Marburgvirus Genomics and Association with a Large Hemorrhagic Fever Outbreak in Angola

Jonathan S. Towner,[1] Marina L. Khristova,[2] Tara K. Sealy,[1] Martin J. Vincent,[1] Bobbie R. Erickson,[1] Darcy A. Bawiec,[1] Amy L. Hartman,[1] James A. Comer,[1] Sherif R. Zaki,[3] Ute Ströher,[4,5] Filomena Gomes da Silva,[6] Fernando del Castillo,[7] Pierre E. Rollin,[1] Thomas G. Ksiazek,[1] and Stuart T. Nichol[1]*

Special Pathogens Branch,[1] DVRD, Biotechnology Core Facility Branch,[2] SRP and Infectious Disease Pathology Activity,[3] and DVRD, NCID, and Global Aids Program,[7] Centers for Disease Control and Prevention Atlanta, Georgia 30333; Special Pathogens Program, National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba R3E 3R2,[4] and Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba R3E OW3,[5] Canada; and Instituto Nacional de Saúde Publica, Ministro da Saúde, República de Angola, Luanda, Angola[6]

In March 2005, the Centers for Disease Control and Prevention (CDC) investigated a large hemorrhagic fever (HF) outbreak in Uige Province in northern Angola, West Africa. In total, 15 initial specimens were sent to CDC, Atlanta, Ga., for testing for viruses associated with viral HFs known to be present in West Africa, including ebolavirus. Marburgvirus was also included despite the fact that the origins of all earlier outbreaks were linked directly to East Africa. Surprisingly, marburgvirus was confirmed (12 of 15 specimens) as the cause of the outbreak. The outbreak likely began in October 2004 and ended in July 2005, and it included 252 cases and 227 (90%) fatalities (report from the Ministry of Health, Republic of Angola, 2005), making it the largest Marburg HF outbreak on record. A real-time quantitative reverse transcription-PCR assay utilized and adapted during the outbreak proved to be highly sensitive and sufficiently robust for field use. Partial marburgvirus RNA sequence analysis revealed up to 21% nucleotide divergence among the previously characterized East African strains, with the most distinct being Ravn from Kenya (1987). The Angolan strain was less different (~7%) from the main group of East African marburgviruses than one might expect given the large geographic separation. To more precisely analyze the virus genetic differences between outbreaks and among viruses within the Angola outbreak itself, a total of 16 complete virus genomes were determined, including those of the virus isolates Ravn (Kenya, 1987) and 05DRC, 07DRC, and 09DRC (Democratic Republic of Congo, 1998) and the reference Angolan virus isolate (Ang1379v). In addition, complete genome sequences were obtained from RNAs extracted from 10 clinical specimens reflecting various stages of the disease and locations within the Angolan outbreak. While the marburgviruses exhibit high overall genetic diversity (up to 22%), only 6.8% nucleotide difference was found between the West African Angolan viruses and the majority of East African viruses, suggesting that the virus reservoir species in these regions are not substantially distinct. Remarkably few nucleotide differences were found among the Angolan clinical specimens (0 to 0.07%), consistent with an outbreak scenario in which a single (or rare) introduction of virus from the reservoir species into the human population was followed by person-to-person transmission with little accumulation of mutations. This is in contrast to the 1998 to 2000 marburgvirus outbreak, where evidence of several virus genetic lineages (with up to 21% divergence) and multiple virus introductions into the human population was found.

Marburgvirus was first identified in 1967, when infected monkeys, imported from the Lake Kyoga region of Uganda, transmitted the virus to laboratory workers and scientists at facilities in Germany and the former Yugoslavia (56). Marburgvirus and its close relative ebolavirus comprise the family *Filoviridae* and are known to cause severe hemorrhagic fever (HF) in humans and nonhuman primates (38). Marburgvirus is a single species, consisting of viruses differing from one another by up to 21% at the nucleotide level. In contrast, four distinct species of ebolavirus (Zaire, Sudan, Reston, and Ivory Coast) have been defined, which differ genetically from one another by approximately 37 to 41% (15). Marburgvirus is a

single-strand, negative-sense RNA virus whose seven gene products are, in order, nucleoprotein (NP), VP35, VP40, glycoprotein (GP), VP30, VP24, and the polymerase (L) (14). Ebola and Marburg HF outbreaks are generally thought to involve the relatively rare introduction of the virus into the human population followed by waves of human-to-human transmission (usually through close contact with infected individuals or their body fluids), with little if any evolution of the virus during the course of the outbreak (42, 48). Symptoms of marburgvirus infection include general malaise, acute fever, abdominal cramping, bleeding disorders, and shock (29). Case fatality rates in previous large Marburg HF outbreaks have ranged from 23 to 83% (5, 30).

Historically, sources of marburgvirus have been confined to East Africa. They have been centered almost exclusively within 500 miles of Lake Victoria, with the exception of a single case

---

in Zimbabwe in 1975, when a traveler became infected and, seeking medical treatment, subsequently transmitted the virus to a health care worker in South Africa. This previous close association of marburgvirus with East Africa contrasts with the observed distribution of ebolavirus, which has caused human HF outbreaks throughout tropical Africa, ranging from Cote d'Ivoire to Uganda. The largest Marburg HF outbreak previously on record occurred from 1998 to 2000 in Durba, Democratic Republic of Congo (DRC) (57), infecting 154 people and with a case fatality rate of approximately 83% (5). The majority of these patients were involved directly or indirectly with illegal gold mining activity in the region. The Durba outbreak was somewhat unusual compared to previous marburgvirus outbreaks in that it continued for almost 2 years, during which time multiple distinct genetic lineages were found to be circulating (37), likely indicating several independent introductions of virus into the human population from the unknown natural reservoir.

The recent emergence of marburgvirus in West Africa is described in this report. Based upon clinical presentations that meet the viral HF case definition (10), the outbreak likely started in October of 2004 in Uige Province in northern Angola yet continued undiagnosed until March 2005, when transmission of HF to health care workers alerted the community to the possibility of Marburg or Ebola HF (9, 58). The outbreak persisted until July 2005 (32), raising the possibility that multiple genetic lineages could have been introduced during this long epidemic period. In this study, we sequenced the entire marburgvirus genomes obtained from multiple laboratory-confirmed cases throughout the Angola outbreak and compared these sequences to those obtained at different times during the 1998 marburgvirus outbreak in Durba, DRC, as well as the complete genome of Ravn marburgvirus isolated in 1987. Collectively, these 16 newly derived complete genome sequences, along with those previously determined for the Popp, Musoke, and Ozolin strains, afforded us the opportunity to carry out the most comprehensive survey to date of the sequence spectrum circulating among marburgviruses. We determined that the newly emergent Angola marburgvirus is closely related to the majority of the previous East African isolates.

MATERIALS AND METHODS

Specimen processing, antigen capture, IgM and IgG ELISA, virus isolation, and IHC staining. Blood and serum specimens from suspected viral HF (VHF) patients were received unfrozen at CDC in International Air Transport Association-approved containers. Samples 200501377 to -88, 200501411, and 200501412 were opened under biosafety level 4 conditions, and virus isolation using Vero E6 cells was performed as previously described (42). The Angola marburgvirus reference sample is clinical specimen number 200501379, or "1379c," while the virus isolated from 1379c is designated "1379v". The samples were inactivated with gamma-cell irradiation or chaotrope for RNA extraction and then tested for marburgvirus antigen, specific immunoglobulin M (IgM), IgG, and nucleic acid by using modifications of existing techniques for Ebola hemorrhagic fever (22, 23, 48, 49). The modifications for the IgM and IgG enzyme-linked immunosorbent assays (ELISAs) consisted of (i) a replacement of the pooled monoclonal antibody mix with one directed to marburgvirus (Musoke strain) NP and (ii) a replacement of the Vero cell lysates with ones in which Vero cells were infected with marburgvirus (Musoke strain). The modifications for the nucleic acid detection were to use a marburgvirus-specific quantitative reverse transcription-PCR (Q-RT-PCR) assay designed to detect all known species of marburgvirus (described below). Immunohistochemical (IHC) staining was performed on skin biopsies fixed in 10% formalin as previously described (60) and adapted for detection of marburgvirus. Specimens 2005000126 to -998, processed

in the field lab, were not gamma irradiated but were treated additionally with 0.5% Triton X-100 and heated at 56°C for 30 min to greatly reduce infectivity.

RNA extraction. RNA was extracted using an Applied Biosystems 6100 nucleic acid preparation station according to a modified protocol (49) using noncellular lysis buffer. Briefly, 50 μl of sample was mixed with 300 μl of 2× noncellular lysis buffer under biosafety containment conditions. Following >15 min of incubation, plastic ware containing samples in lysis buffer was externally decontaminated with 3% Lysol, and RNAs were transported out of biosafety containment and extracted according to the manufacturer's instructions, with the modification of adding two extra 1× noncellular lysis buffer washes at the beginning of the extraction process.

PCR diagnostic assay. Briefly, a Q-RT-PCR assay was designed to detect a region in VP40. Q-RT-PCR assays were carried out by first random-primed conversion of the RNA to cDNA using the High Capacity cDNA synthesis kit (Applied Biosystems) according to the manufacturer's instructions. Ten microliters of the cDNA synthesis reaction mixture was then used to program probe-based real-time PCR assays, using Universal Master Mix (Applied Biosystems) in 50-μl total reaction volumes. The marburgvirus broadly reactive Q-PCR assay used the following forward and reverse primers and probe, respectively: 5′-GG ACCACTGCTGGCCATATC-3′, 5′-GTCGGCAGGAGGIGAAATCC-3′, and 6-carboxyfluorescein-5′-CTCTGGGACTT**T**TTCIACCCTCAGTTGATGA-3′. The quencher, Black Hole, was internally placed on T indicated by boldface within the probe oligonucleotide. Amplifications were carried out for 40 cycles by denaturing for 20 seconds at 94°C followed by 30 seconds at 60°C degrees. Samples exhibiting cycle threshold (C_T) values of less than 35 were considered positive. The VP40 gene sequence of the Angola-05 marburgvirus isolate was determined and allowed development of an Angola-specific Q-RT-PCR assay. The slightly modified primers and probe were 5′-GGTCCACTGCTGGCCATA TC-3′, 5′-GTCGGCAGGAAGCGAAATCC-3′, and 6-carboxyfluorescein-5′-T TCTGGGACTT**T**TTCGACTCTCAGTTGATGA-3′, where the quencher was again internally placed on the T indicated by boldface. The modified Q-RT-PCR assay provided two to fourfold (1 to 2 C_T units) better sensitivity in samples assayed side by side with the original assay (data not shown).

Virus passage histories. Ravn isolate (1987), strain M/Kenya/Kitum Cave (20), was passaged once in SW13 cells, followed by four passages in Vero E6 cells. 05DRC, 07DRC, and 09DRC viruses (Durba, DRC, 1998) were passaged three times each in Vero E6 cells.

Genome fragment amplification and sequencing. The reference sequence, Ang1379c, was determined from RNA isolated directly from clinical material and also from the virus isolate (Ang1379v) after one passage on Vero E6 cells. Multiple RT-PCR fragments were generated using the Superscript III one-step high-fidelity RT-PCR kit (Invitrogen) according to the manufacturer's instructions. The fragments, amplified using the following combinations of primers (listed 5′ to 3′) and designed based upon a consensus sequence derived from an alignment of Musoke, Popp and Ozolin marburgviruses (GenBank ascension numbers DQ217792, NC001608, and AY358025, respectively), were as follows (the oligonucleotide names reflect the approximate nucleotide position of each primer): 21F (GACACACAAAAACAAGAGATGATG) and 810R (GAAGTC CTGAGAATCTAGTTTG), 132F (GTACAAAACCCACTGCCCCTCATG) and 3489R (AACCGTAGTCCCCTTACTAC), and 5363F (AGAIGGIATGAT GAAGAAAAG) and 7781R (CAGGTCCIAGCACITTGCATGTTCC), followed by nested amplification using 5420F (TGAGAAITTICCTTTGAATGGC TTC) and 7671R (TCCAAGGATTIGCAGTTTG), 7567F (TGGCCCTGGA ATIGAAGGACTC) and 10571R (AGCATATGAACAATAGATC), and 10401F (TGTTCCAGAGTGGCAACAAAC) and 14933R (CTCAGAIAAATT AATCCCACTTTAC), followed by heminested amplification using 10401F and 14882R (TGTGGCACCAATTAGCCTTTTCCC) and 14763F (TGGACGATG ATTTATCTGAGTC) and 19113R (GGACACACTAAAAAGATGAAG). The region between nucleotides 3489 and 5363 was amplified by using primers VP35F1 (GCTTACTTAAATGAGCATGG) and GP170R (TGCCCACTCAG TGTAAATCC). The full-length genome sequences from clinical specimens from this outbreak (samples 1380, 1381, 1386, 1411, 0126, 0181, 0214, 0215, 0754, and 0998) and isolates 05DRC and 07DRC were similarly determined with primer modifications by amplifying six genome fragments, designated A to F, as follows: fragment A was amplified with 63F (TGACATTGAGACTTGTCAGTC) and 4997R (GCTTGATTTCCTTCACGC) followed by nested amplification using 375F (CAAGAATGCAGATGCAACC) and 4641R (GTTTGCACCGTGGTC AG); fragment B was amplified using primers 3781F (GTTCCTCCAGTGATA AGAGTC) and 8478R (GATTGAAGTAAGGCAAGTTGTTAA) followed by nested amplification using 4155F (CTAACAAAAGGGGTCTAACCTA) and 8304R (CTTAACTATGTCTTCAAGCCTA); fragment C was amplified using primers 7567F and 10571R followed by nested amplification (sample 1386 only) using 7621F (TTTITGITGIAGGTTGAGGCG) and 10521R (TTGATCCTTC
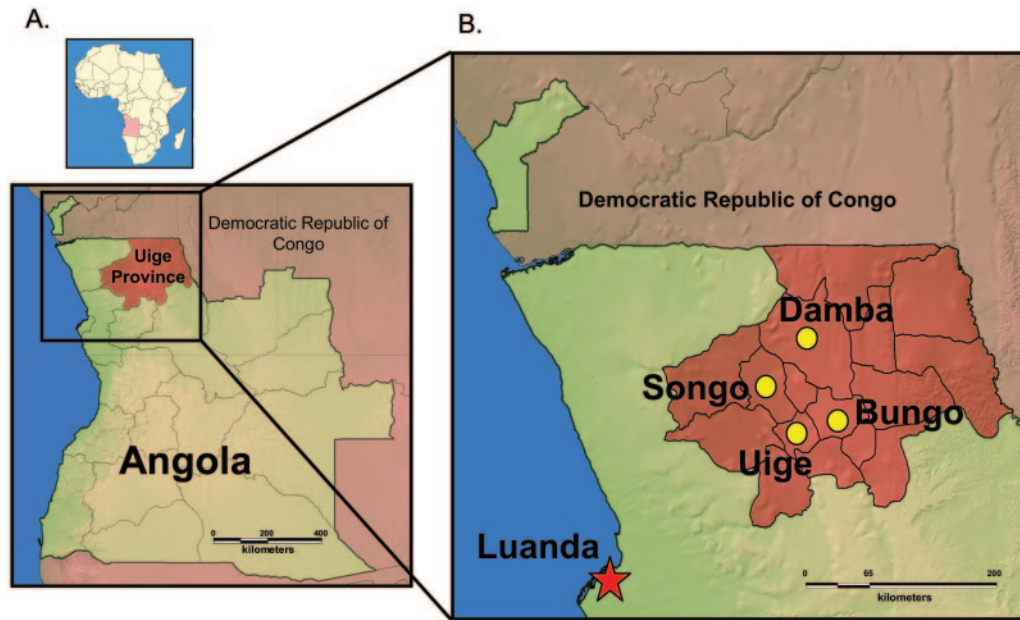
FIG. 1. (A) Map of Angola, with Uige Province highlighted in burgundy. (B) Expansion of the inset in panel A to show the local municipalities from which individual cases were analyzed. The majority of cases (~90%) in the outbreak were in the Uige municipality located within Uige Province.

AAAGCAACTCC); fragment D was amplified using primers 10000F (GTACC TCTAAGGAAAACCATGAAG) and 15460R (GTTGATATAATTGCACGT GTGG) followed by nested amplification using primers 10287F (GACCTCAA TTCCACAGCA) and 15279R (GTCTGCAGTGCCTTGAGT); fragment E was amplified using 14763F and 19113R followed by heminested amplification using 14817F (CTAATTTCTTGAGGGCATATTC) and 19113R; and fragment F was amplified in a heminested reaction using 3F (CACACAAAAACAAGAGATG ATG) and 925R (CTCAACGCCTGCTTGAAAG) followed by heminested amplification using 3F with 801R. The RT-PCR fragments for the Ravn and 09DRC virus isolates required considerable primer modifications. Nucleotides ~3 to 6000 were generated in four fragments using primers 3F with 721R (ACCTGT CACCAAATTTACTCC), 617F (CACTGGCTTACTACAGGCCA) with 1128R (TGTTCATGTCGCCTTTGTAG), 818F (ATTGCATCCACTTGT GCG) with 2117R (AGACAGGATTTTGGTTATTCC), and 1581F (CACTAA GGCAAACTCAGGAC) with GP170R. The remainder of the genome was generated using primers described above, namely, 7567F with 10571R, 10401F with 14882R, and 14763F with 19113R. If necessary, RT-PCR products of the correct predicted sizes were gel purified using either the Qiaquick gel extraction kit (QIAGEN, Valencia, CA) or the NucleoTrap gel extraction kit (BD Biosciences Clontech, Palo Alto, CA). The purified DNA was sequenced by primer walking of both strands, using ABI BigDye 3.1 dye chemistry and ABI 3730XL automated DNA sequencers (Applied Biosystems, Foster City, CA). Sequencing primers representing the most conserved regions were designed encompassing the entire genome fragment. Sequencing primers were synthesized by Integrated DNA Technologies (Coralville, IA). Chromatogram data were assembled using Seqmerge (Accelrys Inc, San Diego, CA), Phred/Phrap base calling and assembly software, and Consed for sequence editing (University of Washington). The terminal ~20 nucleotides were not experimentally determined but were based on the termini of the previously published Popp, Musoke, and Ozolin sequences.

Phylogenetic analysis. Genomic analyses and alignments of filovirus sequences were performed using the Wisconsin package of GCG version 10.3. (Accelrys Inc., San Diego, CA). Phylogenetic analysis was performed using Paup 4.0b10 (Sinauer Associates Inc.).

Nucleotide sequence accession number. The GenBank Accession numbers for the newly determined full-length genome sequences are as follows: Ravn, DQ447649; 07DRC, DQ447650; 05DRC, DQ447651; 09DRC, DQ447652; Ang1379c, DQ447653; Ang1381, DQ447654; Ang1386, DQ447655; Ang0126, DQ447656; Ang0214, DQ447657; Ang0215, DQ447658; Ang0754, DQ447659; and Ang0998, DQ447660.

RESULTS

Discovery of marburgvirus HF in West Africa. In early 2005, reports were received by the CDC and World Health Organization (WHO) of a large HF outbreak in Uige Province in northern Angola, West Africa (Fig. 1). On 15 March, specimens were collected and sent to CDC in Atlanta, Ga., for testing for evidence of infection with viruses associated with VHF which were known to be present in West Africa. To be thorough, the VHF-associated Marburg virus was also included despite the origin of all earlier outbreaks being restricted to East Africa. Testing included virus isolation, antigen capture ELISA, IgM and IgG ELISA, RT-PCR and Q-RT-PCR assays, and in one case IHC staining. Contrary to expectations, evidence of acute marburgvirus infection was found in 12 of the 15 initial specimens (Table 1), all of which were from fatal cases.

Clearly the most sensitive diagnostic test was the Q-RT-PCR assay, which was designed to detect a conserved sequence of VP40 present in all known species of marburgvirus, including the more distantly related Ravn strain. All positive Q-RT-PCR results were obtained in duplicate, with each set yielding amplification curves within 0.4 C_T units. Virus isolation proved to be nearly as sensitive as the Q-RT-PCR assay, although limited clinical material in some cases made virus isolation attempts difficult. The antigen capture assay confirmed the RT-PCR results in four cases yet overall proved less sensitive. None of the initial samples were positive in either the marburgvirus-specific IgM or IgG assays performed in parallel. In those ELISAs, the positive and negative controls in each assay performed as expected, thus suggesting that the negative results with the human specimens were not the result of any systematic errors. In no instances did specimens test positive by an-

TABLE 1. Summary of marburgvirus diagnostic test results for samples sent to CDC from Angola in March 2005

Specimen no. (last four digits)	Patient age	Specimen type	Diagnostic test results					
			Ag^a	IgM	IgG	Q-RT-PCR[b]	Virus isolation[c]	IHC staining
1377	12 yr	Serum	−	−	−	++	+	ND[d]
1378	16 mo	Serum	−	−	−	++	+	ND
1379	8 mo	Serum	+++	−	−	++++	+	ND
1380	3 yr	Serum	++	−	−	++++	+	ND
1381	9 mo	Serum	−	−	−	+++	+	ND
1382	NA[e]	Serum	−	−	−	−	−	ND
1383	NA	Serum	−	−	−	−	−	ND
1384	NA	Serum	−	−	−	−	−	ND
1385	24 yr	Serum	−	−	−	+	+	ND
1386	5 yr	Serum	+	−	−	++	+	ND
1387	5 yr	Serum	−	−	−	+++	+	ND
1388	NA	Serum	+	−	−	++	−	ND
1411[f]	51 yr	Blood	−	−	−	+++	+	ND
1412[f]	51 yr	Serum	−	−	−	++	+	ND
1413	NA	Skin	ND	ND	ND	ND	ND	+

[a] Antigen capture assay. +++, sum optical density (23) of greater than 2.0; ++, sum optical density of 1.0 to 2.0; +, sum optical density of 0.45 (threshold) to 1.0.
[b] ++++, C_T values of less than 20; +++, C_T values of 20 to 25; ++, C_T values of 25 to 30; +, C_T values of 30 to 35.
[c] Virus isolations were considered positive after cultures of Vero E6 cells showed cytopathic effect and subsequent scrapes were positive by immunofluorescence assay.
[d] ND, not determined.
[e] NA, not available.
[f] Case in the capital, Luanda.

tigen capture, IgM, IgG, or virus isolation and not test positive by RT-PCR (data not shown).

Three samples (samples 1411 to 1413 in Table 1) arrived the following week. One specimen was a skin biopsy from a patient in Uige, while the other two samples, 1411 and 1412 (a blood specimen and a serum specimen, respectively, drawn at the same time), were from a physician in Luanda with a recent contact history with the suspected VHF patients in Uige Province. The physician traveled directly to the capitol city of Luanda only days before her death. Both samples tested positive by Q-RT-PCR and virus isolation, while the skin biopsy tested positive by IHC staining.

Complete genome sequence determination and phylogenetic analysis. To establish the genetic relationship of the Angola marburgvirus relative to prior East African viruses, and to determine the extent of sequence variation circulating within the Angola VHF outbreak, complete marburgvirus genomes were sequenced from 11 clinical samples collected from fatal cases throughout a time period of almost 3 months and encompassing the known geographic distribution of the outbreak. There were four municipalities within Uige Province from which virus-positive samples were collected (Fig. 1). The municipalities were Uige (samples 1379 to 1386 and 0998), Songo (sample 0754), Bungo (samples 0181, 0214, and 0215), and Damba (sample 0126). In addition, we included the sample from the Angolan capitol, Luanda (sample 1411) (Fig. 1). RNAs were extracted from multiple types of clinical specimens, including oral swabs, which collectively contained a wide range of viral loads. It should be noted that following the initial Marburg VHF diagnosis, field labs were established by CDC and the Public Health Agency of Canada on site in Luanda and Uige, respectively, at the request of the Ministry of Health of Angola and WHO. Samples 0126, 0181, 0214, 0215, 0754, and 0998 (Table 2) were obtained and tested

within 24 h of specimen collection and then stored frozen until nucleotide sequence analysis could be initiated.

To provide a more complete representation of the overall potential for genetic diversity within a single marburgvirus outbreak and to provide a basis to which to compare the genetic diversity (or lack thereof) circulating in the Angola outbreak, we determined the complete genome sequences of four marburgvirus isolates obtained from the two most recent marburgvirus outbreaks prior to the one in Angola. Three of

TABLE 2. Summary of marburgvirus sequencing from cases representing the geographic, temporal, and viral load distributions seen in the outbreak

Specimen no. (last four digits)	Specimen type	Q-RT-PCR[a]	Date of:		Region
			Onset	Collection	
1379[b]	Serum	++++	24 February	13 March	Uige
1380	Serum	++++	10 March	13 March	Uige
1381	Serum	+++	1 March	13 March	Uige
1386	Serum	++	6 March	11 March	Uige
1411	Blood	+++	16 March	20 March	Luanda
0126	Blood	++	2 April	8 April	Damba
0181	Oral swab	+++	Unknown	13 April	Bungo
0214	Oral swab	+	Unknown	15 April	Bungo
0215	Oral swab	++	2 April	15 April	Bungo
0754	Oral swab	+++	13 April	22 April	Songo
0998	Oral swab	+++	15 May	17 May	Uige

[a] The scale for the Q-RT-PCR assay is the same as that described in Table 1, footnote b, with the caveat that the Q-RT-PCR assay used on samples 0126, 0181, 0214, 0215, 0754, and 0998 had been redesigned, as described in Materials and Methods, to correct nucleotide mismatches between the primers and probe to match the Angola-specific target sequences. The minor modifications resulted in a decrease of 1 to 2 C_T units for clinical samples tested side by side with the original (more generic) Q-RT-PCR assay.
[b] Reference specimen.

A.

B.

	Rav	09DRC	05DRC	07DRC	Ozo	Pop	Mus	Ang	EboZ	EboS
Rav	0.00	0.52	21.16	21.22	21.03	21.33	21.09	21.34	68.27	67.88
09DRC		0.00	21.24	21.30	21.11	21.41	21.18	21.41	68.28	67.87
05DRC			0.00	0.80	1.48	7.50	7.22	7.80	67.19	67.42
07DRC				0.00	1.51	7.51	7.27	7.77	67.25	67.51
Ozo					0.00	7.07	6.85	7.47	67.22	67.55
Pop						0.00	5.94	7.14	67.60	67.61
Mus							0.00	6.83	67.43	67.54
Ang								0.00	67.46	67.45
EboZ									0.00	40.47
EboS										0.00
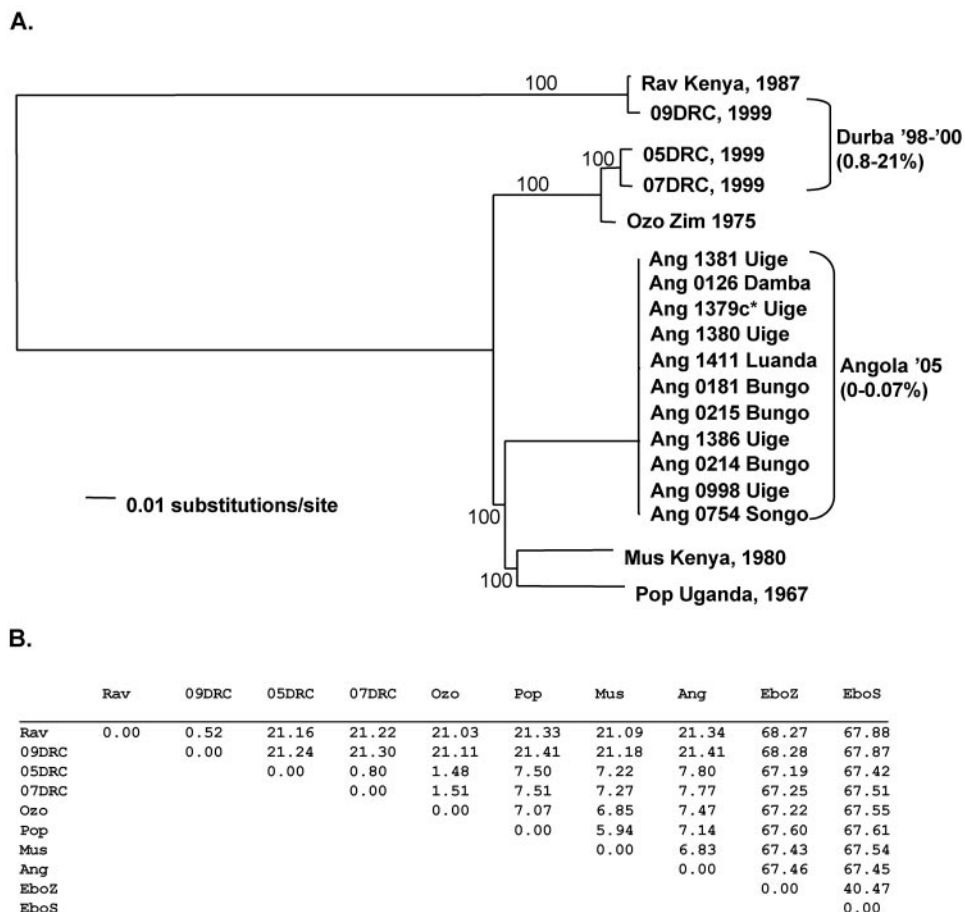
FIG. 2. (A) Maximum-likelihood analysis of full-length genomes with indicated bootstrap values. *, reference specimen. (B) Nucleotide diversity, comparing full-length genome sequences of the indicated filoviruses. The Angola marburgvirus sequence was represented by that determined from the reference specimen 1379c.

the isolates, 05DRC, 07DRC, and 09DRC, were from epidemiologically unlinked cases that occurred at different times during the 1998 outbreak in Durba, DRC, while the fourth isolate, Ravn (Rav), was obtained in 1987 from a 15-year-old Danish boy with a recent history of travel to Kitum Cave in Mount Elgon, Kenya, 9 days prior to disease onset.

All 16 newly determined complete marburgvirus sequences (19,114 nucleotides in length), in addition to those of Popp (Pop), Musoke (Mus), and Ozolin (Ozo), were analyzed to determine their phylogenetic relationship (Fig. 2A) and nucleotide distances (Fig. 2B). A maximum-likelihood analysis placed the Angola (Ang) isolates firmly (100% bootstrap support) within the clade containing the majority of the East African isolates. This result was surprising given the large geographic distance, ~1,000 miles, of Uige, Angola, from the locations of all other known sources of marburgvirus. The branching pattern of the phylogenetic tree shows five overall branches, with the Rav/09DRC branch being the most divergent, showing nucleotide differences greater than 21% relative to viruses in other lineages, including Ang. 05DRC, 07DRC, and Ozo comprise a second well-defined lineage, differing from viruses in all other lineages by greater than 7%. Finally, using a cutoff of ~5% nucleotide difference, Ang, Mus, and Pop

form three additional lineages, with the Angola sequence diverging by 6.8 and 7.1% from Mus and Pop, respectively, and by greater than 7.4% from the Ozo, 05DRC, and 07DRC lineages. Mus and Pop differ from each other by 5.9%. For comparison, the difference between any of the marburgvirus genomes and those of either Zaire ebolavirus or Sudan ebolavirus is greater than 65%, while the Zaire and Sudan ebolaviruses differ from each other by greater than 40%.

Among the 11 clinical samples selected from the Angola outbreak, the genetic sequences were well conserved throughout the entire length (19,114 nucleotides) of the genome (Fig. 3). Ten of 11 genomes had five changes or fewer compared to the reference isolate Ang1379c. Remarkably, four of the sequences, obtained from clinical samples collected over a month and a half, showed no nucleotide differences for the entire 19,114-nucleotide genome. This confirmed the high fidelity of the RT-PCR-based sequence analysis performed and demonstrated that human-to-human passage of marburgvirus could occur in the absence of virus evolution. In addition, our reference sequence, Ang1379c, was also 100% identical to the sequence obtained from the corresponding virus isolate, Ang1379v, indicating the lack of selection occurring during culture of the virus from the clinical specimen. The most ge-

Sample

Consensus nt	T	G	A	G	A	T	T	G	G	C	A	A	T	T	G	A	A	T
1379c*	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1379v (0)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1380 (0)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1381 (2)	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A
1386 (1)	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
1411 (0)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
0126 (2)	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	A
0181 (0)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
0214 (3)	.	.	T	.	.	.	.	A	.	G	.	.	.	.	.	.	.	.
0215 (2)	C	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.
0754 (11)	.	T	.	A	G	C	C	.	A	.	.	.	C	C	A	.	G	A
0998 (5)	.	.	.	.	.	.	.	.	G	G	.	.	.	C	.	G	.	A
Position	27	1500	1589	3576	4087	4620	4645	7259	8202	8523	8763	9006	9038	10852	13099	16011	16273	17248
Location	UTR	NP	NP	VP35	UTR	VP40	VP40	GP	UTR	UTR	UTR	VP30	VP30	VP24	L	L	L	L
Amino acid change†			Q to L									S to G		L to S		K to R		
PCR Fragment	F	A	A	A	A/B	B	B	B	B/C	C	C	C	C	D	D	E	E	E

Fragments (nt positions)
A 375-4641
B 4155-8304
C 7567-10571
D 10287-15279
E 14817-19113
F 3-801

*Reference specimen
Numbers in parentheses () indicate total number of changes relative to the reference specimen.
† Except those noted, all nucleotide changes within coding sequence resulted in silent mutations.

FIG. 3. Summary of single-nucleotide (nt) changes among the 11 indicated marburgvirus sequences from the Angola outbreak. Beneath the position descriptions of each change are the RT-PCR fragment(s) used to generate the sequence.

netically diverse genome came from a specimen collected in the Songo municipality (specimen 0754), which had a total of 11 nucleotide changes (out of 19,114 bases) relative to the reference isolate (0.07% variation). Each of the unique changes were independently confirmed by generation of small (<1-kb) RT-PCR fragments followed by sequence analysis. In contrast, the analysis of three isolates from within the earlier Durba, DRC, outbreak showed at least 10 times greater sequence diversity, ranging from 0.8 to 21% nucleotide difference. The 05DRC and 07DRC isolates differed by 0.8% (>150 nucleotide differences) and differed by over 21% when each was compared to the 09DRC isolate. The 09DRC isolate is noteworthy because it represents the second member of the most distinct lineage within the marburgviruses, first defined by the Rav isolate (20). The minimal genetic diversity observed in the Angola outbreak (maximum of 0.07% variation) relative to that seen in the Durba (0.8 to 21% variation) is consistent with the Angola marburgvirus outbreak being the result of a rare introduction of virus into the human population from the unknown reservoir followed by direct human-to-human transmission.

Comparative analysis of marburgvirus genetic elements. The determination of the Ang, Rav, 05DRC, 07DRC, and 09DRC full-length sequences almost tripled the number of marburgvirus genomes available for analysis. Therefore, using
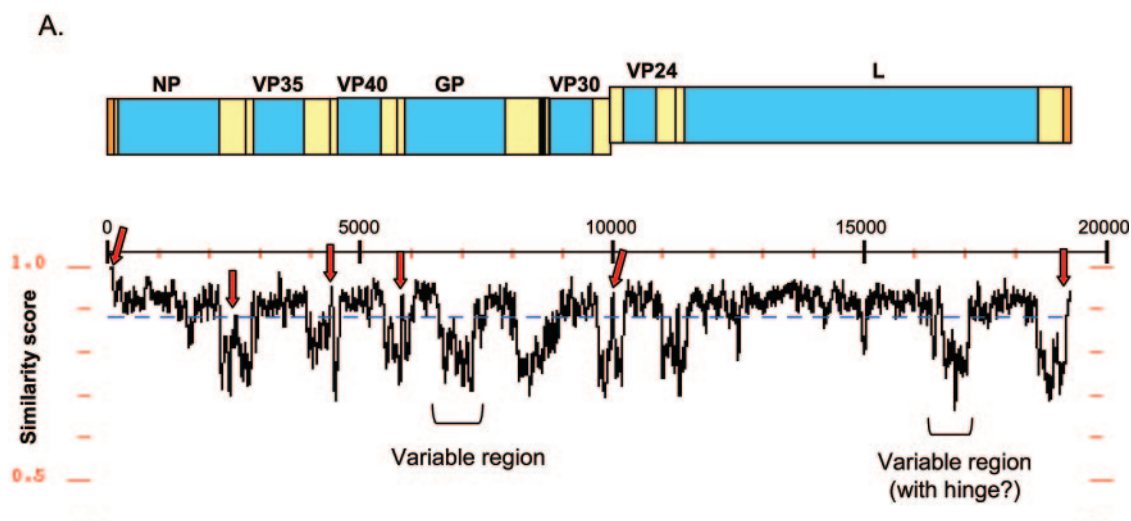
FIG. 4. Similarity plot of an alignment of eight marburgvirus sequences (Ang, Pop, Mus, Ozo, Rav, 05DRC, 07DRC, and 09DRC), using a sliding window of 50 nucleotides. The Angola marburgvirus sequence was represented by that determined from the reference specimen 1379c.

Gene NP	Virus	Start 3'-5'	Stop 3'-5'	IR (NP-VP35)	3'UTR nt	5'UTR nt
	Ang	CUUCUUAUAAUU	UAAUUCUUUUU	GAUUUUU	57	653
	Mus	CUUCUUAUAAUU	UAAUUCUUUUU	GAUUUUU	57	653
	Pop	CUUCUUAUAAUU	UAAUUCUUUUU	GAUUUUU	57	653
	Ozo	CUUCUUAUAAUU	UAAUUCUUUUU	GAUUUUU	57	653
	05DRC	CUUCUUAUAAUU	UAAUUCUUUUU	GAUUUUU	57	653
	07DRC	CUUCUUAUAAUU	UAAUUCUUUUU	GAUUUUU	57	653
	09DRC	CUUCUUAUAAUU	UAAUUCUUUUU	GUUUCUC	57	653
	Rav	CUUCUUAUAAUU	UAGUUCUUUUU	GUUUCUC	57	653

Gene VP35	Virus	Start 3'-5'	Stop 3'-5'	IR (VP35-VP40)	3'UTR nt	5'UTR nt
	Ang	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476
	Mus	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476
	Pop	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476
	Ozo	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476
	05DRC	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476
	07DRC	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476
	09DRC	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476
	Rav	CUUCUUAUAAUU	UAAUUCUUUUU	GA	93	476

Gene VP40	Virus	Start 3'-5'	Stop 3'-5'	IR (VP40-GP)	3'UTR nt	5'UTR nt
	Ang	CUUCUUGUAAUU	UCAUUCUUUUU	GAU	155	340
	Mus	CUUCUUGUAAUU	UCAUUCUUUUU	GAU	155	340
	Pop	CUUCUUGUAAUU	UCAUUCUUUUU	GAU	155	340
	Ozo	CUUCUUGUAAUU	UCAUUCUUUUU	GAU	155	340
	05DRC	CUUCUUGUAAUU	UCAUUCUUUUU	GAU	155	340
	07DRC	CUUCUUGUAAUU	UCAUUCUUUUU	GAU	155	340
	09DRC	CUUCUUGUAAUU	UUAUUCUUUUU	GAU	155	340
	Rav	CUUCUUGUAAUU	UUAUUCUUUUU	GAU	155	340

Gene GP	Virus	Start 3'-5'	Stop 3'-5'	IR (GP-VP30)	3'UTR nt	5'UTR nt
	Ang	CUUCUUGUAAUU	UAAUUCUUUUU	$CUG(X)_{92}$	118	684
	Mus	CUUCUUGUAAUU	UAAUUCUUUUU	$CUA(X)_{92}$	118	684
	Pop	CUUCUUGUAAUU	UAAUUCUUUUU	$CUA(X)_{92}$	118	684
	Ozo	CUUCUUGUAAUU	UAAUUCUUUUU	$CUA(X)_{92}$	118	684
	05DRC	CUUCUUGUAAUU	UAAUUCUUUUU	$CUA(X)_{92}$	118	684
	07DRC	CUUCUUGUAAUU	UAAUUCUUUUU	$CUA(X)_{92}$	118	684
	09DRC	CUUCUUGUAAUU	UAAUUCUUUUU	$CUA(X)_{92}$	118	684
	Rav	CUUCUUGUAAUU	UAAUUCUUUUU	$CUA(X)_{92}$	118	684

Gene VP30	Virus	Start 3'-5'	Stop 3'-5'	IR (VP30-VP24)	3'UTR nt	5'UTR nt
	Ang	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	103	302
	Mus	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	103	302
	Pop	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	102	302
	Ozo	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	103	302
	05DRC	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	103	302
	07DRC	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	103	302
	09DRC	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	103	302
	Rav	CUUCUUGUAAUU	UAAUUCUUUUU	Overlap	103	302

Gene VP24	Virus	Start 3'-5'	Stop 3'-5'	IR (VP24-L)	3'UTR nt	5'UTR nt
	Ang	CUUCUUGUAAUU	UAAUUCUUUUU	AGU	210	317
	Mus	CUUCUUGUAAUU	UAAUUCUUUUU	AGU	210	317
	Pop	CUUCUUGUAAUU	UAAUUCUUUUU	GGU	210	317
	Ozo	CUUCUUGUAAUU	UAAUUCUUUUU	AGU	210	317
	05DRC	CUUCUUGUAAUU	UAAUUCUUUUU	AGU	210	317
	07DRC	CUUCUUGUAAUU	UAAUUCUUUUU	AGU	210	317
	09DRC	CUUCUUGUAAUU	UAAUUCUUUUU	CGU	210	317
	Rav	CUUCUUGUAAUU	UAAUUCUUUUU	CGU	210	317

Gene L	Virus	Start 3'-5'	Stop 3'-5'	IR	3'UTR nt	5'UTR nt
	Ang	CUACCUAUAAUU	UAAUUCUUUUU		192	561
	Mus	CUACCUAUAAUU	UAAUUCUUUUU		192	559
	Pop	CUACCUAUAAUU	UAAUUCUUUUU		192	561
	Ozo	CUACCUAUAAUU	UAAUUCUUUUU		192	561
	05DRC	CUACCUAUAAUU	UAAUUCUUUUU		192	561
	07DRC	CUACCUAUAAUU	UAAUUCUUUUU		192	561
	09DRC	CUACCUAUAAUU	UAAUUCUUUUU		191	574
	Rav	CUACCUAUAAUU	UAAUUCUUUUU		191	574

FIG. 5. Alignment of the transcription start and stop signals, IR sequences, and lengths of the 5' and 3' untranslated regions (UTRs) for each marburgvirus sequences. Shown in red are nucleotides (nt) or UTR lengths that are different from the consensus. The Angola marburgvirus sequence was represented by that determined from the reference specimen 1379c.

this more extensive database of full-length sequences, a comprehensive effort was undertaken to reexamine many genetic features throughout the genome and determine the degree to which these elements are maintained across all eight marburgvirus strains. An initial macroscopic perspective of genome similarity is shown in Fig. 4, in which full-length genomes were analyzed for similarity using a sliding window of 50 nucleotides. The similarity plot reveals a striking pattern of sequence conservation among the open reading frames flanked by regions of much greater variation in the noncoding sequences. However, the more variable noncoding regions are punctuated with spikes of high identity corresponding to the regions containing transcription start and stop sequences (arrows). Alignments of the individual *cis*-acting regulatory features demonstrate that 12 of 14 transcription start and stop sequences are 100% identical (Fig. 5). Of the two start/stop sequences that show variation, the differences are merely single-nucleotide transitions and are within the Ravn/09DRC lineages. In all

A

NP nt

		1	2	3	4	5	6	7	8
Rav	1	0.00	0.43	14.46	14.46	14.32	14.27	14.42	13.70
09DRC	2		0.00	14.56	14.56	14.42	14.37	14.51	13.89
05DRC	3			0.00	0.57	0.86	5.12	5.08	5.46
07DRC	4				0.00	1.15	5.41	5.36	5.75
Ozo	5					0.00	5.27	4.98	5.27
Pop	6						0.00	4.69	5.65
Mus	7							0.00	5.80
Ang	8								0.00

VP35 nt

		1	2	3	4	5	6	7	8
Rav	1	0.00	0.20	14.65	14.95	14.55	14.85	14.44	15.25
09DRC	2		0.00	14.65	14.95	14.55	14.85	14.44	15.25
05DRC	3			0.00	0.40	0.81	5.76	4.14	5.45
07DRC	4				0.00	0.81	5.76	4.14	5.45
Ozo	5					0.00	5.15	3.54	5.25
Pop	6						0.00	4.04	5.56
Mus	7							0.00	3.94
Ang	8								0.00

VP40 nt

		1	2	3	4	5	6	7	8
Rav	1	0.00	0.33	12.83	12.61	12.50	14.14	14.04	14.14
09DRC	2		0.00	12.72	12.50	12.39	14.04	13.71	13.82
05DRC	3			0.00	0.66	0.88	5.59	5.48	6.03
07DRC	4				0.00	1.10	5.81	5.70	5.70
Ozo	5					0.00	4.93	4.82	5.15
Pop	6						0.00	3.95	5.15
Mus	7							0.00	5.37
Ang	8								0.00

GP nt

		1	2	3	4	5	6	7	8
Rav	1	0.00	0.73	22.29	21.99	21.85	21.55	22.19	22.04
09DRC	2		0.00	22.48	22.24	22.09	21.85	22.29	22.24
05DRC	3			0.00	1.08	1.81	7.97	8.75	9.09
07DRC	4				0.00	1.71	8.02	8.70	9.19
Ozo	5					0.00	7.58	8.31	8.99
Pop	6						0.00	6.45	7.62
Mus	7							0.00	7.62
Ang	8								0.00

VP30 nt

		1	2	3	4	5	6	7	8
Rav	1	0.00	0.35	16.08	16.31	16.31	17.38	16.90	16.90
09DRC	2		0.00	16.08	16.31	16.31	17.38	16.90	16.90
05DRC	3			0.00	0.71	0.47	7.09	5.56	6.38
07DRC	4				0.00	0.95	7.21	5.79	6.38
Ozo	5					0.00	6.62	5.08	5.91
Pop	6						0.00	4.61	5.91
Mus	7							0.00	4.73
Ang	8								0.00

VP24 nt

		1	2	3	4	5	6	7	8
Rav	1	0.00	0.39	13.52	13.65	13.12	13.39	13.25	13.91
09DRC	2		0.00	13.91	14.04	13.52	13.52	13.65	14.04
05DRC	3			0.00	0.13	0.66	4.59	5.12	5.64
07DRC	4				0.00	0.52	4.46	4.99	5.77
Ozo	5					0.00	3.94	4.46	5.25
Pop	6						0.00	3.02	4.86
Mus	7							0.00	5.12
Ang	8								0.00

L nt

		1	2	3	4	5	6	7	8
Rav	1	0.00	0.52	21.16	21.22	21.03	21.34	21.09	21.34
09DRC	2		0.00	21.24	21.30	21.11	21.42	21.18	21.41
05DRC	3			0.00	0.80	1.48	7.50	7.24	7.80
07DRC	4				0.00	1.51	7.51	7.29	7.77
Ozo	5					0.00	7.07	6.87	7.47
Pop	6						0.00	5.94	7.14
Mus	7							0.00	6.84
Ang	8								0.00

FIG. 6. Percent nucleotide (nt) (A) or amino acid (B) differences for each of the seven marburgvirus gene products. The Angola marburgvirus sequence was represented by that determined from the reference specimen 1379c.

genomic elements examined, the Angola sequence shows 100% identity with the consensus sequence. Well conserved genomic features include the lengths of the 5′ and 3′ untranslated regions of all seven predicted mRNAs, some of which differ slightly from previous reports (8, 15), as well as the length and composition of the six intergenic (IR) sequences. A few notable variations within the IRs are a requirement for a purine residue in the first position of the trinucleotide IR between VP24 and L and conserved differences within the Ravn/09DRC lineage at three of seven positions in the IR between NP and VP35.

An examination of the nucleotide and amino acid distances for each of the seven marburgvirus gene products is shown in Fig. 6A and B, respectively. At the nucleotide level the most conserved genes are, in order, VP40, NP, VP24, and VP35, showing 0.2 to 15.2% variation, closely followed by VP30, showing 0.3 to 17.4% variation. The gene with the greatest nucleotide differences is GP (0.7 to 22.5%), consistent with previous alignments using fewer marburgvirus strains (8, 20, 45). Surprisingly, the nucleotide difference in the polymerase (0.5 to 21.4%) is almost as much as that seen in the GP

B

NP aa	1	2	3	4	5	6	7	8
Rav 1	0.00	0.14	5.61	5.76	5.47	5.04	5.61	5.18
09DRC 2		0.00	5.76	5.90	5.61	5.18	5.76	5.32
05DRC 3			0.00	0.14	0.14	2.01	1.87	1.73
07DRC 4				0.00	0.29	2.16	2.01	1.87
Ozo 5					0.00	1.87	1.73	1.58
Pop 6						0.00	1.73	1.29
Mus 7							0.00	1.73
Ang 8								0.00

VP35 aa	1	2	3	4	5	6	7	8
Rav 1	0.00	0.00	5.17	5.17	4.86	5.47	5.17	5.47
09DRC 2		0.00	5.17	5.17	4.86	5.47	5.17	5.47
05DRC 3			0.00	0.00	0.30	1.82	0.61	1.52
07DRC 4				0.00	0.30	1.82	0.61	1.52
Ozo 5					0.00	1.52	0.30	1.22
Pop 6						0.00	1.82	2.74
Mus 7							0.00	1.52
Ang 8								0.00

VP40 aa	1	2	3	4	5	6	7	8
Rav 1	0.00	0.00	1.65	1.65	1.32	1.65	1.32	0.99
09DRC 2		0.00	1.65	1.65	1.32	1.65	1.32	0.99
05DRC 3			0.00	0.00	0.33	1.65	1.32	1.65
07DRC 4				0.00	0.33	1.65	1.32	1.65
Ozo 5					0.00	1.32	0.99	1.32
Pop 6						0.00	0.33	0.66
Mus 7							0.00	0.33
Ang 8								0.00

GP aa	1	2	3	4	5	6	7	8
Rav 1	0.00	1.17	22.91	22.91	22.76	22.32	21.88	22.17
09DRC 2		0.00	23.49	23.49	23.35	22.47	22.32	22.61
05DRC 3			0.00	1.03	1.91	8.22	8.37	8.22
07DRC 4				0.00	1.62	8.37	8.37	8.52
Ozo 5					0.00	8.52	8.66	8.66
Pop 6						0.00	6.46	7.05
Mus 7							0.00	7.20
Ang 8								0.00

VP30 aa	1	2	3	4	5	6	7	8
Rav 1	0.00	0.00	8.19	8.19	8.19	9.61	9.96	8.90
09DRC 2		0.00	8.19	8.19	8.19	9.61	9.96	8.90
05DRC 3			0.00	1.07	0.00	4.63	4.27	2.49
07DRC 4				0.00	1.07	4.98	4.63	2.85
Ozo 5					0.00	4.63	4.27	2.49
Pop 6						0.00	4.63	3.20
Mus 7							0.00	1.78
Ang 8								0.00

VP24 aa	1	2	3	4	5	6	7	8
Rav 1	0.00	0.00	4.35	4.35	3.95	4.35	3.95	3.95
09DRC 2		0.00	4.35	4.35	3.95	4.35	3.95	3.95
05DRC 3			0.00	0.00	0.40	1.58	1.19	1.58
07DRC 4				0.00	0.40	1.58	1.19	1.58
Ozo 5					0.00	1.19	0.79	1.19
Pop 6						0.00	1.19	1.58
Mus 7							0.00	1.19
Ang 8								0.00

L aa	1	2	3	4	5	6	7	8
Rav 1	0.00	0.56	11.69	11.86	11.65	12.29	11.17	11.17
09DRC 2		0.00	11.82	11.99	11.77	12.29	11.22	11.26
05DRC 3			0.00	0.39	0.73	4.33	3.78	3.65
07DRC 4				0.00	0.69	4.25	3.90	3.60
Ozo 5					0.00	4.08	3.65	3.43
Pop 6						0.00	3.95	3.82
Mus 7							0.00	3.22
Ang 8								0.00

FIG. 6—*Continued.*

region despite having some stretches of very high conservation (Fig. 7G).

At the amino acid level, the percent differences among the marburgvirus strains for each of the seven gene products are quite different from those seen within the same genes at the nucleotide level. The degrees of variation within all the open reading frames, except GP, are decreased by 2- to 10-fold. The most striking example is the matrix protein VP40, in which the decrease is about 10-fold, demonstrating a distinct intolerance for amino acid changes (1.65% maximum variation). This intolerance is suggestive of tight physical constraints for VP40 in the virus assembly process. At the other end of the spectrum, GP showed no decrease at all in the level of amino acid variation (~23%) relative to that seen at the nucleotide level, suggesting a selective pressure for nonsynonymous changes, most likely exerted by the immune system of the natural reservoir host(s).

We next examined known protein domains and motifs in each of the seven open reading frames by comparative alignment of the amino acid sequences (Fig. 7A to G). The analysis of GP (Fig. 7D) shows that the area of greatest diversity is a continuous 300-amino-acid (aa) stretch from residue 201 to

A

NP alignment

```
        1                                                                                          100
Rav                                                                                           p
09DRC                                                                                         p
05DRC
07DRC
Ozo
Pop                                                                                           v
Mus
Ang
con     MDLHSLLELG TKPTAPHVRN KKVILFDTNH QVSICNQIID AINSGIDLGD LLEGGLLTLC VEHYYNSDKD KFNTSPIAKY LRDAGYEFDV IKNADATRFL
        101                                                                                        200
Rav      e
09DRCe
05DRC
07DRC
Ozo
Pop
Mus
Ang
con     DVIPNEPHYS PLILALKTLE STESQRGRIG LFLSFCSLFL PKLVVGO5DRCS IEKALRQVTV HQEQGIVTYP NHWLTTGHMK VIFGILRSSF ILKFVLIHQG
        201                                                                                        300
Rav                                             a
09DRC                                           a
05DRC
07DRC
Ozo
Pop                                        i
Mus
Ang
con     VNLVTGHDAY DSIISNSVGQ TRFSGLLIVK TVLEFILQKT DSGVTLHPLV RTSKVKNEVA SFKQALSNLA RHGEYAPFAR VLNLSGINNL EHGLYPQLSA
        301                                                                                        400
Rav                                                                                           a
09DRC                                                                                         a
05DRC
07DRC
Ozo
Pop
Mus
Ang
con     IALGVATAHG STLAGVNVGE QYQQLREAAH DAEVKLQRRH EHQEIQAIAE DDEERKILEQ FHLQKTEITH SQTLAVLSQK REKLARLAAE IENNIVEDQG
        401                                                                                        500
Rav          q                    s              n         te                        en     a    l  pkev  s  l  t
09DRC        q                    s              n         te                        en     a    l  pkev  s  l  t
05DRC                       v                              g                                      s  g
07DRC                       v                              g                                      s  g
Ozo                        v                                                                      s  g
Pop          q                                  n                                          a
Mus                              m                    s                             g
Ang                                                                                                             i
con     FKQSQNRVSQ SFLNDETPVE VTVQARPINR PTALPPPVDD KIEHHSTEDS SSSSSFVDLN DPFALLNEDE DTLDDSVMIP STTSEEFQGI PEPERQSQDL
        501                                                                                        600
Rav      i  k   gn      d ar                        d                          v              d
09DRC    i  k   gn      d ar                        d                          dv             d
05DRC                       r                                      t           a
07DRC                                                             t           a
Ozo                                                              t           a
Pop     n                                                                     v
Mus     n                                                                     anp
Ang                   l  p                                                    v
con     DNSQGKQEDE STNPIKKQFL RYQELPPVQE DDPSEYETDS QESIDQPGSD NEQGVDLPPP PLYAQEKRQD PIQHPA-SSQ DPFGSIGDVN GDILEPIRSF
        601                                                                             695
Rav                      mg           s
09DRC                    mg           s
05DRC                                                                  v
07DRC                                                                  v
Ozo                                                                    v
Pop                                                            n
Mus                   m            d                           n
Ang                                                            n
con     SSPSAPQEDT RAREAYELSE DFTNYEDNQQ NWPQRVVTKK GRTFLYPNDL LQTSPPESLI TALVEEYQNP VSAKELQADW PDMSFDERRH VAMNL
```

FIG. 7. Comparative amino acid alignments for each of the seven marburgvirus gene products. The Angola marburgvirus sequence was represented by that determined from the reference specimen 1379c. (A) NP. Shown are conserved cysteines (underlined and in boldface) and serine/threonine phosphorylation motifs (boxed) residing in domains previously defined (28) (gray) when queried with the Angola marburgvirus NP amino acid sequences for the kinase substrates casein kinase II (S/TxxD/E), protein kinase C (S/TxK/R), protein kinase A (RxS/T), glycogen synthase kinase 3 (S/TxxxS/T), casein kinase I (SxxS/T), and mitogen-activated protein kinase (S/TP). (B) VP35. Shown are the predicted oligomerization domain (35) (gray) with hydrophobic residues in the first and forth positions of the heptad repeats (+), type I IFN antagonism domain (boxed) with critical basic residues (asterisks), and conserved cysteines (underlined and in boldface). (C) VP40. Shown are the L domain (PPxY) (boxed) and conserved cysteines (underlined and in boldface). (D) GP. Shown are conserved cysteines (underlined and in boldface); N-glycosylation sites (N x T/S), conserved (boxed in gray) and nonconserved (boxed only); furin cleavage sites (R x K/RR) (double underlined); fusion domain (+); immunosuppressive domain (asterisks); and transmembrane domain (parentheses). (E) VP30. Shown are phosphorylated serine residues (34) (asterisks), conserved cysteines (underlined and in boldface), Cys$_3$-His residues of the zinc binding motif (+), and tetraleucine motif for VP30 oligomerization (boxed). (F) VP24. Shown are conserved cysteines (underlined and in boldface). (G) Polymerase. Shown are conserved cysteines (underlined and in boldface, with the double-cysteine motif [36] double underlined and in boldface), ATP binding motifs (asterisks), D-D diresidue motifs (boxed), MGDNQ motif (+), box A* to C* regions (36) (light gray and labeled), domains of higher variability (light gray and labeled), and motifs A** to C** (53) (dark gray).

B

VP35 alignment

```
        1                                                                          100
Rav               t                        i  s        v                    m       m  k
09DRC             t                        i  s        v                    m       m  k
05DRC                                                                                i
07DRC                                                                                i
Ozo
Pop           s       h                         v                               n
Mus
Ang               t                                  v  r
con  MWDSSYMQQV SEGLMTGKVP IDQVFGANPL EKLYKRRKPK GTVGLQCSPC LMSKATSTDD IIWDQLIVKK TLADLLIPIN RQISDIQSTL SEVTTRVHEI
                                                                        +   +  +   +  +  +    +  +     +  +
                                                                        4   1  4   1  4  1    4  1     4  4
        101                                                                         200
Rav            d      v     k                                          y qk        q
09DRC          d      v     k                                          y qk        q
05DRC
07DRC
Ozo
Pop                                                                            e
Mus
Ang                                                                   t
con  ERQLHEITPV LKMGRTLEAI SKGMSEMLAK YDHLVISTGR TTAPAAAFDA YLNEHGVPPP QPAIFKDLGV AQQACSKGTM VKNATTDAAD KMSKVLELSE
     +   +      +   +   +
     1   4      1   4   1
        201                                                                         300
Rav                                                              a        k
09DRC                                                            a        k
05DRC
07DRC
Ozo
Pop
Mus
Ang
con  ETFSKPNLSA KDLALLLFTH LPGNNTPFHI LAQVLSKIAY KSGKSGAFLD AFHQILSEGE NAQAALTRLS RTFDAFLGVV PPVIRVKNFQ TVPRPCQKSI
                                                                                                 *    *
        301                329
Rav
09DRC
05DRC
07DRC
Ozo
Pop
Mus
Ang  *
con  RAVPPNPTID KGWVCVYSSE QGETRALKI
```

C

VP40 alignment

```
        1                                                                          100
Rav
09DRC
05DRC
07DRC
Ozo
Pop                        q                             p
Mus                        q
Ang                        q
con  MASSSNYNTY MQYLNEPPYA DHGANQLIPA DQLSNQHGIT PNYVGDLNLD DQFKGNVCHA FTLEAIIDIS AYNERTVKGV PAWLPLGIMS NFEYPLAHTV
        101                                                                         200
Rav                                      v                                         v
09DRC                                    v                                         v
05DRC                                                                      i
07DRC                                                                      i
Ozo
Pop
Mus
Ang                                                                       v
con  AALLTGSYTI TQFTHNGQKF VRVNRLGTGI PAHPLRMLRE GNQAFIQNMV IPRNFSTNQF TYNLTNLVLS VQKLPDDAWR PSKDKLIGNT MHPAVSIHPN
        201                                                                         300
Rav                                               g
09DRC                                             g
05DRC                                             s
07DRC                                             s
Ozo                                               s
Pop                                               n
Mus                                               n
Ang                                               n
con  LPPIVLPTVK KQAYRQHKNP NNGPLLAISG ILHQLRVEKV PEKTSLFRIS LPADMFSVKE GMMKKRGE-S PVVYFQAPEN FPLNGFNNRQ VVLAYANPTL
        301
Rav
09DRC
05DRC  i
07DRC  i
Ozo    i
Pop
Mus
Ang
con  SAV
```

FIG. 7—*Continued.*

D

GP alignment

```
        1                                                                                                   100
Rav     iy l        s      v       s      d
09DRC   iy l        s      v       s      d
05DRC   w    f
07DRC   r    f
Ozo     r    lf
Pop      lf                              n
Mus      l          t  n                n
Ang     11          v           i                                                        a
con     MKTTCF-ISL ILIQGIKTLP ILEIASNDQP QNVDSVCSGT LQKTEDVHLM GFTLSGQKVA DSPLEASKRW AFRTGVPPKN VEYTEGEEAK TCYNISVTDP
        101                                                                                                 200
Rav             s              v                                  v                          r
09DRC           s              v                                  v                          r
05DRC              v
07DRC              v
Ozo                v
Pop                                                      r
Mus
Ang
con     SGKSLLLDPP TNIRDYPKCK TIHHIQGQNP HAQGIALHLW GAFFLYDRIA STTMYRGKVF TEGNIAAMIV NKTVHKMIFS RQGQGYRHMN LISTNKYWTS
        201                                                                                                 300
Rav       e  r         i        n   p l k  s    vt  s hs n qin ta sg m  s          mi         g h  lnv e kq    ils  l
09DRC     e rr         i        n   s l k  s    vt  s hs n qin ta sg m  s          mi         g h  lnv e kq    ils  l
05DRC            r              t   p   h                      r
07DRC                           t   p                          r
Ozo              h              t   p                          r                  l
Pop     n             a         i   s              t    d   n    i
Mus                  a          i         l        d s       a       r  h
Ang                    k l      h   v l   s                               a
con     SNGTQTNDTG CFGTLQEYNS TKNQTCAPSK -PPPLPTARP EIKPTSTPTD ATKLNTTNPN SDDEDLTTSG SGSGEQEPYT TSDAVTKQGL SSTMPPTPSP
        301                                                                                                 400
Rav     h    s heq  s p rh      hng dp t   atlln  n tp y  l y     p p tr  i  nd  rel a s    n ql n    d      gqd  s t
09DRCh       s heq  s p rh      hng dp t   atlln  na tp y  l y     p p tr  i  nd  rel a s   an ql n    d          qd  s t
05DRC   g                a     n      pm s               e              m      k   p          s    e         n
07DRC   g                a     n      lpt s              e              m      k   p      is  e         n
Ozo     g                t     n      pt s              a              m      k   p a         e         n
Pop           e    d  gt  pnk        sm p    a      n f    vsl     yd  t             s       g l             n
Mus             kn              sm p         f    a l     d  ti         s i      g            ssk
Ang                 gv   pgk          sm p               p v i  a y   t p           s      l                s
con     QPSTPQQGGN NTNHSQDAVT ELD-TNTTAQ P--P-HNTTT ISTNNTSKHN LSTLS-PPQN TTNPNTQS-A TENEQTSAP- KTTLPPTENP TTAKSTN-TK
        401                                                                                                 500
Rav     niim tsdi  sk p ns  d ssp  rpp     k    f k   i            te       i  e         p fnt tn  e tp       f   dk g
09DRCnivm tsdi  sk p ns  d ssp  rpp     k    f k   i            te       i  e         p fnt tn  e tf       f   dk g
05DRC     m   k           s                                              s          l   s
07DRC     m                s                                             s          l   s
Ozo       m                s                              t              s          l   sg
Pop     g    a  m      l       p    v   k                                           ni
Mus     g a  a       e     p   s  a    v                                            ni
Ang         v      kys        a  v      n                                           kv
con     SPTTTEPNTT NGHFTSPSPT PNSTTQHLIY FRRKRSILWR EGDMFPFLDG LINAPIDFDP VPNTKTIFDE SSSSGASAEE DQHASPNISL TLSYFPHTNE
        501                                                                                                 600
Rav     d
09DRCd          d
05DRC
07DRC
Ozo
Pop
Mus                                            v
Ang       h                                                                                       a
con     NTAYSGENEN DCDAELRIWS VQEDDLAAGL SWIPFFGPGI EGLYTAGLIK NQNNLVCRLR RLANQTAKSL ELLLRVTTEE RTFSLINRHA IDFLLTRWGG
                              +++++ ++++++++++                                                    ****** **********
        601                                             681
Rav               k       k r       e
09DRC             k       k r       e
05DRC
07DRC
Ozo
Pop
Mus               k
Ang
con     TCKVLGPDCC IGIEDLSRNI SEQIDQIKKD EQKEGTGWGL GGKWWTSEWG VLTNLGILLL LSIAVLIALS CICRIFTKYI G
        **********                                    [             ]              ]
```
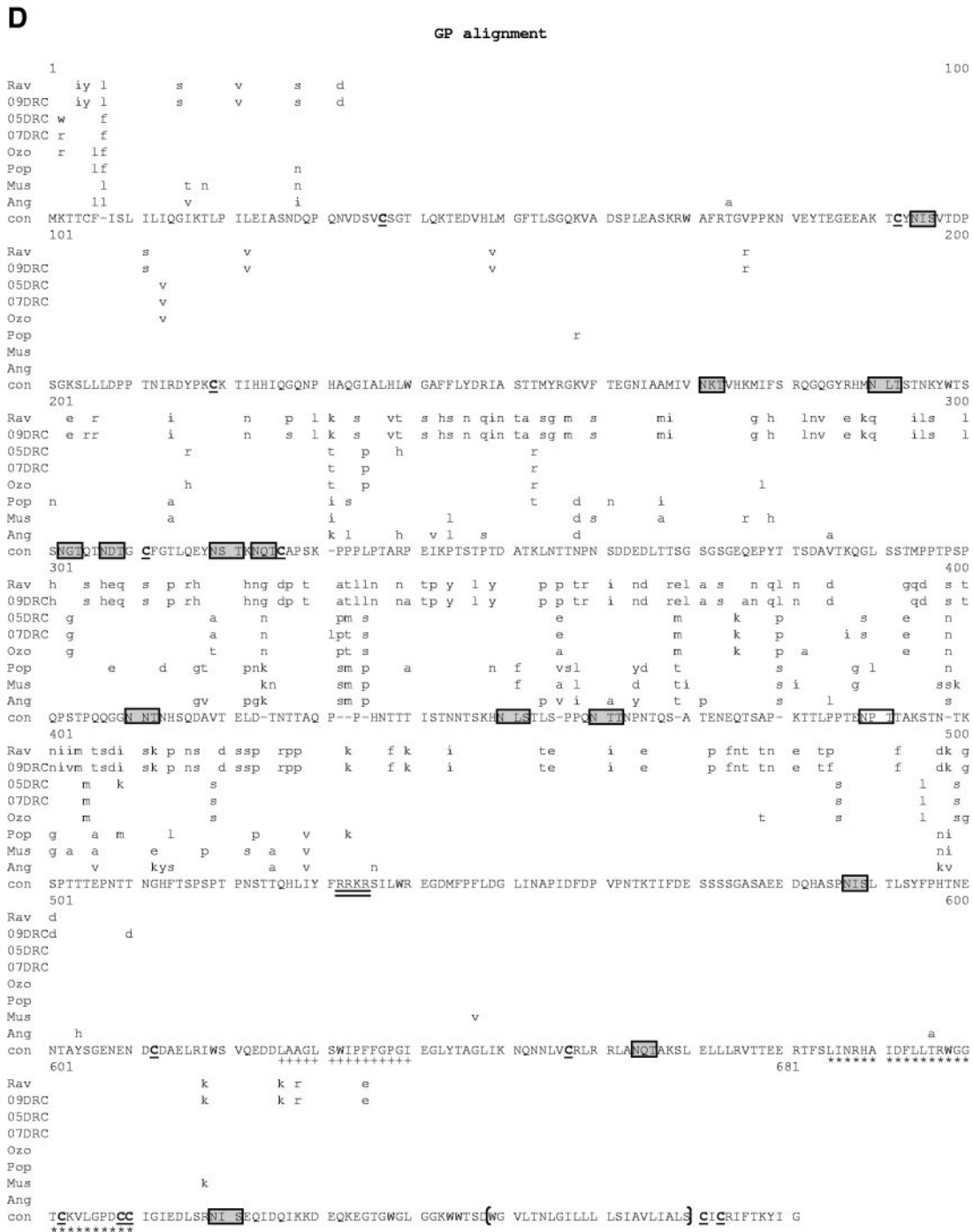
FIG. 7—*Continued*.

501, a region previously divided into two smaller domains (45). Despite this diversity, a number of previously described features remain well conserved, many of which reside within this central variable domain. These features include 13 of 14 proposed N-glycosylation sites (N-x-T/S) and 12 of 12 cysteines (45). Elsewhere in GP, the transmembrane domain (aa 649 to 670) and proposed fusion domain (aa 526 to 540) show 100% identity among all eight strains of marburgvirus, as does the furin cleavage site R-X-K/R-R (aa 432 to 435) (54).

Volchkov et al. (52) proposed the presence of an immunosuppressive domain (ISD) in GP2 of ebolavirus based on analogy to retroviruses. This 26-amino-acid motif is also present in marburgvirus (55) and shows a single Thr-to-Ala substitution at position 12 of the Angola sequence. Lending importance to the function of the ISD motif, recent studies have shown that 17-residue monomers of filovirus ISDs are capable of suppressing T-cell activation and Th1-related cytokine production in activated human and nonhuman primate peripheral blood

E

F

FIG. 7—*Continued.*

mononuclear cells (59). All other positions of the ISD show complete conservation. Another posttranslational modification ascribed to GP is the potential for phosphorylation at serine residues in two independent motifs between amino acids 260 and 273 (46). The GP alignments demonstrate that only one phosphorylation site, encompassing amino acids 268 to 273, is conserved among all eight marburgviruses. The other site, a diserine motif at amino acids 260 to 261, is not present in five of eight strains, thus discounting this motif as a general feature of all marburgviruses.

Within the alignment of marburgvirus L amino acid sequences (Fig. 7G), the regions showing the greatest variation are clustered into five domains, i.e., amino acids 114 to 135, 262 to 348, 1143 to 1206, 1623 to 1645, and 1677 to 1866. The last domain, encompassing residues 1677 to 1866, may contain a hinge region within the polymerase based on analogy to morbilliviruses, which have been shown to tolerate the insertion of the green fluorescent protein within the proposed hinge (13, 31).The areas of greatest conservation within the marburgvirus L sequences are in three large blocks, amino acids 349 to 1142, 1207 to 1622, and 1867 to 2322. The first of these three areas, amino acids 349 to 1142, contain the box A, B, and C sequences common to paramyxo- and rhabdovirus L proteins (3, 36). Boxes A and B share 100% identity among the aligned
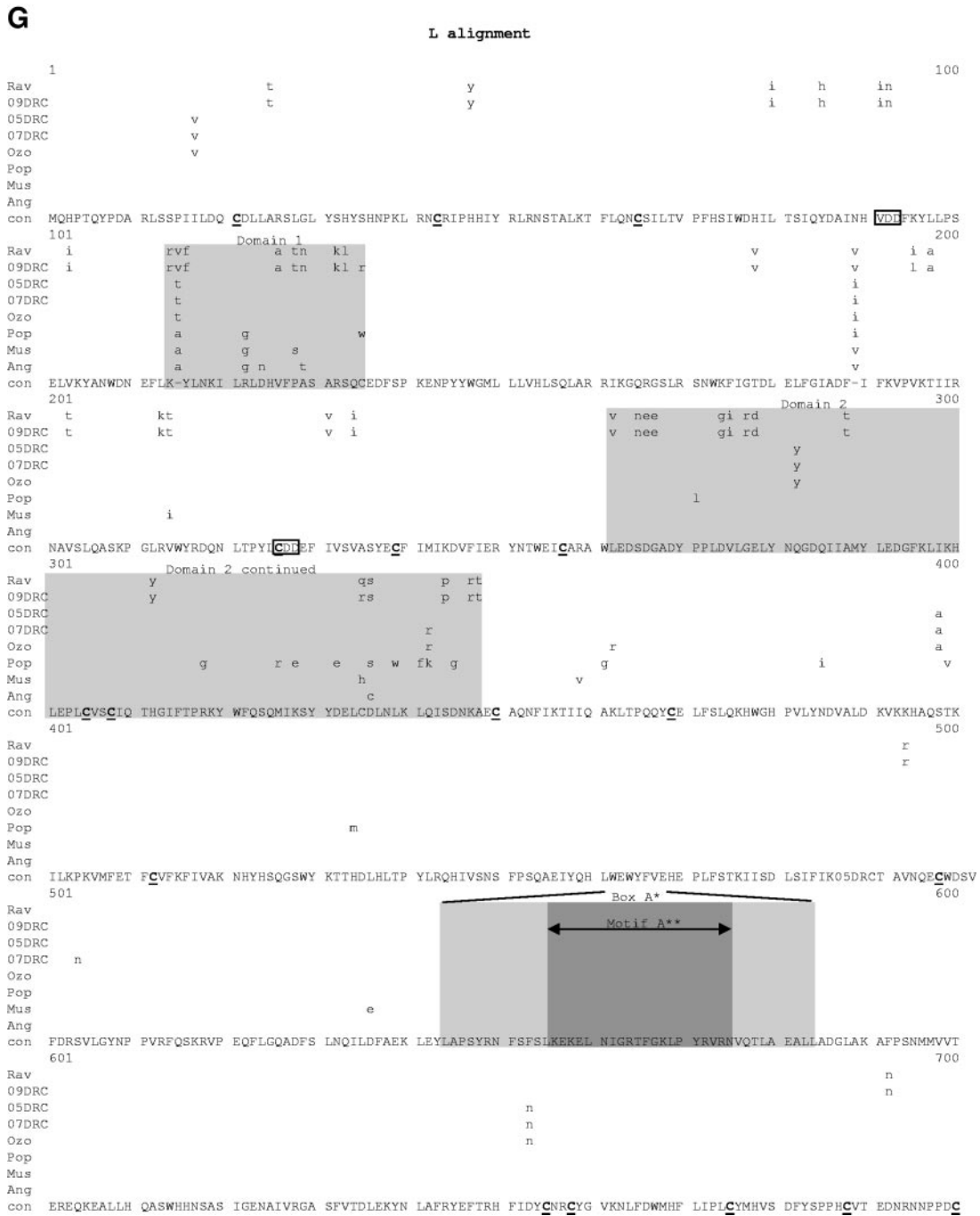
G

L alignment

```
          1                                                                              100
Rav                          t                    y                        i      h      in
09DRC                        t                    y                        i      h      in
05DRC            v
07DRC            v
Ozo              v
Pop
Mus
Ang
con   MQHPTQYPDA RLSSPIILDQ CDLLARSLGL YSHYSHNPKL RNCRIPHHIY RLRNSTALKT FLQNCSILTV PFHSIWDHIL TSIQYDAINH VDIFKYLLPS
          101                    Domain 1                                                 200
Rav   i             rvf           a tn     kl                              v          v        i  a
09DRC i             rvf           a tn     kl r                            v          v        l  a
05DRC               t                                                                          i
07DRC               t                                                                          i
Ozo                 t                                                                          i
Pop                 a       g              w                                                   i
Mus                 a       g       s                                                          v
Ang                 a       q n     t                                                          v
con   ELVKYANWDN EFLK-YLNKI LRLDHVFPAS ARSQCEDFSP KENPYYWGML LLVHLSQLAR RIKGQRGSLR SNWKFIGTDL ELFGIADF-I FKVPVKTIIR
          201                                                              Domain 2      300
Rav   t           kt                    v  i                    v  nee         gi rd           t
09DRC t           kt                    v  i                    v  nee         gi rd           t
05DRC                                                                                   y
07DRC                                                                                   y
Ozo                                                                                     y
Pop                                                                      l
Mus
Ang
con   NAVSLQASKP GLRVWYRDQN LTPYICDIEF IVSVASYECF IMIKDVFIER YNTWEICARA WLEDSDGADY PPLDVLGELY NQGDQIIAMY LEDGFKLIKH
          301              Domain 2 continued                                             400
Rav            y                         qs         p  rt
09DRC          y                         rs         p  rt
05DRC                                                                                           a
07DRC                                   r                                                       a
Ozo                                     r             r                                         a
Pop               g        r e    e  s  w   fk  g            g              i                   v
Mus                                     h                 v
Ang                                     c
con   LEPLCVSCIQ THGIFTPRKY WFQSQMIKSY YDELCDLNLK LQISDNKAEC AQNFIKTIIQ AKLTPQQYCE LFSLQKHWGH PVLYNDVALD KVKKHAQSTK
          401                                                                            500
Rav                                                                                        r
09DRC                                                                                      r
05DRC
07DRC
Ozo
Pop                            m
Mus
Ang
con   ILKPKVMFET FCVFKFIVAK NHYHSQGSWY KTTHDLHLTP YLRQHIVSNS FPSQAEIYQH LWEWYFVEHE PLFSTKIISD LSIFIK05DRCT AVNQECWDSV
          501                                            Box A*                          600
Rav                                                     Motif A**
09DRC
05DRC
07DRC n
Ozo
Pop
Mus                                     e
Ang
con   FDRSVLGYNP PVRFQSKRVP EQFLGQADFS LNQILDFAEK LEYLAPSYRN FSFSLKEKEL NIGRTFGKLP YRVRNVQTLA EALLADGLAK AFPSNMMVVT
          601                                                                            700
Rav                                                                                        n
09DRC                                                                                      n
05DRC                                               n
07DRC                                               n
Ozo                                                 n
Pop
Mus
Ang
con   EREQKEALLH QASWHHNSAS IGENAIVRGA SFVTDLEKYN LAFRYEFTRH FIDYCNRCYG VKNLFDWMHF LIPLCYMHVS DFYSPPHCVT EDNRNNPPDC
```

FIG. 7—Continued.

marburgvirus sequences, while box C has a single K-to-R substitution in the Rav and 09DRC lineages. Two other motifs purported to be present in all polymerases of negative-sense RNA viruses are the diresidue D-D and QGDNQ motifs (3, 21). These were previously identified in the Mus marburgvirus L sequence (36) and are thought to be essential components of the polymerase catalytic core. In this alignment, these potential catalytic core motifs are all 100% conserved, with the exception of one diresidue D-D motif at amino acids 91 to 93, which shows degeneracy at two positions in the Ravn and 09DRC lineages. In

addition, there is a putative ATP and/or purine ribonucleoside triphosphate binding domain found in all known L proteins of single-strand negative-sense viruses (amino acids 1931 to 1956) (36), referred to as motif C in an alignment of filovirus L sequences (53), that shows 100% identity at the core consensus glycine residues and identity at 23 of 26 positions overall within the motif. Three other potential ATP binding motifs at residues, 1325 to 1360, 1390 to 1420, and 1560 to 1593 (36), show complete identity among all taxa analyzed. Finally, 46 of 52 consensus cysteine residues are completely conserved, including the dicys-
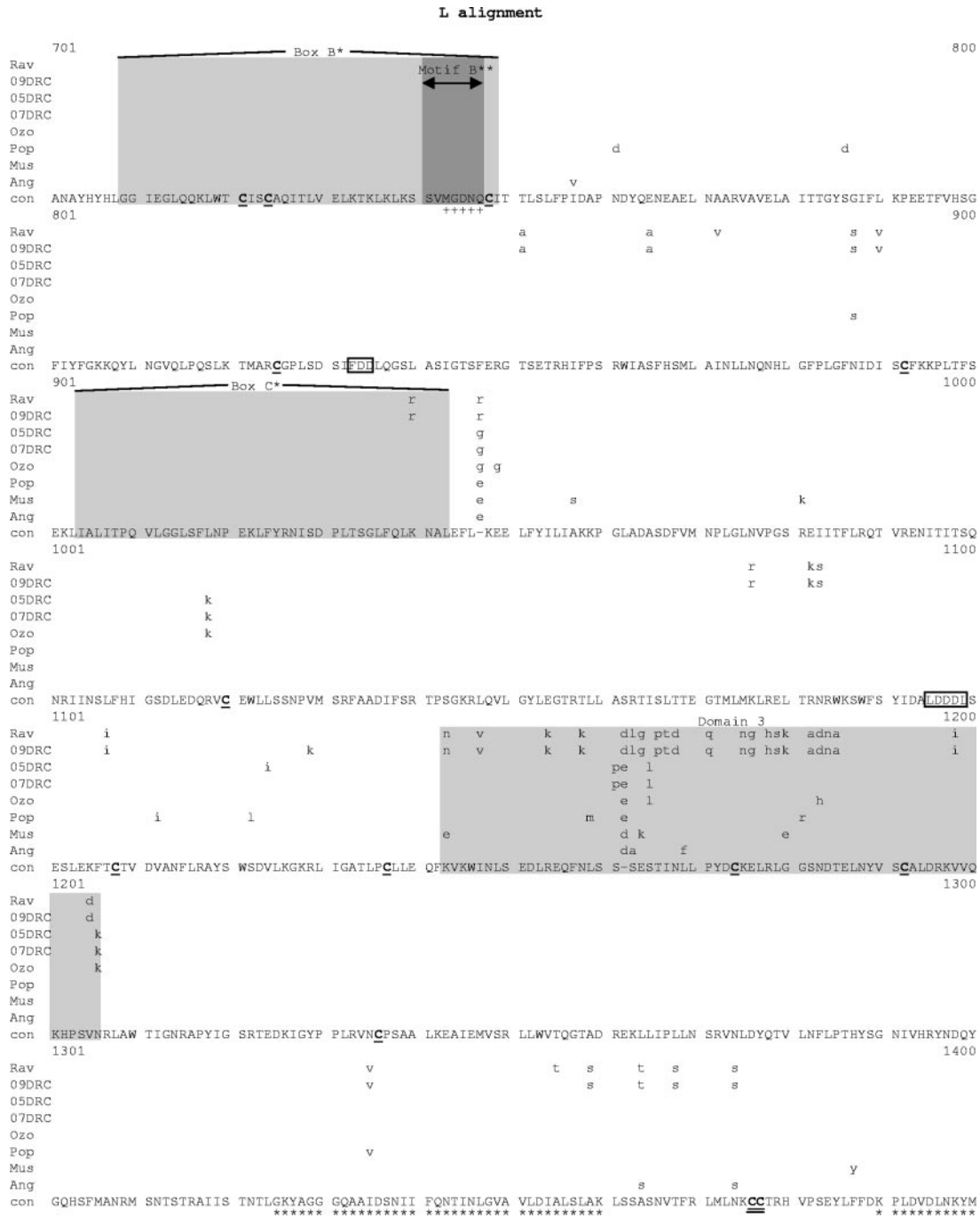
FIG. 7—*Continued.*

teine motif at positions 1376 to 1377, as previously noted (36), which are found in most L proteins at similar locations and are believed to anchor protein secondary structure to maintain the necessary conformation of the putative active sites.

Outside of the polymerase and glycoprotein, a few notable genomic elements reside in the NP, VP35, and VP30 genes. An alignment of NP amino acid sequences (Fig. 7A) shows that nearly all the variability is in the C-terminal half of the protein, similar to that seen in a recent comparison of Zaire and Sudan ebolavirus full-length genome sequences (44). Within this

same half of the protein are seven unique phosphorylation domains (28), each containing one or more serine/threonine kinase substrate motifs. Yet, despite the overall variation, the majority of individual kinase recognition motifs are conserved to the point that, regardless of the marburgvirus strain examined, at least one motif is present within each of the seven domains. The maintenance of these motifs highlights the potential role for phosphorylation in this region of NP, an area postulated to effect protein-protein interactions.

VP35 protein of marburgvirus plays an essential role in tran-

L alignment

```
        1401                                                                                    1500
Rav                                                            i
09DRC                                                          i
05DRC
07DRC
Ozo
Pop
Mus                  t
Ang
con     DNELVYDNDP LCSGIKGRLG RVSRSTLSLS LNVSDIGSYD FPTIAAWTLG ETIVGSIFSD ESSQSTDPIS SGCTKTFVTH FLVYPVESIF YAFGANLIVE
        **********  *T********                                                              C
        1501                                                                                    1600
Rav          n    s                                                                         r  n  p
09DRC        n    s                                                                         r  n  p
05DRC                                                                             h
07DRC                                                                             h
Ozo                                                                              h
Pop
Mus
Ang          n
con     SLSLSRIKSI KNLSDLTFLI SSTIRNLSHR SLRILQSTFR HELVLTRLAH HIPLISLMLG GSAGEKSSSD AVRLFLTASY QNFINNFSCL MKKGQSSLPV
                                                                          * ********** ********** *******T* ***
        1601                  Domain 4                                                   Domain 5    1700
Rav                        c ttk    nhrpv    flt n                                  i   s       y  sp   l a i   s
09DRC                      c ttk a  nhrpv    flt n                                  i   s       y  sp   l a i   s
05DRC                        rc t            s
07DRC                        rc t            s
Ozo                          rc t            s
Pop                          khqi            p
Mus                        i kr i            p g                                              p f
Ang                        f         ypv                                                             i  p
con     WLYFPSEGQQ LKPILKILQR LSDLLSPDKV Q--K-LADTC C-IDSFWVYP SKSTRTNHYY ASLNYWRDKA NKVKNTPFSH LINCSFLELS SHTSSVSSNQ
        1701       ld t nmse  aqst hep    Domain 5 continued                                          1800
Rav     e v l htsr ld t nmse  aqst hep    evctei y    dpaksyl llent  r d    lr  ka  rg pl l v  srqcl a t cp hp
09DRC   e v l htsr ld t nmse  aqst hep    evctei      dpaksyl llent  r      lr  na  rg pl l v  srqcl a t cp hp
05DRC  h      f           t        e         r    p       g       d    g            l                 cm
07DRC  h d    f           t        e         r    p       g       d    g            l                 m
Ozo            d     t            e          r    p               d                 q                 m
Pop                 n     a        d         g                g           g     p t    gf vg d  prp         p  i s p y
Mus         y      nt     a  r            g t            eg                      c  re   p     f       p his   f i
Ang     m          n  t a    em    t      g                   g  c        i           g         l k   p  i g
con     QVTNSKYIVH PEEIPEIN-R TKLINYGSTA LQ-MDIKMPL SEQNLVENCR PSK-IRFKDN QKITKHDQRY EKKESSSQQM SPEDNMQTLA Y-HNSSPSQT
        1801                 Domain 5 continued                                                       1900
Rav     ttep smlrn c  i aa r    t dpr mss i dmrs k   mri srn  l    p ecl    kgk v          s     s          r
09DRC   ttep smlrn c  i aa r    t dpr mss i dmrs k   mri srn  l    p ecl    kgk v          s     s          r
05DRC        i         r       k  f  n  i n        f                 np     p                               i
07DRC        i         r       k  f     i n        f                 p      p                               i
Ozo          i         r       k  f     i n        f                 p                                      i
Pop     lt p              l         ks         i gi  fr         r                         q
Mus     l     a       i     l  ey     k     r  i                                             s
Ang     i              t  p     kf  l    i g       p         k
con     FIKSLDVHED FDASKVILNS EINN-NLTDC TLNT-LLTTP TGTE-LDTSP LQSSRYSSTS RERSLLSREQ ASYLYVDCSN IPSISLDPGF RNMSDQNQVQ
        1901                            Motif C**                                                     2000
Rav        t    yt              a  k  g                                                                k
09DRC      t    yt              a  k  g                                                                k
05DRC      a                                                            h
07DRC      a                                                            h
Ozo        a                                                           h
Pop        t                                                       k                                          n
Mus        t                       k k
Ang        a                                                                                          k
con     MLIN-YKRDL HACFDSNQFC RFTGVVSSMH YKLYDLLPPG ELRKAICLAE GEGSGARLLL KWKETDYLFF NTLATDSQQE AEILSGRVIP RMLYNIDRLS
        2001                                                                                         2100
Rav     v    k                            e          i   a       t          g
09DRC   v    k                 h          e          i   a       t          g
05DRC
07DRC                                          i                                     r
Ozo
Pop              s                                                                   r
Mus
Ang     v                                                                           r
con     ALLESRRLIL NNLTIQITDI TNPLWLDSVI QYLPEDSDIL TMDAETTKDE TREQLYKTIV NIWTRTSPNI PKISIIKVFL LDYEGTLFLM KNAIQYYGQV
```

FIG. 7—*Continued.*

scription and replication of viral RNA. A predicted coiled-coil domain with residues 70 to 120 may effect VP35 oligomerization, an interaction which in turn may be necessary for VP35 to bridge NP with L to form the active polymerase complex (35). At the heart of the presumed coiled-coil domain are heptad repeats containing hydrophobic residues at the first and fourth positions. This domain shows variability at nine positions (Fig. 7B). Yet despite this variation, the spacing of the hydrophobicity is strictly maintained among all the aligned sequences. The Angola sequence shows no variation whatsoever from the consensus sequence throughout the 50-amino-acid domain. In addition to its role in RNA replication, VP35 of ebolavirus has recently been shown to contain an 11-amino-acid motif (Zaire ebolavirus residues 304 to 314) that is thought to be essential for type I interferon (IFN) antagonism (19). This motif, which is possibly involved in RNA binding, is also present in marburgvirus (residues 293 to 303) and has identity at 9 of 11 positions with that of ebolavirus, including the three basic amino acids experimentally demonstrated to be important. Further highlighting the potential importance of this VP35

L alignment

```
          2101                                                                                  2200
Rav                                v    rv                                      i             t  t
09DRC                              v    rv                                      i             t  t
05DRC                 s
07DRC                 s
Ozo                   s
Pop                        k    s                                 k
Mus                             s
Ang                             s
con   QLKKPYSSNA KNSEWYLCCG KRRIQRLQID FPDQVGIFLI CKAMSRQRQA IPYWLKHIEK NYPASLHEFF LTLGFPSLES SFCHRYTIPF SEGKALFHKV
          2201                                                                                  2300
Rav          r                                        vr   r          s          t      i    r        v
09DRC        r                               a        s   vr   r      s          t      i    r        v
05DRC
07DRC
Ozo
Pop
Mus                                                                  i                s
Ang
con   QSYVRQGKQH LHSLMLDYEN NSPLLDLRNH FICSLRGKIT KYYNDILKLN LVIKAVEKGK NWSQLVETLP NMESVCIVEV DHECFGCEKR LLLKLDFIRN
          2301                  2331
Rav                      fsrpk... .
09DRC                    fsrpk... .
05DRC
07DRC
Ozo                      s
Pop                        c  t
Mus                        g
Ang                        g
con   TKIAEQKLLN RVIGYILFFP FGLFKSRSLR A
```

FIG. 7—*Continued.*

domain, the alignment (Fig. 7B) reveals that this domain is a general feature of all known marburgviruses.

VP30 of Zaire ebolavirus has been shown to contain an unconventional Cys_3-His zinc binding domain whose integrity has been shown to be required for VP30 function in virus transcription (33). The alignment in Fig. 7E shows the high conservation of all four zinc-coordinating residues among all eight marburgvirus taxa. Adjacent to the zinc binding domain is a well-conserved tetraleucine motif which could facilitate VP30 oligomerization, similar to that observed with ebolavirus VP30 (18).

Finally, VP24, whose alignment is shown in Fig. 7F, is highly conserved throughout the protein. Recent studies of marburgvirus VP24 have implicated it to be involved in nucleocapsid assembly and interactions between nucleocapsids and budding sites at the plasma membrane (2). In addition, ebolavirus VP24 has been shown to bind Karyopherin α1 and to block STAT1 nuclear accumulation, thus implicating VP24 as a virulence determinant that allows ebolavirus to evade antiviral effects of IFNs (41). Consistent with the idea that VP24 may be a virulence determinant, both mouse- and guinea pig-adapted ebolaviruses, each of which is capable of 100% lethality in its respective animal model, have amino changes that map to VP24 (7, 51).

DISCUSSION

In this report we describe the laboratory confirmation of the first known marburgvirus outbreak in West Africa, along with an extensive comparative genomic analysis of marburgviruses. Virus associated with previous Marburg HF outbreaks had all originated in East Africa, and up to 21% virus genetic diversity had been found based on analysis of partial genome sequences. Consequently, the discovery of Marburg HF in Angola was a surprise, as was the finding that the Angolan marburgvirus was not more distinct, differing from the main group of East Afri-

can marburgviruses by only ~7%. Presumably the absence of a greater genetic difference reflects virus natural host reservoir similarities between the East and West African locations. However, due to the resource-poor nature of the affected area, an index case was never identified, and thus the importation of the virus from East Africa, while unlikely, cannot be ruled out. By the same token, the large virus genetic difference seen between the Ravn/09DRC lineage and the other marburgviruses may represent reservoir host differences, such that viruses within this lineage may be associated with a different host species or subspecies than the other marburgviruses. Interestingly, earlier ecological niche modeling approaches had predicted that parts of northern and eastern Angola were within the range for potential filovirus reservoirs (39, 40).

The natural reservoir for marburgvirus remains unknown, but marburgvirus emergence in Angola will likely extend the scope of the reservoir search beyond East Africa. As recently reported (5), the marburgvirus outbreak in Durba, DRC, was closely associated with illegal mining activity. Epidemiological data linked over 70% of the cases with mines or caves, suggesting that the natural reservoir could well be associated with such environments. With the Angola outbreak, difficulties in surveillance and contact tracing, combined with the delay in the identification of the outbreak, led to poor epidemiological linkage of marburgvirus cases and ultimately to a lack of success in identifying a point source or mounting any ecological study. Filovirus outbreaks in general are relatively rare events. A recent report has suggested that bat species with rather wide geographic distributions may be a potential reservoir for ebolavirus (24). If the natural reservoir of marburgvirus is similar to that of ebolavirus, the emergence of marburgvirus in western Africa should not be surprising, as the sites of multiple large ebolavirus outbreaks are less than 500 miles away, including areas which have experienced almost yearly activity over the last decade (25, 26).

The initial diagnosis was based on 9 of 12 clinical specimens

testing positive by virus isolation, antigen, and/or PCR-based methods, including a newly designed Q-RT-PCR assay designed to detect the VP40 genome region of all known strains of marburgvirus. This Q-RT-PCR assay outperformed all other assays designed to detect acute marburgvirus infection, including the "gold standard" virus isolation assay. The assay was also sufficiently robust to allow deployment into a field setting in Angola. The extensive virus genomic analysis presented here also confirms that the VP40 gene target was an excellent choice for a broadly reactive marburgvirus detection assay, as it is the most conserved virus gene. These features suggest this should be a highly useful assay for detection of potential naturally occurring or deliberate Marburg HF outbreaks in the future.

Following the initial diagnosis, we determined the full-length genome sequence directly from clinical material from 11 patients representing the temporal and geographic distribution of the outbreak. Our purpose was to fully characterize the level of sequence variation generated during the course of a large marburgvirus outbreak with a high case fatality rate and to use this information to determine whether multiple marburgvirus introductions into the human population had occurred. To generate a context with which to compare the Angola marburgvirus sequence data, we determined the complete genome sequence of the Ravn marburgvirus, first isolated in 1987 in Kenya, as well as three additional, uncharacterized marburgviruses from Durba, DRC, the location of the largest outbreak on record prior to the one in Angola. Our data increased the number of the complete genome reference sequences from three to eight. Therefore, we used this expanded database to review the known genetic elements throughout the genome in an effort to assess their importance based upon the degree to which the features are maintained across the eight marburgvirus strains. One of the most striking features is the high degree to which nearly all the genetic elements previously identified were maintained. To some extent this was expected, but we predicted that perhaps the Ravn and 09DRC lineages, differing by greater than 21% from all other marburgviruses at the nucleotide level, would show more versatility as to how the viral proteins could potentially carry out their functions. Diversity within marburgviruses is still, even with this more complete data set, considerably less than that seen with the ebolaviruses, to the point that the Ravn/09DRC lineage would not be considered a different species. In fact, the three most conserved genes, VP40, VP24, and L, among the major species of ebolavirus (Zaire, Sudan, and Reston) show 20 to 25% protein diversity (44). In contrast, those same genes within the marburgvirus lineages shown here have, respectively, 1.6, 4.3, and 12.3% maximum diversity.

Given the general error-prone nature of RNA virus polymerases, the fact that there were comparatively few changes observed among the 11 Angolan marburgvirus isolates was somewhat surprising. In fact, four of the complete genome sequences (19,114 nucleotides in length) from specimens collected over a month and a half and representing at least two to three human to human transmission cycles were 100% identical. In the absence of precise knowledge of transmission events from the natural reservoir, possibly bats (24, 47), this analysis attempts to answer the question of what level of sequence divergence constitutes a distinct lineage versus the degree of

sequence variation to be expected in a large outbreak of single origination. To answer this question, we felt it important to compare the diversity seen in the Angola outbreak with that in the Durba, DRC, outbreak. Among the sequences from the Angola outbreak, the representative from the Songo municipality (specimen 0754) had 11 nucleotide differences from the reference isolate (0.07% variation). However, this maximum level of variation is 1/10 of the minimum diversity seen between the two closest lineages within the Durba outbreak, 05DRC and 07DRC, respectively (0.8% variation). Based on this comparison, we do not consider the Songo lineage to be indicative of a second introduction, although such a possibility cannot be ruled out. For the Angola marburgvirus outbreak, it would be interesting to measure the rate of accumulation of nucleotide changes over time, but unfortunately, the epidemiological links between the 11 patients could not be established and there is no Angola marburgvirus isolate from the presumed beginning of the outbreak in October 2004.

Genome plasticity and rapid evolution in response to positive selection are general features of RNA viruses which possess error-prone polymerases and lack proofreading mechanisms (12). Such features have been well documented in recent outbreaks of emergent RNA viruses, such as severe acute respiratory syndrome coronavirus and human immunodeficiency virus (HIV), in which genetic differences have been shown to accumulate during human-to-human transmission events and, in some instances, within the same tissues of a single host (27, 43). Severe acute respiratory syndrome coronavirus has recently been calculated to accumulate approximately two mutations per genome per human transmission event (50). Why, then, was so little variation found within this marburgvirus outbreak? The answer likely involves at least two related components, namely, (i) the time of progression within the host from initial infection to disease outcome and (ii) the host immune response to marburgvirus infection. With HIV, highly diverse quasispecies have been shown to develop from homogeneous populations in response to the development of host neutralizing antibodies and cytotoxic T-lymphocyte responses (6). Yet, in the later stages of HIV infection towards the development of AIDS, the immune system is no longer able to exert strong selective pressure on the replicating quasispecies and, as a result, HIV evolution dramatically slows (11). For marburgvirus, like Zaire ebolavirus, disease progression is so rapid that most individuals die before an effective immune response can be mounted (1, 22, 48). For ebolavirus the median time to death after onset of symptoms is only 8 to 9 days (22, 48). In addition to the rapid course of infection, ebolavirus actively suppresses the immune system by (i) directly antagonizing the host interferon system through an uncharacterized mechanism involving the virus VP35 protein (4) and (ii) preferentially infecting dendritic cells and macrophages as sites of primary infection, thus diminishing host cell populations that are critical for the establishment of innate and adaptive immune responses (16, 17). These features together allow for explosive virus growth within the host in a manner virtually unchecked by the immune system. The lack of IgM and IgG responses shown in Table 1 is consistent with this view. The lack of genetic diversity within the Angola marburgvirus outbreak is therefore not surprising and is in fact consistent with the absence of genetic differences seen among genome fragments

of viruses analyzed in previous Zaire and Sudan ebolavirus outbreaks (42, 48).

A trial sampling procedure that emerged during the outbreak was the use of oral swabs for sampling suspected Marburg VHF corpses. RNA extracted from oral swabs sometimes gave low C_T values in the Q-RT-PCR assay (Table 2 and data not shown), indicating the potential for high viral loads in these secretions. Unfortunately, definitive interpretations of C_T values from oral swabs are difficult, since the swabbing technique and the volumes recovered are inherently variable. In general, very high viral loads are seen in sera of individuals infected with filoviruses. For instance, during human infections with Sudan ebolavirus, viral loads in patient serum can vary from 10^5 RNA copies/ml at the time of fever onset to greater than 10^{10} RNA copies/ml at the time of death (48). However, the timing and extent to which this range of viral load is reflected in human oral/nasal secretions are currently unknown. Of particular concern is the use of this sampling technique as a means of establishing a particular etiology of a VHF outbreak when multiple agents must be considered and tested, many of which may not ever produce viral loads that can be detected by oral swabs. Blood-based sampling as a reliable source of virus for serology-, antigen-, and PCR-based detection assays is well documented and should remain the method of choice for sampling suspected VHF patients.

ACKNOWLEDGMENTS

REFERENCES

1. Baize, S., E. M. Leroy, M.-C. Georges-Courbot, M. Capron, J. Lansoud-Soukate, P. Debre, S. P. Fisher-Hoch, J. B. McCormick, and A. J. Georges. 1999. Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in ebola virus-infected patients. Nat. Med. **5:**423–426.
2. Bamberg, S., L. Kolesnikova, P. Moller, H. D. Klenk, and S. Becker. 2005. VP24 of Marburg virus influences formation of infectious particles. J. Virol. **79:**13421–13433.
3. Barik, S., E. W. Rud, D. Luk, A. K. Banerjee, and C. Y. Kang. 1990. Nucleotide sequence analysis of the L gene of vesicular stomatitis virus (New Jersey serotype): identification of conserved domains in L proteins of nonsegmented negative-strand RNA viruses. Virology **175:**332–337.
4. Basler, C. F., A. Mikulasova, L. Martinez-Sobrido, J. Paragas, E. Muhlberger, M. Bray, H.-D. Klenk, P. Palese, and A. Garcia-Sastre. 2003. The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. J. Virol. **77:**7945–7956.
5. Bausch, D. G., M. Borchert, T. Grein, C. Roth, R. Swanepoel, M. L. Libande, A. Talarmin, E. Bertherat, J.-J. Muyembe-Tamfum, B. Tugume, R. Colebunders, K. M. Kondé, P. Pirard, L. L. Olinda, G. R. Rodier, P. Campbell, O. Tomori, T. G. Ksiazek, and P. E. Rollin. 2003. Risk factors for Marburg hemorrhagic fever, Democratic Republic of the Congo. Emerg. Infect. Dis. **9:**1531–1537.
6. Borrow, P., H. Lewicki, X. Wei, M. S. Horwitz, N. Peffer, H. Meyers, J. A. Nelson, J. E. Gairin, B. H. Hahn, M. B. Oldstone, and G. M. Shaw. 1997. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat. Med. **3:**205–211.
7. Bray, M., K. Davis, T. Geisbert, C. Schmaljohn, and J. Huggins. 1998. A mouse model for evaluation of prophylaxis and therapy of Ebola hemorrhagic fever. J. Infect. Dis. **178:**651–661.
8. Bukreyev, A. A., V. E. Volchkov, V. M. Blinov, S. A. Dryga, and S. V. Netesov. 1995. The complete nucleotide sequence of the Popp (1967) strain of Marburg virus—a comparison with the Musoke (1980) strain. Arch. Virol. **140:**1589–1600.
9. Centers for Disease Control and Prevention. 2005. Outbreak of Marburg virus hemorrhagic fever—Angola, October 1, 2004-March 29, 2005. Morb. Mortal. Wkly. Rep. **54:**308–309.
10. Centers for Disease Control and Prevention and World Health Organization. 1998. Infection control for viral haemorrhagic fevers in the African health care setting, p. 1–198. Centers for Disease Control and Prevention. Atlanta, Ga.
11. Delwart, E. L., H. Pan, H. W. Sheppard, D. Wolpert, A. U. Neumann, B. Korber, and J. I. Mullins. 1997. Slower evolution of human immunodeficiency virus type 1 quasispecies during progression to AIDS. J. Virol. **71:**7498–7508.
12. Domingo, E., and J. J. Holland. 1997. RNA virus mutations and fitness for survival. Annu. Rev. Microbiol. **51:**151–178.
13. Duprex, W. P., F. M. Collins, and B. K. Rima. 2002. Modulating the function of the measles virus RNA-dependent RNA polymerase by insertion of green fluorescent protein into the open reading frame. J. Virol. **76:**7322–7328.
14. Feldmann, H., and M. P. Kiley. 1999. Classification, structure, and replication of filoviruses. Curr. Top. Microbiol. Immunol. **235:**1–21.
15. Feldmann, H., E. Mühlberger, A. Randolf, C. Will, M. P. Kiley, A. Sanchez, and H. D. Klenk. 1992. Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. Virus Res. **24:**1–19.
16. Geisbert, T. W., L. E. Hensley, T. R. Gibb, K. E. Steele, N. K. Jaax, and P. B. Jahrling. 2000. Apoptosis induced in vitro and in vivo during infection by Ebola and Marburg viruses. Lab. Investig. **80:**171–186.
17. Geisbert, T. W., L. E. Hensley, T. Larsen, H. A. Young, D. S. Reed, J. B. Geisbert, D. P. Scott, E. Kagan, P. B. Jahrling, and K. J. Davis. 2003. Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. Am. J. Pathol. **163:**2347–2370.
18. Hartlieb, B., J. Modrof, E. Muhlberger, H. D. Klenk, and S. Becker. 2003. Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide. J. Biol. Chem. **278:**41830–41836.
19. Hartman, A. L., J. S. Towner, and S. T. Nichol. 2004. A C-terminal basic amino acid motif of Zaire ebolavirus VP35 is essential for type I interferon antagonism and displays high identity with the RNA-binding domain of another interferon antagonist, the NS1 protein of influenza A virus. Virology **328:**177–184.
20. Johnson, E. D., B. K. Johnson, D. Silverstein, P. Tukei, T. W. Geisbert, A. N. Sanchez, and P. B. Jahrling. 1996. Characterization of a new Marburg virus isolated from a 1987 fatal case in Kenya. Arch. Virol. **11:**101–114.
21. Kamer, G., and P. Argos. 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucleic Acids Res. **12:**7269–7282.
22. Ksiazek, T. G., P. E. Rollin, A. J. Williams, D. S. Bressler, M. L. Martin, R. Swanepoel, F. J. Burt, P. A. Leman, A. K. Rowe, R. Mukunu, A. Sanchez, and C. J. Peters. 1999. Clinical virology of Ebola hemorrhagic fever (EHF): virus, virus antigen, IgG and IgM antibody findings among EHF patients in Kikwit, 1995. J. Infect. Dis. **179**(Suppl. 1):S177–S187.
23. Ksiazek, T. G., C. P. West, P. E. Rollin, P. B. Jahrling, and C. J. Peters. 1999. ELISA for the detection of antibodies to Ebola viruses. J. Infect. Dis. **179**(Suppl. 1):S191–S198.
24. Leroy, E. M., B. Kumulungui, X. Pourrut, P. Rouquet, A. Hassanin, P. Yaba, A. Delicat, J. T. Paweska, J. P. Gonzalez, and R. Swanepoel. 2005. Fruit bats as reservoirs of Ebola virus. Nature **438:**575–576.
25. Leroy, E. M., S. Baize, E. Mavoungou, and C. Apetrei. 2002. Sequence analysis of the GP, NP, VP40 and VP24 genes of Ebola virus isolated from deceased, surviving and asymptomatically infected individuals during the 1996 outbreak in Gabon: comparative studies and phylogenetic characterization. J. Gen. Virol. **83:**67–73.
26. Leroy, E. M., P. Rouquet, P. Formenty, S. Souquiere, A. Kilbourne, J.-M. Froment, M. Bermejo, S. Smit, W. Karesh, R. Swanepoel, S. R. Zaki, and P. E. Rollin. 2004. Multiple Ebola virus transmission events and rapid decline of Central African wildlife. Science **303:**387–390.
27. Liu, J., S. L. Lim, Y. Ruan, A. E. Ling, L. F. Ng, C. Drosten, E. T. Liu, L. W. Stanton, and M. L. Hibberd. 2005. SARS transmission pattern in Singapore reassessed by viral sequence variation analysis. PLoS Med. **2:**e43.

28. **Loetfering, B., E. Muehlberger, T. Tamura, H.-D. Klenk, and S. Becker.** 1999. The nucleoprotein of Marburg virus is target for multiple cellular kinases. Virology **255:**50–62.

29. **Martini, G. A.** 1971. Clinical syndrome, p. 1–9. *In* G. A. Martini and R. Siegert (ed.), Marburg virus disease. Springer-Verlag, New York, N.Y.

30. **Martini, G. A., and R. Siegert.** 1971. Preface, p. v. *In* G. A. Martini and R. Siegert (ed.), Marburg virus disease. Springer-Verlag, New York, N.Y.

31. **McIlhatton, M. A., M. D. Curran, and B. K. Rima.** 1997. Nucleotide sequence analysis of the large (L) genes of phocine distemper virus and canine distemper virus (corrected sequence). J. Gen. Virol. **78:**571–576.

32. **Ministry of Health, Republic of Angola.** 2005. Declaration of termination of Marburg viral hemorrhagic fever in Angola. Ministry of Health, Luanda, Angola.

33. **Modrof, J., S. Becker, and E. Muhlberger.** 2003. Ebola virus transcription activator VP30 is a zinc-binding protein. J. Virol. **77:**3334–3338.

34. **Modrof, J., C. Moritz, L. Kolesnikova, T. Konakova, B. Hartlieb, A. Randolf, E. Muhlberger, and S. Becker.** 2001. Phosphorylation of Marburg virus vp30 at serines 40 and 42 is critical for its interaction with NP inclusions. Virology **287:**171–182.

35. **Moller, P., N. Pariente, H. D. Klenk, and S. Becker.** 2005. Homo-oligomerization of marburgvirus VP35 is essential for its function in replication and transcription. J. Virol. **79:**14876–14886.

36. **Muhlberger, E., A. Sanchez, A. Randolf, C. Will, M. P. Kiley, H. D. Klenk, and H. Feldmann.** 1992. The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses. Virology **187:**534–547.

37. **Nichol, S. T., and R. Swanepoel.** Personal communication.

38. **Peters, C. J., and A. S. Khan.** 1999. Filovirus diseases. Curr. Top. Microbiol. Immunol. **235:**85–95.

39. **Peterson, A. T., J. T. Bauer, and J. N. Mills.** 2004. Ecologic and geographic distribution of filovirus disease. Emerg. Infect. Dis. **10:**40–47.

40. **Peterson, A. T., D. S. Carroll, J. N. Mills, and K. M. Johnson.** 2004. Potential mammalian filovirus reservoirs. Emerg. Infect. Dis. **10:**2073–2081.

41. **Reid, S. P., L. W. Leung, A. L. Hartman, O. Martinez, M. L. Shaw, C. Carbonnelle, V. E. Volchkov, S. T. Nichol, and C. F. Basler.** 2006. Ebola virus VP24 binds karyopherin α1 and blocks STAT1 nuclear accumulation. J. Virol. **80:**5156–5167.

42. **Rodriguez, L., A. De Roo, Y. Guimard, S. Trappier, A. Sanchez, D. Bressler, A. J. Williams, T. G. Ksiazek, C. J. Peters, and S. T. Nichol.** 1999. Persistence and genetic stability of Ebola virus during the outbreak in Kikwit, Zaire 1995. J. Infect. Dis. **179**(Suppl. 1)**:**S170–S176.

43. **Ruan, Y. J., C. L. Wei, A. L. Ee, V. B. Vega, H. Thoreau, S. T. Su, J. M. Chia, P. Ng, K. P. Chiu, L. Lim, T. Zhang, C. K. Peng, E. O. Lin, N. M. Lee, S. L. Yee, L. F. Ng, R. E. Chee, L. W. Stanton, P. M. Long, and E. T. Liu.** 2003. Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection. Lancet **361:**1779–1785.

44. **Sanchez, A., and P. E. Rollin.** 2005. Complete genome sequence of an Ebola virus (Sudan species) responsible for a 2000 outbreak of human disease in Uganda. Virus Res. **113:**16–25.

45. **Sanchez, A., S. G. Trappier, U. Ströher, S. T. Nichol, M. D. Bowen, and H. Feldmann.** 1998. Variation in the glycoprotein and VP35 genes of Marburg virus strains. Virology **240:**138–146.

46. **Sanger, C., E. Muhlberger, B. Lotfering, H.-D. Klenk, and S. Becker.** 2002. The Marburg virus surface protein GP is phosphorylated at its ectodomain. Virology **295:**20–29.

47. **Swanepoel, R., P. A. Leman, F. J. Burt, N. A. Zachariades, L. E. O. Braack, T. G. Ksiazek, P. E. Rollin, S. R. Zaki, and C. J. Peters.** 1996. Experimental inoculation of plants and animals with Ebola virus. Emerg. Infect. Dis. **2:**321–325.

48. **Towner, J. S., P. E. Rollin, D. G. Bausch, A. Sanchez, S. M. Crary, M. Vincent, W. F. Lee, C. F. Spiropoulou, T. G. Ksiazek, M. Lukwiya, F. Kaducu, R. Downing, and S. T. Nichol.** 2004. Rapid diagnosis of Ebola hemorrhagic fever by reverse transcription-PCR in an outbreak setting and assessment of patient viral load as a predictor of outcome. J. Virol. **78:**4330–4341.

49. **Towner, J. S., T. K. Sealy, and S. T. Nichol.** 2006. Unpublished data.

50. **Vega, V. B., Y. Ruan, J. Liu, W. H. Lee, C. L. Wei, S. Y. Se-Thoe, K. F. Tang, T. Zhang, P. R. Kolatkar, E. E. Ooi, A. E. Ling, L. W. Stanton, P. M. Long, and E. T. Liu.** 2004. Mutational dynamics of the SARS coronavirus in cell culture and human populations isolated in 2003. BMC Infect. Dis. **4:**32.

51. **Volchkov, V. E., A. A. Chepurnov, V. A. Volchkova, V. A. Ternovoj, and H. D. Klenk.** 2000. Molecular characterization of guinea pig-adapted variants of Ebola virus. Virology **277:**147–155.

52. **Volchkov, V. E., V. M. Blinov, and S. V. Netesov.** 1992. The envelope glycoprotein of Ebola virus contains an immunosuppressive-like domain similar to oncogenic retroviruses. FEBS Lett. **305:**181–184.

53. **Volchkov, V. E., V. A. Volchkova, A. Chepurnov, V. M. Blinov, O. Dolnik, S. V. Netesov, and H. Feldmann.** 1999. Characterization of the L gene and 5′ trailer regions of Ebola virus. J. Gen. Virol. **80:**355–362.

54. **Volchkov, V. E., V. A. Volchkova, U. Stroeher, S. Becker, O. Dolnik, M. Cieplik, W. Garten, H.-D. Klenk, and H. Feldmann.** 2000. Proteolytic processing of Marburg virus glycoprotein. Virology **268:**1–6.

55. **Will, C., E. Muhlberger, D. Linder, W. Slenczka, H. D. Klenk, and H. Feldmann.** 1993. Marburg virus gene 4 encodes the virion membrane protein, a type I transmembrane glycoprotein. J. Virol. **67:**1203–1210.

56. **World Health Organization.** 1967. Outbreaks in laboratory personnel working with Cercopithecus monkeys from East Africa—Europe. Wkly. Epidemiol. Rec. **42:**479–480.

57. **World Health Organization.** 1999. Marburg fever, Democratic Republic of the Congo. Wkly. Epidemiol. Rec. **74:**145.

58. **World Health Organization.** 2005. Marburg virus disease, Angola. Wkly. Epidemiol. Rec. **80:**115–117.

59. **Yaddanapudi, K., G. Palacios, J. S. Towner, S. T. Nichol, C. Sariol, and W. I. Lipkin.** Implication of a retrovirus-like glycoprotein peptide in the immunopathogenesis of Ebola and Marburg viruses. Submitted for publication.

60. **Zaki, S. R., W.-J. Shieh, P. W. Greer, C. S. Goldsmith, T. Ferebee, J. Katshitshi, K. Tchioko, M. A. Bwaka, R. Swanepoel, P. Calain, A. S. Khan, E. Lloyd, P. E. Rollin, T. G. Ksiazek, C. J. Peters, and and the EHF Study Group.** 1999. A novel immunohistochemical assay for detection of Ebola virus in skin: implications for diagnosis, spread, and surveillance of Ebola hemorrhagic fever. J. Infect. Dis. **179**(Suppl. 1)**:**S36–S47.