CHARACTERIZATION OF THE GENOME OF BABOON CYTOMEGALOVIRUS STRAIN (OCOM4-37) ISOLATED FROM THE OLIVE BABOON, *PAPIO CYNOCEPHALUS ANUBIS*.

By

Susan R. Neubauer

Bachelor of Science in Biological Sciences
The University of Tulsa
Tulsa, Oklahoma
1990

Master of Science in College Teaching
Northeastern State University
Tahlequah, Oklahoma
1996

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2011

CHARACTERIZATION OF THE GENOME OF BABOON CYTOMEGALOVIRUS STRAIN (OCOM4-37) ISOLATED FROM THE OLIVE BABOON, *PAPIO CYNOCEPHALUS ANUBIS*.

Dissertation Approved:

Earl L. Blewett

Dissertation Adviser

Richard Eberle

Gregory W. Sawyer

Alexander J. Rouch

Mark Payton

Interim Dean of the Graduate College

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS USED IN THIS WORK

2YT	2x yeast/tryptone
7TM	Seven transmembrane
aa	Amino acid
ADCC	Antibody-dependent cellular cytotoxicity
AGM	African green monkey
AMP	Ampicillin
BAC	Bacterial artificial chromosome
BaCMV	Baboon cytomegalovirus
BLAST	Basic local alignment search tool
bp	Base pair
BPB	Bromophenol blue
Brn-3.0	Brain-3 (a transcription factor/POU domain)
BSA	Bovine serum albumin
C-terminal	Carboxy terminus
CCL22	Chemokine ligand 22
CCMV	Chimpanzee cytomegalovirus
CHO	Chinese hamster ovary
COX-2	Cyclooxygenase-2
cpe	Cytopathic effects
CRE	Cyclic AMP response element
DIG	Digoxygenin
DMEM	Dulbecco's modified Eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DrCMV	Drill monkey cytomegalovirus
dUMP	Deoxyuridine monophosphate
dUTPase	Deoxyuridine triphosphatase

E	Early (genes)
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
EtBr	Ethidium bromide
EtOH	Ethanol
gB	Glycoprotein B
gC	Glycoprotein C
gD	Glycoprotein D
gH	Glycoprotein H
gL	Glycoprotein L
gM	Glycoprotein M
gN	Glycoprotein N
gO	Glycoprotein O
GCPR	G-coupled protein receptor
GPS	Genome priming system
h	Hour(s)
HCF	Host cell factor
HCMV	Human cytomegalovirus
HFF	Human foreskin fibroblasts
HSPG	Heparan sulfate proteoglycan
HSV-1	Herpes simplex virus type 1
HveA	Herpesvirus entry mediator A
HveB	Herpes virus entry mediator B
HveC	Herpesvirus entry mediator C
ICP	Infected cell polypeptide
IE	Immediate early (genes)
IL-10	Interleukin-10
INR	Insulin-like receptor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KAN	Kanamycin
kbp	Kilobase pairs
kDa	Kilodalton

KSHV	Kaposi's sarcoma herpesvirus
L	Late (genes)
LAP	Latency-associated promoter
LAT	Latency-associated transcript
LB	Luria broth
LTR	Long terminal repeat
MCMV	Murine cytomegalovirus
MCS	Multiple cloning site
MDBK	Madin-Darby bovine kidney
MHCI	Major histocompatibility complex 1
m.o.i.	Multiplicity of infection
MW	Molecular weight
N-terminal	Amino terminus
NCBI	National Center for Biotechnology Information
ND10	Nuclear domain 10
NFB	Neural F-box
NF-κB	Nuclear factor kappa B (cells)
NLS	Nuclear localization signal
NS	Not significant
OBP	Origin binding protein
Oct	Octamer
ORF	Open reading frame
oriL	Origin long
oriS	Origin short
OUHSC	Oklahoma University Health Sciences Center
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
Pfu	Plaque-forming unit
PI	Post infection
PK	Protein kinase
RE	Restriction enzyme

RhCMV	Rhesus cytomegalovirus
RR	Ribonucleotide reductase
RT	Room temperature
r*Tth*	Recombinant *Thermus thermophilus* DNA polymerase
SDS	Sodium dodecyl sulfate
SOC	Super optimal broth with catabolite repression
SPLF	SalI-P fragment near L/S segment junction
SSB	Single-stranded DNA binding protein
STR	Short terminal repeat
TAE	Tris-acetate EDTA
TAF	TBP-associated factor
TBE	Tris-borate EDTA
TBP	TATA-box binding protein
TE	Tris-EDTA
TFIID	Transcription factor IID
TG	Trigeminal ganglia
TK	Thymidine kinase
TNR	Tumor necrosis factor
TNFR	Tumor necrosis factor receptor
UL	Unique long
UNG	Uracil N-glycosylase
US	Unique short
USF	Upstream transcription factor
UV	Ultraviolet
VAP	Viral attachment protein
vCOX-2	Viral cyclooxygenase-2
vFcγR	Viral fragment of crystallization gamma receptor
vhs	Viral host shut-off protein
vICA	Viral inhibitor of caspase-8-induced apoptosis
vMIA	Viral mitochondria-localized inhibitor of apoptosis
VP	Viral protein

CHAPTER ONE

INTRODUCTION

In this project I cloned, sequenced, and analyzed the genome of baboon cytomegalovirus (BaCMV) strain OCOM4-37. Genomic analysis is an important aspect of basic science that allows us to examine the coding sequence of an organism, and predict the protein coding regions and possible protein function. In addition, we can compare the genomes of different organisms, in this case primate cytomegaloviruses, and determine how closely related they are to one another and what characteristics they share. Analysis of other cytomegalovirus (CMV) species shows that these viruses evolve along the same lines as their primate hosts. In other words, the more related two primate species are to one another, the more related are the viruses they carry. After isolation, cloning and sequencing the coding sequence of the BaCMV OCOM4-37 genome, comparisons were made with other CMV genomes. These analyses showed that the OCOM4-37 strain is most closely related to CMVs isolated from primates most closely related to baboons.

After sufficient growth was observed (demonstrated by cytopathic effects), BaCMV was isolated from infected cells and the viral DNA purified. The DNA was cut with restriction enzymes and cloned into *Escherichia coli* using plasmid and bacterial artificial chromosome (BAC) vectors to create gene libraries. These libraries were

screened to identify unique clones that were then sequenced. Sequences generated from these clones were assembled using software designed for this purpose, and in this way large contiguous DNA sequences were generated. Sequence analysis was performed using the online database provided by the National Center for Biotechnology Information (NCBI) using the Basic Local Alignment Search Tool (BLAST) to identify other primate CMV sequences that were closely related to the BaCMV sequence. Sequence identified as homologous with human cytomegalovirus (HCMV) strain AD169 was used to construct a genomic map for BaCMV OCOM4-37. Sequences of other primate CMVs (such as clinical strains of HCMV or rhesus CMV) also showed similarities with BaCMV that were not shared with HCMV AD169. In its entirety, this research has confirmed that BaCMV is a member of the CMV family and that it is most closely related to the drill monkey CMV. When compared to herpesviruses whose entire genomes have been sequenced, BaCMV is most similar to rhesus CMV (RhCMV).

Since herpes simplex virus type 1 (HSV-1) is the most-intensely studied herpesvirus, the following review focuses primarily on this virus as the prototype for all herpesviruses. Where applicable, HCMV is also described since it is more closely related to the BaCMV which is our virus of interest.

CHAPTER TWO

REVIEW OF LITERATURE

I. THE HERPESVIRIDAE

Herpesviruses are large, double-stranded DNA viruses that are classified in the family *Herpesviridae*. This family encompasses a wide range of viruses that are ubiquitous in the animal kingdom, and is divided into alpha (α), beta (β), and gamma (γ) sub-families. Recently, the family *Herpesviridae* has been split into three families: *Herpesviridae*, *Alloherpesviridae*, and *Malacoherpesviridae* (Davison et al. 2009). The mammalian, bird and reptile herpesviruses were retained in the new *Herpesviridae* family. Within the α-herpesvirus subfamily are the genera *Simplexvirus* and *Varicellovirus*. These viruses are grouped together based on their wide host range, short reproductive cycle, rapid spread in cell culture, efficient destruction of infected cells, and ability to establish latent infection in sensory ganglia (Roizman 1996). The β-herpesvirus subfamily includes the genus *Cytomegalovirus*. Viruses in this sub-family typically have a restricted host range, grow slowly in cell culture, and establish latency in monocytes (Pignatelli et al. 2006; Taylor-Wiedeman et al. 1991), smooth muscle, and/or endothelial cells (Jarvis and Nelson 2006). Cells infected with these viruses often enlarge, hence the name cytomegalovirus (literally, large-cell virus). The two genera of the γ-herpesviruses are the *Lymphocryptovirus* and *Rhadinovirus*. Characteristics of viruses in this subfamily include replication in lymphoblastoid cells in

vitro, specificity for either T or B lymphocytes *in vivo*, and often latency in lymphoid tissue (Roizman 1996).

II. VIRUS STRUCTURE

Compared to most other viruses, the herpesviruses are large and have complex genomes. All herpesviruses have four basic structural components: 1) a core consisting of a single, double-stranded DNA molecule, 2) a capsid, which is an icosahedral structure composed of protein subunits, 3) a tegument which is an amorphous region located between the envelope and capsid, and 4) an envelope that contains virus-specific glycoprotein spikes (Roizman 1996). The CMVs have the largest genomes of all herpesviruses.

A. The Genome

While herpesviruses are comparatively large, the size of the genome varies from 120 to 230 kilobase pairs (kbp) within the family. In intact virions, the single linear dsDNA molecule is packaged in the shape of a torus, i.e. doughnut-shape (Furlong et al. 1972) while DNA extracted from a virus takes on a linear conformation. After infection in cell culture, however, the linear DNA enters the host cell nucleus and rapidly circularizes (Poffenberger and Roizman 1985). Depending on the size of the genome, between 60 and 120 genes may be present. These genes are usually classified as being essential if they are necessary for growth in cell culture, or nonessential if they are expendable. However, genes categorized as nonessential for *in vitro* replication are often required for survival of the virus *in vivo* (like immune system suppressor genes or genes that enhance cell-to-cell spread). In herpesviruses, genes may be coterminal, overlapping and/or coded on opposite strands. A linear map of the HSV-1 genome depicts

covalently-bonded unique long (UL) and short (US) sequences which represent most of the coding sequence of the genome (Figure 1A).

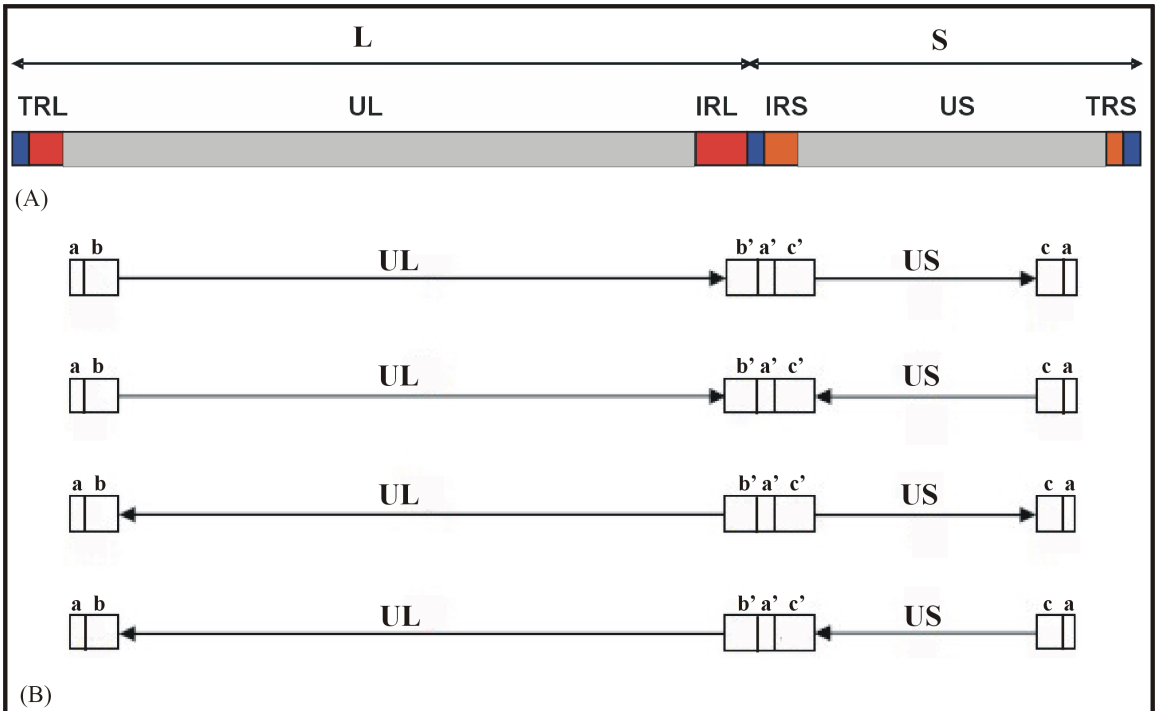


Figure 1. Linear Map of the HSV1 Genome. (A) Arrows designate long (L) and short (S) regions (top). Shaded/colored boxes show unique long (UL) and unique short (US) coding sequences, terminal repetitive (TR) long and short regions, and internal repetitive (IR) long and short regions. (B) Illustrations of the isomers of HSV1. Letters a, b, and c designate repeat sequences and arrows show the orientation of the UL and US segments. (Adapted from Shen and Nemunaitis 2006).

The L and S sequences are both flanked by direct or inverted repeat sequences which are either located within the genome at the L/S junction (internal) or at the genome ends (terminal) (Wadsworth et al. 1975). Because the L and S sequences invert at repetitive sequence junctions during replication, four isomers exist for herpesviruses as shown in Figure 1B (Hayward et al. 1975). Additionally, the number of repeats within junctions varies between individual herpesviruses and may account for size differences between closely-related species. Often, the ratio of guanine and cytosine to adenine and thymine is used to determine the relatedness between organisms. For herpesvirus

genomes, the G+C ratio ranges from 32-75% (Roizman 1996). Within the repetitive regions are sequences that are involved in DNA packaging called *pac1* and *pac2* (Figure 2). These sequence motifs are conserved in herpesviruses and located within the terminal repeat sequences of the genome (Deiss et al. 1986). The *pac1* motif is composed of an A-rich region flanked by poly-C runs while *pac2* consists of a CGCGGCG near an A-rich region (McVoy et al. 1998).

Figure 2. *pac1* and *pac2* Sequences in HSVI. Boxes represent long and short direct repeats (DR) terminal sequences of the 152261 bp genome. Expanded detail shows location and nucleotide bases of the *pac1* and *pac2* sequences. (Data acquired from the published HSVI genome, accession #X14112).

B. The Capsid

The viral capsid is a proteinaceous icosahedral shell that provides protection for the genome. In herpesviruses, the capsid consists of 150 hexameric and 12 pentameric protein subunits called capsomers, and ranges in size from 100-110 nm diameter (Wildy and Watson 1962). Three categories of capsids have been identified and are designated as type A, B or C. Type A capsids have neither DNA nor an envelope, type B capsids contain DNA but are unenveloped, while type C capsids have both DNA and an envelope (Gibson and Roizman 1972; 1974). Some examples of these capsid types are shown in Figure 3. Through electron microscopy, capsid types have been subdivided into eight different forms that probably represent various stages of assembly (Roizman 1974). However, type A capsids are not intermediate forms of B or C capsids, but rather are

disfunctional capsids (Sherman and Bachenheimer 1988). Each of these capsid types varies in the composition of their viral protein (VP) subunits, their content, and their molecular weight (Homa and Brown 1997). While A capsids are entirely non-functional, B capsids are capable of producing mature virions (Perdue et al. 1976) and C capsids are most similar to mature virions (Homa and Brown 1997). The outer capsid shell capsomer hexons and pentons are composed of the major capsid protein, VP5. In HSV-1, these capsomers are linked together by a tripartite complex of two VP23 (UL18) and one VP19C (UL38) minor capsid proteins (Newcomb et al. 1993).

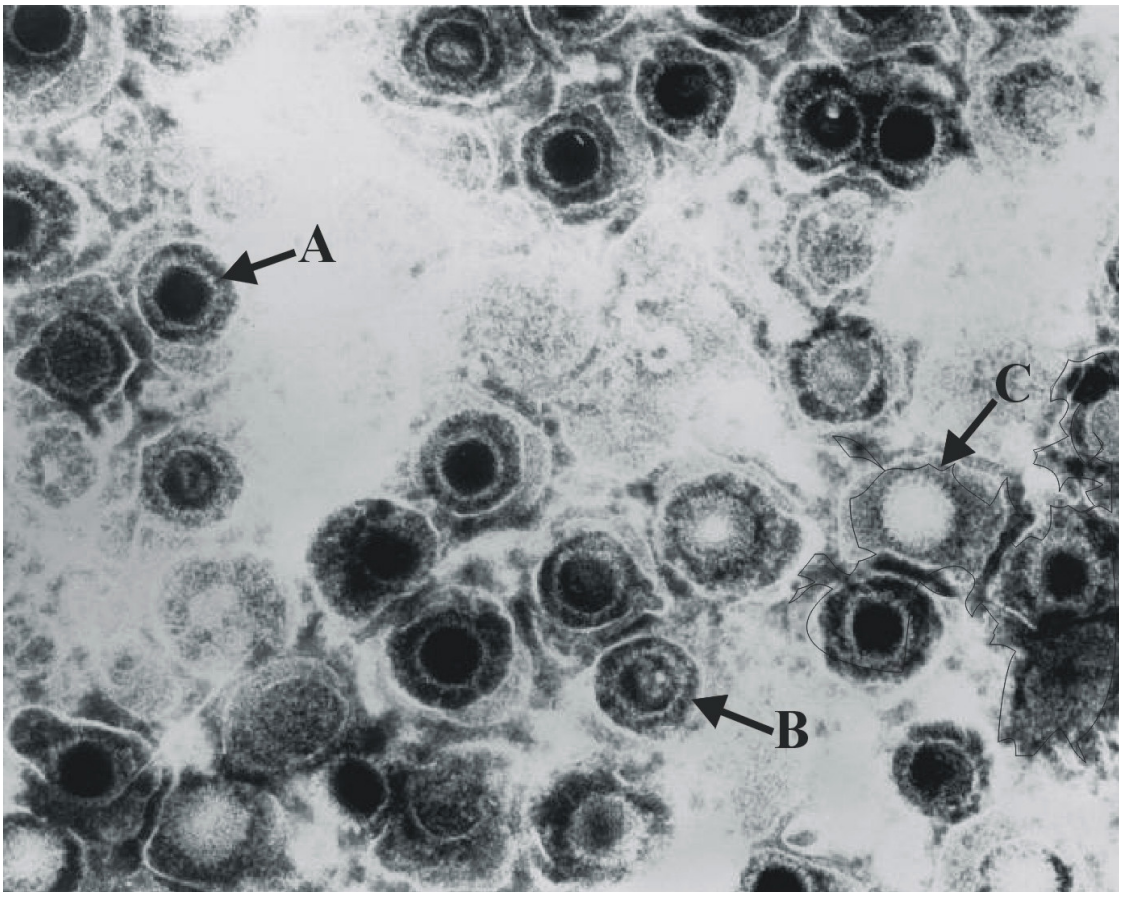


Figure 3. Transmission electron micrograph of herpes simplex virus. Some nucleocapsids are empty, as shown by penetration of electron-dense stain. Letters indicate capsid type (see text for detail). (This micrograph comes from the Centers for Disease Control and Prevention's Public Health Image Library (PHIL), identification number #281. Dr. Erskine Palmer, 1981).

Similar capsid types are present in BaCMV OCOM4-37 as well (Figure 4).

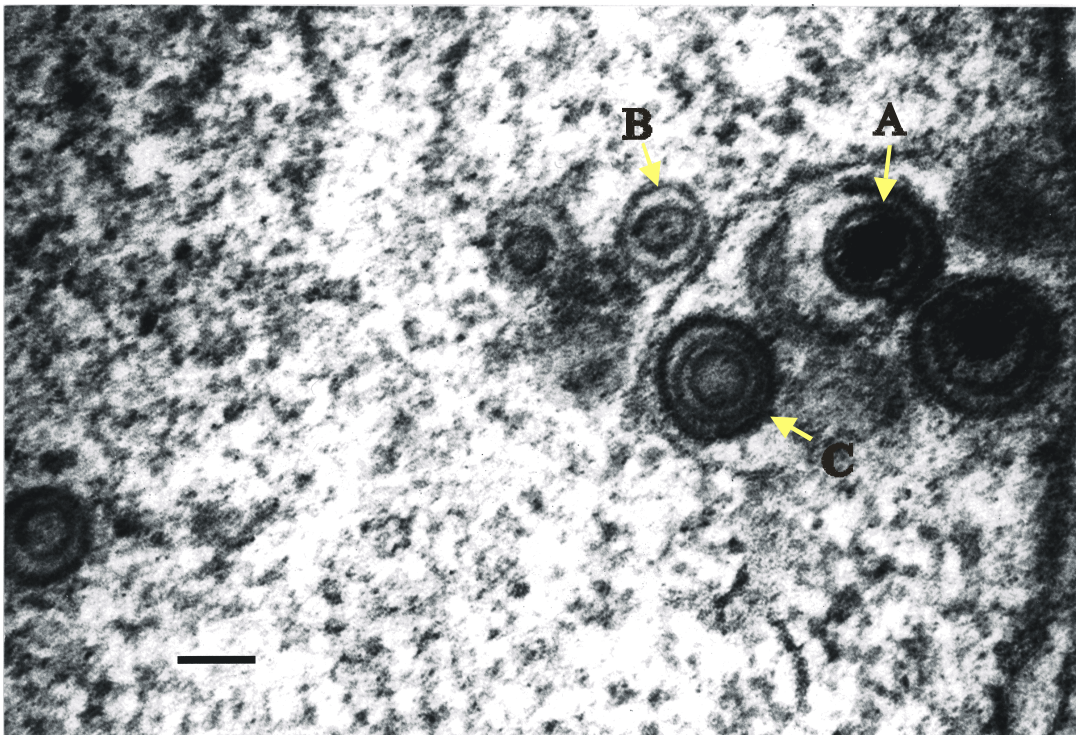


Figure 4. Transmission Electron Micrograph of BaCMV Strain OCOM4-37. Letters indicate capsid type. Bar at lower left represents 100 nm.

C. The Tegument

The tegument in herpesviruses is a collection of proteins that are located between the capsid and viral envelope. The herpesvirus tegument makes up about one-third of the volume of the virion (Desai et al. 2001) and is roughly analogous to the matrix of other viruses. However the herpesvirus tegument is more complex, structured, and contains a number of functional proteins. For some time, the tegument was considered to be an unstructured proteinaceous region. More recently cryoelectron microscopic analyses have shown that the innermost layer of the tegument (composed of the UL36 gene product in HCMV) appears to have icosahedral symmetry similar to that of the capsid (Zhou et al. 1999) suggesting that this protein physically interacts with the major capsid protein (McNabb and Courtney 1992). This is an example of a tegument protein that

links structural components of the capsid with viral glycoproteins to maintain the integrity of the virion. Another HCMV protein, UL37, has been shown to interact with UL36 and is proposed to be the second layer of the tegument (Klupp et al. 2002). Various other tegument proteins have been identified in herpesviruses to date and a number of these have been characterized (Mettenleiter 2002; Zhou et al. 1999). Some known functions of tegument proteins include unpackaging and transport of capsids to the nuclear pore (product of HCMV UL47 and HSV-1 ICP1-2), (Batterson et al. 1983; Bechtel and Shenk 2002; Knipe et al. 1981), suppression of host cell protein synthesis (product of HSV-1 UL41) (Everly et al. 2002), transactivation of other viral genes (product of HCMV UL32) (Liu and Stinski 1992), virion production genes (Meyer et al. 1997), egress from the nucleus (products of HSV-1 UL31 and UL34) (Fuchs et al. 2002; Reynolds et al. 2001), and immunomodulation genes (Child et al. 2004; Trgovcich et al. 2006). An example of immunomodulation is seen where HCMV pp71 (UL 82) causes disruption of MHCI cell surface expression to prevent antigen presentation (Trgovcich et al. 2006). In addition to this, pp71 plays a role as a transactivator of IE genes in much the same way as the HSV-1 ICP0 protein (Preston and Nicholl, 2005). Tegument protein homologues of HCMV conserved in most herpesviruses are UL31, UL34, UL36, and UL37. Table 1 shows functions of some HSV-1 tegument proteins (for HCMV tegument proteins see Table 10).

D. The Envelope

There has been much speculation regarding acquisition of the herpesvirus envelope since they undergo envelopment twice. Therefore, the origin of the herpesvirus envelope has been proposed by two different mechanisms. One model (referred to as the re-envelopment pathway), begins with assembled nucleocapsids that bud through the

inner nuclear membrane (Vlazny et al. 1982). At this point, the envelope acquired from the inner membrane fuses with the outer membrane to release the nucleocapsids into the cytoplasm (Siminoff and Menefee 1966). The nucleocapsids are then finally enveloped

HSV-1 gene	Protein name	Function	Reference
UL7		Interacts with adenine nucleotide translocator 2; nonessential *in vitro*	(Tanaka et al. 2008)
UL11		Virion egress and secondary envelopment	(Baines and Roizman 1992; Loomis et al. 2001)
UL14		Interacts with UL11; associated with nuclear targeting of capsids	(Yamauchi et al. 2008)
UL16		Interacts with UL11	(Yeh et al. 2008)
UL36	VP1/2	Release of DNA from capsids during entry	(Desai et al. 2008)
UL37		Interacts with UL36; required for virus maturation	(Desai et al. 2001)
UL41	vhs	Inhibits host cell protein synthesis	(Smiley 2004)
UL46	VP11/12	Modulates transactivating tegument protein VP16	(Liu et al. 2005)
UL47	VP13/14	Modulates transactivating tegument protein VP16; RNA-binding protein	(Donnelly et al. 2007)
UL48	VP16	Transactivates immediate early genes	(Ellison et al. 2005)
UL51		Unknown; nonessential *in vitro*	(Daikoku et al. 1998)

Table 1. Herpes Simplex Virus 1 Tegument Proteins and Their Functions

at the trans-Golgi face.

Another model, the luminal pathway, proposes that virions acquire a double envelope as they pass through the inner and outer nuclear membranes. They then travel from the inner nuclear space directly to the Golgi apparatus or within the lumen of the endoplasmic reticulum (ER) (Enquist et al. 1998). In either case, viral glycoproteins undergo modification within the Golgi apparatus and mature virions are released from the cell via typical secretory pathways. Based on recent studies using immunogold electron microscopy to determine the distribution of gD targeted to the ER, the re-envelopment

theory seems the most likely model (Skepper et al. 2001). Table 2 describes some CMV glycoproteins.

HCMV Gene	Protein name	Function	References
UL4		Contains signal peptide; nonessential in cell culture	(Chen and Stinski, 2002)
UL16		Inhibits NK cell cytotoxicity; involved in immune regulation	(Vales-Gomez and Reyburn 2006; Wilkinson et al. 2008)
UL18	MHC homologue	Inhibits NK cell cytotoxicity; involved in immune regulation	(Prod'homme et al. 2007; Vales-Gomez et al. 2005)
UL55	Glycoprotein B (gB)	Involved in cell entry; mediates cell-to-cell spread	(Jarvis et al. 2004)
UL75	Glycoprotein H (gH)	Involved in cell entry; mediates cell-to-cell spread	(Jarvis et al. 2004; McWatters et al. 2002)
UL110	Glycoprotein M (gM)	Integral membrane protein that complexes with gN; nonessential in cell culture	(Mach et al. 2000; Mach et al. 2005)
UL115	Glycoprotein L (gL)	Contains a signal peptide and is involved in cell entry; mediates cell-to-cell spread	(Kaye et al. 1992)
US10		Delays trafficking of MHC-I; involved in immune regulation	(Furman et al. 2002)
US11		Causes selective degradation of MHC-I; involved in immune regulation	(Oresic and Tortorella, 2008)

Table 2. Cytomegalovirus Envelope Glycoproteins

III. LIFE CYCLE

The HCMV replication cycle is similar to that of most herpesviruses. The most basic virus replication cycle consists of attachment, penetration, biosynthesis, assembly, and egress. Herpesviruses have large genomes compared to other viruses, and as might be expected from this, their life cycles are more complex.

A. Attachment

1. Viral Envelope Glycoproteins

Because viruses are non-motile, the attachment process is random and depends on Brownian movement of the virus particles for contact with a host cell. In addition, attachment is energy independent and may involve several steps involving multiple viral and cellular proteins. Once attachment occurs, penetration quickly ensues as these are closely-integrated processes.

An important consideration when investigating viral attachment mechanisms is host cell polarity. *In vivo*, epithelial cells possess both an apical (or outer) surface and a basal laminal (or interior) surface, and are therefore described as polarized. Each of these surfaces has different cell-surface receptors that may enhance or inhibit viral attachment depending on whether or not they can bind a particular viral glycoprotein. Given this fact, an important consideration with *in vitro* experimental conditions is the use of non-polarized mammalian cell lines whose receptors may be randomly distributed on the cell surface. Therefore, experimental models may not precisely reflect viral attachment processes *in vivo*, but can certainly provide valuable information regarding the types of cellular receptors and viral proteins involved.

Experiments with HSV-1 and nonpolarized cells have shown that glycoprotein C (gC) and, to a lesser degree, glycoprotein B (gB) are involved in the initial stages of attachment to cell-surface heparan sulfate (Shieh et al. 1992; WuDunn and Spear 1989). In a subsequent step, glycoprotein D (gD) interacts with a cellular coreceptor (described below), and finally fusion occurs between the viral envelope and the cellular membrane. In addition to the role gB and gD play in fusion with the host cell membrane, they have

also been linked to HSV-1 tropism for the central nervous system (Izumi and Stevens 1990).

Likewise, attachment of HCMV requires initial binding of heparan sulfate proteoglycans (HSPG) on the host-cell surface to glycoprotein gC-II and/or gB to enhance stable attachment of the virion to cell-surface receptors (Compton et al. 1993). While gC-II has been identified as the major HCMV heparan-binding protein (Kari and Gehrz 1992), gB-mediated binding has been demonstrated in HSPG-deficient CHO (Chinese hamster ovary) and fibroblast cells where gB binding was diminished, but not entirely eliminated (Boyle and Compton 1998; Norkin, 1995). The fact that some gB binding occurred in these experiments seems to indicate that there must be alternative receptors for gB in addition to heparan sulfate. Indeed evidence exists (in HSV) for a specific gB entry receptor (Bender et al. 2005). Similarly in HCMV, the gB protein interacts and binds with cellular annexin II which could explain why gB also binds to cells lacking HSPG (Pietropaolo and Compton 1997). These data indicate that there may be additional receptors, and other glycoproteins (like gH) that play a role in the attachment/penetration process. What is known, however, is that gB is essential for viral entry and cell-to-cell spread based on its function as a ligand and fusion protein (Navarro et al. 1993; Singh and Compton 2000) and that it is one of the most conserved proteins in all herpesviruses. This, of course, would indicate how important gB is to the entire herpesvirus family.

Structurally, HCMV glycoprotein B has a surface exposed (SU) component represented by a 116 kDa subunit and a transmembrane region (TM) domain consisting of a 55 kDa subunit. Four possible hydrophobic (transmembrane) regions have been

predicted based on the amino acid (aa) sequence of the open reading frame (ORF). One is in the 116 kDa subunit and the remaining is in the 55 kDa subunit. The first of these domains, located between aa 1-24 at the N-terminus, acts as a cleavable signal sequence (Britt and Mach 1996). The second hydrophobic domain is between aa 714-747. gB deletion mutants for this domain are not secreted and are incompletely processed. This indicates some function for this region in gB secretion, folding and/or processing (Reschke et al. 1995). The most hydrophobic region is found between aa 751-771 and serves as a membrane anchoring sequence. The function of a fourth hydrophobic region between aa 784-792 remains undefined. There is also a series of 12 cysteine residues that are conserved in the gB sequence of CMVs (Chou, 1992; Rapp et al. 1992; Schleiss 1994; Spaete et al. 1988), and a heparan-binding site exists within the C-terminal fragment of a soluble form of gB (aa 492-692) (Carlson et al. 1997).

There are three antibody-binding sites on HCMV gB (Meyer et al. 1992), and two of these induce virus-neutralizing antibodies (Kniess et al. 1991; Wagner et al. 1992). One antigenic domain (AD-1) is located toward the C-terminal end between aa 560-640 and is immunodominant to the other domains (Wagner et al. 1992). AD-2 spans a region at the N-terminus between aa 50-86 and contains two separate binding sites: one at aa 50-54 (site I) and the other at aa 67-86 (site II) (Kniess et al. 1991). Site II of AD-2 is the only other region known to induce virus-neutralizing antibodies (Meyer et al. 1990). The third antigenic domain, AD-3, is also at the C-terminus and spans aa 798-805 (Silvestri et al. 1991). These three domains are shown in Figure 5. Interestingly, via aa 1-750, gB has been shown to directly induce the interferon pathway through interaction with an

unknown cell surface receptor (Boehme et al. 2004). This indicates a role for gB in establishing the antiviral state within fibroblasts and endothelial cells.

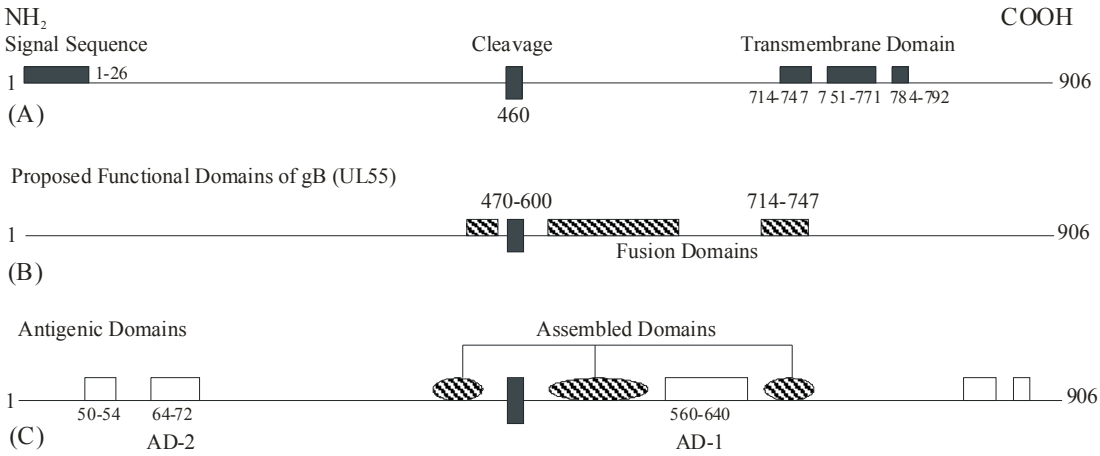


Figure 5. Linear Representation of HCMV AD169 Strain gB (UL55) Gene. (A) Shows the signal sequence (aa 1-26), the cleavage site (aa 460) and the transmembrane domain made up of three hydrophobic regions. Cleavage of the 906 aa precursor at aa 460 creates the surface (SU) and transmembrane (TM) sections. (B) Proposed fusion domains of gB based on homology with glycoproteins from other viruses that have fuseogenic domains. (C) Antibody-binding sites are represented as rectangles in the AD-1 and AD-2 sections. Ellipses represent conformation-dependent antibody binding sites. (Adapted from Britt and Mach, 1996).

2. Host Cell Receptor

The host cell receptor is a cell-surface molecule that undoubtedly performs some cellular function, but which also interacts with the virus to permit attachment. These receptors are sometimes referred to as viral receptors. Attachment to the host cell receptor is critical to initiate infection, so it is common for multiple host cell surface proteins to serve as receptors for a given viral species. There are three distinct families of cellular molecules that act as viral receptors for HSV-1 attachment. One, a member of the tumor necrosis factor (TNF) family designated herpesvirus entry protein A (HveA) is found primarily in lymphoid tissue and plays a limited role as a receptor (Warner et al. 1998). Another group of receptors are members of the immunoglobulin family and act as adhesion molecules in a variety of tissues including skin, brain, and spinal ganglia

(Cocchi et al. 1998). These molecules were formerly designated HveB and HveC but have been renamed nectins (from the Latin, to connect) based on their cellular function (Takahashi et al. 1999). 3-O-sulfated heparan sulfate (which HSV-1 binds to) is the sole representative of a third viral receptor family (Liu et al. 2002; Shukla et al. 1999). These three cell surface receptors are referred to as co-receptors because binding of two or more receptors enhances viral entry. In HCMV, a cell-surface HSPG has been identified as the receptor. This has been demonstrated in HPSG-deficient human fibroblasts and CHO cells as the virus does not infect these cells (Compton et al. 1993; Neyts et al. 1992). Additional evidence for HSPG as a receptor is that sulfated polysaccharides compete with HCMV in vitro to inhibit infection of fibroblast cells (Neyts et al. 1992).

B. Virus Entry

Attachment and penetration are closely linked events that involve a number of viral glycoproteins. In α-herpesviruses, glycoproteins gB and gD, and the gH-gL heterodimer are required for penetration (Spear et al. 2000). In CMVs, which are β-herpesviruses, only gB and the gH-gL complex are required for viral entry (Boyle and Compton, 1998). For the α-herpesviruses, gD has been shown to bind either the herpesvirus entry mediator (HVEM) receptor (Montgomery et al. 1996) or HveC (i.e. nectin-1) (Campadelli-Fiume et al. 2000), and after this primary interaction gB and gH-gL generate fusion with the host cell membrane. As previously discussed, heparan sulfate is usually bound by gC prior to stable attachment of gD. While heparan sulfate receptors are not necessarily required for fusion to occur, HSV-1 entry diminishes in their absence (Pertel et al. 2001).

1. Virus Uncoating and Transport

After fusion, the herpesvirus capsid is released into the cytoplasm and must be transported to nuclear pores where uncoating takes place. Cellular transport mechanisms are used to move herpesvirus components to various compartments within the cell and this requires nuclear localization signals (NLS). The NLS represents an aa motif that can be recognized by cellular binding proteins. Capsid transport occurs via microtubular networks existing within the cell. This has been demonstrated by electron microscopy which shows capsids bound to microtubules in neurons (Penfold et al. 1994), and also through antibody labeling which revealed attachment of dynein, the microtubule-dependent motor, to capsids (Sodeik et al. 1997).

Some tegument proteins such as virion host shutoff protein (vhs) and US11 are released into the cytoplasm after fusion of the viral envelope and a host cell membrane. Other tegument proteins either remain associated with the capsid or are transported independently to the nucleus (O'Hare 1993; Yang et al. 2002). An example of a CMV tegument protein that is immediately transported to the nucleus is phosphoprotein 65 (pp65), which is encoded by the UL83 gene. Not surprisingly, the primary structure of pp65 has a bipartite NLS as shown through experiments using deletion mutants (Schmolke et al. 1995a). The pp65 protein may play a role in transcription activation since it has been shown to compromise the activity of at least one viral-associated protein kinase (Schmolke et al. 1995b). As microtubule transport of the capsid ensues, it is believed that various tegument proteins dissociate and are carried to areas of the cell in which they function.

Once the capsid reaches the nucleus, it binds to the nuclear pore complex (NPC) which is a protein-lined aqueous channel that exists in the double-membraned nuclear envelope (Whittaker and Helenius 1998). Immediately after binding the NPC, the viral genome is ejected into the nucleus leaving the empty capsid behind (Ojala et al. 2000; Sodeik et al. 1997). It has also been suggested that the VP1-3 tegument protein may play a role in nuclear localization and/or DNA release at the NPC. This is because the protein sequence of HSV-1 VP1-3 contains four, weak bipartite NLS and a number of arginine-rich regions. By comparison, the human immunodeficiency virus type I genome contains regions that are high in arginine which are also involved in nuclear targeting (Truant and Cullen 1999). The VP1-3 sequence is also believed to function in DNA extrusion since VP1-3 mutant HSV-1 strains are unable to release DNA at the NPC at nonpermissive temperatures (Ojala et al. 2000). In addition to this, the VP1-3 protein is tightly bound to the capsid (Spear and Roizman 1972), and since the capsid penton is believed to be the site of DNA release (Newcomb and Brown 1994), interaction between the capsid and VP1-3 to facilitate this process is plausible. Once viral DNA is released into the nucleoplasm, immediate-early (IE) gene transcription begins.

C. Biosynthesis

1. Genome Transcription

The biosynthetic strategy used by herpesviruses involves sequential events in which different transcription units are transcribed at different times. Because these events occur as a cascade, the transcribed genes are categorized as immediate early (IE), early (E), or late genes (L). L genes are further subdivided into the leaky late (γ_1) and true late (γ_2) categories. This cascading regulatory system allows for efficient temporal

control of enzyme synthesis, DNA synthesis, and structural (late) protein production. These transcriptional events occur as a defined sequence designed to inhibit cellular gene expression in favor of virus replication. One critical step involves the destabilization of host cell mRNA by the vhs protein which enters the cell as a component of the infectious virion (Everly et al. 2002). Following this, transcription of IE genes begins and the IE proteins activate the E genes. Transcription of E genes is not dependent on DNA synthesis, whereas L gene transcription is (Conley et al. 1981). But once the E transcripts are produced, they promote viral DNA replication. Finally, DNA synthesis stimulates transcription of the early-late (γ_1) or true-late (γ_2) genes which code for the structural proteins of the virus.

For each class of genes, host cell RNA polymerase II is used for transcription (Costanzo et al. 1977), and transcription is initiated at promoter regions in the 5' non-coding region of the gene. However, the arrangement of promoter and regulatory sequences for each gene class is somewhat different. For all classes, the TATA box, to which transcription factors bind to facilitate binding of RNA polymerase, is located 20-35 bp upstream from the initiation codon. For most classes, binding sites for one or more cellular transcription factors are also found upstream of the TATA box. An exception to this is seen in the γ_2 gene subclass which lacks a transcription factor binding site, but has an initiator and activator element downstream of the TATA box (Guzowski and Wagner 1993).

In HSV-1, the IE genes code for proteins designated as ICP0, ICP4, ICP22, ICP27, ICP47 (infected cell polypeptides), and $U_s1.5$. With the exception of ICP47, these are regulatory proteins that are involved in regulation of viral gene transcription.

The functions of ICP47 appear to be retention of MHCI class proteins in the ER (Hagmann et al. 1995; York et al. 1994) and inhibition of antigen presentation to $CD8^+$ T cells (Fruh et al. 1995; Hill et al. 1995). These activities contribute to evasion of the host's immune system (see Latency and Persistence section). Nevertheless, synthesis of all these IE proteins requires the transcriptional activating protein VP-16 (UL48) (Triezenberg et al. 1988) which is brought into the host cell as part of the tegument. To function as a transcriptional activator, VP-16 must somehow enter the nucleus where transcription occurs. There is evidence that translocation of VP-16 into the nucleus is promoted through some interaction with cellular protein HCF (host cell factor) within the cytoplasm (LaBoissiere and O'Hare 2000). Once in the nucleus, VP-16 binds to host cell protein octomer 1 (Oct-1) to enhance activity of cellular transcriptional proteins already bound to viral DNA (Stern and Herr 1991). Subsequently, ICP4 becomes the major transactivating protein for the E and L genes (Carrozza and DeLuca, 1996; DeLuca et al. 1985). ICP4 is also able to repress its own transcription (DeLuca and Schaffer 1985; O'Hare and Hayward 1985). It has been demonstrated that IPC4 activates transcription by interacting with transcription factor IID (TFIID) which requires interaction with TBP-associated factor (TAF250) mediated by the C-terminal region of ICP4 (Carrozza and DeLuca, 1996). ICP4 has a number of functional domains including a DNA-binding region, a nuclear localization region, and two transactivating regions (DeLuca and Schaffer 1988; Shepard et al. 1989). ICP27, contains an RNA-binding sequence, a leucine-rich nuclear export signal, and an export control sequence, and so is able to shuttle between the cytoplasm and nucleus of the cell (Soliman and Silverstein 2000). In its role as a shuttle protein, ICP27 is able to facilitate the export of viral RNA transcripts.

ICP27 binds RNAs via an arginine-glycine region referred to as an RGG box (Sandri-Goldin 1998). These transcripts must contain the same 5' cap and 3' poly A tail that cellular transcripts possess in order to allow export from the nucleus. ICP27 also plays a role in inhibiting host cell protein synthesis by redistributing splicing factors that modify cellular RNA transcripts (Phelan et al. 1993; Sandri-Goldin et al. 1995). In HCMV, UL69 is the homologue of ICP27, but it cannot substitute for ICP27 in HSV-1-infected cells due to some functional differences (Winkler et al. 1994; Winkler and Stamminger 1996).

Viral mRNA is translated by cellular systems under the influence of specific regulatory signals. Since the IE transcripts initiate both E and L gene transcription, control mechanisms exist to ensure temporal regulation of expression. One means of doing this is to shutdown IE and E gene expression at appropriate times. There are two processes currently known to promote this. First, ICP4 can bind to repressor elements in its own promoter to repress its own transcription (Faber and Wilcox 1986; Kristie and Roizman 1984). Since ICP4 is essential for transcription of IE and E genes, this would decrease expression of these genes. Second, the major DNA-binding protein ICP8, which is involved in replication, (Bayliss et al. 1975; Chen and Knipe 1996; Gao and Knipe 1991) represses L gene expression from the parental viral genome (Chen and Knipe 1996; Godowski and Knipe 1983; 1985; 1986) but enhances L gene expression from progeny genomes (Gao and Knipe 1991).

2. DNA Replication

When HSV DNA is released from capsids into the nucleus, it quickly changes from linear to a circular form (Uprichard and Knipe 1996). Actual replication occurs in unique

structures described as replication compartments (Quinlan et al. 1984). These compartments form near nuclear domain 10 (ND10) sites, which are cellular structures with which viral DNA and viral replication proteins associate (Ishov and Maul 1996; Quinlan et al. 1984). To begin DNA synthesis, the HSV-1 origin binding protein (OBP) may bind to one of the three origins of replication (ori) in the genome. There is one site (oriL) located between the viral replication proteins ICP8 (UL29) and DNA polymerase (UL30) genes (Weller et al. 1985), and two copies of the oriS site in the repeat sequences flanking the US region (Frenkel et al. 1975). Only one of the three sites is required to begin viral DNA replication (Igarashi et al. 1993; Polvino-Bodnar et al. 1987). In HCMV, these lytic-phase DNA replicator (oriLyt) sequences are found near the middle of a structurally-complex area of the UL region (Anders et al. 1992; Masse et al. 1992). Initially, replication begins in both directions on the circular DNA molecule resulting in bidirectional replication. Shortly thereafter, replication probably continues by rolling-circle replication which produces head-to-tail concatemeric DNA (Jacob et al. 1979). Genetic and biochemical techniques have identified seven viral proteins (UL5, UL8, UL9, UL29, UL30, UL42, and UL52: see Table 3) involved in HSV-1 DNA replication (Challberg, 1986). In addition to these virally-encoded proteins, DNA synthesis also depends on cellular factors such as topoisomerase and DNA ligase.

a) Replication Proteins

(1) Origin Binding Protein

The OBP encoded by UL9 consists of 851 aa and is involved in theta replication of DNA (i.e. circular DNA molecules). This protein has two functional domains, ATP-binding and DNA helicase motifs, required for DNA replication. To initiate replication, OBP

Gene location	Protein name	Function	Reference
UL5	Helicase-primase component	Helicase, primase, and ATPase activity	(Dracheva et al. 1995)
UL8	Helicase-primase component	Enhances nuclear localization	(Barnard et al. 1997)
UL9	Origin –binding protein	Unwinds double-stranded DNA	(Fierer and Challberg, 1992)
UL29	Single-stranded binding protein	Allows origin unwinding and progression of the replication fork	(Boehmer and Lehman, 1993)
UL30	DNA polymerase holoenzyme	Acts as the catalytic subunit	(Purifoy et al. 1977)
UL42	DNA polymerase holoenzyme	Acts as a processivity factor	(Gottlieb et al. 1990)
UL52	Helicase-primase component	Helicase, primase, and ATPase activity	(Klinedinst and Challberg 1994)

Table 3. Viral Proteins Involved in HSV-1 DNA Replication

binds to an origin of replication as a dimer at the conserved sequence 5'–CGTTCGCACTT-3' (Elias and Lehman 1988; Koff and Tegtmeyer 1988). Once the OBP unwinds the double-stranded DNA, it recruits single-stranded DNA binding protein (ICP8), and together these two proteins recruit the remaining five viral proteins to complete the replication complex.

(2) Single-Stranded DNA Binding Protein

The single-stranded DNA binding protein (SSB) is encoded by UL29 and produces a protein of 1196 aa. It was originally called the major-binding protein (Bayliss et al. 1975). As its name implies, ICP8 binds more efficiently to ssDNA than to dsDNA (Lee and Knipe 1985; Ruyechan 1983). This protein is essential for viral DNA synthesis and optimal expression of late genes. While ICP8 is able to destabilize the DNA helix to allow origin unwinding and replication fork progression (Boehmer and Lehman, 1993), it has also been shown to stimulate both the rate and degree of helicase activity for UL9 (Boehmer et al. 1993). Additionally, this protein has been shown to promote renaturation

of complementary single-stranded DNA which could account for the frequent recombination events that occur in HSV-1-infected cells (Dutch and Lehman, 1993). Since ICP8 is large and interacts with many other replication proteins, it may function as scaffolding protein to assemble the HSV DNA replication complex (de Bruyn Kops and Knipe, 1988). ICP8 also plays a role in recruitment of replication proteins to small prereplication sites before viral DNA replicates (Bush et al. 1991; Liptak et al. 1996). Once DNA replication starts, ICP8 is required for formation of large, globular viral replication compartments in infected cell nuclei (de Bruyn Kops and Knipe, 1988). Evidence of these functions is demonstrated by immunocytochemical and biochemical analysis of ICP8 which shows binding to other HSV proteins or DNA causes conformational and localization-associated antigenic changes to occur in the ICP8 protein (Uprichard and Knipe 2003).

(3) DNA Helicase-Primase Complex

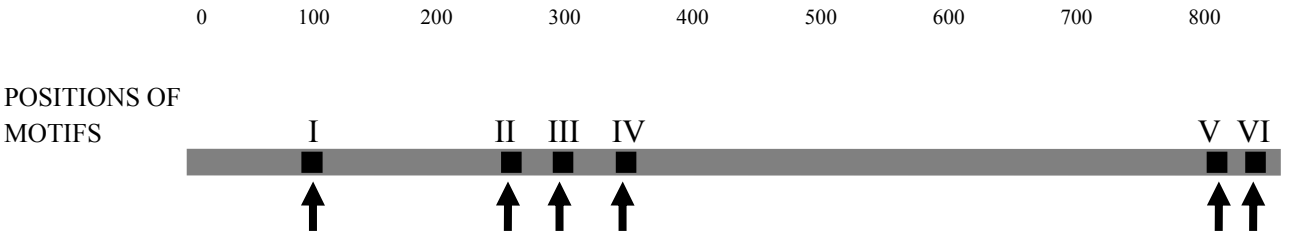
The DNA Helicase-Primase complex is a heterodimer consisting of the UL5, UL8, and UL52 proteins (Crute et al. 1988). UL5 contains 882 aa and two conserved sequences which are essential for DNA replication: the ATP-binding and helicase motifs (Zhu and Weller 1992). There are six conserved sequence motifs in UL5 that are found in all members of the DNA and RNA helicase superfamily (Figure 6).

The product of UL52 contains 1058 aa, has a divalent metal-binding motif common to other DNA primases, and is required for cleavage and packaging of viral DNA (Borst et al. 2008). Like UL5, it is also required for DNA replication. Together, UL5 and UL52 form what is referred to as the core enzyme sub-complex which conducts helicase, primase, and ATPase activities. The last protein of the helicase-primase complex (UL8)

is 750 amino acid residues and functions to enhance nuclear localization and also stimulates core enzymatic activities.

(4) DNA Polymerase Holoenzyme

The fact that HSV-1 encodes its own DNA polymerase was not definitively established until genetic experiments mapped the gene to the HSV-1 genome in the late 1970s (Purifoy et al. 1977). Further support for this finding was provided later when the sequence of the HSV UL30 protein was compared to that of mammalian DNA polymerase α and found to have several shared regions of sequence similarity (Gibbs et al. 1988). The HSV DNA polymerase holoenzyme is a heterodimer composed of a 1235 aa catalytic subunit, UL30, and a 488 aa accessory factor, UL42 (Purifoy and Powell 1976). The interaction between the two is required for viral DNA replication. The UL30 subunit (Pol) has both polymerase and intrinsic 3'-5' exonuclease activity that allows proofreading as the enzyme replicates DNA. The UL42 subunit is a double-stranded DNA binding protein that acts to increase processivity of UL30 (Gallo et al. 1989; Gottlieb et al. 1990). Because HSV-1 DNA polymerase has broader substrate specificity than cellular DNA polymerase, it has been a good target for drug therapy. Another recently-investigated target for drug therapy involves inhibition of the interaction between UL30 and UL42 to prevent DNA replication (Pilger et al. 2004). The HCMV homologues of HSV-1 DNA polymerase are the UL54 polymerase (Heilbronn et al. 1987) and the UL44 accessory protein (Ertl and Powell 1992). The DNA polymerase is the most conserved gene among the herpesviruses.



	MOTIF SEQUENCE	I		II		III		IV		V		VI	
UL5	94	ITGNAGSGKSTCVQ	136	VIVIDEAGLLG	26	LVCVGSPTQTAS	44	NNKRCVEHE	442	AMTITR SQGL	SLDKVAICF	8	SAYVAMSRT
uvrD	26	VLAGAGSGKTRVLV	174	NILVDEFQNTN	16	VMIVGDDDQSIY	26	QNYRSTSNI	267	LMTLHS AKGL	EFPQVFIVG	23	LAYVGVYRA
rep	19	VLAGAGSGKTRVIT	175	YLLVDEYQDTN	16	FTVVGDDDQSIY	26	QNYRSSGRI	271	LMTLHA SKGL	EFPYVYMVG	22	LAYVGITRA
recB	20	IEASAGTGKTFTIA	345	VAMIDEFQDTD	18	LLLIGDPKQAIY	24	TNWRSAPGM	286	IVTIHK SKGL	EYPLVWLPF	44	LLYVALTRS
recD	164	ISGGPGTGKTTTVA	82	VLVVDEASMID	16	VIFLGDRDQLAS	24	QLSRLTGTH	198	AMTVHK SOGS	EFDHAALIT	11	LVYTAVTRA
PIF	255	YTGSAGTGKSILLR	46	ALVVDEISMLD	25	LIFCGDFFQLPP	29	KVFRQRGDV	219	MQTIHQNSAGKRRLPLVRFKA		33	QAYVALSRA

Figure 6. Conservation of the UL5 Protein. The UL5 protein is represented by a shadowed box, and the six conserved motifs are represented by dark bars. The amino acid sequences of the six conserved motifs from six proteins (UL5, Rep, Uvr, RecB, RecD, and PIF) are shown. UvrD, rep, recB, and recD are *Escherichia coli* proteins and PIF is yeast helicase. Shaded aa are identical. The numbers between the aa sequences refer to the number of the residues separating each motif (Zhu and Weller 1992).

(5) Other Viral Replication/Modification Proteins

In addition to viral proteins essential for replication *in vitro*, a number of other proteins are non-essential in cell-culture, but probably are essential in non-dividing cells like neurons. This may be because several of these proteins are important in metabolism of nucleotides, which are less abundant in resting cells, so these viruses have adapted to produce their own enzymes to enhance viral DNA replication.

One of these enzymes, thymidine kinase (TK) encoded by the UL23 gene, is involved in phosphorylation of pyrimidine and purine nucleosides. Like HSV-1 DNA polymerase, TK has a broad substrate specificity so serves as a target for antiviral drugs such as the nucleoside analog acyclovir (Fyfe et al. 1978). TK recognizes acyclovir as a nucleoside and adds phosphate groups to the molecule. This molecule then is incorporated into the growing DNA chain and leads to chain termination since it lacks an attachment site for the next nucleotide. HSV TK is a homodimeric complex of the UL23 gene product that most likely produces nucleoside triphosphate precursors for DNA synthesis and is reportedly necessary for HSV reactivation from latency (Chen et al. 2004). Additionally, efficacy of the antiviral drug acyclovir requires phosphorylation by TK, and drug-resistance may develop due to mutations in the viral TK gene (Coen, 1994). CMVs do not produce TK but instead, another protein kinase (the UL97 gene product) serves the same function and acts to phosphorylate acyclovir (Talarico et al. 1999).

Two other proteins encoded by herpesviruses are the DNA repair enzyme uracil-DNA glycosylase (UNG/UL2) (Caradonna et al. 1987) and deoxyuridine triphosphatase (dUTPase/UL50) (Preston and Fisher, 1984). UNG is an enzyme that acts to remove

uracil from DNA when it is accidentally integrated during replication or when it occurs after cytosine deamination (Lindahl and Nyberg 1974) or incorporation of deoxyuridine monophosphate (dUMP) into DNA (Caradonna et al. 1987). Like many other HSV replication or modification proteins, UNG is required for DNA synthesis in non-dividing cells (Chen et al. 2002). dUTPase proteins are found in all eukaryotes, bacteria, and a number of DNA viruses, and helps maintain low levels of uracil in the cell to decrease its misincorporation into DNA. Phylogenetic analysis of herpesvirus mammalian dUTPase genes indicates that the viral gene probably originated through capture of a host gene. This produced a herpesvirus dUTPase which has a larger size than most dUTPases, but stills functions as a monomer (McGeehan et al. 2001).

A third enzyme, ribonucleotide reductase (RR, UL39, ICP6), reduces the ribose in ribonucleotides to deoxyribose (Bacchetti et al. 1986; Dutia 1983). It is a complex of UL39 [the small subunit (RR2) having peak expression 6 to 8 h post-infection (PI)], and UL40 [the large subunit (RR1) having onset of synthesis at 2 h PI] (Frame et al. 1985; Honess and Roizman 1974). Studies have shown that the large subunit is required for viral replication in non-dividing cells *in vitro* in both HSV-1 (Goldstein and Weller 1988) and HSV-2 (Smith et al. 1992). In addition to its function as a ribonucleotide reductase, the large (RR1) subunit demonstrates intrinsic serine-threonine kinase activity (Cooper et al. 1995) which has been linked to IE gene expression and viral growth, at least in HSV-2 (Smith et al. 1998). HCMV UL45 is a homologue of HSV ribonucleotide reductase, but is dispensable for growth in human fibroblasts and endothelial cells (Hahn et al. 2002).

Finally, alkaline nuclease (UL12) is a phosphoprotein required for efficient virus reproduction. It functions as both an endonuclease and exonuclease to facilitate DNA

maturation and encapsidation. It also appears to interact with ICP8 collaboratively to affect strand exchange between linear double-stranded DNA and circular single-stranded DNA (Reuven et al. 2003, Thomas et al. 1992) This suggests that ICP8 and alkaline nuclease catalyze homologous recombination in a similar fashion as a recombinase (Bortner et al. 1993; Gourves et al. 2000).

D. Assembly and Packaging

Assembly of herpesviruses occurs in the host cell nucleus following DNA replication. Capsid proteins, however, are synthesized in the cytoplasm of the cell. This necessitates translocation of capsid proteins into the nucleus where they form mature capsids. Capsid assembly proteins (UL80 proteins) possess NLS sequences which facilitate translocation of the major capsid protein into the nucleus (Nguyen et al. 2008). Once capsid proteins are in the nucleus, they are capable of self assembly. Self-assembly has been demonstrated *in vitro* using purified viral protein subunits (Newcomb et al. 1999). Once capsids are assembled, concatemeric viral DNA is cleaved at precise distances from the *pac1* and *pac2* sites (Deiss et al. 1986) and concurrently packaged into capsids (Ladin et al. 1982). Within the direct repeat (DR1) is a recognition signal which allows capsid binding, and at a second location a DNA cleavage site ensures packaging of the entire genome (Varmuza and Smiley 1985). Also, eight viral proteins (HSV-1 UL6, UL15, UL25, UL28, UL32, UL33, UL36, and UL37) function in genome packaging, and each has a specific function. During encapsidation, twelve UL6 protein molecules are incorporated into the capsid by interaction with the UL26.5 capsid protein to make up the portal complex through which DNA passes (Newcomb et al. 2003). Two other genes, UL15 and UL28, code for a viral terminase, the enzyme that drives DNA

into the capsid (White et al. 2003). The UL25 protein is believed to be an anchoring protein that stabilizes capsids to prevent loss of DNA that has already been packaged (Newcomb et al. 2006; Ogasawara et al. 2001). Functions of the other proteins (UL32, UL33, UL36, and UL37) are less well understood.

E. Egress

After viral transcription, replication, capsid production, and DNA packaging are complete, viral nucleocapsids must leave the nucleus and undergo further processing before leaving the cell. Prior to egress, the tegument is added, envelopment occurs, viral glycoproteins are processed in the ER and Golgi, and then mature virions are transported to the plasma membrane. The first step in the envelopment process involves the UL31 and UL34 proteins, which are needed for acquisition of an envelope from the inner nuclear membrane (Reynolds et al. 2001). These two proteins form a complex that is involved in primary envelopment of capsids by budding through the inner membrane, probably by partially dismantling the nuclear lamina (Scott and O'Hare 2001). In HSV-1, an additional protein (encoded by the US3 gene) functions as a protein kinase and appears to act as a modulator of the UL34 gene product which targets it to the nuclear membrane (Purves et al. 1992). A second capsid protein, UL37, facilitates exit of capsids from the nucleus and may also play a role in re-envelopment at the Golgi (Desai et al. 2001). At the outer nuclear membrane, de-envelopment occurs and the capsid is released into the cytoplasm. This is where tegumentation is completed before the virion buds into trans-Golgi vesicles containing processed viral glycoproteins (Mettenleiter 2002). In HCMV, these vesicles contain viral gB, trans-Golgi network 46 protein (a Golgi marker), mannosidase II (a Golgi resident enzyme), and the Rib 3 secretory vacuole marker

(Homman-Loudiyi et al. 2003). The HCMV tegument protein, pp150, is also crucial for virion egress and may play a role in tegumentation and envelopment as well (AuCoin et al. 2006).

1. Glycoprotein Processing

Glycoprotein processing is another critical event that must occur prior to release of progeny virions. An example of such processing is illustrated by gB. As previously mentioned, gB is a highly-conserved glycosylated envelope protein common to all herpesviruses. It is the major envelope glycoprotein of the virion, is found in the membranes of infected host cells (Farrar and Oram, 1984), is recognized by neutralizing antibodies (Britt 1984; Nowak et al. 1984b; Rasmussen et al. 1985), and is highly immunogenic in humans (Britt, 1984). Functionally, gB promotes virion penetration into cells, cell-to-cell transmission, fusion of infected cells, and is the primary target for the cell-mediated immune response (Britt et al. 1990; Navarro et al. 1993; Utz et al. 1989).

The HCMV gB transmembrane protein is expressed early in the infectious cycle (Smuda et al. 1997) and consists of 906 aa encoded by the UL55 gene (Chee et al. 1990; Cranage et al. 1986). gB is translated as a 105 kDa precursor protein which undergoes processing in the ER and Golgi to yield a mature 150 kDa glycoprotein (Britt and Auger, 1986; Britt and Vugler, 1989). High mannose carbohydrates are added in the rough ER. As the precursor protein is transported through the Golgi apparatus, the carbohydrates are further processed, and the 150 kDa gB is then cleaved into 116 kDa and 55 kDa subunits. Cleavage occurs by a cellular furan protease which recognizes the consensus sequence RXK/RR (Britt and Vugler, 1989; Vey et al. 1995). Proteolytic cleavage of gB has been shown to occur in the passaged HCMV AD169 strain as well as the clinical HCMV

Towne strain (Spaete et al. 1988). Although gB cleavage is considered essential for the HCMV life cycle *in vivo*, cleavage is dispensable for growth in cell culture (Strive et al. 2002). Once cleavage has occurred, the gB subunits remain linked together by disulfide bonds.

In both the RER and Golgi apparatus, the gB molecule undergoes prolonged folding, extensive disulfide bond formation and rearrangements (Billstrom and Britt, 1995). Finally, additional trimming or modifications may occur in the rough ER or Golgi as well as addition of carbohydrates (Huber and Compton 1999). The cleaved gB molecules associate to form the gp55-116 complex on the viral envelope and surface of infected cells (Spaete et al. 1990; Spaete et al. 1988).

F. Latency and Persistence

For viral persistence or latency to develop, herpesviruses must somehow evade host immune system surveillance. Typically in persistent infections, small numbers of virus progeny remain in tissues from which they are not quickly cleared. This occurs in immunologically "privileged" tissues where low levels of virus are constantly produced and shed. In HCMV infections, virus can be shed in immunocompetent individuals for years. In latent infections, which are typical of herpesviruses, the virus lies dormant in a non-replicating state in certain tissues until reactivated.

CMVs may remain latent in peripheral blood lymphocytes, bone marrow pre-cursor cells, and monocytes-macrophages (Hahn et al. 1998; Kondo et al. 1994). In monocytes and macrophages, the mechanism for latency has been linked to suppression of the major IE promoter/regulatory (MIEP) region of the genome (Sinclair and Sissons 1996). In HSV-1 and HSV-2, latency is established in nerve cells and this confers the

added benefit of allowing them to be easily transported around the body via neuronal fibers. For HSV-1, the replication initiator protein, UL9, which is bound by neural F-box protein (NFB42) is believed to be responsible for neuronal latency (Eom et al. 2004). However, the primary site of infection and viral replication for HSV-1 and HSV-2 is mucosal or epidermal cells, and the early signs of infection are lesions which are full of virus. Once the initial symptoms subside, the virus remains latent in non-dividing neurons of sensory or autonomic ganglia. During this time, the viral genome remains intact and there is little or no production of viral proteins. In this state, immune detection of the infected cell is blocked by various viral evasion mechanisms. Reactiv-ation of herpesviruses usually occurs due to some type of stress imposed on the host such as ultraviolet light, illness, etc., and results in reentry of the virus into the lytic replication cycle. This pattern of latency-reactivation is recurrent in herpesvirus infections and is grouped into three operational events: establishment, maintenance, and reactivation. Establishment of HSV-1 latency begins with acute infection in mucosal epithelia followed by entry of the virus into a sensory neuron. Because HSV-1 infection often involves the oral, optic or nasal mucosa, the trigeminal ganglia (TG) are a primary site for latency (Baringer and Swoveland, 1973).

In the latent state, two observed conditions likely have an effect on viral gene expression: 1) the viral genome becomes a circular episome (Rock and Fraser 1985) and 2) the viral genome associates with cellular histones to develop a chromatin-like appearance (Deshmane and Fraser, 1989). Another likely mechanism involves regulation of IE promoters in sensory neurons. Evidence of such regulation is apparent where the ICP4 promoter is functional in Schwann cells but not in sensory neurons of the TG (Taus

and Mitchell 2001). This indicates that neuron-specific factors differentially regulate IE promoters. In addition to this, certain factors may prevent formation of the VP16 transcriptional activating complex in sensory neurons. One such factor could be a decrease in host cell factor (HCF) since it is required to induce IE promoter activity. Such a decrease in HCF may be attributed to the activities of cellular transcription factors that bind HCF (Lu and Misra 2000a; Lu and Misra 2000b). Sensory neurons are also known to produce transcription factor Brn-3.0, which can bind to non-coding regions of the HSV-1 genome, (Turner et al. 1997) and other proteins that may play a role in expressing IE genes (Hagmann et al. 1995).

As previously mentioned, there is minimal expression of viral mRNA during latency and during this time no structural proteins are produced. Instead, transcription of a special set of RNAs referred to as LAT (latency-associated transcripts) occurs. LAT are restricted primarily to the nucleus, and these RNAs map to the long terminal repeat region of the genome (Stevens et al. 1987). Unlike promoters for other HSV-1 genes, the LAT promoter is activated within sensory neurons. This promoter consists of a strong promoter region (LAP1) and a weaker promoter region (LAP-2) (Chen et al. 1995; Goins et al. 1994). There are numerous binding sites for cellular transcription factors within the LAT promoter including two cyclic AMP (cAMP) responsive element (CRE)-binding sites and three Sp1 sites (Figure 7). The CRE site nearest the TATA box has been shown to influence LAT expression in neurons and is also involved in reactivation from latency (Bloom et al. 1997), (Goins et al. 1994; Leib et al. 1991; Rader et al. 1993). The LAP-2 sequence exhibits long-term promoter activity and is required to prevent shutoff

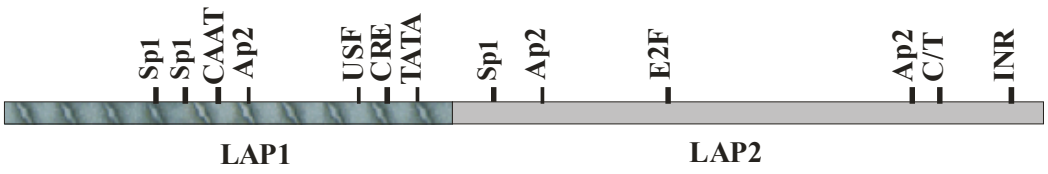


Figure 7. Representation of LAP1 and LAP2 Regulatory Regions and Associated Cis-acting Elements. LAP is the latency-active promoter, Sp1 and E2F are transcription factor consensus sequences, USF is an upstream transcription factor, CRE is a cyclic AMP response element, INR is an insulin-like growth factor receptor, CAAT and TATA are transcriptional factor binding sites. Adapted from Goins, et al. 1994.

of LAP-1 in latent infection (Lokensgard et al. 1997; Lokensgard et al. 1994). LAT itself does not encode for a protein, but is instead an RNA molecule that functions to regulate latency (Stevens et al. 1987). This has been demonstrated in transiently transfected cells where LAT is able to decrease the promoter-transactivating ability of ICP0 (Farrell et al. 1991). While most studies have shown that LAT plays an important role in establishment of latency, it does not appear to be essential (Sawtell and Thompson 1992; Thompson and Sawtell 1997). However, spontaneous reactivation of HSV-1 from latency is drastically impaired in LAT-negative mutants (Perng et al. 1999; Perng et al. 1994). Another reported function of LAT is inhibiting apoptosis in transiently transfected cells (Ahmed et al. 2002). Apoptosis (one type of programmed cell death) is observed in virus-infected cell culture (Teodoro and Branton 1997) and presumably helps prevent viral spread *in vivo*. While LAT does not appear to be necessary for reactivation, virally-encoded TK does seem to be (Coen et al. 1989; Watkins et al. 1998). Another viral enzyme that is probably important in reactivation is RR. This enzyme provides protein kinase (PK) activity which must be present for IE gene expression during reactivation (Smith et al. 1998). Given these factors, the mechanism of reactivation is not completely understood, but it is clear that viral or cellular activating proteins must somehow "overcome" latency-associated down-regulation for productive viral replication to occur.

IV. THE CYTOMEGALOVIRUSES

Like other herpesviruses, the CMVs are enveloped, double-stranded DNA viruses. They possess the largest genomes of all members in the herpesvirus family. The HCMV genome consists of 230 kilobase pairs (kbps), while other CMVs such as guinea pig CMV, are somewhat larger at 240 kbps. The HCMV genome has been extensively studied and serves as the prototype for the primate CMVs. As seen in HSV-1, four isomers of the HCMV genome exist due to inversion of the long and short genome components. Other primate CMV genomes however, appear to exist as only one isomer.

A. Pathogenecity

Of all viruses in the family *Herpesviridae*, HCMV is one of the most prevalent in humans. Approximately 50-90% of all individuals in developed countries will serologically test positive for HCMV, and seroprevalence increases with age (Staras et al. 2006). Since the virus is often latent in kidney tubules, mammary glands or salivary glands of infected individuals, it is often acquired in infancy or childhood from infected adults or children (Homa and Brown 1997). The latter is especially true in day-care centers where infected children in close contact with uninfected children efficiently spread the virus. In individuals with normal, healthy immune systems, HCMV infection is usually subclinical. However, individuals who are immunocompromised may be severely affected by the virus (Hill et al. 1995). AIDS patients suffer numerous CMV-related diseases. Organ transplant recipients, who use immunosuppressive drugs to avoid tissue rejection, are at risk of developing symptoms from reactivation of their own latent HCMV and are also at risk of acquiring HCMV from the transplanted donor organ. In addition to causing disease in immune deficient individuals, HCMV is also capable of

crossing the placenta and infecting a fetus in utero. This often leads to stillbirth, miscarriage, or severe congenital complications (Hayward et al. 1975; Ishov and Maul 1996; Michaels et al. 2003).

The prototype HCMV laboratory strain AD169 has a relatively large genome of 230 kbp. By 1989, the entire genome of AD169 had been mapped and sequenced (Chee et al. 1990). The map reveals unique sequences (where most genes are located) and repetitive sequences that contain very few genes. The genome map is divided into regions designated the terminal repeat long (TRL), unique long (UL), internal repeat long (IRL), internal repeat short (IRS), unique short (US), and terminal repeat short (TRS) similar to that shown in Figure 1 for HSVI. As seen in HSV-1, there are four different isomers of the genome that represent inversions of either the UL or US regions of the genome (Figure 1). Additionally, this particular strain contains over 200 potential ORFs (Chee et al. 1990). Protein products for many of these ORFs have been identified and expressed *in vitro* to discover the structure and function of the encoded proteins. However much of the genome is still being evaluated to determine the functions of other proteins.

The genes expressed by the virus are categorized as immediate early (IE) early (E) or late (L). In general, IE genes produce functional proteins and enzymes that suppress cellular activities and stimulate E gene expression. E genes are associated with viral DNA synthesis and nucleic acid metabolism, and L genes primarily produce structural proteins. This appears to be a sequential process involving transcriptional regulation in which IE and E genes must be expressed to allow L gene expression.

V. RESEARCH OBJECTIVES

The goals of this research project were to sequence the unique regions of the BaCMV genome, identify important genes, and to compare these genes with homologous genes of related CMVs. When this project began, the only completely sequenced and published CMV genome was that of HCMV strain AD169 (accession #X17403). During the course of this research, the chimpanzee (CCMV) and rhesus monkey (RhCMV) cytomegalovirus genomes were sequenced and published as well (AF480884; AY186194). The BaCMV sequence provides an additional model system for research on CMVs. Also, comparison of BaCMV strain OCOM4-37 sequence data with closely-related primate CMV genomes will provide new information regarding similarities and differences between these viruses. Sequencing, in and of itself, is a valuable component of basic science, but it is also useful because it provides a "roadmap" to identify genes and hence speed up research. Consequently, the BaCMV OCOM4-37 sequence has been used to develop a diagnostic PCR assay (Ross et al. 2005) and to identify genes for expression of recombinant proteins for diagnostic purposes.

Because the HCMV, CCMV, and RhCMV genomes have been completely sequenced and published, they were compared to the BaCMV sequence. The BaCMV genome was expected to show greatest homology with CMV genomes of most closely-related primates since this has been observed for other herpesviruses. Additionally, by aligning conserved gene sequences from primate and other animal CMVs with homologous BaCMV sequences, we were able to show phylogenetic relationships that cluster herpesviruses into groups depending on the subfamily in which they belong. It is common for CMV strains passaged in the laboratory to lose genes that are critical for

infection in an animal host, but are not essential for growth in cell culture. Therefore, a difference in gene number was expected between clinical isolates and laboratory strains. HCMV strain AD169 is a laboratory strain that has previously been shown to lack genes found in clinical isolates of HCMV such as Toledo and Towne (Murphy et al. 2003).

CHAPTER THREE

METHODOLOGY

I. CHEMICAL REAGENTS AND MATERIALS

A. Sources of Chemical and Biological Reagents

Amersham
Hybridization bags

Biorad
Xylene cyanol

BDH Incorporated
Chloroform, [ethylenediamine] tetracetic acid (EDTA), EtOH, ether, formamide, H_2O (HPLC grade), methanol, SDS, PEG, phenol, Tris[hydroxymethyl]-aminomethane (Tris)

Boehringer Mannheim GmbH
Digoxigenin (DIG) Oligonucleotide Tailing Kit

Fisher
Boric acid, calcium chloride ($CaCl_2$), disodium phosphate (Na_2HPO_4), potassium phosphate monobasic (KH_2PO_4), monosodium phosphate (NaH_2PO_4), sodium bicarbonate, ($NaHCO_3$), sodium citrate ($Na_3C_6H_5O_7$), sodium chloride (NaCl), sodium hydroxide (NaOH)

Gibco
Delbecco's modified Eagle medium [high glucose 4500 mg/L, with L-glutamine catalogue number 320-1965AJ] (DMEM), fetal bovine serum (FBS)

Kodak
1D Image Analysis Software ver. 3.5

New England Biolabs
GPS-1 Genome Priming System Kit, 1Kb plus DNA ladder, λDNA/ *Hind* III fragment ladder, Klenow enzyme

Pharmacia
Dideoxy-, and deoxy-nucleotide triphosphates, Lambda DNA, Pancreatic RNase

Polaroid
DS-34 camera and 3000 ISO type 667 film

Qiagen
GelPilot 1 kb plus ladder

Schleicher & Schuell
Protran nitrocellulose transfer membrane

Sigma
Bovine serum albumin (BSA), bromophenol blue (BPB), ethidium bromide (EtBr), ampicillin, salmon sperm DNA, proteinase K.

Stratagene
pBluescript II SK(+) (phagemid vector)

TAKARA Bio Inc.
LA PCR Kit ver. 2.1

B. Enzymes

Pharmacia
DNA polymerase (Klenow fragment),and T4 DNA ligase

Pancreatic RNase
Pancreatic RNase was boiled to remove DNase activity, and suspended in 10 mM Tris-HCl pH 7.5, and 15 mM NaCl at a final concentration of 10 mg/mL.

Sigma
Proteinase K (Sigma P 0390) 20 mg/mL stock solution in sterile dH_2O.

Restriction Endonucleases (R.E)
R.E's were obtained from Boehringer Mannheim, BRL, New England Biolabs, Pharmacia, and Promega. They were used with the supplied buffers at recommended concentrations and temperatures.

C. Preparation of Buffers, Solutions and Media

Cell Culture Media
If the media was older than 3 months, L-glutamine was added as a supplement to a final concentration of 0.1 mM. Gentamycin 25 mg per mL stock solution was used at a final concentration of 25 mg per L.

1X Phosphate-buffered Saline Buffer (1X PBS)
A solution of 0.8% NaCl, 2% KCl, 1.5% NaH_2PO_4, 2% KH_2PO_4 was dissolved in H_2O and sterilized by autoclaving.

Pre-hybridization Solution
A solution of 5X SSC, 50% de-ionized formamide, 0.1% SDS, 1mM EDTA, 25 mM NaH_2PO_4, 1 mg Salmon Sperm DNA was prepared.

20X Standard-Sodium Citrate (20X SSC) buffer
A solution of 3 M NaCl, 300 mM Sodium citrate was dissolved in H_2O for use as a stock solution. 1X and 5X SSC buffers were prepared by dilution of this stock solution.

6X Stop (Agarose Gel Loading Buffer)
A solution of 2.5% Bromophenol Blue, 2.5% xylene cyanol, 30% glycerol was dissolved in H_2O.

DIG Wash Solution (2X SSC)
A solution of 10 ml of 20X SSC, 1.0 ml of 10% SDS was mixed with 89 ml of distilled H_2O.

DIG Wash Solution (0,1X SSC)
A solution of 0.5 ml of 20X SSC, 1.0 ml of 10% SDS was mixed with 89 ml of distilled H_2O.

DIG Buffer 1
A solution of 0.1 M maleic acid and 0.15 NaCl was dissolved in 98.5 ml distilled H_2O and pH adjusted to 7.5 with solid NaOH.

DIG Blocking Solution
A solution of 1.0 g of blocking reagent (Carnation powdered milk) and 100 ml of DIG buffer 1 was mixed with stirring at 65 ºC to dissolve.

DIG Buffer 3
A solution of 0.05 ml of 1 M Tris-HCL pH 9.5, 1.0 ml 5M NaCl and 0.1 ml of 1 M $MgCl_2$ was added to 48.85 ml of distilled H_2O.

DIG Color-substrate Solution
A solution of 10 ml of DIG Buffer 3, 45 µl of NBT-solution (vial 9 from DIG kit) and 35 µl of X-phosphate solution (vial 10 from DIG kit) was mixed.

Resuspension Buffer
10 mM Tris 7.4
3 mM MgCl2
10 mM NaCl

Proteinase K digestion solution
2 ml cell lysate (1 ml water/150mm plate)
1925 µl water
40 µl 1M Tris
40 µl 0.5M EDTA
200 µl 10% SDS
25 µl proteinase K (20 mg/ml)

SOB Media (without magnesium)
20 g tryptone
5 g yeast extract
0.584 g NaCl
0.186 g KCl
H_2O to 1 liter
pH adjusted to 7.4 with NaOH

SOC Media
98 ml SOB medium
1 ml 2 M Mg^{2+} Stock
1 ml 2M glucose

2M Mg2+ Stock Solution
20.33 g $MgCl_2$ -6 H_2O
24.65 g $MgSO_4$ -7 H_2O
H_2O to 100 ml
Autoclave (or filter sterilize) and store at RT

2M Glucose
36.04 g glucose
H_2O to 100 ml
Filter sterilized and stored at RT

Wash Buffer for Electroporation
100 ml glycerol
900 ml H_2O

0.5X TE
5 mM Tris-HCl
0.5 mM EDTA (pH 7.5)]

Saturated NaI
Solid NaI was added to 100 ml 10 mM Tris/5mM EDTA in a dark glass container while stirring. Additional NaI was added until no more would dissolve. The solution was stirred overnight to ensure saturation. The NaI solution was stored at RT and allowed to stand overnight before use.

Tris-Borate-EDTA (TBE) Buffer
A solution of 0.089 M Tris-HCl pH, 0.089 M Boric acid, 0,002 M EDTA was dissolved in H_2O.

Tris-EDTA (TE) buffer
A solution of 10 mM Tris-HCl pH 8.0, 1 mM EDTA was dissolved in H_2O and sterilized by autoclaving.

2YT/AMP media
A solution of 16 g Bacto Tryptone, 10 g Bacto Yeast Extract, 5 g NaCl was dissolved in approximately 900 ml of distilled H_2O and pH adjusted to 7.0 with NaOH. Total volume was increased to 1L and the solution was sterilized by autoclaving. The autoclaved media was cooled to room temperature (RT) and then ampicillin was added to 100 μg/mL.

Agar Plates
2YT + 100 μg/ml Amp
1.5 % agar

D. Sources of Materials

Biorad
1.5 microfuge tubes

Corning
2 x 3 well cell culture plates, 96 well cell culture plates, 15 x 60 mm petri dishes, 20 mm x 100 mm petri dishes, 750 cm^2 cell culture flasks, 1500 cm^2 cell culture flasks

Falcon
15 mL conical centrifuge tubes

Perkin Elmer
GeneAmp PCR 2400

Promega
Wizard-Prep suction apparatus model #9993

Qiagen
Midiprep rack, catalog #19014

Schleicher & Schuell
Dot Blot apparatus

Whatman
3mm filter paper, nylon mesh (75 μM)

1. Kits

GPS™-1 Genome Priming System

Kit Components:

- Control Target Plasmid (LITMUS 28)
- pGPS1.1 Transprimer Donor
- pGPS2.1 Transprimer Donor
- PrimerN (30-mer)
- PrimerS (30-mer)
- Start Solution (20X)
- TnsABC* Transposase
- GPS™ Buffer Pack (10X)

Promega Wizard PCR Prep

Kit components

- Wizard® PCR Preps DNA Purification Resin
- Direct Purification Buffer
- Wizard Minicolumns
- Syringe Barrels (3cc)

Qiagen Midi and Maxi Prep

Kit Components

- QIAGEN tips (midi or maxi)
- QIAfilter Midi or Maxi cartridges
- Caps for QIAfilter
- Buffer P1, Buffer P2 and Buffer P3
- Buffers QBT, QC, QF
- RNase A (ribonuclease)

TaKaRa LA PCR 2.1

Kit Components

- *TaKaRa LA Taq*™ (5 units/µl)
- dNTP Mixture (2.5 mM ea.)
- 10 × LA PCR Buffer II (25 mM Mg^{2+} plus)
- 10 × LA PCR Buffer II (Mg^{2+} free)
- $MgCl_2$ (25 mM)
- Control Template (100 ng/µl genomic DNA derived from HL60)
- Control Primer LA3 (10 µM)
- Control Primer LA4 (10 µM)
- λ- *Hind* III digested MW Marker (100 ng/µl)
- 2 × GC Buffers I and II (5 mM Mg^{2+} plus)
- Control Primers GC1 and GC2 (10 µM)

E. Bacterial Strains

Gibco
MAX EFFICIENCY DH10B Competent Cells
E. coli F- mcrA Δ(mrr-hsdRMS-mcrBC) Δ80dlacZ ΔM15 ΔlacX74 deoR recA1 araD139 Δ (ara, leu) 7697 galU galK rpsL endA1 nupG.

Invitrogen
DH5αF'IQ^{TM} Competent Cells
E. coli F'φ80d*lac*ZΔM15 Δ*(lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *hd*R17 (r_k^-, m_K^+) *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1/F' *pro*AB$^+$ *lacI*^{q}ZΔM15 *zzf*::Tn5 [Km^R]

TOP10F' DH5αF'IQ^{TM} Competent Cells
E. coli F´[*lac*Iq Tn*10* (TetR)] *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G

F. Vectors

BACPAC Resource Center
Children's Hospital Oakland Research Institute
Oakland, California, USA.
pBACe3.6 (bacterial artificial chromosome)

Invitrogen
pCR2.1-TOPO and pCR-XL-TOPO

pSL301 (Brosius, 1989)

New England Biolabs
pNEB193

Stratagene
pBluescript II SK + (Alting-Mees and Short, 1989)

II. BACTERIAL CELLS

A. Propagation and Storage of Bacterial Cells

Stocks of *E. coli* were prepared by growing cells to late logarithmic phase in 2YT media, in a New Brunswick air incubator at 37 °C with vigorous shaking. Sterile glycerol was added to an approximate final concentration of 25%. Cells were aliquoted

into 1.5 ml microtubes (Biorad) and stored at -80 °C. DH5αF' cells were obtained from Gibco/BRL as a glycerol stock and stored at -80 °C.

Cells containing a plasmid were grown to logarithmic phase in appropriate selective media and stored at -80 °C after addition of glycerol to an approximate final concentration of 25%.

B. Preparation of Frozen Competent *E. coli* by the Rubidium Chloride Procedure

This procedure produces cells with a transformation efficiency of ~ 10^7 colonies/μg. The protocol and solutions are described in Sambrook et al. (1989). Competence was determined by transformed colonies produced per μg of supercoiled plasmid DNA. These cells were used for most laboratory cloning experiments. However, when transforming cells with very small quantities of DNA, commercial super-competent cells were used (Stratagene 5 x 10^9 colonies/μg, Invitrogen 1 x 10^9 colonies/μg).

C. Transformation of *E. coli* Using Frozen Competent Cells Prepared by the Rubidium Chloride Procedure

The basic protocol described in (Sambrook et al. 1989) was used. DNA samples were brought to 50 μl with HPLC H_2O and placed on ice. The frozen, competent cells were quickly thawed, 3 μl DMSO added and mixed by inverting the tube. Then 50 μl of competent cells were added to each tube of DNA and mixed by inversion. After a 30 min incubation on ice, the transformation mixtures were heat-shocked by placing the tube at 42 °C for 90-120 sec. 2YT broth (400 μl) was added and the mixture incubated at 37 °C for 30-45 min to allow the antibiotic resistance gene to be expressed. If β-GAL blue/white selection was required, 10 μl 1M IPTG was added to each tube. To obtain individual colonies, 100-200 μl of the transformation mixture was spread using a sterile

glass "hockey-stick" onto each agar plate. The agar plate contained the appropriate level of antibiotic and if β-GAL blue/white selection was desired, the plate also included X-GAL. The plate was incubated at 37 °C for 12-18 hr until colonies were visible.

D. Transformation of BACs Using Electroporation

Transformation of BACs was carried out by electroporation. The transformation efficiency was about 40 to 1,500 transformants from one μl of ligation product, or 20 to 1000 transformants/ng DNA. One to 2.5 μl of the BAC ligation material was transformed into 20-25 μl of *E. coli* DH10B competent cells (Gibco BRL, USA) using the BRL Cell-Porator system (Gibco BRL) according to the manufacturer's protocol. The parameters for electroporation were as follows: Voltage: 400V, Capacitance: 330 μF, Impedance: low ohms, Charge rate: fast, Voltage Booster resistance: 4 Kohms. The electroporated cells were transferred to 15 ml culture tubes with 0.4-1 ml SOC and shaken at 220 rpm for 50 min at 37 °C. The SOC medium was spread on one or two LB plates containing 12.5 μg/ml chloramphenicol and incubated at 37 °C for 20-36 hr. White colonies were selected and transferred to 96-well microtiter plates containing LB freezing media. The microtiter plates were incubated at 37 °C overnight then stored at 80 °C.

III. CELLS AND VIRUSES

A. Mammalian Cells

Human foreskin fibroblast (HFF) cells were obtained from Dr. J. Waner, Oklahoma University Health Science Center (OUHSC) (Oklahoma City, OK). Human embryonic lung fibroblasts (MRC-5) were purchased from ATCC (Manassas, VA). The Towne strain of HCMV was generously provided by Dr. J. Zaia, City of Hope National

Medical Center (Duarte, CA) while RhCMV 68-1 and HCMV AD169 were provided by Dr. P. Barry, University of California at Davis (Davis, CA).

B. Isolation of Baboon CMV

The BaCMV OCOM4-37 strain was isolated from an oropharyngeal swab of an olive baboon (*Papio cynocephalus anubis)* at the OUHSC colony (Oklahoma City, OK). The isolation and characterization has been described in detail elsewhere (Blewett et al. 2001). The isolate was plaque purified by two rounds of limiting dilution. The supernatant from positive cultures was added to a subconfluent monolayer of fibroblasts. At 4 hr PI the medium was removed and replaced with medium containing 2 % methylcellulose. At 5-7 days PI, isolated plaques were aspirated with a Pasteur pipette and inoculated onto fresh sub-confluent fibroblast monolayers. These infected cell cultures were passaged until viral cytopathic effects (cpe) was evident in >75% of the monolayer, harvested, and used to grow working virus stocks.

IV. BASIC TECHNIQUES

A. Agarose Gel Electrophoresis and Visualization

Agarose gel concentrations varied between 0.7% and 1.5%, with lower concentration gels used for high molecular weight DNA. Gels were cast in a Biorad horizontal gel apparatus. 1X TAE was used in the gel and as a running buffer. A loading solution of 6X Stop was added to samples containing 0.5-2.0 μg of DNA. Bacteriophage Lambda DNA, digested with *Bst* EII or *Hind* III, was used as size markers. DNA was loaded into 1 mm wide wells and electrophoresed at 40-80 volts to separate DNA fragments. DNA was stained in a dilute solution of EtBr (0.01 μg/uL) and rinsed with deionized H_2O. The DNA was visualized under short-wave ultraviolet (UV) light

(Fotodyne-3-3000). Images were captured using a KODAK 1D Image Analysis Software version 3.5.

B. Oligonucleotide Synthesis

Oligonucleotide syntheses were performed at the Recombinant DNA/Protein Resource Facility at Oklahoma State University, Stillwater, OK.

V. DNA PURIFICATION AND ANALYSIS

A. Purification of Herpesvirus DNA

Ten T100 flasks of MRC-5 cells at 25% confluency were infected with BaCMV at an m.o.i. of 0.1 - 0.5 pfu per cell. These cultures were maintained, trypsinized, and re-plated until the majority of cells in the culture showed cpe. Uninfected MRC-5 cells were then added to the flasks to produce 100% confluency. After 36 - 48 hr the flasks were subjected to a 37 ºC - 80 ºC freeze/thaw cycle twice, followed by scraping of the flasks. Cell debris was concentrated by centrifugation and the supernatant containing cell-free virus was removed and pooled. The pellet of cell debris was resuspended in resuspension buffer and cells gently ruptured with a Dounce homogenizer. Nuclei were collected by centrifugation and the supernatant containing virus saved. The supernatants were then centrifuged at 24,000 rpm at 25 ºC for 60 min in a SW27 rotor to concentrate the virus particles. Virus particles were resuspended in 2 ml of TE, 2 ml of a Proteinase K digestion solution added, and the mixture incubated at 37 ºC for at least 2 hr. Saturated NaI (6 ml) and 100 µl of 10 mg/ml EtBr were added (in a dim room), mixed, and transferred to a 15 ml heat-seal Beckman ultracentrifuge tube. After balancing and sealing, the tubes were centrifuged at 44,000 rpm for 64 hr. After removal from the centrifuge rotor, the DNA bands were visualized under UV light to identify the wide, top

cellular DNA band and the lower, narrow viral DNA band. Viral DNA was collected using a syringe with an 18 gauge needle, and the DNA extracted 5 times with equal volumes of chloroform to remove EtBr. The aqueous solution was transferred to dialysis tubing and dialyzed 6 -18 hr against 4 liters 10 M Tris/0.01 mM EDTA at 4 ºC to remove the NaI. The DNA in the resulting solution was precipitated with (NH_4)OAc and EtOH, collected by centrifugation, dried, resuspended in 200 µl TE buffer, and stored at 4 ºC.

B. Circular DNA Preparations

During replication in host cells, herpesvirus genomes circularize. To enhance the likelihood of recovering circular viral DNA, a specific protocol was used to selectively extract large circular viral DNA molecules from virus-infected mammalian cells. The procedures are described fully in Hirt (1967) and were adapted as described (Messerle, et al. 1997). To prepare circular BaCMV OCOM4-37 DNA for cloning, human fibroblasts were infected and incubated until cpe was general in each of eight T150 flasks. The medium was removed and scraping solution (0.6% SDS, 10 mM EDTA pH 7.5) was added at a volume of 5 ml for every 75 cm^2 of cells in the flask. This was incubated at RT for 10 – 20 min, the cell monolayer scraped off and transferred to centrifuge bottles. A ¼ volume of 5 M NaCl was added to give a 1M NaCl final concentration and the bottles were mixed gently inverting the tube 10 times. Cellular DNA and proteins were precipitated by incubating the bottles overnight at 4 ºC in an ice bucket. After centrifugation at 17,000 g for 30 min, the supernatants were transferred to clean bottle, gently extracted with phenol:chloroform, and the DNA precipitated with NaOAc and ETOH. After 200 µl of TE was added to the pellet, the DNA was allowed to rehydrate overnight in a 4 ºC refrigerator.

C. Purification of Plasmid DNA

Small-scale purification of plasmid DNA was performed using the alkaline-lysis procedure (Maniatis et al. 1982). Pancreatic RNase was added to miniprep DNA at a final concentration of 0.1μg/ml to remove contaminating RNA. Large scale preparation of plasmid DNA by alkaline-lysis and Qiagen maxiprep purification was performed essentially as described by the manufacturer's protocol. Lysozyme was not used in either purification procedure.

D. Recovery of DNA from Agarose Gels

DNA fragments (1-10 μg) were separated by agarose gel electrophoresis and visualized. Gels were stained in dilute EtBr and visualized briefly under long-wave UV to avoid nicking the DNA strands. The desired band was excised and placed in a 0.5 ml microtube with a punctured bottom covered with silicanized glass wool (Heery et al. 1990). This tube was nested inside a 1.5 ml microtube, centrifuged at 10,000 g for 10 min at RT. The buffer in the 1.5 ml tube containing the DNA was extracted with phenol:chloroform, and then precipitated using 7.5M (NH_4)OAc and EtOH (Maniatis et al. 1982). The DNA was resuspended in TE and the quality and quantity determined by running a sample on an agarose gel and comparing it to a known standard such as lambda virus DNA.

E. Transfer of DNA to Membrane

Southern blotting was used to transfer DNA from an agarose gel to a Nytran membrane (Maniatis et al. 1982, Southern 1975). DNA was separated by electrophoresis on a 0.8% agarose gel, stained and photographed. DNA fragments were depurinated by incubating the gel in 0.25 M HCl treatment for 15 min. The DNA was then denatured by

soaking the gel for 3 x 20 min periods in 0.4 N NaOH. A sponge was cut to the same size of the gel, placed in a glass Pyrex dish, and two pieces of 3 mm filter paper were placed on top. The gel was placed upon the filter paper. A Nytran membrane was cut to the same size as the gel and placed directly onto the gel. Two similarly-sized sheets of 3mm filter paper were placed on this membrane. A 6" stack of paper towels 2 mm narrower than the gel was placed on the filter paper and a light weight (250 g) set on top. The dish was filled with 0.4 N NaOH to the top of the sponge. During 12 - 16 hr, the paper towels were removed and replaced, and transfer buffer was replenished as necessary. After transfer, the position of the wells was marked on the Nytran membrane and the DNA cross-linked to it by a 30 sec burst of short-wave UV light. The membrane was stored at -20 °C until the hybridization step.

F. Southern Hybridization

Filters were incubated at 45 °C for 6 - 8 hr in hybridization bags containing prehybridization solution. Digoxigenin (DIG)-labeled probe preparations (concentrations varied between 0.1 μg - 5 μg) were boiled for 2 min, added to the bag and incubation continued overnight. The filters were removed from the bag and washed once for 15 min with 5X SSC at RT, three times for 15 min with 2X SSC, 0.1% SDS at RT, three times for 15 min at 45 °C and dried. DIG-labeled probe binding was detected using the manufacturer's protocol and is described in more detail on page 64.

VI. DNA SEQUENCING AND COMPUTER ANALYSIS TOOLS

DNA sequencing was performed at the Recombinant DNA/Protein Resource Facility at Oklahoma State University, Stillwater, OK. Sequencing of random fragments

from some clones was performed at University of Oklahoma ACGT Facility (Dr. Bruce Roe) in Norman, OK.

DNA sequencing traces were assembled using Seqman (DNAStar, Madison, WI). DNA and protein sequences were manipulated and aligned for similarity using EditSeq and MegAlign (DNAStar, Madison, WI). Primer sequences were designed using PrimerSelect (DNAStar, Madison, WI.). MapDraw (DNAStar, Madison, WI.) and ORF Finder (http://www.ncbi.nlm.nih.gov) were used to locate potential coding sequences. The Basic Local Alignment Search Tool (BLAST) was used to search Genbank for similar DNA or protein sequences (Altschul et al. 1990). CLUSTAL was used for aligning DNA and protein sequences for phylogenetic analysis (Higgins and Sharp 1989). MEGA and PHYLIP were used to infer phylogeny amongst aligned sequences (Felsenstein 1993; Kumar et al. 2001; Tamura et al. 2007). CorelDraw 8 (Corel Corporation, Ottawa, Ontario) was used in the preparation of figures.

VII. SCREENING FOR TERMINAL CLONES OF THE BaCMV GENOME

A. Directional Cloning of Terminal Fragments

1. Viral DNA End-Filling

NaI purified BaCMV OCOM4-37 DNA was used for directional cloning. The 5' and 3' terminal ends of viral DNA were filled in using dNTPs and Klenow fragment of DNA polymerase to create blunt ends. This was accomplished by mixing 5 µl (0.5 µg – 1.0 µg) of viral DNA with 1 µl of Klenow enzyme (2U/µl), 2 µl Klenow buffer, 0.66 µl 10 mM dNTPs and water to make a 20 µl total volume. The mixture was incubated for 15 min at 25 ºC. Klenow enzyme was inactivated by heating to 75 ºC for 10 min.

2. Vector DNA Preparation

A plasmid vector had to be prepared that would accept blunt-ended terminal fragments of the OCOM4-37 genome. *Sma*I is a RE that cuts DNA to create blunt ends at the recognition site. Plasmid cloning vector pSL301 was used because it has a multiple cloning site (MCS) containing all unique six bp RE recognition sites including *Sma*I. The pSL301 vector has a single cutting site for *Bam*HI, *Hind*III, and *Not*I, so using each of these enzymes in combination with *Sma*I created vectors that would only accept fragments from blunted termini (e.g. blunt-*Bam*HI, blunt-*Hind*III and blunt-*Not*I).

Eight μg of pSL301 was digested with 10 units of *Sma*I for 2 hr at 25 ºC to create blunt ends. Enzymes were denatured by phenol:chloroform extraction, and the plasmid DNA was purified by ethanol precipitation. After drying and resuspension in TE (pH 8.0), a sample of the digested DNA was electrophoresed on an agarose gel and visualized to ensure the enzyme had digested the vector. A second series of digests were performed to create ends compatible with three different restriction enzymes, *Bam*HI, *Hind*III or *Not*I. *Sma*I digested vector DNA (2 μg) was cut with 10 units of appropriate RE and allowed to digest at 37 ºC overnight then extracted, precipitated, washed and dried as previously described. Each vector DNA pellet was resuspended in 30 μl TE. Four μl of vector was run on a 0.8% agarose gel and a single band was visualized at the expected DNA size (3.2 kb).

3. Digestion, Ligation and Transformation of the BaCMV OCOM4-37 Insert DNA

The *Hind*III cut vector, religated without insert, produced an excess of vector colonies so *Bam*HI and *Not*I -digested vectors were used for ligation reactions with BaCMV OCOM4-37 digests. Two tubes of viral DNA (0.3 μg) with filled-in termini

were digested for 2 hr at 37 ºC with *Bam*HI or *Not*I. This generated terminal genome fragments having one blunt, filled end and one sticky end (Figure 8). Each digest was extracted with phenol:chloroform, precipitated and dried. The air-dried digests of BaCMV DNA were resuspended in the following ligation solutions. One ligation reaction (5 µl) was prepared by adding 1 µl of *Sma*I/*Bam*HI cut pSL301 DNA (0.15 µg), 0.5 µl of ligation buffer, 1 µl of high concentration ligase (20U/µl) and 2.5 µl of distilled H2O. A second similar ligation was prepared using *Not*I instead of *Bam*HI. Both reactions were incubated for 4 hr at 16 ºC. All 5 µl of each ligation reaction was used in transformation of high efficiency commercially-competent Sure Shot TOP10 F' cells using manufacturer's protocol.

4. Preparing Miniprep Digests for Terminal Sequence Screening

After transformation, 7.1 µl of 1M IPTG and 250 µl of pre-warmed SOC media were added to each tube. Each transformation reaction (100 µl) was spread onto 2YT/AMP/XGAL agar plates and incubated overnight at 37 ºC. The number of blue colonies (representing self-ligated vector) was higher than expected but white colonies (containing inserts) were present on each plate. White colonies were selected for DNA purification and grown overnight in 2YT/AMP broth with shaking. Minipreps were performed using the standard protocol and resuspended in 100 µl of TE. For the digests of potential blunt- *Bam*HI clones, plasmid DNA (5 µl) was digested with *Eco*RI and *Bam*HI. These REs cut the pSL301 multiple cloning at sites 5' and 3' of the expected insert DNA. The potential blunt-*Not*I clones were digested with *Eco*RI and *Xho*I since these REs were expected to excise the insert. The digested DNA samples were electrophoresed on an 0.8% agarose gel, stained with EtBr and visualized under UV light.

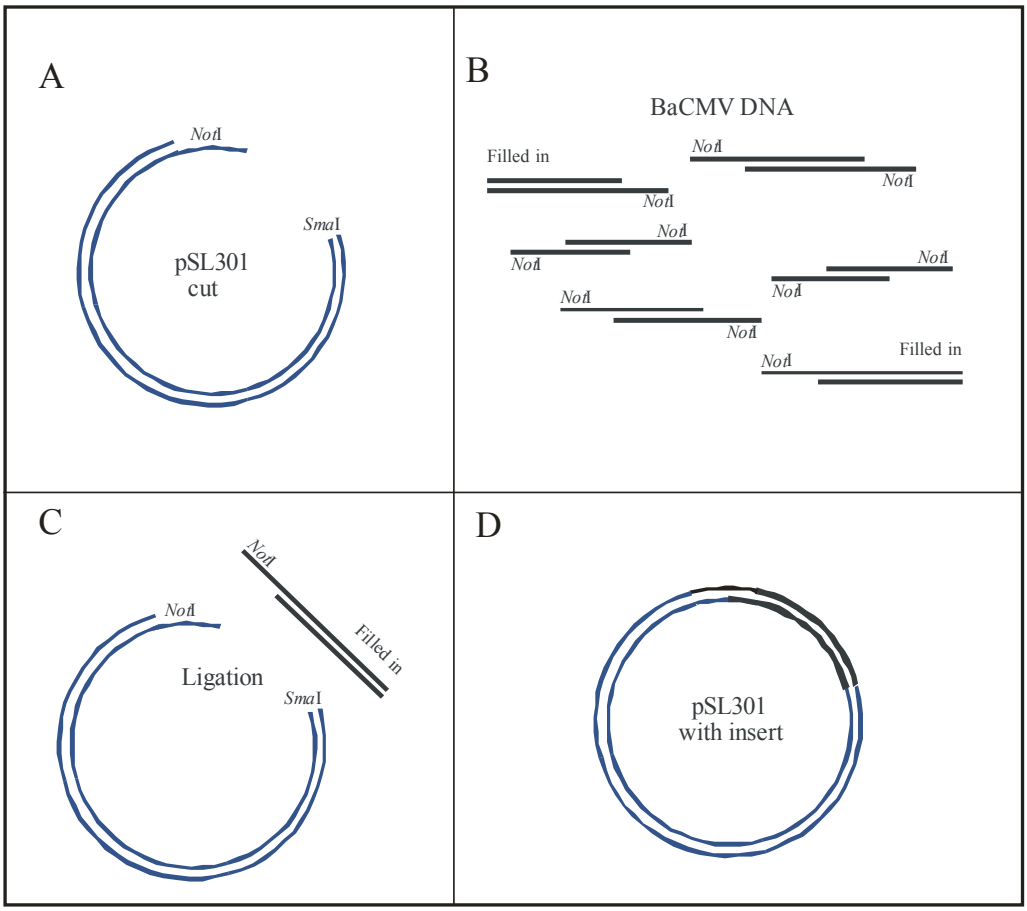


Figure 8. Directional Cloning of Terminal Fragments from the BaCMV Genome. Panel A shows pSL301 cut with *Sma*I and *Not*I restriction enzymes to produce a blunt-*Not* I linear vector. Panel B shows the BaCMV OCOM4-37 DNA genome with filled-in termini and then cut with *Not*I restriction enzyme. Panel C shows the successful ligation of a blunt-*Not* I BaCMV DNA fragment into the vector. Panel D shows the desired BaCMV clone.

B. Circularization of the BaCMV Genome

The previously-described attempt to obtain clones at each terminus using directional cloning generated a clone at the 3' terminus of the genome but no clone was found at the 5' end. Therefore a different strategy was undertaken. This strategy was to use circular BaCMV DNA cut with *Not* I. Ideally, the *Not* I digest would produce a fragment containing the ligated DNA from each terminus (Figure 9). Circular BaCMV DNA was obtained from infected cells as described in Methodlogy Section VII-B (circular DNA preparation) or by the method below.

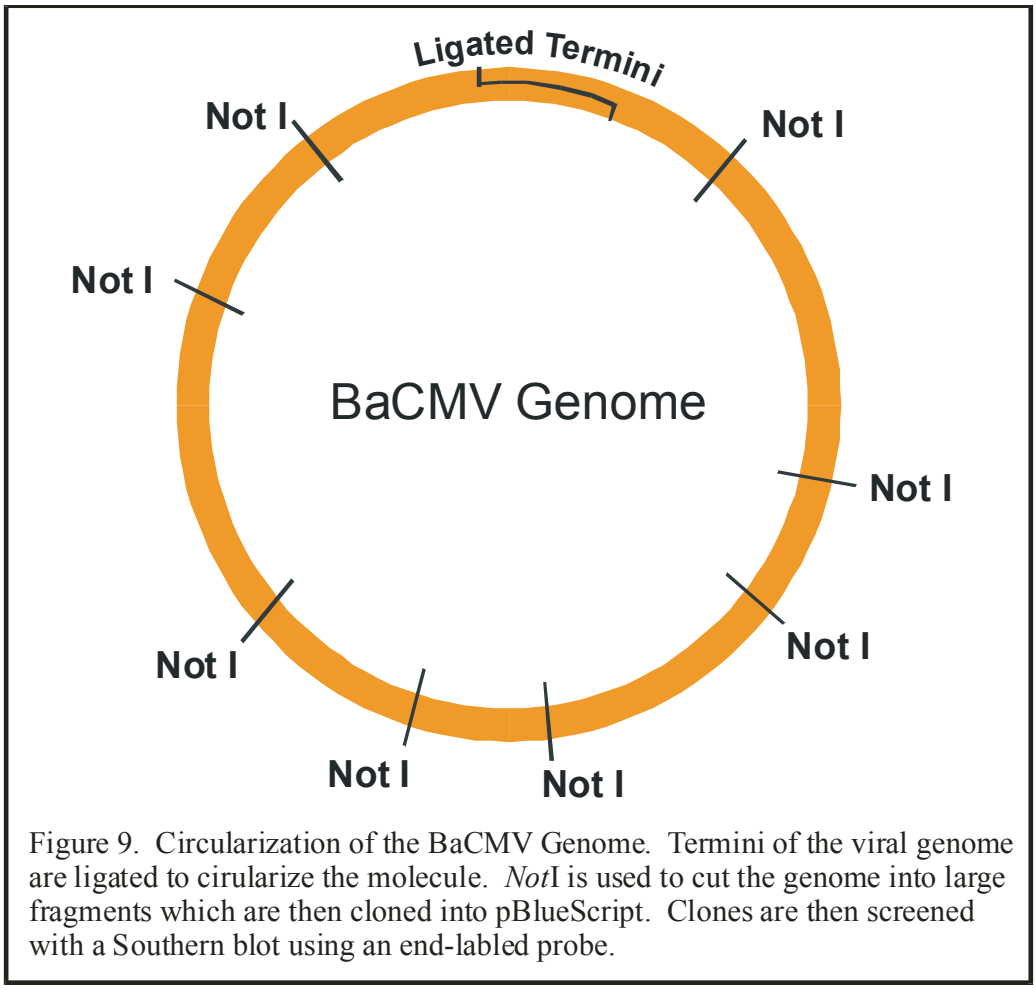


Figure 9. Circularization of the BaCMV Genome. Termini of the viral genome are ligated to cirularize the molecule. *Not*I is used to cut the genome into large fragments which are then cloned into pBlueScript. Clones are then screened with a Southern blot using an end-labled probe.

1. Ligation and Circularization Reaction

Eight µl of BaCMV OCOM4-37 DNA (1.0 µg/µl) was mixed with 1 µl of 10X T_4 DNA ligase buffer and 1 µl of T_4 DNA ligase (1.3 U/µl) for a final volume of 10 µl. The mixture was incubated in a PCR thermocycler at 16 ºC for 2 hr then at 65 ºC to denature the enzyme.

C. Cloning Circular BaCMV DNA

Circular BaCMV DNA digested with *Not* I would generate a fragment containing the annealed termini. The other *Not* I fragments would be useful genomic clones since they are likely large in size.

1. Not I *Digests of Ligated BaCMV DNA*

Ten µl of the ligation reaction was mixed with 50 units of *Not* I, 5 µl of React 3 buffer and 30 µl of distilled H_2O and incubated at 37 ºC for 2 hr. The mixture was then phenol:chloroform extracted, precipitated with EtOH, dried, and resuspended in 12 µl of TE.

2. Ligation of Digested BaCMV Circular DNA (Created by Ligation) and pBS/ Not I *-Cut Vector*

Two µl of pBS/ *Not* I -cut vector (0.25 µg/µl) was added to 6 µl of *Not* I -cut DNA (0.05 µg/µl), 1 µl of 10X T_4 ligase buffer, and 1 µl of T_4 ligase (1.3 U/µl) for a final volume of 10 µl. The mixture was incubated at 16 ºC overnight and then used to transform TOP 10F' cells using the standard protocol. Two hundred µl of the transformation reaction was spread on three plates containing 2YT/AMP/XGAL and grown at 37 ºC overnight. Twelve colonies were selected for growth overnight and plasmid purification using the standard protocol. Restriction digests were performed on

all white colonies and run on an agarose gel. All *Not* I clones having a vector and an insert band were saved for screening with a *Not* I end-labeled probe.

3. Not I *Digestion of BaCMV Circular DNA Prepared from Cell Culture*

Five µl of circular BaCMV DNA prepared from infected cells (approximately 0.8 µg) was combined with 1 µl of *Not* I (10 units), 2 µl of React 3 buffer, and 12 µl of distilled H_2O and incubated at 37 ºC for 2 hr. The mixture was then gently phenol:chloroform extracted, precipitated with EtOH, dried, and resuspended in 20 µl of TE.

4. Ligation of Not I *Digested BaCMV Circular DNA into the pBACe.3.6 Vector*

There are several reasons why cloning DNA into BACs was advantageous. First, large pieces of DNA exceeding 200 kbp can be cloned and maintained in a BAC. Cloning the 5' terminus of BaCMV proved difficult. It is possible that the BaCMV genome may not have useful RE recognition sites near the 5' terminus. In earlier projects, attempts were made to clone BaCMV DNA fragments larger than 15 kbp into plasmids. This is a low efficiency cloning strategy and may explain the lack of success. This approach used circularized DNA, thereby avoiding the need to manipulate the viral genome to "blunt" the ends. The joined termini would likely be a large piece of DNA, thus needing a BAC for efficient cloning. Another good reason to utilize BAC cloning is that DNA which is "difficult" to clone by other means can be maintained in a BAC. This is because there is only one or a very few copies of the BAC in each bacterial cell. Therefore we planned to ligate *Not* I digested circular BaCMV DNA into the vector pBAC e3.6 (Frengen et al. 1999), and electroporate the ligation reaction into commercially electrocompetent *E. coli*.

The pBACe3.6 vector contains a pUC origin, so the vector DNA can be easily prepared. This insert is in the MCS and flanked by RE sites. The vector was digested with *Bam*HI or *Not*I using standard procedures and the reaction then electrophoresed on an 0.8% preparative agarose gel. This digestion cuts out the pUC insert leaving the pBACe3.6 vector linearized with *Bam*HI or *Not*I compatible ends. The linearized vector was purified from the gel and dephosphorylated using the manufacturer's suggested procedure.

Two µl of pBACe.3.6/ *Not*I-cut vector (0.4 µg/µl) was added to 10 µl of *Not*I-cut DNA (0.4 µg/µl), 1 µl of 10X T_4 ligase buffer, 1 µl of T_4 ligase (1.3 U/µl) and 5 µl of distilled H_2O for a final volume of 20 µl. The mixture was incubated at 16 ºC overnight. Ligation reaction was drop dialyzed for 2 hr against 0.5X TE using a membrane with a 25,000 MW cutoff. The ligation reaction was electroporated into *E. coli* DH10B cells as described in section II-D. Over 50 colonies were obtained. Six colonies were randomly selected for growth overnight and BAC DNA purified using the standard BAC miniprep protocol. Restriction digests were performed on all six preparations and run on an agarose gel to determine cloning efficiency. All clones had inserts indicating high cloning efficiency. As many isolated colonies as possible were archived for later screening, and are represented as clones in the 3800 series.

D. Screening Pooled Clones and BAC Clones Using Dot-Blot Method

Existing BaCMV libraries were screened to find clones from each terminus that had not been previously identified. Clones that were not already sequenced were pooled in groups of 9 – 12 for screening. These grouped clones were used for dot blots to allow large numbers of clones (1,000+) to be screened simultaneously. The probes used for

screening were produced from existing clones already mapped to a site adjacent to regions of interest. Two sets of dot blots were prepared. The first blots contained the pooled groups mentioned above. The second blot contained individual clones that were positive from the first blot.

1. Probe Preparation for Dot-Blot Screening

A good deal of internal BaCMV genome sequence had been produced, but little sequence existed for the termini. To identify BaCMV clones from either terminus of the genome, probes were produced for screening purposes. One method used to produce probes was end-tailing of the genomic DNA using terminal transferase and DIG-labeled dUTP. The other method used was to label stock clones thought to be close to the termini with DIG using random hexamers as primers.

End-tailing reactions were prepared containing 0.2 µg of NaI purified BaCMV OCOM4-37 DNA, 8 µl of 5X reaction buffer, 8 µl of $CoCl_2$ solution, 2 µl alkaline-constant DIG-labeled 0.35 mM dUTP, 2 µl 1.0 mM ATP, and 2 µl terminal transferase (100 units) for a total volume of 40 µl. The mixture was incubated at 37 ºC for 15 min. Distilled H_2O (60 µl) was added for a final volume of 100 µl and the reaction was centrifuged for 60 sec. Qiagen PB buffer (500 µl) was added to the mixture and the DNA was purified by Qiagen miniprep technique (previously described). The dried DNA was resuspended in 50 µl of TE buffer. Two probes (*Not* I and *Pme*I) were prepared using 25 µl (0.1 µg) of end-tailed BaCMV DNA, 1 µl of *Not* I or *Pme*I, 3 µl of the appropriate 10X RE buffer and 1 µl dH_2O, and incubated for 2 hr at 37 ºC. Both of these REs recognize and cut at an 8 bp site producing large DNA fragments. Large fragments are likely to overlap and hybridize with more clones than smaller fragments would. Ideally, this

overlap could allow detection of clones at or near the genomic termini. Although all the fragments in the digestions would bind to BaCMV clone DNA, only the end fragments would be DIG-labeled. Ten µl of each probe was used to screen dot blots and the rest was stored at -20 ºC.

To screen our BaCMV library for clones at the left terminus, two additional probes were produced (from stock clones H639 and B165) using random hexamers as primers. Clones H639 and B165 had previously been mapped to the left terminus of the BaCMV genome. To generate the H639 probe, 10 µg of the plasmid clone DNA was digested with 3 µl of buffer, 1 µl of *Hind* III (10 U/µl) and 6 µl of distilled H_2O for 1 h. The digested DNA was run on an agarose gel overnight and the 5.5 kb BaCMV insert band was cut from the gel. The gel fragment was purified by centrifugation in a 0.5 ml test tube containing siliconized glass wool followed by phenol:chloroform extraction, ETOH precipitation, drying, and resuspension in 20 µl TE. Three µl of the H639 DNA was diluted to 15 µl using dH_2O and then denatured in a boiling water bath for 10 min. The test tube was immediately placed on ice and 2 µl DIG-labeled hexanucleotide mix, 2 µl 10 mM dNTP mixture and 1 µl Klenow enzyme (2 U/µl) was added. The solution was mixed, centrifuged briefly, and incubated for 60 min at 37 ºC. Two µl of 0.2 M EDTA (pH 8.0) was added to stop the reaction. The labeled DNA was precipitated by adding 2.5 µl 4M LiCl and 75 µl of -20 ºC EtOH. The solution was mixed and kept at -70 ºC for 30 min. Finally, the mixture was centrifuged for 15 min, washed, dried, and resuspended in TE. The same protocol was used to generate the B165 probe except that the amount of DNA was increased to 15 µg, the buffer to 4 µl, *Bam* HI to 3 µl, and H_2O to 3 µl in the

digest step. In the DIG-labeling steps, all components were increased 4X except for the Klenow enzyme. All other steps were completed per the manufacturer's protocol.

2. Pooling Clones and Preparing Membranes

All clones used for dot blot screening were from the BaCMV library of clones. Most of these clones had been pooled and designated as follows:

Pool numbers	Restriction enzyme used	Clone numbers
S1 – S70	*Bam*HI or *Hind*III fragments	Various (~710 total)
S71 – S78	*Eco*RI or *Xba*I fragments	2320 – 2399
S79 – S85	Blunt/ *Not*I fragments	3031 – 3101

Table 4. Pooled Clones Used for Dot Blot Screening

All clones were pooled in 1.5 ml tubes containing 9 – 11 clones each. Fifteen µl of each clone (0.2 µg/µl) was added to the pool and TE was added to increase the volume to 200 µl. Pools S1–S78 were screened with the *Not*I-cut probe and groups S1 – S85 were screened with the *Pme*I-cut probe. Positive controls included 20 µl of miniprep DNA (0.2 µg/µl) from QN3063 and 5 µl of midiprep (1.0 µg/µl) DNA from H639 (previously identified terminal clones). Negative controls included 20 µl of miniprep DNA (0.2 µg/µl) from QN3058, 5 µl (1.0 µg/µl) of midiprep H589 (both from internal regions) and 20 µl (1.0 µg/µl) of the pBlueScript SKII+ vector. Each of these control samples was increased to 200 µl with TE. After screening pooled clone groups S1 – S85, pooled clones that showed high color intensity for both dot blots (*Not*I and *Pme*I probed) were selected for individual screening.

Dot blots for all clone groups, or individual clones, using either probe were prepared as follows: 20 µl of each sample was added to 180 µl of TE, and 5 µl of 10 N NaOH added to denature the DNA. This mixture was incubated at 37 ºC for 5 min and chilled on ice. Two hundred µl of chilled 0.125X SSC was added to each tube to a final

volume of 400 µl. Two hundred µl of the DNA solution was loaded into separate wells of a dot blot apparatus overlaid with a nylon membrane cut to size. The corner of the membrane was notched to designate orientation. Two hundred µl of 0.125X SSC was added to wells without sample as controls. All samples were allowed to sit 30 min at RT and wells were vacuum-suctioned for 2 min to dry. The membrane was removed from the dot blot apparatus and air-dried at RT for 15 min. The membrane was placed face down on plastic wrap and exposed to UV light for 60 sec to cross-link DNA.

3. Hybridization and Washes

The nylon membrane was placed in a tupperware container and prehybridized in a 20X SSC solution for 1.5 hr at 65 ºC on a rocking platform. Ten µl (20 % of preparation) of the probe and 190 µl of TE were added to a 0.5 ml test tube. The cap of the tube was perforated to allow steam to escape and the probe was boiled for 5 min to denature the DNA and immediately placed on ice. The prehybridization solution was removed from the membrane and replaced with hybridization solution. The hybridization solution was heated to 65 ºC and the probe added. The membrane and probe were incubated at 65 ºC overnight with shaking, then removed and rinsed with 100 ml 2X SSC + 0.1% SDS wash solution. Two additional washes were performed at RT with shaking for 15 min using 100 ml of the same wash solution. Two more washes were then performed using 100 ml of 0.1X SSC + 0.1% SDS at 65 ºC for 15 min, each with shaking.

4. Colorimetric detection of DIG-labeled probe.

The membrane was soaked in 10 ml of DIG buffer 1 for 5 min then incubated with rocking in 100 ml of blocking solution. The solution was removed and the manufacturer's suggested volume of anti-DIG antibody (alkaline phosphatase-

conjugated) diluted in 20 ml of blocking solution was added. The membrane was incubated with rocking for 30 min then drained. Two washes were performed using 100 ml of DIG buffer 1 with rocking for 15 min each. Wash solution was removed after each wash, and finally the membrane was soaked in 20 ml of DIG buffer 3 (see p. 42). This soaking step was performed in the dark without rocking and checked periodically for color change. Development was stopped by washing the membrane in a large volume of TE. The developed membrane was placed in the dark to dry then photographed. After drying, 86 clones were selected from pools for individual screening using the same hybridization screening. Twenty-five of the 86 individual clones were digested with the appropriate RE to excise the insert and the digested DNAs run on an agarose gel. After eliminating duplicates or aberrant clones, a total of twelve individual clones were sent for sequencing (see Results II-C-1).

Dot blots using the H639 and B165 probes and pooled clones S1-S85 (and later individual clones) were performed using the same methods described above. Two clones (E2290 and E2363) were positive in the H639 screening so were chosen for additional testing.

Screening pooled clones with the B165 probe yielded positives from 4 wells. Clones from these 4 wells were used to prepare a blot to screen the following individual clones: B356 – 365, B614 – 623 and E2286 – 2290. From this dot blot two clones, B359 and E2290, were chosen for further analysis.

5. Screening Not *I Clones Using Southern Blot Method*

The following *Not* I clones generated by circularization of the BaCMV genome were selected for screening by Southern blot: *Not* I clones 3208, 3210, 3211, 3212, 3215,

3217, 3218, 3219, 3223, 3231, 3232 and 3234. Restriction digests were prepared using 3 µl of plasmid DNA (0.2 µg/µl), 0.5 µl of *Not* I (10 U/µl), 1 µl of buffer and 5.5 µl of distilled H_2O. Positive control DNA from terminally-located clones B165, H620, N1655 and QN3063 were also digested for screening. Each of the 3 digests contained 3 µl of DNA (0.5 µg/µl), 1 µl of buffer, 1 µl of the compatible restriction enzyme (10 U/µl) for that clone and 5 µl of distilled H_2O. The mixtures were incubated at 37 ºC for 2 hr and then run on an agarose gel overnight. Southern blot was performed. The blot was developed and clones 3208, 3210 – 3212, 3217, 3218, 3223, 3231, 3232 and 3234 were identified as positive. These clones were digested with *Hind* III and *Sac*I, electrophoresed on an agarose gel, and visualized to identify clones that were unique. The following five clones appeared to be unique and were sent to the OSU core facility for sequencing: 3208, 3211, 3217, 3231 and 3232.

VIII. SCREENING POOLED CLONES FOR CLONES IN GAP REGIONS USING SOUTHERN BLOT

There were three areas of the genome that the BaCMV library clones did not cover. A different strategy was used to find clones within these gaps. Two gaps were in the US region of the genome and the other was at the UL right terminus. To identify clones containing DNA from these regions, screening by Southern blot was performed using probes produced from clones bordering the gaps. Two libraries were chosen for this: the pooled clones (previously used for dot-blot screening) and the BAC clones 3826, 3828, 3854 and 3870. The BAC clones were quite large (>12 Kb) so were more likely to overlap gap areas than some of the smaller pooled clones. DIG-labeled probes were made from clones adjacent to the gaps and included B110, B315, H596 and H659.

A. Probe Preparation for Southern Blots

Probes from clones B110, B315, H596 and H659 were all produced by PCR using primers specific for each clone. PCR reactions contained the following components: 1 µl of plasmid DNA (0.2 µg/µl), 15 µl of 3.3 x L buffer, 5 µl of 25mM MgOAc, 1 µl of primer, 2 µl 10mM dNTPs, 1.7 µl DIG-labeled 1mM dUTP, 1.2 µl r*Tth* polymerase (2U/µl) and 15.7 µl dH_2O. Each PCR-generated probe was extracted, precipitated and washed. Then 0.5 µl of each probe was dotted on a nylon membrane and tested with anti-DIG antibodies. The intensity of color after development demonstrated the degree of digoxigenin-labeled dUTP incorporation into the probe. The B110 probe was darkest and the H569 probe showed the least color development. Labeling of the H569 probe was sufficiently poor that it was not used for screening individual clones.

B. Southern Blot Screening of pBlueScript Clones

Three separate screenings were performed using the B110, B315 and H659 probes. Individual clones were selected for screening based on the degree of color intensity seen on previously-screened dot blots of pooled clones. The following clones were digested with compatible restriction enzymes and loaded onto an agarose gel for electrophoresis: *Bam*HI clones 347, 349, 350, 351, 386, 419, 420, 422; *Eco*RI clones 585, 2293, 2302, 2303, 2311; *Xba*I clones 2246, 2247, 2262, 2263, 2270, 2273; and *Not*I clone 1709. The digested DNA was then transferred to a nylon membrane using the Southern blot method and each membrane was probed using one of the three probes and then developed.

C. Southern Blot Screening of BACs

BAC clone screening was performed somewhat differently than for individual pBlueScript clones. First, all BAC clones were screened on a single membrane. Second, all four probes (B110, B315, H569 and H659) were used. All BAC clones were double-digested with *Eco*RI and *Not*I and then electrophoresed on an agarose gel. The DNA fragments were transferred to a nylon membrane using the Southern blot method. The membrane was then probed and developed.

IX. GENERATING BaCMV CLONES OR PCR PRODUCT FOR SEQUENCING

A. Genome Sequencing by Transposon Priming

The GPS1.1 transposon mutagenesis system was used for sequencing clones having a large DNA insert. A good sequencing reaction produces 500 – 800 bp of DNA sequence. When newly identified clones were sent for sequencing, primers that bind near the vector MCS were used so sequence data from the ends of the clone were generated. Therefore most of the clones chosen for transposon mutagenesis already had data generated for the ends. These clones ranged in size from 1 kbp to >15 kbp, so a considerable amount of the internal region still needed to be sequenced. Transposon mutagenesis introduces a transposon containing universal priming sites in random locations on target DNA (Figure 10).

In a given reaction containing many copies of a plasmid, transposons insert in different locations on each individual plasmid thereby providing sequence priming at different points along the length of the clone. Transposon, target DNA, and transposase ratios are adjusted to favor insertion of a single transposon per plasmid. This technique was very effective and was used to provide internal sequence from at least 21 different

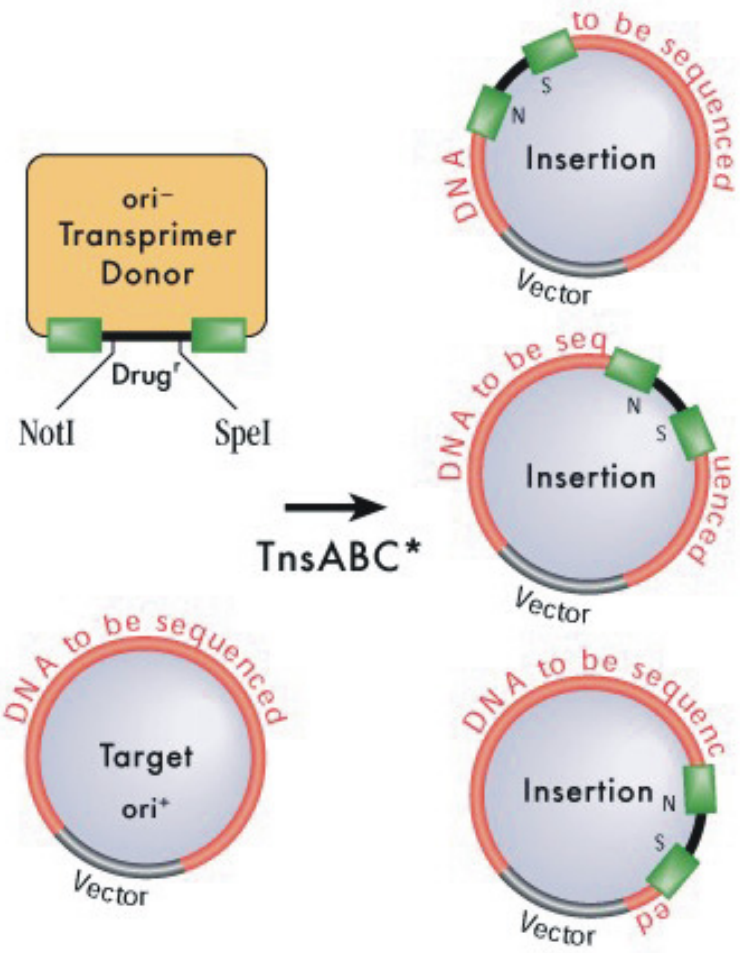


Figure 10. The GPS-1 Genome Priming System. The transposon (Transprimer) inserts DNA randomly into the DNA target. The recombinant transposon can then be transferred to bacteria via transformation. Genes for kanamycin and chloramphenical confer selection properties. Primers N and S allow sequencing from either strand at the insertion point. Figure taken from manufacturers product insert.

plasmid clones of BaCMV OCOM4-37. The following clones were chosen for GPS1.1 transposon mutagenesis: *Bam*HI – 97, 110, 112, 119, 129, 131, 150, 161, 201, 241, 301, 372, 406, 1612 and 3287; *Hind*III – 602, 619, 631 and 659; *Not*I 1690; and *Sma*I/*Not*I 3094.

For all transposon mutagenesis reactions, the concentration of target DNA was adjusted to 0.08 µg/ml. Concentrations were determined by gel electrophoresis of target

DNA with Lambda DNA of known concentration. In each reaction, 2 µl of GPS buffer, 1µl of GPS1.1 transprimer, 0.08 µg/ml of target DNA and dH_2O to bring the volume to 18 µl were mixed together in a 0.5 ml test tube. The contents were mixed by pipeting the solution a few times. One µl of transposase enzyme (mixture of TnsA 7µg/ml, TnsB 10µg/ml, TnsC 20µg/ml) was added and the solution again mixed. The mixture was incubated at 37 ºC for 10 min before 1 µl of start solution was added. The contents were mixed by pipeting and incubated for 1 hr at 37 ºC. The mixture was heated to 75 ºC for 10 min to inactivate the enzymes. The entire reaction (20µl) was transformed into commercially-competent TOP10 F' cells and placed in a shaking water bath for 30 min at 37 ºC. Seventy-five µl of the transformed cells were spread on 4 LB KAN/AMP plates and incubated overnight at 37 ºC. Results were variable but most reactions yielded enough colonies to pick to provide DNA for sequencing. Colonies were selected for minipreps, the DNA digested, and DNA fragments separated by gel electrophoresis

Based on the known sequence of the transprimer, useful *Bam*HI clones were expected to have four bands while *Hind*III clones were expected to have three bands. Digestion with *Bam*HI produces one 3 kb vector fragment and one 1.4 kb fragment containing the transprimer. The sizes of the remaining two bands are dependent on the insertion site of the transprimer. For *Hind*III digestion of the *Hind*III clones, a 3 kb vector fragment is also expected. Again, the sizes of the two remaining bands depend on the location of the transprimer insertion site. An example of RE digests, electrophoresis and a gel of clones from a GPS1.1 transprimer reaction is shown in Figure 11. Clones *Bam*HI 4416, 4418, 4428, and *Hind*III 4434, 4435 show the desired bands. Initially,

clones *Bam*HI 4428, and *Hind*III 4434, 4435 were sent for sequencing. The other clones were reserved for later sequencing.

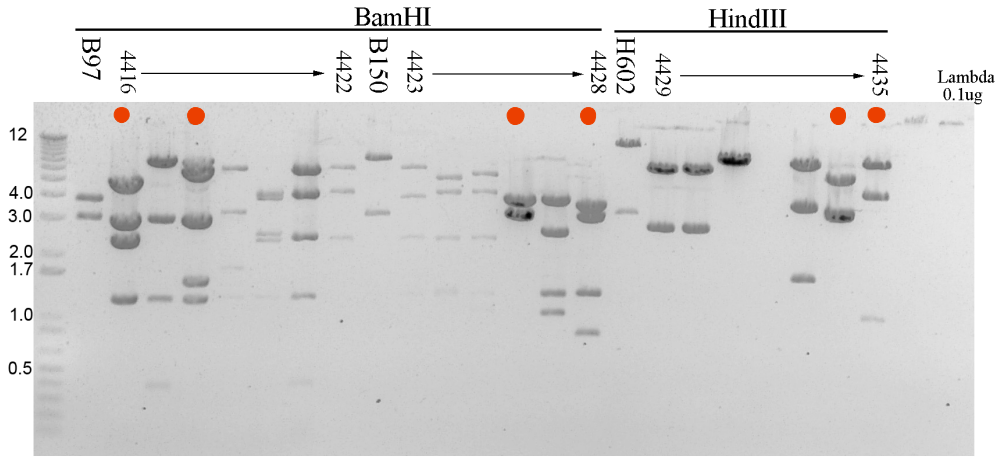


Figure 11. Transposon Mutagensis of pBS *Bam*HI 97, pBS *Bam*HI 150 & pBS *Hind*III 602. Using NEB GPS1.1. DNA size markers are in lane 1. Fragment sizes in bp: 100, 200, 300, 400, 500, 650, 850, 1000, 1650, 5000, 6000, 7000, 8000, 9000, 10000, 11000 & 12000. Useful clones are marked with dots.

B. Cloning and Sequencing the UL1 Region Using HCMV Primers SP216 and SP217

None of the preceding approaches yielded sequence from the UL1 region of the genome. Sequencing the coding regions of the genome was a primary goal of this study. The UL1 gene in HCMV is the first coding sequence in the genome. By amplifying and sequencing the UL1 region of the BaCMV genome, data adjacent to the terminus would be available as a starting point for primer walking into the repetitive terminal sequences upstream of the UL1 sequence. Primer pairs (SP216 and SP217) were designed for the UL1 region of the HCMV genome, and these primers were used to attempt amplification of the BaCMV UL1 region using PCR. It was expected that although the primers would not bind specifically to the BaCMV UL1 sequence, sufficient homology would exist between the two genomes to allow binding at low annealing temperature to produce a BaCMV UL1 PCR product. The PCR product was cloned into pCR2.1-TOPO,

transformed into competent cells, DNA purified, and then sent to be sequenced. Direct cloning PCR products into the TOPO vector was possible because the vector DNA has been linearized and has topoisomerase attached, which allows the PCR product to be cloned more efficiently.

PCR reactions were prepared as follows:

	HCMV (+)	BaCMV (+)	HCMV (-)	BaCMV (-)
DNA	5 (0.25 µg)	5 (0.25 µg)	0	0
3.3x r*Tth* buffer	15	15	15	15
100mM SP216 primer	1	1	1	1
100mM SP217 primer	1	1	1	1
10mM dNTP	8	8	8	8
25mM MgOAc	5	5	5	5
r*Tth* polymerase (0.5U/µl)	2	2	2	2
dH_2O	13	13	13	13

Table 5. PCR to Amplify Sequence from the UL1 Region of the BaCMV Genome
Reactions included positive and negative controls. (All volumes are in µl).

The HCMV DNA reaction was included as a positive control. Two negative control reactions (labeled HCMV- and BaCMV-) had no input DNA and were included to test for DNA contamination. Each reaction was performed at the following temperatures: 99 °C to denature (30 sec), 55 °C to anneal primers (30 sec) and 72 °C for extension (1 min). The thermocycler was set at 30 cycles and the final extension step was increased to 20 min. Five µl of HCMV and BaCMV positive reactions were run on an agarose gel and both reactions showed a band at approximately 600 bp as expected (Figure 12). The remaining 45 µl of the BaCMV PCR reaction was electrophoresed, then the band of amplification product was cut out of the gel and phenol:chloroform purified. A TOPO cloning reaction was performed using 2 µl of DNA, 1 µl of salt solution, 1 µl of TOPO vector and 2 µl of distilled H2O. Two µl of the TOPO cloning reaction was transformed into TOP 10F' cells and then 100 and 200 µl was spread on two separate

plates of 2YT/AMP media and grown at 37 °C overnight. This yielded numerous colonies; 12 were selected for miniprep purification and designated clones numbers 3597 – 3608. Clone 3599 was sent for sequencing (see Results II-C-2).

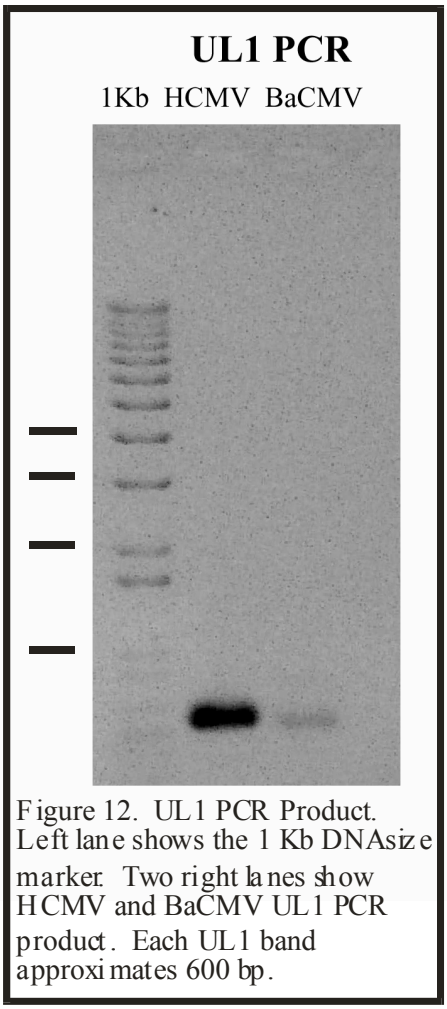


Figure 12. UL1 PCR Product. Left lane shows the 1 Kb DNA size marker. Two right lanes show HCMV and BaCMV UL1 PCR product. Each UL1 band approximates 600 bp.

C. Generating DNA for Sequencing Using PCR

Where gaps in sequence were believed to exist between sequenced clones, PCR was also used to generate DNA for sequencing. Using this method can be problematic since PCR is prone to errors. To minimize this problem, r*Tth* (recombinant *Thermus thermophilus*) DNA polymerase was used for DNA amplification since Perkin-Elmer states that it has higher fidelity than Taq polymerase. Primer pairs were designed using sequence data adjacent to the gap to generate a PCR product spanning the sequence gap. The PCR product was then cloned into the pCR2.1-TOPO vector or sent directly to the OSU sequencing facility along with the primers used to generate the sequence. (Sequencing of PCR product was necessary when TOPO cloning failed). Prior to TOPO cloning, the PCR product was purified using the Promega Wizard Prep kit. The manufacturer's protocol was followed and the DNA was eluted by adding 50 µl of TE and centrifuging the minicolumn for 20 sec at 10,000 x g.

Most PCR reactions were performed using the standard protocol (described elsewhere). However, some clones that did not yield good product using standard PCR showed better results using TaKaRa LA (long and accurate) PCR. This kit is designed to generate PCR products up to 40 kb and can also be used for DNA having a high GC content. It was suspected that the gap regions had been difficult to amplify due to possible high GC DNA content in these areas. G:C base-pair bonds are more stable than those of A:T bonds so may be difficult to denature during PCR. This resistance can decrease PCR and/or sequencing efficiency.

The TaKaRa kit was used to PCR into one gap site that existed between clones apBS *Bam*HI 97 and pTOPO5252 to create clone 5548 which was > 5kb. A forward primer (SP286) and reverse primer (SP367) were designed and used for this reaction. Two µl of BaCMV OCOM4-37 circular DNA (10ng/ml) was used as template. The PCR reaction also contained 5 µl of 10X LA-PCR buffer, 8 µl of 10 mM dNTPs, 1 µl of 100 mM SP286, 1 µl of 100 mM SP367, 0.5 µl of TaKaRa LA Taq polymerase (5U/µl) and 32.5 µl of distilled water. The reaction was run for 30 cycles at 98 °C (10 sec), 64 °C (60 sec) and 68 °C (15 min). All 50 µl of the product was run on an agarose gel, the DNA band was cut out of the gel, phenol:chloroform purified and cloned into the TOPO XL vector. This vector is designed to accept large inserts (up to 10 kb).

To generate DNA sequences for the region between clones B110 and H659, various primer pairs were designed and used for PCR (Figure 13). The following sets of primer pairs yielded results: SP469 and SP467; SP469 and SP266. SP469 was the

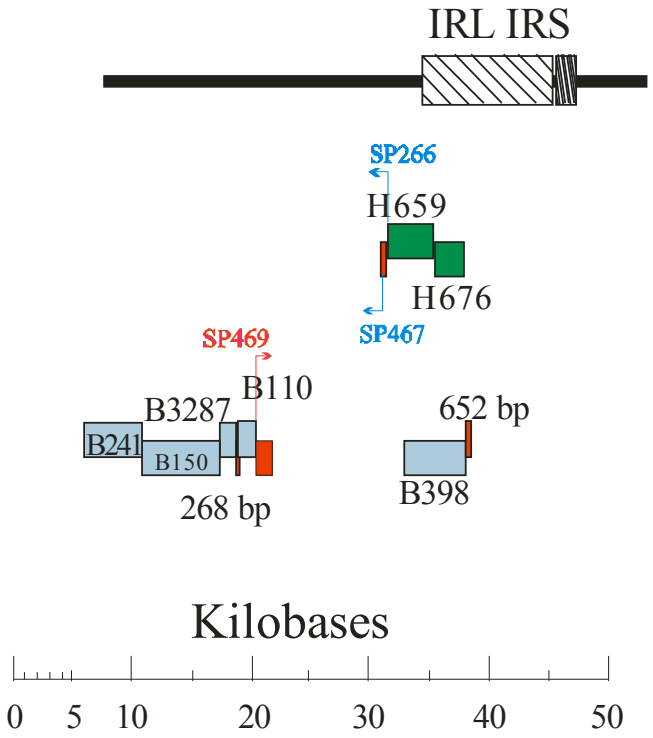


Figure 13. BaCMV (OCOM4-37) B110 - H659 Gap. The upper bar represents the sequenced HCMV (AD169) genome. The boxes represent cloned BaCMV DNA fragments and have been placed on the map based on their similarity with HCMV. Blue = *Bam*HI clone, Orange = *Eco*RI clone, Green = a *Hind* III clone, and Red is uncloned but sequenced PCR fragments. The red and blue arrows show the location of DNA primers.

forward primer in each reaction while SP467 and SP266 were the reverse primers (see Appendix B).

The PCR product from each reaction was cloned into the pCR2.1-TOPO vector, transformed into XL1-Blu cells, plated on 2YT/AMP, and grown overnight. Clone DNA was then purified using the standard miniprep protocol and digested with *Eco*RI. The digests were run on an agarose gel and those clones having appropriate DNA bands and the greatest DNA concentration were sent for sequencing (see Results section II-E).

PCR primers SP378 and SP471 were used to amplify DNA across the un-sequenced region that existed within the BAC clone N3851. These primers generated a product having three distinct bands. The largest was approximately 5.5 kbp, the middle band was 1.6 kbp and the smallest one was about 600 bp (Figure 38A). All three bands were excised from on agarose gel and purified. The large band was TOPO cloned and the two smaller ones were stored for future consideration. Only 8 colonies grew when plated on 2YT/AMP media. All were selected for minipreps but only 7 of the clones grew. These 7 were purified using the standard miniprep protocol and then digested and electrophoresed on an agarose gel (Figure 38B). Clones 7272, 7274 and 7275 were sent for sequencing.

D. Generating Sequence Data Through Primer Walking

Because the GPS1.1 transposons insert randomly into clone DNA, sequences do not always overlap and there are unsequenced regions where gaps exist. This leaves some areas with insufficient coverage. Additionally, there were some DNA inserts that were difficult to sequence, perhaps due to repetitive sequences, high GC content or secondary structure. For these unfinished clones, primer walking was the best strategy to

produce new data. This method had not been used previously because it is both time-consuming and more expensive. When sequencing DNA, it is ideal to sequence both strands and to have two or more sequence data files (traces) for each strand (Figure 14).

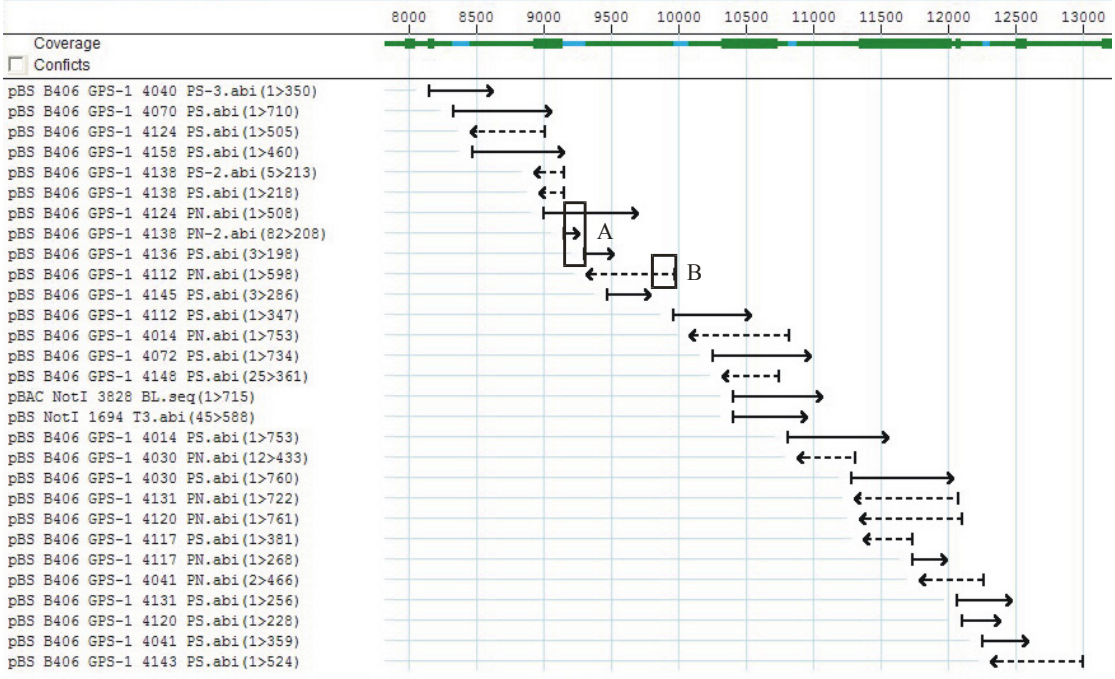


Figure 14. Strategy View of BaCMV OCOM4-37 Clone B406 Alignment. Shows alignment of several GPS-1 clones, orientation, origination point and areas of overlap. Solid arrows represent forward primer sequence, dashed arrows represent reverse primer sequence. The box at 9133 - 9304 bp (A) shows DNA only sequenced on one strand and the box at 9773 - 10055 bp (B) shows DNA only sequenced once.

Overlapping sequencing data is used to assemble a consensus sequence and sequencing both DNA strands twice helps verify the accuracy of the data. In primer walking, the DNA sequence of a newly-generated sequence strand is used to design a new primer. This primer is then used in sequencing to generate more DNA sequence data that extends the previous sequence data. New DNA sequence is used to design yet another primer which is then used to further extend the DNA sequence and so on until the entire region of interest is sequenced (Figure 15). For clones that had gaps or that were difficult to sequence, new primers were designed (Primer Select, DNAStar, Madison, WI) adjacent

to the region to be sequenced. Purified Qiagen miniprep DNA was used for each sequencing reaction and a typical sequencing reaction would provide from 500 – 800 bp of new data.

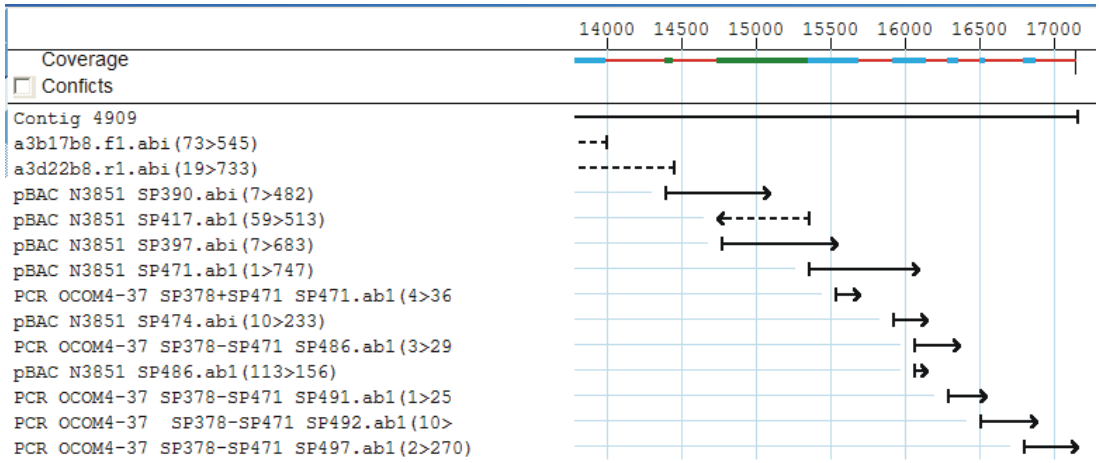


Figure 15. Primer Walking of the pBAC N3851 Clone. Each name under Contig 4909 corresponds with an individual sequencing reaction. The names designate the origin of the DNA being sequenced and the primer used for sequencing. Exceptions to this naming system represent DNA sequences produced by other facilities (OU-ACGT). The arrows on the right side of the figure show the direction of the primer and length of the resulting sequence. Primers used in primer walking are all directed toward the right end of the alignment.

CHAPTER FOUR

RESULTS

In this investigation, over 217,000 bp of DNA from the OCOM4-37 strain was sequenced. It is projected that there are about 8,000 - 12,000 bp are left to complete the BaCMV genome. The HCMV genome consists of 229, 354 bp and RhCMV contains 221,459 bp. Therefore, the projected size of the BaCMV genome is probably somewhere between HCMV and RhCMV. The entire US region of the genome has been cloned and sequenced. However, in the UL component of the genome there are two gaps which have proven difficult to sequence (Figure 16). A clone in the BaCMV library has been identified that spans one of these gaps but subcloning DNA and sequencing of this region has proven difficult. This difficulty is probably due to either secondary structure or repetitive sequences within this region. The terminal UL contiguous sequence (left contig) is contained within a plasmid clone and contains the terminal repetitive sequence (TRL) and ORFs BaUL01 through BaUL25 (partial).

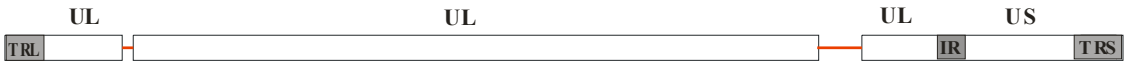


Figure 16. Gaps in Unique Long Sequence: The UL sequence is composed of BaUL28 through BaUL122 (partial). And the US contig contains the internal repetitive sequence (IR), the terminal repetitive short (TRS) sequence and BaUL123 through BaUS34 ORFs.

I. PRELIMINARY DATA

This preliminary data was generated by Dr. Blewett and R. Preston Rogers in 1997 – 1999. Initially, a large quantity of BaCMV OCOM4-37 was grown in cell culture, and DNA purified on NaI gradients. This DNA was digested with restriction enzymes, fragments separated by agarose gel electrophoresis, and the gel stained and photographed to determine the quality of DNA recovered and to choose REs for cloning.

The preparation contained high molecular weight DNA, as the uncut DNA lane showed little degradation (Figure 17). The REs that cut at 8 bp recognition sites, *Asc*I, *Not* I , *Pac*I and *Pme*I either did not cut or cut the DNA into a small number of pieces. The restriction enzyme *DraI* (with a 6 bp AT-rich recognition sequence) also cut the genome into a small number of pieces. Since the cloning project was to use plasmid/phagemid vectors of the pUC lineage (in which DNA fragments > 15 kbp are difficult to clone), REs that digest the genome into smaller pieces were desired. The enzymes *Bam* HI and *Hind* III digested the majority of the BaCMV genome into fragments < 15 kbp (Figure 17). The enzymes *Eco* RI (data not shown), *Ssp* I and *Xba* I cut the genomic DNA into less useable fragments. *Bam* HI was chosen for the first random library.

A. Viral DNA Restriction Enzyme Digest

Genomic BaCMV OCOM4-37 DNA was digested with *Bam* HI, ligated into *Bam* HI -cut plasmid vector (pBluescript II SK +), transformed into *E. coli* TOP10F' cells, and spread on agar plates containing 100 μg/ml ampicillin plus IPTG and X-GAL.

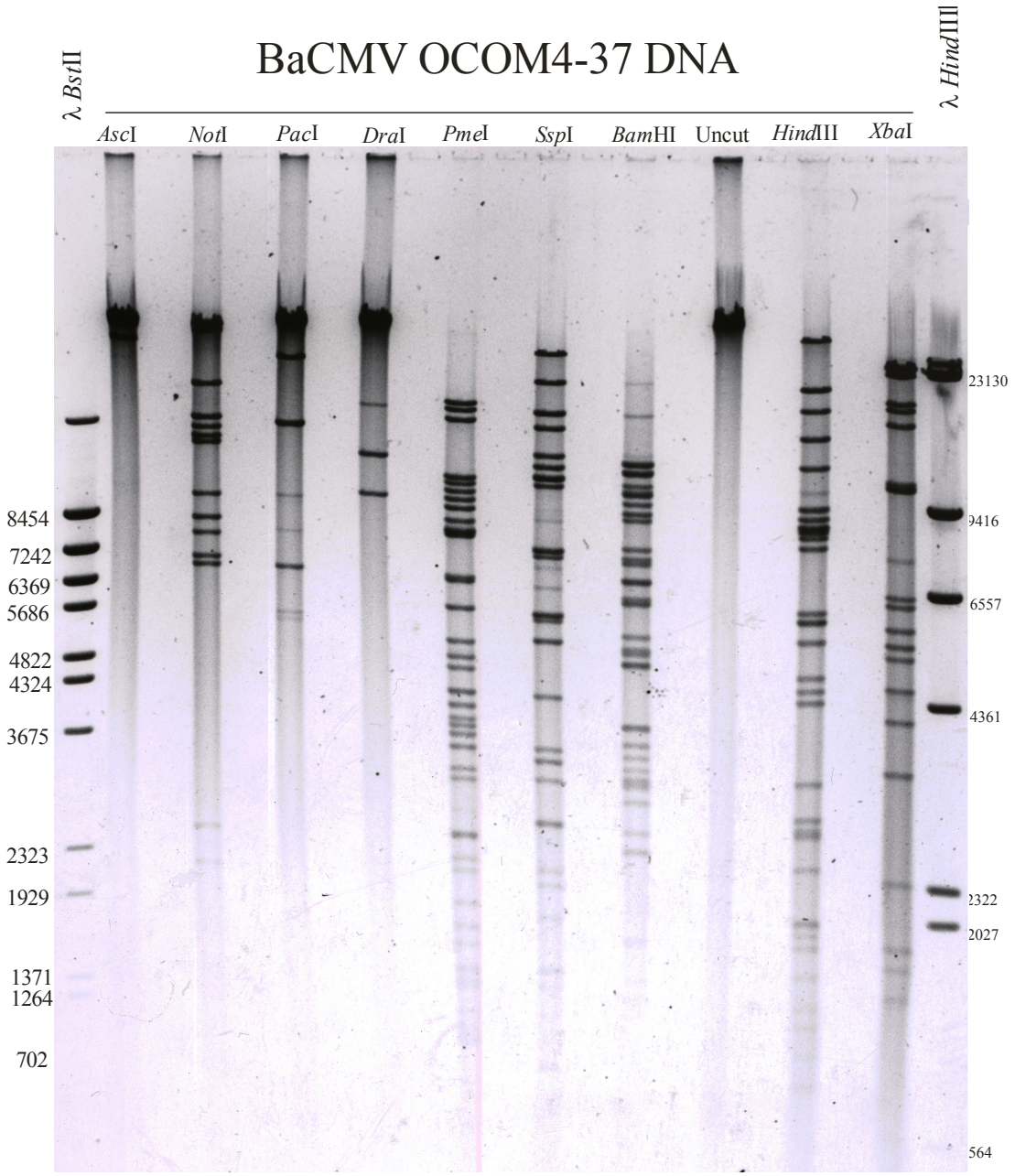


Figure 17. Restriction enzyme digestion of BaCMV OCOM4-37 DNA. Lambdavirus DNA was digested with *BstE*II or *Hind*III to provide size markers. The numbers on each side of the gel are the DNA size standards in base pairs. Equal amounts of genomic DNA were digested.

White colonies were used to inoculate 2 ml of 2YT plus Amp and incubated overnight.

B. *Bam*HI Miniprep Example

Plasmid DNA was isolated using the miniprep procedure, digested with *Bam*HI, electrophoresed and photographed (Figure 18). After digestion and electrophoresis, all clones should have a plasmid DNA band (approximately 2.9 kbp) and a unique viral DNA insert band. Numerous clones containing unique *Bam*HI fragments were identified. To differentiate unique clones with similar sized inserts, clones with similar-sized inserts were digested with *Xba*I plus *Hind*III, electrophoresed on an agarose gel, stained and photographed (Figure 19).

There is one *Xba*I and a *Hind*III in the MCS of the pBluescript II SK (+) vector (one on each side of the *Bam*HI site). Therefore, digestion with *Xba*I plus *Hind*III will cut the insert out of the vector. If the insert contains sites *Xba* I and/or *Hind*III, multiple bands will result on the gel thereby allowing duplicate clones to be identified. Lanes 2-6 and 9 in figure 18 show examples of multiple clones containing the same *Bam*HI fragment. Clones containing unique DNA inserts were sent to the Core Facility at OSU in Stillwater for end sequencing.

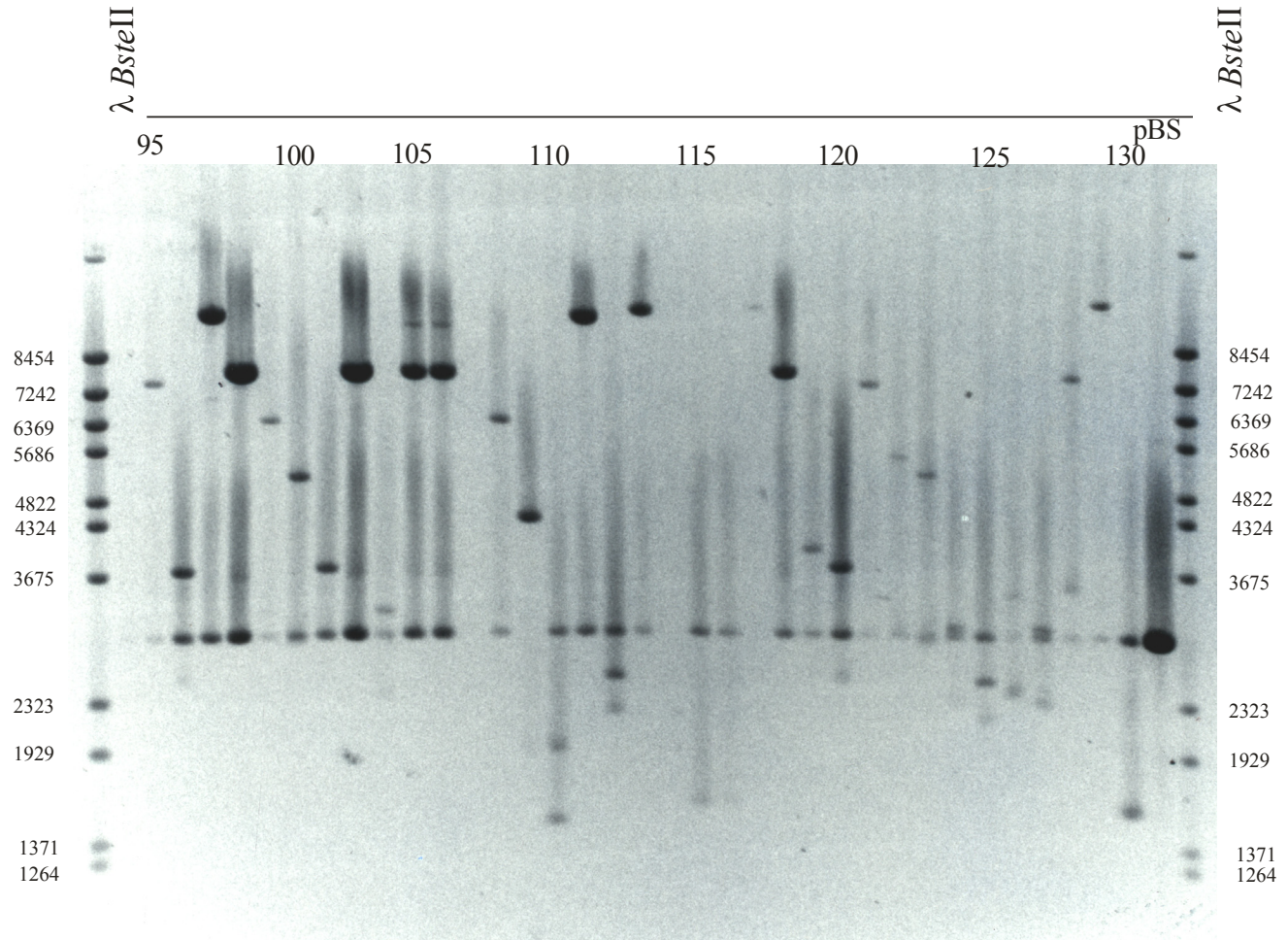


Figure 18. Potential BaCMV OCOM4-37 *Bam*HI DNA Clones. DNA fragments cloned into pBluescript II SK+ digested with *Bam*HI. Minipreps 95-130 are indicated by number. The pBS lane is vector DNA. Lambdavirus DNA was digested with *Bste*II to provide size markers. The numbers on each side of the gel are the DNA size standards in base pairs.

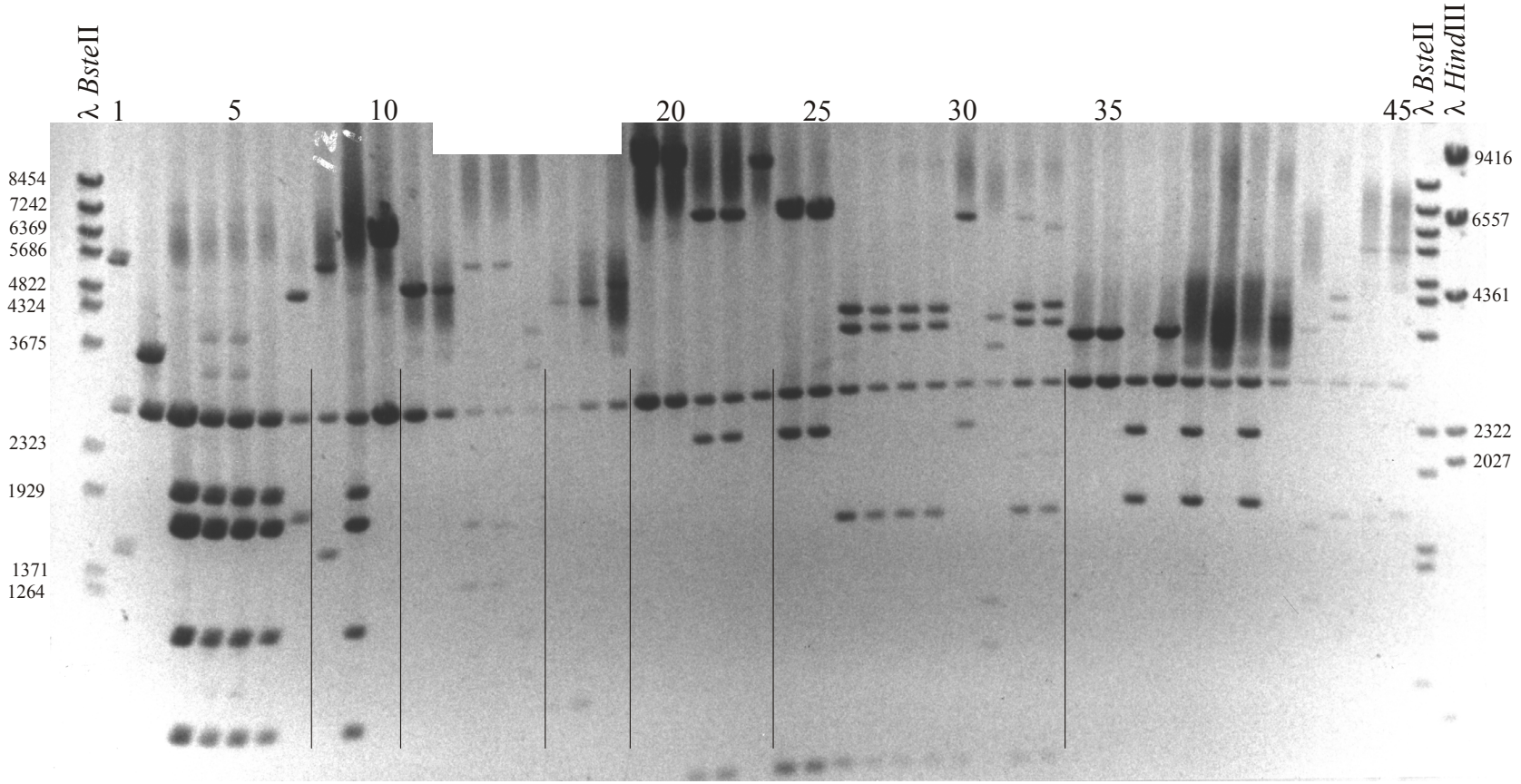


Figure 19. Identifying Unique BaCMV OCOM4-37 *Bam*HI Clones. *Bam*HI DNA clones were digested with *Xba*I and *Hind*III. Lanes 1 - 45 are Minipreps (96, 97, 103, 105, 106, 118, 121) (128, 131, 150_ (101, 123, 179, 194, 197) (95, 122,172) (132, 140, 143, 144, 153) (98, 111, 113, 129, 149, 154, 161, 171, 181, 199) (97, 102, 119, 120, 136, 139, 151, 159, 189, 193, 195, 201). Lambdavirus DNA was digested with *Bst*EII or *Hind*III to provide size markers. The numbers on each side of the gel are the DNA size standards in base pairs.

C. BLAST SEARCH

Digital DNA sequencing data in the form of Abi trace files were received from the sequencing facility. Sequence data was edited by the removal of vector MCS sequences at the 5' end and poor quality chromatogram peaks at the 3' end. The resulting DNA sequence data was used to search GenBank with BLAST. A typical BLAST result is shown in Figure 20.

```
Query=  97 RP
        (899 letters)

Database:  Non redundant GenBank+EMBL+DDBJ+PDB sequences
           312,067 sequences; 581,488,900 total letters.
Searching.................................................done

                                                                    Smallest
                                                                      Sum
                                                              High  Probability
Sequences producing High scoring Segment Pairs:                Score  P(N)       N

emb|X04650|HEHCMVU  Human cytomegalovirus (HCMV) short un...   617  9.8e 41    1
emb|X17403|HEHCMVCG Human cytomegalovirus strain AD169 co...   617  9.9e 41    1


>emb|X04650|HEHCMVU Human cytomegalovirus (HCMV) short unique region, short
            repeats, and part of long repeat.
            Length = 43,275

  Minus Strand HSPs:

 Score = 617 (170.5 bits), Expect = 9.8e 41, P = 9.8e 41
 Identities = 213/327 (65%), Positives = 213/327 (65%), Strand = Minus / Plus

Query:    361 CCCAGCAGCA GCAAGGGATCGGCTTCCGCCCGCCCTGGAGTTTTGATTACCAAGTTGATC 302
              |||||||  ||||| || ||  | || |  || || || ||   ||| |||| |||||
Sbjct:  25395 CCCAGCATGAGCAACGGCTCCACCTCGGTGCGACCCGGCGTGCGGATGACCAGGTTGAGA 25454

Query:    301 CCGTGAAATTTTTCCGCAACGTTNGCCACGGCCCGTGCGTCCGTGGGATCGAGGCAGAGG 242
              |||   | ||     ||    || |||||||||   ||||| ||||| |||  |||||||
Sbjct:  25455 CCGCTCATTTCGCGGGCCGTCTTGGCCACGGCCGCAGCGTCAGTGGGGTCGGTGCAGAGG 25514
```

Figure 20. Typical BLAST Search. GenBank was searched using the BLASTN program and the unique DNA sequence generated when the *Bam*HI Clone 97 was sequenced using reverse primer. The searches results are truncated for display purposes. This BaCMV clone is homologous to the unique short region of the HCMV genome.

Homology of the BaCMV clones with HCMV AD169 was used to position BaCMV inserts on a map of the BaCMV genome (Figure 21).

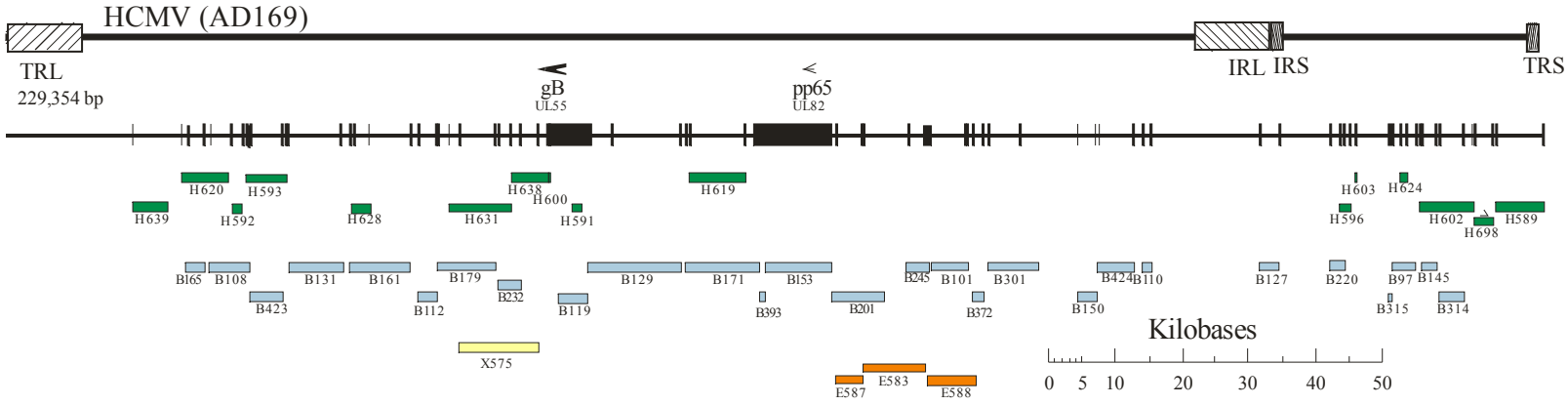


Figure 21. BaCMV (OCOM4-37) Clone Homology Map. The upper bar represents the sequenced HCMV (AD169) genome. The gB and pp65 genes are shown as arrowheads. The thin black bars represent the BaCMV (OCOM4-37) genome. The thick black boxes represent the previously sequenced portions of the BaCMV genome. The colored boxes represent cloned BaCMV DNA fragments and have been placed on the map based on their sequence similarity with HCMV. Blue = *Bam*HI clone, Green = *Hind*III clone, Orange = *Eco*RI clone and Yellow = *Xba* I clone.

D. PCR Amplification and Sequencing Across Clone Junctions

Once unique *Bam*HI clones that had homology to HCMV or RhCMV had been placed on a map, primers were designed to sequence or amplify across *Bam*HI site junctions or across gaps between nearest *Bam*HI fragments. If an overlapping DNA fragment from another RE digest was cloned and mapped to that area, the primer was used to sequence that template. In most cases, PCR amplification of viral DNA was necessary. The end sequences from *Bam*HI clones and PrimerSelect software (DNAStar, Madison, WI) were used to select appropriate primers sequences. Purified BaCMV DNA was used as the template for PCR. Amplification products were sent for sequencing or were cloned into pCR2.1-TOPO. Miniprep DNA was prepared from the TOPO clones and digested with *Eco*RI (which cuts on each side of the PCR product insertion site). The resulting DNA fragments were separated by gel electrophoresis, stained and photographed. Those clones that had good quality inserts of the expected size were sent for sequencing. The resulting DNA traces were placed in the master assembly after the MCS and primer sequence were removed from the 5' end and poor quality data removed from the 3'end.

Sequencing of non- *Bam*HI clones (*Hind*III, *Eco*RI or *Xba*I clones) or PCR amplification products across the ends of *Bam*HI clones confirmed that many clones were indeed adjacent as mapped (Figure 22). In some cases, small fragments, some as small as 16 bp, were found between *Bam*HI clones. As this project continued, all *Bam*HI junctions were confirmed.

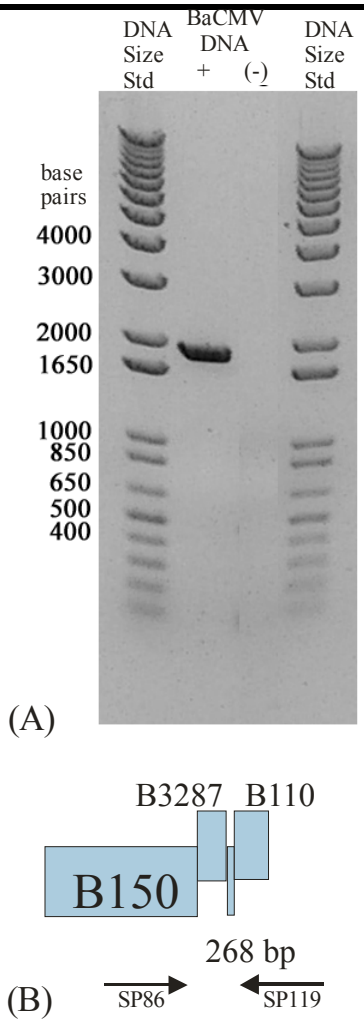


Figure 22 PCR Amplfication of BaCMV OCOM4-37 DNA between clones B110 & B150. Part A Primers SP86 & SP119 were used to amplify viral DNA or a negative control reaction with no DNA. The amplification products were separated by gel electrophoresis, stained and photographed. Part B. A schematic diagram showing the *Bam* HI clones, PCR primers and the B110 - B150 gap. *Bam*HI clones are shown as blue boxes in proportional sizes.

II. DISSERTATION FINDINGS

At the time that this project began, approximately 70% of the BaCMV genome had been cloned but little sequence data had been produced. Therefore, the aims of this research were 1) to clone the remaining 25 – 35% of the genome, 2) to sequence internal regions of clones for which only end sequence data was available, 3) to identify clones containing the genome termini by screening existing clones, 4) to PCR across gaps

between clones, and 5) to primer walk through clones that were difficult to sequence. Once the genome was cloned, sequenced and assembled, the final consensus sequence would be analyzed and compared with other herpesvirus genomes. This involved performing BLAST searches on sequences to identify homologues from other herpesviruses and then creating independent alignments with other primate virus homologues for which published genomes were available. Several BaCMV homologues were grouped into herpesvirus gene families and further analysis revealed distinguishing features seen in other herpesviruses.

The results described below cover several categories: 1) genomic cloning 2) screening existing clones 3) generating PCR product for sequencing 4) trace file assembly, and 5) genome analysis.

A. Directional Cloning to Capture Terminal Fragments

1. Evaluating Vector and Insert DNA by Gel Electrophoresis

After digesting both the pSL301 vector DNA and mammalian DNA with *Bam*HI, *Hind*III or *Not*I (see Methodology VII-A-2), the purified vectors and test inserts were run on a gel and appeared as seen in Figure 23.

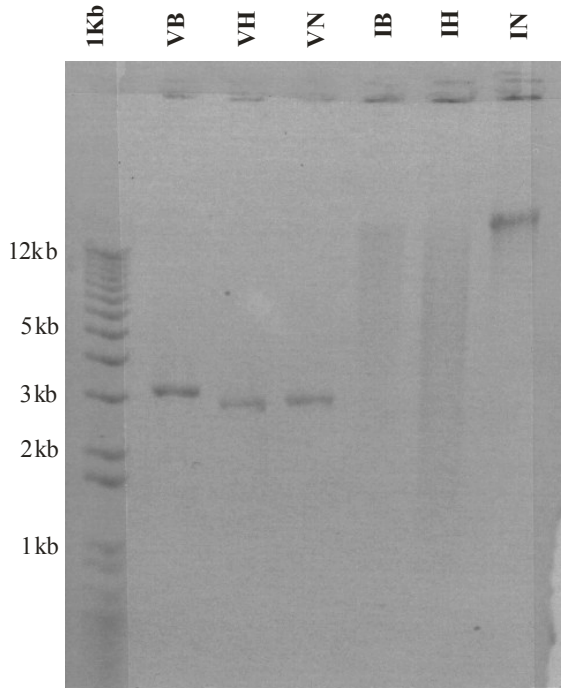


Figure 23. Digests of pSL301 Vector and Mammalian DNA. First lane shows the 1 Kb DNA size marker. VB, VH and VN lanes are pSL301 (vector) cut with either *Bam*HI, *Hind*III or *Not* I. IB, IH and IN lanes are mammalian (insert) DNA cut with the same restriction enzymes.

Vector bands were visible for each restriction enzyme digest and the mammalian insert DNA was smeared due to numerous cuts (as expected). The vector DNA bands should be similar sizes but migrate

at slightly different rates. This is likely due to uneven wells, not a difference in vector size. The vectors were ligated with appropriate insert DNA and transformed into TOP10F' cells and the resulting colony numbers are shown in Table 6. Each vector had been dephosphorylated to reduce vector self ligation. The *Hind* III plates had too many vector only colonies (lacking an insert) indicating the vector was self ligating. Therefore, only the *Bam* HI and *Not* I vectors were used for ligation with BaCMV DNA.

Restriction Enzymes	*Sma* I/*Bam* HI		*Sma* I/*Not* I		*Sma* I/*Hind* III	
Colonies per plate	4	15	89	129	175	115

Table 6. Colonies Resulting from pSL301 Vector Cloning and Transformation

Most of the potential blunt-*Not* I clones had both vector and insert bands (Figure 24) and are possible termini. Two that seemed to have the most concentrated DNA (QN3058 and QN3063) were sequenced. Both clones yielded good sequence data and mapped to the US terminus of the genome. Another 24 minipreps were prepared from additional colonies but, since they appeared to be duplicates of previous clones, none were sequenced. Clones with small inserts were not sequenced since we were expecting end fragments to be large. All clones made in this experiment, whether apparent duplicates or containing a small insert, were later screened by DNA hybridization. A clone made in this experiment (pSL301 QN3094) turned out to be the 3' terminal fragment in later screening experiments.

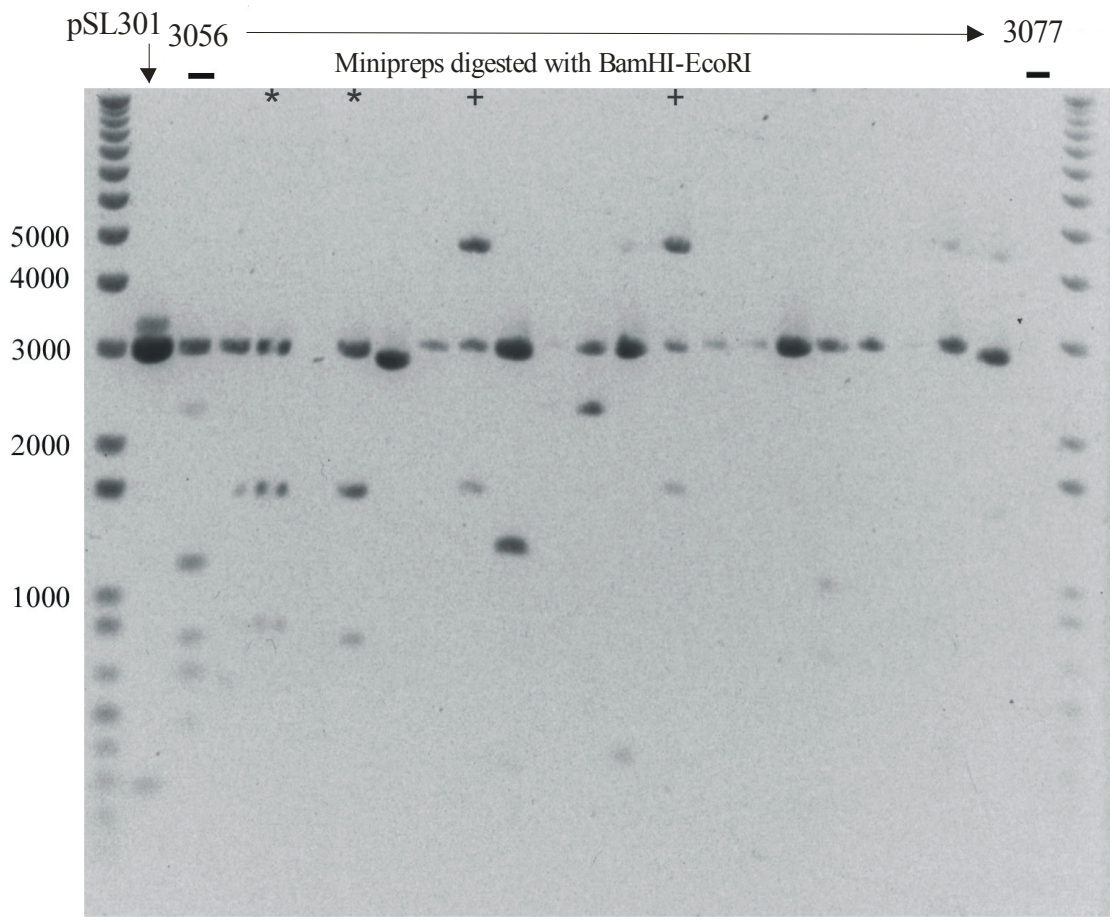


Figure 24. Miniprep Digests of Clones QN3056 - QN3077. Clones QN3058 and QN 3060 (identified by an *) appear to be similar and QN3063 and QN3068 (identified by a +) appear to be duplicates. QN3058 and QN3063 were selected for sequencing because they had more concentrated DNA.

B. Using Circular BaCMV DNA to Generate Clones Containing the Termini

Previous screening methods had not identified sequence data from the UL terminus so cloning fragments from circularized genome DNA was attempted. This strategy was to circularize the normally linear viral genome and then digest the DNA with *Not* I to create fragments that could be cloned into plasmid or BAC vectors. One of the *Not* I fragments should contain the joined US and UL termini. The other genomic

fragments may also represent uncloned sections of the genome. Two strategies were used to circularize the genome (as described in Methodology).

Linear, NaI-purified BaCMV DNA was self-ligated to produce circular molecules. This DNA was digested with *Not* I , purified and ligated into a plasmid vector also digested with *Not* I . After growing bacteria transformed with plasmid clones overnight, the plates looked promising as each plate had more white colonies than blue ones. Unfortunately, this method did not yield data from the UL terminus as hoped. None of the clones that were screened by Southern Blot using a *Not* I end-labeled probe mapped to the termini. Results for the clones that were sequenced are as follows: 3208 aligned near B153 at the center of the UL region, 3211 was a duplicate of N1655 and aligned in the US region, 3217 and 3232 were duplicates and aligned at the 3' end of the BAC 3851 clone and 3231 aligned to the 3' end of 3208 in the UL region. All of the clones generated with this method appear to be fragments or concatamers of genomic pieces but none were end fragments nor were any clones recovered that contained large genomic fragments.

The second strategy was to purify circular BaCMV DNA. Herpesvirus genomes are in a circular state during viral replication and these molecules were preferentially obtained by the Hirt DNA purification method. Figure 25 shows circular BaCMV DNA prepared by this method, digested with *Bam* HI, and then electrophoresed on an agarose gel. These results show the DNA can be digested with *Bam* HI and there is not a lot of cellular DNA contamination (as shown by only slight smearing).

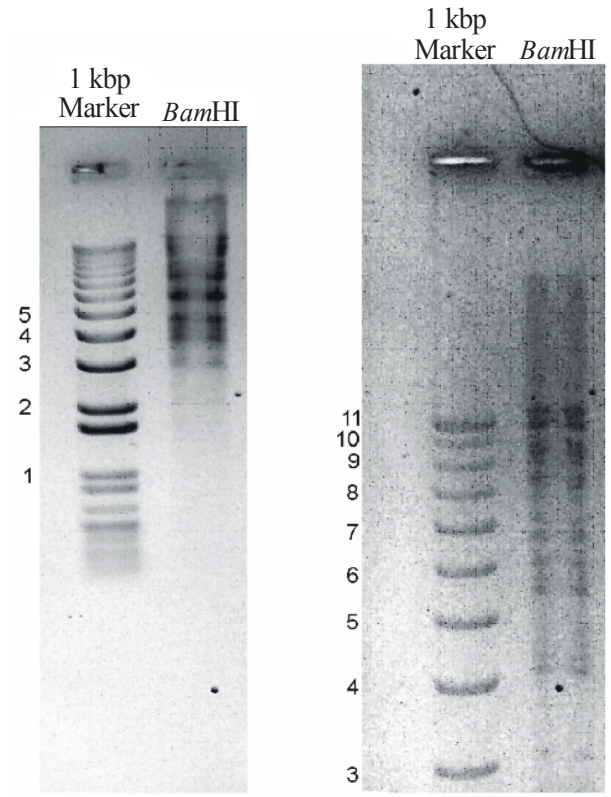


Figure 25. Circular BaCMV DNA Preparation. A *Bam*HI digested DNA preparation was separated by electrophoresis for 2 hrs (left) or overnight (right). Numbers to the left indicate the DNA standard marker in kb.

At this time, the two *Not* I terminal fragments were assumed to be large, (>10 kbp each). Therefore, after the BaCMV DNA circular DNA was digested with *Not* I and purified it was cloned into a BAC vector. This vector, pBACe3.6, had been digested with *Not* I and purified. The digested DNAs were then introduced into *E. coli* DH10B by electroporation.

BAC clone DNA was purified from isolated colonies, resulting in 58 clones. Figure 26 shows a typical gel electrophoresis of a *Not* I digestion for some of these clones. The large, ~9 kbp vector fragment can be seen in each lane. Insert DNA sizes

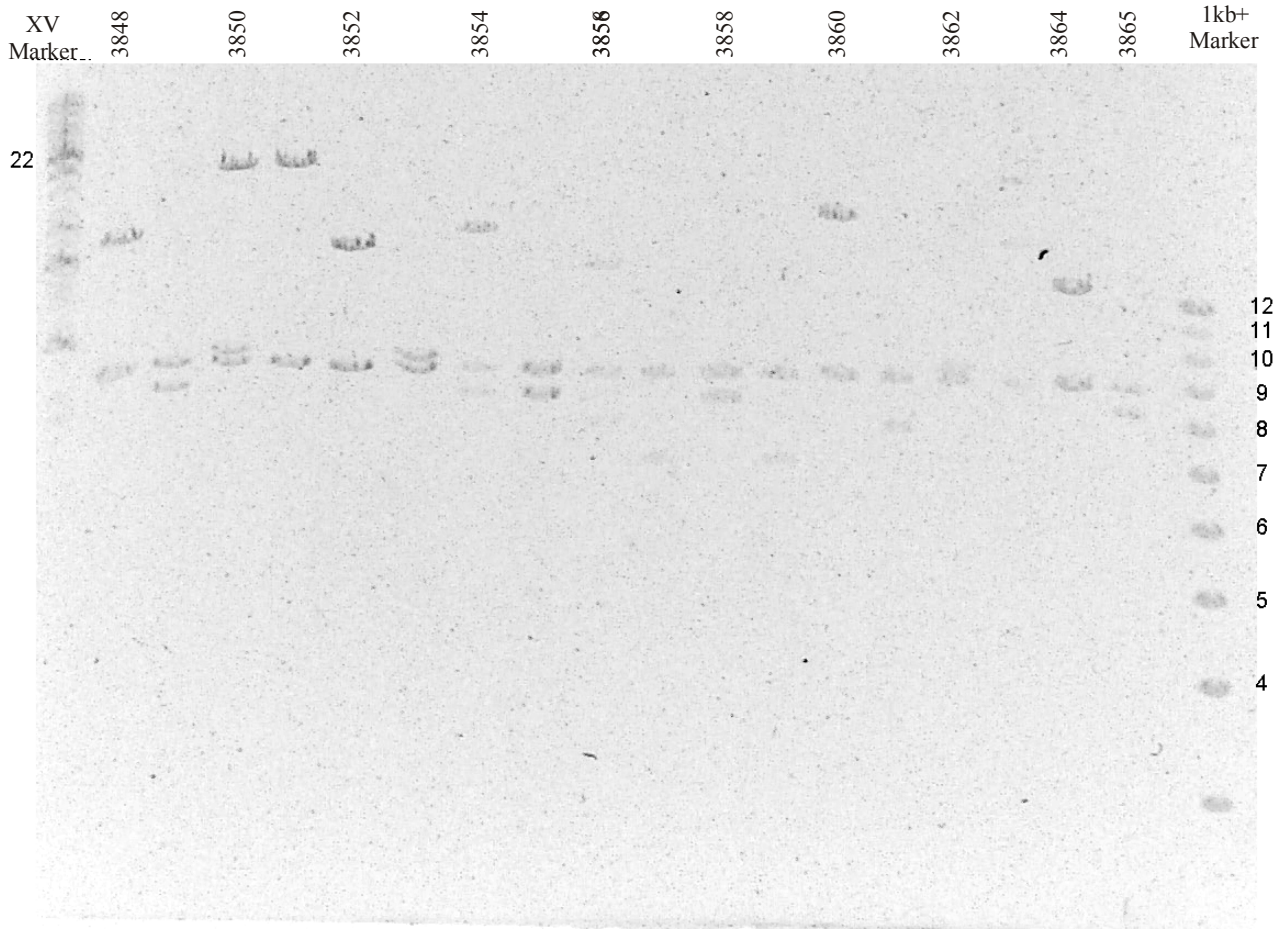


Figure 26 pBACe3.6 BaCMV OCOM4-37 CircDNA Minipreps 3848 - 3865 digested with *Not*I. A large size DNA marker is on the left, a 1kb+ marker is on the right and the numbers are fragment sizes in kbp. The pBACe3.6 vector band is expected to be ~9 kbp.

range from > 20 kbp to < 7 kbp. Some clones, i.e. 3850, were chimeras containing multiple *Not* I fragments. However, most of the clones appeared to contain one large *Not* I fragment and were considered useful in genomic sequencing. At this time, it was believed the terminal clone would be large, so the only clones that were characterized were those that contained a > 9 kbp insert. Clones 3851 and 3852 were end sequenced, and mapped to regions near the 3' terminus and near the 5' terminus respectively.

C. Screening Existing and New BaCMV Clone Libraries by DNA Hybridization to Identify US or UL Terminal Fragments

1. Dot Blot Results

The purpose of dot blot screening was to find clones in the BaCMV libraries that contain DNA located near either terminus of the genome. Both *Not* I and *Pme*I are rare cutters of BaCMV DNA. These REs were used to create probes by digesting end-tailed genomic BaCMV DNA to generate large, labeled end fragments. The longer probes are more likely to overlap terminal fragments and should hybridize to them when screening the clone library. After 8 – 10 min of development, a number of positives were visible on both the *Not* I and *Pme*I-probed dot blots and those clone pools with greatest intensity of color were selected for individual clone screening. Both positive and negative control results were as expected (Figure 27). Pooled clones shown in Table 7 were selected for individual screening. Those clones which had already been evaluated as terminal clones by other methods were omitted from individual screening as were clones that had previously been sequenced and mapped. After two dot blots of individual clones were performed using either the *Pme*I or *Not* I probe, the following clones were selected for restriction digestion and run on an agarose gel: *Bam* HI clones 297, 369, 370, 374, 378, 379, 381, 383, 384, 387, 388, 390 and 394, *Xba* I clones 2320, 2322, 2323, 2325, 2328,

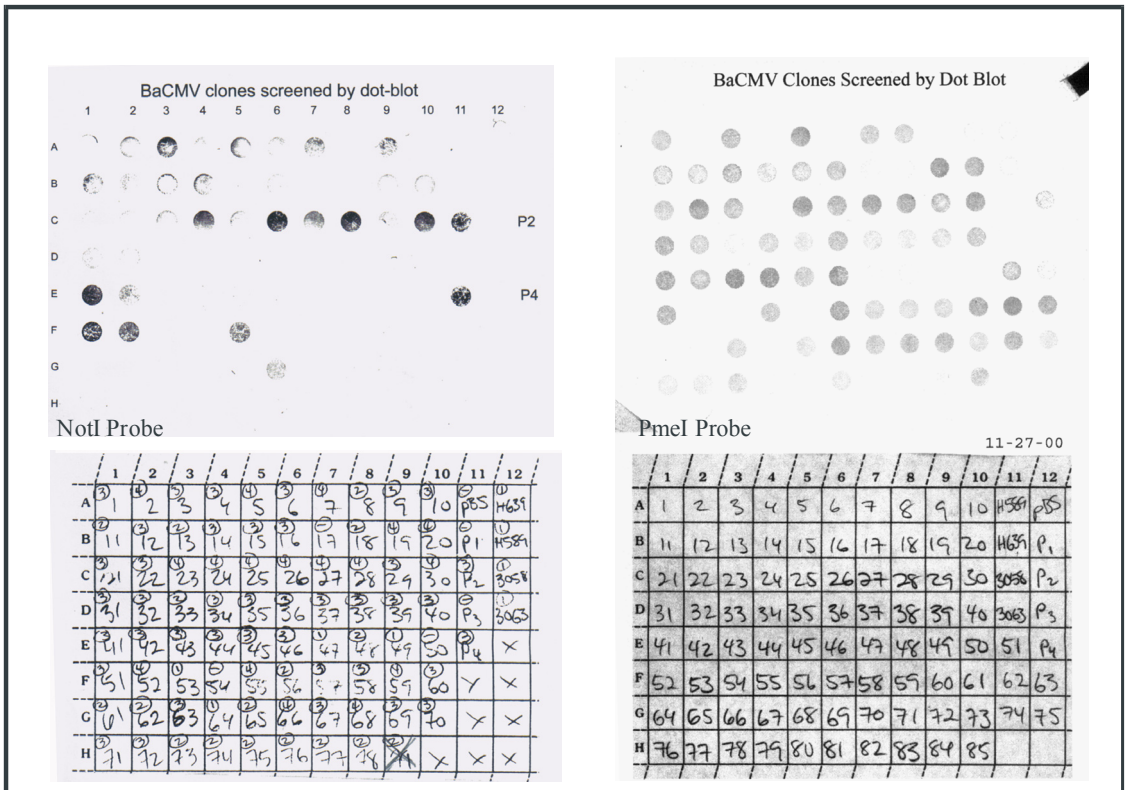


Figure 27. Dot Blot Screening of Pooled BaCMV Clones. Blot on the left was screened with *Not*I probe and blot on the right with *Pme*I probe. Only the strongest positives are visible here. Clones 3058, 3063, 589, and 639 were weakly positive so are obscure in this figure. The pBlueScript negative control does not show positive while the P2 and P4 positive controls do.

Pool #	RE used	Clones in pool	Total number	Omitted
S5	*Bam* HI	130 – 139	9	131
S19	*Bam* HI	296 – 305	7	296, 301, 305
S26	*Bam* HI	366 – 375	8	372, 368
S27	*Bam* HI	376 – 385	10	None
S28	*Bam* HI	386 – 395	3	389 – 395
S71	*Xho* I	2320 – 2329	10	None
S72	*Xho* I	2330 – 2338	9	None
S81	*Not* I /*Sma* I	3051 – 3060	10	3058
S84	*Not* I /*Sma* I	3081 – 3090	10	None
S85	*Not* I /*Sma* I	3091 – 3101	10	None

Table 7. Individual Clones Selected from Pooled Clone Groups
Restriction enzymes (RE) used to generate the clones are listed, the total number included in the dot blot pools, and clones that were omitted are listed.

2329, 2330, 2331, 2333, 2334 and 2335; *Sma*I/*Not*I clone 3052. Based on these gel analyses, some clones were eliminated as duplicates and others were culled due to poor DNA quality. The following clones were selected for sequencing: 297, 370, 378, 381, 384, 387, 388, 2322, 2330, 3052, 3094 and 3095. Clones 297, 370, 381, 384, 387, 388, 2322, 3052 and 3095 were all duplicates of previously-sequenced clones. Clone 378 gave poor sequence and 2330 was all vector DNA. However, QN3094 was successful at adding approximately 4.3 kbp of new sequence to the 3' terminus. When this clone was eventually sequenced, we found it included the BaCMV gene homologous to the terminal RhCMV gene. Due to the repetitive nature of the remaining QN3094 DNA sequence, this clone is believed to represent the most terminal region of the genome at the 3' end and to include some of the terminal repeats.

The H639 and B165 probes were generated from the endmost clones that had been identified at the time. These clones were chosen since H639 and B165 both appeared to be closest to the UL terminus. Dot blots using the H639 probe did not identify any clones near the UL terminus. The pooled clones that were screened using this probe showed a number of positives, but these later mapped to internal regions. While positive and negative controls yielded the expected results, the color was faint for the positive control. Ninety clones had been selected for individual screening from pooled clones. Some of the individual clones that showed a strong positive reaction, especially B424, were actually from the center of the genome map. Therefore, another probe, B165, was produced since this clone is more clearly associated with the UL terminus. One clone, E2290 showed positive results and it was reserved for further analysis.

Dot blot results for pooled clones probed with B165 showed 4 positives, and 30 clones were selected for individual screening from these 4 pools (Figure 28). Using the B165 probe, E2290 was also seen as a strong positive along with another clone, B359.

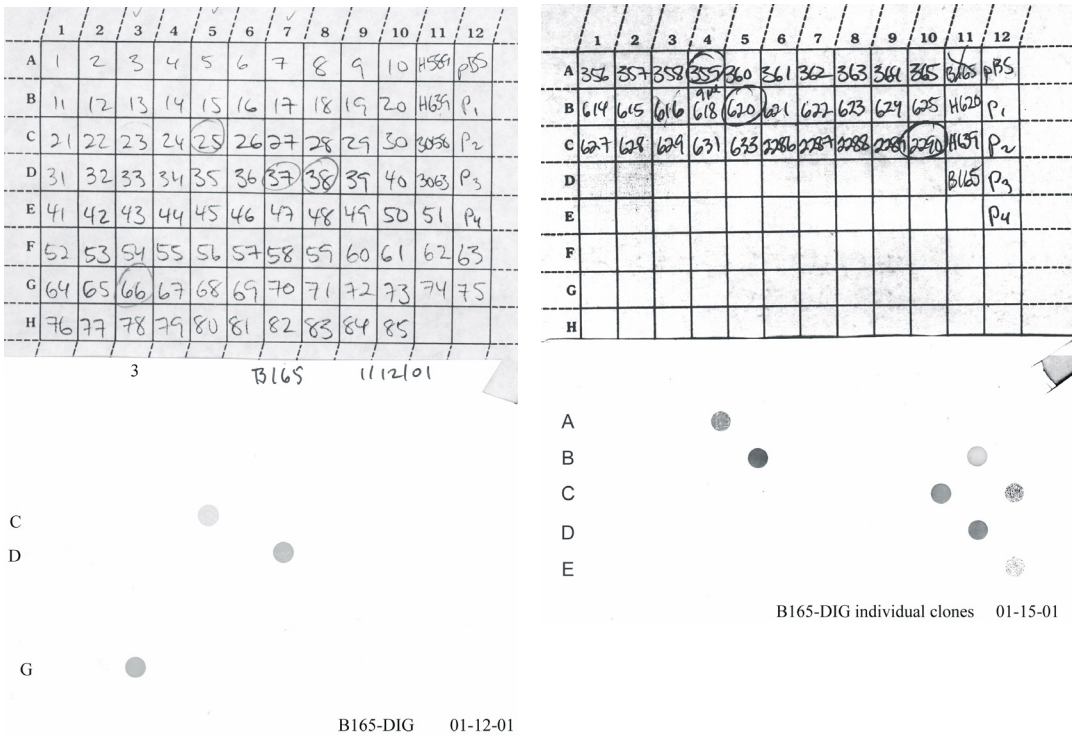


Figure 28. Dot Blots Screened with B165 Probe. Blot on the left was performed using pooled clones while blot on the right was performed on individual clones. Numbers in grid represent clone pool number while image below is the actual blot. Positive and negative controls are located in the two right columns. Positive controls were H639 (right terminus) and QN3063 (left terminus) while negative controls included H589 and QN3058 (both in US region) and pBlueScript.

Both clones were analyzed further and showed vector and insert bands after restriction. However, little DNA was present in the B359 digest so new DNA preps were made using the Qiagen method. The new plasmid DNA did not resemble the original clone for either B359 or E2290. The original B359 clone DNA prep had a single dark insert band on an agarose gel. Therefore the remaining original B359 DNA was sent for sequencing. Good sequence data was obtained and aligned with B165 near the UL terminus but only the KS

(reverse) primer seemed to work; the forward primer yielded no useful sequence. This suggests that the region in the vector MCS where the forward primer binds was probably damaged, missing, or has features that make it difficult to sequence. While these dot blots showed some true positives that yielded good sequence data, they did not provide sequence data from the 5' terminus. Clones believed to be near the termini of the genome were used to produce probes. It is now known that these were not terminal clones. Potential terminal clones described in sections A and B above were also screened. Clones representing the UL and US termini were in the libraries being screened. The US terminus was eventually identified in clone QN3094 while N3851 was shown to be located near the UL terminus.

2. Evaluating Clone pBS E2290

Clone E2290 was positive when screened with both H639 and B165 probes, and therefore was expected to be located near the left terminus of the BaCMV genome. Also, after digestion with *Eco*RI and electrophoresis, clone E2290 appeared to have 3 bands. These band sizes were approximately 14 kb, 3 kb and 1.3 kb, and were likely the result of two unique *Eco*RI fragments ligating together to form a chimeric clone (the 3kb band being vector DNA). To subclone each *Eco*RI fragment, the E2290 plasmid clone DNA was digested with *Eco*RI and religated to see if clones with the individual *Eco*RI fragments could be obtained and used for sequencing. It was expected that clones containing the vector band and either the 1.3 kb or 14 kb insert bands would be recovered. The original clone with the two insert bands could also be recovered. Clones containing some combination of the desired sizes were obtained (Figure 29). Re-cloning

the large insert is a low efficiency event and did not occur. Clone 3185, which appeared to contain the small *Eco*RI insert was sent for sequencing.

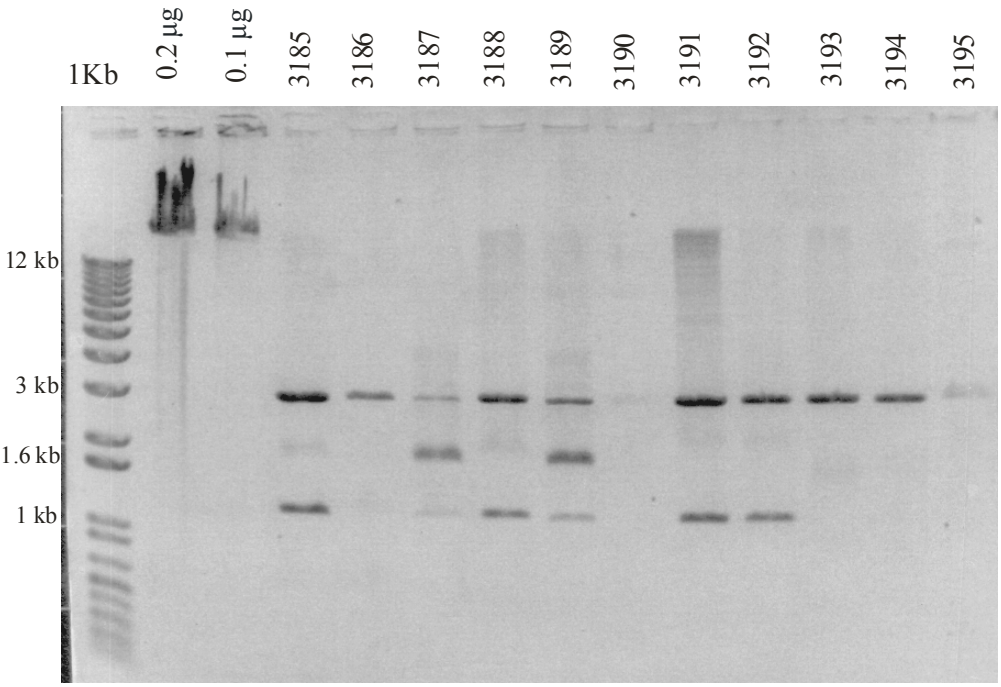


Figure 29. Restriction Enzyme Digests of E2290 Subclones 3185 - 3195. The first lane is the 1 Kb size marker. Lanes 2 and 3 are Lambda DNA used to determine concentration. All other lanes contain E2290 subclones.

When the E3185 subclone data was analyzed, the DNA sequence aligned at positions near the UL terminus where the probes were expected to bind. BLAST searches of this DNA sequence showed significant homology with the HCMV UL19 gene. Therefore data that was generated from this screening was useful, but was located about 25 kbp into the genome. DNA sequence from the other end of the clone DNA (representing the 14 kbp *Eco*RI fragment) did not align with any BaCMV sequences and a BLAST search for this sequence showed significant homology with human DNA. These results show that E2290 is a chimeric clone made up of BaCMV and human DNA *Eco*RI fragments ligated together. The human DNA is likely from the MRC-5 cells used

to grow the virus. This indicates that the viral DNA preparation is not completely pure and contains some cellular DNA. Eventually, BaCMV BAC clones were produced and sequence from the UL terminus was generated.

3. Screening Pooled Clones for Sequence in Gap Regions Using Southern Blot

All gels containing clones used for Southern blots were loaded in the same sequence as shown in Figure 30 so only the gel probed with the H596 probe is shown below. Some Southern blots did not show many positives or showed only weak

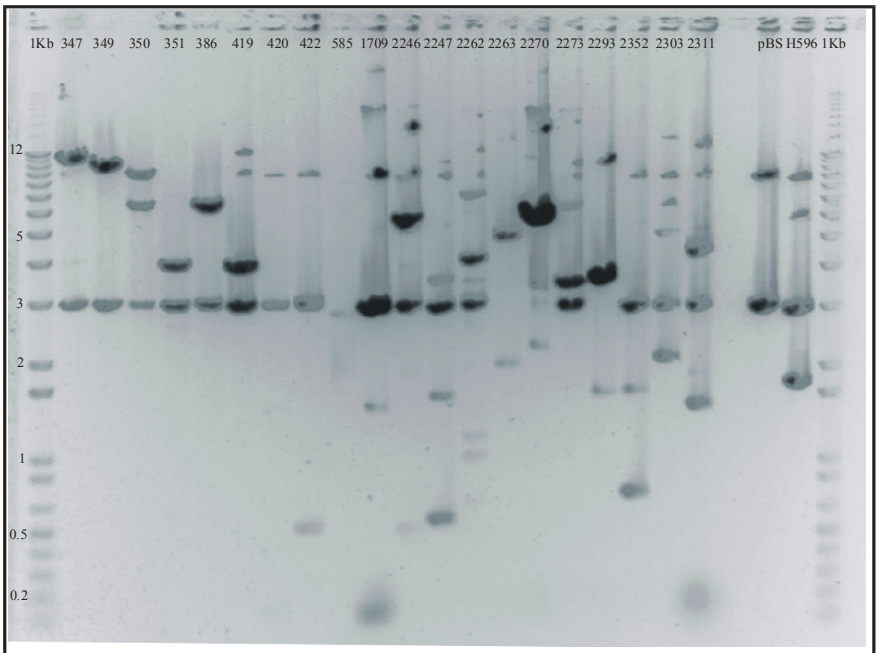
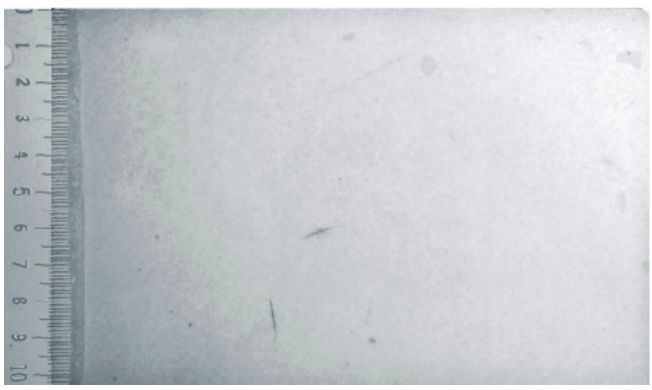


Figure 30. Gel Electrophoresis of Clones Used for Southern Blotting. This blot was screened with the H596 probe. The negative control is pBlueScript DNA, the positive control is H596 and a size marker (numbers are in kilobases) is loaded in the first and last lanes.

positives as seen in Figure 31. Weak positives probably result from spurious binding of probes to partially-matched sequence. Positive controls for the H596 and B315 probes did not develop on the respective membranes even though the probes had been tested prior to performing the Southern blots. The B110 positive control was positive and a few clones were positive as well. One of these, clone B422, is seen in the lower left-hand

quadrant of Figure 31. This clone was sent for sequencing and was found to be a duplicate of clone B127. Clone B127 had already been mapped to the beginning of the US region which is far from the B110 clone location.



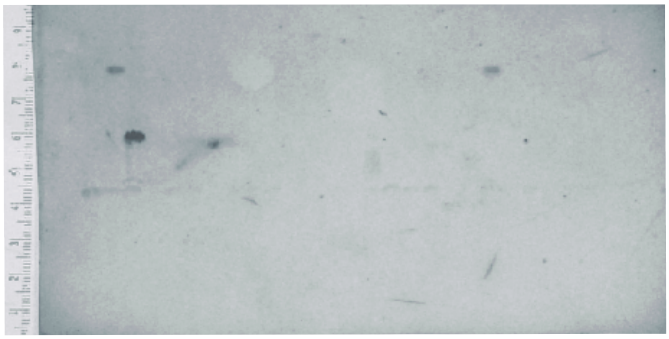


Figure 31. Southern Blots Using H659, B315 and B110 Probes. Rulers on either side of each gel show orientation of digests. Positives are only visible on the B110 gel.

4. Screening BAC Clones for Sequence in Gap Regions Using Southern Blot

The EtBr-stained gel prepared for screening BAC clones is shown in Figure 32.

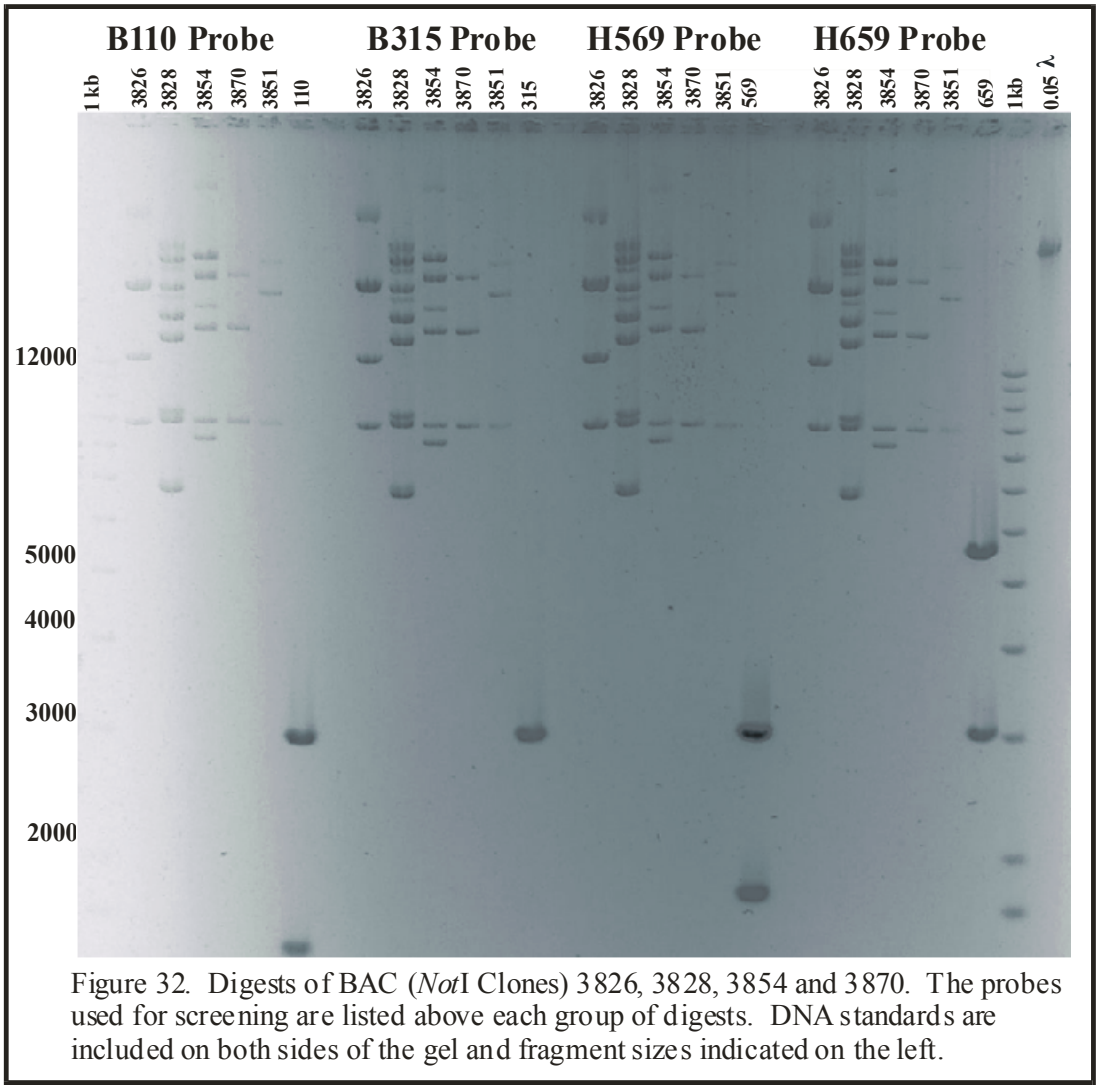


Figure 32. Digests of BAC (*Not*I Clones) 3826, 3828, 3854 and 3870. The probes used for screening are listed above each group of digests. DNA standards are included on both sides of the gel and fragment sizes indicated on the left.

Results for the Southern blot showed that positive and negative controls worked for blots probed with B110, B315 and H596. The H659 probe did not work at all. For each of the probes that worked, bands were visible for BAC clones 3870 and 3854 (Figure 33). The ends of both of these clones were sequenced. The BAC 3870 sequence aligned with the left end sequence of clone H596 and the right end sequence of clone B97. This filled in a large gap near the US terminus. The B315 probe showed positive because 3870 is quite

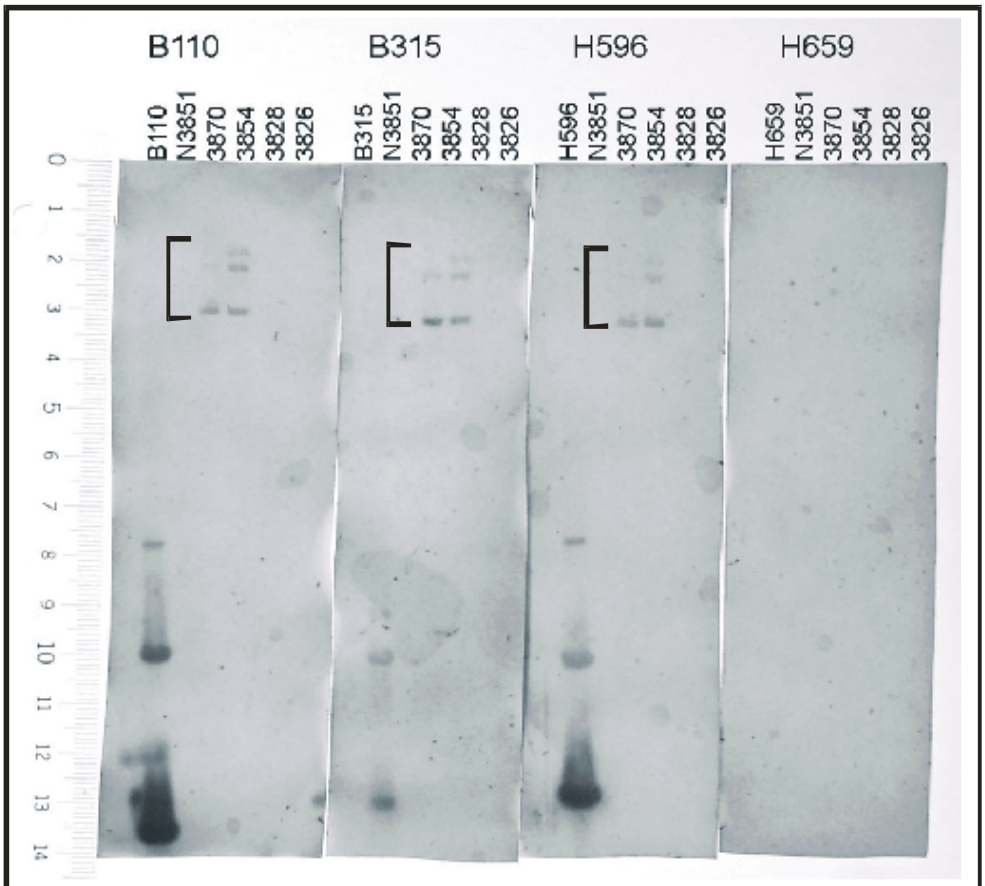


Figure 33. BAC Clones - Southern Blot Screening. Clone numbers on the horizontal axis indicate probe used for the blot. Clone numbers on the vertical axis indicate clone DNA being probed. Clones N3870 and N3854 were positive using probes B110, B315 and H596. Bands are indicated with brackets.

large (approximately 14 kbp) and the B315 clone lies within 3870. The BAC 3870 clone provided a considerable amount of new sequence data and overlapped with a number of clones (in the US region) that had not been accurately mapped (Figure 34). The 3854 BAC sequence duplicated an existing *Not* I clone, 1645, which is adjacent to the left end of 3870. It is not understood why these two clones showed positive results for the B110 probe since this clone was from a distant location on the BaCMV genome. However, there are blocks of repetitive DNA scattered throughout most herpesvirus genomes, and the B110 area could be repetitive as well. This could explain the homology of B110

which comes from an area of the genome with repetitive DNA sequences to other regions of the BaCMV genome.

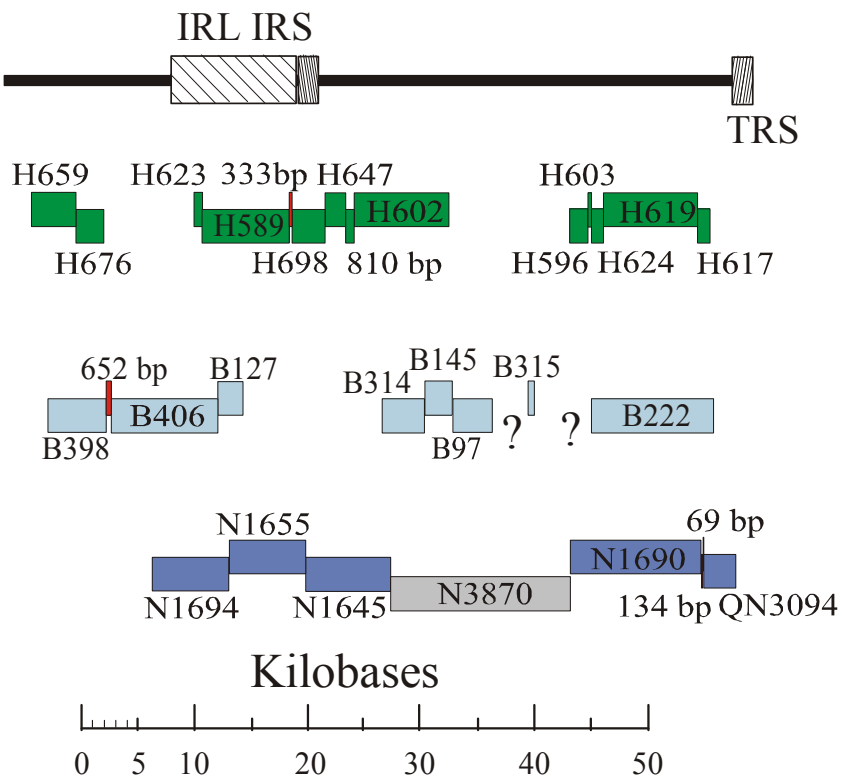


Figure 34. Alignment of BAC Clone *Not*I 3870 in US Region. Top figure shows a generalized representation of the Internal Repetitive, Unique Short and Terminal Repetitive Short Regions of the BaCMV Genome. H = *Hind*III clones, B = *Bam*HI clones, N = *Not*I clones, red boxes are PCR product and the grey box represents BAC clone N3870 and all are drawn to scale. Question marks represents previously unsequenced gaps in genome.

D. Generating BaCMV Clones or PCR Products for Sequencing

1. Genome Sequencing by Transposon Priming

At least 800 trace files from 21 different clones were generated through transposon mutagenesis. Of those shown in Figure 11, GPS clone 4428 (from B150) provided 1440 bp and GPS clone 4434 (from H602) produced 1420 bp of new sequence. While this represented only partial sequence for these two clones, for many of the other clones in the BaCMV library the internal DNA sequence was produced entirely using

transposon priming. Sequencing of GPS-1 clones yielded almost 102,000 base pairs which represents about one-half of the total genome.

2. Cloning of the UL1 Region Using HCMV Primers SP216 and SP217

Dot blot and Southern blot screening had not provided any sequence from the UL terminus, so PCR was used to amplify BaCMV DNA using HCMV primers at the UL1 region. The UL1 region is the coding region in HCMV adjacent to the UL terminus, so amplification of this sequence would likely provide sequence near the BaCMV UL terminus. After PCR, both the HCMV and BaCMV products showed a single band at approximately 600 bp as expected (Figure 12). However, the results of TOPO cloning of the BaCMV product were poor. The number of colonies on the plates was much greater than is normal for TOPO reactions, and many of the clones had multiple, unexpected bands. A second PCR amplification and TOPO cloning showed better results. Fewer colonies were observed on the two plates as is typical of TOPO cloning reactions. The size of the DNA inserts was the expected 600 bp, so clone 3599 was sent for sequencing. Although both primers yielded sequence, only one DNA trace would assemble into the existing contiguous DNA sequence because the other trace was all vector DNA. The trace that was BaCMV PCR product did not assemble into the UL1 region where it was expected to, but instead assembled near the US terminus overlapping N1690. This was probably because the HCMV primers were degenerate (similar, but not identical to, other sites in the genome) and bound at a location other than the desired UL1 region in BaCMV.

E. Generating DNA Across the B110/H659 Gap Using PCR

PCR amplification was used to span the gap between clones B110 and H659 at the 3' end of the UL region using various primer pairs. The approximately 1 kb product

generated with the primer pair SP467 and SP469 is shown in Figure 35 (A) and was TOPO cloned. Eight colonies were selected for miniprep purification and assigned clone numbers 7316 – 7323 (Figure 35 B).

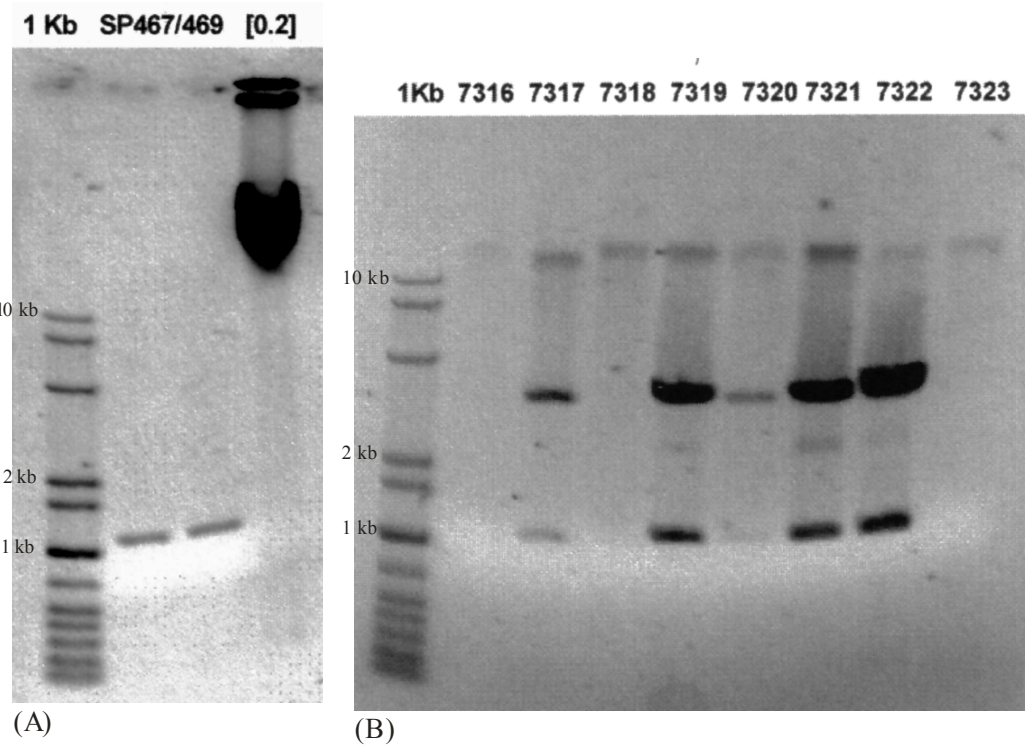


Figure 35. Gel Electrophoresis of SP467 and SP469 PCR Product and *Eco*RI Digests of TOPO Clones 7316 - 7323. (A) Left lane is size marker, two center lanes are PCR product and right lane is lambda DNA for determining concentration. (B) Left lane is size marker and other lanes are clone DNA as indicated.

As expected, five of the TOPO clones had a 3.9 kb vector band and a 1 kb insert that could be the PCR product. Clone 7321 was sent for sequencing and yielded approximately 800 bp of useful sequence. This did not span the gap between the clones, but did add new sequence data overlapping the 3' end of the B110 clone. The amplification product was generated by primer SP469 binding where it was designed to (in the B110 clone area) and binding at another site within the gap. The SP467 primer did not work so the PCR product did not span the gap.

Since no PCR product was generated by the other primer to span the gap, another primer (SP266) was used with the SP467 primer. The PCR product of primer pair SP469 and SP266 was about 1.2 kb and is shown in Figure 36. Twelve colonies from TOPO clones were selected for miniprep purification and assigned clone numbers 7336 – 7347 (Figure 37). Two clones, 7338 and 7342, were sent for sequencing and yielded good data from the ends of each clone. Both clones contained a 1234 bp insert which overlapped clone 7321 and added 351 additional bps to the 3' end of the gap.

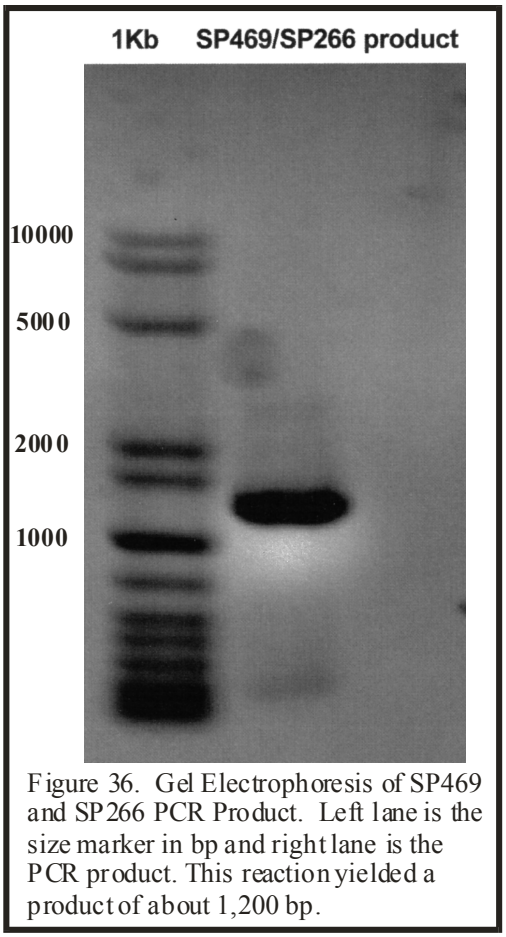


Figure 36. Gel Electrophoresis of SP469 and SP266 PCR Product. Left lane is the size marker in bp and right lane is the PCR product. This reaction yielded a product of about 1,200 bp.

While this sequence extended further outward from the B110 clone, it did not join with the H659 clone sequence to cover the gap as expected. This outcome was surprising since the lower primer, SP266, was complementary to sequence from clone H659. One explan-ation for the smaller PCR product could be that SP266 binds to a region closer to the B110 end of the gap. In fact, a good portion of the sequence from the 1,234 bp insert appears to be highly repetitive. The repetitive nature of this region could also account for the difficulty cloning and sequencing this portion of the genome.

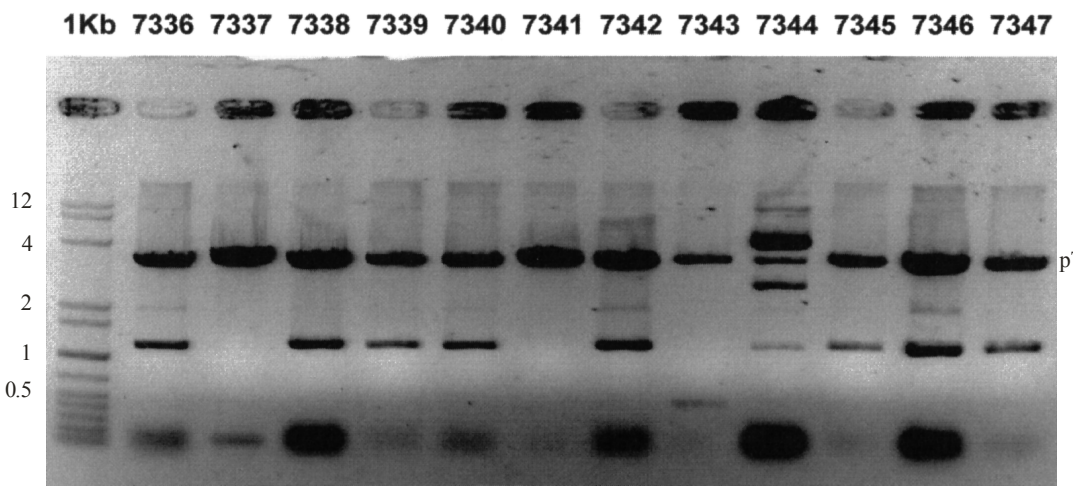


Figure 37. Gel Electrophoresis and *Eco*RI Digests of TOPO Clones 7336 - 7347. Left lane is the size marker, numbered lanes are TOPO clone DNA as indicated. All clones have a vector band at 3.9 kb and eight others have the expected 1.2 kb band.

F. Generating DNA Across the Internal N3851 Gap Using PCR

PCR was also used to amplify DNA for sequencing from the BAC clone N3851. Primer walking is difficult with BACs, and although we had some success, primer walking with BAC N3851 we did not generate useful sequence data for this area. New PCR primers (SP378 and SP471) were designed from sequence on either side of the gap for PCR. The 5.5 kb PCR product generated using these primers was used to create TOPO clones 7272 – 7278 (Figure 38). Three of these clones, 7272, 7274 and 7275, were sent for sequencing using the SP378 and SP471 primers but yielded no useful data. Most of the data files had either multiple traces or ambiguous DNA sequence. Though TOPO cloning was not successful, the PCR primers were used to sequence the ends of the 5.5 kbp product. Sequences generated from the SP471 and SP378 primers aligned within the BAC N3851 clone. The SP471 sequence aligned near the left end of the gap while the SP378 primer sequence aligned near the right end of the gap. These two sequences provided an additional ~1,300 bp of the original 5.5 kbp PCR product.

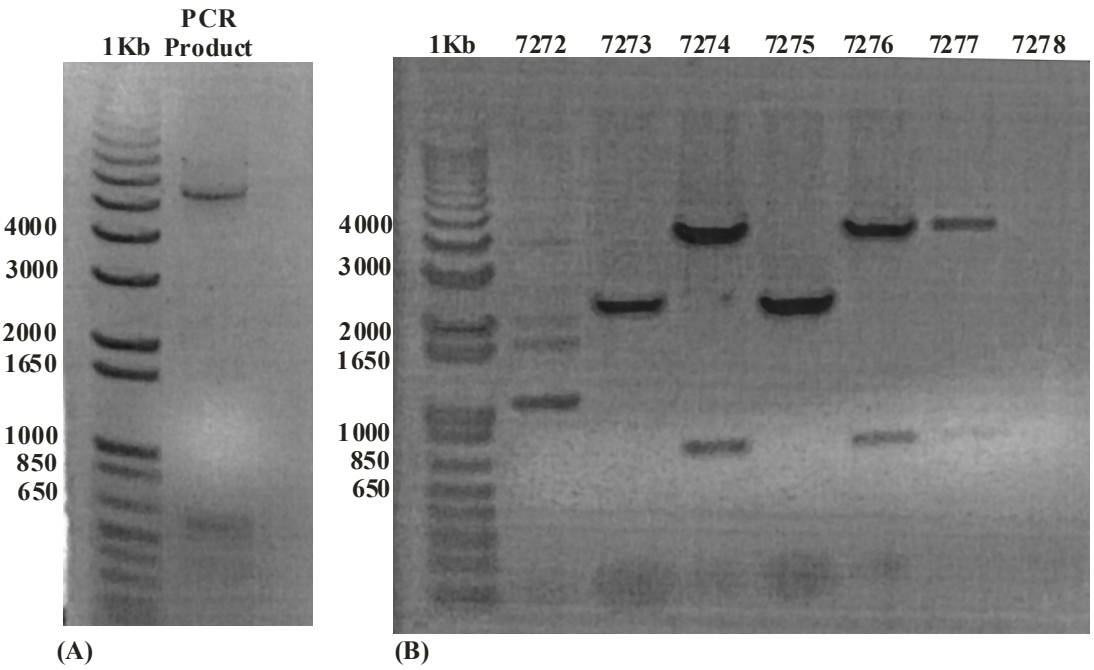


Figure 38. Gel Electrophoresis of SP378 and SP471 PCR Product and TOPO Clones 7272-7278. (A) Left lane is the DNA size standard and right lane is PCR product. The middle band is approximately 1.6 kb and is not prominent in this image. (B) Left lane is the DNA size standard and other lanes are TOPO clones of PCR product as indicated by number.

G. Generating Sequence Data Through Primer Walking

Because TOPO cloning of the PCR product from primers SP378 and SP471 was not successful, primer walking was used to generate new sequence from either the PCR product itself or from pBAC N3851 DNA. Primers 474, 486, 491, 492 and 497 were used for primer walking from the left end of the gap and yielded about 1,700 additional bp. Primers 475, 485, and 490 were used at the right end of the gap and yielded approximately 1,240 new bp. Sequence generated by PCR is not as reliable as sequence from clone DNA as PCR amplification is error prone. Therefore, the sequence produced by primer walking of PCR product was not as accurate as that of sequence generated from clone N3851. The sequence remaining to be determined within BAC clone N3851 is about 2,900 bp.

H. Mapping Existing and New Genomic Clones

BaCMV clones were initially mapped based on their homology to the HCMV AD169 genomic sequence. During the course of this project, the BLAST algorithm was improved and other alignment algorithms became available. Also, DNA sequences from rhesus and African Green monkey CMVs, began to appear in Genbank. This allowed many existing clones whose ends were sequenced to be mapped. Some of the genomic clones produced in attempts to clone the termini (such as N3870) were found to contain previously-uncloned sections of BaCMV DNA. Finally, this led to the discovery of clone N1670 which contained the 5' terminus of BaCMV. This was a large *Not* I clone (produced early in this project) which initially showed no significant homology to anything in Genbank.

I. Final Clone Homology Map

The preliminary clone map consisted of 46 individual clones. The final clone map consists of 83 individual clones and all but one of these was completely sequenced (see thick black bar in Figure 39). In the final version, the mapped clones are assembled to form 3 contiguous sequences representing about 217,000 bp of DNA. Within the BAC clone N3851 there is a gap of missing sequence which is approximately 2,900 bp in length. The projected amount of unsequenced DNA in this region is based on the known size of the BAC clone (as observed on an agarose gel) excluding the amount of DNA that has already been sequenced. Another gap near the UL terminus, between clones B110 and H659, is less-well defined. This is because there is no identified clone that spans this gap. Therefore the estimated range for this gap is between 4,000 and 9,000 bp based on comparison with closely-related primate CMV genomes.

III. DATA ANALYSIS

In the areas of the BaCMV genome that are sequenced, the overall GC content is 48.5%. This is very similar to the GC content for RhCMV which is 49%. In the RhCMV genome, there is a region (genes Rh20 – Rh27) that has a lower GC content (between 20 – 40%) and this sequence corresponds to the HCMV RL11 gene family. The sequenced BaCMV genes include homologues for Rh20 through Rh25, and also show a reduced GC content at about 40%. BaCMV sequence containing potential homologues to Rh26 – Rh28 has yet to be produced as this is the location of one of the remaining gaps. Putative BaCMV genes are numbered as shown in Table 8.

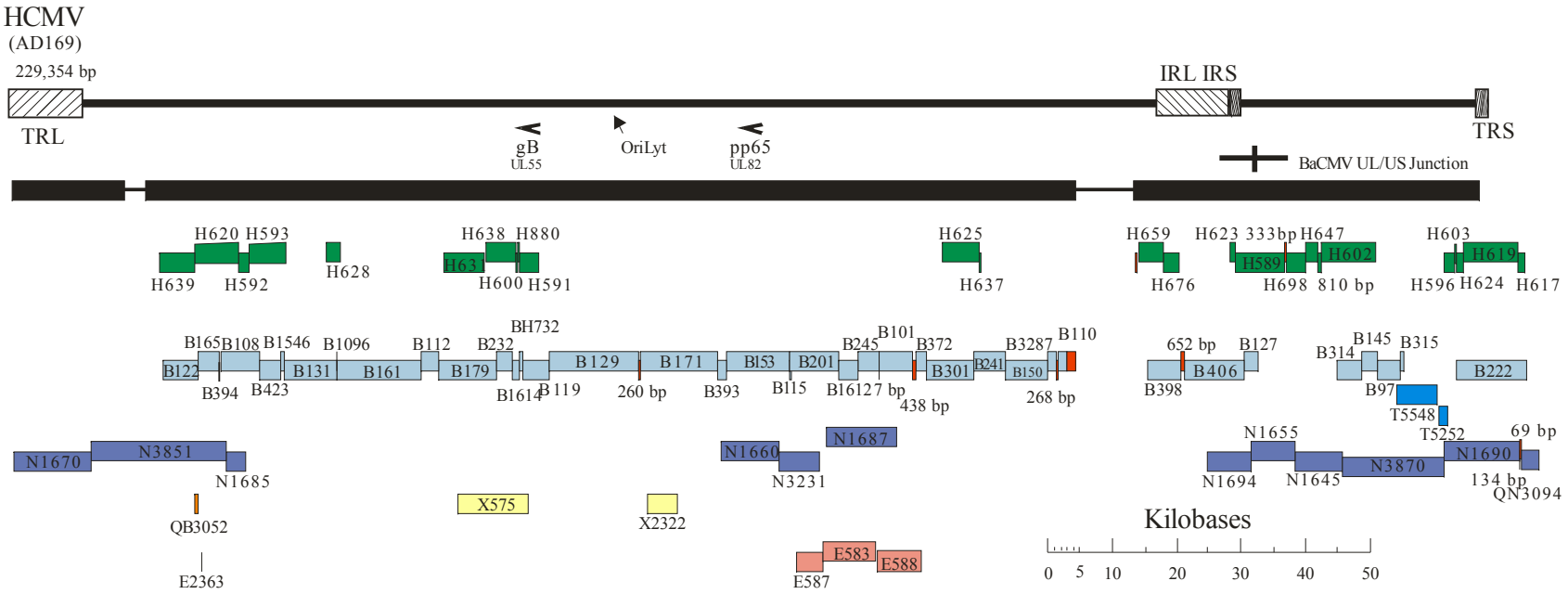


Figure 39. BaCMV (OCOM4-37) Final Clone Homology Map. The upper bar represents the sequenced human CMV (AD169) genome.The gB and pp65 genes are shown as arrowheads. The thin black bar represents the BaCMV (OCO M4-37) genome. The thick black bar represents the sequenced portions of the BaCMV genome.The boxes represent cloned BaCMV DNA fragments and have been placed on the map based on their similarity with HCMV. B = *Bam* HI clone, E = *Eco*RI clone, H = a *Hind* III clone, QB = a *Bam* HI-blunt clone, QN = a *Not* I-blunt clone, T = Topo-cloned PCR fragment, X = a *Xba* I clone and red is uncloned PCR fragments.

Table 8: Properties of the BaCMV (OCOM4-37) Genome

BaCMV ORF	Length (aa)	Predicted MW (kDa)	HCMV name		RhCMV name		Gene Family	Comments:
			Homologue	% Identity	Homologue	% Identity		
BaUL01	166	18.7	TRL1	33.7	Rh01	50.6		
BaUL02	105	11.5						
BaUL03	121	13.2						
BaUL04	220	23.9						
BaUL05	251	28.0	UL153		Rh05	41.8	RL11	Membrane protein
BaUL06	222	25.6						
BaUL07	217	24.7			rh07	23.8		
BaUL08	176	20.0			rh06	17.9		
BaUL09	173	20.5			rh07	26.0		
BaUL10	231	26.0			rh08	24.6		
BaUL11	173	20.5			rh07	21.4		
BaUL12	182	20.6			rh08	17.5		
BaUL13	185	21.0			rh08	18.1		
BaUL14 Ex1	135	14.2			rh10 Ex6	63.0		Cyclooxygenase-2
BaUL14 Ex2	89	10.1			rh10 Ex5	60.0		Cyclooxygenase-2
BaUL14 Ex3	113	13.4			rh10 Ex4	76.0		Cyclooxygenase-2
BaUL14 Ex4	47	5.5			rh10 Ex3	40.0		Cyclooxygenase-2
BaUL14 Ex5	115	12.2			rh10 Ex2	74.0		Cyclooxygenase-2
BaUL14 Ex6	65	7.4			rh10 Ex1	43.0		Cyclooxygenase-2
BaUL15	190	22.3			rh10	40.0		
BaUL16	200	22.6			rh12	18.5		
BaUL17	141	15.8			rh13.1	54		
BaUL18	186	20.3			rh14	47.9		
BaUL19	119	13.1			rh16			
BaUL20	296	33.6	UL04	19.1	Rh19	33.8	RL11	
BaUL21	184	20.5	UL06	17.4	Rh20	39.7	RL11	
BaUL22	236	27.1			Rh21	26.2	RL11	
BaUL23	236	27.1	UL11		Rh22	29.4	RL11	Membrane protein
BaUL24	242	27.6	UL11		Rh23	46.3	RL11	Membrane protein
BaUL25	152	17.1	UL09		Rh25	48	RL11	Membrane protein
BaUL28	217	23.9	UL11		Rh29		RL11	Membrane protein
BaUL29	193	21.5	UL11		Rh29		RL11	Membrane protein
BaUL30	164	18.0			rh30	55.1		
BaUL31	360	41.3	UL13	16.7	Rh31	44.4		
BaUL32	307	35.5	UL14	28.7	Rh33	57.4		
BaUL33	109	12.6	UL17	21.2	Rh35	33.6		
BaUL34	98	10.9	UL19	38.8				
BaUL35	465	52.1	UL20	23.8	Rh36	42.4		
BaUL36	145	16.4	UL21A		rh37	59.5		
BaUL37	312	35.8	UL23	34.9	Rh40	57.1	US22	
BaUL38	206	230	UL24	46.5	Rh42	63.1	US22	
BaUL39	584	67.4	UL25	41.3	Rh43	56.8	US25	
BaUL40	238	27.0	UL26	44.1	Rh44	72.3		
BaUL41	577	65.6	UL27	54.8	Rh46	71.4		
BaUL42	337	38.7	UL28	66.5	Rh47	82.2	US22	
BaUL43	336	38.7	UL29	60.7	Rh50	76.8	US22	Immediate early protein
BaUL44	542	61.2	UL31	55.4	Rh54	77.2		
BaUL45	777	85.9	UL32	28.2	Rh55	48.1		pp150
BaUL46	318	35.9	UL33	56.9	Rh56	76.7	7TM	G-protein-coupled receptor
BaUL47	286	32.7	UL34	58.7	Rh57	79.3		
BaUL48	597	68.7	UL35	44.2	Rh59	61.4	UL25	

Continued on next page:

Table 8: Properties of the BaCMV (OCOM4-37) Genome (continued)

BaCMV ORF	Length (aa)	Predicted MW (kDa)	HCMV name		RhCMV name		Gene Family	Comments:
			Homologue	% Identity	Homologue	% Identity		
BaUL49	450	51.8	UL36	33.5	Rh60	49.8	US22	
BaUL50	69	7.9	UL36	53.7	Rh61	62.3	US22	
BaUL51			UL37	37.0	Rh62	51.0		
BaUL52	306	34.8	UL38	47.1	Rh64	76.3		
BaUL53	101	11.7	UL37		Rh66			Nuclear egress lamina protein
BaUL54	139	15.3			rh67	35.3		
BaUL55	143	16.0	UL42	32.5	Rh68	53.5		Glycoprotein
BaUL56	340	39.2	UL43	42.4	Rh69	56.5	US22	Tegument protein
BaUL57	388	43.6	UL44	68.1	Rh70	90.3		DNA polymerase processivity factor
BaUL58	833	94.8	UL45	57.6	Rh72	70.0		Ribonucleotide reductase-1
BaUL59	290	33.1	UL46	71.0	Rh75	84.1		Minor capsid binding protein
BaUL60	960	110.7	UL47	43.0	Rh76	73.7		Capsid assembly protein
BaUL61	793	89.3	UL48	51.0	Rh78	73.0		Large tegument protein
BaUL62	470	53.5	UL49	74.0	Rh80	83.6		
BaUL63	303	33.6	UL50	52.8	Rh81	66.0		
BaUL64	104	11.7	UL51	82.0	Rh82	92.0		
BaUL65	556	63.3	UL52	60.4	Rh83	79.3		Virion protein
BaUL66	288	33.0	UL53	72.2	Rh85	86.5		Virion protein
BaUL67	1039	116.5	UL54	72.0	Rh87	81.4		DNA polymerase
BaUL68	877	99.9	UL55	56.9	Rh89	73.6		Glycoprotein B
BaUL69	773	88.8	UL56	79.3	Rh91	91.4		Transport capsid assembly protein
BaUL70	1043	116.5	UL57	71.1	Rh92	92.5		Major DNA binding protein
BaUL71	42	4.5			rh93	69.0		
BaUL72	293	31.1			rh96	62.3		
BaUL73	746	43.4	UL69	43.4	Rh97	58.5		
BaUL74	207	22.6			rh99	35.7		
BaUL75	972	111.9	UL70	68.0	Rh100	81.3		DNA helicase-primase component
BaUL76	234	26.3	UL71	55.6		0.0		Tegument protein
BaUL77	348	39.9	UL72	49.1	Rh101	67.1		
BaUL78	76	8.4	UL73	26.3	Rh102	52.6		Glycoprotein N
BaUL79	427	49.4	UL74	35.0	Rh103	59.5		Glycoprotein O
BaUL80	719	81.2	UL75	48.9	Rh104	64.7		Glycoprotein H
BaUL81	293	32.6	UL76	53.2	Rh105	76.8		
BaUL82	592	67.0	UL77	68.6	Rh106	83.8		Pyruvoyl decarboxylase
BaUL83	363	42.2	UL78	25.2	Rh107	41.4	7TM	GCR homologue
BaUL84	265	30.5	UL79	77.4	Rh108	90.6		
BaUL85	652	70.0	UL80	44.4	Rh109	71.0		Capsid assembly protein
BaUL86	532	60.0	UL82	37.2	Rh110	59.6	UL82	Major late antigen; pp71
BaUL87	553	63.0	UL83	32.7	Rh111	47.9	UL82	Phosphorylated matrix protein (pp65)
BaUL88	525	59.8	UL83	38.9	Rh112	51.2	UL82	Phosphorylated matrix protein (pp65)
BaUL89	467	52.7	UL84	53.7	Rh114	76.4		Early nonstructural protein
BaUL90	120	66.4			rh115	71.0		
BaUL91	302	34.2	UL85	75.2	Rh117	84.2		Minor capsid protein; mCP
BaUL92	1343	151.3	UL86	77.2	Rh118	87.8		Major capsid protein; MCP
BaUL93	104	11.8			rh119	84.6		
BaUL94	846	96.0	UL87	69.9	Rh122	85.9		Virion protein
BaUL95	385	46.8	UL88	56.5	Rh123	71.5		Virion protein
BaUL96	423	48.1	UL89	78.4	Rh124	87.8		
BaUL97	114	12.1	UL91	55.0	Rh126	73.0		
BaUL98	252	28.5	UL92	88.6	Rh127	81.1		Virion protein
BaUL99	409	47.0	UL93	52.8	Rh128	72.1		Virion protein

Continued on next page:

Table 8: Properties of the BaCMV (OCOM4-37) Genome (continued)

BaCMV ORF	Length (aa)	Predicted MW (kDa)	HCMV name		RhCMV name		Gene Family	Comments:
			Homologue	% Identity	Homologue	% Identity		
BaUL100	345	37.9	UL94	55.1	Rh129	78.8		Virion protein
BaUL101	431	47.6	UL95	68.2	Rh130	80.3		Virion protein
BaUL102	127	14.5	UL96	53.9	Rh131	72.4		Virion protein
BaUL103	503	57.4	UL97	57.1	Rh132	65.4		Phosphotransferase; phosphorylates ganciclovir
BaUL104	559	63.2	UL98	69.2	Rh134	83.1		DNase; exonuclease
BaUL105	112	65.0			rh135	75.0		
BaUL106	131	14.6	UL99	29.0	Rh136/137	57.3		Phosphoprotein (pp28)
BaUL107	359	41.4	UL100	54.9	Rh138	70.8		Glycoprotein M
BaUL108	737	80.9	UL102	62.4	Rh139	79.4		Helicase-primase;
BaUL109	187	21.4	UL103	47.1	Rh140	55.1		
BaUL110	655	77.5	UL104	72.4	Rh141	86.1		Structural protein; DNA packaging/ cleavage protein
BaUL111	877	99.0	UL105	77.1	Rh142	92.2		DNA helicase
BaUL112	139	15.7	UL111		Rh143	42.4		IL-10-like protein;
BaUL113a	274	29.0	UL112	38.1	Rh144	53.4		
BaUL113b	271	28.1	UL113	39.9	Rh145	58.3		
BaUL114	248	28.4	UL114	67.7	Rh146	80.6		Uracil-N glycosylase
BaUL115	257	29.4	UL115	42.9	Rh147	75.0		Glycoprotein L
BaUL116	340	36.2	UL116	25.9	Rh148	43.2		
BaUL117	388	43.2	UL117	49.5	Rh150	69.6		
BaUL118	200	23.6	UL118	32.5	Rh151	55.3		Large splice transcript
BaUL119	216	21.8	UL119	23.2	Rh152	44.4		
BaUL120	204	23.3	UL120	33.3	Rh154	64.1		
BaUL121	183	21.2	UL121	26.3	Rh155	52.0		Serine/alanine-rich glycoprotein
BaUL122	320	35.1	UL122/3	53.4	Rh156	70.3		Immediate early protein 2 (MIE2)
BaUL123					rh157			
BaUL124	144	16.8	UL147	28.0	Rh158	42.9		
BaUL125			UL148		Rh159			
BaUL126			UL132		Rh160			BaCMV UL seq incomplete
BaUL127	136	15.6	UL146	30.7	RhUL146	56.1		
BaUL128	99	11.1	UL145	49.5	Rh162	85.9		HCMV Toledo
BaUL129	177	22.9	UL144	27.3	Rh163	62.0		HCMV Toledo; TNF receptor
BaUL130	443	50.3	UL141	30.6	Rh164	52.8		HCMV Toledo
BaUL131	193	21.7			rh165	30.8		
BaUL132	193	21.7	UL133		Rh166	39.1		
BaUL133	221	24.0			rh167	38.0		
BaUL134	208	23.0			rh168	26.0		
BaUL135	187	20.7			rh169	40.0		
BaUL136	149	17.1			rh170	34.0		
BaUL137	319	34.6	UL133		Rh171	42.9		*Cercopithecine* HV5 SPLF5 glycoprotein; Cys-rich
BaUL138	176	20.0			rh172	57.6		*Cercopithecine* HV5 SPLF4 glycoprotein; Cys-rich
BaUL139	397	42.0			rh173	44.8	RL12	*Cercopithecine* HV5 SPLF3 glycoprotein
BaUL140	168	17.1			rh173	57.0	12508-13013	*Cercopithecine* HV5 SPLF2 glycoprotein; Ser/Thr rich

Continued on next page:

Table 8: Properties of the BaCMV (OCOM4-37) Genome (continued)

BaCMV ORF	Length (aa)	Predicted MW (kDa)	HCMV name		RhCMV name		Gene Family	Comments:
			Homologue	% Identity	Homologue	% Identity		
BaUL141	339	38.4			rh174	32.7		*Cercopithecine* HV5 SPLF1 glycoprotein
BaUL142	185	20.8			rh176	31.4		
BaUL143	204	22.5			rh178	31.9		
BaUS1	699	77.1	TRS1	36.0	Rh230	59.8	US22	Transcriptional transactivator; tegument protein (AGM SPLF5)
BaUS2	189	22.3	US32	39.3	Rh226	79.7	US1	(AGM SPLF4)
BaUS3	171	19.5	US31	32.0	Rh225	50.3	US1	Site-specific DNA methylase (AGM SPLF3)
BaUS4	205				Rh224	51.0		(AGM SPLF2)
BaUS5	296	33.6	US30	26.8	Rh223	46.9		(AGM SPLF1)
BaUS6	337	37.9	US29	37.8	Rh221	54.1		
BaUS7	494	55.0	US28.5	42.5	Rh220	66.7	7TM	GCR homologue
BaUS8	359	41.0	US28.1	27.5	Rh218	42.2	7TM	GCR homologue
BaUS9	330	38.1	US28.2		Rh216	63.6	7TM	GCR homologue
BaUS10	330	38.1	US28.3	23.7	Rh215	34.3	7TM	GCR homologue
BaUS11	326	36.1	US28.4	25.5	Rh214	67.8	7TM	GCR homologue
BaUS12	483	55.9	US26	52.0	rh211	71.5	US22	Early nuclear protein; tegument protein
BaUS13	484	56.9	US24	64.5	rh209	80.9	US22	
BaUS14	599	69.8	US23	58.7	Rh204	76.8	US22	
BaUS15	368	42.2	US22	55.4	Rh203	73.9	US22	ICP22
BaUS16	170	19.3	US21	49.7	Rh202	69.6	US12	Uncharacterized protein family, UPF0005
BaUS17	274	31.4	US20	45.0	Rh201	62.1	US12	Probable transmembrane protein
BaUS18	261	29.5	US19	26.8	Rh200	44.1	US12	
BaUS19	264	29.7	US18	29.8	Rh199	54.9	US12	
BaUS20	264	29.5	US17	34.5	Rh198	68.3	US12	
BaUS21	227	25.9	US14		Rh197	34.6	US12	
BaUS22	270	30.6	US14	29.8	Rh196	56.5	US12	
BaUS23	242	26.9	US14		Rh195	34.3	US12	
BaUS24	277	31.5	US14		Rh194	38.0	US12	
BaUS25	251	28.6	US13	21.5	Rh192	51.8	US12	
BaUS26	248	28.3	US12	34.7	Rh190	70.7	US12	
BaUS27	260	30.3	US11	33.5	Rh189	41.8	US6	
BaUS28	187	20.8	US8	29.5	Rh187	36.4	US6	
BaUS29	209	25.4	N/A		rh186	23.7	US6	
BaUS30	166	18.3	US6	28.3	Rh185	46.5	US6	Functional homologue of US6
BaUS31	190	21.4	US3	21.2	Rh184	35.4	US3	Immediately-early glycoprotein
BaUS32	354	40.8	US2	31.4	Rh182	31.2	US6	
BaUS33	158	18.3	US1	57.0	Rh181	69.5	US1	
BaUS34	129	14.2			rh179	51.2		

Table 8. Properties of the BaCMV (OCOM4-37) Genome

Most ORFs are numbered based on homology with HCMV genes. SPLF genes, however, are numbered based on homology with AGMCMV. Shaded areas represent different contiguous sequences.

Match with UL146 was produced with undesignated Rhesus CMV, accession #AAO40076.

NS, not significant; aa, amino acids; MW, molecular weight; SPLF, SalI-P fragment near L/S segment junction; GCR, G-protein-coupled receptor; 7TM, seven-transmembrane; ICP, infected cell protein.

% Identity is based on amino acid sequence homology.

Typically, the genes in herpesviruses are collinear with respect to related herpesviruses genomes. Co-linearity of genomes becomes more prominent within herpesvirus subfamilies. HCMV is in the beta-herpesvirus subfamily, and four different genomic isomers form as the DNA replicates. However, the animal CMVs, which are also in the beta-herpesvirus subfamily do not form such isomers and are found in a single UL/US conformation. An unusual feature of the BaCMV OCOM4-37 genome is that the US region is inverted relative to the genomes of CCMV and RhCMV. There are three individual BaCMV clones (B127, H589 and N1670) that separately contain the UL/US junction (Figure 40). These clones were individually obtained at different times with different REs using genomic BaCMV OCOM4-37 DNA. The *Bam*HI and *Hind*III clones were made using DNA prepared in 1997, immediately after the virus was isolated. The *Not*I clone was made using DNA prepared from virus grown in 1998. Having three independent clones that contain the same sequence suggests that this is the correct orientation of the viral genome and is unlikely to be an artifact of cloning. It is possible that other isomers also exist in BaCMV, but we have never obtained a clone supportive of that.

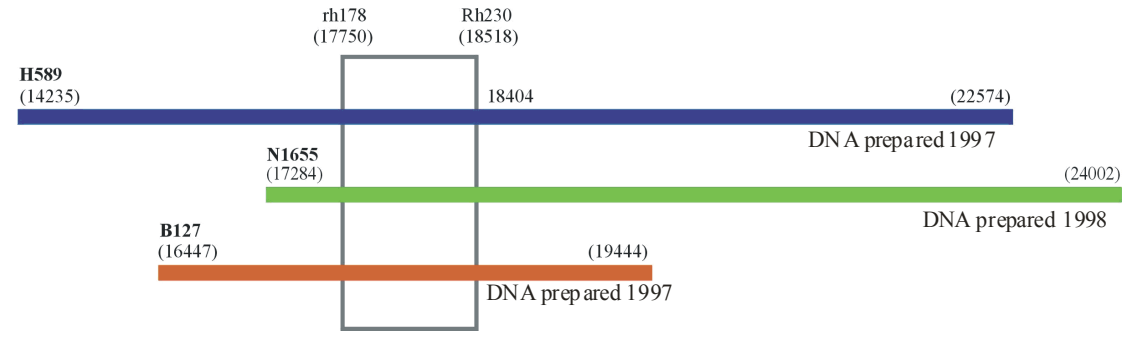


Figure 40. Map of Clones *Hind*III 589, *Not*I 1655 and *Bam*HI 127. Clone H589 is 8,340 bp, clone N1655 is 6,918 bp and clone B127 is 2,997 bp. Numbers at ends indicate location in unique short alignment of the genome. Rectangular box shows region where the three clones overlap at UL/US junction.

While the BaCMV ORFs are co-linear with the US region of CCMV and RhCMV, the entire US is inverted with respect to CCMV and RhCMV (Figure 41). In

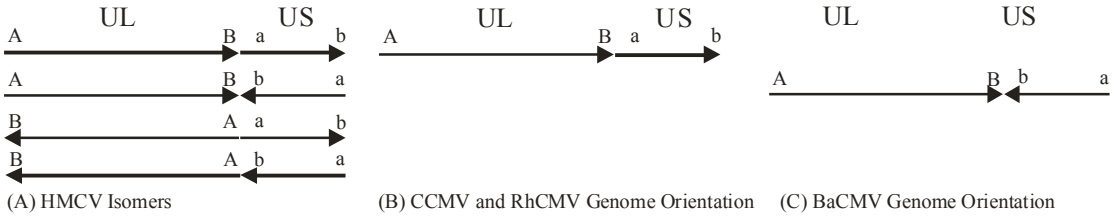


Figure 41. Comparison of Four Cytomegalovirus Genomes. (A) Shows four different isomers that exist for HCMV. Letters A and B designate ends of the UL segment and letters a and b designate ends of US segments. (B) Shows orientation of the chimpanzee (CCMV) and rhesus (RhCMV) genomes. (C) Shows orientation of the BaCMV OCOM4-37 genome. Direction of arrows shows orientation of UL and US regions. CCMV, RhCMV and BaCMV genomes are aligned horizontally with respect to the corresponding HCMV isomer.

the HCMV genome, the sequence where inversion of the UL and US segments occurs is the UL/US joint region and is characterized by highly repetitive DNA containing direct and inverted repeats. This is also characteristic of the BaCMV inversion site (Table 9). Though similar in this respect, the BaCMV repetitive region is considerable smaller (1,354 bp) than it's counterpart in HCMV (10,277 bp).

Repetitive Sequence	Number of Repeats in Junction Site
GGGGG	3
GAAGA	4
ATAAA	4
TTTG	5
GGGGA	5
GGGGC	8
CACAC	8
AAAA	10

Table 9. Example of Repetitive Elements in the BaCMV UL/US Region

A. Individual ORFs: Functional proteins

1. DNA replication proteins.

There are six essential replication proteins in HCMV, CCMV and RhCMV that are also found in BaCMV, and these are the most conserved genes in all herpesviruses.

In HCMV, these proteins are the DNA polymerase (UL54); accessory protein UL44, also referred to as the DNA polymerase processivity factor; DNA helicase (UL105); single-stranded binding protein (UL57); and the helicase-primase components (UL70 and UL102) (Appleton et al. 2004). In HSVI, there is a seventh required protein, the origin binding protein, which has not been found in any cytomegalovirus.

HCMV DNA polymerase enzyme is in the type B polymerase family which functions in base excision repair and replication. The sequence for this gene homologue in BaCMV (BaUL67) shares identity with other primate DNA polymerases as follows: 81.4% with RhCMV, 71.7% with CCMV and 72% with HCMV.

The gene for the HCMV accessory protein is also found in BaCMV (BaUL57) and shares identity with homologous genes of RhCMV (89.5%), CCMV (72.8%) and HCMV (73.5%). In clinical strains of HCMV, the N-terminal 309 aa of this protein are essential for binding DNA polymerase (Weiland et al. 1994) and this is where greatest conservation within these four sequences is found (Figure 42).

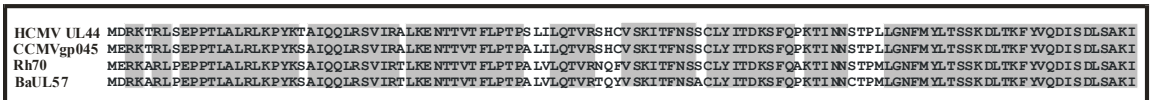


Figure 42. Conservation of the CMV Accessory Protein. Figure shows the first 110 amino acids of the alignment. Shaded gray areas s how sequence having 100% identity. Sequences were aligned with MegAlign (DNA Star) using ClustalW algorithm.

Six conserved motifs are required for proper function of DNA helicase and these have been identified in HSV, HCMV & RhCMV; these motifs exist in BaCMV as well (Figure 43). Within cells, DNA helicase is involved in unwinding the DNA double helix during replication. In herpesviruses, DNA helicase (HCMV UL105) interacts with the helicase- primase components (HCMV UL70 and UL102) to form a heteromeric complex that is required for viral DNA synthesis.

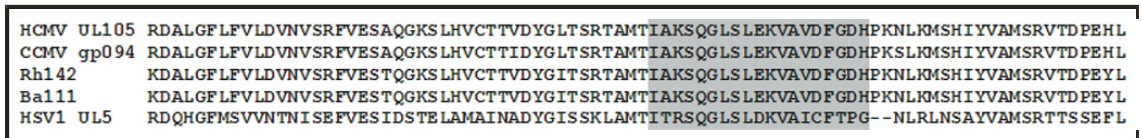


Figure 43. Conservation of the Helicase Protein. Alignment for the DNA Helicase sequence of 5 herpesviruses. AAs enclosed in gray box correspond to motif V. Note 100% conservation in these 19 aas for all cytomegaloviruses represented here and approximately 75% conservation with HSV1.

Additional *trans*-activating factors seen in HCMV and RhCMV are also found in BaCMV. In HCMV, these proteins include UL36-UL38 (BaUL49, BaUL50, BaUL51, BaUL52, BaUL53), TRS1 (BaUS1), IE1 (BaCMV homologue only partially sequenced), IE2 (BaUL122), UL84 (BaUL89) and UL112/113 (BaUL113a, BaUL113b) all of which are required for transient complementation of ori-Lyt-dependent DNA replication (Pari and Anders 1993). In BaCMV, there are two homologues for each of the HCMV UL36 and UL37 genes (see Table 8). HCMV ORFs UL36-UL38, TRS1, IE1 and IE2 are all recognized as regulatory proteins. HCMV UL84 encodes a protein that initiates DNA replication from the origin of replication (Sarisky and Hayward 1996) and the major protein of UL112/113 (pp43) interacts with IE2 as a co-transcriptional activator of the UL54 (DNA polymerase) promoter (Li et al. 1999). HCMV UL36 and UL37 have been characterized as anti-apoptotic proteins which inhibit cell death and thereby promote viral replication (Andoniou and Degli-Esposti, 2006). An origin of lytic replication (ori-lyt) in BaCMV has been identified between homologues of the HCMV UL57 and UL69 genes. In the large UL contig (Figure 16), this would be located between nucleotides 67,068 to 69,382. One feature of the ori-lyt sequence is the Y-block which consists of an oligopyrimidine stretch of 31 nucleotides in HCMV AD169 (Huang et al. 1996). A similar stretch of 33 nucleotides is present in BaCMV OCOM4-37. Additionally, the HCMV ori-lyt resides in a complex repetitive region of the genome (Masse et al. 1992)

which is also true for the BaCMV ori-lyt site. As a point of interest, this region of the genome has been particularly difficult to sequence in all herpesviruses. Sequencing the ori-lyt region in BaCMV required subcloning fragments as small as 400 bp since sequencing attempts on larger fragments often stopped at the ori-lyt site.

Sequence motifs designated *pac1* and *pac2* are conserved in herpesviruses. These motifs are typically located about 30-35 bp from each terminus of the genome and are believed to function in cleavage of concantemeric viral DNA prior to packaging. The *pac*1 DNA sequence motif is characterized by an A-rich region flanked by poly-C runs, and *pac2* is characterized by an A-rich region which is often near a CGCGGCG sequence. For murine cytomegalovirus (MCMV), the *pac1* poly-C run was found to be critical for cleavage site function while the A-rich region was not; for *pac2*, the A-rich region and a distal CGCGGCG motif were both required (McVoy et al. 1998).

In BaCMV OCOM4-37, a motif that matches the *pac1* site is located between 196,182 and 196,207 and compares with the HCMV sequence as seen in Figure 44.

```
CGGGGGGGTGTTTccgcgGGGGGG (BaCMV)
CGGGGGGGTGTTTttagcGGGGGG (HCMV)
```

Figure 44. Comparison of BaCMV OCOM4-37 and HCMV *pac1* Sequences. The A-rich region and poly-C runs are on the complementary strand. Lowercase letters on sequence indicates non-homologous bases.

This location is not 30-35 bp from the terminus as is seen in HSV-1. In HCMV Towne, both the *pac1* and *pac2* motifs are in the US_c region (see Figure 1 for location). However in HSV-1, *pac1* is in the UL_b region and *pac2* in the US_c region (Kemble and Mocarski 1989). This places the *pac1* site for the Towne strain between 30912-30935 bp and the *pac2* site between 226792-226821 bp. Therefore the proposed site of *pac1* in BaCMV is comparable to *pac1* in RhCMV.

The *pac2* motif in HCMV is not well-defined since sequences that match the criteria for *pac2* are not found in the expected orientation to *pac1* (McVoy et al. 2000). Essentially this means that *pac2* elements have not been found on the ends of HCMV genomes. Similarly, a good candidate for a BaCMV *pac2* site has not been identified.

2. Additional Genes Functioning in DNA Replication and Nucleotide Metabolism

In HCMV, uracil-DNA glycosylase (UL114) is involved in excision repair of inappropriate uracil nucleotides. The homologous gene in BaCMV shares 7l% identity with HCMV and 79.8% with RhCMV. Another HCMV enzyme, dUTPase (UL72) plays a role in preventing misincorporation of uracil into DNA during replication (Chen et al. 2002) and this BaCMV homologue shares 49.1% identity with HCMV and 67.1% with RhCMV.

The HCMV UL45 gene codes for a tegument protein which corresponds to the large subunit of ribonucleotide reductase. Normally, ribonucleotide reductase helps regulate dNTP concentrations within the cell, but UL45 lacks a number of catalytic residues, so its role is not entirely understood. Nevertheless, it has been shown to affect virus growth at low m.o.i. (Patrone et al. 2003). The BaCMV homologue to UL45 is conserved, having 47.9% identity with HCMV and 73.3% with RhCMV.

In HCMV, the phosphotransferase gene (UL97) has been shown to function as a protein kinase (Chou and Meichsner, 2000; Rawlinson et al. 1997). The UL97 gene product functions in phosphorylation of ganciclovir (one of the few drugs available to treat HCMV infections) which is necessary for efficacy of the drug (Michel et al. 1998; Sullivan et al. 1992). HCMV strains that are resistant to ganciclovir contain mutations between codons 591-607 of the UL97 gene which impairs phosphorylation and hence

sensitivity to the drug (Chou et al. 2002). The BaCMV gene shares 50.2% identity with HCMV and 72.9% with RhCMV. For the 17-base-pair sequence between codons 591-607, BaCMV shares 82.4% identity with HCMV AD169. None of the aa changes in HCMV leading to ganciclovir resistance were present in the collinear region in BaCMV. This would follow since drug resistance is usually seen in strains of virus exposed to the drug, and administration of ganciclovir would be more common in human than in baboon CMV infections.

Finally, the HCMV pyruvoyl decarboxylase (UL77) has a putative role in polyamine biosynthesis (Yoakum 1993) and the BaCMV homologue shares 70.3% identity with HCMV and 80.4% with RhCMV.

B. ORFs That Encode Regulatory Genes.

HCMV ORFs that regulate viral gene expression include the major immediate-early exon UL122-UL123 (BaUL122), UL36-UL38 (BaUL49-BaUL53), TRS1 (BaUS1), UL69 (BaUL73) and US3 (BaUS31). BaCMV homologues exist for each of these genes. In HCMV, the TRS1 transcript works in conjunction with the IRS1 transcript to transactivate UL44 and other genes (Stasiak and Mocarski 1992) and also plays a role in evasion of antiviral host cell responses (Child et al. 2004).

C. ORFs That Encode Structural Proteins.

Betaherpesvirus genomes code for a variety of capsid and tegument proteins and also viral glycoproteins. When compared with the HCMV AD169 prototype genome, the BaCMV genome possesses homologues for the majority of these genes.

Capsid gene homologues include the major capsid protein, UL86 (BaUL92), minor capsid binding protein UL46 (BaUL59) (Gibson et al. 1996) and minor capsid protein, UL85 (BaUL91).

There are also a number of HCMV tegument protein homologs that have been identified in BaCMV (Table 10). First, the UL48 large tegument protein (BaUL61) is

HCMV Gene	BaCMV Homolog	Protein name	Function	Reference
UL31	BaUL44		Nuclear egress	(Davison et al. 2003b)
UL32	BaUL45	pp150	Basic phosphoprotein	(Liu and Stinski, 1992)
UL36	BaUL49, BaUL50		Maintains structural integrity	(Zhou et al. 1999)
UL37	BaUL53		Interacts with UL36 and likely comprises second layer of the integument	(Klupp et al. 2002)
UL47	BaUL60		Unpackaging and transport of capsids to nuclear pore	(Hyun et al. 1999)
UL48	BaUL61	pp212	Complexed with tegument protein UL37; ubiquitin-specific protease (N-terminal region); involved in capsid transport	(Ogawa-Goto et al. 2002)
UL65	---		Protein kinase	(Chee et al. 1990)
UL80	BaUL85		Capsid assembly protein	(Nguyen et al. 2008)
UL82	BaUL86	pp71	Upper matrix protein; IE gene transactivator	(Ruger et al. 1987)
UL83	BaUL87	pp65	Lower matrix protein; protein kinase	(Nowak et al. 1984a)
UL99	BaUL106	pp28	Required for trafficking	(Jones and Lee, 2004)

Table 10. HCMV Cytomegalovirus Tegument Proteins and BaCMV Homologues

important for intracellular transport of virions (Ogawa-Goto et al. 2002). Second, the UL80 capsid assembly protein (BaUL85) transports the major capsid protein into the nucleus (Nguyen et al. 2008). Third, UL82 (BaUL86) the upper matrix protein pp71 functions as a transcriptional activator (Liu and Stinski 1992). Fourth, UL83 (BaUL87) is the lower matrix protein pp65 and a transactivator of IE gene expression (Preston and Nicholl, 2005) and disruption of the MHC class I antigen presentation pathway

(Trgovcich et al. 2006). Fifth, UL32 (BaUL45) the large phosphoprotein pp150 that is critical for virion egress, possibly at the stage of final envelopment (AuCoin et al. 2006). Sixth, UL99 (BaUL106) the small phosphoprotein pp28 that plays a role in the assembly (Sanchez et al. 2000) and cytoplasmic envelopment (Seo and Britt 2007) of the virion. And finally, UL56 (BaUL69) is a transport protein which is likely involved in cleavage and packaging of the capsid (Bogner et al. 1998). Clearly, there are a number of HCMV tegument genes for which homologs exist in the BaCMV genome. See Table 8 for conservation information between BaCMV and HCMV.

There are genes for various glycoproteins in all primate CMVs that are also present in the BaCMV genome. GenBank searches identified the following HCMV glycoproteins that are also found in both RhCMV and BaCMV (Table 11).

Glycoprotein	HCMV	BaCMV	RhCMV	BaCMV Conservation
---	UL37	BaUL53	Rh66	57.1% with RhCMV
---	UL42	BaUL55	Rh68	53.5% with RhCMV
gB	UL55	BaUL68	Rh89	53.5% with RhCMV
gN	UL73	BaUL78	Rh102	41.4% with RhCMV
gO	UL74	BaUL79	Rh103	59.5% with RhCMV
gH	UL75/	BaUL80	Rh104	81.3% with RhCMV
gM	UL100	BaUL107	Rh138	70.8% with RhCMV
gL	UL115	BaUL115	Rh147	75.0% with RhCMV
---	UL116	BaUL116	Rh148	43.2% with RhCMV
---	UL121	BaUL121	Rh155	52.0% with RhCMV
---	US02	BaUS32	Rh182	31.4% with HCMV
IE	US03	BaUS31	Rh184	35.4% with RhCMV
---	US11	BaUS27	Rh189	41.8% with RhCMV

Table 11. Homology of HCMV, RhCMV and BaCMV Glycoproteins
Comparison of predicted aa identity in ORFs from these three primates shows BaCMV glycoproteins usually share strongest conservation with RhCMV glycoproteins.

Not surprisingly, the HCMV glycoproteins lacking homologs in RhCMV (gpUL04, gpUL16 and gpUS10) are absent in BaCMV as well.

Glycoprotein B is an important envelope glycoprotein in that it acts as both a ligand and fusion protein, and it is critical for viral entry and cell-to-cell spread (Singh and Compton 2000). Furin protease cleaves proteins at a specific recognition motif, RXK/RR (Hosaka et al. 1991). Typically, gB homologs contain this motif since the gB protein undergoes proteolytic processing possibly using cellular furin protease or a similar enzyme. This recognition cleavage site is found in a number of alpha-herpesviruses and virtually all known beta-herpesviruses. The furin protease cleavage site is located between aa 451-454 in the BaCMV UL55 protein (Figure 45). This aa sequence is RTKR which aligns directly with the HCMV RTRR cleavage site. The K to R aa alteration is a conservative change since both aa are basic and share similar properties. Overall homology of BaCMV gB with HCMV is 56.8% and 73.5% with RhCMV.

```
HCMV UL55 FETSGGLVVFWQGIKQKSLVELERLANRS--SLNITHRTRRSTSDNNTTHLS--SMESVHNLV
CCMVgp057 FETTGGLVVFWQGIKQKSLADLEDFAKIQNETLERMRNGTRHKRDTSTSQLSYNSTDVIHNVM
Rh89      YETTGGLIVFWLPVKEKSILEMEELAVAY---NNTNSSTRRKRSTDSASDSNKTSEEVLKSIV
BaUL68    YETTGGLIVFWLPVKEKSIWDMKKLADGYANITNATTSSRTKRSTDDSTLLNNT--EVIHNIV
```

Figure 45. Alignment of Glycoprotein B. Black bar shows furin protease cleavage site for human, rhesus and baboon cytomegalovirus genes. This corresponds (in this alignment) to amino acids 456-459 for HCMV, 456-459 for CCMV, 370-373 for RhCMV and 451-454 for BaCMV.

Three glycoproteins, gH/gL/gO, form a tripartate complex linked to membrane fusion, cell-to-cell spread and virion maturation in HCMV (Huber and Compton 1998; Theiler and Compton 2001). For the BaCMV glycoprotein H homologue, identity is shared at 48.9% with HCMV (UL75) and 64.7% with RhCMV. The BaCMV homologue for HCMV glycoprotein L (UL115) has two presumed N-glycosylation sites (NXS/T with X being any residue) between aa 21-24 and 57-60 (Figure 46).

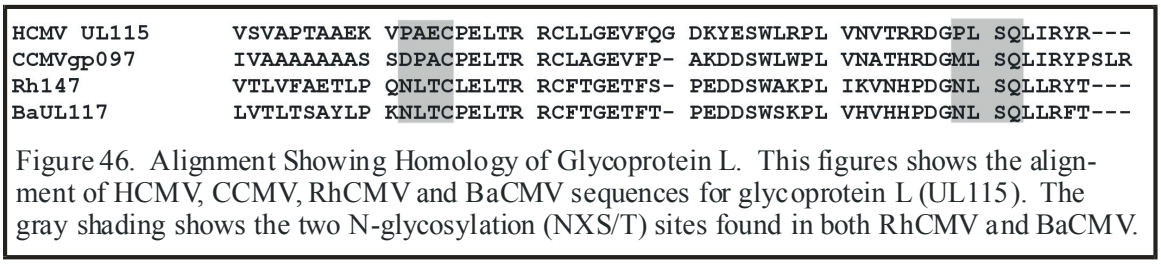


Figure 46. Alignment Showing Homology of Glycoprotein L. This figures shows the alignment of HCMV, CCMV, RhCMV and BaCMV sequences for glycoprotein L (UL115). The gray shading shows the two N-glycosylation (NXS/T) sites found in both RhCMV and BaCMV.

RhCMV gL also has two glycosylation sites, so it shares greater homology with the BaCMV glycoprotein than that of HCMV which only has one. This BaCMV aa sequence shares 47.1% identity with HCMV and 67.7% with RhCMV. In the low-passaged HCMV clinical strain TR, gO has been shown to act as a chaperone protein that promotes incorporation of gH/gL into virions without being present therein itself (Ryckman et al. 2009).

In HCMV, glycoproteins gM and gN are joined by a disulfide bond to form a complex which is transported to the viral assembly compartment (Mach et al. 2005). The cytoplasmic portion of glycoprotein gM contains trafficking signals (which certainly play a role in transport) needed for viral assembly and replication. One of the trafficking motifs consist of a tyrosine-based motif (YXXL) and the other is a cluster of 12 acidic aa (EEEDDDDDEDFED) (Krzyzaniak et al. 2007). The tyrosine motif is found in both RhCMV and BaCMV, but the acidic aa clusters are very different from that of HCMV. The BaCMV gM only contains only 5 acidic aa residues and the RhCMV genome only contains 3. The significance of this difference is uncertain. The BaCMV counterpart of gM shares 54.9% identity with HCMV gM and 70.8 % with RhCMV gM. For glycoprotein gN, modification of the C-terminus by palmitic acid is required for the protein to function in secondary envelopment and replication of HCMV (Mach et al.

2007). The BaCMV homologue of glycoprotein gN is 27 aa shorter than HCMV and RhCMV but still shares 26.3% identity with HCMV and 52.6% with RhCMV.

Glycoprotein gO (UL74) homologs are found in beta-herpesviruses but not other herpesviruses (Huber and Compton 1998). Comparison of glycoprotein gO homologs show they each contain a single hydrophobic domain that functions as a cleavable signal peptide as in HCMV UL74 (Theiler and Compton 2001). This domain is located at the N-terminus and spans aa 14-34 in HCMV, 5-20 in RhCMV and 6-19 in BaCMV (Figure 47A). Also, as seen in other CMVs, the BaCMV gO sequence has a number of N- and O-linked glycosylation sites as would be expected in an envelope glycoprotein. Also, the BaCMV UL74 gene product contains a bipartite NLS, RKQKPERSA-QKTKKR (Figure 47B),

```
HCMV UL74 MGRKEMMVRDVPKMVFLISISFLLVSFIN---CK---VMSKALYN
Rh103     MG---SEIKMWCIWLVIISCSYLTCSK-----KK----KAKKTTY
BaUL81    MG---SVTIKWYALLIIIIISELADTNNTSTTKKPATSKTSTSKT
A.

HCMV UL74 SRNLFRVPKYINGTKLKNTMRKLKRKQAPVKEQFEKKAKKTQSTTT
Rh103     SRNLFKFPH-ITLWKLKQGMRRLHRRRKP-KK--NNTQKRSRRSAN
BaUL81    SRNLFRAPK-ISLWKLKQGMRRLKRKQKPERESAQKTKKRNKRSNN
B.
```

Figure 47. Alignment Showing Homology of Glycoprotein O. A. Locations of hydrophobic cleavable signal peptide sequences are shown in gray boxes. B. Shows a bi-partite nuclear localization signal.

located between aa 252 and 267. BaCMV shares 35.6 % aa sequence identity with HCMV and 59.5% with RhCMV.

D. Immunomodulatory Proteins.

There are 13 reported immunomodulatory proteins found in HCMV that are also present in BaCMV OCOM4-37. Of the 13, several share considerable identity with

HCMV homologs including BaUL112 (UL111), BaUL129 (UL144), BaUS32 (US2), BaUS31, (US3) and BaUS27 (US11). Functions of HCMV UL111 and UL144 are well characterized and discussed below.

Immunomodulatory proteins that show high identity with RhCMV homologs, but somewhat lower identity with HCMV, include UL36, UL37 and the UL118/119 spliced transcript. The HCMV UL36 gene acts as a cell death suppressor, designated vICA (Skaletskaya et al. 2001). Cell death suppression is probably an anti-apoptotic event necessary to avoid immune clearance of virally-infected cells by cytotoxic T or natural killer cells. There are two homologs in BaCMV (BaUL49 and BaUL50) for the HCMV UL36 gene. These two homologs may actually require formation of a spliced transcript to be functional. The product of the HCMV gene UL37 plays a similar role during infection as it encodes an inhibitor of apoptosis, denoted vMIA, which operates by a different mode of action than vICA (Goldmacher et al. 1999).

For the HCMV UL37 gene, there are also two homologs in BaCMV; BaUL51 and BaUL53. Duplicate transcripts for the HCMV UL37 gene are seen in RhCMV as well (Rh62 and Rh66). The HCMV large spliced transcript, UL118/119 shares good identity with BaUL118 and BaUL119. In HCMV, these spliced transcripts are known to encode a Fcγ receptor homolog, vFcγR (Atalay et al. 2002). Other viruses that have similar Fc-binding capacity are able to block antiviral activities such as complement activation or ADCC, both of which are mediated by the Fc domain of IgG (Dubin et al. 1991), BaUS29 shows identity with the HCMV US6 family of genes; however, no HCMV homologue has been identified. BLAST analysis also identified BaUL112 as an interleukin-10 (IL-10) homologue that shares good identity with the RhCMV Rh143

counterpart, but identity with the HCMV UL111 IL-10 sequence is much lower. In RhCMV, the IL-10 gene has been expressed in tissue culture and also in RhCMV-infected rhesus macaques. Based on conservation of this gene in primates, it is hypothesized to play a role in primate CMV persistence and pathogenesis (Lockridge et al. 2000). The BaUL129 gene is homologous with the HCMV Toledo strain tumor necrosis factor receptor (TNFR) gene, UL144. This gene product may be involved in preventing apoptosis of virally-infected cells. HCMV UL144 activates NF-κB (nuclear factor kappa B) at early times of infection and induces expression of CCL22. NF-κB exhibits anti-apoptotic activity (De Smaele et al. 2001) and CCL22 (chemokine ligand 22) is a TH_2 chemo-attractant which inhibits the TH_1 immune response (Nakayama et al. 2004). Four HCMV genes that function in downregulation of major histocompatibility complex class 1 (MHC1) molecules are US2, US3, US8 and US11. The BaCMV homologues of these genes are BaUS32, BaUS31, BaUS30, and BaUS27 respectively.

An additional immunomodulatory protein sequence found in BaCMV (gene BaUL14) is a homologue of the primate COX-2 protein. In RhCMV, there are six exons designated for the COX-2 gene and it is proposed these are spliced transcripts. Analysis of the BaCMV genome revealed all six exons for the COX-2 gene and showed that each of these were collinear with the RhCMV sequence. However, the primate COX-2 sequence that is most identical to that of the BaCMV sequence was from the Cercopithecine herpesvirus strain Colburn (Figure 48). The BaCMV vCOX-2 gene shows good identity with the human COX-2 aa sequence (about 55%) and with the RhCMV aa sequence (about 66%).

```
Identities = 20/33 (60%), Positives = 23/33 (69%), Gaps = 0/33 (0%)
Frame = +1/-2

BaUL14  211   AFSNLSYYTITPFLVPRDRLNPFGVKGE*TSLA  309
              AFSNLSYYT T   VPRD   P GVKG+ + +A
COX-2  10552  AFSNLSYYTRTLPPVPRDCPTPLGVKGDSSPIA  10454

Identities = 13/20 (65%), Positives = 14/20 (70%), Gaps = 0/20 (0%)
Frame = +3/-1

BaUL14  363   YQVRRYCLIRHRLWKKLFLG  422
              +QVRR C IRHRLW K   G
COX-2  10397  HQVRRVCPIRHRLWTKFLFG  10338

Identities = 94/118 (79%), Positives = 104/118 (88%), Gaps = 0/118 (0%)
Frame = +2/-1

BaUL14  482   LSVLPLHHQVVLGHIYGETLEKQHQLRLFKDGKMKYQIIDGEVYPPTVKEAQVHMVYEPS  661
              LS + +  QV L HIYGETLEKQH+LRLFKDGKMKYQ+IDGE+YPPTVKE QVHM+Y P
COX-2  10160  LSTVFVLSQVDLSHIYGETLEKQHKLRLFKDGKMKYQVIDGEMYPPTVKETQVHMLYAPE  9981

BaUL14  662   VPEKLRFALGNEMFGLVPGLMVYAIIWLREHNRVCDVLKQEHPEWNDEQLFQITRLII  835
              VP+ LRFA+GNE+FGLVPGLM+YA IWLREHNRVCDVLK EHPEW DEQLFQ TRLII  9807

Identities = 36/63 (57%), Positives = 43/63 (68%), Gaps = 0/63 (0%)
Frame = +2/-1

BaUL14 929   SRGKRIPPALKDTSREASEQAPKKRYRSLNEYKKRFGFKSHKLLNELTGNTVLQFSYHFT  1108
             S G+ IPPAL+   RE  E   K RY+SLNEY+KRF  K ++   ELTGNT   FSY+FT
COX-2  9512  SGGRNIPPALRRVFREGIEHGRKMRYQSLNEYRKRFRLKPYESFEELTGNTTFGFSYNFT  9333

BaUL14 1109  HFF  1117
             HFF
COX-2  9332  HFF  9324

Identities = 76/111 (68%), Positives = 87/111 (78%), Gaps = 0/111 (0%)
Frame = +3/-1

BaUL14  1185  KEIAVELEAFYGDIEAVELYAGFLAEKPCPDAILDEGVLEPGAPFSLRGLAANVICSPGY  1364
              KEIA  LEA YGD+EAVELY GF+ EKP P AI  E ++E GAPFSL+GL ANVICSP Y
COX-2  9248   KEIAAGLEALYGDVEAVELYTGFIVEKPRPGAIFGESIMELGAPFSLKGLMANVICSPAY  9069

BaUL14 1365  WKRNTVGSSAGFNIVRSATIQSLICSNVKGCPLAAFRAPNKELMTALNGSS  1517
             WK +T G   G +IV++ATIQSLIC+NVKGCPLAAFR  N EL+ A NGSS
COX-2  9068  WKPSTFGGDVGLDIVKTATIQSLICANVKGCPLAAFRVSNAELLKAFNGSS  8916

Identities = 23/56 (41%), Positives = 37/56 (66%), Gaps = 0/56 (0%)
Frame = +2/-3

BaUL14 1649  TLHIAIQGHGVYVRQPLERTAAVIIAEFMINT*VQEIQQYG*DNHSGREGRSDNKQ  1816
             +L + +  +GV+V Q LE TAAVIIA+F++   V E+Q YG +N + R    ++K+
COX-2  8829  SLGLNVNRYGVFVGQSLEGTAAVIIAKFLVRVQVYEVQCYGQNNTNSR*HCGEDKE  8662
```

Figure 48. Six Exons for the BaCMV Homologues of SCMV Coburn COX-2 Gene Identities are percent identical match between aa, positives include conserved changes and gaps are missing aa. Frame indicates which reading frame the sequence is located within. Data acquired from BLASTX (Basic Local Alignment Search Tool), at NCBI (National Center for Biotechnology Information) website.

E. Gene Families Found in BaCMV.

Gene families are groups of genes sharing sequence homology and usually cluster together. The proteins encoded by these genes may also have similar biochemical functions. Gene families are often conserved across different species. There are eight different gene families in HCMV that are also present in BaCMV, as shown in Table 12.

BaCMV gene family members	Characteristics of family*	HCMV family	References
BaUL5, BaUL20 - BaUL29	Genome location, structure, glycosylation sites	RL11	(Chee et al. 1990)
BaUL39 and BaUL48	Genome location, structure	UL25	(Rawlinson et al. 1996)
BaUL46, BaUL83, BaUS7 – BaUS11	GPCR motif; 7TM domains, S-T-rich carboxy terminus	7TM	(Vischer et al. 2006)
BaUL86 – BaUL88	Genome location, structure, phosphoprotein	UL82	
BaUS2, BaUS3, BaUS33	Genome location, structure	US1	
BaUS27 – US 32	Genome location, structure, function	US6	(Jones and Muzithras 1991)
BaUS16 – BaUS26	Genome location, structure, 7TM domains	US12	(Lesniewski et al. 2006)
BaUL37, BaUL38, BaUL42, BaUL43, BaUL49, BaUL50, BaUL56, BaUS1, BaUS13 – BaUS15	Two stretches of hydrophobicity, C-terminal acidic residues	US22	(Chee et al. 1990)

Table 12. BaCMV Gene Family Members and Homologous HCMV Gene Families
*GPCR, G-protein coupled receptor; 7TM, seven-transmembrane.

1. RL11 family.

The HCMV RL11 family consists of 9 genes (UL1, and UL4 - UL11) and they are most likely involved in producing membrane glycoproteins (Davison et al. 2003a). BaCMV homologs have been identified for some of these HCMV genes including BaUL5 (UL153), BaUL20 (UL4), BaUL21 (UL6), BaUL25 (UL9), BaUL23, BaUL24, BaUL28 and BaUL29 (four homologues of UL11) and BaUL22 (matches Rh21). None

of these genes are required for growth in cell culture (Mocarski and Kemble 1996) and yet portions of these genes are fairly well-conserved in beta-herpesviruses (Hansen et al. 2003).

2. UL25 Family.

The HCMV UL25 gene encodes a tegument phosphoprotein (Battista et al. 1999) and UL35 encodes two proteins, one of which localizes to the nucleus while the other is packaged in the viral particle (Liu and Biegalke 2002). There are only two BaCMV genes that belong to the UL25 family. These are genes BaUL39 (UL25) and BaUL48 (UL 35), both of which share significant identity with the HCMV and RhCMV homologs.

3. Seven–Transmembrane Family.

Proteins in this family are so-named because they contain seven membrane-spanning domains. In HCMV, there are four of these: US27, US28, UL33 and UL78. The best characterized, US28, has been shown to act as a G-protein coupled receptor (GPCR) in arterial smooth muscle cells (Streblow et al. 2003). BaCMV shares identity with the same HCMV 7-TM genes that RhCMV does, namely, UL33, UL78, and US28 (Penfold et al. 2003). In BaCMV, the homologs of the four RhCMV 7TM genes are BaUS10, BaUS9, BaUS8, and BaUS6 (which correspond to Rh214, Rh215, Rh216 and Rh220 respectively). As is seen in RhCMV, BaCMV has no counterpart to the HCMV US27 7-TM protein.

4. UL82 Family.

This family of HCMV proteins includes the upper matrix protein, [pp71 (UL82)], the lower matrix protein [pp65 (UL83)], and UL84. There is a single homologue for

UL82 in BaCMV (BaUL88) and two copies of the UL83 homologue (BaUL89 and BaUL90). This is analogous with the RhCMV genome which also has two copies of a UL83 homologue (Rh111 and Rh112; Figure 49). These genes probably arose by tandem duplication and encode a 71,000 mw tegument phosphoprotein (pp71) and a 65,000 mw tegument phosphoprotein (pp65) respectively. In HCMV, UL82 (pp71) has been shown

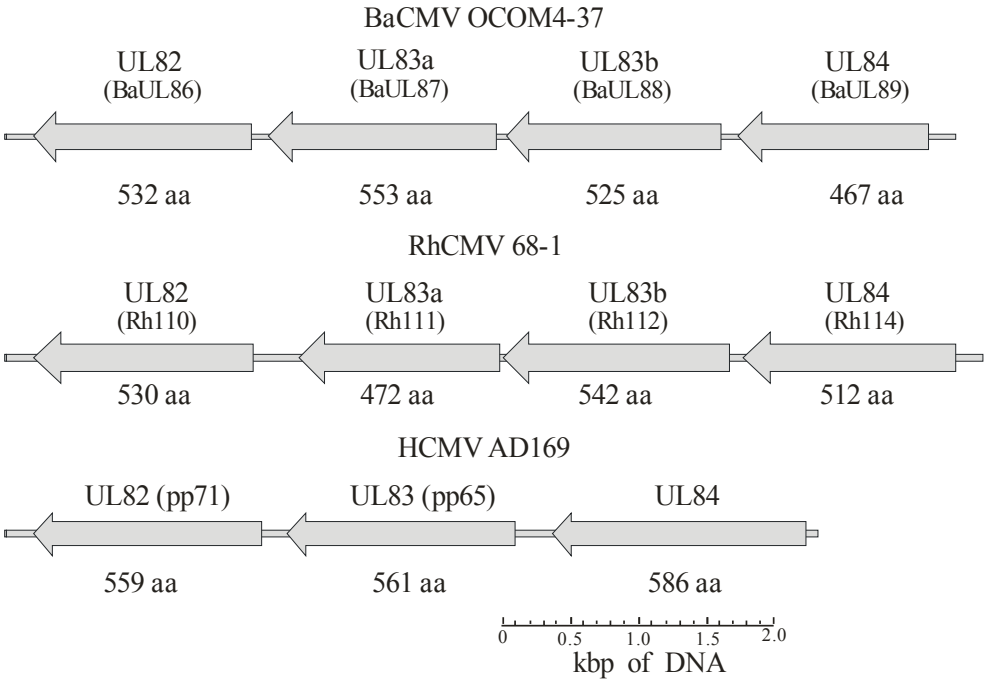


Figure 49. UL82 Gene Family from three Primate CMVs. The orfs are drawn to scale and indicated by wide arrows. The size of the encoded protein is shown below each orf. Based on the HCMV AD169 (X17403), and the RhCMV 68-1 (AF411695) genomic sequences. RhCMV and BaCMV orfs are shown in brackets.

to localize to the nucleus immediately PI (Shen et al. 2008), disrupt the MHCI antigen presentation pathway (Trgovcich et al. 2006), and accelerate the progression of cells through the G_1 phase of the cell cycle (Kalejta et al. 2003). Furthermore, this gene has been observed to interact with the cellular protein, hDaxx, to regulate IE gene expression and viral replication (Cantrell and Bresnahan, 2005; 2006). The HCMV UL83 protein

product (pp65) is also transported to the nucleus early after infection, correlates with virus associated protein kinase activity, is dispensable for growth in cell culture and is required for incorporation of other viral proteins into the viral particle (Chevillotte et al. 2009; Gallina et al. 1996; 1999; Schmolke et al. 1995a; Schmolke et al. 1995b; Yao et al. 2001). NLSs are a necessary component of HCMV tegument proteins (like pp65) that are translocated into the nucleus following penetration. Figure 50 shows three sites where NLSs in HCMV align with homologous protein sequences in RhCMV and BaCMV. In addition to some of the previously described characteristics, pp65 has been observed to elicit a humoral immune response in lupus patients and autoimmune-prone mice (Chang et al. 2006) and to inhibit antiviral gene expression in virally-infected human fibroblast cells (Browne and Shenk, 2003).

When the predicted aa sequence of the UL82 family genes are examined, BaCMV is more closely related to RhCMV. The BaCMV UL82 orf shares 37.2% aa sequence identity with HCMV UL82 and 59.6% with Rh110. BaCMV UL83a and UL83b share 32.7% and 38.9% identity, respectively, with the HCMV counterpart and 39.2% with each other. Identity with BaCMV UL83a and Rh111 is 47.9% and with BaCMV UL83b and Rh112 is 51.2%.

5. US01 Family.

HCMV has three members in this family (US1, US31 and US32) and BaCMV has homologs for each of these (BaUS33, BaUS3 and BaUS2). The BaCMV genes are colinear with the homologous genes of both HCMV and RhCMV (Rh181, Rh225 and and Rh226) and share significant identity with them.

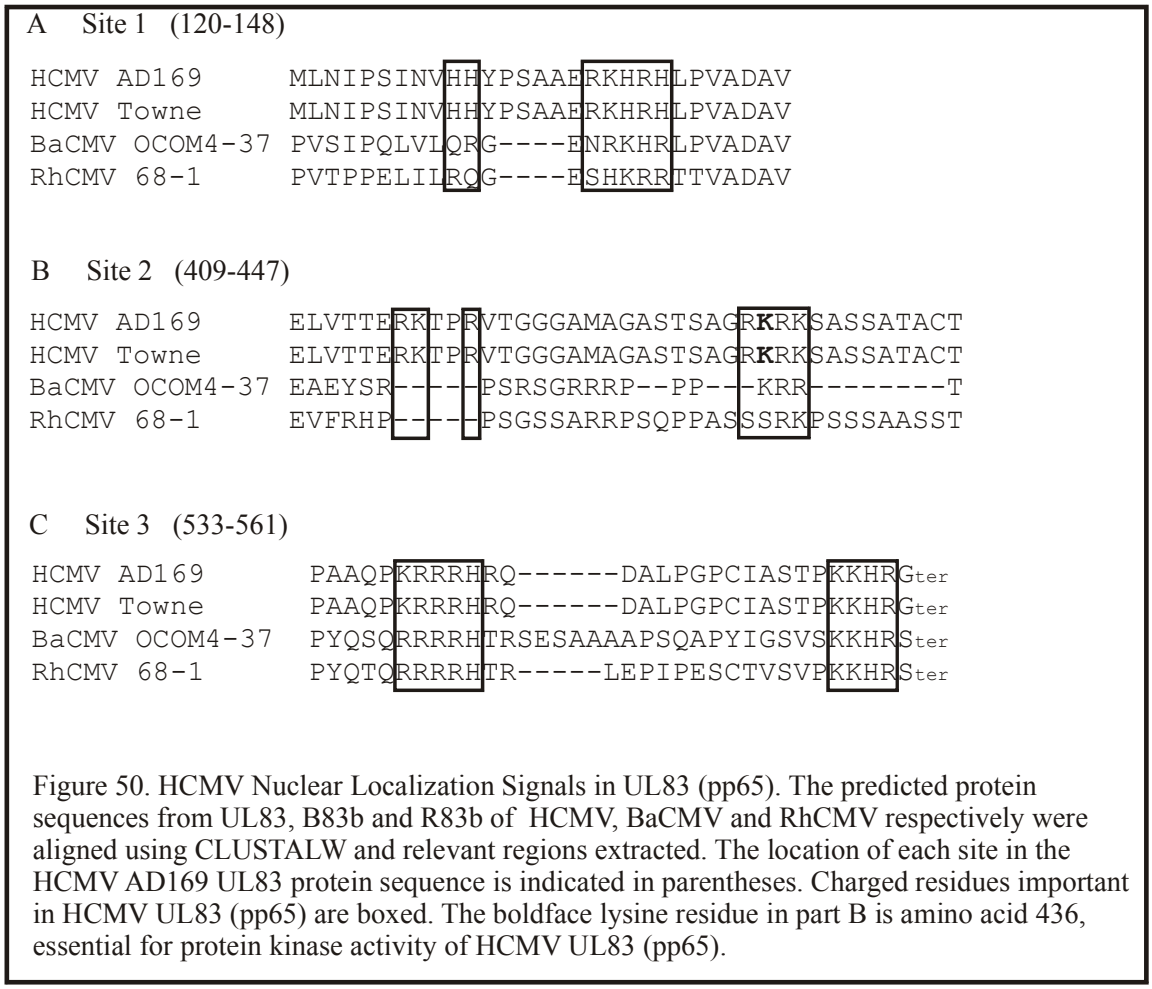


Figure 50. HCMV Nuclear Localization Signals in UL83 (pp65). The predicted protein sequences from UL83, B83b and R83b of HCMV, BaCMV and RhCMV respectively were aligned using CLUSTALW and relevant regions extracted. The location of each site in the HCMV AD169 UL83 protein sequence is indicated in parentheses. Charged residues important in HCMV UL83 (pp65) are boxed. The boldface lysine residue in part B is amino acid 436, essential for protein kinase activity of HCMV UL83 (pp65).

6. US06 Family.

Most of the HCMV US06 family members (US2, US3 and US6 – US11) encode glycoproteins which are structurally and functionally analogous. Five ORFs in BaCMV belong to the US06 family: BaUS2, BaUS3, BaUS6, BaUS8 and BaUS11. In HCMV, US6 has been shown to elicit downregulation of MHC class I molecules through inhibition of the cellular transporter associated with antigen processing (Dugan and Hewitt 2008; Jun et al. 2000). Another HCMV US06 family protein (US2) facilitates degradation of MHC class I heavy chains through utilization of chaperone proteins (Oresic and Tortorella 2008). HCMV US3 and US8 have both been shown to transiently bind MHCI molecules (Gruhler et al. 2000; Tirabassi and Ploegh 2002). HCMV US11

reportedly diverts MHCI molecules in mouse embryonic fibroblasts from the ER to the cytosol where MHCI degradation ensues (Tirosh et al. 2005). Although it is customary to include the HCMV genes US2 and US3 in the US02 gene family, the homologous BaCMV genes are more similar to those of the HCMV US06 family so they are included here.

7. US12 Family.

The HCMV US12 family consists of 10 tandemly-arranged genes spanning the range from US12 to US21. All proteins in the HCMV US12 family are putative seven-transmembrane domain proteins that are related to G-coupled receptors (Lesniewski et al. 2006) and some are proposed to play a role in virion maturation and egress based on their association with the viral assembly compartment (Das and Pellett, 2007). Analysis of the BaCMV US region revealed the same 11 genes that are found in RhCMV for this family. Four of these are duplicates of the HCMV US14 ORF. The remaining seven share homology with the following HCMV genes: US12 (BaUS26), US13 (BaUS25), US17 (BaUS20), US18 (BaUS19), US19 (BaUS18), US20 (BaUS17) and US21 (BaUS16).

8. US22 Family.

Proteins of the HCMV US22 family have two stretches of hydrophobic and charged residues and as many as four conserved motifs that are common to beta-herpesviruses. The functions of most members of this gene family are unknown, but two MCMV homologs of the HCMV US22 family (m140 and m141) have been shown to affect MCMV replication in macrophages (Hanson et al. 2009; Menard et al. 2003). In HCMV, there are at least two proteins in the UL22 family for which functions have been established *in vitro*. These include the HCMV UL23 protein product which acts as a

temperance (or repressing) factor for HCMV replication in HFF cells (Dunn et al. 2003). The other is theTRS1 protein which is able to counteract host cell antiviral responses (Child et al. 2004). In BaCMV there are 10 ORFs that share homology with both HCMV and RhCMV genes. These include BaUL37, BaUL38, BaUL42, BaUL43, BaUL49, BaUL50, BaUL56, BaUS1, BaUS14 and BaUS15. They are expected to share similar function, whatever that function may be.

F. Phylogenetic Review.

Analysis of the phylogenetic relationship between BaCMV and other herpesviruses is presented in Figures 51 and 52. Alignments of the predicted aa sequences from six ORFs that are conserved in herpesviruses were performed using CLUSTALW and then used to construct phylogenetic trees (Figures 51 and 52). These ORFs were BaUL67 (DNA polymerase), BaUL68 (glycoprotein B), BaUL75 (helicase), BaUL91 (major capsid protein), BaUL70 (single-stranded binding protein), and BaUL114 (uracil N-glycosylase) gene sequences. Herpesviruses included in these analyses were HCMV, RhCMV, CCMV, RatCMV, MCMV, KSHV and HSV-1. Using these eight strains provided comparison between alpha-, beta- and gamma-herpesvirus families. These strains were chosen for comparison because they have completely-sequenced genomes. Trees were rooted at the mid-point between the two most divergent operational taxonomic units (i.e. alpha-herpesvirus HSV-1 and gamma-herpesvirus KSHV). All of these trees indicate that the most closely-related virus to BaCMV is RhCMV. The next closest relation is CCMV, followed by HCMV, MCMV, RCMV, KSHV and HSV-1 respectively. The drill monkey CMV (DrCMV) is actually more-closely related to BaCMV than any of the eight strains used for comparison (Blewett et

al. 2003). Unfortunately very little sequence data exists for DrCMV so it was not included in the trees. However, when the aa sequences that are available for DrCMV were compared with three primate herpesviruses (RhCMV, BaCMV and HCMV) then DrCMV shares greatest identity with BaCMV.

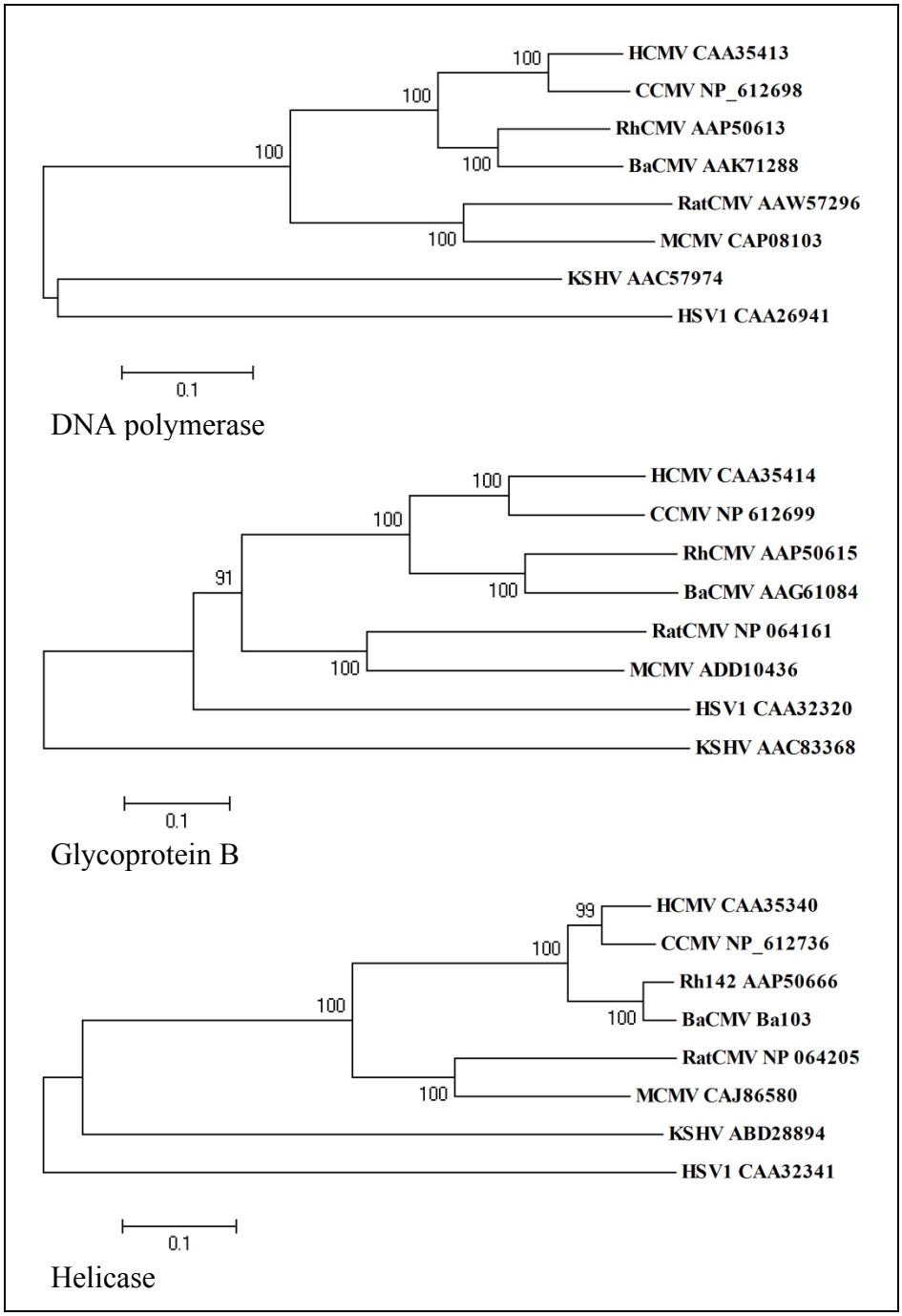


Figure 51. Phylogenetic Relationships for DNA Polymerase, Glycoprotein B and Helicase Proteins. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl 1965) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura 2007). Accession or protein ID numbers are provided for published DNA data.

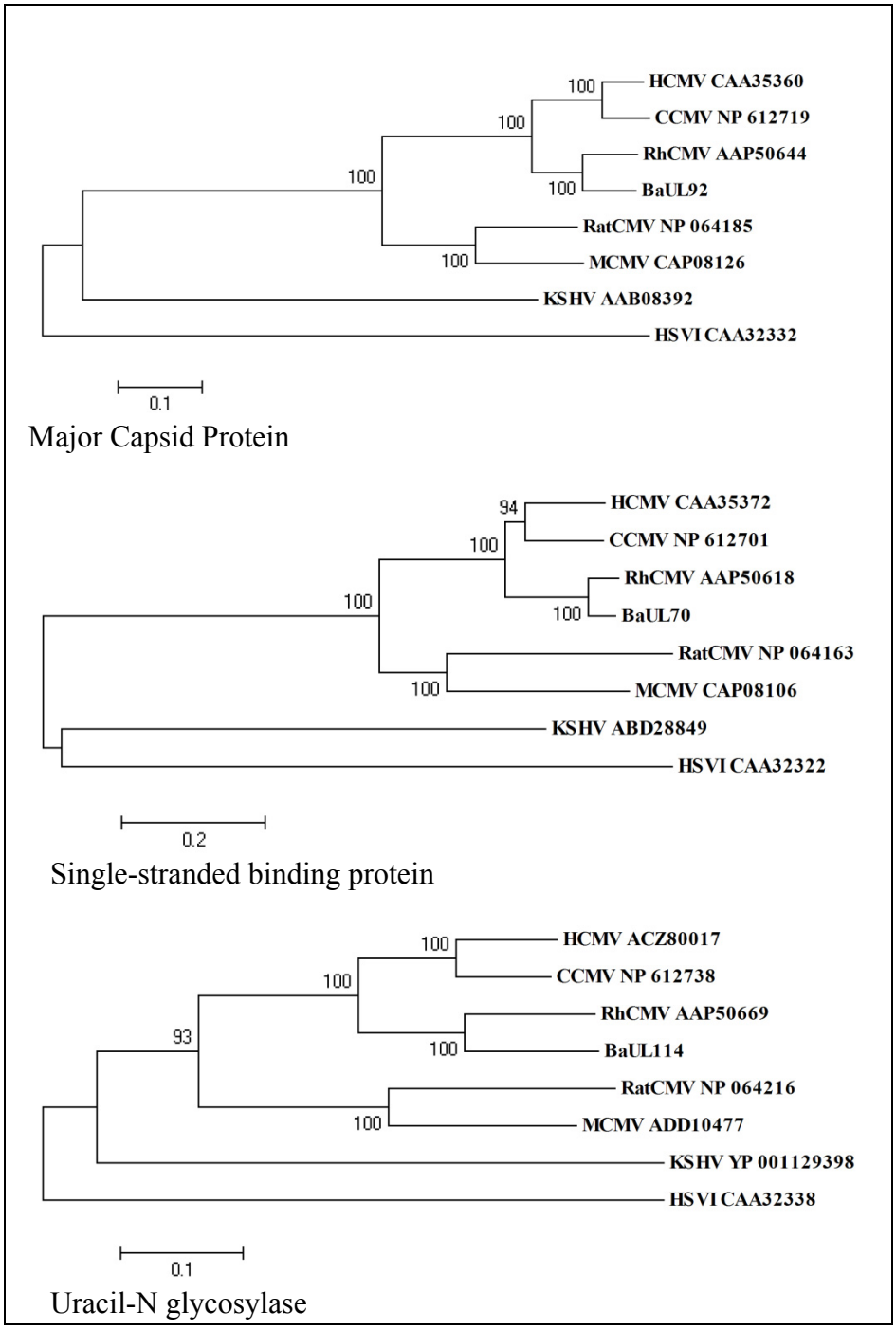


Figure 52. Phylogenetic Relationships for the Major Capsid Protein, Single-Stranded Binding Protein and Uracil-N Glycosylase. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein 1985). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl 1965) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset. Phylogenetic analyses were conducted in MEGA4 (Tamura 2007). Accession or protein ID numbers are provided for published DNA data.

CHAPTER FIVE

CONCLUSIONS

Over 217,000 bp of the BaCMV OCOM4-37 genome were sequenced. The genome is projected to contain between 221,000 and 229,000 bp and to potentially encode over 170 proteins. Homologues of genes found in other herpesviruses have been identified in BaCMV within the UL and US sequences. There is considerable homology between the BaCMV ORFs and those of other animal CMVs. In most respects, the BaCMV genome is more similar to RhCMV than to either HCMV or CCMV. Some general similarities between the RhCMV and BaCMV genomes are (1) a lack of extensive repeats between the UL and US regions, (2) genes that are homologous to those found in clinical strains of HCMV but not the prototype strain AD169, and (3) duplications of the UL83 gene. One difference between the BaCMV genome and those of other sequenced animal CMVs is that the entire US region is inverted.

BLAST searches were conducted using BaCMV ORFs, and those matching published CMV genes were considered authentic. Highest matches were with either RhCMV or, in a few instances, African green monkey CMV genes. The next highest matches were with genes from CCMV and then HCMV. The overall G+C content for HCMV is 57% as compared to 49% for RhCMV. For the part of the BaCMV genome sequenced thus far, the

G+C content is 48%. As expected, this number is more similar to the RhCMV G+C ratio than HCMV.

Sequence analysis of the BaCMV genome showed that partially-sequenced homologues of HCMV IE genes are present in BaCMV. These include IE2, UL36 – UL38, US3, TRS1, and a portion of IE1. The BaCMV IE1 gene sequence is incomplete because it has not been completely sequenced. In HCMV, the IE proteins are produced very early in the infection cycle and function as regulatory proteins, and so are likely to do so in BaCMV as well (Colberg-Poley, 1996; Colberg-Poley et al. 1992). A common feature seen both in BaCMV and RhCMV is duplication of the HCMV UL37 gene homologue.

There are additional BaCMV genes that show identity with HCMV gene sequences and may function as regulatory proteins. One of these is BaUL121 which shares 26% identity with HCMV UL121 and 52 % identity with Rh155. In BaCMV, this sequence contains a highly-charged region at the C-terminus between aa 167 - 184. This charged terminus is characteristic of some *trans*-acting transcriptional activators, and may indicate a regulatory role for this gene. Another HCMV transactivator (UL69) shares 43% identity with BaUL73 homologue while the Rh97 gene shares 58% identity with BaUL73. The UL69 gene in HCMV is homologous to the HSV IE ICP27 gene which is essential for transition from E to L gene expression (Winkler et al. 1994). In summary, all HCMV regulatory and transactivator proteins found in BaCMV are well-conserved when compared with both HCMV and RhCMV genes.

Genes that code for enzymes in HCMV are also found in BaCMV. The gene sequence encoding BaCMV DNA polymerase (BaUL67) has been previously submitted

to GenBank (accession #AAK71288). The BaUL105 ORF is the homologue of HCMV UL97, which exhibits protein kinase activity. UL97 mediates phosphorylation events during DNA replication, capsid maturation, nuclear egress, and cytoplasmic assembly (Azzeh et al. 2006; Krosky et al. 2003). Additionally, the UL97 protein phosphorylates, and thereby activates, the antiviral drug ganciclovir (Littler et al. 1992). In many herpesviruses, prodrug phosphorylation is accomplished by the enzyme thymidine kinase. However in HCMV and other CMVs (including BaCMV) a thymidine kinase gene has not been identified. Other important ORFs found in BaCMV are homologs for HCMV DNA helicase-primase (UL70) and two DNA repair enzymes, dUTPase (UL72) and DNase (UL98). Each of these genes is well-conserved among HCMV, RhCMV and BaCMV sequences (see Table 6). The HCMV UL77 gene codes for the pyruvoyl decarboxylase enzyme which is involved in cellular polyamine synthesis (Yoakum 1993); a critical process in HCMV replication. This essential gene (Kalejta 2008) is unique to beta-herpesviruses and has also been identified in MCMV (Rawlinson et al. 1996), and Rat CMV (Vink et al. 2000) in addition to the primate CMVs.

Structural proteins are also highly conserved between HCMV and BaCMV. These include the major and minor capsid proteins (BaUL91 and BaUL92, respectively) and the capsid assembly protein (BaUL85). Tegument proteins are other structural proteins that are well-conserved, and BaCMV has homologs of most HCMV tegument genes. Some of the tegument proteins in HCMV are UL32 (pp150), UL48 (the large tegument protein), UL82 (pp71), UL83 (pp65) and UL99 (pp28). As previously mentioned, both BaCMV and RhCMV have duplicate copies of the HCMV AD169 UL83 gene. In BaCMV, these duplicates are distinguished as UL83a and UL83b. Of the two,

the UL83b gene shows the greatest homology with HCMV UL83 (38.9%). In RhCMV, these two genes are designated Rh111 (BaUL87) and Rh112 (BaUL88). As seen in BaCMV, the RhCMV Rh112 gene shares highest identity with the HCMV gene. Immunological analysis of Rh112 has demonstrated that this protein shares antigenic sites, or epitopes, with HCMV pp65 (Yue et al. 2006). The BaCMV (BaUL88) protein shares 51.2% aa identity with Rh112, so it is likely that immunological analysis of the BaCMV protein would produce a similar outcome. Support for this was seen when pooled baboon sera from BaCMV-positive baboons were used to screen a lambda BaCMV genomic expression library. Though many positive clones were obtained, >75% of these clones were the UL83b gene. No clones expressing UL83a were recovered, suggesting that the UL83b gene is a much more important target of the baboon humoral immune response in natural infections.

HCMV strain AD169 lacks numerous genes when compared to clinical HCMV strains such as Towne and Toledo (Cha et al. 1996). To determine whether duplicates of UL83 exist in these clinical strains, PCR amplification of the UL82 – UL84 region was performed. PCR analysis showed that neither HCMV strains Towne or Toledo contained this extra gene (Blewett, unpublished observations). Duplication of the UL83 gene may have developed during evolution of the virus or the genes may have different functions (Hansen et al. 2003).

The HCMV genome contains a number of gene families consisting of members having functional or structural similarity. Nine families have been identified in HCMV and representatives for each of these gene families are present in BaCMV. Eight of these families are listed and characterized in Table 6. The HCMV US2 and US6 family

members have similar structure and function as downregulators of MHCI expression on cell surfaces (Pande et al. 2005). Due to these similarities, the BaCMV US2 homologs have been included in the US6 family.

There are at least 22 genes in a 13 kb region of sequence (designated the UL/b' region) that are found in certain clinical strains of HCMV like Toledo or Towne, but not in HCMV AD169 (Prichard et al. 2001). These genes include UL133 – UL154 and many of these are hypothesized to be transmembrane glycoproteins or to function as chemokines (Rigoutsos et al. 2003). In BaCMV, homologues of seven of these genes (BaUL05, BaUL126, BaUL129, BaUL130, BaUL131, BaUL132, and BaUL134) have been identified and all show high homology (see Table 10) with either RhCMV or clinical HCMV strains. RhCMV has an additional gene in this region related to the HCMV UL148 gene. Consequently, an eighth BaCMV gene related to HCMV UL148 may exist since this region is in one of the sequence gaps adjacent to BaUL126.

For CMV to survive and proliferate within a host cell, the viral genome must code for a number of immunomodulatory proteins. The proteins encoded by these genes usually function in evasion of host cell defenses. There are a number of genes in BaCMV that share identity with HCMV immunomodulatory genes and probably function in that capacity. One of these genes is an IL-10 homologue which shows good aa sequence identity with the RhCMV (Rh143) homologue but no significant sequence homology with the HCMV (UL111A) gene homolog. The exact function of the RhCMV Rh143 gene product has not been determined, but in HCMV there is evidence that the vIL-10 protein inhibits dendritic cell maturation and function (Chang et al. 2004). There is also a BaCMV homologue for a TNF receptor (BaUL131). Since the HCMV TNF

(UL144) suppresses apoptotic progression in virally-infected cells, this would be a likely function for the BaUL131 encoded protein if it is actually expressed and functions as a TNF receptor.

Virally-infected cells often undergo apoptosis as a result of the infection. Therefore, to survive, viruses must inhibit processes leading to apoptosis. HCMV encodes two apoptosis inhibitors vICA (UL36) and vMIA (UL37). vICA acts to inhibit apoptosis by suppressing caspase-8 activation. Caspase-8 protease is necessary for inducing apoptosis and also controlling differentiation of monocytes to macrophages (Droin et al. 2008; Lamkanfi et al. 2006; Rebe et al. 2007). Mutation analysis of the HCMV UL36 gene has shown that vICA also controls a caspase-independent cell death pathway at late stages of macrophage differentiation; caspase-dependent cell death responses occurred in early stage of macrophage differentiation (McCormick et al. 2010). In BaCMV, there are two putative genes (BaUL49 & BaUL50) that code for an HCMV UL36 homologue. The first gene (BaUL49) is considerably larger than the second (BaUL50) and shows lower identity with HCMV homlogs. This distinction exists for RhCMV as well. It is uncertain whether these genes are spliced transcripts or function separately. The HCMV UL37 product (vMIA) is another immunomodulatory protein that is targeted to mitochondria where it is capable of inhibiting the proapoptotic Bcl-2 family members Bax and Bak (Norris and Youle 2008). As seen with UL36, there are two homologs (BaUL51 and BaUL53) for HCMV UL37 in BaCMV, but their functions are unknown. While the functions of BaCMV homologs for the UL36 and UL37 genes have not been examined, similar cell- death suppressor proteins have been

identified in MCMV (Cam et al. 2010; Manzur et al. 2008) and it would follow that all CMVs must express these proteins to complete the infection cycle.

A viral cyclooxygenase-2 (vCOX-2) homologue has been identified in RhCMV and is probably also important in immunomodulation. A homologue for this gene, however, has not been identified in HCMV. This may explain why in HCMV, cellular COX-2 inhibitors compromise viral replication *in vitro* (Tanaka et al. 1988; Zhu et al. 2002). Also, COX-2 inhibitors also block cell-to-cell spread in HCMV which makes them possible candidates for control of HCMV infection (Schrőer and Shenk 2008). Cellular COX-2 is an enzyme that is important in the eicosanoid synthesis pathway. Eicosnaoids are a group of cytokines that regulate various immune responses. Many viruses upregulate the eicosanoid pathway early in infection and this increases COX-2 which is required for accumulation of the transcriptional regulator IE2. In RhCMV, studies have shown that when viral COX-2 is expressed during infection, cellular COX-2 is not, and that COX-2 deletion mutants do not replicate normally in endothelial cells (Rue et al. 2004) The function of this protein has not been established in BaCMV however the degree of identity it shares with the human COX-2 gene would suggest a similar function.

G protein-coupled receptors (GPCR), also known as seven-transmembrane receptors (7TM), are a large family of proteins found in cell plasma membranes. When these receptors are bound by a ligand, they function in signal transduction leading to regulation of intracellular activities (Bourne et al. 1990). There are three genes (UL33, US27 and US28) in HCMV that are homologous with cellular GPCR (Welch et al. 1991). The HCMV US28 GPCR homologue is known to regulate both chemokine dependent

and independent signaling in infected cells via the C-terminal tail of the protein (Stropes et al. 2009). In BaCMV, two homologs exist for the HCMV genes UL33 and US28. For the HCMV US28 gene, there are actually five tandem BaCMV sequences that show identity (average of 30%) with this homolog. These five US28 homologs exist in the RhCMV genome as well, although they are divergent from one another. Characterization of these five genes has shown that one of them (RhUS28.5) shows a ligand-binding profile similar to that of HCMV US28, and analysis of northern blots showed that paired genes RhUS28.1/RhUS28.2 and RhUS28.3/RhUS28.4 are each expressed as a single transcriptional unit (Penfold et al. 2003). No functional capacity has yet been designated for the RhUS28.1-RhUS28.4 genes. The BaCMV homologue of RhUS28.5 (BaUS07) shows greatest aa sequence identity (66.7%) of all the RhUS28 homologs found in BaCMV.

Phylogenetic analysis using eight herpesviruses showed that BaCMV groups with other beta-herpesviruses and is most closely related to those CMVs that infect primates. Based on the limited amount of drill CMV DNA sequence available (fragments of the glycoprotein B and DNA polymerase genes), drill CMV is the herpesvirus most closely related to BaCMV. Of the viruses examined in detail, BaCMV is most similar to RhCMV, and then to CCMV and HCMV. More distantly-related CMVs (such as MCMV and RatCMV) group separately from the primate CMVs. A significant amount of research has been conducted on the structure and function of the HCMV genome and, to a lesser extent, the RhCMV genome. These studies have revealed the location and function of structural genes, the temporal expression of IE, E, and L proteins and functions of proteins that are essential to complete the viral life cycle *in vitro*. Given the

high homology and colinearity between these two primate virus genomes and that of BaCMV, it is reasonable to expect that exploration of OCOM4-37 will reveal similar results.

REFERENCES

Ahmed, M., M. Lock, C.G. Miller, and N.W. Fraser. 2002. Regions of the herpes simplex virus type 1 latency-associated transcript that protect cells from apoptosis in vitro and protect neuronal cells in vivo. *J. Virol.* 76:717-729.

Alting-Mees, M.A., and J.M. Short. 1989. pBluescript II: gene mapping vectors. *Nucleic Acids Res.* 17:9494.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.

Anders, D.G., M.A. Kacica, G. Pari, and S.M. Punturieri. 1992. Boundaries and structure of human cytomegalovirus oriLyt, a complex origin for lytic-phase DNA replication. *J. Virol.* 66:3373-3384.

Andoniou, C.E., and M.A. Degli-Esposti. 2006. Insights into the mechanisms of CMV-mediated interference with cellular apoptosis. *Immunol. Cell. Biol.* 84:99-106.

Appleton, B.A., A. Loregian, D.J. Filman, D.M. Coen, and J.M. Hogle. 2004. The cytomegalovirus DNA polymerase subunit UL44 forms a C clamp-shaped dimer. *Mol. Cell.* 15:233-244.

Atalay, R., A. Zimmermann, M. Wagner, E. Borst, C. Benz, M. Messerle, and H. Hengel. 2002. Identification and expression of human cytomegalovirus transcription units coding for two distinct Fc{gamma} receptor homologs. *J. Virol.* 76:8596-8608.

AuCoin, D.P., G.B. Smith, C.D. Meiering, and E.S. Mocarski. 2006. Betaherpesvirus-conserved cytomegalovirus tegument protein ppUL32 (pp150) controls cytoplasmic events during virion maturation. *J. Virol.* 80:8199-8210.

Azzeh, M., A. Honigman, A. Taraboulos, A. Rouvinski, and D.G. Wolf. 2006. Structural changes in human cytomegalovirus cytoplasmic assembly sites in the absence of UL97 kinase activity. *Virology.* 354:69-79.

Bacchetti, S., M.J. Evelegh, and B. Muirhead. 1986. Identification and separation of the two subunits of the herpes simplex virus ribonucleotide reductase. *J. Virol.* 57:1177-1181.

Baines, J.D., and B. Roizman. 1992. The UL11 gene of herpes simplex virus 1 encodes a function that facilitates nucleocapsid envelopment and egress from cells. *J. Virol.* 66:5168-5174.

Baringer, J.R., and P. Swoveland. 1973. Recovery of herpes-simplex virus from human trigeminal ganglions. *N. Engl. J. Med.* 288:648-650.

Barnard, E.C., G. Brown, and N.D. Stow. 1997. Deletion mutants of the herpes simplex virus type 1 UL8 protein: effect on DNA synthesis and ability to interact with and influence the intracellular localization of the UL5 and UL52 proteins. *Virology*. 237:97-106.

Batterson, W., D. Furlong, and B. Roizman. 1983. Molecular genetics of herpes simplex virus. VIII. Further characterization of a temperature-sensitive mutant defective in release of viral DNA and in other stages of the viral reproductive cycle. *J. Virol*. 45:397-407.

Battista, M.C., G. Bergamini, M.C. Boccuni, F. Campanini, A. Ripalti, and M.P. Landini. 1999. Expression and characterization of a novel structural protein of human cytomegalovirus, pUL25. *J. Virol.* 73:3800-3809.

Bayliss, G.J., H.S. Marsden, and J. Hay. 1975. Herpes simplex virus proteins: DNA-binding proteins in infected cells and in the virus structure. *Virology*. 68:124-134.

Bechtel, J.T., and T. Shenk. 2002. Human cytomegalovirus UL47 tegument protein functions after entry and before immediate-early gene expression. *J. Virol.* 76:1043-1050.

Bender, F.C., J.C. Whitbeck, H. Lou, G.H. Cohen, and R.J. Eisenberg. 2005. Herpes simplex virus glycoprotein B binds to cell surfaces independently of heparan sulfate and blocks virus entry. *J. Virol.* 79:11588-11597.

Billstrom, M.A., and W.J. Britt. 1995. Postoligomerization folding of human cytomegalovirus glycoprotein B: identification of folding intermediates and importance of disulfide bonding. *J. Virol.* 69:7015-7022.

Blewett, E.L., G. White, J.T. Saliki, and R. Eberle. 2001. Isolation and characterization of an endogenous cytomegalovirus (BaCMV) from baboons. *Arch. Virol*. 146:1723-1738.

Blewett, E.L., Lewis, J., Gadsby, E. L., Neubauer, S. R., Eberle, R. 2003. Isolation of cytomegalovirus and foamy virus from the drill monkey (Mandrillus leucophaeus) and prevalence of antibodies to these viruses amongst wild-born and captive-bred individuals. *Arch. Virol.* 148:423-433.

Bloom, D.C., J.G. Stevens, J.M. Hill, and R.K. Tran. 1997. Mutagenesis of a cAMP response element within the latency-associated transcript promoter of HSV-1 reduces adrenergic reactivation. *Virology*. 236:202-207.

Boehme, K.W., J. Singh, S.T. Perry, and T. Compton. 2004. Human cytomegalovirus elicits a coordinated cellular antiviral response via envelope glycoprotein B. *J. Virol.* 78:1202-1211.

Boehmer, P.E., M.S. Dodson, and I.R. Lehman. 1993. The herpes simplex virus type-1 origin binding protein. DNA helicase activity. *J. Biol. Chem.* 268:1220-1225.

Boehmer, P.E., and I.R. Lehman. 1993. Herpes simplex virus type 1 ICP8: helix-destabilizing properties. *J. Virol.* 67:711-715.

Bogner, E., K. Radsak, and M.F. Stinski. 1998. the gene product of human cytomegalovirus open reading frame UL56 binds the pac motif and has specific nuclease activity. *J. Virol.* 72:2259-2264.

Borst, E.M., K. Wagner, A. Binz, B. Sodeik, and M. Messerle. 2008. The essential human cytomegalovirus gene UL52 is required for cleavage-packaging of the viral genome. *J. Virol.* 82:2065-2078.

Bortner, C., T.R. Hernandez, I.R. Lehman, and J. Griffith. 1993. Herpes simplex virus 1 single-strand DNA-binding protein (ICP8) will promote homologous pairing and strand transfer. *J. Mol. Biol.* 231:241-250.

Bourne, H.R., D.A. Sanders, and F. McCormick. 1990. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature.* 348:125-132.

Boyle, K.A., and T. Compton. 1998. Receptor-binding properties of a soluble form of human cytomegalovirus glycoprotein B. *J. Virol.* 72:1826-1833.

Britt, W.J. 1984. Neutralizing antibodies detect a disulfide-linked glycoprotein complex within the envelope of human cytomegalovirus. *Virology*. 135:369-378.

Britt, W.J., and D. Auger. 1986. Synthesis and processing of the envelope gp55-116 complex of human cytomegalovirus. *J. Virol.* 58:185-191.

Britt, W.J., and M. Mach. 1996. Human cytomegalovirus glycoproteins. *Intervirology.* 39:401-412.

Britt, W.J., L. Vugler, E.J. Butfiloski, and E.B. Stephens. 1990. Cell surface expression of human cytomegalovirus (HCMV) gp55-116 (gB): use of HCMV-recombinant vaccinia virus-infected cells in analysis of the human neutralizing antibody response. *J. Virol.* 64:1079-1085.

Britt, W.J., and L.G. Vugler. 1989. Processing of the gp55-116 envelope glycoprotein complex (gB) of human cytomegalovirus. *J. Virol.* 63:403-410.

Brosius, J. 1989. Superpolylinkers in cloning and expression vectors. *DNA*. 8:759-777.

Browne, E.P., and T. Shenk. 2003. Human cytomegalovirus UL83-coded pp65 virion protein inhibits antiviral gene expression in infected cells. *Proc. Natl. Acad. Sci. U. S. A.* 100:11439-11444.

Bush, M., D.R. Yager, M. Gao, K. Weisshart, A.I. Marcy, D.M. Coen, and D.M. Knipe. 1991. Correct intranuclear localization of herpes simplex virus DNA polymerase requires the viral ICP8 DNA-binding protein. *J. Virol.* 65:1082-1089.

Cam, M., W. Handke, M. Picard-Maureau, and W. Brune. 2010. Cytomegaloviruses inhibit Bak- and Bax-mediated apoptosis with two separate viral proteins. *Cell Death Differ*. 17:655-665.

Campadelli-Fiume, G., F. Cocchi, L. Menotti, and M. Lopez. 2000. The novel receptors that mediate the entry of herpes simplex viruses and animal alphaherpesviruses into cells. *Rev. Med. Virol.* 10:305-319.

Cantrell, S.R., and W.A. Bresnahan. 2005. Interaction between the human cytomegalovirus UL82 gene product (pp71) and hDaxx regulates immediate-early gene expression and viral replication. *J. Virol.* 79:7792-7802.

Cantrell, S.R., and W.A. Bresnahan. 2006. Human cytomegalovirus (HCMV) UL82 gene product (pp71) relieves hDaxx-mediated repression of HCMV replication. *J. Virol.* 80:6188-6191.

Caradonna, S., D. Worrad, and R. Lirette. 1987. Isolation of a herpes simplex virus cDNA encoding the DNA repair enzyme uracil-DNA glycosylase. *J. Virol.* 61:3040-3047.

Carlson, C., W.J. Britt, and T. Compton. 1997. Expression, purification, and characterization of a soluble form of human cytomegalovirus glycoprotein B. *Virology*. 239:198-205.

Carrozza, M., and N. DeLuca. 1996. Interaction of the viral activator protein ICP4 with TFIID through TAF250. *Mol. Cell. Biol.* 16:3085-3093.

Cha, T.A., E. Tom, G.W. Kemble, G.M. Duke, E.S. Mocarski, and R.R. Spaete. 1996. Human cytomegalovirus clinical isolates carry at least 19 genes not found in laboratory strains. *J. Virol.* 70:78-83.

Challberg, M.D. 1986. A method for identifying the viral genes required for herpesvirus DNA replication. *Proc. Natl. Acad. Sci. U. S. A*. 83:9094-9098.

Chang, M., M.R. Pan, D.Y. Chen, and J.L. Lan. 2006. Human cytomegalovirus pp65 lower matrix protein: a humoral immunogen for systemic lupus erythematosus patients and autoantibody accelerator for NZB/W F_1 mice. *Clin. Exp. Immunol.* 143:167-179.

Chang, W.L.W., N. Baumgarth, D. Yu, and P.A. Barry. 2004. Human cytomegalovirus-encoded interleukin-10 homolog inhibits maturation of dendritic cells and alters their functionality. *J. Virol.* 78:8720-8731.

Chee, M.S., A.T. Bankier, S. Beck, R. Bohni, C.M. Brown, R. Cerny, T. Horsnell, C.A. Hutchison, 3rd, T. Kouzarides, J.A. Martignetti, and et al. 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.* 154:125-169.

Chen, J., and M.F. Stinski. 2002. Role of regulatory elements and the MAPK/ERK or p38 MAPK pathways for activation of human cytomegalovirus gene expression. *J. Virol.* 76:4873-4885.

Chen, R., H. Wang, and L.M. Mansky. 2002. Roles of uracil-DNA glycosylase and dUTPase in virus replication. *J. Gen. Virol.* 83:2339-2345.

Chen, S.-H., A. Pearson, D.M. Coen, and S.-H. Chen. 2004. Failure of thymidine kinase-negative herpes simplex virus to reactivate from latency following efficient establishment. *J. Virol.* 78:520-523.

Chen, X., M. Schmidt, W. Goins, and J. Glorioso. 1995. Two herpes simplex virus type 1 latency-active promoters differ in their contributions to latency-associated transcript expression during lytic and latent infections. *J. Virol.* 69:7899-7908.

Chen, Y.-M., and D.M. Knipe. 1996. A dominant mutant form of the herpes simplex virus ICP8 protein decreases viral late gene transcription. *Virology*. 221:281-290.

Chevillotte, M., S. Landwehr, L. Linta, G. Frascaroli, A. Luske, C. Buser, T. Mertens, and J. von Einem. 2009. Major tegument protein pp65 of human cytomegalovirus is required for the incorporation of pUL69 and pUL97 into the virus particle and for viral growth in macrophages. *J. Virol.* 83:2480-2490.

Child, S.J., M. Hakki, K.L. De Niro, and A.P. Geballe. 2004. Evasion of Cellular Antiviral Responses by Human Cytomegalovirus TRS1 and IRS1. *J. Virol.* 78:197-205.

Chou, S. 1992. Comparative analysis of sequence variation in gp116 and gp55 components of glycoprotein B of human cytomegalovirus. *Virology*. 188:388-390.

Chou, S., and C.L. Meichsner. 2000. A Nine-Codon Deletion Mutation in the Cytomegalovirus UL97 Phosphotransferase Gene Confers Resistance to Ganciclovir. *Antimicrob. Agents Chemother.* 44:183-185.

Chou, S., Rachel H. Waldemer, Anne E. Senters, Kevin S. Michels, George W. Kemble, Richard C. Miner, and W.L. Drew. 2002. Cytomegalovirus UL97 phosphotransferase mutations that affect susceptibility to ganciclovir. *J. Infect. Dis*. 185:162-169.

Cocchi, F., L. Menotti, P. Mirandola, M. Lopez, and G. Campadelli-Fiume. 1998. The ectodomain of a novel member of the immunoglobulin subfamily related to the poliovirus receptor has the attributes of a bona fide receptor for herpes simplex virus types 1 and 2 in human cells. *J. Virol.* 72:9992-10002.

Coen, D.M., M. Kosz-Vnenchak, J.G. Jacobson, D.A. Leib, C.L. Bogard, P.A. Schaffer, K.L. Tyler, and D.M. Knipe. 1989. Thymidine kinase-negative herpes simplex virus mutants establish latency in mouse trigeminal ganglia but do not reactivate. *Proc. Natl. Acad. Sci. U. S. A*. 86:4736-4740.

Coen, D.M. 1994. Acyclovir-resistant, pathogenic herpesviruses. *Trends Microbiol*. 2:481-485.

Colberg-Poley, A.M. 1996. Functional roles of immediate early proteins encoded by the human cytomegalovirus UL36-38, UL115-119, TRS1/IRS1 and US3 loci. *Intervirology*. 39:350-360.

Colberg-Poley, A.M., L.D. Santomenna, P.P. Harlow, P.A. Benfield, and D.J. Tenney. 1992. Human cytomegalovirus US3 and UL36-38 immediate-early proteins regulate gene expression. *J. Virol.* 66:95-105.

Compton, T., D.M. Nowlin, and N.R. Cooper. 1993. Initiation of human cytomegalovirus infection requires initial interaction with cell surface heparan sulfate. *Virology.* 193:834-841.

Conley, A.J., D.M. Knipe, P.C. Jones, and B. Roizman. 1981. Molecular genetics of herpes simplex virus. VII. Characterization of a temperature-sensitive mutant produced by in vitro mutagenesis and defective in DNA synthesis and accumulation of gamma polypeptides. *J. Virol.* 37:191-206.

Cooper, J., J. Conner, and J.B. Clements. 1995. Characterization of the novel protein kinase activity present in the R1 subunit of herpes simplex virus ribonucleotide reductase. *J. Virol.* 69:4979-4985.

Costanzo, F., G. Campadelli-Fiume, L. Foa-Tomasi, and E. Cassai. 1977. Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase B. *J. Virol.* 21:996-1001.

Cranage, M.P., T. Kouzarides, A.T. Bankier, S. Satchwell, K. Weston, P. Tomlinson, B. Barrell, H. Hart, S.E. Bell, A.C. Minson, and et al. 1986. Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. *Embo. J.* 5:3057-3063.

Crute, J.J., E.S. Mocarski, and I.R. Lehman. 1988. A DNA helicase induced by herpes simplex virus type 1. *Nucleic Acids Res*. 16:6585-6596.

Daikoku, T., K. Ikenoya, H. Yamada, F. Goshima, and Y. Nishiyama. 1998. Identification and characterization of the herpes simplex virus type 1 UL51 gene product. *J. Gen. Virol.* 79:3027-3031.

Das, S., and P.E. Pellett. 2007. Members of the HCMV US12 family of predicted heptaspanning membrane proteins have unique intracellular distributions, including association with the cytoplasmic virion assembly complex. *Virology.* 361:263-273.

Davison, A.J., P. Akter, C. Cunningham, A. Dolan, C. Addison, D.J. Dargan, A.F. Hassan-Walker, V.C. Emery, P.D. Griffiths, and G.W.G. Wilkinson. 2003a. Homology between the human cytomegalovirus RL11 gene family and human adenovirus E3 genes. *J. Gen. Virol.* 84:657-663.

Davison, A.J., A. Dolan, P. Akter, C. Addison, D.J. Dargan, D.J. Alcendor, D.J. McGeoch, and G.S. Hayward. 2003b. The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. *J. Gen. Virol.* 84:17-28.

Davison, A.J., R. Eberle, B. Ehlers, G.S. Hayward, D.J. McGeoch, A.C. Minson, P.E. Pellett, B. Roizman, M.J. Studdert, and E. Thiry. 2009. The order Herpesvirales. *Arch Virol.* 154:171-177.

de Bruyn Kops, A., and D.M. Knipe. 1988. Formation of DNA replication structures in herpes virus-infected cells requires a viral DNA binding protein. *Cell*. 55:857-868.

De Smaele, E., F. Zazzeroni, S. Papa, D.U. Nguyen, R. Jin, J. Jones, R. Cong, and G. Franzoso. 2001. Induction of gadd45beta by NF-kappaB downregulates pro-apoptotic JNK signalling. *Nature.* 414:308-313.

Deiss, L.P., J. Chou, and N. Frenkel. 1986. Functional domains within the a sequence involved in the cleavage-packaging of herpes simplex virus DNA. *J. Virol.* 59:605-618.

DeLuca, N.A., A.M. McCarthy, and P.A. Schaffer. 1985. Isolation and characterization of deletion mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory protein ICP4. *J. Virol.* 56:558-570.

DeLuca, N.A., and P.A. Schaffer. 1985. Activation of immediate-early, early, and late promoters by temperature-sensitive and wild-type forms of herpes simplex virus type 1 protein ICP4. *Mol. Cell Biol.* 5:1997-1208.

DeLuca, N.A., and P.A. Schaffer. 1988. Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. *J. Virol*. 62:732-743.

Desai, P., G.L. Sexton, E. Huang, and S. Person. 2008. Localization of herpes simplex virus type 1 UL37 in the Golgi complex requires UL36 but not capsid structures. *J. Virol.* 82:11354-11361.

Desai, P., G.L. Sexton, J.M. McCaffery, and S. Person. 2001. A null mutation in the gene encoding the herpes simplex virus type 1 UL37 polypeptide abrogates virus maturation. *J. Virol.* 75:10259-10271.

Deshmane, S.L., and N.W. Fraser. 1989. During latency, herpes simplex virus type 1 DNA is associated with nucleosomes in a chromatin structure. *J. Virol.* 63:943-947.

Donnelly, M., J. Verhagen, and G. Elliott. 2007. RNA binding by the herpes simplex virus type 1 nucleocytoplasmic shuttling protein UL47 is mediated by an N-terminal arginine-rich domain that also functions as its nuclear localization signal. *J. Virol.* 81:2283-2296.

Dracheva, S., E.V. Koonin, and J.J. Crute. 1995. Identification of the primase active site of the herpes simplex virus type 1 helicase-primase. *J. Biol. Chem.* 270:14148-14153.

Droin, N., S. Cathelin, A. Jacquel, L. Guery, C. Garrido, M. Fontenay, O. Hermine, and E. Solary. 2008. A role for caspases in the differentiation of erythroid cells and macrophages. *Biochimie.* 90:416-422.

Dubin, G., E. Socolof, I. Frank, and H.M. Friedman. 1991. Herpes simplex virus type 1 Fc receptor protects infected cells from antibody-dependent cellular cytotoxicity. *J. Virol.* 65:7046-7050.

Dugan, G.E., and E.W. Hewitt. 2008. Structural and functional dissection of the human cytomegalovirus immune evasion protein US6. *J. Virol.* 82:3271-3282.

Dunn, W., C. Chou, H. Li, R. Hai, D. Patterson, V. Stolc, H. Zhu, and F. Liu. 2003. Functional profiling of a human cytomegalovirus genome. *Proc. Natl. Acad. Sci. U. S. A*. 100:14223-14228.

Dutch, R.E., and I.R. Lehman. 1993. Renaturation of complementary DNA strands by herpes simplex virus type 1 ICP8. *J. Virol.* 67:6945-6949.

Dutia, B.M. 1983. Ribonucleotide reductase induced by herpes simplex virus has a virus-specified constituent. *J. Gen. Virol.* 64:513-521.

Elias, P., and I.R. Lehman. 1988. Interaction of origin binding protein with an origin of replication of herpes simplex virus 1. *Proc. Natl. Acad. Sci. U. S. A*. 85:2959-2963.

Ellison, K.S., R.A. Maranchuk, K.L. Mottet, and J.R. Smiley. 2005. Control of VP16 translation by the herpes simplex virus type 1 immediate-early protein ICP27. *J. Virol.* 79:4120-4131.

Enquist, L.W., P.J. Husak, B.W. Banfield, and G.A. Smith. 1998. Infection and spread of alphaherpesviruses in the nervous system. *Adv. Virus Res.* 51:237-347.

Eom, C.-Y., W.D. Heo, M.L. Craske, T. Meyer, and I.R. Lehman. 2004. The neural F-box protein NFB42 mediates the nuclear export of the herpes simplex virus type 1 replication initiator protein (UL9 protein) after viral infection. *Proc. Natl. Acad. Sci. U. S. A*. 101:4036-4040.

Ertl, P.F., and K.L. Powell. 1992. Physical and functional interaction of human cytomegalovirus DNA polymerase and its accessory protein (ICP36) expressed in insect cells. *J. Virol.* 66:4126-4133.

Everly, D.N., Jr., P. Feng, I.S. Mian, and G.S. Read. 2002. mRNA degradation by the virion host shutoff (vhs) protein of herpes simplex virus: Genetic and biochemical evidence that vhs is a nuclease. *J. Virol.* 76:8560-8571.

Faber, S.W., and K.W. Wilcox. 1986. Association of the herpes simplex virus regulatory protein ICP4 with specific nucleotide sequences in DNA. *Nucleic Acids Res*. 14:6067-6083.

Farrar, G.H., and J.D. Oram. 1984. Characterization of the human cytomegalovirus envelope glycoproteins. *J. Gen. Virol.* 65:1991-2001.

Farrell, M.J., A.T. Dobson, and L.T. Feldman. 1991. Herpes simplex virus latency-associated transcript is a stable intron. *Proc. Natl. Acad. Sci. U. S. A*. 88:790-794.

Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle WA.

Felsenstein, J.D. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.

Fierer, D.S., and M.D. Challberg. 1992. Purification and characterization of UL9, the herpes simplex virus type 1 origin-binding protein. *J. Virol.* 66:3986-3995.

Frame, M.C., H.S. Marsden, and B.M. Dutia. 1985. The ribonucleotide reductase induced by herpes simplex virus type 1 involves minimally a complex of two polypeptides (136K and 38K). *J. Gen. Virol.* 66:1581-1587.

Frengen, E., D. Weichenhan, B. Zhao, K. Osoegawa, M. van Geel, and P.J. de Jong. 1999. A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. *Genomics*. 58:250-253.

Frenkel, N., R.J. Jacob, R.W. Honess, G.S. Hayward, H. Locker, and B. Roizman. 1975. Anatomy of herpes simplex virus DNA. III. Characterization of defective DNA molecules and biological properties of virus populations containing them. *J. Virol.* 16:153-167.

Fruh, K., K. Ahn, H. Djaballah, P. Sempe, P.M. van Endert, R. Tampe, P.A. Peterson, and Y. Yang. 1995. A viral inhibitor of peptide transporters for antigen presentation. *Nature*. 375:415-418.

Fuchs, W., B.G. Klupp, H. Granzow, N. Osterrieder, and T.C. Mettenleiter. 2002. The interacting UL31 and UL34 gene products of pseudorabies virus are involved in egress from the host-cell nucleus and represent components of primary enveloped but not mature virions. *J. Virol.* 76:364-378.

Furlong, D., H. Swift, and B. Roizman. 1972. Arrangement of Herpesvirus Deoxyribonucleic Acid in the Core. *J. Virol.* 10:1071-1074.

Furman, M.H., N. Dey, D. Tortorella, and H.L. Ploegh. 2002. The human cytomegalovirus US10 gene product delays trafficking of major histocompatibility complex class I molecules. *J. Virol.* 76:11753-11756.

Fyfe, J.A., P.M. Keller, P.A. Furman, R.L. Miller, and G.B. Elion. 1978. Thymidine kinase from herpes simplex virus phosphorylates the new antiviral compound, 9-(2-hydroxyethoxymethyl)guanine. *J. Biol. Chem.* 253:8721-8727.

Gallina, A., E. Percivalle, L. Simoncini, M.G. Revello, G. Gerna, and G. Milanesi. 1996. Human cytomegalovirus pp65 lower matrix phosphoprotein harbours two transplantable nuclear localization signals. *J. Gen. Virol.* 77:1151-1157.

Gallina, A., L. Simoncini, S. Garbelli, E. Percivalle, G. Pedrali-Noy, K.S. Lee, R.L. Erikson, B. Plachter, G. Gerna, and G. Milanesi. 1999. Polo-like kinase 1 as a target for human cytomegalovirus pp65 lower matrix protein. *J. Virol.* 73:1468-1478.

Gallo, M.L., D.I. Dorsky, C.S. Crumpacker, and D.S. Parris. 1989. The essential 65-kilodalton DNA-binding protein of herpes simplex virus stimulates the virus-encoded DNA polymerase. *J. Virol.* 63:5023-5029.

Gao, M., and D.M. Knipe. 1991. Potential role for herpes simplex virus ICP8 DNA replication protein in stimulation of late gene expression. *J. Virol.* 65:2666-2675.

Gibbs, J.S., H.C. Chiou, K.F. Bastow, Y.C. Cheng, and D.M. Coen. 1988. Identification of amino acids in herpes simplex virus DNA polymerase involved in substrate and drug recognition. *Proc. Natl. Acad. Sci. U. S. A*. 85:6672-6676.

Gibson, W., M. Baxter, and K. Clopper. 1996. Cytomegalovirus "missing" capsid protein identified as heat-aggregable product of human cytomegalovirus UL46. *J. Virol.* 70:7454-7461.

Gibson, W., and B. Roizman. 1972. Proteins specified by herpes simplex virus VIII. Characterization and composition of multiple capsid forms of subtypes 1 and 2. *J. Virol.* 10:1044-1052.

Gibson, W., and B. Roizman. 1974. Proteins specified by herpes simplex virus X. Staining and radiolabeling properties of B capsid and virion proteins in polyacrylamide gels. *J. Virol.* 13:155-165.

Godowski, P.J., and D.M. Knipe. 1983. Mutations in the major DNA-binding protein gene of herpes simplex virus type 1 result in increased levels of viral gene expression. *J. Virol.* 47:478-486.

Godowski, P.J., and D.M. Knipe. 1985. Identification of a herpes simplex virus function that represses late gene expression from parental viral genomes. *J. Virol.* 55:357-365.

Godowski, P.J., and D.M. Knipe. 1986. Transcriptional control of herpesvirus gene expression: gene functions required for positive and negative regulation. *Proc. Natl. Acad. Sci. U. S. A*. 83:256-260.

Goins, W.F., L.R. Sternberg, K.D. Croen, P.R. Krause, R.L. Hendricks, D.J. Fink, S.E. Straus, M. Levine, and J.C. Glorioso. 1994. A novel latency-active promoter is contained within the herpes simplex virus type 1 UL flanking repeats. *J. Virol.* 68:2239-2252.

Goldmacher, V.S., L.M. Bartle, A. Skaletskaya, C.A. Dionne, N.L. Kedersha, C.A. Vater, J.-w. Han, R.J. Lutz, S. Watanabe, E.D.C. McFarland, E.D. Kieff, E.S. Mocarski, and T. Chittenden. 1999. A cytomegalovirus-encoded mitochondria-localized inhibitor of apoptosis structurally unrelated to Bcl-2. *Proc. Natl. Acad. Sci. U. S. A*. 96:12536-12541.

Goldstein, D.J., and S.K. Weller. 1988. Herpes simplex virus type 1-induced ribonucleotide reductase activity is dispensable for virus growth and DNA synthesis: isolation and characterization of an ICP6 lacZ insertion mutant. *J. Virol.* 62:196-205.

Gottlieb, J., A.I. Marcy, D.M. Coen, and M.D. Challberg. 1990. The herpes simplex virus type 1 UL42 gene product: a subunit of DNA polymerase that functions to increase processivity. *J. Virol.* 64:5976-5987.

Gourves, A.-S., N.T. Le Gac, G. Villani, P.E. Boehmer, and N.P. Johnson. 2000. Equilibrium binding of single-stranded DNA with herpes simplex virus type I-coded single-stranded DNA-binding protein, ICP8. *J. Biol. Chem.* 275:10864-10869.

Gruhler, A., P.A. Peterson, and K. Fruh. 2000. Human cytomegalovirus immediate early glycoprotein US3 retains MHC class I molecules by transient association. *Traffic.* 1:318-325.

Guzowski, J.F., and E.K. Wagner. 1993. Mutational analysis of the herpes simplex virus type 1 strict late UL38 promoter/leader reveals two regions critical in transcriptional regulation. *J. Virol.* 67:5098-5108.

Hagmann, M., O. Georgiev, W. Schaffner, and P. Douville. 1995. Transcription factors interacting with herpes simplex virus {alpha} gene promoters in sensory neurons. *Nucl. Acids Res.* 23:4978-4985.

Hahn, G., R. Jores, and E.S. Mocarski. 1998. Cytomegalovirus remains latent in a common precursor of dendritic and myeloid cells. *Proc. Natl. Acad. Sci. U. S. A*. 95:3937-3942.

Hahn, G., H. Khan, F. Baldanti, U.H. Koszinowski, M.G. Revello, and G. Gerna. 2002. The human cytomegalovirus ribonucleotide reductase homolog UL45 is dispensable for growth in endothelial cells, as determined by a BAC-cloned clinical isolate of human cytomegalovirus with preserved wild-type characteristics. *J. Virol.* 76:9551-9555.

Hansen, S.G., L.I. Strelow, D.C. Franchi, D.G. Anders, and S.W. Wong. 2003. Complete sequence and genomic analysis of rhesus cytomegalovirus. *J. Virol.* 77:6620-6636.

Hanson, L.K., J.S. Slater, V.J. Cavanaugh, W.W. Newcomb, L.L. Bolin, C.N. Nelson, L.D. Fetters, Q. Tang, J.C. Brown, G.G. Maul, and A.E. Campbell. 2009. Murine cytomegalovirus capsid assembly is dependent on US22 family gene M140 in infected macrophages. *J. Virol.* 83:7449-7456.

Hayward, G.S., R.J. Jacob, S.C. Wadsworth, and B. Roizman. 1975. Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. *Proc. Natl. Acad. Sci. U. S. A.* 72:4243-4247.

Heery, D.M., F. Gannon, and R. Powell. 1990. A simple method for subcloning DNA fragments from gel slices. *Trends Genet.* 6:173.

Heilbronn, R., G. Jahn, A. Burkle, U.K. Freese, B. Fleckenstein, and H. zur Hausen. 1987. Genomic localization, sequence analysis, and transcription of the putative human cytomegalovirus DNA polymerase gene. *J. Virol*. 61:119-124.

Higgins, D.G., and P.M. Sharp. 1989. Fast and sensitive multiple sequence alignments on a microcomputer. *Comput. Appl. Biosci.* 5:151-153.

Hill, A., P. Jugovic, l. York, G. Russ, J. Bennink, J. Yewdell, H. Ploegh, and D. Johnson. 1995. Herpes simplex virus turns off the TAP to evade host immunity. *Nature*. 375:411-415.

Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Molec. Biol*. 26:365-369.

Homa, F.L., and J.C. Brown. 1997. Capsid assembly and DNA packaging in herpes simplex virus. *Rev. Med. Virol.* 7:107-122.

Homman-Loudiyi, M., K. Hultenby, W. Britt, and C. Soderberg-Naucler. 2003. Envelopment of human cytomegalovirus occurs by budding into Golgi-derived vacuole compartments positive for gB, Rab 3, trans-Golgi network 46, and mannosidase II. *J. Virol.* 77:3191-3203.

Honess, R.W., and B. Roizman. 1974. Regulation of herpesvirus macromolecular synthesis I. Cascade regulation of the synthesis of three groups of viral proteins. *J. Virol.* 14:8-19.

Hosaka, M., M. Nagahama, W. Kim, T. Watanabe, K. Hatsuzawa, J. Ikemizu, K. Murakami, and K. Nakayama. 1991. Arg-X-Lys/Arg-Arg motif as a signal for precursor cleavage catalyzed by furin within the constitutive secretory pathway. *J. Biol. Chem.* 266:12127-12130.

Huang, L., Y. Zhu, and D.G. Anders. 1996. The variable 3' ends of a human cytomegalovirus oriLyt transcript (SRT) overlap an essential, conserved replicator element. *J. Virol.* 70:5272-5281.

Huber, M.T., and T. Compton. 1998. The human cytomegalovirus UL74 gene encodes the third component of the glycoprotein H-glycoprotein L-containing envelope complex. *J. Virol*. 72:8191-8197.

Huber, M.T., and T. Compton. 1999. Intracellular formation and processing of the heterotrimeric gH-gL-gO (gCIII) glycoprotein envelope complex of human cytomegalovirus. *J. Virol.* 73:3886-3892.

Hyun, J.J., H.S. Park, K.H. Kim, and H.J. Kim. 1999. Analysis of transcripts expressed from the UL47 gene of human cytomegalovirus. *Arch. Pharm. Res*. 22:542-548.

Igarashi, K., R. Fawl, R.J. Roller, and B. Roizman. 1993. Construction and properties of a recombinant herpes simplex virus 1 lacking both S-component origins of DNA synthesis. *J. Virol.* 67:2123-2132.

Ishov, A.M., and G.G. Maul. 1996. The periphery of nuclear domain 10 (ND10) as site of DNA virus deposition. *J. Cell Biol.* 134:815-826.

Izumi, K.M., and J.G. Stevens. 1990. Molecular and biological characterization of a herpes simplex virus type 1 (HSV-1) neuroinvasiveness gene. *J. Exp. Med.* 172:487-496.

Jacob, R.J., L.S. Morse, and B. Roizman. 1979. Anatomy of herpes simplex virus DNA XII. Accumulation of head-to-tail concatemers in nuclei of infected cells and their role in the generation of the four isomeric arrangements of viral DNA. *J. Virol.* 29:448-457.

Jarvis, M.A., T.R. Jones, D.D. Drummond, P.P. Smith, W.J. Britt, J.A. Nelson, and C.J. Baldick. 2004. Phosphorylation of human cytomegalovirus glycoprotein B (gB) at the acidic cluster casein kinase 2 site (Ser900) is required for localization of gB to the trans-Golgi network and efficient virus replication. *J. Virol.* 78:285-293.

Jarvis, M.A., and J.A. Nelson. 2006. HCMV tropism for endothelial cells: Not all endothelial cells are created equal. *J. Virol.* 81: 2095 - 2101.

Jones, T.R., and S.-W. Lee. 2004. An acidic cluster of human cytomegalovirus UL99 tegument protein is required for trafficking and function. *J. Virol.* 78:1488-1502.

Jones, T.R., and V.P. Muzithras. 1991. Fine mapping of transcripts expressed from the US6 gene family of human cytomegalovirus strain AD169. *J. Virol.* 65:2024-2036.

Jun, Y., E. Kim, M. Jin, H.C. Sung, H. Han, D.E. Geraghty, and K. Ahn. 2000. Human cytomegalovirus gene products US3 and US6 down-regulate trophoblast class I MHC molecules. *J. Immunol*. 164:805-811.

Kalejta, R.F. 2008. Tegument proteins of human cytomegalovirus. *Microbiol. Mol. Biol. Rev.* 72:249-265.

Kalejta, R.F., J.T. Bechtel, and T. Shenk. 2003. Human cytomegalovirus pp71 stimulates cell cycle progression by inducing the proteasome-dependent degradation of the retinoblastoma family of tumor suppressors. *Mol. Cell. Biol.* 23:1885-1895.

Kari, B., and R. Gehrz. 1992. A human cytomegalovirus glycoprotein complex designated gC-II is a major heparin-binding component of the envelope. *J. Virol.* 66:1761-1764.

Kaye, J.F., U.A. Gompels, and A.C. Minson. 1992. Glycoprotein H of human cytomegalovirus (HCMV) forms a stable complex with the HCMV UL115 gene product. *J. Gen. Virol.* 73:2693-2698.

Kemble, G.W., and E.S. Mocarski. 1989. A host cell protein binds to a highly conserved sequence element (pac-2) within the cytomegalovirus a sequence. *J. Virol.* 63:4715-4728.

Klinedinst, D.K., and M.D. Challberg. 1994. Helicase-primase complex of herpes simplex virus type 1: a mutation in the UL52 subunit abolishes primase activity. *J. Virol.* 68:3693-3701.

Klupp, B.G., W. Fuchs, H. Granzow, R. Nixdorf, and T.C. Mettenleiter. 2002. Pseudorabies virus UL36 tegument protein physically interacts with the UL37 protein. *J. Virol.* 76:3065-3071.

Kniess, N., M. Mach, J. Fay, and W.J. Britt. 1991. Distribution of linear antigenic sites on glycoprotein gp55 of human cytomegalovirus. *J. Virol.* 65:138-146.

Knipe, D.M., W. Batterson, C. Nosal, B. Roizman, and A. Buchan. 1981. Molecular genetics of herpes simplex virus. VI. Characterization of a temperature-sensitive mutant defective in the expression of all early viral gene products. *J. Virol*. 38:539-547.

Koff, A., and P. Tegtmeyer. 1988. Characterization of major recognition sequences for a herpes simplex virus type 1 origin-binding protein. *J. Virol.* 62:4096-4103.

Kondo, K., H. Kaneshima, and E.S. Mocarski. 1994. Human cytomegalovirus latent infection of granulocyte-macrophage progenitors. *Proc. Natl. Acad. Sci. U. S. A*. 91:11879-11883.

Kristie, T.M., and B. Roizman. 1984. Separation of sequences defining basal expression from those conferring alpha gene recognition within the regulatory domains of herpes simplex virus 1 alpha genes. *Proc. Natl. Acad. Sci. U. S. A*. 81:4065-4069.

Krosky, P.M., M.-C. Baek, and D.M. Coen. 2003. The human cytomegalovirus UL97 protein kinase, an antiviral drug target, is required at the stage of nuclear egress. *J. Virol.* 77:905-914.

Krzyzaniak, M., M. Mach, and W.J. Britt. 2007. The cytoplasmic tail of glycoprotein M (gpUL100) expresses trafficking signals required for human cytomegalovirus assembly and replication. *J. Virol.* 81:10316-10328.

Kumar, S., K. Tamura, I.B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*. 17:1244-1245.

LaBoissiere, S., and P. O'Hare. 2000. Analysis of HCF, the cellular cofactor of VP16, in herpes simplex virus-infected cells. *J. Virol.* 74:99-109.

Ladin, B.F., S. Ihara, H. Hampl, and T. Ben-Porat. 1982. Pathway of assembly of herpesvirus capsids: an analysis using DNA+ temperature-sensitive mutants of pseudorabies virus. *Virology*. 116:544-561.

Lamkanfi, M., N. Festjens, W. Declercq, T.V. Berghe, and P. Vandenabeele. 2006. Caspases in cell survival, proliferation and differentiation. *Cell Death Differ*. 14:44-55.

Lee, C.K., and D.M. Knipe. 1985. An immunoassay for the study of DNA-binding activities of herpes simplex virus protein ICP8. *J. Virol*. 54:731-738.

Leib, D.A., K.C. Nadeau, S.A. Rundle, and P.A. Schaffer. 1991. The promoter of the latency-associated transcripts of herpes simplex virus type 1 contains a functional cAMP-response element: role of the latency-associated transcripts and cAMP in reactivation of viral latency. *Proc. Natl. Acad. Sci. U. S. A*. 88:48-52.

Lesniewski, M., S. Das, Y. Skomorovska-Prokvolit, F.Z. Wang, and P.E. Pellett. 2006. Primate cytomegalovirus US12 gene family: a distinct and diverse clade of seven-transmembrane proteins. *Virology.* 354:286-298.

Li, J., T. Yamamoto, K. Ohtsubo, M. Shirakata, and K. Hirai. 1999. Major product pp43 of human cytomegalovirus UL112-113 gene is a transcriptional coactivator with two functionally distinct domains. *Virology*. 260:89-97.

Lindahl, T., and B. Nyberg. 1974. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochem*. 13:3405-3410.

Liptak, L., S. Uprichard, and D. Knipe. 1996. Functional order of assembly of herpes simplex virus DNA replication proteins into prereplicative site structures. *J. Virol.* 70:1759-1767.

Littler, E., A.D. Stuart, and M.S. Chee. 1992. Human cytomegalovirus UL97 open reading frame encodes a protein that phosphorylates the antiviral nucleoside analogue ganciclovir. *Nature.* 358:160-162.

Liu, B., and M.F. Stinski. 1992. Human cytomegalovirus contains a tegument protein that enhances transcription from promoters with upstream ATF and AP-1 cis-acting elements. *J. Virol.* 66:4434-4444.

Liu, J., Z. Shriver, R.M. Pope, S.C. Thorp, M.B. Duncan, R.J. Copeland, C.S. Raska, K. Yoshida, R.J. Eisenberg, G. Cohen, R.J. Linhardt, and R. Sasisekharan. 2002. Characterization of a heparan sulfate octasaccharide that binds to herpes simplex virus type 1 glycoprotein D. *J. Biol. Chem.* 277:33456-33467.

Liu, M., J. Tang, X. Wang, T. Yang, and A.I. Geller. 2005. Enhanced long-term expression from helper virus-free HSV-1 vectors packaged in the presence of deletions in genes that modulate the function of VP16, U L 46 and U L 47. *J Neurosci. Methods.* 145:1-9.

Liu, Y., and B.J. Biegalke. 2002. The human cytomegalovirus UL35 gene encodes two proteins with different functions. *J. Virol.* 76:2460-2468.

Lockridge, K.M., S.-S. Zhou, R.H. Kravitz, J.L. Johnson, E.T. Sawai, E.L. Blewett, and P.A. Barry. 2000. Primate cytomegaloviruses encode and express an IL-10-like protein. *Virology*. 268:272-280.

Lokensgard, J.R., D.C. Bloom, A.T. Dobson, and L.T. Feldman. 1994. Long-term promoter activity during herpes simplex virus latency. *J. Virol.* 68:7148-7158.

Lokensgard, J.R., H. Berthomme, and L.T. Feldman. 1997. The latency-associated promoter of herpes simplex virus type 1 requires a region downstream of the transcription start site for long-term expression during latency. *J. Virol*. 71:6714-6719.

Loomis, J.S., J.B. Bowzard, R.J. Courtney, and J.W. Wills. 2001. Intracellular trafficking of the UL11 tegument protein of herpes simplex virus type 1. *J. Virol.* 75:12209-12219.

Lu, R., and V. Misra. 2000a. Potential role for luman, the cellular homologue of herpes simplex virus VP16 (alpha gene trans-inducing factor), in herpesvirus latency. *J. Virol.* 74:934-943.

Lu, R., and V. Misra. 2000b. Zhangfei: a second cellular protein interacts with herpes simplex virus accessory factor HCF in a manner similar to luman and VP16. *Nucleic Acids Res.* 28:2446-2454.

Mach, M., B. Kropff, P. Dal Monte, and W. Britt. 2000. Complex Formation by Human Cytomegalovirus glycoproteins M (gpUL100) and N (gpUL73). *J. Virol.* 74:11881-11892.

Mach, M., B. Kropff, M. Kryzaniak, and W. Britt. 2005. Complex Formation by Glycoproteins M and N of human cytomegalovirus: structural and functional aspects. *J. Virol.* 79:2160-2170.

Mach, M., K. Osinski, B. Kropff, U. Schloetzer-Schrehardt, M. Krzyzaniak, and W. Britt. 2007. The carboxy-terminal domain of glycoprotein N of human cytomegalovirus is required for virion morphogenesis. *J. Virol.* 81:5212-5224.

Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Manzur, M., P. Fleming, D.C.S. Huang, M.A. Degli-Esposti, and C.E. Andoniou. 2008. Virally mediated inhibition of Bax in leukocytes promotes dissemination of murine cytomegalovirus. *Cell Death Differ*. 16:312-320.

Masse, M.J.O., S. Karlin, G.A. Schachtel, and E.S. Mocarski. 1992. Human cytomegalovirus origin of DNA replication (oriLyt) resides within a highly complex repetitive region. *Proc. Natl. Acad. Sci. U. S. A.* 89:5246-5250.

McCormick, A.L., L. Roback, D. Livingston-Rosanoff, and C. St. Clair. 2010. The human cytomegalovirus UL36 gene controls caspase-dependent and -independent cell death programs activated by infection of monocytes differentiating to macrophages. *J. Virol.* 84:5108-5123.

McGeehan, J.E., N.W. Depledge, and D.J. McGeoch. 2001. Evolution of the dUTPase gene of mammalian and avian herpesviruses. *Curr. Protein Pept. Sci.* 2:325-333.

McNabb, D.S., and R.J. Courtney. 1992. Characterization of the large tegument protein (ICP1/2) of herpes simplex virus type 1. *Virology*. 190:221-232.

McVoy, M.A., D.E. Nixon, S.P. Adler, and E.S. Mocarski. 1998. Sequences within the herpesvirus-conserved *pac1* and *pac2* motifs are required for cleavage and packaging of the murine cytomegalovirus genome. *J. Virol.* 72:48-56.

McVoy, M.A., D.E. Nixon, J.K. Hur, and S.P. Adler. 2000. The ends on herpesvirus DNA replicative concatemers contain pac2 cis cleavage/packaging elements and their formation is controlled by terminal cis sequences. *J. Virol.* 74:1587-1592.

McWatters, B.J., R.M. Stenberg, and J.A. Kerry. 2002. Characterization of the human cytomegalovirus UL75 (glycoprotein H) late gene promoter. *Virology.* 303:309-316.

Menard, C., M. Wagner, Z. Ruzsics, K. Holak, W. Brune, A.E. Campbell, and U.H. Koszinowski. 2003. Role of murine cytomegalovirus US22 gene family members in replication in macrophages. *J. Virol.* 77:5557-5570.

Messerle, M., I. Crnkovic, W. Hammerschmidt, H. Ziegler, and U.H. Koszinowski. 1997. Cloning and mutagenesis of a herpesvirus genome as an infectious bacterial artificial chromosome. *Proc. Natl. Acad. Sci. U. S. A*. 94:14759-14763.

Mettenleiter, T.C. 2002. Herpesvirus Assembly and Egress. *J. Virol.* 76:1537-1547.

Meyer, H., Y. Masuho, and M. Mach. 1990. The gp116 of the gp58/116 complex of human cytomegalovirus represents the amino-terminal part of the precursor molecule and contains a neutralizing epitope. *J. Gen. Virol.* 71:2443-2450.

Meyer, H., V.A. Sundqvist, L. Pereira, and M. Mach. 1992. Glycoprotein gp116 of human cytomegalovirus contains epitopes for strain-common and strain-specific antibodies. *J. Gen. Virol.* 73:2375-2383.

Meyer, H.H., A. Ripalti, M.P. Landini, K. Radsak, H.F. Kern, and G.M. Hensel. 1997. Human cytomegalovirus late-phase maturation is blocked by stably expressed UL32 antisense mRNA in astrocytoma cells. *J. Gen. Virol.* 78:2621-2631.

Michaels, M.G., D.P. Greenberg, D.L. Sabo, and E.R. Wald. 2003. Treatment of children with congenital cytomegalovirus infection with ganciclovir. *Pediatr. Infect. Dis. J.* 22:504-508.

Michel, D., P. Schaarschmidt, K. Wunderlich, M. Heuschmid, L. Simoncini, D. Muhlberger, A. Zimmermann, I. Pavic, and T. Mertens. 1998. Functional regions of the human cytomegalovirus protein pUL97 involved in nuclear localization and phosphorylation of ganciclovir and pUL97 itself. *J. Gen. Virol.* 79:2105-2112.

Mocarski, E.S., Jr., and G.W. Kemble. 1996. Recombinant cytomegaloviruses for study of replication and pathogenesis. *Intervirology.* 39:320-330.

Montgomery, R.I., M.S. Warner, B.J. Lum, and P.G. Spear. 1996. Herpes simplex virus-1 entry into cells mediated by a novel member of the TNF/NGF receptor family. *Cell*. 87:427-436.

Murphy, E., D. Yu, J. Grimwood, J. Schmutz, M. Dickson, M.A. Jarvis, G. Hahn, J.A. Nelson, R.M. Myers, and T.E. Shenk. 2003. Coding potential of laboratory and clinical strains of human cytomegalovirus. *Proc. Natl. Acad. Sci. U. S. A*. 100:14976-14981.

Nakayama, T., K. Hieshima, D. Nagakubo, E. Sato, M. Nakayama, K. Kawa, and O. Yoshie. 2004. Selective induction of Th2-attracting chemokines CCL17 and CCL22 in human B cells by latent membrane protein 1 of Epstein-Barr virus. *J. Virol.* 78:1665-1674.

Navarro, D., P. Paz, S. Tugizov, K. Topp, J. La Vail, and L. Pereira. 1993. Glycoprotein B of human cytomegalovirus promotes virion penetration into cells, transmission of infection from cell to cell, and fusion of infected cells. *Virology.* 197:143-158.

Newcomb, W.W., and J.C. Brown. 1994. Induced extrusion of DNA from the capsid of herpes simplex virus type 1. *J. Virol.* 68:433-440.

Newcomb, W.W., F.L. Homa, and J.C. Brown. 2006. Herpes simplex virus capsid structure: DNA packaging protein UL25 is located on the external surface of the capsid near the vertices. *J. Virol.* 80:6286-6294.

Newcomb, W.W., F.L. Homa, D.R. Thomsen, B.L. Trus, N. Cheng, A. Steven, F. Booy, and J.C. Brown. 1999. Assembly of the herpes simplex virus procapsid from purified components and identification of small complexes containing the major capsid and scaffolding proteins. *J. Virol.* 73:4239-4250.

Newcomb, W.W., D.R. Thomsen, F.L. Homa, and J.C. Brown. 2003. Assembly of the herpes simplex virus capsid: identification of soluble scaffold-portal complexes and their role in formation of portal-containing capsids. *J. Virol.* 77:9862-9871.

Newcomb, W.W., B.L. Trus, F.P. Booy, A.C. Steven, J.S. Wall, and J.C. Brown. 1993. Structure of the herpes simplex virus capsid molecular composition of the pentons and the triplexes. *J. Mol. Biol.* 232:499-511.

Neyts, J., R. Snoeck, D. Schols, J. Balzarini, J.D. Esko, A. Van Schepdael, and E. De Clercq. 1992. Sulfated polymers inhibit the interaction of human cytomegalovirus with cell surface heparan sulfate. *Virology*. 189:48-58.

Nguyen, N.L., A.N. Loveland, and W. Gibson. 2008. Nuclear localization sequences in cytomegalovirus capsid assembly proteins (UL80 proteins) are required for virus production: inactivating NLS1, NLS2, or both affects replication to strikingly different extents. *J. Virol.* 82:5381-5389.

Norkin, L. 1995. Virus receptors: implications for pathogenesis and the design of antiviral agents. *Clin. Microbiol. Rev.* 8:293-315.

Norris, K.L., and R.J. Youle. 2008. Cytomegalovirus proteins vMIA and m38.5 link mitochondrial morphogenesis to Bcl-2 family proteins. *J. Virol.* 82:6232-6243.

Nowak, B., A. Gmeiner, P. Sarnow, A.J. Levine, and B. Fleckenstein. 1984a. Physical mapping of human cytomegalovirus genes: identification of DNA sequences coding for a virion phosphoprotein of 71 kDa and a viral 65-kDa polypeptide. *Virology*. 134:91-102.

Nowak, B., C. Sullivan, P. Sarnow, R. Thomas, F. Bricout, J.C. Nicolas, B. Fleckenstein, and A.J. Levine. 1984b. Characterization of monoclonal antibodies and polyclonal immune sera directed against human cytomegalovirus virion proteins. *Virology*. 132:325-338.

O'Hare, P., and G.S. Hayward. 1985. Three trans-acting regulatory proteins of herpes simplex virus modulate immediate-early gene expression in a pathway involving positive and negative feedback regulation. *J Virol*. 56:723-733.

O'Hare, P., 1993. The virion transactivator of herpes simplex virus. *Sem. Virol*. 4:145-155.

Ogasawara, M., T. Suzutani, I. Yoshida, and M. Azuma. 2001. Role of the UL25 gene product in packaging DNA into the herpes simplex virus capsid: location of UL25 product in the capsid and demonstration that it binds DNA. *J. Virol.* 75:1427-1436.

Ogawa-Goto, K., S. Irie, A. Omori, Y. Miura, H. Katano, H. Hasegawa, T. Kurata, T. Sata, and Y. Arao. 2002. An endoplasmic reticulum protein, p180, is highly expressed in human cytomegalovirus-permissive cells and interacts with the tegument protein encoded by UL48. *J. Virol.* 76:2350-2362.

Ojala, P.M., B. Sodeik, M.W. Ebersold, U. Kutay, and A. Helenius. 2000. Herpes simplex virus type 1 entry into host cells: reconstitution of capsid binding and uncoating at the nuclear pore complex in vitro. *Mol. Cell. Biol.* 20:4922-4931.

Oresic, K., and D. Tortorella. 2008. Endoplasmic reticulum chaperones participate in human cytomegalovirus US2-mediated degradation of class I major histocompatibility complex molecules. *J. Gen. Virol.* 89:1122-1130.

Pande, N.T., C. Powers, K. Ahn, and K. Fruh. 2005. Rhesus cytomegalovirus contains functional homologues of US2, US3, US6, and US11. *J. Virol.* 79:5786-5798.

Pari, G.S., and D.G. Anders. 1993. Eleven loci encoding trans-acting factors are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA replication. *J. Virol.* 67:6979-6988.

Patrone, M., E. Percivalle, M. Secchi, L. Fiorina, G. Pedrali-Noy, M. Zoppe, F. Baldanti, G. Hahn, U.H. Koszinowski, G. Milanesi, and A. Gallina. 2003. The human cytomegalovirus UL45 gene product is a late, virion-associated protein and influences virus growth at low multiplicities of infection. *J. Gen. Virol.* 84:3359-3370.

Penfold, M.E., P. Armati, and A.L. Cunningham. 1994. Axonal transport of herpes simplex virions to epidermal cells: evidence for a specialized mode of virus transport and assembly. *Proc. Natl. Acad. Sci. U. S. A*. 91:6529-6533.

Penfold, M.E.T., T.L. Schmidt, D.J. Dairaghi, P.A. Barry, and T.J. Schall. 2003. Characterization of the Rhesus Cytomegalovirus US28 Locus. *J. Virol.* 77:10404-10413.

Perdue, M.L., J.C. Cohen, C.C. Randall, and D.J. O'Callaghan. 1976. Biochemical studies of the maturation of herpesvirus nucleocapsid species. *Virology*. 74:194-208.

Perng, G.-C., S.M. Slanina, A. Yukht, B.S. Drolet, W. Keleher, Jr., H. Ghiasi, A.B. Nesburn, and S.L. Wechsler. 1999. A herpes simplex virus type 1 latency-associated transcript mutant with increased virulence and reduced spontaneous reactivation. *J. Virol.* 73:920-929.

Perng, G.C., E.C. Dunkel, P.A. Geary, S.M. Slanina, H. Ghiasi, R. Kaiwar, A.B. Nesburn, and S.L. Wechsler. 1994. The latency-associated transcript gene of herpes simplex virus type 1 (HSV-1) is required for efficient in vivo spontaneous reactivation of HSV-1 from latency. *J. Virol.* 68:8045-8055.

Pertel, P.E., A. Fridberg, M.L. Parish, and P.G. Spear. 2001. Cell fusion induced by herpes simplex virus glycoproteins gB, gD, and gH-gL requires a gD receptor but not necessarily heparan sulfate. *Virology*. 279:313-324.

Phelan, A., M. Carmo-Fonseca, J. McLaughlan, A.I. Lamond, and J.B. Clements. 1993. A herpes simplex virus type 1 immediate-early gene product, IE63, regulates small nuclear ribonucleoprotein distribution. *Proc. Natl. Acad. Sci. U. S. A*. 90:9056-9060.

Pietropaolo, R.L., and T. Compton. 1997. Direct interaction between human cytomegalovirus glycoprotein B and cellular annexin II. *J. Virol.* 71:9803-9807.

Pignatelli, S., P. Dal Monte, G. Rossini, D. Camozzi, V. Toscano, R. Conte, and M.P. Landini. 2006. Latency-associated human cytomegalovirus glycoprotein N genotypes in monocytes from healthy blood donors. *Transfusion*. 46:1754-1762.

Pilger, B.D., C. Cui, and D.M. Coen. 2004. Identification of a small molecule that inhibits herpes simplex virus DNA polymerase subunit interactions and viral replication. *Chem. Biol.* 11:647-654.

Poffenberger, K.L., and B. Roizman. 1985. A noninverting genome of a viable herpes simplex virus 1: presence of head-to-tail linkages in packaged genomes and requirements for circularization after infection. *J. Virol.* 53:587-595.

Polvino-Bodnar, M., P.K. Orberg, and P.A. Schaffer. 1987. Herpes simplex virus type 1 oriL is not required for virus replication or for the establishment and reactivation of latent infection in mice. *J. Virol.* 61:3528-3535.

Preston, C.M., and M.J. Nicholl. 2005. Human cytomegalovirus tegument protein pp71 directs long-term gene expression from quiescent herpes simplex virus genomes. *J. Virol.* 79:525-535.

Preston, V.G., and F.B. Fisher. 1984. Identification of the herpes simplex virus type 1 gene encoding the dUTPase. *Virology*. 138:58-68.

Prichard, M.N., M.E.T. Penfold, G.M. Duke, R.R. Spaete, and G.W. Kemble. 2001. A review of genetic differences between limited and extensively passaged human cytomegalovirus strains. *Rev. Med. Virol.* 11:191-200.

Prod'homme, V., C. Griffin, R.J. Aicheler, E.C. Wang, B.P. McSharry, C.R. Rickards, R.J. Stanton, L.K. Borysiewicz, M. Lopez-Botet, G.W. Wilkinson, and P. Tomasec. 2007. The human cytomegalovirus MHC class I homolog UL18 inhibits LIR-1+ but activates LIR-1- NK cells. *J. Immunol.* 178:4473-4481.

Purifoy, D.J., R.B. Lewis, and K.L. Powell. 1977. Identification of the herpes simplex virus DNA polymerase gene. *Nature.* 269:621-623.

Purifoy, D.J., and K.L. Powell. 1976. DNA-binding proteins induced by herpes simplex virus type 2 in HEp-2 cells. *J. Virol.* 19:717-731.

Purves, F.C., D. Spector, and B. Roizman. 1992. UL34, the target of the herpes simplex virus U(S)3 protein kinase, is a membrane protein which in its unphosphorylated state associates with novel phosphoproteins. *J. Virol.* 66:4295-4303.

Quinlan, M.P., L.B. Chen, and D.M. Knipe. 1984. The intranuclear location of a herpes simplex virus DNA-binding protein is determined by the status of viral DNA replication. *Cell*. 36:857-868.

Rader, K.A., C.E. Ackland-Berglund, J.K. Miller, J.S. Pepose, and D.A. Leib. 1993. In vivo characterization of site-directed mutations in the promoter of the herpes simplex virus type 1 latency-associated transcripts. *J. Gen. Virol.* 74:1859-1869.

Rapp, M., M. Messerle, B. Buhler, M. Tannheimer, G.M. Keil, and U.H. Koszinowski. 1992. Identification of the murine cytomegalovirus glycoprotein B gene and its expression by recombinant vaccinia virus. *J. Virol.* 66. 4399-4406.

Rasmussen, L., J. Mullenax, R. Nelson, and T.C. Merigan. 1985. Viral polypeptides detected by a complement-dependent neutralizing murine monoclonal antibody to human cytomegalovirus. *J. Virol.* 55:274-280.

Rawlinson, W.D., H.E. Farrell, and B.G. Barrell. 1996. Analysis of the complete DNA sequence of murine cytomegalovirus. *J. Virol.* 70:8833-8849.

Rawlinson, W.D., F. Zeng, H.E. Farrell, A.L. Cunningham, A.A. Scalzo, T.W.M. Booth, and G.M. Scott. 1997. The murine cytomegalovirus (MCMV) homolog of the HCMV phosphotransferase (UL97(pk)) gene. *Virology*. 233:358-363.

Rebe, C., S. Cathelin, S. Launay, R. Filomenko, L. Prevotat, C. L'Ollivier, E. Gyan, O. Micheau, S. Grant, A. Dubart-Kupperschmitt, M. Fontenay, and E. Solary. 2007. Caspase-8 prevents sustained activation of NF-{kappa}B in monocytes undergoing macrophagic differentiation. *Blood*. 109:1442-1450.

Reschke, M., B. Reis, K. Noding, D. Rohsiepe, A. Richter, T. Mockenhaupt, W. Garten, and K. Radsak. 1995. Constitutive expression of human cytomegalovirus glycoprotein B (gpUL55) with mutagenized carboxy-terminal hydrophobic domains. *J. Gen. Virol.* 76:113-122.

Reuven, N.B., A.E. Staire, R.S. Myers, and S.K. Weller. 2003. The herpes simplex virus type 1 alkaline nuclease and single-stranded DNA binding protein mediate strand exchange in vitro. *J. Virol.* 77:7425-7433.

Reynolds, A.E., B.J. Ryckman, J.D. Baines, Y. Zhou, L. Liang, and R.J. Roller. 2001. UL31 and UL34 proteins of herpes simplex virus type 1 form a complex that accumulates at the nuclear rim and is required for envelopment of nucleocapsids. *J. Virol.* 75:8803-8817.

Rigoutsos, I., J. Novotny, T. Huynh, S.T. Chin-Bow, L. Parida, D. Platt, D. Coleman, and T. Shenk. 2003. In silico pattern-based analysis of the human cytomegalovirus genome. *J. Virol.* 77:4326-4344.

Rock, D.L., and N.W. Fraser. 1985. Latent herpes simplex virus type 1 DNA contains two copies of the virion DNA joint region. *J. Virol.* 55:849-852.

Roizman, B. 1996. Herpesviridae. *In* Fields Virology. Lippincott-Raven, Philadelphia.

Roizman, B., D. Furlong. 1974. The Replication of Herpesviruses. *In* Comprehensive Virology. Plenum Press, New York.

Ross, T.G., R.P. Rogers, N. Elfrink, N. Bray, and E.L. Blewett. 2005. Detection of baboon cytomegalovirus (BaCMV) by PCR using primers directed against the glycoprotein B gene. *J. Virol. Methods.* 125:119-124.

Rue, C.A., M.A. Jarvis, A.J. Knoche, H.L. Meyers, V.R. DeFilippis, S.G. Hansen, M. Wagner, K. Fruh, D.G. Anders, S.W. Wong, P.A. Barry, and J.A. Nelson. 2004. A cyclooxygenase-2 homologue encoded by rhesus cytomegalovirus is a determinant for endothelial cell tropism. *J. Virol.* 78:12529-12536.

Ruger, B., S. Klages, B. Walla, J. Albrecht, B. Fleckenstein, P. Tomlinson, and B. Barrell. 1987. Primary structure and transcription of the genes coding for the two virion phosphoproteins pp65 and pp71 of human cytomegalovirus. *J. Virol.* 61:446-453.

Ruyechan, W.T. 1983. The major herpes simplex virus DNA-binding protein holds single-stranded DNA in an extended configuration. *J. Virol*. 46:661-666.

Ryckman, B.J., M.C. Chase, and D.C. Johnson. 2009. HCMV TR strain glycoprotein O acts as a chaperone promoting gH/gL incorporation into virions, but is not present in virions. *J. Virol.* 84:2597-2609.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Cold Springs Harbor Press., Cold Springs Harbor, N.Y.

Sanchez, V., E. Sztul, and W.J. Britt. 2000. Human cytomegalovirus pp28 (UL99) localizes to a cytoplasmic compartment which overlaps the endoplasmic reticulum-golgi-intermediate compartment. *J.Virol*. 74:3842-3851.

Sandri-Goldin, R., M. Hibbard, and M. Hardwicke. 1995. The C-terminal repressor region of herpes simplex virus type 1 ICP27 is required for the redistribution of small nuclear ribonucleoprotein particles and splicing factor SC35; however, these alterations are not sufficient to inhibit host cell splicing. *J. Virol.* 69:6063-6076.

Sandri-Goldin, R.M. 1998. ICP27 mediates HSV RNA export by shuttling through a leucine-rich nuclear export signal and binding viral intronless RNAs through an RGG motif. *Genes Dev.* 12:868-879.

Sarisky, R.T., and G.S. Hayward. 1996. Evidence that the UL84 gene product of human cytomegalovirus is essential for promoting oriLyt-dependent DNA replication and formation of replication compartments in cotransfection assays. *J. Virol.* 70:7398-7413.

Sawtell, N.M., and R.L. Thompson. 1992. Herpes simplex virus type 1 latency-associated transcription unit promotes anatomical site-dependent establishment and reactivation from latency. *J. Virol.* 66:2157-2169.

Schleiss, M.R. 1994. Cloning and characterization of the guinea pig cytomegalovirus glycoprotein B gene. *Virology*. 202:173-185.

Schmolke, S., P. Drescher, G. Jahn, and B. Plachter. 1995a. Nuclear targeting of the tegument protein pp65 (UL83) of human cytomegalovirus: an unusual bipartite nuclear localization signal functions with other portions of the protein to mediate its efficient nuclear transport. *J. Virol.* 69:1071-1078.

Schmolke, S., H.F. Kern, P. Drescher, G. Jahn, and B. Plachter. 1995b. The dominant phosphoprotein pp65 (UL83) of human cytomegalovirus is dispensable for growth in cell culture. *J. Virol.* 69:5959-5968.

Schrőer, J., and T. Shenk. 2008. Inhibition of cyclooxygenase activity blocks cell-to-cell spread of human cytomegalovirus. *Proc. Natl. Acad. Sci. U. S. A.* 105:19468-19473.

Scott, E.S., and P. O'Hare. 2001. Fate of the inner nuclear membrane protein lamin B receptor and nuclear lamins in herpes simplex virus type 1 infection. *J. Virol.* 75:8818-8830.

Seo, J.-Y., and W.J. Britt. 2007. Cytoplasmic envelopment of human cytomegalovirus requires the postlocalization function of tegument protein pp28 within the assembly compartment. *J. Virol.* 81:6536-6547.

5, Y., and J. Nemunaitis. 2006. Herpes simplex virus 1 (HSV-1) for cancer treatment. *Cancer Gene. Ther*. 13:975-992.

Shen, W., E. Westgard, L. Huang, M.D. Ward, J.L. Osborn, N.H. Chau, L. Collins, B. Marcum, M.A. Koach, J. Bibbs, O.J. Semmes, and J.A. Kerry. 2008. Nuclear trafficking of the human cytomegalovirus pp71 (ppUL82) tegument protein. *Virology*. 376:42-52

Shepard, A.A., A.N. Imbalzano, and N.A. DeLuca. 1989. Separation of primary structural components conferring autoregulation, transactivation, and DNA-binding properties to the herpes simplex virus transcriptional regulatory protein ICP4. *J. Virol.* 63:3714-3728.

Sherman, G., and S.L. Bachenheimer. 1988. Characterization of intranuclear capsids made by ts morphogenic mutants of HSV-1. *Virology*. 163:471-480.

Shieh, M.T., D. WuDunn, R.I. Montgomery, J.D. Esko, and P.G. Spear. 1992. Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. *J. Cell Biol.* 116:1273-1281.

Shukla, D., J. Liu, P. Blaiklock, N.W. Shworak, X. Bai, J.D. Esko, G.H. Cohen, R.J. Eisenberg, R.D. Rosenberg, and P.G. Spear. 1999. A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell*. 99:13-22.

Silvestri, M., V.A. Sundqvist, U. Ruden, and B. Wahren. 1991. Characterization of a major antigenic region on gp55 of human cytomegalovirus. *J. Gen. Virol.* 72:3017-3023.

Siminoff, P., and M.G. Menefee. 1966. Normal and 5-bromodeoxyuridine-inhibited development of herpes simplex virus : An electron microscope study. *Exp. Cell Res.* 44:241-255.

Sinclair, J., and P. Sissons. 1996. Latent and Persistent Infections of Monocytes and Macrophages. *Intervirology*. 39:293-301.

Singh, J., and T. Compton. 2000. Characterization of a panel of insertion mutants in human cytomegalovirus glycoprotein B. *J. Virol.* 74:1383-1392.

Skaletskaya, A., L.M. Bartle, T. Chittenden, A.L. McCormick, E.S. Mocarski, and V.S. Goldmacher. 2001. A cytomegalovirus-encoded inhibitor of apoptosis that suppresses caspase-8 activation. *Proc. Natl. Acad. Sci. U. S. A*. 98:7829-7834.

Skepper, J.N., A. Whiteley, H. Browne, and A. Minson. 2001. Herpes simplex virus nucleocapsids mature to progeny virions by an envelopment → deenvelopment → reenvelopment pathway. *J. Virol.* 75:5697-5702.

Smiley, J.R. 2004. Herpes simplex virus virion host shutoff protein: immune evasion mediated by a viral RNase? *J. Virol.* 78:1063-1068.

Smith, C.C., M. Kulka, J.P. Wymer, T.D. Chung, and L. Aurelian. 1992. Expression of the large subunit of herpes simplex virus type 2 ribonucleotide reductase (ICP10) is required for virus growth and neoplastic transformation. *J. Gen. Virol*. 73:1417-1428.

Smith, C.C., T. Peng, M. Kulka, and L. Aurelian. 1998. The PK domain of the large subunit of herpes simplex virus type 2 ribonucleotide reductase (ICP10) is required for immediate-early gene expression and virus growth. *J. Virol.* 72:9131-9141.

Smuda, C., E. Bogner, and K. Radsak. 1997. The human cytomegalovirus glycoprotein B gene (ORF UL55) is expressed early in the infectious cycle. *J. Gen. Virol.* 78:1981-1992.

Sodeik, B., M.W. Ebersold, and A. Helenius. 1997. Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. *J. Cell Biol*. 136:1007-1021.

Soliman, T.M., and S.J. Silverstein. 2000. Identification of an Export Control Sequence and a Requirement for the KH Domains in ICP27 from Herpes Simplex Virus Type 1. *J. Virol.* 74:7600-7609.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.

Spaete, R.R., A. Saxena, P.I. Scott, G.J. Song, W.S. Probert, W.J. Britt, W. Gibson, L. Rasmussen, and C. Pachl. 1990. Sequence requirements for proteolytic processing of glycoprotein B of human cytomegalovirus strain Towne. *J. Virol.* 64:2922-2931.

Spaete, R.R., R.M. Thayer, W.S. Probert, F.R. Masiarz, S.H. Chamberlain, L. Rasmussen, T.C. Merigan, and C. Pachl. 1988. Human cytomegalovirus strain Towne glycoprotein B is processed by proteolytic cleavage. *Virology.* 167:207-225.

Spear, P.G., R.J. Eisenberg, and G.H. Cohen. 2000. Three classes of cell surface receptors for alphaherpesvirus entry. *Virology*. 275:1-8.

Spear, P.G., and B. Roizman. 1972. Proteins specified by herpes simplex virus V. purification and structural proteins of the herpesvirion. *J. Virol.* 9:143-159.

Staras, Stephanie A.S., Sheila C. Dollard, Kay W. Radford, W. Dana Flanders, Robert F. Pass, and Michael J. Cannon. 2006. Seroprevalence of cytomegalovirus infection in the United States, 1988-1994. *Clin. Infect. Dis*. 43:1143-1151.

Stasiak, P.C., and E.S. Mocarski. 1992. Transactivation of the cytomegalovirus ICP36 gene promoter requires the alpha gene product TRS1 in addition to IE1 and IE2. *J. Virol.* 66:1050-1058.

Stern, S., and W. Herr. 1991. The herpes simplex virus trans-activator VP16 recognizes the Oct-1 homeo domain: evidence for a homeo domain recognition subdomain. *Genes & Dev*. 5:2555-2566.

Stevens, J., E. Wagner, G. Devi-Rao, M. Cook, and L. Feldman. 1987. RNA complementary to a herpesvirus alpha gene mRNA is prominent in latently infected neurons. *Science*. 235:1056-1059.

Streblow, D.N., J. Vomaske, P. Smith, R. Melnychuk, L. Hall, D. Pancheva, M. Smit, P. Casarosa, D.D. Schlaepfer, and J.A. Nelson. 2003. Human cytomegalovirus chemokine receptor US28-induced smooth muscle cell migration is mediated by focal adhesion kinase and Src. *J. Biol. Chem.* 278:50456-50465.

Strive, T., E. Borst, M. Messerle, and K. Radsak. 2002. Proteolytic processing of human cytomegalovirus glycoprotein B is dispensable for viral growth in culture. *J. Virol.* 76:1252-1264.

Stropes, M.P., O.D. Schneider, W.A. Zagorski, J.L.C. Miller, and W.E. Miller. 2009. The carboxy-terminal tail of human cytomegalovirus (HCMV) US28 regulates both chemokine-independent and chemokine-dependent signaling in HCMV-infected cells. *J. Virol.* 83:10016-10027.

Sullivan, V., C.L. Talarico, S.C. Stanat, M. Davis, D.M. Coen, and K.K. Biron. 1992. A protein kinase homologue controls phosphorylation of ganciclovir in human cytomegalovirus-infected cells. *Nature.* 358:162-164.

Takahashi, K., H. Nakanishi, M. Miyahara, K. Mandai, K. Satoh, A. Satoh, H. Nishioka, J. Aoki, A. Nomoto, A. Mizoguchi, and Y. Takai. 1999. Nectin/PRR: An immunoglobulin-like cell adhesion molecule recruited to cadherin-based adherens junctions through interaction with afadin, a PDZ domain containing protein. *J. Cell Biol.* 145:539-549.

Talarico, C.L., T.C. Burnette, W.H. Miller, S.L. Smith, M.G. Davis, S.C. Stanat, T.I. Ng, Z. He, D.M. Coen, B. Roizman, and K.K. Biron. 1999. Acyclovir is phosphorylated by the human cytomegalovirus UL97 protein. *Antimicrob. Agents Chemother.* 43:1941-1946.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol*. 24:1596-1599.

Tanaka, J., T. Ogura, H. Iida, H. Sato, and M. Hatano. 1988. Inhibitors of prostaglandin synthesis inhibit growth of human cytomegalovirus and reactivation of latent virus in a productively and latently infected human cell line. *Virology*. 163:205-208.

Tanaka, M., T. Sata, and Y. Kawaguchi. 2008. The product of the herpes simplex virus 1 UL7 gene interacts with a mitochondrial protein, adenine nucleotide translocator 2. *Virol. J.* 5:125.

Taus, N.S., and W.J. Mitchell. 2001. The transgenic ICP4 promoter is activated in schwann cells in trigeminal ganglia of mice latently infected with herpes simplex virus type 1. *J. Virol.* 75:10401-10408.

Taylor-Wiedeman, J., J.G.P. Sissons, L.K. Borysiewicz, and J.H. Sinclair. 1991. Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. *J. Gen. Virol.* 72:2059-2064.

Teodoro, J.G., and P.E. Branton. 1997. Regulation of apoptosis by viral gene products. *J. Virol.* 71:1739-1746.

Theiler, R.N., and T. Compton. 2001. Characterization of the signal peptide processing and membrane association of human cytomegalovirus glycoprotein O. *J. Biol. Chem.* 276:39226-39231.

Thomas, M.S., M. Gao, D.M. Knipe, and K.L. Powell. 1992. Association between the herpes simplex virus major DNA-binding protein and alkaline nuclease. *J. Virol.* 66:1152-1161.

Thompson, R.L., and N.M. Sawtell. 1997. The herpes simplex virus type 1 latency-associated transcript gene regulates the establishment of latency. *J. Virol*. 71:5432-5440.

Tirabassi, R.S., and H.L. Ploegh. 2002. The human cytomegalovirus US8 glycoprotein binds to major histocompatibility complex class I products. *J. Virol.* 76:6832-6835.

Tirosh, B., N.N. Iwakoshi, B.N. Lilley, A.-H. Lee, L.H. Glimcher, and H.L. Ploegh. 2005. Human cytomegalovirus protein US11 provokes an unfolded protein response that may facilitate the degradation of class I major histocompatibility complex products. *J. Virol.* 79:2768-2779.

Trgovcich, J., C. Cebulla, P. Zimmerman, and D.D. Sedmak. 2006. Human Cytomegalovirus protein pp71 disrupts major histocompatibility complex class I cell surface expression. *J. Virol.* 80:951-963.

Triezenberg, S.J., R.C. Kingsbury, and S.L. McKnight. 1988. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev*. 2:718-729.

Truant, R., and B.R. Cullen. 1999. The arginine-rich domains present in human immunodeficiency virus type 1 tat and rev function as direct importin beta - dependent nuclear localization signals. *Mol. Cell. Biol.* 19:1210-1217.

Turner, E.E., J.M. Rhee, and L.T. Feldman. 1997. The POU-domain factor Brn-3.0 recognizes characteristic sites in the herpes simplex virus genome. *Nucleic Acids Res*. 25:2589-2594.

Uprichard, S., and D. Knipe. 1996. Herpes simplex ICP27 mutant viruses exhibit reduced expression of specific DNA replication genes. *J. Virol.* 70:1969-1980.

Uprichard, S.L., and D.M. Knipe. 2003. Conformational changes in the herpes simplex virus ICP8 dna-binding protein coincident with assembly in viral replication structures. *J. Virol.* 77:7467-7476.

Utz, U., W. Britt, L. Vugler, and M. Mach. 1989. Identification of a neutralizing epitope on glycoprotein gp58 of human cytomegalovirus. *J. Virol.* 63:1995-2001.

Vales-Gomez, M., and H.T. Reyburn. 2006. Intracellular trafficking of the HCMV immunoevasin UL16 depends on elements present in both its cytoplasmic and transmembrane domains. *J. Mol. Biol.* 363:908-917.

Vales-Gomez, M., M. Shiroishi, K. Maenaka, and H.T. Reyburn. 2005. Genetic variability of the major histocompatibility complex class I homologue encoded by human cytomegalovirus leads to differential binding to the inhibitory receptor ILT2. *J. Virol.* 79:2251-2260.

Varmuza, S.L., and J.R. Smiley. 1985. Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. *Cell*. 41:793-802.

Vey, M., W. Schafer, B. Reis, R. Ohuchi, W. Britt, W. Garten, H.D. Klenk, and K. Radsak. 1995. Proteolytic processing of human cytomegalovirus glycoprotein B (gpUL55) is mediated by the human endoprotease furin. *Virology*. 206:746-749.

Vink, C., E. Beuken, and C.A. Bruggeman. 2000. Complete DNA sequence of the rat cytomegalovirus genome. *J. Virol.* 74:7656-7665.

Vischer, H.F., R. Leurs, and M.J. Smit. 2006. HCMV-encoded G-protein-coupled receptors as constitutively active modulators of cellular signaling networks. *Trends Pharmacol. Sci.* 27:56-63.

Vlazny, D.A., A. Kwong, and N. Frenkel. 1982. Site-specific cleavage/packaging of herpes simplex virus DNA and the selective maturation of nucleocapsids containing full-length viral DNA. *Proc. Natl. Acad. Sci. U. S. A*. 79:1423-1427.

Wadsworth, S., R.J. Jacob, and B. Roizman. 1975. Anatomy of herpes simplex virus DNA. II. Size, composition, and arrangement of inverted terminal repetitions. *J. Virol.* 15:1487-1497.

Wagner, B., B. Kropff, H. Kalbacher, W. Britt, V.A. Sundqvist, L. Ostberg, and M. Mach. 1992. A continuous sequence of more than 70 amino acids is essential for antibody binding to the dominant antigenic site of glycoprotein gp58 of human cytomegalovirus. *J. Virol*. 66:5290-5297.

Warner, M.S., R.J. Geraghty, W.M. Martinez, R.I. Montgomery, J.C. Whitbeck, R. Xu, R.J. Eisenberg, G.H. Cohen, and P.G. Spear. 1998. A cell surface protein with herpesvirus entry activity (HveB) confers susceptibility to infection by mutants of herpes simplex virus type 1, herpes simplex virus type 2, and pseudorabies virus. *Virology*. 246:179-189.

Watkins, A.M., P.J. Dunford, A.M. Moffatt, P. Wong-Kai-In, M.J. Holland, D.S. Pole, G.M. Thomas, J. Martin, N.A. Roberts, and M.J. Mulqueen. 1998. Inhibition of virus-encoded thymidine kinase suppresses herpes simplex virus replication in vitro and in vivo. *Antivir. Chem. Chemother.* 9:9-18.

Weiland, K.L., N.L. Oien, F. Homa, and M.W. Wathen. 1994. Functional analysis of human cytomegalovirus polymerase accessory protein. *Virus Res.* 34:191-206.

Welch, A.R., L.M. McGregor, and W. Gibson. 1991. Cytomegalovirus homologs of cellular G protein-coupled receptor genes are transcribed. *J. Virol.* 65:3915-3918.

Weller, S.K., A. Spadaro, J.E. Schaffer, A.W. Murray, A.M. Maxam, and P.A. Schaffer. 1985. Cloning, sequencing, and functional analysis of oriL, a herpes simplex virus type 1 origin of DNA synthesis. *Mol. Cell Biol.* 5:930-942.

White, C.A., N.D. Stow, A.H. Patel, M. Hughes, and V.G. Preston. 2003. Herpes simplex virus type 1 portal protein UL6 Interacts with the putative terminase subunits UL15 and UL28. *J. Virol.* 77:6351-6358.

Whittaker, G.R., and A. Helenius. 1998. Nuclear import and export of viruses and virus genomes. *Virology*. 246:1-23.

Wildy, P., and D.H. Watson. 1962. Electron microscopic studies on the architecture of animal viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27:25-47.

Wilkinson, G.W., P. Tomasec, R.J. Stanton, M. Armstrong, V. Prod'homme, R. Aicheler, B.P. McSharry, C.R. Rickards, D. Cochrane, S. Llewellyn-Lacey, E.C. Wang, C.A. Griffin, and A.J. Davison. 2008. Modulation of natural killer cells by human cytomegalovirus. *J. Clin. Virol.* 41:206-212.

Winkler, M., S.A. Rice, and T. Stamminger. 1994. UL69 of human cytomegalovirus, an open reading frame with homology to ICP27 of herpes simplex virus, encodes a transactivator of gene expression. *J. Virol.* 68:3943-3954.

Winkler, M., and T. Stamminger. 1996. A specific subform of the human cytomegalovirus transactivator protein pUL69 is contained within the tegument of virus particles. *J. Virol.* 70:8984-8987.

WuDunn, D., and P.G. Spear. 1989. Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *J. Virol.* 63:52-58.

Yamauchi, Y., K. Kiriyama, N. Kubota, H. Kimura, J. Usukura, and Y. Nishiyama. 2008. The UL14 tegument protein of herpes simplex virus type 1 is required for efficient nuclear transport of the alpha transinducing factor VP16 and viral capsids. *J. Virol.* 82:1094-1106.

Yang, W.C., G.V. Devi-Rao, P. Ghazal, E.K. Wagner, and S.J. Triezenberg. 2002. General and specific alterations in programming of global viral gene expression during infection by VP16 activation-deficient mutants of herpes simplex virus type 1. *J. Virol.* 76:12758-12774.

Yao, Z.Q., G. Gallez-Hawkins, N.A. Lomeli, X. Li, K.M. Molinder, D.J. Diamond, and J.A. Zaia. 2001. Site-directed mutation in a conserved kinase domain of human cytomegalovirus-pp65 with preservation of cytotoxic T lymphocyte targeting. *Vaccine*. 19:1628-1635.

Yeh, P.-C., D.G. Meckes, Jr., and J.W. Wills. 2008. Analysis of the interaction between the UL11 and UL16 tegument proteins of herpes simplex virus. *J. Virol.* 82:10693-10700.

Yoakum, G.H. 1993. Mapping a putative pyruvoyl decarboxylase active site to human cytomegalovirus open reading frame UL77. *Biochem. Biophys. Res. Commun.* 194:1207-1215.

York, I.A., C. Roop, D.W. Andrews, S.R. Riddell, F.L. Graham, and D.C. Johnson. 1994. A cytosolic herpes simplex virus protein inhibits antigen presentation to CD8+ T lymphocytes. *Cell.* 77:525-535.

Yue, Y., A. Kaur, S.S. Zhou, and P.A. Barry. 2006. Characterization and immunological analysis of the rhesus cytomegalovirus homologue (Rh112) of the human cytomegalovirus UL83 lower matrix phosphoprotein (pp65). *J. Gen. Virol.* 87:777-787.

Zhou, Z.H., D.H. Chen, J. Jakana, F.J. Rixon, and W. Chiu. 1999. Visualization of tegument-capsid interactions and DNA in intact herpes simplex virus type 1 virions. *J. Virol.* 73:3210-3218.

Zhu, H., J.-P. Cong, D. Yu, W.A. Bresnahan, and T.E. Shenk. 2002. Inhibition of cyclooxygenase 2 blocks human cytomegalovirus replication. *Proc. Natl. Acad. Sci. U. S. A*. 99:3932-3937.

Zhu, L.A., and S.K. Weller. 1992. The six conserved helicase motifs of the UL5 gene product, a component of the herpes simplex virus type 1 helicase-primase, are essential for its function. *J. Virol.* 66:469-479.

Zuckerkandl, L.E.a.P. 1965. "Evolutionary divergence and convergence in proteins.". *In* Evolving Genes and Proteins. V.B.a.H.J. Vogel, editor. Academic Press, New York.

APPENDIX A

BaCMV Sequence

UNIQUE LONG LEFT TERMINAL SEQUENCE:

GCGGCCGCGGCTACGGCTACCGCGCCATCCCGGGTCCCGATGGCGAGTCCCGCCCCGTGTGGACCCAGCACGTGGTGTTC
CTTCTCGGCGGCCACGGCCCCCGCGTGCACCTGAACCGCCCGTCGGCCAGGGAGGCTGAGGCTCGCGGGCTGCTGCCACG
CTGGCCCATGCGTCCCCAGACGCACACAGGTGCGCAGTACACACGGTACACCATCCACACGACCGAGCTGCAACACGTCC
CCAGGGTGCTTTTGAAACATGAGATTCTACGAGCCGTGCCGCAGACAGGTACGAGCCGCCTGGCGGAGCGCGGAGAGCGC
GGCGCCTGGATCGTGACTGACACGTTTACCGCCCCGGGGCGCAAACTGACTGCCAACACACGGCCGGCGGGCGAGTGGTG
GAAGCGAGACTCGAGCGTGGACCGGCTGGAGGACACGCTGGTACAGCTGCGCGAGACAGCCAGTGCGCCCGGGGCGCCTC
GCTTGACACGCAGAAGAATCATTCACACAGACACTAGTAGCAGCTACTACAGTAGCGCAGACAGCCTGCCCTCGAGCGAC
ACGGAATCGGCGGACGAGTCCGAGGGCAGTGAGGAACAAGAAGACAGAGATTGGGACCGGGAGCAGGACCGGGGAGAAGA
GAGCAGTGAGGAAGAGCGGGATAGGTATCAGATTGAGAGCTACCAACAGGAAAGCTACGAACAGGATGCAGAGGCACTGG
CGAGCGGAGACAGCATAAGTGCAGCAAGTGCAGCAGGGGCCGCGGACGAACGTCACGACATCCCCCGCCACAGCAGTTCT
GACAACAGTAGCCATGATAAAGAGAACACATAGAAAACGCGAGTGAACGGTGCTGTAAACAAACCTTTTATTGCAATAAA
AAGGCGCTACATCCAGTGTATCGTGTGTCCGTGCCTACTTTGCGCGTGCCCTGTGTTTGTTTCCCCGTGTGTCCAGTGCC
TCGACTGAGTCTGCTACAACTACGCTGCGGGTCCATTATTTTGCAGGCCGGGGGGCCCGACTTGTACCTTGTCATACTTC
CGCGTTACCCGGAGTTACATCATGCCAGTCCCAGACGGAGCCGGTGGCGGCGGCGTTCTGCAAGATGGCGTCTGGCGAGT
GTGCCTGTGCCTCTCGCTGTCCCTCGCCGTGGTCGCCATTCCCCTCGTCGGCTGGCTGATGGAGGCACGGAGCAACATCG
TGGAATTCATAGGGCTCATCGAGCTCGTCAACAGCATGATCAAGCCCGTCGCCGAGTGGCTGCTGGAAGCGTTCGTGTGG
GTCCTCCCGCGCGCCCTGTTCCTCGGCGGCTACCTGCTGCTGCTGGCGTGGCTGACCCTGGTGTTCGTGTTCATCGGCCT
CCGCATAGTCGGCGTCTTCCACGGACCACCACCGCTGCATGAAGTATTCTACGACGAAGGTCATTAGCACCTCCCCGAGC
CCGTCCCCGCGTACCCGGCCGTGGATGATCCGGTGGCCCTGCCACAGGAGCGTCTCGAAGGCCGCCAGCAGCACGGCCCG
CCAGCGCGGGGTAGAGAGGGCCGCGGAGAGCAGCTCCAGCAGCCACGAAGTCGAAGAGCGCTCCTCGGAGGAAGAGGCGG
AGACTCGCGCAAGACGCCTTGCCGTGACGGAGGCGTTGGCCGCAGACGCCGCCGCCGACGAAGAAGAGGAGTTGCCCTTG
GAGGCGTCCATGGCAGGTGTCTGGAGAGAGATGGGTGCGCGGATCGGGGGGAGAACGACTTCCTTGGTGGTCAAAGGGGA
GCGGGCGAGCAAAGATCCCCGGAAAGCAGGGGTTGTGCCAGCGGTGTGTGAGGAGAAGTGAAATATGGGAGAAGCGGCGA
GACTGCTTATTATAAAGGGTGCACGCATTGTTGATTCTGCAGTTTTTCAAGTGGTTTGGAAAAACCCGGAGCATCATGTG
TCCGGGACTTTTTTTGTTTCTGGAGATTACGGGAATAGCGATGACAGCAGCATCCGGTTCGGCGACGGGGTCAACGAGGA
CACAGCCGAGCATGACGCAAGTGGCACTGTGCCCGGGCGGGAATGTAACGTTCAACTACAGCAGACCTCAAGGTCACTCC
GTGTTTTGGAAATATACCAATTTAACAAAACCGGCACACAAACACCTGCATCAATACGTTATTTGCACTCTGACAGGCAG
CTACATACTGAAAGAGACCCGGAACTCCATGAACATGAAATGTAACAATCGATCACTCCAACTGTACAATGTGAGACCAC
AGGATGCGGGACTGTACGAACTGCACGACCACACAAATAACAGCGTGTTGATGGTTTTTAACGTGACTGTGAGAACCGTG
GTGGCACCACAAGTGACTGGAATGATTATCTACACGGTGTCTCGAGTGTACCATACATCAACCCACGAGAACGGCGTTAC
TAAACACAGAATCGGGAACGGCTGGGACACATGGATGGTGCATCTGTCATTCGCCACGGTGGCCATGACCTGCTTTGCCC
TGGCGGTGATTCTGTCCGGCTGCGTGTGTGCTCGGTCCATACGCGCCTGGTCTAACAATTACCGTCAGCTGAAGGATCAG
CCCGACTCGTGCGACGTGATTAAGCTGCCTGAAGAAAAGAAAGTGCCTATTGATGTGCTGACAGCTGTGACGGATGACAA
GCAACCCGCCACGCTGTGGCTGACCAAGTAACGGGGTCCCCACAATAATTGCCAATCATGCCCACGAGTCTGTTGTGTCA
TTTGTACGTTGAATAAAGAGGGCCAGTATGTAATCTTAGCAGTGCGCTGTGTTTTGTCGGTCACTGGGCGGGCGGTTTCG
ACGAAACTACCTGTGGTTTTGAGTCAGTGAGATTTCCCGGCAAGGGGATGAAATCATGGTTCTGTTGAAATGACTAATAT
TAACGGGAAATCTGCGTTTTCCAGAAAGTTATAAGCCACAACTTTTAATACAGCTCTAAACTACCGATACAAAACACCAA
AAATGTACACATATACAATGATAAATACCAAATTCGAAATGCAACCAAATACAAACCCGTATGTTACAATATGTACAATT
TTATTTGTCACAATTGTTTACACTAACGCATTTCAGATTCATGTTAAGGTTGGTGATAACCTAACGCTACAAGATAACTC
GTCAAAACATGAACAGCATATATCAGGATTATGGTATTATCAAACCAAATGTACAAATGGCTTACGTTACACTGCGTTAG
GCATACAGTTATGCCAAATACCTTACACAAATATCAGTATACACACATTATCCAACAAATGTGGAGACTTTGTAAGATAT
CAGTGTAAAAATACAACGCTACATTTGTATAACATTACCACAACAACGCCAACATCATACACATTATCAATAACATACCC
ACCAGATTCATTTTCAGTAGAGCATTTTTTTCTAAACATAACGTGTGTGACATGTTCCAAACCTACCAAAACGAGTCAAC
GTAGCACACATCCTACCACTATATCAAATTTATACAGTTCCATGGTAACATTTAAGGAAAATAATAGACTCCGTTTACAA

TCCTCTAGTGCAGCCACTGCTACGCCAATCATTGTAACTGTATCAATTGTGATTGTTATGACATGTTTTTGGTATCTGTA
TTATCGCCACAGACGCCGGCCGATGTTGTAAAAACAAAATACATATACTAAATATTGAAAACCAAGCTATAATGATACAG
AATAAAAACCACATACTATATTCTATATATACAATACTTTTATTATGTACAACTGGTTACACAGACACATCATTTAGAAA
CATTTCTGCTCACGTTGGTGACAATGTAACGTTACCAGATGTTCCGACAAAGTCAACTCCTCTTAAAGGCGAATGGTTTT
ATTTTAAATCTGGATGTTCTTCTCATAAACATTCCGATAATGGATATTGTAAATTGTGTTATGTGATTTACATAATGGAA
CACATCATTACCTATCAACCACCTGAAAGACATAAACCTTTGAGATATACATGTAACAAAACAGGTTTACATTTATACGA
CGTTAGCACGCACACACCATCTTTATATAAATTGATACATGAATACGACAGTGAAGATACAACATCTCATTATTTTCTAA
ACATAACGTATGTCAAACCTACGACACCTACAAAATGTACAGAAACTACCACATATATAACTACACATACAGCACCTACG
TTACATTCAACCGAGCCAGCGCCGATAGCATCCAACGCAAGTTTTGGTGCTTTACAATCTTCTGTCCACTCTAGGCCGAT
AATTGTGACTGTAATCATTGCAACATCGTTAGCGACATTTAGTTTGTACATGTATTATCAATACACACACCGATCAATGC
TGTAACATTGTGTGTTTTCAGCTGCATATAACCTTTTTTACGAACAAATATACCAAAGAACGCCGTGAAAATGCCCCCCC
CCCCCCGAAAATAAGCACTGGACACTGCTGCTCTGGTTACATGGAATGTATTTATTGCAATCTACAAAGAATTCACAAAC
ATGCACCCCTGTTGGTGATTCTGCAACGTTTTCTGTAAATTCTACAAATACAGTGAAAGTTGAATGGTTTTTTCAAAACG
GTTTGTGTCCCACTGGTTTTTGTACGCTTTGTTCTATAAATTATTTGGCCGATTATAACCCCGTTACTAAACATGAATTA
CAACATAGTATATTACATATGTTTAACAACACGTGCAATAAAACGGCGTTACACATTTACAACATTACCATACATAATGA
AGAAATGAGATTCAAATTACTTCAGACAACATATAACAAATTGGAAAAATGCTACTGTATAAAATCAGTAGCACACACAA
CATCTACACCGCATGTGCAAACATTTACCGCTTCGTCATTACATGAACATTTTAGCCATCTACAATCCACTGACCTGGGT
GTTACTACACCAATTGTTGTAACCGCAACACTGGGAATTATATTCACAGCTATAGGTTATTCATTCTATAGCCGAAGACG
CCGACGAATGCTGTAACAATAATGTTCCTCCACTCCATAAGCTACATAGCAACACATCCTTTATGGTTGTATTCATTTTA
TATTATTCGTCACGTACATGCCTTATCATACACTTATTCGGTGGAACCTGGTGATAGTGTAAGTTTTGTAGATGAAAGCT
ACACAAACCGTTTTGCATCAGGAGTATGGTATTTTAACATGACAACATTTAATCCATGTGAATCTCATATTCAGCAAAAT
ATACAACTACATATGCTATGTATTATATATTATCAAGTTAATTTTACACTACCAATAGACTTGCCCGAATGTTACCGACC
GTACTTTAATTATACATGTAACGAAACATCTCTACATATGTACAACATAAGCGCAAATGCACCCACAACATACTGGTTGA
AAAAATTTGGTGGACAAGAAAACGAAACATTCTATTACCTTAACATTACATTAAAAATACCCACTATACTCACAACACAG
CTAGAATACTCTCCCCCCCCCCCCACTGGACAGTTTCATCTTTTACAATCCAGTAATTCAGTGAATCCCACTCTGTTAACT
ATGTCTGGACTATTACTGCTTGCATGTTTCGGTTTAGGTTACATGTACTACCGATATACGCACAAACACAGAATCCCGCA
AAAGTGCTTCAAAGGAAATGAATATTTGTAATTCTAAATTACATTTTGCTACTCATTCTATCTTACCGAAACTCCTAACA
GGACTATCATACCTAATATTGCGAAACTTATGCAAAAAACATGCAACGCTAATATTACTACCAACATGCTTGGCGTCTAA
CCTAACGGTTATTTACATTACCAAAAATATAACAACCGTTTTCGGAGGTAATGTTAGTTTACTAACACCTGATAATATTC
CATTTCAATTTCAAACATGGTTCGCTAAAAATCCACAAGATCAAAACTGTCCAACAGCACAGTTATGCAAACAATCCCAT
AAGGATCTCGTAGAACTTACACGATTTGATAATAAATGTCCTCAATCTGACCATTATACATGCAACGACACGGGATTATA
TCTCTATGATATATCTAACACAACAGCTGCAACGTATATACTATCACAATACAACGGTAATGACTCCAAACACATAGACA
TAAGAAACATCACTCACAATGTAATTATTGTTAATTCAATCCCTCCATCATCGGTACTGGTTGCCACTAAGTACACCGAT
ACTAAAAATAGAACCATTATTGTACTAGTAGTAATATGCGCTATTGTTACACTAATGATAGCGGGATGTCTATGCTACCG
TACAAAAACACAACGTTTTGTTGAGGAATATGATAAACCGCATTATACCACAGTTCAACCAGGTAATCATTAAAAACATG
TCCCCACGTTATAATAGCATATATAGAACAATCACGTTACTGTTTTGCAATATTTTTTCCAAACAAAGCACCGCAATATG
GTCCTGTTGCATATATTTTTTTCACAAAGCAACTTGTCTAAAGGTTTCGCAAGCGTTCTGGGTTCCAATCGGTGGTACCG
TAACGTTTATAGATTTAAACGTCACAGAGAATACACAGTTCAATATGGGTGTATGGTACATTGTAAATGACCACTGTAAT
TGGATGGAGTTTAAACCCATAATGTGCAAATTATGCGAAGTAGATTACAACGCCGAAGAAAACCATCAAATTACTCAATA
TCCACCGAGTTCATACTATAAAGAAACATTCAATCATACATGTAACCTAACATCAATGTACCTATATAATATTACCGAAC
AAACACCAACAACCTTCAAATTTGTAAAAAAAAAATATCCAATCAAACGTATTTCTATCATCTTAACATAACCACCTGTA
AAAATTCTCCCCATCAAACATCATTATTAGTACCAATTACCGAAAAACACACCAATACAGAACATACAACTTTTATTGTT
CCAGTTATTATTTCCGTTGTTATTGCACTAGTAGCGGGCTGCCTGTATTACCGGTATACAAAAAAAAAAACCATTTACGC
CACTGGTTTTATTATCCTAACATCAAACAAATGATATTACCGACATATTGTAATTTGTGTTTTACGCTTATACCACTACA
GACATGCCCTAACCTAATTGTTAACAACATCACCAAAAACGGAACAACTACTTTTGAACGTAATGTTAGTTTAACACAAT
ACAATACTTTATTTGATTCAAATGCATGTTACACTGAAAATAAATGTGAAAACTGTTCAACATACTTATGTGACATACTA
AATAACACTGATTTAAACATAACTATGATACCTTCTCGCTGTCCTCAATCTGGAAACTATACATGCGACAAAACAGGATT
ATATTTCTATGATACATATAACATAACAACATATATATTGGCGCAATGCGACATGACTTGCAACATAAACACCAAAAACA
CTAACCTGATACTTTTTAATTTATCCCAGTCGGTACTAATTGCCGAAAAACACACCAATACCAAACATGCAACTTTTACT
GTTCCAGTTATCATTTCCGTTGTTATTGCACTAATAGCAAGCTGCCTGTATTACCGGTATACACACAAACCATTTACAGC
ACTGGTACTATTATCCTAACAAACAAATGACATTAACAGATTGTGATTTTACACTCATTACACTACAAACATATGTTATC
TACGTTACAGGAAATAGAACAAATACCATTGGAAGTAATGGTAGTTCATTAACACCTAATATTTCATGTTATTTACAATT
ATATTATCAAAACCTAAAAGATAAAAATTGTTCAACAGACTTATGTATATTATTAGAAAACAACGGTGGCGTGAAATACA
GAACACCAGATACTTATAGAAACTATACATACAACGAAACAAAATTACTTCTATATGATATATCTAACATAACAACCATT
ACATATATATTATTACAATGTAACGGTAGTGGCGACATAAACACAACAAACTACAACGTAACTGTGGCTTTTAACCCATT
TATATCGTTACTACCAATTACCGAAAAATACACCAATACCAAAGACGCAACCATTATTGCGCCGGTCATACCCTGTGTCC
TTATTGTCGCACTCATGATAGCGGGCTATCTGTACTACCGTGCGAGAATAAAGACACGTTCTGATCACGAATATAAAGAA
CCACAATACACCACGGTTGACTGAACATGACGTTATTTATTATTAGCACACGGTCCCAGATTGTGACATATGTTCTCATT
TTATAGCTACCTCGTGTACTACAAGCCCTTAACTATTGCACTATTGAGTTATTCCACCGTACTTTATTGCTCCCGTCACT
CACCACTGCACTAAAATTATTCTGCACATGGGACCGCAATACATGACGGTGGAATAAACTTGTAATAGCAGGAAACATTA
TAATTAATCTTACATGCTGGATACTAATTATATGCTGGAAAGCTAACCATGAGCCAAAATGGAACGTGCTGGGACAATGA
CGACCTTATCGATGGTGGGGCAATGTCTGTACTTGTATTAATTCTGGTCGTCTTGTCGATTATACTGGAAATTCAGTGTA
TTGTTCATTTCTGCCGCTTAGGCAGACTGTATCGCTGCTGCTGCCCGCGCGGTCCGCGCTCGCACGGCGAAGCATCATTT
CCGTTATCGTAAACAACCGTTGAATTTGCACTTTGCTATAAAGAACCGAGTCCAGGGTAATGGAAAACAGTAATAGCACT
ACAGTTGATGGTGCGGACACCATCCCAGTCTGGTCATGCTTTATACTCCTTATAATTCTAAGCACCATGTTCATAATTTT
AGAATGTATGTGTATTACAAAACTGTGCAGGAGAATGTTTAATCGTTGTCAGTGGCAATTCAGAACCAAGGGCATTGCAA

TCACATAAACCCGTTGAAACATGTCTTTCGCGCTGCATGTTTTACTCCTATGTGTGGTTATCGCTTTTGTGGTTGGAGCG
GTGACTCTTGTCTTGGAAACTTTATATGTAATTCTTTACTGCCAGATACTCTGTAACTGCTGCTGCTACTTTACCCGCTG
CTATCGACACAAAGAATCGCGCCGTTTGCGCGATCCTGATGTAATGAAAAATTGATTATGCATGCCACCCCCTGACGGCC
CATAACGCACATATCCATATACGCTGTTAAATATGAAGTTAACAGCAGTTATGTCTCACCCGAGAATCGCGAGCCTCGGG
ATCCTAGCCATGTCGAGTCAGCTAGTTCTGTTTATTATCGCTGCGGCCTTCTCGGCCACTGTGATTATCCTAGCCATACT
GCTGGATCTCCTGTACCTACGTGTTTATCATAAATTCTGCAATTATTACTGCTGCTGTCCGTTCCAAAGGCTGACGGACA
TAGACACCGTGTCCCTGGATGGCGATGTGTAGTGTGAATGATTGATGGATAGGTTTACGGATGGGTTTATGGACGGATGG
ATTGATTGACAGCGTGACGTCATGGATTACAGCTCGGTATGCTTTTTTATTTCAACCGCGGGAGCGGCCTTCGGTTTCGA
GTGCAACGAGGAGCCGTTGAGAGCCGTCATGAGTTCCTTGTTAGGAGCTCGGAATGCGGCTAACGGACATCCCTTGACGT
TGCTACAGATGAGAGACTGAATGGTGGCGCTCCTGACGATGTTGAACCCCGCGGAGCTACCGACTGTGTTGCGTTTCCAG
TAGCCCGGGGAGCAGATGACGTTCGCCGCGAGTCCCCTCAGGGAGAACGGAGCCCCCGGCTCGAGGACGCCTTCGTCGAG
GATGGCATCGGGACACGGTTTTTCAGCGAGGAATCCGGCGTAGAGTTCAACGGCTTCGATATCACCGTAGAAGGCCTCCA
GCTCGACAGCGATCTCCTTACCTACGAGCGCCATATAGACTTTTTTTAATTCATGTTTGTAACCCCCTGCTTATCGCAAT
TAAACCAAAAAAAATGAGTAAAATGGTATGAAAATTGCAAAACTGTGTTACCTGTGAGTTCGTTCAACAACTTGTGGGACT
TAAACCCGAAGCGTTTTTTATACTCATTTAGGGAGCGATAACGCTTTTTTGGAGCCTGTTCTGAGGCCTCCCGGGAGGTG
TCTTTGAGGGCGGGGGGTATGCGTTTACCACGCGAGGAGATCTGCGGGCGAGAAAAACATAAGGCGCTGAGATTGGGGAT
ATGGAATAATACATAACACACAGACATGAAAAAGCGTACACGTACGTTAATGATTAAACGAGTGATCTGGAACAGCTGTT
CGTCGTTCCACTCCGGGTGCTCTTGCTTGAGCACGTCGCAAACCCTGTTGTGCTCTCTCAGCCAGATAATGGCGTACACC
ATGAGACCGGGCACGAGTCCAAACATCTCGTTCCCGAGCGCGAAACGTAACTTTTCGGGGACGCTGGGTTCGTACACCAT
ATGCACTTGGGCCTCCTTGACGGTGGGAGGGTAGACTTCCCCGTCAATAATCTGGTACTTCATTTTACCGTCCTTGAACA
GACGCAACTGATGCTGTTTTTCCAGCGTCTCGCCATAGATGTGGCCGAGGACGACCTGGTGGTGAAGGGGGAGAACCGAA
AGACACAAAGAGAAGCCGTGACTCACTGCTTTCGTGAAAGGCGGTCGTGTCGCATGCTGGTTCCTAAAAACAGCTTTTTC
CACAATCTGTGTCGAATCAGGCAATAGCGCCTTACCTGATAAGCGGCAAAAGCTAGTGTCCACGCTAATGTCGCCAGCAA
TCGCATGATTGACATGCAAGGGAAGTTTATTCACCTTTCACCCCGAAAGGATTTAAGCGATCCCGAGGGACTAGAAACGG
TGTTATGGTATAATAGGATAGGTTGGAAAAAGCTGTCAATCATCTTGGGAGTTGACTCTGAGCGGCAAAAGCCCGGTGTA
ACGCTGTATGAGGTGACGAGCAATAAAAACAATTGCACGTTTGAAATTGTAAAGTGTTTTTATTAGAAAAGTTGGCCGAC
AATCAGGCGGAAAAAATATTGAATAATAGTTGACTCAAGGTTACAGTATGAATTCTGCAAACATAGGGAAGAATGTTTTA
AACAATGCAATATATGCACATTACTTGTTATACCTTGATTCCCACGTTGTTCCCTTCTCAAGCAATGGAACAAAAAACAC
ATAACATTTCGGTTGCATTATACAGCTCAGTAAACCTGACAGCTCCAGATAACGGTGAAAACTTCCATTGGTACGAATGG
ACAACAAATGCTACATCATGGCACTGTGACATTAATTTACACTTATGCTACATTACGAGCACATCAGATTTAATATATCC
AATTAATTGCCGAAATATGTATTTGCAAATAAACTGCAGTAAAACACAATTACATATTAAAAATATAACAACTCAAACCC
CATTATACTACACGCTAAATAAAGCTGGTATTAATTTAGAAATAGCACAACGACATAACTTTTCTCTAAAAGTAATACAT
CCCACCACAATTTCACCCTCTCTACAATCACTTACTAAACGCACATTTCATGCTACTGGCACGGCTCACGCAGAGGCTAG
GCGACAGTCAGAATCCACCAACAACGTCGTTACTCCGGTTATGGTTGTATTTGGAGGTGCGGCTCTCGTGGGCATTGGTT
ATCTATATGGCCGAAAGCCAAAAATACCCCTCTTGAATGCTTTCTAAGCACGCGGTTTAGAAACCGCAGATACCAATTCA
CCGTACCCAGGGAAATGGGAGGGGTAGGATAAAAAAGAATGCACTGTACATTAGTGACCGTCAGAACCGCGTTACCAACG
ACCACCACCATAACCTCGCGTACCGGACGATTGCCAACACGGTGAATGTGGCGGAGAACTGGTAGGTAAACAGTATGTTA
CGGAGTTTATGGTACTTTGGATACATGCAAACGGTTCTGGCGGTTTCAGCGGCCAGCGATCCAACATGCAACTTGGACTG
CAACTGTAACAGTACCTGCGGATTCCTCTACAACGTGACTAATGCGGTCGGTTACTATCGGAATAACGTGACTTTGCATA
CAAGCATCAGCCACAGCAACCACTCTAACATGCATGTGGGTATGTGGATACGTTACAACTTTCCAGCCGCGTCGTATCCA
TTATGCAGCGTATCTGGCTACAAAGTCGCCAAGGAGCATCACAAAGGATGGTGTTTCGAGTGCAACGAAACGACTTTAAC
TCTCTGTGACTTGGATACCAATCAGACTGGGAGTTATATTTTTAAAAATCTCGCTGGATTGACACAGCACTACACTGTGA
CCGTTGTACCTGTCCCGCGGCCGCCGGCTCCTAAGATTACCACCATCACAAACTGTTCTGTGATATTTTTCAACGAATAC
CTGTGGAAAAATGCGAGTCTCGCTTATGTTACCACCGTTCCACCCACGACCACAACGACAACCACCACCACTTCGAAGCC
AACGAGCACCACCCACCGCACGACCGCTGGTAGAACGATGACAACACACACATACACTACCACAGCCGGCAGCAGCGCCA
CCACGTGGTCAACAAACATGCCAAAATTTATTTCCAAATACTCTAAACTGGCCACGTTCGCGTCCATCTCGGCCGGATTA
TTCAGTTTCGCCCTGGTGGTGTTCCTCCTGGTGTTTCTTTTTGCGCTCTTTAATCTGAAAAAGCAAGGCGACAAGGAGGA
TTCTGGTCCTCCCAAGAAGTCTTCGGTGAAAGACAGTTGTAAAAAGAAAAAAGGTCGACCGGGAGAACCCATCTATCACC
TAATCACGGAGAACATTCAAACGTCCACATCTTGTGCCGTGGAAAAGTCAATGTTTTCCTAGAGTAGCTGCTCGGTCCCG
CCCAAGGTTTCGACGTGAGCGCGGATCGACGAGCTACCCTGAGAAGCCATCGCGAATGTTACATGACCAGCCACATGATG
ATGAACACCAATCCTATCCAGATTCCGACGCCCATCAGCACTCGCAGTAGAATGATAACATCCCAAATAGTAATGGTGCC
CAACGGAGGAGAAGGCGACGTTGGGATAACAAGGGAAACTGCGTCCATGCCACCAGACGACGTCATTAATCTGAAACAAT
AGAAAGGCAGACAGAAACCACATCAAAAGAGCTGCTTGGATAAAACCCCAAAGCCCATGGAGCGCCGCCGCCAAGCATGG
AATACACAGAGATGACATTACACCGAGCTCCGAGACTCAGACTGGTTTTTTACTTATTTTGGATACCTCTCAGCGGTCCC
GTCGTACGCGGAGCAACGGCCGATTCAAGCGTGAACACAACCGAGGCAACACAGACGAATACAAGTAACGCCGGCGTCCG
AACCACCACACTGCCTCCGTGGCCTGTTGAAGTCCAAAGCGAATATTCACTGTGGTCGACAAATACGAAGTGTTACTGCG
GGGCCTTAACTGCGTCGGTGGGATCTACCGTCACATTAAATGGAACGGAACGAAGTAACGAGACGTACACAGCGTGGTTT
TTTGCACACAAATGGAGCGACCTACTGTGTGCGTACTGGGAAGATGGTAGCGGTGATAGCATGCACCGCGAATACCTGCG
TTATAACTGTACAGCAGAACAAATCACTCTGTTTAACTTGTCTACGGCTAACAACGGGGTTTACTACAATCGCAAAGGCC
CTTCGAGCAAAATCATGAACTATACATATATGTGCTACAATGTTACGGTAAGCGAGAACGCCACGAACGCGACCGATCCG
CCGGTTATTACCACTACCAGATGCGGACCTCCGCCATGGAAACTCCTGAAGTACGAATATGACCTAGATAGCATTACAGA
CGATCCAACGTATGCTATAATAAGCCTTGAGGACATGATTAAATCACGAAGCCACGTACGAAAAGGCTTTGTTCTCGAAC
ATACTGCAATGGTTATGATACAATATAGCTGGTTTTGTGGTATTCTGTTCGGGGCGATGCTAACGATTTTGATTGTTTTA
CGGGTTCCGCAGAAACTTTGTTATTTGTGTTTGAATGCAAAACCTCTAATTCGCGGTTACAAAAAGTTAAACGGTTATGA
TTACTAGCGACTGTCCTAATAAAAAAACCATTTCCCTTTGTATGTGGTGAATAAACGTACTTTAGATACATTATACTGCG
TCTGTTTTTTGTTATCTTTGGCGCATAAAGAAAAGAGACCAGGAGAAGGCATGTTCACATTCACTTTTTTAATTGCTTCA

ACTCAATACTGTCACAGCTCTTTCTACATGCCCACATGCACACTAAAAACGTGCTGCGTTGGCGATGAGGTTGTCCTGGA
ATCCCATCTTCCATCAGCATGCAGCGTAATTACGTGGTATCGACATAGGAACAAAAGCAACATATTACTATGTCGATTTA
CGGGAGCATCTGTAACTTTTACTTCAAAGGAACGGATTTCTAGTACGTGTTCGCGGCAATCATTTACATTTCATGGTATC
CAACTACCCTCAACGGGAACTTATTACACCGTGGGCAACTCTTGTAACCATCATCCTACAATGTCCGCCTGTTACAACGT
AACGGTTCACTCAAAATCAACAACTCCACAGAGTTTAAAATTAACTTCGATGATGGCGCCGATGATTCAACGATATACCG
GCAACAATACCCAAATGATTCAGCAATATACCGACAACAATACCAAACTTGATCGAACTTTGGATGTTCACGGCCGCTTG
GGGACTTGTGATTGTTGCTCTGTTGATGCTGTGGGTTGCGATTGAGTTTCGTTTGCCTCAAAAAATGCTTCGATATTTCA
ATACTCGATTTCGAGCATCAACCCGCTCAACACAAACAGTTTAACGTGCCTTCCATGTTAAGATTTTTATTAAACATGTA
TCGACAACTTTTAACATTGGTGTTCATTGTTTCTTTACCGCTAATTAACGCAATTTGTGTAAACGTCAGCAATGGTAAAG
AACCATGTGAATACAAAGACTATACTCCCACTTGCTTGCTACACTGCAATCGTACAATAGCCACAACAGGACAAAACGTA
ACTCTTGGCGTAGCAATTAGCGCTAACAGCTATGTTTTCTGGGAACGAACAGTGAAGAACGAAACGATAATGCTTTGTCA
GTTTGGTGACGACTTTACATTATGTGGATTCCGGCACAAAATTCAATATAAATGTTTATGTAACTACTCTTTAGTGCTCA
TCAATGTGACTACAGAAAATAATGGCCCATACTGCATGTCGTATATTGAACACAGTAATAATTATGTGGAACTTTGCTAT
AACCTCACTGTACWACCGTTTCCCGCACCAAAAACAACTACAAGGCGGCCCAGTAAGTCGCTAACATCCACTACGACAAC
AAGATCATCCCTAACAAGCACCACCACACAAATCTCAATGGAGCCACTTGGTTTTGATAAAATACCATTGAACCATACAT
TTGAAGCACAGCGAGCGACGGCCAGCCACATGTTTTGGTTACTTCTTTTACTGTTAATACTACTCATAGTTATGCTATGG
TGTTGCCGTCTACCTCGACAACACAAGTATCGTCAATCTGTATGATGTTTGTATTTGTAAGCATACATCGCGACGTGTTC
ACAGCAAAACCTTTAGAATGCAAACCCATTAACGGCACCAGAGGACGCAWTGTAACAATGAGTCCACCAACGTTCAATTA
TCACACAACGTATGTATACTGGYATCGACATTTAGGCAWAAATCGATAWGAATTATGTCGATACGACCCCCGAAAGAAAC
AATCTAACGCTCAAATAAAATTTGCATGTTTGAGCAACTACAGBTTATTACTAATTAATGTAACAGCATACCCTTACAAT
GGAAACTACACCGTCAGTTTAGATCTTGGATTAAGACATTTATCTGAAACTTGTTATAAATTAACCGTGAGGAACGTACT
CGCAAAAGTACCCACAACCACCCGAAAAACCACAATCGTWWCTACAACTGCCACGACATCTGKCCACACAACATGCACAA
CTGAATATCCTACTATCACTCTGYCGTCAGAAACTGCACCCCTTATTGCATTATTGGAAAACCAGACAWGCTCTGCGCAG
ATTCAAAATCACTCCGCTATGCATGCCATGTGGATCCTTCTGTTCATTATCATTATTGTCCTTTTGTGTTTTTTCTGCAT
GAGGCTGAAAAATAAGCATAAACCATGCAGCACTGTTTTAACACCTATGGYGCAGTCACCTGAACCTCCCGACCGATATT
TTTAGAAGAATGAATTTGGATAATTGGACATGTATAATCTGGGTTGTTTGCACAATCACCATGTGTATTCACAACACGAA
AGGTAKAAATCAGTCTGACAGCAACAGTACAGAACCTGTACCCAGYAACATGAGAGTTCGTACTGAAGCCAATTATGTAG
TAACTGGGCGGTGTACAACCATCACCCCACCGATTATCGCATACAAACACACGGACTCTAGTTTTTGAACAAAAAGAAGA
GGACGGTGCACACATAGTTGGTGTTGSTGYCTTCGKAACGYGCTTTTATTATNATTGTCATACTGTTGYTCTTTAGGATC
CCRTCAGCAGAATTTTTGACAGCTATCAGAGAATTAGGATCAYTGTCGCCATACTAAGCACATtCACCGTACAAGCTCGA
TAACGGYGAAAACACTGTTTGAACGTTGAAAATGGAGCTTCGCACAACCCTGCACTATGGTCCTGACAACAATACAATGG
AAGCACATTTCGTACGCTTGCTTCACTTTTGTACTCTTGTGGACCGTTTGCAATGCAGCTGTATCACGTGTAACATGCAA
CACCACAAAAACGTATACAGGACGGGTTGGGGACAATATCACCCTTAAAACATCACCTCTAAATGGTAAGGGTTTTCTAT
ACTGGTATCAGTACCCTTGCGGTAGTGACAAAAGACTATGCTGAYTATTCAcGGTTACGATGGCAAAACCGATACGGTCA
ATAACACCAAATTAAAATTTGCTTGCATAAATAACGGCTCACAATTAGTACTCATCAGTGTACAAGCTAATTATTCAGGG
CAATATTGTAGCTCATTGTGGTCATCGTCTGGCAATAGTCATTCTTATACTTGTCATAATGTAACCGTAAAACTACCAAC
TCCATCAACNTCCGTCTCCCTTGCAWAGACCTCCAAAAGGTACGGCAAATAGTTACACAACTACATCAGGATTACAGYTA
CAAATACCCTCCAACGCACCGTATACAGGTAATAWTGGAACACACAATCCAGAGGCACAACATCTWGATGCTACACA

UNIQUE LONG INTERNAL SEQUENCE:

GCTGGCTAGTCCACTGTTTGAACATAAACCGGTTTNTTNAACCGAAAGAACATGGCCTTGAGCATTTGCCAACACTATGC
TCAAAGAAGAATCRCACATTAAARATGACGTGTGGAAAGTTCTAGTAATCGTGTTTATAGCGTTTACAATACATCTAAAC
CCTTTGTATATTGGCAACGTACCAGTCACAGTTAAAACAGGAGACAACATAACCCTAAATGCCACTATACCGCTAGACCA
TGGAGTGGGCTACTGGTATCGCAACGAAACTACATTATGTAGATGTACCATGAAAAACCCAGAATATTCACAGAGTGTAG
AACGCACATTTATATGTCACAAAAATTGTTCACTAACCCTTCTGAATGTAACTAGCCAGAGTTCTGGAAACTACTGCGTA
ATGGCTGAGCCAATGAAGGACAACCCAGAAAACAAATCGTGTTACGACGTAATTGTTCTGCCAACAAAATTTAGTAACGA
TGTCACTTCACAACCATCAAAAACAACGTTAACCAGCAGAGCATCTCCAACAACTAACCTTAAAACTACCACGCGTACAA
CTACGCAACCTTTAACTACTAATGTGTCTACCCACCAGCCAACAACGCCTATCATTACTGTGAATACACATACAACTCCA
ACAACCACGCAAGGTGTGACTACCAATATACCAACCCATCCAACGACTACTGCTGCAACAACCACTACCGGTAACAACGA
CCACGACTAAAACTACAAGAAAAAAGAGGAACAAGAAAAACAGAAGCACCACAACCACTACGAATGCCCCAACTACTGCG
CCATATACCACCACCCCCTTAAAAACGACATTAACTACTGCTGCTACTACAAAAGCGCCATTAATGTTAACTAATACTGT
GACCACTTGTGCACATACTACTGAATCCCCAACAACCAGGTTAACTACCACTGAAACGATAACTACCCCTTTGACAACTA
CTACCCATCCAACGACTACTACCCCTTTGACAACTACTGTCCATCCAASGACTACTACGTCGACAACCACGTCACCTACC
ACTGTGcTTACAACGACATTCACTACGACTATCCCCACGACAACTAAATTAACTACCACAACGACGTTGACCACAACCAC
CACCACTAGGCGGCACACGACATTGACCAGTACCTCCACCACATTCACCGCCCTCGCCACGACTACACCAACTTCCCGCG
CAAACACACCTTCCACAAACTCTCCCGCTCACACGCCCCATAAGTATAACGAAACAAACCACCACCACCTGAAACACCAT
CATCACCATCACGCAGCCACACATGCGTTTTGGGTTTTGGTACTAATTATTGTAGTAATCATCATCATTCTGTGGTGGTA
CAAGATTCCTCAACGACTTCTACAACATTATAGGGATCAAAAGTATGGCGGCGTGGCCGTAAGCGAAGAGTGGCAGTTTG
CTTAACACCAGTTTGACATGACAGCCGAAACACTACCAAAATAAACAAAAGCTTCTCCGTCTTGGTCGCCACACACTTAT
TTGAACTGCGTATACAAAACCGTCTTTGGATCGGCGCCCTCGCCCTGTCACGGGGCCGCCGTTGGGGGCCATGGCCCTGC
ACAAACTCCATAAAAAGTGGCGTCTGTGGCTCCCTGTAATTCTGGGAAGCTGGATCACGCTACTGAATGCTCTAGAAATT
GTGGACGAGTGTTCGGAGGAACAGAGCATCATGCAAGTGGCGCTAAGGCAAGCCGCGGTGAGAATGGCAGCGATACCGGT
GACAGCTAGGCGATCGATGGCGGAGAGACCGAGACCGCCATCCATTCCACAATATCTCGACGTCCGATCTCCGCGACCCC
CGCGACACCACCACCATGACGAAGATCCTGTGTTGGCCAATATTCGCTGGCAACACGAGCGAATGCGAGTTCTTATGCAG

ATACAACGTTCCCGCTCGCAATCGCGACCCTGGGCACCGAATCTCGAGCTGCCATCGGTCAGGCTGCCATCTCAATTCGC
CAGACCTACGTTGGGTTACGGCCCGATTCCAAGTTTTCCACCGGCCTCATTGCAAGCAGCGGCACCACCGATGGCAGCAG
CAGCGGCGGCGAACGGCTCTCAAGAAACTCCGAGGGTAACGATTGTGCGCAACAAACGGCCCGACAATGCGCGTCCCGGG
CCGCGGCTTCCACACCATCGCCGCAACGCCCTCACTCTCACGGGATCCACGGATTCCCTGGATTTGACACGCTTGCGCCT
GTGTCATGAAAGGCAAAATCGCAATTCCAGCTATCGTTGGAAGCATCATTCCACGAGGACACCTCTACGGGCCATGTCCA
TGTGCGTCTCACCGCCGCCACCATCACCAATTCCGCTGCAAAGGCCCCATATTTCGGAACCCATGCAATTCCGACAGCAA
GAATCACCAGAACCACCAAAAAAACGCAATTTGAAAAACCGCCTTCAAGACAAGTGTAGGCAAGCCTTGCATCTCAGGGA
AAAAGACCCAAATCGACATCGGTGGCACGTACGACGCCGTGATTCGCTTGACTCTCTGTCACTCTACGAACTTGACCATC
GACACCAGGCGGCCCACTCTTCTGGCAAGACACCCAATCGTGGCTTCGAACCCACATGCCAGGGAATCATGTGTGGCGTC
TGCGTGAATCCAGCACGCGCGGCAACGATCTGAGCACACCATTATCTGCAGACACCCCAGCGACTCCCACCGCCCCACGT
TTCACTTCCAGCCGATCCAGACCGCCTTCGCCAACCGAAAGCGTGACGATTGATCTTGCGGCCGGAGAAGTTGTAGTGAC
ACCACTGGATGGGAACAACAATCCCAACGCCGATGCCGAACGCCTGCTGCGTAACGAAGAAGCGGCCACTACGGCCTCCA
GGCAGATGTGCAGCGCGGAATCCATCACGACACTAAATCCATGCAACGGCGAGAGGCATTCGCGACGCAAAACGCTCGGT
GTCAGTGTCGCAATGTGGCTTTCCAGCCACTGGCTGCGCCATAGCCGCAACTCGTAAACCGACAACTTTGACACAAAACA
CTGACAGGGAGAGGTGAATCGGTGAAGACGAGAGACGTTTTTCCACTCTCCTTCTGATGATCCGCACCGCCCTAGGCTCT
GGACAACAGGATCGCGGTTACCTGACTCCACATATTAGCTCACCGGGTCCCGCCCTAGCACCTCCCCTACCTTTTCGGCT
ACAATATATATCTCACTTTAAACACACTTTCACGGACTAACTGCACTTGAACTGTTGACTGACAAAAGTAACCGAAAAGA
TGTCACTACAATCTATAAACGTCTTTTTGCTCACCGTACTAACGACCGTTCACTGCATCCCCCCTTCGTACTGGGATCGA
TTGATGCGAGTTTCGCCGAGCGACACGGACATCGTTTCAGCCCCCACCATGCCGCCGTACGGGGACCGTTTCAATTTAAC
CTGTGAATTCCCCACCGGCGGCTGGACGGAGGCCGCGGTTCAGATTCGCTTTTGCAACACCCCGAACTGCCGCAGCCTTC
TGTTAGTTAATTCCAAACAGCTGACCGGGGAGTTTCAGAACAAAAGCCACGAACAGCTCCAAGAACTGCGGTGGAAGCTG
GAGACCGAGGGTCCCGTACAGAGGTTGACGGTTTCCTTTTTGGTAACCAGCAACGTCTCCGGAATCTACCAGTGCGCCAT
TGTAAACAAGATGGACATGGACGTTTTAAAGACCATGGTAGTAATCGCAGAAGTGGAGCTGCGGAGACGACCCCACGTCT
ACTGCGATCTCCAGTTTGGCGAGCAATCGCAGACGCAATACCTGTGGACCCCCGACCCAGAGAAGGTGCGATTGGTGAAC
TGCGGAGAAGTTGGTGGACTAGTGAAACACACGTGGCATAAGCACTTGCATTATGCGGTGGGGAACTCCGGCATGCAGCC
GCCATGTGATTTAGAAAGAGATATGTCGTTGTGTCACAGTTATGTGTTTCAGAGCATGAGGGACAATAATTGCACGAGGG
CCGCCGAGCAGGAACGCGCACTTATCCTGCGCGGCGAGCCGCTCGAGAAGGACGAGTGGTCATGGATCTGTGTGTCAGGT
CTGCCCACCCTGTGTCTGTTGGGCATCCTGACGTCGTGCGTGAGGTTCCGGCAGAGGAGAAACCGACGCTGGCGCCGTAA
GAGAGAAGATGAAGAATCCAGAAAGTCGAAGAAAACGAATTGAGTGGGTACACTGTGGTGGGTATGTGGTGGGTATGTGA
TGAGAGTGCAACACTATTGTAGAGAGAGATTGATGGACAGTTGTCAATGATATAAGCTATTAGTACTAGTTTAGTTTTAA
ATGGAATCAGGGTAATGCGTCAAGGTGACGGGCTGCTCATATTTTGTAGAAGCAAAAAAAGTCTGAACGTGACTAATCGG
GTACGTGGAGGGGATTCTATGTCTACATGATTCACAAATAAAAATTTATACTGTTAAAGAATGAGTGCGAAGGAGTCTTT
TATTCGCCAGCACATAAGTAGGGGGTAACTCGGAAAAATCATAATAAATCATGGTTGTCCAGAGGCTACCGGGACGGTCG
CCGTGGATCTGGAAATGGACAACACGCCCCCCACACACACCGCCGAAAAACGCGGTTATGGATAATCTGTCCACTCTGTA
CAGAATGAACGCGGCCCAAGCGCCGCAATGGCAGACGGAGGACAATGTCCAGGAGCTCATTAGCGTTTTGATACTGTGTA
TCGCCGTGTTGCTATTTTTTTGTTTTCGAATTCCTCAGAAGATCGTGTTCTTTTTTCGGTCGGGGACTTTCCGTCACTCC
GGGAATCGTTAGTGCAACCATGGGAACGGAGTCTGCGCAGGGGATTATGTTTCAGAATGACAGCCAGGCGCGGCTCCCAG
CTGATTTCCGAGAAATGCCGTCCCCGGACCGCATGGCGACTATTGTGATGGTGATTGGAGTGATGTTGATTTTTTATTG
TGGAGTATGCGTATTCCACACCACTTTTTTCGGACATTGCGGAATTCACAGGGGGATTAGCGGACCGGCAGACACCCTCC
TGGGGGGTCCCAGATGGACTGCGCGATACTATCTGATTTCCGAGGCACGCCATGAAGCGCCAGTATAAGAAACTGGCTGC
AATTCAAACCGGACTCATGCATCTGTAGCCCGTCCAGCGAAACATGGATCAGGCTCTGTATGTCCACTTTTTTGGGCGCC
CGCAGAATGAATCCATGATGAAACACGTGCCGTTGCGGGCCGAGGGCGATGAAGCAGTGGAAGACGACTACCCACCGATC
CCGCAGGTCCCGCAGACGGCCCCGCAGCGCTCCTGCATTAGCAGCGGCTGGCAGACCCTGAGAATGTACACCCGCGAGAC
GTACTTCAAATGGATCAGCGGAGGAGAGGATGATAGAGTGTCGCTGACAAAACTACCCAACGAACAGAAGAAAAACGGGA
ATAAGCTGAGCACCATCGCCCCGAAGCTGCGAAGAATGCTCAGTAACAATTGAAATCCATGACATGTCAAGATAACGAAT
GATTCACAAATGATTGACAGGTTCAACGAGGCGCCATGCGATGGTTTTCTTATGCTAATAAAATCGCTCCAAAAGGCTCA
TCTGTGAAGTGTTGTTATTGGCTTGTGATTTATTTCGGTGGTCTGCCAACCATTGTGCGCAACGTCCTGTCGTTCTGTAC
GCTAATGAGACACGATGTCGCCTATCAGCCACTCCTACGTCGTTAGATATAAAAGTAGCGTGGAGGTCGTCATTTCTCAG
ACCACCCACGTTGCGGTCAGACGGTGGTGGCAAGACCCCGGGGATCGAGCGGAGTCGGAACGCCATCTGGTGGGGAACAG
TGAATATGTCTCACGTCCTTTACGAATACTTTAAACGAAAGCTGCCCAGAAAGCAGGTGAACAGCGTGGTGTTGCTCAAC
CGCAGGACGCTGACGGGGGGCTTAGCGCCCATGACCTCCATGGCCACCAACACCTACAAGACCATCATAGTCCTGGGCCA
CGGGAAGTTGCGGTTGGAGAGGCCCAGTAACACGGAAACCATCCAGATGGCGCTGACGCTGTCATTGCAGCGAGCCTCCC
TAAACAACCCACCGGGTGAATCGGACGAATGTCCATCAAGCGGTGACAATGATCATGATTAATATGCAATGAGTCAGAAT
GATAGTGATGAATAACGATAATGAAACCTCTAAATGACTGCACCGTCACTGATGCGTCAAAAGTATGTCGAAACGGTGCC
CGTGCCAACACTTAAAAAGAAAACACCGTCAAGAGAGGGCACTGCGGCAAGATGCCTCATCAGACCTGGCTTGTGATCCT
GTCGCTTTGTGCCCTGACGGTGCCAACCTGGAGCAGCTATGTGTACGTCTCTGTGTTGACTCGTTCCCAGGAGCTGACCT
GGACGGCCATCGGGGGAGTGAACGTGACTAACACCAGCGAGCTGGTAGGCACATACCAGTTCACCCGTAACAACGAAACG
CAAGGTGAAAATGTTTCGTCCGTCGGCACTGGCAACGGCAGTGTGTCCCAGCTACCTGCTGAATGTTTACAGCGTCAGAG
CAACTATCTGGGAACTGTGTGGATTTTTTGCTGCAACCAAAATCAAACGACGATTAATGAGTTCTGGTACCGAGGACACG
GGTACAATCACAGCAAAGAAAACTGCACTTCGCCCTTGGTTACTTACCTACAAACAATGTGTGGCCTCTGGAAAAACCGC
ACCGATGCCAACAAAACGTTTGATAAGGCTACATTATCAGAGTCGGAGGTACGCTGTATTGTTTCCCAAAACTACTCAAT
CAACGACACCATCAAACATTACAATGAAACTGGACTGTCACGGGACGATCAAAAAACGTACCGAGATAGGTTTCCGCACA
ACAGCACGTGCAGCTTGACTGAAGCTTtgatagtggatgaactgaaAAGCTTTATTGCCATGTCCAAATTGTTAATGTTA
ACTATCGAAGGCCTGTCATTTGTTGCCGTTGCGATCTTCGCCACCATCGTCTTACTGCGCTATTATTATTGGAACAAACC
TCCATGTGCTTGGATGGTACCTCGCAGCCCGCGCGGCGCCGACTGTGAAGAAATTGTGATAGAGTCCAAAGCGGAACAGA
CTGAAAGCGGGGAAGATGATACAGCCCCCAACAGTCCTGCGGAACCTGAGAATACAGAAACCACACCCATACTGTCCGCT
TCCATGAAACCCCACAGAGATAAGCCCCCCGCCGTGCCACCTAAACCGACCACCTTTACTATTGACGCAGCCAAGTTAGC
GCAGAAACCGCCGCTGCCACCAAAACCCAAGCCGAAGTCGTTGGTGTTCCAGTATCCCTCACCCAGATCAGCGACCAAAT
CGGCCAAGGCTACCAAAAGTCCAAAGGTATTCACATTCAACGACCACGATATTAGAAAGCACAGAGAAGAAATCACTAAA
AGCAGAGTCTTGTATCCTGCACAGCAGCCAACGACACCTGTGGACAGCGTGCTGACTCCACTCCTGCCGCCGCCACCACC

TGGGCCTTCTATTCGACGGGTATCCGGATCGCCACGAGTGCCAGTCATGATTCCCTGGGACATTACGCAACCTACGTCTC
CCATGTCAATCAACCCATGGACACAGAGAGACCGACCGCTGCCATCAGAACTGGATCCCTGGGCGTTGAAGGCCAACCGG
TGGCTGTGAGGACTCAGATAACTGTCAATCATCCCACATGCACACACTACAGACTGTATATAGCTAATTTAATTTTATTT
CTGTATATAAGATTTATGGGACGGAATGAATAAATGTATATTTTTCTGATATCAAAGTCTAAGCCTCTGCTTCTTGGCCC
TCGGCGTTCAACAGCTGCAGCGCTAATCGGGGCCCGGGAGGACGTATCACGGGTTGGGGTTGAATTTCTTCCATCACCAA
GATGATATGATTGCCCATGGCCGGCGGCTGGTTCTCTTGGTCTTCAATAATTTCTATTTCCATAGGAATCTCCATGACAT
CCATGACATCCATGACTTCGGGCACTTGGGGCACTTGTAGTTGAATCTGCTGCCGTTGCATGTAGTTTCGGAACAGTCGC
CGCCGGGCTCGGGGAGCGTATATGCCCAGACGTCCATGAGCGGGTTCATTGAAACGAAAGAACGATGCCCTCACGGAAGG
CATGAGACCCCGCACGGCCTCCGACAAGGTGGGCTCGGTGCTGCCTCCCATAGTGAGTACCAAAGGTGGTGATAATAAAT
TCTGAGCGCTCCACATTTCCTAAATACTAGTTTGTGTCTGACGTCAACAGCGTCCATCACGGTATATTGGACAGGGAACG
GTGCCAGGAGTTGCGACAGCGCCCCCGTTAGTAAGGAAGCCTTCTAAGATATTTAATGACACGAGATTGACAGCCGCCTC
ACTTGAATCGCGTCCAATACAGACATCATGCGGAACTTAAAGCTAACCGCAGCTCTGGTAATCATCGGTTTGGTGGCAGT
CCACGCCATCCCCAGGTGGGATTGTGTAACGTTCTATTTCCTACCTAGATGGTGACCATGTTCGCATAATATGCTAACTG
TGCTTTTCCCACACCCGTACAGGAGAACCACCACCGATGGCGATAAAAGCACACAACACACAGAAGTTGGTACTCAAACC
GACGACTCCGACTCACACAGCACCCATATAACAACAGAAGACGGCAGCGGAGGAGATACGAGCGGAGACTCCAGTGGAGA
CAGCAGCGGAGAAGGCAGTGGAGAAGACGAATATGACGTTCTAATTGAAGGAGAGTTGGAAGGTGATAATTCCGAAGACA
ACAGCGAAAACGATGACAGCTCCGCAGAGTAATGAGGACAAAAAAAAAACCCACCCGAACCGCCGTTCTAATGCTGGGAC
TTATAATCCTGGCATTCTACCTGACATCAGAGTCTACACTCCATCTTGTACTTAGACTGTAACTTTTCAATTGCAATTTC
CATCTTTGCATATGAGACTAAGGAATCCATGACACCTTGTACTGTGATGAATAAAAACATATTGGGATTCCTACCATGTG
TTCTGTGGAATTATTTTCACGACCTCCGAATCACTGTACATACTAAAAACCACAACACACAATAAAATTATGAATATAAA
AAATAACACTTTATTGCTAACATATGATGTACCGTATGAAAAACGATTATTATTGTGGTATTGCGGGGAAAGGGCAAAGG
CAGTGCGGAAGTACAAAGTAAACGTTAGTTGCTCTGACAGCGTTTGGAGCTGGACCGCCGCAGGGATCTCCGGAGCACTT
CAATCTTGTTGCGAATATGATTTTGTCCGAATGGACACACCAGGTTGTCGGTCAGCTTTCGCAATTTCTTGTCCCCGTAG
AACACATTGCTCCAGCGAAAGCGAAACATGCCATGTGCCACAAACTCCATGACGGTATCAGCCACCACATTAATAACATC
TCCCGAGGTGCCATAAAAGCGTAAATACTGGTCGGCAAACACGGTCACACATTTCTTTCCAATCACGACGCGACAAAAAA
CTTCCACTCGGCTGGTGCCCACCGTACGTTTCCAATAATACTTTTCCAGTGCGCTGTAAATACCAGCCGTAGACACGATC
ATCATGTTCTGGTCCCCCAGCTTGAACAAACAGCAAACGTTCTCATCCCTCCACTTGCACAGGTCATTGAGATTTTCGCA
AGTGACGAGACTTCGCAGCTCCTTTTTCCGCAGCACGTTTGAATTGGGAACTGTCATGAAGGTGCATCTCCGGAGCCCGA
GTCTCATAAAGTCCTCTATgTCATCAGCGATCCAGGTGAGGGCGTCGGCGTACACCCAGTCCACATAGAGGTACACGTTA
CCCGTCGCCCCTACATACAGTTCGGAAGGCAGATTGCTACCTCCTTTGTGCTTCAGAAAAGTTTTGGTGCACACGCTGCC
CACAAACTGCAACTTATCGGGACAGCAGAGGTAACTGGCGACGCGTTCTTCTTCATTCTGGCCTTCGCGAATGCAGCGCC
AGCGACGCCCAACGACAAGTACACGGTCTTTGGGCCAAGGCAGGGGAATGTGCTGACCCCATTTCCACTCGCTCGAGCCG
ACGCACATCACAACAGAACAGGCGGCGGTGCCGCTGCTGCAAGCCAtGGTGCCAAAACAAAAGGAACAGAGCCAAGTCGT
GACAGTTATTTATGTCTTTTGCCACGTCAGACGTTTCATGTGCCAAGCCGCACGCTCAGTCGTCATGCTGACGCCCCCTG
GGACACTCGGGTATAGTGTCGTGGATTTCGTGTTCCCCGGTGAATTTGCGATTTTCCAGCAGCTGCCGCAAGCCCAGGCA
CATAAAAGTACTGAGGTTGTCGGCCAGACGGACGTATCTGAGGTTAATGAGGTCCATGCAGAAGACGCGCCCGCCGTCGG
CCATAACGATGGGGACTCTAGTAGCAGACCATCTGAGGGCGGGTGGAGAAGAGGGCGCTTCGCCTCGAAGGTCCACGGTT
CCCAGGCAAACAACTTCATCTCCAAACAGCTGCCGCAGCGAGCGCAGCATGTGTCCGCCATTAGAACTGTTCTCGAAACA
CTTGAGGTTGCACAATCTGATAACAGCCCCCTCGGGGTAGGCCATGTGAAGACATTCGCCCTGATGACGGGTGACATAAC
GTGCTACGGCTTCTGGGCCAACGTTCCCAGCGCGACACAGGCCTCGAAAAGCGCGCTTGACAGCGGAAGAGCGTAGGTAG
CGGGGGCAGTAGAGGACGTCAATGGCACtGAAACCCCACCCGCTCCAGTTCACAGAAGGTTTTGGCCACGAAGTAGAGAG
CAGCATCCTCCTCGACGTAAGTGAACACCTGTCCGTCCTCGGCTGCGAGAACAATGAGGGTTTTCTTTTGCTGACCGTAC
ACCAAAACGGGATCCCCAACGACACGCCCCACGAAGGTGAGTTTCATTTCACAACACGGATACTGGCCAAGACGTCTGAT
GTCTTCCTCGTTGTATTCATCATCGGCCTCCAGCTCCTGCAGCACCACATACCAGCCGCGCGGCCAAGGCACAGGGATCC
TCTGCCCCGAATTGAGACGCACGTAGGCGTTGAGGCACCCGGGCCCGAAATCGGCAGTCATGGAGAGACTCCCCAGGTCG
CTGAGCACCGTGTAGGACATCCCGCTGCGGTCGTTCACACCTTGCGGTTCAAAGAGCTCACACTATATAACTCTCTAGGC
GCGTCACTCACAACAGCCTCGACGCCCCAATGCACCGCAGACGTGGACACCAGTATAGGGCGCACGAAGCCTCTGCTAGG
CTTTATACCTACGGCGATGGCGACAACGATGTGTTTTCACAATCAGATTCACGCAGACCGACCAATGAAATGGATGACAT
TAACCGCATGGAAGCCGGTTTACTGACACCCAGCGATGTATCTGACCATACTAAATCATCCTTCGAACTGGTGTCCGACA
CGAACAGCGACTCAGAAGCAGACGTAGAGCGCGGTCGACGCGGCGGCCTGGCCACCCACAGAGGCAGAATACCAGCCTCT
CAGTACTCGCGACATCAGCAGTACGTGTCCGCACACCACCGCACTAGGAGAGACAGCACAGTGAGCCTCCCCACCAGCGA
AGAGCGGGAGATTGTCAGTCAGAGTTCCGATGAACAGGATTATTGTCAAGAGACAGATGACTTTAGACGCCGCAGCATTT
CCCTGTTTGACAAGATCTGCTTTGAGAAGTCGGTGACCCGAAACCTGGGCTTTGAACCTAACGTGATTTCTCCCCCAAAC
ACCGAGTTTATGAAACTGCTGCTGTACGAATTTTCTATCCAGCACATTTCCAGAGTCAATCACTGCGTCCCCATGCCGGC
CTTTGCACTGATAAGCCTGATTGACCCCGTGGTGAAGATTACCTCTGCCGGCGAGCGCGACCTGACAAGCAGTATCGTGA
CGCACGCGGTGCTGGTGAATTATTATTATAATTCGAAAATAAAGATTCGAAGTATGCAACATTCGCTGACGACCACTATC
AAAGCAGCCACCGTACGAAATATCGCAGCTTATGTTACCAATTTGACTCCGGCGGGCCGTGCGGCCGCCCTGGCTCTGCA
TTTTCTAACGCACGAAAAAGTTGTTTACGAGCCGCAATATCAAGTGTGCTTGCGACGCTTGCACGATGAGTTGCGCAAAC
GGCGAAGCTCAGAATCACCCAGTACCGGGGAGATTTATGAGTGCATCCGTGACTACAACTTAATATTTCTGAGTGCACCC
TACACGACCCGAGGGGCACTCTACACGTACCGTAACAACCTACAGAAGCTCACCAATGGTTATAAGAGCGATCTCAAATT
GCTGTGTAGTGAAAAACTCAACGAAGAACATTTCCTAAATGATTTGACGTTTTTAATCGCGTCCCAAATCATGCTCATCC
AGTTCCAGCGCAATATCGACATTCTTCGTTTGTATTTAATGCATCATTTACAAAATCTGTCCAACTTGATATATTTGATT
TATATCCAGTTCCCCTCCTTACGCAGTGACTATCTACAGGCCACTGAGGCCATCTACCGCACTACCAACAGCCCGGATGA
CGAACCCTTATTTCAAGTAAACACCTGGCTATTTGACTTTTTACGGGCAGTCAGACATCTGGACGCCTTTGTGAGCCCGG
ATTACGTGATGTGCGCTTTGCGACAGTGGACAATGGGCCGGGACAACGAATCCCGAACTGCAGACAATATGCTGCGTGAT
GCAGCTTTACAAGACCCGCTCACCAAACACGGAAACGTTCATTCCAAATGTCTAACTCGTGATGTGTGTCTAACCAGTCT
GCATTCTCGCCCGACCTCGCGCCACACCCACGTAGAACACATTATGGTGCCCTTCCAAACTATCCAGAAACTGTATCAAG
ACAATATTTCAGCGGACGAACGCGAGTGCGACAGCCGCCGTCAAAGTGCACTGCGAACGTTCTTAGAACTAAACGGAAAT
CACCCCGCGAGCCGTTTTACTTAGAAAATATAGCTTTtATTGAGGCGTATATGTACACATTTGGGTTTTTTTTTATTGGG
GTGATGTTTCATCACGATCGCTCGGTCTGAAATCCGGCTGGCCGGTTCCTGGTGCAAATAAAGATGAGGACATGAGTCCT

TCCACAAACATGTCGTGACTGAGTTCTTCTCCCTCTACCATCGCGTTCTGTCTCGATAACCGGGGTCGGGGAACTAGGGA
AGCTGACACATCCACGGGTGGCTGACTGCTTCTCCTAAGTCCTCTGCGCTGGGCTTGCTGCAGGTTTCCAGCCCTGGCCC
GACGAACATCATCAAAATAGCCTGGACGTATACCTTCCTTAAATAGCGCAGATAGGGATTCAGCCGCCTTCACCACATAG
TTAGTATCTCCTACCTCGTAGCAAAATACTTCACCACCTCCGCCGATAGCGGCCACGGGTCGGTCCCAAGAAAACATTTC
ATCAGTTCCGACGGTGCCAATCACACGTAACTGCCCCTTCATGATTGGCAAAAGCCTCCAGTAATCATAGTCACGGGAAG
TTAGGATGGTGTCGTCCGCGTCACACACAAACAGGGTGTATCCCCGGGGCCACATAATGTCAAAATGCCGTCCGCGGTTA
CGTCGCACATAGTCTTCGACATTAAAAACGCAGCCAGTTTCCCGACTGGTGGCCATCCTGATTGTCGCGGTGTCAATGAC
CTCATCCTGCCTTTGATTGATTTCATATTGCACAACCGCCCACTTCGTGTGGTCGAAAACGACGAACATGGAGGCGGTTC
AGTTTTTCCAGATGGCCCTGATAACGCACGGTGACCGCGGCGGAGGCTGGAATCTCAAAGGGCACCTTGCGGCGGCAACG
GTTAGCGTTGTTACTCAGCTCGCAGTCCGCCAAAAGCGAGTTGAAGATCATTAAAAGCCCTGCATCTAGGGGTACGTCAC
CGTGACGAAGGTAGTACCAGGCCAACCTTTCGGATTTCCTAGTGTCTGCGGAGGAAAAGTCATAGACGCGGCGCAATTTT
AAAAGATGTAGCAGGACTAGCTCGCGTTTTATGTTTATGGACACACACTGGTAGTCAGGAATCTTGTTCACTATCTTCAG
GAACTCGCAAAACGTCTCCACCAAGTCATTGACAGTGAATTTAATCTGTTTTAAATCCTGAACGTGTCTCCGAAATCGGG
ACAGCTCCATGCGTTGCGCCGCGTTATGCACTAGTCCCTTGCGGATGCTGAAAGATTCGTAGCGAAACATTTTGATAAGA
CACTCTACATCCCTACACAAGGCTCGCACACACGCGTTTTGCTCGTGGCGAATAGTAATGTACACCAACCTGATAACTCT
TAGGGCCAACACAGCGTTAGACATGATATACATGGCATAACCTTGACAGGGGGTACCGCAGGGACATGCCCACGGGCAACA
GAAAGAAGTCATCAACTAGGGTGTTTGACCAGTCCGCGAAACCCAAGTCTCGAGCCACACTCTGTTGCACACGCAATTGT
TCAGCCGCCGTCAGATGACGAATGGCCGGAAGATGGAATTTTCCCAACTGCGTGTAGTTGAGTAAGCGCAACTCGGCGTC
CTCCAGTCGCGGCCGCGGCGACTCCGTTGAGGTTGCGGCCGCGGCATCTGAACCATTTGGCAAAGGCGACTCGTGACAGT
CGCCGCAAGTACAGAACTTGGCGTACAGAGCCTTTAGAGCCTCTTCGGCTCTCACATGAAATCTCGGGTGAAAGAACACC
GCATTCTTTAGCGCACGATAGTTCAGAGAAAGCTTTTTACAGATAGACACGGGTGAACACTTGTCAATCAACTCTCGCTC
ATCGGGGTTAAGGAACAAACACATGGTGCCGGAAAGACGGCGAAACAATTCATCAGTGAGGACACATAACTGACGCACGT
AATGAGTCTTAGTGCCAAAGATTTTATAATTCAGCAGCGTGCAAACCGCATATTGTCCCGTAGCCACCGCCACAGCGATC
CGCTTAGCACTTTTGCTGATTTCGGGCAAAAACGGTCCAGCGGCCACTACGCGCCTAAACTGAGAAACGTTCATCCAGCT
GAAGCTGCCAGGCTCGGGTCTCGAAGAGACGCCGCGCAGCAGTCCCCACACGGCCAAGTCTGCTTGGAGACAGAGGCCCC
TGATGTTGTACTTTCCGTGCGCGGTAGCAAAAGCGGAGTTGGCTAGAAAACTGCGACGGACGTACACGATGGGATTTTTG
TTGCTATTGTTTTTGGGACGCTGAGATCCCCCCGTCCGCTTCTCCTTGGACTTCTTCTTGGGACCCGGCGCTTTGCACAA
GCCCATCAGGACGGCCTGACGCCGGTCACGCACGGGCGCGAGGTACGTGCGCATAAAGTTGTGACAGAAATCATCGTCGG
GACAGTCCTCGACAAACTCGAACGCGGTGCGGGGATTGGAGTAAAGATCCTCTGTGATCACGCCACCCGTTGCCATGATT
GCACTGTGCAGATGCGTAGAGCCAGTTTGACTATATATCCCAAGCGCACGGGGCAAAGCCAAGATAGAAAGAAAAAAAAA
GAACGAGCGCGTTGGCCTTAGGACACGACCGTGGATGAATTGTCCTGCAACAGATTGTGAGCCGAATACACGTATTCGCT
CACGTCGCAAGGCTGTTCTCTCTGGCTGCTGACAGGTTTGACGACAGAACCCACAGGAGAGAAACAGAAACGGCGCGAAC
AGCGAAACCAGAAGGAGCCCTGTCTGATGAACCCCCGTAAGGTCTCTGCCAGCTTCAGCATCACGGTCCCGTAGAGCAGA
ACATGCACACCACCCGTGCCGTCCAGAAACACGGCCCTGTTAACAGGCTCCAGCCAGCCAATCACGGCATATTTGGCCTT
TTCCAGAATGCTGATGACAAAGCGGTTGGCCTCGTCCGCCTCGGCCACGTTCCATGAACCGAAAGTGAAAGTACATGCCG
GCCGTCCCCCCAGGCGAATCTTGGTGCCAGCGTGCAGCTGGCAGAAACGAATCACGGCGTTGCGGTCGTGCAGCACTTGA
GACAGATCGTACATGTCAGTAAAATTGTGCTGGTACCAAGCATTCTCCACCACCTTGTCCACGTACTCCCTTTCGAAATA
GCTACAGACGGCAAAGAGGCCGGTCTCAAAGAACTCGTTGAACGAGATGGCCAGCGCGTACACCCGATCCTCTGAAGGCA
CGTAGCAAAACACGTGACCATGGTTGGACACCCAGATCTCCGAGGACTGCGACGTGTGTGGCAGGCACTCGTAGACGGTG
ACCATGCCCAAAAAGTACAGTCTCGAAGCTTCTCGAAAGCAGGAAAAACGCTGATACGTCCTGCGATCGCGCGTCAACCC
GAAGTAGTTGGCATCACCCAAGAGGAGGCCGTGCAGGGGGGGCCAGTTTAGAGGCAGCCAACGACCGGCGTTGGCACGCA
CAAACCGCTGCAGATGCAGGCCGCCGGGGCGCTGTCGGGGCCGTCCGCGCGTGCTGATCCATGCaTCTCGCAGGCCTCTC
CAGAgGGCGCTGAATCTGGGCTGACCGCGCAGGTACAGAGACGTAAGCGGCTCAAAGTTCCCCACGACGACCCTGGAAGG
AACGTGAAAGCAATACAATGAATGTGGTCCAACAGCGTAACACACTCAGAATCCCATGTATCGTTTCATGCCCCTACCTG
CGTTTCTTGAACGGCAGTTTGGTCTCGGGATAAACCGCGTTCAGCGCGCGGGCAATCCACGGGTCCGATGTTTTCGTAACG
ATAGTCATCAATCAACGTAGACAGGCCGGTGCGCACAAAAGTCTTGAGGTCCCGGGCCAGGCGCGTGAGGCGATCCTGGC
CCAGCAACAGCGCGTACACCGTGCCGGAGTCGCCGCAGAACACTTTGACATGTTGAAAGAATGTTTTCTCCCCCTCTCCG
TCAATCATTCCCAGCAGCAACACAGGTTCGCGCCCACTGCTATACCTCTGCAAACGGCGCAGAGTAATGGGCGACACGTC
CAGATCCTGCGCGTTGGCCATCAGCAGATAGTGATTATCTGGCCAGATGAGAGCCAATTTGGCGTTGCGGTGTTCCTCCA
CGTAACGGCTCCTGGCCTCATTATCGGCACAGGCCAACAGCTCGCGAAAACAGCTCCACGAGCGCAGGTACCGCATGCGG
GTGGTGGTGCGCGAGCGCTGGTACATCTCGGCCCGCTCTTCGGGCGGCAAGGGGGTGACATCCGGCTGTACGCGCCGACA
GATGCTCTCGTGCACGGTATCGCATCGTATCATGCCTATGATAAAAAAGTCATTCAAGTCCGAGGCGACCTCGTACAGAC
AGTCCTCTCTGGCGTCGTACACAATGACCGCCCCAGAGGGATCCAAAAACACGGGATACTTGTTGGCAAAATGCGACACG
CTCCCCAACGGCTGTAGCGGCCTGAACTCCCCAAAGAAATTATTCTCGCAGTGCTCCAGATCTATCTCCGTGACAGAGTA
TGGCGAGTACAGGCCGCCCACGATAAACACGCATTGCGCGGGCCATCCCAACGATACGCAGGAGCCCTCAAAGCGCAACA
AGTAGCGCTTGAGCCCCTCCGTATCGCGCCTCATACACAGTTCCGATAAACTGCGCGAGCAAAAGGTCTCGGGGATGGAC
GCCCTTTTTCGCCGCACCGAGCGCGTTTGACTCGCGTCGCCCATGATACACCCCCCAGTCCAAACGTTTCAGTCGCACCG
CGAAAAACTCAGAACGCGACTGAGCATTGAGTCAGCCCGTAGCCTATCATTTATAGCCCATCCTACAGAGACGAAACAAA
CGACAACACACTTTTTGATCATCTGCCATCTCGTCTTTATTTAACTCCGCTTCGACTTCATCTTGGGCCGTTTTCCAGAA
CACAGGCTGGTCTTCGGAGGAGAGAGTGTGGTAATACTCGAGCATTCGCTGGAGATCAATCACGGCAGCGTGTCCTCGCA
TCATACGCTTCAAACTTTCCAGATGGCCGGCGCCGTTCTGATGCGTCGCCTTCGGACTTCATACTTATTCAGCCTCCGAA
CAGATGCGGTCGCATTTGAGTTTCTGTCTGGCTTGTCGCAACTGGCCTAATCGCTCACGCTGCACTTCCTGGTCCACCTC
TCGAGCAAAGTCCTCTCGATCGCGGATAGAAGGGTGCTATAAAACTCGCAAAAGCTTCAACATCACAAGCGTTCATTTGG
GGCCTAAGACACCATTGTACTCGGAGTTCCCAGAACCGCGTCCACTGTTCGCTCCGGGCCGCCGTTGTGGCTCCGGCGCG
TGACACCGGTTCGGGGTTAAAAGGGTATTCGGTATTATAAACCGGAGGCTTATATAGAGAACGACATGTCGAGCCCTGTC
ACCCCAGACGAAAACAGCCAATCCACGACTTCCTCCACGGCTCCGTCGACATCCACCGTCGCCTGTACCGACGCCAGCAG
CACCACCGTGGCGTTGCCCGCCTCGCCTCCCCAAACCACGCAGAAAATAGATCTGCTGAGTCAGCTGTTGCCACCGGCAC
CCCGGCCCGAAGCCTTTGTGGTGGTGCCAGCACCTGGCCCCGAGCCAGATTATCGCCTGCCCGACATTCGAGTGGTGCAA
CTGTGGCTGGAAACCCTCAAAAACGCCGACAGGGACAACTTTGGACGATGCGTGCGACACGCCAAAGCGCACCGCTCGTC
CTATCACCTGACAGCTTACGAACCGTACCTCATCGCCCTCACCGAGCAGTACAACGCGGCCCCGTGCGTGACGGAGAAAG

CCTCGTACCTGCACGGATGCCTGTTCATGTCGTTCCCCGTGGTTTACAACAACACCCAGGGAGGCGTGTACAAACACGAC
TGGAGTAACATTGTGGTCCCGCGGACCATGGTCCGCGAACTGTTCTTCCTACTGTGTTCCACCGGCGAAAGTCCGGCCAT
CTTTCAGCCGATCATCACCAAAGGAGGGCTGTGTTCCTCCATGGTGGTGTACCCCGACGAAACGGCGAAAAACTTTACCA
ATCCCCAGCACATCGGGTTCGTGTACAGCCAACTGGCCATGATACCCTTTGTGCCGCACGCTTTTCCGTCGTACGGTGTG
CCCTTTCTGACCTTGCCCACCCACGTGAGCTACTGTTGGATCGAAGAAATCCCGTTTGGCCGGGCCACCCTGACCCGTCA
GGATGAACACTTGATCATCCACGCGGAACAGCTTACGTGGACAAGCAAAAGGATCACAAACCACGGCCACAAAAAAATGG
CGCGCTACACCGCGCAGTTCAAGGGCGCCCTGCCGGATGACTTCTTAAATGGAACTCAAGTCTGGAAGTCGTCCAAAAAC
GTGCAGTATGAGTTTATGGGAATTCTGTTTACAGTGAACATCGACACCCTGTGTCTGGACCTGCATCACAAGCAATGCAT
CGGCTCCATCTCGACCTCCCACTGCCACCGTGTGGCGGATAACAAACTCACAGCCAGAAACATGCCACGATCATTTGCCT
TTTATTTGTTGCTGGAAAATCAGCAGCAGTGTAACTTGCAATTCAACCGTAACCCCTACTTGTTTTTCAGCGGGGACGGC
CTCAACAGCCCGCTCCTGAACGAACCCAACCCATTCTACTTGACTGTCCACGCTCCTTACGACATTAGTTTCAGTCAACA
CCCCCGACAGACGATTGAGGTTGATATTCGCTACGTGCAAAACGGCAGCCGTTGCTTTCTGGTCGCCAACGTGCCGCACG
AAGACTCCTTTTATACTGGAATGTGCTTGTGGCGGTCCGAGGCGCTGAAGATCACCCTGTGGTCCAGACTGCACACCACC
ATCATCCCCCAGGGCACCCCCATCGCCGCGTTGTATCAGATCAACGAAACCGACGGCAATCTGTACGCGTACAACCATAA
CACGGTCTTCCGGCGAGTCCATTACGCCGATGCAACAACGTTCTACCTCAGCGATTTGAAACTCCCCACGAACAATTTCC
TAACCTCCCCCTGACGGACAGCACTATCATTCTTCATTCTCGCGCGCATCGCCCTTAATTTTGTCTATGATTGCACTGAT
AACGCTTTTAGCACTACCGGCGTCTGAATCTTTTGACGGCGATGGGGTTGGTAGCGGTTTTTTATTACTGTCAGTGGGTT
TCAATGGACTAGTAGCGGGTGAGACTAACATTTTAGCTTGCGATTTTTCCGCAGCTGGTATTAACACCTGCGTGGACGAT
ACAGTTTGTGGTTTTGCGACCTTTTCCTTGATGCTCGTGGCCGGCAATTGAATGGGTGCCGAAGTGGGCGTCTTCGCAAC
ACTATCCTTAGGAATTTTCGTGGATGCGGGCGGTGCCGTGACAGTAGGTTTTGCGGACGATGTCGCAATAGTTGTAGCCG
GCACTTGAAGAGTGGCAGGTTTGAGCACCACGGGTTTAGTGGCTATGGGAGGTTTCACGGCCGCAGGCTTCACCACGGGC
TTTACCACAGGCTTCAGGATGTGTCTCGTATCTATACTGGCCAATGGATCGACAGAAGGGAACGGCCCTAAACTGGGTCT
AGGTTTTGGCTGCGGTGTAATCTTCTGGGGTGCATCCACTGCTAATTTACGTTTAGCTTTGGATGGAGGTGCTTCTGGCA
CAGTAACAACACTAGCTACGGACTCGGCATCGGAGCCGTCATCAGAGAACAAGTCGGCTACTTCGCTTTTCACGTCGTCG
TCATCATCACTAGCTTCTGTTTCACTTGGGTCACGCCTTTTGGCAGGCGAGAAATTGCTATTGGGATCCGTAAAATCAAC
ATCACTCAAATCGGATGCAGCGCTCTTACCAGTTTCGAAATCACTCTCCGGAGCAATGACAGtGGACAATAGTAGGTGTC
AAATCAAATTGGGAGTCAGCCACTTTATTATTCCAATCGATGAGTTGCGATGTATCAACCGTGACAGGCTTAAACAAAGG
TACTGTAGGATCCAGAGGACGTATGGGTTGTATAAATGGATTGTGCTCAGTAGACGACGTTAACTTTGGAGGCTCAAACT
TGTATGCAGGTGTTTGTTCTGTAGAAGAATCCAGTTTCAAACCCTCCAATCCACTACCAATATCCACTAAACTATCCACC
TTCACATTGTCTGATTCGGCATTTTTCCAATCTTTCTCTTCATCGTCAGGCCAGAAAGTTACACGCTTACTAGTACTGGG
TGTATGGCTAGTTACAGGTTTGCTCCAGGGTATGGTGGTGGATACAGGCTTCTCCCcAGGGTGCGGTGGTAGTACcTTGC
GGGTTTCTcCCCAGGGTGCGGCGCTAGTAGGTTTTTCCCAAAAGTCAGCCTTGGTATCGGTAGGCGACAACATGCTAAAC
GCCATCTGCCTATCAGCtCTTTGTAGGTGTGGGTTTCTTAATGTCCCATTTCTCATATGTAGGAAAATCATCGTMTCGTC
GAGCACMTTTCAACCACTTTTCCCATTCCACACGCAATTCCTGGTCATCATACTCCATGGCACCTATGGGATCCCCAGAA
GTAGAAGAcCGCGGAAAAATCTCTTCATAGATGTGAAAATCTTCCATGAGCATTTCTAGCAACATGTCTGCGGTCTCTCC
GTTATGAACCCCGTGCTTAATGATATTAGAACAATAATTGGACTCAAAAGACCTGCATTTTTTCAGCTCTTTCATGAGTG
ACTTACATAAATCTTTAATGTTGATAATGCACTTCTCCGTCAAGGAATCCCAACAGATTCTCACCGTGCTGATAACCCTA
CCGAGATACACCAACTTGTTAGCAGCATTTACCAACGAGCTGTTACAGGGCCGCTCAATGTGCAGCTGCGCCGGTCGTAT
TTCAAACATGTTACTACATATATCCCGCAAAGAATCTTCTAACAGCATGACGTCCTTCTCACGTAGTTGGATAAAAGCCA
CCTCACCGACGGTCAGAGGCTGGCACTCTGACAACCTTGCCAAGGCGGTGGAACTCTCCATTGCACTTTGTATTACCTGG
AGATCCGTACGGGCATCATCTCCGAAAGCACTTACTTCTATAGTTCTCAGGTGAGAACGGGCATCCACTGGCTTTCGAAA
CGGATACGCGTGCGCTCGACGTGaCACTTcACTCACTTGTACCTCAAATTCACTCAAAAGGGAGGACAGGTCAGGTGCGT
GATACCGCAACTCACGATAGTAAGCCAACCAAAGGACAAGCTTGTTGAATATGACGCTGCGCCGGTGTAAGGTTTTTTCG
CCACATCGTCGTAAGATTTTTGCGTTGCGTTCTAAATCGACGTTTGGCCTTTTTGACAATGTCTCTAAAAACTTGTGGAT
AGAAGCCACATGCTTTGGCGGCAGCCCAATAAACTTTAAACTCATGGTTCTTCACCACATATACAGGAAATCGCGCAACG
TGACGCTGTTTCAACGTTCTGTTATTCGTCAGAACGGGGCATGGAAACCATCCTTCATAACCGCACGGTAAGTAACACCT
GGGGTAGTTTAATAGCCGAGGTGAGCAACGTGTCACAAACTGATAAGTTTCGTTTTTGCGCAGGTACGAAACGACAGCAT
TTTGCACATCAACAGCACCTGCAATGTAACCGATTCGCTGCACGCCGCCAAGACTGGCGAAGCTCTGATCAACAGCGCGC
TGGTTTTATTCGGTACCCCCATCAACGCCGTCGTGCTCGTCACTCAGCTGCTGGCCAACCGCGTCCATGGATACTCTACC
CCGATAATTTACATGACCAATCTATATTCAGCGAATTTCCTCACCTTAATCGTTCTCCCCTTCATTATTTTGAGCAATCA
GCACCTGCTTCCGGCTAACGCCGTATTTTGCAAAATTATCTCTCTATTGTATTACTCGAGCTGCAGCGTGGGCTTTGCCA
CGGTAGCACTTATCGCGGCGGACAGGTACAGAGTTATCCACCGCAAAACCCAACTACGCCAGTCATACCGCAACACATAC
CTCATTGTACTGCTAACCTGGGGTATTGGTATCATTTGTGCTACTCCAGGTGGCGTGTATACAACTATTGTAGCCCATAC
CGATGGTAACGAGAAAGAACACACTACTTGTATTATGCACTTTGCATACGATGAACTTTACNGTGCTTGCACTCTGGAAA
ATTCTTATTGTCCTGGCCTGGGGCGTGGTGCCAGTCGTCATGATGATTTGGTTCTACGCGTTCTTCTATAGTACCGTTCA
GAGAACGGCTCAGAAGCAGCAACGCACGCTGACATTTGTGAAAGTGCTGCTCCTGTCATTCATTGTGATTCAAACCCCAT
ACGTGTCAATCATGATCTTCAACGCTTATGCAACCATTGGTTGGGGACTCGACTGCTCTGACTTAACTAAACGCCGAGTG
ATCAACACATTCTCTCGCATAGTTCCCAACCTGCACTGCTTCGTCAACCCCATCCTATACGCTCTCCTGGGAAATGACTT
TGCAGGCAAAGTTCGCCAATGTTTCCGGGGAGAGCTCATGAGCCGCCGGAGCTTTCTCAGAGCCCGGCAACAAGCCCAGA
GCTCGGACGAAGTCCAAACGGCGGTGACTCACGCCACCGCTACGGAAGCTGGCAAACCAAAAGTGAGGCGCGGCGTGTCC
TTTGGCGCCAGGCCCTCCCCCGATCTCCTGTCAGCCAAGAAAAATAGCAAGCGTCTGTCCCAGTCTCATCAGAATCTACG
TCTAGCGTAAATTTCCCCGAACTGTGAACATTTGCCGCTAACCACGCGGTGCAGCATCTCCGGTACATAAGTTTACACAC
CCTGGTACTTGTCACTCACTTGATCACTTTACAGCGGACGCGACGGCCAATCGCTTGGGCTAGGGGCGGCGAGACGGCAT
TATAACGAAACGCCGACCGCTGCCACACGCCACTTGAGAACGCCGCCGTTAGTCACTTTTTCTCAACGGTACGAAATTCG
CAACAGCCATGGATATGAACATAATTATTACTACCCGCGACTTCTCTAATGACGAATCCACGGAGGAGAGCGCGGACCTG
CACAACAGCAACGTCGCCAGCAATCTGTCTAAAGCTTACCGAGGTACGATTCGCGCGGAAGGCAAGAAGAAGCTGCTGAT
TCGCAACCTGCCGGCAACCTTCGGGTGCACGCGCCGCAACAGTAATTTATTTATTTTTGTAACGACCGGGACTATCGCA
AGATTCATCAGGGCATCGTACAGCTGAAGCGAACTCGAACGCAAGTGGACCCCTCGGAGATTGTTAACGTGACCAAAAAC
ATAAAGTGTCGATTACAGCCTCACACGCAGGACCCGCCGCTAACCGGCGGACAGATTCAGACCATCATCTCGCACATCTG
CACCCTGTTTAACCAGCTGGTTTTCACGGCGCAGCTCCGCCACTACTGCGAACATCACGAATTAGTGGTCCTGTACACAC

GAGACGAGCTAACCCGACGCTGCGGAGAAAGATCGGCACTGGGAACGAACATCCACCGGCTCATCTCGCTCTTAGATCAC
AACCAGCACCATGCGCTCTGCAGTGTGCTCGTTGGTCTGCTGCACCAGACGCCGCACATGTGGGCCCGCTCGATCCGCCT
CATCGGCCGATTGAGAAACTTCTTGCAGCAGAGGTTCCTCCATGTTCTGGTGGAAAGCGGACTCCAGGTCGATAGTATTT
TTGAGAGCTGCTACCACAGCGAGGCCTACCGCCTGCTGTTTCAGATCGACAAAAGCAACCAGACGGCCGGTTCTCTGCCC
TGTTCGAGCACCGTTTTACCTGTCAGTGAAAACGAAACTGAAGACACACCTGTCCCGCCGTGTATTTAATGAAATAATAA
AATGTTTCGTTACGAATAAGATGGTTTTGGTTTCGTTTTTGAAAGAGCAACCATGAGCGTGCCCCCGCACATCCCGGCCC
TGCCCGTAGACCACAGCCTCAACTTTCGTCCGGACCTGTTTTCGGATAACAACAGACGTATGTTAGTTGACATGCTCACC
CAGAGCTGCAGCGGTTATGTGGGATTATGTAACTACGGACTTCCCATGCCCGTTTATGTGCTTGATACGCTTATCGACTT
TCAGTTAAAAACGCCTTACATAAAAATTAAAAACATTGCCAGAAAAATCATTCATATTTGCATTTGTGCCAATCATCTCA
GTAACAGTCAAGATCTGTTGACTGATTTTCAGACGAATCTAGACGACGTGAACGTTGACGAAGACAGACTGAGACTGCGT
CGAGGATTTCACAGACTCTGCCGCCAACACAACATGATGGACATTTACCGGAATATTACTCAACTGCTCGATGCGAGGGA
CGTGACCTTTGGCACCTATAAACAATGCCTGTCCCAGATACGCGCGTTAATGGAAGCTTGTCATGTGCGTCCCGATCCCA
GCGTGCAGCCTCTATACGAAAGCCTCACCACNGTACAACCTCCTCTACTTGCCTCCACCCTTCACGACCGCCGAAGCCGT
TCAACTGTACAGTAAAAACATTAGCGATCTTACAAAACGCAATAACAGACCGTTTTATCTACTCACTGAAACACACCGAT
CGAAACGACCAGACTACGTCCTTAACGACATACTGTTCATGCTGTCGGCCCTACACTTACAGTTCCGACACAGACTAGAA
ATACAGCTTACACAATCTTGGATCTTAAGTCAATGAATAATCTCTGTAACGACCTGTACTTTGCGTACACACAGGTCCC
GGAATGTCATCAAGAGTTTGTACGTTTGGTAGAAGGTTTGCAGATTTTGCGCACAGTCCAGCACCCGGTTTTCTCTCCCA
CGCTGAACGGGTTGTGCAACTTGATGAGAACGTTTCACAACGCTAATGTCTACATCTGCCCTGGGTATTTACACGCGTCC
ATCTTTCATTTACTACGTAGAGTGTCGCAAGTGGCCGAGGGTCACCCATCAGACTCGGACTCTGAAAACCAAGAAGCCAC
CAGCGATCACGACACTATATTGAGACGCTTTCAAGAAGGCGACCGCATTGATCTGGAAGCCGACCCCACAGCTATCCCTT
CCCAGAACTTTTATTTCTCCAGAAATCCTTTTGGGAATCGAGAGCTTTTCTGTATCCCGGATTACCCTAGTCAGCATCTT
ACAAGAGACCGGTGGATTGGATATCCCCACATGACATATACGCAATTCAACATGGAAAACGGGCAATTTGATACAGAAGT
CCATCATTTACAGAGAGTGTACAGCCTAGTCATGGAAGGCGCTTCCCGGCAAGCCGGCACCACCGCCAGACGATTTGCCC
AAGCCATGGAGCAAGCCACAGCGGGACACGATGAACCCCAGCCTGTAGATCACGGCGACTTGTTTGCCGACGTGGAAGTG
AGAACCTCACAGGCGTCTTCGGACTCTTCCTCAGCGCCGTCTACCTCCTCGACTCATCATTCCTCATCATCACCCCGTCA
CCGCGAAGGACTGGAACAAGCTCCTAGACAGAGGCGGTTTCTGAGCACGCATCACTTCTCGCCCTACTCCGTGGCCCGTC
AACACCACCAACACCGTCATCGTAGGAGGCGCCCCCCTCCACCCCCGCGGGGACCGGCCCATACTTCTCGCCAAGGACCA
GACATTACCTCCGCATCCGCTTCTGCGTCTGCTTCTGAGGATGACCCAAGAGACAACCTAGCTGAAGATTTGCAAAACCG
ACTGTAAGGGGGGAGGGATGGTGACGATTTTGAAATGAGACAATAAAAACTCATTTTTGTACATTCGGGATAAGCGTGTT
TATTGATATAACAAAAGTACCAGATAATAATAAAAAACACTTTATGGCTTCACTCTACACATAAGAAACTCGGGCACGGG
CTTCGCGCCGAATGCCGCGACCCACACGGAATAGGAAAAGACGCCGGCTAGACACACAGCGCCGACAGACACTGTCATCA
AATAAATAGCCATGAACTGTGTCATCTTCGCGATCTGAATCATCATTGGCACCGTTGTAGTGCATCTCGAAGGCTCTCAT
GTTGGCAACGCTACCAAGCAGTTTAGCCTTCACGTCATCCACATCGTACACCGAGGATGCGGATAACAAGCATGATTCAG
CAAGGGCTTCCAGTTGCTGAGGCCACACTCTCTGACTAAAAGGCAGGTTTAGCATCATAATATCCACGTGGTACTGCACA
TCATTCTTGATGGCCTCGTAAATGCGCCTACACGTTTCGTCCGGCGTTTTCTCAGGCGGTTGGGGGTCATTATCTTCCTG
GAACATGTTCATGGGGTCGGCGGGCTCCCCGGTGTGACAGCACCAAGGAACTTTTTCCAATCTCGCAATTCCAGTGCGTT
TGAGATCGTACCGATGATTTCTGTAGCTCTTTCGACCCATCATACAAGTAAACATTTTGATATTCTCGGCCAAGCGGTAG
AGGCCGTTCATGGCTGGATCGAATCCGTAAAGCACGCCAAACTTGTCCAGTAGGACGTAAATTCTCACCAGAAATTCACC
ATTTTCTAGAAGACATCCTAGTGCCCCAAATATTTGATGCATACACACCATGCGAGAAACAACTTTCACCATAATTTCAA
ATGACCTGTGTCCTTCACCCCGCATACAACTAAAGCAAGTATCTTTAAACGGCACACTAGAATTACTACACCACAGAACA
AAATGAGTATCTATGTTGTCACCCGTGGGATCATGAATTTCGTAGACATTGCTATTTTCCGTCACAATTTCATGAAGCTG
ATTGTCATCCCAATTAAACGTGATTAGCAGCCGCTTCATGATATCATCATATTCAATCAATGAACTCTGATAAGCCTCTT
CTGCCTGAAGGTATTCGCAGCGACGACATCCCAGCTGCATAAAGGTTTTCATGTCCTCGGCCAGGCAATAGACACATTTT
TCCACATCTTCATAACAGTAAATGCCACCGTTGCGACCAATCATTAAAATAATTTGATACGGGGCCAGTTTGGGAAGTTC
CAGCACTCCAAAAGGTTCCATGGTTTCTCCACAGCACAGATACTTTCCTATGGAATGAAAACAAAGTGCAAAAATTTCAC
AATCACCGTTAGCACACACACCCTCCACTTCCCTCGGGCATTCCTCACTTACCATCCAAACACACGTCGTAATACTTGGA
ACTGAAATGGCCCATAAGATTACGAGGACGAAGCTGCAGACGAGCATTTTCGGGCCACGCCACCTGCAACCGGGTTCCAC
ACTCCCGATTCAAGAATTCCACGAGTTTGTCGGGATTGCCGGCGTGCACCGCGATGCAGCCATAAGACATCAGAGTATCG
TGTATATCTTCCATGTTGATTAGCACCAGACCGTTGCAGCGAATATGGCAAGTTGAAGTAGCGCTTCAGACAGAAGCAGT
GAAGTGTGACCGAGCTACGCCGCACGCTATTTATATGCGACATGGAGGCTTAGTCCTGCAAATCAGCTGACTCATTAATG
TAGAAAAGATCACGCTTCGCGTCCATACGGTTATATTGAGATTTAACGCACATTTCGTGAAGATCGTTCTGAATGTGACG
GCGGCTTCGCCCCACGATTCCGCAGAAGAGACTGAGCAGCACAACAGCGCCAAGGACCACGGAAGCAACGCTGAACACAC
TCCAGGCACATAAAACAAGTCGAGGAAACGCGTTGACACCTGTCAACCTTTGAGTGATTTCAGTCCATGTGGGCCATTCC
GGCAGACTTGCGTCGGTAACATTTTCTGCATAGTCATTGGTTACATTCACCAGCGATGTTTGGAATTTTTTCAAGCGGTT
AAACAACCGCATGCACACATCACTTAGAAAAAAGCGATCCGTGCCCGCATGATTGTCCAGCCAAATAGACACACTGGTAG
CGAACGGTTTGGAAACGTTCTTGTTGTGGCGCCACCGACGTTTCTCATTTCGTCGCTGGCGACTCGTCACATTCAACATA
TGTCGCCCCCAAGACGTGCTATGTAACACAATGTTTCCCTGAGTCACGTTTTCATTGGAAATATTCAAACCACATGTCAA
ACTGCCCGTAAAGGGCATAAATGTTGATGCATTACCCAGCAAATAAAAACGCTGCATAATTCTGCCAGCAAAATAATACT
CCAGGCCGTTTCGCATGATTGTGTGCTGCCAATCACGACCATGCTTGAGGAGTGTCACATTGAGTTTCAAGTTTGTTATA
GCCGTGAAACAAGGCCCTTGATAGATGAGTGTCTGGTTGTATGTACAGTTCAGCCGACACGTTATCGTGCCGTTCGCTTC
GCCGCTGTATTGACAGTTGAGAGACCAGCCATGAACCACCAGGCAGAACACCACGGATCCCAGGCAAATCTAAACAACAG
ATGGAATAGGACACATTGAACATTCCACCCACCCGGGAAAACACCTCAACCCTCCCAGTTGTTCCATTACCTGCCAAACA
CGATGGTAGGTGCTCATCACCCCACAAGATGAAAGACACGACTGCAGGTTTACACACAAAACATGTTTATTACCAAAGCA
ACATCGTTATAAACAGTGAGGGCGTGTCTATGTACATGTAACAGTGCGGTTTACAAACACATGCAGTGCAATTTACAAGG
CGGCTACGCCTTCCATCGTTAAATCGATAATCTCTTCGTCGCTACTATCGGAATTGGCTTCAGCCCTTTTCTTAGGGTTG
GAAGTAGGGAGCGAACACTGCCGAGTCAGTGCTTTCTTTCCTGGCGCGGTGCCAGGTTCTGGCGGATCCTGTTTAGGGGT
CGGAGGAGATATCACGTTGAGCATCAGGCTGTCCCGATCTACAAACGGCACCGGTATCTTTCCTTCTCGGACACGTCTCT
CGGCATCCAAGAGGGCGCGCACATAACCGCAGAGGAATCTGGCAAAAGTCCGAGGACTGTTGTGGTTTTTCAGACRACAA
GCTTCCAAGACGTCCCGCTGATGGTACACCTGGTGATAAATGACAATCAGGTCGTCTATGGTCTTGGCGTACTGCCGAAC
CCGGTGCCTGATYTCCAAAGGAAAAATCCCGTTCCAATACTCAAAAACAACGGGCGCAGTCCAAAAGTCCGACATGGTGG

GTGCCATATAAAACAGGATGCCGGCATCATAAGCGAATATATCGTAATTGTCACTCATAAACAGCACCACCTGTCTGGAA
GGCCTAGGGGCAGCCTCCTCCCAAGCTTCAATCACCCCAAACATAGTCAGTCCGTGCTGCAGAGGCGGACAGTTACGCTC
CAGATGAGCGATTTTGGACACGGTGATGCACAGAAACGGATAGGTACAATCGCAGACACAGACGGTCACACCAAACTCAT
GTCGAATGCGTGTCTCGCGCGCCATCAGCTTCTTCACTTGAGCCCTCAGCGCCGTAGGACCCAAGATAGAAGCCTGAACG
AGATATCCTATGTTCAGCTGTGCCTGCTGCCTCTCCCGCTCCTGCTCATATTCTAAACATGCAGCCATGACTCCGATAAG
CGTCTCTTACCGAGCGGGGCTGTCTGCAGCTAATAAGTAGTCAGATACCGACGTCGCTGGAGCCGGGCCAAACTTGACTT
CAGAGCATCGTACGCCTCGCGAACCAGGCCCACAACTCCTTCAGCAGGAGACAGAGTCGAAACAGGAGCAGCAACGTCCT
CCTCTTCTTCCAATTCGCCCCGTTTGCGCTTTACCTCTCTCACATATTCTCTGAACCATCGAGGCAGAGGATCTTCGTGT
CTCTTGCGCTCGATCCACTGGTAACTAAAGtAGCCAAAGGCCACCAGACCTATGGCGCCAGCAGTATTCACGTACTGGAT
GGGAATGGCAGGCATGGTGAAAGCTGCGTGCTGTGTTGAAGGTTGTTAATCTTATATACACACAGGTTTCGCGTCCCGGA
CATGACGGGCGTTAATTTCACCAGCGTCACACCTGCCGCGTCAAATTGCACAAGGTGAGTAACGCGTACACACTGACGCG
TGTCAATCATGAGCCTAACCGCAATCACCCCGATCCCCTACTAACTTCCCCTTTCCCACCCGCTACCGTTGTGTATGTTT
CCCGTGAGAAATAAAAACGTATCACAACATATAATCATGTCAACGTTTTTTATTGGGGGCAGTATCATACAAGGGTGTTT
TAGGCAAAGCATTATCCACATACGCCATCTGGGGGTAAAAGGCAGCAAAAACACTGAGAGCCAATGTACACAAAATATAC
AGTCCCAAAGTTTCCAAGCCGAGAGACAGCGTGGGGTACAGTCCGAGGCTGTGCCGAATCCAATCGAAGAGCTGTGAGCC
TTTGCTTTCAGCACAGGAAATCATGTCAATCATCGTCCAATTTTGTTGAAGTCCATGAACAGTCAGGTTTGTGTAGCGCG
TAATGTTGGATAAGACTRAAGAGTTCGAACATCCCTTCGCACCCGCAGTGGCATTCAACAAGATGTTTAGCGTAGCTCCC
AACGCGCAGTAGCGTCCCTCGACAGTTACATTAGTCAAGGTCGTCCTGGTTCTGGCCATTCAACACGTACAGCAAACCAC
GATACACCWCCGAACAATTCACAAAAGCGTGTGTCTTACTCGCCGATGCGACAGTGACGGTCACCAGCAGAGTGAGGATG
CCAACTACCAAGAGCAATGTCCTGGGAGCCATCACCGCTACAACGAACAGCGCCATGGTGAGTTGTTTCTGGCACCGCGT
CCCGCGTCGTCTGAAGTTTTTGAGCCAAGTCTGCCCTTCATGCAAATTGTTTTCCATTTTTAGAAGTCAATGTCTTTCCC
TGCTGGCTTTTTCTCCCACAACACGTCTTCGTCGTCCTCGGAACTCAGTAAGGCAGCATACTTTCTCCGATCCCACACCG
CTCGCAGCAGACAGGAAAAGACTATAAGGAAAAACCCAAATACAAAACAGGCAACCCCTACCAGCCAAGGACACGGTAAA
AAGTACTGGTTCTCGGGGTTCGCAAAAGTCCCATTACACCACAGCATGAGCATGTTGAAAAGAGAGCAGTGCGTGCCAAC
CGTCACATCACCCCGCAAAAAAAGACACAGACACAGGTTGCAGTGGTAGTTATAAGTGTTATTGACGCAAATGCAGGCAC
AGTTCGGTTTGATTTCAGTTTCACTAACTTCCAGATTTTTTCATTTGCATAAGCACGGAGACCAGGACAGCAAGTAAGAT
AATCCCTAATACAGCAATAATAAAGACCCCTACAGTCCATTTCCATCCATCTTCGTAATACGCTTCAATCGGAGGATGGG
AGGTAGCTGGAAACTCTGACATGTCCATTTGAAAGTTATCAGGGCGCGGGGGGACTTCCCCGACCACCAGGTAATAAGGA
GGAGGTGCAGAATCATCTGAGGTGGGTGGAACATGTCTCGTGGCAGTGACAGACACGGTTGCCAGTCCCAGAGCTTCTTC
ATAACTGGGGGGCGCATCATCATCGCTGGGCTGCCGATGATGCTGTGGTGACTGCTGCTGCTGCGGTGGCTGCTGCTGTT
GTGGTCGTCGCGGCTGCTGTTGCTGCCGCTGCTGCTCATTCGTCGCTGACATGCTTTGAGCCGAGACAGAGAAAGCG
GTCTTTAAAGCTTCCCCTTTTTTGCTCGCAGCGATGTCCCTTCTTATGCAGCTTGCGACAGAAGAGTTCTCTGGGACAGC
CGGTGGGCCGGTAGGCAGGTGCGGCGTCGCGGTCTTTCTTGGGATTGAAGCAGTCACGCCAGTACTGGTGTTGATTTCCC
AAGCGAAGCAGTCCCACTCGGAAAAACGCTGGCAGGTTGAGGGCTAGGTATTCAGGGCGCCCCGTCTCCGGATTGACTCC
GAAGACTCTCAGCTGCTGGTTCACGAAGATGGTAATTTCGCGCCACTTCTGCTTCCACTGCGCTAGCAGCAGAAACACCT
GCAGCACGTTGATAGAGCCAGTACTTTTTATCCAGTCCTGATGAGCCGGATTTTCGTAGGCAACTACCGATACGACCAAG
GCGCTCATGAGATCCTCGTTGCACGCCCAGGCAAACTCTCTTCCATCGTATTTCTGGCAGAGACGAAACAGTCCCATAAA
ACTATGGTCCTTCTGGAAATCATAAGCCAGAGATTCCGCAATGCTGTAGCGAACAGGTCCACACTTTTCTCGAAGAGGGC
AGTAGTTCATAAGACCCTCGGTGCGCAGAGCATGCAAACCGGTGCGCGAGATCAGGTAAACGCTGTCATGGCAGGGGTCG
TGGATGAAAACGCTCCCACCTTCGCCAACGAGAACGACGAGGACGCGGCGGACAATGGCCAGCTTGCCACTGTCGCTCAC
TGTACCCACAGGCACCAGATAATTGTCACAGCACAGGTATTTCTTCTCCAAATCTTTAATATTGAGATGGCCGAAAAAGG
TCTGTCGTAAAAAACTACAAAAGGTAAAACACCATCCCTCTGGCCAAGCGATAGGAAAGGTCTCTCCATTCACTTTCCGA
ACAAACTTCATCAAGCGCCCTGACGGGTCTTTGCTACTCACGGAGCAGTCAAAGGCTTGACGGGCAAGCATTTCATAGCG
ATCCATGTCAGAGACAATCATGATGATGATTGACAGCACACAGCGTTGGACGGAGTCCACGCGCCCGTTCCCGGGGGAAC
CAGATATAGTTCAGTGCCAACTGACGTACCGACCGCAAAATCTGCATCTGCGTCATGTTTTTAGGTGCACTTTTGCTTTT
TGTTGCTGGGTACAAACTCGAAATTGACAGATTCGTCGCTGTCACTCTCTTCCTTGGCGTCGTTGAAGTACCCCGAGTTT
CTCTCTCCGGAGTTGGGAGCCCCCTTCGACGTCAGGTAGCTAGTAATCTTGTGCTGCTCGTATTTCTCCTTGGCGTTGTG
CTCGTGCTCATTAATTTTTCGGCTGGACAGCCGTTCGTGTTTGCGACCCACGGGTTCCGAAAGCATCTCGCTATCCATTC
CCGAGCCGTCATCATGCATAGGACCTGAATTGGTAACCGCGACCGCCCTCTGCTGTTTGAAACTCTTATCGAAGCTCAGG
TCACCGCGCTGGAAAGGCTCCTCCGTGAGGAAGTTGTCCACGGCGAAGAGACCGTTGTGACTACTGACGTGCAGCACCGT
GTCGTGCTCCGTCACGATGCGAACCGTGCAAGGCAGCTTGGTGACCGCGCAATTCATCAGCGCCTGATAGAGATTTTTCA
GTTGCACGTGCACGCGCATATTTTTCACCCCGTGAAAGCTGACTTTGGTGTTAGCCGTAAACTCCAGTTCGCTGCCGCTG
CTGAGGAGGAACTTAATGGCCGGGGGGGCGGAATGTACCAGAACCTGCACCGTACCGGTGGGACAAGGGGCCTTTTTGAC
GCTGCGTTTCACGCGCGTGTGAGGCGAGATCCACTTGATGAGCTCGCTCACGACGGCGTGATCGAAATCCACGTGCACGG
YCGCGCTGTCGCTCTCGCGCACCAGGTCTTGGCCGTGCACGCAGGGAGAGCTGAACTCCATGTTGAAATCGGGAGCGCAC
ATGTGAATTTTGGCCGACAGGTCAGAGATGTCTTGCACGTAAAATTTGGTTAAGTCCTTGCTAGACGTCAGATACATGAA
ATTCCCCAGCATAGGCGTGCAGTTATTGATGGTTTTTGGTTGAAACGACTTGTCGGTAATGTACAGGCAAGCGCTGTTGA
ACGTAATTTTCGACACGTACTGAGTGCGTACCGTTTGCAAAACCAGCGCGGGAGTAGGCAAAAAAGTGACAGTGGTGTTC
TCCTTGAGAGTGCGAATGACCGACCGGAGCTGCTGGATAGCCGACTTGTAAGGCTTAAGCCTCAGAGCCAGCGTCGGCGG
CTCGGGTAAACGGGCTTTTCTATCCATACCTGCCGGTCGTTAGGTGAGCCTGGGACACTATGGATACCCACCGCCAGCTC
CCAGATGTGTCCTCGTCCGCCACCAGGAGTCAAATAAGCTGAAACGCGTTAAATAATTGACTTTTATAGACGACAGACGG
CGCCGCAAAATAAGCACGCTGAAAGTGGGGCAATTGCAGCACACAGCAGGGCTCGATAGCGTTTAATACAGCTCTGCGAA
TCTCTACCCTAGCATACAAGTCCCATCTAGGAAGCACAGTACTTATATACTCCAACCTTACACCGCAAGTGGTACCATTT
CAACATCCTCTTCATCAGATCATCGGGGGTCATGGAGGACTCGCAATAGAAAGGCAACCCTAGACACTGGTCCACGTACA
TAGCCATATTGACACTGGCCATTTCGACGTCGGAATGGAGTTCCGAGGCCCGCTGAAAGACACAATCGGAAACTTGGGAC
ACCCGGGGGTTCTTAGACTTGAGTTCGTCCTGGTGTTCGATCCAGGCGTAGTTGACGACGGGGACTCTGAGGCGTTTCTG
CATGACCGTCGGAAAATAACAGTCATCGTTAGGCGGACTGACGGGCAGCACGGTGGGCTGTTCCTCCGACACCCTGTCCC
GGGTGGCCCAGACCGATGGGGTGACGCTCCATAGATAGGCCACCTCCTCATCGGGACCGAGAGCTAGAAACTGCGCGTTG
CGTGTCCCGTACTGTTGCATGTCAGCCCGCAGGGTTTCCCACTGTCCCACTGAAAGCGACAGTCGAGGCTTCCGAAACAG
TTCGAAAATGAACCTGCCATCCTTGTAGATGGTCCGTTCGAACCAGTCACAGGGAGGCAGGCCAGATTTACACAGATCCA

CGCTGGTGCGCAACGCGGCAAAGTACACATTTTCAAAGATTCGTTCGATCAGGTCCCAGGACGCGAAATAGGTGAACCCC
AAACGCATTAACGTAGTGTGCAAACCCACCACCCCGATACTCAAACTCCGCAATTTACACACGGCCGACTCTAGTGACCA
CTCGCCGGTCGCTATCAGGGCGTCCAGACGCGCGTTAGCCCAGATGACGGACTCGGTCACCAGGTTCCTAAGAATCTGCA
GGTCGAAATAACGCCTGGTGTTGCCCAAAACCACATCTTTTTTGCTCAGCGCGTCTTTGTCGCTGCGCACGAACAAACAA
CGTGTCAGGTTGATTGACAGGCGCTGAACCGGCAGATCGCAGTCCGCGAAATGGCACGATGCCAAGTCCGGGCTCAGACA
GGGCGGAAGGGGAATGTTTCCCACTAGCGGCGAATATTTCACCACGTTATGCACAAACACGATGCCACACTTCCCCTTGA
GAGAACAAAACAGTATGTTACGCAAAAAGTCACTGGCTTTCACCGTGACTTTGGCAGTTGACTGCTCATAACGCACATAC
TCTTTCTCAAAAGTCATCTCGTTGCCGTGCTCGGACAGTTTCCGAGAGATGGCACGCCCAAACAGCGAGTATTGAGAGTC
GCCGGTGCGATAGCGCTTCATGAGCACTCCAGGCACGTTAAACGCGTAATACACCCCACAGATGTCATTCTCTTTCAAGA
TAAAGTTCAAAACCCGAACCGCCGCTACGTCCCACATGTCCACGAAAATGCGCGCCATGGGCTTGTGCACCTCTCTCTCG
CGAACCATGTCGCAGAAACCACCGATGCAGCGGATGACGCTAACGGCTTCAATGTTCAGTCGAGTGACATTGATGGACAC
GGACACGCCGCGAGCCAGATTCTCGTACAGCCACTTGTGCAGAGTGATCCAGTCCACATCACGAGAAAACGTGTCCGAGG
CAATGAAATAATCATATTTGGGAAGCTCGACGAAGCCCATGGACGACATGCAGAAAGGCGACATGGAGATATGTCCGTCG
CGAACGCGCCTGACGATACCCTCCAAGACTTCTCCCTCGTGTCTACAGTCCAGGTGGAAGCGGTACACCCTGGCACAGGG
ACTGTTGACTATGGCCAACGTGAGCGCTATGGCAGACCTGACAAAAGCGCCCCCTATCGATTCATCTTGACCACAGATCT
GACGACAAAWTTTATCGGACAGCACGATGGCGGAGTAAAACCCCCAGCAGCCTGCCAAGTCCAGATARCGATTGGTGTCC
TCCAGGAGCTTGGGCAATTCGTCCGAATGTCGCAAGACAAAGTCGCGCACGGCCTCCTCGCAGAGCGAGGAAGAGTACAT
CATCATGTTTCGACTCTCCCGCACCCAGTTGATTTTGAGGTAGAAGCGTCCGATCAAGAGATCCGAACGGGGCACTTGCT
TAATCAGGGCCGGTCCAAACTTCATGAGCATCTGAAAGAGGCGATACGCGCCCCTGATGCCGGGCTGGCGCGAGATGACA
TCGATCTCGCGCCACGTGCTAGCGCGATCGTCCACTTGCTTGACGGGGTCAGGCACCCCGCAGAAGCGAAACTTGCCCCA
ACGACCGTGACAGTGGCATTTTTTGCCCATCAGGAGCTGCACGGAATTCATCCCTGGCATGTGAACCGCCAACGAACACG
GACCTCCGGGCTTCAGCGTCCCCTCATCGCCTGTGGACCAGCGGGCGGTCGCGTTACTCGCCCTGGACGGCGGACGGTTT
TTATCCTCTTCTGTGGTTGAAGAAACTGTCGTATCCATAGCCTCCCCGTTGTCGTCCGAGCCCTCAGATTCCTCCTCCAT
AATCGCTTCGAGCCCTCCGATACCGTTTTTATGAGATGCGCGAGCCATGAGAACAGCACCACCCAACGACGATGGCCAGG
TTACACAAACTCCACAAACTCGCCCAATTCCACGTTGAAGTTGTCCAGACGGAACGTTGATAGTTTCACCTCTTGGTGGG
TCACCACCGAGGCGTGAATATCGTCATTGAGACACCAGCCAAGTTTCTGAATCACCCCTGTACATAAAGGAATATGTCGC
GGCCTAGTGATTCGCCTGCTGACAAAGCTACAACGTTCTCGGGTGAATTTCTGCCGCACCACGTCCGTTAAGGTATGGCG
ATTCGAGATGCTGGACGTAACTACATAAATCGCAGCCTTGGTTTCCAAGGGCTCGTGGTCGTACACAATAATTAAATACA
CATTTTGCAGTCTCTCATGGGTTTCTTCCAAAGTCATGTACATGTCCTTCGGGATGGACAACGAAAAAATCTGTGTTTTT
AACGTATTGTAAACATAGCCTTGCATAACGCAGGTCAATAACGACGTAATGCCCAGCGACACCGATTTCACACAACCTAT
AGTGTGCAGACGGTGATGCAACAGTTTGTTTCCCATGTCCAGCCAGGTGAGGGCGGCCTTAGCCGGCCCCGAAGTGTAAA
TGCCTTCCAACTGACACCGAGCAGCCACGGACACGTTGGAGGCTTTGATGGCCAGCACATACAGACTGACGTACAGGTAC
CGCGGTGACTCCGCGCCGGTCACCATGCGAAACATGAGCGAACCCGGGTGTACTTCGAACTCCACCAGGCCCTCCGGCAC
GAAGAAACGCGATGTGATGGCCTGATGATCGGCCTCGTACAGGTAATGCACCGCTAGTGTATTTACTTCTCTCTTGAGCT
TAGTAGCGAGCACGAGTTCGTTATCTTCATCGTTCGGGTCCCTCGGTCGCTTGGCGACCTCGCGAACATCCATGATGAGG
AAACGCACAATACACTTTCAGACGTTCCTAAGCGAGCTGAAACAGAATTCAACGGATCATCAGGCCGTATTAAATATCTT
AAGCCACATAGAGATCGGCGCGCTCGATGTTGAGACCATCAACGCCACTATAATACAACAGTTCCTGCAACTTTTACCGA
AGGACGGCTACCACTTTTCCTTTATTCGCCAAAATGCCGTCTACTACTTGCTGAGTCACACCACGTTACAAACGGCCCGT
GACCCCCTCGCCGCGGCTGAAGACCTTTACAGAGAATTGCACAAACACACCCAAAGAAAGAACGTCATCAGTCCATGCGA
GATCCTGAACAACCAGGCCACGCTTTCGAGCTTCAAAAAATTCCTCGAAAAGGTGCAAGAACTGCAGACGTCGTTTAACA
CCAATCAGTTTTTAGCCGCCACTCTACCGGACCGTAAAAACAAAAAGTCACAGCGAGACCTCAGCGACAATGGCACGCAA
TACCTGGTGCGGAACGTGTTTACACAGATCGAAACCATCACAGAAGAAACACGACCCTACAGCAACTGTCGGCTGATTGC
GGAACTGTTGGATATTTTGTACCACAGCACACTGCAATGGGCTAGCGGGTTGATGCAATACCATGCATTCGCGAATCCGA
ATGACAGTGAACTTGACATCATGCTCAAAATGAATTACTTTTACAAATACTACAAATCTGTCAACAACGACCTCAACAAC
GACTTTCAAACGTTTGTGACTAGCGAAAAGCGAAATCTTGCCGCCTTTATGGTCTCTGAAGCGGACGGAGACCAGCAGAT
AGGAGCGGAGCACATCAAAGATTTTAGTTTCAAAATTTTTAATGCCAACCTCAACGACCGCGACACAACAGGTCTAGCGT
TTCCCATATTGTTAACCTCGCTCTCCGTCCTAAACCACATGGGAGTAGAGAACATTTTTTTTCACCAAGGTCTCATTTTT
AGACTCCTGAGTGACGAATCTCAACTGAGTACCAGCGAAGTCACTCGTTTAGACCTCGTATCGCAGATGTGCACTCATAT
GACCAACGACTTTTTCCGCACCACAACAAAGAAACCCATGAAACTACGGGACCTCATGACACGCATTCGTACGTTCACGG
CTCTAGGCCTAAACCGAGACACTTCCAAAATATACGCCCAGATGGCACTTTTCCAACCTAGCCGTATCGCACCCAAAAAA
GGAGGAGAGATGTGGTTTGAGATTCGACAACAACTACTGATTATGATTTACAACACTTATGTGTTTTTTATGTGTCTCTG
GGTGTACAGTCCTACATTTTTATTTTTACATCGACGGAAAATAATCTTAGAGCAGCAAAAATCTACACTGATCGGATCCC
GTCACGAACTCCAATTTATCTGGGACAATGTTACCCGTAATATCTACCAAGATTTCAACGTTCTATTCACCGAAGAGGAG
TTTAACGTGTGTACACAAGGAACCACAGAAACGGAGAAAGATTACCTGTACCGAGACTTAATTAACAAATGGGGTGACAT
TTTGTTTGTGTTAAAACCGCAACTCCTCATCCCCGATGCACAATCACCGCCCCTAGAGGATGTTACTACCGCGGATGTTA
TCAAACGTTGTGCCATGATTAACCTGTCCGAGACAAGCGTACCCTACGAATCACTGCTGCCCCTTATCCATCATTCCAGT
TTTCTGGACACGTTCATGAACCTGGTTATCGTGCCAGAGTTTTTACAAATTCTCACTATCCCGTATAATCAATTTGTGGC
CATCGGGTCCCCTCGTCTGCTGAGGCTCATTCACGCGTGCCGTCTTTTGGCACCGGATCAGATGACGTTATACAACAAAC
TAGTGTCTCTATACAATCTAATCACCTTTGTATCAGAGATTGACGCGGGCATGTTCCGCACCGTGTACAGTCTAACCCTG
GAAATATCCGCCATACTGGAAAACCTCTGCGGTGAACCCGTGTCCCCAGAACTAGACCTGTTAGTAGAACTCATGGCCCA
ATCACTGTCACACAACTTGGAAACCACAGTAAACCCCCTACTCGAAGAAGTGATGCAAAACAATAGTGCGAGTATCGCCA
ACTATCTGGAGCACACGCATCTGTGTTTTTGTTTTGCACACACCACGGGACAACTGTCCCCAGACATGGCTACCGTAGAA
CTGAGTTTGCAATCCAAACTCATCGCAAAAATACCCATGACCCAGTTTCTGGAACTCACCAAGACTCTGATCGCAAAGGA
TCAGAGGTTAACTGAGAATTTACAGAGAATTGAAGAACGGTTACAGCACATTAATCAGCGAATGAAACAGATTAAGGAAG
ATTTCACCCAACTGAATGAGTACTCCAAAACTCACCCCCTGACCATGGAAATCGTTAAGTCCATCAAAAAAACCACCCGA
ACTCTGAAAACCCTAGAACAACGTCTGCTGGCGATTTTGGAGCAAGCCAAAAAGTCCAACAAACTCATCGTGAACTCGCT
GAAACGTGTGGAAAAGACCTGTACAGCCCTCAGTAACAGCAATATCCAGGACCACGGCATACGCTACTGCATTGCCGAAG
CTCGAGGTCTACTCCGCACACATCAAGGAATGGCAGCCATCCCTCCGTCGGCTCTGGAACACATACCGAGCAACGCGGAA
TCCAAGCTGCGCGCCTTCCTTAGAGCGTTTCGCGAAGCAAACGTGGAACACAGCCCGTTGCAGATGAGCGCTGTCACGGA
TAAGCAACAACAGATTCCCATACCCGATGTGTTAACAAACCGCTACCAACCGTCGAACCAGACAAACGACCTATTAAACT

GGTATATTACGTCAAAAGACCAGGCGCAGACGGACATCTTAACTTCGATCTACCCCGAAGGCCGGCATGCGTGATCGCCA
CGCGAGCCTGAAACATGAAAATCTATCAAGCACGATCCGACCAAACCCACAGCCGATTCGGTCCGCGAGCGGGAAGTCAA
TGCGTCTCCAACTGTTTTCTGTACCTGCACGCCCTTCACCTGAAGGGCGCGCACACCACGCTCAACACAGACACCCTGGA
TATGATTTTGACCGAAGGCGCGCGCCTGGACAGCGAGGTGGAAGCGCTTCTCAAGCAAAAACGGCCAGGGGAACGGCTGC
CCGTGTTTCGATTGGGCGAAGAGATTCCTAACATCATCACCAGCAGCTTCGGCAAAACGGCCCACGCCCTCTCGAGGCCG
TTCAACGGTACCGCGGAGACTCGGGACCTAGACGGCTACACGTGCTATGGCATCTTCGACTTTCTCATGCACGCCTACCA
AAAACCCAGACCCACGTACATCCTGGTGACCGTGAACGCCCTAGCACGCGCGGTGATCCTCCTGGAACATGACATCTTTG
TCTTCGACCCGCACGCTACCGAGCGTACTTCCAACGCCGCGGTGTATCAGTGTGATTCGTTGTACGATGTGGTGATGCTG
CTCACCTACTTCGGAACGCGTCTGGCCGACTTTTACTACGACGCCCTGTTCGTGTACATGATAGATTTGTCCATTCAACC
GGTGCCGGAATCAGAGATTAACGGCATCATCATCTCGCTGTTCCGCGACCCCGACATCGCCCTACCCGCGTCCGTTCTGC
CACCACTACCTGGCTCACCCGTGACCATCCTCCCGAGCCTCCCGCCACCGGATACCGCTGCTACACCCAAGAAAACCCCA
GAGAAGCGAAAAGCCACACCCAACGTGTCCCACGGTGGGAAAAAGAAAGCCACAGCAGCCAAAGAGAAGGGTTCGTCACG
CGTCACTACAACCGCTCCTTACGACTGTAGCGACGCTCTGACAGCCCTCAGTAGGTACGATCGATTAATTACACAAGCCG
AACGTGAAATTCAAAACATGACCCTCAAAGCTCCCGCCACATCGGGATGGATCCTGTACTCCAGTTCTGGATTGCCTTTT
GATGAAGCGTTCCTTAAAGACAGAATGGAGCAAATCGTTATGAGTCACATAGAACACACGTGCTGTCTTCACATCCGCTG
GCGAACATCCTCACCGATGTACCAGCAAGTACGTCTCCTCCGGGGCTTCTCAGAAGAAATCGATAGATTCCTCAACATGT
TGCTACACCACCAACTGCACGTCATAGACGTGTACAACGCTATGAAAAACTTTGCACCATCACACCTCACATCGTTACAA
CGCGCCATGCTAGATAAATTACGTGCCGTTTTCCAACATTACGGTTCTATTCACGGACCTAAGGTGATTATGTGGGTGAA
AAACATCCTCAAAGCCATTGAGAAAACAAAATATACACACCTTTCCATGCGGCTCATGGAATATGCAGAAGAAAATCCAT
TGAATATAGACGAAACGTTTGTTTGCCTTCGACAGCAGGACTTGTCGGCCGTATCCAACGCGGTCAACAGCAGACGACAG
CTCACCAATCAGCAACAAGAACAGGTACAGCAAGCGCTGCAGCAACTATCGGCCGCCATCTACGGAATCGACAGCCACGA
CCTAGAAAGAACATCTTTTGACAACAAAGACACAAAACAACTGCTGGCTGTGCTAGATAGTGACGCGCGAACAACATTGT
TACAACGGGGAAACACAAAAATAACGGAGCTGCTGGACGATCTCAAACGCCAGGTGAGCGCGCTCATGCATAGAAGATAC
AATCAAATCATTAGCGGCTCCCTTCCCGTTGAAGAAATCCAGGCCATGCAAAAACGACTGAATCAAGCCGCCAACCTGGC
CCAAGAGATGGCCGGCCTTCATCTGTGCGACGTCAATCTGTCCACGCCATTCAGAGACGTGAATGAGCAATTGAGTTACC
TGGTTACCGGACATGCCTCTGCCAACTCCATGTCCTTCTCCGACGAACTCCTGCAGCTCCGTTCGCAGTTCATGTACGCC
ACTCAGTTGAAGGAAGACACCGAAAGCAAAATTCACGACCTGATGGTTGGTATCGAGACTGCTGTTCAAGATCCCGTGGC
CCGCAACTCCAACATTGCCATGGCCCTAGTCCAGGAACAGCTGAACGAGCTCCGGCAACTCGGAGGCGCCGGAATCTCAG
ACATAGCCCGGCCGCTTAGACGGCGGTACAGAAAACACTACAGTCCCTGTATCAAGGAAGAACAGGCGGCCCAGATYATT
TCGTCCACGAACCTCAGCTTACGACAACCTTACCAAANCGACCAAGCGCTCAAGAGACAAGGGgcGCCcTCCAACCACTG
CTGCGCGAAGACGACGGCCTGCACGAACTGTACATTCAAAAAATTCTCGACATCTTCAACGATCTGCTGAATCGTCTATC
TGAAAAAACATTCCCGAAACCTCAGGTGTTCGATGTCATTAATTCACTCATTAATCAGCTACCCAATACGTCAGCCGTGG
TGAAAGAGCTCCACACCTCCAACGCCGCCCTGTGTCAGTTAGCTAAACAACTGGAAGCCTTGACCAAAGCGCCCCGAGAC
AAGCGGATGGAGATACTCACCGAACTCGTACAATACTTTGTGTCCAACAGTACCTTGCTTAGCCATCTGATGAACCTACA
GGTAGGAAAAACAACGCTCCCGTTACTGTACGAGAATCYAAGCTTACGGTCACTTCAGTGGAAATGGTGCAACAATTCCT
TCAGACGGCCCCTAGCTCCACCGCCGCGGAGCTCGCACGACCCGAACTTCAGGCCAAACTCCAGGCTTTCCTGGAACAAG
AGGCCAAGAAACAAGAGGAAGACAGGAAAACCATCATGAAAGAGCAACAAGGCATGGTAACCTCGGAACTCGCTCGCATT
ACCGACGCCGTGAAAGCCCAAACCCTGTCAGTCATCCCCAACCTCAAACTGGAGTCGATCCAAGATGTAATTAACTCACT
GGGAGCGGACGGTCGCGACATATTGGAAAAGTTCAACCGCAACCTACTGGCCTCGTTGTCCACTCTTGTGAAACTTATGG
AAGAGCGAGTCTCCTTATGTATCCGGGATCTGCTAACTGGCCAGGATACACATTACCTACAGTRTCAGGAAGAAGCCATA
ACCATGCGCCWAGCACTKAACCAYGTGCGGATTATGATGGAGTCCCAACTATCACAAGAGACARTCCGMGCTCTGTCCAT
TTTGGTGCGACGATCCATGTTTATTGAAAAGTGTCAACTCCGAGATCCTACCACGGTGTTTTTTGACACCGACTACGCGG
AAGATTACAAACGTTACAGGAATGCTCAGAAGCAACTGGAACTGCAATTGACGCAGGCCAGAATGGAACTGTACCAACAG
TCCGCCGTGGCACAGCAGACGCTGCACCAGCCCACCGCGCGCATTACGCCTCAAACTCTGACGGTAGGCAAAGACCTCCA
GGACAGATTGGAGACGGAGTCAGCAGGGCCATTCCAACGACAGGTCTTCCAGGACGTGCTGCATCAGCAGCTGGACGCGT
ATAAAAAAATCATTAAAGACGAAACGGAGTTAACAAACACCAAACTGAAATCTGAAAACGAGCTCCGACAGGCCAAACTG
ACCTCCCTGTCGGAGCAATGGAGCGATCTGGTGACACGACTCAAAATGGACGCCATGGAGATTACCACTCCAGACGCCAA
GGTTTTGATTCAAGATCCCGTAGGTGCCATCACACAATTGCTAGCCAAGGCTTCCACACAAATGTCGTACCTGGAGGCCG
AAAAAACGTTACAGTGGTGTCTATTGTTTTTGCGTGAAGCACATGCCCAGATCGAAAGCAACCTAGGGCACCCCTGTTAC
GCACAGCTGCCCAACTTTCCCACACTGATGCAGCAGGCTCAKAGCASACTGGACACTGTGTCCGTACACGTGAATAACAA
CGCTTCGTGCGAAAACTTTGTAGCGCAACACGACTCGGCCTCCTCGGCCGCGGATCGCGAAAATATCGAAACAATCGAAA
CGGTGTGGGCGACTCTAGAGGCCAAACGCGTCGCCGGAGGAGAGGCCCGTTACAAAAAAGTGAAGGAAGTGCTGCTTCGC
ATGAAGCAATCGCTATCAGATGTGGAACTGCAGGATACATTGACCACGGAATACTATGATCTGTTGACCGCCATTCAGGC
CTTTGGTTACAGTCTGGACTTTTCTACGCCGTTGCAGAAGATCCGCGACTTGAAAATGCGCTTCAACGATCTCATCAAGC
AGCAACACCTAAACACGTCCGAACAAGCGCCCCTTCCCATGGCTCTCCTATCGGGAGATACGGCTAACACGTCCCCCGTG
TCGTTCTTGAGAGGCCTAGACGCCTTGGAAAGATACGTTTTGGCAGGCTACCAATATCTCACAGAATGTATCAACCGCCA
GCCGCTCATCTGCCAGCGCATAGATGACATTCCGGCCGTGCTGCCATCGACCGACCTCGATCACAAGAGGGCAGCAATGG
ACCGCCTAAAGCGCCTGTCCTTCAGTAAAAAGAACGACACGGTGTATGAGGTGGTGGACGTGTTCGGTCTTCATCAATTC
ATGACCAGAACGGGGGTTCCCCTCCACCTAATGATAAGTTATGGAAACGTGTTTTTCAAGTACCTGGCACTACATCACGA
CGATAAGCACTTGGCTAAAAAGTTTGCCCAGACGAAAAACGTGGTGACGGGCAGATACAAAGTGGTCATGGTGAATGTCG
CCGTGGCCCAGACCCTAAAAACCTTCTGGTCCCAGATATCCCAGTACGATCTGAAACCGCTGCTGTCCGGTCAGGCAATG
ATAGGCTTAGGAGAAACGAACAGTCTGGTGAACCTGAAAATCTTCATTTATATTGTGGTTTCGGCCTGGAATCTGCAACT
GGATATGACGTTGCAAGATTACAACGGCCCCGTAATTCGCATGAACATTGAGGATTTGTGTATTGCTATTACCACGGTGT
ACCCGGAGTACATCTACGGTATTGTGAAACACCCCATCCAGACCACCCTGTCCTCGCTAGCGCGGCACCTGAAAAAGGAC
ATTATGCACGAGGCCATTAACAACGTGGTTCAGATGCCACCCGCGTACAACGCAGACGAGATGAAAGGCTTCTGCATTAA
CCCAAAAACATGGCCGTCTGTCAATCTGAGCCGCGTTATGTGGGACCAGAACCTCGTGCGTCAGTTGTGCGATGTCGGTC
CTCGTAAGAACGGGGCTCAGAAGCTGTGGCAGTACGCCGTAGCTATCGTCATTTTCCCTCAAGACCTGTTACAGTGCCTG
TGGCTGGAACTGAGGCCCAAGTTCGCCGAGGAGTTCGCCACGCTGTTTGACTTCTTTCAGGCTCTCTTTGTGATGTTCAC
CCACCAGTACGATATCACCCGGGAATCCAATGCCCAGAGCCACCTGGCCACGGGAGAACCCATCACACAGACCGTTGGCA
TTCGACGCAAAGACTACACCGACAAAACTCTGCTGGATGTTTTTATTGAGACCGACACCGCCATCGACTTCGTCTTAGGC

AGTTGGGTGTTCGGCATTCCGGTGTGTTGCGCCATCTACGTGTCCGAAATCCTCGGTGGCAACCGACTGCTGCTGGCTCG
TCACATCGAATACAGCAGTCGCGACCCAGACTTTATGCACGTCCAACGGGCCAAAGATTTGAATCTGAACCATATGGTCA
CACAGACGTGGACCAACACCCCGCTCGAGCAATGCTGGTTTCAGGCACAAATCCAACGCATTCGAGAGAGTTTGAGAACC
CCCATGGAGTTAGATTTTATTCCTTTAATCATTTACAACTCCAACGACAGAACCGTACACTCTGTTGTGCGTCCCCCTAC
CAGCAGCGAACGAGACACGTCTCGTATCGTAGTGGAAAACCCGTTCCCCACCCTGCCACTGGTTGATGTCCCCGAGTCAG
ATCTCGTGTCTTTCACCCGCGTTCCCATCAACACCGATTTTCTTCGGGAAGATCCACCACCAGTGACGCGCCCCAGAGGT
CGCGCTAGTGGTAGTACACACGTAAGCACAGGCAGCAAGGAGAGAAAAATTCTCAAAGCAGTTTCCAAAAAGAAAACGGC
TCAGGCATCCGCCATTCCTCAGGAGGTGCGGGTTAGTAGACCCACGAGCTCGTCGGCGGTGGACGTTTCCGAAGTGACAG
AGGAGGAAGTCACCACGTCCGAGGATGAGGAAGAGGTACTAGCGGCGTTGGAAGAGAGACGACCCTCGGAAGAGGAACAA
CAATCGCTTCCCTCCACGCGCACGTTCGCGGAACCGGTGTTCCAACAATCGGCCGTGCAAACTATTGTCCGCAGGGAAGA
TGTCATTACTGTATCTCCCGCCGCCGCCACCTCCAGCTTCCAAAATCCCCCCATAACCACCCTAACTCAAAACGTGCTGA
ATGCCGTTCAAATCCTGAGATCCGTGCGCGTTGATTTGCGAGGCATGGCGAGATCAGTCAACGAAACCATTAACCGTYGC
GCTTCCTCTATCTGCTGTATACTTGGTACAACATGGTGCCTATCACGGGGTGTTCCATGGACTCCTGTGGCAGATCCAGC
ACGTTGAGGCAGACGTGACGTCGTTTTTCTTCATCCTTCTTGGCAGCGGCGCCCGTGTTGCTGGAGCTAGACATGGTTAC
TGCCGTGACAGGCCGTCTGCTTCACCAAGGCGCAGCCCCTGCACAGCGACTCGCAATCCTCGAGTTCCAGACAGCTGTCG
CGATACTCGCCCGTCTGGCAGCGCGAGCACACCGGCCCGTTGCCGTGACTCGTGAGcTTGAACAGCCGCGGGATGGTGAC
GGAGCGCACCACCACGGTGGAATCGCAGGTCCGCGCGGCGCACGGCAGAATGATGTCGAACGTAACAACGTGGTCATACA
CCGCGCAGGCGGCGTTGAGCCCCAAAACCGCCTTCCAACCAACCCGCAAGCAGCGATGCCCACAAATGGTTTCAGACACG
ATCTCATAAACACGCTTCTTTAATGACACGCTGACTTCCGCAAAGGGAACAGTGCACCAGTCGAGAGGAAAATTCTGCTC
GCAGTGCAGCTTCTCCTTGCCCAGGTTCAGGCAGTGGCCGCAGTGGCGACACACCACGGCGATGAGCGACCTGTCGTCTA
GGTGGCGCTCGCACTTCAGGCGACACATGCACCAGAGCGGAAGCTCCAGGACGCTGCCAATCAACCCTCCACGCAACGCC
GCCGCCAGCCCATACATGACGATCTTGGTGGGCTCCAGAGGCCGCGGCTGCTTGTAGGCGCGTACCTCGTCGTTTAGTCC
CAGCGCGCGCAAGGGCACTAAGTACCGCGGACACGTGTCGCAAAACACCTGCATGGCCTGACGAGCCATCACGTACAGGG
GATTTCCACATGGCACGGGTCCAGCGTGACGGCGCAATGCCGCCACCAGCAACCCGAGCTGCGTGGAGCGCGGTTTGCCG
CGCCGCGTCCACTGGTTCCGATGATAGACAGCCAGGTCATTGCGGCCGTCGAACCTTTTGGCGCGCAACAGGGAGAGGCA
GGGAGAGCTCGTCTGGTTGAGCACGGGGAGTCGCAAGTCCACCCCGAAGAGAAAGCGATGCAGATCGCCCAGGGCAGACG
AGGTCATGCCATCCAAGAGCGCGCTCAGCTTTTCCACCGGATACTGTTGCAAGCGCACATACACAAAGTTAGCATCGTAG
CGCTCGAAACGCAGAAAAGACACGATCCTGGAAACTCCGTGAAAGCAGCTCATGCCGTACACGTAGGCCAGGAACATGAA
GTAGGGCTTCGTCATCACCATCCGCTGAAACAGCGTGGTCACCGCCTCTCGCTCCCCGGGTTCGCAGGCTAGCCATTCGT
GCAGAAAGCGCTGGTACAGACGATCCCCGAGTTCGCGGTTGAGAAAGCGCTTGTCGGTGACAAAGAGCTGCAGGGGGCAG
ACCCGCTCCTCGGTGTGCCAGGCGCCGCGCCCGATCACCATCTGCACGTGGAAGTGATCATCGTGCACGCAGAGGGGAAC
GAATAGCCGCCGGAGCGCGCGGTGAAAGTCACTGGACATAGCGATAGAGCAGGAAGAAGACAATTCCCATCGCCACGCCG
ATTAGAAGCCAGCGCAGGTCCACCAGGCGCGTCACGCCTAGACGCACCTGCGCGAACGAGCTATTACTCCCCCTGGCCAC
CGCCCCGTGGAGCGGCTGCTGCTGGTGCTGGTGGTGCTGCTCCTCCTCCCGCCGAGCGCTACAGCACCCGTGGTCGGAGG
TTCGGCAGCGACACACGCCATAGAAGCCGCCCGGAGGAGGTTGCGCCGCCACTCCCGCCGCTCTGGTATCGCGAGTGAAC
CAGCCTGCTGGTTCGGAGAACGGCACCCGCCCGCGGCTGCCGCCTTCTTGGCACAGAGTTCCTTGATGAGGTTCTCGTAC
TCGGTCTCGTTCTCCGGGCCAAAGGCGATGAGCTCGATGTTGAAGACCGAGGGCGAGTTGGACTGCGTACCACGCACTTA
GTGAGCACCCCGTAGGCCGAGGGCTTAATATCCTCGATGTCCTTTAGGGTGACGATGAGCGACTCGTTCACCTTGAGCAC
GTTGAACTCGCCGACGTGTCGAGACGAGAGCAGTTTCAACGGCGCTCGAACGAAACAGCAGAGAGAAACAGCGCAGCCGG
TGTTTTTAAAAATGAAGCAGGGCACGTGATCAGTGCGACTTTCCCAATAGCTCAAGAGATACTCCACGCAATAAACCGTG
TCAGTCTTTAACATAGCGTCACACACAGAGTAGTTGGGGTTCTTGCAGATGAGTCCCGCGTCCGTGACCCGCAGTTCGGC
GGGCCCTAGTTTCAAAATGCGTCGCGTGGCCGAAACGAGGTCGTGATGCAACCCCTTGCTGATTTCCATGACGTGAAGCG
CGCTGGCGCACGTTTTATTTACCCGCAGATTCCTCTTTTCCCTCCAAAATCCCGTTAATGTCCATGAGCTTGTTGACGAT
GGCCGTTAATAGTTTCGTCTTTTCCTGCAGAATTTGCGCATGACTACAAGAGGAGCACTCCCCGTGCACATACTTGAGAA
ACGCCGCGTACTTTTGACCCGCGTTGACGAAATTCAACCTCGCGTCCAGTGACGGCAGCAGCAGCTCGTACACGCGCGGC
AGCATCGGCTCGAAGCTGACCTCGAAGTCGTCGTTGGGATCGAGGAAGCCGTCCTCGTCGACGTCCTCCCGATCCCGACG
CGCCGCCATTGTTTTTTCTACCTCACCACACGCGACCCCTCCTTGCACTTTATTACTACATACACATCACCCGGCGACAT
GAACCACCAACCTCCACAGCAACCACATCAACCACAGCAACCGCAGCTTTCGTCCCGCAGCTACATCTACGGCGGCTGGG
AAACGGAGCTCACGCAGCCGCTGAACGAACTGACAAATGAACTGCTATGGCACGCGCACCCTCGTCATGTACCCGTCGAC
ATGGCCAGCGCCACCGGCGCGTCGCAATCCGGATCCTCCACCACGGGCGGCGGCGGTGGTGGCGGAGGGGGAGAAACACG
CGCGCTCCACAACTTCCTCTTCAGCCAGTCCCAAACCACCGATCAGCTAGCCGGCCCTGGCGGCCTGCAGTTCTGCGATC
TCGACTCGGAACTAGCCAATTTCGAGTTCTCGGACACGGAACGGCTCAAGAGACTGTGCACCCCGCTCGACATCGACACG
CGATGCAATCTCTGCGCCATCATCAGCATCTGTCTGAAGCAAGATCCCGATCAGGCCTGGCTGCTCGACTACAGCCTGCT
GTGCTTTAAATGCAATTCGGCACCGCGCACCGCCCTCAGCATGCTAGTCACGATGTCCGAGTTCATTCATCTTCTGCGTC
AGTACTTCCCGGACCTGCGCGTGGACCAGCTTTTTAGCCAGCATATTTTGACGGTGTTTGATTTTCATCTGCACTTTTTC
ATCAACCGGTGTTTCGAAAAAcAGCTGGGGGATGCGGTGGAGAACGAAAACATCACCCTCAGCCATCTGGCCGTGGTGAA
AGCCCTGGTAATGGGCGACGAGGCGGTGCCGTACACCAAACACCGGCGTTTTAGTCACCACAAGCAGAAAATCGTGGTGA
AAATTCCCGAGGCCCCCAAACAGCTCCTGCAAACGTTCACGGAGCACAGCACCCCGTCTCCTGACCGCTTTGCGTACCTG
CTCTTTTACATGTGGGCGGGCACCGGAGTGATGGCGAGCACCCCCATGACAGAACTCACCCAAAGCAAGCATCACCTGCT
CGACAGCGTGGCGGACGCCGACGACCCGGAGATTAACAAAGACGTAGGTCCCATCTACCTGTGCCCCGTGCCCGTCTTCC
AAGTGAAGAACCAGACCACCACCGTGTGCCTGCTGTGCGAGCTGATGGCCTGCTCGCACTACGACAACCTGGTGCTACGT
AACTTGTACGAACGCATCACCTCGTACTGCGAAAACAACGTGAAAATGGTGGACCGCATCCAGCTGGCGCTGGCCGACAT
GCTGCGAGAGTGTTCGGTGCCCGCGGGCCCCCAGTCCGACGACATGTCGCGATTCACGCGCCTCCCGATCCCCACCGACT
ATCTCGCCGTCCATAACGTTCGCCCCGACCCCGTCTTCTGCACCATCCTGCGCCAGGCCGGCGTGACCGGCATCTACAAG
CATTTCTTTTGCGATCCGCAGTGCGCGGCCAACATCCGCATCACTCGCGAGGAGGTTCTCTTCGGGAAAATTCATTTGGA
GCACCtGGCCGAACTAAAAGTGTTGATTTGTCACGACAATTACTATATAAGCCAGATGCCCAGACGCGTGTGGATGTATA
TCACGACCTTCAAGGCCTTCCAGATTACAAAACGCAGCTACAAGAACAAAACCCAGCTGTCGGACTTTATGCGCGAGTTC
ACGCAGCTCTTGGAAGCTTGCGAAATCAAGCTAGTGGACCAGTCGTTTGTCGTCGACAAGTATGTTTAGTGGCGTGCGTT
CCACCGAGCGGCGCTCTGCGCTGCGCTCGTTTACCCGCAAGCGCCGACCGCGACGCGGCAGCCTCCTGTCCCCCACCACC
CAGAACAAAAAGCCCAATACCGGCACGGCCATCTCCGCCCCCCGGCCGCGACTGACGCTGCATGAGCTGCACGACATCTT

TAAGGACCACCCGGACCTTgAGCTCAAGTACCTCAACATGATgAAGATGGCCATCACCGGCAAAGAATCCATCTGCCTGC
CCTTTAACTTCCAATCCCATCGCCAACACACGTGCCTGGACATCTCGCCGTACGGAAACGAGCAGGTGTCGAAAATCGCC
TGCACGTCGTGCAGCGAGAACCAGGTGTTCCCCACCGCCTCGGACGCCATGGTGGCGTTCATTAACCAGACGTCCAACAT
TATGAAAAACAGAAAtTTTTATTATGGGTTCTGTAAGAGCAGCGAACTGCTGCGAATGTCTACGAATCAGCCGCCCATCT
TCCAGATCTACTACATCATCCACGCCGCCAACCACGACATTATGCCCTTCATGCACGCCGAGCGCGAGCGGCTGCACATG
CACATCATCTTCGAGAACCCCGATGTGCATATCCCCTGCGAGTGCGTAGCGCAGATGCTGTCGGCGGCGCGCGACGACTA
TCACGTCACCCTCGATATCATCAAAGGCCACGTGGTCATCAGCGTGGTGTGCCACAACCTGGTCGCCAACAGCGTCAAGA
TTGACGTGTCTGTGCTGCAGCGCAAGATCGACGAGATGGACATCCCGAACGACGTCAGCGAGTCTTTCGAGCGCTACAAG
GACCTCCTGAATGAACTCTGTCATGTCAACAGTGCGCGTTCGCCTTGACACAGTACGGTAGAAACGAGGACTCTAGGTAC
ACGCGGCGTGGTAAAATGCCTACAAGAAACTTGTCCTTTTTCAGGGCTCCGGGCGGAAACACAGGGGAAAGAACGTTGGT
CACGGCCTTCACCACCTGATCGAAATACTTGTCCGCGTGGATCGGCACCCCGTGCTCGCGCACGTAGGTGGGATCCTCCG
CGATTTCGTAGTTGCACACCCCGGTGGTCCGGCCGTCGTGAGGCGCCGTTAGCACATAGAACACGCGATCCCCCACCGTC
GGCAGCTCCTCCGAACGCGCCGCCAACCGCTTAATGACCGCAATGTGCGGTAGGTGCGCCTGCTTGTAAAGCGAAATGTC
TTTTGACAGGACAGAGGAGAGAACTAACTCCTCGATCCGGGCCCGCGACGTGTAGAGGTCGTCCCGAGCCTTACAGATCC
GCTGGAGGATACGCCAAAACCCCGAGGGCACGCCCTGCGTCTTGAGCTCCTCGAACGTGAGGTGAGACAGCTGGACGGCA
GCCCGAGACACCTCGGGATCGTCGAAAAGCAGCTGAATGACATCGCGCGTGACATTCTTGACGAACTCACACGACGTCTT
TCGCACTAGATCCACCCCCTTCATACTAAGCGTGGACTCGCCCGCCACCTTGCCGATGTAGCGCTTTTTGCAGATCATCA
TTAGGGACACAAAGACCTTTTCGAATTCTAGCTTAATGGGATCTATGAAAAGACACTGTGTGATGTAATCAGCCAGGCTG
GAGCCGCGGTTAATGAGAGCCGTGGCGGTAATGCCGCGGTAGCACACGAACACGCTGTCCGTGTCCCCATATATCACCTT
GACTTCCACCGGATCTCCCGAGTAGTCCTCCCGATTAAAAAAATTGTGTAACAAACACGGCTCCGCCAGGTTTTCTTCCA
CAAACTTTGACGTGGACGTCAGCATATCGCGACCAATGCGCGTGATGCTGGCGGCGATCGGCAGGCACGGCATCATGCCG
TTGATGACCCCGGTGAAACCATAGAAGGCGTTGCAGGTGACCTTGAGCGCCAGCTGCTCTTTGTCCAACAGCATGCGCTT
CACAGGATCCTGGCAACCCTTCATGGTTTCGCGCACGGCTCGGCGCTGCGACACCCACTTGGTTAGCAGCTCCGAGAGCA
CGGACAGACGCGCGGCGCTTTTCGCAAAGCGATGCGTCACCCCGTTCTCCAGCTCCACCGTGAACACTTGGTCGTCCGGC
GGACTCTCGCCGTCGGGCATAATGAGCGTAGAATAGCAAAGGTTATGGGCCATGATGATTGACGGGTACAGGCTGGCGAA
GTCGAACACCGCCACGGGGTCATTGTAGTACCCGATGGCGGGTTCGAACACCGTCGCGCCCTGGTAGGAGGAGACCTCCG
GGGCGTCGGACGAGTCCGCAGCGGCCCCTTTGTGGTTCGGCAAGATGAAACCCCGGCAAGCGCACTCGTCCAACAGGGAC
GTGTAGATGCGAATCTGCTGTCCGTCAAAAACGACGCGCCGCATGGGGATCTTGGCCAACCGCGCTATGGCCCCAGCCTC
GTAATGGAAATTAATGACGTTAAAAAGATCCCTGACCAGCAGGGCATCCTGCATACAGTACTTTCCCACGTCCGCCCGCC
CCTCTGAACCCGCAATGAACTTCACGGGGATTTCCTTGTAGGACAGGTCATCCTTTTGCTGCTTGAGGTACAGCTCGGCC
ATGGTGTTGAGTTTGTAGTTCGGAGAGTTGGTCTTGGCCATACACACGGGGTACATGTCAATCACCACGCAACCCGAAAT
GAACACCTTGGTAGCCGCGGTGATGGCATGGTGTTTTTTGGCAGGAGTGTATGTGTGAAAACGGCCGTTCAAGGCTAGCT
TGCTATAATGATTGGCGCTGATTTTGTACACCGTGTCCAGCCGCGTCAGAATGTACTTGAAATCGAAAGCGTTGATATTG
TAGCCGGTCACGAAACACGGTGAATACCGCCTGAAAAAAATGAGAAACCCCACCAACATCTCGTACTCCGAGGGGAACTC
GTACACGTCGGCGGTCCCCACTCGCCCACAATCGCCGATCGTGAACAAGTGCATTCCCGAGGTCCCATACACCGTATTCg
CCTCATTACCCGTCCCTCCCGTTTCATAACATACACACGAAATCTGAATAACGATATCTTCCACTTGCTCGGCCGCGGGG
AATCCGCCGCTGGCACTCATACACTCAATATCAAATGATAGACAGCGATAGACAGGCCACGATGCGTCGTCGGGTATGGC
CATAAGGTCAGCCACGTCGCAATCAATCTCTAGATCGCAGGTCGAAGACTTGTTCTTGAGCCGCCACTCAAAGCGGTTCA
CGCAGCACCACCCGAAGGTCGTCATCTTCTTGTCAATCACCAGTCTGGTCAGGGGGTCCACGCGGATTTCGTACACCGGC
ACCCCCTGGTCCAGGAGATACTCCCCAATCTTTCTGGCCATGGTCCAATTGCTCATGGACACCTGGTACAGATTAGGTAC
GGGTCGTGTCCCGTACCCGTAGATGGATGTCTTAGTGACCTGCACAATGCTTAAGGCATAAGACATCCTGGGCTCAGACA
CCAGCTCCGAAGCACTGGCCAGCAGGTCCCGGAGCAGGTCGCCATCCTGATACTCACAGTAAAAGTAGCTCCTCTGCCCA
AACACATTGACACAGATACTGTGTCCCTGCTCGGTAGCACCGAAGAAACGCAAGACGTTTCCCGATGGTGTGATGTGATG
TCGATACCGAGGAGATATTCCCTCGGCTGTATCGTGAGGCAACACCCCATCCACCTGGTCGTAAGTGTGAAACTTGAGAG
GTCCAGACGCTCGTTTCAACAATGTTTCCAGCGGTGGCGACGACGAAGGTAAAGGACACGGCCACGCCATGTCATCACTC
AGCACATGTTTAATGTTCTGATAAAACATAATCGGCGCCCGACTAGTTTTATGCTTAATCAGCCCAGATTGTCCATCATA
CATAGCACCGCGAGGGACGATCTCCAGAAACGTCTTCTGCACAGGACGCTTGACAGCCCACGCTGGCGATTTTCGGGCTC
CACCCAGGTAAGGATTGAAGAACATGTTCGCTGCCGAAGAAGCTCCCTCACACCTCGTACTCGTTCTCCTCCACCGGTTT
GTAGCCCCGACGGCGGTTGCGAACACGGTCCAGAAGGCCAACCTTACCGCTCTGTGAGGCAGCCTTCTCCGATTCCTCCT
TGAGCTTTTCACGTTTTTCAACGTCCAGTTTATGCAAAGCCAGCAGCATCTGAAAAGCCTCCTCCGAGCTGAATTCTTTA
ACCGAAGCAGACGATTTCTCTTTTATGCTGGCATAAAGGCTTtCCTCGTACGAGGGTGGAGGCGTGGTCTCTTTGGCTGG
CGCGGGCGGTGGAGCAATATACGGGAAAAGTTGCTGGATAGGCTGCTGGTACACGCGTTTTTGCCGCATGTAAATCAAAT
AAATAACACCCAATACAGCAAGCAGAAAGAGTATGACAGTTAAAGAACCAAACGGGTTTTTGATGAAACCAACAATGCCC
TCCACAAAGGAAGCCACGGCACCGCCAACGGCTCCGATGGCCACACCCAAAGCTTTGCCCGCTGCACCAAGACCGCTCAT
CAAGTCGTCCAGGCCCCTGAGATATGTAGGGACGGTGTCGAACACCTTTCCTTCCACGTGAACGACCCGTTGCTTGTAAG
CGTTGAACTCTTGCATAATCTCTTCCAGATTGAAAACATTGCTGGCACGCAGTTCATCCTCGGAGTACAGTTCCAAAGCT
TTGAAATCTGTGTTTTCCAAAGGGTCGATATCCAAGGAAATCATGGTTTCTACCACTGGAATACTATCCAGGGGAACTTC
TCTCTTGAACAGATAGTCCACATATTCGTACGAAGTGTTACCAGCAATAAAAATTTTCACGCTGGGAGTTTCACACGCTT
CAGTACGATGACGCCCCAACAGAATTTCGTTGTTTTCTCCCAGCTGACCATACTGAACATGCGAACTATTGTAAAATCTG
TACAACACCACAGGCCTAGAATAACATAGATTGCTTTCACCTTTCACATTGCATGTCACGCATCACTTGAACTGTAGTTTG
GTTAACTTCAACACACTTAGCTAAGCTTATAACATCTCCAGCATATCGAGCAGCAACTGGCTTATCGTAGATGGCAGACA
ACATGGCAGTAGGATTTATTTTACTGAGTTCTTTCAGGACTTCGAGTGTTCTTTTTTGGTCTTTACACCAGGCTTCAGCG
ATCTGTCTCAGAGCACGGTTAATGTAGTTTCTCAAAGTGTCATAAGTGAACTGCAATTGCGCATACACAATGTTATGTAT
TACTTCAGTATTGTTCAATAATGTACTGTCGTCTGTAGATCGCTTAGTACGTGAGGATGTAGTAGCATTTGTTATGTTTG
CGTATCCATCAGCAAGTTTTTTCATGTCCCAAATTGATTTTTCTTTTACTGGTAGCCAAAATACAATCAATCCACCAGTT
GTTTCATAGATAGTTACATTGCCTGCTTTTACATATGTAGAATTATATGTGGTATTAAAAATTTCTTCAATCTTCTTTTC
AGCCTGTTCTTTAATGCAATTTAGGGCAGGATCCGAATCGTTCACCGGTGTTTTTTTTGTCAAAAAAGTTGCCGTCATAC
TGTTAGAAGTAAAATGGTAGGTATCCTCAGCTTCAGAGCGAATAGTGCGGTCAGACTTTTCCCAAAACGTGAATTGACAT
GTACTGTTACCTTCATCCAACACATCCCAAGACATCATAGAATCAGCTCTTTCAAAAAAAGCCACCAGATCGTGGGCAAC
CAGTGGAACATTTTCCCTTCCGTAATATTGCACCATGGAATAGTTTTTAATAACATGAAATTTATCCCTATTTTCACCAA

AATGTTTGTCATTGCTTCCATTGTAGAAAGGAGAAATGTCCACCACTTCACCAGACGACGTTACAAAATAATCATATGGA
TATTTTGATCTGGCGGTCGTTACAGTCACCATGCAGTTTACATTAGAACTTGTTGTATACAGCCACGTGCTACCGGGTTT
ATGCCAAAGTTCTTTCACGGTCACAAATCGCGAACTGTGCGTATTGGAGTAATCTTCCAGCAGAAGCATAGTTTCATTCA
CATAGCTGTCTTTATGGTACACTTCGTACCACTCGCCTGCTAACTGTCTCTTTACAGAGTTGTAGCACCTGTTaATTCTA
TTCACATAATGCACTTCCCACATTGGAATAGGCAGATGCTCCACAGTTTGTCCCAGCATATAATTCATATAGATGTAGGA
ATAACTCTGGCGTTGCGTCAAAACCTTTCGATACACGTTCACTTTGAACGTATGCGCACGAATGTCGCGCTTGTAGACCA
CCATAATACCCTCGTCAAAATCATCTTTGGTGGGTTTGAATGATTCACACTGAATGTTCTGTTCAAATCGAAGAAAGTCC
GTTCCTTTAGCTATTGAACACACGCGATATGGGAATTTTGAAACGTTAGCCTTGATGACTTCATGCGTTCTCAAAGTTGT
GTTTTCTATGAGAGCTCCAGTAGTCGTGTTGATGGTAGGCCTTGGTGTAGTTGACGATAACTGTGAAGCAGAGTTCGAAG
ACGACGCGTGAGTTGATGATGTGGCCACAGACGCATTAGACGTATTTGTAGTAGAAGAAACACTTACCACGCAAGCCGCC
CACAGGCTAACATAGAGAAGCCGGAACCAGTTCTTGTTCATCTTGATCTCGCACCAAATCCAGCAACAGAGAAGCGTTTC
TCGGACGCTTCCGAGACACTCTCACCCTCCGAGAAACGCTAGTTGGCACGCAGGCACTGTCGGGTTCACGAAACACCTCG
CGTTGTTCATCCGGCCTGGCGAGATGCCTGTACAGGAGGAGATATAAATCCTTTGATTTATAAAGTGTGCCTTCGTGACG
CAACAAAAGCGGCCACTCCTTGTTGTACGTCAATATGAGATCCTCTTCGGTAGCAAAGTCGTCACGCAAACACGCAATCG
TGAGCTGTTTTCCAAAGGTTTCGTTATATAGTGCGACCGAAAGTACCAGTTCCCGGACGTGTTTCCACATCTCCTTCTGC
AACATGTTGAGATCACTACAGTTACTAAAATTAAAAAAGCTCTGATACTTGACCACCATCCACTCGCTAGGGTGTACCGT
CCCCTCCGCGCACTTCACTAGATCATCCTTAATGTGAGGCAACACTCCGGCGTTATCACAGGCGTATGCCATGTCTACAT
TGTGAGACAGAGGATAACGATCGGTGCAGTGCGTAAAAAGGGGACCGTTACACAGTTGATAGATCTGCTGACCTAACAGA
GGAAGGGACTCTGCAGGGAGACTTTTGTGAATGATGTTGTTCACTACATATAGGTGTTCGTCATAGCTAAACATTTGTGA
CaCGTCCACTACGTCTGGGTTTTCTCGGTACTGGGTGCGGTAAATGAATCCATTCATGAGTCTGGAGATGAAGTCCAGGC
AGACCGTGCCTATGAGGTTCACGTCTATGAGCTTGCTGGTCAGTCTCTCCTGGGACTTGATGCAGCGAATGACTTTGGCG
TAGCCCACCTCGGAGACCTTCTGTAGGTAGGCCCGCTTGCGTATGTTCACCTCGCGCGTCACGTTATGCGCCCGCGTCCG
CGCGTCCAGGTAATCCAGAAAGTCCTCGTCTTCGTCCTGCGCGGCGCCGTCACTGTTGTTGGACGTATTCTGCGGGTCCT
GCCCACGAATGGCGGACCCGCTCCCCGCCTCTGCCGCCCCCGTTGCCGCCCCATCGGCCGCGGGGCGCCTGATCTCGTCC
AGGAGGGACTTGATTTTGGTATACATTTCGTTGCTTTCGTGGAGCTTGTTGAAAACCGGGTTGTCCTCGAAAGCTTGAAT
GCTGAGGGACGTGATGAGGTCGATGATCCTGCTGGGGGCGGCAAAGACCGACCCCACGAAAAGACGCTCCTCGCGAGAGA
GCGCCTTTTCCCCGACCACGAAAATGTCCTCCACGTCCTCCCCGTACAGGTAGTTGCTGAGGGTGTTCATGAGGACCCTG
CACAGGTGGTGATAGATGTTTAGCTGCTGAATACTTATCCCCACCCGCTTGACTATAACCTCCGAGGTACGGGACCAGTA
GGTAAAATCCGACAAGGAATATATTCGTGCCGGTATATCCGTAAACAGGTTGTACTCCTTCAGCGCCTCCTCCGCCTCCT
GGATATAGCTGTGGTAGACCGATGAAGAAGAGAAGAGGTTTTTGAGGGCCGTGAGCACCCCAGAAAGGCTGGGGATGCGC
TGCGTCAACTCCAGGAGGTCCTGCTCCACCGTCTGGATATTCACATCGCACTGACTCGACGGCCGGTGGACCACGATATG
GTTGCAGAGCAAGCCCTGGAGACGCTTGTTCAGCGAGCGGCCCTGGTTGGGGATGATAGTCAGCTCCTCGTAGCATTGGG
CGCACGTCGTTCCCTCGACGTACGCCTCTTGGCAGGCTACCGGAGACACTCCGCACAGGCGACGCAGCATCTGCAGGAGC
TGCGAGCACACCTCCAGGCCGCTCTCCGGCGACAGGATCGCGTACACGTAGTTCATCTTGCACAGGAAGCGCTCGATATC
GTTGAGCGTTGCTAGATTGACGCTGAACTTTACGTGATCCGTGAAAGTGAGCTCCACCGTGTGATGCCTGTCGCACCGGT
CTAAGCGGAGGATGGTACGGTAGTACTCCATTCGGTCCGGCTTCCTGGAGTACTCCTTGAGGGCGCGGTCCAGCAGCTCC
GTATCCTCGTACAGAGCCTTCAAAATCATCTCCAGGTACAGGCTCAACAACGAGGTCTGCGTACGGTTCTGCCGCACCAC
CTCCGGGTAGATCCTGCGGTACAGATACGCGATAGCCAGTGCGTTCCTCCTGAACGGACTTGACTCCACCAGCAACACGT
TCGGATCGCAGTACTTTAGACACTCCAGTTCCATGGCGTATTCGTTGCATTTGGAACACACTACGCACAGTTTCTGCAAC
AAATTCATCTCCCGTGTTCTGCTCGTCAGAGTAGCACACGGCGTCCTCCGTGATGTGCTCGGAGAGAGACATAGGGACGG
CACACAAATAACTCGCTCGTTGCACAGTCTGCCGCGACGCCGAACGTTAATATATAACGGCGTACACGGCGTTTATTACA
GCCGGCCCCTCTTGGCCGGCAAGAGGCCACAGTCGTCGTCGAACGCGTACTCCCCCCCGCCCGCGTTCATGCCCACAGAG
TTCCCCACCGCTCCCCGGCCGCGGCGGGAAACAACGCCGAGAGGTCATGCACCTCGCCCCCGCCGCCGGGATCTGGCCGG
CTGTCCAGCGCCTGCTGCAAGTTATCCAGCGAGAAGTCCTCCACACCCATTTCCCGCAGCCGCGCAAACTTCTCCATCAG
GGAATCGGCCACCACCTGACAGCCGTCCACGAAAAACAGGACGTCGTCCCGCTCGGGAATCTCCTCCCGAGGTGCTGAGGA
TCTCGTACACGGCCATAAGCTCGGGGATCCACGTCCAGATTCTCGGAATCCAGGGCCGCCATGACCCGGTTCTTAATAAT
GTCCTCGTCAAACAGCACATTTTCGCGGCGCGAGCGCTTAATGAGGACGTCCGAGAGCTTGGTGGCCAGAATGCAGCGCT
GGCGCATGAATTTATAATCCTGACCCCCTATGGAGTTGGTGTTCAGGCTGCGCTCCACCCCGTTCCCCGAAAAGTAACCA
ATCTGCCCAAACTGGTAGATTTCCTTGCCACTGTTAACCCCCGCGTACTTCTCAATGGAAAAACTTACAGTAACCAGGGA
CCGCTGCTCAAAAATGGTGCGCATGAGAGACTCTCGCGTGACCGTAGACGCCAGCGGCGCCGTAACGCCGGTCGCCCTGA
GATCTCTCCAGCGCAGGGCAAAGTCGGAGGGCGACGCAAAATTCGGTGTTTCCCCTAGCAGGTAGGGAAACTGGGCATGG
TTCTGATAGGTGAGCCTGGTGGTGAGGCCGTGTAAGGTCCCAATGTAATCTTTGAAACCGTAAAAACAGAGGAACTTATT
GTGAAAGCGGTTCTCGAGATAGTTGAGCATGCAATCAGGTGACACGTCAATCAAATCATGCTCGTTGTAATCAGCGGTAA
CTGACATGACGAACTTGACGAAATTGTTAAATTCGCCCATGTCCCCGATATTGACGTTCTTGGGCAGTGCGTTAGCGCAG
ACTTTCTGCCAAAACTGTAAACACGAGAGACCGCAATCGGGAAAGAGCTTGTCGTGGTACTTGTATAGCAGGAAGCCCAG
ACAGCCCTTGACCGGGTTTTTCTTGGGCACGTGCTCACGGCGAAAAAACACGTGAGCCGTTGTTCCTTTACTGGCCTTGG
AGATGGGACGGTTTTTGACCCTAAACATGAGGAGACTGGGGACAGAGAGCCTGGAGAGCTTAATCTCCATGGATACCACG
GAGTACGTCCCGCGGGCATAGAGCTTAAAGTCAGGGAGGCGGTCATGGTCCGCGTTCTTGGATGATTTCACATTCCTAAT
AAACAGAAAACCCTTTCGGGAAGAAGTGGTTGAAAAAGCCCCGTGTATGGATTGGAAATGTTGAGTCACCCAACGGGAAA
CCAGGCTGGTGGTGAGAGGGTTGTCTACAATGTATGACGTGGCTGTAATAAGAGCCACGTTTTGGATGATAGCCAACACA
ACTCGGTAATACTCGTACGCCAGAATGGGGTTAAACGCCACGGAGTAAGGGTTCAGATCGAGGTTGAAAGCCTGCGTGGC
CCCCGTCACCTCGTCGCGACTGCTTTTCATACGAACCTCGGACACGAAGCTCATGGCTTCGTCGTCCACAAACTTATTCA
ACGCGGACACCAAGTTCACAAAATCGGCCCTACTATTCACAGTCATAGTGTCTTCCCCCGTTACGGAGTCTATGAGTCTC
ATTTTTTTACAGTAGTCTATGATGCGGTTTAGCCGGTGAGTGCGATCCACGCTGGAACTGGCAGTGCGATTCGACGTGGA
CGCGGAGCCCTCGTCGGCTTTCGCGTCCAGGCTGGCTTCTTTGGACTCCGCGTTGTAGCGGCGCCCGAACGTACCCAGCG
TGTCCACATCGTTCATAAAGCGAGACTGCATTGCCACGACGCAGGGTTCCTTCTTCGGCAGCCTCGGAATCTGAGGCAGG
CGAGTCTGCACTCGCACGAACGCCGTGTTGTAACACGTGTGACAACATGTCCCGCCGCAGGCCCCGCACAGACTAGACGA
ACACCCGATTAAGTGATTACTCAGGGTGGACGAGCCGGTCAGCCCGGTGTTGTAAACAGAAACTCGATTGAGGTACCACA
CGAAAGAGGAAAACAGTTGAGGACACGTACCGCAAATCAGGGCCAGGTGATGGGGAGCGTAACGCTCGTCCTTGGCAACG
GTGCCGGGACTCTTAATGAGCTTTCGGGCGTCGTTGTAAGTGTCCTCGCACAGCCCCGTGAGGCCGTTGGCGAAGTGAAT

GGACCTGAGCAACGCCTCCTGACTGGCCCCGGTACCGCTTGCATTTTTACTGGAGGTCACGTACACCAACCTGTTCTGAT
ATAGAATAGAATTACAACTGAAAAGCAGTGCGCTCATGTGGGTGGAGAGATGAAGTTTGAGCTCGGTGAGAGCTCGGATG
AGGTCAGCGTGGTCCGTAGCGTTCACCACCACGGGCCATTCGGAATACACCATGGAATCCGAAGTCCCGTAGGCCGAGTC
CAGAAACACCGAGGCGAAGCTAAACCCTAACTCGCAGATAACGGCGTCGCTGAGCATCAGGTGGTCTTTCTCCAGGCTGC
TCAGCTTCTGCGTCGTGTATCCGAAGTACTTCTTGTGCGGCGCCAGCTTGACGGACTGCTGGCTGTCGTGCACGAACTGC
TTGAGGGCCGCTTCGATGAGCACCTTTGTCTCCGAGAAGCGCAGAGCTTGGCACCACGAAGTGTAAACATAGTAGAACAG
GGTCTCACTTACCGCGGGCACGTAGAAACCCCTGGACTCCACGAAGGGACTGCGTTTGTCTAGGGACAGCTCGTCCGAAT
CGGCGTCCGCCTGCAGAGAACCAAAGAGAGTAGGAGGGTATAGCGGCACACGCACCGCGTCACATTCATTCAGTCGAGTA
GGCACCGCCTCCTCGAAACACGGGATGAGCTGGCCCGCGTACAGAAACTCTTTGAGGCCGTTGCCCACAACAACGTGCGC
CAGCACTTCTTGGGGGTCACAGCCGGCCGCCGTGCAGAGCGCCGCCAGGTCGGTTGGCACCCGTCCATCCTCCGGTACGT
AGCTTTCGACAGCGTACCTTTGGCGCGTCTCCTCGCACAGCCGCTGAACGTCACCGTGCGGCTCCGAGAAACTCACTATG
TCATGGGTGTTATGGAAGTACATGACAAACGGGCAAAAAGAGGTCACCTTGGTCAGCACCCCGCCGTCGTAACAGACAAC
CGGGGTGCGGACAGAGGTGCAGAAATCTTGATCCACGGTGAGCCCCATCAAGAGAGGGGCAATCACCACGGTGGAAGAAC
GATCGCATAGCGACAGGGTGCTCAAAACCTCGTTGTTCTCGGGGTTTGTCTTCGTAAAGTAGAGATAGGCGGCGGGGCCC
ACGGGAGCGAGAGCACTGAGTTCCTCGTTGCTCATAGTGGTCGTGTAACAACGACGCTAATACCAAGTGAACTTTATACC
GAGGAGGAGCTAAGACGTCACGGTTGACGCATTACGTAGCCCCCCATCGCGGAAATGTACACACTCTCGGCCGCGCACCG
CCCCCTTCCACGGTTAGGGGACCGATTTTTGACCCAAACAGGCTCCGCGCCACCCTCCGCGGGCTCGGAGAACTGTCCCG
CGGAGGCACCTTCGAAAATAGCGTGGTGCCCGATGGTGTTCCATGCGCCATCTGGTGTCCGAAAAAAGCAACGTGGTGCA
ACGTTGTGTTCATGCTGGTTTCAACGTTTGGATGATTGACAGGTCAATAAAACGGTACTTTTACTGTCAAACCGCGTGTG
TCTGGCGTCGTTTTTCATAATTAACGGTCGGTGTACGTGAAAAGTGTCACATTCCACGTCTAACAGTTTAGGAATCTATA
AAGCAGCGTGAGTATTCTCAGCCCCCGCACGTGTGTTTTAGGTTTTGCCCCCCCCTCGCGCGCCATCGCAGTGTCATTTT
CACCGTCCGCGCGAGCGAAATCGCCGTTCGACAGCCGCAAAACGCCGTGCTGCACGCGCGTTTCAAAAACACCGTGGTTC
GACGGAGCCGACCGCGCCGATCGAAACGTTGAAATCCGCGTCGAAAACGGCGgTTTGCACCACGTGACCGCGAAAACGCC
GTCTCGCGCGCGCGTTATGAATTATGATAATCAGCAGTGATACTCAAATTTGAATATGCATTAGGTTTTCGGATACAGGT
AGGGACTGCGCATACCGTAAATCTAGGTGTAGCTGGACTGATATGACGTCACCTGATATACGTATGTATATCAGCGGTAG
CGAATACTGTAGGTACCGTAAGGATATTACGGTAACGCTCATTAGAATACGATACCTTATAAGGCTTAATTAACACTCGC
TAATTAACATACGAATATAAACAAATGCTAATTAATTAGGACACATACGGAAACGTAAACATTCCTACCCTGTATACGCC
CATAGGCTTTCAGGTATCGGGGTCGGAAAAGGGGGCGTATCCGTATATGCATAGCCACACCCCGTTTTGTTTACTGGATA
TAGACACACCCCCTTGATATGACGTAACGGTATACCACGCCCCTTTTGATATGACGTAGGTAACTACGTCACAGAGGATT
ACTATGATCTCAGCCCTAGTCAATGCAGACCGTTCATTGACAGTGCATATATAAAAGCTAATGATTTAGGAACCTGATAT
TGATAATCTAGAACCCCTCACCCCCCAAACCCCCCTCTCCCCAGTGTGGGGGGGTTCCGGTTTCGTGCGCACGCGCACTG
TGCACCGCGACGCCGCGGAAGGCGGGTTCCGGACGTTGGGCAGCGCGTTGTGCGCATGCGGTGCCCTCACTTCAATGCAT
CTCTATCGGACGGTGTTTTCGGTGCCCAGCCCCCACTCCGGGCTCCCAGTATGGGTCGGTGGTCGCGGTTGGCCGGTGAC
CCCAGACCCATACCGCGTGCTCCCAGCCCCCAATTCGGGGTCAAACCCAGGGCAGCGCGCCCCGGGCGCCGGGACCGCCC
CTCCCCCACGCCTCCGCGAGGACGGCTCCATAGGGTTCCATTGGTTTTACCCTCCCCCATTTTTTTACCCAGCGGCCACC
GCTCTCCGGTTTGCAATCCGAAAGGGGCGTGGTTCAGGAAATTGGGGGCGGGATAGAACTTTCTAGAACTTCCGATTTCG
TCTTTAAAAGCCGGCGCACCCGGAGCTCGGTGGCGCAGTCCGCGGCGCCGTCCGGGCAATCGCCAGTGGCGACCTGGGCA
TAGCCGCGGCTTGCACACTGCGCCCACCCCCCGCCCACTGGGGCCTCCACCTCGGTGGGGGGCGGAGCCGGTACAAAAG
GTAGACCTCGCCCCCTCCCCCCACCCCTCCCCCTCCCTCCCCTAGCCCGGGGGGATGGGTTGGCGCCCCGGCGCGAGCCC
TTCCCCCAGCGGGTCCCCTCCCCCCTGGTCGGGGGGGACGGGGTGGCGTCCAGGCGCGACCCCCTCCCCCCTCSGGCACCT
CCCCCCTCCTCCAGCTCCCGGTTTCGCTTTCGATTCCGGGAGCATGGGCCCAGCGAGGCCTGAGCCGCCCCCCTGGATCG
TTCCAGGTGGGCGTGGCCCTGCGCAgCCGCCTAGGGAGCtGGTCGCGGAGCCCCGCCCGCGGGAAGCGGCCTGCGAGCCC
GGCCGGKCGCCACGGAGGCCTGGCAGCGACCAGGGGACCGAGGCCCTGCACCCCGGCTAGGCCGCAGGATGGCCCCGCAG
GGCGGCGGCGACGCGGCTCCGGAACCGCGTTCCGGCTCGGCTCCCAGGCCCGCGCAGGCCGCGCTCCGCCTGGCCCTACG
ATCCAGCGACCCGGCCACGGCCGTGGTGGTGGCGGCGGCGACGGCATCGCGGCCCAGCCTCCCCCGGGCCTCGCTCCCTG
GCTCCGGAGCGGCGGTCGCATCCAGCCCCGATGCGACCCCGCTTTCCGGGGCGACAACCGGGGATCACCCAGCGGGCCCG
CGACGACCCGGGTCCGGAGCTCCTCGCCCGGGCTTCGGGGGCCGGCGGCTGCCACAGCAGCAGCGGCCACGTCGTTTTCG
AGCAGGACCCGGCGGGACGCAAGCCGGCACCGCTACTACCACCACCACAACCCCTACAACCAGCACCAGCACTGCTACAA
CCACCGTCGCCGTCACCACAACCAGCGTCGTCACGTTGGATTTCGAGCCCTCCGCACGGCGTTCGTACAGCCGGCGGGTC
GGGGCCGACCTGGACGCCGGACGCGGCTCCAGGCTCGACGTGGGTTTAGATTTTAAATCGCGTTTCGGGTGTCATATAGC
GGGGTGCTGTGGCCACTACCACGGTGGTTTGGGACCGGGGCAGGTCGCGCCGGTCGGACTTTTAGGAACTGCTTGAGAAA
TGTGCGCTCTAAAAACAACAAAAAAAACCGTCAGAGATCTAGCTTTTACCAGAAGACACGTACGTACAGAACAAAAACAC
TCATCCACTCCTCCGTTCGGCGATTACAACGTCGGACTTGTGTGTCCCCCCATATGGCCGGGTTTTCAAGCTGGCTCGTT
AAGCAAAAAAAGCAAAGTAAAAAAGCGAAACGGACTTGTAGCTTTGTAACTGCGTGAGAAAAAAGATGATACGAAAAAAC
TGGAACTGTTGAATAACGATTTATTACATATATTTGTAGCAATGTGTCTGCACCAGTAAAATGTCTACGAACCCTTTTTT
AAGCCCTTTCATAGGAAGGCCGGCTTCTCGGGACGCCAGGAGCCAACGTTCATTCGCTCATCTAGAGAAAAATGACGACT
CAATGGACGCGCGGGAGTCCCCCCGCGCGTCATCGTGACAGATATTTGTAAATTCACTTGTTCTTGTGTATTCTGTTTAG
ATATCACTGCCACAGCCGTCGACAACCGGTCGCGCGGTGCGCTCTCTCTGCGCGGGTCAGCCGTCGGGATCGCGCGGTGG
CCGCTACTGTGTGTTTTTCTCGCGGCCTCCGCCGTCCCGGCGTCTGGCCTCCCTTGTCAAGATTAGCAAAAAGAGAAAAA
CGACTATAGCTCTCTACCGCAATGATGCAATCCCCACTAGCATGTGCTTTAACTGTGTGAACAAATCGTCTGTCTTAAAA
AAAAACATATCGGGCCAAAGCCAGCATTTTTAGCCACGTCAAAAAATAAGTAAGACATGATAGTATTTTGCTTATAGATC
CGCATCGTCATGGCCACTGATCATGATAATATAGAAATCTATCAATTGTACAACCACCAGTCGTTTCGCCATCGTGTTGC
CCTTTTGCGCGATGGTGACGCTGCATCTTATATGACATGTTGAAATATAAATGTGTTGTTGCCTGTGTTGTTTTCCGCCTT
TTCCCTGCAACGTTTTTTATTATGTGCCTCATCTGTCATTCGAGATTTCGAAACACTTGAGTCATGCTTGTACTCATTGT
CCTGAATATTACTACTGATACTACTCAGACATCTTATAGTCAAGTGTAAGTTCATCTTCATCCTCGTGTTCTCGGCCCAT
TATCTGGTAGAAGTGTGCTGTGAAAGAAACTCACTCGCCTTCCCTGCCTCATTTTCCCCCCCATGTGCGCAGTCACCTAT
CAAACCTGTCATTCACACAACACTCAGTTTGTAACTTAATCCAGAATAGCAGTAACAAGACAAAATAAAACCAAGCTCGT
TCACATTCCACAACAGCGTGCATCTTTATTAGCGTACAATCCGCTCCCACCACCGCCTGTCAATCATCTCCGCCTGTTGT
CCCTCAAAAGCCCTCATCGTCCTCCTCCTCGTCGTCATCGTCGCTGTAGCACATGGCCTCGCAGTCGGGCGTCTGGCAGC
GCAGCGAGATGGAGTCTGTAAAGAGTTCAAATGTGGTCTGTTGTGGCTGCTGCTGCGTGCCCAGGGCCGAGTCGGGCAGG

GTCTCKAKGGCAGAGATGGCAGCCATCCAGTCCGGCTCCGCGGCCGGCTGCGCTGTTTGTGAGCAAATCACCGTACCGGT
CTCCGGAGTGCCGCTCAGCGGCGGACTCTGGAGTTGCGGCCACGTCAATTGTGGCTGAGGCAGCTCGTGCTGTTCCACAA
ACCCCGTCAAGTCACTCATAAACATGTCCGTGGTGGTAGCAGCAGCAGCAGCAGCGCTGTCGTCTGCGGTCCCGGAC
ATCTGAGGCTCGGCCACGGCTGGGGGCTGAACAGTCACGGGGGCATGACACACTTCACTGTTCTGAGTCTCAGAAACCGT
TCGGTTCCGATTCAACGGAGACACCCCTGACGTGTATTGGGAGGGAATGACAGCCGGTTCTGTGTATGACGGCTCTGCTG
GCACGGGTACAGCGGTGGCCGCCGGCGTCATGGGGACCGCTGGTGAGGGCAACGGCGTGATGGTCTGGAGTGAGGTATCA
TGAGCACTAGTGCTTTCAAAAGCGGCAACGTCCAGAAAGCACTTGGTGTCTGAACGGGCAGTACGCCAGCATAAACCGGT
AGGCGCTGAGGCCTTTTTGGATGCGCATCTCGCAAATCTCGTTCTGACACTGGTGCTGGTGCAGCATCTGGAGAACGAGC
TCCGAAACCATGCCGCAGCGGTAAGATTCCATAGCGCCGTTGTCGTCATACTGGTCCAGAAAGTGCGGCACGTTGGGTAT
GTGCCTGCTGCGACGCCGCAGCGCCGCCACGCATCGTGCAAAGGAACACACGTAGGCCACGGCACGGTGAATGGGGTTTG
TGAACGACTCCGGCGGCTGCTTCTGCAGGTCCAGCGTGTCGTCGCGCTGCCAATCATCCGCCATGCCGACGGGCTTTTTA
ATCAGCGTTTTCAGGATGTGGCAGTTGTGCGACCACTCGGGGCTGTTGTCCATGATGCAGGGCAGGTCCAAGTTGTTGAG
TTTGTTGAACAACAACTGCCGCAGGCCCGACACGGTGGCCATGTAGGGGTCGTACGGGTTCACGGGCAGGCTGTGCACGA
AGTGGTACTTGATGTAACTCAGCGTTTCCTCCGTCATGGCCAGAAGCATGTGCAGGCTAGGGGAGTTGAACACGCAGAAC
GTCCTTTCCACCACCAAACGTCTCAGCAGAGACGGTGGCAGCCAGTCAAACTGCTGACGGATGTTCAGCAGGTAGTCGGT
GCTCATGAGTTCCTCGTGCGAGATCAGCAAGCGGCCCGGCTGCTTCGTGTCCTGTGGAAACGCCGTGGGTTTTTTAGGT
CATCGGAGTAGGGAGTGAGGTGCAGGGCTTCGGCCGTGTACCTCTGCAGGTCGTACACCGGGGAGGTGGAACTGGGTAGG
GATGCGTCGTCCCTCACCTcgccactcttgccgactttctggcgttcctctttTTTCTTTTCAATCATGCGGCTCAGGTG
CTGCAGCTTGCACGTCGTCCATGGTATCAAGTTCATGATTGATGATCGCCAGCATCTGCCGGTTGCTGCTGGTGGTGGCC
GGCGAAACGGTGACACGTTCGCCGTGGCTGTGATTGACAGCAGCAGCGGTGGCGGCGGCGGCGCTGTGCCGATGCGACGA
CGGCTGGCGGACGGCGCTTGCGTGCCTGAACACGGTGGTATAGTGGGGTCTCKCAGGCGCTGACCGACGTCCAGAATCAG
GACCTGGTGGCGGGGACTCGGGCGAGTCGCGCACTGGAGGCGCAGCCGGGCTCGGGCGCCGCGTCGTGGTCTGCGACGAC
GGCTGTTGTTGCTCGTACCTGGGTCGCTTGAACGAGCGCGGGACATCATCCTCGCCGTCCAGGCAGAAACGCCGTGCCCG
CCGAGCACGGCGTTCTCGGTCGCTCAGCTGCGACAGCGCCGGCCCGCTACGGTGTCTCTGGTGCAGGTCCATGCTGTACC
GGGGGTGATGATTGGCGTGTTCTGAATTCTGCGTGCTCGGCACGCTGTATATATGCGAATGGGCCGAGCGCATGTGACGT
GTGTTGTCGTTGCTTGCGGTCAGTGATCGCTGTTGGGAGGCCGGATTTTCACGGAAATGTGCACGTTCTTCGCGTTCGAG
TGACATTTCTTAGTGAAACAGCGGCTCCACAGGCTCGCCCACAACGCGTAGCTATGTTCGGTGCGCAAATCTATGAACAC
TTGCACGGTCTCCTGAGGGTTGCGGTGGGTGTAGTTAAGGCAGCGGAAGTCTCTGGTGCGCGAGCCGTCCAGCCGCTTCA
CGGAGACGCAGTGGGGGGCGTGGGGCTGGAAAGTCAACGTGTGCGGCGAGGTGAACTGCTCGCTCACCTTGGAGTCGTAG
TGCTGCGTCAGATGTTCCAGCAGCAGCGGCCACACGCGTGTGACCACCAGCCTCTGGAGGTCGTTGGTGTCATCTGAGTC
GTTGTGACTTGGAGGGCGGCCAACCGCGCCGTGTAAATGATAGCTAATGACGTCAGACAGGCTACGCTTCTCGTATCGCG
TCACGTTCACGGACCGCGTTTCCATAAAGTTAATGTCAGGAGAGGGGCACAGCTTATCGCGGACACTGAGAACCACGCGA
AGACCGCTCGATTGCGGTGGGGTCGCGGGGGGCGAGGAACCTGGCGGGTAGCGCGGGGGGTGATGGTGCAGCAGGTTGCG
GACGTCCAGCTGGGCGCGGACGAAGTTCAAGGGATGGTCTCGGAGTCCCTCGGGGATAATGAAGATGGGAAGCAAGCGGC
GGTGCATGAAGTAGCCGTCTTCTTGGTCCATCTTGTACATGAACGGCAGGCGGACCGAGCGGCCATGGTGATAAATTCCC
GTGTCGAGGCTGTTTTCGGGATGCGAGATGGGATCCAGGTGCTGGTGCAGCGGGATATCCAGGCAGATGGCGTGGTTCAA
CACCTGTGCGATGGCCCGCAGAACCGAGGGGTTAATAACCAGCGTTCGCTCGGGGAAGGGAGTGACGATGCGCAGGCCTA
ACTTGTCTGTACATTTACAATAATCCGTTCGGGCTTCGTAGGACGGCGGCGGGTTGTCGTCGTAGAAGTCCGGAGAG
GCGGTTTTGCAGGCGCTTTTAAAGAAGTAAACTGGGTGGCGCTCTGGATCCGTGTCGCCAAACAGGCTCTGCCAGACCGT
CACCCACACGCGACGCAGTCCGCGACAGATCTCGAAAACAACTTCTTTGGACAGCCCGCGGACGACGCCGTCTCTCAGGC
GCAGATCGAAATCTGCCACGAAGTTGTGTGTGGGCAGCCGGTCGTTAAAGATCTCGTGGCGCGCGTAGTAGAACTGGGTT
TCGGGATTGGAGCTGGCCACATCGTCATCGTGGAGCCAGACCATGTCAGTCAGGGCTTCGTCCGTAAATTGGCTGTCGGG
GACGTGTTTGAGAAGGTCGCTGGGAAACATGCTCTCGTGCCAGTTTTCCGACGTCACCGCGCAAAATACATGATGGTCGC
TGGGAAGCTGGACGCGGAACAGAGGGATGGGGTACTCCGACTGCGCCTCCCTGGGAAGCAAAAACTGGCGCTGCGAGTAG
CGGGTTAGGGTGTCGGTGGCGTACGTGTTGGCGGTCTTCATGGACGCGCACAGGCGCAGCAGTCCGGAGAGAGACCTCTC
CAGAGGGGGGAACACGGAATCGGACATGCTATTGATGTGATCCACCTTGCGCATCGCCTGCGATGAGGTGCCAAAGAAAC
CCGTGAGGGCGGGGCCGTCGGTCCGGTAGCTGTATCCCGAGAGGCGCAGGTGTTGGCTGGGCTCGATGCACATCAGGCGG
CTCTCGATGTACTCCTGCAGGTAGTTGGGGAGCGAAAAGTACTTGTTCATGACGTTCAGCAGGTCGTCCTCTAGGTGGCG
ACCCAGGAACAGGGGCTGCTTCTTGCATTGCTCCTCGCTGGCAGTCTTGTCGTCGTAGGCCACCACTTCGCGGTACTTGA
GGAGGCGGCTTCTAGACAGGGCCGTGCGGTAGGCCAGGTAGATGTAGTGCACGCACACGGTGTCTGGCAAGCGCGCGTGC
TCGCGGAACGTGTTGATTTGCGTGTCGACATGGACCAGTTCGTCGTAGTCTCGCTCGTTGCGCGCCATGGCGTAGCGCAC
GAATTGGGGGATGCGGGAACGGAAGGGGGAGCCAAGGAGCAGTCGGGTGAACTCGCTCATGGTGGCGTGAGTGGGGATAA
TGGCCCCTAGGTCGTGCGAAAAACTGCGTACGTATTCTTCGACGGTGGAGATAGTACTGTACTGTCGCTCGAGCAGATAG
TAGGACATGGTGAGCAACACCTGACCTTCGGTGTGCCCAAAAACGCTGATGAACCATGACGGAGAGGTAGGGCACAGAAA
GATGAGGTTCAGGTAGCGCACGATAGCGCTGTGGTGGAAGTACACCAGGTGTTTGAACTCGCGCACCTCGCCGCCGTGCT
CGGGAGAGATGACGGTGCGGAACAGATGGTTGTAGATGGGTCGCGTTTCGTCTTCGTTGAGGCTGTCCACCAGGGCCGAG
AGCGGGATGGGCTGCCGGGCGTGCAAGTAGTGCGCGAGATTCAAAACGTCGTTGTTGACGACGAACACCGGAGATACGCG
CTGCGAATCGGTGCATTTTTGGGTCTGCAAACAGAAGTAGACGAGGTTGGAGGCGTGGTGTTTAACCAGTAAGGGAAACA
CACAGTGGTCTGCCGCCGACTTGGATAGCACGTTGGCGACTATATGAGCAGAATCATACTCTGTCGCGAACAGAACTAGC
GTCATTGCCGGTGGAGTATGGAGCTGGCACGGCGCCTGTGCGGTTTCCTGGGCTGCAGCGGCGGATCTCGGAGAGACG
GACTACGTGCTGCTGCAGCCCTCGGAGGATATGAACCTCCAGGAGATGCAGGGCTTCTTTGAGGAGAACTTCAGCACACT
GGGGATTACTCCCGAGGACTTGAAGACTTTTGCCAAGGATCAGGAAGTGGTGAAGCATCTGCTCAAGCTCATACCTCTTT
ACCGCCAGTGCCAAATGAAATGTCGGCTGCTGCAGGACTACTTGGCGGATAATTGCCAGCCGCACACGCGGCCCGCCGCC
GAGGTGGAGAACCAAAAGTCTCAGCGGATTCTGCACGCTCTGGACGTAATGATGTTGAAACTGGTGGTGGGCGAGTTCTC
CATGTCGGAGGACGACACGTTAGAGGTGTTGCTCAATAAATTCTCCACGGACCAATTGGCCCTCTGTGAAGTTCAGAAGG
TGATGGGCCTGGTGGATATGGACTGCGATCAGAGCGCCTCCATTTTAGATGCGGCTGCAGCTGCCAGTGCGGACAAGGTG
GTGGTGGAGGAGGACATTTTGAACGATGAGGTGCTGGCCATGATTCCTCATGTGCCAGATCACGAGGTGTCGTCTGATGC
CGAACAGGTGGTCACGACGCCGCGTCCTCCTCCTCCCGCCAAACCCAAGCGAAAACGGTCGTCGGCGCATCCCTCTGTTT
AAAAGATACCCGTTCCCGCATGCTATAATAAAAACATATACAAAAAACAAACAACGCATTCGGTTTCTTTGGTTCGTTTT
ATTCGTGCGGCATGAGTTGTTGTAAGATGTGGTGTTCGGGCGTATGGGTAAAACCGGTGTCTGGGATAGGGGGTACCAAC

AGGTGGCCTATGGGGTGGTTCTCTAGTATGAGATAGGCAAAGACGCGGGTGGTAAAGAACTACCTTGGCTATCGTGGCAC
TGCTGGACACGTGCACGGCAGTGTTGGAGGTGTTGACTAGCGTGACTGAGGCTACGGTTCCTGGCAGCCAAATGGTGGGG
CGGGTGATGAGTCCCAGCTGGGCCATCTGGCGGGTTCCTACGATGAGAGCGCAGGTGGTGGGAGCGTGGATGTAGGCAGC
GTCCAGATACACTGTGTGTGAGTGTGAGGGTTGAATCGTAGTCCGTTGAGGTAGGCGCAGAAAAAAAAAGTTGTTTTCGT
TAGGGGGCAGCATGTTCATCATCTGCCAGGGTTCCGGACAGAAAGAGGGAAAGATGCAGATGTCGCCCTCGATGCTTCCT
GGCGTGATGGCCTGGAACGTGTAGCTGAGATTGATGACCTCCATGCTGAGGTTGCGGAGACCGGGTTCGATGAGATCCGG
CATGCAGACGAATTGTGCGTCCAGACATTTGTAAAACGTGATTCCGAAGTATCCCAGCGGGATGTAGAGTCCTATTCCTA
GGGGTGTACAAGTGTGCGGTTGTTCGGAGAGCCAGATGATGTGCTTGTTGAAAAACGTGAGCTGCGTGTTCTGTTGTCGG
ACGATGAAGCTGGAAGGGCTCCATTTGTACGGAATACTTCGTCTGGGTGAGAAGATATCCACCGGCGGATGTCGGAAGGC
GTTTTGGAAATTGAGGTTGCTGGCGAAGAGCAGGCAGTATCGCAGGTTGAGGCTTTGGTTTTCCTGCGGTTCTTGACGTT
CGCCATTCTTTATTAAAGACGCCTTTTGCCACAAGCGTTGGAAGCGGGATTTGTGTGGCTGCGAGACGATCAGCTGTTGA
TCTAGCTTGTCCGTGAGCACGCTCAGCATTGTGGCTCGGTGCCGAGTAGCGGAGCAGTCATGAGGAACAAGACAGCAATG
AGACTGTTTGTGTGGACATGGATGGCAGTTTTATGTGTGGCGAACTCCTCCACATCGAGTCCAACGAGCACCAGCACAAC
CGCTACGACAACTTTGAAACCAATGAGGAAGCATGATAACAATGACTTTTTTAGTGCGCAGTGCACTTCTCATCTGTATG
AAGTGTCTATGAGATCGTTTGCGGCTTTCGGGTTTGGAGCTTTGGTCTGAACATATTGCTGTTGGTTATTTCTTTTGCTG
TGGTGTTGCGACATTGCTGTTTCCAAAATTTTACTACCACTACGGTGGCAGGATACTGAGAGACAAGTTGGTATGTCTTT
GATATGTTTGTAAATTACTGTTGTTTAGTCGTATTTGTTCTTTAAATGCTAAACTGTCTATGTAGTCCCATAAAGGGGTG
ATGAAATAGTTATTGGTAATTGTTGTATTGTAAGTTGGTGTTKGGTTATATAAATTGGTAARGTTTATGTCACCATGAAT
AGTTTCTACAGTGTGATTGTATAATATTTGTCGGGTATGATTCGCCCAAGCTGATTGACGGTTTCTATCTCTTTTTTCGT
CGCAAAATGGTTTTTGTGCATATCGTAAAGTTGTATATATAAAAGTAACCATGTTTTGTAACTGTGTAATAATGACGGGC
GGCTTAAAgTTAAGTGTGTCTGATGTATTTACTAATATTGTTGTGTTTGGTTCGGTTTTGTTGTATGTTTCCGTTGTTAT
TGTGGGTGTCATGGTGGTGGTTGAGTTGTTTGACCGTTTATTTCTTTTTTTAGTTTTTTGTGCGGATTCTCTTTCCGGTT
TTTGTTTTCGTTTAAGACGTCGCATACCTTGTTTTAATTTCCATAAACTAATTTTGGGTGCTCTAAACAGGTTGCGACTC
ATGGCGTTGGTTAAATAAAAGCTTTTGGTACAGTTGTGCCGTAAAAGATTTCTAAATATAGCTTCGCTAAGTCCTAAAAA
ATATATAGTTTGGCTGTCCATTAGTGTGTAATCCTCACCTAGATATATTTTTGTTACCCATCTGGGAACGTTATATAGCA
TGGGATTGAAAGTGGTTGTGAGGTTACAGGTTTCTTCTCCTGTATCGTTGTGTAGTGATATGTTAATCATTTCTGACAAG
CATTTCATTGATGGCACAGAGCCACATGGGGGTGGTATTATGGTTAAGGATGTGTTTGTATGGTTATATTTACACATGAC
ATAGTTGCTAGGTTTACGGTATGATGGGCCGTAAAAATCAAACCAGAAATACGTAATATTATGTTTATTAGTTGGACCGG
CTAATATTACGTGGTTATCTGGCATGGTTGGCCATTGAAATGTGTATGGATAGTTTCTGAAACTTAGATGTTTAAGTGCA
TTAACAACTTCTACACGTGGTGTACTGGGTTTTTTTGGTTTTGCGGCTTTCGGCTTTGCAGCTTTAGGTTTGGGAGCAGT
TTTAGATGTAGATGTTTTACTCGTTGCAGGTTTTTTGGTTGTTGAAGTATTATTTGTGTCCGCAAGTTCACTAATAATAA
TTATAATTAACAAAGCATACCATTTTATCGTTACAGAACCCATGACGAGTTTGACAACCAGTTTACTGGTTTGTTTTGTC
TACGCTTTGATGTTGACTACGCTACTGCTGGTTAGCTATCGTTGTATGATCGGCTTTCAAGATGATATTGTGACTCGATC
ATTAATGGTTTTTAAGGCGTGTCACCAAGGCCTGTATAATACGTCTTGCTAAAATAAATACTCAAACACACAAATCAGGT
GTCAGTGTTTTTATTTTTACAACAGTCTGATTATACGATACAAAAGATATAGTCCAAATATTGCAGCCAAAGCATAAATA
ATAATCAGTGCGAGGTGAGTGTCCGTCATGTCAACCACAGCTTCTGTGACTTCAACAACTGTACCGTTTTTTAGGAGCAT
GAGGTAATGAGTTCGAGGAGATTTTTTAAAGATCTGGTTGCTCGCATCTAGAGCGAAAAGCAGGTCTTGTCTGTCATGAA
TGTAGAGAACGTTAACAATACCCTGAGAGTCATCGTATTCCACCATACCGCTTTGACAGAAATCACATTGTTCTAGGGTG
ATGTTTCGAACCATAACAATGGGGAAAGTGTCATGTTTGCTAGAGCTGGGCGAACAGGTACTGTTGACGAGAACTTGGGT
AACAATAATGGCGCGGCCGACAACAGTGGTCAAAACTGGAAAAGATGTTCCTTTTACAACATATTCGCTAGAGATTGTGT
AGGTAACTTTGGGCGAAACGGTAAGGACGGCAAACGATTTGTCAGTAGATATACATTTAACTTCTGGGAACGTCATGAAC
GTTTGTGGTCGGAGTGTTTGCAGAACGGATAGAGCTGTTTTCACTGTTCGGGAGTTGGTGGCTTTTGGAAACATGTTAGA
CAAGTGTTCTAGTGCGTGGTCTCGGCGACCACTGCCAGCACACGGACTGTAGATGTCGCTGACGTGTACGTGTTCGTCGC
TTCCAATGAGTTGCGACCAATGTGATAGTTCTGCTAGGGAGCACAGTCCTGTTTCTAATACAAATATTTCTCTTCTTTCA
GTGGAATAATTTAACATGGAGTGAATAATACTGCTTGCTAAGTACAGTTCTTGTCGAGCGAATCGAGAGAGGAAGGACGC
CAGATGCGTTTTGTGTATGAGCAGGATGAGGTCTCGAAGCTGTTTTATTGCCGCGTGTGTAATCaAATCGTAGCGTTGTT
GTTTGGCCAGTATGTATATTAATCTTGATATTGTAAGGGCACTGGAAACTTTCCCATTGGTGTTTAGAGTTGTATGGGCG
ATTTCCGCTGCGGTGGGATACAAAAGCAAAGTTGTTCGAGGTTGCATTTTGGATATGCAGCTGGCTAATAAATCTTGAGC
TAGGAGGATGGATGCCTGTTGATTTATTGCGCGTGAAACGGAGGCTGTAGTGACTTTAGACATGGAGGCTGTGTGAACAA
TTAGTCCGTACGTGAAAGCAGTTTCTACTGTTGCAGTTGACATACTGTGACATTGTCCGTGTTGAAGAACTTTAACGGCG
TATTTGTCCCAAACCGATAGTGCTGTGTCCAAACTCTGATAATTTGCAGAAAGTATTGATTCTAGGAAGTCCTCTTGATT
GATGTAGCTATGAAGGTTTAGATGTTCTTGTTTTACCAGAACCATGAGTGTGTGTTTTGTGGTTTGTCTCAAGATAAAGT
TTTCCCGTCGGTATGGAGCTTTAAACGCGACTCTGGTGAGGTTTCCAAATAGCAGAATTAAAGGTGTTTGTCTGACAACT
ACTGTTAGGGCAAAAAAGTGTTCCGTGACGGTAAGTTGCATGTAGTCATGTTCTGTGACGTAGAAATCTTGATAAAGGCA
GGTTTCTCTGAAGGTGAAAAGCAGGTCACTGTCTTTAAAGAGTTCACATGTACCATTAAAAGTAGGTCTGTGAAGTCCGG
ATGCGTTAAGAGTTGGTGATGTACTGTACGATATGATAGTTTGCAGATTCAAGTTTGTAGGGGGTGGTACAGTAGATGGT
TGATCTTGAAGGTGGTCTTGATCAGTTAAGTTTTGGGCATAAAATTTGTATCGGAGTTTGCCTTGAGAGACGGTGGCGTA
CGTATTAAGTTTGCTGAGATATTGTTCCAGAGTTTGTGAGAGATCCACTTGATTTAGGAAACTTTCGGCCAGGTCGCCAG
TAAAAAGGCATCTAGGTAATTGAAAGGTCATGTATTTGCTTGCCGCTGTGTAAAAATTAAAGCTAATAGCGCTTTCCTTC
ACCGTGCTGATGTTTCGTTCGGTGCCGTTGTTGTCGCATTTTGTGGAGTTCTGTCGTAAGAATCTGATGGGTCTGTAGTT
TAATTGCAGAGTTAACTGTGGGTAAGGTTGTGGTAAGTTATTTGCTTGTACCACGAGCCATGTCACCGCGCAGCCCAACA
ACAGGCTTCGAAGCATGCTGGGTCTCGGCAGCAGATTCACAGATAAGACCGGGACGGGTTCCAACTTGCGCTCTGGCCTA
TTTAATATGACGCAATGCGCGGTCACGTCGCGTTGGGTCGCTATTGTCCTGCAATGAAACGCGCGTCATGTCGGTGAGAA
GCGGAGTTATCACCTCTCTGGATGCGTTGCCCAGCTTTTGCAAGCGAAACGGACGAAAAAAGCACCTAGACATTTATCGG
CGCATGCTACGGGCCTTTCCCTCGTTTGCGGCGTTCAATCGTCTCCTGGGGGGGCTCTTTCCAGCGTACTGTCAGAACTA
TCGCCGCTGGCTCTTTTTCGAGGTGCGCCTGACTCAGCGGATTCCAGACTGCGTGCTGCTGTTTGTGTCTCCCGACACCC
CGCGGCGCGCTCTCTGCTACGTAATAGAGTTTAAGACCACTTGCTCGGACGCCACCGGGCAGTCCGTGCGAGAACACGCA
ACACACAACTTGCAGTACGTACAGGGCCTGAAGCAGCTCAAGGGCGCGCTGACGGACTTTGACGCGCTCAAGGTGCCTCG
GGGTGACTCATGGGCTATCATCCCTACCATTATCTTTTTTCAGCAGCAGGCTGCGCAACCCTCGTTCGCTCGCGCTTTCC
GCTCGGCGCCCTTTACCCTTCGCACTGACTTGGTTATTGACTACCTGAGACGCCGTCAGGATGAGTCTGTTGCAGCCTTA

CTTTCGGCTACCTATCGTCGTGGCyTtCGAGCCTCACGCAGACAACATTCTAAAGTGTCCTCCTCACGTGCTGCAAAGGC
TGGTGGACGACTCCTTCGCGGGGCTCAACAAAGAGGAAGTGGTCGCCAATCGAGTGCGAAAGCGGTACCTGCGCGAGGAG
CTCAGCGCGCTCAACGAAAAGGTGCAAACTTACTGCGAAGATCTGGAAACGCGCGTGTCCGAAGCAGAGGCGCTGCTCAA
ACAACAGTGCGAGGTCGAGCCGCCGCCGTGCCAGGAGCCTCCCGGACAGGCAGGAAGGGAGTCTTATAGTGCTGGCAACG
GCTCTGAGGCGGCGCCTCAGGTCACGCGTGACCGGGGCCAGACCGCAACCTGGGTTGCGCAGTGTTCCGACCAGGGCAAG
GATGTTGTATTTTTTGGCATCACAAAGAATGACCCCTTTATTCGTTTCCACACTGACTTCAGGGGTGAGCTGGTCAACAC
CATGTTCGAGAATGCGTCCACCTGGACCTTCACCTTCGGGGTCTGGTATTACCGTCTCAAACGCAGCCTTTATACACAAC
CCCGGTGGAAAAAAACCTTTCGGTTGGCGCAGATGGAGAATTTCTCCATTTCCCAGGAGTTGCTCATAGGGGCCGTGAAC
GCTTTAGAAAACGTAACAGTGTACCCTGCCTACGACTGCGTGGTCTCCGATTTGGAGGCCGCGATGTGCCTTTTAGCTGC
CTATGGCCAGAACCACTGGGACGGACGCGACCTCCCGAGCTCCATCCAGGCGGTCCTGACCGAGCTGCCTCATTTGTTAC
ATAAACTGTCAGACGAGGTGAGTCGCGAAATCGCGACGTGGGAGGGCGCACCCACCGTGAATTACTATGCCTATCGTGAC
CCCGCGGATCTCAAGTACTACATGCCCTTGAGTAGCGGTCGCCATTACTCTCCGGGAACGTTCGATCAGAATGTGTTGGT
GCGCGTGCTTCACCGGAGACAGGTCATCCAACACTTGCCTGGATATGAGGCGCAAACTGCAACAATGGTGCAAGAGCGCC
TGTCGGGGCAGATGTCGGATGACAGCCTCACGCTGTGGTGTAGGCGGCTACTGGCGAGCAAGGTGGGCCGTGATGTACCT
ATTTTTGTACACGAACAACAGTACCTTCGTTCTGGGATAACCTGTGTGGCCACCCTGCTGCTCATCTGGAAGGTGGTCAA
CTCGGAGAGCGTTTTCGCGCCCCGCTCCGGCAAGTTTACCCTCGCCGACATCCTGGGGCACGATGCGTTCCGCGATCCT
CTCAGGACGATGGCGGCTACGGTTATGGGAATCGGGTGAGGAACTTTGAGTTTCTTTTGGAACATTACATCATTCCCTGG
TATACTCGCGACTCCACCGTTACCATTTCTCAACTTTTCCCCGGGATGGCGCTGTTGGCCATCACTGAGAGCGTCCGAAG
CGGTTGGGATCCGTCGCGGCGAGGAGACAGCCAGGCGGGAGACAGCGGCGGTACCGTGCTGATGCAGATCAGCAAGGTGA
ATCCCGTGGCGGACTTCATGTTCGCGCAGAGTTCTAAGCAGTACGGCGACCTGAAGCGTCTGGAACTGCACGATGCGCTG
CTGTTCCATTATGAGCACGGCCTCGGACGACTTCTTTCGGTGGCCTTGCCCCGTCTTCGCGTTTTCGCCCTGGGGACGGC
GTTATTCAACGTCAATGATATTTACGAGTGCATTTACTTTCTGGTATTAGGTTTTCTTCCGGTAGTGGCCGTGATGTAAG
CATTTTGCATGAGTGAACAGCATAAAGGACGCTCTATTTGGCTCCCAAGAGCCGGGGACAACGCTCTTTGCTCTTCGGAG
CTGCTCGTTTGTGGACGGCGCGGCGGATTCTCTCGGCCTCGAGATGATCGCCTTCACGGGTACCATGAGTCTCAACTTGC
TCAACATGATGATAGGGGCATCAAGTTTTATGGCGTTCGCCTTCGTGCTCATGTGTATGCTGATATTGAGGAGAACGCGA
TTGCCGACGAGCGTCACCATTTTTGTATGGAATCTCATGGGCGTCCAGTTCGTGGCTATCTGTGTCATGATGCTGTCTAG
GTACATCGCTGCGGACAAAACTGCGGCCGCGTCTAGCTTCTGCAGATTGACTGTTCTGTTCGAACACGGCTGCGTGTATA
TCACGTCACTCATATTCATGTTTTTAGTCCTGGACCGCCTGGCCGCTCTTCTTCACGATTGCTATTCCTGGAAACATCAG
ACGAAGTACAACAAGGAACTCAGTTACTATGCCGTGGTGTTTAGCTACATCATGGGCTTGGTGGCCGCTATTCCTACCGC
CTCGGTGGCCATTCCCAGCGGTCGCAGCGGGTTCGGCTGTCAAACGCCTGCAGGCTACACAGCCGTTGAGGTGGCCTTGA
AGGTCGTGTTTTTGTTTTTGGCTCCCGTTGTAACGGTTGTGGCTTTGATCATTCAGGTTTCCTATCATCGAAACCGTGAT
TTCATGTGGAGATACGCATCGAGAGCGCTGGTGTTTTACACGCTTTTGTTCTTTTTGCTGTTCCCCATTAACTACGTGAG
AGCCGTGCGCAGTGGCCTGTTCTCAGTCAACAACACTGCCGTGATTACTCCGTACCCAGACTACGTGGATTACATACTCT
TTTGTTGTGAAGTGCTCGCCGATTTCCGTCTGACGGTGTTTTCAATGTTCATCTTGGCTCTGTGTGACATGGATCCACTG
AGGCGAGTCGAACAGGCTATCGACCAGAGCGCTGTGGAGCTGCAGGTGCCGGAATCTCTCCTCAGGAAATGGCTGCAGTT
CATGGATAAGCTAAGTGCCATGGTGACATGGCGTTTTACCAAGCGTGACGATACCGCGCATCTCACTACCAATGAAGAAC
CCATGGCTACCAATGACTCCTCTGCGGTCGAGGTGACCGTCACGCCGGCTTGTTCGGCTGTGTTCAGTTCTGGATTGTCT
CCAGCTTAATTCGACGTGACAGGGACATGTGTGGGGAAAACGGCGTATTGTGTTACATGTCACCGCAGGCCGTGGGCATG
TGCAAGGCTCTGGCTGACGCCACGACCTCGGGGTGCAGTGTGACTCGGGGGGTGAGAGAGAGTGTACCTGCAAATGCTTT
ATAATCTTCGATAATAAACGTGAAAATACCCATCTCGTGTTTGCTTCTCTGAATTTCGAACGCGTCGCTACTGATTTTCA
TTTTCCCCGTTCCTATCAGGTAAAAGTACCACATTTTCTGGCAGATGATGCGGATGAGCGGCTCGTAGGTCGCCGTGCCC
CAGTGGCGGGTGAAGAAGGCGGCCAGGCGAAACAGCCTGTGTCCGTACAGCGTGCCCAAGGAAAACAACAAGTTGCCGTC
GCGTGCCAGATCTTCCGGGAAAGAGATGTGCAGGCCCGTGTGGCGTTGCACGAAGCGGCTTAACAGCGTCGTGCTGAGTC
GCGGTTGGCTGAGTCGCTGRCTCAGCTGGGCGGCCTGTTCCTCGTGGAAAACCACGTGGAAGTCCACCTGGGGGAACGTC
TGCTGCAGCTCGCGGTACAGATTCGGCCAGTACGAGGTCGGCGTCTTGCGGCTCAGAACGGCGTTGTCGGACACTCCCAG
GTTGTTGGTGGTCTCGCGCAACAGCAGTGTTTCCAGGCCGCGGTGAAAGAGGAGCACGCAGATGAGTCTCAGGATCTTGA
GTTCTTCCAGCCGCAGCGTGTCGAGCGGCTGTCCCCGCGACATCTTTTCGCTGATCTGTAATATTAGATGATTGGCACAA
GTAAAGGAGAATTTGCCCGTTCGGACCTGGTCTTCGTCCGTGTGGGACATGGCCGATCCCGTCTACGTCGGGGGTTTTCT
GGTGCGCTACGACGAACCGCCCGGCGAAGCGGAGCTCTTCTTGCCCTCCGCCGTGGTGGACCGGTGGCTGCGCGATTGTC
GCGGTCCGCTGCCTCTGAATGTCAATCACGACGAATCTGCCACCGTCGGCTCCGTGGCAGGGCTCCAAAACGTCAAAGCC
GGTCTGTTCTGCTTGGGTCGCGTCACGTCGCCCAAGTTTCTGGATATCGTCCAAAAAGCCTCGGAAAAGTCCGAGCTGGT
TTCGCGAGGACCGCCCTCGCAGTCCTCGCTGCAGCCCGACGGCGTGGTGGAGTTTCTGAGCGGCAGTTATTCCGGGCTGT
CTCTCTCCAGCCGCCGCGATATAAACGCGGCCGATAGCGCCGCGGGCGATTCAGAAACAGCGTGCTTCAAACATGTGGCT
CTGTGCAGCGTGGGCCGCCGGCGTGGCACGCTGGCGGTGTATGGCAGGCAGCCAGATTGGGTCATGGAACGTTTCCCGGA
TCTCACGGAGGCCGACCGGGAAGCCTTGCGGGCGCAGGTTCTGGGAAGTAGGGAAGTGGCGGAGACTAAGGCAGACGACG
ACTCGTCGGCAGCCGCCGACCCCTTTCAGTCCGACTCGTACGGCCTCTTGGGAAACAGTGTGGACGCGCTTTACATCCAG
GAGCGGCTCCCCAAGTTGCGTTACGATAAGCGCCTGGTCGGGGTCGCCGGTCGCGAGTCGTACGTGAAAGCCAGTGTTTC
GCCCGGCGAGCAGGAGCCGTGCGATATTAAAGCAGATCAAGAGAGGCCCAAGGAGCCAGAGCAGAGCCGCGTACCTGCCG
AGACCATGTCTCACCCTCTGAGCGCCGTGGCTACTCCGGCGGCTTCGTCCGTCGCGCCCTCTCAGGCCTCCCTGGCACTG
GCCCCGGACGGTGTTTATTTACCTAAAGACGCCTTCTTCTCGCTCATCGGAGCCAGTCGGCCCATGGCCGAGGCGTCGGG
AGCGCGCACAGCGTATCCGGCTGCCCATCCGCCGCCCGCTTATCCTGTCATGAGTTACGAGGACCCCTCCCGTCATTATG
ATTATGGCGCTTGGCTACGACGCCCTGGCTATGACGCCGGGCCTTCTNTGCCTCCCCACACCGCCATTCCTCTGCCATAT
CGAAGACGGGACCCGATGATGGACGAGGTGGAGCGGGCCGCCTGGGAGCGCGGCTACGCGCCTTCAGTGTACGACCACTA
CGGCGCTAGCGGCGGCAGCGGTTCGTGGTCGCGCGCTCGCGGCGGCGCGCTGAAAAGGCGCAGGGAGCGAGACGCCGCGT
CTTCCGACGATGACGAGGACATGAGTTTTCCCGGCGAAGCTGACCACGGCAAGGCTCGAAAGCGGCTGAAGGCTCATCAC
GCGCGGGACGGGGCGGCCGTGGACTCTAAAGGAGGAGATCGCTATGACGAGATTCGAGAGGCGTTACAGGAGTTGAGAAC
GCGAAATGTTGGCCGTGCGGCAGGTGGCGCCGCCGGCTCTCCTGACTCCTACCCAGTTAGCGACGCCCGTAGTGTCTTCC
CCCACCACCACGTCCCACCAGCCCGAGGTGGCCACGAGCGAACCTCCTGCCAAGACTCCCTCTGCTGCTGCCGCTGCATC
GGCCGTTTCGGCCCATAGCAGCAAGTCGGTGGACCGCGGGGTGGTGAACGCCTCTTGTCGTGTGGCTGTACCGGGGGAAG
GCATTCCTCCTCCTAAGGACATGGTGGACCTAAACCGTCGCTTGTTTGTGGCGGCGTTGAATAAAATGGAATAAAAGCAC

CTATGGTACAAACAGATCTTGAGTGGCTATAACTTTATTATGATTGATTGACAGGTACATGACACAGACAAACAGTGATA
GCTATACATGTGTGTCGTATGGTTATTAGATTCGGGGTCGGGAGGGGGATCTGTGAGACAATGGGAAGGGGGCGAGTAGG
GTTTGATAGGCCACGTCGATGTCGGCGGCGGTACAGATGTCCCCGTCCATTTCGCTTTTGAGGTAAGAGGTGTTTTTCAC
CGCGTTGTCCTCGACGCTGGCGGTGATGGGCATTAGGTTTTCTAGGTGCATGTACATGTTCCAACAGGGCAGGGTGAGCA
GCATAGATGTGGGTGAGAGTTGGTCGGAGATGTGAGGCGTAGCGGGTCTTCCGCCACGTTTTTTGCGCTCATCATCTTCG
TCTGAGGTAGATTCGTTTTCGCTTTGGAACGGATTGAAGATGGCCTCGGTGACAATCGGCGGGGTGGTGCCGGGTAATTC
GTCATCTTCGTCTTCACTGGATGATGAATGTTGCTCGGCGAACTGAGAAATGCCGCGAGCGATGGACTCTCTGGGACAGA
GGCCGGCGCGCAGGCGGATTTGGGAGCACGCGGAAGTGAGCGTTTTGTTGGGCATGGTGTACAACGTACCTCGAGGAAAG
AAGTGGATCAGGCCGATGGTTTGCATGAAGGACACGCTAATCTCTGTCTTGGTGAGGTTTTGCATCTTGATAGTAATCTT
CATGCGGTTCTGCAATGGGCCGCATTCTATGCTGAGGCCCGGGATGTCAACGGGAAAGAAGAGGAAAGAGCAGGTGTTCT
GCGTCTCGAACGCGTTGTTGAAATGCGCCACGTGGGTGTGATGCGCAGGGATGGTTACACCTTTTGTGTTCTTGACGATG
AAGCCGTTGTCCGAGTGTCTTTGGAAAAACGGTTCTGGATTGTGGCGGCACATGACCTCGGCTCGGTGGGAGTACACGGA
TAGTTGCAAAAACAGTTCTTGCGGTGGGTGGTGAGACAGATTGATGAGAAACACTCTTACCAGGTTTTCGCTTCCCACGG
TTTCGGCCTTCTGAACATGGGTGTAGCCGTCCGAGCATGCTAGCTGGTCCATGGCATCCACCACCGTTAGCGGAATAAGC
TCTGTGGAAAATATAAAATCGGTAGTGAAAAACTGACCGCCTTCGATGTTCGGGTCTCTGCGGCGGTTCCAGATAATCTT
ATCCAGGTTGAGTTTCACTTTCCACATGTATTGGCACCGGCGAATGGTGGCTCGAGCGTTGGCTTTGAGCGGCCGGTTTT
CGTTCTTTGCTTTCCCGCGGAACAGGTGAATGCCGGTGACTGGCACTCGTAAGAGGGGCAGTGCAAAAACCAGCATGGAT
TTCGTGTGCGTGCCGCTCTTCAGTCGTTTATTGGCATAGTTGCCTACGTCCGTTCGCAATTGATCCCTGAGAGTAGAGGA
CGAACGTCCTTTGATGCTTATGTCCGTGAGCTGAAGAGCGTCGCGGGTGCATGTGGACTCTTGGAAGGCACAAATCACGG
CTGGGTTGGCGACTTGTACGTGGGTGTGCCAGTTGACCACCTTCACGGCATCTGCTTCTAGGGTGGTGTTTTGGGTAGAG
ATGACTCGCATGACCTGACATGACAGGCGTTCGAAAGTGGCGGACAGACTTAAGGGTCGAGAAGTAGACGGTCTGGGTTC
TCTTTCCTCTTGAGGACGATCCATGTCGAACGGTATACGAGAGTTGGCGGAGTGCCCCGGCCGATGACGTTGTCCGCAAA
TACAGCACCCGCCGCTGGCCCGGGAAACAGGCGCGTGATCTTAGGGTGTCGGGGTTTTATAGCGGTTCGGGTTCACGGG
CAGGGCTTTTTGGTATTCGGATCGTCATCCGTCACCGATATCTCTTCCGCGGGTAGTTGACGACGGCGTCTGGGAGTGTG
GAGTGACTGCCAGGCGCCAGATAGGCCATAGAAACGGCGCCAGTCGTCCTTCGCGTCCCACTTGAACTCTTGGAAAGGCA
GCTGATCGCCGTGAACGCAGGCAACGATAGGAGTGAGATCTATGCACCTGATACCGCATTTCCAGATGGGCCAGTAAATA
ACTGGGAATTGTTTATCGGGATCGTCGTCGTCAGAATCGGTTGCCGTATGGGATTTGGTGAAACGTTTGCGTCGAATGAT
CGCTGACAGTTTAGGTCTCTGGATCGCAGCAGAGGTGGAAGCGGTGGGTTCGGGGGGAGGGAACACTTCCATGATCTCAC
ATTCTGAGTCTGAACTGGAAGACGACGAAGAATCACTAGACTTGCTGGATTCGGATGAACTGGATGAGCTGCTGGAGGAG
CTGGATGAACTGGTGTCATCATCGCGTGGATGAGCGTCGTAAAAGTTATAGTATTCCAACTTGGCTAGGACGCGGAACTG
GTCCGTGTAGATGCAATCGGCCTTTTTGGGGCGGAACAAGAGTTTACGTTCGATGAAATGTAACGATGCCAGAGGCTGGA
AGAGTTTAATGTCTGCGTTTCTACTCATGCTTCGAACTTCCATGTAAATGGGCTGCGCTGACATGAGTGGATTGCAGGAG
ATGCTGACTCCGGGAATGTTTTTGGGACAGATGAGGCCGATATATTCATCGGATTCAAAGTACATATCAATAAACACCTG
CGCGATTTTTTTGGGCCTTAGATGCAATCGTTTGGGAGACAAGATGGTGTACCCATTCCTTTCGTTTGGTCTCAGGAAGG
GTTTGGGATTGCGACTCATGGTGATGGACGCGTCGCCTTGTTCCCAGGCTAGGTGAATGAAGGCTGTGGTCTTGGGGGTT
TCCTGTTGAAGGCATTCCATATAAATCTTGACGAGCCCCTTGTCCTTGTCAACCAGTTGAACTTTCGTGACTCGCGCGTC
GCGCAAGGAGCACACCAGTTCCTTGGCGGTATTCATAAACTGTATCGGCATCGGATAGATCTTGACAAGAAACGACGTGT
AAAAAAATGTGCCGTTCTGTAGCCAGTGGCTTTCGCGCTGGGTCCATAGCAGGTCTTTCACATAAATTTGAGTGTGCCAT
CCTCCGGTGAACGCTTCCACCGACACCCGAGCTCTGGCCGGATTGTGTTTCTTCGCCGATCCGGATGGATACAGTGTCAG
TTCGGGAGGGATGACGGGTTTGAGTGGCAAGGCGTACACGTAAAAAGACAGTGGTTCCTCCGAGGGGATGATGGGTTTGC
TGGTGGGATTATACACGTGCAAGGTGAGATTCTTCACTTCTTGCTCTTCAAACACCGTGTGTTTAATCTGGAGGTTCGGT
TCATGGTAGTATTCAGCCTGTGATCTGGAGGTGAGTTGGGTGACGAAGATAATGGAGTGTTGACTGACCTGCACGTCGAT
GCCTGTCTTGACGACCTTGGTCTCTTCCGAATTTACAGGCTCATGGGCTTCGCTCAAGACGTTTTTAAAAATTTTTCCCG
AGATGGGACCGATTTCGGTTACCAATTCGGGCTCGCGCGGCGCTGCGAGAGCCATGTCGACTGGTTACCAGGGTGCTTTC
AACTGAAGTCGGTGTCGTCATTGTGGGCCAGGTACAAGGCGCGAATCTTAGCTCCGGTGCTTTTTGCTAACCGAGCCGAT
GTACGGCGCCTGACTGGGCGCCGCCGCCGCCGATTCTGAACGCGTGTGGCGACGGCGTCTCTGCGACTGGTAGGGGCGCC
ATTCGCCTTCCATGGAGTGGAAGGTTCGGTATTCGTCATCCTGGCCCCAGAAAAACTCCTTCTTTGGGAAACATTCGCTA
CTCACGGTGGCGATGATCGGAGTCAGGCTTTGCACTTTCATTCCGCATTGCCAGTTGGGCCAGCAGAGCATGAGATTTTC
TCGAGCTTCCTCGTCGTCTGAAAGATCGTCACTAGCCTTGCGTTTGCCGCCTCTGACGGTGGATGGTGTTCTTCTTTTGG
GTGGCGGCCTTCGTCTACCGCTCCTACTGGGACGGGAATATTCCGCCTCGGACTCCGAAGACGATGAGTCGCTGGACTCT
TCGTCCGCGCGTTCGCTGTAAATGTCATCGTATTCCAGTTTAGCCATGATTCGGTGCTGGTCTATGAAGCCGGGACACTC
GGCGGCCGTGGGATTTTTCACGAGGAGCGACTTGCGTTTTATAAAGTGCAGCCAGCCGATGAGTTCGAATTTTTCCAGGT
AGATGCTCTCGTGCTCGGCTCTCACTTCTAAAAACACCGGTTGTGTGTTCATGAGCGGGTTGCAGGTGAAACTAACTCCG
GGAATGGATTTCGGACAGATCAGGCCCACGTAATCGTCGGATTCAAAGTACACGTCCAACATAATGTGGGAGATTTTGCC
TGGTTTCAGGTGGAAAGTTTGAGGACACAAAATCGCGAATCCATTTCTGTCGTGTGGGACCAAGAAGGGCTTCGGGTTGC
GGCTCGGAGTGATGTCATAGCGGCCCTGTGTCCAGGACAGGTGAATAAACGCTTTGGTGTCTGGAGCCTCTTCTTGAAGG
CATTCCAAATAAACTTTGATCTCTTGTCTGCTGACTTTCTGAACGTTCGTGACGTGCGTGTCGGGCAGAGAGCACACGAg
CTCGCTGGCCATGTTCACGGCGTCCAGCGGCATGTTCTGCGGATTGAAGACGAACGATGTGGTGTGAAACACGCCGTGCG
GTCTGTAGCGGCTCTGGTTCCGCGTCCATATCAGACGAGAGACGGTGAGACGGCTGTGCCACGCGGAGTCTATCTGTTGG
ATCACTGCATCGGCCACGGGGAGGCGATGCTTGCGGTTTTCTCCGCGTTGCAGCACGAGTTGGGGAATGGAGACGGGTTC
GAGCGGAATGGCGTAAACGTACAGCGACAGCGGCCCCGTAGCGGGAGATAGAGGCCGGTCGGTGGGGTTGTGCACCTGTA
CGGTCACGTTTTCAATGCTCCGAGCCTCCAAGGCCGCATGTTTAATCTGAATGTCGGTGTCGTAGCGCTGACACGGCTGC
GATTCTGGCGTAAACTGTGTAGCGAGGATAATGGCCGGTTGACTAACTTCGACCTTAATTCCCGTGCGAAGCGTATAACT
TTGATGCGGTTGTATCGTGACTGTGCTTTTGTTAAACACGGCTTTCACCACGTTCCCCGAGATGACGGTTAGGAATTCGT
CTTCTTCTGGTGCGTGGCTCGCCATGTTGACTGACCCGATGGGTTCTGCCGACTGAGCCCGTTCCGTTCGTGTCGCTATT
TAGATACAATATAGATGGCGCTACACACACAACACGGGTATGGTGGATAACAAAACGTTTTTATTACGATGGGGATACAG
TTTGATGGTCGCAGACCATGGCGAACGTAACCGGTTTCTGACTCGCTTGCTGTTGTTGCCGTTCCGCCTCAGGATCTGGG
CGGGAGACCTGGATGGGGACAAGGTCCTGGAGGGAGAGGGAGGCACCGAGGAGGGACAGCTGGTGCGGCACGTTGTGCAC
CGCCGACTTCACCTGGTGCGTGAGGGCGGACATGCGGCCGCGGTTCAGGGTGTGTTTGGAAGAGATAAAGTACAGCTGTC
CCAGAACGGTGTTTTTGGGGATGTGCGCGTTCTCGCCCACCACGTTGATTTCCAACCAAGTCCGGGGTAGCCAGATGGTG

ACTGGCATGACGAGGTCTTGGCGGTTGTGAGGGAGAAAGAGTCCCAGAAAGGGACCGTAAAAGCGGCGGTAAATCCGAAT
CAGGTGGGACGATTTCAGGGACAGATCGTACGGCAGGCACACGTCGAGGCCGCCGTTATCGTCGTGTCGGGGAAGAAAGT
AAGGGAGAGGTGTGTCCGAGAAGAAGGGCCGCGCCACAGTGTGTCGTCGGAGGCGCAATCGGAACGACACGATGGGATTC
GACGGCAACTGGTGCTGCGCGCTGCGGAAGGTCACGTTGAATCGGAGGAGGTTGTCCTGATACAGGATCAGCGAGGAGAG
AACTAGTTCCGGAGAGTCTTCGTCTGTGTAGAGCTGTCCGAAGATGGCAGGCCCCATGTTCTGAATCGCGCTTTCGGAGA
GCTTCACGGAGATGCGGCTGGTGAGGCCGCGCGGTCCTGAGATGCCGCCCTCTGCGGTATGCCAGCAGAGGGCGGACTTG
ACGCGGATGGTGAGTTCGGGTGTGCTAGTGTCGTTGCCGTTGTTGTTTTCGTCGGTCGTAAAGTCGA*g*GTCGGGCATGTT
GAGCGAGACGCAGTCGTGTAGTGTGGTGTGTTGGGCGATTTTGCGCCGGAGAGGCAGCTGGTGCCTCAAGGGCGGCATAA
TCTTGAGGGCGAAGATGTTGAGAGACAGCCGGATGATGGACATTTGGGTCATGTGTTCGTCCACGAGGGCTCGGATGATA
TTGCAGCTCCCGGCCTGTGCGGGGAACAGGGTGAGGTTGGGGAAAATGTCCGAGTCGCCATTGGCTACGGCGCAGAGGAG
GCCTCGTTCTTCGGGATCGCGGTGCGGGGTGATGGACTCGAGCTTCAGTGGGGCGGAGAGCAGGATCTGGAACGGATGCT
GCAGGTCGGCGGGCTGGAAGTCCACGTGCAGCTGAATCAGACGAAACGTGCCATGGCGCGTCAGGGGATTCTTGATGTGA
TGCTGGTCCGTGAGGGTCAGGATGGTCTCGTCTGACTGGAGGAGAACGGAGGCCTTTCGGGGGCGGCCTCGGGGTCCATT
CCGCATGGTTTTGTTCGCTCGCGGCATGGTTCGAGACTGATCTCAGTGGCACGGGTCAGCCTTTAAATATGCAGGTCGCG
GATTTGTTATCGGGTGAGACGTCACAGACCACGAAAACCGGTTGCTCGTGAATTAGGTCATCCAGTTGCCTCAGCATGAC
AAACAAGGCCGAGAGTCTAGAAATTTCCTCGTCAGGTGACACGTTGTTGGGTCTTCCGCGGAGCCCGTCGGCGGCGTGTT
CACGGTCCATCCGCAACAACAGTTCTTGGCACTTGACCAGCAGCATGGAGCTGTCTTCCAAGGCTAACTTTCGAACGTAG
TTCATGGTCAAGTCGGAGGCCAGGTTGGTGACCATGGACAAGGAGAGGCACGCGGTCTTCATATCGAGGAGGAGGTGGTT
GTCTGTCACGGGATCCGGAATGGTAAATGTGGCGTCGCGGAACGTGATGTGCTGCAGTTGGTTGACGGCCGCCTTGACCT
CTTCGTAAGAGCGGTCCAGTGAAAACAGTGCCATAATGAGAATTCTCTGGTTAATTTCCACGGCCAGTTGAGAGGGTAGG
AGCCATGGGAGCACCAGATTCCAGGATCCCACGGACACGGGATACTCGCGCGCCACCGGACCTTGGAAGAGCGGGGGTAG
GACGCACAATCTGTCGCCTTTTTCCCAGGTGACGGGTCCTGTGTTGAGAACCGTGTAGAGTTGCCCGTGGGTGGGCACCT
GCATGGCGAGCTGGTTGCCCTCGATGCGTCGCAGGATGGTGGGCGTCATATTTCGCAACAGACTGCGAAGCCGCACATAT
CCGCGGGTGTGATCGATGTATTGGTGAAGACCCAGGTGATGGTGCTTGATGAGATGATGTCTGCTGGGGACGGGGATAAT
GGCTCCCGTGAGCTTGGTGAGTTTACCCACGTCGGTGAGGCTCAGTTTTTGCTCGAACGTGCAAAAGATGGAGAAGTCCG
CCATGTCGCCGTCGATTTTGCGAGCTTTAACTAGCGTGCTGCATCAAATAGTGGAATTCGTTCTTTACGAGTTAAATAAC
ATGTGTTGATGCAGCGGGATGCGTTCGCCTATCACATAGTTGCCAAAGTGAGTTTCCGACGCGGCGAACGAGGCCGACGA
GTTTTTGAGCCTGGTCTCCATCAGCGCCTGCGTGGTGGTGGACAGGATGGGCAGCGCCTCCTGTAGCATGCGGCACGGGT
TTTCGATCAGCTGATCGGTGCCCTCCACGCAGATGTACTGAGTGTCGGTCTCTCCCCGGATACAGTCTTTGGCTCGGAGG
AGGTACTCGTCGATGGTCTTAAACAACGTTTTGTTGGCGTTGATAATATCATCCGTGTTGAAGAATTGCGCGCACGGGCT
GTAGAATTGCGAATTGTAACCTAGGCGTTCGCGGTTCCGAATGTTGTAAAGGACGTCGCCTAGCGAACCCCTGTTGGAAG
CCCAAGGATTATGACTGGACGCGAACGTTTGGGCGTCGGGCTCGGTGTGATCGTAGAGGGTTTTCATGGCCGCGTCCGTG
TCGTACGGATCGACGCCTAGCATGCATGACGTGCGGCCACGCGGGTTGTTGGGGATTTTGAAATAATTGATGTCCATGCT
GACGGGTGTGATGATGAGTTCACATACAGCTTTCTGTCCGTGCAGGTTGGCCGAGGAGTTTTTCTCCGACATGCTGCCAA
AAGTCAGCATGGAGATGGCTTCGGTGTCCAGCAACTGCGGACGGTCTACACCGGCTGCGTGGCGTATCCACCGATCCACC
TCGTCGTGTCGATGCACGTGCATGGGGAAGACGCGGAAAAGGTTCTGTACATGCACGCCCATGTCGCTGCGGGAGCGGTT
CAGGTATGCGATGCAGGTGGCGGACGTGTACCCCAAACCCATGTCCACGGAGTTAATGTTCTGAGTGACGTGATACGTGG
TATTGATGTCCCGCTCCTCTTTGGTGACGTGAGGCTTGTTGATGATGAGTGACGTGCAGGATTTGCCGCTATAGAGCAGC
ATGTCCACCTCAAAAGTGTCGGTCCGTACGGCTGTTAAAGCAAAGCCGGGGTGCACGCCGGCTTTGGCCATCAACGCGAT
GGAGACGGGAGAGATTTTGTACAGCATGCAGGCGATCGTAAGGACGGATTGCAGGGTGTTGGGGCATCCCGCCGTGTAGC
GCGAAAAGGGTGATCGCAGCCAGTTCCAGTATTCGTGCCAGAGGGCGGGGGCCAGAGGAAAACCACCATCGTGCCGGTGG
TAGTGAGGGAACTCGGTCACGTATCGCTTAATGTCGTCGCTGATGGCGGAGCAGATGGTGGGATCGGAGTATAAACGGTG
GAAAGGTATGGCGATGCAGTACTCGGCCAGGATTTTGGGGGCATTGATGGCGCAGCAGCCGTTGTACAACACGTGTTGCA
GGGACGTAAAGTCTGGGGAGCCGTGGCGCTGGGCCGGGTCCAGAGGCGCGCGCACTTCCAAGAGTTTGGCGTGCTCGGGC
ACGAACTGCAGGACCAGGAAGAGTTCTCGACAGGCCTCATAGAGGTTCACGTCGTCGCTCGTGCCGATGTCTTGCACCAG
CTGCGGCAGCAGTTCGCCCAGAGCCTCGGGTTCGGCGGGCATAGTGAAAGCGGGACGGTAGAAGAGGTCTACCAAAAGCG
TCTTGAGATTCAATCCGATGCCGCAGGCGCGGTTGTTGGTCATGGCGGGGAGAAAGCACAGGTAGAAGATCTTCTGCATG
GTCCATTCGTCGTCGGTAGCTCGGTGTTCGTCGACGAACAGGGGTTCGTCGGCGTCCATCGCCGCCAGACGCGGCACGTC
TGATATGCCATGGTGACGGGATTCGATGTTGGCGTTGTTCAGCGGCTGGCGATCGGTGACGATCTGGACGCCTTCTTGGT
TCCGGGGCAGGTGAGTCACGAAGGGCGGCCAGAGACGGTGATCGTGGAGCGCGTTGGCGTAGGCGCTCAGAGGTTCTTCG
ACCAGCTGGCCGTTGTTGAGGCCTGGGAGCGAAGACACGCGGGTTACCAGTCGCAGGAGAACGATGATCATGCGATAATG
TGCGTGAATGTTGGAAGGGATGAGGCCGTCGCCGAGGTGTTCGACGATCATAGCGATCATGGCGTAGCTGTGGCCGAAAG
CCAACAGCTGGCGGTTTTGGAACATGTTGACGATACATCGCGCGATAAACGGTCGAATGGTCAGAAAAGCGTCCACGTTG
CCGTGCACCAGCGCCTCGATCAGGTAGAACAGTTCGGGGTAGGTGGCACAGGTGACGGTACTTCGGAACAATTGCAGGGT
AGCTTCGTAGTCCGGAGGCGGTCGGAGCCGAACCATCTCGGCGATGTGTGCGGCGCGGAGTTCTTGAAAGCCGCTGGGGG
CCAGTCCGTCCGGTATGTTGCCAATGACGATGCGTGGCGTGCAGAGGGCTGCGGTCTCCCCATTCAATTGGTGGTGTGTG
AAGTCGAAAAAGGGATGGACTTCGGTGTAGAGTGTTAAGTTACCTACCTTGTAAAAATCAGTAACCGTGTAGTCCTGCTT
G[W]TCTCGTTGGTGGTGCGAGGCACTTCTCGGCGCACGCGGTAAAAGTGCACGAGGCGCCGGGCCATGCCGCCCATCTGTT
CTTGCTGGAAACGGCATTCGGTGAGTCTGGCTACGGCCGGGTCGGTGGGGACTCCCCGTTCCATGAAGGTCTGCAGGCAC
GCGTTGGATTCGTGAATGACAGGGTGACAGAGAGTGGCCAGCGCATCGACAAAGTCGAGTCGCTGGAGGGCGCGCTCGCG
ATTCAGGAGGAAAGCGGAGGTGGGGAGCGCATTCTCTTTGGTTTCGCTGAGTTTCACTTTATTTTCCACCGTAGTGTAAG
CCTTGTCTTCGGGAAGGTACAGTCCCACCGGGAAAAAGAAGGTGAGATCAACGTTTCTTTCTAGGGGATTTTTGGTGTCG
GTGTTTTTGTACACGTCCCGATGGTGTTCTATGGCTACCGCGTGTTCGCCCAGTTGAATGACATCCATGGCTAACGACGT
GTGACTGTTTTGTCCGGCTTTCTCTAGCGAGGAGCCGGTGAGATCGGGGTTACCCGAGGTCAGGTGAGCGGTGTAACTGT
TAAAGTCGGCTAAGATGGAGTGGTGTGCGATCGCCGTGACGGCGTTTTCGGGACTGAGGACAAAGTTGCCGTAAGTGGCG
GGGGCGGTCACGTTCTGTTTGGAGATGCTGCTTTTTAGGACTGACATGAGGGTCTGGACCACCTGCGCGGTGCTAATCAG
GACGCCCCGGAGCTGTTTGCCGGAGCTAGTCACGTATGTGCTGGGGTTTTCGAGGATGCTGCCGGTGGTGGCATCTACCA
TTTTAGTGAGCAATTTTGTAATGTAGTCGCGATCGGACGTTCGGTTCAGCAGGAACATGGTGCTAAGCATCTTGGTCTTG
AAACTCTGCAGGATGTTGCTTCGTTGGATGCGGTTCAGGGCCTGTCGACACAGCGTGGCGTTCTCCACCAGCGTCTGAAC
CACAAAGTACGGCGGAGCCTTGCGCAATAGCACCTGGATGAAGCAATGAATCAGACCCCGCTCCATAGCGTCGGCGGTGT

TTTTCAGCGCCCGCATCACGGTGTGCATGGCCTCCACGTTGAGAATCTTGTCCAGGATCGTGTTCTCGAAGGTTTCGCGC
AGGTGCAGCAGACAGGCGGCGCTGAGTTCAAAGGGAATGGTAATGGGGCTTTTTTCACTGTATTTGGTCACCATGATGGT
GGTCTGGCGGGAGGTGGGCATGCCGGCGCCGCTGGCCACGCGCGGGACTTGGATGTGAAACAGCATCTTCCCTGTGGTCA
TTTTGTTCAGGTCGTCGAACTTGATGGCGTGCGCGGCCGTGGCCAGGGCCGTCTGAAGGAAATAAACCCATTCCAGACGG
TTGCAGAAGGTGCCGAAGATGGCCTCGAAGTGGAGGTTGTAGCGTTCCGGATCGTCCCCGTAGTAGATGCGCAAGCTGTC
GAACATCTCCTCTCCCGCGCTCGTTTTCACGTGGGTGAGAAAGTCGGCTGGGATTCCAACTTTCGGGAGTAACTCGACCG
CCGTCCAGTTCTCCATGACGACGGCGTGTGGCGATGGCGTCCTCGCCGACCGTCTTTGGGTGCCGCGCCGCGCAAGACGC
TCTAATAGTGGATTCGGTGGCGGAACCCCAGGTAGTCAATGTGCCCGTACACGTGAATTCGTACAATCTCACCCAGGAGC
TGTCGTCGGTCGAAGACGCTCGCTTCTGCCAAACGCGGCCCGTGGACGCGGAGCGCGTGCGCGCTGTTTTCGGCGCGCTC
TATCGATCCGCCTCGCCGCATATGCGGGAGGAGAGTGACCGCATCAAGCTGATTTTGGGACGTTTGTTACTGGGGCCGGT
GGCCGTGCCCTGCTACTGTGACGAATGGGAGGCGAATGACTATATGGTGGAGTCGCCTCAGTTTTGCGCCGGCCCTCTGC
TCTATGTGCACAGGCGCTGTCCCTGTCCCAACGCGGGACGCTCGCTCGCTTTCACCGTCATGGAAGGGCACATCGCGACG
CATGTTTTTAGAGGGCTGCTGTCACTCACCGAGTGGAACCAGCACCTGCCCGACATGTTTTGTCCGTGCGGGACAGTGAG
TCGGCGGGATCGGTACAGTATGGCCTGCTTGCCGCGTCATCTCACGTTGTATATTACTGATTACCCTTACTTCATGGTGG
AGATCGGGCGCGTGCTCACCGTGAGCGAGGTGGACGACCATGTCAGCCTCATGACGAATTACTTGGGGGACGCCGTGACA
CCTCGTCTGCAGGTTCACTACAAGTTGTTGTTTGGGGTGAACGTCAGGCCGCTGGCCCCGCTAGCCCTGAGCTCCACGTG
CGAGTTCTTTCTTCTGGAATTGCAGAAGATGTGGCTCGCCGTGGAGTATCACAACGAGGTGACGGCGGATTTTTTTGGCC
GCGTCTTTACCGAACTGCATAGGGATCGGGGGCGCGTCATGATGGCCCTGCGAGTGCCCGACCAGACCGTGTGTCATCTC
AACCACTTTGCCCTGGGCCGTTTCAAACGGCAAGTGTTGTACTTTAAGCTGTCAATCATCTACGGCAAGTACAAAGCCGG
CGTGGCGCGGAACACGCTGTGTTATCGACGTCTGAGTCTAACTTTCGGGGACAATGACACTGTCTGGAGGAATATGTTTT
ACGTGTATTACGAAATAGGTCGCGGCGAACCCCTGAGTCGCAGCCGTCAGCCGCCCCAGATCACCCCTCCTCCCCGTTGG
GCCTCGGGCGCGTGTCTGGTGGCGGCTCGGGTGCCCCCTTCCGTGGCGGCCGCGGCCGCCGCGAATCGCCGAGATTCCAC
CGGCGACCGCCTCAAGCGGTACGTGTGTATCATCTCTAGACTCATGTTTGTTCGCTACGGGCAGTTGTGGCGCCGAGAGA
AGCAGCAGAGCGCTTTTGATTTCACCGGCCGCACTCTGAGACAGGCCGCTCTGGAGCACCATCGCCAGAAACAGCTGTCG
GCTCAGCGGTACGCGGCCAGTCGCCGGAAATTCATCGGGGGGATGGAGTTCTCGGAAATTACGGSCGTCAGCCTGGATCG
CGTGGCGGTCAACGCTTTCAACACCAATCGGGTGATTAATATGAAGGCGGCTCTATCGGGACGTGTGGGCTGCCGCGTCA
ATCGGCTTCCCAAAAACATGACCCACAGCTTTGTCATGTACAAGCACACGTTCAAAGAACCCGCGTGCACCGTCAGCACG
TTCGTGTCGAACGATGCCGTCTACACGAACTCGCTGAATGTGAACATACGAGGCTCGTACCCTGAGTTTCTTTACTCGTT
GGGGGTGTATCGCTTGCATGTCAATATCGATCACTTTTTTCTCCCTGCTGTGGTTTGTAACAGCAATTCGTCACTAGATG
TGCACGGCCTGGAGGACCAGACGGTCATCCGGTCGGAACGCAGTAAAGTGTACTGGACTACTAACTTTCCGTGTATGATT
TCTCACACTAATAATGTGAACGTGGGCTGGTTTAAGGCCGCCACGGCTATCGTGCCTCGGGTTTCGGGCTCAGATCTCGA
GAATATACTGCTCCAAGAGCTGTCGTGCATTAAAAGCATGAGGGACGTGTGTATCGATTACGGCCTACACCGTGTGTTTA
CCAAGATGGAGTTGCGAAACTCCTATCAAATTCCTTTTCTGTCCAAGCAGTTAATTCTGTTTATTCGCGCTTGCCTGCTC
AAGCTCCACGGCCCGGACAAACGGCTCTACTTGGATCGATTAGTATTTGAGGCGGTGCAACGGGGCGTGTTCGATTACAG
TAAAAACATTACGGCGCACACCAAGATTAAACACACTTGCGCTTTAATCGGAAGCCGTTTGGCCAACAACGTTCCGAAAA
TTTTGGCGCGCAACAAAAAGATTAAGTTGGACTACCTCGGGCGCAACGCCAACGTCTTGACGGTCTGTCGTCACTTGGAG
GTCAACAAAATCTCTCGTGCGCGTCTTAAGGTGCTGGTAGACGTCTTGTCCACTCTCCAGGCGTTGAGTGAGACGGCTCA
CACCCGAAACGTCATTCGCCATACACTCCTGCGCCTCTGCGGCTCGTTGTCGGCGACAGCGGCGTCTTCCTCCTCCTCCT
CCGTTTCGATATGAGCACCGCTCTTTCCGCGAACGGTGACGGTCGCGCTGCGGATGAGTTGCCAGGACCCGGCTGGCGGG
ATGCGGCTCTGGTTTTGCCCAACGGGCTGGTGCGCGAACACATGATCCGGCACTCCGACGTGGCCGACATgGTTCGCCGC
ACGATTCCGTCACCACCCGATGCCGAGGACGGCTGTGTGTTCGCTTCGGAGCTGGCCTTCCATACGAGCAGCCGTTTTAA
TCGAATATCGTCTGTATTTTCCATTTATTGGCAGAATCACAGCCAGCTCATTTATGCTCTAACCGGTATCACCCATTGTA
CTAAAATCGTCATCGAATGCGGCCAGGTGGGATCCAAGGGCGATGACAACAAGTTTTATGACCGTCCCGGTATTTATGTC
ATTCGACCGTCGGATGGGACTGTCACGCCTCGAAATGTGGTCTGGCCGGGAACCAGTGTGCACTGGCTGGACAGCGTCGA
GATCCGTACGGTACAGCGTCGCCTGAATGTGGCTCGTTCCTTTATTAATTACTTTCGAGATTCTATGTTCTGGATAAACC
AGAAACCTAATGACGATTTATGTCCTTGTCCGGTCGAGGTGGAAGAACGCCTGTTTCCTCTCTTGAACCTGTCTCGTGGC
GACGTGAAAGTGTTTGATGCGCGTGTAGGCGCTGCCTACAACCGACTGTCAACGAACAATCTTCCGCGATGTGGACGTCG
TCTGTTGGATCACTGTGTGAATCTGGCCGCTTCGAAAAAACTGTTACTACTGAACTTGCCGCGACTGGAAAATTTTTTCT
TATGTCAGGTGTGTTTGTACGAGCTGGATGAAGATGAGGTTGGTGAAGAACTTTTGGGCATGTTGTGCGGTAAGGTGAGC
GACGGAGACGCTAAATTCCTGTTACATCGCAAAACCATGAAGGTGGCGGCCTGTATGGCGTTTATTTTAAATTGTTTACA
CAAGCATCAGGATCGTCTGCCGGAGGTGAACCAGCGGGTCGATGAGTGTGACCTCTTGACCATCGCTTTGCGGCGTTACT
ACCGACATCATGCGGGAGTGCAGTCGCGTGTGGTGGCCGCGGCTTGCAAGTTTTTGGAGCACTACGGAGATTCGTTTTCT
CCGTTCGCCAGTCTTACGAGCCTGGGGGTTCACATTCCGCTGGACGCCAACGTGACTCGAAAGCAGATGATTGCTATTCT
GCGGACCTAACTGATTTTAAACTTGGCGGTGTACAGCTCGTCTGACAGGTACGTAGCCATAATGACAGCCACCACCAAAT
CGTCGGAGTAGCGGTGTTGTTTTTTGGCGCTGTACGTGGTGTTGCCTTCTGCCAGCGTGATTCGGTGCAGATTCTGGATC
TGATCCAGGAGGTATTCGATGGGATCGTGGTTGAGCTTGATGGTGTAGGAGACAATCTCCTGCGAGGCTTTAATGTAGCC
CGAGTTGAAACGGGAGATGAATTGTTCCACGGCCAGGCGCTTGTCACGTCCCATGAGGTAGAAGGGTTGCTCGATGTTGT
TTTGATCGGGGGTGTGGTAAAACAGCACGGGGACCAGGCTGTTGCATTGCACGTTCTGTCGGATGAGGCAAGCTATGCGC
ACCGCGGCGGCCTGGTTGGAGTTGCCTTCCACCGCCACGCGCAGCTCGCTGATGTAGGGGTGTAGGTTGATGACCGAGGT
CAGCATGTGTGCGGCGCATTCCGCGATGGCCGTCTCGGAGCTCTCTGACAGGTCACGCAGAAAGTAGTGTTCCAGGCCGT
AGATGATGAATTGGTGACGGTACAGCCCCACGGCGGCAATGCCCGTGCCGGAGGCTTTCCGATTTGTGGTGAAGGCCGGG
TCGAGATACACGTAGATTGTTTTGCCAAAATGCTCCTGCGCGTTCACGTTCAGGGTGCTGTAACGTAGGATGTCGAACTC
GTCGCGGCTCTGGTCCGTGATCAGGACGTTGCTTTCGCTGATCTTGTTGGTACCGCCGATGATCTCGTCCATGAACGAGC
CTGGCATGAACATGTTGGCGGTTTTTCTGACCTGCGAGTTGAGGCTAATGAAGGTGGGCTTGTGCAGGCGGTAGCAAGGG
CAGGCGGTGGCGTCGCCCTTCTCCGTGAAGCTGTGCAGGTGCTCCTCGCACACATAGGATACCACGTTAAGCATGTCAAA
GGGCGCATTGTTGAGGCGCGTGAGGAAGCAGGTGGCGTCGCTGGTGGTGTTGGTGGAGGAGATGAAGATGATTTTGGTGG
TGTTCTGCGCCAGAAAGCCCAGGATGGTGTTAAAGGCCTCCTTTTTGATGAAATGCGCCTCGTCCACTAGCAGGAGGTGG
AAGTTTTGGCCGCGAATGCTCTGCGGAGAGAAAACAGAAAGGCGGGGATTAGTGTGGGTGTGTCGGCGGTGTGTTACGCG
AACGCCCATGGGTGTCAGGTGGCAACATAAAAGGTCAGGTTCCTTCCGTCTCGCGCGAGAAGACGGCCATGAATGCATTG
TTGGCAGAACTGAATCGACTGGGTGTCGCGCACCTGACGACCGAGGATGTGTTTACTTTCGCCGAGAACGTGTTTCACAA

CTTTGGGTTTCTCTTTCAGACCGAGGAATCTGGCCCCCGGAAGATAGATTTGGTCTCTTCTCTCTTCGAACACCTCACGG
TGGAGTGCGTGAATGACATCTTGGACGCGTGCAGCAGCGATGAGCAGCCACCCTCACCAGCAGCACCACCAACACCAGCA
GCAGCAGCCACAACAGCACCCAAGACTGGCGCGGGGMCGGCCGCCAAAGAGCGGGACGCCCGGCACCACGCCGGGCGTGC
CGGACAGTTGTGACGTCTCGGGGGCTTGTGACATGCGGCACGTGCAGAACGTGTTTACCGAGGAGATTCAGCTCCACTCG
CTGTACGTGTGCACACGTTGTTTTCGGACTCACCTGTGCGATCTCGGCAGCAGCTGCGCCCTCATTTCCACGTTAGAGGG
CTCGGTGTGTGTCAAGACGGGGCTGGTGTACGAGGCCCTGTACCCCATGGCGCGGGAGCACGCGCTGGAGCCCATCGAGG
AGACGTCGCTGGACGACGTCAATGTGATCGGGGCTGTGTTGACCGAGGTGTATCGCTACCTCATGAGCCATGCCGCGCGG
TACGCCGATGTCATTCAGGAGGTGGTGGAACGGGACCGCCTCAAAAAGCAGGTGGAGGATAGTATTTATTTCACTTTCAA
TAAAGTCTTCCGTTCCATGCAGAACGTGCACCGGATCTCTGTGCCTATTATTAGTCAATTGTTTATCCAGCTTATTATCG
GCATCTATTCGAAACAAACTAAGTACGACTCGTGCGTCATCAAGGTGAGTCGCAAAAAGCGAGAGGACGCGTTGCTCAAG
CAGATGCGTTCCGAATATGGAAACGCACCTGTTTTCGGACCTGGCGTTTGAGCAGGAGTTCGACGATCAGCTGCAGTTTC
CCCTGCATGTGGTTCTGGACCCCGTCGACCTGTCGTTCGAGGAGGCCAAAACCGTGCGGTACGTCTATTACCGTAGCGAC
CCGGATCCCGTGGAAAACTGGCGACGGGCCGCGTTTCGCGTGTTTCTGCGGCCGTTGGAGCTGCTCACCTATCTGCGGGA
GGCAGGGGTGACGGTGCCCCCTATCGGTCAAGTGGTTTCACTACACCATGCGTTATTCACCACACTGGGAATCCGCTGCC
TGGAGCGGCGCATTTCGGGGGCGCAGATGTTGTATCTGCGTTTGGTGTGGGAGGACGAGCGTGGCGGCGCCGAGGGGGAT
CGGCGGCGCTTtCGGGACTGGGAGTTTCTGGTTCGCGACTTGTTGCAAGGGGAACTAGAGATTtCGGAGTCGGAGGAGGC
GAGCGCTGAGAACAACGTCTCGTCGCTCATCACAGACTCGGCTGTGGAGATAGCCAAGGCCAAATCGGCCGCCACTCCGT
TTTTTGAGATGCCCACGAGGAGCGGCGAGGGGGAGGCCGCGGAAGTGAGTGAGGCAGAGCGCCGCGCGGTGCACTTTCTG
CGCGGCGACTCTGAATTTACGTACCACGCGGGCCCTCTGGAGCCTCCTTCGAAGATTCGGGGGCACGAGATTGTGCAGCC
GCGTCTAGAAGTGAACCCCGATGTCATCTACGCCTCTGGGCCGCACGACGACGACCGCGTCTCCAAGACGGATGAGTGGC
AGCAAGGCGGCTTGCTGCGACTTCGCTCCGTCTGGGATGTGCAGCAGCGTCTGCGCCTCCACGTCCTCTGGTACGCGCAC
TCGTTTTGGAGGTCGCGCGGACTGCGCTACGAGGACCGCGAGGAGGATCTGCGCCTCACTCTGGACGCCTACTTTGATAG
CGTGGCCGTGGAGTACGAACTGGCCCGTGAGGTGTACCGCGAAATCAAGGCGGTGCTCAGAACGGACCGTATGTTACCTC
AGAAGTTTTCCTGTCATCTGTCTATCGAGACCTCCTGGCTCTTACTTTGGGAACTGTTCGATTGCGCCCTGGAAATTTGG
AGGGACCGAGCGGATGTCAACAGCTGCATTATTAAGGCGCTAACCCACAGGTTGCGCAGCGGCAAACAGGTCCACGGAGA
CTCTGCCGCGGTGGATAAAACGAATAATCCCTACGAGACTTGGTATGCCGATGTGGTGCGCTGCGTGCGAGCGGAGGTGG
ACCTGGGGGTGGAGGTACGAGTGGAGACGTGTCCGCACAGCAATCTGTGGATCACCAGAGGTCGTGACGGACAGCTAAGG
AAGTGGGTGACGCAGCCTGACACTTATGTGCTGTATGTCACCCCGGGACTAGTCTTTCACTGGGTCTCGCCCGGCGGCTT
CGCAGTCTCCTCTCGAGTGCGCCTCGATGGCGTCGGGAGGGATCACTTCGAACGATTCCAGATGTCTGCGCCAGTTCTTG
CAAAGAGAGTGCTGCTGGAGGGAGGTGGGCCGAGGCCGAAAGTTCCGGGAGTTTAGTGTATAGCCTGTAGATCGGCCATC
TTCTCCGTGGAGGGGGRCAATGCGTGCATGGCCTGTCAACTGTATCTGTTCAAGCGAGGCGGTGAAAGCATCATCTGTTT
GGCGTGCAATGGCCACTTCATGGGTGTGTACTCCTGTCCTCGAGTGCAAAGAGTACGACAAGAGGTTGTGGGACTGCCGG
CCACGTACAAGTTGTTATTCTTGGGGCCCTTGGGTCCGGCCCGCATAGACTTTGTTCCAGCATTTTCTCCTATGATGAGC
GTAGTGCCTCGCTGCTGTATCACTCCGCAGCTCATTTACGACGTGTGTACCCTGGTGCCTCCGGAAGAAGCTGAGAGTAT
TCGAGTGAAAGGCTGTACCGGATCGTACGAGCCCGGAGTGGAGAAAGCTATCAGTCTGGGCGGTGCCGGTGCTTGGCTGG
TGAACAGAAGCAGCGGCTACGCTTTGTACTTTTATCTGCTGTGTTATGATTTATTCACGGTGTGTGGCAACGACGAAGAG
TTGCCTTCCATGGCGAGACTGATGGCTCTAGCCACCGCTTGTGGACAGGTTGGATGCAACTACTGTAAGGATCACGGAGG
GCATACAGATCCTACTGGATGTTATGTGGGATGTGTTCCGGATCGGGGCAACTGTCTGTGTTACACCTTGTGCAACTCTC
CGACCATGAATCGTATCACTCACGAAAACCCCGTGGCATTCTTTTGCGATGTGGATAGGGCTACTTACCTGTGTTCTGTG
GGCTCCAAGACTAAAGGGAAAATTACTTTGGCGGATGACCTGGATTATCACATAGGCGCCGAAGACGCGGATGGGGACTG
GGTGCCGTTGCATGCCAACGCTTGGCGAGTGGTGAAACTGGAAGAGCCGATCGGTCGGATGATTGTGTGCGCCTGCCCGG
TGTTGGAAGACCTAGTGCACTGAAGGCGGCCACCTGTCGAAAATGGACGACATGATGATGGATGGGTGGACACAGATGGA
CATACTCACGTTGGTGTTGTAGCAGCTGGCGAACAGCGCCGTGCTSTTGGATCCCTTGTGGTMGATGCTGATCACGAAGT
CCTTGTTNTCCACCACGTAATCCTTGGCGAACGTGTGCSGACATCGGAATTCTACCTCTTTCAGCACGAACTGAGACACG
TGTTTCTGGTGCGCCACGTAGCCGATGCTGATGCCAATCATGTTCTTGAGCAGAAAGCAGATGATGGGGATGATGAACCA
GGTCTTGCCGTGGCGCCGAGGCACCAGAAACACCGTGGCCTTCTGCTTGAACACTTCCACGGAAGTTTGCGACAGGAAGT
CAATCTCGAAGGCGTGGACCAGATACTGAAGCACGCGATTGGCCAGCACCGGGATCTTGGTCACTGCTATAAAAAAGATA
ACGTGAATCAATAAATTCTTTTGAAACGGTTCGAGTCGAATGGGTTTGGCGTCTCCCTCCGCGGTGGTGAAGCCACCGTC
TAACCACTTTTTGAAGTCTGTCATGAAGTTGTTAATCTGGACGAACTGAGGGTCCCGATACAGTTCAGTCAGGGCGTCCA
GTTTCTGGTAGGACTGCCTCTGCTCCTCGGGGCACGGCAGAAAGTTCAGGTCATCCAGGGCCGTCTTCAGGTTCTCGTGA
AACACGAGCGTCTGTCTGTCTTCGTCAGAGTGATAGTCGCGATAACGGCCGCAGAACGTCATCAGGGGTAAAAACGCCTC
GTTGCAGCAGTGCGCCAGCCCCAACTCGGGGTGCATCATCTGGTAGCGCTTGCGGCATAGGGCAGCCACGTTAGTGAAAG
AGGTAGAGATGCAGGAGCTGGGGTTGTTTTTGCGTTTTTGCAACTCTTCGTAGCGCTGCTGGATCCGCGAGGCGGATTCC
CCGCGCAGCATGGCGACAACCACCAACTTTTCTTCTTTTTCCTCTCCCGACAACAACAATTCTCCTCCTTCTCCTCCCTC
GGAAATGCTGTTGTCAGTGGAGGAGGAGGAGGACGGGGAGGGAGGCTGTCACCAGCAGCAGTCACAACAACAGCCGCGGT
CGGGCGCGGCGCCAGCAGCGGCTGCGTGCTCCTCGATTTCCACGTCGCTATGTGCGGTGGAGCGGATGGTTGAGCTCTCC
ACACAGACTCCGTCGTCCGATTTCGAGGTGCAGGAAACGTTGCGTTTTGACGAAGCCGTAAATATGGCTCTAATCGCCTG
CGAAGCCGTGTCTCCTTACGACCGGTTTCGCTTAATCGAGACCCCCAACGAAAATTTCTTGCTGGTAACTAACGTGGTTC
CCAGGGAGCCCGCCGAGATGCCCGCGATCGGCGCGGAGGCGGAGCGCGGCTTTTCCTCACGGTCTCAGGACGTGTTGGAT
CGCTTCGCCTGCACGTCCGCGCCTCTACCTGCGCTCGGGGTTCGTGATTACGCGCTGCACAATGCGGATCGCCTCACCTA
CGAGGGGGATTTGGTGTACGGGAGCTACCTGATGTACCGCAAAAGCCACGTGGAGCTCTCCCTGTCTAGCAACAAGGCGC
AGTACGTGGAAGCCGTGTTGCGTCAGGCATACACACCTGGCCTGCTGGATCACCACAACGTGTGTGATCTGGAGGGGCTG
CTCTACCTGCTGTACTGCGGACCCCGCAGCTTCTGTTCGCGGGACACGTGTTTCGGGCGGGAGAAACACGGGTGTCCGTT
CCCCGCTCTGTTGCCCAAAATCTTTTACGAACCCGTTCGCGATTATATGACGTATATGAATCTGGCGGAACTGTACGTTT
ATGTGTGGTATCGCGGCTATGACTTTCCCTCGGCTTCAGATTCGCCCGCGGGGGTCGGTGACGACGAGCTGCGCTTGCCT
TTAAAGGCCGTGTCGCTCGATCGCCTGCAAGAGGTACTGAAAGCCGTCCGCGGGCGTTTCTCGGGACGCGAGGTCCCGAC
GTGGCCCGCGTCGTCGAGAACCTGTCTCCTCTGCGCGCTCTACAGTCAGAACCGCTTGTGTCTGGACCTCGCTCGTGACG
AGGCGCGGACGGTGCACTACAGCCCGATCGTCATCCAGGATTGCGCCGCGGCCGTGACTGACGTCACGCTCAGCCATATA
TTGCCCGGGCAGACCGCGGTCACCCTCTTTCCTGTGTATCACGTAGGGAAGCTACTGGACGCGCTCACCGAGACCGAAGC
GGGCATCGTGTCTCTTAATCTATGACGTCGGCCAACAAGGACCTGTTGAAGGCGGTAATGCGCCTGGAGCTCGAGAAGAA

ACAGAATCTGTTTTTGCGACGCGCTTATGGAGCTCAGCACCGAGTCACCAAACACCACCATCTGAAGGTGATGCAAACGG
CGGCTCGCGAGCAAACCAGATACAGCCAGGTGGCGGTGACACAGGTGGCTGCTAACGTGGTACGGGAACGTGAAAGCCTT
AAGAAGGAGCTCCAGCGCGCTCGCGCCCTGCAGAAGGCCACGGACGCGGACGAGCTTCTGGACTCTTTAATGGAACTGAA
GGATACGGTGGAGGACGTGTGTGACACCTTTGTCGACTCCGTGTCCAATACCTGTGAGGTGGATCTGTTGGGGGAAGAGG
ACTGAAATTGTGTCAATAAAAACCAAGTTGTCACCACCATGGCAACTCGATCCAAAGTTAAGTCTCGTCCTCGTTCTGTG
TCGGCCGAGTCGGGTCGCCCGTGGCAGCCCGCGCCCTTGCGACGTTCCAGGGCGTCTCACAGGCGAGCTTGGATGCGGCA
GGCCGTGGCCGCGACCACGACGCTAGAGAACCAAGCCGCTCTCGTGTGTTCTTCTTCCTCCTGTACCTCTGCTGCTACGC
AAGACGCTGCCGTTTCTACCTCTGCCGCTTCCGCCCAGTCTGAAAAGGCGGCGGCTAGTCCGCCTTCTCCTCCAACGCCT
TTGACGTCTCGGGGTAGAGCGGGGTCGCGGAAATCGAGTGTGGATACTGTTGTTGACACCTGTCACGAACAGCAGCTGCA
GTGCCGCGAGGATGCGTCGCACTGGTCTTTCGAGTACGAGACGGCTGAACTCCGAGGAGGAGGAGGAGGGGGGGAAGGAG
TGCAGGCTCGGGCGCCTTTGCGATCTCGCAACAACAGCCACGACGATGCTGCGACGACTCCTCAGCAGCACGAGGAGTAC
GATAACTATCCGGCCTCCAGTCAGACTTCGGAGGCGTCGGTGGCTGATGATCACGAGGCCGTGCATTGTACGTGTTCGAA
CGATCAGGTCATTACCACCTCTATCCGGGGTTTGGTCTGCGATGCCTCGTTCTTTGCGCGAATTACGCAACCCGAGCTCT
GTGAGATGGCCGTCACCTTCTTTGGTGTACGTGCCTAAGGATGAACATTTCTGCCaGCGGGTCTGTTATGCCGTGGACAT
GAGCGATGCCCACAAGCGCATCGGGTGCGGTTCGTTCGGGGAAGTGTGGCCTCTGAGCCGGCATCACGTCGTTAAGATTG
CCAGAAAACACAGCGAGACGATCCTCAGCGCCTACATGTCCGGCCTGATCCGCACGAGGGCGGGACCGCGGGCCGAGGAG
GCCGACGGTGTGTACCACGGACTCCTGACGGCCACGGGTTGTTGTCTCATGCACAACATCACGCTGTACACGCGTTTTCA
CACTGACCTGTTCCACTACGAGGACTGGGACATGACGTCCATCGCTAGCTACCGACGGGCATTCCATCACCTGGCCGAAG
GGATTAAGTTCCTGAACCACACCTGCAAAGTGTGTCATTTTGACATCTCCCCCATGAATGTGCTCATCAATGTCAACCCT
CACTACCCCaGCGAGATCCTGCACGCTGTTCTCTGCGACTACAGTCTCAGCGAACCCTACCCCGTGTGCAACGACCGCTG
CGTGGTGGTGTTTCAGGAAACCGGCACGGCTCGTCGCATCCCCAACTGCACGCACCGCGTTCGCGAGTGTTACCACCCCG
CTTTCCGGCCCCTGCCGCTGCAGAAACTCATGATTTCCAATCCCCACGCCCGTTTCCCCGACAGCAGTCTGCTACGGCGT
TTCTGCATGGCCGAGCTGGCCGCTCTGGGGAATGTGCTGGGGTTCTGCTTGGTACGACTGCTGGACCGTCGGGGCATCCG
AGAGGTCCAGGGCGGTCCCGAGGCCATGCTGTTCAAGCACGCCTCCACCGCGTGCCGCGCGTTGGAGAGCAACCGCCTGA
CGGCCTGCTCGGACGCGTGCATGCTCATCATGGCCGCGCAACTGGCCTACCATGCCGGTCTCCTGGGGGACGAGGGCCTG
CCCATCGTGAACAGGACGGTGAATTTTGTGGAGAACAAGCTGTCCTCCTGCCGGGTTCGCGCCTTTCGACGCTACTATCA
CGAATGTCTTGGCAACCTGCCACATGATTACGTCAGAAAGAATCTGGAGCGGCTACTGGCCGTGCCAGACGGTGTGTATT
TATACACTTCGTTTCGGCGCACCGCGAACATTAGTGCCGAGGAGGACCTGGACGGTGATTGTCGGCAGGTTTTCGTGGAC
TGAGTCCCCGGAGAGGACCGTCGAGGGATACGCCACGGCGGTCCGCTCGCAGCAGCAGCAGGCATGTGGGACGAGACGGG
AGATGCGAACGAGGACCTAGCGCGCTTGCTAGCTTTGCGGGACGATGAGCCCCTCAGCTTGTTTATCATGAACACGTTTC
TGCTCAGGCAGGAGGGGTTTCGCAATCTGCCCTATTCTATTCTGCGCCTGGCCTATGCTTATCGCTTGTTTACCAGAGCG
GGTCGCACCCATGGGGTGGCGGtGGCCGAGGAGTTTGTGTCCGCGGTGGCCGCTCTGGCTCGAGACGAGCTGCTTCGCGA
CATTCTCGGCAATCGGTACGCGGAGTCGCGCGAAGAGATGCACACGGCGctggaccgtctggtccggTGTCATGTGGATC
GAGACAATGATCAGGAGGAGATTGACGTGTCCGAAGACTACGTGTGGCTCAGTAAGCTGCTCGACCTGGCTCCTAACTAC
CGTCAGTTGGAACTGTACCAACTGCTGGAGAAGGAGTCGCGCGGCCAGTCCCGCAACGTTGTTTGGCACGTGCTGCGCTT
GGABTCGGTGTCGGCCACCARGATTTACGATGCTTTTGTATCTGGCTTTCCGCCAAGCCCCGTGCACCAGCCCTCCAGGG
CCGGGGCCTCGCCCGGCATCCAGTTTGGCCTGCAGCACGAAGGTGTGGTGAAAACCATGGTCGAGTTTCATGTGACGCAC
GGGCGAGAGCCGGTACGAGACGGACTGGGACTCCTCATAGATCCTACCTCTGGCCTCATTGGCGCCTCTATCGACCTGTG
CTTCGGAGTTCGCAAATCTAGCAGCGGGGACGCGCGCTCAAGCCTGCGCATCTCCCCCTGCGCCAGTGTGTATGAAATCA
AGTGTCGGTACAAGTACCTCCGCAGCAAAACGGACCCATTTGTGCAGAGCGTGCTCCAGCGCGCGGACGCGGCGGCCGTG
GCCAAGCTGCTGCTCTCCCACCCCGTGCCCGGTGTGGAGTTTCGCCACGATAGCGAAATTCCATCGGCGCGAGAGTTCCT
CGTGTCCCACGACCCCCTGTTCAAGGCGACACTGAAACGCGGTCGTCCGATTAAGCCTCCGGACGCTTTGCGCGACTACA
TCCACGATTTACTCTACCTGAATAAAACCGAACGCTCGGAAGTTATCGTTTTTGACGCCAAAAATGTGGATGACACTATT
AACGCTTCGCTTTCCCCTACGGGCCCGGTGGCGGAGGACGCCGTGGAAACCGGCAGCGACTCGCAGTCGAGTCTGCACGA
CGAAAACACCCCAGATCTGCTGGGGAAGCTCACCATTTACGAGGTGTCGCGCTTTTCGCTGCCCGCTTTCGTCAATCCGC
GCCACCAATACTACTTTCAGCTTTTGATTCAGCAGTATGTACTGAGCCAATACTATATAAAGAATCATCCGGATCCAGAG
CGCATCGATTTTCGAGATCTGCCTAGCGCGCACCTGGTCTCGGCCATCTTTCGCGAGCGCGACGAGAGCGAAGTGGGGCG
GCCGCTGCATCTGGGCGGTCGGGTTTTTCACTGCGATCACATTCCTCTGTTGCTCATCGTCACTCCGGTGGTTTTCGATC
CCAATTTcACGAAACATGTGGTTTGTACTGTTGTAGATCGCTGGAGCCGTAATCTGTCCCGAAAGACCAATCTGCCGGTA
TGGGTGCCCAATGCTGCAAACGAATATGTTGTTTCTTCGGTTCCCCGACCCGTGAGCCCTTGAGGGACGCCATGGGCAGA
CCGGTGTGTCTCAGATCTTTTGCCGATCTGGACAACACGTCGTCGGAAGATGAAGAAAACCTGGACCTCCTTAGTATCAC
ATCGGAGGTGTCGCTGACTGCTCAGGATCCTGTCGTGGCTGAGCCCAGACCCAAAAAGAAAAAAAAGAAAAAAGACTACA
AGGAAAGGCAGCTGACATTAGCTCTCATGCAGGATGAGGGTGGTTACCATTCCCCTGCCAGGGCCACGAGAACTATGCGT
CCGAAAACGTCGTCGAATGCTTCCTCCTCCTCATCATCAAGGCTTTTGTCGCGCCCCCCAACTCCCAAGCATCCCCGAAC
CCGTCGTATATCTCCTACAGACGCTAGTGTGTATTAAGCGGCCGTGTTCCACCACCACCGTGCTATGAAAATAATAATAG
AACAGAACAAAGATTAGGACCTTACAACTCGTTTCATTTGTATCTCATGTGTGTTTAATAAAACTATGTTACAATTGTAC
AGACGTGTGCGCGCGTTCGTTTGTCCATAGCATATATGCATGCGGTGTTGGCCTCAAGCCATCTCGTCCGTGTCGTCCGC
GTCCGAGGTCACTAAACCGTCTTTCAAAGATGAGATCTCCTCCGAGGTAGTCAGACTCTTGTATTTGATGGTGTGGCGGG
CGCGGAAGTAGCGGATGCAACGACAGGCTGTGAAGCCGGCCCATATGATGAAGCACAAAGCAAAAGCGGTGGCCACTTGT
GAGGTGTAGTCCGGTGTGTACAAGAGGTCGTACTTGATAATAGGGTAGATCATGCCGCACAGACCAAAGAAGGTCCCGGC
GTGATACCCGAACTGTACTTTCACGTAGCGCCAGAACACCACCTCCAACATCAGAAAGTAAATAATGACGATGAACGCGT
ACACGAAGAACACGGCTATGACCATGTGTCCGGTCTGGACAAAAAAATTGTTACCAAAGCCCAGACACAGGGACATGGCC
ATGACCGCCGTGTTAAAGCCTAGCAGCATTTGAACCAGGTTCACGATCAGGGTGCGGTACTGCAGGGTACCATGCAGCTT
GGGGTGAATTCGCTGCAGCTGAAACATGCTGTTCCTGTGGCTCTGGTAGTTAGTGACCATGTTGATGTTGAAGAAGATCG
TCAGGCAGAAAAAATGCACGCAGGTCATGAAGGCGATCACGCTCGGGAGGCGGAAGGTCATGGCCAAGGTCATCAATTCG
TAGGTGTCCATCACCAGGATGAACACGAAGATTGACATGCTGTCCCCCATGTACGCAATGTCTCGCGTCCGCTGGTTGAG
GCTCAACCCCTCCTCCTTGCGGTACGTGATTTTCAGCCAGCAACACACATAGTAGGCGATCACACACAAAAACACAATCA
TGCTAAAAAACACGTACAAGATGGTCTGGACGGCGTCCATGTACAGCATGGGCGTGAAGAGGTGCATCACATTGAACGTA
GACAGGTTCGTCTCCGTGTGATCCACGACGTTGTAGTACGAGCACGGGTAGCCCACTCCGGGGACGTTCATCATCGTCGT
GTGTCCGCAGATGTTGGCGAAGGTCAGGAACGCCAGGAACACGGAGATGCCCCAGATGCGGCTGTTCACCGAGTCCACTT

TGGAGGGCGCCATGGTACCGCCGCGGTGGACGTCGAGGCGGTGCCAGGCGGGGTGCGCGCTGATGACGAAAAAACCGATG
AGTCGTTAAATACTATGGTGCGTGGTGTTCTGCGGGGCTTTTCGCACGGCCAGGTCCGCGAACGCAAAACCCCTGTGGCC
GTTTGCGGACGGTGGCGCACGGCGTTCGCGCTCTCGGGCTCCGGATTCGGGACTCGCGGGCCAGGCTCGCCGCACGAGAT
GTCAGCCGGCACGACCCATGTGGGTGCCAGCTGTCATCTCAGCTGGTACGGGATCCTCGAGTCCGCCGTGCCCGTGGTGC
AGTGCCTATTTATACACTTGGCCCGTGGCCACAGCGAGCCACGACTCCAAACCTTTATCGTCAGAGCCGATCCCCTGCCG
CCTGCGGATTTGCGGGCTGTTCATCGAGGGAACTATGCTAATTTGGCCGCCACCACCGACGGCGACGAGCGCCAGCAGAG
TTTGGACCGCCGCAGCGCGGTCTTGGCGCGCGTGCTGCTGGAGGGCAGCGCGGTCATCCGCGTGTTGGCACATCCGTTCA
GCCCCGTGCAGATTCGCACCGATCACCACGGTCTAGAGATTGTGGAAGCGGTGCCCGCGCTGGACGTGGACTCCGCCGCC
CTGTCGAACGCACTCAGCCTCTTTCATGTGGCCAAACTGGTGGTCATAGGCTCCTACCCTGAACTGCCCGAACCCCGTGC
CGTTCCGGTGGGCGAGGAGCGGGTGAGCGATGAATACGGAACGCATGCCAACAAAAAGATGCGCAGGGGATACTACGCTT
TCGATGTCGCTATGTCGTTCAGGGTGGGGTCCCACAAGTATGTGCTGGAGCAGGACGACGAGGCCGTGGTCTCGCGCCTG
TTTGAAATCCGGGAGGTCTCCTTCCTCCGCACCTGTTTGCGGCTGGTGACCCCCGTGGGATTCGTGGCGGTGGCCGTCAC
CGACGAACAGTGCTGCCTGTTACTGAAATCCGCCTGGACTCGACTGTACGAGCTCTTGTTTCGGGGCTTTGCGGGGCAGG
CGCCGCTGCGGGACTACCTGGGGCCCGACCTTTTCGAGGCCGGGGGCACGCGCTCCTTCTTTTTTCCCGGCTTCCCTCCC
GTCCCCGTGTACTCTGTCAATAGCCTGGCCGTGTTGTTGCGGGAGACCGCCATCGATGCCGCCGGGGAAATCATGTCGTG
GTGTGGGCTGCCAGACATCGTGGGGGCTGCGGGCAAaGTTGGAAGTGCAGCCGAGCGCGCTGTCGCTCGGCGTGCCCGAC
GATGAGTGRCGGATCTTTGGCGRGGAGCTCACCCTGCCCATGCGACTCAACGGCACCGCGTTTCGCGAAGCCAGCGACTC
GCTGTGTACCCAGCTGTGGATGCGTGCGGGGGACGAAGAGGAGCTCACGGCCCACGTGCACGCCAGGGACTCCACCCTGC
ACCACCTGGTGATCGTGGACCTGGTGGAGTGTGTGCTGAGCAAGTGTATTCGAACCCGGGACTTTAACCCCCACCTGCGA
TTCAGTAACAGACTGCACGCCAAGAATGTGTGTCAACAGTTTATTGAGAATCTGCGGTATCGCTCGGGTTGTGCCTTTAG
GCAGATTCAGAGTCTGCTGGGTTACATCTCGGAGCATGTCACCTCCGCTTGCGCGTCCGCGGGCTTGTCGTGGATTCTGT
CGCGAGGTCACTGTGAGTTCTACGTCTACGACGGCAATCTGCCCCACGGGCCGGTCTCGGCGGAGGCGTGCGTGCGGACC
GTCATCGAGTGCTACTGGCGCAAGCTGTTTGGCGATGCGCCAGGTCCCACTTGCCGCTTGCAGGCCAGCGCTCCCGGAAT
CATGCTGGTGTGGGGTGGCGAGCGCGTGGTCAGTCCTTTCGGGTCTCCGGGGGATGATGACTGGTGGTCTGTGGTGGGGA
AGGTGCTGGGGCGCGTGCCCGCTCTGCTGCGGGAGCGGCCGACCGCTGCCGATCTAGAGGCCGTGTACCGTGAAGTGCTC
TTCCGCTTCGTCTCTCGCCGCAACGATGTGGAGTTCTGGTTACTGCAGTGCCAGCCCGGGGAACACGAGGTGCGCCCTCA
CGCCGGAGTCATCGACTGCGCTCCTTTTTACGGAGTGTGGTCCGAACAGGGCCAGATTATCGTGCAGTCTCGAGACTCGG
CGCTGTCCGCAGAGGTCGGTTACGGCGTGTACCTCAAGAGAGCGTTTGCTATGGTGACCGCGTGCATCGAGGCTATGTGT
TTAGAGGCCTCGCCCGCTGTGCCTCCTTGTTCTTCGACGGGCTCTTCGTGCTCGTCGTCCTCTTCCTCGGCCTGGCCGGG
CTCGGCTACCACAACGACGACAACAACAACGGACGAACGCGAGGATTGGACCAAGGCGCTGGGCGGGGAGGCGCGCCAGG
CGGCAGAACGCGTTGCCGCAGACGGGTTGCGGTTTTTCCGCCTCAATGCCTGATACCTAATGCTGCATGATTGATTGACA
GGTGGCTTAAAGCTTTTCAGAACGAGGGTTGTCTGAAAGGCAGGCTGGGAAAGTTCGAGTTGAGTACCTGTACGATGGCC
GTTTCGACCAGGGTCCCGTAAAGAGTTTTGCGTTCAATGAGTCGCATGGAAAATTTGTACATGTCCTCCTGATTTGTAAT
GCGGAACCAAAACCGCATCACAGTTGGAATGAGGCCGTGTCGGCTCTCCTGGAATAAAACAAAGGCACTGGGAACATTCT
CCACAATCTCTGGTGTTCTGTCGTCAAAAAACAGGTACAGGGTGGTGCAGATGAGTTTTTTACAACTACTGCGCAGTAAC
TTTGCCAAACGAGTCGCGGGACGTTGGTCTTTTTCAGATATTCCAGCAGGTCGCCGAGGCACTTGCGGGTGAGGGTAAGG
TTTTCCACACAAGAGAGTATAGTACAGAGCGTAAAGTGCATGAGGTTGTGGGGCATTACAATCCCGCAACGATGTTTCAG
TCGCATAGACGACAGTTTCACATGGGACATGACATCCTCGTTTTCCAAAAGAAACACTGTAGAGTAGCCCAGAAACGAGG
GTTTCAGTTGCAACTCACTGCGGTATTCCGCGATAGAGCACAGCATGCCGTGGTCGGTGTGCACCCATAGCCGGTCCCCC
TTGACGGTCAGGTCCACCACCTCTGGGAACGTGATGATCGCGTTGCTGCCAGTGGCGTCAGTATGGACTTCCAACACCCC
GCGGATCATGAGGATTTCCATGACGGAATCGCAGCGGACCTCCCGCCGCGTCGGGATGTGGATTTGGACGTTCAGCGTTT
GGAGGCCGGTGCGCGACTGGCTGGCTAGTCATCTGTCGGCGTCTTGCCTTTTCGCGATTGCAGATGAGTTCGACGTACTG
TCGAAACTTGCTGTCGTCATACGCGGCGGCAACGATTTCCCCAAAAGACAGGTTGAGATAGTCTTCGGGAATGGGTTCGA
GGGTGGACAGGGAGAGAACGCTAAACAAGTAGGGCAGGGTCACGCGTTCCACTGTGTAGTTGGAGTACACGATGGCTTCC
TCGGCGCCCTGTGCGGTGACGATGCGTTTCATTCGGAACGTGCGGAAGTATTCGTTTTCCCACAGCTCGGTCAGCAGGCT
TTCGAGCTTTTCCGTATTGGGGATAAACTGCGAGAAGAAGCTGTTGGCCACCACGCGGTTATCTTCCACCGCCAGGGGCG
TGAAGGACACCTCCTGCGCTTTGCGAATCGCTTCGGTGACTGACAGGTGGTACAGCTTGGGGTCGCGTTCGAAGGGGCGC
CCAGAGCGTTGGCGCGCGTACCGGGTCAGCAGGTCGCGCAGTTCGCGGACCCGATGTTCCCACGTCTGATTGAGGCTTTT
CAGGTCTTGGATCTCGTCCACCTGGGCTTGGATCTGTTCTTCCAGGCATTTTATCATCTGCTTCTTAAACAGATCACGAA
TATCACGGTCTGGCGCGGCGCCGGCCGCCGAAATTCCTCCGGGGGCGGACGGTGTGCTGCCGGATGCGCGGGAGGGCATC
AGGCGGCTCTGGTCCACCAGCGAAGGGGTGAGGTCCTGCAGGAAGGATTCCACGCTGTCCTCGAGCCGATGCGCGATTTG
CTGTCGGACACGTTCAGGAGGAATTTCATGATGGCCTTTTTGGCGTCGCTGCCGCGGTCGTGCTTGTCCATCATCTCCAC
GATCTTCTTGCAGTTCATCTCGTGGCGGCTGGTGGTGACGGCCTTCACGGGGTACGTGTTGAGTAACTGACAGATCTTCT
GATGGCGGCACAGGCTGTCGTAGCGCAGGATCTCTTCGTGCAGGTGGGCGATCGGGGTGGTGAAAAACAGCTGGTCGCGC
TCGTAGGCCAGGGGCTCGGTCACCACGTAGCTTCCGGTGGCGCGTTCCTTCAGCTGCGCCTCGATCTTCTCCGACTTGAT
CCTCTTGAAGATGCGCAGCTCAGGTGAGTTGACGATGCGCACCGCGCACAGGTGCGTGACCACTTGATGCAGCAGCGGGA
GATTGGCGTTGAGCAAGTCTTGGGCCAGCGGATGAGCCTCCGTGGATTTGTGGATGGCGGCGCGCACGAAGGCCGCGCGC
TGCGCGAGGTCCGCGGGCTTGGACTTGACGTCCATCAGAGGGATGATGCCGACGGTGACGCACCAGTCCACGTAGAGGCC
ATACTCGTCGTTATCGGCGAACTGGTATAAAATGTCACGCAGCGCCCCGAGCACTCCCGTCTGCACGCTCTGGCGCACCG
AGGCGCTCCACACCGCCAGATAGTTGTCCAGATCCTCGTCGTCCAGGGCCCTGTACGGGAAGAGAGCCGAGTGGATCTTC
CATTCGTCGGAGACCTGTTTAATGGTAATGGTGTCAAAAATGGTTTTGCACAGCCCGTAGAGCAGTTGTTTGCGCAGCAC
GCAGGGGTCCCGCAGGACTTGGTGGATACTCTTGCCGCGGCACGTGCCGGTAAAGCCGTGCAGCAGGCGGATGTAGCCCA
GGGTCTGGGTTGTGGGGAAGATGTCGATGACCGTGTCATCGTCGCGGCCGGTCAAGTACGAGGACGCCTTAATTCTCAAC
CTGTCGTCCGCCGCTAAGATCGAGCGGATCGTCGACAAGGTCAAGTCCTTATCCAACGTGCGCTTCGAGCCGGAGGATTT
CTCGTACCAGTGGTTCAAATCCATAAGTCGCGTGGAGCGCTCCGCGGACCAGGACGTCCGCGTGCCCTGTACGGATCAGT
GGCCCTTCTTCCCCTTTCGGGTGCTGCTGGTGACTGGCACGGCCGGCGCCGGCAAGACGTCCAGCATTCAGGTCCTGGCT
GCCAATCTCAATTGCGTAATCACCGGCACCACGGTGATCGCCGCGCAGAACCTTAGCGCGATCCTCAACCGCACGCGCTC
GGCGCAGGTCAAGACCATCTACCGCGTGTTCGGCTTCAACAGCAAGCACGTGCCCCTGGCCGAGGGCGGGAACGACACCT
TAAAGCAGTACCGGATATGCGAGCCGCGGGACGAGACCACCATCCAGCGCGTGCAAGTCAACGATCTCTTGGCCTACTGG
CCCGTGATCGCCGACATTGTGGAAAAACACTTAAATATGTTGGAGCGCAAGGCGTCCTTGGAGGACGTTTCCGAGATGTG

CGAGAGCAACATCATCGTGATTGACGAGTGCGGCCTGTTGCTCAAGCACATGCTGCAGGTGGTGGTGTTTTTCTATTACT
TCTATAACGCCGTGTGCGACACGCGGCTTTACCGCGAGCGGCGCGTGCCCTGCATCATCTGTGTGGGCTCGCCTACGCAG
ACGGAGGCGCTCGAGAGCCGGTACGATCACACCACGCAGAATAAGAACATTCGCAAGGGAATCGACGTGCTCTCGGCCCT
CATCCAGAACCAGGTGCTCATGGAGTACTGTGAGATAGCTGACAACTGGGTCATGTTTATCAATAACAAGCGCTGCACTG
ACTTGGACTTTAGCGATCTGCTCAAGTACATGGAGTTTGGTATTCCTCTGACCGAGGAGCACATCACCTACGTGGATCGC
TTTGTGAAACCTCCGAGTTCCATTCGAAACCCGTCCTACGCTTCCGATATGACCCGCCTGTTTCTCTCGCACGTCGAGGT
GCAGGCGTATTTCAAAAGGCTGCACGAGCAGATCCGGCAGAGCGAGAGGCACCGCCTGTTCGATTTGCCCGTGTATTGCG
TGGTCAACAACAGGGCGTACCAGGAGCTCTGCGACCTCAGCGACCCACTCAGTGGGTCGCCGCAGCCGGTGGACATCTGG
TTCCGCCAGAACCTGACGCGAATCATTAATTACTCGCAGTTTGTAGATCACAACCMTGTCTAACGAAATACTCAAAGAGC
CGTTGTACNGGGGAAGACATCACCGATGAAGCATCGGAAACTTTATTAACACTGCGAATCACGTACATCCAGGGTAGTTC
GGTGGGGGTGAACTCTAAGGTCAGGGCGTGTGTGGTGGGATACCACGGCACGTTCGAGCAGTTTGTGGAAATTTTGCAGA
AGGACACCTTCATCGAAAGGACCCCTTGCGAGCAGGCCGGGTACGCGTACTCGCTGGTCTCCGGCCTGCTCTTCTCCGCT
ATGTACCTCTTTTACTCCTCGCCGTACACCGACGAGGACTCCCTGCGTGAGTTGGCCCGAATCGAGTTGCCCCACATAGC
TTCCCTCTGTGTGGATGACAATAATAAAGGTGTGGATGGTGGTGTGGATGGTGAGGATGAGGGGGTAGCAGCAGCAGCGG
CGGCGACGGTCGTGTCGGGGTCTGGGTTGGGGCCCACGCGGGTGGCGGTTACTGGGTACGAGGGGGACGATGACCATTCC
CTGATTCTGGGGATAGATGAGGATACCTACGACATTCCCTGTATGGCCGAAGAGGTTTCGGACTCGGACCTGCTGGCGCA
AACCAGCTTGTACGAGGACCCCTTCTTTGCCAAGTACGTAAAACCCCCGTCCATCAGCCTGCTGTCTTTCGAGGACACCG
TTCAGATTTACACGACCTTTCGCGAAATCTTTCTGAAACGGTACCAGGTGATGCAGCGGGCGACCCACGGCCGCTTCGCC
GCCTTGCCCATGATCACCTACAACCGTCGCAATGTCGTCGTCAAGTCCAACTGTCAGATCGCCTCGCAGACGGGCTCTTT
CGTGGGGATGCTCTCGCACGTgTCGCCGGCACAGACGTACACGCTCGAGGGCTACACCACAGACAACATCCTGAGCCTGC
CCGGCGGACACCGCATCCACCCCGAGGTGGCGCAGCGCGGTCTGTCGCGGCTAGTGCTCAAGGACGCGCTGGGCTTTCTG
TTCGTGTTGGACGTGAACGTGTCCCGCTTCGTGGAGTCCACACAAGGTAAGAGCTTACACGTGTGCACCACGGTGGACTA
TGGCATCACGTCGCGGACCGCCATGACGATCGCCAAGAGTCAGGGGCTGTCGCTGGAGAAGGTGGCGGTGGACTTTGGCG
ACCaCCCCAAGAACCTGAAGATGAGCCACATCTACGTGGCCATGTCGCGGGTGACGGATCCCGAGTACCTGGTGACGAAC
GTGAACCCTCTGCGCCTGCCGTACGAGAAGAACACGACGATCACGCCGTACATCTGCCGCGCGCTCAAGGACCGGCGGAC
CATGCTGATCTTCTAGAGGAGGAAACACAAAAAACAGGGGAAAACGAGGGGGAGCCACCCTGCCTGGCGTTAGGGACCAG
AACCTGGCGGGCGGCTCCGCTGGGTCTCATTCTGCTCCACATAATGATTTCAGTATGCACTGGTAATAATACACATCCGT
GATCCCGTGTGGAAGAGAGAGGATAATAAAACATGTGTTGAGAAAAATTGCCAACGGCTATGTGATGTGTCTCTCACGTA
GCAGAATAGACCTTTTCCGAATCAGTGTTGACTATCCACAAGAGAACGCTCTTAAAAGCTGAGCTGAGCTGTTATAAGGG
ATGTGGTGATACATTCATAAGTATTTCCTTAACATACAACAAATAGTTCTGTATTCGTTTTCTAAATGAGAGGAGAGAAG
TATGAGGTGTTTATTTTTCTAACATTCCTTAATAACAATACAGTTAGGAAGCGTAGCGTTTTCAGCGTGCATCCTTTGTC
TCTATGCAACTGAAACTGTTCGATAGGCAGAGACGAAGCGGTCTAAAAAAAGCTCGTTCAGATCGTTCGTGGCCTGAAAG
TGGACGCGCCGACGCGCGCCGGCGTATAGCGATTTGTGTATATAGCGTGCGATCTGCTAATCATGTAGACCAATCGACAT
GGATATCCAGCTTTTGATACGGGTATTTTAACAGTTGTTATGGAAGCTTTCGAACTTCATGATTGAGGATAAAGGGGCTA
GCAAAGGAAACTAAGTTTTCTCCTAAGAGCAATTTTCTTAAGAGCTATAACATCATGGTAAATTCAGAATCCTATGGTGT
CTTTCTTTAAGAGAAAAAAATCAACATTACGGTCTAACGTTCGTTTGATGATCTTTAGGTCATCACTGGCGACCAACACC
TACAGTTGTTTGTTTTGTACGAGTCCTAAAGTACGATATTGTTTGTTCTTATGAGTAACTAACTTATCTGCTATGTTTTG
TTTTTCTTTTTTTTTATTTACTTTACACGTGAGGATTGTTACCCGTCTGCGGCGGCAGACGACTCTGGGCCAGATGACAG
AGAACTCACATGAGCTTGGCTTGCATGACAAAGACATCATCTCCCTCGAACGGGGATGGCGCAACTTACTTAAAATGGGC
TTATGAAATATTACAAATTACAACGAAAAGAAAACGAGAAGGAGAAAAGATGGTAGAAAAGGAAGCTTAACTGAGACGAC
TTTATCCCTAATCTTCTACTATCACTTACTTAACCTACGAGATTACCGATTACTTGTTATGGCAACAAGAGACAAAATAG
TAGTCATAGAATTAAAAAAAAAACAGCTTAACCACACTGACCAGACCATCTGGAAAAAAAATGTCGTGGCGGCCAAAGAA
GAAGAACTGAGTGACGATGGCCGGAGAACAACCTCTCGGTCTCTTTCAGGACTGCGGCCTAAACGTTTGTGATGATGATG
TAGTCGTAATAGTCTTCAGATATAGATTTAAGTCATTCATGGTTGTGTTAAACTAGATGATACAGATGAATGTACACAGC
TAGGTGTGAGACTTACAGCTATCCATCCATTCTTTATTGTGATTAGCTGAGAGATTATAGACGCGTAACTATTTATCGCT
ATGGTATGTGCACTATGGAGTAGCCCTATAGGGTTGGGGAGTAGGGTGTTGTTTATCGCTATGGCTCTTTTATCACTACG
GTTGTCGCTAATGATTTCTAGGGAGCGGAGACACAAGAACTTAAACTCTATACCTATATGATAGCTTAGCGGATAACCCT
TATATAGATTGGTTGTCACTATACCTTATTGTCCTATAGTCACTATGAGAATATCAAACACGAAACAAAAACGAAGATAC
GTGATGAAGATAAACGAAAGGAAACTTACGGCAAAAACATGTACTTACGTTCTATATGTACCTTATCACTATGTGTACTT
ATGGAATATTGTGTTATGAGGCTTTTCTATCGATTTATCTAAAGGTAGTATCAATAAAAAGAAGAAAAAAAAAAGTTCGT
TCACGCGGACTTGTCTACAGCTTGGCGGTTACGGCTTGCGGTTGCAGAGCATCGTCGACAACGGCGCAAGAGAGAGTTCG
TCGCGGACCGGCCGCGAGGCGCGTACAAGCGTGGCGGTTGCTAGAAATAACTCTGTGCATTGGCTCTTGTGGGTGAAGAC
GCCCGACAGTGTGCCCGGGACTCCTCGATCGCGAACCCGCGCTCACCGGAGAGTCCGGTCGTGGGGGACCCTGTGGTTGT
TCGTTTTGATGCTTGATTAGATATCCTGTTGGTGGGGGTGGGGGGCAGGCTGTTGGTTGAGTAGTTTTACGTTTTCGAAA
GGTTAGCCTTATTTTTCTTCTACCCCTATCACATGCGTACTTGTTTCTTGAGTTAAGTGTTTGTTTTTTTACACAGTTAT
CGTTTTTACAAGTCTGGCTTGATTGTTCAGGGGGGGTTGCATCTTTAAGTGAAAGTGTGTGAGTGACACCACTCTCTTTA
TTCTTTTTTCTTTAGTCTTCTCTATTCATCTGAGCGGTGTCTCGTTCGTGTAGTGTAAAATAGCGTAACTGTAATATGGA
TATCGGTTGGTTAGGAGAACTAAATGTGAATTAGGTATTATTCTCTTATGCTGCTAACAGAATTGCTTCTCCGTAACTAT
TATCGTCATACAGATAGATTGCGTTCGTGTTTTCAAGTTCCCCAGCAGAAACAGGGAGTCTGTGGCTTTTTGGTTCGTGT
ACATCCGTGTTTATTCGCGTGTCAAATCTGATCTTCCCGCGATATGATGTAGGGTCCTCGATGTAGGATCTCGAGTATCT
CTTTTTTAGCARAGTGAGGGTTCAATCTGAGTCTGTGTACAAACGTACGTGAAGTTTGTGACGTTAGATCGAGAGCGGCT
CGAACCTTCTGTAGAGCTTTATTTAGTGCAACTTTACAAAAGAGTTAGAAGCTAAACGAATCTCTGAAGGTGCTACCCAT
TTACACTTCTTCTTCGAAGAAACATCCAGTCCGTAAAAAAATAAACAAGTGTCTTCAAATTACATTTCCACTATTTTTTG
CAATACATCTGTGAAAGTAGGCAGTAGATACCAGATTCTTTTCGAGTTTTTAAAAATGGACGACGATGTTCTTTTTTCTC
AAGAGTATGTCGTAAGCTTGTCCGTGGTGTAGATCGAGTTGCTATTGTTATTTCCTTTTGCATACACAGTTTTTTTTCAG
TAAGATGTTTGACCGTGATTTTGCCCACCCGGCCCGGGAGACAAGGCAAGGAGATATTTTTGCGTTTGTTTGCCATTTAT
CATCGTTATAATGTGTCGTAGCTATTTAGTGGTGACAGTGACCTGTCAAATTTGTTATCTGTTATCAATAATGTATCGGT
TATACAGACTAATAAATCTTCAAGGATTAGGATTGAGGAAAAGCAAACAAAAAGAAAGATCATAGGGGACCACCTTACCT
GTGACCGAAAACCATTTTCTGTTGCAGCTCCGCAGCGCGGCTCGGCAGTCGAGGATGGGGTCTCGACCGGCGCGGATGTG
CGGTGTTTCCAATCTGCTGTGCCTTCTCCTGGTGGTGCTGGTGGCGGTGGTCATCGATCGGGGCTGTGGGGCTTCCAAGC

CCTCCAAGCCCCCGGTAGACTGCGATCCCATCCATGGCACTCTGTCGCGCATTATCAAAGAAGTGCGGACTGGATACGGT
TCCATTAAGCACGCGTTGGTAGGTTTCGGGACGTTTCTCTCATGTCACACGATTCATTTCGTGGAAATATAACGTGCTGT
CCTGTTTTTTTCCCTTTGCAGCAAAGCAAGGACACCGTGTACTATGTTTCATTGTTCCATGAAAACTTGTTAAATGAAAT
GCTGAGTCCAGTTGGCTGTCGCGTGACTAATGAGCTTATGCAGCATTATCTGGATGGCGTGCTGCCACGGGCATTTCAGT
GCGGTTACGATAACACGACCTTGGATGGCCTGCATTCTTTGGTTTCCTCTTTGGACGCTCTGTATAAACACATGTTGAAA
TGCGTAAGTCATTATTTTCTCGGAAAGTTGGCGGCAGAATGTGTTGACAGTCAAAATTTCCTTGGGAAATTGGCAGTCAA
ACGTTTTGGCAGCCAGTGTTTCCTTGAAAAGTTGGCAGTACAACTTGTTGACATCCTCTAGAACTTGGCAGGCGACGTTT
CCTTGAAAAGTTGGCAGTACAACTTGTTGACATCCTCTAGAACTTGGCAGGCGATGTTTCCGTTGGAATTTGGCAAGCCG
TACGAGCCTTCAAAATGCGATTTAAGCTCCGCAGATGTCAGCAGAAACCTTTCAAAACCGTAAAAACGGCTCCAATCGCA
GCATGTCGCGATATATCGCTGCTAGCCGTTCCAAAACGGCAGAGAAATCGAAATCAGGCTTCAGCTTGGCTCCAATCGCC
GTATATCGCCATATATCGCGCAATTGATTCGACTATCGCGCGATATAGGCACGATATCGGCACGATACGATCGCATATCG
CCAGATGTCGGCCCATATCGCCCAATATCGCCATATATCGCGTCATGTCGCCGACATAATAGCCATGTCGCGACTATCGC
CTCATGTCGCCACTATCGTGCCATATCGCCACTATCGCGACATGACCTCTAGTTGACCCCATGCCCGATACGACCCCTCC
GTGGGCCTCTGCGTTTTAGAAACTGCGTGGTTCCATTTCACCGTTTGCTTAAACGGTGTTTGTTGTCTCTTGGCAGCCTG
CCCTGGCGTGTACCGGGCAAACGCCGGCGTGGACGCAGTTTCTGGAAACGGAGCACAAGGTCAGTTGAGGTCGCCGTGTT
CCCCAGCATTGGCCGCGTGCCATTTCGGTTTAGTAAAACGCTGAACGTTATTTTTGCAGCTGGATCCGTGGAAGGGCACG
ATTAAGGCCACGGCGGAGATGGACCTGCTCGTCAACTATTTGGAGACCTTCCTCGCGCAGTCCTGAGGGAGCTGTCACTC
AAACTAACAACGTGCCGGTTAATTCAACGTGCTCCGGTTTACAACTCTCAGCCTGCCTGGCTGGTTCGGCTTTAATAAAC
GTACTGTTTAACCACGTTGCGTCGTGACGTTGTTTATGGGTGTTTCAGGGCGGGAGCGGTTCATCATGTATAAATAGAAT
GAGCGAGCCGCGGCAGCTCTCTGAAGCTGCCTACGCGCGGACAAGGGCTGCCACAGCGCTCGGGAGCGGTGCTGCAAGAT
GGATAACAACATGCTGACGCGAAAATACTGGACTTTCTATGGCGTTACGCGCGCGCTGCATCAGAGCGTGAATACGGATT
TCGACGTGCAGCAGTTCTGCTTCGAGAGCGCGCGTCTCGTCAAGTGTATAGACGGCGATGGCACCACCACGTCGTTCGGG
AAGGGCTGGTTGTGCGTGACCGTTATTCAGAATGGGGGAACGAGCGGCGAGAAGCCGCAGCAGAGGCCGGGCTGTCTCTC
TTTGGACGTTACCATCGAAAACACGGTGGAAACCCTGTCTGGTCGAGGCATGGTCTTGAAAAACAAGAACGTTTCGTCCG
TGGTGGGCAATTCGGGACCCCATGACAGCTCCTTGCTAACGGTTTTGATTGAGGGCAGCAACATCCAAGTGACTCAAGTC
AAGCACTGCGTTAGGTCTGTGGACGCGGGTTCACATGCGTCGTCCTCTGGGGCGGCTTCGTCGTCTTCTATTGCTGCCGC
TGCCTCCTCCGCAAGTTCGGAGGAGAAGCGGGCGCGTAGCGAGCATCGCCGCGAACACAAAAAGCAGATGCTGCAGCAAC
AGCATCAGCAGGCTAACTCTATGGTAGATGGCAGCAGCAGCGGCGGCGGCGGCAGCAATGGACAGGTGCGGGTTTCGGGG
GGACCTGACGGGCGCGATCCACGCATGCTGAATCGCCCCAAAGAACGGCGACCAGACGCGGACGGTAAGTCGATCCGTGT
CGCGGCCCCTGTGACCGCGGAGTGTGCTGGTGGTGTGGATGATGGTGATGATGATGATGACGCTTTGGGTGTGTGGCAAC
TCGCAGGCGGAACACATCATTAATTTCCATTTTCGTGCGGTGTCCGTCTAGGCTCTCCCTTAGACAAAAGGCAAAAGACA
CACCACGATCACTCGAAGAATGATTCGGAGAATCCGCCGCGTCTGGCGCCAGAATCTGACCCTGTGGCTTTTCTAAATTT
TGCACACGCGCCCTCGTTGATGGATGCAGGTGGCGACGGTGCGTTACTCAATGACACATCTGGGTTTGGAGACGCCGCCG
GAGAGCCCGAAGAGGCTCAGTCCGCGGCTCCCGCACTTGGGAATAGCGGGGCCGAGGTGGGACCTATAATTGATGATGGG
GTCAGGGCATCCACTAGTACTGCTGCGCATTTGTCGCCTGCTACTTCGGAATCGTCTTCTAATTGCCCTGTTCCTCCGAA
TCCGCAGGCCTCCACCACCATCGATCCAATATCACCTCTGCCTCCTCCTCCTCGCAGCCCTCCTTTTGAGGACATTATTC
AATCGTTAACCAGACTGATCAATGAGTGTAGTAAGGACGTGCCTAACATTCCCCCGCTTCCCACCAGACGTGTTTCAGAG
TGCGAgTCGCGGGCCGCCACCCCTGTTACCAGTACCAGCACCACCAACGCCGGCGCCAGCACCAGTAATGCCAACATTGC
TCCCGACGTACCCCGTCCCGTGTGCGAGATCCGACCCTATATGGTGACTCCCGCCGCTCCCCCGCCGCAGCAGCAGCAGC
CTGAGCAAGCCGCGACCGCCAGACGGACTAGGGGAGGTGCTCGCCCTCGTAACGGCTCTCGCGGTGGGGGTGGCGCTCGG
CGGCCTTCAGCCTCCTCCTCGTCTCGAAGGCGGCGGAGAAACCGGGACGACGACGAGTCCGAGGACGAGATTTTGCCAGG
CCCCAGCCGGAGACGAGCCGTCCCCCCATCCTTTGTCGAGGACGGACTCGAAATTATTGACGCCGAGGAAGAGGTAGCTT
TGGCGGCTGCCTCGATCGCGTCGTTTTTCGACTAAAGGAAACAGAGGTCCACAAGCGGGTGTGTATGTTTGTATCATATA
CGTTATTATTATTGATATTATTGATATTATTATTGCCATTCTTTTATCTTGTACAGTTTTTAATAAAAAACATTCACAAG
CTCCAATCTATGTTTCTCTTGTGGTGTTTTTCCAGATATTCGTTGCACAGCACAAAGTGATCGTTGCCCACAAAGGACTT
GGTCGTATTCTTAGGCGAGGGATGACAGGATTTCAGGATCAGGTGGCGGTTTTTGTCGATCAGGTACTCCAGACTCTGGG
CCTCGGCGCCCCACAGCATAAACACCAAGTGCTCGCGGCAGTCTGACAGTCTCATGATCACACGCTCGCTCAGCAGCTGC
CAGCCCAGATGCCGGTGCGAACCCGGCTGGCCGCGGGCAACGGTAAACACGGTATTGAGCAACAGCACGCCTCGGCGGCA
CCAGTCGTCCAGGCAGCCGCCCACGGGTCGCTGGAAGCCGCTCACGGTGCGAGACAGTTCGCGGTACACATTCTGCAGGG
AGGGCGGCGCGACCTGCTCGGCGAAGGTGCCGAACGCCACACCGCTAGCGCTCCCGTCGCAATACGGGTCTTGACCGACA
ATCACCACGTACACATCGTTGGGCGCGCACAGCTGGCTCCAGCGATGCACGTGCTCGTGCGGAGGGTACACGGTCTCGCG
CCGGCGAGCCTGCTGGACCGCATTCACCGTCTCCTGAAGCAAGGCCAGGTCGCTAGACGGTATGCTTAGAAACACTAGCC
ATTCTTTATTAAGAGACAGAAGGCGAGTCTGTTGATCATAGCTTAACACACAAGGTTTATTATCGGGAATGTTGTCTAAC
ATCCATTGCTTTAAGGCCATTATGCAGCAAAGCGGGGTTGTTTACCAAGGGCTGTTTCCACTGGTCTGAGAGTTCGTGGT
AGTATTTATCGAGGCGGTCAATCAGGGCGGTCTCCAGGTCGTGACGTATGAAGAATTCTTTCAGAAAGTTGTAAATTGCA
AAAAACAGCGGCAGTCCATCTTTTTGGGCTTTGGTCGACACCGGGATGCGGATAGGGCGATGAACCTTGGAATCGAGGTT
GCGCAAAGCCAGGATAATGGAGAACTGTCCCTCTGCTCCAAATTCAAACCCCAACACGTTTTCAGTAAAAATACTGCGTT
GATATGTGAGTTGGCGAAGGTCGTAGGCGCGGCACACATCCTTAACGCAGGTGTAAATGGGGCCCGTGTCTCCGCACTCG
CTGTAGCCTTGCATCAGAGTCATCCAAACCGGCGCAGCGTCAGATTTTAGTAAGGCCAGCAAAGTTCTTAGCTGATGAGG
GTTGTTATGGAGCAGAGTCAGGTAATCCATAAAGTTGTCGTCCAGCGAGATTAGGGAGGAGGCATGCTTTGTGCTAGCGG
TAAAACGCAGAAGTTGAGAGAGATTGCCATCGGGATGATGGACGTGCACCAGAGGCTTAGACCAGCTGTCATCTTCAGGT
GTAAAAGTTTCCCCGGTGAAGCATCTGCGTGTGAGCTCAGGACAGGTAAGGTTTTTCGGCAGGTAGGCAGAGGTTAGCGT
TACCAACAGACTCACCGCCAATATCTTCTGCATGGGTTTTCTTACTCGACAGTTGCTCACTTCACTTTTTAGGTTTAGGC
TGTCCTTGTATGATCAGATGCTGGAACTTTGCGAGTTCTTTGAGAAAGCTGGAGTTCAGGTCTGGAATTCCTAAAACTTT
ATATATTCCGTCCACTAGGTTCGAGGTAGCTTCTGCACTAACTACCAAACTGCTGTTGCCAACTTCAATGATTGCCACGG
GTTTACTGCAGTTTGTTAGCTCCGCAGGGTAGAAAGACACACTGGTAACTAGTGACCAGGAGGTGAGTGGCTGTTTAGTA
TAGTTGCATGAGGTGTTGCTCAAAGTGTCATTGCGCTCGCAATTGGCATAGAAGTACATGACGTCATCAGAGTCAGATAA
TAGGCTGTAACTGGGGTTTTGGGGGTAATATATGTCCTCTGGCAATTCCAAACGCTGTTCTTCACTGCTTCTTTCATCAC
TTTCCTCAGTTTCCTCGTCTGAGTCTGTAAAAATGATAGAAATAGTTGCATTGCACAAGTCTACAATTTGTAGAACTTGG
TAGGGATCTGTAGTGGCGTTGAAGGTTTCGTTAGGGTTCCCGGTGGTAAAGTTAGAGTCCAAAGTAGAAGTTAAATTATC

TAAAGATGTAGGTGTCACTGTATCCGTAACGTTTGCAGTTAAGTTGGTTGTGATTGATGGTGTGCTTGTATTAGTTGCTG
TTTTAGATGTGCTGGTAGCTGTCTTTGTTGATGTAGTGCCGCTCGAGGTAATATTTTTGCTGGAGGTTGGTGTAGGTGTA
GTAGTTGTACTGTTTGTAGTGGCTGTGGCTGGGCTAGAAACCGGTGACACTGATGAGTTGGTAGTGGCAGGAGACGTGGT
TGATGACTTGGTTGTGGTATTGCTTGAGGGAGCTGTAGTGCTACTGCTGCTCACCGTGCTGATGCTGGTAGAGGTGGTAT
TTGGTTTACTGCTAGTAGAGGCGTTTGTGGTACTGCTAACGGTGGAGGTATTGTTGGAAACACAGTCCCTGGTGATCACA
AACAGCACAGCCAGCAACACCCAAGTTTTCATCGCCTGTTGCATACCTAGGGATGGGATTCCGTGAAGCTGTTGTACAGA
GGCAGATGAGGCGCGGCGATAGACGGATATGACAGGGCGCAGATGTTAATATTAAATACGTCTGGCTGCCTTTCTGCCAC
CGCGGCACACAGCGTCGCCGCGTCAAACAGATCTTTCTTGGTCGCGCCCGACACATACAGCTGGTGCACTCCCGTTTCAT
AGTACGACACGGCCCTGGGGAACACGGACCGCACGGTGTCTAGCACGTTGTGCGAGTTGGCGTGGCACTGCTTGAAGGCT
GCAATGGCTTGAGGAGTGTGAGGCACTGTTTGGGTAATGATCTGGGGCCCTGCCAGCACGGTGATATTGTGCAGTTTCAG
CAGAGTGATTTTTAGGTTCTGGAACGCGTTGCCGCAGGCGGCGCTAACTTGGATCTGAATGGCGGTGGAGTCGTACACAC
GGCTCACGCGAGACAGTTCTTCAAATATGTATTTGTACTTTATTACACAGCTGGGAGCAAAGCGACAGTCGCTGGGGAGA
AAGTAAGGCGAGTTGTGGAATTCTTCCGGCAGGTCCGTGGATTCCAGCAGCACGCGGTGGTCCGCCAGGTCTATGGTCGG
GACGATGGGCTCTTCGGACACCTCTGACACGGAGGGTGAAGGCGACAGCGGCGAAGCGGGTCCGGGACAGCTGGCGGTCT
CCTCATCGTCGTCTAGCGGGGGGTAGCGGGTGTTTATCTGGATGTTAATCTCTTCGAATGCTCGGTCTCCATTGCAGTTG
GGGGCCAGGTACATTTCCCGCTGGAAAGTGTCTAGGATGTCGTCCATGTCTTCCATCCTGTTGCTGGGTCCGGCTGGTGT
GTCGCTCTGGGACGCGGGCGGAGCTGTATTGAAGACGATGAGGCGCTTCTGCTGTTTAACATCTCTGGAGCACTTGCGAA
TCAGCGAAGTGCCGTACGGGTACACGCTCTTGATCAGCTTGGGAGCCGCTGTGGGCTGACGCGCGTCACGTGATGTTGAT
GGACGGGGGTGCATAATCCTGACATGGTTCACGTCCTGGTTGTCCGGAGGGATTCGCAGGAATCGCTTCTGTGGCACTAC
CACGATGGTGCGGTGGGTCTTTCTTTTGTTAGCAGGGGCGAGAGACTTGGAGACGCGGTGCGAGCTGTGGACGATGCTGA
TGTTGGAATACATGGTGAGGCCTGTGTGTCCTGTCTGCTGCTGACTACCACTGCTTATAGTACGGAGGCGGTGGGTAGTC
TTTTTCCTGGTCCAGCCTTTGGTACGCCCAAGGATTCTGTTGGCGCATCAGGATGTAAGCGATGTACAGGAGAGCAAATA
GCATAAAGACAATGAAGCCGAGGATGGCATAGCCCAGCAGCCCGATGATGGGCCATGCGTCCGGATCGCGCCTGAGGAAA
ATGTATGTAAGAAGAGAGTGCTTGTTGCCAAGGTGTAGTTTCAAGCGAACTGTGTCAGGAACATAATCTTTTAGATGGAA
AATGTATTCCCAAAAGTTCTGGTTATCTCTGGTAACGTTGCGAACTGTATGTGATGTATGAAGAACGTCGGCGATGTATA
CATATGCTGTAGTGTTTCTGTATATACTTTCATTGAGGCACGACCTGTCATCCTTCAGAATCAAACGGAGTTGAGTCACA
GTGCTGTTCAGTGGAGAAGCAAAGAAAGCTTTCACGGGAACGGTAATAAACTCCCAGGTTTTGTTAACGATTTGTCCTCT
CCCAATAAATTGGCAAGTATACTGTGTGGTGACATTTGGAGTTGGTATTGTTACAGTAAAGTTTGCAGTGCTGCTGTTGT
AATACATTCCTTGTGCTGTAATGAAAATAATTTAAGATGTTATTAGTCAGGGTTAGGATATGCACATGCAGCTACATACG
TATACATAAGCTATAAACACACAAAAGTAACTACTTACACGCATTGTTGGTGGTGATGGATCCAATGCTAGTGGAGTTCC
CATTGCATCCAACTCGCTGAAAGCACTCTATGGATATAAGGTCACGCGTAACGTTTAGGTGATCCATGTTGTAAATCTTG
CAACTGGTAACTGTCATCAGGTAGTTATAACTGTAAGCATTTTCACATGTGACGTTGATGTCAGGTAACCGAGTGCTGTT
GATTGTTGTGTTGGCGGTAGGAGCAATGGTGGTGCTGGTGTTACCAGAAGATGCGTTGGCGACTGTAGTCACTGGAGATT
CTGTTGTATTACTGGACGAATTGCTTGTTGTTTGATTCTCAGCAGTGGTGCTGTTGGATGTTGGAGTGCTGGTGGCGTTA
TTCGAAGAGGTTGTAGAAGCTGCAGTTGAGTTACTAGTTGCATTGGTGGAAGTAGAGTTGGGACTTGCGGAGCTGCTTGT
CAAACTGGTCGGTGAACTTGTGATAGTTGATGTTGACGGAGCAGTTGTAGGTGAAGGAGACGTAACAGTCGATGTGTTTG
TAGCGGTGGTTATACTCGTAGTGCTGCTTGCAGTGGTAGCGGAGATTCCACCCATAATCGACTGACTAACAAGCAGAGCC
ACATAGGCCACGAGGGCCACGGGAGACCCGGTGCCAAGCATGCTTGTCTGACAGAAAACTGTGCCAACACGCTATATGTA
TACCTTATACGTCACTTATGACGCTTGAACTTATAGCTCCATATAAGGCATGCGTTAATAATAAATAGTACAAATACAAG
TTGTGATACTATAACACCCATTACTATTTGGTTTTGTCTCTTCTTAGGCCACCACTGAAGAGCCACCTCGGGGGTTTCAA
CTAAGTCTGTTTCATTAATTACAATTGGAGCAATCTCAGATACAAGTTTGGTGGATAAGGAAGAAACATTGGTAATATTT
TTAATGCGAAAGGATCGTATTCCATATGTACCATTAGAGCCCCAACACGTTACAGTACTGTTTAAATAAAAACTGTATGG
CGCACGTACAGTTACATTCAACTTTACATAAAAAGGTTGATAATGTTTACTATATCCTTTATATCCACCTTTAATAAAAG
TGGTATTGCGTTTTGTTGTTAGCATAATTCCACTCGCTAAAAAGCGGGATCGCTTAGCTGATGTATTGCAATAAATAGTT
AAGTTAGTTGCATTGTAATAAATATAAGCCATTTTCACAGTAGGTGCGTATGGATCACCCTTGTAGAAAGGAATCCCGCA
CACAAGCGTCATTATCACCAGCCCGCACCCTAGGCACAGCCAAACAGCGTCAGAAAAACGTAGCATCTTACCACCTCTTC
ACCGCGGGCTCAAAATATCGCGTTCGACGTGATAGCCATAATTTATAGTTGTACAACTGACCTGTGGTGTCATGGTCCGT
CTTCAGGATCCCGGCCACGATTACTAGCACTGACAGCAGCACACAGATGGAGCACAGGAGGCGCGCGTCAAACTCCGGGT
CGCACTGCCTCAACTGTTCTCTGTGGGGCACCACGCCTGCACACAGCAGGTCCGGGCGCCCTGGGGTCTCCTCCAGCAGC
ATCCCTACGGTGGTGTGGGGCAAGACAGGGGTGATTAGCATCTCAAAGTACGGGGACCACAACGCCTGACGGACATCCAA
CTCGTATGGATCCAAGGTACTATGGTTAAAGGCCCAAATCCTGTCGGTTAAATTCATGAACCAATAAAGGTGATTGTCTA
ATTTGGGTACCAAGCATTTCACTTTCACGCCTTCCACGTCTGCACACATCTCTGGGGTGTTCATTCCGTCAGCTGCGTTC
ACAACCACGGTGATGAGTGTCATTATTGTTAACCACCACACCACATTCACACCACGCAGCATGGCAGCACACCACCAAAC
AGATGTCTCACAGAACAAAAGAGGCCAGCACACCGGAGTCAGGTTATCTACAAATGATTGATATGATTTATTAGCTAGCT
TATTATAATAATATATATATACATACACATACAAACTTGATAATACTGATATGGTATACATATAACACTGTTGGGTAAAA
TCAAGTATCAACTAACTCTGCTTGTGCTACACAGCATCACAGAGCCACCTACAGTTCATCAGTACCTAGGGCAGCTACTG
TGCCTGCTGGCTTGCACACTCTCATCATCTAATAACTCTTCTTCACTGGGACTCCTGGTTCATGTCATTGATGGCTTTCT
CGATGGCCATGCTCAGGGTCTCCAGTTCCTCGTCTTTGGGTTCATCAAACTGGCCCACCGCGTACATCTTGGAGGCGTTA
TCATAGATGGGGAGCATGAAGGAGTTGGTGTTCTCTCCAGAGAAGATGCGGATGGCCACCTGTTTGGGGAACTGCTGGTT
CAGGGGAATCAGCGTTTGGATGGCTCCCAGGAAGTCCACGGGGGTGGCGGCATGTATAATCACCGTGCGGTAGTCAGATG
AGCGAGGGCACAGTTCGTTGGTGTGCACGGTCTTTACATCCCATGCGGCTCTGAGTCCTTGCTGACAGGCCTCGGCGGTC
TTGCGGGTGGCCTCAGCATTGTGCACATGGGGCATGGTGTGTTCCATCAGGAAAGGAGTGGAGATGGCTAAGTTACAGAT
CCTTCCCAACTTCACTCTGGCCTGATCCACGGCCTCTTTCACTTCATGGGTGCGGGTATAGATCAGCATGATCCCTCTGT
TGTTCACTTGCATGCTTTTGCAGATGTTCACGGTTTCAGACAGCACCTGGTGCACGTTGGTCTGAGTGAAGGGCAGGTTC
TTGTACTCCAGTGACCTGCTGGTGGCACGGAACATGCGAGACACTTCACCAGCTTTCACTCTGCCTCTCTTGGTCTCCAC
TGTAGGAGTGGTGAAGACAATTCCAGCTTTCTCCTTCATGGCACTTCTCACCTTCTCACAGTCCAGGGAGGTCATTCTCT
TGGTGGCCTTAGTCTTGGGAATCTTGGGCTTCTTGGCACTACTGGTAATGCCAGATGAGGTGCTGGGCATCACCACAGGG
GGAGGTGGTGTGGGAGGAAGAGGGGAGATGGGGGAGGCCAGTGTGCTGGGTGAAGACCCCCCTCCACACTCATCACAGCA
CTCCGAGTCACTGCTGGAGGAGGAAGAGGAACTCTCCACAGACATTTGACCCAAGGATCCAGTGCTCTGGGTGGTCATTG
TGAGCTTCACACCTGGTTGTGAGGAAAACTCTTGAGAGACTTGGTGTTCCTCCTCGTCTTCGCTGTCTGAGATCACAATA

AAGCCAGAAGGAGAGGTTGGTTTTCTGCTGGACTTGTGTTTGGACTTGGCTTGCAAGCAATCTTCACCAGGCTCTGTCTT
GATCTCCGAAGCGGGCAGCATGCACGTGGGAGGCAGGGGCGGTTTAATAATGATCTTGGTAGGGTGATTAGCCTTACGGG
ATTTTTTCCGTTCAGTGCTGCTGGTCTCAGGTGGAGAAACAGTTGCAACGTTGGGCAGCGTCTGGGGCTGAGTGTTGACA
CCCTGAGGGGTGCCTGGGGCGGTGGTGGTGAAAGTTGGAGTGGTGATAGTCATACATGCTGAAGTCTGGTCAATGCCAGC
TTGATTGACAGCCTCAGCAAGAATGTCATCATCATCATCATTACCACCACCACCTGAGTCAAAGAAAGACACGGGACACA
CTTAGTACTTTTTGGAACAGGGAAACATTTATTGATCCACGTACACAGGACAAGGACACAATATCTTCCCAAGTTCCCAT
CAGACTAAGTCTCAGACACCAACAATACAGTGTCATATGTGAGTATATGTTGTATATATGGAGTATAACATAAAGTATAT
AGAGTATAAATGGCAATACTTTGTTAATACAGTTAAACAACCAGGGGGTTACAATACAATTAGTACATTAATCCTAACTG
ACTGACTAAATAGGGGCACAATGTTTGAGTTATAGGGACCAGCCCATATAACTATCTGTCTTTACTGGTGGCTCTTCTTG
CTTCTGGTCACCATGGGATGAATGCTGGTGGAGGGTTCTGGTTCCTGCTTCACTTGAGGCTTCCCAGCTTTGAACAGGAT
TATCTCTACCTCACTCTCTGTCAAGGATTCCTCCTCCTCACTCTCCTCCTCAGCCTGCTCTTCAGCCTGCTCTTCAGCCT
GCTCTTCAGCCTGTCCTGCCTCAGCCTGCCCTTCCTCTGCCTGTTCAGCCTGCCCTTCCTCTGCCTGTGCAGCCTGCCCT
TCCTCTGCCTGTGCAGCCTGTTCTTCCTCCACCTGAGTCTCCTCCTCAGCCTGAGTGTCAGCCTGTTCCTCTTCCTCCTC
CAGCAGTAaGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCACAGAGGCCATCgT

UNIQUE SHORT CONTIGUOUS SEQUENCE:

GAATTCGCCCTTGCCCCAAGCAAAACCTGTAAATACTAAAATAGCGAGATAGAACGGATCGTTATGGGTAGATATAGCGC
ATGATGTGGTGACGAGGTCCGGAGGGCCAGCTCGACAGACCTCCCGCGAAGCGTAGCATGTGCAGTGCAAAATAAGCAAG
CCGAGGCAGCACAGAAGAATCTTCATCCCAGCAGCCACCATCTGCTACGCGAATACAGTATAGCTTGGTTGTACGATGCG
GCCTTTTTCTTTAGATGTTCCAAGTAGAACTTCCTAGTTACGGTCCAGCCGGGTAATTCTAATTCGATGTCCACAACCTG
ATAGAGGCCATCATCGTCTTGAGCAAGCTTGTAGAACAGTTGATTGAAAAACACCCTAACGTCTTGGACGTTAGGGTTTA
AGCATACTTCGGATCGGTTCTGCAGCACAGCAATCACTTGCTGTCTTTTACAGGTTTTGCTTGGGGCAATAGGCCTTACA
TAAAGTGTGTCCTTGGGAATGGCCCGTAACAAACGACTTTTTAAACATGTACAGGGTTTTGTTTCCGTGACGTTTTTGGC
GTGAAGCCAACATGTAAACATAACCAAAATAACACTGGTGAGGTGTAGTTGTAAACCGAGTATGCGTTTCATATTTATGT
AAGTACCGTTTAGGATTCATCGTACCGTGGAATACGTAGTGCCTTTTACATTCCGTAGTTTCGTACCGTTTAGTAAAAAT
GGTATTGTACCATAGTTAAGAGGTAGCGCTTCGGTTATGGAACGACAGCACATGATGTATGATGTTTCCGTAAACGTATT
AATTTACCCGGTCCTTTTGTAACCGTAAACCATGTTGGTGTGGTAATTGCATTAAATTGAGATTGATTGCTGATGTGTAC
ACAAGTTTTCTTTTCAGACGCGTCGTAGACAATCAGTTCGTATTTATTACAATGATAACTTTTTGGTTCTAGGTAAACAC
AAGTTGCCGTCCACGGTATACCGTGGTAATACTGAACACAACGACAGCGCAGTTCCTTGCTCCATGTTTTTTGTACACAT
AATAAAGCCAATATGTAAAGCGCAGTTCCTAATAAATAACGCGATGTACTCATGTCGGATAAAAGTACGCATGTAATTAA
ATTTAAAAAAGATACAGTACAACTAATTAAATTATAAGAATGCATTTATTGAGTACTGGATTTTCTTCTGGCTACTTCTA
CAATTTTCTTAGTCAGGGGAGCGTCGGGATTCATACAAACTGGAAGCCCAGTCTTGAGCGTAACTCTGCAAACGTAATCA
TAAAATTATAAGAGCAACATATCTATATAAAATGTACATCAACACAACGTTTAACACAACTTACACAACTTCTGTTTTTT
CACAATTTCCATTTGGAGGTGTAATTTCCACGCGCGTAACCATCTTAGGGTTAATTGTATCCGTAGTTGCTACACAGCGG
CAACGCGGTATGGTCATATAACTATTTGCAAAAGTACCAGCAATAAGTAACATAACTACTACAAACGACATAATCTTAAC
TTTAAACATAACTGGCTGCATTTCGATTTCCTGTTGTACAACTGTAATTTACTGTCAGAAAACTATCCGTCTATTGTTGA
GGCGGTTATGATTGATGATTAATTATCTTCATACTCATGTTTTATGATGTTGTTTCAGGAAACGCCATAGTTCCATGTGA
AAACGATGCGGATACATTTTATTAGATGGTGTTATCTTCATTTTTATGGGTACCTATTACGAAACACGTATACCAATTGT
AACTCTTTTAAATTCGAAGTTTTTTCAGAATCTTGTTCTTCACCATTGGTGATTCTGGATCCAAACATACTTGACGTCCA
TCGCGTAGCGTGACGATAACTTCCACAACATTGCAATTGGGTCCTGGCAGATTCACCTTCACGGTTTTAATGCGGGCATG
GGAAATTCCCTTAGCAACGTTAGCGCATTCACAGTGTAGTTCACGTACAGACTCAGCGTAAGCAATTAGGCTCATCAGTA
GTAGAGCCACTCCTAAAAACCGGGGATTGAATACAGCGTTCATGGCGGTTCACTTTTCTCAATCTGTCGTTTTTCTGTCA
CAAGGGTAACCGTGAATTGTTGAGTCTTCGGCGATTCGCTAGAGATAAGCATCTCAACGTTATATTGTGATGATATCAAT
TATGAGCGTCGAGAGGATACGATTTAAATCTGTGACACAGTCTGCAGTCCCAGTCATGTAGCCTATGTAACGTAACTTTT
TTTTTTTTATCTTTATCGTTTCTGATATAACACCAGGCTCCAGAAGATGTGGGAAGTTTATCTTTGTATGTTTCATAATC
TAAACATAGTGGTGTGTCTCTGAGTGTATGATTGAAATATGCTACAACTTCTGTGTGTCGGCATCCAATGTGGGGGTTTT
GAATCCAAACACATTGAACGCGACCCGATGGTACGCGTCCTTTGTAACCGTGCATACAGTTGCACCGGAATCCTATATGA
CAGTCACTTAACATAATCACACACATTACAACACATAGTACATAGATGCGCATCACGATTCAGTTACCGGTAACTCTTTA
CGTTATAATTTTTCCCTGGACAATCATTATCGTACCTGTTTGCCTTTCACAATATTTCACCCTGCATTCCCCAGTCCGGG
AGCCACCACCAAGGATCTCCATATGAGTCCGATCTCAGTTTATCGTACAGATTATGTGCAGTGCAATTCGAACATTTTAC
TGGCACGTGGGTACCGTTGATGCGTTGTACGCAGTAAGCTCCATTCCCATCGTCTCCAGATCCATGGCGGTGAATGGTCT
CGTTCACGAAGCCGTAGCTCAGACATATTGAATGCTGCGGTTTTCCATTCCTCGCCTTCACCGACGGTGGAAAGTGAGCT
AGGGCTTCGTTGCCGCATGGTGGTCCGGATCGATGCAGCCACAGGCAGTCCGCTTTCGGTGGTATGAACTTAGCGTGCCT
GCTGTGCTTGCATGGACAACGCCTCACCTTCGCCACGGCTCGTACCAGAGATAAGTAGCAGCAACAAAAAACAAGACAGG
TTAGCAGCAGACAGTATTGCATCGCCCCGGTCAGATCAACGGAAATCAGATAAGTGCTCTGGGTCTGGTCCAAAAGGTTA
GCAGCTAGAACACAGGTGCACATTATATGTGTGCCTCTGCGAGTGCAATATGTGCGTCACGGGCTGACGTCTGAGGTAGG
GAAGGTCTCAGATGCGATCCACGCACCCGGGACCCAGTCCGAGACTTCAAGTTTCTTCAGGATTTTATCAGGGATAAGGC
GCAGGGTGACAGCAGATATGGCGCTATGCCTCCGGAGGACGAGGGAGCGTTTGCGTGAAGTAGCCATCAGTACGGTGCAG
GCCCCATGGTCGGCGGTGTAGGACGCGTGGTGCGTCAGGTGCTCCCATCGTCCGGGTCCGTCGCGAGGTCGCGCGGACAG
TAACTGCACGGCGTTGTGGTACTGAATGAAGACCATGATGTTAGTTGTTTCGAGGGGACCGATGCGATTTGCAAACGGGC
TTATAAGGCCAACCAGACGGAAATATTCCGCGTCACGGTTTCCGCGGATTGGCCACCTCGGATCACGCGTTTCCCTCTAA
TATATAAAGTGCTCACGGGACGGATCGCGCAGGTAGCGGGCGACACGCGCACAACACGAATAAACACAGATATCAATCAC
CAACTCGCAACATCATGCTGAGATTGCTTTATTATTATCATGTGTGCAATGTGGTTGTGTGAACGGATCATGGTTACTAT
ACAGAAACATCAATCAATCTGACAGACAGCCGGGATAGGCATGTTGCTATGGCGGGAGGTGGCCCAAATGGGGAACGCAG
AACAGAGGAGAACACATGTGTGTTTCAGGCGATGCTGTGGTTATGCCGTGGTCATGCCATGGTTACGACATCCTTTCGTA
CATGTACGCCGTCAGCGTTTTGCCGTACGCGGTGTTCAGCAGGAGCAGTCTGCAGCCCAGGATCAGCAGGATGAGGGAAA
ACAGGAACGCCACAAACACCAGCAGGAACCACCAGTCGCGGCTGGCGGAGTCGTCGTCCTCATAGTCCGAGGCCAGCGAG

TGGTTGCCAGAGTGGTGCTGGTTGTGGTGGCCCTGGCCCGGTTTCAGATGCGCGGTGCCGTTCACCGATCGGCACCTGAC
GTTGTGAGTGGTGGTACAGTTGGCAACCACCACGGACGAGGAGGCGTTGCACGTGTCGCATGCTGTGCAGTTAATGTTCT
CGGTGGCGTTGTTGTACGATCCCGGGGGACAGGGCTCGCAGACGGTCCCTGTGGTATCGCTACAGTTTGTCTTTACCCGA
TAGCCCGTATTGCTGCACTTCGGGCAGCACTGTGGGCCCGCGGGATACTCGTCCGGCTTGCACGGGGGTGCGGGGAACCC
ATCTGCCAAGACAACGGCCCAGAGGACCGACAGTAGCAGCACAATTTTCCAGGTTGGCTCATGGTTGCGATAGGCGCGCC
GCCGTCTCGTGGCTGACCCTATTTCGCGGCGCTGTTGACGGGCCAGCGGGTGTGGTATGAAACAGCGTCAGGTAAATACA
CGGTGTTTGTCCGCCATATTGTGTACCCTCGTCTCCTCCCCCACCCACGGGCATACCCATCCACACACCCGCGGACACCC
ATGGACACCCAGCGACCCCAGCTGGAAGAACACGCGCACACGCTCATTTCTATTATCATAAAAAGCTTACAAGCGTCTCT
TTATTTCACTGTACAAAGGGGGAGGGTAAGGGGGATCATCGTCGTCGTCGGGGTAAGGCGGCTCGGGAGACTCGCACTCA
AACAGCGCCTCGTACGACTTGCTGCACTCGTCCCCGTCCACGTCGCTGTCCATGTCGTCGTCAATGCTGCAGTACAGCTC
GTCGTACTCGGCCGGGCGCACAAACAGTCCAGACACAGGTCGGTAGCCGTGCTTCTTCTTTTGGAAGCGCTCGTAACTGC
GGCAGCACTTTCCGCACGTCATGCGGATAAGCAACATGAGTGCCGCGCACAGGGCTCCGATCAAAAACACAGTCAGCCCC
AGTGTCACCCAAGGCAGTTCAATAAAGTCATCATCCTCCTGTCCATCATGAGGAAACAGATGGGGGTGTATCTTGCGGCG
CCGGGGAATCTGGATGCGAAAGCAGTCCCCGTCCAAGATATACTGCTGCACGACGGGCCAGCAGCGATCCGGTGCACTCG
CAGGGCAAGGCAGATTGTGCTTGGAGTGGTTAGTCTTGGCCAAGTAGTAGAACACTCGGTGAAAGCCGTAAAAAGTCCCG
CACGTTTGGCGCTGCCACTCATTGTGCTCGGAGGTCCAGATCCTGTCACCGTCGGCGTTGAATCCCAACTTGGGTTTGCA
ATCGGGTTCCCCGGGCCGCTGCAGAGTCTCTACAAAAGTGATCACCGTAAACCGTTTCTCGACGCGTCCGTGAGTGTCAT
TGCGCAACACGCAGTCAAACATGGCCGTGGTGTTGTTAGTTACGTTCAGAGTCAGGGTTATGTAATGGGCAAGCCCCAGG
TTTACAAAGTCCCAGCTAGCATTTGATAGACGGCTAGTCAGATCATCCGTCACCAGCGTCCCGTTTGTCACATTCACTAT
GAGTTCATGACCTGCATACTCCGATTGACAGAAACGGGCGCGGATGGAAACCATGGGCCACGAGTGAGGCAGAAAGGTGC
AGGGAATGGTGATCTGCGATCCCAGCGGCATGATGAGGGGCGCCGGGGAGTTGTCCTCGTAGCATTGGTCCCATGACAGA
TGGTCGATGTCCGTCGACCTCGCTCCCGAGGCCAGGCTGATCAGGACGACGGCCAGAGCCGCTACGGCGCGCCGCACTGA
CATTGCTGCGACCCGTGGCACCATCAGTCGTGGTACGGAGAACTCCACGCTGGGAAACCTGCCCTGAATAGTCCGTTCCC
GGCCAAATCCAACGTGCGCGGGGCGTTTGGAAGAGCGACCCCGCTCAGTGATCAGCTGCGCGGTTTCATCCTCCGTTCCG
TTCCCCGCCGACCCTCTCGGAAACCGCACGGGGCGCATGGTCGCGGCGACCGTCGGCCACCGCCGACCACCGCCGACCAC
CGCCGACCACCGCCGACCACCGCCGACCACCGCCGACCACCGCCGACCACCGCCGACCACCGCCGACCACCGCCGACCGG
CCGCACCGTGTCCGGCTCCGCTTCGCTTCACTGCGGCACGGGCTGCGGGCTCGTGGGCTCCAAGATCAACACAGTGTCCT
CAAAGTTGAAATCCTGGTCTTCTTCCAGCCTTCTCGCTCCCGATCTCTTTTTAAACACCAGGAACGGAGGGTTCAGATAG
CGCGAGTAGCGGTTTTTCCAGGGTTGGGCGCAGCAGATGGCGATGAGGATCATCAACAGGATAATGCCGGCCAGTATAGT
GGCTACCAGCGTAGGAAGTTCCATAGCAAAAGAGCGCTCGCTCGGGCGGCGAGTCCTTTGGCCAAAGGCAGCGCAGACGG
GGAACGGTGAAAACTTTTATCGGGATCTGGTGCCGTTTGAACATGGCCCGTCCCACGTCCGCTCGTGTCTTCAGGGGGT
GTTCTCAGTAGCTGCTTGGTGTTCCAAGGTGTTTTCCACATCTTCTTGGATGTACAGATGGATATGCGCTGCGGTTTGCA
GAATCAGTGGCTGCCGAGTTGTTGGCTGCTCAGTTGTTGGTTGAGTTGGCTGCTCGGTCAAGAGTTTGGTGACAGCGGAT
CCACCCTCGATGGTTACCACTGTCTGTTTAGGTTGCGATGACCGTATGTTCACCATCTCTTCGCCTCCTTCGCAAGGCGG
CGAGTAGTACTGCACGGGATCGGGTTTGCCGCCGGGCCCGTAGCGCCAACGGTCCATCCACTTGCTGTAGTGTCTGGTGA
ACCATTGTCCCACTGTCTCGTACGTGCCATCTTCCAGCATGTTCAAAGTCTTAAATTTCCTAACGTCGCAGCCATACCTG
TAGCAGCTAACTATGCACTGACAGGTGAAGACATCAATTATCGCCGCCCAAGCATCTCGCGGGTAGTGGCAGAGAAAGAT
TATACCTGCTATGATGACTAGCAAGCAGATGGCAGAGAACACTAGCAGGCCGATGTTGACAGGGTCGCTCCAGTCGCCGC
TGCCGTCGCTCATGGTGACGGTGAACGGTGTTACTGCTCACACGAGTGCGAGTGAGTCCCGAAACACGACGTGGCTTTTT
ATGGCCCATTCTCCGCCTCAGCTTGCGTCACCACTCAGTGTTTCATTCTGGGTTGGGTAACCGCTACGGATTGCAGGCTG
GAGGAGATGCCGCTGTCATCTGAAAACACGGAGTCGTCGCTGGACTGGCTTTCCCTCTCCGTCCCGGCATCATCCATATT
GGGATCCAGCAGCTTTTCTGTCTCGTTCCCAGACGATGCCTTCAAAGCTTTGCTTTTCGGTTTCTTGATTTTGGTCTTGG
TCTTCATCTTCGTTTTAGCTTTGAGGTCGTTCACGTTGTCAATGTCAATCAACACATCGTTGGCTTGGTCTCTGTACAGC
CGCTGTTTGGACCGCGCACCCGACCGCTTGAAATCATCGGGCCACACGCGGCCGTTCTTCTTGCGGCAGCACTTGGAAAA
GATGGCACCCAGCCAGGGACAGCTAAAGAAGTTGTAGCACACCAGGATAACATACATGGGGTAGCAGAACATGGCCAATG
TCATGATGGCCACAAAGATAGCGCCCACAACAACCAAAACTATAAAAACCATATCCATGGCGGCTCCCGAAGCGCAAAGT
CCAGTCAGGGCTGCCCCGCGTTCACAGACGACACTGTTGAGTCGTGCGCGCTGTGTCTGTCAATCAAGCGAGCAATCAAT
CACTCCTGTACACAACAACAAAGCTGTCTGGCACCGAGCCTTCTGGCACACGCAGACCTCGGGGAATTGTTTTTTGTACA
GTCCCATAGGAGTTGCAAAGTCATAGCTGATGATGCGGCATCACGTCGCGGGCATTTCCAGGTGACGTCCGGGGTGTACA
CACTACCTGGAATCAGTCACTACTCTTGAGACATCCCCCCGGTTTGCATGCTGGTGATGATCCGCAGGGTGCTGTGCACG
GCGGCGTCTATGACGGCGGCCATCTCGGGCATCACGCCCGGAAAGCTAACCGCAAGCTCCGTAGCCACCGCGGCGGTGAC
ACTTTGCTGGTTTTCCATGGTCACCGTCACGCTTTCGGGTGGCGGCGTGGCAGGGGGTCTCTCAGGCAGGCTCGGCGACA
GTGGTGGCGGTCTGTCAGCGCCCCCGAGCTCCGAACAAACCAGAATAGGAAGTTCCATGTAGATGGTCTGCATCTTGTCC
GCTGGGGCAATGTCACGCATCTCTGGCATCACCATCTCTTCCCGCTCCGAGGCAGTCTTCAATTCAGTCTTTGTCTTCTT
TTGAGGTTTGGGGCAGATCAGCAGCACCCTGTTGCAATCTGAACATAATTTGTCTCGTAAGTTCTCAGGTCTGTTCCCGC
AGAGTCGGGCCGCGAGGGATGGAGAGCAGCACCAGTCAAAGATGTCATCGTGAACCCTGAGGGTTCGTTTCCTATGGCAC
ATCAGAAAGAGTAACAGGGCAAGGACTGTACAAAATGAAATGCAACCAAATATACCGATGAAAATGTGTCTGTCCATCAT
GTCCATGTATGCGCAAAGTGCAGTCCGGTAGAATGGCGCGTGGTCGGCAGCAGAGAGGGGTGTTGGGGTGTTGGCGCGGA
GCCTTCTGGCGCAAGCCCAGGGCAGCGGGCCTTTTATGTCATCACATGATTGATGTTAGTCGCGGGTCACGACGGATGGG
GAACACGTGGTGGTGTCGCGTTTCTCAGGTATCTGCTGCTTTTTCTTTCTGAGACCTTCATGCTTGGCGTGCTCGGGCAA
CGTGGCGTACACATTTTCCAGAATGGTTTGCAGAACCTTCTGAAATCTGTCGGCCATGTGATGGGAGAAGTTTTCGGTCT
CCACCATGCCGGCCAGCGTCACGGGGTAACCGCAGGCGATCACGGTCACCTTCACAGCAGTGTCTACTCTTCCCGTCAAC
ATGACTCCCCGCAGACACGCCTTTACCCCGCGTCTCAGAAGTGCGGCGAGGTGATTAGCAAACGCGCTGGTGTTTTGCAG
ACCGACCTTCTGGGCTCTAAAGTTGGACACGTCCACGCTCACGGTGATGAGGGTGTGATCCACGGCGGGCAAGCAGTAGC
TTAGGTCGGTCTGCACCGCAACAGTGGTCGTGGCGGCATGCTTCCCGGACCGCCGCGGCAGGCAGCAGCAGAATGCGTTC
TTCCAGCACCTTTTCCATCTCCCACAGGCAGACTTCATTGCGTTCCACAGCCTCATGTTGCAAAAGAGTCTTTCTAAACT
TGTTCAAGAGTTGCGAGAGGTATCTCCCTTGAAGCAATGTGCAAAAACTGAGCAGGGTTGGCTATATATGACGGGCAGTC
CAGTTCCTGGTTCATTGTGTTGTGGTTAGCGGTCAGGCTCTGGGGTTATTACGCAACATCACGGAATCCATGGCAGCCTG
CAGTTTGACGTCCATGGTGCGAGAAAAGTGTTCACTGTCTAGCACGCCGGCCAGACCCACCGGACAGTTTGTGACATAGA
TTAATACCTCGGCCGCCGTTTGTCGTTTCCCACTCGACTTGACAATTCGCAGGATGTTGCACACCTCCTtGGCTAAAATG

TTGTTAATCTGATGCGAAAACTTTTGTGGGTCAACCAAATAGAGCCTTCTGTGTAGCAGATCAGACAAGTCGAGGATGAG
TTCCATGTAAGTAACACCAGCATGGGAGGTGGTATAGGTACCTGTGGTTCTAATCTTAATGGTGGCATTCTGCGGAGGCC
GAGCGAAGTACTTTCGGATGTTGTTGGTTAGTCTTCTCAGAAATAGTTTGAAGCGCTCGATAAATATGGAAAACAGATCC
ATCTCGGCGAAGCGCGCTGCGATGTGGTTAACTATCGAGCGGAGAGTGCTGTATATATAGTAACCAGGTCCAATGATATC
GCTCGGTGTTCACTGGCGCGGGGTGTTGGCCAGTGCGGCTTTCAGTGCGGCCTCCACGGCCGCCTGCACAACGGCGGCAA
TGTCAGGCGATAGAGTAGACAAGTCTACGGCGGGTGTGGACTCCGGTGCGGGAGCTGGCGCTGGTGCGGGAGCGGGCGTT
GGCTCAGGTGAGCGATCAGGTGTTGCTTCAGGTGTGGCTTCAGGTGCGGTGTTCGGGCCAGGTAACCCGGAAGACGCAGA
TGCGTGGTTAGGTAAGTTGTTAGGTAAATTGTCAGGTAAATGATTAGCTTTCGGTTCGTCCTCCTCACCCGTTTCTTTCA
AATTGCAATTTGATGGACACCTGTGAACCAGGACCTCAGTCTCATTGGCGTTTGCCGCCTCCTCCTCGTTCTTCGGAGGC
GAACATTCGATCTTCGTTATCTGATGTTCGCACGGCAGCTCCTCCAGATCAACTGTCACGCAATCGACCTGCGTGTCCTT
CTTAGTTTTGGGCTGTCCCCGGCGGCTCTGGCGCGGTTCCTTTGTGCACAGCTTGGCACTCAGCCAGTTCCAGAACTGGA
AGCAACAGGGGAAGTAGAACTTGAGGTTTTGTTTGGCAACTTGGTTAGAGCAGCGTTCCCCTCTCTTGACCTCCGCGACA
AACTTCCCGATCTGAGCGAAGCATGTCGAAAAGCATCTGGAAAAGCAATGGCAAACCTGCGCCCATGCGGCCGCGCACCA
CAGGGGACACACACAGCAACGTCCGTCATCTTTCTTGTCCACGGCGCCCTCTTCTGGATCGGCAGGTTCTTCTGAATCCG
CAGGTTTTTGCAGGGAGGCTATCAGCGAGGCTACCTTTTCCAGGCAAGGCTTGAGGCAGCAAGGTGCGTATTTGCGGTTA
CGATAAACCCAGACGAAAAAGCATATTACCCCGATGATCGTTAAGGGTAATACACCGCAAGCATCACAATGCTCGTCGT
ATCAGTCATGATGTTATGCCTTTAATCCTCCACGAAGCCACAAAACCACAAAAGAACGACTCGCACACCTGTCTGAGTGC
AAAGCAAGCGCCAAGCGGCCTTTTTATTTACTGACAGTGTATTAGGTCATGTGAAACACTCGAATCATCACACAACATAG
GTAAAGCGCGTGAGGTCCTGGCAAAAACCCGAGCCAAACCGAGAATGAAGCAAGATACATTTTCACACAACCACATATTA
CATCACGGTAGACAGCATATGTCAATCAACCATCACAGTTCTGAAACCAGCTCAGCGTGGCACCGGTGAGCCTGAGGTCG
CTCTTCATCTCCACCCACTATAGTCTCAACGAAGACCGTCTCGGTGGCCGCGCAGATGGGAGTGCATAACACGGTGGGCC
GTCGAGTAACGTTTTCAACGACGATCTCGTCGCCCGCACCGTCCTTCGTTTCGTTCGGGGTCCAGGTGTTTACCACTTCC
CACTCCACTCTCTGGTTGCTTTCCTCCTCGCGGCGACGGCGCCAGTGTGATGACCGTCGGCTGAACTGCGTGCGACCCCG
CCGCCAACCGTTGAGACAATAGTAGCAATCGTAACAGCATATACGACATTCGTGACAGTTACAGCACACATCGCGACAGA
CTTTACAGGTAAAATTCCAGTCATAGATCAGGTAAAAGAGTCCCACCACCGAGGTCATGCCACCCGCCCCGCAAATCATC
ACACCTAGTGGCCACATATCTAACACGCTGCCAGCCTGGGTGTAGGCAGGTACCGATGTGCCATTGGTACAGTTGCCGCT
CATGGTCGCGGCGATATATTTGGCAAAAACTAACTTTTCCAGTTGTACAGGTTTGCAAATTAGGTAGAATCATACGCGTC
ACTGGAATCATCATGAGCAGATGTTAACCTTCTTCCTTCTGGATAATAACGAACGATTAGCTGTTCCCTGTCCTCGCTCC
GATGACGGGTGTATAACGAACGTTTTACTTTTCCTTGTCCAAAGATTAGGGCCGCAACAACTGTGAGGACTAGTCCCAGG
GCCCATACGGCATTTGTACCAGCGGTTGCCGCAACTGTCATAAATTGATCAGTGCTAGGTGGAGGAGTTGTACTGGTAAA
ACTACAGAACGACGGTGTGCTATTTGGAAAGGTTCCATTTCCATCCAATACTGTTAAATCAAAAGATTCGGGTTTTGTGT
TACCGTTTTCAGTTTTTAAAGTATATCTTCCACTATCATTCTTTGTGACATTATACAACCATAACGTTGTATGATTACAG
TTATAACAGAGGCTTTTATAATGATGTTTAGTGTTATGGGAACCCGATGAACGACATATTTCTTGGTCTTTTGGTTTTCC
AGTGCTGTTATCAAGTAAAACCCAATGAGTGTGATAATGACTAGTGACAGTAACATTTGTTGCCGTTAAATTTACGGTTT
CGCCTTCGGTAGCATTAACAGATACAAGTGTATACGCCAAAATAGGTGCTGCTGTAGTAGTTGTTGTACTTGCATTTAGA
GTTGTGCTTGTTATGTTGGTCGTTGAACTAGTATAATTTGGAGTTGACGTTCCGGTGGTATAATTTGTAGATACGTTGGA
TGTGCTGGTAATTGCAGCGGTGGGTGATGACATGCTGGATGAAGCATTTGGTGAGCTAGTATTAGGGGTACTGGCTAATG
TGGTATTGTTCGAAGACGTATATGATACAGTCGTTGAACTGGAATTATTGCTAGTACTGGATTGATTACTGGTCTCTCCT
GTGGACGTAACAACTGTGGATGGTGATGAAACGTTTTGGCTAGTTGTTGTGAACGTTGTACTTGTAGTTGTGTTGGTAGT
AGACAAAACCGTGGATAGAGCAGATGTTGCATTCGTGGTTGTAGGTGTAGTTGTAGTTAGCGAAGTAGAGGTGGTTACGC
TGGTAGTAGTATTTGGTGCTGCAGTTGAAGATGTATTCGATGATGGAGTTGTATTTGCAGCGGAGCGAGATAGGATGGCG
ATGAAAATGTAGATGGTGGTATCTGCTATGATGTAGCTTATCCTGAACTGCATCTTGCTGCAACCACCTGGAGCACAACC
ACGTCGTCTCGGTGTTTCAACGTACGTGACCTTGAGTGCAGCATTCGGTGGATTGCGCGGACCATCTCGTTCGCAACACC
TTTATACATTGCTTGTGAAACTTCCTGCGTGGCATCATCGGAGGCTTCGGGTTTCTATCGGGACGCATTCCAATCGAAGT
GGTGAATTTAGTGTTTTACATGGACAGGCACACTGCTGGATACAAATGACAATCATGTGTATGGCATTAAGATGTTAAAC
TGCGGAATACGCTGTAAAGACATAAACATAACGTACATAAGAGTGCGGTGCTTTCCATATGTGGCGGCTGCGCGTTAGAA
GGCTGCAAGTTCTGCTGGCACATAATGAAGATATGTTCTTATATCACCATAACAATGCTGCTAATAAAAGCTTACCGTTA
GACGCTAGGACACTCTGTCTGTAAAAATATAAATGGAAAACGTACCTATTCATGTAGTTGCTTGCGCTATACATATTTGC
AGTTTAGATTGTAGGCACACTGCTGAGACACGTTCACAGGTGGAAGTTAAACACTTAAAACGCAGCTGCTAAATTAAACA
AATAAAGCTACTGCTATAATTCATTTAGCAACCGTCACGTATTAAAGCACTTTTTCATAATCCCGCAGTATCCCCATTAC
CCGGCTTCGTGCTGTTAGTGTTGTATACAACAGGCAGCCCATGCGACGTTGTGTTAGTTATTTTTTAGTACATAGTTTAT
GGTAGTAGCTCACAGATAAGAACGCACGCTAATTAAAATTTACCACACACTGAAATGGTACTCTTGTACTGACTTTGACC
CATCATGTAGGCAGATACACACACACACACACACACACACACACACACACACACACAGTGCGTGCGATAAGCTGTGTGGT
GCTGCGCAGATTCATGCTTGCACAGTCGTAAACACTTCGGTTCTGAATGCAATGAGCCAAATGAACGTTCACCGTATCTG
AGCGCTCTAGCTTATATTCTCACATATACATAAGGAAATTAAACCTTTCCCGTTGTCATACTTCTGGGAAACCTTACAAT
CAACAAGTATTAGAAACTGGACATACTCAACTTCATATTCATACCAATTTTCGCACCTTCCATCAGCTCGTAGGAACACA
CACACACACACACACACACACACACAGTGTATGTGATGAGTTGTGTTATGGAGCTTCGAAAATCCATTCTTTGTTAAAGCTT
AGCGCTGTGATAGAATGATTTTACAAACTAACGTAGTTTGTTAATGCGCATTGCATGCCAATATTAAGAAAGCTGATAAT
CAGACAGAATGACGTAAAAAGGAAATAAGTTTCCGTTCACCGTACTGCTGTCTACCTGCATCCATAGTTTTGCATACATA
CACACACACACACACACAGTGTATGTGAGTAGGCGTGTGGTAATGCTGTAAAGGCTCATCTTGGACTCTGCAGCAATTTT
GATTGACATGCCCATCACGACTATTGCAGTTATTACATAAGGAAGTAATCGGTTTCCGAAACACTGAGAACTTACGACAG
TCCATCCGTTTCTGCATTACACACACACACACACACACACACACACACACACACACACACACAGCGTATGCGGCGATGTTCACG
AATGTCTGCAGCATGTTCATCTTTGAAGATGCTTCAACGAAACTACGGTTATCACACACTCTCACTTATTTATATTGGAT
GTGTTGAGAGCTTCAGAACTCTGATTGTAACTATTACAACATAATAAAACTCATACCAGACACATTTGTTTTACGAGTTT
ATTCATCATTCGCTATGCGCAATGGCTTTCTGTACCCTCTCATTTTTTCCATCAGTGCTTGGGGTATACGTAGAAACCAT
AGTCCTAATACAGTAGCAGCTATTACTATGAGCATCCACAAGCTGTGTAATTCACGTGACGCTTTGCGGGCCTCATGGTT
ATAGTTGTCAGATGAATTCAGTGTGTTAAATGGTACTGTTAGTAACGGTTGAACGTTAGTTGTAGGTTGCATTGTTGTAC
GCGTTGGATACGAGTAGTTCTTACAGTTTTCTGTATTAAGTCTCCATGATTTATGTTTACTACTGGTGCTGTTCCATATC
TGTGAATGACTTGTAAGCGAACATTGGGTAGAGTTAATTCTGCATTGAAAAGTAGGGTTGGCCGTTCTGGTAGAGGCTAC
ATATTCTGAAACGACGTTGGTACCATTACATAGTCGGACGGTAGGTTCAAACACAGTTGCGCTACCCGTGAGTAAAGATA

CCATACCCGGGAAACACACGTGAAAAACACACTGTACTGCAAAAAAATTTTTACCGTTCAAAATCAATACGTATATGTCA
CCCATTGGACTTAGCTGTGTTGATGTTGTAATGTTCATAGTGTTGTTCACATAAACAAAACAGTCGTAGGTTCCTGTATG
ATTTACTTCATATTTTCTAATACCCAGCGCTATGTATGATACATAACCATTATTTGTACGTTGCGCATGACCGTTTGGTG
TTCTTGAAGTATCTTTCACTTGTACGTCATATTCGAGTGTGTCTCTACCACCAACAAAATGATACATCCAAGTATCCACC
ACCGAACCGTTACAAGAAGTGAATCGTATTCCTCGCGTACAATTGCAATCAAATTGACATGTAAATAAAAACTGATTACC
CAACGGCAAAACTGTTTGCGTTGTGGTGGTCCAAGACATCTTTGGAGCATCTTTTGGAATTTTAAAGTCATATTCATTAA
TGTTAGCTAGTAGGTAACAAGGTAGAACCAGAAGAGTTACTGCAGTTTTACTATTCATACTTGCTGATTGACAGGACTTC
AGTTACATTGACTAATCCTTTATATCCTAACACCTGTAAATCATAATAAACCATATGACAACCTTCAGTTAGGTGCATAA
TCATCTTTATTCAAAATCATTACACGAACGCCGTCCTTTTCCGGGGCACGTCTGTGTACACCACGTTCGTAAATCTCGGC
TTGCACGTCCACCAAACGCAAGTACAAAGAACCATGACAAAAGCAACCACACAACCCGTAACCGCTGCAATGTCCCAGGG
ATTCATAGTGACTGTGCTCCGCTACCCCGTCAAACCCGTTCTGATACACAGCACGCTCCCTTATCCGATACTCACTGTTC
GGATGGCCTGTTTGATACCACGTTGTTACACAACACCCGAGTCCCAGTTTCCGCGAACATTTCCTAGGCAAACAAACGAA
ACTCCACCGTGGCCACCATGAAGACAGACGATCGGATGATTGTTGCCGTACGAACCACCACGGGCGCGCTAATTATCGTC
GGGCTCATTAGCCTGCTAATAGCCCTCATGTGGTGGTTACTCGCCAAGTAGGCATGGAATGCAAACCCGTGCCTACGGCA
TGCATTCTCTGTTTAGAAATGTTGGGGATCAGCGTCGTCCTCGTCGGGATCCTGGGAGTCGCGTGGACCTGCCTGCAAGC
GCTCGATCAGAGAGCGGCGAGCAACCGCACCGCGGCGGTTACAAACCGCACAGAGTAGCGAGCTGGGGCGCGGGTCGTCT
CGGTGGCGGTGGACGAGGAGGACGCGGAGGCAGGGAGGACCTAGCAGTGCCAACGGCTAGGACCTTGCTACCTTCGTTTC
TTCCCTTCTGATTACGACATTCCTTAGTGGTGGCGGTAGCTTTCTCCGTCGTCTTGGAGGAGCGCTTGGTTTTGCTGCGA
CAAAACAGCCGTGTGCAGAGAGACGCCTTTTTTTTGGCCTTCGGAGCGGCGGCACCGTCGTCATCGCCGAGTTTCTGCTC
TGCGGAGCCTTTCGCGTCCTTGGCCTTCTTGCTGCTGTCTTGCCGTCGAAGCTCGGGAAACACCACAAAGCGGCGACGCG
GACTCTCCCTGCACACCCTGTACTTCAAGCGGCCGGACCATAACGGCGGGCGGTAGCGTCGTGGAGATAGAGTGGTGTAA
AACTTGTTAAACATGTACAGGCCGAGGCCCGTGAACACGGCTGCCAGGAAAACGATACATAACGTAGTGTGAATAAACGT
GCTGATGAAGTCCATCATCCAAATCACGTTGTCCATTGTGGCCTGCTGGGGAAAGTGCTCCAAGAAATCTGCATCATGTG
CATCACACATCCATTTTATAACGCAACCGGTGATGCGGCAGGGGGGTTAGGCTGCGCGGGCTGGGCCGCGGCGCCACGCG
CAGAGTGGGTTCTCTTGCGCGGGGGCGGCAGTAGCTCACGTGGCGGTCGGGGAGGGCGAGAGAGACGGCTGCCCTCCGCC
GCCGCGGCCGCGTGGTGAAACAAAACCAGTGATTCCAGGTGAGATTCAGGCGGGACAGGCGAAGCAGGCGCGGGGCTCGG
AGTCGTAGCAGTTACATGGGGACTGCCGCCGCTGCTGCCGCTGGGTTCTTCGTCGCACACGGGGATATACGTGAGGCGCG
AGCGCTTGCGTCTAGCGGGGCTGCCGTGCGGACAGGGTTCGTGCGCCGCTTGGTAAGCTCGCACGATCGCCTTCACGTGC
AGGTAGTAGTGGTAGCAGTCATGCAGCGTCTGTTTGTAAAAGCGGTAAGCGGGGCCCAGCGGCCACGTGCGTTGCAAGGA
GAAGAAGTGGTACCTGCGGCGAAGAGAAAACCCGCAGGCCCTGAAGAGCTGGTCGAGCAGAGCCAGAAGGAACGCGGTGA
GCAAGGTGCAAAGAGCGGTGGCGAAGACCACGCGCAAGAACATGCTCAGGATCAGCAGAAAGGTCAGCATGATGGCGGTG
GCGAGAGGGCTACATGCATAGCAATGGTGACAGTAGTAGCGATAGTGACAGCAGCGCCGCAGCGGCACGCGCGTGCATGG
GATATTGATGTGCCCTGGGGAAAGGCGACACCGGGTTTTGTGTCGCCTGGGGCGTGGTGACAAGGAATTGTGGCGGGCAG
CGGTTGGAAAGGGCCAAGCTGACGCGAGGGGCCGTGCTTCGGCCTGGAAAGTCCCTGCGACGTCACACCGTTGACCGTTA
CGGGACTTTCCCGAACCCTTCCTTTCTCACTCTGGGTGCGCGCGGGCCAAAAAGCGTGCATTTGGCGCGCGGAGCTCCGG
AGACGAGGGCGCCGGGCGCTGCGCGCAGCGCTCCGGTGAAAATAAACCGGGGGATAAAAACAACGTGCTTTGGCATTGTG
CCAACTGTCAGGAGCACCCAGGGGCCCGTTAAGGGGGCACCGAGACGGCGACGGTTATCTGTCAATCAAAATGATTGACA
GATAACGACACCCACGGGCACGCGCAGTGTCCTCGCGTGGCGGAGCTGTCAATCACACTCGTGTGCGCGCCCCGTGTGTG
ATTGACAGCTTCAATGACGGCTCCGTCAATTCCCCCGACCGCCTGTCTATCCATCTGTTCCATAAAAATGCCGCCCATTG
CCACACACACCTAACGTGCCCCAATTCGTGTATTTGCAGCCCGCTCGCTCAACACGCCTTCTATAAAAGCCGCCTCCGTA
TACCACACTCACACCACTCGTCCGCTCCGCTGTCATCATGAGCGGACACCGCCCGCCCCGCGGCGTGGTCTCAAAAGCGCT
ACGCCGGTCTGCCCTGCAGCCCCAGCGTCCAGTCTGGACTGCTCCCACTGCCGCCGCGCCACTGGCACCCGGGTAACTTT
GCACAGAACATGGATATGTCTTCCCTCGCCTGCGCGCTGCCCGCCACGAATCCTGCGCCGACCAGCACTCCGATGCACAA
CCCGCCCCACGCTCCCGGAGCTCCCACGGCACAGGCTCCTTGCACGGCGGCGCAACAGCAGCTGTTGGCCCAGGTGCCGG
CGCAGCGTCACGCGCTGACCAGAGAAGAATTCTTAGTGATGGCTTCTGCATGGTCAGGCTGCTTCTTGGCCAACACCACC
CCGCGTGCCACTCGCAGGTGGTGCCAACGTGAAAGCGGCCAGGTGCTCCCGCTCGGAAAACCGAGCGGTTACTATGCCCT
CGTCACACCTCGCGCGCAGATGCGGGACGTGGGAGCCACGGATCTTCGACAACTGTCACCCACCGACGACTGGATTGTCC
TCATAGCCACCATTGTGCACGAAGTGTGTTCCGAACCTGGCCAACAGCCAACACTCTGTCATCACGAAGGTCTGTTCATG
GCCGTGGCTAGCCGTTTCCGTGTCTTTCTGTTTGACCTGCCTCGCATGACCTTGCATCTCATAGCTGCCGACGCAGATGA
GTTCTTTAGGTACGGTTTCGGCGATTGCTGCCGAATGTACGTGTGTTCAGTGCTTCCGCCACTCACCACACAGCCTCCTG
ACGTGGTTGAGCTCTTAACGCGCTCCTGGCCCACACCCGCCGCCGCAGCCGCGGCTATCAATCACAAGGCTGGAGGCCAG
ACCGTGCATTTGGAAACCCCCGGCCGACTTCCACAGCCCTGGCTGCTCACAACATGCTGGGAAGAACTCGAACGCTGGGA
TCCCTTTCGGAAATGGCCGCACTCGGCGAGCATTGTAGCGGCCGTCCGGAATTACATCATTCAGCGCCTCTGCTGTACCT
GCCATCTCCTAGGCGTGGTGACCACCGCCCCCTGGCTATACACTGTGAAAGAGTGGTCGGCTGCCAAGAACAACTATTAC
GCCACGCGCTCGCGCGGTCCCGGTGGAGGAGTAGGACGTAGGCCCCCGGACGTGCAGGGAGTAGAGACGGCCACTGTGAA
GCGTCGGCCGAAATCCAACAAGGGCCGTCACGGTGGTTCTGGCGCGCAACACCAGCAGCAACCCCAGCCCGGCGCAGCCA
CCGCACAGAAAAGTCCCACCAACGCGGCGCAGGCTAAGAAATCGGAAGTGGATGAAGATGGTCTCCTGACCGCCTCGCAC
CCCATGATTATGGTCATCGTCATAATGGACGAACTGGGCAGCATCTTTGGGTACTGTACTCAAGACGGCTTAATCTACCC
TCTGGCCGAAGACTTACCCCATTTCCTCCGCGTGGGGCTTCTAGGTACCTTCACCATCGGTCGCATAGCCACCACCGACG
AACAAGCGGCCACGCGCCTGATACCTGACAGGGATCCCCTAGTCTGGGAGCGCCCGCGCGGACGCGCTGTACCTGTGG
CCTCGCAGCGGAGGTGGCCCACATGATCTGCCCAGTCAACTATCCTTCCTCATGCGCCCGGGCCGCTGGCGCTCCCAACC
CATAGGAGAATGCGGCGGTGAAGGCTACCAACAACAAACGTTTTCGCCCGACGAAGACGGCTGGCGCGCGCGCGACGCAC
AAGCGCTGGGCCTACCGCCCAACGCCGGCTGCGACACCCCTCTCAGCGACCAGGCTCAGCTCAGTCTAGCCCTGCGCCAG
GACAATCGATACTGGAGACTCAGCTCGGCCCCGGAAGAGGACGAAAATCCTGGCGAAGACGGTCACTTTAACGAGCTACC
TTACACCACTTGGGCTCCCACGGACTATAATCCTCAGTGGAACCACCCGCCGCCCGCTGCCTTTGGTGCCAATGCTATTG
TTGCCCGCCGCATGTCTCATCACCACTCTCGGGCCGGTCGCCCTCACCGCCGTCCTCTGCCCGTCCAACCCAGTGACCGA
GATCCTTCCCATTTTCCAATGTACACCACAGAAGGAGACCTCGTAGTGGCGTTCTAAAATAAAAACTGTTTTTACACCTG
AATCCATTCACCTAGTCATCCGTTCATTATGTGTCGGCGCAGCGCGCTGCAGAGGCTCCGCGCCGCTCTACAATACTGGT
GGCACAACTATTTCACACAGTGTATACGCCGCCAATGCCAACAACAGCCGGAGGATGAGGAGGAAAGCCAGACGAAAACC
GACGTCGCGAACGAACACCTGTCAATCTACTTGGCGGTTCTGAAGGACGCCATGCACCACCAACAAAACGCGTCCCTCGC

CACGATTGAAGGCTTCGTAAACCCCTGCTTTATGACCAACATGTTTACACCCACCCTCACCCGACGCTACAGCCTATAGC
CCTCTGTATCTTTTCCCAAAAACTGTTTGGTTCTCCTGTGAATAAAACACGCACACGTGCGATTCTTTCATGTTATCGTC
TGATTTATTTACACAGAAACGAGTCGTTAAAGACACATCACGAGTAATATAACTAGGGACGCGATCCAGATTCCCAAACT
TAAGCGCTTCGTGGGGAACAGCATGGCGCGGCTCTTCTGGGTCTCGTACGGTACACGTACACGCAAACAGTCTTTCCTTT
GTTTGTTTTGTGTTCCCCCCTTAGCGAAACAGGGGCACCGGTGGAGGGCAGGCGCCGAGGTTCGCGGCTACCAACTGCTG
TCCGATGACAACCAGGGCGCGACCCACGAGCGAGGTGTTGGCAGACGGCTCCTTCTGGTTCTCTCTAGTCTGACGCCGCT
GGCTGGCCTCCGATTGGGACGACCGCGCGTTGTTGGAATCGCTCTCAGTCGCGGGCCGGCGTCGGCAGCGGGCACACAGC
ATGGAGAACCGGTGGTAGGTCACAGAAGAGATGATAGCGTCGCCGAGAGCACGGGCCGCAAACGCCGTCAATGTGGGTTC
GTAGAGGATGATGGCCGCGTGCCACATGCACGGAGGCACAGTGCCAGTGGTCGCCGAACCCCGAGTGTGGCGTGCCAGAA
TCCCTGGGGGCGCAAACATTTATACGACTTTAGGGTGATGTCAGTAACCACTGTCTGTCGTCAGCGCCTCTTGCCCAACA
ACTCGGTGACGCCCATGCTAGACCGCGGTTTCCCGGTAGGCAGGACGCACGTCCATAACGATGGGGCAGTCCTCGGGTCG
CCAAATCTCCACCTCGGCCGGGCACGGCGTTCCGTCTTCTTCGGAATCCGAGTCCGAAAAGGTGGGTGCCATGCGTTCGG
AGGCGGGGGGCACGGGCATCTGGAAGTACTCGTAGCTCTGGCGCCAGAGAACGTTCCACCTTCGCAAGCATTTCCACCAC
GCGAACGAGTTTTCCGCGGACTCCCCCGACGAGGACGCGGAAGACCCCGACGACGAGTCCGTCAGCCGCCCCCGGCGGTA
CAGGGCGTGCTCCCGCCAGTTACCGCACATGCACCAGAGGTTGTGCAAGCGATGCATCTGCCTGCGCCAGCGCAGGGCCG
CACGTTCGCAACGGTTGCGCCAGCGCAACTGAAAATGAGGCTCTTCCACGCACTCCACGTGCGGGGTCCAGTCCCCGCAG
GGGCACCAGAGCGCGTGGGACTCGCTCAGCACCGCCCGCCAGCGGAGCTCGGTTTGAGCGCAGGAGTCGCCGAGTCTGGG
AATCAAGTCACGTGGCTGCCGGCTCATGGTCTAGCGGCGTTTTCCAAAACGCCGCGTACGCCCTGTAAACGTTTTCGTTT
TATTTTTTTGTTTGCTTTGCCAAATCTTATGACCGGTCCTGAGGATGCGGGGGACGCGCCGTCAGGAAATTTCCCGCGGC
TCTCCCCTCTGTCGTGTGTTGCTACCGCTGTCGTCTCACAGGTTGTTCGAGGGTCGGTACACGGGCAGCATGCCGAACCA
TTGCCAGCTGCGGCTGCTAACGGTGCTAACGCCGCTGTGGGCGCCGGTGGACACCATGGACGAACGCGCGGTATTCCGCC
GCCTTTTAAGCTTTTTAATGAACCGACTGAAACACGACGCCCTGGCGGTGTTGCTGGGGCGTGCGCGAGGCTTTTCTGTC
AGATACACGGGCAGGCTCGCCAGAGTGATGCGATTTTCATCTAGGCTTAGGGCGTGTTCTCTCCAATTGGCACATAAACA
CCACAGGTCGTGGTTCTCTCGTGCCTGGAGCAGCCAGCGCATGGCCGCTCGCTCCATCCAGTCGCTGCCGCTGATCTCCT
CGCTGTTGGCCTCGTCGCTCATAACAATGTGCGATTTCCAGTGTCCGCAGACGCACCACAGACCGTGGGATGAGAGGAGC
GAGGACCGCCAAGCTTCCTCATTTTCCATGTGTTCTTGCGCGCTTTTCGCCCGGGGCCCCGCCGCGGTCATCGGCACAGA
CGCGAAATGACAAGGAATGACAAGGAATGGCAAATGGACACGATCTCGCAAATGGACAACTCTCTTTATTCGGGGAAGGG
AGGAAACACTGAGGAATAACTGGGGGGTGTGGATGAAATGGGAGATTCGGGCGATAGGGACGAGGGTGTCAACGAAGAGG
ACGGTGCAGTCTCGTCCACCGAGGATGGTTTGGCAATCATCATCATGTCCTCAGCAGCCTTCGGACTCCAGGGGGTAGTA
AACGCGTAGCGCGGGGGAACCGCGGGCAGAGGCGTTTCGGAAACACCTCCAAAGCGCGCCCGCCGTCTCTTCCAGAACAT
TTCCCTGTTATTGTGCAACATCTCAGCACGCAGCAGAGTCGAAGACCGGCGCCCCCTGTCAATCTTCCCTAGCCACCGCC
AGGCGGTCCGCGCGAACTCCACCGCGAAAAACGAGACGGCGAACCACCAGAGACATTGCCCCACATAGTGAAAAAAGCTC
CACGAGTCCGGTTCCTCGAGTCCGCCATCGCGTCCCACGATGATTTCACACCGTCTAGAGGGAACGTCCCGGTCAGAAAT
GAGCCTCAGCTTAGCCTTGTGGGTGGTTGCGGGGTCCAGGTCGTTGCGTCTGGTAAAGACCCGCCAAGTGACAGGTTGAT
TAATGGTAGTCTCGTTGTTCAGCGACAGGCCGCTGATATGCAGTTGAATCCGCTGACTGAACCCCAGGCGGTCTTTAATT
TTGTAATTGAATTCGATGACGTGAGGTTGATGGCTTTTTGGTACCAGACGACCCGTGGCTACCATGAGGTTGTTTTGCTG
CCGGCACTCCTGCTCCACGATGACGTCTTCGCCGGGTTCCACACGACACCCGTGAACCGTACCGGTGATGTGGAGCACGA
TAACCATGGCGGTGACCACCACCCACGTGCATGTTGACCAAACCACAGATGTGTGGATCATCGGCGATGAAGCCATGGGC
GTGTGGTTCGTTTGGTGAACTGGACGCACCGGACAGTGTTGTTCGGTGGAGTGAATCAGGGGAAAGCTGGCAGGCGGCCG
CGCGTGGCGCTCATATGTACCACGTCTGTCAGCTTCGCTCAGCCGGTTGGTACTGCGAGCGTGACGTGGTTAACCACAAA
ATGAATCCAAAGGCGGGACGTCTAAGAGTCCGCGGGGAAGGCCGTGACAAATCCGTGTAAACACATGGAAGAGGCGCCGT
GTAGGTGATGTGACAGACGTGACTGGGAGGCTTCTTGTTTTCTAGGATCGTGTACCGAGGCCACGGCCACAGGTATTGCA
AACCGCTTTCGATCCAGTGGCGCCAATACAGGACGGGACGACGAACAACTTGGCGGGGAATGAGTTCCACGGAAAGGATG
ATAATTAATACTAGACAGAGGAAGACATCCAAATACAGTGTTGTTCGTAATCGCTCATCCAGAAGGAAGAGTCGTAGAGA
AAGGGGTCGAATGGAGAGAACGGGGGGGCACCGATGTGCATCCACCACTCCGGAGGCGGAGAATGCCAGCTGCGAGGAAC
ACTGAAGGCCGTAAAAATACAGGCGACGCGCTCCGCACCGGCTTGGAACCTCCGCCGTAAACGTGATTTCCAGCCGGGAT
GGAAACGGAACACGAGGCCGTGAGCAGTTAGTAACTGCATAATCCACGCGCAGAGCATGGCCGTGGACGGCAGGGCCGTA
ACCGTGAGCGAAAATGATGTACGCATTCTCGCCGTGACTCAGACGGCAGCAGCTGCCGTAGTAGGAGGTTTCTCCCAGCA
CCTGTCCTTTTAAAGTCCAGCGCCGCACGCCGTGGAGCTTCACGAAGCTCGGACTGTTTGTACAGTCGCTGAACGCTAGG
TCCGTGGTCCAATGAGTATCAGTCACGCAATAGCTGTACGGTTTCTGTCGCTGGCGCCTGCTGCGCCTGCTAAAGGCGCT
GATGCGCGAGCGGGTGGGACCCGTGGTGCGCGGGGAGCGAGGAGCAGGGGAGTTAAACGGAACGGCTGAGGATAAGCGCG
AGGGTGGCAGCGGCGATGACGACAACGACGTCTTATGGACCACGATGGATTGTGCCCCAACCGTGGGTGATGTTAGAAAC
ACGGCCACATTCGCGGCTGAGGCGCATCCCAGCGTGATCCACCACGCGGCGGTGTGCCAGGGCGGGGCTGGCATGTTGCC
ATCTGCCCGCTACCAACATGTGTGCCCTCTCACCCCCGGAATAGGGGTTTTGGGAAACGTGTCACTCACGTACGTTGTTT
TAATACCGTGCCCCCGAAACCGTGCCCGCCTCAAAATCACACACAACACACACGTTTTAGTATGGCGCACTAACCGGCAC
TTGTTTAATGAGAACACAGGTTTAGCTTTCATATGAGAAACTGCAGCTCGTCGGTCTGCTCGCAGTCCTCACTCTCGTAG
GTACATGCCACAGCCACCTGCATGTGACTAAGATCCAGGGAAATAGCATCCGGACGCAGAGAGCCCCAAGACCTGGACGG
ACAGAGTTTGTTGAGCAGAGCGCGCAGCAAGTAACAGAATTCGCTGCGACACCGGTGCCCAAAGATGACATAGATGAAGG
GATTCACAAAACAGTGACAGAGAGCAATAGATTCCGTAATAACTAGGCCTCTCCTCAAAATTCTGTCACTGGCACAATTC
CACGGAATCTGCATGTGCTCTAAACTGTGAACGAAAAGCAGAAGGTTGTAGGGCAACCAGAAAATAGCAAATACAATAGC
TACCGCCAAAATAATATTCAGAGAGGTCAGGCGATAACGCATCCGTTGGTACGATTTGGTAAGCTTAGTGAGTTTCAAGT
ACCAGTAGATAATGGCTGCGGTCGGTAGAATTAATGTGCAGAGATTGACTTCGGTGTTAATAAAAATAACGTAAAAATTG
GTAAGTTCTTGATACTCGGCCACACAGACGTTGTCTCCTTTTCTTAGAAACACAAAGTGAGGCAAGGCTAGGATGATACA
CAAACCCCACATGAAAACGCACCCCGCCGTGGCATTTCGTAGCGAACGGTAATCAGCCCTGTCTACACCCCAGAGGATAG
CGTAGCATCGATCCAGAACAATGAGAAGCAGAAAGACAGTGCTGGCAAACAGGGCACAGTAGAAGATAGCCGTCATGACC
ACACACGAGACATGTGACATGATGTTGTGATCAAGGAGATACATGATCCAGAACGGCAGCGTGCACACAAACAGCAAGTC
AGAAATAGCTAAGTTGAGGAAATAGATGTCACATGCAAACGTGATGCGGCGATTCCATACGATGGTCATGAGAACCATAA
AGTTTCCTAGAAGTCCGAGTATGAAGATGATGGTGTAGAAAGTAACCGTGACGGACTTGGACATGTGCACAATGTCGGTC
ATGTTGCATGCCTCTGCAGCATCATCATATTCAAAGTCGGTACTAGTTTCAAAGAATGTGGTTAGCGTAGTACTAACAGT
GGTTAAACTTGTTGTAGTTGCATTGCTAACTGTTGTAATTAAACTGGTTGTAATGTTAGGTGTTGAGTTCGACACAATAG

ATGAAAGCGTTGACGATAAAGTGCTTGTAATGTTTGAAGATGTTAGCGACGATGTACTCAAAGTTGTTGTTACATTACGG
GTGCTGTTGCCAGTACTAGTGAGCGTAGATGCTAAACTGCTCGTGACGGTCGATCTGGACGTTGAAGTAGATAGTGAGGT
TACATTGGCAGGGGTGGAACTCGTGCGAATCGTTGTAGTTGTCGTAGTATTCACGGGGGTACTACTTTTGCTTGTGAGAG
TGGTGGTGGTAGAGCTTGATGTAGACGGAGTTGTGGTAGCATTTGTGGTGGCGGTCATGGTGGTAGCAGTCATTCTCCAC
AGCGGTCTGATGAGTCGCAGTGTACGTGTGTGCTGTGGAGTTCTCAGAAACCTTTTATTTGTTTTCTAGAACCCGCCCCA
AAGAACACAACGCAAAGCCAACGAACAGGGTTGTCTGAAAATAGTTTATTCTTTTTTGGCACGTATTTACAAAAAGCATT
CGTAGGATTTCGGAAAGGTCTGATCAAACCAAGTCACCCTCTTTTTACGTTTTTTACCCTTTGTGTTTCTTTCCTGTACA
CGGAGCTCCACGGTATTCTTGGCAGCCTCCAGGTCCAGCTCGTCGTACGGATGATACCGAAACAGGGAGTACACTTGCTC
CCGAACCTTATCGCTAATCCCAAAGTAAATAATGGGATTAACGACTGCGCGAAGGTACACCGCAGTCGCCGAACCCATGG
AAATCAAGTGTTTGATCTCCTCGACCGTACAGTTGGAATTGCCGTAGATCCTTTGTAATATGTTGTCGATGATGAGAACG
ACGTGAAAAGGTCCCCAGAACAAAATCGCTATAATGGCGATCAAGAGCATCATGTCACCGGTTTCAGCGTTCAGTCTTTT
GTTGCCCCACGTCGCTCGCCTCATCTTGTAGCTCACCGTCAGAACAATGCCCACCGGGCTGATGAACACACAGGCTATCA
TCAAGAAGTCTAGAAGAGTGTGCGCGGGCTCGTTGAAATGCCAGGTGTAGTTGCCTAAGATACATGTGTGCCCGTCGTGG
GCATTCCTAAAGACGTAGTAGGGCGAGGCGACAAAGGCGCACACCACCCACGTTCCTAGAGAGTTCTTCATAGCTGCCTT
TGTGCTGATAGGCACCATCCAGACCAGGTTGGAGTAGCGTTCCATGGCGATGGAGATGAGGAACCAAGCTTGCATAAAGA
GGGGAACGTAGAAAGTCACGGATAAGCCGATACAGGTCGCTCGATCCAACGCCGTGTAGTTCAGTTCGTAGTACATCCAT
AATGGCAACGTTCCAGTAATAATCAGATCCACTAGAGTAGTGTGCATAAAGTATACATCGCTGGCGAAACTGAGACCGCG
TCTTTTCACAAAAACGAGCAACAGCAGAATGTTGCCCGTCAAGCCCACGAAAATGGCTATGCTGTAAATGATGATTTCCA
AATAGCGTAGGGGCCCCACAGTTCTATAGTTGTCAAAGCCGCAAGATGAGTTATTCATGTTTCTGTCAAGCAGTTTGCAA
ATCGCGCAAGCTGTGGAGTCCGCACACCCGCAGTTTCTCACTTCACTCAAACATGTGTGTGACGAGTGATTCAATTTCGG
TCAGTGCATGTGTGTTTTTACAACATGGATTCGCATTCTCTAGGAAAAGGGTCAGGGACCTGTTCGTTGTTAATTTTCAT
GTCAACCATGAGCTGAGCATGTTCACCGTCCAGCGTTTGGTAAGGTATTCGCATGAACAAATGAAAAATGCTGCGTAAGA
GTTTTTTGCTGATGGATATGTAAAGGATGGGATTGAAAATGGGCCGGGCGAACACGAACGCGATACCCATTATGCGAAAA
AAGTGTTCTTGCTTAAGGTAGTAGCAACTAAGTTCGCTTGAGTCGGTAACTATGTCCCGAAACATGACGAGATTAAACAG
TCCGTTAAAGAATAACATGTTGATGGCCATGGCGTCCAGGTACCGGCTGGTTCTGGTATTAAGCTTTCTGTTACCCCAGG
AAGATGAATGAATCTTAATGGACAAAAACAATACCACACCGGCCGGGACCACAAACGTCCAGAGGTTAATGGCAACGTCC
ATCACGGTGCGGTACGGACGGGACACATGCCAAGTAAAATTCCCTAGAATGCATGAGTCCGGTTCTTGGGTTTCTGTTAA
CATGTAGTAAGGCGATGTTACCAGGGTCACAATGGTCCAAAGCAACGCGCAGCTGATGGCGGCGGTTTTACAACTCACGG
GTGTGTTTCTAACTAGGTTCTGATAGCGTTCCACGGCTATGGCAACTATCATATCGCCATGCATGAATAATGGGACGTAA
AACGCGAAAGATAAGCCGATACAGTCCCAACGCGATGCTTGGGTATAGTTCAGTAGATAGTAAACCCATCCGGGTATGGT
TACTAACATCATCAAGTCCACGAACAACATGTTCATAAACAAGATATCATTCGCGAACCACTGAAGGCGTCTCATGTAGA
TAACGCAGAGTATCAGCAGGTTCCCGCCGATACCGAGAATCATACCCAGGCTGTACAAACTGATGGTGAGGTCAGGGGTC
ATGCCGTAGGCTTTCATCGTGTCGTTTACGTGGCAGACTTCTACAGTCATTTTGCGGCCGTTGTTCGCTCGCTCCCCAAA
ATCCCGCTACGTGGTTTGCGACCGGCAGTTTTTATGTGGTTCCAAGGATCATATACATCACAAAGGTTGTTATCCTGTTC
ATTTTTTACAAGAAACAAGTGTAGTCACGGGGCGTTACGTCGGCACCCTGGATGATGAGCGTAATGTCCTCGCCGTCTTT
GGCGCGCTTGTTCTTACCGCCCTTGCCCGTCAGTTGTACGGCCTCGGTCTCTGTATCGCTCTCTGGTGACGGCTGTCGAT
GAGTACCTTCCATGTCAATCTCGCTGTAAGGTACGCGTCGGAACAGATTGCACACTTGCGTGCGAAATTTTTGGCTGATG
ATGAGGTAGACGATGGGGTTCCACACCGAGCGCAGAAACACCAGGCTTTCCGTCAGCAGCCCAATAAGATGTCTTATGTA
ACAATGTGATGTCTTATACAATGTCTGTATAATGTTATCTATCACCGTAAACAGGTGGAAAGGTACCCAGAAACACAAAG
CGACAATGGCCAACTGAAGCAAAATCAAGCTGTTTTTGTTATTTAGTTTTCTGCTACCCCAAAACGATCTGCTCATCTTA
AAACTATAAGCAAACACAATAGGTAATGGAATTATAAAAGTCAAAATGGTGATAACGCCACTCAGTATGCTGTGGTACAG
TCCCGTCACGTGCGACGCGTAATTCCCCAGAAGACATTCGTCAGAATGGTTTGCTTTTCTGAAAATGGAATACGGTAACG
CCAGGAACACGGCCATGAACCAGGCCCATACGCATCTTTTAATCGCGGTGTTTTTCTTCACCGGGGTTAAGTGAATCAGG
TACGAATAACGCTCCAAGGCGATCAAGATGAGCATCCAAGATTGCAAAAACAAAGAGACGTAGAAACCGTACGAAAAGGT
GATACAGGCGAATTGGTTTAGTGTCACAGATTGCAGCGCATAGTGTACCCAGGCCGGCAACATGCAGATGCAAAACAGGT
CCGCTATGGCAGCATTCAAAAAGTAAACGTCGCTGGCAAATTTAATCTGACGCTTGGTAGTTAACACGAGAACTAGCACA
AAATTCCAAAACAGTCCAATGATGAGGAACACACTATAGACGACGATCGTGGCGGAATGATCTATTCCAAAGTGACTTAG
AGTCCCATTGAGTTTGCACGTGGCGTTTGTGCTGTTATTCGCCATGGTTCCCGGCTGGCGCGACACACACTCGATCTAAC
GAAACGTTTTAACCGTCGATTTCTAACGCCGACTAACGGCGAATTCCAGTGTCACCGACCTCGCTTTATATAATTAAATC
GCACGTGTTTATCCGAGGATGTCGCTGTTTTTTATCATTGACGATCACAAAACAGAAGACCTTGATGACATATTACATCC
TGTTATTCGGAGGTGGCTGCCGCTGCTGCCTCTTCTTCGTCCAGCTTCGTCGGAGGCTGTGTGGCGACCGGAACGTAGCC
TCCCCACTGTTTACCCGAAATGCGCCAGAAGAACGGGCGTAGACTGACTATGAACTCGTGGCGAACCACGGGACAGGCCA
CAAAGTACACCAGAGGGCTGATGCCGGCGTACGCAAAGGCGAGAGACTCGCACATGACGACAATGGGTGCCCAGTGGATG
GCGCTAGAGGGCTCTACGTGACTAAGACTGGCCATCGCGTCAATCAACAACACCAGATGAAAGGGGAACCACACGATGGC
AAACACCATGATAATGGTGACCAGCAGGGCGCATCGGTGCAGGTCGGCAATCTGGTCGGCCGCTGCCCGGCGCTTGGCGT
GGTAATAACAGGCGGCTATCATGATGAGCGGCAACCAGAAAGTGCACACATGCACCTCCACGTTGAGGATGGCAGAGGAT
TGACTGGGCATGACGCCCATGTTGCCGAGGCATTGCTGCACGTCGCCGCTGAAGATGGCAAAGGGCGAGGCGGCCAACAC
GGAAATTAACCAGACCGACAGTCCCACCACCTGATTGTAGCTGGCGTGGACGGGCAGCGGCTGGCGGCACACGATCACGC
TGTAGCGATCCCAGGCCAACAAAAAGTAGAAGGTGGCGCGGGCAAACGCCGTCAGATAAAACACAAAGGTAAGACCCCGG
CACCACGGATCCGGTAAGTGTCCGCCGGTGCTGAGGTGGTAGGTCCACATGGGAATGGCGAAGGTAAAAACCCCCTCGGT
GATCATGAGATGTCGAAATAAAGTGTCGCTAAAAGCACCCGCATGGTGTCGTCTCTTGCTGTTCTTGAAGCCCAGATACA
AATGTCCCGTGATCCCGAACAGAAAAACGCACGTGTAGCCGGCGATAATGGGCCCCGGGCCAAGAGGGGTCAGGGCGGTC
ACGTTACGAACGCTCGAATTCATGTCGGCCTCGAGGTTTCTAACGTGAGGTCCAACGGGTCCCTGTAGAATAGACGGCTT
TCGGAAGATCCACATCATGATTAATAAGCCGTGATCTTGGCACACGGCTGCGACGTCATTTCTCCCAGGGCATAAAGCCC
GTTACATTCCAAAGTGCGCCAGCACACCACCATGCTGCGCCAATCCTATCGCTACGCCTCGGGTCCCGAGATCAGGCGGA
CGCTGTGTGCGCTCAAAGACGTTTTCACACACCAGGATTCGCAGACCGCGCTGCGAACCCTGGTGTCTGTCAAGGCGGAA
TGCCGCAGCAGCGTGCCGGTGTCCGCCCCGTCCGGCTGGCGGTTGGTCTTCGCGAGCTTCCCCAGCATCTTTGGGAGCAG
CACCCCGTGTCACGGTCCCAATTCCCCATGGGGCAAGCTGATCTGCTGCGAAGAGCCGCTAGAGGTTCTGGGCTTCCTGC
AGTTTTGGAACACTGCCCACCATCCGCAACAGAACGACGGGCTGTCGGACGAAGACGATGATGACTCCGACGACGAATGC
TGCGACGACAGCTGCCTGTACGCGGAAGAGCTGCGGCCGCCGTGCAAAGCCAGGCGGGTCGTGATGATGGGACGCTACGA

GAGCATGTGGTTGCTGGATCGCGACCAAAACGTTCTCTACTATCTCGCCAGCGGCCTGGATGACTTTGCGCGGCACGGTC
TGCTGCACTGCGAGAGCATCTACACCGGCAAGTTTCGCATGCCCATGTTAACCACAGATCCCGATGACATGATTTCCATG
CTTCGCATGAACGATCACTCCACCAGCCATCTCAAGGCGGCCATGGCCAACTACCGCTGGCAATGCGTACCGCTCAGGAC
CCCCGGGGAAATGACCCGTCCCTTGATGATCTGCGGAGAGACCGACGACCTGAAATCATGTTGGCCCTTTCTGTGCATGG
ACGAGTGCATGTTTAACAATCTAATACAGTTTTTCAGGGACCGCCTGTGCTGTGAGTTTCTAACCTTTGGCGTGCTGGGC
GAAGTTCTGCCGTCGGGAGTGTTTCACTCGGAATGGGTGGTGCTCATGGACGTGTTTGGAGAGTTTTATACTTTCGACGT
GTATCGTCGCGAAATGTGGAGATTGGCAGATGACCTGGACTCGTTCCTGACGATAGGCATGTTGAAAATCTATCAGGCCG
GTCGGCGTTTCAGCTGCGCGGTGTCTGACGCCGAAAGGCTGGAAGTGCCCGCGCGCTGCCCCCACGGCACCTATCAGTTC
TGGGATCGCTATCTGTACATGGACCGGGGAACGGCGCAGCACAGTTTTCTCACCCGTTTCAAGTGGCTAATACGTAAAGA
TCGTTTCAAGAAAAAGGCCGTCGGCGGCTGTCACGTCCGCAACAGCTACCACGAAGTGGCCGGAACGGCCGATGGCAGTT
GGGAGCCCCATCTGCGCTGCGATTTTCCAAAGGCCACCGAGGTGGAGGCCGCGCACAATTTTTGGCGTCGCGTTTCAGAA
CATGTGCAGTCCAGTGAGGTTAAGCAGGGTGTGCGGTGCTATTCCATCTGGAACCATCTGACCCCCCCACTAGAAAAAGA
GATTGCACGACTGAGAGCGACGCTGACGGACGAGAACGGTAAGTCCTCCTCCGCCTCATCATCCAGAGATCTCTTGACCG
AACAAACAGTCACGCCAACGGCCAGTCCTCATGGCACATTAACCAGACCTCATAAACAACAACAACCTGCTGAGAAACCT
CCAATACCTCAGAAACCACAGCGCTATGAAAGTCGTCACAGTGATGACGATCTGTTAGAGATTGACGGGCCCAGCACTAG
TCACCGCCAGAAAGAGGAACACGAAACCTCCGATGACACACAAAGCAACGATAGCGGTGAAAGAAGAGAGGAGGCAGATG
ATGAGGACAGGGACGAAACGGACGGAGAAGAAGAGCCCACATGCGATTACTATACTGAAGTCCTCATTCGGAGAACTGCC
AAAGCCGCCCTTCTGGAAGGACGGCCATTTCCTCGCCCAGCCGTCCCCCGCCCCGGTTCGTATTTACCCCCTTGGCTGTA
AAAACCCCAAAGTGAATAAAGAGAACATATGAGCGGTCTGATCCCATGTGTGGACAATGATTGACTGATTGATGGATGGA
TACCTTGTGTGCTTGCTTGTTTGCTTGTTTGCGTGCTTACACTCGTCTCCTGAGAGCTTTCTAGTATAAACAATACATCC
CGGGGAGTGTCCTTTGGCGCCATTGGCTCACACGCCACGGGCTCACAATCCACGTCGGTATATAGCAACGCGCTGTGATG
ATGGTCCCGAGCTTGGCAAAAACGCCGAGAACCCCAAACCTGATCTGTGAACCGCTCAGCAGTTCGGACCGAGATCTCAA
CACATCGGTCCAAACGCTAGATCGGTTAACAGAACGAGGTGAGTATCGTCGCTCGTCGTGGTATTTTTTCTCCTGGTGTG
GCGTCCCGCGCGGCGATGGCGGCCCGGCGGCACTGCGAGCGGGACGCGGCGGCGGGGAACGATGATGATGATGGCATGAA
CTTCTACAGGCAGGTGCTCACGGACTTTCGGGACTTGTTCTTTTGCATGGAAACCTCCCAGATCGAACAGTACGTGCGCC
ATAACCACGGGCGCCGACTGTGCCTGGGACCTCCCGAAGGCTGGTACGTGCAACTGCAGCGCGACCAGCAGCTGCTGCAA
GCGCAGCAGGCGGCTAGAAAACTAATCTGTTGCGACGAACCGCTGTGCGCCCTGGGCTATGCCATTAAACTCCTGCCCGA
GCCTCACCGCGACCAAGATCACGACCACTTACGAACGGATTACGTAATCCTCCTGGGTCGTTTCTGCCGCCTCTACGCCT
TTGAGCCTCGCGAAAACTACATGGTGATCGTGGCTCACCGGCTGGAGGACCTGGCCAAGTACGGGCTGGTGCGCTCCGAG
TTCATCTACCGCGACACCATCCACAACCAGCTGCGCCGCATGTCGCCGAGCCTCATACGCTCCGGCCCGCGCCCCTCGCG
CACCATGCACATACTGTTTCTGAACGAAACCACCCCGGACAGCTTTTACGTCACGGCTGAGCGGATCGTGGAACACGACG
TGAAGCTGTACACTCCGGGGTACGGAGCGGTTATCTTGCGCCTCATGAAAACCGTCGCCGAGCTGCGGGCCGTGTGGCCC
TTCAGCAGTCTGACGGAAACCGAGGCCAAACGCTGGTGGTGGGGCATCAGGGCCAACCTGGCCACACCCTGGTATCTGTT
GGGGGCCACGGGACGGCCGCGCCAGGGCCGCACGTTTGTGGCCGAGGTGCTAGTGATGCTGGACTGGTTCGGGGCGGTGT
ACGCCATCCAGCTCGACGAGCCCAACCACCACATTCGGCGCGTGGCCAACACCATCACGGAGTTTTTCCGCATGGGGCTC
CTGAAACTGGTGTTCGGGTATCGACGCTTCGAGCCCGAGCGCCAGAAAAAAATCCGCATGGAACATCAATGCATTTGCCC
GCACGCGGACGAGCGCGAGATGGACCATCATCACGAGCTCTTAGCCCCTTACGATCGTGAGGAGCGTCGAATGCACTGTC
TGATTAAGCATTACGACTGGCTCTGTTTAACGGACCGCTTCGACCCGCACCAAGGCGCCTGGGAAAGGCTGGATCCCAAT
ACCCTCGTGCTGCATCGCTATGACACGTTGAGCCAAACCTACGTGCTCGAGTCTGAGATCGTCGGTGTCGAGGCAGCGGA
GCGCGAGGCAGCGAGCCATCAGGAAGACACTGGCCCTCGGCTCCACTGCCTCGTCTCGACGAAGTCCACGACGCGGGAGC
GGTGCACCGAACGCGTGCTGGTCGCGCTTGTCCATCAATCGCGGCTGGTGACGTACAACGATCCCTTTCCTTTTAAAAGC
CTAACGGGGTTACGGGAGTTTATCAGCGTGTGAGCAGCAGCGGCGACGACAAAGAAGAAGGAGCGAGCACTGAGGAGAGA
GAAAGAAAGAGGGAGACCATCACTATCATTATCATGTGGCGCACTCGCTGGGAAGACGGGGCCCCGGTCTTCACCCGCAC
CGACGACTTTCCATACTGCCACACTAAATATGAAACTTATCTGCGCGTCATGGGGGACTTTCGCGAGCTGTTTCTGCGCC
AGGCCAGCATCGACTGGGTTCGCACGTGGGTGCGCGAGCACATCGATCAAGTGATGTGCCTGGGCATTCCCTGCAACTGG
TTTGTGCACGTAAGGCCGGGCGGCCATCTCCCGGAGCTCCGAGACCAGCTCCTGGACGACATCACGTGCTGCCGCGAGCG
CATGGCGGTTATCGGCAAGTGCATCGTGTGGGTGCAAGACCGCTTCGTGGAGACCGAACTCGTTATGTGCATGGCGGGCG
ACACCAAGATGTACATCTACGAACCCAGCCAGGACATCCTGATCCTGGCGGCCGAGCACCTGGACGAACTGGCGCGCTAC
GGGCTGCTGCACGCCGAGTGCGTGTACCGGCAGCCCCACACGCCCTACGCCACGCAAGTGCCGCGGGACGTGGTGGAAGC
CCTGCTGCAGAGCGCGGGCGACGCGGAAAGCCTGGCCGAATGCGTGGGCGCTCACCACGATCAGGACGTGAACCTGCACA
CGCCCGGGCGCCACGCCGAGACGCTGAAACTGCTCGCCAACTTTGGTTGTCTGAGCGAGCTGTGGCCGTTCGAGGCCATC
GATCCCTACCACTTGACAGAATGCCAGATGTATCTCAGCCTGCGCCTGCGGTGTCGCTGGTATATTATGGGCGCCGTGGG
TTCCTACCGTCCCGGGGGTTTTTTTGAGGTTAGCAGCGTGATTTTTTCGATCGCTTCGGACGCATTTACGCGGTGATTG
TGCGGCCCTACACGGATCGCACGTCGCATGTGTGCCTGGCCCCGGGGAACGTGTACCGCCTGGCGGACAACCTTTCGGAG
TTTTTCAAGGGTGGGTTAATGAAACTGTATGTTCGGCGACGCCACGAACACCACCTGCGCCAGCTAGCGCGTCTGGAACG
CGGGGTGCGTTGCATTCACCTGGACGAAACTTACCGCCTGCACTTTACCGCCTTCGATCACGGGGTGTCTCGGGACTTTG
CCCGTCAGTATCGCTGGCTGTGTCGGCCGGACCGCTTTCGGGCGGACATGTTCACTACCTGGGACGGCTGGGACGCGTTT
ACTATCTGGCACTCACGCGTGATGCGCGGGGACTTCGTGGACAGACGGCGCCGCCACGAAAATGGAGACGAAGATGCTCC
TCCCCGCCGTCGAAGAATGTTCGATCCCGCCAACCTGCCGCCACCTCCACCCCCGCCTCCACCTCCCTCGCCGCCTCCTC
CGGCGGACAGAGAGCCCACACCCCCGCCCGAGGCTCCCGCGCCCCAGCCGGCGCCGCTCGCGCCCGAAGCGCCCGCCAAC
CAGCAGGCGCCCACGCCGCCACCTATCCCGCGGGTGCTGGAAAACAATAACAACAACTTTGATGAAGTCAACGAGGACGA
CATTGAAGTCCTCGGGCGCGACAACGAGGCTTTCGACATGAACGGCGAAGACGAAGACGACGATGACGACGGCGAGGACC
TGGGGTTCGACCTAGACGAAGACGGCATGTTCGACCTAAAGGACCTAGATGAGTGGCACGCCGAAAGAGAGGCGGCTCAA
GCTAACCGCTGGCGCCTGGGCCAGACGTTGATGAACGCGTACAACACCGAAACCCACGTGTTGGAGGCTGAGGTGGAGAG
CCGGCGGATACACTTGAACCGTGATCTGTCACCCGACTGGATTCACAGTTTTGACTTTCGAGCCTTCTTCCAGTAAGGGA
TGACTGAGACACACACGTCACATGTTAATCTTCCAGGAAGCTTTGGGGTGCTCGTATATTAAGCTTTGGTGTGAGGCGTC
GTGTTCAGATTTAAAAAATCCAGCCGCGGAGCTCGCAGCCCGCACACGCTCCCTCTTTTATAATCCCCCTTTTTGTGT
TATTCTTTTTTTTTTCCCTGACTATGGCACTGCTCTCCAGGGCCGCCGCCGAGGCCTGGGGAACGTATCTGAAGCAGCGA
GATGACCGTCCCGAGGACGTGATCCGTTGCGACTATGGGGTGTTCAATCTGCGGAACATTGTTTTTCAGCGTGCTATGAC
CATGTTGCAGGGCATTTTTCTGCGGCAGTACGATCACGTCTGCCTGCGTGACTATGTGAAAAAACACTCCGGCACCATGA

TCCCCTTACGCAACCCCAGCAGTTGGTTTCTCATTTTTCGCTGCGGCGCCGACATTCCGCAAGTGGAAGGCCGGCAGTTT
GGCACTGATTACATGTGTTGTGATGAAAAGCTGGAGGCGGTGGGGATCCTGGCCGTGAAACTGCCGGACGGGGATCGCAA
AAGGGTGCAGGAAACGACTTGTGTTCTGCTGATCGGGCACTATGGGTTTGTGTACGTGTACGACTGGGATACGGACGGGC
TTTTCCAGGTCACTAACACGGTGAAAGAACTGGCGCAATATGGGCTGCTGATGTGCGAATCCGTGTACCGACATTCGCAC
ACCCCTTTCTCCACCACCGAACCGCGCTACCAGGTGGAGAAGTTCCTCTGCCTCGATCCCACGGACGCACGGGCCGTGGC
GAACGTTGCGGAAGAATTTCACGGGATCAACTTGGTAATCAAAACTCCAGGGCGGGCGGAAGCCGATCCCTTGCTGCTGC
TGGGTAGCGTGGAAAACCTGCGTGCGGTCTACCCGTTCGCCAAGATGAGACCCACTAACTTCGAAGACCTCATCAAGTAC
ATAAACCAACGTCTGTGCTGTCGGTGGTACGTGCTCGGCATCACCGGCCGCTATGCCGCGTTCGGTGTACTCATGACCTG
CGGGGTCATCATACTGGACGAGTGCGGCGTGTGCTACGCCCTGCGGGTCGATGAGTCCGACGTGTACCGCATAGCCGACA
ACGTCCACATGTTATTCAAATGCGGGTTCCTCAAACTGCGCGGCATGCTGCGCTTCGACCGCGGAATGCGCGGCGAGGCC
CGCCTGGAGCGGGCGCCCGCCTGCGATCACTTTAACCGGCAAGAAGATCTCTCGTGGTGTTCCATGCTAGGTACGGTAAG
TCGCCAGCAGCTCGACAGCGCGCACGACTGGTTGACGCGCAACAGCCGCTGCTACGACAAGGTACCCACGAGCCCCTGGT
GGGGACGCACTGATTTGCTGCCAACCGGCGTGCTGCAAGAAAACCAAAACTGGTGCTTTCCGTGGACCAACGTCGCGGCC
TTGGCAGCTCCTCAACACGGTGGCTGGGAAGACCACGACGAGATCACCAATAAGGTGCCTCATCATCGCGTGTTCCGCGT
GCACAAAGGCTACCCGCCCAGAACCCTCAGGGGCGACCAGGACACGGACGAAGAGGAGGACGAGGATGACGACGAAGACG
ACGAGAACGAGCCGCAGGAGAACGACGAGAATCAAACACAATCTCCGTCGCCTCCGCTGGTAAATGGCAGTTTAGGAGAC
AGCGCGGACGAAACATCCTCGCTGTCGGGTGACGAGGCACCCTCCGAACAACCGCCTCAAGAACTGCTAGACTTAATTAA
AAAGGAATGGCAAAACGAGGACGAGCCGGTGACCTGTGAGAAGGAACTCTGGTTACGACGCGCGCGCAGGGCCGATGCTA
TGGTGACCTGTTCTTGTCGTCTCCAACGCGTTAGCATTTATGATCCCCTAAACATTGGTGTTGACTATTTTGAATAAACA
GTCTTGTAAGTTCGTATGTTCATGTGCGTGTCTCATTTCATGTGTGTGTTGCCTGCATGTTTGGAGCGGGGTAAACAATT
CACAGCGTATTTAGACGTGTCTACCCTCCTCCACCCACCAAAGTCGTGTGTGCACATCCCACACAGACACAAACAAACAT
GCAAGTCCAGTCCGATGAGATCGAAACGGTGCGCTCCATGTTTCTCCTGCGAGTCTACATCCTCATCTGGTTTCATAACC
TCCTACTACTGAGCGCGTGTGCAGTGTGCTGGCTCATGTTCCCGGAACGTTTCGCGCAGCTCTTACCGAGCATACGCATG
ACGCTTTTCTGTCTCATGCTGTCAATCATCTGCCTGGGTGTGCTGCACTGGGCCGAGCCGGAATTCCCCAAAGATATGTA
CATGTTACTGACCTACACGTTCCTCATAGCCGTGGCGGTCGTCGCGTCGGGGTTTCAATTTGAAAACCGCTCGATCGCGT
ACGCCACGCTGTCTATCGTCCTCCTGTTTGTCGTGCTAAGCGTTTCCACGCACCGCTTCGACAACGATGTGGAGTTCCAC
AGACCGTTTGTCATAGGCTCGTCTAACCTTCTTATCTGTCTCTCCGTTATTTTTTACTACTTTCCCGGTGAGGTTGGGGA
GCTAAGTGTCATGTTCGGAGGCTTGCTGGTCCTAGTGGCCTCCGTCATGCGTGACAGTCAGCACATGTTACACAGAATCG
ACTACGAAAGCTATATCCCGGGCGCCATCTGCCTGTATCTCGACATGATGTATCTATTCCTAACCCTGGTGTACTTTTTG
TCCACGCCCAACTCCCAGGAAGTCTCTAACGTCGTGTATAAGAGTTAGATCAGCCAGCGAGCGACGGCTAGAAGCCGTCT
CCGTCCACAACCAACATGGTTCAGCAGGAGAACGCCGAAATGGAGGCGTCCGATTCTCCGTTACTCGTGTCGGTTGACGA
GGCCATAGCGTGGTTCAAGCAATTCGTGCTCTGGATGCGAATCTACGGCATCTTCGTCTTTGCTCTGAGTATAACCTTTG
GGGTCTCCGGCCTGATCTGGCTGGGATACCCGCAGGCCGAGAACTTTTGCGTGGTAAACCACAGCGTGGTGCTAACCATG
TTCATCCCGATCCTCTGCATGCTCACGCTGTACACGCTGGGCGCACAGCATCCCAGCAACCTGACCACCTTGTTTTTTTA
CATGACTGTCAATCTCGTGCCGGCGACGATGTTCGACATCTGCTCCGACGGTTACGCCATCACCACTGCGTACGTTATGA
GCGTGGCGATCTTCATCTCCTTCACCGGGTTAACCTACATCGGAGGAATGGATGCCAAACGTTGGAAATGCATCGCCGGA
ATCTACACCATCCTCTTCCTCCTGTTTCTCATCTCCACGGGATTGCAATCCGTGCCGTGGTACACCAAGCTCGTGGTGAT
CATCTCTGCTTTCTCCATCACCTTCTTCTCCTTCATCCTGTGCTACGATACCTCCGTGGTCATCTACCAGGCTTCCTACC
TGTATTGTATTCGAGCGGCTCTTCGTCTCTACGTCACCACAGTGTCCATCTTTCTAGCTATTCTTTTTATGTTGTCCATG
CCCAAATGGGTGGAGCACGCTGCCAACAGCACTCGGACAACGTCTTAAACGACAAGACAGAACAGACCGCCGCAAAACGC
CACGCTCTTCCCCATCGACACCATGAACTTGTCGCCTGACCGTTCTCTTGACGACTATGTTTTCGTCCTTGGCAGTCGAG
TTCTCTTTTACCTCAAAGTCTTCGCCACCCTCGTACTTCAAGTCACCGCCACCGTCATCGCATGCATTATCAACTGGATT
CTCTGCCCGGCGTTTCTCGAGGCGTACTGCTCAAAAGTACCTCGTCTTGCCGCTCTCTGGCTTTTCGTCCCCATCGTCTC
CTTGTTTACCTTGCACCTTTGGGGTAAACAAAGATGGAGTCGCCAACTCACTTTCTCCATCATCTACGTCATCCCCAACA
CCATAGCCCTGTATATGATGACCGCCTGTCGCCCGTTGGACAAACTCTTGCTGGCCACTATCGTCCCTCTCGCTTTTTTC
ACCGGCGGCGTGGGGCTCGTGTTCCTCGTTCGTATTCAACCCTCCAACGGAGCCTACGGTCTGCACCGATGGCTGTTTCC
CGTCACAGTGGCCGGAAGCGCACTGGTTTTAGTGGCGCTGACCATTGTGCCGGTTTCGGACGCGGCCTGGAACGGTTGCT
ACCTAGCCCTGCTACTGATTCTCGTCGCCGGACTCTGGATCCACGATCTGGCCTGCATCGTGCACCATGATTGTTTCGAA
GCGGCTTTTCCGCTCACCATCCGCATCTACATAGAAAACCTCATCATGTATATTATCGGACTCATGATCTTCGACCCCGT
CTTTTGGCAGATGACGGACATGTCCGACATCAGTCTAAAGTTCACCCTGACCAAGTGGTGGCTGGAGCCCGTTGTCCCTC
AAAGCTAATGACGTGTAAAATCCACCCCGCTGGTCGCACCATAAGATAGACTGTTTGCCGCGTAGCATTTAGAACGATCG
GAGCGCAGCATCTGAAGCGAACAACACACGCGCGACCCACATCATGAACCACCGCAGCTATGCAGACGCGGAACTGGCCG
AAAACTGGATGCACGAGAACCTGGTAGAATGGATTGACAAATTCCGCAGCATCGTGGCCGTCTACAGCAACGCTCTCTTC
GAACTCGCGGGCACCTTGAGCATCTGCGTGCTCTTTTGGTTTGCTTATCCCGGCGTGACAGAGAAATGCCTGGCCAGCGT
GATTCCCACGGGATCCATCCTGATTCCCACTCTCTGCCTCGCTATAGCCTGCTGTTGCCAGAAAGAACTACTGCGCTACT
CGGGATCTTCCACGATCGCCTGCGTTATCATCGAAACCGGCATCACTGTCATGACAGGGTTTTGCTGTCATCGTGTCAAT
CTCGGTCTAGGCTTTGCTCTCACCGTACTGGTGTTCATCGTGTGCAACACCGTTACGTTCTACGCCGGTCGCAACGCTAC
GCGCTGGCGAATATTTTTCTGCGGCTACAGCTGGTGCCTCCTCTGCTTCTTCTTTTTCATTACGTTCGCCAAAGCGGCGA
TTATCTACAAAATCTTCACCGTCATGTACTTTTGCTTGGTTAGCGGCGTAACCTATCTACTGGTGTATCAGCTGTTGCAC
ATCATGTACCCATCGCCTTCTGACAGCCTCAGCGAGGCCGCCGAACTCGCTCGCAGCCTGGTCGTGTACACGGCCGTCAT
TGCCATGTTCAACGCTCTCACCATGGTGTTTTCCGTCAATCAATGGATGGGCGAAATCATTAAAGCCATGAACCAAACAA
AACCGTCAGCTTGGGATTGGCTGCTACTCAACAAGAAATAGGCCATGGAGCAGGATGTGTACCCGCATAACTGTATGATT
CATGTTGTATATAAACGTGATGAGTGTGCATGTCGCCCATATAACTTGTGCTTGAAACATTGAGAACACTATGACAAACC
CTTCTGTAATAAAGCTGGTTAAACAAACCCCAATAAAAAACCACTCGTGCAAAGTGTTTGATGGCTACAATCTCTTCACT
CTCCGTCTGTATCGTACGGTCTCGCTACAGGCGTCTTGCACGTTTGTCATTGCCGCCATGATGATCCTACTCACACCTTA
CGGGGACACTTTCGAGGGCATCTGCCGCACGGGCTTTCTGCCAGCTCTCAGTCTACTGATCCCCAACGCCTGCCTCACTC
TACTTCATGCTAAAAAGCAACACTCCGACAGCTGGGCCGTCCTGACCTCTTACACGATAGTCACCTCAGCTGCCGTGGTT
CTAGCAAACGTCTGTGTAGACGAGATTCAATTCATCACATGGGCAGCCCTCATGTCGGGAATGCTATTCATCACCTGCAC
GGGTCTGTCCTGTCTGGGAGGCCTTTACTATAATCGCTGGCGACTCCTCGTCCTCATTTTCCTAATATTCGTGGGCATCG
TTTTGGTAGTTTTGTCGTTTCAGCCGCTGTCCCTACCCAATAAAATCTTACTGGGGTACTACGTGATCGTCCTGGCGTTT

ATGTTAGGGGTCATTGCTTTCGATACCTCACAACTCTTCGAGCTTTCGTCCAACCGAGAAACACACGTCCTGGGCCTCTG
CCTCTACGAGGATTTAGTCTACTGCTACTTGCTAATCCTTCTCATTCTGACCACCGAAAGCTCGCTGGCAAAGCTCACCG
ACTGGATGCAAGAGTTTTCTGTCTCCGACAGCGAAACAAACGAAACCTCTCCATTCTACACATAAACTCGCCAGCATGAT
AGCGACTCCGCACGTTCCGCTCTGGCTACGCCGTCAGATTCTCGTGTGGTCCGTGTATGCTTGGACATTTTTCGGGATCG
CGGTGACCGTGTTCGTCTATGGAGCCGTGCGGATGATAGAACCCAGCTGTTCCACCGGTCCGTGCGTGCTGTATCCCGCG
AGACAGATGGTCGTGTTGAGTCCCATCCTCATTCTGGTGCGGGAATCGTACGAACACGACGACGACCAACATAACTCCCT
AGCAGCAACACGGAATCTGGGCTTTCTGCTGTACGTCATTCATCTCGCCCTCTCATACGCCCTTTTCGGCGTGTGCATCA
ACCCGGTGAACATCGTGTACAGCTGTGCCGTCATAACTGTGCTGTTGCTACCCTCCGTCTTCATCCGATGGCACGCACGT
GGACTGCAATACTACAAAGCCGCTTTCGAACATCTCTTCCATCGTTTGCCTCACCATCAGTCTTGTTCTTCGCAACGTTCT
GGAAACAAGCCAAGCTCTTATACCCATGTTGGTTTTCCACGTGCTATCCATGATCAGTTTCGTGGGCGCCTTCACGATCG
CGCTTTCGGAAATGGAGCAATACATGACGGTGGCGCGTATCAAACGAACGGCCATTGTGCTGTATTGCCATCTCAACCTG
ATGCATTACTCCATCCTGTACATGTGGTATTCACCCTGGTAAACTCCTTCATGATGCTTCCCAGGAGTTTCGGTGTATAA
AAACTCGCTTACTAACCATCAGACTATCACTCCCAATCCGACTGCGGAGAGATATGCAAAACTGTAGCATGTGCGCGTTT
AGTATCAACTGGCTGCGGCGCTTCAGCACCATCTTCAAAATCTACGGCTGGTTCGTGCTACAAGTGTCCGTCACTGTCGC
GCTCTATGGCGTCTATCGCACGGCGTGTTCGGGATTGACCGTGGACTGCAAGGCGGACCCCGCACCCATGTTGATGGTGG
CCATTCCGGCCATCATCCTGGTTCTGGAAAACACCAGCAAACGCGCTTCGAACACGCTCAACGGAGTGGTTTGGGAAACA
TACTACATTACAGCGTCCTGTCTCCTCATCAGCATCTGCGCGGACGACTGGACGATCCTGCGCAGCTTCGCACTGACCTC
ACTCATGTTTATTTTGCAATCCGGGACGGCCGTCTTTGAAAGCATGAACCTGTATCGCCGAAAGCTGCTGGCTGTGGTGT
TCAGCTTAATTGTTCTCTTCACCATGACCATCGTGTACACGTTCGGTCAGCTGACGCCTACTCACATGCTGTTCATGACC
GCATACATGATGTTTCTGGTTTTAACATCCATTACTGTCCACCTCGAAACCGTTGCCATCCGACATGTACACCAGTCATG
CGAAGTCACATGGCCGGCCATGATGCTGTACATCAATTTCATGTTACTACTCCAGTGTAACACTTTCCTCCTCACTCCCA
ACCTCTGGTCAGCCCGGTGGGACAGCACGTTTAGTTCGCTGATTGCCTATTTTCACAGCGACTCCACCGTATCGCCCACA
GCAGCTCGGAAGTTGCCATAAGTATAAGGTGCCGGTGTAGCGTTTCGCCGTATTCGGACGCTGAGCGCTTTGGGGAGAAG
ACATGGATGCAAAGACTCGTTCTACGCTGCGATGGCTGCCGCGCACGTGCCTCGTCATCCGAATACACGGCTGGTTCGCG
GTGGCCCTGGCCTGGAGCGTGGTGATCTACGGAATTGCGCGCTTCACTTTTCCCGACTTCTATGTGTTTGCCAACTGTGA
GGCTTGCCCGCTTCCAGTGCTTTTGGTGATTGTGCCTGCCCTTTTGTTTTTCGTGGTTGATTTTAAAGCAGATAGATTTA
ACAACACGGGTGCGTGGGTGTACTGTGTGCTGCTGGTCTCTGTGTCTAACGCCGTGTTTAATACCTGTACACATGATGGC
ATACCCGTGTCCATGTTCTGCATGTCTCTGCTCATGTTTACCACCTATACCGCTACGGGTGTTGTATGTGCACGGGGGGC
CAGATGTTTGTATGTGTGTTACTTTTTGTGTATTGTGGTTGCTCTCACATTGGTGGGAGTGTATGCGAAGACCTTCATCG
TCACGGCCGTCATCTACGCCATACTTCACATCATTCTACAGGTCACCGTAGGCGCGGAAATGTATTATAACCTAAACTTG
GCGCCGTTAACCAGAACTAGTACTAGCGCATCGCTGTTGCTTTATTTTCTCTTCATCATTTTGTACGAAATGTCCATTAT
CACTTGGACATCGGGACTCAGATTTTCCATACTCAGACATCTATCCAGCAGTGCAGTGAACGCAACCGCCACGAACAACT
CGAGTGGATAAAGATCTGCACGGAGCATCCTGTGGACTCACGCCTTAAAAGCTTCCTGAAGACAACCTAAGCGAACCATT
TCAACGTCCGCGAACAACCAGAAAGCATCGGCAGTTCCAGCCACCATGCCGACCGTCAAGTACAAGCGTCTCAAACCCGA
GCATCATGAGGACGCAGCACCCCTTCCTGAAATCCGGTTGAACCATGAACACGCGGCCTGGCTGAATCGCTATCTTCCAC
TCATCAGAATGTACATCTTTTGCTTGTTTATCGTCGCTCTATCTGTGTGTACGCACATGGTCTTGATCGTTTTATCTTCT
AATGTGCCGTTGAAACCAACTATCTGTGTCAGGGAACCTAGCGTTCTTCTAGTGGTGGTCGTTGTGGCTTTCATGTTCTT
CAGCAACATGTGTGGCACGGAAGACTCCAGCCTCAAAGGTGTGGCGGCCGCCTTTGCTGTCATGTTTGCGCCTCTCAGCG
CGTCCCTGATTAACTTCTGTACAGACGCCCCTACGGCCGCATTCGCTTTGCTACTGACGCTATGCTTTTTCATCACCGTT
AGCGGCGTGGCCATTTGGATCGCGGCTCCATTGTCTTTTAGTTATCGCATTGTCGCCTGCATCCTCGTTCCTATTACCCT
GCTTGTATTCTATTTCGGACAATTACCGCGAGTGGCACGGAACATACTATCGGCTGTCACTCATGCCGCCATTGGAATCA
CCATCTGGGACACCACGTGGAAAAGCATCTACAGAACTCGTGTACACTTAATATTAATGGCCTTTGAATATTGTGAGATC
ATGCTTGCGGCGTATTACATGTACGTGTACCTGCTGACGCCCACCCTCTGGACTGTAGACCCCAGCAAAATGTTCACGGG
AATATCGCAACTCATGAACGTCACGGGCAACCAGACCTTTTGTCCGGACCCCTTTAAATGTAATTGTGTTACCATTGGGT
GTTAGACGTAATTGTATAGCCATTGGATGTTAAACGTTATATATGATCACCTGTACGCCTTATTCTCTCAGATGCAATGT
ACCTGTACGCAATAAACGCCTTACAACCTGTGAGCCATGGAGGTCCTTCGATTCCAAGTCCGGTTCACCGAAAGCATTAT
CTGGATACAACGGTTTAAGATCCTCATACAGCTATACAGCTACTGGCTCCTTCAGGTCACTGTTACTTCCACCCTCTCCA
CCCTCCTGTGGTTAGCCTATCCTCAATGGCAGAACCTCTGCCAGACAGACCCTGCCCCAATTTTGATCATTATCATACCC
GTACTCAATCTGGCCTGTCTGGTGGTGCGCTGGGAGAAACACGCCAGCGACCTCCTGACCATGACGGGGTACGTCTTCTG
GATCACATCGCTCATCACCCTGTTCGGTTTCTGCGTGTCATTTTCCACCACCATCACCGCTCACATAGCGTCCGTGACGC
TTTTCGTGATAGTCACCTGCGGCGCGTTCTTTGAGCGCGGCGGGAGCGGCTGCACCCGTCGCTATACCCTGGCTGGTATC
TGGACGATAGCCAGCGCCGCGGCGCTGCTGATCATCGCCGTAGCCTCCCAATCGTCTAACCTCACCACCCACATCATGCT
AAGCGTGTATGTGGTCTTTTTAACCTGTTTCATCGTCCTAAACTTCTGGGTCACGCAGCAACACCAAAAACACCTGCACC
CAGAACAGCTGATCCGGGGGGCGTTGACCCTTTACGTAATGTATTTTGTGTTATTCCAAGTCACGCTTCTCATGCTCAAT
TCGAAAATCTGGGATGTAAGCTTTAGTAAGATTTTCAGTTCGTTCACCACCGCAAACAAGACTGTTTCGTAACACCAAAG
GGGCACATGGGGTTCCAGCACGCTATAAAACCCCACGCGCGAGTGTGTTTCACCAGACCGTCGTGAAAGATGGGCCGCCA
GGACACGTCCCGCGAGGACAACGAGGACTACGAAGAAATTATGAGATGGGTGCGTCGTTTTGTGTGGCTAACGCGGGTGT
ACACGGTCCTGGCCGTACAGATGGCCGTGACCCTGGCCTTCTGCTTAGTGTGCATCATGTGTGCTTGGCGAATACATGCA
CCATACGTGAGAGAAACCCTGCCGATCTGGATTATGATCGTCCCCAGCATACTGCGTTTCAAGCTTCGTAAGAAAACCTA
CCACGCCACCAGCGTAAAGACGGCCGTACTCTACACTGTGATCAACAGCTGCGCCCTGGCGATATGGAGCATGTGTCTGG
AACGCAACGTACTGTGGCAGGCATACATCCTGTCCCTGGCTTTGGAGTTGGGCTGCACCATCATGGCCTGCGTTCTAGCC
TCCACTCGCCCGCGCGGATCCATCATCGCCGCGTTCCTAATTCTCGCCCTGCCTCTATTCTGCACCATCGTGTACTACCA
GCCCTGGACTCCGGCTCAAAAGTGGCTGGCAGTGTTGACCGCCACCGTGGTGGACCTGATAACCCTGGCTCTGCTCCACG
ACACGCTCCTGGTTCTGTGCCATGCTCCCCGATCGTTGTTCGATCGACATGCCGTGCGAGCCGCGCTATTGCTGTATGTG
GACCAAGTGCTAGTTCTGATGATGAGTGTGGTGCCTCTGACAGCCGACGGCTGGTATCCAGATTACTTCCTGGGCTCTAA
CCAACCCGTTGTGTGAATACGTGTGTAAATAAAAAAAAAGTTGTACCCGTGCGTTTGGCGTGTGGTGTATTGCGCAATCG
CAGACTCTGAAGGTCAAGTGAAGGTCACCCACACACATCATTTAGCTAAACCAACGTTTTTAAAAACGACCGTTAAAACC
GTTCATGTGGGCGCGCTTTTCGAAAATCAACAGTTTTGGGATTCAGGCCAGGCAGACAGCTTGGCGCCACTGGTGTCATC
TGCGCGTCACGATGTAGCTACATAAAGCCGGCTCGTTTGGCACTTCAGCACAGCTACTGAAGCAGTTTCAGCCGACCGCA
ACCATGAAACTGCTCTTTGCTTGCCTGTTACTGTGGTGCGCACACCACTTGGTGAGGGAAAGCCATGCTCACCTACCAGA

ATGCGCCCCTTTCTTATGGGGGGAACCCAGCGAAGAGGACAAACGGGAGCTGGGGCTTCTCGACATGCCTCCCATGCCTC
CTGCACCTCCGGAATCTGAAGATTTGGACAATAGGAAGATCAACATCATCAATGTCAACCATACCTACGAGATCATCGAA
AGTCCAGCGCCCGTCTATCGCGTGCCAGCACCACCGAAACCTCCATCGCCCGTGTATCGTCTGGACACTTTGGACGCCCA
ATGCCGTTTACTGAACGCTCACACACTAAAAGCTTGGTGGGCATTCACTGGTAACGCTACCAGTCCAGATCTGCGTGTGA
GAGCCGTGTGGTACGGCACAACGGGATATGGAACCATGCCCACTCCACAAAATGTGTCCACGGACTCTGAACAGCGGTAC
CACCTGCACACCGACCAGTTCCACAAATACTTTACCGACATCTGGTTTGATGTGTATCCTTGCAAGACGTGTGCGGCTCA
GAGCGTGAGATGCCATCCACAGATCGAAATCTCCTGGTTCCCGATGTGGGCTTCCATATCGGACATTGACAGACTGTTGT
TCGAAGGCCGGAACTGGACGGCCTACTGGGCTCTGCTGTTAAAGTTTGTGCAGTGGACATTCGCCGCGTCCATGTTCATC
CAGATATACTGGAGCACGGCTGGAAGAGACTACTGGCCGCACGGCTTCTTTGACAACATGCACTGAGTTGACCCGTGAGA
AATGCGGAACTGTGATTTTGCAAGTGCAGGTATGAAGCAAGTGAATGTGCGTGTTTCATGACGCGTGGACAAACCCGGAT
CGCGTCAACGAATTGCAACACACAACGAGCCAGGCGCATATAAGCCGTGCCAGGTTTGTCCTGGGCACAGCTCTACCAAG
CTCTCTTGGACCAGAGCCTCGCTAACTGTTGATTGACAAGCACTTGTGTCAGTCAACGCTTGACGGACAGGCGCCTTGAC
AGACAGCCTTCACAGCCTCGTCAACATGAAGCAGCTCGTTGTTGTCGTTACCCTGGTGGCCCTGCTGGGGGTGTCTGCTA
TCTCCGCGGAATTCTGCTCATCCGCTAACCGGAGCGTGTCAATCTACCGTGACTTCAACGAGACCCAGTGGCTGAACGAG
ATGGCGCAAATCTACGCCGAAGCCGAGACCAAGGCTATCGGCAACGCCGAGAGTCCCAGGTGCTTCGTGTGGCTGAACTG
GCGCTGGAGCATGTACCTGGCGAGACGGGCTGACAGCTACCCCGTCGTGAAGCCTTTCTACAGACTGCCTTGCCCATTCA
GCACCAACGCCAGCAGCGCGAACGAGATCGTGAAGGTCGAACTGCCTGCCTTCTGGATCTGAGTCCAGACCCCCCCCCTC
GCCCCAGACAACGGACTGATTGTATTAATAAAACGTTTATTTTTCTACACTTTCTGAGTCTCGTGGTTTGTTTAAGAAAC
GGGTGGGACATGAAACCTGGTGTTCGAGGGTGGTGGGCGTATGCACGACAGACTGCGAATATAAGCAGTGGGTCACATTA
GGCGTTCTCAGACACTCTTAACACCACTAACACATAACTCGGTGTTAACATGGGGATAGGTTCCATCGTACTGTGTCTAC
TGTGCGCGTTGGTGCAGGTTCAGGGGGACCACAGTTGGGCACCCGTTACGGTACACAGCACCAATGCTACATGTACAGTT
ACCGCAGGCATTCTGGACGCTACATGGTGCGTGATGGGAGAGCTGTCACCGTGTACCATGAGTTATAAGTTTTTATCTTA
CAGCCGAAGAAACGGAGAACTACTCGTCCTGCCGGGAAAGCCATCGTATCTGAGCGAGATGGTCATGTGTTTTACAATCT
ACCACCCCGTACCCACGGATCTCGGGGTTATAGACTTACACATCAAACCCTGTAAGGAACGTTGTGTACTACCGTGTACT
CCGACGACTGCCATCTCCTGGAGCTCGTTATTCAATGGCGACGTGCAGAGACTGCTTTACGAAGATAGACAGTGGTGTAT
CGTGAAAGTAGTGTGCATGAATGTGGTGGCCGTGTGCTTTGTGGCATTTGTACTCTGGCGGTTTGGCACACGCCGCGGGA
GTATGAATTCGCACGGCGCGTACTGTGTGCATTCCAATAAAAAGCAACAGTGATTAATGACAAAGCTGTGGACTTTCTTA
TGATGGCTACTTGTAGTTTCATTATCGCGTACGTTTTGTTTTATGCAACTCGCTTTGCGCAAGGGGGTATGGAAAAATAG
TGGGTGACGACACACATGCCACGTCAATGCAAAGTCGCCCACGCCTTTCCTGTGTTTACCCCACTTCTGCTTTAGTAAGG
CATTTAAAAACCACGTTCGGCTTACTAAAGCACTTCGCGGTGCCCGTCAAGATGTTGCGGTACCTGTGGCTGCAGTTGCT
AGCGACCTTCCTGCTCTTTGAAGTTTCCTTCTGCTGCTTCTTTTCCAAAAAGGGACTTACCACATCATATAACCGAAGAT
TCTTTAACAGATGGATACAACTCAATAGGCACTGCAAAAAAGACCCTGTGATACCACGCATGTATCAGGAATTCAGCAAT
ACATGCGAATTGAAGAACGATCATTTGTACGCCAGTGGCATTATCGCGGGCAACTTCACGGACATTGCCTGGTTGCATGT
TCAGATCATTCGAAAGCGTTATGACAACTGGTTCTTTGTGCTATCCAGCGATGCGCAACATCCGCCCATCGTTATGGACT
CTAAGAAATACAGATACGGAATTCAACCTGAAGCGCTGATAGCTGAAGCTCAATGGCAAAATGAATTTGGGATTACAAAA
GAGGTACGCGTGTATAATCAGCGCTTGTTTCACTATAAGTTTCAAGTGCCCGCTTTCGAACATGATGTGCATATTACCAT
GCACATAGAACCTGAAACCAAACCTCTGTTCATCACATGTAATCCTACCTATGGGCTGGACTGGAGTCCTCAGCTTTTGA
AAAACTACATAAGAGAAATGTATCTGAAATTCGGATACATCTTCTTAAATATTACCATGATGGCAACGGCATGGGTTATC
GTTATGCTGTTCTTTTGGTGGGTTTGGCGTCTGTTGTATAAAGATGCCATGGAGCACGCAGCGCGTTGCCGAATGCTGGT
GAAACTGCACCGGGAAAGATATCGTCGCATCGATCATAAAGATTTGTAATCAACTTACGACATCGATGTGGCGAAAACCT
CTACAGATACTGAGAACGGAACACCCGTCACCGCAGCGGTTCACTCGCCGACGTGATGCAAGACAGACGTCTCGTGAGGC
TTATCACGAACGCACAACTAACCGCGATTAACGTCACGGGAACGATTGTCAACACCCATTCCGTTCGGTCATGGTGGAAA
CGTCACTTCCAGCTAAGGTATATAATCAACGTTCGTCTCGGCTCCAGCACGAGTGACAAATTAGTTGTCAACGCAGCAAC
ATGAAGCCCCTGTGCACGCCGTGGATCGACGGGACCATCCTGTGCTCGCTGCTGTTTCTGCTGGCCTTCTCTGGTGTCAG
CTCCGCATGGAGCAACGACACAAAAACTGCTCCAGAGAAGGTCTCAACATACCTGAGGGACCATGCCAAAACGCGTCCGC
TCTTTGGCAGCGACAGCAGCAGTGCGACCTGCGATCGCCAGTGCCCGTTGCGAGACGACAGCCCCGCCTGCAACGCCTTT
CTGGAAAGCTACGAACGAGCCGTCCAGCTCCTGGGGTGCGATCCCCAGCTAAAGCACCGGCACGTGCTGGGGCCTCGCGG
ACTGTTCGGGTGCCCGAAAGCGCCCAGAGGTGTGCTCGTACTGCTGGAGCATTACGGAGGACTCACGCTGATGGTGTACA
TGGCGCTCGTTCTGTTCGCGGTGGTGTTCACATCTCTCATGCTGTATGTGATGGACGAGTGCGACGTTGCTCGGAACGCG
AGACGCTGCGGAAGCCATTGACGCTAAAGCAACATTGAACTTGTCAATCAATCTCTTTAATTAAAAGATGTACATATTTA
TTGTTACACTGTTTACGTTGTGGTTTCTTACGCCTCGAGTAGATCTGTGTCAGTGGGCAAACTGTGTCACGATAAGGGTG
GGATATGCGCTCACTATCATGACATAGTCTGACCATTGTAACAGCATGAGCGGTCGGCAGGCAGTGCGCGCGTCTAGATA
GCCGCGGAGGTCTATCGACAACATGTTGTCACCTGGAGGTGCACAGGGCTGTTGATGGAAATACACGAACATACAACCCC
TTGCGCGTCCAGGTTTTCACATGCCCGATTCGGCTGGATGCGTGACCGTACATCTCGCGTTCCATTCGCGCCGCCCAGCA
TGCACCAACCCCGCACCCCCAACCTCCATTCCTTCTCCCTCTCGGCGTGTGCATGATTGACAATCACCAAACCGACCGGA
CAGCTCCTGCACATACATCCGTTTTTTTTACATGTCTATGTGCGCAAGCCGTGCGGTAAGGTCGGATGATCCATCATCATC
ACCATTATCACCATCATTCTCCGTTCGTACGAGGCCGGTATCTATCTGACCTCCTCAGCAGACCCCCTGCACTTAGATTC
GTATACAGATCTGTGCCAGAGGGCGGCCGAGAAAGTCAAAAGACACATTCACGGAATCCGTTTTGCGATCCGACTCGCAA
CCGCCCCAACCCTGCCCAGATGCCGCGGAATACGTCCAACTTCCCCAGAGGCCAGGCGCCATAGATAGTCATAGTAAAGT
CTATAGCGCCGAGCCTTTGGGAAAGTCGAACGACACATCCGAACCAGCGGAGATCCGAACGACAAATGAGCGGGTGGACG
CGGGCTCAGATATACACAAGACGGATTCTCGGAAATCTCGGATGGGTGTGGCTCGCATTCGGAGGAAAAACCGTAACCGG
TGCCCGGGTACAAATTCGCTGACGCGTGGCGGTTCAGAAACCGCAACGAACATCTTTCCATACACGTCAATCATTCTGAC
GCTGGTTTTTCCTGGGCGGTTCTGTAAGAATGGTGTATATATCAAGAAACACCGTGCAGTTCACTCAGCGCGTCTGGAGA
CGAGCTACCGCTACATAGACTTCATTGACAGCTCGGGCTAAAAGCGCCTCCACCGCTATCGACATGAAGTTTACTTGCGC
TGTCTGCGGCCTCCTGGGCCTCATCCTAATGGGCCTAATACGAACAACAGATTCTGCACCGTATCAAAGGATGTCTTCAG
ACTATAGAATGCTGCATTCATCATGCTTTCTGAAGGCCGGTAAAGTGTACGGCCGCTTCGTCGTGGAAAGCGATTTTGTG
GATACAAAATATGTGTTCGCTATGGCTCATCGCCAGGGTTCATATGGAACGCTCCACGCGACCAACACACATTCAACAAA
GACCACCAAAACCTACATGGTGGAGGACTTTGTGCACAGCACTACGTATGGCATCGATTTGGTGATAGTGGGCCTGTACG
ACGTGTACCATACCTTCCAGTGCAAAGTGGAAACGCACATTGACTTATCGCCAGAAAGCGTGCTGTGGTACGCGGAGAAA
ATCTCACTGGAAAACAATTGGTGGCTGTTTGCCAAGATCGCAATGTACCTCGTGGTGACCAGCGCGCTGTTAACGTACCT

AACACTGATGGTGCATTTTCGTCTTCGCGTCACGGGAGAAGTGGCTGTGAGCAACCATGACTCCAAGGTGTGATGACAGG
TAAATGTTTGTCTCGGTATGTTTGTTTTCATGTTCGTTATTGTACGTCTGCATGTAATCCACTTGAATTAACTACCGTAA
CGATAACACACTTGACCTTTTGTGTTTGTAGGTCACTGACAATAAACAACGGAATGATTGGCAGGTAGCCATGATGGACG
AGGAGCTGTCCGTCAAGGAGACGTGGTGCGATGGCGTGCAGACCGTTTGCATCTCAAAGAAGCACAGAGTGCAGCACTGT
TAACGAAACCCTCAGGGAACAACGGAACAGTTTACGAACTACCTCAGTGGCACCAACAGAGTTCACACCTGTTCGGTGGC
CAACCATGAACTCATTACCTCGGTGGTTACCGAGATTGACATGCACTTCGTATTTTATTAATAAAACTAATTGTGATTGT
AACACATTACAGTTGGTTGCGTATTTTGTTTCCGTTTAGATTGCACATCGGGCTGGGGCCGTCATGGGCTGGGCTTGGCT
CTGGCAAAATGCGGACGCGCATCCCGTGCGTCAAGGGTTTCTTGTATATATCCCCAAACACCGTTCAGAATCTCAGCTAT
TGCTCGGACTTACTGCGTTCGCGATATGTCCTGGTTGATTTGCGCCTGTTTGCTGAGCACTTTGATACGGACACAAGGCT
ATGGAGCTAGGCATACTGCATGGTTTCACGATTCCAATGATGATTTACAGCATCCGCGTCCGTATGAACCAGCTGCAAGG
CAGCTTCACAGAAGCCAGCAGAGACCGAAGCAGCCACCTCCGACAACACCAACCACACCAACAGCAACCGTGAAAGACGA
CGGATTCAGAGTGCTCGAAAATGTCTGCTACATACAAGGCACGAGACTGTTTGTGCGTGGGGCCGTTGAAGGGCGCATAC
AAGATTACCATCTGGCCGTTGAGTTACACGGTGCCAGCGGGGAATGGGAGAGATTTATATTTGCTCAAAAGGACGTTCAA
TTTAAACGAGGACAGGGAGTCTATTGGAACTTTCCAGGTTACGAAACACCACCAGAATTAAAAAGGGTGGTCATCCATAT
GTACTTCAACACGTTCGCAGAACTCTGGGTTACGTGCGAACCCGAACTGAGGGTGGACTTTGGAAGTGAAAACATAGCAT
GGCTGGTTATGAGAATGGTGACTTTTTCGGAGTACTGGTACTTAGCGTGGTGCACCTTTGTCTATTTCATTCATGTCTGC
ATTATGGCATTATTTGTGGATGAGATCAATGTCCCGAGTTTGAAAATCGTGTGCCACGTATTCTGTGCGTAAAATAAAAC
CCTGGCAACTCACATCTGTGTGCACGATTTATTTATAGCGTTTTTGGCAGTACAAGCGAACACAGAGATAGAGCACGGGA
CACTCCACGACCGACACGCAAAACATACCCGCACTTACGCGAATTCATACAAACTAGTAAGGGCCATGCTGAGGGGGGTC
GTTTACCAGTGCGGTGGGCTGCATCGGGAGGAATTCAGAGTCCGGCTTACAGTTCGTAACCGCGATCATGGAGCAAGGCA
GCATCGTTGGCGCCGTGGGACACACGCCCATGCGCGAGCTGGCATGGCGCCGGGTCGCTGACGACTCGCACGATTTGTGG
TGCGCCTGCATGGATTGGAAAGCGCATGTGGAATACGTGCAGCCTCCCACCGATGAGCTGTTTCCCGCGAGTGGCGGCTG
GCCGGCGCAGGTTGAAGCCCAATGGCGCCACCAGGTCAAGGTGGCGCACGACGCGTGGTGTCAGTGCCAAGACTGGCGCG
GCCACGCTCTGCGAAGTCGCTCCCTCACGCTGGACTCCGACTCGTCTCGCTCGTCGATGTCGTCCTACACGTCGGCGGTG
TCGTGTCCTCCGCTGTCGGAGAACTCCAGGGTGTCGTGGTGGAAGCGGATGAAAGGGCGATTTCACCTGCGCGGCTGGTG
GGGGCGATGGGTTCAGCGACGCAGGCAGCAGGACACGCAGAATCAGTCGTCGGCGGCATCGTGACTGCTCCGGGTGAACG
CGGCTTACAACACAGGGTGGGAACTGGAGGGGACGTCATACTCTGGGGATGCATCCTCAGGATTAAAACATGGCATTGTC
ATCCACACCATGTTTGGTTTTCTAAGAGACAATTTCAACTCGGGACGTTTCTTGGGAGGAACGGGCGGGGGACAGGTGGG
CATGGGAGAGGGCGGTGGGCTGATGGTCATGGTGACAGTGACGGGATTGGCAGGGGCTGGGGGAGGCGGGGGAGGGTGGC
TAGGACGTGTAAACGCCGTCTTCTTGCGTTTGCGTCTCGGTGGCAGAATGCAGGGTTCCCTTGGGTTGCACTGGCCGTCT
GGGGGTTGTCTCGTGCTCAGCGTGCGGCCCCGACCCTTGTGGAGTCGAGGCGGTTTCGGGGGCGCGTGAGGAGACGTCTC
GATGGGTTCCTGGCGCACGTTCTGGGGTGGTGGCGCAGCAGTGTGGCACGGTCCATCGACGGGCGTGGACGAGTCAGGAG
ATGGAGGCTTACCCTGGCAGGCACAGACCAGACAACAGAACACACATGTCAATCAATACAGTGCAGCACAGCCAAGGTAA
ATACACGAAGCATAACAGAGCTTACCTCGGGGACAACTTGGTCTTGGTGGCGCATGTTCGGCAAGTCGGCGTTAGCGGAG
TGTTGCGAAGTGTCCTTGGTGCCCGTGGTTTGGATAGTGATGGTTCGCTGGCCACAAAGATGGGCTCAAAATGTGGCAAG
GTCAACGAAGCAAGTGACGAGTTTTATACCATTCGCATGCTCAGGGAGCGGTGGATGTGGAAGCTGACGTAAATGCGTAA
GACTGGTGCATGCGTGTCCAAAGCTCGTGCACACGGCGTCCGCGCAGACGGTTTACTGACTGCGAGTGTTCAGCAAGTTT
TTCGGTAGCATATTGGCGAGTGCGTCGGCGAGTGCTTGCGCGATTTCAGTCTCGCTTTCGGTTAGCGTTAGAGCATCCTC
GGTGTCGCGCGGCTCTTGAGTGTCGTTGTGGCCTTTGGTAGCAGACTCGACATCGCTGTCAATCACCACCTCTGTGAGGT
CTTGGCTTTCTCCGGTTTCGTAGTCAGAAGCGGCACACGTTGGCTCGTCAACATCCACAGTGATCACAGTAGGTTTACGG
CTAGGCGTCGTGGACCGAGGCTGTTTGGTGATAATGCGTTTGCTAAGAGCTTTGGGTTTGCTAATAGTAGCGAGAATGTC
GTCTCCGCTGAGGGCCGAGTAGATGGCTGCGGCGAAGATGGCAATCATGGTCATTACGCACGCTGTGCCGCAGCACAAGG
CTGCCTTTAGCGCGGAGCGCAAGCTGCAACGCATGGCTGCGGGCTGAAAATGTGTAGCACGCAGTTAATGTTGATAATGC
GTTAGATGCTGCGGGCAGCGCGCGCTGTGCTGTGCGGGTTGCGGTTGGTGGCGCAGGCGTTTATGTACGTGTTCAGTGTG
TGTGTGGCCAATGGTGTACACAGAAAAAGCATTACAAGGTGCGAACTGTACGGCAGGAGGAAAAGCATTAGTCACAGCAG
CTTCCCTCCAGCAGGGGAGCCGTTAAGTGGAGCGGTGGTGGTTTTGGCTGTGTGCCAACATCAGCAGACGCACGAGCTGG
AGAGTTTGGCAGTAAGCCATGAGGACACCGCATTAGGGCCTGATGGTTTGGCAGTGTGCCAGCCGAGGAGAAGTGGGAAG
TTGGCATATCGCCAGCGGTGGCGTGGCTGGAAATTGGCATTCAGCCAGCAAATTGGGAATTCGGCAGCAGCTTTCCTCCC
CCGGCACCAGAGTGAGCCCTCTGCTGGGTCTCCGGCCTCCTCTCCCGGTTTGATGCCCCGGGGGGGTGTTTCGGCGGGGG
GGGGTAAAAATTGGCACGTTGCCAAGTTTGCGGAAGCTTGGGTGGGCGTGGCCTTGGCACCGTGCCAGGAGTTGGCGCCG
TGCCAGGTTAGCGTGTGGGTAGCATGGGCTGGGCGTGGTGCTGGCGCGGCGCCAGGTTTCAGAGGCGCGCGGACCCCAAG
AGGAAGAGCTGCGGCCTGCTGGCGGCCAGCTCGGCGGCCGTGGCGCTGCTGACGCTGGCGGCGCTGGCCTCTTTTTCTTC
TTCCGCCGGCGGCTCGGCGCGGTCGCTGCTGTTCGCCGAGACCCACGACAACAACAGGCGCACGAAGAACAAGACGAAGA
AGAGCGGGAGAAGGAGAAACACGCGCTAGCTACATACAGATGATGATAGCGAGCGGCCGCCGGCCGTCACAACACCTACA
TACATGCTAAGGACACAGACATACTACATGTGTTGCTGGCACAGCGCCAGGGCTGGAACATGTGTTGCTGGCACAGTGCC
GGGACTCGGATGAATGACGGCGGCCGCGGTTTGATTGACATGTCCGGGTTTGAGACACACAGACACACACAGAGAAACAT
GGCGGCCGCCCTCCTCCCTCCTGTGGACGTTTGCTTGATGGACAGATCCGGCCGAACGCAGCGCGGAACGCATCCCAAAC
CGGCGCCCGCACCCCCGCCGGGGCCCGGGAGGGGGGGCGTCTCGAGGCCCGGGGCGTCTTTTTGGCCGTTCCGCTGGCCC
TCTGGCCTCCTTTCGCCCGGCGCGCCGCCAGAATCGCCTTCCCTTTTTCTCCCCCCGGCGCGAAACCTGGCACGCTGCCA
ACCGAAGCGTGACTCAGCGGCTCGGGGCGAACACGGAGCCGGCCAAGTGATGACTCAGCGGGAGGCCGGCTGACTCACAC
TCGCACGCTCGCTCGCCGCAGCTCTCAAAAGCACGCGCGCGCAGCGCGACACACACCACAGCACGCGCGCAACGCAAGGA
CTCAGCACAGGGAGTGCGAAAAGCCCCGGGAGAGACGAGGGTGAGAGGCGGAGAATCCACCTTCCTCGGCCTCCCATAGT
GCTGGGATGAGCCATCTTGCCTGACTGTATTATTTCTTTTCTTCTGTTTACTTCAGAATTAATTTGCTATTTTTTCTAGT
TTTCTAAAAAGGAAGCTTAGATTATTGATTTTAGATCTTTCTTCTTTTCTAATATATGTATTCAAGGCTATAAATCTTCC
TCTATGCACTGCTTTTGCTGCATCCCACAAAATTTGATAACTCACACTTTCATCTTCATTTAGTTCAAAATATTTTTATT
TCTCTTGAGATTATTTTATTGAATTTTAATTTTTTAAAGTGATAAAATTGTGTGTATTATTGTGTGTAACATGATGTTTT
GAAATATATATTTTGAAATATATATTATATATATATTTCAATATATTATATTTCTAGGTATATTCCACTATATACATA
CAGTAGAATGGCTAAATTGAGCTAATTAACATACATATGACCTCACAAAATTATCATTTTTTTGGTGAGAACACTTACAA
TCTACTTCACATTTTCAAGAATTTGATACATTGTTATTAACTACAGTCACTAGGTTGTACAACAGGTCTCTTGAATTTAT
TTCTCCTGCCTAACTGAAATTTTGTATCTTTTAACCAACTTCCTCCCCCTACCCCCATATCTCTTGAGATTTCTTCTTTG

AGCCGTGTGTTATGTAAAAGTGCAGTATTTAATCTCCAAGTATTTCGAACACTTCCAGCTTTCTTTCCATGACTGGTTTT
TAGTTAAATTCCATTTTAGTCTGAGAGCAGACGTTTTATAACTTTTGTTTTTAAAAAATTGTTAAGGTGTGTTTTATGGC
CCAGGATGGGATCTATCTTGGTGAATGTTCCATGTGAGCCTGAGAAGAATGTGTATTGTGTTGTTGCTGGAATGAAGTAG
TTGATCGATGTCAATTACAGCCAGATGACGGATGGTGTTATTGAATTCCACTATTCCCTTATGATTTTCTGCCTGCTGAG
TTTGTCCATTTCCGATAGGAGGTGTTGAAACCTCCAACTATAATAGTAGATTTATCTATTTCTCTCTGCAGTTCTGTCAG
TCTTTGTCTCGTATATTTTGATGCTGTTGTTAGGTGCATACACATTAAGAATTATTATGTTTTCTTGGAGAACTGACCCC
TTTATCATCATGCAATGCTTCTCTTTATCCCTGATAACTTTCCTTGCTCTGAAATCTGCTCTGCCTGACATTAATGTAGC
TACCCCTGCTTTCTTTCGATTAGTGTTAGGATGGTATATTTTTCTTCATCCCTTTGTTTAATTGATTTCCTGTAGTTTCA
ATACAATATACCTAGTATAGTTTTTTTTTTTTTTTTGTATTTATCCAGGTTGGTGTTCTCTGAGTTTCCTAGATCTGTGG
TTTTGTGTCTGACATTAATTTGGGAAATATTCTCAGTAATTACTGCTTCAAATGTTTCTTCTCCTCCTTCTTCTCCTTCC
TTTTTTTTTGAGACAGGGTTTCCTCAGCTGCCGAGGCTGGAGTTTGGTGGGGTGGCTGAGCTTGCTGCAACCTCTGCCTT
CTGGGATCAAGTGATCCTCCTACCTCAGCCTCCTGAGTAGCTGGGACTACAAGGGCGTGCCACCACACCTGGCTAATTTT
TGTATTTTTTGTGGAGATGGGATTTCACCATGTTGGCCAGGCTGGCGTCGAACTCCTGAGCTCAAAAGATACGCCCGCCT
CGGCCTCCTGAAGAACTGGGATTATAGGTGTGAGCCACCATGCCCAGCCTCAAGAATGTTTCTTCTGTTTTTTTTTTTG
CTCTCTTTCCTCTCCTTCTGGTATTCCCATTATGTGTCTGTTACATCTTTTGTAGTTGTACAACAGTTCTAGGATAATCT
TCTGTTTTTCAGTCTTCTCTCTCTTTGATTTTTCATTTTGGAAGTTTTTGTTGTCAGATCCTCAAGCTCAGACAGTCTCT
CCCCAGTTGAGTCCATTGAAGGAATTCTTCATTTCTCTTACAGTATTTTTAGTCTCTAGCATTTCTTTTTGATTTTTTTC
CTTAGAATATCTATCTCTGTGCTTACACTGCCTATTTGTTCTTCCATGTTGTCTGTTTTATCCATTAGATTTCTCAACAT
GTTACTCATAATTATTTAACTGGCCGGGTGCCATGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGCAG
ATCACCTGAGCTCAGGAGTTCAAGACCAGCCTGGGCAACCAGGTGAAACTACAGGAAATAAAAAATAAATAAATCTTAAA
AGAATCCAGAGGGAAAAAACACTTTATGAGGAGCAAAGATAAGAATCATATTTACTTTCTCCTCAGAAACCATGTAAGGA
AGAGGAGAATGGAGTGAAATACTTAAAGTGTTGAGAAAAAAAAAAAAACCCTAAACACCAGCCCAGAATGTTCTGCCCTGT
GAAATTATTCTTCAAAAGCGAAGAAAAAAATAGACTTTTTTTGTTTTTTAGACGGAGTCTTGCTCTGGTGCTCAGCCTGG
AGTGCAATGGCATCATATCGGCTCACTGCAACCTCTGCCTCCAGGATTCAAGTGATTCTCTTGTCTCAGCCTCCTGAGTA
GCTGGGATTACAGGCATGTGCCACCACGCCTGGCTAATTTTTTGAATTTTTAGCAGAGATGGGGTTTCTCCATGTTGGCC
AGGCTTGTCATGAACTCCTGGCCTCAACTGATCCACCG

APPENDIX B

Primer	Sequence
SP86	GGTGACATTTGGAGTTGG
SP119	TTTAACCGACAGGATTTGG
SP216	TACAGTAACGCAGGGAGACACAGT
SP217	GAGAACCGCACGGAGTAATCAAAT
SP266	ACGGAATGTAAAAGGCACTA
SP286	GCTGGTCCTAGTGGCCTCCGTCAT
SP367	GTTTCGAGGTGGACAGTAATGGAT
SP378	CACTGCTGCTGTCGCTGCCAAAGA
SP390	AACTCTTGTAACCATCATC
SP397	CATTGGTGTTCATTGTTTC
SP417	GTAGACGGCAACACCATAGC
SP467	CCTCGGCTTGCTTATTTTG
SP469	ACTGCTGGAGGAGGAAGAGGAAC
SP471	GCCACATGTTTTGGTTACTTCTTTTACTG
SP474	AGAAACTGCACCCCTTATTG
SP475	CGTAGTGAATGTCGTTGTAAGC
SP485	GCACAAGTGGTCACAGTATTAG
SP486	GGCTGAAAAATAAGCATAAACC
SP490	CGTTGTTTTTGATGGTTGTG
SP491	ATTATGTAGTAACTGGGCGGTG
SP492	GTCGCCATACTAAGCACATTC
SP497	TGCGGTAGTGACAAAAGAC

VITA

Susan Rae Neubauer

Candidate for the Degree of

Doctor of Philosophy

Dissertation: CHARACTERIZATION OF THE GENOME OF BABOON CYTOMEGALOVIRUS STRAIN (OCOM4-37) ISOLATED FROM THE OLIVE BABOON, *PAPIO CYNOCEPHALUS ANUBIS*

Major Field: Biomedical Science

Biographical:

Education:
Completed the requirements for the Doctor of Philosophy in Biomedical Science at Oklahoma State University, Tulsa, Oklahoma in May, 2011.

Experience:
2003 – Present: Associate Professor of Biology, Tulsa Community College, Tulsa, Oklahoma
1996 – 2003: Associate Professor of Biology, Rogers State University, Claremore, Oklahoma

Professional Memberships:
2000 – Present: American Society for Microbiology
1996 – Present: National Association of Biology Teachers

Name: Susan Rae Neubauer Date of Degree: May, 2011

Institution: Oklahoma State University Location: Tulsa, Oklahoma

Title of Study: CHARACTERIZATION OF THE GENOME OF BABOON CYTOMEGALOVIRUS STRAIN (OCOM4-37) ISOLATED FROM THE OLIVE BABOON, *PAPIO CYNOCEPHALUS ANUBIS*

Pages in Study: 225 Candidate for the Degree of Doctor of Philosophy

Major Field: Biomedical Science

Scope and Method of Study: Molecular Biology of Herpesviruses

Findings and Conclusions: This project involved cloning, sequencing, and analyzing the genome of baboon cytomegalovirus (BaCMV) strain OCOM4-37. After isolation, cloning and sequencing the coding sequence of the BaCMV genome, comparisons were made with other CMV genomes. These analyses showed that the OCOM4-37 strain is most closely related to CMVs isolated from primates most closely related to baboons.

ADVISOR'S APPROVAL: Earl L. Blewett Ph.D.