TRANSCRIPTIONAL PROFILING OF *ARABIDOPSIS THALIANA* RESPONSES TO TNT AND RDX: A ROUTE TO PLANT IMPROVEMENT FOR PHYTOREMEDIATION

by

DREW ROBERT EKMAN

(Under the Direction of Jeffrey F.D. Dean)

ABSTRACT

Serial Analysis of Gene Expression (SAGE) was used to profile transcript levels in *Arabidopsis thaliana* roots and assess their responses to 2,4,6-trinitrotoluene (TNT) and 1,3,5-hexahydro-1,3,5-triazine (RDX) exposure. SAGE libraries representing control, TNT, and RDX-exposed seedling root transcripts were constructed, and each was sequenced to a depth of roughly 30,000 tags. More than 19,000 unique tags were identified overall. Upon TNT exposure, the second most highly induced tag (27-fold increase) represented a glutathione S-transferase. Several cytochrome P450 enzymes, as well as an ABC transporter, were also highly induced by TNT exposure. These analyses also revealed an oxidative stress response upon TNT exposure. RDX exposure also induced a number of transcripts including: molecular chaperones and transcription factors, as well as vacuolar proteins and peroxidases. Comparison of these transcripts to those induced by TNT, identified significant differences, suggesting drastically different mechanisms for the metabolism of these two compounds. Preliminary DNA microarray analyses were also used to verify the expression levels of transcripts affected by TNT as

identified previously by SAGE. Identification of transcriptome-level responses to TNT and RDX exposure will better define the metabolic pathways plants use to detoxify these xenobiotic compounds, which should help us improve phytoremediation strategies directed at these and other nitroaromatic compounds.

INDEX WORDS: Phytoremediation, Serial Analysis of Gene Expression, Microarray, Munitions,Explosives, Bioremediation, SAGE

TRANSCRIPTIONAL PROFILING OF *ARABIDOPSIS THALIANA* RESPONSES TO TNT AND RDX: A ROUTE TO PLANT IMPROVEMENT FOR PHYTOREMEDIATION

by

DREW ROBERT EKMAN

B.S., The University of Georgia, 1997

M.S., The University of Tennessee, Knoxville, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

TRANSCRIPTIONAL PROFILING OF *ARABIDOPSIS THALIANA* RESPONSES TO TNT AND RDX: A ROUTE TO PLANT IMPROVEMENT FOR PHYTOREMEDIATION

by

DREW ROBERT EKMAN

Major Professor: Jeff Dean

Committee: Alan Przybyla
Claiborne Glover
Russell Malmberg
Scott Merkle

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

DEDICATION

To Lis

ACKNOWLEDGEMENTS

There are several to whom I owe a sincere debt of gratitude. First and foremost I would like to thank my major professor Jeff Dean and my EPA sponsor Lee Wolfe, both of whom were constant sources of help and encouragement. Their mentorship was instrumental in the successful completion of this work. I'd also like to thank Alan Przybyla and Walt Lorenz who in addition to sharing their keen scientific insight and technical knowledge to problems I encountered also shared their genuine enthusiasm for science and maybe most importantly their good humor. Claiborne Glover, Scott Merkle, Alan Przybyla, and Russell Malmberg were kind enough to serve on my committee and for that I am grateful. Their wisdom and critical thinking served me well. Sincere thanks also to Macarthur Long at the EPA who not only provided guidance and support during my tenure with the EPA, but in addition to Lee, ensured that I always had what I needed to complete my work. I owe a debt of gratitude to Steve McCutcheon also at the EPA for being my initial link with the NNEMS (National Network for Environmental Management Studies) fellowship program and for his continued support of the project. That being said, I'd also like to thank the NNEMS program and the U.S. EPA for supporting me for the majority of my time spent pursuing this degree. Deepest thanks also to my family who supported me throughout my years in graduate school and who always displayed genuine interest in my work. Thanks also to Katie Ekman, a canine of 14 years who was a constant fixture at my feet during the writing of this dissertation. And finally, to my beautiful wife Lisa, who has been there for the entirety of both of our

undergraduate and graduate degrees, sharing many late nights up studying. Thank you for being the incredible that person you are. Thank you for everything.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

EXPLOSIVES 2

TOXICITY 4

ENVIRONMENTAL CONTAMINATION BY EXPLOSIVES AND METHODS OF REMOVAL 6

BIOREMEDIATION 8

BIOREMEDIATION OF TNT AND RDX 9

PHYTOREMEDIATION 12

PLANT METABOLISM OF TNT AND RDX 15

GENOMICS 20

TRANSCRIPTIONAL PROFILING 21

PHYTOREMEDIATION AND GENOMICS 29

REFERENCES 32

2 SAGE ANALYSIS OF TRANSCRIPTOME RESPONSES IN *ARABIDOPSIS THALIANA* ROOTS EXPOSED TO TNT 49

ABSTRACT 50

INTRODUCTION 51

RESULTS....52
DISCUSSION....56
EXPERIMENTAL PROTOCOL....63
REFERENCES....69
3 EFFECTS OF THE EXPLOSIVE RDX ON TRANSCRIPT EXPRESSION IN *ARABIDOPSIS THALIANA* SEEDLING ROOTS....82
ABSTRACT....83
INTRODUCTION....83
MATERIALS AND METHODS....85
RESULTS AND DISCUSSION....90
CONCLUSIONS....96
REFERENCES....105
4 MICROARRAY ANALYSIS OF *ARABIDOPSIS* TRANSCRIPTOME CHANGES INDUCED BY TNT....110
INTRODUCTION....111
MATERIALS AND METHODS....113
RESULTS AND DISCUSSION....118
REFERENCES....126
5 CONCLUSIONS....128
REFERENCES....135

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

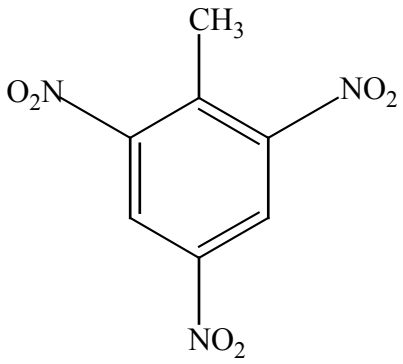
2,4,6-trinitrotoluene (TNT) and hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) are compounds widely used as explosives by the military. The great utility of these compounds is a result of their high stability (they are relatively insensititive to the addition of heat, flame, spark, or shock) and the characteristically large amounts of energy released upon their detonation. During the First World War, little consideration for the potential toxicity of these compounds to the environment was exercised. However, during World War II, when the scale of production increased dramatically, concern over the toxicity of these compounds emerged. Unfortunately by this time, extensive contamination of land and water surrounding the sites of their manufacture, storage, and disposal had already occurred. Research to assess the toxicity of these compounds to humans and other organisms soon revealed the need to remove contaminating explosives from the environment. In response, large-scale methods of remediation were developed. However, these methods proved environmentally destructive, laborious, and costly. For this reason, alternative methods of remediation are being sought. Alternatives using plants (phytoremediation) have been shown to remove explosives effectively from contaminated environments with the advantage of producing little impact on the surrounding landscape.

EXPLOSIVES

In contrast to primary explosives and low explosives, both of which are easily detonated by the addition of heat, flame, spark, or shock, TNT and RDX are considered high explosives due to their relative insensitivity to these conditions. It is through the detonation of primary or low explosives (e.g., gunpowder) that high explosives are detonated to release large amounts of energy. Although TNT and RDX are both often classified as nitroaromatic explosives, TNT is the only one of these to possess true aromaticity (Figure 1.1). Originally synthesized by Wilbrand in 1863, TNT is produced by the sequential nitration of toluene to nitrotoluene, dinitrotoluene (DNT), and finally TNT, using a mixture of nitric and sulfuric acids (Palmer et al., 1996). The relatively high solubility of TNT in water (~156 mg/L at 25^0C) leads to the production of large volumes of contaminated wastewater during TNT production and munitions loading operations. In the past, this contaminated water was collected in holding lagoons, or in some cases released directly into surface waters. Current contaminated sites typically contain average concentrations of 10,000 mg/kg TNT in soil and 156 mg/L TNT in water (Fernando et al., 1990).

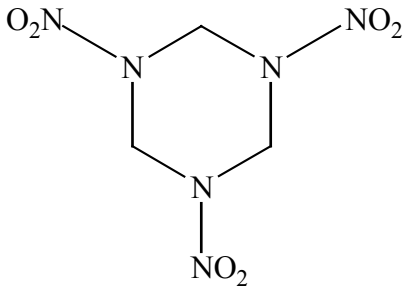
Used as the main component in various explosive compositions for artillery shell charges or plastic explosives, RDX (an acronym for **R**oyal **D**emolition E**X**plosive or **R**esearch **D**epartment E**X**plosive) is technically a cyclic aliphatic nitramine (Figure 1.1). It is considered the most powerful of the military high explosives and, as such, is widely used. During the late 1970s, RDX production was second only to TNT among the high explosives (Ryon et al., 1984). Developed during World War II, RDX is synthesized through either the direct nitration of hexamine using the Woolwich process or through

indirect nitration using the Bachmann process (Yinon, 1990). The solubility of RDX is less than that of TNT (~60 mg/L), which has helped minimize its spread through aquatic systems.



2,4,6 trinitrotoluene

(**TNT**)



hexahydro-1,3,5-trinitro-1,3,5-triazine

(**RDX**)

Figure 1.1. Explosive nitroaromatic compounds of concern to the Department of Defense. TNT and RDX are highly recalcitrant, toxic compounds that are widespread in soil, as well as ground and surface water, at military installations.

TOXICITY

During the early stages of World War I, large-scale manufacture of TNT resulted in thousands of cases of "TNT intoxication" and hundreds of deaths (Palmer et al., 1996). With the onset of World War II, production was increased substantially. However, far fewer TNT-related illnesses and deaths were reported because of an alteration in the production process that removed a hazardous byproduct, tetranitromethane (TNM), from the final product (Bucher, 1990). Despite this, TNT toxicity remains a significant problem as exposure to even low doses can have severe long-term consequences due to the compound's mutagenic properties.

Low-level TNT intoxication in humans is characterized by gastrointestinal disorders, sometimes with associated tightening of the chest. In cases of severe intoxication, cyanosis (a bluish or purplish tinge to the skin and mucous membranes) is produced as a result of methemoglobin formation. (Methemoglobin is a brownish-red, non-functional form of hemoglobin that occurs when hemoglobin is oxidized either during decomposition of the blood or by the action of various oxidizing drugs or toxic agents). Later stage symptoms include liver damage, aplastic anemia, cataract formation, menstrual disorders, and neurological disorders (Ermakov et al., 1969). As noted above, TNT has been reported to be mutagenic, even in the absence of metabolic activation, producing frame-shift reversal mutations in *Salmonella typhimurium* (Dilley et al., 1978; Ellis et al., 1978b). Not surprisingly, TNT is a carcinogenic substance producing various tumors in the bladders of rodents (Furedi et al., 1984a). Its toxicity to aquatic organisms, including marine copepods (*Tigriopus californicus*), oysters (*Crassostrea gigas*), and

freshwater unicellular green algae (*Selenastrum capricornutum*), is well established (Won et al., 1976), as toxicological studies using these organisms revealed that TNT concentrations as low as 2.5 to 10 mg/L could inhibit physiological functions. As a result of this, the recommended maximum drinking water concentration for TNT is 140 μg/L (Stahl and Aust, 1995).

Numerous studies have confirmed adverse effects of RDX exposure on the nervous system, liver, and kidneys, as well as induction of testicular degeneration and inflammation of the prostate in mice and rats. Inhalation is known to induce nausea, irritability, confusion, and convulsions in exposed humans (Kaplan et al., 1965; Ketel and Huyhes, 1972). In addition, hypotriglyceridemia, behavioral changes, and mortality have also been observed in F344 rats fed toxic doses of RDX (Levine et al., 1990a). Due to this acute toxicity in rodents, RDX has often been employed as a rat poison. Although not thought to be mutagenic, one study reported a dose-related increase in hepatocellular carcinomas and adenomas in female mice (Lish et al., 1984) suggesting carcinogenic potential of RDX. Moreover, hexahydro-1,3,5-trinitroso-1,3,5-triazine, a degradation product of RDX has been used as an experimental tumorigen (Lewis, 1992). RDX has also been found to be quite toxic to various aquatic species (Bentley et al., 1977). Bluegill used in acute exposure tests, displayed an LC_{50} of 3.6 mg/L after 96 hours of exposure. In the same report, chronic exposure studies using fathead minnows produced LC_{50} values of 6.3 mg/L. The toxicity of RDX to plants has also been evaluated (Cataldo et al., 1990). At soil concentrations of 50 mg/kg RDX, bush beans (*Phaseolus vulgaris*), wheat (*Triticum aestivim*), and bromegrass (*Bromus inermis*) became chlorotic and displayed reduced plant heights, indicating a toxic response. As a result of this

widespread toxicity, water concentrations of RDX above 105 μg/L are considered unacceptable.

ENVIRONMENTAL CONTAMINATION BY EXPLOSIVES AND METHODS OF REMOVAL

The widespread use of explosives by the military has led to extensive contamination of soil, as well as surface and ground water at processing and manufacturing installations throughout the United States and Europe. At least 50 sites identified by the U.S. Army Environmental Center (USAEC) (Figure 1.2) are in need of remediation (Jerger and Woodhull, 2000), and a large number of other contaminated sites have been identified in the U.K., Canada, and Australia. Clean-up efforts are underway at many sites in the U.S., but relatively little clean-up is in progress in other countries where the problem has only recently been recognized. In Germany the situation has been further complicated by the demolition of explosives manufacturing facilities after World War II, followed by the reconstruction of industrial and residential complexes over these sites. Very little characterization has been done on these sites, and only a few projects aimed at cleaning up these locations have been initiated.

Contamination became widespread at manufacturing and processing facilities due to the rush to produce these explosives during the World Wars. In many instances, the explosives leached from disposal lagoons into the surrounding soil. In other cases, TNT-laden wastewater or "pink water", so named for the color TNT imparts to aqueous solutions, was released into surface waters, thereby spreading the contamination.

Figure 1.2: Facilities in the United States contaminated with explosives. (Spain 2000).

Obviously, removal of these compounds from the environment is of high priority to the Department of Defense. Methods for removal have been developed, but all possess undesirable characteristics and are in need of improvement. Typically, soil at highly contaminated sites is treated through excavation followed by incineration or composting. These strategies are not only laborious and costly, but concerns regarding toxicity and mutagenicity of the transformation products resulting from these processes remain (Tan et al., 1992; Jarvis et al., 1998). In addition, incineration produces unusable ash and has been met with public opposition due to concern over air emissions. Estimated costs for using incineration range between $200 and $1000 per cubic yard (Vanderford, 1996). On top of this, excavation of contaminated soils is an ecologically destructive practice, leaving sites heavily disturbed and adversely impacting nearby ecosystems.

Water polluted by the presence of explosives can be effectively treated using carbon adsorption columns, but the spent carbon cannot be safely regenerated. This prevents reuse of these materials and consequently drives up the cost of treatment. Due to the inherent destructiveness and/or costs associated with these current methods, researchers are seeking less intrusive, more environmentally friendly strategies to remove TNT and RDX from the environment.

BIOREMEDIATION

An effective and potentially less destructive, as well as less costly, alternative to excavation is bioremediation--the use of organisms to remove or degrade pollutants from contaminated sites. In terms of costs, bioremediation offers a significant reduction for

clean-up projects. It has been estimated that microbial bioremediation in general (not explosives exclusively) can cost as little as $25/cubic yard, compared to methods, such as pump/treat or soil excavation, which typically cost $50-$100/cubic yard (Watanabe, 2001). Despite being a relatively new technology, many organisms have been identified that possess abilities for contaminant degradation and removal. These include numerous species of bacteria, certain fungi, and a number of plants.

BIOREMEDIATION OF TNT AND RDX

Bacteria

A large number of microbes have been found capable of metabolizing explosives and both aerobic and anaerobic degradation mechanisms have been observed. Aerobic metabolism of TNT often occurs through the removal of nitro groups from the toluene ring following the formation of hydroxylamine aromatic compounds (e.g., 2-hydroxylamino-4,6-dinitrotoluene). Alternatively, the formation of a Meisenheimer complex, in which the ring is reduced via addition of a hydride ion (contributed by reduced pyridine nucleotides) followed by nitro group loss (Esteve-nunez et al., 2001), can also lead to nitro group removal. In general, aerobic bacteria have difficulty metabolizing TNT to a significant extent, and few studies have reported the complete mineralization of TNT in aerobic microbial cultures (Traxler et al., 1974; Bae et al., 1995). Only a handful of studies have reported the use of TNT as a carbon or nitrogen source by bacteria (Duque et al., 1993; Montpas et al., 1997; Oh and Kim, 1998).

Anaerobic metabolism of TNT by prokaryotes involves the complete reduction of nitro groups to form 2,4,6-triaminotoluene (TAT), typically using H_2 as an electron donor. TAT formation in aerobic systems is highly unlikely due to the requirement for a reduction potential around –200 mV. Although it is widely believed that TAT is further metabolized through reductive elimination of the amino groups (Schnell and Schinck, 1991) there are no reports of deaminating reactions or toluene formation in anaerobic cultures supplemented with TNT. However, strains of *Desulfovibrio sp*. have been observed to grow on TNT as the sole nitrogen source (Boopathy and Kulpa, 1992; Preuss et al., 1993).

Whereas extensive research has been conducted on the metabolism of TNT in bacterial systems, very little information has been obtained for RDX. However, it appears as though some microbes are capable of effectively degrading this explosive. Under anaerobic conditions, RDX was metabolized through successive reduction of the nitro substituents to hydroxyl-amino derivatives (the formation of hydrazine and dimethylhydrazine, both known mutagens, was observed in the culture medium) (McCormick et al., 1981). RDX degradation was also observed in liquid cultures of *Desulfovibrio sp*. in which ammonia derived from the compound's nitro groups served as the sole nitrogen source (Boopathy et al., 1998a). Although there has been much speculation on the metabolic mechanisms responsible for the degradation of RDX by bacteria, virtually nothing has been confirmed.

Fungi

The discovery that certain fungi have the ability to mineralize TNT has made this class of organisms a major contender for the remediation of sites contaminated with TNT, RDX, and other explosives. The first observation that fungi possess this ability occurred in a study of *Phaenerochaete chrysosporium* grown under nitrogen-limiting conditions (Hawari et al., 1999). Subsequent work showed that TNT may be mineralized under a variety of nutrient starvation conditions (carbon, nitrogen, or sulfur limitation), and that the reactions were carried out by enzymes secreted as part of the lignin-degrading system of this white rot fungus (Stahl and Aust, 1995). As was seen in studies with bacteria, reduction of the TNT nitro groups was the initial step in fungal degradation of this compound, and in *P. chrysosporium*, this step appears to occur through the activity of a membrane-bound nitroreductase that requires NADPH as a cosubstrate (Rieble et al., 1994). The reaction is sensitive to molecular oxygen. Degradation continues under lignolytic conditions, ultimately resulting in complete mineralization of TNT, most likely through the activity of lignin peroxidases (Michels and Gottschalk, 1995; Hodgson et al., 2000). Despite this ability to mineralize TNT, *P. chrysosporium* tolerates relatively low levels (<20 mg/kg) of the explosive, making field application of this organism unrealistic. Other fungal species display a greater level of tolerance to TNT (e.g., *Cladosporium resinae*), but do not appear capable of mineralizing the compound (Bayman and Radkar, 1997).

Little is known of the potential for fungi to degrade RDX. One study reported the mineralization of RDX under aerobic conditions, although the mechanism of degradation was not described (Fernando and Aust, 1991).

Plants

Since the mid-1990s, phytoremediation, or the use of plants for the restoration of contaminated soils, groundwater, or surface water, has been recognized as a potentially effective technology (Burken, 2000). Although a relatively new technology, phytoremediation is already showing promise for the clean up of a wide variety of pollutants and may well prove effective for the removal of explosives from contaminated soil and water. Not only is phytoremediation an efficient method for contaminant removal, it has the advantage of being environmentally friendly and offers substantial financial savings. For example, where excavation and incineration may cost $200-$1,500/ton, and *in situ* bioremediation may cost around $50-$100/ton, phytoremediation may cost as little as $10/ton of soil (Watanabe, 2001). The promise of this technology is significant and has stimulated extensive research into its potential for addressing a number of the most threatening and widespread pollutants.

PHYTOREMEDIATION

Phytoremediation has become an attractive alternative for environmental cleanup for several reasons: 1) plants quickly stabilize surface soil, limiting the runoff of

contaminants into nearby streams; 2) plant roots can penetrate soil to reach and extract contaminants buried in the subsoil; 3) plants provide their own energy for the requisite degradative processes; 4) plant root exudates can stimulate the growth of soil microbes that further contribute to the breakdown of organic contaminants; and 5) plants have the potential to be deployed on disturbed sites in a cost-effective manner that minimizes further disruption of the surface soil and consequent spreading of the contaminant (e.g., aerial seeding).

The number of contaminant attenuation mechanisms possessed by plants makes their use in remediating contaminated land and water feasible. As a result of their sedentary nature, plants have evolved diverse abilities for dealing with toxic comounds in their environment. It is thought that plants have the ability to "synthesize, rearrange, and detoxify the most complex array of biochemicals and biopolymers of any living organisms" (Meagher, 2000). Furthermore, "plants control most of the energy in an ecosystem, and usually account for several orders of magnitude greater biomass than any few bacterial species in the soil" (Meagher, 2000).

Indirect mechanisms involving plant-mediated modifications to the immediate surrounding environment add to these abilities. It is well documented that plant-induced changes to the contaminated environment (alteration of redox conditions, shifts in microbial community composition, and modificiation of site hydrodynamics) serve as an indirect means of remediation. For example, plants may secrete 10-20% of their photosynthate in root exudates to stimulate the growth of diverse bacterial and fungal communities in the rhizosphere (Lynch, 1982; Haselwandter, 1983; Shimp et al., 1993; Anderson et al., 1994). Populations of these organisms are significantly greater (two to

four orders of magnitude) in vegetated versus non-vegetated soils. In addition, these specific rhizosphere-associated organisms typically possess a wide range of metabolic capabilities, including those required for the degradation of recalcitrant xenobiotic compounds (Anderson et al., 1994; Walton et al., 1994). Typical root exudates include phenolic compounds, organic acids, alcohols, and various proteins that serve as carbon and nitrogen sources for the bacteria and fungi. These plant products provide not only for the growth and long-term survival of soil microbes, but also serve as a supply of co-metabolites, such as catechin and coumarin, that facilitate microbial breakdown of compounds like poly-chlorinated biphenyls (PCBs) (Bedard et al., 1987; Donnelly et al., 1994; Fletcher and Hedge, 1995; Hedge and Fletcher, 1996). It is for this reason that phytoremediation technology is often described as "plant-assisted" remediation (Burken et al., 2000)

With respect to their direct roles in remediation processes, plants use several different strategies for dealing with environmental chemicals: phytoextraction, phytodegradation, phytovolatilization, and rhizodegradation (Schnoor, 1997). Phytoextraction involves the removal and subsequent storage of contaminants by the plant. This term is frequently applied to the removal and storage of heavy metals that may undergo electrochemical transformation in plants, but cannot be degraded. However, certain organic chemicals may also be treated in this manner due to inherent resistance to degradation. Conversely, phytodegradation describes processes in which plants metabolize the contaminants they take up. Components of this mechanism are often utilized by plants exposed to herbicides and thus have been researched extensively. The metabolic processes involved in phytodegradation have strong similarities to those

used by animals for modification and degradation of drugs and other toxins. This has given rise to a conceptual model for phytodegradation known as the "green liver" model, which will be discussed below later in more detail (Sandermann, 1994). A further attenuation mechanism, referred to as phytovolatilization, involves the release of contaminants to the atmosphere following their uptake from the soil or water. This mechanism has been observed for both organic and heavy metal contaminants, including trichloroethylene (TCE), which has been observed in the off-gas from plant leaves in the laboratory and field (Compton et al., 1998), and in the production of volatile, elemental mercury by genetically-engineered *Arabidopsis thaliana* grown in the presence of ionic mercury (Rugh et al., 1996; Bizily et al., 1999).

An indirect mechanism, rhizodegradation refers to the transformation of contaminants by resident microbes in the plant rhizosphere (i.e., the microbe-rich zone in intimate contact with the root vascular system). As mentioned above, the presence of plants on contaminated sites can drastically affect soil redox conditions and organic content (often through the secretion of organic acids from roots), as well as soil moisture. Rhizodegradation is the dominant mechanism in the removal of total petroleum hydrocarbons from soil by deep-rooted trees (Carman et al., 1998), as well as annual species (Schwab and Banks, 1994).

PLANT METABOLISM OF TNT AND RDX

Numerous studies have been conducted to understand the fate of TNT in terrestrial and aquatic plants (Dacre and Rosenblatt, 1974; Smock et al., 1976; Palazzo

and Leggett, 1986; Cataldo et al., 1990; Harvey et al., 1991; Gorge et al., 1994; Wolfe et al., 1994; Medina et al., 1996; Medina, 1996; Best et al., 1997; Larson, 1997; Vanderford et al., 1997; Lauritzen, 1998; Scheidemann et al., 1998; Thompson et al., 1998; Best et al., 1999; Bhadra et al., 1999a, b; Larson et al., 1999b; Larson et al., 1999a; Wayment et al., 1999). TNT was never recovered stoichiometrically in these studies, suggesting the existence of metabolic pathways for its transformation and possible degradation. In contrast, RDX typically accumulates in plant tissues with little evidence of degradation.

TNT Metabolism

Although the metabolism of TNT in microbial systems has been studied extensively, until recently little was known about its metabolism in plants. Studies of TNT metabolite mass balances confirmed the existence of plant mechanisms for the uptake and transformation of this explosive (Hughes et al., 1997). Researchers employed uniformly labeled ^{14}C-TNT to determine the fate of TNT in *Catharanthus roseus* hairy root cultures, axenic *Myriophyllum aquaticum* plants, and wild-grown *Myriophyllum aquaticum*. In all three systems, the TNT was completely transformed, although mineralization was not observed. The aminodinitrotoluenes (2-amino-4,6-dinitrotoluene and 4-amino-2,6-dinitrotoluene) were observed in the growth media, as had been observed in microbial studies. However, unidentified ^{14}C-labeled soluble and extractable products were observed, but could not be identified, as reduction products (i.e., aminonitrotoluenes, hydroxyl-aminonitrotoluenes, or azoxynitrotoluenes). In addition, radiolabel also ended up in bound residues (i.e., a non-extractable plant-associated

material that could only be quantified after combustion of the plant tissue). The soluble and solvent-extractable products comprised the majority of TNT derivatives. The production of unidentifiable soluble products hinted at the involvement of plant conjugation reactions. This notion was later confirmed through the identification of conjugates in which various six-carbon sugars were bound to the TNT transformation products (Bhadra et al., 1999a).

The identification of various TNT transformation products and conjugates lead to conceptualization of the "green liver" model for plant metabolism of explosives (Sandermann, 1994). This model was born out of the idea that plants use "detoxification" processes similar to those observed in the metabolism of xenobiotics by animals. This is in contrast to microbial heterotrophic metabolism in which foreign compounds are typically broken down for energy via the diverse catabolic pathways available to microbes. Being autotrophic, plants do not possess this range of catabolic pathways and consequently tend to detoxify xenobiotics rather then utilizing them as carbon or nitrogen sources. It is through the multi-phase mechanism proposed by the green liver model that this detoxification is often achieved.

Studies of plant metabolism of herbicides and other xenobiotics first revealed the existence of this multi-phase detoxification mechanism (Klein and Scheunert, 1982). In the initial phase of detoxification, following entrance of the compound into the cell, transformation reactions introduce chemical groups amenable to subsequent conjugation reactions. This is frequently achieved through either enzyme-catalyzed oxidation reactions involving cytochrome P450 enzymes and peroxidases (Bhadra et al., 1999b) or through enymze-catalyzed reduction or hydrolysis reactions. Once transformed, the

metabolized derivative may be conjugated to glutathione (GSH) or six-carbon sugars by glutathione S-transferases (GSTs) or UDP:glucosyltransferases, respectively. This constitutes the second phase of detoxification. Introduced moieties (OH-, NH_2-, SH-, and COOH) on transformed molecules are typically conjugated to sugars (Frear, 1976), while conjugated double bonds, halogen- or nitrofunctions typically promote conjugation to glutathione. In the third phase of detoxification, conjugates are sequestered into plant storage organelles and/or released to the extracellular space. At this point, the conjugate may be either immobilized through the addition of further substitutions or degraded. In most cases, the conjugates are rapidly removed from the cell, likely to prevent the inhibition of conjugating enzymes as it has been found that glutathione conjugates inhibit both GST and glutathione reductase (Schröder and Wolfe, unpublished). Transportation of conjugates is often achieved through the action of ABC-type ATPases that appear to recognize the sugar or glutathione groups, rather than the xenobiotic compound itself (Schroder and Collins, 2002).

RDX Metabolism

This multi-phase mechanism for TNT detoxification in plants has not been observed for metabolism of RDX. Although few studies have examined the fate of RDX in plants, it appears that the compound is much more resistant to transformation than TNT. Using *Catharanthus roseus* and *Myriophyllum aquaticum*, the disappearance of RDX occurred at a much slower rate than for TNT (Bhadra et al., 2000). Mass balance studies in *C. roseus* showed that the majority of ^{14}C-labeled RDX remained

untransformed after entering the plant (Bhadra et al., 2000). This phenomena was previously noted by Harvey et al. (1991), in a study of bush beans grown for several days in a hydroponic solution containing 10 ppm RDX. The plants took up 60% of the available RDX in this time, but the compound remained virtually untransformed in the plant (Harvey et al., 1991). However, after 60 days, a fraction (30%) of the RDX appeared to have undergone some sort of transformation as various polar derivatives were detected in root tissues, and 50% of the RDX label was no longer extractable. Analysis of foliar tissues revealed that RDX concentrations were in excess of 90 mg/kg, suggesting a mechanism for phytoextraction of this explosive. Similar results were seen in hybrid poplars where a large percentage of ^{14}C-radiolabeled RDX was identified in the leaves (Thompson et al., 1999). Accumulation of RDX in the aboveground portions of terrestrial plants growing in RDX-contaminated soil is of some concern for phytoremediation processes. Ingestion of these tissues by insects and herbivores could introduce the compound into foodchains, thereby reducing the benefits of a phytoremediation strategy for cleanup. The use of plants to remediate RDX-contaminated soil may also be limited by the inability of plants to remove RDX efficiently from soils having high organic content. It has been observed that while uptake rates are relatively high in soils having low organic content, uptake is drastically reduced in the same soils when organic matter is added (Chen, 1993).

While many plants have innate abilities to remove explosives and other organics from aqueous solutions (Best, 1997, 1999; Bhadra, 1999), few are robust enough to survive and remediate the high levels of contamination common to most of the sites known today. In an attempt to address this situation, researchers have designed

transgenic plants with enhanced abilities to tolerate and remove TNT from their environment (French, 1999; Hannink, 2001). However, current understanding of the biochemical mechanisms involved in the processes of uptake and degradation of TNT and RDX in plants is rudimentary and limits our ability to manipulate metabolism efficiently. Studies of plant-explosives interactions have been limited to toxicity evaluations and identification of metabolites, which allows only for inference of the enzymes that might be involved. If phytoremediation is to become an effective technology for the removal and degradation of explosives, a more refined knowledge of the mechanisms used by plants for these processes must be obtained. Elucidation of these mechanisms would facilitate engineering of plants designed for not only removal and degradation of explosives, but also for tolerance and persistence in environments where contaminant levels are too high to permit survival of wild-type plants. One promising strategy for identifying the pertinent metabolic pathways involves use of genomic techniques to determine changes in gene expression induced by contaminant exposure. The induction of specific plant genes in response to explosives should highlight specific enzymes responsible for conferring tolerance or degradative abilities to plants.

GENOMICS

The characterization of complete genome sequences for a number of organisms, and the emergence of new technologies has created a situation where the expression of large numbers of genes can be studied simultaneously under any chosen condition. For

many organisms (e.g., *S. cerevisiae*) total genome expression analysis is now routine (Lashkari et al., 1997; Wodicka et al., 1997), and for many higher eukaryotes, expression analyses covering a significant portion of the genome are possible. The massive amounts of data generated by these analyses can present difficult bioinformatics challenges, but the information that can be gleaned from such studies provides new insights into cellular and genome function (Meltzer, 2001). In addition to speeding discoveries of fundamental biological functions, such studies are leading to significant advances in our understanding of human disease, particularly cancers (Khan et al., 1998; Alon et al., 1999; Golub et al., 1999; Alizadeh et al., 2000; Bittner et al., 2000; Ross et al., 2000; Scherf et al., 2000). Currently, the most commonly used genomic techniques are aimed at measuring changes in the transcriptome. As transcription represents the initial expression of the genome, comparisons of transcriptome profiles for different tissues (e.g., treatment versus control) can be a powerful approach to the study of metabolism and cellular function.

TRANSCRIPTIONAL PROFILING

A comparison of transcript abundance between different tissues or treatments has proven to be an effective method for identifying differentially expressed genes. While it must be recognized that transcript levels do not always reflect the levels of corresponding proteins in the cell, and that post-translational modifications are common (Gygi et al., 1999), it is generally accepted that in the modulation of gene expression, the primary level of control resides at the level of transcription (Donson et al., 2002). Based on this assumption, a number of technologies have been developed to allow for the rapid

quantification of different gene transcripts simultaneously. Current methods are based on direct analysis of the transcriptome using DNA sequencing or fragment sizing, or indirect analysis using nucleic acid hybridization.

Methods for Direct Analysis of the Transcriptome

Nucleotide Sequencing-based methods

Methods of transcriptional profiling that rely on DNA sequencing include large-scale expressed sequence tag (EST) sequencing, serial analysis of gene expression (SAGE), and massively parallel signature sequencing (MPSS).

Large-scale Expressed Sequence Tag (EST) Sequencing

EST sequencing projects provide both the potential to discover previously unidentified genes and the ability to assess differential gene expression. By measuring the frequency of occurrence of a particular EST in each cDNA library, the induction or repression of that gene can be determined in relation to other libraries. This approach also provides long sequences that facilitate identification of gene function. Thus, this technology is not as dependent as other techniques on established sequence data for the organism under study. However, the large amount of DNA sequencing required makes this technology costly and its routine use for profiling unrealistic at present. Furthermore, size and sequence composition biases in cloning and cDNA synthesis limit

the comprehensiveness of this approach (although these limitations are not unique to EST technology).

Serial Analysis of Gene Expression (SAGE)

Issues of sequencing costs for in-depth analyses of EST libraries was addressed in part through development of the SAGE technology (Velculescu et al., 1995). Like EST sequencing, SAGE is an open-ended technique that allows for the analysis of all transcripts in a given sample and thus is an excellent tool for gene discovery. In brief, SAGE generates short (14-mer) cDNA fragments known as tags from the 3' end of transcripts. Although short in sequence, these tags usually provide enough information to identify the original transcript expressed in the tissue under study (Figure 1.3). The tags are then linked together in long chains called concatemers, which are subsequently cloned and sequenced. Using custom software, the tags can then be identified and counted to determine the number of times a specific tag appears within all of the sequences from a particular tissue sample.

One of the major drawbacks of the SAGE technology is the short length of the tags. This characteristic can sometimes make unambiguous gene identification difficult, although the potential for misidentification is drastically reduced by the fact that the tags are generated from a specific location within each transcript (the NlaIII restriction site closest to the 3' end of the transcript). SAGE is also most effective where a sequenced genome is available for tag identification. In addition, as SAGE measures the abundance of individual tags in a large pool of tags, the extent of sequencing can have a profound

effect on the validity of the data based on sampling effects. For example, an analysis of 1,000 tags will be much less reliable than an analysis of 20,000 tags (which can cost considerably more).

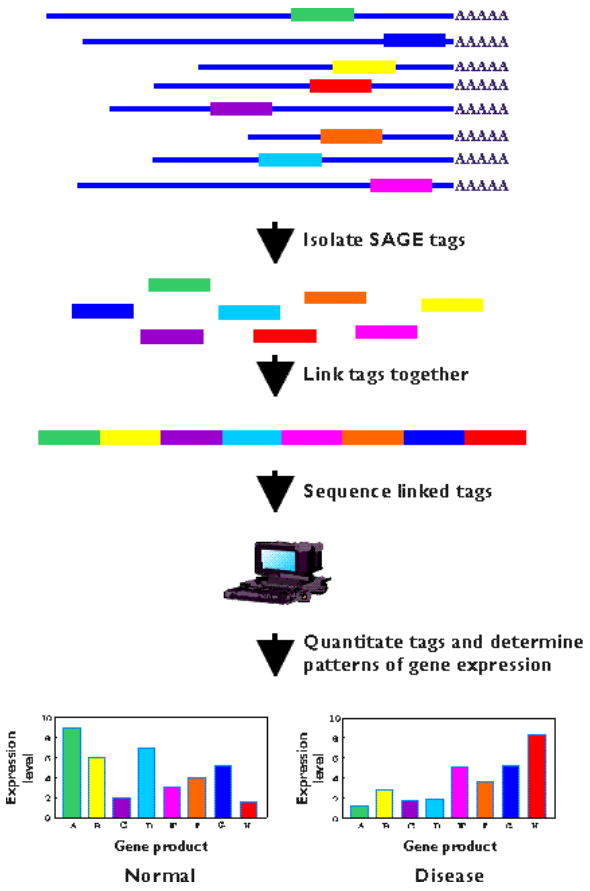


Figure 1.2. General diagram of the SAGE concept showing the linkage of tags together to form concatemers followed by sequencing and computational analysis. Diagram from the sagenet website: (http://www.sagenet.org/home/Description.htm)

Massively Parallel Signature Sequencing (MPSS)

Massively Parallel Signature Sequencing or MPSS provides a solution to both the time required for depth of analysis in EST sequencing projects and the lack of sequence data included in SAGE tags (Brenner et al., 2000a; Brenner et al., 2000b). The technique is based on use of microbeads for in vitro cloning of cDNA fragments generated from an mRNA population. This is achieved by synthesizing random oligonucleotides (32-mers) or "tags" from a defined set of 4-mers, and then attaching them to the generated cDNA fragments such that each fragment possesses a unique tag. These cDNA-tag conjugates are then amplified using PCR and subsequently hybridized to complementary sequences (anti-tags) attached to microbeads. As each microbead possess approximately 100,000 copies of a single anti-tag sequence, roughly the same number of cDNA-tag conjugates can be bound to one microbead. The cDNA fragments are then sequenced in a flow cell, and the number of beads displaying the same cDNA fragment are counted and summed for comparison. Typically sequences of 16-20 bases from each cDNA fragment are recorded, thereby increasing the amount of sequence data derived from each transcript over that obtained using the SAGE technique (Donson et al., 2002). As the procedure can be completed in a matter of a few days, MPSS combines both the digital nature of EST sequencing projects with a vastly superior throughput. In addition, the large number of sequence tags that can be measured (1,619,000 in a single study of human acute monocytic leukemia cells (Brenner et al., 2000a)) provides for more robust statistical analyses, producing more reliable data. The major drawback to this technology is its

proprietary nature. The technology is both owned and executed by Lynx Therapeutics, Inc., and the cost of access keeps it out of reach for most academic researchers.

Fragment sizing-based methods

Methods of direct analysis that rely on DNA fragment sizing include differential display (DD) and cDNA-AFLP (and variations of these). These methods utilize physical separation of cDNA fragments representing mRNAs through gel electrophoresis and selective PCR amplification.

Differential Display (DD)

The first application of fragment-sizing based methods for transcriptional profiling was differential display (DD) (Liang and Pardee, 1992). Differential display requires the reverse transcription of mRNA primed by anchored oligo dT primers containing a specific nucleotide at the 3' ends. These cDNA fragments are then amplified by PCR using the anchored primers and an oligonucleotide primer of arbitrary sequence. The PCR products are separated on a polyacrylamide gel and visualized using autoradiography (radioactive dATP is supplied during PCR amplification to allow for product visualization). Comparing the bands for mRNA populations originating from two different cell populations (i.e., different tissues, treatments, etc.), differences in banding patterns reveal differential expression of specific transcripts (Zhang et al., 1998). Although a relatively cost-effective method for the analysis of differential gene

expression, differential display possesses shortcomings that deter its continued use given alternative technologies. In particular, the PCR strategy approach attempts to amplify specific cDNAs using arbitrary primers at low annealing temperatures so as to allow for priming at multiple sites. Thus, amplification of cDNAs is based not only on their initial concentrations, but also on the quality of match between the primers and the template (Matz and Lukyanov, 1998). This situation results in detection of numerous false positive bands after electrophoresis (Sun et al., 1994; Sompayrac et al., 1995) and limits reproducibility (Haag and Raman, 1994; Zhang et al., 1998). In addition, the technique does not quantify the degree of differential expression, although its sensitivity does allow for detection of changes in relatively rare transcripts.

cDNA-AFLP

A modification of the original differential display protocol, cDNA-AFLP (**a**mplified **f**ragment-**l**ength **p**olymorphism) attempts to counteract the problems mentioned above through the use of more stringent PCR conditions provided by the ligation of adaptors to restriction fragments, and the use of specific primer sets (Vos et al., 1995; Bachem et al., 1996). Unlike differential display, cDNA-AFLP (and other similar techniques) allows for the systematic survey of an organism's transcriptome through the use of selective fragment amplification. cDNA-AFLP has been shown to produce both good reproducibility and sensitivity (Donson et al., 2002). However, like differential display, the technique still requires extensive band isolation and DNA

sequencing, making it less suitable to high-throughput analyses than other transcriptional profiling approaches.

Methods for Indirect Analysis of the Transcriptome

Methods for indirect analysis of the transcriptome rely on the hybridization of complementary sequences. While this characteristic has been exploited for decades, (Gillespie and Spiegelman, 1965) until fairly recently this has been on only a relatively small scale. Due to the recent explosion in the amount of sequence data and routine methods for cloning, hybridization-based approaches now provide a means to study the expression of thousands of genes simultaneously. These approaches most often employ DNA microarray technology. DNA microarrays consist of specific DNA elements ("probes") affixed to a solid support to which labeled cDNAs ("targets") derived from cellular mRNAs expressed under the test condition(s) may be hybridized (Figure 1.4). The probes may be either cDNAs amplified by PCR from cloned sequences or synthesized oligonucleotides representing unique gene sequences. Unlike previously described technologies, DNA microarrays only allow expression measurement for genes corresponding to those fabricated into the array. However, high-density microarrays now commonly in use may contain more than 400,000 elements, and the fabrication of 60 million elements on a single array is now possible (Perlegen Sciences, Mountain View, CA). This provides the promise of routine analysis of the genetic expression of entire higher eukaryote genomes in the near future.

Microarrays, although close-ended and, therefore, less comprehensive, offer some advantages over the other technologies mentioned. Foremost, microarrays allow for fast analysis, as the spotting of genes or oligos is a rapid, automated process, and subsequent hybridization and scanning require little time. Thus, microarrays can be used to quickly profile transcripts across numerous conditions, which would be prohibitively time consuming and costly using other techniques. Although not as vulnerable to sampling effects as some other techniques (e.g., SAGE and EST sequencing projects), microarrays are affected by hybridization efficiencies that can vary considerably across array elements (Meltzer, 2001). Finally, microaray technology is relatively affordable and open to the general scientific community, making its widespread use a reality.

PHYTOREMEDIATION AND GENOMICS

In the case of plant responses to explosives, virtually nothing is known with regard to the genes that respond to these toxic compounds. Thus, both SAGE and microarrays have great potential to contribute to our understanding of the processes plants use in dealing with these compounds. As the plant model organism, *Arabidopsis* is ideally suited to these types of studies. With a sequenced genome (The *Arabidopsis* Genome Initiative, 2000), extensive bioinformatic resources, ease of growth and handling, established methods for transformation, and a wealth of biochemical and molecular understanding in the scientific community, *Arabidopsis* provides a means for taking full advantage of the strength of transcript profiling technologies. Genomic studies of *Arabidopsis* responses to explosives will accelerate the discovery of genes of

potential use in the engineering of plants with enhanced abilities for explosives tolerance, removal, and degradation.

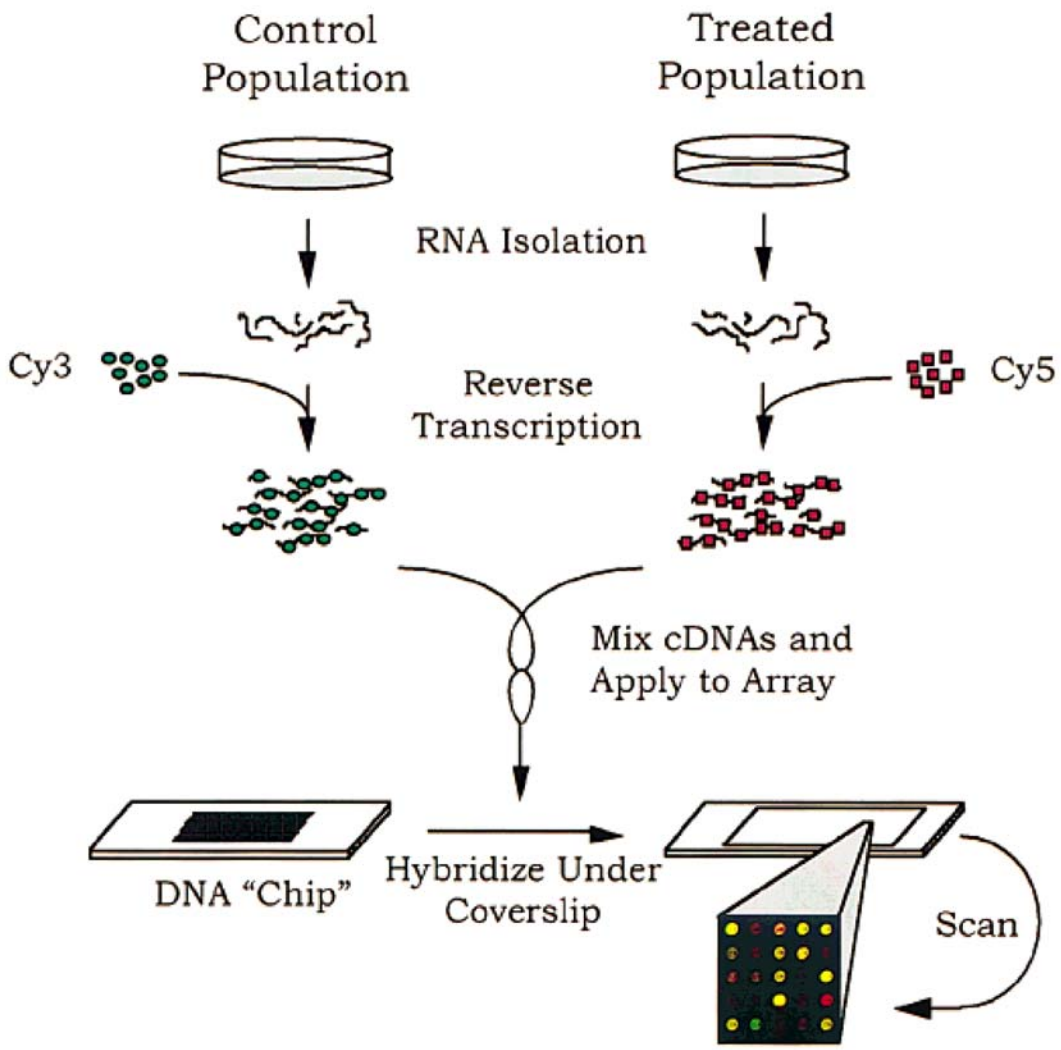


Figure 1.4. Simplified overview of the method for sample preparation and hybridization to cDNA microarrays. For illustrative purposes, samples derived from cell culture are depicted, although other sample types are amenable to this analysis (Nuwaysir et al., 1999).

REFERENCES

Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A., Boldrick, J.C., Sabet, H., Tran, T., Yu, X., Powell, J.I., Yang, L., Marti, G.E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D.B., Tibshirani, R., Sherlock, G., Chan, W.C., Greiner, T.C., Weisenburger, D.D., Armitage, J.O., Warnke, R., Staudt, L.M., et al. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature **403,** 503-511.

Alon, U., Barkaai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., and Levine, A.J. (1999). Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc. Natl. Acad. Sci. USA **966,** 6745-6750.

Anderson, T.A., Kruger, E.L., and Coats, J.R. (1994). Enhanced degradation of a mixture of three herbicides in the rhizosphere of a herbicide-tolerant plant. Chemosphere **28,** 1551-1557.

Bachem, C.W.B., van der Hoeven, R.S., de Brujin, S.M., Vreugdenhil, D., Zabeau, M., and Visser, R.G.F. (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. Plant J. **9,** 745-753.

Bae, B., Autenrieth, R.L., and Bonner, J.S. (1995). Aerobic biotransformation and mineralization of 2,4,6-trinitrotoluene. In Bioremediation of recalcitrant organics, R.E. Hinchee, Hoeppel, R.E., and Anderson, D.B., ed (Columbus, OH: Battelle Press), pp. 231-238.

Bayman, P., and Radkar, G.V. (1997). Transformation and tolerance of TNT (2,4,6-trinitrotoluene) by fungi. Int. Biodeterior. Biodegrad. **39,** 45-53.

Bedard, D.L., Wagner, R.E., Brennen, M.J., Haberl, M.L., and Brown, J.F.J. (1987). Extensive degradation of aroclors and environmentally transformed polychlorinated biphenyls by *Alcaligenes eutrophus* H850. Appl. Environ. Microbiol. **53,** 1094-1102.

Bentley, R.E., Dean, J.W., Ellis, S.J., Hollister, T.A., LeBlanc, G.A., Sauter, S., and Sleight III, B.H. (1977). Laboratory Evaluation of the Toxicity of RDX to Aquatic Organisms. Contract DAMD17-74-C-4101. EG&G Bionomics, Wareham, MA. DTIC AD-A061730.

Best, E.H.P., Miller, J.L., Fredrickson, H.L., Larson, S.L., Zappi, M.E., and Streckfuss, T.H. (1997). Screening of aquatic and wetland plant species for the phytoremediation of explosives-contaminated groundwater from the Iowa Army Ammunition Plant. Ann. N.Y. Acad. Sci. **829,** 179-194.

Best, E.H.P., Sprecher, S.L., Larson, S.L., Fredrickson, H.L., Bader, D.F. (1999). Environmental behavior and fate of explosives from groundwater from the milan army ammunition plant in aquatic and wetland plant treatments. Chemosphere **38,** 3383-3396.

Best, E.H.P., Zappi, M.E., Fredrickson, H.L., Sprecher, S.L., Larson, S.L., Ochman, M. (1997). Screeing of aquatic and wetland plant species for the phytoremediation of explosives-contaminated groundwater from the Iowa Army Ammunition. Plant. Ann. N.Y. Acad. Sci. **829,** 179-194.

Bhadra, R., Wayment, D.G., Hughes, J.B., and Shanks, J.V. (1999a). Confirmation of conjugation processes during TNT metabolism by axenic plant roots. Environ. Sci. Technol. **33,** 446-452.

Bhadra, R., Wayment, D.G., Hughes, J.B., and Shanks, J.V. (1999b). Characterization of oxidation products of TNT metabolism in aquatic phytoremediation systems of *Myriophyllum aquaticum*. Environ. Sci. Technol. **33,** 3354-3361.

Bhadra, R., Williams, R., Barman, S., Stone, M.B., Hughes, J.B., and Shanks, J.V. (2000). Fate of RDX and HMX in axenic plant roots. Chemosphere **44,** 1259-1264.

Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Ben-Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, C., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D., Sondak, V., Hayward, N., and Trent, J.M. (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature **406,** 536-540.

Bizily, S., Rugh, C.L., Summer, A.O., and Meagher, R.B. (1999). Phytoremediation of methylmercury pollution: *mer*B expression in *Arabidopsis thaliana* confers resistance to organomercurials. Proc. Natl. Acad. Sci. USA **96,** 6808-6813.

Boopathy, R., and Kulpa, C.F. (1992). Trinitrotoluene as a sole nitrogen source for a sulfate-reducing bacterium *Desulfovibrio sp.* (strain B) isolated from an anaerobic digester. Curr. Microbiol. **25,** 235-241.

Boopathy, R., Gurgas, M., Ullian, J., and Manning, J. (1998a). Metabolism of explosive compounds by sulfate-reducing bacteria. Curr. Microbiol. **37,** 127-131.

Brenner, S., Williams, S.R., Vermass, E.H., Storck, T., Moon, K., McCollum, C., Mao, J.I., Luo, S., Kirchner, J.J., Eletr, S., DuBridge, R.B., Burcham, T., and Albrecht, G. (2000b). In vitro cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. Proc. Natl. Acad. Sci. USA **97,** 1665-1670.

Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermaas, E., Williams, S.R., Moon, K., Burcham, T., Pallas, M., DuBridge, R.B., Kirchner, J., Fearon, K., Mao, J.I., and Corcoran, K. (2000a). Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat. Biotechnol. **18,** 630-634.

Bucher, J. (1990). Toxicology and Carcinogenesis Studies of Tetranitromethane (CAS No 509-14-8) in F344/N and B6C3F1 Mice (Inhalation Studies). Technical Report Series No. 386. National Toxicology Program, U.S. Department of Health and Human Services, Public Health Service, Research Triangle Park, N.C.

Burken, J.G., Shanks, J.V., and Thompson, P.L. (2000). Phytoremediation and Plant Metabolism of Explosives and Nitroaromatic Compounds. In Biodegradation of Nitroaromatic Compounds and Explosives, J.C. Spain, ed,pp. 239-275.

Carman, E., Crossman, T., and Gatliff, E. (1998). Phytoremediation of no. 2 fuel-oil contaminated soil. Journal of Soil Contam. **7,** 455-466.

Cataldo, D.A., Harvey, S.D., and Fellows, R.J. (1990). An Evaluation of the Environmental Fate and Behavior of Munitions Material (TNT, RDX) in Soil and Plant Systems: Environmental Fate and Behavior of RDX. Final Report, Army Project Order 88PP8853, PNL-7529, Pacific Northwest Laboratory, Richland, WA.

Chen, D. (1993). M.S. thesis. University of Illinois, Urbana-Champaign.

Compton, H.R., Harosi, D.M., Hirsch, S.R., and Wrobel, J.G. (1998). Pilot-scale use of trees to address voc contamination. In Bioremediation and Phytoremediation, Chlorinated and Recalcitrant Compounds, G.B. Wickramanayake and R.E. Hinchee, eds (Columbus: Battelle Press).

Dacre, J., and Rosenblatt, D.H. (1974). Mammalian toxicology and toxicity to aquatic organisms of four important types of waterborne munitions pollutants-an extensive literature evaluation. AD-A778725. TR-7403. U.S. Army Medical Bioengineering Research and Development Laboratory, Fort Detrick, Fredrick, MD.

Dilley, J.V., Tyson, C.A., and Newell, G.W. (1978). Mammalian Toxicological Evaluation of TNT Wastewaters. Volume 2: Acute and Subacute Mammalian Toxicity of TNT and LAP Wastewaters. Final Report, Contract DAMD17-76-C-6050. SRI International, Menlo Park, CA. DTIC AD-A080957.

Donnelly, P.K., Hegde, R.S., and Fletcher, J.S. (1994). Growth of PCB-degrading bacteria on compounds from photosynthetic plants. Chemosphere **28,** 981-988.

Donson, J., Fang, Y., Espiritu-Santo, G., Xing, W., Salazar, A., Miyamoto, S., Armendarez, V., and Volkmuth, W. (2002). Comprehensive gene expression analysis by transcript profiling. Plant Mol. Biol. **48,** 75-97.

Duque, E., Haidour, A., Godoy, F., and Ramos, J.L. (1993). Construction of a Pseudomonas hybrid strain that mineralizes 2,4,6-trinitrotoluene. J. Bacteriol. **175,** 2278-2283.

Ellis, H.V., Hodgson III, J.D., Hwang, S.W., Halfpap, L.M., Helton, D.V., Anderson, B.S., VanGoethem, D.L., and Lee, C.C. (1978b). Mammalian Toxicity of Munitions Compounds Phase I: Acute Oral Toxicity, Primary Skin and Eye Irritation, Dermal Sensitization, Disposition and Metabolism, and Ames Tests of Additional Compounds. Report 6, Contract DAMD17-74-C-4073. Midwest Research Institute, Kansas City, MO. DTIC AD-A060333.

Ermakov, E.V., Aizenschtadt, V.S., and Bentsenostsev, B.B. (1969). Chronic trinitrotoluene poisoning (clinical picture and pathogenesis of neurovegetative disturbances). Sov. Med. **32,** 119-122.

Esteve-nunez, A., Caballero, A., and Ramos, J.L. (2001). Biological degradation of 2,4,6-trinitrotoluene. Microbiol. and Mol. Biol. Rev. **65,** 335-352.

Fernando, T., and Aust, S.D. (1991). Biodegradation of munition waste, TNT (2,4,6-trinitrotoluene), and RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine) by *Phanerochaete chrysosporium*. ACS Symposium Series. Ind. Eng. Chem. **486,** 214-232.

Fernando, T., Bumpus, J.A., and Aust, S.D. (1990). Biodegradation of TNT (2,4,6-trinitrotoluene) by *Phanerochaete chrysosporium*. Appl. and Environ. Microbiol. **56,** 1667-1671.

Fletcher, J.S., and Hedge, R.S. (1995). Release of phenols by perrennial plant roots and their potential importance in bioremediation. Chemosphere **31,** 3009-3016.

Frear, D.S. (1976). Pesticide conjugates-Glycosides, In: *Bound and Conjugates Pesticide Residues*. In ACS Symposium, D.D. Kaufman, G.G. Still, G.D. Paulson, and S.K. Bandal, eds (Washington, DC: American Chemical Society), pp. 35-54.

French, C.E., Rosser S.J., Davies, G.J., Nicklin, S., and Bruce, N.C. (1999). Biodegradation of explosives by transgenic plants expression pentaerythritol tetranitrate reductase. Nature Biotechnology **17,** 491-494.

Furedi, E.M., Levine, B.S., Gordon, D.E., Rac, V.S., and Lish, P.M. (1984a). Determination of the Chronic Mammalian Toxicological Effects of TNT. Twenty-four Month Chronic Toxicity/Carcinogenicity Study of Trinitrotoluene (TNT) in the Fischer 344 Rat. Phase III, Vols. 1-4. Final Report, Contract DAMD17-79-C-9120. IIT Research Institute, Chicago, IL. DTIC AD-A168637.

Gillespie, D., and Spiegelman, S. (1965). A quantitative assay for DNA-RNA hybrids immobilized on a membrane. J. Mol. Biol. **3,** 829-842.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D., and Lander, E.S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science **286,** 531-537.

Gorge, E., Brandt, S., and Werner, D. (1994). Uptake and metabolism of 2,4,6 TNT in higher plants. Environ. Sci. Pollut. Res. **1,** 229-233.

Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol. **19,** 1720-1730.

Haag, E., and Raman, V. (1994). Effects of primer choice and source of Taq DNA polymerase on the banding patterns of differential display RT-PCR. Biotechniques **2,** 226-228.

Hannink, N., Rosser, S.J., French, C.E., Basran, A., Murray, J.A.H., Nicklin, S., and Bruce, N.C. (2001). Phytodetoxification of TNT by transgenic plants expressing a bacterial nitroreductase. Nature Biotechnology **19,** 1168-1172.

Harvey, S.D., Fellows, R.J., Cataldo, D.A., and Bean, R.M. (1991). Fate of the explosive hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) in soil and bioaccumulation in bush bean hydroponic plants. Environ. Toxicol. Chem. **10,** 845-855.

Haselwandter, M.M. (1983). Physiological Plant Ecology Vol. 3. Responses to the Chemical and Biological Environment. (Berlin: Springer-Verlag).

Hawari, J., Halasz, A., Beaudet, S., Paquet, L., Groom, C., Ampleman, G., and Thiboutot, S. (1999). Biotransformation of 2,4,6-trinitrotoluene with *Phanerochaete chrysosporium* in agitated cultures at pH 4.5. Appl. and Environ. Microbiol. **65,** 2977-2986.

Hedge, R.S., and Fletcher, J.S. (1996). Influence of plant growth stage and season on the release of root phenolics by mulberry as related to the development of phytoremediation technology. Chemosphere **23,** 2471-2479.

Hodgson, J., Rho, D., Guiot, S.R., Ampleman, G., Thiboutot, S., and Hawari, J. (2000). Tween 80 enhanced TNT mineralization by *Phaerochaete chrysosporium*. Can. J. Microbiol. **46,** 110-118.

Hughes, J.B., Shanks, J.V., Vanderford, M., Lauritzen, J., and Bhadra, R. (1997). Transformation of TNT by aquatic plants and plant tissue cultures. Environ. Sci. Technol. **31,** 266-271.

Jarvis, A.S., McFarland, V.A., and Honeycutt, M.E. (1998). Assessment of the effectiveness of composting for the reduction of toxicity and mutagenicity of explosive-contaminated soil. Ecotoxicol. and Environ. Safety **39,** 131-135.

Jerger, D.E., and Woodhull, P.M. (2000). Applications and costs for biological treatment of explosives-contaminated soils in the U.S. In Biodegradation of nitroaromatic compounds and explosives, J.C. Spain, J.B. Hughes, and H.-J. Knackmuss, eds (Boca Raton: CRC Press), pp. 395-424.

Kaplan, A.S., Berghout, C.F., and Peczenik, A. (1965). Human intoxication from RDX. Arch. Environ. Health **10,** 877-883.

Ketel, W.B., and Huyhes, J.R. (1972). Toxic encephalopathy with seizures secondary to ingestion of composition C-4: A clinical and electroencephalographic study. Neurology **22,** 871-876.

Khan, J., Simon, R., Bittner, M., Chen, Y., Leighton, S.B., Pohida, T., Smith, P.D., Jiang, Y., Gooden, G.C., Trent, J.M., and Meltzer, P.S. (1998). Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res. **58**, 5009-5013.

Klein, W., and Scheunert, I. (1982). Bound pesticide residues in soil, plants, and food with particular emphasis on the application for nuclear techniques. IEIA-SM-263/38 Interatomic Energy Agency.

Larson, S.L. (1997). Fate of explosive contaminants in plants. Ann. N.Y. Acad. Sci. **829,** 195-201.

Larson, S.L., Weiss, C.A., Escalon, B.L., and Parker, D. (1999a). Increased extraction efficiency of acetonitrile/water mixtures for explosives determination in plant tissues. Chemosphere **38,** 2153-2162.

Larson, S.L., Jones, R.P., Escalon, B.L., and Parker, D. (1999b). Classification of explosives transformation products in plant tissue. Environ. Toxicol. Chem. **18,** 1270-1276.

Lashkari, D.A., DeRisi, J.L., McCusker, J.H., Namath, A.F., Gentile, C., Hwang, S.Y., Brown, P.O., and Davis, R.W. (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc. Natl. Acad. Sci. USA **94,** 13057-13062.

Lauritzen, J. (1998). M.S. thesis. Rice University, Houston.

Levine, B.S., Furedi, E.M., Gordon, D.E., Barkley, J.J., and Lish, P.M. (1990a). Toxic interactions of the munitions compounds TNT and RDX in F344 rats. Fund. and Appl. Toxicol. **15,** 373-380.

Lewis, R.J. (1992). Sax's Dangerous Properties of Industrial Materials. (New York: Von Nostrand Reinhold).

Liang, P., and Pardee, A.B. (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science **257,** 967-971.

Lish, P.M., Levine, B.S., Furedi-Machacek, E.M., Sagartz, E.M., and Rac, V.S. (1984). Determination of the Chronic Mammalian Toxicological Effects of RDX: Twenty-four Month Chronic Toxicity/carcinogenicity Study of Hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) in the B6C3FI Hybrid Mouse. Phase VI. Vol. I. Contract No. DAMD 17-79-C-9161. U.S. Army Medical Research and Development Command. DTIC AD-A160774.

Lynch, J.M. (1982). Interactions between bacteria and plants in the root environment. In Bacteria and Plants. Soc. Appli. Bacteriol. Symp. Ser., M.E. Rhodes-Roberts and F.A. Skinner, eds (London: Academic), pp. 1-23.

Matz, M.V., and Lukyanov, S.A. (1998). Different strategies of differential display: areas of application. Nucl. Acids Res. **26,** 5537-5543.

McCormick, N.G., Cornell, J.H., and Kaplan, A.M. (1981). Biodegradation of hexahydro-1,3,5-trinitro-1,3,5-triazine. Appl. and Environ. Microbiol. **42,** 817-823.

Meagher, R.B. (2000). Phytoremediation of toxic elemental and organic pollutants. Curr. Opin. in Plant Biol. **3,** 153-162.

Medina, V.F., Rivera, R., Larson, S., and McCutcheon, S.C. (1996). Phytoremediation: modeling of TNT and its breakdown products. Soil Groundwater Cleanup **Feb/Mar,** 19-24.

Meltzer, P.S. (2001). Large-scale genome analysis. In Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. 2nd ed., A.D. Baxevanis, and Ouellette, B.F.F., ed (New York City, NY: John Wiley and Sons, Inc.), pp. 393-412.

Michels, J., and Gottschalk, G. (1995). Pathway of 2,4,6-trinitrotoluene (TNT) degradation by *Phanerochaete chrysosporium*. In Biodegradation of nitroaromatic compounds, J.C. Spain, ed (New York, NY: Plenum Press), pp. 135-149.

Montpas, S., Samson, J., Langlois, E., Lei, J., Piche, Y., and Chevenert, R. (1997). Degradation of 2,4,6-trinitrotoluene by *Serratia marcensens*. Biotechnol. Letters **19,** 291-294.

Nuwaysir, E.F., Bittner, M., Trent, J., Barrett, J.C., and Afshari, C.A. (1999). Microarrays and toxicology: the advent of toxicogenomics. Mol. Carcin. **24,** 153-159.

Oh, K., and Kim, Y. (1998). Degradation of explosive 2,4,6-trinitrotoluene by s-triazine degrading bacterium isolated from contaminated soil. Bull. Environ. Contam. Toxicol. **61,** 702-708.

Palazzo, A.J., and Leggett, D.C. (1986). Effect and disposition of TNT in terrestrial plants. J. Environmental Quality **15,** 49-52.

Palmer, W.G., Small, M.J., Dacre, J.C., and Eaton, J.C. (1996). Toxicology and Environmental Hazards. P.L. Marinkas. (New York: Nova Science Publisher, Inc.).

Preuss, A., Fimpel, J., and Dickert, G. (1993). Anaerobic transformation of 2,4,6-trinitrotoluene (TNT). Arch. Microbiol **159,** 345-353.

Rieble, S., Joshi, D.K., and Gold, M. (1994). Aromatic nitroreductase from the basidiomycete *Phanerochaete chrysosporium*. Biochem. Biophys. Res. Commun. **205,** 28-304.

Ross, D.T., Scherf, U., Eisen, M.B., Perou, C.M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S.S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J.C., Lashkari, D., Shalon, D., Myers, T.G., Weinstein, J.N., Botstein, D., and Brown, P.O. (2000). Systematic variation in gene expression patterns in human cancer cell lines. Nat. Genet. **24,** 227-235.

Rugh, C.L., Wilde, D., Stack, N.M., Thompson, D.M., Summer, A.O., and Meagher, R.B. (1996). Mercuric ion reduction and resistance in transgenic *Arabidopsis thaliana* plants expressing a modified bacterial *mer*A gene. Proc. Natl. Acad. Sci. USA **93,** 3182-3187.

Ryon, M.G., Pal, B.C., Talmage, S.S., and Ross, R.H. (1984). Database Assessment of the Health and Environmental Effects of Munition Production Waste Products. Final Report, Army Project Order 83PP3802. ORNL-6018, Oak Ridge National Laboratory, Oak Ridge, TN. DTIC AD-A145417.

Sandermann, J.H. (1994). Higher plant metabolism of xenobiotics: the 'green liver' concept. Pharmacogenetics **4,** 225-241.

Scheidemann, P.A., Klunk, A., Sens, C., and Werner, D.J. (1998). Species dependent uptake and tolerance of nitroaromatic compounds by higher plants. Plant Physiol. **152,** 242-247.

Scherf, U., Ross, D.T., Waltham, M., Smith, L.H., Lee, J.K., Tanabe, L., Kohn, K.W., Reinhold, W.C., Myers, T.G., Andrews, D.T., Scudiero, D.A., Eisen, M.B., Sausville, E.A., Pommier, Y., Botstein, D., Brown, P.O., and Weinstein, J.N. (2000). A gene expression database for the molecular pharmacology of cancer. Nat. Genet. **24,** 236-244.

Schnell, S., and Schinck, B. (1991). Anaerobic aniline degradation via reductive deamination of 4-aminobenzoyl-CoA in *Desulfobacterium anilini*. Arch. Microbiol. **155,** 183-190.

Schnoor, J.L. (1997). Phytoremediation. Technology Evaluation Report TE-97-01. National Environmental Technology Applications Center, Pittsburgh, PA.

Schroder, P., and Collins, C. (2002). Conjugating enzymes involved in xenobiotic metabolism or organic xenobiotics in plants. Int. Journ. of Phytoremed. **4,** 247-265.

Schwab, A.P., and Banks, M.K. (1994). Biologically mediated dissipation of polyaromatic hydrocarbons in the root zone. In Bioremediation through Rhizosphere Technology, T.A. Anderson and J.R. Coats, eds (Washington, D.C.: American Chemical Society.

Shimp, J.F., Tracy, J.C., Davis, L.C., Lee, E., and Huang, W. (1993). Beneficial effects of plants in the remediation of soil and groundwater contaminated with organic materials. Environ. Sci. Technol. **23,** 41-77.

Smock, L.A., Stoneburger, D.L., and Clark, J.R. (1976). The toxic effects of trinitrotoluene (TNT) and its primary degradation products on two species of algae and the fathead minnow. Water Res. **10,** 537-543.

Sompayrac, L., Jane, S., Burn, T.C., Tene, D.G., and Danna, K.J. (1995). Overcoming limitations of the mRNA differential display technique. Nucl. Acids Res. **23,** 4738-4739.

Stahl, J.D., and Aust, S.D. (1995). Biodegradation of 2,4,6-trinitrotoluene by white rot fungus *Phanerochaete chrysosporium*. In Biodegradation of nitroaromatic compounds and explosives, J.C. Spain, ed (Boca Raton: Plenum Press).

Sun, Y., Hegamyer, G., and Colburn, N.H. (1994). Molecular cloning of five messenger RNAs differentially expressed in preneoplastic or neoplastic JB6 mouse epidermal cells: one is homologous to human tissue inhibitor of metalloproteinases-3. Cancer Res. **5,** 1139-1144.

Tan, E.L., Hoe, C.H., Griest, W.H., and Tyndall, R.L. (1992). Mutagenicity of trinitrotoluene and its metabolites formed during composting. Journ. of Toxicol. and Environ. Health **36,** 165-175.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408,** 796-815.

Thompson, P.L., Ramer, L., and Schnoor, J.L. (1998). Uptake and transformation of TNT by hybrid poplar trees. Environ. Sci. Technol. **32,** 975-980.

Thompson, P.L., Ramer, L., and Schnoor, J.L. (1999). Hexahydro-1,3,5-trinitro-1,3,5-triazine translocation in poplar trees. Environ. Toxicol. Chem. **18,** 279-284.

Traxler, R.W., Wood, E., and Delaney, J.M. (1974). Bacterial degradation of alpha-TNT. Dev. Ind. Microbiol. **16,** 71-76.

Vanderford, M. (1996). Aquatic phytotransformation of trinitrotoluene in contaminated media (Houston: Rice University).

Vanderford, M., Shanks, J.V., and Hughes, J.B. (1997). Phytotransformation of trinitrotoluene (TNT) and distribution of metabolic products in *Myriophyllum aquaticum*. Biotech. Lettters **19,** 277-280.

Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. (1995). Serial analysis of gene expression. Science **270,** 484-487.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. Nucl. Acids Res. **23,** 4407-4414.

Walton, B.T., Hoylman, A.M., Perez, M.M., Anderson, T.A., Johnson, T.R., Guthde, T.E., and Christman, R.F. (1994). Rhizosphere microbial community as a plant defense against toxic substances in soils. In Bioremediation Through Rhizosphere Technology, T.A. Anderson, Coats, J.R., eds (Washington, DC: Am. Chem. Soc.), pp. 82-92.

Watanabe, M.E. (2001). Can bioremediation bounce back? Nat. Biotechnol. **19,** 1111-1115.

Wayment, D.G., Bhadra, R., Lauritzen, J., Hughes, J., and Shanks, J.V. (1999). A transient study of formation of conjugates during TNT metabolism by plant tissues. Int. Journ. of Phytoremed. **1,** 227-239.

Wodicka, L., Dong, H., Mittman, M., Ho, M.H., and Lockhart, D.J. (1997). Genome-wide expression monitoring in *Saccharomyces cerevisiae*. Nat. Biotechnol. **15,** 1359-1367.

Wolfe, N.L., Ou, T.Y., Carriera, L., and Gunnison, D. (1994). Alternative methods for biological destruction of TNT: a preliminary feasibility assessment of enzymatic degradation. Technical Report IRRP-94-3. U.S. Army Corps of Engineers Waterways Experiment Station, Vicksburg, MS.

Won, W.D., Disalvo, L.H., and Ng, J. (1976). Toxicity and mutagenecity of 2,4,6-trinitrotoluene and its microbial metabolites. Appl. and Environ. Microbiol. **31,** 576-580.

Yinon, J. (1990). Toxicity and metabolism of explosives. (Ann Arbor, MI: CRC Press).

Zhang, J.S., Duncan, E.L., Chang, A.C.M., and Reddel, R.R. (1998). Differential display of mRNA. Mol. Biotechnol. **10,** 155-165.

CHAPTER 2

SAGE ANALYSIS OF TRANSCRIPTOME RESPONSES IN *ARABIDOPSIS THALIANA* ROOTS EXPOSED TO TNT[1]

[1]Ekman, D.R., Lorenz, W.W., Przybyla, A.E., Wolfe, N.L., and Dean, J.F.D. Submitted to *Plant Physiology*.

ABSTRACT

Serial Analysis of Gene Expression (SAGE) was used to profile transcript levels in *Arabidopsis thaliana* roots and assess their responses to 2,4,6-trinitrotoluene (TNT) exposure. SAGE libraries representing control and TNT-exposed seedling root transcripts were constructed, and each was sequenced to a depth of roughly 30,000 tags. More than 19,000 unique tags were identified overall. The second most highly induced tag (27-fold increase) represented a glutathione S-transferase. Several cytochrome P450 enzymes, as well as an ABC transporter, were highly induced by TNT exposure. These analyses also revealed an oxidative stress response upon TNT exposure. Although some increases were anticipated in light of current models for xenobiotic metabolism in plants, evidence for unsuspected conjugation pathways was also noted. Identifying transcriptome-level responses to TNT exposure will better define the metabolic pathways plants use to detoxify this xenobiotic compound, which should help improve phytoremediation strategies directed at TNT and other nitroaromatic compounds.

(Note: the datasets from these analyses are available at the ncbi Gene Expression Omnibus (GEO) at: http://www.ncbi.nlm.nih.gov/geo/)

INTRODUCTION

Soil and groundwater at sites throughout the U.S. and Europe were contaminated in the past century by manufacturing, processing, and storage of explosives, such as 2,4,6-trinitrotoluene (TNT)[1]. Unlike many other nitroaromatic compounds, including pesticides and various feedstock chemicals, the energetic nitroaromatics (Figure 2.1) are highly resistant to degradation and may persist in the environment for decades. Certain plant species have the ability to accumulate TNT from their surroundings and, thus, offer a potential means for removing these compounds from the environment[2]. Unfortunately, few of these species are capable of tolerating the contamination levels common to sites most in need of remediation. Transgenic plants have been developed with enhanced abilities to tolerate and remove TNT from soil under laboratory conditions[3,4]. However, lack of information on the biochemical mechanisms involved in TNT uptake and metabolism limits our ability to modify and adapt plants specifically for this task.

To obtain a more complete picture of the metabolic processes plants employ to cope with nitroaromatic agents, Serial Analysis of Gene Expression (SAGE)[5,6] was used to identify transcriptome-level responses in *A. thaliana* seedling roots exposed to TNT. Although this technique has been widely used to study gene expression changes in human cancers[7], only two published reports describe the use of SAGE in plants: a study of gene expression in rice seedlings[8], and an examination of wood formation in loblolly pine[9]. The identification of plant genes and pathways responding to TNT exposure at the transcriptional level should facilitate a more reasoned approach to the development of plants better suited for use in phytoremediation of TNT-contaminated sites.

RESULTS

SAGE Libraries.

SAGE libraries representing transcripts expressed in *Arabidopsis* root tissues grown in the absence or presence (15 mg/L) of TNT were sequenced to characterize about 30,000 tags from each. The datasets from these analyses are available online in spreadsheet format (http://www.arches.uga.edu/~jeffdean/SAGE/SAGEData.html). Of the 32,203 tags characterized from the library for TNT-treated tissue, 12,005 represented unique transcripts, but 7,900 of these were singletons. Similarly, 12,719 unique tags were encountered amongst the 31,973 characterized in the control library, with 8,322 of these representing singletons. A double-reciprocal plot of the total unique tags identified versus the total number of tags sequenced after each sequencing run was used to estimate the rate of transcript discovery (Fig. 2.2). From this, transcriptome sizes of approximately 21,000 and 15,000 were estimated for control and TNT-treated root tissues, respectively. Only 25% of all unique tags (5,084 out of 19,640) were detected in both libraries.

BLAST Analyses.

Of the 1,045 most abundant tags in both SAGE libraries (which included all tags seen 10 or more times amongst the 64,176 tags characterized) about 70% (739), could be matched to a single model gene in the AGI (**A**rabidopsis **G**enome **I**nitiative) database.

Most of the remaining tags (233) matched sequences found one or more times among all *Arabidopsis* sequences deposited in GenBank. Many such tags were found in expressed sequence tags (ESTs), but were not positioned adjacent to the CATG sequence closest to the 3' end of the transcript. This suggests they might have been generated from alternatively spliced transcripts. A good example of this was the fourth most abundant tag in the control tissue library (AGGTCTTGGT, counted 134 times). This sequence appears only one time in the *Arabidopsis* genome, and falls immediately 3' of the penultimate CATG site on the annotated transcript of gene At3g09260 (β-glucosidase). The full-length transcript for this gene annotated in the AGI database was responsible for second most abundant tag in the control library (ATTTGCCAGA, counted 286 times) and the fourth most abundant tag in the TNT-treatment library (counted 259 times).

Of the remaining unidentified tags in the top 1,045, one that did not correspond to anything listed in the AGI database (AGTAACGATA) matched a sequence found on the *Arabidopsis* mitochondrial genome, and numerous ESTs incorporating this sequence have been deposited in GenBank. About 5% of the tags (57) matched more than one model gene, while a few tags (14) contained at their 3' end a contiguous stretch of "A" residues suggesting that they incorporated part of a poly(A) tail. Two tags, one of which (TCCCCGTACA) was the 37th most abundant tag overall, could not be matched to any *Arabidopsis* sequences in GenBank.

Differentially Expressed Transcripts.

TNT treatment induced an apparent increase of at least five-fold in 242 tags, while 287 tags decreased in abundance at least five-fold in response to this treatment. For tags that were relatively abundant in one library but not observed in the other, a minimum change in expression level was estimated by assuming that a single copy of the tag was found in the library from which it was absent. Many of the tags displaying the greatest induction in response to TNT exposure represented gene products known to be involved in plant responses to oxidative stress, such as monodehydroascorbate reductase and phospholipid-hydroperoxide glutathione peroxidase (Table 2.1). The identity of the tag most highly induced by TNT remains uncertain. Its sequence appears twice on Chromosome 5, and in one of these instances the tag lies in the middle of a computationally predicted transcript that has never been isolated. However, a third occurrence of the sequence resides on Chromosome 2 at the penultimate CAGT site on the transcript for LKP2 (At2g18910), a signaling protein involved in the *Arabidopsis* circadian clock[10]. The evidence suggests the most likely source of this tag to be a splice variant of the LKP2 transcript. In contrast, the second most highly induced tag, which increased nearly 28-fold in TNT-treated roots, definitely represented a glutathione S-transferase (GST). Several of the induced tags represented cytochrome P450s, a large family of enzymes known to be involved in the detoxification of xenobiotic compounds by plants and other organisms (Table 2.2). As further indication of the stress imposed upon the seedlings by TNT exposure, tags for various transporter proteins, transcription factors, signal cascade proteins, and heat shock proteins were also among those most

highly induced by TNT exposure. Other strongly induced tags representing gene products known to be involved in detoxification reactions or in the protection of plant cells against oxidative stress are noted in Table 2.3. Prominent among the enzymes likely to protect against oxidative stress are GSH-dependent dehydroascorbate reductase, 1,4-benzoquinone reductase, peptide methionine sulfoxide reductase, and glutathione reductase.

Tags that showed the greatest decrease upon exposure to TNT are listed in Table 2.4. The largest effect was seen for a tag representing AIR1A, a transcript that has been associated with lateral root formation in response to auxin[11]. AIR1A and AIR1B, a nearly identical transcript that appears to arise via alternative splicing, encode proteins that are similar to a group of cell wall-plasma membrane connector proteins widely distributed in plants[12]. However, neither of the AIR1 proteins harbors the N-terminal proline- or glycine-rich extracellular domain thought to mediate cell wall interaction in other members of the connector protein family. Interestingly, the AIR1B transcript is represented by a different SAGE tag, which happens to be the tag showing the fourth most precipitous decline in response to TNT. Another member of the cell wall-cell membrane connector protein family is pEARLI-1, an *Arabidopsis* homolog that was also severely repressed in response to the TNT treatment (Table 2.4).

TNT exposure also depressed levels of tags representing three different latex protein homologs, another group of widely distributed proteins whose functions are as yet unknown[13]. Other depressed tags included those for various transcription factors, membrane channels, cytoskeletal elements, and ribosomal proteins.

Quantitative PCR.

SAGE has previously been shown to provide an accurate reflection of gene expression levels for medium- and high-abundance transcripts[14], but the induction or repression of selected *Arabidopsis* transcripts in response to TNT exposure was independently verified in this study using real-time quantitative PCR. Genes that SAGE analysis suggested were induced (At3g28740; cytochrome P450), repressed (At2g36830; aquaporin) or remained unaffected (At2g39460; 60S ribosomal protein L23A) by TNT exposure were tested. As shown in Table 2.5, the quantitative PCR data showed general agreement with the SAGE data.

DISCUSSION

A major goal of this study was to identify plant-specific enzymes and metabolic pathways that might previously have been overlooked for their importance in conferring tolerance to TNT [seedlings were grown in sterile liquid culture where microbial influences could be avoided]. The tolerance of *Arabidopsis* to TNT under these conditions (15 mg/L) was comparable to those seen in other plants grown under sterile culture conditions[4]. For example, TNT was shown to induce chlorosis and cell death in cultures of *Anabena* at a concentration of 10 mg/L[15], while cell suspension cultures of *Datura innoxia* were capable of tolerating TNT at up to 29 mg/L without an affect on growth[16]. Although the liquid culture system used in this study exposed all portions of the *Arabidopsis* seedlings to TNT, only the root tissues were sampled for SAGE analyses

because under conditions of normal terrestrial growth TNT and its derivatives tend to remain associated with root tissues[17]. Thus, efforts to improve plant tolerance to TNT would seem most likely to benefit from detailed study of the metabolic responses in these tissues.

SAGE Analysis of the *Arabidopsis* Transcriptome.

The *Arabidopsis* genome has been estimated to harbor approximately 25,500 genes[18], but given the possibilities for alternative splicing and polyadenylation during transcript maturation in eukaryotes, the transcriptional space of *Arabidopsis* should be substantially larger. This inference was supported by SAGE results identifying more than 19,000 unique tags in a sampling of only 60,000 tags from root tissues alone. Just as was seen in a similarly sized SAGE study of rat fibroblast cells[19], the discovery of new tags in the *Arabidopsis* SAGE libraries had not yet begun to level off at this depth of sampling, and estimates suggested an eventual transcriptome size in the range of 35,000-40,000 for the root tissues. However, as noted by Velculescu et al.[20], to obtain a reasonably complete sampling of a somewhat larger transcriptome (ca. 56,000 transcripts) using SAGE, one would need to sequence on the order of ca. 650,000 tags. Thus, our data likely represent a low estimate of transcriptional capacity in *Arabidopsis*.

The size of this SAGE study was similar to a hypothetical situation (15,720 genes, 62,178 sampled tags) modeled by Stollberg et al.[21] to estimate the maximum likelihood of generating accurate transcript number and transcript copy frequency given different levels of randomness and non-uniqueness in the transcriptome. Modeling suggested that

in a study of this size 1.5 - 6% of SAGE tags should represent two or more genes depending on the non-randomness of sequences represented in the transcriptome. This value matched closely with the 5% figure noted for the 1,045 most abundant tags in this study. However, the modeling study did not take into account the further complexity that might be introduced by variations in transcript maturation processes, such as the alternative splicing. Adding further uncertainty to estimates of the *Arabidopsis* root transcriptome size was the observation that among tags categorized as having ambiguous origin, several matched the antisense sequences for SAGE tags predicted for certain transcripts having numerous ESTs on deposit in GenBank. SAGE previously demonstrated that antisense transcripts for certain genes were abundant in *Caenorhabditis elegans*[22] and *Plasmodium falciparum*[23], and although not verified in this study, SAGE tags from antisense transcripts have been verified by RT-PCR in stressed *Arabidopsis* plants (G. May, personal communication). Perhaps this phenomenon is a reflection of the up-regulation of endogenous reverse transcriptases, such as that encoded by At2g166680, which was induced seven-fold by TNT exposure (Table 2.1).

Multiphase Mechanisms of TNT Detoxification in *Arabidopsis.*

Studies of the mechanisms plants use to metabolize herbicides and other xenobiotics have previously pointed to a multiphase process for detoxification[24,25,26]. Following entrance of the compound into the cell, transformation reactions introduce chemical substituents amenable to conjugation, for example, hydroxyl groups added via enzyme-catalyzed oxidation reactions involving cytochrome P450 enzymes and other

mixed function oxidases. The modified compound is subsequently conjugated to glutathione (GSH) or any of several six-carbon sugars through the action of glutathione S-transferases (GSTs) or UDP:glucosyltransferases, respectively. In the third phase of the process, conjugates are sequestered in the vacuole, where they may be inactivated by further modifications or enter into degradation pathways, or secreted into the apoplasm where they may be covalently coupled into the cell wall.

SAGE analysis indicates that this multiphase process also functions in the metabolism of TNT by *Arabidopsis*. Although this has been suggested previously for two other plants, *Catharanthus roseus* and *Myriophyllum aquaticum,* based on mass-balance studies of TNT disappearance and metabolite production in axenic cultures[27,28], the enzymes most likely involved in the process have not previously been determined. The apparent induction of several cytochrome P450 genes by TNT exposure supports oxidation as the initial transformation step in TNT metabolism. Oxidative transformation of TNT in plants has been suggested from analyses of metabolic products[28], but reductive metabolites, such as 2-amino-4,6-dinitrotoluene or 4-amino-2,6-dinitrotoluene, have been recovered most often in mass-balance studies of plant cultures exposed to ^{14}C-TNT. Typically, however, the fraction of TNT derivatives recoverable from such studies as small molecules has been only a small percentage of the added TNT, and most of the isotopic label recovered from such studies remained in unidentified conjugates or bound to high molecular weight materials. This suggests that after oxidation of TNT, subsequent conjugation reactions occur too quickly to allow for routine measure of the oxidized intermediates.

While hydroxylation is an effective means of modifying xenobiotic compounds for subsequent conjugation, it rarely serves to decrease toxicity of the parent compound and can sometimes make the compound more toxic. In such cases, hydroxylating enzymes must be coordinately expressed with conjugating enzymes that will decrease the toxic nature of the derivatives as quickly as possible. The SAGE data suggest that glutathione S-transferases were the enzymes shouldering primary responsibility for conjugation reactions involving TNT metabolites. Although previous studies have not provided evidence for the conjugation of glutathione to TNT, its conjugation to herbicides and other phytotoxins is well known. In fact, detoxification of herbicides by glutathione conjugation is used advantageously in agriculture where compounds that elevate glutathione and glutathione S-transferase levels, known as "safeners," are applied to crops prior to herbicide application[29]. Many safeners are aromatic compounds with nitrogen-containing functional groups, not unlike TNT, and their metabolism in plants has been relatively well studied. Enhanced glucosylation reactions in response to safener treatment have also been noted[30], and may explain the TNT induction of a SAGE tag representing sucrose-UDP glucosyltransferase. From these observations it would seem that many of the mechanisms known to function in safener metabolism may be directly applicable in understanding the response of *Arabidopsis* to TNT exposure.

On the other hand, the seven-fold induction of anthranilate N-benzoyltransferase (Table 2.3), an enzyme that catalyzes the first committed step of a phytoalexin biosynthetic pathway and that has been shown to be adept at utilizing hydroxycinnamoyl-CoA esters to modify anthranilate[31], may suggest a conjugation pathway for TNT derivatives that has not been previously suspected. Bhadra et al.[28] found that oxidation of

the methyl group converted nearly 20% of the TNT added to plant cultures to various benzoyl derivatives. As benzoate:CoA ligase is a member of an enzyme family that includes 4-coumarate:CoA ligase[32], the near three-fold induction of 4-coumarate:CoA ligase seen in the TNT-treated roots may reflect a mechanism for activating oxidized TNT derivatives prior to anthranilate conjugation.

Conjugation of glutathione with cellular toxins is generally thought to confer increased solubility, decreased toxicity, and increased transport competency on the toxin derivatives[33]. The transport of GSH conjugates across membranes is an ATP-dependent process handled by ABC transporter proteins[34,35,36]. As noted in Table 2.3, SAGE detected five-fold increase in an ABC transporter (At3g53480) upon TNT exposure. This ABC transporter is of a class related to the *S. cerevisiae* PDR5 gene and, as such, has been suggested to have potential function as a toxin-conjugate efflux pump[37]. Treatment of liquid-grown *Arabidopsis* seedlings with the nitroaromatic cytotoxin, 1-chloro-2,4-dinitrobenzene (CDNB), elicited similar increases in transcript levels for three other ABC transporters, including the AtMRP1 (At1g30400) and AtMRP4 (At2g47800), that are thought to handle transport of glutathione conjugates and anthocyanins into the vacuole[38].

In addition to undergoing oxidation reactions, the TNT taken up by plants is also subject to reduction reactions that target the aromatic nitro-groups, as evidenced by the amino, dinitro-derivatives of TNT identified in plant tissues by Bhadra et al.[28]. SAGE suggests that such reduction reactions may be catalyzed by NADPH-dependent flavoenzymes, such as 12-oxophytodienoate reductase (At1g76680), which was induced 10-fold by TNT treatment. Similar in sequence to the yeast Old Yellow Enzyme (OYE), 12-oxophytodienoate reductase is related to the nitrate ester reductases[39], and another

member of this family, pentaerythritol trinitrate (PETN) reductase, has been shown to increase TNT tolerance when expressed in transgenic tobacco plants[3]. Thus, 12-oxophytodienoate reductase may prove a useful target for over-expression experiments aimed at increasing plant tolerance to TNT and other nitroaromatic compounds.

Although there is no published evidence for mineralization of TNT by plants, the strong induction by TNT of carbamoyl phosphate synthetase (23-fold) suggests that nitrogen metabolism changed in the treated plants, possibly reflecting an unappreciated mechanism for removing nitrogen from the aromatic ring. Isotopic labeling studies have been so far limited to carbon-14, but tracking of labeled nitrogen might help to further clarify the metabolic pathways in plants for TNT.

Gene Repression in Response to TNT.

Inference of metabolic function for genes whose expression is induced by stress can be relatively straightforward given sufficient information on the activity of the encoded proteins. However, understanding the down-regulation of other genes under the same conditions of stress is much more difficult since their repression may be the indirect outcome of regulatory shifts necessary to induce protective genes. Among the transcripts most repressed by TNT exposure were several whose products have previously been associated with the normal structure and function of plant cell walls, e.g., the cell wall-plasma membrane linker proteins, AIR1a, AIR1b and pEARLI-1, an extensin, and an arabinogalactan protein (Table 2.4). As a group, latex protein homologs were also highly repressed. Although these proteins have no known function, they are generally thought

to accumulate in vacuoles. Thus, it is possible that their repression reflects changes in vacuole metabolism brought about by sequestration of TNT-conjugates in this organelle. Changes in vacuolar metabolism might also explain the repression of aquaporin and membrane channel proteins noted in Table 2.4 if these particular family members are associated with the tonoplast.

This study was undertaken to improve our understanding of the mechanisms involved in plant tolerance and metabolism of xenobiotic compounds, particularly TNT. Results from these experiments suggest the involvement of previously unappreciated enzymes and, at the same time, strengthen some existing theories of how plants cope with toxic compounds. Identification of the genes involved in the metabolism of TNT should provide for focused engineering attempts to create plants better suited to remediation.

EXPERIMENTAL PROTOCOL

Plant Material, Growth Conditions, and Root Tissue Isolation.

A. thaliana ecotype Columbia seeds (WT-2, Lehle Seeds, Round Rock, TX) were surface-sterilized and placed in sterile Murashige and Skoog liquid medium prepared according to the manufacturer's protocol (Invitrogen Life Technologies, Carlsbad, CA). Plants were grown for 14 days at 25^0C under a 16-hour photoperiod with constant shaking at 85 rpm in a growth chamber. TNT was obtained from the U.S. Army Center for Environmental Health Research (Fort Detrick, MD). Toxicity was assessed by adding TNT from a stock solution in DMSO to yield final concentrations of 0, 5, 10, 15, 20, 25,

30, and 40 mg/L in MS medium, and cultures were run in triplicate. The cultures were returned to the growth chamber for five days, during which time the seedlings in each flask were examined for signs of stress (leaf chlorosis and necrosis). A final concentration of 15 mg/L TNT was judged to produce notable stress in the plants without causing death. Root tissues for SAGE library construction were isolated from seedlings grown 14 days in liquid MS medium and dosed with TNT to a final concentration of 15 mg/L. Control tissues were isolated from seedlings grown under the same conditions and dosed with an equivalent volume of DMSO. The seedlings were grown in the presence of TNT or DMSO for 24 hours, after which they were submerged briefly in dH_2O to remove excess medium, and excised roots were immediately frozen in liquid nitrogen. To obtain sufficient biomass, five flasks of seedlings (70-80 seeds in 200 ml medium in a 500 ml flask) received each treatment, and treatments were replicated on three separate occasions. Root tissues were pooled by treatment and stored at -80^0C prior to RNA extraction.

RNA Isolation and cDNA Synthesis.

Total RNA from root tissues was extracted using the LiCl precipitation technique of Chang et al.[40]. Poly(A) RNA was isolated from total RNA using Dynabeads oligo-$dT_{(25)}$ magnetic beads (Dynal Biotech, Lake Success, NY) at a ratio of 0.25 mg of total RNA per 250 μL of Dynabeads and following the manufacture's instructions. Double-stranded cDNA was synthesized from 5 mg of poly(A) RNA using the Superscript Choice cDNA synthesis kit (Invitrogen Life Technologies) and following the

manufacturer's protocol, except for the substitution of a 5'-biotin $dT_{(18)}$ primer in the first-strand reaction.

SAGE Library Construction.

SAGE libraries were constructed according to the SAGE Detailed Protocol, Version 1.0c[6], a brief description of which follows. Biotinylated cDNAs from each tissue sample were bound to streptavidin-coated magnetic beads and digested with NlaIII, a restriction enzyme recognizing the four-base sequence, CATG (anchoring enzyme). DNA released by this digestion was washed away and the beads, with the adherent 3' ends of each cDNA, were split into two pools. Linkers containing a binding site for BsmF1 (a Type-II restriction endonuclease – the tagging enzyme), but different sites for PCR primers, were ligated to the NlaIII cleavage site at the 5' ends of the bead-bound cDNA fragments in each pool. Both pools of cDNAs were digested with BsmF1 to release SAGE tags from the beads, after which the pools were combined, and 102 bp linker-flanked ditags were formed by blunt-end ligation. Following amplification of the ditags by PCR, the linkers were removed by NlaIII digestion, and ditags were ligated to form concatemers. The concatemers were subsequently size-fractionated, ligated into the pZero vector (Invitrogen Life Technologies), cloned, and sequenced.

SAGE Data Analysis.

Sequence files were compiled and analyzed using the SAGE Software, ver. 3.03, provided by Dr. Kenneth Kinzler (Johns Hopkins University, Baltimore, MD). Tags containing linker sequences and repeated ditags were excluded prior to analysis. Because the library representing TNT-treated roots was sequenced to a slightly greater extent (32,203 tags for TNT treatment vs. 31,973 tags for the control), values for the control library tags were normalized prior to making comparisons of relative gene expression. Ratios were used to compare the relative expression of tags between the two libraries (e.g., TNT/control), and in instances where a particular tag was absent from a library, a value of 1 was substituted to avoid division by zero. Using the SAGE software, Monte Carlo simulations were performed to estimate the statistical significance of any differential expression. The null hypothesis for these analyses was that the abundance, type, and distribution of transcripts were the same in both libraries.

Gene Identification.

To identify the genes from which tags were derived, each 10-base tag plus the 4-base NlaIII recognition sequence was first compared against the Arabidopsis Gene Initiative (AGI) database of model genes using the Patmatch analysis tool available on The Arabidopsis Information Resource (TAIR) server (http://www.arabidopsis.org). If the tag was found to match exactly the NlaIII site closest to the 3' end of a model gene, this identity was accepted for the tag. Tags that could not be found in the model gene

database were compared against all *Arabidopsis* sequences in GenBank using the same Patmatch tool. Exact matches were annotated accordingly.

Quantitative PCR.

Total RNA from *Arabidopsis* plants grown under conditions identical to those used to generate RNA for the SAGE studies was used to independently verify the expression of selected genes by quantitative RT-PCR. Messenger RNA isolated using Dynal Oligo dT_{25} magnetic beads served as template for single-stranded cDNA synthesis using Taqman Reverse Transcription Reagents (Applied Biosystems, Foster City, CA). Real-time fluorescent detection of RT-PCR products was performed using an ABI Prism 7700 sequence detection system and Sybr Green PCR Master Mix (Applied Biosystems). The following PCR primers were designed for this study using Primer Express v. 1.0 (Applied Biosystems): cytochrome P450 (At3g28740) forward, 5'-TTGATGCCTTTTGGGATTGG-3' and reverse, 5'-CAAGGTCACTAGCCGTTGAGC-3'; 60S ribosomal protein L23A (At2g39460) forward, 5'-TCCAGACCAAGAAAGTGAACACA-3' and reverse, 5'-CATAGTCTGGTGTAAGCCTCACGT-3'; and aquaporin (At2g36830) forward, 5'-GCTTCTCGGCTCCGTCG-3' and reverse, 5'-GGCACAGCCAAGCCACC-3'. Amplification reactions were carried out according to the manufacturer's specifications as follows: two minutes at 50^0C followed by a 10 min activation of the enzyme at 95^0C, and 40 subsequent cycles consisting of 95^0C for 15 sec followed by 60^0C for 1 min. All amplification reactions were run using a dilution series of cDNA from either control or

TNT-treated tissues using the same PCR master mix. Amplimers were checked for purity and size by gel electrophoresis to ensure that the correct sequence was amplified. Controls reactions omitting reverse transcriptase were run for all samples to ensure that genomic contamination did not contribute to the amplified products.

Acknowledgements

This work was supported by NNEMS Fellowship U-91587201-0 from the U.S. Environmental Protection Agency to D.R.E. We thank Arthur Karnaugh for help with DNA sequencing protocols, and Caroline Stevens for assistance with data analysis. Thanks also to MacArthur Long and Steve McCutcheon for help in initiating the project and its continued support, and Jeff Dangl for comments on the manuscript.

REFERENCES

1. Walsh, M.E., Jenkins, T.F., Schnitcker, P.S., Elwell, J.W., & Stutz, M.H. USA Cold Regions Research and Engineering Laboratory CRREL Special Report 93-5, Hanover NH, pp 1-17 (1993).
2. Hannink, N.K., Rosser, S.J., & Bruce, N.C. Phytoremediation of explosives. Crit. Rev. Plant Sci. **21**, 511-538 (2002).
3. French, C. E., Rosser S.J., Davies, G.J., Nicklin, S., & Bruce, N.C. Biodegradation of explosives by transgenic plants expressing pentaerythritol tetranitrate reductase. Nat. Biotech. **17**, 491-494 (1999).
4. Hannink, N., et al. Phytodetoxification of TNT by transgenic plants expressing a bacterial nitroreductase. Nat. Biotech. **19**, 1168-1172 (2001).
5. Velculescu, V.E., Zhang, L., Vogelstein, B., & Kinzler, K.W. Serial analysis of gene expression. Science **270**, 484-487 (1995).
6. Velculescu, V.E., et al. Characterization of the yeast transcriptome. Cell **88**, 243-251 (1997).
7. Riggins, G.J., & Strausberg, R.L. Genome and genetic resources from the Cancer Genome Anatomy Project. Hum. Mol. Genet. **10**, 663-667 (2001).
8. Matsumura, H., Nirasawa, S., & Terauchi, R. Technical advance: transcript profiling in rice (Oryza sativa L.) seedlings using serial analysis of gene expression. Plant J. **20**, 719-726 (1999).
9. Lorenz, W.W., & Dean, J.F.D. SAGE Profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). Tree Physiol. **22**, 301-10 (2002).

10. Schultz, T.F., Kiyosue, T., Yanovsky, M., Wada, M. & Kay, S.A. A role for LKP2 in the circadian clock of Arabidopsis. Plant Cell **13**, 2659-2670 (2001).
11. Neuteboom, L.W., et al. Isolation and characterization of cDNA clones corresponding with mRNAs that accumulate during auxin-induced lateral root formation. Plant Mol. Biol. **39**, 273-287 (1999).
12. Goodwin, W., Pallas, J.A., & Jenkins, G.I. Transcripts of a gene encoding a putative cell wall plasma membrane linker protein are specifically cold-induced in *Brassica napus*. Plant Mol. Biol. **31**, 771-781 (1996).
13. Stromvik, M.V., Sundararaman, V.P., & Vodkin, L.O. A novel promoter from soybean that is active in a complex developmental pattern with and without its proximal 650 base pairs. Plant Mol. Biol. **41**, 217-231 (1999).
14. Evans, S.J., et al. Evaluation of Affymetrix gene chip sensitivity in rat hippocampal tissue using SAGE analysis. Serial Analysis of Gene Expression. Eur. J. Neurosci. **16**, 409-13 (2002).
15. Pavlostathis, S.G., & Jackson, G.H. (1999). Biotransformation of 2,4,6-trinitrotoluene in *Anabaena* sp. cultures. Environ. Toxicol. Chem. **18**, 412-419 (1999).
16. Lucero, M.E., Mueller, W., Hubstenberger, J., Phillips, G.C., & O'Connell, M.A. Tolerance to nitrogenous explosives and metabolism of TNT by cell suspensions of *Datura innoxia*. In Vitro Cell Dev. Plant **35**, 480-486 (1999).
17. Cataldo, D.A., Harvey, S.D., Fellows, R.J., Bean, R.M., & McVeety, B.D. An Evaluation of the Environmental Fate and Behavior of Munitions Material (TNT,

RDX) in Soil and Plant Systems (Report AD-A223, 546). Pacific Northwest Laboratories, Richland, WA (1989).

18. The Arabidopsis Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408**, 796-815 (2000).
19. Madden, S., Galella, E., Zhu, J., Bertelsen, A., & Beaudry, G. SAGE transcript profiles for p53-dependent growth regulation. Oncogene **15**, 1079-1085 (1997).
20. Velculescu, V.E., et al. Analysis of human transcriptomes. Nat. Genet. **23**, 387-388 (1999).
21. Stöllberg, J., Urschitz, J., Urban, Z., & Boyd, C.D. A quantitative evaluation of SAGE. Genome Res. **10**,1241-1248 (2000).
22. Jones, S.J., et al. Changes in gene expression associated with developmental arrest and longevity in *Caenorhabditis elegans*. Genome Res. **11**, 1346-1352 (2001).
23. Patankar, S., Munasinghe, A., Shoaibi, A., Cummings, L.M., & Wirth, D.F. Serial analysis of gene expression in *Plasmodium falciparum* reveals the global expression profile of erythrocytic stages and the presence of anti-sense transcripts in the malarial parasite. Mol. Biol. Cell. **12**, 3114-3125 (2001).
24. Ishikawa, T. The ATP-dependent glutathione S-conjugate export pump. Tr. Biochem. Sci. **17**, 433-438 (1992).
25. Ishikawa, T., Li, Z-S., Lu, Y-P., & Rea, P.A. The GS-X pump in plant, yeast, and animal cells: Structure, function, and gene expression. Biosci. Rep. **17**, 189-207 (1997).
26. Kreuz, K., Tommasini, R., & Martinoia, E. Old enzymes for a new job - Herbicide detoxification in plants. Plant Physiol. **111**, 349-353 (1996).

27. Bhadra, R., Wayment, D.G., Hughes, J.B., & Shanks, J.V. Confirmation of conjugation processes during TNT metabolism by axenic plant roots. Environ. Sci. Technol. **33**, 446-452 (1999a).

28. Bhadra, R., Wayment, D.G., Hughes, J.B., & Shanks, J.V. Characterization of oxidation products of TNT metabolism in aquatic phytoremediation systems of *Myriophyllum aquaticum*. Environ. Sci. Technol. **33**, 3354-3361 (1999b).

29. Davies, J., & Caseley, J.C. Herbicide safeners: a review. Pestic. Sci. **55**, 1043-1058 (1999).

30. Kreuz, K. , Gaudin, J., Stingelin, J. & Ebert, E. Metabolism of aryloxyphenoxypropanoate herbicide, CGA 184927, in wheat, barley and maize:differential effects of the safener CGA 185072. Z Naturforschung **46**, 901-905 (1991).

31. Yang, Q., Reinhard, K., Schiltz, E., & Matern, U. Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA : anthranilate N-hydroxycinnamoyl/benzoyltransferase from elicited cell cultures of carnation, *Dianthus caryophyllus* L. Plant Mol. Biol. **35**, 777-789 (1997).

32. Beuerle, T., & Pichersky, E. Purification and characterization of benzoate:coenzyme A ligase from *Clarkia breweri*. Arch. Biochem. Biophys. **400**, 258-264 (2002).

33. Schröder, P., & Collins, C. Conjugating enzymes involved in xenobiotic metabolism of organic xenobiotics in plants. Intl. J. Phytoremed. **4**, 247-265 (2002).

34. Martinoia, E., Grill, E., Tommasini, R., Kreuz, K., & Amrhein, M. ATP-dependent glutathione S-conjugate export pump in the vacuolar membrane of plants. Nature **364**, 247-249 (1993).

35. Li, Z-S., Zhao, Y., & Rea, P.A. Magnesium adenosine 5'-triphosphate-energized transport of glutathione-S-conjugates by plant vacuolar membrane vesicles. Plant Physiol. **107**, 1257-1268 (1995).

36. Theodoulou, F.L. Plant ABC transporters. Biochim. Biophys. Acta **1465**, 79-103 (2000).

37. Mitterbauer, R., & Adam, G. *Saccharomyces cerevisae* and *Arabidopsis thaliana*: Useful model systems for the identification of molecular mechanisms involved in resistance of plants to toxins. Eur. J. Plant Pathol. **108**, 699-703 (2002).

38. Tommasini, R., et al. Differential expression of genes coding for ABC transporters after treatment of *Arabidopsis thaliana* with xenobiotics. FEBS Lett. **411**, 206-210 (1997).

39. Schaller, F., & Weiler, E.W. Molecular cloning and characterization of 12-oxophytodienoate reductase, an enzyme of the octadecanoid signaling pathway from *Arabidopsis thaliana* - Structural and functional relationship to yeast old yellow enzyme. J. Biol. Chem. **272**, 28066-28072 (1997).

40. Chang, S., Puryear, J., & Cairney, J. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. **11**, 113-116 (1993).

41. Paquette, S.M., Bak, S., & Feyereisen, R. Intron-exon organization and phylogeny in a large superfamily, the paralogous cytochrome P450 genes of *Arabidopsis thaliana*. DNA Cell Biol. 19, **307**-17 (2000).

FIGURE LEGENDS

Figure 2.1. Chemical Structures of Nitroaromatic Explosives.

The nitroaromatic explosives of greatest concern for environmental contamination, TNT, RDX, and HMX, are depicted.

Figure 2.2. Estimates of Transcriptome Size in *Arabidopsis* Roots.

Double-reciprocal plots of new tag discovery versus total tags sequenced are shown for the TNT-treated (top panel) and control (bottom panel) root libraries. The estimated transcriptome size was taken as the 1/y-intercept from the linear regression depicted in each graph.

Table 2.1. SAGE tags induced at least 10-fold by exposure to TNT

	Tag Abundance		Fold Increase[a]			
tag sequence	TNT	Control	TNT/ Control	p-chance[b]	Locus	Annotation
GTGAGTTTGA	30	0	>30.0	0	At2g18910 *	Possible signaling protein
CCAAATTCTG	55	2	27.5	0	At1g17170	Putative glutathione transferase
TAGCCAATTA	72	3	24.0	0	At5g11510	MYB-like protein
GGAGTTTGTA	23	1	23.0	0	At3g27740	Carbamoyl phosphate synthetase (small subunit)
ATGTTTCGCG	21	1	21.0	6.68E-06	At5g13750	Transporter-like protein
GGAGAAGTCC	20	1	20.0	1.67E-05	At1g60940	Putative serine/threonine-protein kinase
GCAATTCTAC	35	2	17.5	0	At5g03630	Monodehydroascorbate reductase-like protein
AGAAGTTTAT	17	1	17.0	8.67E-05	At5g63790	*No apical meristem* (NAM)-like protein
TGAGTTTCAA	17	1	17.0	8.67E-05	At4g01870	Unknown function
					At5g60880	Putative protein
GGTTAGTCGA	15	1	15.0	3.13E-04	At1g05050	Putative transcription factor
AGAGAAAGTG	14	0	>14.0	5.34E-05	At3g28740	Putative cytochrome P450 (CYP81D11-A-TYPE)
CGGGGAAAAA	13	1	13.0	9.70E-04	?	Polyadenylation sequence?
AGTTGAGTTC	38	3	12.7	0	At5g45020	Putative glutathione transferase
GTGAAGTTTG	12	0	>12.0	1.90E-04	At3g45270	Putative protein
GGAAAAGGTG	23	2	11.5	2.34E-05	At3g21720	Putative isocitrate lyase
ACCAAAATTG	11	0	>11.0	5.10E-04	At2g29490	Putative glutathione S-transferase
CCACAGTTTT	30	3	10.0	0	At1g78080	Putative AP2 domain protein (RAP2-like transcription factor, TINY)
AACGCAGAAA	10	0	>10.0	1.05E-03	At1g01550	Hypothetical protein
CAGGATGTGT	10	0	>10.0	1.05E-03	At1g05680	Putative indole-3-acetate beta-glucosyltransferase
GCACTCTTGA	10	0	>10.0	1.05E-03	At2g28570	Unknown protein (protein-tyrosine kinase motif)
CTTCTCTAGT	10	1	10.0	6.37E-03	At1g75270	Putative GSH-dependent dehydroascorbate reductase
CTTGTCCTCA	10	1	10.0	6.37E-03	At1g76680	12-oxophytodienoate reductase (OPR1)

[a]For the calculation of expression ratios, a value of 1 was substituted where tag counts were zero.
[b]P-chance values are averages of three Monte Carlo simulations
*If tag is positioned at penultimate CATG site.

Table 2.2. Cytochrome P450 genes induced by exposure to TNT

Tag sequence	Locus	TNT	Control	TNT/control	p-chance	P450	Type[a]
AGAGAAAGTG	At3g28740	14	0	>14.0	5.43E-05	CYP81D11 (putative)	A
ACACCAAAGC	At1g78490	11	2	5.5	1.17E-02	CYP708A3 (like protein)	NA
GAAAAGATTT	At2g30750	5	0	>5.0	3.10E-02	CYP71A12	A
GCTGAGAGAC	At4g22690	5	0	>5.0	3.10E-02	CYP706A2 (like protein)	A
ATGCGAAGCT	At2g30490	41	9	4.6	3.34E-06	CYP73A5 (type enzyme)	A
AAGCATCCGC	At1g64950	4	0	>4.0	6.08E-02	CYP89A6 (hypothetical)	A

[a]Cytochrome P450 classification as described elsewhere[41].

Table 2.3. TNT-induced SAGE tags representing potential detoxification pathway components.

	Tag Abundance		Fold increase[a]			
tag sequence	TNT	Control	TNT/ Control	p-chance[b]	Locus	Annotation
TTGAGAAATT	9	1	9.0	1.06E-02	At5g54500	1,4-Benzoquinone reductase-like
ATTCTGAGAA	9	1	9.0	2.04E-03	At1g78340	Glutathione transferase-like protein
TGATGAGTTT	94	11	8.5	0	At3g09390	Metallothionein-like protein
TGGCGGATTA	40	5	8.0	0	At4g19880	Putative protein (related to glutathione S-tranferase)
AAGATCCAAG	8	1	8.0	1.96E-02	At2g18160	G-box binding bZIP transcription factor
TGCAAGTTAT	8	1	8.0	1.96E-02	At3g28480	Prolyl 4-hydroxylase, putative
TAGAATTCTC	7	1	7.0	3.50E-02	At4g03430	Putative pre-mRNA splicing factor
TGATTCAAAA	7	1	7.0	3.50E-02	At2g16680	Putative non-LTR retroelement reverse transcriptase
AAACTGTTTG	7	1	7.0	3.50E-02	At5g48930	Anthranilate N-benzoyltransferase
TCACTCCTAT	6	1	6.0	6.52E-02	At3g44300	Nitrilase 2
CAAATCAGTT	33	6	5.5	1.34E-05	At1g76930	Extensin
TCTCGAACCT	11	2	5.5	1.17E-02	At5g20830	Sucrose-UDP glucosyltransferase
GCTGTTTTTG	155	30	5.2	0	At1g07890	L-ascorbate peroxidase
TGTTTGGCTG	5	0	>5.0	3.10E-02	At3g58500	Phosphoprotein phosphatase 2A isoform
CCAATTAGTC	10	2	5.0	2.08E-02	At3g53480	ABC transporter - like protein
GAAACGCTCA	10	2	5.0	2.08E-02	At5g07460	Peptide methionine sulfoxide reductase-like protein
GTTTCGAGAT	47	10	4.7	0	At4g11600	Phospholipid hydroperoxide glutathione peroxidase
AAAACTCGGT	9	2	4.5	3.39E-02	At3g24170	Glutathione reductase, cytosolic
CTTGGTGCAA	11	3	3.7	3.07E-02	At2g44350	Citrate synthase
ACGAAGGTCG	9	3	3.0	7.08E-02	At3g44320	Nitrilase 3
TAACTTGTGC	20	7	2.9	1.06E-02	At1g51680	4-Coumarate:CoA ligase
GAAACTTAAA	31	11	2.8	1.45E-03	At4g30170	Peroxidase ATP8a

[a]For the calculation of expression ratios, a value of 1 was substituted where tag counts were zero.
[b]P-chance values are averages of three Monte Carlo simulations.

Table 2.4. SAGE tags repressed at least 14-fold by exposure to TNT

	Tag abundance		Fold decrease[a] Control/			
Tag sequence	Control	TNT	TNT	p-chance[b]	Locus	Annotation
TTTTCTTACC	85	0	>85.0	0	At4g12550	Putative cell wall-plasma membrane linker protein (AIR1A)
TATCCTTGTT	52	1	52.0	0	At2g01520	Major latex protein homolog-like
TTGTATGTTT	45	0	>45.0	0	At1g70850	Major latex protein homolog-like
TTTTCTTATC	37	0	>37.0	0	At4g12550	Putative cell wall-plasma membrane linker protein (AIR1B)
GACCAACCAC	29	1	29.0	0	At2g36830	Putative aquaporin
TGTCTTAGCT	29	1	29.0	0	At1g09690	Putative 60S ribosomal protein L21
TTTATGCTTT	50	2	25.0	0	At4g22212	Expressed protein
TGTTTATTTT	24	1	24.0	0	At5g23710	Unknown protein
TTTTCTTCTT	23	1	23.0	3.34E-06	At3g54580	Extensin precursor-like protein
TTACAATAAC	21	1	21.0	0	At5g19510	Elongation factor 1B alpha-subunit
GGAACATATA	20	0	>20.0	0	At4g38740	Peptidylprolyl isomerase ROC1
GTGTTGTATG	20	0	>20.0	0	At1g14960	Putative major latex protein
TTGGTTATGT	20	0	>20.0	0	At5g56540	Arabinogalactan-protein (AGP14)
AACCCGGCCA	16	0	>16.0	2.34E-05	At4g17340	Membrane channel-like protein
TTTTTCTTTG	15	0	>15.0	2.00E-05	At4g12520	pEARLI 1-like protein
TCTTTGTCTT	15	1	15.0	1.87E-04	At5g10280	Putative transcription factor MYB92
AGTTTATCAC	14	0	>14.0	4.00E-05	At5g04750	F1F0-ATPase inhibitor-like protein
TTATCTCTCT	14	0	>14.0	4.00E-05	At1g04820	Tubulin alpha-2/alpha-4 chain
TTTTGCTATC	56	4	14.0	0	At5g26260	Putative protein
CAACATTGTA	42	3	14.0	0	At5g14200	3-isopropylmalate dehydrogenase
TTTCTGGTAA	14	1	14.0	4.90E-04	At4g14960	Tubulin alpha-6 chain (TUA6)

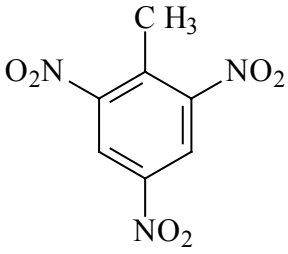
[a]For the calculation of expression ratios, a value of 1 was substituted where tag counts were zero.
[b]P-chance values are averages of three Monte Carlo simulations.

Table 2.5. Comparison of transcript expression levels as determined by SAGE and qPCR.

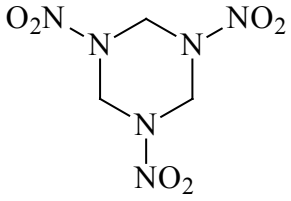
	Control		TNT		Ratio
Amplicon/SAGE tag	qPCR C_t value[a]	SAGE count	qPCR C_t value	SAGE count	qPCR[b]
Cyt P450	35	0	29	14	28X increase
Ribosomal Protein	23	31	24	34	2X decrease
Aquaporin	25	29	29	1	11X decrease

[a]C_t is a function of the starting number of templates present.
[b]Expression ratios relative to control.

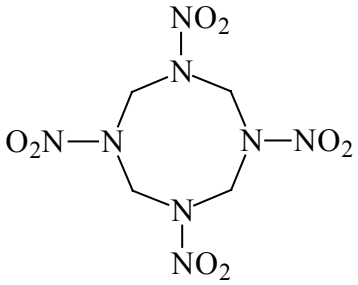
Figure 2.1.



TNT
(trinitrotoluene)

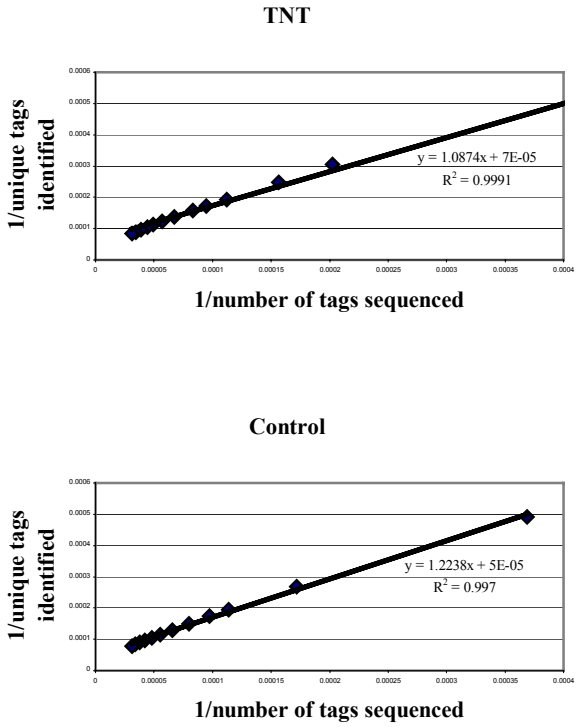


RDX
(hexahydro-1,3,5-trinitro-1,3,5-triazine)



HMX
(octahydro-1,3,5,7-tetranitro-1,3,5,7-tetrazocine)

Figure 2.2.


TNT
1/unique tags identified
y = 1.0874x + 7E-05
R2 = 0.9991
1/number of tags sequenced
Control
1/unique tags identified
y = 1.2238x + 5E-05
R2 = 0.997
1/number of tags sequenced

CHAPTER 3

EFFECTS OF THE EXPLOSIVE RDX ON TRANSCRIPT EXPRESSION IN *ARABIDOPSIS THALIANA* SEEDLING ROOTS[2]

[2]Ekman, D.R., Wolfe, N.L., and Dean, J.F.D. To be submitted to *Environmental Toxicology and Chemistry*.

ABSTRACT

Arabidopsis thaliana root transcriptome responses to the munition 1,3,5-hexahydro-1,3,5-triazine (RDX) were measured using serial analysis of gene expression (SAGE). Sequencing of SAGE libraries from both control and RDX-exposed root tissue to approximately 30,000 tags each revealed the induction of a number of general stress response transcripts. Induced genes included ones encoding molecular chaperones and transcription factors, as well as vacuolar proteins and peroxidases. Among the strongly repressed genes were ribosomal proteins, a cyclophilin, a katanin, and a peroxidase. Comparison of the transcriptional profile for the RDX response to that induced by trinitrotoluene (TNT) exposure revealed significant differences. This suggests that *Arabidopsis* employs drastically different mechanisms for coping with these two compounds. With respect to the goal of engineering plants better capable of tolerating and remediating explosives at contaminated sites, this study suggests that different genes will be necessary to deal effectively with each type of explosive. **(Note: the datasets from these analyses are available at the ncbi Gene Expression Omnibus (GEO) at: http://www.ncbi.nlm.nih.gov/geo/)**

INTRODUCTION

Extensive contamination of a large number of munitions manufacturing and disposal facilities throughout the U.S. and Europe over the past several decades by the polynitramine, hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) (Figure 3.1), has created a

significant environmental problem. An important military explosive, RDX is a compound of recognized toxicity (Cholakis et al., 1980). Often used as a rat poison, RDX ingestion produces adverse effects in the central nervous system, gastro-intestinal tract, and kidneys. Moreover, biological transformation of RDX can frequently serve to increase its toxicity. Hexahydro-1,3,5-trinitroso-1,3,5-triazine, a degradation product of RDX, has been used as an experimental tumorigen (Merck & Co., 1983; Sax, 1989).

As a cost-effective alternative to currently used methods for clean-up (i.e., excavation of the contaminated soil followed by incineration), phytoremediation, or the use of plants to aid in the removal and/or degradation of environmental pollutants, is being assessed for its potential to address RDX contamination (Best et al., 1997b; Best et al., 1997a; Thompson et al., 1999; Bhadra et al., 2000). Unfortunately, although the RDX in the surrounding environment was taken up rapidly by the plants in these studies, most of it remained untransformed and ended up in the arial portions where it could be introduced into the foodchain (Harvey et al., 1991). Therefore, plants capable of tolerating high levels of RDX while simultaneously restricting the compound to the roots and promoting its degradation there would be highly desirable. Unfortunately, virtually nothing is known about the plant genes involved in RDX metabolism. Consequently, this study was undertaken to identify key enzymes in the metabolism of this energetic compound using serial analysis of gene expression (SAGE) (Velculescu et al., 1995; Velculescu et al., 1997), a technique for transcriptional profiling. The SAGE technique provides a quantitative snapshot of transcript levels in sampled tissues and has received its widest application in the study of gene expression changes in human cancers (Riggins and Strausberg, 2001). In this study we report changes in root transcript levels for liquid-

grown *Arabidopsis* seedling roots exposed to RDX and compare them with the transcript profiles previously measured in similar tissues exposed to TNT.

MATERIALS AND METHODS

Plant Material, Growth Conditions, and Root Tissue Isolation

The following experimental descriptions were developed to conform as closely as possible to the MIAME guidelines being promulgated for describing microarray experiments (Brazma et al., 2001). *A. thaliana* ecotype Columbia seeds (WT-2, Lehle Seeds, Round Rock, TX) were surface-sterilized by immersion in 70% ethanol for two min, followed by immersion in a solution of 30% bleach, 0.2% triton X-100 for 40 minutes with gentle shaking. The sterilized seeds were then placed in sterile Murashige and Skoog liquid medium prepared according to the manufacturer's protocol (Invitrogen Life Technologies, Carlsbad, CA). Plants were grown for 14 days at 25^0C under a 16-hour photoperiod with constant shaking at 85 rpm in a growth chamber. RDX was obtained from the U.S. Army Center for Environmental Health Research (Fort Detrick, MD). Toxicity was assessed by adding RDX from a stock solution in DMSO to yield final concentrations of 0, 5, 20, 40, 80, 100, and 150 mg/L in MS medium with cultures run in triplicate. The cultures were returned to the growth chamber for five days, during which time the seedlings in each flask were examined for signs of stress (leaf chlorosis and necrosis). A final concentration of 150 mg/L RDX was judged to produce notable stress in the plants without causing death. Root tissues for SAGE library construction

were isolated from seedlings grown 14 days in liquid MS medium and dosed with RDX to a final concentration of 150 mg/L. Control tissues were isolated from seedlings grown under the same conditions and dosed with an equivalent volume of DMSO. The seedlings were grown in the presence of RDX or DMSO for 24 hours, after which they were submerged briefly in dH_2O to remove excess medium, and excised roots were immediately frozen in liquid nitrogen. All root tissues were stored at -80^0C prior to RNA extraction. The root tissues for each SAGE library were pooled from approximately 1400 seedlings recovered from 36 treatment flasks. Prior to seeding, the 36 flasks were divided into three replicates of 12 with each group being started on different days.

RNA Isolation and cDNA Synthesis

Total RNA from root tissues was extracted using the LiCl precipitation technique of Chang et al. (1993). Poly(A) RNA was isolated from total RNA using Dynabeads oligo-dT(25) magnetic beads (Dynal Biotech, Lake Success, NY) at a ratio of 0.25 mg of total RNA per 250 μL of Dynabeads and following the manufacture's instructions. Double-stranded cDNA was synthesized from 5 mg of poly(A) RNA using the Superscript Choice cDNA synthesis kit (Invitrogen Life Technologies) and following the manufacturer's protocol, except for the substitution of a 5'-biotin dT(18) primer in the first-strand reaction.

SAGE library construction

SAGE libraries were constructed according to the SAGE Detailed Protocol, Version 1.0c (Velculescu et al., 1997a), a brief description of which follows. Biotinylated cDNAs from each tissue sample were bound to streptavidin-coated magnetic beads and digested with NlaIII, a restriction enzyme recognizing the four-base sequence, CATG (anchoring enzyme). DNA released by this digestion was washed away and the beads, with the adherent 3' ends of each cDNA, were split into two pools. Linkers containing a binding site for BsmF1 (a Type-II restriction endonuclease – the tagging enzyme), but different sites for PCR primers, were ligated to the NlaIII cleavage site at the 5' ends of the bead-bound cDNA fragments in each pool. Both pools of cDNAs were digested with BsmF1 to release SAGE tags from the beads, after which the pools were combined, and 102 bp linker-flanked ditags were formed by blunt-end ligation. Following amplification of the ditags by PCR, the linkers were removed by NlaIII digestion, and ditags were ligated to form concatemers. The concatemers were subsequently size-fractionated, ligated into the pZero vector (Invitrogen Life Technologies), cloned, and sequenced.

SAGE Data Analysis

Sequence files were compiled and analyzed using the SAGE Software, ver. 3.03, provided by Dr. Kenneth Kinzler (Johns Hopkins University, Baltimore, MD). Tags containing linker sequences and repeated ditags were excluded prior to analysis. Because

the library representing control roots was sequenced to a slightly greater extent (31,973 tags for the control versus 31,209 tags for the RDX treatment), values for the RDX library tags were normalized prior to making comparisons of relative gene expression. Ratios were used to compare the relative expression of tags between the two libraries, and in instances where a particular tag was absent from a library, a value of 1 was substituted to avoid division by zero. Using the SAGE software, Monte Carlo simulations were performed to estimate the statistical significance of any differential expression. The null hypothesis for these analyses was that the abundance, type, and distribution of transcripts were the same in both libraries.

Gene Identification

To identify the genes from which the tags were derived, each 10-base tag plus the 4-base NlaIII recognition sequence was first compared against the Arabidopsis Gene Initiative (AGI) database of model genes using the Patmatch analysis tool available on The Arabidopsis Information Resource (TAIR) server (http://www.arabidopsis.org). If the tag was found to match exactly the NlaIII site closest to the 3' end of a model gene, this identity was accepted for the tag. Tags that could not be found in the model gene database were compared against all *Arabidopsis* sequences in GenBank using the same Patmatch tool. Exact matches were annotated accordingly.

Quantitative PCR

Total RNA from *Arabidopsis* plants grown under conditions identical to those used to generate RNA for the SAGE studies was used to independently verify the expression of selected genes using quantitative RT-PCR. Messenger RNA isolated using Dynal Oligo dT_{25} magnetic beads served as template for single-stranded cDNA synthesis using Taqman Reverse Transcription Reagents (Applied Biosystems, Foster City, CA). Real-time fluorescent detection of RT-PCR products was performed using a DNA Engine Opticon System (MJ Research, Inc., Boston, MA) and Sybr Green PCR Master Mix (Applied Biosystems). The following PCR primers were designed for this study using Primer Express v. 1.0 (Applied Biosystems): nitrilase2 (At3g44300) forward, 5'-GGACATTACTCGAGACCAGATGTTT -3' and reverse, 5'-TCGAAATGAATGTGACCGGTT -3'; peptidylprolyl isomerase (At2g39460) forward, 5'-CGAGATGGAGCTTTTCGCTG-3' and reverse, 5'-CCGGTACAGAGAGCACGGAA-3'; and putative HMG protein (At2g17560) forward, 5'-TGGTAAGGCTGCTGGAGCTAG -3' and reverse, 5'-GCTCTCAGCCTTAGCGACGT -3'. Amplification reactions were carried out according to the manufacturer's specifications as follows: two min at 50^0C followed by a 10 min activation of the enzyme at 95^0C, and 40 subsequent cycles consisting of 95^0C for 15 sec followed by 60^0C for 1 min. All amplification reactions were run using a dilution series of cDNA from both control and RDX-treated tissues in the same PCR master mix. Amplimers were checked for purity and size by gel electrophoresis to ensure that the correct sequence was amplified. All samples were treated with DNAse prior to reverse

transcription using DNA-free™ (Ambion, Inc., Austin, TX) to remove any contaminating genomic DNA. Control reactions omitting reverse transcriptase were run for all samples to ensure that genomic contamination did not contribute to the amplified products. Melt curve analysis confirmed the absence of both primer-dimer and nonspecific product formation.

RESULTS AND DISCUSSION

Toxicity of RDX to sterile hydroponic cultures of *Arabidopsis thaliana*

Sterile *Arabidopsis* seedlings were grown for 14 days in liquid Murashige and Skoog medium before treatment with RDX at concentrations ranging from 5 to 150 mg/L. Visual signs of toxicity (mainly leaf chlorosis and necrosis) were not evident in any of the cultures, except at the highest dosage (150 mg/L) after 72 hours of exposure. While this concentration was in excess of the aqueous solubility of RDX (approximately 40 mg/L) it was nonetheless necessary to induce signs of toxicity in these plants. This may be due to a requirement for accumulation of the compound within the plant tissues. This would require a pool of RDX in the media that, although initially insoluble, would enter into solution as portions of the soluble fraction were removed from the media by the plants.

SAGE Libraries

SAGE libraries representing the transcripts expressed in *Arabidopsis* root tissues grown in the presence or absence of RDX were sequenced until just over 30,000 tags were characterized from each. The datasets from these SAGE library analyses are available online in spreadsheet format (http://www.arches.uga.edu/~jeffdean/SAGE/SAGEData.html). As shown in Table 3.1, 12,122 of the 31,209 tags characterized in the library for RDX-treated tissue represented unique transcripts, and 7,909 of these were seen only once (singletons). Similarly, 12,719 unique tags were encountered amongst the 31,973 characterized in the control library, with 8,322 of these representing singletons.

Differentially Expressed Transcripts

Significant differences in transcript expression were observed between the SAGE libraries. In the presence of RDX, 135 tags displayed at least a five-fold increase in expression over control. Moreover, a relatively large number of tags observed in the RDX library (6672) were absent from the control library. Among the induced tags were those representing transcripts from several different functional classes, including defense-related proteins, transporters, transcription factors, and vacuolar proteins (Table 3.2). The most highly induced tag corresponds to an NPR1-like protein previously shown to respond to pathogen attack (Cao et al., 1994; Delaney et al., 1995; Glazebrook et al., 1996; Shah et al., 1997; Cao et al., 1998; Chern et al., 2001; Friedrich et al., 2001). The

second most highly induced tag represented a DnaJ-like protein known to act as a molecular chaperone or co-chaperone in association with Hsp70 (Kushnir et al., 1995; Zhou et al., 1995; Kroczynska et al., 1996; Best et al., 1997b), which potentially signifies stress-related problems with protein folding and/or stability in the cells exposed to RDX. Also noteworthy is the induction of a vacuolar processing enzyme (γ-VPE) responsible for the maturation of various cellular proteins, including a number that function in cellular defense (Kinoshita et al., 1999). Cytochrome P450 enzymes are often utilized for transformation of cellular toxins, either to make them more amenable to conjugation (e.g., attachment of glutathione or six-carbon sugars) or to increase solubility. The involvement of cytochrome P450 enzymes has been reported previously in the degradation of RDX by two different strains of *Rhodococcus* sp., YH11 (Tekoah and Abeliovich, 1999) and DN22 (Coleman and Duxbury, 1999). Interestingly, three *Arabidopsis* cytochrome P450s were induced at least five-fold in the presence of RDX. Also worth noting is the strong induction (12X) of a putative peroxidase transcript. This may indicate oxidative stress in the plant roots or, alternatively, a mechanism for oxidation of RDX. The fungus, *Phanerochaete chrysosporium*, is known to degrade the nitroaromatic explosive, TNT, as well as a number of other organic pollutants using oxidative enzymes such as peroxidases. Degradation of RDX by this fungus has been reported to proceed by an oxidative mechanism, presumably catalyzed by lignin peroxidases (Fernando and Aust, 1991).

Over 100 tags were repressed at least five-fold in response to RDX exposure (Table 3.3). One of the tags showing the greatest repression corresponds to a protein involved in cytokinesis (cell division), possibly indicating a reduced growth rate in the

presence of RDX. Other tags strongly repressed by RDX included a putative HMG (high mobility group) protein. This class of proteins (five have been identified in *Arabidopsis*) is thought to act as architectural elements in chromatin structure, and through their impact on chromatin are suspected to influence such processes as transcription and recombination (Crothers, 1993; Ner et al., 1994; Grosschedl, 1995; Grasser, 1998). Decreased expression was also noted for a tag representing a glutaredoxin protein. This protein is thought to function with ribonucleotide reductase in the production of deoxyribonucleotides (Holmgren, 1989). Alternatively, rice glutaredoxin was found to have dehydroascorbate reductase activity (Minakuchi et al., 1997), which along with the potential for glutaredoxin to participate in general thiol reduction, suggests that oxidative stress could result from repression of this gene. Interestingly, the tag for a putative monodehydroascorbate reductase was also observed to decrease in abundance upon RDX exposure.

Quantitative PCR

SAGE has previously been shown to provide an accurate reflection of gene expression levels for medium- and high-abundance transcripts, but the induction or repression of selected *Arabidopsis* transcripts in response to RDX exposure was independently verified in this study using real-time quantitative PCR. Genes that SAGE analysis suggested were induced (At3g44300; nitrilase 2), repressed (At2g17560; HMG protein), or remained unaffected (At2g39460; peptidyl prolyl isomerase) by RDX exposure were tested. As shown in Table 3.4, the quantitative PCR data were in general

agreement with the SAGE data. Transcripts measured as induced or repressed using SAGE were observed to follow similar trends when quantitated using qPCR. However, the transcript unaffected by the presence of RDX as measured using SAGE displayed a 3-fold decrease in abundance according to qPCR. The relatively high expression levels for this transcript indicated by SAGE (tag counts of around 40 in both libraries) would suggest that a sampling effect is not responsible for this contradiction. It is more likely that transcripts for other peptidyl-prolyl isomerase enzymes with similar sequences may have been detected by the qPCR technique. In fact, the SAGE data revealed repression of three other transcripts for variants (At4g38740, At1g265502, At2g47320) of this enzyme.

Differences In Transcript Expression between TNT-exposed and RDX-exposed *Arabidopsis* Roots

A previous study of gene expression changes induced in *Arabidopsis* by the presence of another toxic explosive, 2,4,6-trinitrotoluene (TNT), identified several highly induced transcripts (Ekman et al., 2003). A comparison with the RDX treatment data reveals significant differences in the responses to these two compounds (Table 3.5). For example, the tags representing different cytochrome P450 enzyme transcripts vary considerably between the treatments. Whereas a putative CYP81D11 transcript was induced at least 14-fold (14 to 0) in the TNT SAGE library, it was not detected in RDX-treated tissues. Also absent from the RDX-treated tissue was a transcript encoding a glutathione S-transferase, the most highly induced transcript (27X) in the TNT library.

Tags representing oxidative stress enzymes, such as monodehydroascorbate reductase, were low in the RDX library while their expression was highly induced in the TNT-treated roots (data not shown). Thus, while metabolism of TNT in plants probably requires a multiphase oxidative mechanism, as was described previously (Ekman et. al. 2003) and as has been observed in plants for a variety of other compounds, this may not be the case for RDX. This will be a major consideration in the development of plants for remediation of sites contaminated with RDX. Furthermore, if plants are to be able to remediate both of these explosives simultaneously (TNT and RDX are often found together at polluted sites) it may require the introduction of suites of genes specific for each compound.

Analysis of transcripts expressed commonly among all three libraries (control, TNT, and RDX) and those unique to each treatment gives an indication of the specificity of gene expression induced by each compound. As indicated in Figure 3.2, it appears as though a number of genes (3,572) cannot be dispensed with despite adverse conditions. Most likely, the majority of these are housekeeping genes or those specifically employed for growth under the culture conditions used. On the other hand, a significant number of genes were specific to each library. The largest number of genes expressed specifically for any given library was that for the control library (2,185). This potentially indicates the reduced expression of genes that are not completely necessary for survival in the munition-treated plants (1,994 for TNT and 1,745 for RDX). This sacrifice may be a result of the increased demand for higher expression levels of specific genes utilized in responding to toxicity. It should be noted that these numbers are derived from a depth of

sequencing of around 30,000 tags in each library. Further sequencing will likely alter these numbers.

CONCLUSIONS

From these analyses it appears as though the mechanism for RDX metabolism in *Arabidopsis* is vastly different than for TNT. This is somewhat expected from the results of previous studies analyzing the fate of RDX in plants (Harvey et al., 1991; Best et al., 1997b; Burken and Schnoor, 1998; Thompson et al., 1999). Identification of induced cytochrome P450 transcripts suggests a potential route to the tranformation of RDX, although the functions of these enzymes may not be directly related to RDX metabolism, but simply reflect a response to the general stress imposed by toxicity.

Future studies must include analyses of transcriptome responses in leaf and stem tissues. As these tissues are commonly sites of RDX accumulation in plants, discovery of differential gene expression in these locations is crucial to achieving a full understanding of this compound's metabolism in plants. Such studies are underway using high-density DNA microarrays to measure expression changes in these tissues and to observe the effects of longer periods of exposure and different concentrations of RDX. Eventually, transcriptome analyses of soil-grown plants exposed to RDX will be necessary to determine responses likely to be found under field conditions.

FIGURES LEGENDS

Figure 3.1. Explosive compounds of particular concern as environmental contaminants. TNT and RDX are highly recalcitrant, toxic compounds found extensively in soil, as well as in ground and surface water, at or near many military installations.

Figure 3.2. Diagrammatic representation of the tags unique to each library (control, TNT, or RDX) as well as those shared among subsets of the libraries.

TABLES

Table 3.1. SAGE library statistics.

	Frequency distribution[a]				
	≥20	19 to 5	4 to 2	=1	Totals
RDX library					
Unique tags	153 (1.3)[b]	1,021 (8.4)	3,039 (25.1)	7,909 (65.2)	12,122
Tags sequenced	7,063 (22.6)	8,498 (27.2)	7,739 (24.8)	7,909 (25.4)	31,209
Control library					
Unique tags	178 (1.4)	1,005 (7.9)	3,214 (25.3)	8,322 (65.4)	12,719
Tags sequenced	7,276 (22.8)	8,219 (25.7)	8,156 (25.5)	8,322 (26)	31,973

[a]Frequency distributions were calculated based on the total number of unique or sequenced tags in each library shown in the totals column.
[b]The percentages of tags in each frequency group are shown parenthetically

Table 3.2. SAGE tags induced at least 7-fold by exposure to RDX

	Tag Abundance		Fold Increase[a]			
tag sequence	Control	RDX	RDX/Control	p-chance[b]	Locus	Annotation
GTGGTAACGG	1	30	30.0	0.00	At4g26120	NPR1 like protein
TGCTTACCGT	1	14	14.0	0.00	At4g36040	DnaJ-like protein
GGATAACATC	1	14	14.0	0.00	At4g32940	gamma-VPE (vacuolar processing enzyme)
GGTTAGTCGA	1	13	13.0	0.00	At1g05050	putative transcription factor
CGCTGACATA	1	12	12.0	0.00	At1g49570	peroxidase, putative
TTCAAGTCCA	1	11	11.0	0.00	At5g66120	3-dehydroquinate synthase, putative
GCCGTTCTTA	3	32	10.7	0.00	At3g42050	Vacuolar H+-ATPase subunit H (VHA-H)
TCTTCTCGAA	1	10	10.0	0.01	At4g38920	H+-transporting ATPase 16K chain P2, vacuolar
GTGATGCTCT	1	10	10.0	0.01	At5g10300	alpha-hydroxynitrile lyase-like protein
TCCCCTATTA	1	9	9.0	0.01	********	no matches in genome
GGAGACAGTG	1	9	9.0	0.01	At4g18930	putative protein
GATGTCTGGC	1	9	9.0	0.01	At1g35516	myb family transcription factor
TTTGGAGGAG	2	17	8.5	0.00	At4g19200	similarity to glycine/proline-rich protein GPRP
TTCCTATTCT	2	17	8.5	0.00	At2g45960	aquaporin (plasma membrane intrinsic protein 1B)
TTTCTGGATA	1	8	8.0	0.02	At5g20150	ids4-like protein
TAATGTAATG	0	8	≥8.0	0.00	At5g04590	sulfite reductase
CCTGGAATCA	1	8	8.0	0.02	********	no matches in genome
ATCCTTGTCT	1	8	8.0	0.02	At2g37110	unknown protein
AGAAAGCAGG	1	8	8.0	0.02	At5g35200	putative protein
TTTTTTTTCT	0	7	≥7.0	0.01	2 HITS	AT1G33190, auxin-response factor (fragment) and AT3G11650.1, hairpin-induced protein, putative (HIN1)

[a]For the calculation of expression ratios, a value of 1 was substituted where tag counts were zero.
[b]P-chance values are averages of three Monte Carlo simulations.

Table 3.3. SAGE tags repressed at least 8-fold by exposure to RDX

	Tag Abundance		Fold Increase[a]			
tag sequence	Control	RDX	Control/RDX	p-chance[b]	Locus	Annotation
GTTCTGCAAA	13	1	13.0	0.00		hits on chromosomes 2 & 5
TTTAAAAAAA	11	1	11.0	0.00	At5g27030	putative protein; cytokinesis (cell division)
TCTGAAAGAG	11	1	11.0	0.00	At3g05590	putative 60S ribosomal protein L18
TTTCACACTT	10	1	10.0	0.01	At4g30170	peroxidase ATP8a
TTGGTGTTTC	10	1	10.0	0.01	At5g63030	glutaredoxin-like protein
GCCCTGCGAT	10	1	10.0	0.01		one hit on chromosome 4
AAAGGCGGCG	10	1	10.0	0.01	At2g17560	putative HMG protein
GGACAGATTC	9	0	≥9.0	0.00	At5g62300	ribosomal protein S20
GGAAATGAAC	9	0	≥9.0	0.00		genome hits chromosome 1, 2, 4, 5
GATCGACCAA	9	1	9.0	0.01	2 hits	At1g20310; hypothetical protein At1g74030; putative enolase
ATCATCATCA	9	1	9.0	0.01	*******	penultimate CATG At4g27840 ??? (numerous genome hits in both orientations --SSRs)
TAATTATGTA	8	0	≥8.0	0.00	At4g39675	Expressed protein
TAAGAAATCT	8	1	8.0	0.02	At4g17530	ras-related small GTP-binding protein RAB1c
TAAAAGCTTT	8	1	8.0	0.02	At2g47320	putative peptidyl-prolyl cis-trans isomerase
GTCTTAATGA	8	0	≥8.0	0.00		genome hits chromosome 1 & 2
GGAAATAAAA	8	1	8.0	0.02	At1g27030	expressed protein
GACGCTAGCG	8	1	8.0	0.02	At2g10410	pseudogene
ATGTGTTGGT	8	0	≥8.0	0.00	At2g34560	putative katanin
ATCAAAAAAA	8	0	≥8.0	0.00		one hit on chromosome 3 (1 ESTs)
AGGAAGGAAG	8	1	8.0	0.02	At5g03660	putative protein

[a]For the calculation of expression ratios, a value of 1 was substituted where tag counts were zero.
[b]P-chance values are averages of three Monte Carlo simulations.

Table 3.4. Comparison of transcript expression levels as determined by SAGE and qPCR.

Amplicon/SAGE tag	Control		RDX		Ratio
	qPCR C_t value[a]	SAGE count	qPCR C_t value	SAGE count	qPCR [b]
nitrilase 2 (At3g44300)	18	1	17	7	2X increase
peptidyl prolyl isomerase (At2g39460)	15	44	17	42	3X decrease
HMG protein (At2g17560)	20	10	21	1	2X decrease

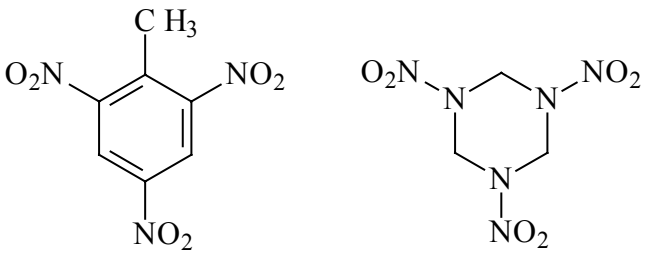
[a]C_t is a function of the starting number of templates present.
[b]Expression ratios relative to control.

Table 3.5. Differential expression of selected genes in *Arabidopsis* roots dependent upon the type of explosive (either TNT or RDX) used for exposure.

	Tag Abundance			Fold Increase for RDX[a]		
tag sequence	Control	TNT	RDX	RDX/TNT	Locus	Annotation
TNT-specific cytochrome P450s						
AGAGAAAGTG	0	14	0	0.07	At3g28740	cytochrome P450, putative; CYP81D11
GCTGAGAGAC	0	5	0	0.20	At4g22690	cytochrome P450 like protein; CYP706A2
GAAAAGATTT	0	5	0	0.20	At2g30750	cytochrome P450; CYP71A12
RDX-specific cytochrome P450s						
TATGCCGCCC	1	0	5	5.00	At1g16400	cytochrome P450-like protein; CYP79F2
cytochrome P450 common to both						
ACCTAACTGA	0	5	5	1.00	At4g13310	cytochrome p450-like protein; CYP71A20
ACACCAAAGC	2	11	3	0.27	At1g78490	cytochrome P450-like protein; CYP708A3
AATAAAACTT	2	9	1	0.11	At2g24180	cytochrome P450, putative; CYP71B6
AAGCATCCGC	0	4	1	0.25	At1g64950	cytochrome P450, hypothetical protein; CYP89A6
ATTGCGCGTG	1	1	5	5.00	At3g20940	cytochrome P450, putative; CYP705A30

[a]For the calculation of expression ratios, a value of 1 was substituted where tag counts were zero.

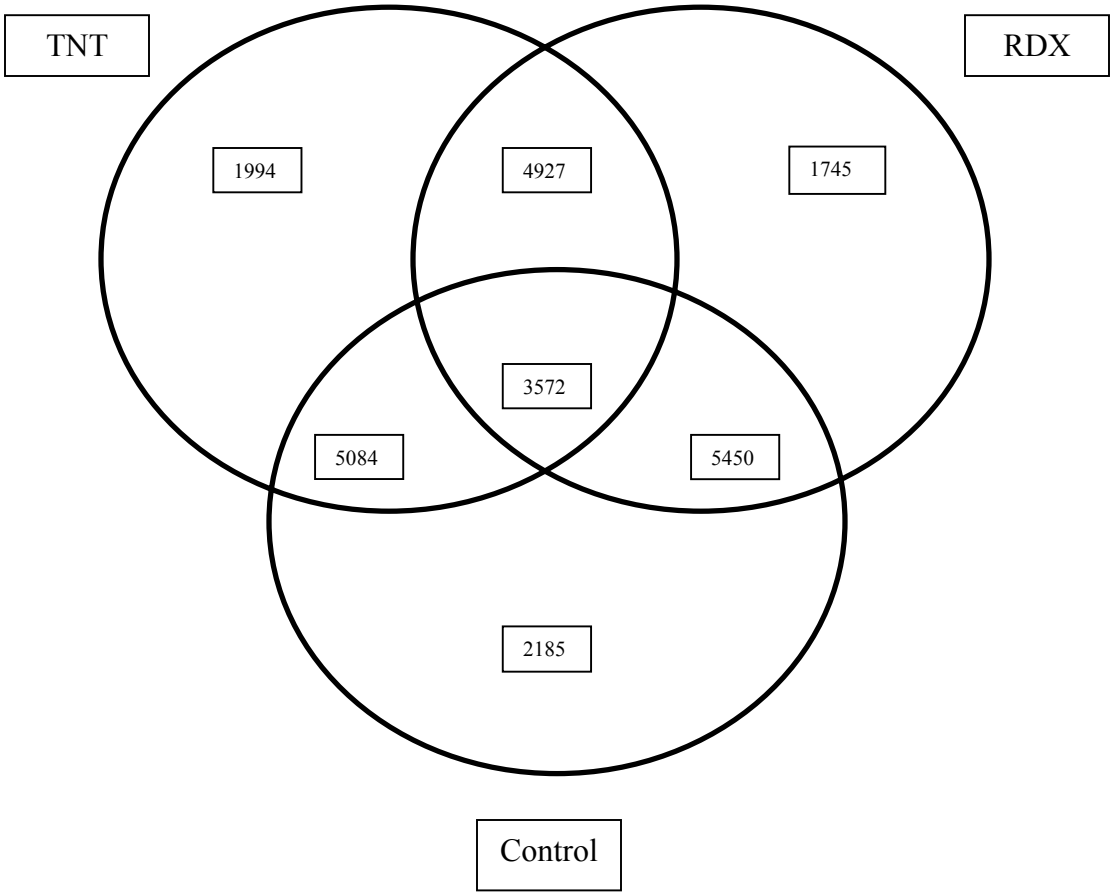
Figure 3.1



TNT (trinitrotoluene)

RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine)

Figure 3.2



REFERENCES

Best, E.H.P., Sprecher, S.L., Fredrickson, H.L., Zappi, M.E., and Larson, S.L. (1997a). Screening submersed plant species for phytoremediation of explosives contaminated groundwater from the Milan Army Ammunition Plant, Milan, Tennessee. Technical Report EL-97-24. U.S. Army Engineer Waterways Experiment Station, Vicksburg, MS.

Best, E.H.P., Miller, J.L., Fredrickson, H.L., Larson, S.L., Zappi, M.E., and Streckfuss, T.H. (1997b). Screening of aquatic and wetland plant species for the phytoremediation of explosives-contaminated groundwater from the Iowa Army Ammunition Plant. Ann. N.Y. Acad. Sci. **829,** 179-194.

Bhadra, R., Williams, R., Barman, S., Stone, M.B., Hughes, J.B., and Shanks, J.V. (2000). Fate of RDX and HMX in axenic plant roots. Chemosphere **44,** 1259-1264.

Burken, J.G., and Schnoor, J.L. (1998). Predictive relationships for the uptake of organic contaminants by hybrid poplar trees. Environ. Sci. Technol. **32,** 3379-3385.

Cao, H., Li, X., and Dong, X. (1998). Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. Proc. Natl. Acad. Sci. USA **95,** 6531-6536.

Cao, H., Bowling, S.A., Gordon, S., and Dong, X. (1994). Characterization of an *Arabidopsis* mutant that is nonresponsive to inducers of systemic acquired resistance. Plant Cell **6,** 1583-1592.

Chern, M.S., Fitzgerald, H.A., Yadav, R.C., Canlas, P.E., Dong, X., and Ronald, P.C. (2001). Evidence for a disease-resistance pathway in rice similar to the NPR1-mediated signaling pathway in *Arabidopsis*. Plant J. **27,** 101-113.

Cholakis, J.M., Wong, L.C.K., Van Goethem, D.L., Minor, J., Short, R., Sprinz, H., and Ellis III, H.V. (1980). Mammalian toxicological evaluation of RDX. Contract No. DAMD17-78-C-8027. U.S. Army Medical Research and Development Command, Fort Detrick, Frederick, MD.

Coleman, N., and Duxbury, T. (1999). Biodegradation of RDX by Rhodococcus sp. strain DN22. Second Int. Symp. Biodegradation of Nitroaromatic Compounds and Explosives. Leesburg, VA, Sept. 8-9, Abstract p. 13.

Crothers, D.M. (1993). Architectural elements in nucleoprotein complexes. Curr. Biol. **3,** 675-676.

Delaney, T.P., Friedrich, L., and Ryals, J.A. (1995). *Arabidopsis* signal transduction mutant defective in chemically and biologically induced disease resistance. Proc. Natl. Acad. Sci. USA **92,** 6602-6606.

Fernando, T., and Aust, S.D. (1991). Biodegradation of munition waste, TNT (2,4,6-trinitrotoluene), and RDX (hexahydro-1,3,5-trinitro-1,3,5-triazine) by *Phanerochaete chrysosporium*. ACS Symposium Series. Ind. Eng. Chem. **486,** 214-232.

Friedrich, L., Lawton, K., Dietrich, R., Willits, M., Cade, R., and Ryals, J. (2001). NIM1 overexpression in *Arabidopsis* potentiates plant disease resistance and results in enhanced effectiveness of fungicides. Mol. Plant-Microbe Interact. **14,** 1114-1124.

Glazebrook, J., Rogers, E.E., and Ausubel, F.M. (1996). Isolation of *Arabidopsis* mutants with enhanced disease susceptibility by direct screening. Genetics **143,** 973-982.

Grasser, K.D. (1998). HMG1 and HU proteins: architectural elements in plant chromatin. Trends Plant Sci. **3,** 260-265.

Grosschedl, R. (1995). Higher-order nucleoprotein complexes in transcription: analogies with site-specific recombination. Curr. Opin. Cell Biol. **7,** 362-370.

Harvey, S.D., Fellows, R.J., Cataldo, D.A., and Bean, R.M. (1991). Fate of the explosive hexahydro-1,3,5-trinitro-1,3,5-triazine (RDX) in soil and bioaccumulation in bush bean hydroponic plants. Environ. Toxicol. Chem. **10,** 845-855.

Holmgren, A. (1989). Thioredoxin and glutaredoxin systems. J. Biol. Chem. **264,** 13963-13966.

Kinoshita, T., Yamada, K., Hiraiwa, N., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (1999). Vacuolar processing enzyme is up-regulated in the lytic vauoles of vegetative tissues during senescence and under various stressed conditions. The Plant J. **19,** 43-53.

Kroczynska, B., Zhou, R.G., Hayman, G.T., and Miernyk, J.A. (1996). AtJ1, a mitochondrial homologue of the *Escherichia coli* DnaJ protein. Plant Mol. Biol. **31,** 619-629.

Kushnir, S., Babiychuk, E., Kampfenkel, K., Bellesboix, E., Vanmontagu, M., and Inze, D. (1995). Characterization of *Arabidopsis thaliana* cDNAs that render

yeasts tolerant toward the thio-oxidizing drug diamide. Proc. Natl. Acad. Sci. USA **92,** 10580-10584.

Merck & Co. (1983). The Merck Index, 10th ed. (Rahway, NJ).

Minakuchi, S.S., Higaki, N., Sato, K., Ohtsuki, K., Kurata, A., Yoshikawa, H., Kotaru, M., Masumura, T., Ichihara, K., and Tanaka, K. (1997). Purification and characterization of glutaredoxin (thioltransferase) from rice (*Oryza sativa* L.). J Biochem (Tokyo) **121,** 842-848.

Ner, S.S., Travers, A.A., and Churchill, M.E.A. (1994). Harnessing the writhe: a role for DNA chaperones in nucleoprotein complex formation. Trends Biochem. Sci. **19,** 185-187.

Riggins, G.J., and Strausberg, R.L. (2001). Genome and genetic resources from the Cancer Genome Anatomy Project. Hum. Mol. Genet. **10,** 663-667.

Sax, N.I., and Lewis, R.J. (1989). Dangerous Properties of Industrial Materials. (New York, New York: VanNostrand Reinhold).

Shah, J., Tsui, F., and Klessig, D.F. (1997). Characterization of a salicylic acid-insensitive mutant (sai1) of *Arabidopsis thaliana* identified in a selective screen utilizing the SA-inducible expression of the tms2 gene. Mol. Plant-Microbe Interact. **10,** 69-78.

Tekoah, Y., and Abeliovich, N.A. (1999). Participation of cytochrome P-450 in the biodegradation of RDX by a Rhodococcus strain. Second Int. Symp. Biodegradation of Nitroaromatic Compounds and Explosives. Leesburg, VA, Sept. 8-9, Abstract p. 7.

Thompson, P.L., Ramer, L., and Schnoor, J.L. (1999). Hexahydro-1,3,5-trinitro-1,3,5-triazine translocation in poplar trees. Environ. Toxicol. Chem. **18,** 279-284.

Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. (1995). Serial analysis of gene expression. Science **270,** 484-487.

Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Hieter, P., Vogelstein, B., and Kinzler, K.W. (1997). Characterization of the yeast transcriptome. Cell **88,** 243-251.

Zhou, R.G., Kroczynska, B., Hayman, G.T., and Miernyk, J.A. (1995). AtJ2, an *Arabidopsis* homolog of *Escherichia coli* DnaJ. Plant Physiol. **108,** 821-822.

CHAPTER 4

MICROARRAY ANALYSIS OF *ARABIDOPSIS* TRANSCRIPTOME CHANGES INDUCED BY TNT

INTRODUCTION

Microarrays

Measurement of changes in gene expression can be an effective method for studying numerous biological phenomena, including disease states, morphogenesis, and toxicological effects (Deyholos and Galbraith, 2001). Although many methods exist to aid researchers in these analyses (see Chapter 1), high-density DNA microarrays provide for the parallel quantification of expression levels for large numbers of genes at a relatively low cost. Analyses on this scale provide for the characterization not only of individual genes, but also the observation of multigenic patterns of expression (Brown and Botstein, 1999), which can aid in the elucidation of gene function.

The two major types of microarrays used for transcript expression profiling are DNA microarrays, which utilize larger DNA elements (400-2000 bp) that are amplified from cloned sequences using PCR, and oligonucleotide microarrays that employ smaller DNA elements (20-90 bp). These may be synthesized prior to spotting onto the glass support or built directly on the support using a photolithographic process (McGall et al., 1996; Lipshutz et al., 1999). Although the longer length of PCR products has potential for greater sensitivity, such microarrays can suffer from nonspecific hybridization due to similarities between gene family members or protein domains. However, specificity can be lost when using oligonucleotides at the low end of the size range, forcing the need to strike a balance between the two extremes (Lashkari et al., 1997). For this reason longer

oligonucleotides (60-90 bp) are gaining in popularity because of their balance between sensitivity and specificity (Figure 4.1).

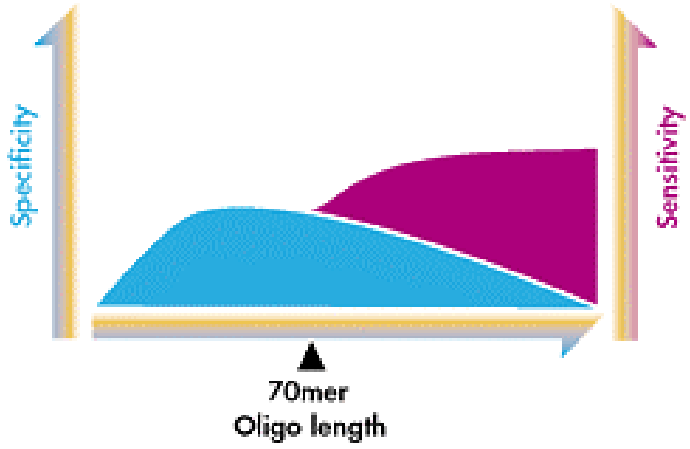


Figure 4.1. The ability of longer oligonucleotide DNA elements (specifically 70-mers) to strike a balance between both specificity and sensitivity as compared to both shorter oligonucleotides and longer oligonucleotides/cDNA fragments. Diagram from the Qiagen website: (http://www.qiagen.com).

Arabidopsis Oligonucleotide Microarrays

To assess changes in the *Arabidopsis* transcriptome induced by energetic compounds, previous studies utilized serial analysis of gene expression (SAGE). However, to fully appreciate plant responses to these compounds, changes must be measured under a variety of conditions and within different plant tissues. Whereas costs are prohibitive to use SAGE for such studies, microarray technology is better suited for this type of analysis. Consequently, the use of high-density oligonucleotide microarrays representing 26,090 genes (well characterized genes and predicted open reading frames) from the Qiagen-Operon Arabidopsis Genome Oligo Set Version 1.0 (Qiagen Operon, Alameda, CA) was initiated. Preliminary analyses using conditions identical to those used for the TNT SAGE studies (15 mg/L TNT) have been conducted. These initial studies have revealed striking similarities in the identification of differential expression by both techniques. However, the limitations characteristic of each technique were also revealed confirming the need to avoid using either one in isolation.

MATERIALS AND METHODS

Plant Material, Growth Conditions, and Root Tissue Isolation

The following experimental descriptions were developed to conform as closely as possible to the MIAME guidelines for describing microarray experiments (Brazma et al. 2001). *A. thaliana* ecotype Columbia seeds (WT-2, Lehle Seeds, Round Rock, TX) were

surface-sterilized and placed in sterile Murashige and Skoog liquid medium prepared according to the manufacturer's protocol (Invitrogen Life Technologies, Carlsbad, CA). Plants were grown for 14 days at 25^0C under a 16-hr photoperiod with constant shaking at 85 rpm in a growth chamber. TNT was obtained from the U.S. Army Center for Environmental Health Research (Fort Detrick, MD). Root tissues for microarray analyses were isolated from seedlings grown 14 days in liquid MS medium and dosed with TNT dissolved in DMSO to a final concentration of 15 mg/L. Control tissues were isolated from seedlings grown under the same conditions and dosed with an equivalent volume of DMSO. The seedlings were grown in the presence of TNT or DMSO for 24 hr, after which they were submerged briefly in dH_2O to remove excess medium, and excised roots were immediately frozen in liquid nitrogen. All root tissues were stored at -80^0C prior to RNA extraction. The root tissues for each condition (i.e., TNT-dosed or control) were pooled from approximately 1400 seedlings recovered from 20 treatment flasks grown in parallel.

RNA Isolation

Total RNA from root tissues was extracted using the LiCl precipitation technique of Chang et al. (1993). Poly(A) RNA was isolated from total RNA using Dynabeads oligo-dT(25) magnetic beads (Dynal Biotech, Lake Success, NY) at a ratio of 0.25 mg of total RNA per 250 μL of Dynabeads and following the manufacture's instructions.

Labeling cDNA for Oligonucleotide Microarray Hybridization

The following protocol was adapted from a Galbraith lab protocol (http://ag.arizona.edu/microarray/protocol1.doc). Messenger RNA isolated from each condition under study was labeled for microarray analysis as follows: 2 μg of poly A+ mRNA (in a volume of 50 μL) was combined with a mixture of dNTPs (3.0 μL of a mixture of 10 mM dATP, dCTP, dGTP (350 μM final), 2 mM dTTP (70 μM final)) (Sigma-Aldrich, St. Louis, MO), either Cy5-dUTP or Cy3-dUTP (2.0 μL) (24 μM final) (Amersham Biosciences, Piscataway, NJ), and random 15-mer primers (2.0 μL of 0.5 μg/μL solution for a 10 ng/μl final concentration) (Qiagen Inc., Valencia, CA) in a total volume of 57 μL. The solution was heated for five min at 65^0C, after which 5X first strand buffer (17 μL) (BD Biosciences, Clontech, Palo Alto, CA), 0.1 M DTT (8.0 μL), RNAse inhibitor (1.0 μL) (Invitrogen Life Technologies, Carlsbad, CA), and reverse transcriptase (1.0 μL) (BD Biosciences Clontech, Palo Alto, CA) were added to yield a total volume of 84 μL. The solution was then heated for an additional two hr at 42^0C, after which the reaction was terminated by the addition of EDTA (5 μL of 0.5 M) and the RNA was removed using NaOH (5 μL of 1 M). The solution was then heated to 65^0C for a period of 10 min, followed by the addition of Tris-HCl (pH 8.0) (25 μL of 1 M) and TE (100 μL), prior to storage in the dark at -20^0C until use. Prior to hybridization, all labeled targets were purified using Microcon YM-30 nitrocellulose spin columns (Millipore Corp., Billerica, MA).

Preparation of the Oligonucleotide Microarrays

The following protocol was adapted from a Galbraith lab protocol (http://ag.arizona.edu/microarray/protocol1.doc). Printed in the Galbraith laboratory at The University of Arizona, the high-density arrays used in these experiments consist of 70-mer oligonucleotides printed on aminosilane-coated slides (Telechem) arrayed in a format to accommodate the presence of 27,216 elements.

Spotted oligonucleotides are not immobilized on the glass slides prior to distribution, and this must be done prior to their use in experiments. To accomplish this, the microarrays were held over a 50^0C water bath for 10 seconds to rehydrate, followed by snap-drying on a 65^0C heating block for 5 seconds. For each array this process was repeated four times. Following rehydration, the arrays were crosslinked using a Stratalinker UV Crosslinker (Stratagene, La Jolla, CA) set at 130 mJ. After crosslinking the slides were washed in a solution of 1% SDS for five minutes at room temperature with gentle shaking. SDS was removed from the array by gentle shaking in 100% ethanol for 30 seconds. The arrays were then dried by centrifugation at 1000 rpm for two minutes. Slides were stored desiccated in the dark until use.

Hybridization of Microarrays

The following protocol was adapted from a Galbraith lab protocol (http://ag.arizona.edu/microarray/protocol1.doc). Prior to hybridization, 20X SSC (12.0 μL), Liquid Block (7.2 μL) (Amersham Biosciences, Piscataway, NJ), 2% SDS (4.8 μL),

and dH_20 to a final volume of 120 μL was added to the purified sample. This solution was heated over boiling water for 2 min and rapidly transferred to ice. Hybridizations were carried out using a Genomic Solutions Automated Hybridization Chamber programmed as follows: O-ring stuff (75^0C for 2 min), probe introduction (65^0C), hybridization (55^0C for 8 hr with agitation), 2X SSC, 0.5% SDS wash (55^0C; flow for 16 sec; hold for 5 min), 0.5X SSC wash (25^0C; flow for 16 sec; hold for 5 min), 0.05X SSC wash (25^0C; flow for 16 sec; hold for 5 min). Following hybridization, slides were rinsed with gentle shaking in dH_2O for 10 sec and spun (1,000 rpm for 5 min) or blown dry (filtered air).

Scanning of Microarrays

Microarrays were scanned using a Scanarray 5000 spectrometer (GSI Lumonics, Billerica, MA) with appropriate laser and filter settings. Data was extracted using Quantarray software (Applied Biosystems, Foster City, CA) selecting background subtraction and quantification via the fixed circle method. Data were normalized by measuring the total intensity of each array on the assumption that the majority of spots represented genes whose expression was unaltered by treatment. Spots whose intensities were altered by local effects (e.g., dust, residual wash buffer, etc.) were flagged and removed from analyses.

RESULTS AND DISCUSSION

Differentially expressed transcripts

Microarray analysis identified numerous *Arabidopsis* transcripts that were differentially expressed in response to TNT (Table 4.1). Striking were the increased levels of several glutathione transferase and glutathione peroxidase transcripts (Table 4.1). Transcripts for glucosyltransferase enzymes were also highly induced, suggesting an up-regulation of conjugating enzymes by TNT. Other highly induced transcripts included those for transcription factors, oxidative stress response enzymes, quinone oxidoreductases, and cytochrome P450s. The most highly induced transcript was that for nitrilase 2, an enzyme thought to be responsible for the conversion of indoleacetonitrile to the plant growth factor indole-3-acetic acid (auxin) (Normanly et al., 1997).

Several transcripts were also observed to decrease upon TNT treatment (Table 4.2). A transcript for a pEARLI-1-like protein showed the greatest repression in response to TNT. Repression of this transcript along with those for a cell wall plasma membrane-disconnecting CLCT protein (AIR1A), extensins, and an arabinogalactan protein suggests that TNT has a strong impact on plant cell wall formation. Repression was also observed for transcripts related to vacuolar functions, such as gamma tonoplast intrinsic proteins and major latex proteins.

Comparison of Data Derived From Microarrays and SAGE

Although the reliability of data from microarray analyses hinges on the use of replicates for statistical power, it is interesting to compare these preliminary microarray data to the results obtained using SAGE to analyze *Arabidopsis* seedling roots treated in an identical fashion (exposure to 15mg/L TNT for 24 hr). As the fundamental processes underlying these techniques differ completely (i.e., DNA sequencing vs. differential hybridization), the identification of the same differentially expressed transcripts with both techniques strongly suggests that these gene products have biological relevance for the plant's response to TNT.

In general, genes that showed strong induction by TNT as measured using SAGE, were strongly induced in the microarray analyses (Table 4.3). For example, a glutathione-S-transferase (GST) transcript (At1g17170) was the tag most highly induced in the SAGE studies. Transcripts for this enzyme were the third most highly induced in the microarray analyses, and transcripts for other GSTs comprised six of the top 20 most highly induced by TNT. Similarly, the most highly repressed transcripts were also comparable for the two techniques (Table 4.4).

This effect (general agreement between the two techniques for the most highly differentially expressed transcripts) was also observed in another study in which the SAGE technique was compared to GeneChip technology (Affymetrix, Inc., CA) for measuring changes in gene expression in human blood monocytes and granulocyte-macrophage colony-stimulating factor induced macrophages (Ishii et al., 2000). Furthermore, it is known that the validity of microarray data increases with the number of

replicates used (Lee et al., 2000). Thus, agreement between the techniques is likely to increase as the experiment is repeated.

While many of the other high ranking differentially expressed transcripts were not identical matches between SAGE and microarrays, many were similar in function. This may result from the use of a different pool of plants for the two studies. However, it may also be that cross-hybridization occurred in regions where family members possess sequence similarities. This may have been the cause of the apparent disagreement between the techniques regarding the most highly induced GST. While SAGE indicated the GST encoded by gene At1g17170 as being the most highly induced, the microarray results pointed to gene At1g17180, a family member located immediately downstream from At1g17170 (Table 4.3).

General Comparisons of SAGE and Microarrays

This work also allowed for head-to-head comparison of two fundamentally different transcriptional profiling methods: SAGE and high-density microarrays. While both of these methods provide broad analyses of the transcriptome, each has distinct strengths and weaknesses. In terms of time, microarray analysis is far superior. Hybridization, scanning, and data extraction can all be achieved within days. Since hybridizations can be run in parallel, the experimental replicates required for statistically significant analyses can be performed simultaneously (Lee et al., 2000). In contrast, the construction of SAGE libraries typically consumes a period of two weeks for each library, and sequencing can consume at least another two weeks depending on the

required depth of analysis and the availability of high-throughput sequencing equipment. This time factor is a major consideration in the design of experiments for measurement of transcript levels under several different conditions. Furthermore, these time estimates only include the work required to obtain measures of transcript levels and do not take into account the time required for bioinformatic analyses and data mining. In terms of cost, microarrays also have the advantage, provided the required equipment (e.g., hybridization chambers, scanner, spotting robot, etc.) is available. These cost differentials hinge predominantly on the lack of need for extensive DNA sequencing and the costly reagents it consumes. Finally, the requirement for polyadenylated transcripts by SAGE precludes its use for analysis of prokaryotes, while microarrays are not limited by this requirement. Despite significant advantages, unlike SAGE, microarrays cannot aid in gene discovery. Moreover, as SAGE does not require extensive genome analysis, changes in transcriptome metabolism such as alternative splicing of nascent messenger RNA molecules, can be observed allowing for greater depth of understanding of transcriptome responses. Thus, care must be taken not to assume that even high-density microarrays can provide a complete picture of the transcriptome.

Table 4.1. Top 20 transcripts induced by TNT exposure

Control	TNT	TNT/Control	SAGE rank[a]	Gene I.D.	Annotation
1379	12631	9.16	97	At3g44300	nitrilase 2
572	5212	9.11	NP[b]	At1g54090	hypothetical protein
136572	1105662	8.1	NP	At1g17180	glutathione transferase,
133914	894671	6.68	487	At1g02920	glutathione transferase, putative
575	3818	6.63	20	At1g76680	12-oxophytodienoate reductase (OPR1)
60806	389368	6.4	16	At3g21720	putative isocitrate lyase
784	4510	5.75	NP	At4g34138	similar to glucosyltransferase -like protein
1135	5744	5.06	NP	At1g72900	disease resistance protein (TIR-NBS class), putative
900	4149	4.61	NP	At1g26970	protein kinase, putative
146471	665160	4.54	NP	At3g15450	expressed protein
879	3958	4.5	110	At1g02850	glycosyl hydrolase family 1
229008	1004476	4.39	NP	At2g29420	glutathione transferase
53984	222163	4.12	NP	At1g21000	expressed protein
156249	614698	3.93	6	At5g13750	transporter-like protein
208573	814582	3.91	NP	At3g50970	dehydrin Xero2
221705	843792	3.81	247	At4g11600	glutathione peroxidase
323149	1189150	3.68	489	At1g78380	glutathione transferase, putative;
49920	183112	3.67	17	At2g29490	glutathione transferase,
363743	1314001	3.61	1014	At1g02930	glutathione transferase, putative
783	2808	3.58	NP	At1g53850	20S proteasome alpha subunit E1 (PAE1)

[a] Ranked according to degree of induction (i.e., descending from most highly induced)
[b]NP=not present in SAGE libraries.

Table 4.2. Top 20 Transcripts Repressed by TNT Exposure

Control	TNT	Control/TNT	SAGE rank[a]	Gene I.D.	Annotation
580557	66334	8.75	16	At4g12520	pEARLI 1-like protein
345396	46140	7.49	NP[b]	At5g58850	myb family transcription factor; protein id: At5g58850.1
23128	3952	5.85		**********	b/t hypothetical protein; protein id: At4g12540.1 and putative cell wall-plasma membrane disconnecting CLCT protein (AIR1A) ; protein id: At4g12550.1
1022035	196517	5.20	NP	At2g23830	unknown protein
20234	4309	4.70	NP	At3g20950	cytochrome P450, putative
38607	7978	4.84	NP	At5g47990	cytochrome p450 family
797955	172184	4.63	7	At2g36830	putative aquaporin (tonoplast intrinsic protein gamma)
4480641	1032280	4.34	NP	At5g57660	putative CONSTANS B-box zinc finger protein
3379	804	4.20	NP	At3g02620.1	putative stearoyl-acyl carrier protein desaturase
17168	4372	3.93	NP	At3g29360	UDP-glucose dehydrogenase, putative
127380	31517	4.04	NP	At4g02270	hypothetical protein similar to extensin-like protein
1628890	431025	3.78	NP	At5g10430	arabinogalactan-protein (AGP4)
952744	251759	3.78	NP	At2g29750	UDP-glycosyltransferase family
5390606	1407372	3.83	6264	At5g44550	putative protein
1220910	314425	3.88	NP	At5g21950	similar to putative hydrolase
3137398	805884	3.89	1170	At1g14450	expressed protein
1981036	553714	3.58	1121	At3g26520	gamma tonoplast intrinsic protein
65535	19254	3.40	NP	At5g06640	putative protein
973194	280122	3.47	2	At4g12550	putative cell wall-plasma membrane disconnecting CLCT protein (AIR1A)
6476741	1973671	3.28	4025	At1g26250	unknown protein

[a] Ranked according to degree of repression (i.e.,x descending from most highly repressed)
[b]NP=not present in SAGE libraries.

Table 4.3. Comparison of the 20 Most Highly Induced Transcripts as Measured Using SAGE and Microarrays*.

Microarray				SAGE		
Fold	**Gene I.D.**	**Annotation**		**Fold**	**Gene I.D.**	**Annotation**
9.16	At3g44300	nitrilase 2		27.5	At1g17170	**putative glutathione transferase**
9.11	At1g54090	hypothetical protein		24.0	At5g11510	MYB like protein
8.10	**At1g17180**	**glutathione transferase*,**		23.0	At3g27740	carbamoyl phosphate synthetase small subunit (AGI); pyrimidine nucleotide biosynthesis--NOT IN 3' MOST EXON
6.68	At1g02920	glutathione transferase, putative		21.0	**At5g13750**	**transporter-like protein**
6.63	**At1g76680**	**12-oxophytodienoate reductase (OPR1)**		20.0	At1g60940	putative serine/threonine-protein kinase
6.40	**At3g21720**	**putative isocitrate lyase**		17.5	At5g03630	monodehydroascorbate reductase (NADH) - like protein
5.75	At4g34138	similar to glucosyltransferase -like protein		17.0	**2 hits**	At4g01870; protein of unknown function
5.06	At1g72900	disease resistance protein (TIR-NBS class), putative		17.0	At5g63790	putative protein; NAM, no apical meristem, - like protein A.thaliana
4.61	At1g26970	protein kinase, putative		15.0	At1g05050	putative transcription factor
4.54	At3g15450	expressed protein		14.0	At3g28740	cytochrome P450, putative; CYP81D11-A-TYPE
4.50	At1g02850	glycosyl hydrolase family 1		13.0	**polyA??**	does not appear in the genome (the truncate w/o A's does not match any putattive genes either)
4.39	At2g29420	glutathione transferase		12.7	At5g45020	putative protein--glutathione transferase
4.12	At1g21000	expressed protein		12.0		one hit on chromosome 3
3.93	**At5g13750**	**transporter-like protein**		11.5	**At3g21720**	**putative isocitrate lyase; C-compound and carbohydrate utilization**
3.91	At3g50970	dehydrin Xero2		11.0	**At2g29490**	**putative glutathione S-transferase**
3.81	At4g11600	glutathione peroxidase		10.0	At1g78080	putative AP2 domain containing protein RAP2--transcription factor TINY A.thaliana
3.68	At1g78380	glutathione transferase, putative;		10.0	At2g28570	unknown protein; Protein-tyrosine kinase
3.67	**At2g29490**	**glutathione transferase,**		10.0	**At1g76680**	**12-oxophytodienoate reductase (OPR1)**
3.61	At1g02930	glutathione transferase, putative		10.0	At1g75270	GSH-dependent dehydroascorbate reductase 1, putative
3.58	At1g53850	20S proteasome alpha subunit E1 (PAE1)		10.0	At1g05680	putative indole-3-acetate beta-glucosyltransferase

*Genes observed in both the microarray analysis and SAGE above are indicated in bold.

Table 4.4. Comparison of the 20 Most Highly Repressed Transcripts as Measured Using SAGE and Microarrays*.

Microarray				SAGE		
Fold	**Gene I.D.**	**Annotation**		**Fold**	**Gene I.D.**	**Annotation**
8.75	**At4g12520**	**pEARLI 1-like protein**		85	**At4g12550**	**putative cell wall-plasma membrane disconnecting CLCT protein (AIR1A)**
7.49	At5g58850	**myb family transcription factor**		52	At2g01520	major latex protein homolog-like; A.thaliana
5.85	xxxx	**putative cell wall-plasma membrane disconnecting CLCT protein (AIR1A)**		45	At1g70850	major latex protein homolog-like: A. thaliana
5.20	At2g23830	unknown protein		37	AF098631	**putative cell wall-plasma membrane disconnecting CLCT protein (AIR1B)**
4.84	At5g47990	cytochrome p450 family		29	At1g09690	putative 60S ribosomal protein L21
4.70	At3g20950	cytochrome P450, putative		29	**At2g36830**	**putative aquaporin (tonoplast intrinsic protein gamma)**
4.63	**At2g36830**	**putative aquaporin (tonoplast intrinsic protein gamma)**		50	At4g22212	Expressed protein
4.34	At5g57660	putative CONSTANS B-box zinc finger protein		24	At5g23710	unknown protein
4.20	At3g02620.1	putative stearoyl-acyl carrier protein desaturase		23	At3g54580	**extensin precursor -like protein**
4.04	At4g02270	**hypothetical protein similar to extensin-like protein**		21	At5g19510	elongation factor 1B alpha-subunit
3.93	At3g29360	UDP-glucose dehydrogenase, putative		20	At5g56540	**arabinogalactan-protein AGP14**
3.89	At1g14450	expressed protein		20	At1g14960	major latex protein, putative; stress response
3.88	At5g21950	similar to putative hydrolase		20	At4g38740	peptidylprolyl isomerase ROC1; protein folding and stabilization
3.83	At5g44550	putative protein		16	At4g17340	membrane channel like protein; other transport facilitators
3.78	At2g29750	UDP-glycosyltransferase family		15	**At4g12520**	**pEARLI 1-like protein**
3.78	At5g10430	**arabinogalactan-protein (AGP4)**		15	At5g10280	**putative transcription factor MYB92**
3.58	At3g26520	**gamma tonoplast intrinsic protein**		56	At5g26260	putative protein;UNCLASSIFIED PROTEINS
3.47	**At4g12550**	**putative cell wall-plasma membrane disconnecting CLCT protein (AIR1A)**		42	At5g14200	3-isopropylmalate dehydrogenase; C-compound and carbohydrate metabolism
3.40	At5g06640	putative protein		14	At4g14960	tubulin alpha-6 chain (TUA6); cytoskeleton
3.38	At3g02610.1	putative stearoyl-acyl carrier protein desaturase		14	At1g04820	tubulin alpha-2/alpha-4 chain; cytoskeleton

*Genes observed in both the microarray analysis and SAGE above are indicated in bold.

REFERENCES

Brown, P.O., and Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. Nat. Genet. **21,** 33-37.

Deyholos, M.K., and Galbraith, D.W. (2001). High-density DNA microarrays for gene expression analysis. Cytometry **43,** 229-238.

Ishii, M., Hashimoto, S., Tsutsumi, S., Wada, Y., Matsushima, K., Kodama, T., and Aburatani, H. (2000). Direct comparison of GeneChip and SAGE on the quantitative accuracy in transcript profiling analysis. Genomics **68,** 136-143.

Lashkari, D.A., DeRisi, J.L., McCusker, J.H., Namath, A.F., Gentile, C., Hwang, S.Y., Brown, P.O., and Davis, R.W. (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc. Natl. Acad. Sci. USA **94,** 13057-13062.

Lee, M.L.T., Kuo, F.C., Whitmore, G.A., and Sklar, J. (2000). Importance of replication in microarray gene expression studies: statistical methods and evidence form repetitive cDNA hybridizations. Proc. Natl. Acad. Sci. USA **97,** 9834-9839.

Lipshutz, R., Fodor, S., Gingeras, T., and Lockhart, D. (1999). High density synthetic oligonucleotide arrays. Nat. Genet. **21,** 20-24.

McGall, G., Labadie, J., Brock, P., Wallraff, G., and Nguyen, T. (1996). Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. Proc. Natl. Acad. Sci. USA **93,** 13555-13560.

Nadon, R., and Shoemaker, J. (2002). Statistical issues with microarrays: processing and analysis. Trends in Genetics **18,** 265-271.

Normanly, J., Grisafi, P., Fink, G.R., and Bartel, B. (1997). *Arabidopsis* mutants resistant to the auxin effects of indole-3-acetonitrile are defective in the nitrilase encoded by the NIT1 gene. Plant Cell **9,** 1781-1790.

CHAPTER 5

CONCLUSIONS

The viability of phytoremediation as an approach to reducing environmental contamination is still somewhat in question. While there is a dire need of affordable methods of remediation--the world market for remediation was estimated at $15-$18 billion and growing in 1998 (Glass, 1999)--several hurdles prevent phytoremediation from supplanting classical methods of clean up. One major obstacle is the lengthy time that is often required for complete phytoremediation. Site remediation using plants can take years, during which time sites must be monitored with respect to contaminant levels and the trees or plants in use may require care and/or harvesting (Watanabe, 2001). Furthermore, sites containing high levels of contaminants may inhibit growth and prevent successful introduction of plants. There are also concerns over the potential for introduction of contaminants into the food chain. The question of how to dispose of plants that extract heavy metals or immutable organics is also a consideration.

For these reasons, researchers have endeavored to engineer wild-type plants with genes that can confer greater tolerance and/or degradation abilities (Rugh et al., 1996; Rugh et al., 1998; Bizily et al., 1999; French et al., 1999; Hannink et al., 2001). Although effective under controlled conditions in the laboratory, engineered plants remain unproven in the field, and some of the species used are unsuitable for field

applications. Furthermore, although less burdened by regulatory requirements than microbial-based remediation processes, phytoremediation using transgenic or genetically-modified (GM) plants faces significant public and government opposition. Major concerns over field release of such plants include increased invasiveness and decreased genetic variability of native plants due to interbreeding.

Despite these misgivings, researchers continue to pursue the development of genetically-engineered plants for phytoremediation. Intense interest by federal agencies and commercial institutions that control numerous sites in need of remediation (there are about 12,000 contaminated "Superfund" sites listed in the U.S., and 400,000 sites in Western Europe (Glass, 1999)) is keeping the technology at the forefront, and many argue that solutions to public concerns exist. For example, the use of sterile clones has been suggested as a solution to the problem of invasiveness and interbreeding.

As researchers become more adept at introducing genes into a wider variety of plant species, additional properties desirable for remediation processes such as rapid growth, deep root structures, and high water uptake, will be available in modified plants (Dietz and Schnoor, 2001). Yet, there remains a general lack of knowledge with regard to the metabolic mechanisms used by plants to cope with toxins, which prevents focused engineering attempts. It has been argued that "increased understanding of the enzymatic processes involved in plant tolerance and metabolism of xenobiotic chemicals will provide new potential for engineering plants with increased phytoremediation capabilities" (Dietz and Schnoor, 2001), but the lack of enzymological information for many contaminants prevents informed decisions on which genes to engineer. This was

certainly the case for explosives remediation, and it was for this reason that the project described in this dissertation was undertaken.

Using genomic techniques to measure changes in gene expression in *Arabidopsis thaliana* upon exposure to toxic concentrations of TNT and RDX has proven useful for identifying some of the metabolic mechanisms plants may use for detoxifying explosives. For example, it appears as though conjugation via glutathione is the major means by which *Arabidopsis* deals with the toxic effects of TNT. In addition, the induction of several cytochrome P450 transcripts reveals the potential for transformation prior to conjugation. The five-fold increase of a transcript representing an ABC transporter protein hints at a possible mechanism for conjugate sequestration in keeping with that observed for herbicide metabolism (Martinoia et al., 1993). Using SAGE, specific gene family members involved in TNT and RDX metabolism were identified. This information will be of great utility in the design of transgenic plants for phytoremediation. While TNT appears to stimulate the function of a multi-phase mechanism for detoxification in *Arabidopsis*, RDX appears to rely on a completely different detoxification mechanism. As noted above, several of the most highly induced genes in the TNT SAGE library were not even expressed in the presence of RDX. Moreover, while an aquaporin transcript was highly repressed in the TNT library, another transcript representing aquaporin was highly induced in the RDX library. This may indicate different mechanisms for vacuolar metabolism or different strategies for affecting munitions uptake. These differences will require the engineering of multiple enhanced pathways for the metabolism of both compounds.

There are also limits to the application of these data. As it is transcriptionally based, it is also limited in scope. It is well known that transcript levels are not always related directly to downstream protein levels or enzyme activities. Therefore, measures of protein expression and activity for genes that appear to be differentially expressed should be pursued before substantial efforts are expended on creation of transgenic plants. Knockouts of genes suspected to be involved in the metabolism of TNT and/or RDX, created using RNA interference (RNAi) or antisense technologies, will aid in confirming these findings. Studies conducted under field conditions (e.g., growth in contaminated soil with concentrations of explosives more similar to those at contaminated sites) are also necessary to better assess the critical physiological changes induced by exposure to explosives. Finally, transgenic plants with increased or decreased expression of relevant genes will be necessary to fully test hypotheses arising from the research presented in this dissertation.

Further studies using transcriptome profiling to fully elucidate plant responses to explosives will also be necessary. One way to accelerate this process is to combine the strengths of both techniques. When studying specific processes or responses in which the genes involved have already been discovered using SAGE or other open-ended techniques, microarrays provide a powerful method for rapid analysis of cellular responses. This approach has been used effectively in the study of breast cancer to elucidate major changes in metabolic pathways during its progression (Nacht et al., 1999). Using SAGE, researchers identified genes that were differentially expressed and then used these candidates to fabricate custom microarrays for screening clinical breast

tumor samples to find genes that are consistently expressed at different levels in diseased and normal tissues.

Custom microarrays have also been used for rapid assessment of drug or pollutant mechanisms of toxicity (Nuwaysir et al., 1999). This approach can be an extremely efficient means of screening drug candidates for toxicity prior to conducting more costly and time-consuming studies. In environmental research, identification of genes responding to toxicant exposure can be useful in risk assessment. If a compound whose mechanism of action is unknown induces expression of genes previously shown to be induced by exposure to a class of toxicant whose mechanism of action has already been elucidated (e.g., endocrine disruptors), the unknown compound's toxicity may be rapidly classified and its threat to the environment more accurately assessed (Nuwaysir, 1999) (Figure 5.1). While this method has been employed in human toxicity studies (e.g., ToxChip v1.0 developed by Nuwaysir and colleagues), it has yet to be employed in the study of plant responses (Nuwaysir et al., 1999). It will be useful for future toxicological studies to develop a similar microarray or "ToxChip" for *Arabidopsis*.

While microbial bioremediation has already achieved moderate success, phytoremediation is still mostly in the proving phase. Although it has shown great promise in the laboratory, there are relatively few examples of its successful transition into the field. However, as this new technology develops, the limitations responsible for the delay in its successful application will be overcome. The ecological benefits afforded by phytoremediation provide the impetus for pursuing its widespread implementation. The use of "natural" solutions, such as phytoremediation, encourages the reestablishment of biological systems on contaminated sites, while current methods

(excavation/incineration) do just the opposite. Thus, public acceptance of this technology as an alternative to more destructive processes will help to bring it to the forefront of remediation strategies.

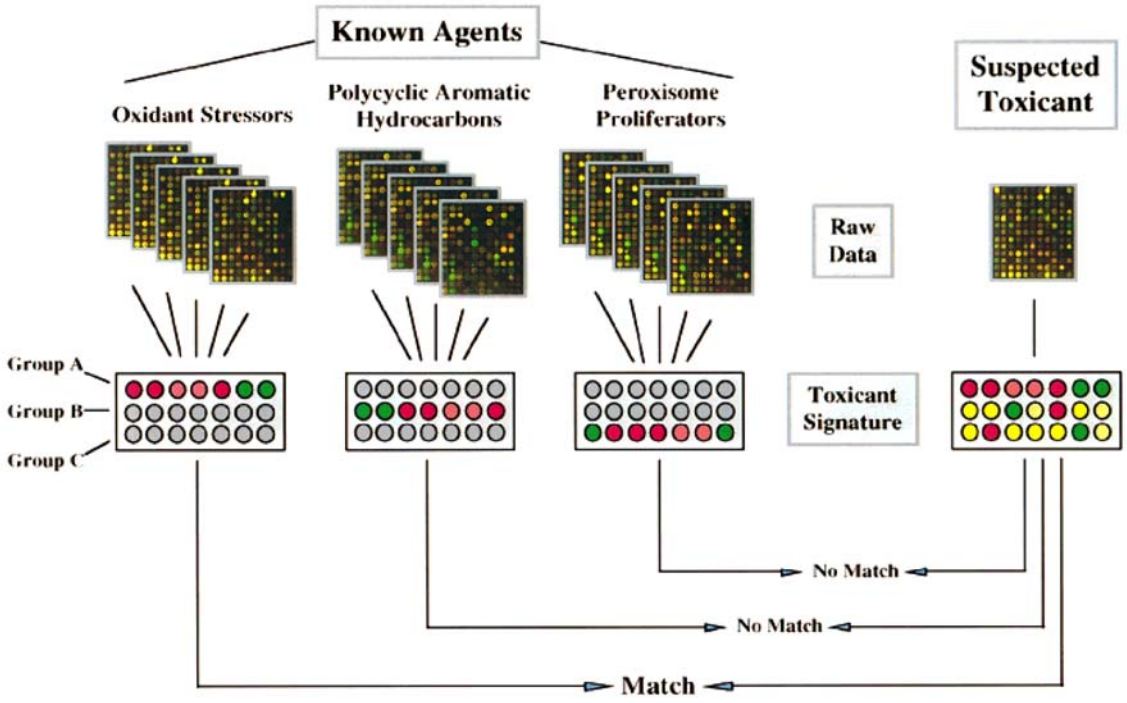


Figure 5.1. Schematic representation of a microarray-based method for identifying a toxicant mechanism of action. In this method, gene-expression data derived from exposure of a model system to known toxicants is compared to changes induced by novel compounds (termed the toxicant signature) (Nuwaysir, 1999).

REFERENCES

Bizily, S., Rugh, C.L., Summer, A.O., and Meagher, R.B. (1999). Phytoremediation of methylmercury pollution: *mer*B expression in *Arabidopsis thaliana* confers resistance to organomercurials. Proc. Natl. Acad. Sci. USA **96,** 6808-6813.

Dietz, A.C., and Schnoor, J.L. (2001). Advances in phytoremediation. Environ. Health Persp. **109,** 163-168.

French, C.E., Rosser, S.J., Davies, G.J., Nicklin, S., and Bruce, N.C. (1999). Biodegradation of explosives by transgenic plants expression pentaerythritol tetranitrate reductase. Nat. Biotechnol. **17,** 491-494.

Glass, D. (1999). U.S. and international markets for phytoremediation. (D. Glass Associates, Inc.; 1999). http://www.channel1.com/users/dglass/INFO/phy99exc.htm.

Hannink, N., Rosser, S.J., French, C.E., Basran, A., Murray, J.A.H., Nicklin, S., and Bruce, N.C. (2001). Phytodetoxification of TNT by transgenic plants expressing a bacterial nitroreductase. Nat. Biotechnol. **19,** 1168-1172.

Martinoia, E., Grill, E., Tommasini, R., Kreuz, K., and Amrhein, M. (1993). An ATP-dependent glutathione-S-conjugate 'export' pump in the vacuolar membrane of plants. Nature **364,** 247-249.

Nacht, M., Ferguson, A.T., Zhang, W., Petroziello, J.M., Cook, B.P., Gao, Y.H., Maguire, S., Riley, D., Coppola, G., Landes, G.M., Madden, S.L., and Sukumar, S. (1999). Combining serial analysis of gene expression and array

technologies to identify genes differentially expressed in breast cancer. Cancer Res. **59,** 5464-5470.

Nuwaysir, E.F., Bittner, M., Trent, J., Barrett, J.C., and Afshari, C.A. (1999). Microarrays and toxicology: the advent of toxicogenomics. Mol. Carcin. **24,** 153-159.

Rugh, C.L., Senecoff, J.F., Meagher, R.B., and Merkle, S.A. (1998). Development of transgenic yellow poplars for mercury phytoremediation. Nat. Biotechnol. **16,** 925-928.

Rugh, C.L., Wilde, D., Stack, N.M., Thompson, D.M., Summer, A.O., and Meagher, R.B. (1996). Mercuric ion reduction and resistance in transgenic *Arabidopsis thaliana* plants expressing a modified bacterial *mer*A gene. Proc. Natl. Acad. Sci. USA **93,** 3182-3187.

Watanabe, M.E. (2001). Can bioremediation bounce back? Nat. Biotechnol. **19,** 1111-1115.