttcattgt	ttgatacatc	6240
ggcgtgtgta	tggacttcag	acgttgtctg	tgtatcaaaa	aaccatacat	cctctgtata	6300
attctctctt	acacacgtgt	ataattcgcc	attttctatg	taaaaatcga	tgtcagaatg	6360
gctggttata	tccaataaat	tatcatcatc	caacacctca	acggtaggtt	caggacatgc	6420
agttttataa	aaataacatg	ggtctttggt	agggtttacc	acggcctttg	gaaaaagtaa	6480
ttgcatggcc	gttaaaatac	catgacgaaa	tgctcgcatg	ccggcatgta	aaatacccaa	6540
tgggatgggt	tttcttatat	gaaagtctac	atcaagtatg	aggtttgtga	ttataagatt	6600
tgtattaaat	agctcattcc	tgtttatata	aagctgatct	ttgggtatgt	ttgatgaaat	6660
tttagaaaacg	tttttaacag	acgtagataa	tagtaaagtc	aactgcataat	ctcgtagtga	6720
agcgcaaca	aaattacatg	gattaatttg	tttaaggtcc	tccgcaatta	atcgagcctc	6780
gtgcggtaaa	gtgtaacggt	ttgttattga	tgaccacgta	tcattagcaa	taacagcaaa	6840
tgcttggggc	ccgtgaggca	aggctaccgg	atatacaggc	attggtccag	ttacctcaga	6900
atggccgatg	agggcttcta	atggagtttt	ataactcagg	atggatacat	catgtgtggc	6960
tatcccagtg	gcagcagaga	aaaacagtaa	tagttttgta	atccccgggc	tcgtatcaaa	7020
accagtacga	ccactttggt	taggtgtatc	gtttgcaaag	ttggctgctc	gtaacgcctc	7080
cgcgaaaaca	cccgaatcct	caaaattaga	caattcgtca	aaaccgggtg	gatttgaggg	7140
aatagtgagg	gaccatccat	atggactaaa	ttgtttttca	atgtttttcca	cacgacgagt	7200
tagcgttgta	gctaggtcac	atacgcctat	aaacttgcta	ggttttgagg	catacgtaa	7260
acttaaagta	tatgttttag	taattgtata	tttatgtcca	atctcaggtc	caagttcagt	7320
gacatcacaa	attacgttct	tttttatata	gtcacgcatg	ttgagacgag	aacgtacatg	7380
attaaaaaaa	ttagcagtag	ctctttttcc	caggttggtat	gattttaaga	ggaccggttt	7440
attcacaaaa	tctgagtatg	taaccgcttg	taggtgtctc	gcgatctggt	tccgattgaa	7500
acattcaaaa	tgtgccagat	aaatataatc	aacaaattca	cggtctggaa	ctttaaggcc	7560
ttttctatcg	ttggtaatat	actccgatac	tgcggtgatt	tccggtgtgt	ctgtatgtat	7620
tcgctgtaaa	atgtacgata	gagcattttt	ggctgtcaaa	cctcgtgtat	atggtgagga	7680
acaacaaaa	atggaagtt	tatcaaaaga	caacaagtcc	gaaatattgt	accactaca	7740
attaggtaat	gccgggactt	ggtaagttaa	aaacaaatct	ttaattgcct	gtaagtcata	7800
taagggggtt	tccaacgtat	tgaacttgt	gtccgtttgt	aacaagtaat	agcgtgtagc	7860
caacactagc	gttttttcag	agggctcaaa	togaacaata	taccaaaacg	gcgagcatcc	7920
atacccccag	tagagtcgct	gatatgcagc	caatacttga	cgttcgtaat	gggcatataa	7980
tgatgttagc	tcctgacgac	caacggatgt	tttaactaac	ttgcagagtg	ttgcctctgt	8040
gatgcatagg	ccgttgcgc	ataatccctt	tcggtttaaa	tggtgtgttg	ttaccatcag	8100

-continued

agtttgata	acttccgagt	gaatgtcaaa	cgtctccgat	atacataggg	tatcagatat	8160
tatatgcgga	tttaggggtg	ctccatacca	taacgcctta	tataaagctt	taaaatcagt	8220
ttgggtttta	aaacaacaaa	aaaatatagg	ccagaccogg	gatcgtacat	ctccagttga	8280
aaatccacca	attaataaaa	aaataacggt	gacgtcccta	ctacaaaata	aatgcattat	8340
ttggttttct	tcacgttttt	cagttacttc	acgtgggctg	ttagttggga	ttacttgcgt	8400
gatctcttcc	ctcccatttt	tgacaagac	gtcatctaag	tcgggagtcc	aagtataact	8460
caccacatac	agaggttctg	tgcttatctg	cccggtaagc	aacaacagcg	agtgaggagat	8520
tgccatccc	tttggtgcaa	ataataaccg	aatcgtcggg	ttggaggatt	tatccatagt	8580
tcaatacgtt	ggaaagccag	tcaatcatgc	agacggtgtg	tgccagctta	tgtggatatg	8640
ctcgaataac	aactgaagag	ccatcttatg	aagaggtgcg	tgtaaacacg	cacccccaaag	8700
gagccgcctc	gctccgcctc	caagaggctt	taaccgctgt	gaatggatta	ttgcctgcac	8760
ctctaacggt	agaagacgta	gtcgttctg	cagataatac	ccgtcgtttg	gtccgcgccc	8820
aggctttggc	gcgaacttac	gctgcatggt	ctcgtaacat	tgaatgttta	aaacagcacc	8880
atcttactga	agataacccc	ggtcttaacg	ccgtggctcg	ttcacacatg	gaaaactcaa	8940
aacggcttgc	tgatatgtgt	ttagctgcaa	ttaccocattt	gtatttatcg	gttgccgcgg	9000
tggtatgtac	tacggatgat	attgtcgcac	aaacctgag	aatgaccgct	gaaagtgaag	9060
tggtcatgtc	tgatgttgtt	cttttgaga	aaactcttgg	ggtcgttgc	aaacctcagg	9120
catcgtttga	tgtttcccac	aacctgaat	tatctatagc	taaaggggaa	aatgtgggtt	9180
taaaacatc	acctatataa	tcggaggcga	cacaattatc	tgaattaaa	ccccactta	9240
tagaagtatc	ggataataac	acatctaacc	taacaaaaaa	aacgtatccg	acagaaactc	9300
ttcagccctg	gttgacccca	aaacagacgc	aagatgtaca	acgcacaacc	cccgcgatca	9360
agaaatccca	tgttatgctt	gtataaatat	tgaataaaa	actaaaaacg	tttctggtgt	9420
atgtttttat	tttgatata	aaatataaac	attgctggct	ggcgtgggta	ttacatttaa	9480
tgttttagta	gaaaatcgac	atcgtttggt	tctttatcag	ttgaacccaa	tccacgcggt	9540
ccccgttcgc	tggtgtggc	tattgatctc	aacgttttag	taaaatacca	ttgtacacc	9600
ggatgcccac	atttaccgcg	gatagcataa	ggaaatgcaa	tattacttaa	aacgttgtgt	9660
tttaagtgt	tttgggtgtt	gtgatctatt	aacaggacct	gtgcaagacg	atctcccgtt	9720
tttatacgt	tgatcatcacc	cgtgagatta	tatacgtaga	atctacagt	ttctcctgca	9780
ggccatgccc	ttggacacac	gataatgcct	gatcggcttt	tcgatgatct	tccaaaaata	9840
taagcgttta	tactcggatg	ttgtaagtcc	cagtctctta	taatcggtaa	gacaattttt	9900
ataaattcat	tcctttttaa	atatagggta	tatggtacac	aaatatcata	tcccgcgtct	9960
tcttggcgtt	ttgattgat	gatatgtttg	taggttaagg	gaacatcgat	atggtattct	10020
gcagaatccc	tatgtaaagg	ttgccctgc	tgtaccgtgg	aaatatcagc	aaattcaggt	10080
ataacggggt	ttcataaatt	tgacggcgag	tttgataagg	ggtgaacttg	tatcgattta	10140
aaaattggat	ccagatgttt	aagaacgttt	tttgggagaa	ggcgactttg	tcttaatttt	10200
accgggaaca	agtagattgt	taaatgtccg	ggtaaaataa	cggttactcc	tgcccggtaa	10260
tacaaaaggg	ctgaaattac	tcctctgtaa	cccgcaccaa	taactccggt	ggcgacaaaa	10320
aaattgtctt	catcagcaag	ggcagtatct	ttgcattgaa	ttaacaacag	tgcgatttca	10380
ttgggaggog	ccgacttaac	caacagctcc	aactgtgca	tataaaaacc	gccccgtgtt	10440
acagattttt	cagatggcag	ttcgagtttc	ttgtggttcc	ggagtaacaa	cggttgatgt	10500

-continued

cgacttactt tatcgtctaa cacgcattgc agcgtatctg cacattcagc ttgaacttct 10560
 attaaaattg tatcttttaa acaccgatcc ggaatagttt ggctacaaaa catatcacct 10620
 gtatttactg ccgtttccaa gatgggatca attaccgctt cgttcatatt aataacgatg 10680
 caaattttat ttttttgta agacagcagt ggggagccaa actttgcaga acggaatttt 10740
 tggcatgcca gctgttcgcg tcgtggagtt tataatcgac gatcaatgat caccaccctt 10800
 ttcttctacg catccctttt ggggtgtgt gtagccctta ttcgtagc ttatcatgcg 10860
 tgtttccggt tatttactcg ttctgtatta cgcagcacgt ggtaaaccg tttgcctata 10920
 aaaggggagc gcgtgtataa gagggccctt gtttaatacg cgtctgccc tgtttggata 10980
 tttcacgacc ctatcgttta ttacgtaat ggcatcttc gacggtgaca gactttgctg 11040
 ctctaataca gtgcgtcgta aaacaacgcc tagttattcc ggacaatc gaaccgccc 11100
 gcgaagtgtg tgcgtaggac ccccgatga ttcagacgac tcgttgggtt acattaccac 11160
 agttggggcc gattctcctt ctccagtgt cgcgatctt tttttgaa ataaaaatac 11220
 gaccctcgc gtacatcaac caaacgactc cagcggatcg gaagatgact ttgaagacat 11280
 cgatgaagta gtggccgctt ttcgggagc cgtttgaga catgaactgg ttgaagatgc 11340
 tgtatatgaa aaccgctaa gtgtagaaaa accatctaga tcttttacta aaaatgccc 11400
 ggttaaacct aaattagagg attcaccgaa gcgagctccc ccgggagcag gcgcaattgc 11460
 cagcgggaga ccaatttctt tcagcactgc accaaaaacc gcaacaagct cgtggtgccc 11520
 tctacgcca tcatataaca aacgcgtctt ttgtgaagcg gtcgggccc tagccgccat 11580
 gcaggcaca aaggctgcc aagcggcttg gaatagtaat ccccaagga ataacgcca 11640
 attagaccgt ttgttaaccg gagccgttat tcgtattacg gtgoatgagg gtttaattt 11700
 aatacaagcc gctaatagag cagacctagg tgaaggagca tcggtatcca aacgtggaca 11760
 taatcgaaaa actggagatt tacagggggg catgggtaat gaacctatgt acgcacaagt 11820
 tcgtaagcca aaaagtcgaa cgatacaca aacgactggg cgtataacta atcgaagtag 11880
 ggcccgttct gcatcaagaa ctgatacgcg aaaatagga tataattacg cagtaacggt 11940
 ttaccgggta ttatgtataa taaataaacg tataaaagac agtcgtggtt tgtgtttatt 12000
 ataaatggtt attatattgc acatattata aactgtttaa atagtaccac gtggtattat 12060
 gaacagttta taatcagttg ctaccaaaca aaccocatta gacggcgggt tttgataaag 12120
 ggaatcgctt atttaacta aagattttac tctataagta tggagtgtaa tttaggaacc 12180
 gaacatccta gtacagatac gtggaatcgt agtaaacgag aacaagcggg tgtggacgca 12240
 tttgatgaat cgttgtttgg tgatgtagca tcggatattg gatttgaaac gtcgttatat 12300
 tcacatgacg ttaaaactgc tccgtctccg ccttgggtag ctagccctaa aattttatat 12360
 caacagttaa tacgggatct tgatttttca gaagggccc gtttactatc atgtcttgaa 12420
 acctggaagc aggtattatt ctcatgtttt cctattaatg aggacctata ttcgatatg 12480
 atggttttat ccccgatcc agatgacgtt atctcaaccg tttcaaccaa agacctggt 12540
 gaaatgttta atttaacaac ccgggttcc gttcgattgc ctagtccacc aaagcaaccg 12600
 acggggcttc cagcttactg tcaggaggtc caggattcgt ttaccgtaga actacgccc 12660
 cgggaagaag catacacaaa actactagtt acttattgta aatcgattat acgttatctc 12720
 caaggaacgg cgaaaaggac gacaataggt cttaataatc aaaaccctga ccagaaagct 12780
 tacacgcaac tcaggcaag tattctactt agatattatc gtgaggtggc aagtttggcg 12840

-continued

cgtcttctgt	acctacattt	atatttaacc	gtaacgcgtg	aattttcctg	gcgtttgtac	12900
gccagtcaat	ctgcacaccc	ggacgtgttt	gcggctttaa	aattcacctg	gaccgaacgt	12960
cgacagttoa	cggtgctgtt	tcatcctgta	ttatgcaacc	acggcattgt	gttattagaa	13020
gggaaaccac	taacagcgtc	tgcttgagg	gaaataaatt	accgccgccg	agaactggga	13080
ctgcctctag	ttagatgtgg	tcttggtgaa	gaaaacaaat	ctccgttgg	tcaacaaccc	13140
tcattttcgg	ttcatttacc	acggtcggtg	ggttttctta	cccaccacat	taagcgtaag	13200
ttagacgcat	atgcggtcaa	acatcctcaa	gaaccgagac	atgtacgagc	ggatcatcct	13260
tacgcaaaag	ttgttgaaaa	tagaaactac	ggtagtagca	tcgaagctat	gattttagca	13320
cctcctgccc	catccgagat	cctgccgggg	gaccaccac	gcccaccac	gtgtgggttt	13380
ttaacgcgtt	aaacgtcatt	gggtagagg	gtgtaataa	attacgaaa	cgatgatgag	13440
ttttttat	ttacaatgag	ccgtatatgg	tatgtctg	atgtgctcta	aagtccata	13500
tataaaagaa	gcccacaga	gtgtatgctg	attgctgacc	gcgacctg	gatgttttac	13560
aggcgcgttt	gtttgtctcg	gttataagta	tcagtcggg	tcattataac	cgagggaat	13620
cccgcgcaca	gcgatatc	tctaatacca	cagactcccc	ccgtcacaca	cacggaacac	13680
gttatcggtc	aaccaattgg	tatacacacc	caccocagat	attgtccaat	tcagaaacat	13740
tagttgcggt	tcaagaacta	ctgaactccg	agatggatca	ggacagcagt	tctgacgcat	13800
cggatgat	tccgggatac	gccttacatc	attctacata	taatggatcc	gaacaaata	13860
catcaacttc	cagacatgaa	aatcgcata	ttaaattaac	ggagaggga	gctaagagg	13920
aatcaacat	caatcggac	gcgatcgacg	acgagggaga	ggcggaggag	ggagaggcgg	13980
aggaggacgc	gatcgacgac	gagggagagg	cggaggagg	agaggcggag	gaggacgcga	14040
ttgacgacga	gggagaggcg	gaggaggagg	aggcggagg	ggacgcgatt	gacgacgagg	14100
gagaggcgg	ggaggagag	gcggaggagg	gagaggcgg	ggaggagag	gcggaggagg	14160
acgcgatcga	cgacgaggga	gaggcggagg	aggacgcggc	ggaggaggac	gcgatcgacg	14220
acgagggaga	ggcggaggag	gattat	ctgtaagtca	agtttgcagt	cgagacgcgg	14280
atgagggtta	ttttacgtta	gaccggaaa	taagtacag	taccgatctt	cgattgcaa	14340
aggttatgga	gcctgcggta	tcaaaggaa	ttaatgtatc	aaaacgttgt	gttgaacctg	14400
ttaccctaac	aggctctatg	ttagcgcata	atgggtttga	tgagtctctg	ttgctatgc	14460
gcgaatgtac	ccgtcgcgaa	tatattacgg	tccaaggatt	atacaccaca	attcatttac	14520
ggtatcagtt	tgatacttcc	cggatgacac	ccccacagat	tttgagaact	ataccagccc	14580
ttcctaacat	gacacttgg	gaacttttat	tgatttttcc	tattgaattt	atggcccagc	14640
caatttctat	agaacgtatt	ttagttgaag	atgtat	agataggcgg	gcttccagta	14700
aaacacataa	atacggccc	cgttggaatt	ccgtctacgc	acttccatat	aatgcgggta	14760
aatgtatgt	acaacacatt	cctgggtttt	atgacgtg	cttacgtgct	gtggccaag	14820
gaacggccat	ttggcatcac	atgatattat	ccacagcagc	atgcgctatt	tctaactcga	14880
tttcacatgg	agatggatta	ggatttttgg	tagacgcggc	aattcgtatt	agcgaaact	14940
gtattttttt	gggacgtaac	gataat	gcgtggggga	tccatgttgg	ttagaagacc	15000
atcttgccc	attaccacga	gaagccgtac	cgcacgtact	ccaagtgaca	cagttgggtt	15060
tgccaaatgg	gggtccaacg	gttgccatta	tcgctgggtt	ttttggggcg	ttggcatatt	15120
ggcccgaact	aagaattgct	ataagtgaac	catctacatc	tttgggtgga	tatgctaccg	15180
gtcacatgga	acttgccgaa	tggtttttat	tttcacgtac	acatagt	taagccaacat	15240

-continued

ttacccaac ggaacggaa atgtagcgt cttttttac gttgatggt actcttggtg	15300
gaggatggt gaactggatc ttagagcaa ctgcaatgta ttagctgct ccttaccatt	15360
cccgttcgac ttacatcgcg gtctgtgaat ctctgcccta ttactatata ccggttaata	15420
gtgacctggt atgtgattta gaggtattac tgttaggcga ggtcgacctc ccaactgttt	15480
gtgaatccta cgcaactatt gcacacgaat taaccggata tgaggctggt cgcacagcag	15540
ccacaaattt tatgatagag ttgcccatt gttataagga aagttagacc gatttaattg	15600
taagcgcgta cctgggggcc gttttattgt tacaacgggt gttgggtcat gcaaatcttc	15660
ttttgttgt tctctccggt gctgctgtt acggaggatg ttcaatttac atccccgag	15720
gtattttaga tgcatataat actttaattg tggcagcaag tcctctttac gctcaccaaa	15780
ctttaacatc cttttgaaa gaccgcgatg atgcaatgca aactttgggg attcgaccga	15840
caacggacgt tttacccaaa gagcaagaca ggatagttca ggcacacct atagagatga	15900
acttccggtt tgtgggattg gagaccatct atccccgaga acagccatt cctcctggtg	15960
acctagccga aaatcttatg caatacagga atgaaattct gggtttggtg tggaaaagcg	16020
tagccatgca tttactacga aaatattaag ggtgtgatt tttttcatta ggatgaaaag	16080
aacgtttcct agccacacc acaaaggagt ttgtaaaata aaatctctgt ttagacctta	16140
aaattgttg tgtgtgtgt gtggggggtc cgtgagatc gacctttaca agatataatt	16200
gtccatata gcaatgttt ctoggtttgc gogttccttt tccagcgatg atagaacgag	16260
taaatcttat gatgtagtt accaaagttt taatgccgcg gaacgtgatt tgcccacacc	16320
taccgggac tgggtttcta tttcccaacg cataaccagc gagcgcgta gggatggatg	16380
tcttattcca acgcccggcg aggttttga gacggcggtg aaggctttat ctgaaaagac	16440
cgacagccta acatcgccg ttttcaaaag taccgaaaga cacagtgttc tgcttgatt	16500
acaccataat aatgttctg aatcgttggg ggtctcgtgt atgtctaacg atgttcatga	16560
cggtttatg cagcgttata tggaaacaat tcaaagatgt ttggatgacc tgaactttc	16620
tgggatgga ctttgggtg tttatgaaaa tacatattgg cagtatctca aatacaccac	16680
aggagccgag gtaccgtgta cttcagagaa ggtaataaaa aagtctaaat ccacggttt	16740
gttgttttca tccgtagttg ccaataaacc aatatccaga catcctttta aatctaaagt	16800
tataaattcg gattaccggg gaatatgtca ggagctactg gaggcgttag gagctgtgca	16860
aaagtatatg tattttatgc gtccagatga tctacaaaac cccagcccgg atacaagaat	16920
acgtgtacaa gaaattgctg cttacacggc tactggctac ggttgatgt tatggttctt	16980
ggacgttggt gacgccaggg tatgtcgcca tctcaaaact caatttcgac ggattcgagg	17040
gccgcgagcg tctgttattc cagatgattt gcttagacga ctttaaaaa cgggtcctgc	17100
ggtctcagcg ggcacaggag ttgcttttat tttagcagca acaactgccg gcgctcttac	17160
tgcgcttttg cgtattagt tattatggcg aaaggagag tggcgggatg gtttaaatgg	17220
aaccgcagct gcaattgttg cggcggttga acttattacg cttttgcacc accattttca	17280
atacttaatt aatgatgac ttattggata tgcattgttg ggggatgggg gattaaacga	17340
tccttatata ttaaaggcgc tacgtgccca gggacggtt ttatattttg cgggtcagtt	17400
ggtcagaaca atgtcaacac acagttgggt tgggttagag accagcacc atagtgggt	17460
ttcccgggcc gtggcgagca gtattttagc acatgggggt aaaccacaa agtattatgc	17520
tcaggttctt gccgccagta aacggtatc tccgttacct ttaagcgtg tatccgaacc	17580

-continued

atcgagtgtg	tctgatcagc	cgtatatctg	ttttaatcga	ctgggatctc	caatagggac	17640
aggatatagg	aatttggaat	gtgtctgttt	aacgggaaat	tatttatctg	acgacgtaaa	17700
tgcaagtctg	catgtaatta	atacagaagc	accgttaaac	agtatagcac	ccgatacaaa	17760
tagacagcgg	acttctcgcg	ttttagtctg	tccagacacg	ggtttgatg	taactgtccg	17820
aaaaaaccc	tgctctggaca	taggccatac	ggacggtagt	ccagttgacc	caacgtatcc	17880
tgatcattac	accgggataa	aggcggaata	tgaaggtccg	gttcgggatg	aatcaaacac	17940
aatgtttgac	caaagatcgg	atttacgtca	catagaaacc	caagcatctt	taaatgatca	18000
cgtatatgaa	aataaccac	ccaaggaagt	gggttttaac	tcactctcag	acctggatgt	18060
ggatagcctt	aacgggtaca	cctccggaga	catgcataca	gacgatgact	tatcaccaga	18120
ttttataccc	aacgacgttc	ccgttagatg	taaaaccacg	gttacgttta	ggaaaaatac	18180
gcctaagagt	catcattaag	tacagcggtt	aatagatagt	tatggactag	gcactttggc	18240
ggtcatttcc	acaaccagg	taaaattggg	ggatttgga	gaaaatagtc	tattgcgtat	18300
tttctgttca	ataattggac	tgctttat	aaaggtctga	ttggttgatt	gggttataaa	18360
aggaattact	cctttaaatt	ttacttaatg	taccacaat	atcaagtggt	cgtttgatt	18420
taacgattat	taccggtacc	atgggagact	tgatcatgtt	gacaaagg	ccgggtttta	18480
cgtaaacgg	gaaacttcag	tacttaaac	aagtggatga	tattttaagg	tatggagtcc	18540
ggaaacgcga	tcgaacagga	atcggaacgt	tatctttatt	tggaatgcaa	gctcgataca	18600
atctgcgaaa	tgaatttcc	cttttaacta	caaagcgtgt	ttttggagg	gccgtcgtgg	18660
aagagtgtgt	atggtttatc	cggggtcaa	ccgattccaa	agaactcgcc	gctaagata	18720
tacacatatg	ggatatatac	ggatcgagca	aatttctaaa	taggaatggc	ttccataaaa	18780
gacacacggg	ggaccttggc	cccatttacg	gcttccagtg	gagacattt	ggagcggaat	18840
ataaagactg	tcaatcaaac	tatttacagc	aaggaatcga	tcagctgcaa	actgttatag	18900
atacaattaa	aacaaaccca	gaaagccgac	gaatgattat	atcgtcttgg	aatccaaagg	18960
atatcccctt	aatggtaacta	cctccatgtc	acacgttatg	tcagttttac	gttgcaaacg	19020
gtgaattatc	ctgccaagta	taccagagat	cgggggatat	ggccttggg	gtaccgttca	19080
acattgtctg	atatgcactt	cttacctaca	tagtagcgca	tgttacagga	cttaaaaccg	19140
gagatttaat	tcatacaatg	gggatgcac	atatttactt	gaatcatata	gatgctttaa	19200
aagtgcagct	agctcgatcc	ccaaaacctt	ttccttgctt	taaaattatt	cgaaatgtaa	19260
cagatataaa	cgactttaa	tgggacgatt	ttcagcttga	tgatataat	ccacaccccc	19320
ccctaaaaat	gaaatggct	ctttaatgga	tttttaaag	ttgtcaagac	agtagatgtg	19380
ttgcgaatgt	aataaaatga	tatacacaga	cgcttgggt	tggttctgt	ttatgaacag	19440
caacggatgc	atagggttgc	gataactcgc	ataagacca	atgtccaag	gatagatatac	19500
acaccaatta	taactgtctac	aacggaaaat	gtagtggcgt	aggtagatgc	atcgtaggta	19560
taaacggccg	aaaacggagg	gaatttttta	gggtaacat	ctagatgaca	cgaataggtg	19620
ataggtccgt	cgagtccga	tggtggacaa	gaactttgca	tgtttcaaaa	ccgttggttt	19680
tgatcacaca	cccagtaat	ctcactgttt	tcgtggttaa	tgggagaatc	gttaaccac	19740
catacgaat	gtacaacgcc	acgtggcaca	cattttgccc	tacatactat	gtgtccatca	19800
ataataccta	tagacacgtt	gggaaatgga	tagacgtcag	gggtaacgac	agcagaatat	19860
ttcatattag	agacccatc	ccgaatccat	aaaacattac	attggatggc	tgggggtggg	19920
taatccattt	gttttctgt	tggaaatcgt	accgccgaaa	cataactaaa	taatccattg	19980

-continued

gcatattcctt gtattgcacg ggtataaaa tttttccga tgttacaaa ccttgaagtc 20040
caccgaacac gtaccgagtg cggggataa tactttgata cgttacagta ggctgcgtat 20100
gtctgtccgg ttaagactgg atcgccgaca acggaatat ttggacgata atacgttgta 20160
actgtaatac tegtgtccga tatgacgttc ttagtttttg tattaacgac tcgccaata 20220
tacgttccct ccggtgtagc atccatagat aaaattgta cagaaaaatc agacgttggt 20280
ttaacatctg gtattacata attttcccta gcgtgtgtaa atatctcagg gttgtttatt 20340
aagtttaaat cggcactggt gctatataac ataaccgta aatctggcat gcgtattaac 20400
gcattgccca gttgacggtg cggatctata aggtgacgcg taaaccaaac ttcaatatga 20460
agatcggggc gtataagcga ctccacctt gttatatttg aaccttcggt atctaaagaa 20520
tattgttcat atgttttttg ttgctgctta aaggccgctt gttgtccggt cgttagacgc 20580
atgtaacaag gcatgataaa tegtgtgaaa tagggatgg attgtattcc gccgtgaacg 20640
cattgtatat tttcatatag aaaagtggtt tegtgaatggt ggggtgtggc tcggggatcg 20700
ggctttcggg aagcggccga ggtgggccc acgcccggat cgggctttcg gtagcggcc 20760
gaggtggggc cgacggccgg atcgggcttt cgggaagcgg ccgaggtggg cgcgacggcg 20820
ggatcgggct ttcgggtagc ggccgagtg ggcgcgacgg cgggatcggg ctttcgggaa 20880
gcggccgagg tgggcccgcac ggccggatcg ggctttcggg aagcggccga ggtgggccc 20940
acggcgggat cgggctttcg ggaagcggcc gaggtgggccc cgaagcgggg atcgggcttt 21000
cgggtagcgg ccgaggtata taattcagtt ataactacgg gtgtggggtt agattcagtc 21060
gataattgta tacacgcgat cgttaaaatt aaatttattt gtatccgctt catcctggtt 21120
tttattgaca catccacgct ccccttaaat aaaagattaa aacaccacc cgggaattta 21180
aatgatgaa acgttttttt cgacattggg aataataaaa acggcttttg caactttaa 21240
aactttattt atctcgatta cgatacatat gtaccacata gatagcatag atttattata 21300
atataaacac acacgtgata tacttttagt atatgagatg ccataaaaca gtcaataggt 21360
ttaacgctta gtctcatcat ctgaatacac gtcaaaccg ccgcaactgt tgatgttaga 21420
attataatag ctccccatga aatgcccgca aatgttacag ctatacccgt caccgaggtc 21480
gttgatata atacaattac ccataggttt ttttttctt gatataaac gccaaaacc 21540
tgtaaccгаа atgtataat atgacctcct attgaaactg ctaacgttac ttgtgtaagt 21600
ttgataaaat gatttaattt aattatattt gagattgcc acattaatgg ggttaactata 21660
tataacaccg ggggtataac agacattata cgaattcctt taaacacgcg ttaagggtc 21720
cgggaacttt ctgatggtc acatactctc ccgcggtcat tttgtgtata tacaacggca 21780
aaacctaaat ctgtataagt gtttaattgc ttatggcgat ttttacgata tatacacgta 21840
tcttgcaaat cggtgccgac atcgacaatt gaaactagtg tgacaataga tatacacaat 21900
ccaataagaa cctcatattt actgacatac atataaaaa taacggttag taaacctccc 21960
aaccagttc ccaacatcat aacataaaaa taaatattgc gtccattgaa tgcgtaaca 22020
aagttgtagt aatggatatt cacagcagcc actgttccgg taatcgcgga tatggaaatt 22080
cccagtaatt ctacaaatgg aagatcccgg gatattgggc aaccaaccg ccataacaca 22140
gcaaaacca acacgaccac cgtctgcaa catcgtcca attttgctaa tgtgcgtaga 22200
aatttcacgg atgttgcca taaccccgaa acgacgatca accccataat agttgcattg 22260
acggcagctt cgcagacgtg atattgtaa attaaccgg acgtgataac gcttgcttgt 22320

-continued

```

agtcccacga gaaacaaccg cgatgctgag gttattgcac acgaattaca ttcttgaggg 22380
tttccgacac atccttggat tgattgagcg cggattaatt ctctgtctaa cacaccaggg 22440
ttttcatcat ggacagctct ttcaccatc acggccatgt cctaagttta ataattcaaa 22500
acaaataaaa atgtgttcat ctatggtaca cacaagtttg tatgtaaaat ataagcaaaa 22560
gttgacttta ttaactgta catattacgt cagattcacg tgataattca gaataatcca 22620
gggttcctgc aggttccact ggaggagcca cacaatattc gogaattccg attccctcct 22680
gccatgtggt ttcggggagt ttcccccca ttttatttcc ggtatttttt tcgtttcttt 22740
ttgttaataa attgctctct tttttaatg gtggttcac cttcacagat tccatgttcg 22800
caaataattg catcgaggtt aattttctt taaggctttt gggacttaag aacgttgcat 22860
aaaaaaaaa atgcacgggt gcggaacgtt ggatatacaa tccaaccatg ggggagttag 22920
ttaaggcgag ataaaaatta atataacacg tctcatcccg tgtaactta agattttgta 22980
cggcagaacg gaatccactg tgtgtttcca ataatactcc aaattcacgc atactccgc 23040
tgccataaac aacattatta aggatccttt ttgaattgtg gattgagcgt attaaattat 23100
atggtgtagg cttgcttccg tttatatcca aggaacatt aaatgagata aaaccacccc 23160
cggcggctcg gatgtacata tccgtggctg ttagaatgaa gcatgttgta aaccctaaag 23220
ttttaagtag tcgctgtaaa cgggtgaatt gatcgcgttt taagcaaatg cttatatctg 23280
gagtttagatt tggaacatc attgtataac aagcgagttc acgttttaca acttgttgt 23340
aacattgtac ttgatcatct ggaccacaat caccggggcg ttgccatacc atcgtttgga 23400
taatactcgc ctcggggggt tgtccggtaa atttaaaata taaccgtgtt ggggtgcagc 23460
gatcttttgt atggcgaaac gcgtcaataa gcgaggaccg tccctcgtt gccgcgagta 23520
caaccattct cggcccagtc caattatact ggtcaaacat atttgccgtt ataggaatat 23580
acagttgttc tgtttccaaa ctacagttaa taattaatcc ttcgtcgtg aatattaaaa 23640
tagaatccct tagtctatta accagagtg atatagacga aattaaacca gtaagcgttt 23700
tttccgttaa aacagctctg gcgatttctg gggcgtcaaa acccgcatgc aattccatgt 23760
ccaaagcatc gtctgtacgc gaactcaaat ccataattta ctacttaaaa tgtttactat 23820
agaaaaagta atcatatgta aacacacgag tttcgttaat atgtttgttt aaccgatcc 23880
ggtgacttaa gtacataaac aggcattgata tttgaatagt acggccatg ggaggaaca 23940
tttccacgtg ttccaatata ggggtgttcc cttaataggg actgtgcaat aaaatcgtta 24000
agaagtacc agatttgatg taatgtttgt cataaaaaat atgtacatca ttatatacgt 24060
ctgtaattaa cacaagatca catcgaagaa ttactgaagc cgctgtgaaa ctttcacaa 24120
gacgatataa acttggttaa gtgtattgat ggggctcttt ggactgacac gctttatcca 24180
tgaacataaa ctggttaaac ccagcatcat ttcaacgcca cccggagttt taacccccgt 24240
ggcggtagac gtatggaacg tcatgtacac attgttgaa cgtttatacc ctgtgggtaa 24300
acgcgagaat ttacacggac catctgtaac gatacattgt cttggagtct tattgoggct 24360
attaacacaa cggctactact atccgatatt tgtattgaa cgttgtagc acggccatt 24420
atcacgtgga gccaaagcaa ttatgtcacg ggccatgaac cacgatgaaa ggggaacctc 24480
ggacttaaac cgtgttctac tatcatccaa cacatcatgt tctatcaagt ataacaaaac 24540
atcgaaacaa tatgacagtg tgtttcgaaa ctctccacg agttgtattc ctacggaaga 24600
aaacaaatcc caggatatgt ttttgacggt ttgtccacga caaactgaca agacgatctg 24660
cctgcgcgac caaaacgtat gcagctctac ctctacaatg ccatcccgag gacatcctaa 24720

```

-continued

ccatcgatta tatcacaat tgtgtgcaag tcttattaga tggatggggt atgcatacgt 24780
 cgagggcgtt gacattgagg cggacgaggc atgtgcaaac ttatttcata cgcgtacagt 24840
 ggctttgggt tatakacag atactgattt actcttcag ggctgtgata ttttgtaga 24900
 tgcaattcct atgtttgctc cagtagtacg atgtcgcgat ttgcttcaat atttaggaat 24960
 tacataccct gaatttttgg ttgcctttgt tcgctgtcag accgatttgc atacaagtga 25020
 caacctaana tctgttcagc aagttattca ggataccggc ctgaaagttc cacatcaaat 25080
 ggacacttca acgcgctccc cacttacga ctcgtggaga catggcgagg ttttcaaaag 25140
 tcttaccgta gccacgtcgg gtaaaacaga aaacggagtg tccgtttcca aatagcatc 25200
 taaccgatcg gaggtgacag tagacccag ttgggcttta aacctctgc caccctcatc 25260
 ctccccattg gataatttgg aacgcgcat ttgtgaacat ataatcgccg tggtaactcc 25320
 attgaccgcc ggtcgcctaa agttaatgaa acgtgtaaat attatgcaaa atacggcaga 25380
 cccatataat gttattaaca cttatatca taacttaag ggggaaaaaa tggctcgcca 25440
 atacgcacgt atttttaaac agtttattcc tactocactc cactaaaca ctgtattaac 25500
 aaaatattgg aattaaaaca cacataagag cgacttaatg gttcattgtt ttattttgct 25560
 cgtatataca tgtataaat cgtttatcac tgtgcccga taagatgtac tgtgtctctc 25620
 aaaaaattt gtgtttttat ctgcaatcat aatgcaagt ggaagtcgg aatcgggagg 25680
 tggggtgta aatagttttg gtacattaat cgctgataaa agcctgtccg cgctgaattt 25740
 cacgtattgt gtaattgcat cgacgttcc caaacgggtt ttgggtgcat gggattttaa 25800
 aaacgcacac tcgatttcaa cggcttccga aaacagttga tgtattctgg tgatagcggg 25860
 tttttcgggt acatagttat tttatataca acacgatgcg ctggtatgta tggcttcatc 25920
 tcggttata aggtcgttaa attgacaagt tacaacaaat agtccgttat tgcgtaata 25980
 tgcaatagcc gcgaacgatg atacaaaaaa aatgccctct ataagaatca ttagtatata 26040
 tttttctgoa acgtaggggt tgtcccgtag cttttcttcc aaccattgta ctttttgttg 26100
 gatcgacgga ttattaatag tgacatttac gtattgtacc cgcaacgatt catcccctct 26160
 gaacaacatt agttgaattt gactatagac acgcgctggg acaacctcga tgcactcttg 26220
 ttcaatgtag taatggtgaa tatccttttg ggaaaagagt tgggtagag agcccaaat 26280
 aacatttacc agatcatctg ccgccgataa aatgtaaaa ataaatctgt agaatattag 26340
 ttcatcttcc gttaaacagt ccaagtattg ataactatct tcaatgataa aatcgctttc 26400
 taaccaacga ttcgaaatgc tcagggcagc taaattgttt atatctggac actccggcct 26460
 gtaaaaaaaa tgactgcaat cttctgtag cattttgaa tagtttcccg tgtaaattha 26520
 taaagcaca ctggtacag ttaattcgcc tcccgaac agtccgctgt tcgtagcttt 26580
 acgaatttta cagtagtaca taccgtttt aaggccgct ttataggcac gtataagcaa 26640
 attcattatt ttgaggcgg gaattgtccc gtctggcgt tcctcaataa ataaagtcatt 26700
 tgattgactt tggcaataa atggcgcct ttctgcacac atatcaacga gatcctcttg 26760
 ctcatattca aacgctgttt tatattttta gagtgggtga ctattagata aacagccaaa 26820
 cgaacgtatt actgaccatt ggtttttctc aagtatgttt ataactcca gtcgtttttc 26880
 ttcacatgaa tacatatctc ttagttcgtc cataaggtct aagttgggtc taagtaactc 26940
 acccgagggt gtgaccttac taaacatatt attataaatt ggagagaaac cctcactgca 27000
 ctccgttacc tgtgcagatg aaactgtggg cattaacgct aagaactgcg agttgtataa 27060

-continued

```

cccataagcg caaatatcat ctgcgagggt acaccatggt aaatctaaat aacttatcgt 27120
agaaaaccca tcttgggtga accatccctt agcatattta ctttcggtaa aacccttaaa 27180
cggggctaag ccgccaatct tacacatttc catgcttggt ttcattgtct catacaacat 27240
taactccgct atttgtacat ttaaccgtct agctggttgg gaagttaaata caaatcctaa 27300
gcgagacaaa gttgtatgta acccttgtat gccaatgcca agtgatcggg tgttttttac 27360
accctttacat gattttttac atggaaaagt cccagccgcc aggaccccg ttaaaaaaat 27420
aacagtcggt cttgctgtca attgaaggtc gtttaaatata aatgacactg ggcctttgga 27480
taagcacggt gtaagattta tgctggcaag attacatagc ccatgttgat gagcgtctgc 27540
cttttgaaca atttccgtac acaaatttga ccccgatgata gcatttcctt gggatttcat 27600
atgataatta cgattacagc catctttgaa cattaanaag gggcttcctg ttacagcagc 27660
actgctgatg attgtgaatg cgatatcttg aatgggaaca gaagaaacgc ctaatccttc 27720
tctctctaaa cgtaaatagg ttgaagtga tgcctcccg tgtaatgttc gaaggatatac 27780
ggctctgtta tcaaaaagag tccactgaac attactagcc ccttttagat agcttaggta 27840
tctttcaaaa aataaatctg ggtccataa acaacaaaat atgttatcac atcgaatat 27900
ttcatcacga accaacattc cacgtgtggc caaacagtt tgtagatcga cgtgccatgg 27960
ttctatgtaa acacaaactc cagttggtcg ttcacaatca ctgttaattg ccataacct 28020
gcaatctaaa agttttaaaa ctgcaagaag accttcctg tgattttccg taggtattaa 28080
attcagactc tgtagagaaa ttcccactcc acctcgactt tgtaataccg ttcccacatc 28140
gcctgtgata gctcgaacag ctctcccaac agtgatggat tccgggtcca ttaaataaca 28200
actgcccgtt gccccgtct ctgcacctaa aaacatcata accggtgtag ccgggacaat 28260
ttcttgacat gccaacgctg tgaaaaatac ccgacagaca tcagtccatg tataacctc 28320
atattatcgg ggaataagag ttgcatgttt aggcaggttt acgatttctg ttgtcacggt 28380
ggccgccagt cttaaaaaga attggcaaa cgactctaata ttaccttcct ctaacttagt 28440
taaaataaaa tctctgtact ttaaagcaga ctgtagtcca aggtagtca aagcggggta 28500
ttgatctttc aaaaacggtt ctaatatagc ccgacgaatt tcgtccctcc gcccttcaat 28560
tgcttgccg actcggggag ttaaacagag aattgggaa gtcaaccacg tttccatgga 28620
aacggatcgt aggttaatac ggcaatggat aagttctcca caacatcgg acaactcgtc 28680
atcttctcgc gtcaccgct taagttttga gacgatagtg ctaataact cattaattc 28740
caccggtgtg gttgattcgg gcggaatgat gtattcctg tagccatgtt gacataatcg 28800
gtttataatg tcatgaaccg tattaanaat tcttttgaac tccataacgg ataactgatt 28860
taggctccgg aataaacctt taaacctaa actcacagct gagttagttc tacaatattg 28920
tagactccct tatatatggt tacgtacagc ctgcccctcc ccagtatata atatcacgca 28980
aaaccacgc tatgttaaat tcagtttatt ttacatacat gctttaataa taacattcgt 29040
tccatgtatt tgtaccccc cacacaaccc cctctaacca aatagttggc acgttataac 29100
ctccgaacgg ttccatcgtt cttgtataac gcacagactc tgatggaatt gttccaatta 29160
acgtatatgc cgatacatg caggataatt gtgtgggaag tccccgaaa tcgccggtcc 29220
attgatacaa tcgctgtcta gccaagtcc aattactcc tgtaatttcg ccaatactac 29280
atcgagggtg tgcgggtca ttggataact gcacaagcgg caacgcctt gtgttatatg 29340
gctggtgggt atttgaacc ccttcagtc cccagcggc attttcagct cgtatgcgtc 29400
ctaacaggaa gccaatacca cgacaaaac attgttcgtt tagttggctt aatgcaagat 29460

```

-continued

gcaagtcttac accttctcgt tggcgtcgt gtgtatatac aaaaaccaag aacacatgct 29520
tcagtcgcgc cgcggaaga tgtaaatctt tgtcaacgtc caaaatacag caggccggga 29580
tgttggctgt gaccctgcga gttgaagttt tgtctgtacg tgcagcttct tggggacctt 29640
tggccacggc ggttatattg cataaattat cctgaatggt atattccagc agggacccaa 29700
aaaaacttat aaatcgatgt gaaatacat gacattgtac catcgacgct aaacactccg 29760
aaaaccttat gagccgcgtt tccatcacgac tgcacccata ggcagaaaca attgctgttc 29820
tgttggcatc cgctgcctgt ttatccgcat attcttctgc cggcatgcg gcgatgaaac 29880
ttaatgacgt tacatatgct ctaagcccc caccttctcc aacggccaa ggagccgtgc 29940
aggcattgaa taggttctgt aaacctcta gtagtacatc ggggtcacgt ccagcctgtg 30000
taagtgtatt agcttctcca atcatgtcag atggatgacg aaggattaag acgattgacc 30060
cagcatgctc aatgtccgga cgaaaaaat cggttaatga cacttgttgg attagctgtg 30120
tcgttgattt aaaattatct aacgggagtc taatggtaac ttgocgggta ccaattgaag 30180
ttggatttat ttgaatggtt tccatcacgac taataacaat tgaacggggg gttacttgaa 30240
tagacgcggt ttctgtacgt tttggtggt catgtatcgg ttgtttgttc agacctcaa 30300
agcgagggcc aattgttaaa tcgcgactcc aattccgaa gaagccgga gcataagtca 30360
tatgaagccc gttccctatt tgaataaaac ggttatttcc taaagactg atattagttc 30420
cacatagcgt ttgtctgctt aaagtaaaat gcgagttggt tggttgactc cccatagctg 30480
aggggttaaa ttcacacaat gcaatcgtga cgtggtacta tctgaaatgt tgcctggggt 30540
atgtgtacac attatacagt cgtagtaccg tttatataat gttaggtagg aggagcctat 30600
aaaaatattt tgattggcgt taaaaggttc ttcaacttac cgtgacgtcc tttttattaa 30660
catgcgtttt tattgatggt acatttatgt ctttctcttc cggacggatg tagctttttc 30720
atatcacggt ataaagttaa gtcagcgtag aatataccat ggaagaacca atttgttatg 30780
atacacaaa acttttggat gatttaagta acttgaagt acaagaagcg gacaacgaaa 30840
gaccatggtc accagagaaa acagaaatcg ccagagttaa ggtagttaa tttttacgat 30900
ctaccagaaa aattccagct aacatttta ttcagatag ggaaccctg cattctaata 30960
tctgttttgt atattccaat acatttttgg cggaggctgc tttcaccggc gaaaatttac 31020
ccgactggtt gttttggaga ctagatctag actggacgat agaggagcca ggtaatagct 31080
taaaaatttt aaccagccta tcaagtgtag tacaagattc cgagacgta catcgtttat 31140
cggccaataa attacgaacc tcgtctaaat ttggaccctg ttcgatacac ttcattataa 31200
cggactggat aaatatgtac gaggtcgcct taaaggatgc aacaacagcc attgaatcac 31260
cattcactca cgctcgtatt ggaatgttg aaagcccat tgcagcttta acacaacata 31320
aatttgcat catttacgat atgccattg ttcaagagg gattcgtgtt ttaacacaat 31380
atgcaggatg gcttctccg ttaatgtta tgggaaatca gattcaaat agctcactca 31440
ctcctctaac acgagccctt tttataatct gtatgattga tgaatatctc acggaacgc 31500
cagtacatag catatcagaa ttatttgcag atactgtaaa ttaattaaa gatgaggcgt 31560
tcgtatccat cgaagaagcg gtaacgaatc cacgaacggt gcacaggtca cgaatttctc 31620
cagctctggc ttatcgagac cttatgttt ttgagacatc cccgggaatg cttgctagga 31680
gacttagatt agacaatggt atatgggaaa gcaacctctt atcgttctcc accccggaa 31740
ttcatattga ggcgctgta catttactaa actccgacc ggaagcgaa accacatctg 31800

-continued

gaagtaatgt	agcagaacac	acccgtggca	tttgggaaaa	ggttcaggct	agtacatcgc	31860
ctagtatggt	aataagcacc	cttgccgaat	cggggtttac	aagattttca	tgcaaattgc	31920
tacgtcgggt	tattgctcac	cacacactcg	cgggttttat	tcacggaagc	gttgtagcag	31980
acgagcatat	tacagatttc	caacaaacac	taggatgtct	cgcttttagt	ggtggactgg	32040
cataccaatt	agtggaacg	tacgctccta	ctaccgagta	tgtgttaaca	tatacacgga	32100
cagtaaacga	gaccgaaaa	cggtatgaaa	cgctattacc	cgcttagga	ttaccaccgg	32160
gaggcctggg	acaattatg	cggcgtggt	ttgctccacg	acccttatt	gaaagtatac	32220
aagcgacacg	cgtaatacta	cttaatgaaa	ttcacatgc	agaagctaga	gagacaacat	32280
atthtaagca	aacacataat	caatcctcag	gtgcgttatt	accacaagca	ggacaaagtg	32340
ccgtacgcga	agccgtacta	acctggttg	acctacgtat	ggattcaaga	tgggttatta	32400
ctcccccggt	ggatgtgggt	atgacaccto	ctatttgtgt	tgatccaccg	gctacagggt	32460
tggaagctgt	catgataaca	gaagcactaa	agattgcata	tcctaccgaa	tataatcgct	32520
ctagcgtggt	tgtggaaccg	tcgtttgtgc	cttatattat	tgcaacaagc	acgcttgatg	32580
ccctttcggc	aacaatagct	ttgtcttttg	atacacgggg	aatacagcaa	gccttgccta	32640
ttcttcagtg	ggctcgcgat	tatggatccg	gaaccgtgcc	caatgcagat	ggatatcgca	32700
caaaactatc	tgctcttata	acaatattag	aaccttttac	ccgtacacac	ccccagtac	32760
ttttaccatc	tcacgtttct	actatagatt	cccttataat	cgaacttcat	cggactgttg	32820
gcattgcctg	tgacctgctt	ccccagcacg	tccgtccttt	ggttcctgac	cgctcttcta	32880
ttacaaatag	cgttttttta	gaaactctct	attatgatga	actttacggt	cggtggacc	32940
gactggataa	aacatcgcag	gcgttggttg	aaaattttac	atccaacgcg	ttagtggttt	33000
ctcggatcat	gttaatgtta	caaaaatttt	ttgcgtgtcg	tttttatcca	acgccagatc	33060
ttcaggctgt	tggtatctgt	aaccacaaag	ttgaacgcga	tgaacaattt	gggtatggc	33120
gtttaaacga	tctgtctgat	gcggttggtc	atattgttgg	gacaatacaa	ggaatccgaa	33180
cgcaaatgag	agtgggaata	tccagcctgc	gcacaattat	ggccgatgct	tcctcagccc	33240
ttagggaaat	tgaaaattta	atgactaaaa	cctccacttc	tgctattggg	cctctttttt	33300
caacgatggc	ttcccggtat	gcacggttta	cacaggatca	aatggacatt	ttaatgcgtg	33360
ttgacaaact	aacaacagga	gaaaatatac	ccggtcttgc	aatgtagag	atthttttta	33420
ataggtggga	acgaatagca	acagcttgta	ggcatgccac	ggcagtccc	tcggccgaat	33480
ctattgcaac	cggtgtgta	gaattgagc	gcggtttaaa	aaatatacaa	gaggatcgtg	33540
taaatgcccc	aacctcatat	atgagtcacg	ccgaaatct	ggaagatcac	aaggcagcag	33600
tttcattcgt	tatggactcc	aggcaacagt	ttattgtgga	ttctggacct	cagatgggcg	33660
cggttttaac	ttcacaatgt	aatataggaa	catgggagaa	tgtaaagca	acgtttttac	33720
atgataatgt	taaaataacg	acaacggtca	gagacgta	ttcagaggct	ccgacgctga	33780
taataggaca	aagatggctt	cgccagatg	agatthttat	taatgtagat	ttgcgtcttg	33840
gcgtaccggg	gaatacaagt	gggagtgacc	cttaataata	aacaggcgtg	tttatgtaca	33900
ttaaagtatt	tgtggttttt	attgactggg	cgtttcgttt	gtataacgct	gttgttgcta	33960
gtatthttat	aacctcctag	gtttttggag	ctacacgtgc	ttattcaacg	ctctttggga	34020
tttgaatcat	cgtaaacgta	gcgtccctac	cagttgagcg	cgtaatthtc	gtaagcaata	34080
aatggatat	aattccgcct	atagctgtca	ctggtgcggg	agtgggaagc	cgtaatcaat	34140
ttgacggtgc	cctgggaccg	gcgtcaggtc	tgctatgttt	aagaacatct	ttatcgtttt	34200

-continued

tgcatatgac	atatgcgcat	ggaattaatg	caaccctgtc	atcagacatg	attgatggat	34260
gtttacaaga	gggtgcagca	tggactacgg	atctgtctaa	tatggggagg	ggtgtccag	34320
atatgtgtgc	tcttgttgat	ctcccacatc	gaatttcata	tattaaactg	ggggacacta	34380
ccagtacgtg	ctgcgttttg	tctagaatat	acggcgatag	ccattttttt	accgttccag	34440
acgagggttt	tatgtgcaca	caaattcccg	ctagagcgtt	tttcgatgat	gtgtggatgg	34500
gacgtgaaga	gtcgtataca	attataactg	tagactcaac	gggaatggcc	atctatcgtc	34560
agggaaacat	atcttttatt	tttgatccac	atggccatgg	gactatagga	caggctgtag	34620
ttgttcgggt	gaataccacg	gatgtgtact	cttatatcgc	atcggagtat	accacccgcc	34680
ccgataacgt	agaatcccaa	tgggcccgtg	cattagtttt	ttttgtcacc	gcaaacgacg	34740
gtcccgtaag	cgaagaagcg	ctatcttcgg	cagtaacgct	tatatacggg	agctgtgata	34800
catattttac	agatgaacaa	tattgcgaaa	aactggttac	agctcaacat	ccgttgcttc	34860
tttcacctoc	taattccacg	acaattgtgc	ttaataaatc	gtctatagta	cctcttcacc	34920
aaaacgttgg	tgaaagtgtg	tcottggaag	caaccctaca	ttcaacgtta	accacacggy	34980
ttgcaactgg	ccctagatgt	agttacacgg	aggttgatcc	ttggcatcgc	gttctagaaa	35040
caacctcgac	tgggtctggc	gttttgattg	gtcgtcgtag	acgccctcct	tcatggactc	35100
ctccttcaag	cgaggaaaat	ttagcttcta	tcgacgatgg	cttggtaaat	aatacacatt	35160
ccacggataa	ttacataaaa	cccgctaaaa	aggttctcaa	atttaaacca	actgtagacg	35220
tgccggataa	aacacaagtg	gcacatgtat	taccccgctt	acgagaagtt	gctaacaccc	35280
cagacgttgt	gttaaatgta	tccaatgtag	atacgcctga	atccagtcct	actttttcac	35340
ggaacatgaa	tgtaggaagc	agtttgaaa	atcgggaagc	atttctattt	gaacagagtg	35400
gtgatgtcaa	catggttgtc	gaaaaactac	tacaacatgg	gcatgaaatt	agcaatggat	35460
acgtacaaaa	tgcggtgggt	acgttgata	ctgttattac	cggtcataca	aatgttccca	35520
tttgggtaac	aaggcccttg	gttatgccag	acgaaaagga	tccattggag	ctttttatta	35580
acctcaccat	tttgcgttta	acgggatttg	tgggtgaaaa	tggaacacgt	acacatcatg	35640
gtgctacaag	cgttgatata	gactttatag	gtccccttgg	ggaaatttta	acagattttc	35700
cctccgcccg	ggaacttata	cgcgttacaa	gtttgatatt	aacaaacatg	ccggggggcgg	35760
aatatgctat	taaaactggt	ctccgaaaaa	aatgtacaat	tggcatgctc	attatcgcta	35820
agtttggctc	agttgccatg	cgggttcagg	atacaaccgg	cgttttacat	gccgaactag	35880
atgtgttaga	agcggatcta	ggaggttcgt	cgcccataga	cctctattct	agactgtcga	35940
caggtcttat	aagtatacta	aattcgccta	ttattttcca	tcccggactt	tttgccgagc	36000
ttattccaac	ccgtacaggg	tccctgtctg	aacgaatacg	tcttctttgt	gaattagtct	36060
cggcccgga	gacacgctat	atgctgtaac	acaccgcgct	tgtttctagt	gtaaaggctt	36120
tagagaatgc	attacgtgct	acccgcaata	aaattgatgc	cattcaaata	ccagaagttc	36180
cccaggaaac	cccggaagaa	accgacatc	caccggaaga	gttaattcgg	cgtgtatatg	36240
agatacgcac	cgaagttaca	atgctattga	cctcggctgt	tacagaatac	ttcaccgcg	36300
gagtggtata	tagcacacgg	gcottgatcg	ctgaacaatc	ccctaggcgt	tttcgggtcg	36360
cgaccgcaag	tacggcacc	attcaacg	ttttagattc	tcttccgga	ttcgaogcta	36420
aattaacggc	aatcatatcg	tccctgtcta	tacaccctcc	tcctgagact	atacaaaatc	36480
tcccgtcgt	atctctgtta	aaagagctta	ttaaagaagg	ggaagattta	aacacagaca	36540

-continued

cggtctcgt	atcgtggtta	tctgtagtgc	gggaagctca	aaccgcaggt	tacttatcca	36600
gacgagagtt	cgatgaatta	tcacgtacaa	ttaaaccat	taatacacgc	gcaacgcaac	36660
gggcttccgc	ggaagcagag	ttgtcttgct	ttaatacgt	aagcgcggcc	gtagaccaag	36720
ccgtaaagga	ctatgaaaca	tataacaatg	gtgaggtcaa	gtatcctgaa	ataacacggg	36780
atgatttatt	agcaacaatt	gtacgtgcta	cagacgattt	ggtgcgacag	ataaaaattt	36840
taagtgatcc	aatgatccaa	tccggtttac	aaccttcgat	taaaagacga	ttggaacaa	36900
ggcttaaaga	ggttcagacg	tatgcaaacg	aggcccgaac	cacacaggac	acaataaaga	36960
gtcgaaaaca	ggcggcatat	aataaactcg	gggggttact	tcgcccggtg	accggttttg	37020
tgggacttag	ggctgcagta	gattttattac	cggaacttgc	ttctgagtta	gatgtccaag	37080
gagccctggt	aaatctcag	accaaagtct	tagaggcgcc	ggtagagatc	cgttctcaac	37140
ttacgggtga	tttctggcgc	ttatttaacc	aatatcgaga	cattttagaa	catcccggaa	37200
acgcacgcac	atctgtctta	ggaggactgg	gagcttgttt	tacagctatt	atcgaaattg	37260
tgccgatacc	tacggagtat	agaccatcat	tgcttgctgt	ttttggtgac	gtggcagatg	37320
tgcttgcatc	cgacatcgcg	accgatctca	ctaaccggga	aagtgagtcc	gccataaacg	37380
ctgttgttgc	aaactcttagt	aaagcgacgt	tagtttcatc	tacagtgcc	gccttatcct	37440
ttgtgttgc	gttatataaa	aaatatcagg	ctttacaaca	agaaattacg	aataccata	37500
agttgactga	attacaaaa	caacttgag	atgacttctc	caccctagct	gtctcatctg	37560
gacacttgaa	gtttatatca	tcttcaaatg	tagatgatta	tgaataaac	gatgcgatat	37620
tatcaataca	aaacaatgtg	caagccctaa	tggatacgg	taaacttgtt	gaagtgaac	37680
tgcaaaagct	accccccat	tgatttgctg	ggacatctac	cttatctcga	gtagtaaagg	37740
atcttcataa	actcgtcaca	atggcacatg	agaagaagga	acaggcaaaa	gtgttaatta	37800
ccgattgtga	acgtgcatat	aaacaacaaa	cgactcgggt	tttgtatgag	cgttggacac	37860
gtgatattat	agcatgtctg	gaggcaatgg	aaacgcgcca	tatatttaac	gggacagaac	37920
tggcacggtt	gcgagatatg	gccgctgcgg	gagggtttga	tatacacgca	gtttaccac	37980
aagcacgtca	ggttgtagcg	gcatgtgaaa	ctacagccgt	tacggcatta	gatactgtgt	38040
ttccccacaa	tccatatacc	cccgaataa	caaatattcc	accacctttg	gctttgttaa	38100
gagggttaac	atggtttgat	gatttttctg	ttacggctcc	cgtattcacc	gttatgtttc	38160
caggtgttag	tattgagggg	ctccttctgc	ttatgcgtat	tcgcgcggtt	gtgttattat	38220
ccgccgatac	gtctattaat	ggaataccta	actaccgaga	tatgatatta	cgaacctcgg	38280
gggatctatt	acaaatacc	gcattggctg	ggtatgttga	tttttacaca	cggtcttatg	38340
atcagtttat	aaccgaaagt	gtaacgttaa	gtgaacttag	agcagacatc	agacaggctg	38400
ccggggctaa	acttacagaa	gcaaataagg	ctttggagga	agtaactcat	gttcgggcac	38460
acgaaacggc	taaacttgca	cttaaagaag	gtgtcttcat	tacattacca	agcgaaggtt	38520
tattgattcg	ggctatagag	tattttacaa	ctttcgatca	taaacgattt	ataggaacgg	38580
catatgaaa	agttttacaa	acaatggtag	accgcgatct	aaaggaggcc	aacgcagagc	38640
ttgcacagtt	tcgtatggtg	tgacagcga	caaagaaccg	tgcaatacaa	attttacaaa	38700
acattgttga	tacggccaat	gccactgagc	aacaagaaga	cgtggatttc	actaacctga	38760
agacgttatt	aaaactaacc	ccccctcca	aaacaattgc	attggccatt	gatagatcta	38820
cttccgttoa	ggacattgtc	acgcagtttg	cattgctgtt	agggcgtctg	gaagaagaaa	38880
ctggtacggt	ggacattcag	gcgggtgact	ggatgtacca	agctcgcaat	attattgact	38940

-continued

```

cccatccact aagtgtcgt atagacggta cggccccct gcatacttat aaagataggg 39000
tggataaaact ttatgcgtta cgaactaaat tagatctcct acgacgacga atagaaccg 39060
gtgaggttac gtgggacgat gcatggacaa catttaaaag agaaacgggg gatatgttg 39120
catcggggga cacgtacgct acttccgtag atagtataaa ggcaactccag gcatcggcgt 39180
ctgtggttga catgctttgt tccgaaccg aatTTTTTTT attgctctg gaaacgaaaa 39240
accgtctcca aaaaaagcaa caggaacgta aaacggcgtt ggatgttgtg ttgcaaaaac 39300
aaagacagtt tgaagagacc gcgtctcgt tacgagcttt aattgaacgt attccaacgg 39360
agagtaccac tgacgttctt cgtatgttat tacgtgattt cgatcaattt acacatttgc 39420
ctatatggat aaaaacacag tatatgacat ttcgaaattt actcatggta cggttaggct 39480
tgtatgcaag ttatgctgag atttttccac cgcgctctcc aaacggagta ttgctccta 39540
ttcccccat gtcgggtgta tgtctagaag accaatccc atgcattcgc gcgcggttg 39600
cgcgctttat gggggaggcg tctgtggtgc aaacgtttag ggaagccaga tcttctatag 39660
acgctttgtt tggaaaaat ttaacctttt acttggtac tgatggggtt ccacttcgat 39720
atagagtgtg ttataaatca gttggggtta aacttgaac catgctatgc agtcaggggtg 39780
gattatcttt acgaccggca cttcccgatg aaggtattgt ggaagaaact acatatcgg 39840
cattacgcgt ggccaatgag gtcaatgagc tacgcattga atacgaatcc gctataaaat 39900
cggggttttc tgccttttcc acctttgta ggcatcgcca cgcogaatgg ggtaaaacca 39960
acgcacgcag agccattgca gagatatacg cggccttat aacaacaaca ttgacacgac 40020
aatacggggt tcattgggac aagcttattt attcttttga aaaacaccac ctaacttctg 40080
taatgggcaa tggactaaact aaaccaatcc agagaagggg tgatgtacgc gtattagagt 40140
taacctatc tgatattgta actattttgg ttgccacaac cccgtacat cttctcaatt 40200
ttgctagatt ggatttaatt aaacagcatg agtatatggc cgtaccctc agaccgtaa 40260
tcgagggccg atttagaggt cgtttactcg ttcgctcatt ggatggagac ccgaaaggca 40320
atgcccgggc cttttttaat gccgccccat ccaaacataa actcccgtta gctcttgat 40380
caaaaccaaga tcctaccggc gggagaatat ttgcatttcg gatggcagat tggaaacttg 40440
ttaaaatgac acagaaaata acggatcctt ttgcgccatg gcaactttcc ccccccccg 40500
gggtaaaggc caatgtcgtat gcagttaccc gtataatggc aacagatcgt cttgcgacca 40560
ttactgtaact tgggcgcatg tgtctccgc caatttcctt agtgtcaatg tggatacgc 40620
tgcaaccgga ggaattcgca tacagaacac aagatgatgt ggacattata gttgatgcga 40680
gactggattt gtatccacg cttaatgcaa gatttgatac cgtcccagc aataccacgt 40740
tagagtggaa tacagaccgt aaagtaatta cagatgctta tattcaaacc ggggcaacga 40800
cagtttttac agtaacgggg gcggcaccaa ctacgcttc taatgtaaca gcgtttgaca 40860
tagcaactac ggctatttta ttggggctc ctttggttat tgccatggaa cttacatccg 40920
ttttttcaca aaattccgga cttacttttg ggttaaaatt attcgattcc cggcatatgg 40980
ctacagattc gggatatatc tcagccgat ctccgatat tgtttcttg gggttacgtt 41040
tactgcatat ggatcctcac ccaattgaaa atgcatgttt aattgtocaa ctagaaaaac 41100
tgtccgcgct cattgcaaac aaacctctta caaacaatcc cccgtgttta ctgctattgg 41160
acgaacatat gaatccctct tatgttttat gggaacgaaa agactcgatt ccagctccgg 41220
attatgtggt cttttgggg ccagaatctc ttattgattt gccgtacatc gactccgatg 41280

```

-continued

aggactcttt cccctcgtgt cccgatgac cattttactc gcaaattatt gccggttatg 41340
 cgccccaaagg cccccaaac ctogacacaa ctgattttta cccaacggag ccaactattta 41400
 agtctcccgt tcaagttggt agaagttcca aatgtaaaaa aatgcccgtc cggcccgcgc 41460
 agccccgcga gccccgcgag cccgcgcagc cgcgcgagac cgtccagccc gcgcagccca 41520
 tagaacccggg cacacaaata gtggtacaaa attttaagaa accccaaagc gtaaaaaaa 41580
 cccttagcca aaaagatatt cccttgatg tggaaaccga atcagaaacg gctgtgctta 41640
 tacctaagca attaaccacc tccattaaaa caaccgttg taaaagtatt accccaccaa 41700
 ataaccaatt gtcggattgg aaaaataatc cacagcaaaa ccaaactgta aaccaagcgt 41760
 tcagtaaacc aatacttgag attacctcca ttccgacaga tgactcgata tcttaccgga 41820
 cttggattga aaaaatcaat caaacacaaa aacggcatca aatgaccct cgaatgtata 41880
 actccaaaac agtattccac cctgtaaata accaattacc ttcttgggtt gacacggcag 41940
 ccgatgcccc ccaaactggac ctattgacaa actataaaac aagacagccg tcgccaaaact 42000
 ttccgcggga cgtacacaca tggggcgtat cttctaacc gtttaactca ccgaacagag 42060
 acctatacca aagtgtttt agtgaacct ctgacggcta tagcagtgag agtgaaaatt 42120
 ctatcgtact aagtctcgac gaacatcggc catgtcgcgt tcctaggcac gtacgcgttg 42180
 ttaatgccga tgtagtacc ggtcgacgtt atgtccgag gacgccttg ggagcactgg 42240
 cactgttaag ccaggcatgt cggcgtatga tcgacaactg tagatataca cgtaaacttt 42300
 taatggacca cacggaagat atatttcaag gcctggggta tgttaaattg ttattagatg 42360
 gaacatatat ataaagtagc gcctattaaa gaaaaaaaa aaacaacgat tttttctgt 42420
 gtatttttat ttacacccca cgaactcttg aagcgtttcc agattgtccc gtgtgtgaca 42480
 aggtctgtcc cttaccccc tgggggggat tttgggttg gggcggggta gactgtggca 42540
 cgccttgggc cgcgggcggt gatccggttg ttggctggac agtgcctgac tgtgctccct 42600
 gttgcggtt ttgtccagaa gaccocgaca ccactgttg ctgttgcca acggatgccg 42660
 acgtcgttg aggtgggggg tgttcgggg atgatccga aaacgccaac gcggcgggct 42720
 gttgtaaagc agactgatgc gcgctctgtg ttttttgcg caatatagta gccccgaga 42780
 ttccaaaact catggatgga tttgggggtt gtgctgcat aatagcggg ttaaacgtac 42840
 gtttaagcc aaccgttggc cttaacctag tcatagggc agtctcgca aacatggccg 42900
 ttccgctgat cgtatttga ttatggttag cgcgtgcac gcggcactg gccgcggctc 42960
 ccacggtgta aatgctctg gcatcagcga tgtccacacg gtgaccaggt tgcaaagtc 43020
 cactggcgtt taaaagtcgt attaaagcaa cgggggtgta agccgcaatt gcttccaccg 43080
 aaaaatggtt ggggttgctg ggatcaaaga ctacacgaga cgatcgggt tgtgtcatcg 43140
 tttattagtt tacgggacaa tcgataaacg catacacgta catctgcgca ggatatttac 43200
 ggaaaggcaa tttatttcca gaaaagcacc gccctaata caactaccag tacaattaca 43260
 atgaacaggc catatgtcac gttagctac ggtagagcaa gttccagac acgcgtagtt 43320
 tgggtatcgg gtaacgcagg ttaaatgtca ctttgcatth gaacagacgt gtttgactt 43380
 ccgttctcgg gtgggatct gaatgaagc cgcagcgta tatattcatc caaattattg 43440
 ccagtttctt tatacatgta tgcacccgtg gcgcgggcca taagttaat ggtgcgagat 43500
 ggatcttcgg gtcccataaa acgaaaggat aactgaacat atggcattcg cacaaagcag 43560
 ttcaccaca ttaaagcctg gagaggtcgg cggtaatac cccacctcg ttaattgat 43620
 tccaaagcag ataggttgat accggtactt aacgttgaac taagaatcac gttattactg 43680

-continued

tcaatggaca cttcagccac tgggtcgtta gtcggacgaa aaaaaaac ttgaaatagc 43740
acagacaccc ccgtattttg aatttttatg taagggtcac aatctacttg cgccaattc 43800
gccattaaac gcataatata ctctaccgga aaggcttcgg atacgttgtc ttcgccgta 43860
aactgaaaaa cacaacgggc gggggggcgt tgtggatcaa atattggaag atccccatcg 43920
caacattgaa gagcgttgg taccaccaac cgaatacgtt gtaaaagatt atctccgcaa 43980
cccctcctgc gttcactcgg tacatacgtt ctccgtgaca tattgatcta aggttgcaaa 44040
ccaaggcaca cgcgtgaagt atttagacca tttatcgtgg gatataggag gagtttgag 44100
tgatccacc cctgacgact tattaatgcy tttattttcc ccatgtatta agcatccttc 44160
aatatttcat gcaaatctag aaatttggcc atgactcccg caaagcgttc acggcgacgg 44220
gtcacgctgg cactatgttc acatggaaca acataagcag atttttctga atcgttactt 44280
tctttatgtt ttaaaccgga cgcagcgca ctggtaaag atataaatt taattgagcg 44340
tcagttgtag gtagaattgc ttctatttcc gggggaatta aattttcaaa ccaaaccgaa 44400
agagtaaagg tgctatcagc aggaaaatac tttgactcca gtgcacgat atttaataga 44460
ttaacatcgg tgctgtaat taaatcgcgg gccctcatcc cagagatgga tcgggtagaa 44520
tcagaagaac ccatgtagg attcgaatcg cccgtattct ccgaaaatac atcttotaat 44580
tccggatggt gttccgacgc attttccgat tcgtacatcg cttataatcc agcccttctg 44640
ctaaaaacg atttgttatt ttcagaattg ttatttgcct cccacttaat aatgttccc 44700
cgtgcaatag aaaacaacgt cacttatgag gcctcttcgg cggtaggtgt ggataatgaa 44760
atgacctcaa gtaccactga atttatagaa gaaattggag acgttttggc gttagacaga 44820
gcctgtttgg tctgcagaac gcttgatttg tataaacgta aatttggact gacaccgaa 44880
tgggttgccg actacgccat gttatgatg aaaagtctgg catccccgcc ctgtgcagtt 44940
gtcaacttta gcgctgcctt tgaatttgg tatcttatgg atcgttacta cctgtgccgt 45000
tataacgta ctttggttgg gtcctttgcc aggcgcacgc tttccctgtt agatatacaa 45060
agacattttt ttttgcattg atgttttctg accgatggag ggttaccag tatacagccg 45120
ccccccgta aggaaatggc caacaaagta agatattcca attactcctt ttttgtacag 45180
gcggtagtta gggctgcatt actatcgatc agcacgtctc gtttagacga aaccgaaacg 45240
cgtaagtcat tttactttaa tcaggacgga ctgactggag gccctcaacc tttagcggcc 45300
gccttggtta attgaaaga ttgcccgcgg atggttgact gttcatcadc ggaacatcgc 45360
acaagtggga tgattacctg cgcggaacgt gcattaaaag aggatataga gtttgaagat 45420
atattaatag acaaaactaa aaaatcgtct tacgtagaag cagcttgggg ttacgcagac 45480
ttggctttat tattactgag tggggttctg acttggaaag tagacgagcg taaaaattgt 45540
gctatagaaa ctgcgcttgg atgtgttaaa tcatactggc aggcgaaccg gattgaaac 45600
tccagggagc ttccaaaaca attttccaaa tttacgagcg aggatgcctg tcccgaagta 45660
gcatttgggc ctattttgtt aactaccta aaaaacgcaa agtgccgtgg tcgcaogaat 45720
accgaatgca tgttatgttg tttattaacc atagggcact attggatcgc tttgcggcag 45780
tttaaaagg atataatagc atactcagca aataacaaa gtttatttga ctgtatcga 45840
cctgtaatoa atgcatggag cctagataac cccattaac ttaaatttcc atttaatgat 45900
gagggtcgat tcataaccat tgtaaaagca gcaggttccg aggcgtata taaacattta 45960
ttttgcgac tcctatgccc tctctcgaa ttacagaaa accctaaaat tttatttgc 46020

-continued

catcctacaa ccgcgataa ggaagtgtg gagttatata aagcccaact ggctgcacaa 46080
aacagatttg aaggctgtgt atgtgctggc ctgtggacat tggcgtatgc atttaaagcc 46140
taccagattt ttccacgcaa accaaccgcc aatgccgat tcatacgaga tggaggactt 46200
atgcttcgac gacatgcaat atcgctggc tccctcgaac acaccctatc gaagtatgtc 46260
taggcgatat aaatccgtat ctcgagcgg gccttcgatg cgtgtacgct ccagaacgcc 46320
atgcccgcgt caaacattc gaggaaaact tatgtcaaag gagcggctctg tgtacogcca 46380
ttattttaat tacatcgcaa ggtcccccc agaagaacta gctaccgta gaggcttaat 46440
cgtgccaaat attaagacga ccctgtcac ccttcgctt aacttgggtc agacagtggc 46500
ggataactgc ctgtcgttat ccggaatgg ttatcattta ggtctcggag gttattgtcc 46560
gacatgcaat gcatctggag aaccgcgtct atgtcgaacc gatcggggcg ctctgatact 46620
agcatatggt cagcagctta acaacatata cgaatatcgt gtgtttcttg catccatttt 46680
ggcgctatca gaccgagcca acatgcaagc agcgtccgct gaaccctat tgcgcagcgt 46740
attggcacia ccgaattat tttttatgta tcatattatg agggaggggg gcatcgagaga 46800
tatacgcgta cttttttatc gtgatggaga tgcgggaggg tttatgatgt atgttatatt 46860
tccggggaaa tctgttcacc tccattacag actaatcgt catatacag ccgcgtgtcg 46920
ggggtataaa atagtcgcac acgtttgga gacaacattt ttaactgctg tatgtcga 46980
cccagaacia caaacagaga ctgtggtgcc atccattgga acatcggacg tttactgtaa 47040
aatgtgtgac ctttaacttg atggagaatt gcttttgaa tacaaaagac tctacgcatt 47100
atgtgatgac tttgttcctc ctogtgatt tcagcttcag tgttcatttt attatcccag 47160
cacggggcgt gtatacaaac aaagcctgcc gcctgcaagc ggttagcat tttaacgta 47220
acaactcgtg tctctggaat aaaacgcttt aaaagcgtt ctgtgagttt agtgcgttt 47280
ccaaataacg ccttaaaagt tacactcgc gtcccaatga gatgagaaa ataatagtca 47340
atgtttaaag acagcccgtg tgatgttac tgaatgggat cttccgctaa gtcagatatt 47400
attaacttac gctttgcttc cccacaccgt ttactcgcgg tattctgtaa aggatctcca 47460
cgtagcaaac ctacactttt tgcacagcc tccactcgt ctgtgggggc cacaataaca 47520
taagggatgc gttctcgaac gtttgggatt tgaccctgct tcattactaa tttataatat 47580
actgttaagt gagccaagcg acggtttatg taggcggatg gtggacgact aagctcggcc 47640
gtcataacia acttattaat atccaatttg ggtgatgtaa tctggcgatg tgcactgca 47700
attatgcgct caaacccggc catcccagac ggcatggccc gtctattcca ttcagcaatg 47760
gaaacacag acgcctccgc cgcagcacgc gagacggtgt cgtcatataa caacagttct 47820
acaagtttg ggcataaat gtaataaat tgacagttgt tttttctaac caagtogact 47880
cccttcatta aaacctttcc gccgtaaat accccaatgt actttttctt tgtataaagc 47940
aaaagttta taaaagttt ttcacactcc aactttatag gaggacaaa cagagccgtt 48000
gaaattatat gtgccatttt ctgcgcgatt ttagctatcc cctcaacact aacacccttg 48060
aatcgataa acacagaatc cgtatctcca tatataacct ttacctgta cgctttttgg 48120
gagagaacgc tactttcaat gtctggaac gctgtaataa aacgttcaa tgcggcccag 48180
ttattatgaa tataatctct ggtacttaat aacatttgac ggccaattgt agtgacagt 48240
gccgctacgt ataacatg cagaaatccc tgcgcaactc cagtaaaacc gtacacggaa 48300
ttacaaacta cttttatcgc ggcttgtgt ttgtctaata acaactgctt atctgaagaa 48360
cttccgggta tgcgcgctct aatagccttg cgcatagcca accagcttt taaaagaaca 48420

-continued

```

cccagcagac tttctcgaac gttagagcgc acaaaaaaaaa gacgttttcc tccaactgta 48480
aaggtggcat aatcggatgg attcaaacgt ttaaccgtct caaaatttaa cgtagcgtg 48540
gtaaacata agttatgggc ctgaattata cttggatata aacttgcaaa atccaatacg 48600
accaccggat cgatataaaa tcccgtatca gggtcacaaa ccctggctcc tttatatcct 48660
acatttcgcc cacttgacgt accagtgga gaaacgctct cgtcttcac catctctcc 48720
tcaacatccc cgacatcggg aataacatcc ttatattcaa aagtagctgg gtatcccca 48780
tcgggtaaaa taaatctcgc agacgaagcc agtcctaata aacaggtgta aatcctaacc 48840
tgctgtccgc cgtaaatagc cttggttaaa gtaattctag ctagccttgc aaccgggat 48900
aactcaaggt gtggtaata tttaaaaa agtttcccca caagagccga gtcttgata 48960
caatattcac caataattcc tcgtgtatcc ggtccactag cgtaatatcc cggaatgtct 49020
ttgtaggga aatctctctt ggactcattt agagcttcac gtgcaaccga atctaattta 49080
taactcgaga gttttaattt ttcagttgca attgcataca tatccagaga tatgagaccg 49140
ttgatcttta ccttgcttcg tcgctgaaat ccggatttgc caacatocca tatcttaaac 49200
agacccccac ggtttatact gccataacca tcaagcttga gactgtatat agaattaagt 49260
ttctccataa taaacgcccc atcaaaaata acaatgttat aacctgtgac aaactogga 49320
gcgtactggt ttacgagggc cataaatgca attaatagct cgaattcact atcaaacctc 49380
agcacagctg gctccggtaa ccccgctcc ttcatttctt gtacatacct ttgtggaag 49440
tcacaagagc caagggaaaa cagtaaaatg tgttctaaag actgtcgagg gattgaatat 49500
aatagacaag aaatttggat tacaagatcc tccagatgty ttgcatcggg aaacgccagc 49560
tcattagatc ctctgattt acattcaata tcgaaacata acaacttga gtcaggccat 49620
gagtcacgtg ttggtatagc ctgcagatta tccgacatgc agtcaattc aacgtcgtt 49680
aacgttaatt ggcgacttgc cggtcgaact cgaacacgtt ccccatcaac tccaggttt 49740
agttgatacc aacaaaaact acaaaagccg ggattatcca ttagaaaacg agtggtagcg 49800
tctaccgcac cttcatactt tttcaactcc gggtgaaagt tatcacaag ataatttga 49860
aatthagatg agggagaata caccctgtaa aacgcacatg gctgtgtatc gtagtaataa 49920
acatctgtgc gctcaataac ctcaacgcga aagctttctg gagatgcgct ttaaacgag 49980
gtaccatgaa aagcgttctt gtctccattt aacgttgcat cttttgtgt tatcatagaa 50040
ctgcgtaaac actcggcaag taatacagat aactcgctac cggaacgat gccacaagcg 50100
gtatccacct cggctttggt tatataaaaa tattgacaga tgccgtatac atgaaactgcc 50160
acccttttcc cacatcggga catgccaagt aaagtaataa cggtagcaag cggctgtgtt 50220
gcagttgcaa accgggatac atctccatta gacgcggctt ctgttgttcc gacaatatca 50280
tatacatgga atgtgttaaa gcgggggtca aacttatccc cagcaagtc gatttcccc 50340
caaatattca cgcgtctagc ccaggggctg gaacaacgaa aatccagaat cggaaactct 50400
tttccattac agtaaaactt agggggtcga ctaagtgtac cgaagtgaac cccctttcgt 50460
tcttccatg gcacatcttc atctaacaat ttaggggcca aaaattgaa cgatgacatg 50520
gtagtttgt aactatgaag aaattctctg ttactaccgc gcccggttct tgggttatat 50580
ttaatccctg atgcttgggt taaaaaggga ttacaaaacc ccgttctgat cgccatttta 50640
tgtaaacgat tgataatctt gtaaaaagcc agtgttactg agtaacacaa cccacgccc 50700
ttctaataca taaagtgtaa tcacgtgatt tgttgtggtt tccgcatatg taataccgt 50760

```

-continued

ttaaaagcct	ctcttcttaa	tgtatcgaca	gactggggtt	tgggtggtca	tttgaccctg	50820
ccaacaaccc	cccattatta	cgagtacttc	acccaaatgg	aaaatactca	gaagactgtg	50880
acagtgccca	cggggcccct	gggttacgtt	tatgcgtgcc	gggttgaaga	tttgatctg	50940
gaggaaattt	catttttggc	cgctcgtagc	acggactctg	atttggtttt	attacctttg	51000
atgcgtaatt	tgaccgtgga	aaaaactttt	acatccagcc	tgcggtggtt	ttctggagca	51060
cgactactgg	gtcttgccgg	agctggtatt	accttaaac	tcaactaccg	tcatttctat	51120
ccatctgtct	ttgtctttca	cggaggcaaa	cacgttttac	ccagctccgc	ggcccaaat	51180
ctcacacgcg	cgtgtaacgc	ggctcgagaa	cggtttgggt	ttcacgctg	ccaagggcct	51240
cctgttgacg	gtgctgttga	gacgaccggc	gctgagatat	gcacccgcct	tggattagag	51300
ccagaaaata	caatattata	cttgggtgct	acggcattgt	ttaaggaagc	cgtatttatg	51360
tgcaacgtgt	ttctgcatta	tggaggactc	gatattgttc	atattaacca	tggggatggt	51420
atacgataac	cgttattttc	ggtacaactt	ttcatgcccg	atgttaaccg	tctggtacc	51480
gaccattcca	acactcatca	caggctctatc	ggagagggtt	ttgtataccc	aacacccttt	51540
tataacacgg	ggtgtgcca	tttaatacat	gactgtgtta	ttgctcccat	ggccggtgcc	51600
ttgcgcgcca	gaaatgtaac	tgccgtcgcc	cgaggagcgg	cccaccttgc	ttttgatgaa	51660
aatcacgagg	gggcagtact	ccccctgac	attaagtaca	cgtattttca	gtcctcttca	51720
agtgaacca	ctaccgcccg	tggagcgcgt	cgaaacgatg	tcaactccac	gtctaagcct	51780
agcccatcgg	gggggtttga	aagacggttg	gcgtctatta	tggccgctga	cacagccttg	51840
cacgcagaag	ttatattcaa	cactggaatt	tacgaagaaa	ctccaacaga	tatcaaagaa	51900
tggccaatgt	ttataggcat	ggaggcact	ttgccaaggc	taaacgctct	ggggtcatat	51960
accgctcgtg	tggccggggt	cattggtgcg	atggttttca	gccccaaatc	tgcgttgat	52020
ctaactgagg	tggaggatag	cgggatgacc	gaagccaagg	atgggggacc	gggtccatca	52080
tttaatcgat	tttaccagtt	tgccgacct	catttagctg	cgaatcccca	aacagatcga	52140
gatggccaag	ttctatccag	tcagtctacg	ggttcatcaa	acacagagtt	tagcgtggat	52200
tatttggcac	tcatttggtg	atttggagca	cccctgttgg	cgcgactgct	tttttatcta	52260
gaacgctgtg	acgtggtg	gtttacaggg	ggtcacgggg	atgctttaa	atatgttacg	52320
gggacctttg	actctgaaat	tccatgtagt	ttatgtgaaa	aacacacg	gcccgtatgc	52380
gctcacacaa	cagtacaccg	acttagacaa	cgcatgccgc	gatttggaca	agccaccctg	52440
caacctattg	gggtgtttg	aacaatgaac	agccaatata	gcgactgcga	tcctctagga	52500
aactatgctc	catatttaat	ccttcgaaaa	cccggggatc	aaacggaagc	agcaaaggca	52560
accatgcag	acacttatag	ggctacacta	gaacgcttgt	ttatcgatct	agaacaagag	52620
cgactactgg	atcgcggtgc	cccatgttct	tccgagggac	tatogtctgt	cattgtggat	52680
catccaacgt	ttcgtcgcat	attagacaca	ctgcgtgccc	gtatagaaca	gacaacaaca	52740
caattttatg	aagtgttgg	tgagaccgc	gattataaga	tccgtgaag	attatccgaa	52800
gccaccatt	caatggcgtt	aacgtttgat	ccactactcag	gagcattttg	tccattacc	52860
aattttttag	ttaaaccgaac	acacctagcc	gtgttacaag	acttagcatt	aagccaatgt	52920
cattgtgtat	ttaccggaca	gcaagttgag	ggcggaact	ttcgtaacca	attccaacct	52980
gttttgcggc	ggcgttttgt	tgacctgttt	aatggggggt	ttatatcaac	acgctctata	53040
accgtaacat	tatctgaag	tcctgtatcc	gccccaaatc	cgacattggg	acaagacg	53100
cccgcggggc	gtacctttga	tggggattta	gcgcgcgtaa	gcgtggaagt	tattcgggat	53160

-continued

atcacgagtta aaaatagggc cgttttttca ggtaactgta caaatctctc tgaggcagcc 53220
cgggcaaggc ttgtaggcct tgcaagtgcg taccaacgcc aagaaaaaag agtggatatg 53280
ttacacgggg ccctaggggt ttgcttaaa cagtttcacg gcctgttatt tcctcggggc 53340
atgccaccaa acagtaaact ccccaaccg cagtggtttt ggaccctgtt acaacgcaac 53400
cagatgccgg cagataaact tacacacgaa gagattacca ctattgcagc tgtaaacgg 53460
tttaccgagg aatatgcagc aataaacttt attaatctac cccaacctg cataggagaa 53520
ttagcccagt tttatatggc aaatcttatt cttaaatact gcgatcattc acagtacctt 53580
ataaatacct taacttctat aattacgggt gccaggcgcc cgcgtgacct atcatccgtt 53640
ttgcattgga ttcgtaaaga tgtcacgtcc gccgcggaca tagaaacca agcaaaggcg 53700
cttcttgaaa aaacggaaaa cttaccggaa ttatggacta cggcttttac ttcaactcat 53760
ttagtccggc cggccatgaa tcaacgtccc atggtcgttt taggaataag cattagtaaa 53820
tatcacggag cggcaggaaa caaccgcgto tttcaggcag ggaattggag cggtttaaac 53880
gggggtaaaa atgtatgcc gctatttaca tttgatcgca ctgcgccgtt tataatagca 53940
tgtcctagag gaggttttat ctgcccgtta acaggtcct cgtcgggaaa tcgagaaacc 54000
accctatccg accaagtctg cggataatt gtcagtggcg gggccatggt tcaattagcc 54060
atatacgcca cggttgtgcy tgcagtgggc gctcggagcac aacatagggc atttgacgac 54120
tggttaagtc ttacagacga tgagttttta gccagagact tggaggagt acacgaccag 54180
attatccaaa ccctggaac gccctggacc gtagaaggcy ctctagaagc agtaaagatt 54240
ctagatgaaa aaacgacagc gggagatggg gaaaccccca caaacctagc atttaatttt 54300
gattcttgty aaccaagcca tgacaccaca tctaactgat taaacattc agggtaaac 54360
atctcagggt caactgtccc tggctttaa cgacccccg aagatgacga actctttgat 54420
cttagtggtt tccccataaa acatgggaac attacaatgg aatgatgta acctccctct 54480
ttatccaatt aaagcccaca cgcgggtgag tgtacgtaat aaacaagtca atattacata 54540
ttctgttgty tttctttttt ttgtgtgtag tccttaccga tatgacctgt aatagatgt 54600
gtctccaacc attcagctta cagtccagty gacagtaaca gcccgataac atggaattgg 54660
atattaatcg aacattgtty gttctactgy gtcaagtta tacgtacatc tttcaggtty 54720
aactgctacy tcgatgtgat ccaagggty cgtgtcgtt tttatatcgy ttagcggcta 54780
actgtttgac agttcgttat ttattaaagc tgtttctccg gggatttaat acccagctaa 54840
aatttgaaa cactcccacy gtttgytcac tgcattggc attatgtat gtaaaggag 54900
aaggtagcgy tttgttgag ttgctacaac attttaaac gcgtttgtt tatggtgaga 54960
ctaaagactc aaactgtatc aaagattact ttgtctcagc gtttaactta aaaacctgcc 55020
aataccacca tgagctgtcy ttaacaacat acggaggtta cgtatcgagt gaaattcagt 55080
ttttacacga cattgagaat tttttaaagc agcttaatta ctgctatatt atcacgtctt 55140
ctcgtgaggy gctaaacaca ttgaaaccg tgacgcgggt tatgacagat actataggaa 55200
gcggtctaatt accaccgty gagttgtyt atccggcga tccatgtgct atatgtttg 55260
aagaattatg tataacagct aaccaagty agacctaca tcgtagatta ttaggatgta 55320
tctgcgatoa cgttactaag caagttcgyg ttaacgtgga tgttgacgat attattcgyt 55380
gtttaccata tatccctgat gtaccggata tcaaacgtca atccgacctg gaagcgttac 55440
gaacacttca aaccaagacy gtagtcaatc ccatgggacy aaagaacgat acgtttgacc 55500

-continued

aaacatacga	aattgcgagc	accatgcttg	attcttataa	tgTTTTTaaa	cctgcccctc	55560
ggtgatgta	cgccatcagc	gagcttaaat	tctggttaac	gtctaattcc	actgaaggac	55620
cccaacgtac	tttagacgtg	tttgttgata	atttggatgt	attaaacgaa	catgaaaaac	55680
acgcagaact	tacagccgta	acggttgagt	tggcgttatt	tggaaaaact	cccatacact	55740
ttgatagggc	gttttctgaa	gaactcggat	ctctggatgc	aattgatagt	atTTTggttg	55800
gcaatcgctc	atcctcacca	gacagtcaga	tagaagcatt	aattaaagcc	tgttatgccc	55860
atcatctatc	gtcgcctctc	atgcgtcaca	tttctaacc	gagtcatgat	aacgaagccg	55920
cottacgcc	acttttagaa	agagtgggt	gtgaggatga	tttaaccaa	gaggcgagtg	55980
acagcgctac	agcatccgaa	tgtgatctga	acgatgatag	tagcataact	tttgcTgttc	56040
atggatggga	aaactgtta	tccaaagcaa	aaattgacgc	tgcggaaaga	aaacgagtat	56100
atcttgaaca	tctgtctaag	cgctctctaa	ccagcctcgg	tagatgtatc	cgcaaacagc	56160
gccaaagact	agaaaaaaca	ctcagggtaa	acgtttatgg	agaggcotta	ttgcagacat	56220
ttgtttcgat	gcaaaatggg	tttggggcac	gaaacgtgtt	tttagctaag	gtttccaggg	56280
caggggtgat	tatcgacaat	cgcatcagc	aagcggcctt	tgatgcacat	agatttataa	56340
ggaatacctt	agttcgacat	acagtagatg	cggctatggt	acctgcactt	acacataaat	56400
tttttgagtt	ggtcaacggc	ccattgttta	atcacgatga	acaccgtttt	gcacaacccc	56460
ctaacaccgc	cttatttttt	accgtggaaa	acgttgccct	atttccgcac	ttaaaagagg	56520
aattggcaaa	gtttatgggc	ggtgtcgttg	gttccaactg	gcttctcagt	ccatttaggg	56580
gcttttattg	cttttctggg	gtagaaggcg	ttacttttgc	acagagactt	gcctggaaat	56640
atattaggg	gcttgtgttt	gcaaccacac	tattcacctc	tgTTTTtccat	tgtggggagg	56700
tgcggttatg	tcgcgttgac	cgtctaggta	aggatccacg	cgggtgcacg	tctcaacct	56760
aaggtatagg	cagttccac	ggacccttag	acggcattta	tttaacgtac	gaagaaacat	56820
gtccccctgt	ggctattatt	caaagtggag	aaacagggat	cgaccagaat	accgtcgtaa	56880
tctacgatcc	agacgttttt	tctcttctat	acacccta	gcagcggctg	gctccggatt	56940
caacggacc	ggcgttttca	taacctccgt	tacgggggtg	tggttatgct	ttttatgcat	57000
atTTTctatg	tttTttacgg	cggttgtgtc	ggtctctcca	agctcgtttt	atgagagttt	57060
acaagtagag	cccacacaat	cagaagatat	aaccgggtct	gctcatctgg	gcgatgggtga	57120
tgaaatcaga	gaagctatac	acaagtccca	ggacgcccga	acaaaacca	cgttttacgt	57180
ctgcccaccg	ccaacaggct	ccacaatcgt	acgattagaa	ccaactcgga	catgtccgga	57240
ttatcacctt	ggtaaaaact	ttacagaggg	tattgctggt	gtttataaag	aaaacattgc	57300
agcgtacaag	tttaaggcga	cggtatatta	caaagatggt	atcgttagca	cggcgtgggc	57360
cggaagtctc	tatacgcaaa	ttactaatag	atatgcggat	agggtaccaa	ttcccgtttc	57420
agagatcacg	gacaccattg	ataagtttgg	caagtgttct	tctaaagcaa	cgtacgtacg	57480
aaataaccac	aaagttgaag	cctttaatga	ggataaaaat	ccacaggata	tgccctc	57540
cgcatcaaaa	tataattctg	tgggatccaa	agcatggcat	actaccaatg	acacgtacat	57600
ggttgccgga	acccccggaa	catataggac	gggcacgtcg	gtgaattgca	tcattgagga	57660
agttgaagcc	agatcaatat	tcccttatga	tagttttgga	ctttccacgg	gagatataat	57720
atacatgtcc	ccgttttttg	gcctacggga	tgggtcatac	agagaacatt	ccaattatgc	57780
aatggatcgt	ttcaccagc	ttgaggggta	tagacaaaag	gatcttgaca	ctagagcatt	57840
actggaacct	gcagcgcgga	acttttttagt	cacgcctcat	ttaacgggtg	gttggaaactg	57900

-continued

gaagccaaaa cgaacggaag tttgttcgct tgtcaagtgg cgtgagggtg aagacgtagt 57960
 tcgcgatgag tatgcacaca attttcgctt tacaatgaaa acactttcta ccacgtttat 58020
 aagtgaaaca aacgagttta atcttaacca aatccatctc agtcaatgtg taaaggagga 58080
 agcccgggct attattaacc ggatctatac aaccagatac aactcatctc atgttagaac 58140
 cggggatatac cagacctacc ttgccagagg ggggtttggt gtggtgtttc aacctctgct 58200
 gagcaattcc ctgcccgcgc tctatctcca agaattggtc cgtgaaaaca ctaatcattc 58260
 acccaaaaa caccgactc gaaataccag atcccagcga agcgtgccag ttgagttgcy 58320
 tgccaataga acaataacaa ccacctcatc ggtggaatth gctatgctcc agtttacata 58380
 tgaccacatt caagagcatg ttaatgaaat gttggcacgt atctcctcgt cgtggtgcca 58440
 gctacaaaa cgcgaacgcg ccctttggag cggactatth ccaattaacc caagtgcctt 58500
 agcgagcacc attttgatc aacgtgttaa agctcgtatt ctggcgagc ttatctccgt 58560
 ttctaattgt ccagaactgg gatcagatac acgcattata cttcaaaact ctatgagggt 58620
 atctggtagt actacgcggt gttatagccg tcctttaatt tcaatagtta gtttaaatgg 58680
 gtccgggagc gtggagggcc agcttggaac agataacgag ttaattatgt ccagagatct 58740
 gttagaacca tgcgtggcta atcacaagcg atatthtcta tttgggcatc actacgtata 58800
 ttatgaggat tatcgttacg tccgtgaaat cgcagtcctat gatgtgggaa tgattagcac 58860
 ttacgtagat taaacttaa caottcttaa agatagagag tttatgccgc tgcaagtata 58920
 tacaagagac gagctgcggg atacaggatt actagactac agtgaaattc aacgccgaaa 58980
 tcaaatgcat tcgctgcggt tttatgacat agacaagggt gtgcaaatatg atagcggaac 59040
 ggccattatg cagggcatgg ctcagttttt ccagggactt gggaccgcgg gccaggccgt 59100
 tggacatgtg gttcttgggg ccacgggagc gctgctttcc accgtacacg gatttaccac 59160
 gtttttatct aaccatttg gggcattggc cgtgggatta ttggttttgg cgggactggt 59220
 agcggccttt tttgcgtacc ggtacgtgct taaacttaaa acaagcccga tgaaggcatt 59280
 atatccactc acaaccaag ggttaaaaca gttaccgga ggaatggatc cctttgccga 59340
 gaaacccaac gctactgata ccccaataga agaaattggc gactcacaaa acactgaacc 59400
 gtcggtaaat agcgggtttg atcccgataa atttcagaaa gcccaggaaa tgattaaata 59460
 tatgacgtta gtatctgcg ctgagcgcca agaacttaaa gcccgcaaaa aaaataagac 59520
 tagcgcctt ttaacttcac gtcttaccgg ccttgcttta cgaatcggc gaggatactc 59580
 ccgtgttcgc accgagaatg taacgggggt gtaaatagcc aggggtttg ttttaattta 59640
 ttaataaaaa tegtattac gttactcatg tgtctccatt acgcatcaca ggggtatth 59700
 ataccgata atatacaaaa cgcgttttg acctctacc caccgatat cttaacgggg 59760
 ttattatgga atcgtctaac ataacgcgc tacaacaacc gtcgtctatc gcacatcatc 59820
 cgtccaaaa gtgcgctta agtctcaatg aaacagtaaa agattctccc cccgcgattt 59880
 atgaagatag gttagaacac acgccgtac aattaccccg cgacgggtaca ccccgagacg 59940
 tatgttctgt gggacagcta acctgtcgag catgtgcaac gaaaccttt cgccttaacc 60000
 gcgacagcca atacgactac ttaaacacat gtccaggggg cgtcatatth tcaactggcac 60060
 tggagattat aacgggtcga tgggtttgca tcccgcgtgt gtttccggat accccagagg 60120
 aaaaatggat ggcgccatat attatccag accgagaaca accatcatca ggggatgaag 60180
 attctgacac cgattaaath taacttaath aaaaccttac caccataaa aacgccttct 60240

-continued

gtttgtttaa	cacgacaccg	cttaacaaaa	aaaaaaaaac	caaacacgcc	ttttatgaat	60300
gtaatacttt	tatttgttgg	ttaacaccgc	cccaccatca	tctgatttgc	aaacatatcg	60360
gcgtcgtctg	ccgtggaccc	ctgtattaaa	ggggccttgg	aactcgcctc	cactgcattt	60420
acatcttgtc	caactgtatc	tgtatgtggg	gtgcttgttg	tattttggga	tgagcataga	60480
cccgaaacgc	tttgaagctg	ttttaataaa	atcgatattc	gaggatcccg	tgccccctct	60540
ggtatatttg	tatgggtcga	caaaggcatt	tgtgtcccat	tttgtgattt	tagctctgta	60600
acctcctgtt	gcagttttgc	cacaacccca	gcaagctcct	cgtgctgacc	attagaaact	60660
ctgtgtctcc	tctgccaata	tgatggagaa	actcgacgtc	tccgatgctg	tatatacgtt	60720
ggttccaccg	gaaaatatat	atttgaggga	aactctccgt	ccatttgaga	ctccccacta	60780
taaaaagaat	ccaattccct	ttgatccatg	ctcttgaat	cccgttttcc	tggaagacgg	60840
acatcggttt	tgctcggaaa	atttacacac	ggggtctgca	agtcaatacc	ccgttcggcg	60900
gccaatgcgt	tcataaatgc	ggacatttgc	atttccaaac	gattgggtgg	tggtatccc	60960
ggaaacccgt	acggtccccc	gaagtgtccc	ggagggcaac	cataaccccc	tgtattaggt	61020
gggaaggcag	gcgggtgtgg	agatccatat	ggcccgcga	tatactgtcc	gttatttggga	61080
gctccaattg	atacctgcgg	atTTTTtagtc	tgcccggtta	acagctgtga	ataatacgcg	61140
gtaggatata	gtacaaattc	ccctccggtt	ggaacgcccg	acgggggctg	tggtgagata	61200
ttactagcgt	tacctgtctc	agaagccata	tcgctgtcgt	tcctacacaa	ctgcgtaacc	61260
tttaaatgcg	gaacagtctt	ttcacaatct	tcatttgatt	ccccaacacc	caacgcgaga	61320
tcgtatatgg	gcccgcgggg	gtggaatgtg	gogtttataa	cacccgctt	gggtaattta	61380
gactccaccc	cattaacggt	ggttatccga	gcaagtccat	atccggtgct	agcctgaaga	61440
taaaacgtgac	ccataattcc	ggcttcgctg	ctacgttttg	caaccacgtc	ccatctatct	61500
cttaaaagca	tattgtttcac	ggctgtggat	aataacacct	tgggagagtt	atcttcgcta	61560
accttccata	ctttatttaa	acccgcgtag	tctttaacca	gcgacaataa	ccgcgcttta	61620
ctttccatcg	ataaaacccg	gaatggttca	attgaagatt	ccggggtaca	gtcataattg	61680
accactgttc	caacgcgtct	tccaacaaca	cataacgcaa	catgggtaaa	aaaattaccg	61740
tctggtatct	cattcgggga	caatcgtttt	gaagacaggg	atacggaggg	taagtaattt	61800
gtgaccaagt	ataacgcacg	ttctagcggga	gataatacag	aatctctatt	tccaaaaaaa	61860
ttcgaatggg	ccgcttcaaa	cagcaccgca	tgtagttgag	ggcatctaac	gatacccaaa	61920
aaaaaaggtc	cgcgtatgtc	ctcaatgatt	gcgattactt	caccacgac	acagtctttt	61980
cgatgatcga	tgtttattgg	tattttacta	gtaggcggca	aagcggaccg	cacaatctct	62040
ggggtaatat	ttaattcccc	ttcgtccttt	gaatataagg	ctaaatacc	agccacgtat	62100
aacgcttcac	agttctcttc	gtcagcttca	gcagccatta	taaacacccc	acggaccgga	62160
tagtgaatac	tcacggtgtg	gaggcaaaact	gaggaatgac	acccaaacag	acaaaatata	62220
gaagatcata	gtcactgtta	acgttgaact	gcgcaaggcg	gcgactttct	tccaatgccg	62280
cccttacacg	cggttgtgtc	attaacattc	caagtccccg	ttcatattgc	aacataacac	62340
tgatcatgat	tgataccacg	gcggctatgg	gtagggatgt	aacattttgt	cggcgggtgt	62400
ctaattccaa	tgcaattaag	cttatgagcc	gatcttggtg	ctgtccagaa	gaaatatcta	62460
ttacggttct	tcctaaactt	ccacgactaa	gctgggtatg	cgcgtctaaa	caaagagcaa	62520
ctaattccagg	aaacatttca	gtcagctctg	tggtccgatt	taacgtatac	agtggtgcta	62580
tatatcgttc	acataaaaaa	tgaagttat	tattaccgct	tttaaaactt	ccatcaaac	62640

-continued

ccgtcgtcc gcgcaagatt acattgttg taggggtcc tgttgettct gacacaatca 62700
 aaccagttg aaaattatth tttagttat ctccgtatac gttcccgttc cataataagc 62760
 gccttaataa taataacgcc gtaatcgtgt caattgttaa ccttaataga gtttggcttt 62820
 ccataagaaa cacgttttgg gcccgttcta aatagccgc gcccgctgt tgaatcttgt 62880
 ccacatagc ggtatgattg cgatcaataa tgtcattaac cccaggatta aactgtccag 62940
 gtgcagcgg taggacctgc aaccgtataa gcgcatccat aacagaatgt gacgttaagg 63000
 cgccttgatc ataccgccc ccacgagcat gaaactggtc gcgtggtaga cgatcatagc 63060
 aaaattgata actgttttta ttttcgtgtg ttgtcatata attcacaat gtctcagtat 63120
 attccggtag gtgctctata aggttcccga aggacgaaac ttgaggttcg tggacactat 63180
 tagatgtcct atacattaaa tataaacata ataccgcaca ctggaacgcg gactacgctc 63240
 tatctccaac atacattctc ccggcggact gtagacatgt tacogttgtg ttcataaacg 63300
 tacgggaaat gcgcccgtct ttacaatcaa ctccgcgtgc agctacgggc ctatctaaca 63360
 caagccgttc ctgcagagta cgataccatg gcccgaaaac aatccctgga gaggttattgc 63420
 cccttgccct tcccaagtac accaggtgta taaaatccac ttgaaagttt gtatcgtact 63480
 gcaacggtgc atcatttttg gcaactctgta cctcggggtg tatagactca ttgcgtatta 63540
 tttctgtacg tgtacattcc tcagattgtg catctgcttc tccgcctcg gcagcagccg 63600
 tctccaggga atccaaaacc ttggccatgc gcgttagttg ttcttcgagg gctttaaac 63660
 gacgatctat ttccgttgg aacgtaatcg tttcccgcg aaggttgtct aatgcggcaa 63720
 cggccgcgc attttttaac gttaacgtat ttttttcaa atcgggattc ataccgctc 63780
 ttaactcaaa cgcgggagcc gtccagtagt gtatggggaa gttgggggct ataaagttct 63840
 tagtggtaga caaaaatc cccacattat tcgaaaacga gatagatccg aaccatc 63900
 tcgcgctcat ggtgtctgca gcaaaacaaag tcaactggcg tgaatataaa ccggtactgc 63960
 tttaaaagot gttttcttac ccattggaaa acatcccgtt tatactttgt aaaattccac 64020
 cacaagcacc taaagaaggc cttctaaggg gtaaatccac cccacaagct gcaatttctt 64080
 caaactttgt taaagcggaa cgatggcatg atttcgcacg ctttttcgca agagaacata 64140
 cgtgaatttt ctttttgcat agacgtcttc gctctctaac ggacctatc gggggggtat 64200
 attccgctac attctccaaa tgcgacgcta gcataacaag gtttccatga atcaccttg 64260
 ggggtaaccg agttacctgt aacaggttca gacccggtg agatacaaac acaaggaggg 64320
 gggtcacctat tatttcatca gatcccgtgg gtgtggtttc cttattaaa gccatggat 64380
 cctcagctg gcgcataccc tcgcaaaact ggtgatactt agtaggggta tgtatattag 64440
 cgctaaaacg gcaagatttt aattccacta taaaacaaac ggtctttccg gcaccactgg 64500
 attccgtttg tataatacaa acacaatcgg ggcgtcggcg tcccaaattt acttcaaacg 64560
 acattgatat gcgtacagcc ctttgaacat ccacgtggga taacggcgac aggagtttg 64620
 ccagcctcgg gttgaacgcg tccgcgaaac ctgcagctac gttatcaata tcctttttga 64680
 gtacatcgta aaaacgagtg tggcaacgtt gtcccaaacg aaaacacttg gccgaattc 64740
 gactagcgg catatttgaa gttccgtccc agaagataac ctaagacgcg tttgtotaca 64800
 ataaacatgt caacgataa aaccgatgta aaaatggcg ttttgcgtat ttatttgac 64860
 gggcgctatg gaattggaaa aacaaccgcc gccgaagaat ttttacacca ctttgcaata 64920
 acaccaaac ggatcttact cattggggag cccctgtcgt attggcgtaa ccttgcaggg 64980

-continued

gaggacgcca	tttcggaat	ttacggaaca	caaaactcgc	gtcttaatgg	agacgtttcg	65040
cctgaagacg	cacaacgcct	cacggtcat	tttcagagcc	tgttctgttc	tccgcatgca	65100
attatgcatg	cgaaaatctc	ggcattgatg	gacacaagta	catcgatct	cgtacaagta	65160
aataaggagc	cgtataaaat	tatgttatcc	gaccgacacc	caatcgctc	aactatatgt	65220
tttcccttgt	ccagatactt	agtgggagat	atgtcccag	cgcgcttcc	tgggttattg	65280
tttacgcttc	ccgtgaacc	ccccgggacc	aacttggtag	ttgtaccgt	ttcactcccc	65340
agtcatttat	ccagagtaag	caaacgggcc	agaccgggag	aaacggtaa	tctgcccgtt	65400
gttatggttc	tgagaaatgt	atatataatg	cttattaata	caattatatt	tcttaaaact	65460
aacaactggc	acgcggtctg	gaacacactg	tcattttgta	atgatgtatt	taaacagaaa	65520
ttacaaaaat	ccgagtgtat	aaaactacgc	gaagtacctg	ggattgaaga	cacgttattc	65580
gccgtgctta	aacttccgga	gctttgcgga	gagtttgaa	atattctgcc	gttatgggca	65640
tggggaatgg	agaccctttc	aaactgctca	cgaagcatgt	ctcgttctgt	attatcgta	65700
gaacagacac	cccagcatgc	ggcacaagaa	ctaaaaactc	tgctacocca	gatgaacctg	65760
gcaaactatg	cctccggtgc	atggaatata	ttgaaagagc	ttgttaatgc	cgttcaggac	65820
aacacttctc	aaatatacct	agtatttacg	tatgtaccag	taaaaagatg	atacacattg	65880
tcactactgc	gtgtacgtgt	ttttcttttt	tatatatgcg	tcatttatta	ccacatcctt	65940
taatcccgcc	tttatctccc	taaaacggag	tggtaatatt	aaaagccgcc	aagcctgttg	66000
gtgggtgagg	aggggtaaag	gcacgctgtg	tgcataacgt	tgcggtgata	ttgtagcgca	66060
agtaacagcg	actatgtttg	cgctagtttt	agcggtggtg	attcttctc	tttgaccac	66120
ggctaataaa	tcttacgtaa	caccaacccc	tgcgactcgc	tctatcggac	atatgtctgc	66180
tcttctacga	gaatattccg	accgtaatat	gtctctgaaa	ttagaagcct	tttatcctac	66240
tggtttcgat	gaagaactca	ttaaactcact	tcactgggga	aatgatagaa	aacacgtttt	66300
cttggttatt	gttaaggtta	accctacaac	acacgaagga	gacgtcgggc	tggttatatt	66360
tccaaaatac	ttgttatcgc	cataccattt	caaagcagaa	catcgagcac	cgtttcctgc	66420
tggacgtttt	ggatttctta	gtcacctcgt	gacacccgac	gtgagcttct	ttgacagttc	66480
gtttgcgcgc	tatttaacta	cgcaacatct	tgttgcgttt	actacgttcc	caccaaacc	66540
ccttgatgag	catttgaaa	gagctgagac	cgcagcaact	gcagaaaggc	cgtttggggt	66600
aagtctttta	cccgtcgcgc	caacagtccc	caagaatact	attctggaac	ataaagcgca	66660
ttttgctaca	tgggatgccc	ttgcccgaca	tacttttttt	tctgccgaag	caattatcac	66720
caactcaacg	ttgagaatac	acgttcccct	ttttgggtcg	gtatggccaa	ttcgatactg	66780
ggccaccggt	tcggtgcttc	tcacaagcga	ctcgggtcgt	gtggaagtaa	atattggtgt	66840
aggattttat	agctcgtcta	tttctttatc	ctctggacca	ccgatagaat	taattggtgt	66900
accacataca	gtaaaactga	acgcggttac	aagcgacacc	acatggttcc	agctaaatcc	66960
accgggtccg	gatccggggc	catcttatcg	agtttattta	cttggacgtg	ggttgatata	67020
gaatttttca	aagcatgcta	cggctgatata	atgcatat	cccgaagaga	gtttggatta	67080
ccgctatcat	ttatccatgg	cccacacgga	ggctctcgcg	atgacaacga	aggcggatca	67140
acatgacata	aacgaggaaa	gctattacca	tatcggcga	agaatagcca	catcaatttt	67200
tgcgttctgc	gaaatgggcc	gtaccacaga	atattttctg	ttagatgaga	tcgtagatgt	67260
tcagtatcaa	ttaaaattcc	ttaattacat	tttaatgcgg	ataggagcag	gagctcatcc	67320
caacactata	tccggaacct	cggatctgat	ctttgccgat	ccatcgcagc	ttcatgacga	67380

-continued

actttcactt ctttttggtc aggtaaaacc cgcaaatgtc gattatttta tttcatatga 67440
tgaagcccgat gatcaactaa agaccgcata cgcgctttcc cgtgggtcaag accatgtgaa 67500
tgcactttct ctccgagcgt gtgttataat gagcatatac aaggggctgc ttgtgaagca 67560
aaatttaaat gctacagaga ggcaggtctt attttttggc tcaatgattt tattaatatt 67620
ccgcaagaga ctgaaaaatt catctcgggt attagacggt cgcacaactt tgcttttaat 67680
gacatccatg tgtacggcag ctccacccac gcaagcagca cttaacatac aagaaggcct 67740
ggcatactta aatccttcaa aacacatggt tacaatacca aacgtataca gtccttgtat 67800
gggttcctt cgtacagacc tcacggaaga gattcatggt atgaatctcc tgcgccaat 67860
accaacacgc ccaggactta acgaggtatt gcatacccaa ctgacgaat ctgaaatatt 67920
cgacgcgca tttaaaacca tgatgatttt taccacatgg actgccaag atttgcatat 67980
actccacacc catgtaccag aagtatttac gtgtcaagat gcagccgagc gtaacggaga 68040
atatgtgctc attcttccag ctgtccaggg acacagttat gtgattacac gaacaaaacc 68100
tcaaaggggt ttggtatatt ccctggcaga tgtggatgta tataacocca tatccgttgt 68160
ttatttaagc agggatactt gcgtgtctga acatggtgct atagagacgg tcgcaactgcc 68220
ccatccggac aatttaaaag aatgtttgta ttgcggaagt gtttttctta ggtatctaac 68280
cacggggcg attatggata taattattat tgacagcaaa gatacagaac gacaactagc 68340
cgctatggga aactccacaa ttccaccctt caatccagac atgcacgggg atgactctaa 68400
ggctgtggtg ttgtttccaa acggaactgt ggtaacgctt ctaggattcg aacgacgaca 68460
agccatacga atgtcgggac aataccttgg ggcctcttta ggaggggctt ttctggcggg 68520
agtgggggtt ggtattatcg gatggatggt atgtggaat tccgccttc gagaatataa 68580
taaaatacct ctgacataaa aaacatggt aataaaaagt cactataaac gtattctcta 68640
caatacttta ttcgcaata atacacacta cctttgggtt tttttccgt ccccaaatgg 68700
tgtttggtgc actctaccaa aaaatagagc gctaaatat gctatataac gcctccagc 68760
aaaatacgtt tcaaaggcat taccgatgat tgtattgtag tacagggcaa tgggaattga 68820
tgatcccaat aaacggcata gacgcacagc gccgttatag caggggtctc cagagtacag 68880
ggatctaaag taccgggata tctcactctc atgcctttcc gtgacagaaa catcaaccgg 68940
aacagtatcc gataaaacca ctctgtttt tgcaaggcgt aaaattcgca caccttcctt 69000
ttttgcaaga tgtgacgttt ccttgaaca gggaaactgg gggagtggta agaacaaca 69060
agtttcagcc aactgcca taaagcccac ttccctcaag aggotgtttg ctgtatccac 69120
aatggtccgt attaaatctt gagcaacttg atccgtgtca tcatcactgg gtaacgctt 69180
aacataacta cgcgttaaat ctcaataac ggcataaca ttaaacgctt cccaccgaga 69240
cagtataat tgaacaatca cgaaccgttg acaggacgtc agatcacgtc cgtaagcatg 69300
cccgaataat ggaagttccc ccggttcgcc atataccgca acaactgag tatatatcgt 69360
ctcagggct tcattaagtt catcttcaag tccagccat tttctggctt taaatataac 69420
ctcgtccgca aaaaaaccg cacatgataa cgcgcgata caatgagtag tggctttatg 69480
gcgaggatcc caaatgtcca ttaccgggg gatggtccta atctgtaca agttacttag 69540
tgtaatatga tcggacttct taocccgtct aggtgtttc tcagaatacg gttcaaccga 69600
aatcggcaca tcatctgctt ttacgtcttc cgtaaccaca tcagcagcgc gccgactaac 69660
aattatactt gtttttcat cgtcgttact tccgttaagc gcgtctgta tctcgggct 69720

-continued

cccgtcgaat aatccaactca cttagctcctg caaactttct ggtaactcca acatacgcac	69780
atacaccaat gaaaaactgg ctctcgtttgg tacgtacata aagccatttg tggatattaat	69840
ggcggtaggg gttggaaaca attttagctt attctcgcgc gtaacatcta cccccccac	69900
caatgttaaa tgcgtcacgg ggaggacac gagataatct gcgagcgtag ggtcctccac	69960
ttcaacatca aatgttccgc aaaggtcgcg atccaccgcc cccgatcccg ctgcaagtaa	70020
ggccactcga tccaaaaaca cgcagttatt attggatgat accgcccacg tcttcccggg	70080
gcgattgagc tcaacttcgaa cgtaactggc aacagatctg tcaccgggtc cgaccccgcg	70140
aacaacatgt ccaaattttg cgatctcgc tccatgtttg cggggtatgg aaattaagca	70200
tccccgcac ataaaaaac ccctggtagc acgctcgtta aaataaacg ttacgccgtt	70260
ataagatacg gttgaatgat atggaaatc catattaaag cgtttatcgg aacattaacc	70320
tcgaacttgc cgtcccgtga tcgtgtgato gccaacctta ggtccacacc gaatatgaga	70380
aatatataac tacacgcaa cattcaaac accgtggtat cattaacgtc atatgaaaag	70440
atccaatcaa tccaatcaac cacacctcct accgtttagc acgtcagcta tgtgacatgc	70500
tccaaacata cgtaaacatt tagagagggg gttataacag tctgtcaggc ggggtatatt	70560
ctacataata caaggatcgg cttaacttt gtcaacattt ttactttgga ctataaactg	70620
cgactgaaoc ttatgaaccc acccaagcc cgcgtctcgg aacagacaaa ggacttgctt	70680
agcgttatgg ttaaccagca ccccgaagag gacgcaaaag tgtgtaaatc cagtgataat	70740
tcaccgcttt ataaccacat ggttatgta tcgtatggg gtgatacggc cttactatta	70800
agctctgcac gtacccgcac atctaccgta aacaggtcgg cgtttacgca acactccgtg	70860
ttttatatta tatccacggt gttgattcaa ccaatatgtt gtatcttctt tttttttac	70920
tataaagcga cacgctgtat gctcttatc acagccggg tacttctgac gattctacat	70980
cactttcgcac ttattattat gttattgtgt gtctacagaa atatacgatc agacctgcta	71040
cccttatcta catcccagca actgctgctt ggaattattg ttgtgactcg aacaatgcta	71100
ttttgtatta cggcgtatta tactctttt atagacacc ggggttctt tttgattacc	71160
ggacacttgc aaagttaggt tatttttcca gatagcgtt caaaaatact tcctgtgtcg	71220
tggggtccaa gtccagccgt gttactggta atggcggcag ttatttacgc tatggactgt	71280
ttggtggaca cggtatcctt tattgggcca agggtgtggg tccgtgttat gttaaaaaca	71340
tctatttctg tttagtcac ttcaataaat gtactataat tgttcagtct aaaaataatg	71400
ttgggtatatt ataattaccg cccccgtgtt acttggaac acccatacat atgttccact	71460
ctacatcaaa cttctcgcag tttcttgtt cccgcacacg tttacacgtc cggattcaag	71520
tcgcaacgct gctgacaaaa tgacaacggt ttcattgtcc gctaacgtga ttactacaac	71580
ggaatctgat cgtattgctg ggtattttaa catcccagcg gggatcattc caactggaaa	71640
tgtgctgtoa accatagagg tgtgtgcaca ccgttgcat tttgattttt ttaaacaaat	71700
acgatcagat gataacagcc tttactcggc tcaattcgat attcttttg ggacatactg	71760
caatacatta aactttgtgc gttttctaga acttgactg tctgtcgtt gcatctgtac	71820
taaatttccg gagctggctt acgtgcgaga tggcgttatt caatttgagg tacaacaacc	71880
catgatagca cgtgatggcc cacatcccgt cgatcagcct gttcataat atagggttaa	71940
gcggatacac aagcgttctg taagcgtgc gtttgaatt gcatcggaag cgttgagttt	72000
gttaagtaac acatattgct atgggacaga gattgactca tcgttacgta taagagctat	72060
ccaacagatg gctcgaat taccgaccgt tttggactca tttgaacgag gcactgccga	72120

-continued

tcaacttctt	ggtgttctat	tggagaaagc	cccaccgcta	tcgctgcttt	caccaattaa	72180
taaattccaa	cccgagggac	atctaatacg	tgttgacgc	gcgccctac	tttcggacct	72240
caaacgtaga	gtctgtgcg	atatgtttt	tatgaccoga	cacgccagg	aacctaggct	72300
gatctctcgg	tatctgtcgg	atatgtttc	gtgcaccaa	ccatcgggta	tggtatcacg	72360
aataactcat	acaacactc	gcgacggca	ggttgacggt	gtgttggtaa	caacagcaac	72420
cttaaacg	caactattac	aggaatfff	acaaattgac	gacaccgocg	ctgacgtacc	72480
agtaacatat	ggcgaatgg	ttctacagg	gacaaacttg	gtaaccgccc	ttgtgatggg	72540
aaagcccgct	cgcggaatgg	atgatgtagc	ccgccatctc	cttgatataa	ccgaccctaa	72600
cacgttaaac	ataccgtcta	tacccccaca	atccaactcc	gattcaacga	cagctgggct	72660
tccggttaac	gcccgtgttc	ctgcggtatt	agtgattggt	ggggataaac	ttgtattctt	72720
agaagcatta	gaacggcggg	tctaccaagc	tacgcgcggt	gcctaccctc	ttattgaaa	72780
tatagatatt	acgtttatca	tgccaatggg	agtgtttcag	gcaaaactca	tggacagata	72840
tacacgacac	gcccggcatt	tttcaactgt	atccgaacag	gatccacgtc	aatttcacc	72900
ccaagggatt	tttttttata	ataaagatgg	gatattaaca	cagttgactc	ttcgtgatgc	72960
aatgggtacc	atctgccaca	gttcattgct	tgatgtcgag	gccacacttg	ttgccctccg	73020
ccaacaacat	ttagatcgtc	agtgttattt	tggtgtatac	gtggccgagg	gtacagagga	73080
cacattggat	gttcaaatgg	ggaggtttat	ggaaactggt	gcagatatga	tgccctatca	73140
ccctcattgg	gtaaacgaac	atttaacaat	tctacagttt	atagctccga	gcaaccocg	73200
tctaaggttt	gaattaaacc	ccgcctttga	tttttttgtt	gcaccggggg	acgtagacct	73260
tcccggacog	cagcgtcccc	cggaagccat	gccaaccggt	aacgcaacat	tacggattat	73320
caacggaaac	attcccgtgc	ctctatgtcc	catttcattt	cgagactgtc	gcggaacca	73380
actcggtttg	ggaagacata	caatgacccc	ggcaaccatt	aaagccgtaa	aggatacatt	73440
tgaagaccgc	gcatacccaa	ctattttcta	catgctagag	gctgttattc	atggaaacga	73500
aagaaacttc	tgtgcgttac	tgcgactggt	aacacagtgt	attcgcgggt	attgggagca	73560
atcccacag	gtggcatttg	taaataactt	tcacatgtta	atgtacataa	ctacatatct	73620
cggaaacggt	gagcttcccc	aagtctgtat	taatataat	cgggatttac	tgcaagatgt	73680
aagagcatta	cgccaaacta	taaccgattt	tacaatacaa	ggagagggcc	ataacggcga	73740
gacctcggaa	gcgctaaata	acatccttac	ggatgacacg	tttattgac	ctattctatg	73800
ggattgtgat	gcgttaatat	accgtgatga	agccgcccga	gaccgactcc	ccgcaattcg	73860
tgtaagcggg	cgaaacggat	accaagccct	tcactttgtg	gatatggcgg	ggcataactt	73920
ccaacgacgc	gataatgtgt	taatccacgg	gagaccggtt	cggggagaca	cgggtcaggg	73980
tattcccatt	actccacacc	atgaccgtga	atggggattt	ctctccaaga	tttactacta	74040
tattgtcatt	cctgcatttt	cccgcggttc	ctgttgatac	atgggcgtgc	gttatgatcg	74100
cctataccct	gcgttacagg	cagttatcgt	tccggaaatt	cccgtgatg	aagaagcccc	74160
aactacccca	gaagatccaa	gacacccctc	tcacgcacac	caactcgttc	cgaactctct	74220
taacgtttac	ttcataatg	cacacctaac	cggtgatggt	gatgcattgc	tcacactaca	74280
agagttaatg	ggagatatgg	ctgaacgaac	gacggccatt	ttagtatcaa	gcgccccga	74340
tgccgggagc	gccacggcaa	caaccagaaa	tatgagaata	tatgacggag	cgctttacca	74400
tggccttatt	atgatggcat	atcagcgcta	cgatgaaacc	attgcaacgg	gtactttttt	74460

-continued

ttatcccgtt	ccggtcaacc	ctctgtttgc	atgtccgga	catttggcat	cattgcgtgg	74520
aatgacaaat	gctagcgagg	ttttggcaaa	aatggtacca	ccaatccctc	cttttctggg	74580
agccaaccac	cagcaacta	tacgccaacc	cgttgcctac	catgtaacgc	atagtaagtc	74640
ggattttaat	actottacat	attctcttct	tggagggtat	tttaagtta	caccaatatac	74700
tcttacacat	caactacgaa	cgggatttca	ccccgggatt	gcctttaccg	tagtgcgcca	74760
ggatcgcttt	gccacagagc	aacttttata	tgccgagcgt	gcttctgaat	cgtactttgt	74820
cggacaaatc	caagtacacc	atcatgatgc	tattgggggg	gtaaacctta	ccctaaccaca	74880
accagagcct	cacgtggacc	tgggagtcgg	gtatacagct	gtatgtgcca	cagcagccct	74940
gcgatgcctc	ctcacggata	tgggcaatac	tgcccaaaat	cttttttttt	cacgaggagg	75000
agtgcacaat	ttacatgata	acgttaccga	atcgttgcgt	cgtataacag	catcgggggg	75060
tcgcttaaat	cccaccgaac	ccctaccat	cttcggcgga	ctacgtctctg	ctacatcggc	75120
aggaattgca	cgagggcaag	cctctgtgtg	tgagtttgtg	gccatgccgg	tgtccactga	75180
cctacaatat	tttagaactg	catgcaatcc	tagaggtcga	gcatctggaa	tgttatatat	75240
gggtgaccgt	gacgccgaca	tagaggctat	aatgtttgat	cacacacaat	cggatgttgc	75300
ttatacagat	cgagcaactc	ttaacccatg	ggcatcaca	aaacattcat	acggtgacag	75360
gctatacaac	ggaacatata	acottacagg	cgcttctcct	atctacagcc	catgctttaa	75420
gttttttaca	ccagcggagg	ttaacactaa	ttgtaataca	ctggatcggc	ttctaattga	75480
ggcaaaggct	gtggcgtcgc	aaagctccac	cgacactgaa	tatcaattta	aacgccctcc	75540
cggttctaac	gaaatgacac	aggatccgtg	tggccttttt	caagaagcat	atccaccact	75600
atgctcaagc	gatgcggcca	tgttacgaac	ggctcacgcg	ggagaaaccg	gggcagatga	75660
agttcaacta	gcccaatatac	tgattcgaga	cgcgctgccc	cttaggggat	gtcttctctc	75720
tcgcgataat	tttaccacag	cccacatacc	cactcccaat	aaaagcctctg	tagagcgcac	75780
tggcatctta	cttgagattt	ggatacgcctc	ggccgacttg	gtctgtttca	cgtctcttca	75840
aacaacatgg	ctatgccatt	tgagatagag	gtattgttac	caggagaact	atccccggcg	75900
gaaacatctg	cattacagaa	atgtgagggg	aaaattatta	ccttctcaac	cctgcgtcat	75960
cgagcttcac	tgggtgatata	agcgcgtctg	tcataattaca	ttaacgggtc	tccaccagac	76020
acgctctcgc	tgttagaggc	ataccgaatg	cgattcgcgg	cagttataac	acgggtcatc	76080
ccgggaaagt	tgttggcgca	tgccattggc	gtgggtactc	ctacaccocg	gttgtttatt	76140
caaaatacat	cccccgttga	tctttgtaat	ggcgattaca	tctgcttact	tctcccggtt	76200
ttcgggtcog	cagactcaat	tcgcttggac	tctgtaggac	tggaaattgt	tttcccttta	76260
accatcccc	agaccttaat	gcgagaaatc	atcgccaaag	tggttgcacg	ggccggtgag	76320
cgcacggcgc	cgggtgctca	aattttaccc	cacgaagtcc	tacgagggcg	ggatgtcatt	76380
tgttacaatg	gaagcgtta	tgaactcgaa	acaaatttac	aacatcggga	cggatcggat	76440
gcggctattc	gcacattggt	tttaaatcta	atgttttcca	taaacgagg	atgtctgctt	76500
ttattggcgc	tgattccaac	ttgttagtgc	caaggagcac	acgacgggta	tgtaaattta	76560
ttgatacaaa	cggccaattg	cgtagagaa	accggccagt	taattaatat	accgccaatg	76620
ccgcggtatc	aagacggcca	tcgccgattt	cccatatatg	aaactatttc	atcttgata	76680
tcaacatcat	ctagactggg	ggataccttg	ggaactcgcg	caattttacg	cgtctgtgtg	76740
tttgatggac	cctctactgt	tcacccggga	gaccgcacgg	ccgtgattca	agtgtaaaca	76800
gggtgtaata	aaaacacaac	cagctagtgt	acatttcacg	cgtcttgttt	ttatttaata	76860

-continued

ggcataaaca cggaatccgg tatacatgaa ctgccaatat acacggacat aattaatgca 76920
 accatcagat catctgacat tgttcccgtg gtacctttac ccggtgaagt ttttgtgtct 76980
 agattaccca taccgccttt aattacctct gtcaggttat ccaactgttt acatagatac 77040
 tccacggggg ctacacctaa ctttactgtt agggatacaa gctcctgtga ggctattata 77100
 tttccggagt taaatcgttt aacaaaatag tctacggccg gcgttttttg tttttgtaat 77160
 aaaaaaaaaag ggtacgccac gctacatccg ggaggatgg aatgataaaa cagtaacact 77220
 ggagcggaag atagcacgtt tcccttttcg aggacagcaa actggtgtgc tatagccaac 77280
 gatattgcaa ctgcagaatc ctggctgctg tttccctcta tagaaactg tacgtttgta 77340
 aatgtattgg ggtgtaaagc gagtatgtgg cctaagcatt gagtaacgca acgcccatac 77400
 tcaactggaag acgtgccagt taaagctcta agaaaaagt gctccaatcc aaatataatc 77460
 caatccgact tataacgacc aacaatcgct acaccagtac cagacgctcg tgtatttgag 77520
 gtaaactgag ggtctacgta aacgtacaac actgacgata atatagcaca attcgcaacg 77580
 gttgacggcc gatataaaat aaacctctca cgggcagttt ttgtaataa tggccgggca 77640
 aacccccacc cccagaatc ctgtttacgc ccacctaca tttcctgcac gaaggagtgc 77700
 gccataaata aatctgcagt gcgccgatg gctccatcca ttgtgatgaa aaccggctta 77760
 ttttaatacat aacacgaaca agctgtgaca tcgctatgtg ctaaaacacg cggcatgtga 77820
 tcgtcgcata catatgtaac aacgtttaa aactgatccg acgatccacg taagtatac 77880
 aaaaaacttg tacttgcttt tccggtattt gttgatgaaa caaaaaataat tttacaattg 77940
 gtttgattta aaaaaccgac tatagtttgt acagcatcag gtcgaataaa attagcttca 78000
 tccacaaaaca gaagattaaa atcttgacct cggataccct ggaacgatag aaagatatat 78060
 agttaccoca ccaaagttta aatgtatcct taaataccac gtacgtaaaa aatgtttgaa 78120
 tacgtacata tttctttttt ttttccagta caaccatata cgggtgataa tgggaagcca 78180
 tttggcaaat gaaaccaaac atgcactttg gcataatgat cacacaaaag gattactaca 78240
 cgttgatgata cctaacgcgg ggcttattgc ggccggaata gatcccgcat tactgatattt 78300
 aaagaaacc ggacaacgct tcaaggttga agtacaaca agatatcatg ctacaggca 78360
 atgcgaaccg tgggtgcaag ttttcgccgc gtacattccc gataacgcct taacaaatct 78420
 cttaatacca aaaacggaac catttgtttc acacgttttt tccggccacgc ataattcagg 78480
 gggattgatt ttatcattgc ctgtttatct tagccccggt ttattctttg atgcatttaa 78540
 cgttgtagcg atacgaataa atactggaag ccgcaagcac cgtgatattt gtattatgta 78600
 tgcaagaacta atccaaaacg gaacgcgta ttttgctgat ggacaacggg tacttttatt 78660
 atgcaaacag ctgattgcgt atatccgatg caccctcctg cttgcatcgt ctataaaaat 78720
 atacgcagag catatggtgg cagccatggg tgaatcacac acgtcaaatg gggacaatat 78780
 tggaccctgt tcatccataa tcgatcttga tcgacagta acttctggag gtattgatga 78840
 ctcccctgct gaaacacgca tacaggaaaa taatcgggac gtccttgagc taataaaacg 78900
 ggccgtaaac attgttaact ccaggcacc cgtccgacct tctagtccc gcgttgcatc 78960
 tgggttgctt caaagtgcaa agggccacgg agcgcaaac tccaacacag atccgatcaa 79020
 taacggttcc tttgatggcg tccttgagcc goctggacaa gggogattta cgggaaagaa 79080
 aaacaattcg tccgccagca tcccaccttt acaagacgtt ctattgttta cccagcttc 79140
 gacagaacc caaagtctta tggaaatggt cgacatctgt tatgcccaat tagttagcgg 79200

-continued

ggacactcca	gcagatttct	gaaaacggcg	tcccctatca	attgtaccgc	gacattacgc	79260
agaatcccc	agtccgttga	ttgtagtata	ttacaacgga	tcctctgcct	ggggaggacg	79320
tattaccgga	agtccaatth	tatatcactc	tgacacggct	attattgatg	ctgcgtgtat	79380
aaatgcccgg	gttgacaatc	cccaaagcct	acatgtgaca	gctcgccaag	agctagtcgc	79440
gcgtttaccg	tttttgcta	acgtcctaaa	taatcaaacc	cccttaccgg	cctttaaacc	79500
aggcgccgaa	atgtttttaa	accaggatth	taaacaagcg	tgtgtgacat	cgctaaccga	79560
aggctctata	acggagttag	aaacgaaccc	gactctacaa	caactcatgg	aatatgatata	79620
tgacagattct	tcccacacgg	ttattgatga	aattgtagcc	cgacacaccg	acctgattca	79680
gactatagtt	tcggtgttaa	cggaaatgth	aatggatgth	ttttataaca	gctccttgat	79740
gtatgcgggt	ttggcgtatc	tgctcatctg	atatacacga	ccacaaggth	gggggtatata	79800
accctacctt	cacgcttctc	tcccattgctg	gtaggtaaat	cgttctataat	atthatttga	79860
ctattataat	tcaggagggg	aaatacttaa	gctttccaag	gtccccgthc	ccgtagcctt	79920
agaaaaggth	ggtattggta	attccacaca	actgaggggt	aaatttatac	gcagcgcgga	79980
tattgttgat	attggaatth	gttctaagta	tttaccggth	caatgttacg	cgtaacattg	80040
tctaggatth	aaccagcaat	tacaatccat	tttagtttth	ccgggggghat	ttgcggcatg	80100
tttttgatth	accgataacc	tacaggcagc	actactgca	tcgttaatcg	gacctattct	80160
agacagatth	tgcttctcta	ttcccacacc	ccataaataa	attagtgtca	ctataaaaac	80220
ataacaccag	aatctcttca	tatgtaatth	tacgtcattth	ctcccgtthc	cacccctct	80280
taaaatataa	ataaaccggg	tggtgggcat	taaaccaca	agtaccggg	cggaatccg	80340
ctagactgth	ttctgtctca	tggaattaca	acgcataatth	ccgctgtaca	ccgctacggg	80400
tgacagcgc	aaattaaccc	ccgaggcagth	tcagagactc	tgcatgcat	taacgctgga	80460
tatgggatta	tggaagtcca	tcctgaccga	tcccgggthg	aaaataatgc	gatcaactgc	80520
ttttataact	ttaagatgth	ctccgtthtth	cccccttcaa	acggataacta	ctaataatgth	80580
cgthgttgta	gccacaatth	acatcacgth	cccacgthcag	atgaacttac	ctccgaagac	80640
ttttcatgta	attgtaaatth	ttaattacga	ggtctcgtac	gcaatgacgth	cgactttaaag	80700
aatthtaccg	gttgaataa	tagaccatgth	ttttggagca	acgtttaaaga	accgcatgth	80760
gtacccctth	ccaacatcta	ttccggatth	tcgagcagat	cccacccctg	cagatcttac	80820
accaacgcca	aacttaagca	actacttaca	accccgctgth	cttccgaaaa	atccatacgc	80880
atgtaaaagth	atthctccgth	gagtgthgth	gtcagacgaa	cgaagcgtth	tatatgtact	80940
ggctatggaa	cctaatttaa	tagggctatg	tcccggcggth	tgcatgctc	ggatacttgth	81000
ctctgtatth	aatgactctc	tcagccatgth	ggacggatgth	gatgaaatgth	atcatagagth	81060
tcacgtgggth	gcactgtatg	cgthacccca	tgthacaaaat	catgcggaag	gttgthgthgth	81120
ttgggctcgg	tgatgtgga	gaaaggccgth	tcagcgggaa	ttaaaagthgth	agthtagacat	81180
tgggccacg	caggttctth	ttgtagatgth	caccacctgth	attogaatth	cgactactaa	81240
aaatctctgth	attaccgcaa	atcttgggca	cgthtatagth	ggaaccaacg	ccagthgthct	81300
ctctgtacga	gtaaatthcat	ctgggtggca	gctthtatatg	ttggagaaa	cattaagcgg	81360
ggctattatth	aacggctgth	gtctgcttca	gogaattgth	ttcccggaga	cacaagatth	81420
atcgggtgaa	ccggaaccta	caaccaccta	gtatacctta	actcaaccgth	cgthgtggaa	81480
agthtatatgth	caacattthac	agthaatatth	taaaagthta	atthataaaa	cactcacgth	81540
tgthgttgta	ctgacgcgga	acaccgctgth	gctgthaaagac	ccgthcggthaa	atgaaaacgth	81600

-continued

aatagattcg ccttttacat gatccacgta atttgcccca aaccactggt ccaggcgaga 81660
cttgataacc tcaaacacgg gttccgttgc tttgcgtata tgagccgatat aaccactttt 81720
aattcctcta aacgtggcca ttactaaagc tattaatggt acaagaaacc atgttttccc 81780
atgtctacgt ggtaccaaaa acacagttga tttttgtttg aagtgttcta aaacactgtc 81840
agaaacactt ggcgtgttaa aactgtacg cagaaagcag tcaactctgt cggcatgatac 81900
gcccaatagc accgatgaaa taaaatgctg ggtgtgcatg aggatcattt tttgaaacag 81960
ttccaaacgc cccttatatc tgccatagat tggaacgtca acctttgctc gtttgccatg 82020
acttccacac tcttcaatac tctcaaaaga tgtttccaca aggtacgaaa accgttgtgt 82080
aaaggtagac aactgacaga aactatccga cagagaaaac gcgcgaaatg tgttcataac 82140
accgctatac gcatttcgat gaggtgctgc ttcttccggt gaatattcat aaaactgtac 82200
actactgaca gcctttttta attcagggct tacgtttgca tttaccgaat atcgccatgg 82260
tttcaaaact acattggggg tacagttgta ccctgttgac gatagaaacg cgccaaacat 82320
tgcccgctga gcagtagccg agaacagtgg aatatattca caacagttgt gaagcgttcc 82380
aattccggga ataacggcct gatgacgtcg ggttacatct atagcaaaat tcagaaacgg 82440
gatttgggtt gcgtttccca gagacccttg ccgcgtggaa cacgggtagt gggactccaa 82500
cgtcccaaa gcgttcatccc tacgacgctt tagacgttca aaatatctta cagattcttc 82560
accaagcgta cgaccaaaaca ttatcaatga catttaacat caattcacgg aatccgcctc 82620
atctcttgta agcagtaaaa caggaagccg cgtcatctta cgtactcgtt acgtatata 82680
cataaacatt ttcagggccg cattcattca ctttggctcat gtcaggccac actccaacct 82740
acgcttctca taggcgtaac cgtgtcaaac tagttgaggc gcataaccgc gcgggggttat 82800
ttaaagaacg gaccctcgat ctaatccgtg ggggtgagag tgtacaagat ccagcatttg 82860
tgatgctctt tactgctgca aaagaggcct gcgcggattt aaataaccag ctccgctctg 82920
cagctcgcat agcttcagtt gaacagaaga ttcgtgatat acaatccaag gttgaggaac 82980
aaacaagtat tcaacagatt ttaaatacaa acagacgcta tatagaccc gattttattc 83040
gcggtttgga taaaacagaa gacgataata ccgataatat agacagactg gaagacgcgg 83100
taggaccgaa catcgaacac gaaaatcata cttgggttgg agaagacgac gaagcgttac 83160
ttacacaatg gatgctgacg acacaccccc caacctcaa atatctcaa ctgcaggacc 83220
tttgcgttcc caccacaata ccgacggaca tgaaccaa at gcaaccgcag ccgatcagca 83280
agaacgagaa tccaccaaac ccacacacgg atgtgtaaat catccatggg ccaatccgctc 83340
aactgcaaca tgcatggaat caccagaacg atcacaacag acaagcttat ttttattaaa 83400
gcacggctta acgagagatc caatacatca acgcgaaagg gtggacggtt ttccacaatt 83460
taacaaaccc ccatgggttt ttagaatttc caaattatcc cgtttaattg taccatctt 83520
cacgctcaat gaacagttat gtttttctaa attacagatt cgagatagac ccaggtttgc 83580
gggacgggga acgtatgggc gtgttcatat ataccatcg tcaaaaatag ctgtaaaaac 83640
catggacagt cgtgttttta atagagagtt aattaacgcg attttagcga gtgagggttc 83700
tatacagaca ggggaaagcg taggtatttc tagcatagtt tgccttttag gtttttcggt 83760
acaaacaaa cagctactgt ttccggcata cgacatgat atggatgaat acattgttcg 83820
cctgtccaga cggttgacaa tacctgatca catagacaga aaaattgcc atgtattttt 83880
agatttggct caagcgttga cgtttttaaa tgaacgtgc ggcctgacct acctagatgt 83940

-continued

gaaatgtggc aatatttttc ttaacgtcga caactttgcc tcgttggaaa taaccacagc	84000
agtaatcggg gactatagcc tagtaacatt aaatacgtat tccctttgta ctcgagcgat	84060
at ttgaaagt ggaaatccat cccacccgga gcacgtacta cgcgtacccc gggatgcatc	84120
gcagatgtca tttcgtttgg tgttgagtca tggaaacaaac caaccocctg aaatcttgct	84180
tgattatatt aatggaacgg gccttactaa atatactgga accttgcccc aaagagttgg	84240
acttgcgatt gatctttatg cattgggcca agcactctta gaagttatcc tgctaggacg	84300
tcttcccggg caactgcccc tttcagtaca tcggaccccg cattatcact actacggta	84360
taagttatca ccagatttgg cgttgatagc gctggcatat cgatgtgtcc tggcgccata	84420
tatactccca tctgacatcc cgggggactt aaattataat ccctttatac acgcccggaga	84480
gctgaacacc cgtatttccc ggaattcttt acgcccggata ttccagtgtc acgcagtgcg	84540
ttacggcgtg acgcactcaa agcttttcga aggcatacgc attccgcct cattataccc	84600
agccactgtt gttacatcgt tgttgtgtca cgataattca gaaatacgtc cggatcacc	84660
tttattatgg cacgatcggg attggatagg atcgacataa gcccccagcc agccaaaaaa	84720
attgcccggt tgggaggtct acagaccctt tttgtaaaaa cggatattaa cacgattaac	84780
gttgaaacacc attttataga cacgctacag aagacatcac cgaacatgga ctgtcgcggg	84840
atgacacggg gtatttttat tcgtttatcc cacatgtata aaattctaac aactctggag	84900
tctccaaatg atgtaaccta cacaacaccc ggttctacca acgcactgtt ctttaagacg	84960
tccacacagc ctcaggagcc gcgtccggaa gagttagcat ccaaattaac ccaagacgac	85020
attaacgta ttctattaac aatagaatcg gagactcgtg gtcaggggcga caatgccatt	85080
tggacactac tcagacgaaa ttaatacacc gcatcaactc ttaaatggag tgtatctgga	85140
cccgtcattc cacctcagtg gttttaccac cataaacta cagacacata cgggtgatgcg	85200
gcggcaatgg cgtttggaaa aaccaacgaa cggcggcac gagcagatgt tgaagcattg	85260
tttatagatc cggctgatat ccgtactcct gatcatttaa cgcagaagc tacaactaag	85320
ttttttaatt ttgacatgct caataccaaa tctccaagtc tccttgtggg tacaccaaga	85380
atcggaacgt atgaatgtgg acttttaatc gacgttcgaa cgggacttat aggcgcgtcg	85440
ttggacgttc ttgtatgtga cagggaccct ttaactggca ccctaaatcc ccacctgca	85500
gaaaccgaca tttcattttt tgaaattaaa tgtcgtgcta aatacctctt tgatccagat	85560
gacaaaaata acccgctcgg tcggacgtac accacgttaa taaatagacc tacaatggca	85620
aatctacggg actttttata tactataaaa aacctatgtg taagcttctt tggaccctca	85680
gcaaacccaa gtacacgcca ggccctaata acggatcacg ttgaatggaa acgtttagga	85740
tttaaggtg ggagggccct tacagaactc gacgccatc atttgggcct caatcggaca	85800
atctcatccc gagtgtgggt atttaatgat ccggacatac aaaaggggac aattacaacc	85860
attgcatggg ccaactggaga tacggctctt caaattcctg tatttgccaa tccgcggcac	85920
gtaacttta acaaaattgc cgtacaaacc tatgtattat cgggttactt tccagcgcta	85980
aaactacggc cttccttgt cacctttata ggacgtgtgc gccgaccaca cgaggtggga	86040
gtcccattgc gcgtcgatac acaagcggct gccatttacg aatataactg gccgactatc	86100
ccaccccact gtgcggttcc ggttatagcc gttctaacgc ctatcgaagt tgatgtgcct	86160
agagtgcac aaataactaa agacacagga aacaacgca ttacatcagc attgcggta	86220
ttgcgatggg acaatcttca tccagcggtc gaggaggaat ctgtggattg tgcaaacgg	86280
acaacgagct tgttacgtgc aacggagaaa ccgttgcttt gaactcagag ttctttgaag	86340

-continued

actttgactt tgatgagaat gtaacagagg acgccgataa atccacacaa cgccgcccac 86400
gagtgatcga tgtaacacca aaacgaaaac cttcgggaaa gagctcccat tccaaatgcg 86460
caaaatgtta aaccctgata aaccctgata aacgttctaa taaaaacatc aaatcatggt 86520
tggttactgt gaatgtttgt tttattgctt gggggtttac aagtacaacc cagctactc 86580
cccccactg tttgatcgct cgtataacag ctcatcctcg cgtccggtt catatgttga 86640
gtcattttca tagacgtagc cgtagccttg tgatgggtaa tttgtcggc gagaatttct 86700
atgtgcaggt tttacttttc gtatgtatcc cgtaccgcg tcgggtactc ttcttacggc 86760
accgtagaac cgactgcggt tctgtcgtg atacacatat gcacgcatca atctgagaag 86820
caacatgaca acggaanaa cggccaggca agccaagggt ccccgagttg tgggaattaa 86880
ccgtggagat tgaaccgata tagggtcata taatcgggtcc atatacgagt gcgcgggcgt 86940
tcccacgta gcacaggcca cgagcgttcc cagggacggt cctattaaca cgtgtatata 87000
atgcgcaaaa attaatctg atactataag atatacaact gacaatgtac taaatgtaga 87060
catggccacg gacaccgatg accacagtcc cgtatgtaga tgattcgcca ccacaagttc 87120
cagcattaat gatacaata ggatacatat cggcatcaac gcagccatca aattcacgaa 87180
cactgcgcgc gtaggccccg caagcgata taaaaagacg ctctgctgtc gtaaatttgc 87240
gaccgctttt atgttcggtt cgtccaattt tcccggtcca caaaaatcg ttgtaaatat 87300
tacactgtgc gcaaatgtc caagatataa tgtagcagcc acgocgattt gcttghtaagc 87360
taataataac acaacggcgt ttaataacca caatgacaaa agaccccaaa aaagtgttgt 87420
gggatctaca actaaccatg caacaccgga gctttgcccg acacgttgat ttttcgtttc 87480
tcggtgtata atcgcggccg tgatcagtg atataccgcc atggccattg ccgttaaagc 87540
cgtgtagtaa gtaaatgcca caacgctatg tggttccaaa aacaaaaccg gggcgctgta 87600
tccaccteta tttccggacc ataccccc atctaggggtg gcgttaaata actcataatc 87660
aactacggoa gcataaaaac aagggatccc ggtatattca gaagaggcgg caattaacgt 87720
agccaggagc attaccgcac ccaaagtga catcatcacc tgaattatcc aaattcgcca 87780
attaagcgta tccatttgat gatctaacgc ttccacctcg ggtgtcgtgg tgtcgtacgg 87840
cgagactttt tcagaacgcg gcccttctt ttgagttccc atgtctocca acaccgggga 87900
gagcaacgcc gccgtctatg cgtccagtac acagctcgcg cgggcgttat atggagggga 87960
tctggtttog tggattaaac acaccaccc gggaaatagc ctggaactgc aattggatgt 88020
tccagtaaaa ctaataaaa ctggtatgtc acaactcgc cggtaaccg tcgtacgtgc 88080
ccctatggc tctggtaaaa caacgcctt gcttgagtgg cttaaacag cgttaaaggc 88140
agatattagc gtactggtg tctcatgtcg ccgtagcttt acccagacgt tgattcaacg 88200
gtttaacgat gcaggcctct ccggattcgt aacatattg acatccgaga catatattat 88260
gggttttaaa cgtttgattg tgcaacttga aagcctacac cgcgtatcca gcgaagctat 88320
cgacagctac gacgtattaa tactggatga ggtaatgtca gtgattggac aattatactc 88380
ccccacaatg agacgtcttt ccgcggttga tagcctatta tatcgtcttt taaatcgtg 88440
ttctcaaat atcgcgatg atgtacagt aaactcgcag tttattgatt taatctccgg 88500
attgcgtgga gatgaaaaca tacacacaat tgtgtgtaca tacgcccggag ttgggttctc 88560
cggaagaact tgcacgatcc tgcgtgatat gggcatcgac acgcttgtgc gagtcattaa 88620
acgatctcct gaacacgagg atgtacgtac catacacaa ctacgtggaa cattttttga 88680

-continued

cgaactagca	ctacgattac	aatgtgggca	taacatctgt	atattttcat	caactttatc	88740
gttttcggag	ctagttgctc	agttttgtgc	aatatttaca	gactctattc	ttattttaaa	88800
ctcaactcgg	cccctatgta	atgtaaacga	atggaaacat	tttcgcgtgt	tggtgtacac	88860
taccgctcgtg	accgttggat	tgagttttga	catggctcat	tttcatagca	tgtttgctta	88920
cataaagcca	atgcatatg	ggccgatat	ggatcggtc	taccagtc	tagggcgtgt	88980
acgtttattg	ctacttaatg	aagttttgat	gtacgtcgat	ggctcaagga	ccagatgcgg	89040
accctgttc	tcgccaatg	tactaaactt	taccatcgca	aataaatttc	aatggtttcc	89100
tacacacacc	caaataacta	acaaactgtg	ctgtgcattt	aggcaacgat	gtgcaaatgc	89160
atttacacgc	tcgaaccccc	atctcttctc	aagattttaa	tacaaacacc	ttttcgagag	89220
atgctctctt	tgagtttag	ccgatagcat	taatatctta	caaactcttt	tgccctctaa	89280
ccaaattttg	gttgatttgg	atggcatggg	tccaataacg	gacgtttccc	cagttcaatt	89340
ttgtgcattt	atacacgac	tcagacatag	cgtaacgcc	gtagcttcc	gtatgcgttc	89400
tcttagacag	gacaatgaca	gctgcttgac	cgattttggc	ccttccggat	ttatggccga	89460
taacattacc	gcgtttatgg	aaaagtatct	tatggagtca	attaataaccg	aagaacaaat	89520
taaagtattt	aaagcccttg	catgtccaat	agaacagcct	agactagtca	atacggcaat	89580
attggggggg	tgtatacgaa	taoctgaagc	gttggaaagc	tttgacgtat	ttcaaaaaat	89640
atacacgcac	tacgcttccg	gttggtttcc	cgctctggac	aaaaccgggg	aatttagcat	89700
cgcgactata	actaccgccc	caaatttaac	cacacattgg	gagctgtttc	gccgttgtgc	89760
ctatatgtoa	aaaacactca	agtggaatcc	gtccaccgaa	ggctgtgtaa	cacaagtttt	89820
ggatacggac	attaatacac	ttttcaatca	acacggggat	tcgctggctc	aactaatatt	89880
tgaggttatg	cgctgtaacg	ttactgacgc	taagattata	ttaaaccgcc	cggtttggcg	89940
aacaaccgga	ttcttagatg	gatgccataa	tcaatgcttc	cgccaatcc	ctacaaaaca	90000
cgaataaac	attgctctat	ttcgtttaat	ttgggaacaa	ttatttggcg	cccgcgtaac	90060
taaaagtaac	cagacccttc	cggaaggtac	tcgtgtgaaa	aacctaaaaa	aaaaagatct	90120
agaaacttta	cttgattcaa	ttaacgtgga	tcgttctgca	tgtcgtacct	accgccagtt	90180
gtataacctg	cttatgagcc	agcgcatttc	gttctctcaa	cagcgttaca	aaattactgc	90240
ccccgcttg	gcacgccacg	tgtattttca	agcacatcaa	atgcacttgg	ccccgcatgc	90300
cgaagccatg	ctacaattag	cgctatcgga	actgtccccg	ggatcgtggc	cgcgataaaa	90360
cggggcggta	aattttgaaa	gtttataacc	cgtaataacc	atatatggac	atccataggg	90420
ggggttacat	aaataactaag	cctctgtaca	acacaaaggg	cctctaacaa	tgcaactgaa	90480
cacaaccaag	ctatggacgc	aacgcagatt	accttgggta	gagaaagcgg	acacatttgt	90540
gccgcaagca	tatacacatc	ctggacacag	tccggacaat	taacacagaa	cggcttttcc	90600
gtgttatact	acttattatg	caaaaactca	tgtgggaaat	acgtccctaa	gtttgocgaa	90660
attaccgtac	aacaagagga	tttatgtcgc	tactccaggc	atggggggag	tgtttctcgc	90720
gcaacgtttg	cgctatctcg	cagggcggcg	tcctcggctg	cgttagacgc	ctggccctt	90780
gaaccactgg	gtaacgcaga	caactggcgt	tgtctccatg	gcactgccct	ggccaactta	90840
cggcgctgat	tagggtttaa	atcgttttat	togccagtaa	catttcgagac	tgatacgaat	90900
acaggtcttc	tgtaaaaaac	aatccccgat	gaacacgcgt	tgaataatga	caacacgcca	90960
tctaccggag	tattgagggc	taattttccc	gtggccattg	atgtttcagc	agtcagcgca	91020
tgtaacgccc	acacgcaag	tacgtcgcta	gcctacgccc	gcctgaccgc	acttaaatct	91080

-continued

```

aacggtgaca cccagcaaca aacaccttta gacgtggagg taattacacc aaaggcctac 91140
atacgtcggga aatataagtc tacgtttttcc cccctatag agcgggaagg ccaaacctcc 91200
gatttgttta accttgaaga acgccgcttg gttcttagtg gcaatcgcgc aattgtggta 91260
agggtactct taccggtgta ttttgactgt ttaacaacgg attccaccgt tacatcttcc 91320
ctttcaatat tagcaacata tagactgtgg tacgcggcgg cgtttggaaa acccggggtt 91380
gtccgtccaa tctttgcgta tttaggcccg gaactcaatc cgaaggggta agacagagac 91440
tacttttgta ctgtcggatt tcccggatgg accactcttc ggacacaaac tccagccgtc 91500
gaatctattc gcacggctac ggagatgtac atggaaacgg atgggttggt gccagtaacc 91560
ggtattcagg cctttcatta tctagccccc tggggacagc atccccctt acctccgcgg 91620
gtgcaggatc ttattgggca aatccctcaa gatactggac atgcagatgc aactgtcaat 91680
tgggacgcgg gccggatata taocgtcttc aaacagcctg tacaactaca agatcgttgg 91740
atggcaaatg ttgatttcag cgctttttt cccacgatata actgcgctat gttccccatg 91800
cattttagat taggcaaaat cgtcctggct agaatgcgtc gaggaatggg gtgcctaaaa 91860
cccgcgttgg tgtctttttt tggggggtta cggcacatac tcccgagtat atacaaagct 91920
attattttta tagccaatga aattagcctt tgcgtcgaac aaacgcctt ggaacagggc 91980
tttgcataat gtacttatat aaaagatgga ttttgggaa tcttcaccga tttacatacg 92040
cgcaatgtat gttcagatca ggcacgttgt tcggccttaa atttagcggc cacctgcgaa 92100
agagcagtca cgggccttatt acgaattcaa ctaggcttta actttacacc cgccatggaa 92160
ccggtactcc gggctcgaggg tgtgtacact cacgcattta cctgggtgtac cacgggaagc 92220
tggctgtgga atttacaac aaacagcct cgggatttag ttggcgtgcc atggcgaagt 92280
caggcggcgc gagattttaa ggagcgtctt tcaggactcc tatgtaccgc aacaaaaatt 92340
cgagaacgga tacaggaaaa ttgcataatg gaccatgtcc tatacgacat atgggcccga 92400
caagttgtgg aggtcggcag aaaaacatac gtcgattttt ttgaacatgt ttttgatcgc 92460
cgttatactc cgttatactg gagtcttcag gagcaaaatt cggaaacaaa agcaataaccg 92520
gcatcttata tgacatacgg acacatgcaa gataaggatt ataaaccaag acagataatt 92580
atggttcgta atcccaacc acatggacct cctactgttg tttactggga attgctacca 92640
tcgtgtgcct gtattcccc catagactgc gctgctcacc tcaagcccct tatacacacg 92700
tttgtcacta ttattaacca tcttctagat gctcataatg atttttcaag tccatcattg 92760
aaatttactg acgatccctt tgcttcata aacttcttgt ttttatgaca aaaaacacg 92820
ccgcaacaac ccaccttaa aataaaagg ttatttactt tacaaccctt ggtgaatttt 92880
tatacgtttc aaataactga acatttttcg gtgttaccat ggtgcgattt aaccacaaa 92940
aatatacgtc cttctgatat tccgaatctc gtaaaggctc atttaacaat cccgggggta 93000
cttgaccacc accatctgga cagggggggg ttccgtgggg caggcaaaa cgctgacca 93060
ccccacatga atatatagcc tttataatat tgggggcccgt tccaggctga gggttcagta 93120
acttaacaaa catataatgc ggcaatcgc gggtttttgt aaaggggttg ttatcaacga 93180
catacattag agtgtttaac aaccataaaa ctccctcata taaaaaccga cgcatttttt 93240
ccaaaggctc tatttgacac tcaacgcgtc taagatatac agacaattgt acaaacagcg 93300
atggagatgc cccggagggc ccaatgcctt ccagatacat taaaataaca cataaggtaa 93360
aatctaggac attatccggg cggaatagag tcatccgata gattaacagc cgcggaggca 93420

```

-continued

```

ccccaccgt ataccacctt tcttcaaccg cagttaatac ggaaaaaata aatccgcgga 93480
acgctggttg agtaaacacac tccatgtagt aacgatcaca ggacacctca cttgaatcac 93540
cattcaacac tactaaaaacg gtctcttggg gttccggttt tacgcgcagt gatacaacag 93600
agtttgccaa aaagcgtggc ttcaaaccgg ttacctcccg cgcctcgcag acgaatcttg 93660
gtattgcttg tattctaaga tcttcgatca cgtcgctcac atccaacccc tcttcggctc 93720
gtgttagtaa gttgctgacg gttacgctgc aacctaaaat gctgggtata tttattccgg 93780
acatcccacg gcccatcccc gcgcctcccg tttgctcga ttttatccag taaggtcgaa 93840
tccgctgcat ttaccttgty taccgtaac ctctcagggg ggtgtccttt cataaaatgg 93900
gataggtttt tatatccaac atgcatgtag tggttattta ttttattggg ttccgggatt 93960
ctttcgtcat cttctgtagg gtcaggcaaa cccaggaag gacttgggtg tctccgtggg 94020
ccccgtttta ttacctctgc gcgaacctgc atttcatata atattcggat ttgggataaa 94080
taggactctg ttctcgctt ttaaaaaata gcctggcata actcttctc tgacctatgt 94140
acctcgcttt gagttaccaa gaactcctaat cgggtggccc gtaatatgaa tgaaaaatac 94200
ggcgcaacta gtaatgagat tgacgcattt gaatatgata cagaaatttc ctggccttga 94260
ttattgttta cccggtgaag cttaaacacg cgaacaagtt cctgtttcca tagctcagac 94320
aaacgtttta tatcatctcc ataagggggg atataacgag attgaaaact attggcaata 94380
tatgcatcat cccctattat gccggaaga tctataacct cgtgatttaa atcggaata 94440
cgtgtttctt ctgccattgt aatatgtgac ccttagatg gctttatttt taccctctct 94500
tcccgtaacg gtttcagctc tccttcttg aactggagcc tttcggtcag atcgtgttc 94560
acatccttga gacctcaat ggttttgaat aaattattca cataaccctc gagcatgccg 94620
ttgatactgt taaccaccga agttttaaac gcactttgaa cgtttgtgtg tccggacatt 94680
gccccccgt taaaggattg gttggccttg ccaaaccocg gttgtgatgt gtccaccgat 94740
ccacttcctt ccagaatgty attgcccgtt tctctagat aggaacgtac ggttccgta 94800
atatctccaa catgtctcat gttttttaag ttaactatta gctttacaag tctagacgcy 94860
gccgatccag cccgtgtgtg atcgttctcg cccattatac gatcaaccgc acgtgtgctg 94920
tgagatctat catcttcatt ccggcgacct attaacacgc gcaaaggggc tgtatttaaa 94980
acttggcaga cgcgagcatg ttcacgtaat gcataacagg ccaacacctc cccagaaagc 95040
cgctgtaagg gtgagcaaa tactacacct tccccacata caacggggcg ccacacgacc 95100
aaacactctc cctcatgccc cgttacatca tcctttgcca taattaatct tcggttataa 95160
ttataataaa gacgcgtcct atcataatcc ataatagcaa cattttgcat aactcaact 95220
aggcttgta caaccgccc tcctctggcc aacgttgcat cggcaacttt taacatctgg 95280
gacagttctg ccgcttgacc catatacgtt ttaaatgtyg caggggttcc attctgttct 95340
gatcgtacct ttcttacaac gggcacaata cctacacagg ctatccagtc cacgtatttg 95400
gcaaaaccga cccttcatt taaaccactg gtatagagac aaccggttat tccacgcaga 95460
aactcaagta acgatgacty taatgtttga cgccaggttt caaaaacctg atgtgcaagc 95520
cgtacggctt ctgattctcc acatagccca taacgttccg cttagagccc gccatgcagg 95580
ttacattggt ggatgtggty ttcccaatct gotgctaggt cctcataccg agttgcatcc 95640
aacgcgttca tcaaaacggt tgctgaact tggcgaatta cagtttccgt agaccgtaca 95700
gcgctatata tgcttctgccc atcgttatat ccaaagtcac cggctaggat ttttcgaaac 95760
aacatacttt gcgtggttyg gtgtattaac atccagccat ctctctccg aaatgtacaa 95820

```

-continued

```

aaccctatat cggggcgta ctcatccag tatatatcga acatgttctt gtattggca 95880
tttgggttac ttccattcaa gccctggca atagaaacag aacttgctat cctttttct 95940
tcactaccgg aactgttatt aaaaagagac gttatttcgg ccattgaaaa ccacgatgaa 96000
aagatcaatt tctgtagaca gttcttcacc caaaaacgtt ttaaatccag agacgccaa 96060
tggatttgat gacagtgtat atttaaactt cacctctatg catagcattc aacctatcct 96120
ctcacggatt cgagaacttg ccgcaattac gattccaaaa gaacgtgttc cgcggttg 96180
ttggtttaa cagtactcg aactgcaagc gcctcctgaa atgcagagga atgagctccc 96240
cttctccgtt tatttaatta gcggaatgc cggctccgga aaaagcacgt gtatccaaac 96300
gcttaacgaa gctatcgatt gcattattac cggatccacc agggttgctg cccaaaatgt 96360
tcatgctaag ttatcaacgg cttatgagag tcgtccgata aacacaatct ttcatgaatt 96420
tggttttcgc ggaatcaca ttcaggtcga gctgggcccgt tacgcatata actggactac 96480
gacccccct tctattgagg acctgcaaaa aagagatatt gtatactact gggaaagttt 96540
aattgatata acaaacgagc tgtttcaaat gggggacgac ggtcgcggag gaacatcgac 96600
atntaaacc ctgtgggcaa ttgaacgttt gottaataaa cctacaggct caatgtccgg 96660
aaccgcgttt atcgcatgcy gttcccttcc ggcttttacc cggagcaacg ttattgttat 96720
tgatgaagca ggattgctag ggcgctatat tctcacggcc gttgtttact gttggtggct 96780
tttgaatgot atatatcaaa gccctcagta cataaacggt cgaanaaccg tcatagtatg 96840
cgtcggttcg cccaccocaa ctgactcgtt agaatctcat ttcaacatg acatgcagcg 96900
ttcacacgta actcctagtg aaaatatact cacgtatata atctgcaatc aaactctcgy 96960
tcaataatac aacatctcac ataactgggc aatctttatt aataacaaac gatgtcaaga 97020
ggacgatatt ggaatcttt taaaaacgct tgagtacggg ctacctatta ccgaagcaca 97080
tgccgctctg gtcgatacat ttgtgttacc tgcacctat attaacaatc ctgctaactc 97140
tcccggatgg acgctctctg attcgtcgcga taaggaggty agcgcgtata tgagtaagtt 97200
acacgcgcac ttaaaactat cgaanaaatga ccattttct gtgtttgcct taccgactta 97260
tacattcatc cggctaaccg catttgatga ataccgcaa ttaacgggac aaccggact 97320
ttctgttgaa cattggatac gggcaaacct cggctgtttg cacaattatt cccaagccg 97380
agatcatgac atgggaacag ttaaaatcga aacacattca aatcgcgact taattgtagc 97440
ccgtacagac atcacttacg tgctaaatag tctcgtagtt gtaaccacaa gactacgtaa 97500
gttagttatt ggattcagty gtacatttca atcgtttgca aaggttttac gtgacgactc 97560
ctttgtgaag gctcaggagg agacatccat cgaatatgct taccggttcc tgtcaaacct 97620
aatctttgga ggcttgatta acttttaca tttttgtta aataaaaacc tacatcccga 97680
taaggtatcg ttagcataca aacggttagc tgccttaacc ctggagtat tgtctggaac 97740
aaacaaagcc cccttacagc aagcagcggc taatggggcg ggtgccggga ttgactgtga 97800
tggtgacgct acttctgccg ataaagcctt ctgctttacc aaagccccg agtccaaagt 97860
aacgcccctc ataccogaag acccgatga tgtaattttt acggcactta acgacgaggt 97920
tattgactgt gtatactgcc agtacgaatt ttcctatccc aaatcatcca atgaggcca 97980
tgctcagttt ctgttaatga aagctattta cgatggcga tatgccatat tagcagagct 98040
tttcgaaagc agctttaca cgcgccctt tagcgcgtat gtcgataatg ttaatttcaa 98100
cggaagcgag cttttgatcg gcaatgtgcy ggggggctg ttatctttg cattacaaac 98160

```

-continued

agatacgtat acccttttgg ggtatacttt tgcaccctgt ccagtccttg tagaggaact 98220
 gacccgaaaa aagctgtacc gcgaaactac cgaatgtta tatgctctac acgtacctct 98280
 tatggtctta caggatcaac atgggtttgt gtccatcgta aacgtaacg tatgtgaatt 98340
 taccgagtct atagaggatg cagaattggc aatggccacc acggtggact atggccttag 98400
 ttctaaacta gccatgacaa ttgcacgctc acagggctctg agtttagaga aggtagctat 98460
 ctgttttacg gcgataaac tgcgcctaaa tagtgtgtat gttgccatgt cgcgtacggt 98520
 ctctcttagg ttcttaaaaa tgaatctaaa ccctctacgg gaacgatatg aaaaatccgc 98580
 agaaattagc gatcacattc ttgccgtctc acgtgatccc aacgtacacg ttgtgtatta 98640
 aagcattgta taaaacacg catgcgggct tgctgttctc atttctaggt tttgtcttaa 98700
 atacaccgcg catgagcadc tctggacccc caacgacgtt tattttatat aggttacatg 98760
 gggttaggcg ggttcttcac tggactttac cggatcatga acaaacactc tacgcattta 98820
 cgggtgggct aagatcaatg gcggtgaaga cggacgctcg atgtgatata atgagcggtg 98880
 gtatgatcgt cttcaaacac accatcacag tgaccctgct aacctagac tgttctactg 98940
 acttttcatc atacgcattt acgcaccggg atttccactt acaggacaaa cccacgcaa 99000
 catttgcgat gccgtttatg tcctgggtcg gttctgacc aacatctcag ctgtacagta 99060
 atgtggggcg ggtactatcc gtaataacg aagatgacct atccatgtgt atctcaattg 99120
 ttataacg tttacgggta aacagacctg acgatcagac cacaccaaca ccaacccgcg 99180
 accagtatac atcgaaagg cggcagcctg aaaccaactg tccttcttca ccacaaccg 99240
 cctttttcac atcagacgac gaogttcttt cgtaaatatt acgggacgcc gcaaacgct 99300
 aaagacagat tcaagactaa catttatccc aactgattac atttcatac cgaataaacg 99360
 acacaaaaaa tttatattta acggctttta atttgaagac acctatctc ttaacgttga 99420
 tgagccttgc aggttgggtg ccgctcttca ccggtattat acataaccga tttaccgtgt 99480
 ttaccgcagt ctgaccattt accagtgtat gtctgtaata cgacgttgtt gtgtccgac 99540
 aaaattaact cgcgtacaaa tttctgatgt tccccggcg tggcaacgct ggcatttcca 99600
 aacacattac gttctcgtac gtccatgacc gctattttca gtattaattg gttggtcgg 99660
 caaagtattt tccttatgta aaaggacacg atctaaagcc gtaaacctcat acacaaacac 99720
 tggtaaccaac ggacgcgatt ttccgtccgt tgagcgggtg taatatcggc gaggtcttct 99780
 tgcacgaata ctctcgtaca gtaggtttct gacacggggt gcatggggtt tttgacacaa 99840
 cacaaacatt tgcaggctct tatgactgga tggattgaat ttatttttag atagggtcac 99900
 gtgtttttgt cgtgacacgc ctcgaccaga aaaggctgcg gttttcgtac acgacaccgt 99960
 tatttcacag gcgttcataa ccaagctgcg gcggatgggt tcggtaatt gtctccgcc 100020
 aagttcgtca atagatgata ccatgaacaa cgtatcaaat ggtacatagt cgtctttggt 100080
 tttctcaata cagcccgcgt gcccaatcgg aaatttttca tttgcatcaa cgctattttc 100140
 tgtaaaatcg ttctgaacac tgtgttggct ggctacctgt ttaaaatttg ggtcgaaca 100200
 cgggtccacga tgcaatcccc aaccocattg aagcaatgcc gtcggtacg aaggaggcaa 100260
 ctccgaaac attatggtac gcaagagggt cgattggagt gttatataac actccaatcg 100320
 atctcgggtt cgcctttacg cgtaaaatac tcattggctt gaacgaaatg tcgacaattc 100380
 cgaaatggaa cacgggacaa tggcgacgga tgcgctgtg ttagcaccag atgacatctt 100440
 gaattcgggt gggttgtctt ctgtgcacgc gcacccaca gcataaaaac taaccctgta 100500
 cggttctcgc ataacctctg tagcacgcgt tgcaccagcc gcccccagcc taagtataca 100560

-continued

tgcgaccccg gagtcccgcg acgaaccgta agcgtggtat tcagcaataa caccocctgc 100620
 cttgcccaac tctccaggca tccgtgagtg ggcggagtca tatttgggta tgattccatg 100680
 agggccgcaa aatattttt aagactagac ggtgggtgta tgccacgttt taccataaac 100740
 gctagcccat gtgcatgtcc cgcggtaggg tatggatcct gaccaataat tacaacgcga 100800
 atgctctggy gtcgcaaaa tcgctccat gcaaaaatat cgcctgtaga tggaagtatt 100860
 tcttccocctg aatttaaaag acgattgtat tctaaaaaa tacctttcgc gtacggctct 100920
 ttaagttcgt ccgacaacag gtcataccac tcaggggaaa tgttaactt gctgaaaact 100980
 tcaaccgaat ccagttgcga agagacgggg gtgaacgttt ccgtgtcgta atgatgtgac 101040
 atgttattta acttgaaggt tggggggtct agcttaacc ccaaaggcag cccgcggggt 101100
 cgcttgcggg ttttttgggt aaccggatgg gccaaaacat aatgtcctt tgaatccgat 101160
 agtttcattt cattggcata cgcgttgaa caaacgctcg gctcccaga cacatccatt 101220
 ttccgggata tttgtggaag atggagtaga gtctaccat acaccgaaa gggcatccaa 101280
 caaagcatcg cgtatgtccc cgttttatg ttcttacca acagattgtg ccagcccctt 101340
 taagtgagc tagtgatttg tccagtacgc catttgttg tctttaaacc aaagtataac 101400
 ttccggtact ggacattttg tcttaaccac gattcccgat agcgcctcgc tgaggtttga 101460
 taccgggggt gccgcatagt cccacgcctc atataccgat gacacgcacg gttccggtat 101520
 aatcaaactc acatccgata gcggtttggc tcaaaaaaac aacggagtgt cgtcttgag 101580
 atgaagacia tacgcatggt tgatagtttt taaaaaaact atctgcagta accatttatg 101640
 tgatgccatg acgcttgtgt tttcccttca ctacgacgtt gtcgtatcct ttgaaaaact 101700
 tgaccactot aatggaagca tggacaagta tgagttttat atatacagtt ggcctttagt 101760
 taaactctgt gtgcatatc tcattttcct aaaaaggcgc atcttaatat gtcaaacgtc 101820
 acggcgtgcc gacaaaagca atttccatgc aagatttggg ttagtattt atacaccaa 101880
 tcacatgtoa cgtattaagc tttacagtc cccgttatct gatataatca cttttottaa 101940
 cacgtcatcg gaaaaacaga tgtttatatt atacctctcg cggtcattta cggcaaac 102000
 ttagaccgtt tcaagcggg ctgaaaacgc tcaaatgcc ttttgaggc ctgcccaacg 102060
 gccattatcc cttggatcta agattgattt gcggtaacgt ttgccaatca agctttaaaa 102120
 acgtacccca aacttaaac gctcaaatg ccttttgag gcctgccca cggccattat 102180
 cccttgatc tgagattgat ttacggtaac gtttgccaaa cccacgcatt tcagtttaaa 102240
 tatttctaag cattcttagt gcgtacttgg cagcgtgctt aaaatatcaa ccaatatcca 102300
 ttatgctaca cgtttccttc tatccgtttc aatccattaa aagtcatta acaaaaatga 102360
 tgcatcatac ctaattcacc taaaaacctg actcattgca gcagcgtttc ctccttgag 102420
 actatccagt tggcatttta aacgggtccg gctgcctaaa ccgaaaacac cgttgctttt 102480
 actgtaagta caaaactaaa atttatattt gcgtgcgtat tttgtaacat atatgccttt 102540
 tatcccccg caagtttget ttaccctcgc cttcaccacc cccgccacct tccggccatt 102600
 ttaataactt taattgctat aagacatacc caaacggat gatttttgcc gctggaaaaa 102660
 cagcttctaa ttttcccgtc tcaactcggc cttggttga tctcaagta tacctttagt 102720
 ttgctcccg agaggtgtat aaatacaaac ggtgacaagt attgagcgtat atctcaaatt 102780
 tttgtaattt agggcggagc gcttacgaca gcacatcgt actgtagac tgttatgttt 102840
 attgtatttg cagagcagga tgcccgggtt actccgagac cggattgcgg gcattccgaa 102900

-continued

tcgtgtacgg acttaccag gggcagtatt tacaccttgg gttccagata taccaaccct 102960
 tacgaccaat agcaacactc aggtatTTTT aaaatgcacg tttaatgatc ataatttaca 103020
 tacagttggt aataaagcag actgtggatg ttttaaggcat ttcttcccc ctccaacaa 103080
 actaggactt cttcatcttg tttggaatac ctttaccgcg tttaccggca gagctTTTT 103140
 tggtaagggt tttcagtga cctgatgttg atccggagggt ggagggggta ttggactccc 103200
 cctgtggaga ggcaactttg cgggttttac ttcccttaca tgccgaatca gactcagatg 103260
 tcaggtctat tgtaagcat cgtttaacgt ctctgccggt atgaaataaa cggcgcttag 103320
 caccocctgc gcttcccggt ttaatccccg gtaacacaga aaaaagcctg actTTTTggg 103380
 gtgtatttac caatcgggta tccctttcat cgccacgaga ggtctccccg gttgaggtgg 103440
 tttctggtct tacaattgga cctgtaatta gttggatggc tgtatcttcc caggccagg 103500
 tttgcatggt taggcggggt ggtcgggtac atcgatcaa caagaataac atgtttgta 103560
 caaacggctc tgttgaatca tgcaaaagac aacgcaggga tgtttttaa cccgcctcat 103620
 cagccccgta aatacctata tagtttaata tcaacatTTT tgtaggctct acaatttcgg 103680
 gttgatacag ttccgcaagt tgatcatcaa gccatccgag taaagggtgc atgtaacacg 103740
 ggaatctcgc gtttccctct gttcctctat ccgtggctcg aaaaggcagt ctgtccatgg 103800
 ttcgtgggtc ttgattaatt cccacagata ctggacgatc acggtagtcc tgccccccgg 103860
 tccgggggtg ctgtgcagat tcaatcgagc catacaccac cggggtcgcc gatcgaacag 103920
 caggttggtc tttaaaaat accttccgta aaaatgatgc ggtagagcat gttttggta 103980
 caccagggtc cgagtctcgg gtcgggtggt gtatagaatc ctggtgagag tcaactggtg 104040
 actctgctgt gggctctcta gccgacgatt gaaggggcc agggtttggg gattgaatgg 104100
 gctccccact cgatcttgat gttggctggt ggatggactc ccgactcggg cctgggcttg 104160
 gtggcagaag atctatgaca tctcccggta ggatgtcgat ggaatcttca aatgacggct 104220
 cagaaaaacc atcgtcgtcg gatgggtgca cttcatattc cttgtaactt gtatcactta 104280
 cgatcttatg caggatggat tgcactggac accggcagag aggacactgg acgctgggtg 104340
 aggtccatgc ccgaatacaa acaaagcaga agtcgtgcaa acacggcatg gttttccga 104400
 gatcggaaac ggtgctcatg catatggtgc aggtattatc cgaagcgtcg gagtgccgc 104460
 taccgccgcg taatatggta tccatggtaa caactggctg tattctaag tccgggcac 104520
 caaacacgta gcagaactgc catgcttctt aaattgtgag ttgtggcgag tacattttta 104580
 taattggtac caacgaagac acaccctat atccctccac ccatttcttt taagtccac 104640
 ccactaaaac gtgggtataa aatgtgtatt ggggtaggcg gacagtcca acaaacaggg 104700
 aagttgattg gtataacctt gggccgggta tacagctaag tgacatttta gattctgtct 104760
 ttatttagat aaagagcgt acgaagacat ttctccacc ccctgtaata cccgtaaata 104820
 aagtaagtc cacaaacaaa agcactgtat ataggaagtc ggggtattg ggacagttac 104880
 tccattagag gcgtacaac aatactggga tagggtaatg caagtcccc ccgatggctg 104940
 ccccgcaaac gcgcggggag gtgggtcgcg ttttttttt ctctctcgag gggcgccgca 105000
 gagggctggc ctctctccc ggggtccgcg gggcgcccag aaaccgggg ggggttattt 105060
 tcgggggggg gtccgaccag ccgcgccgct gcccgcccgc acagacagac agacactttt 105120
 ttcataaaaa ccgttccgct tttattaaca acaaacagtc cgcgcgccag tggcgctcac 105180
 gagaaaagga ggggactccg tcacccccga ctctgcgggg ggtcctccc cccgcgcct 105240
 cccacacat cgtcctcgtc ctcgaggac gaggacgagg acaacagctc caccttgacc 105300

-continued

gcccggcgca aaccacccc gcggtctcgc agcacaccgc gggccaccga cacgatgctc 105360
acccc aaagg atgaccccc tgcgtcccgc tgcgtcccgc ccccctctc gctgtccac 105420
gcgtcttcac accccacctc ccaatcgtcc agctccaaag cgtgttctct gtcgtctcgc 105480
gtgcgccgct gtcgccccgc ctgggtttct gacggcgtt ccgagccccc gtggtgtccg 105540
aacacgaacc gtgtccgc gtccccccc aacaccgtct ccgcccgc aaaaccgggc 105600
ggccacatta ctctgggaat cggggggagg gcattccgag cctcgtccgc cgacgcatac 105660
agccccaccg accgaccggc cacgggtgga agcacagtg gttctcggc agggctcggg 105720
tccagcaggc cgtggcgga aacaccctc gccaggtgg gtacgtcgc gccctccggc 105780
ccggcggccc ccggtctcc tccctcggga aggaagcgc gtcgaagcgc ggcaccagg 105840
ccccatcgtt ttgtcgcgc gtggctatgt gccgcctcgt ccacaaagtc ggctgcccc 105900
agccccagac ccgagactc tcgcccagg tccttgcaac cgtcaaaacc cggcagcacg 105960
tactgccggt attcacgggg cgacaggggg acgcccgtct tggggcccgc gcgggtacac 106020
acggtgtatg cgacgttccc accgcggcac aaacacagg gttgttcgcc cgggtacagg 106080
ttggcaaacg cagtctcgt acgagcaaaa ctgcctggc caaaggtgc cgacgatgca 106140
aacacggccc gggcagatcc ttctgtgacc gccgagctc gccatcggac gacgcctgg 106200
gcgtccggtc gcgcccgggc ccggacgtac acgtgatact gagacaaagc gggccatcc 106260
ctgggccacc tctcaggggc caccgcctc aacaccagca accggcggcc ggcagaggcc 106320
aaccgcgagc ctagatactc gacggcccc gcaaaggcca ggtctcgggt cgacagtaat 106380
aaaacgcccc gggcgttcaa agcggacacg tccggcgggc cggctcagtt cccggcccag 106440
gcatgagtgc tcggcagga caaccggtta ctacgggctc ccaggaccac agacagtccc 106500
cctcgggatg gactccatga cggctccgga tctgtcgcga ggggtctctc gagggggccg 106560
ttgatgtcct ctccgggcaa cggatcgtag atgatcagaa gcctcacatc ctccgggtct 106620
gggatctgoc gcatccaggc gcacctccgt cgcagccct ccaactccgt ggggggacca 106680
aaccgtcgtt ctctccgcc cggaccccga gcgcccgttt ccgccaaggc gccgggatca 106740
aagcttagcg cagggcgcca ggcctggga aacaatgggt cgtcgaccag acggcggatg 106800
gttccggggg tacagtacgc ctgcccagcc tgggtccgac ggaccggggt atgcagggcc 106860
ccccgggaa taccccgaaa tccccggtt gggcccggtc cgtcaagtgg catcgttatt 106920
acggcggggg gatccaccac agggcccag gtgatggtca cgggctcggg taccgcctc 106980
ttggccttgg aaaccacatg atcgtctgca acccggcgt ccgcccaggg tgtctcccta 107040
atcttctcga ggagcctct gctctcact ggtgggact tgcgcttgc cggagtctg 107100
aaacgatcat ccggtggaca cacagaaaga gagcgtcgc cggcccaggc ctgagggctg 107160
ggagcctgtg tggcccgggt tgttgagaa gggtgaccgc gggagatccg ccccgcggga 107220
ctggagccc ttgcctcgg gtatgccatg ctggcaagg ctctcggag actctgtagg 107280
ataaagtgtt tttggcccgc gtcgtatcga cggctcatag ccacggccc ggcgcgtgg 107340
gggagagccc agagggcctc ccccgtggc atggtctcgc ctacatcggg aacgggagac 107400
gctacgtcc ccgtaaccgc ggtaccgcc cgtcccgtg gcaacagctt ttggtagaac 107460
tggttcaggc ccgagttgac accggtcagc ttgggttct ggagccatgc tataggtct 107520
ctgtctggac agtagatcag gtaaatcagc gcgcccact gctagccg atctccaac 107580
tccggcacgt aaagcggcac gggttccgtt gaggcctcgt aacgagccc cgcgcctctc 107640

-continued

acagcctcat cctcccagtg accctctctg gtctccccgg acggtccaaa ccgcaccctg 107700
 ttgatggga ggggtgccga tccgggccaa gggcttccgt cgggcatcat gagcggcccc 107760
 gacaccgggg gaattatcgg ggttctggat cgcggcaggg aaaatgattt ctgtctctgg 107820
 cgccccggtt cccccgcaag acgtttggtc ttacgaatcc tcggatcggg accgctgatg 107880
 gatcgatata ccggttggat attttgttct gtcgaccac catcatttga gtccgaatca 107940
 tccgaatttg acggggaag ggcgtgttcg cgtccggacc tgctgcctgt agtttcaact 108000
 cccaccgaaa cgcgccgggg ttcactctct tcactctccg atgacgatcc ccacgacgag 108060
 gaagaggatg aagacgaaac aaactcacga ctctttggct ttttctccac tgggctgtca 108120
 tctcaatcgg ggtctgtgtc gtgggatctt cccggcaggg ccaaaaacgc tctaggtttg 108180
 cccccgcaag aacgtccagg gacgcgaggt gttatacccc gggcatcatg tttccttggg 108240
 cgggtatcat cgtctcaaa cggcaggtcc gcctttgccc ccttagcggg aacgctgtcc 108300
 gaaaggacgt ggtacaattg ctcaaccggg ccgggtacag gtccaccggg tttccgcgcc 108360
 gggagtggga ccttaacctt caaagtcttt ttcttcgggc tctttcctct agcgggcccgt 108420
 tgagttttct ggagaactac tccgtccccg gatgcatgcg catgaccgcg ttgctcatcg 108480
 cccgcctttt taccgagat ggactgagtt tgtctgtctc gatggaccac cgacggcaaa 108540
 cctggtgaat ttctctctct cgtttgtcgg ggtatagacc gctggtcttc ccggtgatcg 108600
 ttcccggcgg cgtctccaac aggagacgcg ggggatacag gggagaagc ctgcgggaac 108660
 ggaggggtcg tacctctgcc cgtttcccca tcgttcatcg gtggtttttg agacctagca 108720
 agcttcgttc cgagagagac tgtctcaagg gagcgcgctc ctctgtttg tttctcgcgcg 108780
 ccggcctcgg agaatcgggt gtggaagacc tcggccagcg ggattacag cgagcccatt 108840
 agatcctgac cgtcctcgca tacgtagtcg tcttgtgtta gctcttcgcc aacatcttc 108900
 gttctgggtt ctggttgaag tcccgatcag gaggaattg aaacgatctc gtgttcccgt 108960
 cccaccatga ccccgttctc tccaaatagt agatcgtcag gctgactcga ggtgaccacc 109020
 cgggcccctg gttcggcggc cgcgcgggcc gcgtccaaca ggtccattaa ctccaaagta 109080
 tcaggcgacc ccgcgcgctt ggggtgtagag cgtgcatcgc gcggcgtatc catcgcactg 109140
 ggggtgaatt agacgtaccg gagttttcca aacgctctcg cagccttcaa aggattgcga 109200
 ttgcggttgg tgagggagtt ccaacagtac ttaaacgtg ttgtgcccc ccctcgaccg 109260
 catatttctt ccccgctgct tcaccgtgta aatattctta atgataagac gatgtagtga 109320
 ttggacgaga ctcgaggcgg gaagtctcat gaccatagta tgcgtttaag gagagaccgc 109380
 tggttggcga tgtacgcccg gtgtctatct ccgcatacct tacaacatca taacaaggga 109440
 taccagacat gtgaatttca ttacatatg tttaaataac aaccaatcat cgtgtgtcta 109500
 cagacgatata ataataata taaacacaat tggggttgtc tcacatgcaa aacatcttat 109560
 ataacacggg ttgtttccac ccacccggca tctagttaat caaatgcacg tcgacggtgt 109620
 gtttgggtcc ctctccgctg tcattacgtt cgcgcaatca acaagcgtat acaccaccac 109680
 cctcccaac gattatgtca ggcggcacga agcccgcgat aaccataaa atacacacgg 109740
 ggttgtggtg ttcacgtaac cccccgccga tggggagggg gcgcggtacc ccgccgatgg 109800
 ggagggggcg cggtaacccc ccgatgggga gggggcgcgg taccocgccg atggggaggg 109860
 ggcgcggtac ccgccgatg gggagggggc gcggtacccc gccgatgttt ataaccataa 109920
 ttctctaaac cgtttagtaa aatcacaaaa aaatttattc aaaaacaagt cgaagaactt 109980
 catatctgag gcatgtaaac ccggttcgcac ttcctggggg ggaatggggg ggggtggggg 110040

-continued

ggtgaaaaag ggggggggtt aaattgggcg tccgcatgtc tgtggtgtac gccaatcggg 110100
 tacactcttt tgatctgcat tcgcacttcc cgttttttca ctgtatgggt tttcatgttt 110160
 tggcatgtgt ccaaccaccg ttgcactttt ctttctatat atatataat atatataat 110220
 atatataag aaagagagag agtttcttgt tcgcgctgt tcccgcgatg tcgcggtttt 110280
 atgggggtgt ggcgggcttt tcacagaata tatatattcc aaatggagcg gcaggctttt 110340
 taaaatcgat ttgactgtat aaaaaaaaa acacggggcc cccccctttt tttggtgtta 110400
 taaaggcaac ccaatcgaag gtctcccgc cgggaatccc ccattgccat tttaccaag 110460
 tagccttatt catagatgta aacgtttggg tgtgtgtttt gttgtgcagg gttcgtccga 110520
 ttcataacgc gacagcgtcg agtcggtttt aagggaaaag gttactacgg cccaaggac 110580
 atgttttga cctcaccggc tacgcggggc gactcgtccg agtcaaaacc cggggcatcg 110640
 gttgatgta acgaaaagat ggaatatgga tctgcaccag gacccctgaa cggccgggat 110700
 acgtcgcggg gccccggcgc gttttgtact cggggttggg agatccacc gcccaggctc 110760
 gttgaggaca tcaaccgtgt ttttttatgt attgcacagt cgtcgggacg cgtcacgcga 110820
 gattcacgaa gattgcggcg catatgcctc gacttttata taatgggtcg caccagacag 110880
 cgtcccactg tagcgtgctg ggaggaattg ttacagcttc aaccaccca gacgcagtgc 110940
 ttacgcgcta ctttaatgga agtgtcccat cgaccccctc ggggggaaga cgggttcatt 111000
 gaggcgcga atgttcttt gcataggagc gcaactggaat gtgacgtatc tgatgatggt 111060
 ggtgaagacg atagcgacga tgatgggtct acgcatcgg atgtaattga atttcgggat 111120
 tccgacgcgg aatcatcggg cggggaagac tttatagtgg aagaagaatc agaggagagc 111180
 accgattctt gtgaaccaga cggggtaccg ggcgattgtt atcgagacgg ggatgggtgc 111240
 aacaccctgt ccccaaagag accccagcgt gccatcagc gatacgcggg tgcagaaacc 111300
 gcggaatata cagccgcaag agcgcctacc gcgttgggcg aggggggtgt agattggaag 111360
 cgacgtcga acgaagcccc gcgccggcat gatataccgc ccccccattg cgtgtagtct 111420
 ttataaataa atacaatggt ttggctcgtg tctttttttg atgtctgtct gtgggggagt 111480
 ggggtgttgt ggatattaga gggtagaggg tgctggtttg aacgtctcca ttaaccacg 111540
 gggccccac acgggccgtg ttggtatgaa ctctcgggat cccgcgggta gcaccgggc 111600
 ggtgaatatg ccggaactta ctgcacacga cacgataccc ccgcgacca ggtctctatg 111660
 aacgacgcgg aacgggtactt cgcgcggcgt ctatgcgcca tatctaccga ggcctacgag 111720
 gcttttata acagcccctc cgagagaccg tgcgcgagtt tgtgggggag ggcaaaggac 111780
 gccttcggac ggatgtgcgg ggagctcgca gcgtagac aacgtccacc ctcggttcg 111840
 ccgatccgca gagcgggtgt atcgttatta cgcgagcaat gcatgccgga tccacaatcg 111900
 catctgagc tcagcgagcg gctgatattg atggcatatt ggtgctgttt gggacacgcc 111960
 ggacttccga ctattggatt gtcgcccgat aataaatgca tccgcgccga attatatgac 112020
 cgccccggg gaatttgca caggcttttt gacgcgtacc tgggctgagg gtcccttggg 112080
 gtccaaagaa cctacgagag atcctgacac cccatccctt tatatagaaa aaaaaataa 112140
 atttaaaaca tacaccggat aaaagcgtac tgttttttat ttaaatttac acgctcggcg 112200
 ttgccccggt tcggtgatca ccgggtctta tctatataca ccgtgtaact cgaacccccg 112260
 tgactcccct caatcgcggt accaaactct tcttcgctat ccgtagattc cgagtccctg 112320
 aaatcgtcca cttatccaac aaattgtgac gttatataac ccaaggcaaa ggcgctccc 112380

-continued

gtcatagcaa atacaaagac aattattagc gtaatataac agaatttttt acgatgatat 112440
 attttatggt gatattttcc aattcgagc aaaaattcat ctgccgtttc attttcgcta 112500
 tcactataat aacacttttc agccgaacgg ctcggttgta tggctgttat cgttgattata 112560
 ttggttgctg ctcgagggtg taccaccgct tccatcagta aggccacggc ctcaccctcc 112620
 atggtgtttt gtccggccat agaaatccag attgtaaggc cagcaggcta gtttaaaagt 112680
 gtttaatacc acaccttttg atatttatat acatgcaaga ttctagatta ttcatcaata 112740
 ggtcgtttta agcgcgtttt cataaacgtt gtcagctata cgcacattct cacaaagagg 112800
 taaagttacc ttacgttatt attaaataaa acatgtagac attattaata atcctaggaa 112860
 caatcaaacc catatttgta agttatgttt aaccctccc ctttttgca ttatctccgc 112920
 cctctataat tcggatcact ttataagtgt gtcggtgagt atattttgta cagttgttg 112980
 acaacagggt ttggttcat taacactatc aacataagtc ggggtataca agtataatga 113040
 acgacgttga tgcaacgagc acctttgttg gacaaggaaa gttccgtggc gccatctcaa 113100
 catcaccgct acatattatg caaacatgty ggtttataca acagatgttt ccagttgaaa 113160
 tgtcgcocgg catagaatct gaggatgato ccaattatga cgttaacatg gatatacagt 113220
 cttttaatat atttgatggt gtacacgaaa ctgaagccga agcctctgtg gcattgtgcy 113280
 cagaagcagc cgttggaatt aataaagcgg gatttgtaat attaaaaacg ttacaccag 113340
 gggcggaagg ttttgcgttt gcgtgtatgy acagtaaaac atgtgaacat gtggtcatta 113400
 aagcgggtca acgtcaagga acggccaccg aggcaaccgt gtttaagagcy ttaaccacc 113460
 catcogttgt acagcttaaa ggaacgttta cgtataacaa aatgacatgt cttatattac 113520
 cacgttaccg aacagattta tactgtatc tagctgcaa gcgcaacctc cccatagtgy 113580
 acattttagc aattcagcga tctgtattac gcgcgttaca gtatcttcat aataacagta 113640
 ttattcacgc tgatataaaa tctgaaaata tatttattaa ccaccaggt gatgtttgtg 113700
 tgggagactt tggagcagcy tgtttccocg tggatattaa tgccaacagc tattatggct 113760
 gggctggaac aatcgcaca aactctcctg agttattggc tagagatcca tatggacctg 113820
 ccgtggacat atggagtgcc gggattgtat tatttgaat ggctacagga cagaactcgt 113880
 tatttgaacy agacggttta gatggcaatt gtgacagtga gcgtcaaatt aaacttatta 113940
 tacgacgac tggaactcat cccaatgaat ttcccattaa ccctacatca aatcttcgcy 114000
 gacaatacat tggtttgcca aaacggtctt ctcgaaaacc cggatccagc ccattgtgga 114060
 caaatctata tgagttgcca attgatttgg agtatttgat atgtaagatg ttatcgtttg 114120
 acgcacgtoa tcgaccatca gcagaggtgt tgcttaacca ctctgttttc caaactcttc 114180
 ccgatccata tccaaatcca atggaagttg gagattaaaa ttcattaagc ctgttaataa 114240
 aatattgtat aaattgtgtt tataacgtat aaccggttaa ggcaaatagc gtacaaacgc 114300
 gcaatgtttt gaaatactaa tataaataac ataaccaata gaaacttaac acagagtcac 114360
 gccccattac aacaaggata aaacacggga tcattttctt aacattgtag tagcgtgaa 114420
 aagcgtcccc tccccggct cacagagctg ctctcgtgcy tagttgggta tactggtgcy 114480
 cctcatttaa tcgcatggtt tttaatccaa tgtttgatc cggccgttat atttacata 114540
 caagtgacca acgctttgat cttcaagggc gaccacgtga gottgcaagt taacagcagt 114600
 ctcacgtcta tcctatttcc catgcaaat gataattata cagagataaa aggacagctt 114660
 gtctttattg gagagcaact acctaccggg acaactata gcggaacact ggaactgtta 114720
 tacgcggata cgggtgocgt ttgtttcccg tcagtacaag taataagata gcacgagatg 114780

-continued

```

ccccggatta gaacgagcgc ttttatttcg tgtagggtaca aacattcgtg gcattatggt 114840
aactcaacgg atcggatatac aacagagccg gatgctgggtg taatgttgaa aattaccaa 114900
ccggaataa atgatgctgg tgtgtatgta cttcttgctt ggttagacca tagcagatcc 114960
accgatggtt tcattcttgg tgtaaatgta tatacagcgg gctcgcata caacattcac 115020
ggggttatct acacttctcc gtctctacag aatggatatt ctacaagagc cttttttcaa 115080
caagctcgtt tgtgtgattt acccgcgaca cccaaagggt cgggtacctc cctgtttcaa 115140
catatgcttg atcttcgtgc cggtaaatcg ttagaggata acccttgggtt acatgaggac 115200
gttgttacga cagaaactaa gtccgttggtt aaggagggga tagaaaatca cgtatatcca 115260
acggatatgt ccacgttacc cgaaaagtcc cttaatgatc ctccagaaaa tctacttata 115320
attattccta tagtagcgtc tgtcatgatc ctaccgccca tggttattgt tattgtaata 115380
agcgttaagc gacgtagaat taaaaaacat ccaatttata gcccaatac aaaaacaaga 115440
aggggcatac aaaatgcgac accagaatcc gatgtgatgt tggaggccgc cattgcacaa 115500
ctagcaacga ttcgcgaaga atcccccca cattocgttg taaaccggtt tgttaaatag 115560
aactaattat cccgattttt atattaaata aactatatgc gttttattta gcgttttgat 115620
tacgcgttgt gatatgaggg gaaggattaa gaatctccta actataagtt aacacgcca 115680
catttgggag gggatgtttt atgaagcctt aaaggccgag ctggtatata cgagagcagt 115740
ccatggtttt agacctcggg cgaattgcgt ggttttaagt gactatattc cgagggtcgc 115800
ctgtaatatg gggacagtta ataaacctgt ggtgggggta ttgatgggtt tcggaattat 115860
cacgggaacg ttgctgataa cgaatccggt cagagcatcc gtcttgcgat acgatgattt 115920
tcacaccgat gaagacaaac tggatacaaa ctccgtatat gagccttact accattcaga 115980
tcatgcggag tcttcatggg taaatcgggg agagtcttcg cgaaaagcgt acgatcataa 116040
ctcaccctat atatggccac gtaatgatta tgatggattt ttagagaacg cacacgaaca 116100
ccatgggggt tataatcagg gccgtgggat cgatagcggg gaacggtaa tgcaaccac 116160
acaaatgtct gcacaggag atcttgggga cgatacgggc atccacgtta tccctacgtt 116220
aaacggcgat gacagacata aaattgtaaa tgtggacca cgtcaatac gtgacgtgtt 116280
taaaggagat cttaatccaa aaccocaaag ccaaaagactc attgagggtg cagtggaaga 116340
aaatcacccg tttactttac gcgcaccgat tcagcggatt tatggagtcc ggtacaccga 116400
gacttggagc tttttgccgt cattaacctg tacgggagac gcagcgcggc ccatccagca 116460
tatatgttta aaacatacaa catgctttca agacgtgggt gtggatgtgg attgcgcgga 116520
aaatactaaa gaggatcagt tggccgaaat cagttaccgt tttcaaggta agaaggaagc 116580
ggaccaacog tggattgttg taaacacgag cactactgtt gatgaactcg aattagacc 116640
ccccgagatt gaaccgggtg tcttgaagt acttcggaca gaaaaacaat acttgggtgt 116700
gtacatttgg aacatgcgag gctccgatgg tacgtctacc tacgccactg ttttggtcac 116760
ctggaagggt gatgaaaaa caagaaccc tacgcccga gtaactctc aaccaagagg 116820
ggctgagttt catatgtgga attaccactc gcactgtatt tcagttgggt atacgtttag 116880
cttgcaatg catcttcagt ataagataca tgaagcggca tttgatttgc tgttagagtg 116940
gttgatgtgc cccatcgatc ctacatgtca accaatgcgg ttatattcta cgtgtttgta 117000
tcatcccaac gcacccaat gcctctctca tatgaattcc gttgtacat ttacctcggc 117060
acatttagcc cagcgtgttg caagcacagt gtatcaaaat tgtgaacatg cagataacta 117120

```

-continued

caccgcatat tgtctgggaa tatctcatat ggagcctagc ttgggtctaa tcttacacga 117180
cgggggcacc acgttaaagt ttgtagatac acccgagagt ttgtcgggat tatacgtttt 117240
tgtggtgtat tttaacgggc atgttgaagc cgtagcatac actggtgtat ccacagtaga 117300
tcattttgta aacgcaattg aagagcgtgg atttccgcca acggccggtc agccaccggc 117360
gactactaaa cccaaggaaa ttacccccgt aaaccccgga acgtcaccac ttctacgata 117420
tgccgcatgg accggagggc ttgcagcagt agtactttta tgtctcgtaa tatttttaat 117480
ctgtacggct aaacgaatga gggttaaagc ctatagggta gacaagtccc cgtataacca 117540
aagcatgtat tacgctggcc ttccagtgga cgatttcgag gactcggaat ctacggatac 117600
ggaagaagag tttggtaaag cgattggagg gagtcacggg ggttcgagtt acacggtgta 117660
tatagataag acccgggtgat caccgaaccg gggcaacgcc gagcgtgtaa atttaaataa 117720
aaaacagtac gcttttatcc ggtgtatggt ttaaatttat ttttttttc tatataaagg 117780
gatgggggtg caggatctct cgtaggttct tgggactcca agggaccgc agcccaggta 117840
cgcgtaaaa agcctgtgac aaattcccc gggcgggtca tataattcgg cgcgatgca 117900
tttattatcg ggcgacaatc caatagtcgg aagtccggcg tgtoccaaac agcaccaata 117960
tgccatcaat atcagccgct cgtgagctc cagatgcgat tgtggatccg gcatgcattg 118020
ctcgcgtaat aacgataaca ccgctctcgg gatcggcggg accgaggggtg gacgttgtct 118080
atccgctcgg agctccccgc acatccgtcc gaaggcgtcc tttgcctcc cccacaaact 118140
cgcgacaggt ctctcggagg ggctgtgtat aaaagcctcg taggcctcgg tagatatggc 118200
gcatagcggg gggcggaagt accgttcggc gtcgttcag agagcctggt gcgcggggggt 118260
atcgtgtcgt gtgcagtaaa gtccggcata ttcaccgcc ggggtgctcac cgcgggatcc 118320
gcagagattc ataccacacg gccctgtgtg ggaccccgtg ggttaatgga gacgttcaaa 118380
ccagaccctc ctaccctcta atatccacaa caccocactc cccacagac agacatcaaa 118440
aaaagacacg agccaaacca ttgtatttat ttataaagac tacacgcat gggggggcgg 118500
tatatcatgc cggcgcgggg ctctgtgtcg acgtcgcttc caatctacac ccccctcggc 118560
caacgcgggt agcgccttcg cggtgtata ttccgcggtt tctgcaaccg cgtatcgctc 118620
gatggcacgc tggggtctct ttggggacgg ggtgttgcac ccatccccgt ctcgataaca 118680
atcgcggggt acccgtctg gttcacaaga atcgggtgctc tcctctgatt cttcttcac 118740
tataaagtct tcccgtccg atgattccgc gtcggaatcc cgaattcaa ttacatccga 118800
tggcgtagac ccatcatcgt cgctatcgtc ttcaccacca tcatcagata cgtcacattc 118860
cagtgcgctc ctatgcaaag gaacattcgg cgcctcaatg aaccctctt cccccgagg 118920
gggtcgatgg gacacttcca ttaaagtagc gcgtaagcac tgcgtctggg tgggtgaaag 118980
ctgtaacaat tcctcccagc acgctaactg gggacgctgt ctggtgcgac ccattagata 119040
aaagtcgagg catatgcgcc gcaatcttcg tgaatctcgc gtgacgcgtc ccgacgactg 119100
tgcaatacat aaaaaaacac ggttgatgtc ctcaacgagc ctggccgggt ggatctcca 119160
accggagta caaaacgcgc cggggccccg cgacgtatcc cggccgttca ggggtcctgg 119220
tgagatcca tattccatct ttccgttaac atcaaccgat gccccgggtt ttgactcggg 119280
cgagtcgcc cgcgtagccg gtgaggtgca aaacatgtcc ttggggccgt agtaaccttt 119340
tcccttaaaa ccgactcgac gctgtcgcgt tatgaatcgg acgaacctg cacaacaaaa 119400
cacacacca aacgtttaca tctatgaata aggtacttg ggtaaaatgg caatggggga 119460
ttccggggcg ggagacctc gattgggttg cctttataac accaaaaaaa ggggggggccc 119520

-continued

ccgtgtgttt ttttttatca cgtcaaatcg attttaaaaa gcctgcccgt ccatttgaa 119580
 tataatatatt ctgtgaaaag cccgcccaca ccccataaaa ccgcgacatc gcgggaacac 119640
 gcgcgaacaa gaaactctct ctctttctct atatatatat atatatatat atatatatat 119700
 agaaaagaaag tgcgaacggt ggttgacac atgccaaaac atgaaaaccc atacagtga 119760
 aaaacgggaa gtgcgaatgc agatcaaaag agtgtatccg attggcgtag accacagaca 119820
 tgcggacgcc caatttaacc cccccctttt ttcaccccc caccaccacc cattccacc 119880
 caggaagtgc gaacgggttt acatgcctca gatatgaagt tcttcgactt gtttttgaat 119940
 aaatTTTTTT gtgattttct acaacgggtt agagaattat ggttataaac atcggcgggg 120000
 taccgcgcc cctcccctc ggcggggtac cgcgcccct ccccatcggc ggggtaccgc 120060
 gcccctccc catcgcggg gtaccgcgc cctcccctc cggcgggta ccgcccccc 120120
 tcccctcgg cggggggtta cgtgaacacc acaaccctg gtgtatttta tgggttatcg 120180
 cgggctctgt gccgcctgac ataatcgttg ggaggggtgg tgggtatac gcttgtgat 120240
 tgcgcgaacg taatgacgac ggagaggac ccaaacacac cgtcgactg catttgatta 120300
 actagatgcc ggatgggtgg aaacaaccg tgttatataa gatgtttgc atgtgagaca 120360
 accccaattg tgtttatgta tattatata cgtctgtaga cacacgatga ttggttgta 120420
 tttaaacata tgtaaatgaa attcacatgt ctggtatccc ttgttatgat gttgtaagg 120480
 atgcggaat agacaccgg cgtacatgc caaccagcg tctctccta aacgcatact 120540
 atggtccatg aacttcccgc ctcgagtctc gtccaatcac tacatcgtct tatcattaag 120600
 aatatttaca cggtagacac acggggagga aatagcggg cgagggggg gcacaacacg 120660
 ttttaagtac tgttgaaact ccctaccaa ccgcaatgc aatccttga aggctcgag 120720
 agcgtttgga aaactcgggt acgtctaaat tcaccccagt gcgatggata cgcgcccgat 120780
 gcagcgtct acacccaac gcgcggggtc gcctgatact ttggagtaa tggacctgt 120840
 ggacgcggc gcgcgggccc ccgaacacag ggcccgggtg gtcacctga gtcagcctga 120900
 cgatctacta tttggagaga acggggtcat ggtgggacgg gaacacgaga tcgtttcaat 120960
 tccctccgta tcgggacttc aaccagaacc cagaacggaa gatgttgcg aagagctaac 121020
 acaagacgac tacgtatgac aggacggtca ggatctaag ggctcgcctg taatcccgt 121080
 ggccgaggtc ttccacacc gattctcgga ggccggcgc cgagaaccaa caggagccga 121140
 tcgctccctt gagacagtct ctctcggaac gaagcttgct aggtctcaa aaccaccgat 121200
 gaacgatggg gaaacggca gaggtacgac ccctccgtc ccgacgcct tctcccctgt 121260
 atcccccg tctcctggtg gagacccgc cgggaacgat caacgggaag accagcgtc 121320
 tataccccga caaacgacga gaggaattc accaggttg ccgtcgggtg tccatcgaga 121380
 cagacaaact cagtccatct cgggtaaaaa gccggcgat gagcaagcg gtcacgcga 121440
 tgcatcggg gagcgagtag ttctccagaa aactcaacgg cccgctcag gaaagagccc 121500
 gaagaaaaag actttgaag ttaagggtcc actcccggc cggaaacccg gtggacctgt 121560
 acccggccc gttgagcaat tgtaccact ccttcggac agcgttccc ctaagggggc 121620
 aaagcggac ctgcccgtt agaccgatga taccgccca aggaaacatg atgcccggg 121680
 tataacacct cgcgtccctg gaogttcgtc gggggcaca cctagagcgt ttttggccct 121740
 gccgggaaga tcccacgac cagaccgat tgaggatgac agcccagtgg agaaaaagcc 121800
 aaagagtcgt gattttgtt cgtctcctc ctcttcctc tcgtgggat cgtcatcgga 121860

-continued

ggatgaagac gatgaacccc ggcgcgtttc ggtgggaagt gaaactacag gcagcaggtc 121920
 cggacgcgaa cacgcccctt ccccgtcaaa ttcggatgat tcggactcaa atgatgggtg 121980
 gtcgacgaaa caaaatatcc aaccgggata tcgatccatc agcgggtccc atccgaggat 122040
 tcgtaagacc aaacgtcttg cgggggaacc ggggcgccag agacagaaat cattttccct 122100
 gccgcatcc agaaccocga taattcccc ggtgtcggg ccgctcatga tgcccgacgg 122160
 aagcccttgg cccggatcgg caccctccc atccaacagg gtgcggtttg gaccgtccgg 122220
 ggagaccaga gagggtcact gggaggatga ggctgtgaga gcgggcggg ctcgttacga 122280
 ggcctcaacg gaaccocgtc cgtttactgt gccggagtgg ggagatccgg ctagacagta 122340
 ccgcgcgctg attaacctga tctactgtcc agacagagac cctatagcat ggctccagaa 122400
 cccaagctg accggtgtca actcggccct gaaccagttc taccaaaagc tgttgccacc 122460
 gggacgggcy ggtaccgccc ttacggggag cgtagcgtct cccgttccgc atgtaggcga 122520
 agccatggcc acgggggagg ccctctgggc tctccccac gcggccgcy cctgggtat 122580
 gagccgtcga tacgaccggg cccaaaaaca ctttatccta cagagtctcc gcagagcctt 122640
 tgccagcatg gcatacccc aggcaacggg ctccagtcgg gcggcgcgga tctccgcgg 122700
 tcacccttct ccaacaaccc cggccacaca ggctcccac cctcagccgt cggccgcgc 122760
 acgctctctt tctgtgtgtc caocggatga tcgtttacga actccgcgca agcgaagtc 122820
 ccagccagtc gagagcagaa gcctcctcga caagattagg gagacaccg tcgcygacgc 122880
 ccgggttgca gacgatcatg tggtttcaa ggccaagagg cgggtatccg agcccgtagc 122940
 catcacctcg gccctgtggy tggatcccc cgcgtaata acgatgccac ttgacggacc 123000
 ggccccaaac gggggatttc ggcgtattcc cggggggccc ctgcataccc cggctccgtc 123060
 ggaccaggct cgaagcgtc actgtacccc cgaaccatc gcccgctctg tcgacgacc 123120
 attgtttccc acgcccctgg gcctcgcct aagctttgat cccggcgcct tggcggaaat 123180
 cggcgtcgg cgtccgggcy gaggagacc acggtttggt ccccccagcy gaggggaggc 123240
 gctgcgacgg aggtgcgctt ggatgcggca gatcccagac ccggaggatg tgaggcttct 123300
 gatcatctac gatccgttgc ccggagagga catcaacggc ccctcagaga gcacctcgc 123360
 gacagatccg ggaccgtcat ggagtccatc ccgaggggga ctgtctgtgg tccctggcagc 123420
 cctgagtaac cggttgtgcc tgccgagcac tcatgcctgg gccgggaact ggaccggccc 123480
 gccgacgtg tccgcttga acgcccggg cgttttatta ctgtcgacc gagacctggc 123540
 ctttgccggg gccgtcagat atctaggctc gcggttgccc tctgcccgc gccggttgct 123600
 ggtgttgac gcggtggccc tcgagaggty gccagggat ggaccgctt tgtctcagta 123660
 tcacgtgtac gtccgggccc cggcgcgacc ggacgcccag gccgtcgtcc gatggccaga 123720
 ctcggcggtc acagaaggac tcgcccgggc cgtgtttgca tcgtcgcgca ccttggggc 123780
 agcgagtttt gctcgtatcg agactgcgtt tgccaacctg taccgggcy aacaaccct 123840
 gtgtttgtgc cgcggtggga acgtcgcata caccgtgtgt acccgcgcy gcccaagac 123900
 ccgctcccc ctgtcgcgcc gtgaataacc gcagtagctg ctgccgggtt ttgacggtyg 123960
 caagacctc gcgcgacagt ctcgggtctt gggctcggg gcagccgact ttgtggacga 124020
 ggcgccacat agccaccgcy cagcaaacgg atggggcctg ggtgccgcy ttcgaccgt 124080
 cttccttccc gagggacgga gaccggggc cgcgggccc gaggccgcy acgtaccac 124140
 ctggcgagg gtgttttgc gccacgcct gctggaacc gacctgcy cagaaccact 124200
 cgtgcttcca cccgtggcy gtccgtcgtt ggcctgtat gcgtcggcy acgaggctc 124260

-continued

```

gaatgccctc cccccgattc ccagagtaat gtggccgccc ggttttgggg ccgcgggagac 124320
ggtgttgagg gggagcgacg gaacacggtt cgtgttcgga caccacgggg gctcgggaacg 124380
gccgtcagaa acccaggcgg ggcgacagcg gcgcaccgca gacgacagag aacacgcttt 124440
ggagctggac gattgggagg tggggtgtga agacgcgtgg gacagcgagg aggggggagg 124500
ggacgacggg gacgcaccgg ggtcatcctt tggggtgagc atcgtgtcgg tggccccggg 124560
tgtgtctgca gaccgcccgg tgggtttgag cccggcggtc aagggtggagc tgttgcctc 124620
gtcctctgoc tccgaggacg aggacgatgt gtggggaggg cgcgggggga ggagcccc 124680
gcagagtcgg gggtagcgga gtcccctcct tttctcgtga gcgccactgg cgcgcggact 124740
gtttgttgtt aataaaaagc gaacggtttt tatgaaaaa gtgtctgtct gctcgtgcgg 124800
gcgggcgacg ggcgggctgg tcggaccccc ccccgaaaat aacccccccc cggtttctgg 124860
gcgcccggcg gacccccgga gagg                                     124884

```

```

<210> SEQ ID NO 77
<211> LENGTH: 12
<212> TYPE: DNA
<213> ORGANISM: Varicella zoster

```

```

<400> SEQUENCE: 77

```

```

ctgcagatag tt                                     12

```

```

<210> SEQ ID NO 78
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Varicella zoster

```

```

<400> SEQUENCE: 78

```

```

Leu Gln Ile Val
1

```

```

<210> SEQ ID NO 79
<211> LENGTH: 12
<212> TYPE: DNA
<213> ORGANISM: Varicella zoster

```

```

<400> SEQUENCE: 79

```

```

ctgcagatga ta                                     12

```

```

<210> SEQ ID NO 80
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Varicella zoster

```

```

<400> SEQUENCE: 80

```

```

Leu Gln Met Ile
1

```

What is claimed is:

1. An isolated polynucleotide comprising the nucleotide sequence of nucleotides 115,808 to 117,679 of SEQ ID NO:76, with the proviso that nucleotide 116,255 is an adenine.

2. The polynucleotide of claim 1 wherein the polynucleotide is isolated from a varicella zoster virus.

3. An isolated polynucleotide consisting essentially of the nucleotide sequence of nucleotides 115,808 to 117,679 of SEQ ID NO:76, with the proviso that nucleotide 116,255 is an adenine.

* * * * *

UNITED STATES PATENT AND TRADEMARK OFFICE
CERTIFICATE OF CORRECTION

PATENT NO. : 6,528,066 B1
DATED : March 4, 2003
INVENTOR(S) : Grose et al.

Page 1 of 1

It is certified that error appears in the above-identified patent and that said Letters Patent is hereby corrected as shown below:

Column 8,

Line 4, please delete "other wise" and insert -- otherwise --.

Column 19,

Line 35, please delete "Avall" and insert -- AvaII --.

Column 22,

Line 50, please delete "subdloning" and insert -- subcloning --.

Column 23,

Line 21, please delete "S4000" and insert -- S-4000 --.

Column 34,

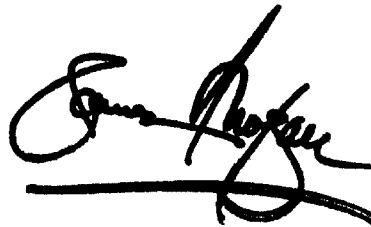
Line 6, please delete "striing" and insert -- striking --.

Column 36,

Line 54, please delete "linases" and insert -- kinases --.

Signed and Sealed this

Twenty-second Day of July, 2003



JAMES E. ROGAN
Director of the United States Patent and Trademark Office