Analysis of Differential Gene Expression by Boolean Selection: Identification of Transcripts Associated with the Formation and Reorganization of Neuronal Networks.

Thesis

UNRESTRICTED

Analysis of Differential Gene Expression by Boolean Selection: Identification of Transcripts Associated with the Formation and Reorganization of Neuronal Networks

A THESIS PRESENTED BY

ALEXANDER DETTMAR CRAWFORD

TO THE

OPEN UNIVERSITY
MILTON KEYNES, ENGLAND

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

MASTER OF PHILOSOPHY

MAX PLANCK INSTITUTE FOR DEVELOPMENTAL BIOLOGY
TÜBINGEN, GERMANY

SEPTEMBER 1996

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

The quantitative comparison of gene expression represents one of the major approaches for the analysis of gene function. Recently, the discovery of unknown genes based only upon their expression pattern has become possible through the application of various differential cloning methods, including subtractive hybridization, differential display, and expressed sequence tag analysis. Due to numerous technical and procedural limitations, however, none of these technologies is currently capable of rapidly identifying differentially regulated transcripts based on multiple expression-related criteria. Therefore, an alternative approach was devised which allows for the sequential application of numerous selection criteria during the initial screening process, thereby enabling the rapid identification of transcripts with highly specific expression patterns. Because of its use of multiple positive and negative selection steps, this approach is called Boolean selection.

One of the first applications of this new approach was the search for novel genes associated with the formation and reorganization of neuronal networks in the mammalian brain. An ideal model system for this analysis is the retinocollicular system of the rat, where numerous aspects of neurite outgrowth, axon guidance, synapse formation and reorganization, and regulation of neuronal survival can easily be studied. Boolean selection was applied for the identification of transcripts preferentially expressed in the superior colliculus during late prenatal development as well as after the loss of afferent input in the adult animal. An initial analysis of several of these selected transcripts revealed that all satisfied at least one selection criterion and that over half satisfied both, indicating the usefulness of Boolean selection for the rapid identification of differentially regulated genes.

Although additional analysis by *in situ* hybridization will still be necessary to verify the extent of differential expression and to provide information on the cellular localization of the transcripts identified so far, it was possible with this preliminary screen to gain new insight into some of the molecular and cellular activities in the superior colliculus during periods associated

with the formation and reorganization of synaptic connections. Some of the identified transcripts encode proteins associated with post-transcriptional protein synthesis and modification events, including ribophorin (an essential subunit of the protein glycosylation complex oligosaccharyltransferase), calnexin (a membrane-associated molecular chaperone of the endoplasmic reticulum), and translation initiation factor, which together suggest enhanced levels of protein production and secretion. Other transcipts include rSec6 (part of the protein complex involved in the docking and fusion of synaptic vesicles to the plasma membrane) and Fyn (a nonreceptor tyrosine kinase found at high concentrations in axonal growth cones and also known be involved in synaptic plasticity), Fas-associated protein factor (a recently-identified molecule potentially involved in Fas-mediated signal transduction during programmed cell death), and a novel member of the EGF superfamily of transmembrane growth factors (which presently includes the EGF/TGFα family and the neuregulins). Further analysis of these molecules may provide additional information about the signaling mechanisms and cell-cell interactions at work during the establishment and modification of neuronal networks.

ACKNOWLEDGMENTS

The research presented in this thesis was performed at the Max Planck Institute for Developmental Biology in Tübingen, Germany under the supervision of Prof. Friedrich Bonhoeffer and under the external supervision of Dr. Rob Krumlauf of the National Institute for Medical Research in London, England.

The first phase of the project, the development of the Boolean selection technique, was done in collaboration with Camila V. Esguerra of the Max Planck Institute for Developmental Biology. The second phase of the project, the application of this technique to the analysis of differential gene expression in the retinocollicular system of the rat, was done in collaboration with Dr. Mathias Bähr and his colleagues of the Department of Neurology of the University of Tübingen.

I would like to thank Elke Hoffmann for her excellent technical assistance throughout the second phase of the project. I also thank Dr. Mathias Bähr for his enthusiasm and optimism. In addition, I am indebted to several members of his group, past and present, for their surgical support.

Last but not least, I am very grateful to Prof. Friedrich Bonhoeffer for his generous support and patience, especially with regard to my numerous, diverse projects.

INTRODUCTION

An effective strategy for the analysis of complex embryological events and for the identification of key components involved in regulating these events has been the use of genetic screens (Nüsslein and Wieschaus, 1980; Horwitz et al., 1983; Jürgens et al., 1995; Haffter et al., 1996). This type of functional approach has certain limitations, however, because many genes have multiple roles throughout development. This limits the effectiveness of such screens to the analysis of processes occuring relatively early in development, as mutations in genes with essential functions in these early processes can often result in such significant perturbations that the systematic analysis of subsequent events is not possible. In addition, certain genes cannot be identified by genetic screens because of inherent functional redundancies or other compensatory mechanisms that come into effect when their encoded proteins are missing or nonfunctional, as has been suspected to occur for numerous targeted mutations in mice (Shastry, 1994). Finally, many developmental processes, especially later events such as the formation of functional neural circuits, are difficult to analyze by genetic means because of the inefficient screening procedures necessary for the determination of phenotype. Large-scale screens are also restricted to very few model organisms because of numerous issues regarding generation time, fecundity, cost of maintainance, and suitability for systematic analysis.

Other functional approaches for the identification of genes with specific biological roles include protein purification and expression cloning, yet these require the use of sensitive *in vitro* assays, which are often quite difficult or sometimes even impossible to establish for certain biological activities and *in vivo* systems. In addition, these strategies are ineffective for the identification of genes that encode the different constituents of protein complexes consisting of multiple, independent factors, as they are usually limited to the identification of single genes. These disadvantages also hold true for descriptive approaches that involve comprehensive comparisons of protein expression profiles, such as two-dimensional gel electrophoresis and methods based on the differential generation of antibodies.

With the exception of expression cloning and genetic screens that rely on insertional mutagenesis strategies, one additional limitation common to all approaches mentioned above is the additional time required for the actual cloning of individual genes after the initial screening effort. In the case of protein-based methods, this usually requires the microsequencing of peptides for subsequent cDNA cloning steps, which can be an error-prone and technically difficult process. In the case of standard genetic screens using chemical or radiation-induced mutagenesis, the identification of individual mutations can usually only be accomplished by positional cloning, which requires relatively detailed genetic map information in addition to significant effort even with invertebrates such as *Drosophila* or *Caenorhabditis*, which have relatively short generation times. In addition, because of the low resolution of the current genetic map and its relatively long generation time, it is not yet possible to readily identify mutant genes in the zebrafish, which is currently the only vertebrate model organism amenable to systematic genetic analysis.

An alternative approach with significant potential for the process of gene discovery is one based on the analysis of differential mRNA expression. This strategy relies on the fact that specific cellular responses to external and internal stimuli are determined by corresponding changes in genetic activity. These changes are reflected by alterations in the transcription rates of various genes and can therefore be monitered through the quantitative examination of messenger RNA populations. There are several different methods that have been developed over recent years that employ this strategy for the identification of differentially regulated transcripts, albeit with varying degrees of efficiency.

Identification of Differentially Regulated Transcripts

Current approaches for the analysis of differential gene expression, including subtractive hybridization (Sargent and Dawid, 1983; Hedgewick et al., 1984), differential display (Liang and Pardee, 1992), and expressed sequence tag (EST) analysis (Okubo et al., 1992), as well as various PCR-based procedures such as suppression subtractive hybridization (Diatchenko et al., 1996) and representational difference analysis (Hubank and Schatz,

1994), all have numerous disadvantages that make them inadequate for the rapid identification of genes based on multiple expression-related criteria. Differential display, also known as mRNA fingerprinting, involves the global PCR amplification of mRNA with combinations of arbitrary primers, resulting in the generation of amplicons of discrete length which are then visualized by polyacylamide gel electrophoresis. Some of the primary disadvantages associated with differential display and other, similar procedures include (1) low reproducibility because of an inherent susceptibility to minor changes in reaction conditions, (2) low sensitivity because of an inherent bias for abundantly expressed transcripts, and (3) significant requirements in terms of labor and reagent consumption, which together preclude the use of these approaches for the targeted identification of specific genes. Some of these points are also true for the analysis of differential gene expression by tag sequencing. This approach involves the generation of primary cDNA libraries and the single-pass sequencing of several hundred to several thousand randomly chosen clones, allowing the subsequent comparison of transcript abundance profiles based on the frequencies of different sequences. Because of the significant costs associated with this method, and also because of the low sampling depth that is due to the limited number of different clones that can be sequenced (given the high redundancy of primary cDNA libraries), EST analysis is not an option for the rapid identification of tissue-specific or differentially regulated transcripts.

One of the main disadvantages associated with subtractive hybridization procedures is an inherent procedural limitation that results in the generation of cDNA populations that only partially fulfill the original selection criteria. This limitation is common to all standard protocols where the enrichment process involves physical subtractions using mRNA or cDNA that has the original transcript abundance profile of the tissue or cell type from which it was isolated. This results in the effective C_0t of most moderately-expressed to weakly-expressed transcript species remaining well below the values needed for the efficient removal these transcripts from the sample that is to be enriched for differentially expressed genes. This reduction of only the more abundant transcripts results in the relative enrichment of all weakly-expressed transcripts, regardless of whether they are common to both samples or actually differentially expressed. The subtracted sample is therefore only partially made up of differentially expressed transcripts,

increasing the time and effort needed for the relatively labor-intensive secondary screening procedures.

In addition, due to the large number of genes that are differentially expressed in different cells or tissues, regardless of whether the difference in question is of spatial, temporal, or conditional nature, it is extremely difficult to identify the few genes that may encode proteins responsible for any specific functional differences between these cells or tissues. Most protocols require significant amounts of mRNA and therefore relatively large amounts of source material, which usually results in the mRNA sample representing the genetic activities of many different cell types. It is for these reasons that subtractive cDNA library screens for genes encoding defined protein activities, or even for genes directly associated with their expression or activity (such as transcription factors, receptors, or components of specific signal transduction pathways), have proven notoriously ineffective in isolating these genes.

Multiple-Parameter Expression Screens by Boolean Selection

For this reason an alternative approach is proposed which allows for the rapid identification of transcripts that fulfill multiple selection criteria. This is accomplished by the generation, in parallel, of several highly-enriched pools of differentially regulated transcripts by an improved subtractive hybridization strategy, and by the subsequent isolation of cDNAs from two or more of these enriched pools by the newly-developed rescue hybridization procedure. Because of the ability of this approach to selectively purify differentially expressed transcripts by the application of several positive and negative selection steps, it is designated as Boolean selection (see Figure 1). This selection scheme represents the first effective method for the rapid cloning of specific genes on the basis of *multiple* expression-related criteria.

For example, for the identification of transcripts encoding proteins thought to be expressed only in a specific region of the brain (region A) at a specific timepoint during development (timepoint X), the Boolean selection strategy would initially involve two negative selections achieved by the separate subtractive hybridizations {region A NOT region B} and {timepoint X NOT

timepoint Y}. These two subtractions can be abbreviated as {A NOT B} and {X NOT Y}. In this case NOT is understood to mean "greater than by factor *n*", where *n* reflects the stringency of subtraction. Standard subtraction approaches would be restricted to the generation and analysis of only one of these libraries because of their inherent limitation to negative selection schemes, thereby necessitating the subsequent analysis of large numbers of clones selected on the basis of only this one parameter (spatial or temporal). However, the introduction of a positive selection step, achieved by a rescue hybridization that allows for the enrichment of cDNAs common to both subtracted pools, enables the rapid selection of those cDNA clones that fulfill *both* criteria {A NOT B} AND {X NOT Y}. With this selection scheme it is therefore possible to efficiently identify transcripts with highly regulated expression patterns simply through the application of multiple expression-based selection criteria.

Formation and Reorganization of Synaptic Connections in the Rat Superior Colliculus

An ititial test of the effectiveness of the Boolean selection strategy for the identification of differentially regulated transcripts was its use for the analysis of gene expression changes associated with periods of synapse formation and reorganization in the mammalian brain. An excellent model system with which to approach this question is the superior colliculus (SC) of the rat, one of the primary targets of retinal ganglion cells (in addition to the lateral geniculate nucleus) and the mammalian equivalent of the optic tectum of lower vertebrates. The superior colliculus is ideal for the analysis of the various molecular mechanisms involved in the establishment of neuronal networks, in part because of the well-characterized, topographically ordered projection of retinal ganglion cell axons onto the retinorecipient layers of the SC during development (allowing the analysis of invariant processes such as the guidance of retinal axons by target-specific molecules) and because of the accessibility of the visual system for experimental manipulation (allowing the analysis of plastic processes that are more dependent upon function, such as activity-dependent synaptic strengthening).

In addition to the normal establishment of synaptic connections during development, which is initiated shortly before birth and is largely complete within the first three postnatal weeks, there is an artificial reorganization of synapses that occurs after eye removal or after axotomy of the optic nerve in adult animals (Lund and Lund, 1971a, 1971b). This reorganization takes place as the result of the lesion-induced degeneration of retinal axons, which leads to the loss of afferent inputs for neurons within the retinorecipient layers of the SC. After the loss of optic terminals, the synaptic sites of these neurons are vacated and reoccupied by local terminals (Maters, 1977), indicating that although retinal fibers are unable to regenerate through the optic nerve because of myelin-associated neurite growth inhibitors (Schwab, 1990), their target neurons are still capable of inducing axonal sprouting and reestablishing synapses with these axons. This raises questions as to the molecular nature of this response and its relation to events that occur during development.

The target-specific factors that are involved in modulating the deafferentation-induced sprouting response are still unknown and may be similar to those that regulate the initial formation of synaptic connections during development. Recent evidence points to the involvement of insulin-like growth factor-1 (IGF-1) in stimulating reactive axonal growth in the adult hippocampus following the removal of afferent inputs (Guthrie et al., 1995), yet it is not yet clear what role this factor has during development, although additional results also indicate its ability to promote the survival of both embryonic and adult motoneurons (Neff et al., 1993). Expression of IGF-1 and other, perhaps unknown growth and/or trophic factors in the superior colliculus during phases of synapse formation and reorganization may be involved in the regulation of both axonal outgrowth and the survival of presynaptic neurons.

Other mechanisms that are likely to have significant roles during the formation of synaptic connections in the superior colliculus include interactions between target-specific chemoattractants and their corresponding receptors on the ingrowing axons of afferent neurons (von Boxberg et al., 1993). Despite the recent cloning of repulsive factors that regulate the guidance of retinal ganglion cell axons (Drescher et al., 1995), no attractive guidance molecules have yet been identified in the superior colliculus. Recent evidence, however, appears to indicate an upregulation in

the expression of attractive components in the superior colliculus following the lesion-induced loss of afferent inputs (Wizenmann et al., 1993; Bähr and Wizenmann, 1996). Nevertheless, it has not yet been determined whether these factors are also expressed during development, raising the possibility that this response is associated only with deafferentation.

In order to identify target-specific molecules involved in the various cellular processes associated with synapse formation and reorganization, transcripts preferentially expressed in the superior colliculus during late prenatal development as well as after the loss of afferent inputs were isolated by Boolean selection. The identification of such transcripts may enable the elucidation of the cellular responses induced by target-specific factors that are expressed by post-synaptic neurons during the formation and reorganization of neuronal networks.

MATERIALS AND METHODS

Collection of Tissue Samples

Superior colliculi from rat embryos at embryonic day 18 (E18), from normal adult rats, and from deafferented adult rats were obtained by standard surgical procedures and were immediately frozen in cryotubes in liquid nitrogen. Sprague-Dawley rats were used in all cases. For the deafferentation of adult rat superior colliculi, adult rats were anaesthetized with 7% chloralhydrate (420 mg/kg body weight) and one optic nerve was unilaterally transsected intraorbitally immediately behind the eye. The contralateral superior colliculus was removed at 21 days post-axotomy.

Isolation of poly(A)$^+$ RNA

Polyadenylated RNA was isolated with oligo$(dT)_{25}$-coupled paramagnetic latex microspheres (Dynabeads, Dynal) according to the manufacturer's specifications with several modifications. Briefly, for each tissue sample 200 µl Dynabeads suspension (5 mg/ml Dynabeads, 250 mM Tris-HCl pH 8.0, 20 mM EDTA, 0.1% Tween 20, 0.02% NaN_3) were washed once with, and subsequently resuspended in, 200 µl lysis/binding buffer (0.5 M LiCl, 100 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% LiDS, 5 mM DTT). Frozen tissue samples were quickly transferred from cryotubes in liquid nitrogen to 1.5-ml Eppendorf tubes containing 600 µl lysis/binding buffer and 3 µl 20 mg/ml Proteinase K (Boehringer Mannheim) and homogenized by drawing them several times though successively smaller needles (18-gauge, 23-gauge, and 27-gauge) on 3-ml syringes. Any remaining cell debris was pelleted by centrifuging the tubes at 4°C in an Eppendorf microcentrifuge at maximum speed for 1 minute. The resuspended Dynabeads were added to the supernatants and these mixtures were then briefly vortexed and incubated on ice for 10 minutes. After concentrating the Dynabeads on a magnetic particle concentrator (MPC, Dynal) for 5 minutes, the particles were washed several times by resuspension and reconcentration, twice with 0.5 ml 150 mM LiCl, 10 mM Tris-HCl pH 8.0, 10 mM EDTA, 0.1% LiDS and three times with 0.5 ml 150 mM LiCl, 10 mM Tris-HCl pH 8.0, 10 mM EDTA.

The poly(A)$^+$ RNA was eluted from the Dynabeads by resuspending them in 80 µl 0.2 mM EDTA pH 8.0, incubating the suspension at 65°C for 5 minutes, and removing the particles by magnetic concentration.

After spectrophotometrically determining the nucleic acid concentration using quartz cuvettes pretreated for 30 minutes with 1:1 concentrated HCl / methanol, the poly(A)$^+$ RNA samples were subjected to DNase treatment for the removal of any remaining trace amounts of genomic DNA prior to cDNA synthesis. The reaction mixture, consisting of 80 µl poly(A)$^+$ RNA sample in 0.2 mM EDTA pH 8.0, 10 µl 10X DNase buffer (0.4 M Tris-HCl pH 7.5, 60 mM $MgCl_2$, 20 mM $CaCl_2$), 1 µl 10 mg/ml tRNA (Boehringer Mannheim), 2 µl 40 U/µl RNase inhibitor (Boehringer Mannheim), 1 µl of a 1:10 dilution in 1X DNase buffer of 10 U/µl DNase I (Stratagene, La Jolla), and 6 µl DEPC-treated dH_2O, was incubated at 37°C for 30 minutes, then extracted twice with 100 µl 25:24:1 phenol (acidic) / chloroform / isoamyl alcohol (Amersham) and once with 100 µl 24:1 choroform / isoamyl alcohol (Amersham). The DNase-treated poly(A)$^+$ RNA was recoupled to 100 µl fresh Dynabeads in lysis/binding buffer with 0.5 µl 20/mg Proteinase K, and was then washed and eluted as described above, with the wash volumes reduced to 250 µl and the elution volume to 10 µl.

Synthesis of cDNA and Adaptor Ligation

Synthesis of double-stranded cDNA was carried out with several components of the Superscript Plasmid System (Life Technologies) according the manufacturer's specifications with several modifications. Briefly, 1.5 µg DNase-treated poly(A)$^+$ RNA in 9 µl 0.2 mM EDTA pH 8.0 was combined with 2 µl 1 µM random hexadeoxynucleotides (Pharmacia Biotech), incubated at 70°C for 10 minutes, then cooled on ice for 2 minutes. To this mixture was added 4 µl 5X first-strand buffer (250 mM Tris-HCl pH 7.5, 375 mM Kcl, 15 mM $MgCl_2$), 2 µl 100 mM DTT, and 1 µl 10 mM dNTPs (Pharmacia Biotech). After pre-incubating this mixture at 37°C for 2 minutes, 2 µl 200 U/µl M-MLV RNase H$^-$ reverse transcriptase (Superscript II, Life Technologies) was added and this reaction was incubated at 37°C for 30 minutes, at 45°C for 20 minutes, and at 50°C for 10 minutes. The reaction was then placed on ice and the following components were added: 91 µl DEPC-treated dH_2O, 30 µl 5X second-strand buffer (100 mM Tris-

HCl pH 6.9, 0.45 M KCl, 23 mM $MgCl_2$, 0.75 mM β-NAD^+, 50 mM $(NH_4)_2SO_4$), 3 µl 10 mM dNTPs, 4 µl 10 U/µl *E. coli* DNA polymerase I, 1 µl 10 U/µl *E. coli* DNA ligase, and 1 µl 2 U/µl *E. coli* RNase H (all enzymes from Life Technologies). After incubation at 16°C for 2 hours, 2 µl 3 U/µl T4 DNA polymerase (New England Biolabs) were added and the reaction was kept at 16°C for an additional 5 minutes, after which it was stopped by placing it on ice and adding 10 µl 0.5 M EDTA pH 8.0.

The double-stranded cDNA was purified from this reaction mixture by adsorbtion to silica-gel microparticles (Qiaex, Qiagen) according to the manufacturer's specifications. The cDNA was eluted in 30 µl 10 mM Tris-HCl pH 8.0 and divided into 2 15-µl aliquots, one of which was restricted with *Alu* I (AG/CT) and the other with both *Alu* I and *Rsa* I (GT/AC) in 30-µl reaction volumes with 20 U of each restriction enzyme (New England Biolabs) and including 10 mM Bis Tris Propane-HCl pH 7.0, 10 mM $MgCl_2$, and 1 mM DTT. Both restriction digests were incubated at 37°C for 4 hours, after which the restricted DNA was then purified with Qiaex particles and eluted in 20 µl 10 mM Tris-HCl pH 8.0.

Synthetic adaptors (Wang and Brown, 1991; Balzer and Bäumlein, 1994) were generated by annealing pairs of complementary oligodeoxynucleotides (custom synthesis by Pharmacia Biotech) to form two different double-stranded DNA molecules, each with one blunt end and one four-base non-complementary overhang at the other end. One pair, CTCTTG-CTTGAATTCGGACTA and pTAGTCCGAATTCAAGCAAGAGCACA, annealed to form adaptor R, which included an *Eco*R I restriction site (G/AATTC) near the blunt end, whereas the other pair, AGTTAC-ACGTCTAGAATGGCT and pAGCCATTCTAGACGTGTAACTGATA, annealed to form adaptor X, which included an *Xba* I restriction site (T/CTAGA) near the blunt end ('p' represents phosphate). The annealing reaction was carried out by combining equimolar amounts of the two complementary oligodeoxynucleotides in dH_2O for a total DNA concentration of 1 µg/µl, heating this mixture to 75°C, and allowing it to slowly cool to 25°C over at least 30 minutes.

Adaptors were ligated to the *Alu* I and *Alu* I/*Rsa* I restriction fragments of the different cDNA pools according to the following scheme: the X adaptor was ligated to the embryonic superior colliculus cDNA pool (E_X) and to one

of two separate normal adult superior colliculus cDNA pools (N_X), and the R adaptor was ligated to the other normal adult superior colliculus cDNA pool (N_R) and to the deafferented adult superior colliculus cDNA pool (D_R). The ligation reactions included 10 µl 1 µg/µl adaptor, 20 µl Qiaex-purified *Alu* I or *Alu* I/*Rsa* I-restricted cDNA in 10 mM Tris-HCl pH 8.0, 10 µl 10X ligation buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, 25 µg/ml BSA), 5 µl 400 NEB U/µl T4 DNA ligase (New England Biolabs), and 55 µl dH_2O, and were incubated at 16°C for 12 hours, after which the cDNA was purified by adsorption to silica-gel membrane spin columns with an exclusion limit for double-stranded DNA of 75 base-pairs (Qiaquick PCR Purification Columns, Qiagen). The purification procedure was carried out according the manufacturer's specifications and the cDNA was eluted in 50 µl TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). At this point the *Alu* I and *Alu* I/*Rsa* I-restricted cDNA fragments for each pool were combined.

Amplification of cDNA Fragments

PCR amplification of the adaptor-ligated cDNA fragments was carried out using the 21-mer component of each adaptor as the respective primer (R-21 and X-21). The PCR reactions consisted of 5 µl purified and adaptor-ligated cDNA fragments in TE, 10 µl 10X reaction buffer (0.5 M KCl, 100 mM Tris-HCl pH 8.8, 15 mM $MgCl_2$, 1% gelatin), 1.5 µl 25 mM dNTPs, 4 µl 25 mM $MgCl_2$, 1 µl 1 µg/µl primer in dH_2O, 0.5 µl 5 U/µl Taq DNA polymerase (Pharmacia Biotech), and dH_2O up to 100 µl. The reactions were thermocycled (Omnigene, Hybaid) according to the following program: 3 minutes at 95°C, 30 cycles of {15 seconds at 95°C, 30 seconds at 50°C, 2 minutes at 72°C plus 5 seconds extension per cycle} followed by 10 minutes at 72°C. The amplified cDNA fragments were purified with Qiagen spin columns as described above and eluted in 50 µl TE. The DNA concentration of each sample was determined spectrophotometrically and adjusted to 100 ng/µl. Each pool was then reamplified essentially as described above, with the modifications that 100 pg cDNA were used as template material and that both an unmodified and a biotinylated primer (custom synthesis, Pharmacia Biotech) were used for separate amplification reactions for the generation of tracer and driver cDNA, respectively. Each

successive round of reamplification (after each subtraction and rescue hybridization) was carried out in the same fashion.

Subtractions and Rescue Hybridizations

Subtractions were carried out by combining 20 µl 100ng/µl biotinylated driver cDNA amplicons with 1 µl 100ng/µl unmodified tracer cDNA amplicons, incubating this mixture at 95°C for 5 minutes, then adding 20 µl 2X HB^+ buffer (2 M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 20 mM cetyltrimethylammonium bromide (CTAB, Fluka)) and overlaying this mixture with mineral oil and incubating it at 68°C for 16 hours. Driver-driver and driver-tracer hybrids were removed from this mixture by transferring the aqueous solution under the mineral oil to an Eppendorf tube containing 100 µl 10 mg/ml streptavidinylated Dynabeads (Dynal) washed once with, and resuspended in, 1X HB^+ buffer. This suspension was then incubated at 68°C with vigorous shaking (Thermomixer, Eppendorf) for 15 minutes. The Dynabeads were quickly removed by magnetic concentration and the supernatant was transferred to another Eppendorf containing fresh Dynabeads, and the procedure was repeated. Aliquots (5 µl) of the final supernatant were used directly as templates for successive rounds of PCR amplification.

Rescue hybridizations were carried out in essentially the same fashion as the subtractions, with the difference that the Dynabeads were saved, combined, washed twice with 1X HB^+ buffer and twice with TE, and finally incubated twice for 10 minutes at 95°C in 100 µl TE to elute the captured tracer cDNA fragments.

Filter Hybridizations

Slot blot filter membranes (Hybond$^+$, Amersham) were generated with 1 µg of each cDNA pool using a vacuum manifold (Life Technologies) according the manufacturer's specifications. Hybridizations were performed at 68°C for 16 hours in 0.5 M Na_2HPO_4 pH 7.2, 1 mM EDTA, 7% SDS, and 1% BSA, and the filters were washed twice at 65°C for 30 minutes with 2X SSC and 0.1% SDS and once at 65°C for 30 minutes with 0.1X SSC and 0.1%

SDS. Radioactively labeled hybridization probes were generated by randomly primed polymerase extension (T7 Quickprime, Pharmacia) using α-^{32}P dCTP (Amersham). Results were quantified by PhosphorImager (Molecular Dynamics) using the ImageQuant program.

Analysis of Library and Individual Clones

After the final reamplification following the rescue hybridization, 1 µg cDNA was restricted (in the case of the EN_7 pool with *Xba* I), purified with Qiagen spin columns, and ligated to 100 ng *Xba* I-restricted, dephosphorylated pBluescript KS^- (Stratagene). This ligation was used to transform XL-2 *E. coli* (Stratagene), which were plated at a density of 300 ampicillin-resistant colonies per 11-cm square agar plate. DNA midipreps (Qiagen) were carried out for 24 randomly picked clones, and the cDNA inserts were isolated by restriction with *Xba* I and agarose gel purification with Qiaex particles. These were then used as hybridization probes as described above with slot blot filters including 1 µg of each original cDNA pool. Those clones exhibiting appropriate differential expression levels were sequenced as double-stranded plasmids by dideoxy chain termination (Autoread, Pharmacia Biotech) using fluoroscein-labeled SK and KS primers (Stratagene, custom synthesis by Pharmacia Biotech) on an automated sequencer (ALF, Pharmacia Biotech) according the manufacturer´s specifications. Sequences were analyzed using GCG software (Genetic Computing Group) and database searches performed using BLAST algorithms (National Center for Biotechnology Information).

RESULTS

Negative Selection: An Improved Subtraction Strategy

One of first prerequisites for being able to perform multiple-parameter selections for the isolation of transcripts that fulfill numerous expression-based criteria was the development of an approach for the highly effective removal, or subtraction, of common transcripts (i.e., transcripts present at approximately equal levels in two or more samples). As mentioned previously, all published subtraction protocols are, however, insufficient in this regard because their inherent inability to generate suitably enriched cDNA pools representing transcripts that are actually differentially regulated. One major reason for this problem is that these methods are based on multiple or even single rounds of subtractive hybridization with driver cDNA or mRNA that has the original abundance profile of its source tissue. Because of the large variation in abundance of individual transcript species within any cell or tissue (from 10^{-1} to 10^{-6} for homogeneous cell populations, and from 10^{-1} to 10^{-12} or possibly even much lower for highly complex tissues such as the brain), the effective C_0t for the rare (or even moderately abundant) common transcript species during the subtractive hybridization is so low that these are not sufficiently removed. In fact, because of the more effective removal of abundant common transcripts (in relative terms), the relative abundance of these rarer common transcripts is increased, leading to high false positive rates during the subsequent screening steps. This problem precludes the use of these subtracted cDNA pools for any additional positive selection strategies, which would increase the background levels to unacceptable levels.

Nevertheless, one recently reported method that employed multiple rounds of subtraction, albeit with limited effectiveness (Wang and Brown, 1991), seemed amenable to the extent that several technical and procedural modifications could rectify these limitations. Briefly, this procedure involves the generation of relatively short, amplifiable cDNA fragments flanked by synthetic adaptors created by the fragmentation of full-length cDNA with four-base restriction restriction endonucleases and the ligation

of adaptors, followed by the PCR amplification of the fragments with an adaptor-specific primer. In this manner, cDNA pools can be created for two samples A and B, where B is then used as the driver cDNA to subtract common fragments from A, the tracer cDNA, thereby enriching for A-specific fragments. This is done through the restriction endonuclease-mediated removal of the adaptor sequences from the driver cDNA, the photobiotinylation of these fragments, and the subsequent hybridization of an excess of driver cDNA with tracer cDNA under conditions appropriate for DNA renaturation. Driver-tracer duplexes are removed after this hybridization by several streptavidin incubations and phenol-chloroform extractions and the remaining fragments from the tracer cDNA are reamplified with the adaptor-specific primer.

This procedure is repeated several times using several different driver cDNAs and the following enrichment scheme: both A and B undergo subtractions in parallel and in complementary fashion, meaning that B driver cDNA is used to enrich for A-specific fragments and A driver cDNA is used to enrich for B-specific fragments. After each round of subtraction, the subtracted cDNA (denoted as A_0, A_1, A_2 . . . or B_0, B_1, B_2 . . .) is reamplified and used in an additional hybridization with another driver cDNA. For example, B_0 is the driver that is hybridized with A_0 to generate A_1, which then undergoes the next subtraction with B_1 as the driver. The rationale for this approach takes into account the relative enrichment of the common cDNAs that were originally less abundant, but because of the depletion of the more abundant cDNAs during the first subtraction, were enriched along with most A-specific cDNAs. Using the original driver B_0 for the second round of subtraction where A_1 where is now the tracer would only result in the continual depletion of abundant common cDNAs from A_1 but in the additional enrichment of the common cDNAs that were originally less abundant. Because this latter class of cDNAs is similarly enriched in B_1, however, it is possible with an appropriate excess of B_1 as the driver cDNA to achieve their effective depletion from the A_1 tracer.

Nevertheless, because this priniciple is not consistently applied throughout the enrichment scheme (two additional rounds of subtraction are carried out with A_0 and B_0 as driver cDNAs), this protocol does not achieve the complete enrichment of differentially expressed transcripts. Originally this efficiency was reported to be approximately 40% (Wang and Brown, 1991),

implying that over half of the cDNA clones of the subtracted library are not actually differentially expressed, but independent results point to somewhat lower efficiencies of about 10% (A. D. Crawford, unpublished results) or even 5% (J. Casal, C. V. Esguerra, and M. Leptin, unpublished results). This and other technical limitations precluded the use of this approach in its original form for the negative selection strategy and necessitated the development of an altered protocol that could overcome these limitations.

Several conceptual and technical changes were therefore introduced which allowed the subtractive enrichment of differentially expressed transcripts to become significantly more efficient, both in terms of time and with regard to the degree of enrichment (see Figure 2). First, driver and tracer cDNAs were generated with different adaptor sequences, which eliminated the contamination caused by the amplification of trace amounts of driver cDNA with unrestricted adaptors remaining in the subtracted sample (Balzer and Bäumlein, 1994). Second, biotinylated driver cDNA was generated not by photobiotinylation, an inefficient, time-consuming procedure, but rather by the amplification with primers that had already been biotinylated as part of the chemical synthesis process and that had been purified by reverse-phase HPLC to insure an effective biotinylation rate of over 99.9% (MWG-Biotech GmbH, personal communication). Third, driver-tracer duplexes were removed after the hybridization step by capturing them onto streptavidinylated paramagnetic latex microspheres (Dynabeads®) rather than through repetitive incubations with free streptavidin and phenol-chloroform extractions, an improvement that proved to be very effective at removing all detectable traces of biotinylated driver cDNA and driver-tracer cDNA hybrids (data not shown). Fourth, the driver-tracer hybridization reactions were carried out in the presence of cetyltrimethylammonium bromide (CTAB), one of several cationic detergents that have been shown to increase the rate of DNA renaturation by more than 1000-fold (Pontius and Berg, 1991). This allowed the cDNA concentrations for these hybridizations to be reduced by 100-fold while still maintaining an improvement in effective C_0t values as compared to the original method, thereby significantly increasing the throughput in terms of the numbers of hybridizations that could be carried out in parallel.

One final change was an alteration of the hybridization strategy so as to insure similar abundance profiles of common cDNA fragments for both

driver and tracer cDNAs during every hybridization. This was done by maintaining the sequential order of driver-tracer matches such that A_0 was paired with B_0, A_1 with B_1, A_2 with B_2, and so on (see Figure 3). The rationale for this approach was to insure the complete removal of all common cDNA fragments from each pool undergoing enrichment for differentially regulated transcripts, regardless of their abundance in the original sample. An initial measure of the overall effectiveness of this new approach can be seen by the complete lack of cross-hybridization between two final cDNA pools after several rounds of subtractive hybridization (see Figures 4 and 5), indicating that both subtractively enriched cDNA pools represent transcripts preferentially expressed in their original source materials, which in this case are embryonic superior colliculus and normal adult superior colliculus, respectively.

Positive Selection: A Rescue Hybridization Strategy

Given the apparent effectiveness of this new subtractive hybridization strategy in generating cDNA pools that were significantly enriched for differentially regulated transcripts, as well as the dramatic improvement in throughput in terms of the number of subtractions that could be carried out in parallel, it was now possible to introduce an additional selection mechanism based on positive selection. Because of the large number of differentially regulated transcripts in any given system, regardless of the experimental parameter involved, additional selection criteria are needed in order to restrict the selection process to the identification those transcripts that may have some significant role in that system. This is especially important when the overall selection and screening process cannot be done systematically, as is the case when cDNA clones that remain after an enrichment procedure are characterized individually by procedures such as Northern analysis, *in situ* hybridization, or sequencing.

For this reason it was attempted to integrate part of the usual post-selection screening process into the original selection procedure. This was achieved by adding an additional selection criterion to the enrichment strategy, in this case by introducing a positive selection step for the specific isolation of cDNA fragments that had been enriched in two separate subtraction procedures. By performing a subtractive hybridization in reverse so that it

functions as a rescue hybridization, it becomes possible to use the driver cDNA to capture identical fragments from the tracer cDNA instead to remove them. This is useful when the fragments common to two separate cDNA pools represent the transcripts of interest, as is the case when two independent subtractions are performed to isolate transcripts that fulfill two different selection criteria.

Boolean Selection of Differentially Regulated Transcripts in the Rat Superior Colliculus

In order to identify transcripts that are associated with the formation and reorganization of neuronal networks and which therefore might be upregulated in the superior colliculus during development and following the loss of afferent inputs, two separate subtractions were carried out in order to isolate (1) transcripts preferentially expressed in embryonic superior colliculus as compared to normal adult colliculus and (2) transcripts preferentially expressed in deafferented adult superior colliculus as compared to normal adult colliculus. After the negative selection process (the subtractive hybridizations) had generated cDNA pools enriched for differentially regulated transcripts fulfilling these two criteria, the positive selection process (the rescue hybridization) then selected those cDNA fragments common to both pools which might therefore represent transcripts fulfilling both criteria. The rescue hybridization was carried out using the deafferentation-specific cDNA pool as the driver and the embryonic-specific cDNA pool as the tracer. These selected fragments were then reamplified and subsequently cloned into a univeral cloning vector to generate the enriched cDNA library EN7.

The effectiveness of the positive selection procedure was measured by an initial expression analysis of 24 clones that were randomly chosen from the first low-density plating of the enriched cDNA library (of about 3,000 clones). For these first 24 clones, the extent of differential expression was determined by using the individual cDNA fragments as probes in slot blot hybridizations with the original cDNA pools (see Figure 6). This cDNA Southern procedure is considerably more effective than Northern analysis in screening large numbers of independent cDNA clones to verify and quantify their differential expression. Although not as direct as Northern analysis

and therefore potentially prone to artifacts such as the nonlinear PCR amplification of individual cDNA fragments, the quantification of transcript expression through the hybridization analysis of cDNA pools has been shown to a reliable indicator of actual mRNA expression levels (Wang and Brown, 1991; Ko et al., 1990; C. V. Esguerra, unpublished results). Based on this analysis, it was determined that all 24 of these randomly picked clones were upregulated in the adult superior colliculus after deafferentation (data not shown), and 14 of these 24 clones were also upregulated in the embryonic superior colliculus.

Although the verification of the differential expression of these transcripts by *in situ* hybridization and Northern analysis is still in progress, these numbers appear to indicate the effectiveness of the rescue hybridization procedure in generating cDNA pools that are highly enriched for differentially regulated transcripts that satisfy multiple expression-based criteria. These results also suggest that the improved subtractive hybridization procedure is effective in its enrichment of cDNAs that represent transcipts that are actually differentially expressed, which has been verified by other applications of this method as well (C. V. Esguerra, unpublished results). An additional analysis of the degree of enrichment provided by both the improved subtractive hybridization strategy and the rescue hybridization procedure is still in progress, yet initial results indicate that the final rescue hybridization dramatically increases the frequency of those cDNA fragments that are present in both final pools (data not shown). This suggests that the positive selection step is quite useful for the rapid selection of genes that fulfull multiple expression criteria.

Sequence Analysis of Selected cDNA Clones

The clones that showed increased levels of expression in both embryonic SC and deafferented adult SC were analyzed further by sequencing and by sequence homology searches in the electronic databases. This sequence analysis showed that of the 14 clones, 13 were unique, with 2 clones consisting of the same sequence (EN7-12 and EN7-13). In addition, it was revealed that of these 13 unique clones, 6 encoded known proteins, 5 encoded proteins with no significant homology to any known protein, 1 was identical to an unknown gene with NGF-induced expression, and 1 clone

(EN7-10) encoded a novel transmembrane protein with significant homology to EGF-like growth factors (see Figure 6).

Further sequence analysis of EN7-10 enabled the identification of the full-length sequence of its human homolog based on unpublished sequence entries in GenBank. This homologous human cDNA displayed approximately 95% homology on the amino acid level to the corresponding region encoded by EN7-10 (data not shown). Interestingly, this same region from an unpublished partial *Xenopus laevis* cDNA sequence also displayed over 89% homology to both rat and human sequences, suggesting a significant conservation in the function of this molecule. Additional sequence comparisons using the full-length sequence of the human cDNA revealed that in addition to its single EGF domain (see Figure 7), it also contains two domains with significant similarity to both follistatin and agrin (see Figure 8). This novel structural configuration appears to define a new class of EGF-like molecules in addition to the EGF family and the neuregulins (see Figure 9). Because of the similarity of EGF domain of EN7-10 to that of betacellulin and because of its neuronal expression pattern (data not shown), this molecule was named neurocellulin.

DISCUSSION

Analysis of Differential Gene Expression by Boolean Selection

Based on these preliminary results, it appears that the Boolean selection method is capable of rapidly identifying differentially expressed transcripts that satisfy multiple selection criteria. Although additional analysis will be necessary to determine the overall efficacy of Boolean selection strategies for this and other applications, especially with selection schemes that involve more than two experimental parameters, it can already be concluded that this type of approach may be useful for the process of gene discovery in many different biological systems.

Nonetheless, several technical limitations have already become apparent that will need to be overcome before this methodology can become universally applicable. With single-parameter selections such as {deafferented adult superior colliculus NOT normal adult superior colliculus} or {D NOT N}, essentially all selected cDNA clones can be shown to be differentially expressed according to these criteria, although this will still need to be verified with larger numbers of clones. (As mentioned previously, NOT is understood to mean "greater than by factor of n", where n is at least 2 when the ratio of driver cDNA to tracer cDNA is 20.) However, this efficiency is significantly decreased by the addition of an additional selection criterion such as {deafferented adult superior colliculus AND embryonic superior colliculus NOT normal adult superior colliculus} or {D AND E NOT N}. In this case the efficiency drops from 100%, where 24 out of 24 clones satisfy the criteria {D NOT N}, to only 58%, where 14 out of 24 clones satisfy the criteria {D AND E NOT N}. This efficiency drops still further by the addition of an additional, third selection criterion (data not shown). This decrease in selection efficiency necessitates an increase in effort during the post-selection analysis steps. Because the actual differential expression of the clones must be verified and quantified by blot-based methods that use the individual clones as hybridization probes, any significant loss in efficiency in the initial selection process increases the redundant effort in this latter phase of the screen. The most likely reason for this decrease in

enrichment fidelity is the use of double-stranded cDNA fragments as driver and tracer populations during the subtractive hybridizations, which cause trace amounts of common cDNA fragments to remain in the subtracted cDNA pools when they should have been removed entirely. This does not affect the efficiency of the subtractive hybridization as much as it does that of the rescue hybridization, because these trace amounts of common cDNA fragments are only enriched to the point of detection during the rescue hybridization process. However, an additional effect of the use of double-stranded cDNA fragments is the need to perform multiple rounds of subtraction and reamplification in order to create cDNA pools that are suitably enriched for differentially expressed transcripts and that can be used for additional rescue hybridizations. These multiple rounds of PCR amplification result in an inherent representational bias because of the dependence of the amplification process on the length of the cDNA fragments. Longer fragments are therefore reduced in abundance to the point of nondetection.

Another limitation of the methodology as it now stands is the sequence-dependent fragmentation and adaptor ligation process, which also inherently leads to the misrepresentation of certain transcripts in the original cDNA pools. Although most evidence to date points to the fidelity of this process in creating representative cDNA populations that reflect the original transcript abundance profile of the source tissue (Wang and Brown, 1991; Ko et al., 1990; C. V. Esguerra, unpublished results), it is obvious that any process that depends on the presence of distinct sequence motifs (in this two different four-base restriction sites) will result in the inefficient amplification of some transcripts because of their length and their lack of these sequence motifs. This will also be true for significant portions of many other transcripts that do contain these sites, but not at regularly spaced intervals.

The solution for both of these problems would be to substantially alter the standard cDNA synthesis procedure by using biotinylated oligonucleotide adaptors linked to random primers for the first-strand synthesis. By carrying out this reaction with an appropriate excess of these random primer-adaptors, the length of the resulting single-stranded cDNA fragments would be consistantly limited so that all fragments could be amplified with approximately equal efficiency, given the additional assistance of

thermostable polymerase mixtures with significantly higher processivities (Barnes, 1994). As the generation of amplifiable fragments from all transcripts using this approach would be an entirely random, sequence-independent process, there would no longer be any significant chance for length-dependent bias to occur during the PCR amplification steps. In addition, by adding an independent, second primer to the distal end of the first-strand cDNA by RNA ligase, thereby allowing the amplification of the fragments with two separate primers, it would become possible to generate single-stranded fragments of opposite complementarity for the different hybridization reactions. For example, by amplifying the driver cDNA fragments with one biotin-labeled primer and one digoxygenin-labeled primer so that the sense strand is labeled with digoxygenin, it would be possible to completely remove the biotinylated antisense strand with streptavidinylated beads, thereby generating digoxygenin-labeled, single-stranded driver cDNA. The same strategy could also be used to generated unlabeled, single-stranded tracer cDNA consisting of the complementary antisense strand. This would enable both the subtractive hybridizations and the rescue hybridizations to be carried out with complementary single-stranded driver and tracer cDNAs, thereby greatly increasing their efficiency and reducing the number of subtractive hybridizations needed for the sufficient enrichment of differentially expressed transcripts.

Possible Physiological Relevance of Identified Transcripts

Although additional analysis by *in situ* hybridization will still be necessary to verify the extent of differential expression and to provide information on the cellular localization of the transcripts identified so far, it has already been possible with this preliminary screen to gain new insight into some of the molecular and cellular activities in the superior colliculus during periods associated with the formation and reorganization of synaptic connections. Some of the identified transcripts encode proteins associated with post-transcriptional protein synthesis and modification events, including ribophorin (an essential subunit of the protein glycosylation complex oligosaccharyltransferase), calnexin (a membrane-associated molecular chaperone of the endoplasmic reticulum), and translation initiation factor, which together suggest enhanced levels of protein production and secretion. Other transcipts encode proteins that appear to indicate increased neuronal

activity, such as rSec6 (part of the protein complex involved in the docking and fusion of synaptic vesicles to the plasma membrane) and Fyn (a nonreceptor tyrosine kinase found at high concentrations in axonal growth cones and that is known be involved in synaptic plasticity). Two of the more interesting transcripts include Fas-associated protein factor (a recently-identified molecule potentially involved in Fas-mediated signal transduction during apoptotic cell death), and a novel member of the EGF superfamily of transmembrane growth factors (which presently includes the EGF/TGFα family and the neuregulins). Further analysis of these molecules may provide additional information about the signaling mechanisms and cell-cell interactions at work during the establishment and reorganisation of neuronal connectivities in the superior colliculus.

Clone 10: Novel EGF-like gene with follistatin-like domains

The EGF family of membrane-anchored growth factors includes epidermal growth factor (EGF), transforming growth factor α (TGFα), heparin-binding EGF-like growth factor (HB-EGF), amphiregulin, betacellulin, and teratocarcinoma-derived growth factor (TDGF, also known as cripto). These proteins all contain one EGF domain and with the exception of TDGF are known to be ligands for the EGF (erbB) receptor (EGFR). The neuregulins, an additional family of EGF-like proteins that encompasses numerous splice variants of a single transcript, alternatively called glial growth factors (GGFs), heregulins, neu differentiation factor (NDF) and acetylcholine receptor inducing activity (ARIA), are characterized by an immunoglobulin domain in addition to the single EGF domain and are known to interact with the EGFR-like receptors erbB2 (neu), erbB3, and erbB4 (Carraway and Burden, 1995). The protein encoded by clone 10 appears to define an additional, third class of EGF-like molecules that is characterized by two follistatin domains in place of the immunoglobulin domain present in the neuregulins. Because of the predomininant expression of this transcript within the central and peripheral nervous systems (data not shown) and the similarity of its EGF domain to that of betacellulin, the protein it encodes was named neurocellulin.

It is not yet clear whether the EGF domain of neurocellulin is capable of binding the EGF receptor or any of the three EGFR-like receptors. Both

neurocellulin and TDGF differ slightly from the consensus sequence of the remaining EGF superfamily members (including the neuregulins), which is $CX_7CX_2GX_{1-2}CX_{10-13}CXCX_3Y/FX_1GX_1RC$. This structure, which folds into an open hairpin-loop configuration with three disulfide bonds, contains an arginine immediately proximal to the final cysteine that has been shown to be necessary for EGF receptor binding and activation by TGFα (Defeo-Jones et al., 1989). This arginine is likely to be similarly important for the interactions of other superfamily members with their respective receptors and is substituted by histidine in neurocellulin, raising the question of whether neurocellulin interacts with other, still unknown EGFR-like receptors. In the EGF domain of TDGF this single arginine is substituted by asparagine, yet TDGF was shown to possess mitogenic activity, indicating that these conservative substitutions either do not interfere with binding to the known members of the EGF receptor family, or that they specify interactions with other receptors.

Growth factors of the EGF superfamily have widely divergent actions on numerous cell types. TGFα was originally found in the culture medium of various retrovirus-transformed fibroblasts and was identified by its ability to stimulate the proliferation of normal fibroblasts (DeLarco and Todaro, 1978), an effect later shown also to be dependent on the action of TGFβ (Anzano et al., 1983; Massague, 1983). Since then it has been found to be expressed in numerous tissues, including various regions of the nervous system (Kaser et al., 1992; Wilcox and Derynck, 1988), where it was recently shown to stimulate the survival and neurite outgrowth of mesancephalic dopaminergic neurons (Alexi and Hefti, 1993). EGF is also widely expressed throughout the nervous system (Birecree et al., 1991) as is the EGF receptor (Seroogy et al., 1995; Birecree et al., 1991) which has also been found to be upregulated in Schwann cells after lesions in the sciatic nerve (Toma et al., 1992), suggesting its potential role in mediating events related to neuronal regeneration.

As the prototype molecule of the EGF/TGFa class of transmembrane growth factors, TGFα and the molecular mechanisms of its actions have been studied in considerable detail. Although initially assumed to be active only in soluble form after release from the cell membrane by proteolytic cleavage (Pandiella et al., 1991), it was demonstrated that TGFα is also able to stimulate the growth of neighboring cells while still anchored to the cell

membrane (Wong et al., 1989; Brachmann et al., 1989). This mode of action has been termed "juxtacrine" activity. Another intruiguing finding was that TGFα may itself act to transduce signals from the cell surface to the cytoplasm in a fashion similar to that of classical transmembrane receptors (Shum et al., 1994). This is suggested by the association of the transmembrane form of TGFα with a protein kinase complex, raising the possibility that an interaction between TGFα and the EGF/TGFα receptor leads to reciprocal signaling events. Given the high degree of homology (over 98% on the amino acid level) within the transmembrane and cytoplasmic domains of neurocellulin between its mammalian and amphibian homologs, it is possible that this part of the protein is involved in similar protein-protein interactions during the initation of specific signal transduction events. An additional or perhaps alternative explanation for this significant conservation of sequence homology may be that neurocellulin is targeted to distinct locations within the plasma membrane by virtue of the interactions of these regions with components of the cytoskeleton or other membrane-associated proteins.

In addition to the single EGF domain adjacent to the transmembrane region, neurocellulin contains two follistatin-like regions in the N-terminal part of the protein. These follistatin-like domains might be also be indirectly involved in activating growth factor receptors. It has been speculated that the follistatin-like regions of agrin and agrin-related proteins serve to bind and localize molecules of the transforming growth factor β and platelet-derived growth factor families, thereby creating localized matrix-associated concentrations of them (Patthy and Nikolics, 1993). There are numerous lines of evidence to support this hypothesis. First, it is well known that follistatin binds activin with high affinity (Nakamura et al., 1990), thereby blocking its action on gonadotrophs and inhibiting the release of follicle-stimulating hormone (FSH). Follistatin is also known to bind inhibin, another member of the TGFβ superfamily that shares one of two subunits with activin (Shimonaka et al., 1991). Furthermore, osteonectin/SPARC, another protein containing the follistatin module, has been shown to bind the platelet-derived growth factors (PDGF) AB and BB, thereby preventing these from binding the PDGF receptor (Lane et al., 1991; Raines et al., 1992). This raises the possibility that the follistatin-like regions in neurocellulin, agrin, and SC1, an osteonectin-related glycoprotein found in brain extracellular matrix (Johnston et al., 1990), are capable of binding

members of the PDGF family as well. Given the known roles of agrin and SC1 in synapse formation (Gautam et al., 1996; Johnston et al., 1990) and the neuronal expression of PDGF (Yeh et al., 1990; Sasahara et al., 1990), it may be the case that the modulation of PDGF activity by agrin or SC1 is important for the process of synaptogenesis.

It remains to be seen which growth factors are bound by neurocellulin and how these signaling activities are modulated by the interaction with neurocellulin. Enhanced expression of members of the transforming growth factor β superfamily has been documented following lesions and/or the loss of synaptic input (Nichols et al., 1991; Finch et al., 1993). It has been suggested that the increase in TGFβ expression may serve in part to induce the resynthesis of extracellular matrix that is lost as the result of degeneration of afferent fibers (Pasinetti et al., 1993). Some results point to the role of TGFβ-like molecules in inducing apoptotic cell death (de Luca et al., 1996), whereas in other experimental contexts it appears that these factors are able to promote survival of neurons (Schubert et al., 1990; Krieglstein et al, 1995) and to protect them against excitatory amino acid-mediated neurotoxicity by stabilizing calcium ion homeostasis and inducing expression of the Blc2 oncoprotein, an inhibitor of apoptotic cell death (Prehn et al., 1994). Clearly the scope of actions of these diverse molecules depends on the cellular context in which their action takes place, so it cannot be predicted what growth factor-mediated effects neurocellulin is able to modulate.

Additional intriguing results that point to the potential interaction between neurocellulin and TGFβ-like molecules include the potent inhibition by TGFβ of hepatocyte growth factor (HGF) secretion in fibroblasts (Gohda et al., 1992) and recent *in situ* hybridization data that indicate partially overlapping expression patterns of neurocellulin with HGF throughout the rat brain (S. Isenmann, personal communication). One possible explanation for these findings could be that the follistatin domains in neurocellulin act in an inhibitory fashion to block the action of TGFβ and similar molecules, thereby enabling HGF secretion to take place. It should be noted, however, that the observed interaction between HGF and TGFβ was seen with human fibroblasts and that similar results cannot be assumed for the various neurons that express neurocellulin and HGF. Further analysis will be necessary to correlate the neuronal expression patterns of the various

members of the TGFβ superfamily with the expression of neurocellulin and HGF. Any functional interactions that may occur between these different classes of molecules will then have to be determined in an *in vitro* environment.

As mentioned previously, it will be necessary to determine the binding affinities of the two follistatin domains of neurocellulin for all TGFβs, growth/differentiation factors (GDFs), bone morphogenic proteins (BMPs), activins and inhibins, and glial cell line-derived neurotrophic factor (GDNF), as well as for the PDGFs and the neurotrophins. It is quite possible that the two follistatin domains of neurocellulin bind several of these different growth factors with widely divergent affinities, thereby giving rise to quite different forms of molecular interaction, such as complete inactivation in the case of higher affinities and limited binding and release (or presentation to receptors) in the case of lower affinities. The proximity of the two follistatin-like regions in neurocellulin suggests that they may indeed be involved in the presentation of these growth factors to their receptors, given the need for most cell surface receptors to dimerize in order for the signal transduction process to be initiated (Heldin, 1995). In the case of receptors with either serine/theonine or tyrosine kinase activity, the proximity or local concentration of ligands is a critical factor in determining whether receptor dimerization occurs. Factors that bind these ligands and thereby create localized matrix-bound concentrations of them, such as heparan sulphate proteoglycans in the case of TGFβs or fibroblast growth factor (Lopez-Casillas et al., 1993; Spivak-Kroizman et al., 1994; Ruoslahti and Yamaguchi, 1991), serve in part to promote receptor dimerization and the subsequent initiation of signal transduction. It remains to be seen whether the two follistatin-like regions in neurocellulin are capable of simultaneously binding different growth factors, leading to the heterodimerization of the corresponding receptors and perhaps to the initiation of specific signal transduction pathways.

Another interesting aspect of follistatin-like protein domains is their own ability to bind to heparan sulphate proteoglycans (Nakamura et al., 1991). This property may enable the indirect attachment of follistatin-bound growth factors to components within the extracellular matrix, thereby providing receptors on neighboring cells with localized reservoirs of ligands. It has been shown that the binding of follistatin to heparan sulphate

does not interfere with those molecular interactions that are necessary for binding growth factors, despite the fact that these different interactions are mediated by the same protein region (Sugino et al., 1993; Inouye et al., 1992). It is not unlikely that high-affinity associations of the two follistatin-like regions in neurocellulin with glycosaminoglycans allow for the localization of the EGF domain of neurocellulin to the extracellular matrix, enabling its accumulation within extracellular spaces in the vicinity of neurocellulin-expressing neurons. An intruiging possibility is that neighboring neurons and glia secrete specific proteases that cleave the EGF domain from matrix-bound neurocellulin, thereby allowing the activation of its receptor on these cells. Similar mechanisms have been shown to be involved in the release of ARIA, one of several neuregulin variants responsible for the induction of acetylcholine receptor expression in postsynaptic neurons, from the extracellular matrix (Loeb and Fischbach, 1995). It has previously been demonstrated that the EGF-like domain of neuregulin is sufficient for receptor binding and activation (Holmes et al., 1992; Wen et al., 1994), leaving open the possibility that this is the case for neurocellulin as well.

Aside from the potential molecular mechanisms of action of neurocellulin, there remains the question of the scope and nature of this action. Expression of neurocellulin has been observed in embryonic retinal ganglion cells throughout the phase of axonal outgrowth and is undetectable in adult retinal ganglion cells (S. Isenmann, personal communication), but is again observed during the axonal regeneration of adult retinal ganglion cells into peripheral nerve grafts after transection of the optic nerve (C. V. Esguerra, personal communication). It may be the case that neurocellulin either acts in an autocrine fashion to promote neurite outgrowth or that it acts in the same fashion as glial growth factor, one of several neuregulin variants, in promoting the survival of glial cells or perhaps even of target neurons and in stimulating their secretion of factors needed for trophic support and neurite outgrowth (Mahanthappa et al., 1996; Trachtenberg et al., 1996). Functional studies, such as co-culture experiments with transfected cell lines and primary neurons and/or glia, will be needed to address these questions in more detail. An additional approach to determine the *in vitro* effects of neurocellulin on various cell types could be done at the molecular level, through the identification of neurocellulin-incuced transcripts.

Clone 22: Fas-associated protein factor, FAF1

Fas, a member of the tumor necrosis factor (TNF) receptor family, is known to trigger cell death by apoptosis when activated by Fas ligand (FasL) binding or by anti-Fas antibody crosslinking (Nagata, 1994). Fas-mediated apoptotic cell death has been shown to have an important function in the development and function of the immune system (Nagata, 1994; Lowin et al., 1994), but the molecular mechanisms by which Fas transduces apoptotic signals from the cell surface to the cytoplasm are not yet well understood. Fas-associated protein factor was recently identified by a yeast two-hybrid screen for proteins that specifically interact with the wild-type cytoplasmic domain of Fas, but not with the altered form of Fas found in *lpr^{cg}* (lymphoproliferation) mutant mice (Chu et al., 1995). The *lpr^{cg}* mutation results in a single base change within the death domain, replacing isoleucine with asparagine and thereby completely abolishing the signal-transducing properties of Fas (Watanabe-Fukunaga et al., 1992). This and other naturally occuring mouse mutants such as *lpr* and *gld* (generalized lymphoproliferation disease), all characterized by the loss or inactivation of Fas or FasL, develop lymphadenopathy and systemic lupus erythematosus-like autoimmune disease, implicating Fas-mediated apoptosis in the deletion of autoreactive lymphocytes. Gene targeting studies have also shown the Fas pathway to be involved in the elimination of tumor cells and virus-infected cells mediated by cytotoxic T lymphocytes and natural killer cells (Lowin et al., 1994).

Subsequent functional analysis of Fas-associated protein factor by transient overexpression in Fas-expressing COS cells revealed that FAF1 potentiates Fas-mediated cell death induced by antibody-crosslinking (Chu et al., 1995). This result, in combination with the verification that FAF1 specifically interacts with the wild-type death domain of Fas but not with the mutant form deficient in signal transduction, suggests that FAF1 is part of the Fas signal transduction pathway and that it acts downstream of Fas in the induction of apoptosis. It is not yet clear, however, whether this interaction is ligand dependent, i.e. whether FAF1 preferentially binds dimeric as opposed to monomeric Fas. Other proteins recently shown to specifically associate with the wild-type cytoplasmic domain of Fas include FADD/MORT1 and RIP, yet these contain death domains homologous to that of Fas and are capable of inducing apoptotic cell death on their own, which

is not the case for FAF1 (Chinnaiyan et al., 1995; Boldin et al., 1995; Stanger et al., 1995). It remains to be seen whether FAF1 functions downstream of multiple death domain-containing proteins during apoptotic signal transduction or whether it associates exclusively with Fas.

Another interesting question is whether FAF1 is also involved in non-apoptotic signal transduction mediated by Fas or other death domain-containing receptors. In addition to Fas, the TNF receptor family includes the TNF receptors (P55 and P75), CD40, OX40, CD27, CD30, 4-1BB, and the low affinity NGF receptor. These receptors are capable of mediating very divergent cellular reponses, including not only apoptosis but also cell proliferation and differentiation (Baker and Reddy, 1996; Smith et al., 1994). Fas has also been shown to mediate cell proliferation in certain cell types (Owen-Schaub et al., 1993; Alderson et al., 1993), yet the signal transduction pathway for such non-apoptotic responses is poorly understood as well, leaving open the possibility of any involvement of FAF1 in these processes. An intriguing connection between the molecular mechanisms involved in apoptosis and those involved in axonal guidance was recently made with the observation that the death domain motif is also present in proteins such as unc-5 and unc-44, an ankyrin-related protein (Hofmann and Tschopp, 1995). Both proteins are necessary for the guidance of pioneer axons and migrating cells in *Caenorhabditis elegans*, raising the possibility that the death domain mediates protein-protein interactions involved in cytoskeletal rearrangement, which is also an early event of apoptosis.

The role of receptor-mediated apoptosis in the development and function of the nervous system has not yet been specifically examined (Rabizadeh and Bredesen, 1994). Nonetheless, cell death has long been recognized as an important phenomenon during neuronal differentiation and in the establishment of neuronal connectivities (Rubin et al., 1994; Gould and McEwen, 1993). Although most work on neuronal death has focused on the interactions between neurons and their targets, there are now several lines of evidence underscoring the importance of afferent inputs in regulating neuronal survival as well, both during development and in the mature nervous system. For example, in the ciliary ganglion (CG) of the chick embryo, which is made up of peripheral autonomic neurons that project to the eye musculature and that receive afferent inputs exclusively from the accessory oculomotor nucleus (AON), as many additional neurons die after

the removal of the input, the AON, as do after the removal of the target, the optic vesicle (Furber et al., 1987). Whereas usually 50% of CG neurons die by the end of the normal cell death period (E14 to E15), the early (E4) surgical removal of either AON or optic vesicle results in the degeneration and death of up to 90% of CG neurons. The cell death resulting from the loss of preganglionic input is not due to the failure to CG neurons to innervate their target, as this was shown to still occur. Moreover, the combined loss of both input and target results in the death of virtually all remaining neurons, indicating the ability of the afferent inputs to sustain the survival of those CG neurons deprived only of their targets. Comparable results have been obtained for the parabigeminal nucleus of the rat (Linden and Renteria, 1988; Linden and Pinon, 1987), indicating the role of afferent supply in determining the survival of neurons.

In summary, the observed increases in FAF1 expression may be an indication of the occurance of apoptotic cell death, mediated by Fas or other death domain-containing receptors, in the superior colliculus during embryonic development as well as after the loss of afferent input. It might be possible that this increased expression of FAF1 and perhaps also of other components of the cell death pathway occurs during periods of synapse formation in order to prime postsynaptic neurons for apoptosis, should the required synaptic connections not be established. Additional investigations of cell death in the superior colliculus are warranted, especially an analysis for the presence of apoptosis-associated markers, including DNA fragmentation as visualized by TUNEL and the expression of FasL and other functionally related molecules such as tumor necrosis factors.

Clone 4: Fyn tyrosine kinase

Fyn is one of several well-characterized nonreceptor tyrosine kinases (including also Src, Yes, and Abl) that participate in numerous signal transduction pathways in neurons and other cell types. It was originally identified on the basis of its sequence homology to Src (Kawakami et al., 1988) and has subsequently been shown to be developmentally regulated in different neuronal cell types (Ingraham et al., 1992; Maness, 1992) as well as in various cells of the haematopoietic lineage (Toyoshima et al., 1992). Its primary function appears to be the mediation of signal transduction

events that are initiated by the activation of cell-surface receptors, including the EGF receptor (Margolis et al., 1992), the receptor for macrophage colony-stimulating factor (Courtneidge et al., 1993) and Sek, of the Eph receptor family (Ellis et al., 1996). Its interaction with the PDGF receptor (Twamley et al., 1992) suggests that Fyn might be involved in the regulation of oligodendrocyte proliferation. Recent analysis of Fyn expression has shown it to be abundant in neuronal growth cones, implicating it in the regulation of neurite outgrowth and perhaps axon guidance (Bixby and Jhabvala, 1993).

There are several lines of evidence which suggest that Fyn is also involved the regulation of synaptic plasticity. First, the protein was found to be one of the major components of the post-synaptic density (PSD) fraction prepared from mouse forebrain, and was implicated in the tyrosine phosphorylation of two other proteins in this fraction, the epsilon subunits of the NMDA receptor (Suzuki et al., 1995). Second, the analysis of mice homozygous for the targeted deletion of the *fyn* gene demonstrated that Fyn appears to be required for the NMDA-receptor mediated induction of long-term potentiation, an important form of activity-dependent synaptic plasticity (Grant et al., 1992). Mutant mice display an impairment in spatial learning, as well several behavioral deficits including the disturbance of instinctive suckling behaviour in newborns (Yagi et al., 1993).

In addition, there have been several recent results which indicate that Fyn might be involved in signal transduction through glycosyl-phosphatidylinositol-anchored cell-surface proteins. In the immune system, it could be demonstrated that the induction of lymphocyte proliferation through the activation of the membrane-associated decay-accelerating factor (DAF, also known as CD55) is dependent upon the interaction of Fyn with its GPI anchor (Shenoy-Scaria et al., 1992). In the nervous system, the immunoglobulin superfamily member F3, a plurifunctional GPI-anchored molecule predominantly expressed on axons and a known modulator of neurite outgrowth, was found to associate with both Fyn and L1, another GPI-anchored cell adhesion molecule, suggesting that this complex may be involved in the mediation of F3 signaling events (Olive et al., 1995). Fyn was also found to interact with the GPI-linked neuronal cell adhesion molecule contactin/F11 (Zisch et al., 1995), indicating that the signaling pathway through which GPI-anchored cell-surface molecules respond to

external stimuli such as receptor binding involves Fyn and possible other member of the Src family of nonreceptor tyrosine kinases.

Given the ability for GPI-linked proteins to be transferred between cells (Ilangumaran et al., 1996), this indicates the possibility of a novel form of intercellular communication involving the transfer of GPI-anchored signal-transducing molecules to cells that may not actually express these proteins. This acquisition of signal transduction capacity is qualitatively different from other forms of ligand-induced signaling because its potential for more rapidly initiated, relatively stable, and highly localized alterations in cellular phenotype. One possible example of this might be the numerous interactions between axonal growth cones and other neurons or their processes, which could involve the transfer of GPI-anchored molecules in either direction, giving the growth cone (or its counterpart) an additional mechanism by which to acquire and process positional information. Considering the high concentrations of Fyn in embryonic growth cones, it would not be surprising if, in addition to its modulation of receptor-mediated signal transduction, this tyrosine kinase is responsible for signaling events mediated by GPI-linked proteins transferred from other cells.

In summary, it is difficult to determine what cellular events and functions might be indicated by the observed increases in Fyn expression in the superior colliculus during during phases of synapse formation and reorganization. One of the more likely explanations could be the increased degrees of axonal outgrowth in the embryo and of axonal sprouting in the lesioned adult, yet this would still have to be determined by an immunohistochemical analysis of Fyn protein expression.

Clone 2: rSec6

Several proteins involved in the life cycle of the synaptic vesicle have been identified as homologues of yeast proteins responsible for vesicle trafficking between Golgi bodies and the plasma membrane. rSec6 was recently identified by homology cloning as the mammalian equivalent of Sec6p (Ting et al., 1995), one of several components of the soluble 19.5S complex in yeast that interacts with the vesicle-associated GTP-binding protein Sec4p (Potenza et al., 1992). There is increasing evidence that this particle is

involved in the formation of the 7S complex (consisting of VAMP, synaptotagmin, syntaxin, and SNAP-25) and the 20S complex (consisting of VAMP, syntaxin, SNAP-25, a-SNAP, and NSF) which were recently identified as putative synaptic vesicle docking/fusion particles in the presynaptic terminal (Söllner et al., 1993a, 1993b).

From genetic studies with the yeast homologs of these proteins, it has become apparent that rSec6 and rSec8 may associate with Rab3a, the mammalian homolog of Sec4p, to mediate the delivery and docking of the synaptic vesicle to the plasma membrane (Aalto et al., 1993; Brennwald and Novick, 1993). The 7S complex is formed during the docking step by the association of the two vesicle-associated proteins VAMP and synaptotagmin with the plasma membrane-associated proteins syntaxin and SNAP-25. Neurotransmitter release caused by the calcium-induced fusion of synaptic vesicles with the plasma membrane is preceded by the formation and dissociation of the 20S complex. Although the exact molecular functions of rSec6 within this cascade of events are only beginning to be elucidated, it is clear that this factor plays an important role in synaptic vesicle-mediated neurotransmitter release.

Enhanced expression of proteins such as rSec6 that are involved in the life cycle of the synaptic vesicle may therefore be an indicator of increased neurotransmitter release, both in neurons of the embryonic colliculus as well as of the adult colliculus after the loss of afferent inputs. This increase in synaptic activity may well be associated with the target-associated synaptic plasticity and reorganization that is known to occur in the embryonic brain during the phase of innervation and in the adult brain as the result of deafferentation. The loss of afferent input to the superior colliculus has been shown to result in the elimination of most affected synapses, subsequently leading to significant synaptic reorganization and reinnervation that may enable some deafferented neurons to be resupplied with alternate synaptic inputs (Lund and Lund, 1971a, 1971b). This reafferentation might prevent these neurons from undergoing cell death, which usually results from the loss of afferent inputs (Furber et al., 1987). Further studies will need to be untertaken in order to examine more closely these potential relationships between neurotransmission, synaptic reorganization, and cell death in the superior colliculus during development and following deafferentation. One possible approach would be examine the effects of a

chemically-induced blockage of synaptic transmission on synapse formation through electrophysiological recording and dye-labelling, and to combine this with an analysis for markers of programmed cell death.

It should be mentioned that an alternative explanation for the increase in neurotransmitter release after deafferentation might also simply be the spurious activity of existing neuronal circuits after the loss of their afferent inputs, or of the additional activity of aberrant connections formed as the result of the synaptic reorganization and reinnervation mentioned above. There has been much speculation as to the potential causes of deafferentation pain in human patients (Davar and Maciewicz, 1989; Sweet, 1988), and at least one possible contributing factor to this phenomenon might be the formation and spurious activity of such abberant synapses that lead to the disfunctional activity of various sensory neurons.

Clone 21: NGF-induced EST

Recently, an expressed sequence tag analysis of differential gene expression in PC-12 cells was carried out for the identification of transcripts whose expression was affected by treatment with nerve growth factor (Lee et al., 1995). The rat pheochromocytoma cell line PC-12 acquires neuron-like characteristics when grown in the presence of NGF, including the elaboration of neurite-like processes, an arrest of cell division, and the acquisition of electrical excitability (Greene and Tischler, 1982; Halegoua et al., 1991), and can therefore be considered an appropriate model system for the analysis of certain aspects of neuronal differentiation. By obtaining expressed sequence tags of over 6,000 randomly picked clones from primary cDNA libraries of both untreated and NGF-treated PC-12 cells, it was possible to generate expression profiles of these cells that could be compared to determine the changes induced by NGF. Over 600 differentially regulated transcripts were identified by this approach, including many transcripts encoding proteins involved in cellular pathways not known to be regulated by NGF, as well as numerous transcripts that exhibited no significant sequence homologies to known genes. Clone 21 is identical to one of these novel ESTs, raising the possibility that this transcript is one that is associated with neuronal growth and differentiation.

Clone 23: Ribophorin

Ribophorin was originally identified as an abundant integral membrane glycoprotein restricted to the rough endoplasmic reticulum (Kreibich et al., 1978) and is thought to be one of several components regulating the cotranslational translocation of nascent polypeptides into the RER lumen (Yu et al., 1990). Recent evidence also indicates that ribophorin is also part the oligosaccharyltransferase complex, an abundant N-linked glycosylation activity that is also associated with the RER and which is responsible for the transfer for the transfer of high mannose oligosaccharides to asparagine acceptor sites within these nascent, unfolded polypeptides during the translocation process (Kelleher et al., 1992). Enhanced expression levels of this protein in the superior colliculus during development and following deafferentation indicate increased levels in the synthesis of membrane-associated and secreted N-linked glycoproteins.

Clone 7: Calnexin

Calnexin is also an integral membrane protein of the endoplasmic reticulum and has been shown to associate with numerous nascent membrane and secretory proteins via their carbohydrate moieties (Bergeron et al., 1994). This interaction takes place during the translocation process and constitutes activities previously thought to be restricted to soluble chaperones of the heat shock protein families. In addition to its ability to capture unfolded glycoproteins through its lectin-like affinity for early intermediates in the processing of N-linked oligosaccharides, calnexin has also been demonstrated to bind misfolded and incompletely folded proteins, indicating that it may function as part of the quality control apparatus that prevents proteins from continuing along the secretory pathway unless they have acquired the proper structural conformation. Upregulation of calnexin may therefore also indicate higher rates of synthesis of secretory and membrane-bound glycoproteins. These proteins may include molecules involved in mediating cell-cell interactions during the establishment of neuronal connectivities in the superior colliculus, such as trophic factors and their corresponding receptors, extracellular matrix proteins, and cell adhesion molecules. Further analysis of the cDNA clones in the EN_7 library may

provide additional clues as to which additional molecular interactions are associated with this process.

Clone 12: Translation initiation factor

Although this and most other studies on the regulation of gene expression have focused on transcriptional control, there are clearly additional mechanisms that allow cells to regulate the synthesis and activity of proteins. One of these mechanisms is the regulation of translation initiation (in addition to the control of mRNA conformation and stability, as well as the regulation of protein folding, post-translational modification, and secretion). One of the major criticisms of transcript-based methods for the analysis of differential gene expression has been that changes at these other levels of control cannot be monitored by examining changes at the mRNA level. Nevertheless, at the very least it may be possible to determine whether non-transcriptional mechanisms are involved in certain cellular responses by examining the expression of transcripts that encode essential components of these other control systems. Translation initiation factors are involved in the regulation of eukaryotic translation initiation (Lindahl and Hinnebusch, 1992) and are themselves subject to various levels of control, including phosphorylation by signal transduction pathways activated by extracellular factors (Frederickson and Sonenberg, 1992) as well as the post-translational conversion of specific amino acids (Park et al., 1993). Because initiation is the rate-limiting step in protein translation, the regulation in the rate of initiation is an essential factor in determining the overall level of cellular activity. Increased levels in expression of an initiation factor, therefore, correlate well with the observed changes in ribophorin and calnexin transcription and suggest an elevated degree of protein synthesis and cellular activity in the superior colliculus during the formation and reorganization of synaptic connections.

An Approach for the Systematic Analysis of Gene Expression

In addition to the possible physiological significance of the transcripts that have been identified in this initial screen, several conclusions can already be drawn as to the use of this type of strategy for the analysis of differential

gene expression. There several inherent limitations to this type of approach, including above all the number of experimental parameters that can be analyzed and the amount of gene expression data that can be generated for any given system. Because of the need to analyze the expression of the different transcripts on an individual basis after the initial selection process, the sampling depth becomes limited to such an extent that saturation cannot be reached in any individual screen, thereby strictly limiting the amount of new expression information that is generated. Therefore, although Boolean selection may represent an improvement in terms of differential cloning as far as the identification of individual transcripts is concerned, it is not an answer to the need for an approach that is capable of quantitatively analyzing the expression of many transcripts in parallel.

This type of systematic analysis is, however, the only possibility for reaching an understanding of the entire spectrum of molecular mechanisms that are the basis of all cellular activities. If individual cells are regarded as a complex, dynamic integrators of multiple input signals - regardless of the physical nature of these signals - then it is clear that many cellular components can either directly or indirectly influence cellular responses to these inputs. It would therefore be of considerable interest to be able to quantitatively analyze the *complete* genetic activity of *individual* cells in order to determine their overall biochemical composition, and to be able to compare this data from many different cells. It is only in this manner that the analysis of differential gene expression would be able to significantly contribute to an understanding of cellular phenomena.

Traditional methods for the analysis of gene expression are, however, not suitable for this type of approach. All involve the use of simple probes representing single genes in order to determine the presence of these probes in complex targets using procedures that are relatively labor-intensive (such as quantitative RT-PCR, Northern blot hybridizations, RNase protection assays, and *in situ* hybridizations). Although whole-mount *in situ* hybridizations are being optimized for high-throughput approaches in both zebrafish and mouse embryos (W. Driever and B. Hermann, personal communication), this is not an option for the *comprehensive* analysis of gene expression for the simple reason that the number of possible experimental parameters vastly outweighs the estimated number of genes in the average vertebrate genome. For example, each new experimental parameter not

included in the original collection of *in situ* hybridization targets (new mutations, transgenes, pharmacological agents, combinations thereof, as well as different developmental or experimental timepoints) would necessitate tens of thousands of additional *in situ* hybridizations with the same probes used in the original analysis. Additional limitations of whole-mount *in situ* hybridization include the lack of cellular resolution, quantitive capacity, and access to later developmental stages (including adult), which could of course be offered by the analysis of tissue sections, yet only at the cost of significant decreases in efficiency.

Similar arguments apply to approaches involving the generation of large numbers of expressed sequence tags for the quantification and comparison of gene expression (Adams et al., 1995; Okubo et al., 1992). In addition to being prohibitively expensive, EST analysis is limited to comparisons of very few samples and only of relatively abundant transcripts within these samples. Other technologies for the identification of differentially expressed genes, such as the Boolean selection scheme described here, as well as classical subtractive cloning approaches, differential display, and related methods, are also limited to the examination of very few samples, while novel genes must be characterized further individually. One final disadvantage common to EST analysis, most differential cloning methods, and blot-based RNA quantitation is that the requirement for relatively large amounts of source material severely restricts the cellular resolution of the expression data that is obtained.

An alternative approach would be the high-throughput, high-resolution, and comprehensive analysis of gene expression by combining aspects of the recently developed reference library system (Zehetner and Lehrach, 1994; Gress et al., 1994; Lennon et al., 1996) with the saturation offered by normalized cDNA libraries and the cellular resolution offered by the generation of cDNA probes from single cells. This would involve the generation of one master reference library that contains cDNA clones of most expressed genes of an organism, the robotic arraying of this library and production of multiple copies of high-density gridded library filters, and the hybridization analysis of these filters with complex cDNA probes (Zhao et al., 1995; Auffray et al., 1995; Harrison et al., 1995; Gress et al., 1992) generated from large numbers of tissue-specific and cell-specific mRNA samples. This system would have the advantage of being able to

easily produce and distribute these library filters to laboratories worldwide, allowing for the subsequent coordinated integration of additional experimental data with the gene expression information.

In addition to the technical and logistical advances offered by the reference library system, there are two other recent developments that would enable this type of analysis. The first is an approach for the generation of cDNA libraries that contain all clones at approximately equal levels of abundance (Soares et al., 1994; Sasaki et al., 1994; Takahashi and Ko, 1994; Patanjali et al., 1991; Ko, 1990). The composition of standard cDNA libraries reflects the profile of gene expression of the tissue from which they are made, creating the problem that abundantly expressed genes make up the majority of the clones for any given library. Normalized cDNA libraries, however, have reduced levels of these abundant clones and therefore enable the detection of more weakly expressed genes that might otherwise not be included in finite collections of clones. Similar to most subtraction procedures, normalization involves the renaturation of tracer mRNA or cDNA with driver cDNA and the subsequent removal of driver-tracer duplexes, with the difference that driver and tracer samples are identical. The rate of duplex formation is according to second-order reaction kinetics and is therefore dependent upon the abundance of individual mRNAs or cDNAs, which in this case is identical between driver and tracer pools. This principle allows for the specific removal of mRNA or cDNA species at a rate that is proportional to their original abundance, thereby creating pools with approximately equal abundance levels of all clones.

The second recent development that would enable this type of approach is the reliable PCR amplification of mRNA from single cells (Brady et al., 1995, Karrer et al., 1995; C. V. Esguerra, unpublished results). By modifying these methods with the introduction of solid-phase techniques for the manipulation of mRNA and cDNA, this technology would enable the generation of high-quality single-stranded hybridization probes from single cells, which could then be used to probe high-density cDNA library filters to quantitatively determine the comprehensive expression profile of these cells. Such procedures could of course also be adapted to analyze other small samples, such as groups of cells removed from cryostat sections (Luqmani and Lymboura, 1994) or other microdissected tissues. The detection of weakly expressed transcripts within these individual cells could be achieved

through the controlled normalization of the cDNA probes to specific C_0t values.

In summary, the systematic analysis of cell-specific gene expression, carried out through the hybridization of high-density filter arrays of normalized cDNA libraries with cell-specific and tissue-specific probes, would allow for the rapid generation of high-resolution expression data for the majority of genes in an organism. The recent development of glass microarrays (Shena et al., 1995) in place of polymer filters might allow for an even greater thoughput and sensitivity for this process. Once sufficient numbers of probes have been analyzed, this expression data could be used to correlate the activity of numerous different transcripts and to identify functionally associated genes, such as signaling molecules and their corresponding receptors, components of specific signal transduction pathways, transcription factors and their target genes, and many others. In this manner it would be possible to rapidly and quantitively determine the comprehensive gene expression profile of any given sample and to systematically correlate this data with previously obtained data, thereby allowing for the efficient characterization of novel transcripts. In combination with more specific methods for the systematic manipulation of gene expression (Spradling et al., 1995) as well as strategies for the systematic analysis of protein phosphorylation and glycosylation, this system could over the longer term enable the elucidation of the molecular mechanisms by which cells process and store information.

REFERENCES

Aalto MK, Ronne H, Keranen S: Yeast syntaxins Sso1p and Sso2p belong to a family of related membrane proteins that function in vesicular transport. *EMBO J* 1993, 12:4095-4104.

Adams MD, et al: Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* 1995, Suppl:3-183

Alderson MR, Armitage RJ, Maraskovsky E, Tough TW, Ros E, Schooley K, Ramsdell F, Lynch DH: Fas transduces activation signals in normal human T lymphocytes. *J Exp Med* 1993, 178:2231-2235.

Alexi T, Hefti F: Trophic actions of transforming growth factor α on mesencephalic dopaminergic neurons developing in culture. *Neurosci* 1993, 55:903-918.

Anzano M, Roberts AB, Smith JM, Sporn MB, De Larco JE: Sarcoma growth factor from conditioned medium of virally transformed cells is composed of both type a and type b transforming growth factors. *Proc Natl Acad Sci USA* 1983, 80:6264-6268.

Attisano L, Carcamo J, Ventura F, Weis FM, Massagué J, Wrana JL: Identification of human activin and TGFb type I receptors that form heteromeric kinase complexes with type II receptors. *Cell* 1993, 75:671-680.

Auffray C, Behar G, Bois F, Bouchier C, Da Silva C, Divignes M-D, Duprat S, Houlgatte R, Jumeau M-N, Lamy B, Lorenzo F, Mitchell H, Mariage-Samson R, Pietu G, Sebastiani-Kabaktchis C, Tissier A: IMAGE: intégration au niveau moléculaire de l'analyse du génome humain et de son expression. *C R Acad Sci Paris* 1995, 318:263-272.

Bähr M, Wizenmann A: Retinal ganglion cell axons recognize specific guidance cues present in the deafferented adult rat superior colliculus. *J Neurosci* 1996, 16:5106-5116.

Baker SJ, Reddy EP: Transducers of life and death: TNF receptor superfamily and associated proteins. *Oncogene* 1996, 12:1-9.

Balzer H-J, Bäumlein H: An improved gene expression screen. *Nucl Acids Res* 1994, 22:2853-2854.

Barnes WM: PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. *Proc Natl Acad Sci USA* 1994, 91:2216-2220.

Bergeron JJM, Brenner MB, Thomas DY, Williams DB: Calnexin: a membrane-bound chaperone of the endoplasmic reticulum. *Trends Biochem Sci* 1994, 19:124-128.

Birecree E, King LE, Nanney LB: Epidermal growth factor and its receptor in the developing human nervous system. *Dev Brain Res* 1991, 60:145-154.

Bixby JL, Jhabvala P: Tyrosine phosphorylation in early embryonic growth cones. *J Neurosci* 13:3421-3432.

Boldin MP, Varfolomeev EE, Pancer Z, Mett IL, Camonis JH, Wallach D: A novel protein that interacts with the death domain of Fas/APO1 contains a sequence motif related to the death domain. *J Biol Chem* 1995, 270:7795-7798.

von Boxberg Y, Deiss S, Schwarz U: Guidance and topographic stabilization of nasal chick retinal axons on target-derived components in vitro. *Neuron* 1993, 10:345-357.

Brady G, et al: Analysis of gene expression in a complex differentiation hierarchy by global amplification of cDNA from single cells. *Curr Biol* 1995, 5:909-922.

Brachmann R, Lindquist PB, Nagashima M, Kohr W, Lipari T, Napier M, Derynck R: Transmembrane TGFα precursors activate EGF/TGFα receptors. *Cell* 1989, 56:691-700.

Brennwald P, Novick P: Interactions of three domains distinguishing the Ras-related GTP-binding proteins Ypt1 and Sec4. *Nature* 1993, 362:560-563.

Carraway KL, Burden SJ: Neuregulins and their receptors. *Curr Opin Neurobiol* 1995, 5:606-612.

Chinnaiyan AM, O´Rourke K, Tewari M, Dixit VM: FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis. *Cell* 1995, 81:505-512.

Chu K, Niu XH, Williams LT: A Fas-associated protein factor, FAF1, potentiates Fas-mediated apoptosis. *Proc Natl Acad Sci USA* 1995, 92:11894-11898.

Courtneidge SA, Dhand R, Pilat D, Twamley GM, Waterfield MD, Roussel MF: Activation of Src family kinases by colony stimulating factor-1, and their association with its receptor. *EMBO J* 1993 12:943-950.

Davar G, Maciewicz RJ: Deafferentation pain syndromes. *Neur Clin* 1989, 7:289-304.

Defeo-Jones D, Tai JY, Vuocolo GA, Wegrzyn RJ, Shofield TL, Riemen MW, Oliff A: Substitution of lysine for arginine at position 42 of human transforming growth factor-a eliminates biological activity without changing internal disulfide bonds. *Mol Cell Biol* 1989, 9:4083-4086.

DeLarco JE, Todaro GJ: Growth factors from murine sarcoma virus-transformed cells. *Proc Natl Acad Sci USA* 1978, 75:4001-4005.

Diatchenko L, Lau Y-FC, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD: Suppression subtractive hybridization: a method for generating differentially regualted of tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* 1996, 93:6025-6030.

Drescher U, Kremoser C, Handwerker C, Löschinger J, Noda M, Bonhoeffer F: In vitro guidance of retinal ganglion cell axons by RAGS, a 25-kDa tectal protein related to ligands for eph receptor tyrosine kinases. *Cell* 1995, 82:359-370.

Ellis C, Kasmi F, Ganju P, Walls E, Panayotou G, Reith AD: A juxtamembrane autophosphylation site in the Eph family receptor tyrosine kinase, Sek, mediates high affinity interaction with p59fyn. *Oncogene* 1996,

Finch CE, Laping NJ, Morgan TE, Nochols NR, Pasinetti GM: TGFβ_1 is an organizer of responses to neurodegeneration. *J Biol Chem* 1993, 53:314-322.

Frederickson RM, Sonenberg N: Signal transduction and regulation of translation initiation. *Sem Cell Biol* 1992, 3:107-115.

Furber S, Oppenheim RW, Prevette D: Naturally-occuring neuron death in the ciliary ganglion of the chick embryo following removal of preganglionic input: evidence for the role of afferents in ganglion cell survival. *J Neurosci* 1987, 7:1816-1832.

Gautam M, Noakes PG, Moscoco L, Rupp F, Scheller RH, Merlie JP, Sanes JR. *Cell* 1996, 85.

Gohda E, Matsunaga T, Kataoka H, Yamamoto I: TGFβ is a potent inhibitor of hepatocyte growth factor secretion by human fibroblasts. *Cell Biol Int Rep* 1992, 16:917-926.

Gould E, McEwen BS: Neuronal birth and death. *Curr Opin Neurobiol* 1993, 3:676-682.

Grant SG, O'Dell TJ, Karl KA, Stein PL, Soriano P, Kandel ER: Impaired long-term potentiation, spatial learning, and hippocampal development in *fyn* mutant mice. *Science* 1992, 258:1903-1910.

Gress TM, Hoheisel JD, Lennon GG, Zehetner G, Lehrach H: Hybridization fingerprinting of high-density cDNA library arrays with cDNA pools derived from whole tissues. *Mammalian Genome* 1992, 3:609-619.

Gress TM, Müller-Pillasch F, Adler G, Zehetner G, Lehrach H: European pancreatic cancer reference library system, EPCRLS. *Eur J Cancer* 1994, 30A:1391-1394.

Guthrie KM, Nguyen T, Gall CM: Insulin-like growth factor-1 mRNA is increased in deafferented hippocampus: spatiotemporal correspondence of a trophic event with axon sprouting. *J Comp Neurol* 1995, 352:147-160.

Haffter P, Granato M, Brand M, Mullins MC, Hammerschmidt M, Kane DA, Odenthal J, van Eeden F, Jiang Y-J, Heisenberg C-P, Kelsh RN, Furatani-Seiki M, Warga R, Vogelsang L, Beuchle D, Shach U, Fabian C, Nüsslein-Volhard C: The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio. Development* 1996, in press.

Harrison SM, Dunwoodie SL, Arkell RM, Lehrach H, Beddington RSP: Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo. *Development* 1995, 121:2479-2489.

Hedrick SM, Cohen DI, Nielsen EA, Davis MM: Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. *Nature* 1984, 308:149-153.

Heldin C-H: Dimerization of cell surface receptors in signal transduction. *Cell* 1995, 80:213-223.

Hofmann K, Tschopp J: The death domain motif found in Fas (Apo-1) and TNF receptor is present in proteins involved in apoptosis and axonal guidance. *FEBS Lett* 1995, 371:321-323.

Holmes WE, Sliwkowski MX, Akita RW, Henzel WJ, Lee J, Park JW, Yansura D,Abadi N, Raab H, Lewis GD, et al.: Identification of heregulin, a specific activator of p185^{erbB2}. *Science* 1992, 256:1205-1210.

Horvitz HR, Sternberg PW, Greenwald IS, Fixsen W, Ellis HM: Mutations that affect neural cell lineages and cell fates during the development of the nematode Caenorhabditis elegans. *Cold Spring Harbor Symp Quant Biol* 1983, 48:453-63.

Hubank M, Schatz DG: Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucl Acids Res* 1994, 22:5640-5648.

Ilangmuran S, Robinson PJ, Hoessli DC: Transfer of exogenous glycosylphosphatidyinositol (GPI)-linked molecules to plasma membranes. *Trends Cell Biol* 1996, 6:163-167.

Ingraham CA, Cooke MP, Chuang YN, Perlmutter RM, Maness PF: Cell type and developmental regulation of the fyn proto-oncogene in neural retina. *Oncogene* 1992, 7:95-100.

Inouye S, Ling N, Shimasaki S: Localization of the heparin binding site of follistatin. *Mol Cell Endocrinol* 1992, 90:1-6.

Johnston IG, Paladino T, Gurd JW, Brown IR: Molecular cloning of SC1: a putative brain extracellular matrix glycoprotein showing partial similarity to osteonectin/BM40/SPARC. *Neuron* 1990, 2:165-176.

Jurgens G, Mayer U, Busch M, Lukowitz W, Laux T: Pattern formation in the Arabidopsis embryo: a genetic perspective. *Phil Trans Royal Soc London B: Biol Sci* 1995, 350:19-25.

Karrer EE, Lincoln JE, Hogenhout S, Bennett AB, Bostock RM, Martineau B, Lucas WJ, Gilchrist DG, Alexander D: *In situ* isolation of mRNA from single plant cells: creation of cell-specific cDNA libraries. *Proc Natl Acad Sci USA* 1995, 92:3814-3818.

Kaser MR, Lakshmanan J, Fisher DA: Comparison between epidermal growth factor, transforming growth factor α and EGF receptor levels in regions of adult rat brain. *Mol Brain Res* 1992, 16:316-322.

Kawakami T, Kawakami Y, Aaronson SA, Robbins KC: Acquisition of transforming properties by FYN, a normal SRC-related human gene. *Proc Natl Acad Sci USA* 1988, 85:3870-3874.

Kelleher DJ, Kreibich G, Gilmore R: Oligosaccharyltransferase activity is associated with a protein complex composed of ribophorins I and II and a 48 kd protein. *Cell* 1992, 69:55-65.

Ko MSH, Ko SBH, Takahashi N, Nishiguchi K, Abe K: Unbiased amplification of a highly complex mixture of DNA fragments by 'lone linker'-tagged PCR. *Nucl Acids Res* 1990, 18:4293.

Ko MSH: An 'equalized cDNA library' by the reassociation of short double-stranded cDNAs. *Nucl Acids Res* 1990, 18:5705-5711.

Kreibich G, Ulrich BL, Sabatini DD: Proteins of rough microsomal membranes related to ribosome binding. I. Identification of ribophorins I and II, membrane proteins characteristic of rough microsomes. *J Cell Biol* 1978, 77:464-487.

Lennon G, Auffray C, Polymeropoulos M, Soares MB: The IMAGE consortium: an integrated molecular analysis of genomes and their expression. *Genomics* 1996, 33:151-152.

Liang P, Pardee AB: Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 1992, 257:967-971.

Lindahl L, Hinnebusch A: Diversity of mechanisms in the regulation of translation in prokaryotes and lower eukaryotes. *Curr Opin Genet Dev* 1992, 2:720-726.

Linden R, Pinon LG: Dual control by targets and afferents of developmental neuronal death in the mammalian central nervous system: a study in the parabigeminal nucleus of the rat. *J Comp Neur* 1987, 266: 141-149.

Loeb JA, Fischbach GD: ARIA can be released from extracellular matrix through cleavage of a heparin-binding domain. *J Cell Biol* 1995, 130:127-135.

Lopcz-Casillas F, Wrana JL, Massagué J: Betaglycan presents ligand to the TGFβ signaling receptor. *Cell* 1993, 73:1435-1444.

Lowin B, Hahne M, Mattmann C, Tschop J: Cytolytic T-cell cytotoxicity is mediated through perforin and Fas lytic pathways. *Nature* 1994, 370:650-652.

de Luca A, Weller M, Fontana A: TGFβ-induced apoptosis of cerebellar granule neurons is prevented by depolarization. *J Neurosci* 1996, 16:4174-4185.

Lund RD, Lund JS: Synaptic adjustment after deafferentation of the superior colliculus of the rat. *Science* 1971, 171:804-807.

Lund RD, Lund JS: Modifications of synaptic patterns in the superior colliculus of the rat during development and following deafferentation. *Vision Res* 1971, Suppl 3:281-298.

Luqmani YA, Lymboura M: Subtraction hybridization cloning of RNA amplified from different cell populations microdissected from cryostat tissue sections. *Anal Biochem* 1994, 222:102-109.

Mahanthappa NK, Anton ES, Matthew WD: Glial growth factor 2, a soluble neuregulin, directly increases Schwann cell motility and indirectly promotes neurite outgrowth. *J Neurosci* 1996, 16:4673-4683.

Maness PF: Nonreceptor protein tyrosine kinases associated with neuronal development. *Dev Neurosci* 1992, 14:257-270.

Margolis B, Silvennoinen O, Comoglio F, Roonprapunt C, Skolnik E, Ullrich A, Schlessinger J: High-efficiency expression/cloning of epidermal growth factor-receptor-binding proteins with Src homology 2 domains. *Proc Natl Acad Sci USA* 1992, 89:8894-8898.

Massagué J: Epidermal growth factor-like transforming growth factor. I. Isolation, chemical characterization, and potentiation by other transforming factors from feline sarcoma virus-transformed rat cells. *J Biol Chem* 1983, 258:13606-13613.

Maters LH: Effects of neonatal deafferentation on the superficial laminae of the superior colliculus. *Brain Res* 1977, 126:19-30.

Milner JJ, Cecchini E, Dominy PJ: A kinetic model for subtractive hybridization. *Nucl Acids Res* 1995, 23:176-187.

Nagata S: Mutations in the Fas antigen gene in lpr mice. *Semin Immunol* 1994, 6:3-8.

Nakamura T, Takio K, Eto Y, Shibai H, Titani K, Sugino H: Activin-binding protein from rat ovary is follistatin. *Science* 1990, 247-836-838.

Nakamura T, Sugino K, Titani K, Sugino H: Follistatin, an activin-binding protein, associates with heparan sulfate chains of proteoglycans on follicular granulosa cells. *J Biol Chem* 1991, 266: 19432-19437.

Neff NT, Prevette D, Houenou LJ, Lewis ME, Glicksman MA, Yin QW, Oppenheim RW: Insulin-like growth factors: putative muscle-derived trophic agents that promote motoneuron survival. *J Neurobiol* 1993, 24:1578-1588.

Nichols NR, Laping NJ, Day JR, Finch CE: Increases in transforming growth factor β mRNA in hippocampus during response to entorhinal cortex lesions in intact and adrenalectomized rats. *J Neurosci Res* 1991, 28:134-139.

Nusslein-Volhard C, Wieschaus E: Mutations affecting segment number and polarity in Drosophila. *Nature* 1980, 287:795-801.

Okajima A, Miyazawa K, Kitamura N: Characterization of the promoter region of the rat hepatocyte growth factor/scatter factor gene. *Eur J Biochem* 1993, 213:113-119.

Okubo K, Hori N, Matoba R, Niiyama T, Fukushima A, Kojima Y, Matsubara K: Large-scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. *Nature Genet* 1992, 2:173-179.

Olive S, Dubois C, Schachner M, Rougon G: The F3 neuronal glycosylphosphatidylinositol-linked molecule is localized to glycolipid-enriched membrane subdomains and interacts with L1 and fyn kinase in cerebellum. *J Neurochem* 1995, 65:2307-2317.

Owen-Schaub LB, Metrissian S, Ford RJ: Fas/APO-1 expression and function on malignant cells of hematologic and nonhematologic origin. *J Immunother* 1993, 14:234-241.

Pandiella A, Bosenberg MW, Huang EJ, Besmer P, Massagué J: Cleavage of membrane-anchored growth factors involves distinct protease activities regulated through common mechanisms. *J Biol Chem* 1991, 267:24028-24033.

Park MH, Wolff EC, Folk JE: Hypusine: its post-translational formation in eukaryotic initation factor 5A and its potential role in cellular regulation. *Biofactors* 1993, 4:95-104.

Patanjali SR, Parimoo S, Weissman SM: Construction of a uniform-abundance (normalized) cDNA library. *Proc Natl Acad Sci USA* 1991, 88:1943-1947.

Patthy L, Nikolics, K: Functions of agrin and agrin-related proteins. *Trends Neurosci* 1993, 16:76-81.

Pontius BW, Berg P: Rapid renaturation of complementary DNA strands mediated by cationic detergents: a role for high-probability domains in enhancing the kinetics of molecular assembly processes. *Proc Natl Acad Sci USA* 1991, 88:8237-8241

Potenza M, Bowser R, Müller H, Novick P: SEC6 encodes an 85 kDa soluble protein required for exocytosis in yeast. *Yeast* 1992, 8:549-558.

Rabizadeh S, Bredesen DE: Is $p75^{NGFR}$ involved in developmental neural cell death? *Dev Neurosci* 1994, 16:207-211.

Raines EW, Lane TF, Iruela-Arispe ML, Ross R, Sage EU: The extracellular glycoprotein SPARC interacts with platelet-derived growth factor (PDGF)-AB and -BB and inhibits the binding of PDGF to its receptors. *Proc Natl Acad Sci USA* 1992, 89:1281-1304.

Rubin LL, Gatchalian D, Rimon G, Brooks SF: The molecular mechanisms of neural apoptosis. *Curr Opin Neurobiol* 1994, 4:696-702.

Ruoslahti RE, Yamaguchi Y: Proteoglycans as modulators of growth factor activities. *Cell* 1991, 64:867-869.

Sargent TD, Dawid IB: Differential gene expression in the gastrula of Xenopus laevis. *Science* 1983, 222:135-139.

Sasahara M, Fries JW, Raines EW, Gown AM, Westrum LE, Frosch MP, Bonthron DT, Ross R, Collins T: PDGF B-chain in neurons of the central nervous system, posterior pituitary, and in a transgenic model. *Cell* 1991, 64:217-227.

Sasahara A, Kott JN, Sasahara M, Raines EW, Ross R, Westrum LE: Platelet-derived growth factor B-chain-like immunoreactivity in the developing and adult rat brain. *Dev Brain Res* 1992, 68:41-53.

Sasaki YF, Ayusawa D, Oishi M: Construction of a normalized cDNA library by introduction of a semi-solid mRNA-cDNA hybridization system. *Nuc Acids Res* 1994, 22:987-992.

Shastry BS: More to learn from gene knockouts. *Mol Cell Biochem* 1994, 136:171-182.

Schena M, Shalon D, Davis RW, Brown PO: Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995, 270:467-470.

Shenoy-Scaria AM, Kwong J, Fujita T, Olszowy MW, Shaw AS, Lublin DM: Signal transduction through decay-accelerating factor. Interaction of glycosyl-phosphatidylinositol anchor and protein tyrosine kinases p56lck and p59fyn. *J Immunol* 1992, 149:3535-3541.

Schwab ME: Myelin-associated inhibitors of neurite growth. *Exp Neurol* 1990, 109:2-5.

Seroogy KB, Gall CM, Lee DC, Kornblum HI: Proliferative zones of postnatal rat brain express epidermal growth factor receptor mRNA. *Brain Res* 1995, 670:157-164.

Shimonaka M, Inouye S, Shimasaki S, Ling N: Follistatin binds to both activin and inhibin through the common subunit. *Endocrinol* 1991, 128: 3313-3315.

Shum L, Reeves SA, Kuo AC, Fromer ES, Derynck R: Association of the transmembrane TGFα precursor with a protein kinase complex. *J Cell Biol* 1994, 125:903-916.

Soares MB, Bonaldo MF, Jelene P, Su L, Lawton L, Efstratiadis A: Construction and characterization of a normalized cDNA library. *Proc Natl Acad Sci USA* 1994, 91:9228-9232.

Smith CA, Farrah T, Goodwin RG: The TNF receptor superfamily of cellular and viral proteins: activation, costimulation, and death. *Cell* 1994, 76:969-972.

Söllner T, Whiteheart SW, Brunner M, Erdjument-Bromage H, Geromanos S, Tempst P, Rothman JE: SNAP receptors implicated in vesicle targeting and fusion. *Nature* 1993, 362: 318-324.

Söllner T, Bennett MK, Whiteheart SW, Scheller RH, Rothman JE: A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell* 1993, 75:409-418.

Spivak-Kroizman T, Lemmon MA, Dikic I, Ladbury JE, Pinchasi D, Huang J, Jaye M, Crumley G, Schlessinger J, Lax I: Heparin-induced oligomerization of FGF molecules is responsible for FGF receptor dimerization, activation, and cell proliferation. *Cell* 1994, 79:1015-1024.

Spradling AC, Stern DM, Kiss I, Roote J, Laverty T, Rubin GM: Gene disruptions using P transposable elements: an integral component of the Drosophila genome project. *Proc Natl Acad Sci USA* 1995, 92:10824-10830.

Stanger BZ, Leder P, Lee T, Kim E, Seed B: RIP: a novel protein containing a death domain that interacts with Fas/APO-1 (CD95) in yeast and causes cell death. *Cell* 1995, 81:513-523.

Sugino K, Kurosawa N, Nakamura T, Takio K, Shimasaki S, Ling N, Titani K, Sugino H: Molecular heterogeneity of follistatin, an activin-binding protein. *J Biol Chem* 1993, 268:15579-15587.

Suzuki T, Okumura-Noji K: NMDA receptor subunits epsilon 1 (NR2A) and epsilon 2 (NR2B) are substrates for Fyn in the postsynaptic density fraction isolated from the rat brain. *Biochem Biophys Res Comm* 1995, 216:582-588.

Sweet, WH: Deafferentation pain in man. *Appl Neurophysiol* 1988, 51:117-127.

Takahashi N, Ko MSH: Toward a whole cDNA catalog: construction of an equalized cDNA library from mouse embryos. *Genomics* 1994, 23:202-210.

Takayama S, Sasahara M, Iihara K. Handa J, Hazama F: Platelet-derived growth factor B-chain-like immunoreactivity in injured rat brain. *Brain Res* 1994, 653:131-140.

Ting AE, Hazuka CD, Hsu S-C, Kirk MD, Bean AJ, Scheller RH: rSec6 and rSec8, mammalian homologs of yeast proteins essential for secretion. *Proc Natl Acad Sci USA* 1995, 92:9613-9617.

Toma JG, Pareek S, Barker P, Mathew TC, Murphy RA, Acheson A, Miller FD: Spatiotemporal increases in epidermal growth factor receptors following peripheral nerve injury. *J Neurosci* 1992, 12:2504-2515.

Toyoshima K, Yamanashi Y, Inoue K, Semba K, Yamamoto T, Akiyama T: Protein tyrosine kinases belonging to the Src family. *Ciba Found Symp* 1992, 164:240-253.

Trachtenberg JT, Thompson WJ: Schwann cell apoptosis at developing neuromuscular junctions is regulated by glial growth factor. *Nature* 1996, 379:174-177.

Twamley GM, Kypta RM, Hall B, Courtneidge SA: Association of Fyn with the activated platelet-derived growth factor receptor: requirements for binding and phosphorylation. *Oncogene* 1992, 7:1893-1901.

Wang Z, Brown DD: A gene expression screen. *Proc Natl Acad Sci USA* 1991, 88:11505-11509.

Watanabe-Fukunaga R, Brannan CI, Copland NG, Jenkins NA, Nagata S: Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature* 1992, 356:314-317.

Wen D, Suggs SV, Karunagaran D, Liu N, Cupples RL, Luo Y, Jansen ZM, Ben-Baruch N, Trollinger DB, Jacobsen VL, et al.: Structural and functional aspects of the multiplicity of Neu differentiation factors. *Mol Cell Biol* 1994, 14:1909-1919.

Wilcox JN, Derynck R: Localization of cells synthesizing transforming growth factor α mRNA in the mouse brain. *J Neurosci* 1988, 8:1901-1904.

Wizenmann A, Thies E, Klostermann S, Bonhoeffer F, Bähr M: Appearance of target-specific guidance information for regenerating axons after CNS lesions. *Neuron* 1993, 11:975-983.

Wong ST, Winchell LF, McCune BK, Earp HS, Teixido J, Massagué J, Hermann B, Lee DC: The TGFα precursor expressed on the cell surface binds to the EGF receptor on adjacent cells, leading to signal transduction. *Cell* 1989, 56:495-506.

Wrana JL, Attisano L, Carcamo J, Zentella A, Doody J, Laiho M, Wang XF, Massagué J: TGFβ signals through a heteromeric protein kinase receptor complex. *Cell* 1992, 71:1003-1014.

Yagi T, Aizawa S, Tokunaga T, Shigetani Y, Takeda N, Ikawa Y: A role for Fyn tyrosine kinase in the suckling behaviour of neonatal mice. *Nature* 1993, 366:742-745.

Yeh HJ, Ruit KG, Wang YX, Parks WC, Snider WD, Deuel TF: PDGF A-chain gene is expressed by mammalian neurons during development and in maturity. *Cell* 1991, 64:209-216.

Yeh HJ, Silos-Santiago I, Wang YX, George RJ, Snider WD, Deuel TF: Developmental expression of the platelet-derived growth factor α-receptor gene in mammalian central nervous system. *Proc Natl Acad Sci USA* 1993, 90:1952-1956.

Yu YH, Sabatini DD, Kreibich G: Antiribophorin antibodies inhibit the targeting to the ER membrane of ribosomes containing nascent secretory polypeptides. *J Cell Biol* 1990, 111:1335-42.

Zehetner G, Lehrach H: The reference library system - sharing biological material and experimental data. *Nature* 1994, 367:489-491.

Zisch AH, D'Alessandri L, Amrein K, Ranscht B, Winterhalter KH, Vaughan L: The glypiated neuronal cell adhesion molecule contactin/F11 complexes with Src-family protein tyrosine kinase Fyn. *Mol Cell Neurosci* 1995, 6:263-79.

Zhao N, Hashida H, Takahashi N, Misumi Y, Sakaki Y: High-density cDNA filter analysis: a novel approach for large-scale, quantitative analysis of gene expression. *Gene* 1995: 207-213.

APPENDIX

Index of Figures

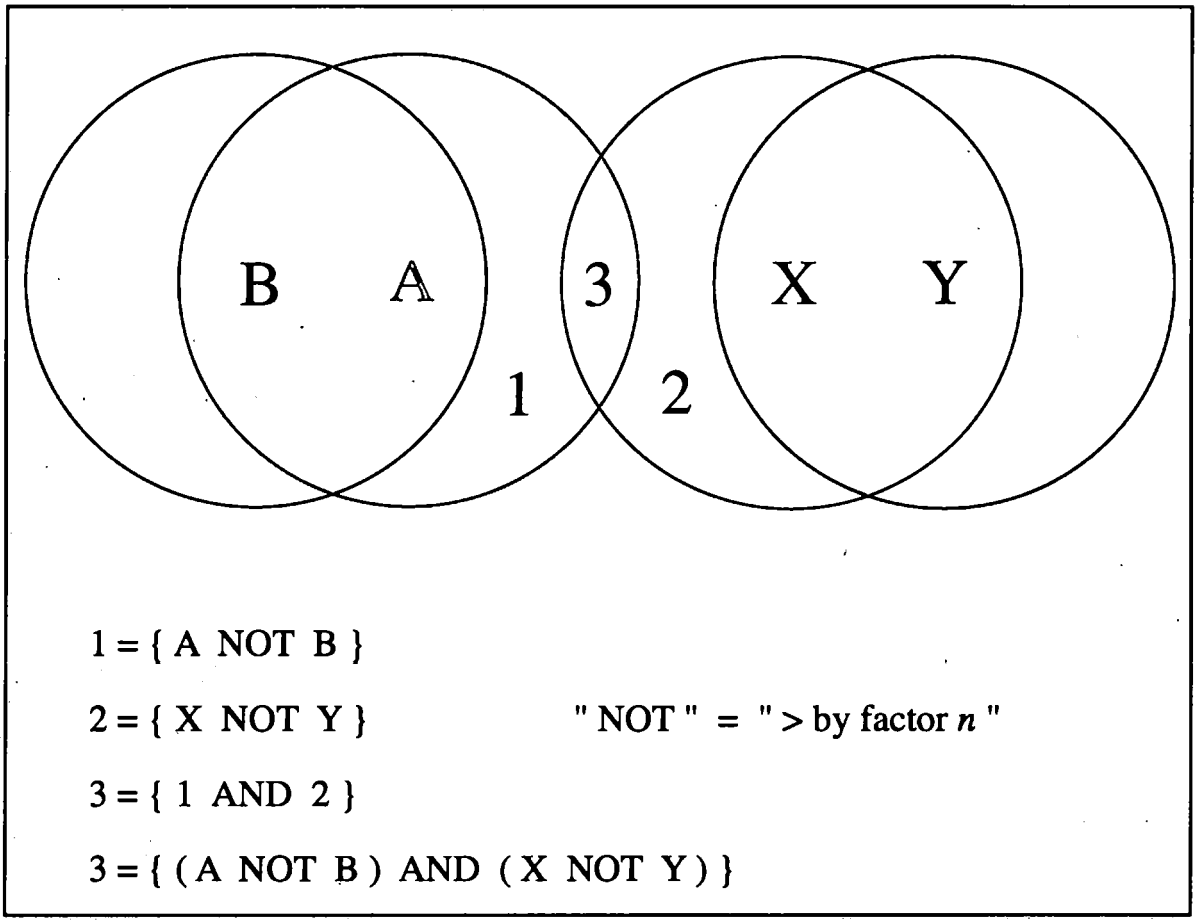


Figure 1. Gene Identification by Boolean Selection

For the rapid identification of transcripts specifically expressed in tissue A at timepoint X, Boolean selection involves two initial negative selections through the subtractive hybridizations 1 = {region A NOT region B} and 2 = {timepoint X NOT timepoint Y}. Here NOT is understood to mean "greater than by factor *n*", where *n* is dependent upon the ratio of driver cDNA to tracer cDNA. This is followed by one positive selection through the rescue hybridization 3 = {1 AND 2} = {(A NOT B) AND (X NOT Y)}. This allows for the enrichment of cDNAs common to both subtracted pools and enables the rapid selection of those cDNA clones that fulfill *both* criteria {region A NOT region B} and {timepoint X NOT timepoint Y}.

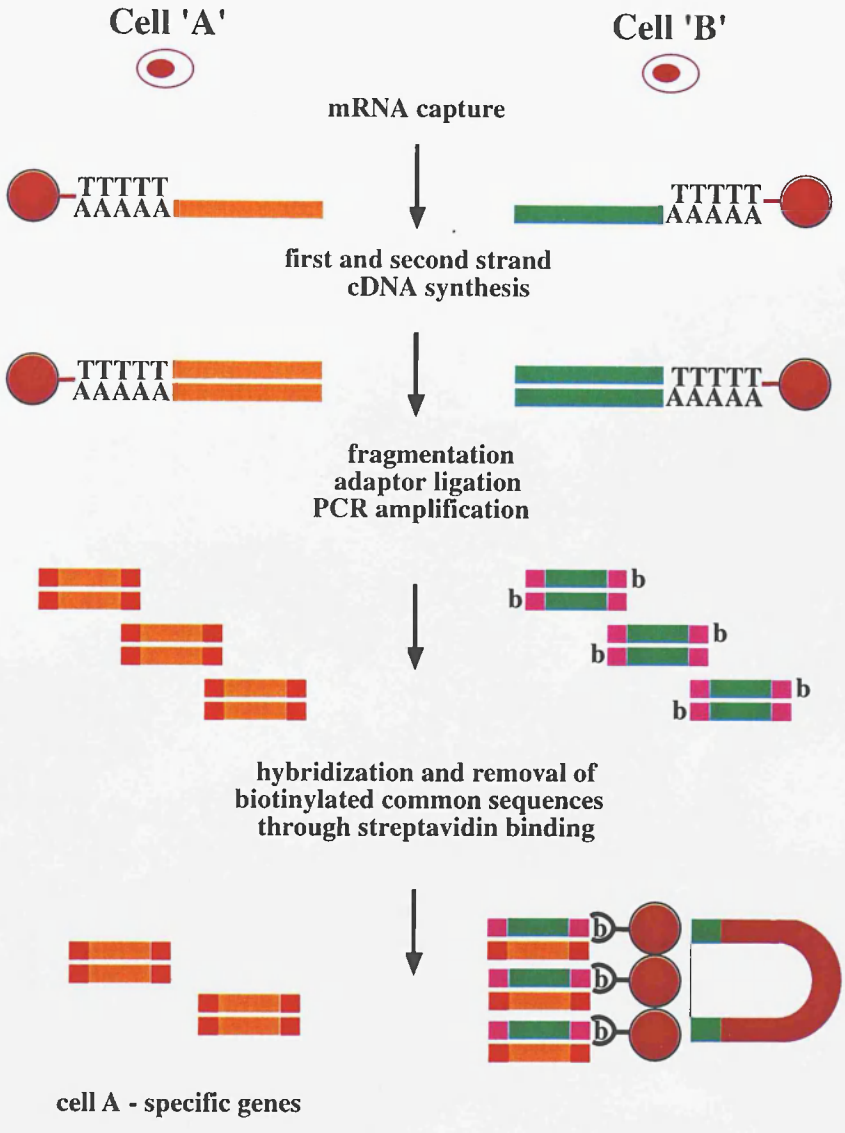


Figure 2. cDNA Synthesis, Amplification, and Subtraction

After mRNA purification and cDNA synthesis, the cDNA is fragmented with 4-base restriction endonucleases, ligated to adaptors, and PCR amplified with adaptor-specific primers (biotinylated primers are used for the driver cDNA). Driver and tracer cDNAs are combined at the ratio 20:1, denatured and allowed to renature in an optimized hybridization buffer. Driver-tracer hybrids and excess driver cDNAs are removed with streptavidinylated magnetic beads and the remaining tracer cDNA is reamplified for the next round of subtraction.

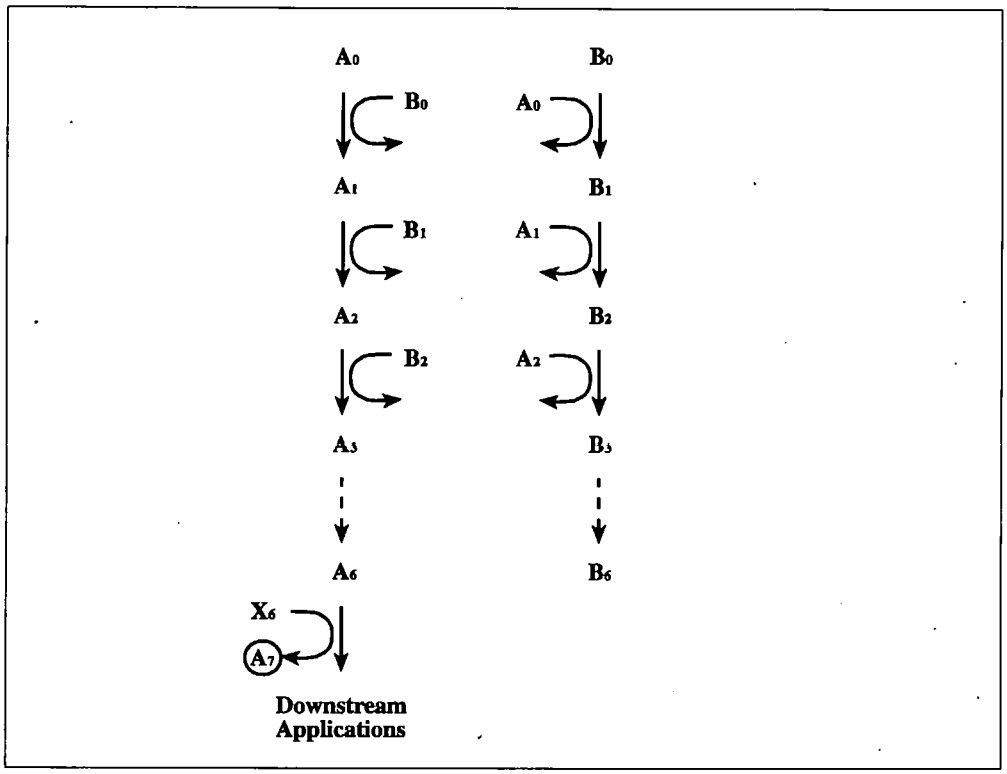


Figure 3. Subtraction and Rescue Hybridization Scheme

For the efficient enrichment of differentially expressed transcripts, the subtraction strategy was designed so as to remove all common transcripts, regardless of their original abundance. As each subtraction is essentially a normalization with regard to these common transcripts, repetitive use of the original driver cDNA does not efficiently remove weakly to moderately expressed transcripts, which increase in relative abundance after each round of subtraction. The efficient removal of these transcripts can therefore only be achieved with the use of driver cDNAs that have the same transcript abundance profile, which holds true for the corresponding cDNA pool of the opposite sample when serial subtractions are done in parallel. This is done by maintaining the sequential order of driver-tracer matches such that A_0 was paired with B_0, A_1 with B_1, A_2 with B_2, and so on. The rescue hybridization is carried out with the final subtracted cDNA pool (A_6) as the target and another subtracted cDNA pool (X_6) as the bait (the bait cDNA pool is biotinylated and used to select identical amplicons from the target cDNA pool).

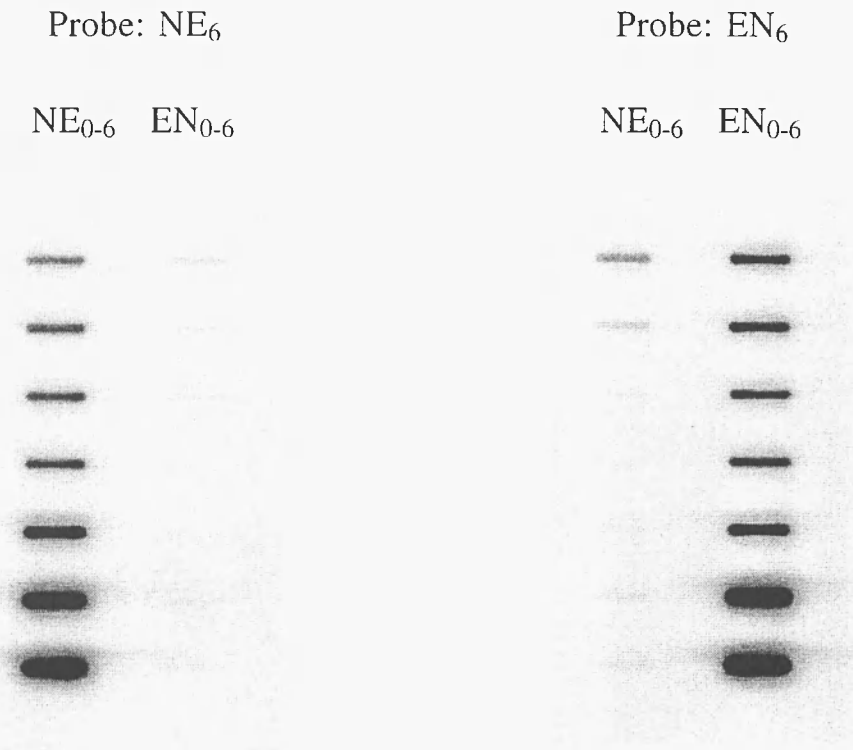
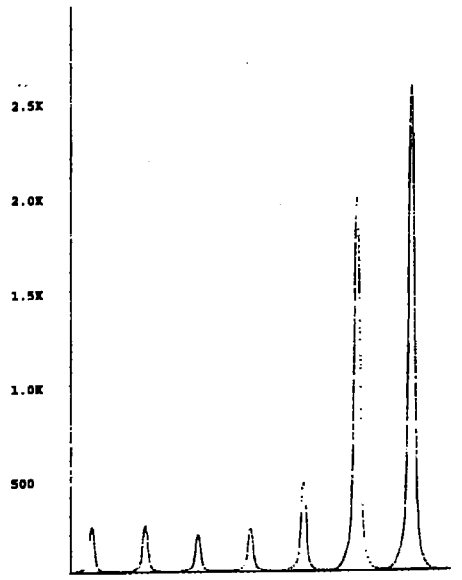
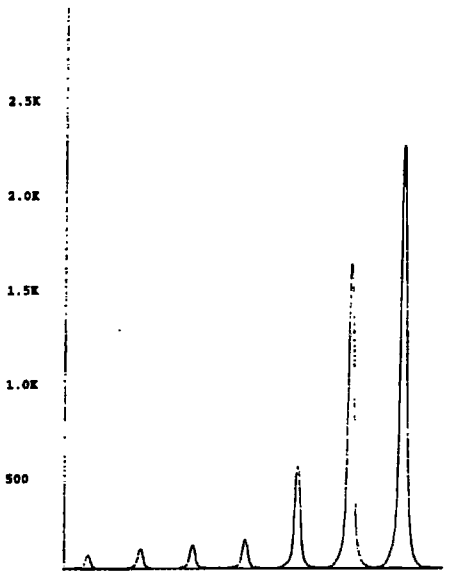


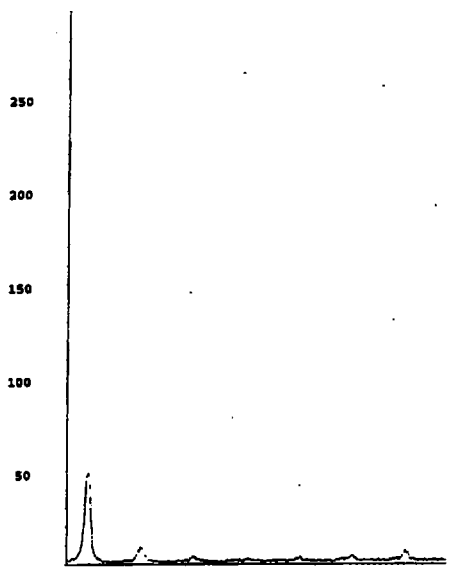
Figure 4. Enrichment Efficiency of the Improved Subtraction Procedure
In order to test the degree of enrichment of differentially regulated transcripts as provided by the improved subtraction prodecure, the enriched cDNA pools after the final round of subtractions were used as hybridization probes against all cDNA pools after each round of subtraction. Depicted above are the slot blot results obtained by carrying out these hybridizations using as probes the EN_6 pool (right panel) and the NE_6 pool (left panel). In each panel the right column of slots contains 1 μg each of EN_0, EN_1, EN_2, ..., and EN_6 in descending order, and the left column of slots contains 1 μg each of NE_0, NE_1, NE_2, ..., and NE_6, also in descending order. No cross-hybridization is seen between the two final pools EN_6 and NE_6 , indicating that nearly all common cDNAs were efficiently removed. The increased level in expression of the enriched transcripts can be seen by comparing the signal intensities of the two original cDNA pools EN_0 and NE_0 for each probe.



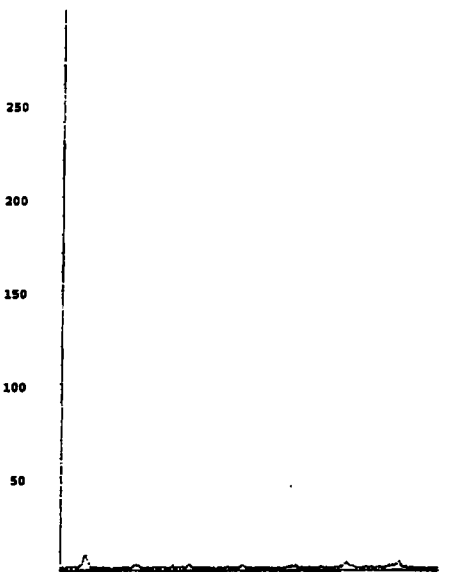
Probe: EN_6, Targets: EN_{0-6}



Probe: NE_6, Targets: NE_{0-6}



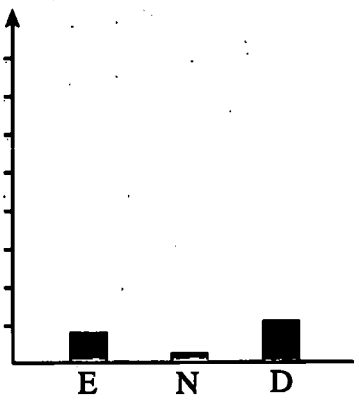
Probe: EN_6, Targets: NE_{0-6}



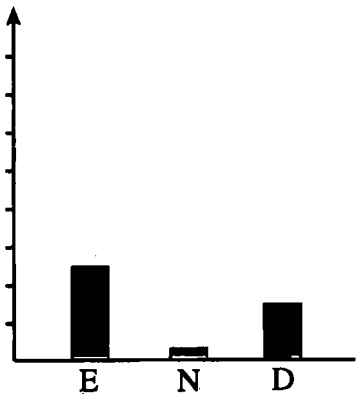
Probe: NE_6, Targets: EN_{0-6}

Figure 5. Quantification of Enrichment and Subtraction Efficiencies

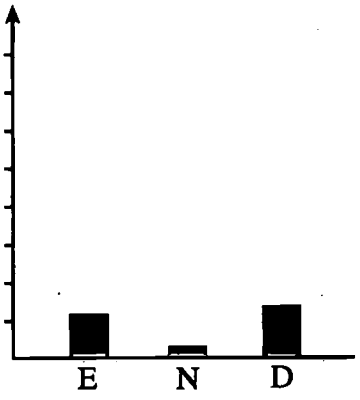
For the quantitative assessment of the subtraction results depicted in Figure 4, the signal intensities of all target cDNA pools were determined for each probe using quantitative autoradiography. The y-axis of each plot represents arbitrary units for the determination of relative signal intensities, and the x-axis represents the 7 different cDNA pools of one sample after each round of subtraction (EN_{0-6}, for example). The two lower panels have 10-fold lower y-axis values in order to depict the little remaining cross-hybridization between each probe and the cDNA pools of the opposite sample.



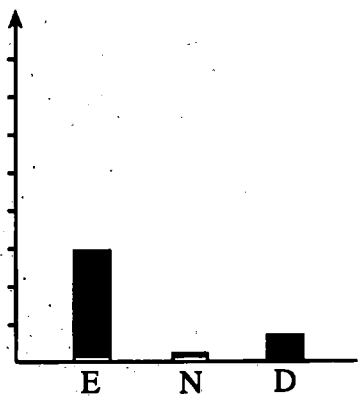
clone EN7-1
novel



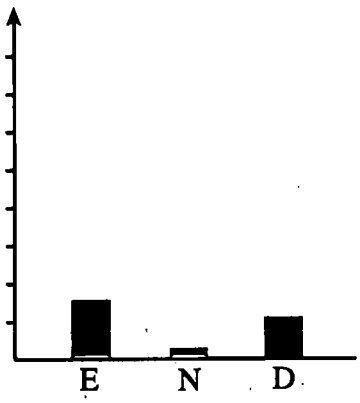
clone EN7-2
rSec6



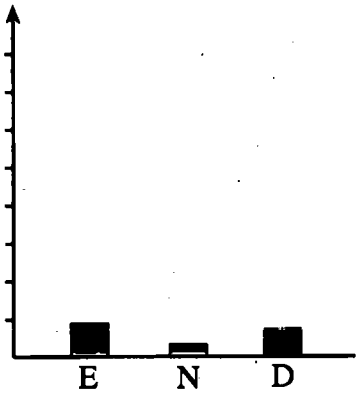
clone EN7-3
novel



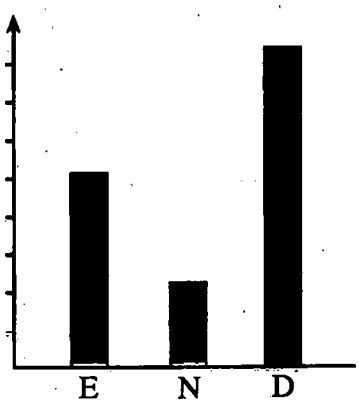
clone EN7-4
Fyn tyrosine kinase



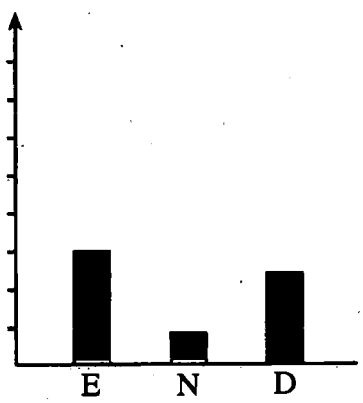
clone EN7-5
novel



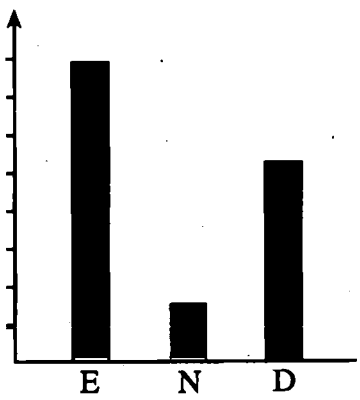
clone EN7-6
novel



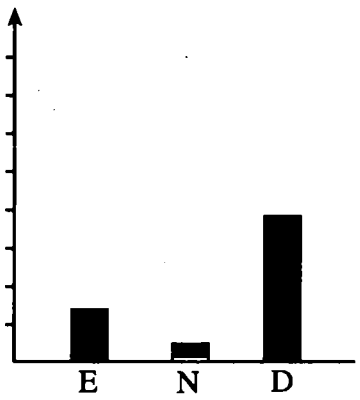
clone EN7-7
calnexin



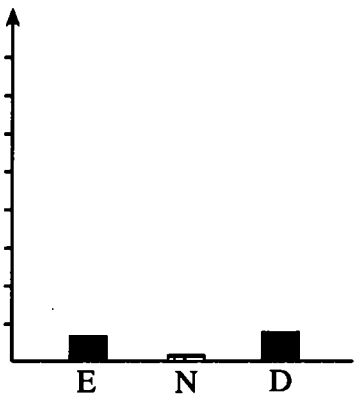
clone EN7-10
novel EGF-like gene



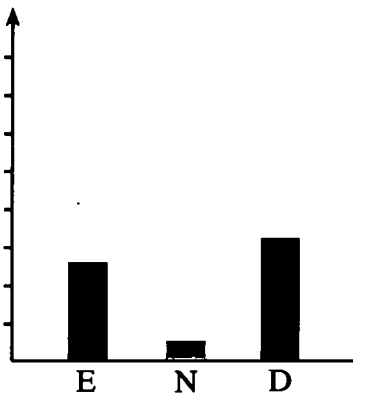
clone EN7-12
initiation factor



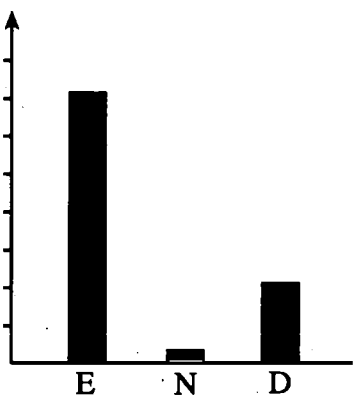
clone EN7-14
novel



clone EN7-21
novel NGF-induced gene



clone EN7-22
Fas-associated factor



clone EN7-23
ribophorin

Figure 6. Quantification of Expression Levels of Selected Clones

The actual differential expression of the first 24 clones was determined by using these cDNA fragments as hybridization probes with slot blots containing 1μg each of the original cDNA pools (E, N, and D represent embryonic, normal adult, and deafferented adult superior colliculus, respectively). Shown above and on the previous page are only those 14 clones that were upregulated in both embryonic and deafferented SC (clones EN7-12 and EN7-13 were identical, therefore only one is shown). The remaining 10 clones were only upregulated in deafferented SC as compared to normal adult SC (data not shown) and were not analyzed further. The 14 clones that fulfilled both selection criteria were sequenced and the putative sequence identification determined by BLAST searches (on both DNA and protein level) in all accessible electronic sequence databases.

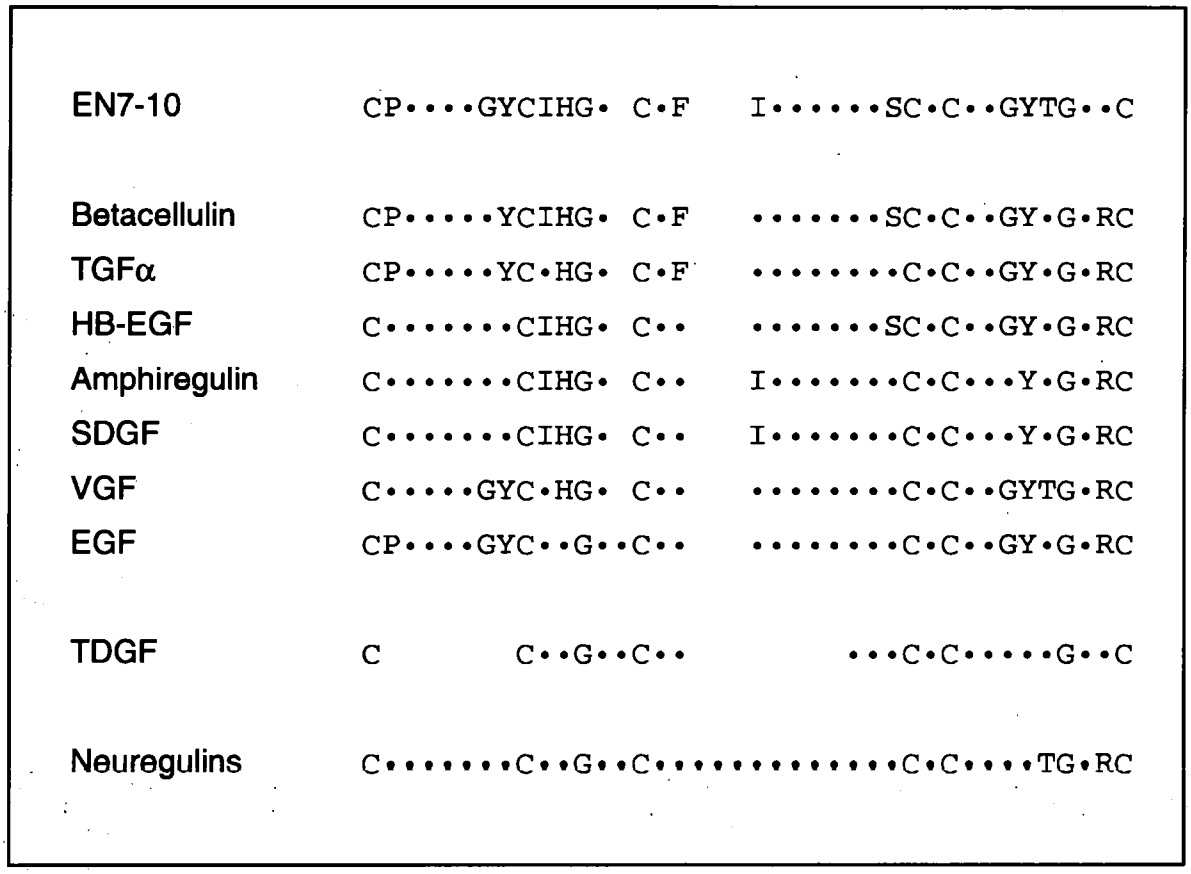


Figure 7. EN7-10: Sequence Homologies within the EGF Domain

The amino acid sequence of the EGF domain of clone EN7-10 was compared to that of other EGF superfamily members. This comparison was limited to the sequence between the distal and proximal cysteines of the six-cysteine motif that makes up most of the EGF domain. Abbreviations are: TGFα (transforming growth factor α), HB-EGF (heparin-binding EGF-like growth factor), SDGF (Schwannoma-derived growth factor), VGF (Vaccinia virus growth factor), and EGF (epidermal growth factor). The sequence shown for the neuregulins is the consensus sequence of several different splice variants (heregulins, glial growth factors, neu differentiation factor (NDF), and acetylcholine receptor-inducing activity (ARIA)).

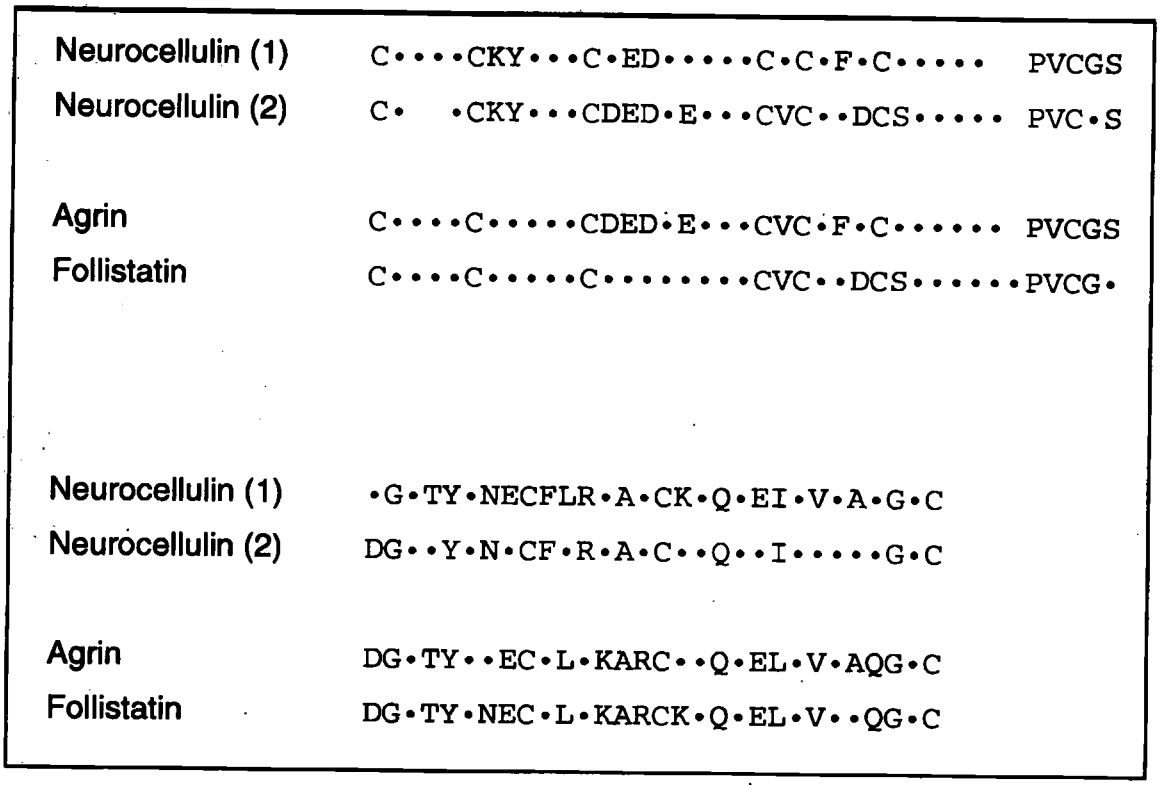


Figure 8. EN7-10: Sequence Homologies within the Follistatin Domains
The amino acid sequence of the two follistatin domains of the human homolog of clone EN7-10 was compared to those of follistatin and agrin. This comparison was limited to the sequence between the distal and proximal cysteines of the ten-cysteine motif that makes up most of the follistatin domain. Shown here is only one of the nine follistatin domains of agrin; the two follistatin domains of the human EN7-10 homolog are shown in sequential order. Conserved cysteines are depicted in pink, other residues conserved between all four domains are shown in green, and additional residues only conserved in two or three of these domains are shown in black.

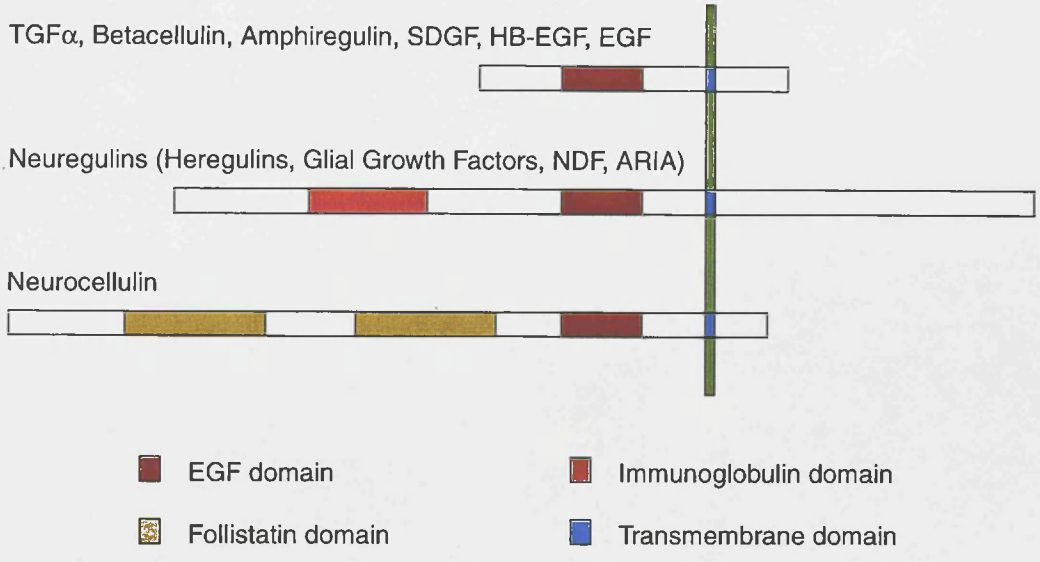


Figure 9. Structural Comparison of EGF Superfamily Members

The overall structure of the human homolog of clone EN7-10 was compared with those of the EGF family of transmembrane growth factors and the neuregulins. Abbreviations are: TGFα (transforming growth factor α), HB-EGF (heparin-binding EGF-like growth factor), SDGF (Schwannoma-derived growth factor), VGF (Vaccinia virus growth factor), and EGF (epidermal growth factor). The structure shown for the neuregulins is only of several different splice variants (heregulins, glial growth factors, neu differentiation factor (NDF), and acetylcholine receptor-inducing activity (ARIA)), yet most of the other variants are essentially similar with respect to the overall structure.