Transcriptomics and Proteomics: Tools for Optimising Phytoremediation Activities

Julian O. D. Coleman[a,*], Richard P. Haslam[b], and Andrew L. Downie[b,c]

[a] School of Biological and Molecular Sciences, Oxford Brookes University, Oxford, UK.
E-mail: jcoleman@brookes.ac.uk
[b] Crop Performance and Improvement, Rothamsted Research, Harpenden, UK
[c] Current address: AstraZeneca, Alderley Park, Cheshire, UK

* Author for correspondence and reprint requests

Z. Naturforsch. **60 c**, 544–548 (2005)

Negligent industrial development has greatly contributed to environmental pollution through the contamination of water and soil by xenobiotic organic chemicals. Remedial strategies to deal with chemical pollution require reliable methods to identify and monitor contamination, as well as effective procedures to attenuate or to eliminate the pollutant. In the food chain, plants are ideally placed as early bio-indicators of environmental pollution as they experience and respond to environmental toxicants sooner than organisms at higher trophic levels. Furthermore, some plants are capable of detoxifying anthropogenic chemicals by metabolic transformation and could prove useful for the remediation of contaminated water and soil: so-called phytoremediation.

So far research technologies aimed at developing plants for bio-indication/bio-monitoring and for remediation have largely relied on standardised chemical and biochemical procedures to evaluate phytotoxicity, metabolic fate and persistence of organic pollutants in plants. The next stage in the evolution of these plant-based technologies is the improvement and optimisation of any innate phytoremediation activities identified in selected plants.

In general, uptake followed by metabolism and compartmentation is responsible for the detoxification of organic xenobiotics in plants. These are complex cellular systems that may be organised in well-defined pathways and are often controlled by large families of genes. In order to elucidate complex traits such as detoxification, an emerging idea is to make use of global approaches such as the new "omic" technologies to identify molecular changes in plant tissues exposed to specific organic xenobiotics. From expression profiles at the messenger RNA level, transcriptomics permit the identification of function-related gene clusters and at the protein level proteomics provide information on where, when and at what level specific proteins accumulate.

We conclude that these global approaches may be a useful way of widening screening capability to identify appropriate molecular markers that can be used to improve detoxification activity.

Key words: Phytoremediation, Transcriptomics, Proteomics

Introduction

For more than a century the manufacture and use of synthetic organic chemicals have made significant contributions to improvements in living standards and human health. The chemical industry is an important part of the global economy and both demand and output are predicted to markedly increase over the next decade. The number and range of chemicals is immense, with over 300,000 chemicals in common commercial use and several thousands are added annually. However, one of the inevitable repercussions of the manufacture of synthetic chemicals is their release into the environment either as a direct consequence of manufacture and use or as a result of incautious disposal. As a significant and increasing proportion of the synthetic chemicals in use are xenobiotic compounds *i.e.* compounds that are potentially toxic to living organisms, extensive areas of land and large bodies of water are polluted with organic chemicals that pose major environmental and human health problems. The existing pollution and the threat of contamination from new xenobiotics suggest that there is urgent need for effective and affordable technological solutions to these problems. An attractive technological option is to harness the biochemical activities of plants for monitoring and remediating (phytoremediation) contaminated sites.

As primary producers at the base of the food chain, plants are ideally placed to act as bio-indicators of environmental pollution, as they experience and respond to environmental toxicants

0939–5075/2005/0700–0544 $ 06.00

sooner than organisms at higher trophic levels. In addition plants can be more sensitive to some environmental contaminants than animals (Wang and Freemark, 1995). Furthermore, some plants are capable of detoxifying anthropogenic chemicals by metabolic transformation and could prove useful for the remediation of contaminated water and soil: so-called phytoremediation.

In general, uptake followed by metabolism and compartmentation are responsible for the detoxification of organic xenobiotics in plants. These processes are mediated by complex cellular systems that are organised in well-defined biochemical phases (Coleman *et al.*, 1997) that are controlled by large families of genes (Edwards *et al.*, 2000; Sánchez-Fernández *et al.*, 2001). In general, at least three major phases of metabolism can be recognised: Phase I reactions usually involve hydrolysis or oxidation and are catalysed by esterases and cytochromes P450 respectively; Phase II metabolism consists of conjugation reactions, which are mediated by enzymes such as glucosyl and glutathione transferases; in Phase III, conjugates are transported to the vacuole by the action of ABC transporters.

The exploitation plants for bio-monitoring and phytoremediation require identification of a plant species that will respond and/or detoxify organic chemical contaminants. Furthermore, the development of phytoremediation as a viable commercial technology will almost certainly necessitate further screening to identify biotypes with the highest detoxification capacities or optimising these processes by developing new plants either through conventional plant breeding or genetic modification. So far research technologies aimed at developing plants for bio-indication/bio-monitoring and for remediation have largely relied on standardised chemical and biochemical procedures to evaluate phytotoxicity, detoxification and metabolic fate of organic pollutants in plants. The large number of chemicals involved coupled with the number of potential plant species make this approach an enormous task. An emerging idea aimed at overcoming the complexities of the problem is to make use of global approaches such as the new "omic" technologies to identify molecular changes in plant tissues exposed to specific organic xenobiotics. From expression profiles at the messenger RNA level, transcriptomics permit the identification of function-related gene clusters and at the protein level proteomics provide information on where, when and at what level specific proteins accumulate (see Fig. 1).

In transcriptomic analyses many genes can be studied in parallel. Essentially, thousands of cDNAs or oligonucleotides designed from expression sequence tag (EST) information are used to create DNA microarrays, which can then be hybridised with the mRNA from plants of interest that have been exposed to different treatments and the gene expression compared. Genes (sequences) that show significant differences in expression between treatments can then be selected for further study. Generally, analyses are carried out with same species microarrays, however despite the limitations associated with specificity, useful results can also be obtained by cross-species analyses.

In proteomic analysis, denatured proteins are separated by two-dimensional gel electrophoresis according to two independent properties, isoelectric point in the first dimension and apparent molecular weight in the second. This results in gels with hundreds of protein spots (gene products). These proteins can be identified by mass spectro-

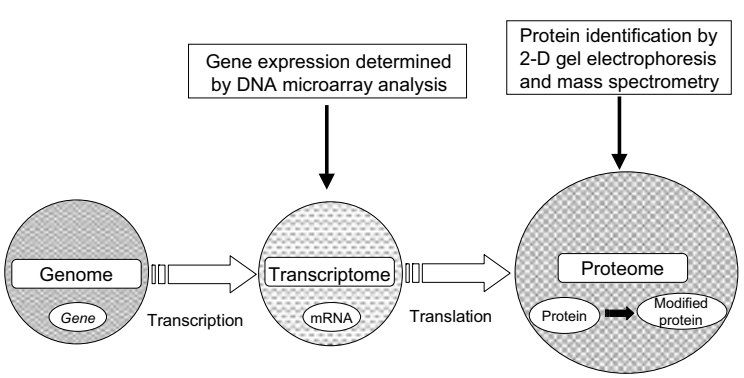

Fig. 1. Global approaches to study gene function.

metry; for example MALDI-TOF (Matrix Assisted Laser Desorption Ionisation-Time of Flight) spectrometry gives the masses of tryptic peptides obtained from the protein spots and the proteins can be identified from information held in data banks.

Here we review the initial results of proof-of-concept studies that are relevant to phytoremediation by using these "omic" technologies.

Transcriptomic study of the response of *Arabidopsis thaliana* to methanol

Formaldehyde is a volatile organic compound (VOC) that has been classified as a potential human carcinogen. Formaldehyde is widely used in the manufacture of building materials and water soluble compounds of formaldehyde are frequently present in the industrial wastewater. Furthermore because of its volatility formaldehyde is a common pollutant of indoor and outdoor air and the removal of formaldehyde from contaminated air by plants has been reported (Sandermann *et al.*, 1997). Plants also encounter formaldehyde from a number of other sources including the metabolism of endogenous methanol as shown in Fig. 2. The metabolism of formaldehyde involves its spontaneous binding to glutathione to form *S*-hydroxymethylglutathione, which in turn acts as a substrate for formaldehyde dehydrogenase forming *S*-formylglutathione as the product. As shown in Fig. 2, *S*-formylglutathione is hydrolysed by the action of *S*-formylglutathione hydrolase to form glutathione and formate. Further metabolism of formate can occur either via the C-1 pathway or it can be oxidised to carbon dioxide by the action of formate dehydrogenase. This pathway is known to be present in *A. thaliana* and the principal enzymes involved have been cloned and characterised (Haslam *et al.*, 2001).

With this background of information on the metabolism of methanol/formaldehyde, transcriptomic studies were employed to examine the response of *A. thaliana* to exogenous methanol treatment (Downie *et al.*, 2004). The rationale being that expression profiles at the mRNA level would provide valuable information on the response of the plant to methanol and its metabolic product formaldehyde.

Gene expression in response to methanol treatment of *A. thaliana* leaves analysed with a 26,090 element microarray showed that a relatively high number, 484 transcripts (1.9%), were regulated by methanol. As might be expected for this level of transcriptional activity, many were genes associated with transcription and RNA processing. However, based on functional analysis the largest category of regulated genes was metabolism. Both the number and overall proportion of this group

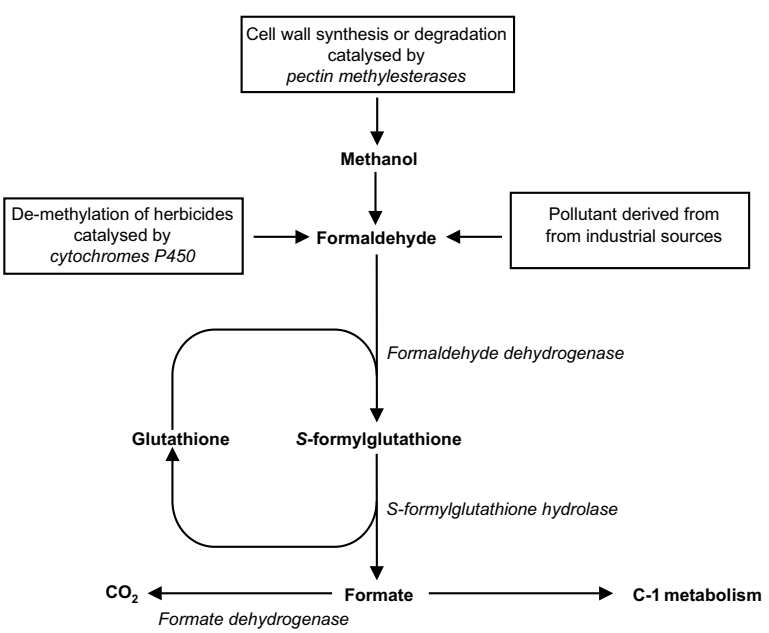

Fig. 2. Outline of the pathways for the metabolism of methanol and formaldehyde in plants.

increased over 72 h. However, of particular interest was the finding that within the various functional categories, genes encoding proteins normally associated with detoxification were by far the most strongly represented group. These proteins include cytochromes P450, the conjugation enzymes glucosyl and glutathione transferases and ABC transporters.

Flavonoids are stress-related compounds that are thought to have a protective function. A feature of plants treated with methanol was the obvious increase in anthocyanin pigments. The microarray data showed that one of the responses to methanol was at a key branch point in the flavanoid pathway with metabolism directed towards anthocyanin synthesis and away from flavonol biosynthesis. Further to this it has been previously emphasised that the terminal steps of anthocyanin and their coupled deposition in the vacuole involve glucosyl and glutathione transferases and ABC transporters, proteins also associated with detoxification (Goodman *et al.*, 2004).

Data from the microarray analysis and from RNA gel blot analysis, revealed that transcripts of formaldehyde dehydrogenase and formate dehydrogenase key enzymes in the metabolism of methanol and formaldehyde, as shown in Fig. 2, were unaffected by methanol treatment. This unexpected result suggests that there are alternative pathway(s) for the metabolism of methanol or that the enzymes are regulated by post-translation modification or other means. If the latter conclusion is true, it is a reminder that attempts to improve the flux through detoxification pathways by the facile increase of the level of the protein of key enzymes (*e.g.* by genetic modification) may not always produce the desired result.

Proteomics of leaf apoplastic extracts of *Arabidopsis thaliana*

Regulation is not always amenable to transcriptomic analysis. Furthermore, the correlation between mRNA abundance and protein levels may be poor and also the regulation of some proteins frequently takes place at the post-transcriptional level (*e.g.* protein modifications). Proteomic analysis is therefore necessary to know how and to what extent a messenger will be translated, where the corresponding protein will accumulate and to identify post-translation modifications. In the context of phytoremediation, proteomics also offer the opportunity of studying the multiple gene effects that may be elicited by chemical stress.

The apoplast is the first tissue compartment that an organic chemical encounters during the initial stages of its absorption by a plant. This compartment consisting of the extracellular matrix including the cell wall, is a dynamic environment made up of several chemical components including carbohydrates, proteins, lignin, water and metabolites. The apoplast is implicated in the plant's response to biotic and abiotic perturbations and plays a role in a variety of functions including signalling, defence and detoxification. Furthermore, esterases (Phase I enzymes) that act on herbicide esters (Haslam *et al.*, 2001) and GST (Phase II) activity (Flury *et al.*, 1996) have been identified in apoplastic extracts. Although a number of proteins have been designated as apoplastic proteins, there is a paucity of studies describing the full complement of proteins present in the apoplast. The successful development of procedures for the extraction of soluble apoplastic proteins free of contamination by proteins from other compartments coupled with proteomic analysis offers a new and rapid way of identifying an apoplastic protein if the appropriate genomic data is available for the species under study.

In a recent study (Haslam *et al.*, 2003), leaf apoplast extracts from three species, *Arabidopsis thaliana*, *Triticum aestivum* (wheat) and *Oryza sativa* (rice), were subjected to proteomic analysis. Proteins were separated by two-dimensional electrophoresis and the dominant proteins were excised and then sequenced and identified by MALDI-TOF mass spectrometry. For *A. thaliana*, 56% of the selected proteins were identified and in the case of *O. sativa*, 39%. The proteins identified belonged to several different functional groups with microbial defence and cell expansion the principal classes. It is interesting to note that the patterns of proteins revealed by the two-dimensional analysis were different for each species, perhaps indicating that there might be a characteristic "fingerprint" or pattern for each species at the apoplast level. However it must be borne in mind that many of the proteins were not identified and that the dynamic nature of the apoplast must come into consideration. However this study has demonstrated that proteomic analysis can be used to resolve the soluble apoplastic protein complement and to identify at the protein level any responses induced by organic chemical effectors.

Summary and outlook

Plants can be developed as effective tools for the sensing, cleanup and post-remediation control of water and soil polluted by organic chemicals. The accomplishment of these objectives requires harnessing and optimising the inherent mechanisms that plants use to deal with chemical toxicants.

Use of test plants as effective early warning devices for pollution control depends on identifying rapid-response events at the molecular level that can be used as reliable markers. Ideally these markers should respond at low levels of contamination so that remedial action can be taken before significant ecological damage is done.

In general phytoremediation of organic pollutants in soil and water relies on metabolic reactions in the plant that can transform the parent toxicant to detoxified products that are either accumulated in the plant's biomass or are vaporised to the atmosphere. Identifying the metabolic pathways involved and determining the control mechanisms that regulate the pathways is a necessary step towards the rational optimisation of the detoxification process. As some of these pathways overlap or parallel pathways plants use for the synthesis of endogenous compounds (as discussed above for anthocyanins), understanding how the pathways are controlled is of paramount importance. The relevance of understanding control mechanisms is highlighted by failed attempts to improve the yield of commercially important products in transgenic plants by overexpressing so-called rate-limiting enzymes of metabolic pathways.

The global techniques such as transcriptomics and proteomics provide plant scientists with the ability to integrate information from the genome, expressed mRNAs, their respective proteins and their subcellular localisation. These approaches will lead to important insights into the mechanisms that plants use to deal with toxic organic chemicals and ways of improving phytoremediation activities.

Coleman J. O. D., Blake-Kalff M. M. A., and Davies T. G. (1997), Detoxification of xenobiotics by plants: chemical modification and vacuolar compartmentation. Trends Plant Sci. **2**, 144–151.

Downie A. L., Miyazaki S., Bohnert H. J., John P., Coleman J. O. D., Parry M. A. J., and Haslam R. P. (2004), Expression profiling of the response *Arabidopsis thaliana* to methanol stimulation. Phytochemistry, **65**, 2305–2316.

Edwards R., Dixon D. P., and Walbot V. (2000), Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. Trends Plant Sci. **5**, 193- 198.

Flury T., Wagner E., and Kreuz K. (1996), An inducible glutathione S-transferase in soybean hypocotyl is localized in the apoplast. Plant Physiol. **112**, 1185–1190.

Goodman C. D., Casati P., and Walbot V. (2004), A multidrug resistance-associated protein involved in anthocyanin transport. Plant Cell **16**, 1812 –1826.

Haslam R., Raveton M., Cole D., Pallett K., and Coleman J. O. D. (2001), The identification and properties of apoplastic carboxylesterases from wheat that catalyse de-esterification of herbicides. Pestic. Biochem. Physiol. **71**, 178–189.

Haslam R., Downie A. L., Raveton M., Gallardo K., Job D., Pallett K. E., John P., Parry M. A. J., and Coleman J. O. D. (2003), The assessment of enriched apoplastic extracts using proteomic approaches. Ann. Appl. Biol. **143**, 81–91.

Sánchez-Fernández R, Davies T. G. E., Coleman J. O. D., and Rea P. A. (2001), The *Arabidopsis thaliana* ABC protein superfamily, a complete inventory. J. Biol. Chem. **276**, 30231–30244.

Sandermann H., Nasse B., and Langebartels C. (1997), Luftreinigung durch Zimmerpflanzen. Eine Beurteilung aus wissenschaftlicher Sicht. In: Reinhaltung der Innenraumluft (FGU Berlin, ed.). Seminar 33. Utech '97, Berlin, pp. 77–88.

Wank W. and Freemark K. (1995), The use of plants for environmental monitoring and assessment. Ecotox. Environ. Safety **30**, 289–301.