Available online at www.sciencedirect.com
SCIENCE DIRECT
VIROLOGY
Virology 350 (2006) 251–257
www.elsevier.com/locate/yviro

Minireview

Phylogeny and evolution of old world arenaviruses

Sébastien Emonet [a], Jean-Jacques Lemasson [a], Jean-Paul Gonzalez [b], Xavier de Lamballerie [a], Rémi N. Charrel [a,*]

[a] *Unité des Virus Emergents (EA3292, IFR48, IRD UR178), Université de la Méditerranée, Marseille, France*
[b] *Conditions et Territoires d'Emergence des Maladies (UR178), Institut de Recherche pour le Développement, France*

Received 20 December 2005; returned to author for revision 11 January 2006; accepted 20 January 2006
Available online 21 February 2006

Abstract

The intention of this study was to investigate the genomics, phylogeny and evolution of the Old World arenaviruses based on sequence data representing the four viral genes. To achieve this aim, we sequenced the complete S and L RNA segments of *Ippy virus* (IPPYV), *Mobala virus* (MOBV) and *Mopeia virus* (MOPV). Full-length sequences of the NP, GPC, Z and L genes were used to reconstruct phylogenetic relationships and to compare resulting tree topologies. Each of the five Old World arenavirus species (namely *Lassa virus* [LASV], IPPYV, MOBV, MOPV and *Lymphocytic choriomeningitis virus* [LCMV]) are monophyletic; seven selected strains of LASV showed a similar topology regardless of the gene under analysis; IPPYV rooted the three other African arenaviruses; the four African arenaviruses are rooted by the ubiquitous LCMV; and the tree topologies of the three African arenaviruses other than LASV are identical regardless of the gene used for analysis. No evidence for significant evolutionary events such as intra- or intersegmental recombination was obtained.

Keywords: Arenavirus; Evolution; Recombination; Phylogeny; Viral hemorrhagic fever; *Lassa virus*; BSL-4; Africa; *Ippy virus*; *Mobala virus*; *Mopeia virus*

Introduction

The family Arenaviridae comprises a unique genus (*Arenavirus*) that currently contains 22 recognized viruses. Arenaviruses possess single stranded bi-segmented RNA genomes. The large (L) genomic segment (~7,200 nt) encodes the viral RNA-dependent RNA polymerase and a zinc-binding matrix protein, acting as a bona fide matrix protein (Perez et al., 2003; Strecker et al., 2003) The small (S) genomic segment (~3500 nt) encodes the nucleocapsid protein (NP) and the glycoprotein precursor (GPC) in two non-overlapping reading frames of opposite polarities. The GPC is secondarily cleaved into the envelope proteins G1 and G2. The genes on both S and L segments are separated by an intergenic non-coding region with the potential to form one or more hairpin configurations. The 5′ and 3′ ends of each RNA segment possess a relatively conserved reverse complementary sequence spanning 19 nucleotides at each extremity. The arenaviruses have been classified according to their antigenic properties into two groups: the Tacaribe serocomplex (including viruses indigenous to the New World) and the Lassa–Lymphocytic choriomeningitis serocomplex (including the viruses indigenous to Africa and the ubiquitous *Lymphocytic choriomeningitis virus* [LCMV], recognized as the Old World group) (Charrel and de Lamballerie, 2002; Clegg et al., 2000). Specific rodents are the principal hosts of the arenaviruses. Humans usually become infected through contact with infected rodents or inhalation of infectious rodent excreta or secreta. At least 9 arenaviruses are associated with human disease, of which five – *Lassa* (LASV), *Junin*, *Machupo*, *Guanarito* and *Sabia* – are known to cause severe hemorrhagic fever in western Africa, Argentina, Bolivia, Venezuela and Brazil, respectively. The other four arenaviruses are LCMV (causing acute central nervous system disease (Barton, 1996) and congenital malformations (Barton and Mets, 2001)), Flexal and Tacaribe viruses (febrile illnesses in laboratory workers (Peters et al., 1996; Buchmeier et al., 1974)) and, more recently, Whitewater Arroyo virus associated with fatal cases of infection in California (CDC, 2000). The five arenaviruses causing viral hemorrhagic fever are included in the

* Corresponding author. Fax: +33 491 324495.
E-mail address: Remi.Charrel@medecine.univ-mrs.fr (R.N. Charrel).

0042-6822/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2006.01.026

Category A Pathogen List, considered as Select Agents as defined by the CDC, and listed as Biosafety Level 4 agents. In the past 5 years, our understanding of the phylogeny and evolution of arenaviruses has made considerable progress due to the availability of complete genomic sequences; evidence for recombination has resulted from the analysis of full-length genomes and showed that significant evolution events may play a substantial part in arenavirus evolution (Archer and Rico-Hesse, 2002; Charrel et al., 2001, 2002). To date, most studies have focused on New World arenaviruses and neglected Old World arenaviruses. Our objective was to study the evolution of these viruses with a strategy similar to that previously used for New World arenaviruses (Charrel et al., 2003) to explore the respective role of the 3 mechanisms driving the genetic diversity of Arenaviridae (specifically, the accumulation of mutations, intrasegmental recombination and intersegmental recombination or reassortment). To achieve this aim, we first determined the full-length genome sequences for *Ippy* (IPPYV), *Mobala* (MOBV) and *Mopeia* (MOPV) viruses in order to complete sequence data sets for African arenaviruses; we then used a comprehensive set of sequences for phylogeny reconstruction and compared tree topologies obtained independently from the four genes. Finally, different methods were used to search for significant evolutionary events such as recombination and reassortment.

Results and discussion

To date, little data were available for the Old World arenaviruses, and the most comprehensive phylogenetic study was based on analysis of a partial region of the nucleoprotein gene (Bowen et al., 1997, 2000). Therefore, it was essential to obtain a complete genetic data set in order to examine the phylogenetic relationships of arenaviruses within the Old World complex. Here, we report the first comprehensive phylogenetic study based on independent analysis of full-length sequences of

Table 1
Characteristics of complete sequences of Old World Arenaviruses S (A) and L RNA (B)

(A)

Virus species	Acronym	Strain	S RNA						
			GenBank accession nos.	5′ NCR (nt)	GPC (aa)	IGR (nt)	NP (aa)	3′ NCR (nt)	Length (nt)	GC (%)
Lassa virus	LASV	Josiah	NC_004296	55	491	67	569	100	3402	44.15
		Josiah	AY628203	55	491	67	569	100	3402	44.12
		NL	AY179173	nc	491	67	569	nc	nc	44.62
		z148	AY628205	54	491	66	569	96	3396	43.96
		Macenta	AY628201	54	491	66	569	99	3399	43.98
		AV	AF246121	53	491	67	569	91	3391	44.35
		CSF	AF333969	nc	490	65	569	nc	nc	44.05
Lymphocytic choriomeningitis virus	LCMV	WE	M22138	77	498	70	558	60	3375	44.77
		Armstrong	NC_004294	77	498	70	558	61	3376	46.12
Mopeia virus	MOPV	AN20410	AY772170	53	489	129	570	68	3427	45.55
		Mozambique	**DQ328874**	53	489	86	570	68	3384	45.39
Mobala virus	MOBV	Acar 3080	**AY342390**	84	491	69	568	50	3380	43.82
Ippy virus	IPPYV	Dak An B 188 d	**DQ328877**	nc	495	64	570	58	nc	45.07

(B)

Virus species	Acronym	Strain	L RNA						
			GenBank accession nos.	5′ NCR (nt)	Z (aa)	IGR (nt)	L (aa)	3′ NCR (nt)	Length (nt)	GC %
Lassa virus	LASV	Josiah	NC_004297	65	99	106	2218	157	7279	41.12
		Josiah	AY628202	65	99	106	2220	157	7285	40.99
		NL	AY179172	nc	99	105	2220	139	nc	41.3
		z148	AY628204	66	99	104	2219	156	7280	42.2
		Macenta	AY628200	66	99	104	2219	156	7280	42.2
		AV	AY179171	nc	99	105	2220	nc	nc	40.27
		CSF	L: AY179174	nc	99	nc	2217	nc	nc	L: 40.31
			Z: AY179175							Z: 50.51
Lymphocytic choriomeningitis virus	LCMV	WE	AF004519	88	90	202	2209	32	7219	42.06
		Armstrong	L: J04331	88	90	200	2210	32	7220	L: 40.36
			Z: M27693							Z: 53.06
Mopeia virus	MOPV	AN20410	AY772169	79	103	116	2237	56	7271	42.5
		Mozambique	**DQ328875**	80	103	111	2229	55	7242	40.87
Mobala virus	MOBV	Acar 3080	**DQ328876**	116	99	102	2220	150	7325	39.59
Ippy virus	IPPYV	Dak An B 188 d	**DQ328878**	148	101	170	2208	71	7316	39.76

Accession numbers in bold refer to sequences determined in this study.