IDENTIFICATION OF HOST PROTEINS REQUIRED FOR VESICULAR STOMATITIS VIRUS INFECTION

Debasis Panda
University of Nebraska-Lincoln, debasisvet@gmail.com

Follow this and additional works at: http://digitalcommons.unl.edu/vetscidiss
Part of the Veterinary Medicine Commons, and the Virology Commons

Panda, Debasis, "IDENTIFICATION OF HOST PROTEINS REQUIRED FOR VESICULAR STOMATITIS VIRUS INFECTION" (2011). Dissertations & Theses in Veterinary and Biomedical Science. 8.
http://digitalcommons.unl.edu/vetscidiss/8

This Article is brought to you for free and open access by the Veterinary and Biomedical Sciences, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Dissertations & Theses in Veterinary and Biomedical Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
IDENTIFICATION OF HOST PROTEINS REQUIRED FOR VESICULAR
STOMATITIS VIRUS INFECTION

By

Debasis Panda

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillments of Requirements

For the Degree of Doctor of Philosophy

Major: Integrative Biomedical Sciences

Under the Supervision of Professor Asit K. Pattnaik

Lincoln, Nebraska

August, 2011
Viruses usurp host cell pathways for different stages of their infection. Understanding virus-host interaction will be invaluable to elucidate molecular mechanisms of virus infection and to identify drug targets. In order to identify such critical cellular genes in vesicular stomatitis virus (VSV, a model non-segmented negative strand RNA virus) infection, we developed a stable cell line constitutively expressing replication proteins of VSV. Attempts to establish a cell line replicating a sub-genomic replicon was not successful because of induction of interferon response by replication of viral genomic analog. Subsequently, we used siRNA technology and conducted a genome-wide siRNA screen in HeLa cells to identify host factors regulating VSV infection. A total of 23,000 human genes were knocked down individually, and their effect on viral infection was interrogated using a high-throughput cell-based assay. Our study identified several previously unknown host proteins required for VSV infection. Bioinformatics analysis predicted enrichment of several biological functions among these proteins and some of them are commonly utilized by other pathogens such as human immune deficiency virus (HIV), hepatitis C virus (HCV) and Influenza virus. We also noted that 35% of these genes (25 out of 72) are required for lymphocytic choriomeningitis virus (LCMV) and human parainfluenza virus type 3 (HPIV3) infection, suggesting evolutionary conserved mechanisms of virus-host interactions. Further studies
focusing on host coatomer complex 1 (COPI) identified a role of COPI in early stage of VSV infection. The effect of COPI is mediated at the level of viral RNA synthesis. COPI functions are required not only for VSV but also for LCMV and HPIV. ADP ribosylation factor 1 (ARF1), the immediate upstream modulator of COPI was found as a required factor for VSV RNA synthesis. ARF1 is activated by the Golgi-associated brefeldin A resistant guanine nucleotide exchange factor 1 (GBF1) which was found to be a critical determinant of VSV RNA synthesis. These studies suggested that the components of the cellular secretory pathway are required for VSV RNA synthesis.
ACKNOWLEDGMENTS

I would like to thank many people for their excellent support and invaluable assistance both professionally and personally.

I gratefully acknowledge the help of my Ph.D. supervisor, Dr. Asit K. Patnaik. During the past five years his encouragements, inspirations, thoughtful advice to tackle research problems were of immense help to complete this task. He not only provides guidance and approach but also provides hope and perspective. Proper design of experiments and critical analysis of results, two things leaned in this lab, I will treasure in my future career. In this lab, I also learned to build long term aspirations in research career. I also want to take this opportunity to thank my supervisory committee members Dr. Peter Angeletti, Dr. Fernando Osorio and Dr. Jay Reddy for their valuable guidance during my research. The efforts that Dr. Osorio and Dr. Reddy have made to read carefully my thesis are acknowledged. Special thanks to Dr. Reddy for his suggestions to build a successful career.

I extend my sincere gratitude to Dr. James Pearson, Nick Barrows, Dr. You Zhou, Dr. David Kelly and Terri Fangman for being tolerant to my numerous questions. Expert opinions from these people have contributed greatly to my research. I am thankful to my current lab members Phat Dinh and Anshuman Das for providing a wonderful collaborative environment to work. Sincere thanks to former lab members Subash Das, Israrul Ansari, Debasis Nayak and Phani Das for teaching me several techniques. Thanks to Lalit and Sakthi for their generous help. Members of Dr. Clinton Jones lab and Dr.
Charles Wood lab have also generously helped me with reagents which greatly accelerated my research.

I am thankful to my friends in Lincoln and Omaha for providing an environment full of fun which at times was required to ease the pressure.

I am indebted to my parents for their love, sacrifices and support throughout my life. Without their continuous encouragement it would not have been possible for me to perform the research. Last but not least, I owe thanks to my beloved wife Rakhi for proofreading this thesis and whose love and motivation stood beside me as strength to complete this work.
# Chapter I: Introduction

1.1 Significance of vesicular stomatitis virus as a model pathogen 1

1.2 Classification of VSV 3

1.3 Structure of the virion 3

1.4. Viral genome organization 5

1.5 Viral genome products 5

1.5.1 Nucleocapsid protein 5

1.5.2 Phosphoprotein 6

1.5.3 Matrix protein 9

1.5.4 Glycoprotein 10

1.5.5 Large polymerase protein 11

1.6 VSV life cycle 11

1.6.1 Virus attachment 11

1.6.2 Virus entry and uncoating 12

1.6.3 Transcription of Genome 14

1.6.4 Replication of Genome 17

1.6.5 Virus assembly 19

1.7 Defective interfering particles of VSV 20

1.8 Reverse genetics system of VSV 20

1.9 Host factors required for VSV infection 21

1.10 Sub-genomic replicon system to study viral replication 25
1.11 High-throughput siRNA screen to identify host factors required for virus infection 26

1.12 Overall objective and experimental plan 28

   1.12.1 Objective 1: Hypothesis and experimental approach 28

   1.12.2 Objective 2: Hypothesis and experimental approach 30

Chapter II: Materials and methods

2.1 Cell culture and maintenance of cell lines 31

2.2 Preparation of stock of VSV 31

2.3 Preparation of stock of LCMV and HPIV-3 31

2.4 Preparation of recombinant vaccinia virus (vTF7-3) stock 32

2.5 Stocks of defective interfering particles 32

2.6 Plaque assay 32

2.7 Nucleocapsid preparation and transfection 33

2.8 Molecular cloning 34

   2.8.1 Preparation of chemically competent bacteria 34

   2.8.2 Recombinant DNA technique 34

   2.8.3 Plasmid purification 35

2.9 Construction of plasmids 35

   2.9.1 Construction of N, P, L plasmids under the CMV promoter 35

   2.9.2 Construction of a plasmid expressing N, P and L 36

   2.9.3 Construction of mini replicon 37

   2.9.4 Construction of a sub-genomic replicon of VSV 38

2.10 Plasmid DNA transfection 38
2.11 Establishment of stable cell line expressing N, PeGFP and L

2.12 Metabolic labeling and immunoprecipitation of protein

2.13 Metabolic labeling and analysis of RNA after DI particle infection

2.14 Luciferase assay

2.15 Reverse transcription (RT) PCR to detect DI particle genome

2.16 Quantitative RT PCR

2.17 Immunoblotting

2.18 High-throughput assay for primary screen
    2.18.1 siRNA Library Design
    2.18.2 Primary screen assay optimization
    2.18.3 Automated determination of cell number and percentage infection

2.19 Protocol for primary Screen
    2.19.1 Transfection
    2.19.2 Virus infection
    2.19.3 Fixing and staining
    2.19.4 Statistical analysis

2.20 Validation screen assay design

2.21 siRNA transfection in six well or twelve plates

2.22 Immunofluorescence

2.23 Antibodies

2.24 Chemicals

2.25 Treatment of cells with chemicals
Chapter III: Generation and Characterization of Cell Lines Expressing VSV Replication Proteins and Attempts to Establish A Cell Line Carrying Selectable Sub-Genomic Replicons of VSV

3.1 Construction of a cell line constitutively expressing VSV replication proteins 54

3.2 Efficient replication of DI particle genomes in cells stably expressing the VSV replication proteins 58

3.3 Failure to demonstrate persistent DI particle replication using the NPeGFPL stable cell line 62

3.4 Failure to demonstrate replication of a minimal replicon of VSV 62

3.5 Attempt to develop a cell line supporting replication of an autonomous replicon 64

3.5 DI particle RNA replication activates IFNβ promoter 69

3.6 DI RNA replication activates ISRE promoter and NFκB promoter 70

3.7 Discussion 74

Chapter IV: Genome wide RNAi screen identifies host proteins required for VSV infection

4.1 Genome wide siRNA screen 82

4.2 Statistical analysis to obtain primary hit list candidates 84

4.3 Secondary validation screen 84

4.4 Identification of host genes required for LCMV and HPIV3 87

4.5 Bioinformatics analysis 88

4.6 Discussion 103
Chapter V: Host Coatomer complex I function is required for VSV RNA synthesis

5.1 COPI is required for early stage of VSV infection in HeLa cells 108
5.2 COPI is required for early stage of LCMV and HPIV3 infection 112
5.3 COPI functions to regulate VSV RNA synthesis 113
5.4 ARF1 mediated COPI assembly is required for VSV and LCMV but not for HPIV3 gene expression 118
5.5 GBF1, the GEF for ARF1 is required for VSV, LCMV and HPIV3 infection 120
5.6 GBF1 and COPI are not part of VSV replication complex 125
5.7 Discussion 130

Chapter VI: Summary and conclusion 134
List of Figures

Figure 1.1 Cryo-electron microscopic architecture of virion of VSV 4
Figure 1.2 Genome organization of VSV 8
Figure 1.3 Schematic representation of VSV life cycle 13
Figure 1.4 Schematics of reverse genetics system of VSV 23
Figure 3.1 Transient expression of VSV proteins by CMV promoter 55
Figure 3.2 Schematic representation of the plasmid expressing the VSV N, PeGFP, and L proteins 57
Figure 3.3 Detection of VSV proteins in stable cell clones by immunofluorescence 59
Figure 3.4 Expression of VSV proteins in different cell clones 60
Figure 3.5 Replication of DI particle RNA genome in 293-NPeGFPL cell clones 61
Figure 3.6 Long term replication of DI genome in NPeGFPL cells 63
Figure 3.7 Schematic of bicistronic mini replicon 65
Figure 3.8 Schematic of autonomous replicon 67
Figure 3.9 GFP expression in 293 cell line transfected with autonomous replicon 68
Figure 3.10 Induction of interferon by replication of DI RNA 71
Figure 3.11 Replication of DI genome in cells expressing N-PeGFP and L proteins 72
Figure 3.12 Kinetics of IFNβ promoter activation 73
Figure 3.13 Activation of ISRE and NFκB promoter by DI particle replication 75
Figure 3.14 Activation of ISG-56 protein by DI RNA replication 76
Figure 4.1  Genome wide siRNA screen design 83
Figure 4.2  Schematic of secondary validation screen 85
Figure 4.3  Effect of gene depletion on VSV, LCMV and HPIV3 infection 99
Figure 4.4  Biological process classifications for the identified candidates 100
Figure 4.5  Comparison of genome-wide screens between VSV and other viruses 102
Figure 5.1  Effects of COPI depletion on VSV protein and RNA synthesis 110
Figure 5.2  Depletion of COPI sub-units affects VSV gene expression 111
Figure 5.3  Inhibition of VSV gene expression by CBM 114
Figure 5.4  COPI silencing decreases gene expression of LCMV and HPIV3 115
Figure 5.5  COPI silencing affects VSV gene expression which is independent of VSV entry and budding 117
Figure 5.6  ARF1 is required for VSV gene expression 119
Figure 5.7  GBF1 controls VSV RNA synthesis 121
Figure 5.8  VSV RNA synthesis is sensitive to GBF1 inhibitors 123
Figure 5.9  GBF1 is the GEF required for VSV RNA synthesis 124
Figure 5.10  GBF1 affects a post entry stage of VSV infection 126
Figure 5.11  Requirement of GBF1 for LCMV and HPIV gene expression 127
Figure 5.12  GBF1 and COPI do not co-localize with VSV replication complex 129
Figure 6.1  Integrated model of VSV-host interactions 138
## List of tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1</td>
<td>Sequences of Primers and Probes</td>
<td>43</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>Optimized concentrations of primers and probes</td>
<td>43</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>List of genes identified as required factors for VSV</td>
<td>89</td>
</tr>
</tbody>
</table>
### Abbreviations used

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSV</td>
<td>Vesicular stomatitis virus</td>
</tr>
<tr>
<td>LCMV</td>
<td>Lymphocytic choriomeningitis virus</td>
</tr>
<tr>
<td>HPIV3</td>
<td>Human para influenza virus 3</td>
</tr>
<tr>
<td>Hr(s)</td>
<td>Hour(s)</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>COPI</td>
<td>Catomer protein complex 1</td>
</tr>
<tr>
<td>ARF1</td>
<td>ADP ribosylation factor 1</td>
</tr>
<tr>
<td>GBF1</td>
<td>Golgi specific brefeldin resistant factor 1</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>RdRp</td>
<td>RNA dependent RNA polymerase</td>
</tr>
<tr>
<td>ISRE</td>
<td>Interferon stimulated response element</td>
</tr>
<tr>
<td>NFκB</td>
<td>Nuclear factor of kappa light polypeptide gene enhancer in B-cells</td>
</tr>
<tr>
<td>DI</td>
<td>Defective interfering</td>
</tr>
<tr>
<td>eGFP</td>
<td>Enhanced green fluorescent protein</td>
</tr>
<tr>
<td>GFPrr</td>
<td><em>Renilla reniformis</em> GFP</td>
</tr>
<tr>
<td>M protein</td>
<td>Matrix protein</td>
</tr>
<tr>
<td>N protein</td>
<td>Nucleocapsid protein</td>
</tr>
<tr>
<td>P protein</td>
<td>Phosphoprotein</td>
</tr>
<tr>
<td>L protein</td>
<td>Large polymerase protein</td>
</tr>
<tr>
<td>G protein</td>
<td>Glycoprotein</td>
</tr>
<tr>
<td>WB</td>
<td>Western Blotting</td>
</tr>
</tbody>
</table>
Chapter I

INTRODUCTION

1.1 Significance of vesicular stomatitis virus as a model pathogen

Vesicular stomatitis virus belongs to the family *Rhabdoviridae* and order *Mononegavirales*. It is an enveloped virus with a negative-stranded RNA genome of 11,161 nucleotides (Lyles and Rupprecht, 2007). Members of the *Rhabdoviridae* family can infect animals, plants as well as insects. VSV spreads between ungulates and rodents via insect vectors. Arthropod vectors such as flies and mosquitoes transmit VSV in natural or laboratory conditions. VSV causes a self-limiting disease termed “vesicular stomatitis” in cattle, horse and pigs. Vesicular stomatitis is rarely fatal but is associated with significant economic loss to live stock producers. Infection in humans by VSV is asymptomatic. Mild flu like symptoms has been reported after VSV infection.

VSV has served as an excellent model pathogen to study fundamental aspects of virus biology and viral pathogenesis. Discovery of RNA dependent RNA polymerase (RdRp) in VSV (Baltimore *et al.*, 1970) led to the classification of single-stranded RNA viruses into positive-and negative-strand RNA viruses. VSV can infect almost all vertebrate cells and many invertebrate cells and has a short infection cycle. These characteristics have earned appreciation for VSV as an excellent model virus to study virus entry, replication and virus budding as well as innate and adaptive immune defense mechanisms. In fact, much of our current understanding about negative strand RNA viruses has come from studies with VSV. The cell-free system provided a major tool to study virus transcription and replication as well as RNA encapsidation (Emerson & Yu, 1975). Sequential nature of transcription and single polymerase entry site has been demonstrated using the cell-free reconstitution system with VSV (Emerson, 1982;
Iverson & Rose, 1981). VSV glycoprotein has been widely used in cell biology to probe the secretory pathways (Bonifacino & Lippincott-Schwartz, 2003). Because of the error prone nature of VSV polymerase and short life cycle, VSV also serves as an elegant model to study virus evolution (Novella, 2003).

Absence of a suitable system for direct manipulation of genomes of negative strand RNA viruses was a major hurdle in the advancement of the field till early 1990s. Reverse genetic system for manipulation of a negative strand RNA virus genome was initially established using VSV defective interfering particles (Pattnaik et al., 1992). Based on this approach, infectious rabies virus and VSV were subsequently recovered from cDNA clone (Lawson et al., 1995; Schnell et al., 1994; Whelan et al., 1995). The reverse genetic system for VSV is a powerful tool to dissect the cis- and trans- acting elements on VSV genome replication and transcription. Much of our understanding of the roles of leader, trailer, and intergenic gene junction for non segmented negative strand viruses has been derived using VSV as a model (Hwang et al., 1998; Li & Pattnaik, 1999; Pattnaik et al., 1995; Stillman & Whitt, 1997; 1998; Wertz et al., 1994; Whelan et al., 2000; Whelan & Wertz, 1999b).

In the 21st century research using VSV as a model has taken a new dimension. The observation that VSV replicates well in transformed cell line but its replication is attenuated in primary cells led the way to use VSV as a selective anti-tumor agent (Barber, 2005; Lichty et al., 2004). Excellent reverse genetic system to manipulate VSV genome to develop a better oncolytic agent has given significant impetus to tumor therapeutics. With the advancement modern molecular biology tools, the researchers have begun to use VSV as a vaccine vector. Vaccines based on live attenuated VSV are
effective in animal models. Strains of VSV that induces interferon response have also been in use as potential vaccine platform. For over four decades, research using VSV has made a significant impact on many aspect of basic virology. The ongoing research using VSV as model virus still holds the promise for exciting discoveries in the coming years.

1.2 Classification of VSV

VSV belongs to the family \textit{Rhabdoviridae} and the genus \textit{Vesiculovirus}. It is placed in the order \textit{Mononegavirales}. The order \textit{Mononegavirales} comprises of four families namely \textit{Rhabdoviridae}, \textit{Paramyxoviridae}, \textit{Filoviridae} and \textit{Bornaviridae}. Members of \textit{Mononegavirales} are enveloped viruses with negative sense RNA genome. Rhabdoviridae also includes other genera such as Lyssavirus, Ephemerovirus, Novirhabdovirus, Cytorhabdovirus and Nucleorhabdovirus. VSV has two different serotypes, Indiana (VSV\textsubscript{I}) and New Jersy (VSV\textsubscript{NJ}). Though certain differences exist, both the serotypes exhibit similar properties (Martinez & Wertz, 2005). VSV\textsubscript{I} is the most studied serotype from molecular biology point of view.

1.3 Structure of the virion

Mature virion of VSV appears bullet shaped with a rounded tip and flattened bottom. VSV is approximately 180 nm long and 80 nm wide. The genomic RNA of VSV is 11,161 nucleotides and lacks the 5’ cap and 3’ poly A tail. Recently a three dimensional structure of trunk of VSV virion has been determined with the use of cryo electron microscopy (Ge \textit{et al.}, 2010).
Figure 1.1 Cryo-electron microscopic architecture of virion of VSV (From Ge et al., 2010, Science)

N molecules are represented in green whereas M proteins are shown in blue. The inner and outer leaflets of the envelope are represented in purple and pink.
The outer most layer is a lipid envelope decorated with the glycoprotein (G). The lipid envelope is derived from the host cell plasma membrane from which the virus buds and is composed of phospholipids and cholesterol. The middle layer consists of the matrix (M) protein forming a bridge between envelope and genome. The innermost layer is the core of the virion that contains the virus genomic RNA tightly wrapped with the nucleocapsid protein (N). This N-RNA also known as the nucleocapsid or ribonucleoprotein core to which the viral RdRp, a complex of the large polymerase protein (L) and the phosphoprotein (P) associate. The interaction of N subunits together with the M and nucleocapsid interaction imparts the classic bullet shape to the virion.

1.4. Viral genome organization

VSV contains a non-segmented negative-strand RNA genome. Because of the negative-sense nature, the viral genome does not function as an mRNA to synthesize proteins. The extreme 3’ end contains a non-coding leader RNA gene of 47 nt while the 5’ end contains the complement of the non-coding trailer sequences of 59 nt. The leader and trailer regions contain sequences essential for transcription, replication and virus assembly. The protein-coding genes are organized in a modular fashion in the order 3’-N-P-M-G-L-5’. Each gene is flanked by sequences important for generation of capped and polyadenylated mRNAs. The untranslated regions of mRNAs do not contain sequences required for translational control or mRNA turnover. Translation of viral mRNA is dependent on host cell translational machinery.

1.5 Viral genome products

1.5.1 Nucleocapsid protein
The nucleocapsid (N) protein is 422 aminoacids long and is a basic protein with a net positive charge. VSV genome inside the virion as well as in the infected cells is always wrapped by the nucleocapsid (N) protein. Approximately 1200 copies of N molecule cover the entire genome with each N molecule covering 9 nucleotides (Green et al., 2006) and this Encapsidation is important for its biological activity. N protein when expressed alone, can encapsidate cellular RNA, but when expressed along with P protein becomes competent to bind viral RNA (Masters & Banerjee, 1988a). Chemical probing analysis suggest that the N proteins sequester the phosphodiester backbone whereas the bases are exposed for the polymerase to recognize them for transcription and replication (Iseni et al., 2000). Crystal structure of N protein bound to the RNA revealed that N protein consists of two separate lobes and the RNA is sequestered inside the grove generated by these two lobes (Green et al., 2006). Each monomer of N protein also interacts with three adjacent N protein monomers and this interaction help stabilize N-RNA decameric disc like structure found in cells expressing the N protein as well as those in virus-infected cells. Recent evidences also favor a model where N protein reversibly encapsidates and releases the RNA and thus allows the polymerase to gain access for transcription and replication (Green & Luo, 2009). N protein is present in the replicase complex but is not present in transcriptase complex (Qanungo et al., 2004) suggesting differential role for N protein in modulating virus RNA synthesis. Recently we and others have identified a role of N protein in the regulation of transcription and replication of the N-RNA template (Harouaka & Wertz, 2009; Nayak et al., 2009).

1.5.2 Phosphoprotein
Phospho (P) protein is 265 amino acids long and is post translationally modified. The P protein is found with different degrees of phosphorylation in infected cells and in the virions (Barik & Banerjee, 1992; Chen et al., 1997). Although early mutagenesis studies have identified domains in the P protein, based on the recent structural and functional studies, we proposed a new domain-organization of this protein. The central region of the molecule (P central domain or P<sub>CD</sub>, spanning residues 107-177) containing parts of the hinge region folds independently and is involved in oligomerization of the protein (Ding et al., 2006). The crystal structure P<sub>CD</sub> predicts that an α-helix and two β-hairpins in this domain are critical for homodimer formation and that the N and L binding domains likely reside outside of this central domain. The C-terminal domain of P (P<sub>CTD</sub>, spanning residues 183-265) also folds as an autonomously folding unit, whose structure has been recently solved alone (Ribeiro et al., 2008) or in complex with the N protein (Green & Luo, 2009). Indeed, the structure of P<sub>CTD</sub> in complex with N protein confirms that P<sub>CTD</sub> interacts with N protein. For viral RNA synthesis to ensue, the P protein must interact with N and L proteins. The structure of the N-terminal domain (P<sub>NTD</sub> spanning residues 1-106) of the protein is not known at this time; it is presumed to fold independently and likely encode critical determinants of P protein functions like transcription activity and L protein binding activity. This domain also has been shown to interact with and maintain the N protein in an encapsidation-competent form (Chen et al., 2007).
Figure 1.2. Genome organization of VSV

Schematic representation of VSV genome is shown. 3’ end contains leader and 5’ end contains trailer. Corresponding positions of genes also shown.
The P protein forms a complex with N protein known as N\textsuperscript{0}-P complex and this is competent to bind viral RNA (Masters & Banerjee, 1988a). The binding of N\textsuperscript{0}-P complex to the RNA also involves interaction of P protein with the L protein. Recognition of RNA by the L protein is mediated through the P protein (Emerson & Yu, 1975). Thus, P protein plays a crucial role in RNA recognition by the L protein. P protein also participates with the L protein in different ratio to form transcriptase and replicase complex (Qanungo et al., 2004). Recently the structure of VSV nucleocapsid in association with the P protein has been solved (Green & Luo, 2009). It was suggested that N protein and P protein undergo conformational change upon interaction with the RNA to accommodate the L protein for transcription and replication.

1.5.3 Matrix protein

The matrix (M) protein is synthesized in the infected cell as a soluble protein and also is the most abundant protein in the mature virion (1600 copies). M protein consists of 229 amino acids and the amino terminal region constitute signal for membrane binding (Lenard & Vanderoef, 1990). It plays a critical role in virus assembly by binding with the nucleocapsids (Chong & Rose, 1994). The classic bullet shaped structure of VSV was due to interaction of M with the nucleocapsids. VSV budding occurs at the plasma membrane and M protein forms membrane microdomains at the plasma membrane where there is co-localization of nucleocapsids and the M protein (McCreedy & Lyles, 1989; Ohno & Ohtake, 1987; Swinteck & Lyles, 2008). Nucleocapsids are selectively recruited by the M protein for assembly (Flood & Lyles, 1999). VSV M protein contains a late domain (PPPY), which appears to be responsible for recruiting the host cellular machinery for final release of virus particles (Harty et al., 2001; Jayakar et al., 2000). In
addition to its role in virus assembly and budding, M protein also plays a role in virus pathogenesis by inhibiting interferon synthesis and inducing apoptosis (Ahmed et al., 2003; Kopecky & Lyles, 2003a). Several lines of evidence also indicate that M protein mediated cytopathic effect is due to the suppression of host gene function (Ahmed & Lyles, 1998; Ferran & Lucas-Lenard, 1997; Jayakar & Whitt, 2002). By interacting with Rae1, M protein inhibits nucleocytoplasmic transport of host genes (Enninga et al., 2002; Faria et al., 2005). Cytopathogenesis by M protein is also due to inhibition of host translation (Connor & Lyles, 2002). Induction of apoptosis in VSV infection is due to M protein alone (Kopecky & Lyles, 2003b; Kopecky et al., 2001). VSV-M protein mediated apoptosis is due to induction of intrinsic apoptotic pathway (Balachandran et al., 2000; Kopecky et al., 2001).

1.5.4 Glycoprotein

Glycoprotein (G) is the only protein present in the outer envelope of the virion. G protein is a type I transmembrane protein. On the surface of the virion, G proteins exist in the form of trimers. Majority of the G protein is located outside the virion and constitute the ecto domain. The crystal structure of G protein has been solved recently (Roche et al., 2006; Roche et al., 2007). VSV can infect almost all the cell types and G protein-mediated attachment to the receptor is a critical determinant for this. VSV G has affinity for negatively charged phospholipid protein and this is important for attachment and membrane fusion (Carneiro et al., 2002). Low pH-dependent conformational change of the G protein and fusion inside the endosome is necessary to release the virion content (Puri et al., 1988). The cytoplasmic amino acid sequences are required for incorporation of G protein into the virion (Whitt et al., 1989). G proteins form specialized membrane
micro domains on the plasma membrane for virus budding and these are formed independent of other viral proteins (Brown & Lyles, 2003). Recent evidence suggest that G protein participate with M protein in virus budding (Swinteck & Lyles, 2008).

1.5.5 Large polymerase protein

Large polymerase protein (L) is the catalytic subunit of viral polymerase and is required for transcription and replication. In addition to RNA synthesis, L protein also catalyzes mRNA cap addition and cap methylation (Ogino & Banerjee, 2007; Sleat & Banerjee, 1993). The L protein is a 250 kDa protein and composed of six conserved region (Poch et al., 1990). The RNA polymerase activity maps to conserved region III (Sleat & Banerjee, 1993), whereas the mRNA capping and cap methyl transferase activities maps to region V and region VI respectively (Li et al., 2005; Li et al., 2008). Electron microscopy and deletion analysis showing the molecular architecture of L protein provided more detail understanding of the organization of different functional domains of the L protein (Rahmeh et al., 2010). L protein is also a component of both replicase and transcriptase (Qanungo et al., 2004).

1.6 VSV life cycle

VSV is an arthropod-borne virus. VSV infection in the arthropod remains asymptomatic where as in the animals the virus causes disease. The lifecycle of VSV is typical representative of most non-segmented negative strand RNA viruses. Infection starts with entry of virus to susceptible cells and it involves several stages.

1.6.1 Virus attachment

VSV can infect almost all cell lines in culture condition but the exact receptor for VSV has not been identified. However because of wide tissue tropism, the receptor for
VSV has been suggested to be a ubiquitous molecule. Phosphatidyl serine, a negatively charged membrane phospholipid has been suggested to be the receptor for VSV (Schlegel et al., 1983). However others have provided evidence that phosphatidylserine is not the receptor for VSV (Coil & Miller, 2004). Phosphatidyl serine exists on the inner side of plasma membrane and may not be a suitable receptor for VSV attachment. Experimental evidence also suggests that endoplasmic reticulum chaperone Gp96 is a factor required for VSV G-mediated attachment of virion to the host cell surface (Bloor et al., 2010). Non-specific electrostatic and hydrophobic interaction between VSV and host cell may mediate attachment of virion to the cell surface (Bailey et al., 1984). Also the pH-dependent conformation of G protein of VSV is required for attachment of virus to cell membrane (Fredericksen & Whitt, 1998).

1.6.2 Virus entry and uncoating

After attachment, virus enters to the cell by receptor mediated endocytosis. VSV was one of the model system used to investigate the mechanism of virus entry as early as 1980s (Matlin et al., 1982). In fact, VSV entry was shown to be dependent on clathrin mediated endocytosis (Cureton et al., 2009; Cureton et al., 2010; Sun et al., 2005). Once attached, virions are internalized by clathrin-coated vesicles.
Figure 1.3: Schematic representation of VSV life cycle

Model of VSV infection cycle showing important stages starting from virus adsorption to virus budding. (Adapted from The Rhabdoviruses, R.R. Wagner Ed., Plenum Press, New York)
These coated vesicles after losing their coat form early endosomes. From which, virus is presented to the multi vesicular bodies. As virus travels through the endocytic pathways, the pH of the carrier vesicles progressively drops and there is a pH-dependent conformational change of the G protein of VSV. At pH below 6.5, G protein fuse with the endosomal membrane and the internal contents of the virus is trapped in the endosome. A back fusion event in the multivesicular body is necessary for release of viral RNA to the cytoplasm (Le Blanc et al., 2005). Using a dual fluorescent VSV, we have also demonstrated that it takes approximately 28 minutes for the virus to deliver its content to the cytoplasm after attachment to the cell surface (Das et al., 2009).

1.6.3 Transcription of Genome

Dissociation of M protein from the input nucleocapsid is the primary event of uncoating and is necessary for transcription of the viral genome (Li et al., 1989; Rigaut et al., 1991). VSV genome is always associated with the N protein. The viral RNA dependent RNA polymerase (RdRp) which is a combination of the L protein and P proteins is also associated with the N-RNA template for transcription and replication. VSV contains a negative strand RNA genome and cannot initiate protein synthesis directly. Therefore VSV packages its own RdRp to initiate RNA synthesis. Primary transcription from the input RNA is the first biosynthetic step after infection. Viral polymerase directs the synthesis of viral mRNA without the help of viral proteins (Baltimore et al., 1970; Emerson & Yu, 1975; Moyer & Banerjee, 1975). A general feature of VSV as well as other non-segmented negative strand RNA viruses is the sequential nature of transcription (Abraham & Banerjee, 1976; Ball & White, 1976). As a result of sequential nature of viral genome transcription and attenuation of transcription at
each gene junction, a gradient of mRNA abundance is seen in infected cells as well as under in vitro transcription reactions. The genes present at the 3’ promoter proximal end are transcribed more frequently compared to the genes present in the 5’ promoter distal end (Villarreal et al., 1976).

The common features in all NNS RNA viruses is that transcription is sequential (Abraham & Banerjee, 1976) and that the genes located at the 3’-end of the genome are transcribed more frequently than the genes present at the 5’-end, resulting in a gradient in mRNA abundance (Iverson & Rose, 1981). Although several models, such as precursor-cleavage model, multiple entry site model, were proposed to account for the polar and sequential nature of transcription of VSV genomes, the most favored model (Emerson, 1982) posits that the RdRp begins transcription from the extreme 3’-end of the genome to generate Le RNA and then stops and reinitiates transcription at the N gene start site. In this single-entry, stop-start model, as the RdRp moves along the template, it stops and starts at various gene junctions, generating the individual mRNAs. Transcriptional attenuation at the gene junctions due to the inability of a fraction of the RdRp to reinitiate transcription could result in the gradient in mRNA abundance. As the N protein becomes available, its association with the nascent Le RNA allows the RdRp to read-through the Le-N and other gene junction sequences, generating the N protein-encapsidated full-length replication products. How the N protein binding to Le sequences allows the polymerase to readthrough the gene junction remains an open question. Although the model accommodates many of the available data, more recent data are incompatible with this single-entry model. If one considers single-entry for the RdRp at the 3’-end of the genome, Le RNA should be synthesized in excess over the N mRNA because
transcription is sequential and polar. However, polR1 mutants of VSV synthesize excess N mRNA over Le RNA suggesting that the RdRp initiates transcription from the N mRNA start site independent of Le RNA synthesis (Chuang & Perrault, 1997). The results implicate the presence of two entry sites for the RdRp, one at the extreme 3’-end and the other at the Le-N junction, although they do not rule out the possibility of RdRp entering only at the 3’-end and reaching the N mRNA start site at the Le-N junction by a nontranscriptive scanning mode. The demonstration that transcription and replication initiate at separate sites in VSV (Whelan & Wertz, 2002) and the purification and characterization of two distinct RdRp complexes from VSV-infected cells performing transcription and replication functions separately (Qanungo et al., 2004) also implicate the two-entry site model.

The leader and trailer sequences not only play roles in encapsidation and packaging of the genome into nascent virion, but also play indispensable role in the regulation of transcription and replication (Li & Pattnaik, 1999; Whelan & Wertz, 1999b). Deletion analysis in the leader region identified that first 24 nucleotides are responsible for replication and transcription whereas nucleotides 25 to 47 are required for transcription (Li & Pattnaik, 1999). The leader and N gene junction sequences also play important role in transcription (Whelan & Wertz, 1999b). Using a bipartite VSV replicon system, the role of conserved gene junction sequences has also been demonstrated (Stillman & Whitt, 1997; 1998). First three residues in the mRNA, “UUG” are important not only for transcription but also for mRNA capping and polyadenylation (Stillman & Whitt, 1999). Alteration of these signals affects VSV polymerase processivity, and results in premature transcripts. Large polymerase protein also has additional functions of
capping and methylation of the nascent messages. Capping of the nascent transcripts is important for transcription elongation and failure of correct 5’ capping results in abortive transcription (Stillman & Whitt, 1999). VSV polymerase-mediated capping is different from that of host (Ogino & Banerjee, 2007). All the VSV mRNAs have conserved sequences 5’ AACAG 3’. VSV polymerase adds the cap to the mRNAs but not to the leader sequence. The histidine-arginine motif in the L protein is responsible for transfer of 5’ monophosphorylated RNA onto the GDP to form the cap structure. Highly conserved sequences present at the end of the each protein-coding gene also regulate transcription termination and polyadenylation. Conserved sequences at the intergenic region (3’ AUAC UUUUUU -G/C A- UUGUC nn UAG, n is not conserved) also regulate transcription. Mutagenesis studies have identified a critical role of these sequences in transcription termination as well as efficient transcription re-initiation of each gene (Barr et al., 1997a; b; Hwang et al., 1998). Transcription initiation requires efficient termination of preceding gene (Stillman & Whitt, 1997; 1998). The mechanism by which the AUACUUUUUU sequence regulates polyadenylation is not completely understood. It has been proposed that viral polymerase adds the polyadenylate residues to the mRNAs by reiterative transcription. The intergenic di nucleotides G/CA act as a spacer for transcription termination and re-initiation. The relative distance between the gene start and gene end also plays a role in transcription termination (Whelan et al., 2000).

1.6.4 Replication of Genome

During replication, full-length viral genome is replicated to produce anti-genome. The anti-genome subsequently serves as a template to produce full-length genome. The
fundamental difference between replication and transcription is the requirement of newly synthesized viral proteins for replication. In the presence of cycloheximide, replication is inhibited whereas transcription still continues (Wertz & Levine, 1973). In a cell free system, N protein was found to be required to encapsidate the genome for replication (Patton et al., 1984). P protein along with N protein is required for replication of genome in infected cells (Peluso, 1988). It has been also suggested that the P protein maintains the N protein in a soluble encapsidation competent form (Masters & Banerjee, 1988a; b).

Both the genome and anti-genome contains signal at their 3’end for replication. The genome also contains signal for packaging into progeny virion. Studies using minigenome and defective interfering particles have been instrumental in mapping the sequences required for efficient replication (Li & Pattnaik, 1997; 1999). The 3’ termini of genome and anti-genome are identical at 15 of 18 position and these terminal complementarities were found to be essential for genome replication (Wertz et al., 1994). However the sequences differ significantly at positions 19-46. In the genomic promoter, these sequences are required for transcription whereas in the anti-genomic promoter these sequences are required for replication (Li & Pattnaik, 1997; 1999). Because the anti-genomic promoter replication is more efficient than the genomic promoter replication, virus infected cells more genomic copies are found than the anti-genomic copies.

Recently, separate VSV transcriptase and replicase carrying transcription and replication functions respectively has been identified (Qanungo et al., 2004). The replicase is a complex of L, N and P proteins whereas transcriptase consists of L and P protein along with host guanylyltransferase, translation elongation factor 1α and heat shock protein 60. The precise role of these cellular proteins in genome transcription is unknown. Although
there is considerable advancement of the research in identifying the cis-and trans-acting viral elements in genome replication and transcription, little is known about the role of host proteins modulating transcription and replication.

1.6.5 Virus assembly

VSV assembly occurs at the plasma membrane. Individual viral components are transported to the site of assembly through different routes. The G protein is transported to the plasma membrane through the secretory route and forms microdomains (Brown & Lyles, 2003). The sequence of G protein facilitating budding was shown to be at the membrane proximal region as well as at the cytoplasmic domain (Robison & Whitt, 2000; Schnell et al., 1998). Viral M protein is synthesized as a soluble protein and is transported to the plasma membrane by a yet unknown mechanism and form separate microdomain independent of G protein (Swintek & Lyles, 2008). The amino terminal 20 amino acids of M protein is required for membrane association (Lenard & Vanderoef, 1990). Nucleocapsids once synthesized in the cytoplasm are transported to the plasmamembrane in microtubule dependent manner (Das et al., 2006). The nucleocapsids are selected to bind with the M protein for budding because majority of intracellular nucleocapsids do not interact with M protein (Flood & Lyles, 1999). VSV genome contains a cis-acting signal at the 5’ end of the genome which is required for efficient assembly into progeny virion (Whelan & Wertz, 1999a). However the exact mechanism by which the nucleocapsids are selected for assembly is not known. Following assembly, virus budding occurs. M protein recruits cellular proteins at the site of budding for final release of virus particles (Harty et al., 1999). Proline rich sequence in the M protein has been identified to be responsible for budding (Irie et al., 2004a). Ubiquitination of M
protein was shown to be required for final release of virions (Jayakar et al., 2000). Cellular proteins involved in membrane trafficking such as TSG101 or VPS4 have been implicated in budding of HIV, but are not required for budding of VSV (Irie et al., 2004b).

1.7 Defective interfering particles of VSV

Defective interfering particles (DI particles) are sub-genomic virus particles generated by large deletion of virus genome during the process of replication (Schubert et al., 1979). Because DI particles do not encode for any protein, replication of their genome depends on the wild-type virus to provide the replication proteins. Several classes of defective interfering particles for VSV have been well characterized (Meier et al., 1984; Perrault & Leavitt, 1978a; b). Because the DI particle genomes are shorter and they contain a stronger promoter for replication, they inhibit the replication of wild type virus (Pattnaik et al., 1995; Perrault & Holland, 1972). In laboratory conditions, DI particles can be generated by a high multiplicity of infection of wild type virus. The sizes of these DI particles are also small; therefore, they can be easily separated from the wild type virus by density-gradient centrifugation. Because of the above unique properties, during the 1970s and 1980s, DI particles provided an excellent tool to study the process of virus replication and much of our understanding of VSV replication comes from studies with DI particles.

1.8 Reverse genetics system of VSV

Genetically engineered non segmented negative strand RNA viruses offer a unique tool to investigate several aspect of virus biology. For positive strand RNA viruses, recovery of infectious virus from the cDNA was established in the early 1980s
(Racaniello & Baltimore, 1981). However for negative strand viruses, because of the typical replication cycle of these viruses, development of reverse genetics system faced several hurdles. The first system to recover an infectious viral genome was established with VSV DI particles (Pattnaik et al., 1992). In this system, cDNA clones expressing N, P, L, M and G proteins were transfected along with cDNA clone of DI genome. The cDNAs were under the control of bacteriophage T7 promoter. Thus the expression of the viral proteins and the DI RNA was carried out by the bacteriophage T7 RNA polymerase, which was provided by infecting the cells with a recombinant vaccinia virus expressing the T7 RNA polymerase (Pattnaik et al., 1992). Use of a hepatitis delta virus ribozyme to generate authentic 3’ end was required for replication of the viral genome (Pattnaik et al., 1992). Subsequently infectious full length rabies and VSV were generated following the above system (Lawson et al., 1995; Schnell et al., 1994). Negative strand RNA viruses from other families such as paramyxovirus, Bornavirus, Ebolavirus were also recovered entirely from cloned DNA using similar approach (Radecke et al., 1995; Schneider et al., 2005; Volchkov et al., 2001). With the advent of reverse genetics system, it has been possible to engineer changes in viral genomes as well as to sub-genomic replicons to probe the role and requirement of cis-acting signals in viral genome replication and transcription (Conzelmann, 1998).

1.9 Host factors required for VSV infection

Viruses are obligate intracellular parasites. Viruses such as VSV encode for limited number of proteins and the functions of those proteins are limited to replicate the genomes. Many steps in the life cycle of VSV such as attachment, entry and uncoating,
genome replication and assembly depend on specific host cell functions. Virus and host interaction is also complex. Host has developed strategies to eliminate the invading viral pathogen, whereas viruses have developed strategies to counter the host defense systems. Viruses have also historically been used to discover and characterize several cellular mechanisms such as DNA replication, membrane fusion, RNA capping, internal ribosomal entry during translation etc. Comprehensive knowledge about such critical players in virus lifecycle will provide a valuable understanding about molecular mechanism of virus infection. Such understanding will also facilitate generation of novel drugs that target an essential host protein involved in virus replication and will help to avoid selection of drug-resistant mutant viruses. Until recently global analysis of host proteins required for virus infection was hindered by non-existent of suitable system. With the advancements in genomics and knowledge of RNA interference mechanisms, it has been possible to develop genome-scale loss-of-function screens to identify host factors that influence virus infection. In the past, VSV has led the pathway to understand molecular aspect of virus genome replication and transcription as well as immune recognition. However little is known about the host factors involved not only in VSV but also in other negative strand RNA virus infections.
Figure 1.4: Schematics of reverse genetics system of VSV

BHK-21 cells are infected with recombinant vaccinia virus expressing T7 RNA polymerase. Subsequently these cells are transfected with full length cDNA clone of VSV and support plasmid expressing viral N, P and L proteins. At 48 hrs post transfection, supernatant is collected and passaged onto naive BHK-21 cells. Recovery of virus is evidenced by cytopathic effects and immunofluorescent staining for viral M protein. Also, using this system replication of DI genome can be examined when DI particles are infected to cells expressing viral N, P and L proteins.
Recently, elegant studies by Gruenberg and colleagues demonstrated that uncoating and delivery of VSV genome occur by a back fusion event inside the multivesicular endosomes and host proteins Alix and TSG101 are required for this event (Le Blanc et al., 2005; Luyet et al., 2008). VSV transcriptase contains cellular guanylyl transferase, heat shock protein 60 and eukaryotic elongation factor 1-alpha along with viral N, P and L protein (Qanungo et al., 2004). The exact role of these host factors in viral replication and transcription is not known. Involvement of other cellular factors in VSV replication is also not known at this time. VSV matrix protein contains late domain and utilizes host cell factors such as Nedd4 like ubiquitin ligase (Harty et al., 1999) for the release of the assembled virus. In contrast to HIV and Ebola virus, VSV budding does not require activity of ESCRT III complex (Irie et al., 2004b). Involvement of other proteins in the budding of VSV is not known.

A limited number of restriction factors are also known to modulate VSV and other negative strand RNA virus infection. Innate immune system is the first line of defense against the invading pathogen. In an in vitro system, using VSV and Rabies RNA, cytoplasmic pattern recognition receptor for RNA such as Retinoic acid inducible gene I (RIG-I) was identified (Hornung et al., 2006; Pichlmair et al., 2006). Recognition of RNA by the cytoplasmic sensors leads to induction of interferon signaling pathways. Subsequent studies have identified the ligand for the RIG-I as base paired RNA with 5’ triphosphate (Schlee et al., 2009; Schmidt et al., 2009). RNAi interference was also recently identified as an evolutionarily conserved antiviral mechanism in insects, plants and animals (Ding & Voinnet, 2007). VSV was instrumental as a model pathogen to decipher the antiviral mechanism mediated by RNA interference in C. elegans (Wilkins
et al., 2005) Drosophila (Sabin et al., 2009) and mammals (Otsuka et al., 2007). The above studies have identified several key elements in the RNAi pathway acting as host resistant factors for VSV. Nevertheless the molecular mechanism by which VSV subverts these antiviral defense pathways is poorly understood.

1.10 Sub-genomic replicon system to study viral replication

Understanding the virus-host interaction and identifying host proteins involved in virus replication pose a major challenge in the field of virology. Sub-genomic replicon bearing cell lines harboring a drug resistant marker for several positive strand RNA viruses have been instrumental to study different aspects of virus biology such as viral genome replication, identification of host cell factors in virus replication, innate immune regulation and identification of drug candidates. Genomes of positive strand RNA viruses are infectious and thus it has been possible to establish sub-genomic replicon system by deleting structural gene and replacing a drug resistance gene. Cell lines continuously supporting replication of such replicons have been reported for several positive strand RNA viruses such as Bovine viral diarrhoea virus (BVDV), hepatitis C virus (HCV), dengue virus, west nile virus (WNV) and sub acute respiratory syndrome corona virus (Behrens et al., 1998; Ge et al., 2008; Lohmann et al., 1999; Ng et al., 2007; Rossi et al., 2005). Similar to establishment of reverse genetics system for negative-strand RNA viruses, development of replicon system for negative- strand RNA viruses is difficult because the RNA genome itself is not infectious and the replication and transcription of genome requires assembly into a functional nucleocapsid complex. Consequently, research in negative strand RNA viruses is lagging behind.
1.11 High-throughput siRNA screen to identify host factors required for virus infection

RNA interference was recently identified as a mechanism of post transcriptional gene silencing where the RNAi machinery knocks down the endogenous or exogenous RNA in a sequence dependent manner (Fire, 1999; Fire et al., 1998). Completion of human genome sequencing project has provided new impetus to the field of genomics (Lander, 2011). Coupling RNAi with knowledge of human genome, researchers have developed tools to silence gene functions. Also, advances in bioinformatics analysis and high-throughput cell-based screening and imaging platforms have made it feasible to probe the impact of each and every human gene not only in viral system, but also in other pathophysiological conditions (Mohr et al., 2011). RNAi experiments offer loss-of-function studies. For example, if a particular gene is involved in a virus life cycle, RNAi knockdown of that gene gives an altered phenotype after virus infection. In the last five years genome-wide siRNA screens using mammalian cell-based systems have led to several exciting discoveries in many biological conditions previously seemed impossible (Mohr et al., 2011). First genome-wide RNAi screen to identify human host factors involved in a virus infection was reported for HIV (Brass et al., 2008). Subsequently genome wide RNAi screen has been reported for several other viral systems such as influenza (Brass et al., 2009; Karlas et al., 2010; Konig et al., 2010), West Nile virus (Krishnan et al., 2008), HCV (Li et al., 2009; Tai et al., 2009). However, before the arrival of human genome-wide siRNA screens, drosophila siRNA screen provided researchers the opportunity to expand the knowledge of virus-host interaction. Drosophila RNAi is robust because there is no interferon system which provided simplicity for data
analysis and in addition knockdown of genes is efficient. Most of the Drosophila genes have their human orthologs and thus data obtained from Drosophila screens can be extended to human system. Easy genetic manipulation to generate knockout flies in short period of time offer a unique advantage to study Drosophila as a model organism. Studies have been published using Drosophila screen, identifying host proteins involved in virus infection as early as 2004 (Cherry et al., 2005; Cherry & Perrimon, 2004).

Though the RNAi technology has helped the researchers to achieve critical milestones, the field of RNAi is still in its infancy. Multiple groups have completed RNAi screen in the same biological system. Ironically the results are dissimilar from screen to screen (Bushman et al., 2009; Goff, 2008). Several factors such as types of reagents used, type of cell line, assay, degree of depletion of gene and statistical analysis could contribute to the disparate results. Recently, there has been an increasing effort to identify the reproducibility between the genomic screens and to better understand the RNAi technology (Barrows et al., 2010). In a high-throughput screen, there are off target effects resulting in false positives. Many bioinformatics tools are available to design effective RNAi reagents, yet the available reagents are not 100% effective in knocking down the targets (Tilesi et al., 2009). Statistical analysis is also one important component of a high-throughput screen and the type of analysis used could greatly influence the results (Barrows et al., 2010; Birmingham et al., 2009). A new trend of coupling other “omics” technologies such as protein-protein interaction map and mRNA microarray with the RNAi is also emerging (Mohr et al., 2011). Nevertheless, researchers could harness the power and utility of RNAi technology for several landmark discoveries in wide variety of fields in the past several years. Though much of our current knowledge on the
biological system exists at the gene level, emergence of new technologies as well as refinement of existing technologies of integrative screening and data analysis will provide new dimension to the research field. Additional genome-scale data collection will shift the paradigm towards a system level understanding rather than at gene level understanding.

1.12 Overall objective and experimental plan

Identification of host genes affecting virus replication will be valuable to understand virus biology and virus-host interaction. Negative-strand RNA viruses include several human and animal pathogens. Many of these viruses cause significant cytopathic effects in cultured cells after infection. For some of these viruses, specific containment facility is required to study the virus biology. Development of a system which supports replication of sub genomic replicon will be a major impetus for studies on virus-host interactions and drug discoveries and thus the need for specific containment can be alleviated. The overall objective of this study was to establish a replicon system for VSV as well as to identify host factors required for VSV infection in cell culture system.

1.12.1 Objective 1: Hypothesis and experimental approach

Development of replicon systems for positive strand RNA viruses has greatly contributed to our understanding of virus biology. Development of such replicon system poses a major challenge for negative-strand RNA viruses because the RNA of negative-strand RNA viruses exist as nucleocapsids bound to viral proteins and must be transcribed to produce proteins for replication. Availability of such a replicon system for
VSV will help to facilitate a detailed understanding of complex nature of virus-host interaction as well as rapid drug discovery.

Since VSV genome replication depends on presence of the replication proteins such as N, P and L, it was hypothesized that cells constitutively expressing replication proteins of VSV can be used to replicate a minimal replicon of VSV encoding a drug resistant marker or defective interfering particle genome of VSV. Also a modified sub-genomic replicon of VSV lacking M and G proteins can be replicated in the cells and if positive selection is applied a cell line can be developed harboring the replicon. Based on these hypotheses, the objective was to establish a cell line supporting continuous replication of a sub-genomic replicon of VSV. To achieve this objective, two different approaches were used. In the first approach, a cell line which constitutively expresses VSV replication proteins was generated. Long term replication of DI genome was attempted to establish in this helper cell line. In the second approach, sub-genomic replicon of VSV where M and G protein coding regions were deleted from the full length cDNA was used to generate a cell line. This sub-genomic replicon also encoded a drug resistance marker [neomycin phosphotransferase gene (Neo\(^r\))] under VSV specific promoter. For initial round of replication, proteins expressed from the helper plasmids are required. It was expected that subsequent rounds of replication will then become independent of the proteins provided through the helper plasmid as this replicon will produce the proteins necessary to sustain its own replication. Cells were expected to confer resistance to the drug G418 when there is replication dependent expression of Neo\(^r\).
1.12.2 Objective 2: Hypothesis and experimental approach

During the course of infection, viruses exploit the sub cellular milieu and orchestrate several key events in order to replicate and propagate. Host cells provide the building blocks for viral metabolism as well as in some cases serve as platform for genome replication. Viruses usurp several cellular pathways to complete each step of their life cycle. Identification of these host cell functions has been a long standing interest in virology. So far there has not been significant effort to identify host factors required for VSV or any other non-segmented negative strand RNA viruses. To fill the gap in the knowledge, the objective was to identify the host proteins involved in VSV infection in cell culture condition. To achieve this objective, a genome-wide siRNA screen was employed to silence each of the known and putative genes encoded in the human genome. A total of 23,000 human genes were knocked down individually, and their effect on viral replication was interrogated using a high-throughput cell based assay. A total of four separate siRNAs targeting each gene using a format of two different pools, each containing two siRNAs was used. The hits generated after the primary screen were validated in a secondary screen using different set of siRNAs to minimize the off-target effects and false discoveries.
CHAPTER II
MATERIALS AND METHODS

2.1 Cell culture and maintenance of cell lines

Baby hamster kidney (BHK-21) cells were obtained from ATCC (CCL-10). BHK-21 cells were maintained in minimal essential media (MEM) containing 5% heat-inactivated foetal bovine serum (FBS) and supplemented with 100 units of Penicillin, 20 units of streptomycin and 20 units of kanamycin (1X PKS) per one ml of media. HeLa cells (ATCC: CCL2) and Human embryonic kidney (HEK 293), HEK-293T cells were maintained in Dulbecco Modified Eagle’s media with 10% heat-inactivated FBS and 1X PKS. Madin Darby canine kidney (MDCK) and Vero cells were obtained from Veterinary Diagnostic Center, UNL. BER40 cells were obtained from George Belov (University of Maryland).

2.2 Preparation of stock of VSV

Plaque purified single clone of viruses (VSV-eGFP, VSV PeGFP or wild type VSV) were grown in BHK-21 cells. Confluent monolayer of BHK-21 cells in 100 mm dishes were infected with 0.01 multiplicity of infection in 1 ml of media. After 1 hr of adsorption, unbound viruses were removed by washing with DMEM and fresh DMEM ith 2% FBS and 1X PKS was added. After 16-18 hr post infection, supernatant was collected, clarified by centrifugation and stored in -80°C in small aliquots. The titers of these viruses were determined by plaque assay using BHK-21, HeLa and HEK-293 cells.

2.3 Preparation of stock of LCMV and HPIV3

GFP expressing LCMVr3-GFP (Emonet et al., 2009) was grown in BHK-21 cells. Fluorescent focus forming unit was calculated after infecting HeLa cells. GFP expressing
HPIV3-GFP (Zhang et al., 2005) was grown in HeLa cells. Fluorescent focus forming unit was calculated after infecting HeLa cells with HPIV3-GFP.

2.4 Preparation of recombinant vaccinia virus (vTF7-3) stock

Stock of a recombinant vaccinia virus expressing bacteriophage T-7 polymerase (Fuerst et al., 1986) was generated in BHK-21 cells using a protocol as described previously (Pattnaik & Wertz, 1990). The titer of the stock vTF7-3 was determined by plaque assay in BHK-21 cells.

2.5 Stocks of defective interfering particles

Defective interfering particles (DI) particles used in these studies were DI-T particles with a panhandle type genome (Meier et al., 1984; Pattnaik & Wertz, 1990). Stocks of these DI particles were generated as described earlier (Pattnaik & Wertz, 1990).

2.6 Plaque assay

Plaque assay for VSV was conducted in BHK-21 and HeLa cells. A 10-fold serial dilution of the virus stock was prepared in cold DMEM. BHK-21 cells or HeLa cells in 12-well tissue culture plates were infected with 100μl of the diluted virus stock as inoculum per well. After 1 hour of adsorption, unbound viruses were removed by washing two times in DMEM. Equal mixture of 1.2% low melting point agarose and 2X DMEM containing 10% FBS and 2X PKS was prepared and added to the monolayer. Cells were incubated undisturbed inside the hood for 15 minutes allowing the agar to solidify. Cells were further incubated at 37°C for 16-20 hours for plaques to develop. Subsequently, cells were fixed (2% glutaric dialdehyde in PBS) for 1 hr. Agar plugs were removed and fixed plaques were stained with 0.1% crystalviolet in 30% methanol.
Numbers of plaques were counted for each dilution and virus titer was determined by multiplying the number of plaques with dilution factor.

### 2.7 Nucleocapsid preparation and transfection

VSV-eGFP was grown in BHK-21 cells in four 100 mm tissue culture plates. At 16 hour post infection, approximately 6 ml of supernatant was obtained per plate. The supernatant was clarified at 2000 RPM for 15 min at 4°C to remove the debris. Approximately 11 ml of supernatant was loaded onto 2 ml of 15% sucrose cushion (15% sucrose in NTE) per tube and pelleted by ultracentrifugation for 2 hours at 36,000 rpm using a SW41 rotor. The pellet was resuspended in 2 ml of nucleocapsid generation buffer (10 mM Tris-HCl pH 8.0, 0.4 M NaCl, 1.85% Triton X, 5% glycerol, 0.6mM DTT) per tube and kept on ice for 2 hrs with intermittent shaking. The nucleocapsids from both the tubes were pooled, loaded on a 15% sucrose cushion made with nucleocapsid generation buffer and pelleted by centrifugation at 36,000 rpm for 4 hrs. The pelleted nucleocapsid was then resuspended in 100 μl resuspension buffer (5% glycerol and 10 mM Tris-HCl, pH 8.0), aliquoted and frozen in -80°C. Long-term storage resulted in approximately 50% loss of activity. Hence, nucleocapsids were used within one week after generation.

Nucleocapsids were transfected using Lipofectamine 2000 (Invitrogen). To 100 μl OptiMEM (Invitrogen) thawed at room temperature, 4 μl of lipofectamine was added and the mixture was incubated for 5 minutes. In a separate tube, 4 μl of nucleocapsids were added to 100 μl of OptiMEM. The diluted nucleocapsid and lipofectamine mixtures were combined together and incubated at room temperature for 20 minutes for complex formation. HeLa cells grown in one well of a 6-well plate were washed once in DMEM
and 600 µl of OptiMEM without antibiotic was added. To this, the nucleocapsid and
Lipofectamine complexes were added drop wise, mixed gently and incubated for 6 hours.
Absence of virus particles was examined in the nucleocapsid preparation by adding the
nucleocapsids without Lipofectamine and monitored for cytopathic effect.

2.8 Molecular cloning

2.8.1 Preparation of chemically competent bacteria

Chemically competent *Escherichia coli* (E. coli) DH5α was prepared by Inoue
method (Sambrook & Russell, 2001). Competent bacteria were frozen in small aliquots in
-80°C.

2.8.2 Recombinant DNA technique

Restriction endonuclease enzymes were purchased from New England Biolabs
(Ipswich, MA). Restriction digestions of the plasmids or PCR products were done with
the restriction endonucleases as per the manufacturer’s recommendations. Purification of
the digested fragment from the low melting point agarose gel was done using phenol and
chloroform method (Sambrook & Russell, 2001). Purified DNA was resuspended in
water. DNA ligation was performed by ligating the insert and vector with T4 DNA ligase
(Invitrogen). Vector and insert ratio was kept 1:3 to 1:5. Ligation reaction was carried
out by incubating the mixture at 2-4 hours at room temperature or 16-18 hours at 16°C.
For transformation, 5µl of ligation mix was added to 50 µl of competent DH5α. Mixture
was incubated for 30 on ice. Bacteria were heat shocked at 42°C for 90 seconds followed
by incubation for 4 min on ice. Subsequently, 950 µl of room temperature LB broth was
added and further incubated at 37°C for 1 hour. Finally bacteria were plated on LB agar
plates containing appropriate antibiotic for selection of transformants containing recombinant plasmid.

**2.8.3 Plasmid purification**

Plasmids were purified from the bacteria in small scale by alkaline lysis method. Individual bacterial colonies were grown for 8 hr or 16 hr in 1.5 ml LB broth containing appropriate antibiotic (Ampicillin or Kanamycin or Chloramphenicol). Pelleted bacteria was lysed by 300 μl of TENS buffer (10mM Tris-Hcl pH 8.0 1mM EDTA, 0.1M NaOH, 0.5% SDS). The cell suspension was neutralized by adding 150 μl of sodium acetate, pH 5.2. Supernatant was collected into a fresh microfuge tube after pelleting the cell debris. Plasmid DNA was precipitated by adding 900 μl of 100% ethanol followed by centrifugation for 5 minutes. Precipitated plasmid DNA was dried after washing with ice cold 70% ethanol. Dried plasmid was resuspended in 50 μl of tris-EDTA buffer containing 100 μg/ml of RNaseA. Large-scale plasmid purification was performed by gravity flow column purification method using Qiagen’s column Tip 100 as per manufacturer’s instructions.

**2.9 Construction of plasmids**

**2.9.1 Construction of N, P, L plasmids under the CMV promoter**

pGEM-P, pGEM-N and pGEM-L plasmids, which encode the P, the N, and the L proteins of VSV, respectively, have been described previously (Pattnaik & Wertz, 1990). The plasmid pGEM-PeGFP, which encodes PeGFP fusion protein has been reported previously (Das et al., 2006). All these plasmids were in pGEM-3 vector (Promega) under the control of T7 RNA polymerase promoter. For expression of these proteins by cellular RNA polymerase II (Pol II) promoter, the entire protein coding sequences from
the above plasmids were released by restriction digestion and sub-cloned in plasmids where the expression of the viral proteins is under the control of cytomegalovirus immediate early (CMV IE) promoter. The N, P, and the L coding sequences released from the respective plasmids were cloned in pHygEGFP vector (Clontech, Mountain View, CA), replacing the hygromycin-eGFP fusion protein coding sequences. These plasmids were termed pHE-N, pHE-P and pHE-L, respectively. Additionally, PeGFP, P, and L coding sequences were sub-cloned in pcDNA 3.1-neo vector (Invitrogen) under CMV promoter. These plasmids were named pc-PeGFP, pc-P, and pc-L, respectively.

2.9.2 Construction of a plasmid expressing N, P and L

To construct a plasmid encoding all three proteins (N, PeGFP, and L) of VSV in tandem, the coding sequences of P along with upstream CMV promoter and downstream poly (A) signal sequences were released from pHE-P plasmid by digesting with BglII and NaeI and this fragment was made blunt by treating with Klenow (New England Biolabs). This blunt-ended fragment was inserted in pHE-N plasmid at NaeI site, which is located downstream of poly (A) signal sequences and the resulting plasmid was named pHE-NP. The coding sequences of P in pHE-NP were replaced with coding sequences of PeGFP from pGEM-PeGFP. The resulting plasmid was termed pHE-NPeGFP. Subsequently, the N and PeGFP coding sequences with their corresponding CMV promoters and poly(A) signal sequences were released from pHE-NPeGFP by digestion with DraIII and XmnI and inserted at the unique BstZ17I site located downstream of the SV40 poly(A) signal sequences in the pc-L plasmid. This final plasmid was termed pc-NPeGFPL, which contains the coding sequences for all the three VSV proteins. The expression of all the three VSV proteins is under the control of independent CMV promoters and poly (A)
signals in the pcDNA vector that also carries the neomycin phosphotransferase gene for selection of transfected cells with G418.

2.9.3 Construction of mini replicon of VSV

Previous works from our lab reported the construction and use of positive sense mini replicons of VSV (Das & Pattnaik, 2005). To circumvent the issue of transcription from cryptic promoter, negative sense minireplicon with eGFP as reporter gene was constructed (Nayak et al., 2009). The eGFP mini replicon was under the control of T7 promoter. Two unique restriction sites (*BsiWI* and *KasI*) were engineered in this negative sense mini replicon to place two different reporter genes. These restriction sites flank VSV gene specific initiation and termination signals as well as leader and trailer sequences. To generate authentic 5’ and 3’ termini after transcription by CMV promoter, hammerhead ribozyme and hepatitis delta ribozyme were placed before trailer and after leader sequences respectively. The base replicon without any reporter gene to pHygeGFP plasmid was transferred to pHyg-eGFP vector where the transcription of the minireplicon was under CMV promoter. *EcoRI* and *NotI* enzymes were used to cut the Hyg-eGFP portion from the vector plasmid and base replicon was cloned back at these sites. The plasmid “p-base replicon” does not contain any gene in both the *BsiWI* and *KasI* sites is considered as empty base replicon. The hygromycin gene was PCR-amplified with KasI site and cloned into the *KasI* site in the base replicon. This plasmid was named as p-Hyg replicon. Subsequently the mRFP gene was amplified by PCR and cloned into the *BsiWI* site. The resultant plasmid was named as p-RFP-Hyg Replicon. Sequences and orientations were then verified by sequencing the plasmid.
2.9.4 Construction of a sub-genomic replicon of VSV

The VSV sub-genomic replicon was also constructed based on the negative sense minimal replicon. In this sub-genomic replicon, VSV-G and M protein coding sequences were removed to disable infectious virus production. Full length VSVPeGFP construct generated in our laboratory (Das et al., 2006) was used as a template for further modification. The M and G coding sequences together were deleted from VSVPeGFP using SnaBI and NheI, and Neomycin phosphotransferase gene was cloned using these sites. Sequence of NeoR was verified by sequencing. To generate a negative sense replicon, N-PeGFP-Neo sequences were PCR amplified and cloned into the BsiWI site of the negative sense base replicon. The resultant plasmid was named as p-N-PeGFP-Neo replicon. L coding region was cut from the plasmid and made blunt. p-N-PeGFP-Neo replicon was digested with KasI, made blunt and blunt ended L fragment was inserted into this site. Orientation was verified by restriction enzyme digestion and sequencing. This final construct was termed as pN-PeGFP-Neo-L replicon.

2.10 Plasmid DNA transfection

Plasmids were transfected using Lipofectamine 2000 (Invitrogen) or Fugene (Roche) according to manufacturers’ instructions.

2.11 Establishment of stable cell line expressing N, PeGFP and L

HEK-293 cells in 35mm tissue culture plates were transfected with pc-NPeGFPL plasmid or empty vector pcDNA3.1-Neo using Lipofectamine 2000. After 48 hours of transfection, cells were trypsinized and 25% of the cells were plated in 60 mm cell culture dishes. After 16 hours, the cells were subjected for drug selection with 1mg of G418 per ml of growth medium. Media in the culture dishes were replaced every 48
hours with fresh media containing the same concentration of G418 until drug resistant cells appeared. Empty-vector transfected G418 resistant cells not expressing the viral proteins served as control cells. Roughly 15% to 20% of cells in the culture transfected with pc-NPeGFPL plasmid expressed green fluorescence as judged by fluorescence microscopic examination of the culture. The green fluorescent cells were enriched by multiple rounds of growth and cell sorting by fluorescent assisted cell sorter (FACS, FACSVANTAGE SE, Becton Dickinson) until greater than 90% of the cells in the culture expressed green fluorescence. Subsequently, for isolation of stable cell clones, individual cells positive for PeGFP expression were sorted by FACS into 96-well tissue culture plates. Individual cells were then grown, expanded, and analyzed by immunofluorescence for expression of the N and the L proteins. Several cell clones expressing all the three proteins were then identified for further studies.

2.12 Metabolic labeling and immunoprecipitation of protein

HEK-293 cells grown in 60mm cell culture dishes were transiently transfected with appropriate plasmids. At 48 hours post-transfection, cells were starved in cysteine and methionine free media for 1 hour and radiolabeled with 50 μCi of Expre$^{35}$S$^{35}$S protein labeling mix per ml of cysteine- and methionine-free medium for 6-8 hours. Preparation of cell extracts, immunoprecipitation with various antibodies, sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis (SDS-10%PAGE) analysis, and detection of proteins by fluorography were carried out as described previously (Pattnaik & Wertz, 1990). Stable cell clones expressing the viral proteins were grown in 6-well culture plates and the proteins were radiolabeled for 6-8 hours as described above and detected by immunoprecipitation and SDS-PAGE.
2.13 Metabolic labeling and analysis of RNA after DI particle infection

NPeGFPL stable cells grown on 60mm culture dishes were infected with DI particles of VSV (5 µl of DI stock per 60 mm dish) and immediately radiolabeled with 50 µCi \( ^3 \)H-uridine per ml of medium for 20 hours. Cell extracts were immunoprecipitated with anti-N antibody, radiolabeled RNAs were recovered from immunoprecipitated complexes by phenol:chloroform extraction and analyzed by electrophoresis in acid-agarose-urea gels, and detected by fluorography. DI RNA replication in BHK-21 cells transiently expressing the viral proteins by T7 RNA polymerase was performed as described previously (Pattnaik & Wertz, 1990).

2.14 Luciferase assay

Stable cells expressing N, PeGFP, and L proteins or control cells not expressing the viral proteins were co-transfected with 0.4 µg of IFN\( \beta \)-Luc or ISRE-Luc or pGL4.32 (Luc2p/NFκB-RE/Hygro) along with 10 ng of pRL-TK using Fugene 6 (Roche Applied Science) as per manufacturer’s protocol. The pRL-TK vector (Promega), which contains renilla luciferase reporter gene under the control of herpes simplex virus thymidine kinase (TK) promoter, serves as an internal control for transfection efficiency. At 24 hours post-transfection, the cells were either mock-infected or infected with DI particles. At 16 hours post-infection, the cells were lysed in lysis buffer (Promega) and luciferase assay was performed using Dual Luciferase assay kit (Promega). Luciferase activities were expressed as fold change over control uninfected cells after normalizing with renilla luciferase activity.

2.15 Reverse transcription (RT) PCR to detect DI particle genome
For RT-PCR detection of DI RNA, DI particle infected 293-NPeGFPL cells grown in six well plates were lysed in Trizol (Invitrogen). RNA was extracted as per manufacturer’s recommendation and resuspended in 10 μl of water. Four μl of RNA was used for reverse transcription in total 10 μl volume. Superscript III (Invitrogen) was used for reverse transcription. The reaction mixture contained 4.0 μl of RNA, 0.5 μl of SS III RT, 1 pmole of primer, 2 μl of RT buffer, 0.25 μl of RNAsin (New England Biolab). First-strand synthesis was carried out at 50°C for 1 hour. Because both 3’ and 5’ termini sequences are complementary, only one primer was used for first strand synthesis as well as PCR amplification of the cDNA. The primer sequence is 5’ACGAAGACCAC AAA ACCAGATAAAAA 3’. The PCR amplification was carried out using Taq DNA polymerase (New England Biolab). The reaction mixture contained cDNA 5 μl, 10 X Taq polymerase buffer 5 μl, Primer 4 μl (4 pmole), 1 μl of Taq DNA polymerase in a 50 μl reaction condition. The cycling conditions were as follows; initial denaturation at 94°C 5 minutes, denaturation at 94°C for 30 sec, annealing 52 °C for 30 sec, extension 72°C for 2 minutes 30 sec for 35 cycles and final extension at 72°C for 10 minutes.

2.16 Quantitative RT PCR

Quantitative reverse-transcription PCR (qRT-PCR) was used to measure the VSV P mRNA as well as genome and anti-genome in experiments involving siRNA mediated gene knockdown. Total RNA was extracted from one well of a 12-well tissue culture plate by Trizol (Invitrogen) according to manufacturer’s protocol. First-strand cDNA was synthesized using M-MLV RT (Invitrogen) and 200 ng of total RNA was used in 20 μl reaction volume. For quantification of P mRNA and beta-actin mRNA, 1 μl of oligo – (dT) was used. For anti-genome, VSV 2955-R primer (2pmole/ μl) and B-actin1 R primer
(2pmole/ μl) was used. First strand synthesis was performed according to M-MLV RT protocol provided by the manufacturer. The sequences of the primers and probes are given in Table 2. 1. qRT-PCR was optimized with different concentrations of primers and probes in Cepheid Smart cycler (Cepheid, CA, USA). Final concentrations of primers and probes are listed in Table 2. 2. The thermal set up for the PCR was done as follows, initial denaturation was done at 95°C for 5 minutes, denaturation at 95°C for 30 seconds, annealing and extension at 60°C for 30 sec. Total amplification cycles were 40. Fold change was obtained using a ΔΔCT method after normalizing with the non targeting siRNA treated sample.

2.17 Immunoblotting

For western blotting, cells (12 well or six well format) were washed with cold PBS once. Cells were collected in cold PBS after scrapping with a cell scraper and pelleted at 10000 RPM for 5 minutes. Supernatant was aspirated out and cells were pelleted and lysed using 20 μl RIPA buffer (10Mm Tris pH 8.0, 140 mM sodium chloride, 1.0% Triton X 100, 1.0% sodium deoxycholate, 0.1% sodium dodecyl sulfate along with protease inhibitors such as 2mM PMSF, 1mM Leupeptin). To facilitate lysis, cells were frozen and thawed, vortexed and again pelleted at 13000 RPM for 20 minutes. Supernatant was collected into a fresh tube. Protein was quantified using Bradford reagent (Biorad, Catalog # 500-0006). Equal amounts of proteins were separated by SDS-polyacrylamide (SDS-12% PAGE) gel except for GBF1 experiment where 4-20% gradient gel (BioRad) was used. After separation, proteins were transferred to a PVDF membrane by semi-dry transfer or wet transfer.
### Table 2.1: Sequences of Primers and Probes

<table>
<thead>
<tr>
<th></th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSV2795F</td>
<td>GTGACGGACGAATGTCTCATAA</td>
</tr>
<tr>
<td>VSV2860R</td>
<td>TTTGACTCTCGCTGATTGTAC</td>
</tr>
<tr>
<td>VSV2955R</td>
<td>TGATGAATGGATTGGGATAACA</td>
</tr>
<tr>
<td>BActin-F</td>
<td>CAAGTACTCCGTGTGTTGAT</td>
</tr>
<tr>
<td>BActin-R</td>
<td>CATACTCCTGCTTGCTGAT</td>
</tr>
<tr>
<td>VSV 2825 probe</td>
<td>6-FAM/CCATCCTGCTCGCTGAGATAC/TAMRA</td>
</tr>
<tr>
<td>B Actin probe</td>
<td>Cy5/TCGCTGCCACTCCAGCAGAT/BHQ2</td>
</tr>
</tbody>
</table>

### Table 2.2: Optimized concentrations of primers and probes

<table>
<thead>
<tr>
<th></th>
<th>Primers</th>
<th>Primer Concentrations</th>
<th>Probe</th>
<th>Probe Concentrations</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSV P mRNA</td>
<td>VSV2795-F</td>
<td>500 nM</td>
<td>VSV2825</td>
<td>200 nM</td>
</tr>
<tr>
<td></td>
<td>VSV2860-R</td>
<td>500 nM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSV Anti-genome Or Genome</td>
<td>VSV2795-F</td>
<td>1000 nM</td>
<td>VSV2825</td>
<td>200 nM</td>
</tr>
<tr>
<td></td>
<td>VSV2955-R</td>
<td>1000 nM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human B-actin</td>
<td>B-actin-F</td>
<td>500 nM</td>
<td>B actin Probe</td>
<td>1000 nM</td>
</tr>
<tr>
<td></td>
<td>B-actin-R</td>
<td>500 nM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Membranes were blocked with blocking buffer (5% skim milk with PBS and 0.05 % Tween 20). Membranes were probed with primary and corresponding secondary antibodies. The same membranes were stripped and re-probed with another primary antibody. VSV M monoclonal antibody was used to detect viral M protein. Anti-actin antibody was used to detect beta-actin as a loading control.

2.18 High-throughput assay for primary screen

2.18.1 siRNA Library Design

The screen was performed at the Duke University using the Qiagen genomic siRNA library v 1.0. The genomic siRNA library was arrayed in 384-well micro plate format. The library consisted of 4 distinct siRNAs (A, B, C and D) targeting 22909 known and putative human genes. Four siRNA were divided into pool of 2 with each pool containing 2 unique siRNA duplexes (set AB and set CD). This format resulted in 74 micro plates for each set and total 148 plates. This 2 X 2 pool design allowed each gene to be tested by two independent siRNA sets. All stars non targeting siRNA (Qiagen) was used as a negative control which did not affect the levels of any of the known or putative human gene products. Also siRNAs targeting to N and L gene of VSV were synthesized and used as additional control siRNA to verify transfection efficiency. The sequences of the siRNAs targeting VSV genes are as follows

N1 5’ CUGCAAGGCCUAAGAGAGA 3’
N2 5’ UGGAAUACCCCGGCAGAUUA 3’
L1 5’ GCAGUUAUCCAGCAUAUCAU 3’
L2 5’ GAGAAACGUUGUAAGAAUU 3’
All the control siRNAs were manually added to the designated wells during screen.

2.18.2 Primary screen assay optimization

Assay optimization was the longest part of the screen. Corning 384-well black tissue culture plates (Catalog # 3712) were pre-arrayed with 1 p mole of siRNA using a velocity Bravo liquid handling system (Agilent technologies, CA). Reaction conditions for the screen such as the cell line, cell number, siRNA concentration, amount of Lipofectamine RNAi MAX to be used, control siRNAs to be used in the screen, multiplicity of infection (MOI), incubation period after infection, desired percent infection were optimized empirically by series of experiments conducted at the Duke University RNAi Screening center. A test screen with one genomic siRNA library plate and a pilot screen with three sets of genomic screen plate (three plates from AB pool and corresponding three plates from CD pool) were conducted to assure the validity of the screen. HeLa cells (CCL2, ATCC, Manasass, VA) was used for the assay. After optimization, 3000 cells per well was used for transfection. Transfection of the cells was done with 15.4 nM final concentration of siRNAs in a 65 μl of total volume. Lipofectamine RNAiMAX (Invitrogen, CA) was used as transfection reagent in the amount of 0.05μl per well in 10 μl of OptiMEM (Invitrogen, CA). Each ptate contained siRNAs targeting VSV N and L gene as control siRNAs to determine transfection efficiency. VSV-eGFP virus grown in BHK-21 cells and titrated in HeLa cells was used at 0.5 MOI. This particular MOI when infected in 25 μl volume in one well of a 384 well plate, resulted in approximately 60% infection after 18 hrs of infection. To identify factors involved in all stages of VSV life cycle such as entry, uncoating, replication and budding, the screen was standardized with a multiplicity of infection and time to allow
multiple rounds of virus infection. During assay optimization, it was observed that 0.5 MOI of virus in 25 μl volume resulted in second-round of virus infection and gene expression after 10 hours post infection. Therefore, percent infection (60-70%) at 18 hours post infection was considered a result of multiple rounds of virus infection.

2.18.3 Automated determination of cell number and percentage infection

At the end of the infection, cells were fixed with 4% paraformaldehyde in PBS for 15 minutes, permeabilized with 0.1% triton-X 100 in PBS for 15 minutes and stained with Hoechst 33342 in PBS for 30 minutes. Stained cells were imaged with a Cellomics ArrayScan VT I automated microscope. Images were analyzed with vHCS Scan Target Activation software v 5.1.2 to identify infected cells. Cells without VSV-eGFP infection served as control population for background fluorescence. Four fields per well of a 384 well plate were imaged at 10X magnification. First, cells were identified by their nuclei staining in channel 1 of Cellomics. Cells those passed the criteria were analyzed for GFP expression in channel 2. GFP intensity was calculated for each and every individual cell in the field. Frequency distribution of GFP-intensity was plotted and compared between NT controls infected with VSV-eGFP and uninfected control. For selection of infected cells, gates were set by manual inspection of representative images from VSV-eGFP infected and uninfected cells. Cells appearing at least 5 standard deviations away from the average intensity of uninfected control population in the same plate or same batch were considered positive for GFP expression. Once the gate was determined, all the plates were subjected for data analysis by vHCS Target activation software. Finally, data was obtained using vHCS View software v 5.1.2 and the numbers of cells present in the
well were identified as “Valid Object Count” whereas the percentage of infection was determined as “% selected”.

2.19 Protocol for primary Screen

2.19.1 Transfection

Primary screen was conducted using HeLa cells (CCL2, ATCC). HeLa cells between 30 and 40 passages were used. Per day, 14 microplates were transfected and the procedures for screening all the 148 plates were completed within a span of 14 days. Cells were continuously cultured to maintain a supply of cells for the transfection. The following protocol describes the work flow for transfection for each day. Before thawing the assay plates, cells were examined for viability. Assay plates were thawed for 15-20 minutes at room temperature. Then the plates were centrifuged at 500 g for 4 minutes. Control siRNAs 5 μl each of All stars NT, N1N2, L1L2 (200 nM stock) were added to the designated wells manually. OptiMEM and Lipofectamine RNAi MAX mix was prepared in a sterile container. Ten μl of mix was added to each well using a Well Mate dispenser (Thermo Scientific). Plates were centrifuged at 300 g for 30 sec and incubated at room temperature for siRNA and Lipofectamine complex formation. While the plates were being incubated, cells were trypsinized, counted and resuspended as per the requirement. Diluted cells at a concentration of 3000 cells in 50 μl volume per well were added using the Well mate dispenser (Thermo Scientific). Cells were further incubated for 52 hrs to allow the knock-down of genes.

2.19.2 Virus infection

After 52 hours of transfection, VSV-eGFP infection was done at an MOI of 0.5 MOI. Virus suspension was prepared in DMEM 2% FBS and PKS in a sterile bottle.
Temperature of DMEM was approximately 18-20°C at the time of addition of virus. Transfected media from the wells were aspirated by the Biotek ELx 405 automated plate washer (Biotek, VT, USA). Virus inoculum (25 μl) was added per well with the help of a Matrix automated multichannel pipette (Thermo Scientific). Plates were centrifuged at 300 x g for 30 sec and infection was done for 18 hrs.

2.19.3 Fixing and staining

After 18 hrs of infection, infected media was aspirated using the Biotek ELx405 plate washer. Cells were fixed with 4% paraformaldehyde in PBS using a Matrix automated multichannel pipette (Thermo Scientific). After 15 minutes of fixation, cells were washed with PBS using the plate washer. Triton X-100 (0.1% in PBS) was used to permeabilize the cells. Cells were subsequently stained for 30 minutes with Hoechst 33342 to stain the nuclei. Finally all the wells were filled with 50 μl of PBS after washing the Hoechst stain and the plates were sealed and were imaged with the Cellomics ArrayScan.

2.19.4 Statistical analysis

Sum rank statistics was applied to identify the hits from the primary screen. Because there was significant toxicity associated with the knockdown of some genes, wells where the cell number was less than 800 (VOC less than 800) were discarded from further analysis. The average cell number in the non targeting siRNA treated wells was 3000. Below 800 cell number, the data analysis became erroneous and it could have potentially created difficulty for further validation of genes. After comparing control siRNA, batch-wise analysis (with plates used in one day) rather than genomic analysis was used to get the primary hits. The reason for this is that the non-targeting siRNA
treated controls were showing similar percent infection with a minimal standard deviation in a batch. There were batch to batch (i.e., day to day) variations in percent infection of non-targeting control samples and hence, batch wise analysis was more accurate in this context. Percent infection values for set AB and set CD were ranked from lowest to highest and given ranks from 1 to n (n= number of samples analyzed) respectively. Sum rank is summation of ranks for corresponding pairs of set AB and set CD. The \( p \) value for each pair was determined and finally samples were aligned according to the \( p \) value. Two hundred genes with the lowest \( p \) values were identified as primary candidate hits (\( p<0.01 \)). Another issue was emergence of numerous ribosomal subunit genes as potential candidates affecting VSV infection. Since ribosomes are required in general for cellular and VSV mRNA translation, their impact on virus infection was obvious. The ribosomal sub units were therefore omitted from the data analysis.

### 2.20 Validation screen assay design

The hits obtained by the primary screen had at least two independent siRNAs showing similar effect in VSV infection. In a high-throughput siRNA screen a gene is called a true hit if two independent siRNAs show similar phenotype. However in order to avoid false discoveries and minimize off target effects, siRNA from a different source was used for secondary validation. Thus the validation screen was done with Dharmacon On Target Plus pool of 4 siRNA (Dharmacon, CO, USA). Using siRNA from a different manufacturer, bias in siRNA design or target seed match can be avoided. The sequence of Dharmacon siRNAs did not overlap with that of Qiagen siRNAs. The assay plate format was also changed to 96-well format to increase the number of cells transfected and infected to increase the statistical power.
Six μl of pooled siRNA (500nM) was pre-arrayed per well in 96-well plates using the Biomek FX liquid handling system at UNMC Eppley Cancer Center (Omaha, NE). Before transfection, plates were thawed and centrifuged at 500 g for 2 minutes to collect the siRNA at the bottom of the wells. Lipofectamine RNAi MAX and OptiMEM mix was prepared (0.25 μl of Lipofectamine RNAi MAX in 20 μl OptiMEM per well). Lipofectamine mix was added to the siRNA by a multichannel pipette. Plates were centrifuged at 500 g for 1 minute and incubated at room temperature for 30 minutes. HeLa cells were trypsinized and counted. For VSV infection 15,000 cells per well were added to each well in 75 μl of DMEM with 10% FBS and 1X PKS to yield a final concentration of 30 nM siRNAs. For LCMV and HPIV infection, 10,000 cells per well were added in 75 μl of growth media. Cells were incubated for 50 hrs for knockdown of the genes. For VSV-eGFP infection, 0.05 MOI of virus was added in 100 μl of DMEM 2% FBS and 1X PKS per well using a multichannel pipette. Plates were centrifuged at 300 g for 30 sec and incubated for 16 hrs for the infection to progress. LCMVr3-GFP (0.05 MOI) and HPIV-GFP (0.2 MOI) was used to infect HeLa cells. Virus inoculum containing required MOI of viruses were prepared in 40 μl of DMEM (with 5% FBS and PKS) per well. After 1 hr of infection, 60 μl of DMEM with 5% FBS and 1X PKS was added to each well and further incubated for 41 hrs for the infection to progress. Row A of the plate was not infected and this served as uninfected control for image analysis. Infected cells were fixed as per the protocol described in the primary screen. Cells were stained with DAPI to stain the nuclei. Image analysis was done in Cellomics ArrayScan VTI at UNMC, Eppley Cancer Center to obtain the “valid object count” and “% infection”. Experiment was repeated four times for VSV and two times each for LCMV.
and HPIV. Candidates were identified using Wilcoxon Mann Whiney U test (Rank sum test) using median normalized value.

2.21 siRNA transfection in six well or twelve plates

Reverse transfection of siRNA was performed in 6-well or 12-well tissue culture plates. Final concentrations of siRNAs were 5nM (COPI sub units and GBF1) or 30 nM (ARF1). Required amount of siRNAs were plated in the 12- well of 6-well plates. Lipofectamine RNAiMAX in OptiMEM was prepared (2 μl of lipofectamine in 300 μl of OptiMEM per well) and added to the wells. Plates were incubated for 30 minutes for complex formation. HeLa cells were trypsinized, counted and 200,000 cells per well in 500 μl of DMEM 10% FBS+ PKS were added for COPI experiments. For GBF1 and ARF1 experiments 100,000 cells were added. Cells were further incubated for 42-44 hours for COPI depletion and 66-68 hours for GBF1 and ARF1 depletion.

2.22 Immunofluorescence

Cells were either fixed with 4% paraformaldehyde in PBS at room temperature for 15 minutes or cold methanol and acetone (50:50) for 10 minutes. Permeabilization was done by 0.1% Triton X 100 in PBS for 10 minutes. Permeabilized cells were blocked with 2% bovine serum albumin in PBS Tween 20 (0.05% Tween 20 in PBS). For labeling, cells were incubated with primary antibody inside a humidified chamber for 1 hour. After three successive washings, appropriate secondary antibodies were added and incubated for 45 minutes. After five washings, cells were stained with DAPI whenever required and mounted in aqueous mounting media. Cells were imaged by Olympus Inverted Confocal microscope at UNL microscopy core facility.
2.23 Antibodies

Anti-VSV antibody has been reported previously (Li & Pattnaik, 1999). Anti-N monoclonal antibody and anti-M monoclonal antibody producing hybridoma cells 23H12 were obtained from Douglas Lyles (Lefrancois & Lyles, 1982; Lyles et al., 1992). Anti-L rabbit polyclonal antibody was a gift from Manfred Schubert (Schubert et al., 1985). Anti- human ISG56 antibody was obtained from S. Sarkar (Guo et al., 2000). Commercial antibodies such as anti-actin (Sc-47778, Santa cruz biotechnology), anti-eGFP (Sc-9996, Santacruz), anti-*Renilla reniformis* GFP (K 17, Santacruz), anti-beta COP (maD, Sigma), zeta COP (31-45, Sigma) and anti-ARF1 (Sigma) and anti-GM130 monoclonal (BD science) and anti-GBF1 monoclonal (BD science) were purchased.

2.24 Chemicals

Bafilomycin A1, Brefeldin A, Golgicide A, Tyrphostin AG1478 and CBM were purchased from Sigma.

2.25 Treatment of cells with chemicals

Stock concentration of Brefeldin A was prepared in ethanol where as DMSO was used to prepare the stock solutions of Bafilomycin A1, Golgicide and AG. Working stock solution of CBM was prepared in water. HeLa cells were treated with drugs before or after VSV infection.
CHAPTER III

GENERATION AND CHARACTERIZATION OF CELL LINES EXPRESSING VSV REPLICATION PROTEINS AND ATTEMPTS TO ESTABLISH A CELL LINE CARRYING SELECTABLE SUB-GENOMIC REPLICONS OF VSV

Part of the work described in this chapter was published in Journal of Virology, 2010.

3.1 Construction of a cell line constitutively expressing VSV replicon proteins

N, P and L are the replication proteins for VSV. Presence of N, P and L is sufficient to replicate the sub-genomic particles for VSV (Pattnaik & Wertz, 1990). Plasmids where the expressions of N, P and L were driven by cellular pol II (CMV) promoter were constructed. Protein expression from these plasmids was verified by labeling with $^{35}$S and immunoprecipitation with anti-VSV antibody (Fig 3.1). To determine if the proteins are functional in supporting replication of VSV DI particle genomes, the 293 cells that had been transiently transfected with the three plasmids were infected with DI particles and replication of DI particle genome was examined. Our results show that the expressed proteins are functional in replicating the DI particle RNA genome (data not shown). The level of DI RNA replication supported by the proteins expressed by cellular Pol II was reproducibly less than that obtained in the vaccinia virus-T7 polymerase expression system. We reasoned that it might be possible to obtain higher levels of replication, if all the VSV proteins are expressed from a single vector resulting in the expression of all three proteins in the same cell. Therefore, we constructed a plasmid in which the coding sequences for N, PeGFP, and L are under the control of separate CMV promoter and poly (A) sequences. This plasmid (pc-NPeGFPL, Fig. 3.2) expressed N, PeGFP, and L proteins in readily detectable amounts (Fig. 3.1, lane 5) and supported replication of DI particles (data not shown) at levels higher than that of three plasmid system but at a level approximately 15-20 % of that seen in the vaccinia virus-T7 polymerase system.
**Figure 3.1: Transient expression of VSV proteins by CMV promoter.** Cells were transfected with plasmids pHE-N (lane 2), pc-PeGFP (lane 3), or pHE-L (lane 4) encoding the individual viral proteins or the plasmid pc-NPeGFPL (lane 5) encoding the three viral proteins or the empty vector (lane 1) as shown on top of each lane. After 48 hours of transfection, cells were radiolabeled with Expre$^{35}$S$^{35}$S label for 8 hours and the radiolabeled proteins were immunoprecipitated with anti-VSV antibody, analyzed by SDS-PAGE, and detected by fluorography. Total lysates from cells infected with VSV-PeGFP (lane 6) or wild-type VSV (lane 7) were labeled and analyzed similarly to identify the virus-specific proteins. Positions of VSV proteins are shown on the right.
In order to generate stable cell lines expressing the three replication proteins of VSV, 293 cells were transfected with pc-NPeGFPL plasmid. This plasmid also encoded neomycin phosphotransferase gene for selection of transfected cells in the presence of G418. At 48 hours post-transfection, the cells were treated with G418. The drug resistant culture was then enriched for cells expressing PeGFp fluorescence by FACS. Immunofluorescent staining of these cells for expression of the N and L proteins demonstrated that the majority of the green fluorescent cells also expressed the N protein but only a small percentage (less than 5%) of the cells expressed the L protein (data not shown). To isolate cell clones expressing all three proteins, single cell sorting of cells positive for PeGFp expression was performed by FACS into 96-well tissue culture plates, with each well having one cell. Each expanded clonal population of cells was then analyzed for expression of L protein as well as N protein by immunofluorescent staining. These cells were also examined for expression of the viral proteins by radiolabeling and SDS-PAGE analysis. Additionally, most cell clones that expressed all the three viral proteins were also screened for their ability to support replication of the DI particle genome (data not shown). Based on these initial screening results, three cell clones (293-pcNPegFpL cell clone # 204, 206, and 211) were selected for further studies. Immunofluorescent staining of the cells derived from the clone # 204 showed that all the three viral proteins were expressed in majority of the cells in the culture (Fig.3.3).

Although, the levels of PeGFp or the N proteins in individual cells in the culture appeared to be similar, the L protein expression was somewhat variable from cell to cell. Similar variability in the expression of L protein was also seen in cells derived from the two other cell clones (data not shown). Examination of the expression of N, PeGFp, and
Figure 3.2: Schematic representation of the plasmid expressing the VSV N, PeGFP, and L proteins. Coding sequences of N, PeGFP and L were cloned under the control of CMV immediate early promoter and poly(A) signal in pcDNA 3.1-neo vector. The CMV promoter and poly(A) signal sequences for N and PeGFP were derived from pHygEGFP vector. Coding sequences of N, PeGFP and L are shown in blue, green, and red, respectively. Ampicillin resistant gene (Amp<sup>R</sup>), neomycin resistant gene (Neo<sup>R</sup>) and SV40 promoter and origin of replication (O), CMV promoter (pCMV), and polyadenylation signal sequences (pA) are shown.
L proteins in these clones by radiolabeling, immunoprecipitation, and SDS-PAGE analysis revealed that each of these cell clones expressed readily detectable levels of the three viral proteins (Fig. 3.4).

3.2 Efficient replication of DI particle genomes in cells stably expressing the VSV replication proteins

To test the functionality of proteins expressed by these stable cell clones, the cells were infected with DI particles of VSV and checked for the replication of DI particle RNA. Results show that each of the cell clones supported efficient replication of the VSV DI particle genome (Fig. 3.5 A). The level of DI genome replication in the three cell clones (lanes 3-5) was comparable to that seen in the vaccinia virus-T7 system (lane 1). These results suggest that the replication proteins of VSV can be constitutively expressed in a cell and they can form functional polymerase complex to support the replication of DI particles of VSV.

To examine the stability of expression of all the 3 VSV proteins and long term effects of expression of the VSV proteins on cell growth, the clonally isolated cells were monitored for growth under culture conditions. After continuous passaging over six months (approximately 60 passages), examination of the stable cell clones # 204 and # 206 showed relatively similar level of protein expression in these cell clones as compared to early passage cells (data not shown). Importantly, even after 6 months of passages of the cells in culture conditions, these cell lines supported efficient replication of DI particle RNA (Fig.3.5 B). These results suggest that constitutive expression of N, PeGFP, and L proteins has no adverse effects on cell viability or growth and that the cell lines can be maintained for long time without any adverse effect on the levels of protein expression.
Figure 3.3: Detection of VSV proteins in stable cell clones by immunofluorescence.

Cells from stable cell clone # 204 were simultaneously immunostained with anti-N monoclonal antibody (10G4) and anti-L antibody and subsequently with donkey anti-mouse Alexa 647 and goat anti-rabbit Alexa 594 secondary antibodies. Cells were mounted and observed under inverted confocal microscope. Epifluorescence microscopy of cells expressing PeGFP (panel a); immunofluorescent images of the same cells expressing the N (panel b, pseudo color blue) and the L protein (panel c, red). Panel d shows the merged images of panels a, b, and c. DIC image of the cells is shown in panel e.
Figure 3.4: Expression of VSV proteins in different cell clones. G418 resistant stable cell clones # 204, 206, and 211 (lanes 2-4) as well as pcDNA3.1 vector transfected stable 293-pcDNA cells (lane 1) were radiolabeled with Expre\textsuperscript{35}S\textsuperscript{35}S protein labeling mix. The radiolabeled proteins were immunoprecipitated with anti-VSV antibody, analyzed by SDS-PAGE, and detected by fluorography. Radiolabeled proteins from HEK293 cells infected with VSV-PeGFP (lane 5) or wild type VSV (lane 6) cells are also shown. Viral proteins are identified on the right.
Figure 3.5: Replication of DI particle RNA genome in 293-NPeGFPL cell clones.

(A) Cell clones # 204, 206, or 211 (lanes 3-5, respectively) or 293-pcDNA control cells (lane 2) grown in 60 mm dishes were infected with DI particles and radiolabeled with $^3$H-uridine. Radiolabeled DI RNA was recovered from cell extracts by immunoprecipitation with anti-N antibody and analyzed by agarose-urea gel electrophoresis and detected by fluorography as described previously. Lane 1 shows DI RNA replication products obtained in BHK-21 cells using the vv-T7 system. DI RNAs (− and + sense) are identified on the right. (B) DI RNA replication in stable cell clones after 60 passages (approximately six months) in culture. Experimental protocol was the same as in panel A.
or protein function.

3.3 Failure to demonstrate persistent DI particle replication using the NPeGFPL stable cell line

After generating a cell line which supports DI particle replication efficiently, next objective was to establish a cell line where DI particles replicate on a long-term basis. Previously, for Sendai virus it has been shown that DI genomes can be maintained in helper cell lines expressing N, P and L protein of sendai virus (Willenbrink & Neubert, 1994). To establish similar type of cell line for VSV, the 293-NPeGFPL cells were infected with DI particles. Cells were further passaged every 2 days up to 12 passages. At each passage, samples were collected to examine DI genome product by reverse transcriptase PCR. DI RNA could be detected at 4th passage by RT PCR whereas no DI genome could be detected by 6th passage (Fig. 3.6). Repeated attempts to maintain DI genome replication persistently in such helper cell line failed. DI genome replication occurs in the cytoplasm. So in each cell division DI genomes also should be partitioned into the daughter cells. There was also no overt cytopathic effect exhibited on the cells due to replication of DI RNA. Failure to establish such a system indicates that DI genome can be maintained only for short term duration.

3.4 Failure to demonstrate replication of a minimal replicon of VSV

Another possibility of failure to establish a persistent DI RNA replication in helper cell line is lack of positive-selection. Defective interfering particle genome did not encode any drug resistance reporter gene, so cannot be used for selection. To circumvent
Figure 3.6 Long term replication of DI genome in NPeGFPL cells

NPeGFPL cells (# 204) were infected with DI particles. After 24 hours of infection, cells were trypsinized and passaged further. Subsequently, cells were passaged after 48 hours. Cell lysate was collected in Trizol at each passage. RNA was extracted and presence of DI genome was examined by reverse transcriptase PCR.
this issue, a minimal replicon of VSV with two reporter genes as two separate transcription units was generated (Fig. 3.7 A). This plasmid contained mRFP gene as visual marker and hygromycin gene for drug selection. Figure 3.7 B depicts the rational for generating and using the minimal replicon. The replicon was under the control of CMV promoter. 293-NPeGFPL cell clone 204 was transfected with the mRFP-Hyg replicon. At 2 days post transfection, when the cells were observed under microscope, no red fluorescence could be observed suggesting low level or no replication of mini replicon. However after two days, Hygromycin B was added at 500μg/ml along with G418 500μg/ml. Cells were further passaged until drug-resistant colonies appear. Recovery of cells resistant to both the drugs would be an indicator of possible replication and transcription of mini replicon. The cells were found dead by 6-8 days of drug selection. No drug-resistant colonies were observed even after 18 days of drug selection.

3.5 Attempt to develop a cell line supporting replication of an autonomous replicon

Parallel to the above approach, an autonomous replicon of VSV was also constructed and attempts were made to launch the replicon in 293 cells. The rational for generating the autonomous replicon was that the replicon encoding the viral replication proteins once assembled by the proteins provided by the helper plasmids, will subsequently be replicated by the proteins encoded in the replicon. Thus, the replicon will become independent of the helper plasmids and persistent replication will be established. The replicon should encode a drug selection marker; therefore positive selection can be applied to isolate the cells harboring the replicon. To generate such a cell line, a sub-genomic replicon without M and G protein coding regions and with a Neo R gene was
Figure 3.7: Schematic of bicistronic mini replicon. A. Relevant components of the minimal replicon are shown. CMV, Cytomegalo virus immediate early promoter; pA SV40 polyadenylation signal; HH, hammerhead ribozyme; hepatitis delta virus ribozyme; Tr, VSV trailer; Le, VSV leader; mRFP, monomeric red fluorescent protein coding seq; Hyg, hygromycin resistant gene. B. Replication of bicistronic mini replicon. Resultant mini replicon transcript by CMV promoter is cleaved by HH and HDR. Subsequently this is encapsidated by VSV N protein and assembled into functional RNP. Replication by VSV polymerase will generate complementary sequence and transcription by VSV polymerase will express mRFP and confer hygromycin resistance.
generated (Fig 3.8). The principle of replication of sub-genomic replicon was similar to that of minimal replicon. The sub genomic replicon was transfected to HEK-293 cells along with N, P, and L expressing plasmids. The expression of N, P and L from the support plasmids are driven by CMV promoter. After 48 hrs of transfection, transfected cells were treated with G418 at a concentration 1mg/ml. Cells were sub cultured at 1:4 ratio and maintained in the presence of G418 at 1mg/ml. After 10 days of drug treatment few GFP expressing cells were observed. GFP-expressing cells were isolated from the plate by local trypsinization and grown further. Furthermore, GFP-expressing population was enriched by FACS. After 3 rounds of FACS sorting, a cell population where more than 50% of the cells were expressing GFP was achieved (Fig. 3.9). To examine possible replication of genome, cells were radiolabeled with $^3$H after actinomycin D treatment. Actinomycin D treatment blocks cellular mRNA transcription but has no effect on viral genome transcription. Repeated attempts to show VSV mRNA products were unsuccessful. However, PeGFP protein and N protein was readily detected in these cells by western blot analysis (data not shown). The reason for the failure of VSV specific replication in presence of Actinomycin-D is unclear. It can be speculated that the replicon plasmid was integrated into the host chromosome and there was expression of PeGFP through a cryptic promoter despite the fact that PeGFP was in negative sense orientation to the CMV promoter. To alleviate this issue, the sub-genomic replicon was transferred to an episomally replicating vector (pCEP4, Invitrogen). However, attempts to establish a replicon cell line using pCEP4 based replicon was also unsuccessful.
Figure 3.8: Schematic of autonomous replicon. A. Relevant components of the autonomous replicon are shown. CMV, Cytomegalo virus immediate early promoter; HH, hammerhead ribozyme; hepatitis delta virus ribozyme; Tr, VSV trailer; Le, VSV leader; N, nucleocapsid protein; PeGFP, phosphoprotein fused in frame with eGFP; L, large polymerase protein. Neo, neomycin phosphotransferase.
Figure 3.9: GFP expression in 293 cell line transfected with autonomous replicon

HEK-293 cells were transfected with replicon plasmid along with N, P and L expressing plasmid. Drug selection started after 48 hours of transfection. GFP-expressing cells were enriched by FACS analysis. A cell population expressing PeGFP protein is shown in the right. Left panel is phase contrast image of the same cells.
3.5 DI particle RNA replication activates IFNβ promoter

Activation of dsRNA signaling and subsequent IFN production by both VSV and VSV DI particle has been well documented (Marcus & Gaccione, 1989). Matrix (M) protein is the only protein encoded by VSV known to shut off host cellular mRNA synthesis and nucleo-cytoplasmic transport of RNAs and thereby inhibit IFN production (Ahmed & Lyles, 1998; Faria et al., 2005). The availability of vaccinia virus-free system that supports DI RNA replication provided us a unique opportunity to examine the effects of DI RNA on host cell functions, particularly, the activation of innate immune signaling pathways in the absence of the viral M protein. It was hypothesized that replication of DI particles of VSV might be activating the interferon response and thus after few passages, DI genomes are lost from the cell. To test this hypothesis, luciferase reporter system was used. DI particle infection of the control cell line not expressing the viral proteins (293-pcDNA) exhibited only basal level (1.3-fold) of IFNβ promoter activity over uninfected cells (Fig. 3.10). However, in 293-pcNPeGFPL cells expressing the replication proteins, DI particle infection and subsequent DI RNA replication resulted in activation of IFN β promoter by as much as 67-fold over uninfected cells and 50-fold over DI particle infected control 293-pcDNA cells (Fig. 3.10). Additionally, activation of IFNβ promoter was not detected in 293-NP cells (expressing only the N and P proteins) infected with DI particles (Fig. 3.11). Replication of DI RNA in 293-NPeGFPL cells, but not in 293-pcDNA control cells or in 293-NP cells could be demonstrated (Fig. 3.11). These results suggest that replication of DI RNA, not entry and uncoating of DI RNA or mere expression of the viral replication proteins per se potently activates IFNβ promoter.
A kinetic analysis of IFN β promoter activation further revealed that indeed replication of DI RNA induced activation of IFN β promoter. In 293-pcDNA cells, DI particle infection resulted only basal level of IFNβ promoter activity while in 293-pcNPeGFPL cells infected with DI particles, IFNβ promoter activity increased with time post-infection (Fig. 3.12). The level of DI RNA at 2 hpi (as determined by RT-PCR in our hand) was below detection (Fig. lower panel), but a significant increase in IFN β promoter activity (7-fold) over similarly infected control cells could be readily seen. As time post-DI particle infection increased, IFN β promoter activity also increased, concomitant with increased DI RNA replication (Fig. 3.12 lower panel). Since, in cells expressing only the viral N and P proteins and infected with DI particles (a condition in which DI RNA replication does not occur), activation of IFNβ promoter was not observed, these studies demonstrate that IFNβ promoter activation requires DI RNA replication. As mere expression of the three replication proteins did not activate IFN β promoter and the fact that we could not detect RNP formation with cellular RNAs in the presence of N, PeGFP, and L protein (data not shown), these data strengthens our conclusion that only replication of DI RNA induces IFN β.

3.6 DI RNA replication activates ISRE promoter and NFκB promoter

In contrast to IFNβ promoter, ISRE promoter contains two ISRE elements which can be activated by IFN regulatory factor 3 (IRF3), IRF7, or both (Sarkar et al., 2004). To elucidate the effect of DI RNA replication upon IRF3-mediated signaling, an ISRE-promoter driven luciferase reporter plasmid was used. DI particle infection of 293-pcDNA control cells led to a meager 1.7-fold increase in luciferase activity over the corresponding uninfected cells (Fig. 3.13 panel A).
**Figure 3.10: Induction of interferon by replication of DI RNA.** Stable cell 293-NPeGFPL clone # 206 expressing N, PeGFP, and L proteins or 293-pcDNA control cells not expressing the viral proteins were co-transfected with 0.4 μg of IFNβ-Luc along with 10 ng of pRL-TK (renilla luciferase plasmid from Promega, which served as an internal control for transfection efficiency). At 24 hours post-transfection, the cells were either mock-infected or infected with DI particles. At 16 hours post-infection, the cells were lysed in buffer (Promega) and luciferase assay was performed using Dual Luciferase assay kit (Promega). Luciferase activities were expressed as relative fold induction over uninfected 293-pcDNA control cells after normalizing with renilla luciferase activity. Data from three independent experiments is presented as mean±SD. Samples from a similar experiment radiolabeled with ³H-uridine were also analyzed for replication of DI RNA genome as described in legend to Fig. 2A and shown in lower panel of Fig 3A. DI RNAs (– and + sense) are identified on the right.
Figure 3.11: Replication of DI genome in cells expressing N-PeGFP and L proteins.

HEK-293 cells stably expressing N and P proteins or N, PeGFP and L proteins were transfected with luciferase reporter plasmids as described in panel A and then infected with DI particles. Luciferase activity in the samples were analysed at 12 hours post DI particle infection as described in panel A and expressed as fold change over 293-pcDNA control cells. Mean±SD from three independent experiments is presented. Parallel set of samples were analyzed for DI RNA replication product by RT-PCR (lower panel). As control for similar amounts of total RNA being used in the analysis, RT-PCR products for ribosomal protein L32 are also shown. Representative data for DI RNA replication and ribosomal protein L32 mRNA quantitation is shown.
Figure 3.12: Kinetics of IFNβ promoter activation.

Experimental protocol was the same as described in panel B, except that the samples were collected at various times (0, 2, and 12 hours for 293-pcDNA cells and 0, 2, 4, 8, and 12 hours for 293-NPeGFPL cells) post DI particle infection. Luciferase activity in the samples was measured and expressed as fold change over 293-pcDNA control cells.
However, DI particle infection of 293-NPeGFPL cells resulted in greater than 200-fold induction of luciferase activity over the uninfected 293-NPeGFP cells (Fig. 3.13, panel A). Activation of IFN β gene transcription requires co-ordinate action of IRF3, NFκB and ATF2/c-jun transcription factors (Wathelet *et al.*, 1998). Nearly 40-fold induction of NFκB promoter activity was observed in 293-NPeGFPL cells infected with DI particles as compared to similarly infected 293-pcDNA cells (Fig. 3.13 panel B). For a direct verification of IFN-stimulated gene expression, we examined expression of interferon stimulated gene-56 (ISG56), one of the viral stress-inducible genes that are induced by IFNs, dsRNA, and virus infections (Terenzi *et al.*, 2006). Results (Fig. 3.14) showed strong induction of ISG56 protein expression in 293-NPeGFPL cells infected with DI particles (lane 4), whereas in uninfected 293-NPeGFPL cells and in 293-pcDNA cells with or without DI particle infection, ISG56 protein was undetectable. In each of the above studies, DI RNA replication products were readily detected (data not shown). Taken together, the results suggest that replication of DI RNA in the 293-NPeGFPL cells potently activates IFN and IFN-signaling.

3.7 Discussion

Current experimental systems to examine negative-strand RNA virus genome replication and transcription largely depend on transient expression of the viral proteins, which mostly use the recombinant vaccinia virus expressing the T7 RNA polymerase. Long-term effects of replication of viral genomes on host cell functions and the effects of host cell functions on viral genome replication is difficult to address using these systems.
Figure 3.13: Activation of ISRE and NFκB promoter by DI particle replication. The experiment was conducted as described in Fig. 1.10 but (A) using the ISRE promoter driven luciferase gene. (B) using the NFκB promoter driven luciferase gene.
Figure 3.14: Activation of ISG-56 protein by DI RNA replication. For detection of endogenous interferon stimulated gene 56 (ISG56), cell lysate was prepared in lysis buffer. Equal amount of total protein was separated by SDS-10% PAGE, transferred to PVDF membrane and probed with polyclonal rabbit anti-ISG 56 antibody. A non-specific band detected with anti-ISG 56 antibody is identified with an asterisk. The level of actin in each sample to serve as loading control was analysed after stripping and re-probing the same membrane with anti-actin antibody.
because vaccinia virus causes significant cytopathic effects. Additionally, since infection of cells with VSV and many other negative-strand RNA viruses result in cell death, it is not possible to study the effects of virus replication on host cell functions in the context of virus infection. This study aimed at establishing a replicon system for VSV to further our understanding of virus-host interaction. Though the original objective was not achieved, the results uncovered several interesting aspects of VSV biology. Successful generation of stable cells expressing N, PeGFP and L proteins together suggests that these proteins may not play any observable role in VSV-induced cytopathogenesis. The cells expressing the VSV replication proteins appear morphologically normal, and grow as do the un-transfected 293 cells or the empty vector-transfected control cells. These cells had been maintained in culture for at least 60 passages over a 6 months period and no loss of viral protein expression or loss of polymerase functions had been detected, indicating that these stable cells can maintain functionality of the replication proteins for a long time.

In a previous study using separate plasmids encoding Sendai virus NP, P, and L proteins, it was possible to establish stable cells expressing the viral proteins (Willenbrink & Neubert, 1994). These stable helper cell clones supported and maintained Sendai virus DI RNA replication for at least 90 days. In contrast to the above study, this study failed to establish a persistent DI genome replication in NPeGFPL cells. Also potent activation of IFN signaling was observed in cells replicating the DI genome. VSV M protein is the only protein known to provide counter measure against cellular defense system (Ahmed & Lyles, 1998; Her et al., 1997). Failure to counter measure the cellular defense mechanisms activated by the replicating DI RNA genome could be the major
contributor for loss of DI genome from the cells. Consistent with the notion, induction of ISG56 protein in cells replicating the DI genome at 4th passage was observed and in such condition, DI replication product was also observed (data not shown). However, ISG56 induction was not detected in the cells which are devoid of DI genome during the later passages (data not shown). In case of Sendai virus, N, P or L protein might contribute towards counteracting IFN system thus persistent DI genome replication could be established. For another member of paramyxoviridae, respiratory syncytial virus (RSV), a cell line supporting autonomous replication of RSV replicon was recently established (Malykhina et al.). However, these authors deleted only three glycoprotein genes to generate the replicon and coding sequences for other viral genes were intact (Malykhina et al., 2011). VSV M protein is also involved in significant cytopathic effects (Black & Lyles, 1992; Blondel et al., 1990). In our lab, we had generated a VSV lacking the G protein (described in detail in the next chapter). This virus undergoes single round of infection in infected cells and cannot undergo budding yet show cytopathic effects similar to wild type VSV, thus strengthening the conclusion that M protein greatly contribute for the cytopathic effect. In association with N, P and L proteins, VSV M protein also plays a critical role in regulating a balance between virus transcription and replication by inhibiting viral transcription (Pal et al., 1985; Ye et al., 1985). Presence of M protein might be absolutely necessary to regulate the transcription and replication of VSV genome and absence of M protein could be detrimental for virus life cycle. So the failure to establish a replicon system is confounded by limitation to remove the matrix protein from the genome. A rabies virus without M and G protein has been generated and this virus has been used to study the regulation of transcription and replication by M protein
(Finke & Conzelmann, 2003). However, long-term replication of that mutant rabies virus has not been shown. So far, a generation VSV without M protein has not been reported. It is tempting to speculate that such a virus may not be viable.

Nevertheless, generation of stable helper cell lines gave us a unique opportunity to investigate several aspect of virus life cycle. Using this cell line, it was shown that replication of DI RNA activates IFN and IFN- signaling pathways. It has been reported previously that snap-back DI (±) particles of VSV activate interferon signaling. Furthermore, a preexisting molecule in the snap-back DI genome (presumably double-stranded RNA structure) has been proposed to be responsible for induction of IFN, as heat-inactivation or UV treatment of DI particles did not inhibit interferon induction in chick embryo fibroblast and mouse L cells (Marcus & Gaccione, 1989; Sekellick & Marcus, 1982). Contrary to these findings, our results revealed that replication of DI-T RNA with a panhandle-type DI genome (Meier et al., 1984) is required for IFNβ activation as well as IFN signaling. Mere entry and uncoating of DI particles or expression of only the viral replication proteins was not sufficient to induce the IFN β or IFN signaling. VSV DI genomes are synthesized in the form of nucleocapsid and possible formation of dsRNA structures in infected cells has not been reported. Viral replication intermediates such as dsRNA or ssRNAs with triphosphorylated 5’-ends are sensed by both cytoplasmic sensors, e.g. RIG-I, MDA-5, and Nod2 (Barral et al., 2009; Sabbah et al., 2009) and endosomal receptors (Toll-like receptors, TLR3, TLR7/8, TLR9) (Thompson & Locarnini, 2007). However, the HEK293 cells are deficient in TLRs and Nod2. Therefore, in these cells, IFN induction and signaling might be mediated by RIG-I and/or MDA-5 through recognition of DI RNA. RIG-I was shown to be involved in
detection of rhabdoviruses and paramyxoviruses and subsequent induction of IFN response (Hornung et al., 2006; Kato et al., 2006). However, it has been demonstrated recently that MDA-5 is also involved in recognition of measles virus (Ikegame et al.) and sendai virus DI particles (Yount et al., 2008). Whether the activation of IFNβ and NFκB as a result of VSV DI RNA replication involves one or more of these cytoplasmic sensors will require further investigation.

Since infection of cells with VSV and many other negative-strand RNA viruses result in cytopathogenesis and cell death, it is not possible to study the long term effects of virus replication on host cell functions in the context of virus infection. VSV-M protein mutant viruses have been in use to study IFN activation following VSV infection (Ahmed et al., 2008; Ahmed et al., 2009; Kato et al., 2006). However, this virus is also cytopathic and long term effect of VSV RNA replication cannot be studied using this virus. The system established in this study will provide the opportunity to study global effects of viral genome replication on host cell functions in the absence of the cytopathogenic effects of VSV M and G proteins.
CHAPTER IV

GENOME WIDE RNAi SCREEN IDENTIFIES HOST PROTEINS REQUIRED FOR VSV INFECTION
4.1 Genome wide siRNA screen

To identify host proteins modulating VSV infection, a genome wide siRNA screen was conducted. Qiagen genome wide siRNA library was used in 384 well microplate format. In this library, four independent siRNAs targeting each gene were divided into two pools each containing two siRNAs per well (two siRNAs per well and two wells per gene). Thus, this design created a total of 148 plates (Fig. 4.1 A). HeLa cells were used for the screen. Moreover, most of the siRNAs are validated in HeLa cells. These cells can also be readily infected with VSV. VSV-eGFP which expresses enhanced green fluorescent protein as an extra cistron (Das et al., 2006) was used in a cell based assay. The level of eGFP expression can be correlated with VSV gene expression. AllStars non targeting control siRNA (Qiagen) and siRNAs targeting to VSV N and L mRNAs (Dharmacon) were used as control siRNAs to verify transfection efficiencies. Screen was optimized using the control siRNAs. Optimized conditions for the screen were determined empirically. Briefly, 3000 HeLa cells were reverse transfected using Lipofectamine RNAiMAX. At 52 hrs post transfection, cells were infected with VSV-eGFP and at 18 hrs post infection fixed and processed for automated image analysis using Cellomics ArrayScan VTI. Cell number (Valid Object Count) and a percent infection was obtained for each well. Under these conditions, siRNAs targeting to N and L genes reproducibly showed less than 2 % infection.
Figure 4.1: Genome wide siRNA screen design.

(A) Genomic library design. In the Qiagen genomic library, each gene is targeted by 4 independent siRNA duplexes. Two independent siRNAs targeting the same gene were plated in each well of a 384 well plate. This design created a paired plate set (set AB and set CD) and each set has 74 assay plates. Set AB and set CD are identical in terms of the position of the target genes. (B) Schematic representation of the primary screen. Arrayed siRNAs were reverse transfected to HeLa cells in 384 well plate format. After 52 hours post transfection, cells were infected with VSV-eGFP. Eighteen hours post infection, cells were fixed and processed for automated image analysis. Percent infection and total cell number for each well were measured.
4.2 Statistical analysis to obtain primary hit list candidates

Some of the wells showed significantly reduced cell number after transfection and infection. Reduction in cell number could be due to knock-down of critical cellular genes required for cell growth and/or could be due to VSV infection. Samples showing cell numbers 800 or less were omitted from further analysis. Thus, a total of 5536 genes were discarded (24% of total genes). Sum rank analysis was used to obtain the primary hit list candidates. Many ribosomal proteins were identified in this screen. We decided to omit the ribosomal genes from the list because of two reasons: (i) ribosome function is required for VSV protein synthesis and these are expected to show in the screen (further strengthening the robustness of the screen) and (ii) we were interested to identify some other unknown genes. So we ran the sum rank analysis after deleting the ribosomal genes from the database. A ‘p value’ was given to each gene and a hit list was obtained (p<0.01). This list contained 174 genes. Examination of percent infection for these genes showed that all these genes had percent infection below 12% whereas the percent infection of non-targeting control siRNA was approximately 60% and thus genes in this hit list showed approximately five fold inhibitory effects on VSV infection when knocked down.

4.3 Secondary validation screen

Primary candidates obtained from the genome scale RNAi screens are often populated with off target effects. To reduce the false discoveries and off target effects, the primary screen was conducted in 2 X 2 formats. This format offered significant advantages compared to the pool of four format and a recent report successfully applied
Figure 4.2: Schematic of secondary validation screen. Arrayed siRNAs were reverse transfected into HeLa cells in 96 well plate format. After 51 hours post transfection, cells were infected with VSV-eGFP. Sixteen hours post infection, cells were fixed and processed for automated image analysis. Percent infection and total cell number for each well were measured.
this format to identify genes involved in circadian clock and genes required for Yellow fever virus infection (Barrows et al., 2010; Zhang et al., 2009a). This 2 X 2 format resulted in high percentage of validation in secondary deconvolution assays (Zhang et al., 2009a). In our primary screen, the genes identified had at least two independent siRNAs showing the same phenotype on VSV infection. In siRNA screens using a pool of four siRNAs, if two siRNAs in the subsequent validation screen show the same phenotype, it is generally considered to be the true effect of the gene knockdown and not to be due to off-target effects (Karlas et al., 2010; Konig et al., 2010; Sharma & Rao, 2009).

Nevertheless, we took a more conservative approach and decided to use siRNA from a different manufacturer for the secondary validation screen rather than deconvoluting the siRNA from the primary screen. This approach ameliorated any bias in siRNA target prediction and synthesis of siRNA by a particular manufacturer. So for the secondary validation screen, Dharmacon On Target plus pool of four siRNA was used. We had total 174 genes for the secondary screen. The assay plate format was also changed to 96-well format to increase the number of cells transfected and infected and thereby increasing the statistical power. A schematic of secondary validation screen is shown in figure 4.2.

Similar to primary screen, a cell based screen was used for the secondary assay. Percent infection and valid object counts were obtained for each well. The secondary validation screen was repeated four times. For statistical analysis, data was normalized to plate wise median, and Wilcoxon rank sum test was applied to identify genes inhibiting VSV infection after knockdown. Using this method, we were able to identify at least 72 genes as required factors for VSV infection, thus yielding a validation rate of 40% from the primary list. These genes showed statistically significant difference ($p < 0.01$) in percent
infection compared to NT control wells. Although some of the genes (vATPases) were previously known to be involved in VSV infection, majority of genes have not been implicated in VSV infection. Table 4.1 represents the list of genes, their known and putative functions and their sub-cellular localizations.

### 4.4 Identification of host genes required for LCMV and HPIV3

In recent years, several groups have completed genome wide siRNA screens for HIV, HCV and other positive strand RNA viruses. However, there are no such reports for negative strand RNA viruses except for Influenza. The negative strand RNA viruses also comprises of several medically important pathogens. In order to gain a better understanding of virus-host interaction and identify genes and pathways utilized by a diverse family of negative strand RNA viruses, the involvement of 174 genes (identified in the screen using VSV) was investigated in infection of cells with Human Parainfluenza virus 3 (HPIV3) a member of Paramyxoviridae family and Lymphocytic choriomeningitis virus (LCMV) a member of Arenaviridae family. These viruses are cytoplasmically replicating RNA viruses but belong to different families and have different genome organizations. These viruses show common as well as unique strategies for entry, uncoating, genome replication, and virus assembly processes in infected cells. Therefore, it was with great interest that we examined the involvement of host factors identified from the VSV screen in LCMV and HPIV3 infection. For the screen, GFP expressing LCMV and HPIV-3 were used (Emonet et al., 2009; Zhang et al., 2005). Similar to VSV, for these viruses the screen was conducted in 96-well plate format and percent infection and valid object count was obtained for each well. Hit list was identified using Wilcoxon rank sum test. Seventy one and thirty genes were identified as
candidates required for LCMV and HPIV respectively. Figure 4.3 A depicts the degree of overlap among the three viruses. Twenty five genes were found to be required by all three viruses. Among these required factors, VSV and LCMV shared maximum number of genes (54) where as VSV and HPIV shared only 27 genes. Genes such as FAAH, NUP160, SLC12A9 and SLC43A2 were identified consistently as required factors in LCMV infection only. Genes such as GLYATL1 and SLC39A7 were found to be required for HPIV-3 infection only. The normalized percent inhibition compared to NT siRNA treated wells for 72 genes (positive for VSV) were shown as a color coded map and compared to those in LCMV and HPIV infection (Fig. 4.3 B). Majority of the genes, when depleted, reduced VSV infection by two fold or more. This screen also identified several previously known candidates such as vacuolar ATPases involved in VSV and LCMV infection whereas HPIV-3 was insensitive to vATPase depletion. VSV and LCMV enter the cells through endocytosis and require the vATPases whereas HPIV enters through fusion at the plasma membrane. These results lent support for the robustness of the assay.

4.5 Bioinformatics analysis

Over representation of biological functions was analyzed by Panther classification system (Mi et al., 2005). Gene ontology biological function enrichment analysis identified 8 statistically significant functional groups based on hyper geometric p values (Fig. 4.4). Biological functions such as vesicle mediated trafficking, intracellular protein transport, exocytosis were enriched in the dataset. We were interested in examining whether other viruses such as influenza, HCV, HIV shared these mammalian genes for
Table 4.1 List of genes identified as required factors for VSV

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene ID</th>
<th>Description</th>
<th>Uniprot function</th>
<th>Uniprot sub-cellular location</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAL</td>
<td>161823</td>
<td>Adenosine deaminase-like protein</td>
<td>Putative nucleoside deaminase. May catalyze the hydrolytic deamination of adenosine or some similar substrate</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>ANKRD53</td>
<td>79998</td>
<td>Ankyrin repeat domain-containing protein 53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARCN1</td>
<td>372</td>
<td>Coatomer subunit delta</td>
<td>Coatomer complex is required for budding from Golgi membranes, and is essential for the retrograde Golgi-to-ER transport of dilysine-tagged proteins</td>
<td>Cytoplasm and Golgi apparatus</td>
</tr>
<tr>
<td>ARF1</td>
<td>375</td>
<td>ADP-ribosylation factor 1</td>
<td>Involved in protein trafficking among different compartments. Modulates vesicle budding and uncoating within the Golgi complex. Deactivation induces the redistribution of the entire Golgi complex to the endoplasmic reticulum, suggesting a crucial role in protein trafficking</td>
<td>Golgi membrane, cytosol, perinuclear region of cytoplasm, plasma membrane</td>
</tr>
<tr>
<td>ARHGAP23</td>
<td>57636</td>
<td>Rho GTPase-activating protein 23</td>
<td>GTPase activator for the Rho-type GTPases by converting them to an inactive GDP-bound state</td>
<td></td>
</tr>
<tr>
<td>ATP6AP2</td>
<td>10159</td>
<td>ATPase H(+)-transporting lysosomal accessory protein 2</td>
<td>Encodes a component of vacuolar ATPase (V-ATPase). May mediate renin-dependent cellular responses by activating ERK1 and ERK2</td>
<td>Membrane; Single-pass type I membrane protein</td>
</tr>
<tr>
<td>Gene</td>
<td>RefSeq ID</td>
<td>Description</td>
<td>Function</td>
<td>Location</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>ATP6V0B</td>
<td>533</td>
<td>ATPase, H+ transporting, lysosomal 21kDa, V0 subunit b</td>
<td>This gene encodes a component of vacuolar ATPase (V-ATPase)</td>
<td>Membrane</td>
</tr>
<tr>
<td>ATP6V0D1</td>
<td>9114</td>
<td>V-type proton ATPase subunit d 1</td>
<td>Subunit of the integral membrane V0 complex of vacuolar ATPase. Vacuolar ATPase is responsible for acidifying a variety of intracellular compartments in eukaryotic cells</td>
<td>Endosome membrane</td>
</tr>
<tr>
<td>ATP1F1</td>
<td>93974</td>
<td>ATPase inhibitor, mitochondrial</td>
<td>Endogenous F_{1,F(o)}-ATPase inhibitor limiting ATP depletion when the mitochondrial membrane potential falls below a threshold</td>
<td>Mitochondrion</td>
</tr>
<tr>
<td>BBC3</td>
<td>27113</td>
<td>Bcl-2-binding component 3</td>
<td>Essential mediator of p53-dependent and p53-independent apoptosis</td>
<td>Mitochondrion</td>
</tr>
<tr>
<td>BCL2L2</td>
<td>79370</td>
<td>Bcl-2-like protein 2</td>
<td>Promotes cell survival</td>
<td>Mitochondrial membrane</td>
</tr>
<tr>
<td>C14ORF111</td>
<td>51077</td>
<td>rRNA-processing protein FCF1 homolog</td>
<td>Essential protein involved in pre-rRNA processing and 40S ribosomal subunit assembly</td>
<td>Nucleus, nucleolus</td>
</tr>
<tr>
<td>C2ORF28</td>
<td>51374</td>
<td>Apoptosis-related protein 3</td>
<td>May play a critical role in inducing the cell cycle arrest via inhibiting CCND1 expression in all-trans-retinoic acid (ATRA) signal pathway</td>
<td>Cell membrane; Single-pass membrane protein</td>
</tr>
<tr>
<td>C9ORF96</td>
<td>169436</td>
<td>Protein kinase-like protein SgK071</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Accession</td>
<td>Description</td>
<td>Function</td>
<td>Localization</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>-------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>CARD14</td>
<td>79092</td>
<td>Caspase recruitment domain-containing protein 14</td>
<td>Activates NF-kappa-B via BCL10 and IKK. Stimulates the phosphorylation of BCL10.</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>CDK15</td>
<td>65061</td>
<td>Cyclin-dependent kinase 15</td>
<td>Serine/threonine-protein kinase involved in the control of the eukaryotic cell cycle, whose activity is controlled by an associated cyclin</td>
<td></td>
</tr>
<tr>
<td>CEACAM7</td>
<td>1087</td>
<td>Carcinoembryonic antigen-related cell adhesion molecule 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COPA</td>
<td>1314</td>
<td>Coatomer subunit alpha</td>
<td>Coatomer complex is required for budding from Golgi membranes, and is essential for the retrograde Golgi-to-ER transport of dilysine-tagged proteins</td>
<td>Cytoplasm and Golgi apparatus membrane</td>
</tr>
<tr>
<td>COPB</td>
<td>1315</td>
<td>Coatomer subunit beta</td>
<td>Coatomer complex is required for budding from Golgi membranes, and is essential for the retrograde Golgi-to-ER transport of dilysine-tagged proteins</td>
<td>Cytoplasm and Golgi apparatus membrane</td>
</tr>
<tr>
<td>COPB2</td>
<td>9276</td>
<td>Coatomer subunit beta'</td>
<td>Coatomer complex is required for budding from Golgi membranes, and is essential for the retrograde Golgi-to-ER transport of dilysine-tagged proteins</td>
<td>Cytoplasm and Golgi apparatus membrane</td>
</tr>
<tr>
<td>COPG</td>
<td>22820</td>
<td>Coatomer subunit gamma</td>
<td>Coatomer complex is required for budding from Golgi membranes, and is essential for the retrograde Golgi-to-ER transport of dilysine-tagged proteins</td>
<td>Cytoplasm and Golgi apparatus membrane</td>
</tr>
<tr>
<td>COPZ1</td>
<td>22818</td>
<td>Coatomer subunit zeta-1</td>
<td>Coatomer complex is required for budding from Golgi membranes, and is essential for the retrograde Golgi-to-ER transport of dilysine-tagged proteins</td>
<td>Cytoplasm and Golgi apparatus membrane</td>
</tr>
<tr>
<td><strong>CRMP1</strong></td>
<td>1400</td>
<td>Collapsin response mediator protein 1</td>
<td>Necessary for signaling by class 3 semaphorins and subsequent remodeling of the cytoskeleton. Plays a role in axon guidance, invasive growth and cell migration</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>CUL5</strong></td>
<td>8065</td>
<td>Cullin-5</td>
<td>Core component of multiple SCF-like ECS (Elongin-Cullin 2/5-SOCS-box protein) E3 ubiquitin-protein ligase complexes, which mediate the ubiquitination and subsequent proteasomal degradation of target proteins.</td>
<td></td>
</tr>
<tr>
<td><strong>DCAF13</strong></td>
<td>25879</td>
<td>DDB1- and CUL4-associated factor 13</td>
<td>Possible role in ribosomal RNA processing May function as a substrate receptor for CUL4-DDB1 E3 ubiquitin-protein ligase complex</td>
<td>Nucleus, nucleolus</td>
</tr>
<tr>
<td><strong>DCD</strong></td>
<td>117159</td>
<td>Dermcidin</td>
<td>DCD-1 displays antimicrobial activity thereby limiting skin infection by potential pathogens in the first few hours after bacterial colonization</td>
<td>Secreted</td>
</tr>
<tr>
<td><strong>DDX60L</strong></td>
<td>91351</td>
<td>DEAD box protein 60-like</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EIF3A</strong></td>
<td>8661</td>
<td>Eukaryotic translation initiation factor 3 subunit A</td>
<td>Component of the eukaryotic translation initiation factor 3 (eIF-3) complex, which is required for several steps in the initiation of protein synthesis</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td><strong>EIF3G</strong></td>
<td>8666</td>
<td>Eukaryotic translation initiation factor 3 subunit G</td>
<td>Component of the eukaryotic translation initiation factor 3 (eIF-3) complex, which is required for several steps in the initiation of protein synthesis</td>
<td>Cytoplasm, nucleus</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Gene ID</td>
<td>Description</td>
<td>Function</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------</td>
<td>-------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>EIF3S2</td>
<td>8668</td>
<td>Eukaryotic translation initiation factor 3, subunit 2 beta, 36kDa</td>
<td>Endoplasmic reticulum membrane, single-pass type I membrane protein</td>
<td></td>
</tr>
<tr>
<td>ERN2</td>
<td>10595</td>
<td>Endoplasmic reticulum-to-nucleus signaling 2</td>
<td>Induces translational repression through 28S ribosomal RNA cleavage in response to ER stress. Pro-apoptotic.</td>
<td></td>
</tr>
<tr>
<td>FAM167B</td>
<td>84734</td>
<td>Protein FAM167B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAU</td>
<td>2197</td>
<td>40S ribosomal protein S30</td>
<td>This ribosomal protein is synthesized as a C-terminal extension protein (CEP) of a ubiquitin-like protein</td>
<td></td>
</tr>
<tr>
<td>FBL</td>
<td>2091</td>
<td>rRNA 2'-O-methyltransferase fibrillarin</td>
<td>Involved in pre-rRNA processing. Utilizes the methyl donor S-adenosyl-L-methionine to catalyze the site-specific 2'-hydroxyl methylation of ribose moieties in pre-ribosomal RNA. Site specificity is provided by a guide RNA that base pairs with the substrate</td>
<td></td>
</tr>
<tr>
<td>FOXF2</td>
<td>2295</td>
<td>Forkhead box protein F2</td>
<td>Probable transcription activator for a number of lung-specific genes.</td>
<td></td>
</tr>
<tr>
<td>FRS2</td>
<td>10818</td>
<td>Fibroblast growth factor receptor substrate 2</td>
<td>Adapter protein that links FGR and NGF receptors to downstream signaling pathways. Involved in the activation of MAP kinases</td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Gene ID</th>
<th>Description</th>
<th>Function</th>
<th>Subcellular Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>GABBR2</td>
<td>9568</td>
<td>Gamma-aminobutyric acid type B receptor subunit 2</td>
<td>Receptor for GABA. The activity of this receptor is mediated by G-proteins that inhibit adenylyl cyclase activity, stimulates phospholipase A2, activates potassium channels, inactivates voltage-dependent calcium-channels and modulates inositol phospholipids hydrolysis</td>
<td>Cell membrane</td>
</tr>
<tr>
<td>GBF1</td>
<td>8729</td>
<td>Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1</td>
<td>Promotes guanine-nucleotide exchange on ARF5. Promotes the activation of ARF5 through replacement of GDP with GTP</td>
<td>Golgi membrane</td>
</tr>
<tr>
<td>GRM2</td>
<td>2912</td>
<td>Metabotropic glutamate receptor 2</td>
<td>Receptor for glutamate. The activity of this receptor is mediated by a G-protein that inhibits adenylate cyclase activity. May mediate suppression of neurotransmission or may be involved in synaptogenesis or synaptic stabilization</td>
<td>Cell membrane</td>
</tr>
<tr>
<td>H1FOO</td>
<td>132243</td>
<td>Histone H1oo</td>
<td>May play a key role in the control of gene expression during oogenesis and early embryogenesis, presumably through the perturbation of chromatin structure. Essential for meiotic maturation of germinal vesicle-stage oocytes</td>
<td>Cytoplasm, Nucleus</td>
</tr>
<tr>
<td>HNRPK</td>
<td>3190</td>
<td>Heterogeneous nuclear ribonucleoprotein K</td>
<td>One of the major pre-mRNA-binding proteins. Binds tenaciously to poly(C) sequences. Likely to play a role in the nuclear metabolism of hnRNAs, particularly for pre-mRNAs that contain cytidine-rich sequences. Can also bind poly(C) single-stranded DNA.</td>
<td>Cytoplasm, nucleus</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>HSPA9B</strong></td>
<td>3313</td>
<td>Stress-70 protein, mitochondrial</td>
<td>Implicated in the control of cell proliferation and cellular aging. May also act as a chaperone.</td>
<td>Mitochondrion</td>
</tr>
<tr>
<td><strong>IFI35</strong></td>
<td>3430</td>
<td>Interferon-induced 35 kDa protein</td>
<td></td>
<td>Nucleus</td>
</tr>
<tr>
<td><strong>KCNH4</strong></td>
<td>23415</td>
<td>Potassium voltage-gated channel subfamily H member 4</td>
<td>Pore-forming (alpha) subunit of voltage-gated potassium channel. Elicits an outward current, but shows no inactivation. Channel properties may be modulated by cAMP and subunit assembly.</td>
<td>Membrane</td>
</tr>
<tr>
<td><strong>LAMC3</strong></td>
<td>10319</td>
<td>Laminin subunit gamma-3</td>
<td>Binding to cells via a high affinity receptor, laminin is thought to mediate the attachment, migration and organization of cells into tissues during embryonic development</td>
<td>Secreted</td>
</tr>
<tr>
<td><strong>LYSMD1</strong></td>
<td>388695</td>
<td>LysM and putative peptidoglycan-binding domain-containing protein 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MAT2A</strong></td>
<td>4144</td>
<td>S-adenosylmethionine synthase isoform type-2</td>
<td>Catalyzes the formation of S-adenosylmethionine from methionine and ATP</td>
<td></td>
</tr>
<tr>
<td><strong>NHP2L1</strong></td>
<td>4809</td>
<td>NHP2-like protein 1</td>
<td>Binds to the 5’-stem-loop of U4 snRNA and may play a role in the late stage of spliceosome assembly. The protein undergoes a conformational change upon RNA-binding</td>
<td>Nucleus</td>
</tr>
<tr>
<td><strong>NOL5A</strong></td>
<td>10528</td>
<td>Nucleolar protein 56</td>
<td>Required for 60S ribosomal subunit biogenesis.</td>
<td>Nucleus</td>
</tr>
<tr>
<td><strong>NRL</strong></td>
<td>4901</td>
<td>Neural retina-specific leucine zipper protein</td>
<td>Transcription factor which regulates the expression of several rod-specific genes, including RHO and PDE6B</td>
<td>Nucleus</td>
</tr>
<tr>
<td>Gene</td>
<td>RefSeq</td>
<td>Description</td>
<td>Function</td>
<td>Location</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>OR7A10</td>
<td>390892</td>
<td>Olfactory receptor 7A10</td>
<td>Odorant receptor</td>
<td>Cell membrane</td>
</tr>
<tr>
<td>P2RY10</td>
<td>27334</td>
<td>Putative P2Y purinoceptor 10</td>
<td>Putative receptor for purines coupled to G-proteins.</td>
<td></td>
</tr>
<tr>
<td>PIWIL3</td>
<td>440822</td>
<td>Piwi-like protein 3</td>
<td>Acts via the piRNA metabolic process, which mediates the repression of transposable elements during meiosis. Directly binds piRNAs, a class of 24 to 30 nucleotide RNAs that are generated by a Dicer-independent mechanism.</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>PKMYT1</td>
<td>9088</td>
<td>Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase</td>
<td>Acts as a negative regulator of entry into mitosis (G2 to M transition) by phosphorylation of the CDK1 kinase specifically when CDK1 is complexed to cyclins. Mediates phosphorylation of CDK1 predominantly on 'Thr-14'. Also involved in Golgi fragmentation</td>
<td>Endoplasmic reticulum membrane, Golgi membrane</td>
</tr>
<tr>
<td>PPP1R15B</td>
<td>84919</td>
<td>Protein phosphatase 1 regulatory subunit 15B</td>
<td>Maintains low levels of EIF2S1 phosphorylation in unstressed cells by promoting its dephosphorylation by PP1</td>
<td></td>
</tr>
<tr>
<td>PPP5C</td>
<td>5536</td>
<td>Serine/threonine-protein phosphatase 5</td>
<td>May play a role in the regulation of RNA biogenesis and/or mitosis. In vitro, dephosphorylates serine residues of skeletal muscle phosphorylase and histone H1.</td>
<td>Nucleus, cytoplasm</td>
</tr>
<tr>
<td>PWP2H</td>
<td>5822</td>
<td>Periodic tryptophan protein 2 homolog</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAB5C</td>
<td>5878</td>
<td>Ras-related protein Rab-5C</td>
<td>Protein transport. Probably involved in vesicular traffic</td>
<td>Cell membrane, lipid anchor,</td>
</tr>
<tr>
<td>RIT2</td>
<td>6014</td>
<td>GTP-binding protein Rit2</td>
<td>Binds and exchanges GTP and GDP</td>
<td>Cell membrane</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Accession</td>
<td>Description</td>
<td>Function</td>
<td>Cellular Location</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>-------------</td>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td>SDF4</td>
<td>51150</td>
<td>45 kDa calcium-binding protein</td>
<td>May regulate calcium-dependent activities in the endoplasmic reticulum lumen or post-ER compartment</td>
<td>Golgi apparatus lumen, cytoplasm</td>
</tr>
<tr>
<td>SLC46A1</td>
<td>113235</td>
<td>Proton-coupled folate transporter</td>
<td>Act both as an intestinal proton-coupled high-affinity folate transporter and as an intestinal heme transporter which mediates heme uptake from the gut lumen into duodenal epithelial cells</td>
<td></td>
</tr>
<tr>
<td>SLC9A3</td>
<td>6550</td>
<td>Sodium/hydrogen exchanger 3</td>
<td>Involved in pH regulation to eliminate acids generated by active metabolism or to counter adverse environmental conditions</td>
<td>Membrane</td>
</tr>
<tr>
<td>SMAD2</td>
<td>4087</td>
<td>Mothers against decapentaplegic homolog 2</td>
<td>Receptor-regulated SMAD (R-SMAD) that is an intracellular signal transducer and transcriptional modulator activated by TGF-beta (transforming growth factor) and activin type 1 receptor kinases.</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>SYVN1</td>
<td>84447</td>
<td>E3 ubiquitin-protein ligase synoviolin</td>
<td>Acts as an E3 ubiquitin-protein ligase which accepts ubiquitin specifically from endoplasmic reticulum-associated UBC7 E2 ligase and transfers it to substrates, promoting their degradation</td>
<td>Endoplasmic reticulum membrane</td>
</tr>
<tr>
<td>TAS2R48</td>
<td>259294</td>
<td>Taste receptor type 2 member 19</td>
<td>Receptor that may play a role in the perception of bitterness and is gustducin-linked</td>
<td>Membrane</td>
</tr>
<tr>
<td>TIPIN</td>
<td>54962</td>
<td>TIMELESS-interacting protein</td>
<td>Required for normal progression of S-phase. Important for cell survival after DNA damage or replication stress.</td>
<td>Cytoplasm</td>
</tr>
<tr>
<td>TOX</td>
<td>9760</td>
<td>Thymocyte selection-associated high mobility group box protein TOX</td>
<td>May play a role in regulating T-cell development</td>
<td>Nucleus</td>
</tr>
<tr>
<td>UQCRC1</td>
<td>7384</td>
<td>Cytochrome b-c1 complex subunit 1, mitochondrial</td>
<td>A component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is part of the mitochondrial respiratory chain</td>
<td>Mitochondrion inner membrane</td>
</tr>
<tr>
<td>Code</td>
<td>Accession</td>
<td>Description</td>
<td>Function</td>
<td>Location</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>--------------------------------------------------</td>
<td>---------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>UTP11L</td>
<td>51118</td>
<td>Probable U3 small nucleolar RNA-associated protein 11</td>
<td>Involved in nucleolar processing of pre-18S ribosomal RNA</td>
<td>Nucleus</td>
</tr>
<tr>
<td>UTP6</td>
<td>55813</td>
<td>U3 small nucleolar RNA-associated protein 6 homolog</td>
<td>Involved in nucleolar processing of pre-18S ribosomal RNA</td>
<td>Nucleus</td>
</tr>
<tr>
<td>WDR46</td>
<td>9277</td>
<td>WD repeat-containing protein 46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZNF446</td>
<td>55663</td>
<td>Zinc finger protein 446</td>
<td>May be involved in transcriptional regulation.</td>
<td>Nucleus</td>
</tr>
</tbody>
</table>
Figure 4.3: Effect of gene depletion on VSV, LCMV and HPIV3 infection

A. Ven diagram depicting host factors shared by VSV, LCMV and HPIV.

B. Relative fold inhibition of infection in siRNA treated wells compared to non-targeting siRNA well was calculated. Data represents average values from 4 independent experiments for VSV and 3 independent experiments for LCMV and HPIV3. Percent of non-targeting control calculated for each of 72 genes was depicted as color coded map. Blue represent strong inhibition whereas red represent weak inhibition.
Figure 4.4: Biological process classifications for the identified candidates

Gene ontology biological process terms significantly overrepresented by the genes in the screen were identified using panther classification system. The terms were ordered by the number of genes represented in a category. Values in the parenthesis denote number of genes and the p values.
their infection. We analyzed the candidate genes identified in other human genomewide siRNA screens by biological function using Panther database. Gene ontology terms were identified for each of these screens and statistical significance was obtained by hypergeometric $p$ values. Functions with $p$ values $< 0.05$ were considered significant. Figure 4.5 shows the biological function enriched in VSV dataset and also identifies involvement of these gene ontology terms in screens with other viruses. This analysis revealed that several biological functions such as exocytosis, vesicle mediated transport and intracellular transport are shared by other human pathogens such as influenza, HIV and HCV. Some functions such as segment specification and gut mesoderm development are uniquely required by VSV. The limited overlap of identified candidates between screens for the same pathogen is a growing concern in the field. The precise reasons for this disparity are not known. Till now there is no such screen for VSV or any other cytoplasmic replicating negative strand RNA viruses. Therefore, a direct comparison of data is not possible. The replication mechanism for Influenza, HIV and positive strand RNA viruses are different from VSV. However, identification of pathways shared by all these viruses implicates evolutionary conserved mechanisms of virus-host interactions.
Figure 4.5: Comparison of genome-wide screens between VSV and other viruses

Candidates genes identified in the VSV screens were compared for the biological function with human genome wide screens for HIV, HCV, Influenza and WNV.

Predicted gene ontology for biological functions was obtained and hypergeometric $p$ value was obtained. Results from human genome wide screen for HIV (Brass et al., 2008, Zhou et al. 2008), HCV (Tai et al., 2009, Li et al., 2009), Influenza (Konig et al., 2010, Karlas et al, 2010, Brass et al., 2009) and WNV (Krishnan et al., 2008) were used for classification of biological functions.
4.6 Discussion

In the previous years, considerable advancements have been made in our understanding of the roles of VSV proteins in viral genome replication and pathogenicity. However, the role of host proteins in VSV infection is poorly understood. During the last 2-3 years, RNAi technology has been successfully applied to identify factors involved in infection of several viruses including Influenza, HIV, HCV, WNV and Dengue virus. However, so far there has not been such investigation for any of the nonsegmented negative strand RNA viruses. This study describes a genome-wide siRNA screen to identify genes required for VSV infection. This approach allowed us to identify factors required for VSV infection in an unbiased manner. In the primary screen, we used Qiagen siRNA to knockdown genes in HeLa cells. HeLa cells were used because these cells can be easily transected with the siRNA and VSV readily infects these cells. Four independent siRNAs were used per gene in a pool of 2 format. The 2 X 2 pool genomic screen design required that at least two independent siRNA generate the same phenotype. Thus the 174 genes identified as required factors for VSV infection by sum rank analysis from the primary screen showed significant reduction of VSV gene expression when knocked down and atleast two independent siRNA show the same phenotype. The off-target effects are generally associated with hit identification in such genomic screens, thus, we used a more robust approach for the validation screen. Because in the primary screen, at least two independent siRNAs produced the same phenotype in VSV infection, the likelihood of obtaining the same phenotype is strong had we deconvoluted the original Qiagen siRNA pools. Instead, we changed the source of siRNA for the validation screen. This approach has two advantages; (i) factors associated with siRNA design and manufacturing can be avoided, (ii) change in the source allowed us to test more
independent siRNAs targeting one particular gene. Therefore, for the validation screen we used Dharmacon “On Target Plus” pool of 4 siRNAs. To provide more statistical power, we also changed the plate format to 96 well format and the validation screen was done in quadruplicate. In this manner, the 72 genes identified in the secondary validation screen had altogether at least three independent siRNAs showing the same phenotype (at least two siRNAs from the primary screen and at least one siRNA from the validation screen). Our primary and secondary validation screens were designed to identify factors required at all the stages of VSV infection such as virus entry and uncoating, gene expression, and virus budding. This screen also identified genes which were previously known to be involved in VSV gene expression such as vATPases further advocating the robustness of the screen.

Although there have been great advancement in the field of RNAi technology, the output from a genome-wide screen requires stringent validation. Several laboratories have made considerable efforts to identify host factors for other pathogenic human viruses such as HIV, Influenza and HCV. However, the limited number of overlapping genes identified from these screens for the same virus is a concern. Several factors potentially contribute to the outcome of such genomic screens e.g. types of cell line, types of siRNA, different strain of virus, type of statistical analysis performed etc (Barrows et al., 2010; Goff, 2008; Watanabe et al., 2010). Several other technical limitations of this approach also can not be ignored. For example, the stability of mRNA and knockdown efficiency for each gene is different. It is not possible to standardize the knockdown efficiency for each and every gene in the genome-scale screen. Also if the function of a particular gene is redundant with other genes, it may not be identified
because this approach knocks down a single gene at a time. Depletion of critical genes shows toxic effects on the host and these genes are also omitted from the analysis. Despite several limitations, the genome-wide siRNA screen provides an unparalleled advantage to probe for function of unknown genes not only in viral infection, but also in other biological conditions. Detailed investigation of functions of these genes will allow us to find novel drug targets against viral diseases. Also several drugs targeting host proteins are in clinical trials for influenza and HIV (Coley et al., 2009; Hayden, 2009; Ludwig, 2009; Moss et al., 2010). In our screen, because of the more stringent approach, the possibility of false positive genes in our hit list may be lower. Moreover, the high stringency of our hit identification might have compromised our ability to identify additional genes required for VSV infection.

We also tested the involvement of these genes in LCMV and HPIV3 infection. LCMV and HPIV3 belong to Arenaviridae and Paramyxoviridae family respectively. These viruses also contain negative strand RNA genome and replicate in the cytoplasm. There is also no comprehensive analysis of virus-host interaction for these viruses. Results obtained with VSV can be extrapolated to other negative strand RNA viruses. We show here that many of the genes identified as required factor for VSV are also involved in LCMV and HPIV3 infection. For example, COPI is required for VSV as well as for LCMV and HPIV3, further underscoring the notion of using VSV as a model virus. Genes involved in endosomal function such as vATPases are required for VSV and LCMV entry, but not for HPIV3 entry. In the secondary validation screen we also positively identified vATPase subunits for VSV and LCMV but not for HPIV3. Some genes are uniquely required by VSV such as BCL2L2, C2ORF28 and IFI35. The detailed
mechanism by which these genes impact VSV infection is not known. Some genes for example C11ORF42 are required by LCMV only where as GLYATL1 is required only by HPIV3. These results suggest that though these viruses use a subset of genes for their infection, there may be some mechanisms unique to each family of viruses.

Biological process classification revealed that the genes required for VSV infection were enriched in several functional categories such as exocytosis, vesicle mediated transport, protein transport among others. Many of these processes are uniquely identified for VSV, whereas some of the functions are shared by other virus families such as Influenza, HIV and HCV emphasizing the evolutionarily conserved as well as unique adaption displayed by these viruses.
CHAPTER V

HOST COATOMER PROTEIN COMPLEX I (COPI) FUNCTION IS REQUIRED FOR VSV RNA SYNTHESIS
5.1 COPI is required for early stage of VSV infection in HeLa cells

Network and biological function analysis of the validated candidates identified vesicle trafficking pathway as the top scoring pathway. Among these genes, there are six sub units of coatomer 1 (COPI) vesicular transport complex. The well characterized function of COPI is the retrograde transport of luminal and membrane proteins from Golgi to ER (Beck et al., 2009; Lee et al., 2004). COPI complex also plays a role in steady state distribution of proteins in the Golgi thereby maintaining anterograde transport (Tu et al., 2008). We focused on COPI pathway because knockdown of all the COPI sub units except COPE resulted in strong inhibition of VSV infection. Further examination showed a concomitant decrease in VSV gene expression when the COPI sub units were depleted by siRNA (Fig. 5.1 A). Six out of seven COPI sub-units (except COPE) were found to be required for VSV infection as determined by eGFP expression in a multi round infection (Fig. 5.1A). Similar results were obtained when we examined VSV M protein expression (data not shown). There was also a more than 50 fold reduction of infectious VSV production after depletion of COPI (data not shown). VSV G protein is processed through early secretory pathway and VSV assembly is affected if this process is disturbed (Orci et al., 1997; Ostermann et al., 1993). So, COPI depletion might have affected VSV G protein processing and consequently affected infectious virus production. Nevertheless, COPI depletion had a strong effect on VSV infection. Therefore, we were interested to examine if COPI plays a role in any other stage of VSV infection. To this end, we examined if the COPI components play a role in entry and uncoating as well as gene expression of VSV. The effect of depletion of COPI sub units on VSV gene expression was examined at 4 hrs post infection. Four hrs of infection is
considered as early stage of VSV infection and at this time point VSV gene expression can be readily detected but there is minimal virus budding and re-infection. Depletion of all the COPI subunits except COPE reduced VSV protein expression by 5-10 fold as evidenced by western blot analysis of M protein (Fig 5.1B). Decrease in protein expression could be due to reduced translation of protein from the mRNA or other steps prior to translation. To determine if there is any difference in RNA production, we examined VSV mRNA, genomic RNA and antigenomic RNA (replication product) after COPI knockdown. We found reduced level of transcripts as well as antigenome and genome RNA (Fig. 5.1C and data not shown). These results suggest that COPI functions at the level of VSV RNA synthesis by facilitating transcription and/or replication from the genome or enhancing the stability of RNA. Because depletion of multiple COPI subunits produced same phenotype on VSV infection, we considered that the effects of siRNAs are true effects due to protein depletion. Moreover, if two independent siRNAs produce the same phenotype, then the phenotype is generally considered on target. We verified protein expression of endogenous COPZ1 and COPB1 after gene knock-down by individual siRNA transfection. Figure 5.2A shows that two siRNA duplexes strongly reduced COPZ1 protein expression (> 90% inhibition of protein) whereas one siRNA duplex partially reduced COPZ1 expression. Three siRNA duplexes also significantly reduced COPB1 protein. When VSV M protein expression was examined under these conditions, we found that the more the gene is knocked down, the stronger is the inhibition of VSV M protein expression. To verify if other COPI subunits also affect similarly, we deconvoluted the pool of 4 siRNA and found that for each COPI sub units (except COPE), knock-down by at least two independent siRNAs
Figure 5.1: Effects of COPI depletion on VSV protein and RNA synthesis

(A) HeLa cells were transfected with pool of four siRNAs targeting each of seven COPI subunits, and then infected with VSV-eGFP for 16 hours. VSV gene expression was assessed by western blot analysis for eGFP. (B) HeLa cells were transfected as described in (A) but infected with VSV-eGFP for 4 hours. VSV M protein expression was examined by immunoblotting. Lower panel shows relative level of M protein expression normalized to non-targeting control sample. Data represent mean ± SD from three independent experiments. (C) Quantitative real time PCR measurement of VSV P RNA and VSV anti-genomic RNA. Values were obtained from duplicate measurements of each sample from two independent experiments and presented as mean ± SEM after normalizing to NT control.
Figure 5.2: Depletion of COPI sub-units affects VSV gene expression

HeLa cells were transfected with individual siRNA duplex targeting COPZ1 (A) and COPB1 (B) subunits, then infected with VSV-eGFP for 4 hours. VSV gene expression was assessed by western blot analysis for M protein. (B) HeLa cells were transfected with individual siRNA duplex targeting each of seven COPI sub units, then infected with VSV-eGFP for 4 hours. VSV gene expression was assessed by western blot analysis for M protein. Average relative fold expression of M protein normalized to NT control wells was presented from three independent experiments.
resulted in reduced VSV gene expression (Fig 5.2 B). It is of note that depletion of ARCN1 and COPA by all four siRNAs strongly reduced VSV gene expression. These results argues strongly infavor of involvement of COPI function in VSV gene expression. To further confirm the role COPI in VSV infection, we disrupted the function of COPI using a pharmacological agent 1, 3-cyclohexanebis methylamine (CBM), which has been shown previously to inhibit COPI function (Hu et al., 1999; Zhang et al., 2009b). Similar to siRNA results, we also observed that inhibition of COPI function using CBM affected VSV M protein synthesis at early stage of viral infection (Fig. 5.3 A). CBM treatment not only reduced VSV protein expression but also reduced VSV mRNA and anti genome levels at 3 hrs post infection (Fig. 5.3 B). When CBM was treated at 1 hr after VSV infection, more than 10 fold inhibition of VSV RNA expression was observed. This result suggests that COPI predominantly functions to regulate VSV RNA synthesis at the level of transcription and/or replication rather than affecting the primary transcription of input genomic RNA. The above experiment cannot completely exclude the possibility that COPI is involved in primary transcription of VSV. Collectively the results suggest involvement of COPI in early stage of VSV infection primarily regulating VSV RNA synthesis.

5.2 COPI is required for early stage of LCMV and HPIV3 infection

Our validation screen also identified involvement of COPI in arenavirus (LCMV) and paramyxovirus (HPIV3) infection. Similar to VSV, these viruses also replicate in the cytoplasm. Because COPI was found to be required for early stage of VSV infection, we examined the role of COPI in early stage of LCMV and HPIV infection. We knocked down COPI function by siRNA and subsequently infected with LCMV and HPIV3. Cell
lysates were collected at 8 hr post infection for LCMV and 14 hr post infection for HPIV3. As a measure of replication, level of GFP protein was examined by western blot analysis. Similar to VSV, depletion of the 6 sub-units of COPI also significantly affected LCMV and HPIV3 gene expression at early stage of infection (Fig. 5.4). Taken together, COPI is required not only for VSV RNA synthesis but also is required early in infection in case of other cytoplasmic replicating negative-strand RNA viruses such as LCMV and HPIV3.

5.3 COPI functions to regulate VSV RNA synthesis

Previous reports have identified a role for COPI in VSV G protein processing (Orci et al., 1997; Ostermann et al., 1993). Also βCOP is involved in endosomal transport (Aniento et al., 1996; Whitney et al., 1995) and VSV entry is mediated by endosomal acidification (Le Blanc et al., 2005). COPI has multiple functions inside the cell and might have effects on multiple stages of VSV infection. Therefore, we wanted to examine in detail how COPI is modulating VSV infection. To rule out the possibility that the effect of COPI in early stage of VSV infection is dependent on its function to process G protein, we used a recombinant VSV without encoding a G protein (VSVΔG) (Fig. 5.5 A). Infectious viruses were generated by supplementing G protein by exogenous expression during infection. VSVΔG virus can undergo single round of infection but can not bud and re-infect new cells as the viral genome lacks G gene. Reduction inVSV gene expression was also observed when the COPI-depleted cells were infected with VSVΔG (Fig. 5.5 B). These results further strengthen the conclusion that COPI plays a role in an early stage of VSV infection possibly by facilitating entry and/or gene expression of
Figure 5.3: Inhibition of VSV gene expression by CBM

(A) HeLa cells were infected with VSV-eGFP at 0.5 MOI. After 1 hr of infection, cells were washed and CBM was added at the indicated concentrations. Cell lysate was prepared at 4 hours post infection and immunoblot was performed for M protein. (B) HeLa cells were infected with VSV-eGFP. At 1 hr post infection cells were washed and treated with CBM. At 4 hours post infection, RNA was extracted and quantitative RT PCR was performed for VSV P mRNA and anti-genome. Value represents mean ± SEM of duplicate measurements from two independent experiments.
Figure 5.4: COPI silencing decreases gene expression of LCMV and HPIV3

(A) HeLa cells were transfected with pool of 4 siRNAs against the indicated COPI subunits and 42 hours post transfection, infected with LCMV for 6 hours. Level of eGFP was detected by western blot. (B) HeLa cells were transfected with siRNAs against the indicated COPI subunits and then infected with HPIV for 14 hours. Level of GFP was detected by western blot.
VSV. To further dissect the role of COPI in VSV entry, uncoating or VSV gene expression; we prepared nucleocapsids from VSV virions and transfected these nucleocapsids to COPI-depleted cells. Nucleocapsids are functional units for VSV gene expression and can initiate gene expression after being delivered into the cells. The transfection of nucleocapsids also bypasses the endosome mediated entry and uncoating steps of VSV. There was a 2-3 fold reduction of VSV gene expression from the nucleocapsids when COPI function was depleted as compared to NT siRNA transfected cells (Fig. 5.5 C). These results suggest a role of COPI in VSV RNA synthesis. However, because the transfection of nucleocapsids bypasses the entry and uncoating steps, our results cannot exclude a role of COPI in VSV entry and uncoating. Furthermore, when VSV infected cells were treated with CBM 1 hr post infection, strong inhibition of VSV protein synthesis and RNA synthesis was observed (Fig. 5.3). We and others have also shown that VSV entry and uncoating is rapid and it takes place within 30 minutes to 1hr after virus adsorption (Das et al., 2009; Johannsdottir et al., 2009). To further verify that VSV entry and uncoating is rapid and occurs within 1 hour of infection, we used Bafilomycin A1, an inhibitor of vATPases and thereby block endosomal mediated VSV entry (Le Blanc et al., 2005). Pre treatment of cells (30 minutes before infection) with Bafilomycin A1 blocked VSV infection completely whereas VSV infection is insensitive to this drug when treated simultaneously with infection or 1 hour post infection suggesting that VSV entry and uncoating is rapid. CBM treatment 1 hr post infection and subsequent inhibition of VSV gene expression suggest a role of COPI in VSV RNA synthesis. Collectively these results suggest that COPI is required for VSV gene
Figure 5.5: COPI silencing affects VSV gene expression independent of VSV entry and budding

(A) Schematic representation of VSV-PeGFPΔG virus. (B) HeLa cells were transfected with the indicated siRNAs, then infected with VSV PeGFP ΔG virus. At 5 hr post infection, cells were harvested and M protein expression was assessed by immunoblotting. Lower panel shows relative level of M protein expression from three independent experiments. (C) VSV nucleocapsids prepared from infectious viruses were transfected to cells depleted of COPZ1 and ARCN1. At 6 hrs post nucleocapsid transfection, M protein expression was examined by western blot analysis. Lower panel shows relative level of M protein expression from three independent experiments.
expression and this function of COPI is independent of effect of COPI on G protein processing or virus entry through endosome.

5.4 ARF1 mediated COPI assembly is required for VSV and LCMV but not for HPIV3 gene expression

COPI assembly on the Golgi membrane requires activation of ARF1 (Hsu et al., 2009; Hsu & Yang, 2009; Lippincott-Schwartz & Liu, 2006). Upon activation, ARF1 recruits a pre-formed COPI complex from the cytoplasm to the Golgi membrane. Because ARF1 is the immediate upstream effector of COPI, we were interested to examine involvement of the ARF1 in VSV infection. We knocked down ARF1 by siRNA, infected with VSV and analyzed VSV M protein expression. VSV protein expression was reduced by 40% as examined by western blot analysis (Fig. 5.6 A). Under this knock-down condition using the siRNA available to us, ARF1 protein level was also decreased by approximately 50% and this correlated with moderate reduction of VSV protein expression (60% of NT control). To re-confirm the role of ARF1 in VSV gene expression, we transfected HeLa cells with a dominant negative mutant of ARF1 [ARF1 T31N, (Peters et al., 1995)] and examined VSV gene expression. Similar to ARF1 knock-down, moderate inhibition of VSV gene expression was observed after disrupting ARF1 function by expression of the dominant negative ARF1. Further evidence for the involvement of ARF1 in early stage of VSV infection came from experiments with VSV-ΔG virus. Depletion of ARF1 also affected gene expression of VSV-ΔG virus suggesting a role for ARF1 mediated COPI assembly on VSV gene expression rather than G protein processing (data not shown). Similar to COPI knockdown, ARF1 depletion also affected VSV mRNA level as well as genome replication (Fig. 5.6 C). When we examined the role of ARF1 in LCMV and
Figure 5. 6: ARF1 is required for VSV gene expression

(A) HeLa cells were treated with the pooled siRNAs against ARF1. At 72 hrs post transfection, cells were infected with VSV-eGFP at 0.05 MOI. Levels of M protein and ARF1 expression were examined at 6 hrs post infection. Values represent mean ± SD from three independent experiments. (B) HeLa cells were transfected with the plasmids expressing ARF1\textsubscript{wt} or ARF1\textsubscript{T31N}. Twenty four hrs post transfection, cells were infected with VSV-eGFP for 4 hrs. Levels of M protein and ARF1 expression were examined by western blot. (C) Quantitative real time PCR for VSV P mRNA and antigenome after ARF1 depletion. Values were obtained from duplicate measurements of each sample from two independent experiments and presented as mean ± SEM after normalizing to NT control. (D) HeLa cells were transfected with pooled siRNA against ARF1 and then infected with LCMV for 6 hrs. Level of eGFP expression was examined by western blot analysis. (E) \textit{Renilla reniformis} GFP expression after HPIV infection was assessed in cells depleted of ARF1.
HPIV-3 gene expression, we found that ARF1 depletion affected LCMV gene expression by 50-60% of that of NT control cells where as HPIV3 gene expression appeared to be independent of ARF1 function (Fig. 5.6 D and 5.6 E).

5.5 GBF1, the Guanine Nucleotidase Exchange Factor (GEF) for ARF1 is required for VSV, LCMV and HPIV3 infection

ARF1 is post translationally modified by myristoylation at its amino terminus. In the GDP bound (inactive) form, the myristoylated motif is sequestered but in GTP bound (active) form this motif is exposed. This myristoylated motif is required for membrane targeting of ARF1 (Hsu & Yang, 2009). GBF1 is the guanine nucleotide exchange factor for ARF1 and GBF1 maintains COPI assembly on the Golgi membrane through ARF1 activation (Garcia-Mata et al., 2003; Kawamoto et al., 2002). Because of the crucial role of GBF1 in COPI pathway, we investigated involvement of GBF1 in VSV gene expression. We silenced GBF1 by siRNA and then infected with VSV. Treatment with siRNA reduced endogenous GBF1 protein level by 90-95% (Fig 5.7 A). When GBF1 was knocked down, a 2.5 fold reduction in VSV protein synthesis was observed (Fig. 5.7 A). Additionally a concomitant 2-2.5 fold reduction of VSV mRNA and anti-genome level were observed following GBF1 silencing (Fig 5.7 B). Taken together, these results suggest that VSV gene expression critically depends on GBF1 functions.

Furthermore, we used pharmacological inhibitors of GBF1 to probe for the role of GBF1 in VSV gene expression. To this end, we used brefeldin A (BFA), Golgicide A (GCA) and Tyrphostin AG1478 (AG1478) to inhibit function of GBF1. BFA is a well studied pharmacological inhibitor of GBF1, Brefeldin Inhibited Guanine nucleotide exchange factor -1 and -2 (BIG1 and BIG2) (Peyroche et al., 1999)
Figure 5.7: GBF1 controls VSV RNA synthesis

(A) HeLa cells were transfected with pooled siRNAs against GBF1. At 72 hours post transfection, cells were infected with VSV- eGFP for 4 hours. Levels of M protein and GBF1 were examined by western blot. Lower panel shows relative level of M protein expression from 3 independent experiments. (B) HeLa cells were transfected with GBF1 siRNA for 72 hrs, then infected with VSV-eGFP for 3 hrs. Quantitative real time PCR was performed for VSV P mRNA and anti-genome. Values represent mean ± SEM from duplicate measurements from two independent experiments.
whereas, GCA and AG1478 were recently identified and characterized as being more specific inhibitors of GBF1 (Pan *et al.*, 2008; Saenz *et al.*, 2009). Similar to siRNA mediated depletion of GBF1, VSV protein synthesis (Fig. 5.8 A) and RNA synthesis (Fig. 5.8 B) were also sensitive to BFA, GCA and AG1478 treatment. All the three drugs decreased VSV protein level and RNA level by 2-3 fold (Fig. 5.8). To exclude the possibility that the effects of these drugs on VSV gene expression were nonspecific due to unknown off-target effects rather than specific inhibition of GBF1 function, we used Madin Darby canine kidney (MDCK) cells. Dog genome has a naturally occurring mutation (M832L) in sec7 domain of GBF1 and therefore MDCK cells are resistant to BFA, GCA and AG1478 (Lanke *et al.*, 2009; Pan *et al.*, 2008; Saenz *et al.*, 2009; Verheije *et al.*, 2008). However, in these cells other Guanine nucleotide exchange factors regulating trans-Golgi network functions are sensitive to BFA action (Saenz *et al.*, 2009; Wagner *et al.*, 1994). When MDCK cells were infected with VSV and then treated with the GBF1 inhibitory drugs, no effect on VSV gene expression was observed even at higher concentrations of BFA, GCA and AG1478 (Fig. 5.9 A). In another approach, we used BER40 cells. BER40 cell line is a derivative of Vero cell and has a mutation (A795E) in GBF1 (Belov *et al.*, 2008) which confers resistance to BFA in the BER40 cell line. A 2-fold suppression of VSV gene expression was observed in Vero cells after BFA treatment whereas in Ber40 cells VSV gene expression was insensitive to the effect of BFA (Fig. 5.9 B). To further provide support for this result that these pharmacological inhibitors indeed affected VSV gene expression, time-course experiment was performed. In this context, BFA, GCA and AG1478 were added at every 1 hr interval to the cells before or following VSV infection and the level of VSV M protein expression
Figure 5.8: VSV RNA synthesis is sensitive to GBF1 inhibitors

(A) HeLa cells were infected with one MOI of VSV-eGFP and at 1 hr post infection, cells were treated with indicated concentrations of the drugs. Cell lysate prepared at 4 hr post infection and M protein expression was assessed by immunoblotting. Lower panel shows relative levels of M protein expression from 3 independent experiments. Values represent mean ± SD. (B) HeLa cells were infected with VSV-eGFP and 1 hpi treated with BFA (2 μg/ml), GCA (10 μM) and AG1478 (30 μM) for 2 hr. Fold change in P mRNA and anti-genome was examined by quantitative RT PCR. Values represent mean ± SEM from duplicate reactions from two independent experiments.
Figure 5.9. **GBF1 is the GEF required for VSV RNA synthesis**

(A) MDCK cells were infected with VSV-eGFP. At 1 hr post-infection cells were treated with the indicated concentrations of the drugs. M protein expression was examined by immunoblotting at 4 hrs post infection. Lower panel shows relative levels of M gene expression from three independent experiments. Values represent mean ± SD. (B) BER 40 and vero cells were infected with VSV-eGFP. After 1 hr, cells were treated with BFA at indicated concentrations. M protein expression was analyzed by western blot analysis.
was determined by immunoblotting (Fig. 5.10). Results showed that VSV M protein synthesis was affected when the drugs were added up to 3 hours post infection suggesting that VSV RNA synthesis requires GBF1 activity. Experiments with VSV-ΔG virus and nucleocapsids transfection also showed that GBF1 is required for VSV gene expression (data not shown). Collectively, the data unequivocally suggest that GBF1 function is required for VSV RNA synthesis.

We next examined the requirement of GBF1 in LCMV and HPIV 3 gene expression. For this, control or GBF1 depleted cells were challenged with LCMV and HPIV-3 and level of GFP expression was examined by western blot analysis. Depletion of GBF1 reduced gene expression from both LCMV and HPIV-3 by more than 10 fold (Fig. 5.11A and 5.11B). Similar to the results obtained with siRNA treatment, we observed that LCMV and HPIV-3 gene expression were also sensitive to pharmacological inhibitors of GBF1 (Fig 5.11 C and 5.11 D). Thus from the studies reported here, we conclude that GBF1-ARF1-COPI mediated pathway regulates gene expression of diverse families of cytoplasmic replicating RNA viruses.

5.6 GBF1 and COPI are not part of VSV replication complex

Many positive strand RNA viruses re-organize host intracellular membranes and utilize them as sites of their genome replication (Hsu et al., 2010). Because of the critical role of ARF1, GBF1 and COPI in formation of replication complexes for positive-strand RNA viruses, we wanted to investigate if these factors play similar roles in VSV infection. We examined sub-cellular localization patterns of endogenous βCOP and GBF1 in VSV infected cells. In uninfected cells, βCOP was predominantly localized to a
Figure 5.10: GBF1 affects a post entry stage of VSV infection

HeLa cells were infected with VSV-eGFP and these cells were treated with Brefeldin A, Golgicide A or AG1478 at the indicated time points pre or post infection. M protein expression was analyzed by western blot analysis at 4 hours post-infection.
Figure 5.11: Requirement of GBF1 for LCMV and HPIV gene expression

(A) HeLa cells were transfected with GBF1 siRNA and subsequently infected with LCMV for 6 hrs. LCMV gene expression was measured by immunoblotting of eGFP.

(B) HeLa cells were infected with HPIV3 after depleting GBF1. HPIV gene expression was assessed by western blot for GFP.

(C) HeLa cells were infected with LCMV. At 1 hour post-infection, cells were treated with indicated concentrations of the drugs for 6 hours. LCMV gene expression was assessed by western blot examination of eGFP.

(D) HeLa cells were infected with HPIV3 and 1 hr post-infection treated with indicated concentrations of the drugs. HPIV3 gene expression was assessed by western blot for GFP.
perinuclear site and co-localized to Golgi apparatus (Fig. 5.12). In VSV infected cells the localization of βCOP was not altered significantly. A previous report from our lab as well as others have shown that VSV replication occurs throughout the cytoplasm and the punctuate structures (marked by PeGFP) are the sites for genome replication those also contain viral N, L proteins as well as viral RNA (Das et al., 2006). In VSV-infected cells, βCOP staining did not absolutely co-localize with VSV replication sites (Fig. 5.12). However, punctate structures marked with PeGFP were found in close proximity to βCOP. Similar results were also obtained when we examined distribution of GBF1 and ARF1 (Fig 5.12 and data not shown). These results suggest that although GBF1 and ARF1 mediated COPI activation is required for VSV infection; these proteins may not be components of VSV replication complex rather they participate indirectly in modulating VSV infection.
Figure 5.12: GBF1 and COPI do not co-localize with VSV replication complex

HeLa cells were mock-infected or infected with VSV-PeGFP. At 3 hours post infection, cells were fixed and stained with endogenous COPB1 (A) or GBF1 (B). Images were captured in Laser scanning confocal microscope. All fluorescence images were merged Z stacks of 0.5μM thickness.
5.7 Discussion

This screen identified involvement of COPI complex in VSV gene expression. COPI complex was originally identified 25 years ago (Malhotra et al., 1989; Orci et al., 1986). This complex consists of seven subunits such as α, β, β’, δ, ε, γ and ζ (Hsu et al., 2009). COPI complex has been shown to mediate transport of cellular proteins and cargo from the Golgi to the endoplasmic reticulum as well as intra Golgi transport (Beck et al., 2009; Orci et al., 1997; Rabouille & Klumperman, 2005). ER resident membrane protein cargo can be retrieved back to ER from Golgi by interaction with COPI (Letourneur et al., 1994). Recent evidences also suggest a role of COPI in conjunction with COPII in the ER to Golgi transport of proteins (Scales et al., 1997). VSV-G protein has been used widely as cargo to probe anterograde transport mechanisms. Depletion of COPI perturbs the steady state distribution of the Golgi enzymes (Tu et al., 2008) and in turns inhibits transport of VSV G protein to the plasma membrane for budding. However, we observed reduction of VSV gene expression at early stage of infection in cells depleted of COPI. Similar effect was also observed when we used a recombinant VSV without G protein. These results suggest a role of COPI in early stage of VSV infection and the effect is independent of VSV G protein processing and infectious virus production. VSV infection involves entry of virus particles through endosomal compartments and subsequent release of nucleocapsids to cytoplasm. COPI sub units such as β COP was shown to be present on endosomes suggesting involvement of a subset of COPI vesicles in regulation of endosome function (Aniento et al., 1996; Whitney et al., 1995). The effect of COPI depletion on early stage of VSV infection could be due to blockage of VSV entry and/or RNA synthesis. Further verification of the role of COPI in VSV gene expression came
from nucleocapsid transfection experiments. Nucleocapsids transfection bypasses the endosomal entry steps and thus they are delivered directly to the cytoplasm. Moreover, results from the drug (CBM, BFA, GCA and AG1478) treatment studies further strengthen the conclusion that COPI function is required for VSV gene expression. Furthermore, treatment of HeLa cells with 2 μM Bafilomycin A1 30 minutes before infection completely blocked VSV infection. However, treatment of Bafilomycin A1 during infection or 1 hour after infection did not exert its inhibitory effect on VSV RNA synthesis. These results confirm that VSV entry and uncoating is rapid, and the effect we observed with CBM, BFA, GCA and AG1478 on VSV infection is in fact due to interference with VSV RNA synthesis.

We did not find components of COPII complex either in the primary screen or in the validation screen as required factors in VSV infection. Depletion of one of the COPII subunits resulted in statistically insignificant reduction of VSV infection. VSV G protein transport is dependent on COPII function (Rowe et al., 1996). The failure to identify COPII components as required factors could be due to improper knockdown or other compensatory mechanisms. However, COPI subunits emerged strongly in the primary screen as well as in the validation screen. The requirement of COPI for genome replication has been well-documented for positive strand RNA viruses such as poliovirus, HCV and Drosophila C virus (Cherry et al., 2006; Maynell et al., 1992; Tai et al., 2009). The studies presented here revealed the requirement of COPI for gene expression of negative-strand RNA viruses further providing evidence for critical role of secretory pathway in virus infection. Not only COPI, but also ARF1 and GBF1 were also found to be required for VSV RNA synthesis. All positive strand RNA viruses replicate in close
association with membranous complexes (Ahlquist et al., 2003; Salonen et al., 2005). Several positive strand RNA viruses such as polio virus, HCV, CVB3 modulate the activities of ARF1, GBF1 and COPI for formation of the intracellular sites of replication (Belov et al., 2008; Goueslain et al.; Hsu et al., 2010; Matto et al.; Verheije et al., 2008). However, the nature of replication organelles for cytoplasmically replicating RNA viruses is not known. It has been demonstrated in earlier reports that VSV genome replication occurs throughout the cytoplasm (Das et al., 2006; Heinrich et al., 2010). In this study, we did not observe co-localization of VSV replication proteins such as P, N and L with ARF1, GBF1 or COPI suggesting that these cellular proteins may not be directly involved with the VSV replication complex formation. Moreover, localization of endogenous ARF1, GBF1 and COPI was not altered in VSV infected cells, as compared to uninfected cells. Several cellular signaling pathways emanate from GBF1/ARF1/COPI pathway including phophotidyl ionositol pathway, actin cytoskeleton organization and lipid droplet formation (Donaldson, 2005; Guo et al., 2008; Myers & Casanova, 2008). It is possible that the downstream signaling pathway is affected when the COPI pathway is perturbed and thus COPI may indirectly regulate VSV gene expression.

LCMV and HPIV3 also replicate in the cytoplasm. LCMV entry occurs by receptor-mediated endocytosis whereas HPIV3 enters by fusion of the viral envelope with the host cell plasma membrane that is independent of endosomal function. So, the effect of COPI depletion on early stage of HPIV3 infection rules out the possibility that COPI affect entry of HPIV3. However, LCMV entry also could be mediated by COPI function. Drug treatment (BFA, GCA and AG1478) experiments suggest a role of COPI in gene expression of LCMV and HPIV3 which is independent of effect of COPI on
endosome or glycoprotein processing. Moderate effect on VSV gene expression was observed when ARF1 was silenced whereas HPIV3 gene expression remained insensitive to ARF1 depletion in our experimental setting. We also observed moderate inhibition of ARF1 protein expression after knockdown of ARF1. Multiple ARF proteins exist in the cell and it is possible that depletion of ARF1 function can be complemented by other ARF proteins. Further experiments are required to probe in detail the mechanism of COPI requirements for LCMV and HPIV3 replication.

Though multiple siRNA screens have identified COPI complex as a host cofactor for influenza virus infection (Brass et al., 2009; Karlas et al., 2010; Konig et al., 2010), these studies differ in their conclusion regarding role of COPI in influenza virus infection. One study has pointed out the role of COPI in HA protein surface expression (Brass et al., 2009), whereas another study has suggested a role of COPI in influenza entry (Konig et al., 2010). The discrepancies in their results might be due to types of assay used. However, for cytoplasmic replicating negative strand virus such as VSV, our results point towards a role of COPI in RNA synthesis. It is not clear however, whether COPI functions are directly required for VSV RNA transcription and/or replication or are required to maintain stability of transcripts and genome.

To conclude, this study provided a resource to explore further the role of the candidate cellular proteins in VSV infection. It is further expected that this resource will serve as a meaningful starting point for investigators studying virus-host interaction not only for VSV, but also for other negative strand RNA viruses such as LCMV and HPIV3. Mechanistic understanding of involvement of these host factors in VSV infection will illuminate potential new therapeutic targets against viral diseases.
Chapter VI
SUMMARY AND CONCLUSION

Viruses utilize the host factors to complete their life cycle. Using VSV as a model pathogen studies reported in this dissertation have identified several unique host factors required for VSV infection. Future studies may unravel the mechanistic involvement of these host factors in VSV replication and gene expression. The following conclusions can be derived from these studies.

1. **Expression of VSV replication proteins is not toxic for the cells and a cell line stably expressing replication proteins can be generated**

   VSV M has been implicated for the cytopathic effects of VSV. However M protein deletion mutant viruses also exert cytopathic effects suggesting involvement of other viral proteins. This study showed that expression of replication proteins of VSV such as N, P and L has no adverse effect on cell viability. A stable cell line expressing N, PeGFP and L could be established and maintained in culture for at least three months without loss of the replication functions of the proteins. Functionality of these proteins expressed in the stable cell line was evidenced by efficient replication of DI particle genome.

2. **Long term replication of DI genome or sub-genomic replicon in stable helper cell line could not be established**

   We were interested in establishing a replicon system for VSV sub genomic replicon. Maintenance of DI particle genome in the stable helper cell line on long term basis was not successful. After 5th passage the cells lost the DI particle genome. This study could not establish a cell line supporting replication of a subgenomic replicon of VSV with a drug resistant marker.
3. Replication of DI genome activates IFN signaling

While investigating the reasons for failure to establish a long term DI genome replicating cell line, we found that DI particle replication in these helper cell line activated IFN response. Activated IFN response was observed in cells replicating DI RNA but not in cells cured of DI RNA genome. However, this cell line provided us a unique tool to examine genome replication dependent IFN activation. We further showed that not mere entry but the replication of DI RNA genome activates IFN response. It has been reported previously that rhabdovirus genomic RNAs (VSV and rabies virus) induce IFN in replication-independent manner (Hornung et al., 2006; Pichlmair et al., 2006). VSV DI genomes are synthesized in the form of nucleocapsid wrapped by the N protein and possible formation of dsRNA structures in infected cells has not been reported. The mechanism by which VSV DI RNA replication activates IRF3 and NFκB is unknown at this time.

4. Genome wide siRNA screen identified host proteins required for VSV infection

To identify the host proteins influencing VSV infection, we conducted a genome wide siRNA screen approach. So far there has not been such report identifying host genes implicated in VSV infection. Functional genomics approach identified host genes in an unbiased manner. Several novel candidates were identified in this screen. Our screen also identified several known regulators of VSV infection lending further support about strength of our screen. These results provide a strong platform to develop a comprehensive virus-host interaction map for VSV.

5. Host COPI complex is involved VSV RNA synthesis
This study identified involvement of COPI in early stage of VSV infection. Further experiments suggested a role of COPI in early stage of VSV infection and the effect is independent of VSV G protein processing and infectious virus production. To further dissect the role of COPI in VSV entry and gene expression, we transfected functional nucleocapsids to COPI depleted cells. Nucleocapsids are functional units for VSV gene expression and the transfection of nucleocapsids also bypasses the endosome mediated entry and uncoating steps of VSV. Nucleocapsid transfection experiments confirmed a role of COPI in VSV RNA synthesis. Pharmacological inhibitors of COPI also revealed a role of COPI in VSV RNA synthesis.

6. COPI is involved in RNA synthesis of LCMV and HPIV3

To further examine possible involvement of COPI in other cytoplasmic replicating RNA viruses, we used LCMV and HPIV3. We showed that, similar to VSV, early stage of LCMV and HPIV3 infection also depends on COPI function. It remains to be determined whether COPI regulates RNA synthesis for these viruses or not. Nevertheless, these results suggest a conserved function of host COPI in cytoplasmic replicating negative strand RNA viruses.

7. ARF1 and GBF1 mediated COPI assembly is required for VSV RNA synthesis

ARF1 and GBF1 are upstream regulators of COPI function. Upon activation, ARF1 mediates assembly of a preformed COPI complex onto Golgi membrane. When we reduced ARF1 expression by siRNA and subsequently infected with VSV, we found a moderate inhibition of VSV RNA synthesis. Furthermore, a dominant negative mutant of ARF1 also inhibited VSV RNA synthesis. These results suggest that ARF1-regulated COPI function is required for VSV RNA synthesis. GBF1 is the guanine nucleotide
exchange factor for ARF1. Our results also revealed a role of GBF1 in VSV RNA synthesis. Collectively GBF1-ARF1-COP1 pathway regulates VSV RNA synthesis.

8. COPI is not a part of VSV replication complex

Positive strand RNA viruses replicate in association with double membrane vesicles. Some of the positive strand RNA viruses require COPI complex for formation of their replication complex. Because ARF1, GBF1 and COPI play critical roles in formation of replication complexes for positive strand RNA viruses, we examined their possible involvement in VSV replication complex formation. In contrast to positive strand RNA viruses, ARF1, GBF1 and COPI did not co-localize with VSV replication complex. The sub-cellular distribution of these proteins also did not change following VSV infection compared to non-infected cells. These results suggest that COPI may not be involved directly in VSV replication complex assembly. Further studies addressing downstream effectors of COPI functions will be needed to have a better and mechanistic understanding of the involvement of COPI in VSV and other negative-strand RNA virus replication.

9. Integrated model of VSV-host interaction

Based on the results obtained in this dissertation, I propose an integrated model of VSV-host interaction (Fig. 6.1). The candidate genes were placed at the positions most likely relevant to VSV lifecycle. The annotations were gathered from Entrez Gene and Uniprot and refined by manual examination of gene functions using Entrez Pubmed. Host factors required for all three viruses are identified in green whereas host factors required by VSV are shown in red.
**Figure 6.1:** Integrated model of VSV-host interactions
REFERENCES


**Kawamoto, K., Yoshida, Y., Tamaki, H., Torii, S., Shinotsuka, C., Yamashina, S. & Nakayama, K. (2002).** GBF1, a guanine nucleotide exchange factor for ADP-
ribosylation factors, is localized to the cis-Golgi and involved in membrane association of the COPI coat. *Traffic* 3, 483-495.


