Journal of Neurochemistry, 2003, **85**, 1279–1288 doi:10.1046/j.1471-4159.2003.01753.x

Gene profiling of hippocampal neuronal culture

Michal Dabrowski,*[,1] Stein Aerts,†[,1] Paul Van Hummelen,‡ Katleen Craessaerts,* Bart De Moor,† Wim Annaert,* Yves Moreau† and Bart De Strooper*

*Laboratory for Neuronal Cell Biology, Center for Human Genetics, Katholieke Universiteit Leuven and VIB (Flemish Interuniversity Institute for Biotechnology), Leuven, Belgium

†Department of Electrical Engineering ESAT-SCD, Katholieke Universiteit Leuven, Heverlee (Leuven), Belgium

‡VIB Microarray Facility, Leuven, Belgium

Abstract

We performed mRNA expression profiling of mouse primary hippocampal neurones undergoing differentiation *in vitro*. We show that 2314 genes significantly changed expression during neuronal differentiation. The temporal resolution of our experiment (six time points) permits us to distinguish between gene expression patterns characteristic for the axonal and for the dendritic stages of neurite outgrowth. Cluster analysis reveals that, in the process of *in vitro* neuronal differentiation, a high level of expression of genes involved in the synthesis of DNA and proteins precedes the up regulation of genes involved in protein transport, energy generation and synaptic functions. We report in detail changes in gene expression for genes involved in the synaptic vesicle cycle. Data for other genes can be accessed at our website. We directly compare expression of 475 genes in the differentiating neurones and the developing mouse hippocampus. We demonstrate that the program of gene expression is accelerated *in vitro* as compared to the situation *in vivo*. When this factor is accounted for, the gene expression profiles *in vitro* and *in vivo* become very similar (median gene-wise correlation 0.787). Apparently once the cells have taken a neuronal fate, the further program of gene expression is largely independent of histological or anatomical context. Our results also demonstrate that a comparison across the two experimental platforms (cDNA microarrays and oligonucleotide chips) and across different biological paradigms is feasible.

Keywords: cDNA microarray, gene expression, gene profiling, hippocampal neuronal culture, hippocampus, neuronal differentiation.

J. Neurochem. (2003) **85,** 1279–1288.

The progressive differentiation of neuronal precursor cells towards polarized, electrically active and synaptic transmission competent neurones is a fundamental aspect of brain development. The molecular analysis of this process is difficult because of the anatomical complexity of the developing brain and the multitude of different proteins and metabolic pathways involved in this process. High-density oligonucleotide or cDNA arrays (Schena *et al*. 1995; Lockhart *et al*. 1996) allow for the analysis of this complex process at the RNA expression level. The isolation of sufficient RNA from specific populations of cells at different stages of differentiation from the brain remains, however, technically challenging. One way to circumvent this problem is to isolate a certain subpopulation of neuronal precursor cells and to let them differentiate *in vitro*. A classical and well-studied example of such a cell culture system is the primary culture of hippocampal neurones as developed by Banker and collaborators (Goslin and Banker 1991). Hippocampal cells are isolated from late stage embryos and can be grown for weeks. One particular advantage of this culture system is that cells at different stages of development in the embryonic hippocampus become apparently resynchronized upon isolation (Fletcher and Banker 1989). They then differentiate again *in vitro* in a quite stereotypical way along five morphologically well-defined stages (Dotti *et al*. 1988). The neurones progressively develop neurites that differentiate into dendrites and one axon. In a later stage of the culture, active synapses are generated. The expression and subcellular distribution of proteins or RNA can be studied at any specific stage (Goslin and Banker 1991). A wealth of information on almost every aspect of neuronal differentiation, particularly

Received December 4, 2002; revised manuscript received February 17, 2003; accepted February 17, 2003.

Address correspondence and reprint requests to Bart De Strooper, Center for Human Genetics, VIB4, Herestraat 49, 3000 Leuven, Belgium. E-mail: Bart.destrooper@med.kuleuven.ac.be

[1]These authors contributed equally to this work.

Abbreviations used: GO, gene ontology; Syp, synaptophysin.

© 2003 International Society for Neurochemistry,

neurite outgrowth, neuronal polarization and synapse formation and function, has been obtained using this system (Fletcher *et al*. 1994; Boyer *et al*. 1998; Bradke and Dotti 2000a,b; Skutella and Nitsch 2001; Hopf *et al*. 2002). Axonal growth and development of polarity is blocked by inhibitors of RNA or protein synthesis indicating changes in gene expression during the neuronal differentiation process (Jareb and Banker 1997). Although a number of genes has been studied in this developing system (Bradke and Dotti 2000a), no attempt has been reported yet to document the global changes in gene expression during neuronal differentiation. We used microarrays containing 21 439 cDNA clones to analyse gene expression in primary hippocampal neurones differentiating in culture. In addition to the data described in this report, we make our full data set and annotation available at our web site http://www.esat.kuleuven.ac.be/neurdiff/ in a downloadable file format or via a purpose-built web application. The web application implements the functional gene ontology (GO) (Ashburner *et al*. 2000) annotation and allows investigators to select sets of genes of interest and to cluster and visualize expression profiles. We use throughout the paper the official gene symbols as they appear in the Mouse Genome Database (Blake *et al*. 2002).

Materials and methods

Preparation of neuronal culture

Hippocampal neurones were prepared from E17 CD1 strain mouse embryos, as described in Goslin and Banker (1991) and Annaert *et al*. (1999). Hippocampi from one litter were pooled and the dissociated cells were plated on poly-L-lysine-coated plastic dishes in minimal essential medium containing 10% horse serum, 600 mg/L of glucose and antibiotics. The number of cells corresponding to one hippocampus was plated onto one dish, resulting in a cell density of approx. 10 000 cells/cm^2; only for the culture lasting 12 days, twice as many cells were plated (approx. 20 000 cells/cm^2). After 6 h the medium was changed to Neurobasal medium containing 2% of B27 supplement (both from Life Technologies, Merelbeke, Belgium). Cytosine arabinoside (5 μM) was added after 24 h.

Immunofluorescence

Neurones were plated on poly-L-lysine-coated glass coverslips and cultured as described above. Fixation and indirect immunofluorescence were performed as described in Annaert *et al*. (1999) using monoclonal antibodies against Tau (1 : 200, Roche, Mannheim, Germany), Map2 (clone HM-2, 1 : 200, Sigma, Bornem, Belgium) and synaptophysin (clone C7.2, 1 : 100, Synaptic Systems, Göttingen, Germany). For detection, Alexa 488-conjugated secondary goat anti-mouse antibody (1 : 1000, Molecular Probes, Leiden, Netherlands) was used. Pictures were captured using Biorad MRC1024 confocal microscope (Bio-Rad Laboratories, Hercules, CA, USA).

RNA isolation and experiment design

Total RNA was isolated at 7 h, 18 h, 33 h, 72 h, 8 days and 12 days from the start of the culture. Total RNA from brains of newborn CD1 mice was used as a common reference for all the six time points. Minimum three independent cultures, each from one litter consisting of 10–12 embryos, were used for every time point. Typically, we pooled RNA from several neuronal cultures at the same time point of differentiation for a single hybridization.

Microarray hybridization

Construction of microarrays

The mouse gene set consisted of five microarray slides containing in total 21 492 cDNA fragments, each spotted at two distant positions. The clone set was composed from the 8K collection of Incyte (Mouse Gem I, Incyte, Palo Alto, USA) and from the 15K collection of National Institute of Ageing (NIA, HGMP Resource Centre, Hinxton, UK). The complete set can be found at *http://www.microarrays.be* (/currently available arrays).

Probe synthesis and hybridization

The probes were prepared according to Puskas *et al*. (2002). Briefly, 5 μg of total RNA was reverse transcribed, converted to double-stranded cDNA and amplified by *in vitro* transcription, resulting in the amplified RNA. The single-stranded fluorescently labelled cDNA probe was prepared from the amplified RNA by a reverse transcription, in the presence of Cy3-dCTP or Cy5-dCTP. The description of microarray procedures is given in the Supporting Methods that can be downloaded from our web site http://www.esat.kuleuven.ac.be/neurdiff/.

The probe synthesis and hybridizations were repeated twice, with inversion of the dyes for the experimental samples and the reference. Given the duplicate spots this resulted in four replicate ratio profiles for each clone.

Data management and analysis

Pre-processing

A pronounced dye-effect in the data (scatter plots showing the raw data are provided at our web site) was corrected by a slide-by-slide smoothing spline fitting of the background-corrected and log-transformed data (smoothing parameter $f = 0.2$) using Matlab's curve fitting toolbox (The Mathworks, MA, USA) (Yang *et al*. 2002). Other effects that result from systematic errors (e.g. array effects, pin effects, sample effects and gene effects) were negligible (data not shown).

Filtering

The log-transformed dye-normalized values were used to calculate the Pearson correlation coefficient for all six pairwise combinations of the four individual profiles that are available for each clone. Clones were filtered using a lower threshold of 0.6 for the minimum of the six correlation coefficients. An independent filtering was applied on the four profiles of each clone to retain only significantly changing genes by ANOVA analysis on each clone with one factor (time) and six treatments, using as H_0-hypothesis that the means at all time points are equal.

Annotation of our data set

The total number of distinct clones that are spotted on a set of five cDNA microarray slides is 21 439. These represent 13606 distinct

Unigene clusters (http://www.ncbi.nlm.nih.gov/UniGene/), or 8984 entries in the Mouse Genome Informatics Database (http://www.informatics.jax.org/), or 9502 Locuslink IDs (http://www.ncbi.nlm.nih.gov/LocusLink/). For 3233 MGI IDs we retrieved one or more GO terms, using the MGI annotation from October 2002 (http://www.geneontology.org). This annotation is used in our web application (http://www.esat.kuleuven.ac.be/neurdiff/) to retrieve the members of a functional class (e.g. a query for cytoskeleton, GO:0005856 retrieves all genes annotated with GO:0005856 and all genes annotated with one of the child terms of GO:0005856). All the annotations, however, are dynamic as we update them regularly.

A web application for functional gene expression analysis

A web application was developed for the data analysis and is available at http://www.esat.kuleuven.ac.be/neurdiff/. The following functions can be performed: (i) construct a gene group (a) by searching for genes using keywords or gene identifiers, (b) by Gene Ontology class, or (c) by cluster; (ii) apply a filtering on the genes within a group based on their internal correlation or p value from the ANOVA analysis; (iii) visualize all replicate measurements for one clone to identify false positives; (iv) visualize the average time profiles of a gene group in a heat map or in a profile chart using different representations of the measurements; and (v) save the gene group as a gene set for later investigation. The application architecture consists of three tiers: (i) a data layer using a MySQL database, (ii) a business layer consisting of several Java classes controlled by a Java servlet and of C executables for calculations (e.g. hierarchical clustering) and (iii) a visualization layer using Java Server Pages and HTML.

Clustering

Several clustering runs were performed using our implementation of the K-means algorithm with different values for the parameter K. The profiles that were used for clustering were the averages of the four replicate profiles of each clone that passed the correlation filter described above, in log scale and normalized by subtracting the mean and dividing by the standard deviation of the six measurements.

Comparing two microarray data sets

To compare our neuronal time series experiment (N) with the hippocampus time series experiment of Mody *et al*. (2001) (H) we first downloaded the H data set from http://braingenomics.princeton.edu. Then we selected the 475 Unigene clusters that overlap between the filtered data sets, some clusters represented by more than one profile per data set, thereby creating 686 distinct pairs of time profiles for the same genes in both data sets, all log transformed and normalized by subtracting the mean and dividing by the standard deviation (over all the time points).

N contains six time points and H contains five slightly different time points. We determined the optimal mapping of the time points using two approaches. In the first approach, we calculated the Pearson's correlation coefficients for the six possible order-preserving mappings of five time points between N and H, and identified the mapping for which the median correlation between the two data sets was highest. In the second approach, we used the published Aach and Church (2001) implementation called genewarp of an established time-warping algorithm (Kruskal and Liberman 1999), which minimizes the accumulated distance between the pairs of the mapped time points. The permuted data sets were constructed by random permutation of the time points of each gene profile separately.

Results

Neuronal differentiation *in vitro*

In preliminary experiments, we established that the neurones cultured in the absence of glial cells at cell densities of approx. 20 000 cells/cm^2 remained healthy and reached 'Stage 5' of differentiation, as defined by Dotti and Banker, after 10–12 days in culture (Dotti *et al*. 1988) (Fig. 1). Cells at 7 h in culture displayed initial outgrowth of neurites (Stage 2). The future axon was identifiable in most neurones at 18 h in the culture, and we chose 18 h as the first of the two time points representing 'Stage 3'. Only this neurite (future axon) stained for the Tau (Mapt) protein at 24 h (Fig. 1b). Around 33 h many axonal growth cones reached neighbouring cells and we used these cells for the second 'Stage 3' time point to document possible changes in gene expression following neurite–target interaction. Future dendrites increased their length from 72 h onwards. We took the 72 h and 8 days time points as early and late 'Stage 4'. It is important to note that, during 'Stage 4', axons continued to grow and branch. They are more difficult to identify morphologically, because of the greater number of outgrowing and branching dendrites. Transition to the mature morphology (Stage 5) occurred gradually and was already visible in some neurones at 8 days in the culture. At 9 days the dendrites and cell bodies showed punctate synaptophysin (Syp) staining, representing newly formed synaptic contacts (Fig. 1b). After 12 days in culture, practically all neurones displayed mature morphology: greatly enlarged soma, tapering and extensively branched

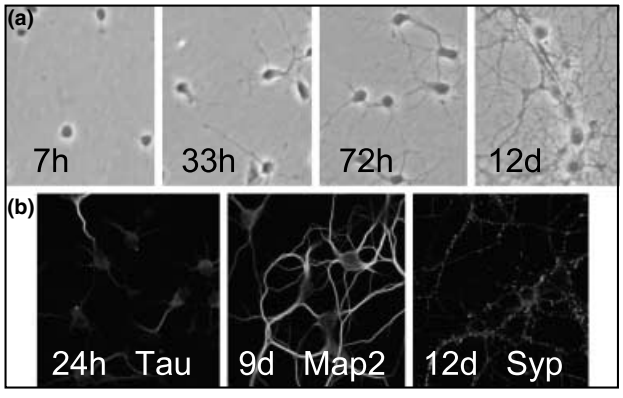


Fig. 1 Timing of the stages of neuronal differentiation. (a) Contrast phase pictures of neurones at the indicated times in the culture. (b) Confocal pictures of neurones fixed at the indicated times in the culture and stained for the proteins Tau (Mapt), MAP2 (Map2) and Synaptophysin (Syp).

dendrites with varicosities, and axons forming a dense mesh on the culture dish (Fig. 1, 12 days). The amount of contaminating glia cells in our cultures was very low, in accordance with previously published work (Dotti *et al.* 1988).

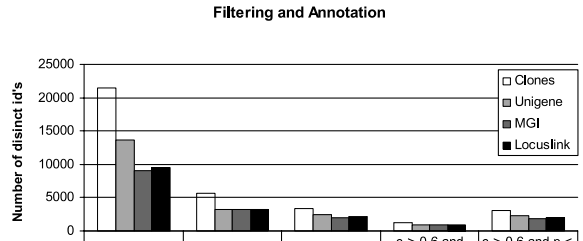
Quality of the data and its annotation

We selected the reliable (reproducible for the four replicate profiles) average profiles by a correlation filter described in the Methods. This filtering retained 3341 clones (15.5% of all) representing 2510 distinct Unigene clusters in the differentiating neurones (Fig. 2). Genes with significantly changed expression were identified by an independent (i.e. with no correlation filter applied first) ANOVA filter also described in the Methods. Almost all (92%) of the 3341 reliable profiles showed highly significant ($p < 0.01$) changes in expression during the time course.

For many practical purposes it is important that a cDNA clone is assigned to a known entry in the existing databases, describing transcript (Unigene), gene (MGI), genomic location (LocusLink), or function (GO). We reannotated our clone set (as described in the Methods) and made the annotations easy to use via our web application. Figure 2 summarizes also the annotation.

Global changes of gene expression in neuronal differentiation

We grouped the genes according to their expression patterns by *K*-means clustering (Tou and Gonzalez 1979). Optimal results (tight clusters with biologically interpretable profiles) were obtained for $K = 20$ (Fig. 3). Several of the clusters



	Total	GO	c > 0.6	c > 0.6 and GO	c > 0.6 and p < 0.01
Clones	21439	5617	3341	1241	3083
Unigene	13606	3233	2510	902	2314
MGI	8984	3233	2008	902	1853
Locuslink	9502	3231	2118	902	1955

Fig. 2 Filtering and annotation of our data set. The columns represent: the whole clone set (Total), its subset annotated with GO, the clones that pass the correlation filter (reliable profile) ($c > 0.6$), the intersection of the previous two subsets ($c > 0.6$ and GO), and the clones with reliable profile and significant change by ANOVA ($c > 0.6$ and $p < 0.01$). The rows give numbers of distinct identifiers from the specified transcript (Unigene), gene (MGI, LocusLink) and genomic (Ensemble) databases. The description of the reannotation process and the URLs of the databases are given in the Methods.

are enriched for genes involved in similar functions (as compared to the other clusters). Based on these functions, four larger groups of clusters were identified, indicated in Fig. 3. Group A contains Clusters 1 and 15, both strongly down regulated during neuronal differentiation (yellow box in Fig. 3). This group contains many genes involved in the cell cycle: DNA replication (Top2b, Prim1, Prim2, Lig1, Rev3l), chromatin assembly (H1f0, H3f3a, H3f3b, Chaf1a, Nasp), cell-cycle regulators (Ccnb1-rs, Ccna2, E2fb, Bub1b). Group A also contained several components of the actin cytoskeleton (Actb, Myo1b, Vil2), indicating that these genes also become down regulated as a part of the reorganization of the cytoskeleton during the neuronal differentiation. Group B, encompassing Clusters 3, 6, 7, 10 and 11 (green in Fig. 3), contains moderately down-regulated genes. This group is highly enriched in genes involved in RNA metabolism: ribosomal proteins (at least 29, a full lists of genes in each cluster can be found at our web site), RNA binding proteins (Pcbp2, Refbp1, Nsap1), nuclear ribonucleoproteins (Snrpa1, Snrpd1, Snrbp, Hnrpa1, Hnrpab, Hnrpk, Hmrph1, Hnrpa2b1), RNA splicing factors (Sfrs2, Sfrs3, Sf3b1) and regulators of translation (Eif1b2, Eif4a1, Eif4g2, Eef1a1, Naca). Overall these changes indicate a progressive switch from biosynthetic activity towards more functional activity in the developing neurones (also reflected in the up-regulated genes, see below). Group B also contains several genes of the TGF-β signalling pathway (Bmp1, Tgfb2, Tgfbr1, Madh4, Sin3b). Notch1 can be found in Cluster 6 and the kinase Rock1 in Cluster 10, both proteins involved in axonal and dendritic outgrowth (Berezovska *et al.* 1999; Redmond *et al.* 2000; He *et al.* 2002). Group C, composed of Clusters 2, 4, 5 and 8 (blue in Fig. 3), contains all genes that were moderately up regulated. This group is enriched for genes playing a role in the secretory and endocytic pathways and it includes genes for ER localized proteins (Hspa5, Erp29, P4ha2, Noe1-pending, Aldh3a, Sec61a2-pending), Golgi proteins (Grs2-pending, Vcp, Siat6, Cope, Ap1b1), lysosomal proteins (Ptp, Tpp2, Smpd1, Lyst), a series of intracellular protein trafficking regulators (Rab12, Cop7a, Cop9, Arl6ip, Prnp and Ly6e) and vacuolar sorting proteins (Vps29, Vps41, Vps45). A second function clearly overrepresented in group C is energy metabolism, as exemplified by the presence of ATP synthases, cytochrome subunits and soluble mitochondrial enzymes. Also several synaptic function-related genes are represented in group C, including Syt11, Synj2bp and Gad1. Group D, comprising the strongly up-regulated Clusters 18 and 20 (purple in Fig. 3), contains many genes encoding proteins involved in neurone-specific functions and energy metabolism. The latter included ATP synthase genes (Atp61, Atp6a2, Atp6b2, Atp6s1) and glycolysis enzymes (Aldo1, Gpi1, Ldh1, Pgk1, Ckmt1). Neurone-specific functions were represented by genes of the synaptic vesicle cycle (Syngr3, Vamp2, Nsf, Syn2, Arf2), a GABA

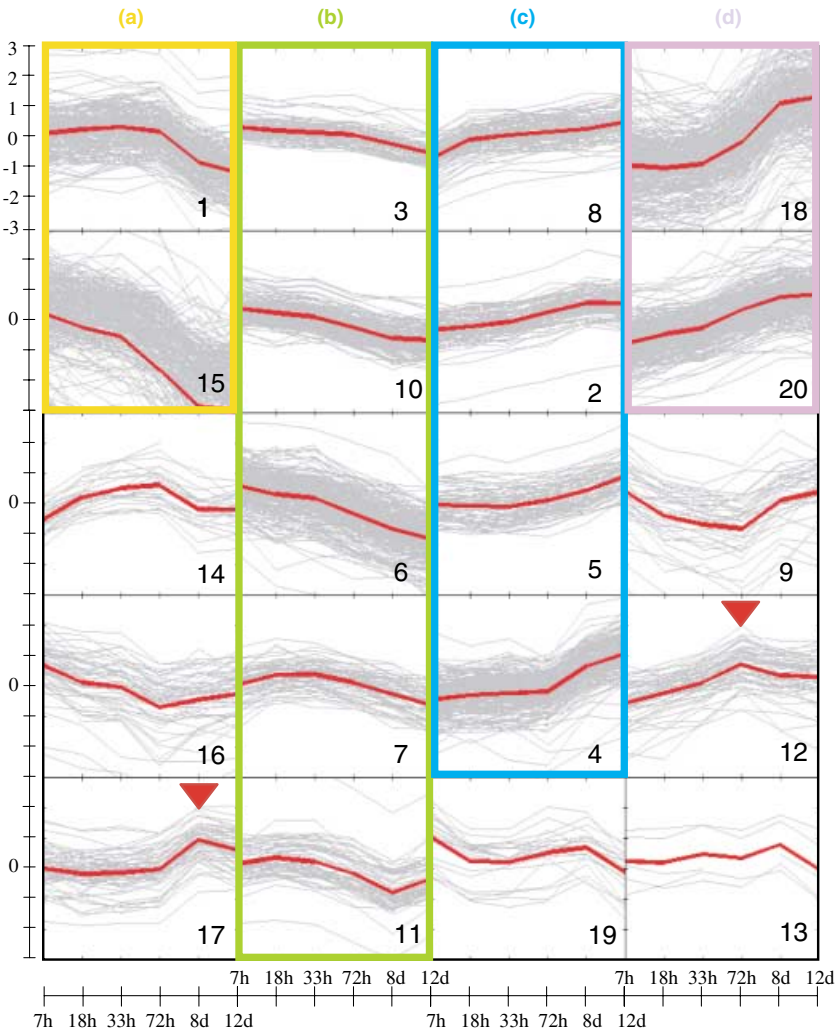


Fig. 3 Results of a *K*-means clustering for *K* (the number of clusters) equal to 20. Each of the 20 squares represents one cluster. On the *x* axis are the six time points (7 h, 18 h, 33 h, 72 h, 8 days, 12 days) and on the *y* axis the log-transformed filtered and normalized (see the Materials and methods) expression ratios. Each grey line represents a profile for a single clone line and the average profile of a cluster is shown as a bold red line. Each cluster or a part of it can be visualized separately as a profile chart or as a heatmap using our web application. The red triangles mark the maxima at 72 h and 8 days.The clusters in which we found an overrepresentation of genes with stated functions were grouped together and boxed in colour. The clusters are not numbered consecutively to maintain the correspondence between the cluster numbers in this figure and in the original clustering results on our web site. The functions overrepresented in the boxed groups of clusters: (a) DNA replication, chromatin assembly; (b) ribosomal proteins, RNA binding, translational regulation; (c) Golgi/ER/lysosome, protein traffic, energy, vesicular transport; (d) energy, synaptic, App-related.

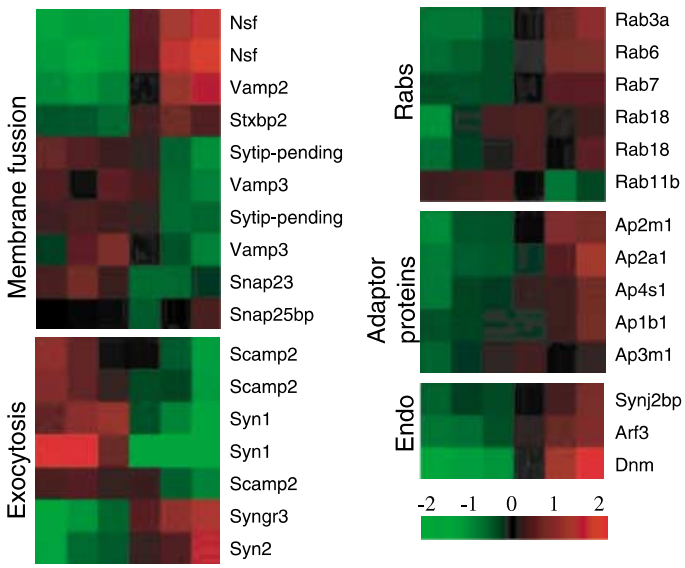


Fig. 4 Expression of the synaptic vesicle cycle genes. The genes were chosen as described in the main text. Within each subgroup the similar profiles were grouped together by hierarchical clustering (implemented at our web site). The colours indicate level of expression at a given time-point, relative to the average expression of this gene over all time-points, in log2 scale, with red for expression higher and green for lower than the gene average.

receptor Gabrg2, ion channels (Kcnd2, Clcn3) and cytoskeletal genes (Tuba4, Nfl, Kifc2, Kifc3, Kifap3). Interestingly in the light of ongoing research on the function of the amyloid beta (A4) precursor protein (App) and its role in Alzheimer's disease, the majority of Alzheimer's disease related genes also belonged to this group. This included App, Aplp2, Icam5, Adam9, Psen2 and Snca.

In the remaining clusters we were not able to identify a clearly overrepresented function, possibly because they contain a large proportion of ESTs representing unknown genes. Nevertheless, these clusters are of interest. The genes in Cluster 12 for example have a maximum expression at 72 h in culture, correlating with the beginning of dendritic outgrowth. Members of this cluster include Rac3, Kif3 and Apc. Cluster 17, with maximum of expression at 8 days in culture (which correlates with the early stage of full polarization of the neurones), contains genes with diverse functions.

Expression of synaptic vesicle cycle genes during neuronal differentiation

The previous approach classified genes according to their expression profiles. The data can also be analysed from a different perspective, by investigating shifts in expression in defined functional classes of genes. We illustrate this second approach for a group of genes important for the synaptic vesicle cycle. The web interface to our data set permits easy access to expression data for other functional groups.

The regulated release of neurotransmitters is a key function of the mature neurone. Many proteins have been implicated in the process of targeting and docking of the synaptic vesicle at the pre-synaptic membrane, the priming of the vesicle and its fusion (Lin and Scheller 2000; Lloyd *et al*. 2000). Many of these proteins are also involved in membrane trafficking outside the synapse (Chen and Scheller 2001; Parlati *et al*. 2002). Conversely, for many members of the same protein families the subcellular localization and function is not yet known. We decided to select all the genes in a gene family (e.g. Vamp) if at least one member of this family (e.g. Vamp2) is known to play a role at a particular stage of the synaptic vesicle cycle, essentially following Lloyd *et al*. (2000). During differentiation a gradual shift in expression of members of the different families was observed (Fig. 4). The most strongly regulated gene in the membrane fusion group was Nsf, involved at the disassembly step (Tolar and Pallanck 1998). The early expression pattern of Vamp3, Snap23 and Syt4 coincided temporarily with the outgrowth of the early axon (Stage 3). These findings are in good agreement with published data showing that Vamp3 is not found in synaptic vesicles (Chilcote *et al*. 1995) and that Snap23 is almost undetectable in adult brain (Wong *et al*. 1997). Vamp2 in contrast had a late expression pattern, with a maximum at 12 days (Stage 5) in the culture. Vamp2 is involved in the synaptic vesicle fusion step as part of the SNARE complex and is essential for secretion (Ahnert-Hilger *et al*. 1996; Mochida 2000). It has been postulated that the general exocytosis machinery (Jahn and Sudhof 1999) is used in the outgrowth of both axons and dendrites (Martinez-Arca *et al*. 2001; Tang 2001). A high level of expression of the ubiquitously expressed Scamp2 coincided with axonal outgrowth. Syn1 was expressed at a high level until 72 h in culture (early Stage 4) and then became strongly down regulated. The Syn1 protein (together with Syp) is known to preferentially localize to the distal axon and the growth cones at the stage of early axonal outgrowth (before cell–cell contact) (Fletcher *et al*. 1991), which is consistent with its early expression pattern in our experiment. Of the two transcripts of Syn1 described in rat cerebellum, the longer one is expressed only until P7 (Haas and DeGennaro 1988), similarly to the profile we observed. Expression of Syn2 and Syngr3 (Sugiyama *et al*. 2000) reached their maximum at 12 days in the culture (Stage 5), suggesting a role in mature synapses. Exocytotic events in the nerve terminals are compensated by endocytosis (Di Fiore and De Camilli 2001; Jarousse and Kelly 2001). In keeping with this, the genes in the endocytosis group were also up regulated in the late Stages 4, 5 of the culture, with a maximum at 12 days in culture when synapses have been generated (Fig. 1). In developing rat brain the expression of Dnm starts to increase from P7 and reaches adult levels at P23 (Faire *et al*. 1992). Also in the chick embryo's retino-tectal system dynamin is up regulated only after synapse formation (Bergmann *et al*. 1999). In rat brain the mRNA expression of Arf3 increases postnatally, from P2 to P27 (Tsai *et al*. 1991). The clustering analysis discussed above revealed coexpression of many synaptic and mitochondrial genes. Therefore it is interesting to note that Synj2bp recruits synaptojanin to mitochondria, which may affect their intracellular distribution (Nemoto and De Camilli 1999). The early expressed *Rab*11b has a role in the apical membrane recycling systems in polarized epithelial cells and in growth cone mobility (Casanova *et al*. 1999; Lapierre *et al*. 2001). The late-expressed Rab3a is up regulated in development only after synapse formation. Expression of this gene in the developing barrel field occurs later than that of Sv2a, Syn1 and Syp, and coincides with the onset of adult-like physiological activity (Stettler *et al*. 1996).The majority of the genes encoding adaptor proteins were expressed late in our culture system, including Ap2a1. Similarly to what is seen in *Drosophila* development (Dornan *et al*. 1997), the expression patterns of Ap2a1 and Dnm were highly similar. Adaptor proteins and Rab proteins are important at several stages of intracellular vesicular transport and their predominant late-expression pattern is likely to reflect increased transport needs in mature neurones. The observed up regulation of Dyn, Rab3a, Arf3 and Ap2a1 at 12 days in the culture is in agreement with the fact that by Stage 5 the neurones had functioning synapses with active secretion and compensatory endocytosis.

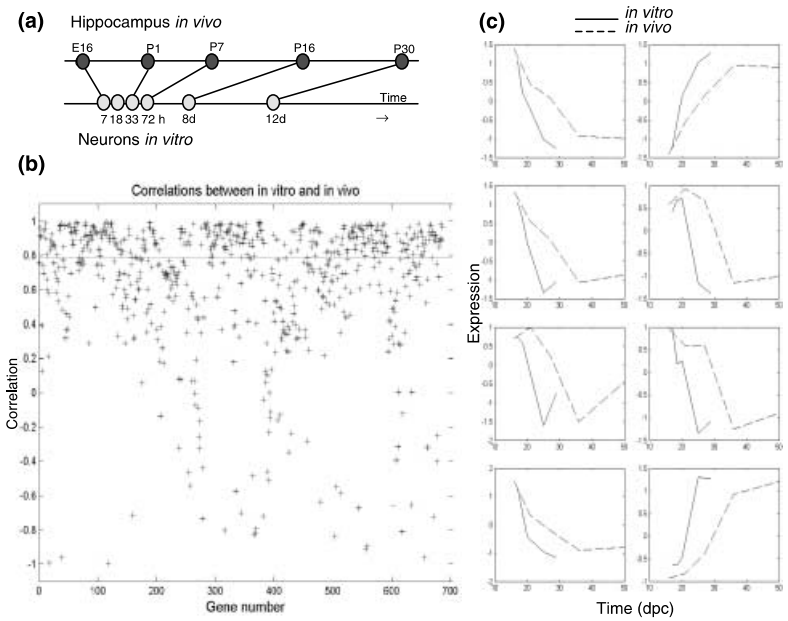


Fig. 5 Comparison of gene expression in the developing hippocampus and in the neurones differentiating *in vitro*. (a) Experiments and time points used for the comparison of gene expression in the developing mouse hippocampus (data set from Mody *et al.* 2001) and in the mouse hippocampal neurones differentiating in culture (our data). The mapping between time points that is shown is the optimal mapping as it results from a time-warping analysis. (b) Scatter plot of the Pearson correlation coefficients, between the log-transformed normalized profiles composed of the five time points indicated in (a), for each distinct pair of profiles representing the same transcript (Unigene ID) in both data sets. (c) Representative profile charts for eight genes with a high correlation coefficient showing the profiles composed of the five time points indicated in (a). The *in vitro* and *in vivo* profiles are plotted on the same time scale. The *x* axis represents time in days from the conception (dpc) and the *y* axis the expression values.

Comparison with the developing mouse hippocampus

Mody *et al*. (2001) reported a data set consisting of 1926 significantly changed expression profiles in developing hippocampus *in vivo*.

The 3341 expression profiles in the neurones differentiating *in vitro*, resulting from our experiment, provided us with an exciting opportunity to compare the expression of genes in both systems. There were 475 distinct genes (defined here as distinct Unigene IDs) represented in both data sets. Because some of the genes were represented more than once, the number of possible comparisons (686) was higher than the number of common genes. Because the number of measurements and their timing in the two experiments were not identical, we had to choose which time points to use to make the comparison. We performed a data-driven analysis (described in Methods) to find the optimal mapping of five consecutive time points between the two experiments. Both methods we used resulted in the same optimal mapping maximizing the similarity between expression profiles in the two systems. This mapping is illustrated in Fig. 5(a) and the Pearson correlation coefficients between expression profiles *in vivo* and *in vitro* calculated corresponding to this mapping are plotted in Fig. 5(b). The mean and median correlation coefficient are 0.646 and 0.787, respectively. The median correlation coefficient for the 'optimal' mapping between the same data sets with permuted time profiles was 0.0394, which illustrates the statistical significance of the correlation analysis. The high correlation obtained for the shown mapping, between an experiment *in vitro* lasting 12 days and an experiment *in vivo* spanning 34 days, suggests that the program of gene expression was accelerated *in vitro* when compared to the situation *in vivo*. This is clearly the case for the highly correlated genes, which is illustrated in Fig. 5(c) where profiles are plotted on the same time scale. The link P30–12 days contributes highly to the overall high similarity between the two expression profiles. When we calculated, for instance, the correlation coefficients for all the possible mappings of four consecutive time points (thus removing in some mappings the P30) we found that the six highest ranking mappings included the link P30–12 days (data not shown), indicating that high similarity extends to the latest points in the two experiments.

Discussion

We provide here the first genome-wide analysis of changes in gene expression accompanying the differentiation of

hippocampal neurones *in vitro*. This culture system has been used extensively in the past and is considered as an excellent model to study neuronal cell biology (Fletcher *et al*. 1994; Boyer *et al*. 1998; Bradke and Dotti 2000a,b; Skutella and Nitsch 2001; Hopf *et al*. 2002). We extend this claim to the transcriptome level (Fig. 5). We have demonstrated that neuronal differentiation is characterized by changes in the expression of genes from many different functional families. At least 2314 genes show a change in expression with a statistical certainty ($p < 0.01$). This indicates that the rebuilding of the rounded, unpolarized cell observed at Stages 1 and 2 (Fig. 1) into the highly complicated polarized and electrically active neurone in Stage 5 requires an orchestrated change in expression of thousands of genes. This change is remarkably smooth as the dominant pattern in the gene expression profiles is a gradual up or down regulation of gene expression over several stages of differentiation. This pattern was seen at the global level of analysis, but also within most functional groups (e.g. synaptic vesicle cycle in Fig. 4), resulting in replacement of early genes by the late ones with (seemingly) similar function.

The change in expression was significant at the *p*-level 0.01 for 2314 genes out of 2510 genes that pass the correlation (reproducibility) filter. Apparently, filtering for reproducible profiles results in a bias towards the genes with a change in expression. We hypothesize that the genes that change expression during differentiation are regulated in a more robust way, resulting in a higher biological reproducibility, and/or are expressed at higher absolute levels, which leads to more reproducible measurements.

One word of caution is indicated. When interpreting our data set, it is important to take into account that the less-than-perfect developmental-phase coherence among neurones could partially smoothen the slope of gene profile curves. On the other hand, the neurones differentiate in a remarkably synchronized way (see Fig. 1) and data were sampled in duplicate at six different time points in two independent experiments, providing a quite high level of resolution and reliability.

The global picture of gene expression patterns during neuronal differentiation as it emerges from the clustering analysis makes remarkable teleological sense. In a first phase of the culture (Stages 2, 3, early Stage 4), a high level of expression of genes characteristic for DNA and protein synthesis is observed, which then becomes progressively down regulated (Fig. 3, group A, Clusters 1 and 15). The later Stages 4 and 5 of differentiation are characterized by a strong enhancement of protein transport (Fig. 3, group C: Clusters 2 and 5) and energy generating systems (Fig. 3, group D: Clusters 18 and 20) and the turning on of specific neuronal functions, such as synaptic vesicle cycling (Fig. 3, group D: Clusters 18 and 20, see also Fig. 4). The high morphological resolution of the *in vitro* system permitted us to identify gene expression patterns characteristic for the axonal (Stage 3) and dendritic phase (Stage 4) of neuronal differentiation. Therefore we were able to resolve the gene expression patterns described in Mody *et al*. (2001) as 'differentiation and synapse formation', into two very different patterns: early – characteristic for axonal outgrowth, and late – characteristic for dendritic outgrowth and maturation. The difference can be appreciated, for example, by choosing 'cytoskeleton' as the GO group at our web site. Classification of the approximately 1000 genes in our data set for which no function is known into these two classes may be a useful first step towards the further elucidation of their function in neurones.

The high similarity of the expression profiles for the 475 common genes, in our data set and in the data set published by the group of Joe Z. Tsien (Mody *et al*. 2001), has several important consequences. First, each data set can be considered as an independent confirmation of the other one. Second, we demonstrate that the results obtained with two different experimental platforms *can* yield a good agreement. Assuming that the measurements from both platforms are reliable the effect of which platforms is used should be small and this is essentially what we have found. In a recent comparison (Kuo *et al*. 2002) of previously published data sets obtained with cDNA microarrays (Ross *et al*. 2000) and with Affymetrix oligonucleotide chips (Butte *et al*. 2000) the mean correlation between the two platforms was 0.278. The authors' conclusion (Kuo *et al*. 2002) was that 'corresponding measurements from the two platforms showed poor correlation […] implying a poor prognosis for a broad utilization of gene expression measurements across platforms'. The correlations we report in this paper between the measurements obtained with cDNA microarrays (our data set) and with oligonucleotide chips (Mody *et al*. 2001) are much higher than those reported by Kuo *et al*. (2002). Several factors contribute to this difference: (i) differences in pre-processing (Lowess fit vs. constant dye normalization), (ii) use of averages across replicates and (iii) filtering of reliable expression profiles. In our case, the use of (i) and (ii), prior to comparison with the data from Mody *et al*. (2001), resulted in the mean and median correlations of 0.385 and 0.542. The use of only the filtered profiles increased the mean to 0.646 and the median to 0.787. Stringent filtering has therefore a major contribution in improving the quality of the data. This improvement comes at the cost of reducing the number of measurements available (stringent filtering selects 3341 clones out of 21 439) but this situation is similar to what happens with the Affymetrix platform where the use of replicates allowed the selection of 1926 clones out of about 11 000.

Third, and most important, the high overall similarity (median correlation 0.787) obtained between expression profiles *in vivo* and *in vitro* demonstrates that expression profiles of at least 50% of genes during neuronal

development *in vivo* and *in vitro* were remarkably similar, most likely reflecting the same genetic program of neuronal differentiation. Diaz *et al*. (2002) showed also recently that a group of genes in isolated cultured granule cells exhibited very similar temporal expression patterns as those observed in the cerebellum *in vivo* between P6 and P20. Apparently once the cells have taken a neuronal fate, the further program of gene expression is, for a period of time, largely independent of histological or anatomical context. Interpretation of our results shown in Fig. 5 has to take into account the differences in timing of the two compared experiments shown in Fig. 5(a). From the results for the optimal mapping of five time points shown in Fig. 5 and also from the results of mapping of four consecutive time points (data not shown), we conclude that not only was the programs of gene expression *in vivo* and *in vitro* largely the same, but also that its execution *in vitro* was faster than *in vivo*. It is clear that culture conditions (cell density, contact with glia) can affect the rate of neuronal differentiation [e.g. duration of the initial outgrowth of an axon (before it contacts a target cell), or the balance between dendritic elongation and branching (Berezovska *et al*. 1999; Redmond *et al*. 2000)]. It is therefore understandable that the transition from the situation tissue to the cell culture conditions may have a strong effect on the rate of differentiation. We do not know what causes the observed acceleration. We are tempted to speculate that the network of connections between neurones in culture is much less elaborate than in the brain (Pokorny and Yamamoto 1981) and thus it may take much less time *in vitro* to complete the 'wiring' and to establish active synapses. It is also clear that the similarity between the situation *in vivo* and *in vitro* brakes at some point past Stage 5, as the neurones in the culture ultimately die.

Taken together our findings clearly demonstrate that hippocampal neurones *in vitro* are a remarkable relevant biological system to investigate hippocampal neuronal differentiation.

Acknowledgements

The authors thank K. Coddens, R. Maes and K. Seeuws from the VIB microarray facility for their excellent technical help, T. Bogaert for database assistance.

This work was supported by the VIB, the K.U. Leuven (research council, GOA Mefisto 666, IDO), the FWO-Vlaanderen, IWT (STWW, GBOU), AWI (Bil.Int.Coll.), the EU, FP5 DIADEM and the DWTC (IUAP V-22 and IUAP V-19). MD is a Marie Curie fellow (QLK6-CT-2000–52154). WA and YM are postdoctoral fellows of the FWO-Vlaanderen.

References

Aach J. and Church G. M. (2001) Aligning gene expression time series with time warping algorithms. *Bioinformatics* **17**, 495–508.

Ahnert-Hilger G., Kutay U., Chahoud I., Rapoport T. and Wiedenmann B. (1996) Synaptobrevin is essential for secretion but not for the development of synaptic processes. *Eur. J. Cell Biol.* **70**, 1–11.

Annaert W. G., Levesque L., Craessaerts K., Dierinck I., Snellings G., Westaway D., George-Hyslop P. S., Cordell B., Fraser P. and De Strooper B. (1999) Presenilin 1 controls gamma-secretase processing of amyloid precursor protein in pre-golgi compartments of hippocampal neurons. *J. Cell Biol.* **147**, 277–294.

Ashburner M., Ball C. A., Blake J. A., Botstein D., Butler H., Cherry J. M., Davis A. P., Dolinski K., Dwight S. S., Eppig J. T., Harris M. A., Hill D. P., Issel-Tarver L., Kasarskis A., Lewis S., Matese J. C., Richardson J. E., Ringwald M., Rubin G. M. and Sherlock G. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29.

Berezovska O., McLean P., Knowles R., Frosh M., Lu F. M., Lux S. E. and Hyman B. T. (1999) Notch1 inhibits neurite outgrowth in postmitotic primary neurons. *Neuroscience* **93**, 433–439.

Bergmann M., Grabs D. and Rager G. (1999) Developmental expression of dynamin in the chick retinotectal system. *J. Histochem. Cytochem.* **47**, 1297–1306.

Blake J. A., Richardson J. E., Bult C. J., Kadin J. A. and Eppig J. T. (2002) The Mouse Genome Database (MGD): the model organism database for the laboratory mouse. *Nucleic Acids Res.* **30**, 113–115.

Boyer C., Schikorski T. and Stevens C. F. (1998) Comparison of hippocampal dendritic spines in culture and in brain. *J. Neurosci.* **18**, 5294–5300.

Bradke F. and Dotti C. G. (2000a) Establishment of neuronal polarity: lessons from cultured hippocampal neurons. *Curr. Opin. Neurobiol.* **10**, 574–581.

Bradke F. and Dotti C. G. (2000b) Changes in membrane trafficking and actin dynamics during axon formation in cultured hippocampal neurons. *Microsc. Res. Tech.* **48**, 3–11.

Butte A. J., Tamayo P., Slonim D., Golub T. R. and Kohane I. S. (2000) Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. *Proc. Natl Acad. Sci. USA* **97**, 12182–12186.

Casanova J. E., Wang X., Kumar R., Bhartur S. G., Navarre J., Woodrum J. E., Altschuler Y., Ray G. S. and Goldenring J. R. (1999) Association of Rab25 and Rab11a with the apical recycling system of polarized Madin-Darby canine kidney cells. *Mol. Biol. Cell* **10**, 47–61.

Chen Y. A. and Scheller R. H. (2001) SNARE-mediated membrane fusion. *Nat. Rev. Mol. Cell Biol.* **2**, 98–106.

Chilcote T. J., Galli T., Mundigl O., Edelmann L., McPherson P. S., Takei K. and De Camilli P. (1995) Cellubrevin and synaptobrevins: similar subcellular localization and biochemical properties in PC12 cells. *J. Cell Biol.* **129**, 219–231.

Di Fiore P. P. and De Camilli P. (2001) Endocytosis and signaling: an inseparable partnership. *Cell* **106**, 1–4.

Diaz E., Ge Y., Yang Y. H., Loh K. C., Serafini T. A., Okazaki Y., Hayashizaki Y., Speed T. P., Ngai J. and Scheiffele P. (2002) Molecular analysis of gene expression in the developing pontocerebellar projection system. *Neuron* **36**, 417–434.

Dornan S., Jackson A. P. and Gay N. J. (1997) Alpha-adaptin, a marker for endocytosis, is expressed in complex patterns during Drosophila development. *Mol. Biol. Cell* **8**, 1391–1403.

Dotti C. G., Sullivan C. A. and Banker G. A. (1988) The establishment of polarity by hippocampal neurons in culture. *J. Neurosci.* **8**, 1454–1468.

Faire K., Trent F., Tepper J. M. and Bonder E. M. (1992) Analysis of dynamin isoforms in mammalian brain: dynamin-1 expression is spatially and temporally regulated during postnatal development. *Proc. Natl Acad. Sci. USA* **89**, 8376–8380.

Fletcher T. L. and Banker G. A. (1989) The establishment of polarity by hippocampal neurons: the relationship between the stage of a cell's development in situ and its subsequent development in culture. *Dev. Biol.* **136**, 446–454.

Fletcher T. L., Cameron P., De Camilli P. and Banker G. (1991) The distribution of synapsin I and synaptophysin in hippocampal neurons developing in culture. *J. Neurosci.* **11**, 1617–1626.

Fletcher T. L., De Camilli P. and Banker G. (1994) Synaptogenesis in hippocampal cultures: evidence indicating that axons and dendrites become competent to form synapses at different stages of neuronal development. *J. Neurosci.* **14**, 6695–6706.

Goslin K. and Banker G. (1991) Rat hippocampal neurons in low-density culture, in *Culturing Nerve Cells* (Goslin K., ed.), pp. 251–282. MIT Press, Cambridge, Massachusetts.

Haas C. A. and DeGennaro L. J. (1988) Multiple synapsin I messenger RNAs are differentially regulated during neuronal development. *J. Cell Biol.* **106**, 195–203.

He Z., Wang K. C., Koprivica V., Ming G. and Song H. J. (2002) Knowing how to navigate: mechanisms of semaphorin signaling in the nervous system. *Sci. STKE* **119**, RE1.

Hopf F. W., Waters J., Mehta S. and Smith S. J. (2002) Stability and plasticity of developing synapses in hippocampal neuronal cultures. *J. Neurosci.* **22**, 775–781.

Jahn R. and Sudhof T. C. (1999) Membrane fusion and exocytosis. *Annu. Rev. Biochem.* **68**, 863–911.

Jareb M. and Banker G. (1997) Inhibition of axonal growth by brefeldin A in hippocampal neurons in culture. *J. Neurosci.* **17**, 8955–8963.

Jarousse N. and Kelly R. B. (2001) Endocytotic mechanisms in synapses. *Curr. Opin. Cell Biol.* **13**, 461–469.

Kruskal J. B. and Liberman M. (1999) The symmetric time-warping problem: from continuous to discrete, in *Time Warps, String Edits, and Macromolecules: the Theory and Practice of Sequence Comparison* (Kruskal J., ed.), pp. 125–161. CSLI Publications, Stanford.

Kuo W. P., Jenssen T. K., Butte A. J., Ohno-Machado L. and Kohane I. S. (2002) Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* **18**, 405–412.

Lapierre L. A., Kumar R., Hales C. M., Navarre J., Bhartur S. G., Burnette J. O., Provance D. W. Jr, Mercer J. A., Bahler M. and Goldenring J. R. (2001) Myosin vb is associated with plasma membrane recycling systems. *Mol. Biol. Cell* **12**, 1843–1857.

Lin R. C. and Scheller R. H. (2000) Mechanisms of synaptic vesicle exocytosis. *Annu. Rev. Cell Dev. Biol.* **16**, 19–49.

Lloyd T. E., Verstreken P., Ostrin E. J., Phillippi A., Lichtarge O. and Bellen H. J. (2000) A genome-wide search for synaptic vesicle cycle proteins in Drosophila. *Neuron* **26**, 45–50.

Lockhart D. J., Dong H., Byrne M. C., Follettie M. T., Gallo M. V., Chee M. S., Mittmann M., Wang C., Kobayashi M., Horton H. and Brown E. L. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14**, 1675–1680.

Martinez-Arca S., Coco S., Mainguy G., Schenk U., Alberts P., Bouille P., Mezzina M., Prochiantz A., Matteoli M., Louvard D. and Galli T. (2001) A common exocytotic mechanism mediates axonal and dendritic outgrowth. *J. Neurosci.* **21**, 3830–3838.

Mochida S. (2000) Protein–protein interactions in neurotransmitter release. *Neurosci. Res.* **36**, 175–182.

Mody M., Cao Y., Cui Z., Tay K. Y., Shyong A., Shimizu E., Pham K., Schultz P., Welsh D. and Tsien J. Z. (2001) Genome-wide gene expression profiles of the developing mouse hippocampus. *Proc. Natl Acad. Sci. USA* **98**, 8862–8867.

Nemoto Y. and De Camilli P. (1999) Recruitment of an alternatively spliced form of synaptojanin 2 to mitochondria by the interaction with the PDZ domain of a mitochondrial outer membrane protein. *EMBO J.* **18**, 2991–3006.

Parlati F., Varlamov O., Paz K., McNew J. A., Hurtado D., Sollner T. H. and Rothman J. E. (2002) Distinct SNARE complexes mediating membrane fusion in Golgi transport based on combinatorial specificity. *Proc. Natl Acad. Sci. U S A* **99**, 5424–5429.

Pokorny J. and Yamamoto T. (1981) Postnatal ontogenesis of hippocampal CA1 area in rats. I. Development of dendritic arborisation in pyramidal neurons. *Brain Res. Bull.* **7**, 113–120.

Puskas L. G., Zvara A., Hackler L. Jr and Van Hummelen P. (2002) RNA amplification results in reproducible microarray data with slight ratio bias. *Biotechniques* **32**, 1330–1334, 1336, 1338, 1340.

Redmond L., Oh S. R., Hicks C., Weinmaster G. and Ghosh A. (2000) Nuclear Notch1 signaling and the regulation of dendritic development. *Nat. Neurosci.* **3**, 30–40.

Ross D. T., Scherf U., Eisen M. B., Perou C. M., Rees C., Spellman P., Iyer V., Jeffrey S. S., Van de Rijn M., Waltham M., Pergamenschikov A., Lee J. C., Lashkari D., Shalon D., Myers T. G., Weinstein J. N., Botstein D. and Brown P. O. (2000) Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* **24**, 227–235.

Schena M., Shalon D., Davis R. W. and Brown P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470.

Skutella T. and Nitsch R. (2001) New molecules for hippocampal development. *Trends Neurosci.* **24**, 107–113.

Stettler O., Tavitian B. and Moya K. L. (1996) Differential synaptic vesicle protein expression in the barrel field of developing cortex. *J. Comp. Neurol.* **375**, 321–332.

Sugiyama T., Shinoe T., Ito Y., Misawa H., Tojima T., Ito E. and Yoshioka T. (2000) A novel function of synapsin II in neurotransmitter release. *Brain Res. Mol. Brain Res.* **85**, 133–143.

Tang B. L. (2001) Protein trafficking mechanisms associated with neurite outgrowth and polarized sorting in neurons. *J. Neurochem.* **79**, 923–930.

Tolar L. A. and Pallanck L. (1998) NSF function in neurotransmitter release involves rearrangement of the SNARE complex downstream of synaptic vesicle docking. *J. Neurosci.* **18**, 10250–10256.

Tou J. T. and Gonzalez R. C. (1979) Pattern classification by distance functions, in *Pattern Recognition Principles* (Gonzalez R. C., ed.), pp. 75–109. Addison-Wesley, Reading, Massachusetts.

Tsai S. C., Adamik R., Tsuchiya M., Chang P. P., Moss J. and Vaughan M. (1991) Differential expression during development of ADP-ribosylation factors, 20-kDa guanine nucleotide-binding protein activators of cholera toxin. *J. Biol. Chem.* **266**, 8213–8219.

Wong P. P., Daneman N., Volchuk A., Lassam N., Wilson M. C., Klip A. and Trimble W. S. (1997) Tissue distribution of SNAP-23 and its subcellular localization in 3T3-L1 cells. *Biochem. Biophys. Res. Commun.* **230**, 64–68.

Yang Y. H., Dudoit S., Luu P., Lin D. M., Peng V., Ngai J. and Speed T. P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30**, e15.