The complete unique long sequence and the overall genomic organization of the GA strain of Marek's disease virus

Lucy F. Lee*†, Ping Wu*, Dexin Sui*, Delin Ren*, Jeremy Kamil‡, Hsing Jien Kung‡, and Richard L. Witter*

*United States Department of Agriculture-Agricultural Research Service, Avian Disease and Oncology Laboratory, East Lansing, MI 48823; and ‡University of California Davis Cancer Center, Sacramento, CA 95810

Contributed by Richard L. Witter, March 13, 2000

We have determined the DNA sequence of the unique long (UL) region and the repeat long (RL) region in the genome of serotype 1 GA strain of Marek's disease virus (MDV), a member of the α-herpesvirus family. With this information, the complete nucleotide sequence of GA-MDV is now known. The entire GA-MDV genome is predicted to be about 174 kbp in size, with an organization of TRL-UL-IRL-IRS-US-TRS, typical of a α-herpesvirus. The UL sequence contains 113,508 bp and has a base composition of 41.7% G + C. A total of 67 ORFs were identified completely within the UL region, among which 55 are homologous to genes encoded by herpes simplex virus-1. Twelve of them are unique with presently unknown functions. The sequence of RL reported here together with those published earlier reveal the major structural features of the RL. Virtually all of the ORFs encoded by RL are specific to serotype I of MDV. These ORFs are likely to contribute to some of the unique biological properties of MDV. Among the proteins encoded by MDV-specific ORFs are Meq, a jun/fos family of transcriptional factor implicated in transformation and latency, virus-encoded interleukin-8, a CXC chemokine, and pp38 and pp24, two phosphoproteins with undefined functions. There is also a putative lipase gene (LORF2) that has homologies in HPRS-24 (serotype II) strain of MDV and in various avian adenoviruses. An additional unique feature of MDV is the presence of long terminal repeat remnant sequences of avian retrovirus reticuloendotheliosis virus. These remnant sequences are derived from the U3-enhancer region through ancestral insertions by reticuloendotheliosis virus proviruses.

Marek's disease virus (MDV) is an oncogenic herpesvirus, which causes a highly contagious neoplastic disease in chickens (1, 2). Marek's disease is characterized by the development of T cell lymphomas, neurological disorders, immunedeficiency (3), and for some strains, atherosclerosis (4). This disease can be successfully prevented by vaccination with antigenically related nonpathogenic or attenuated virus strains (3). Three serotypes of MDV can be recognized by respective mAbs raised against these viruses (5). The oncogenic MDV, the prototype of this group, is designated as serotype 1. Serotype 2 and 3 designate nonpathogenic but antigenically related herpesviruses from chickens and turkeys, respectively. The exceptionally short latency of the MDV-induced lymphomas makes it a valuable model to study herpesviral oncogenesis and to define viral genes involved in T cell transformation. A comparative analysis of the genome structure and sequences between the oncogenic and vaccine strains would be most valuable in identifying genes that are responsible for the pathogenic phenotypes of the virus. The genomic structure of MDV is similar to that of herpes simplex virus (HSV) and consists of a long and short region, each flanked by inverted repeat sequences and terminal repeats (6). Early work, based on partial DNA sequences of MDV, confirmed that its genome is collinear with and closely related to that of α-herpesviruses (7). This is surprising, as MDV exhibits biological properties more closely resembling γ-herpesviruses. This provides impetus to identify genes unique to MDV, which may be responsible for the lymphotropic and oncogenic phenotypes. The sequence of the unique short and flanking repeat regions has been published (8, 9), as have the sequences containing genes of potential interest (10–15). Yet before this work, <25% of the MDV genome sequence was identified. In this paper, we present the complete unique long (UL) sequence of the GA strain of MDV. In addition, we have determined the sequence of a portion of the repeat long (RL) region, which was previously uncharacterized. With this new information, the entire GA strain of the MDV sequence is completed. Furthermore, we present a preliminary analysis of the organization and function of MDV genes, some of which are homologs of genes present in other herpesvirus whereas others are unique. This synthesis of new and previous sequence information should greatly facilitate progress in our understanding of this important disease and its causative virus.

Materials and Methods

Methods for DNA Sequence and Putative Gene Product Analysis. The GA strain of MDV (16) is a prototype strain representing the vMDV pathotype (17) and has been widely used for molecular and biological studies. *Bam*HI and *Eco*RI libraries of MDV GA strain (6, 18) were used to generate the original DNA sequences. The junctions between *Bam*HI fragments were amplified by using PCR. DNA sequencing was performed on double-stranded plasmid by the dideoxy chain termination method using [^{35}S]dATP (New England Nuclear, Life Science Products, Boston, MA) and the TAQuence version 2.0 DNA sequencing kits (United States Biochemical) as suggested by the manufacturer. Some fragments were sequenced by using an automated sequencer (373A DNA Sequencer, Applied Biosystems, Foster City, CA) and dideoxy-sequencing methods (Prism, Applied Biosystems). DNASTAR suite of programs (DNASTAR, Madison, WI) was used for data input and sequence assembly.

Initial analyses of the coding content of the genome and individual ORF were performed by using MacVector (Scientific Image Systems, New Haven, CT) and Gene Construction kit (Textco, West Lebanon, NH). The ORFs with a consensus start methionine and that were at least 100 aa in length were reported as putative proteins. These ORFs were translated by using DNASTAR suite. The putative protein sequences were down-

Abbreviations: UL, unique long; RL, repeat long; US, unique short; RS, repeat short; MDV, Marek's disease virus; REV, reticuloendotheliosis virus; HSV, herpes simplex virus; LTR, long terminal repeat; TRL, terminal repeat long; ICP, infected cell protein; LORF, ORF in UL region; SORF, ORF in US region; IL8, virus-encoded interleukin-8.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF147806).

†To whom reprint requests should be addressed. E-mail: leelu@pilot.Msu.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

loaded to the National Center of Biotechnology Information (NCBI) BLAST server for homologous search by BLASTP. Unless otherwise specified, default settings (without filter) were used for BLAST searches. The identity was determined and reported as "number of amino acids in a gene of MDV-1 identical to its homolog divided by the number of overlapped sequences between these two genes." The percentage identical value also was provided within the quotation.

Assignment of ORFs and Nomenclature. All of the ORFs with significant homology to HSV were named as "UL" after the HSV-1 counterpart name. Those genes without significant homology to HSV-1 were named as ORF plus number. They are further referred to as LORFs or R-LORFs depending on whether the start codon initiates within the UL region or the RL region, respectively (ORFs of the UL and RL regions). The numbering of the R-LORFs and LORFs begins from the left terminus of the known extent of the terminal repeat long (TRL) sequence and proceeds to the right. The complete UL sequence is deposited in GenBank with an accession no. of AF147806.

Results and Discussion

The Genomic Organization of MDV. The majority of the sequence data reported here is derived from the cloned fragments of a *Bam*HI-based and overlapping *Eco*RI-based genomic libraries of GA-MDV (6, 18, 19). PCR amplification of the genomic DNA was used to generate amplicons that encompass the junctions of these cloned fragments and to independently confirm the linkage map and the sequence data. With this strategy, the entire UL and the great majority of RL sequences were determined. Combining the new data generated by this study and those published previously (8–15), an overall genomic organization of GA-MDV can be constructed. The UL region of GA-MDV is 113,508 bp in length and the RL region is ≈12,584 bp. The sizes of the repeat short (RS) and unique short (US) regions, based on published work, are ≈12,121 bp and ≈11,160 bp respectively. These data still lack the *Bam*HI O2 fragment that covers the junction between the Rs and RL but taken together yield a total size of just more than 174 kbp for GA-MDV. This genome size is ≈20 kbp longer than HSV genome and is comparable in size to EBV genome. Although MDV is classified as a α-herpesvirus, the extra-coding capacities perhaps contribute to some of the γ-herpesvirus properties (see below). When one compares the overall structural features of GA-MDV and HSV-1, the major size differences are localized in the repeat regions. The RL of MDV is 30% larger than that of HSV-1 and the RS is twice as long. The organizations of UL genes in these two genomes are generally collinear with one another, excepting the presence of several ORFs unique to MDV. In the MDV genome, there are two stretches of sequences located at the boundaries of TRL/UL and IRL/UL, which encode MDV-specific ORFs. These are also areas where ancestral insertions by retrovirus reticuloendotheliosis virus (REV) long terminal repeat (LTR) were identified (20). Both the repeat and the UL/RL junction regions also show significant divergence among the different serotypes of MDV, underscoring the potential importance of these MDV-specific ORFs in conferring the oncogenicity of the viruses (GenBank accession no. AB024414) (21, 22). In the ensuing sections, we will first describe MDV genes that have homologs in HSV-1, followed by a more detailed discussion of MDV-specific ORFs.

MDV UL/RL Genes and ORFs. Table 1 summarizes genes and ORFs deduced by the sequence data of RL and UL. A map of the genetic organization of MDV based on this and other sequence data are shown in Fig. 1. MDV genes that are homologous to all of the HSV-1 genes are named after the UL genes of HSV. The degree of homology in the conserved regions between the MDV-I UL and those of serotype-2, and HSV-1, are also indicated in Table 1. MDV-specific ORFs are referred to as LORFs or R-LORFs depending on whether the start codon initiates within the UL region or the RL region, respectively (ORFs of the UL and RL regions). The numbering of the R-LORFs and LORFs begins from the left terminus of the known extent of TRL sequence and proceeds to the right. Based on the sequence data, there are 55 HSV-1 homolog genes and 28 unique ORFs. Given the evolutionary distance between MDV and HSV-1, it is striking that MDV carries all except one of the 56 genes encoded by the HSV-1 UL region. The HSV-1 homolog genes also appear to be in the same orientations and at comparable positions. It must be noted, however, that the extent of homology varies and the corresponding proteins may not necessarily function in exactly the same way. Of the 28 MDV unique ORFs, 12 were present in the unique region and 16 were present in the known extent of the RL region sequence. It must be noted that the known extent of each RL region carries only 14 unique ORFs. However, two of these ORFs, R-LORFs 13 and 14, have variances in their 3′-coding sequence depending on whether their coding sequence initiates within the TRL or the IRL, as their stop codons are located within the UL region. These two R-LORFs are thus counted twice and numbered, depending on if the copy is at the TRL or the IRL, as R-LORFs 13 or 13a and 14 or 14a, respectively. The nature of the UL genes with HSV-1 homologs and the UL LORFs and R-LORFs specific to MDV are described below.

UL Genes Involved in DNA Replication. UL genes involved in DNA replication are remarkably conserved between MDV and HSV (23, 24). Only UL30, which encodes MDV DNA polymerase, has been characterized in detail. This protein with a molecular weight of 135 kDa was found in all three serotypes of MDV and shares 54% identity and 70% similarity with that of HSV (25). Like HSV DNA polymerase, this enzyme is sensitive to phosphonate acetate, a potent inhibitor of viral replication (26). Other enzymes involved in nucleic acid metabolism such as uracil DNA glycosylase (UL2), subunits of ribonucleotide reductase (UL39 and UL40), thymidine kinase (UL23), and deoxyuridine triphosphatase (UL50) also are encoded by the MDV genome. Thus, the overall strategy of DNA replication appears to be very similar to HSV.

Replication Origins. HSV-1 contains three origins of DNA replication, two of which, referred to as OriS, are located in the RS and a third one, OriL, located in UL. MDV has two putative replication origins with a structure similar to HSV OriS. They are however located in RL rather than RS. These two putative replication origins have identical sequences, each containing an AT-rich hairpin loop, and three UL9-binding sites, similar to the HSV OriS replication origin. In addition, they serve as the divergent promoters for *pp38* (or *pp24*) and *Bam*HI gene family transcription (26), and carry c-myc and *Meq*-enhancer binding sites (27, 28). In HSV-1, OriL is located between UL29 (DNA-binding protein) and UL30 (DNA polymerase). Despite the sequence and positional conservation of these two genes in MDV, we failed to identify a structure similar to that of HSV OriL in the intergenic region of UL29 and UL30. BLAST search of the entire UL region did identify three consensus UL9-binding motifs, but they are located in the midst of UL36-coding sequence, without the accompanying AT-rich sequence and palindrome structure (29). Whether these motifs serve as replication origin remains to be established.

UL Genes Encoding Virion Proteins. A large cluster of tegument genes (UL18, UL19, UL25, UL36, UL37, UL48, and UL49) is present in MDV genome (30). Of note are UL48 and UL36. UL48 encodes a protein homologous to VP16, or αTIF, which initiates the transcription of other immediate early genes. The

Table 1. Putative MDV-1 gene products and homologs present in other herpesviruses*

ORFs	Location*, bp	Size, aa	Identity/overlap of aa, % MDV-2	Identity/overlap of aa, % HSV-1	Designation	ORFs	Location*, bp	Size, aa	Identity/overlap of aa, % MDV-2	Identity/overlap of aa, % HSV-1	Designation
R-LORF1	333–1052 (c)	239			ICP0	UL16	37996–39078 (c)	360	216/352 (61)	109/338 (32)	
	137624–138343					UL17	39082–41313 (c)	743	402/727 (55)	216/754 (28)	
R-LORF2	Join:(c)	142			vIL-8	UL18	42745–43704 (c)	319	251/320 (78)	129/313 (41)	Capsid protein
	1437–1665					UL19	43830–48005 (c)	1391	1184/1393 (84)	708/1368 (51)	MCP
	1764–1899					UL20	48307–49011 (c)	234	128/228 (56)	53/194 (27)	Membrane protein
	2075–2138					UL21	49272–50912	546	314/546 (57)	157/544 (28)	
	136538–136601					UL22	51058–53499 (c)	813	464/809 (57)	157/689 (22)	gH
	136777–136912					UL23	53682–54740 (c)	352	257/350 (73)	99/322 (30)	TK
	137011–137239					UL24	54695–55669	323	152/310 (49)	76/189 (40)	
R-LORF3	1722–2030	102				UL25	55701–57452	583	430/581 (74)	258/570 (45)	
	136646–136954 (c)					UL26	57493–59409	636	322/533 (60)	191/538 (35)	
R-LORF4	2334–2711 (c)	125				LORF5	59557–59916	136			
	135965–136342					UL27	59616–62213 (c)	865	715/865 (82)	411/847 (49)	gB
R-LORF5	2686–3033	115				UL28	62288–64669 (c)	793	561/785 (71)	350/773 (45)	
	135643–135990 (c)					UL29	64861–68436 (c)	1191	949/1189 (79)	504/1119 (45)	DNA binding pro.
R-LORF6	4200–4817 (c)	205				UL30	68705–72361	1218	880/1180 (74)	612/1155 (52)	DNA polymerase
	133859–134476					UL31	72285–73187 (c)	300	230/300 (75)	126/255 (49)	
R-LORF7	4270–5289 (c)	339			Meq	UL32	73204–75129 (c)	641	440/641 (68)	281/671 (41)	glycoprotein
	133387–134406					UL33	75128–75490	120	67/115 (58)	40/110 (36)	
R-LORF8	5615–6028	137				UL34	75623–76453	276	173/271 (63)	93/188 (49)	Virion protein
	132648–133061 (c)					UL35	76537–76929	130	88/127 (69)		
R-LORF9	8281–8604 (c)	107				UL36	76990–86967	3325	1552/2703 (57)	840/3050 (27)	LTP
	130072–130395					LORF6	86615–87082	155			
R-LORF10	9009–9311	100				UL37	87182–90322 (c)	1046	624/1025 (62)	282/1029 (27)	Capsid assem. pro.
	129365–129667 (c)					LORF7	90033–90395	120			
R-LORF11	10152–10463	103				UL38	90695–92107	470	294/472 (62)	138/169 (37)	
	128214–128525 (c)					UL39	92332–94800	822	603/792 (76)	370/776 (47)	RR large unit
R-LORF12	11601–11948 (c)	115				UL40	94853–95884	343	256/323 (79)	187/304 (61)	RR small unit
	126727–127076					UL41	95932–97257 (c)	441	288/439 (65)	110/285 (38)	vhs protein
LORF1	11893–12894 (c)	333				UL42	97914–99023	369	272/365 (74)	85/279 (30)	Subunit of DNA pol
R-LORF13	12280–12594	104				UL43	99183–100445	420	228/422 (54)		
R-LORF14	12389–12856	155	38/91 (41)		pp24	UL44	100665–102170	501	364/501 (72)	60/242 (24)	gC
LORF2	Join:	752	398/763 (52)		v-lipase	LORF8	102767–103493 (c)	208			
	13091–13186					UL45	103033–103668	211	155/211 (73)		
	13257–15431					UL46	103801–105507 (c)	568	231/559 (41)	110/383 (28)	
LORF3	16351–17547	198				UL47	105649–108075 (c)	808	409/844 (48)	129/586 (22)	
UL1	17737–18224	195	102/169 (60)	22/88 (25)	gL	UL48	108314–109597 (c)	427	248/409 (60)	131/335 (39)	VP16
LORF4	17745–18173 (c)	142				UL49	109705–110454 (c)	249	129/252 (51)	43/135 (31)	VP22
UL2	18206–19147	313	203/361 (63)	131/268 (48)		UL49.5	110597–110890 (c)	95	67/95 (70)		Membrane protein
UL3	19173–19858	228	153/215 (71)	110/221 (49)		UL50	110867–112177	436	254/389 (65)	60/162 (37)	
UL4	20374–21180 (c)	268	144/268 (53)	40/128 (28)		UL51	112265–113014 (c)	249	126/181 (69)	69/151 (45)	
UL5	21234–23810 (c)	858	667/849 (78)	476/849 (56)	Helicase	UL52	113016–116243	1075	698/1073 (65)	405/1071 (37)	
UL6	23882–26038	718	481/680 (70)	269/609 (44)	Virion protein	UL53	116222–117286	354	227/351 (64)	90/343 (26)	gK
UL7	25875–26792	305	171/303 (56)	80/264 (30)		UL54	117434–118855	473	260/485 (53)	82/230 (35)	ICP27
UL8	26826–29132 (c)	768	480/766 (82)	192/774 (24)		LORF9	118985–119794 (c)	269	118/192 (61)		
UL9	29146–31671 (c)	841	604/877 (68)	403/846 (47)	OBP	UL55	119962–120459	166	92/162 (56)	39/120 (32)	
UL10	31770–33044	424	325/424 (76)	104/402 (29)	gM	LORF10	120818–121399 (c)	193			
UL11	33098–33352 (c)	84	27/51 (52)	12/28 (42)		LORF11	122031–124742 (c)	903	480/900 (53)		
UL12	33328–34902 (c)	524	265/468 (56)	142/392 (36)		LORF12	124944–125390 (c)	148			
UL13	34893–36434 (c)	513	336/512 (65)	142/417 (34)	PK	R-LORF14a	125416–126288 (c)	290	64/158 (40)		pp38
UL14	36152–36898 (c)	248	123/229 (53)	42/139 (30)							
UL15	Join: 36913–37956	737	570/737 (77)	405/743 (54)		R-LORF13a	126023–126397 (c)	124			
	41457–42629										

*A homologous search was conducted with advanced BLAST from the National Center of Biotechnological Information with nonfilter option. The result was expressed as number of identical amino acids/number of overlapped amino acids of MDV-1 gene product with other herpes virus homolog (percentage of identity).

sequence homology between MDV and HSV-1 VP16 is 39% over a stretch of 335 aa (30). UL36 is a large tegument protein of 3,304 aa in length and has certain unusual structural features. In addition to the three UL9-binding motifs, it carries 18 units of 18-bp repeat and 7 units of 30-bp repeat (29). The function of these regular repeat structures is not clear.

UL Genes Encoding Glycoproteins. All of the glycoproteins homologs to those of HSV-1 in the UL region, such as gL (UL1) (31), gM (UL10), gH (UL22) (32) gB (UL27) (14), gC (UL44) (13), gK (UL53) (33), as well as the glycoprotein encoded by UL32 (34) have been identified. They all carry signal peptides at the N terminus and membrane-anchoring hydrophobic sequences with an exception of

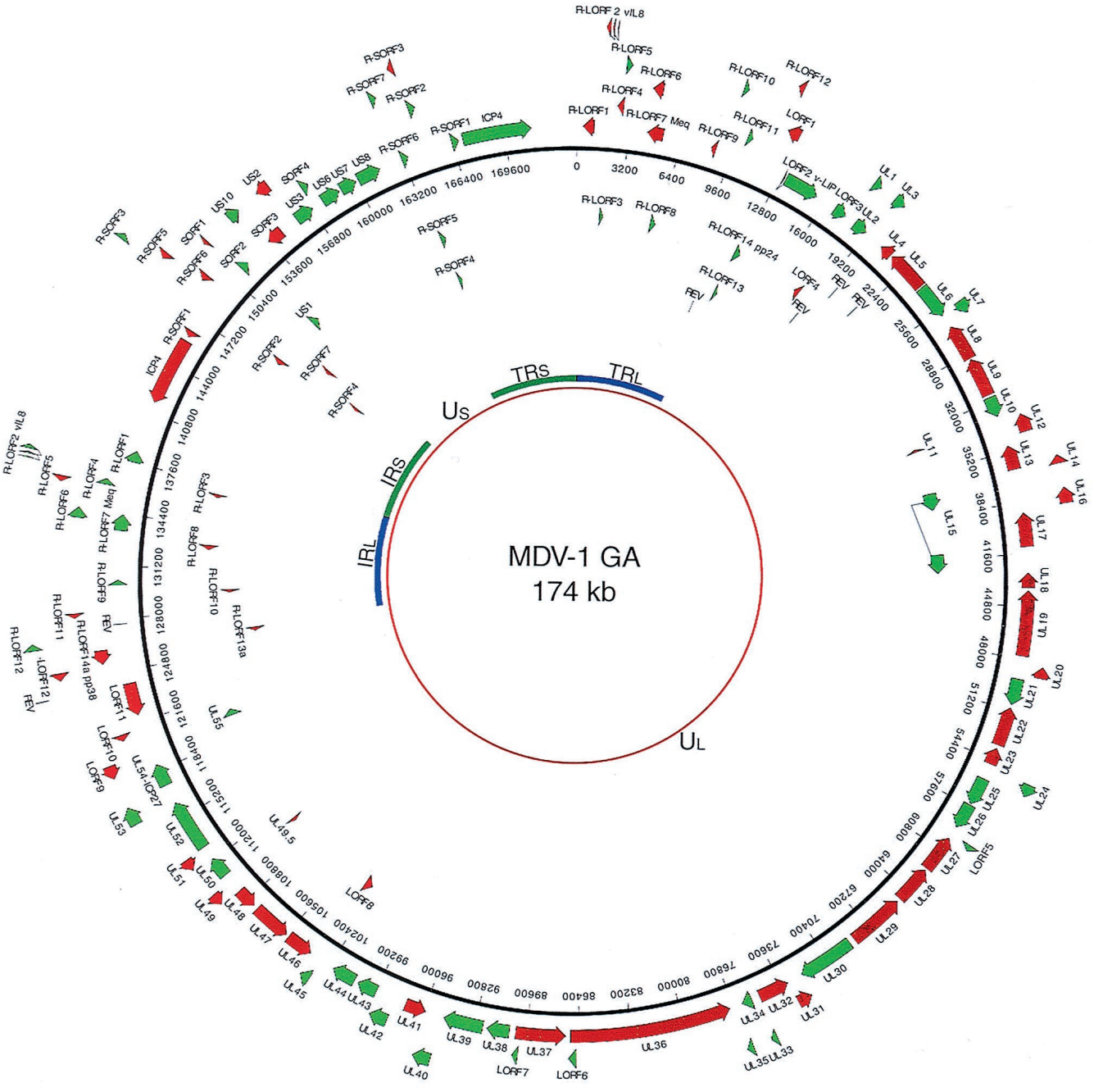

Fig. 1. Genetic organization of GA strain of MDV. The map is based on the recent UL sequence data and the rest is predicted from the other MDV sequence data available as described in the text. Retroviral insertion sites near the repeat/unique junctions were indicated in the map.

gL. The sequence identity between MDV and HSV in their homologous regions ranges from 22% in gH to 49% in gB. Interestingly, like HSV, gC was not found to be essential for MDV replication *in vitro*, and gB offers significant protective immunity to challenge with virulent strains (35).

UL Genes Involved in Regulation. MDV regulatory genes homologous to HSV infected cell protein (ICP) 0 and ICP27 are present in the repeat and UL regions, respectively. Interestingly there is no ICP47 homolog encoded by MDV-1 and R-LORF1 only has limited homology with HSV-2 ICP0. The MDV UL54 (ICP27) homolog has been reported to be an immediate-early phosphorylated nuclear protein (33).

LORFs: MDV-Specific ORFs. Aside from genes related to HSV, there are several interesting ORFs, which are unique to MDV-I or to the MDV family of viruses in general. We have named these unique ORFs in the UL region as LORFs, and those in the long repeat as R-LORFs. (This nomenclature follows the system used by Brunovskis and Velicer (8) in describing SORFs in unique small region.) There are 16 R-LORFs and 12 LORFs. Most of these LORFs are completely uncharacterized and in most cases a BLAST search does not reveal significant homology to known genes or known sequence motifs. However, there are several exceptions in which homology to other genes has been detected and in some cases, important properties related to MDV functions were identified.

Meq, R-LORF7, is perhaps the most extensively characterized among the unique genes of MDV. This protein has a basic leucine zipper domain that shares significant homology with jun/fos family of transcriptional factors. Meq is localized in the nucleus and nucleolus (36). It binds specific DNA motifs, dimerizes with itself and with jun/fos family of basic leucine zipper proteins, and exhibits transactivation and repression activities (37). *Meq* is one of the few genes consistently expressed in all MDV transformed and tumor cells, implicating its functions in latency or transformation (38). Consistent with its role in oncogenesis are the observations that (*i*) when overexpressed, *Meq* can transform a rat fibroblast cell line and protects transformed cells from apoptosis (36); (*ii*) transfection of antisense *Meq* reverts the transformed phenotype of MDV tumor cell line (38); (*iii*) *Meq* knock-out MDV mutant exhibits no oncogenic potential (R. Morgan, personal communication); and (*iv*) full-length Meq is encoded by only the oncogenic, serotype 1 MDV. However, like the oncoprotein of other DNA tumor viruses, Meq also may play a significant, yet unidentified, role in MDV replication.

vIL8, R-LORF2, was initially discovered as a gene fused to an alternatively spliced *Meq*. In its natural form, it has three exons as revealed by the cDNA structure. It is among the first CXC chemokines identified in herpesviruses (39); most other herpesviruses encode CC chemokines or their receptors. The functions of this chemokine remain to be explored. It is tempting to speculate that virus-encoded interleukin-8 (vIL8) is involved in recruiting target cells for infections or serves as a decoy to diminish the cellular interleukin-8-mediated immune response.

pp38, R-LORF14a, is a 290-aa phosphoprotein, highly expressed in both the lytically and latently infected cells as well as in tumor cells (11, 12). This gene spans the junction of UL/IRL. The function of this protein is still not clear but has been suggested to be involved in the maintenance of transformation and immunosuppression. The promoter of this gene coincides with the putative replication origin of MDV. pp24, another phosphoprotein, which shares the same 65 N-terminal amino acids with pp38, is located at the junction of TRL and UL (40). pp24 has the ability to associate with pp38. As such it may function in the same pathway as pp38 or may serve as a regulator for pp38. Other than this, very little is known about pp24.

v-LIP, LORF2, is predicted to encode a 756-aa protein with a region of remarkable degree of amino acid identity with the GXSXG motif containing catalytic domains of hornet venom phospholipase A1 and vertebrate lipoprotein and triacylglycerol lipases. cDNA sequencing has revealed that the *v-LIP* transcript is made up of two exons; a signal sequence is conferred from a very short first exon and the catalytic domain is encoded by the much longer second exon. This pattern is conserved in the close homolog present in MDV-2 HPRS-24 (22). Remarkably, aside from the MDV-2 homolog, the closest homologs to *v-LIP* are ORFs from several avian and fowl adenoviruses (GenBank accession nos. AAB88667 and AF155911-9) (41), which may suggest ancestral recombination events between these two DNA viruses. To our knowledge this is the first herpesviral lipase ever reported. The function of this lipase in MDV infection greatly deserves further investigation. Interestingly, in a retrovirus insertion mutant of the JM strain of MDV-1, JM-Hi3, a REV LTR was inserted in the *v-LIP* coding sequence such that lipase activity should be abrogated (20). The REV LTR is inserted in the *v-LIP* coding sequence at amino acid 123. This mutant not only is competent at *in vitro* replication but also replicates much better *in vitro* than the parental JM virus (42). This suggests that *v-LIP* is not required for *in vitro* replication and its function is more likely to be manifested *in vivo*. It is conceivable that *v-LIP* may contribute to the atherosclerosis phenotype for some strains of MDV.

Retrovirus Ancestral Insertion Sites in MDV-1. One of the fascinating features of MDV is the presence of retroviral-like sequences in the genome of serotype 1, but not SB-1 (serotype II) or turkey herpesvirus (serotype III) (20, 43–45). There are several stretches of retrovirus REV LTR sequences clustered near the UL/RL junctions, which are believed to be due to ancestral insertions. These sequences, derived from the U3 region of REV LTR are involved in transcriptional regulation and tissue tropism of retroviruses (46) Some of the ancestral insertions may have been assimilated into the coding sequences, e.g., the insertions in UL5 and in UL2. Others such as the insertion near LORF12, which is 21 nt preceding the AUG codon, the insertion near UL3, which is 51 nt preceding the AUG codon could potentially contribute enhancer functions. Detailed mutational analyses are required to reveal the possible functions of these LTR remnants.

Summary. The UL sequence of MDV reported here, together with the US sequence (8, 9) and the various sequences (8–15) described previously, constitute virtually the entire genome of MDV. The UL sequences of MDV are highly homologous to HSV with corresponding genes arranged in exactly the same order and presumably the same transcriptional units. These genes are mostly involved in viral replication and virion assembly. Consequently, a similar infection and replication scheme for these viruses is predicted. There are however interesting differences. For instance, MDV replication produces mostly cell-associated virions and unlike HSV, gD (US6) is found not to be required for *in vitro* MDV growth (47–49). *In vivo*, MDV infections target the bursa of Fabricius and the thymus, resulting in the transformation of T cells, which are distinct from the routes taken by HSV. Some of these MDV-specific properties may be attributed to the presence of extra-coding sequences, defined as LORFs or SORFs in the MDV genome. There are four SORFs, SORF1–4, present in US of MDV (8). Only SORF2, which encodes a 132-aa protein, has been partially characterized. SORF2 is believed to be analogous to the HCMV US22 transactivator family of proteins (44). A MDV carrying retroviral LTR insertion upstream of SORF2, with consequent SORF2 activation is attenuated in oncogenicity and causes thymic atrophy (43, 44, 50). This suggests that SORF2 may be involved in modulating MDV replication in thymus. Another intriguing feature of SORF2 is that very close homologs are found in fowlpox virus and, as in the case of *v-LIP*, in fowl adenovirus (8). It is emblematic of the likely importance of SORF2 in viral pathogenesis that 3 distinct families of double stranded DNA viruses of chickens conserve such close SORF2 homologs. This also may suggest ancestral recombination events between these different DNA viruses. There are two classes of MDV unique genes. One class is comprised of the MDV genes that lack HSV-1 homologs and are coded for by the unique genome regions; these genes seem to be conserved between all serotypes, which may indicate their common functions in the life cycle of this family of viruses. The other class of MDV unique genes are those which are coded by the repeat regions and which appear to differ greatly between serotypes and may account for the differences in pathogenicity and biological properties between the major serotypes of MDV. The LORFs and R-LORFs in the UL may represent entire genes or segments of spliced genes. Although infrequent, spliced transcripts are produced by MDV genome. The analysis of R-LORFs and R-SORFs, most of which are unique to serotype I MDV, should provide important clues to the oncogenic properties of this virus. Several genes, such as *Meq*, *vIL8*, and *pp38* are described in this report and have already contributed to our understanding of MDV pathogenesis. Studies of other MDV-specific gene products will undoubtedly continue to shed light on the cellular and viral processes behind Marek's disease.

This work is supported by a National Institutes of Health Grant CA 46613 and United States Department of Agriculture Grant 97–35204.

1. Churchill, A. E. & Biggs, P. M. (1967) *Nature (London)* **215,** 528–530.
2. Nazerian, K., Solomon, J. J., Witter, R. L. & Burmester, B. R. (1968) *Proc. Soc. Exp. Biol. Med.* **127,** 177–182.
3. Calnek, B. W. & Witter, R. L. (1997) in *Diseases of Poultry*, eds. Calnek, B. W., Barnes, H. J., Beard, C. W., McDougald, L. R. & Saif, Y. M. (Iowa State Univ., Press, Ames), pp. 369–413.
4. Fabricant, C. G., Fabricant, J., Litrenta, M. M. & Minick, C. R. (1978) *J. Exp. Med.* **148,** 335–340.
5. Lee, L. F., Liu, X. & Witter, R. L. (1983) *J. Immunol.* **130,** 1003–1006.
6. Fukuchi, K., Sudo, M., Lee, Y. S., Tanaka, A. & Nonoyama, M. (1984) *J. Virol.* **51,** 102–109.
7. Buckmaster A. D., Scott, S. D., Sanderson, M. J., Bournsnell, M. E. G., Ross, N. L. J. & Binns, M. M. (1988) *J. Gen. Virol.* **69,** 2033–2042.
8. Brunovskis, P. & Velicer, L. F. (1995) *Virology* **206,** 324–328.
9. McKie, E. A., Ubukata, E., Hasegawa, S., Zhang, S., Nonoyama, M. & Tanaka, A. (1995) *J. Virol.* **69,** 1310–1314.
10. Jones, D., Lee, L. F., Liu, J. L. Kung, H. J. & Tillotson, J. K. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4042–4046.
11. Cui, Z. Z., Lee, L. F., Liu, J. L. & Kung, H. J. (1991) *J. Virol.* **65,** 509–515.
12. Chen, X. B., Sondermeijer, P. J. & Velicer, L. F. (1992) *J. Virol.* **66,** 85–94.
13. Coussens, P. M. & Velicer, L. F. (1988) *J. Virol.* **62,** 2373- 2379.
14. Ross, L. J. N., Sanderson, M., Scott, S. D., Binns, M. M., Doel, T. & Milne, B. (1989) *J. Gen. Virol.* **70,** 1789–1804.
15. Ross, L. J. N., Binns, M. M., Sanderson, M. & Schat, K. A. (1993) *Virus Genes* **7,** 33–51.
16. Eidson, C. S. & Schmittle, S. C. (1968) *Avian Dis.* **12,** 467–476.
17. Witter, R. L. (1997) *Avian Dis.* **41,** 149–163.
18. Gibbs, C. P., Nazerian, K., Velicer, L. F. & Kung, H. J. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3365–3369.
19. Silva, R. F. & Witter, R. L. (1985) *J. Virol.* **54,** 690–696.
20. Isfort, R., Jones, D., Kost, R., Witter, R. & Kung, H. J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 991–995.
21. Igarashi, T., Takahashi, M., Donovan, J., Jessip, J., Smith, M., Hirai, K., Tanaka, A. & Nonoyama, M. (1987) *Virology* **157,** 351–358.
22. Kato, K., Jang, H. K., Izumiya, Y., Cai, J. S., Tsushima, Y., Miyazawa, T., Kai, C. & Mikami, T. (1999) *J. Vet. Med. Sci.* **61,** 1161–1165.
23. McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., NcNab, D., Perry, L. J., Scott, J. E. & Taylor, P. (1988) *J. Gen. Virol.* **69,** 1531–1574.
24. Roizman, B. & Sears, A. E. (1990) in *Virology*, eds. Fields, B. N. & Knipe, D. M. (Raven, New York), 2nd Ed., pp. 1795–1841.
25. Sui, D., Wu, P., Kung, H. J. & Lee, L. F. (1995) *Virus Res.* **36,** 269–278.
26. Lee, L. F., Nazerian, K., Leinbach, S. S., Reno, J. M. & Boezi, J. A. (1976) *J. Natl. Cancer Inst.* **56,** 823–827.
27. Bradley, G., Hayashi, M., Lancz, G., Tanaka, A. & Nonoyama, M. (1989) *J. Virol.* **63,** 2534–2542.
28. Qian, Z., Brunovskis, P., Lee, L. F., Vogt, P. K. & Kung, H. J. (1996) *J. Virol.* **70,** 7161–7170.
29. Mao, Y. P., Wu, P., H.-J. Kung, H. J. & Lee, L. F. (1996) in *Current Research on Marek's Disease*, eds. Silva, R. F., Cheng, H. H., Coussens, P. M., Lee, L. F. & Velicer, L. F (Am. Assoc. Avian Pathologists, Kennett Square, PA), pp. 176–181.
30. Yanagida, N., Yoshida, S., Nazerian, K. & Lee, L. F. (1993) *J. Gen. Virol.* **74,** 1837–1845.
31. Yoshida, S., Lee, L. F., Yanagida, N. & Nazerian, K. (1994) *Virology* **200,** 484–493.
32. Wu, P., Reed, W., Yoshida, S., Sui, D. & Lee, L. F. (1999) *Acta Virologica* **43,** 152–158.
33. Ren, D., Lee, L. F. & Coussens, P. M. (1994) *Virology* **204,** 242–250.
34. Lee, L. F., Wu, P. & Sui, D. (1996) in *Current Research on Marek's Disease*, eds. Silva, R. F., Cheng, H. H., Coussens, P. M., Lee, L. F. & Velicer, L. F (Am. Assoc. Avian Pathologists, Kennett Square, PA), pp. 245–250.
35. Nazerian, K., Lee, L. F., Yanagida N. & Ogawa, R. (1996) *J. Virol.* **69,** 1409–1413.
36. Liu, J. L, Ye, Y., Lee, L. F. & Kung, H. J. (1998) *J. Virol.* **72,** 388–395.
37. Qian, Z., Brunovskis, P., Rauscher, F., III, Lee, L. & Kung, H. J. (1995) *J. Virol.* **69,** 4037–4044.
38. Xie, Q., Anderson, A. S. & Morgan, R. W. (1996) *J. Virol.* **70,** 1125–1131.
39. Liu, J. L., Lin, S. F., Xia, L., Brunovskis, P., Li, D., Davidson, I., Lee, L. F. & Kung, H. J. (1999) *Acta Virologica* **43,** 94–101.
40. Zhu, G. S., Iwata, A., Gong, M., Ueda, S. & Hirai, K. (1994) *Virology* **200,** 816–820.
41. Becker, Y., Asher, Y., Tabor, E., Davidson, I. & Malkinson, M. (1994) *Virus Genes* **8,** 55–69.
42. Witter, R. & Offenbecker, L. (1979) *J. Natl. Cancer Inst.* **62,** 143–151.
43. Isfort, R. J., Qian, Z., Jones, D., Silva, R. F., Witter, R. L. & Kung, H. J. (1994) *Virology* **203,** 125–133.
44. Jones, D., Isfort, R., Witter, R. L., Kost, R. & Kung, H. J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 3855–3859.
45. Jones, D., Brunovskis, P., Witter, R. L. & Kung, H. J. (1996) *J. Virol.* **70,** 2460–2467.
46. Coffin, J. (1996) in *Fundamental Virology*, eds. Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B. & Straus, S. E. (Raven, Philadelphia), pp. 763–843.
47. Ross, L. J. N. & Binns, M. M. (1991) *J. Gen. Virol.* **72,** 939–947.
48. Parcells, M. S., Anderson, A. S. & Morgan, R. W. (1994) *Virus Genes* **9,** 5–13.
49. Parcells, M. S., Anderson, A. S. & Morgan, R. W. (1995) *J. Virol.* **69,** 7888–7898.
50. Witter, R. L., Li, D., Jones, D., Lee, L. F. & Kung, H. J. (1997) *Avian Dis.* **41,** 407–421.