Invertebrate Neuroscience, 1, 341-349 (1996)

Expression profiling of mRNA obtained from single identified crustacean motor neurons: determination of specificity of hybridization

ROMAN PEKHLETSKI, ROBIN L. COOPER[1], HAROLD L. ATWOOD[1] and DAVID R. HAMPSON

Faculty of Pharmacy and Department of Physiology[1], MRC Group in Nerve Cells and Synapses, University of Toronto, Toronto, Ontario, Canada

ABSTRACT The purpose of this study was to determine if the technique of expression profiling would allow us to determine the changes in the abundance of certain mRNAs in identifiable, single neurons as a result of heightened electrical activity. In doing so we developed an approach to test the specificity of hybridization in expression profiling. Messenger RNA from single identified crayfish motor neurons was amplified by ligation-mediated reverse transcription PCR and hybridized by dot-blotting to 45 target cDNAs from different species. As a test of specificity, the hybridization was repeated using unlabelled cDNAs, the dots were excised, and the hybridized nucleic acids were re-amplified, cloned, and sequenced to confirm their identity. By cloning and sequencing the re-amplified product for each cDNA examined, one can determine the degree of background hybridization as compared to homologous hybridization. False positive results were also observed when a species-specific cDNA and highly stringent hybridization conditions were used. Our results demonstrate that ligation-mediated PCR is a useful technique for checking the specificity of expression profiling. This approach can easily be adapted to any situation when confirmation of the specificity of nucleic acid hybridization is required. During this study, part of a novel crayfish neuronal actin cDNA was cloned and sequenced.

KEY WORDS: mRNA; crustacean; cross-hybridization; expression profiling

Introduction

Expression profiling or reverse hybridization is a technique that is used to detect differences in mRNA levels between cell types or in a given cell type that has been experimentally altered (Van Gelder *et al.*, 1990). This procedure is based on the hybridization of a probe synthesized from the mRNA of a single cell, to target cDNAs which are fixed to a nucleic acid transfer membrane. This technique has been used as a tool for monitoring differences in the levels of expression of known mRNAs in neurons that have been subjected to increased levels of electrical stimulation (Mackler *et al.*, 1992; Haisenleder *et al.*, 1993). In addition, this approach allows one to detect, in a single cell, differences in the levels of expression of up to a few hundred previously characterized mRNAs in a given experiment. Also of interest, is the application of this rapid dot-blotting procedure to screen for the regulation of potentially interesting mRNAs, under various experimental conditions, in species different from those from which the known cDNAs are derived (Kravitz *et al.*, 1992).

However this hybridization-based method is plagued by several potential artifacts. These artifacts stem from the presence of non-specific cross-hybridizing sequences (NCSs) in both the probe and target nucleic acids. These NCSs often belong to one or more of the following classes: (a) sequences with extremely rich GC content (Varshney *et al.*, 1988); (b) functional motifs which are relatively short, evolutionarily conserved, and thus widely distributed homologous sequences (Karlin and Altschul, 1993); and (c) repeated sequences which are commonly present in many mRNAs (Skinner, 1977; Browx, 1983; Claverie and Makalowski, 1994). Repetitive DNA sequences are particularly abundant in crustaceans (Fowler *et al.*, 1985; Cruces *et al.*, 1986). The presence of NCSs in nucleic acids could lead to stable hybridization between two phylogenetically distant nucleic acids. In order to estimate the degree of such non-specific background hybridization in reverse hybridization experiments, we have modified the original approach by using ligation-mediated reverse transcription PCR instead of the mRNA amplification procedure originally used by Van Gelder *et al.* (1990).

In our study, we used the amplified and labelled cDNAs derived from a large identified motor neuron

Corresponding author: Dr Roman Pekhletski
E-mail: roman.pekhletski@utoronto.ca

of crayfish as probes and 45 previously characterized known cDNAs for screening. Of these 45 cDNAs initially analyzed, five were picked for further study. The nucleic acids hybridizing to these five cDNAs were re-amplified by PCR and cloned. After screening and identification of positive colonies based upon hybridization to the five cDNAs used as probes, the primary sequences of hybridized molecules were determined.

A preliminary report of this work has been presented previously in abstract form (Pekhletsky *et al.*, 1994).

Materials and methods

Motoneuron identification and *in situ* first strand cDNA synthesis

The somata from identifiable single neurons were isolated from the second abdominal ganglion of the ventral nerve cord from specimens of *Procambarus clarkii*, measuring 5–6 cm in body length (Atchafalaya Biological Supply Co., Raceland, LA). The identified neuron was determined to be the common excitor motor neuron which innervates the deep abdominal extensor muscles (Parnas and Atwood, 1966; Wine and Hagiwara, 1977). The somata of the five phasic excitatory motor neurons exist in a tight grouping close to the formation of the second root. In order to identify the common excitor for physiological recording, the ventral nerve cord from abdominal segments 1 to 4 was removed and pinned out in a Sylgard coated dish. This bathing medium was exchanged twice with an RNAse free crayfish saline (Cooper *et al.*, 1995). The second abdominal ganglion was gently de-sheathed and the exposed ganglion was treated with collagenase/dispase (Sigma, 2 mg/mL) for 15 minutes. In the de-sheathed ganglion, the somata retain their characteristic location as in the ganglion (Otsuka *et al.*, 1967; Wine and Hagiwara, 1977; Wine and Krasne, 1982). The soma of the common excitatory motoneuron is a large cell (diameter 80–120 μm) in the ganglion, and occurs in a characteristic location. With rapid intracellular recording (60–90 mΩ electrode resistance), identification of the common excitatory soma can be made. This provided an additional confirmation of its identity (Wine and Krasne, 1982). After physiological identification was made and location carefully recorded, the first stage of the first strand cDNA synthesis (cDNA synthesis *in situ*) was made according to Mackler *et al.* (1992), modified by using 2.3 pmoles of random hexanucleotides instead of T7-oligo-$(dT)_{24}$ to prime reverse transcription. This step was followed by aspiration of the cell cytoplasm into an electrode (4 to 10 μm diameter) and further synthesis of single-stranded cDNA (Mackler *et al.*, 1992), again using 2.3 pM of random hexanucleotides to prime first strand synthesis.

cDNA synthesis and ligation of adaptor primers

Double-stranded cDNA (ds-cDNA) synthesis was carried out as described by Mackler *et al.* (1992). The ds-cDNAs were treated with the Klenow fragment of *E. coli* DNA polymerase I (Pharmacia) in the presence of all four deoxynucleotide triphosphates (dNTPs). This reaction was followed by ligation-mediated single-sided PCR (Mueller and Wold, 1989; Pekhletsky *et al.*, 1992). Pre-annealed primers A (5′-GATCCGTGAAGCTTGCGTACTAGTT-3′) and B (5′-AACTAGTACGCAAGCTTCACG-3′) were used as the adaptor complex, which was ligated to cDNA fragments. Primer A was ligated to 5′ ends of the cDNA strands due to the fact that only the 5′ ends of the cDNAs had 5′ phosphates (see Fig. 1). A large molar excess of pre-annealed adaptor complex was used to minimize ligation between cDNA fragments. The DNA was precipitated twice with isopropanol in 0.3 M sodium acetate in order to remove salts and unincorporated oligonucleotides. The filling in of the single

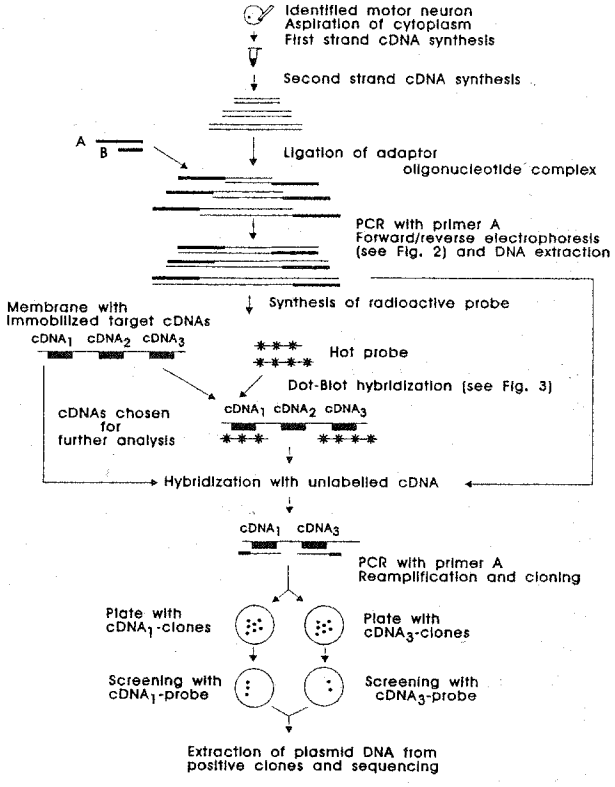

Fig. 1. Schematic representation of the experimental approach used to determine the specificity of hybridization in expression profiling. *Abbreviations: A and B are oligonucleotides A and B (see text); cDNA1, cDNA2, and cDNA3 are different target cDNAs.*

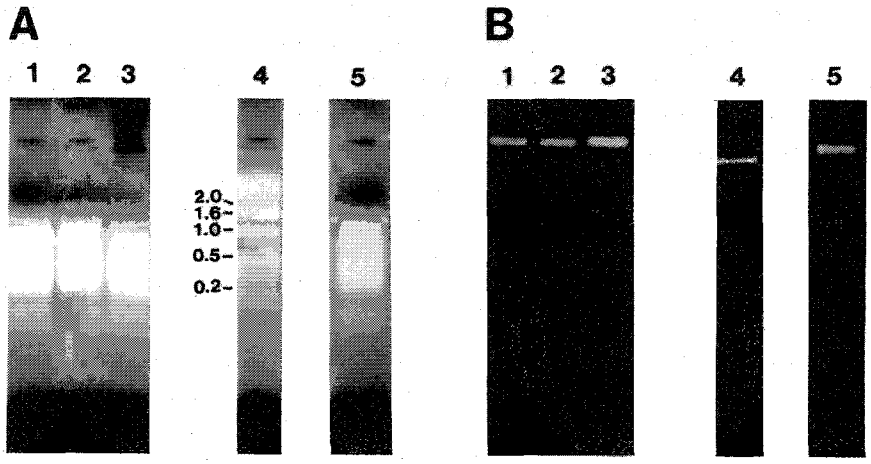

Fig. 2. Results of the separation of the products of ligation-mediated RT-PCR by forward (A) and reverse (B) agarose electrophoresis. *Lanes 1, 2, 3 and 5 are amplified products from four identified homologous motoneurons from different individual crayfish specimens. Lane 4 contains 0.8 μg of the molecular weight markers (1 kb ladder, Gibco/BRL).*

stranded ends generated by the ligation of primer A was accomplished during ramping of the thermocycler to 95 °C prior to the first cycle of the subsequent PCR. Thus, primer A was the only primer needed for amplification.

Polymerase chain reaction

The PCR was carried out in 50 μL containing Taq DNA polymerase buffer (1.5 mM $MgCl_2$; Pharmacia), 80 pmol of primer A, all four dNTPs at a concentration of 250 μM each and 2 units of Taq DNA polymerase (Pharmacia). Amplification was carried out in a Programmable Thermal Controller (PTC-100, Model 60; MJ Research, Inc.) for 35 cycles through 95 °C for 40 s, 68 °C for 30 s, and 72 °C for 2.5 min. After the 35th cycle, a final extension step (5 min. at 72 °C) was carried out to complete the PCR.

Forward–reverse agarose electrophoresis and cDNA extraction

Amplified cDNA fragments of the appropriate size were recovered in 1% low melting point (LMP) agarose (Electrophoresis Grade, BRL) as follows: after 40 min. of electrophoresis, the bottom part of the gel was removed to exclude the smallest PCR products (less than 200 bp) and unincorporated primers (Fig. 2). To concentrate the amplified material in a small piece of the LMP agarose and obtain an estimate of the amount of cDNA produced, the direction of electrophoresis was reversed. After 45 min. of the reverse electrophoresis, the cDNA was cut out and extracted by standard procedures (Sambrook *et al.*, 1989).

Probe synthesis and dot-blot hybridization

Radioactive probes were synthesized using a Random Primer DNA Labelling kit (Boehringer Mannheim Biochemica). 25 ng of DNA, extracted from LMP agarose, were used as a template for preparation of each probe. Radioactive probes were separated from unincorporated [α-^{32}P]dCTP (Amersham) by gel-filtration through NICK Columns (Pharmacia). 100 ng of the corresponding cDNA preparation, denatured by boiling, were used as an unlabelled probe. Forty-five target cDNAs (60–100 ng) were immobilized on Hybond membrane (Amersham) by UV fixation. Hybridization was carried out at 54 °C for 12–18 hours in 20 mL of hybridization solution, containing 1 mM EGTA, 0.55 M Na^+ (included as part of the sodium–phosphate buffer; pH 7.2, 20 °C), 7.4% SDS with gentle shaking (55 rpm). Hybridization was followed by five washes in 100 mL of washing solution under high stringency conditions (60 mM Na^+, 1 mM EDTA) or at low stringency conditions (120 mM Na^+, 1 mM EDTA) at 50 °C with orbital shaking (50 rpm).

Re-amplification of hybridized cDNA fragments

The dot-blot hybridization was repeated under the same conditions as stated above, but with an unlabelled DNA. A small area of the membrane, which contained the corresponding immobilized cDNA and the hybridized material, was excised and used as the template for PCR amplification under the conditions described above.

Post-PCR treatment and cloning

PCR products were analyzed by agarose electrophoresis, extracted from corresponding slices of LMP agarose, digested with *Hin*dIII and ligated into pBlueScript KS(+) (StrataGene).

Determination and analysis of nucleotide sequences

DNA sequences were determined using the Sequenase 7-deaza-dGTP Sequencing kit (US Biochemical). The BLASTN program of the BLAST E-Mail Server (Altschul *et al.*, 1990) supplied by the National Institutes of Health, USA, was used to search for homologous sequences.

Results

The experimental strategy is schematically outlined in Fig. 1. The use of ligation-mediated RT-PCR allows re-amplification of any cDNA fragment hybridized to a target cDNA immobilized on a membrane. Furthermore, the amplification of the immobilized target cDNA is excluded because it does not possess the adaptor oligonucleotide A ligated to the ends of the cDNAs.

Samples subjected to forward and reverse agarose electrophoresis are shown on Fig. 2. The sizes of amplified cDNAs were between 0.2 and 2 kb. This size range is optimal for both random probe synthesis and re-amplification of the hybridized cDNA fragments. The total amount of PCR product was estimated to be approximately 0.8 μg from the intensity of the 1 kb DNA ladder in lane 3 (see Fig. 2B). DNA from lane 2 was extracted after reverse electrophoresis for radioactive probe synthesis.

The results of dot-blot hybridization to the 45 immobilized target cDNAs with the radiolabelled cDNA mixture derived from crayfish neuron are shown in Fig. 3. Some of the cDNAs did not produce any visible signal, while others produced very strong signals. Only one of the target cDNAs (crayfish muscle actin) was species-specific to the probe. This actin cDNA did not produce a strong hybridization signal, even under relatively gentle washing conditions (120 mM Na^+, 50 °C). Based on this observation, we concluded that using a stronger washing condition (60 mM Na^+ at 50 °C) (see Fig. 3B) was sufficiently stringent for conducting a second hybridization with the unlabelled cDNAs.

Five signals, representing signal intensities ranging from very weak to very strong, were selected for assessment of hybridization specificity. The five selected cDNAs included: (1) crayfish muscle actin (accession number D14612); (2) *Drosophila melanogaster* vesicle associated membrane protein VAMP (L14270); (3) 1.75 kb genomic fragment (exon 5 to exon 13) of *Drosophila melanogaster* dunce gene (M14977); (4) rat metabotropic glutamate receptor mGluR4a (M90518); (5) rat metabotropic glutamate receptor mGluR1α (M61099). The hybridization for these five cDNAs was repeated under the conditions stated above, but using the unlabelled probe (see Materials and methods).

Dots containing hybridized nucleic acids from each of the five cDNAs were excised, and the hybridized DNA was re-amplified by PCR using oligonucleotide A as the primer, digested with *Hin*dIII, and ligated into pBlueScript KS(+). *E. coli* colonies were transferred to Hybond membrane and screened by hybridization with ^{32}P-labelled probes synthesized by random priming of the template of the corresponding target cDNAs. Inserts of the clones producing positive signals were sequenced.

The sequencing data are summarized in Table 1. From two to five positive colonies were analyzed for each of the five hybridizations. Sequence comparisons revealed that, except for CA repeats which were found in the dunce and mGluR1α clones, none of the sequences derived from the five target cDNA groups overlapped with sequences from the same group. In addition, a search of the non-redundant nucleotide sequence database revealed that, with the exception of clone 1A, all of the sequenced DNA fragments were unrelated to the corresponding target cDNAs. These observations indicated that multiple species of nucleic acids hybridized to a given cDNA in each case.

We found that a number of sequences contained one of the following: (a) AC-rich domains; (b) AC repeats; or (c) tandem repeated sequences of 12–82 bp. Based on the results from the alignment of each cloned DNA fragment with the corresponding target cDNA, we conclude that there is sufficient homology to participate in the production of a positive dot-blot signal for each sequenced DNA fragment. However in each case (with the exception of the alignment of clone 1A (Table 1) with crayfish muscle actin), the homologous region did not extend through the entire length of the cloned DNA fragment. In most cases, a short homologous region containing 30 to 200 bases was flanked

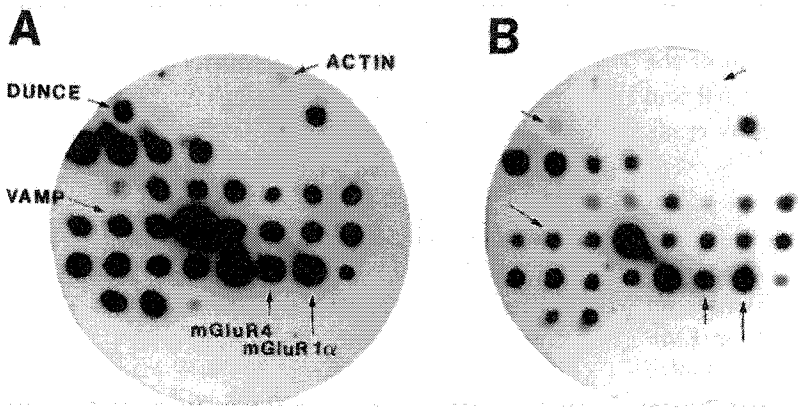

Fig. 3. Expression profile of amplified cDNAs hybridized to 45 different target cDNAs under low (A) and high (B) stringency conditions. *The five target cDNAs which were utilized to determine the hybridization specificity are pointed by arrows.*

by non-homologous sequences.

Two observations indicated that our approach to examine the specificity of hybridization in expression profiling was working as anticipated. The first observation concerns the sequences of the flanking regions of the cloned fragments. DNA fragments cloned using the strategy shown in Fig. 1 should be flanked on both ends by part of the adapter oligonucleotide A, including the region extending from the *Hin*dIII site to the 3′ end of the adaptor (*Hin*dIII-3′ fragment). It should be noted however, that the presence of an internal *Hin*dIII site within a DNA fragment would result in the loss of the *Hin*dIII-3′ fragment on one end of the insert. Except for clone 1A, which was the novel neuronal actin, all of the sequenced inserts were flanked on both ends by the *Hin*dIII-3′ fragment of adapter oligonucleotide A. Clone 1A was flanked by the *Hin*dIII-3′ fragment of oligonucleotide A at one end and by a *Hin*dIII site at the other end. This *Hin*dIII site is located at the same position in the crayfish muscle actin cDNA. This cDNA was a species-specific target cDNA for the dot-blot hybridization.

The second finding which indicates that our strategy was working properly, was the successful identification of the novel homolog of the crayfish muscle actin (Accession Z54232, see Table 1). Since the protocol described here includes a number of stages which involve DNA polymerization (reverse transcription, PCR) and modification (restriction, ligation), there are many opportunities for the introduction of artifacts. The fact that we were able to identify a true neuronal homolog of crayfish muscle actin (Kang and Naya, 1993) confirms that all these stages were working as expected.

Discussion

The use of expression profiling by dot-blotting known cDNAs and hybridizing with amplified, labelled mRNA extracted from whole tissue, or more currently single cells, is an approach that has become increasingly utilized. In addition, the use of cross-species cDNAs in such procedures is becoming more prevalent, as in lobster single identified neurons (Kravitz *et al.*, 1992).

We have also used this procedure on single identified neurons that have well characterized physiological and anatomical changes due to heightened electrical activity (Lnenicka and Atwood, 1985a, b). The physiological changes have been shown to require protein synthesis and axonal transport (Nguyen and Atwood, 1990). This evidence may indicate that increased electrical activity alters mRNA levels. In order to determine if our results were reliable we stopped to question, 'How can we test the specificity of the hybridization to the amplified mRNAs?'.

In this report, we present a rapid and accurate procedure to determine the sequences of nucleic acids which hybridize to immobilized target cDNAs in single cell expression profiling experiments. In order to determine true hybridization from non-specific (background) hybridization in expression profiling, it is essential to sequence the material hybridizing to the target cDNA. The approach outlined here can be easily adapted to any situation when information on the sequence of hybridizing material is required. Our data have indicated that, among each of the five cDNAs examined, the majority of the signals produced from the amplified mRNA were due to non-specific cross-hybridization. The results precluded us from utilizing expression profiling to determine if certain mRNAs were up- or down-regulated in transcription in identifiable single neurons as a result of heightened electrical activity.

An earlier report (Tsykin *et al.*, 1990) indicated that purification of RNA followed by Northern blot analysis was more likely to provide reliable estimates of mRNA expression levels than dot-blot hybridization of cytoplasmic extracts because contaminants are removed. Northern blots of purified RNA are only possible when relatively large amounts of mRNA are available and are therefore unsuitable for determination of expression levels in a single cell (White and Bancroft, 1982; Van Gelder *et al.*, 1990). A number of studies have appeared which make use of expression profiling at the single cell level (Moro *et al.*, 1990; Eberwine *et al.*, 1992; Kravitz *et al.*, 1992; Surmeier *et al.*, 1992; Bargas *et al.*, 1994; Miyashiro *et al.*, 1994). We are not aware that any of these reports have unequivocally confirmed the identity of the hybridized nucleic acids to the target cDNAs.

With the use of ligation-mediated RT-PCR, we have shown that PCR amplification followed by cloning and sequencing of the hybridized DNAs is possible. By comparison, the use of traditional PCR with two primers which are specific to a target cDNA (Mackler and Eberwine, 1993) would result in amplification of only the fragments which are truly homologous to the target DNA and not other fragments present in the mixture but contributing to the signal intensity. Thus, the use of specific-primer PCR would lead to the amplification of only those fragments which are truly specific to the tested cDNA, even if the ratio of these truly specific sequences to the NCSs is very low. Furthermore, the use of specific primers can result in the amplification of target cDNA that is

Table 1. Sequence analysis of DNA fragments cloned from the 5 different dot-blot signals.[a]

Clones	Length bp[b]	Sequenced bp[c]	Most homologous % in bp[d]	Homology to target cDNA % in bp[e]	Type of sequence	Accession #
ACTIN						
1A	185	185	dbj D14612. Crayfish mRNA for muscle actin. 82 in 185.	82 in 185	cDNA	emb Z54232
3A	~750	88 T7	gb M12658. Mouse 4.5s RNA gene. 73 in 41.	59.2 in 76	46 bp SR[f]	emb Z49265
5A	235	235	gb M55089. *Rhodobacter capsulatus* hydrogen oxidation gene. 78 in 37.	68 in 47	–	emb Z54263
mGluR4a						
1R	~500	147 T7 83 T3	emb X80760. *Coturnix coturnix* clusterin gene. 71 in 146.	60.6 in 104	32 bp SR	emb Z54250
2R	~370	129 T7	emb X59140. Human mini-satellite MS630 5′-flanking region. 71 in 105.	71 in 69	–	emb Z54273
3R	~280	106 T7	gb L36092. Human germline T-cell receptor b-chain gene. 67 in 58.	66.3 in 104	12 bp SR	emb Z54259
4R	~420	113 T7	gb M11160. Mouse Thy-1.2 gene. 71 in 92.	63.2 in 57	SR-like[g]	emb Z54262
mGluR1α						
1α	259	259	emb X81325. Human H12A gene. 72 in 154.	56.1 in 214	CA rich[i] SR-like	emb Z54249
2α	~330	181 T3	gb M97546. Rat satellite repeats. 14.65 in 55.	56.3 in 96 or 67.6 in 37	CA rich	emb Z54254
		143 T7	emb Z28202 Yeast chromosome XI ORF YKR106w. 63 in 122.	94.4 in 18 or 62.8 in 86	CA rich	emb Z54255
4α	~450	236 T3	gb L36914. Pig microsatellite (CA repeat). 93 in 60.	76.7 in 43	CA repeat[h]	emb Z54261
		211 T7	gb L36911. Pig microsatellite DNA (CA repeat). 93 in 60.	75 in 50	CA repeat	
5α	~400	220 T3	gb L16861. Human (clone cos 146BIO) sequence tandem repeat region. 88 in 52.	78.4 in 37	CA repeat	emb Z54264
		175 T7	emb Z28297. Yeast chromosome XI reading frame ORF YKR072c. 94 in 53.	75.3 in 38	CA repeat	emb Z54265
6α	~250	203 T7	gb L36911. Pig microsatellite DNA (CA repeat). 74 in 95.	71.8 in 39	CA repeat	emb Z54267

VAMP						
1V	~360	155 T3	gb M32515. Rat simple sequence DNA, clone 8. 77 in 35.	57.3 in 117	82 bp SR	emb Z54251
		147 T7	dbj D15553. Rice cDNA partial sequence (C0814A). 60 in 85.	72.4 in 29 or 62 in 71	82 bp SR	emb Z54252
2V	~30	182 T3	gb L39064. Human interleukin 9 receptor (IL9R) gene. 50 in 294.	54.6 in 108 or 67.9 in 28	28 bp SR	emb Z54256
		335 T7	gb U08979. *Plasmodium vivax* isolate CH-5 circumsporozoite gene, partial cDNA. 55 in 112.	59.3 in 118 or 61.5 in 39	SR-like	emb Z54257
DUNCE						
1D	~450	67T3	gb U13369. Human ribosomal DNA complete repeating unit. 65 in 72.	186.7 in 15 or 100 in 8	23 bp SR	
		75T7	same as above	same as above	same	emb Z54248
2D	~250	106 T7	gb M30836. Yeast telomeric DNA. 67 in 105.	100 in 12	CA rich	emb Z54253
3D	~200	132 T7	emb Z27493. Pig microsatellite DNA. 67 in 103.	62.9 in 35	CA rich	emb Z54258
4D	~310	112 T7	gb M98479. Human transglutaminase mRNA, 3′-untranslated region. 68 in 64.	66.7 in 14	23 bp SR-like	emb Z54260
6D	~230	137 T7	gb M77199. Ovine dinucleotide repeat polymorphism (MAF70 locus) DNA. 97 in 77.	85.7 in 14 or 64.1 in 39	CA repeat	emb Z54266

[a]*Flanking oligonucleotide A is excluded from all of the analysis results.* [b]*Length is estimated by* Hind *III-restriction analysis or by sequencing (exact number).* [c]*A number followed by T3 or T7 means that only ends of an insert were sequenced using either the T3 or the T7 primer.* [d]*Database accession number is followed by name of the most homologous sequence retrieved after a BLASTN search; numbers indicate % (first number) of homology in overlapping region (second number; bp).* [e]*This column indicates the percentage of homology (first number) and length of homologous region (bp; second number) aligned in each sequenced fragment with the corresponding target cDNA.* [f]*Tandem repeat (the number of bp in each period is shown).* [g]*Tandem repeat with variations in the length and/or sequence of the period.* [h]*This refers to the presence of a long tandem repeat with 2 bp (CA) period.* [i]*This implies that relatively short CA repeats are interrupted by other sequences.*

still bound to the test filter and/or target DNA that has been released from the filter. It is difficult to remove all traces of the target cDNA in this situation, even when purification from target cDNA immobilized on the membrane is carried out. In the approach outlined here, this problem was circumvented because the immobilized target cDNAs were not flanked by the adapter oligonucleotide used in the amplification of the hybridized DNA. The fidelity of PCR amplification in treating all sequences equally is an issue that is commonly addressed in PCR based approaches, but considering that the mRNAs must be amplified when working with a single cell, this is an unavoidable problem. Likewise with the aRNA amplification procedure (Van Gelder *et al.*, 1990), which has been less thoroughly investigated in its fidelity as compared to PCR, there are known problems with the accuracy in replication by reverse transcriptase (Boyer *et al.*, 1992; Temin, 1993).

Taken together, our results lead to the following conclusions: (1) the approach described here can be used to test the specificity of hybridization in an expression profiling experiment; and (2) 94.7% of the signals tested (18 clones from 19 sequenced) were artifacts of the reverse hybridization technique. The use of very complex mixtures for probe synthesis is likely to be the primary reason why so many non-homologous sequences were encountered using the expression profiling approach. In complex mixtures such as whole-cell mRNA, each individual low-abundance mRNA represents an extremely small proportion of the total whole-cell mRNA. The abundance of each individual message of low-abundance mRNA is less than 0.004% of the total mRNA (Alberts *et al.*, 1983) and therefore only about $1/(4 \times 10^5)$ of incorporated radioactivity during probe synthesis would be specific to a particular target cDNA. Thus it is advantageous to use relatively large amounts of the radioactive probe to monitor the expression level of a low-abundance mRNA. Contaminants such as genomic DNA, rRNA, tRNA, etc. would further increase the complexity of the probe and the contribution to the background signal. The use of DNAse pretreatment prior to cDNA synthesis could eliminate the amplification of genomic DNA, although this may be problematic at the single-cell level.

In order to amplify different regions of the mRNA, rather than only 3′-untranslated sequences, the T7-oligo-$(dT)_{24}$ oligonucleotide used in the original expression profiling report (Van Gelder *et al.*, 1990) was replaced by random hexanucleotides. This may result in more annealing to genomic DNA as well as rRNA and tRNA. However it is difficult to avoid the annealing of an oligo-dT containing primer to any six or more 'A's occurring together in these nucleic acids (Pekhletski, R., unpublished observation). Therefore, some of the sequences shown in Table 1 could be derived from non-mRNA molecules due to use of the random primers. Finally, as noted above, some species of crustaceans are particularly enriched in repetitive DNAs such as satellite DNA; the presence of repetitive DNA sequences may also have contributed to the high level of non-specific hybridization.

The sequence of the cDNA to be tested in an expression profiling experiment can dramatically affect the background level of hybridization. Not all regions of the tested cDNAs are equal in their ability to cross-hybridize to different mRNA sequences. For example, in the case of mGluR4a, all hybridized DNA fragments sequenced were complementary to two clearly defined regions of the cDNA. These two regions are probably derived from a phylogenetically amplified sequence, which is common for a number of different mRNAs. In instances where such sequences have been previously identified, the elimination of such cross-hybridizing sequences can be carried out to reduce the level of background hybridization. An initial database search for homologous sequences for target cDNAs to be used in expression profiling would be helpful in identifying such cross-hybridizing fragments. After localization of such fragments, one could easily eliminate them from the cDNAs to be tested by standard recombinant DNA methods. The use of such 'optimized' cDNAs in conjunction with larger amounts of labelled probe may dramatically decrease the level of non-specific cross-hybridization in an expression profiling experiment.

Another approach for identifying false positives is to use a target sequence in which two or more regions within a given gene are fixed to the filter in which the mixed probe is hybridized. One can then compare the two known fragments for relative intensity differences. A method to reduce the number of false positives would be to hybridize to oligonucleotide targets due to their greater sensitivity to mismatches. Nevertheless, on the basis of our present observations, we suggest that for accurate quantitation in expression profiling, the sequences of the hybridizing nucleic acids should be confirmed.

Acknowledgments

The authors thank Dr J. Eberwine (University of Penn.) for advice on techniques and fruitful discussions, Dr Eric Hull (University of Toronto) for editorial comments, and Dr L. Wolszon and Dr E. Macagno for providing a working environment at Columbia University for pilot studies in amplifying mRNA from leech neurons. A hearty thanks go to those who unselfishly supplied cDNAs to us (L. I. Grossman,

W. A. Caterall, T. P. Snutch, L. K. Kaczmarek, T. C. Sudhof, W. Trimble, W. K. Kang, R. Davis, B. Dauwalder, and C. Thomsen). This work was supported by a Medical Research Council (Canada) Group of Nerve Cells and Synapses (H.L.A. and D.R.H.) and by a Network Centres of Excellence (Canada) postdoctoral fellowship (R.L.C.).

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. (1983) *Molecular Biology of the Cell*, p. 409. New York: Garland Publishing, Inc.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403-404.

Bargas, J., Howe, A., Eberwine, J., Cao, Y. and Surmeier, D. J. (1994) Cellular and molecular characterization of Ca^{2+} currents in acutely isolated, adult rat neostriatal neurons. *J. Neurosci.*, **14**, 6667-6686.

Boyer, J. C., Bebenek, K. and Kunkel, T. A. (1992) Unequal human immunodeficiency virus type 1 reverse transcriptase error rates with RNA and DNA templates. *Proc. Natl. Acad. Sci. USA*, **89**, 6919-6923.

Browx, S. D. M. (1983) Insertion sequences and tandem repetitions as sources of variation in a dispersed repeat family. *J. Mol. Biol.*, **165**, 249-256.

Claverie, J.-M. and Makalowski, W. (1994) Alu alert. *Nature*, **371**, 752.

Cooper, R. L., Marin, L. and Atwood, H. L. (1995) Synaptic differentiation of a single motor neuron: Conjoint definition of transmitter release, presynaptic calcium signals, and ultrastructure. *J. Neurosci.*, **15**, 4209-4222.

Cruces, J., Wonenburger, M. L. G., Diaz-Guerra, M., Sebastian, J. and Renart, J. (1986) Satellite DNA in the crustacean *Artemia*. *Gene*, **44**, 341-345.

Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M. and Coleman, P. (1992) Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci. USA*, **89**, 3010-3014.

Fowler, R. F., Bonnewell, V., Spann, M. S. and Skinner, D. M. (1985) Sequences of three closely related variants of a complex satellite DNA diverge at specific domains. *J. Biol. Chem.*, **260**, 8964-8972.

Haisenleder, D. J., Yasin, M., Yasin, A. and Marshall, J. C. (1993) Regulation of prolactin, thyrotropin subunit, and gonadotropin subunit gene expression by pulsatile or continuous calcium signals. *Endocrinology*, **133**, 2055-2061.

Kang, W.-K. and Naya, Y. (1993) Sequence of the cDNA encoding an actin homolog in the crayfish *Procambarus clarkii*. *Gene*, **133**, 303-304.

Karlin, S. and Altschul, S. F. (1993) Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc. Natl. Acad. Sci. USA*, **90**, 5873-5877.

Kravitz, E. A., Schneider, H., Brooks, B. and Eberwine, J. H. (1992) Monitoring gene expression in single neurons from the lobster CNS. *Soc. Neurosci. Abstr.*, **18**, 1413.

Lnenicka, G. A. and Atwood, H. L. (1985a) Long-term facilitation and long-term adaptation at synapses of a crayfish phasic motoneuron. *J. Neurobiol.*, **16**, 97-110.

Lnenicka, G. A. and Atwood, H. L. (1985b) Age-dependent long-term adaptation of crayfish phasic motor axon synapses to altered activity. *J. Neurosci.*, **5**, 459-467.

Mackler, S. A. and Eberwine, J. H. (1993) Diversity of glutamate receptor subunit mRNA expression within live hippocampal CA1 neurons. *Mol. Pharmacol.*, **44**, 308-315.

Mackler, S. A., Brooks, B. P. and Eberwine, J. H. (1992) Stimulus-induced coordinate changes in mRNA abundance in single postsynaptic hippocampal CA1 neurons. *Neuron*, **9**, 539-548.

Miyashiro, K., Dichter, M. and Eberwine, J. (1994) On the nature and differential distribution of mRNAs in hippocampal neurites: Implications for neuronal functioning. *Proc. Natl. Acad. Sci. USA*, **91**, 10800-10804.

Moro, L., Colombi, M., Crespi, S., De Lernia, R. and Barlati, S. (1990) Study of fibronectin expression in tumour cells by dot-blot and in situ hybridization: quantitative evaluation by image analysis. *Cell Biol. Int. Rep.*, **14**, 701-715.

Mueller, P. R. and Wold, B. (1989) In vivo footprinting of muscle specific enhancer by ligation mediated PCR. *Science*, **246**, 780-786.

Nguyen, P. V. and Atwood, H. L. (1990) Expression of long-term adaptation of synaptic transmission requires a critical period of protein synthesis. *J. Neurosci.*, **10**, 1099-1109.

Otsuka, M., Kravitz, E. A. and Potter, D. D. (1967) Physiological and chemical architecture of a lobster ganglion with particular reference to gamma-aminobutyrate and glutamate. *J. Neurophysiol.*, **30**, 725-752.

Parnas, I. and Atwood, H. L. (1966) Phasic and tonic neuromuscular systems in the abdominal flexor muscles of the crayfish and rock lobster. *Comp. Biochem. Physiol.*, **18**, 701-723.

Pekhletsky, R. I., Chernov, B. K. and Rubtsov, P. M. (1992) Variants of the 5′-untranslated sequence of human growth hormone receptor mRNA. *Mol. Cell. Endocrinol.*, **90**, 103-109.

Pekhletsky, R., Cooper, R. L., Hampson, D. and Atwood, H. L. (1994) Changes in crayfish mRNA expression during motor neuron adaptation. *Amer. Zool.*, **34**, 66A.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning*, 2nd ed. New York: Cold Spring Harbor Laboratory Press.

Skinner, D. M. (1977) Satellite DNA's. *Bioscience*, **27**, 790-796.

Surmeier, D. J., Eberwine, J., Wilson, C. J., Stefani, A. and Kitai, S. T. (1992) Dopamine receptor subtypes co-localized in rat striatonigral neurons. *Proc. Natl. Acad. Sci. USA*, **89**, 10178-10182.

Temin, H. M. (1993) Retrovirus variation and reverse transcription: abnormal strand transfers result in retrovirus genetic variation. *Proc. Natl. Acad. Sci. USA*, **90**, 6900-6903.

Tsykin, A., Thomas, T., Milland, J., Aldred, A. R. and Schreiber, G. (1990) Dot blot hybridization using cytoplasmic extracts is inappropriate for determination of mRNA levels in regenerating liver. *Nucleic Acids Res.*, **18**, 382.

Van Gelder, R. N., von Zastrow, M. E., Yool, A., Dement, W. C., Barchas, J. D. and Eberwine, J. H. (1990) Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. USA*, **87**, 1663-1667.

Varshney, U., Jahroudi, N., van de Sande, J. H. and Gedamu, L. (1988) Inosine incorporation in GC rich RNA probes increases hybridization sequence specificity. *Nucleic Acids Res.*, **16**, 4162.

White, B. A. and Bancroft, F. C. (1982) Cytoplasmic dot hybridization. Simple analysis of relative mRNA levels in multiple small cell or tissue samples. *J. Biol. Chem.*, **257**, 8569-8572.

Wine, J. and Hagiwara, G. (1977) Crayfish escape behavior. I. The structure of efferent and afferent neurons involved in abdominal extension. *J. Comp. Physiol.*, **121A**, 145-172.

Wine, J. J. and Krasne, F. B. (1982) The cellular organization of crayfish escape behavior. In *The Biology of Crustacea*, ed. H. L. Atwood and D. C. Sandeman, pp. 241-292. New York: Academic Press.

Accepted 14 March 1996