Current Topics in Microbiology and Immunology

Larry J. Anderson
Barney S. Graham *Editors*

Challenges and Opportunities for Respiratory Syncytial Virus Vaccines

Springer

Current Topics in Microbiology and Immunology

Volume 372

For further volumes:
http://www.springer.com/series/82

Larry J. Anderson · Barney S. Graham
Editors

Challenges and Opportunities for Respiratory Syncytial Virus Vaccines

Responsible series editor: Richard W. Compans

Springer

Editors
Larry J. Anderson
Department of Pediatrics and Children's Healthcare of Atlanta
Emory University
Atlanta, GA
USA

Barney S. Graham
Vaccine Research Center
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, MD
USA

ISSN 0070-217X ISSN 2196-9965 (electronic)
ISBN 978-3-642-38918-4 ISBN 978-3-642-38919-1 (eBook)
DOI 10.1007/978-3-642-38919-1
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013955565

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Respiratory syncytial virus (RSV) was first isolated from chimpanzees in 1955, and shortly thereafter from young children, and recognized as an important cause of serious lower respiratory tract disease in infants and young children, i.e., pneumonia and bronchiolitis. Shortly after its discovery efforts to develop a vaccine began. It is now over 50 years since its discovery and no safe and effective vaccine is yet available. The fact that the peak of disease occurs in young infants despite the presence of maternally acquired antibody and that humans have multiple disease-associated infections througout life, foretell the challenge in inducing protective immunity with a vaccine. The first RSV vaccine, formalin-inactivated tissue culture grown virus (FI-RSV) formulated with alum, was ineffective, associated with enhanced respiratory disease, and raised concerns that other vaccines might also predispose to FI-RSV vaccine enhanced respiratory disease. Older children did not experience FI-RSV vaccine enhanced respiratory disease suggesting that prior infection establishes a safe immune response pattern. Therefore, development of RSV vaccines for the RSV naïve infant has focused on live-attenuated RSV or other live viruses expressing RSV genes, while other types of vaccines are being developed for older children and adults. Subsequent to the FI-RSV vaccine trial, a number of live-attenuated RSV strains, live virus and other gene expression systems, and protein subunit vaccines in different platforms have been developed and tested in animals and a few have also being studied in humans. No vaccine has yet shown sufficient promise to move toward licensure. However, the prospects for many candidate vaccines, as well as different vaccine platforms, will remain unknown until clinical trials in the selected target population for the vaccine are performed. Though the lack of success to date highlights the biological difficulties in developing an RSV vaccine, the efficacy of immune prophylaxis suggests a safe and effective vaccine is achievable. The availability of ever more powerful tools to study the immune response and pathogenesis of disease and ability to construct a wide variety of vaccines using different vaccine platforms suggest that an RSV vaccine should be within reach. Caren Hall who made so many contributions to our understanding of RSV wisely noted in a poem one of the ongoing challenges to achieving an RSV vaccine (Anderson and Heilman 1995).

Immunity and RSV
What is this thing we call immunity?
Does it exist for ills from RSV?
Perhaps for mice within their splenic soul,
But in the babe, the old, what is its role?
Is there a pattern of response we can discern?
Or is it like night skies that change with each earth's turn,
With season, angle viewed, and light years passed?
Has each from different, fluid molds been cast?
Thus, is it only solitary stars we see.
And not a constellation called immunity?

Caroline Breese Hall

We feel this book brings together in one place what we know about RSV and helps to organize the constellation of facts about the virus and host factors that can guide successful development of an RSV vaccine. We feel that a better understanding of the of clinical and epidemiologic features of infection, functional and structural features of the virus, pathogenesis of the associated disease, and the immune response it induces provide the underpinnings for success.

Finally, we mourn the passing of Caren Hall. She has been the queen of RSV and authored most of the seminal studies on the clinical and epidemiologic features of infection and transmission of the virus. Equally important, she has been

Caroline Breese Hall (center) teaching and mentoring

a wonderful friend to so many in the field and an example to all of an outstanding scientist with a kind and nurturing spirit. In recognition of what she accomplished and the person she was, we dedicate this book to her. We refer the reader to the obituaries published in the Journal of Pediatric Infectious Disease Society in May 2013 Plotkin et al. (2013); Caserta and Long (2013); Englund et al. (2013). These three tributes provide a glimpse into why we felt so honored to have known and worked with her.

Larry J. Anderson
Barney S. Graham

References

Anderson LJ, Heilman CA (1995) Protective and disease-enhancing immune responses to respiratory syncytial virus. J Infect Dis 171:1–7

Caserta M, Long C (2013) Caroline Breese Hall Obituary. J Ped Inf Dis Soc 2:2–3

Englund J, Karron R, Dennehy P, Kimberlin D (2013) Caroline Breese Hall Obituary. J Ped Inf Dis Soc 2:3–4

Plotkin SA, Peter G, Katz (2013) Tribute to Caroline Hall. J Ped Inf Dis Soc 2:1–2

Contents

Part I
Virology, Pathogenesis, and Clinical and Epidemiological Features of Disease

Respiratory Syncytial Virus: Virology, Reverse Genetics, and Pathogenesis of Disease

Peter L. Collins, Rachel Fearns and Barney S. Graham

Abstract Human respiratory syncytial virus (RSV) is an enveloped, nonsegmented negative-strand RNA virus of family *Paramyxoviridae*. RSV is the most complex member of the family in terms of the number of genes and proteins. It is also relatively divergent and distinct from the prototype members of the family. In the past 30 years, we have seen a tremendous increase in our understanding of the molecular biology of RSV based on a succession of advances involving molecular cloning, reverse genetics, and detailed studies of protein function and structure. Much remains to be learned. RSV disease is complex and variable, and the host and viral factors that determine tropism and disease are poorly understood. RSV is notable for a historic vaccine failure in the 1960s involving a formalin-inactivated vaccine that primed for enhanced disease in RSV naïve recipients. Live vaccine candidates have been shown to be free of this complication. However, development of subunit or other protein-based vaccines for pediatric use is hampered by the possibility of enhanced disease and the difficulty of reliably demonstrating its absence in preclinical studies.

P. L. Collins (✉)
Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Building 50, Room 6517, 50 South Dr MSC 8007, Bethesda, MD 20892-8007, USA
e-mail: pcollins@niaid.nih.gov

R. Fearns
Department of Microbiology, Boston University School of Medicine, Boston, MA 02118, USA

B. S. Graham
Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

L. J. Anderson and B. S. Graham (eds.), *Challenges and Opportunities for Respiratory Syncytial Virus Vaccines*, Current Topics in Microbiology and Immunology 372,
DOI: 10.1007/978-3-642-38919-1_1, © Springer-Verlag Berlin Heidelberg 2013

Contents

1 Introduction

Respiratory syncytial virus (RSV) was first isolated in 1955, but its biochemical and molecular characterization remained rudimentary for many years due to its relatively inefficient growth in cell culture, pleomorphic and cell-associated nature, and physical instability. Detailed characterization began in 1981 with the molecular cloning and sequencing of RSV RNA. The resulting molecular tools have greatly increased our understanding of RSV and revolutionized research towards treatment and prevention. For example, candidate live-attenuated RSV vaccine viruses designed by reverse genetics are presently in clinical studies (see chapters by H.Y. Chu and J.A. Englund, and by R.A. Karron et al., this volume). Vectored vaccines and recombinantly expressed viral antigen vaccines also are under development (see chapters by T.G. Morrison and E.E. Walsh, and by R.J. Loomis and P.R. Johnson, this volume). The murine monoclonal antibody that was the basis for palivizumab, used clinically for passive immunoprophylaxis in high-risk infants and children, was produced using recombinantly expressed antigen.

2 Classification

RSV is the type species of Genus *Pneumovirus*, Subfamily *Pneumovirinae*, Family *Paramyxoviridae*, Order *Mononegavirales*. Human RSV exists as two antigenic subgroups, A and B, that exhibit genome-wide sequence divergence (Table 1). The other members of this genus are bovine RSV (BRSV), ovine RSV(ORSV), and pneumonia virus of mice (PVM) (see Table 1 for amino acid sequence

Table 1 Amino acid sequence identity between the proteins of RSV subgroup A (RSV-A) and the indicated members of subfamily *Pneumovirinae*[a]

Viruses compared		Amino acid sequence identity for the indicated protein (%)										
		NS1	NS2	N	P	M	SH	G	F	M2-1	M2-2	L
RSV-A versus	RSV-B	87	92	96	91	91	76	53	89	92	72	93
	BRSV	69	84	93	81	89	38	30	81	80	42	84
	PVM	16	20	60	33	42	23	12	43	43	10	53
	HMPV-A	–[b]	–[b]	42	35	38	23	15	33	36	17	45
	AMPV-A	–[b]	–[b]	41	32	38	19	16	35	37	12	43

[a] Viruses are listed in order of decreasing relatedness to HRSV-A. RSV-B is RSV subgroup B; HMPV-A and AMPV-A are subgroup A of human and avaian human metapneumovirus. Viruses in *Paramyxovirinae* are not shown for comparison because the percent identity is <20 % at the level of the entire protein

[b] This virus does not have this gene

Table 2 Notable features of RSV

Replication and budding in vitro are inefficient, infectivity is unstable, particles grown in vitro are mostly large filaments
RSV encodes additional proteins that are either unique to the genus (NS1 and NS2) or found only in a subset of viruses in *Paramyxoviridae* (SH, M2-1, and M2-2)
Two genes, NS1 and NS2, are dedicated to expressing proteins that interfere with the host type I interferon system, among other functions
Overlapping ORFs in the M2 mRNA encode factors that confer transcription processivity (M2-1) or shift viral RNA synthesis from transcription to RNA replication (M2-2)
The M2 and L genes overlap and L mRNA is expressed by a backtracking mechanism
The small hydrophobic (SH) protein forms a pentameric ion channel, but its function is unclear
The F protein precursor is activated by cleavage at two furin recognition sites
The F protein activates TLR-4 signaling pathways, but this is inhibited by the G protein
Viral attachment appears to involve both the F and G proteins, but F fuses independently of G
The G protein is heavily glycosylated, nonglobular, and highly variable
The G protein bears a CX3C fractalkine-like motif that may modify the cellular immune response
The G protein is expressed in membrane-bound and secreted forms; the latter interferes with antibody-mediated neutralization, and interacts directly with antigen presenting cells to modify their function

relationships). More pneumoviruses remain to be identified: recent wide-ranging fieldwork provided sequence evidence of RSV-like viruses in African bats (Drexler et al. 2012). Subfamily *Pneumovirinae* contains a second genus, *Metapneumovirus*, which consists of human and avian metapneumoviruses (HMPV and AMPV). The other subfamily of Family *Paramyxoviridae*, *Paramyxovirinae*, includes animal and human parainfluenza viruses (PIVs), mumps, measles, Nipah and Hendra viruses, and numerous other viruses whose number continues to expand (Drexler et al. 2012). Some of the notable features of RSV are summarized in Table 2.

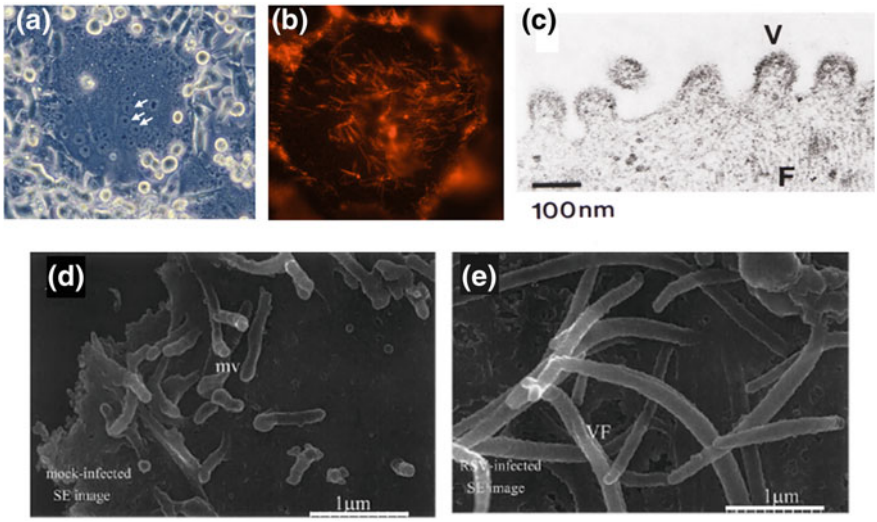

Fig. 1 Photomicrographs (**a** and **b**) and electron micrographs (**c**–**e**) of RSV-infected cells and associated viral structures. **a** is a photomicrograph of a syncytium in an RSV-infected cell monolayer (several nuclei are indicated with *arrows*; courtesy of Dr. Alexander Bukreyev). **b** is a fluorescence photomicrograph of a syncytium in an RSV-infected cell monolayer (not the same one as in **a**) stained with an antibody specific to the F protein, showing filamentous viral projections (courtesy of Dr. Ursula J. Buchholz). **c** is an electron micrograph of negatively stained budding RSV virions: V indicates a budding virion and F indicates filamentous cytoplasmic structures that likely are nucleocapsids (courtesy of Dr. Robert M. Chanock) (Kalica et al. 1973). **d** and **e** are field emission scanning electron micrographs of the surface of uninfected (**d**) and RSV-infected (**e**) cells, illustrating viral filamentous structures (VF in **e**) that are thought to form at sites of virus budding and may yield filamentous particles; also shown are microvilli (mv in **d**) that are found in uninfected cells (courtesy of Dr. Richard Sugrue) (Jeffree et al. 2003)

3 Virion

The RSV virion consists of a nucleocapsid packaged in a lipid envelope derived from the host cell plasma membrane (Figs. 1, 2). Virions produced in cell culture consist of spherical particles of 100–350 nm in diameter and long filaments that usually predominate and are 60–200 nm in diameter and up to 10 μm in length (Jeffree et al. 2003) (Fig. 1). In vitro, 95 % of progeny virus remains associated with the cell surface as particles that seemingly have failed to fully bud. In preparing virus stocks, infected cells typically are subjected to freeze-thawing, sonication, or vortexing to release attached virus, although this reduces infectivity and increases cellular contamination. RSV readily loses infectivity during handling and freeze-thawing due to particle instability and aggregation, although this can be partly overcome by excipients such as sucrose (Ausar et al. 2007). There is indirect evidence that the surface glycoproteins, especially F, are factors in the instability (Sastre et al. 2007; Rawling et al. 2011). The long filamentous shape of the particle likely also confers fragility.

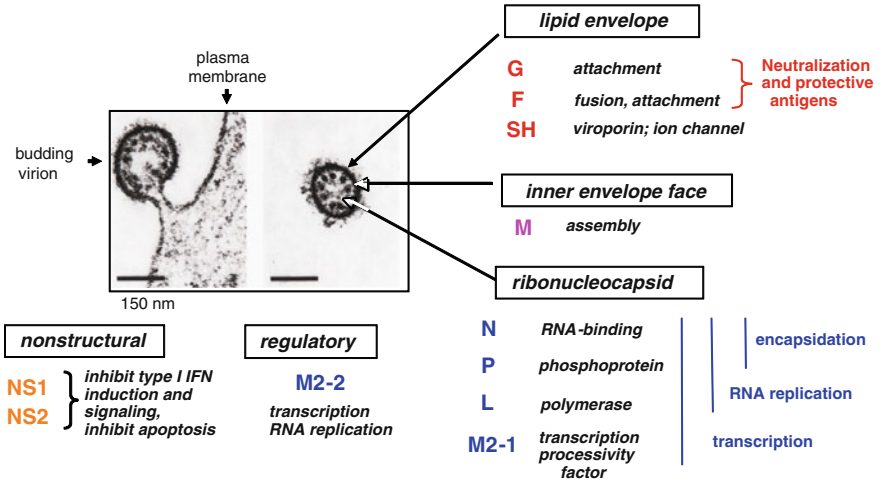

Fig. 2 RSV proteins and their functions and location in the virion, shown in reference strained electron micrographs of negatively stained budding (**a**) and free (**b**) virions (courtesy of Dr. Robert M. Chanock)

The RSV envelope contains three viral transmembrane surface glycoproteins: the large glycoprotein G, the fusion protein F, and the small hydrophobic SH protein (Fig. 2). The nonglycosylated matrix M protein is present on the inner face of the envelope. The viral glycoproteins form separate homo-oligomers that appear as short (11–16 nm) surface spikes. RSV lacks neuraminidase or hemagglutinin activity and the F is known to be heavily sialylated, presumably because of the lack of a neuraminidase. There are four nucleocapsid/polymerase proteins: the nucleoprotein N, the phosphoprotein P, the transcription processivity factor M2-1, and the large polymerase subunit L (Fig. 2).

4 RNAs

The RSV genome (Fig. 3) is a single-stranded nonsegmented negative-sense RNA of 15,191–15,226 nt for six sequenced strains (subgroup A strain A2, 15,222 nt, GenBank accession number M74568, is the reference strain). RNA replication involves a complementary copy of the genome called the antigenome (Fig. 4). The genome and antigenome lack 5′ caps or 3′ polyA tails. The first 24–26 nt at the 3′ ends of the genome and antigenome have 88 % sequence identity (Fig. 3b), representing conserved promoter elements that will be described later. The genome and antigenome are bound separately for their entire length by the N protein to form stable nucleocapsids. These are the templates for RNA synthesis and remain intact throughout the replicative cycle and in the virion. In addition, encapsidation

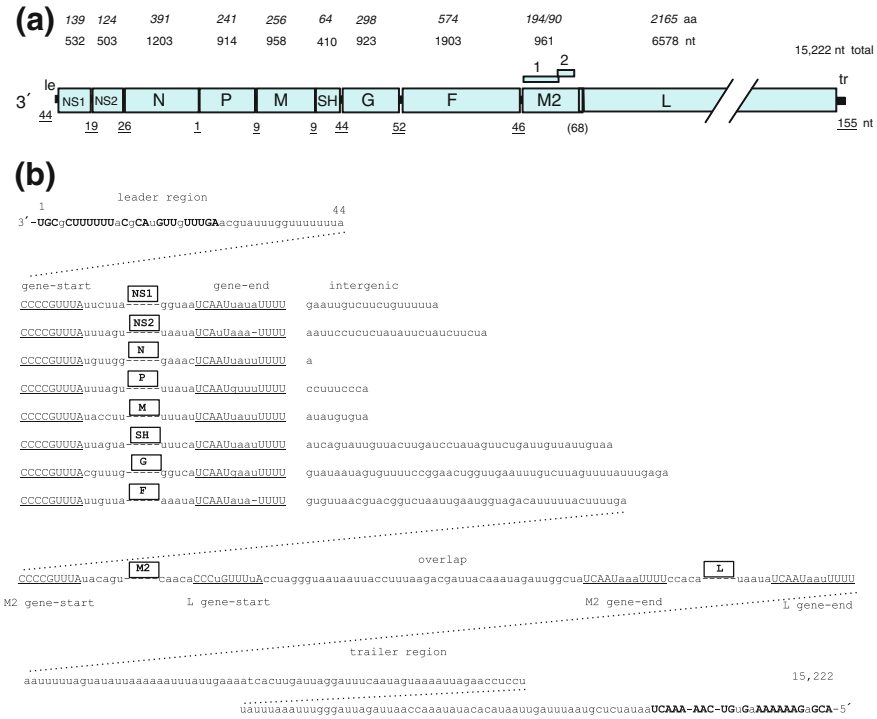

Fig. 3 Diagram of the 3′ to 5′ negative-sense RSV genome (approximately to scale, strain A2), and sequences of the leader region, gene junctions, overlap, and trailer region. **a** Genome diagram: each *box* represents a gene encoding a separate mRNA. The *first row* of numbers over the diagram indicates the nucleotide lengths of the genes, and the second, *upper row* of numbers (*italicized*) indicates the amino acid lengths of the primary, unmodified proteins. The overlapping ORFs of the M2 mRNA are illustrated over the M2 gene. The numbers under the diagram indicate the lengths of the leader, intergenic, and trailer regions (*underlined*) and the gene overlap (*parentheses*). **b** Sequences of the leader region, gene junctions, overlap, and trailer region (3′ to 5′, negative-sense). This shows the leader region, followed by the NS1, NS2, N, P, M, SH, G, and F genes and their intergenic regions, followed by the overlapping M2 and L genes and the trailer region. The main body of each gene is deleted and is represented by a *box* with the gene name. Nucleotide assignments that are conserved between the 3′ ends of the genome and the antigenome (represented here as the reverse complement in the trailer region) are in *bold capitals*. The gene-start and gene-end signals of the gene are *underlined*, and conserved asignments are in *capitals*

protects the RNA from degradation and shields it from recognition by host cell pattern recognition receptors that initiate innate immune responses.

The genome has 10 genes in the order 3′ NS1-NS2-N-P-M-SH-G-F-M2-L (Fig. 3). Each gene encodes a corresponding mRNA (Fig. 4) (Collins et al. 1986). The mRNAs have methylated 5′ caps and 3′ polyA tails. Each mRNA encodes a single major protein except for M2, which has two separate ORFs that overlap slightly and encode the M2-1 and M2-2 proteins. The downstream M2-2 ORF is

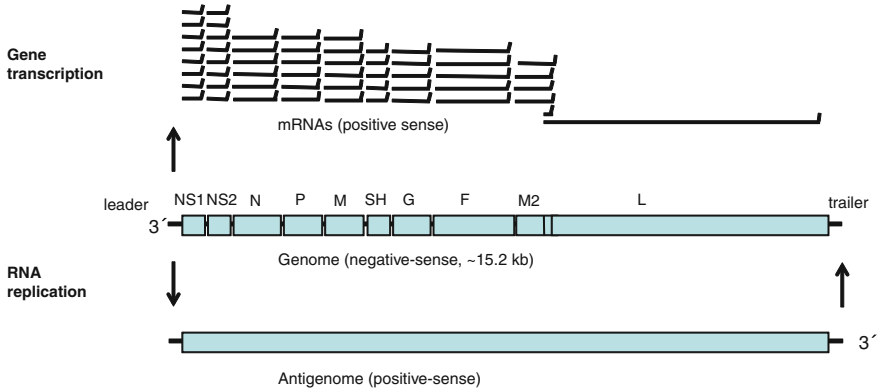

Fig. 4 Overview of RSV transcription and RNA replication. The polymerase enters the negative-sense genome at its 3′ end executes transcription to yield positive-sense subgenomic mRNAs (in a polar gradient) or executes the first step in RNA replication to yield full-length positive-sense antigenome. The polymerase enters the antigenome at its 3′ end and executes the second step of RNA replication to yield full-length progeny genomes. Note that the L gene yields two polyadenylated mRNAs: a very short species due to termination in the gene overlap, and full-length L mRNA

accessed by ribosomes that exit the M2-1 ORF and reinitiate, a process that is influenced by upstream structure in the M2 mRNA (Gould and Easton 2007).

The 3′ end of the genome consists of a 44-nt extragenic leader region that precedes the NS1 gene. The 5′ end of the genome consists of a 155-nt extragenic trailer region that follows the L gene (Fig. 3). Each gene begins with a highly conserved 9-nt gene-start (GS) signal and terminates with a moderately conserved 12–14-nt gene-end (GE) signal that ends with 4–7 U residues (genome-sense) that encode the polyA tail by polymerase stuttering (Fig. 3b). The first nine genes are separated by intergenic regions that vary in length from 1 to 58 nt for the strains sequenced to date. These lack any conserved motifs, are poorly conserved between strains, and appear to be unimportant spacers, except that at some gene junctions the first nucleotide of the intergenic region is important for mRNA termination (Bukreyev et al. 2000; Harmon and Wertz 2002). A tolerance for intergenic variability is illustrated by the finding that incrementally increasing the length of an intergenic region in recombinant RSV up to 160 nt had little effect on gene expression or viral replication in vitro; however, this was moderately attenuating in mice, indicating that excessive length is restrictive (Bukreyev et al. 2000). The last two genes, M2 and L, overlap by 68 nt: specifically, the L GS signal is located 68 nt upstream of the end of the M2 gene (Collins et al. 1987) (Fig. 3b). The same overlap occurs in BRSV, and gene overlaps occur for some genes in some members of *Rhabdoviridae* and *Filoviridae*.

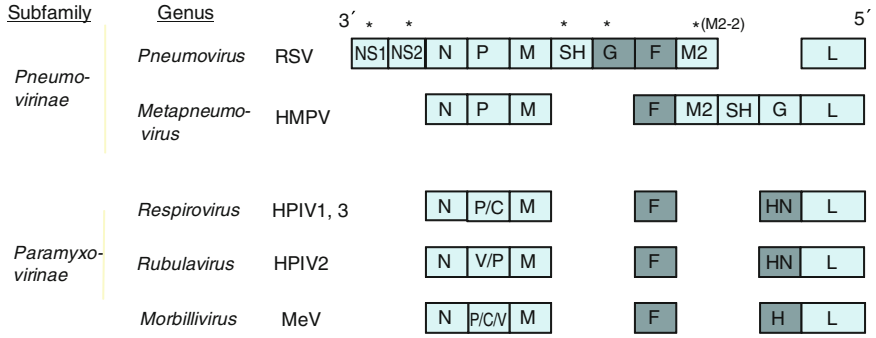

Fig. 5 Comparison of the genes and gene order of RSV with those of selected members of *Paramyxoviridae*: HMPV, human parainfluenza virus (HPIV) serotypes 1, 2, and 3, and measles virus (MeV). The genes are shown in their 3′ to 5′ order in genomic RNA, which is the direction of transcription. Genes are not to scale, and orthologous genes are aligned vertically as much as possible (the only genes that could not be appropriately aligned are SH and G of RSV and SH of HMPV), with gaps introduced to maximize the alignments. Genes encoding major protective antigens are in *dark shading*. *Asterisks* indicate proteins that can be deleted from RSV without loss of replication, although this may be reduced. Proteins that have no direct ortholog in RSV include: *C* small accessory protein, *V* cysteine-rich accessory protein, *HN* hemagglutinin-neuraminidase glycoprotein, *H* hemagglutinin glycoprotein. The *Henipavirus* and *Avulavirus* genera and a number of unclassified viruses within *Paramyxovirinae* are not represented

5 Proteins

RSV encodes 11 separate proteins, and thus is more complex than most members of *Paramyxovirinae*, which typically have 6–7 mRNAs encoding 7–9 separate proteins (Fig. 5). The N, P, M, F, and L proteins of RSV have clear orthologs throughout *Paramyxoviridae*, and their relative genome order is conserved (Fig. 5). Amino acid sequence relatedness between RSV and *Paramyxovirinae* is low and is evident primarily for the F and L proteins and segments in the *C*-terminal region of N. The NS1, NS2, M2-1, and M2-2 proteins of RSV have no counterparts in *Paramyxovirinae*, and an SH protein (which is found in all members of *Pneumovirinae*) is present only in a few members of *Paramyxovirinae*.

The RSV F and G glycoproteins are the only viral neutralization antigens and are the major protective antigens. These proteins are reviewed elsewhere (see chapter by J.S. McLellan et al., this volume) and will only be briefly described here. The 574-amino acid F protein directs viral penetration and syncytium formation, like a typical F protein of *Paramyxoviridae*. The RSV F protein has general structural similarity to the F proteins of *Paramyxovirinae*, and similarly is synthesized as an inactive F0 precursor that is activated by cellular endoprotease to yield two disulfide-linked subunits, NH2-F2–F1-COOH. However, RSV F has two, rather than one, cleavage sites (Gonzalez-Reyes et al. 2001): one site (KKRKRR↓F-137) corresponds to that found in other paramyxoviruses, and the

second site (RARR↓E-110) is located 27 amino acids upstream. RSV F0 is readily cleaved intracellularly by furin-like protease and is not a limiting factor for viral infectivity and tropism. The RSV F protein also binds to TLR-4, initiating signal transduction and innate immune responses (Haynes et al. 2001).

The 298-amino acid RSV large glycoprotein G, involved in attachment, appears to be unrelated to the *Paramyxovirinae* HN, H, or G attachment proteins. RSV G has a membrane anchor near its *N*-terminus, with the *C*-terminal two-thirds of the molecule being external. G also is produced as a secreted form—estimated to account for 80 % of released G protein (Hendricks et al. 1988)—that lacks the membrane anchor due to translational initiation at the second AUG in the ORF followed by proteolytic trimming. The ectodomain of G consists of two large divergent domains flanking a short central conserved segment. The divergent domains have a high frequency of amino acid differences among RSV strains; in addition, variants of G have been noted in nature containing partial intra-gene duplications (Eshaghi et al. 2012), small frame shifts, and *C*-terminal extensions. The central conserved region has a cysteine noose stabilized by two disulfide bonds, and this includes a CX3C motif. Surprisingly, this conserved region can be deleted with little effect on replication in vitro or in mice (Teng and Collins 2002). The large divergent domains have a high content of serine, threonine, and proline residues, as well as a high content of N-linked and, especially, O-linked sugars, and these large domains are thought to have extended, unfolded structures. These features also are characteristic of mucins, suggesting possible mucin mimickry by G, although the significance of this is unknown. The sugar side chains increase the estimated Mr of G from 32,000 for the polypeptide backbone to 80,000–90,000, and possibly 180,000 (Kwilas et al. 2009). The amino acid divergence and the presence of a sheath of host-specified sugars are thought to reduce immune recognition. Surprisingly, given its involvement in attachment, RSV lacking the G gene replicates in some cell lines as efficiently as wt RSV. A live-attenuated RSV vaccine candidate lacking most of the SH and G genes due to spontaneous deletions during passage in vitro appeared to be competent for replication in children, although it was highly restricted (Karron et al. 1997a). RSV isolates have been found, from infants exposed to or infected with human immunodeficiency virus, with deletions spanning most of the G ectodomain, indicating that loss of most of G can occur in nature in some situations (Venter et al. 2011). Thus, G is a malleable, variable protein that is absolutely not essential for replication.

In addition to its role in attachment, G helps RSV evade host immunity. The region of G containing the CX3C motif noted above has been reported to mimic the CX3C chemokine fractalkine, with the effect of reducing the influx of immune cells into the lungs of RSV-infected mice (Tripp et al. 2001). The secreted form of G has been shown to interfere with antibody-mediated neutralization, acting as an antigen decoy as well as impeding cell-mediated neutralization of RSV by Fc receptor-bearing immune cells (Bukreyev et al. 2008). G has been speculated to mimic the receptor for tumor necrosis factor alpha (TNF-α), with the possible effect of inhibiting the antiviral effects of that cytokine (Langedijk et al. 1998). G can interact with DC-SIGN on human dendritic cells (DCs) and alter signaling

pathways associated with antigen presentation (Johnson et al. 2012). Also, the central conserved domain of the G protein has been shown to inhibit the activation of several TLRs including TLR-4, thus countering the effect of the F protein (Polack et al. 2005).

The 64-amino acid SH protein is a transmembrane protein that is anchored near the *N*-terminus, with the *C*-terminus oriented extracellularly. Most of the SH protein is unglycosylated (Mr $\sim$ 7,500), but SH also accumulates in a variety of forms from Mr 4,500 to up to 60,000 or more due to differences including N-linked sugar, polylactosaminoglycan, and translational initiation at the second methionine codon. This array of isoforms is conserved but their significance is unknown. SH forms pentameric pore-like structures that confer cation-selective channel-like activity (Carter et al. 2010; Gan et al. 2012), although the significance of this for RSV is not clear. Thus, the SH protein appears to be a viroporin, a class of small viral proteins that can modify membrane permeability and can affect budding and apoptosis. SH was reported to reduce apoptosis, but the effect was small (Fuentes et al. 2007). SH also appeared to inhibit signaling from TNF-α, an antiviral cytokine (Fuentes et al. 2007). Recombinant RSV lacking SH can replicate somewhat more efficiently in vitro than its wt parent—presumably due to its smaller genome size and smaller number of genes—and was slightly attenuated in mice and chimpanzees (Whitehead et al. 1999).

The 256-amino acid M protein plays key roles in virion morphogenesis. Early in infection, M is detected in the nucleus and may be responsible for the modest inhibition of host transcription during RSV infection, whereas later M is found associated with cytoplasmic viral inclusion bodies—thought to be the site of viral RNA synthesis—and the plasma membrane—the site of virion formation (Ghildyal et al. 2006). M appears to silence viral RNA synthesis by nucleocapsids, presumably in preparation for their packaging into virions (Ghildyal et al. 2006), and appears to be required for the transport of nucleocapsids from viral inclusion bodies to the plasma membrane (Mitra et al. 2012). M is not required to initiate the formation of viral filaments (thought to be the precursor to infectious virus), but in the absence of M the filaments remain stunted and immature (Mitra et al. 2012). Crystallography revealed an M protein monomer that is organized into compact *N*-terminal and *C*-terminal domains joined by a short linker (Money et al. 2009). The monomer surface contains a large positively charged area that extends across the two domains and may mediate association with nucleocapsids and the negatively charged plasma membrane (Money et al. 2009).

The 391-amino acid N protein binds tightly to the genome and antigenome to form helical nucleocapsids, creating the templates for RNA synthesis. N protein expressed in bacteria bound to host RNA to form decamer rings that resembled one turn of the helical nucleocapsid (Tawar et al. 2009). Determination of the atomic structure of the N-RNA rings indicated that each N monomer consists of *N*-terminal and *C*-terminal domains separated by a hinge. Each N monomer was associated with seven nt of RNA, with the RNA groove at the hinge. Adjacent monomers were oriented in the same direction and loosely connected, providing flexibility that would allow polymerase access without disassembling the helix.

Of the seven nt associated with each N monomer, nt 2–4 are oriented into the groove while the other four nt face outward. Passage of the polymerase may induce a transient hinge movement that makes the three buried nucleotides flip out to be accessible (Tawar et al. 2009). The N protein also has a role in antagonizing host innate immunity: N binds to the dsRNA-regulated protein kinase PKR and prevents it from phosphorylating eIF-2a and inhibiting protein synthesis (Groskreutz et al. 2010).

The 241-amino acid P protein is an essential polymerase co-factor. It also acts as an adapter that binds to the N, M2-1, and L proteins to mediate interactions in the nucleocapsid/polymerase complex. In addition, P binds to free N protein monomers and delivers them to nascent genomes/antigenomes, thus preventing N from self-aggregating or binding to nonviral RNA (Castagne et al. 2004). The expression of N and P alone are sufficient to form viral inclusion bodies, which are large, dense cytoplasmic structures that are thought to be the sites of viral RNA synthesis. P exists as a homotetramer formed through a multimerization domain in the middle of the molecule, which is flanked by intrinsically disordered domains (Castagne et al. 2004; Llorente et al. 2006). The *C*-terminal region of P was shown to interact with the nucleocapsid by binding to a pocket on the surface of the N protein that includes discontinuous residues from positions 46–151 brought together in the folded structure (Galloux et al. 2012); other P-N binding sites may also exist, perhaps depending on conformation. P may contribute to conformational changes that help the polymerase access the RNA template (Castagne et al. 2004) and appears to be necessary for promoter clearance and chain elongation by the viral polymerase (Dupuy et al. 1999). It also appears to have a role in dissociating the M protein from the nucleocapsid during uncoating to initiate infection (Asenjo et al. 2008). P is the major phosphorylated RSV protein and contains phosphate at more than 10–12 sites, with different sites exhibiting differing rates of turnover due to interplay between cellular kinases and phosphatases (Asenjo et al. 2005). Many of the activities of P described above appear to be affected by dynamic phosphorylation/dephosphorylation at a subset of these sites, apparently involving a small percentage of the total phosphate content (Asenjo et al. 2006, 2008). Experiments in which P phosphorylation was reduced by mutational ablation or the use of inhibitor against cellular kinase supported the idea that phosphorylation is important for RSV replication, but that much of the low-turnover phosphate is not essential (Lu et al. 2002).

The 2,165-amino acid RSV L protein is very similar in length to its *Paramyxovirinae* counterparts and shares low but unambiguous sequence relatedness along nearly its entire length. Specific segments are conserved within and beyond *Mononegavirales* that are thought to represent catalytic domains involved in polymerization. Analysis of RSV mutants has provided preliminary identification of functional regions in L, including the polymerization domain (Fix et al. 2011), a putative nucleotide-binding site involved in capping (Liuzzi et al. 2005) as well as residues that affect the efficiency of recognition of GE signals (Cartee et al. 2003).

The 194-amino acid M2-1 protein is an essential transcription processivity factor (Fearns and Collins 1999b; Collins et al. 1996, 1999). M2-1 accumulates in phosphorylated and nonphosphorylated forms and forms a homotetramer via an

oligomerization domain at residues 32–63 (Tran et al. 2009; Cartee and Wertz 2001). The M2-1 protein binds RNA: the specificity of this interaction remains somewhat unclear, but M2-1 may preferentially bind RSV mRNAs (Cartee and Wertz 2001; Blondot et al. 2012). M2-1 also interacts with the P protein: binding to RNA or P involves partially overlapping domains in the center of the molecule (Blondot et al. 2012). M2-1 can be found in viral inclusion bodies and its presence there depends on interaction with P (Blondot et al. 2012). Interactions with RNA or the P protein are essential for the ability of M2-1 to support RNA synthesis and are competitive (Tran et al. 2009; Blondot et al. 2012), suggesting that P delivers M2-1 to the nucleocapsid and is then displaced. M2-1 also binds to the M protein and mediates its transport to inclusion bodies and interaction with nucleocapsids (Li et al. 2008). M2-1 contains a CCCH zinc finger motif near its *N*-terminus (residues 7–25) that is essential for its activity in viral RNA synthesis (Hardy and Wertz 2000). This unusual CCCH motif is also found in tandem in a family of cellular zinc finger proteins that includes tristetraprolin (TTP), which binds to AU-rich elements present in a number of host response mRNAs including cytokine mRNAs and affects their stability. Like TTP, M2-1 was recently shown to associate with cellular stress granules, which are involved in translational regulation under stress conditions, but the significance of this possible similarity is unclear (Fricke et al. 2013). M2-1 is unique to *Pneumovirinae*, although it shares structural homology with the VP30 transcriptional activator of *Filoviridae* (Blondot et al. 2012).

The M2-2 protein (88 or 90 amino acids, depending on the start site; Chang et al. 2005) is expressed at a low level in infected cells, and its status as a virion component is not known. Deletion of M2-2 from recombinant RSV results in a virus that exhibits delayed and reduced RNA replication and increased "runaway" transcription; this contrasts with wt RSV, for which transcription appears to be downregulated later in infection in favor of RNA replication (Bermingham and Collins 1999). These results suggest that M2-2 plays a role in regulating RNA synthesis; specifically, as the level of M2-2 increases during the time course of infection, it reduces transcription and promotes RNA. Consistent with a direct effect on RNA synthesis, over-expression of M2-2 inhibited RNA synthesis by mini-replicons and inhibited replication of complete RSV (Collins et al. 1996; Cheng et al. 2005). In addition, in experiments designed to produce virus-like particles, expression of M2-2 increased the efficiency of packing: this might reflect its effects on RNA synthesis or might be an unrelated activity (Teng and Collins 1998). Replication of ΔM2-2 RSV in vitro is delayed but reaches titers comparable to wt RSV, whereas in mice and chimpanzees the virus was restricted approximately 500- to 1000-fold compared to wt RSV (Bermingham and Collins 1999; Teng et al. 2000).

The NS1 and NS2 proteins (139 and 124 amino acids, respectively) are thought to be nonstructural. While the two proteins can function separately, they appear to form complexes and may have synergistic effects, but this is poorly understood (Spann et al. 2005; Swedan et al. 2011). NS1 and NS2 interfere with innate immune responses including interferon induction and signaling (Spann et al. 2005;

Swedan et al. 2011). They also inhibit apoptosis, thereby prolonging the life of the cell and increasing viral yield (Bitko et al. 2007). In a mini-replicon system, co-expression of NS1—and, to a lesser extent, NS2—inhibited transcription and RNA replication, affecting both the genomic and antigenomic promoters (Atreya et al. 1998). These effects remain to be further investigated, but they suggest that NS1 and possibly NS2 might downregulate and restrain viral RNA synthesis. This may be comparable to effects shown for the C and V proteins of some members of *Paramyxovirinae* that, by downregulating viral RNA synthesis, avoids the accumulation of viral dsRNA that otherwise activates innate immunity. Recombinant RSV lacking the NS1 and/or NS2 genes have increased sensitivity to interferon, cause increased apoptosis, and replicate with reduced efficiency in cultured cells and experimental animals, with the effect of deleting NS1 being greater (Teng et al. 2000; Whitehead et al. 1999) (see chapter by S. Barik, this volume).

6 Transcription and RNA Replication

For transcription, the polymerase enters at the 3′ end of the genome and copies the genes into their corresponding mRNAs (Fig. 4) by a sequential stop–start process guided by the GS and GE signals (Kuo et al. 1996). Synthesis of each mRNA initiates opposite the first nucleotide of the GS signal. Surprisingly, the initiating nucleotide can be selected by the polymerase independent of the template (Kuo et al. 1997). Thus, the GS signal triggers the start of mRNA synthesis; in addition, its complementary sequence present in the nascent transcript is thought to act as a signal for capping and cap methylation by the polymerase, based on analogy with other nonsegmented negative-strand RNA viruses (Wang et al. 2007). Capping and/or methylation appears to be essential for mRNA elongation: when RSV capping was blocked by a specific inhibitor, transcription produced uncapped abortive RNAs of ~45 to 50 nt (Liuzzi et al. 2005). Polymerase that is engaged in mRNA synthesis is unresponsive to encountering an additional GS signal, but encountering a GE signal triggers polyadenylation/termination of the mRNA and makes the traversing polymerase responsive to a GS signal (Kuo et al. 1996; Fearns and Collins 1999a). The triggering of polyadenylation/termination at the various GE signals is not completely efficient, and the polymerase occasionally continues synthesis through the next gene. This produces various readthrough mRNAs that account for approximately 10 % of total mRNA (Collins and Wertz 1983). RSV transcription has a polar gradient in which gene transcription decreases along the gene order (Fig. 4). This is typical for *Mononegavirales* and occurs because some of the transcribing polymerases disengage and exit the genome at the various gene junctions.

Studies with mini-replicons showed that N, P, and L are the viral proteins necessary for transcription, but under these conditions transcription terminates prematurely and nonspecifically within several hundred nt, and genes that are further downstream are not significantly transcribed (Collins et al. 1995, 1996,

1999; Fearns and Collins 1999b). Fully processive transcription requires in addition the M2-1 protein, which can be present in relatively low relative molar amounts (Fearns and Collins 1999b; Collins et al. 1996). In mini-replicon experiments, M2-1 also decreased the efficiency of termination at the GE transcription signals—possibly a reflection of the same processivity activity—resulting in increased production of readthrough mRNAs (Hardy and Wertz 1998). These activities raise the possibility that M2-1 might affect the relative levels of expression of the various viral genes, such as by promoting sequential transcription and reducing the transcriptional gradient, but this has not been observed in infected cells (Fearns and Collins 1999b).

One-way sequential transcription does not provide for initiation at the L GS signal, since it is located upstream of the M2 GE signal, as noted (Fig. 3a). Studies with mini-replicons showed that, upon completing transcription of the M2-1 gene, the polymerase backtracks by retrograde scanning to initiate at the L GS signal (Fearns and Collins 1999a). Apart from the M2 GE and L GS signals, the overlap region did not appear to contain any other *cis*-acting elements needed for this activity. Furthermore, the polymerase was found to scan in both directions. This led to the realization that scanning may occur at each gene junction and may be the mechanism by which the next GS signal is located, and could explain the tolerance for intergenic variability noted above. The presence of the M2 GE signal within the L gene (due to the overlap) causes 90 % of newly initiated L gene transcripts to add polyA and terminate at the signal, producing a 68-nt polyadenylated RNA that does not appear to encode a protein and is not known to have any further significance (Collins et al. 1987). The synthesis of full-length L mRNA depends on the "error" of polymerase readthrough at the M2 GE signal (Collins et al. 1987). Whether the level of M2-1 protein affects this process remains to be evaluated. Although 90 % of L gene transcripts terminate at the M2 GE signal, there is not a steep drop in the transcriptional gradient at the L gene (Fearns and Collins 1999a; Kwilas et al. 2010). This indicates that backtracking is very active. The gene overlap appears to be an accidental arrangement that may provide no advantage but can be tolerated due to the scanning function of the polymerase.

The relative level of expression of the various RSV genes is determined mostly by the polar gradient of transcription. Thus, the most abundant mRNAs are for the NS1 and NS2 proteins that antagonize host responses. The expression of the L mRNA, the last gene in the order, is further reduced by an unidentified effect that appears to be post-transcriptional (Kwilas et al. 2010), possibly mRNA stability. Differences in the efficiency of polyadenylation/termination by the various GE signals due to natural sequence variation may affect the relative levels of gene expression by changing the amount of transcriptional readthrough (Harmon and Wertz 2002). On the one hand, increased readthrough spares the polymerase from disengagement at the gene junctions, providing more polymerase to downstream genes. On the other hand, ORFs that are at internal positions in readthrough transcripts are not efficiently translated, reducing the synthesis of proteins from genes downstream of inefficient GE signals. Thus, the overall effect is complex, and the impact of GE variation is not clear.