

Identification of diverse nerve growth factor-regulated genes by serial analysis of gene expression (SAGE) profiling

James M. Angelastro, Lars Klimaschewski, Song Tang, Ottavio V. Vitolo, Family A. Weissman, Laura T. Donlin, Michael L. Shelanski, and Lloyd A. Greene*

Department of Pathology, Center for Neurobiology and Behavior and Taub Institute, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032

Communicated by Joseph Schlessinger, New York University Medical Center, New York, NY, July 10, 2000 (received for review June 6, 2000)

Neurotrophic factors such as nerve growth factor (NGF) promote a wide variety of responses in neurons, including differentiation, survival, plasticity, and repair. Such actions often require changes in gene expression. To identify the regulated genes and thereby to more fully understand the NGF mechanism, we carried out serial analysis of gene expression (SAGE) profiling of transcripts derived from rat PC12 cells before and after NGF-promoted neuronal differentiation. Multiple criteria supported the reliability of the profile. Approximately 157,000 SAGE tags were analyzed, representing at least 21,000 unique transcripts. Of these, nearly 800 were regulated by 6-fold or more in response to NGF. Approximately 150 of the regulated transcripts have been matched to named genes, the majority of which were not previously known to be NGF-responsive. Functional categorization of the regulated genes provides insight into the complex, integrated mechanism by which NGF promotes its multiple actions. It is anticipated that as genomic sequence information accrues the data derived here will continue to provide information about neurotrophic factor mechanisms.

Neurotrophins, exemplified by nerve growth factor (NGF), exert a variety of actions on their targets, including regulation of proliferation, differentiation, neurite growth, neurotransmission, plasticity, repair, and survival (1, 2). Good progress has been made in uncovering initial steps in the receptor-dependent signaling mechanism by which NGF and other neurotrophins work (3, 4). Beyond the initial signaling events, NGF promotes its actions by means of both transcription-independent and -dependent pathways (3, 5, 6). Understanding the mechanism and consequences of neurotrophin responses therefore requires a description of the genes that are subject to regulation by these factors.

Detection of neurotrophin-regulated genes necessitates cellular models that can be compared before and after factor exposure. Because many neurotrophin-responsive cells require the factors for survival, they are not optimally suited for such experiments. For this reason, a large percentage of NGF gene regulation studies have used the PC12 line of rat pheochromocytoma cells (7, 8). These do not require NGF in serum-containing media, but respond to NGF by changing their phenotype from that of proliferating chromaffin-like cells to that resembling nonproliferating, neurite-bearing sympathetic neurons. Application of a variety of approaches has identified on the order of 50 genes that respond to NGF (9–13). These include immediate early genes as well as those that are regulated relatively late in the differentiation process and that encode proteins with clear roles in neuronal function (1–4).

Despite such progress, present data suggest that many additional NGF-responsive genes remain to be identified. It has been estimated that 5–10% of the genes expressed in PC12 cells may be NGF-regulated (11, 12), which would suggest regulation of at least 1,000 transcripts. Detecting and identifying such transcripts is important because it provides key

insight at the molecular level as to how trophic factors exert their many actions. Although the PC12 cell model is not without limitations, the extent and utility of information regarding NGF mechanisms and neuronal function that it has furnished strongly support its continued exploitation for defining transcriptional actions of neurotrophins.

To carry out a comprehensive analysis of NGF-dependent gene regulation in PC12 cells, we chose the SAGE (serial analysis of gene expression) technology first described by Velculescu *et al.* (14). Among the advantages of SAGE is that it has the potential to provide detection of all genes expressed in a given cell type, provides quantitative information about the relative expression of such genes, permits ready comparison of gene expression in the same cell type under different conditions (e.g., before and after NGF exposure), and yields sequence information that can be used to identify the detected genes. Thus far, SAGE methodology has proved itself to reliably detect expression of regulated and nonregulated genes in a variety of cell types (15–17). Here, we describe results garnered from comparative SAGE profiling of approximately 157,000 tags from PC12 cells before and after long-term exposure to NGF.

Materials and Methods

Cell Culture. PC12 and PC12-derived cells were cultured as described (8) on 100-mm plastic dishes coated with rat tail collagen in complete medium (85% RPMI medium, 10% horse serum, and 5% FBS). For NGF treatment, cells were plated at approximately 20% confluency and maintained for 4 days in serum-free RPMI medium and 50 ng/ml recombinant human NGF (a kind gift of Genentech) and then for 5 days in complete medium with the same NGF concentration. The initial exposure to serum-free conditions ensured death of any NGF-nonresponsive cells, and the reintroduction of complete medium assured that both naive and NGF-treated cultures were exposed to the same serum concentrations.

SAGE. Total cellular RNA was isolated from naive and NGF-treated cells after two washes with ice-cold PBS and extraction with 1 ml of Tri-reagent per dish (Molecular Research Products, Cincinnati). Polyadenylated mRNA was isolated by oligo(dT) chromatography using the MessageMaker kit (Life Technologies, Grand Island, NY). Five micrograms of mRNA was used to construct each SAGE library, using SAGE protocol version 1.0c (14). In brief, mRNA transcripts were converted to cDNAs with

Abbreviations: NGF, nerve growth factor; SAGE, serial analysis of gene expression; NCBI, National Center for Biotechnology Information; EST, expressed sequence tag; GST, glutathione S-transferase.

*To whom reprint requests should be addressed. E-mail: lag3@columbia.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

biotinylated oligo(dT)₁₈ as the primer (cDNA synthesis kit, Life Technologies). The cDNAs were digested with *Nla*III, and the 3' ends were bound to streptavidin-coated magnetic beads and then ligated to linkers 1 and 2. For the naïve cell library, we modified the protocol to use 1/5 the amount of the linkers to reduce subsequent nonspecific incorporation of partial linker sequences into concatemers. *Bsm*F1 IIS then was used to release the tags, which then were blunted at their 3' ends and ligated to form ditags, which then were amplified by PCR. The amplified ditags were isolated by gel electrophoresis and, after extraction, purified through a silica membrane column and redigested with *Nla*III to remove the linkers. The glass-membrane purification step was used to enhance the reliability of digestion with *Nla*III (18). The ditags were concatemerized at their *Nla*III overhangs (14). Concatemers with minimum size of 500 bp were obtained by gel purification. These were ligated into cloning vector pcZERO plasmid (Invitrogen) and transformed into DH10B bacteria (Life Technologies) by electroporation. For both libraries, plasmids with concatemer inserts were cycle sequenced with Big Dye terminator chemistry and analyzed either on a 377 or 3700 Applied Biosystems DNA sequencer. The tags were extracted from the sequences and analyzed by using the SAGE 300 software package (14).

Matching Tags with Transcripts. We considered only tags encountered two or more times for the two libraries combined. Tags were initially analyzed with the National Center for Biotechnology Information (NCBI) rat SAGE tag to gene mapping database (<ftp://ncbi.nlm.nih.gov/pub/sage/map/Rn/Nla3>), which matches possible 14-mer tags with known rat genes and expressed sequence tags (ESTs). With the use of sequences present in the NCBI UNIGENE rat database, potential matches were further scrutinized to determine whether there was a match at the 15th base and to determine whether the matched sequence could be confirmed as being at the most 3' end of a known rat transcript or EST. We considered only cases in which a clear poly(A) tail and/or polyadenylation signal was present on the transcript or EST. Matched ESTs were further analyzed by an advanced BLAST search for homology with known genes.

Northern Blot Analysis. Ten micrograms of total RNA (isolated as described above) from each sample was analyzed as described (19). The blots were exposed and quantified with a Storm PhosphorImager (Amersham Pharmacia–Molecular Dynamics). Probes were kindly supplied by R. Dalla-Favera, Columbia University (human β -actin), R. Vallee, University of Massachusetts, Worcester (LIC-2; ref. 20), P. Cserjesi, Columbia University (MASH-1), and D. Park, University of Ottawa (CDK4). Full-length probes for rat cofilin and galectin-1 were generated by PCR using primers that flank their ORFs. Probes for 14-3-3- γ , ATFx, MAP4, PYBP2, BTF3, pleiotrophin, CD9, Phox2a, and Tsc-22 were rat ESTs produced by the University of Iowa rat genome project and purchased from Research Genetics, Huntsville, AL. All probes were sequenced to verify identity.

Western Blotting. Western blots were carried out as described (19) with a mouse anti-cofilin antibody (1:400) and rabbit anti-ADF antiserum (1:5,000) kindly provided by J. Bamburg, Colorado State University, Fort Collins (21).

Results and Discussion

Generation and Analysis of PC12 Cell SAGE Tags. The present study used matched cultures of PC12 cells (passage 27) harvested at comparable cell density (25–50% confluency) with or without 9 days of NGF exposure. This relatively long treatment was chosen because it elicits extensive neuritogenesis, cessation of proliferation, and acquisition of a neuronal phenotype. The cultures

were used to construct appropriate libraries that were subjected to the SAGE protocol.

The present analysis considered (after exclusion of duplicate ditags, tags of less than 15 bases, mitochondrial genes, and repetitive elements) 74,880 15-mer tags from NGF-untreated (naïve) cells and 82,743 from NGF-treated cells. These represented 75,317 different transcripts for the two libraries combined. This number is unlikely to reflect extensive sequencing errors. Zhang *et al.* (16) reported a maximal sequencing error rate of 0.7% per base in their SAGE study; repeat analysis of the same concatemers in our study indicated an error rate per base of less than 0.2%. By comparison, single libraries analyzed by the Cancer Genome Anatomy Project (CGAP) SAGE consortium (as of January 2000) consisting of 50,000–100,000 tags reported up to 33,000 novel transcripts (<http://www.ncbi.nlm.nih.gov/CGAP/>). The higher level of transcripts detected here may in part reflect our use of 15-mer tags rather than the 14-mer tags analyzed by CGAP (see below), as well as apparent stimulation of low-abundance genes by NGF (see Table 2, which is published as supplemental material on the PNAS web site, www.pnas.org). Additionally, PC12 cells may express basal levels of many transcripts. This possibility is supported by the observation that even at the present level of analysis, 3/4 of the transcripts are represented by only a single tag.

If one considers transcripts represented by tags detected at least twice, then 16,199 were observed for the naïve cell library and 12,761 for the library of NGF-treated cells. For both libraries combined, 21,086 unique transcripts were represented two or more times. These data compare well with those of Velculescu *et al.* (17) who reported 23,580 unique genes expressed in brain libraries totaling 202,448 tags represented two or more times. Of the 21,000 unique transcripts here, approximately 2,400 (11%) were matched to known, named genes and an additional 3,500 (17%) to undefined rat ESTs, based on sequence information present in NCBI databases.

Although several major SAGE projects used 14-mer base tags (CATG+10 bp), we found that 15-mer tags (CATG+11 bp) are more reliable. The NCBI rat SAGE tag to gene mapping database, which matches possible 14-mer tags with known rat genes and ESTs, in a number of cases correctly matched the same tag with 3' sequences representing four different transcripts. However, this ambiguity was in almost all cases resolved when the 15th base was added. As one example, four different 15-mer tags [CATGGTGGCACA (T)/(C)/(A)/(G)] were matched by the 14-mer database as ribosomal protein S13. Consideration of the 15th base led to an unambiguous match (CATGGTGTGGCAG) with the published S13 sequence.

Reliability of Quantitative SAGE Analysis of PC12 Cell Transcripts. There are several criteria by which to judge the fidelity of our SAGE data to the status of transcripts in our cells. One is that housekeeping genes should not undergo major NGF regulation. As anticipated, four different EF-1 proteins (α , β , γ , and δ) showed little, if any, differential expression. Several additional examples of such nonresponsive genes are given in Table 3, which is published as supplemental data, including cytoplasmic β actin and alcohol dehydrogenase. Fig. 1 illustrates the distribution of fold-changes in tag number between naïve and NGF-treated cells for tags detected two or more times. This reveals that approximately 90% of the tags show less than a 3-fold difference in expression for the two libraries, which is another indication that most expressed genes are nonresponse to NGF. The 10% of NGF-responsive transcripts detected by our analysis is in line with previous estimates (11, 12).

Another measure of the reliability of our analysis is that it should detect changes in expression of transcripts already known to be NGF responsive. Our data revealed 17 such transcripts previously reported as NGF-responsive in PC12 cells or other

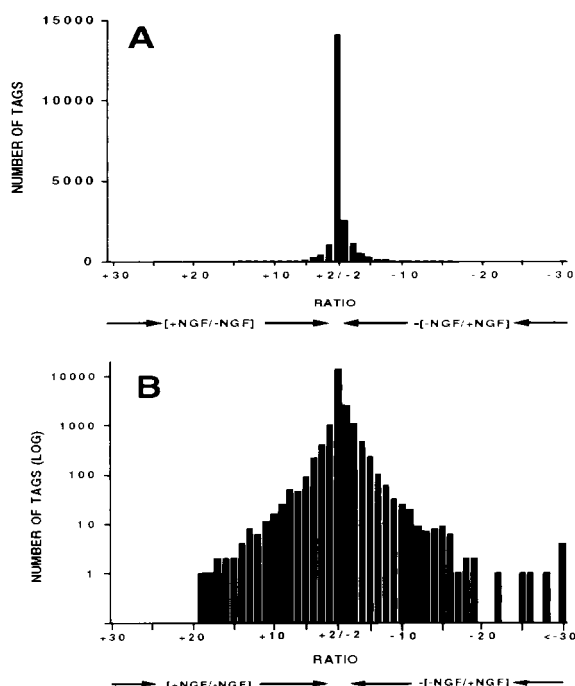


Fig. 1. Distribution of NGF-promoted changes in expression of transcripts in PC12 cells as revealed by SAGE. Transcripts for which two or more tags were detected are included. Up-regulated transcripts are to the left, down-regulated transcripts to the right. Tag numbers are expressed either on a linear (A) or log (B) scale.

systems. In all cases, the changes in expression shown by the SAGE data were consistent with the changes previously reported (see supplemental Table 4).

Gene Regulation by NGF. Approximately 10% of the 21,000 tags detected two or more times changed in abundance by 3-fold or greater in response to NGF. The data in Fig. 1 indicate that NGF-down-regulated transcripts exceed the number that are up-regulated. Fig. 1 further shows that the numbers of regulated transcripts diminished at greater levels of change so that approximately 3.8% (800) changed by 6-fold or greater whereas 149 (0.7%) changed by 10-fold or more. Of tags regulated by 6-fold or more, 150 were assignable (see *Materials and Methods*) to known, named genes and 238 to currently novel, unmatched rat ESTs. Approximately 130 of the identified genes have not been previously reported to our knowledge as targets for NGF regulation; these are listed in Table 1.

To confirm our analysis, a set of regulated transcripts was subjected to Northern blotting using RNA from naïve and NGF-treated cultures independent from those used to establish our SAGE libraries. Fourteen named genes and 13 genes corresponding to ESTs were probed by using β -actin (which SAGE revealed as nonresponsive) as a loading control. In all cases, the blots showed changes consistent with those revealed by SAGE. Several such blots for named genes are shown in Fig. 2. In one case our SAGE data were inconsistent with published findings. It was reported that the transcription factor MASH-1 is up-regulated by NGF in PC12 cells (22). In contrast, our SAGE data indicate a drop in expression (tags +NGF/–NGF, 0:6). This was confirmed by our Northern blot analysis (Fig. 2A) and would be the anticipated response in that MASH-1 influences early specification of autonomic lineage and is not expressed by postmitotic sympathetic neurons (23). Also consistent with this, expression of the transcription factor Phox2a, which appears to be

downstream of MASH-1 in specification of the sympathetic lineage (24), also is down-regulated by NGF in our SAGE analysis (tags +NGF/–NGF, 5:27).

Two distinct NGF receptor types have been recognized; Trk is a receptor tyrosine kinase required for regulation of at least several NGF-stimulated genes (4, 19) and p75 is a tumor necrosis factor α receptor superfamily member that appears capable of evoking death and whose signaling mechanism is at best partially understood (1, 4). To carry out an initial analysis of Trk and p75 involvement in regulation of the genes described here, Northern blots were performed on RNA isolated from PC12nr5 cells expressing only p75 receptors or from Trk-transfected PC12nr5-T14 cells (19). For the six transcripts tested [TSC-22 (Fig. 2B), MAP4, galectin 1 (Fig. 2B), tetraspanin CD9, 14-3-3 γ , and pleiotrophin], regulation by NGF was observed only for the cells expressing Trk. Thus, p75 alone is not sufficient to regulate the NGF-responsive genes we assessed to date, whereas the presence of Trk is necessary for this action.

Table 1 lists approximately 130 transcripts of named genes indicated by SAGE to undergo significant levels of long-term regulation by NGF in PC12 cell cultures. To facilitate evaluation of these changes from a biological perspective, the transcripts have been loosely grouped into functional categories. Although it is beyond the scope of this report to discuss each change in detail, a few points can be made about how these data begin to illuminate our understanding of how NGF-promoted transcriptional changes contribute to neuronal differentiation, plasticity, and function.

Cytoskeleton. Prior studies have drawn attention to the importance of microtubules in NGF-promoted neurite outgrowth. Aside from their structural role, microtubules play a key role in supporting retro- and antero-grade neuronal transport. In this respect, the striking up-regulation of transcripts encoding LIC2, a cytoplasmic dynein light intermediate chain putatively involved in organelle transport (20), is especially interesting.

Dynamic regulation of the actin network, particularly at the growth cone, is another important element of neuronal differentiation and plasticity. SAGE reveals substantial up-regulation of transcripts encoding proteins involved in actin gelation (a transgelin-like protein, ref. 25); actin depolymerization (cofilin, ref. 26); formation and organization of actin filaments (ARP 2/3 34-kDa subunit, ref. 27); linkage of signaling pathways to f-actin reorganization (cortactin B, brain isoform, ref. 28); and alteration of actin organization in dendrites and growth cones (drebrin A, ref. 29). The diversity of the roles of these proteins underscores the complexity of the mechanisms by which NGF regulates actin function and points to several new candidates that may mediate NGF-promoted synaptic plasticity.

Because cofilin regulates neuronal actin polymerization (26) and NGF causes rapid cofilin dephosphorylation and relocalization (21), we carried out Western blotting with anticofilin in extracts of naïve and NGF-treated PC12 cells to determine whether the observed transcriptional changes led to changes in protein expression (see Fig. 3, which is published as supplemental material). This revealed an increase in cofilin protein expression similar to that for cofilin transcripts. In contrast, levels of the closely related actin depolymerizing factor (ADF) protein (26) were unchanged, thus indicating the selectivity of the actions of NGF on this class of molecules.

Regulation of Proliferation. A striking response of PC12 cells to NGF is cessation of proliferation (7). Understanding how this occurs is relevant both to the neuroblast-neuron transition and the regulation of tumor cell growth. In this context, it is of interest that NGF up-regulates transcripts for a putative tumor suppressor (DOC1, ref. 30); for an enzyme that suppresses nucleotide synthesis (phosphoribosylpyrophosphate synthetase-

Table 1. Named transcripts identified by SAGE as NGF-regulate in PC12 cells

Tag	+/- NGF	Identity	Tag	+/- NGF	Identity
Intracellular signaling			Protein/vesicle trafficking (cont.)		
AATGTGAGTCA	14:0	14-3-3 protein γ -subtype	AGAATGAAGTT	6:0	MUM-2
GATTGTCTTGA	12:0	25-kb FK506 binding protein (FKBP25)	CTGTCTGACTC	0:12	β adaptin
AAGCCTTGCTG	9:0	Ash-m protein	ACCTTGCCCTC	3:18	Translocon-associated protein β subunit
TCCTGTGTCCT	7:1	Inositol trisphosphate receptor 3 (IP3R-3)	Granules/synaptic vesicle components		
ATCAACACCCG	100:27	Guanine nucleotide BP G α /Gal α 8	CTAGACACTG	7:1	Secretogranin III (SgIII)
ATAGCTGGGGC	18:5	MAP kinase kinase (MEK1)	TCCTGTGTCCT	7:1	IP3R-3
TGCCCAATAAA	0:15	NIPK-death-induced kinase	Calcium binding proteins		
TTCTGCCTCCA	1:8	Similar to novel Ser/Thr kinase PKL12/MP5K	CCCCCTGGATC	24:3	Calcyclin/S100A6
TCTCACCCACT	0:6	Phosphoprotein HASPP28	TCATCTTTAAC	1:7	Calreticulin
ACGAGCTTTAA	1:6	FK506-binding protein FKBP23	ACGAGCTTTAA	1:6	FK506-binding protein FKBP23
Transcription factors/regulators			Ubiquitination/proteasome components		
CCCTTCACTC	8:1	Drl-associated polypeptide (DRAP1)	GTGCTGGACCT	10:1	Proteasome activator PA28 β
CGAAGTCAGGC	8:1	CIIDBP/HMG 1 protein homolog	AGACGCCTGTG	9:0	26S proteasome-associated PAD1 homolog
TTCCCCACACA	8:1	Transcriptional regulator TSC-22	TTCCTGTGCC	6:0	Anaphase-promoting complex subunit 2
CAAATAAGTTT	7:0	Transcriptional cofactor pirin	ACAAGTATACT	6:1	Proteasome ϵ -chain precursor
AGCCTGTTTTT	6:1	TFIID 30-kDa subunit (TAFII-30)	ATCAGTGGCTT	21:4	Proteasome β -type RN3 subunit
AGAACCTAGTC	5:126	ATFx/ATF5	Intercellular/substrate interactions		
TGCTCCGTGTA	1:9	Supt5hp homolog	CGCGCGGATGG	209:15	β -gal-binding lectin, Galectin 1
GAGAACATCAC	0:6	BTF3	AAATCCTTTCA	11:0	Pleiotrophin
TGATCTTTTTG	2:11	ATF-4/CREB2	GGTGTGCCAGG	10:0	Putative integral membrane receptor
AGAAGCGCAAG	5:27	Arix1/PHOX2A homeodomain protein	ATTCTGTGCTG	10:1	Tetraspanin family member CD9
Proliferation/DNA synthesis			ATCTAAGCCAG	9:1	Galectin 3
ACTGAGTGCTT	11:1	Deleted in oral cancer homolog/DOC1	TCTGTCTGCT	8:0	Degen spermatocyte (des) homolog
AAGGGTCCCG	8:0	PRPP5-associated protein 39	CCCTTCCCTGC	7:0	Cortexin
TTCCTGTGCC	6:0	Anaphase-prom complex subunit 2 (APC2)	CGCGCGCGCGC	7:0	BIT/MFP/SHPS-1/P84
CAAAGTGCATT	0:9	Cyclin-dependent kinase 4 (CDK4)	CTCTGACTTTA	26:4	OX47/basigin/neurothelin/M6/CE9
GGGCAGACAGG	1:9	Replication protein A p32 (RPA32)	AGCTCTACAGG	6:0	G10 protein homolog (edg2)
CGCAAGAAGGT	1:8	DNA polymerase δ catalytic subunit	TGAGCCAGCC	6:0	Ephrin type-B receptor 1 precursor
GAGCATTAGT	1:6	DNA primase small subunit	GAGTGGACTCT	12:2	Tetraspanin AD1/ME491/CD63
RNA processing/splicing/stability			GAGACCCTCAG	6:1	Brain sigma receptor
AAGCTGGTTTA	9:0	snRNP B	GCTACAGGGAG	6:1	Rod outer segment membrane protein 1/ROM1
ATGAGGAACTT	7:0	U1 snRNP C (U1-C)	TACTTGTGTTT	6:1	IgG superfamily member gp65
GTCGCTTCTGA	1:9	Similar to splicing factor U2AF p35	CTACAGTTCTC	0:13	Latrophilin/CIRL/CL1
GAAGAGTGTA	2:15	Pyrimidine-rich tract binding protein	Neuroprotection		
CACAAGTGTGA	0:7	SR-related protein SRp129	GTGGCCCACTT	22:2	Cyt c oxidase assembly protein COX17
GAAATGTAAGA	2:12	Poly(rC) binding protein 2 (PCBP-2)	TATGCACAGGC	8:1	DNA-repair protein XRCC1
TTGATCGAAGT	0:6	Poly(A) polymerase, isoform VI	GTCTAGGTCAC	7:0	Hepatitis B virus X interacting protein
GCCCGAAAGAT	0:6	Suppressor of white apricot homolog 2	GAGAGGCATTT	6:0	Glutathione S-transferase 8/GST8
Protein folding/chaperones			ACCGCCTTAG	11:2	GST-microsomal form
AGCCTCCCTTG	12:1	C terminal of Hsp70-interacting protein	Energy metabolism		
TAGAGCTGCT	11:0	Nucleosome assembly protein 2/NAP2	GCATACGGCGC	18:0	Mitochondrial ATP synthase subunit e
TGGGTTAGACC	8:1	Prefoldin subunit 1	CCAAGGAAAAC	51:8	Lactate dehydrogenase A
AGAAGTTCAGA	15:2	Heat shock protein 60 (HSP60)	ATTAAGTGGT	6:1	Glutamate dehydrogenase
GAGCCTTTTGG	159:37	Cyclophilin A	CGCTGAGGCCT	6:1	Hom 64-kDa ubiquinol-cyt c reductase
Microtubule function related			CTCTGTTTTCC	6:0	Similar to 14-kDa NDUF55
TGGAACCTTGC	16:1	Cytoplasm dynein light intermediate chain/LIC2	CATCCTTGATG	11:2	Brain creatine kinase B
GAGGAGGGGGA	8:0	α -6 tubulin	TTCCAGCTGCT	26:5	Brain phosphoglycerate mutase B
GTATGTGGATC	11:2	Homologous with MAP 1A/B light chain 3	TACTAGAAAAG	1:7	NADH-ubiquinol oxidoreductase NDUF52
Actin cytoskeleton regulation			Miscellaneous metabolism		
AGTTTGCTGAT	38:3	Transgelin-like protein	TGCTCCTGTGA	13:0	β -hexosaminidase α -subunit
AGCTTTCCTGT	9:1	Cortactin B (brain isoform)	GCTGGAATTGA	10:1	Farnesyl pyrophosphate synthetase
TGGGTTAGACC	8:1	Prefoldin subunit 1	GAGAAGAAGGA	6:1	Peroxisomal enoyl-CoA hydrat-like protein
GCAATAAATGG	6:0	Actin binding protein drebrin A	ACCTTGTGAT	2:30	D-dopachrome tautomerase
TGTATAATCAG	6:1	Myosin regulatory light chain 2B	ACCTACAGGAT	1:12	Branched chain α -ketoacid dehydrogenase E1 α
TTGCACCTTCT	6:0	ARP2/3 complex 34-kDa subunit	CTTGTGACAGG	0:9	Monoglyceride lipase
GAAAGCAGACC	44:10	Cofilin	Ion pumps/transporters		
Protein/vesicle trafficking			TTCTAGCATAT	11:0	Na ⁺ , K ⁺ -ATPase β -1 subunit
TCAGGCATTTT	9:0	Ras-related protein RAB-1B	TCCCCTGCTA	6:0	Organic cation transporter BOCT
TGTGAAGTAGC	17:2	ADP-ribosylation factor ARF1	TTGGTGAGGTA	1:16	LAT4 Na-independent neutral aa transporter
TGTGCAGTGAA	17:2	Signal peptidase 12-kDa subunit	ATTCTCTGGAT	1:7	Sarco/endoplasmic reticulum Ca-ATPase
GCACACTGTGT	7:0	Endophilin III/SH3p13	CTGGAGCTGGG	1:8	Mitochondria dicarboxylate carrier
TGGTGACTAAG	7:1	ADP-ribosylation factor BP GGA 2			

associated protein, ref. 31): and for a subunit of a complex with ubiquitin-protein ligase activity that triggers degradation of proteins required for mitosis (anaphase-promoting complex

subunit 2, ref. 32). In addition, there is substantial down-regulation of transcripts encoding the DNA replication protein RPA32 (33), the proliferating cell nuclear antigen-associated

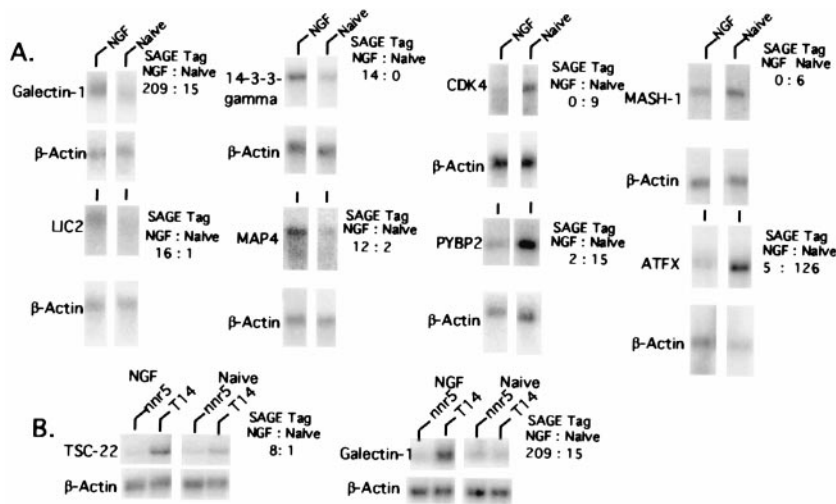


Fig. 2. Northern blot analysis of NGF-promoted gene regulation in (A) PC12 cells and (B) PC12-cell-nnr5 derived cells with (T14) and without (nnr5) Trk expression. Total cellular RNA was isolated from cells treated without (naive) or with NGF for 7–14 days (NGF). Ten micrograms of RNA was analyzed as described in *Materials and Methods* with the indicated probes. Blots were stripped and reprobed with β -actin to indicate relative loading. Ratios to side of blots indicate numbers of tags in PC12 cell SAGE analysis +/- NGF.

catalytic subunit of DNA polymerase delta (34), and the small subunit of DNA primase (35). These observations identify multiple genes whose regulation may participate in generating the neuronal postmitotic state. In addition to roles in mitosis, cell-cycle associated molecules have been implicated in apoptotic neuronal death evoked by loss of support by trophic factors such as NGF (36). Cdk4 and cyclin-dependent kinase p130-PITSLRE are two such examples considered as possible elements in the apoptotic mechanism (36, 37). These are down-regulated by NGF and hence would be expected to increase expression upon NGF deprivation, and thereby potentially contribute to the apoptotic mechanism.

Membrane and Vesicle Trafficking and Formation. Axon generation entails significantly enhanced formation and trafficking of membranes and their components. Within this context, the up-regulation of transcripts for Rab-1B (38), ARF-binding protein GGA2 (39), MUM-2 (40), and the 12-kDa signal peptidase subunit (41) all appear to be functionally significant. NGF-promoted neuronal differentiation in PC12 cells also includes formation of synaptic vesicles (7). The up-regulated transcripts for ARF1 (42) and endophilin III (43) presumably contribute to vesicle genesis and recycling, respectively, whereas those for secretogranin III (44) and IP3R-3 (45) would provide vesicular components.

Transcription Factors. A number of NGF-regulated genes were found to encode transcriptional regulators, which suggests that such factors themselves may play important roles in mediating NGF's actions on gene expression. Among those that are up-regulated, TSC-22 may be especially notable in that this evolutionarily conserved transcriptional repressor also is induced by other growth inhibitory molecules including transforming growth factor type β (46). TSC-22 is expressed in many developing mammalian tissues including sympathetic neurons (47) and in the fly, its homologue shortsighted/bunched, plays a required role in development of the nervous system (48). The up-regulated CIIDBP, a high mobility group I protein (HMGI) homologue, is also a potentially significant mediator of NGF actions in that HMGI proteins appear to play important roles in coordinating proliferation and differentiation during development (49). Of those transcription factors that are down-

regulated, ATFX/ATF5 is reported to suppress cAMP-regulated gene expression (50), indicating that NGF may effectively increase neuronal responsiveness to cAMP and other effectors that use CREB-dependent signaling.

14-3-3 Family. Aside from revealing transcript levels for diverse proteins involved in a common function, SAGE provides information about relative expression and regulation of transcripts within protein families. One example is the 14-3-3 family of highly homologous proteins that bind to and affect the distribution, availability, and/or activity of other proteins, thereby regulating a variety of enzymatic activities and signaling pathways (51). The differences in regulation and activities of the seven major mammalian 14-3-3 isoforms are not well understood. Six isoforms were detected in PC12 cells by SAGE (see supplemental Table 5); all are highly expressed in brain (52, 53). Transcripts for the γ form were highly up-regulated in response to NGF (14:0); η , the most abundant form, was up-regulated by about 3-fold as was the less abundant ζ isoform. In contrast, the β and ϵ isoforms underwent little or no regulation whereas the θ form was significantly down-regulated. These observations indicate that even closely related isoforms are differentially regulated by NGF and that they may play specific roles in the NGF response.

NGF and Glutathione S-Transferases (GSTs). Protection of neurons from oxidative and toxic stresses is an important action of NGF (54). GST, which comprises a multigene family as well as microsomal and mitochondrial forms, is among the major enzymes involved in such protection, and it is reported that NGF up-regulates GST activity in brain (55). SAGE reveals seven distinct GST transcripts in PC12 cells (see supplemental Table 6) of which three (GSTb, GST YB3, and microsomal GST) appear to be significantly up-regulated. These observations not only illuminate a potential mechanism by which NGF confers neuroprotection, but also illustrate selective actions of the factor on a group of genetically and functionally related molecules.

Concluding Remarks. Our aims in undertaking the present analysis have been several-fold. One has been to generate a database of the genes expressed in naive and neuronally differentiated PC12 cells. This cell line has been and continues to be an extensively used model for a variety of aspects of neuroendocrine/neuronal

behavior and function, and such information should enhance its utility in this regard. The similarity of PC12 cells to chromaffin cells and sympathetic neurons also should provide useful information regarding genes potentially expressed in these and other neuronal/neuroendocrine cell types. Approximately 2,400 transcripts for named genes have been identified in the cells to date and this number should grow significantly as additional genomic information enters public databases. A second aim has been to provide a comprehensive analysis of genes that undergo long-term regulation in response to NGF and thereby to gain insight about the molecular basis of neurotrophic factor actions. Aside from confirming known NGF responsive genes, we identified approximately 130 additional named genes that appear to undergo significant levels of regulation by NGF. Even this fragmentary list of regulated, known genes has begun to yield insights about the NGF mechanism and it can be anticipated that its size and impact will grow as a consequence of advancing

genomic information. The next steps will be to test the functional consequences of manipulating the expression of these genes in neuronal systems and to determine temporal and causal relationships between the regulated genes. Further analyses also will be required to identify genes whose regulation by NGF is transient in nature. A final goal has been to provide information for discovery of new neuronal genes, particularly those subject to NGF regulation. The several hundred currently unmatched ESTs and roughly 500 unmatched tags regulated by more than 6-fold provide fertile ground in this respect.

We thank Claudine L. Bitel for outstanding technical assistance and Drs. R. Vallee, R. Dalla-Favera, P. Cserjesi, D. Park, and J. Bamburg for providing critical reagents. We are also grateful to Drs. A. Chalazonitis, N. Papadopoulos, R. Parsons, K. Kinzler, V. Velculescu, and L. Zhang for helpful advice. This work is supported in part by grants from the National Institute of Neurological Disorders and Stroke (L.A.G. and M.L.S.).

- Frade, J. M. & Barde, Y. A. (1998) *BioEssays* **20**, 137–145.
- Tessarollo, L. (1998) *Cytokine Growth Factor Rev.* **9**, 125–137.
- Segal, R. A. & Greenberg, M. E. (1996) *Annu. Rev. Neurosci.* **19**, 463–489.
- Friedman, W. J. & Greene, L. A. (1999) *Exp. Cell Res.* **253**, 131–142.
- Burstein, D. E. & Greene, L. A. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 6059–6063.
- Greene, L. A. (1984) *Trends Neurosci.* **7**, 91–94.
- Greene, L. A. & Tischler, A. S. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2424–2428.
- Greene, L. A., Farinelli, S. E., Cunningham, M. E. & Park, D. S. (1998) in *Culturing Nerve Cells*, eds. Banker, G. & Goslin, K. (MIT Press, Cambridge, MA), 2nd Ed., pp. 161–187.
- Greenberg, M. E., Greene, L. A. & Ziff, E. B. (1985) *J. Biol. Chem.* **260**, 14101–14110.
- Leonard, D. G., Ziff, E. B. & Greene, L. A. (1987) *Mol. Cell. Biol.* **7**, 3156–3167.
- Lee, N. H., Weinstock, K. G., Kirkness, E. F., Earle-Hughes, J. A., Fuldner, R. A., Marmaros, S., Glodek, A., Gocayne, J. D., Adams, M. D., Kerlavage, A. R., et al. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 8303–8307.
- Mayumi, K., Yaoi, T., Kawai, J., Kojima, S., Watanabe, S. & Suzuki, H. (1998) *Biochim. Biophys. Acta* **1399**, 10–18.
- Brown, A. J., Hutchings, C., Burke, J. F. & Mayne, L. V. (1999) *Mol. Cell. Neurosci.* **13**, 119–130.
- Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. (1995) *Science* **270**, 484–487.
- Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E., Jr., Hieter, P., Vogelstein, B. & Kinzler, K. W. (1997) *Cell* **88**, 243–251.
- Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Vogelstein, B. & Kinzler, K. W. (1997) *Science* **276**, 1268–1272.
- Velculescu, V. E., Madden, S. L., Zhang, L., Lash, A. E., Yu, J., Rago, C., Lal, A., Wang, C. J., Beaudry, G. A., Ciriello, B. P., et al. (1999) *Nat. Genet.* **23**, 387–388.
- Angelastro, J. M., Klimaschewski, L. & Vitolo, O. V. (2000) *Nucleic Acids Res.* **28**, e62.
- Loeb, D. M. & Greene, L. A. (1993) *J. Neurosci.* **13**, 2919–2929.
- Hughes, S. M., Vaughan, K. T., Herskovits, J. S. & Vallee, R. B. (1995) *J. Cell Sci.* **108**, 17–24.
- Meberg, P. J., Ono, S., Minamide, L. S., Takahashi, M. & Bamburg, J. R. (1998) *Cell Motil. Cytoskeleton* **39**, 172–190.
- Johnson, J. E., Birren, S. J. & Anderson, D. J. (1990) *Nature (London)* **346**, 858–861.
- Lo, L. C., Johnson, J. E., Wuenschell, C. W., Saito, T. & Anderson, D. J. (1991) *Genes Dev.* **5**, 1524–1537.
- Hirsch, M. R., Tiveron, M. C., Guillemot, F., Brunet, J. F. & Goridis, C. (1998) *Development (Cambridge, U.K.)* **125**, 599–608.
- Prinjha, R. K., Shapland, C. E., Hsuan, J. J., Totty, N. F., Mason, I. J. & Lawson, D. (1994) *Cell Motil. Cytoskeleton* **28**, 243–255.
- Bamburg, J. R. (1999) *Annu. Rev. Cell Dev. Biol.* **15**, 185–230.
- Welch, M. D. (1999) *Trends Cell Biol.* **9**, 423–427.
- Ohoka, Y. & Takai, Y. (1998) *Genes Cells* **3**, 603–612.
- Shirao, T. (1995) *J. Biochem. (Tokyo)* **117**, 231–236.
- Tsuji, T., Duh, F. M., Latif, F., Popescu, N. C., Zimonjic, D. B., McBride, J., Matsuo, K., Ohyama, H., Todd, R., Nagata, E., et al. (1998) *J. Biol. Chem.* **273**, 6704–6709.
- Kita, K., Ishizuka, T., Ishijima, S., Sonoda, T. & Tatibana, M. (1994) *J. Biol. Chem.* **269**, 8334–8340.
- Zachariae, W. (1999) *Curr. Opin. Cell Biol.* **11**, 708–716.
- Iftode, C., Daniely, Y. & Borowiec, J. A. (1999) *Crit. Rev. Biochem. Mol. Biol.* **34**, 141–180.
- Yuzhakov, A., Kelman, Z., Hurwitz, J. & O'Donnell, M. (1999) *EMBO J.* **18**, 6189–6199.
- Foiani, M., Lucchini, G. & Plevani, P. (1997) *Trends Biochem. Sci.* **22**, 424–427.
- Park, D. S., Levine, B., Ferrari, G. & Greene, L. A. (1997) *J. Neurosci.* **17**, 8975–8983.
- Lahti, J. M., Xiang, J. & Kidd, V. J. (1995) *Prog. Cell Cycle Res.* **1**, 329–338.
- Dugan, J. M., deWit, C., McConlogue, L. & Maltese, W. A. (1995) *J. Biol. Chem.* **270**, 10982–10989.
- Hirst, J., Lui, W. W. Y., Bright, N. A., Totty, N., Seaman, M. N. J. & Robinson, M. S. (2000) *J. Cell Biol.* **149**, 67–80.
- Chiari, R., Foury, F., De Plaen, E., Baurain, J. F., Thonnard, J. & Coulie, P. G. (1999) *Cancer Res.* **59**, 5785–5792.
- Kalies, K. U. & Hartmann, E. (1996) *J. Biol. Chem.* **271**, 3925–3929.
- Faundez, V., Horng, J. T. & Kelly, R. B. (1997) *J. Cell Biol.* **138**, 505–515.
- Ringstad, N., Gad, H., Low, P., Di Paolo, G., Brodin, L., Shupliakov, O. & De Camilli, P. (1999) *Neuron* **24**, 143–154.
- Dopazo, A., Lovenberg, T. W., Danielson, P. E., Ottiger, H. P. & Sutcliffe, J. G. (1993) *J. Mol. Neurosci.* **4**, 225–233.
- Blondel, O., Takeda, J., Janssen, H., Seino, S. & Bell, G. I. (1993) *J. Biol. Chem.* **268**, 11356–11363.
- Shibanuma, M., Kuroki, T. & Nose, K. (1992) *J. Biol. Chem.* **267**, 10219–10224.
- Dohrmann, C. E., Belaousoff, M. & Rafferty, L. A. (1999) *Mech. Dev.* **84**, 147–151.
- Treisman, J. E., Lai, Z. C. & Rubin, G. M. (1995) *Development (Cambridge, U.K.)* **121**, 2835–2845.
- Zhou, X. & Chada, K. (1998) *Keio J. Med.* **47**, 73–77.
- Nishizawa, M. & Nagata, S. (1992) *FEBS Lett.* **299**, 36–38.
- Skoulakis, E. M. & Davis, R. L. (1998) *Mol. Neurobiol.* **16**, 269–284.
- Watanabe, M., Isobe, T., Ichimura, T., Kuwano, R., Takahashi, Y. & Kondo, H. (1993) *Brain Res. Mol. Brain Res.* **17**, 135–146.
- Martin, H., Rostas, J., Patel, Y. & Aitken, A. (1994) *J. Neurochem.* **63**, 2259–2265.
- Semkova, I. & Kriegstein, J. (1999) *Brain Res. Rev.* **30**, 176–188.
- Guegan, C., Ceballos-Picot, I., Chevalier, E., Nicole, A., Onteniente, B. & Sola, B. (1999) *Neurobiol. Dis.* **6**, 180–189.