Title: Hippocampal gene profiling: Toward a systems biology of the hippocampus

Authors and affiliation: Luis M. Valor and Angel Barco

Instituto de Neurociencias de Alicante (Universidad Miguel Hernández-Consejo Superior de Investigaciones Científicas). Campus de Sant Joan. Apt. 18. Sant Joan d'Alacant. 03550. Alicante, Spain

Abbreviated title: Hippocampal transcriptomics

Number of pages: 34

Number of figures and tables: 3 / 1

Corresponding authors: Angel Barco, Instituto de Neurociencias de Alicante (UMH-CSIC), Campus de Sant Joan, Apt. 18, Sant Joan d'Alacant 03550, Alicante, Spain. Phone: +34 965 919232. Fax: +34 965 919492. Email: abarco@umh.es. Luis M Valor, Instituto de Neurociencias de Alicante (UMH-CSIC), Campus de Sant Joan, Apt. 18, Sant Joan d'Alacant 03550, Alicante, Spain. Phone: +34 965 919531. Fax: +34 965 919492. Email: lmv@umh.es

Keywords: Transcriptomics, proteomics, microarray, functional genomics, hippocampus

Abstract

Transcriptomics and proteomics approaches give a unique perspective for understanding brain and hippocampal functions but also pose unique challenges because of the singular complexity of the nervous system. The proliferation of genome-wide expression studies during the last decade has provided important insight into the molecular underpinnings of brain anatomy, neural plasticity and neurological diseases. Microarray technology has dominated transcriptomics research, but this situation is rapidly changing with the recent technological advances in high-throughput sequencing. The full potential of transcriptomics in the neurosciences will be achieved as a result of its integration with other "-omics" disciplines as well as the development of novel analytical bioinformatics and systems biology tools for meta-analysis. Here we review some of the most relevant advances in the gene profiling of the hippocampus, its relationship with proteomics approaches, and the promising perspectives for the future.

Introduction

An array of 'omics' terms has bloomed in the biomedical jargon in the last decade, including genomics, epigenomics, transcriptomics, proteomics, phenomics and metabolomics to name a few examples. The impact of the 'omics' era is noticeable in almost every research field, and the neurosciences are not an exception. Neuroscientists are starting to implement global strategies to answer traditional but still far from well-understood questions, such as how neuronal circuits are formed during embryonic development, how a memory is formed at the molecular and cellular levels, how a neurodegenerative process is ignited in the neuron, or how the environment can predispose an individual to behave in a particular manner. The focus on individual genes has provided enormous success in the past and still represents the main stream of research, but it is also now showing its limitations. Physiologists long ago reached the conclusion that the brain cannot be explained in terms of individual cells but, rather, in terms of neuronal circuits and cellular ensembles. Similarly, molecular biologists during the last decade have started to address molecular processes as a whole rather than focus on specific genes, leading to the prominent emergence of systems biology. Not a single gene product but a complex gene program accounts for the response to a particular environmental stimulus or intracellular condition. In consequence, the need for global approaches to define the molecular mechanisms involved in brain processes is becoming more widely acknowledged. Such approaches permit the interrogation of any specific experimental perturbation with the maximum possible coverage and minimal researcher's bias. A prime manifestation of this shift is the growing relevance of a relatively new trend in molecular research, as part of systems biology, called functional genomics, which attempts to describe the global dynamics of gene expression and protein networks (Geschwind and Konopka, 2009). In this review, we will give a brief

overview of transcriptomics and functional genomics studies conducted in the hippocampus, and we will discuss the main strategies for interpreting the resulting lists of genes.

Tackling the neural transcriptome

Similar to the molecular biology revolution of a few decades ago, the neurosciences have been lagging behind other biomedical areas, such as cancer biology, in the adoption of transcriptomics and functional genomics approaches (Figure 1A). Some reasons for this are the extreme cellular heterogeneity and complexity of the neural tissue; the difficulty, in the case of human studies, in accessing non-fixed brain tissue in the appropriate developmental or disease stage; the general preference of neuroscientists for hypothesis-driven rather than discovery-based approaches; and the relative lower funding that may have in the past delayed the early access to these novel technologies (see reviews by (Geschwind and Konopka, 2009; Nisenbaum, 2002). Still, the ability to monitor in parallel the expression of tens of thousands genes is extremely attractive and neuroscientists have now widely adopted these powerful technologies.

Diverse methods have been used for transcriptomic profiling in the brain. Microarray technology, compared with other transcriptomics technologies (Figure 1B), is by far the most popular approach because of its early automation and commercial availability at an affordable cost. The early reports indicating poorly overlapping results between different microarray platforms and research groups (Kothapalli et al., 2002; Kuo et al., 2002; Tan et al., 2003), have been largely surmounted by constant improvements in the design of array platforms (de Reynies et al., 2006). The most reliable platforms dominate the market, the public repositories of gene expression profiles are exponentially growing (Table 1), and we have entered a new phase of

reduction in costs and customization. Thousands of experiments have demonstrated good general agreement with alternative measurements of transcript levels using single-gene techniques, like in situ hybridization (ISH) and quantitative RT-PCR (see (de Reynies et al., 2006; Mieczkowski et al., 2010) for specifically designed comparative studies).

This does not mean that transcriptomics does not have important challenges ahead. Microarray technology still confronts some technical issues, like cross-hybridization and the use of fluorescent/luminescent dyes that influence the dynamic range of detection. Other important concerns, not directly related to the technology itself, are the frequently poor descriptions of experimental and analytical procedures and the limited open access to raw datasets and analytical proceedings, which hamper repeatability (Ioannidis et al., 2009). Additionally, the strong dependence on valid gene annotation in databases limits transcriptome surveys (Asmann et al., 2008). Until recently, most commercial arrays for mammalian samples only interrogated a fraction of the transcriptome, frequently excluding many neuronally specific expressed genes, as a consequence of a bias in public databases toward more abundant genes for which more information has been compiled. Indeed, the need to detect less abundant neuronal transcripts has fueled the use of alternative technologies (Asmann et al., 2009; Li et al., 2004). The new generations of arrays can detect rare transcripts and even alternatively spliced mRNA isoforms and non-coding RNAs, permitting a more comprehensive analysis of the brain transcriptome. This is particularly important because breakthrough discoveries in the last decade have revealed that the transcriptome world is much more complex than expected (Carninci et al., 2005; Fejes-Toth et al., 2009; Kim et al., 2010a).

Although microarrays have dominated the field since 2000, the situation may be changing. Technological advances in the last few years have allowed the strong re-emergence of sequencing-based technologies. Although the first attempts to tackle the neural transcriptome were based on large-scale sequencing, they required laborious cDNA library construction and systematic clone sequencing, and in practical terms they were expensive, time-consuming and often accessible only to a few research groups (e.g., serial analysis of gene expression (SAGE), sequencing of subtraction libraries, and massive parallel sequencing (MPSS); for a brief description see (Ginsberg and Mirnics, 2006; Valor and Grant, 2007)). In contrast, the novel next-generation sequencing platforms allow sequencing of hundreds of millions of base pairs in a single reaction within a few hours at accessible cost (Mardis, 2008). For instance, the development of RNA-seq (Mortazavi et al., 2008) makes it possible to directly sequence a pool of pre-processed cDNA fragments using little starting material for a few thousand dollars, and it is reasonable to assume that the costs will be reduced even more in the near future. RNA-seq technology presents two important advances compared with microarrays: 1) it does not require previous knowledge of the transcriptome and allows a fully unbiased survey, and 2) it produces a digital count of transcript abundance rather than an analogue non-linear luminescent or fluorescent dye signal. Sequencing-based technologies also open up other interesting methods of investigation into the regulation of the transcriptome that were previously unavailable, such as sequencing-based methylation analysis and ChIP-seq (chromatin immunoprecipitation high-throughput sequencing) (see "Linking transcriptomics with proteomics" section).

We believe that after nearly 15 years, through the work of many authors, transcriptomic techniques have achieved their promise of becoming an essential tool in the neurosciences (Dougherty and Geschwind, 2005). The hippocampus represents a

prime example of the early pitfalls and limitations as well as the current potential for transcriptomics studies in the nervous system. The simultaneous analysis of the expression of tens of thousands of genes can offer access to issues of great complexity, such as the molecular determinants of hippocampal cell identity, the specification of hippocampal circuits and their modification by experience, and the molecular alterations determining the onset and progression of neurological diseases that affect this brain region. In the next sections, we will review some representative experiments and discuss the most recent advances on these questions.

Gene profiling in the hippocampus: from gene catalogues to functional genomics

1. From atlas of expression to circuit ontogeny and evolution

The elaboration of a comprehensive atlas of brain gene expression, including the hippocampus, would represent an important advance toward understanding brain physiology and connectivity. Two main strategies have been used in the construction of such an atlas: (a) high-throughput in situ hybridization (ISH) and (b) microarray analysis of different brain areas (Figure 2).

The first approach is not affordable for a single laboratory. Either a consortium or a dedicated institution is needed to generate such a volume of anatomical data and to create and maintain a public database that could provide access to this dataset. At the moment there are several public repositories, such as the Allen Brain Atlas, or ABA (Jones et al., 2009; Lein et al., 2007), GenePaint (Visel et al., 2004), Brain Gene Expression Map, or BGEM (Magdaleno et al., 2006), Gene Expression Nervous System Atlas, or GENSAT (Gong et al., 2003), in which ISH images for several thousand transcripts can be inspected and downloaded (Table 1). With the exception of BGEM, these databases do not use radioactive riboprobes, which provide better linearity in the

signal than colorimetric-based probes (Carter et al., 2010), and therefore most of these ISH data are qualitative or at best semi-quantitative. Despite this limitation, high-throughput ISH is extraordinarily useful in understanding the connectivity between brain areas since it provides detailed spatial information. As a proof of principle, ABA data have been recently used to identify functional subdomains in the hippocampus and the molecular determinants of their differential connectivity to other brain areas (Dong et al., 2009; Thompson et al., 2008). It should be noted, however, that because of the very high cell density in the hippocampus, many neuronal genes seem more highly expressed in this area than in other brain regions (Figure 2A). This can make the semi-quantitative comparison of gene expression levels in colorimetric ISH unreliable.

In contrast, microarray technology provides precise quantitative data and makes genome-wide studies more accessible and cost affordable for individual laboratories. As a result, there have been several attempts to create gene expression maps based on this technology, either for the whole organism (Bono et al., 2003; Su et al., 2004; Zhang et al., 2004), the whole brain, or restricted to the hippocampus. Different sampling strategies have been used, from gross hippocampus dissection (Sandberg et al., 2000) to controlled microdissection (typically of CA1, CA3 and DG) (Datson et al., 2004; Lein et al., 2004; Zapala et al., 2005; Zhao et al., 2001), voxelation, which does not consider natural boundaries among brain areas (Brown et al., 2002), laser-capture microdissection (LCM) (Bonaventure et al., 2002; Kamme et al., 2003) and cell sorting of fluorescently labeled neurons (Lobo et al., 2006; Sugino et al., 2006). Overall, these studies show that an increase in sampling resolution is extremely important to properly discriminate cellular subpopulations and, in consequence, functional anatomical units. Electrophysiological recordings have demonstrated that even adjacent cells with similar morphologies can have remarkably different electrical properties. Thus, it is possible

that small, though biologically relevant, changes in gene expression in a specific neuronal population can be overlooked due to a dilution effect from the mRNA pool from the surrounding tissue. In the case of the nervous system, proper interpretation of the gene profiling results frequently requires validation by methods that provide cellular resolution to determine whether the changes affect specific cell types or subpopulations.

Given the remarkable properties of the nervous system, the ultimate challenge would be to define the transcriptome of small neuronal ensembles or even single cells. A recent and elegant approach to address the issue of cellular heterogeneity in transcriptomic studies in the central nervous system (CNS) has been presented recently by Heintz's team. They genetically labeled the ribosomes from specific neuronal populations with an antigenic tag that enables affinity purification of polysomal mRNAs for microarray experiments. Consequently, only the transcripts engaged in translation in the targeted cellular population were examined (Heiman et al., 2008). The utility of this approach, referred to as TRAP (translating ribosome affinity purification), was illustrated by the comparative analysis of 24 CNS cell populations, revealing hundreds of differentially expressed genes and identifying cell-specific and enriched transcripts for each population (Doyle et al., 2008). The improvement in the anatomical resolution of gene expression changes may result not only from refinements in the sampling procedure but also in data processing. Thus, a recent bioinformatics analysis of brain single-cell transcriptomes has been able to resolve two transcriptionally distinct mitochondrial modules, one enriched in neuronal processes and synapses and another one restricted to the cell body (Winden et al., 2009).

A major conclusion from the analysis of regional brain transcriptomes is that gene expression differences between brain regions are generally relative (differential levels of transcripts) and not absolute (presence or absence of transcripts) (Gray et al.,

2004; Sandberg et al., 2000; Zapala et al., 2005). This observation can explain why the intense initial search for highly specific promoters for different brain regions may have been less successful than expected. The identity of particular brain regions or neuronal circuits seems to be determined more by unique combinations of gene expression levels, defining a "regional-specific signature", rather than by the expression of unique genes.

After retrieving a catalogue of regionally expressed genes, the next step is to find a biological meaning for such regional variation. Interestingly, it is possible to establish correlations between anatomical patterns of expression with developmental and evolutionary histories. Thus, the expression atlas generated by Zapala and colleagues showed that adult expression profiles are a reflection of the embryonic origin of the tissue and revealed that developmental patterning genes, such as the Hox genes, were significantly enriched among the genes that exhibited strong expression differences between distinct brain areas (Zapala et al., 2005). The study by Sugino and colleagues found a high degree of heterogeneity among neuronal subpopulations affecting genes related to synaptic and carbohydrate metabolism functions. In this group of genes, there was an over-representation of paralog genes, suggesting that gene duplication may be a determinant for cell type diversification (Sugino et al., 2006). Another recent study explored the evolution of the synapse by merging data from proteomics and transcriptomics and concluded that the synaptic components with the most variable profiles among different brain areas correlated with a more recent history in evolution and were enriched in signaling pathway components (Emes et al., 2008). In conclusion, these studies suggest that evolution and ontogeny are primary forces in determining "regional-specific signatures" that could bring specific functionalities to each brain area.

2. From neuronal and synaptic markers to functional circuit assembly

During the development of the nervous system, the coordinated and sequential activation of genetic programs related to cell proliferation and differentiation, and synapse formation and maturation leads to the formation of extremely complex and precise neural circuitries with hundreds of millions of synaptic connections. A better understanding of the development of the hippocampus may result from the examination of these genetic programs. Early reports of gene expression profiling during the development of the hippocampus either *in vivo* (Mody et al., 2001) or in culture models (Dabrowski et al., 2003) provided a first glance into the complexity of the process, revealing the activation of genetic programs related to synaptic transmission, energetic metabolism, signal transduction, and transcriptional regulation among other functions.

Primary cultures of hippocampal neurons have been demonstrated to be a particularly accessible model to study synaptogenesis. These studies have taken different forms. For example, gene expression profiling studies have been used to identify novel proteins involved in synaptogenesis. After an initial screen using transcriptome-wide techniques, a functional test should be carried out to properly assign a specific role for the differentially expressed transcripts in the development of the synapse. Such functional assays can even take a large-scale approach. Thus, Paradis and colleagues coupled a microarray analysis of synapse development in postnatal hippocampus with an RNAi screening in hippocampal cultures (Paradis et al., 2007). By correlating changes in gene expression with effects on the postsynaptic density, the authors enlarged the list of synaptogenic proteins with five novel proteins, most of them belonging to the cadherin and semaphorin families.

Another viable strategy is to undertake a systems biology approach by integrating transcriptomics with information from other sources. As an example, the

study by Valor et al. combined transcriptional profiling using microarrays with electrophysiological recordings using multi-electrode arrays and morphological analysis using conventional immunocytochemistry techniques (Valor et al., 2007). The aim of this work was not to identify novel synaptogenic genes but to correlate transcriptional profiles with other properties of hippocampal cultures, such as network firing activity and synapse maturation. Nearly 20% of the core components of the synaptic machinery, as previously defined by mass spectrometry studies (Collins et al., 2006), exhibited prominent changes in transcription that preceded the coupling of pre- and postsynaptic markers and the onset of spiking activity. Continued expression was followed by maturation of morphology and electrical neuronal networks. Notably, only mature cultures exhibited activity-dependent gene expression, suggesting that the initial genetic program for synapse formation is independent of firing activity.

3. From immediate-early genes to transcriptome-driven neural remodeling

Arguably the most characteristic and remarkable feature of the nervous system is its plasticity. External stimuli can modify the membrane properties, intracellular signaling cascades and gene expression in individual neurons to modulate the responses to subsequent stimuli, changing the properties of the neuronal circuits in a durable manner. It is well-known that long-term synaptic plasticity requires *de novo* synthesis of both RNAs and proteins (Steward and Schuman, 2001). The classical description of the succession of events underlying long-term plasticity considers two waves of gene expression that can be dissected by biochemical means (Platenik et al., 2000). Immediately after stimulation, there is a first wave of expression consisting of the induction of so-called immediate early genes (IEGs), a diverse group of genes that were first identified by subtractive hybridization methods preceding the development of

microarray technology (Loebrich and Nedivi, 2009). Several of these IEGs are transcription factors and can therefore play an active role in the initiation of the second wave of gene expression. The effector molecules, responsible for the rearrangement of synaptic connectivity in response to experience, are produced as part of this complex transcriptional response (see (Abraham and Williams, 2008; Flavell and Greenberg, 2008; Loebrich and Nedivi, 2009). These events are thought to critically contribute to the acquisition and consolidation of hippocampus-dependent memory (Collingridge and Bliss, 1995). More than 20 years after the first reports on neuronal activity-driven transcription, a comprehensive and complete picture of the process in molecular terms is still lacking. How the external signal is propagated from the membrane (through the activation of receptors and channels) to the nucleus and how the nuclear activity in turn alters the future synaptic responses remain major outstanding questions in neurobiology.

To gain a global view of the activity-dependent genetic program, several studies have been conducted to generate genome-wide datasets of transcripts induced by different protocols of neuronal stimulation, such as the use of specific receptor agonists or antagonists, electrical stimulation, seizure induction or training in different learning tasks (Altar et al., 2004; Cavallaro et al., 2002; Coba et al., 2008; French et al., 2001; Havik et al., 2007; Hunsberger et al., 2005; Leil et al., 2002; Leil et al., 2003; Levenson et al., 2004; Matsuo et al., 1998; Pegoraro et al., 2010; Ploski et al., 2010; Ryan et al., 2010; Wibrand et al., 2006; Zhang et al., 2007). A glimpse of the compiled data indicates that dozens to hundreds of genes belonging to a vast variety of functional categories are affected (see for example (Cavallaro et al., 2002; Havik et al., 2007; Hong et al., 2004; Li et al., 2004; Zhang et al., 2009). Moreover, gene expression analysis demonstrates that the transcriptional response to a particular stimulus is highly dependent on its characteristics, including duration, drug administration or animal

training protocols, time of sampling, etc. These procedural differences could explain the apparent low concordance observed in some studies (Wang et al., 2009).

A successful strategy to extract more meaningful information from these sorts of transcriptomics data is to explore the role of specific upstream transcription factors in gene induction. Despite the high rate of false positives and false negatives, *in silico* prediction of transcription factors binding sites (TFBS) can provide interesting clues. It is possible to implement algorithms for TFBS predictions in the regulatory regions of the list of altered genes. Levenson and colleagues have applied this approach in the case of differentially expressed genes during memory consolidation in the hippocampus and have found an overrepresentation of binding motifs in their promoters for several transcription factors, including c-Rel (a member of the NF-κB family) whose association with memory consolidation was novel (Levenson et al., 2004). Further experiments in knockout mice confirmed the role of this transcription factor in hippocampus-dependent memory (Ahn et al., 2008) A recent microarray-based screen for long-term potentiation (LTP)-related genes in freely moving rats went further in its use of bioinformatics tools. Both TFBS prediction and network analysis of functionally related genes and chromosomal clustering of co-expressed genes were used to predict central roles for NF-κB, SRF, CREB and EGR1 in the stabilization of the LTP response. The analysis also revealed potential links with the MAPK signaling pathway, chromatin remodeling, local protein synthesis machinery and neurogenesis (Ryan et al., 2010).

A comparison with previous datasets accessible in public databases can also contribute to extracting additional meaningful information from genome-wide experiments. For example, Park and colleagues conducted a microarray study to determine the temporal gene expression profile in the mouse dentate gyrus in response to LTP induction in hippocampal slices and identified a few hundred genes that

responded to stimulation. The chromosomal location of these genes, which encoded proteins involved in various cellular processes including the structure and function of the synapse, revealed that they were grouped in highly conserved clusters, suggesting that genes located in the same cluster might be regulated by the same transcription factors. To support this hypothesis, the authors used the dataset on genome-wide CREB occupancy in response to activity generated using the serial analysis of chromatin occupancy (SACO) technique (Impey et al., 2004) and identified clusters of activity-regulated genes that were potentially regulated by CREB.

Transcription factors can act as integrative elements for upstream transduction signals originating in the postsynaptic membrane. A less explored approach to investigate activity-driven gene expression is to bypass the stimulus and try to identify directly the target genes of activity-regulated transcription factors (West et al., 2002), using loss- or gain-of-function approaches. Among the examined transcription factors, we can find CREB (Barco et al., 2005; Jancic et al., 2009; Lemberger et al., 2008; Valor et al., 2010), SRF (Etkin et al., 2006; Stritt et al., 2009), Npas4 (Lin et al., 2008) and MEF2 (Flavell et al., 2006; Flavell et al., 2008). However, using this approach, it is difficult to distinguish between direct and indirect targets because the genetic manipulation is generally chronic and may allow emergency of compensatory transcriptional effects that mask the direct effects of the inhibition, depletion or overexpression of the transcription factor. Novel techniques based on chromatin immunoprecipitation (ChIP) can provide a solution to these issues by precise genome-wide mapping of the binding sites for these transcription factors (Kim et al., 2010b; Tanis et al., 2008) (see “Linking transcriptomics with proteomics” section).

4. From markers of neuropathology to the molecular etiology of neurological and psychiatric diseases

Transcriptomics technology, led by microarray platforms, has the potential for commanding personalized medicine. In cancer research, gene expression profiling has been very successful for molecular diagnosis and prognosis because the presence of certain gene signatures ("disease signatures") can predict particular clinical outcomes (Strauss, 2006). Translated to the brain, transcriptome profiling can also represent an important asset in neuropathology studies given its potential to unveil the complex genetic and environmental interactions that contribute to neurological or psychiatric diseases and determine their progression (see extensive reviews by (Altar et al., 2009; Mufson et al., 2006) for specific references on transcripomics studies related to Alzheimer's disease (AD), schizophrenia, bipolar disorder, aging, etc). The ability to monitor whole transcriptomes has already allowed the identification of new biomarkers of disease in brain samples from patients and provided testable hypotheses for further study. Hypothesis-generating research is particularly important in this field because the etiology of many neural diseases remains elusive and appears multifaceted. Moreover, the discovery of most psychiatric drugs has been the result of serendipitous clinical observations and the empirical validation of their efficacy, but their molecular mechanisms of action remain largely unknown.

However, brain samples are much more difficult to obtain than tumor samples, given the difficult access to the diseased neural tissue. *Post mortem* brains represent the main source of samples, but their preservation conditions generally compromise RNA quality. Moreover, given the delay between diagnosis and death in many neural diseases, it is hard to evaluate whether the changes in gene expression observed in patients are related to the etiology of the disease, to its side effects or even to its

prolonged treatment. Although this issue severely constrains the clinical use of microarray technology in neurological and psychiatric diseases, modest success has been achieved using gene profiling of peripheral blood in Alzheimer's and Huntington's diseases, Tourette's syndrome, tuberous sclerosis and others (reviewed in (Sharp et al., 2006). These seminal studies suggest that bringing transcriptomics technologies to clinical labs in hospitals might allow neurologists and psychiatrists in the future to base diagnoses and prognoses on the comparison of expression profiles in samples from patients with defined "disease signatures", similar to what it is now possible in cancer research (Hadd et al., 2005).

An alternative and powerful approach to tackling neurological disorders is to investigate the transcriptional changes in animal models, which permit gene expression profiling in healthy, presymptomatic and terminal individuals. Such analyses lead to a description of the dynamics of the diseased transcriptomes that can be correlated with other distinctive features of the disease (cell loss, morphological changes, aberrant protein inclusions, altered behavior). The availability of an animal model also allows the rapid assessment of the efficacy of new therapeutic approaches for the treatment of the condition. An additional value of transcriptomics studies in animal models is the complementation of genome-wide association studies (GWAS) in creating the new discipline of integrative genomics. GWAS examine DNA variation in large human populations to discover genetic loci that influence the onset and/or the progression of the disease, but their results show poor correlation with gene expression changes (see reviews by (Le-Niculescu et al., 2007; Schadt, 2009). Animal models have the advantage of controlled genetic and environmental conditions (unlike human populations) that minimize the sources of gene expression noise and can provide

sufficient information to prioritize candidate targets to be extrapolated to patients (Le-Niculescu et al., 2007; Le-Niculescu et al., 2008).

Similar to the studies discussed in previous sections, further inspection of the expression profiles using bioinformatics tools can provide functional meaning and testable hypotheses. For example, a novel algorithm that assumes that co-expressed genes are functionally related (Weighted Gene Coexpression Network Analysis, or WGCNA) was applied to microarray data of the hippocampal CA1 subfield from AD patients. The algorithm was not only able to confirm a number of known key genes in AD but also identified novel ones as putative hubs (i.e., central and highly interconnected gene products that are predicted to be essential in a molecular network system). The study also organized the data into functional categories that pointed to mitochondrial and synaptic disruption as relevant events in the pathophysiology of the disease (Miller et al., 2008). More recently, a similar approach was used to compare the transcriptional network profiles of the human and the mouse brain, including hippocampal tissue, resulting in the definition of modules of gene co-expression that were preserved between the two species. Interestingly, the study identified among the transcriptional divergences a human-specific correlation of presenilin 1 with oligodendrocyte markers, and a human-specific module that strongly correlated with AD progression and contained several poorly characterized genes likely related to this disease (Miller et al., 2010). Such network modeling can be useful to clarify the complexities of the molecular dynamics in disease and to identify the most relevant targets for therapeutics in the molecular network.

Linking transcriptomics with proteomics: Toward a systems biology of the hippocampus

Despite the great advances in the field of high-throughput proteomics (reviewed in (Bayes and Grant, 2009; Chen and Yates, 2007; Deutsch et al., 2008; Kline et al., 2009), the generalized use of these techniques has been limited primarily due to the large dynamic range of protein abundance that makes quantification and time course analysis technically difficult (Deutsch et al., 2008; Kline et al., 2009), particularly if compared with microarray experiments. For most genes, there is a good correlation between transcript and protein levels (Cardozo et al., 2003; Chin et al., 2007; Kislinger et al., 2006; Mijalski et al., 2005). Moreover, interacting proteins or functionally related proteins tend to be co-expressed at the transcript level (Bhardwaj and Lu, 2005; Jansen et al., 2002; Oldham et al., 2006; Oldham et al., 2008; Winden et al., 2009), suggesting a coordinated regulation of certain processes at the transcriptional level. Consequently, most researchers have taken advantage of the larger accessibility of transcriptomics techniques to make reasonable inferences about the protein world by analyzing transcriptional profiling upon different experimental conditions.

Although transcriptomics and proteomics techniques frequently complement each other, there are unique applications for both approaches. Thus, whereas transcriptomics techniques have the important feature of examining non-coding RNAs, proteomics is required to investigate, in a global manner, the posttranslational modifications of pre-existing proteins that control many aspects of neuronal physiology. These modifications refer primarily to phosphorylations (Collins et al., 2005), but recent studies have also highlighted the relevance of other posttranslational modifications (Scheschonka et al., 2007; Sharma, 2010). Proteomics approaches are also needed to investigate protein turnover and subcellular localization because proteins can be purified from specific subcellular organelles or macromolecular complexes using immunoprecipitation and fractionation procedures, providing essential information

about the molecular composition of subcellular compartments. Thus, proteomic studies have provided the first opportunity to study the molecular composition of the synapse from a global perspective (Collins et al., 2006). Combined mass spectrometry and peptide array technology have revealed the complex combinatorial phosphorylation patterns of postsynaptic proteins triggered by the synaptic activation of specific neuronal receptors (Coba et al., 2009).

These studies have also demonstrated a high degree of interrelatedness between proteins of diverse functions, including different types of receptors, kinases, cytoskeletal components and scaffolding proteins among others. Hence, the activation of a particular receptor promotes the orchestrated modulation of a network of synaptic and extrasynaptic proteins that will initiate different downstream processes, including transcription (Pocklington et al., 2006). While the activation of the synaptic complexes triggers specific genetic programs in the cell nucleus, the molecular composition and activity of the complex itself would be, in turn, modulated by the activation of specific genetic programs, as discussed in previous sections. This dialogue between the synapses and the nucleus seems to be a fundamental mechanism underlying brain plasticity and cognitive functions and is a target of different disorders that affect the nervous system (Kandel, 2001; Valor and Grant, 2007). The combination of global transcriptomics and proteomics approaches is therefore essential for being able to *listen* to both parts of this dialogue and translate it into a meaningful interpretation of brain function.

Novel technical developments may pave the way in this direction. Chromatin immunoprecipitation (ChIP) allows the recovery of DNA from chromatin after crosslinking and precipitation, using specific antibodies against the transcription factor or DNA-interacting protein of interest. This opens up the possibility of mapping DNA-protein binding events in the genome accurately and quantitatively. The

immunoprecipitated DNA fragments can be analyzed in four ways: (1) by qPCR for single regulatory region studies (locus-specific ChIP assay), (2) by microarray technology to investigate the binding to a predetermined set of promoters displayed in the array (ChIP on chip), (3) by creating a library that can be sequenced (ChIP-SAGE or ChIP-SACO), and (4) by direct high-throughput sequencing (ChIP-seq), the latter being the most powerful approach to date (see (Schones and Zhao, 2008) for a detailed description of ChIP-related technologies). As an example of the power of this technique, a recent ChIP on chip analysis compared genome occupancy by total CREB and activated CREB (phosphoCREB) in different brain areas after electroconvulsive stimulation (Tanis et al., 2008). Whereas the number of promoters occupied by total CREB was similar in cortex, hippocampus and striatum, the number of loci occupied by phosphoCREB differed among regions, showing the lowest occupancy in striatum. Interestingly, phosphoCREB was bound to different promoters in the three regions, suggesting differential CREB functions throughout the brain.

Combining genome-wide occupancy mapping using ChIP for specific transcription factors with transcriptomics analysis allows for correlating gene transcription with local changes in the chromatin of promoters or regulatory sequences, providing unprecedented insight into gene expression mechanisms and DNA/RNA/protein interactions (Karlic et al., 2010; Visel et al., 2009; Welboren et al., 2009). A first attempt to apply this approach to the hippocampus has been recently reported (Peleg et al., 2010). In this study, the acquisition of contextual fear conditioning was correlated with an increase in the acetylation level of several histone residues and the induction of a complex gene expression program in the hippocampus. Importantly, the acquisition of the behavioral task was impaired in aged animals, which

also showed a dramatic attenuation of the acetylation of K12 in histone H4 in genes that were induced by the task in young animals, but not in the aged ones.

The full integration of proteomics and transcriptomics would require a detailed description of the upstream signaling pathways that activate each transcription factor in a given condition, a challenging task given the complex cross-talk and the combinatorial contribution of multiple signaling pathways to protein phosphorylation and gene expression (Coba et al., 2009; Coba et al., 2008; Michaelevski et al., 2010). As a first approach, merging occupancy data for several transcription factors can provide a full landscape of genome-wide gene regulation (Barski et al., 2007). A recent study described, for the first time, such a landscape in cortical cultures after membrane depolarization. The genome-wide information on the neuronal activity transcriptome obtained by RNA-seq was correlated with the genome-wide mapping of chromatin binding for several proteins: activity-regulated transcription factors (CREB, SRF and NPAS4), the acetyltransferase CBP, the RNA polymerase II and several histone tags (Kim et al., 2010b). Such descriptions have an enormous potential for changing our current views regarding gene regulation and genome organization. As a proof, the aforementioned work led to the discovery of a novel and unexpected RNA species associated with regulatory enhancers (eRNA), which likely play an important role in activity-driven gene expression.

Perspectives

Global approaches represent an important complement to the traditional reductionist single-gene studies. Whereas single-gene studies are hypothesis-driven, global approaches are hypothesis-generating or discovery-based and therefore have the potential to yield novel insight into brain function and dysfunction. Current techniques

allow the relatively easy production of a large volume of transcriptomics and proteomics data. Continuous technical improvements in combination with a reduction of costs, the exponential growth of public databases that archive and organize this information, and the development of novel bioinformatics tools for the analysis of large biological datasets will ensure a new era of progress characterized by the application of interdisciplinary global approaches to the neurosciences (Figure 3).

The major challenge for these global approaches is the proper interpretation and analysis of large descriptive datasets of candidate molecules. The lack of immediate experimental validation tools has frequently led to a crude description of genes catalogued by known function (Cao and Dulac, 2001). This is especially true for neurogenomic studies because public resources are strongly biased toward the most abundant gene products (overlooking the scarcely expressed neuronal genes), and functional annotations and pathway networks are biased toward extensively studied processes, such as cancer and cell cycle. As a symptom of the difficulty in data interpretation, many publications based on a genome-wide analysis first present a catalogue of altered transcripts or proteins and subsequently step down from the global to the single-gene level to describe in detail the role of one or a few molecules in the examined process or perturbation. Frequently, the selection of the gene for further study is based on previous knowledge, detracting therefore, at least partially, from the utility of the initial global approach. Fortunately, open access to global data may prevent the loss of disregarded information.

This situation has been changing in the last few years thanks to improvements both in the technology for the detection and quantification of changes in gene expression and in the bioinformatics tools available for their analysis. Transcriptomics studies can greatly benefit from data mining, meta-analyses and comparison with

datasets from other sources, especially from proteomics and single-protein studies, to formulate new hypotheses and discover emergent properties of the nervous system. Due to the prominent role of the hippocampus in cognition and disease and its relatively accessible structure, this brain region has become a prime target for systems biology-based studies of the nervous system that combine proteomic, transcriptomic and other "-omic" data toward the ambitious goal of understanding brain function.

Acknowledgements

We thank Eva Benito, Eloisa Herrera and Jose P. Lopez-Atalaya for a critical reading of the manuscript. Research at AB's lab is supported by the Spanish Ministry of Science and Innovation Grants BFU2008-00611, SAF2008-03194-E and CSD2007-00023 and grants from Fundación Ramón Areces and Fundació la Marató de TV3.

Figures and Tables

Figure 1: Citations of transcriptomics technologies in the literature from the year 1995 to the year 2009. A. Comparison of all of the microarray citations in the PubMed database (white squares) with cancer (black triangles) and hippocampal (black circles) specific citations. For the hippocampus-specific literature, the term "hippocampal" was used because it produced more hits than "hippocampus", probably because of the inclusion of neuronal culture studies. Citations of hippocampal studies have the same profile as the general publications but with a few years of delay, indicating the late implementation of this technology to research in the hippocampus compared to other fields. **B.** Search of the PubMed database using the terms indicated in the legend box. Only unequivocal terms were used, and reviews were excluded. In contrast to other transcriptomics technologies, the citations of microarray studies show an exponential profile.

Figure 2: Examples of brain atlases. A search for the AMPA receptor subunit transcript Gria1 was performed in the Allen Brain Atlas (**A**) and in repositories of microarray data (**B,** GEO Profile using the GDS1490 series: mean ± s.d.; **C,** ArrayExpress Gene Expression Atlas using the E-GEOD-4734 experiment: direct output from the web site). The numbers in **A** correspond to the regions indicated in the x-axis of **B**. High-throughput ISH and microarray analyses are complementary: ISH data provide spatial information, whereas microarray data provide more precise quantification of transcript abundance.

Figure 3: Gene profiling is key in systems biology studies of the hippocampus. Gene expression profiles are represented as heat maps of the genetic signatures for the connectome (full set of neural connections between different brain areas and over development), the proteome (full set of proteins and their interactions in networks), the

genome and epigenome (full set of genes and modifications of the chromatin) and the phenome (full set of phenotypes in an organism).

References

Abraham WC, Williams JM. 2008. LTP maintenance and its protein synthesis-dependence. Neurobiol Learn Mem 89(3):260-8.

Ahn HJ, Hernandez CM, Levenson JM, Lubin FD, Liou HC, Sweatt JD. 2008. c-Rel, an NF-kappaB family transcription factor, is required for hippocampal long-term synaptic plasticity and memory formation. Learn Mem 15(7):539-49.

Altar CA, Laeng P, Jurata LW, Brockman JA, Lemire A, Bullard J, Bukhman YV, Young TA, Charles V, Palfreyman MG. 2004. Electroconvulsive seizures regulate gene expression of distinct neurotrophic signaling pathways. J Neurosci 24(11):2667-77.

Altar CA, Vawter MP, Ginsberg SD. 2009. Target identification for CNS diseases by transcriptional profiling. Neuropsychopharmacology 34(1):18-54.

Asmann YW, Klee EW, Thompson EA, Perez EA, Middha S, Oberg AL, Therneau TM, Smith DI, Poland GA, Wieben ED and others. 2009. 3' tag digital gene expression profiling of human brain and universal reference RNA using Illumina Genome Analyzer. BMC Genomics 10:531.

Asmann YW, Wallace MB, Thompson EA. 2008. Transcriptome profiling using next-generation sequencing. Gastroenterology 135(5):1466-8.

Barco A, Patterson S, Alarcon JM, Gromova P, Mata-Roig M, Morozov A, Kandel ER. 2005. Gene expression profiling of facilitated L-LTP in VP16-CREB mice reveals that BDNF is critical for the maintenance of LTP and its synaptic capture. Neuron 48(1):123-37.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129(4):823-37.

Bayes A, Grant SG. 2009. Neuroproteomics: understanding the molecular organization and complexity of the brain. Nat Rev Neurosci 10(9):635-46.

Bhardwaj N, Lu H. 2005. Correlation between gene expression profiles and protein-protein interactions within and across genomes. Bioinformatics 21(11):2730-8.

Bonaventure P, Guo H, Tian B, Liu X, Bittner A, Roland B, Salunga R, Ma XJ, Kamme F, Meurers B and others. 2002. Nuclei and subnuclei gene expression profiling in mammalian brain. Brain Res 943(1):38-47.

Bono H, Yagi K, Kasukawa T, Nikaido I, Tominaga N, Miki R, Mizuno Y, Tomaru Y, Goto H, Nitanda H and others. 2003. Systematic expression profiling of the mouse transcriptome using RIKEN cDNA microarrays. Genome Res 13(6B):1318-23.

Brown VM, Ossadtchi A, Khan AH, Yee S, Lacan G, Melega WP, Cherry SR, Leahy RM, Smith DJ. 2002. Multiplex three-dimensional brain gene expression mapping in a mouse model of Parkinson's disease. Genome Res 12(6):868-84.

Cao Y, Dulac C. 2001. Profiling brain transcription: neurons learn a lesson from yeast. Curr Opin Neurobiol 11(5):615-20.

Cardozo AK, Berthou L, Kruhoffer M, Orntoft T, Nicolls MR, Eizirik DL. 2003. Gene microarray study corroborates proteomic findings in rodent islet cells. J Proteome Res 2(5):553-5.

Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C and others. 2005. The transcriptional landscape of the mammalian genome. Science 309(5740):1559-63.

Carter BS, Fletcher JS, Thompson RC. 2010. Analysis of messenger RNA expression by in situ hybridization using RNA probes synthesized via in vitro transcription. Methods.

Cavallaro S, D'Agata V, Manickam P, Dufour F, Alkon DL. 2002. Memory-specific temporal profiles of gene expression in the hippocampus. Proc Natl Acad Sci U S A 99(25):16279-84.

Chen EI, Yates JR, 3rd. 2007. Cancer proteomics by quantitative shotgun proteomics. Mol Oncol 1(2):144-59.

Chin MH, Geng AB, Khan AH, Qian WJ, Petyuk VA, Boline J, Levy S, Toga AW, Smith RD, Leahy RM and others. 2007. A genome-scale map of expression for a mouse brain section obtained using voxelation. Physiol Genomics 30(3):313-21.

Coba MP, Pocklington AJ, Collins MO, Kopanitsa MV, Uren RT, Swamy S, Croning MD, Choudhary JS, Grant SG. 2009. Neurotransmitters drive combinatorial multistate postsynaptic density networks. Sci Signal 2(68):ra19.

Coba MP, Valor LM, Kopanitsa MV, Afinowi NO, Grant SG. 2008. Kinase networks integrate profiles of N-methyl-D-aspartate receptor-mediated gene expression in hippocampus. J Biol Chem 283(49):34101-7.

Collingridge GL, Bliss TV. 1995. Memories of NMDA receptors and LTP. Trends Neurosci 18(2):54-6.

Collins MO, Husi H, Yu L, Brandon JM, Anderson CN, Blackstock WP, Choudhary JS, Grant SG. 2006. Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. J Neurochem 97 Suppl 1:16-23.

Collins MO, Yu L, Coba MP, Husi H, Campuzano I, Blackstock WP, Choudhary JS, Grant SG. 2005. Proteomic analysis of in vivo phosphorylated synaptic proteins. J Biol Chem 280(7):5972-82.

Dabrowski M, Aerts S, Van Hummelen P, Craessaerts K, De Moor B, Annaert W, Moreau Y, De Strooper B. 2003. Gene profiling of hippocampal neuronal culture. J Neurochem 85(5):1279-88.

Datson NA, Meijer L, Steenbergen PJ, Morsink MC, van der Laan S, Meijer OC, de Kloet ER. 2004. Expression profiling in laser-microdissected hippocampal subregions in rat brain reveals large subregion-specific differences in expression. Eur J Neurosci 20(10):2541-54.

de Reynies A, Geromin D, Cayuela JM, Petel F, Dessen P, Sigaux F, Rickman DS. 2006. Comparison of the latest commercial short and long oligonucleotide microarray technologies. BMC Genomics 7:51.

Deutsch EW, Lam H, Aebersold R. 2008. Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics. Physiol Genomics 33(1):18-25.

Dong HW, Swanson LW, Chen L, Fanselow MS, Toga AW. 2009. Genomic-anatomic evidence for distinct functional domains in hippocampal field CA1. Proc Natl Acad Sci U S A 106(28):11794-9.

Dougherty JD, Geschwind DH. 2005. Progress in realizing the promise of microarrays in systems neurobiology. Neuron 45(2):183-5.

Doyle JP, Dougherty JD, Heiman M, Schmidt EF, Stevens TR, Ma G, Bupp S, Shrestha P, Shah RD, Doughty ML and others. 2008. Application of a translational profiling approach for the comparative analysis of CNS cell types. Cell 135(4):749-62.

Emes RD, Pocklington AJ, Anderson CN, Bayes A, Collins MO, Vickers CA, Croning MD, Malik BR, Choudhary JS, Armstrong JD and others. 2008. Evolutionary

expansion and anatomical specialization of synapse proteome complexity. Nat Neurosci 11(7):799-806.

Etkin A, Alarcon JM, Weisberg SP, Touzani K, Huang YY, Nordheim A, Kandel ER. 2006. A role in learning for SRF: deletion in the adult forebrain disrupts LTD and the formation of an immediate memory of a novel context. Neuron 50(1):127-43.

Fejes-Toth K, Kapranov P, Sotirova, Assaf, Dumais, Gingeras. 2009. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. Nature 457(7232):1028-32.

Flavell SW, Cowan CW, Kim TK, Greer PL, Lin Y, Paradis S, Griffith EC, Hu LS, Chen C, Greenberg ME. 2006. Activity-dependent regulation of MEF2 transcription factors suppresses excitatory synapse number. Science 311(5763):1008-12.

Flavell SW, Greenberg ME. 2008. Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. Annu Rev Neurosci 31:563-90.

Flavell SW, Kim TK, Gray JM, Harmin DA, Hemberg M, Hong EJ, Markenscoff-Papadimitriou E, Bear DM, Greenberg ME. 2008. Genome-wide analysis of MEF2 transcriptional program reveals synaptic target genes and neuronal activity-dependent polyadenylation site selection. Neuron 60(6):1022-38.

French PJ, O'Connor V, Voss K, Stean T, Hunt SP, Bliss TV. 2001. Seizure-induced gene expression in area CA1 of the mouse hippocampus. Eur J Neurosci 14(12):2037-41.

Geschwind DH, Konopka G. 2009. Neuroscience in the era of functional genomics and systems biology. Nature 461(7266):908-15.

Ginsberg SD, Mirnics K. 2006. Functional genomic methodologies. Prog Brain Res 158:15-40.

Gong S, Zheng C, Doughty ML, Losos K, Didkovsky N, Schambra UB, Nowak NJ, Joyner A, Leblanc G, Hatten ME and others. 2003. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature 425(6961):917-25.

Gray PA, Fu H, Luo P, Zhao Q, Yu J, Ferrari A, Tenzen T, Yuk DI, Tsung EF, Cai Z and others. 2004. Mouse brain organization revealed through direct genome-scale TF expression analysis. Science 306(5705):2255-7.

Hadd AG, Brown JT, Andruss BF, Ye F, WalkerPeach CR. 2005. Adoption of array technologies into the clinical laboratory. Expert Rev Mol Diagn 5(3):409-20.

Havik B, Rokke H, Dagyte G, Stavrum AK, Bramham CR, Steen VM. 2007. Synaptic activity-induced global gene expression patterns in the dentate gyrus of adult behaving rats: induction of immunity-linked genes. Neuroscience 148(4):925-36.

Heiman M, Schaefer A, Gong S, Peterson JD, Day M, Ramsey KE, Suarez-Farinas M, Schwarz C, Stephan DA, Surmeier DJ and others. 2008. A translational profiling approach for the molecular characterization of CNS cell types. Cell 135(4):738-48.

Hong SJ, Li H, Becker KG, Dawson VL, Dawson TM. 2004. Identification and analysis of plasticity-induced late-response genes. Proc Natl Acad Sci U S A 101(7):2145-50.

Hunsberger JG, Bennett AH, Selvanayagam E, Duman RS, Newton SS. 2005. Gene profiling the response to kainic acid induced seizures. Brain Res Mol Brain Res 141(1):95-112.

Impey S, McCorkle SR, Cha-Molstad H, Dwyer JM, Yochum GS, Boss JM, McWeeney S, Dunn JJ, Mandel G, Goodman RH. 2004. Defining the CREB regulon: a genome-wide analysis of transcription factor regulatory regions. Cell 119(7):1041-54.

Ioannidis JP, Allison DB, Ball CA, Coulibaly I, Cui X, Culhane AC, Falchi M, Furlanello C, Game L, Jurman G and others. 2009. Repeatability of published microarray gene expression analyses. Nat Genet 41(2):149-55.

Jancic D, Lopez de Armentia M, Valor LM, Olivares R, Barco A. 2009. Inhibition of cAMP response element-binding protein reduces neuronal excitability and plasticity, and triggers neurodegeneration. Cereb Cortex 19(11):2535-47.

Jansen R, Greenbaum D, Gerstein M. 2002. Relating whole-genome expression data with protein-protein interactions. Genome Res 12(1):37-46.

Jones AR, Overly CC, Sunkin SM. 2009. The Allen Brain Atlas: 5 years and beyond. Nat Rev Neurosci 10(11):821-8.

Kamme F, Salunga R, Yu J, Tran DT, Zhu J, Luo L, Bittner A, Guo HQ, Miller N, Wan J and others. 2003. Single-cell microarray analysis in hippocampus CA1: demonstration and validation of cellular heterogeneity. J Neurosci 23(9):3607-15.

Kandel ER. 2001. The molecular biology of memory storage: a dialogue between genes and synapses. Science 294(5544):1030-8.

Karlic R, Chung HR, Lasserre J, Vlahovicek K, Vingron M. 2010. Histone modification levels are predictive for gene expression. Proc Natl Acad Sci U S A 107(7):2926-31.

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S and others. 2010a. Widespread transcription at neuronal activity-regulated enhancers. Nature.

Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S and others. 2010b. Widespread transcription at neuronal activity-regulated enhancers. Nature 465(7295):182-7.

Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, Scott MS, Gramolini AO, Morris Q, Hallett MT and others. 2006. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell 125(1):173-86.

Kline KG, Finney GL, Wu CC. 2009. Quantitative strategies to fuel the merger of discovery and hypothesis-driven shotgun proteomics. Brief Funct Genomic Proteomic 8(2):114-25.

Kothapalli R, Yoder SJ, Mane S, Loughran TP, Jr. 2002. Microarray results: how accurate are they? BMC Bioinformatics 3:22.

Kuo WP, Jenssen TK, Butte AJ, Ohno-Machado L, Kohane IS. 2002. Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 18(3):405-12.

Le-Niculescu H, McFarland MJ, Mamidipalli S, Ogden CA, Kuczenski R, Kurian SM, Salomon DR, Tsuang MT, Nurnberger JI, Jr., Niculescu AB. 2007. Convergent Functional Genomics of bipolar disorder: from animal model pharmacogenomics to human genetics and biomarkers. Neurosci Biobehav Rev 31(6):897-903.

Le-Niculescu H, McFarland MJ, Ogden CA, Balaraman Y, Patel S, Tan J, Rodd ZA, Paulus M, Geyer MA, Edenberg HJ and others. 2008. Phenomic, convergent functional genomic, and biomarker studies in a stress-reactive genetic animal

model of bipolar disorder and co-morbid alcoholism. Am J Med Genet B Neuropsychiatr Genet 147B(2):134-66.

Leil TA, Ossadtchi A, Cortes JS, Leahy RM, Smith DJ. 2002. Finding new candidate genes for learning and memory. J Neurosci Res 68(2):127-37.

Leil TA, Ossadtchi A, Nichols TE, Leahy RM, Smith DJ. 2003. Genes regulated by learning in the hippocampus. J Neurosci Res 71(6):763-8.

Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ and others. 2007. Genome-wide atlas of gene expression in the adult mouse brain. Nature 445(7124):168-76.

Lein ES, Zhao X, Gage FH. 2004. Defining a molecular atlas of the hippocampus using DNA microarrays and high-throughput in situ hybridization. J Neurosci 24(15):3879-89.

Lemberger T, Parkitna JR, Chai M, Schutz G, Engblom D. 2008. CREB has a context-dependent role in activity-regulated transcription and maintains neuronal cholesterol homeostasis. FASEB J 22(8):2872-9.

Levenson JM, Choi S, Lee SY, Cao YA, Ahn HJ, Worley KC, Pizzi M, Liou HC, Sweatt JD. 2004. A bioinformatics analysis of memory consolidation reveals involvement of the transcription factor c-rel. J Neurosci 24(16):3933-43.

Li H, Gu X, Dawson VL, Dawson TM. 2004. Identification of calcium- and nitric oxide-regulated genes by differential analysis of library expression (DAzLE). Proc Natl Acad Sci U S A 101(2):647-52.

Lin Y, Bloodgood BL, Hauser JL, Lapan AD, Koon AC, Kim TK, Hu LS, Malik AN, Greenberg ME. 2008. Activity-dependent regulation of inhibitory synapse development by Npas4. Nature 455(7217):1198-204.

Lobo MK, Karsten SL, Gray M, Geschwind DH, Yang XW. 2006. FACS-array profiling of striatal projection neuron subtypes in juvenile and adult mouse brains. Nat Neurosci 9(3):443-52.

Loebrich S, Nedivi E. 2009. The function of activity-regulated genes in the nervous system. Physiol Rev 89(4):1079-103.

Magdaleno S, Jensen P, Brumwell CL, Seal A, Lehman K, Asbury A, Cheung T, Cornelius T, Batten DM, Eden C and others. 2006. BGEM: an in situ hybridization database of gene expression in the embryonic and adult mouse nervous system. PLoS Biol 4(4):e86.

Mardis ER. 2008. The impact of next-generation sequencing technology on genetics. Trends Genet 24(3):133-41.

Matsuo R, Kato A, Sakaki Y, Inokuchi K. 1998. Cataloging altered gene expression during rat hippocampal long-term potentiation by means of differential display. Neurosci Lett 244(3):173-6.

Michaelevski I, Segal-Ruder Y, Rozenbaum M, Medzihradszky KF, Shalem O, Coppola G, Horn-Saban S, Ben-Yaakov K, Dagan SY, Rishal I and others. 2010. Signaling to transcription networks in the neuronal retrograde injury response. Sci Signal 3(130):ra53.

Mieczkowski J, Tyburczy ME, Dabrowski M, Pokarowski P. 2010. Probe set filtering increases correlation between Affymetrix GeneChip and qRT-PCR expression measurements. BMC Bioinformatics 11:104.

Mijalski T, Harder A, Halder T, Kersten M, Horsch M, Strom TM, Liebscher HV, Lottspeich F, de Angelis MH, Beckers J. 2005. Identification of coexpressed gene clusters in a comparative analysis of transcriptome and proteome in mouse tissues. Proc Natl Acad Sci U S A 102(24):8621-6.

Miller JA, Horvath S, Geschwind DH. 2010. Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. Proc Natl Acad Sci U S A 107(28):12698-703.

Miller JA, Oldham MC, Geschwind DH. 2008. A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. J Neurosci 28(6):1410-20.

Mody M, Cao Y, Cui Z, Tay KY, Shyong A, Shimizu E, Pham K, Schultz P, Welsh D, Tsien JZ. 2001. Genome-wide gene expression profiles of the developing mouse hippocampus. Proc Natl Acad Sci U S A 98(15):8862-7.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5(7):621-8.

Mufson EJ, Counts SE, Che S, Ginsberg SD. 2006. Neuronal gene expression profiling: uncovering the molecular biology of neurodegenerative disease. Prog Brain Res 158:197-222.

Nisenbaum LK. 2002. The ultimate chip shot: can microarray technology deliver for neuroscience? Genes Brain Behav 1(1):27-34.

Oldham MC, Horvath S, Geschwind DH. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. Proc Natl Acad Sci U S A 103(47):17973-8.

Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, Geschwind DH. 2008. Functional organization of the transcriptome in human brain. Nat Neurosci 11(11):1271-82.

Paradis S, Harrar DB, Lin Y, Koon AC, Hauser JL, Griffith EC, Zhu L, Brass LF, Chen C, Greenberg ME. 2007. An RNAi-based approach identifies molecules required for glutamatergic and GABAergic synapse development. Neuron 53(2):217-32.

Pegoraro S, Broccard FD, Ruaro ME, Bianchini D, Avossa D, Pastore G, Bisson G, Altafini C, Torre V. 2010. Sequential steps underlying neuronal plasticity induced by a transient exposure to gabazine. J Cell Physiol 222(3):713-28.

Peleg S, Sananbenesi F, Zovoilis A, Burkhardt S, Bahari-Javan S, Agis-Balboa RC, Cota P, Wittnam JL, Gogol-Doering A, Opitz L and others. 2010. Altered histone acetylation is associated with age-dependent memory impairment in mice. Science 328(5979):753-6.

Platenik J, Kuramoto N, Yoneda Y. 2000. Molecular mechanisms associated with long-term consolidation of the NMDA signals. Life Sci 67(4):335-64.

Ploski JE, Park KW, Ping J, Monsey MS, Schafe GE. 2010. Identification of plasticity-associated genes regulated by Pavlovian fear conditioning in the lateral amygdala. J Neurochem 112(3):636-50.

Pocklington AJ, Armstrong JD, Grant SG. 2006. Organization of brain complexity--synapse proteome form and function. Brief Funct Genomic Proteomic 5(1):66-73.

Ryan MM, Mason-Parker SE, Tate WP, Abraham WC, Williams JM. 2010. Rapidly induced gene networks following induction of long-term potentiation at perforant path synapses in vivo. Hippocampus.

Sandberg R, Yasuda R, Pankratz DG, Carter TA, Del Rio JA, Wodicka L, Mayford M, Lockhart DJ, Barlow C. 2000. Regional and strain-specific gene expression mapping in the adult mouse brain. Proc Natl Acad Sci U S A 97(20):11038-43.

Schadt EE. 2009. Molecular networks as sensors and drivers of common human diseases. Nature 461(7261):218-23.

Scheschonka A, Tang Z, Betz H. 2007. Sumoylation in neurons: nuclear and synaptic roles? Trends Neurosci 30(3):85-91.

Schones DE, Zhao K. 2008. Genome-wide approaches to studying chromatin modifications. Nat Rev Genet 9(3):179-91.

Sharma SK. 2010. Protein acetylation in synaptic plasticity and memory. Neurosci Biobehav Rev.

Sharp FR, Xu H, Lit L, Walker W, Apperson M, Gilbert DL, Glauser TA, Wong B, Hershey A, Liu DZ and others. 2006. The future of genomic profiling of neurological diseases using blood. Arch Neurol 63(11):1529-36.

Steward O, Schuman EM. 2001. Protein synthesis at synaptic sites on dendrites. Annu Rev Neurosci 24:299-325.

Strauss E. 2006. Arrays of hope. Cell 127(4):657-9.

Stritt C, Stern S, Harting K, Manke T, Sinske D, Schwarz H, Vingron M, Nordheim A, Knoll B. 2009. Paracrine control of oligodendrocyte differentiation by SRF-directed neuronal gene expression. Nat Neurosci 12(4):418-27.

Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G and others. 2004. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A 101(16):6062-7.

Sugino K, Hempel CM, Miller MN, Hattox AM, Shapiro P, Wu C, Huang ZJ, Nelson SB. 2006. Molecular taxonomy of major neuronal classes in the adult mouse forebrain. Nat Neurosci 9(1):99-107.

Tan PK, Downey TJ, Spitznagel EL, Jr., Xu P, Fu D, Dimitrov DS, Lempicki RA, Raaka BM, Cam MC. 2003. Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res 31(19):5676-84.

Tanis KQ, Duman RS, Newton SS. 2008. CREB binding and activity in brain: regional specificity and induction by electroconvulsive seizure. Biol Psychiatry 63(7):710-20.

Thompson CL, Pathak SD, Jeromin A, Ng LL, MacPherson CR, Mortrud MT, Cusick A, Riley ZL, Sunkin SM, Bernard A and others. 2008. Genomic anatomy of the hippocampus. Neuron 60(6):1010-21.

Valor LM, Charlesworth P, Humphreys L, Anderson CN, Grant SG. 2007. Network activity-independent coordinated gene expression program for synapse assembly. Proc Natl Acad Sci U S A 104(11):4658-63.

Valor LM, Grant SG. 2007. Integrating synapse proteomics with transcriptional regulation. Behav Genet 37(1):18-30.

Valor LM, Jancic D, Lujan R, Barco A. 2010. Ultrastructural and transcriptional profiling of neuropathological misregulation of CREB function. Cell Death Differ.

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F and others. 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature 457(7231):854-8.

Visel A, Thaller C, Eichele G. 2004. GenePaint.org: an atlas of gene expression patterns in the mouse embryo. Nucleic Acids Res 32(Database issue):D552-6.

Wang YY, Smith P, Murphy M, Cook M. 2009. Global Expression Profiling in Epileptogenesis: Does It Add to the Confusion? Brain Pathol.

Welboren WJ, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, Stunnenberg HG. 2009. ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. EMBO J 28(10):1418-28.

West AE, Griffith EC, Greenberg ME. 2002. Regulation of transcription factors by neuronal activity. Nat Rev Neurosci 3(12):921-31.

Wibrand K, Messaoudi E, Havik B, Steenslid V, Lovlie R, Steen VM, Bramham CR. 2006. Identification of genes co-upregulated with Arc during BDNF-induced long-term potentiation in adult rat dentate gyrus in vivo. Eur J Neurosci 23(6):1501-11.

Winden KD, Oldham MC, Mirnics K, Ebert PJ, Swan CH, Levitt P, Rubenstein JL, Horvath S, Geschwind DH. 2009. The organization of the transcriptional network in specific neuronal classes. Mol Syst Biol 5:291.

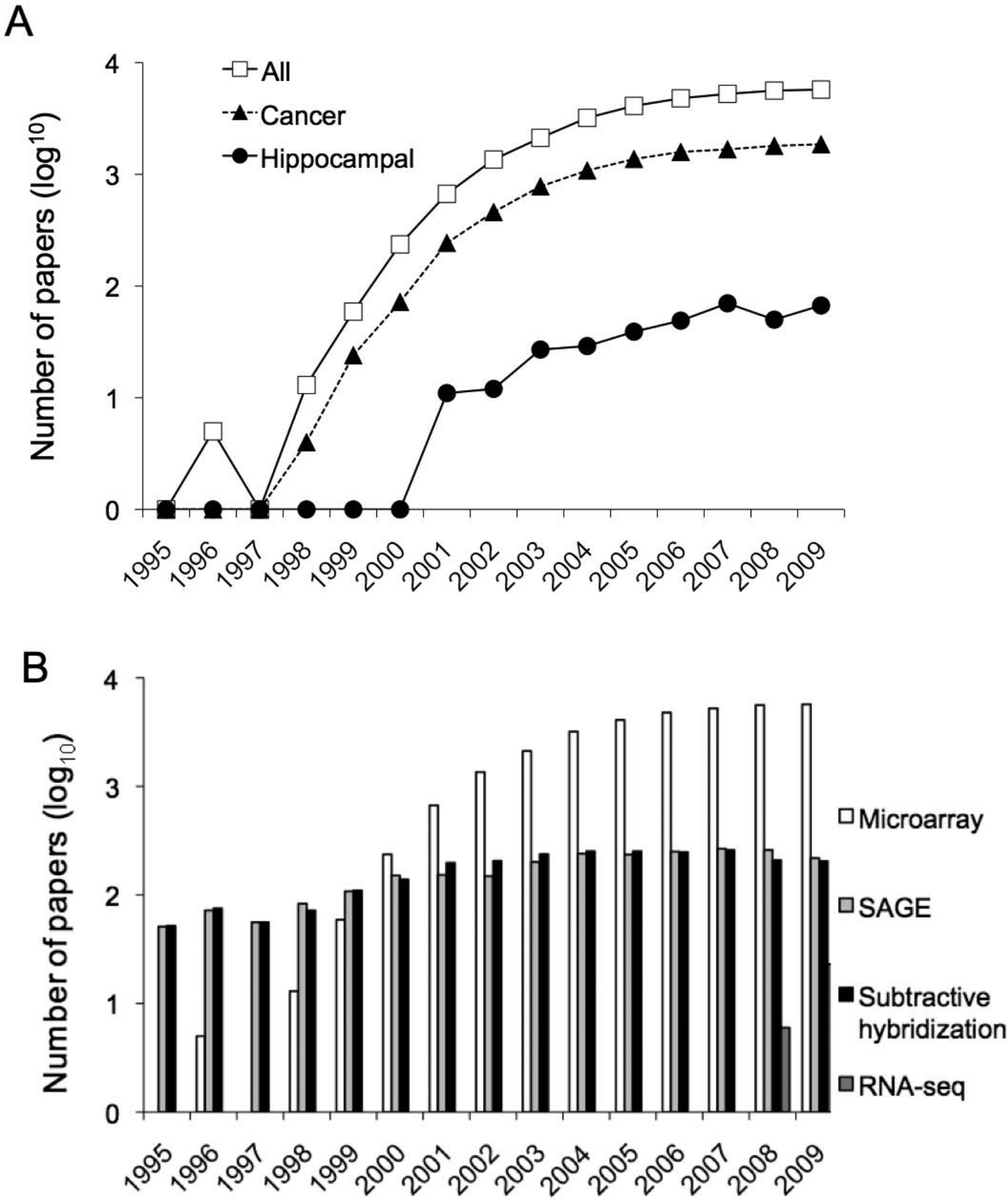
Zapala MA, Hovatta I, Ellison JA, Wodicka L, Del Rio JA, Tennant R, Tynan W, Broide RS, Helton R, Stoveken BS and others. 2005. Adult mouse brain gene expression patterns bear an embryologic imprint. Proc Natl Acad Sci U S A 102(29):10357-62.

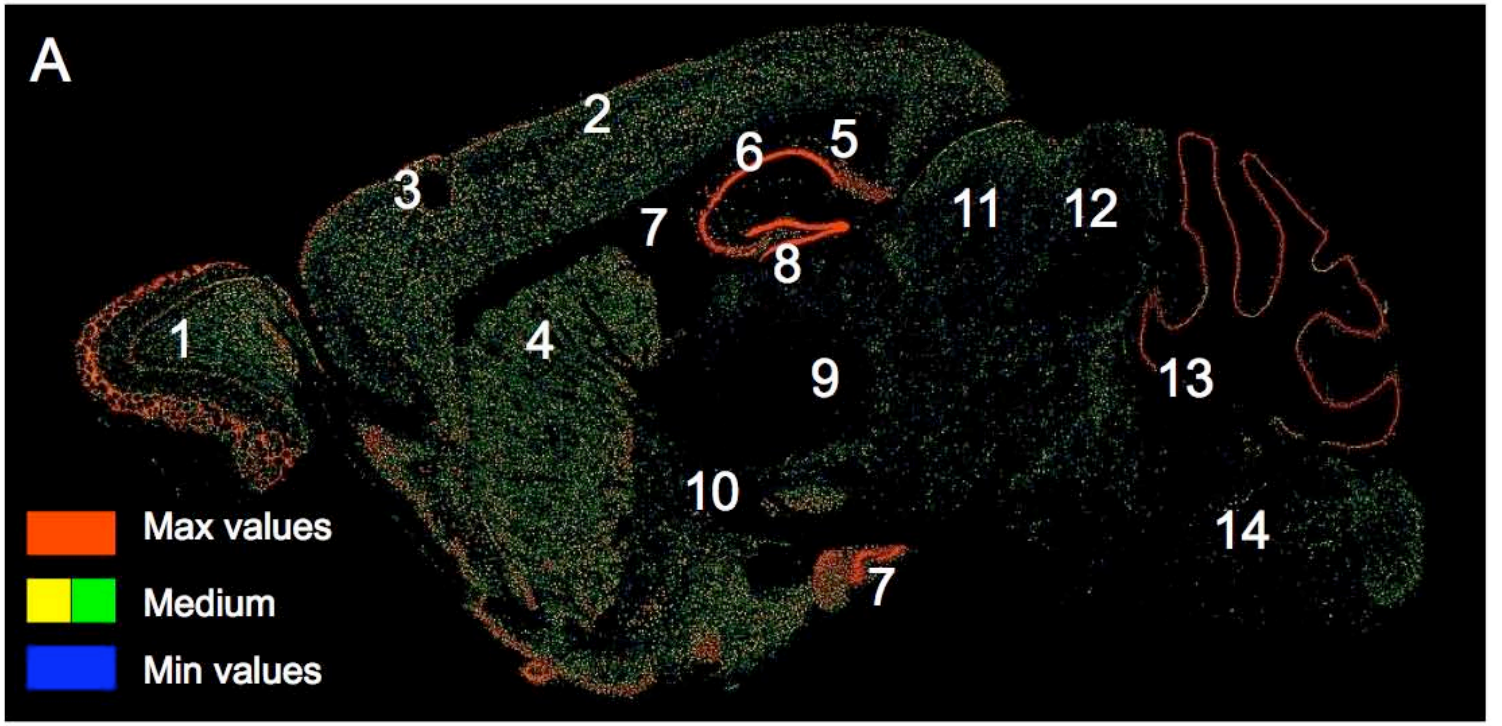
Zhang SJ, Steijaert MN, Lau D, Schutz G, Delucinge-Vivier C, Descombes P, Bading H. 2007. Decoding NMDA receptor signaling: identification of genomic programs specifying neuronal survival and death. Neuron 53(4):549-62.

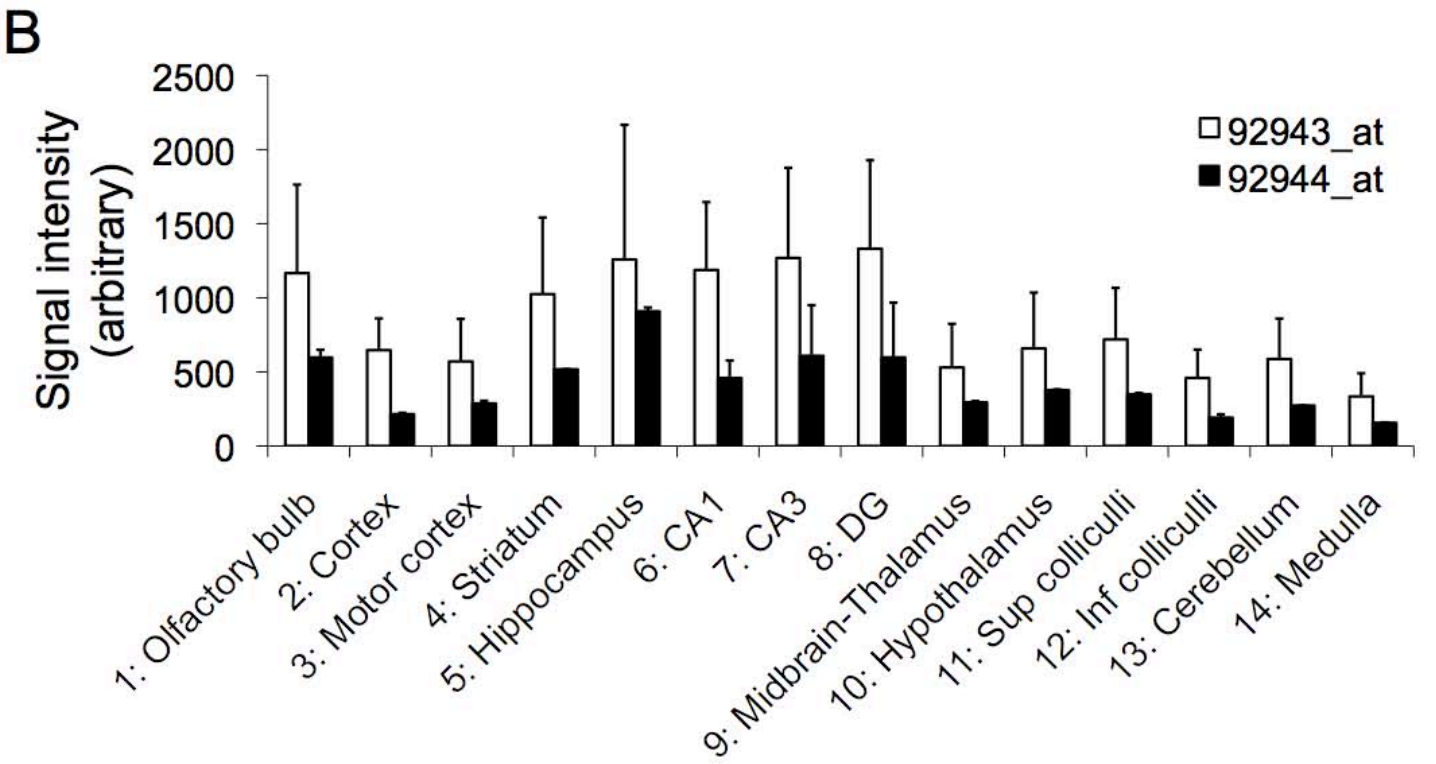
Zhang SJ, Zou M, Lu L, Lau D, Ditzel DA, Delucinge-Vivier C, Aso Y, Descombes P, Bading H. 2009. Nuclear calcium signaling controls expression of a large gene pool: identification of a gene program for acquired neuroprotection induced by synaptic activity. PLoS Genet 5(8):e1000604.

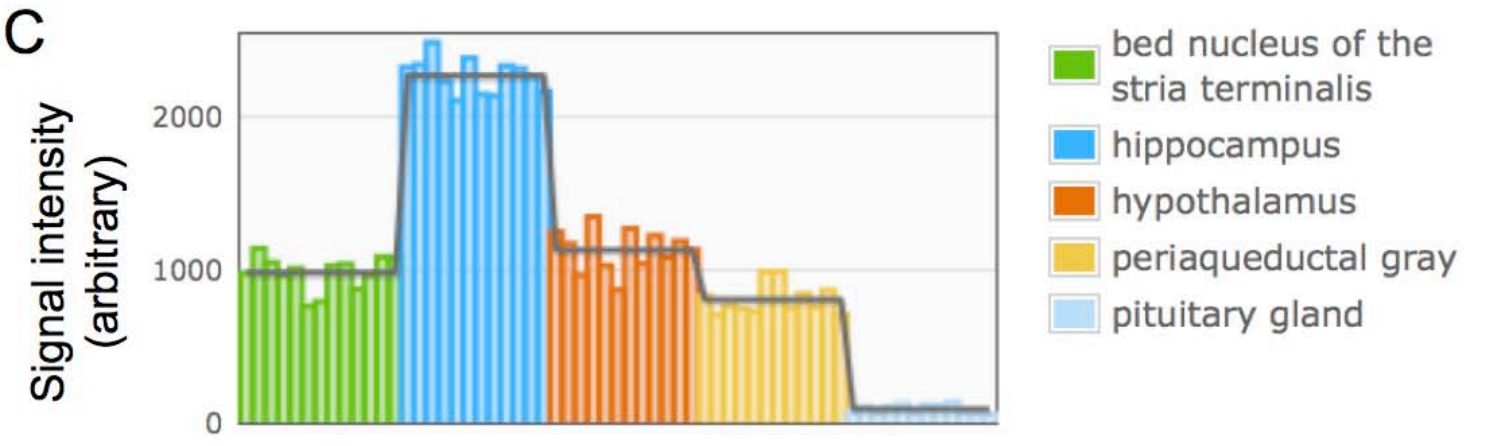
Zhang W, Morris QD, Chang R, Shai O, Bakowski MA, Mitsakakis N, Mohammad N, Robinson MD, Zirngibl R, Somogyi E and others. 2004. The functional landscape of mouse gene expression. J Biol 3(5):21.

Zhao X, Lein ES, He A, Smith SC, Aston C, Gage FH. 2001. Transcriptional profiling reveals strict boundaries between hippocampal subregions. J Comp Neurol 441(3):187-96.


A
All
Cancer
Hippocampal
Number of papers (log10)
0
1
2
3
4
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
B
Number of papers (log10)
Microarray
SAGE
Subtractive hybridization
RNA-seq


A
1
2
3
4
5
6
7
8
9
10
11
12
13
14
7
Max values
Medium
Min values


B
Signal intensity (arbitrary)
2500
2000
1500
1000
500
0
92943_at
92944_at
1: Olfactory bulb
2: Cortex
3: Motor cortex
4: Striatum
5: Hippocampus
6: CA1
7: CA3
8: DG
9: Midbrain-Thalamus
10: Hypothalamus
11: Sup colliculli
12: Inf colliculli
13: Cerebellum
14: Medulla


C
Signal intensity (arbitrary)
2000
1000
0
bed nucleus of the stria terminalis
hippocampus
hypothalamus
periaqueductal gray
pituitary gland

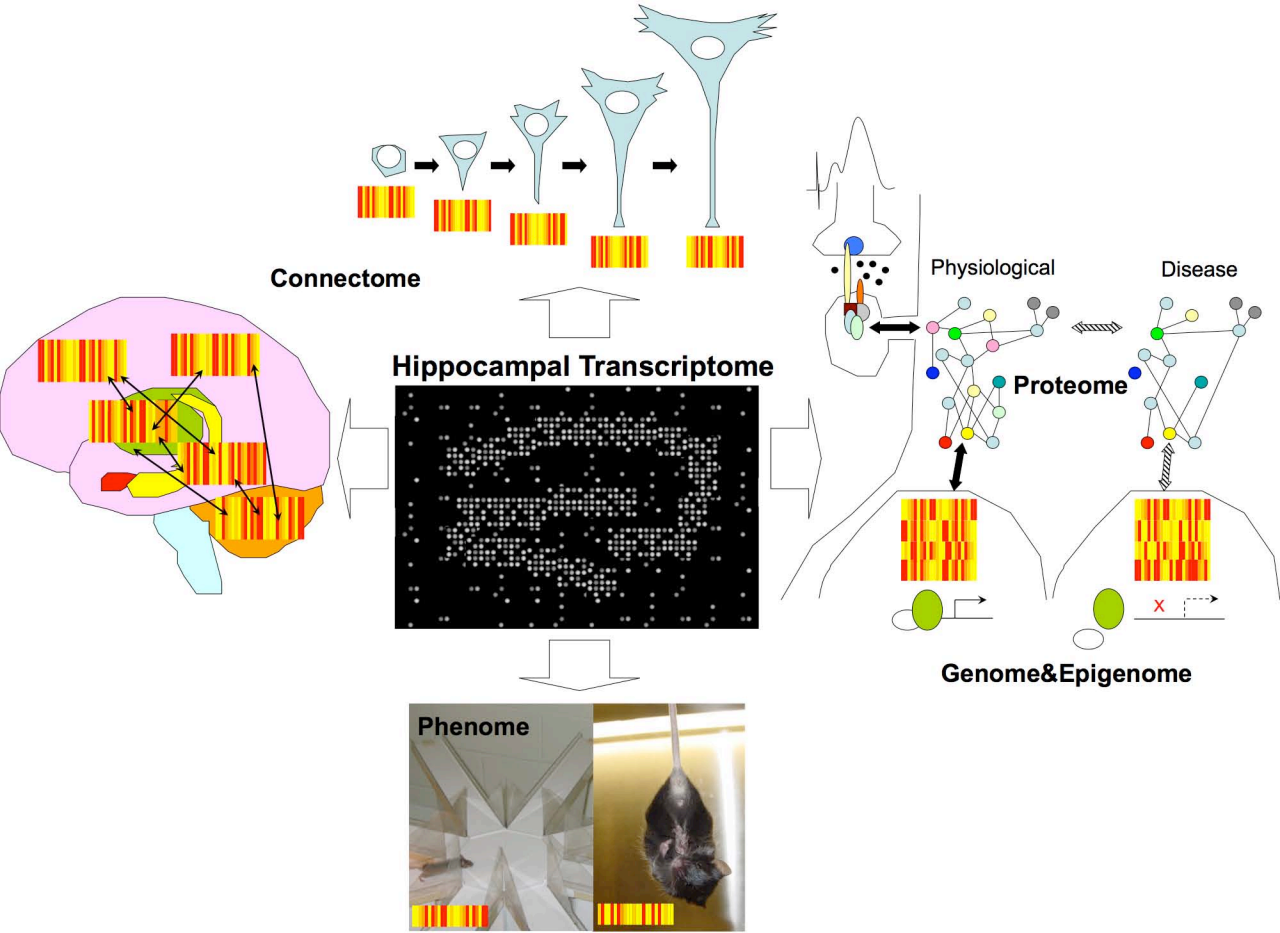

Connectome
Hippocampal Transcriptome
Physiological
Disease
Proteome
Genome&Epigenome
Phenome

Table 1. Open access data repositories of gene expression profiles

Name	Link	Brief description
GEO DataSets	http://www.ncbi.nlm.nih.gov/gds	Depository of microarray and high-throughput sequencing data.
GEO Profiles	http://www.ncbi.nlm.nih.gov/sites/entrez?db=geo	Search of gene profiles in deposited experiments
ArrayExpress	http://www.ebi.ac.uk/microarray-as/ae/	Depository of microarray and high-throughput sequencing data. Search of gene profiles in deposited experiments
BioGPS	http://biogps.gnf.org/	Search of gene profiles for tissue and cell types in microarray data
GENSAT	http://www.ncbi.nlm.nih.gov/sites/entrez?db=gensat	Search of gene profiles in ISH data
Allen Brain Atlas	http://www.brain-map.org/	Search of gene profiles in ISH data
GenePaint	http://www.genepaint.org/	Search of gene profiles in ISH data
BGEM	http://www.stjudebgem.org/web/mainPage/mainPage.php	Search of gene profiles in ISH data