Review article

Pneumoviruses: Molecular Genetics and Reverse Genetics

Gholamreza Ahmadian¹, Mehdi Shamsara¹,² and Andrew J. Easton*²

¹Department of Molecular Genetics, National Institute of Genetic Engineering and Biotechnology, P.O. Box: 14155-6343, Tehran, I.R. Iran ²Department of Genetics, Faculty of Basic Sciences, Tarbiat Modarres University, P.O. Box: 14115-111, Tehran, I.R. Iran ³University of Warwick, Coventry, United Kingdom

Abstract
Pneumoviruses are responsible for significant respiratory disease in their hosts and represent a major problem for human and animal health. Pneumoviruses are members of the family Paramyxoviridae, subfamily Pneumovirinae and the virus particles consist of a negative-sense, nonsegmented RNA genome within a helical nucleocapsid structure enveloped in a lipid membrane derived from the host cell. Over the past four decades much work has extended our understanding of the molecular biology and pathogenesis of pneumoviruses but despite this only limited treatments and prophylaxis are available. The human pathogen, respiratory syncytial virus (hRSV) which belongs to the genus of Pneumovirus is the best characterized of the subfamily. HRSV is the major cause of hospitalisation of very young children with respiratory disease worldwide. No vaccine is available though new treatments offer some respite for children in the highest risk groups, the immunocompromised and children with congenital heart disease. The recently discovered human pathogen human metapneumovirus (hMPV) belongs to the genus Metapneumovirus and recent data indicates that this virus is second only to hRSV in terms of disease impact. The pneumoviruses also include agents of veterinary importance such as bovine respiratory syncytial virus (bRSV), ovine and caprine RSV, and pneumonia virus of mice (PVM: all in the genus Pneumovirus) and avian metapneumovirus (APV: genus Metapneumovirus). The development of reverse genetics systems for negative strand RNA viruses has opened the possibility of manipulating the virus genomes to identify genes involved in pathogenesis and to explore the biological consequences of specific mutations. This information is informing the rational design of new vaccines. These plasmid-based systems have shown that for all paramyxoviruses the N, P and L proteins are necessary and sufficient for RNA replication. However, the pneumoviruses differ from the other family members in that fully efficient transcription from the virus genome requires the presence of an additional protein encoded by the M2 gene. The present article reviews pneumovirus biology and molecular genetics including a discussion of current concepts of Pneumovirus reverse genetics.

Keywords: Paramyxoviridae; Pneumovirus; Reverse Genetics; Replication; Vaccine.

Table of content:
Introduction
Virion structure
pneumovirus genome organization
Pneumovirus Gene Products
- Nonstructural (NS1 and NS2) proteins
- Nucleocapsid (N) protein
- M2 proteins
- SH proteins
- Phosphoprotein (P)
- Matrix (M) protein
- Attachment (G) protein
- Fusion (F) protein
- Large (L) polymerase protein
Transcription and replication of the pneumovirus genome
Packaging and Assembly
Reverse genetics
Applications of reverse genetics in vaccine strategies

*Correspondence to: Andrew J. Easton, Ph.D.
Tel: +44 (0) 2476 524727; Fax: +44 (0) 2476-523701
E-mail: a.j.easton@warwick.ac.uk
INTRODUCTION

Viruses which synthesis a double stranded DNA molecule during replication such as the retroviruses are easy target for genetic manipulation. Transfection of full-length cDNA molecules leads to establishment of replication virus particles (Wei et al., 1981). Of the other viruses with an RNA genome, most of the other positive strand RNA viruses are also suitable to genetic engineering approaches. Full-length genomic RNA has been show to be infectious when transfected into cells. Plus-strand RNA serves as mRNA for the synthesis of viral proteins as well as template for viral RNA replication. Thus, transfection of cloned cDNA of Poliovirus RNA into permissive cells results in the formation of infectious virus particles (Racaniello and Baltimore, 1981). The life cycle of negative strand RNA viruses differs from that of the other RNA viruses in many ways. The genomic RNA of negative-strand RNA viruses is not infectious, replication does not involve a DNA intermediate and infectious virus particles must also deliver their own RNA-dependent RNA polymerase into the infected cell to start the first round of virus specific mRNA synthesis.

Thus, approaches different from those used for positive-strand RNA viruses have had to be developed to allow the rescue of genetically engineering viruses of negative strand RNA viruses. These include the rhabdoviruses (such as rabies virus), paramyxoviruses (measles virus, mumps virus, respiratory syncytial virus), orthomyxoviruses (influenza virus) and filoviruses (Marburg and Ebola viruses). The genome of these viruses is found in both virion and infected cells to be complexed with the viral nucleoprotein (NP or N) as ribonucleoprotein (RNP) complexes, and it is these RNP complexes, rather than naked viral RNA, that are the templates recognized by the viral RNA polymerase. The minimal unit of a negative-strand RNA virus that is able to initiate infection inside the host cell is genomic or anti-genomic RNP complexed with the viral RNA polymerase. Attempts to generate recombinant negative strand RNA viruses have thus required the establishment of methods to generate such complexes from synthetic cDNA. Genetic manipulation of a negative strand-RNA virus was first made possible in 1990 for the segmented influenza-A virus using biological active viral RNP complexes that were reconstitute in vitro (made of synthetic RNA and purified nucleoprotein and polymerase protein) and then transfecting the complex into cells previous infected with a fully functional, helper, virus (Enami et al., 1990).

Pneumoviruses (family Paramyxovirinae, sub-family Pneumovirinae) are enveloped viruses with negative-sense, nonsegmented RNA genomes. The human pathogen, respiratory syncytial virus (hRSV) remains the best characterized of this group. This agent is most frequently isolated in acute respiratory tract infection in children and is now regarded as the major aetiologic agent associated with bronchiolitis and pneumonia in infancy, and is the single main reason for the hospitalisation of children in their first year of life. HRSV is also responsible for annual winter epidemics of serious illness in infants in temperate climates (Collins et al., 1996b; Holberg et al., 1991). The disease in children is not life-threatening unless other contributory factors are present such as congenital heart problem, immunosuppression or malnutrition. RSV isolates are divided into two antigenically distinguishable subgroups, A and B, based on the reactivity of monoclonal antibodies to the virus proteins (Gimenez et al., 1986; Anderson et al., 1985; Mufson et al., 1985). The recently discovered human pathogen human metapneumovirus (hMPV) also belongs to this virus subgroup, as do the veterinary pathogens avian metapneumovirus (APV), bovine respiratory syncytial virus (bRSV), ovine and caprine RSVs, and pneumonia virus of mice (PVM). The text begins with a review of pneumovirus biology and molecular genetics, continues with a discussion of current concepts related to reverse genetics of pneumovirus, and concludes with a discussion of live-attenuated virus vaccines.

**Virion structure:** Electron microscopic analysis of the virions of all four members of the subfamily Pneumovirinae demonstrates that they are structurally similar to one another and similar to other members of the family Paramyxoviridae. Virions are irregular in shape and range from 150 to 300 nm in diameter. However, they are very pleomorphic and filamentous forms that are 60-100 nm in diameter and up to 10 μm in length can be observed when virus is grown in tissue culture. The significance of these filamentous particles is unknown (Berthiaume et al., 1974; Bach and Howe, 1973; Norrby et al., 1970). A diagrammatic representation of a typical pneumovirus is shown in figure 1. Virions form within the cytoplasm of infected cells, assemble at the cell membrane and mature by budding, during which they acquire the lipid envelope which is derived directly from the plasma membrane of the host cell. The envelope contains the three virus glycoproteins, G, F, and small hydrophobic (SH) proteins (Huang et al., 1985; Collins and Mottet, 1993). The G and F glycoproteins are transmembrane spikes and serve for attachment and fusion, respectively. The precise localization and function of the SH glycoprotein
The F and G glycoproteins are believed to interact with the virus matrix (M) protein layer on the internal surface of the membrane. The helical nucleocapsid is located within the M-protein layer, and includes the 13- to 15-kb single-stranded, non-segmented RNA genome, together with the nucleocapsid (N), phosphoprotein (P), and large (L) proteins (Wunner and Pringle, 1976; Huang et al., 1985). The pneumovirus nucleocapsids (diameter, 13 to 14 nm) are significantly smaller than those that have been described for other paramyxoviruses (18 nm) (Berthiaume et al., 1974). By analogy with the other members of the Paramyxoviridae, the L protein is considered to be the major component of the RNA dependent RNA polymerase which directs synthesis of virus genomic RNA and mRNA. The L protein is also thought to contain some other enzymatic activities including those of protein kinase, mRNA capping, methylation and polyadenylation. The M2-1 transcriptional enhancer protein is also thought to be associated with the nucleocapsid, although this has not yet been determined directly.

pneumovirus genome organization: The genomes of the pneumoviruses are single stranded negative sense RNA ~15,000 nucleotides in length. While the precise sizes of the genomes are strain dependent, the number of nucleotides does not follow the “rule of six” (Bukreyev et al., 2000; Calain and Roux, 1993; Samal and Collins, 1996).

Although, the genomes of the members of the family Paramyxoviridae are approximately the same length, the pneumoviruses encode a larger number of mRNAs (RSV and PVM encode ten and APV encodes eight mRNAs) compared with the six or seven for the other paramyxoviruses. This is achieved mainly by pneumoviruses encoding shorter mRNAs and hence smaller polypeptides, than those of other members of the family. A schematic diagram of the genome organisation of representative members of the Paramyxoviridae is shown in figure 2. The mammalian pneumoviruses direct the synthesis of three extra mRNAs encoding the non-structural NS1 and NS2 proteins and 22K protein. APV and human metapneumovirus (hMPV) does not encode genes equivalents of the NS1 and NS2 proteins. While all members of the family Paramyxoviridae encode a standard set of structural proteins (Pringle and Easton, 1997), one feature which differentiates the Paramyxovirinae from the Pneumovirinae are different coding potentials of P genes and the use of an RNA editing mechanism to express multiple proteins from this gene (Curran and Kolakofsky, 1999). Furthermore, the Pneumovirinae lack counterparts of the C and V proteins found in Paramyxovirinae, although the PVM P gene does direct the synthesis of a polypeptide by using a second open reading frame in
a manner analogous to the expression of the paramyxovirus C proteins (Barr et al., 1994). Another unique feature the Pneumovirinae is the gene encoding the M2 proteins which has not been identified in any other virus genome.

The 3’ end of the genome, referred to as the leader region is the point at which the virus polymerase binds to initiate both replication and transcription. The leader region of non-segmented negative strand viruses is thought to contain at least two signals, the first is vRNA sense and is promoter for the synthesis of positive sense RNA, the second is in the nascent positive sense transcript and is involved in initiating its encapsidation (Blumberg et al., 1983). Each virus gene is located between these two regulatory elements. Each RSV gene is initiated by a conserved nine nucleotide sequence called the gene start signal and transcription of each gene starts at the first nucleotide of this sequence. At the end of each gene there is a gene end signal, directing polyadenylation and transcription termination activity which, in RSV, shows significant sequence variation compared to the highly conserved gene start sequence. In PVM, the gene end and gene start sequences are well conserved; however they are not identical to those of RSV (Table 1).

<table>
<thead>
<tr>
<th>Virus</th>
<th>Gene start</th>
<th>Gene end</th>
</tr>
</thead>
<tbody>
<tr>
<td>hRSV</td>
<td>GAG (g/a) CAE (u/a)</td>
<td>uAGTsAn</td>
</tr>
<tr>
<td>PVM</td>
<td>AGGA (c/u) aA (a/g) u</td>
<td>UA g UUA uu</td>
</tr>
<tr>
<td>APN</td>
<td>GGGAGCAAGU</td>
<td>uaguaauuu</td>
</tr>
</tbody>
</table>

Table 1. Consensus gene start and gene end signals of pneumoviruses shown in mRNA sense. Letters in upper case indicate compete conservation and lower case indicates preferred bases.

All of the genes of pneumoviruses are separated by short intergenic regions that are not transcribed into mRNA and in contrast to the gene start and gene end sequences; do not show any sequence homology. The intergenic regions in the RSV genome vary in length from 1 to 52 nucleotides and lack any obvious con-
served feature of primary and secondary structure other than that each sequence ends with an adenosine residue in the genome RNA (Collins et al., 1986). The lack of common sequences in the intergenic regions suggests these sequences do not contain putative signals for transcription events such as capping or polyadenylation. It is likely that the viral polymerase translocates, without synthesis, across the intergenic regions (Collins et al., 1987).

**Pneumovirus Gene Products:** As indicated in figure 2, the nucleocapsid, phosphoprotein, matrix, fusion, and polymerase proteins are common to all members of the family Paramyxoviridae. Much of our understanding of the functions of these proteins has been derived by extrapolation from experiments performed with one or more specific viruses, particularly Sendai virus. More recently, it has been possible to confirm these functions directly by targeted mutagenesis and reverse genetics.

*Nonstructural (NS1 and NS2) proteins:* The 3′ proximal genes of pneumoviruses (in genome orientation) encode two short, nonstructural proteins (NS1 and NS2) that are missing in other paramyxoviruses. The hRSV NS1 protein interacts with the M protein in infected cells, and two-hybrid analysis identified an additional interaction between NS1 and the P protein (Evans et al., 1996; Hengst and Kiefer, 2000). The hRSV NS1 protein was identified as a regulatory protein since it inhibited both transcription and replication in minireplicon systems (Atreya et al., 1998). Deletion of the NS1 gene resulted in virus that was attenuated in chimpanzees (Teng et al., 2000), confirming the role of NS1 in virus growth. Deletion or insertion of a stop codon in the RSV NS2 gene resulted in recombinant viruses that were attenuated in cell culture (Teng and Collins, 1999) and in chimpanzees (Whitehead et al., 1999a). Passage of the NS2 stop mutant in cell culture yielded revertant viruses whose artificial NS2 stop codon had been replaced with a sense codon, thus restoring expression of a (mutant) NS2 protein (Teng and Collins, 1999). This finding demonstrated that, although not essential, NS2 confers a growth advantage to hRSV replication in cell culture. Schlender et al. (2000) generated bRSV mutants that do not express NS1 and/or NS2. These viruses were highly attenuated in MDBK cells but only moderately attenuated in IFN-deficient Vero cells, suggesting that NS1 and NS2 antagonize the IFN-mediated antiviral response. For bRSV, both the NS1 and NS2 proteins are required to control the IFN-α/β-mediated cellular response (Schlender et al., 2000). Although hRSV continues to replicate in cells pretreated with type I IFN (Atreya and Kulkarni, 1999), it does not block either type I or type II IFN signaling (Young et al., 2000). These findings suggest that hRSV has developed an alternative strategy to counteract the cellular IFN response, one that most likely interferes with events downstream of IFN signaling.

**Nucleocapsid (N) protein:** The N protein forms an integral part of the nucleocapsid complex of the virion and is an essential component of the polymerase complex. The N protein is thought to be responsible for giving the RNA genome its helical structure through a tight association with the virus genome RNA; interestingly, hRSV N protein expressed in insect cells spontaneously formed nucleocapsid structures containing RNA (Bisgaard, H., and Study Group on Montelukast and Respiratory Syncytial Virus, 2003). Deletion mutant analysis of N protein expressed in cells suggests that a large segment of this protein is required for interacting with the P protein (Garcia-Barreno et al., 1996; Stokes et al., 2003).

**M2 proteins:** Expression of the hRSV N, P and L proteins, together with a plasmid-encoded minigenome, resulted in premature termination of mRNA synthesis (Collins et al., 1996a). Coexpression of low levels of the M2 gene yielded full-length mRNAs and polycistronic read-through transcripts, whereas higher amounts of the M2 gene inhibited replication and transcription (Collins et al., 1996a). Similarly, small amounts of the M2 gene enhanced transcription of a bRSV minigenome (Yunus et al., 1998). Further study of the M2 gene, which contains two overlapping ORFs (Fig. 3). The nucleotide sequence of the M2 gene of pneumonia virus of mice (PVM) show that the gene encoded a protein of 176 amino acids from a major ORF, which is smaller than the equivalent proteins encoded by human, bovine and ovine RSV. In common with the M2 genes of the RS viruses and avian pneumovirus (APV), the PVM mRNA also contained a second ORF (ORF2) that partially overlaps the first ORF and which is capable of encoding a 98 residue polypeptide (Ahmadian et al., 1999). The product of the first ORF, the M2-1 protein, is required to synthesize full-length mRNAs and read-through transcripts (Collins et al., 1996a; Fearns and Collins, 1999b) by functioning as an elongation factor and an antitermination factor (Collins et al., 1996a; Hardy and Wertz, 1998; Fearns and Collins, 1999b; Hardy et al., 1999). In contrast, the M2-2 protein, from the second ORF, inhibits RNA replication (Collins et al., 1996a; Jin et al., 2000a) and may play a role in the switch from tran-
scription to replication (Bermingham and Collins, 1999). Consequently, recombinant RSVs that lost the ability to express the M2-2 protein were restricted in their growth in cell culture and in animal models (Bermingham and Collins, 1999; Jin et al., 2000a; Teng et al., 2000).

Translation of ORF2 of hRSV M2 gene initiates at the one of the three initiation codons located upstream of the termination codon for the first ORF. Mutagenesis studies have shown that the location of the termination codon of ORF1 protein play an important role in directing translation of ORF2 from upstream initiation codons in vivo. This indicates that the second ORF is accessed by the ribosomes that are departing from the first ORF and that these ribosomes reinitiate on one of the AUG codons 5' to the point of translation termination (Ahmadian et al., 2000). The kinetics of synthesis of the M2-2 protein in RSV infected cells have been analysed by Western blot analysis. The RSV M2 ORF2 protein was initially detected at 24h post-infection and continued to accumulate throughout the infectious cycle to 48h. The increasing amount of this protein over a long time during infection suggests that it is stably accumulated in the infected cells and may indicate that this protein is needed in the late stages of the virus life-cycle (Ahmadian et al., 1999). The ability to recover HMPV lacking M2-1 contrasts with human respiratory syncytial virus, for which M2-1 is an essential transcription factor. Meanwhile, expression of the downstream HMPV M2-2 ORF was not reduced when translation of the upstream M2-1 ORF was silenced, indicating that it is initiated separately. The r∆M2-2 mutants exhibited a two- to fivefold increase in the accumulation of mRNA, normalized to the genome template, suggesting that M2-2 has a role in regulating RNA synthesis and M2-1 appears to be essential for significant virus replication in vivo (Buchholz et al., 2005).

**SH proteins:** The hRSV SH protein is a short transmembrane protein of unknown function. Deletion of SH in a recombinant RSV did not affect RNA replication but slightly increased virus titres in certain cell lines (Bukreyev et al., 1997; Techaarpornkul et al., 2001) and caused moderate attenuation in animal models (Bukreyev et al., 1997; Whitehead et al., 1999a). Deletion of the SH gene in the distantly related a paramyxovirus SV5 did not significantly alter the growth characteristics of a recombinant virus in cell culture (He et al., 1998). However, the SV5∆SH virus induced greater DNA fragmentation and higher caspase-2 and caspase-3 activities than did wild type virus, indicating that the SH protein interferes with apoptosis (He et al., 2001). In one study, expression of the RSV M2-2, NS1, NS2 and SH proteins was abrogated in various combinations (Jin et al., 2000b). Virus in which all four genes were deleted could not be generated and deletion of the M2-2 gene together with the NS1 gene also proved detrimental to virus replication. However, recombinant viruses, such as RSV∆SH, ∆NS1 and ∆NS2, were viable, although they were attenuated in cell culture and in cotton rats. Replication of recombinant viruses was more affected in HEp-2 than in Vero cells, suggesting that one or more of the deleted genes encode proteins that counteract the cellular IFN response. Taken together, these

---

**Figure 3.** Diagrammatic representation of the organization of the pneumovirus M2-1 and M2-2 genes. The grey bar represents the M2-1 open reading frame, and the white bar represents the M2-2 open reading frame. The positions of the translation initiation and termination codons of are identified. aa, amino acids.
findings indicate that although the RSV M2-2, NS1, NS2 and SH proteins are not essential for virus replication, they encode supporting functions required for efficient virus replication. hMPV ΔSH, ΔG, and ΔSH/G have been readily recovered and have been found to replicate efficiently during multicyle growth in cell culture. Thus, the SH and G proteins are not essential for growth in cell culture. When administered intranasally to hamsters, the ΔG and ΔSH/G mutants replicated in both the upper and lower respiratory tracts, showing that HMPV containing F as the sole viral surface protein is competent for replication in vivo (Biacchesi et al., 2004).

Phosphoprotein (P): The P protein is an essential component of the replication and transcription complexes of the pneumoviruses (Barr et al., 1994; Lambden, 1985; Li et al., 1996; Satake et al., 1984; van den Hoogen et al., 2002). Analysis of the interaction between the P and N proteins has shown that the carboxy terminus of the P protein contains most of the elements necessary to bind to the N protein (Garcia-Barreno et al., 1996; Hengst and Kiefer, 2000; Khattar et al., 2001; Lu et al., 2002; Malik Peiris et al., 2003; Slack and Easton, 1998). The P-protein genes of the Pneumovirinae encode several different proteins using alternative start codons. In addition to the full-length sequence, the P gene of hRSV encodes an mRNA for an additional, short polypeptide of unknown function (Caravokyri et al., 1992). The P protein is phosphorylated at specific serine residues (Dupuy et al., 1999). Villanueva et al. (2000) demonstrated that most of the phosphorylation, occurring at residues 116, 117, 119, and 232, is not essential for transcription or replication.

Matrix (M) protein: The M protein is associated with the inner face of the lipid membrane of infected target cells (Peeples and Levine, 1997; Wunner and Pringle, 1976). All M proteins include a hydrophobic carboxy-terminal domain. The M proteins of nonsegmented minus-strand viruses are thought to function by rendering the nucleocapsid transcriptionally inactive before packaging and also by promoting the association of the nucleocapsid complex with the nascent envelope (Ghildyal et al., 2002; Teng and Collins, 1998).

Attachment (G) protein: The hRSV G protein was identified as the viral attachment protein based on the observation that G-specific polyclonal antibody blocked absorption of the virus to the surface of target cells (Levine et al., 1987). The G proteins of pneumoviruses are heavily glycosylated with both N- and O-glycosylated sites, adding approximately 55 kDa to the mass of the hRSV polypeptide. While the roles of these glycosyl groups have not been completely clarified, extensive glycosylation of the extracellular domain of the G protein could reduce its antigenicity by shielding the virus protein with host specified sugars. Alternatively, partial utilization of glycosylation sites can result in antigenic heterogeneity, representing a source of diversity within a genetically homogenous virus population. The immune response to hRSV correlates with and provides selection pressure for G-protein variation (Garcia-Barreno et al., 1994; Melero et al., 1997). The hRSV G protein in infected cells is present in two forms. The first is translated from the entire open reading frame and contains an amino-terminal sequence prior to the putative signal/membrane anchor region. A second, secreted, form of the protein has been observed (Roberts et al., 1994) which is generated by the initiation of translation at an internal AUG initiation codon that is in the same reading frame as that used for the full-length G protein. The function of the secreted protein is not known, but it may alter the immune response of the host, possibly by acting as a “decoy” or by altering specific aspects of the host inflammatory response (Johnson and Graham, 1999; Johnson et al., 1998).

Fusion (F) protein: The F protein of hRSV was first identified by immunoprecipitation using monoclonal antibodies which inhibit the formation of multicell syncytia in cell culture. The F protein of hRSV is synthesized as an inactive precursor (F0) that assembles into a homo-oligomer in the rough endoplasmic reticulum (ER), forming a structure that is presumed to represent a single virion spike (Collins et al., 1995). When the F0 protein reaches the trans Golgi network, it is activated by cleavage into two disulfide-linked subunits, F1 and F2, at two furin consensus sites (Gonzalez-Reyes et al., 2001; Zimmer et al., 2002; Zimmer et al., 2001).

Large (L) polymerase protein: By analogy to other viruses, the pneumovirus L protein is thought to be the major component of the viral RNA-dependent RNA polymerase complex, which is responsible for the synthesis of all viral RNA, including mRNA, replicative intermediates, and the progeny RNA genomes. The L protein is also thought to be responsible for mRNA methylation and capping, although this has not been shown directly. As the name indicates, the L protein is very large—the hRSV L protein contains 2,165 amino acids, and the APV and hMPV L proteins 2,000 amino
products of viral transcription (Collins, 1986). A viral transcriptional map was deduced from UV inactivation of gene transcription, the viral genome appears to be transcribed sequentially from a single promoter in the leader region (Dickens, 1984). There have been few direct studies of the specifics of the L proteins of the Pneumovirinae. Most recently, Cartee et al. (2003) demonstrated that a single amino acid change (N1049D) resulted in aberrant transcriptional termination without altering the rate of virus replication.

Transcription and replication of the pneumovirus genome: The details of the current model of transcription and replication of the pneumovirus genomes are consistent with those proposed for all nonsegmented negative-sense RNA viruses. Based on the kinetics of UV inactivation of gene transcription, the viral genome appears to be transcribed sequentially from a single promoter in the leader region (Dickens et al., 1984). A viral transcriptional map was deduced from UV mapping studies and from analysis of the relationships between the 10 mRNAs and a number of polycistronic readthrough mRNAs identified as minor products of viral transcription (Collins et al., 1984 and 1986). According to this model, the viral polymerase recognises a sequence in the 3′ leader region of the genome RNA and begins transcription at the first nucleotide. A short RNA transcript representing the leader region is released. No evidence for the existence of the RSV leader RNA has been presented to date. Transcription is then re-initiated at this point on the conserved gene start sequence and continues to the conserved termination signal when a poly A tail is added to the mRNA. At some point during this process, the mRNA is capped to ensure translation on host cell ribosomes; there are no reports directly addressing this process. The signal marking the end of transcription among the pneumoviruses contains a short run of U residues, and this marks the position at which the polyadenylated tail is added to the mRNA. The mRNA then dissociates from the transcription complex (Barik, 1993). The RNA polymerase now either dissociates from the template or, moving over the intergenic region, reinitiates at the beginning of the next gene for the process to occur again. If the polymerase detaches from the genome it may only reinitiate at the 3′ leader region. As a result of this mode of transcription the 3′ proximal genes are transcribed into mRNA abundantly and there will be a progressive attenuation in the abundance of viral gene transcripts in the 3′ to 5′ direction (Krempl et al., 2002). This process is called sequential transcription. The relative abundances of the mRNAs will be reflected in the abundances of the respective proteins. For nonsegmented negative strand viruses in general, attenuation of sequential transcription at the intergenic regions appear to be a major factor in determining the intracellular molar ratios of the different mRNAs and proteins and variation in the intergenic regions might be a mechanism for obtaining different degrees of attenuation (Bukreyev et al., 2000; Chambers et al., 1990).

The process of termination and reinitiation is poorly understood, but the M2-1 protein of RSV has been shown to be essential for production of full-length virus mRNA (Collins et al., 1995 and 1996a; Fears and Collins, 1999b; Hardy and Wertz, 2000). The antitermination activity during transcription permits the viral polymerase to remain associated with the template, thereby increasing the production of full-length mRNA and enhancing the production of products of genes that are distal to the 3′ leader region. The eight RSV gene junctions, which vary in their ability to terminate transcription, all direct the production of more read through mRNA in the presence of M2-1 protein, although they vary in their sensitivity to the presence of the M2-1 protein (Hardy et al., 1999).

Elements within intergenic regions play a vital role in the regulation of transcription (Fears et al., 2002). The intergenic regions of the Pneumovirinae show considerable sequence diversity and vary from 1 to 56 nucleotides in length. The transcriptional start sequences are thought to be 10 nucleotides in length, and the consensus sequences, while conserved, are slightly different for each pneumovirus (Collins et al., 1986; Kuo et al., 1997). The PVM transcriptional start signals are more variable than those of RSV or APV, which are absolutely conserved with the exception of the L gene (Chambers et al., 1990; Randhawa et al., 1997; Stec et al., 1991).

The transcriptional stop signals of hRSV consist of a conserved pentanucleotide sequence together with a 1- to 4-nucleotide AU-rich region and a 4- to 7-nucleotide poly (U) tract. Saturation mutagenesis of the hRSV gene transcriptional start sequences showed that residues 1, 3, 6, 7, and 9 were critical in directing transcription initiation, although there was some variability in efficiency, particularly with the NS1 and NS2 gene start sequences, which were approximately 40% less efficient than the others (Kuo et al., 1996a, 1996b and 1997).

Deletion of a transcriptional stop signal from an
upstream gene resulted in transcriptional read through to the next gene without termination. Harmon et al. (2001) analyzed the hRSV M-gene transcriptional stop sequence and determined that the integrity of the pentanucleotide and AU-rich regions was essential for efficient termination. The residue following the poly(U) tract might also be important for efficient transcription termination (Sutherland et al., 2001).

No significant difference in transcriptional activity or efficiency was seen with different sized RSV intergenic sequences (Hardy et al., 1999; Kuo et al., 1996b). Bukreyev et al. (2000) described the recovery of a recombinant hRSV containing an intergenic region of 160 nucleotides. This very long intergenic region had little effect on sequential transcription.

The genome structure of hRSV contains an exception to the process described above. The gene end signal for the M2 gene is located 68 nucleotides downstream of the gene start sequence for the L gene (Collins et al., 1987), and much of the mRNA that initiates at the L transcriptional start site terminates at the M2 gene end sequence (Collins et al., 1987; Fearns and Collins, 1999a). The role (if any) of this truncated L transcript remains to be elucidated.

RNA replication involves a switch from the stop-start model of transcription to an antitermination read-through mode which results in synthesis of full-length replicative intermediate (RI) and this in turn serves as template for synthesis of vRNA. Transcription and replication promoters of hRSV overlap but are not identical (Fearns et al., 2002). In the replicative mode, the polymerase binds to the 3′ end of the genome RNA and initiates RNA synthesis de novo. Once initiated, the polymerase is committed to continue to the end of the template to produce an antigenome. The polymerase complex then uses the antigenome as a template, binding to and initiating RNA synthesis, at the 3′ end of the antigenome RNA. In the genome sense RNA, the trailer region contains the antisense copy of the sequence necessary for initiating replication to produce more genome RNA from the antigenome.

Although the balance between the synthesis of mRNA versus antigenome is thought to be a central event in the growth cycle of these viruses it is not well understood. An important distinction between transcription and replication is that the latter is coupled with encapsidation of nascent RNA, though the exact nature of this is not understood. According to the simplest model, the polymerase in its default state is a transcriptase that is switched to replication by the rapid binding of the N protein to the leader RNA at the 5′ end of the nascent positive sense antigenome. The general model for the switch from transcription to replication, developed for VSV (Kingsbury, 1974) is dependent on the encapsidation of the nascent chain by the N protein (Vidal and Kolakofsky, 1989). This is believed to inhibit the polymerase recognising the gene end and gene start signals and thus forcing the synthesis of the full-length antigenome RNA molecule. Fearns et al. (1997), by expressing increasing amounts on the N protein in a minigenome model with or without an equivalent increase in levels of the P protein, found an increase in the production of the replicative form of the RSV genome. However, this did not cause any decrease in mRNA production. Instead, the ratio of mRNA synthesis to antigenome synthesis remained unaffected and the production of viral mRNAs was the predominant event. More recently, a role for M2-2 protein has been proposed (Bermingham and Collins, 1999).

The 3′ and 5′ termini of the pneumovirus genome contain the sequences that direct replication. Reverse genetics studies showed that the 3′-terminal 44-nucleotide leader region and the 5′-terminal 40 nucleotides of the trailer region of hRSV were necessary for replication, encapsidation, and assembly (Collins et al., 1991). The immediate termini of the pneumovirus genomes have complementary sequences, as has been described for other negative-sense RNA viruses (Mink et al., 1991; Randhawa et al., 1997; van den Hoogen et al., 2002). The trailer sequences are more efficient than the leader sequences at directing replication, as would be expected from their role in replication and the requirement for the leader region to direct both replication and transcription (Fearns et al., 2002). Analysis of the leader and trailer sequences suggested that nucleotide positions 1, 2, 3, 5, 6, and 7 were particularly important for replication, while position 4 was found to be tolerant of alteration (Peeples and Collins, 2000). Using synthetic chimeric minigenomes with terminal sequences from both APV and RSV in virus infected cells, Marriott et al. (2001) showed that paired termini from the same virus were required for replication to occur. It is not clear whether the critical step is replication or encapsidation, but the data suggest that an association between the leader and trailer is necessary for productive infection (Marriott et al., 2001).

**Packaging and Assembly:** Reverse genetics experiments suggest that the accumulation of the M protein may be a component of the trigger to initiate the assembly of virions (Teng and Collins, 1998). It is anticipated that the nucleocapsid complex/M-protein structure associates with the internal tails of the major glycoproteins that are inserted into the membrane of
the infected cell after processing through the Golgi complex. In parallel with studies performed with other viruses, a progressive series of interactions between the glycoproteins and the M protein, which acts as a bridge with the nucleocapsids, results in the progeny virions budding from the surface of the cell in the form of filament structures (Parry et al., 1979). In polarized cells, such as those of the respiratory tract, hRSV buds from the apical surface and is shed into the airways, from which it can be transmitted to new hosts (Roberts et al., 1995).

**Reverse genetics:** Reverse genetics provides a powerful means for analysis of the role of cis- and trans-acting elements in the control of RNA replication and transcription and for recovery of infectious virus from the cDNA clone in order to study the roles of individual gene products in viral replication and pathogenesis. Employment of reverse genetics is important for development of live attenuated vaccine by introduce attenuating mutations into infectious recombinant virus.

The production of recombinant RSV requires the intracellular coexpression, from transfected plasmids, of antigenomic RNA and the N, P, L and M2-1 support proteins (Collins et al., 1999a; Collins et al., 1995). The requirement for M2-1 as a fourth support protein distinguished RSV from the paramyxoviruses (Fig. 4) (Collins et al., 1999a).

The second type of reverse genetics system involves minireplicons, which are short cDNA encoding version of genomic or anti-genomic RNA in which some or all of the viral genes have been deleted and replaced with none or more genes encoding marker proteins such as bacterial chloramphenicol acetyl transferase (CAT), luciferase and/or green fluorescent protein (GFP) (Fig. 5). Minireplicons usually do not contain the genes necessary for their own replication and must be complemented by viral protein provided by cotransfected plasmids or by coinfecting standard, helper, RSV. They are useful, because it is easier to perform extensive mutational and biochemical analyses with a short replicon containing only one or a few genes instead of the complete genome. The fact that the minireplicon and support proteins are encoded by separate plasmids also means that each component can be manipulated separately without affecting the synthesis of the others.

**Figure 4.** Schematic diagram for the generation of nonsegmented negative-sense RSV viruses. Cells are cotransfected with protein expression plasmids for the N, P, L and M2-1 proteins and with a plasmid containing a full-length viral cDNA, all under the control of the T7 RNA polymerase promoter. Following infection with recombinant VV encoding T7 RNA polymerase, vRNA is synthesized and the virus replication cycle is initiated.

**Figure 5.** Schematic representation of a minigenome construct. A CAT or GFP reporter gene is flanked by the viral leader sequence (Le) and gene start signal (GS) of the nucleocapsid gene on one end and the gene end signal (GE) of the polymerase gene and the viral trailer (Tr) on the other end. This minigenome is cloned in the context of a T7 RNA polymerase promoter-terminator (P-T7, T-T7) cassette. The authenticity of the transcribed ends of the negative-sense minigenome are determined by the position of the T7 promoter sequence and a HDV ribozyme sequence (δ).

The recovery of recombinant virus entirely from cDNA is better by using a plasmid directing transcription of antigenome (positive strand) RNA, rather than genome RNA. If naked negative strand RNA genomes are produced in the cytoplasm of cells that also producing complementary mRNAs encoding viral proteins, the two can be hybridize and prevent the critical assembly of the genome into the RNP. In starting with an anti-genome RNP, only one successful round of replication driven by the plasmid encoded support proteins is required to produce an infectious genome RNP.

The replacement of helper virus by protein expressed from transfected plasmids with T7 promoters in cells infected with a recombinant vaccinia virus encoding bacteriophage T7 RNA polymerase (Fuerst and Niles, 1986) markedly enhanced recovery rate and allowed for investigation of viral trans-acting factors required for replication, assembly, and budding of virus like particles (Pattnaik and Wertz, 1991; Curran and Kolakofsky, 1991).
The anti-genome ribozyme sequence of hepatitis delta virus (HDV) has generally been used to much more efficient production of virus. This has the advantage of requiring only sequences downstream of the cleavage site for autocatalytic activity (Perrotta and Been, 1990). RNAs ending with correct 3′ nucleotide can be generated by autocatalytic cleavage from primary transcripts containing the HDV ribozyme sequence immediately downstream of viral sequences. According to the ribozyme cleavage mechanism, 3′ terminal ribose of the (upstream) genome analog should possess a cyclic 2′-3′ phosphate instead of hydroxyl group (Schneider-Schaulies et al., 1990). This modification might contribute to the success of the approach, in preventing polyadenylation of the RNA or in delaying degradation of the RNA 3′ terminus.

**Applications of reverse genetics in vaccine strategies:** Reverse genetics provides a powerful means to define the genetic basis of attenuation of existing biologically derived RSV and parainfluenza virus (PIV) vaccine candidates that had been attenuated by conventional methods, such as cold-passage or chemical mutagenesis (Collins et al., 1999b; Skiadopoulos et al., 1999). As a first step, the complete consensus nucleotide sequence must be determined for each virus in question and compared with that of its wild-type parent. The identified mutations or sets of mutations must then be inserted into wild-type recombinant virus via reverse genetics to determine whether they confer an attenuation phenotype (Collins et al., 1999b; Skiadopoulos et al., 1999).

The second source of attenuating mutations is the generation of novel mutations not found among existing virus mutants or vaccine candidates. RSV and the PIVs encode several nonessential proteins whose major function is to facilitate virus replication in vitro, for instance, by antagonizing host defense factors. Deleting or silencing such genes or ORFs is attractive because the resulting virus frequently retains replicative efficiency in cell culture. If these viral accessory factors enhance pathogenesis or down regulate host immune responses, their ablation might improve the safety and efficacy of the resulting vaccine (Collins et al., 1999b). Such mutations can be made by deleting a complete transcriptional unit encoding an mRNA and protein, such as has been done for the NS1, NS2, and SH genes of RSV, or by deleting or silencing a single ORF contained within a complex mRNA encoding more than one protein, as has been done to eliminate expression of the M2-2 protein of RSV and the C, D, and V proteins of PIV3 (Durbin et al., 1999; Nagai and Kato 1999). Deletion of a complete gene is a desirable mutation because it would be expected to be more genetically and phenotypically stable than point mutations.

Another method of attenuation is based on bRSV, a closely related animal counterpart of human RSV that is attenuated in primates due to a natural host range restriction (Buchholz et al., 2000). bRSV itself is over-attenuated and insufficiently immunogenic in seronegative chimpanzees (Buchholz et al., 2000). Growth was improved by replacing the G and F genes of bRSV with their human RSV counterparts, although the resulting chimeric virus remained over-attenuated (Buchholz et al., 2000). The systematic replacement of additional bRSV genes with their human RSV counterparts will yield viruses with a range of growth properties that can be evaluated as vaccines.

Additional means of attenuation of recombinant RSV can be devised. To date, a number of attenuating mutations have been identified by trial and error, including point mutations in the L protein, a point mutation in the leader region, and C-terminal deletions of the M2-1 protein (Tang et al., 2001). Based on studies with vesicular stomatitis virus (VSV), attenuation can be achieved by changing the gene order (Wertz et al., 1998). This strategy could be applied to RSV provided that it does not reduce the efficiency of growth in vitro, which otherwise could interfere with vaccine production. An attenuating mutation identified in one mononegavirus sometimes can be transferred to another, particularly if it involves a conserved residue. For example, an attenuating point mutation in the Sendai virus C protein was transferred to the corresponding position in the C protein of recombinant human parainfluenza virus type 3 (hPIV3), resulting in attenuation of the HPIV3 recipient (Durbin et al., 1999).

Reverse genetics was used to expedite development of an RSV subgroup B vaccine candidate, which faced the obstacles of a lack both of promising vaccine candidates and of a recombinant recovery system. Replacement of the G and F genes of recombinant RSV A2 (subgroup A) with their counterparts of strain B1 of subgroup B produced a wild-type-like chimeric AB virus bearing the protective antigens of subgroup B (Whitehead et al., 1999b). Attenuated AB viruses were readily produced using attenuated versions of the A2 backbone and are in preparation for clinical evaluation. Alternatively, the subgroup B G gene, representing the more divergent protective antigen, can be expressed as an added gene from a recombinant subgroup A virus (Jin et al., 1998).

In animals infected with r∆M2-2, virus has been recovered from only 1 of 12 animals and only in the
nasal turbinates on a single day. However, all of the animals have developed a high titer of hMPV-neutralizing serum antibodies and have been highly protected against challenge with wild-type hMPV. The hMPV rΔM2-2 virus is a promising and highly attenuated hMPV vaccine candidate (Buchholz et al., 2005). However, both viruses have been at least 40-fold and 600-fold restricted in replication in the lower and upper respiratory tract, respectively, compared to wild-type rhMPV. They also induced high titers of hMPV-neutralizing serum antibodies and conferred complete protection against replication of wild-type hMPV challenge virus in the lungs. Surprisingly, G is dispensable for protection, and the ΔG and ΔSH/G viruses represent promising vaccine candidates. In contrast, ΔSH replicated somewhat more efficiently in hamster lungs compared to wild-type rhMPV (20-fold increase on day 5 post infection). This indicates that SH is completely dispensable in vivo and that its deletion does not confer an attenuating effect (Biacchesi et al., 2004).

References


Chambers P, Matthews DA, Pringle CR, Easton AJ. (1990). The nucleotide sequences of intergenic regions between...
nine genes of pneumonia virus of mice establish the physical order of these genes in the viral genome. Virus Res. 18: 263-270.


Gonzalez-Reyes L, Ruiz-Arguello MB, Garcia-Barreno B, Calder L, Lopez JA, Albar JP, Skehel JJ, Wiley DC,
Ahmadian et al.


Melero JA, Garcia-Barreno B, Martinez L, Pringle CR, Cane PA (1997). Antigenic structure, evolution and

Teng MN, Whitehead SS, Bermingham A, St Claire M, Elkins WR, Murphy BR, Collins PL (2000). Recombinant respiratory syncytial virus that does not express the NS1 or M2-2 protein is highly attenuated and immunogenic in chimpanzees. *J Virol.* 74: 9317-9321.


