VIROLOGY 238, 64–67 (1997)
ARTICLE NO. VY978814

Phylogenetic Relationship of the Complete Rauscher Murine Leukemia Virus Genome

View metadata, citation and similar papers at core.ac.uk
brought
provided by Elsevier

Anis H. Khimani,*,† Michael Lim,* Thomas G. Graf,‡ Temple F. Smith,‡ and Ruth M. Ruprecht*,†,[2]

*Laboratory of Viral Pathogenesis, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115; †Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115; and ‡BMERC, Boston University, Boston, Massachusetts 02215

Received April 25, 1997; returned to author for revision June 2, 1997; accepted September 2, 1997

We report the complete nucleotide sequence of the genome of Rauscher murine leukemia virus (R-MuLV), the replication-competent helper virus present in the Rauscher virus complex, and its phylogenetic relationship with other murine leukemia virus genomes. An overall sequence identity of 97.6% was found between R-MuLV and the Friend helper virus (F-MuLV), and the two viruses were closely related on the phylogenetic trees constructed from either *gag, pol,* or *env* sequences. Moloney murine leukemia virus (Mo-MuLV) was the next closest relative to R-MuLV and F-MuLV on all trees, followed by Akv and radiation leukemia virus (RadLV). The most distantly related helper virus was Hortulanus murine leukemia virus (Ho-MuLV). Interestingly, Cas-Br-E branched with Mo-MuLV on the *gag* and *pol* trees, whereas on the *env* tree, it revealed the highest degree of relatedness to Ho-MuLV, possibly due to an ancient recombination with an Ho-MuLV ancestor. In summary, a phylogenetic analysis involving various MuLVs has been performed, in which the postulated close relationship between R-MuLV and F-MuLV has been confirmed, consistent with the pathobiology of the two viruses. © 1997 Academic Press

INTRODUCTION

The Rauscher virus complex contains the replication-competent helper virus, termed Rauscher murine leukemia virus (R-MuLV), and the replication-defective Rauscher spleen focus-forming virus (R-SFFV); spleens of infected mice also contain a third component, Rauscher mink cell focus-inducing (R-MCF) virus (Pluznik and Sachs, 1964; Weiss *et al.,* 1985). This virus complex induces a clinical disease spectrum similar to that of the Friend virus complex, which is also composed of a replication-competent helper virus, the Friend murine leukemia virus (F-MuLV), and the defective anemia strain of the Friend spleen focus-forming viruses (Weiss *et al.,* 1985). Both the Rauscher and Friend virus complexes cause erythroleukemia in newborn and adult mice. Earlier studies from this laboratory have shown that live, attenuated Rauscher virus complex generated immune responses that protected against high-dose challenge with live virus (Ruprecht *et al.,* 1990a, 1990b). By adoptive transfer, we determined the correlates of immune protection in this system; immune T cells alone were able to confer protection (Ruprecht *et al.,* 1990a; Hom *et al.,* 1991). Presently, we are seeking to determine the epitopes that are recognized by cytotoxic T cells. To generate the necessary reagents, the complete nucleotide sequence of R-MuLV was needed.

The similar pathogenicity of the Friend and Rauscher virus complexes suggested a strong degree of homology at the primary nucleotide sequence level. The components of the Friend virus complex have been sequenced (Obata *et al.,* 1984; Remington *et al.,* 1992; Friedrich *et al.,* 1991; Perryman *et al.,* 1991a), whereas only the long terminal repeat (LTR) sequences of R-MuLV clone-9 (Van Der Feltz *et al.,* 1986), R-SFFV (Bestwick *et al.,* 1984), R-MCF (Vogt *et al.,* 1985), and sequences of the envelope (*env*) gene of R-SFFV (Bestwick *et al.,* 1984) and R-MCF virus (Vogt *et al.,* 1985) have been reported. The full-length genomes of other ecotropic, replication-competent helper murine leukemia viruses that have been characterized fully are: Moloney murine leukemia virus (Mo-MuLV) (Shinnick *et al.,* 1981), Akv (Herr, 1984; Etzerodt *et al.,* 1984), radiation leukemia virus (RadLV) (Merregaert *et al.,* 1987), and Cas-Br-E (Perryman *et al.,* 1991b). Here, we report the complete nucleotide sequence of R-MuLV, a comparison of this sequence with those of other ecotropic replication-competent helper MuLVs, and their phylogenetic relationships. As expected from the similar biological characteristics, we have found a high degree of sequence identity between R-MuLV and F-MuLV, and both viruses appeared on the same branch by phylogenetic analysis of the *gag, pol,* and *env* sequences.

MATERIALS AND METHODS

The integrated proviral DNA of R-MuLV clone RV-1, generated by Habara *et al.* (1982) using a bacteriophage

[1] The nucleotide sequence data reported in this paper have been submitted to the GenBank nucleotide sequence database and have been assigned Accession No. U94692.

[2] To whom correspondence and reprint requests should be addressed. Fax: (617) 632-3112. E-mail: ruth_ruprecht@dfci.harvard.edu.

0042-6822/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

TABLE 1

Sequence Homology between the R-MuLV and Other Murine Leukemia Virus Genomes

Strain	U3	R	U5	5′ Leader[a]	*gag*		*pol*		*env*		Complete genome
					nt[b]	a.a.[c]	nt	a.a.	nt	a.a.	
F-MuLV											
PVC-211[d]	97.3	98.5	100	96.4	98.0	97.4	98.0	98.5	96.9	95.4	97.6
57[e]	97.3	92.6	100	96.8	97.8	96.8	97.5	98.0	96.9	96.3	97.3
FB29[f]	97.6	98.5	98.7	96.0	97.6	96.7	98.1	98.3	96.9	95.7	97.6
Mo-MuLV[g]	86.3	94.1	97.4	91.6	84.8	90.7	90.9	96.1	82.1	85.9	87.6
Cas-Br-E[h]	89.0	97.1	97.4	90.3	84.8	92.5	91.7	96.6	72.1	78.1	85.7
Akv[i]	74.8	95.6	89.5	74.2	79.2	85.3	84.7	89.8	75.9	80.7	80.8
RadLV[i]	69.4	94.1	90.8	73.9	79.5	84.5	84.7	90.2	75.5	80.6	80.5
Ho-MuLV[k]	79.8	92.6	90.8	80.2	79.8	85.3	ND[l]	ND	69.3	71.6	ND

[a] 5′ leader sequence from the 3′ end of U5 to the start codon of *gag.*
[b] nt, percentage nucleotide identity.
[c] a.a, percentage amino acid identity.
[d] Accession No. M93134 (Remington *et al.,* 1992).
[e] Accession No. X02794 (Friedrich *et al.,* 1991).
[f] Accession No. Z11128 (Perryman *et al.,* 1991a).
[g] Accession No. J02255 (Shinnick *et al.,* 1981).
[h] Accession No. X57540 (Perryman *et al.,* 1991b).
[i] Accession No. J01998 (Herr, 1984; Etzerodt *et al.,* 1984).
[j] Accession No. K03363 (Merregaert *et al.,* 1987).
[k] Accession No. M26527, M26528 (Voytek and Kozak, 1989).
[l] Not determined, sequence not available.

Charon 4A vector and further subcloned into the *Eco*RI site of the plasmid pBR322, was kindly provided by Dr. S. Aaronson. The resulting plasmid, p3028N, containing the R-MuLV provirus with flanking mouse genomic DNA, was used for the DNA sequence analysis. The primers used for sequencing the R-MuLV genome were designed based on the reported F-MuLV sequence (Remington *et al.,* 1992). DNA sequence analysis was performed on double-stranded p3028N DNA templates by the Sanger method (Sanger *et al.,* 1977), using dye-labeled dideoxy nucleotides as terminators. Samples were analyzed on an automated DNA sequencer (Applied Biosystems Model 373A automated DNA sequencer) (Smith *et al.,* 1986). Both strands of the entire genome were sequenced by using primers in both directions.

The criteria for selecting murine leukemia viruses for sequence alignments and phylogenetic analyses were: (1) ecotropic, replication-competent helper viruses; and (2) at least fully sequenced *gag* and *env* genes. Sequence alignments were performed by using the GAP alignment program in GCG (Genetics Computer Group, Madison, WI). Phylogenetic trees were constructed by using the PIMA (Pattern-Induced Multi-Sequence Alignment) algorithm (Smith and Smith, 1992) and PAUP (Phylogenetic Analysis Using Parsimony) software (Swofford, 1985). Based on the amino acid sequence alignments of the Gag, Pol, and Env regions and the third codon position variation in these genes of R-MuLV and other murine leukemia viruses, phylogenetic trees were constructed. The use of the third codon position eliminates any selective bias. The protein sequences were aligned initially with PIMA. Subsequently, the protein alignments with the corresponding nucleotide sequences were entered into the PAUP program to generate the phylogenetic trees. Bootstrap analysis using the PHYLIP software package (Felsenstein, 1989, 1990) was also performed to generate phylogenetic trees based on the third codon position and the complete sequence of the *gag, pol,* and *env* genes of all MuLVs.

RESULTS AND DISCUSSION

The Genbank accession number of the complete nucleotide sequence is U94692. The entire R-MuLV genome is 8282 bases in length, which is the same length as that of the F-MuLV strain PVC-211 (Remington *et al.,* 1992). This full-length helper R-MuLV sequence was aligned with the sequences of other murine leukemia viruses at the nucleotide and amino acid levels (Table 1). The nucleotide sequence alignment showed an overall identity between R-MuLV and F-MuLV of 97.6% and of 87.6% with the T-cell tropic Moloney murine leukemia virus (Mo-MuLV), respectively. R-MuLV is closely related to F-MuLV in each region of the genome, whereas the greatest sequence divergence between R-MuLV and other MuLVs was seen in the U3 and *env* regions. GAP

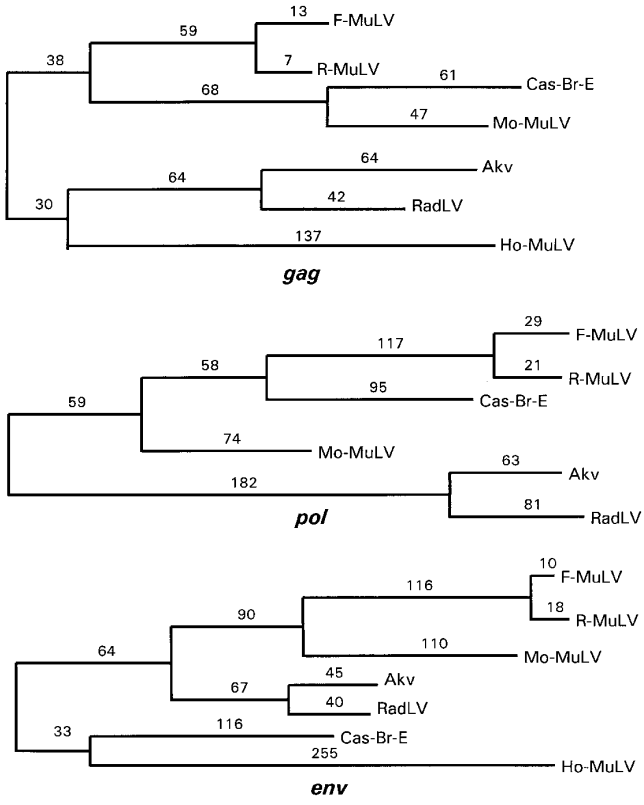

FIG. 1. Phylogenetic relationship of murine retroviruses based on *gag, pol,* and *env* sequences. The sequences for each virus were obtained from the Genbank by the accession numbers listed in Table 1. The phylogenetic analysis was carried out by initial amino acid sequence alignment of the Gag, Pol, and Env sequences followed by imposition on the nucleotide sequence of the respective *gag, pol,* and *env* genes using the PIMA algorithm (Smith and Smith, 1992). The three trees were constructed using the PAUP software (Swofford, 1985) based on the PIMA alignments. The numbers on the tree branches represent the number of implied third codon position substitutions.

analysis of the R-MuLV *env* region with the *env* sequence of R-SFFV showed 55% homology in the 5′ half and 98% homology in the 3′ half, which is consistent with the known origin of the R-SFFV *env* from MCF-like sequences at the 5′ end and truncated ecotropic *env* sequences at the 3′ end (Bestwick *et al.,* 1984). R-MuLV *env* and R-MCF (Vogt *et al.,* 1985) *env* sequences showed an overall identity of 78%; again, the 3′ region was highly homologous (99.2%) to ecotropic sequences (not shown). Furthermore, the overall sequence identity between the LTR region of the R-MuLV sequence reported here and the published LTR sequences of R-MuLV clone-9 (Van Der Feltz *et al.,* 1986), R-SFFV (Bestwick *et al.,* 1984), and R-MCF (Vogt *et al.,* 1985) was >98%.

The three phylogenetic trees shown were constructed by the PAUP program (Swofford, 1985) based on codon position 3 variation in the *gag, pol,* and *env* regions of all the murine leukemia viruses (Fig. 1). The Doolittle amino acid trees (Feng and Doolittle, 1987) were the same as those obtained from the third codon position analysis (data not shown). Phylogenetic trees, which were generated by Bootstrap analysis (Felsenstein, 1989, 1990) from the third codon position and the complete nucleotide sequence of the *gag, pol,* and *env* genes also showed nearly identical branching of each MuLV as that obtained by the other methods (data not shown).

R-MuLV and F-MuLV fall into one very closely related subgroup in all trees. The *gag* genes of these two viruses appear to have evolved more slowly than those of all other MuLVs based on the shorter branch lengths on the *gag* tree. This may be due to some selective constraint on the *gag* gene of R-MuLV and F-MuLV. Mo-MuLV, a virus that causes T-cell leukemias/lymphomas in neonatal mice, is the next closest relative to R-MuLV and F-MuLV on all trees, followed by Akv and RadLV. The latter two viruses form a separate subgroup and cluster together on all three trees. The two viruses are similar biologically; both are present as integrated proviruses in the germ line of their inbred host mouse strains, and both cause thymic leukemias (Weiss *et al.,* 1985). The most distantly related helper virus is Hortulanus murine leukemia virus (Ho-MuLV), a virus isolated from wild European *Mus hortulanus* that causes various hematological malignancies (Voytek and Kozak, 1988, 1989). Another wild mouse isolate, the neurotropic Cas-Br-E (Gardner *et al.,* 1976), branches with Mo-MuLV on the *gag* and *pol* trees. In contrast, Cas-Br-E exhibits the highest degree of relatedness on the *env* tree to Ho-MuLV. The different topology of Cas-Br-E on the *env* tree relative to its *gag* and *pol* positions is also supported by the consensus tree generated by the Bootstrap analysis of the third codon position and the complete sequence of the three genes (data not shown). The unexpectedly close phylogenetic relationship between Cas-Br-E and Ho-MuLV *env* sequences may be explained by a recombination event that may have occurred with a putative ancestor of Ho-MuLV.

In summary, we have performed an extensive phylogenetic analysis of ecotropic, replication-competent MuLVs. This analysis confirms the postulated close relationship of the complete R-MuLV nucleotide sequence with the F-MuLV genome.

ACKNOWLEDGMENTS

We thank Christine Bogle and Noele Nelson (Molecular Biology Core Facility, Dana-Farber Cancer Institute) for assistance with the DNA sequencing, Alice Ng for technical assistance, and Anne DiSorbo for the preparation of this manuscript. Supported in part by U.S. Army Grant DAMD17-94-J-4431 and the Center for AIDS Research (CFAR) Core Grant IP30 28691-01 awarded to the Dana-Farber Cancer Institute. This work was also supported by Grant P41 LM05205-14 from the National Library of Medicine. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the granting agency.

REFERENCES

Bestwick, R. K., Boswell, B. A., and Kabat, D. J. (1984). Molecular cloning of biologically active Rauscher spleen focus-forming virus and the sequences of its *env* gene and long terminal repeat. *J. Virol.* 51, 695–705.

Etzerodt, M., Mikkelsen, T., Pedersen, F. S., Kjeldgaard, N. O., and Jorgensen, P. (1984). The nucleotide sequence of the Akv murine leukemia virus genome. *Virology* 134, 196–207.

Felsenstein, J. (1989). PHYLIP-Phylogeny Inference Package (Version 3.2). *Cladistics* 5, 164–166.

Felsenstein, J. (1990). PHYLIP (Phylogeny Inference Package) version 3.5c. Department of Genetics, University of Washington, Seattle.

Feng, D. F., and Doolittle, R. F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25, 351–360.

Friedrich, R. W., Koch, W., Von Maydell-Livonius, U., Schrewe, H., and Zimmerman, W. (1991). European Molecular Biology Laboratory (EMBL) Data Library, Accession No. X02794.

Gardner, M. B., Henderson, B. E., Estes, J. D., Rongey, R. W., Casagrande, J., Pike, M., and Huebner, R. J. (1976). The epidemiology and virology of C-type virus-associated hematological cancers and related diseases in wild mice. *Cancer Res.* 36, 574–581.

Habara, A., Reddy, E. P., and Aaronson, S. A. (1982). Rauscher murine leukemia virus: Molecular cloning of infectious integrated proviral DNA. *J. Virol.* 44, 731–735.

Herr, W. (1984). Nucleotide sequence of Akv murine leukemia virus. *J. Virol.* 49, 471–478.

Hom, R. C., Finberg, R. W., Mullaney, S., and Ruprecht, R. M. (1991). Protective cellular retroviral immunity requires both $CD4^+$ and $CD8^+$ immune T cells. *J. Virol.* 65, 220–224.

Merregaert, J., Janowski, M., and Reddy, E. P. (1987). Nucleotide sequence of a radiation leukemia virus genome. *Virology* 158, 88–102.

Obata, M., Amanuma, H., Harada, Y., Sagata, N., and Ikawa, Y. (1984). Env-related leukemogenic genes (gp55 genes) of two closely related polycythemic strains of Friend spleen focus-forming virus possess different recombination points with an endogenous mink cell focus-forming virus *env* gene. *Virology* 136, 435–438.

Perryman, S., Nishio, J., and Chesebro, B. (1991a). Complete nucleotide sequence of Friend murine leukemia virus, strain FB29. *Nucleic Acids Res.* 19, 6950.

Perryman, S. M., McAtee, F. J., and Portis, J. L. (1991b). Complete nucleotide sequence of the neurotropic murine retrovirus Cas-Br-E. *Nucleic Acids Res.* 19, 1707.

Pluznik, D. H., and Sachs, L. (1964). Quantitation of a murine leukemia virus with a spleen colony assay. *J. Natl. Cancer Inst.* 33, 535–546.

Remington, M. P., Hoffman, P. M., Ruscetti, S. K., and Masuda, M. (1992). Complete nucleotide sequence of a neuropathogenic variant of Friend murine leukemia virus PVC-211. *Nucleic Acids Res.* 20, 3249.

Ruprecht, R. M., Mullaney, S., Bernard, L. D., Gama Sosa, M. A., Hom, R. C., and Finberg, R. W. (1990a). Vaccination with a live retrovirus: The nature of the protective immune response. *Proc. Natl. Acad. Sci. USA* 87, 5558–5562.

Ruprecht, R. M., Chou, T. C., Chipty, F., Gama Sosa, M., Mullaney, S., O'Brien, L., and Rosas, D. (1990b). Interferon-α and 3′-Azido-3′-deoxythymidine are highly synergistic in mice and prevent viremia after acute retrovirus exposure. *J. AIDS* 3, 591–600.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.

Shinnick, T. M., Lerner, R. A., and Sutcliffe, J. G. (1981). Nucleotide sequence of Moloney murine leukemia virus. *Nature* 293, 543–548.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H., and Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature* 321, 674–679.

Smith, R. F., and Smith, T. F. (1992). Pattern-induced multi-sequence alignment (PIMA) algorithm employing secondary structure-dependent gap penalties for use in comparative protein modeling. *Protein Eng.* 5, 35–41.

Swofford, D. L. (1985). PAUP (phylogenetic analysis using parsimony) user's manual. Illinois Natural History Survey, Champaign, IL.

Van Der Feltz, M. J. M., Kranendonk-Odijk, M. E., Stark, J., and De Both, N. J. (1986). Nucleotide sequence of the Rauscher murine leukemia virus long terminal repeat. *J. Gen. Virol.* 67, 2785–2790.

Vogt, M., Haggblom, C., Swift, S., and Haas, M. (1985). Envelope gene and long terminal repeat determine the different biological properties of Rauscher, Friend, and Moloney mink cell focus-inducing viruses. *J. Virol.* 55, 184–192.

Voytek, P., and Kozak, C. A. (1988). Ho-MuLV: A novel pathogenic ecotropic virus isolated from the European mouse, *Mus hortulanus. Virology* 165, 469–475.

Voytek, P., and Kozak, C. A. (1989). Nucleotide sequence and mode of transmission of the wild mouse ecotropic virus, Ho-MuLV. *Virology* 173, 58–67.

Weiss, R., Teich, N., Varmus, H., and Coffin, J. (Eds.) (1985). "RNA Tumor Viruses," 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.