# A gene expression atlas of the central nervous system based on bacterial artificial chromosomes

**Shiaoching Gong**[1], **Chen Zheng**[1], **Martin L. Doughty**[1], **Kasia Losos**[1], **Nicholas Didkovsky**[2], **Uta B. Schambra**[4], **Norma J. Nowak**[5], **Alexandra Joyner**[6], **Gabrielle Leblanc**[7], **Mary E. Hatten**[2] & **Nathaniel Heintz**[3]

[1]GENSAT Project, [2]Laboratory of Developmental Neurobiology and [3]Laboratory of Molecular Biology, Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, Box 260, New York 10021, USA
[4]Department of Anatomy and Cell Biology, East Tennessee State University, Tennessee 37614, USA
[5]Roswell Park Cancer Institute, Buffalo, New York 14263, USA
[6]Developmental Genetics Program, Skirball Institute of Biomolecular Medicine, Department of Cell Biology, New York University School of Medicine and Howard Hughes Medical Institute, New York 10016, USA
[7]National Institute of Neurological Disorders and Stroke, NIH, Bethesda, Maryland 20892, USA

...................................................................................................................................................................................................................................................

**The mammalian central nervous system (CNS) contains a remarkable array of neural cells, each with a complex pattern of connections that together generate perceptions and higher brain functions. Here we describe a large-scale screen to create an atlas of CNS gene expression at the cellular level, and to provide a library of verified bacterial artificial chromosome (BAC) vectors and transgenic mouse lines that offer experimental access to CNS regions, cell classes and pathways. We illustrate the use of this atlas to derive novel insights into gene function in neural cells, and into principal steps of CNS development. The atlas, library of BAC vectors and BAC transgenic mice generated in this screen provide a rich resource that allows a broad array of investigations not previously available to the neuroscience community.**

Since the revolutionary work of Ramon y Cajal[1], the structure of the nervous system has been interpreted in the context of specific neuronal and glial cell types, and their inter-relationships. The remarkable range of cell types present in the CNS and the complexity of their interconnections complicate efforts to provide a thorough analysis of CNS gene expression, and demand that expression data be collected at cellular resolution. The development of transgenic mouse methodology for reporter gene analysis using precisely modified BACs offers a new method for correlation of gene expression with cell type, because the elaborate specializations of morphologically complex cells can be visualized[2,3]. We describe here a large-scale project using BAC transgenic mice to provide detailed gene expression information for CNS-expressed genes, to identify a library of BAC vectors for manipulation of specific CNS cell types, and to provide a collection of mouse lines carrying enhanced green fluorescent protein (EGFP) reporter genes in a large variety of CNS cell types. We demonstrate the use of this approach to formulate detailed hypotheses concerning individual gene functions, to analyse cell migration, to discover novel topographic features of CNS projections, and to define BAC vectors that provide experimental access to cells within the circuitry of neural systems.

## The GENSAT BAC Transgenic Project

The ability to use reporter gene strategies in transgenic mice to map gene expression at cellular resolution was first demonstrated using modified BACs carrying ~150 kilobases (kb) of genomic DNA surrounding the mouse *Zipro1* gene[3]. Subsequent studies have established the general utility of BAC transgenic mice for analysis of mammalian gene expression and function, for the isolation of specific cell populations, and for the identification of effective vectors that can allow reproducible experimental access to specific cell types[2]. Visualization of EGFP expression in specific neuronal or glial cell populations has also provided compelling evidence that this strategy can result in the identification of expressing cell types, as the detailed morphology of EGFP-expressing cells is

apparent. The Gene Expression Nervous System Atlas (GENSAT) BAC Transgenic Project is designed to take advantage of these features of EGFP reporter genes to provide detailed expression maps for thousands of genes. The basic strategy for the project is to: (1) select genes to be analysed; (2) prepare modified BACs in which an EGFP reporter gene is substituted for the gene of interest; (3) prepare transgenic lines carrying each reporter construct; (4) analyse systematically reporter gene expression in these lines during development of the CNS; and (5) annotate the expression data and report to a public database.

### Methodological considerations

This project has required critical decisions regarding the overall strategy of the project, refinements of existing methodologies, and development of new resources for data management and presentation. In particular, new methodology was developed and tested for the production of BAC reporter gene constructs, for efficient production of BAC transgenic mice, and for presentation of the anatomical data using a web-administered database (http://www.gensat.org/). These decisions and methodological improvements are described in the Supplementary Information.

The GENSAT BAC Transgenic Project is the first large-scale effort to use both the mammalian genome sequence and the BAC clones that provide the basis of the genome physical map to generate anatomical data and experimental resources for use by the neuroscience community. The usefulness of the GENSAT data, vectors and animals to the neuroscience community depends in part on the accuracy of the expression data obtained and the ability to target gene expression in a reproducible manner using the BAC vectors described in the GENSAT project. Comparative analysis of data collected for known genes targeted using BACs selected from the mouse genome databases versus expression data available in the literature or in our laboratories has highlighted several important points regarding the BAC transgenic approach to gene expression mapping.

First, in each of the BAC transgenic vectors, the endogenous messenger RNA and protein coding sequences have been replaced by sequences encoding the EGFP reporter gene. As in any gene-replacement experiment, the stabilities of the reporter gene mRNA and protein can be different from those of the endogenous gene products. Thus, our results identify the relative rates of transcription for each gene in cells of the nervous system; they are not a direct measure of mRNA accumulation or protein abundance for the endogenous gene products. Furthermore, the increased sensitivity of the reporter gene assays, particularly in BAC transgenic lines carrying multiple copies of the BAC transgene, may allow detection of sites of expression that are not evident from *in situ* hybridization experiments. For these and other reasons, one can expect some differences between the GENSAT expression data sets and *in situ* hybridization or immunohistochemical data, particularly for dynamically regulated genes.

Second, to achieve a high rate of success in reproducing the expression pattern of a gene of interest, it is necessary to include the entire transcription unit and all of its associated regulatory information. There is at present no way of predicting the locations of these regulatory elements. However, as the average transcription unit for mammalian genes is less than 100 kb, and the carrying capacity of BACs is several hundred kilobases, it is possible in most cases to identify a BAC covering the transcribed unit and including 50–100 kb of 5′- and 3′-flanking intergenic DNA. In our experience so far, ~85% of BACs chosen using these general parameters express reproducibly in multiple transgenic lines. Those that do not express reproducibly owing to position effects on integration into the genome are easily identified, because the sites of expression vary significantly between lines.

Third, the release of the mouse genome databases[4,5] has markedly improved the ability to identify and target an appropriate BAC for this screen. For example, our original BAC transgenic lines for choline acetyltransferase (Chat) were prepared before release of the genomic databases using a BAC (RP23-51F19) that we subsequently learned contained ~85 kb of 5′-flanking DNA, but truncated the gene ~25 kb before the polyA addition site. As shown in Fig. 1a, this BAC expressed well in transgenic mice, correctly revealing the main cholinergic brainstem nuclei (Fig. 1a, panels 4 and 5) and spinal cord motor neurons (Fig. 1a, panel 6). However, no expression was detected in the basal forebrain (Fig. 1a, panels 2 and 3), and moderate ectopic expression of the BAC transgene was detected in layer 2–3 of the cerebral cortex (Fig. 1a, panel 1). On release of the genomic databases, BAC transgenic reporter lines were redone using BAC (RP23-431D9) covering the entire Chat locus, including ~85 kb of 5′-flanking DNA and ~60 kb of 3′-flanking DNA. As illustrated in Fig. 1b, expression of the reporter gene in these mice faithfully reproduced that of the endogenous gene, revealing the appropriate brainstem cholinergic nuclei (Fig. 1b, panels 4 and 5), spinal cord motor neurons (Fig. 1b, panel 6), and basal forebrain (Fig. 1b, panels 2 and 3) and cortical interneuron populations (Fig. 1b, panel 1). Critical regulatory elements must reside within the 3′-flanking DNA of the mouse Chat locus.

Fourth, for very large genes (>200 kb), the BAC approach to gene expression mapping has a low probability of success. Thus, about 10% of the lines we have analysed so far were for large genes in which the reporter gene construct contained ~100 kb of 5′-flanking DNA and ~100 kb of the transcription unit. In only one case has this approach been successful.

Finally, for most genes, multiple copies of the EGFP reporter gene are required for a sufficient signal-to-noise fluorescent ratio in tissue slices.

Thus, multiple lines carrying the BAC transgene were routinely screened to choose an optimal line for in-depth expression analysis.
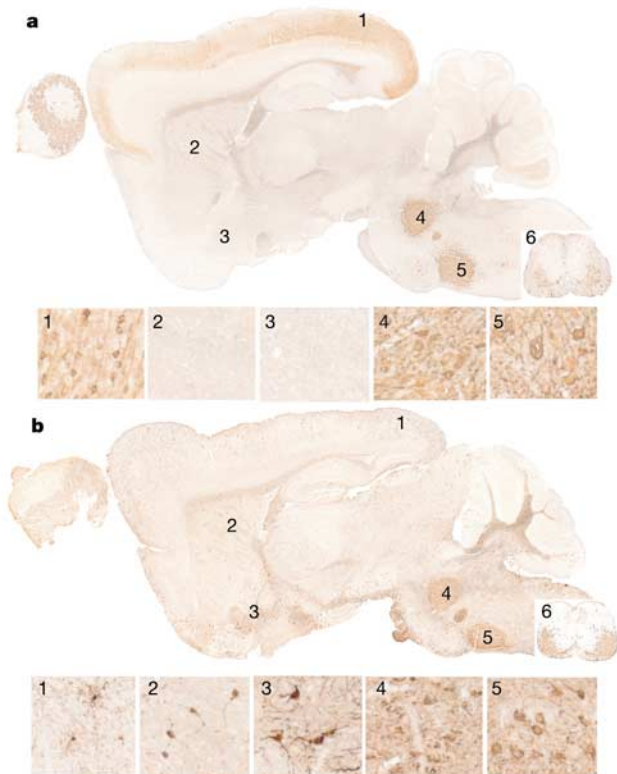


**Figure 1** Expression of *Chat* from different BAC constructs. Expression in adult sagittal sections from a BAC transgenic line prepared using BAC RP23-51F19 (**a**) or BAC RP23-431D19 (**b**). Panels 1–6 show higher-magnification images of the cerebral cortex (panel 1), basal forebrain (panels 2 and 3), brainstem (panels 4 and 5) and spinal cord (panel 6).
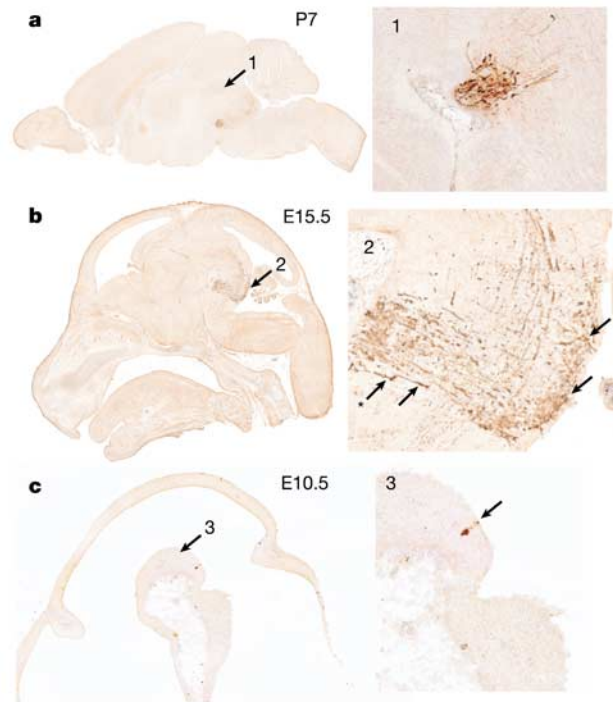


**Figure 2** Cell lineage revealed by expression data from reporter gene analysis in BAC transgenic mice. **a**, Labelling of P7 *Gscl* BAC transgenic brain; the only structure labelled is the IPN of the brainstem (panel 1). **b**, *Gscl* expression at E15.5. Higher magnification reveals individual migrating neurons (panel 2). **c**, Labelling of E10.5 reveals two labelled cells (panel 3). Higher magnification shows the two committed progenitor cells of the IPN.

Data from the chosen line are included in the database using the following descriptors: 'confirmed' refers to genes in which multiple lines yield matching data sets that agree with the available literature, or in which a single line has been produced that matches the available literature. In those cases where there is a discrepancy between the literature and the reporter gene data (missing or additional sites of expression), this is noted in the database; 'to be confirmed' refers to genes in which multiple lines produce matching data sets, but for which no independent expression data are available. 'Anomalous' refers to data collected from BAC reporter constructs producing different expression patterns in different lines (that is, position effects).

## Biological insights from GENSAT expression data

To illustrate the advantages of systematic analysis of gene expression using BAC transgenic reporter analysis to formulate precise hypotheses concerning gene function in development, we have chosen specific examples relevant for cell specification, axon–target interactions, cell migrations and studies on neural systems.

### The *Gscl* gene is involved in neural specification

One of the most striking advantages of the BAC constructs is the ability to visualize small cohorts of cells that express a particular gene and to follow their development over time. *Gscl* (goosecoid-like) was discovered during genomic analysis of the DiGeorge syndrome/velocardiofacial syndrome (DGS/VCFS) critical region in the human genome[6]. It is a member of a small family of homeodomain genes that includes goosecoid and *GSX*, and binds a conserved DNA sequence consistent with its proposed function as a transcription factor[7]. Inspection of BAC transgenic reporter lines

for *Gscl* at postnatal day 7 (P7) reveals robust expression in a subset of neuronal cell bodies within the interpeduncular nucleus (IPN) (Fig. 2a, panel 1). Lower levels of axonal expression (and an occasional cell soma) are evident in the tegmental nucleus, which receives afferent input from the IPN (data not shown). At embryonic day 15.5 (E15.5), *Gscl* expression in the developing IPN and in the ventricular zone adjacent to the developing IPN is evident, as are profiles of migrating neurons transiting from the ventricular zone to the area of the presumptive IPN (Fig. 2b, panel 2). The exquisite specificity of *Gscl* expression in the developing CNS, and the ability of BAC transgenic reporter gene studies to reveal this specificity, are most dramatically illustrated by analysis of these mice at E10.5 (Fig. 2c, panel 3). Only two cells expressing EGFP were detected in serial sections of a whole *Gscl* BAC transgenic embryo. These cells are present in the ventricular zone at the border of the developing mesencephalon and metencephalon, as expected from visualization of the expanding IPN progenitor pool in E15.5 embryos. These results suggest that *Gscl* is a lineage marker for a specific class of neurons that arise in the ventricular zone at E10.5, establish a progenitor pool that is still differentiating and migrating into the presumptive IPN at E15.5, and eventually settle into the IPN during the first postnatal week. Although previous *in situ* hybridization data are consistent with this proposal, the ability to visualize *Gscl*-expressing cells and their projections with such precision has allowed us to suggest a hypothesis that could not be formulated based on the *in situ* hybridization data. Furthermore, the identification of *Gscl* expression in a small population of neurons in the IPN suggests a functional role for Gscl that has not been tested in any of the studies done to assess the consequences of loss of Gscl function *in vivo*. For example, loss of this small cell population as a
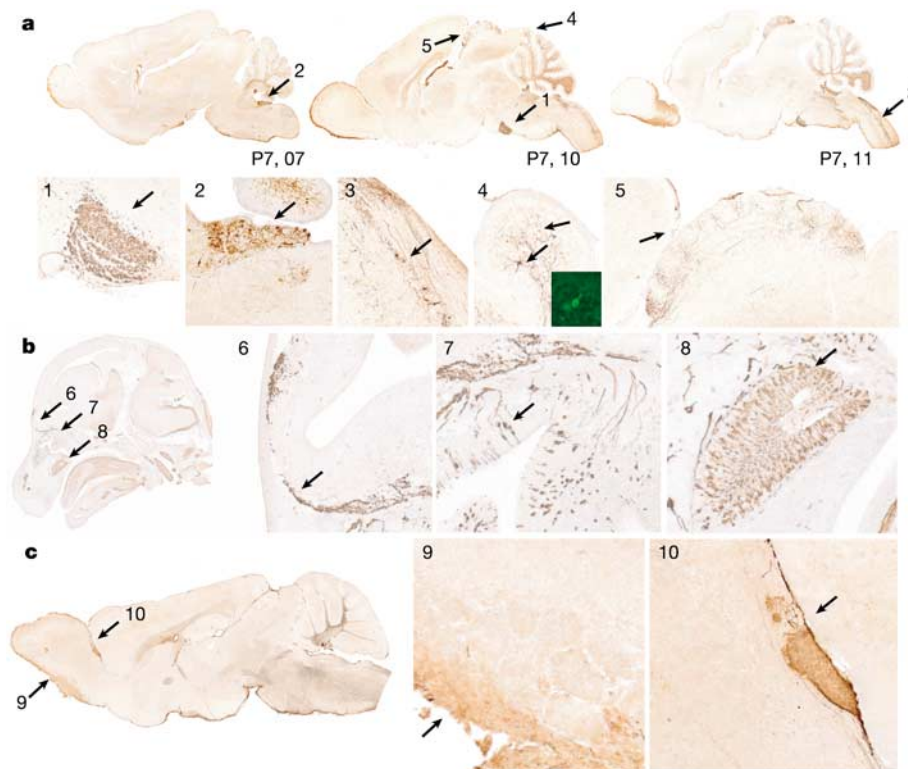


**Figure 3** Predicting developmental functions for specific genes on the basis of expression data in BAC transgenic mice. **a**, Serial sagittal sections of P7 *Sema3b* BAC transgenic mice. Numbered arrows indicate areas shown in the corresponding panels. Neurons that project mossy fibre afferents to the cerebellar cortex and cerebellar Golgi neurons are labelled: pontine nucleus (panel 1), vestibular nucleus (panel 2), the spinocerebellar tract (panel 3) and the cerebellar cortex (panel 4). Confocal microscopy of Golgi neurons (panel 4, fluorescent insert). A novel topographic map of retinal ganglion axons is seen in the superior colliculus (panel 5). Confocal microscopy reveals labelling of cells in the cerebellar cortex (panel 4, fluorescent insert). **b**, *Sema3b* expression in the developing olfactory system (E15.5): MOB (panel 6), olfactory epithelium (panel 7), vomeronasal organ (panel 8). **c**, *Sema3b* expression in the adult: external plexiform layer (panel 9); AOB (panel 10).

consequence of *Gscl* deletion would be missed unless specific markers that identify this cell type were examined in the knockout animals. Given the important functions of the habenular/interpeduncular system in the generation of hippocampal theta rhythms and the control of rapid-eye-movement sleep[8], one might expect electroencephalographic recordings in freely behaving, *Gscl* null mutant mice to reveal important, yet subtle, behavioural phenotypes. These studies would now seem warranted, particularly given the presence of this gene in the minimal DGS/VCFS critical region[6] and the clinical complexity of this syndrome[9]. Further investigation of these issues will be aided by the ability to visualize the EGFP-expressing neurons of the IPN using the BAC transgenic line presented here.

## The *Sema3b* gene is involved in axon–target-interactions

Another key step in the development of the CNS is the outgrowth of specific sets of axons and their interaction with particular targets. In our screen, the semaphorin 3B (*Sema3b*) BAC transgenic mice provide an example of the ability of BAC constructs to reveal the morphology and axon–target projection patterns of neurons that express a particular guidance molecule. Sema3b is a member of a small family of secreted semaphorins, first characterized as repulsive axon guidance signals[10–12]. Although the specific roles of Sema3b in vertebrates are not yet clear, Sema3b can act to repulse axons expressing neuropilin 2 (*Npn2*), and can antagonize the repulsive action of Sema3a on axonal growth cones expressing neuropilin 1 (*Npn1*)[13].

In the developing cerebellar system Sema3b marked the projection neurons from the pontine nucleus, the vestibular nucleus and the spinocerebellar tract (Fig. 3a). Mossy fibres emanating from the pons (Fig. 3a, panel 1), the vestibular nuclei (Fig. 3a, panel 2) and the spinocerebellar tract (Fig. 3a, panel 3) were labelled clearly (panel 4). In addition, the large Golgi interneurons (Fig. 3a, panel 4, lower arrow and fluorescent inset) present in the granule cell layer also expressed the EGFP reporter. This distribution is particularly interesting, as each of the *Sema3b*-expressing cell types in the brainstem, spinal cord and cerebellum provide presynaptic contacts

to cerebellar granule cells. Moreover, it is noteworthy that *Sema3a* is expressed in cerebellar Purkinje cells and can collapse growth cones of basilar pontine axons *in vitro*[14]. As mossy fibres grow through a field of immature granule cells and Purkinje cells before establishing synapses with granule neurons, these data suggest that Sema3b might transiently antagonize the repulsive action of Sema3a during targeting of *Sema3b*-expressing mossy fibre axons to, and ramification of Golgi cell axons within, the developing cerebellar internal granular layer.

Visualization of the expression of the EGFP reporter in *Sema3b* BAC transgenic mice also predicts an important role for Sema3b in the formation of the circuitry of the olfactory system. In E15.5 embryos, *Sema3b* expression in the developing olfactory system is widespread (Fig. 3b). Individual cells expressing the reporter are evident throughout the olfactory epithelium (Fig. 3b, panel 7), and labelled axon fascicles extend from these cells to the surface of the developing main olfactory bulb (MOB; Fig. 3b, panel 6). Intense expression is also present in the developing vomeronasal organ (VNO; Fig. 3b, panel 8). Two features of the staining patterns suggest specific roles for Sema3b in the formation of axonal projections to the MOB and the accessory olfactory bulb (AOB) from the sensory epithelia. First, we do not observe *Sema3b*-expressing axons exiting from the external plexiform layer into the glomerular layer of the MOB at any stage of development (Fig. 3b, panel 6; Fig. 3c, panel 9). Given the strong expression of *Sema3a* in the developing olfactory epithelium (Fig. 3b, panel 7), it is possible that transient *Sema3b* expression is required to antagonize the repulsive effects of *Sema3a* on olfactory sensory neuron growth cones as they exit from the nasal epithelium and project to the MOB. In both the MOB and the cerebellar cortex, we suggest that *Sema3b* must be turned off as axons reach their targets so that *Sema3a* expression in mitral cells and Purkinje cells, respectively, can prevent axons from overshooting their normal target fields. Second, *Sema3b*-expressing neurons in the VNO (Fig. 3b, panel 8) project specifically to the posterior AOB (Fig. 3c, panel 10). In *Npn2* knockout mice, axons from apical VNO neurons overshoot their normal terminal field in the anterior AOB, and invade the posterior
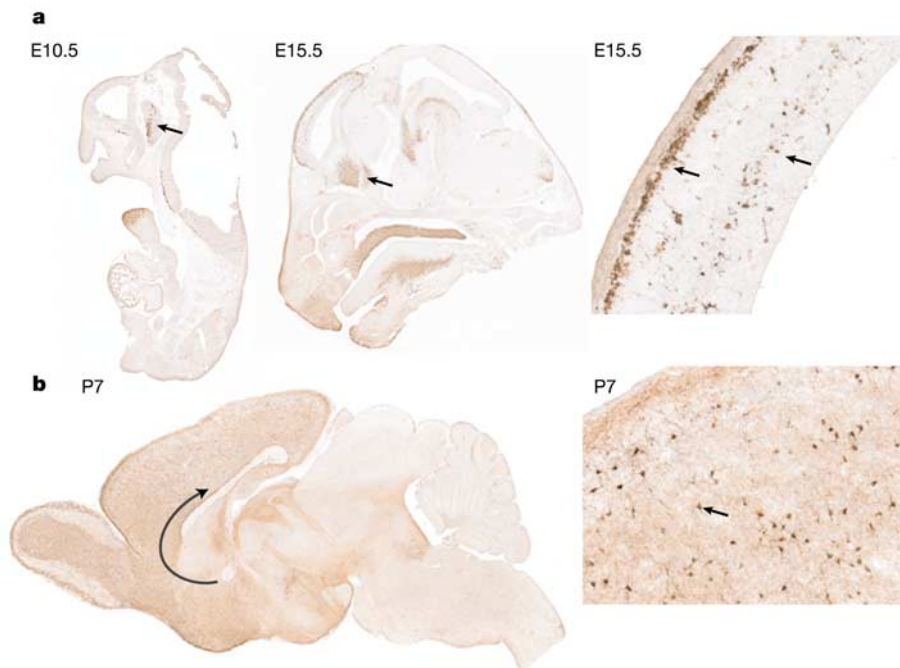


**Figure 4** Genesis and migration of cortical interneurons is revealed with *Lhx6* BAC transgenic mice. **a**, Interneurons are born in the ventral forebrain (E10.5, left panel, arrow), which gives rise to the MGE (E15.5, middle panel, arrow). By E15.5, interneurons arrive in the cerebral cortex (right panel, arrows). **b**, At P7, interneurons have spread through the cortex (arrows); higher magnification is shown in the right panel.

AOB[15]. As Sema3b can bind to Npn2-containing receptors with high affinity, these results suggest that the release of Sema3b from axon terminals in the posterior AOB may have a direct role in repelling *Npn2*-expressing axon terminals from the posterior AOB.

Finally, the *Sema3b* BAC transgenic mice revealed a novel pattern of axon–target interactions in the superior colliculus (Fig. 3a, panel 5). Axons from the retina arrive at the superior colliculus embryonically, although their invasion of the superficial grey layer is delayed until shortly after birth[16,17]. During the first few postnatal weeks, significant remodelling of retinocollicular projections occurs to establish the adult retinotopic map[18,19]. Sema3a (collapsin 1) has been shown to bind to retinotectal fibre tracts in chick embryos, and to induce retinal ganglion cell (RGC) growth cone turning *in vitro*[15], suggesting a role for semaphorins in establishment of the retinotectal (retinocollicular) projection. The pattern of *Sema3b*-expressing RGC axons in the superior colliculus at P7 is provocative (Fig. 3a, panel 5). In medial sagittal sections from *Sema3b* BAC transgenic mice, an interesting striped pattern of RGC projections to the superior colliculus is evident (Fig. 3a, panel 5). These results predict a role for Sema3b in the formation or refinement of the retinocollicular map, and reveal a novel topographic feature of the map that is not explained by the reported actions of ephrins[20–22] and Eph receptors in establishing the anterior–posterior map of RGC terminal fields on the superior colliculus. Taken together, the *Sema3b* expression data illustrate the advantages of reporter gene assays to reveal novel topographic features of the CNS that cannot be visualized by conventional methodology. Moreover, the *Sema3b* BAC transgenic animals offer the means to examine the behaviour of growing axons in real time as well as to carry out further genetic studies on the role of Sema3b in axon guidance.

### *Lhx6* and *Pde1c* genes reveal tangential migratory patterns

The migration of immature neurons from germinal zones to positions where they form synaptic connections is one of the hallmarks of cortical development. The BAC transgenic mice provide a means to visualize these migrations in real time and to follow the same cells over the developmental epoch of neural layer formation. Genetic analysis of mice lacking the transcription factors Dlx1, Dlx2 and Nkx2.1 has identified an important pathway of cortical interneuron migration that runs from the basal forebrain into the cortex[23,24]. An early population of cells migrates from the medial ganglionic eminence (MGE) up into the cortex, giving rise to GABA (γ-aminobutyric acid)-containing interneurons. Later, cells from the lateral ganglionic eminence (LGE) migrate dorsally into the cortex, also generating GABAergic interneurons. As some 20% of the neurons in the cortex are these GABAergic interneurons, these ventrodorsal cell migrations provide a large proportion of cortical neurons[25]. Recent studies indicate that antibodies against Lhx6 stain migrating MGE cells. In the current screen, *Lhx6* provided a marker for these cells[26]. Although antibody labelling studies showed expression of *Lhx6* in this migratory stream, they did not provide direct access to the development of cells expressing *Lhx6*. As shown in Fig. 4, *Lhx6*-positive cells were visible in the ventral forebrain as early as E10.5, with a large population of cells evident in the MGE by E15.5. At E15.5, cells can be seen migrating from the MGE up into the cortex. Higher-magnification views at E15.5 reveal small, labelled cells in the cortex in the marginal zone and scattered throughout the cortical plate. This pattern is in good agreement with studies of the migration pattern of DiI-labelled cells. By P7, the migration is nearly complete and labelled cells are absent from the marginal zone. Dense bands of EGFP–Lhx6 cells are seen across the cortex; on the basis of their morphology, these are apparently interneurons. Thus, the *Lhx6*BAC transgenic mouse provides an ideal system for dynamic studies on the genesis and migration of cortical interneurons.

In addition to generating lines of mice that confirm existing models of cell migrations in the cortex, lines were identified that revealed new pathways of migration. One of these was the phosphodiesterase 1C (*Pde1c*) gene, a cyclic nucleotide-hydrolysing phosphodiesterase that is important in regulating some cyclic AMP responses of postsynaptic receptors[27]. Although Pde1c has not been studied extensively in the CNS, experiments in smooth muscle cells support a role in cell proliferation and differentiation[28]. In the E10.5 embryo, staining was seen along the dorsal ridge of the



**Figure 5** Expression of *Pde1c* reveals a novel and distinct migratory pathway in developing brain. **a**, At E10.5, EGFP-positive cells are visible along the dorsal aspect of the neural tube. At E15.5 (middle panel), labelled cells are present in the anterior portion and superficial aspect of the cortex. At higher magnification (right panel), cells above the cortical plate are labelled. **b**, At P7, cells in layer 1 of the cortex are labelled (arrow). The right-hand panel shows higher magnification with confocal microscopy (inset); layer 1 Cajal Retzius cells express the reporter gene. Cells that undergo a similar pattern of migration, the cerebellar granule cells, are also labelled, as are cells in the olfactory bulb and pontine nucleus.

spinal cord, within the rhombic lip of the midbrain–hindbrain territory and in the anterior aspect of the telencephalic vesicle (Fig. 5). By E15.5, Pde1c marked a major population of dorsally migrating neurons, the external germinal layer of the fourth ventricle, and another population of dorsally migrating cells in the cerebral cortex. Along the fourth ventricle, the labelled cells were the precursors of the cerebellar granule neurons, which spread across the surface of the cerebellar anlagen at this developmental stage as a population of mitotic, migratory cells. In the cortex, Pde1c marked a population of cells in the retrobulbar region of the cortex. By E15.5, labelled cells had spread across the cortical surface in a tangential migratory pattern that matched the pattern proposed for neurons in the subpial granular layer (SGL). Although earlier evidence held that these cells were present only in human cortex, recent studies[29] have proposed that this cell population migrates across the surface of the cortex, intermingling with 'pioneer' neurons that arise from the cortical ventricular zone and migrate radially to form the marginal layer[30]. Previous studies have not clearly distinguished these two populations as they both express the reelin gene[31] and only occasionally calretinin. By P7, Pde1c-expressing cells persisted in layer 1 of the cortex, suggesting they are the SGL cell population because these cells persist longer than

pioneer neurons, which are thought to disappear by birth. Thus, Pde1c marks this previously unstudied cell population and provides a means to study the dynamics of the proliferation and migration of these cells.

## Access to cells within specific synaptic circuits

In addition to revealing developmental processes, many of the genes studied in the BAC transgenic screen marked specific cell classes within particular systems of the CNS. This is one of the most important results of the screen, as EGFP-labelled cells can be used directly for anatomical, physiological and genetic analysis of the particular cell population. For example, the EGFP-labelled BAC transgenic mice provide reproducible access to specific cell populations for electrophysiological studies of CNS cells and circuits, for vital imaging, or for a variety of cell isolation and tissue culture studies. The cerebral cortex is an important example, where the ability to impale and/or image live cells can allow more detailed studies on both the nature and plasticity of cortical circuits, and the characterization of specific and functionally distinct populations of pyramidal cells and interneurons. Although the cellular composition of the striatum is significantly less complex than that of the cerebral cortex, experimental access to each of the striatal cell types can significantly accelerate studies of striatal physiology, and the pathophysiological events reproduced in mouse models for Huntington's and Parkinson's diseases. We have chosen the cerebral cortex and striatum as examples of the ability of BAC transgenics to provide access to specific neurons within CNS circuits.

The cerebral cortex consists of six cell layers containing pyramidal cell neurons (the primary circuit neurons, layers 2–6) and interneurons (the local circuit neurons), as well as several glial cell types. Analysis of the first 100 genes screened in the GENSAT project revealed both layer- and cell-type-specific expression in the developing and adult cerebral cortex. For example, specific subsets of layer 2–3 pyramidal neurons are labelled in semaphorin 6D (Sema6d) BAC transgenic mice (Fig. 6a). Sema6d is a recently identified member of the semaphorin gene family whose expression and functions in the mouse CNS are not known. At P7, labelled cells in layer 2–3 have completed their radial migrations from the collapsing ventricular zone of the telencephalon, and are in the process of forming their connections with other cortical neurons. A novel gene, P271, is expressed predominantly in layer 2 and 5 pyramidal cells (Fig. 6b). This is a novel protein with a single proline-rich domain that is closely related to the human KIAA1399 gene. The morphology of expressing cells is clearly evident, including the ramifications of basal dendrites within layer 5 and the extension of the apical dendrite and its elaboration within layer 1 of the developing and adult cortex. Inspection of the P271 data sets also reveals staining of axons as they pass through the striatum en route to subcortical targets. Dopamine receptor D4 (Drd4) BAC transgenic lines express at high levels in the prefrontal cortex (Fig. 6c). At high magnification, these cells are identified as layer 5 pyramidal cells. This is consistent with previous data showing enrichment of this receptor in the prefrontal cortex, a principal target for antipsychotic drugs[32]. In contrast to the examples presented above, Otx1 is expressed in most layer 5 and 6 pyramidal cells throughout the developing and adult brain, as previously reported[33,34]. In this case, the very high levels of Otx1 expression in the developing and adult cerebral cortex, and the presence of EGFP in the dendritic processes of layer 5 and 6 pyramidal cells extending throughout the cerebral cortex, prevents effective visualization of cell somata by diaminobenzidine immunohistochemistry. To present convincing evidence of the restricted expression of Otx1 in the correct cortical layers, direct confocal imaging is used to reveal the morphology and distribution of expressing cells in the Otx1 BAC transgenic line (Fig. 6d, right panel). The presentation of data from both immunohistochemical and direct confocal imaging of the EGFP reporter gene in these lines ensures that one can
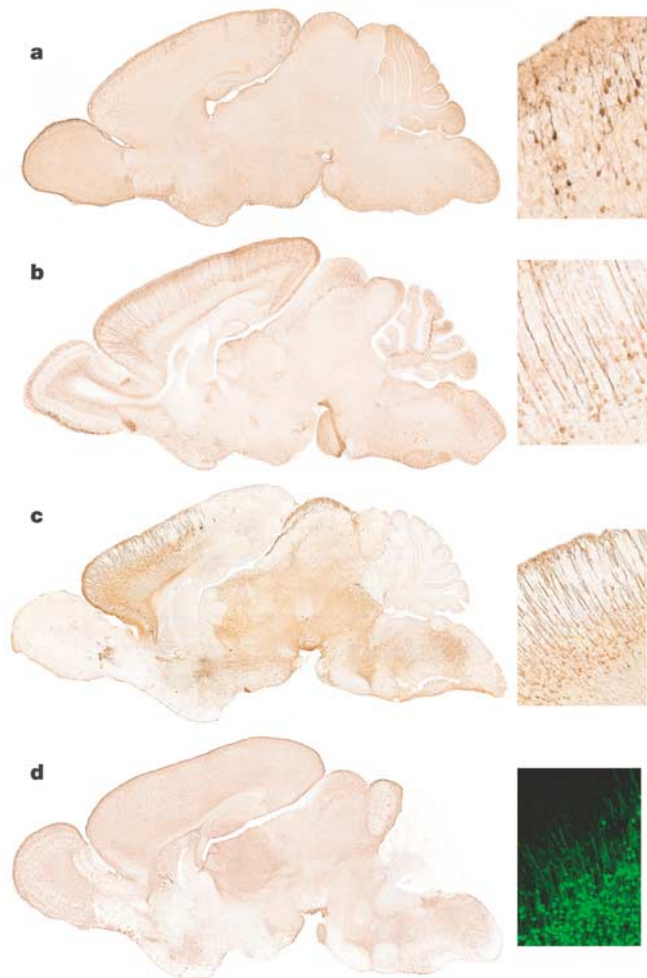


**Figure 6** Layer-specific gene expression in the developing cerebral cortex. **a–d**, P7 BAC transgenic mice expressing Sema6d (**a**), P271 (**b**), Drd4 (**c**) and Otx1 (**d**). The right-hand panels show higher magnification of cells in the same sections. Sema6d is expressed in cells of layer 2–3. P271 is expressed in pyramidal neurons of layers 2 and 5. Drd4 is expressed in layer 5 pyramidal cells in the frontal cortex, and Otx1 is expressed in pyramidal cells of layer 5–6. For Otx1, pyramidal cells are evident in the confocal image of unstained sections (**d**, right-hand panel).

properly assess the specificity of expression for all BAC transgenic lines. These data identify BAC vectors and transgenic animals that provide unprecedented tools for further analysis of the development, function and degeneration of the mouse cerebral cortex.

The striatum, a non-layered structure, is the largest component of the basal ganglia. Striatal projection neurons, approximately 90% of the neuronal population, provide the output from the striatum and receive nearly all of the input from extrinsic afferents and from striatal interneurons. The small population of striatal interneurons is thought to regulate afferent activity within the striatum[35]. The development of this system is complex, with cells deriving principally from the LGE and MGE. In the adult, all of the psychomotor behaviours ascribed to the striatum are linked to the dopaminergic innervation from the substantia nigra and the ventral tegmental area. Similarities in the morphology and neurochemistry of the cells in the striatum have confounded studies on the development of the neurons and their target interactions. In the BAC transgenic screen,

several genes revealed the cell classes and projection patterns of the striatum. As illustrated in Fig. 7a, medium spiny neurons projecting to the globus pallidus are labelled in the dopamine receptor D2 (*Drd2*) BAC transgenic mice[36,37]. In dopamine receptor D1a (*Drd1a*) BAC mice (Fig. 7b), medium spiny neurons projecting to the substantia nigra are labelled. Anatomical data suggest that the *Drd2* EGFP cells express enkephalin, whereas the *Drd1a*-positive cells express substance P[36,38]. Functional studies on acutely isolated EGFP-labelled neurons can now clarify whether there is co-localization of the two dopamine receptors in a subset of medium spiny neurons. A subpopulation of medium spiny neurons projecting to the substantia nigra, presumably also expressing *Drd1a*, expresses the M4 muscarinic cholinergic receptor (*Chrm4*) (Fig. 7c). These data are in agreement with electron microscopic studies in rats showing *Chrm4* localization to the somata and dendrites of medium spiny neurons[39]. The distribution of the *Chrm4*-expressing subpopulation across the striatum, and their specific localization to the striatal matrix, is of particular interest. Large, cholinergic aspiny neurons of the striatum are specifically labelled in the *Chat* BAC transgenic mice (Fig. 7d). Finally, clusters of medium spiny neurons that project to the substantia nigra are revealed in prodynorphin (*Pdyn*) BAC transgenic mice (Fig. 7e). These cells correspond to the striatal 'patches' that have been documented in a number of anatomical studies of the developing and adult striatum[40]. Taken together, this set of BAC transgenic mice reveals the principal cell types and projection pathways of the striatal system. These mice offer new tools for the analysis of the development of the striatum, and for a clearer understanding of how the dopamine and cholinergic environments influence synaptic function of specific subclasses of striatal neurons. Thus, they should facilitate novel neurophysiological and cell degeneration studies critical for our understanding of Parkinson's disease as well as neuropsychiatric disorders including Tourette's syndrome, schizophrenia and drug addiction.

## Implications for biological research

We report here selected findings from an ongoing large-scale screen of CNS gene expression in BAC transgenic mice. The screen uses novel vectors for the preparation of BAC reporter gene constructs, as well as methods for high-throughput transgenesis, histology and microscopy. Moreover, a MySQL database of the full-scale TIFF images, which can be viewed in browser format online (http://www.gensat.org/), has been generated to allow public access to the GENSAT data and for the design of further experiments. The information and materials generated so far provide immediate opportunities for further investigation. However, the full impact of this project will not be realized until many more reporter gene mice are analysed, and until the reagents and information from this screen are combined with further methods for genetic manipulation of the mammalian CNS[2].

The use of BAC transgenic reporter genes to study gene expression and provide experimental access to specific cell types can stimulate many fields of biological research. Thus, the capabilities for cell-specific genetic and physiological experimentation in the CNS that issue from these studies can be readily extended for the analysis of other cells, organs or tissues expressing the genes chosen for analysis by GENSAT, or in the context of similar future efforts. Given the general utility of this approach, there are several issues that warrant further discussion.

First, the use of a reporter gene that can reveal the morphology of complex cells and the systematic analysis of gene expression is necessary to obtain detailed insights and formulate precise hypotheses about gene function or development. The ability to present these large data sets from a systematic analysis for the benefit of other scientists is critically dependent on the use of a web-based browser that can present the captured data at full resolution. Although generation of large numbers of transgenic or gene-
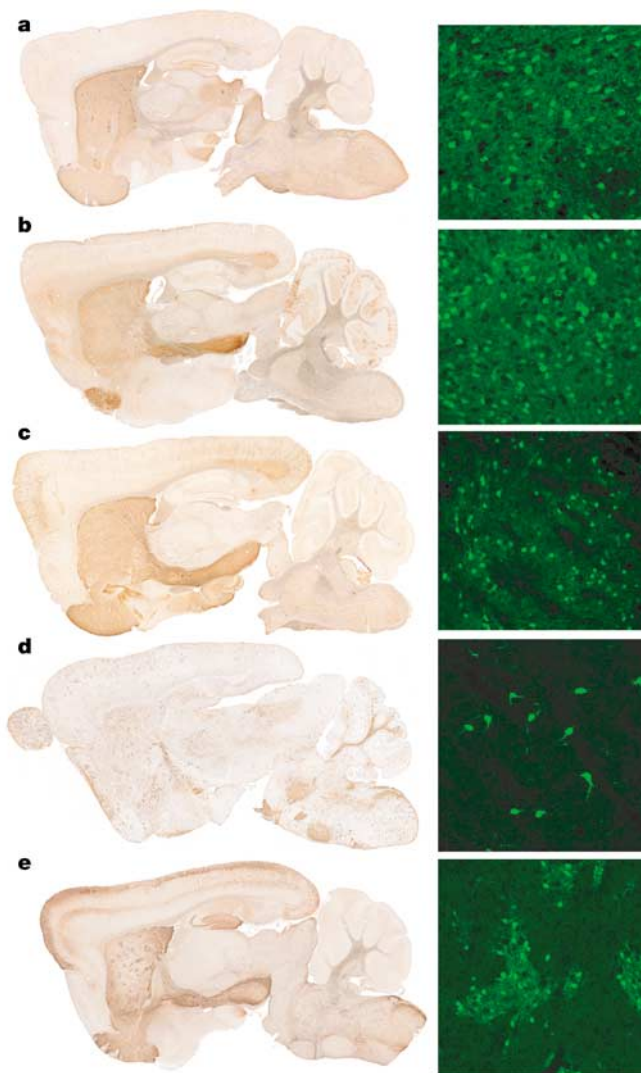


**Figure 7** Cell-specific EGFP marker expression in the adult striatum of BAC transgenic mice. Sagittal sections of adult mice are shown immunostained with antibodies against EGFP (diaminobenzidine). **a**, In dopamine receptor *Drd2* BAC transgenic mice, the medium spiny neurons (right panel, confocal microscopy) are labelled. **b**, Dopamine receptor *Drd1a* transgenic mice. **c**, In *Chrm4* transgenic mice, the nigrostriatal pathway is stained and medium spiny neurons are labelled (right, confocal microscopy). **d**, In *Chat* BAC transgenic mice, cholinergic interneurons of the striatum are labelled. **e**, *Pdyn* expression in striatal patches.

targeted mouse lines requires significant effort, in our experience it is the detailed analysis of these lines and the presentation of the data to the public that is rate limiting.

Second, the ability to visualize individual cells using the EGFP reporter has revealed that a large number of the genes studied so far are expressed in novel subsets of classically defined cell types. It is apparent from these data that these subsets of cells can also share other important features; for example, they may all project to the same area of the CNS, or expression may mark distinct subpopulations in multiple CNS structures. We believe that, in many cases, these differences portend distinct functional properties, and anticipate that similar subdivisions of classical cell types will be observed as these methods are applied to other organ systems. This will require careful reconsideration of the meaning and definition of specific cell types *in vivo*.

Third, our experience so far has highlighted the advantages and limitations of the BAC reporter gene approach. The main advantages are the relative ease of producing BAC transgenic animals, the ability to identify lines carrying multiple copies of the BAC transgene in the genome, and the use of transgenic techniques in other vertebrate species. Several methods are now available for manipulation of BACs by homologous recombination[3,4,41,42], which can allow the construction of complex cassettes for *in vivo* manipulation of genes, cell types and tissues. A single day of pronuclear injection will normally yield multiple BAC transgenic founders (see Supplementary Information) that all transmit the transgene to subsequent generations. For these reasons, the use of well-characterized BAC vectors for *in vivo* manipulation can often proceed more efficiently than conventional gene targeting. Furthermore, anecdotal evidence from our laboratories and those of our colleagues points to many cases in which targeting EGFP into an endogenous genomic locus using homologous recombination ('knock in') has failed to produce sufficient reporter gene expression to be directly visualized in tissue slices or *in vivo*. This is consistent with the general observation reported here that, in many cases, EGFP fluorescence cannot be detected unless more than five copies of the BAC transgene are present in the genome. The ability to produce BAC transgenic lines expressing a wide range of dosages of the BAC transgene has many important applications for both vital imaging and genetic analysis. Finally, gene targeting is not yet feasible in many species that can be manipulated by pronuclear injection (cows, rats, zebrafish and so on). The main limitations to this approach to gene expression analysis are that it cannot be routinely used to study very large genes ($>150$ kb), and that it is considerably less efficient than lower-resolution techniques such as *in situ* hybridization. For this reason, the GENSAT project now includes a pre-screen using radioactive *in situ* hybridization to examine a large number of genes each year (T. Curran, unpublished data). From this pre-screen, 250–300 genes are chosen for in-depth analysis using BAC transgenic mice.

Fourth, the use of transgenic constructs as the basis for a gene expression screen provides valuable insights into eukaryotic gene regulation. Thus, the data obtained in using this approach identify individual BACs as either sufficient or insufficient to mediate correct gene expression *in vivo* (Fig. 1), and precisely identify the cell types that can be accessed using this segment of genomic DNA. This sort of experimental information is critical in trying to decipher the code for transcriptional regulation. It also identifies a starting point for dissection of transcriptional regulatory regions, and for the construction of artificial, small promoters to target gene expression *in vivo*.

Finally, the ability to specifically and reproducibly target expression to specific CNS cell populations using well-defined BAC vectors offers the opportunity to dramatically expand the capability for cell-specific genetic manipulation of the mouse genome. Thus, BAC transgenic mice expressing site-specific recombinases, increased levels of wild-type proteins, dominant-activating

or negative alleles, and other proteins and RNAs of interest in specific CNS cell types, have been or are being produced using BAC vectors identified in the GENSAT project.

## Concluding remarks

The publication of the initial results of the GENSAT BAC Transgenic Project marks the beginning of a major effort to use genome mapping and sequence information in a directed, large-scale endeavour to advance mammalian neuroscience research. For the first time, the framework for completion of a high-resolution atlas of CNS gene expression, for the generation of a comprehensive library of BAC vectors to access defined CNS cell types, and for the creation of a repository of mice carrying labelled CNS cell populations has been established. Over the past 2 years, improvements in the methodology used in this screen, and in the availability and accuracy of the mouse genome databases, have markedly improved the efficiency of this project. Continued improvements in reporter gene technology, in the elements required for effective expression of polycistronic transcripts *in vivo*, and in inducible recombinases for genetic manipulation will further enhance the utility of this effort. We view the GENSAT BAC Transgenic Project as complementary to ongoing, large-scale directed and non-directed phenotypic screens to test CNS gene function. It is our hope that the examples we have discussed and the data we have released will accelerate research in all areas of neuroscience, and that access to the BAC vectors and transgenic lines we have produced will stimulate the fusion of mammalian molecular genetics and functional analysis of CNS cells and circuits. □

## Methods

### BAC transgene construction

The plasmid (pLD53.SC2) used in the homologous recombination is a derivative of pLD53 (ref. 43). pLD53 was digested with *Bam*HI and *Sac*I to get rid of the *tetAR* and *oriT* origin and replaced by *Not*I–*Sal*I–*Spe*I adaptor. A 1.1 kb EGFP.PA was cloned into *Not*I and *Sal*I sites, and an *Asc*I–*Not*I–*Swa*I–*Sma*I multiple cloning site was then cloned into the *Not*I site, which was knocked out thereafter. This shuttle vector (pLD53.SC2) was digested with *Asc*I/*Sma*I and purified by running on a 1% low-melting agarose gel. A 300–500 bp 'A Box' fragment used for the homologous recombination was amplified by polymerase chain reaction (PCR), digested with *Asc*I and cloned into this shuttle vector.

BAC host cells were streaked on an agar plate supplemented with chloramphenicol (Chlr) ($20 \mu g$ ml$^{-1}$) and incubated overnight at $37 °C$. A single colony was picked, inoculated with 5 ml of Luria broth (LB) supplemented with Chlr ($20 \mu g$ ml$^{-1}$), and grown to optical density (OD) 600 of 0.5. The cells were harvested by centrifugation at 3,000 r.p.m. for 10 min at $4 °C$. The pellet was suspended with 5 ml of ice-cold 50 mM CaCl$_2$, placed on ice for 5 min and harvested as above. The pellet was suspended with $300 \mu l$ of ice-cold solution of 50 mM CaCl$_2$ and 20% glycerol. An aliquot of $5 \mu l$ of PSV1.RecA plasmid ($20 ng \mu l^{-1}$) was transformed into $100 \mu l$ of competent cells; 1 ml of SOC was added, cells were incubated at $30 °C$ for 1.5 h with shaking at 225 r.p.m., and selected in 5 ml of LB containing tetracycline (Tet; $5 \mu g$ ml$^{-1}$) and Chlr ($20 \mu g$ ml$^{-1}$) at $30 °C$ overnight with shaking at 300 r.p.m. The overnight culture was diluted 1:50, inoculated with 50 ml of LB containing Tet ($5 \mu g$ ml$^{-1}$) and Chlr ($20 \mu g$ ml$^{-1}$) in a 500 ml flask, and grown to OD 600 of 0.6–0.8 at $30 °C$ (4–5 h). The cells were harvested by centrifugation in a cold rotor at 3,000 r.p.m. for 10 min (in a Beckmann J6-MI centrifuge). The competent cells were prepared according to published protocols[4]. An aliquot of $2 \mu l$ of shuttle vector ($500 ng \mu l^{-1}$) was transformed into $40 \mu l$ of competent cells by electroporation. After the recovery, the cells were selected in 5 ml of LB containing Tet ($5 \mu g$ ml$^{-1}$), ampicillin (Amp; $50 \mu g$ ml$^{-1}$) and Chlr ($20 \mu g$ ml$^{-1}$), and incubated at $30 °C$ overnight. The overnight culture ($10 \mu l$ and $100 \mu l$) was spread onto LB plates containing Amp ($50 \mu g$ ml$^{-1}$) and Chlr ($20 \mu g$ ml$^{-1}$), and incubated overnight at $43 °C$. Cointegrates were identified as reported previously[4].

Modified BAC DNA was prepared by double acetate precipitation and CsCl gradient separation. The quality and concentration of the DNA were checked on pulse field gel. DNA was diluted to $2 ng \mu l^{-1}$ and dialysed on a 25 mm, $0.025 \mu m$ filter (VSWP02500, Millipore) by floating it on 20 ml of injection buffer with the shiny side up for 3–4 h. An aliquot of DNA was removed to another tube and mixed with the same amount of ×2 polyamine (50:50) 48 h to 1 week before the injection. The BAC DNA ($0.15–1 ng \mu l^{-1}$) was injected into 200 pronuclei of fertilized oocytes of FVB/N mice.

### Histology

Neonatal and adult animals were anaesthetized (Nembutal) and perfused with 4% paraformaldehyde. The brain and spinal cord were removed and post-fixed (4% paraformaldehyde, 1 h). Embryos were removed from deeply anaesthetized dams by laparoscopy and placed in ice-cold phosphate-buffered saline (PBS), and post-fixed by immersion in 4% paraformaldehyde (12 h, $4 °C$). Serial sagittal sections ($16 \mu m$) were generated with a Micron or Leica cryostat, post-fixed in methanol (1 min, $4 °C$) and

washed in PBS. Slides were either used directly for confocal imaging or immunostained using antibodies against EGFP (gift of M. Rout, Rockefeller University), horseradish peroxidase (using the TSA Biotin System, NEN) and ABC Vectastain (Vector Laboratories) methods. Primary antibody was added (1:2–10,000) overnight at 4 °C. For each neonatal and adult BAC transgenic animal, a set of 10–11 serial sections was generated and processed as above. Adult sections were chosen to match the series presented in the Paxinos atlas[44], with sections 1–10 matching Paxinos sections 129, 126, 123, 120, 117, 114, 111, 108, 105 and 102, respectively. Transverse spinal cord sections were C4–C6, T3–T6 and L6–S1. P7 sections were matched to sections from the Valverde atlas[45], with sections 1–11 matching Valverde figures 25–35, respectively. Transverse spinal cord sections were C4–C6, T3–T6 and L6–S1. E15.5 sections were chosen to match the Schambra atlas[46] with sections 1–10 matching Schambra figures 1–10, respectively. Transverse spinal cord sections were C4–C6, T3–T6 and L6–S1. For E10.5 embryos, the maximum number of sections possible was generated.

Images were acquired with a Zeiss Axiocam camera on a Zeiss Axioskop2 microscope with a ×5 and ×10 (1.5 numerical aperture) objective and an automated $x,y$ stage (Marzhauser scan8) controlled and hosted by a PC with Zeiss KS400 software running a macro written by us. Adult sagittal sections were approximately 200 MB. The final resolution of the ×10 images was approximately 1.33 μm per pixel. Confocal images were acquired with a Zeiss Axioskop2 using a BioRad Radiance confocal imaging system.

1. Ramon y Cajal, S. *Histology of the Nervous System* (Oxford Univ. Press, New York, 1911).
2. Heintz, N. BAC to the future: the use of BAC transgenic mice for neuroscience research. *Nature Rev. Neurosci.* **2,** 861–870 (2001).
3. 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).
4. Gong, S., Yang, X., Li, C. & Heintz, N. Highly efficient modification of bacterial artificial chromosomes (BACs) using novel shuttle vectors containing the R6Kγ origin of replication. *Genome Res.* **12,** 1992–1998 (2002).
5. Copeland, N. G., Jenkins, N. A. & Court, D. L. Recombineering: a powerful new tool for mouse functional genomics. *Nature Rev. Genet.* **2,** 769–779 (2001).
6. Gottlieb, S. *et al.* The DiGeorge syndrome minimal critical region contains a goosecoid-like (*GSCL*) homeobox gene that is expressed early in human development. *Am. J. Hum. Genet.* **60,** 1194–1201 (1997).
7. Gottlieb, S., Hanes, S. D., Golden, J. A., Oakey, R. J. & Budarf, M. L. Goosecoid-like, a gene deleted in DiGeorge and velocardiofacial syndromes, recognizes DNA with a bicoid-like specificity and is expressed in the developing mouse brain. *Hum. Mol. Genet.* **7,** 1497–1505 (1998).
8. Valjakka, A. *et al.* The fasciculus retroflexus controls the integrity of REM sleep by supporting the generation of hippocampal theta rhythm and rapid eye movements in rats. *Brain Res. Bull.* **47,** 171–184 (1998).
9. Emanuel, B. S., McDonald-McGinn, D., Saitta, S. C. & Zackai, E. H. The 22q11.2 deletion syndrome. *Adv. Pediatr.* **48,** 39–73 (2001).
10. Raper, J. A. Semaphorins and their receptors in vertebrates and invertebrates. *Curr. Opin. Neurobiol.* **10,** 88–94 (2000).
11. Nakamura, F., Kalb, R. G. & Strittmatter, S. M. Molecular basis of semaphorin-mediated axon guidance. *J. Neurobiol.* **44,** 219–229 (2000).
12. Chen, H., He, Z. & Tessier-Lavigne, M. Axon guidance mechanisms: semaphorins as simultaneous repellents and anti-repellents. *Nature Neurosci.* **1,** 436–439 (1998).
13. Takahashi, T., Nakamura, F., Jin, Z., Kalb, R. G. & Strittmatter, S. M. Semaphorins A and E act as antagonists of neuropilin-1 and agonists of neuropilin-2 receptors. *Nature Neurosci.* **1,** 487–493 (1998).
14. Baird, D. H., Hatten, M. E. & Mason, C. A. Cerebellar target neurons provide a stop signal for afferent neurite extension *in vitro. J. Neurosci.* **12,** 619–634 (1992).
15. Walz, A., Rodriguez, I. & Mombaerts, P. Aberrant sensory innervation of the olfactory bulb in neuropilin-2 mutant mice. *J. Neurosci.* **22,** 4025–4035 (2002).
16. Sachs, G. M. & Schneider, G. E. The morphology of optic tract axons arborizing in the superior colliculus of the hamster. *J. Comp. Neurol.* **230,** 155–167 (1984).
17. Sachs, G. M., Jacobson, M. & Caviness, V. S. Jr. Postnatal changes in arborization patterns of murine retinocollicular axons. *J. Comp. Neurol.* **246,** 395–408 (1986).
18. Edwards, M. A., Caviness, V. S. Jr & Schneider, G. E. Development of cell and fiber lamination in the mouse superior colliculus. *J. Comp. Neurol.* **248,** 395–409 (1986).
19. Edwards, M. A., Schneider, G. E. & Caviness, V. S. Jr. Development of the crossed retinocollicular projection in the mouse. *J. Comp. Neurol.* **248,** 410–421 (1986).
20. Brown, A. *et al.* Topographic mapping from the retina to the midbrain is controlled by relative but not absolute levels of EphA receptor signaling. *Cell* **102,** 77–88 (2000).
21. Feldheim, D. A. *et al.* Genetic analysis of ephrin-A2 and ephrin-A5 shows their requirement in multiple aspects of retinocollicular mapping. *Neuron* **25,** 563–574 (2000).
22. Yates, P. A., Roskies, A. L., McLaughlin, T. & O'Leary, D. D. Topographic-specific axon branching controlled by ephrin-As is the critical event in retinotectal map development. *J. Neurosci.* **21,** 8548–8563 (2001).
23. Anderson, S. A., Eisenstat, D. D., Shi, L. & Rubenstein, J. L. R. Interneuron migration from basal forebrain to neocortex dependence on *Dlx* genes. *Science* **278,** 474–476 (1997).
24. Marin, O. & Rubenstein, J. A long, remarkable journey: tangential migration in the telencephalon. *Nature Rev. Neurosci.* **2,** 780–790 (2001).
25. Hatten, M. New directions in neuronal migration. *Science* **297,** 1660–1663 (2002).
26. Lavdas, A. A., Grigoriou, M., Pchnis, V. & Rubenstein, J. G. The medial ganglionic eminence gives rise to a population of early neurons in the developing cerebral cortex. *J. Neurosci.* **19,** 7881–7888 (1999).
27. Ross, C., MacCumber, M., Glatt, C. & Snyder, S. Brain phospholipase C isozymes: differential mRNA localizations by in situ hybridization. *Proc. Natl Acad. Sci. USA* **86,** 2923–2927 (1989).
28. Rybalkin, S., Rybalkina, I., Beavo, J. & Bornfeldt, K. Cyclic nucleotide phosphodiesterase 1C promotes human arterial smooth muscle proliferation. *Circ. Res.* **90,** 151–157 (2002).
29. Meyer, G., Soria, J. M., Martinez-Galan, J. R., Martin-Clemente, B. & Fairen, A. Different origins and developmental histories of transient neurons in the marginal zone of the fetal and neonatal rat cortex. *J. Comp. Neurol.* **397,** 493–518 (1998).
30. Derer, P. & Derer, M. Cajal-Retzius cell ontogenesis and death in mouse brain visualized with horseradish peroxidase and electron microscopy. *Neuroscience* **36,** 839–856 (1990).
31. D'Arcangelo, G. *et al.* Reelin is a secreted protein recognized by the CR-50 monoclonal antibody. *J. Neurosci.* **17,** 23–31 (1997).
32. Wang, X., Zhong, P. & Yan, Z. Dopamine D4 receptors modulate GABAergic signaling in pyramidal neurons of prefrontal cortex. *J. Neurosci.* **22,** 9185–9193 (2002).
33. Frantz, G., Weimann, J., Levin, M. & McConnell, S. Otx1 and Otx2 define layers and regions in developing cerebral cortex and cerebellum. *J. Neurosci.* **14,** 5725–5740 (1994).
34. Weimann, J. *et al.* Cortical neurons require Otx1 for the refinement of exuberant axonal projections to subcortical targets. *Neuron* **24,** 819–831 (1999).
35. Nicola, S., Surmeier, J. & Malenka, R. Dopamine modulation of neuronal excitability in the striatum and nucleus accumbens. *Annu. Rev. Neurosci.* **23,** 185–215 (2000).
36. Le Moine, C. & Bloch, B. D1 and D2 dopamine receptor gene expression in the rat striatum: sensitive cRNA probes demonstrate prominent segregation of D1 and D2 mRNAs in distinct neuronal populations of the dorsal and ventral striatum. *J. Comp. Neurol.* **355,** 418–426 (1995).
37. Surmeier, J., Song, W.-J. & Yan, Z. Coordinated expression of dopamine receptors in neostriatal medium spiny neurons. *J. Neurosci.* **16,** 6579–6591 (1996).
38. Hersh, S. *et al.* Electron microscopic analysis of D1 and D2 dopamine receptor proteins in the dorsal striatum and their synaptic relationships with motor corticostriatal afferents. *J. Neurosci.* **15,** 5222–5237 (1995).
39. Bernard, V., Levey, A. & Bloch, B. Regulation of the subcellular distribution of M4 muscarinic acetyl choline receptors in striatal neurons *in vivo* by the cholinergic environment: evidence for regulation of cell surface receptors by endogenous and exogenous stimulation. *J. Neurosci.* **19,** 10237–10249 (1999).
40. Gerfen, C. R. & Young, W. S. III Distribution of striatonigral and striatopallidal peptidergic neurons in both patch and matrix compartments: an *in situ* hybridization histochemistry and fluorescent retrograde tracing study. *Brain Res.* **460,** 161–167 (1988).
41. 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).
42. Muyrers, J. P., Zhang, Y., Testa, G. & Stewart, A. F. Rapid modification of bacterial artificial chromosomes by ET-recombination. *Nucleic Acids Res.* **27,** 1555–1557 (1999).
43. Metcalf, W. W. *et al.* Conditionally replicative and conjugative plasmids carrying *lacZα* for cloning, mutagenesis, and allele replacement in bacteria. *Plasmid* **35,** 1–13 (1996).
44. Paxinos, G. & Franklin, K. *The Mouse Brain in Stereotaxic Coordinates* 132 (Academic, San Diego, 2001).
45. Valverde, F. *Golgi Atlas of the Postnatal Mouse Brain* (Springer, Vienna, 1998).
46. Schambra, U., Lauder, J. & Silver, J. *Atlas of the Prenatal Mouse Brain* (Academic, San Diego, 1992).