

Watching Every Step of the Way: Junín Virus Attenuation Markers in the Vaccine Lineage

Betina Inés Stephan*, Mario Enrique Lozano and Sandra Elizabeth Goñi

Área de Virosis Emergentes y Zoonóticas, Laboratorio de Biología Celular y Molecular, Universidad Nacional de Quilmes, Argentina

Abstract: The Arenaviridae family includes several hemorrhagic fever viruses which are important emerging pathogens. Junín virus, a member of this family, is the etiological agent of Argentine Hemorrhagic Fever (AHF). A collaboration between the Governments of Argentina and the USA rendered the attenuated Junín virus vaccine strain Candid #1. Arenaviruses are enveloped viruses with genomes consisting of two single-stranded RNA species (L and S), each carrying two coding regions separated by a stably structured, non-coding intergenic region. Molecular characterization of the vaccine strain and of its more virulent ancestors, XJ13 (prototype) and XJ#44, allows a systematic approach for the discovery of key elements in virulence attenuation. We show comparisons of sequence information for the S RNA of the strains XJ13, XJ#44 and Candid #1 of Junín virus, along with other strains from the vaccine lineage and a set of Junín virus field strains collected at the AHF endemic area. Comparisons of nucleotide and amino acid sequences revealed different point mutations which might be linked to the attenuated phenotype. The majority of changes are consistent with a progressive attenuation of virulence between XJ13, XJ#44 and Candid #1. We propose that changes found in genomic regions with low natural variation frequencies are more likely to be associated with the virulence attenuation process. We partially sequenced field strains to analyze the genomic variability naturally occurring for Junín virus. This information, together with the sequence analysis of strains with intermediate virulence, will serve as a starting point to study the molecular bases for viral attenuation.

Received on: July 08, 2013- Revised on: August 26, 2013- Accepted on: August 26, 2013

Keywords: *Arenaviridae*, Junín virus, natural variability, point mutations, vaccine lineage, viral attenuation.

INTRODUCTION

The *Arenaviridae* family comprises only one genus formed of 25 species of arenaviruses, as categorized by the ICTV [1], and listed in (Table 1). A further classification is based on the geography of origin and distribution of the viruses and their hosts: Old World Arenaviruses (OWA) and New World Arenaviruses (NWA). The NWA group harbors all arenaviruses from the American continent and the OWA group includes the only arenaviral species present in all continents (Lymphocytic Choriomeningitis Virus – LCMV) as well as the African arenaviruses. To date, there are no reports of any native viral species from Europe or Australia. The precise geographic distribution pattern has its origin in a very narrow host range that is characteristic of all arenaviruses. The principal hosts for the NWA are several specific members of the rodent family *Cricetidae*, subfamilies *Sigmodontinae* and *Neotominae* [2]. In the case of LCMV, the rodent reservoir is the worldwide distributed *Mus musculus*, which explains the unusual ubiquity of this virus. One NWA, the Tacaribe Virus (TACV), was isolated from bats [3], in what was thought to be an atypical occurrence for an arenavirus. However, a recent discovery of new arenaviral

species in boas [4] opens up new avenues of inquiry regarding the viral host range and biodiversity.

This viral family is of medical interest because 6 of the 25 members have been consistently associated with clinical symptoms and described as the etiological agents of human diseases. All pathogenic NWA produce hemorrhagic fevers, with very similar physiopathology (Table 2). The phases of the illness are well characterized and progress from prodromal to neurologic-hemorrhagic to convalescent [5]. There is a short incubation period before the onset of symptoms. In the case of Argentine Hemorrhagic Fever (AHF), the syndrome is characterized by renal, neurological, vascular, immunological and hematological alterations. If it remains untreated, the mortality rate can be as high as 15-30%.

Since it was described, AHF has been reported to cause annual outbreaks during the autumn-winter season [6]. However, the number of actual cases in each outbreak has significantly diminished since the application of the vaccine to the risk population within the endemic area. For the 1992-2000 period the efficacy of the vaccine was estimated at 98% [7].

The development of the vaccine against AHF was the product of a collaboration between the US and Argentine governments. A greatly attenuated strain of Junín virus (JUNV), Candid#1, was extensively tested in rhesus monkeys and human volunteers with good results. Afterwards, a comprehensive clinical trial was conducted in the AHF

*Address correspondence to this author at the Área de Virosis Emergentes y Zoonóticas, Laboratorio de Biología Celular y Molecular, Universidad Nacional de Quilmes, Argentina; Tel: 00541143657100; Ext: 5651; Fax: 00541143657132; E-mail: igastephan@gmail.com

Table 1. Arenaviridae family. NR: not reported; OWA: Old World Arenavirus; NWA: New World Arenavirus; the lineages are listed as: A, B, C and Rec A/B. The distribution is based on the country of virus isolation. Table modified from [6, 7]

Virus	Acronym	Evolutionary Lineage	Distribution	Reservoir	Human Patogenicity
Flexal	FLEV	NWA-A	Brazil	<i>Oryzomys spp.</i>	Yes
Pichinde	PICV	NWA-A	Colombia	<i>O. albigularis</i>	NR
Paraná	PARV	NWA-A	Paraguay	<i>O. buccinatus</i>	NR
Allpahuayo	ALLV	NWA-A	Peru	<i>Oecomys bicolor</i>	NR
Piritál	PIRV	NWA-A	Venezuela	<i>Sigmodon alstoni</i>	NR
Junín	JUNV	NWA-B	Argentina	<i>C. musculinus</i>	Yes
Machupo	MACV	NWA-B	Bolivia	<i>C. callosus, C. laucha</i>	Yes
Guanarito	GTOV	NWA-B	Venezuela	<i>Z. brevicauda</i>	Yes
Sabiá	SABV	NWA-B	Brazil	Unknown	Yes
Chapare	-	NWA-B	Bolivia	Unknown	Yes
Latino	LATV	NWA-C	Bolivia	<i>Calomys callosus</i>	NR
Tacaribe	TCRV	NWA-B	Trinidad	<i>Artibeus spp. (bat)</i>	Yes
Cupixi	CPXV	NWA-B	Brazil	<i>O. capita</i>	NR
Amapari	AMAV	NWA-B	Brazil	<i>O. capita-Neacomys guianae</i>	NR
Oliveros	OLVV	NWA-C	Argentina	<i>Bolomys obscurus</i>	NR
Whitewater Arroyo	WWAV	NWA-Rec A/B	USA Southwest (NM,TX,UT,OK)	<i>Neotoma albigula, N. mexicana, N. micropus, N. sinerea</i>	Yes
Lassa	LASV	OWA	Nigeria, Ivory Coast, Guinea, Sierra Leone	<i>Mastomys sp.</i>	Yes
Tamiami	TAMV	NWA-Rec A/B	Florida, USA	<i>Sigmodon hispidus</i>	NR
Bear Canyon	BCNV	NWA-Rec A/B	California, USA	<i>Peromyscus sp.</i>	NR
Lymphocytic Choriomeningitis	LCMV	OWA	Worldwide	<i>Mus musculus</i>	Yes
Mobala	MOBV	OWA	Central African Republic	<i>Praomys sp.</i>	NR
Mopeia	MOPV	OWA	Mozambique	<i>Mastomys natalensis</i>	NR
Ippy	IPPYV	OWA	Central African Republic	<i>Arvicanthis sp.</i>	NR
Lujo	LUJV	OWA	South Africa	Unknown	Yes
Luna	LUNV	OWA	Zambia	<i>Mastomys natalensis</i>	NR

Table 2. Arenaviral Hemorrhagic Fevers Pathological Characteristic. The Name of the Disease (When Existing), Along with the Principal Clinical Symptoms Caused by Old World Arenaviruses and New World Arenaviruses are Briefly Explained. Modified from [13]

Agent (Disease)	Pathological Characteristics
Old World Lassa (Lassa Fever)	<p>Incubation: 3 to 21 days. Fever, headache, myalgia, backache, trembling and sickness.</p> <p>Generalized infection: hemorrhagic dissemination of virus to several organs and systems via bloodstream, lymphatic system, respiratory and digestive tract.</p> <p>Black vomit, aqueous diarrhea (dehydration), quantity decrease of lymphocytes and platelets, mild thrombocytopenia, abdominal, pleuritic and hepatic pain.</p> <p>Extensive reticulo-endothelial compromise: capillary injuries causing stomach, small intestine, kidney, lung and brain bleeding.</p> <p>Multifocal hepatocellular necrosis with Councilman-like bodies, hepatocytes cytoplasmic degeneration and minimal inflammatory response.</p> <p>Adrenal focal necrosis and cytoplasmic inclusions.</p> <p>Respiratory system: interstitial pneumonia, cough, dyspnea, bronchitis, pneumonia and pleurisy.</p> <p>Cardiovascular system: pericarditis, tachycardia, bradycardia, hypertension, hypotension, thrombocytopenia, leukopenia and hiperuricemia, lymphadenopathy, elevated aminotransferases, decreased prothrombin level, disorder of blood circulation and bleeding through the skin, lungs, gastrointestinal tract and other mucosa membranes.</p> <p>Nervous system: encephalitis, meningitis, uni- or bilateral hearing decrease, convulsions.</p>

(Table 2) contd....

Agent (Disease)	Pathological Characteristics
New World Junín (Argentine Hemorrhagic Fever)	Incubation: 6 to 14 days. First 4 days: decaying, fever, anorexia, nausea and vomiting, headache and myalgia. Second stage: acute hemorrhagic syndrome (epistaxis and hematemesis), or acute neurologic syndrome. General: malaise, high fever, severe myalgia, anorexia, back pain, abdominal tenderness, conjunctivitis, retro-orbital pain, photophobia, and usually constipation.
Machupo and Chapare (Bolivian Hemorrhagic Fever)	Acute phase of infection: peripheral blood viral active replication. Oropharyngeal enanthem.
Guanarito (Venezuelan Hemorrhagic Fever)	Gums swollen, congested and bleeding (gingival border). Proteinuria high dehydration and hemoconcentration. In women, early menorrhagia.
Sabiá (Brazilian Hemorrhagic Fever)	Multifocal hepatocellular necrosis with formation of Councilman-like bodies, nuclear pyknosis, cytoplasmic eosinophilia, cytolysis, inflammation and a mild cellular infiltration composed of mononuclear cells and neutrophils.
Whitewater Arroyo Flexal Lujo	Kidney damage: distal tubular cells and collecting ducts. Glomeruli or proximal tubules. In a few cases, renal failure was reported. Cardiovascular system: postural hypotension and relative bradycardia, arrhythmias are transient and benign. Different degrees of dehydration, uremia, proteinuria, hematuria and oliguria. Respiratory: dry cough, without sore throat. Pharyngeal enanthem. Interstitial or bronchial pneumonia, pulmonary edema and hemorrhage.

endemic area [8, 9]. A diagram with the passage history that led to the attenuated JUNV Candid#1 strain is shown in (Fig. 1). The intermediate strains were named XJ#, in reference to the original strain and the number of passages in mouse brain. The vaccine was originally produced in the USAMRIID (United States Army Medical Research Institute for Infectious Diseases) facilities. In 2005, the National Institute of Human Viral Disease (INEVH) located in Pergamino, Argentina, finished the certification process and took over the vaccine production for local use.

Other vaccines against Arenaviruses have not been released. Significant advances were made to develop an effective protection against Lassa fever, which is a disease endemic to the West-African region with 3-5x10⁵ cases reported annually [10]. Its apparently expanding territory, together with the imported cases reported yearly all over the world, make this virus the biggest health threat of the Arenavirus family [11]. Several different approaches were used to try to achieve protective coverage with a Lassa virus vaccine. However, the necessity of a single-dose treatment and the great diversity among the many viral strains converge to increase the difficulty of this task [12].

Arenaviruses are enveloped viruses with a bipartite, single stranded RNA genome [14, 15]. The virion contains three structural proteins (L, N and Z) while the lipid envelope carries the three-piece glycoprotein (SSP, G1 and G2). The larger genomic segment (L RNA, ~7kb) encodes the RNA dependent RNA polymerase (L protein), which is the minor component of the virion, as well as Z, a Zinc finger-protein of 11 kDa that could be the arenaviral counterpart of a matrix protein [16]. The other genomic segment (S RNA, ~3,5 kb) codes for the nucleocapsid protein N and the glycoprotein precursor (GPC), which is processed by cellular proteases into three parts. The arrangement of the two open reading frames in opposite orientations, separated by the stably structured, non-coding intergenic region, gave rise to the term “ambisense coding” [17]. The first 19 nucleotides at each end of one segment are complementary, which allows a base-pairing that forms a panhandle structure highly conserved among arenaviruses [18].

SEQUENCE ANALYSIS OF THE VACCINE-RELATED JUNV STRAINS

In order to identify and characterize changes directly related with the attenuated phenotype of the JUNV vaccine strain, the complete nucleotide sequence of all available intermediate strains (XJ13, XJ17, XJ34, XJ39, XJ#44, XJ48 and Candid#1) from the vaccine genealogy were analyzed, along with a small number of field strains [19, 20].

A comparison of the sequence of the vaccine genealogy strains revealed a set of differences that might be associated with the decrease in virulence (Fig. 2). An alignment of the S RNA-derived coding sequences disclosed eleven individual changes which translated into amino acid substitutions. The first of these changes was located in the stable signal peptide (XJ13 I₃₅→XJ17 I→XJ34 I→XJ39 I→XJ#44 V₃₅→XJ48 I→Candid#1 V), the following three were in the mid-part of G1 (T₁₆₈→A→A→A→A→A→A; E₁₈₆→E→E→E→E→E→G and S₂₀₆→S→S→S→S→S→P), two more were found at the carboxyl terminus of G2 (F₄₂₇→F→F→F→F→F→I and T₄₄₆→T→S→S→T→S→S), and the last five were at the amino half of N (V₄₇→V→V→V→V→V→E; R₅₉→K→K→K→R→K→R; I₁₅₈→I→I→I→V→I→V; E₂₆₈→E→E→E→E→E→D and T₃₂₂→T→T→T→I→T→I). The same analysis, performed on the coding sequences derived from the L gene, revealed only nine nucleotide changes implied in amino acid substitutions. Two of these are reversions (XJ13 R₈₈₁→XJ#44 G→Candid#1 R; S₉₂₁→G→S), while the other seven might be associated with the attenuated phenotype (H₇₆→Y→Y; V₄₁₅→V→A; D₄₆₂→N→N; L₉₃₆→L→P; R₁₁₅₆→K→K; S₁₆₉₈→S→F and I₁₈₈₃→I→V). The Z gene showed no nucleotide changes among these three JUNV strains.

This analysis was extended to include two other completely sequenced JUNV strains, Romero and MC2, as well as a partial sequence from the Cba-IV4454 strain (S RNA and Z gene). The first strain is classified as of high virulence and was originally isolated from an AHF patient [21, 22]. In the Genebank its name is mistakenly spelled “Rumero”. The second strain (MC2) is of intermediate virulence and was isolated from a rodent captured in the AHF endemic

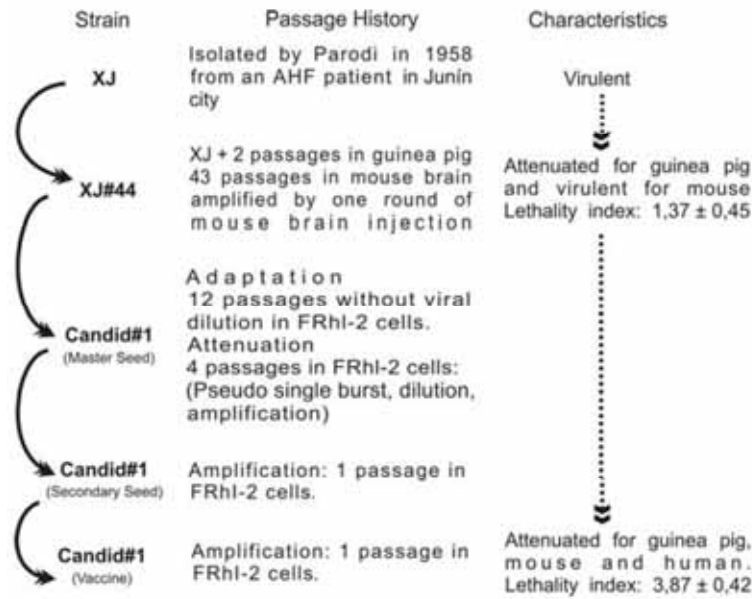


Fig. (1). History of passages for the Junin virus vaccine strain Candid#. Adapted from [11].

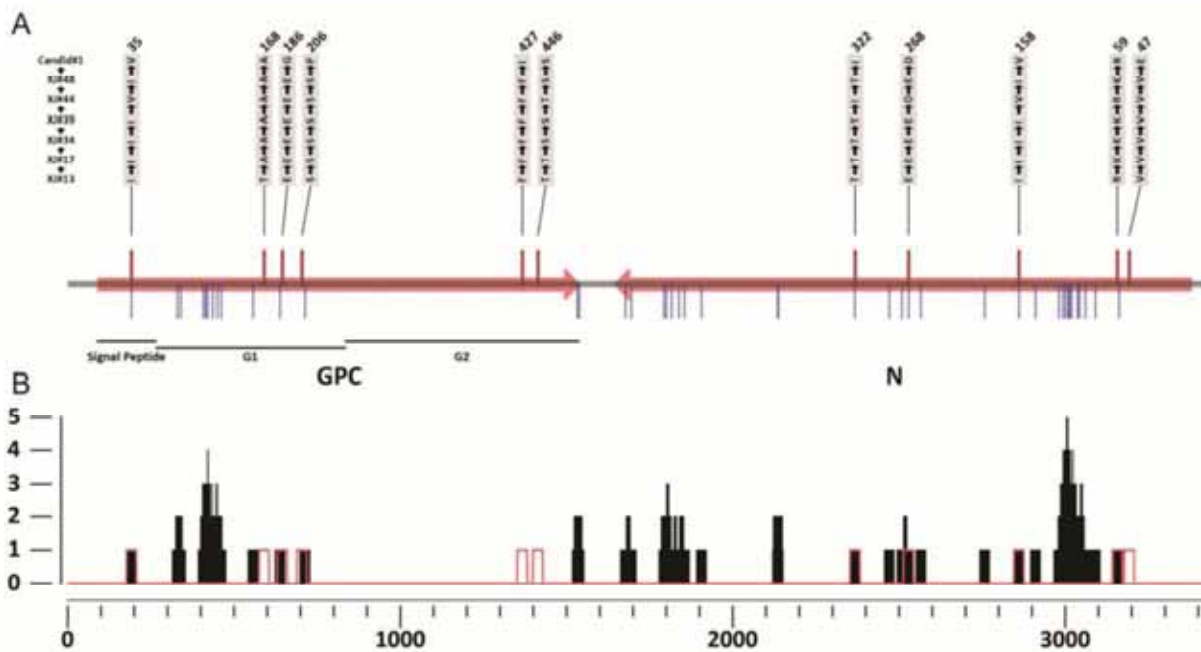


Fig. (2). Summary of the changes found in Junin virus proteins. **A.** S RNA; the N and GPC ORFs are shown as arrows, indicating the translation sense. The strains used for the comparison are shown on the left side. The parts of GPC are marked below the diagram. Amino acid changes among the strains (XJ#13, XJ#17, XJ#34, XJ#39, XJ#44, XJ#48 and Candid #1) are drawn as red lines above the gene, followed by the identity of the change and its position. Variations among the field strains (XJ#13, Romero, IV4454 and MC2) are drawn as lines below the gene. **B.** Relative frequency of mutation for the different types: black is for type-1 and red for type-2.

area [23-25]. The Cba-IV4454 strain was also obtained from a patient and is considered of intermediate virulence. The analysis was performed on deduced amino acid sequences for the six aforementioned JUNV strains. Two distinct types of mutations were identified, according to a specific definition. Type 1 mutations are those where only one of the JUNV field strains (XJ13, Romero, IV4454, and MC2) differs from all others. Type 2 mutations refer to the positions

with changes among the vaccine-related strains (XJ13, XJ#44, and Candid #1). This last type allows identifying key positions that might be involved in the attenuation process, while the Type 1 mutations take into account the naturally occurring variations. An overlap of these two types of changes uncovered several mutations with probably little significance for the viral virulence. Positions GPC₃₅, N₁₅₈, N₂₆₈ and N₃₂₂ all had changes among the six sequences that

varied between two particular amino acids, regardless of the strain's origin. Something similar happened with positions L₈₈₁ and L₉₂₁, where the only difference was found in XJ#44.

A mutation frequency index was calculated, defined as the number of mutations per amino acid in an overlapping-window strategy, with the resulting value assigned to the central residue. The analysis was made for both types of mutations for the complete JUNV genome sequence, using in-house software. This global plotting approach gives an overview of the distribution of the mutations along the genome. Several possible hot-spots were identified, that is, regions of the sequence with a marked tendency towards natural variation. Also, a small number of type 2 mutations were found outside the hot-spots, namely GPC₁₆₈, GPC₄₂₇, GPC₄₄₆ and N₄₇ for the S RNA and positions L₇₆, L₉₃₆, L₁₁₅₆ for the L RNA. These changes could more confidently be assigned to the virulence attenuation process.

CODING REGIONS

A further search for Type 2 mutations was performed on a set of field samples, isolated from human cases or rodents, captured in the endemic area during the period of 1963-91. Out of the genome, four regions were selected for this analysis: i) G fragment, comprising positions 303 to 941 within the GPC coding sequence; ii) N fragment, spanning nucleotides 1632 to 2095 of the nucleoprotein coding sequence and two regions within the open reading frame of the RNA dependent RNA polymerase [13]. The fragments were amplified by PCR and sequenced for 13 different JUNV field strains. Afterwards, the sequences were translated *in silico*, aligned and scanned for positions with differences between strains. A part of the Type 2 mutations obtained for the G fragment are shown in (Table 3). The selected region comprises most of G1, the consensus site for the arenavirus glycoprotein processing [26] and the beginning of G2. Some of the variants present among the field samples (from both humans and rodents) were also found among the attenuated JUNV strains. Consequently, the importance of these positions in the attenuation process should be re-evaluated.

The development of a reverse genetics system made it possible to directly test the role of the different variable positions in viral fitness and virulence. Albariño *et al.* used chimeric genomes, bearing a combination of XJ13 and Candid#1-derived sequences, to assess the importance of each portion of GPC. They found that only the Candid#1-derived G2 part was critical for the attenuated phenotype. A further characterization mapped the significant points to the transmembrane domain within G2 [27]. However, a number of variable positions remain to be evaluated by this experimental method.

Previous studies described four conserved regions in the L protein of all arenaviruses [28]. The polymerase domain was mapped to the region III, with the presence of four conserved motifs: A, B, C and D [29]. The variation analyses showed only one change in this region, at position 1156 (R→K→K). Albeit the mutation has a conservative character, it is probably related to the attenuation process because there are no natural variants in this region. The other changes in the L ORF are located between regions II and III (L936, L→L→P) and in region I (L76, H→Y→Y). The modifica-

tion of a Leucine for a Proline brings about a mayor structural change and is present only in the Candid#1 strain, in spite of being immersed in a region of high type 2 mutations index. The change at position L76 is near a predicted ATP/GTP binding site (www.expasy.org). Even though these results are preliminary, they suggest an important role for the polymerase in the virulence attenuation, in accord with reports for other viruses [29-31]. Also, some of the mutations in the structural proteins (nucleoprotein and the two glycoproteins) could be associated with the attenuation of virulence. The N protein has a highly conserved carboxyl-terminus, containing a zinc binding domain [32]. The mutation at N47 (V→V→E) lies outside that region, but could be related to other characteristics of the protein, such as the oligomerization capability [33], possibly interfering with one or both of N's functions in the viral life cycle. In the genome replication, N promotes the synthesis of antigenomic, full-length copies of the S RNA. In the assembly of the virion, the protein binds the genomic RNA, forming the nucleocapsid. The other structural protein, GPC, sported a mutation at the carboxyl end of G1 (GPC168, T→A→A), directly affecting the conserved sequence (N₁₆₆R₁₆₇T₁₆₈K₁₆₉) for a predicted N-glycosylation site (NETGLYC 1.0). For the G2 portion, three domains have been described. First, the outer domain, located at the carboxyl end, which is outside the virion and interacts with G1; second, the transmembrane domain, which interacts with the stable signal peptide, and last, the inner domain, situated at the amino end, which could interact with the nucleocapsid protein or the Z protein, on the inside of the virion [26, 34]. The changes within G2 are in the transmembrane domain (GPC427, F→F→I) or in the cytoplasmic tail (GPC446, T→T→S) and, in both cases, could affect these important interactions.

The potential attenuation markers were found to be unevenly distributed in the genomic segments, with a higher degree of conservation in the L RNA, which indicated a faster evolution rate for the polypeptides encoded in the S RNA. In a comparison of genomes of virulent and avirulent strains of Pichinde arenavirus, Lan and collaborators [29] found that attenuation related mutations mapped to comparable genomic regions, but were fewer in number. If the comparison is limited to only the field strains of Junín virus (XJ13, Romero, MC2 and IV4454) the mutation distribution pattern became markedly different for the large and small genomic segments. In a protein sequence analysis of the field strains, 45 divergent sites were found for the S RNA derived ORFs and 48 sites for the L RNA derived ORFs. Interestingly, for the L RNA, 46 out of the 48 sites were XJ13-specific changes, while for the S RNA, this was only the case in 4 of the 45 sites.

As stated before, the G fragment included the sequence for the cleavage site between G1 and G2. The recognition site sequence has been published as QLPRRSLK₂₅₁AFF [35] with the cleavage occurring between K₂₅₁↓A or L↓K₂₅₁ [36]. Most of the analyzed samples didn't have mutations in that region. Only isolate H_FHA5054 presented changes at positions 244 and 245 (Q→H and L→F respectively), but it is not certain if these modifications would have any detectable outcome. A portion of the variation observed for the G fragment, namely the part corresponding to G1, could be associated with the host jump of a rodent strain to humans (recep-

Table 3. G Fragment Positions with Variations in Amino Acid Sequence. The Changes with Regard to the XJ13 Strain are Annotated in Bold Letters. Type 2 Mutations Among Vaccine Related Strains are Shaded in Gray. Adapted from [13]

Residue Position	XJ13	XJ17	XJ34	XJ39	XJ#44	XJ48	Candid#1	MC2	IV4454	Romero	AN_8640	AN_5185	AN_13365	AN_16501	AN_17058	AN_17246	AN_17116	AN_17064	H_Lye63	H_FHA5069	H_p1879	H_FHA5054	H_8027	
107	S	S	S	S	S	S	S	S	T	S	S	S	S	S	S	S	S	S	S	T	T	T	S	
109	Q	Q	Q	Q	Q	Q	Q	K	M	M	Q	Q	Q	M	Q	Q	M	M	Q	M	M	M	Q	
111	S	S	S	S	S	S	S	S	T	S	S	S	S	S	T	S	S	S	S	T	T	T	S	
116	A	A	A	A	A	A	A	A	A	E	A	A	A	A	A	A	A	A	A	A	A	A	A	
121	Q	Q	Q	Q	Q	Q	Q	E	Q	Q	Q	Q	E	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	
125	I	I	I	I	I	I	I	I	V	I	I	I	I	I	V	I	I	I	I	V	V	V	I	
133	S	S	S	S	S	S	S	S	S	S	S	S	S	N	S	S	N	N	S	S	S	S	S	
143	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	W	R	R	R	W	
151	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	V	V	V	A	
157	H	H	H	H	H	H	H	Y	Y	H	H	H	H	H	H	H	H	H	H	H	H	H	H	
168	T	A	A	A	A	A	A	T	T	T	A	A	T	T	A	A	T	T	A	T	T	T	A	
184	V	V	V	V	V	V	V	I	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	
186	E	E	E	E	E	E	G	E	E	E	G	G	E	E	G	G	E	E	G	E	E	E	G	
206	S	S	S	S	S	S	P	S	S	S	P	P	S	S	P	P	S	S	P	S	S	S	P	
208	P	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	
209	N	N	N	N	N	N	N	D	S	N	N	N	N	N	N	N	N	N	N	N	S	S	S	N
244	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	Q	H	Q	
245	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	F	L	

tor affinity differences). They could therefore be directly linked with changes in the strains virulence (increase or decrease) through a cellular tropism change. Another key feature of this region was the N-glycosylation target sequence N₁₆₆R₁₆₇T₁₆₈K₁₆₉, as determined by NETGLYC 1.0. This precise sequence is only present in the 65% of the field strains, while in the remaining 35% there is an alanine residue (A₁₆₈). Complementary bioinformatics studies performed on this region predicted only slight changes to the protein properties.

On the other hand, the importance of two positions, strongly related to the N-glycosylation of JUNV and MACV was established. Also a high conservation degree of the residues involved, along with their environ was found to be necessary [37]. In recent years, an analysis of LCMV proved the implication of each N-linked glycan of the glycoprotein with GPC expression, fusion activity and infectivity [38]. Interestingly, all the attenuated strains derived from XJ13 have an A₁₆₈ mutation. The hypothesis was that mutations present in field strains (natural variants) would be of minor importance in virulence attenuation. In fact, none of the type two muta-

tion analyzed here were absent in field samples. Therefore, the data presented in (Table 3) could signify that the attenuated phenotype is not influenced by mutations in the G1 region. That result concurs with Albariño and collaborators [39], who mapped the essential attenuation-related mutations to G2.

NON CODING REGIONS

Previous studies suggested that sequence changes within the intergenic region could play a role in the attenuation processes of arenaviruses [40]. Conversely, our analysis of the intergenic regions from Candid#1, XJ48, XJ#44, XJ39, XJ34, XJ17 and XJ13 revealed a total conservation among the strains for both genomic RNAs. If there is a low tolerance for nucleotide changes in the region, it would suggest an evolutive constraint, probably related to the calculated secondary structure and the associated transcription regulation function [32, 41]. Also, there was a high degree of conservation of the non-coding regions at the genomic ends of the analyzed strains, with nucleotide homology values ranging from 93% to 97%.

However, when viral RNAs obtained from infected cells were analyzed, a great variability could be observed at the 3' and 5' non-coding regions of RNAs derived from the same parental strain. The analysis was designed to evaluate the genomic and antigenomic forms separately. A number of 5' non-templated bases were found for both Candid#1 genomic segments. When comparing the 5' end of the genomic RNAs with the 3' end of the antigenomic RNAs (which are the template for the former), at least one additional guanine was present at all 5' ends of genomic clones with additional bases, as has been described for other arenaviruses [42]. On the contrary, the 3' RACE of genomic forms yielded several clones with short deletions. A calculation of RNA secondary structure from Candid#1 S and L RNAs predicted the formation of a panhandle between the 5' and 3' ends of each genomic segment. The observed 3' deletions all mapped to regions within the panhandle (Fig. 3).

Taken together, these results support the employ of cell-derived RNAs as primers for the synthesis of viral RNA, as well as 5' terminal sequences from virus-derived RNAs into incomplete panhandle structures as templates for the completion of the 3' terminal sequences. The resulting heterogeneity may well be a consequence of the action of different RNA editing mechanisms, as reported for other arenaviruses [43]. Still, the role of these regions in the process of attenuation remains to be established.

Interestingly, a comparison of the sequence of the ends (3' and 5') of both genomic segments revealed several positions that were highly conserved and could be part of the minimal promoter. The untranslated sequences on the extremes of the S RNA, as well as the 5' end of the L RNA, have an approximate length of 80 nucleotides. This is not the case for the 3' end of L RNA, which has only 30 nucleotides. Considering that the genomic forms serve as templates for the antigenomic forms and vice versa, all non-coding terminal regions have to include the promoter sequence that is the

recognition site of the viral polymerase, in order to achieve the completion of the replicative cycle. Since one of these regions only has a size of 30 nucleotides, it is reasonable to conclude that the minimal promoter sequence is contained in that range. In support of this, if a sequence comparison is made of the 80 nucleotides of the ends of S and L RNAs, the homology value is 60% for the 5' ends and only 50% for the 3' ends. But, if the analysis is reduced to only 30 nucleotides, the value of the 3' region ascends to 71%. In fact, there is a 19 nucleotide sequence motif highly conserved among arenaviruses (*arena region*) that can be found in all JUNV genomic ends. The remaining 11 bases of the 3' end of the L RNA (GCTCAAGTGCC) have an elevated homology degree with two regions within the S RNA 3' end. These repeated elements could be part of the translation or transcription processes. They have been found in other arenavirus species, in similar distributions. A comprehensive analysis of the genomic sequence of subclade B1 New World arenaviruses [44] showed that the element that mapped to positions 38-46 (GCUCAAGUG for JUNV L RNA and GCUCAGUG for S RNA) was conserved within the group [20]. Therefore, it could be possible to identify a sequence motif at the 3' end of both genomic RNAs of most arenaviruses. Also, in the particular cases of Machupo, Tacaribe, and JUNVS, the motif described as GSYC(A)₁₋₂GUR, has a relatively high positional conservation within the RNA secondary structure when calculated by bioinformatic tools.

CONCLUSIONS AND SUMMARY

This works aim was to present a condensed report of the studies about the molecular determinants of viral virulence in JUNV. A set of putative attenuation markers have been identified within the GPC, Z and L ORFs of JUNV vaccine strain, Candid#1. Some of these positions were more confidently linked to the attenuated phenotype. The main reason for this is that these changes were found inside regions with

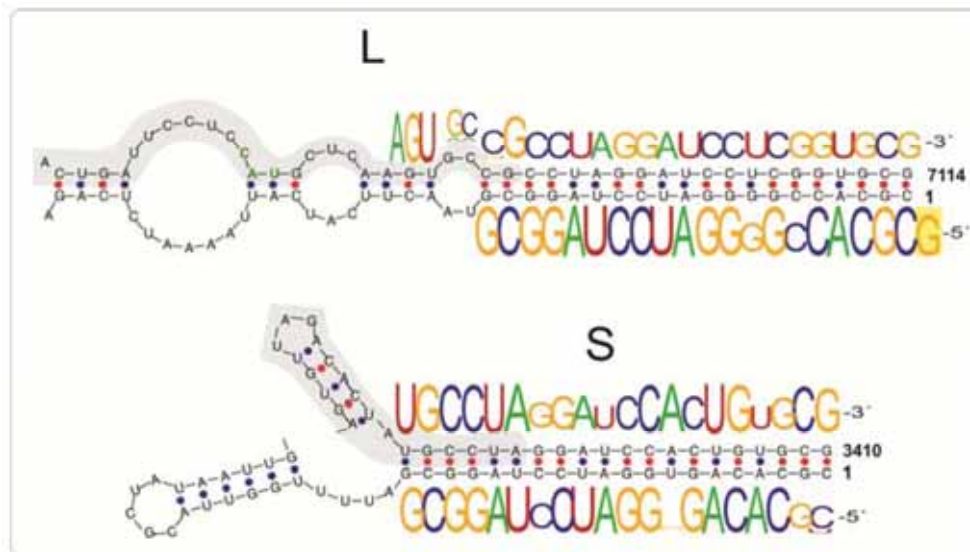


Fig. (3). Junin virus Candid#1 genomic end sequences. Predicted panhandle structures for the L and S RNAs. The region found in all 3'end sequences is shadowed in gray. The sequence-logos show the nucleotide frequency determined by a RACE assay. A non-templated G is highlighted in yellow. Taken from [13].

a low natural mutation frequency (GPC168, GPC427, GPC446, N47, L76, L936 and L1156). They were analyzed with phylogenetic and bioinformatic tools.

Increasing complete genome sequence information availability, particularly from JUNV strains, has been a significant contribution for the identification of further putative virulence attenuation markers. In recent years several sequences were published for members of the vaccine strain lineage [20, 27]. They were included in previously done informatic analyses, confirming the results observed before, that is, only a few changes are necessary to obtain an attenuated phenotype. To further study the role in the virulence attenuation process of different point mutations, Albariño and collaborators developed a JUNV reverse genetics system [39]. Their results showed a direct involvement in the attenuated phenotype of certain point mutations within the GPC ORF of Candid#1.

Moreover, in order to analyze the natural mutation distribution, Goñi and collaborators designed a RT-Nested-PCR based technique and applied it to human and rodent derived samples from the Argentine Hemorrhagic Fever endemic area [45]. The sequences thus obtained were translated *in silico* and analyzed with different tools. A comparison was instituted between the GPC derived sequence data (G fragment) and the results published by Albariño and collaborators [25]. (Table 3) summarized the findings of that analysis. As is expected, the greater variability appears mainly in the portion of the glycoprotein complex that is exposed to the exterior and therefore comes in contact with the immune system. Changes to the sequence and conformation of epitopes could explain an increase in virulence due to escape from the natural host defenses. Also, slight variations within the receptor-binding region have been shown to confer the capability of infecting human cells as well as the animal host [46]. However, there are certain constraints acting on the outer part of the glycoprotein, related to its capacity of efficiently performing the original function within the viral cycle.

Studies addressing similar questions for other arenaviruses produced interesting results. Zhang and collaborators [47] created reassortants using an attenuated and a virulent PICV parental strain. They found that the virulence in guinea pigs could mainly be attributed to the S segment. Nevertheless, the virulence level of the parental strain was not attained, pointing to determinants being present in the L segment. Similar studies made with LCMV strains mapped the attenuation markers of this virus principally to its L segment [31]. Although both studies determine the virulence in an animal model of guinea pigs, the results obtained for each virus are different. This suggests that each attenuation process produced a distinct set of mutations in different regions. It would be interesting to assay what these mutations are for JUNV, principally taking into account the available intermediate strains of the vaccine genealogy, and the immense combination probabilities.

The sequence information accumulated through the analysis of a number of strains with varying virulence degrees will be the springboard for natural biodiversity studies. The mutations identified as potentially central to the attenuation process are ideal candidates for mutagenesis and reverse

genetics assays. However, more complex studies will have to be made to determine the minimal changes that render an attenuated phenotype, since it is likely that many variations act synergically to decrease the observed virulence. Also, compensatory changes might arise to make up for structure or activity losses that hinder viral fitness [48, 49]. Therefore, even if a particular change is found in both field and vaccine strains, it does not necessarily negate its importance in virulence attenuation. Conversely, some variants present in natural and vaccine strains might have a cumulative effect upon the activity of a single viral protein or on the capacity of interaction with other viral or cellular proteins. The Z protein shows a low number of variations. This protein is responsible for the budding of the virus and the correct virion assembly. It also has several functions related to the modulation of the host immune response [50]. All of the roles of Z within the viral cycle are performed through interaction with proteins. The apparently low tolerance of changes within the Z ORF suggests that its functions greatly depend on the correct structure/sequence. Possibly, the regions of interaction of its partners are also submitted to a similar constraint. It will be interesting to analyze the precise regions once they have been positively identified. A great step towards identifying possible co-variation sites will come with the elucidation of the crystallographic structure of the viral polypeptides and their interaction partners. The beginning has been made with the first glycoprotein and partial nucleoprotein and L [51-53] structures and we are hoping for a steady advance in this area.

Summing up, this work presented a set of mutations that are probably related with an attenuated virulence phenotype. Also, most of the mutations analyzed here (Type 1 and 2) were mapped to a few discrete genomic regions. The development of a RT-PCR based method followed by sequencing, was the basis for a broad analysis of JUNV strains in search of natural genomic variability. This method could furthermore be useful if a program for the monitoring of vaccinated individuals is implemented.

CONFLICT OF INTEREST

The author(s) confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENTS

We would like to acknowledge the current and previous members of the LIGBCM-AVEZ laboratory at National University of Quilmes for their contributions. The work is supported by CONICET.

REFERENCES

- [1] Adams, M. J.; Carstens, E. B. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2012). *Archives of virology* **2012**, *157*, 1411–22.
- [2] Cajimat, M. N. B.; Milazzo, M. L.; Hess, B. D.; Rood, M. P.; Fulhorst, C. F. Principal host relationships and evolutionary history of the North American arenaviruses. *Virology* **2007**, *367*, 235–43.
- [3] Downs, W.; Anderson, C. Tacaribe virus, a new agent isolated from Artibeus bats and mosquitoes in Trinidad, West Indies. *Am J Trop Med ...* **1963**, *640–646*.
- [4] Hetzel, U.; Sironen, T.; Laurinmäki, P.; Liljeroos, L.; Patjas, A.; Henttonen, H.; Vaheri, A.; Artelt, A.; Kipar, A.; Butcher, S. J.; Vapalahti, O.; Hepojoki, J. Isolation, identification and

- characterization of novel Arenaviruses, the etiological agent of Boid Inclusion Body Disease. *Journal of virology* **2013**.
- [5] Enría, D. A.; Mills, J. N.; Bausch, D.; Shieh, W.-J.; Peters, C. J. CHAPTER 68 - Arenavirus Infections. In *Tropical Infectious Diseases: Principles, Pathogens and Practice (Third Edition)*; W.B. Saunders: Edinburgh, **2011**; pp. 449–461.
- [6] Ambrosio, A. M.; Saavedra, C.; Riera, L. M.; Fassio, R. M. La producción nacional de vacuna a virus Junín vivo atenuado (Candid # 1) anti-fiebre hemorrágica argentina. *Acta Bioquím Clín Latinoam* **2006**, *40*, 5–17.
- [7] Ministry of Health *National AHF Control Program*; Fourth Edition; Azul, Buenos Aires, Argentina, **2007**; Vol. 5, p. 12.
- [8] Barrera Oro, J. G.; Eddy, G. A. Characteristics of candidate live attenuated Junín virus vaccine. In *Fourth International Conference on Comparative Virology*; Banff, Alberta, Canada, **1982**; pp. 17–22.
- [9] Maiztegui, J. I.; Feinsod, F.; Briggiler, A. M.; Peters, C. J.; Enría, D. A.; Lupton, H. W.; Ambrosio, A. M.; Tiano, E.; Feuillade, M. R.; Gamboa, G.; Conti, O.; Vallejos, D.; MacDonald, C.; Barrera Oro, J. G. Inoculation of Argentine volunteers with a live-attenuated Junín virus vaccine. In *VII International Congress of Virology*; Edmonton, Canada, **1987**; p. 69.
- [10] Ogbu, O.; Ajuluchukwu, E.; Uneke, C. J. Lassa fever in West African sub-region: an overview. *Journal of vector borne diseases* **2007**, *44*, 1–11.
- [11] Sogoba, N.; Feldmann, H.; Safronetz, D. Lassa Fever in west Africa: evidence for an expanded region of endemicity. *Zoonoses and public health* **2012**, *59 Suppl 2*, 43–7.
- [12] Lukashevich, I. S. Advanced vaccine candidates for Lassa fever. *Viruses* **2012**, *4*, 2514–57.
- [13] Goñi, S.; Lozano, M. Molecular Attenuation Process in Live Vaccine Generation for Arenaviruses. In *Point Mutations*; Logie, C., Ed.; Intech, **2012**; pp. 1–25.
- [14] Martínez Segovia, Z.; Grazioli, F. The nucleic acid of Junín virus. *Acta Virologica* **1969**, *13*, 264–268.
- [15] Riviere, Y.; Ahmed, R.; Southern, P. J.; Buchmeier, M. J.; Dutko, F. J.; Oldstone, M. B. A. The S RNA segment of lymphocytic choriomeningitis virus codes for the nucleoprotein and glycoproteins 1 and 2. *Journal of virology* **1985**, *53*, 966–8.
- [16] Pérez, M.; Craven, R. C.; de la Torre, J. C. The small RING finger protein Z drives arenavirus budding: implications for antiviral strategies. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 12978–83.
- [17] Auperin, D. D.; Romanowski, V.; Galinski, M.; Bishopt, D. H. L.; Bishop, D. H. Sequencing studies of pichinde arenavirus S RNA indicate a novel coding strategy, an ambisense viral S RNA. *Journal of virology* **1984**, *52*, 897–904.
- [18] Young, P. R.; Howard, C. R. Fine structure analysis of Pichinde virus nucleocapsids. *The Journal of general virology* **1983**, *64 (Pt 4)*, 833–42.
- [19] Goñi, S. E.; Iserte, J. A.; Ambrosio, A. M.; Romanowski, V.; Ghiringhelli, P. D.; Lozano, M. E. Genomic features of attenuated Junín virus vaccine strain candidate. *Virus genes* **2006**, *32*, 37–41.
- [20] Goñi, S. E.; Iserte, J. A.; Stephan, B. I.; Borio, C. S.; Ghiringhelli, P. D.; Lozano, M. E. Molecular analysis of the virulence attenuation process in Junín virus vaccine genealogy. *Virus genes* **2010**, *40*, 320–8.
- [21] McKee, K. T.; Mahlandt, B. G.; Maiztegui, J. I.; Eddy, G. A.; Peters, C. J. Experimental Argentine hemorrhagic fever in rhesus macaques: viral strain-dependent clinical response. *The Journal of infectious diseases* **1985**, *152*, 218–21.
- [22] Yun, N. E.; Linde, N. S.; Dziuba, N.; Zacks, M. a; Smith, J. N.; Smith, J. K.; Aronson, J. F.; Chumakova, O. V.; Lander, H. M.; Peters, C. J.; Paessler, S. Pathogenesis of XJ and Romero strains of Junin virus in two strains of guinea pigs. *The American journal of tropical medicine and hygiene* **2008**, *79*, 275–82.
- [23] Berria, M. I.; Gutman Frugone, L. F.; Girola, R.; Barrera Oro, J. G. [Immunological studies with Junin virus. I. Formation of antibodies in guinea pigs inoculated with live viruses]. *Medicina* **1967**, *27*, 93–8.
- [24] Candurra, N. A.; Damonte, E. B.; Coto, C. E. Antigenic relationships between attenuated and pathogenic strains of Junin virus. *Journal of medical virology* **1989**, *27*, 145–50.
- [25] Weissenbacher, M. C.; Laguens, R. P.; Coto, C. E. Argentine hemorrhagic fever. *Current topics in microbiology and immunology* **1987**, *134*, 79–116.
- [26] York, J.; Romanowski, V.; Lu, M.; Nunberg, J. H. The signal peptide of the Junín arenavirus envelope glycoprotein is myristoylated and forms an essential subunit of the mature G1-G2 complex. *Journal of virology* **2004**, *78*, 10783–92.
- [27] Albariño, C. G.; Bird, B. H.; Chakrabarti, A. K.; Dodd, K. a; Flint, M.; Bergeron, E.; White, D. M.; Nichol, S. T. The major determinant of attenuation in mice of the Candid1 vaccine for Argentine hemorrhagic fever is located in the G2 glycoprotein transmembrane domain. *Journal of virology* **2011**, *85*, 10404–8.
- [28] Vieth, S.; Torda, A. E.; Asper, M.; Schmitz, H.; Günther, S. Sequence analysis of L RNA of Lassa virus. *Virology* **2004**, *318*, 153–68.
- [29] Lan, S.; McLay, L.; Aronson, J.; Ly, H.; Liang, Y. Genome comparison of virulent and avirulent strains of the Pichinde arenavirus. *Archives of virology* **2008**, *153*, 1241–50.
- [30] Endres, M. J.; Griot, C.; Gonzalez-Scarano, F.; Nathanson, N. Neuroattenuation of an avirulent bunyavirus variant maps to the L RNA segment. *Journal of virology* **1991**, *65*, 5465–70.
- [31] Riviere, Y.; Ahmed, R.; Southern, P. J.; Buchmeier, M. J.; Oldstone, M. B. Genetic mapping of lymphocytic choriomeningitis virus pathogenicity: virulence in guinea pigs is associated with the L RNA segment. *Journal of virology* **1985**, *55*, 704–9.
- [32] Tortorici, M. A.; Ghiringhelli, P. D.; Lozano, M. E.; Albariño, C. G.; Romanowski, V. Zinc-binding properties of Junín virus nucleocapsid protein. *The Journal of general virology* **2001**, *82*, 121–8.
- [33] Levingston Macleod, J. M.; D'Antuono, A.; Loureiro, M. E.; Casabona, J. C.; Gomez, G. a; Lopez, N. Identification of two functional domains within the arenavirus nucleoprotein. *Journal of virology* **2011**, *85*, 2012–23.
- [34] Capul, A. a; Perez, M.; Burke, E.; Kunz, S.; Buchmeier, M. J.; de la Torre, J. C. Arenavirus Z-glycoprotein association requires Z myristoylation but not functional RING or late domains. *Journal of virology* **2007**, *81*, 9451–60.
- [35] Beyer, W. R.; Pöpplau, D.; Garten, W.; von Laer, D.; Lenz, O. Endoproteolytic processing of the lymphocytic choriomeningitis virus glycoprotein by the subtilase SKI-1/S1P. *Journal of virology* **2003**, *77*, 2866–72.
- [36] Lenz, O.; ter Meulen, J.; Klenk, H. D.; Seidah, N. G.; Garten, W. The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 12701–5.
- [37] Bowden, T. a; Crispin, M.; Graham, S. C.; Harvey, D. J.; Grimes, J. M.; Jones, E. Y.; Stuart, D. I. Unusual molecular architecture of the machupo virus attachment glycoprotein. *Journal of virology* **2009**, *83*, 8259–65.
- [38] Bonhomme, C. J.; Capul, A. A.; Lauron, E. J.; Bederka, L. H.; Knopp, K. A.; Buchmeier, M. J. Glycosylation modulates arenavirus glycoprotein expression and function. *Virology* **2011**, *409*, 223–33.
- [39] Albariño, C. G.; Bird, B. H.; Chakrabarti, A. K.; Dodd, K. a; White, D. M.; Bergeron, E.; Shrivastava-Ranjan, P.; Nichol, S. T. Reverse genetics generation of chimeric infectious Junin/Lassa virus is dependent on interaction of homologous glycoprotein stable signal peptide and G2 cytoplasmic domains. *Journal of virology* **2011**, *85*, 112–22.
- [40] Wilson, S. M.; Clegg, J. C. Sequence analysis of the S RNA of the African arenavirus Mopeia: an unusual secondary structure feature in the intergenic region. *Virology* **1991**, *180*, 543–52.
- [41] Franze-Fernández, M. T. Iapalucci, S.; López, N.; Rossi, C. Subgenomic RNAs of Tacaribe virus. In *The Arenaviridae*; **1993**; pp. 113–132.
- [42] Polyak, S. J.; Zheng, S.; Harnish, D. G. 5' termini of Pichinde arenavirus S RNAs and mRNAs contain nontemplated nucleotides. *Journal of virology* **1995**, *69*, 3211–5.
- [43] Garcin, D.; Kolakofsky, D. Tacaribe arenavirus RNA synthesis *in vitro* is primer dependent and suggests an unusual model for the initiation of genome replication. *Journal of virology* **1992**, *66*, 1370–6.
- [44] Flanagan, M. L.; Oldenburg, J.; Reigier, T.; Holt, N.; Hamilton, G. a; Martin, V. K.; Cannon, P. M. New world clade B arenaviruses can use transferrin receptor 1 (TfR1)-dependent and -independent entry pathways, and glycoproteins from human pathogenic strains are associated with the use of TfR1. *Journal of virology* **2008**, *82*, 938–48.

- [45] Goñi, S. E.; Stephan, B. I.; Iserte, J. A.; Contigiani, M. S.; Lozano, M. E.; Tenorio, A. Viral diversity of Junín virus field strains. *Virus research* **2011**, *160*, 150–8.
- [46] Abraham, J.; Kwong, J. A.; Albariño, C. G.; Lu, J. G.; Radoshitzky, S. R.; Salazar-Bravo, J.; Farzan, M.; Spiropoulou, C. F.; Choe, H. Host-species transferrin receptor 1 orthologs are cellular receptors for nonpathogenic new world clade B arenaviruses. *PLoS pathogens* **2009**, *5*, e1000358.
- [47] Zhang, L.; Marriott, K. A.; Harnish, D. G.; Aronson, J. F. Reassortant analysis of guinea pig virulence of pichinde virus variants. *Virology* **2001**, *290*, 30–8.
- [48] Rimmelzwaan, G. F.; Berkhoff, E. G. M.; Nieuwkoop, N. J.; Smith, D. J.; Fouchier, R. a M.; Osterhaus, a D. M. E. Full restoration of viral fitness by multiple compensatory co-mutations in the nucleoprotein of influenza A virus cytotoxic T-lymphocyte escape mutants. *The Journal of general virology* **2005**, *86*, 1801–5.
- [49] Parks, C. L.; Lerch, R. A.; Walpita, P.; Wang, H. P.; Sidhu, M. S.; Udem, S. A. Comparison of predicted amino acid sequences of measles virus strains in the Edmonston vaccine lineage. *Journal of virology* **2001**, *75*, 910–20.
- [50] Fehling, S. K.; Lennartz, F.; Strecker, T. Multifunctional nature of the arenavirus RING finger protein Z. *Viruses* **2012**, *4*, 2973–3011.
- [51] Igonet, S.; Vaney, M.-C.; Vonhrein, C.; Bricogne, G.; Stura, E. A.; Hengartner, H.; Eschli, B.; Rey, F. a X-ray structure of the arenavirus glycoprotein GP2 in its postfusion hairpin conformation. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 19967–72.
- [52] Morin, B.; Coutard, B.; Lelke, M.; Ferron, F.; Kerber, R.; Jamal, S.; Frangeul, A.; Baronti, C.; Charrel, R.; de Lamballerie, X.; Vonrhein, C.; Lescar, J.; Bricogne, G.; Günther, S.; Canard, B. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. *PLoS pathogens* **2010**, *6*, e1001038.
- [53] Zhang, Y.; Li, L.; Liu, X.; Dong, S.; Wang, W.; Huo, T.; Guo, Y.; Rao, Z.; Yang, C. Crystal structure of Junin virus nucleoprotein. *The Journal of general virology* **2013**.