

Negative Strand RNA Virus Lecture I (VSV, Ebola)

Wagner and Rose, Chapter 37 of *Fields Virology*, 3rd Edition.

OVERVIEW: Pringle, C.R. Chapter 29 (p426-437). In: *Molecular Basis of Virus Evolution* (Eds.: Gibbs, Calisher and Garcia-Arenal), Cambridge Univ. Press, 1995

EBOLA: <http://www.cdc.gov/ncidod/dvrd/spb/mnpages/dispages/ebola.htm> (see links also, and <http://www.journals.uchicago.edu/JID/journal/contents/v179nS1.html>)

The **Mononegavirales** (*mono* - single; *nega* - negative; *virales* - viruses) are a taxonomic order, which includes several families of viruses with similar genomic organization and replicate strategies -- the *Filoviridae*, *Paramyxoviridae* and *Rhabdoviridae*, plus **Borna disease virus**. These viruses probably diverged from a single common ancestor as recently as the last ice age. They are also frequently associated with **emerging infections** and/or cross-species transmission events (eg, Ebola).

Use of negative sense (-) RNA genomes means, by definition, that the viral genome is of opposite polarity to mRNA. Thus, the viral genome cannot be used to make proteins until it has first been transcribed to produce mRNAs. This has the following implications:

1. purified virion RNA is not infectious (as noted above, it cannot encode protein)
2. the viruses must bring their own RNA polymerase into the cell in order to make mRNA (ie, the viral polymerase must be incorporated into the viral particle, or virion)

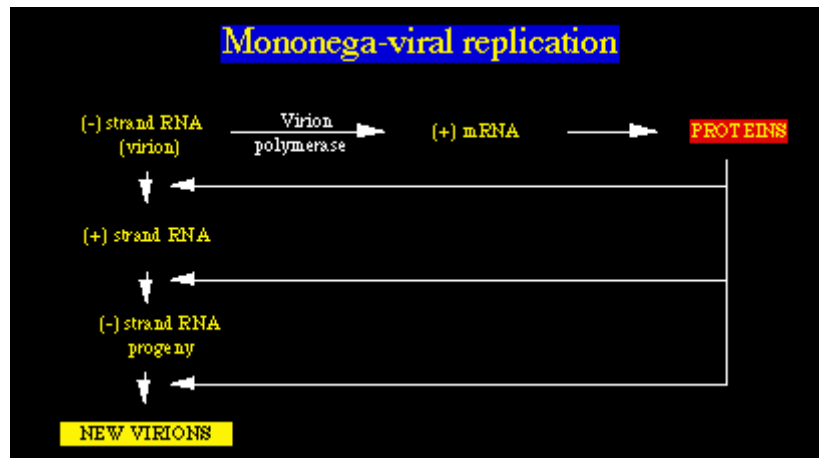
The other key feature of these viruses is that they make **gene-unit length mRNAs** (ie, each mRNA encodes only a single protein). This is achieved by the use of transcriptional stop and start signals, which are located at the boundaries of all of the viral genes. **Stop/start transcription** has two major results:

1. Since there is only a single promoter, located at the 3' end of the viral genome, the polymerase can only load onto its RNA template at one site. As it moves along the viral RNA, the polymerase encounters stop/start signals at the boundaries of each of the viral genes. This results in pausing of the enzyme, which often falls off the template. The result is that more mRNA is made from genes that are located close to the promoter, and less mRNA is made from genes located far from the promoter. This means that there is a **polarity of transcription** (see Figure below). The viruses use this to regulate the expression of their genes, since highly expressed proteins are encoded close to the promoter (eg, structural proteins such as the nucleocapsid protein, N), while proteins that are needed in only small amounts (eg, enzymes such as the RNA polymerase, L) are encoded far away from the promoter.
2. The other major consequence of stop/start transcription is that it **complicates genome replication**. The only way that the complete viral RNA genome can be copied is if the transcriptional stop/start signals can be ignored or over-ridden. This means that the critical decision during viral RNA synthesis occurs very early on -- at the first gene boundary (located between the leader RNA and the N gene). If the stop/start signals here are obeyed, then only subgenomic mRNAs will be produced. However, if the stop/start signal here is ignored or over-ridden, then a complete copy of the viral genome can be made.

Transcriptional polarity

[Sudan]				
Paramyxo-	10-12 proteins; 15-18 kb	Pleomorphic	Vertebrates	Mainly respiratory
Subfamily: Paramyxovirinae				
• <u>Genus</u> : Morbillivirus (e.g., measles virus, canine distemper)				
• <u>Genus</u> : Respirovirus (e.g., parainfluenzaviruses [Sendai = PIV-1])				
• <u>Genus</u> : Rubulavirus (e.g., mumps virus)				
• <u>Possible new future genus</u> : Henipavirus (e.g., Hendra virus, Nipah virus); largest of paramyxoviruses				
Subfamily: Pneumovirinae				
• <u>Genus</u> : Pneumovirus (e.g., respiratory syncytial virus)				
Rhabdo-	5 proteins; 11-15 kb	Bullet shape	Animals, plants	Fever, neurologic
• <u>Genus</u> : Lyssavirus (e.g., rabies virus)				
• <u>Genus</u> : Vesiculovirus (e.g., vesicular stomatitis virus)				
Borna-	5 proteins; 9 kb		Animals	Neurologic
• similar to member of plant virus genus, Nucleorhabdovirus				

Generic mononegaviral replication scheme



Filoviridae

History/Outbreaks.

In 1967 simultaneous outbreaks of hemorrhagic fever occurred in Yugoslavia and in Germany, in lab workers who were processing kidneys from African green monkeys. There were 31 cases and

7 deaths. The virus was first characterized in **Marburg**, Germany and traced to a single shipment of Ugandan monkeys. Sporadic additional cases showed up in 1975, 1980, 1982 and 1987.

In 1976 there were epidemics of severe hemorrhagic fever in Zaire and Sudan. In Zaire, there were approximately 300 cases with an 80% fatality rate (due to Ebola-Zaire; EBO-Z). In Sudan, there were a roughly similar number of cases, with a fatality rate of roughly 50% (due to Ebola-Sudan; EBO-S).

Ebola virus was originally isolated in Zaire (now Democratic Republic of the Congo), and it was named after a small river in N.W. Zaire. Ultrastructurally the virus resembled Marburg virus but it was antigenically (and genetically) distinct. It now appears that at least three and probably four EBO viruses exist -- EBO-Z (Zaire), EBO-S (Sudan), EBO-CI (Côte d'Ivoire) and EBO-R (Reston). The first two are known to be highly lethal in humans and are spread via bodily fluids and by close (nonsexual) contact. The Reston virus *appears* to be less lethal in humans (0 deaths in 6 cases), although it is lethal in nonhuman primates.

Outbreaks of Ebola occurred in 1995 in the Kikwit area of Zaire (over 315 cases, with 80% fatality; due to EBO-Z) and in the Gulu region of Uganda in 2000 (over 400 cases, but with roughly 50% fatality; due to EBO-S). It is uncertain how the Kikwit and Gulu outbreaks started. However, a smaller outbreak in 1996 in Gabon was traced to a group of 20 young Gabonese who trapped and caught a Chimpanzee that was sick. It is believed that exposure to Ebola occurred during the preparation of the Chimpanzee, prior to cooking and consumption of the animal. Interestingly, Ebola was isolated only from meat-eating Chimps, and not from strictly vegetarian members of the same troupe of animals.

Outbreaks of EBO-Reston have occurred in US primate colonies in the Washington area (Reston, 1989) and in Texas (1990, 1996). These outbreaks were contained by destruction of all animals within the affected area of the facility. The outbreaks all appear to trace back to shipments of macaques from a single Philippine exporter. A total of 6 humans have become infected by EBO-Reston, but none has died.

Finally, while the major route of Ebola transmission is clearly close contact with bodily fluids and blood (eg, during health care, preparation for burial, etc), it is *possible* that some Ebola viruses *might* be transmissible via an aerosol route in some cases. One piece of evidence to support this idea is the fact that EBO-Zaire has been shown to infect rhesus monkeys that did not have direct contact with experimentally inoculated monkeys held in the same room (*Jaax et al. Lancet 346:1669, 1995*).

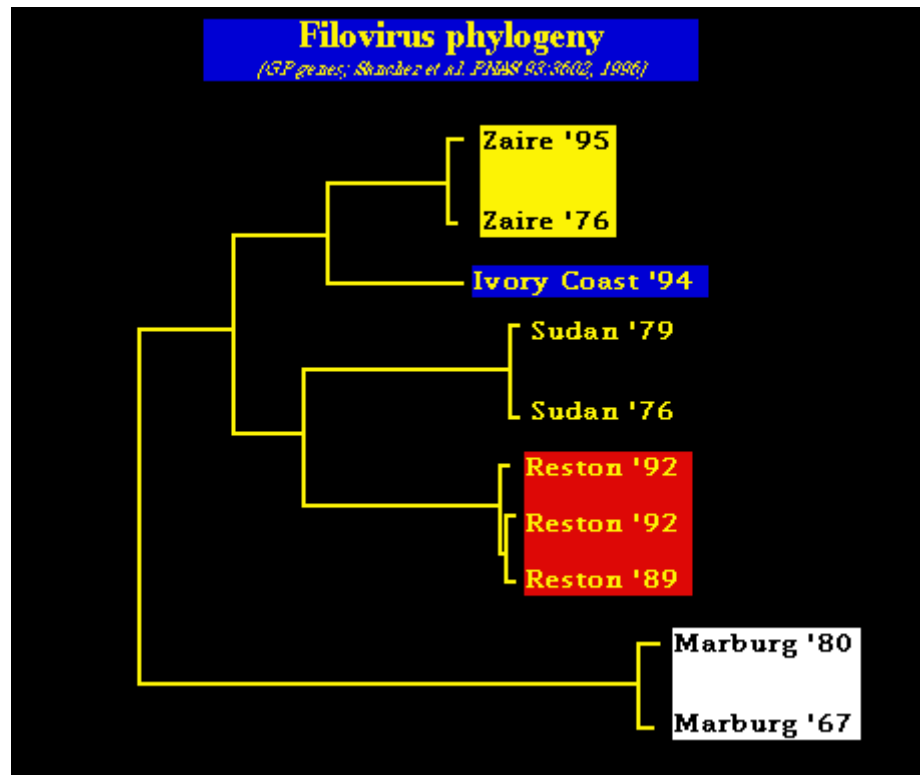
Filoviruses are classic emerging infections. Filoviruses are Biosafety Level 4 agents (cf. HIV is only 2+). They are filamentous with a linear ~13-19kb genome. They can infect mice, hamsters, guinea pigs and monkeys -- although the **viral reservoir in the wild is not known**. Human epidemics seem to be related to blood-borne nosocomial spread (often due to re-use of needles in hospitals; *nosocomial = hospital infection*) and to close contact with infected persons (since this is a hemorrhagic disease, this presumably would involve exposure to large amounts of blood). Primary infections with Marburg and Ebola are 25-90% fatal. Death is thought to be due to visceral organ necrosis (eg, liver) due to viral infection of tissue parenchymal cells. It is uncertain what role hemorrhage has in death.

Wild-caught monkeys are now quarantined before release to US primate centers.

Cloned viruses will help the development of diagnostic serologic tests for infection, and work is also progressing to try to develop a vaccine for Ebola virus. The first successful vaccination against this virus was reported in 1998, by Gary Nabel's group at the University of Michigan. In this report, a DNA vaccine encoding the Ebola virus glycoprotein was able to elicit a T-cell based

immune response in guinea pigs, which was sufficient to protect the animals against infection with a live-Ebola virus (*Xu et al. Nature Medicine 4:37, 1998*). Subsequent studies in nonhuman primates have confirmed that a DNA vaccine can represent an important component of an effective Ebolavirus vaccine. Specifically, a combination of DNA immunization and boosting with adenovirus vectors encoding viral proteins resulted in the protection of cynomolgus macaques from an otherwise lethal dose of highly pathogenic, wild-type Ebola Zaire virus (*Sullivan et al. Nature 408:2000*).

Filovirus genetics: Sequence analysis of Ebola viruses from outbreaks in 1976 and 1995 revealed a surprisingly high degree of genetic conservation for an RNA virus. One interpretation of this is that EBO viruses have coevolved with their natural reservoirs and do not change substantially in the wild (see below).

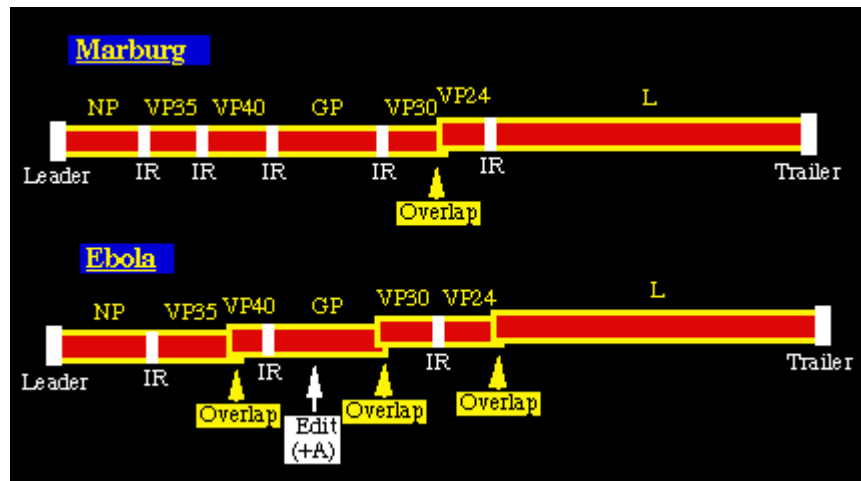


Overall, Filoviridae are more closely related to paramyxoviruses than to rhabdoviruses. Based on genetic analysis, two distinct groups identified (Marburg and Ebola). Unique properties of the filoviruses, compared to other mononegaviruses are (i) largest genome size, (ii) 3'- UAAUU is found between most genes, (iii) several overlapping genes; (iv) putative immunosuppressive domain in GP gene product. There is also one important difference between Marburg and Ebola-- in Marburg, the GP is encoded in a single open reading frame, while in Ebola, GP is encoded in two open reading frames. Expression of GP therefore involves a **site-specific RNA editing** event that is analogous to one which occurs in Measles virus. Specifically, a non-templated A residue is added to the mRNA, which allows joining of the two open reading frames. This results in the production of both a truncated, soluble form of the Ebola virus glycoprotein (sGP; 50-70 kD in size) and a full-length, transmembrane anchored version of the same protein (GP; 120-150 kD in size).

Filovirus genome structures

IR: intervening regions; GP: viral glycoprotein; VPxx: viral proteins; Editing site: addition of a

nontemplated A



Ebolavirus sGP and GP have different functional properties, which may be important in disease pathogenesis. The functional subdomains of these molecules are shown below.

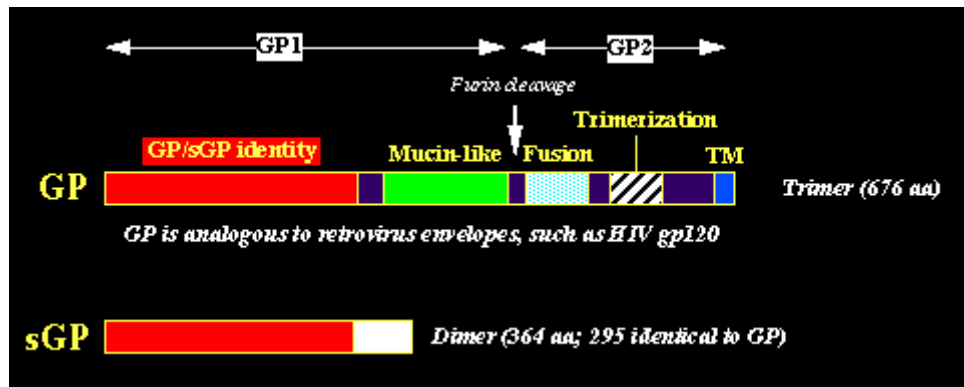
sGP: The soluble sGP molecule is secreted as a trimer, and is identical at its N-terminus to the homologous region of the transmembrane glycoprotein (GP). sGP interacts with neutrophils through CD16b, the neutrophil-specific form of the Fc γ receptor III, whereas the transmembrane glycoprotein (GP) interacts with endothelial cells but not with neutrophils (*Yang et al. Science 279:1034, 1998*). It is possible that interaction of sGP with neutrophils results in the blockade of early events in the activation of these cells, thereby inhibiting inflammatory responses which might contribute to innate protection against viral infection. sGP may also act as a "decoy" for antiviral antibodies.

GP: The transmembrane glycoprotein is produced as a long precursor, which undergoes cleavage by a cellular protease (furin), to produce GP1 and GP2. These can be viewed as being somewhat analogous to HIV-1 gp120 and HIV-1 gp41 (which are produced by cellular proteolytic cleavage of the gp160 precursor).

Ebolavirus GP2 remains in the membrane (due to its transmembrane domain) and is responsible for mediating fusion between the virus and the plasma membrane, via its fusion domain. The GP1 component is attached to GP2 via a non-covalent linkage, and is thought to mediate virus attachment to its host cell(s), which include vascular endothelial cells.

Ebolavirus GP is also cytotoxic for vascular endothelial cells *in vitro*, and this is thought to contribute to the virus' ability to trigger vascular leakage (hemorrhage) *in vivo*.

Ebolavirus glycoproteins



Legend:

2. GP: transmembrane glycoprotein (subsequently cleaved into GP1 and GP2 subunits)
3. sGP: soluble glycoprotein
4. GP/sGP identity: region shared by sGP, GP
5. Mucin-like domain: highly glycosylated domain of GP that is essential for cytotoxicity
6. Fusion domain: responsible for membrane fusion; located within GP2
7. Trimerization domain: allows GP2 to form stable trimers, like other viral fusion proteins
8. TM: transmembrane domain: anchors GP2 in the membrane

Borna Disease Virus

Pathogenesis: Borna disease virus (BDV) is a **neurotropic** agent that naturally infects horses and sheep, and which is capable of infecting primates. The disease induced by BDV resembles neuropsychiatric illnesses (*schizophrenia*). Viral replication is unusual in that it is non-cytolytic and there is little cell-free virus released. *Recently, BDV RNA has been detected in some human brain samples from neuropsychiatric patients (de la Torre et al., Virology 223:272, 1996).*

Genetics: The complete BDV genome has been sequenced and three conserved sequence blocks were identified, similar to those found in other mononegaviruses. From left-to-right, these are: (1) Nucleoprotein and polymerase cofactors; (2) Matrix and envelope and (3) Viral polymerase. Unique molecular aspects of BDV are that (1) there is **nuclear replication and transcription**, with a high level of spliced mRNAs and (2) BDV is related to the plant virus genus, *Nucleorhabdovirus*.

Rhabdoviridae

Greek "*rhadbo*": rod-shaped

Over 100 rhabdoviruses exist & they infect almost all animals. Two genera affect mammals:

Genus	Features	Example
Lyssa-	invade CNS; (<i>fr. Greek "lyssa": frenzy</i>)	Rabies virus
Vesiculo-	invade epithelial cells (usu. tongue) & cause vesicles	Vesicular stomatitis virus (VSV)

Rabies: Causes encephalitis in animals and in humans they bite. Rabies virus can infect all warm-blooded animals, but some are more susceptible than others (eg, foxes, wolves > dogs, skunks, raccoons > opossum). It is spread to humans via animal bites. In the US it is most prevalent in skunks, but also found in raccoons and sometimes in bats. Elsewhere rabies is more common in humans, due to its presence in dogs.

Bats and rabies: In the U.S. about 1-2 cases of rabies occur each year. Since 1990, 20 of 22 domestically acquired human rabies infections in the United States have resulted from infection with bat rabies variants, and in only one of these cases was there a clearly documented bat bite (<http://www.wadsworth.org/rabies/bat.htm>). Many of these bat rabies strains were silver-haired bat (SHB) rabies virus. SHBRV is carried both by silver-haired bats (relatively rare and solitary) and also by other strains of bats (*overall, much less than 1% of all bats test positive for rabies, and the bats most often found around humans -- brown bats -- have never been shown to have cause human disease*). The rarity of SHBRV is strongly suggestive that something unusual is going on here. In addition, data show that the SHBRV variant replicates with unusually high efficiency in cultured epithelial cells, particularly at low temperatures (34°C). This may allow the virus to replicate more efficiently in the skin (*Morimoto et al. PNAS 93:5653, 1996*).

CLINICAL PICTURE: Whether disease results reflects the location and severity of bites (typically ~15% rate of infection). Disease onset is slow, with an unusually variable incubation period (can be over a year) during which virus replicates in muscle near the entry site. Thereafter, the virus enters peripheral nerves. From here, it travels to spinal ganglia & enters the brain. It is then disseminated to all tissues (including salivary gland). Death is inevitable if the virus enters nerves, but post-exposure intervention before this is generally successful.

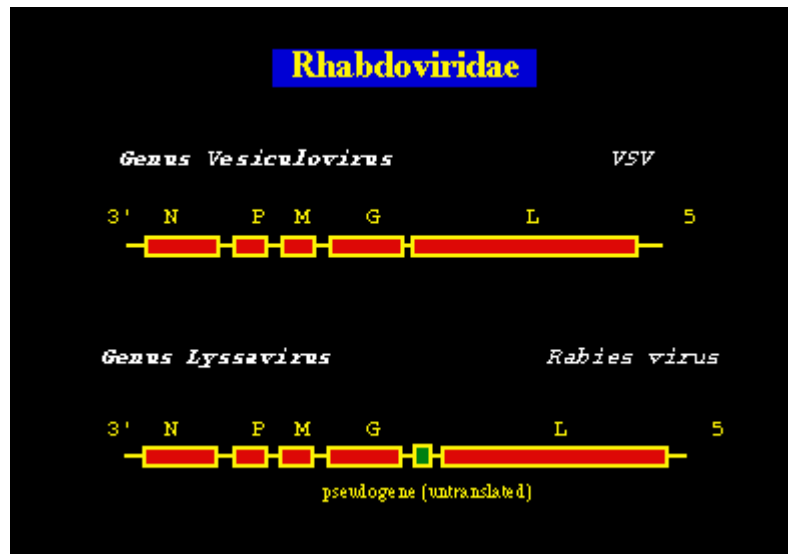
Control of rabies: is achieved by controlling its animal reservoir -- ie, by vaccinating domestic animals and also by the use of vaccine-containing bait to target wild animals. For exposed humans, there is a vaccine.

TROPISM: Rabies virus is neurotropic. It enters nerves in part via the acetylcholine receptor, but the virus can also infect AChR-negative cells, suggesting that infection may involve more than one receptor. Rabies is only very mildly cytopathic *in vitro*.

VSV: Causes epidemic but self-limiting vesicular disease of cattle. Also infects swine, horses, humans & even insects (**very broad host range**). In humans, it causes a mild flu-like illness that's fairly common in lab workers. In keeping with its broad host range, the VSV receptor is not a protein (prolonged trypsinization of cultured cells doesn't block infection). It may be phosphatidyl serine.

CYTOPATHICITY: Induces rapid CPE *in vitro*. Virus can be assayed conveniently by plaque assay (ie, exposure a monolayer of cultured cells to VSV, then wash & overlay with semi-solid media - this permits diffusion of virus to neighboring cells, but prevents convection to other regions of the plate).

Rhabdovirus genomes

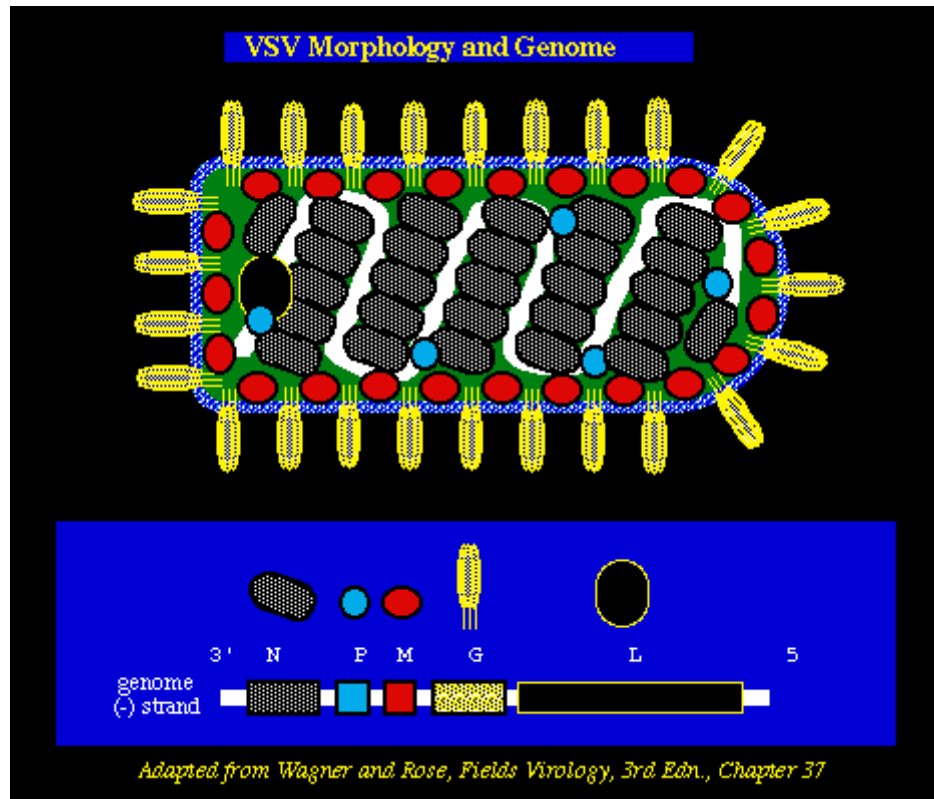


RABIES: The rabies genome is very similar to that of VSV. The biggest difference is that the intergenic regions are longer and more divergent in rabies virus, and the virus also contains an untranslated pseudogene. The best studied rabies protein is the G-protein, which is important for vaccination since it is the target of neutralising antibodies. It is also important for pathogenesis, since attenuated strains of rabies contain a mutation at amino acid residue 333 relative to wild-type strains. This results in a loss of pathogenicity, a reduction in kinetics of spread within the CNS, and a decreased efficiency of infection of nerve cells *in vitro*.

Molecular Biology of VSV

Overall, the molecular biology of VSV is considerably better understood than that of rabies virus

The morphology and structure of VSV is similar to that of rabies virus. The particles are bullet-shaped and are composed of two major structures -- a **nucleocapsid** or ribonucleoprotein (RNP) core and a lipoprotein **envelope** which surrounds that core.



Viral Proteins/Structure

VIRAL RNP CORE: The nucleocapsid or RNP core is the infectious component of VSV and all other rhabdoviruses. As shown in the diagram, this core includes the viral genomic **RNA** which is tightly associated with the highly abundant **nucleocapsid protein (N)**. The RNP core also contains less abundant proteins -- the **phosphoprotein (P)**, and the **viral RNA polymerase (L)**. *The relative abundance of these proteins, on a per virion basis is as follows: approx. 1250 molecules of N, 500 molecules of P and only 50 molecules of L (not surprising, in that L is an enzyme whereas the other proteins are structural in nature).*

N protein: The function of the N protein appears to be (1) to promote RNA encapsidation or packaging and (2) to allow genome replication, by favoring antitermination of transcription (ie, by allowing the viral polymerase to read-through the stop/start signals located between the viral genes).

P protein: The function of the P protein, which is highly acidic due its phosphorylation, appears to be (1) to act as a polymerase co-factor, possibly by helping to displace N protein from the viral RNA and (2) to bind to the N protein, perhaps allowing it to encapsidate the viral RNA (this is an important step in RNA packaging).

L protein: This is the viral RNA-directed RNA polymerase. It is not active on its own, however, since P protein is needed for catalytic activity.

Viral RNA: As shown in the diagram, the VSV genome contains 5 genes in the order 3' N-P-M-G-L 5'. *The P gene also encodes two smaller products from a second reading frame, although their function remains unclear.* Each of these genes is separated by a very short intervening sequence of only two nucleotides.

VIRAL ENVELOPE: The major components of the VSV envelope are (1) the membrane-anchored viral **glycoprotein (G)** and (2) the **matrix protein (M)**. Roughly equivalent amounts of the two proteins are found in each virion (approx. 1500 molecules per virion).

G protein: The glycoprotein, G, forms trimeric spikes on the surface of the viral particle and it forms both the major antigenic determinant on the virus, as well as the major receptor-binding molecule on the virus. G protein undergoes a conformational shift at mildly acidic pH (< 6.0), which stabilizes the trimer and exposes a hydrophobic domain that can insert into cellular membranes and allow membrane fusion to occur. Thus, VSV fusion is activated in the endocytic vesicle, in response to acidic pH.

M protein: This is a highly basic protein. Its functions are believed to include (1) an involvement in virion assembly; (2) a role in inhibiting viral RNA synthesis, in order to allow RNA packaging to start (this may occur as a result of condensation of the RNP core) and (3) a role in inhibiting cellular gene expression by shutting off transcription by cellular RNA polymerases I and II, and also by preventing the nuclear export of cellular mRNAs (*von Kobbe et al., Mol. Cell 6:1243, 2000*).

Viral gene expression

After entry into its host cell, and uncoating of the RNP core, VSV begins to express its genes. Since the viral genome is of negative sense (ie, of opposite polarity to mRNA), the very first step is transcription of viral mRNAs. *Note that the RNP core is transcriptionally active, and that is not inhibited by actinomycin-D (unlike cellular RNA polymerases, which use a DNA template to direct mRNA synthesis).*

Viral transcription begins at the 3' end of the viral genome, at a **single promoter element**, and proceeds sequentially across the genome. It is generally believed that the individual gene-unit-length mRNAs are produced by a **stop-start** transcription mechanism (see below). One result of this is that the transcriptase pauses and transcription is attenuated about 30% at each gene junction. This in turn produces a gradient of mRNA production, such that N>P>M>G>L.

Stop/start transcription is achieved by the presence of transcriptional signals at gene boundaries. There is a 5'-initiation signal, as well as 3'- polyA and termination signals, which are ordered: [polyA signal/terminator]--[intergenic region]--[initiator]. *Note that the intergenic region (2 nucleotides) is not transcribed during viral mRNA synthesis.*

Viral RNA replication

Unlike viral mRNA transcription, viral RNA replication requires the virus to form a single complete copy of its genome. Thus, replication differs from mRNA transcription in that transcriptional "start/stop" signals must be ignored somehow.

The decision to replicate the viral genome must therefore be made when the first intergenic region is encountered (this is located between the region that encodes the short untranslated leader RNA and the gene encoding the N-protein). This **intergenic region must be read-through** in order for viral RNA replication to occur.

Interestingly, viral RNA replication **requires active translation** (this was proved experimentally, since viral RNA replication, but not viral mRNA synthesis, was blocked by inhibition of protein synthesis using cycloheximide). This observation is consistent with a model in which newly formed viral N protein selectively binds to the viral leader RNA. By doing so, N prevents the

recognition of transcriptional termination signals. Thus, the switch from mRNA synthesis to RNA replication is regulated principally by the **anti-termination activity of the N protein**.

Note that the full-length (+) sense viral RNA does not contain any transcription regulatory sequences, and thus functions only as a template for full-length (-) strand production (ie, for production of new viral genomes).

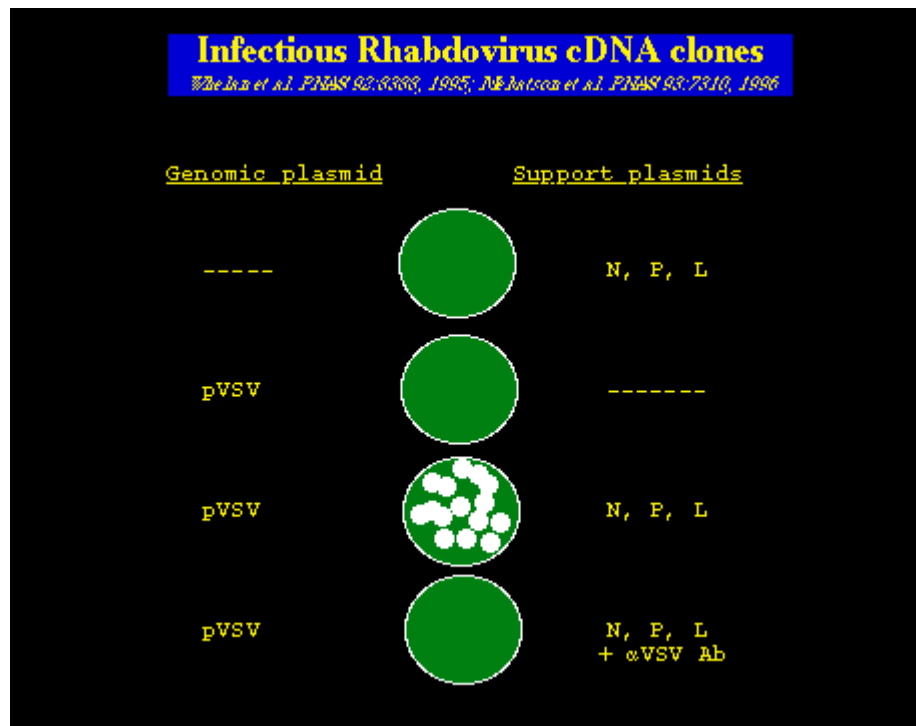
Molecular Genetics and Vectors

Replication of VSV has been difficult to study using recombinant DNA methods. This is because deproteinized RNA is not infectious. Likewise, RNA transcribed from cDNA clones is not by itself competent to initiate infection. The virion RNA-polymerase is needed for infectivity, and in addition the RNA must be encapsidated to be a functional template for the polymerase. Finally, constant synthesis of N protein is needed for replication.

Recently, great progress has been made in this area. Methods for the production of infectious virus from cDNA clones of both VSV and Rabies virus have been developed. This is allowing the development of novel vector systems based on VSV and Rabies virus, and has also permitted investigators to perform genetic analyses on the genomes of these viruses.

Infectious Rhabdovirus cDNA clones

Schell et al. EMBO J. 13:4195, 1994; Lawson et al. PNAS 92:4477, 1995; Whelan et al. PNAS 92:8388, 1995; Mebatson et al. PNAS 93:7310, 1996



KEY: The dark circles represent cell monolayers, and the white areas are plaques in these monolayers, which were caused by infectious virus. You can see that infectious virus was produced **ONLY** when an intact cDNA copy of the VSV genome was introduced into cells together with support plasmids encoding the viral RNP constituents (ie, N, P and L). The bottom

panel proves that the plaques seen were due to VSV, since plaque formation was blocked by the addition of a neutralizing antiserum directed against VSV (anti-VSV Ab).