International Journal of Genetics 7(2): 31-37, 2017
ISSN 2222-1301
© IDOSI Publications, 2017
DOI: 10.5829/idosi.ijg.2017.31.37

Current Techniques and Applications of Reverse Genetics: An Overview

Kibruyesfa Bayou

Wollega University, College of Medical and Health Science,
School of Veterinary Medicine, P.O. Box: 395, Nekemte, Ethiopia

Abstract: Reverse genetics is a gene-driven approach that deduces, from directed mutations, the full range of phenotypes controlled by each gene. This investigative process proceeds in the opposite direction of so-called forward genetic screens of classical genetics. Genome sequencing projects have identified large numbers of organisms' genes for which no role has yet been defined. To address this problem a number of techniques have been developed, to enable researchers to identify animals, plants and microorganisms with mutations in genes of known sequence. The reverse genetics techniques or approaches include RNA interference (RNAi), insertions and deletions into the genomes, interference using transgenes and point mutations by targeting induced local lesions in genomes (TILLING), each with its own strengths and weaknesses. Reverse genetics applications include study of RNA viruses, by which it is possible to reveal viral induced expression of a foreign gene and characterize viral accessory proteins using recombinant knockout viruses. Another potential applications (using RNAi) involve in the treatment of high cholesterol induced diseases and cancer, for live-attenuated virus vaccine development and also they are used to study about plants. Finally, to add more on the current genetic information of different organisms advanced studies should be done on the possible alternative techniques and applications of reverse genetics to solve the already existing biomedical related problems of the globe.

Key words: Genome · Mutagenesis · Reverse Genetics · RNA Interference

INTRODUCTION

Reverse genetics is a discipline that involves the use of genetic manipulation and modification to study an organism's altered phenotype. Genetics relies on mutants. In conventional or "forward" genetics, naturally occurring or artificially induced mutants are first identified on the basis of their specific appearances or properties, their phenotypes. Forward genetics, which essentially goes from the phenotype to the gene, has been an active field of research for more than one-hundred years and huge collections of spontaneous or induced mutants in several species have accumulated throughout the 20th century. Even now, trying to identify the genes whose mutations are causal in phenotypes such as complex human diseases remains a challenging goal [1].

The development of a reverse genetics system for a range of different RNA and DNA viruses has revolutionized the field of virology by making it possible to introduce designed mutations, insertions and deletions into the viral genome of live viruses. It has by now been used in a range of applications that include the attenuation of viruses, the modification of host specificity and the generation of replication-deficient viruses. These strategies have also been applied to the development of new vaccine strategies and are widely used in the characterization of the structure and function of individual viral genes and coding sequences [2].

In reverse genetics, the starting point is the gene. A "Reverse geneticist" aims to specifically modify a gene, or its expression and characterize the phenotypic consequences of this modification. Since the late 1990s, tens of eukaryotic genomes have been sequenced. Genome-wide sequencing revealed a large number of genes whose functions are unknown and cannot be predicted [3].

The number of sequenced genes whose function remains unknown has soared in recent years with the advent of affordable genome sequencing. Even in those species whose entire genome has not been sequenced,

Corresponding Author: Kibruyesfa Bayou, Wollega University, College of Medical and Health Science, School of Veterinary Medicine, P.O. Box: 395, Nekemte, Ethiopia. E-mail: kibruyesfab@gmail.com.

there are genes whose sequence is known but whose function is not yet well understood. Bioinformatics has been invaluable in investigating the function of these genetic components and powerful *in silico* techniques have been developed for the analysis of genome sequence information, but the elucidation of gene function must always be verified *in vivo* using genetic analysis. Reverse genetics is a powerful tool in the post-genomic era that establishes a direct link between the biochemical function of a gene product and its role *in vivo* [4].

Therefore, the objective of this seminar paper is to overview the current major techniques and applications of reverse genetics.

Major Techniques in Reverse Genetics

Directed Deletions and Point Mutations: Site-directed mutagenesis is a sophisticated technique that can either change regulatory regions in the promoter of a gene or make subtle codon changes in the open reading frame to identify important amino residues for protein function. Alternatively, the technique can be used to create null alleles so that the gene is not functional. For example, deletion of a gene by gene targeting (Gene Knockout) can be done in some organisms, such as yeast, mice and moss. Unique among plants, in *Physcomitrella patens*, gene knockout via homologous recombination to create knockout moss is nearly as efficient as in yeast. In the case of the yeast model system, directed deletions have been created in every non-essential gene in the yeast genome. In the case of the plant model system, huge mutant libraries have been created based on gene disruption constructs. In gene knock-in, the endogenous exon is replaced by an altered sequence of interest [5].

In this method, a single-stranded oligonucleotide is produced that differs from the target sequence by one or a few bases. Because they differ in only a few bases, the target DNA and the oligonucleotide will pair. When successfully paired with the target DNA, the oligonucleotide can act as a primer to initiate DNA synthesis, which produces a double-stranded molecule with a mismatch in the primer region. For instance, when this DNA is transferred to bacterial cells, the mismatched bases will be repaired by bacterial enzymes. About half of the time the normal bases will be changed into mutant bases and about half of the time the mutant bases will be changed into normal bases. The bacteria are then screened for the presence of the mutant gene [6].

Insertional Mutagenesis: Transformation DNA (T-DNA) or transposon insertion has been exploited to create disruptions in target genes of interest, introduce new genes, or activate endogenous genes, for example in the plant genome. For insertional knockouts this technology has been used in both monocotyledonous and dicotyledonous plants and remains one of the most effective ways of performing reverse genetics in many species. A population of plants each having an insertion(s) at a unique site in the genome is generated either by transformation (T-DNA) or transposon activation. Such populations with genome-wide insertions have been generated for several plant species as a reverse genetic service. Plants carrying an insertion in a gene of interest can be identified by screening the population with PCR using one gene-specific primer and one insertion based primer [7].

Gene disruption studies using this technique typically result in a total loss of function and the insertion can be easily followed, using PCR, in a population of plants where it is segregating. Some of the disadvantages include the facts that phenotypes may not be obvious if the gene function is redundant and insertions in essential genes will typically result in lethality making these types of genes difficult to examine using this technique. In addition, with T-DNA insertion libraries, any given plant will carry at most a few insertions; therefore, very large populations (Hundreds of thousands of plants) are needed to achieve genome saturation. This problem is alleviated for species in which endogenous transposons can be activated, but for those that don't have this advantage, generating and screening very large populations for T-DNA insertions is difficult [8].

Chemical Mutagenesis and Tilling: Chemical mutagenesis using ethylmethane sulphonate (EMS) or ethyl nitrosourea (ENU) induces point mutations in DNA in all species in which it has been tested. Targeting induced local lesions in genomes (TILLING) is a method that combines a standard and efficient technique of mutagenesis with a chemical mutagen such as EMS with a sensitive DNA-screening technique that identifies point mutations in a target gene. Generation of mutant populations using these or similar mutagens was an established practice long before the advent of gene sequencing and reverse genetics. Mutations induced using these mutagens are distributed in the genome randomly and because point mutations are less damaging

than large rearrangements, a high degree of saturation can be achieved in a mutant population, more easily enabling the examination of gene function on a genomic level [5].

In addition, while most reverse genetics techniques provide only loss-of-function alleles, chemical mutagenesis can result in either loss of function or gain of function mutations causing null, hypomorphic, neomorphic or hypermorphic phenotypes. While hypomorphic alleles may be most useful for determining wild-type gene function, neomorphic or hypermorphic alleles are more likely to be dominant and so are more likely to cause an observable phenotype in cases of either redundant or essential genes. The frequency of missense alleles is on average three times higher than that of non-sense alleles, but it is difficult to predict how many of these mutations will have an effect on gene function since many of them may not alter the gene product (s) significantly. Examples of dominant point mutations that do have an effect on gene function have, however, been well documented in several cases, for example, the ethylene response pathway, leaf polarity determination and host-pathogen defense [9].

Gene Silencing: The discovery of gene silencing using double stranded RNA, also known as RNA interference (RNAi) and the development of gene knockdown using Morpholino oligos, have made disrupting gene expression an accessible technique for many more investigators. This method is often referred to as a gene knockdown since the effects of these reagents are generally temporary, in contrast to gene knockouts which are permanent. RNAi creates a specific knockout effect without actually mutating the DNA of interest. While RNA interference relies on cellular components for efficacy (e.g. the Dicer proteins; Figure 1) a simple alternative for gene knockdown is Morpholino antisense oligos. Morpholinos bind and block access to the target mRNA without requiring the activity of cellular proteins and without necessarily accelerating mRNA degradation. Morpholinos are effective in systems ranging in complexity from cell-free translation in a test tube to *in vivo* studies in large animal models [10].

Interference Using Transgenes: Another way that gene function can be analyzed is by adding DNA sequences of interest to the genome of an organism that normally lacks such sequences and then seeing the effect that the introduced sequence has on the organism's phenotype. This method is a form of reverse genetics. An organism that has been permanently altered by the addition of a DNA sequence to its genome is said to be *transgenic* and the foreign DNA that it carries is called a *transgene* [11]. Here, we consider the technique for the creation of transgenic mice, which are often used in the study of the function of human genes because they can be genetically manipulated in ways that are impossible with humans and, as mammals; they are more similar to humans than are fruit flies, fish and other model genetic organisms. The oocytes of mice and other mammals are large enough that DNA can be injected into them directly. Immediately after penetration by a sperm, a fertilized mouse egg contains two pronuclei, one from the sperm and one from the egg; these pronuclei later fuse to form the nucleus of the embryo. Mechanical devices can manipulate extremely fine, hollow glass needles to inject DNA directly into one of the pronuclei of a fertilized egg. If several hundred embryos are injected and implanted, there is a good chance that one or more mice whose chromosomes contain the foreign DNA will be born [6].

Major Applications of Reverse Genetics

To Study RNA Viruses: To study RNA viruses by reverse genetics, their complete genomes must first be reverse transcribed into a complimentary DNA (cDNA) clone, which can then be manipulated in many ways. To rescue or create an infectious, recombinant RNA virus, the cDNA clone must be converted back into RNA. Negative-sense RNA viruses are problematic because their viral RNA lacks the signals recognized by eukaryotic host cells for transcription and translation. Therefore, no viral positive or negative-sense genomic RNA strands, viral messenger RNA (mRNA) strands, or viral proteins are produced from transfected naked negative-sense viral genomic RNA alone to induce an infection. To circumvent these problems, viral proteins necessary for viral mRNA transcription and genomic replication, which are supplied by the virus during normal replication, need to be present during the artificial rescue of an RNA virus from a cDNA clone. Once an infectious, negative-sense RNA virus has been rescued, the phenotype of the recombinant virus can be studied [1].

Viral Induced Expression of a Foreign Gene: One application using reverse genetics is to express a protein, which is foreign to the viral host, from a distinct gene on the virus genome. For the foreign gene to be recognized

and expressed from the viral genome it must mimic a viral gene by encoding the necessary viral mRNA regulation sequences. The genomes of the *Paramyxovirinae* subfamily, which are non-segmented, negative-sense, single stranded RNA (ssRNA) viruses, contain a series of 6-7 distinct gene units, which may contain one or more open reading frames (ORF). The main proteins expressed from the distinct gene units are NP, P, M, F, HN or heam-agglutinine (H) and L; also SH for rubella viruses (Figure 2A). The series of gene units are flanked by 3' leader and 5' trailer untranslated regions (UTRs) that are essential for viral transcription and replication regulation [12].

Each gene unit is separated by semi-conserved intergenic regions that consist of gene end, intercistronic and gene start sequences (Figure 2B). Therefore, when inserting a foreign gene as a distinct gene unit, the new gene unit must contain the gene end, inter-cistronic and gene start sequences to be effectively expressed through viral mRNA transcription. During viral mRNA synthesis, the viral RNA polymerase recognizes the gene end sequence and stutters, adding non-templated adenosine (A) residues to create a poly-A tail. The viral RNA polymerase then reengages viral mRNA transcription at the gene start sequence for the next gene unit. The viral RNA polymerase will sometimes fail to reengage mRNA transcription for the downstream gene, which results in fewer mRNA transcripts for downstream genes compared to upstream genes transcribed from the same template. This phenomenon is called *transcriptional polarity* and ultimately leads to less protein expressed from the downstream gene units in a gradient fashion [13].

This phenomenon can be used advantageously for regulation purposes by cloning the new foreign gene unit into various locations on the viral genome. For example, if the expression of the foreign gene is to be up-regulated then the gene unit could be cloned towards the 3' end of the viral genome, but if the foreign gene is to be down-regulated then the gene unit could be cloned towards the 5' end of the viral genome. Furthermore, if the foreign gene unit is to be cloned as the first gene on the viral genome then a second viral regulatory sequence needs to be taken into consideration. Past the 3' leader UTR are three G ribo-nucleotides equally separated by five ribo-nucleotides, which correspond to one complete turn of the 3-dimensional helical encapsidated RNA genome. This location is thought to co-regulate viral replication perhaps through the assembly and binding of the L-P complex with the encapsidated RNA genome [14].

This method has been used to effectively express a foreign gene from the genome of infectious, recombinant ssRNA viruses for many purposes. For example, a recombinant *Sendai Virus* (SeV), which is non-pathogenic to humans thus a prime candidate for vaccine development and gene transfer, has been used as a expression vector in gene transfer to deliver foreign eukaryotic genes to neurons and airway epithelium [15].

In addition, several recombinant *paramyxo-viruses* have been developed that express other viral surface proteins to be used as vaccines. The ebola-virus glycoprotein and the *Respiratory Syncytial Virus* (RSV) fusion protein have been expressed from recombinant *Human Parainfluenza Virus-3* (HPIV-3) and SeV viruses, respectively and have resulted in protective immunity against *ebola-virus* and RSV, respectively. A third area where the insertion of a foreign gene into a recombinant virus has been used is for the expression of a reporter gene for the purpose of tracing the viral infection. A recombinant HPIV-3, strain JS, has been engineered to express the enhanced green fluorescent protein (EGFP) and was used to trace the infection of HPIV-3 exclusively to the apical surface of ciliated airway epithelium by attaching to α2-6-linked sialic acid receptors [16].

Characterization of Viral Accessory Proteins Using Recombinant Knockout Viruses: The other application using reverse genetics is to study the functions of viral genes through mutational analysis. The cDNA clone for a ssRNA virus can serve as backbone for genetic manipulation and infectious, recombinant viruses can be rescued from it, assuming the mutation is not deleterious to the virus. The recombinant parent and mutant viruses can then be measured using traditional and modern virological techniques and compared to each other. If any differences in growth characteristics, for example, are seen between the two recombinant viruses, the change may be attributed to the mutations of the mutant virus [1].

For the most part, this method has been used to study the accessory proteins expressed by ssRNA viruses because they are generally not essential for virus replication, even though they may play part in replication in some way. Most other genes in the viral genome are essential for viral replication and any deliberate knockouts or mutations may render the recombinant virus uninfectious. The accessory proteins are expressed from mRNA transcribed from the P gene of most *paramyxo-viruses* genomes [17].

Reverse Genetics for Live-Attenuated Virus Vaccine Development: Historical methods for developing live-attenuated vaccines, performed extensive passage in cell culture at increasingly suboptimal temperatures or chemical mutagenesis to attenuate virus pathogenisity. This can be very labor intensive and may take years to establish a thoroughly accredited vaccine strain, in which the gene elements or segment(s) contributed to the attenuation phenotype are well defined and characterized. With the advancement of molecular biology and genetic engineering technologies, along with the increased understanding at the molecular levels of immune response, new approaches on live-attenuated virus vaccine development have emerged. One approach for molecular attenuation of RNA viruses is reverse genetics, a process of generating a recombinant virus from a cloned cDNA copy of a viral genome [18].

Reverse genetics is a powerful molecular tool to attenuate virus and understand the molecular determinants related to virus attenuation, tissue tropism and virulence factor(s). In recent years, reverse genetics has markedly accelerated the development of virus vaccines due to the feasibility that the viral genomes can be modified at will through a cDNA intermediate. More stable attenuated mutants can be generated by rational designs such as point mutations, deleted genes, changed virus gene expression order and substituted gene or attenuating mutations from related viruses. In addition, the system provides reliable methods to identify and develop new types of attenuating mutations, including deletion of non essential genes for virus replication and construct marker vaccine viruses containing the specific sequence tag or restriction enzyme site(s) that is essential to distinguish the vaccinated groups from the infected group [19].

Treatment of High Cholesterol with RNAi: Recent research has examined the potential of RNAi for the treatment of disorders of cholesterol metabolism. Although cholesterol is essential for life, too much cholesterol is unhealthy: high blood cholesterol is a major contributor to heart disease. Cholesterol is normally transported throughout the body in the form of small particles called lipoproteins, which consist of a core of lipids surrounded by a shell of phospholipids and proteins. The apolipoprotein B (ApoB) is an essential part of lipoproteins. Some people possess genetic mutations that cause elevated levels of ApoB, which predisposes them to coronary artery disease. Findings from studies suggest that lowering the amount of ApoB can reduce the number of lipoproteins and lower blood cholesterol in these people, as well as in people who have elevated cholesterol for other reasons. In 2006, Tracy Zimmermann and her colleagues at Alnylam Pharmaceuticals and Protiva Biotherapeutics demonstrated that RNAi could be used to reduce the levels of ApoB and blood cholesterol in non human primates [6].

Treatment of Liver Cancer with RNAi: Another potential application of RNAi is the treatment of cancer. Research has demonstrated that human liver-cancer cells consistently express a micro RNA (miRNA) called miR-26a at lower levels than do normal cells. This miRNA normally helps to regulate the cell cycle; its reduced production leads to impaired cell-cycle control, abnormal cell division and cancer. Janaiah Kota and colleagues placed sequences for miR-26a in adeno-associated virus and injected the genetically engineered viruses into mice that are predisposed to liver cancer. Only 2 out of 10 mice that received miR-26a sequences developed cancer, whereas 6 out of 10 control mice that received adeno-associated virus without the added miR-26a sequences developed cancer. These results suggest that miRNAs might one day be used to effectively treat cancer [20].

To Study about Plants: Genome sequencing projects have identified large numbers of plant genes for which no role has yet been defined. To address this problem a number of techniques have been developed, over the last 15 years, to enable researchers to identify plants with mutations in genes of known sequence. These reverse genetic approaches include RNAi and related technologies and screening of populations mutagenised by insertion (PCR), deletion (PCR) and point mutation (TILLING), each with its own strengths and weaknesses. The development of next-generation sequencing techniques now allows such screening to be done by sequencing. In the future, it is likely that the genomes of thousands of plants from mutagenised populations will be sequenced allowing for the identification of plants with mutations in specific genes to be done *in silico* [4].

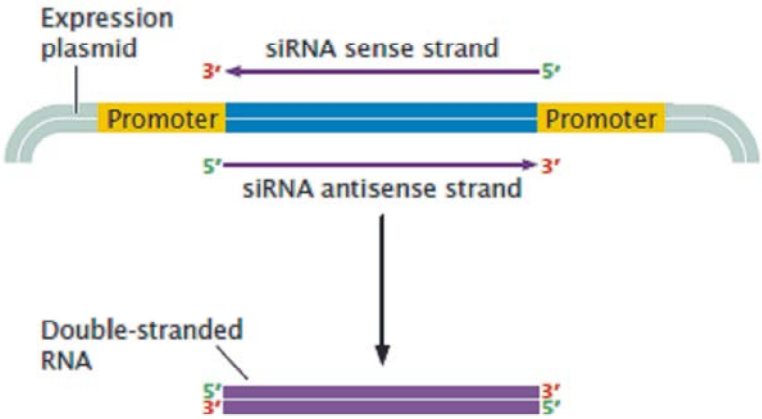

Fig. 1: Small interfering RNAs (siRNAs) can be produced by cloning DNA sequences corresponding to the siRNAs between two strong promoters. When cloned into an expression vector, both DNA strands will be transcribed and the complementary RNA molecules will anneal to form double-stranded RNA that will be processed into siRNA by Dicer [6]

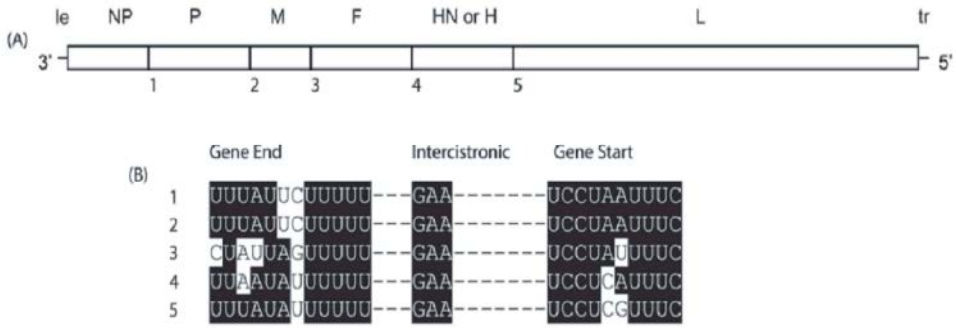

Fig. 2: Organization of the major genes units encoded in the *Paramyxovirinae* genomes. (A) The viral genomes consists of the following major features in order: 3' leader (le), nucleocapsid protein (NP), phosphoprotein (P), matrix protein (M), fusion protein (F), hemagglutinin-neuraminidase (HN) or hemagglutinin (H), large protein (L) and 5' trailer (tr). (B) Semi-conserved gene end, intercistronic and gene start sequences of HPIV-3 at the gene junctions NP-P (1), P-M (2), M-F (3), F-HN (4) and HN-L (5) [1]

CONCLUSION

Reverse genetics is a discipline that involves the use of genetic manipulation and modification to study an organism's altered phenotype. It provides us with significant information about which genes or mutations cause a manifest phenotype we observe. Additionally, a second mutation that suppresses the phenotype governed by the first mutation gives us a clue to the genetic interaction of cellular molecules. Such genetics can be applied to the study of all living organisms because genes and genetic codes are a universal system for all living organisms. There are different popular techniques available for reverse genetics studies in different organisms. All have their advantages and disadvantages depending on the species and researcher and the questions being asked. Finally, to add more on the current genetic information of different organisms advanced studies should be done on the possible alternative techniques and applications of reverse genetics to solve the already existing biomedical related problems of the globe.

REFERENCES

1. Jason, P.R., 2009. The Use of Reverse Genetics to Clone and Rescue Infectious, Recombinant Human Parainfluenza Type 3 Viruses. Utah State University Digital Commons@USU. All Graduate Theses and Dissertations.

2. Sola, I., S. Alonso, S. Zúñiga, M. Balasch, J. Plana-durán and L. Enjuanes, 2003. Engineering the transmissible gastroenteritis virus genome as an expression vector inducing lactogenic immunity. J. Virol., 77: 4357-4369.
3. Serge, H., L. Vincent, A. Yann and P. Luc, 2013. Reverse genetics in eukaryotes IGDR, Institut de Génétique et Développement de Rennes CNRS: UMR6061, Université de Rennes, IFR140, FR.
4. Erin, G. and H, George, 2010. Reverse genetics techniques: engineering loss and gain of gene function in plants. Briefings in functional genomics. Volume 9, number 2. 103^110doi:10.1093/bfgp/elp059.
5. Manis, J.P., 2007. "Knock Out, Knock In, Knock Down - Genetically Manipulated Mice and the Nobel Prize" (PDF). New England Journal of Medicine (Waltham, MA, USA: Massachusetts Medical Society), 357(24): 2426-2429.
6. Benjamin, A., 2012. Genetics: A Conceptual Approach, Fourth edition, Southwestern University, W. H. Freeman and Company, New York.
7. Bazopoulou, D. and N. Tavernarakis, 2009. The NemaGENETAG initiative: large scale transposon insertion gene-tagging in Caenorhabditis elegans, Genetica, 137: 39-46.
8. Wick, L.M., W. Qi, D.W. Lacher and T.S. Whittam, 2005. Classical and Reverse Genetics, J. Bacteriol., 187: 1783.
9. Eckardt, N.A., 2007. Positive and negative feedback coordinate regulation of disease resistance gene expression. Plant Cell, 19: 2700-2.
10. Jose, S., C. Kenneth, J.H. Gregory and V.R. Fabiola, 2004. RNA-interference-based functional genomics in mammalian cells: Reverse genetics coming of age, Oncogene, 23: 8401-8409.
11. Hedrich, H.J., 2004. The Laboratory Mouse (The Handbook of Experimental Animals). Elsevier, New York.
12. Calain, P. and L. Roux, 1995.Functional characterization of the genomic and antigenomic promoters of Sendai virus. Virology, 212: 163-173.
13. 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. U.S.A., 95: 3501-3506.
14. Tapparel, C., D. Maurice and L. Roux, 1998. The activity of Sendai virus genomic and antigenomic promoters requires a second element past the leader template regions, J. Virol., 72: 3117-3128.
15. Shirakura, M., M. Fukumura, M. Inoue, S. Fujikawa, M. Maeda, K. Watabe, S. Kyuwa, Y. Yoshikawa and M. Hasegawa, 2003. Sendai virus vector-mediated gene transfer of glial cell line-derived neurotrophic factor prevents delayed neuronal death after transient global ischemia in gerbils. Exp. Anim., 52: 119-127.
16. Zhang, L., A. Bukreyev, C.I. Thompson, B. Watson, M.E. Peeples, P.L. Collins and R.J. Pickles, 2005. Infection of ciliated cells by human parainfluenza virus type 3 in an in vitro model of human airway epithelium. J. Virol., 79: 1113-1124.
17. Galinski, M.S., R.M. Troy and A.K. Banerjee, 1992. RNA editing in the phosphoprotein gene of the human parainfluenza virus type 3. Virology, 186: 543-550.
18. Chen, H., K. Subbarao and D. Swayne, 2003. Vaccine, 21: 1974-1979.
19. Kun, Y. and W. Zaishi, 2008. Reverse Genetics for Live Attenuated Virus Vaccine Development, Cancer Immunotherapy and Trends In Bio/Pharmaceutical Industry.
20. Eamens, A., M.B. Wang and N.A. Smith, 2008. RNA silencing in plants: yesterday, today and tomorrow. Plant Physiol., 147: 456-68.