Virus Genes 12:2, 107–116, 1996
© 1996 Kluwer Academic Publishers, Boston. Manufactured in The Netherlands.

A Genomic Map of Infectious Laryngotracheitis Virus and the Sequence and Organization of Genes Present in the Unique Short and Flanking Regions

MARTHA A. WILD, STEPHANIE COOK, & MARK COCHRAN
martha_wild@syntro.cts.com

Received September 30, 1995; accepted October 19, 1995

Requests for reprints should be addressed to Martha A. Wild, Syntro Corporation, 3535 General Atomics Court, San Diego, CA 92121, USA.

Abstract. We present a genomic map of infectious laryngotracheitis virus (ILT) and an 18,912 bp sequence containing the entire unique short region and a portion of the flanking short repeats. In determining the genomic map, an 856 bp region repeated as many as 13 times was identified within the short repeats. The unique short sequence contains nine potential open reading frames (ORFs). Six of these ORFs show homology to other known herpesvirus unique short genes. Using the herpes simplex virus nomenclature, these genes are the US2, protein kinase, and glycoproteins G, D, I, and E (ORF 1, 2, 4, 6, 7, and 8, respectively). Interestingly, an open reading frame with homology to HSV-1 UL47 (ORF 3) is found in the unique short. One very large open reading frame (ORF 5) is present and contains a threonine-rich, degenerate repeat sequence. This gene appears to be unique to ILT among sequenced herpesviruses. Two ORFs were identified within the short repeat (SR) region. SRORF 1 is homologous to a gene (SORF3) found in the unique short region in both MDV and HVT, and appears to be specific to avian herpesviruses. SRORF 2 has homology to HSV US10.

Key Words: ILTV, herpesvirus, viral genome, sequence

Introduction

Infectious laryngotracheitis virus (ILT) is a herpesvirus that causes a respiratory illness of varying virulence in chickens. Live attenuated ILT vaccines are available to protect against the disease, but several reports have implicated vaccine viruses in the possible recurrence and spread of the disease (1,2), limiting vaccination to use in uninfected birds early in an outbreak. In order to design a more efficacious, attenuated vaccine, we have studied the genomic organization of the ILT virus.

The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned the accession number U28832.

ILT is classified as an alpha herpesvirus with a type D genome (3) composed of a unique long region and a unique short region flanked by inverted repeats. A genomic restriction map of an Australian ILT isolate (SA-2) was described by Johnson et al. (4). Using this map, Guo et al. (5) isolated and sequenced a DNA fragment from the USDA challenge strain, which appeared to be derived from the unique short region. We have mapped the USDA challenge strain of ILT, and report the characteristics of 11 putative genes present in the short region. Our data indicate that the sequence identified by Guo et al. (5) is part of the short repeat sequence and is not from the unique short. Other reports (6,7) describe the sequences of two genes, one homol-

ogous to PRV gG and the other unlike other reported herpesvirus genes. These two genes were mapped to the unique long region of SA-2. However, these sequences are identical to sequences we have identified as being from the unique short region. Our data indicate that the overall organization of the short region of ILT is similar to other herpesviruses.

Materials and Methods

Growth of Chicken Kidney Cells and ILT Virus

An ILT virus, designated fowl laryngotracheitis challenge virus, lot number 83-2, was obtained from National Veterinary Services Laboratories in Ames, Iowa. ILT viruses were grown in primary chicken kidney cells (CK) from 6 to 9-day-old pathogen-free chicks (Hy-Vac). Trypsinized cells were plated in growth media (GM) consisting of Eagle's basal medium (modified) with Hank's salts, 10% binary ethyleneimine-treated fetal bovine serum (FBS), 2 mM glutamine, 200 units/ml penicillin, 200 mg/ml streptomycin, and 8.9 mM sodium bicarbonate (8). After 24 hr at 39°C, cells were switched to maintenance media (GM with 1% FBS). CKs were inoculated with ILT at a 0.01–0.1 multiplicity of infection, and viral stocks were harvested 4–5 days later by scraping and sonicating. Titers were typically 10^5 to 10^6 PFU/ml.

Preparation of Viral DNA

Cells and media from infected flasks were pelleted at 1700 g for 5 min at 4°C. The supernatant and cell pellet were treated separately. Virion particles were centrifuged out of the supernatant at 23,500 g for 30 min. The original cell pellet was resuspended in 1 ml/flask of 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 1.5 mM $MgCl_2$, and was incubated for 15 min at 4°C. To this was added 25 μl/flask of 20% NP40, and the mixture was dounce homogenized using an A pestle. The preparation was centrifuged at 1700 g for 10 min at 4°C, and the supernatant was retained and the pellet discarded. To the supernatant was added (per original flask) 10 μl of 0.5 M EDTA, 50 μl of 20% $NaDodSO_4$, and 25 μl of 10 mg/ml proteinase K. This mixture was used to resuspend the pellet of viral particles obtained by high speed centrifugation of the first supernatant. The mixture was treated at 65°C for 1 hr, and was phenol-extracted and ethanol precipitated. The resuspended DNA pellet was treated with pancreatic RNase A, followed by phenol extraction and ethanol precipitation.

Creation of the Cosmid Library

The cosmid library of ILT DNA was created following the protocol of van Zijl et al. (9). Briefly, 35–50 kb fragments of sheared ILT DNA were isolated from glycerol gradients, the ends were repaired using T4 and Klenow polymerases, and the fragments were ligated into cosmid pSY1626. Cosmid pSY1626 was made by first removing the *Ava*I to *Hin*dIII fragment from pHC79 (BRL) containing the tetracycline gene. The remaining fragment and the *Eco*RI-digested polylinker from pWE15 (Stratagene) were filled in with Klenow polymerase and ligated together. The resulting cosmid vector, pSY1005, was modified at the *Eco*RI site to create pSY1626 by blunt-ended insertion of a 1.5 kb *Hin*dIII-*Bam*HI fragment from pNEO (P-L Biochemicals) containing the kanamycin resistance gene. PSY1626 was cut and made blunt at the *Bam*HI site for use as the cosmid vector. The ligation mixture was packaged using Gigapack XL (Stratagene) according to the manufacturer's directions. Colonies were selected on LB plates containing kanamycin.

Sequencing

Manual sequencing was performed using ^{35}S-dATP (NEN) with the BRL Sequenase kit. Reactions using both dGTP and dITP mixes were performed to clarify areas of compression. Alternatively, compressed areas were resolved on 8% acrylamide gels that were 40% formamide. Automatic fluorescence sequencing was performed using an Applied Biosystems (ABI) 373A DNA Sequencer. Subclones were made to facilitate sequencing. Internal primers were synthesized on an ABI 392 DNA synthesizer. The sequence was obtained for both strands with an average redundancy of five and was assembled using DNAstar software. Manipulation and comparison of se-

quences were performed with DNAstar programs, and with Superclone and Supersee programs from Coral Software. Comparisons with GenBank (release 85) were performed at the National Center for Biotechnology Information using the BLAST network service (10).

Results

A cosmid library of the ILT genome was created to facilitate restriction endonuclease mapping. Forty-three overlapping cosmids were analyzed by digestion with *Asp*718I and *Not*I. *Asp*718I was known to cut the genome relatively infrequently (11), and we had found that *Not*I cuts the genome less than 10 times. Comparison of these cosmid digests allowed the order of the *Asp*718I fragments covering 85% of the ILT genome to be determined (Fig. 1). On the long end of the genome, seven cosmids were identified that all contained a *Not*I site 0.9 kb from the end of the cloned insert; all other cosmid inserts had heterogeneous ends from shearing. This 0.9 kb fragment was used as a probe (P1 in Fig. 1) to genomic ILT digested with *Asp*718I, *Not*I, or *Bam*HI; the sizes of the genomic fragments that hybridized were identical to the size of the fragments excised from the cloned cosmid insert, indicating that the cloned insert extended all the way or very close to the end of the unique long (data not shown). The 0.9 kb fragment did not hybridize to other bands in the ILT digest, consistent with previous reports that this virus resembles PRV and contains no long repeat (4). Once the cosmid clones were ordered, the restriction sites for a more frequent cutting enzyme, *Bam*HI, were mapped.

Our map indicated that the cosmid library did not include clones from the unique short portion of the genome. Cosmids spanning the unique short region of HVT (12) and PRV (9) have been found to be underrepresented in cosmid libraries. We therefore compared the *Asp*718I fragments found in the cosmid clones with an *Asp*718I digest of ILT and identified fragments of 8.0, 5.1, and 2.5 kb that were not represented in the cosmid library (Fig. 2). These fragments were cloned into plasmids and hybridized to each other and to ILT digested with *Bam*HI. Subsequent cloning, restriction endonuclease mapping, and hybridization of the fragments identified in this manner allowed the map of the entire unique short region and some of the flanking short repeat to be elucidated (Fig. 2). Subclones

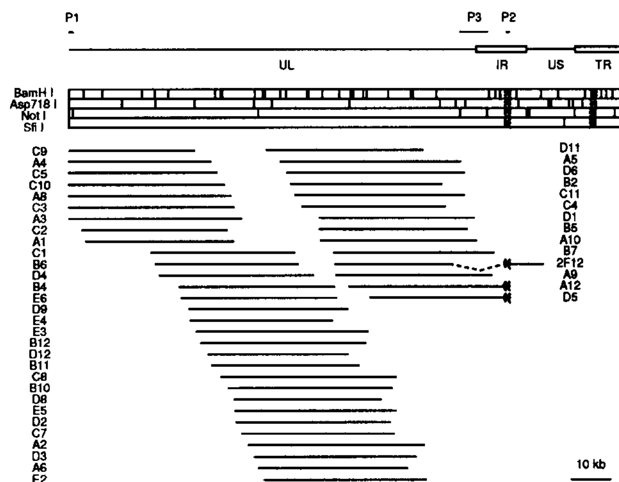

Fig. 1. Genome of the ILT virus, identifying the unique long (UL), unique short (US), internal repeat (IR), and terminal repeat (TR). The *Bam*HI, *Asp*718I, *Not*I, and *Sfi*I restriction maps of the virus are drawn underneath, with the highly repetitive region of the short repeats indicated by a set of wavy lines. The position of the cosmids used to determine the map of ILT are drawn beneath the restriction map. Note that cosmid 2F12 contains two noncontiguous sections. Three probes used to characterize the ILT genome are indicated as P1, P2, and P3. P1 is a 0.9 kb *Not*I fragment found at the terminus of the unique long region, P2 is the 856 bp *Hin*dIII fragment found in multiple copies within the short repeat, and P3 is a 6.6 kb *Not*I fragment used to identify the fragments at the end of the terminal repeat.

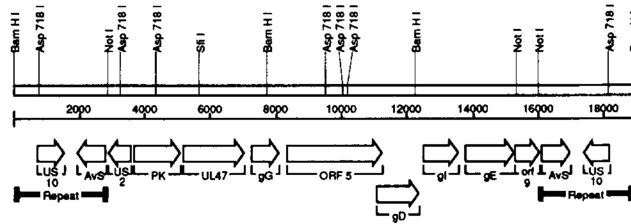

Fig. 2. The region sequenced and the positions of the *Asp*718I, *Bam*HI, *Not*I, and *Sfi*I sites. The extent and orientation of the open reading frames found in the ILT unique short and the flanking short repeat regions are indicated.

of this region were made, and the entire unique short region was sequenced.

To complete the genomic map, we searched for an *Asp*718I or *Bam*HI fragment that spanned the region between the short repeat sequences of the 8.0 or 2.5 kb *Asp*718I fragments mentioned earlier and the unique long region identified in the cosmid map. A 10 kb *Not*I fragment from the rightmost end of cosmid D5 (Fig. 1) was hybridized to genomic ILT digests on Southern blots. Surprisingly, ladders of hybridizing bands were seen when the enzymes *Bam*HI, *Not*I, and *Asp*718I were used (data not shown). The bands corresponding to these ladders were not generally visible in ethidium bromide–stained gels. Subsequent subcloning and mapping of the 10 kb D5 fragment indicated that it contained up to five repeats of an 856 bp segment, and that the cosmid insert ended within a repeat motif. *Hin*dIII, which cuts once within the repeat, was used to clone the 856 bp fragment. When this fragment (Fig. 1, P2) was used to probe ILT digested with *Sfi*I, *Not*I, *Asp*718I, and *Bam*HI, ladders of hybridization were again seen (Fig. 3). These ladders arise from varying numbers of the 856 bp repeat in different viral molecules. *Sfi*I cuts only once in this ILT strain, and a ladder at very high molecular weight can be seen. *Not*I and *Asp*718I cut further away from the repeat, generating ladders beginning at 10.5 or 12 kb, respectively. The *Asp*718I digest should generate two overlapping ladders, because one fragment is bounded by an *Asp*718I site in the unique long, while the other is bounded by the end of the TR_S. In contrast, only one ladder should be generated by the *Not*I digest. Comparison of Fig. 3 lane c (*Not*I) with lane d (*Asp*718I) does suggest that in lane d a second ladder is superimposed on the first, starting somewhat higher. *Bam*HI cuts close to the repeated region, and a ladder beginning at 3.4 kb is found. *Hin*dIII cuts within the repeat and generates a strongly hybridizing 856 bp band, as well as the two flanking *Hin*dIII fragments of about 1.1 and 2.5 kb, which each contain a portion of the repeated sequence.

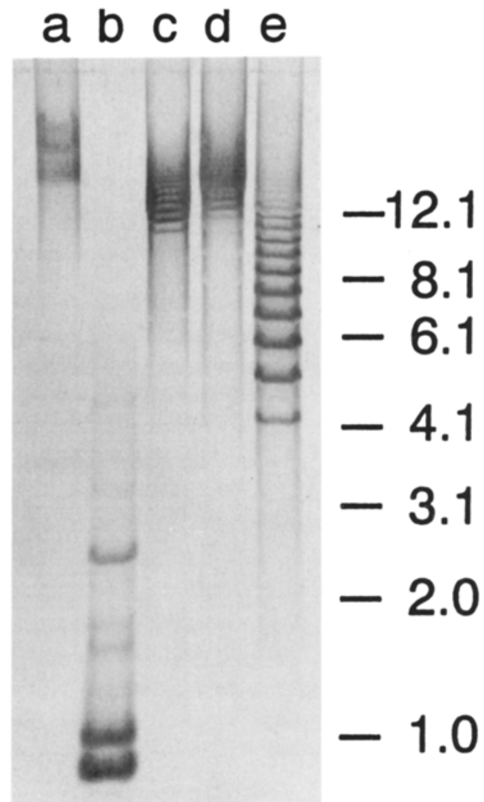

Fig. 3. Southern blot showing the repetition of an 856 bp element within the short repeat. Genomic ILT DNA digested with *Sfi*I (*a*), *Hin*dIII (*b*), *Not*I (*c*), *Asp*718I (*d*), or *Bam*HI (*e*) was probed with an 856 bp *Hin*dIII fragment from the short repeat. Positions of molecular weight markers are indicated.

The presence of the 856 bp repeat accounted for the occasional observation of very fine submolar bands in ethidium bromide–stained *Asp*718I digests. It also accounted for the lack, in ethidium bromide–stained gels, of a molar or half-molar quantity *Asp*718I or *Bam*HI band greater than 10 kb, which we expected to span this region based on our analysis of the cosmid clones. Instead, because of the presence of the 856 bp repeat, this band exists as many submolar bands comprising the ladder. As can be seen in the *Bam*HI digest, there can be 13 or more repeats of the region. The relationship of the 856 bp repeat to the surrounding sequence is depicted in Fig. 4.

To identify the remainder of the short repeat from the 856 bp repetitive region to the *Bam*HI fragments used for sequencing the unique short, we used the 8.0 kb *Asp*718I fragment containing part of the short repeat as a probe to a second cosmid library of ILT. One cosmid, clone 2F12, hybridized to the probe. Restriction endonuclease analysis of 2F12 and comparison to the cosmid map indicated that it was not a single contiguous cosmid but was composed of two large noncontiguous fragments (see Fig. 1). The break in the rightmost fragment was within a repeat of the 856 bp region. This fragment included at least two 856 bp repeats and extended 4.6 kb through the remainder of the short repeat into the unique short.

To identify the end of the TR_s, we used the 6.6 kb *Not*I fragment spanning the unique long and the short internal repeat (IR_s) (P3 in Fig. 1) as a probe (Fig. 5). Hybridization of a *Not*I digest of ILT with P3 identified the 6.6 kb band corresponding to the probe and a 2.9 kb band repre-

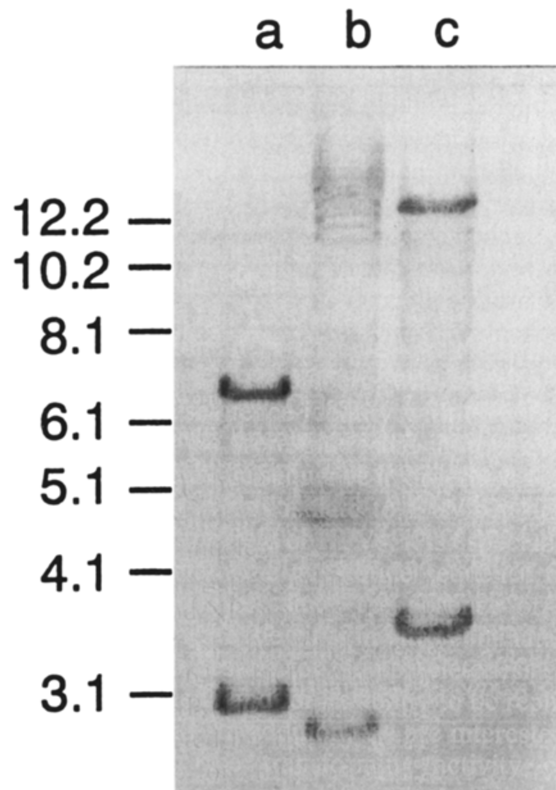

Fig. 5. Southern blot identifying fragments from the internal and terminal repeat that hybridized to a 6.6 kb *Not*I fragment containing the junction of the unique long and the internal repeat. Genomic ILT DNA digested with *Not*I (a), *Asp*718I (b), and *Bam*HI (c) was probed with the 6.6 kb *Not*I fragment. Positions of molecular weight markers are indicated.

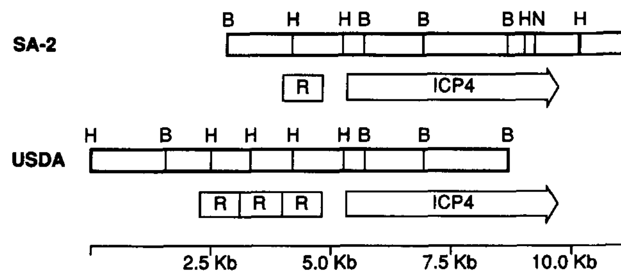

Fig. 4. Depiction of the position of the 856 bp repeat region in the USDA strain, compared with the same region from the SA-2 strain described by Johnson et al. (24) Three repeats are arbitrarily shown in the USDA strain, and the region is not repeated in SA2. B = *Bam*HI; H = *Hin*dIII; R = 856 bp repeat.

senting the end of the TR_S. In the *Bam*HI digest, the predicted 13 kb fragment containing a portion of the IR_S and a 3.5 kb fragment corresponding to the end of the TR_S are evident. In the *Asp*718I digest, an overlapping 2.7 kb fragment from the unique long hybridizes, and the high molecular weight ladder described previously is seen.

Sequencing of the ILT unique short and flanking region indicated that the overall GC content of the area sequenced was 55.6%, with a higher GC content, 62%, in the portion of the short repeat region sequenced, and a slightly lower GC content, 53%, attributable to the unique short region. A total of 2909 bp of short repeat were sequenced on both sides of the unique short; the unique short region itself is 13,904 bp long. Nine open reading frames were identified in the unique region and two (duplicated) in the repeat region, as diagrammed in Fig. 2. Comparison of the proteins encoded by these ORFs to the GenBank database demonstrated homology for most of the potential proteins with other known herpesvirus gene products. Table 1 summarizes the closest homology found for each gene and gives the probability score for each match as generated by the search program. ORF 2, the protein kinase (PK) gene, is the most highly conserved of the ILT ORFs to its herpes homologues. In contrast, the glycoprotein genes are less conserved. It should be noted that portions of the sequences of the ILT protein kinase, gG, and ORF 5 genes have been published (6,7,13), although these genes were identified as being in the unique long region. A description of each of the nine unique short genes and the two genes in the flanking short repeat follows.

Table 1. Table 1 indicates the ORFs of the ILT unique short and the HSV nomenclature for these genes, in those cases where homology is found. The third column shows the best matches from the Blast homology search, and the probability scores assigned by the program for the matches indicated. Smaller numbers indicate less likelihood that the match could occur randomly.

ORF	HSV Homologue	Best Matches	Blast Score
1	US2	EHV1 EUS1	3.1×10^{-13}
2	PK	MDV PK	8.2×10^{-36}
3	UL47	HSV1 UL47	6.0×10^{-1}
4	gG	PRV gG	5.3×10^{-5}
5	ORF 5	EHV1 EUS5	1.9×10^{-45}
		Human mucin	1.1×10^{-25}
6	gD	MDV gD	6.8×10^{-4}
7	gI	VZV gI	4.2×10^{-2}
8	gE	SHV SA8 gE	1.7×10^{-6}
9	ORF 9	EBV BLRF2	5.7×10^{-1}
SR1	no HSV homologue	MDV "SORF3"	4.8×10^{-4}
		HVT "SORF3"	2.6×10^{-1}
SR2	US 10	EHV4 US10	1.2×10^{-1}

The first open reading frame in the unique short encodes a 229 amino acid (aa) protein with homology to other herpesvirus US2 proteins. The coding sequence of the gene ends just within the unique short region, and a potential poly-A addition site is found 115 bases downstream in the short repeat. Two possible TATA promoters are found 37 and 70 bases upstream from the initiation codon.

ORF 2 encodes a protein kinase of 476 aa with strong homology to many other herpesvirus protein kinases and to cellular protein kinases. The organization of the US2 and PK genes, with their 5′ ends close together and their promoters possibly overlapping, is similar to that found in other herpesviruses. Two TATA sequences are present 14 and 49 bases upstream of the PK start codon, and two polyadenylation signals are found, one immediately after the stop codon and one 50 bases downstream.

ORF 3 encodes a 623 aa protein with similarity to the herpes simplex virus UL47 gene. The program comparing this protein with other UL47 proteins projects a poor probability score for this homology. However, at least one of the regions of identity between ILT and HSV UL47 corresponds to a region that is conserved among other herpesvirus UL47 homologues, suggesting that this homology is significant (Fig. 6). Additionally, it should be noted that equally poor probability scores for homology generated by comparisons of the gG or gI genes are also seen for certain homologue pairings, suggesting that these scores are not sufficient for determining homology. It is interesting that the ILT UL47 gene,

```
ILT 277 QHGPMAAVFRNAGAGLFLWPAMRAAFEERDKRLLRACLSSLDIMDAAVLASF
HSV 351 QSGPDAAVFRSSLGSLLYWPGVRALLDRDCRVAARYAGRMTYLATGALLARF
EHV 531 LRTPNSAVFRAFFGSLVYWAELRLALRDPASINCRYVGFHLQTSEIYLLARA
MDV 472 MRDPMASAARASYGSLAYWPELRCALGSENKRIVRYAIVAMLQAEIYLLTRI
```

Fig. 6. Relationship of herpesvirus UL47 proteins to each other and to the ILT UL47 homologue in a conserved region. Amino acids shared between ILT UL47 and the other UL47 proteins are in boldface type. Pairwise comparisons have been made between the sequences as shown. A vertical bar indicates an identical amino acid, two dots indicate a positive probable acceptable mutation rate, and one dot indicates a neutral probable acceptable mutation rate (29).

normally found in the unique long region of other herpesviruses, may have been transposed into the unique short in ILT.

The fourth open reading frame encodes a 292 aa glycoprotein homologous to PRV gG. Four N-linked glycosylation sites with the consensus sequence NXT or NXS are present. The protein has a signal sequence of 26 aa, which could be cleaved at G/AP, but lacks a transmembrane anchor. Analysis of cell lysates and media with a peptide antiserum to ILT gG indicates that this protein is modified and secreted, similar to other herpesvirus gG homologues (data not shown). This gene has a consensus TATA sequence 83 bases upstream from the ATG start and has two potential polyadenylation sites 73 and 166 bases downstream from the stop codon.

ORF 5 could encode a protein of 985 aa. A hydrophobic signal sequence is found at the amino terminus, and a hydrophobic sequence is present at the carboxy terminus. Nine glycosylation sites are found, suggesting that this is a glycoprotein. ORF 5 contains an imperfect repeat, consisting of 30–36 bp repeated approximately 23 times from amino acid 431 to amino acid 677. The hydrophilic amino acid consensus sequence created by this repeat is FTQTPSTEPET/A. Comparison of ORF 5 with other herpesvirus sequences (Table 1) shows similarity to the glycoprotein product from the equine herpesvirus 1 US5 gene (EUS5, ref. 14). The low probability score for this identity arises primarily from the fact that both genes contain threonine-rich repeats. It is not clear whether this reflects homology in form, function, or both. Both of these threonine-rich genes, EUS5 and the ILT ORF 5, are large, have similar positions among flanking genes in the unique short, have signal sequences, and encode glycoproteins, but other sequence similarities are not seen. Interestingly, the ORF 5 repeat region shows similarity to mucin genes, which also contain threonine-rich repeats, but whether this reflects a similarity in function of the encoded proteins is unclear. A TATA sequence is found 560 bases upstream of the start codon; the nearest consensus polyadenylation signal is at the end of the gI gene. This suggests that the ORF 5 transcript may be coterminal with the gD transcript.

The open reading frame for the gD homologue (ORF 6) overlaps the end of ORF 5. Four in-frame methionines are found within the first 58 amino acids of the open reading frame, and it is not clear which is the actual translational start codon. Comparison of the sequences surrounding the four ATGs with the eukaryotic translational initiation consensus sequence A/GCCATGG (15) suggests that the latter two ATG codons may be preferred translational start sites. The protein sequences derived from each of these starts were examined for the presence of eukaryotic signal sequences and signal cleavage sites. A start at aa 58 within the ORF would result in a signal peptide of 26 aa with a predicted cleavage site between two alanine residues. This same signal sequence would be positioned much further from the amino terminus and embedded in a more hydrophilic sequence if the other start sites were used. We have tentatively assigned the start of ILT gD to position 58, which would result in a protein 377 aa long. Of course, it is possible that more than one initiation codon is used in vivo. Experiments of Zelnik et al. (16) suggest that alternate in-frame ATG codons are used to initiate MDV and HVT gD transcription in vitro, although the in vivo situation was not addressed. Additional experiments on gD transcription and translation in ILT are necessary to identify its translational start codon.

The ILT gD homologue has a secretory signal sequence and a transmembrane helix (aa 352–372) at the carboxy terminus. Only one potential glycosylation site is found at position 250–252; this is of the form NPS and may not be glycosylated due to the proline residue. There is some question, therefore, as to whether processed ILT gD contains N-linked oligosaccharides. This would be similar to the gD homologue in pseudorabies virus, gp50, which lacks N-linked glycosylation sites (17). As in other herpesviruses, the gD coding sequence lacks a poly-A addition signal immediately following the gene, and the closest signal is at the end of the gI gene.

The seventh open reading frame encodes a protein of 362 aa and is most homologous to varicella zoster virus glycoprotein I. The encoded protein shows all the characteristics of related gI glycoproteins, including a signal sequence with a potential cleavage site at positions 22 and 23 between a glycine and an isoleucine, a trans-

membrane helix at the carboxy terminus from 272 to 292, and four possible N-linked glycosylation sites. A TATA sequence is present 51 bases upstream from the methionine start codon. Two possible poly-A addition signals are found within the coding sequence for ILT gI, and may be the signals used by the gD and ORF 5 transcription units upstream.

The gE gene (ORF 8) follows the gI. This gene is 499 aa long and contains four N-linked glycosylation sites. A signal sequence of 18 aa is present, and there are two, and possibly three, membrane-associated helices in the carboxy-terminal portion of the protein. The gE gene has a TATA box 86 bases upstream of the start codon, and a potential poly-A addition signal just before the 3′ end of the coding region. This may serve as the polyadenylation site for the gI gene.

The ninth open reading frame extends across the junction of the unique short and the short repeat, and could encode a protein of 260 aa. This protein has no signal sequence or membrane anchor but has one possible N-linked glycosylation site. In a search of GenBank, some similarity is found between this protein and BLRF2 of EBV, but the significance of this similarity is unknown. The poly-A addition signal in the short repeat may be utilized by this gene. A potential TATA sequence is found 178 bases upstream of the first ATG of this ORF.

The first open reading frame in the short repeat (SRORF 1) encodes a 294 aa protein that displays homology to the gene product of MDV SORF3 (18,19) and HVT ORF3 (20). In MDV and HVT, the corresponding gene is found as one copy in the unique short, and its function is unknown. No homology has been identified with mammalian herpesviruses; this gene appears to be specific to avian herpesviruses. MDV SORF3 has been deleted by Parcells et al. (21) and does not appear to be absolutely required for infection in chickens.

SRORF 2 encodes a protein of 278 aa with homology to other herpesvirus US10 genes. A zinc finger motif, found in the EHV-4 US10, is highly conserved in the ILT US10 (aa 201–218); this suggests that the ILT US10 gene is a DNA binding protein. Regulatory sequences include a poly-A addition signal 163 bp after the stop codon; it is unclear where the promoter for this gene resides.

Discussion

The organization of the genes in the short region of ILT is similar to that seen in other herpesviruses. Several genes encoding glycoproteins are present, and the order of these genes is similar to that seen in equine herpesvirus 1, particularly with respect to ORF 5. Similarities to avian herpesviruses are also evident in the presence of the avian-specific gene, SRORF1, and its position relative to US2 and PK, though it differs from HVT and MDV in that it is in the short repeat and is duplicated, also appearing downstream from the ORF 9 gene. The PK gene itself has the most homology to MDV and HVT PK genes; however, other genes are found to be more like their homologues in diverse herpesviruses, such as EHV, PRV, and SHV SA8. Unusual characteristics of the ILT unique short are the inclusion of a gene normally found in the unique long, the UL47 homologue, and the presence of the unique gene, ORF 5, which contains a set of degenerate repeats.

The gene encoded by ORF 5 contains threonine-rich, degenerate repeats. These are similar in composition and in their repetitive nature to repeats found in mucin genes. This repeated region in mucin is modified by O-linked oligosaccharides and is highly hydrophilic. It is interesting to speculate on what the function of this somewhat similar region might be in infection if it is expressed in toto in ILT. At least a portion of this gene is known to be expressed, as Kongsuwan et al. (6) cloned and sequenced a fragment from it by probing a lambda gt11 library with a monoclonal antibody that was known to bind to a 60 kD ILT protein (g60) on Western blots (22). The relationship of such a 60 kD protein to the predicted 985 aa product from ORF 5 is unknown. Comparison of our sequence with the complete sequence of the g60 coding region (13) shows a 98.5% homology between the SA-2 strain and the USDA strain. Interestingly, there is an insertion of a block of 10 aa in g60 relative to the ORF 5 protein; this difference reflects one additional degenerate repeat sequence in the SA-2 strain.

As mentioned earlier, Kongsuwan et al. (7) described an ILT gene that encoded a 32 kD protein with similarity to PRV gG. A comparison of our ILT gG protein sequence with their 32 kD

protein found 10 aa differences in the first 273 residues of the protein. At amino acid 274, a deletion of one base pair in SA-2 relative to the USDA strain created a frame shift, such that 19 additional residues were found in the challenge strain as opposed to 26 in SA-2. A peptide made from our carboxy-terminal sequence elicited antisera in mice that reacted with ILT gG (data not shown); this indicates that our sequence reflects the actual carboxy terminus in the USDA strain. A similar situtation was found when our ILT gD protein was compared with the ILT gD sequence of Johnson et al. (23). Ten differences were found in the first 419 aa, after which a deletion of a base in the SA-2 strain relative to ours caused the predicted carboxy termini to differ, with 15 more amino acids in the USDA strain and 9 in SA-2. These differences could arise from errors introduced during cloning and sequencing of these genes. It is also possible that the carboxy termini of the ILT gG and gD genes are variable between these strains.

The presence of an 856 bp repetitive motif in the short repeat region of the USDA challenge strain of ILT has not been previously noted. Comparison of the motif to the ILT ICP4 sequence of Johnson et al. (24) indicated that it corresponded (99% identity) to nucleotides 1140–1996 of their sequence, which is a region just upstream of the ICP4 homologue (see Fig. 4). Only one copy of this region is found in the SA-2 strain sequence, however. Restriction digests indicate that the region to the right of the repeat as shown in Fig. 4 is similar between these two strains; however, the position of the *Bam*HI site indicated to the left of the repeat differs between them. It is unclear why this particular region is repeated in the USDA challenge strain. The 856 bp motif is slightly less GC rich (46%) than the rest of the sequenced repeat region, and does not appear to be part of a protein coding region. Comparison of the sequence with the consensus sequence for herpesvirus replication origins described by Baumann et al. (25) shows no similarity.

Our analysis of the structure of ILT disagrees with previous reports. Comparison of our sequence with those of the Australian ILT isolate SA-2 indicates that a 32 kD protein described by Kongsuwan et al. (7) is almost identical to our gG, and the sequenced fragment of the g60 protein presented by Kongsuwan et al. (6) is part of our ORF 5 gene. However, they identified the 5 kb *Asp*718I fragment containing both of these genes as coming from the unique long region of SA-2 (4). Recently, Guo et al. (5) reported the sequence of a region from the USDA challenge strain that they ascribed to the unique short on the basis of comparison to the map presented by Johnson et al. (4). No identity was found between this sequence and the unique short sequence we describe. Instead, the sequence described by Guo et al. (5) shows 98% identity to the ILT ICP4 sequence (24). The *Bam*HI sites within the ICP4 coding region generate two contiguous fragments of 1.2 and 1.7 kb (see Fig. 4). In our map, two contiguous *Bam*HI fragments of this size are found within the short repeats (see Fig. 1). In addition, we mapped the 856 bp repeat element, found upstream of the ICP4 gene (see Fig. 4), within the short repeats. This indicates that the ICP4 gene in our strain is present in the IR_s and the TR_s. It is possible, but unlikely, that the Australian SA-2 vaccine strain underwent an unusual rearrangement that altered the relationship of the unique long, unique short, and short repeats. However, Guo et al. (5) used the same challenge strain we mapped, and the sequence they reported is not in the unique short, but in the short repeats, similar to the ICP4 genes of other herpesviruses.

The appearance of the 856 bp derived hybridization ladder is reminiscent of defective interfering particles, but we do not believe this represents a case of defective interfering particles in our viral stock. Our ILT viral stocks were passaged at low multiplicity, and viral stocks originating from a single picked plaque exhibited similar ladders of hybridization. In restriction digests, we do not encounter digest fragments that cannot be accommodated in the linear viral map, or restriction fragments in higher than molar quantities, as might be expected from defective interfering particles (26). From these considerations we conclude that varying numbers of 856 bp units are present in the short repeats of standard viral DNA from the USDA challenge strain of ILT. Since fragments exist that contain 13 or more repeats of the region, genomic DNA from ILT could vary by over 11 kb in the short repeat regions. Repetitive regions have been identified in other herpesviruses; for example,

Marek's disease virus contains a 132 bp repetitive sequence in the long repeat regions (27,28), and expansion of this repeat is associated with reduction of viral oncogenicity. The presence of the 856 bp tandem repeats in ILT, in contrast, does not appear to affect viral pathogenicity, since this strain does cause severe clinical disease in chickens. It would be interesting to examine other ILT strains for the presence of this repeat.

Acknowledgments

This work was funded by Hoechst-Roussel Agri-Vet Company.

References

1. Leib D.A., Bradbury J.M., Gaskell R.M., Hughes C.S., and Jones R.C., Avian Dis *30,* 835–837, 1986.
2. Hughes C.S., Williams R.A., Gaskell R.M., Jordan F.T.W., Bradbury J.M., Bennett M., and Jones R.C., Arch Virol *121,* 213–218, 1991.
3. Roizmann B., Desrosiers R.C., Fleckenstein B., Lopez C., Minson A.C., and Studdert M.J., Arch Virol *123,* 425–449, 1992.
4. Johnson M.A., Prideaux C.T., Kongsuwan K., Sheppard M., and Fahey K.J., Arch Virol *119,* 181–198, 1991.
5. Guo P., Scholz E., Maloney B., and Welniak E., Virology *202,* 771–781, 1994.
6. Kongsuwan K., Johnson M.A., Prideaux C.T., and Sheppard M., Virus Genes *7,* 297–303, 1993.
7. Kongsuwan K., Johnson M.A., Prideaux C.T., and Sheppard M., Virus Res *29,* 125–140, 1993.
8. Wark M.C., Tannock G.A., and Pye D., J Biol Standard *7,* 73–80, 1979.
9. van Zijl M., Quint W., Briaire J., de Rover T., Gielkens A., and Berns A., J Virol *62,* 2191–2195, 1988.
10. Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J., J Mol Biol *215,* 403–410, 1990.
11. Guy J.S., Barnes H.J., Munger L.I., and Rose L. Avian Dis *33,* 316–323, 1989.
12. Reilly J.D. and Silva R.F., J Virol Methods *41,* 323–331, 1993.
13. Sheppard M.G., Prideaux C., Johnson M., Fahey K.J., York J.J., and Kongsuwan K., International Patent publication no. WO92/03554, 1992.
14. Telford E.A.R., Watson M.S., McBride K., and Davison A.J., Virology *189,* 304–316, 1992.
15. Kozak M., Nucleic Acids Res *15,* 8125–8148, 1987.
16. Zelnik V., Ross N.L.J., and Pastorek J., J Gen Virol *75,* 2747–2753, 1994.
17. Petrovskis E.A., Timmins J.G., Armentrout M.A., Marchioli C.C., Yancey R.J. Jr., and Post L., J Virol *59,* 216–223, 1986.
18. Velicer L.F., Brunovskis P., and Coussens P.M., International Patent publication no. WO92/03547, 1992.
19. Sakaguchi M., Urakawa T., Hirayama Y., Miki N., Yamamoto M., and Hirai K., Virus Genes *6,* 365–378, 1992.
20. Zelnik V., Darteil R., Audonnet J.D., Smith G.D., Riviere M., Pastorek J., and Ross L.J.N., J Gen Virol *74,* 2151–2162, 1993.
21. Parcells M.S., Anderson A.S., Cantello J.L., and Morgan R.W., J Virol *68,* 8239–8253, 1994.
22. York J.J., Sonza S., Brandon M.R., and Fahey K.J., Arch Virol *115,* 147–162, 1990.
23. Johnson M.A., Tyack S.G., Prideaux C.T., Kongsuwan K., and Sheppard M., DNA Seq 5, 191–194, 1995.
24. Johnson M.A., Tyack S.G., Prideaux C.T., Kongsuwan K., and Sheppard M., Virus Res *35,* 193–204, 1995.
25. Baumann R.P., Yalamanchili V.R.R., and O'Callaghan D.J., J Virol *63,* 1275–1283, 1989.
26. Rixon F.J. and Ben-Porat T., Virology *97,* 151–163, 1979.
27. Fukuchi K., Tanaka A., Schierman L.W., Witter R.L., and Nonoyama M., Proc Natl Acad Sci USA *82,* 751–754, 1985.
28. Maotani K., Kanamori A., Ikuta K., Ueda S., Kato S., and Hirai S., J Virol *58,* 657–660, 1986.
29. Dayhoff M.O., Barker W.C., and Hunt L.T., Methods Enzymol *91,* 524–545, 1983.