# Analysis of complex brain disorders with gene expression microarrays: schizophrenia as a disease of the synapse

Károly Mirnics, Frank A. Middleton, David A. Lewis and Pat Levitt

**The level of cellular and molecular complexity of the nervous system creates unique problems for the neuroscientist in the design and implementation of functional genomic studies. Microarray technologies can be powerful, with limitations, when applied to the analysis of human brain disorders. Recently, using cDNA microarrays, altered gene expression patterns between subjects with schizophrenia and controls were shown. Functional data mining led to two novel discoveries: a consistent decrease in the group of transcripts encoding proteins that regulate presynaptic function; and the most changed gene, which has never been previously associated with schizophrenia, regulator of G-protein signaling 4. From these and other findings, a hypothesis has been formulated to suggest that schizophrenia is a disease of the synapse. In the context of a neurodevelopmental model, it is proposed that impaired mechanics of synaptic transmission in specific neural circuits during childhood and adolescence ultimately results in altered synapse formation or pruning, or both, which manifest in the clinical onset of the disease.**

**Károly Mirnics**
Depts of Psychiatry, Neurobiology and PittArray, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA.

**Frank A. Middleton**
Dept of Neurobiology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA.

**David A. Lewis**
Depts of Psychiatry and Neuroscience, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA.

**Pat Levitt***
Dept of Neurobiology and PittArray, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA.
*e-mail: plevitt+@pitt.edu

The advent of modern functional genomic approaches holds great promise for revealing novel molecular and cellular mechanisms that contribute to disease[1]. Northern hybridizations have given way to microarrays[2,3], PCR has evolved into a high-throughput quantitative real-time assay[4,5] and serial analysis of gene expression (SAGE)[6] has emerged as an ingenious combination of sequence-independent restriction enzyme digestion with high-volume sequencing efforts. These methods, among many others[7–10] provide information about gene expression in a tissue or cell. However, they differ in their capabilities for sample throughput, in the type of data obtained and in the depth of information analysis[11], which makes each strategy complementary rather than competitive (Fig. 1a).

DNA microarrays, which are based on well-established principles of nucleic acid hybridization, interrogate samples for the presence of thousands of genes simultaneously[2,3,17]. So far, microarrays have been successfully applied in comparative genomic hybridization[8,18–20], on-chip sequencing[21,22], novel gene discovery[23–25] and gene expression profiling[2,3,26]. These studies are often carried out in conjunction with conventional methods for assessment of transcriptome differences, including *in situ* hybridization[27], suppression subtractive hybridization[28] and representational difference analysis (RDA)[25].

Gene expression patterns in the human brain exceed the complexity of other organ systems by an order of magnitude. The degree of difficulty in the analysis of such patterns is magnified in the investigation of psychiatric disorders, which appear to result from the interplay of polygenic and epigenetic factors on multiple brain circuits (for reviews, see Refs 29,30). Examples of microarray-based analysis of gene expression in schizophrenia[27,31], and other studies, are used to highlight particular advantages and limitations that investigators face when designing and carrying out such studies.

## Microarray choices for gene expression analysis

Currently, two widely accepted microarray types are used for gene expression analysis: oligonucleotide and cDNA microarrays[2,20,32–34], although both have weaknesses[12,35]. The major manufacturer of oligonucleotide arrays is Affymetrix (for a technology review, see Refs 33,36), which offers dozens of distinct microarrays. There are several advantages of the Affymetrix 'GeneChip®' technology, including the presence of multiple oligonucleotide features on the chip that interrogate different 3′ regions of the same gene, quantitative results and good nominal sensitivity[33]. Furthermore, data comparison can be performed *post hoc* across many microarrays simultaneously.

cDNA microarrays immobilize longer fragments of genes onto a solid support[2,3,37]. These microarrays are readily customized, their production is not proprietary, and the printing and scanning equipment, as well as the software tools for data analysis, are of modest cost and available from many sources. For comparative gene expression studies, fluorescence-tagged nucleotides [for example, Cyanine 3 (Cy3) and Cyanine 5 (Cy5)] or fluorochrome-couplers are incorporated into the cDNA during a standard reverse transcription reaction. Such labeled samples are combined and hybridized to the same microarray. Each microarray probe will bind its complementary, fluorescently labeled cDNA species with a distinct

http://tins.trends.com 0166-2236/01/$ – see front matter  PII: S0166-2236(00)01862-2

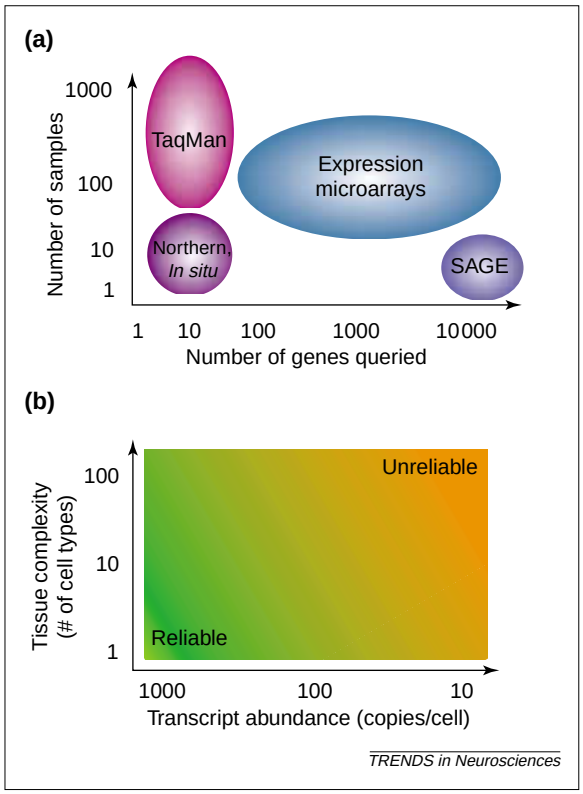


Fig. 1. Microarray throughput and sensitivity. (a) Throughput versus information in gene expression analysis. The most commonly used tools of functional genomics are complementary, rather than competitive. In combination, these novel tools of functional genomics will facilitate the investigation of the most complex questions related to diseases of the nervous system (reviewed in Refs 12,13). Serial analysis of gene expression (SAGE) allows exhaustive analysis of all known and unknown transcripts in a few valuable samples, whereas real-time quantitative RT–PCR can assay thousands of samples for expression of a limited number of genes. Oligonucleotide and cDNA microarrays fill the gap between these two techniques. Conventional assays have limited throughput and analyze the expression of only a few genes; however, they are very sensitive and are usually the best choice for data verification. (b) Dilution versus detection in gene expression analysis. Although the nominal microarray sensitivity is constant, the biological sensitivity of microarrays is positively correlated with the abundance of the mRNA species in the sample and inversely correlated with the phenotypic complexity of the tissue. Specifically, if one is examining gene expression in subtypes of local circuit neurons, the transcripts from these cells (and hence the sensitivity of the assay to detect changes in such transcripts) will be serially diluted at multiple levels. First, as glial mRNA comprises ~50% of the total mRNA in gray matter, the biological sensitivity of the array will be halved. Second, a further dilution occurs by mRNA from projection neurons. Given that the ratio of projection neurons to interneurons in the human cortex is approximately 3:1 (Refs 14,15), this dilution will further decrease the sensitivity 3-fold. Third, if the transcript of interest is expressed in only one-third of local circuit neurons, the biological sensitivity of the array becomes 18 times less than the nominal microarray sensitivity. This serial dilution of rare transcripts (for example, receptors), coupled with less than ideal probe quality, might account for many 'non-expressed' calls on the microarrays[16].

excitation–emission wavelength, and the amount of signal for each fluorochrome can be quantified independently for each microarray spot. cDNA arrays on nylon or nitrocellulose membrane, using individual radiolabeled samples hybridized to separate filters, are the least expensive array type, and are often overlooked as a viable choice[38].

### Limitations of microarray technology vis-á-vis brain tissue

Both cDNA microarrays and oligonucleotide GeneChips® (Affymetrix) can detect as few as one in 250 000 mRNA copies. Theoretically, this sensitivity should detect even the sparsest of mRNA species in a uniform cell population[34,39,40]. Many low-abundance transcripts in brain samples, however, are not detected by microarrays. Unfortunately, increasing the absolute amount of the hybridized target usually will not increase the sensitivity of the microarray – the relative abundance of the transcript within the RNA pool, coupled with microarray probe characteristics, will influence the detection limit for each individual transcript. Thus, for analysis of transcripts expressed only in small subpopulations of neurons within the cerebral cortex, the rare transcripts will be effectively diluted beyond microarray sensitivity by many abundant mRNA species expressed across the tissue. Consequently, even if one had access to the entire genome on a microarray, the cellular complexity of the brain would obviate truly global gene expression (transcriptome) profiling (Fig. 1b).

What is the limit for detecting expression differences? A well-controlled single microarray experiment might detect changes as modest as 1.5-fold at the specimen mRNA level. Unfortunately, it cannot be presumed that functionally significant structural or chemical changes in brain diseases will reach a 50% level, and often they do not[41–44]. Therefore, some important expression differences might be reported as false negatives. Furthermore, alterations of this magnitude would be identified only if changes occurred across the majority of cells within a circumscribed brain region. Because of the cellular complexity of brain, many authentic gene expression changes will be diluted by mRNA from other cells in the sample that express normal levels of the same transcript, potentially masking even a complete transcript loss within neuronal subpopulations. This limitation, combined with the constraints of biological sensitivity, might explain the relatively few differences in gene expression between different brain regions in rodents reported recently[45,46]. As noted by Geschwind[35], the relatively few gene expression differences found between different mouse brain regions[45,46] might say more about the limitation of microarray sensitivity rather than biological phenomena.

Currently, most protocols that use microarrays require a prohibitive amount of starting material[47,48]. Hundreds of milligrams of wet weight human brain tissue are needed to isolate several micrograms of high quality, minimally degraded mRNA, of which at least several hundred nanograms are used in a single hybridization. This sample requirement can be overcome, to some degree, by amplification of the starting material either with PCR-based methods[49,50] or antisense RNA polymerase amplification[51–53],

which enables analysis of single cell gene expression profiles. As an alternative to sample amplification, several labeling modifications focus on enhancing signal intensity, which effectively reduces starting material requirements[54,55].

The source of mRNA is an important factor in all microarray studies. This is particularly crucial in comparisons of highly convoluted cortical tissue from gyrencephalic species, where gyrus shape is individually variable and potentially results in different gray–white matter ratios[16]. For example, the comparison of a sample from a diseased tissue with more mRNA of glial origin (more white matter) to a control sample with more mRNA of neuronal origin (more gray matter) will result in suggestive, but false-positive data. Such data might be interpreted incorrectly as an increased expression of glial markers and decreased expression of neuronal markers in the diseased sample. Real-time quantitative RT–PCR and northern analysis of brain samples are also subject to the same vulnerability in gray–white matter bias. However, *in situ* hybridization, as an anatomical verification strategy, provides precise information about cellular localization and true magnitude of transcript changes in individual affected neurons. It also reveals cell loss and allows co-localization studies, and is not affected by dilution with unwanted species of transcripts.

### Experimental design, comparison paradigms and data interpretation

The success of individual projects depends upon the strategic design of the experimental paradigm, which is in part determined by the available starting material and the choice of array platform. For example, the use of oligonucleotide GeneChips generates microarray data that can be compared *post hoc* with results obtained from many other arrays of the same type. Data quality is highly dependent upon array consistency. By contrast, with fluorescence-based cDNA microarrays, expression between two samples is compared on the same microarray, obviating individual microarray variability. The two-sample comparison strategy can use different pairing paradigms: (1) compare each individual sample to a different, closely matched control[27]; (2) compare each experimental sample to an aliquot of an identical reference (usually pooled control) on the same microarray; (3) compare on individual microarrays each experimental sample or control to a pooled, multi-tissue reference[48]; or (4) compare a pooled sample to a pooled control[56]. Each of these approaches has distinct advantages[16]. In the context of studies that use human material, preserving the individuality of samples permits the discovery of subject-specific gene expression patterns that might emerge as a consequence of a complex polygenic disease, such as schizophrenia[27]. Pooling of closely matched control material and comparing it with the individual experimental samples has the advantage of a uniform, unchanging reference with which the diseased subjects are compared. However, such pooling is vulnerable to normal variation in gene expression that occurs at different life stages and outliers that might create artificial baseline levels of gene expression. In an ideal situation, both matched pairwise and individual subject–pooled control comparisons should be made. Unfortunately, this is often not feasible because of constraints on tissue availability.

The assessment of assay noise is accomplished by dividing the same sample into two aliquots, labeling them with different fluorochromes and hybridizing them onto the same microarray. The obtained variability represents the combined noise of the reverse transcription–labeling, the hybridization procedure and the quality of the microarray. Replication of the microarray experiment with the same starting material (starting from a new sample isolation) is helpful and addresses the reliability of the data[57]. It is also essential in uncovering true gene expression changes in a 1.2–1.6-fold range. Whenever possible, all experiments should be performed at least in duplicate with the labels reversed. However, if resources are limited, same-sample repetitions might not be essential. Biologically, the same gene expression change seen in 12 of 21 individual subject comparisons might be more informative than the same changes seen in four of seven comparisons, repeated in triplicate.

### Data analysis considerations

Reporting the 'most changed genes' reduces the microarray approach to a high throughput northern hybridization strategy, failing to take advantage of simultaneously obtained transcriptome differences between experimental and control subjects. Many complex methods for data analysis are being developed[26,58–65]. In reality, however, for neuropsychiatric disorders that involve multiple circuits and symptoms, informatics methods for discovering truly complex expression patterns, where presence, absence and change of many gene products simultaneously define a disease, are currently not available. Such multifactorial and multilevel analyses should combine massive gene expression datasets with the clinical diagnosis, medication, family history of the illness and genetic heritage of each subject. Expression data eventually will be assessed for functional gene pathways[27,66] and related to disease–model databases and gene sequence databases (for example, single nucleotide polymorphisms), which add individuality to the diagnosis.

### Microarray studies related to human brain disorders

With the above-discussed technical advantages and limitations of gene expression microarrays in mind, it is no surprise that the first successful application of

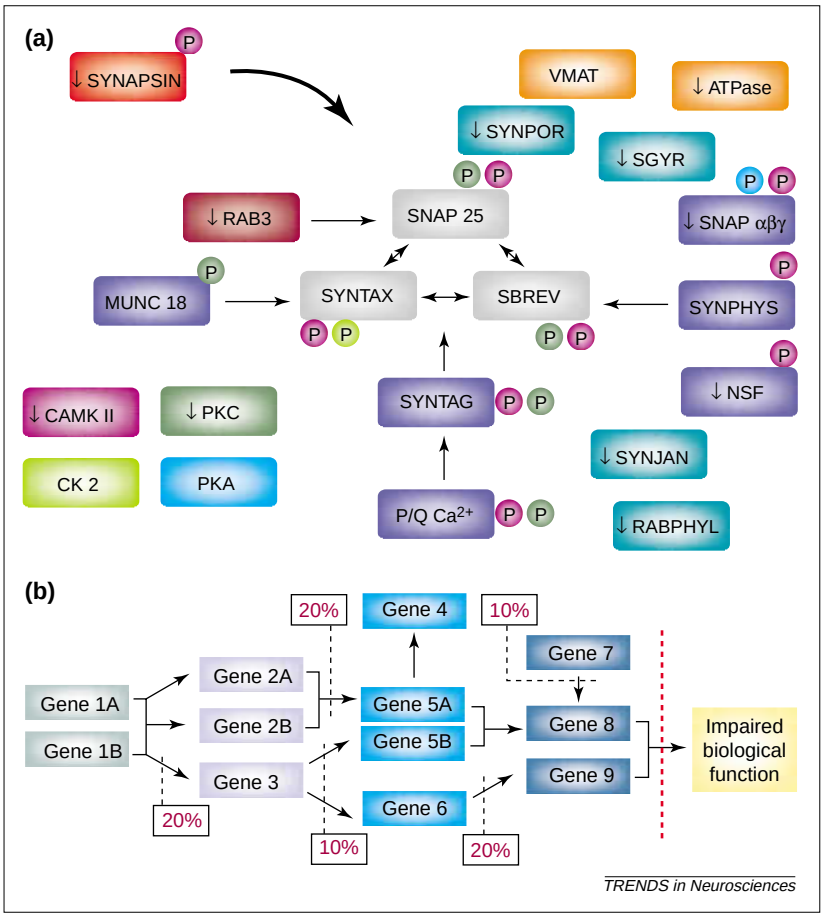


Fig. 2. The presynaptic secretory pathway. (a) At the presynaptic terminal, we found that in all subjects with schizophrenia, multiple PSYN genes interacting [directly (indicated by arrows) or indirectly] with a core complex[69,70] (shown as grey objects in centre) showed decreased expression (indicated by downward arrows in boxes). Surprisingly, the exact combination of PSYN genes with decreased expression was different in each of the comparisons. As a consequence, only a subset of subjects (for example, 20% for Gene 1B, 10% for Gene 7) might exhibit altered expression of a particular combination of PSYN genes, indicated in (b). In total, however, the majority of the subjects with schizophrenia will exhibit a consistent impairment of the PSYN pathway, indicated by broken red line in (b). 'P' denotes phosphorylation by kinases; the colour of the circle labelled 'P' indicates the kinase involved in the phosphorylation. Symbols for genes can denote complete gene families, which often perform redundant functions. Abbreviations: ATPase, vacuolar proton pump; CaMKII, $Ca^{2+}$calmodulin kinase 2; CK2, casein kinase 2; NSF, *N*-ethylmaleimide sensitive factor; PKA, protein kinase A; PKC, protein kinase C; P/Q $Ca^{2+}$, P/Q $Ca^{2+}$ channel; RABAPT, rabaptin; RABPHIL, rabphillin; SBREV, synaptobrevin; SNAP, soluble NSF attachment protein; SYNJAN, synaptojanin; SYNPHYS, synaptophysin; SYNTAG, synaptotagmin; SYNTX, syntaxin; VMAT, vesicular monoamine transporter.

these methods to the study of complex neurological and psychiatric disorders has been reported only recently. Whitney *et al.* compared gene expression in normal white matter with that found in acute lesions from one subject with multiple sclerosis (MS)[67]. In acute lesions, altered expression levels of many genes were associated with genes related to cell metabolism–structure, cytokines and cell adhesion molecules, and the technique also implicated genes that had not been previously associated with MS (for example, the gene encoding the Duffy chemokine receptor).

Ginsberg *et al.* analyzed, on low complexity cDNA arrays, individual Alzheimer's disease (AD) senile plaques (SPs) and normal hippocampal CA1 neurons[68]. They showed that β-amyloid-positive SPs contained numerous, predominantly neuronal species of mRNA. In a follow-up study[53], gene microarrays that contained 18 000 cDNA clones were used to compare the relative abundance of multiple mRNAs in tangle-bearing AD and in normal CA1 neurons. The combined results revealed specific decreases and increases in transcript levels of signaling and structural molecules in subjects with AD. Efforts in complex single-cell transcript profiling are under way in several other laboratories[51], and will probably represent a major focus of disease-oriented neuroscience research over a prolonged period of time.

These initial studies demonstrate that harvested samples from postmortem human brain are of sufficient quality for accurate and reproducible profiling of gene expression in different disease states. In order to discover meaningful expression patterns in the massive gene expression datasets, each of the studies attempted to use data analysis that extended beyond reporting the most changed genes in tabular form. However, the bioinformatics foundation of the known molecular and cellular functional cascades and pathways are not readily available to place the data in their most meaningful context (Fig. 2).

## Synaptic dysfunction in schizophrenia

Polygenic, complex disorders such as schizophrenia might ultimately benefit the most from novel functional genomics strategies. Schizophrenia affects more than 1% of the population worldwide, with clinical symptoms typically showing a late adolescence or early adulthood onset. They include a constellation of delusions, hallucinations, decreased motivation and impaired executive functions[29,30]. The underlying etiology of schizophrenia involves a combination of genetic and epigenetic factors, with over ten chromosomal loci linked to the disease[71–74]. A convergence of observations from clinical, neuroimaging and anatomical studies has focused attention on the prefrontal cortex as a major site of alterations in schizophrenia[75–77].

cDNA microarrays have been used to compare transcriptomes in matched subjects[27], analyzing >7800 genes and expressed sequence tags (ESTs) in the initial study[27]. Each experiment compared an individual with schizophrenia with a different control that was matched for gender, postmortem interval (PMI), age and brain pH. This 'transcriptome difference' was subjected to a data analysis scheme that defined specific, complex patterns of altered gene expression associated with schizophrenia. In one analysis, genes encoding proteins that cooperate to perform specific cellular functions were grouped, which allowed the assessment of changes for over 250 gene groups (Fig. 2). This analysis identified only a few gene groups that consistently differed between subjects with schizophrenia and matched controls. In all subjects with schizophrenia, the most changed gene group was related to presynaptic secretory function (PSYN gene group; Fig. 3), and, in particular,

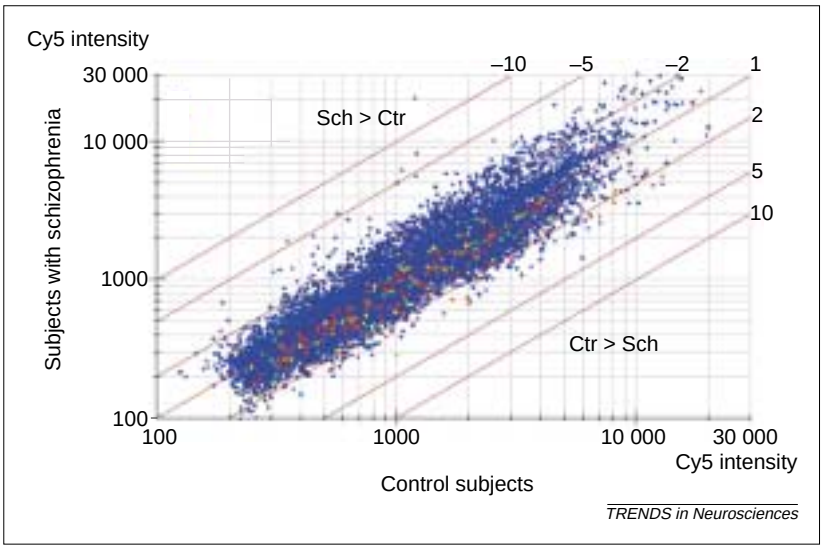


Fig. 3. Schizophrenia as a presynaptic secretory machinery deficit. Experimental data, plotted as Cy3 versus Cy5 fluorescent signal ratio across six pairwise comparisons. Each point represents a single gene expression observation. Any point on the line labelled '1' has equal expression between a subject with schizophrenia and matched control. For all the expressed genes (blue) the expression differences are comparably distributed in both control > schizophrenia (below the line labeled '1') and schizophrenia > control (above the line labeled '1') classes. However, the majority of the transcripts related to presynaptic secretory machinery (PSYN, red), glutamate-mediated neurotransmission (Glu, yellow) and GABA-mediated neurotransmission (GABA, green) were consistently decreased in all subjects with schizophrenia. Within the PSYN gene group, we found that the expression levels of two genes, encoding *N*-ethylmaleimide-sensitive factor (NSF) and synapsin 2 (SYN2) were consistently decreased. In order to verify these findings, *in situ* hybridization was carried out on sections obtained from the same tissue blocks and from a new cohort of subjects[27].

the 'mechanics' of neurotransmitter release. However, the most affected genes in the PSYN group varied across subjects with schizophrenia, which suggests the illness has distinct 'molecular signatures' (Fig. 2). Schizophrenia, therefore, might involve a combination of different sequence-related polygenic susceptibility factors and physiological adaptations that lead to impairment of the same function – signaling between neurons.

The microarrays included targets for both neuron-specific and ubiquitous genes. Using a broad survey, it was found that the expression of regulator of G-protein signaling 4 (RGS4) was reduced in all but one individual with schizophrenia examined[31]. In the brain, RGS4 is one of over 20 RGS family members that serve as GTPase-activating proteins (GAPs), which reduce response duration of postsynaptic neurons after the release of presynaptic neurotransmitters that bind to G-protein-coupled receptors, including metabotropic glutamate, 5-$HT_2$ and D2 receptors (for a review, see Ref. 78). RGS4 had not been studied in schizophrenia previously, but microarray analysis to date has revealed a dramatic (50–84%) decrease in its levels in ten subjects examined. Moreover, RGS4 levels were decreased across motor and visual regions of the same subjects with schizophrenia. These observations were verified by *in situ* hybridization in the original subjects with schizophrenia and in a set of additional subjects.

What might be the physiological outcome of a severe reduction in RGS4 expression? In experimental systems, reduction of RGS levels prolongs postsynaptic signaling[78,79]. In animal models, RGS4 levels are altered in response to acute stress[80]. Thus, the decreased RGS4 expression in the cortices of subjects with schizophrenia might be relevant to the exacerbation of symptoms typically exhibited by subjects with schizophrenia in the face of stress.

Using the already acquired microarray data, the specificity of the RGS4 change was examined retrospectively in two ways: by gene group and by chromosomal location. The gene group analysis showed that the G-protein signaling cluster, with over 200 genes represented on the cDNA array, was unchanged[31]. Given this finding, it was hypothesized that the PSYN and RGS4 expression changes were linked in a crucial way to produce synaptic pathophysiology. Remarkably, RGS4 deficiency, coupled with decreased efficiency of presynaptic neurotransmitter release, is consistent with the therapeutic efficacy of both haloperidol and atypical antipsychotics[81], which target $G_i/G_q$-coupled receptors that might be regulated by *RGS4* (Ref. 78).

RGS4 maps to chromosomal region 1q21–22 (NCBI genomic Accession Number AC031977.6), which has recently been reported as a highly significant schizophrenia susceptibility locus[82]. Of the 71 genes represented with microarray probes from 1q21–22 cytogenetic region, only levels of RGS4 are altered consistently. Genomic screening studies are currently being carried out across a large population of subjects with schizophrenia.

Of course, a variety of confounds related to lifestyle, medical history and treatment can influence any study of human disease, so that data interpretation should be made cautiously. In these analyses, there are several reasons that medication exposure probably did not account for the specific and consistent gene expression changes observed. There was no constant pattern of drug therapy at the time of death and, additionally, the assessment of a unique non-human primate model of chronic haloperidol treatment[83] demonstrated unchanged expression of both the PSYN gene group and RGS4 individually in dorsolateral prefrontal cortex[27,31]. Furthermore, lifetime incidence of alcohol abuse or dependence, which was present in a subset of the schizophrenic subjects, was not associated uniquely with decreases in PSYN or *RGS4* gene expression. In this regard, recent reports clearly show that the expression of PSYN genes was increased, rather than decreased, in a prefrontal cortex study of alcohol abuse[84]. At least one independent group of investigators has reproduced and expanded the essence of these findings[56]. Many of the same PSYN genes (or family members) were consistently changed across the two different platforms and subject populations, indicating that several of the conserved changes might emerge as hallmarks of the disease. Interestingly, preliminary reports[85] have found that

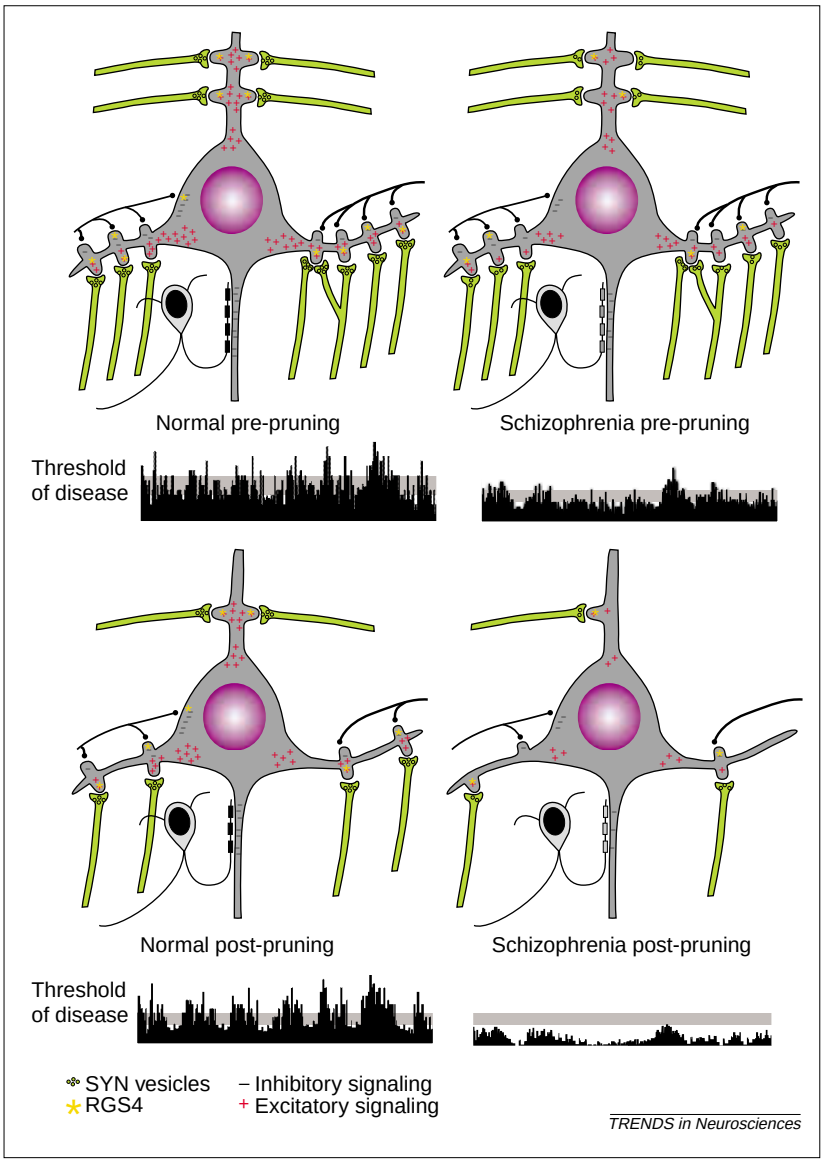


Fig. 4. A synaptic–neurodevelopmental model of schizophrenia. This model proposes an altered development of cortical circuits as a result of disruption in SYN and regulator of G-protein signaling 4 (RGS4) function. During childhood, an over-abundance of excitatory and inhibitory synapses is produced in the cerebral cortex[87,88]. As a result of molecular defects and altered SYN or *RGS4* gene function, synaptic drive is decreased (fewer +, indicating excitatory signaling and –, indicating inhibitory signaling) in subjects who will develop schizophrenia. Although core complex formation, release or recycling of individual vesicles (SYN vesicles shown by green circles), or both, can be impaired, the physiological deficiency remains clinically asymptomatic until synaptic pruning ends. Before pre-adolescence, the initial overproduction of synapses compensates for functional deficits, thus retaining circuit function above the disease threshold. During the period of normal pruning, from adolescence through to puberty, altered synaptic function might contribute to abnormal decreases of specific classes of synapse (for example, on basilar dendrites). In addition, continued decrease in synaptic drive might result in further loss of both inhibitory and excitatory synapses. As a result, the physiological synaptic 'buffer' is lost, resulting in the subthreshold state for normal circuit function. Then, the decreased number of synapses is not able to compensate for impaired synapse function, and the clinical manifestations of the disease emerge. In an attempt to compensate for inefficient presynaptic release during development, postsynaptic changes might follow, including (but not limited to) downregulation of RGS4. This could result in a compensatory increase in duration of signaling through G-protein-coupled receptors. However, these adaptational mechanisms might not be sufficiently effective (or even desirable), and, as a result of the impaired synaptic release (and consequent lack of normal synaptic drive), overpruning of the presynaptic neuropil could occur.

in the rodent perinatal hippocampal lesion model of schizophrenia[86], RGS4 expression is significantly reduced in the prefrontal cortex when these rats reach adulthood.

## A synaptic-neurodevelopmental model of schizophrenia

In light of these findings, a model is proposed[27] suggesting that different sequence mutations or polymorphisms in genes related to synaptic communication, combined with other factors, might lead to the shared clinical manifestations of schizophrenia. The basic tenets of this model have four components:

(1) The etiology of schizophrenia involves a polygenic pattern of inheritance, resulting in altered function of proteins that control the 'mechanics' of synaptic transmission.

(2) The heterogeneity in expression defects within the PSYN group reflects distinct adaptive capacities of different populations of neurons. The alterations in RGS4 expression could reflect a primary gene defect or a postsynaptic adaptation in an effort to enhance neurotransmission, which thus compensates for the deficits in PSYN gene function.

(3) Deficits in PSYN and *RGS4* gene expression might affect the extended postnatal developmental process of synapse formation and pruning[87,88], which ultimately provides a link to the neurodevelopmental timecourse of schizophrenia[72]. In the model (Fig. 4), impaired synaptic transmission in subjects with schizophrenia is present from early life (probably in a subset of synapses), but the exuberant synaptic connections that form during the third trimester of gestation and early childhood compensate for a functional synaptic impairment. These exuberant synapses are pruned by late adolescence, uncovering the existing synaptic impairment. Inadequate synaptic activity might lead to overpruning and manifestations of the symptoms of schizophrenia.

(4) Independent of their possible roles in causing the disease, the deficits in PSYN and *RGS4* expression have physiological and behavioral consequences relevant to the pathophysiology of schizophrenia, and could help explain many of the cognitive deficits that arise from prefrontal cortex dysfunction in schizophrenics. Knockout studies in mice, in which specific members of the PSYN group have been deleted, reveal the complexity of physiological deficits that occur because of presynaptic dysfunction[69,89–91].

## Future directions

Several of the genes identified as being differentially expressed in subjects with schizophrenia have been mapped to loci implicated in the pathogenesis of the disease. However, in schizophrenia, the majority of the consistent gene expression changes might be adaptive in nature. The importance of these non-inherited changes should not be overlooked – they are potentially related to the symptoms of the disease and might prove to be useful drug targets. The key to controlling the most debilitating symptoms, where the predisposing factors can be present

developmentally, before the onset of clinical symptoms, could lie in targeting the epigenetic changes that are clearly associated with the conserved symptoms of schizophrenia. Indeed, the majority of strategies for current disease treatment of many medical conditions are often based on symptomatic therapy (for example, antipyretics, antihistamines, painkillers).

The combined linkage-expression data and the rapid development of proteomics[92], will facilitate the understanding of schizophrenia in a broader, functional context. New candidate susceptibility genes will guide the creation of transgenic and gene-deficient animals, and advances in biobehavioral sciences will allow us to comprehensively dissect the behavioral profiles associated with specific patterns of gene expression changes. Although further away, investigators have the potential to develop reliable, microarray-based diagnostic tools to assess, premorbidly, the risk of each individual for developing the disease. Finally, a much more in-depth understanding of the biological basis of schizophrenia and other neuropsychiatric disorders will use a targeted pharmacogenomics approach to design therapeutic compounds[93] that will result in treatment breakthroughs.

## References


1 Collins, F.S. and Mansoura, M.K. (2001) The Human Genome Project. *Cancer* 91, 221–225
2 Schena, M. *et al.* (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470
3 DeRisi, J. *et al.* (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* 14, 457–460
4 Freeman, W.M. *et al.* (1999) Quantitative RT-PCR: pitfalls and potential. *BioTechniques* 26, 112–122, 124–125
5 Heid, C.A. *et al.* (1996) Real time quantitative PCR. *Genome Res.* 6, 986–994
6 Velculescu, V.E. *et al.* (1995) Serial analysis of gene expression. *Science* 270, 484–487
7 Sutcliffe, J.G. *et al.* (2000) TOGA: an automated parsing technology for analyzing expression of nearly all genes. *Proc. Natl. Acad. Sci. U. S. A.* 97, 1976–1981
8 Drmanac, S. *et al.* (1998) Accurate sequencing by hybridization for DNA diagnostics and individual genomics. *Nat. Biotechnol.* 16, 54–58
9 Ferguson, J.A. *et al.* (1996) A fiber-optic DNA biosensor microarray for the analysis of gene expression. *Nat. Biotechnol.* 14, 1681–1684
10 Walt, D.R. (2000) Techview: molecular biology. Bead-based fiber-optic arrays. *Science* 287, 451–452
11 Kozian, D.H. and Kirschbaum, B.J. (1999) Comparative gene-expression analysis. *Trends Biotechnol.* 17, 73–78
12 Watson, S.J. *et al.* (2000) The 'chip' as a specific genetic tool. *Biol. Psychiatry* 48, 1147–1156
13 Colantuoni, C. *et al.* (2000) High throughput analysis of gene expression in the human brain. *J. Neurosci. Res.* 59, 1–10
14 Winfield, D.A. *et al.* (1980) An electron microscopic study of the types and proportions of neurons in the cortex of the motor and visual areas of the cat and rat. *Brain* 103, 245–258
15 Hendry, S.H. *et al.* (1987) Numbers and proportions of GABA-immunoreactive neurons in different areas of monkey cerebral cortex. *J. Neurosci.* 7, 1503–1519
16 Mirnics, K. *et al.* DNA microarrays and human brain disorders. In *Methods in Neurogenetics* (Chin, H. and Moldin, S., eds), CRC Press (in press)
17 Lockhart, D.J. and Winzeler, E.A. (2000) Genomics, gene expression and DNA arrays. *Nature* 405, 827–836
18 Cheung, V.G. *et al.* (1998) Linkage-disequilibrium mapping without genotyping. *Nat. Genet.* 18, 225–230
19 Hacia, J.G. *et al.* (1996) Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-colour fluorescence analysis. *Nat. Genet.* 14, 441–447
20 Geschwind, D.H. *et al.* (1998) Klinefelter's syndrome as a model of anomalous cerebral laterality: testing gene dosage in the X chromosome pseudoautosomal region using a DNA microarray. *Dev. Genet.* 23, 215–229
21 Kozal, M.J. *et al.* (1996) Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat. Med.* 2, 753–759
22 Gunthard, H.F. *et al.* (1998) Comparative performance of high-density oligonucleotide sequencing and dideoxynucleotide sequencing of HIV type 1 pol from clinical samples. *AIDS Res. Hum. Retroviruses* 14, 869–876
23 Schena, M. *et al.* (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl. Acad. Sci. U. S. A.* 93, 10614–10619
24 Tanaka, T.S. *et al.* (2000) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9127–9132
25 Geschwind, D. *et al.* (2001) A genetic analysis of neural progenitor differentiation. *Neuron* 29, 325–339
26 Golub, T.R. *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, 531–537
27 Mirnics, K. *et al.* (2000) Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. *Neuron* 28, 53–67
28 Yang, G.P. *et al.* (1999) Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes. *Nucleic Acids Res.* 27, 1517–1523
29 Lewis, D.A. and Lieberman, J.A. (2000) Catching up on schizophrenia: natural history and neurobiology. *Neuron* 28, 325–334
30 Carpenter, W.T. and Buchanan, R.W. (1994) Schizophrenia. *New Engl. J. Med.* 330, 681–690
31 Mirnics, K. *et al.* (2001) Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia. *Mol. Psychiatry* 6, 293–301
32 Duggan, D.J. *et al.* (1999) Expression profiling using cDNA microarrays. *Nat. Genet.* 21 (Suppl. 1), 10–14
33 Lipshutz, R.J. *et al.* (1999) High density synthetic oligonucleotide arrays. *Nat. Genet.* 21 (Suppl. 1), 20–24
34 Lockhart, D.J. *et al.* (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 14, 1675–1680
35 Geschwind, D.H. (2000) Mice, microarrays, and the genetic diversity of the brain. *Proc. Natl. Acad. Sci. U. S. A.* 97, 10676–10678
36 Barlow, C. and Lockhart, D.J. Neural gene expression analysis using DNA arrays. In *Methods in Neurogenetics* (Chin, H. and Moldin, S., eds), CRC Press (in press)
37 Bowtell, D.D. (1999) Options available – from start to finish – for obtaining expression data by microarray. *Nat. Genet.* 21 (Suppl. 1), 25–32
38 Bertucci, F. *et al.* (1999) Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples. *Hum. Mol. Genet.* 8, 1715–1722
39 Wodicka, L. *et al.* (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotechnol.* 15, 1359–1367
40 Kane, M.D. *et al.* (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res.* 28, 4552–4557
41 Collinge, J. and Curtis, D. (1991) Decreased hippocampal expression of a glutamate receptor gene in schizophrenia. *Br. J. Psychiatry* 159, 857–859
42 Akbarian, S. *et al.* (1995) Gene expression for glutamic acid decarboxylase is reduced without loss of neurons in prefrontal cortex of schizophrenics. *Arch. Gen. Psychiatry* 52, 258–278
43 Woo, T.U. *et al.* (1998) A subclass of prefrontal gamma-aminobutyric acid axon terminals are selectively altered in schizophrenia. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5341–5346
44 Volk, D.W. *et al.* (2000) Decreased glutamic acid decarboxylase67 messenger RNA expression in a subset of prefrontal cortical gamma-aminobutyric acid neurons in subjects with schizophrenia. *Arch. Gen. Psychiatry* 57, 237–245
45 Sandberg, R. *et al.* (2000) From the cover: regional and strain-specific gene expression mapping in the adult mouse brain. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11038–11043
46 Lockhart, D.J. and Barlow, C. (2001) Expressing what's on your mind: DNA arrays and the brain. *Nat. Rev. Neurosci.* 2, 63–68
47 Hegde, P. *et al.* (2000) A concise guide to cDNA microarray analysis. *BioTechniques* 29, 548–560
48 Alizadeh, A.A. *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403, 503–511
49 Herrler, M. (2000) Use of SMART-generated cDNA for differential gene expression studies. *J. Mol. Med.* 78, B23

50 Zhumabayeva, B. *et al.* (2001) Use of SMART-generated cDNA for gene expression studies in multiple human tumors. *Biotechniques* 30, 158–162

51 Luo, L. *et al.* (1999) Gene-expression profiles of laser-captured adjacent neuronal sub-types. *Nat. Med.* 5, 117

52 Eberwine, J. *et al.* (1992) Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci. U. S. A.* 89, 3010–3014

53 Ginsberg, S.D. *et al.* (2000) Expression profile of transcripts in Alzheimer's disease tangle-bearing CA1 neurons. *Ann. Neurol.* 48, 77–87

54 Stears, R.L. *et al.* (2000) A novel, sensitive detection system for high-density microarrays using dendrimer technology. *Physiol. Genomics* 3, 93–99

55 Adler, K. *et al.* (2000) MICROMAX: a highly sensitive system for differential gene expression on microarrays. In *Microarray Biochip Technology* (Schena, M., ed.), Eaton Publishing

56 Vawter, M. *et al.* (2001) Application of cDNA microarrays to examine gene expression differences in schizophrenia. *Brain Res. Bull.* (in press)

57 Lee, M.L. *et al.* (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl. Acad. Sci. U. S. A.* 97, 9834–9839

58 Eisen, M.B. *et al.* (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14863–14868

59 Hastie, T. *et al.* (2000) 'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns. *Genome Biol.* 1, 0003.0001–0003.0021

60 Hastie, T. *et al.* (2001) Supervised harvesting of expression trees. *Genome Biol.* 2, 0003.0001–0003.0012

61 Brown, M.P. *et al.* (2000) Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc. Natl. Acad. Sci. U. S. A.* 97, 262–267

62 Toronen, P. *et al.* (1999) Analysis of gene expression data using self-organizing maps. *FEBS Lett.* 451, 142–146

63 Claverie, J.M. (1999) Computational methods for the identification of differential and coordinated gene expression. *Hum. Mol. Genet.* 8, 1821–1832

64 Furey, T.S. *et al.* (2000) Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 16, 906–914

65 Tamayo, P. *et al.* (1999) Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. U. S. A.* 96, 2907–2912

66 Kanehisa, M. and Goto, S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30

67 Whitney, L.W. *et al.* (1999) Analysis of gene expression in mutiple sclerosis lesions using cDNA microarrays. *Ann. Neurol.* 46, 425–428

68 Ginsberg, S.D. *et al.* (1999) Predominance of neuronal mRNAs in individual Alzheimer's disease senile plaques. *Ann. Neurol.* 45, 174–181

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

70 Linial, M. (1997) SNARE proteins – why so many, why so few? *J. Neurochem.* 5, 1781–1792

71 Andreasen, N.C. (1996) Pieces of the schizophrenia puzzle fall into place. *Neuron* 16, 697–700

72 Weinberger, D.R. (1995) From neuropathology to neurodevelopment. *Lancet* 346, 552–557

73 Tsuang, M. (2000) Schizophrenia: genes and environment. *Biol. Psychiatry* 47, 210–220

74 Pulver, A.E. (2000) Search for schizophrenia susceptibility genes. *Biol. Psychiatry* 47, 221–230

75 Andreasen, N.C. *et al.* (1997) Hypofrontality in schizophrenia: distributed dysfunctional circuits in neuroleptic-naive patients. *Lancet* 349, 1730–1734

76 Bertolino, A. *et al.* (2000) Specific relationship between prefrontal neuronal *N*-acetylaspartate and activation of the working memory cortical network in schizophrenia. *Am. J. Psychiatry* 157, 26–33

77 Weinberger, D. *et al.* (1986) Physiologic dysfunction of dorsolateral prefrontal cortex in schizophrenia. I. Regional cerebral blood flow evidence. *Arch. Gen. Psychiatry* 43, 114–124

78 De Vries, L. *et al.* (2000) The regulator of G protein signaling family. *Annu. Rev. Pharmacol. Toxicol.* 40, 235–271

79 Berman, D.M. *et al.* (1996) The GTPase-activating protein RGS4 stabilizes the transition state for nucleotide hydrolysis. *J. Biol. Chem.* 271, 27209–27212

80 Ni, Y.G. *et al.* (1999) Region-specific regulation of RGS4 (Regulator of G-protein-signaling protein type 4) in brain by stress and glucocorticoids: *in vivo* and *in vitro* studies. *J. Neurosci.* 19, 3674–3680

81 Lieberman, J.A. *et al.* (1998) Serotonergic basis of antipsychotic drug effects in schizophrenia. *Biol. Psychiatry* 44, 1099–1117

82 Brzustowicz, L.M. *et al.* (2000) Location of a major susceptibility locus for familial schizophrenia on chromosome 1q21-q22. *Science* 288, 678–682

83 Pierri, J.N. *et al.* (1999) Alterations in chandelier neuron axon terminals in the prefrontal cortex of schizophrenic subjects. *Am. J. Psychiatry* 156, 1709–1719

84 Lewohl, J.M. *et al.* (2000) Gene expression in human alcoholism: microarray analysis of frontal cortex. *Alcohol Clin. Exp. Res.* 24, 1873–1882

85 Marcotte, E.R. *et al.* (2000) Gene expression changes in adult prefrontal cortex and nucleus accumbens following neonatal ventral hippocampal lesions. *Soc. Neurosci. Abstr.* 30, 385.314

86 Sams-Dodd, F. *et al.* (1997) Neonatal lesions of the rat ventral hippocampus result in hyperlocomotion and deficits in social behaviour in adulthood. *Psychopharmacology* 132, 303–310

87 Huttenlocher, P. (1979) Synaptic density in human frontal cortex – developmental changes and effects of aging. *Brain Res.* 163, 195–205

88 Bourgeois, J.P. *et al.* Formation, elimination and stabilization of synapses in the primate cerebral cortex. In *The Cognitive Neurosciences* (Gazzaniga, M., ed.), MIT Press (in press)

89 Verhage, M. *et al.* (2000) Synaptic assembly of the brain in the absence of neurotransmitter secretion. *Science* 287, 864–869

90 Schluter, O.M. *et al.* (1999) Rabphilin knock-out mice reveal that rabphilin is not required for rab3 function in regulating neurotransmitter release. *J. Neurosci.* 19, 5834–5846

91 Rosahl, T.W. *et al.* (1995) Essential functions of synapsins I and II in synaptic vesicle regulation. *Nature* 375, 488–493

92 Dutt, M.J. and Lee, K.H. (2000) Proteomic analysis. *Curr. Opin. Biotechnol.* 11, 176–179

93 Destenaves, B. and Thomas, F. (2000) New advances in pharmacogenomics. *Curr. Opin. Chem. Biol.* 4, 440–444