SEQUENCING AND CHARACTERIZATION OF THE POTENTIALLY PATHOGENIC GENES OF GREEN TURTLE HERPESVIRUS

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

MICROBIOLOGY
(MARINE BIOLOGY)

DECEMBER 2003

By:
Olivia Nigro

Thesis Committee:

Yuanan Lu, Chairman
Philip C. Loh
Paul Q. Patek

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to Dr. Yuanan Lu. Without his support, advice and guidance this work would not be possible. I would also like to thank Dr. Vivek Nerurkar, and Dr. Richard Yanigahara for their direction and support in the Retrovirology Research Laboratory. I would also like to thank the Dai Ho Chun Fellowship foundation for providing me with financial support during my graduate studies. Additionally, I would like to thank the staff and researchers at RRL, especially: Deirdre Killebrew and David Troelstrup for answering my questions and helping me during periods of uncertainty during my studies; Linda Lin, Katja Zeigler and Joseph Marino for making the work day enjoyable; and Leeroy Lee and Jason Tongston for showing me what ohana meant when I first moved to the island. I would like to acknowledge Tracy Freitas for his assistance in creating phylogenetic trees. Finally, I would like to thank my family for their unconditional support and encouragement.

ABSTRACT

Green turtle fibropapilloma (GTFP) is a debilitating and frequently fatal disease of green sea turtles. The etiological agent that causes this disease is unknown, however, recent evidence suggests that a novel green turtle herpesvirus (GTHV) is closely associated with GTFP. Because this virus has not been isolated or propagated *in vitro* cell culture to date, it is necessary to obtain more genomic information of GTHV in order to establish its pathogenic role in relation to GTFP. Using a modified genomic walking technique, based on inverse polymerase chain reaction (IPCR), a total of 9,032 base pair DNA sequences of the newfound GTHV was sequenced from tumor tissues of green turtles with fibropapillomas. This newly identified DNA fragment contains three potential pathogenic genes from the unique long (UL) region of GTHV: UL29 (DNA binding protein gene), UL28 (assembly protein gene) and UL27 (glycoprotein B gene). The UL29 gene was sequenced through 4 consecutive genomic walks in which a total viral fragment of 4,054 bp of sequence data was obtained. This viral DNA fragment contains the entire open reading frame of the UL29 gene (ORF), which is oriented in the 3' to 5' direction, and encodes the full-length genomic sequence of the DNA binding protein (DBP) gene (3,585 bp) corresponding to a protein of 1,195 amino acid residues. The viral fragment also encodes a partial sequence of the UL28 gene, which is oriented in the 3' to 5' direction as well. Through 2 subsequent genomic walks, the remainder of the UL28 was sequenced in its entirety, as well as a partial ORF of the UL27 gene. The UL28 gene has an ORF of 2,250 bp with a putative amino acid translation of 750 amino acids. Through 1 subsequent genomic walk the entire ORF of the UL27 gene was

sequenced. This ORF is 2,551 bp long and overlaps the UL28 genome by 1 bp. The UL27 gene is oriented in the 3' to 5' direction and encodes the putative glycoprotein B (gB) transmembrane peptide, which is 851 amino acids in length. Phylogenetic analysis of GTHV DBP, UL28, and gB genes substantiated that this novel *Chelonian* herpesvirus is closely related to the *Alphaherpesvirinae* subfamily. Examination of the translated amino acid sequence further supports this categorization since GTHV DBP comprises a highly conserved zinc finger motif ($CXLCX_4RX_2C$) and a putative DNA binding domain, and exhibits high sequence homology to other alphaherpesviruses. The gB peptide also showed high homology to alphaherpesviruses in the extraviral, transmembrane and C-terminal tail regions of the peptide. Cloning and sequencing the genome of this putative herpesvirus will facilitate current understanding of its role in causing GTFP and the development of molecular- and immuno-based methods for the diagnosis and prevention of this devastating disease of green sea turtles.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Green Turtle Fibropapilloma

Green turtle fibropapilloma (GTFP) is a *chelonian* disease characterized by the presence of benign, lobe -shaped tumors predominant on the fibrous tissue and within the visceral tissues and organs of green sea turtles (*Chelonia mydas*). The incidence of this disease has had serious detrimental effects on the quality of daily life of several species of sea turtles (Herbst 1999). Because of the seriousness of the threat posed by the disease, scientists began to study the disease soon after its discovery in 1938 (Smith and Coats 1938).

1.1.1 The Host Species

Chelonia mydas, or the green sea turtle, is the largest of the chelonids, with adults that can exceed one meter in carapace length and 200 kg in body mass. This species is referred to as "green", not because of the color of its shell, but because of the color of the subdermal fat of the turtles {National Marine and Fisheries Service (NMFS) 1998}. It is a circumglobal species found in tropical seas, and in a lesser extent, in subtropical waters with temperatures above 20°C. This includes the Atlantic, Pacific and Indian Oceans. In the Atlantic Ocean, the species can range from the coast of the Netherlands, southward to the coast of Portugal, and in the Western North Atlantic from Massachusetts through Bermuda, to the Florida Keys and the Bahamas (Ernst *et al.* 1994). The majority of the turtles in the Northern Atlantic are juveniles, or subadults. The species is also found in the Gulf of Mexico, the Caribbean Sea, and in to the South Atlantic through Argentina. *C. mydas*' range in the Pacific begins in Alaska and extends through Chile (Ernst *et al.* 1994).

Several characteristics and behaviors of the green sea turtle may contribute significantly to the effects of GTFP on individual turtles and the species survivorship. The green sea turtle is particularly slow growing and exhibits delayed sexual maturity (NMFS 1998). Studies by Balazs *et al.* indicate that the growth rate of the green sea turtle is on average about 2 cm straight carapace length per year (Balazs *et al.* 1994a). According to this data, it can be estimated that at least 25 years are required for a green sea turtle to achieve sexual maturity (NMFS 1998). This can contribute to the depletion of green sea turtles due to GTFP because if a turtle is struck early in life, it may be unlikely to reach an age at which reproduction can occur.

1.1.2 Epidemiology

Smith and Coates (1938) first described green turtle fibropapilloma as fibro-epithelial growths on the skin of large marine turtles, as observed in the Florida Keys. Although the disease has been documented as being principally an affliction of green sea turtles, similar fibropapillomatosis (FP) diseases are found to afflict loggerhead (*Caretta caretta*), and olive ridley (*Lepidochelys olivacea*) turtles (Quackenbush *et al.* 2001). Recently, a singular case of FP was reported in a leatherback turtle (*Dermochelys coriacea*) (Huerta *et al.* 2002), as well as in a Kemp's ridley turtle (*Lepidochelys kempii*) (Guillen and Villalobos 2000). However, the highest prevalence rates are found among the green sea turtle populations. For example, according to data collected by Schroder *et al.* (1998), 62% of the green sea turtles captured in Florida Bay between 1990 and 1996 exhibited GTFP lesions, while 11% of loggerhead turtles captured in the same area exhibited fibropapillomas. Similarly, 1-10% of olive ridley turtles nesting in Costa Rica

in 1997 exhibited fibropapillomas (Aguirre *et al.* 2000a). GTFP has also been reported in Australia and Brazil (Aguirre *et al.* 2000b), (Matushima *et al.* 2000).

Green sea turtles living and nesting in the waters around the Hawaiian Islands, in particular, are severely affected with GTFP. Annual strandings in Hawaii have significantly increased from 10-20 cases in the early 1980's to 200-300 cases in the late 1990's, with fibropapilloma as one of the major causes of the strandings (Murakawa *et al.* 2000). In waters surrounding Molokai, between 1982 and 1996 it is estimated that 61% of green sea turtles exhibited GTFP (Balazs *et al.* 1998). Green sea turtles inhabiting Kaneohe Bay showed a 43.9% incidence of fibropapillomas between 1989 and 1997 (Balazs *et al.* 2000b). In waters surrounding Maui, annual tumor prevalence is between 78 and 88%, the highest rate of any island in the chain (Murakawa *et al.* 2000).

The individual turtles affected with fibropapillomas may follow certain patterns. Data suggests that there is a female-biased sex ratio of turtles with tumors, compared to turtles without tumors. Recently, a study determined that the sex ratio of non-tumored turtles was 1.01 female to 1.0 male. The tumored turtles exhibited a sex ratio of 1.32 female to 1.0 male (Koga and Balazs 1996). The exact nature of the relationship is unknown and needs to be determined in reference to the etiology of the disease.

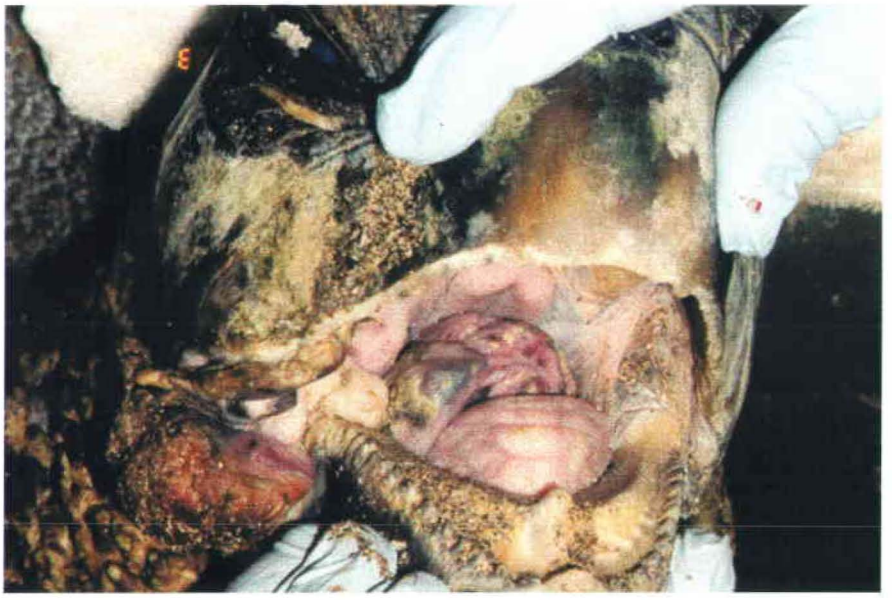
1.1.3 Clinical Symptoms

The most marked clinical manifestation of GTFP is the presence of multiple fibrous tumors that can range in size from a few centimeters to 25 cm in diameter (Balazs *et al.* 1998). The tumors are most commonly found on the external soft tissue of infected turtles including the hind- and foreflippers, eyelids, conjunctivae, neck, shoulders, and mouth (Balazs *et al.* 2000b) (Figure 1.1a). Tumors have also been associated internally

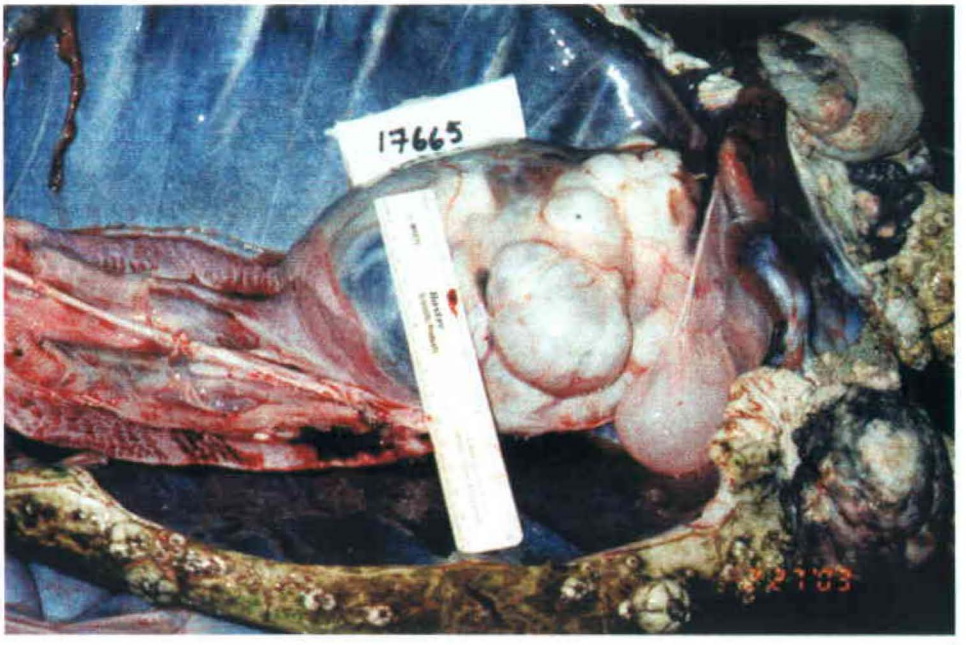
with the heart, lungs, liver, gonads, spleen and kidneys (Jacobson *et al.* 1991)(Figure 1.1b). Histopathologically, the tumors are characterized by marked hyperplasia of the epidermis, with anastomosing rete ridges extending into the dermis. Within the epidermis there is oftentimes a marked cytoplasmic vacuolation and ballooning degeneration of superficial epidermal cells. Many of these spongiotic cells contained eosinophillic intranuclear inclusions. The underlying dermis consists of proliferating fibroblasts with occasional trematode eggs within capillaries (Jacobson *et al.* 1991).

The biological properties of the fibropapillomas have been characterized using flow cytometery, particularly, the ploidy of tumor cells and the growth cycle of these cells (Papadi *et al.* 1995). This study indicated that all GTFP tumors studied had histograms indicative of normal diploid profiles. Other clinical effects of the disease may include nonregenerative anemia, hypoproteinemia, hypoalbuminemia, and several electrolyte abnormalities (Norton *et al.* 1990).

Behaviorally, green sea turtles affected with GTFP have been observed basking, a characteristic not common to green sea turtles. Basking during both the day and the night have been observed in both natural and experimental settings by tumored turtles (Swimmer *et al.* 1996). In the controlled experiments, basking was not observed in healthy turtles. This characteristic is common to freshwater turtles and is involved in thermoregulation. The body temperatures of the turtles was elevated after basking, and has been theorized to aid in the immune response of the turtles (Swimmer *et al.* 1996). The basking behavior may contribute to the large number of strandings of GTFP-afflicted turtles.



A. Oral tumors of a GTFP afflicted green sea turtle



B. Kidney tumors of a GTFP afflicted green sea turtle

Figure 1.1. Green sea turtles afflicted with GTFP. Both turtles were found in the Hawaiian Islands. Photos courtesy of Mr. George Balazs, National Marine and Fisheries Service.

1.2 Evidence for a Viral Etiology

Several theories were established about the cause of GTFP. The presence of spirorchid trematode ova in the vasculature of many fibropapillomas, observed in 1939, led some pathologists to believe that GTFP represents a fibroplastic reaction to trematode larva (Herbst *et al.* 1995). Environmental factors, such as pollution, have long been under scrutiny as the causative agent, or cofactor involved in the onset of GTFP. However, in 1991 preliminary experiments suggested that GTFP could be transferred in some form to disease free recipient turtles (Herbst *et al.* 1994). Cumulative evidence suggests that GTFP is caused by a filterable, subcellular agent, particularly it is most likely a virus (Herbst *et al.* 1994). In recent years, GTFP incidence has increased globally, spurring a great need for identification of the etiology of the disease. Intensive study has been underway to unmask the agent, and a major focal point of study has been the role of viral agents as the cause of GTFP.

1.2.1 Herpesvirus

The first paper presenting data that suggested the involvement of a novel herpesvirus in association with GTFP was written by Jacobson *et al.* in 1991. In this study, tumors from two green sea turtles severely affected with GTFP were surgically removed from the cervical integument, corneas, pericloacal region, and the hard and soft integument. The tumors were examined through various staining and light microscopy techniques, as well as transmission electron microscopy (Jacobson *et al.* 1991). Through light microscopy, previously unobserved areas of epidermal ballooning degeneration with eosinophillic, intranuclear inclusions were seen in several tumors. Further examination

through electron microscopy reveled that the intracellular inclusions contained immature virus particles with electron-lucent, and electron-dense centers. Viral particles were also seen within adjacent areas in the cytoplasm. Immature virus particles ranged from 77 to 90 nm and a viral envelope was obtained from the nuclear membrane. Mature, enveloped particles seen within adjacent areas of the cytoplasm were between 110 and 120 nm. Based upon the size, location and morphogenesis of the intracellular and extracellular particles, the virus particles were determined to be consistent with those of the *Herpetoviridae* family (Jacobson *et al.* 1991).

The direction of much of the research investigating GTFP was shifted to concentrate on identifying this herpesvirus. Infections by certain herpesviruses, in nature, have lead to the appearance of specific neoplasias. Specific elements of herpesviruses can be shown to induce morphological transformation in indicator cells (Wagner *et al.* 1985). In order to establish that a viral infectious agent was responsible for GTFP, a controlled transmission study was done (Herbst *et al.* 1995). In this study, cell-free tumor homogenates derived from tumors of GTFP affected turtles were used to infect live captive-reared turtles. This tissue was observed under a transmission electron microscope prior to infection. Live turtles were infected with homogenates that were filtered through a 0.45 μm syringe tip filter and homogenates in an unaltered state. Both groups of cell-free tumor homogenates were able to induce tumor development in the experimental turtles. Sites that were sham-inoculated did not develop tumors, nor did any uninoculated sites. Filtered GTFP extracts were no more successful at inducing tumors than non-filtered extracts. The transmission electron microscope (TEM) examination of the tumor tissue before and after inoculation showed the presence of an intranuclear,

herpes-like particles in various stages of viral assembly, as well as enveloped virus particles in the cytoplasm. Although these data were not able to clarify the ultimate causal agent of GTFP, the data strongly support the notion of a subcellular, filterable agent as the cause of GTFP. This study also excluded the possible etiological role of trematode ova, and corroborated data suggesting that these ova are not responsible for GTFP tumor development (Herbst *et al.* 1995).

In lieu of a purified virus with which to perform further characterization of the GTFP agent through transmission experiments, methods were imposed to provide basic information regarding the GTFP agent, and its infectivity. Herbst *et al.* (1996b) showed that the agent responsible for causing GTFP is chloroform-sensitive. The cell-free tumor homogenates were subjected to ultracentrifugation at 100,000g for 2 hours and separated into supernantent and pellet components. Turtles inoculated with the supernatant showed no induction of GTFP. The homogenate pellet, which theoretically contained a 5-fold increase in viral concentration, was able to induce tumors in captive –reared turtles. However, the pellet was less effective in inducing tumors than the starting material, suggesting a fragile nature to the GTFP agent. The results of the centrifugation experiment as well as the chloroform extraction support the notion of a large, enveloped virus, such as a herpesvirus, as the cause of GTFP (Herbst *et al.* 1996b).

Several molecular based polymerase chain reaction (PCR) techniques were used to further implicate the role of the herpesvirus as the etiologic agent of GTFP. Highly conserved herpesviral DNA sequences of the DNA polymerase gene (DNA *pol*) were used to develop degenerate primers. These primers were capable of amplifying DNA extracted from green sea turtle tumor tissues (Quackenbush *et al.* 1998). A sequence of

237 base pairs of the DNA *pol* gene of a novel virus, deemed green turtle herpesvirus (GTHV), was amplified. A comparative search of the GenBank database indicated that GTHV was closely related to several other herpesviruses, including pseudorabies virus and bovine herpesvirus. Primers were then designed from the sequences of these two viruses and used to amplify a 483 bp sequence of GTHV (Quackenbush *et al.* 1998).

Using this sequence information, additional studies were done that determined the distribution of the herpesvirus among tumored and non-tumored tissues of GTFP infected, and normal tissues of uninfected green sea turtles. 100% of DNA extracted from fibropapillomas of green sea turtles nesting in Hawaii contained herpesviral DNA sequences (N=23) (Quackenbush *et al.* 1998). A comparative study was performed by Lu *et al.* (2000b), which used a nested PCR technique to assess the presence of GTHV DNA in specific tissues of the Hawaiian green sea turtle. Normal-appearing tissue from GTFP infected animals, with the exception of PBMC's, showed the presence of GTHV sequences in over 80% of the tissue/organ samples. When nested PCR was attempted on Hawaiian green sea turtles not exhibiting GTFP, 0% of the tissues were found to have GTHV sequences. Finally when nested PCR was used to determine if GTHV sequences were present in the tumors, 100% of the tissue samples were positive (Lu *et al.* 2000b).

Several other species of marine turtles are also affected by marine turtle fibropapillomas, including loggerhead turtles, and olive ridley turtles. When tissue and tumor samples from loggerhead turtles, as well as green sea turtles nesting in Florida, were subjected to PCR, using the GTHV specific primers, similar results were seen as reported for green sea turtles nesting in Hawaii (Lackovich *et al.* 1999). 93% of the tumor tissue exhibited herpesviral sequences. This number may be lower than is seen in the

green sea turtles nesting around the Hawaiian Islands because nested PCR is a more sensitive technique, and is able to amplify lower quantities of DNA (Lu *et al.* 2000b). Sequence analysis was performed on sequences obtained from the amplicons from Floridian green sea turtles, and loggerhead turtles. The sequences predicted that the herpesvirus of the loggerhead turtles differed from those of the green sea turtle by only 1 of 60 amino acids (Lackovich *et al.* 1999). This indicates only a minor variation in the two herpesviruses. The data obtained from the PCR experiments indicate that GTHV is regularly associated with GTFP, and not just a coincidental finding.

Taking this technique one step further, Real-Time PCR was used to quantify the levels of turtle herpesviruses (THV) DNA in various tissues. This technique was used to estimate the viral load of herpesviral DNA in samples from green, loggerhead and olive ridley sea turtles from various locations, and using tissues from several different internal organs, skin and tumors (Quackenbush *et al.* 2001). The viral loads of THV DNA in tissues obtained from tumors were estimated to be between 2 and 20 virus copies per cell. The viral load for other organs and skin varied with a range of .001 to 170 copies per cell.

Finally, using reverse transcription PCR, expression of the GTHV DNA *pol* gene was determined in the tumors and normal-appearing nontumored tissues and organs from turtles with fibropapillomas. Amplification from GTHV DNA *pol* RNA from tumor tissues occurred in 100% of the specimens examined, indicating that these genes are expressed in tumor tissues. Limited DNA *pol* gene expression was seen in the normal appearing tissue of tumored animals (Lu *et al.* 2003). These data indicate that GTHV replication occurs in the tumor tissues, and in limited amounts in uninvolved tissues.

1.2.2 Papillomavirus

Papillomaviruses, DNA viruses of the *Papovaviridae* family, are known to cause fibropapillomas similar to those seen in green sea turtles in other many vertebrate species (Lackovich *et al.* 1998), causing papillomaviruses to become an initial, suspected etiological agent of GTFP. In 1989, Jacobson *et al.* performed several techniques to test for the presence of episomal papillomavirus DNA. All fibropapillomas examined were negative for papillomavirus group-specific antigens by the peroxidase-antiperoxidase test. Papillomavirus DNA could not be demonstrated in any of the fibropapillomas examined by ethidium-bromide stained gels, or by low stringency Southern blot hybridization, with a bovine papillomavirus type-2 derived DNA probe. Additionally a reverse Southern blot against 25 different cloned papillomaviruses also failed to demonstrate papillomavirus DNA in all tumors studied (Jacobson *et al.* 1989).

To date papilloma-like particles have not been visualized through light or electron microscopy, and have not been isolated in tissue culture (Aguirre *et al.* 1994a). The difficulties involved in papillpomavirus visualization by light and electron microscopy may have played a role in these findings. In order to further investigate the association between GTFP and papillomaviruses, a PCR-based method was performed using degenerate primers that have the ability to detect human papillomavirus with accuracy greater than 70% (Lackovich *et al.* 2000). While the positive control yielded the expected PCR results, no fibropapilloma was positive for papillomavirus sequences. This evidence, and the fact that the tumor-inducing filtrate of GTFP is sensitive to chloroform inactivation, to which papillomaviruses are insensitive, has lead scientists to look towards other agents as the cause of GTFP.

Recently, however, small, naked virus particles have been identified in tumor-like aggregates in cell lines derived from a green sea turtle (Lu *et al.* 2000a). Tumor aggregates occurred in cell lines derived from turtle tumor, lung, periorbital, and testicular tissues. This cell line was used to make a cell free medium with which healthy green sea turtle lung cells were inoculated. Aggregates were seen in these cells, but not seen in healthy lung cultures that were not inoculated. By thin-section electron microscopy, massive collagen deposits were visualized both intracellularly, and extracellularly. Additionally, virus particles, which were round and naked with an electron dense nucleocapsid measuring 50+-5 nm in diameter, were found exclusively within the nuclei of affected cells, either freely scattered or largely aggregated as viral assays. Viral envelopes and viral budding were not observed. Several distinct viral arrays were present usually in each cell. Each array consisted of viral particles of relatively uniform size ranging from 40-55nm in diameter. Nuclear degeneration, loss of nuclear chromatin and break down of cellular organelles was observed in infected cells (Lu *et al.* 2000a).

Based on the size, morphology, and intranuclear location, this small naked virus most closely resembles a papovavirus, possibly a papillomavirus. PCR cycling was performed using primers derived from highly conserved regions of known papovaviruses. Amplification of a PCR product occurred, but it was significantly smaller that the expected fragment; 235 base pairs instead of 480 base pair expected result. This 235 base pair sequence showed low sequence similarity, 36-44%, with some of the papillomaviruses of humans and animals. Because of the irregular size and low sequence

similarity, it is inconclusive if this sequence is true papovavirus amplification (Lu *et al.* 2000a).

1.2.3 Retrovirus

Retroviruses have the well-established ability to induce proliferitive diseases in birds, mammals and fish (Casey *et al.* 1997). Not very much is known about the role they play in reptiles, invertebrates, or their role in tumor induction. This may be due to the fact that no cell culture systems are available for the propagation of reptilian retroviruses. Due to this deficit, polymerase enhanced reverse transcriptase (PERT) assays were performed to measure the amount of retroviral DNA comparatively in tumored and non-tumored tissues. Strikingly, both tumor and non-tumor tissues were positive for reverse transcriptase activity, indicating the presence of a retrovirus. The number of virons per sample was estimated to be high. However, the samples of skin from asymptomatic turtles were reverse transcriptase positive as well (Casey *et al.* 1997). The presence of retrovirus in certain tumor samples was assessed using sucrose gradient centrifugation. Tumors from the skin and heart banded at a level characteristic of retroviruses, while the eye sample did not (Casey *et al.* 1997).

Electron microscopy was then performed on the samples and numerous virus-like particles ranging in size from 96 to 122 nm were seen. Surface projections corresponding to envelope glycoproteins were apparent on these particles, morphology similar to the prominent spikes seen on the retrovirus, WDSV (Casey *et al.* 1997). Additional evidence of retroviral activity is provided by the sodium dodecylsulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the sucrose-gradient purified tumor samples and uninvolved skin samples. These data showed a conserved profile among four unrelated

tumors and showed seven prominent proteins. Most importantly, the tumor-specific patterns differed from the non-specific background staining observed in uninvolved skin samples (Casey *et al.* 1997).

The questions regarding the involvement of a retrovirus as the causative agent, or an agent associated with GTFP infection still remain largely unanswered. This is because first, reverse transcriptase assays cannot distinguish between different retroviruses. It remains possible that green sea turtles are hosts to more than one retrovirus, capable of different disease potential. Secondly, only preliminary data on the involvement of retroviruses with GTFP have been done. Elucidation of the relationship is complicated by the issues of not knowing how the virus is transmitted, the tissue tropism, mode of expression and pathogenicity associated with infection (Casey *et al.* 1997). It is apparent through these studies that green sea turtles are retroviral hosts, but the relationship between these viruses, and GTFP is still unclear.

1.3 Non-Viral Cofactors

Although many scientists feel that GTFP has a viral etiology, there are several other etiologies, and cofactors currently being investigated. Chemical pollution, as previously stated, may play a role in the etiology of GTFP. For example, polychlorinated biphenyls (PBC's) and PBC congeners have been shown to affect sex ratios in several turtle species (Miao *et al.* 2001). The levels of PBC's were measured from tissues of turtles stranded with invasive tumors, and the toxicicity levels were calculated. It was found that hexachlorobiphenyls, which are toxic, were found in high concentrations in the liver and adipose tissues of the turtle specimens (Miao *et al.* 2001). Although these studies are not conclusive in linking PBC concentrations to GTFP, they indicate

abnormalities that have the potential to effect the green sea turtle populations. Additional studies, with larger data sets are required to further analyze the problem.

The role of natural tumor promoters in GTFP has been investigated also. Toxic benthic dinoflagellates form the *Prorocentrum* species are distributed worldwide and have the potential to produce a tumor promoter called okadaic acid (OA). These dinoflagellates are epiphytic on macroalgae and seagrasses that are normal components of green sea turtle diets. It has been shown that green sea turtles in the Hawaiian Islands consume *Prorocentrum* and that high-risk FP areas are associated with areas where *P. lima* and *P. concavum* are both highly prevalent and abundant (Landsberg *et al.* 1999). The presence of presumptive OA in tissues of green sea turtles nesting around the Hawaiian Islands indicates that these turtles may have been exposed to such tumor promoters. These data, although also not complete, suggest a potential role for OA as a tumor promoter in the etiology of GTFP (Landsberg *et al.* 1999).

1.4 GTHV Genome

GTHV needs to be isolated and Koch's postulates fulfilled in order to determine this herpesvirus is the etiological agent causing GTFP. Unfortunately, isolation and propagation of GTHV has not been achieved *in vitro* to date. In lieu of isolation and purification of GTHV for further transmission experiments, cloning and sequencing of the GTHV genome have been done in order to characterize potential pathogenic genes of this herpesvirus. A modified genomic walking technique mediated by IPCR was used to obtain the full-length gene sequence of the GTHV DNA *pol* gene (UL30), as well as partial sequences from the upstream GTHV DNA binding protein gene (UL29), and the downstream UL31 gene (Yu *et al.* 2001). The entire length of the sequence obtained from

that study was 4837-bp long, including 3507-bp of the DNA *pol* gene, 357-bp of the DNA binding protein (DBP) gene and 723-bp of the UL31 gene. The genome information obtained in this previous study is shown in Figure 1.2. Phylogenetic analysis of the sequence of the GTHV DNA *pol* gene showed that GTHV is most closely related to the alphaherpesvirus subfamily.

The genome of members of the herpesvirus family is one molecule of linear double-stranded DNA. The total genome length among the family is between 120- 200 kilobases (kB) in length. The genome of alphaherpesviruses, particularly the well-studied HHV-1 genome, can be divided into six regions: The unique long (UL), unique short (US), a long repeated sequence (RL), a short repeated segment (RS), the origins of replication, and the ends of the linear molecules. The sequencing in this experiment targets genes in the UL region of the GTHV genome. The UL region in HHV-1, which is 108 kb long, encodes at least 56 distinct proteins. It contains genes for the DNA replication enzymes and capsid proteins, and many other proteins (MeGeoch *et al.* 1989).

The specific goals of this study are:

1. To modify and use the genomic walking technique to obtain the full-length nucleotide sequence of the UL29, UL28, and UL27 genes of GTHV.
2. To analyze the nucleotide sequence of the UL29-UL27 genes in order to identify transcriptional elements of the sequences.
3. To translate the UL29-UL27 genes into putative amino acid sequences and identify functional portions of these sequences.
4. To analyze the obtained nucleotide and amino acid sequences of the GTHV UL29-UL27 genes in order to aid in the classification of GTHV.

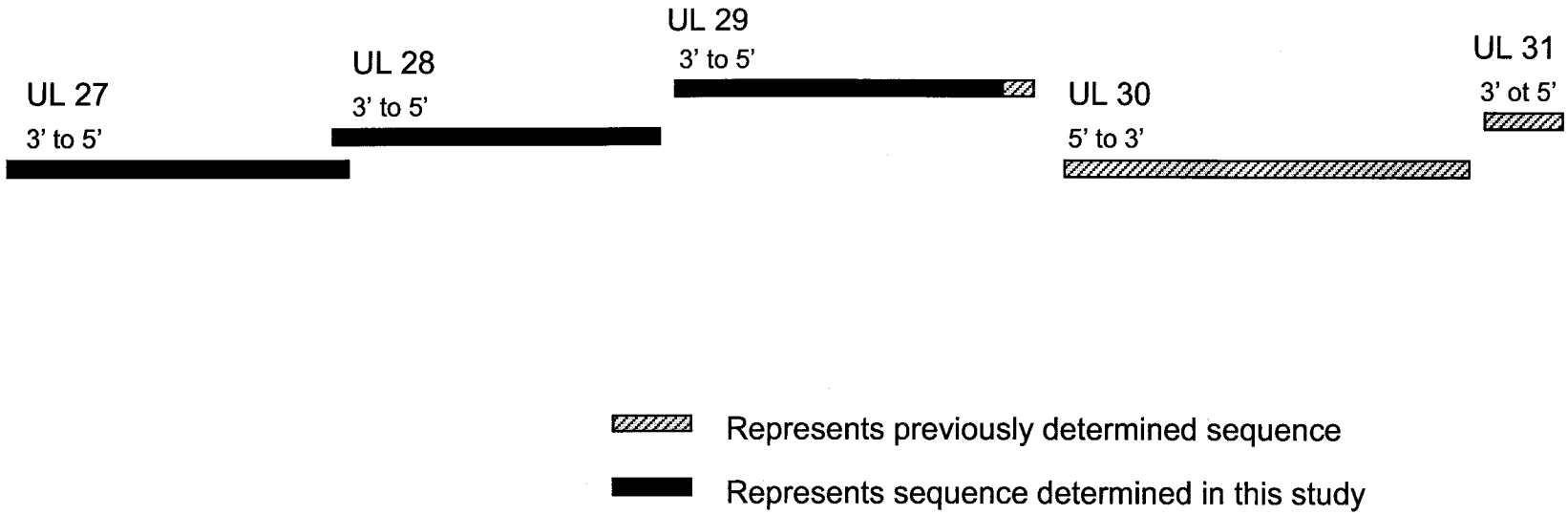


Figure 1.2. Known GTHV gene sequence information. The dashed boxes indicate sequences obtained during previous studies. The solid boxes indicate sequence information obtained through this study. The direction of translation of the gene is indicated below the gene name.

CHAPTER 2: METHODS

2.1 DNA Isolation from Tumor Specimens

Approximately 1 gram (g) of tumor tissue from two tumor samples (A17-14a and A18-14e) was surgically removed from the tongues of two green sea turtle with GTFP. Turtle A17-14a (ID# 15095) was an immature female turtle with multiple tumors on the eye, skin, mouth and glottis. No internal tumors were seen. The straight carapace length of the turtle was 47.5 cm and it weighed 25lbs. The turtle carcass was found on Oahu. Turtle A18-14e (ID# 15096) was a sub-adult male turtle with numerous tumors on the eyes, skin and mouth. Wounds indicating a shark attack lead to the stranding of this turtle in Maui. The straight carapace length of the turtle was 76 cm and the turtle weighed 83.4 lbs. The tumors were separately minced into small pieces using a scalpel and then ground into fine powder using a pre-chilled mortar and pestle. Some sea sand was added to facilitate the grinding. The tissue was placed in a 50 ml propylene tube (Falcon). Next, 12 µl of digestion buffer (100 mM NaCl; 10 mM Tris-HCl, pH 8.0; 25 mM; 25 mM EDTA, ph 8.0; and 0.5% sodium dodecyl sulfate) was added per g of tissue, and buffer was used to rinse the mortar and pestle to ensure maximum recovery. Next, 120 µl/(g of tissue) of proteinase K was added to the tube. The tumor homogenate was incubated for 18 h at 55°C. Protein was removed from the digests by successive phenol/ chloroform/ isoamyl alcohol (IAA) extractions. Phenol was added at a ratio of 2.5x the total volume of the reaction mixture. The tube was inverted gently, and then spun in a Centra GP8R centrifuge (IEC) at 3000 RPM for 10 min. After centrifugation the top layer was harvested and phenol/ chloroform/ IAA (24:24:1) was added at a ratio of 1x to the remaining volume of solution. The mixture was centrifuged in the same manner, and the

top layer was harvested. After centrifugation, chloroform/ IAA (24:1) was added at a ratio of 1x the reaction volume. The tube was centrifuged in the same manner, and the top layer was harvested. The DNA was recovered by 100% ethanol precipitation (2.5x) at -30°C overnight. The DNA was then dried and resuspended in water. Using a DU 640 Spectrophotometer (Beckman), the concentration of the tumor DNA was determined. The tumor DNA was diluted to a final concentration of 1μg/μl and stored at 4°C.

2.2 PCR Determination of GTHV in Tumor Tissues

Polymerase chain reaction (PCR) was used to determine that DNA extracted from tumor tissues contained GTHV sequences. Forward and reverse oligonucleotide primers were designed from a sequence of the GTHV DNA *pol* gene (Genebank ascension number: AF23984) with the sequences: 5'AGCATCATCCAGGCCCACAATCT 3' and 5'CGGCCAGTTCCGGCGCGTCGACCA 3', respectively. The primers were designed to amplify a 455-bp fragment of the gene. Amplification was initiated by adding 0.5 μg of tumor-extracted turtle DNA to a 25 μl reaction, containing 0.2 mM each dNTP, 10mM Tris-HCL, 50 mM KCl, 1.5 mM MgCl, and 0.25 mM each of the forward and reverse oligonucleotides into 0.2-ml snap-cap PCR tubes. Subsequently, 1.25 U of Ampli*Taq* (Perkin-Elmer) were added to the reaction mixture and it was denatured at 95°C for 5 min. This was followed by 45 cycles of 94°C for 30 sec, 50°C for 30 sec and 72°C for 60 sec in a Perkin-Elmer GeneAmp PCR System 2400. After cycling, the reaction mixture was subjected to a final extension of 7 min at 72°C, and then stored at 4°C. PCR products were size fractionized on 2.0 % agarose gel, and stained with ethidium bromide. Positive amplifications were selected based on size (Figure 2.1) and used for preparation of GTHV DNA for IPCR.

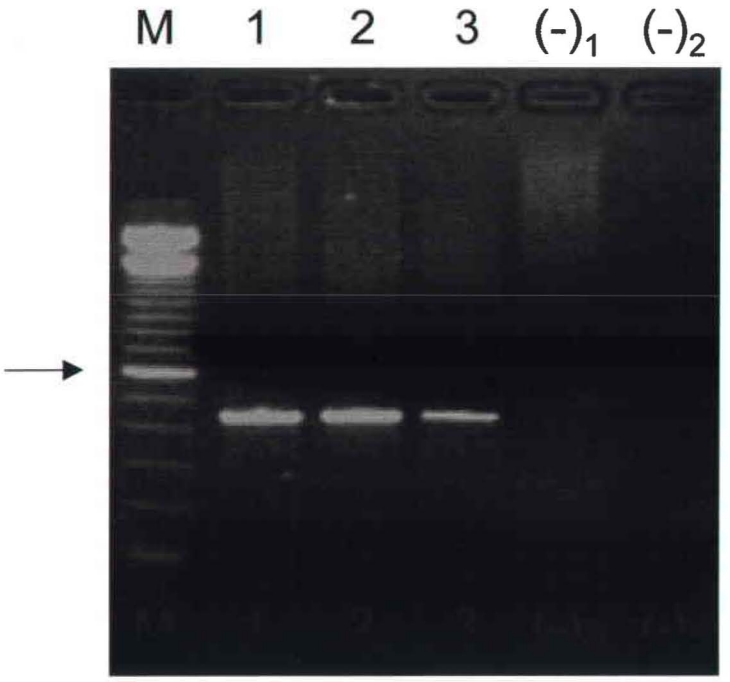


Figure 2.1. Gel electrophoresis of PCR on tumor tissue-extracted GTHV DNA. The DNA was run electrophoretically on 2.0% agarose gel and stained with ethidium bromide. Lane 1 shows the PCR results of sample A17-14e. Lane 2 shows the PCR results of sample A18-14e. Lane 3 shows the PCR results of the positive control A15-14e. Lane $(-)_1$ shows the PCR results of a negative control using A17-14e without *Taq* polymerase. Lane $(-)_2$ shows the PCR results of a negative control using normal reaction conditions and DEPC H_2O in lieu of DNA. The M lane shows a 100 bp ladder (Roche) and the arrow indicates 600 bp.

2.3 Genomic Walking

The genomic walking technique is based on a restriction enzyme digestion followed by self-ligation and IPCR. It was employed in this study to amplify sequences of DNA that are flanking the previously known DNA *pol* gene. Because the virus has not been isolated or purified, this method provides a reliable and efficient means to attain sequence data without isolation of the virus.

2.3.1 Restriction Enzyme Selection and IPCR Oligonucleotide Primer Design

A previously determined sequence of GTHV (Yu *et al.* 2001) was used to design oligonucleotide primers and select restriction enzymes (RE's) to cut the GTHV DNA for the first genomic walk. A priming region was determined to be the 250 bp from the 5' end of the known sequence. IPCR oligonucleotide primers were designed within this section, and RE's that did not cut within the priming region were selected. For each subsequent genomic walk, a priming region was determined. The size of the region varied based on the location of the RE sites. IPCR oligonucleotide primers designed for each genomic walk are shown in Table 2.1. IPCR primers were oriented in such a way that they point away from each other and the G+C content of each primer pair was targeted at 50%.

For each genomic walk, a RE map was created using the MapDraw program of the DNAstar package. RE's were selected based on several criteria. First the RE must not cut within the priming region for the specific genomic walk, and should only cut once in the region adjacent to the priming region. Only RE's that left sticky ends with 4-bp overhangs were selected. These 4-bp overhangs are preferred because they occur theoretically once in 256 bp (Hartl and Ochman 1994) allowing the potential amplicon to

Table 2.1. Table of IPCR oligonucleotide primers. The primers used for each of the three genes are separated by double lines. Primer pairs are indicated by color blocks.

PRIMER PAIRS						
Gene	Walk	Name	Sequence	Length	% G+C	IPCR Annealing Temp
UL 29	1	UL29-3B	5' AGA AGG GCC GCT GGT TTT CAA 3'	21-mer	52.38	60°C
UL 29	1	UL29-3T	5' CGA GCT ACT TTT ACC TTC GG 3'	20-mer	50	60°C
UL 29	2	UL29-4B	5' GAT CGG GAT CGA ACT CGG AAA AAC 3'	24-mer	50	60°C
UL 29	2	UL29-4T	5' ATA TGT CGG CAC GTT CAG ATG GAG 3'	24 -mer	50	60°C
UL 29	3	UL29-5B	5' TTG ACG TCG CTG ACC ACC TTT ACG AAC GAT 3'	30-mer	50	63°C
UL 29	3	UL29-5T	5' AGC GCT TCG TGC AAG GAA TAT CCG AAG ACT 3'	30-mer	50	63°C
UL 29	4	UL29-6B	5' GGC CGT AAC GTC GCG GGT GTT GAA AAA 3'	27-mer	55.56	65°C
UL 29	4	UL29-6T	5' AGC GAA CGG CAG TAC GAG TTT GAG ATC 3'	27-mer	51.18	65°C
UL 28	1	UL28-1T	5' AAG CGG ATT ACG CCC TCT ATA GCC T 3'	25-mer	52	65°C
UL 28	1	UL28-1B	5' AAC CTC CAA GCG ATC GCG CAA GTA A 3'	25-mer	52	65°C
UL 28	2	UL28-2T	5' GCA AAG AGG AGG ACG GCC TTT ATC TGA C 3'	28-mer	53.57	63°C
UL 28	2	UL28-2B	5' GAA GCG AGC ACC AAC TCT CGA ACG TAT C 3'	28-mer	46.43	63°C
UL27	1	UL28-2T	5' GCA AAG AGG AGG ACG GCC TTT ATC TGA C 3'	28-mer	53.57	63°C
UL27	1	UL28-2B	5' GAA GCG AGC ACC AAC TCT CGA ACG TAT C 3'	28-mer	46.43	63°C

be a manageable size for cloning and sequencing. Additionally, sticky-ended DNA is better suited for self-ligation than blunt-cut DNA. Table 2.2 shows the restriction enzymes selected and used in this study.

2.3.2 Restriction Enzyme Digestion

For each genomic walk, 10 μg of tumor-extracted GTHV positive green sea turtle DNA was digested with 20U of the selected RE (*Hae* II, *Ngo*M IV, *Eco*R I, *Btg* I, *Ava* I, *Afl* III, or *Nhe* I) (New England Biolabs), in a total reaction volume of 100 μl containing 10 μl one of the 10x RE buffers (Buffer 2: 50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol; Buffer 3: 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol; Buffer 4: 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol; or *Eco*R I specific Buffer: 50 mM NaCl, 100 mM Tris-HCl, 10 mM $MgCl_2$, 0.025% Triton X-100) in a 1.5 ml snap-cap tube. The selected buffer for each RE is indicated in Table 2.2. The RE digests using *Btg* I, *Afl* III, and *Nhe* I also required 1 μl of 100x bovine serum albumin (BSA). All digests were incubated at 37°C for 2 hours and followed by heat inactivation at the following times and temperatures: *Ava* I, *Btg* I, *Hae* II and *Ngo*M IV; 80°C for 20 minutes, *Afl* III, *Eco*R I and *Nhe* I; 65°C for 20 minutes. For the UL29 DBP gene, *Hae* II and *Ngo*M IV were used for the first walk, *Eco*R I was used for the second walk and *Btg* I was used for the third walk. *Ava* I, *Sac* II, and *Nsp* I were also used to digest DNA for the third walk, but later did not produce bands in IPCR. Circularized DNA prepared previously by *Eco*R I digestion and self-ligation was reused for the fourth walk. For the UL28 gene, *Ava* I and *Btg* I were the RE's selected for the first walk; however only *Ava* I-cut DNA was used for

Table 2.2. Restriction Enzymes selected for GTHV DNA digestion. The sequence cut by each specific enzyme is listed in the second column. Non-productive RE's refers to RE's that did not result in a strong, single band after the DNA was ligated and subjected to IPCR.

Restriction Enzymes				
Enzyme	Sequence Cut	Gene	Walk	NEB Buffer
Hae II	PGCGC'Y Y'CGCGP	UL29	1	4
NgoM IV	G'CCGGC CGGCC'G	UL29	1	4
EcoR I	G'AATTC CTTAA'G	UL29	2, 4	*EcoR* I specific buffer
Btg I	C'CPYGG GGYPC'C	UL29	3	3
Ava I	C'YCGPG GPGCY'C	UL28	1	4
Afl III	A'CPYGT TGYPC'A	UL28	2	3
Nhe I	G'CTAGC CGATC'G	UL27	1	2
Non-Productictive Restriction Enzymes				
Enzyme	Sequence Cut	Gene	Walk	NEB Buffer
Ava I	C'YCGPG GPGCY'C	UL29	3	4
Nsp I	PCATG'Y Y'GTACP	UL29	3	2
Sac II	CCGC'GG GG'CGCC	UL29	3	4
Apo I	P'AATTY YTTAA'P	UL28	1	3
Hae II	PGCGC'Y Y'CGCGP	UL28	1	4
Hinf I	G'ANTC CTNA'G	UL28	1	2

cloning. *Afl* III, *Apo* I, *Hinf* I, and *Hae* II were selected for DNA digestion for the second walk, but only *Afl* III was used for cloning. For the UL27 Glycoprotein B gene, *Nhe* I was used. The cutting sequences of the RE's used in this study are shown in Table 2.2. RE digestion was confirmed by gel electrophoresis in a 0.8% agarose gel with ethidium bromide staining.

2.3.4 Phenol/ Chloroform/ Isoamyl Alcohol Extraction

RE digested DNA was purified by phenol/chloroform/isoamyl alcohol extraction to the following protocol. Equal volumes of 25:24:1 phenol/chloroform/isoamyl alcohol (Sigma) and RE digested DNA were combined in a 1.5 ml snap-cap tube, and incubated on ice for 5 minutes. The mixture was then spun at 14,000 RPM for 20 minutes at 4°C in a 5417R centrifuge (Eppendorf). The top aqueous layer of the mixture was harvested using a 200 µl pipette, and transferred to a new 1.5 ml tube. An equal volume of 24:1 chloroform/ isoamyl alcohol was added to the tube. The mixture was incubated on ice, centrifuged, and the top aqueous layer was harvested as described above. 2.5x of 100% ethanol was added to the mixture and the mixture was then incubated at -20°C overnight. Following incubation, the mixture was centrifuged at 14,000 RPM for 30 minutes at 4°C. The supernatant was removed using a 200 µl pipette. The pellet was then washed with 200 µl of 75% ethanol, and incubated on ice for 10 minutes. The mixture was the centrifuged at 14,000 RPM for 20 minutes at 4°C. The supernatant was poured off, and the pellet was allowed to air dry, and then resuspended in 30 µl of DEPC treated H_20. The concentration of DNA was determined using a DU 640 Spectrophotometer

(Beckman), and the final concentration was adjusted to 0.5 μg/μl. The RE digested DNA was stored at -30°C.

2.3.5 T4 Ligation

One microgram of RE digested DNA was circularized in a reaction volume of 1 ml, containing 50 Weiss Units of T4 Ligase (Promega) and 100 μl of the supplied 10x buffer (300 mMTris-HCl, 100 mM $MgCl_2$, 100 mM DTT and 10 mM ATP). The mixture was incubated overnight at 16°C, after which the mixture was heat inactivated for 10 minutes at 70°C, in order to stop the T4 ligation activity before IPCR. The large reaction volume is used in order to ensure the formation of monomeric circles instead of long, linear concatomers. Ligation reactions were stored at -30°C between uses.

2.3.6 Inverse Polymerase Chain Reaction (IPCR)

IPCR was performed on the self-ligated DNA using the Expand High Fidelity PCR System (Roche). The oligonucleotide primers designed for each walk of each gene are shown in Table 2.1. Primers are designed in such a manner as to be pointing away from each other, differentiating PCR from IPCR. The primers bind to specific sites on the self-ligated DNA therefore eliminating the need for separation of the circularized DNA. The same primers were used for the second walk of the UL28 gene and the first walk of the UL27 gene, because the RE sites of *Afl* III and *Nhe* I were close enough to allow the same priming region. The PCR reaction was performed in a final volume of 50 μl containing 200 μM of each dNTP, 2.6 Units of Expand High Fidelity Polymerase, 300 nM of both the forward and reverse primers, 1.5 mM $MgCl_2$, 5 μl of the supplied 10x buffer and 15 ng of the self-ligated template DNA. When preparing the IPCR reaction mixture, upper and lower reaction mixtures were made. The lower contained both

forward and reverse primers, DNTP's and DEPC treated H_2O to 10 µl. The upper mixture contained Expand High Fidelity Polymerase, $MgCl_2$, 10x buffer, and DEPC treated H_2O to 25 µl. The lower mixture was aliquoted first, followed by 15 µl of solution containing the DNA. The upper mixture was added to the solution just prior to the beginning of the IPCR. This method ensures that no reaction will take place before the initiation of IPCR. The reaction was denatured at 95°C for 5 minutes and cycled 45 times using a Perkin-Elmer GeneAmp PCR System 2400 with the following conditions: 94°C for 30 seconds, variable annealing temperatures for 60 seconds and 70°C for 3 minutes. The annealing temperature varied due to the characteristics of the primers. The annealing temperature used in each walk is shown in Table 2.1. The long extension time of 3 minutes was used to ensure that the DNA circle produced by the digestion and subsequent ligation had sufficient time to be amplified. The size of the DNA circle is unknown prior to IPCR, so the long extension time is necessary for all reactions. Following the final cycle, the IPCR reaction is subjected to a final extension at 72°C for 7 minutes. The purpose of this extension is to add a poly-A tail to the amplicon for subsequent TA cloning. IPCR reactions were then stored at 4°C. Eight microliters of each IPCR product were size fractionalized electrophoretically on 1.5% agarose gel and stained with ethidium bromide, and returned to storage at 4°C. This entire process of RE digestion self-ligation and IPCR is called genomic walking. A schematic drawing of genomic walking is shown in Figure 2.2.

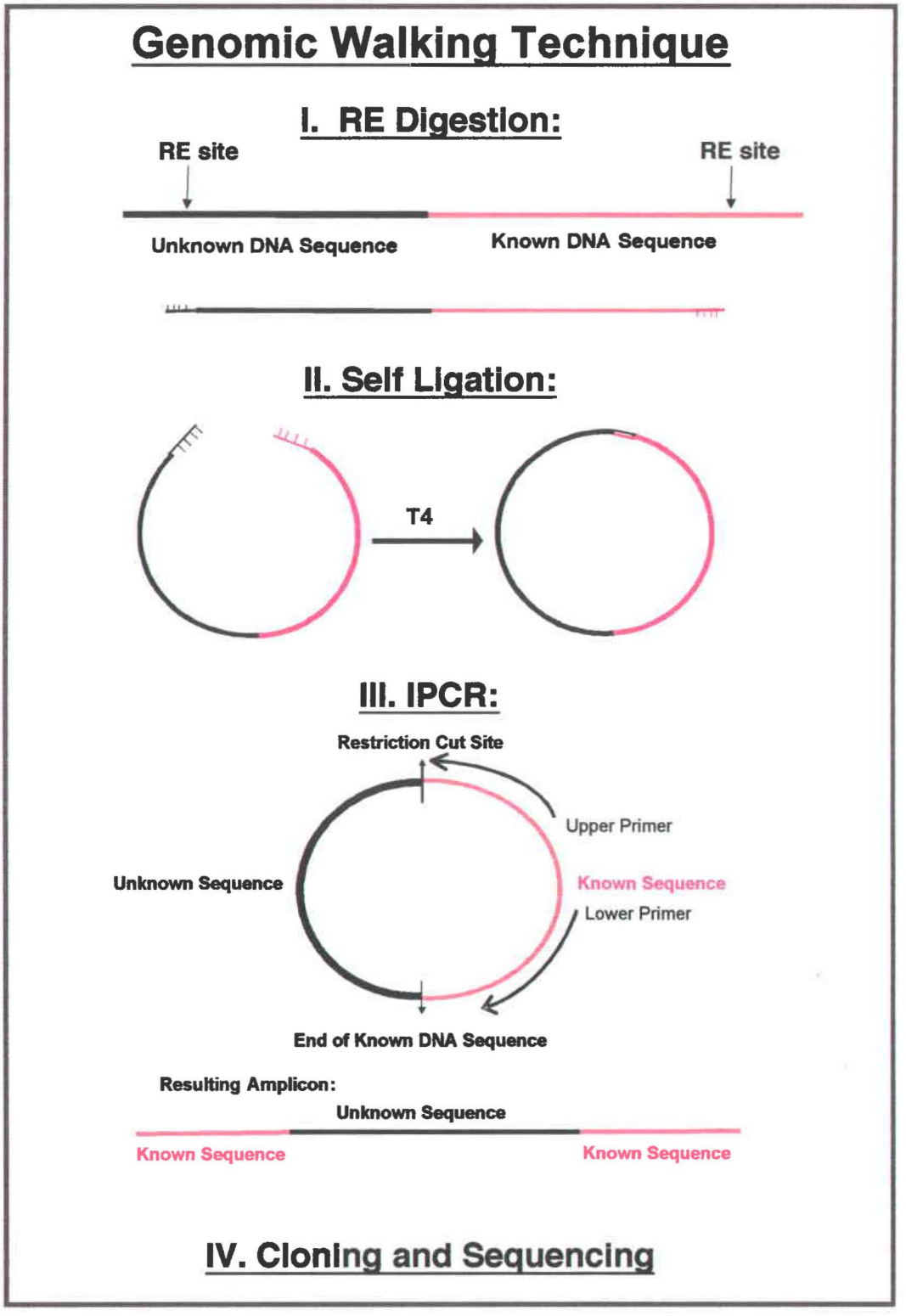


Figure 2.2. Schematic drawing of the genomic walking technique using IPCR.

2.3.7 TA Cloning of IPCR Products

IPCR products that appeared as a single, clear, strong band were ligated to the TA cloning vector. Ten ng of the IPCR product were incubated overnight, in a water bath at 14°C with 50 ng of pCR 2.1 vector (Invitrogen), 4 Weiss Units of T4 ligase (Invitrogen), and 1 µl of the supplied 10x ligation buffer (60 mM Tris-HCl, 60 mM $MgCl_2$, 50 mM NaCl, 1 mg/ml BSA, 70 mM β-mercaptoethanol, 1 mM ATP, 20 mM dithiothreitol, and 10 mM spermidine) in a total reaction volume of 10 µl. After the ligation, 1 vial of Invαf (Invitrogen) competent cells per each ligation reaction was thawed on ice for five minutes. Two microliters of the ligation reaction mixture was added to each vial of competent cells, without stirring or pipetting up and down. The competent cell mixture was incubated on ice for 30 minutes, followed by heat shocking for 30 seconds in a 42°C water bath. After the heat shock, vials were immediately placed on ice for 2 minutes. Two hundred and fifty microliters of SOC media (2.0% Trypone, 0.5% Yeast Extract, 10.0 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$-$6H_20$, and 20 mM glucose) were added to each vial of competent cells. The vials were incubated in a shaking incubator for 1 hour at 37°C and 225 RPM.

LB (Luria-Bertani) liquid media and agar plates containing ampicillin had been previously prepared in the following manner: 1.0% Tryptone, 0.5% Yeast Extract 1.0% NaCl were combined in an autoclavable 1 liter glass bottle. The pH was adjusted to 7.0, using an accumet pH meter 50 (Fisher Scientific). For agar plates, 1.5% agar was added to the solution. Both solutions were autoclaved on the liquid cycle for 40 minutes. The solutions were allowed to cool to 55°C and 50 µg/ml of ampicillin was added. The liquid media was stored at 4°C. The agar solution was poured into 10 cm plates and allowed to

cool. The plates were inverted and then stored at 4°C. Before use, 40 µl of 40 mg/ml X-Gal (400 mg X-Gal dissolved in 10 ml dimethylformamide) was spread evenly on the surface of each LB agar plate, using a glass bacterial spreader, flamed before each use. The X-Gal was allowed to absorb for at least 15 minutes prior to bacterial plating.

Each vial of transformed Invαf competent cells was plated twice, using volumes of 50µl and 100 µl of the competent cell culture. The plates were incubated at 37°C for 18 hours, and then placed at 4°C to allow color development. Colonies were chosen based on blue/white screening. White colonies were picked with disposable inoculators and inoculated into 3 ml of LB-media. The cultures were incubated overnight at 37°C in a shaking incubator set at 250 RPM. After incubation, the cultures were spun in a Centra GP8R centrifuge (IEC) at 2500 RPM for 20 minutes. The excess media was poured into a 10% bleach solution. Plasmids were extracted from the cell using the Qiagen miniprep kit, and following the QIAprep Spin Miniprep Protocol, using a microcentrifuge. According to this protocol, the bacterial cells were resuspended in Qiagen Buffer P1, lysed using Buffer P2, and the reaction was neutralized using Buffer N3. This mixture was spun for 10 minutes in a 5415 C tabletop centrifuge (Eppendorf) at 14,000 RPM. The supernatant was then decanted into a Qiaprep spin column, centrifuged and the flow-through was discarded. The DNA, bound to the membrane of the QIAprep spin column, was subsequently washed with Buffer PB, and Buffer PE, with centrifugation and removal of flow-through after each step. The Qiaprep spin column was centrifuged a final time to remove trace amounts of buffer. Plasmids were eluted from the QIAprep spin column into 50 µl DEPC treated H_2O, in a 1.5 ml snap-cap tube.

After extraction, the amount of plasmid DNA (pDNA) was quantified using a DU 640 Spectrophotometer (Beckman). Plasmid insert specificity was confirmed by PCR, using the oligonucleotide primers designed for the IPCR, and 0.5 μg of the plasmid DNA (pDNA). Reaction conditions were the same as for the IPCR, with the exception of the extension time, which was decreased for inserts below 1000-nt in length to 1.5 minutes. PCR amplification was confirmed by gel electrophoresis in 1.5% agarose gel, stained with ethidium bromide. Amplicon insertion into the vector was also tested using *Eco*R I digestion. Because the sequence of the amplicon is unknown, it is necessary to digest the PCR product as well as the pDNA, in order to interpret the digestion pattern of the pDNA as the same as the original amplicon. 0.5 μg of pDNA or IPCR amplicon was digested using 1U of *Eco*R I (New England Biolabs). One microliter of the *Eco*R I specific10x buffer was added to the solution and DEPC treated H_2O was added to 10 μl. Five microliters of the *Eco*R I digested pDNA solution was run on 1.5% agarose gel adjacent to the digested IPCR amplicon and stained with ethidium bromide. pDNA that were confirmed to have an insert of identical size to the original amplicon through PCR and confirmed to have an identical pattern of *Eco*R I digestion to the original amplicon were selected and the pDNA was sent for sequencing. pDNA was stored at -30°C for further use.

2.4 Sequencing and Sequence Analysis of GTHV DNA

Amplicons were sequenced two times in both directions using an Applied Biosystems 377XL DNA automated sequencer. The ABI Prism BigDye Terminator and Primer Cycle Sequencing Chemistries were used for the pre-reaction. Sequencing was done using M13 forward and reverse primers (Invitrogen) due to the binding site for these

primers on the pCR 2.1 vector (Invitrogen). Color electropherograms were created from the gel image and used in sequence analysis. The electropherograms were analyzed and each nucleotide peak was accounted for.

Sequences were further analyzed using the EditSeq and MegAlign programs of the DNAstar package. Using the EditSeq program, the IPCR priming regions were identified, and the plasmid sequences were removed from the amplicon sequence. The site for restriction enzyme digestion and ligation were identified and the overlapping DNA sequences were confirmed and then removed from the sequence, leaving a linear fragment of GTHV DNA. The MegAlign program was used to align the replicates of the sequence to ensure that the reading was consistent throughout all of the sequencing attempts. Sequencing in the manner described provides accurate sequence data on up to 800 nt of DNA. For the longer amplicons it was necessary to create internal sequencing primers in order to ensure the accuracy of the data. Internal sequencing primers were developed until the sequence could be read accurately in both directions. Internal sequencing primers are listed in Table 2.3.

The Clustal X program was used in order to compare the sequences of GTHV genes obtained in this study to the same genes in other herpesviruses. Clustal X was used to create both amino acid and nucleotide sequence alignments of the DNA *pol*, DBP, UL28, and gB genes of GTHV to the homologous genes in other herpesviruses. The nucleotide sequence alignments created in Clustal X for each GTHV gene were used to generate neighbor-joining phylogentic trees. Percent frequencies of the groupings were determined after 1000 bootstrap evaluations and the tree was viewed using the TreeView program (Page 1996). The amino acid alignment produced by the Clustal X program was

used to generate sequence distances between each GTHV gene and the homologous gene of other herpesviruses using the MegAlign program of the DNAstar package.

Additionally for amino acid sequence analysis, the Protean program from the DNAstar package was used to create a hydrophilicity plot according to the Kyte-Doolittle method (1982). The hydropathy window was 9 amino acids. For additional sequence analysis, the Genetics Computer Group (GCG) Wisconsin Package was used, including the programs: PepPlot, Gap, SPScan, and BestFit.

Table 2.3. Internal sequencing oligonucleotide primers. These oligonucleotide primers were designed for sequencing of large amplicons produced during IPCR. The primers designed for each gene are separated by double lines, and each walk is separated by color groupings.

INTERNAL SEQUENCING PRIMERS					
Gene	Walk	Name	Sequence	Direction	Amplicon Size
UL29	2	UL29 seq 2.1	5' TAC TGG GGC AGA ACC CAC AAG CGT A 3'	Forward	1,554 nt
UL29	4	UL29 seq 4.1	5' GGT AAC CGC TCA ATT GCC TT 3'	Forward	2,155 nt
UL29	4	UL29 seq 4.2	5' TTC GGA AAA GGT CGC TGT CA 3'	Reverse	2,155 nt
UL29	4	UL29 seq 4.3	5' CCT TCA TAG GCG CCC GTA AAG A 3'	Forward	2,155 nt
UL 28	1	UL28 seq 1.1	5' AGA ACG GAA AAC GGG TGT AG 3'	Forward	2,127 nt
UL 28	1	UL28 seq 1.2	5' TGT CGC AAA AGG GCG TTT CGG A 3'	Reverse	2,127 nt
UL 27	1	UL27 seq 1.1	5' CGT CAC CAT GTG GGA CAA CAA CGA CAA T 3'	Forward	3,219 nt
UL 27	1	UL27 seq 1.2	5' GGA CTG AAG GCT CAG ATC TCC AGG AAA T 3'	Reverse	3,219 nt
UL 27	1	UL27 seq 1.3	5' CAA GAT CTC ACT TCC ACT CTT GTG GTG C 3'	Forward	3,219 nt
UL 27	1	UL27 seq 1.2	5' TGG ATA TCG TAG TCG GGT AGA GGG GAA A 3'	Reverse	3,219 nt

CHAPTER 3: RESULTS

3.1 DNA Extraction from Tumor Tissue and PCR Detection of GTHV

Approximately 1 g of tumor tissue from the tongues of two green sea turtles were minced, digested with protinase K, and DNA was extracted from the tissue. By spectrophotometric analysis, 1 g of tumor A17-14a yielded 3.6 mg of DNA. One gram of tumor A18-14e yielded 3.06 mg of DNA. The DNA was resuspended in DEPC treated H_2O to a final concentration of 1 μg/μl. PCR was performed in order to ensure that the tumor tissue was positive for GTHV sequences. PCR, visualized by gel electrophoresis, stained with ethidium bromide, showed that both DNA samples were positive for the GTHV DNA *pol* gene. The positive control used, a previously extracted DNA sample from turtle tumor tissue (A15-14e), was run concurrently to A17-14a and A18-15e. This sample was positive as well for GTHV DNA *pol* gene sequences. The negative controls, A17-14e without *Taq* polymerase, and DEPC treated H_2O in place of GTHV DNA were both negative for GTHV DNA *pol* sequences. The agarose gel from this PCR is shown in Figure 2.1.

3.2 Genomic Walking from the UL29, UL28 and UL27 Genes

3.2.1 Genomic Walking Results from the UL29 Gene

To clone and identify more genomic sequences of the UL29 gene, directly upstream (coding strand) of the DNA *pol* gene, a genomic walking technique was employed, which is based on restriction enzyme digestion, self-ligation, IPCR, cloning and sequencing. Four successive genomic walks were performed using the different REs including *Hae* II and *Ngo*M IV, *Eco*R I, *Btg* I, and *Eco*R I and we obtained a total of 4,260 bp of genomic DNA, including the overlapping priming region (563 bp) and 3,697

bp of newly identified sequence. The first walk using *Hae* II enabled us to obtain an amplicon of 314 bp (Figure 3.1), of which a 56-bp portion was newly obtained. Using *Ngo*M IV, a fragment of 655 bp was amplified, which provides 92 bp of new sequence, in which the first 56 bp were identical to the sequence newly identified through the *Hae* II walk. For the second walk, oligonucleotide primers were designed from the newly obtained sequence since selected RE, *Eco*R I, cleaved the DNA at the beginning of the 92-bp sequence. This time IPCR resulted in a 1,554-bp amplicon, of which 1,456 bp were newly acquired (Figure 3.1). An internal sequencing oligonucleotide primer was designed to obtain accurate sequence data for the internal portion of amplicon (Table 2.2). For the third genomic walk, primers were developed at the upstream portion of the newly acquired sequence, and the restriction enzymes *Btg* I, *Ava* I, *Sac* II, and *Nsp* I, were tested and evaluated to digest the tumor DNA. Of these four enzymes, only the use of *Btg* I resulted in a fragment suitable for cloning. The amplicon was 598 bp in length containing 381 bp of new sequence (Figure 3.1). For the fourth walk, the circularized DNA prepared for the second walk using *Eco*R I, was used with primers developed from the upstream portion of the third walk. This produced an amplicon of 2,155 bp in length with 1,768 bp of new sequence (Figure 3.1). Three internal sequencing primers were designed for this portion of sequence (Table 2.2). Through four successive genomic walks, a total of 4,260 bp of GTHV genomic DNA was sequenced, which contains the entire open reading frame (ORF) of the DBP gene, and a 282-bp fragment of the proposed UL28 gene (upstream).

3.2.2 Genomic Walking Results from the UL28 gene

Two successive genomic walks were employed in order to amplify a fragment of 2,485 bp. This newly identified viral sequence includes the entire ORF for the GTHV UL28 gene. *Ava* I and *Btg* I were the RE's used to digest DNA for the first genomic walk; however, *Btg* I did not produce a band suitable for cloning after IPCR. Digestion using *Ava* I and subsequent IPCR amplification resulted in a fragment of 2,127 bp in length (Figure 3.2). This included 1,852 bp of newly obtained sequence information. For this walk, two internal sequencing primers were designed in order to confirm the accuracy of the newly identified sequence from this large amplicon (Table 2.2). In the second genomic walk DNA was digested with *Afl* III, *Apo* I, *Hae* II, and *Hinf* I. Only the *Afl* III-digested DNA resulted in an amplicon, which was 695 bp in length, including 317 bp of newly acquired sequence information (Figure 3.2). Through two genomic walks, 2,169 bp of new sequence data was obtained, which completed the sequencing of the entire UL28 transport gene, as well as197 bp of the UL27 gene.

3.2.3 Genomic Walking Results from the UL27 Gene

One genomic walk was employed to amplify a fragment of 3,219 bp. This sequence includes the entire ORF of the UL27 gB gene. Using GTHV DNA cut with *Nhe* I, an amplicon of 3,219 bp was obtained (Figure 3.3), which included 2,603 bp of new sequence information. Four internal sequencing primers were developed due to the length of the fragment that was cloned (Figure 2.3). This ensures that the obtained sequence data are accurate. Figure 3.4 shows the entire nucleotide sequence and putative amino acid translation of the UL29, UL28, and UL27 genes.

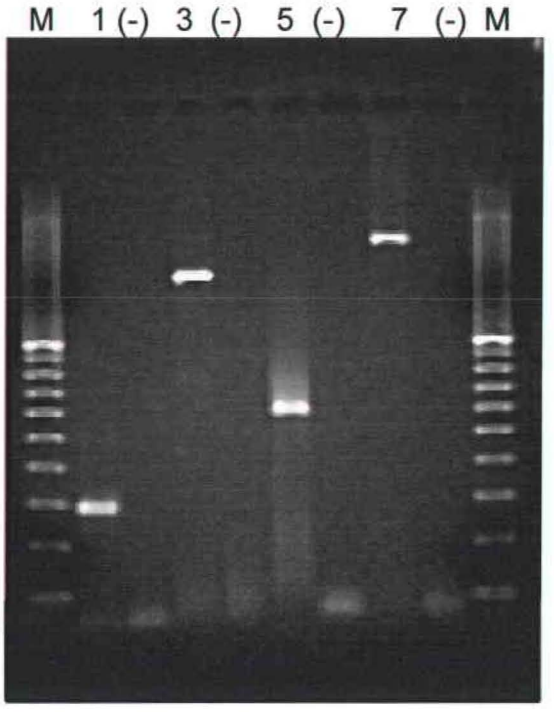


Figure 3.1. IPCR products from the four genomic walks of the UL29 gene. Lane 1 shows the first walk containing an IPCR amplicon from DNA cut with *Hae* II (314 bp). Lane 3 shows the second walk containing an amplicon from DNA cut with *EcoR* I (1,554 bp). Lane 5 shows the third walk containing an amplicon from DNA cut with *Btg* I (598 bp). Lane 7 shows the fourth walk containing an amplicon from DNA cut with *EcoR* I (2,155 bp). The intermittent lanes show the negative control (DEPC H_20) for each IPCR reaction. The IPCR products were subjected to 1.5% agarose gel electrophoresis and stained with ethidium bromide. M lanes show a 100 bp PCR ladder (Promega).

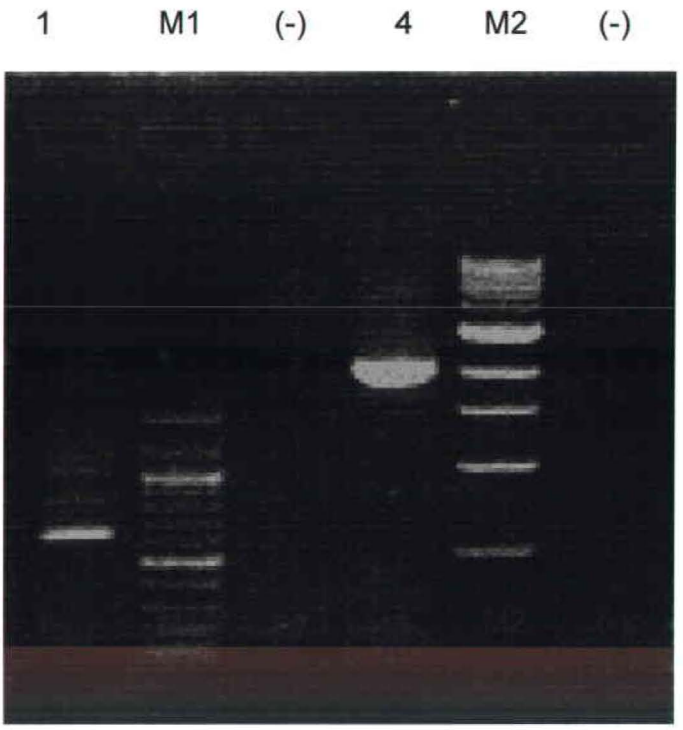


Figure 3.2. IPCR products from the two genomic walks of the UL28 gene. Lane 1 shows the second walk containing an IPCR amplicon from DNA cut with *Afl* III (695 bp). Lane 4 shows the first walk containing an amplicon from DNA cut with *Ava* I (2,127 bp). The (-) lanes show the negative control (DEPC H_2O) for each IPCR reaction. The IPCR products were subjected to 1% agarose gel electrophoresis and stained with ethidium bromide. The M1 lane shows a 100 bp ladder (New England Biolabs). The M2 lane shows a 1 kb ladder (New England Biolabs).

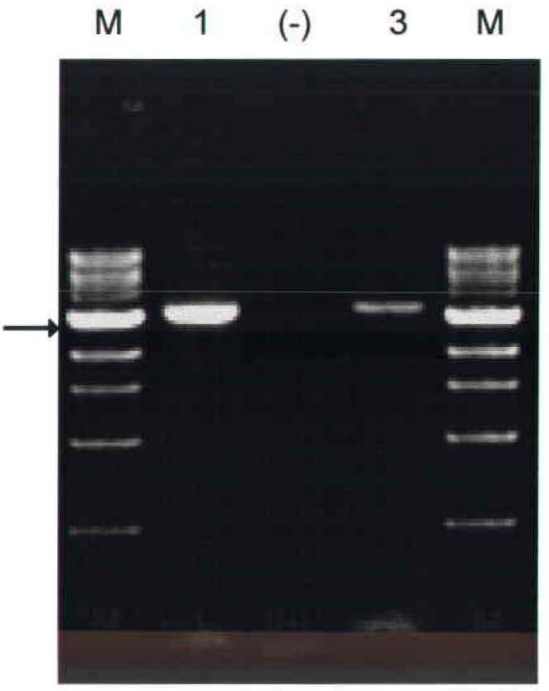


Figure 3.3. IPCR Product from the genomic walk of the UL27 gene. Lane 1 shows the IPCR amplicon from DNA cut with *Nhe* I (3,219 bp) (annealing temperature 65°C). Lane 3 shows the IPCR amplicon from DNA cut with *Nhe* I (3,219 bp) (annealing temperature 67°C). The (-) lane shows the negative control (DEPC H_2O) for the IPCR reaction. The IPCR products were run subjected to 0.8% agarose gel electrophoresis and stained with ethidium bromide. M lanes show a 1 kb DNA ladder (New England Biolabs). The arrow indicates 3000 bp.

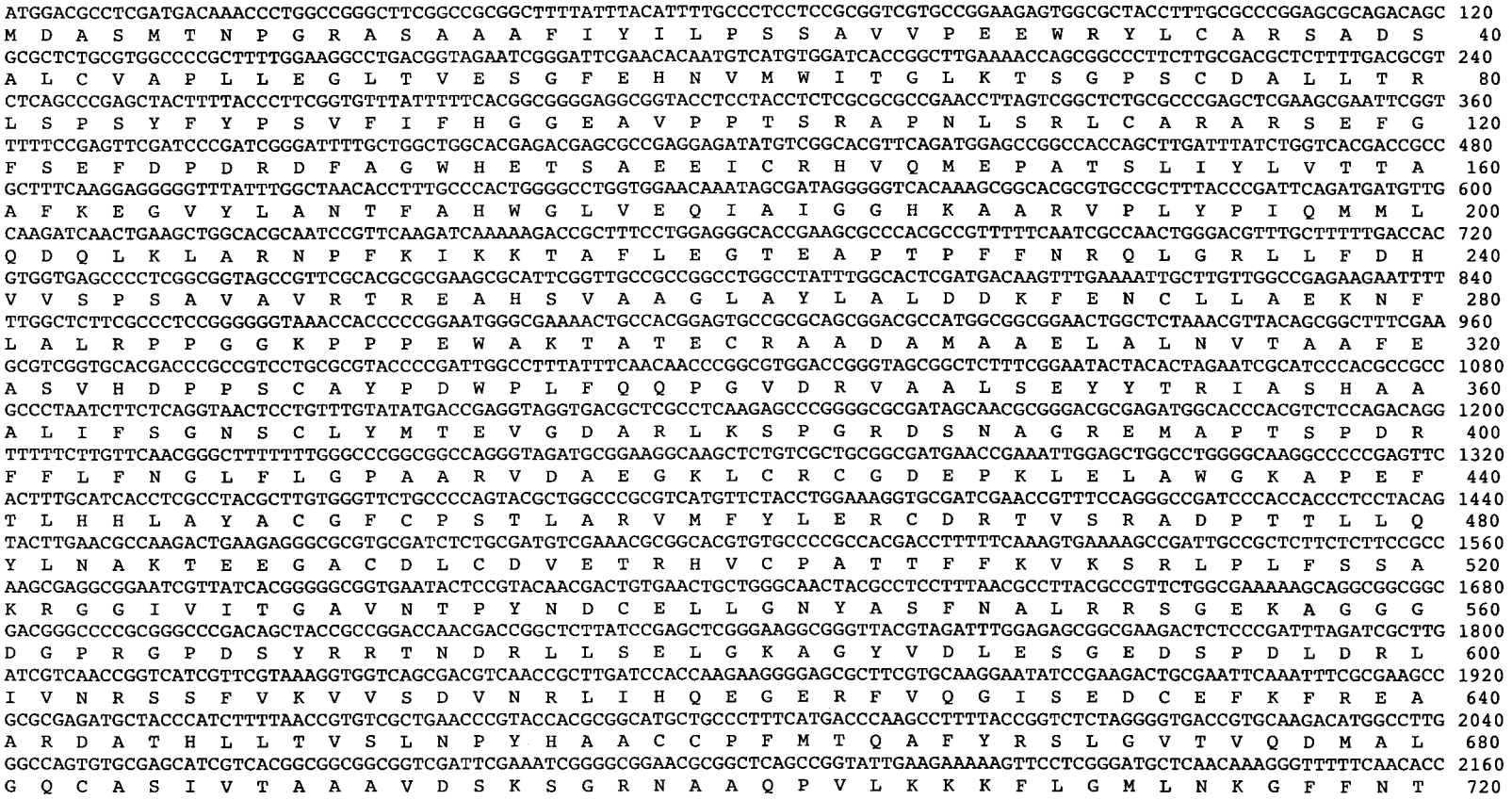


Figure 3.4. Nucleotide sequence and amino acid translation of the UL29, UL28, and UL27 genes. The sequence begins with the UL29 gene. (*) Indicates the stop codon for the UL29 nucleotide sequence. (+) Indicates the start codon for the UL28 nucleotide sequence and (**) indicates the stop codon. (++) Indicates the start codon for the UL27 gene and (***) indicates the stop codon. (+++) Indicates the end of the sequence obtained in this study.

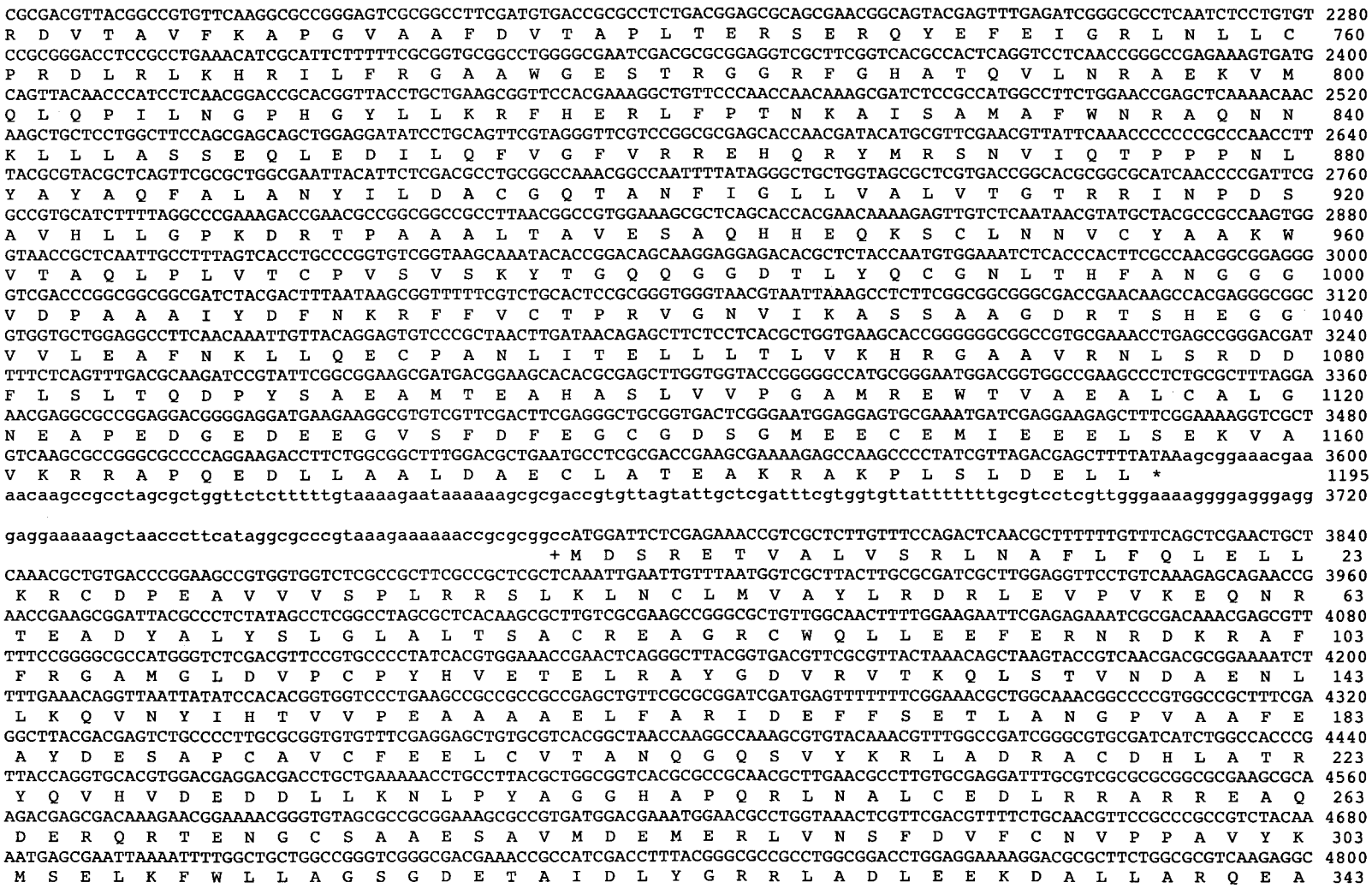


Figure 3.4. (Continued) Nucleotide sequence and amino acid translation of the UL29, UL28, and UL27 genes.

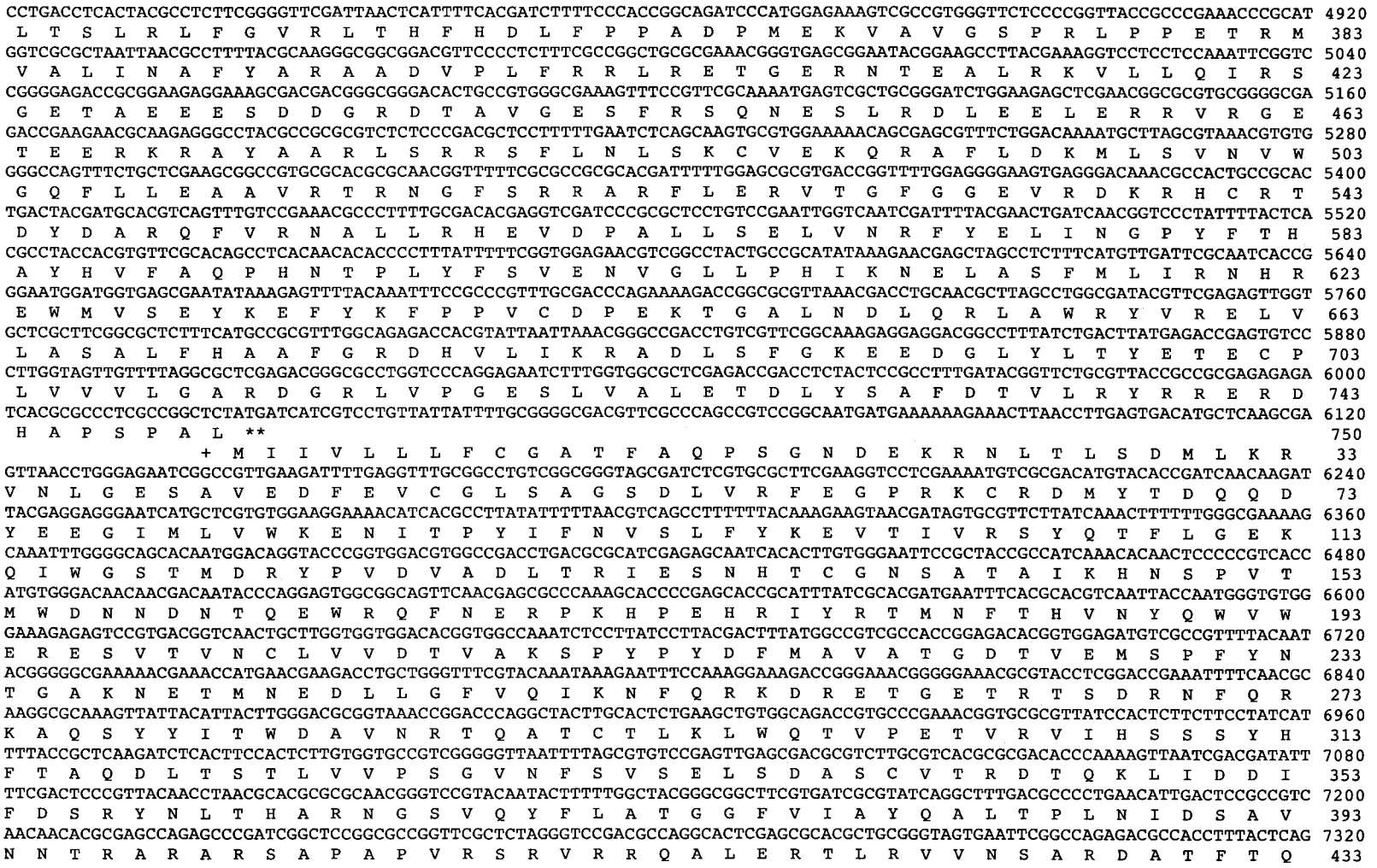


Figure 3.4. (Continued) Nucleotide sequence and amino acid translation of the UL29, UL28, and UL27 genes.

```
GTGCAATTCACCTACGACGTTCTCCGGCAACACATTAACTCGGTCTTCAGCCGCTTGGTCAACGCCTGGTGCGAGCTTCAGAACCGCGACGTAACGGTATGGGAGCTGCTGCTGAACATC 7440
V  Q  F  T  Y  D  V  L  R  Q  H  I  N  S  V  F  S  R  L  V  N  A  W  C  E  L  Q  N  R  D  V  T  V  W  E  L  L  L  N  I   473
AAAGCCTCGTCCGTGATGAGCACGGTTATGGGAAAACGGGTCGCGGCCAAAAAAATGGGGGATTTGGTGGCGGTCAGCAAATGTCTCAAGGTCAACTCCGCCAACGTGAAACTGTGGAAT 7560
K  A  S  S  V  M  S  T  V  M  G  K  R  V  A  A  K  K  M  G  D  L  V  A  V  S  K  C  L  K  V  N  S  A  N  V  K  L  W  N   513
ACGATGTATATCACCGACGGGCCTTACGCGGGCCGGGGCTTGTGTTGGTCCAGGCCGGTAGTCGTCTTTTCTCTTTCCGACAACGCAAATCAGGAAGATTTGATCACCGGTCACTTGGGA 7680
T  M  Y  I  T  D  G  P  Y  A  G  R  G  L  C  W  S  R  P  V  V  V  F  S  L  S  D  N  A  N  Q  E  D  L  I  T  G  H  L  G   553
GAACACAACGAGCTGCTTCCCTACTTGGGGTTGACGGAGAGTTGCGAGCAAAACTCCCGCAAGTATTTTATGTTCGGCGGGGTTTACATGTTGTACGAAAATTATCATTTCGTCCGCCCG 7800
E  H  N  E  L  L  P  Y  L  G  L  T  E  S  C  E  Q  N  S  R  K  Y  F  M  F  G  G  V  Y  M  L  Y  E  N  Y  H  F  V  R  P   593
GTGACTTTGGACGAGATTACGGAAGTGCGCACTTTTGTAGAGTTTAACGTTTCCCCTCTACCCGACTACGATATCCAGCACCTAGAGGTCTATTCCCAGGAGGAAATCAAGCTCTCCAAT 7920
V  T  L  D  E  I  T  E  V  R  T  F  V  E  F  N  V  S  P  L  P  D  Y  D  I  Q  H  L  E  V  Y  S  Q  E  E  I  K  L  S  N   633
TCGGCAGATTACCTGGAAGTGACGCGCTTTCAAAACGCTTACTTGCAACTGGTCAACGATCTGAAACTGGTCGTGTACGGCGACGCCAGCCGCGCCTTTATTGCCCGCGTTCAGGCCTTT 8040
S  A  D  Y  L  E  V  T  R  F  Q  N  A  Y  L  Q  L  V  N  D  L  K  L  V  V  Y  G  D  A  S  R  A  F  I  A  R  V  Q  A  F   673
TTCGGAAGCCTCGGTACCGTGGGCGAAGCGCTGGGAAAGGTCATCGGTGTAGTAGGTGGGGCGCTGGCCGGCGTGCTCAATGGGATCGTTGGTTTTTTGACCAATCCTTTCGGGGGCTTT 8160
F  G  S  L  G  T  V  G  E  A  L  G  K  V  I  G  V  V  G  G  A  L  A  G  V  L  N  G  I  V  G  F  L  T  N  P  F  G  G  F   713
ACCATTTTGCTGCTAGTCGGAGGCGGTCTGATCGCCGCTTTCCTGGCACTGAAATTCGTCCAGACGTTTAAAAAGAACCCCCTTTCGACCATTTTCCCGGTCAGTGCCAACCAAGGGCAA 8280
T  I  L  L  L  V  G  G  G  L  I  A  A  F  L  A  L  K  F  V  Q  T  F  K  K  N  P  L  S  T  I  F  P  V  S  A  N  Q  G  Q   753
TGGGATCCCAAAGAGCCCCTGCCGGAGCTCAGCAAAAAACGGCAGCAAGAAATCAACGATTTCCTGCGAGATCTGAGCCTTCAGTCCGAAGCGGAAAGACGTGAGAAAAAAGCGGCGAGT 8400
W  D  P  K  E  P  L  P  E  L  S  K  K  R  Q  Q  E  I  N  D  F  L  R  D  L  S  L  Q  S  E  A  E  R  R  E  K  K  A  A  S   793
TCCTCCGGATGGTCGCTGGCCAAAGAAAAACTATCTAACGTCAAAGACAACTTTACGGACGCTAAAAATAAAATGATGAAATCTTTACGCCACCGCGGAAACGGATACGCTCCCGTTTCG 8520
S  S  G  W  S  L  A  K  E  K  L  S  N  V  K  D  N  F  T  D  A  K  N  K  M  M  K  S  L  R  H  R  G  N  G  Y  A  P  V  S   833
ACGAGCACGGCCAGTATGTATCTAGACGACGAATTATCGACGTATTCTCAGGTCTGAgggaccgcgcggttctttcttgttaaaaaaagaccaaaagacttttgacgcctatagcactgg 8640
T  S  T  A  S  M  Y  L  D  D  E  L  S  T  Y  S  Q  V  ***                                                                  851
aaaaaaaacgactcagtcagccgtgttttttatgtatcgatttatttattcaagaaataaaggtacggaaatcaaagtacactttatgatttcgttattgcattcgagcccatcatcgc 8760
cgttataaacttctgctttagctgacggccccggggattcttaccgacggatggagccgcgctagcc +++                                                    8827
```

Figure 3.4. (Continued) Nucleotide sequence and amino acid translation of the UL29, UL28, and UL27 genes.

3.3 Sequence Analysis of GTHV Genes UL29, UL28 and UL27

3.3.1 Sequence Analysis of the GTHV UL29 Gene

The GTHV DBP gene has an ORF of 3,585 bp, and the entire length of this sequence has been deposited into GeneBank with the accession number: AY332227. This viral gene is located upstream (coding strand) of the DNA *pol* gene and oriented in the 3' to 5' direction. The two genes are separated by an intron of 160 bp (Figure 3.4). The DBP gene is 59.19% G+C rich. This sequence is predicted to encode a 1,195 amino acid peptide, with an estimated molecular weight of approximately 130.811 kilodaltons (kDa). The amino acid sequence of the putative GTHV DBP gene was determined by a Peptide Blast Library Search Program to contain a conserved domain for the herpesviral single stranded DNA binding protein gene.

Analysis of the nucleotide and amino acid sequences of this newly determined GTHV DBP gene revealed several distinctive characteristics common to other herpesviruses. Transcriptional elements seen in the noncoding regions of this gene include a putative TATA box starting at 53 bp upstream from the start codon of the DBP gene and oriented in the 3'to 5' direction. Two putative startpoints of transcription were identified at 24 and 27 bp upstream from the start codon, respectively. A putative CAAT box was identified at 143 bp from the UL29 start codon. At the downstream end of the gene, a putative polyadenylation signal is situated at 49 bp downstream from the stop codon of the UL29 gene. This polyadenylation sequence is conserved in DPB genes of alphaherpesviruses (Wu *et al.* 1998).

In addition, two other conserved amino acid sequences with putative motifs were observed. The first one is a putative zinc-binding motif with the consensus sequence

$CXLCX_4RX_2C$ (Wang and Hall 1990). This sequence is located between 491 and 504 amino acids and it is highly conserved in alphaherpesviruses (White and Boehmer 1999) (Figure 3.5). The second sequence is a putative DNA binding domain (Figure 3.6). This sequence is characterized by a set of basic and aromatic amino acids (Wang and Hall 1990) and identified between amino acids 627-839 of the newly acquired GTHV DBP sequence.

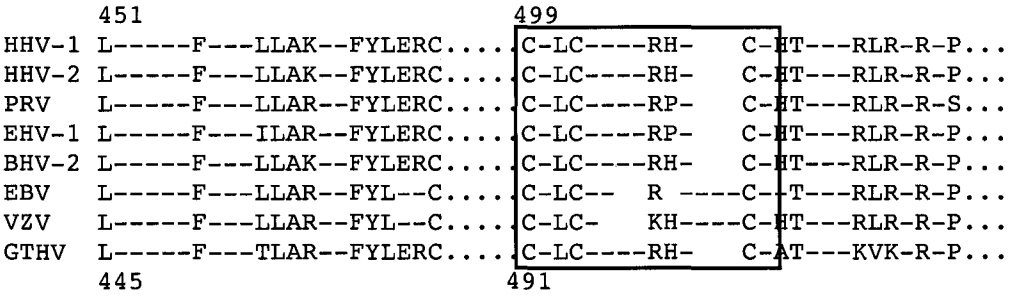


Figure 3.5.Putative zinc finger domain of GTHV DBP and other herpesviruses. The boxed area represents the putative zinc-finger domain.

```
      543
HHV-1 Y-D-D-LGNY--F--- KR-     D-----R----E--R...
HHV-2 Y-D-D-LGNY--F--- KR-     D-----R----E--R...
PRV   Y-D-D-LGNY--Y--- KR-     N-----K----D--R...
EHV-1 Y-D-D-LGNY--Y--- KR-     N-----K----D--R...
BHV-2 Y-D-D-LGNY--Y--- RR-     G-----R----E--R...
EBV   Y-E-D-LGNF--F--  K--     D--   R--- E--R...
VZV   Y-D-D-LGNY--Y----K--     D-----R----D--R...
GTHV  Y-D-E-LGNY--F--- RR-----G-----R----D--R...
      535

      650                                        762
HHV-1 A-H-M-----PY----CP----L--R--L---QD--L--C...RV-----K--R-K-RV-F-G
HHV-2 A-H-M-----PY----CP----L--R--L---QD--L--C...RV-----K--R-K-RV-F-G
PRV   A-H-M-----PY----CP----L--R--L---QD--L--C...RV-----R--R-K-RV-F-T
EHV-1 A-H-M-----PY----CP----L--R--F---QD--L--C...RV-----R--R-K-RV-F-N
BHV-2 A-H-M-----PH----CP----L--R--L---QD--L--C...RV-----K--R-K-RV-F-A
EBV   A-H-M-----PF----CP----L--R--L---QD--L--C...RV-----K--K-K-RV-F-G
VZV   A-H-M-----PY----CP----L--R--L---QD--L--C...RV-----R--R-K-RV-F-G
GTHV  A-H-L-----PY----CP----A---R-L---QD--L--C...RL-----R--R-K-RI-F-G
      644                                        755

      797
HHV-1 L---Y- K--- R-D---  G--- FL-K-FH---FP--K---  S------WFW---QR
HHV-2 L---Y- K--- R-D---  G--- FL-K-FH---FP--K---  S------WFW---QR
PRV   L---Y- R--- A-D---  G--- FL-K-HH---FP--K---  G------WFW---QR
EHV-1 M---Y- R----S N---  G--- FL-K-YH---FP--H---  I------WFW---QR
BHV-2 L---Y- K----G-N---  G--- FM-K-FH---FP--K---  S------WFW---QR
EBV   L---F- K--- R-E-----G----FL-- YH---FP--K-    S----  W------R
VZV   L---Y- R--- R-D---  G--- FL---FH---FP  R-----S------W------R
GTHV  L-  R- K--- L-P---  G--- YL-K-FH---FP--K--   S--    AFW---QN
      793
```

Figure 3.6. Putative DNA binding site of GTHV DBP and other herpesviruses. Amino acid sequence of GTHV DBP and other closely related herpesviruses' DNA binding sites. Areas with highly conserved residues characterized by the presence of aromatic (F, W, Y) and basic (H, K, R) residues are shown below. (-) indicates non-conserved residues. The boxed area is the putative DBP binding sequence motif.

3.3.2 Sequence Analysis of the GTHV UL28 Gene

Analysis of the GTHV UL28 gene indicated it has an ORF of 2, 250 bp, and the entire sequence has been submitted to GeneBank (accession number: AY455977). This viral gene is located upstream (coding strand) of the DNA binding protein gene (UL29). It is oriented in the 3' to 5' direction, and the 3' end of the gene contains an out of frame overlapping (1bp) ORF, which is part of the 5' terminus of the UL27 gene encoding Glycoprotein B (gB). The GTHV UL28 gene, which is 57.24% G+C rich, encodes a peptide of 750 amino acids, with a calculated molecular weight of approximately 85.80 kDa. The GTHV UL28 was determined by a Peptide Blast Library Search to contain a conserved domain for the UL28 herpesvirus homolog.

Transcriptional elements present before the translational start codon for this gene include a potential CAAT element with the sequence 'ATTG', which reads 'CAAT' on the complimentary strand beginning 106 bp from the start codon of the UL28 gene. A potential transcriptional startpoint was identified 30 bp upstream from the translational start codon. No polyadenylation signal was identified after the stop codon of the UL28 gene. Alignment analysis of the amino acid sequence also indicated the presence of a potential nuclear localization signal (NLS) from amino acids 467-469 with the sequence 'RKR'. The potential NLS is well conserved among herpesviruses (Kato *et al.* 1999).

3.3.3 Sequence Analysis of the GTHV UL27 Gene

The GTHV gB gene (UL27) has an ORF of 2,553 bp. This viral gene is located directly upstream (coding strand) of the UL28 gene, overlapping by 1 nucleotide. UL27 is oriented in the 3' to 5' direction. (Figure 3.4). This gene sequence, which is 52.76% G+C rich, encodes a peptide that is 851 amino acids long and has a calculated molecular

weight of 96.18 kDa. The GTHV gB gene was determined by a Peptide Blast Library Search to contain a conserved domain for the herpesviral gB peptide.

A putative TATA box was identified between nucleic acid residues 219 and 213, upstream from the ATG start codon of the GTHV UL27 gene, within the coding sequence of the UL28 gene. The presence and location of this feature is conserved in other herpesviruses including Gallid HV-2 (Kato *et al.* 1999). A putative CAT box was identified from 545 to 541 nucleic acid residues upstream from the ATG start codon of the GTHV UL27 gene. A putative initiation site of eukaryotic mRNA transcription was identified from 24 to 32 nt, and 27 nt to downstream from the TATA box (Breathnach and Chambon 1981). The presence of this transcription factor is conserved in other herpesviruses. A putative polyadenylation signal was found 121 nucleotides downstream from the stop codon of the UL27 gene.

Peptide analysis of the putative gB peptide was performed using the SPSscan program from the GCG's Wisconsin package. This search identifies putative signal sequence regions of a peptide. A potential signal sequence was identified using this program between amino acid residues 14 and 15 of the GTHV UL27 peptide. A hydrophilicty plot of the putative gB peptide of GTHV and the gB peptide of HHV-1 was created using the Protean program of the DNAstar package, in order to determine the hydropathic regions of the peptide. A 70 amino acid peptide fragment of The GTHV gB peptide was seen to be particularly hydrophobic from amino acids 661 to 730 and is hypothesized to be a transmembrane region of the UL27 peptide (Figure 3.7). Further analysis of the peptide sequence of gB is considered in the discussion.

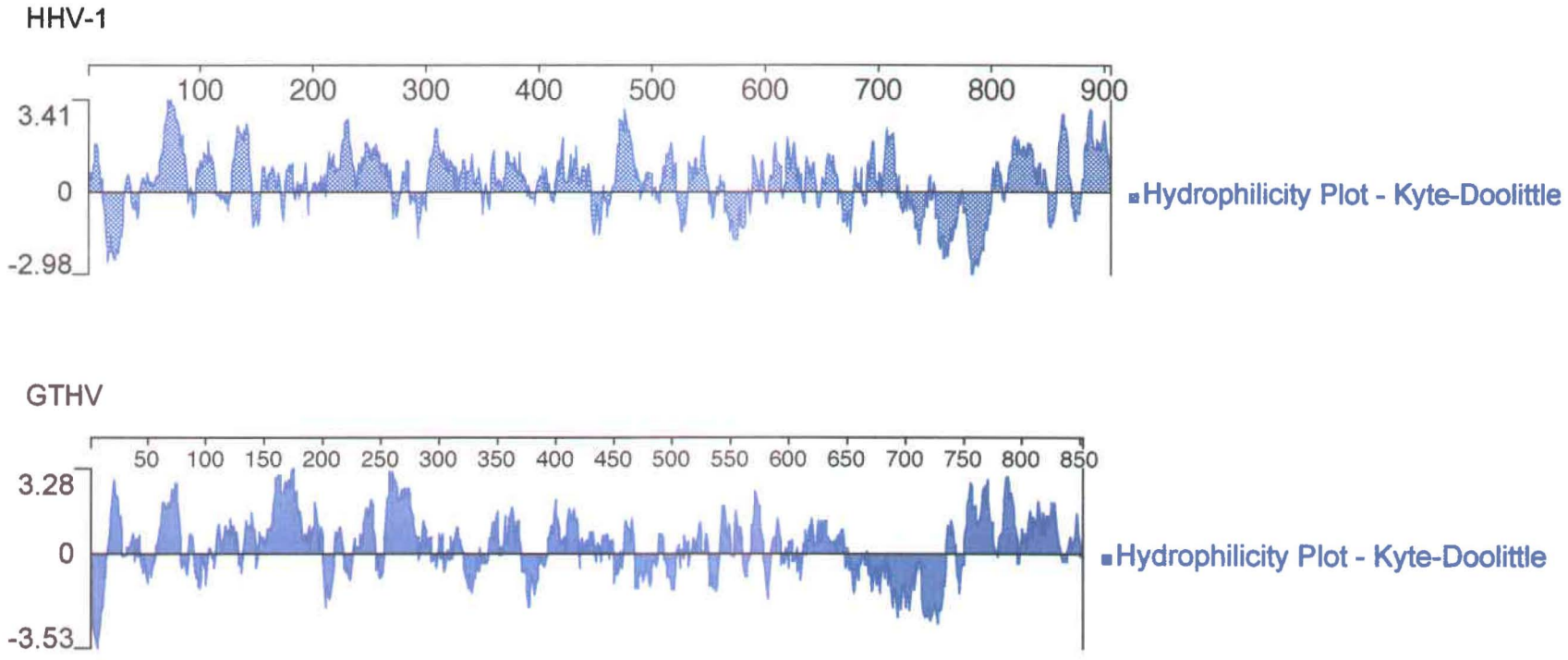


Figure 3.7. Hydrophilicity plot of HHV-1 and GTHV gB peptides. Plot was configured using the Protean program of the DNAstar package, using the Kyte-Doolittle method (1982). The points above the line indicate hydrophilicity and the points below the line indicate hydrophobicity.

3.4 Phylogenetic Analysis of the GTHV UL29, 28, and 27 Genes

3.4.1 Phylogenetic Analysis of the GTHV UL29 Gene

Through a Peptide Blast Library Search, the GTHV UL29 gene showed high sequence homology to other herpesviral DBP sequences, specifically human herpesvirus-2 (HHV-2), human herpesvirus-1 (HHV-1), bovine herpesvirus-1 (BHV-1), bovine herpesvirus-2 (BHV-2), pseudorabies virus (PRV), and equine herpesvirus-1 (EHV-1). A neighbor-joining phylogenetic tree was constructed using an alignment created in the Clustal X program comparing GTHV DBP gene to the nucleotide sequences of 20 closely related herpesviruses. The phylogenetic tree, which is displayed using the TreeView program (Figure 3.8), shows the GTHV DBP is closely related to members of the alphaherpesvirus subfamily, corroborating data reported in previous studies (Yu *et al.* 2001) based on the DNA *pol* gene.

The sequence distances of the amino acid sequence of the GTHV DBP compared to 20 closely related herpesviruses was determined in the Megalign program of the DNAstar package using an amino acid alignment created in the Clustal X program. The percent identity of the GTHV DBP to other herpesviruses is shown in Figure 3.9. The GTHV DBP is most similar to members of the alphaherpesvirus subfamily with a range of sequence identity from 36.2% to 33.2%. The most similar alphaherpesvirus is Human HV-2 (36.2%), followed by Human HV-1 (36.1%), and Bovine HV-1 (36.0 %). Members of the betaherpesvirus subfamily ranged in percent identity to GTHV DBP from 22.9% to 19.5%. Members of the gammaherpesvirus subfamily ranged in percent identity to GTHV DBP from 27.1% to 25.1%.

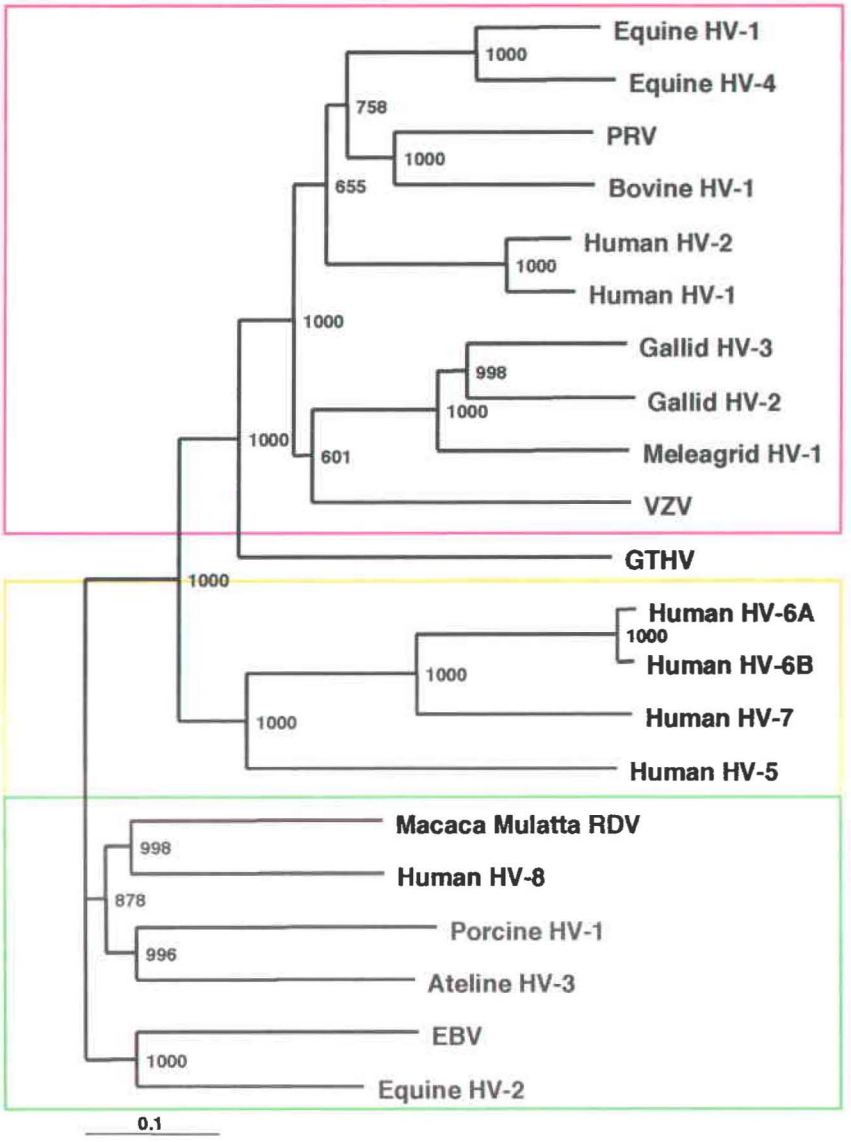


Figure 3.8. Neighbor- joining phylogenetic tree of the GTHV DBP gene and other herpesviruses. The tree is displayed using the TreeView program. The tree was bootstrapped 1000 times and the bootstrap value for the GTHV branch is 1000, indicating the analysis is reliable. All sequences, excluding GTHV, were obtained from the NCBI Pub-Med website using the accession numbers listed in Figure 3.9. The pink box surrounds members of the alphaherpesvirus subfamily. The yellow and green boxes surround betaherpesviruses, and gammaherpesviruses, respectively.

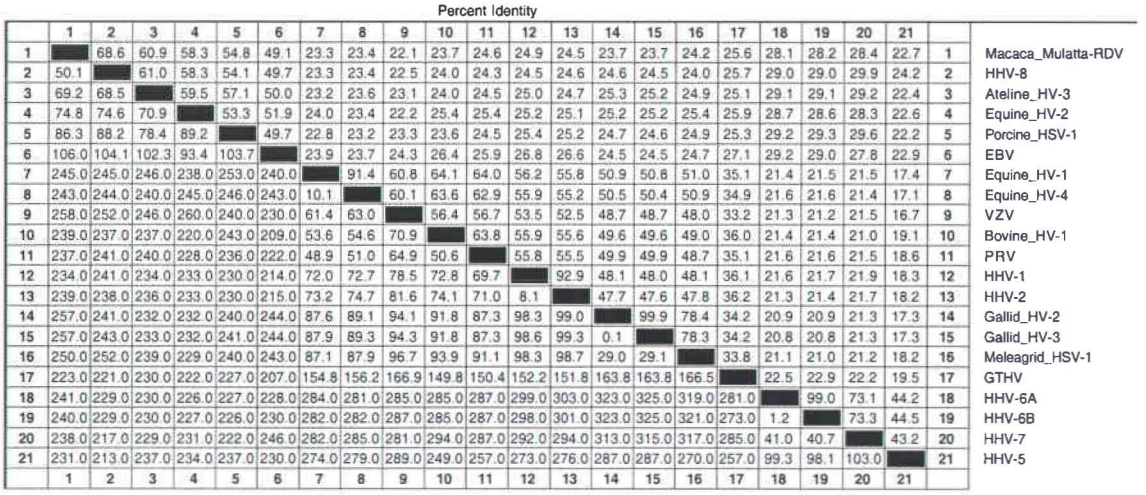


Figure 3.9. The amino acid sequence distances of the GTHV UL29 gene compared with 20 related herpesviruses. All sequences, excluding GTHV, were obtained from the NCBI Pubmed website using the following accession numbers: Ateline Herpesvirus-3: NC001987, Bovine Herpsvirus-1: Z78205, Epstein-Barr virus (EBV): AJ07799, Equine Herpesvirus-1: M86664, Equine Herpesvirus-2: NC001650, Equine Herpesvirus-4: NC001844, Gallid Herpesvirus-2: NC002229, Gallid Herpesvirus-3: NC002577, Human Herpesvirus-1 (HHV-1): NC001806, Human Herpesvirus-2 (HHV-2): D10658, Human Herpesvirus-5 (HHV-5): NC001347, Human Herpesvirus-6A (HHV-6A): NC001664, Human Herpesvirus-6B (HHV-6B): NC000898, Human Herpesvirus-7 (HHV-7): NC001716, Human Herpesvirus-8 (HHV-8): NC003409, Macaca Mulatta Rhadinovirus: NC003401, Meleagrid Herpesvirus-1: NC0026141, Pseudorabies Virus (PRV): L24487, Porcine Lymphotrophic Virus: AF478169, Varcella-Zoster Virus (VZV): X04370.

3.4.2 Phylogenetic Analysis of the GTHV UL28 Gene

Through a Peptide Blast Library Search, the GTHV UL28 gene showed high sequence homology to other herpesviral UL28 sequences, specifically pseudorabies virus (PRV), human herpesvirus-1 (HHV-1), human herpesvirus-2 (HHV-2), bovine herpesvirus-1 (BHV-1), bovine herpesvirus-2 (BHV-2), and equine herpesvirus-1 (EHV-1). A neighbor-joining phylogenetic tree was constructed using an alignment created in the Clustal X program comparing GTHV UL28 gene to the nucleotide sequences of 20 closely related herpesviruses. The phylogenetic tree, which is displayed using the TreeView program (Figure 3.10), shows the GTHV UL28 homolog is closely related to members of the alphaherpesvirus subfamily, corroborating data reported in previous studies (Yu *et al.* 2001) based on the DNA *pol* gene, as well as from the UL29 gene.

The sequence distances of the amino acid sequence of the GTHV UL28 protein compared to 20 closely related herpesviruses was determined in the Megalign program of the DNAstar package using an amino acid alignment created in the Clustal X program. The percent identity of the GTHV UL28 protein to other herpesviruses is shown in Figure 3.11. The GTHV UL28 homolog is most similar to members of the alphaherpesvirus subfamily with a range of sequence identity from 48.8% to 39.9%. The most similar alphaherpesvirus is PRV (48.8%), followed by Human HV-2 (45.7%), and Human HV-1 (45.5%). Members of the betaherpesvirus subfamily ranged in percent identity to GTHV UL28 homolog from 38.6% to 34.0%. Members of the gammaherpesvirus subfamily ranged in percent identity to GTHV Ul28 homolog from 36.6% to 25.5%.

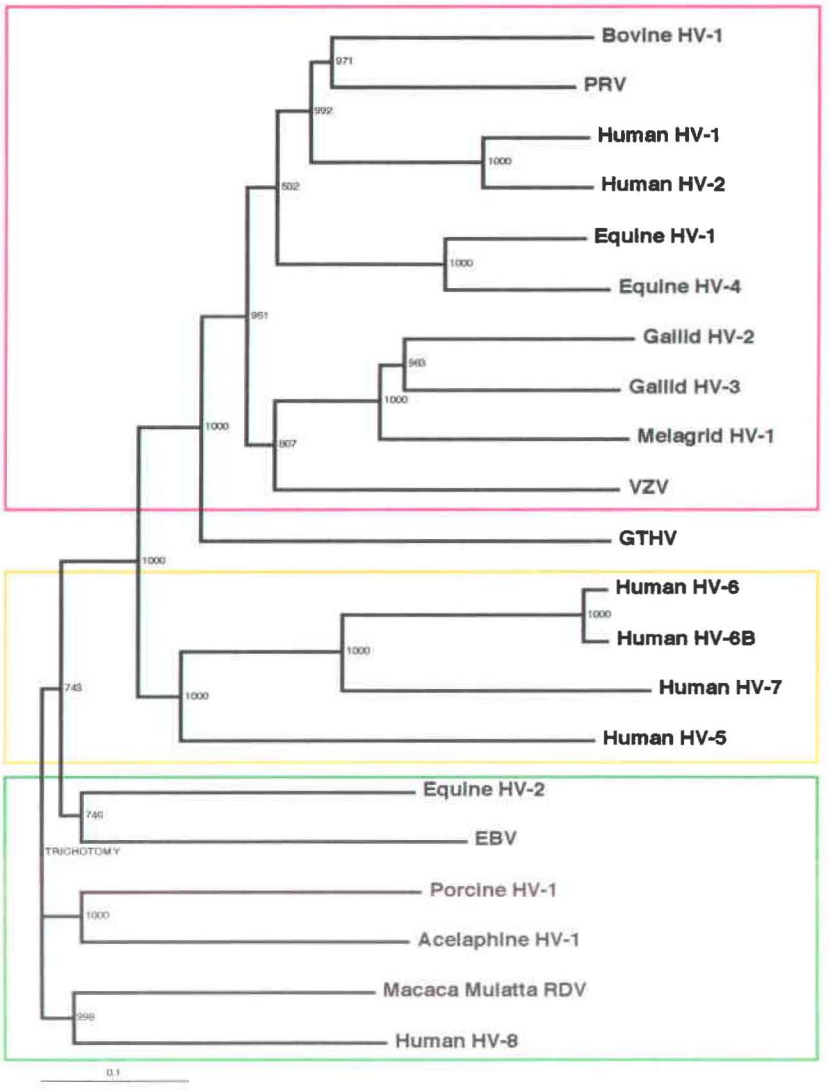


Figure 3.10. Neighbor- joining phylogenetic tree of the GTHV UL28 gene and other herpesviruses. The tree is displayed using the TreeView program. The tree was bootstrapped 1000 times and the bootstrap value for the GTHV branch is 1000, indicating the analysis is reliable. All sequences, excluding GTHV, were obtained from the NCBI Pub-Med website using the accession numbers listed in Figure 3.9. The pink box surrounds members of the alphaherpesvirus subfamily. The yellow and green boxes surround betaherpesviruses, and gammaherpesviruses, respectively.

Percent Identity

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
1		60.6	57.4	60.4	55.9	42.1	27.7	35.0	34.0	33.5	32.9	33.2	34.2	32.8	33.3	32.6	34.1	41.1	41.3	41.0	34.4	1	Ateline_HV-3
2	92.2		56.6	57.4	53.6	41.5	29.6	35.7	35.9	35.1	34.5	34.4	34.5	34.5	34.6	34.0	36.5	41.7	41.8	40.7	35.7	2	Porcine_HV-1
3	101.6	104.2		64.6	57.7	43.8	29.4	35.7	35.7	34.5	34.7	34.5	34.3	32.8	33.6	33.5	36.6	39.4	39.4	39.1	34.3	3	Human_HV-8
4	89.3	101.8	76.0		57.3	44.6	28.7	35.6	35.2	34.3	35.2	35.2	34.4	33.9	33.9	34.1	36.7	40.9	40.9	40.2	34.1	4	Macaca_Mulatta_RDV
5	93.0	104.3	87.5	88.5		42.4	31.9	37.5	35.1	34.2	36.1	36.3	35.0	33.2	35.0	35.0	36.2	37.1	37.2	36.3	32.9	5	Equine_HV-2
6	123.4	125.0	113.7	108.2	106.6		22.2	25.8	23.9	24.3	25.1	25.7	24.6	24.3	25.6	25.6	25.5	29.6	29.6	28.6	23.4	6	EBV
7	241.0	217.0	218.0	227.0	194.2	236.0		58.7	57.7	57.4	54.0	55.4	47.8	47.4	46.9	48.0	39.9	29.0	28.9	28.2	27.1	7	Bovine_HV-1
8	231.0	216.0	213.0	213.0	189.0	213.0	58.9		62.1	60.9	58.6	58.0	54.8	51.7	52.4	52.3	48.8	36.4	36.3	34.4	32.0	8	PRV
9	236.0	210.0	217.0	222.0	212.0	243.0	67.1	63.4		89.6	58.8	58.9	58.1	54.1	54.4	54.2	46.6	35.0	34.9	33.2	30.4	9	Equine_HV-1
10	228.0	209.0	223.0	223.0	211.0	240.0	69.2	64.4	12.9		58.2	58.7	56.5	54.9	55.0	54.8	45.1	34.2	34.2	32.4	29.1	10	Equine_HV-4
11	236.0	214.0	214.0	205.0	189.5	213.0	74.8	68.6	76.1	77.3		90.7	53.7	52.4	53.6	52.7	45.5	32.6	32.4	31.5	29.0	11	Human_HV-1
12	231.0	215.0	217.0	205.0	190.1	211.0	73.5	72.0	77.4	79.3	10.6		54.2	52.5	54.2	53.3	45.7	33.4	33.2	31.9	28.9	12	Human_HV-2
13	234.0	220.0	234.0	230.0	207.0	226.0	91.0	82.4	77.0	78.8	84.7	84.3		50.8	50.7	51.1	43.4	34.5	34.5	34.2	29.0	13	VZV
14	238.0	216.0	248.0	228.0	224.0	245.0	99.9	88.0	89.0	87.6	94.7	96.0	94.7		75.8	74.0	42.3	33.2	33.4	31.2	29.0	14	Gallid_HV-2
15	240.0	223.0	244.0	238.0	210.0	230.0	101.4	86.1	89.0	88.4	90.9	90.8	96.5	34.8		71.3	43.8	32.7	33.1	31.5	28.8	15	Gallid_HV-3
16	253.0	233.0	246.0	234.0	210.0	230.0	96.2	86.1	89.2	88.6	93.5	93.8	94.6	39.0	44.1		44.2	33.9	34.0	32.7	29.1	16	Melagrid_HV-1
17	258.0	222.0	221.0	218.0	216.0	227.0	132.3	114.3	125.1	127.8	123.4	123.9	137.3	141.4	132.5	133.4		38.4	38.6	37.0	34.0	17	GTHV
18	197.3	191.3	218.0	200.0	221.0	211.0	228.0	207.0	213.0	215.0	223.0	211.0	216.0	213.0	219.0	212.0	197.0		99.1	69.0	48.9	18	Human_HV-6
19	194.4	189.9	218.0	200.0	220.0	211.0	229.0	208.0	215.0	215.0	227.0	214.0	216.0	210.0	213.0	210.0	194.3	1.2		69.0	49.1	19	Human_HV-6B
20	196.0	203.0	222.0	210.0	234.0	227.0	240.0	240.0	245.0	245.0	244.0	237.0	222.0	246.0	243.0	232.0	212.0	63.0	63.0		46.5	20	Human_HV-7
21	195.0	189.2	203.0	203.0	195.0	209.0	220.0	203.0	231.0	237.0	248.0	238.0	221.0	226.0	232.0	228.0	186.3	94.4	93.5	104.9		21	Human_HV-5
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		

Figure 3.11. The amino acid sequence distances of the GTHV UL28 gene compared with 20 related herpesviruses. All sequences, excluding GTHV, were obtained from the NCBI Pubmed website using the accession numbers listed in Figure 3.9.

Percent Identity

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22		
1		64.5	54.9	53.9	54.4	51.5	52.0	50.8	49.9	50.6	31.2	35.0	27.7	27.6	27.5	26.5	26.8	28.6	26.7	26.4	27.3	26.2	1	Bovine_HV-1
2	50.7		53.9	53.5	53.7	53.9	54.9	50.4	50.1	50.8	32.4	36.2	29.0	30.1	28.7	28.5	28.3	29.8	26.2	26.2	27.2	28.3	2	PRV
3	63.8	66.1		48.7	48.6	49.8	51.6	54.9	54.6	54.9	34.6	39.2	30.2	30.4	30.6	28.5	28.6	32.0	30.5	30.6	30.9	28.8	3	VZV
4	72.9	68.2	70.7		89.9	46.5	47.1	47.6	46.9	47.3	28.1	34.1	23.5	25.0	23.6	22.0	24.7	24.4	21.7	21.8	22.9	22.6	4	Equine_HV-1
5	71.7	68.1	71.6	11.2		46.4	46.9	47.2	46.6	47.0	28.3	34.1	23.8	24.6	24.0	22.3	24.9	24.5	22.2	22.2	23.6	22.8	5	Equine_HV-4
6	84.3	77.1	87.3	87.7	87.7		87.9	49.6	49.7	49.6	33.8	37.0	29.7	30.5	29.3	27.9	29.3	30.8	27.1	26.9	28.2	27.9	6	Human_HV-1
7	82.3	73.8	84.5	85.0	86.1	13.8		51.4	51.4	51.1	34.1	36.8	30.4	31.4	28.9	28.2	29.8	31.1	27.6	27.4	28.2	28.6	7	Human_HV-2
8	81.3	78.7	81.5	80.2	82.5	84.5	82.1		85.9	84.4	39.0	43.6	34.2	34.1	35.1	31.9	31.7	34.9	33.2	33.1	32.7	30.1	8	Gallid_HV-2
9	84.5	79.7	82.8	82.7	84.8	84.7	81.8	19.8		80.3	39.3	43.4	34.3	34.1	35.0	31.2	31.5	34.7	33.2	33.1	32.9	29.7	9	Gallid_HV-3
10	82.7	78.3	79.4	81.9	84.3	85.5	83.6	20.9	28.2		39.4	44.2	35.3	34.3	35.1	32.3	32.4	35.5	32.8	32.9	32.5	29.7	10	Meleagrid_HV-1
11	166.7	157.1	173.2	165.0	168.6	154.6	154.6	156.3	155.1	150.7		48.6	33.2	34.2	35.7	31.4	31.6	34.9	33.7	33.6	35.4	31.2	11	GTHV
12	153.4	139.0	141.9	136.9	139.1	138.4	144.5	135.6	137.6	133.9	107.3		35.1	34.5	35.3	33.0	33.3	36.3	35.0	35.0	37.1	30.9	12	LETV
13	211.0	207.0	217.0	227.0	224.0	201.0	200.0	203.0	202.0	194.4	203.0	188.4		71.5	62.9	54.7	52.4	47.8	36.7	36.2	36.9	34.2	13	Human_HV-8
14	206.0	192.3	213.0	200.0	209.0	194.9	189.3	195.0	197.0	195.0	196.9	191.8	45.2		61.6	55.7	52.6	49.2	37.2	36.9	37.9	35.3	14	Macaca_Mulatta_RDV
15	209.0	209.0	222.0	216.0	216.0	207.0	218.0	196.0	197.0	195.4	189.0	187.7	65.9	72.2		54.3	51.8	48.4	38.3	38.6	39.5	34.3	15	Ateline_HV-3
16	203.0	197.8	214.0	219.0	219.0	203.0	204.0	206.0	213.0	202.0	209.0	195.8	84.3	82.0	84.1		50.1	47.4	38.9	38.7	38.2	34.6	16	Equine_HV-2
17	209.0	203.0	205.0	212.0	212.0	200.0	200.0	197.0	200.0	193.3	196.6	177.6	94.5	92.2	91.4	91.7		49.2	34.0	34.1	35.0	33.6	17	Porcine_HV-1
18	203.0	197.5	198.0	211.0	215.0	190.0	195.0	201.0	203.0	195.0	188.6	187.3	112.9	106.7	109.7	111.8	95.0		39.8	39.7	39.3	34.5	18	EBV
19	205.0	211.0	210.0	217.0	218.0	228.0	228.0	200.0	200.0	208.0	218.0	185.5	169.7	177.7	168.7	158.5	171.7	150.5		97.0	65.5	45.0	19	Human_HV-6A
20	209.0	211.0	209.0	216.0	218.0	231.0	231.0	201.0	201.0	206.0	219.0	185.5	174.3	180.7	165.1	160.1	170.7	151.9	4.0		65.8	44.5	20	Human_HV-6B
21	196.0	198.0	210.0	200.0	200.0	214.0	220.0	210.0	209.0	210.0	200.0	166.2	168.2	169.2	159.8	163.3	160.9	154.0	64.4	63.6		44.0	21	Human_HV-7
22	208.0	197.0	200.0	209.0	212.0	214.0	211.0	198.1	199.0	201.0	197.5	184.1	175.4	166.5	175.3	171.5	171.6	167.1	113.1	115.4	117.0		22	Human_HV-5
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22		

Figure 3.12.The amino acid sequence distances of the GTHV UL27 gene compared with 21 related herpesviruses. All sequences excluding GTHV were obtained from the NCBI Pubmed website using the accession numbers listed in Figure 3.9,except LETV which has an accession number AY124577.

3.4.3 Phylogenetic Analysis of the UL27 Gene

Through a Peptide Blast Library Search, the GTHV UL27 gene showed high sequence homology to other herpesviral UL27 sequences, specifically lung eye trachea virus (LETV), Gallid HV-2 and Gallid HV-3. A neighbor-joining phylogenetic tree was constructed using an alignment created in the Clustal X program comparing GTHV UL 27 gene to the nucleotide sequences of 21 closely related herpesviruses. LETV was added to the herpesviruses used for comparison. UL27 is the only gene that has been sequenced of this virus to date. The phylogenetic tree, which is displayed using the TreeView program (Figure 3.13), shows the GTHV UL27 homolog is closely related to members of the alphaherpesvirus subfamily, corroborating data reported in previous studies (Yu *et al.* 2001) based on the DNA *pol* gene, as well as from the UL28 and UL29 genes.

The sequence distances of the amino acid sequence of the GTHV UL27 protein compared to 21 closely related herpesviruses was determined in the Megalign program of the DNAstar package using an amino acid alignment created in the Clustal X program. The percent identity of the GTHV UL27 protein to other herpesviruses is shown in Figure 3.12. The GTHV gB gene shows the highest percent identity to LETV (48.6%), followed by Meleagrid HV-1 (39.4%) and Gallid HV-3 (39.3%).

The GTHV gB gene showed varying percent similarity to the three subfamilies of herpesviruses. The alphaherpesvirus subfamily showed both the highest and the lowest percent identities with a range of identity from 39.3% (Meleagrid HV-1) to 28.1% (Equine HV-1). Members of the betaherpesvirus subfamily ranged in percent identity to GTHV UL27 homolog from 35.4% to 31.2%. Members of the gammaherpesvirus subfamily ranged in percent identity to GTHV UL27 homolog from 35.7% to 31.4%.

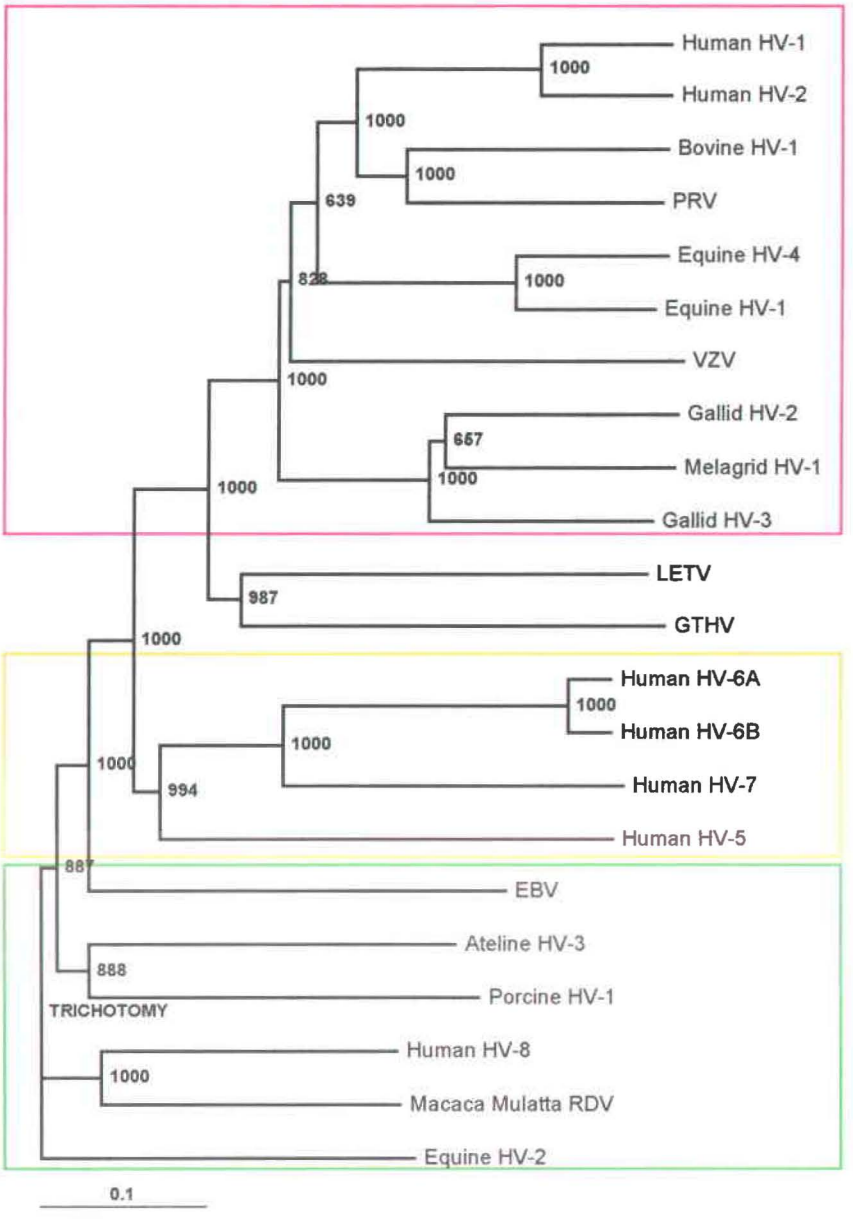


Figure 3.13. Neighbor- joining phylogenetic tree of the GTHV UL27 gene and other herpesviruses. The tree is displayed using the TreeView program. The tree was bootstrapped 1000 times and the bootstrap value for the GTHV branch is 987, indicating the analysis is reliable. All sequences, excluding GTHV, were obtained from the NCBI Pub-Med website using the accession numbers listed in Figure 3.9, except LETV (Lung eye trachea virus), which has the accession number: AY124577. The pink box surrounds members of the alphaherpesvirus subfamily. The yellow and green boxes surround betaherpesviruses, and gammaherpesviruses, respectively.

CHAPTER 4: DISCUSSION

4.1 Summary of Genomic Walking

Employing a modified IPCR genomic walking technique, 9,032 bp of the GTHV genome, a novel *chelonian* herpesvirus, was sequenced. This includes 8,469 bp of previously unknown sequence information. Sequence analysis indicated that this genomic fragment included the full-length sequences of the DNA binding protein gene (UL29), the UL28 transport protein and gB (UL27). The DBP gene has an ORF of 3,585 bp with an expected translation of a 1,195 amino acid peptide. The UL28 gene has an ORF of 2,250 bp encoding a peptide of 750 amino acids. The gB gene has an ORF of 2,551 bp, with an expected translation of an 851 amino acid peptide. The gB gene overlaps the genome of the UL28 gene by 1 bp.

IPCR-based genomic walking and sequencing represents a rapid and effective method to gain unknown nucleic acid sequences and it is designed specifically for amplifying unknown genome sequences adjacent to a known DNA fragment. This cloning and sequencing technique eliminates the requirement for the creation of genomic libraries, which can be time-consuming. The green turtle tumor DNA that was circularized in the second walk of the UL29 gene using the *Eco*R I digestion was also used for the fourth walk. This demonstrates that DNA prepared for earlier walks could be used for subsequent walks, which saves time and research materials.

We have found that successful application of this genomic walking technique for sequencing depends on several steps. First, because the priming region determines which monomeric circles will be amplified, the step of designing and selecting oligonucleotide primer pairs for IPCR is highly critical. We have demonstrated that increasing the length

of the IPCR primers (> 25 nt) facilitates a stringent, specific IPCR amplification. Secondly, it is very important and necessary to use low concentrations of RE-digested DNA in relatively large solution volumes (1μg DNA/ 1ml total solution) during the self-ligation reaction. This setting can promote the efficiency of the formation of monomeric circular DNA and the prevention of production of linear concatemers. The third key step is the selection of a restriction enzyme for genomic DNA digestion. Selected REs are required to cut genomic DNA in such a manner as to generate a 4-bp overhang since this will enhance a more specific self-ligation, and produce products of suitable length for IPCR.

It has been not uncommon to amplify significantly long segments of DNA in IPCR. We chose to use Expand High Fidelity *Taq* directly instead of standard *Taq* DNA polymerase since the latter was previously speculated not to be a suitable polymerase for amplification and sequencing of long DNA fragments for IPCR (Yu *et al.* 2001). Our data showed that the Expand High Fidelity *Taq* –mediated IPCR resulted in large-sized crisp bands during the second and fourth walks of the UL29 sequencing. It is desirable to obtain large-size amplicons in IPCR since such amplification will speed up the time and ease at which sequence data can be obtained. The largest gene fragment we were able to sequence was the UL27 gene, which was sequenced entirely in one genomic walking step. The fragment was 3,219-bp long, and contained 2,603 bp of newly determined sequence data. Additionally, the Expand High Fidelity system was selected because it reduces the amount of PCR-induced mutations compared to standard *Taq*, which ensures accuracy of the sequence data that we obtained.

4.2 Analysis of the GTHV Nucleotide Sequences

The three GTHV genes sequenced in this study are located upstream of the DNA *pol* gene (UL30). The GTHV DBP and DNA *pol* genes are divergently transcribed, with the DNA *pol* gene oriented in the 5' to 3' direction, and the DBP gene oriented in the 3' to 5' direction. The subsequent UL28 is located upstream of the DBP gene, and the two genes are separated by an intron of 184 bp. The GTHV DBP gene is oriented in the 3' to 5' direction, as well. The GTHV gB gene (UL27) is located upstream of the UL28 gene. The gB gene is oriented in the 3' to 5' direction. The unique long (UL) region of the herpesviral genome shows a high level of conservation of gene order. All alphaherpesviruses show a conserved order of these four genes.

The putative origin for viral DNA replication, OriL, which is a 144-bp perfect repeat in HHV-1, was not found in the region between UL29 and UL30 genes of the GTHV genome. OriL has been reported in this region of the genome in other alphaherpesviruses. The genome for HHV-1 contains three origins of viral replication, two copies of OriS and one copy of OriL. The functional significance of these origins of replication is not yet clear, and mutants with deletions of OriL or both copies of OriS are replication competent (Hardwicke and Schaffer 1995). This suggests that the two types of origins can substitute for each other, indicating that the OriL and OriS may be functionally equivalent. OriS and OriL are also highly homologous on the nucleotide sequence level, which strengthens this argument (Hardwicke and Schaffer 1995). OriL is not present in GTHV in the same position as seen in HHV-1 and HHV-2. This has been seen in several other herpesviruses including PRV and BHV-1 (Kupershmidt *et al.* 1991; Meyer *et al.* 1997). Several plausible explanations for the absence of OriL in this

location in GTHV are first that due to the non-essential nature of OriL, viral replication is successful without the origin. Secondly, OriL may be located in another area of the genome, which was observed in other alphaherpesviruses; including PRV, in which OriL is found in the UL2 region of the genome (Kupershmidt *et al.* 1991). This UL2 region has not been sequenced in GTHV. One theory about the significance of the origins is that OriL and OriS initiate different types of viral replication (e.g. theta versus rolling circle) (Hardwicke and Schaffer 1995). If this were true it would limit the replication options of GTHV, permitting OriL is missing from the GTHV genome entirely.

The genomes of herpesviruses, in particular the well-studied HHV-1, share many transcriptional features with both eukaryotic cellular mRNA and the mRNA's expressed by other nuclear-replicating DNA viruses (Wagner *et al.* 1985). Herpesviral genes may contain nucleotide signals for a eukaryotic transcription startpoint, promoter elements, such as a TATA box, or a CAAT box as well as the sequence for a polyadenylation signal (Wagner *et al.* 1985).

The eukaryotic transcription startpoint, which has no extensive sequence homology, has a tendency to be an 'A' residue flanked on either side by pyrimidine residues (Lewin 1997). The elements of the promoter region of the gene can be found upstream of the transcription startpoint. The TATA box, which is involved in the binding of RNA polymerase, is usually found –25 bp upstream of the transcription startpoint. The CAAT box is usually found –75 bp upstream of the startpoint, and is thought to play an important role in the strength of the promoter region (Lewin 1997). The locations of these promoter elements in relation to the startpoint; however, is variable (Lewin 1997). In this study, an attempt was made to use the newly attained sequences of GTHV to identify the

promoter areas of the GTHV genes UL29-UL27, by comparing GTHV to homologous genes in other herpesviruses, as well as finding these elements in relation to each other.

An intron of 160 bp separates the UL30 gene translational stop codon from the translational start codon of the UL28 gene. This intron contains the putative promoter elements as well as the startpoint of transcription. Two potential startpoints of transcription may be the 'A', 24 bp upstream from the translational start codon residue within the sequence 'TCATTC' or the 'A' residue 27 bp upstream from the translational start codon within the sequence 'CCATC'. It is possible that another sequence may act as the startpoint, as well. The putative TATA box is found 53-bp from the translational start codon and is 31 bp and 27 bp away from the two potential startpoints, respectively. A putative CAAT sequence (CAAT) was identified 148 bp away from the translational start codon and 125 bp and 121 bp away from the two potential startpoints, respectively. An illustration of the putative promoter region of the UL29 gene is shown in Figure 4.1.

The UL28 gene is separated from the UL29 gene by an intron of 184 bp. Within this intron, a putative CAAT box was identified on the complimentary strand of the DNA, 106 bp upstream from the translational start codon of the UL28 gene. The sequence of the putative CAAT box is 'ATTG', which is CAAT on the complimentary strand. Although it is not well understood, CAAT boxes are able to function in either orientation (Lewin 1997). A putative startpoint was found 74 bp from the putative CAAT box with the sequence CCCTTCAT. This putative startpoint sequence lies very close to the location that is most common in eukaryotic genes (-75 bp from the startpoint) (Lewin 1997). No TATA box was identified in this intron, before the translational start codon for UL28. There are some A/T rich areas, which may function as the TATA box. A minority

of eukaryotic promoters do not contain TATA elements (Lewin 1997), and it is thought that genes that lack TATA elements, or TATA elements with several mutated nucleotide are not as efficiently transcribed as their counterparts with this element (Lewin 1997). An illustration of the putative promoter region of the GTHV UL28 gene is shown in Figure 4.1.

Because the UL28 and UL27 genes are overlapping, no intron separates the two genes. The upstream transcriptional elements of the UL27 gene are found within the coding region of the UL28 gene. The startpoint for transcription of this gene was identified by comparison to HHV-1 and PRV (Robbins *et al.* 1987). One putative startpoint is 162 bp upstream of the ATG translational start codon for UL27. The sequence in this region is 'CCTTTATCT'. A putative TATA box was identified 35 bp upstream of this sequence with the sequence TAATTA. The putative CAAT box for this gene was identified 367 bp upstream of the startpoint, with the sequence 'GGTCAATC' This is very similar to the eukaryotic consensus sequence "GGCCAATCT' (Lewin 1997). An illustration of this promoter area is shown in Figure 4.1. This promoter region is longer and more spread out than the standard promoter region, however this has been observed in the UL27 gene of other herpesviruses. The UL27 gene of Gallid HV-2 has a TATA box –253 bp from the translational start codon (Kato *et al.* 1999). A possible explanation for this occurrence could be the fact that the putative promoter region is within the coding region of the UL28 gene. The DNA sequence must encode the sequences for both genes, which may result in non-ideal locations for the promoter elements. Secondary DNA structures may play a role in moving the promoter elements

closer together, however data has not been collected on this idea, and research would have to be done to support this claim.

The polyadenylation signal, another eukaryotic mRNA processing element, adds a long string of 'A' residues to the end of an mRNA. For herpesviral genes, the most commonly seen polyadenylation signal is 'AATAAA', which is also the consensus sequence for eukaryotic cells (Wagner *et al.* 1985). This sequence is observed 49 bp downstream of the stop codon for the UL29 gene. This location is conserved among herpesviruses. No polyadenylation signal is seen after the UL28 gene, and it is possible that the 'AATAAA' signal that lies 121 bp downstream from the UL27 stop codon acts as the polyadenylation signal for both genes. This would indicate that as with HHV-1 and several other herpesviruses that the 3' terminus of the GTHV UL28 transcript is coterminal with that of the UL27 transcript (Whalley *et al.* 1989).

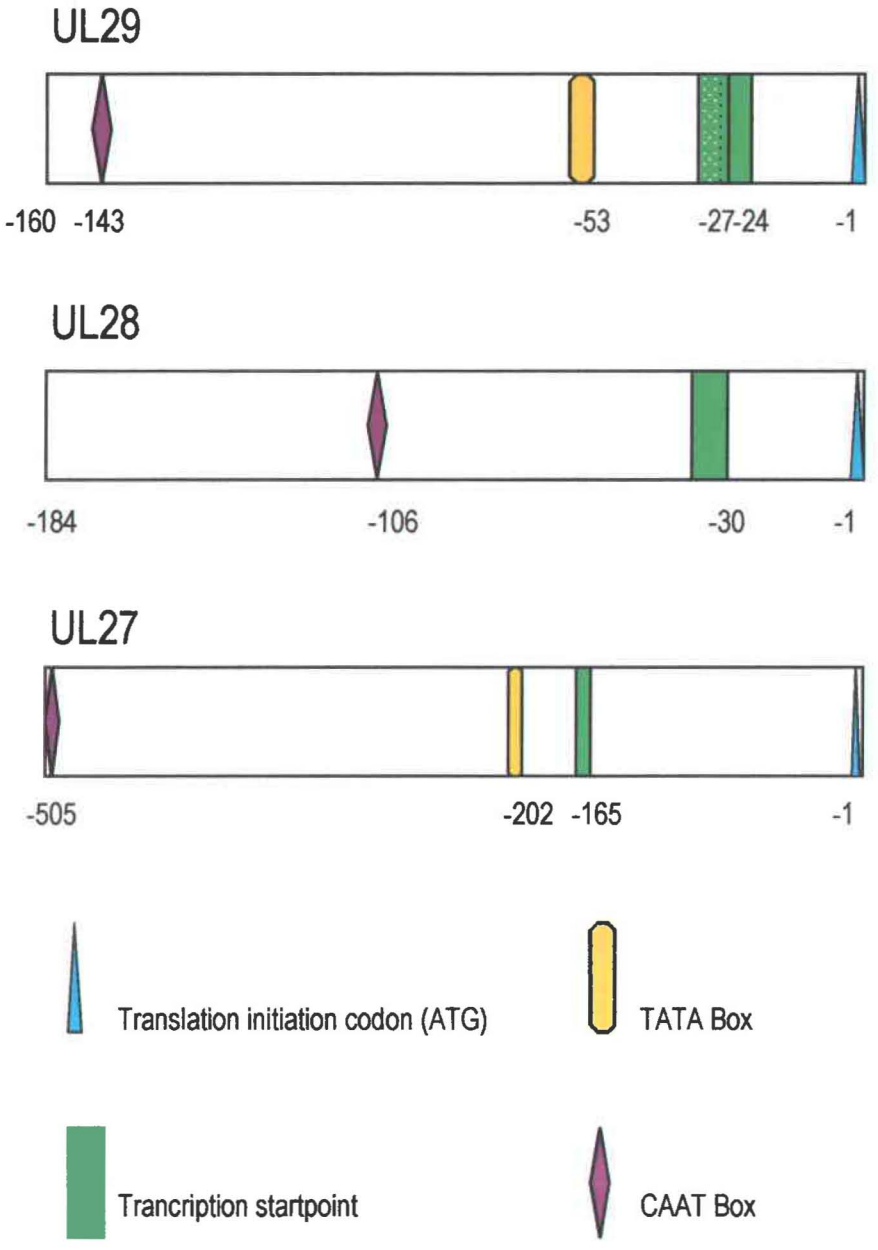


Figure 4.1. Putative promoter regions for GTHV genes UL29, UL28, and UL27. These illustrations indicate only one of the possible promoter sequences interpreted from the sequences of their respective genes. The UL29 promoter illustration represents the 160-bp intron separating UL29 and UL30. The two green boxes indicate two putative transcription startpoints as described in the text. The UL28 promoter illustration represents the 184-bp intron separation the UL28 and UL29 genes. The UL27 and UL28 genes are not separated by an intron, therefore the promoter illustration ends with the putative CAAT box.

4.3 Analysis of the GTHV Amino Acid Sequences

The function of DBP, although unknown in GTHV, has been experimentally determined in other herpesviruses (Ben-Porat *et al.* 1983; Lee and Lehman 1997; White and Boehmer 1999; Cohrs *et al.* 2002; Lee and Knipe 1985). In particular, the homolog DBP of HHV-1, known as infected cell protein No. 8 (ICP-8), has been well studied. This viral-encoded protein has a molecular weight of 130.0 kDa, which is very similar to the DBP of GTHV (approximately 130.811 kDa), and was detected in the nucleus of the infected cell during HHV-1 viral replication (Puvion-Dutilleul *et al.* 1985). This viral protein is also known to bind cooperatively and nonspecifically to single stranded DNA and play a role in DNA replication by binding the DNA of the replication origin and the replication fork after forming a complex with the origin binding protein, UL9 (Wang and Hall 1990; Lee and Lehman 1997; Wu *et al.* 1998). Bortner and his co-workers (Bortner *et al.* 1993) reported that ICP-8 was also involved in renaturation and strand exchange and therefore, it may play a role in DNA recombination. A study on PRV DBP revealed that herpesviral DBP also play a role in the stabilization and protection of progeny DNA in infected cells (Ben-Porat *et al.* 1983). A specific 56-kDa fragment of the ICP-8 protein has been experimentally determined to be the major determinant of DNA binding (Wang and Hall 1990). ICP-8 contains a unique domain with highly conserved aromatic and basic amino acid residues (F, W, Y and H, K, R, respectively). These characteristic residues can form stacking and electrostatic interactions with DNA (Wang and Hall 1990) and are required for single stranded viral DNA binding. The GTHV DBP shows high sequence homology to ICP-8 and other DBP, including the highly conserved

aromatic and basic amino acid sequences and the conserved zinc finger motif, and it is reasonable to expect that the function of GTHV DBP would be the same as ICP-8 and other herpesviral DBP.

The function of the GTHV UL28 gene is unknown, however it has been determined in other herpesviruses. By transcriptional analysis, the HHV-1 UL28 homolog has a molecular weight of 85.6 kDa, which is similar to the molecular weight of GTHV UL28, which has been calculated to be 86.2 kDa. In HHV-1 the UL28 gene, which encodes ICP 18.5 is important for the formation of mature capsids (Addison *et al.* 1990). ICP 18.5, along with the protein products of the UL6 and UL15 genes, are essential for cleavage of replicated DNA into unit length genomes as well as the packaging of DNA into the procapsid (White *et al.* 2003). Capsid assembly of HHV-1 takes place in the nucleus and ICP 18.5 has been observed within the nucleus of HHV-1 infected cells. It has been shown that UL28 and UL15 form a complex (Abbots *et al.* 2000; Koslowski *et al.* 1997; Koslowski *et al.* 1999), which has been hypothesized to be analogous to the terminase of double stranded DNA bacteriophage (White *et al.* 2003). This enzyme is involved in the translocation of DNA into the capsid by cleaving replicated DNA at specific sites, as well as providing energy through the hydrolysis of ATP.

Alignment analysis of the GTHV UL28 gene indicated that the UL28 gene has high sequence homology at both the amino acid and nucleotide levels to the UL28 genes of other herpesviruses, which suggests the function of the GTHV protein encoded by the gene to be similar. Alignment analysis of the amino acid sequence also indicated the presence of a potential nuclear localization signal (NLS) from amino acids 467-469 with

the sequence 'RKR'. The potential NLS is well conserved among herpesviruses (Kato *et al.* 1999). This signal may be important in transferring the protein to the nucleus, where it is hypothesized to function.

The function of gB, although unknown in GTHV, has been well studied in other herpesviruses. In HHV-1, gB is one of several glycoproteins present in the virion envelope. gB is essential for viral infectivity and required for penetration into the host cell. Production of gB in HHV-1 cells is as follows. gB is thought to be synthesized on ribosomes attached to the rough endoplasmic reticulum (RER) (Claesson-Welsh and Spear 1987). Newly synthesized gB is thought to have N-linked oligosaccharides attached soon after gB production. gB is then reassembled into dimers and transported from the RER to the inner nuclear membrane where it is inserted into the virion envelope. The assembled virion is moved via the Golgi apparatus to the cell surface. While in transport, the N-linked oligosaccharides are processed and O-linked oligosaccharides are added to gB and other HHV-1 glycoproteins (Claesson-Welsh and Spear 1987).

The structure of this peptide is very closely linked to its function. Several distinct characteristics of glycoproteins are well conserved. Most envelope glycoproteins, including herpesviral homologs of gB, contain (i) a cleavable N-terminal signal sequence, (ii) a large extraviral region that interacts with the host cell, (iii) an area that anchors the protein in the viral membrane and (iv) a C-terminal cytoplasmic tail that can range in size from 20-150 amino acids (Eckert and Kim 2001). The polarity and hydropathy of the amino acid sequence can often predict the presence of these features, but protein expression studies are required for absolute determination of these features.

The characteristic signal sequence, present in transmembrane glycoprotein genes, such as gB, is thought is used in viral infected cells for translocation of gB across the RER membrane (Claesson-Welsh and Spear 1987). A signal sequence, located towards the beginning of the N-terminus of the peptide, has been found in HHV-1, PRV, EHV-1, and many other gB genes of herpesviruses (Whalley *et al.* 1989; Wolfer e*t al.* 1990; Claesson-Welsh and Spear 1987). The signal sequence is cleaved after transport out of the membrane (Claesson-Welsh and Spear 1987). According to von Heijne (1984) the region around the cleavage site shows strong preferences for specific amino acid residues, in certain positions. The characteristic segments of the signal sequence are a positively charged N-terminal region, a central hydrophobic region and a polar C-terminus (von Heijne 1985).

The SPScan program, which detects signal sequences in peptides, detected a putative signal sequence in the N-terminus of the putative gB peptide. Using this method, the peptide is predicted to be cleaved between amino acids 14 and 15. The score assigned to this putative signal sequence based on the amino acids present was 14.1 out of 15, which is relatively high. The putative signal sequence, however did not pass the McGeoch Scan performed by SPScan. This tests for the presence of both a region of charged amino acids at the N-terminal of the sequence, and the central hydrophobic region. In this peptide the charged N-terminal is absent. This is obvious when looking at the Kyte-Doolittle hydrophilicity plot of GTHV in comparison to HHV-1 (Figure 3.7). In HHV-1, residues 1-12 are polar residues. The graph dips below the line at residue 13 indicating the beginning of the hydrophobic region of the signal sequence. This is followed by the polar C-terminus starting at residue 33. The HHV-1 peptide is cleaved

between amino acids 30 and 31 (Claesson-Welsh and Spear 1987). The GTHV hydrophilicity plot lacks the positive region at the initial amino acid. The hydrophobic region of this putative signal sequence is strong, however, the (-3,-1) rule for signal sequences proposed by von Heijne (1985) is fulfilled. Determining definitively the presence and location of a signal sequence would require further research.

The extraviral portion of HHV-1 gB contains 6 sites for N-glycosylation (Pellett *et al.* 1985). This site is specified by the sequence NXS or NXT where X is any amino acid except proline (Grogan *et al.* 2002). The putative extraviral portion of the GTHV gB peptide contains 12 sites with this sequence indicating 12 potential sites for N-glycosylation. In an alignment made using the BestFit program of the GCG Wisconsin package, two of the N-glycosylation sites align exactly with sites in the HHV-1 gB peptide, and a third site aligns two amino acids away from the site on HHV-1 (Figure 4.2). All 12 N-glycosylation sites may not be utilized in the GTHV gB peptide, due to folding of the protein and space constraints, but these sites are present in the amino acid sequence. Additionally, all of the cytosine residues located in this region of the GTHV gB peptide, are shared with the HHV-1 gB peptide, indicating the potential for common secondary and tertiary structural elements between these peptides. This conservation of cystine residues has been observed in other herpesviruses including PRV (Robbins *et al.* 1997).

The portion of gB that spans the viral membrane has been characterized in HHV-1 and other herpesviruses (Pellett *et al.* 1986; Spatz and Maes 1993; Robbins *et al.* 1987). The hyrophphilicity plots shown in Figure 3.7 show three distinct peaks in the hydrophobic portion of the regions from 726-795 in HHV-1. The transmembrane domain

is predicted to have three hydrophobic segments that transverse the membrane three times and create an anchor for gB. These three peaks have been determined as the three areas where the HHV-1 gB overlaps the viral membrane. The three hydrophobic peaks seen in GTHV from residues 661-730 appear very similar in hydropathic nature to this region of the HHV-1 peptide, indicating a similarity in the structure and function of the two molecules. This region of the GTHV gB gene shows exceptionally high sequence homology to the transmembrane domain of HHV-1 (Figure 4.2).

The C-terminal portion of the GTHV gB peptide, beginning with amino acid residue 730, shows a high degree of hydrophilicity, similar to this portion of the peptide in other herpesviruses (Figure 3.7), as well as having a positive charge (5.054 at pH 7.0) similar to other herpesviruses (Pellett *et al.* 1986; Robbins *et al.* 1987). In HHV-1, the gB peptide is theorized to be placed on the underside of the viral membrane due to the fact that the hydrophobic region of the peptide traverses the membrane an odd number of times (3) (Pellett *et al.* 1986). Due to similar data found upon examination of the GTHV gB peptide a similar location can be inferred.

Figure 4.2. Alignmnents of the extraviral and transmembrane domains of GTHV and HHV-1. A. Alignment of the putative extarviral domain of GTHV gB with HHV-1 gB. N-glycosylation sites of HHV-1 are highlited in blue. Putative N-glycosylation sites of GTHV are highlighted in yellow. N-glycosylation sites common to both viruses are boxed. Cytosine residues common to both viruses are boxed as well.
B. Alignment of the putative transmembrane domain of GTHV gB with HHV-1 gB.

4.4 Phylogenetic Analysis of GTHV Gene Sequences

Previous study on the GTHV *pol* gene and its phylogenetic relationship to other herpesviruses suggested GTHV to be a member belonging to the alphaherpesvirus subfamily. Present analysis of both the nucleic acid and the anticipated amino acid sequences of GTHV UL29, UL28 and UL27 genes add further support to this classification of GTHV.

The alphaherpesvirus classification of GTHV is supported by the conservation of the gene order and orientation of the UL27-UL30 genome block. The order of the UL27 to UL30 genes and the putative proteins they encode is highly conserved among alphaherpesviruses with the genes in the order DNA *pol*, DBP, UL28, gB. The arrangements of this gene block in the betaherpesviruses that were analyzed in this study (listed in Table 4.1) have the gene order DNA *pol*, gB, UL28, DBP. The gammaherpesviral sequences analyzed in this study (listed in Table 4.1), with the exception of EBV, have the gene order DBP, UL28, gB, DNA *pol*. EBV has the same order as the betaherpesviruses. The alphaherpesviruses in this study all have divergently transcribed UL29 and UL30 genes, which is not observed among the beta- and gammaherpesviruses that were analyzed in this study (Table 4.1),

Based on the sequence of GTHV UL29 gene, GTHV appears to be most closely related to the homologous genes of HHV-1 (simplexvirus), HHV-2, BHV-1, BHV-2, EHV-1, and PRV. However, such a relationship branches early in the tree suggesting a possible divergence of GTHV from other genus's of the subfamily. Similar results were seen upon analysis of the UL28 gene. The phylogenetic tree created in this study (Figure 3.9) groups the GTHV UL28 gene closest to the alphaherpesvirus subfamily. Analysis

based on the amino acid sequences of UL28 showed the greatest percent identity of GTHV to PRV, which is a member of the varicellovirus genus. This is similar to the GTHV DNA *pol* gene, which also showed the greatest sequence homology to PRV (Table 4.1).

The percent identity of the GTHV UL27 gene to other herpesviruses did not group as clearly with the alphaherpesviruses as the UL29 and UL28 genes. In order to ascertain that this did not represent a frameshift mutation in the GTHV UL27 sequence, the sequence distances were determined using the nucleic acid sequences of the 22 viruses, and a similar pattern of results occurred (data not shown). The highest percent identity (excluding LETV) and the lowest percent identity were both found within the alphaherpesvirus subfamily, with a range 39.4% to 28.1%. Betaherpesviruses and gammaherpesviruses both ranged from around 35% identity to 31%, with several members of each subfamily showing higher percent identity to GTHV UL27 than members of the alphaherpesvirus subfamily. Surprisingly Meleagrid HV-1, a member of the 'Marek's Disease –like virus' genus showed the highest percent identity, with the second and third highest percent identity to two other members of this genus Gallid HV-2 and Gallid HV-3. One speculation about this difference in grouping from the UL29 and UL28 genes is that because this protein interacts directly with host cell structures, the genetic variance in the viral gene may have been evolutionarily directed by host specificity. The fact that the three most homologous genes are from viruses that infect avian cells, could corroborate this data, as reptiles are closer evolutionarily to avian species than mammals.

The phylogenetic tree created based on the nucleotide sequence of the GTHV UL27 gene compared to other herpesviruses showed similar results to the two previously discussed genes, with GTHV grouping most closely with the alphaherpesviruses, but not within one particular family. This finding is similar to the data for the UL29 and UL28 genes. All three GTHV genes have showed the greatest percent identity to members of three different genuses. This great variance may indicate that GTHV represents a novel genus of alphaherpesvirus.

LETV was also included in the phylogenetic comparisons of GTHV to other herpesviruses. LETV is hypothesized to cause lung-eye-trachea disease in marine turtles, which is associated with conjunctivitis, pharyngitis, tracheitis and pneumonia (Coberley *et al.* 2001). LETV is the only reptilian herpesvirus, aside from GTHV that has been sequenced. Unfortunately, only the UL27 gene of LETV, of the genes involved in this study, has been sequenced. The nucleotide sequence GTHV UL27 gene grouped most closely with LETV in the phylogenetic tree created in this study. Percent identity to the amino acid sequence of LETH to GTHV was the closest of the viruses analyzed. The UL27 gene of LETV grouped outside of the three alphaherpesviral genuses and together the two viruses may represent a novel genus.

Table 4.2. Compilation of the percent identity of herpesviruses to GTHV. The table shows the percent identity of the four genes of GTHV that have been sequenced to date. The orientation of each gene is listed in the column following the percent identity. Alaphaherpesviruses are separated from the betaherpesviruses by a double line, and betaherpesviruses are separated from gammaherpesviruses by a double line. The highest percent identity for each gene is in bold. LETV showed the highest percent identity to gB (not shown) with 48.6% identity.

Virus	Genus	DNA Pol Gene % Identity	Gene Orientation	DBP Gene % Identity	Gene Orientation	UL28 Homolog % Identity	Gene Orientation	gB % Identity	Gene Orientation
GTHV	**N/A**	**N/A**	**5' to 3'**	**N/A**	**3' to 5'**	**N/A**	**3' to 5'**	**N/A**	**3' to 5'**
HHV-1	Simplexvirus	49.7%	5' to 3'	36.1%	3' to 5'	45.5%	3' to 5'	33.8%	3' to 5'
HHV-2	Simplexvirus	49.7%	5' to 3'	**36.2%**	3' to 5'	45.7%	3' to 5'	34.1%	3' to 5'
Bovine HV-1	Varicellovirus	49.6%	3' to 5'	36.0%	5' to 3'	39.9%	5' to 3'	31.2%	5' to 3'
Equine HV-1	Varicellovirus	50.6%	3' to 5'	35.1%	5' to 3'	46.6%	5' to 3'	28.1%	5' to 3'
Equine HV-4	Varicellovirus	50.6%	3' to 5'	34.9%	5' to 3'	45.1%	5' to 3'	28.3%	5' to 3'
VZV	Varicellovirus	53.5%	3' to 5'	33.2%	5' to 3'	43.4%	5' to 3'	34.6%	5' to 3'
PRV	Varicellovirus	**53.6%**	5' to 3'	35.1%	3' to 5'	**48.8%**	3' to 5'	32.4%	3' to 5'
Meleagrid HV-1	Marek's Disease-like Virus	50.6%	5' to 3'	33.8%	3' to 5'	44.2%	3' to 5'	**39.4%**	3' to 5'
Gallid HV-3	Marek's Disease-like Virus	49.1%	5' to 3'	34.2%	3' to 5'	43.8%	3' to 5'	39.0%	3' to 5'
Gallid HV-2	Marek's Disease-like Virus	49.4%	5' to 3'	34.2%	3' to 5'	42.3%	3' to 5'	39.3%	3' to 5'
HHV-5	Cytomegalovirus	32.6%	3' to 5'	19.5%	3' to 5'	34.0%	3' to 5'	31.2%	3' to 5'
HHV-6a	Roselovirus	41.3%	3' to 5'	22.5%	3' to 5'	38.4%	3' to 5'	33.7%	3' to 5'
HHV-6b	Roselovirus	41.3%	3' to 5'	22.9%	3' to 5'	38.6%	3' to 5'	33.6%	3' to 5'
HHV-7	Unclassified	40.9%	3' to 5'	22.2%	3' to 5'	37.0%	3' to 5'	35.4%	3' to 5'
Macaca Mulatta RDV	Rhadinovirus	44.8%	5' to 3'	25.6%	5' to 3'	36.7%	5' to 3'	34.2%	5' to 3'
Ateline HV-3	Rhadinovirus	44.2%	5' to 3'	25.1%	5' to 3'	34.1%	5' to 3'	35.7%	5' to 3'
Equine HV-2	Rhadinovirus	44.8%	5' to 3'	25.9%	5' to 3'	36.2%	5' to 3'	31.4%	5' to 3'
HHV-8	Rhadinovirus	44.4%	5' to 3'	25.7%	5' to 3'	36.6%	5' to 3'	33.2%	5' to 3'
Porcine HV-1	Unclassified	44.4%	5' to 3'	25.3%	5' to 3'	36.5%	5' to 3'	31.6%	5' to 3'
EBV	Lymphocryptovirus	45.3%	3' to 5'	27.1%	3' to 5'	25.5%	3' to 5'	34.9%	3' to 5'

4.5 Future Research

The ultimate goal in all research on GTFP is to first elucidate the agent responsible for causing this devastating disease, and second to establish methods of diagnosis and treatment of the viral infection. Current investigation of GTHV has established this virus as the major candidate responsible for the onset of GTFP in green sea turtles.

Immunological-based studies, using bacteria expressed viral proteins could provide an excellent vehicle to establish a closer relationship between GTHV and GTFP. Sequencing of the GTHV UL30- UL27 genes has provided essential information for such studies. Using the newly derived sequence data of the UL28 gene, we have been able to construct a gene expression cassette for GTHV DBP using a pET vector. With the production of recombinant GTHV DBP from transformed *E. coli*, we are currently preparing polyclonal antiserum against GTHV DBP in New Zealand white rabbits. This antiserum will lead to the establishment of highly sensitive and specific immunoassays for the diagnosis of GTHV. In HHV-1, the DBP is one of seven genes that are necessary and sufficient for viral replication, and also one of three genes to be expressed in relatively large quantities (Olivo *et al.* 1989). These data suggest that herpesviral DBP can be used as a sensitive biological marker for the development of immunological - based assays for early detection of viral infection. Production of anti-GTHV DBP serum makes it possible to conduct immunological studies on GTHV and to generate fundamental information in unveiling a pathogenic relationship between GTHV and GTFP.

Similar studies will be done using the gB gene of GTHV. The gB gene of other herpesviruses has been deemed one of the most antigenic viral-encoded proteins (Franti *et al.* 2002). Because gB is a surface protein, it will be readily available for binding of anti-GTHV serum and may be essential in answering the pivotal questions surrounding this causality of this disease. Recently, regions of the gB peptide of Bovine HV-1 were used in conjunction with portions of the glycoprotein D (gD) peptide in an attempt to create a divalent DNA vaccine (Manoj *et al.* 2003). gB has been targeted for several other vaccines as well (Caselli *et al.* 2001; Loomis-Huff *et al.* 2001). Additional studies with these potentially pathogenic bacteria-expressed genes could include transformation of healthy green sea turtle cell lines. If a causal relationship is determined to exist between GTHV and GTFP studies similar to this would be paramount to fight a disease, which threatens to decimate the worldwide population of green sea turtles.

References

Abbotts AP, Preston VG, Hughes M, Patel AH, Stow ND. 2000. Interaction of the herpes simplex virus type 1 packaging protein UL15 with full-length and deleted forms of the UL28 protein. J Gen Virol. 81: 2999-3009.

Addison C, Rixon FJ, Preston VG. 1990. Herpes simplex virus type 1 UL28 gene product is important for the formation of mature capsids. J Gen Virol. 71: 2377-2384.

Aguirre AA. 2000a. Rescue, rehabilitation, and release of marine turtles with fibropapillomatosis: An epidemiologic perspective. In Abreu-Grobois FA, Briseno-Duenas R, Marquez-Millan R, and Sarti-Martinez L (comps.), Proceedings of the Eighteenth International Sea Turtle Symposium, March 3-7, 1998, Mazatlan, Sinaloa, Mexico, p. 111. U. S. Dep. Commer., NOAA Tech. Memo. NMFS-436.

Aguirre AA , Spraker TR, Chaves A, du Toit L, Eure W, Balazs GH. 2000b. Fibropapillomas in olive ridley turtles in Costa Rica. In Abreu-Grobois FA, Briseno-Duenas R, Marquez-Millan R, Sarti-Martinez L (comps.), Proceedings of the Eighteenth International Sea Turtle Symposium, March 3-7, 1998, Mazatlan, Sinaloa, Mexico, p. 111. U. S. Dep. Commer., NOAA Tech. Memo. NMFS-436.

Balazs GH, Pooley SG. 1991. Research plan for marine turtle fibropapilloma. U.S. Department of Commerce, pp 1-113.

Balazs GH, Murakawa SKK, Ellis DM, Aguirre AA. 2000a. Manifestation of fibropapillomas and rates of growth of green turtles at Kaneohe Bay in the Hawaiian Islands. In Abreu-Grobois FA, Briseno-Duenas R, Marquez-Millan R, Sarti-Martinez L (comps.), Proceedings of the Eighteenth International Sea Turtle Symposium, March 3-7, 1998, Mazatlan, Sinaloa, Mexico, p. 112-113 U. S. Dep. Commer., NOAA Tech. Memo. NMFS-436.

Balazs, GH, Rice M, Murakawa SKK, Watson G. 2000b. Growth rates and residency of immature turtles at Kiholo Bay, Hawaii. In Abreu-Grobois FA, Briseno-Duenas R, Marquez-Millan R, Sarti-Martinez L (comps.), Proceedings of the Eighteenth International Sea Turtle Symposium, March 3-7, 1998, Mazatlan, Sinaloa, Mexico, p. 283-285. U. S. Dep. Commer., NOAA Tech. Memo. NMFS-436.

Balazs GH, Puleloa W, Medeiros E, Murakawa SKK, Ellis DM. 1998. Growth rates and incidence of fibropapillomatosis in Hawaiian green turtles utilizing coastal foraging pastures at Palaau, Moloki. In Epperly SP and Braun J (comps.), Proceedings of the Seventeenth Annual Symposium of Sea Turtle Biology and Conservation, March 4-8, 1997, Orlando, Florida, p. 130-132. U. S. Dep. Commer.,NOAA Tech. Memo. NMFS-SEFSC-415.

Ben-Porat T, Veach RA, Hampl H. 1983. Functions of the major nonstructural DNA binding protein of a herpesvirus (pseudorabies). Virology. 124: 411-24.

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

Breathnach R, Chambon P. 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem. 50: 349-383.

Caselli E. Balboni PG, Incorvaia C, Argnani R, Parmeggiani F, Cassai E, Manseervigi R. 2001. Local and systematic inoculation of DNA or protein gB1s-based vaccines induce a protective immunity against rabbit ocular HSV-1 infection. Vaccine 19: 1225-1231.

Casey RN, Quackenbush SL, Work TM, Balazs GH, Bowser PR, Casey JW. 1997. Evidence for retrovirus infection in green turtles *Chelonia mydas* from the Hawaiian Islands. Dis Aquat Org. 31: 1-7.

Claesson-Welsh L, Spear PG. 1987. Amino-terminal sequence, synthesis, and membrane insertion of glycoprotein B of herpes simplex virus type I. J Virol. 61: 1-7.

Coberley SS, Condit RC, Herbst LH, Klein PA. 2002. Identification and expression of immunogenic proteins of a disease-associated marine turtle herpesvirus. J Virol. 76: 10553-10558.

Coberley SS, Herbst LH, Ehrhart LM, Bagley DA, Hirama S, Jacobson ER, Klein PA. 2001. Survey of Florida green turtles for exposure to a disease-associated herpesvirus. Dis Aquat Org. 47: 159-167.

Cohrs RJ, Wischer J, Essman C, Gilden DH. 2002. Characterization of varicella-zoster virus gene 21 and 29 proteins in infected cells. J Virol. 76: 7228-7238.

Curry SS, Brown DR, Gaskin JM, Jacobson ER, Erhart LM, Blahak S, Herbst LH, Klein PA. 2000. Persistent infectivity of a disease-associated herpesvirus in green turtles after exposure to seawater. J Wildlife Dis. 36: 792-797.

Eckert DM, Kim PS. 2001. Mechanisms of viral membrane fusion and its inhibition. Annu Rev Biochem. 70: 777-810.

Ernst CH, Barbour RW, Lovich JE. 1994. *Turtles of the United States and Canada.* Washington: Smithsonian Institute Press.

Franti M, Aubin JT, De Saint-Maur G, Kosuge H, Yamnishi K, Gautheret-Dejean A, Garbarg-Chenon A, Huraux JM, Augut H. 2002. Immune reactivity of human sera to the glycoprotein B of human herpesvirus 7. J Clin Microbiol. 40: 44-51.

Grogan MJ, Pratt MR, Marcaurelle LA, Bertozzi, CR. 2002. Homogeneous glycopeptides and glycoproteins for biological investigation. Annu Rev Biochem. 71: 593-634.

Guillen L. Villalobos JP. 2000. Papillomas in Kemp's ridley turtles. *In* Kalb H Wibbels T (comps.), Proceedings of the Nineteenth Annual Symposium on Sea Turtle Conservation and Biology, March 2-6, 1999, South Padre Island, Texas, p. 237. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-443.

Hardwicke MA, Schaffer PA. 1995. Cloning and characterization of herpes simplex virus type 1 oriL: comparison of replication and protein-DNA complex formation by oriL and oriS. J Virol. 69: 1377-1388.

Hartl DL, Ochman H. 1994. Inverse polymerase chain reaction. Meth of Mol Biol. 31: 187-96.

Herbst LH, Jacobson ER, Moretti R, Brown T, Sundberg JP, Klein PA. 1995. Experimental transmission of green turtle fibropapillomatosis using cell-free tumor extracts. Diseases of Aquatic Organisms 22: 1-12.

Herbst LH, Moretti R, Brown T. 1996a. Autogenous vaccination as an adjunct to surgery in the rehabilitation of green turtles with fibropapillomatosis. In Keinath JA, Barnard DE, Musick JA, Bell BA (comps.), Proceedings of the Fifteenth Annual Symposium on Sea Turtle Biology and Conservation, February 20-25, 1995, Hilton Head, South Carolina, p. 136. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-387.

Herbst LH, Moretti R, Brown T, Klein PA. 1996b. Sensitivity of the transmissible green turtle fibropapillomatosis agent to chloroform and ultracentrifugation conditions. Dis Aquat Org. 25: 225-228.

Herbst LH, Jacobson ER, Moretti R, Brown T, Klein PA. 1994. Green turtle fibropapillomatosis: Transmission study update. In K. A. Bjorndal, A. B. Bolten, and D. A. Johnson (comps.). Proceedings of the Fourteenth Annual Symposium on Sea Turtle Biology and Conservation, March 1-5, 1994, Hilton Head, South Carolina, p. 55. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-351.

Herbst LH. 1999. Marine turtle fibropapillomatosis: Hope floats in a sea of ignorance. In: Nineteenth Annual Symposium on Sea Turtle Conservation and Biology, South Padre Island, Texas. U. S. Department of Commerce, pp 39-40.

Huerta P, Pineda H, Aguirre AA, Spraker T, Sarti L, Barragán A. 2002. First confirmed case of fibropapilloma in a leatherback turtle (*Dermochelys coriacea*). In Mosier A, Foley A, Brost B (comps.), Proceedings of the Twentieth Annual Symposium on Sea Turtle Biology and Conservation, February 29-March 4, 2000, Orlando, Florida, p. 193. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-SEFSC-477.

Jacobson ER, Buergelt C, Williams B, Harris RK. 1991. Herpesvirus in cutaneous fibropapillomas of the green turtle, *Chelonia mydas*. Dis of Aquat Org. 12: 1-6.

Jacobson ER, Mansell JL, Sundberg JP, Hajarr L, Reichmann ME, Ehrhart LM, Walsh M, Murru F. 1989. Cutaneous fibropapillomas of green turtles, *Chelonia mydas*. J Comp Path. 101: 39-52.

Kato K, Jang HK, Izumiya Y, Cai JS, Tsushima Y, Miyazawa T, Kai C, Mikami T. 1999. Identification of the Marek's disease virus serotype 2 genes homologous to the glycoprotein B (UL27), ICP18.5 (UL28), and DNA-binding protein (UL29) genes of herpes simplex virus type 1. J Vet Med Sci. 61: 1161-1165.

Koga SK, Balazs GH. 1996. Sex ratios of green turtles stranded in the Hawaiian Islands. In Keinath JA, Barnard DE, Musick JA, and Bell BA (comps.), Proceedings of the Fifteenth Annual Symposium on Sea Turtle Biology and Conservation, February 20-25, 1995. Hilton Head, South Carolina, p. 148-152. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-387.

Koslowski KM, Shaver PR, Casey JT, Wilson T, Yamanaka G, Sheaffer AK, Tenney DJ, Pederson NE. 1999. Physical and functional interactions between the herpes simplex virus UL15 and UL28 DNA cleavage and packaging proteins. J Virol. 73: 1704-1707.

Koslowski KM, Shaver PR, Wang XY, Tenney DJ, Pederson NE. 1997. The pseudorabies UL28 protein enters the nucleus after coexpression with the herpes simplex virus UL15 protein. J Virol. 71: 9118-9123.

Kupershmidt S, DeMarchi JM, Lu ZQ, Ben-Portat T. 1991. Analysis of an origin of DNA replication located at the L terminus of the genome of pseudorabies virus. J Virol. 65: 6283-6291.

Kyte J, Doolittle R. 1982. A simple method for displaying the hydropathic character of protein. J Mol Biol. 157: 105-132.

Lackovich J, Brown DR, Klein PA. 1998. PCR confirms absence of papillomavirus from sea turtle fibropapilomas. In: Proceedings of the Eighteenth International Sea Turtle Symposium, Mazatlan, Sinaloa, Mexico. U.S. Department of Commerce. NOAA Technical Memo, pp 273.

Lackovich JK, Brown DR, Homer BL, Garber RL, Mader DR, Moretti RH, Patterson AD, Herbst LH, Oros J, Jacobson ER, Curry SS, Klein PA. 1999. Association of herpesvirus with fibropapillomatosis of the green turtle *Chelonia mydas* and the loggerhead turtle Caretta caretta in Florida. Diseases of Aquatic Organisms. 37: 89-97.

Landsberg JH, Balazs GH, Steidinger KA, Baden DG, Work TM, Russell DJ. 1999. The potential role of natural tumor promoters in marine turtle fibropapillomatosis. J Aquat Animal Health. 11:199-210.

Lee CK, Knipe DM. 1985. An immunoassay for the study of DNA-binding activities of herpes simplex virus protein ICP8. J Virol. 54: 731-738.

Lee SS, Lehman IR. 1997. Unwinding of the box I element of a herpes simplex virus type 1 origin by a complex of the viral origin binding protein, single-strand DNA binding protein, and single-stranded DNA. Proc Natl Acad Sci U S A 94: 2838-2842.

Lehman IR, Boehmer PE. 1999. Replication of herpes simplex virus DNA. J Biol Chem. 274: 28059-28062.

Lewin, B, 1997. *Genes VI.* New York: Oxford University Press.

Loomis-Huff JE, Eberle R, Lockridge KM, Rhodes G, Barry PA. 2001. Immunogenicity of a DNA vaccine against herpes B virus in mice and rhesus macaque. Vaccine. 19: 4865-4873.

Losey GS, Balazs GH, Privitera LA. 1994. A cleaning symbiosis between the wrasse, *Thlassoma dupery*, and the green turtle *Chelonia mydas*. Copeia. 3: 684-690.

Lu Y, Aguirre AA, Work TH, Balazs GH, Nerurkar VR, Yanigahara R. 2000a. Identification of a small, naked virus in tumor-like aggregates in cell-lines derived from a green turtle, Chelonia mydas, with fibropapillomas. J Virol Meth. 86: 25-33.

Lu Y. Wang Y, Yu Q, Aguirre AA, Balazs GH, Nerurkar VR Yanigahara R. 2000b. Detection of herpesviral sequences in tissues of green turtles with fibropapillomas by PCR. Arch Virol. 145: 1885-1893.

Lu Y, Wang Y, Aguirre AA, Zhao ZS, Liu CY, Nerurkar VR, Yanigahara R. 2003. RT-PCR detection of the expression of the polymerase gene of a novel reptilian herpesvirus in tumor tissues of green turtles with fibropapillomas. Arch Virol. In Press.

Manoj S, Babuik LA, van Drunen-Little-van den Hurk S. 2003. Immunization with a dicistronic plasmid expressing truncated form of bovine herpesvirus-1 glycoprotein D and the amino-terminal subunit of glycoprotein B results in reduced gB-specefic immune responses. Virology. 313: 296-307.

Matushima ER, Filho AL, DiLoretto C, Kanamura CT, Gallo B, Baptistotte MC. 2000. Cutaneous papillomas of green turtles: a morphological and immunohistochemical study in Brazilian specimens. In Abreu-Grobois FA, Briseno-Duenas R, Marquez-Millan R, Sarti-Martinez L (comps.), Proceedings of the Ninteenth International Sea Turtle Symposium. March 2-6, 1999. South Padre Island, Texas, USA. p. 237-239. U. S. Dep. Commer., NOAA Tech. Memo. NMFS-SEFSC-443.

Meyer G, Vlcek C, Paces V, O'Hara MK, Pastoret PP, Thiry E, Schwyzer M. 1997. Sequence analysis of the bovine herpesvirus type 1 genes homologous to the DNA polymerase (UL30), the major DNA-binding protein (UL29) and the ICP 18.5 assembly protein (UL28) genes off herpes simplex virus. Arch Virol. 142: 89-102.

Miao XS, Balazs GH, Murakawa SKK, Li QX. 2001. Congener-specific profile and toxicity assessment of PBC's in green turtles from the Hawaiian Islands. Sci Total Environ. 281: 247-253.

Murakawa SKK, Balazs GH, Ellis DM, Hau S, Eames SM. 2000. Trends in fibropapillomatosis among green turtles stranded in the Hawaiian Islands. In Kalb H, Wibbels T (comps.), Proceedings of the Nineteenth Annual Symposium on Sea Turtle Conservation and Biology, March 2-6, 1999, South Padre Island, Texas, p. 239-241. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-443.

National Marine and Fisheries Service (NMFS). 1998. Recovery plan for U.S. Pacific populations of the green sea turtle (*Chelonia Mydas*). National Marine Fisheries Service, Silver Spring, MD.

Norton TM, Jacobson ER, Sundberg JP. 1990. Cutaneous fibropapillomas and renal myxofibroma in a green turtle, *Chelonia mydas*. J Wildlife Dis. 26: 265-270.

Olivo PD, Nelson NJ, Challberg MD. 1989. Herpes simplex virus type 1 gene products required for DNA replication: identification and overexpression. J Virol. 63: 196-204.

Page RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci. 12: 357-358.

Papadi GP, Balazs GH, Jacobson ER. 1995. Flow cytometric DNA content analysis of fibropapillomas in green turtles (*Chelonia mydas*). Dis Aquat Org. 22: 13-18.

Pellett PE, Jenkins FJ, Ackermann M, Sarmiento M, Roizman B. 1986. Transcription initiation sites and nucleotide sequence of a herpes simplex virus 1 gene conserved in the Epstein-Barr virus genome and reported to affect the transport of viral glycoproteins. J Virol. 60: 1134-1140.

Pellet PE, Kousoulas KG, Pereria L, Roizman B. 1985. Anatomy of the herpes simplex virus strain F glycoprotein B gene: primary sequence and predicted protein structure of the wild type and of monoclonal antibody mutants. J Virol. 53: 243-253.

Puvion-Dutilleul F, Laithier M, Sheldrick P. 1985. Ultrastructural localization of herpes simplex virus major DNA-binding protein in the nucleus of infected cells. J Gen Virol. 66: 15-30.

Quackenbush SL, Work TM, Balazs GH, Casey RN, J. Rovnak, Chaves A, duToit L, Baines JD, Parrish CD, Bowser PR, Casey JW. 1998. Three closely related herpesviruses are associated with fibropapillomatosis in marine turtles. Virology. 246: 392-399.

Quackenbush SL, Casey RN, Murcek RJ, Paul TA, Work TM, Limpus CJ, Chaves A, DuToit L, Vasconcelos-P. J, Aguirre AA, Spraker TR, Horrocks JA, Vermeer LA, Balazs GH, Casey JW. 2001. Quantitative analysis of herpesvirus sequences from normal tissue and fibropapillomas of marine turtles with real-time PCR. Virology. 287:105-111.

Robbins AK, Dorney DJ, Wathen MW, Whealy ME, Gold C, Watson RJ, Holland LE, Weed SD, Levine M, Glorioso JC, Enquist LW. 1987. The pseudorabies virus gII gene is closely related to the gB glycoprotein gene of herpes simplex virus. J Virol. 61: 2691-2701.

Schroeder BA, Foley AM, Witherington BE, Mosier AE. 1998. Ecology of marine turtles in Florida Bay: Population structure, distribution, and occurrence of fibropapilloma. In Epperly SP and Braun J (comps.), Proceedings of the Seventeenth Annual Symposium on Sea Turtle Biology and Conservation, March 4-8, 1997, Orlando, Florida, p. 265-267. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-415.

Smith GM, Coates CW. 1938. Fibro-epithelial growths of the skin in large marine turtles, *Chelonia mydas* (Linneaus). Zoologica (NY) 23:93-98.

Spatz SJ, Maes RK, 1993. Immunological Characterization of the feline herpesvirus-1 glycoprotein B and analysis of its deduced amino acid sequence. Virology. 197: 125-136.

Swimmer JY, Whittow GC, Balazs GH. 1996. Atmospheric basking in Hawaiian green sea turtle, *Chelonia Mydas*: Comparison of tumored and non-tumored turtles. In Keinath JA, Barnard DE, Musick JA, and Bell BA (comps.), Proceedings of the Fifteenth Annual Symposium on Sea Turtle Biology and Conservation, February 20-25, 1995. Hilton Head, South Carolina, p. 318-322. U.S. Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-387.

von Heijne, G. 1984. How signal sequences maintain cleavage specificity. J Mol Biol. 173: 243-251.

von Heijne, G. 1985. Signal sequence: The limits of variation. J Mol Biol. 184: 99-105.

Wagner EK, Costa RH, Gayathri Devi B, Draper KG, Frink RJ, Hall LM, Rice MK, Steinhart W. 1985. Herpesvirus mRNA. In Becker Y (Ed.) Viral Messenger RNA: Transcription, Processing, Splicing, And Molecular Structure. Boston: Martinus Nijhoff Publishing.

Wang YS, Hall JD. 1990. Characterization of a major DNA-binding domain in the herpes simplex virus type 1 DNA-binding protein (ICP8). J Virol 64: 2082-2089.

Whalley JM, Robertson GR, Scott NA, Hudson GC, Bell CW, Woodworth LM. 1989. Identification and nucleotide sequence of a gene in Equine Herpesvirus 1 analogous to the herpes simplex virus gene encoding the major envelope glycoprotein gB. J Gen Virol. 70: 383-394.

White CA, Stow ND, Patel AH, Hughes M, Preston VG. 2003. Herpes simplex virus type 1 portal protein UL6 interacts with the putative terminase subunits of UL15 and UL28. J Virol. 77: 6351-6358.

White EJ, Boehmer PE. 1999. Photoaffinity labeling of the herpes simplex virus type-1 single-strand DNA-binding protein (ICP8) with oligodeoxyribonucleotides. Biochem Biophys Res Commun. 264: 493-497.

Wolfer U, Kruft V, Sawitzky D, Hampl H, Wittmann-Liebold B, Habermehl KO. 1990. Processing of pseudorabies virus glycoprotein gII. J Virol. 64: 3122-3125.

Wu SL, Hsiang CY, Ho TY, Chang TJ. 1998. Identification, expression, and characterization of the pseudorabies virus DNA-binding protein gene and gene product. Virus Res 56: 1-9.

Yu Q, Lu Y, Nerurkar VR, Yanagihara R. 2000. Amplification and analysis of DNA flanking known sequences of a novel herpesvirus from green turtles with fibropapilloma Brief report. Arch Virol. 145: 2669-2676.

Yu Q, Lu Y, Nerurkar VR, Yanagihara R. 2001. Rapid acquisition of entire DNA polymerase gene of a novel herpesvirus from green turtle fibropapilloma by a genomic walking technique. J Virol Meth. 91:183-195.