

## Adding Genes to the RNA Genome of Vesicular Stomatitis Virus: Positional Effects on Stability of Expression

Gail W. Wertz,\* Robin Moudy, and L. Andrew Ball

Department of Microbiology, University of Alabama School of Medicine, Birmingham, Alabama 35294

Received 27 February 2002/Accepted 25 April 2002

**Gene expression of the nonsegmented negative strand (NNS) RNA viruses is controlled primarily at the level of transcription by the position of the genes relative to the single transcriptional promoter. We tested this principle by generating engineered variants of vesicular stomatitis virus in which an additional, identical, transcriptional unit was added to the genome at each of the viral gene junctions. Analysis of transcripts confirmed that the level of transcription was determined by the position of the gene relative to the promoter. However, the position at which a gene was inserted affected the replication potential of the viruses. Adding a gene between the first two genes, N and P, reduced replication by over an order of magnitude, whereas addition of a gene at the other gene junctions had no effect on replication levels. All genes downstream of the inserted gene had decreased levels of expression, since transcription of the extra gene introduced an additional transcriptional attenuation event. The added gene was stably maintained in the genome upon repeated passage in all cases. However, expression of the added gene was stable at only three of the four positions. In the case of insertion between the N and P genes, a virus population arose within two passages that had restored replication to wild-type levels. In this population, expression of the additional gene as a monocistronic mRNA was suppressed by mutations at the end of the upstream (N) gene that abolished transcriptional termination. Because transcription is obligatorily sequential, this prevented transcription of the inserted downstream gene as a monocistronic mRNA and resulted instead in polymerase reading through the gene junction to produce a bicistronic mRNA. This eliminated the additional attenuation step and restored expression of all downstream genes and viral replication to wild-type levels. These data show that transcriptional termination is a key element in control of gene expression of the negative strand RNA viruses and a means by which expression of individual genes may be regulated within the framework of a single transcriptional promoter. Further, these results are directly relevant to the use of NNS viruses as vectors and vaccine delivery agents, as they show that the level of expression of an added gene can be controlled by its insertion position but that not all positions of insertion yield stable expression of the added gene.**

Many of the key principles involved in control of transcription of the nonsegmented negative strand (NNS) RNA viruses of the order *Mononegavirales* were elucidated using the prototypic rhabdovirus, *Vesicular stomatitis virus* (VSV). VSV has an 11-kb negative-sense RNA genome. At the 3' terminus of the genome there is a 47-nucleotide (nt) leader RNA sequence followed, in order, by the five genes encoding viral nucleocapsid protein (N), phosphoprotein (P), matrix protein (M), attachment glycoprotein (G), and the viral RNA-dependent RNA polymerase (L), and ending with the 54-nt 5' trailer RNA sequence (23). The leader sequence contains elements essential for polymerase binding and for signaling replication and transcription (16, 35). The trailer contains sequences required in replication and also a *cis*-acting signal for assembly of newly replicated RNAs into particles (15, 35).

It was first shown with VSV that transcription of the NNS RNA viruses initiates at a single 3'-proximal promoter site and that monocistronic mRNAs are generated by obligatorily sequential transcription of each of the genes (1, 3). Within the framework of the single polymerase entry site, there are conserved sequences at the beginning and end of each gene and at

the gene junctions that signal initiation and termination of each mRNA (4–6, 29, 30). Further, the transcriptional termination of each upstream gene is required for initiation at the next gene downstream, consistent with the sequential nature of transcription (5). Transcription is the major step at which expression of the genes of VSV is controlled. The promoter proximal gene is transcribed in greatest abundance, and each successive downstream gene is transcribed in progressively lower amounts (31). This reduction in transcript yield is referred to as attenuation (14). It occurs specifically as the polymerase crosses each gene junction, results in reduction of the downstream product by approximately 25 to 30%, and is thought to be due to dissociation of polymerase during some step in termination or reinitiation (14).

Sequences at each gene end, the intergenic dinucleotide and the downstream gene start, referred to together as the gene junction, are well conserved in VSV (23). Extensive mutagenesis of the gene junction has identified sequences that are required for mRNA polyadenylation, termination and downstream initiation, and modification (5, 6, 13, 28–30). The sequences required for termination and polyadenylation are the conserved 3'...AUACUUUUUU...5' sequence found at the end of each gene and the first nucleotide of the intergenic dinucleotide 3'..GA..5' (5, 6, 29). Alteration of any nucleotide of the AUAC decreases the efficiency of termination, and alteration of the C residue to any other nucleotide abolishes

\* Corresponding author: Department of Microbiology, BBRB 17/Rm. 366, University of Alabama School of Medicine, Birmingham, AL 35294. Phone: (205) 934-0877. Fax: (205) 934-1636. E-mail: gailw@uab.edu.

termination. Similarly, shortening the U7 tract by even a single nucleotide, or interrupting it with a heterologous nucleotide, eliminates termination and polyadenylation (5). Moreover, the base composition of the AUAC tetranucleotide controls the ability of the polymerase to slip on the U7 tract and generate polyadenylate tails (4).

Transcription of each downstream mRNA was initially shown to require that the sequence 3'...UUGUCNN...5' be conserved at the start of each gene and to be affected by the second nucleotide of the intergenic dinucleotide (6, 29, 30). Recent work has shown that other upstream sequences are also involved (9a). As mentioned above, termination of each upstream gene is required for the polymerase to be able to transcribe the gene immediately downstream. If the polymerase fails to terminate at the end of a gene, it then reads through the gene junction and synthesizes a readthrough RNA rather than a monocistronic mRNA from the downstream gene.

The controlled expression of genes is important to the viral life cycle. For example, the amount of N protein is critical for optimal genome replication, as was shown previously by generating engineered VSVs in which the position of the N gene was altered from its 3'-proximal position to successive distal positions (33). We found that transcription of the N gene was reduced in a stepwise manner as the N gene was moved to successively promoter-distal positions and, correspondingly, overall viral genome replication was reduced in a stepwise manner (33). Thus, the replicative capability of these viruses can be controlled by altering the transcription level of this key gene.

In the present work, we tested several aspects of the principles involved in the transcriptional control of gene expression of the negative-stranded RNA viruses by inserting an additional transcriptional unit at each of the VSV gene junctions. This allowed us to test whether the position at which the gene was placed controlled expression levels and whether all of the viral gene junctions could accept an additional gene equally readily.

Heterologous genes have been expressed from VSV and many other NNS RNA viruses from a variety of positions in the genome, with varying results as to expression levels (7-9, 17, 24). Additionally, the effects of adding an extra gene on replication have varied widely, in some cases having little or no effect and in others causing more than a 100-fold decrease in replication. In some cases where replication was impaired, the product of the added gene may have been inhibitory to replication. In other cases, sequence alterations, such as restriction enzyme sites introduced during cloning, may have affected replication or expression levels. To date, however, in no case has the expression and stability of an identical gene been examined from all available positions.

In the work reported here, we compared the effect of inserting an identical transcriptional unit, in an identical manner with no alteration in the gene junction sequences, at each of the gene junctions on genomic stability, stability of expression, and effects on viral replication. Importantly, the added transcriptional unit was designed to express no protein product that might confer selective advantage or disadvantage on the virus. The transcriptional unit added at each of the internal VSV gene junctions contained the conserved VSV transcription start and stop sequences but lacked a translation initiation sequence or a major

open reading frame, thus avoiding the possibility of feedback of a protein product affecting any step of transcription or replication. Further, transcription was assayed by metabolic labeling of mRNA, which allowed us to compare the viral transcripts directly and avoided issues of relative translation efficiency, functionality, or protein stability that can arise if measuring protein products or enzyme activities.

Our results show that the position of insertion determined the expression level of the additional gene and, furthermore, that the added gene was stably maintained at all positions in the genome. However, only certain positions in the genome stably maintained expression of an additional transcriptional unit. Molecular analysis of genomes of the virus population which evolved to suppress expression of an added gene revealed that the mechanism of suppression of the added gene was rapid selection of mutations which abolished termination of the upstream gene and thereby silenced initiation of the downstream monocistronic mRNA. These data furthered our understanding of transcriptional control and showed that termination of transcription is a key step at which control of downstream gene expression can effectively be regulated within the confines of a single transcriptional promoter.

These results have direct relevance to the use of NNS viruses as vectors and vaccine delivery agents, as they show that the level of expression of an added gene can be controlled by the position at which it is inserted; however, they also show that not all positions of insertion are stable for expression of a heterologous gene.

## MATERIALS AND METHODS

**Cells and viruses.** The baby hamster kidney cell line (BHK-21) was used for transfections, growth of virus, single-step growth analyses, and metabolic analyses of viral RNA and protein synthesis. Virus pools were titrated on VERO 76 cells. Virus having the wild-type gene order was derived from an infectious cDNA clone of VSV-IN as described previously (2, 33, 34). From corresponding cDNA clones we also recovered viruses with genomes that contained an additional heterologous transcriptional unit at each of the four internal gene junctions. These cDNA clones were constructed by stepwise assembly of cloned cDNA fragments that represented the individual VSV genes or the heterologous transcriptional unit, as described before (2). This method reconstructed the wild-type intergenic junctions and left the nucleotide sequence of the viral genome unaltered except for the introduction of the heterologous transcriptional unit. This was composed of the first 7 nt of the VSV-IN M gene, followed by 602 nt derived from the genome of bacteriophage  $\phi$ X174 ( $\phi$ X174 nt 1174 to 1775) flanked by *Xho*I linkers (CCCTCGAGGG), followed in turn by the last 32 nt of the VSV-IN (Orsay) G gene. We constructed four different infectious cDNA clones containing this cassette at either the N-P, P-M, M-G, or G-L intergenic junctions, and we named the recovered viruses NIP, PIM, MIG, and GIL, respectively. These viruses were recovered by transfection as described previously (2, 34). As in our previous work, the intergenic dinucleotide was made 3'..GA..5' at all four gene junctions (2).

**Single-cycle growth analysis of viral replication.** Single-step growth analysis was carried out in BHK cells infected at a multiplicity of 3 exactly as described previously (33).

**Analysis of RNA synthesis.** RNA synthesis was analyzed by direct metabolic labeling of cells infected with the indicated virus at a multiplicity of 5 (see Fig. 1) or 10 (see Fig. 3). Cells were exposed for 2.5 h to [<sup>3</sup>H]uridine (33  $\mu$ Ci/ml) in the presence of actinomycin D (10  $\mu$ g/ml). Cells were harvested, cytoplasmic extracts were prepared, and RNA was analyzed by gel electrophoresis as described previously (33). Autoradiographs of fluorogrammed gels were quantitated using a PDI densitometer 320i. Molar amounts of RNA species were quantitated based on their uridine content. The results presented are representative of at least three independent experiments. RNase H directed cleavage of RNA following annealing with specific oligonucleotides ([gene, nucleotides, polarity] N, 336 to 351, -; P, 1420 to 1437, -;  $\phi$ X174 insert, 5'GCGCCAGTATTAACAGTCG3', -; L, 7580 to 7599, +)

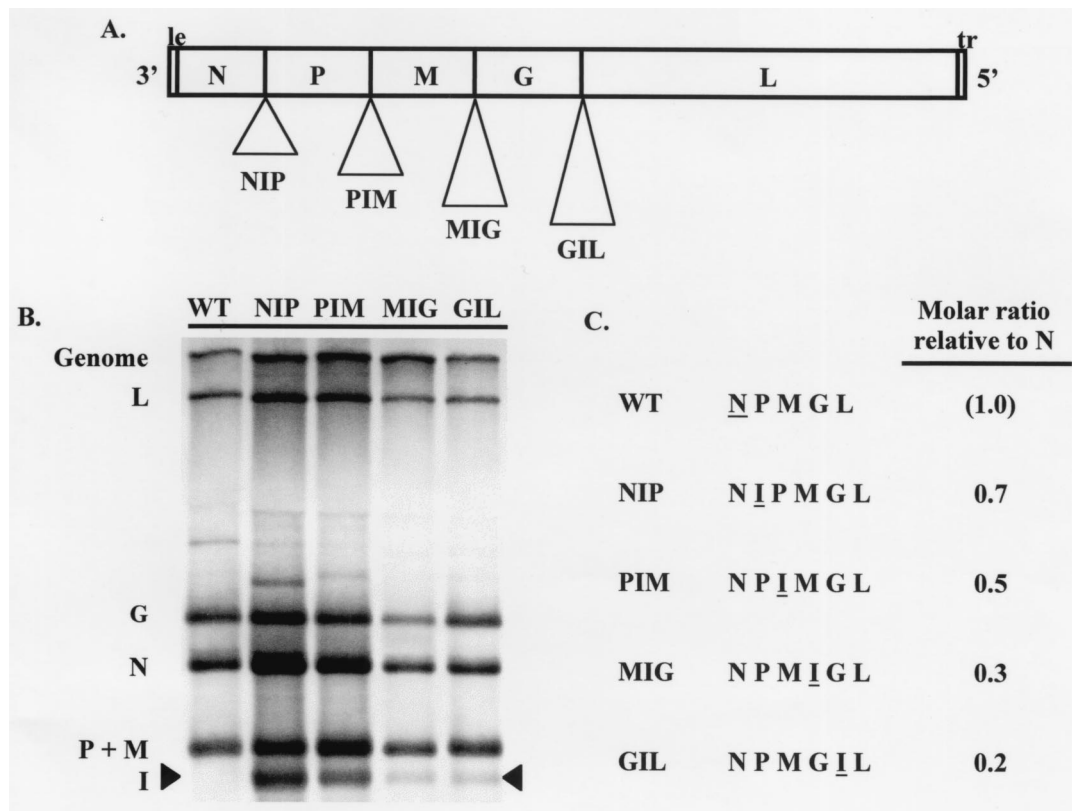


FIG. 1. Insertion and expression of an additional, identical transcription unit at one of each of the VSV gene junctions. (A) Diagrammatic representation of the position of insertion of a 661-nt noncoding transcription unit at each of the four VSV gene junctions. The recovered engineered viruses bearing the additional gene were named according to the position at which the inserted I gene was added: NIP, PIM, MIG, and GIL. (B) RNA synthesis directed by viruses NIP, PIM, MIG, and GIL and the wild-type virus recovered from the parent cDNA clone without an additional gene. VSV-specific RNA was labeled with [ $^3$ H]uridine in the presence of actinomycin D for 2 h beginning at 1.5 h postinfection and analyzed by gel electrophoresis. The positions of the genomic RNA, the five VSV mRNAs, and the mRNA from the inserted I gene are indicated to the left of the figure. The P and M mRNAs comigrated. (C) Molar amounts of the I mRNA expressed from the different positions in the genome. RNAs were quantitated by densitometric analysis of autoradiographs, and molar amounts relative to expression of the N mRNA were calculated.

was carried out as described previously, and cleavage products were recovered and analyzed by gel electrophoresis and fluorography (6).

**Analysis of viral protein synthesis.** Protein synthesis directed by each of the recovered viruses was measured in BHK cells infected at a multiplicity of 20. At 5 h postinfection cells were washed and incubated in methionine-free medium for 30 min. They were then exposed to [ $^{35}$ S]methionine (30  $\mu$ Ci/ml) for 30 min. At the end of the labeling period, cell monolayers were harvested directly into gel loading buffer and separated on a 10% low-bis polyacrylamide gel as described previously (33). Viral proteins were quantitated by phosphorimager analysis of autoradiograms of dried gels and molar amounts of proteins were determined.

**Sequence analysis of viral genomes.** The RNA from virus isolated from supernatant fluids was extracted and the central region of the genome was analyzed by reverse transcription (RT)-PCR using murine leukemia virus reverse transcriptase and oligonucleotide primers from the N gene (1027 to 1046; +) and from the M gene (2271 to 2290; -). The sequence across each of the gene junctions and the surrounding area was analyzed in both directions using the following oligonucleotide primers (gene, nucleotides, polarity): N, 1027 to 1046, +; N, 1321 to 1376, -; PhiX insert, 5'GCGCCAGTATTAACAGTCG3', -; P, 1420 to 1437, -; P, 1464 to 1488, +; P, 154 to 1562, +; M, 2271 to 2290, -. Automated sequence analysis was carried out on a Perkin-Elmer Applied Biosystems 377 sequencer.

## RESULTS

**Effect of position of insertion of an additional gene on expression levels.** We inserted an additional 661-nt transcription unit containing the conserved VSV gene start and gene end sequences at each of the four gene junctions of an infectious

cDNA of the VSV genome. This was done without introducing changes into the conserved gene junction sequences (see Materials and Methods). Infectious viruses were recovered from each of the four cDNAs by transfection of the appropriate plasmid into cells along with plasmids encoding the VSV N, P, and L proteins as previously described (34). The recovered viruses were designated according to the position at which the inserted "I" transcriptional unit (also referred to as the I gene) was placed: NIP, for the virus with the gene inserted between the N and P genes, and PIM, MIG, and GIL for the viruses having the additional gene inserted between the P and M, M and G, and G and L genes, respectively, as indicated in Fig. 1A.

The effect of insertion position on expression of the inserted gene was examined by direct metabolic labeling of RNAs with [ $^3$ H]uridine in the presence of actinomycin D in infected BHK cells. In addition to the synthesis of the viral genomic RNA and the monocistronic mRNAs transcribed from the viral L, G, N, P, and M genes, RNA from the inserted I gene was clearly visible migrating slightly faster than the viral P and M mRNAs (Fig. 1). The relative abundance of the I mRNA varied according to the position at which the gene was inserted into the genome. The most abundant RNA synthesis arose from insertion of the gene at the 3'-most site between the N and P genes,



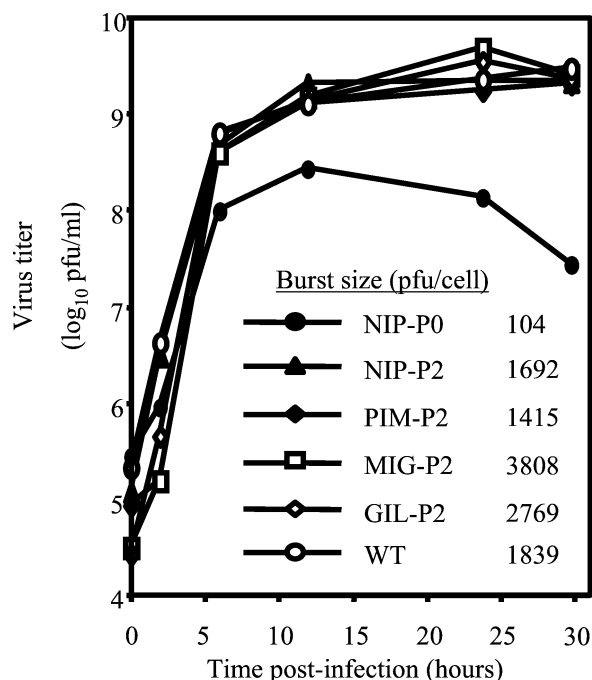


FIG. 2. Replication of viruses with an additional gene inserted at one of each of the VSV gene junctions. The viruses having an additional gene inserted were assayed for replication potential in comparison to wild-type virus by single-step growth in BHK-21 cells infected at a multiplicity of 3 PFU. All virus from P0 and P2 was analyzed. For simplicity, results are shown for NIP-P0, NIP-P2, PIM-P2, MIG-P2, and GIL-P2. The replication levels of the PIM, MIG, and GIL viruses were essentially indistinguishable at P0 and P2. Yields per cell at 24 h postinfection are shown in the inset.

and successively lower amounts arose from insertion between the subsequent gene pairs (Fig. 1). Quantitation of the molar amounts of the I mRNA relative to the 3'-proximal N mRNA showed a progressive decline in synthesis (Fig. 1) that was consistent with an attenuation of transcription across each gene junction of approximately 30% (6, 14). These data confirm that gene expression is controlled by the position of a gene relative to the 3' promoter.

#### Effect of insertion of additional genes on viral replication.

Viral replication was examined by analysis of progeny virus production in BHK-21 cells in culture. All of the viruses having an extra gene replicated to levels similar to that of the wild-type virus (recovered from the parent cDNA with no additional gene), except in the case of virus NIP (Fig. 2). The replication of virus NIP was reduced 10- to 30-fold in comparison to wild-type virus, as reflected in the reduced viral burst size.

One explanation for the reduced replication of the NIP virus was that the insertion of an extra gene to the genome would introduce an additional junction-specific transcriptional attenuation step that should decrease the level of expression of all downstream genes. Appropriate relative molar ratios of the N and P proteins are important for replication in VSV (12, 21). Insertion of an additional gene between the N and P genes should not affect N gene expression but would be expected to decrease the molar amount of P protein from the wild-type level. This was confirmed, as will be shown below.

**Effect of passage on viral replication levels.** Virus recovered as above had been subjected to one low-multiplicity passage to obtain a pool of virus suitable to carry out experiments; this was termed the original passage (P0). To investigate the stability of the inserted gene, the stability of its expression, and the phenotype of reduced replication observed in the NIP viruses compared to the other three viruses, we passaged the viruses two additional times. To do this, the four viruses were passaged once at a low multiplicity of infection (0.005). The viruses from this passage (P1) were then isolated by banding in sucrose velocity gradients to remove any possible defective interfering particles, and a second passage (P2) was carried out at a low multiplicity as above. Viruses from the second passage were characterized directly for replication in single-step growth analyses. In contrast to the NIP-P0 virus, the NIP-P2 virus showed no deficiency in replication; it replicated to the same levels as wild-type virus (Fig. 2). The replication of the PIM, MIG, and GIL viruses was essentially the same as that of the wild-type virus both before and after passaging (Fig. 2).

**RNA synthetic phenotype of passaged viruses.** The patterns of RNAs synthesized by the passaged viruses in BHK cells were analyzed and found to show no differences in the case of the twice-passaged PIM-P2, MIG-P2, and GIL-P2 viruses (Fig. 3). However, the passaged NIP-P2 virus, which formerly had synthesized the highest levels of the I mRNA, now synthesized barely detectable levels of I mRNA and instead made an abundant RNA that migrated between the L and G mRNAs (Fig. 3A). The new, large RNA was characterized by hybridization with specific oligonucleotide probes followed by digestion with RNase H and gel analysis and found to be a bicistronic RNA consisting of the N and I RNAs. The large RNA was specifically cleaved by RNase H following annealing with a negative-sense oligonucleotide specific to N mRNA or to the I mRNA from the inserted gene (PhiX sequence), but not after annealing to negative-sense oligonucleotides specific for the P gene (data not shown) or positive-sense oligonucleotides specific for the L sequence (Fig. 3B). The size of the large RNA and the fact that it annealed to oligonucleotide probes specific for both the N and I sequences, but not to probes for P sequences, indicated that it was a readthrough RNA of the N and I genes. The finding that it did not anneal to the positive-sense L-specific probe ruled out the possibility that it might have been a conventional 3' copyback defective interfering RNA which might have arisen during passage (10). These findings were confirmed by RT-PCR and sequence analysis as described below.

The rapidity of selection of revertants was unexpected; therefore, we returned to the original NIP-P0 virus and repeated the passages independently, exactly as it was done the first time. The virus from the second independent passage series, NIP-P2.2, again replicated as well as the wild type. The RNA synthesis phenotype also displayed the same pattern as that observed above for NIP-P2 in that little or no monocistronic I mRNA was made and, instead, the N-I bicistronic RNA was present in abundance (data not shown).

The PIM, MIG, and GIL viruses were subsequently passaged 10 times at low multiplicity as described above. Their replication levels and levels of expression of the I mRNA remained stable and indistinguishable from that observed at P0.

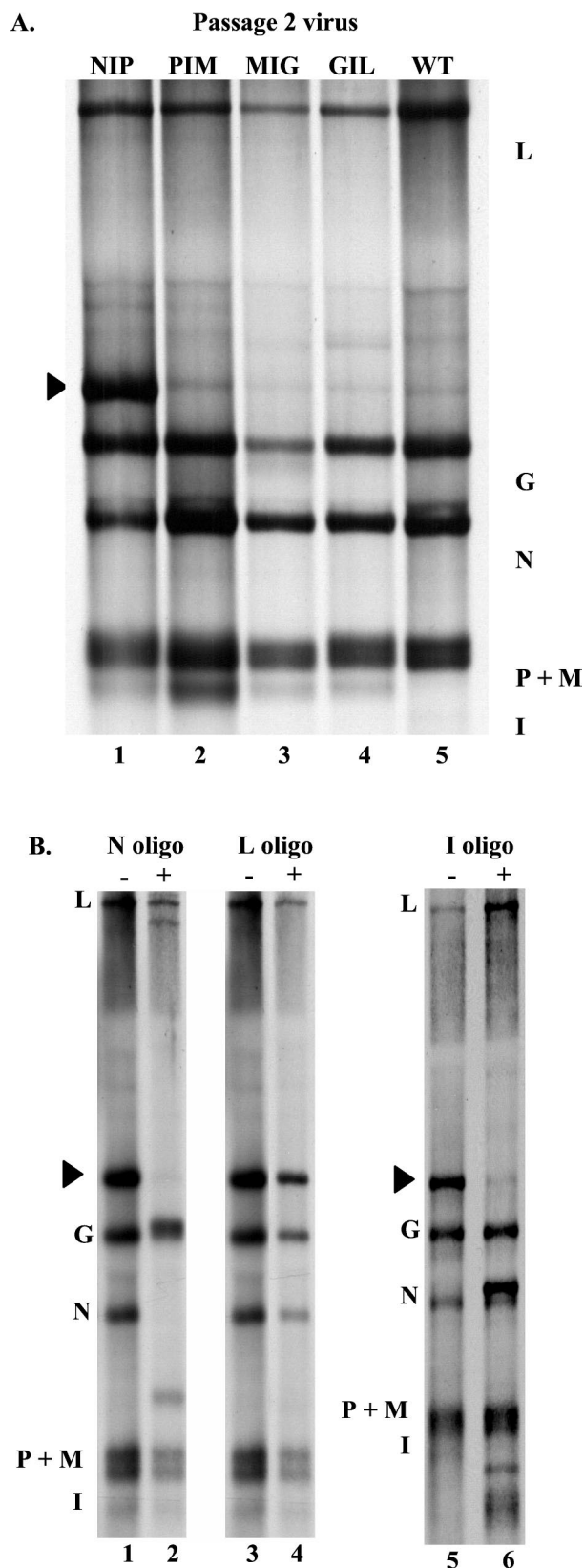


FIG. 3. RNA synthesis directed by viruses having an additional gene inserted at each of the viral gene junctions following two additional passages. (A) Comparison of RNA synthesis directed by viruses NIP-P2, PIM-P2, MIG-P2, GIL-P2, and wild type. RNA was labeled

**Sequence of revertant viral genomes.** We used RT-PCR to determine the sequence of the genomes of the revertant viruses. The segment of the genome that extended from the end of the N gene through the start of the M gene and included the N-I, I-P, and P-M gene junctions was amplified (Fig. 4). The sequence of this product was analyzed in two ways. First, the bulk PCR product was isolated by gel purification and the sequence across each of the gene junctions was analyzed directly to examine the sequence of the virus population. The sequence of the bulk PCR product across the I-P gene junction was found to be wild type (Fig. 4), as was that of the P-M junction (data not shown). However, the sequence at the N-I junction, which in the original virus was clearly wild type with an uninterrupted U7 tract, was distinguishably heterogeneous in the twice-passaged viruses (Fig. 4C).

To determine the sequences at the gene junctions of individual virus species in the heterogeneous population, two independent RT-PCRs were carried out on the genomes of the virus populations derived from each of the two independent passages described above: the NIP-P2 and NIP-P2.2 viruses. The RT-PCR products were isolated and cloned, and 35 independent DNA clones (15 from NIP-P2 and 20 from NIP-P2.2) were isolated and sequenced in both directions across the N-I, the I-P, and the P-M junctions. The summary of the sequence analysis is shown in Table 1. The sequence across the P-M junction was wild type in 34 of 35 clones. The one P/M junction clone that was not wild type had the P-M U7 tract extended to U8. The sequence across the I-P junction was wild type in 33 of 35 clones. In the two clones that were not wild type, the U tract at the I-P junction was U8 in one and U2CU4 in the other. In contrast, the sequence of the N-I junction was wild type in only 4 of 35 clones; in the other 31 clones there were several types of mutations (Table 1). The most frequent change, in 14 of 35, was interruption of the U7 tract by a C residue (U2CU4 or U4CU2). The second most frequent change (11 of 35) was the shortening of the U tract to U6. Previous work has shown that the U7 tract is required for transcriptional termination. Interruption of the U tract by a heterologous nucleotide or shortening of the U tract by even a single nucleotide abolishes slippage and termination (5). The next most abundant change involved alterations of the AUAC tetranucleotide that precedes the U7 tract (Table 1), where any alterations decrease termination (5).

These data show that mutations that are known to inhibit termination of transcription were selected for at the N-I gene junction in virus that was passaged twice. The two downstream junctions examined in each of the same DNA clones, the I-P and P-M gene junctions, accumulated a mutation that would eliminate termination in only 1 of 70 junction sequences examined. Since the sequences at the VSV gene junctions are

with [ $^3$ H]uridine in the presence of actinomycin D (as described in the legend to Fig. 1) and analyzed by gel electrophoresis. (B) Analysis of the identity of the N-I RNA (indicated by arrow) by annealing of negative-sense oligonucleotides specific for the N or I mRNA sequence, or a positive-sense oligonucleotide specific for the L gene, to RNAs from an infection with NIP-P2 labeled as in panel A, followed by incubation in the presence (+) or absence (-) of RNase H and gel electrophoresis.

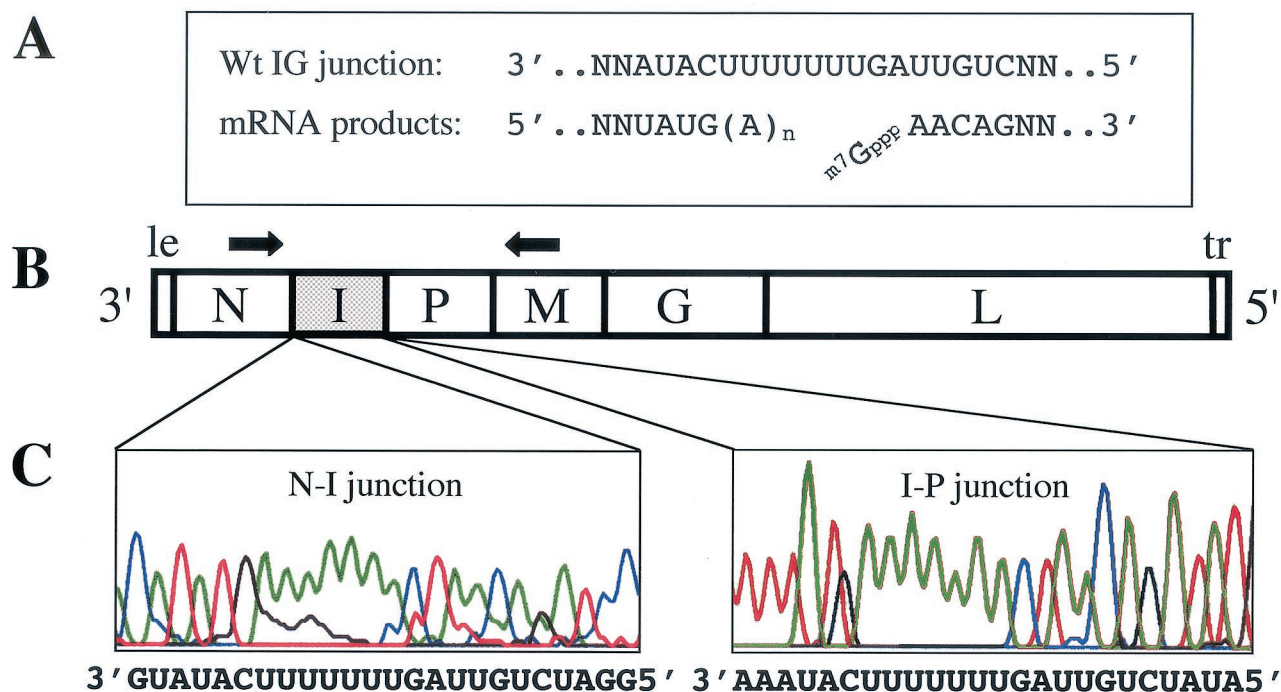


FIG. 4. Analysis of sequence across the gene junctions flanking the inserted gene. (A) The VSV conserved gene junction sequence and its relation to progeny mRNA. (B) Diagram of the position of primers used for RT-PCR analysis of the genomes of the initial and P2 viruses. (C) Sequence analysis of bulk PCR products of NIP-P2 genomic RNA.

highly conserved, these data provide a strong internal control showing that the changes observed at the N-I junction did not arise from slippage or errors on those sequences by the enzymes used in RT or PCR.

In keeping with the obligatorily sequential nature of transcription of the NNS viruses, the failure to terminate transcription of an upstream gene results in polymerase readthrough at that junction and prevents polymerase initiation at the gene immediately downstream. Concomitant with these sequence findings, we observed that synthesis of the monocistronic I mRNA was suppressed within two passages and that a bicistronic N-I RNA was instead synthesized in the passaged viruses (Fig. 3A). These results are consistent with the N gene

end mutations eliminating termination and, as a consequence, the polymerase reading through the gene junction to generate a polycistronic RNA. This should have the effect of eliminating the extra attenuation step introduced by transcription of the added gene. To determine if this was so, we compared synthesis of proteins from the original virus isolates with that of the twice-passaged virus.

**Effect of inserted genes on viral protein production.** The effect of an additional transcriptional unit inserted at each of the VSV intergenic junctions on expression of the viral proteins was examined in BHK cells by metabolic labeling with [<sup>35</sup>S]methionine. Viral proteins were analyzed by polyacrylamide gel electrophoresis and quantitated by phosphorimaging, and the relative molar amounts of each protein were expressed as the percent of the wild-type virus molar ratios. The molar amount of the proteins expressed from genes located downstream of the inserted transcription unit in virus NIP-P0 fell by approximately 35% in comparison to wild-type levels (Table 2). This is as would be predicted for the insertion of an extra transcriptional unit at the indicated position.

In contrast, the molar amounts of all proteins downstream of the I gene insertion in NIP-P2 were restored to almost wild-type levels. These data are consistent with the loss of expression of the added gene as a monocistronic mRNA and elimination of the attendant attenuation that occurs as the polymerase terminates transcription of the upstream gene, crosses the intergenic junction, and initiates at the downstream gene. The molar amounts of proteins expressed from the PIM, MIG, and GIL viruses were found to be unchanged after two passages, with the molar amounts of proteins downstream of

TABLE 1. Intergenic junction sequences of 35 cDNA clones of NIP-P2 virus<sup>a</sup>

Junction sequence	No. of clones with indicated change at IG junction			Change
	N-I	I-P	P-M	
3' . . AUACUUUUUUUGAUUGUC . . 5'	4	33	34	WT
<u>U</u>				
. . AUACUUUUUUUGAUUGUC . .	3	1	1	U <sub>8</sub>
▲				
. . AUAC_UUUUUUGAUUGUC . .	11	0	0	U <sub>6</sub>
. . AUACUU_CUUUUUGAUUGUC . .	13	1	0	U <sub>2</sub> CU <sub>4</sub>
. . AUACUUUUU_CUUUGAUUGUC . .	1	0	0	U <sub>4</sub> CU <sub>2</sub>
. . AU_UUUUUUGAUUGUC . .	1	0	0	AU..
. . A_CAUUUUUUGAUUGUC . .	2	0	0	ACAC

<sup>a</sup> IG, intergenic; WT, wild type.



TABLE 2. Molar ratios of VSV proteins

Virus	Molar ratio of VSV protein (% of wt value) <sup>a</sup>				
	N	P	M	G	L
NIP-P0	100	64	62	43	61
NIP-P2	100	94	90	80	86
PIM-P2	100	100+	62	48	61
MIG-P2	100	100+	100+	61	61
GIL-P2	100	100+	100+	100+	57

<sup>a</sup> Molar ratios of individual VSV proteins were calculated relative to the N protein and expressed as the percentage of the molar ratio of the proteins of the wild-type virus with no added gene. wt, wild type.

the inserted gene still reduced by approximately 30% compared to wild-type levels, consistent with stable expression of the added gene, I, from these viruses.

These data show that the elimination of termination at the end of the N gene resulted in readthrough of polymerase at the N-I junction and suppressed transcription of the monocistronic I mRNA. This eliminated the attenuation step at that N-I gene junction and restored expression levels of all proteins downstream of the inserted transcriptional unit to near wild-type levels.

## DISCUSSION

Foreign genes have been expressed from many of the NNS RNA viruses from different positions in the genome (7–9, 17, 24, 32). In all cases, the added genes have been stably maintained in the genome, but the effect of the added genes on viral replication has varied. In some cases, such as the addition of green fluorescent protein to rinderpest virus between the P and M genes, replication was slowed but eventually titers were comparable to that of wild-type virus (32). In other cases, such as the addition of firefly luciferase to the 3'-most locus of Sendai virus, replication was reduced by two- to fourfold and plaques were smaller, whereas in the case of Newcastle disease virus, addition of the chloramphenicol acetyltransferase (CAT) gene between the last two genes, HN and L, reduced replication by 100-fold (8, 17). No clear correlation emerged from these reports as to the effect of position of insertion of an extra gene and the potential effects on viral replication potential. An additional consideration in evaluating these reports is that in some cases the wild-type sequence was maintained at and around the added gene junctions, whereas in others, changes were introduced within sequences that may have regulatory roles. The stability and expression of an identical gene from all available positions in an NNS genome has not previously been examined.

In the work reported here, we inserted an identical transcriptional unit at each of the gene junctions of VSV in an identical manner, such that no alterations were introduced in the gene junction sequences. We then examined the stability of the added gene in the genome, the stability of expression of the transcriptional unit, and the effects of having a gene added at the various positions on viral replication. Critical to the experiment was the fact that the added transcriptional unit was designed to express no protein product that might confer selective advantage or disadvantage on the virus. The reported results show that position of insertion determined the expres-

sion level of the added gene and that the added gene was stably maintained at all positions in the genome. Further, only insertion of an additional gene between the N and P junction resulted in significant reduction of viral replication potential; insertion at all other junctions had no measurable effect on replication in single-step growth analyses.

Upon passage, the NIP virus recovered replication to wild-type levels within two passages. Concomitant with this, the twice-passaged NIP-P2 virus lost the expression of the inserted gene as a monocistronic mRNA. Instead, it was expressed as the downstream half of a bicistronic, readthrough RNA. Readthrough RNAs are produced when the viral polymerase fails to terminate at an upstream gene junction. When this happens the transcription of the downstream gene as a discrete monocistronic mRNA is inhibited, since downstream initiation requires upstream termination (6). This in turn leads to loss of an attenuation step and increased expression of all downstream genes.

Analysis of sequence across the viral gene junctions that preceded and followed the inserted gene as well as the next junction downstream (the N-I, I-P, and P-M gene junctions) showed that while the sequence at the I gene end remained wild type, the sequence at the N gene end, preceding the inserted gene, rapidly accumulated mutations. Several types of mutations were observed. The most frequently observed mutations either interrupted the U7 tract at the gene end or shortened it by 1 nt. A third sort of mutation that arose was alteration of the gene end tetranucleotide, AUAC. In previous work using subgenomic replicons, mutations which either shortened the U tract or interrupted it were both shown to abolish transcriptional termination, and alteration of any of the four nucleotides of the tetranucleotide reduced the efficiency of termination (5). The accumulation of mutations at the N gene end of the NIP-P2 virus correlated with loss of termination at the N gene end and with loss of expression of the inserted gene as a monocistronic mRNA. These data provide striking forward genetic confirmation by natural selection in virus populations of the previous mutagenic analyses of the role of individual nucleotides in termination (5, 6, 13, 29). The N-I intergenic dinucleotide and the I gene start sequence were not found to accumulate changes. In addition, the sequence of the I-P and P-M gene junctions, the next junctions downstream of the insert, did not accumulate mutations. These findings provide internal controls that the observed mutations at the N gene end were not artifacts of the RT-PCR involved in analysis of the recovered genomes. They further suggest that the intergenic junctions are not intrinsically error prone.

The reason for the instability of expression of the inserted gene from between the N and P genes may be due to a replicative disadvantage caused by the disruption of the molar ratio of N protein to P protein and the effect of this ratio on efficient genome replication (12, 21). Genome replication requires a constant supply of soluble N protein to encapsidate the nascent replicating RNAs (20). Previous work has shown that alteration in the molar ratio of N to P proteins affects the levels of functional N protein, since the P protein serves as a chaperone to prevent N from self-aggregating and becoming nonfunctional in replication (12). Alternatively, the alteration of the relative molar ratios of other proteins which need to interact may also be critical. Previous work in which the order of the

three central genes of VSV was rearranged to all possible permutations did not result in a replicative disadvantage in every instance where the N and P ratios were disturbed. In particular, a virus having the gene order 3'-N, G, M, P, L-5' replicated as well as wild-type virus in cell culture, indicating that there are many complex interactions yet to be understood.

The data presented here show that the insertion of an additional gene at any of the gene junctions resulted in down-regulation of the expression of all the viral proteins downstream of the insertion (Table 2). This was due to the extra attenuation step introduced by transcription of the added gene. While the replication of all the viruses having the insert at the P-M, M-G, or G-L gene junction was comparable to that of wild-type virus with no insert, the replication of the NIP-P0 virus, which had the molar amount of P protein reduced by over 30%, was reduced by an order of magnitude, indicating that it was replicating at a disadvantage. The NIP-P2 virus, which had wild-type replication levels, had approximately wild-type relative amounts of all viral proteins and in particular the N/P protein molar ratio. Thus, the mutations at the N gene end which inhibited transcription termination, and as a consequence produced a readthrough RNA, eliminated the extra attenuation step and restored expression of downstream genes to near wild-type levels.

In contrast, the PIM, MIG, and GIL viruses maintained expression of the inserted gene for up to at least 10 passages. This finding is similar to previous reports which showed that expression of the CAT gene or CD4 from between the G and L genes of VSV remained stable for up to at least 15 or 26 passages, respectively (22, 24). A VSV expressing the measles virus fusion (F) protein gene from between G and L, however, lost expression of the F gene after three passages (22). It was unclear why, but the authors speculated that the F protein was toxic to VSV growth. The loss of F expression correlated with accumulation of mutations at the prior gene end and polymerase readthrough as described here. This led the authors to suggest that the transcription stop sequence might be a hotspot for mutation. However, our sequence analyses of three such signals showed that the gene end mutations accumulated only where they silenced expression of the inserted gene. Thus, our data suggest the gene end is not a mutational hotspot but, instead, that in cases where transcription of an added gene provides a disadvantage to replication of a virus, the most efficient way to silence expression of the additional gene will rapidly be selected. In the NNS viruses where transcription is obligatorily sequential, efficient termination of each gene is key for downstream gene expression. VSV optimal termination requires the wild-type sequence at the 12 nt comprising the gene end and the first nucleotide of the intergenic junction, and alteration of any of these nucleotides can reduce or abolish efficient termination.

It is known that the RNA-dependent RNA polymerase of VSV has a high polymerase error rate of approximately  $10^{-3}$  to  $10^{-4}$  misincorporations per nt per round of copying (25–27). This error rate coupled with a lack of a proofreading mechanism results in these viruses existing in the form of a quasispecies (11). The recovery of a virus from a cDNA clone represents a stringent cloning event, but no more than biological cloning by the picking of a single plaque which is the outgrowth of progeny from a single virus particle. By the time the virus

has replicated to titers of  $10^8$  PFU per ml, it can be expected to exist as a quasispecies. Engineered viruses recovered from cDNA have been amplified many millionfold by the time they have been passaged two to four times. Thus, they are most likely quasispecies in which a variety of mutations exist due to polymerase error events. Mutations which provide an advantage to the virus can readily be selected (11, 18, 19).

The wild-type sequence was the preponderant sequence at the N-I junction of the virus recovered in the first large passage, NIP-P0, as shown by bulk sequence analysis of the PCR product and as evidenced by the fact that a monocistronic I mRNA was synthesized abundantly from this virus population. However, after two passages, the expression of the monocistronic mRNA was almost completely suppressed, and results of the analysis of the sequences of individual clones derived from independent RT-PCR amplifications of the genome was consistent with a viral quasispecies in which the majority of the viruses in the population represented viruses that had a selective advantage for replication due to mutations which eliminated termination at the N gene end and, thus, expression of the monocistronic I mRNA. These changes correlated with elimination of the junction-specific attenuation event and restoration of protein molar ratios and replication of the virus to wild-type levels. Taken together, these data suggest that within the confines of a single transcriptional promoter and obligatorily sequential transcription, the selection of sequences which alter the termination efficiency of an upstream gene is an effective means whereby expression of a downstream gene may be silenced or reduced in the NNS RNA viruses.

The findings reported here have important implications for the use of NNS viruses as vectors and vaccines. On the positive side they show that the level of expression of a heterologous gene can be controlled by where the gene is positioned relative to the 3' promoter. However, they also caution that not all positions on the genome will yield stable expression of a foreign gene.

Finally, although the viruses having insertions at the P-M, M-G, and G-L junctions replicated as well as wild-type virus with no insertions in a single-step growth curve, it will be of interest to determine the fitness of these viruses in comparison to wild-type virus. It will also be important to determine whether the size of an inserted transcriptional unit, as well as the addition of another gene junction with its attendant attenuation step, constitutes a selective disadvantage to the virus.

#### ACKNOWLEDGMENTS

We thank Xiaoling Tang for excellent technical assistance and our colleagues for constructive comments.

This work was supported by Public Health Service grants from the NIAID (R37 AI 12464 to G.W.W. and AI 18270 to L.A.B.).

#### REFERENCES

1. Abraham, G., and A. K. Banerjee. 1976. Sequential transcription of the genes of vesicular stomatitis virus. *Proc. Natl. Acad. Sci. USA* **73**:1504–1508.
2. Ball, L. A., C. R. Pringle, B. Flanagan, V. P. Perepelitsa, and G. W. Wertz. 1999. Phenotypic consequences of rearranging the P, M, and G genes of vesicular stomatitis virus. *J. Virol.* **73**:4705–4712.
3. Ball, L. A., and C. N. White. 1976. Order of transcription of genes of vesicular stomatitis virus. *Proc. Natl. Acad. Sci. USA* **73**:442–446.
4. Barr, J. N., and G. W. Wertz. 2001. Polymerase slippage at vesicular stomatitis virus gene junctions to generate poly(A) is regulated by the upstream 3'-AUAC-5' tetranucleotide: implications for the mechanism of transcription termination. *J. Virol.* **75**:6901–6913.



5. Barr, J. N., S. P. J. Whelan, and G. W. Wertz. 1997. The *cis*-acting signals involved in termination of VSV mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. *J. Virol.* **71**:8718–8725.
6. Barr, J. N., S. P. J. Whelan, and G. W. Wertz. 1997. Role of the intergenic dinucleotide in vesicular stomatitis virus RNA transcription. *J. Virol.* **71**:1794–1801.
7. Bukreyev, A., E. Camargo, and P. L. Collins. 1996. Recovery of infectious respiratory syncytial virus expressing an additional, foreign gene. *J. Virol.* **70**:6634–6641.
8. Hasan, M. K., A. Kato, T. Shioda, Y. Sakai, D. Yu, and Y. Nagai. 1997. Creation of an infectious recombinant Sendai virus expressing the firefly luciferase gene from the 3' proximal first locus. *J. Gen. Virol.* **78**:2813–2820.
9. He, B., R. G. Paterson, C. D. Ward, and R. A. Lamb. 1997. Recovery of infectious SV5 from cloned DNA and expression of a foreign gene. *Virology* **237**:249–260.
- 9a. Hinzman, E. E., J. N. Barr, and G. W. Wertz. 2002. Identification of an upstream sequence element required for vesicular stomatitis virus mRNA transcription. *J. Virol.* **76**:7632–7641.
10. Holland, J. J. 1987. Defective interfering rhabdoviruses, p. 297–360. *In* R. R. Wagner (ed.), *The rhabdoviruses*. Plenum Press, New York, N.Y.
11. Holland, J. J., J. Torre, and D. Steinhauer. 1992. RNA virus populations as quasispecies. *Curr. Top. Microbiol. Immunol.* **176**:1–20.
12. Howard, M., and G. W. Wertz. 1989. Vesicular stomatitis virus RNA replication: a role for the NS protein. *J. Gen. Virol.* **70**:2683–2694.
13. Hwang, L. N., N. Englund, and A. K. Pattnaik. 1998. Polyadenylation of vesicular stomatitis virus mRNA dictates efficient transcription termination at the intercistronic gene junctions. *J. Virol.* **72**:1805–1813.
14. Iverson, L. E., and J. K. Rose. 1981. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell* **23**:477–484.
15. Keene, J. D., M. Schubert, and R. A. Lazzarini. 1979. Terminal sequences of vesicular stomatitis virus RNA are both complementary and conserved. *J. Virol.* **32**:167–174.
16. Keene, J. D., B. J. Thornton, and S. U. Emerson. 1981. Sequence-specific contacts between the RNA polymerase of vesicular stomatitis virus and the leader RNA gene. *Proc. Natl. Acad. Sci. USA* **78**:6191–6195.
17. Krishnamurthy, S., Z. Huang, and S. K. Samal. 2000. Recovery of a virulent strain of Newcastle disease virus from cloned cDNA: expression of a foreign gene results in growth retardation and attenuation. *Virology* **278**:168–182.
18. Novella, I., D. Clarke, J. Quer, E. Duarte, C. Lee, S. Weaver, S. Elena, A. Moya, E. Domingo, and J. Holland. 1995. Extreme fitness differences in mammalian and insect hosts after continuous replication of vesicular stomatitis virus in sandfly cells. *J. Virol.* **69**:6805–6809.
19. Novella, I. M., C. H. Hershey, C. Escarmis, E. Domingo, and J. J. Holland. 1999. Lack of evolutionary stasis of an arbovirus during alternating replication in mammalian and insect cells. *J. Mol. Biol.* **287**:459–465.
20. Patton, J. T., N. L. Davis, and G. W. Wertz. 1984. N protein alone satisfies the requirement for protein synthesis during RNA replication of vesicular stomatitis virus. *J. Virol.* **49**:303–309.
21. Peluso, R. 1988. Kinetic, quantitative and functional analysis of multiple forms of the vesicular stomatitis virus nucleocapsid protein in infected cells. *J. Virol.* **62**:2799–2807.
22. Quinones-Kochs, M. I., M. J. Schnell, L. Buonocore, and J. K. Rose. 2001. Mechanisms of loss of foreign gene expression in recombinant vesicular stomatitis viruses. *Virology* **287**:427–435.
23. Rose, J., and M. Schubert. 1987. Rhabdovirus genomes and their products, p. 127–166. *In* R. R. Wagner (ed.), *The rhabdoviruses*. Plenum Press, New York, N.Y.
24. Schnell, M. J., L. Buonocore, M. A. Whitt, and J. K. Rose. 1996. The minimal conserved transcription stop/start signal promotes stable expression of a foreign gene in vesicular stomatitis virus. *J. Virol.* **70**:2318–2323.
25. Steinhauer, D. A., J. C. De La Torre, and J. J. Holland. 1989. High nucleotide substitution error frequencies in clonal pools of vesicular stomatitis virus. *J. Virol.* **63**:2063–2071.
26. Steinhauer, D. A., and J. J. Holland. 1987. Rapid evolution of RNA viruses: RNA versus DNA genomes. *Annu. Rev. Microbiol.* **41**:409–433.
27. Steinhauer, D. A., and J. J. Holland. 1986. Direct method for quantitation of extreme polymerase error frequencies at selected single base sites in viral RNA. *J. Virol.* **57**:219–228.
28. Stillman, E. A., and M. A. Whitt. 1998. The length and sequence composition of vesicular stomatitis virus intergenic regions affect mRNA levels and the site of transcript initiation. *J. Virol.* **72**:5565–5572.
29. Stillman, E. A., and M. A. Whitt. 1997. Mutational analyses of the intergenic dinucleotide and the transcriptional start sequence of vesicular stomatitis virus (VSV) define sequences required for efficient termination and initiation of VSV transcripts. *J. Virol.* **71**:2127–2137.
30. Stillman, E. A., and M. A. Whitt. 1999. Transcript initiation and 5' end modifications are separable events during vesicular stomatitis virus transcription. *J. Virol.* **73**:7199–7209.
31. Villareal, L., M. Breindl, and J. Holland. 1976. Determination of the molar ratios of vesicular stomatitis virus induced RNA species in BHK-21 cells. *Biochemistry* **15**:1663–1667.
32. Walsh, E. P., M. D. Baron, J. Anderson, and T. Barrett. 2000. Development of a genetically marked recombinant rinderpest vaccine expressing green fluorescent protein. *J. Gen. Virol.* **81**:709–718.
33. Wertz, G. W., V. P. Perepelitsa, and L. A. Ball. 1998. Gene rearrangement attenuates expression and lethality of a nonsegmented negative strand RNA virus. *Proc. Natl. Acad. Sci. USA* **95**:3501–3506.
34. Whelan, S. P. J., L. A. Ball, J. N. Barr, and G. W. Wertz. 1995. Efficient recovery of infectious vesicular stomatitis virus entirely from cDNA clones. *Proc. Natl. Acad. Sci. USA* **92**:8388–8392.
35. Whelan, S. P. J., and G. W. Wertz. 1999. Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J. Virol.* **73**:297–306.