JOURNAL OF VIROLOGY, Apr. 2011, p. 3905–3917 Vol. 85, No. 8
0022-538X/11/$12.00 doi:10.1128/JVI.02372-10


# Key Genomic Changes Necessary for an *In Vivo* Lethal Mouse Marburgvirus Variant Selection Process▽

Loreen L. Lofts,* Jay B. Wells, Sina Bavari, and Kelly L. Warfield†

*U.S. Army Medical Research Institute of Infectious Diseases, Frederick, Maryland 21702*

Received 12 November 2010/Accepted 18 January 2011

**Marburgvirus (MARV) infections are generally lethal in humans and nonhuman primates but require *in vivo* lethal mouse variant selection by the serial transfer (passage) of the nonlethal virus into naïve mice to propagate a lethal infection. The passage of progenitor (wild-type) MARV or Ravn virus (RAVV) from infected *scid* BALB/c mouse liver homogenates into immunocompetent BALB/c mice results in the selection of lethal mouse viruses from within the quasispecies sufficient to establish lethality in immunocompetent mice. Genomic analysis in conjunction with the passage history of each mutation detailed the altered primary and secondary structures of the viral genomic RNA throughout the process. Key findings included the following: (i) a VP40:D184N mutation previously identified in the lethal guinea pig MARV genome was the first mutation to occur during the passage of both the MARV and RAVV variants; (ii) there was biased hypermutagenesis in the RAVV variant genome; (iii) there were two identical mutations in lethal mouse MARV and RAVV variants, VP40:Y19H in the PPPY motif and VP40:D184N in a loop structure between the two VP40 domains; (iv) the passage of wild-type MARV and RAVV in mice resulted in the selection of viral variants from among the quasispecies with different genotypes than those of the wild-type viruses; and (v) a lethal mouse RAVV variant had different tissue tropisms distinct from those of its wild-type virus. These studies provide insights into how marburgviruses manipulate the host for enzymes, metabolites, translation regulators, and effectors of the innate immune response to serve as potential viral countermeasures.**

The negative-sense, single-stranded RNA family *Filoviridae* is comprised of two established divergent genera, *Marburgvirus* and *Ebolavirus*. The genus *Marburgvirus* consists of one species, *Marburg marburgvirus*, which has two members, marburgvirus (MARV) and Ravn virus (RAVV). The genus *Ebolavirus* consists of five species (all of which have one member): *Bundibugyo ebolavirus* (Bundibugyo virus [BDBV]), *Reston ebolavirus* (Reston virus [RESTV]), *Sudan ebolavirus* (Sudan virus [SUDV]), *Taï Forest ebolavirus* (Taï Forest virus [TAFV]), and *Zaire ebolavirus* (ebolavirus [EBOV]) (25). Marburgviruses are the etiological agents of the lethal Marburg hemorrhagic fever (MHF) in humans and nonhuman primates (reviewed in references 23 and 24). Nonhuman primates are extensively used as animal models for MHF based on pathological manifestations consistent with those of human disease (16, 48) and are considered the model of choice for vaccine and therapeutic safety and efficacy studies. However, for studies that require preliminary analyses based on large numbers of samples, the number of primates required is often ethically problematic and prohibitively expensive, especially for the number of animals required to analyze a batched throughput of therapeutic candidates. In addition, mechanistic pathogenesis studies are not feasible due to the lack of genetically modified or transgenic monkeys. Therefore, comprehensive studies were undertaken to define the genomics, pathogenesis, and virulence of the MARV and RAVV mouse models that can be useful to address the above-mentioned gaps.

Human or nonhuman primate marburgvirus variants (wild type [wt]) are not lethal to adult immunocompetent mice but can be selected by the *in vivo* passage of the viruses into naïve mice (46, 47). The identification of lethal mouse (lm) marburgvirus variants was elusive until the development of a novel approach that incorporated an initial propagation of the marburgvirus in *scid* BALB/c mice (here referred to as *scid* mice) prior to the infection and passage of the viruses in immunocompetent BALB/c mice.

Filoviruses are considered "emerging or reemerging pathogens" (reviewed in reference 24) or the etiological agents of MHF as a result of a change or evolution in a previously identified organism or as a known MHF pathogen whose incidence rapidly increases in a given geographic area. The description of key consensus genomic changes within the lm marburgvirus variant quasispecies identified during the selection process of lm marburgvirus variants provides the basis to describe the potential ability of marburgvirus to naturally select variants that will propagate in new hosts for emergence or reemergence. Genomic analyses of lm marburgvirus variants are consistent with previous reports of mechanisms for other lethal host RNA viral variants by *in vivo* selection that augment pathogenesis or virulence either in new hosts or in tissue tropism(s) within a susceptible host (reviewed in reference 35). These mechanisms include unidirectionally biased hypermutagenesis of pyrimidines as a result of Ade-to-Ino editing in multiple genes, a distinct order of mutational occurrence, mutations in multiple genes, extensive hypermutation in untranslated regions (UTRs), and switching to host codon usage. Each of these mechanisms will be described separately; however,

* Corresponding author. Mailing address: USAMRIID, Virology Division, 1425 Porter Street, Frederick, MD 21702. Phone: (301) 619-4926. Fax: (301) 619-2290. E-mail: loreen.lofts@us.army.mil.

† Present address: Integrated BioTherapeutics, Inc., Germantown, MD 20876.

▽ Published ahead of print on 2 February 2011.

one must be cognizant that all of these occurred as an integrated-biosystems approach to establish marburgvirus lethality in mice.

In summary, this report presents the genomic and phenotypic analysis of lm marburgvirus variants generated during the process of viral selection of lethal mouse variants. These analyses provide an identification of key changes in marburgvirus variant genomes to promote the fitness of the marburgvirus quasispecies in a new host. Challenge studies to characterize the lm marburgvirus genomic variant changes were done with mice with different major histocompatibility complex (MHC) *H-2* haplotypes that were infected with lm MARV and RAVV variants to investigate phenotypic changes. In addition, infected mouse tissue virus concentration studies were conducted to determine the tissue tropism(s) of the lm marburgvirus variants to further distinguish MARV and RAVV.

## MATERIALS AND METHODS

**Viruses and host cells.** The lethal nonhuman primate MARV variant Ci67 (wt MARV variant) [USAMRIID designation E6(2)PP#1(3X)E6(1), which denotes two amplification of a viral stock in Vero E6 cells (ATCC DRL 1586), followed by the purification of one viral plaque three times with a subsequent amplification in Vero E6 cells] was amplified from MARV-Germany E6(1), which originated as a human isolate [USAMRIID designation MARV-Germany E6(1)]. The procurement of Ci67 at the USAMRIID was related to us by Alan Schmaljohn. This virus originated from a 1967 outbreak in Germany and was transferred to Peter Jahrling at the USAMRIID from Stephan Becker. The nucleotide sequence of Ci67 is 99.5% homologous to that of Poppinga, which was isolated during the same outbreak (reviewed in reference 24). The lethal nonhuman primate RAVV (wt RAVV variant) was described previously (46). Viruses were enumerated and purified with a standard plaque assay (30) and propagated in Vero E6 cells. The virus concentrations in viral stocks were enumerated and plaque purified three times before use in animal experiments. All procedures with infectious marburgviruses were done in a biosafety level 4 laboratory at the USAMRIID.

**Animals.** The procurement of *scid* and immunocompetent BALB/c mice and the housing conditions for all mouse strains were described previously (46, 47). All C57BL/6 mice used in these experiments were immunocompetent and were obtained from the Frederick Cancer Research and Development Center, National Cancer Institute (Frederick, MD). All mice used in these studies were 6 to 10 weeks of age and were of either gender. Research was conducted in compliance with the Animal Welfare Act and other federal statutes and regulations relating to animals and experiments involving animals and adhered to principles stated in the *Guide for the Care and Use of Laboratory Animals* (33a). The facility in which this research was conducted (USAMRIID) was fully accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International. Mouse protocols were approved prior to study initiation by the USAMRIID Institutional Animal Care and Use Committee.

***In vivo* lm marburgvirus selection.** Procedures for the *in vivo* selection of lm marburgvirus variants by the passage of wt marburgviruses in mice were previously described (46, 47). Briefly, each passage consisted of the infection of a group of 10 mice inoculated intraperitoneally (i.p.) with 1,000 PFU of wt MARV or RAVV variants with initial passages in *scid* mice. On day 7 postinfection (p.i.), two of the mice were euthanized, the livers were homogenized in 10 ml of phosphate-buffered saline (PBS) (pH 7.4) and pooled, and 200-μl aliquots were injected i.p. into naïve mice. The mean time to death (MTD) in days was monitored for the remaining eight mice. After 10 passages of RAVV variants or 15 passages of MARV variants in *scid* mice, viruses in liver homogenates from the last *scid* mouse passage were used to infect groups of 10 immunocompetent BALB/c mice; 2 of the mice from this group were euthanized on days 3 to 4 p.i., and the livers were processed for use as viral inocula for subsequent passages. The remaining eight mice in the group were monitored for patency and death until study termination on day 62 p.i.

**Viral challenge studies with lm marburgvirus variants.** Challenge studies were done to characterize the marburgvirus mouse model for virulence. To compare the virulences of the lm marburgvirus variants, two different mouse strains, BALB/c and C57BL/6 mice, with 5 mice in each group, were injected i.p. with 1,000 PFU of either MARV variant plaque pick 7 (PP7) or the RAVV PP4 variants. Infected mice were observed for 14 days p.i. at least twice daily for morbidity and mortality. Mice considered moribund (based on set criteria, including weight loss, reduced grooming, anorexia, decreased activity, and decreased responsiveness) were euthanized. The i.p. route was selected because only those mice infected by this route succumbed to infection (47) during a time frame consistent with human MHF. (This route of infection was also previously shown to be the only effective route for a lethal immunocompetent mouse EBOV variant [8].) Therefore, the i.p. route was used for all challenge studies presented here.

In addition, infected mouse tissue virus concentration studies were conducted to determine the tissue tropism(s) of lm marburgvirus variants. Groups of 20 mice were infected with wt or lm marburgvirus variants to determine the viral concentrations in immunocompetent BALB/c mouse organs. Five mice were randomly selected from the group at the time point of 1, 3, 5, or 7 days p.i., and the viral concentrations in the brain, gonads, heart, intestines, kidney, liver, lungs, lymph nodes, and spleens were determined by a plaque assay (30).

**Statistical analysis.** Statistical analyses were done with SAS, version 9.2, for Kaplan-Meier survival analysis for the construction of survival curves and to calculate the mean and median survival times. Log-rank tests were used to the compare survival curves. Values of virus concentrations in tissues were transformed to logarithm base 10 values prior to analysis. Where group variances were homogenous, Student's *t* tests were used to compare tissue concentrations between groups. Where group variances were not homogenous, *t* tests with the Satterthwaite approximation were used to compare tissue concentrations between groups. The *P* values were not adjusted further for multiple comparisons due to relatively low numbers of samples; however, a significance level of 0.01 was chosen for all comparisons of tissue viral concentrations.

**Phylogenetic analysis.** A cladogram of 51 filoviruses with complete genomic nucleic acid sequences deposited in GenBank was generated with the Lasergene MegAlign Clustal W v:8.0.2 method (39) to calculate branch distances corresponding to nucleotide sequence divergence. The database included marburgvirus genomes from human and bat primary viral variants from various geographical locations as well as viral genomes isolated from laboratory-infected rodents.

**Viral RNA extraction, cDNA synthesis, and nucleotide determination.** Viral RNAs extracted from TRIzol LS (Invitrogen Corp.) reagent-treated wt marburgvirus variants suspended in culture medium and the lm marburgvirus variants in liver homogenates were annealed with random hexamer oligonucleotides primers to synthesize cDNA using Superscript II (Invitrogen Corp.). Negative-sense genome sequences were determined by the primer-walking method with PCR templates amplified from cDNAs (amplimers) representative of the entire genome annealed to viral gene-specific oligonucleotide primers. Primers were designed by Lasergene Primer Select v:7.2.1(1), 410 (DNASTAR, Inc.), based on the genomic nucleotide sequence of lethal guinea pig MARV Musoke (GenBank accession number EF446131) or the wt RAVV variant (GenBank accession number EF446131). Nucleotide sequences were determined with Applied Biosystems technology (ABI Prism BigDye Terminator cycle sequencing ready reaction kit v1.0 or 3.1 according to the manufacturer's instructions for amplimers) and analyzed with an ABI Prism 3730 DNA analyzer. The 5′- and 3′-noncoding (nc) terminal nucleotide sequences were determined with an amplimer synthesized by annealing virus terminus-specific primers to within a few bases of the respective 5′ or 3′ ends of each terminus. The remainder of the terminus was extrapolated from the genomic nucleotide sequence used for the synthesis of terminus-specific primers. Contigs of the nucleotide sequence data were assembled with the Lasergene SeqMan II v:7.2.1(1), 410, software program for Microsoft Windows (DNASTAR, Inc.), and the amino acids were deduced by the translation of the open reading frames (ORFs) with Lasergene EditSeq v7.2.1(1), 410.

**Nucleotide sequence accession numbers.** The viral genomes of the wt and lm marburgvirus variants described in this report were deposited in GenBank under accession numbers GQ433353 for the wt MARV variant, GQ433352 for the lethal *scid* mouse MARV variant, and GQ433351 for the lethal immunocompetent mouse MARV variant.

## RESULTS

***In vivo* virus selection process for lm marburgvirus variants.** Lethality was deemed the primary endpoint criterion necessary to establish robust marburgvirus mouse models. The endpoint was defined as an MTD of less than 10 days for infected mice to succumb to MHF. As previously reported (46), wt RAVV variant-infected *scid* mice succumbed to MHF at passage 4 in

<10 days, whereas wt MARV variant-infected *scid* mice did not succumb until passage 10. *scid* mice are homozygous for an autosomal recessive mutation (*scid*) (4) located on chromosome 16 (5), which was discovered in C.B-17/lcr mice (4, 11). The *scid* mutation confers an early development arrest of lymphoid cells; however, myeloid cell differentiation and function are normal (reviewed in reference 7). Although the majority of *scid* mice lack functional lymphocytes, 2 to 23% have been reported to develop a limited number of B and T cells between 3 and 9 months of age (6). In addition, EBOV-infected *scid* mice that lack B and T lymphoid cells fail to produce antibodies or cell-mediated immune responses (9). Therefore, wt marburgvirus variant passages were initiated in *scid* mice to allow the selection of viruses within the quasispecies to generate a sufficient viral concentration to establish lethality in immunocompetent mice without interference from a host immune response. Transfers of wt MARV variants in *scid* mice were carried out for 15 passages and transfers of wt RAVV variants were carried out for 10 passages prior to the transfer of viruses in the liver homogenates into immunocompetent BALB/c mice (Fig. 1). No immunocompetent BALB/c mice succumbed to MHF during wt MARV variant passages 16 to 28 or wt RAVV variant passages 11 to 23. One immunocompetent mouse succumbed to lm MARV variant infection at passage 29, whereas lm RAVV variant-infected mice succumbed to infection during passages 24 to 26.

**Biased hypermutagenesis of pyrimides.** The genomic nucleotide sequences of the lm MARV or RAVV variants were compared by alignment with the respective wt marburgvirus variants. The alignments revealed 33 and 75 differences between the wt MARV and lm MARV variant and the wt RAVV and lm RAVV variant genomes, respectively (Table 1). Of these differences, a preponderant 91.0% (57 of 62) were Thy-to-Cyt nonrevertant transition mutations between the wt and lm RAVV variant genomes; in contrast, wt and lm MARV variants had 37.0% (10/27) nonrevertant Thy-to-Cyt transitions. The hypermutagenesis of Thy-to-Cyt transitions was similarly demonstrated for other RNA viruses, including subacute sclerosing panencephalitis (SSPE) virus, which had 80% nucleotide differences between the Biken strain SSPE virus and the putative progenitor ascribed to unidirectionally biased hypermutagenesis (49).

**Mutations in multiple genes.** Further analysis of the alignments indicated genomic sequence changes in the UTRs of the *NP*, *VP35*, *VP40*, *GP*, and *VP30* genes and in all of the translational coding regions or ORFs except for *VP24* for both lm MARV and RAVV variant genomes. The changes were identified as being both nonsynonymous and synonymous mutations after the translation of the ORFs. The location of the mutations within a translational region was used to determine mutation function based on the structure-function of previously described domains determined with the BLAST-N algorithm.

There were five nonrevertant, nonsynonymous mutations in the lm RAVV variant *VP35* gene. All of these mutations were within the first 62 amino acids (aa) of VP35. Multiple mutations in the lm RAVV *VP35* variant coding and noncoding regions and the single Thy insertion at nucleotide (nt) 4353 in the lm MARV variant *VP35* are the first lethal rodent MARV variant *VP35* mutations to be reported for *in vivo* lethal rodent marburgvirus selection studies, although it was not the first

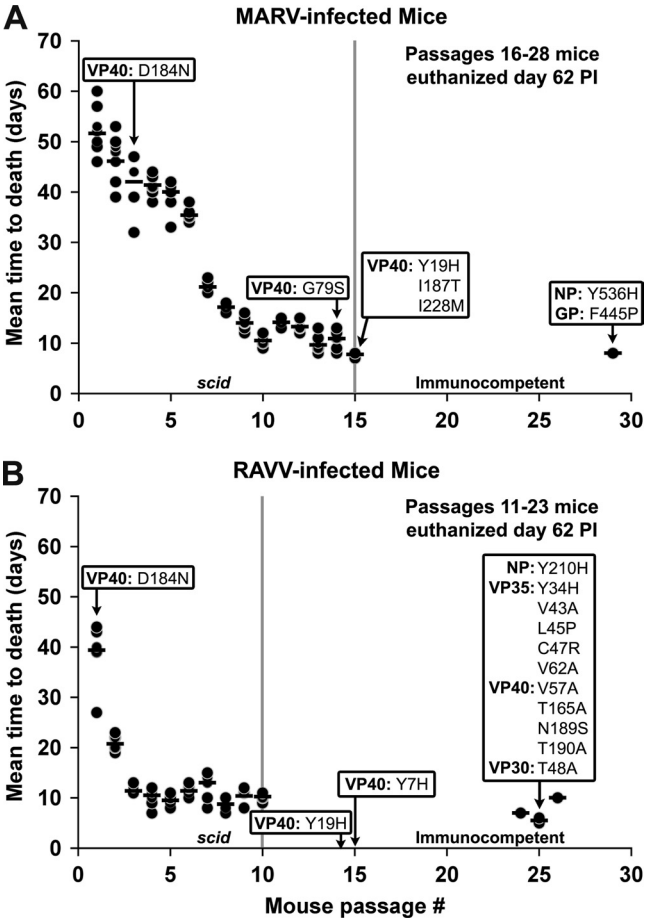


FIG. 1. Order of occurrence of mutations in reference to MTD. Virus-infected *scid* or immunocompetent mouse livers at each passage provided sequential samples to determine the order of occurrence of each of the mutations. Two common nonsynonymous mutations were identified between the lm MARV and RAVV variants. The first common mutation, VP40:D184N, occurred at passage 3 for MARV and passage 1 for RAVV. The other common mutation, VP40:Y19H, occurred at passage 15 for MARV and passage 14 for RAVV. The MTDs at each passage are indicated by a bar for each group of mice. As indicated on the graph, there were no mouse deaths at passages 16 to 28 with MARV and 11 to 23 with RAVV.

*VP35* mutation to be reported for other filoviruses, as a single *VP35* nonsynonymous mutant (Cyt3163Thy) in the lethal mouse EBOV variant was described previously (13). Mutations in *VP40* were prevalent in both lm marburgvirus variants, with approximately twice the nonrevertant mutations in the lm RAVV variant (13 mutations) compared to the lm MARV variant (7 mutations). There were two common nonsynonymous mutations in the lm MARV and lm RAVV variants, both in *VP40*. These mutations were at VP40:D184N and VP40:Y19H. There were unique nonsynonymous mutations in the lm MARV variant in the *NP* and *GP* genes and in the lm RAVV variant in the *NP*, *VP35*, and *VP30* genes; however, previously described domains have not been identified for regions that contain these unique mutations.

**Distinct order of mutation occurrences.** The genomic sequences of viruses isolated from mice infected with wt MARV and wt RAVV variants at each passage were used to determine

TABLE 1. Lethal mouse MARV and RAVV variant mutations throughout serial transfer passages[a]

| Gene | nt | aa | RAVV | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Genomic nucleotide for mouse strain | | | | | | | | | | | | | | | | | | | | | |
| | | | wt | *scid* at passage: | | | | | | Immunocompetent at passage: | | | | | | | | | | | | | | |
| | | | | 1 | 2 | 3 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 15 | 19 | 22 | 23 | 24 | 25 | 26 | 28 | 28 |
| *NP* | 731 | Y210H | T | | T | T | | | | T | T | | T | T | T | T | T | T | T | C | C | C | C |
| | 733 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | T | C | C | C | C |
| | 739 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | T | C | C | C | C |
| *ncNP* | | | | | | | | | | | | | | | | | | | | | | | | |
| *VP35* | 2964 | M7T | T | | T | T | C | C | T | T/N | T | C | N/C | C | C | C | C | C | C | T/C | T | T | T | T |
| | 3032 | L30P | T | | T | T | C | C | C/T | N | N | C | C | C | C | C | C | C | C | C | T | T | T | T |
| | 3033 | | T | | T | T | C | C | C/T | T/N | T/N | C | C | C | T | C | C | C | C | T | T | T | T | T |
| | 3044 | Y34H | T | | T | T/N | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 3072 | V43A | T | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 3077 | L45P | T | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 3078 | | T | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 3083 | C47R | T | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 3085 | | T | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 3129 | V62A | T | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 3391 | | T | | T | T | C | T/C | C | T | T | T/C | C | C | C | C | C | C | T/C | T | T | T | T | T |
| *ncVP35* | 4134 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | C | C | C | C | C | C |
| | 4139 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | C | C | C | C | C | C |
| | 4157 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | C | C | C | C | C | C |
| | 4173 | | T | | T | T | | | | T/N | T | | T | T | T | T | T | T | C | C | C | C | C | C |
| | 4174 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | C | C | C | C | C | C |
| | 4188 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | C | C | C | C | C | C |
| | 4217 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | C | C | C | N | C | C |
| *ncVP40* | 4567 | | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | | T/C | T | T | T | T | T |
| *VP40* | 4586 | Y7H | T | T | | T | T | T | T | T | T | T | T | N | C | C | C | C | C | C | C | C | C | C |
| | 4588 | | T | T | | T | T | T | T | T | T | T | T | C | C | C | C | C | C | C | C | C | C | C |
| | 4621 | Y19H | T | T | | T | T | T | T | T/N | T | T | C | C | C | C | C | C | C | C | C | C | C | C |
| | 4622 | | T | T | | T | T | T | T | T | T | T | C | C | C | C | C | C | C | C | C | C | C | C |
| | 4624 | | T | | | T | T | T | T | T | T | T | T | C | C | C | C | C | C | T/C | T | T | T | T |
| | 4627 | | T | T | | T | T | T | T | T | T | T | T | C | C | C | C | C | C | C | C | C | C | C |
| | 4630 | | T | T | | T | T | T | T | T | T | T | T | C | C | C | C | C | C | C | C | C | C | C |
| | 4633 | | T | T | | T | T | T | T | T | T | T | T | C | C | C | C | C | C | T | T | T | T | T |
| | 4636 | | T | T | | N | T | T | T | N | N | T | T | N | C | N | N | N | C | T/N | N | T/N | T | T |
| | 4647 | Y27S | T | T | | T | T | T | T | T | T | T | T | C | C | C | C | C | C | T | T | T | T | T |
| | 4665 | L34P | T | T | | T | T | T | T | T | T | T | T | C | C | C | C | C | C | T | T | T | T | T |
| | 4737 | V57A | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 4738 | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C | C | C |
| | 4902 | Q112L | T | A | A | T | T | T/A | A | A | A | A | T/A | T | T | T | T | T/A | T | A | A | T/A | A | A |
| | 5060 | T165A | T | A | A | A | A | A | A | A | A | A | A | A | A | A | G/A | A | A | G | G | G | G | G |
| | 5117 | D184N | G | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A |
| | 5133 | N189S | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | G | G | G | G | G |
| | 5135 | T190A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | G | G | G | G | G |
| | 5449 | | A | A | A | N | G | G/A | A | G/N | N | A | G | G | G | G | G | G | G | A | A | A | A | A |
| *GP* | | | | | | | | | | | | | | | | | | | | | | | | |
| *ncGP* | 8268 | | T/C | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8284 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8290 | | T/C | | | | T | T/C | | | | T/C | | | T | | | T | C | C | | | C |
| | 8298 | | T/C | | | | T | T | | | | T/C | | | T | | | T | C | C | | | C |
| | 8301 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8305 | | T/C | | | | T | T | | | | T/C | | | T | | | T | C | C | | | C |
| | 8307 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8308 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8334 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8337 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8339 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8340 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8341 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8342 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8376 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8390 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8391 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8399 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8410 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8411 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8419 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8424 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8426 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8429 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8440 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8447 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8448 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8449 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8454 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8457 | | T | | | | T | T | | | | T | | | T | | | T | C | C | | | C |
| | 8600 | | A | | | | A | A | | | | A | | | A | | | | | G | | | G |
| *ncVP30* | | | | | | | | | | | | | | | | | | | | | | | | |
| *VP30* | 9010 | T48A | A | A | | | A | A | | | | | | | A | | | | | G | | | | T |
| *L* | 16375 | | T | | N | T/N | | | | T | T | | T | T/N | T | T | T/N | T | T | C | C | C | C | |
| *TR* | 19068 | | T | | T | T | | | | T | T | | T | T | T | T | T | T | T | C | C | C | C | |

[a] The negative-sense genome sequence was determined with amplified cDNA templates; therefore, Thy (T) residues are reported. There were 33 differences between wt and lm MARV and 75 differences between wt and lm RAVV variant genomes based on nucleotide sequence alignments. Nucleotides in shaded cells indicate the first passage at which there was a nucleotide difference between the passaged viral genome and that of the wt virus. Both nonsynonymous and synonymous mutations were detected in the translated regions. The mutations were verified throughout the passage history to determine the order of occurrence of each mutation.

TABLE 1—*Continued*

| nt | aa | MARV | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Genomic nucleotide for mouse strain | | | | | | | | | | | | | | | | | | | | |
| | | wt | *scid* at passage: | | | | | | | | | | | | | Immunocompetent at passage: | | | | | |
| | | | 1 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 13 | 14 | 15 | 15 | 16 | 17 | 18 | 30 | 31 | 31 |
| 156 | | C | | | T | | | T | T | T | T | T | T | T | T | C | | | | T | T | T |
| 1708 | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | | | | T | T | T |
| 1709 | Y536H | A | A | A | A | A | A | A | A | A | A | A | A | A/C | A/C | A | A/C | A/C | A/C | C | C | C |
| 2161 | | G | G | G | G | G | G | G | G | G | G/A | G/A | A | A/C | A | A | A | A | A | | | A |
| 2501 | | T | T | T | T | T | T | T | T | T | T | T | T/C | C/T | C | C | C | C | C | | | C |
| 2503 | | T | T | T | T | T | T | T | T | T | T | T | T/C | C/T | C | C | C | C | C | | | C |
| 2545 | | T | T | T | T | T | T | T | T | T | T | T | T/C | C/T | C | C | C | C | C | | | C |
| 4352/3 | | | | | | | | | | | | | | | | | | | | | | T |
| 4620/1 | Y19H | T | T | T | T | T | T | T | T | T | T | T/C | T/C | C/T | C | C | | | | C | C | C |
| 4621/2 | | T | T | T | T | T | T | T | T | T | T | T/C | T/C | C/T | C | C | | | | C | C | C |
| 4801/2 | G79S | G | G | G | G | G | G | G/A | G/A | G/ | T | A | A | A/C | A | A | | | | A | A | A |
| 5116/7 | D184N | G | G | A | A | A | A | A | A | A | A | A | A | A/C | A | A | | | | A | A | A |
| 5126/7 | I187T | T | T | T | T | T | T | T | T | T | T/C | T/C | C/T | C/T | C | C | | | | C | C | C |
| 5250/1 | I228M | A | A | A | A | A | A | A | A | A | A | A/G | G/A | G/A | G | G | | | | G | G | G |
| 7094/5 | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C | C |
| 7272/3 | F445P | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | C | C/T | C |
| 7273/4 | | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T/C | T | T | C | C/T | C |
| 8781/2 | | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G/A | A | A |
| 17404/5 | | G | G | G | G | G | G | G | G | G | G | G/A | A/G | A/G | A/G | A | | | | A | A | A |
| 17917/8 | | C | C | C | C | C | C | C | C | C | C/T | C/T | T/C | T/C | T | T | | | | T | T | T |

the order of occurrence of each mutation. Mutations were selected by the comparison of the respective wt and lm marburgvirus variant full-length genomes. Genomic regions in proximity (~100 to 200 nt upstream and downstream) to each mutation were selected for analysis and subsequent alignments. The order of mutation occurrence was determined based on the passage in which each mutation occurred in relation to subsequent mutations; thus, the passage of wt MARV or RAVV variants of approximately 30 passages for each variant provided viral genomes for analysis at each passage for the determination of the order of occurrence of mutations required to establish lethality in mice. The results showed a distinct order for the two common mutations, i.e., the VP40:D184N mutation was the first mutation to occur, and the VP40:Y19H mutation occurred at halfway (passage 14 to 15) through the sequential serial transfer for both variants. The nonsynonymous mutations are shown in relation to the MTD (days) of the mice for reference (Fig. 1). Also of interest are the multiple mutations in the lm RAVV variant *VP35*, *VP40*, and *GP* gene transcriptional nc units that occurred in passages 24 and 25.

**Mutations in 3′-UTRs.** The secondary RNA structures were predicted for the *VP35* (nt 3935 to 4410) and glycoprotein (*GP*) (nt 7992 to 8670) gene 3′-UTR lm marburgvirus sequences (Fig. 2) using the RNAfold algorithm (18). These specific regions contained multiple mutations that consisted of Thy-to-Cyt transition mutations. RNAfold analysis provided a weighted prediction of the covariance rank of the lowest energy of the consensus nucleotide structure to determine the minimum free energy of the RNA secondary structure (18). The lm MARV variant 3′-UTRs had three *NP* Thy-to-Cyt transitions and one *VP30* Ade-to-Gua transition. The lm RAVV variant *GP* 3′-UTR consisted of 30 Thy-to-Cyt mutations and 1 Ade-to-Gua mutation, and the *VP35* 3′-UTR had seven Thy-to-Cyt mutations, whereas the lm MARV variant had no *VP35* or *GP* nc mutations. The predicted minimum free energies of the RNA secondary structure of the lm RAVV variant *GP* and *VP35* 3′-UTRs indicate a more stable structure for the wt marburgvirus RNA secondary structures than for the lm marburgvirus variant based on the kcal/mol value of the composite structure (Fig. 2).

**Codon usage.** The protein translation codon usages of the wt and lm marburgvirus variants were compared for frequency of use in the *Mus musculus* (house mouse) genome. The house mouse (designation gbrod, 5,3036 coding sequences [CDSs] [24,533,776 codons]) codon frequencies are available through the Codon Usage Database, NCBI GenBank flat file release 160.0 (15 June 2007), and were used for this analysis. Viral codon usage switching to house mouse codon usage was determined based on the greater frequency of usage of the lm MARV or lm RAVV variant codons than that of wt marburgvirus variants (Table 2). Analysis of the codon usage indicates that codon switching had biological significance, as 50% of the mutations (T to C or G to A) occurred in the third codon position switch to laboratory mouse codon usage (seven switch/seven nonswitch). Sixty-seven percent (six switch/six nonswitch) of the second codon positions and 64% (nine switch/five nonswitch) of the first codon positions switched to mouse codon usage, indicating a bias toward the mouse genome in the first and second codon position mutations. In addition, the MARV genome has ~62% A+T bias for nucleotide usage as a probable result of host tRNA availability. These data indicate that codon usage is due to host mutation pressure.

**Phylogenetic analysis.** The wt RAVV variant (GenBank accession number EU500827) and lm RAVV variant (accession number EU500826) were previously reported (47). An evolutionary profile of full-length filovirus genomes available in GenBank was compiled by using MegAlign (Fig. 3). The resultant cladogram indicated two major branches distinguishing members of the filovirus genera *Marburgvirus* and *Ebolavirus* as well as distinct branches representing the four lineages of MARV and the one lineage of RAVV within the marburgvirus clade and the five lineages of ebolaviruses representing individual species (40). The lm MARV variants are included in the MARV lineage, and the lm RAVV variants are included in the RAVV lineage. These two lineages are distinguished by approximately 26.2% nucleotide divergence between the closest members of each lineage.

**Viral challenges with lm marburgvirus variants.** Immunocompetent BALB/c and C57BL/6 mice were infected with lm MARV or RAVV variants. Surviving animals were considered censored at day 14, the last day of the study (Fig. 4). Both mouse strains infected with the lm RAVV variant survived 2 days less than did equivalent mice infected with the lm MARV variant ($P = 0.0014$) (Table 3). There were no survivors after day 8 p.i. for the BALB/c mice or after day 12 p.i. for the C57BL/6 mice, whereas there was a 20% survival rate for the lm MARV variant-infected BALB/c mice and a 40% survival rate for the C57BL/6 mice on day 14 p.i. The lm MARV variant-infected BALB/c mice had a statistically suggestive but not statistically significant lower median survival time than did identically infected C57BL/6 mice ($P = 0.0502$) (Table 3). In contrast, lm RAVV variant-infected BALB/c mice had a significantly lower median survival time than did identically infected C57BL/6 mice ($P = <0.0001$). BALB/c mice infected with the lm RAVV variant had significantly lower median survival times than did the lm MARV variant-infected mice ($P = <0.0001$), and similarly, C57BL/6 mice infected with the lm RAVV variant had significantly lower median survival times than did mice infected with the lm MARV variant ($P = 0.0053$). Additionally, none of the lm RAVV variant-infected BALB/c and C57BL/6 mice survived, whereas the lm MARV variant-infected BALB/c and C57BL/6 mice had 20% and 40% survival rates, respectively. This study indicated that BALB/c mice were clearly more susceptible when infected with the lm RAVV variant (the results were not as decisive for the lm MARV variant) based on survival rates. In addition, the lm RAVV variant demonstrated greater virulence than the lm MARV variant in both mouse strains.

Next, viral propagation in major organs from lm marburgvirus variant-infected BALB/c mice was assessed by viral plaque assays of serial dilutions of tissue homogenates (Fig. 5). Lethal mouse MARV variant concentrations were generally greater than those of the lm RAVV variants on days 1, 3, and 5 but showed a downward trend in most tissues except the intestines on day 7. Specifically, there was a statistically significant ($P \leq 0.01$) viral concentration increase in the lm MARV variant-infected mouse tissues at day 1 in the lungs (3.38 $\log_{10}$ PFU/ml), whereas the lm RAVV variant had no detectible viral concentration in these tissues at this time point. There were statistically significant increases in viral concentrations in

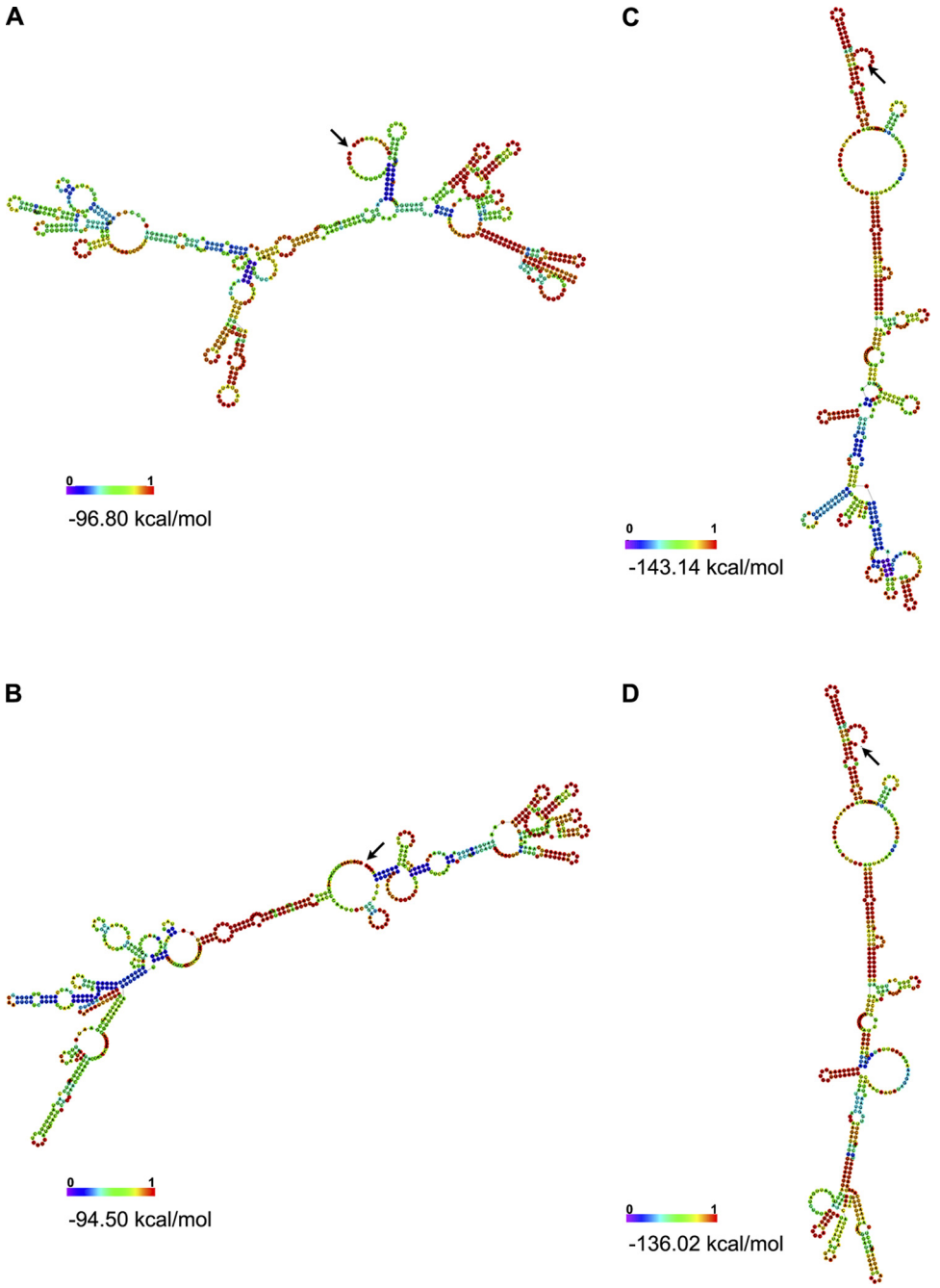


FIG. 2. Comparison of predicted RNA secondary structure minimum free energies. wt RAVV (A) and lm RAVV (B) GPs are shown on the left, and wt RAVV (C) and lm RAVV (D) VP35 3′-UTR secondary structures are shown on the right. The arrows indicate the 3′ and 5′ ends of each UTR. The base pair probabilities as indicated by the normalized scale of 0 (blue) to 1 (red) were determined with the Vienna Websuite RNAfold algorithm with default parameters. The predicted minimum free energies of the lm RAVV variant GP and VP35 3′-UTRs indicate less stable structures than those for the wt variant structures based on the kcal/mol values.

various organs infected with the lm MARV variant versus the lm RAVV variant at day 3, i.e., heart (3.95 versus 0 $\log_{10}$ PFU/ml), spleen (5.41 versus 4.45 $\log_{10}$ PFU/ml), and liver (5.40 versus 0 $\log_{10}$ PFU/ml) tissues, and in the liver at day 5 (6.60 versus 3.33 $\log_{10}$ PFU/ml). The lm RAVV variant concentration had a time course comparable to that of the lm MARV variant concentrations for gonads, heart, and kidney; however, although not an increase of statistical significance, the lm RAVV variant-infected brain tissue concentration was increased over that of the lm MARV variant at day 5 (0 versus 2.47 $\log_{10}$ PFU/ml) and day 7 (1.49 versus 3.10 $\log_{10}$ PFU/ml).

## DISCUSSION

A novel approach to establish MHF mouse models incorporated passages of virus in infected *scid* mouse liver homoge-

TABLE 2. Switching of virus to host codon usage[a]

| Isolate | Gene | Nucleotide | Wild type | | Immunocompetent | |
|---|---|---|---|---|---|---|
| | | | Codon | Frequency[b] | Codon | Frequency[b] |
| MARV | *NP* | 1709 | TAC | 16.1 | TCC | 18.1 |
| | | 2161 | GAA | 27.0 | AAA | 21.9 |
| | *VP40* | 4620/4621/4622 | CTT | 13.4 | CCC | 18.2 |
| | | 4801/4802 | GGT | 11.4 | AGT | 12.7 |
| | | 5116/5117 | GAC | 26.0 | AAC | 20.3 |
| | | 5126/5127 | ATT | 15.4 | ACT | 13.7 |
| | | 5250/5251 | ATA | 7.4 | ATG | 22.8 |
| | *GP* | 7094/7095 | CCT | 18.4 | CCC | 18.2 |
| | | 7272/7273/7274 | TTT | 17.2 | CCT | 18.4 |
| | *L* | 17404/17405 | ACG | 5.6 | ACT | 16.0 |
| | | 17917/17918 | CGC | 9.4 | CGT | 4.7 |
| RAVV | *NP* | 731/733 | TAT | 12.2 | CAC | 15.3 |
| | | 739 | AGT | 12.7 | AGC | 19.7 |
| | *VP35* | 3044 | TAT | 12.2 | CAT | 10.6 |
| | | 3072 | GTG | 28.4 | GCG | 6.4 |
| | | 3077/3078 | TTA | 6.7 | CCA | 17.3 |
| | | 3083/3085 | TGT | 11.4 | CGC | 9.4 |
| | | 3129 | GTT | 10.7 | GCT | 20.0 |
| | *VP40* | 4586/4588 | TAT | 12.2 | CAC | 15.3 |
| | | 4621 | CCT | 18.4 | CCC | 18.2 |
| | | 4622/4624 | TAT | 12.2 | CAT | 10.6 |
| | | 4627 | GCT | 20.0 | GCC | 26.0 |
| | | 4630 | GAT | 21.0 | GAC | 26.0 |
| | | 4737/4738 | GTT | 10.7 | GCC | 26.0 |
| | | 5060 | ACT | 13.7 | GCT | 20.0 |
| | | 5117 | GAC | 26.0 | AAC | 20.3 |
| | | 5133 | AAC | 20.3 | AGC | 19.7 |
| | | 5135 | ACC | 19.0 | GCC | 26.0 |
| | *VP30* | 9010 | ACA | 16.0 | GCA | 15.8 |
| | *L* | 16375 | AAT | 15.6 | AAC | 20.3 |

[a] The protein translation codons of the wt and lm MARV and RAVV variants were compared for frequency of use in the host genome (laboratory mouse). Codon usage switching of the virus to host codon usage was based on the greater frequency of usage between the lm and wt variants. All genes that contained mutations except lm RAVV VP30 demonstrated codons that switched to the host codon. Shaded codons indicate virus codon switching to host codon usage.

[b] Frequency is the number of particular codons per thousand.

nates followed by similar viral passages into immunocompetent mice. This procedure resulted in the generation of a viral population among the quasispecies that target and replicate in the liver, a major site of marburgvirus infection-induced pathology. The critical criterion to determine the utility of mice as models for MHF was the duplication of a MTD in mice equivalent to the MTD in nonhuman primates in conjunction with liver pathology.

Comparative genomic sequence analysis of wt and lm marburgvirus variants provided the basis for the identification of five mechanisms that resulted in alterations to primary and/or secondary RNA structures. The mechanisms were unidirectionally biased hypermutagenesis of pyrimidines, mutations in multiple genes, a distinct order of mutational occurrence, mutations that occurred in untranslated regions, and a switching to new host (laboratory mouse) codon usage. The lm MARV variant genome differed from the wt genome by 11 to 13 nucleotide mutations, which were similar to the numbers of mutations identified previously for other lethal rodent filovirus genomes (8, 26, 44). In addition, the lm MARV variant was able to switch to the laboratory mouse codon usage. In contrast, the lm RAVV variant genome used all five mechanisms described above, including extensive biased hypermutagenesis to generate approximately five times the number of nucleotide changes (a majority of which were Thy-to-Cyt transitions) in the lm RAVV variant genome. Although hypermutation has not been described for other lethal rodent filoviruses, multiple mutations have been reported. Mutations in multiple genes have been identified in lethal mouse immunocompetent EBOV (13) and lethal guinea pig MARV (26) variant genomes. Attempts have been made to implicate individual genes from among those genes involved in the lethal rodent selection process using reverse genetics; however, the complement of all of the identified mutations was deemed necessary to efficiently induce a lethal infection in rodents (13, 14). Also consistent with the EBOV adaptation studies, our previous studies (47) of the viral selection of the lm marburgvirus variants reported the involvement of the transient production of type I interferon (IFN) and may contribute to the ability of lm RAVV to escape viral clearance from the brain as described for tissue tropism studies.

The identification of unidirectionally biased hypermutagenesis or, more specifically, Ade-to-Ios editing by adenosine deaminases acting on RNA (ADARs) of the RAVV variant genome was of particular interest. ADARs bind to double-stranded RNA (dsRNA) structures and catalyze the conversion of Ade to Ino (reviewed in references 2, 20, 34, and 42). Inosine is read as Gua by translational and reverse transcriptional machinery, and bases pair with Cyt; this results in the replacement of a Cyt instead of a Watson-Crick complemen-

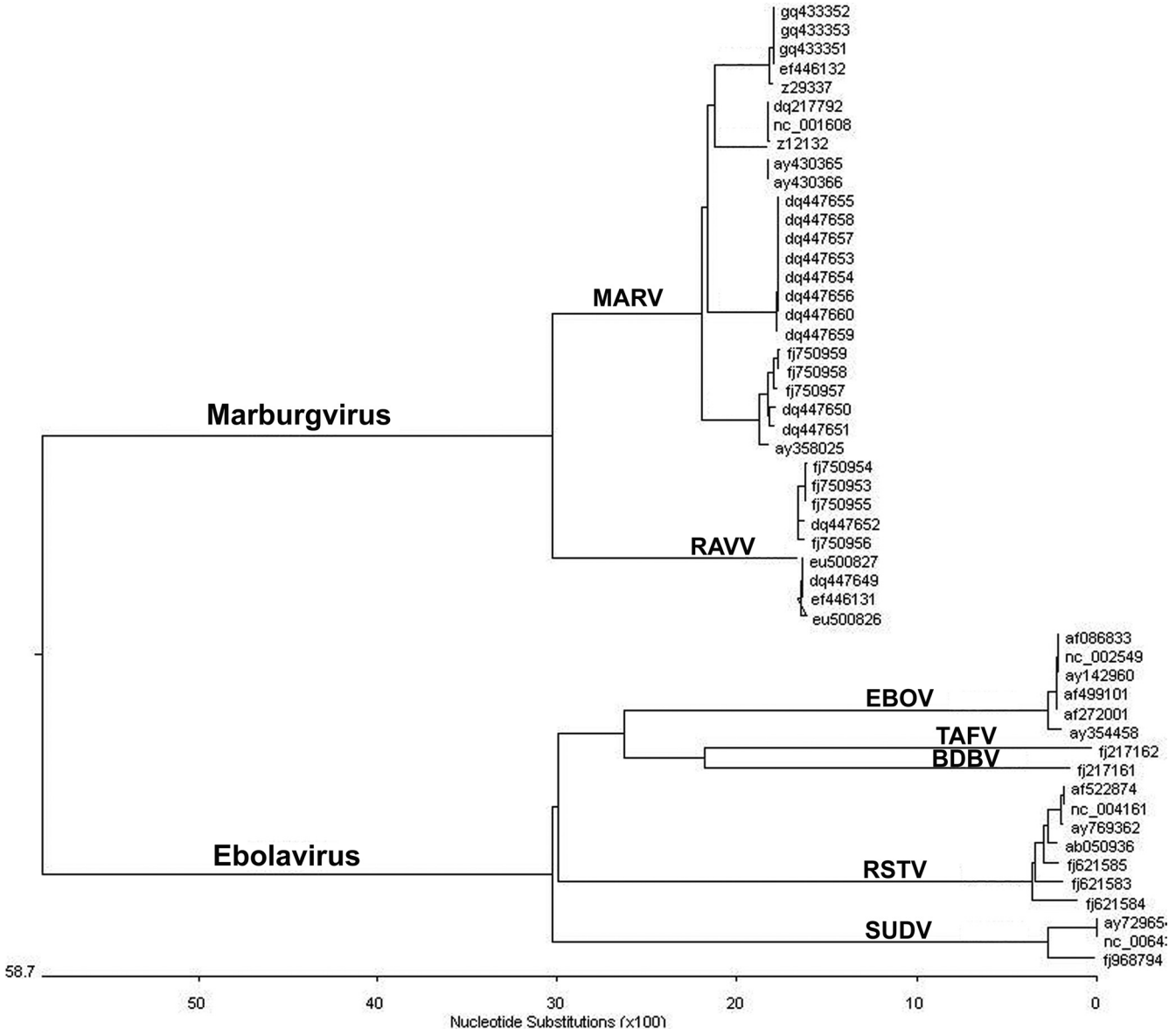


FIG. 3. Cladogram of evolutionary relationships predicted from the multiple-sequence alignment of full-length filoviruses. The length of each branch pairs represents the distance between viruses, whereas the units at the base of the tree indicate the number of substitution events. The lm MARV variants are included in the MARV lineage or upper marburgvirus branch, and the lm RAVV variants are included in the RAVV lineage or lowest marburgvirus branch. GenBank accession numbers are shown at the far right.

tary Thy base pair with Ade. Ade-to-Ios editing has been reported for diverse viral RNAs (reviewed in reference 19); however, despite the recognition of Ade-to-Ios-biased hypermutation in viral transcripts during persistent and lytic infections, the function and consequences of editing are still not known (reviewed in reference 19), although examples of direct ADAR antiviral editing were reported previously for lymphocytic choriomeningitis virus (LCMV) (50) and measles virus (45). *In vitro* and *in vivo* studies of LCMV infection demonstrated high rates of ADAR1-specific Ade-to-Ios mutations leading to dysfunctional glycoproteins and impaired viral infectivity.

Three ADARs have been defined as dsRNA-editing enzymes; two of these (ADAR1 and ADAR2) demonstrate adenosine deaminase activity. ADAR1 has two differentially localized isoforms: the large ADAR-L or ADAR p150 is present in both the cytosol and nucleus and is regulated by an interferon-inducible promoter. ADARs are highly expressed in brain tissue with site-specific ADAR editing in the brain, notably in glutamate and serotonin receptor mRNAs (reviewed in references 2, 20, and 42), which leads to single-codon changes and the resultant translated proteins with altered physiological functions.

The numbers of nonsynonymous mutations deduced from the nucleotide sequence were 8 and 19 for lm MARV and lm RAVV variants, respectively. Two identical amino acid changes occurred in VP40. One of the two mutations occurred in the PPPY motif (VP40:Y19H) that was proposed to mediate

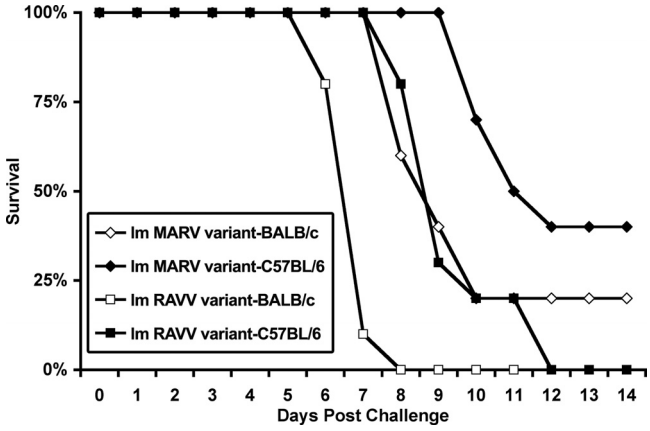


FIG. 4. *In vivo* survival studies with immunocompetent mice with different MHC *H-2* haplotypes. BALB/c and C57BL/6 mice ($n = 5$ for each group) were infected i.p. with 1,000 PFU of wt or lm MARV and RAVV variants and were observed for 14 days p.i. to compare survival rates. Both mouse strains infected with the lm RAVV variant had a survival of 2 days less than that of equivalent mice infected with the lm MARV variant. There were no survivors after day 8 p.i. for the BALB/c mice or after day 12 p.i. for the C57BL/6 mice, whereas there was a 20% survival rate for the lm MARV variant-infected BALB/c mice and a 40% survival rate for the C57BL/6 mice on day 14 p.i. The statistical analysis for this experiment is presented in Table 3.

the cellular release of filovirus-like particles (41). The other VP40 mutation (VP40:D184N) was located at the interdomain connecting loop between the VP40 N- and C-terminal domains based on the crystal structure of EBOV VP40 (12, 15). This analysis was based on the nucleotide sequence alignment of EBOV variant Mayinga (GenBank accession number AF086833) with MARV variant Poppinga (GenBank accession number Z29337) (26). The function of the lm marburgvirus variant VP40:D184 interdomain connecting loop mutation has not been elucidated; however, marburgvirus VP40 is multifunctional (21, 33), and the loop may play a role in the generation of conformational changes for the various functions of this protein. A type I IFN antagonist function was recently reported for the marburgvirus *VP40* gene (43). Type I IFN was hypothesized previously to activate ADAR (3). Thus, the lm RAVV variant may interfere with two IFN-inducible pathways, i.e., the induction of ADAR and the ability to regulate IFN antagonist function via VP40.

The successful effort to establish marburgvirus mouse models not only produced viral variants capable of causing lethality in mice for experiments that require challenge virus but also provided a unique opportunity to investigate the genomic mutations that accumulate throughout the passages of an attenuated virus to generate a lethal mouse variant. The VP40:D184N mutation was indicated as a key change for the *in vivo* virus selection of lm marburgvirus variants, as it appeared before passage 4 in both the MARV and RAVV variants. This mutation was also described previously for the lethal guinea pig MARV variant (26). Taken together, this mutation identified in two lethal rodent marburgvirus models was designated a trigger event to initiate the process to select for the lm marburgvirus variants.

Marburgvirus VP35 acts as a cofactor of RNA-dependent RNA polymerase and is essential for viral transcription and replication (32). VP35 may have a secondary role as a determinant of viral assembly based on the mutation study of a marburgvirus VP35 Δ1-122aa construct, which was able to bind NP but prevented the formation of typical intracytoplasmic inclusion bodies (31). The seven lm RAVV variant VP35 mutations all occur in the first 62 amino acids of the protein, and studies are ongoing in our laboratory to determine the role of a Δ1-62aa deletion in the lm Ravn variant. Marburgvirus genomes consist of seven transcriptional units that encode seven structural proteins in the order 3′-*LE-NP-VP35-VP40-GP-VP30-VP24-L-TR*-5′. These transcriptional units are separated by UTRs that consist of a 3′ region with transcriptional and translational termination signals for the upstream gene, a nontranscribed intergenic region, and a 5′ region that contains transcriptional and translation initiation signals for the downstream gene. Other UTRs include the 3′ leader (LE) and 5′ trailer (TR) regions, which contain signals for replication start and RNA packaging (27). UTRs are known to play crucial roles in the posttranscriptional regulation of gene expression (28). 3′-UTRs can specifically control the nuclear transport, polyadenylation status, subcellular targeting and rates of translation, and degradation of mRNA (reviewed in reference 10). The average length of 5′-UTRs in RNA viruses is constant over diverse taxonomic classes and ranges between 100 and 200 nucleotides, whereas the average length of 3′-UTRs is much more variable, ranging from about 200 nucleotides in plants and fungi to 800 nucleotides in humans and other vertebrates. The lengths of both 3′- and 5′-UTRs vary considerably within a species, ranging from a dozen nucleotides to a few thousand (36). The numbers of nucleotides in the marburgvirus 5′-UTRs range from 1 to 208 and the numbers of nucleotides in the 3′-UTRs range from 317 to 684. The *GP* gene has the largest 3′-UTR, 684 nt. The long 3′-UTR of the matrix (M) mRNA (greater than 1,000 nucleotides) of measles virus (38) and canine distemper virus (1) was previously shown to increase the level of M protein production and viral replication. It is noteworthy that the long 3′-UTR of the SSPV M protein and the 5′-UTR region of the F gene of measles virus both have unusually high Cyt/Urd ratios, and many of the Urd-versus-Cyt differences between the Nagahata and Edmonston M genes occur in the 3′-UTR region. Only one 5′-UTR mutation was discovered in lethal mouse MARV variants, i.e., in

TABLE 3. Virulence of lm marburgviruses in mouse strains with different MHC *H-2* haplotypes[a]

| Comparison | *P* value | |
|---|---|---|
| | Time to death | Survival time |
| BALB/c vs C57BL/6 (lm MARV variant) | **0.0014** | 0.0502 |
| BALB/c vs C57BL/6 (lm RAVV variant) | **0.0002** | **<0.0001** |
| MARV vs RAVV (BALB/c) | **<0.0001** | **<0.0001** |
| MARV vs RAVV (C57BL/6) | 0.092 | **0.0053** |

[a] Statistically significant *P* values for time to death and survival times (days) of immunocompetent BALB/c and C57BL/6 mice infected with either the lm MARV or RAVV variant are shown. Detailed graphs of this statistical analysis are provided in Fig. 4. Statistically significant ($P < 0.0001$) values are in boldface type. This study indicated that BALB/c mice were more susceptible to the lm RAVV variant than were C57BL/6 mice based on time to death and survival rates, and the RAVV variant demonstrated greater virulence than did the lm MARV variant in both mouse strains.

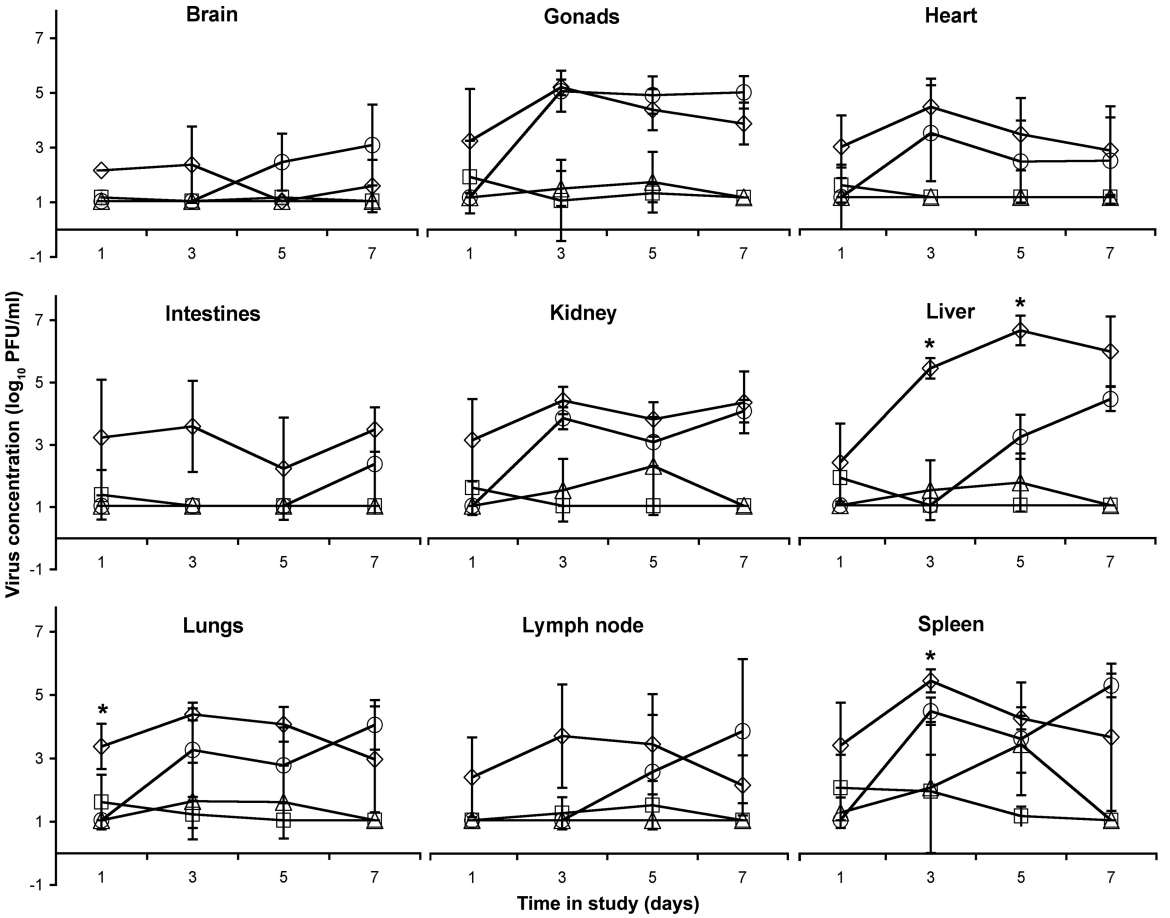


FIG. 5. Tissue tropisms of the wt and lm MARV and RAVV variants (□, wt MARV; ◇, lm MARV; △, wt RAVV; ○, lm RAVV). Phenotypic differences among viruses were based on tissue tropisms of variants in infected immunocompetent BALB/c mouse tissues at 1, 3, 5, and 7 days p.i. ($n = 5$ at each time point). The viral concentrations were statistically significantly increased in lm MARV versus lm RAVV variants (asterisks) in lung tissue at day 1 p.i., the spleen and liver tissues at day 3 p.i., and liver at day 5 p.i. The lm RAVV variant viral concentrations in gonad, heart, and kidney tissues were comparable to those of the lm MARV variant; however, the concentration was increased in the brain, intestines, liver, lungs, lymph node, and spleen at later time points (days 5 and 7 p.i.).

MARV *VP30*. There were three 3′-UTR mutations in MARV *NP*, and there were 7 *VP35* and 31 *GP* mutations in the RAVV variant. The RAVV 3′-UTRs were analyzed by RNA structure prediction of the nucleic acid sequence to serve as a tool to generate a viral RNA structure-function relationship hypothesis (37).

One genomic adaptation mechanism used by both lm marburgvirus variants enabled the virus to substitute the more frequently used codon of the host (laboratory mice) for the codon of the wt viral genome to be used in the lm marburgvirus variants. Codon usage refers to the frequency of codon occurrence in translated DNA sequences and was initially reported by Grantham et al. (17). In some organisms, highly expressed genes have a strong codon preference in relation to the concentration of corresponding tRNAs, whereas genes expressed at a lower level have a more uniform pattern of codon use.

As indicated by the genomic analysis of marburgvirus variants, the lm marburgvirus variants have different mechanisms for *in vivo* virus selection from among quasispecies. In addition, the marburgvirus variant differences were exemplified by the survival times of infected immunocompetent BALB/c and C57BL/6 mice. The lm RAVV variant was more virulent than the lm MARV variant in both mouse strains tested. BALB/c mice trended toward greater susceptibility to both marburgvirus variants, which was not unexpected, as the variants were isolated after serial transfer in BALB/c mice. BALB/c and C57BL/6 mice differ in the MHC *H-2* haplotype; i.e., BALB/c mice are $H\text{-}2^d$, and C57BL/6 mice are $H\text{-}2^b$. There are also differences in the functions of T lymphocytes, as indicated by previously reported studies of *Leishmania major*. *L. major*-resistant C57BL/6 T lymphocytes produce nitrous oxide and kill the parasite, whereas susceptible BALB/c T lymphocytes produce more interleukin-4 that suppresses macrophages (22, 29).

Lethal mouse marburgvirus variant propagation, as determined by virus concentrations in major organs of infected BALB/c mice, also indicated differences in the hosts' abilities to clear these viruses from various tissues. The viral concentrations in gonads, heart, and kidney were similar among the marburgvirus variant-infected mice; however, other organs demonstrated marked differences. The lm MARV variant concentrations in infected brain, gonads, heart, intestines, liver,

lungs, lymph nodes, and spleen were cleared from these organs after day 5. In contrast, the lm RAVV variant concentrations from infected brain, gonads, intestines, kidney, liver, lungs, lymph nodes, and spleen were increased at day 7, indicating no clearance of the variants. Taken together, the survival time for different immunocompetent mice and viral propagation in infected immunocompetent BALB/c tissues indicate phenotypic differences between the lm MARV and RAVV variants.

The lethal *scid* and immunocompetent mouse marburgvirus variants have proven to be beneficial for therapeutic studies (47) and are projected to be useful for pathogenesis studies with bioengineered mouse strains. These genomic studies have identified key changes involved throughout the developmental history of two lm marburgvirus variants that were selected from the quasispecies and were able to propagate in immunocompetent mice. This analysis of the genomic mutations in relation to the serial transfer of the viruses over 30 passages is the first report of the sequential order of mutations during the *in vivo* virus variant selection process in a new host. In addition, five genomic alteration mechanisms were utilized by the two marburgviruses over the course of mouse adaptation to describe the complexity, versatility, and ability of this virus to inflict disease in a host. These findings contribute to the knowledge base of filovirus pathogenicity at the molecular level and provide the basis for how the virus manipulates the host for the provision of enzymes (adenosine deaminase), metabolites (Ino), translation regulators (UTRs), and effectors (cytokines) of the innate immune response to enable the marburgvirus to establish lethality in a new host system.

Future marburgvirus pathogenesis studies with genetically modified mice are now possible due to the availability of lethal variants to define the pathways involved in marburgvirus pathogenesis. In addition, the mutations can be used as focal points for exploitation in vaccinology (*VP40* mutations with the virus-like particle platform) and therapeutic drug design (ADAR pathway and UTR targeting with RNA interference). Whole-genome deep sequencing of the individual passage isolates is under way to decipher the order of the emergence of lethal host viral variants in relation to host innate factors and gain insight into the population genomics of marburgvirus quasispecies.

## ACKNOWLEDGMENTS

In particular, we thank Jens H. Kuhn for sharing his expertise of filovirus taxonomy. We acknowledge Lorraine Farinick and Lewis S. Long for assistance in the production of the figures and graphs, Paul Gibbs and Sarah L. Norris for statistical analysis, and Daniel Lackner, Steven B. Bradfute, Darran J. Wigelsworth, and Michael D. Parker for their constructive reviews of the manuscript.

This work was supported by USAMRIID project plan number 03-4-4J-012 and the Defense Threat Reduction Agency (KK004_07_RD_B).

Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Army.

## REFERENCES

1. **Anderson, D. E., and V. von Messling.** 2008. Region between the canine distemper virus M and F genes modulates virulence by controlling fusion protein expression. J. Virol. **82:**10510–10518.
2. **Bass, B. L.** 2002. RNA editing by adenosine deaminases that act on RNA. Annu. Rev. Biochm. **71:**817–846.
3. **Bass, B. L., and H. Weintrab.** 1989. An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell **55:**1089–1098.
4. **Bosma, G. C., D. M. Custer, and M. J. Bosma.** 1983. A severe combined immunodeficiency mutation in the mouse. Nature **301:**527–530.
5. **Bosma, G. C., et al.** 1989. The mouse mutation severe combined immune deficiency (*scid*) is on chromosome 16. Immunogenetics **29:**54–57.
6. **Bosma, G. C., et al.** 1988. Evidence of functional lymphocytes in some (leaky) scid mice. J. Exp. Med. **167:**1016–1033.
7. **Bosma, M. J., and A. M. Carroll.** 1991. The scid mouse mutant: definition, characterization, and potential uses. Annu. Rev. Immunol. **9:**323–350.
8. **Bray, M., K. Davis, T. Geisbert, C. Schmaljohn, and J. Huggins.** 1999. A mouse model for evaluation of prophylaxis and therapy of Ebola hemorrhagic fever. J. Infect. Dis. **179:**S248–258.
9. **Bray, M., S. Hatfill, L. Hensley, and J. W. Huggins.** 2001. Haematological, biochemical and coagulation changes in mice, guinea-pigs and monkeys infected with a mouse-adapted variant of Ebola Zaire virus. J. Comp. Pathol. **125:**243–253.
10. **Conne, B., A. Stutz, and J.-D. Vassalli.** 2000. The 3′ untranslated region of messenger RNA: a molecular 'hotspot' for pathology. Nat. Med. **6:**637–641.
11. **Custer, D. M., G. C. Bosma, and M. J. Bosma.** 1985. Severe combined immunodeficiency (SCID) in the mouse. Am. J. Pathol. **120:**464–477.
12. **Dessen, A., V. Vochkov, O. Donik, H.-D. Klenk, and W. Weissenhorn.** 2000. Crystal structure of the matrix protein VP40 from Ebola virus. The EMBO J. **19:**4228–4236.
13. **Ebihara, H., et al.** 2006. Molecular determinants of Ebola virus virulence in mice. PLoS Pathol. **2:**705–711.
14. **Enterlein, S., et al.** 2006. Rescue of recombinant Marburg virus from cDNA is dependent on nucleocapsid protein VP30. J. Virol. **80:**1038–1043.
15. **Gomis-Ruth, F. X., et al.** 2003. The matrix protein VP40 from Ebola virus octamerizes into pore-like structures with specific RNA binding properties. Structure **11:**423–433.
16. **Gowen, B. B., and M. R. Holbrook.** 2007. Animal models of highly pathogenic RNA viral infections: hemorrhagic fever viruses. Antiviral Res. **78:**79–90.
17. **Grantham, R., C. Gautier, M. Gouy, M. Jacobzone, and R. Mercier.** 1981. Codon catalog usage is a genome strategy modulated for gene expressivity. Nucleic Acids Res. **9:**r43–r74.
18. **Gruber, A. R., R. Lorenz, S. H. Bernhart, R. Neuboc, and I. L. Hofacker.** 2008. The Vienna RNA Websuite. Nucleic Acids Res. **36:**W70–W74.
19. **Hamilton, C. E., N. F. Papavasiliou, and B. R. Rosenberg.** 2010. Diverse functions for DNA and RNA editing in the immune system. RNA Biol. **7:**1–10.
20. **Keegan, L., A. Leroy, D. Sproul, and M. O'Connell.** 2004. Adenosine deaminases acting on RNA (ADARs); RNA-editing enzymes. Gen. Biol. **5:**209.
21. **Kolesnikova, L., H. Bugany, and H.-D. Klenk.** 2000. VP40, the matrix protein of Marburg virus, is associated with the membranes of the late endosomal compartment. J. Virol. **76:**1825–1836.
22. **Kostka, S. L., et al.** 2009. IL-17 promotes progression of cutaneous leishmaniasis in susceptible mice. J. Immunol. **182:**3039–3046.
23. **Kuhn, J. H.** 2008. Clinical presentation of filoviral disease, p. 99–120. *In* C. H. Calisher (ed.), Filoviruses: a compendium of 40 years of epidemiological, clinical, and laboratory studies. Springer-Verlag, New York, NY.
24. **Kuhn, J. H.** 2008. History of filovirus disease outbreaks, p. 59–96. *In* C. H. Calisher (ed.), Filoviruses: a compendium of 40 years of epidemiologicial, clinical, and laboratory studies. Springer-Verlag, New York, NY.
25. **Kuhn, J. H., et al.** 2010. Proposal for a revised taxonomy of the family *Filoviridae*—classification, names of taxa and viruses, and virus abbreviations. Arch. Virol. **155:**2083–2103. doi:10.1007/s00705-010-0814-x.
26. **Lofts, L. L., M. S. Ibrahim, D. L. Negley, M. D. Hevey, and A. L. Schmaljohn.** 2007. Genomic differences between guinea pig lethal and nonlethal Marburg virus variants. J. Infect. Dis. **196S:**S305–S312.
27. **Mavrakis, M., L. Kolesnikova, G. Schoehn, S. Becker, and R. W. H. Ruigrok.** 2002. Morphology of Marburg virus NP-RNA. Virology **296:**300–307.
28. **Mignone, F., C. Gissi, S. Liuni, and G. Pesole.** 2002. Untranslated regions of mRNAs. Gen. Biol. **3:**4.1–4.10.
29. **Mills, C. D., K. Kincaid, J. M. Alt, M. J. Heilman, and A. M. Hill.** 2000. M-1/M-2 macrophages and the Th1/Th2 paradigm. J. Immunol. **164:**6166–6173.
30. **Moe, J. B., R. D. Lambert, and H. W. Lupton.** 1981. Plaque assay for Ebola virus. J. Clin. Microbiol. **13:**791–793.
31. **Moller, P.** 2005. Die Rolle der Homooligomerisierung des Polymerasekofaktors VP35 im Vermehrungszyklus des Marburg-Virus. Philipps Universitat Marburg, Marburg, Germany.
32. **Muhlberger, E., B. Lotfering, H. D. Klenk, and S. Becker.** 1998. Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. J. Virol. **72:**8756–8764.
33. **Muhlberger, E., et al.** 1996. Termini of all mRNA species of Marburg virus: sequence and secondary structure. Virology **223:**376–380.

33a. **National Research Council.** 1996. Guide for the care and use of laboratory animals. National Academy Press, Washington, DC.

34. **Nishikura, K.** 2006. Editor meets silencer: crosstalk between RNA editing and RNA interference. Nat. Rev. Mol. Cell. Biol. **7:**919–931.
35. **Pe'ery, T., and M. B. Mathews.** 2007. Viral conquest of the host cell, p.

169–208. *In* D. M. Knipe et al. (ed.), Fields virology, 5th ed., vol. 1. Lippincott Williams & Wilkins, Philadelphia, PA.
36. **Pesole, G., et al.** 2001. Structural and functional features of eukaryotic mRNA untranslated regions. Gene **276:**73–81.
37. **Schroeder, S. J.** 2009. Advances in RNA structure prediction from sequence: new tools for generating hypotheses about viral RNA structure-function relationships. J. Virol. **83:**6326–6334.
38. **Takeda, M., et al.** 2005. Long untranslated regions of the measles virus M and F genes control virus replication and cytopathogenicity. J. Virol. **79:** 14346–14354.
39. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
40. **Towner, J. S., et al.** 2006. Marburgvirus genomics and association with a large hemorrhagic fever outbreak in Angola. J. Virol. **80:**6497–6516.
41. **Urata, S., et al.** 2007. Interaction of Tsg101 with Marburg virus VP40 depends on the PPPY motif, but not the PT/SAP motif as in the case of Ebola virus, and Tsg101 plays a critical role in the budding of the Marburg virus-like particles induced by VP40, NP, and GP. J. Virol. **81:**4895–4899.
42. **Valente, L., and K. Nishikura.** 2005. ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. Prog. Nucleic Acid Res. Mol. Biol. **79:**299–338.
43. **Valmas, C., et al.** 2010. Marburg virus evades interferon responses by a mechanism distinct from Ebola virus. PLoS Pathol. **6:**95.
44. **Volchkov, V. E., A. A. Chepurnov, V. Volchkova, V. A. Ternovoj, and H.-D. Klenk.** 2000. Molecular characteristics of guinea pig-adapted variants of Ebola virus. Virology **277:**147–155.
45. **Ward, S. V., et al.** 2011. RNA editing enzyme adenosine deaminase is a restriction factor for controlling measles virus replication that also is required for embryogenesis. Proc. Natl. Acad. Sci. U. S. A. **108:**331–336.
46. **Warfield, K. L., et al.** 2007. Development of a model for marburgvirus based on severe-combined immunodeficiency mice. Virol. J. **4:**108.
47. **Warfield, K. L., et al.** 2009. Development and characterization of a mouse model for Marburg hemorrhagic fever. J. Virol. **83:**6404–6415.
48. **Warfield, K. L., et al.** 2006. Viral hemorrhagic fevers, p. 227–257. *In* J. R. Swearengen (ed.), Biodefense research methodology and animal models. CRC Press, Boca Raton, FL.
49. **Wong, T. C., M. Ayata, S. Sueda, and A. Hirano.** 1991. Role of biased hypermutation in evolution of subacute sclerosing panencephalitis virus from progenitor acute measles virus. J. Virol. **65:**2191–2199.
50. **Zahn, R. C., I. Schelp, O. Utermohlen, and D. von Laer.** 2007. A-to-G hypermutation in the genome of lymphocytic choriomeningitis virus. J. Virol. **81:**457–464.