In: Crusio WE and Gerlai R (eds.)
Handbook of Molecular Genetic Techniques for Brain and Behavior Research
(Techniques in Behavioral and Neural Sciences, Vol. 13) (1999) Elsevier
Chapter 2.2.3, pages 212-224.

Brain region-specific genes: the hippocampus

*Benjamin S. Pickard, Ben J. Davies, Ken A. Rose, Genevieve Stapleton, Muriel Steel and Richard Lathe**

Centre for Genome Research and Centre for Neuroscience, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JQ, UK

*Author for correspondence

How does the brain store information? Such a question presents a formidable challenge, not only because of the undoubted complexity of the brain, but also because we are limited in the analytical techniques we can bring to bear. At one level our understanding of what individual brain regions might do derives from selective lesion studies. For instance, in patients, lesions to the hippocampus lead to a pronounced failure of new declarative or explicit memory, while other types of memory, and recall of events well prior to damage, are substantially unaffected (Scoville and Milner, 1957). In rodents, where the water maze affords a measure of spatial memory (Morris, 1984), that shares common aspects with declarative memory (Cohen and Eichenbaum, 1993), ability to perform this navigation task is also abolished by hippocampal lesions (Morris et al., 1982).

At another level, identifying the role a brain region fulfils is only a first step towards understanding how it might fulfil that role. Neurons intercommunicate chemically and electrically, but what aspects of this intercommunication might underlie memory processes? The celebrated hypothesis put forward by Hebb, on purely theoretical grounds, held that if one neuron regularly communicates with another, then information could be stored by strengthening the synaptic connection between them. The discovery that hippocampal synapses display just such a form of synaptic plasticity, known as long-term potentiation (LTP; Bliss and Lømo, 1973), raises the important question of whether LTP might represent the molecular and cellular embodiment of certain types of memory. The answer will surely not be simple, but in favor of this contention Morris et al. (1986) were able to show that blockade of NMDA receptors prevents LTP and also impairs performance of rats in the watermaze.

Such pharmacologic studies, as well as the lesion studies mentioned earlier, lack complete specificity. Surgical or neurotoxic lesions can damage brain regions adjacent to the target structure, while pharmacologic agents are seldom if ever fully specific in their activities. More recent investigations have brought the powerful technique of transgenesis to bear on the problem. By homologous recombination in cultured embryonal cell lines it is possible to disrupt or modify the activity of a single gene within the mouse genome. Using this highly selective procedure it was found that animals deficient in a subunit of the NMDA receptor show impairments in both LTP and spatial learning (Sakimura et al., 1995). Although mice lacking the core NMDA-R1 subunit of the receptor (without which no NMDA receptor function is observed) fail to survive (Li et al., 1994), their specific deficits argue that this receptor is also necessary for the developmental formation of synaptic contacts.

Synaptic LTP in region CA1 is initiated by an ingress of Ca^{2+} through NMDA receptors (Bliss and Collingridge, 1993; Roberson et al., 1996); others have therefore focussed on generating mutations affecting calcium-dependent signalling molecules acting downstream of the NMDA receptor. Studies on a variety of mutant mice, including those lacking the α subunit of calcium/calmodulin protein kinase type II (CaMKII) (Silva et al.,1992a,b) or a brain isoform of protein kinase C (PKC-γ) (Abeliovich et al., 1993ab), revealed abnormalities of LTP and/or learning (reviewed by Chen and Tonegawa, 1997).

Despite the elegance of these transgenic experiments, the anatomic and temporal non-specificity of the genetic lesions complicates interpretation. The gene disruption is present in all tissues and at all stages of development. Thus, neither the impairments of memory nor those of LTP can be ascribed uniquely to loss of gene function in the hippocampal formation in the adult animal at the time of testing (Lathe and Morris, 1994; Mayford et al., 1995; Lathe, 1996). Behavioral and electrophysiological deficits could also be due to alterations in developmental processes, or indeed to dysfunction of other body or brain systems.

To address some of these problems efforts have been made to exploit inducible transgene expression in brain (Mansuy et al., 1998) or a line of transgenic mice in which expression of the site-specific recombinase CRE was serendipitously restricted to the hippocampus (Tsien et al., 1996a).

This latter approach generated mice in which a target gene (NMDA-R1 subunit) that had been flanked by CRE recognition sequences was deleted only in hippocampus (Tsien et al., 1996b). Nonetheless, in such an approach the site of integration of the transgene expressing CRE is not predetermined, a large number of different transgenic lines will need to be screened for each construct; and with no assurance of reproducing the region-specificity. Importantly, the integration of multiple copies, as usually occurs, leads to unreliable transgene expression (Dobie et al., 1997).

To avoid these difficulties we have sought genes expressed selectively in hippocampus and subregions thereof. Not only would the identity of genes specifically expressed in this brain region provide new insights into hippocampal function but, more importantly, such genes would furnish essential tools for the specific analysis of the role of the hippocampus in learning and memory - they could be used to target reproducibly the expression of probe genes (exemplified by but not restricted to CRE) to hippocampal regions. That such region-specific genes might exist was was encouraged by the previous description of restricted gene expression in other brain regions, for instance of tyrosine hydroxylase in the locus coeruleus and substantia nigra (Berod et al., 1987).

Here we discuss four approaches to the identification of genes and transcripts whose expression is restricted to the hippocampus, and the advantages and drawbacks of the techniques employed. Unexpectedly, each approach highlighted a different and non-overlapping subset of genes, but the expression of at least some of these was largely if not exclusively restricted to the hippocampus.

Differential hybridization

The first and most basic technique we attempted comprised the screening a cDNA library from adult rat hippocampus with probes derived from hippocampus (HPC) RNA or from material prepared from brain from which the hippocampus had been surgically removed ('rest of brain' - ROB). A cDNA library from adult rat HPC (10^6 clones) was prepared in lambda-ZAPII, and plaque-lifts were screened with HPC or ROB cDNA probes (Stapleton et al., 1995). From 10^6 primary cDNA clones, 361 hybridized significantly better to the HPC probe than the ROB probe. Of these, approximately 75% gave no significant signal with the ROB probe while 25% gave a weak signal on rescreening. A representative sample of 41 differentially hybridizing plaques was further analyzed by DNA sequencing. For reasons we do not understand, 11 contained repetitive elements. Why such sequences are particularly abundant in hippocampus is not known, but being inappropriate for transgenic experimentation were not studied further. Of the remaining, 9 were identified in the database while 21 were novel. Analysis of 17 selected clones is presented in Table 1.

Many identify transcripts expressed abundantly in rat hippocampus. Rat amyloidogenic glycoprotein (rAG) is the ortholog of the human amyloid precursor protein (APP) associated with the pathophysiology of Alzheimer's disease. The adult expression pattern of rAG mRNA is widespread, with particularly robust expression in hippocampus (Shivers et al., 1988). Transcripts for SNAP-25 (synaptosomal-associated protein) are particularly enriched in hippocampal mossy fibres and the molecular layer of the dentate gyrus as well as in several other brain regions (neocortex, piriform cortex, anterior thalamic nuclei, pontine nuclei, and cerebellar granule cells; Oyler et al., 1989). Other clones identified, however, did not show marked evidence of enriched

expression in the hippocampus. These include ß-adaptin, the receptor for platelet derived growth factor (PDGF), calmodulin (CaM), and RCM3. We have no explanation why these clones gave clear differential signals on primary and secondary screens, but are not reported to be enriched in hippocampus.

We focussed our attention instead on the novel cDNA species revealed by the differential screen. Northern analysis was performed using RNA from a number of brain regions. Of the clones analysed, several hybridized preferentially to hippocampal RNA, or identified transcript of sizes present only in hippocampus (FIG 1). Expression of Clone 15.13a identified a number of transcripts of which one, at 4.4 kb, appeared exclusively in hippocampus (see FIG 1) while clone 12.10a revealed a hippocampal- specific transcript at 7 kb (not shown). We have not established whether these transcripts are due to an alternative promoter that is only active in hippocampus, or is instead due to hippocampal-specific mRNA processing.

mRNA identified by the clone 14.5a was highly enriched in hippocampus; further analysis of this clone and further cDNAs corresponding to it revealed that 14.5a encodes a novel cytochrome P450 of the steroidogenic type, dubbed Cyp7b (Stapleton et al., 1995). Cyp7b expression in mouse and rat was confirmed by Northern blotting to be very highly enriched in the hippocampus, but was not exclusive to the formation (FIG 1; Stapleton et al., 1995). In situ hybridization using full-length cDNA probles to both mouse and rat brain sections has now shown that the hippocampus is the predominant site of brain expression (unpublished data, see FIG 3).

Subtractive hybridization

Hybridization can be used to remove sequences common to two starting preparations, so enriching for sequences only present in one or other pool. In the first technique we employed, chemical cross-linking subtraction (Hampson et al., 1992), difficulties were encountered: none of the 'enriched' clones demonstrated specificity for the hippocampus. The second technique tried, the multiple hybridization selection and reamplification technique of Wang and Brown (1991), was more successful. HPC and ROB mRNA preparations from adult rat brain were converted to double-stranded cDNA; each cDNA pool was divided in two, the aliquots were digested with frequent cutting restriction enzymes, and repooled. Oligonucleotide linkers were ligated to the cDNA fragments from each separate pool, permitting PCR amplification. The subtraction process was then carried out in tandem, one to identify HPC-specific cDNAs and the other to identify ROB-specific cDNAs. At each stage cDNA from one pool (the 'tracer') was hybridized with an excess of cDNA from the other pool (the 'driver') that had been biotinylated (using photobiotin in the presence of light); excess driver together with the tracer/driver hybrids was removed by streptavidin binding in conjunction with phenol-chloroform extractions, leaving unhybridized (subtracted) tracer

cDNAs in the aqueous phase. Several rounds of subtraction were carried out (three sets of long and short hybridizations that favour the removal of cDNA species common to both pools and present at relatively high abundance, such species are usually refractory to subtraction because of their quantity; Wang and Brown 1991). After each hybridization and subtraction the tracer pools were PCR amplified.

Initial experiments suffered from cross-contamination of driver and tracer pools, and a modification was made (see Balzer and Baumlein, 1994) in which the different cDNA pools are made using different linker/primer sets, preventing coamplification of driver and tracer. This protocol was applied to four rounds of subtraction of hippocampal (HPC) versus 'rest of brain' (ROB) cDNA; the products were cloned into a single-stranded M13 bacteriophage vector taking advantage of restriction sites in the PCR primers.

Duplicate plaque lifts were screened for hybridization to the fourth round subtracted HPC cDNA (H+4) and the primary ROB cDNA pools. 18 individual clones appeared to hybridize the HPC-enriched material, but not to ROB cDNA; these were used as hybridization probes for Northern blots and for the original mixed cDNA pools resolved by agarose gel electrophoresis ('cDNA Southern' analysis).

This yielded contradictory results. According to cDNA Southern data, many clones represented transcripts which were very highly enriched in the hippocampus and, in at least two cases, appeared to be almost specific to this region. Northern analysis, however, for the most part contradicted these results (Table 2), with one exception. Clone H4M10, identified as hippocalcin, appeared relatively specific to the hippocampus, as reported (Kobayashi et al, 1992). cDNA Southern analysis concurred with the hippocampal specificity, but this was very much less pronounced when the clone was used to probe Northern blots of mRNA from different brain regions; expression was also seen in lesser amounts in the cortex (FIG 2), confirming the results of Saitoh et al. (1993).

Our results demonstrate that the subtraction process is sufficiently thorough to enrich for sequences specific to or enriched in one starting pool. However, the protocols are not without their problems. First, specific cDNA species emerged during successive rounds of subtraction, and gave discrete bands on agarose gel electrophoresis. These appear to be due to cDNA species with extensive secondary structure such that they refold upon themselves, so avoiding removal by hybridization with the driver (Pickard, 1996). Second, in cases where the starting pools are not wholly representative of the cDNA populations, particular sequences will be selected for that, purely by chance, were present in one starting pool but not the other. Furthermore, any gene- or sequence-specific alterations that might skew the representation of a given sequence in the starting cDNA populations (due, for instance, to differential splicing, polyadenylation, hybridization to other endogenous RNAs or indeed covalent modification) will generate robust positives that are not confirmed by Northern hybridization (discussion).

Candidate gene approaches

Polymerase chain reaction (PCR) can address directly the spectrum of gene family members in a particular tissue, taking advantage of oligonucleotide primers corresponding to amino acid sequence motifs conserved within the family. Because amino acid sequence motifs may be encoded by a diversity of nucleotide sequences, a mixture of oligonucleotides (so-called 'degenerate' primers) representing all possible coding sequences is usually employed. Often, where any one of the bases A, C, T, or G could be used in a codon, the nucleotide inosine is specified to prevent excessive degeneracy (inosine can potentially hydrogen bond with all bases although it has a preference for G and A). Rat hippocampus cDNA was used as a template in PCR reactions designed to find members of five gene families which were selected for their apparent large sizes and potential roles in neuronal function (Table 3). PCR products were purified by agarose gel electrophoresis and cloned, by the 'TA overhang' method, into plasmid pCRII (Invitrogen).

Gprotein-coupled receptors. This family comprises a diversity of cell-surface receptors characterized by the presence of 7 transmembrane domains and which, on ligand binding, alter cellular metabolism through the modulation of the function of heterotrimeric G (GTP binding) proteins. Using degenerate PCR based on conserved motifs we identified 9 previously described receptors (adenosine-a1 and -a3, adrenergic-α1b and -β2, dopamine-d1, cannabinoid, endothelin-1, rat testis GPCR, R334 receptor, and a candidate vasoactive intestinal peptide receptor). However, Northern blotting revealed that in no case was expression restricted to the hippocampus.

Metabotropic glutamate receptors. These represent a subclass of the G-protein coupled receptors. PCR amplification predominantly revealed mGluR3, though clones were obtained corresponding to mGluR4 amd mGluR5. Because mGluR3 was represented most highly in the products, ways of increasing the abundance of other members were sought. One method attempted relied on the fact that, of the clones obtained, only mGluR3 contained a TaqI restriction site within the PCR product sequence. The PCR products were cut with this enzyme and then reamplified prior to cloning. Nevertheless, no new mGluRs were identified using this approach; none of the known mGluR subunits appears to be restricted to the hippocampus.

Heterotrimeric G proteins: α subunits. These participate in downstream signalling from the receptor groups discussed above and cycle between active (GTP-bound) and inactive (GDP-bound) forms. PCR amplification of G α subunit sequences from rat HPC cDNA subunits α-q, -i1, -s, -olf, and -i3. The presence of the olfactory G protein α subunit was particularly surprising as this had been reported to be specifically expressed in the olfactory bulb (Zigman et al., 1993). However, even though a specifically expressed 3' untranslated region exists, the coding region is very widely expressed in brain (not presented).

Protein tyrosine phosphatases (PTPs). These enzymes are either membrane-associated or cytoplasmic polypeptides and modulate the activation of target proteins by tyrosine dephosphorylation. Nine different PTPs were identified - α, γ, δ, H1, zeta, P19, Tcell PTP, and two novel sequences, PTP-4 and PTP-24 (Pickard, 1996). Sequence analysis revealed substantial homology between PTP-4 and the large multidomain PTPs μ and κ (Gebbink et al., 1991; Jiang et al., 1993) while PTP-24 shows homology within the amplified catalytic domain to a number of PTPs, most convincingly with hPTP-β (Kruger et al., 1990). Northern blotting was performed for the two new species. PTP-4 demonstrates a major transcript size of approximately 5.5 kb, principally localized to brain but with peripheral expression in lung and testis. PTP-24 detects brain transcripts of approximately 6 kb, but either 4 kb or 5 kb in the other tissues examined. Overall, PTP-24 is less restricted in expression than PTP-4 but neither was specific in their expression to the hippocampus (Pickard, 1996).

Of the other PTPs, the candidate tumor suppressor PTP-γ was found to be expressed fairly selectively in hippocampus but, earlier in development, expression is reported to be widespread in septal/thalamic nuclei, cortex and hippocampus (Barnea et al., 1993)

Serine proteases. These enzymes are characterized by a conserved serine residue in the active site and are held to play a variety of roles in the brain, including substratum clearance during neuronal migration and synapse formation, and growth factor activation by proteolytic cleavage of precursor molecules. PCR primers were designed according to sequence motifs conserved within the major (chymotryptic) clan of enzymes. The predominant species obtained by amplification, and confirmed by Northern blotting, were tissue-type plasminogen activator (t-PA), and RNK-Met-1, a lymphocyte protease not previously reported in brain, and two new family members, BSP1 (brain serine protease 1) and BSP2.

In situ hybridization revealed BSP1 transcripts in the hippocampal CA fields, with expression in CA1 and CA3 but only low-level expression in CA2. Expression was also detected in the polymorphic layer of the dentate gyrus and within deep layers of the entorhinal cortex (FIG 3). No significant expression was detected in other regions of the adult rat brain. Overall, expression of BSP1 was substantially restricted to the hippocampal formation while BSP2 expression was more widespread (Davies, 1998).

Gene-Trapping

The final approach was gene-trapping. A promoterless reporter cassette, comprising a splice-acceptor site derived from *en-2* preceding the *'ß-geo'* (a *ß*-galactosidase [*lacZ*]-neomycin resistance fusion) (Skarnes et al., 1992) was introduced into mouse ES cells. The *ß-geo* open-reading frame lacks a translation initiation site: application of G418 selected for insertion events where the

reporter has integrated into an endogenous gene locus, generating a fusion linking the N-terminus of an endogenous protein to the reporter polypeptide. G418-resistant ES cells were introduced into the cavity of mouse blastocysts, reimplanted into foster mothers, and the resulting chimeric males mated to produce offspring heterozygous for the gene-trap insertion. Initially 15 independent lines were produced by this method, as described (Skarnes, 1993, Wilson et al., in preparation; Steel et al., 1998).

To explore expression of the reporter gene, brain sections were incubated with the chromogenic ß-galactosidase substrate X-gal. In 4 lines, *obn, kin*, *hpk* and *glnC*, distinct staining was present in the hippocampus. *glnC* expressed ubiquitously and was not analyzed further. The remaining 11 lines, with the exception of *glnA* (that expressed at barely detectable levels in cerebellum) failed to exhibit reporter gene expression in adult brain (Steel et al., 1998), but of which 6 were negative in all tissues. A frequency of 4/15 lines expressing in hippocampus suggests that the gene-trap procedure is a feasible method to identify genes expressed in this brain structure.

In *obn* and *hpk* reporter activity was in several brain regions, but in brain of *kin* mice staining was substantially, but not exclusively, restricted to the CA1-CA3 subfields of the hippocampus (FIG 3). Weak staining was also detected in the molecular layer of the dentate gyrus and the lateral septum, with further weak reporter activity in amygdala, striatum and olfactory bulb (Steel et al., 1998).

Identification of the trapped genes employed a RACE protocol (rapid amplification of cDNA ends; Frohman et al., 1988) as modified (Skarnes et al., 1992). This involves priming reverse transcriptase with a *lacZ*-specific primer, tailing the product with poly-dA and subsequent PCR. We were initially unable to obtain cDNA clones corresponding to the hybrid transcripts and used a new modification that permits amplification of refractory sequences (Steel et al., 1998). This identified *kin* as an insertion into a gene encoding the Abelson-related non-receptor tyrosine kinase, *arg* (Kruh et al., 1990). In the *kin* line the *ß-geo* reporter is inserted downstream of an alternative 5' exon of the *arg* gene (Steel et al., 1998), and reports only expression from this promoter. Disappointingly, reporter mRNA was widespread in hippocampus and adjacent brain structures. The widespread pattern of mRNA expression, together with the restricted localization of the translation product, argues that the hybrid transcript is only translated efficiently within CA1-3.

Discussion

We have sought genes whose expression is restricted to the hippocampus, with a view to generating transgenic animals in which only the hippocampus is modified. We discuss below the different approaches employed, their merits and drawbacks, and possible means to improve them.

In our hands the most robust method was also the simplest - differential hybridization. This

yielded a gene, Cyp7b, whose expression is largely but not exclusively restricted to the hippocampal formation (Stapleton et al., 1995) and encoding a novel enzyme metabolizing neurosteroids (Rose et al., 1997). Subtractive hybridization, though technically more demanding, highlighted clone H4M10 (Pickard, 1996) that turned out to be identical to hippocalcin, a gene whose expression was previously suggested to be specific to the hippocampus (Koboyashi et al., 1992). However, while enriched in the hippocampus, brain expression outwith the formation was also detected.

The candidate gene family approach yielded a new serine protease, BSP1 (Davies et al., 1998), with a highly restricted pattern of expression that will permit exploitation in a transgenic context, and protein tyrosine phosphatase (PTP) gamma, whose restricted pattern of expression warrants further study (Pickard, 1996). Finally, gene-trapping uncovered an insertion, *kin*, directing reporter gene expression almost exclusively to hippocampus (Steel et al., 1998). However, *kin* transcripts were expressed widely in brain, but without detectable reporter activity, suggesting that the fusion transcript is widely expressed but only translated in the hippocampus. A more recent gene-trap line, *Ex-194*, has identified a gene whose expression is highly restricted to the dentate gyrus (C. Jarvis and W. Skarnes, pers. comm.), underlining the power of the gene-trap approach.

The two latter techniques (candidate gene families and gene-trapping) owe much of their success to serendipity. There are emerging suggestions that the hippocampus is an unusually rich site of gene expression (Steel et al., 1998); further experiments will be required to address whether these approaches can be applied to other brain regions. The gene-trap approach does offer one important advantage. Because gene-trap insertions tend to be near the start of the trapped gene, the technique can identify alternative promoter sites whose activity is restricted to a given tissue. This is illustrated by the *kin* insertion that identifies an alternative 5' promoter/exon of the targeted gene (Steel et al., 1998).

Despite the power of the two former techniques (differential and subtractive hybridization), as employed here, the differential complexity of the tissue samples may be a complicating factor. Here we compared gene expression of one small brain region, the hippocampus, with that of the entire brain devoid of hippocampus ('remainder of brain'). One improvement we suggest would be to perform first differential screening or subtraction using tissues of more similar complexity (for instance hippocampus versus cortex or cerebellum).

The most serious problem encountered is that many of the screens detect cDNAs from genes that are widely expressed. Although differential splicing, polyadenylation site selection, and promoter utilization, can all yield short cDNA segments whose abundance differs between different tissues, this was not a major complicating factor. In contrast, and for reasons we do not understand, cDNAs were routinely detected that are differentially represented in the starting cDNA populations but, on Northern analysis, did not differ significantly in their levels of expression between the different

tissues examined (in this case hippocampus and other brain regions). This result is not understood, but we surmise that any post-transcriptional modification of an mRNA molecule that might differentially affect its representation in cDNA will reveal itself artefactually as a tissue-specific cDNA. We have considered several complex explanations, for instance, the possibility that tissue-specific complementary sequences could hybridize to an mRNA and prevent first strand cDNA synthesis, so producing biased representation of cDNA sequences.

A further caveat concerns the in situ hybridization technique as a means to confirm regional specificity. In the rodent hippocampus the cell bodies are very tightly aligned, while elsewhere in brain are often more evenly distributed. Thus, in situ hybridization, at short exposure times, tends to highlight the hippocampus. Caution is warranted in the interpretation of such experiments: even widely expressed genes can appear to be 'specific' to the hippocampus. Indications of hippocampus-specificity from in situ hybridization must be considered unproven until confirmed by other techniques including Northern analysis.

Our analyses are also limited by their bias towards well-expressed genes. Partly this is an advantage, because genes expressed poorly are less attractive for transgenic exploitation, but at the same time region-specific genes that are expressed at relatively low levels, such as embryo brain kinase (*ebk*), may be missed. *ebk* expression is widespread during development becomes progressively more restricted to the hippocampus during maturation (Ellis et al., 1995). However, our preliminary experiments failed to confirm this specificity (not presented).

Other procedures that deserve consideration include differential display (Liang and Pardee, 1992). We attempted to use this method but found many artefacts: we estimate that a transcript whose abundance differs only by a factor of 5 between two different tissues will yield a candidate 'tissue-specific' band on display. However, differential display in combination with subtractive hybridization may offer a more powerful approach than provided by either technique alone.

For the future two further techniques may be envisaged. The first, 'virtual subtraction', reflects the increasing availability of databanks of randomly sequenced cDNA products from different tissues. This will permit 'subtraction' between two sequence banks, pinpointing sequences abundant in one tissue but sparsely represented in the other. The second involves the use of gridded cDNAs or expressed sequences. In principle, all the genes in the mammalian genome can be represented as 100,000 unique cDNA sequences. With emerging technologies these could be arrayed at very high density on filters for hybridization. Parallel hybridization with cDNA probes prepared from two (or more) different tissues will identify sequences expressed in one tissue but not another.

Finally, the question arises - do truly region-specific genes exist? On the one hand, it may be argued that an anatomically distinct region such as the hippocampus must require specific gene expression. The hippocampus could be, in terms of gene expression, as distinct from cortex or

cerebellum as liver is different from lung. On the other hand, it is likely that all genes are expressed at some level in all tissues, and no gene is 100% tissue-specific. However, even if not totally restricted to a given tissue, a gene expressed 100-fold more strongly in one tissue or brain region than another will still be of great utility. Further, we have uncovered examples where developmental expression is widespread, but later becomes restricted to the hippocampus; the exploitation of such genes will best be performed in combination with inducible systems such as conditional ablation or drug-activated transcriptional activators and recombination enzymes.

Despite these reservations our analysis has pinpointed genes whose exploitation in a transgenic context is underway (work in progress). It may be hoped that similar approaches may be employed to identify genes whose expression is restricted to other brain structures.

Acknowledgements

This project was part funded by the MRC (to RL and R.G.M. Morris), a BBSRC/CGR bursary to BP and by MRC studentships to GS and BD, by a Charitable Foundation, and by BBSRC funding to the Centre for Genome Research. We thank R.G.M. Morris for helpful suggestions on the manuscript.

References

Abeliovich, A., Chen, C., Goda, Y., Silva, A.J., Stevens, C.F, and Tonegawa, S. (1993a). Modified hippocampal long-term potentiation in PKC gamma-mutant mice Cell 75: 1253-1262.

Abeliovich, A., Paylor, R., Chen, C., Kim, J.J., Wehner, J.M., and Tonegawa, S. (1993b). PKC gamma mutant mice exhibit mild deficits in spatial and contextual learning. Cell 75: 1263-1271.

Balzer, H.J. and Baumlein, H. (1994). An improved gene expression screen. Nucleic Acids Res. 22: 2853-2854.

Barnea, G., Silvennoinen, O., Shaanan, B., Honegger, A.M., Canoll, P.D., D'Eustachio, P., Morse, B., Levy, J.B., Laforgia, S., Huebner, K., et al (1993). Identification of a carbonic anhydrase-like domain in the extracellular region of RPTP gamma defines a new subfamily of receptor tyrosine phosphatases. Mol. Cell. Biol. 13: 1497-1506.

Berod, A., Biguet, N.D., Dumas, S., Bloch, B., and Mallet, J. (1987). Modulation of tyrosine hydroxylase gene expression in the central nervous system visualized by in situ hybridization. Proc. Natl. Acad. Sci. USA 84: 1699-1703.

Bliss, T.V., and Collingridge, G.L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361: 31-39.

Bliss, T.V., and Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J. Physiol. 232: 331-356.

Cohen, N.J. and Eichenbaum, H (1993). Memory, Amnesia and the Hippocampal System. MIT Press, Cambridge, MA.

Chen, C. and Tonegawa, S. (1997). Molecular genetic analysis of synaptic plasticity, activity-dependent neural development, learning, and memory in the mammalian brain. Annu. Rev. Neurosci. 20: 157-184.

Davies, B. J. (1998). PhD thesis, The University of Edinburgh.

Davies, B.J., Pickard, B.S., Steel, M., Morris, R.G.M. and Lathe, R. (1988) Serine proteases in rodent hippocampus. *Submitted.*

Dobie, K.W., Mehtali, M., McClenaghan, M., and Lathe, R. (1997). Variegated gene expression in mice. Trends Genet. 13: 127-130.

Ellis, J., Liu, Q., Breitman, M., Jenkins, N.A., Gilbert, D.J., Copeland, N. G., Tempest, H.V., Warren., S., Muir, E., Schilling, H., Fletcher, F.A., Ziegler, S.F., and Rogers, J.H. (1995). Embryo brain kinase: A novel gene of the eph/elk receptor tyrosine kinase family. Mech. Dev. 52: 319-341

Frohman, M.A., Dush, M.K., and Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA 85: 8998-9002.

Gebbink, M.F., van Etten, I., Hateboer, G., Suijkerbuijk, R., Beijersbergen R.L., Geurts van Kessel, A. and Moolenaar, W.H. (1991). Cloning, expression and chromosomal localization of a new putative receptor-like protein tyrosine phosphatase. FEBS Lett. 290:123-130.

Hampson, I.N., Pope, L., Cowling, G.J., Dexter, T.M. (1992). Chemical cross linking subtraction (CCLS): a new method for the generation of subtractive hybridisation probes. Nucleic Acids Res. 20: 2899.

Jiang, Y.P., Wang, H., D'Eustachio, P., Musacchio, J.M., Schlessinger J. and Sap J. (1993). Cloning and characterization of R-PTP-kappa, a new member of the receptor protein tyrosine phosphatase family with a proteolytically cleaved cellular adhesion molecule -like extracellular region. Mol. Cell. Biol. 13: 2942-2951.

Kobayashi, M., Takamatsu, K., Saitoh, S., Miura, M., Noguchi, T. (1992). Molecular cloning of hippocalcin, a novel calcium-binding protein of the recoverin family exclusively expressed in hippocampus. Biochem. Biophys. Res. Commun. 189: 511-517.

Krueger, N.X., Streuli, M. and Saito, H. (1990). Structural diversity and evolution of human receptor-like protein tyrosine phosphatases. EMBO J. 9: 3241-3252.

Kruh, G.D., Perego, R., Miki, T., and Aaronson, S.O. (1990). The complete coding sequence of arg defines the Abelson subfamily of cytoplasmic tyrosine kinases. Proc. Natl. Acad. Sci. USA 87: 5802-5806.

Lathe, R. (1996). Mice, gene targeting and behaviour; more than just genetic background. Trends Neurosci. 19: 183-186.

Lathe, R. and Morris, R.G.M. (1994). Analysing brain function and dysfunction in transgenic animals. Neuropathol. Appl. Neurobiol. 20, 350-358.

Li, Y., Erzurumlu, R.S., Chen, C., Jhaveri, S. and Tonegawa, S. (1994). Whisker-related neuronal patterns fail to develop in the trigeminal brainstem nuclei of NMDAR1 knockout mice. Cell 76: 427-437

Liang, P. and Pardee, A.B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971.

Mansuy, I.M., Mayford, M., Jacob, B., Kandel, E.R., and Bach, M.E. (1995) Restricted and regulated overexpression reveals calcineurin as a key component in the transition from short-term to long-term memory. Cell 92: 39-49.

Mayford, M., Abel, T., and Kandel, E.R. (1995). Transgenic approaches to cognition. Curr. Opin. Neurobiol. 5: 141-148.

McHugh, T.J., Blum, K.I., Tsien, J.Z., Tonegawa, S. and Wilson, M.A. (1996). Impaired hippocampal representation of space in CA1-specific NMDAR1 knockout mice. Cell 87: 1339-1349.

Morris, R.G.M, Garrud, P., Rawlins, J.N.P. and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal

lesions. Nature 297: 681-683.

Morris, R.G., Anderson, E., Lynch, G.S. and Baudry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. Nature 319: 774-776.

Oyler, G.A., Higgins, G.A., Hart, R.A., Battenberg, E., Billingsley, M., Bloom, F.E. and Wilson, M.C. (1989). The identification of a novel synaptosomal-associated protein, SNAP-25, differentially expressed by neuronal subpopulations. J. Cell Biol. 109: 3039-3052.

Pickard, B.S. (1996) PhD thesis, The University of Edinburgh.

Roberson, E.D, English, J.D. and Sweatt, J.D. (1996). A biochemist's view of long-term potentiation. Learning Memory 3: 1-24.

Rose, K.A., Stapleton, G., Dott, K., Kieny, M.P., Best, R., Schwarz, M., Russell, D.W., Björkhem, I., Seckl, J. and Lathe, R. (1997). Cyp7b hydroxylates neurosteroids dehydroepiandrosterone (DHEA) and pregnenolone at the 7alpha Position. Proc. Natl. Acad. Sci. USA 94, 4925-4930.

Saitoh, S., Takamatsu, K., Kobayashi, M., Noguchi, T. (1993). Distribution of hippocalcin mRNA and immunoreactivity in rat brain. Neurosci. Lett. 157: 107-110.

Sakimura, K., Kutsuwada, T., Ito, I., Manabe, T., Takayama, C., Kushiya, E., Yagi, T., Aizawa, S., Inoue, Y., Sugiyama, H., et al. (1995). Reduced hippocampal LTP and spatial learning in mice lacking NMDA receptor epsilon 1 subunit. Nature 373: 151-155.

Scoville, W.B., and Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiat. 20: 11-21.

Shivers, B.D., Hilbich, C., Multhaup, G., Salbaum, M., Beyreuther, K. and Seeburg PH. (1988). Alzheimer's disease amyloidogenic glycoprotein: expression pattern in rat brain suggests a role in cell contact. EMBO J. 7: 1365-1370.

Silva, A.J., Paylor, R., Wehner, J.M., and Tonegawa, S. (1992a). Impaired spatial learning in alpha-calcium-calmodulin kinase II mutant mice. Science 257: 206-211.

Silva, A.J., Stevens, C.F., Tonegawa, S., and Wang, Y. (1992b). Deficient hippocampal long-term potentiation in alpha-calcium-calmodulin kinase II mutant mice. Science 257: 201-206.

Skarnes, W.C., Auerbach, B.A., and Joyner, A.L. (1992). A gene trap approach in mouse embryonic stem cells: the lacZ reporter is activated by splicing, reflects endogenous gene expression, and is mutagenic in mice. Genes. Dev. 6: 903-918.

Stapleton, G. (1994). PhD thesis, University of Edinburgh.

Stapleton, G., Steel, M., Richardson, M., Mason, J.O., Rose, K.A., Morris, R.G.M., and Lathe, R. (1995). A novel cytochrome P450 expressed primarily in brain. J. Biol. Chem. 270: 29739-29745.

Steel, M., Moss, J., Clark., K.A., Kearns, I.R., Davies, C.H., Morris, R.G.M., Skarnes, W.C. and Lathe, R. (1998). Gene-trapping to identify and analyze genes expressed in the mouse hippocampus. *Submitted.*

Sung, L.A., Fowler, V.M., Lambert, K., Sussman, M.A., Karr, D. and Chien, S. (1992). Molecular cloning and characterization of human fetal liver tropomodulin. A tropomyosin-binding protein. J. Biol. Chem. 267: 2616-2621

Takahashi, Y. (1992). Gene expression in cells of the central nervous system. Prog. Neurobiol. 38: 523-569.

Tsien, J.Z., Chen, D.F., Gerber, D., Tom, C., Mercer, E.H., Anderson, D.J., Mayford, M., Kandel, E.R. and Tonegawa, S. (1996a). Subregion- and cell type-restricted gene knockout in mouse brain. Cell 87: 1317-1326.

Tsien, J.Z., Huerta, P.T. and Tonegawa, S. (1996b). The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. Cell 87: 1327-1338.

Wang, Z., and Brown, D.D. (1991). A gene expression screen. Proc. Natl. Acad. Sci. USA 88: 11505-11509.

Zigman, J.M., Westermark, G.T., LaMendola, J., Boel, E. and Steiner DF. (1993). Human G(olf) alpha: complementary deoxyribonucleic acid structure and expression in pancreatic islets and other tissues outside the olfactory neuroepithelium and central nervous system. Endocrinology 133: 2508-2514.

Table 1. Differential hybridization

Clone	transcript sizes	identity	expression	
			HPC	ROB
3.22a	4.8	novel	++	(+/-)
3.28a	nd[1]	PDGF receptor[1]	nd[1]	
3.3a	2.4	novel	++	++
4.5a		mitochondrial gene		
4.5b	1.6	novel	++	++
6.19d		ß-adaptin		
9.9a		calreticulin		
11.2a	2.9	mouse ECA39	++	+
12.10a	7[2], 3.9, 3.5, 2.7, 1.9	novel; homology to human fetal liver tropomodulin	++[2]	(+/-)
13.9c		ribosomal protein S13		
14.4a		calmodulin RCM3		
14.5a	5.0, 2.0, 1.8	novel, dubbed Cyp7b	++	(+/-)
14.15a		rat amyloidogenic glycoprotein (rAG)		
14.15b	2.25	novel	++	++
15.13a	4.4[3], 3.2, 2.5	novel	++[3]	(+)
16.26b		SNAP-25		
16.26c	1.8	novel	++	(+/-)

[1]For known genes the transcript sizes and expression patterns were not determined (nd); [2]the 7.0 kb transcript of 12.10a was specific to the hippocampus; [3]the 4.4 kb transcript corresponding to 15.13a was specific to the hippocampus.

Table 2. Subtractive hybridization

Clone	Identity	cDNA southern		Northern	
		HPC	ROB	HPC	ROB
H4M6,22	novel	++	(+)	nd	
H4M10	hippocalcin	+++	-	+++	+/-
H4M19	novel	+	(+)	nd	
H4M21,25	novel	+++	(+)	nd	
H4M27	Na/H exchange protein 1	nd		nd	
H4M28,29	novel	+	(+)	++	++
H4M33	5-HT1a receptor	nd		nd	
H4M35	novel	++	-	++	++
H4M37	hEST T04934	++	(+)	nd	

nd, not determined

Table 3. Conserved motifs for candidate gene PCR amplification

	Sequence motif 1	Sequence motif 2
1) G protein-coupled receptors	LCVIALDRY	FILCWLPFF
	NLAVADL S L EF	FALCWLPYH VI LQ FV N F
2) Metabotropic glutamate receptors	SGEELSY S	TMYTTC
3) G protein a subunits	RLLLLG K	DVGGQR
4) Protein tyrosine phosphatases	KC-QYWP	HCSAGAG I
	FWRMVIWE I D Q	
5) Serine proteases	VLTAAHC	GDSGGPLV

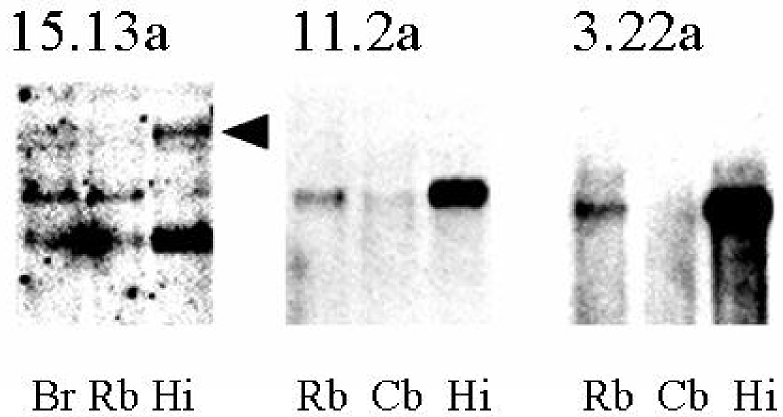

Figure 1. Northern analysis of clones identified by differential hybridization screening. RNA samples were: Br, brain; Rb, rest of brain (ROB), Hi, hippocampus (HPC). 15.13a identifies several transcripts including one specific to hippocampus (4.4 kb, arrowed).

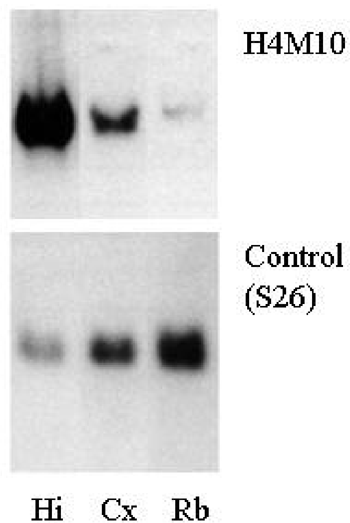

Figure 2. Northern analysis of clone H4M10 identified by subtractive hybridization. RNA samples were Hi, hippocampus (HPC), Cx, cortex; Rb, rest of brain (ROB). The control probe corresponded to ribosomal protein S26, revealing overloading in two lanes that emphasises the hippocampal specificity of H4M10.

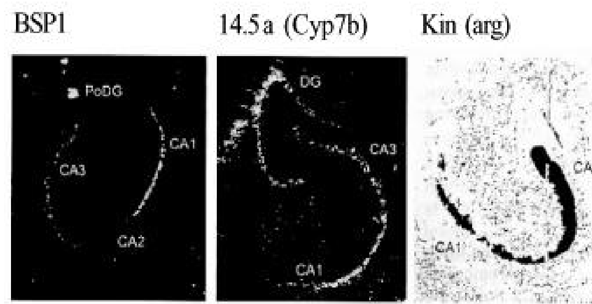

Figure 3. Regional localization of the expression of three transcripts identified in this study. A, in situ hybridization (dark field) to rat brain of novel serine protease BSP1; CA1-3 and are hippocampal subregions, PoDG are polymorphic cells of the dentate gyrus. B, in situ hybridization of the novel steroid metabolizing enzyme Cyp7b to mouse brain showing expression throughout CA1-3 and DG. C, staining (light field) of the hippocampus of *kin* animals with the chromatogenic ß-galactosidase (ß-geo) substrate X-gal, showing highly specific coloration in CA1-3.