doi:10.1006/geno.2001.6653, available online at http://www.idealibrary.com on IDEAL 

# Viral Vectors as Part of an Integrated Functional Genomics Program

Christopher G. Janson* and Matthew J. During

*CNS Gene Therapy Center, 1025 Walnut Street, Suite 511, Philadelphia, Pennsylvania 19107, USA*

**To whom correspondence and reprint requests should be addressed. Fax: (215) 955-9629. E-mail: cgjanson@hotmail.com.*

**Over the past decade, viral vectors have slowly gained mainstream acceptance in the neuroscience and genetics communities for the *in vivo* study of gene function [1]. Using stereotactic techniques, it is possible to characterize neuroanatomical relationships through the delivery of neurotropic viral vectors to specific brain regions. More sophisticated studies combine viral vectors with other methods of genetic manipulation such as germline transgenic mice. As more is learned about the properties of different viral vectors, it has become possible to use viral vectors to test hypotheses about the function of genes, through targeted *in vivo* delivery to the central nervous system (CNS). The effects of gene expression in the brain can be measured on the molecular, biochemical, electrophysiological, morphological, and behavioral levels. We propose that viral vectors should be considered as part of an integrated functional genomics platform in the CNS.**

## What's "Wrong" with Transgenic Mice?

The most definitive approach in assigning gene function in mammals has been the generation of transgenics using embryonic stem (ES) cell technology in mice, and pronuclear injection or nuclear transfer techniques with other animals (for example, rats, sheep, and monkeys) in which ES technology has lagged [2–5]. Although several investigators are developing non-murine ES cells [6] in monkeys and humans, that technology has not been fully realized as of this writing. Nuclear transfer is a promising transgenic technique, but remains technically demanding and inefficient [7,8]. Indeed, the problem with most germline transgenic approaches is that they are very laborious and time-intensive, particularly genetic "knockouts" which involve the construction of homologous targeting vectors as well as the care and breeding of heterozygous animals, and the rate of gene cloning has rapidly outpaced our ability to do functional studies *in vivo.* Moreover, standard transgenic animals are expensive, difficult to breed, and may be sterile.

Another problem with classical transgenics is maintaining tight control over gene expression. A pleiotropic gene may fail to generate a viable animal when it is eliminated or overexpressed early in development, and some promising transgenics have turned out to have lethal phenotypes. On the opposite extreme, transgenics can have completely silent or masked phenotypes due to redundancy of gene functions, which means that the substantial investment of creating a transgenic occasionally results in a phenotypically normal animal. Although transgenics have been generated using brain-specific promoters such as nestin [9,10] and CaM-kinase-II [11,12], expression is often "leaky" and the gene may be expressed outside of the intended region. Somewhat tighter control over gene expression is possible through small molecule response elements [13–15], but the extra time and expense of creating these regulatable elements tend to limit their use to cases in which it is necessary to avoid an early lethal phenotype or to study a particular developmental stage. Simple promoter targeting in transgenics lacks the extra temporal-spatial precision possible with viral-based, somatic cell "transgenics," in which genes can be stereotactically delivered to the brain under control of both capsid proteins and promoter/regulatory elements.

Despite their limitations, classical transgenics remain the ultimate tool for testing a gene's function and provide many useful animal models. These models are generally compatible with various techniques of somatic cell transgenesis, including viral vectors. For example, it is possible to create hybrid germline/somatic hybrids to study brain function, with viral vector-mediated stereotactic delivery of the *cre* recombinase gene to transgenics containing flanking *lox* elements [16]; this method combines the standard germline approach [17,18] with the rapidity and temporal-spatial specificity of viral vectors, allowing investigators to "turn on" the effects of the transgene in a well-controlled manner while avoiding the time and expense of selecting and breeding *cre-lox* double heterozygotes.

In both somatic-cell and germline modified animals, it is possible to exert control over gene expression using drug-responsive elements that can activate an otherwise silent phenotype or repress a constitutively active one. These drugs, typically steroids and antibiotics, are designed to cross the blood–brain barrier and can be peripherally administered. Recombinant adeno-associated virus (AAV) vectors have been used to introduce inducible/repressible genes to the rodent brain [19], regulated both by neural-specific promoters and drug-responsive elements, allowing for multi-level control over expression [20,21]. Alone or in combination with classical transgenics, a range of inducible/repressible gene constructs are available for targeting expression and assessing complex temporal-spatial patterns of gene function in the brain.

doi:10.1006/geno.2001.6653, available online at http://www.idealibrary.com on IDEAL

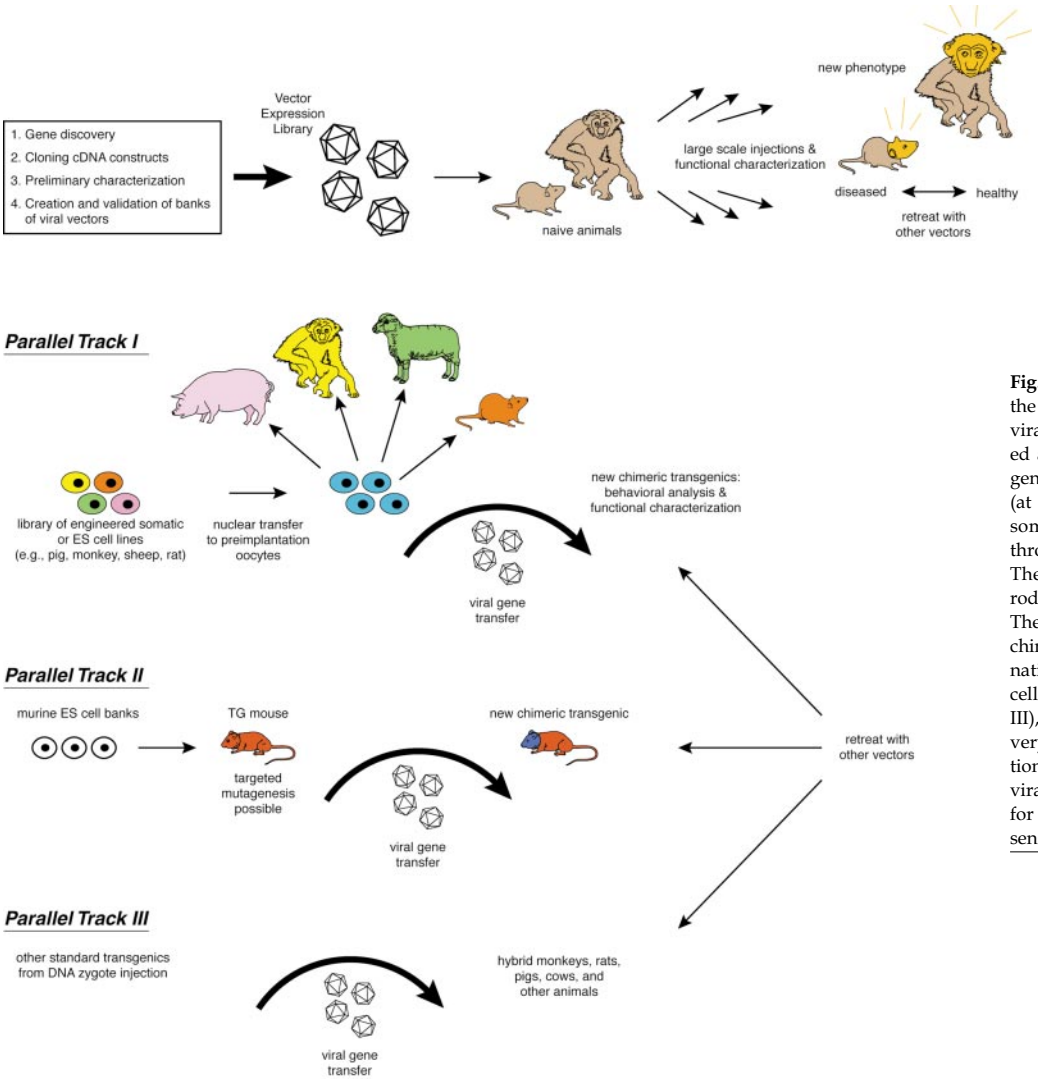


**Fig. 1.** The figure at left represents the manner in which recombinant viral vectors might be incorporated as part of an *in vivo* functional genomics program. The first stage (at top) is simply the creation of somatic cell transgenic animals through viral delivery to the brain. The recipient animals may be rodents, primates, or other species. The next stage is the creation of chimeric animals through a combination of germline and somatic cell techniques (parallel tracks I, II, III), which carries the potential for very complex genetic investigations. In addition to recombinant viral gene vectors, other vectors for retreatment may include antisense and stem cells.

## Toward the Use of Viral Vectors for Large-Scale Expression Studies

At least $3 \times 10^4$ human genes (of which many are likely to be important in the brain) have been identified through the Human Genome Project (HGP), and other sequencing projects are already underway in mouse, rat, dog, cat, sheep, pig, cow, monkey, and kangaroo [22,23]. With the completion of whole-genome sequencing under the HGP and ongoing efforts in other mammalian projects, an important application of viral gene transfer will be *in vivo* characterization of newly cloned genes. As genes are directly sequenced and assigned putative functions based on theoretical considerations, promising candidates are being selected for many disease and developmental genes. Some basic characteristics of a gene may be inferred from internal sequence cues or through database comparison of DNA and protein homologues, but such theoretical analysis is limited on its own, especially given the gaps in current mammalian sequencing projects. The completion of large-scale sequencing projects in a wide variety of organisms will facilitate prospective or comparative genomics [24], and data mining will then provide candidate genes for *in vivo* functional characterization.

Recent advances in techniques of combinatorial expression analysis such as RT-PCR clustering [25], serial analysis of gene expression (SAGE), and particularly gene microarrays [26] are important for identifying structure–function relationships among gene families, yet their main role is correlative analysis of genes rather than proving a function for a given gene in a specific context. Theoretical models offer a provisional and largely static view of gene function, whereas genes are dynamic entities that form a complex, interrelated pattern in living organisms; the operation and function of a gene may vary greatly depending on what other genes are co-expressed in a particular brain region or physiologic condition, sometimes over a very short period of time. For example, genes encoding transcription factors, cytoskeletal elements, or

biosynthetic enzymes are likely to alter the function of many other genes in real time. In a complex and heterogeneous system such as the brain, the ultimate test is to change gene expression and then to assess different physiologic outcomes over time. The latest generation of viral vectors, such as AAV and lentivirus, is ideal for introducing multiple real-time changes in gene expression, due to their high levels of expression and lack of deleterious effects.

Following the isolation of interesting clones from genomic and brain cDNA libraries, it will be possible to administer new genes and splice variants, singly or in combination, to the brain with viral vectors. The advantages of making cDNA clones available to investigators for functional genomics were addressed by the IMAGE consortium [27] and more recently by the "Mammalian Gene Collection" (MGC) of full-length cDNA resources [28]. These clearinghouses will help to accelerate the pace of brain functional studies using gene transfer vectors.

Once a gene is chosen for study, packaged in viral vectors, and stereotactically delivered, it is possible to study effects on animal behavior, neurochemistry, electrophysiology, and regional gene co-expression. Phenotypic changes can be measured through cognitive and motor tests, many of which are now fully automated through the use of laser or infrared grids and data analysis software. Brain chemistry, electrical activity of neurons, and metabolism can be measured in real time with microdialysis and depth electrodes (voltametry, biosensors). Anatomical changes can be observed with non-invasive imaging to follow the expression patterns of brain proteins in live mammals in real time. Techniques for *in vivo* imaging of brain reporter genes using magnetic resonance scanning have existed for several years [29,30] and, more recently, charge-coupled devices (CCD) with image analysis software allowed *in vivo* quantification of light-emitting proteins [31,32]. Positron emission tomography (PET) technologies also have been used to measure *in vivo* expression of certain marker genes in conjunction with pro-drugs [46,47]. These techniques add to our ability to study real-time brain gene expression in mammals, in the same way that *Drosophila melanogaster* development and its temporal-spatial transcriptional gradients were modeled in the past. Alternatively, one can indirectly assess gene function in the brain by serially monitoring the co-activation of other genes at the RNA and protein level using microarray analysis and immunohistochemical techniques.

## Viral Vectors and the "High-Throughput" Dilemma

Viral vectors have never been perceived as good candidates for "high-throughput" or industrial-scale characterization of genomic DNA of unknown function. At this time, they are best suited to detailed analyses of genes that have been partially characterized. Though it generally requires less work than germline transgenic approaches, the production of large quantities of high-titer virus and neurosurgical injection of experimental animals is still a labor-intensive and costly process. One improvement in AAV production uses helper-virus-free production techniques with plasmid co-transfection and column purification [33], which makes high-titer rAAV production faster and less expensive. Compared with other techniques, viral vectors may eventually offer an alternative high-throughput method to address complex neuroanatomical and neurophysiologic questions.

The most high-throughput method of gene capture and analysis is large-scale, systematic disruption of genes, which has been implemented in both mammalian and non-mammalian systems. For example, the Berkeley Drosophila Project has carried out insertional mutagenesis throughout the *Drosophila* genome, the final goal being to catalog the disruption of all ~ 12,000 genes. These disruptions are classified on the basis of both phenotype and direct molecular mapping, and have been shown to share functional similarities with a number of mammalian genes [34,35]. Large-scale genomic disruption also has been performed in mice using ES technology; several promoter-trap or gene-trap methods of targeting have been used [36–39] for tagging and retrieval of genes. When the respective transgenic clones are grown using ES cells, the spatial-temporal localization of "trapped" genes in the brain can be studied through reporter gene expression patterns [40,41]. Both deletional and insertional mutagenesis have been performed, and rapid identification of transgenic animals is possible on the basis of inserted coat color genes [42].

The final goal of "gene trapping" is to create a comprehensive library of cells with readily identifiable disruptions throughout the genome, so that clonal cells with tagged mutations may be applied to the generation of transgenic animal lines through ES or nuclear transfer techniques. Many transgenic clones with a defined genetic knockout may thus be generated. An important limitation of whole-genome mouse knockouts is that the mouse genome is not strictly homologous to the human genome and thus certain genes cannot be tested. Moreover, some knockout regions may be cryptic. Finally, this technology is not widely available to test numerous genes that show immediate potential for generating animal models. When they do become accessible, comprehensive collections of "knockout" cell and animal clones will provide valuable tools that will prove especially useful in conjunction with viral vectors — for instance in phenotypic rescue studies or the generation of new chimeras using viral gene transfer to the brain.

An alternative high-throughput approach to functional genomics involves the systematic addition of genomic sequences using yeast and bacterial artificial chromosomes (YACs and BACs, respectively) to create transgenic mice, either by directly introducing a large piece (~ 150 kb) of genomic DNA with its native promoter and regulatory elements into the pronucleus of a mouse oocyte, or by introducing the YAC/BAC construct to ES cells by lipofection, which are then fused with blastocysts to make chimeras [43]. Animals other than mice may be modified when the non-ES-cell technique is used, which has the advantage of a more direct approach. Although pronuclear injection is still an arduous and inefficient technique, *in vitro* manipulation of the YAC/BAC allows for precise mutagenesis and attachment of marker genes or epitopes to the genes of interest [44].

doi:10.1006/geno.2001.6653, available online at http://www.idealibrary.com on IDEAL

A proposed advantage of the BAC approach for functional genomics is that large native promoter elements can be incorporated, which tend to minimize positional effects on gene expression once the gene integrates into the mouse genome. These native promoter elements may be quite weak or unpredictable, however, and incapable of expressing every gene of interest to the extent that it will demonstrate a detectable phenotype. The United States government has funded a "High-Throughput Expression Project" that aims to create a "gene pipeline" of as many as 5000 genes in a 5-year period, using BAC technology to express human disease genes, developmental genes, and unknown genes (ESTs) in mice. The aim is to catalog detailed neuroanatomical expression patterns that may be used in correlative studies on gene function. In terms of gene cloning and handling of animals, this gene pipeline involves a prodigious amount of work and is certainly no faster than recombinant viral techniques. Viral vectors have already been used to target ES cells, and could be used in a complimentary fashion to BAC transgenesis, both for germline and somatic studies. Although viral vectors such as AAV and lentivirus are much too small to deliver many genomic sequences, especially those with large promoter sequences and introns, newer vectors such as viral-encapsidated artificial chromosomes may supplement transgenic studies with BACs in the future [45].

## Conclusion

Here we have tried to emphasize the complimentary nature of viral-based somatic cell transgenesis. Viral-based methods have several important limitations, most significantly their relative lack of standardization and their inherent size limitations. Viral vectors will never replace transgenic animals as a research tool; however, when used in conjunction with transgenic technology, the full value and power of viral-based transgenesis can be realized (see figure). Further progress in virology and vector design will help to realize this goal of an integrated approach to functional genomics in the brain.

## References


1. Janson, C. G., *et al.* (2001). Viral-based gene transfer to the mammalian CNS for functional genomic studies. *Trends Neurosci.* (in press).
2. Hammer, R. E., *et al.* (1985). Production of transgenic rabbits, sheep, and pigs by microinjection. *Nature* **315:** 680–683.
3. Charreau, B., *et al.* (1996). Transgenesis in rats: technical aspects and models. *Transgenic Res.* **5:** 223–234.
4. Fitchev, P., *et al.* (1999). Nuclear transfer in the rat: potential access to the germline. *Transplant Proc.* **31:** 1525–1530.
5. Wolf, D. P. (1999). Nuclear transfer in the rhesus monkey: practical and basic implications. *Biol. Reprod.* **60:** 199–204.
6. Thomson, J., *et al.* (1995). Isolation of a primate embryonic stem cell line. *Proc. Natl. Acad. Sci. USA* **92:** 7844–7848.
7. Gurdon, J. B., *et al.* (1999). The future of cloning. *Nature* **402:** 743–746.
8. Kubota, C., *et al.* (2000). Six cloned calves produced from adult fibroblast cells after long-term culture. *Proc. Natl. Acad. Sci. USA* **97:** 990–995.
9. Lendahl, U. (1997). Transgenic analysis of CNS development and regeneration. *Acta Anaesthesiol. Scand. Suppl.* **110:** 116–118.
10. Tronche, F., *et al.* (1999). Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. *Nat. Genet.* **23:** 99–103.
11. Mayford, M., *et al.* (1996). Control of memory formation through regulated expression of a CaMKII transgene. *Science* **274:** 1678–1683.
12. Tang, Y., *et al.* (1999). Genetic enhancement of learning and memory in mice. *Nature* **401:** 63–69.
13. Chen, J., *et al.* (1998). Transgenic animals with inducible, targeted gene expression in brain. *Mol. Pharmacol.* **54:** 495–504.
14. Jerecic, J., *et al.* (1999). Studies on conditional gene expression in the brain. *Ann. NY Acad. Sci.* **868:** 27–37.
15. Gao, X., Kemper, A., and Popko, B. (1999). Advanced transgenic and gene-targeting approaches. *Neurochem. Res.* **24:** 1181–1188.
16. Brooks, A. I., *et al.* (1997). Nerve growth factor somatic mosaicism produced by herpes virus-directed expression of cre recombinase. *Nat. Biotechnol.* **15:** 57–62.
17. Smith, A. J. H., *et al.* (1995). A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. *Nat. Genet.* **9:** 376–384.
18. Call, L. M. (2000). A cre-lox recombination system for the targeted integration of circular yeast artificial chromosomes into embryonic stem cells. *Hum. Mol. Genet.* **9:** 1745–1751.
19. Haberman, R. P., *et al.* (1998). Inducible long-term gene expression in brain with adeno-associated virus gene transfer. *Gene Ther.* **5:** 1604–1611.
20. Gossen, M., *et al.* (1995). Transcriptional activation by tetracyclines in mammalian cells. *Science* **268:** 1766–1769.
21. Rivera, V. M., *et al.* (1996). A humanized system for pharmacologic control of gene expression. *Nat. Med.* **2:** 1028–1032.
22. Collins, F., *et al.* (1998). New goals for the US Human Genome Project. *Science* **282:** 682–689.
23. Refer to GenomeWeb at the U.K. Human Genome Mapping Project (http://www.hgmp.mrc.ac.uk), sponsored by the British Medical Research Council, for a full listing of vertebrate genome databases.
24. O'Brien, S. J., *et al.* (1999). The promise of comparative genomics in mammals. *Science* **286:** 458–481.
25. Wen, W., *et al.* (1998). Large-scale temporal gene expression mapping of central nervous system development. *Proc. Natl. Acad. Sci. USA* **95:** 334–339.
26. Fields, S., Kohara, Y., & Lockhart, D. J. (1999). Functional genomics. *Proc. Natl. Acad. Sci. USA* **96:** 8825–8826.
27. Lennon, G., *et al.* (1996). The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33:** 151–152.
28. Strausberg, R. L., Feingold, E. A., Klausner, R. D., and Collins, F. S. (1999). The Mammalian Gene Collection. *Science* **286:** 455–457.
29. Bogdanov, A., Jr., and Weissleder, R. (1998). The development of in vivo imaging systems to study gene expression. *Trends Biotechnol.* **16:** 5–10.
30. Louie, A. Y., *et al.* (2000). In vivo visualization of gene expression using magnetic resonance imaging. *Nat. Biotechnol.* **18:** 321–325.
31. Contag, P., *et al.* (1998). Bioluminescent reporters in living mammals. *Nat. Med.* **4:** 245–247.
32. Yang, M., *et al.* (2000). Visualizing gene expression by whole-body fluorescence imaging. *Proc. Natl. Acad. Sci. USA* **97:** 12278–12282.
33. During, M. J., *et al.* (1998). Peroral gene therapy of lactose intolerance using an adeno-associated virus vector. *Nat. Med.* **4:** 1131–1135.
34. Spradling, A. C., *et al.* (1999). The Berkeley Drosophila Genome Project Gene Disruption Project: Single P-element insertions mutating 25% of vital Drosophila gene. *Genetics* **153:** 135–177.
35. Maleszka, R., *et al.* (1998). Data transferability from model organisms to human beings: Insights from the functional genomics of the flightless region of Drosophila. *Proc. Natl. Acad. Sci. USA* **95:** 3731–3736.
36. Durick, K., Mendlein, J., and Xanthopoulos, K. G. (1999). Hunting with traps: genome-wide strategies for gene discovery and functional analysis. *Genome Res.* **9:** 1019–1025.
37. Zambrowicz, B. P., *et al.* (1998). Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature* **392:** 608–611.
38. Wiles, M. V., *et al.* (2000). Establishment of a gene-trap sequence tag library to generate mutant mice from ES cells. *Nat. Genet.* **24:** 13–14.
39. Whitney, M., *et al.* (1998). A genome-wide functional assay of signal transduction in living mammalian cells. *Nat. Biotechnol.* **16:** 1329–1333.
40. Friedrich, G., and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5:** 1513–1523.
41. Stokova, A., *et al.* (1998). Gene trap expression and mutational analysis for genes involved in the development of the mammalian nervous system. *Dev. Dyn.* **212:** 198–213.
42. Zheng, B., Mills, A. A., and Bradley, A. (1999). A system for rapid generation of coat color-tagged knockouts and defined chromosomal rearrangements in mice. *Nucleic Acids Res.* **27:** 2354–2360.
43. Peterson, K., *et al.* (1997). Production of transgenic mice with yeast artificial chromosomes. *Trends Genet.* **13:** 61–66.
44. Yang, X., *et al.* (1997). Homologous recombination based modification in E. coli and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nat. Biotechnol.* **15:** 859–865.
45. Vos, J. M. (1998). Mammalian artificial chromosomes as tools for gene therapy. *Curr. Opin. Genet. Dev.* **8:** 351–359.
46. Herschman, H.R. *et al.* (2000). Seeing is believing: non-invasive, quantitative, and repetitive imaging of reporter gene expression in living animals, using PET. *J. Neurosci Res.* **59:** 699-705.
47. Doubrovin, M. *et al.* (2001) Imaging transcriptional regulation of p53-dependent genes with PET *in vivo. Proc. Natl. Acad. Sci. USA.* **98:** 9300-9305