Promoting plant promoters

G.C. Machray, P.E. Hedley, D. Davidson, A.F.M. Ibrahim, J.A. Watters & J.W.S. Brown

Agricultural application of plant genetic manipulation is now a reality, with major shares of production of a variety of crops in the USA being given over to new cultivars produced through recombinant DNA technology. The technology is still in its infancy however, being confined largely to the modification of single traits, involving the transformation of plants with single genes. Many of these confer entirely new properties to the plant, additional to and largely independent of the existing plant biochemistry and physiology. Others, while similar to existing functions, have properties which confer insensitivity to the regulatory checks and balances to which their endogenous plant analogue is subject, and hence superimpose the desired effect on a background of normal plant cell metabolism. In the main, reactions at the end of complex pathways have been targeted for manipulation to yield a desired product. Such approaches, while successful, offer limited insight into the grander scheme of plant metabolism, which remains an area of great opportunity for more sophisticated applications of this technology, both for the increased knowledge of plant biology to be gained, and for the subsequent rational and safe exploitation of that knowledge.

Many factors need to be taken into account to refine and improve strategies for the genetic manipulation of plant cell metabolism. Pathway control may be shared among a number of components and require treatment as a quantitative trait. Intricate regulatory networks, mediated by multiple signal transduction pathways, will require dissection and individual manipulation of contributing elements. The effects of the compartmentation of metabolism and metabolic channelling must be considered. All of these require detailed knowledge of relevant gene expression and of the fate of the products of that expression. This information will aid decision-making in the choice of targets for manipulation. In addition to the identification of targets, effective manipulation of the complex metabolic processes which are targeted will require highly specific promoters able to regulate gene expression in a highly controlled manner, to ensure that spatial and temporal constraints are met.

Promoters can be divided into two major classes: constitutive, which are expressed in all cells, and regulated, which can be expressed in particular cells or tissues, or at particular stages of development. In addition, expression levels from each type of promoter can be highly regulated. The source of promoters for use in transgenics is isolated genes. Many plant genes are organised in multigene families where variation in

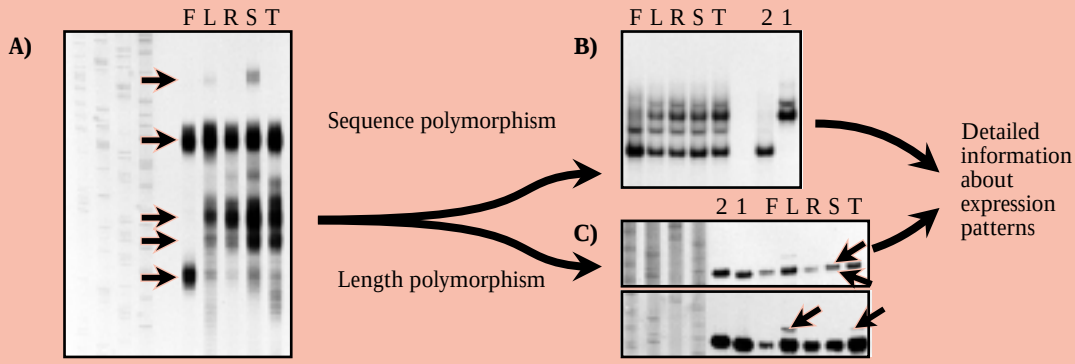

Figure 1 A novel approach for expression profiling of multigene families and isolation of gene-specific promoters
5' RACE is carried out on total RNA isolated from different organs using a [^{32}P] end-labelled primer. Products are separated on a DNA sequencing gel alongside a DNA sequencing ladder marker (A). Products are either directly cloned into pGEM-T Easy and sequenced, or individual bands (arrows) are excised, re-PCR'd and sequenced. The generated sequences of the 5' UTRs allow specific bands to be related to different genes and/or different transcription start sites. If there is sequence polymorphism but no length polymorphism among different gene family members (U1A), SSCP analysis is carried out by running RT-PCR products generated by using end-labelled primers on a mutation detection enhancement (MDE) gel (B). Labelled PCR products from plasmids representing individual members are used as controls (1=U1A-1, 2= U1A-2). If length polymorphism results from different sequences of the 5' UTRs (U2B"), labelled RT-PCR products are separated on a DNA sequencing gel (C) with labelled PCR products from clones of individual members as controls (1=U2B"-1, 2=U2B"-2). Short exposure allows visualisation of single base differences (arrows in upper photograph), while longer exposure and higher loading intensity allow detection of rare transcripts (arrows in lower photograph). The detailed information from these experiments allows primers to be designed for the isolation of gene-specific promoters by inverse PCR. F=flower, L=leaf, R=root, S=stem, T=tuber.

expression pattern and level exists among individual members. We have been actively developing and applying molecular techniques to examine expression pattern and rapidly isolate gene-specific promoter sequences.

Differential spatial and temporal expression from complex gene families can be dissected using technologies such as RT-PCR[1] (reverse transcription - polymerase chain reaction) based on differences in length of transcripts. Where no length variation exists, SSCP (single-stranded conformational polymorphism) based on variation in coding sequence content, can be applied[2]. Many genes also encode transcripts which include non-coding regions - untranslated regions at their 5' or 3' ends (5'-UTR, 3'-UTR). The nucleotide sequences of these can be obtained by approaches such as RACE (rapid amplification of cDNA ends). These sequences are ideal for RT-PCR techniques, because they are more variable, both in length and content, than the coding sequence for genes which are members of gene families. Use of gene-specific sequences for individual gene family members in these approaches can deliver not only increased definition of the tissue-specific expression patterns from multigene families but can also confirm constitutive expression. 5'-UTR sequences, as well as 5' introns, can be used in the design of primers for inverse-PCR (IPCR) from the genome to obtain specific promoter sequences for genes which show constitutive or regulated expression. These promoters then provide the switches to deliver constitutive expression, or to manipulate expression specifically within selected tissues, of heterologous genes when built into recombinant constructs.

Within the Gene Expression Unit, these RT-PCR based methods have been used to characterise expression from several complex plant gene families and to clone promoters from genes which show interesting expression profiles. Figure 1 describes application of these technologies to the analysis of tissue-specific expression of the U1A and U2B" gene families in potato. This has revealed lack of expression of one U1A gene in floral tissue, constitutive expression of another, and constitutive expression of an additional gene not previously detected. Two U2B" genes were shown to be constitutively expressed, while a third was expressed only in leaf and tuber tissue. The confirmation of constitutive expression of four genes is in itself

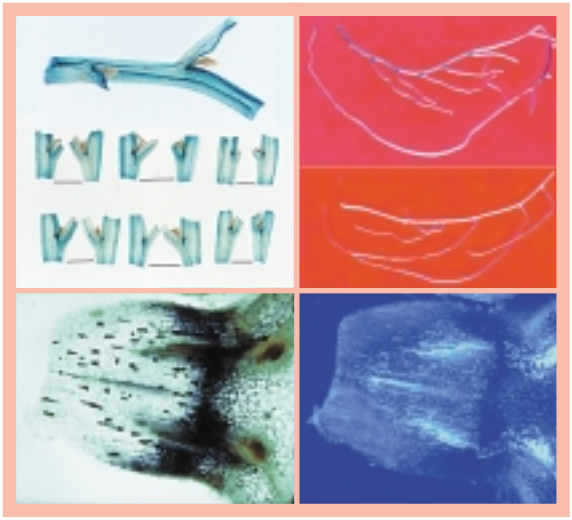

Figure 2 Expression pattern from the promoter of the invertase gene *invGE*. GUS reporter gene expression driven by the *invGE* promoter in a transgenic potato line under axillary buds in stem (top left) and in transgenic v. control roots (top right). The paired images below show the same cross section of a stem axil from this transgenic line stained with GUS (left) and with aniline blue (right) to reveal callose-containing xylem tissue.

an important finding - the promoters of these genes are candidates for genetic manipulations requiring constitutive expression. One which we have cloned is expressed at 65% of the level of the CaMV 35S promoter; promoters such as this are of increasing importance given concerns over gene silencing resulting from promoter duplication and the use of promoters derived from viruses or other plant pathogens.

These methods also have been applied to analyse tissue-specific expression and to obtain promoters from the invertase gene family in potato. We have previously described the characterisation of a pollen-specific invertase promoter[3] and suggested a potential use for the pollen-specific promoter in the generation of male-sterility in cultivated potato. This is now perhaps an even more desirable trait given current concerns over pollen dispersal into the environment on release into field-trialling of genetically-manipulated crops. Immediately upstream of this promoter in the potato genome lies a second invertase gene. The promoter of this second gene has been cloned by the methods described above and a series of transgenic potato plants generated in which it drives expression of the *uidA* gene coding β-glucuronidase (GUS). This promoter also determines expression in floral tissues, including pollen and the calyx. Its expression pattern is more diverse however and further expression is seen in specific vegetative tissues of the potato plant. Notably, these include regions where the vasculature branches, such as where the stem gives rise to lateral leaves or stolons, or where lateral roots branch from a main root (Fig. 2). Expression under the lateral bud of the stem is important because in the tuber, which is a modified stem with shortened and broadened axis,

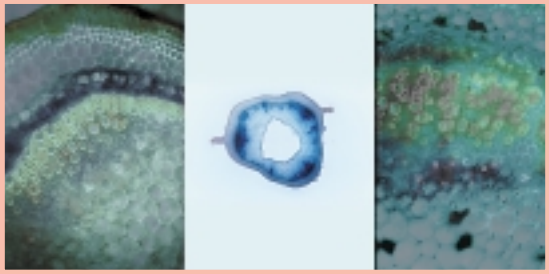

Figure 3 Expression pattern from the promoters of the invertase genes *invCD1* and *invCD4*. Both promoters drive expression in the three major and three minor vascular bundles and associated tissues, as seen in the GUS-stained cross-section (center) of stem tissue from a transgenic potato line carrying an *invCD1* promoter-GUS reporter fusion. Individual vascular bundles from the appropriate transgenic lines co-stained for GUS activity (blue) and with aniline blue to reveal callose-containing xylem (white) showed *invCD1* expression preferentially in external phloem (left) and *invCD4* expression preferentially in internal phloem (right).

the equivalent of this region is the 'eye', which will give rise to the new sprout. Each eye is formed by leaf scars with subtended lateral buds with undeveloped internodes, and expression was also detected from this promoter in eyes excised from tubers. The upstream promoter may have utility in the control of developmental processes in the potato plant since it governs expression in key tissues for micropropagation as well as in the sprouting of tubers. Enhanced control of these processes could have significant commercial benefits.

Two further invertase gene promoters have been cloned using gene-specific IPCR technology. Series of transgenic lines generated for each of these, in which they govern expression of *uidA*, indicate a highly specific expression pattern confined to stem and petiole tissues. Both promoters direct expression in vascular tissue, within the three major and three minor vascular bundles and further vascular tissue interspersed between them (Fig. 3). In potato, each vascular bundle is composed of external and internal phloem sandwiching a layer of xylem. When examined in detail, one promoter expressed GUS preferentially in external phloem while GUS expression was driven preferentially in internal phloem by the second promoter. Both promoters may have application in insect control strategies, particularly against sap-sucking insects. Further roles for these promoters are under investigation - evidence suggests they are, for example, likely to be switched on by pathogen attack and may therefore have further use in control strategies targeted against diverse plant pathogens ranging from nematodes to fungi and bacteria.

We have described the successful application of these methods to the analysis of differential expression in diverse plant multigene families and to provide promoter switches for further biotechnological application. While this is an important objective, there also remain interesting biological questions about the promoters and the genes which they regulate. What are the controlling elements in the promoters and can they be manipulated to generate altered expression profiles? What are the roles of the gene products in the biochemistry and physiology of the plant and why have multiple genes evolved to carry out this function in closely-related environments? Fortunately, the promoters themselves provide a means by which we can begin to address these problems by driving targeted antisense or post-transcriptional gene silencing knock-out of the function of their cognate gene.

References

1 Simpson, C.G., Sawbridge, T.I., Jenkins, G.I. & Brown, J.W.S. (1993). *Nucleic Acids Research* 20, 5861-5862.

2 Hedley, P.E., Machray, G.C., Davies, H.V., Burch, L. & Waugh, R. (1994). *Gene* 145, 211-214.

3 Maddison, A., Meyer, R., Hedley, P. & Machray, G.C. (1997). *Annual Report of the Scottish Crop Research Institute for 1996/97*. Scottish Crop Research Institute, Dundee, 100-101.