BAC TO THE FUTURE: THE USE OF BAC TRANSGENIC MICE FOR NEUROSCIENCE RESEARCH

Nathaniel Heintz

The development of methods for engineering bacterial artificial chromosomes (BACs), and for the efficient production of BAC transgenic mice, has allowed the design of *in vivo* approaches to the analysis of gene expression and function in the brain, which could not be accomplished using traditional methods. These strategies have shed light on the functions of single genes in the nervous system, and will accelerate the use of functional genomic approaches to neuroscience research.

EXPRESSED SEQUENCE TAGS
Short (200–500 base pairs) DNA sequences that represent the sequences expressed in an organism under a given condition. They are generated from the 3′- and 5′-ends of randomly selected complementary DNA clones. The purpose of EST sequencing is to scan for all the protein-coding genes, and to provide a tag for each gene on the genome.

REPORTER GENE
A gene that encodes an easily assayed product. It is coupled to the upstream sequence of another gene, and can then be transfected into cells to identify factors that activate response elements in the upstream region of the gene of interest.

Howard Hughes Medical Institute, Laboratory of Molecular Biology, The Rockefeller University, New York 10021, USA.
e-mail: heintz@mail.rockefeller.edu

The near completion of the human genome project, and the continuing identification of transcripts in the context of high-throughput EXPRESSED SEQUENCE TAGS (EST) sequencing efforts, have revealed the presence of at least 30,000 genes in the mammalian genome[1,2]. Most of these genes are expressed in the central nervous system (CNS), often producing multiple transcripts that encode distinct protein products. A significant fraction of the genome (>5,000 genes) is preferentially or exclusively transcribed in the nervous system. The functions of several hundreds of these gene products are known, and it is possible to guess the biochemical nature of many more of these products by extrapolating from studies in invertebrates or by homology-based sequence comparisons. However, it is difficult or impossible in most cases to predict *in vivo* functions on the basis of information from other well-studied family members. For these reasons, post-genomic studies of CNS development and function can be advanced by new methods that exploit genomic information to contribute to our knowledge of the molecular mechanisms that operate in the intact organism.

An understanding of the functions of an individual CNS-expressed gene requires detailed information on its pattern of expression, on the localization and biochemical activities of its product and the partners it interacts with, on the phenotypic consequences of altering its functions, and on the transcription profile, the properties and the connectivity of the cell types that express it. The ability to collect and interpret this diverse information would be advanced by the availability of a simple and efficient histological method for the analysis of gene expression and function. Here I will focus on the use of bacterial artificial chromosome (BAC) transgenic mice for this purpose. The precise and efficient engineering of BACs can facilitate a variety of studies that cannot be readily accomplished using conventional approaches. I will discuss the molecular methods that are used to manipulate these large DNA constructs, the usefulness of REPORTER GENES and targeted-expression studies to address fundamental problems in molecular genetic studies of the mammalian CNS, and potential applications of these methods to other areas of neuroscience research.

Methods for manipulating BACs

The use of reporter genes and targeted-expression studies has been crucial in the analysis of invertebrate gene expression and function[3]. However, this approach has been limited in mammals by the intrinsic difficulty of identifying key regulatory elements that could direct accurate expression *in vivo*. The original impetus for developing methods to precisely modify BACs came from the desire to solve this problem so that transgenic approaches could be used effectively to investigate mammalian CNS gene expression and function. The proper expression of large DNA constructs (bacterial, yeast and phage artificial chromosomes; BOX 1) in transgenic mammals was shown by their ability to rescue mutant phenotypes after reintroduction into the

Box 1 | **YACs, BACs and PACs**

The use of large DNA constructs was first developed to aid in the genetic analysis of chromosome function in yeast. Vectors that allow the propagation of large exogenous DNA fragments as linear chromosomes in yeast were then used to construct libraries of human genomic DNA comprised of pools of yeast cells, each carrying a single, linear, artificial chromosome (YAC) with up to several megabases of foreign genomic DNA[86]. YAC libraries were crucial in early studies of the mammalian genome, because their large carrying capacity allowed the entire genome to be archived and screened for analysis of specific genetic or chromosomal intervals. The increased simplicity of molecular and genetic manipulation in bacteria led several years later to the development of vectors based on the F EPISOME for propagation of large DNA segments in *Escherichia coli*[87]. The F-factor-based bacterial artificial chromosomes (BACs) are propagated at low copy number in a recombination-deficient strain of *E. coli*, and they can carry up to several hundred kilobases of foreign DNA. The increased stability, the simplicity of DNA preparation, and the improved capability of BACs for high-throughput applications have led to their use as the preferred large-insert cloning system for genomic analysis. They have, for example, served as the principal substrates for the public effort to physically map and sequence several genomes. Phage artificial chromosomes (PACs) are large-insert clones that are very similar to BACs, but based on vectors derived from phage P1 (REF. 88). They are also propagated in *E. coli*, and have been used for several genomic studies. The carrying capacity of PACs is several hundred kilobases. Both BACs and PACs replicate as closed circular, supercoiled DNA, and purified DNA is prepared using protocols similar to those used for conventional plasmid DNAs.

genome[4]. This strongly indicated that the use of markers to report detailed expression information, and the use of gene dosage or dominant alleles to provide functional data, could be achieved using BAC constructs if a precise method were developed to manipulate these large DNAs. BACs were chosen because they have served as the primary source of archived genomic DNA for a variety of genome mapping and sequencing projects.

The first strategy[5] that was developed for the manipulation of BAC DNA took advantage of the precision of HOMOLOGOUS RECOMBINATION in *Escherichia coli*, which had been used extensively for marker insertion into, and excision from, the bacterial genome[6]. This method[5], and all subsequent methods that have been reported[7–10], rely on three basic features (FIG. 1). First, competence for homologous recombination is restored to the BAC host strain by reintroduction of the *E. coli recA* gene, or by the introduction of another enzyme that can restore to the BAC host the ability to perform homologous recombination. Second, a shuttle vector (or DNA fragment) that carries the desired reporter gene or modification cassette, flanked by sequences homologous to the genomic DNA carried in the BAC, is used to target the modification cassette into a precise site on the genomic DNA insert. Appropriate recombinants are selected and screened for precise co-integration into the targeted site on the BAC DNA. Third, unwanted vector sequences are resolved through a second homologous recombination event, or excised from the co-integrants using an appropriate site-specific RECOMBINASE. Negative selection is used to enrich the desired end-product — a BAC that carries the modification cassette inserted into the exact position chosen in the design of the experiment.

The method for BAC manipulation that has been used most frequently in the construction of BAC transgenic mice is the original system described by Yang *et al.*[5] As mentioned above, several other protocols have been reported for BAC modification that take advantage of homologous recombination in *E. coli* to target into the correct site in the BAC, and remove vector sequences using a second recombination step for resolution or excision[7–10]. Although it is beyond the scope of this review to discuss the technical aspects of these and other protocols[11], one point is worth mentioning. If the targeted BAC is to be used in transgenic mice to specifically express large marker genes, recombination enzymes, toxins, tract-tracing proteins, DOMINANT-ACTIVATING or DOMINANT-NEGATIVE alleles, or multiple combinations of these elements using INTERNAL RIBOSOME ENTRY SITES (IRES), then it is advisable to avoid polymerase chain reaction (PCR) amplification of these segments in assembling the modification cassette. This is due to the fact that the transfer of large markers into the BAC as PCR-amplified products introduces the uncertainty that a nucleotide change during the PCR amplification could be carried through to the final modified BAC. For example, sequence analysis of six marker cassettes that were introduced into BACs using PCR-amplified linear DNA resulted in the identification of mutations in three of these clones; further analysis showed that these mutations were introduced during the amplification reaction, leading the authors to conclude that it is necessary to verify the sequence to ensure that each manipulated BAC carries a wild-type marker cassette[10]. This problem is avoided if the marker cassette is carried on a shuttle vector, as the frequency of mutation in *E. coli* is extremely low (1×10^{-6}). Recently, a highly efficient and reproducible BAC modification system that is suitable for high-throughput studies has been developed[12]. This system makes use of improved shuttle vectors, improved selection strategies and 'built-in' resolution cassettes to facilitate the procedures required for BAC modification. It has been used to insert reporter genes successfully into >100 different BACs that carry CNS-expressed genes (GENSAT BAC Transgenic Project, National Institute of Neurologic Disease and Stroke). In this method, preparation of the specific targeting constructs for each gene does not require PCR amplification of the modification cassette to be inserted into the genomic locus on the BAC. So, this system does not require sequence verification of the individually manipulated BACs before their use *in vivo*.

During the past year, the ability to design transgenic experiments using BACs, with high probabilities of success, has greatly improved owing to the data available from the human and mouse genome projects. To achieve proper expression of a BAC transgene, a BAC should be chosen with the gene centred on its insert, and ~50 kb of 5′- and 3′-flanking DNA. As the carrying capacity of BACs is several hundred kilobases, and an average mammalian gene is <50 kb, the information available from the different genome projects allows the selection of a BAC with these characteristics for most genes. However, some mammalian transcription units reach lengths that exceed the carrying capacity of BACs. In these cases, one could turn to yeast artificial chromosomes (YACs), as they can include up to several megabases of genomic DNA[13]. Alternatively, the appropriate

EPISOME
A genetic unit of replication that can exist either extrachromosomally or integrated into the bacterial chromosome.

HOMOLOGOUS RECOMBINATION
The substitution of a segment of DNA with one that is identical or almost identical to it. It occurs naturally during meiosis, but can also be used experimentally for gene targeting to modify the sequence of a gene.

RECOMBINASE
An enzyme that recognizes specific DNA sequences and catalyses the reciprocal exchange of DNA strands between these sites during viral integration, chromosomal segregation and related processes.

DOMINANT ACTIVATING
Describes a mutant molecule that is capable of forming a heteromeric complex with the normal molecule, generating a constitutively active protein.

DOMINANT NEGATIVE
Describes a mutant molecule that is capable of forming a heteromeric complex with the normal molecule, knocking out the activity of the entire complex.

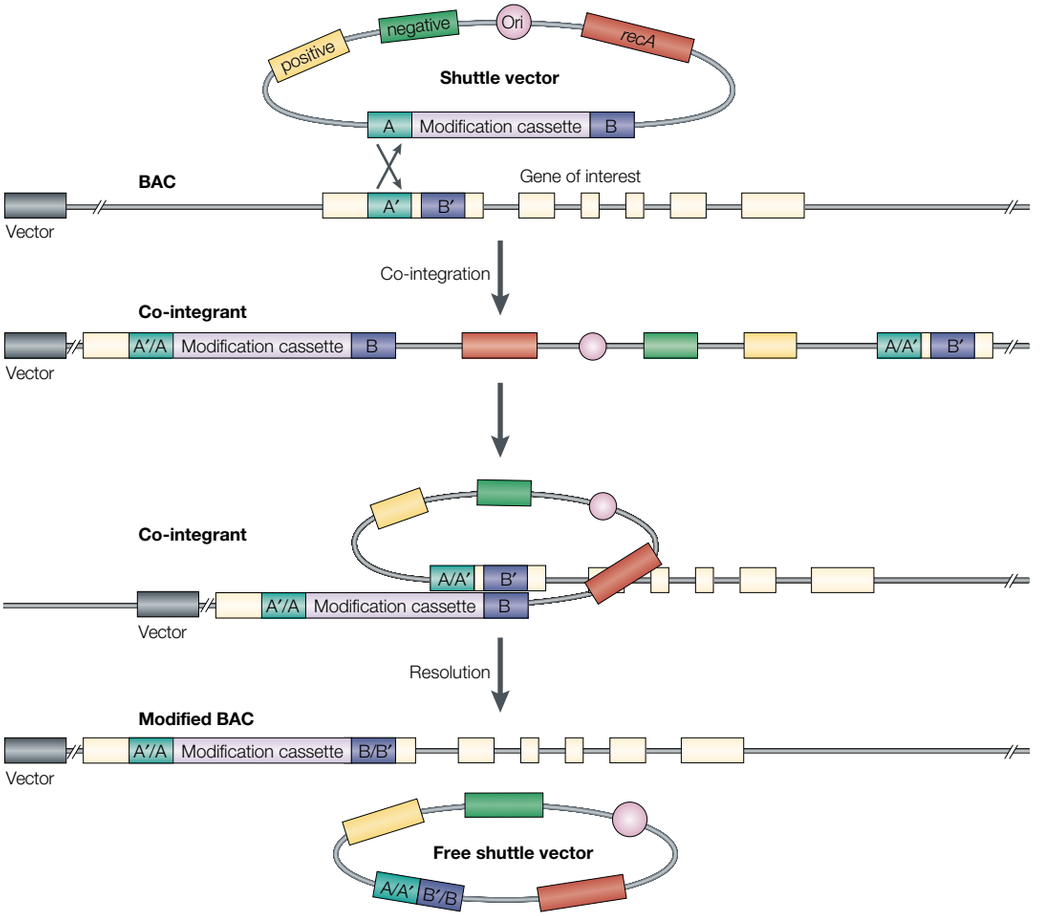

Figure 1 | **BAC manipulation by homologous recombination.** Features of the shuttle vector include: *recA* (or another gene that encodes a recombination enzyme that can restore host competence for homologous recombination); Ori, a conditional origin of replication; a positive selectable marker; a negative selectable marker; A and B homology arms that are identical to sequences A′ and B′ in the bacterial artificial chromosome (BAC); and a modification cassette that carries the functional elements to be inserted into the BAC (see FIG. 2). The first step of the process is referred to as co-integration. It occurs by crossing over between one of the homology arms and the BAC. It is selected for by the positive selectable markers that are carried on the shuttle vector and the BAC. In this case, recombination between A and A′ results in incorporation of the modification cassette and vector sequences into the BAC to yield the co-integrate shown immediately below the first arrow. The second step of the process is referred to as resolution. It occurs by a second homologous recombination event that occurs within the co-integrate. In this case, looping out of the shuttle vector sequences and recombination between B and B′ yields the precisely modified BAC, with the modification cassette inserted at the correct position in the BAC, and free shuttle vector. This event is selected for by the negative selectable marker and the conditional Ori that are carried on the shuttle vector. By correct design of the A and B homology arms, this generic strategy can be used to produce insertions, deletions and point mutations.

regulatory information for correct expression of many of these very large genes might be restricted to the 5′-flanking genomic DNA and the first few introns. In this case, BAC constructs carrying the 5′-end of the gene, and extensive 5′-flanking DNA, could allow accurate transcription of the transgene. My group is now testing this possibility for several large mammalian genes.

Transcription map of CNS-expressed genes

An important first step in investigations of gene function in the CNS is to map expression throughout development. The thousands of cell types in the CNS are generated on a relatively precise schedule, undergo a defined series of movements and steps in their differentiation, project their axons to stereotyped positions for integration into appropriate circuitry, and express a well-defined final morphology that is indicative of their function. Just as these properties led, by detailed histological analysis and comparative anatomy, to the formulation of revolutionary hypotheses on the development and functions of the CNS[14], they also allow us to make inferences about the functions of specific gene products. So, knowledge of exactly which cells express a given gene, what those cells are doing at the time that the gene is expressed, and which signals elicit expression of that gene in normal or abnormal circumstances, can often be very informative in assessing gene function or interpreting CNS phenotypes. This is particularly true if the

INTERNAL RIBOSOME ENTRY SITE
A sequence that is inserted between the coding regions of two proteins, and allows efficient assembly of the ribosome complex in the middle of a transcript, leading to translation of the second protein.

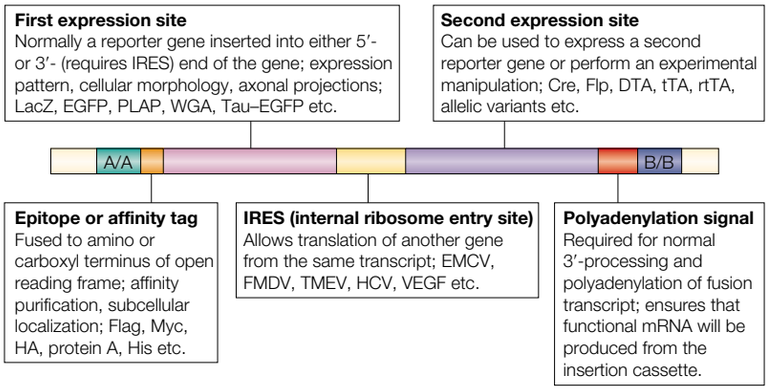

Figure 2 | **A modification cassette.** There are many functional elements that could be included in a modification cassette. Typical elements of a modification cassette are illustrated in this figure. A great deal of flexibility in the design of the cassette can be achieved by the incorporation of internal ribosome entry sites (IRES) into the cassette to allow the expression of multiple proteins from the fusion transcript. As indicated, functional elements that might be expressed from the modification cassette include a variety of reporter genes (encoding LacZ, enhanced green fluorescent protein (EGFP), placental alkaline phosphatase (PLAP), wheat-germ agglutinin (WGA), Tau–EGFP and so on), recombination enzymes (Cre recombinase, Flp recombinase), regulatory components (tetracycline-dependent transactivator protein (tTA), reverse tetracycline-dependent transactivator protein (rtTA) and so on), toxins such as DTA (diphtheria toxin attenuated), allelic variants (dominant-negative or activating mutations), or epitope and affinity tags (Flag, Myc, haemagglutinin (HA), protein A, His and so on). EMCV, encephalomyocarditis virus; FMDV, foot-and-mouth disease virus; HCV, hepatitis C virus; TMEV, Theiler's murine encephalomyelitis virus; VEGF, vascular endothelial cell growth factor.

sequence of its encoded product is informative from a biochemical point of view. For example, the roles of HOX GENES in the developing hindbrain were first proposed on the basis of the functions of Hox clusters in segmentation of the *Drosophila* embryo, and the ordered expression of individual genes in the cluster during development of the vertebrate hindbrain. Interpretation of developmental phenotypes arising from null mutations of these genes has been heavily dependent on detailed expression analysis of the targeted gene and other family members in the developing RHOMBOMERES of normal and mutant mice[15–19]. Similarly, the identity of vertebrate odorant receptors was first based solely on the discovery of a very large family of genes that encode proteins containing seven transmembrane domains, the expression of which was restricted to the olfactory epithelium[20]. It was only later that *in vitro* studies of these putative receptors provided definitive proof that these molecules can respond to odorants and transduce signals[21]. These and many other examples illustrate the importance of precise expression information in the formulation of hypotheses about gene function and in the interpretation of phenotypic information.

The use of BAC transgenic mice to map precisely the cell types that express a gene of interest is based on two simple facts. First, in most cases, large genomic DNA fragments (>100 kb) are expressed independently of the site of integration into the genome of transgenic mice[4]. Second, inclusion of EPITOPE TAGS and marker proteins into endogenous loci of invertebrate and vertebrate genes has not generally altered their patterns of expression[3,22]. As a consequence, the expression of a reporter gene from a modified BAC in transgenic mice can be used to determine the profile of cell types that express that gene throughout CNS development. A typical modification cassette might resemble that shown in FIG. 2. An important point to be made regarding this type of construct is that the use of IRES[22] can allow the creation of POLYCISTRONIC MESSENGER RNAS, from which several proteins can be synthesized under the control of the endogenous transcriptional regulatory sequences. This is tremendously advantageous, because it allows the co-expression of multiple reporter genes (encoding enhanced green fluorescent protein (EGFP), LacZ, placental alkaline phosphatase, wheat-germ agglutinin and so on) to visualize cells that express the gene or combinations of reporter genes and other useful proteins (CRE RECOMBINASE, FLP RECOMBINASE, TETRACYCLINE-DEPENDENT TRANSACTIVATOR, REVERSE TETRACYCLINE-DEPENDENT TRANSACTIVATOR and so on), to mark cells and to perform an experimental manipulation simultaneously.

In the experience of my group, expression of EGFP using this strategy in transgenic mice can allow the visualization of cellular morphology in the CNS; the result is comparable to that obtained by immunohistochemical or immunofluorescence detection. Given the large number of different cell types that are present in the brain, and the classical definition of CNS cell types on the basis of morphology[14], the ability to identify cells that express a gene of interest is enhanced by the use of reporter genes, because it is possible to see the detailed morphology of the expressing cells. For example, in our initial studies of brain lipid-binding protein (BLBP) expression in the developing brain, it was not possible to identify correctly, on the basis of *in situ* hybridization studies, the cell types that express the gene, even though robust signals were obtained using colorimetric protocols[24]. On the preparation of high-titre antisera to BLBP, followed by immunofluorescence detection, it became immediately clear that this gene is transiently expressed in many glial cell types during their differentiation in the developing brain[25]. In contrast to the *in situ* hybridization data, direct visualization of an EGFP reporter gene in tissue from transgenic mice also revealed details of cellular morphology, allowing the correct identification of BLBP-expressing cell types (FIG. 3).

Several observations show that the expression of large DNA transgenes can accurately reflect the pattern of transcription of the endogenous chromosomal gene. The initial finding that large genomic DNA constructs could reproduce the *in vivo* regulation of a gene came from rescue experiments using YACs[26,27]. These and other YAC complementation experiments clearly showed that large DNA transgenes can carry all of the required information for correct *in vivo* expression[4,28]. Studies using BAC constructs to complement mouse mutant phenotypes followed[29–33], confirming the results obtained with YACs, and further supporting the observation that DNA constructs that carry >100 kb of genomic DNA can often carry all of the coding and regulatory sequences required for normal gene function. Furthermore, the use of human YAC constructs allowed direct assays of the expression of the human transcripts in YAC transgenic mice, leading to the demonstration

HOX GENES
Transcription factors expressed in specific patterns that are important for determining regional identity along the anteroposterior axis of the embryo. They are also known as homeobox genes.

RHOMBOMERES
Neuroepithelial segments found in the embryonic hindbrain that adopt distinct molecular and cellular properties, restrictions in cell mixing, and ordered domains of gene expression.

EPITOPE TAG
The immunological determinant of an antigen, which has been fused to a protein of interest for its subsequent localization with specific antibodies.

POLYCISTRONIC MESSENGER RNA
A messenger RNA that codes for more than one protein.

CRE RECOMBINASE SYSTEM
A method in which the Cre enzyme catalyses recombination between *loxP* sequences. If the *loxP* sequences are arranged as a direct repeat, recombination will delete the DNA between the sites.

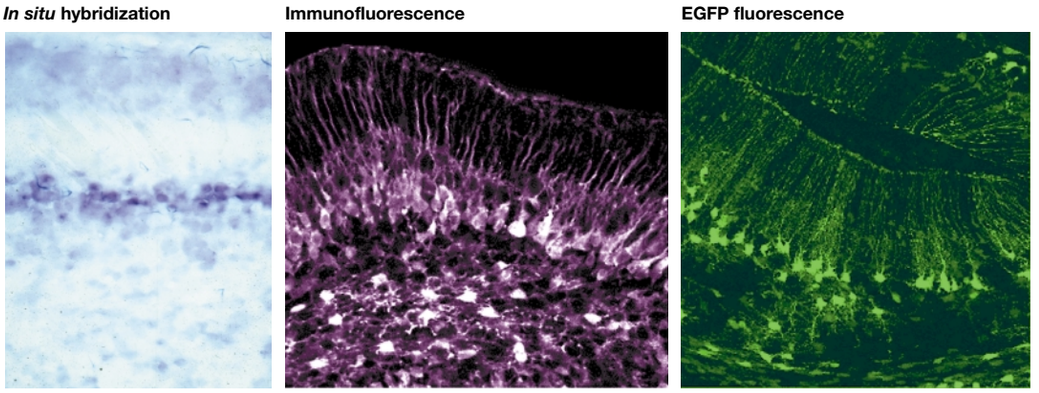

Figure 3 | **Visualization of brain lipid-binding protein expression in postnatal cerebellum.** Colorimetric *in situ* hybridization detection of brain lipid-binding protein (BLBP) messenger RNA; immunofluorescence detection of BLBP protein using monospecific polyclonal antiserum raised against recombinant BLBP; and enhanced green fluorescent protein (EGFP) detection of BLBP-expressing cells from transgenic mice. The detailed morphology that is evident in the second and third panels allows the identification of the BLBP-expressing cells as Bergmann glia and white matter astrocytes (unpublished data provided by T. Anthony).

FLP RECOMBINASE
A protein involved in the amplification of the yeast 2-μm plasmid. It encodes a protein that catalyses site-specific recombination between sites called Flp recognition targets (FRT). The Flp/FRT system has been successfully applied as a site-specific recombination system.

TETRACYCLINE-DEPENDENT TRANSACTIVATOR SYSTEM
A system that allows the precise control of gene expression in eukaryotic systems through the administration of tetracycline. It is based on two key elements: the tetracycline-dependent transactivator protein (tTA) and the target gene under the control of a tTA-responsive element. When these elements are transfected into eukaryotic cells, the tTA binds to the tTA-responsive element to initiate transcription. Tetracycline can then be administered to stop expression of the target gene.

REVERSE TETRACYCLINE-DEPENDENT TRANSACTIVATOR SYSTEM
A system that allows the precise control of gene expression in eukaryotic systems through the administration of tetracycline. It is based on two key elements: a mutant form of the tetracycline-dependent transactivator protein (tTA), and the target gene under the control of a tTA-responsive element. Once these key elements have been transfected into eukaryotic cells, the mutant tTA is expressed, but does not bind the tTA-responsive element. Binding of the mutant tTA to the tTA-responsive element and initiation of transcription is then induced by the addition of tetracycline.

POSTSYNAPTIC DENSITY
An electron-dense thickening underneath the postsynaptic membrane at excitatory synapses that contains receptors, structural proteins linked to the actin cytoskeleton and signalling machinery, such as protein kinases and phosphatases.

that several human genes are appropriately expressed from these large DNA constructs *in vivo*[34–36]. The ability to perform precise modifications of BACs then led to several further studies that have confirmed the accurate expression of reporter genes in BAC transgenic mice; these studies include analyses of zinc finger proliferation 1 (*Zipro1*; REFS 5,37), nicotinic cholinergic receptor α9 (*Acra9*; REF. 38), recombination-activating genes *Rag1* and *Rag2* (REF. 39), renin 1 (*Ren1*; REF. 31), neuropeptide Y (*Npy*; REF. 40), myogenic factors *Mrf4* and *Myf5* (REF. 41), and neuronatin (*Nnat*; REF. 42). Data for many other BAC transgenes support this observation[12].

Epitope and affinity tags: protein function

Individual neurons can make hundreds of thousands of synaptic contacts, and can extend axons over large distances to reach their targets. As a result of this morphological complexity, neurons require compartmentalization of biochemical functions that might not occur in other cell types. For example, recent studies of Ca^{2+} dynamics in individual dendritic spines have shown that these structures are 'biochemical microcompartments' that are isolated from the dendritic shaft[43]. Even functions as fundamental as protein synthesis can be handled differently in these complex cells, as shown by the dendritic localization of specific mRNAs[44] and by the local synthesis of their cognate proteins[45,46]. As our knowledge of the biochemical functions of highly specialized structures present in CNS cell types has increased, the ability to form a working hypothesis about gene function from this information has improved. For example, the localization of a given gene product to synaptic vesicles[47–49], to the POSTSYNAPTIC DENSITY[50–52], or to the mitochondrial membrane[53], has provided insights into the possible roles of specific proteins in the CNS, which were not evident from their expression profiles or sequence. At the very least, analysis of the signalling mechanisms that operate in neurons, and interpretation of subtle CNS phenotypes, will require highly precise information about the specific subcellular domains in which their components are found, and how these domains limit the activities of transduction events within the cell.

Subcellular localization studies have traditionally made use of monospecific antibodies to detect proteins *in situ*. The generation of monospecific antibody probes for *in situ* detection is expensive and time consuming, and might require several attempts to obtain useful reagents. To overcome these difficulties, and provide a simple, inexpensive and efficient alternative to the generation of monospecific antibodies for large numbers of proteins, studies in invertebrate and cell-culture systems have often made use of epitope tags[22]. This strategy relies on the construction and expression of recombinant fusion proteins that carry small peptide (epitope) tags, which are fused to the wild-type protein. These tags can then be recognized by highly specific antibodies. Expression of the tagged fusion protein *in vivo* or *in situ* can then be achieved using a previously characterized antibody that is specific for the chosen epitope tag[22]. The usefulness of this approach for mammalian neurons has been nicely documented in studies of protein localization in organotypic slice cultures[54,55]. However, epitope tags have not yet been widely used for *in vivo* studies of mammalian CNS-expressed proteins, because it has been difficult to express tagged proteins in the appropriate *in vivo* patterns. The ability to precisely manipulate BACs enables simple construction of epitope-tagged fusion proteins from genomic information for most genes, and accurate expression of the tagged proteins in transgenic mice. Using these methods, it will be possible to determine the subcellular distribution of new gene products precisely and efficiently.

The ability to target the expression of recombinant fusion proteins accurately to appropriate CNS cell types can also accelerate the characterization of protein complexes with important roles in CNS function. Recently developed approaches to the identification of protein–protein interactions that are relevant for the function of a given gene product are frustrated by the complexity of the nervous system. So, even robust results obtained in

YEAST INTERACTION SCREEN
A system used to determine the existence of direct interactions between proteins. It commonly involves the use of plasmids that encode two hybrid proteins: one that is fused to the GAL4 DNA-binding domain, and one that is fused to the GAL4 activation domain. The two proteins are expressed together in yeast; if they interact, the resulting complex will drive the expression of a reporter gene, commonly β-galactosidase.

MASS SPECTROMETRY
In mass spectrometry, a substance is bombarded with an electron beam of sufficient energy to fragment the molecule. The cations that are produced are accelerated in a vacuum through a magnetic field, and sorted on the basis of mass-to-charge ratio. The ratio is roughly equivalent to the molecular weight of the fragment.

POSITIONAL CLONING
A strategy for cloning on the basis of location in the genome, rather than the function of the product. It commonly involves linking the locus of interest to one that has already been mapped.

SUPPRESSION SCREEN
A system used to identify genes that, when overexpressed, lead to the suppression of a mutant phenotype. By contrast, an overexpression screen is used to identify genes that, when overexpressed, lead to the appearance of a mutant phenotype, and a misexpression screen is used to identify genes that, when expressed in ectopic regions, lead to the appearance of a mutant phenotype.

YEAST INTERACTION SCREENS[56] must be followed by extensive characterization of candidate molecules to ensure that they are expressed at the correct time and in the appropriate cell type *in vivo*. Moreover, it must be established that the proposed interactions do not represent adventitious associations between members of protein families that have the capacity to associate, but do not normally do so. Furthermore, these methods are optimized for the detection of binary protein–protein interactions, and their use in characterizing novel interactions between membrane proteins is limited. As an alternative, affinity methods for purification and characterization of multicomponent protein complexes have been developed for use in cultured cells, allowing a comprehensive determination of the proteins that comprise even highly complex intracellular machines[57]. The use of standardized affinity matrices for this type of biochemical study, and the development of sensitive methods for the identification of extremely small amounts of purified proteins using MASS SPECTROMETRY[58], indicates that the application of this strategy to the characterization of complex protein assemblies in the CNS will represent an important advance in neuroscience research. The ability to introduce affinity tags into proteins of interest, and to target their expression to the appropriate cell types *in vivo* using BAC transgenic mice, presents the opportunity to develop these methods for the characterization of protein assemblies present in even very rare CNS cell types.

BACs and phenotypic analysis of the CNS

During the past decade, the use of genetics as a tool to understand the development, function and dysfunction of the mammalian CNS has become central to neuroscience research. Many mutations that affect the mammalian brain have been identified by POSITIONAL CLONING, or prepared using homologous recombination. These studies have led to fundamental insights into many aspects of CNS development and function. In spite of the usefulness of loss-of-function mutations for genetic analysis, geneticists studying invertebrate models 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[3]. For example, the use of gene dosage, ectopic expression, and dominant-negative or activating alleles, to study gene function is commonplace in yeast and in invertebrates. High-copy-number SUPPRESSION SCREENS have been useful in identifying further genes that are involved in the cell cycle in yeast[59], and over- or misexpression screens are now being used to uncover functions for the approximately two-thirds of the *Drosophila* genes for which there is no readily observable loss-of-function phenotype[60,61]. The ability to target expression accurately to specific cell types in the CNS, and to obtain reproducible, copy-number-dependent and position-independent expression of transgenic constructs, has been essential for the extension of these and other genetic techniques to mice.

It has now been shown in many studies that gene dosage can provide an important complement to more traditional gene-targeting studies in mice. For genomic regions that are known to be sensitive to dosage effects, very mild overexpression can lead to important functional inferences. For example, YAC and BAC transgenic experiments covering the minimal Down's syndrome (DS) locus have resulted in CNS phenotypes in mice carrying extra copies of the *minibrain* gene. These phenotypes are similar to those present in partial trisomy 16 mouse models of DS, implicating the *minibrain* gene in learning defects associated with this DS[62]. Similarly, complementation studies using BAC transgenes that include the mouse *Clock* locus could rescue the long-period and loss-of-rhythm phenotypes of the original *Clock* mutations, leading to the identification of the first mammalian gene involved in circadian rhythmicity[63]. However, when carried against a wild-type background, it was noted that these BAC transgenes could also shorten the normal circadian period[64], showing that the activity of the *Clock* gene is also rate limiting *in vivo*. As this region of the mammalian genome was not known to be dosage sensitive, these results indicated that increased gene dosage could be generally applied to discover roles for other genes, the activity of which might be rate limiting *in vivo*.

The idea that informative phenotypes might be uncovered by gene-dosage experiments, even in cases in which the gene of interest was known not to be rate limiting, has been tested in studies of the mouse *Zipro1* gene[27]. Initial studies of the expression of this gene in the developing CNS indicated a possible role in cerebellar granule cell proliferation[24]. Gene-targeting studies to produce null alleles of *Zipro1* were disappointing, yielding no readily observed phenotype, and indicating that this gene was functionally redundant in the murine genome. However, transgenic mice that carried multiple copies of a BAC carrying the *Zipro1* gene were observed to have more proliferating granule cell precursors in the developing CNS than wild-type mice, resulting in alterations in cerebellar morphogenesis. Enhanced proliferation was not observed in transgenic mice that carried control BACs in which expression of the *Zipro1* gene was replaced by reporter gene expression, showing that the observed proliferative phenotype was a direct consequence of increased *Zipro1* expression. Further analysis showed that *Zipro1* can also contribute to the control of the proliferation of precursors in the skin. Subsequent studies have confirmed that Zipro1 is expressed in proliferating precursors in both of these sites, and that it can directly activate the transcription of several genes that are known be important in precursor cell proliferation *in vivo* (C. Wynder, unpublished observations).

Several considerations indicate that this approach might yield important information for many genes. First, as BAC transgenes are reproducibly expressed in a copy-number-dependent and position-independent manner, and because they can be easily and precisely manipulated, it is possible to design appropriate controls that are necessary for interpretation of phenotypic effects of increased BAC dosage. Second, the introduction of reporter genes into the BAC construct allows the rapid identification and detailed analysis of the cell types that overexpress the gene of interest, significantly facilitating phenotypic analysis. Third,

although it has been shown that duplication or triplication of most loci in *Drosophila* does not result in an overt phenotype[64], genetic analysis using high-level expression through P-ELEMENT insertion, HEAT-SHOCK PROMOTERS and the GAL4 UAS SYSTEM has revealed relevant functions for many fly gene products[60,61]. Similar results have been obtained in *Caenorhabditis elegans*[65]. So, dosage experiments for the large family of worm G-protein-coupled receptors have revealed functions of several members of this family that did not have an informative phenotype in loss-of-function analyses.

The usefulness of transgenic approaches for the analysis of dominant gain-of-function mutants has been repeatedly shown in studies of mouse models of human neurological disorders[66]. For example, conventional transgenic studies of the pathogenic mechanisms operating in Purkinje cells in spinocerebellar ataxia 1 have led to fundamental advances in our understanding of POLYGLUTAMINE DISORDERS[67,68]. In many instances, however, an understanding of the full spectrum of the mechanisms that contribute to disease progression will require the preparation of animal models that express the mutant alleles in the same pattern as that seen in the human disease. The use of KNOCK-IN strategies to reproduce the human mutation has been an effective solution to this problem[69]. In some cases, a single copy of the mutant allele might not be sufficient to reproduce the sought-after human pathogenic mechanisms during the relatively brief lifetime of a mouse. In these instances, large DNA transgenic studies can be very valuable in developing mouse models. For example, Hodgson *et al.*[70] used YACs that contained the entire human huntingtin locus, engineered to express full-length alleles carrying either 46 or 72 glutamine repeats, and compared these mice with animals that overexpressed the normal huntingtin gene (18 repeats). Mice that carried the mutant alleles developed behavioural, electrophysiological and histological phenotypes that more closely resembled Huntington's disease, including selective striatal neurodegeneration. Results from YAC transgenic mice that expressed familial Alzheimer's disease alleles, alone or in combination with mutant presenilin 1 YAC transgenes, also led to the conclusion that large DNA transgenes can provide insights into pathogenesis that are not evident from conventional transgenic studies[71]. Given the improved facility with which these studies can be carried out using BACs, it is evident that the production and analysis of mice that carry polymorphisms or pathological alleles, or dominant mutations based solely on genetic data from invertebrate, cell-culture or structural studies, will be increasingly important in genetic studies of molecular mechanisms of CNS function.

Cell-specific studies of genes, cells and circuits

The ability to target gene expression accurately to specific cell types in the CNS using BAC transgenic mice can facilitate a variety of experimental approaches that are not directly related to the analysis of the functions of a single gene. The most obvious of these include the enhanced efficiency with which one can pursue several techniques that have now become commonplace in mammalian genetic analysis. For example, the preparation of cell-specific knockout mice[72], and the production of regulated systems for expression *in vivo*[73], can be facilitated using modified BACs to express the appropriate proteins in cells of interest. So far, these experiments have relied heavily on knock-in strategies, or on the use of a few specific promoters that can be accurately expressed in transgenic mice. The use of BACs to drive expression of the recombination proteins or regulatory components of these systems considerably expands the range of cell types that can be accessed, and improves the efficiency with which these methods can be applied.

Another important application of BAC transgenic technology is the use of reporter genes to provide experimental access to purified CNS cell populations. For example, analysis of gene-expression profiles from the normal or the diseased brain is complicated by the fact that even the most precisely microdissected CNS tissue samples contain many cell types. At present, resolving these complex data sets to identify clusters of genes that respond together in a specific cell type after a stimulus requires either single-cell expression profiling[74] or laser-capture microdissection techniques[75]. If we add to this the realization that, for most brain areas, we do not know how many functionally distinct cell types exist or how to identify them, then the considerable challenges in the application of gene-expression profiling, as a discovery tool, to the CNS become apparent. The ability to use FLUORESCENCE-ACTIVATED CELL SORTING to sort cell populations from BAC transgenic mice that express fluorescent marker proteins can provide a source of large numbers of purified cells to aid in these experiments[76] (FIG. 4). Access to specific CNS cell types, or subpopulations of cells that are defined by expression of a molecular marker, can also facilitate many other molecular, biochemical and cell-biological assays.

A main difficulty encountered in studies of molecular mechanisms of CNS development and function derives from the fact that neurons function as complex networks, and that they are dependent on interactions in these networks for their normal functions. Establishing the projection patterns of neurons participating in these circuits can provide important insights into the organization of circuits and their physiological roles. For example, the convergence of axons from sensory neurons in the olfactory epithelium that express a single olfactory receptor onto a specific pair of glomeruli located at stereotyped positions in the olfactory bulb provided the first insights into the organization of the mammalian olfactory system[20,77]. Furthermore, the ability to express fluorescent reporter genes from specific loci for these receptors has allowed the precise mapping of these circuits throughout development, and access to specific sensory cells for functional imaging and electrophysiology[78].

There are several recently developed histological methods for tracing CNS projections that can be combined with BAC transgenic studies to visualize CNS circuitry rapidly. Detailed projection patterns of specific neurons have been visualized by the targeted expression of Tau–EGFP[78] or placental alkaline phosphatase[79,80], using transgenic, knock-in or GENE-TRAP approaches to

P ELEMENT
A *Drosophila* transposable element that has been used as a tool for insertion mutagenesis and for germ-line transformation.

HEAT-SHOCK PROMOTERS
DNA sequences that control the expression of a family of proteins (heat-shock proteins) that are synthesized in response to increases in temperature.

GAL4 UAS SYSTEM
An expression system in which ectopically expressed GAL4 will activate the transcription of a reporter gene or another target gene that is downstream of an upstream activation sequence (UAS).

POLYGLUTAMINE DISORDERS
Diseases characterized at the molecular level by CAG-trinucleotide-repeat expansions in a gene, which translate into an excess of glutamine repeats in the coded protein. A well-known example is Huntington's disease, which is caused by the presence of additional CAG repeats in the gene *huntingtin*.

KNOCK-IN TRANSGENESIS
The insertion of a mutant gene at the exact site of the genome where the corresponding wild-type gene is located. This approach is used to ensure that the effect of the mutant gene is not affected by the activity of the endogenous locus.

FLUORESCENCE-ACTIVATED CELL SORTING
A method that allows the separation of cells that express a specific protein by tagging them with a fluorescent antibody against the molecule of interest. A laser beam excites the fluorescent tag, and the emission of light triggers the cell sorting.

GENE TRAPPING
A mutation strategy that uses insertion vectors to trap or isolate transcripts from flanking genes. The inserted sequence acts as a tag from which to clone the mutated gene.

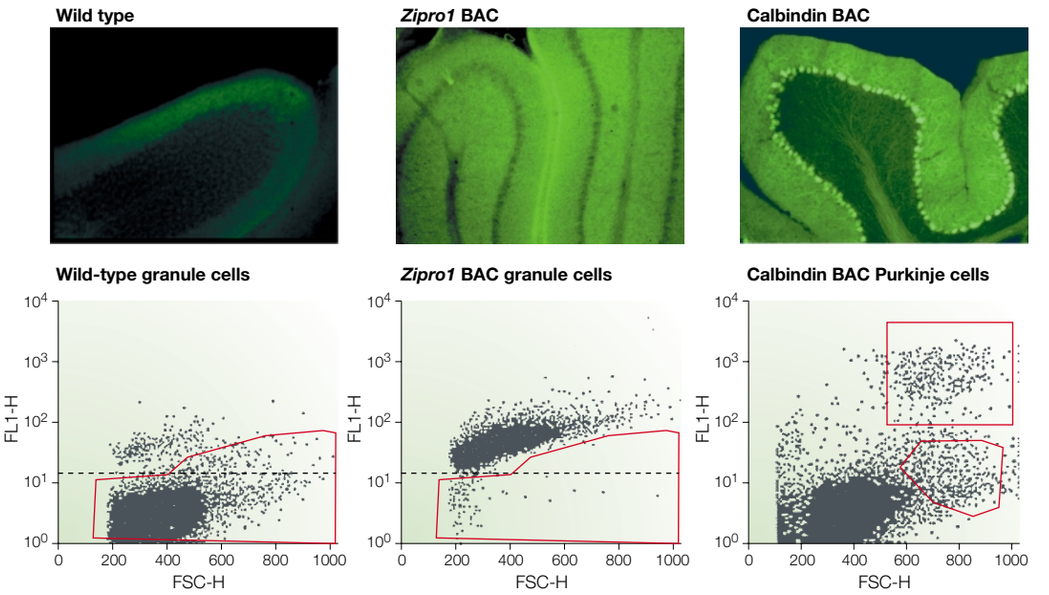

Figure 4 | **Fluorescence-activated cell sorting of neurons from BAC transgenic mice.** The upper panel shows fluorescence images of the cerebellum from wild-type mice, *Zipro1* BAC transgenic mice that express enhanced green fluorescent protein (EGFP) in granule cells, and from calbindin BAC transgenic mice that express EGFP in Purkinje cells. The lower panel shows fluorescence-activated cell sorting profiles of cells isolated from these mice at postnatal day 3. FL1-H, fluorescence intensity; FSC-H, forward light scatter (an indirect measure of size). The shift in the plot from *Zipro1* BAC mice shows moderately increased fluorescence intensity of small cells, consistent with the expression of EGFP in cerebellar granule cells. The shift in the plot from calbindin BAC mice shows greatly increased fluorescence intensity of large cells, consistent with the fluorescence of cerebellar Purkinje cells in these mice (unpublished data provided by X. W. Yang).

target expression to appropriate neurons. Histological identification of secondary and tertiary cells in a specific circuit has been accomplished by transgenic expression of plant LECTINS as trans-synaptic reporters[81,82]. Trans-synaptic neuronal tracing has also been accomplished by the specific expression of Cre recombinase in BAC transgenic mice, in combination with a conditional Cre-dependent pseudorabies virus (PRV), to trace neural circuits involved in feeding[40]. The binary nature of this system allows the simultaneous control of the site (specific cell types in specific brain regions) and timing of virus production and hence precise circuit tracing.

The immediate future

The ability to express proteins in specific cell populations in BAC transgenic mice has stimulated several efforts to develop new methods, or to adapt existing technology for *in vivo* use. Although many of these ideas are still poorly developed, some are of obvious importance and are under active development. Two of these are described here.

Combinatorial transcription systems to target cell types. It is relatively rare to find a gene that is uniquely expressed in a single CNS cell type. In most cases, genes are transcribed in several different cell types or at distinct times in development. The specificity with which these techniques can be applied would be significantly improved by using a binary approach. For this reason, systems are now being developed in which expression in the animal is the result of the production of two (or more) components, each expressed in a distinct pattern from a different BAC. Using this strategy, proper expression of the recombinase or regulatory protein of interest will be achieved only in those cells in which there is an overlap in the expression of the required components.

Transgenic silencing of neuronal activity. Genetic regulation of neuronal excitability would be useful for many *in vivo* studies. Efforts to modify neuronal excitability have included the ectopic expression of K^+ channels to suppress activity. Early attempts to apply this strategy *in vivo* met with limited success[83,84]. Recently, transgenic studies using the GAL4 UAS system to target expression of a recombinant Shaker K^+ channel (EKO) to different cell populations in *Drosophila* have indicated that this strategy might be widely applicable *in vivo*[85]. So, dose-dependent effects on neuronal excitability have been documented *in vitro* and *in vivo*, and behavioural effects that are consistent with suppressed excitability or loss of neuronal function *in vivo* have been observed. Although this method has not yet been used in transgenic mice to modulate neuronal excitability, the application of this method to the analysis of several different cell types in *Drosophila* indicates that it might be readily adapted for use in mammals. The combination of this system with the reproducible and highly specific regulation of transgene expression that can be achieved using BAC constructs could allow the application of this strategy to a wide variety of mammalian CNS cell types.

LECTINS
Sugar-binding proteins that tend to agglutinate cells. Concanavalin A is a widely used example.

Concluding remarks

Here I have argued that the accurate transcription of BAC transgenes *in vivo*, coupled with the ability to easily insert, delete, or alter genes that are resident in these large DNA constructs, has led to the development of novel strategies that are particularly important in neuroscience research. The preparation of transgenic mice that carry a properly engineered BAC construct can provide rapid access to the profile of cell types that express the gene of interest, to the localization of its encoded product within the cell, and to the phenotypic consequences of its overproduction. Variations of this basic construct can be used to introduce affinity tags for biochemical analyses of protein assemblies that are required for CNS function, to map neuronal circuits that include the cell of interest, and to create cell-specific genetic perturbations. I believe that the application of BAC transgenic methodology, in 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 molecular mechanisms that govern the development, function and dysfunction of the mammalian nervous system.

1. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).
2. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
3. Miklos, G. L. & Rubin, G. M. The role of the genome project in determining gene function: insights from model organisms. *Cell* **86**, 521–529 (1996).
   A review of genetic and molecular approaches to gene function in invertebrates that provides insights into the increasing importance of transgenic organisms in understanding genes, pathways and biochemical functions *in vivo*.
4. Giraldo, P. & Montoliu, L. Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic Res.* **10**, 83–103 (2001).
   A technical discussion of large DNA vectors and their use in producing transgenic mice.
5. Yang, X. W., Model, P. & Heintz, N. Homologous recombination based modification in *Escherichia coli* and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nature Biotechnol.* **15**, 859–865 (1997).
   The first paper to describe the use of homologous recombination to manipulate BACs in *E. coli*. Provides the basic logic for manipulation of BACs and the first example of reporter gene expression from a modified BAC in transgenic mice.
6. Eggleston, A. K. & West, S. C. Exchanging partners: recombination in *E. coli*. *Trends Genet.* **12**, 20–26 (1996).
7. Zhang, Y., Buchholz, F., Muyrers, J. P. & Stewart, A. F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genet.* **20**, 123–128 (1998).
8. Jessen, J. R. *et al.* Modification of bacterial artificial chromosomes through Chi-stimulated homologous recombination and its application in zebrafish transgenesis. *Proc. Natl Acad. Sci. USA* **95**, 5121–5126 (1998).
9. Imam, A. M. *et al.* Modification of human β-globin locus PAC clones by homologous recombination in *Escherichia coli*. *Nucleic Acids Res.* **28**, E65 (2000).
10. Lee, E. C. *et al.* A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**, 56–65 (2001).
11. Messerle, M., Crnkovic, I., Hammerschmidt, W., Ziegler, H. & Koszinowski, U. H. Cloning and mutagenesis of a herpesvirus genome as an infectious bacterial artificial chromosome. *Proc. Natl Acad. Sci. USA* **94**, 14759–14763 (1997).
12. Gong, S., Yang, X. W. & Heintz, N. Rapid modification of bacterial artificial chromosomes (BACs) for large-scale BAC-mediated transgenic studies (submitted for publication).
13. Fabb, S. A. & Ragoussis, J. Yeast artificial chromosome vectors. *Mol. Cell Biol. Hum. Dis. Ser.* **5**, 104–124 (1995).
14. Ramón y Cajal, S. *Histology of the Nervous System* (English translation by N. Swanson and L. W. Swanson) (Oxford Univ. Press, New York, 1995).
15. Hunt, P. & Krumlauf, R. Hox codes and positional specification in vertebrate embryonic axes. *Annu. Rev. Cell Biol.* **8**, 227–256 (1992).
16. Wilkinson, D. G. Genetic control of segmentation in the vertebrate hindbrain. *Perspect. Dev. Neurobiol.* **3**, 29–38 (1995).
17. Lumsden, A. & Krumlauf, R. Patterning the vertebrate neuraxis. *Science* **274**, 1109–1115 (1996).
18. Rijli, F. M., Gavalas, A. & Chambon, P. Segmentation and specification in the branchial region of the head: the role of the Hox selector genes. *Int. J. Dev. Biol.* **42**, 393–401 (1998).
19. Trainor, P. A. & Krumlauf, R. Patterning the cranial neural crest: hindbrain segmentation and Hox gene plasticity. *Nature Rev. Neurosci.* **1**, 116–124 (2000).
20. Buck, L. & Axel, R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**, 175–187 (1991).
21. Zhao, H. *et al.* Functional expression of a mammalian odorant receptor. *Science* **279**, 237–242 (1998).
22. Fritze, C. E. & Anderson, T. R. Epitope tagging: general method for tracking recombinant proteins. *Methods Enzymol.* **327**, 3–16 (2000).
    A review of the advantages and practical considerations for the use of epitope tags *in vitro* and *in vivo*.
23. Martinez-Salas, E. Internal ribosome entry site biology and its use in expression vectors. *Curr. Opin. Biotechnol.* **10**, 458–464 (1999).
    A practical discussion of IRES sequences, their discovery, and their use to create polycistronic mRNAs for expression in eukaryotes.
24. Kuhar, S. G. *et al.* Changing patterns of gene expression define four stages of cerebellar granule neuron differentiation. *Development* **117**, 97–104 (1993).
25. Feng, L., Hatten, M. E. & Heintz, N. Brain lipid-binding protein (BLBP): a novel signaling system in the developing mammalian CNS. *Neuron* **12**, 895–908 (1994).
26. Schedl, A. *et al.* Transgenic mice generated by pronuclear injection of a yeast artificial chromosome. *Nucleic Acids Res.* **20**, 3073–3077 (1992).
27. Schedl, A., Montoliu, L., Kelsey, G. & Schutz, G. A yeast artificial chromosome covering the tyrosinase gene confers copy number-dependent expression in transgenic mice. *Nature* **362**, 258–261 (1993).
    An early study documenting correct expression from a large DNA transgene in mice. This study confirmed and extended earlier work from this laboratory, indicating that position-independent and copy-number-dependent expression could be achieved using a 250-kb YAC.
28. Huxley, C. Exploring gene function: use of yeast artificial chromosome transgenesis. *Methods* **14**, 199–210 (1998).
29. Antoch, M. P. *et al.* Functional identification of the mouse circadian *Clock* gene by transgenic BAC rescue. *Cell* **89**, 655–667 (1997).
    A companion paper to the positional cloning study, reporting the identification of the mouse *Clock* gene. It describes the use of gene dosage in BAC transgenic mice to show that *Clock* is rate limiting *in vivo*.
30. Wines, M. E., Shi, Y., Lindor, M. & Holdener, B. C. Physical localization of the *mesoderm development* (*mesd*) functional region. *Genomics* **68**, 322–329 (2000).
31. Mullins, L. J. *et al.* Granulation rescue and developmental marking of juxtaglomerular cells using 'piggy-BAC' recombination of the mouse Ren locus. *J. Biol. Chem.* **275**, 40378–40384 (2000).
32. Means, G. D., Boyd, Y., Willis, C. R. & Derry, J. M. Transgenic rescue of the tattered phenotype by using a BAC encoding Ebp. *Mamm. Genome* **12**, 323–325 (2001).
33. Probst, F. J. *et al.* Correction of deafness in *shaker-2* mice by an unconventional myosin in a BAC transgene. *Science* **280**, 1444–1447 (1998).
34. Sabath, D. E., Spangler, E. A., Rubin, E. M. & Stamatoyannopoulos, G. Analysis of the human ζ-globin gene promoter in transgenic mice. *Blood* **82**, 2899–2905 (1993).
35. Callow, M. J., Stoltzfus, L. J., Lawn, R. M. & Rubin, E. M. Expression of human apolipoprotein B and assembly of lipoprotein(a) in transgenic mice. *Proc. Natl Acad. Sci. USA* **91**, 2130–2134 (1994).
36. Frazer, K. A., Narla, G., Zhang, J. L. & Rubin, E. M. The apolipoprotein(a) gene is regulated by sex hormones and acute-phase inducers in YAC transgenic mice. *Nature Genet.* **9**, 424–431 (1995).
    An interesting study showing the proper regulation of a human gene in YAC transgenic mice, even though this gene does not occur in the mouse genome. So, the regulatory code for transcription of this gene is present in lower vertebrates, even though the gene itself is not.
37. Yang, X. W., Wynder, C., Doughty, M. L. & Heintz, N. BAC-mediated gene-dosage analysis reveals a role for Zipro1 (Ru49/Zfp38) in progenitor cell proliferation in cerebellum and skin. *Nature Genet.* **22**, 327–335 (1999).
    The first deliberate use of gene-dosage analysis in transgenic mice to uncover a function for a gene that does not show an obvious phenotype in the null state.
38. Zuo, J., Treadaway, J., Buckner, T. W. & Fritzsch, B. Visualization of α9 acetylcholine receptor expression in hair cells of transgenic mice containing a modified bacterial artificial chromosome. *Proc. Natl Acad. Sci. USA* **96**, 14100–14105 (1999).
39. Yu, W. *et al.* Coordinate regulation of *RAG1* and RAG2 by cell type-specific DNA elements 5′ of *RAG2*. *Science* **285**, 1080–1084 (1999).
40. DeFalco, J. *et al.* Virus-assisted mapping of neural inputs to a feeding center in the hypothalamus. *Science* **291**, 2608–2613 (2001).
    An interesting new technique for neural tracing that combines BAC transgenic mice with genetically altered pseudorabies virus (PRV) expression to allow cell-specific and temporally controlled mapping of neural circuits.
41. Carvajal, J. J., Cox, D., Summerbell, D. & Rigby, P. W. A BAC transgenic analysis of the *Mrf4*/*Myf5* locus reveals interdigitated elements that control activation and maintenance of gene expression during muscle development. *Development* **128**, 1857–1868 (2001).
42. John, R. M. *et al.* Imprinted expression of neuronatin from modified BAC transgenes reveals regulation by distinct and distant enhancers. *Dev. Biol.* **236**, 387–399 (2001).
43. Sabatini, B. L. & Svoboda, K. Analysis of calcium channels in single spines using optical fluctuation analysis. *Nature* **408**, 589–593 (2000).
44. Steward, O. mRNA localization in neurons: a multipurpose mechanism? *Neuron* **18**, 9–12 (1997).
45. Martin, K. C., Barad, M. & Kandel, E. R. Local protein synthesis and its role in synapse-specific plasticity. *Curr. Opin. Neurobiol.* **10**, 587–592 (2000).
46. Steward, O. & Schuman, E. M. Protein synthesis at synaptic sites on dendrites. *Annu. Rev. Neurosci.* **24**, 299–325 (2001).
47. Lin, R. C. & Scheller, R. H. Mechanisms of synaptic vesicle exocytosis. *Annu. Rev. Cell Dev. Biol.* **16**, 19–49 (2000).
48. Sudhof, T. C. The synaptic vesicle cycle revisited. *Neuron* **28**, 317–320 (2000).
49. Sollner, T. H. & Rothman, J. E. Molecular machinery mediating vesicle budding, docking and fusion. *Experientia* **52**, 1021–1025 (1996).
50. Cohen, R. S., Blomberg, F., Berzins, K. & Siekevitz, P. The structure of postsynaptic densities isolated from dog cerebral cortex. I. Overall morphology and protein composition. *J. Cell Biol.* **74**, 181–203 (1977).
51. Blomberg, F., Cohen, R. S. & Siekevitz, P. The structure of postsynaptic densities isolated from dog cerebral cortex. II. Characterization and arrangement of some of the major proteins within the structure. *J. Cell Biol.* **74**, 204–225 (1977).

52. Kennedy, M. B. Signal-processing machines at the postsynaptic density. *Science* **290**, 750–754 (2000).
53. Green, D. R. & Reed, J. C. Mitochondria and apoptosis. *Science* **281**, 1309–1312 (1998).
54. Arnold, D., Feng, L., Kim, J. & Heintz, N. A strategy for the analysis of gene expression during neural development. *Proc. Natl Acad. Sci. USA* **91**, 9970–9974 (1994).
55. Arnold, D. B. & Clapham, D. E. Molecular determinants for subcellular localization of PSD-95 with an interacting K^+ channel. *Neuron* **23**, 149–157 (1999).
56. Fields, S. & Song, O. A novel genetic system to detect protein–protein interactions. *Nature* **340**, 245–246 (1989).
57. Fondell, J. D., Ge, H. & Roeder, R. G. Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proc. Natl Acad. Sci. USA* **93**, 8329–8333 (1996).
58. Rout, M. P. *et al.* The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* **148**, 635–651 (2000).
59. Levine, K., Tinkelenberg, A. H. & Cross, F. The CLN gene family: central regulators of cell cycle Start in budding yeast. *Prog. Cell Cycle Res.* **1**, 101–114 (1995).
60. Rorth, P. A modular misexpression screen in *Drosophila* detecting tissue-specific phenotypes. *Proc. Natl Acad. Sci. USA* **93**, 12418–12422 (1996).
61. Perrimon, N. New advances in *Drosophila* provide opportunities to study gene functions. *Proc. Natl Acad. Sci. USA* **95**, 9716–9717 (1998).
62. Smith, D. J. *et al.* 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 (1997).
63. King, D. P. *et al.* Positional cloning of the mouse circadian *Clock* gene. *Cell* **89**, 641–653 (1997).
64. Roberts, P. A. & Broderic, D. J. Properties and evolutionary potential of newly induced tandem duplications in *Drosophila melanogaster*. *Genetics* **102**, 75–89 (1982).
65. Jansen, G. *et al.* The complete family of genes encoding G proteins of *Caenorhabditis elegans*. *Nature Genet.* **21**, 414–419 (1999).
66. Heintz, N. & Zoghbi, H. Y. Insights from mouse models into the molecular basis of neurodegeneration. *Annu. Rev. Physiol.* **62**, 779–802 (2000).
67. Zoghbi, H. Y. & Orr, H. T. Glutamine repeats and neurodegeneration. *Annu. Rev. Neurosci.* **23**, 217–247 (2000).
68. Orr, H. T. Beyond the Qs in the polyglutamine diseases. *Genes Dev.* **15**, 925–932 (2001).
69. Wheeler, V. C. *et al.* Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. *Hum. Mol. Genet.* **8**, 115–122 (1999).
70. Hodgson, J. G. *et al.* A YAC mouse model for Huntington's disease with full-length mutant huntingtin, cytoplasmic toxicity, and selective striatal neurodegeneration. *Neuron* **23**, 181–192 (1999).
71. Lamb, B. T. *et al.* Amyloid production and deposition in mutant amyloid precursor protein and presenilin-1 yeast artificial chromosome transgenic mice. *Nature Neurosci.* **2**, 695–697 (1999).
72. Gu, H., Marth, J. D., Orban, P. C., Mossmann, H. & Rajewsky, K. Deletion of a DNA polymerase β gene segment in T cells using cell type-specific gene targeting. *Science* **265**, 103–106 (1994).
73. Mansuy, I. M. & Bujard, H. Tetracycline-regulated gene expression in the brain. *Curr. Opin. Neurobiol.* **10**, 593–596 (2000).
74. Dulac, C. Cloning of genes from single neurons. *Curr. Top. Dev. Biol.* **36**, 245–258 (1998).
75. Emmert-Buck, M. R. *et al.* Laser capture microdissection. *Science* **274**, 998–1001 (1996).
76. Tomomura, M., Rice, D. S., Morgan, J. I. & Yuzaki, M. Purification of Purkinje cells by fluorescence-activated cell sorting from transgenic mice that express green fluorescent protein. *Eur. J. Neurosci.* **14**, 57–63 (2001).
77. Mombaerts, P. *et al.* Visualizing an olfactory sensory map. *Cell* **87**, 675–686 (1996).
78. Bozza, T. C. & Mombaerts, P. Olfactory coding: revealing intrinsic representations of odors. *Curr. Biol.* **11**, R687–R690 (2001).
79. DePrimo, S. E., Stambrook, P. J. & Stringer, J. R. Human placental alkaline phosphatase as a histochemical marker of gene expression in transgenic mice. *Transgenic Res.* **5**, 459–466 (1996).
80. Leighton, P. A. *et al.* Defining brain wiring patterns and mechanisms through gene trapping in mice. *Nature* **410**, 174–179 (2001).
81. Yoshihara, Y. *et al.* A genetic approach to visualization of multisynaptic neural pathways using plant lectin transgene. *Neuron* **22**, 33–41 (1999).
Shows the use of transgenic expression of wheat-germ agglutinin to trace neuronal circuits in transgenic mice.
82. Horowitz, L. F., Montmayeur, J. P., Echelard, Y. & Buck, L. B. A genetic approach to trace neural circuits. *Proc. Natl Acad. Sci. USA* **96**, 3194–3199 (1999).
83. Johns, D. C., Marx, R., Mains, R. E., O'Rourke, B. & Marban, E. Inducible genetic suppression of neuronal excitability. *J. Neurosci.* **19**, 1691–1697 (1999).
84. Nadeau, H., McKinney, S., Anderson, D. J. & Lester, H. A. ROMK1 (Kir1.1) causes apoptosis and chronic silencing of hippocampal neurons. *J. Neurophysiol.* **84**, 1062–1075 (2000).
85. White, B. H. *et al.* Targeted attenuation of electrical activity in *Drosophila* using a genetically modified K^+ channel. *Neuron* **31**, 699–711 (2001).
86. Coulson, A., Waterston, R., Kiff, J., Sulston, J. & Kohara, Y. Genome linking with yeast artificial chromosomes. *Nature* **335**, 184–186 (1988).
87. Shizuya, H. *et al.* 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 (1992).
88. Ioannou, P. A. *et al.* A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet.* **6**, 84–89 (1994).

Acknowledgements

I thank J. Friedman, M. E. Hatten, R. B. Darnell and X. W. Yang for their helpful comments during the preparation of this article. I also thank S. M. Magdaleno and T. Curran for the phrase "BAC to the future", which first appeared in a commentary on gene dosage in mice, published in *Nature Genetics*.

Online links

DATABASES

The following terms in this article are linked online to:

FlyBase: http://flybase.bio.indiana.edu/
Shaker

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/
Acra9 | BLBP | calbindin | *Clock* | huntingtin | *minibrain* | *Mrf4* | *Myf5* | *Nnat* | *Npy* | presenilin 1 | protein A | *Rag1* | *Rag2* | *Ren1* | Tau | *Zipro1*

OMIM: http://www.ncbi.nlm.nih.gov/Omim/
Alzheimer's disease | Down's syndrome | Huntington's disease | spinocerebellar ataxia 1

FURTHER INFORMATION

Encyclopedia of Life Sciences: http://www.els.net/
knockout and knock-in animals | transgenic animals | transgenic mice production | yeast artificial chromosomes

Access to this interactive links box is free online.