

Analysis of mammalian central nervous system gene expression and function using bacterial artificial chromosome-mediated transgenesis

Nathaniel Heintz⁺

Howard Hughes Medical Institute, Laboratory of Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

Received 18 February 2000; Accepted 21 February 2000

The anatomical complexity of the mammalian central nervous system (CNS) presents special problems for the analysis of CNS gene expression and function. The most difficult challenge is presented by the simple fact that there are hundreds of functionally and morphologically defined cell types in the CNS. Given this complexity, the interpretation of CNS phenotypes is often problematic. The preparation of transgenic mice carrying marked bacterial artificial chromosomes (BACs) provides an important avenue for improving our understanding of CNS-expressed genes and phenotypes. This approach can allow efficient analysis of patterns of gene expression, subcellular localization of their encoded products and neuronal projection patterns. BAC transgenic mice can also provide access to information relevant to gene function based on phenotypes arising from increased gene dosage or expression of activating and dominant-negative alleles. This review will concentrate on these issues and their relevance to the analysis of CNS-expressed genes.

INTRODUCTION

As we enter the 'post-genomic' era, it has become increasingly evident that the impact of genetics on our quest to understand the development, function and dysfunction of the mammalian central nervous system (CNS) will be enormous. New approaches in human genetics and the utilization of ever more powerful reverse genetic tools in model systems will provide increasingly frequent opportunities to take advantage of the fundamental information buried in genome sequences and the descriptive data issuing from high-throughput DNA chip or micro-array assays. In the past decade, the identification and preparation of scores of mutations that perturb the mammalian CNS have begun to shed light on the molecular mechanisms that participate in many aspects of CNS development and function. In this review, I will discuss the special problems encountered in applying genetics to the analysis of a highly complex tissue such as the nervous system, and focus on recently developed methods that circumvent some of these problems. Although it is beyond the scope of this brief review to cover the biological insights gained through analysis of the many mutations, I will provide specific examples to illustrate the importance of genetic analysis for the discovery of novel pathways governing the formation and function of the mammalian CNS.

THE COMPLEXITY OF THE MAMMALIAN CNS

From a genetic standpoint, the complexity of the mammalian CNS presents special challenges at every level of analysis. The

fundamental point of genetics is to understand function through the correlation of phenotype and genotype. The functions controlled by the mammalian CNS are many and varied, and often these functions are not accessible in genetically tractable model systems. For example, studies of complex human behaviors, neurological diseases and psychiatric illnesses (1–4) have demonstrated a strong influence of inheritance. The ability to discover the genetic basis for this influence on higher order nervous system function requires precise appreciation of the behavioral variations in the population at large and in individuals who deviate from this norm. In humans, where one has access to evaluation of subtle phenotypes through direct communication with the subject, it is possible to develop quantitative tests for advanced CNS functions. These can then be used to identify genes that play a critical role in these behaviors. With the advent of powerful methods for the dissection of polygenic traits, human genetics is likely to play a crucial role in the initial identification of genes that play critical roles in advanced CNS functions. However, complex behaviors can be very difficult to assess in model systems (5). Thus, the use of reverse genetic approaches to interrogate the system further and discover the mechanistic roles of the encoded gene products can be severely limited by the lack of suitable quantitative tests for advanced CNS functions in experimental organisms. A major challenge, therefore, is to combine insights gained from primate behavioral and clinical evaluation with mouse phenotypic studies so that genetic analysis can have maximal impact on systems-level neurobiological research.

⁺Tel: +1 212 327 7956; Fax: +1 212 327 7878; Email: heintz@rockvax.rockefeller.edu

A second challenge faced in genetic analysis of CNS function comes from the diversity of cell types that are present in the mammalian CNS and their dependence on interconnectivity to generate function (6). Given the hundreds of morphologically distinct cell types present in the brain, and the certain subdivision of these cell types based on functional distinctions or the expression of molecular markers, the interpretation of phenotypes for mutations acting on this organ system is tremendously difficult. Even in cases where the effect of the mutation is severe, resulting in frank cell death that is evident at the histological level, the site of action of the mutant gene is often difficult to assess strictly based on anatomical data. This problem derives from the interdependence of neuronal cell types and their reliance on trophic factors produced by other cell types for vitality. Thus, secondary cell death of afferent neurons often accompanies the functional inactivation or death of the specific class of neurons that are directly affected by the mutation. In most cases, therefore, the cellular basis for a CNS phenotype cannot be assessed without data from experimental embryonic chimeras or at least precise information on the pattern of expression of the mutated gene (7,8).

Genetic studies of CNS development and function are also confounded by the highly articulated structure of individual neurons and the compartmentation of molecular pathways within these very complicated cell types. In some cases, functional deficits that are easily observed at the behavioral level result in significant anatomical phenotypes that are revealed only by in-depth assessment of the number and morphology of specific neuronal cell types or specializations. Thus, mutations that block the final stages of differentiation of specific neurons often result in the failure of afferent neurons to refine exuberant connections on the affected population. An appreciation for this type of phenotype cannot come from routine anatomical studies, requiring instead either advanced neuroanatomical tracing techniques or electrophysiological studies of the appropriate neuronal populations (9). Similarly, mutations that affect the number of dendritic spines present on a given neuron are exceedingly difficult to assess, yet can result in profound CNS dysfunction (10).

THE USE OF MODIFIED BACTERIAL ARTIFICIAL CHROMOSOMES (BACS) FOR ANALYSIS OF CNS-EXPRESSED GENES

In many cases, the identification and interpretation of CNS phenotypes can be advanced by knowledge of the sequence of the affected gene, its precise pattern of expression, the localization of its encoded product within the cell, and visualization of the fine structures of cells expressing the mutant gene product *in vivo*. In vertebrates, this information has traditionally involved isolation of full-length cDNAs, in-depth *in situ* hybridization analysis, preparation of specific antibodies for use in immunohistochemical or immunofluorescence localization studies, and visualization of cellular morphology using available markers and a variety of established anatomical techniques. These methods are well established, and have yielded crucial functional insights. However, in many cases they cannot be applied efficiently and require intensive effort to yield detailed information. Several recent advances have tremendously improved the efficiency

with which these important data for CNS-expressed genes can be gathered.

The maturation of the expressed sequence tag (EST) databases and Unigene databases (11), the completion of the human genome sequence, and the improvement of informatics tools for prediction of transcripts from genomic DNA sequence (reviewed in ref. 11), are already yielding a great deal of information regarding CNS-expressed genes. Although these projects have not yet advanced sufficiently that it is routine to obtain full-length cDNA sequence and predict complete amino acid sequences for novel CNS-expressed transcripts, this important goal is clearly within reach. As these full-length sequences become available, data from the recently announced structural genomics initiatives (12) and rapidly improving methods for prediction of the structure and evolutionary relationships between expressed genes (13) will lead to increasingly accurate predictions of the biochemical functions of expressed proteins. Furthermore, large-scale EST sequencing projects can lead to the first insights into the expression profiles of interesting genes. For example, there are currently at least several thousand novel ESTs in these databases that are expressed exclusively or predominantly in cDNA libraries prepared from the brain (mouse UniGene databases, dbEST:

<http://www.ncbi.nlm.nih.gov/dbEST/index.html> ;

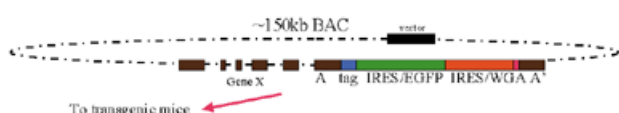
and UniGene Mouse Sequences Collection:

<http://www.ncbi.nlm.nih.gov/UniGene/Mm.Home.html>).

Although the profiles of expression of these transcripts are not sufficiently precise to yield functional inferences, this information is quite useful in selecting candidate genes for genetic analysis.

The development of methods for the simple and accurate manipulation of BACs has allowed the development of an alternative and highly efficient strategy for analysis of CNS-specific genes (14). This approach is based on two simple facts: large genomic DNA fragments (>100 kb) are in most instances expressed independent of the site of integration into the genome of transgenic mice (15–17); inclusion of epitope tags and marker proteins into endogenous loci of invertebrate genes has in most cases not altered the patterns of expression of these genes or the localization of their encoded products within the cell (18). To take advantage of this information, a homologous recombination system was established in *Escherichia coli* that allows for the preparation of BACs with highly precise modifications. Using this system, it is possible to create mutations in BACs that range from single nucleotide changes to deletions of tens of kilobases to insertions of marker genes of several kilobases (15,19–21; unpublished data). One can, therefore, construct BACs that allow very rapid analysis of the expression pattern of the gene of interest, the localization of its encoded product, high-resolution visualization of the morphology of cells expressing the gene and determination of the projection patterns of these cells.

A typical construct might resemble that shown in Figure 1. An important point to be made regarding this type of construct is that the utilization of internal ribosome entry sites (IRES) (21) can allow the creation of polycistronic mRNAs from which several proteins can be synthesized under the control of the endogenous transcriptional regulatory sequences. This is tremendously advantageous because it allows marker proteins (EGFP, lacZ, etc.) and tract-tracing proteins (WGA, other



Results:

- **EGFP:** precise expression information for each gene
 - **eTAG:** the subcellular location of its encoded product, biochemistry
 - **WGA or BL:** projection patterns for newly identified cell populations
 - **Gene Dosage:** function from increased gene dosage phenotypes
 - **BAC vectors:** molecular tools for further manipulation of genes or cells
- identify novel subpopulations of neurons and glia
 - to create accurate models for human neurologic disorders
 - tagged protein for biochemical studies

Figure 1. A typical BAC construct.

lectin, etc.) to be produced in precisely the same cells that express the endogenous gene. In our experience, expression of EGFP using this strategy in transgenic mice can allow visualization of cellular morphology in the CNS that is comparable to that obtained by visualization using immunofluorescence with marker antibodies. For example, direct visualization of EGFP fluorescence in transgenic mice expressing this marker under the control of the calbindin locus reveals the detailed structure of cerebellar Purkinje cells, whereas similar studies using mice expressing EGFP under the control of the BLBP locus reveal the elaborate morphology of Bergmann glia (Fig. 2). Using confocal microscopy, it has been possible to visualize individual Purkinje cell dendritic spines and the detailed morphology of Bergmann glial end feet in these animals (data not shown). In many cases, this detailed morphology allows definitive determination of the cell types expressing the gene based simply on their morphology. This is a great advantage over determination of gene expression patterns using *in situ* hybridization, where it is often extremely difficult to assess the cell types expressing the gene in a given brain region without additional experimentation. The ability to determine efficiently precise patterns of cell types expressing

the gene of interest throughout CNS development can significantly accelerate initial studies of gene function.

The utilization of a second IRES element in the BAC construct can allow the co-expression of yet another protein of interest (e.g. CRE recombinase, WGA, rTTA, etc.). In this context, the expression of tract-tracing proteins is of particular interest. Although we have not yet generated animals expressing tract-tracing proteins from BACs, the use of two plant lectins for this purpose has been reported using conventional transgenic animals (22,23). Given these results, we hope that it will be possible to visualize both the primary cell types expressing a given gene and their synaptic partners in a single animal. This is an exciting prospect for several reasons: in cases where the cellular morphology visualized using the primary marker protein does not allow definitive identification of the cell type, knowledge of its projection patterns might help; for the many genes that are expressed in subpopulations of morphologically defined neurons, determination of the projection patterns of this subpopulation can yield important insights into the circuit functions of these neurons; and the expression of the lectin markers may reveal details of connectivity that have not yet been discovered.

The use of epitope tags for the determination of the subcellular distribution of proteins in invertebrates and in cultured mammalian cells is very well established (18). Because of the precision of homologous recombination in *E.coli*, it is quite simple to introduce an epitope tag into the protein encoded by the gene of interest in the BAC at the same time that one introduces the marker genes (19). Since a variety of epitope tags and their cognate antibodies are now available commercially, one has a wide range of options from which to choose (18). Although the introduction of an epitope tag into the protein can in some cases change its subcellular distribution, this is relatively infrequent and usually can be overcome by changing the location of the tag within the protein. Since the preparation of useful antibodies for a protein of interest is often an expensive and long-term project, the ability to detect the epitope-tagged protein *in vivo* offers a very efficient and useful alternative. In trying to interpret CNS-

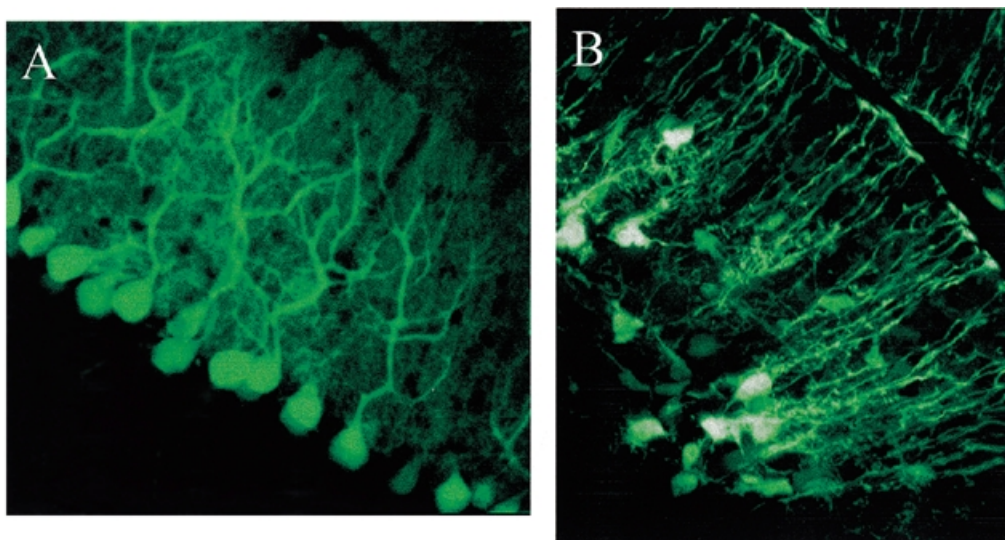


Figure 2. (A) High-resolution photograph of cerebellar Purkinje cells expressing EGFP. (B) High-resolution photograph of Bergmann cells expressing EGFP.

expressed gene function, localization of its encoded product, or correlation of its subcellular distribution in different cell types or under different conditions, can provide crucial information. Obviously, the spectrum of functions that one might consider is significantly different for proteins located in the nucleus than those present at the synapse. Furthermore, the redistribution of the protein in response to a stimulus can also be quite informative. For example, there are many well characterized transcriptional responses that involve regulated release of factors from cytoplasmic complexes and their entry into the nucleus in response to growth factors, cytokines, etc. (24). The ability to obtain this type of information in an efficient manner using epitope tags presents a significant advantage over the time-consuming preparation of sufficiently useful antibodies to the native protein for these studies. The development of peptide tags for affinity purification is also of great utility (25). We have, for example, inserted the His₆ tag into the *Zipro1* locus in BAC transgenic animals for the isolation of Zipro1 containing transcription complexes from cerebellar granule cells (19). It is now possible to utilize Ni²⁺ chelation affinity chromatography to characterize the Zipro1 complexes using whole-brain extracts from the BAC transgenic mice, as has been done very successfully for His-tagged transcription factors in cultured mammalian cells (26). This strategy can be extended for the purification of any macromolecular complex from any cell type in the brain using the BAC transgenic approach.

The utility of this strategy for analysis of CNS-expressed genes is dependent on the BAC transgenes reproducing the endogenous pattern of expression for the gene of interest. The first two cases analyzed in my laboratory were BAC transgenes for the *Zipro1* (*RU49*) locus and the *calbindin* locus. It is noteworthy that in both of these cases a large number of conventional transgenic lines had been produced and in no case had these lines reproduced the endogenous pattern of expression (27). Instead, severe position effects were observed that obviated the use of these conventional constructs for *in vivo* studies. In both of these cases, several lines of BAC transgenic mice were produced that completely reproduced the endogenous pattern of expression. The efficiency with which one obtains founders is comparable to that of conventional DNA constructs. Occasionally, a transgene-positive line is produced which does not express the gene. This may reflect integration of the BAC construct into constitutive heterochromatin, although no experiments to confirm this suggestion have been carried out.

Although the technology for the production of transgenic mice using modified BACs is relatively new, information from >50 lines of mice carrying BAC transgenes for at least 15 different genes is available (14,19–21,28,29; J.M. Friedman, H. Monyer and P.W. Rigby, personal communication, unpublished data). These data demonstrate that the pattern of expression of BAC transgenes in most cases reproduces that of the endogenous locus, that the cell types expressing the genes are the same in different lines carrying the same BAC construct, and that the level of expression of the BAC transgene correlates with the copy number inserted into the mouse genome. Thus, BAC transgene expression is in general position independent and gene dosage dependent. There are, of course, exceptions to this rule. First, for very large genes, the BAC strategy cannot work simply because the locus does not

fit into the BAC. Secondly, the position of the gene within the BAC can strongly influence its expression. If the gene is asymmetrically located in the BAC insert, e.g. if there is very little 5' flanking DNA present, then expression will not be accurate. To circumvent this problem, one should choose a BAC with the gene located near the center of the insert, so that at least 30 kb of flanking DNA are present at both the 5' and 3' ends.

MUTANTS OF THE NERVOUS SYSTEM

During the past decade, a large number of mutations that affect the mammalian nervous system have been identified by positional cloning, or prepared using homologous recombination. Many of these studies have yielded critical insights into the function or dysfunction of the mammalian CNS and these studies have been reviewed extensively in the literature (30–32). Large-scale mutagenesis projects using chemical mutagens or gene disruption strategies are currently under way to generate a mutant mouse resource that can be used to screen for additional mutations, some of which will affect the CNS (33,34). These studies will continue to yield important phenotypic information for neurobiologists, and the subsequent identification of the genes affected by these mutations can provide mechanistic information that is of great benefit for understanding many aspects of CNS development and function.

In spite of the utility of loss-of-function mutations for genetic analysis, invertebrate geneticists have long recognized that other genetic approaches must also be used if the full power of genetics is to be focused on a given biological issue. As elegantly argued by Miklos and Rubin (35): 'Much of the knowledge of developmental processes in the fly, worm, mouse, and zebrafish, and of the cell biology of yeast, has been obtained via loss-of-function perturbations... Nevertheless, the loss-of-function approach quickly reaches a pragmatic limit for several reasons... It is clear from such studies that future work will be driven increasingly by powerful transgenic technologies that allow finer and finer orchestrations of multiple developmental networks *in vivo*.' Although the use of transgenic approaches for genetic analysis in mammals has not been exploited frequently, the examples that have been reported suggest that this type of approach will be of great utility for studies of the CNS.

BAC-mediated gene dosage studies

The use of gene dosage experiments to study gene function is commonplace in yeast and invertebrates. For example, high copy number suppression screens have been useful for identifying additional genes involved in cell cycle traverse in yeast (36), and overexpression or misexpression screens are now being used to uncover functions for approximately two-thirds of the fly genes that have no readily observable loss-of-function phenotype (37,38). The first studies utilizing large genomic fragments for transgenic mice employed yeast artificial chromosomes (YACs) to demonstrate that YACs often carry all of the information necessary for recapitulation of the endogenous pattern of expression, and thus can be used to investigate complex transcriptional regulatory elements *in vivo* and to clone mutant genes by complementation. The use

of YAC transgenic mice to screen for functions by gene dosage analysis was first reported in studies using YACs covering the human Down syndrome locus to produce transgenic mice (39). These studies reported learning and memory defects in mice carrying an extra dose of the human *minibrain* gene, implicating it in the learning defects associated with this Down syndrome. This approach was extended for the identification of genes initially identified as quantitative trait loci (QTLs) in human genetic studies (40). In these instances, fine mapping of the genes conferring these phenotypes took advantage of the fragmentation of the YAC DNA during pronuclear injection, allowing the investigators to map relatively small sub-fragments of the YAC that could influence the phenotype.

The practical advantages of working with BACs versus YACs (41) has resulted in a shift in emphasis from YAC clones to BAC clones for positional cloning, for completion of the Human Genome Project, and for complementation cloning in transgenic mice. The first indication that BAC-mediated gene dosage analysis might be an effective method of screening for novel genetic functions came during the analysis of the murine *Clock* locus (42). In these studies, it was observed both that transgenic mice carrying a BAC from the *Clock* locus could rescue the long-period and loss-of-rhythm phenotypes of the original *Clock* mutations, and that it could shorten the circadian period on a wild-type background. In this case, complementary phenotypes were observed in the loss-of-function versus increased gene dosage situations, indicating that the *Clock* gene is both necessary and rate limiting for production of circadian rhythms.

The demonstration that BAC-mediated gene dosage analysis could uncover functions for genes that reveal no obvious phenotype in the null state came from studies of the mouse *Zipro1* gene (19). Initial studies of the expression of this gene in the developing CNS suggested a role in granule cell proliferation (42). Gene targeting to produce null alleles of *Zipro1* was disappointing, yielding no readily observed phenotype and indicating that this gene is functionally redundant in the murine genome. However, transgenic mice carrying multiple copies of a BAC carrying the *Zipro1* gene were observed to have higher numbers of proliferating granule cell precursors in the developing CNS, resulting in alterations in cerebellar morphogenesis. Observation of the pattern of co-expression of a marker gene inserted into this locus using the BAC modification strategy discussed above revealed that *Zipro1* is expressed in the skin. Further analysis demonstrated that *Zipro1* also controls the proliferation of precursors at this site. The fact that mice carrying the same number of copies of the BAC from which the *Zipro1* gene was deleted displayed none of these phenotypes definitively demonstrated that it is the *Zipro1* gene, and not other adjacent genes carried in the BAC, that plays a role in progenitor cell proliferation.

Several considerations indicate that this approach may yield important information for many genes. First, the average mouse gene is between 30 and 40 kb. Thus, BACs often contain all the information necessary for correct copy-number-dependent and position-independent transcription in transgenic mice (16). For most genes, this ensures that the phenotypes observed reflect increased dosage in the proper cell types *in vivo* and reproducible results in different transgenic lines carrying equivalent copy numbers. Secondly, the ability to construct precisely modified BACs by homologous recom-

ination in *E.coli* (16) allows the insertion of marker genes to confirm rapidly that transgene expression reflects the endogenous locus, and to prepare appropriate controls for dosage of other genes that might be carried on these large genomic fragments. Thirdly, although it has been demonstrated that duplication or triplication of most loci in *Drosophila melanogaster* does not result in an overt phenotype (42), genetic analysis using high-level expression via P-element insertion, heat shock promoters and the GAL4 UAS system has revealed relevant functions for many fly gene products (37,38). Similar results have been obtained in *Caenorhabditis elegans* (43). Thus, dosage experiments for the large family of worm G-protein-coupled receptors revealed functions for several members of this family that did not display an informative phenotype using loss-of-function analysis. Fourthly, the co-expression of a marker gene from the typical construct employed for gene dosage studies (Fig. 1) allows high-resolution analysis of precisely those cells most likely to express an anatomical phenotype. This can present a significant advantage over 'forward genetic' approaches in which evaluation of subtle anatomic phenotypes is often very difficult because the site of action of the mutation is not known. Based on these results, and given the small fraction of genes that yield an apparent loss-of-function phenotype, we believe that BAC-mediated gene dosage analysis can provide an important new tool for functional analysis of mammalian genes. We anticipate that the extension of these techniques to correctly express CRE recombinase in a large variety of CNS cell types, or to express 'knock down', dominant-negative or gain-of-function mutations via BAC-mediated transgenesis, will offer additional important avenues for genetic analysis in mice.

Animal models

The preparation of animal models for human CNS disease has been an extremely important development of the last decade. For dominant mutations, the use of transgenic techniques to model human disease has led to major advances in our understanding of pathogenic mechanisms (32). However, these studies have also highlighted the limitations of conventional transgenic methodology for the production of accurate animal models, and the difficulties associated with modeling human pathophysiology in mice. One major problem faced in the creation of animal models derives from the fact that many of the genes involved are widely expressed, yet result in a stereotyped and relatively cell-specific pathological progression. To understand the factors mediating this progression, or even to evaluate whether this progression can be studied in mice, it is important that the mutant gene product be expressed accurately. For some genes, the use of 'knock in' strategies to reproduce the human mutation has been an effective solution to this problem (42). However, in many cases, in particular for long-term neurodegenerative diseases such as the 'triplet repeat' diseases, a single copy of the mutant allele is not sufficient to reproduce the pathology one hopes to investigate during the relatively brief lifetime of a mouse. In these instances, increasing the dosage of the mutant gene can often accelerate the pathogenic process so that it can be studied in mice.

The facility and precision of the current methods for BAC modification (unpublished data), and the position-independent and dosage-dependent expression of most BAC transgenes, offer a powerful avenue for the production of accurate animal models for human disease. Although this methodology has not yet been extensively exploited for this purpose, the 'hit and run' nature of the BAC modification system presents an important advantage for this type of experiment because it allows serial modifications to the same BAC. One can, for example, introduce epitope tags and appropriate markers into the gene in the first round of modification, and produce different mutant alleles of the protein of interest in a second round of modification. Alternatively, one can produce mice with regulated production of the protein of interest using a single BAC transgene in which the transcriptional regulatory factor of interest replaces the endogenous gene in the first modification step, followed by insertion of the transcriptional response elements driving expression of the mutant protein into the BAC vector in the second step of modification. This strategy allows the production of mice that express the mutant gene product in a temporally controlled manner in only the cells of interest. For example, we have recently used this strategy to insert the rTTA transcriptional regulatory system (44) into the calbindin BAC so that we can express the mouse *Lurcher* gene in a temporally controlled fashion in cerebellar Purkinje cells (unpublished data) for further investigation of the mechanisms of cell death elicited by expression of GluR $\delta 2^{Lc}$.

CONCLUSIONS

In this review, I have focused on recently developed BAC transgenic technology that can accelerate the analysis of CNS-specific gene expression and function. These methods exploit the accurate transcriptional regulation of genes present on BACs in transgenic mice, and the ability to manipulate precisely these large genomic DNA constructs in *E. coli*, to develop a new strategy for investigation of CNS-expressed genes. This strategy presents an efficient alternative to traditional approaches for the analysis of CNS-specific gene expression and protein localization. It also represents an additional and very important tool for genetic analysis in mammals, and for the creation of animal models of human disease. It seems evident that the application of BAC transgenic methodology, and its combination with novel strategies for the functional analysis of genes, cells and neuronal circuits *in vivo*, will have an increasingly important impact on our understanding of the mammalian CNS.

ACKNOWLEDGEMENTS

N.H. is supported by the Howard Hughes Medical Institute. This work was also supported by NIH grants R01 39636-01, P01 NS 30532-08, Project V.

REFERENCES

- Selkoe, D.J. (1997) Alzheimer's disease: genotypes, phenotypes, and treatments. *Science*, **275**, 630–631.
- Bandmann, O., Marsden, C.D. and Wood, N.W. (1998) Genetic aspects of Parkinson's disease. *Mov. Disord.*, **13**, 203–211.
- Brin, M.F. and Koller, W. (1998) Epidemiology and genetics of essential tremor. *Mov. Disord.*, **13**(Suppl. 3), 55–63.
- Pulst, S.M. (1999) Genetic linkage analysis. *Arch. Neurol.*, **56**, 667–672.
- Crabbe, J.C., Wahlsten, D. and Dudek, B.C. (1999) Genetics of mouse behavior: interactions with laboratory environment. *Science*, **284**, 1670–1672.
- Ramon y Cajal, S. (1911) *Histologie du Systeme Nerveux de l'Homme et des Vertebres*. Maloine, Paris, France (reprinted by Consejo Superior de Investigaciones Cientificas, Madrid, Spain, 1955).
- Mullen, R.J., Hamre, K.M. and Goldowitz, D. (1997) Cerebellar mutant mice and chimeras revisited. *Perspect. Dev. Neurobiol.*, **5**, 43–55.
- Rossant, J., Spence, A. and Rossant, J. (1998) Chimeras and mosaics in mouse mutant analysis. *Trends Genet.*, **14**, 358–363.
- Weimann, J.M., Zhang, Y.A., Levin, M.E., Devine, W.P., Brulet, P. and McConnell, S.K. (1999) Cortical neurons require Otx1 for the refinement of exuberant axonal projections to subcortical targets. *Neuron*, **24**, 819–831.
- Kashiwabuchi, N., Ikeda, K., Araki, K., Hirano, T., Shibuki, K., Takayama, C., Inoue, Y., Kutsuwada, T., Yagi, T. and Kang, Y. (1995) Impairment of motor coordination, Purkinje cell synapse formation, and cerebellar long-term depression in GluR delta 2 mutant mice. *Cell*, **81**, 245–252.
- Patterson, M. and Handel, M. (eds) (1998) *Trends Guide to Bioinformatics. Trends Supplement*. Elsevier Science, UK, pp. 1–33.
- Burley, S.K., Almo, S.C., Bonanno, J.B., Capel, M., Chance, M.R., Gaasterland, T., Lin, D., Sali, A., Studier, F.W. and Swaminathan, S. (1999) Structural genomics: beyond the human genome project. *Nature Genet.*, **23**, 151–157.
- Sali, A. (1998) 100, 000 protein structures for the biologist. *Nature Struct. Biol.*, **5**, 1029–1032.
- Yang, X.W., Model, P. and Heintz, N. (1997) Homologous recombination based modification in *Escherichia coli* and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nature Biotechnol.*, **15**, 859–865.
- Schedl, A., Beermann, F., Thies, E., Montoliu, L., Kelsey, G. and Schutz, G. (1992) Transgenic mice generated by pronuclear injection of a yeast artificial chromosome. *Nucleic Acids Res.*, **20**, 3073–3077.
- Huxley, C. (1998) Exploring gene function: use of yeast artificial chromosome transgenesis. *Methods*, **14**, 199–210.
- Peterson, K. (1999) Use of yeast artificial chromosomes to express genes in transgenic mice. *Methods Enzymol.*, **303**, 186–203.
- Jarvik, J.W. and Telmer, C.A. (1998) Epitope tagging. *Annu. Rev. Genet.*, **32**, 601–618.
- Yang, X.W., Wynder, C., Doughty, M.L. and Heintz, N. (1999) BAC-mediated gene-dosage analysis reveals a role for Zipr1 (Ru49/Zfp38) in progenitor cell proliferation in cerebellum and skin. *Nature Genet.*, **22**, 327–335.
- Yu, W., Misulovin, Z., Suh, H., Hardy, R.R., Jankovic, M., Yannoutsos, N. and Nussenzweig, M.C. (1999) Coordinate regulation of RAG1 and RAG2 by cell type-specific DNA elements 5' of RAG2. *Science*, **285**, 1080–1084.
- Davies, M.V. and Kaufman, R.J. (1992) Internal translation initiation in the design of improved expression vectors. *Curr. Opin. Biotechnol.*, **3**, 512–517.
- Yoshihara, Y., Mizuno, T., Nakahira, M., Kawasaki, M., Watanabe, Y., Kagamiyama, H., Jishage, K., Ueda, O., Suzuki, H., Tabuchi, K. *et al.* (1999) A genetic approach to visualization of multisynaptic neural pathways using plant lectin transgene. *Neuron*, **22**, 33–41.
- Horowitz, L.F., Montmayeur, J.P., Echelard, Y. and Buck, L.B. (1999) A genetic approach to trace neural circuits. *Proc. Natl Acad. Sci. USA*, **96**, 3194–3199.
- Chatterjee-Kishore, I., Akker, F. and Stark, G.R. (2000) Association of STATs with relatives and friends. *Trends Cell. Biol.*, **10**, 106–111.
- Jones, C., Patel, A., Griffin, S., Martin, J., Young, P., O'Donnell, K., Silverman, C., Porter, T. and Chaiken, I. (1995) Current trends in molecular recognition and bioseparation. *J. Chromatogr. A*, **707**, 3–22.
- Fondell, J.D., Ge, H. and Roeder, R.G. (1996) Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc. Natl Acad. Sci. USA*, **93**, 8329–8333.
- Pavlou, O., Ehlenfeldt, R., Horn, S. and Orr, H.T. (1996) Isolation, characterization and *in vivo* analysis of the murine calbindin-D28K upstream regulatory region. *Brain Res. Mol. Brain Res.*, **36**, 268–279.
- Payne, C.M., Mullins, L.J. and Mullins, L.J. (1999) Manipulating large genomic clones via *in vivo* recombination in bacteria. *J. Hum. Hypertens.*, **13**, 845–848.
- Zuo, J., Treadway, J., Buckner, T.W. and Fritzel, B. (1999) Visualization of alpha9 acetylcholine receptor expression in hair cells of transgenic mice

- containing a modified bacterial artificial chromosome. *Proc. Natl Acad. Sci. USA*, **96**, 14100–14105.
30. Keverne, E.B. (1997) An evaluation of what the mouse knockout experiments are telling us about mammalian behaviour. *Bioessays*, **19**, 1091–1098.
 31. Mayford, M. and Kandel, ER. (1999) Genetic approaches to memory storage. *Trends Genet.*, **15**, 463–470.
 32. Heintz, N. and Zoghbi, H.Y. (2000) Insights from mouse models into the molecular basis of neurodegeneration. *Annu. Rev. Physiol.*, **62**, in press.
 33. Zambrowicz, B.P. and Friedrich, G.A. (1998) Comprehensive mammalian genetics: history and future prospects of gene trapping in the mouse. *Int. J. Dev. Biol.*, **42**, 1025–1036.
 34. Justice, M.J., Noveroske, J.K., Weber, J.S., Zheng, B. and Bradley, A. (1999) Mouse ENU mutagenesis. *Hum. Mol. Genet.*, **8**, 1955–1963.
 35. Miklos, G.L. and Rubin, G.M. (1996) The role of the genome project in determining gene function: insights from model organisms. *Cell*, **86**, 521–529.
 36. Levine, K., Tinkelenberg, A.H. and Cross, F. (1995) The CLN gene family: central regulators of cell cycle Start in budding yeast. *Prog. Cell Cycle Res.*, **1**, 101–114.
 37. Rorth, P. (1996) A modular misexpression screen in *Drosophila* detecting tissue-specific phenotypes. *Proc. Natl Acad. Sci. USA*, **93**, 12418–12422.
 38. Perrimon, N. (1998) New advances in *Drosophila* provide opportunities to study gene functions. *Proc. Natl Acad. Sci. USA*, **95**, 9716–9717.
 39. Smith, D.J., Stevens, M.E., Sudanagunta, S.P., Bronson, RT., Makhinson, M., Watabe, A.M., O'Dell, T.J., Fung, J., Weier, H.U., Cheng, J.F. and Rubin, E.M. (1997) Functional screening of 2 Mb of human chromosome 21q22.2 in transgenic mice implicates minibrain in learning defects associated with Down syndrome. *Nature Genet.*, **16**, 28–36.
 40. Symula, D.J., Frazer, K.A., Ueda, Y., Deneffe, P., Stevens, M.E., Wang, Z.E., Locksley, R. and Rubin, E.M. (1999) Functional screening of an asthma QTL in YAC transgenic mice. *Nature Genet.*, **23**, 241–244.
 41. Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl Acad. Sci. USA*, **89**, 8794–8797.
 42. Antoch, M.P., Song, E.J., Chang, A.M., Vitaterna, M.H., Zhao, Y., Wilsbacher, L.D., Sangoram, A.M., King D.P., Pinto, L.H. and Takahashi, J.S. (1997) Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. *Cell*, **89**, 655–667.
 43. Jansen, J. (1999) The complete family of genes encoding G proteins of *Caenorhabditis elegans*. *Nature Genet.*, **21**, 414–419.
 44. Lin, X., Cummings, C.J. and Zoghbi, H.Y. (1999) Expanding our understanding of polyglutamine diseases through mouse models. *Neuron*, **24**, 499–502.
 45. Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W. and Bujard, H. (1995) Transcriptional activation by tetracyclines in mammalian cells. *Science*, **268**, 1766–1769.

