AD_____

MIPR NO: 95MM5602

TITLE: Development of Methods that Detect and Monitor Environment Munitions Contaminants Using Plant Sentinels and Molecular Probes

PRINCIPAL INVESTIGATOR(S): Paul Jackson, Ph.D.

CONTRACTING ORGANIZATION: Department of Energy (DOE), Albuquerque
Albuquerque, New Mexico 87185-5400

REPORT DATE: March 1996

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1996	Final (21 Feb 95 - 31 Jan 96)

4. TITLE AND SUBTITLE

Development of Methods that Detect and Monitor Environment Munitions Contaminants Using Plant Sentinels and Molecular Probes

5. FUNDING NUMBERS

95MM5602

6. AUTHOR(S)

Paul Jackson, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Energy (DOE), Albuquerque
Albuquerque, New Mexico 87185-5400

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick
Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Plants accumulate TNT and similar compounds from soil. Their sessile nature requires that plants adapt to environmental changes by biochemical and molecular means. In principle, it is possible to develop a monitoring capability based on expression of any gene that is activated by specific environmental conditions. In the past year, we have identified plant genes activated or repressed upon exposure to TNT. We have cloned portions of these genes, sequenced these and developed DNA probes that measure TNT-induced gene activation or repression. These are being used to develop sensitive gene expression assays to monitor plant responses to growth in munitions-contaminated environments. We have also developed a collaboration with a private company to field test these assays in a system that uses selected plants to remove and degrade munitions from contaminated waste waters. By monitoring expression of TNT-induced genes, we can monitor whether the plants are exposed to this compound and, indirectly, follow the degradation process in a constructed ecosystem.

14. SUBJECT TERMS

TNT, Environmental Pollution, gene expression, mRNA

15. NUMBER OF PAGES

22

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to *stay within the lines* to meet ***optical scanning requirements***.

Block 1. Agency Use Only *(Leave blank)*.

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	PR	- Project
G	- Grant	TA	- Task
PE	- Program Element	WU	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

- DOD - See DoDD 5230.24, "Distribution Statements on Technical Documents."
- DOE - See authorities.
- NASA - See Handbook NHB 2200.2.
- NTIS - Leave blank.

Block 12b. Distribution Code.

- DOD - Leave blank.
- DOE - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
- NASA - Leave blank.
- NTIS - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only)*.

Blocks 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the US Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

PJJ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86 23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Paul J Jackson 3/1/96

PI - Signature Date

TABLE OF CONTENTS

INTRODUCTION

Plants have been tested for use as sensors for environmental pollution because they are inexpensive to produce and maintain and are in direct contact with the environment. Most bioassays measure only the harmful effects of pollutants on plants by scoring survival rates or visible indicators such as leaf damage and growth inhibition. However, an emerging approach to environmental bioassays is to detect changes in gene expression as a result of pollution. Assays based on such genetic responses have been largely limited to detection of bacterial genes encoding pollutant degradation, or uncharacterized stress-responsive genes (Blom, *et al.*, 1993; Erb and Wagner-Döbler, 1993; Van Dyk, *et al.*, 1994). Plants also modify their gene expression in response to environmental changes, activating different metabolic pathways to protect cells and tissues from damage while repressing others. By monitoring expression of genes that respond to specific pollutants, it may be possible to detect these environmental contaminants indirectly, using plants as environmental monitors.

Such a molecular bioassay depends on the discovery of specific probes for genes that respond to particular environmental factors. *Datura innoxia* and certain other plant species tolerate moderate amounts of 2,4,6-trinitrotoluene (TNT) in soil and cell culture media. These plants accumulate and metabolize TNT. The basis of tolerance may include an inducible change in the pattern of gene expression produced by the plant. By monitoring these changes, one may discern a specific "mRNA or protein signature" that is unique to TNT exposure Identification of such a pattern (and associated molecular probes to detect this) should allow development of strategies that use plants as "environmental sentries" to monitor TNT soil contamination and migration.

A molecular bioassay depends on the discovery of specific probes for response to particular environmental factors. Arbitrarily primed PCR RNA fingerprinting (RAP), as described by Welsh, *et al.* (1992), was used to identify and clone cDNA fragments encoding messenger RNA's whose concentrations increase or decrease with TNT exposure. Amplifying different arbitrary subsets of these mRNA's displayed these profiles in an easily interpreted gel electrophoresis pattern. While the majority of cDNA products were unchanged by changing TNT concentrations and exposure times, a small subset showed changes that were TNT dose or time dependent. DNA fragments specifically produced in reactions containing mRNA template from TNT-treated cells were excised, and the cDNA cloned and sequenced. Searches for homologous sequences in nucleic acid and protein sequence databases demonstrated that three cDNA isolates have significant homology to previously described protein or cDNA sequences. The remaining clones appear to represent previously unknown genes whose expression may respond specifically to TNT exposure. Molecular probes for these genes have been developed to further study kinetics of TNT-induced gene expression. Probes developed to detect changes in gene expression in plant cell cultures exposed to TNT are being tested to determine whether the same genes are similarly regulated in intact plants. A collaboration with an independent private company is being developed to determine whether this approach can be used to monitor changing TNT concentrations in a constructed

ecosystem that uses different plants to remove and metabolize TNT from contaminated soils and effluents.

EXPERIMENTAL METHODS

Cell culture growth. Suspension cultures of *Datura innoxia* cells were grown as previously described (Jackson, et al., 1984). Cultures in 250 mL baffled flasks were diluted 4-fold every 2 days by adding 12.5 mL of suspension to 37.5 mL of fresh medium. Suspensions were aerated by shaking (120 rpm) at 30°C in the dark. Stock cultures were expanded to provide cells for experiments and were allowed to grow for 24 hours after passage before exposure to the inducing agent.

TNT Exposure. A 110 mM stock solution of TNT in ethyl alcohol was prepared and stored in the dark at 4°C. This was introduced to final concentrations of 44, 110, 220 and 440 μM TNT into cultures diluted in fresh growth medium 24 hours previously. Previous experiments showed that 44 μM TNT had no effect on long-term cell growth rates, while 110 and 220 μM TNT significantly inhibited growth and 440 μM TNT caused death within 48 hours. Control cultures containing no additives or the same ethyl alcohol concentration as cultures containing 440 μM TNT were also grown and harvested. Cultures were incubated and shaken for 2, 8 and 13 hours following TNT addition before portions of the cell suspensions were harvested for RNA isolation.

RNA Isolation. Cell suspensions (5 mL) were removed from medium by gravity filtration and frozen in liquid nitrogen. Frozen cells were ground to a fine powder with a sterile liquid nitrogen-chilled mortar and pestle, and the RNA was isolated using the procedure described by Chomczynski and Sacchi (1987). Reprecipitated RNA was dissolved in sterile water and used immediately for first strand synthesis or stored at -70°C.

Poly(A)$^+$ RNA Selection. For some samples, the poly(A)$^+$ RNA fraction was isolated from cleared lysates without prior RNA purification. Messenger RNA was isolated using a commercially available kit (PolyAttract, Promega Corporation, Madison, WI) or oligo(dT) cellulose. The latter method was based on the Quick Prep kit (Pharmacia Biotech, Inc., Alameda, CA). Purified poly(A)$^+$ RNA was used in first strand synthesis reactions immediately after spectrophotometric quantification.

First Strand Synthesis with Decamer Primers. First strand cDNA syntheses were performed in sterile "V"-bottom 96-well microtiter plates (Nunc, Inc.). For each 25 μl reaction, 200 ng of total RNA, or 50 ng of poly(A)$^+$ RNA were heated to 65°C for 5 minutes, then chilled on ice. Reaction mixes contained RNA, 1 μM 10-nt arbitrary primer (Table 1), 100 μM each dATP, dTTP, dCTP and dGTP, 50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$ and 200 U M-MLV Reverse Transcriptase (BRL-Life Technologies, Inc.). Reaction wells were sealed with vinyl tape to reduce evaporative loss and the plate was incubated for 1 hour at 37°C. Reactions were further incubated at 65°C for 15 minutes to inactivate the enzyme. The resultant first strand cDNA was stored at 4°C until needed.

Table 1. **Oligonucleotide Primer Sequences**

Primer	Sequence	Primer	Sequence
AP-21	CGA GTA TGA G	AP-26	GAA GCT GAG G
AP-22	TTG ATA CGA C	AP-27	TCC ATA TCG A
AP-23	GCT CAT CAG G	AP-28	CGG AGT CTG T
AP-24	CTC CTG TAG A	AP-29	GGA ATG CCA T
AP-25	GAT TAT CGC C	AP-30	GAG CAC GAC T

Second Strand Synthesis and PCR Amplification. Five mL of each first strand reaction mix (equivalent to 40 ng of beginning total RNA) was added to a 25 mL PCR reaction mix (10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$) containing 200 µM of each dATP, dTTP, dCTP and dGTP and 1 µM additional DNA oligomer. The second strand mixes were capped with 25 mL melted paraffin wax and 2 units of Taq polymerase (AmpliTaq, Perkin-Elmer Corp.) were placed on top of the wax in 5 µl of 1 x PCR buffer. Second strand reactions were incubated using a "hot start" protocol (Bassam and Caetano-Annolés, 1993) in a Perkin-Elmer 9600 thermocycler by heating at 94°C for 5 minutes, 40°C for 5 minutes, then 72°C for 5 minutes. This was followed by 40 amplification cycles of 94°C for 30 seconds, 40°C for 2 minutes, then 72°C for 30 seconds (Bauer, *et al.*, 1993). After a final extension at 72°C for 5 minutes, the samples were stored at 4°C until analyzed by gel electrophoresis.

Polyacrylamide Gel Electrophoresis and Silver Staining. Amplified DNA was separated on a nondenaturing TBE-6% polyacrylamide gel (Protean II, Bio-Rad Laboratories, Cambridge, MA) to resolve fragments between 0.2 and 1.0 kb. Fragments were detected by silver staining (Bassam, *et al.*, 1991). Stained gels were preserved by drying on a Bio-Rad Model 583 gel dryer at 80°C for 2 hours under vacuum.

DNA Recovery and Re-Amplification. DNA fragments that showed increased or decreased intensity following TNT exposure were excised from dried, silver stained gels and placed in 100 mL of PCR reaction mix (10 mM Tris-HCl pH 8.3, 10 mM KCl, 4 mM $MgCl_2$, 200 µM each dATP, dTTP, dCTP and dGTP, 1 µM of the same arbitrary primer used to generate the original fragment, and 10 U of Taq polymerase Stoffel fragment (Perkin-Elmer Corp.). DNA was re-amplified from the dried gel, with no further purification (Weaver, *et al.*, 1994) by 40 cycles of 94°C for 30 seconds, 40°C for 2 minutes, then 72°C for 30 seconds. After a final extension at 72°C for 5 minutes, the reactions were stored at 4°C. Polyacrylamide gel electrophoresis of the reaction products showed DNA migrating at the same rate as the original fragment, although the re-amplified fragments were sometimes accompanied by as many as two additional fragments with slightly different molecular weights.

Cloning and Sequencing. PCR-amplified DNA was ligated to pGEM-T vector (Promega Corp.) without prior purification from the PCR reaction mix. PCR reaction mix (10 µl) was ligated in a 15 µl ligation reaction mix (30 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP) to 5.3 ng of pGEM-T by addition of 1 Weiss unit of T4 DNA Ligase (Promega Corp.). This was incubated at 15°C

overnight. The entire 15 µl ligation reaction was used to transform competent DH5α cells, using a 5 minute 37°C heat shock and 1.5 hour pre-incubation in LB broth before plating the entire reaction onto LB-agar plates containing 50 µg/mL ampicillin, 40 µg/mL 5-bromo-4-chloro-3-indolyl-β-**D**-galactoside (X-gal) and 200 µg/mL isopropylthio-β-**D**-galactoside (IPTG).

Overnight cultures were diluted 1:200 into 10 mL of LB broth containing 50 µg/mL ampicillin, then grown to an OD_{600} of 1.5 - 1.8. Glycerol (150 mL) was added to 1 mL of the liquid culture and the stock was frozen and stored at -20°C. Plasmid DNA was isolated from the remainder using the Qiagen Tip-20 alkaline lysis protocol provided by the company (Qiagen, Inc., Chatsworth, CA). Purified DNA was sequenced by the Sanger dideoxy method using dye-terminators (Dye-Deoxy, Applied Biosystems, Inc., Foster City, CA) in an Applied Biosystems model 373A automated sequencer. The M13 forward universal primer (5'-GTAAAACGACGGCCAGT-3') and reverse primer (5'-AACAGCTATGACCATG-3') were used to sequence both strands.

Northern Blotting or RNAse Protection Assays. DNA sequences were used to design DNA oligomers used as PCR primers to amplify genes encoding the expressed mRNA sequences. The resulting amplicons were ligated directly to pGEM-T DNA (Promega, Inc.) as previously described. The ligation results were used to transform DH5a cells and white colonies picked following ampicillin, IPTG, X-gal selection. Cloned inserts were sequenced and the sequences compared to previous cDNA sequences. Labeled DNA, used to probe Northern blots, was generated by PCR using primers specific to the SP6 and T7 promoter sequences present in the vector DNA. Hybridization was conducted using standard methods.

Sequence Database Searches. DNA sequences, minus any vector and PCR primer sequences, were used to search for sequences of significant homology in the GenBank database using the GCG Wordsearch module (Devereux, *et al.*, 1984). The basic local alignment search tool (BLAST) was also used to search sequence databases (Altschul, *et al.* 1990). The National Library of Medicine electronic mail server (blast@ncbi.nlm.nih.gov) was used for these searches, in which cDNA sequences were sent via E-mail for nucleotide sequence comparisons (BLASTN program) with the non-redundant sequence database compiled from GenBank, GenBank updates, EMBL, and EMBL updates. The cDNA sequences were also compared with protein sequence databases by using the BLASTX program. Only default parameters were used, making the searches relatively non-selective, but only statistically significant matches ($P = 1 \times 10^{-4}$) were investigated.

Plant Growth. Plant tissue samples will also be monitored for content of TNT and its metabolic derivatives and for TNT induced genes using the probes described above. This information will be useful to determine what portion of TNT removal and biodegradation is accomplished by the vegetative portion of the ecosystem and what portion is accomplished by microbial activity. Each system ecology will be monitored for TNT depletion, with the expectation that TNT concentrations will drop as liquid moves through the system. We expect that TNT concentrations will drop at the same rate that other parameters such as total nitrogen or chemical oxygen demand declined in previous pilot studies. Basic operational parameters such as water temperature, pH, ORP, flow, greenhouse temperature, ambient air temperature, pump cycling and aeration cycling will be continuously monitored

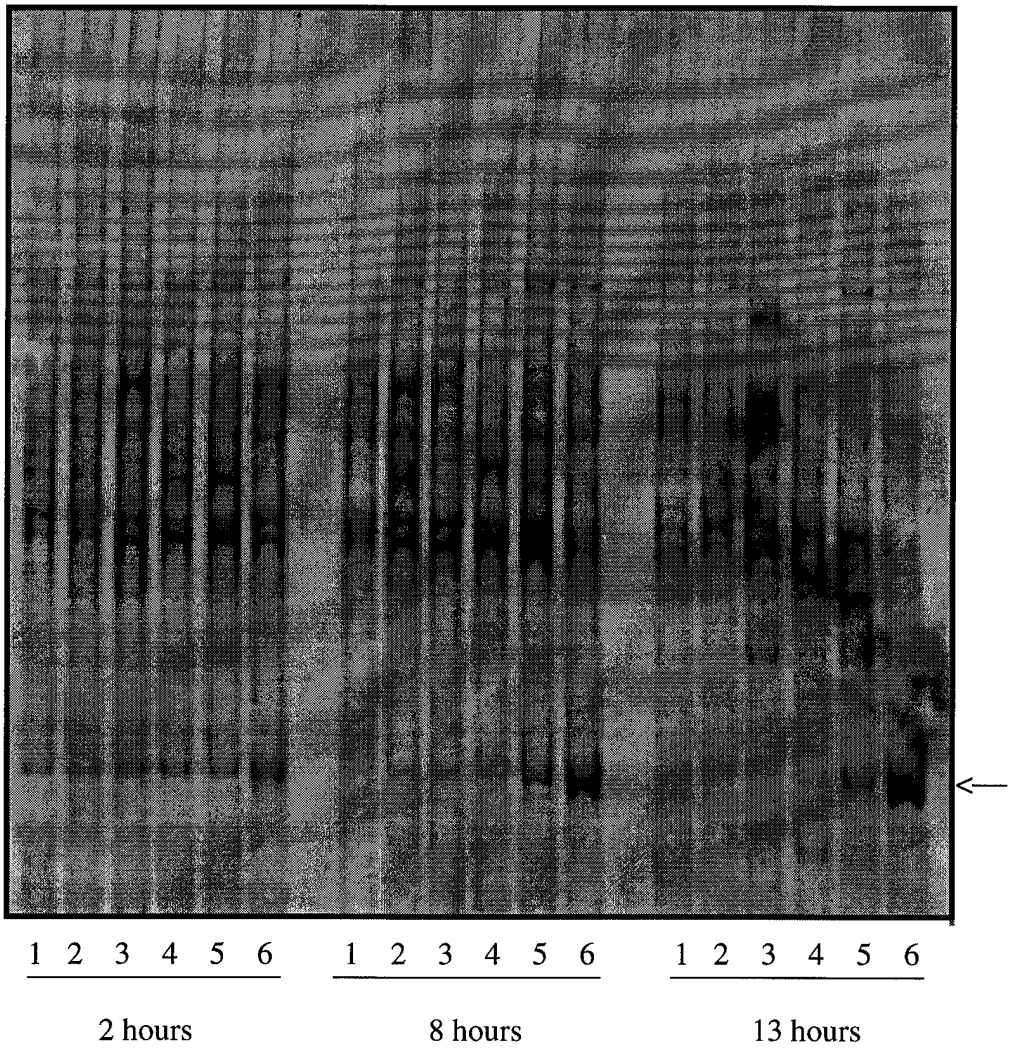

Figure 1. An Example of Time and TNT Dose-Responsive cDNA.

Differential display PCR reactions were performed using RNA from cells exposed to TNT for 2, 8, and 13 hours. Primer 5'-CTCCTGTAGA-3' was used in first and second strand reactions. Within each time point cultures were treated with (1) No change, (2) 1.6 ml/ml ethyl alcohol, (3) 44 μM TNT, (4) 110 μM TNT, (5) 220 μM TNT and (6) 440 μM TNT. Electrophoresis was at 200 V for 14 hours through a 6% TBE-polyacrylamide gel. Arrow shows a 290 bp amplified cDNA fragment (AP24.1) that demonstrates both dose and time dependent expression.

using a computer for data collection. All other samples will be collected on a weekly basis as grab samples and analyzed for remaining parameters of interest. If indications of stress in different portions of the system become apparent, monitoring frequency will be increased. This work is in progress. Each of six different treatment plant ecologies is being established. These will be monitored for bacterial activity using ORP probes, by selecting samples of the liquid to be analyzed for bacterial content (FAME analysis), by performing PCR analysis to study the distribution of non-cultural microbes and fungi present, and by performing plate counts on selected media. This will provide immediate information about changes in microbial content upon introduction of the septage carrier and munitions-containing waste stream. Microbial sub-populations are expected to change relative to one another. However absolute quantities of microbes are expected to remain approximately the same.

RESULTS

Classes of Differential Display Patterns. A representative example of TNT-induced gene expression identified using the methods outlined is shown in Figure 1. More than 20 different DNA fragments were identified as candidates for TNT-responsive cDNAs using this method. Of these, three showed TNT dose-dependent induction. Two other fragments showed time and dose dependence. These were of particular interest because the relationship between their expression and TNT exposure suggested that they might encode TNT-specific response genes. Six fragments were present only at the highest TNT concentration. Several additional fragments decreased or disappeared upon TNT exposure. Finally, four fragments were picked from the poly(A)$^+$ differential display experiment because they appeared only when RNA was isolated from cells that were continuously grown in 44 μM TNT, and were not amplified from RNA isolated from untreated cells.

When the DNA fragments were isolated and re-amplified then examined by electrophoresis, two DNA fragments were present in some reactions. One migrated slightly faster than expected. An attempt was made to re-amplify and clone the reamplification products of both bands. However, this was often unsuccessful. Fragments were ligated to plasmid pGEM-T and ligation products were transformed into *E. coli* (strain DH5α) and plated.

All the sequenced inserts contained complete primer sequences at their termini, confirming their origin as full length PCR products. However, when two clones from the same re-amplified DNA fragment were sequenced, the length and sequence of one clone differed from the other. This is consistent with the observation of several polymorphic DNA fragments present in a single amplified DNA sample (Bauer, *et al.*, 1993.). Northern blots containing mRNA isolated from cells exposed to different TNT concentrations were therefore hybridized to probes to determine which of the cloned sequences is representative of a TNT-modulated gene. One of the two probes hybridized to a gene that was constitutively expressed while the other hybridized to mRNA present only in cells exposed to TNT.

Computer Database Searches. When DNA sequences of the cDNA fragments were compared to the GenBank database using the GCG Wordsearch program

(Devereux, *et al.*, 1984), two clones showed significant homology to a previously described sequence, while others represented novel mRNA sequences. The identified sequences, from fragments AP26.2 (clone M) and AP26.4 (clone Q) share 71% nucleotide identity with a portion of the 3' untranslated region of a tobacco (*Nicotiana silvestris*) ubiquitin cDNA clone (Genschik, *et al.*, 1992). AP26.2 and AP26.4 were generated from RNA extracted from cells that had been growing continuously in 44 μM TNT. The sequences and lengths of these two clones were identical, although they were originally cloned from two different DNA fragments. It is possible that one clone is a contaminating DNA from another fragment. If so, then a single band on a polyacrylamide gel may contain length, as well as sequence polymorphisms. Since DNA was separated on native gels, it is possible that fragments of different sizes might migrate together based on secondary structures of the molecules. Because these clones are identical, they were regarded as a single cDNA isolate.

DNA sequence comparisons using the BLASTN program revealed no significant nucleotide sequence matches other than tobacco ubiquitin. BLAST expresses the statistical significance of a match with a P-value, such that a P-value close to one indicates a high probability that such a match could occur by chance alone. Nucleic acid sequence matches for tobacco ubiquitin (clones AP26.2 and AP26.4) had P-values of 1.1×10^{-6}. Ubiquitin cDNAs originating from other species made less significant matches, with *Arabadopsis thaliana* (P = 0.38), rice (P = 0.77), and *Brassica rapa* (P = 0.98). Other cDNA clones, with relatively high P-values ($P > 1 \times 10^{-4}$) were not considered homologous to known DNA sequences.

BLASTX translates both strands of a nucleotide sequence in all three reading frames, then compares the peptide sequences with protein sequence databases. [Clone AP29.3 (clone I) produced a polypeptide sequence from the plus strand (+1) reading frame that is rich in serine, and was similar to the serine and threonine-rich transcription activation (STA) domains (Theill, *et al.* 1989) of a variety of nuclear zinc-finger proteins such as human AF-9 (Nakamura, *et al.* 1993), a leukemia-associated protein ($P = 1.6 \times 10^{-9}$), and *Drosophila* ecdysone-inducible proteins E75A, B, and C (Seagraves and Hogness, 1990; Feigl, *et al.* 1989) ($P = 1.4 \times 10^{-5}$) that are products of alternative mRNA splicing from the same transcript. In addition, a serine-rich domain is found in the *engrailed*-like protein *smox*-2 of the parasitic flatworm *Schistosoma masoni* (Webster and Mansour, 1992) ($P = 5.8 \times 10^{-5}$). The plus strand (+3) reading frame of clone AP26.3 (clone O) encodes a peptide sequence that was similar to the extracellular dermal glycoprotein (EDGP) of carrots ($P = 2.0 \times 10^{-16}$) (Satoh, *et al.* 1992) and to soybean basic 7s globulin ($P = 1.7 \times 10^{-5}$) (Kagawa and Hirano, 1989). The EGDP gene is transcriptionally activated in response to mechanical wounding and pathogen invasion. The basic 7s globulin mRNA is induced in soybeans exposed to hot water for several hours, and is closely related to EDGP. Although EDGP mRNA is normally abundant in plant suspension cultures, AP26.3 is clearly induced in TNT-exposed cells. These proteins are both associated with wounding, pathogen invasion, and other forms of stress, so AP26.3 may be another general stress-induced cDNA, like the ubiquitin homologues previously described.

The recently initiated whole plant studies take advantage of a system developed by The Southwest Wetlands Group, Inc. of Santa Fe, NM. They have demonstrated that a combination of different ecologies can be used to treat municipal and residential waste water, landfill leachates, industrial wastewater,

acid mine drainage, and septage. The unique approach suggested by SWG incorporates septage as a carrier media and uses several different ecologies -- aquatic, wetlands, and meadow -- that rely on the associated microbial, fungal, algal and plant communities to function as consortia in processes that convert TNT to benign nitrogenous compounds. The use of these complex self-regulating ecologies makes this biotechnology inherently more effective and efficient. The ability to monitor *in situ* TNT concentrations indirectly by monitoring gene expression should provide a method to finely regulate the treatment process by monitoring TNT-modulated gene expression in plants located at different points within the treatment cycle. Completion of such experiments requires establishment of the different constructed ecologies. Ongoing establishment of the appropriate plant species is accompanied by use of *D. innoxia* probes to determine whether TNT-modulated genes from this species are sufficiently similar in other species to allow monitoring using the *D. innoxia* probes.

DISCUSSION

The gradual onset of chronic stress in plants can result in adaptation. Persistent changes occur that allow an adapted plant to endure conditions that would have killed it, had the stress been introduced abruptly. When challenged with an acute stress, many plants halt translation of mRNAs that are not related to repair or adaptation (Gallie, 1993). They do this by selective mRNA degradation or by sequestering transcripts into intracellular compartments. If the source of the stress is not removed, even the sequestered transcripts are degraded. Moreover, many genes are transcriptionally arrested by environmental stress. In this way, the condition of a plant's environment may be reflected in its pattern of gene expression.

While transpiring water out of soil, plants continuously monitor groundwater and soil quality over a large sample area. If there are toxic substances in the soil, plants eventually give evidence of their presence by external signs of stress. However appearance of these changes is not specific to any particular compound. A method is needed that detects changes in plants resulting from the presence of a particular compound, TNT, at concentrations that are lower than those that cause obvious systemic damage.

PCR amplified cDNA fragments generated from *D. innoxia* RNA templates with expression patterns suggesting a role in the plant cell response to TNT were identified. The DDRT-PCR technique is a powerful and practical method for discovering novel mRNAs that may serve as genetic indicators for TNT exposure. By simultaneously viewing the effect of different TNT concentrations at different time points, it was possible to distinguish TNT-dependent responses and general stress responses from cell growth-dependent responses within a subset of cDNA fragments.

Ten different arbitrary decamers were used as primers to select over 20 candidate DNA fragments that showed patterns of expression suggestive of TNT-dependent transcriptional modulation. These were cloned and sequences were determined for ten of these clones. One cDNA isolate, cloned from both AP26.2 and AP2.4, showed significant similarity to a known gene, with 71% sequence identity to the 3'-untranslated region of the tobacco ubiquitin mRNA. Although the coding regions of ubiquitin genes are conserved, there is normally

great diversity in the untranslated regions (UTR). The similarity between the AP26.4 clone and tobacco ubiquitin may imply some shared function that is encoded in this region. Two functions that are determined by 3'-UTR elements are intracellular localization of the mRNA (Hesketh, *et al.*, 1994) and mRNA stability (Green, 1993). Ubiquitin is a highly conserved gene found in all eukaryotes. It is a component of a system in which damaged or short-lived proteins are poly-ubiquitinated and thus marked for degradation (Cheng, *et al.*, 1994). Such a gene product might be expected to appear under highly stressful culture conditions. Another isolate, AP26.3, encodes a polypeptide that is similar to the wound- and pathogen-inducible extracellular dermal glycoprotein (EDGP) of carrot (*Daucus carota*). These two isolates demonstrate the validity of the DDRT-PCR technique to identify and isolate probes for stress-induced genes.

The appearance of ubiquitin and EDGP among the differential display clones demonstrates that this approach will identify genes that respond to TNT exposure. However, it also shows a fundamental obstacle to finding a TNT-specific genetic response. Many genes will change expression in response to stress of any kind. These generalized responses may have roles in repairing damage to the cell, such as scavenging damaged proteins, repairing DNA damage and arresting the cell cycle. Such transcripts can be expected to appear at levels of stress that can also be discerned through changes in growth rate, tissue damage, or other external pathological indicators. However, other genes will be quite specific for particular types of stress and may even be peculiar to certain chemicals. These may owe their specificity to encoding components responsible for catabolism of the chemical. Such genes might be the most sensitive indicators of particular pollutants, since they would appear at very low pollutant concentrations when widespread cell or tissue damage had not yet occurred. In the DDRT-PCR experiments performed here, such transcripts might appear at TNT concentrations that do not interfere with long term cell growth, but increase with exposure to higher TNT concentrations. Some of these transcripts may also be temporally down-regulated, as TNT is absorbed and degraded in the cell. Further screening to demonstrate TNT-specific gene expression modulation is needed to determine whether those genes already identified are modulated only by TNT exposure or whether their response is part of a broader generalized environmental stress response in the plant.

Induction of a catabolic pathway through increased mRNA abundance implies the existence of TNT-modulated regulators of gene expression. These regulatory proteins might contain DNA-binding and transcription-activating domains that can be identified by their sequence homology to known domains in sequence databases. Clone AP29.3 can encode a serine- and threonine-rich polypeptide that is similar to the transcription activation domains of the zinc finger proteins AF-9 (Nakamura, *et al.* 1993), E75A (Seagrave and Hogness, 1990) and *sox*-2 (Webster and Mansour, 1992). Transcription activating proteins might induce the coordinate expression of a variety of genes responsible for different aspects of cell protection and repair.

When selected DNA fragments were isolated and cloned, two isolated recombinant colonies were picked for sequencing. Sequences of two clones derived from the same apparent fragment were different, showing that a single gel band may be composed of more than one cDNA fragment. This has been observed by others, and precludes direct sequencing of DNA isolated from the gel

(Bauer, *et al.*, 1993). To sequence the differentially expressed cDNA, all of the cDNAs in a fragment must be cloned, and the differentially expressed clone must be distinguished from contaminating sequences. This was accomplished by producing DNA or RNA using different cDNA clones as template then hybridizing this to Northern blots containing mRNA isolated from cells exposed to different TNT concentrations. The induced sequences were readily apparent.

Once a single TNT-responsive cDNA clone is isolated, its sequence can be used to design specific primers for PCR amplification, and conventional RT-PCR can be used to detect this transcript. The specific clone can be screened further for TNT specificity by probing RNA from cells treated with different stress-inducing agents. These agents can be chemically related or unrelated to TNT. Moreover, complete cDNA and genomic clones can be isolated from libraries for further characterization. An array of highly specific TNT-induced cDNA probes can then be developed into a molecular bioassay for detecting TNT-induced changes in plant gene expression. This may allow detection of very small amounts of TNT in soils. Messenger RNA markers, especially those that accumulate in easily-sampled portions of the plant, may be detected using sensitive RT-PCR assays. If other specific molecular probes can be found for other chemicals, it may be possible to simultaneously detect different pollutants using tissue from a single plant.

CONCLUSIONS

The results presented to date strongly support our hypothesis that plants can be used as environmental sentries to monitor soil and water for selected contaminants. In this study we have focused on TNT contamination. However, in principle, such an approach could be used to detect any contaminant that is water soluble, taken up by plants and has a specific impact on plant gene expression.

Considerably more research is required to generate more TNT-modulated gene probes and demonstrate that these are TNT-specific, or at least specific enough so that, when used with a battery of different probes, a TNT-specific signature of gene expression is produced. While early efforts focused on identifying TNT-induced gene fragments. We have also identified a number of fragments that represent genes whose expression appears to be suppressed by TNT-exposure. In the future, we will continue testing probes to determine their specificity and isolating the intact genes (and their regulatory elements).

Transfer of this technology to the field requires demonstration that what is found in cell cultures and hydroponically grown plants also occurs in plants growing in their native environment. To demonstrate this, we are establishing a collaboration with The Southwest Wetlands Group, Inc., a small company in Santa Fe, New Mexico. They have independently demonstrated that a septage driven system of constructed ecologies can decontaminate soil and effluent contaminated with munitions and certain other organic compounds. Optimal use of their system requires continuous monitoring of TNT concentrations at different points in the treatment cycle. We hope to demonstrate whether probes for TNT-induced genes from *D. innoxia* can be used to monitor these same genes in selected plants present in their different constructed ecologies. Such a method of monitoring would provide a rapid method of determining what factors are

responsible for optimizing the treatment process. These experiments will also demonstrate whether such an approach can be used in practical field applications. At this point, the results are extremely promising.

References.

- Altschul, S.F., Gish, W., Miller, W., Meyers, E.W. and Lipman, D.J. (1990) Basic Local Alignment Search Tool. J. Mol. Biol. **215**, 403-410.
- Bassam, B.J., Caetano-Annolés, G. and Gresshoff, P.M. (1991) Fast and sensitive staining of DNA in polyacrylamide gels. Anal. Biochem. **196**, 80-83.
- Bassam, B.J. and Caetano-Annolés, G. (1993) Automated 'hot start' PCR using mineral oil and paraffin was. BioTechniques **14**, 31-33.
- Bauer, D., Muller, H., Reich, J., Riedel, H., Ahrenkel, V., Warthoe, P. and Strauss, M. (1993) Identification of differentially expressed mRNA species by an improved display technique (DDRT-PCR). Nucleic Acids Res. **21**, 472-480.
- Blom, A., Harder, W. and Matin, A. (1992) Unique and overlapping pollutant stress proteins of *escherichia coli*. Appl. Environ. Microbiol. **58**, 331-334.
- Cheng, L., Watt, R. and Piper, P.W. (1994) A method for the elimination of false positives generated by the mRNA differential display technique. BioTechniques **16**, 1096-1103.
- Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation from acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**, 156-159.
- Devereux, J., Haeberli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs or the VAX. Nucleic Acids Res. **12**, 387-395.
- Erb, R.W. and Wagner-Döbler, I. (1993) Detection of polychlorinated biphenyl degradation genes in polluted sediments by direct DNA extraction and Polymerase Chain Reaction. Appl. Environ. Microbiol. **59**, 4065-4073.
- Feigl, G., Gram, M. and Pongs, O. (1989) A member of the steroid hormone receptor gene family is expressed in the 20-OH-Ecdyson inducible puff 75B in *Drosophila melanogaster*. Nucleic Acids Res. **17**, 7167-7178.
- Gallie, D.R. (1993) Posttranscriptional regulation of gene expression in plants. Ann. Rev. Plant Physiol. Plant Mol. Biol. **44**, 77-105.
- Genshchik, P., Parmentier, Y., Durr, A., Marbach, J., Criqui, M.C., Jamet, E. and Fleck, J. (1992) Ubiquitin genes are differentially regulated in protoplast-derived cultures of *Nicotiana sylvestris*, and in response to various stresses. Plant Mol. Biol. **20**, 897-910.
- Green, P.J. (1993) Control of mRNA stability in higher plants. Plant Physiol. **102**, 1065-1070.
- Hesketh, J., Campbell, G., Piechaczyk, M. and Blanchard, J.M. (1994) Targeting of c-Myc and β-globin coding sequences to cytoskeletal-bound polysomes by c-Myc 3' untranslated region. Biochem. J. **298**, 143-148.
- Jackson, P.J., Roth, E.J., McClure, P.R. and Naranjo, C.M. (1984) Selection, isolation and characterization of cadmium-resistant *Datura innoxia* suspension cultures. Plant Physiol. **75**, 914-918.

- Kagawa, H. and Hirano, H. (1989) Sequence of a cDNA encoding soybean basic 7S globulin. Nucleic Acids Res. **17**, 8868.
- Nakamura, T., Alder, H., Gu, Y., Prasad, R., Canaani, O., Kamada, N., Gale, R.P., Lange, B., Crist, W.M., Nowell, P.C., Croce, C.M. and Canaani, E. (1993) Genes on chromsomes 4, 9, and 19 involved in 11q23 abnormalities in acute leukemia share sequence homology and/or common motifs. Proc. Natl. Acad. Sci. USA **90**, 4631-4635.
- Satoh, S., Sturm, A., Fuji, T. and Chrispeels, M.J. (1992) cDNA cloning of an extracellular dermal glycoprotein of carrot and its expression in response to wounding. Planta **188**, 432-438.
- Seagraves, W.A. and Hogness, D.S. (1990) The E75 ecdyson-inducible gene responsible for the 75B early puff in *Drosophila* encodes two new members of the steroid receptor superfamily. Genes and Dev. **4**, 204-219.
- Theill, L.E., Castrillo, J.L., Wu, D., and Karin, M. (1989) Dissection of functional domains of the pituitary-specific transcription factor GHF-1. Nature **342**, 945-948.
- Van Dyk, T.K., Majarian, W.R., Konstantinov, K.B., Young, R.M., Dhurjati, P.S. and LaRossa, R.A. (1994) Rapid and sensitive pollutant detection by induction of heat shock gene-bioluminescene gene fusions. Appl. Environ. Microbiol. **60**, 1414-1420.
- Weaver, K.R., Caetano-Annolés, G., Gresshoff, P.M. and Callahan, L.M. (1994) Isolation and cloning of DNA amplification products from silver-stained polyacrylamide gels. Biotechniques **16**, 226-227.
- Webster, P.J. and Mansour, T.E. (1992) Conserved classes of homeodomains in *Schistosoma mansoni*, an early bilateral metazoan. Mech. of Dev. **38**, 25-32.
- Welsh, J., Chada, K., Dalal, S.S., Cheng, R., Ralph, D. and McClelland, M. (1992) Arbitrarily primed PCR fingerprinting of RNA Nucleic Acids Res. **20**, 4965-4970.

Appendix A:
BLASTN Nucleic Acid Sequence Search Results.
List only the five highest scoring matches.

```
Query=  94-0530a20.Seq -20.5 sequence of clone "A" (AP21.1)
        (170 letters, both strands)

Database:  Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep
           9, 1994
           203,796 sequences; 207,906,619 total letters.
                                                                  Smallest
                                                                  Poisson
                                                          High    Probability
Sequences producing High-scoring Segment Pairs:           Score   P(N)        N

emb|X14597|PHCHSJ  P.hybrida chsJ gene for chalcone synthase    115  0.98        1
emb|X76558|SPRAD26 S.pombe rad26 gene.                          112  0.9994      1
emb|X64587|MMU2AF  M.musculus mRNA for splicing factor U2...    111  0.9999      1
**********************************************************************
Query=  94-0630120.Seq -20.5 sequence of clone "L" (AP26.2)
        (264 letters, both strands)

Database:  Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep
           9, 1994
           203,796 sequences; 207,906,619 total letters.
                                                                  Smallest
                                                                  Poisson
                                                          High    Probability
Sequences producing High-scoring Segment Pairs:           Score   P(N)        N

      *** NONE ***
**********************************************************************
Query=  94-0630m20.Seq -20.5 sequence of clone "M" (AP26.2)
        (229 letters, both strands)

Database:  Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep
           9, 1994
           203,796 sequences; 207,906,619 total letters.
                                                                  Smallest
                                                                  Poisson
                                                          High    Probability
Sequences producing High-scoring Segment Pairs:           Score   P(N)        N

gb|M74100|TOBCEP527A Nicotiana sylvestris (clone 6PCEP52-...    194  1.1e-06     1
gb|J05507|ATHUXP1A   A.thaliana ubiquitin extension prote...   128  0.39        1
emb|Z35369|ATTS3808  A. thaliana transcribed sequence; cl...   127  0.39        1
emb|Z34610|ATTS3370  A. thaliana transcribed sequence; cl...   127  0.41        1
gb|J05508|ATHUXP2A   A.thaliana ubiquitin extension prote...   127  0.44        1
**********************************************************************
Query=  94-0530n20.Seq -20.5 sequence of clone "N" (AP26.3)
        (215 letters, both strands)

Database:  Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep
           9, 1994
           203,796 sequences; 207,906,619 total letters.
                                                                  Smallest
                                                                  Poisson
                                                          High    Probability
Sequences producing High-scoring Segment Pairs:           Score   P(N)        N

      *** NONE ***
**********************************************************************
Query=  94-0530o20.Seq -20.5 sequence of clone "O" (AP26.3)
        (242 letters, both strands)

Database:  Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep
           9, 1994
           203,796 sequences; 207,906,619 total letters.
                                                                  Smallest
                                                                  Poisson
                                                          High    Probability
Sequences producing High-scoring Segment Pairs:           Score   P(N)        N

emb|Z34801|CEF59A2 Caenorhabditis elegans cosmid F59A2          127  0.48        1
**********************************************************************
```

Query= 94-0630p20.Seq -20.5 sequence of clone "P" (AP26.4)

Database: Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep 9, 1994
203,796 sequences; 207,906,619 total letters.

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Poisson Probability P(N)	N
emb|X56603|MMPCBP M.musculus mRNA for calcium-binding p...	115	0.998	1
gb|M92988|MUSERP60A Mouse calregulin (ERp60) mRNA, comple...	115	0.999	1
emb|X14926|MMCRP55 Mouse mRNA for calreticulin	115	0.999	1

Query= 94-0630q20.Seq -20.5 sequence of clone "Q" (AP26.4)
(229 letters, both strands)

Database: Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep 9, 1994
203,796 sequences; 207,906,619 total letters.

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Poisson Probability P(N)	N
gb|M74100|TOBCEP527A Nicotiana sylvestris (clone 6PCEP52-...	194	1.1e-06	1
gb|J05507|ATHUXP1A A.thaliana ubiquitin extension prote...	128	0.38	1
emb|Z35369|ATTS3808 A. thaliana transcribed sequence; cl...	127	0.39	1
emb|Z34610|ATTS3370 A. thaliana transcribed sequence; cl...	127	0.41	1
gb|J05508|ATHUXP2A A.thaliana ubiquitin extension prote...	127	0.44	1

Query= 94-0530e20.Seq -20.5 sequence of clone "E" (AP29.1)
(184 letters, both strands)

Database: Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep 9, 1994
203,796 sequences; 207,906,619 total letters.

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Poisson Probability P(N)	N
gb|U08931|NTU08931 Nicotiana tabacum SR1 seed coat-specif...	122	0.69	1
gb|L12467|YSCGCN1A Saccharomyces cerevisiae translational...	115	0.991	1

Query= 94-0530g20.Seq -20.5 sequence of clone "G" (AP29.2)
(213 letters, both strands)

Database: Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep 9, 1994
203,796 sequences; 207,906,619 total letters.

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Poisson Probability P(N)	N
gb|S74758|S74758 mas5=heat shock protein [Saccharomyc...	133	0.16	1
emb|X56560|SCYDJ1 Yeast YDJ1 gene for a protein homolo...	133	0.16	1
emb|X77783|SCACCOAST S.cerevisiae gene for acyl-CoA synth...	119	0.93	1
gb|U02878|RPU02878 Rickettsia prowazekii Madrid E trans...	117	0.98	1
dbj|D14877|CLOORF1 Clostridium perfringens orf, virR an...	114	0.9991	1

Query= 94-0630i20.Seq -20.5 sequence of clone "I" (AP29.3)
(190 letters, both strands)

Database: Non-redundant PDB+GBupdate+GenBank+EMBLupdate+EMBL, 5:16 AM EDT Sep 9, 1994
203,796 sequences; 207,906,619 total letters.

Sequences producing High-scoring Segment Pairs:	High Score	Smallest Poisson Probability P(N)	N
gb|U00431|U00431 Mus musculus HMG-1 mRNA, complete cds.	160	0.00068	1
emb|Y00365|CHHMG1 Chinese hamster HMG-1 gene for high ...	160	0.00070	1
emb|Z11997|MMNHCHMG1 M.musculus mRNA for non-histone chro...	160	0.00070	1
gb|J03998|PFAGAR P.falciparum glutamic acid-rich prot...	160	0.00070	1
emb|X13703|DDPDG1 Dictyostelium plasmid pDG1 DNA	151	0.0042	1

Appendix B:
BLASTX Peptide Sequence Search Results.
List only the five highest scoring matches.

Query= 94-0530a20.Seq -20.5 sequence of clone "A" (AP21.1)
(170 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
sp|P33042|VL03_VARV	PROTEIN L3.>pir|I36844|I36844 M...	+1	41	0.13	2
sp|P29663|CYB_AMICA	CYTOCHROME B (EC 1.10.2.2) (FRA...	-3	39	0.40	2
sp|P34869|CYB_HETFR	CYTOCHROME B (EC 1.10.2.2).>gp|...	-3	37	0.44	2
sp|P34871|CYB_ISUPA	CYTOCHROME B (EC 1.10.2.2).>gp|...	-3	37	0.44	2
gp|U06488|PRU06488_1	cytochrome b [Mitochondrion Poe...	-3	38	0.48	2

****

Query= 94-0630120.Seq -20.5 sequence of clone "L" (AP26.2)
(264 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
pir|A45560|A45560	SPAG-1=sporozoite surface antig...	-2	43	0.0015	4
sp|P35499|CINS_HUMAN	SODIUM CHANNEL PROTEIN, SKELETA...	+2	41	0.022	3
gp|L04236|HUMSCN4A21_1	skeletal muscle sodium channel ...	+2	41	0.022	3
gp|M81758|HUMSKM1A_1	sodium channel alpha subunit [H...	+2	41	0.022	3
pir|S27809|S27809	GTPase-activating protein - fru...	-2	48	0.029	2

****

Query= 94-0630m20.Seq -20.5 sequence of clone "M" (AP26.2)
(229 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
sp|P11000|WAPA_STRMU	WALL-ASSOCIATED PROTEIN PRECURS...	+3	44	0.062	2
pir|PN0528|PN0528	G-protein coupled receptor type...	-2	57	0.085	1
pir|A60095|A60095	larval glue protein 1 precursor...	+3	40	0.13	2
pir|JL0142|JL0142	leukocyte antigen CD31 precurso...	+3	58	0.20	1
sp|P16284|PEC1_HUMAN	PLATELET ENDOTHELIAL CELL ADHES...	+3	58	0.20	1

****

Query= 94-0530n20.Seq -20.5 sequence of clone "N" (AP26.3)
(215 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
gp|L10910|HUMHCCA_1	splicing factor [Homo sapiens]	-3	42	0.070	3
gp|L10911|HUMHCCB_1	splicing factor [Homo sapiens]	-3	42	0.071	3
sp|P17132|ROC_RAT	HETEROGENEOUS RIBONUCLEOPROTEIN...	-1	42	0.12	2
gp|L35899|DROAPROJ_1	Drosophila melanogaster project...	-3	62	0.17	1
gp|U11274|MMU11274_1	A+U-rich RNA-binding protein [M...	-1	42	0.23	2

****

Query= 94-0530o20.Seq -20.5 sequence of clone "O" (AP26.3)
(242 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
gp|D14550|DAREDGP_1	EDGP precursor [Daucus carota]	+3	143	2.0e-16	1
gp|X16469|GMBG7SG_3	Soybean Bg gene for basic 7S glo...	+3	83	1.7e-05	1
sp|P13917|7SBG_SOYBN	BASIC 7S GLOBULIN PRECURSOR.>pir...	+3	83	1.7e-05	1
gp|D16107|SOYSBG7S_1	basic 7S globulin [Glycine max]	+3	83	1.7e-05	1
pir|S21426|S21426	conglutin gamma precursor - narr...	+3	70	0.0027	1

****

Query= 94-0630p20.Seq -20.5 sequence of clone "P" (AP26.4)
(242 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
pir|S22697|S22697	extensin - Volvox carteri (fragme...	+3	47	0.012	3
pir|S38389|S38389	T-cell receptor beta chain precur...	-1	45	0.022	2
pir|A25889|RGMSA1	GTP-binding regulatory protein Gs...	+3	41	0.15	2
gp|M21142|HUMGNAS6_2	Human guanine nucleotide-binding ...	+3	41	0.15	2
sp|P29797|GBAS_PIG	GUANINE NUCLEOTIDE-BINDING PROTEI...	+3	41	0.15	2

****

Query= 94-0630q20.Seq -20.5 sequence of clone "Q" (AP26.4)
(229 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
sp|P11000|WAPA_STRMU	WALL-ASSOCIATED PROTEIN PRECURSOR...	+3	47	0.022	2
pir|A60095|A60095	larval glue protein 1 precursor -...	+3	42	0.022	2
sp|P03111|VE1_HPV1A	E1 PROTEIN.>pir|A03655|W1WLE E1 p...	+1	37	0.052	3
gp|Z34098|SCXLTEL_8	ORF [Saccharomyces cerevisiae]	+1	41	0.096	2
pir|A45155|A45155	mucin FIM-C.1 - African clawed fr...	+3	39	0.30	2

****

Query= 94-0530e20.Seq -20.5 sequence of clone "E" (AP29.1)
(184 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
gp|X79050|GGGFBPB_1	gammaFBP-B [Gallus gallus]	-1	41	0.31	2
gp|X79051|GGGFBPA_1	gammaFBP-A [Gallus gallus]	-1	41	0.32	2
gp|X79011|GGGFBPC_1	gammaFBP-C [Gallus gallus]	-1	41	0.33	2
pir|S07571|S07571	twitchin - Caenorhabditis elega...	-3	44	0.36	2
sp|P35193|YOL1_YEAST	HYPOTHETICAL 47.6 KD PROTEIN IN...	-1	43	0.38	2

****

Query= -20.5 sequence of clone "G" (AP29.2) 94-0530g20....
(345 letters, both strands)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate
128,149 sequences; 36,541,332 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Sum Probability P(N)	N
gp|M15816|SYNTOX1_1	Synthetic diphteria tox gene DNA,...	+2	65	0.088	1
sp|P27031|RHAA_SALTY	L-RHAMNOSE ISOMERASE (EC 5.3.1.14...	+2	65	0.58	1
gp|Z23277|CDTOXBA_1	toxin B [Clostridium difficile]	-3	56	0.65	3
sp|P12526|MAS_RAT	MAS PROTO-ONCOGENE. >pir|A31816|T...	-3	52	0.97	2
sp|P04201|MAS_HUMAN	MAS PROTO-ONCOGENE. >pir|A01375|T...	-3	52	0.97	2

Query= 94-0630i20.Seq -20.5 sequence of clone "I" (AP29.3)
(190 letters)

Translating both strands of query sequence in all 6 reading frames

Database: Non-redundant PDB+SwissProt+SPupdate+PIR+GenPept+GPupdate, 5:02 AM EDT Sep 9, 1994
127,124 sequences; 36,206,892 total letters.

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Smallest Poisson Probability P(N)	N
gp|L13744|HUMAF9X_1	AF-9 gene product [Homo sapiens]	+1	65	1.5e-09	2
gp|U02548|AAU02548_1	vitellogenin [Aedes aegypti]	+1	69	5.0e-06	2
gp|U00691|U00691_4	G5/D6 ORF [Dictyostelium discoi...	-3	48	6.3e-06	3
sp|P17671|E75A_DROME	ECDYSONE-INDUCIBLE PROTEIN E75-...	+1	54	1.4e-05	2
sp|P17672|E75B_DROME	ECDYSONE-INDUCIBLE PROTEIN E75-...	+1	54	1.6e-05	2

Appendix C:
Complete cDNA Sequencs of Differential Display Products.

> -20.5 sequence of clone "A" (AP21.1) 94-0530a20.Seq

CGAGTATGAGCCCAGATCTAGGGGGACTTTAACAGCAACGTGTAAACAAATATGGAACATTCAAGAAAAGGAAATGCTCATAACCCAG
GGAGATCAGAAACATATTACACAACATATNATAAAATGTGTTTCTTCAGTTCTCTTCATTCACATCATACTCCTCATACTCG

> -20.5 sequence of clone "L" (AP26.2) 94-0630l20.Seq

GAAGCTGAGGCAATTGTTACCAATCATTTGGTTGTACGAGGCAGCTATCGAAGCCTTAGCTTGGTTGTGTATGGAAACACCACTGAAG
ATCTTGGGCAATTCAACATAGATGTCGACTTAGATGGCTCTTTGGCTAATACTGTTAGTGTTGTTGAGGGAGACCTTGAAGACCTCCC
ACCTGCACTGCGTCCAAATAATCTATCTACTGAGCAAACATTGTCTTCCTTGAAATCATTATCTTTTGAAAAGTATTCCCTCAGCTTC

> -20.5 sequence of clone "M" (AP26.2) 94-0630m20.Seq

GAAGCTGAGGGTTCACAAATGGATATTAGAATAAGTCACGCTAGTGTAAGCTTGCACTTATCCGATAATATTGACGTTTTAAGACAAC
TACTAGTCTATCCACTACAGCAATTTGGAGCAAAACAGAAAACTGTAAGACCACAGACTGGCAGTATAACTGCAATTCACACAATTGA
TCTACTAGAAAGGATGTCATGCCTACTTGATCTTCTTCTTTGGCCTCAGCTTC

> -20.5 sequence of clone "N" (AP26.3) 94-0530n20.Seq

GAAGCTGAGGATTAACGGAAGGAAGATTCAAGCCACTGCAACAGTGCCATTTCGTACAGCTTTATTTGCATCAGTTTCTTTAGAGGGG
GTCTCTGTTTTTGTTCTTGAGGATCTCTTTGATGCAGAAATTTTCCATTTTCTTTCANTAGNTTCGNCCANTTCTATCTGTTCTTGCC
TGCATTCTTGGCTACAAAATGGTGTATCCCCTCAGCTTC

> -20.5 sequence of clone "O" (AP26.3) 94-0530o20.Seq

GAAGCTGAGGGGTTAAGGGTATGGCTGGTCTTGGAAGAACAAGAATATCTCTTCCTTCTCAATTCTCAGCTGAATTTAGTTTCCCTAG
AAAATTTGCTGTTTGTTTGANCTCTAGCAACTCTAAGGGTGTTGNACTNTTTNGGAGATGGTCCTTATTTTTTCCTTCCAAATAGGGA
ATTTTCAAATGATGATTTTTCCTACACTCCACTTTTTATCAATCCAGTTAGCACAGNCTCAGCTTCAATCACTAGTGCGGGCGCCTGC
AGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCAACG

> -20.5 sequence of clone "P" (AP26.4) 94-0630p20.Seq

GAAGCTGAGGATGANGATGATGATGANGAGATTCANGAAAACAGAGAGAGAATAGTGAANAANAGAGTGATCAAATTGAGAGAGAGAA
GCAGTGTTATTTGTTCANTTANGCCAAACTGAAAACTATGGTCTCACAGAGCAGACAAATCCCATTANTAACCAACAATGGAGAGCTT
GAAGCCAGCAAAGCAGAGGACCATTTTTGGGTTCCACCCCACCACAACTCCCCACCCCTCAGTTTCANTCACTAGTGCGGTCGCCTGC
AGGTCGACCATATGGGAGAGCTCCCANCGCGTTGGC

> -20.5 sequence of clone "Q" (AP26.4) 94-0630q20.Seq

GAAGCTGAGGGTTCACAAATGGATATTAGAATAAGTCACGCTAGTGTAAGCTTGCACTTATCCGATAATATTGACGTTTTAAGACAAC
TACTAGTCTATCCACTACAGNAATTTGNAGCAAAACAGAAAACTGTAAGACTACAGACTGGCAGTATAACTGCAATTCACACAATTGA
TCTACTAGAAAGGATGTCATGCCTACTTGATCTTCTTCTTTGGCCTCAGCTTC

> -20.5 sequence of clone "E" (AP29.1) 94-0530e20.Seq

GGAATGCCATAAAAGGTGCATTTTTACTCGTTATATTGTTGGTTTTGTGGTAAAAAATGAAAACTGAGGATTCTTCCTTCTCGATGGT
TACTGTTATTGGAGGGGCTAATGATATTTGAGGTAAAGTTGGCTATTTTGGGTTGGTGGGGTCTNCCTCTCTTGAGCTATGGCATAAT
GGCATTCC

> -20.5 sequence of clone "G" (AP29.2) 94-0530g20.Seq

GGAATGCCATGGCCACATCAAAGAAAGTAATCACTAGAGATAATTGGGAGAAAAGGCTCAATGATGTAAAGATTAGAAAAGAAGATAC
AGATAGATTGGTAATGAATTTTCTAGTTACAGAGGGTTATGTAGAAGCTGCAGAAAAGTTCAGAATAGAATCTGGGACTGATCCTGAT
ATGGATCTTGCTACCATCACGGATCGGATGGCATTCC

> -20.5 sequence of clone "I" (AP29.3) 94-0630i20.Seq

GGAATGCCATATACTACTTCTAGTTAAACCAATTTGNAATGATCAAAATTTACATAAGCTTAGAGAACTAGTCATTTCAAAAATCTAC
ACCAAAAGCTCTATTCATCATCTACTCCTATCTCATCTTTCACTTCCTCATCTACTGCTATCTCTTTCTCTTCCTTCTCATCATCTTC
ATCAATGGCATTCC